Design of custom CMOS amplifiers for nanoscale bio-interfaces

Siddharth Shekar

Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Graduate School of Arts and Sciences

Columbia University

2019
ABSTRACT

Design of custom CMOS amplifiers for nanoscale bio-interfaces

Siddharth Shekar

The miniaturization of electronics is a technique that holds a lot of potential in improving system performance in a variety of applications. The simultaneous miniaturization of sensors into the nano-scale has provided new ways to probe biological systems. Careful co-design of these electronics and sensors can unlock measurements and experiments that would otherwise be impossible to achieve. This thesis describes the design of two such instrumentation amplifiers and shows that significant gains in temporal resolution and noise performance are possible through careful optimization.

A custom integrated amplifier is developed for improving the temporal resolution in nanopore recordings. The amplifier is designed in a commercial 0.18 µm complementary metal-oxide-semiconductor (CMOS) process. A platform is then built with the amplifier at its core that integrates glass-passivated solid-state nanopores to achieve measurement bandwidth over an order of magnitude greater than the state of the art. The use of wavelet transforms for denoising the data and further improving the signal-to-noise ratio (SNR) is then explored.

A second amplifier is designed in a 0.18 µm CMOS process for intracellular recordings from neurons. The amplifier contains all the compensation circuitry required for canceling the effects of the electrode non-idealities. Compared to equivalent commercial systems and the state of the art, the amplifier performs comparably or better while consuming orders of magnitude lower power.

These systems can inform the design of extremely miniaturized application-specific integrated amplifiers of the future.
Contents

List of Figures vi

List of Tables x

Acknowledgments xi

1 Introduction 1
   1.1 Thesis outline ........................................ 2

2 Background 4
   2.1 Introduction ........................................ 4
   2.2 Nanopores ........................................... 5
      2.2.1 Biological nanopores .............................. 6
      2.2.2 Solid-state nanopores ............................ 8
      2.2.3 Applications .................................... 10
   2.3 Signal and noise in nanopore recordings ............... 12
      2.3.1 Nanopore recordings ............................ 13
2.3.2 Signal-to-noise ratio ............................................. 14
2.3.3 Modeling signals .............................................. 16
2.3.4 Modeling noise .................................................. 20
2.3.5 SNR-constrained bandwidth ................................. 27
2.4 Intracellular recordings ........................................... 29
  2.4.1 Configurations ................................................. 30
  2.4.2 Measurement modes ........................................... 31
  2.4.3 Modeling the interface ....................................... 32
  2.4.4 Need for compensation ....................................... 36
2.5 Multi-clamp amplifiers ........................................... 38
  2.5.1 Commercial solutions ......................................... 38
  2.5.2 Integrated approaches ....................................... 41
3 Design of a Low-Noise Nanopore Front-End ......................... 43
  3.1 Introduction ..................................................... 43
  3.2 Design considerations .......................................... 44
    3.2.1 Input-referred voltage noise .............................. 44
    3.2.2 Unity-gain bandwidth ...................................... 45
    3.2.3 Input and feedback capacitance .......................... 46
    3.2.4 DC gain .................................................... 47
    3.2.5 Other considerations ....................................... 47
  3.3 Chip design ..................................................... 49
    3.3.1 Integrator ................................................... 50
    3.3.2 Voltage amplifier ........................................... 51
    3.3.3 Output buffer ............................................... 53
    3.3.4 Feedback network .......................................... 54
    3.3.5 On-chip electrode ......................................... 55
    3.3.6 Digital logic, ESD and biasing .......................... 55
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.4</td>
<td>Board design</td>
<td>134</td>
</tr>
<tr>
<td>6.4.1</td>
<td>System overview</td>
<td>134</td>
</tr>
<tr>
<td>6.4.2</td>
<td>Chip packaging and assembly</td>
<td>136</td>
</tr>
<tr>
<td>6.4.3</td>
<td>Power domains</td>
<td>136</td>
</tr>
<tr>
<td>6.4.4</td>
<td>Data acquisition</td>
<td>137</td>
</tr>
<tr>
<td>6.4.5</td>
<td>Shielding and assembly</td>
<td>138</td>
</tr>
<tr>
<td>6.5</td>
<td>Electrical characterization</td>
<td>138</td>
</tr>
<tr>
<td>6.5.1</td>
<td>Current-clamp</td>
<td>139</td>
</tr>
<tr>
<td>6.5.2</td>
<td>Voltage-clamp</td>
<td>140</td>
</tr>
<tr>
<td>6.5.3</td>
<td>Comparison to state of the art</td>
<td>147</td>
</tr>
<tr>
<td>6.6</td>
<td>Experimental results</td>
<td>147</td>
</tr>
<tr>
<td>6.7</td>
<td>Summary</td>
<td>152</td>
</tr>
<tr>
<td>7</td>
<td>Conclusions</td>
<td>155</td>
</tr>
<tr>
<td>7.1</td>
<td>Summary of contributions</td>
<td>155</td>
</tr>
<tr>
<td>7.2</td>
<td>Future work</td>
<td>157</td>
</tr>
<tr>
<td>7.3</td>
<td>Final thoughts</td>
<td>160</td>
</tr>
<tr>
<td></td>
<td>References</td>
<td>162</td>
</tr>
</tbody>
</table>
## List of Figures

2.1 Coulter counter ......................................................... 6
2.2 Silicon nitride nanopore .................................................. 10
2.3 Simplified nanopore measurement setup ............................... 20
2.4 Nanopore input-referred noise PSD .................................... 28
2.5 Equivalent electrical circuit for a single ion channel ............... 33
2.6 Equivalent electrical circuit for a cell ................................ 34
2.7 Combined equivalent circuit of a cell and pipette .................... 36
2.8 Basic capacitance compensation circuit .............................. 40
2.9 Basic resistance compensation circuit ................................ 41
3.1 Circuit topology for the low-noise nanopore front-end .............. 49
3.2 Schematic of the OTA used for forming the integrator ............... 51
3.3 Schematic of fully-differential OTA (voltage amplifier) ............. 52
3.4 Schematic of fully-differential OTA (output buffer) .................. 53
3.5 Schematic of the active current-divider ................................ 55
3.6 Photograph of a single channel in the nanopore TIA ................ 57
3.7 Die photograph of the nanopore TIA chip

3.8 Nanopore amplifier board-level block diagram

3.9 Boosting filter schematic

3.10 Fourth-order 10 MHz Bessel filter schematic

3.11 Photograph of the nanopore measurement setup

3.12 AC response of the TIA

3.13 Measured DC response of the TIA as a function of injected current amplitude

3.14 Measured and simulated open-headstage input-referred noise PSD of the nanopore amplifier

3.15 Concatenated open-headstage noise time traces recorded using the nanopore amplifier at different filtering bandwidths

4.1 Packaged nanopore amplifier chip with and without the fluid chamber

4.2 Photograph of the on-chip electrode before and after processing

4.3 Nanopore amplifier chip with silicone passivation

4.4 Characterization of nanopores and STEM-thinning in silicon nitride

4.5 Schematic of the cross-section of the nanopore chip including the silicone passivation

4.6 Input-referred noise PSD of the nanopore amplifier with two different pores

4.7 Concatenated open-pore noise time traces recorded using the nanopore amplifier at different filtering bandwidths

4.8 Nanopore signal variation with different bias voltages

4.9 Nanopore signal variation with different filter cutoff frequencies

4.10 Short intra-event features in ssDNA translocations

4.11 ssDNA translocation events showing two-state features

4.12 ssDNA translocation events showing the deep level in the beginning or the middle of an event

4.13 ssDNA translocation event statistics
6.2 Die photograph of the multi-clamp amplifier chip .................. 122
6.3 Circuit schematic of the dual-input OTA .......................... 125
6.4 Schematic of the capacitance compensation circuit ............... 125
6.5 Simplified block diagram of the current-clamp current injection circuit . 127
6.6 Circuit schematic of the current-clamp current injection block .... 128
6.7 Circuit schematic of the TIA in the voltage-clamp ................... 130
6.8 Block diagram of the state-estimator resistance compensation circuit . 132
6.9 Circuit-level implementation of the resistance compensation circuit . . 135
6.10 Multi-clamp amplifier board-level block diagram .................... 135
6.11 Photograph of the multi-clamp daughterboard with the chip attached . . 136
6.12 Photograph of the multi-clamp amplifier inside the aluminum enclosure . 139
6.13 Noise performance of the current-clamp voltage buffer ............ 141
6.14 Square wave rise and fall time improvement with capacitance compensation 141
6.15 Linearity and step response of the current injection circuit .......... 142
6.16 Linearity of the voltage-clamp TIA as a function of input current amplitude 144
6.17 Noise performance of the voltage-clamp TIA ....................... 146
6.18 Capacitance compensation in the voltage-clamp ..................... 146
6.19 Series resistance compensation measurement (negligible $C_M$) .... 148
6.20 Series resistance compensation measurement ($C_M = 20$ pF) .......... 148
6.21 Current injection through a sharp microelectrode in the bath ......... 149
6.22 In vitro intra- and extracellular current-clamp recordings .......... 150
6.23 Photograph of a patch-clamp measurement from a neuron ........... 151
6.24 Gigaseal formation ................................................. 151
6.25 Multi-clamp current recordings .................................. 152
6.26 Multi-clamp voltage recordings ................................ 153
6.27 Current-clamp recordings of EPSP and IPSP ....................... 153
# List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Average false detection rates</td>
<td>15</td>
</tr>
<tr>
<td>4.1</td>
<td>Contributions of low frequency noise to integrated noise at different band-</td>
<td>79</td>
</tr>
<tr>
<td></td>
<td>widths</td>
<td></td>
</tr>
<tr>
<td>4.2</td>
<td>Comparison of this work with other high-bandwidth DNA translocation studies</td>
<td>93</td>
</tr>
<tr>
<td></td>
<td>through solid-state nanopores</td>
<td></td>
</tr>
<tr>
<td>6.1</td>
<td>Comparison of the integrated multi-clamp amplifier to existing state of</td>
<td>147</td>
</tr>
<tr>
<td></td>
<td>the art</td>
<td></td>
</tr>
</tbody>
</table>
Acknowledgments

There are a number of people who have helped and supported me in this endeavor. I am extremely grateful to each and every one of these people. The first person I would like to thank is my advisor, Kenneth Shepard, who was responsible for convincing me to go down this path and for believing in me and my abilities. Ken’s enthusiasm is infectious and his brilliance sets a lofty goal to aspire towards.

The first project I worked on was the nanopore amplifier and I learned a great deal from Daniel Fleischer and Peijie Ong, who have been wonderful office-mates too. Special thanks to Jianxun Lin and Jacob Rosenstein for teaching me the tricks of the trade. This project would have gone nowhere without help from the wonderful members of Marija Drndić’s lab at the University of Pennsylvania. Marija’s optimism and energy are second to none and she has been an inspirational figure all throughout. David Niedzwiecki and Francis Chien Chen-Chi were the magicians who fabricated the nanopores. I will fondly remember our (many) lunches from Magic Carpet and at Ramen Bar.

The patch-clamp amplifier was devised over one of the innumerable discussions I’ve had with the incredibly knowledgeable Krishna Jayant. Krishna’s hunger for
knowledge is likely exceeded only by his hunger for spice – a fact I learned first-hand over our many lunches at the Chinese restaurants nearby.

I would particularly like to thank Kevin Tien and Girish Ramakrishnan for feeding my cravings for pizza and for the countless hours we spent discussing the philosophical aspects of circuit design.

I would also like to thank the members of my committee Professors John Kymissis, Lance Kam and Andrew Marks for their time and their contributions.

I have had the good fortune of overlapping with many members of the Shepard lab who have all contributed towards making my stay in the lab enjoyable and have been gracious enough to share their knowledge with me. I would like to thank Andy Hartel, Adrian Bradd, Boyan Penkov, Chen Shi, Esha John, Fengqi Zhang, Jared Roseman, Jeffrey Sherman, Jordan Thimot, Kukjoo Kim, Michael Lekas, Rizwan Huq, Scott Kontak, Scott Trocchia, Steven Warren, Tarun Chari, Tiago Costa and Yihan Zhang. There probably are names I missed there and my apologies if I did, but my thanks are to you too.

My parents and my sister have always been beacons of light in my life and I have never managed to find the words to express how blessed I feel about their unwavering faith in me. To them, I will always be a child and to me, they will always be the world. Finally, my wife has seen me at my best and at my worst. Knowing that she’s always by my side has made this journey so much easier, and for that, I remain endlessly grateful.
Chapter 1

Introduction

The year this thesis was written marked the 60th anniversary of the invention of the integrated circuit by Jack Kilby and, independently, Robert Noyce - a feat for which they were jointly awarded the Nobel Prize in Physics in 2000 [1]. Just a year after their invention, the physicist Richard Feynman delivered his seminal talk “There’s plenty of room at the bottom” [2]. Unbeknownst to them (perhaps) at the time, humanity was about to embark on what was to be, until then, the greatest scientific advance made in any six decades of human history. A key enabler for this advance was the rapid miniaturization of transistors as governed by Moore’s Law [3].

We now have the technology to fit billions of transistors on a single chip about the size of a quarter. The advancement in technology that led to this has also allowed us to build ever-smaller structures and sensors in order to probe events at the microscopic scale. This relentless progress in the design of sensors and front-end amplifiers has allowed scientists of various disciplines to ask increasingly specific questions and have
the right set of tools to answer them. In the biological context, we seek to understand how the structure and function of simple molecules influences behavior at the cellular level, at the organism level and at the community level.

The greatest difficulty in these studies at the molecular level invariably links to the proportionally shrinking magnitudes of the signals of interest. While some such difficulties may be addressed by the design of new systems, such as the invention of the patch-clamp amplifier in order to study membrane currents in the giant squid axon, practically, there exists a limit to every recording beyond which the signal cannot be separated from the noise. A deep understanding of the system and the associated noise processes can, however, help us push that limit further.

1.1 Thesis outline

Chapter 2 provides an introduction to nanopores and summarizes the various applications that use nanopores as single-molecule sensors. It also discusses the basics of signal and noise and their optimization in the context of nanopore recordings. Intracellular recordings, including their various configurations and modes are also reviewed. The chapter concludes with a discussion on the current state of the art of integrated multi-clamp amplifiers.

Chapter 3 talks about the design and testing of an amplifier chip optimized for nanopore recordings. The chip is fabricated in a commercial 0.18 µm complementary metal-oxide-semiconductor (CMOS) process and contains twenty-five independently addressable channels. System-level design considerations are also discussed.

Chapter 4 explores the use of the nanopore amplifier discussed in Chapter 3 in the design of a low-noise nanopore recording platform. Translocation experiments with ssDNA in 3 M KCl enable measurement bandwidths with up to 100 ns of temporal resolution. Additional glass-passivation techniques to further reduce parasitic capacitance allow for further improvements in the signal-to-noise ratio (SNR).
Chapter 5 investigates the use of wavelet denoising as opposed to conventional Bessel filtering for improving the SNR in high-bandwidth nanopore and ion channel recordings. When compared to Bessel filters, wavelet denoising offers a superior combination of shape retention and noise suppression and its applicability to simulated as well as measured data is explored.

Chapter 6 discusses the design of a fully integrated patch-clamp amplifier. The chip is fabricated in a commercial 0.18 µm process and includes circuitry for injecting current as well as for canceling the resistive and capacitive non-idealities of the pipette used for recording. The system is validated both electrically as well as experimentally.

Chapter 7 concludes the dissertation with a summary of the original contributions made by this work to the field of microelectronics and front-end amplifiers for biological applications. Avenues for future work are also discussed.
Chapter 2

Background

2.1 Introduction

Biological systems make extensive use of electrochemical signaling while interacting with one another and with the environment. Considering that the membrane of a cell is nearly impervious to ion transport, the primary conduit for these interactions is a class of proteins known as ion channels [4]. The introduction of the patch-clamp technique provided the first reliable way to electrically observe the currents through these ion channels reproducibly and with high SNR [5]. We now know that these channels typically carry currents in the pA range. These relatively weak currents are only observable at bandwidths of a few kHz.

The geometry of ion channels places them in a class of structures known as nanopores. Some naturally occurring and artificial biological ion channels can be configured so that they remain “always on” and are considered to be biological nanopores.
The biomimetic equivalents of these “always on” ion channels fabricated in solid-state membranes are appropriately titled solid-state nanopores. The nanometer scale geometry of nanopores makes them very attractive for use in single-molecule studies. The bandwidths offered here far exceed the bandwidths offered by optical techniques which rely on an additional transduction step of converting the chemical information to photons before converting them into electrons using an image sensor.

In instrumentation design, it is imperative to understand the sensor’s behavior in the environment it will be measured in, in order to design an appropriate front-end amplifier. This chapter first briefly explores the history of solid-state nanopores and discusses the practical limits that set the upper limit on the bandwidth in nanopore recordings. The chapter also covers intracellular current and voltage recordings and their associated design considerations. These insights will guide the design of front-ends that can further push the boundaries of current measurement capabilities.

2.2 Nanopores

A nanopore, as the name suggests, is a nanometer-sized opening in a dielectric membrane. Nanopores are commonly found in nature and play a critical role in biological systems. Over the last two decades, solid-state nanopores modeled after their biological counterparts have also been fabricated. While nanopores have been used for a variety of purposes [6–10], their use as single-molecule sensors is of particular interest. This is because their geometry endows them with extremely high spatio-temporal resolution.

The use of nanopores as single-molecule sensors is essentially a miniaturization of the operating principle of the Coulter counter [11] so that it operates at nanometer-length scales. The basic operation of a Coulter counter is illustrated in Figure 2.1. Consider a system of two reservoirs separated by an impermeable membrane containing a single pore. An electrochemical gradient applied across this membrane will
set up an ionic current through the pore. Any suspended particles in the solution that pass through this pore will create a change in the ionic current so long as the conductivity of the particle is different from that of the solution. The more similar the size of the particle becomes to the size of the pore itself, the larger the percentage of current that is blocked and hence, the easier it is to detect. Thus, the number of particles suspended in the solution can be counted precisely using an appropriately sized pore.

![Diagram of Coulter counter](image)

Figure 2.1: Operating principle of the Coulter counter. As a suspended particle in the reservoir to the left translocates through the pore into the reservoir on the right, it blocks the existing ionic current \( I_{\text{open}} \) to a new level \( I_{\text{blocked}} \). The number of such events directly corresponds to the number of particles that have translocated.

### 2.2.1 Biological nanopores

In his notes from 1989, David Deamer was the first person to document the idea of applying this technique to sequence DNA [12]. Given that the diameter of single-stranded DNA (ssDNA) is approximately 1.3 nm, this would require the use of a pore whose diameter was also of such length scales. Thankfully, nature is replete with examples of such pores in the form of ion channels. Briefly, ion channels are membrane-spanning proteins that allow the passage of ions such as potassium, sodium
and calcium through otherwise impermeable lipid membranes. These channels are usually gated and have a statistical chance of being closed depending on the electrochemical environment around them.

Biological nanopores are a unique subclass of ion channels and are characterized by the fact that they are “always open”. Practically, biological nanopores are often toxins that form nanoscale pores, or outer-membrane protein channels that are incorporated into a lipid membrane [13]. Early work with biological nanopores focused on the use of α-hemolysin [14, 15], a naturally occurring protein channel secreted by *Staphylococcus aureus*. Several other pores have been explored since then, such as *Mycobacterium smegmatis* porin A (MspA) [16, 17], bacteriophage phi29 DNA-packaging motor [18], bacteriophage SPP1 DNA-packaging motor [19], *Escherichia coli* Fragaceatoxin C (FraC) [20], *Escherichia coli* Outer membrane protein G (OmpG) [21], *Escherichia coli* cytolysin A (ClyA) [22], *Escherichia coli* Curlin sigma S-dependent growth subunit G (CsgG) [23, 24] and aerolysin [25, 26].

The greatest strength of biological nanopores is their reproducibility. By taking advantage of processes that have been perfected over millions of years of by nature, copies of biological nanopores can be made nearly identical to each other. However, this also becomes the greatest drawback of these pores in that modification of their geometry is not straight-forward. However, recent advances in bioengineering have paved the way for modifying the molecular architecture of these pores [27]. Modulating the number of charges in the channel [16], introduction of reactive amino acids or introduction of hydrophobic groups that bind organic molecules are now possible [28]. It is worth noting that as of the time of writing of this thesis, the only commercial nanopore-based DNA sequencing platform uses biological nanopores exclusively [24].

Despite being extremely versatile, biological nanopores are not without drawbacks. One of these is the already mentioned relative lack of freedom in modifying their geometry. Another disadvantage is the relatively restricted set of changes that can be made in their electrochemical environment. Changes to the temperature, pH,
salt concentration and voltage can all affect the pore’s behavior and the membrane’s structural integrity.

### 2.2.2 Solid-state nanopores

As discussed in Section 2.2.1, the lack of precise control over the geometry of the nanopores and the desire to improve their robustness led researchers to explore alternative ways of creating these nanopores. Solid-state materials such as silicon nitride and silicon oxide proved to be a straightforward choice because of the direct translation of microfabrication techniques that had been developed over the past few decades in the electronics industry.

**Fabrication techniques**

The first demonstration of true nanometer-scale control over the fabrication of solid-state nanopores was in 2001 [29]. A focused ion beam was used to sculpt the pore as it was being created with an ion detector being used in a feedback loop to control the size of the pore. In 2003, the transmission electron microscope (TEM) was used for the first time to fabricate a nanopore [30, 31]. The advantage of this technique was that the nanopore could be observed immediately after it had been created using the same tool. The TEM drilling technique quickly became the predominant way to fabricate nanopores [32]. Its inability to scale as the number of nanopores to be fabricated increases remains the only major drawback of this technique.

2014 saw the introduction of yet another technique to fabricate nanopores termed “Controlled Dielectric Breakdown” (CDB) [33–35]. In this technique, an electric field is applied across an impervious dielectric membrane suspended in a salt solution. As the strength of the applied field increases, inherent defect sites in the membrane material allow for tunneling currents to start flowing, which erodes the dielectric material in the current path and leads to the creation of a nanopore. For thin membranes,
a pulsed voltage may need to be applied instead of a constant voltage in order to achieve precise control [36]. Alternatively, a current source may be employed instead of a voltage source to create the pore [37]. CDB offers a compelling alternative to conventional techniques because it removes the need for expensive equipment such as the TEM and can, in theory, be tried by anyone who has access to a voltage source and a sensitive ammeter.

While these techniques typically form nanopores that have symmetric shapes [33, 38], conical nanopores are also common [39–41]. These are usually formed when the nanopores are created using track-etching techniques or using laser-pulled glass nanopipettes. The asymmetrical geometry of these pores lends them interesting transport properties [42].

**Materials**

A great strength of solid-state nanopores is the ability to change the dielectric material that contains the pore. Early work focused on using silicon nitride and silicon oxide since these were the materials most compatible with existing semiconductor fabrication processes [29, 30]. Since then, other metal oxides such as aluminum oxide (Al$_2$O$_3$) [43, 44] and hafnium oxide (HfO$_2$) [45, 46] have been explored. It is also possible to coat nanopores made in other dielectrics with other organic [47, 48] and inorganic materials [49] in order to change the functional properties of the nanopore channel itself.

As the nanopore is made thinner, it passes more current in its open state for a given voltage, since the same voltage is applied over a smaller ionic resistance. This, in turn, translates to larger signal levels in single-molecule experiments. An equivalent way of interpreting this is that, for the same diameter, a thinner nanopore is spatially more sensitive than a thicker pore. As a result, decreasing the nanopore’s thickness is of significant interest. Silicon nitride can be thinned using a scanning transmission electron microscope (STEM) down to the theoretical limit of amorphous silicon [50].
This sputtering technique removes nitrogen atoms faster than silicon atoms due to the higher sputtering rate of nitrogen [51]. The eventual thickness of the amorphous silicon is approximately 1 nm. An example of nanopores drilled using this technique is shown in Figure 2.2.

![Nanopores](image)

**Figure 2.2:** Nanopores of several different diameters drilled in thinned silicon nitride membranes

Two dimensional (2D) materials represent the absolute physical limit in terms of how thin nanopores can physically become [6]. Over the last few years, a variety of 2D materials such as graphene [7, 52], hexagonal boron nitride (h-BN) [53], metal dichalcogenides such as molybdenum disulfide (MoS$_2$) [54, 55] and tungsten disulfide (WS$_2$) [56] have been used to fabricate nanopores. The lattice nature of these materials has been exploited to demonstrate atomic control over the size of the fabricated nanopores [53].

### 2.2.3 Applications

Within the category of single-molecule experiments, the nanopore has proven to be an extremely versatile sensor capable of interrogating a range of different properties of the translocating molecule [57]. A few examples of such applications follow:
**Nucleic acid detection**

DNA translocation is, perhaps, the single most common type of nanopore experiments until now [58]. Starting with early experiments with biological nanopores [14, 15], recent experiments have shown detection of both single-stranded and double-stranded DNA [17, 38, 48, 59–63]. Oxford Nanopores now sells a product that can be plugged into a USB drive and sequence DNA using biological nanopores [24]. While solid-state nanopores have not reached this point yet, homopolymer differentiation [38, 64] as well as individual nucleotide discrimination [54] have been shown. Beyond mere nucleic acid detection, nanopores have also been used to detect methylation of specific sites [13, 17, 55, 65–67].

**Protein detection**

Along the lines of the idea of nucleic acid detection, nanopores have also been used to detect many different aspects of proteins, with the ultimate goal being to sequence proteins [68]. Conformational and structural changes in proteins have been investigated [69–72]. Specific proteins and their concentrations can be detected accurately when combined with DNA carriers [62, 73, 74]. Despite these advances, studies have shown that most protein translocations are fast [45] and beyond the measurement limit of commercial instruments [75].

**Force spectroscopy**

Since the constriction of the nanopore sets the upper limit on the diameter of a molecule that is allowed to translocate, a large adapter attached to a smaller molecule can allow for trapping of the molecule. Varying the voltage applied on this trapped molecule can then change the effective force experienced by the molecule in the pore and allow for characterizing molecular forces. The large molecule referred to here could just be secondary structures within the same molecule [76]. Other variations of
this experiment include the use of optical tweezers to measure the force experienced by the molecule [77]. Double-stranded DNA can also be used for such experiments by taking advantage of the DNA unzipping as it passes through the pore [78–80].

2.3 Signal and noise in nanopore recordings

The aim of any measurement is to record the signal with as high fidelity as is physically possible. For the kinds of measurements discussed here, the signal can be described to be a change in the measured current. Invariably, noise contaminates every recording so that the signal is rarely ever as clean as that shown in Figure 2.1. Depending on the exact nature of the signal and the noise, noise may sometimes appear to be like the signal itself. In general, any unwanted addition to the signal may be considered to be noise.

Electromagnetic interference (EMI) is a special class of noise that is most commonly manifested as line noise at 60 Hz (or 50 Hz depending on the geographical location where the experiment is performed) coupling into the measurement setup. EMI can also be caused by fast switching digital logic that might be near the front-end amplifier. Depending on the strength of the coupling, EMI can be large enough to saturate the front-end amplifier circuitry. While it is possible to use notch filters to reduce the effect of EMI, following the old adage of “Prevention is better than cure” is greatly preferable. Through proper grounding and shielding techniques, EMI can be reduced to arbitrarily low levels.

An arbitrary measurement $y(t)$ can, in general, be decomposed into the signal component $x(t)$ and the noise component $n(t)$. These quantities are related by the equation:

$$y(t) = x(t) + n(t)$$

Certain assumptions are typically made about the nature of $n(t)$. Primary among
these is that the instantaneous amplitude of \( n(t) \) is normally distributed with mean 0 and standard deviation \( \sigma \). Noise cannot be deterministically predicted in the time domain. Depending on the nature of the noise, it may or may not be correlated with itself in time. However, another important assumption we make about noise is that it is a wide sense stationary (WSS) process. Further, we assume that its power spectral density (PSD) remains well-defined and constant as a function of time. The PSDs of \( y(t) \), \( x(t) \) and \( n(t) \) are related in a similar manner as Equation 2.1

\[
S_y(f) = S_x(f) + S_n(f) \tag{2.2}
\]

The standard deviation of the noise process \( \sigma \) can be obtained using the following equation:

\[
\sigma = \lim_{T \to \infty} \sqrt{\frac{1}{T} \int_0^T n^2(t) dt} \tag{2.3}
\]

\( n(t) \) can be observed through \( y(t) \) if \( x(t) = 0 \). Through Parseval’s theorem, \( \sigma \) is also related to \( S_n(f) \) (assuming one-sided PSD) by the equation

\[
\sigma = \sqrt{\int_0^\infty S_n(f) df} \tag{2.4}
\]

Practically, the limits of the integral in Equation 2.4 are determined by the measurement bandwidth of the system. The lower limit is usually constrained by the duration of the recording whereas the upper limit is determined by the measurement circuitry. Since the duration of the recording can be made arbitrarily long, we can rewrite Equation 2.4 as

\[
\sigma = \sqrt{\int_0^B S_n(f) df} \tag{2.5}
\]

2.3.1 Nanopore recordings

In a typical nanopore single-molecule experiment, the membrane containing the nanopore is suspended in a solution containing electrolytes. A voltage bias is then
applied across this membrane in order to induce an ionic current flow through the nanopore. The molecule of interest is then added to the appropriate chamber. The focus in this work will primarily be on DNA molecules. Since DNA is negatively charged at pH 8, it is added to the chamber connected to the lower potential. As the DNA molecule diffuses and arrives near the nanopore, it gets electrophoretically pulled through the pore. This causes a transient block in the ionic current and generates the signal of interest.

Let $I(t)$ denote the measured current, $I_{baseline}(t)$ denote the baseline current, $I_s(t)$ denote the signal component and $I_n(t)$ represent the noise component. Similar to Equation 2.1, we can write the following equation:

$$I(t) = I_{baseline}(t) + I_s(t) + I_n(t) \quad (2.6)$$

When this equation is applied to Figure 2.1, $I_{baseline} = I_{open}$ and $I_s$ is a sequence of negative-going pulses of amplitude $\Delta I = I_{open} - I_{blocked}$.

### 2.3.2 Signal-to-noise ratio

SNR is a metric used to represent the relative magnitudes of signal and noise. Depending on the field, it may either use signal and noise powers or signal and noise amplitudes. For example, in fields where information is frequency limited (such as in communications), SNR is commonly expressed as

$$\text{SNR} = \frac{\Delta I^2}{I_{RMS}^2} = 20 \log_{10} \left( \frac{\Delta I}{I_{RMS}} \right) \text{dB}$$

where $I_{RMS}$ denotes the root-mean-squared (RMS) value of the noise and is equal to the standard deviation of the noise process when the mean is 0.

With the assumptions about the nature of the signal made in Section 2.3.1, it is more informative to express the SNR in terms of the amplitudes instead of the powers.
This is also consistent with the established practice in the field of electrophysiology [5].

\[
\text{SNR} = \frac{\Delta I}{I_{RMS}}
\]  

(2.7)

Increasing the measurement bandwidth \(B\) increases \(I_{RMS}\) through Equation 2.5. Since noise is a stochastic process with a normal amplitude distribution in the time domain, the exact amplitude of noise at any given point in time is unbounded. However, given the normal distribution and a minimum SNR, it can be shown that the average rate of false events detected by a simple thresholding algorithm for such a noise process is given by [5]

\[
\lambda_f = kB e^{-0.5 \times \text{SNR}_{min}^2}
\]  

(2.8)

where \(k\) is a value that depends on the spectral characteristics of the noise source and the filtering scheme, \(\text{SNR}_{min}\) is the ratio of the threshold used for calling an event to \(I_{RMS}\), and \(B\) is the measurement bandwidth. For a Gaussian filter, \(k\) ranges between 0.849 and 1.25 and is typically around 1 [5]. Table 2.1 lists the expected false rates for some bandwidths and \(\text{SNR}_{min}\) values used in the experiments described in this thesis.

<table>
<thead>
<tr>
<th>Bandwidth (MHz)</th>
<th>(\text{SNR}_{min})</th>
<th>(\lambda_f ) (s(^{-1}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>8</td>
<td>(7.92 \times 10^{-9})</td>
</tr>
<tr>
<td>1</td>
<td>6</td>
<td>0.02</td>
</tr>
<tr>
<td>2.5</td>
<td>6</td>
<td>0.048</td>
</tr>
<tr>
<td>5</td>
<td>6</td>
<td>0.095</td>
</tr>
<tr>
<td>10</td>
<td>6</td>
<td>0.19</td>
</tr>
</tbody>
</table>

Table 2.1: Average false rate values for different bandwidths and SNRs

As \(B\) is increased, it becomes necessary to increase \(\text{SNR}_{min}\) in order to limit the average false event rate. For example, with \(B = 5\) MHz and \(\text{SNR}_{min} = 6\), there is, on average, a false event every 10 s. If \(\text{SNR}_{min}\) is reduced to 5 instead, this increases to once every 40 ms. Increasing the available SNR of the measurement allows for the
use of larger values of SNR$_{\text{min}}$. This can be achieved either by increasing the signal strength itself or by decreasing the noise levels in the measurement.

2.3.3 Modeling signals

The typical signal in a single-molecule experiment performed using nanopores is a pulse-like current waveform. There are two parameters related to the pulse that need to be considered, the amplitude and the duration.

Signal amplitude

Consider a cylindrical nanopore with diameter $d$ and thickness $t$ suspended in an electrolyte with resistivity $\rho$. The resistance of the nanopore can be expressed as [81]:

$$R_{\text{pore}} = \rho \left( \frac{4t}{\pi d^2} + \frac{1}{d} \right)$$

where the first term in the product represents the contribution of the cylindrical portion of the nanopore and the second term represents the two-sided access resistance of the pore [82]. The access resistance starts becoming a significant contributor to the total resistance when $t/d \approx \pi/4$ which is the case for some of the nanopores tested in this work. It is also common to replace $t$ by $t_{\text{eff}}$ since nanopores formed by TEM drilling often have an hourglass shape and the effective thickness that fits the measured resistance is lower than the actual membrane thickness [38, 81, 83, 84].

The conductance of the nanopore is the inverse of its resistance:

$$G_{\text{pore}} = \frac{1}{R_{\text{pore}}}$$

$$= \frac{1}{\rho} \left( \frac{4t}{\pi d^2} + \frac{1}{d} \right)^{-1}$$

For a bias voltage $V_{\text{bias}}$ applied across the membrane, the baseline current through
the nanopore is then given by

\[ I_{\text{baseline}} = I_{\text{open}} = V_{\text{bias}} G_{\text{pore}} = \frac{V_{\text{bias}}}{\rho} \left( \frac{4t}{\pi d^2} + \frac{1}{d} \right)^{-1} \] (2.10)

When a molecule such as DNA is present in the nanopore, the effective diameter of the nanopore is modified to

\[ d' = \sqrt{d^2 - d_{\text{DNA}}^2} \] (2.11)

where \( d_{\text{DNA}} \) is the diameter of the DNA molecule and is approximately 1.2 nm for ssDNA and 2.2 nm for dsDNA. The current through the nanopore in the presence of DNA can then be written as

\[ I_{\text{blocked}} = \frac{V_{\text{bias}}}{\rho} \left( \frac{4t}{\pi d'^2} + \frac{1}{d'} \right)^{-1} \] (2.12)

From Equations 2.10 and 2.12, we can obtain an expression for the amplitude of the current blockade as

\[ \Delta I = I_{\text{open}} - I_{\text{blocked}} = \frac{V_{\text{bias}}}{\rho} \left( \left( \frac{4t_{\text{eff}}}{\pi d^2} + \frac{1}{d} \right)^{-1} - \left( \frac{4t_{\text{eff}}}{\pi d'^2} + \frac{1}{d'} \right)^{-1} \right) \] (2.13)

\[ = \frac{\pi V_{\text{bias}}}{\rho} \left( \frac{d^2}{4t_{\text{eff}} + \pi d} - \frac{d'^2}{4t_{\text{eff}} + \pi d'} \right) \] (2.14)

Equation 2.14 offers the following insights:

**Increasing** \( V_{\text{bias}} \) **increases** \( \Delta I \) It should be noted that an increase in the bias voltage also increases the capture rate but simultaneously decreases the translocation time due to the larger electric field experienced by the DNA molecule during translocation [61].

**Using a solution with lower** \( \rho \) **increases** \( \Delta I \) The solution used in these experiments is typically aqueous, although some work has been dedicated to looking at ionic and organic liquids [54, 85]. Typically, there is also an inverse relationship between the conductivity of the solution and the translocation time.
Decreasing $d$ increases $\Delta I$ While this might seem counterintuitive at first, decreasing the diameter can increase $\Delta I$. This continues until the limit where $d \approx d_{DNA}$ which is the minimum diameter required to be able to translocate DNA. In this situation, $\Delta I$ is maximized and can be nearly as large as $I_{open}$. At the other extreme, when $d \gg d_{DNA}, t_{eff}$, $\Delta I$ has an inverse relationship with $d$.

Decreasing $t_{eff}$ increases $\Delta I$ In the limit where $t_{eff} \to 0$ and $d \to d_{DNA}$, $\Delta I$ is maximized. As a result, 2D materials are of particular interest for single-molecule experiments involving nanopores [6, 32]. Decreasing $t_{eff}$ is important for another, more subtle reason. The entire channel constitutes the sensing region of a nanopore. The thinner the channel becomes, the higher its spatial resolution. For example, the spacing between nucleotides in DNA is approximately 0.34 nm. A 10 nm thick nanopore would have $\approx 30$ nucleotides contributing to the signal current at the same time whereas only 3 nucleotides would contribute to the signal current if the nanopore were only 1 nm thick. However, a thinner pore also results in a larger electric field for the same bias voltage, which leads to shorter event durations. Simulation studies suggest that the optimal $t_{eff}$ for nucleotide sensing might not be that of the absolute thinnest nanopore that can be fabricated [86].

Signal duration

The duration of the pulse-like translocation event is determined by the amount of time the molecule spends in the nanopore. For nanopores with thicknesses < 10 nm, the characteristic diffusion time is < 500 ns [57]. As a result, most such events should have been invisible at the kHz bandwidths afforded by off-the-shelf amplifiers. However, translocation time is heavily governed by interactions between the molecule and the pore walls [61, 87]. The translocation velocity profile is also non-linear over the
course of the molecule’s stay in the nanopore [76, 88, 89]. Further, DNA is known to translocate orders of magnitude faster through solid-state nanopores than through biological nanopores of comparable dimensions [90]. As a result, significant effort has been directed toward altering the surface chemistry of nanopores in order to increase translocation time. This includes the use of both organic [47, 91–93] as well as inorganic coatings [43, 44] over the pore surface.

An alternative strategy has been used successfully with biological nanopores in order to control DNA translocation. A polymerase enzyme is incorporated directly on top of the nanopore so that it traps the DNA molecule before it enters the pore. The DNA is then “ratcheted” through the pore one base at a time at speeds of the order of a few bases per ms [94]. A superimposed triangle wave on the bias voltage then allows for stretching and compressing the DNA that is in the nanopore [78].

When a voltage bias is applied across the nanopore membrane, it establishes an electrical gradient in the nanopore channel. A non-electrical gradient in the opposite direction could thus be used to potentially slow down translocation. Alternatively, a gradient in the same direction as the field can increase $\Delta I$ without affecting translocation times. This has been successfully exploited by using chemical [95], temperature [61, 63, 96], and viscosity [54, 97, 98] gradients.

KCl is the most commonly used electrolyte in nanopore experiments because of its high conductivity. However, there usually is a trade-off between the conductivity of the electrolyte and the corresponding translocation times [99]. LiCl is a promising alternative and has been reported to slow down DNA translocation without significantly altering the conductance [100]. Other organic salts and ionic liquids have also been investigated [54, 98] albeit the conductance values in these cases are invariably lower than that when KCl is used.
2.3.4 Modeling noise

Consider the simplified measurement setup for a nanopore sensor using an operational amplifier (opamp) as shown in Figure 2.3. Instrumentation amplifiers used for recording nanopore signals are typically linear systems. Thus, the input current $I(t)$ is related to the output voltage $V(t)$ in the frequency domain as

$$V(f) = I(f)H(f)$$

(2.15)

where $V(f)$ and $I(f)$ are the Fourier transform representations of $V(t)$ and $I(t)$ respectively. Thus, a measured output voltage can be referred back to the input as a measured current if $H(f)$ is known. While not required, $H(f)$ is usually designed to be constant as a function of frequency in the bandwidth of interest. Note that here $H(f)$ has units of ohms, and hence, such an amplifier is known as a transimpedance amplifier (TIA).

Recall from Equation 2.6 that $I(t)$ is composed of both the signal of interest and added noise. The noise here reflects the cumulative effect of all the possible noise sources in the system referred to the input. Since different systems can have different gains ($|H(f)|$), it is customary to refer the noise to the input to enable fair
comparisons between systems. Such “input-referred noise” is also useful in order to
determine how much signal strength will be required from the sensor for a given noise
specification.

Given an output noise $v_{n,\text{out}}(t)$ and its corresponding PSD $S_{n,\text{out}}(f)$, we can express
the PSD of the input-referred noise $S_n(f)$ as

$$S_n(f) = \frac{S_{n,\text{out}}(f)}{|H^2(f)|} \quad (2.16)$$

This expression is helpful because it is the output-referred noise that is often the
easiest to measure. Further, it is common for systems to have multiple noise sources
injecting noise at different nodes. Each of these noise sources will have a different
effective contribution to the output noise depending on the transfer function from
the point of injection of the noise to the output. At every node in a system, the
noise sources at that node sum in power if one makes the assumption that these noise
sources are uncorrelated in nature. In general, for a system containing $N$ nodes with
$N$ such noise sources with PSDs given by $X_i(f)$ and transfer functions $H_i(f)$ from
node $i$ to the output, the following equation holds

$$S_{n,\text{out}}(f) = \sum_{i=1}^{N} X_i(f)|H_i^2(f)| \quad (2.17)$$

Note that no constraints have been placed on the dimensions of $X_i(f)$ and $H_i(f)$. Thus, voltage noise sources at certain nodes of a current-measurement setup will
also contribute to the total noise and be referred back to the input as a current noise
through Equation 2.16. Indeed, many noise sources with different origins are common
in nanopore recordings [101].

Note that any noise sources that add directly at the input (such as EMI, poten-
tially) have the same transfer function to the output as the signal itself. Thus, it
becomes critical to particularly reduce the noise PSDs of sources that inject noise at
the input. Further, if the system is designed in such a way that the first stage has
enough amplification, noise added at subsequent stages has less of an impact since
the gain to the output from these stages is lower. This is a direct consequence of Friis’ equation which is well known in the electrical engineering community.

On the basis of their spectral properties, the noise sources common in nanopore recordings can be divided into several categories [101–104].

$1/f^\alpha$ noise

This noise type is dominant at low frequencies and is commonly seen with $\alpha = 1$, although other values for $\alpha$ are also possible. The special case of $\alpha = 1$ is commonly referred to as flicker noise. While the exact physical origins of flicker noise are unknown, it has been observed in a variety of systems [105]. In modern semiconductor processes, it is believed that charge traps at the oxide-silicon interface give rise to flicker noise [106]. Although surface charge fluctuations and nanobubbles [107] have been considered to be a potential flicker noise mechanism in solid-state nanopores, it has been found that flicker noise follows the Hooge phenomenological relationship better [104]. In this sense, the flicker noise mechanism in nanopores seems to be different from that in transistors in modern fabrication processes. The Hooge relationship is expressed as follows:

$$S_{n,flicker,pore}(f) = \frac{\alpha_H I_{baseline}^2}{N_c f}$$

(2.18)

where $\alpha_H$ is the Hooge fitting parameter and $N_c$ is the number of charge carriers in the nanopore. Thus, as the baseline current increases, the flicker noise contribution of the nanopore also increases. It has been found that treating a solid-state nanopore before experiments can help with reducing the overall flicker noise level [108].

There are two other important sources of flicker noise that need to be considered. First, the amplifier itself will introduce flicker noise into the measurement although this reduced through appropriate design choices. Second, the electrode-electrolyte interface is also capable of adding flicker noise [109]. Including these gives the ex-
pression for the total flicker noise as

\[ S_{n,flicker}(f) = S_{n,flicker,pore}(f) + S_{n,flicker,amplifier}(f) + S_{n,flicker,electrode}(f) = \frac{K_{flicker}}{f} \] (2.19)

It is worth noting that flicker noise power as given by the previous equation tends to infinity as the frequency tends to zero. Practical measurements cannot be infinitely long and impose a lower limit on the frequency that is larger than zero just by virtue of the recording length being finite.

We can calculate the RMS value of the noise from a flicker noise source as

\[ I_{RMS,flicker} = \sqrt{\int_{f_1}^{f_2} S_{n,flicker} df} = \sqrt{\int_{f_1}^{f_2} \frac{K_{flicker}}{f} df} = \sqrt{K_{flicker} \ln \left( \frac{f_2}{f_1} \right)} \] (2.20)

Thus, the power of the noise injected by a flicker noise source increases by the same amount for every extra decade of bandwidth.

**White noise**

The PSD for a white noise source remains constant as a function of frequency. This arises because the autocorrelation of the output of a white noise source in the time domain is a delta function. In other words, every time instant of white noise is uncorrelated with every other time instant.

Thermal noise generated by resistors, also known as Johnson-Nyquist noise [106], is an example of a white noise source and is a consequence of the fluctuation-dissipation theorem. The current noise generated by a resistor is given by the equation

\[ S_{n,white,resistor} = \frac{4k_B T}{R} \] (2.21)
where $k_B$ is Boltzmann’s constant and $R$ is the value of the resistance. Since a nanopore is a physical ionic resistor, it generates a proportional current noise. The feedback resistance $R_F$ of the TIA, if implemented using an actual resistor, and the charge-transfer resistance $R_{ct}$ at the electrode-electrolyte interface used for the recordings also generate their equivalent current noise.

Transistors also generate white noise. For a transistor with transconductance $g_m$, its output current noise is given by

$$S_{n,\text{white,transistor}} = 4k_B T \gamma g_m$$  \hspace{1cm} (2.22)

where $\gamma$ is a factor that depends on the region of operation, and is $2/3$ when the transistor is in saturation. The previous equation when divided on both sides by $g_m^2$ gives the equivalent voltage noise at the gate of the transistor. This transformation is useful because the noise performance of opamps, which form the basic building block of TIAs, is frequently described by their input-referred voltage noise PSD.

Shot noise is yet another noise source that is spectrally white. It is a consequence of the quantized nature of charge and occurs whenever the number of charge carriers is low enough that Poisson statistics govern overall transport. Shot noise is common in semiconductor junctions such as those found in diodes or diode-connected transistors, and becomes relevant particularly when the current levels are low. Theory suggests that ion channels may also exhibit shot noise \[110\]. The expression for shot noise is given by

$$S_{n,\text{white,shot}} = 2qI_{DC}$$  \hspace{1cm} (2.23)

where $I_{DC}$ is the steady-state current flowing through the device and $q$ is the elementary charge.

Finally, solid-state nanopores exhibit yet another white noise source attributed to surface charge fluctuations \[111\]. The exact amount of current noise due to this source depends on several factors such as $I_{\text{baseline}}$, the electrolyte concentration and the pH of the solution. Interestingly, the dependence on $I_{\text{baseline}}$ is quadratic.
In general, the total white noise can be expressed as

$$S_{n,\text{white}} = K_{\text{white}}$$

(2.24)

where $K_{\text{white}}$ is the sum of all the white noise sources of concern. The RMS value of such a noise source then becomes

$$I_{\text{RMS,white}} = \sqrt{\int_{f_2}^{f_3} S_{n,\text{white}} df}$$

$$= \sqrt{\int_{f_2}^{f_3} K_{\text{white}} df}$$

$$= \sqrt{K_{\text{white}}(f_3 - f_2)}$$

(2.25)

Compared to flicker noise, the RMS value of a white noise source grows faster still when the bandwidth is increased.

*f noise*

The primary cause of noise PSD with a linear dependence on the frequency has been attributed to losses in the membrane dielectric of the nanopore [103]. In general, this can be expanded to all capacitances present at the input of the TIA, of which the membrane capacitance is usually the dominant contributor. For a capacitor with capacitance $C$ and a loss tangent $\tan \delta$, the voltage noise is given by [112]

$$S_{n,f,\text{capacitor}} = 8\pi k_B T f C \tan \delta$$

(2.26)

From Equation 2.16, we note that flicker noise can also transform to appear as $f$ noise if $H(f) \propto 1/f$, as is the case for a capacitor. Thus, the flicker noise of the TIA can appear in the input-referred noise PSD as $f$ noise too.

In general, the total $f$ noise can be expressed as

$$S_{n,f} = K_{f,f}$$

(2.27)
where $K_f$ is the appropriate scaling factor to account for all the $f$ noise contributors. The RMS value of such a noise source then becomes

$$I_{RMS,f} = \sqrt{\frac{\int_{f_3}^{f_4} S_{n,f} df}{f_3}} = \sqrt{\frac{\int_{f_3}^{f_4} K_f f df}{f_3}} = \sqrt{\frac{K_f}{2} (f_4^2 - f_3^2)}$$

$$\approx \sqrt{\frac{K_f}{2} f_4} \quad \text{if } f_4^2 \gg f_3^2$$

(2.28)

$f^2$ noise

There is no physical noise source in a standard nanopore measurement setup that adds $f^2$ noise directly. However, there are two noise sources whose transfer functions transform them into a noise source with an $f^2$ dependence on frequency. These are the thermal voltage noise generated by $R_{ct}$ and the TIA [59, 101]. Consider the input-referred noise source of the TIA ($v_n$) shown in Figure 2.3. The equivalent current noise PSD generated by this source at high frequencies is

$$S_{n,f^2,TIA} = 4\pi^2 v_n^2 C^2 f^2$$

(2.29)

$C$ in the equation above can be replaced by $\sum C_i$ to include all the parasitic capacitances at the amplifier’s input. This usually includes the following:

- $C_{pore}$ - the pore’s membrane capacitance
- $C_W$ - the wiring capacitance associated with the connection between the pore and the amplifier.
- $C_I$ - the amplifier’s input capacitance
- $C_F$ - the capacitance in the TIA feedback network
The combined effect of all these parasitic capacitances is to modify the previous equation into

\[ S_{n,f,TIA} = 4\pi^2 v_n^2 \left( \sum C_i \right)^2 f^2 \]

\[ = 4\pi^2 v_n^2 (C_{\text{pore}} + C_W + C_I + C_F)^2 f^2 \quad (2.30) \]

In general, the total \( f^2 \) noise can be expressed as

\[ S_{n,f^2} = K_{f^2} f^2 \quad (2.31) \]

where \( K_{f^2} \) is the appropriate scaling factor to account for the dominant \( f^2 \) noise sources. The RMS value of this source then becomes

\[
I_{\text{RMS},f^2} = \sqrt{\int_{f_4}^{f_5} S_{n,f^2} df} = \sqrt{\int_{f_4}^{f_5} K_{f^2} f^2 df} = \sqrt{\frac{K_{f^2} (f_5^3 - f_4^3)}{3}}
\]

\[ \approx \sqrt{\frac{K_{f^2} f_5^3}{3}} \quad \text{if} \quad f_5^3 \gg f_4^3 \quad (2.32) \]

Of all the noise sources, the RMS value of an \( f^2 \) noise source rises fastest with frequency. It is the noise source that invariably dominates high frequency noise and is the limiting factor for further extending measurement bandwidths.

### 2.3.5 SNR-constrained bandwidth

The net current noise as a function of all the noise sources discussed in the previous section can be expressed as

\[ S_n(f) = S_{n,flicker}(f) + S_{n,white}(f) + S_{n,f}(f) + S_{n,f^2}(f) \quad (2.33) \]
This results in a classic bathtub shaped curve when the input-referred current noise PSD is plotted against frequency on a log-log scale and is illustrated in Figure 2.4. Note that since all of these sources are assumed to be uncorrelated to each other, their powers add up. Thus, beyond a certain frequency, the flicker noise, for example, does not disappear, but merely becomes lower in power compared to the white noise such that the white noise ends up determining $S_n(f)$ at that frequency.

If the conditions are appropriate to operate in the capacitance dominated noise regime, then the maximum achievable bandwidth $B_{\text{max}}$ for a given $\text{SNR}_{\text{min}}$ can be written using Equations 2.7 and 2.30 as

$$I_{\text{RMS}} = \frac{\Delta I}{\text{SNR}_{\text{min}}} \approx \frac{2\pi v_n \sum C_i B_{\text{max}}^{3/2}}{\sqrt{3}}$$

$$\therefore B_{\text{max}} = \left( \frac{\sqrt{3} \Delta I}{2\pi v_n \text{SNR}_{\text{min}} \sum C_i} \right)^{2/3} \quad (2.34)$$

For a given $\text{SNR}_{\text{min}}$, there are three ways in which $B_{\text{max}}$ can be increased:

**Increase $\Delta I$** This has been discussed in Section 2.3.3.

**Decrease $v_n$** This can be achieved through careful design of the amplifier and will be discussed in the following sections.
Decrease $\Sigma C_i$ This can be achieved in several ways and will be discussed in the following sections.

2.4 Intracellular recordings

Cells are the basic unit of life. They form the base of the hierarchy consisting of tissues, organs, organ systems and organisms, in that order. Thus, understanding the functioning of the cell itself is crucial to understanding the functioning of the upper levels of the hierarchy.

We know now that the cell consists of an impermeable membrane that interacts with the environment around it through ion channels. These channels are responsible for regulating the relative concentrations of ions inside the cell by virtue of their ion selectivity and gating behavior. These ionic fluxes regulate key cellular functions such as muscle contraction, hormone and neurotransmitter release and growth and death of the cells themselves [4, 5].

This information on ion channels and the critical role they play in cellular functioning was found out through recordings made using the patch-clamp technique. Patch-clamping was developed by Erwin Neher and Bert Sakmann in the late 1970s and early 1980s for which they were awarded the Nobel Prize in Physiology or Medicine in 1991. Despite the several decades that have passed since then, it is still considered the gold standard for intracellular recordings - a testament to the simplicity and elegance of the technique itself.

While it was originally developed to study single ion channels [113], the technique has since been adapted into a variety of different configurations for different kinds of experiments. One such configuration is the so-called “whole-cell” patch-clamp which has enabled studies of intracellular electrochemistry [114].

Although the focus here is on single-electrode patch-clamping, other techniques such as two-electrode voltage-clamping [115], discontinuous single-electrode voltage-
clamp [116] and active electrode compensation have also been explored [117, 118].

2.4.1 Configurations

A typical patch-clamp setup consists of a glass micropipette pulled to diameters of a few µm or lower mounted on a micromanipulator and connected to an amplifier. The pipette is then lowered into a bath containing the cell of interest. The major configurations of the patch-clamp setup are discussed here but several other variations such as the perforated patch and loose patch also exist.

Cell-attached

As the pipette approaches the cell membrane, application of suction pulls the membrane into the pipette forming the so-called “gigaseal” where the leak resistance from the interior of the pipette to the bath becomes in excess of $1\,\text{GΩ}$. This is known as the cell-attached configuration and is the precursor to all the other configurations discussed here. Obtaining the gigaseal is crucial to achieving low-noise recordings because it isolates the ion channel activity from the potentially large background currents [119]. This configuration is used to study currents through ion channels in the patched region of the membrane.

Inside-out

Once the gigaseal is formed, retracting the pipette can tear off the patch of membrane that has been sucked into the pipette without losing the gigaseal. This now exposes the “inside” part of this patch of membrane to the outside, i.e., the bath. Hence, this is known as the “inside-out” configuration. The bath solution can be changed to study the dependence of the behavior of ion channels in the excised patch on the cytosolic environment.
Whole-cell

If instead of retracting the pipette, additional suction or a voltage pulse is applied in the cell-attached configuration, the membrane can be ruptured allowing for access to the internal chemical environment of the cell. This is known as the “whole-cell” configuration. This can be used to study ensemble ion channel response to a particular electrochemical stimulus applied to the cell.

Outside-out

Retracting the pipette after achieving the whole-cell configuration can cause resealing of both the cell membrane and the excised patch. However, now the “outside” part of the excised patch is exposed to the bath giving rise to the name “outside-out”. This configuration can be hard to achieve but is useful for studying ion channel activity for changes in the extracellular medium.

2.4.2 Measurement modes

There are two primary electrical parameters of interest in patch-clamp recordings depending on the configuration. These are the current flowing through the ion channels and the intracellular potential with respect to that of the surrounding bath. Consequently, there are two primary measurement modes.

Current-clamp

In this mode, the current through the pipette and, by extension, the membrane patch or the entire cell membrane, depending on the configuration, is held constant and the pipette’s voltage is measured. In the whole-cell configuration, this corresponds directly to the intracellular voltage. A special case is when the current is held fixed at zero, and is called the $I = 0$ current-clamp. By varying the magnitude and duration
of the current injected into the cell, the cellular response to a current stimulus can be obtained.

**Voltage-clamp**

In this mode, the voltage inside the pipette is held fixed and the current through the pipette is measured. Depending on the configuration, this could either be the current through a single ion channel or through all the ion channels in the cell membrane. As in the case with the current-clamp, the voltage can be stepped to either hyperpolarize or depolarize the cell in order to trigger voltage-gated ion channel activity. Note that as per this description, the measurement is exactly identical to that obtained while measuring currents through nanopores as described in Section 2.2.

### 2.4.3 Modeling the interface

#### Modeling the cell

In the simplest case, the cell membrane can be thought of as a resistor $R_M$ in parallel with a capacitor $C_M$. In the absence of any ion channels, $R_M$ is usually fairly large and $C_M$ is determined by the area and the specific capacitance of the cell membrane which is known to be $\approx 1 \mu F/cm^2$, and is between 0.3 and 0.8 $\mu F/cm^2$ for lipids commonly used for forming bilayers [120].

The ion channels themselves typically have a voltage response associated with them. A simple equivalent circuit for a potassium selective ion channel is shown in Figure 2.5 [4]. $R_M$ is not shown here because ion channel conductances generally dominate the low conductance offered by the membrane itself. In general, the conductance of the ion channel is governed by several different electrical and chemical factors. Depending on the channel, the associated “reversal potential” (shown here by $E_K$) could be either positive or negative.
Another important metric of interest is the “resting membrane potential” (RMP). The cytosol in a cell consists of many different kinds of ions such as Ca$^{2+}$, K$^+$, Na$^+$ and Cl$^-$. These ions also exist in the bath solution that the cell is in, although not necessarily at the same concentration. In fact, there exist ion pumps in the cell membrane that act to regulate the concentration gradient of a particular ion. In equilibrium, each of these ions has an associated Nernst potential that contributes to the net electrochemical potential inside the cell which is known as the RMP. The RMP of a cell can be measured in current-clamp mode with no current passing through the pipette.

Consider the equivalent circuit of a cell membrane containing sodium, potassium and leak channels and its simplified representation shown in Figure 2.6. It is clear that the contribution of the various reversal potentials to the RMP depends on the respective conductances, which is governed by the number of corresponding ion channels in the membrane that are open at rest. All excitable cells have a negative RMP since the proportion of ion channels that are selective to ions with a negative reversal potential is usually higher than those selective to ions with a positive reversal potential [4].
Modeling the pipette

The pipette used in patch-clamp experiments is a glass capillary pulled to diameters of a few µm as is the case with patch pipettes, and 10’s to 100’s of nm as is the case with nanopipettes. The pipette is filled with an appropriate solution containing electrolytes with an electrode immersed in the solution providing a (usually) Faradaic interface. The pipette is commonly made out of borosilicate glass or quartz.

The physical geometry of the pipette lends it two important parameters of interest

- The series resistance of the pipette $R_S$
- The parasitic capacitance to the bath $C_p$

The series resistance is a result of the tapering down of the capillary to small diameters and because the filling solution has finite conductivity. It is common for $R_S$ to be a few MΩ for patch pipettes and can be as large as a few hundred MΩ for nanopipettes. As the pipette diameter is increased, $R_S$ decreases but it becomes difficult to achieve gigaseals and target small structures. On the other hand, decreas-
ing diameters afford the ability to investigate smaller structures but at the cost of increased $R_S$.

A related resistance is the access resistance which seeks to increase $R_S$ once the cell has been patched. The access resistance arises because while the bath acts like an infinite reservoir, the cell does not. Further, the cytosol in the cell is not as conductive as the bath in typical measurements. In whole cell recordings, part of the cytosol mixes with the filling solution in the pipette at the tip and serves to decrease the local conductivity.

Once the cell has been patched, there is yet another resistance of interest. This is the seal resistance from the pipette to the bath $R_{\text{seal}}$. As discussed previously, the formation of a gigaseal implies $R_{\text{seal}} > 1 \, \text{G} \Omega$, but measurements are still possible even when this is not the case. Depending on the value of $R_{\text{seal}}$, these are referred to as the “loose-seal” or the “tight-seal” configurations.

The capacitance to the bath is due to the glass acting as a dielectric between two conductive solutions. Since the glass is usually significantly thicker than the Debye length at the electrolyte concentrations of interest, the capacitance is primarily governed by the geometry of the pipette and the extent to which the pipette is lowered into the bath. Coating the tip of the pipette with Sylgard is also known to decrease $C_p$ by locally increasing the dielectric thickness [5]. $C_p$ is normally of the order of a few pF and can be reduced by decreasing the height of the pipette immersed into the solution. This is also desirable for reducing movements in the pipette tip.

Although $R_S$ and $C_p$ are distributed entities, it is easier from a design perspective to treat them as lumped elements. This approximation is commonly used and holds up well when compared to the distributed model. Figure 2.7 shows the equivalent electrical circuit approximation of a cell combined with a pipette.
2.4.4 Need for compensation

In the ideal scenario, $R_{\text{seal}}$ is arbitrarily large whereas $R_S$ and $C_p$ both tend to zero. Practically, however, it is only $R_{\text{seal}}$ that can be made close to the ideal case by achieving a gigaseal. The impact of the residual $R_{\text{seal}}$ is negligible once the gigaseal has been achieved.

On the other hand, $R_S$ and $C_p$ are both finite and non-zero and their impact on the intracellular measurement must be evaluated for both current- and voltage-clamp modes. The input signal to be measured can be modeled as a voltage source in parallel with $C_M$ in current-clamp and a current source in the same position in voltage-clamp.

Parasitic capacitance

In current-clamp, $C_p$ serves to create an R-C filter along with $R_S$ and attenuate the high frequency content of the input signal. Considering nominal values of $R_S$ and $C_p$ as $10 \text{ M}\Omega$ and $5 \text{ pF}$, the low-pass filter formed by the R-C pair has a cutoff frequency
of $1/(2\pi R_S C_p) \approx 3$ kHz. Further, a large value of $C_p$ requires a proportionally large amount of charge to change its voltage, such as one that might occur when the intracellular potential changes showing cell activity. Since this charge needs to be provided by the cell and given the cell’s finite volume, this might serve to change the internal concentrations of ions within the cell thereby affecting the validity of the measurement.

In voltage-clamp, as the voltage in the pipette is stepped up and down, $C_p$ injects large spikes of current as the voltage across it is instantaneously changed. This can obscure any small signal currents that might be present at the onset of the voltage step. Another problem caused by residual $C_p$ is in loop stability. TIAs are commonly used as voltage-clamp front-end amplifiers and can be destabilized by large capacitors present at their inputs.

**Series resistance**

As discussed previously, $R_S$ in current-clamp acts in conjunction with $C_p$ to filter the input signal and suppress its high frequency content. Further, if a non-zero current is applied through the pipette, it creates a proportional voltage drop across $R_S$ and leads to a difference between the true and measured membrane voltages. In general, any leakage current flowing through $R_S$ will cause such a voltage drop and needs to be compensated for.

The effect of $R_S$ in a voltage-clamp recording manifests in three primary ways [121]:

**Filtered $V_M$** When a step change is applied at the pipette voltage $V_p$, the voltage that appears at $V_M$ is filtered through the R-C filter formed by $R_S$ and $C_M$. Thus, the voltage change experienced by the cell will be slower than the desired step. This may lead to measurements that do not accurately reflect the cell’s response to a step change in $V_M$. 

37
Voltage offset Any bias or signal currents flowing through $R_S$ will cause a corresponding voltage drop as in the current-clamp case. However, since the pipette voltage is held constant by the TIA, this will cause the intracellular voltage to change thus defeating the original intention of clamping it.

Signal current filtering Any desired signal of interest also sees the filtering due to $R_S$ and $C_M$. For large enough signal currents, transient changes in $V_M$ caused by this filtering may be large enough so as to lead to complete loss of clamping ability. This is commonly observed with whole-cell measurements of Na$^+$ activity in neurons, and requires the use of either low-resistance pipettes or appropriate compensation circuits in the amplifier. Even if the clamp is retained, the measured signal will be heavily filtered. For example, with $R_S$ and $C_M$ being 25 MΩ and 30 pF respectively, as is typical in whole-cell experiments, the signal bandwidth is a mere 212 Hz.

2.5 Multi-clamp amplifiers

2.5.1 Commercial solutions

Both current- and voltage-clamp recordings require a front-end amplifier with an extremely low input leakage current since the cell, given its limited volume, will not be able to sustain large currents for very long on its own.

The current-clamp front-end fundamentally consists of a voltage buffer with a high-impedance input and typically has a gain of unity. Intracellular current-clamp inputs are several mV large and do not require much amplification. On the other hand, extracellular voltage measurements have much smaller amplitudes and will need further amplification. If the same amplifier is to be used for both kinds of recordings, some gain in the voltage buffer is desirable. An effective current-clamp
front-end also needs capacitance compensation as discussed in Section 2.4.4. Current injection in the pA – nA range is common [122].

The voltage-clamp front-end consists of a TIA. The transimpedance element typically consists of a resistor although capacitive TIAs have also been explored to eliminate the thermal noise contributed by $R_F$ [121]. However, the capacitive feedback system generally needs to be reset frequently in the presence of bias currents. Thus, it is generally used for cell-attached recordings whereas the resistive feedback is preferred for whole-cell recordings. TIAs with resistive feedback usually also have a $C_F$ in parallel with $R_F$ to aid with stability. Even if not introduced by design, some amount of parasitic $C_F$ is inevitable. This introduces a low-frequency pole in the TIA’s response with the 3-dB cutoff given by $1/(2\pi R_F C_F)$. While large values of $R_F$ are desirable to reduce its input-referred thermal noise contribution, this also decreases the low-pass cutoff frequency [123]. For example, if $R_F = 1 \text{G} \Omega$ and $C_F = 0.1 \text{pF}$, the pole of this R-C filter is at $\approx 1.5 \text{kHz}$. Additional filters are then required after the TIA in order to correct its frequency response and extend its measurement bandwidth. This is achieved by introducing a zero at the exact frequency as the TIA’s pole [124].

Due to the low input-leakage and capacitance requirements of the headstage, the opamp used for implementing the buffer and the TIA often uses discrete junction field effect transistors (JFETs) as its input pair [124]. The feedback and current injection resistors are discrete resistors and are usually in a range of values to allow for configurability depending on the application. The headstage is also sometimes cooled in order to reduce the noise since most noise sources have a linear dependence on the absolute temperature.

A simplified implementation of the capacitance compensation circuitry is shown in Figure 2.8. The parasitic capacitance $C_p$ sinks a current of $C_p dV_{\text{command}}/dt$. In the absence of capacitance compensation, this current is provided by the TIA and flows through $R_F$, thereby corrupting the actual measured current. With the capacitance
compensation circuit activated, it can be shown that

\[ I_{inj} = (\alpha - 1)C_{inj} \frac{dV_{command}}{dt} \]  \hspace{1cm} (2.35)

Thus, when \((\alpha - 1)C_{inj} = C_p\), the current required by \(C_p\) is provided by the capacitance compensation circuit instead of the TIA and the charging transients associated with \(C_p\) are removed from the TIA’s output. The input to the gain block \(\alpha\) is sometimes delayed in order to account for the TIA’s response time [5, 124]. Although it is possible for either \(\alpha\) or \(C_{inj}\) to be varied to achieve the same end result, discrete solutions typically implement a few fixed values of \(C_{inj}\) to provide coarse tuning, while fine tuning is achieved through changes in \(\alpha\).

![Basic capacitance compensation circuit](image)

**Figure 2.8: Basic capacitance compensation circuit [5]**

Most commercial systems employ positive feedback in order to implement \(R_S\) compensation. Pre-conditioning the command voltage to account for the filtering of \(V_M\) by \(R_S\) and \(C_M\) and scaling the command voltage to account for the voltage drop on \(R_S\) account for the first two problems discussed in Section 2.4.4. These solutions
do not require feedback but do not address the third, and perhaps, the most crucial problem of signal filtering.

The basic principle of using positive feedback to compensate for $R_S$ relies on adding a scaled version of the current monitor signal to the command voltage. The scaling factor determines the amount of $R_S$ compensation. Figure 2.9 shows a simplified implementation of positive feedback $R_S$ compensation. Additional filtering is required in order to stabilize the loop. Although the diagram depicts that the TIA is ideal and has infinite bandwidth, practical limitations to the bandwidth discussed previously in this section act to destabilize the loop. Further, $C_p$ must be compensated almost entirely prior to $R_S$ compensation. In practice, the residual $C_p$ must be $< 100\,\text{fF}$ in order for high $R_S$ compensation to remain stable [5]. Loop instability, which results in oscillations on $V_p$ can kill the cell if the oscillations are large enough.

![Figure 2.9: Basic resistance compensation circuit using positive feedback [5]](image)

2.5.2 Integrated approaches

While commercial systems are able to achieve current- and voltage-clamp recordings with high SNR, they are bulky and expensive systems that primarily use discrete components. The sheer size of these systems limits the ability to scale up the number of recording sites for experiments where multiple simultaneous measurements are
required. For example, given the importance of ion channel testing for pharmaceutical studies, having high throughput can enable faster discovery of new drugs as well as reduced testing times for existing drugs. This need for miniaturization has thus spurred the development of integrated implementations of the patch-clamp amplifier.

One of the earliest such implementations consisted of a voltage-clamp amplifier with a capacitive front-end [125], but did not include any compensation circuitry. Resistive-feedback front-ends were also explored [126]. This was later expanded into a full-fledged voltage-clamp setup with capacitance and resistance compensation circuitry and two-channels integrated onto the same die [127], and then further to include current-clamp capability and four-channel operation as well [128]. The primary drawback with this implementation was the extremely high noise levels – as much as an order of magnitude larger than commercial solutions.

One recent integrated multi-clamp solution used a diode in the feedback path of the voltage-clamp in order to accommodate a wide range of input bias currents [129]. The completely digital interface allowed for integration with an external computer to enable automated patch-clamp experiments [130].
Chapter 3

Design of a Low-Noise Nanopore Front-End

3.1 Introduction

The simplest possible circuit that can clamp the voltage across a nanopore membrane is a simple voltage source. A voltage source, however, does not inherently have the ability to measure the current flowing through it. In that sense, a TIA offers a virtual ground – a fixed voltage with respect to the TIA’s ground – and provides any current required in order to clamp the voltage at the virtual ground.

TIAs have been the subject of active research in the analog circuits community for several decades and have found popular use in electrophysiology setups as the voltage-clamp amplifier [5]. Indeed, several high-performance TIAs are available as
commercial products for integration into board-level designs particularly when the sensor is a photodiode.

The need for application-specific designs and miniaturization has directed several efforts over the last few decades towards designing TIAs using integrated circuits (IC), including several designs targeting nanopore recordings specifically [125, 131–141] while other efforts have eschewed custom ICs for custom-designed boards using commercially available ICs [38, 45, 142, 143]. Realization of large-valued resistors in CMOS processes continues to remain an area of active research. Fabricating passive large-valued resistors is impractical both in terms of the area consumption and due to the associated parasitics [127].

Reducing the size of TIAs is favorable for nanopore recordings because it is generally accompanied by a reduction in parasitics which is necessary for reducing $\Sigma C_i$, and thus extending measurement bandwidth as discussed in Section 2.3.5. This needs to be accompanied by a simultaneous reduction in the noise generated by the TIA.

## 3.2 Design considerations

There are several factors that need to be considered while designing a low-noise TIA, some of which dictate opposite design decisions. The TIA discussed here uses an output transconductance amplifier (OTA) as an opamp and implements resistive feedback.

### 3.2.1 Input-referred voltage noise

The noise performance of a TIA is typically characterized by referring its output noise to the reference input as a voltage noise $v_n(f)$. In a well-designed system, the noise is dominated by that contributed by the first opamp and, in particular, by the transistors that form the input differential pair of the opamp. For the circuit shown in Figure 2.3, $v_n$ transforms into a noise current source through the net impedance at
the input to the TIA. As discussed in Section 2.3.4, this is the dominant noise source at high frequencies and hence, reducing $v_n$ is of critical importance.

$v_n$ at high frequencies for a well-designed opamp is governed by the thermal noise generated by the input pair transistors. As discussed in Section 2.3.4, the thermal noise of a transistor depends strongly on its $g_m$. Increasing the $W/L$ (width to length ratio) of a transistor places it in its subthreshold regime of operation which maximizes $g_m$ for a given bias current flowing through the transistor. Beyond this, improvements in $g_m$ for a given technology can only be achieved by increasing the bias current itself.

It is worth noting that flicker current noise increases as the bias current through the transistors is increased. However, when referred to the transistor’s input, flicker noise scales inversely as the area and is largely independent of the current [106]. Thus, input transistors with high $g_m$ and high $W/L$ are desirable.

### 3.2.2 Unity-gain bandwidth

While opamps are often idealized as having infinite gain across all frequencies and zero response time, practical opamps are limited in terms of the input frequencies they can respond to. The unity-gain bandwidth (UGB) of an opamp in a feedback loop is defined as the frequency where the magnitude of the loop gain equals unity. Opamps can be modeled as having a single pole response as

$$A(s) = \frac{A_0}{1 + s/\omega_p}$$

where $A_0$ is the open-loop gain at DC and $\omega_p$ is the location of the pole. The UGB is then equal to $\omega_u = A_0\omega_p$.

The UGB is an important metric because it helps determine the maximum bandwidth that the closed-loop system can operate at. For example, for the feedback loop
shown in Figure 2.3, the loop gain can be written as

\[ L(s) = A(s) \frac{G_F + sC_F}{G_F + G + s(C_F + C)} \]  
\[ \approx A(s) \frac{C_F}{\sum C_i} \]

where the approximation holds at moderate frequencies and if the zero appears before the UGB of the opamp itself – which would be required by stability. Thus, a high \( \omega_u \) would ensure a high operating bandwidth for the closed-loop system as well.

In traditional unity-gain stable opamp design, \( \omega_u \) is often of the form \( g_m/C \) where \( g_m \) is the transconductance of the transistors that form the input pair and \( C \) is some capacitance. While \( \omega_u \) can be increased by appropriately manipulating \( g_m \) and \( C \), in practice, it is often limited by parasitic poles affecting phase margin. Through careful design, it is possible to design opamps that are unity-gain stable up to several hundreds of MHz.

### 3.2.3 Input and feedback capacitance

The input capacitance of the amplifier \( C_I \) is seldom an intentional introduction. It is a parasitic capacitance that is a result of the devices that are connected to the input of the TIA, and of the interconnects themselves. By careful design, \( C_I \) can be reduced to sub-pF levels, but it can never be eliminated completely. Further, the contribution of interconnects to \( C_I \) can be significantly lowered in IC-based TIAs. Larger transistors increase \( C_I \) but may be necessary to achieve other design goals. In general, decreasing \( C_I \) reaches a point of diminishing returns since its effect on noise performance is almost always due to its contribution to \( \sum C_i \). Thus, if other terms dominate \( \sum C_i \), then reducing \( C_I \) will decrease noise but not significantly.

The feedback capacitance \( C_F \) is necessary for loop stability as discussed previously. Thus, \( C_F \) needs to be large enough to ensure stability and not affect the UGB of the feedback loop significantly, but also small enough in order to not increase noise levels.
by increasing $\sum C_i$ and to not adversely affect the closed-loop bandwidth determined by $1/(2\pi R_F C_F)$.

### 3.2.4 DC gain

There are two DC gains of interest while discussing a TIA’s performance. The first is the DC gain of the TIA itself, which, when the opamp is assumed to be ideal, is simply $R_F$. The second is the DC gain of the opamp $A_0$. Larger values of $A_0$ ensure a successful voltage clamp. For example, for a TIA with its reference node set to 0 V, a current that induces a 1 V drop across $R_F$ would require that the input terminals of the opamp maintain a difference of $1/A_0$ V. Thus, the TIA clamps the voltage at a slightly different voltage than the desired 0 V. For $A_0 = 1000$, this would result in a magnitude difference of 1 mV.

Commercial opamps often achieve DC gains that are in the $10^5 - 10^6$ range. At this point, the clamp error becomes of the order of a few µV which is low enough for most applications. The only downside to large DC gains in opamps is that they may become difficult to stabilize.

### 3.2.5 Other considerations

Besides the factors described in the previous sections, there are several other factors that, while not as important, still need to be considered while designing a low-noise TIA.

**Offsets**

All opamps have an offset voltage between their differential inputs which translates to an equivalent difference between the desired and actual clamp voltage. These offsets could be a result of systematic or random mismatches in the transistors that form the input pair of the opamp or through offset voltages induced in the electrode-
electrolyte interface. Random mismatches can be reduced by increasing the size of the transistors while systematic mismatches can be tackled by following good layout practices. The offset voltage is generally not an issue because it is static and can be removed by calibration.

Depending on the kinds of transistors that are used to form the input pair, there might also be a bias current that flows into the input of the TIA. This would appear as an offset current in the measurement. Similar to the offset voltage, this is usually a static quantity and can be removed by calibration. Further, metal-oxide-semiconductor field-effect transistors (MOSFETs) have a primarily capacitive gate and have extremely small leakage currents.

**Dynamic range**

For a TIA implemented using resistive feedback, the upper limit on the dynamic range is usually determined by \( R_F \) and the supply voltage \( V_{DD} \), such that the largest difference between currents that can be measured is \( V_{DD}/R_F \). Thus, a reduced value of \( R_F \) increases the amplitude of the largest current that can be measured.

The lower limit of dynamic range may be determined by measurement noise or by the analog-to-digital converter’s (ADC) resolution. For example, for an ADC with \( N \)-bit resolution, the smallest current that can be measured at the full sampling rate of the ADC is \( V_{DD}/(R_F \times 2^N) \), assuming \( V_{DD}/R_F \) translates to the full-scale of the ADC. This lower limit can be reduced further either by increasing the resolution of the ADC or by oversampling the data and then filtering appropriately.

**Power supply noise**

Noise on the power supply for the TIA can couple into the output in a variety of different ways. Mismatch in the differential branches of the opamp is one such mechanism. Since the reference voltage for the TIA is also usually derived from the power supply, any noise on the reference effectively appears as an additional noise source.
that increases the $v_n$ of the opamp by summing with it in power. Power supply noise can be controlled through the appropriate use of decoupling capacitors and through careful differential design.

### 3.3 Chip design

The TIA implemented here is similar to a traditional opamp-based design with resistive feedback through a feedback resistor $R_F$. Instead of using a passive resistor, the resistance is implemented using an active current-divider scheme [132]. The feedback capacitance $C_F$ is programmable through digital control logic to one of four different values. The single-ended current measurement is converted into a differential voltage signal on-chip in order to increase the dynamic range and resilience to EMI. Figure 3.1 shows a simplified schematic of the amplifier. The chip was designed and fabricated in a commercial 0.18µm CMOS process and contains a 5 × 5 array of TIAs. It achieves $v_n = 2.6$ nV/√Hz while consuming 5.1 mW per channel from a 1.8 V supply.

![Circuit topology for the low-noise nanopore front-end](image)

Figure 3.1: Circuit topology for the low-noise nanopore front-end
3.3.1 Integrator

The integrator is, perhaps, the most critical block in this TIA design. This is because the noise performance of the integrator determines the overall noise levels, if sufficient gain is provided in this stage. The integrator consists of an OTA with a two-bit programmable capacitor in feedback. $C_F$ consists of a fixed 0.05 pF capacitor and a 0.1 pF and 0.85 pF capacitors that can be added in parallel to yield potential $C_F$ values of 0.05 pF, 0.15 pF, 0.9 pF or 1 pF. Including the digital selection logic, the capacitor bank occupies an area of 100 µm × 60 µm.

The circuit schematic of the OTA is shown in Figure 3.2. The OTA is implemented using a two-stage design. The first stage is a telescopic differential OTA with PMOS inputs while the second stage is an NMOS-input common-source amplifier. The load in the first stage consists of current mirrors - with source degeneration using resistors - instead of the classical active load in order to reduce its noise contribution. The inputs are cascoded in order to increase the overall gain of the first stage.

For the differential pair, PMOS inputs were chosen instead of NMOS inputs because of their superior flicker noise performance in this particular technology. In particular, thick-oxide variants of transistors are used in order to reduce input leakage currents to sub-fA levels. The input transistors are sized with $W/L = 800\, \mu m/400\, nm$ and are biased in weak inversion to maximize the $g_m/I_d$, where $I_d$ is the bias current through the transistors. The net input capacitance for these transistors was < 1 pF. Finally, the input transistors are laid out in a common-centroid manner in order to minimize systematic mismatch.

The OTA is unity-gain stable. Stability is achieved through Miller-compensation using a 4.5 pF capacitor. The series resistor provides the zero in the loop gain required to improve phase margin. Under the typical corner, simulations for the OTA showed that it has a UGB of approximately 100 MHz. The OTA occupies an area of 200 µm × 150 µm including the 90 µm × 90 µm compensation capacitor.
3.3.2 Voltage amplifier

The $8\times$ voltage amplifier shown in Figure 3.1 also serves as a single-ended to differential converter. The conversion to differential signaling doubles the maximum current that can be measured by the system. The amplification serves to reduce the effective value of $C_F$ by a factor of eight, which helps improve the closed-loop bandwidth of the TIA. Such reduction is difficult to achieve in practice by physically reducing $C_F$ itself, because capacitor values smaller than the ones used here are prone to significant process variations, and are often dominated by fringe capacitance that can change based on the local routing, rather than the intended parallel plate capacitance.

The amplifier is constructed using an OTA (Figure 3.3) and resistors connected in the standard inverting amplifier configuration. The resistors are sized small so...
that their thermal noise contribution is reduced. This comes at the expense of higher power consumption since a larger current is needed to generate the same voltage drop across a smaller resistance. The OTA is a fully-differential two-stage design. The first stage consists of PMOS inputs and an active NMOS load while the second stage is an NMOS common-source amplifier. Two common-mode feedback loops (CMFB) set the common-mode values at the output of each of the two stages (nominally $V_{DD}/2$). Each CMFB is a simple differential amplifier with active loads.

![Schematic of the fully-differential OTA](image)

**Figure 3.3:** Schematic of the fully-differential OTA used in the voltage amplifier
3.3.3 Output buffer

The purpose of the output buffer is to drive the capacitive load that the PCB will present. Since signal amplification has already been achieved, this stage only buffers the signal and does not further amplify it. This also helps relax the design constraints for this stage. The output buffer is constructed in a similar manner as the preceding voltage amplifier except the resistor ratio is set to provide a gain of unity.

Since this stage might need to drive large capacitances, the OTA is constructed as a single stage differential pair with PMOS input and NMOS active load. The schematic for this OTA is shown in Figure 3.4. A simple CMFB loop sets the common-mode potential of the output to a value that is nominally $V_{DD}/2$.

![Figure 3.4: Schematic of the fully-differential OTA used in the output buffer](image)
3.3.4 Feedback network

It is difficult to realize large-valued linear resistors in CMOS processes. Linear resis-
tors generally do not possess a high enough sheet resistance to implement resistances
larger than 100 kΩ without incurring a significant area penalty. Parasitic capacitance
to the substrate also compromises the high frequency performance of such resistors
[127]. Thus, an active circuit similar to one proposed previously [131] is used to
implement a large-valued resistor in this work.

The feedback network consists of a variable attenuator followed by a resistor and
an active current-divider. The attenuator is formed using an opamp in inverting gain
configuration with gain $1/A < 1$ and is built off-chip. The resistor, nominally 10 kΩ,
is on-chip and converts the voltage output of the attenuator into a proportional current.
The effect of the attenuator, thus, is to make the resistor appear $A$ times larger than
it physically is.

The current-divider is formed using an OTA and diode-connected PMOS transis-
tors as shown in Figure 3.5. When $I_{in}$ is positive, the transistors operate in weak
inversion whereas when $I_{in}$ is negative, the bulk-source junction becomes forward-
biased. In both cases, if the input and output nodes are held at the same voltage,
then the transistors experience the same voltage and the relative current conduction
becomes purely a function of the ratio of the $W/L$ of the transistors. This scheme
requires careful layout of the devices such that mismatches between the transistors
are minimized. Since the smaller PMOS is directly connected to the input, its flicker
noise and shot noise are directly added to the input.

In this work, $M$ was set to 200 so that $A = 3.75$ yields $R_F = A \times M \times 10 \text{kΩ} =
7.5 \text{MΩ}$ and $R_F = 45 \text{MΩ}$ when $A = 22.5$. It is worth noting that this active resistor
implementation has thermal current noise lower by a factor of $M$ than a passive
resistor of the same value. However, this noise benefit comes at the expense of
increased power dissipation. The OTA used in the current-divider is similar to that
used in the integrator except the first stage is a folded-cascode differential pair with PMOS inputs.

### 3.3.5 On-chip electrode

Each amplifier channel has an associated $100 \, \mu m \times 100 \, \mu m$ electrode connected to the TIA input located adjacent to the amplifier itself. This is in addition to an input that is connected to bondpads on the I/O ring located on the periphery of the chip. The area under the electrode itself is kept free of routing and transistors in order to minimize the parasitic capacitance at this node. Further, the proximity of the surface electrode helps reduce $C_W$ to a few fF. The top metal layer in this process is 4 µm-thick aluminum, and is located approximately 9 µm above the substrate. Post-processing of this electrode for interfacing with salt solutions is discussed in Chapter 4.

### 3.3.6 Digital logic, ESD and biasing

Each of the 25 channels is individually programmable. The available programming options include the following:

- Turning the channel on/off through control over the bias voltages
- Connecting the on-chip electrode to the TIA input
• Connecting the TIA input to a bondpad in the I/O ring

• Two-bit programmability of $C_F$ including shorting $C_F$ out

This control is achieved by means of a channel-local eight-bit scan-chain. The scan-chains are implemented using complementary logic which operate from the same 1.8 V supply as the analog circuitry. Since the scan-chain is silent during actual recording, noise from the digital circuitry affecting the analog blocks was not a concern. Buffers were introduced between the scan-chain blocks of each channel to add delay and protect against potential hold-time violations. Although the digital logic operates at 1.8 V, the digital I/O pads available for this process operate at 3.3 V. Level-shifters are added to convert data between the two voltage domains in order to ensure compatibility.

Electrostatic discharge (ESD) protection is necessary in modern CMOS processes for devices connected to I/O pads. For signal pads, these generally consist of reverse-biased diodes connected to the supply and ground rails. These diodes contribute parasitic capacitance as well as leakage currents that are of the same order as the expected signal currents. Thus, despite the on-chip electrode discussed in Section 3.3.5 technically being an I/O pad, ESD protection was not added to this pad.

Several bias voltages are needed to correctly bias the OTAs discussed in the previous sections. Each row of amplifiers shares the same bias generation circuitry in order to reduce the number of associated I/O pads.

### 3.3.7 Layout and packaging

Figure 3.6 shows a photograph of a single amplifier channel indicating the layout of the individual blocks. Each channel occupies an area of $400 \mu m \times 400 \mu m$ and includes all of the circuitry discussed in the preceding sections, excluding the biasing and ESD protection.
Figure 3.7 shows a die photograph showing all twenty-five channels. The die occupies an area of $5\text{ mm} \times 5\text{ mm}$.

Figure 3.6: Photograph of a single channel in the nanopore TIA

The chip is packaged onto a 272-pin ball-grid array (BGA) substrate. The chip has a total of 228 bondpads that connect to the fingers on the BGA substrate via wirebonds. These wirebonds are then encapsulated using dam (Hysol FP4451TD) and fill (FP4450HF) epoxy such that the electrode surfaces are left exposed. This doughnut-encapsulation technique has been used previously for other systems as well [144, 145].
3.4 Board design

3.4.1 System overview

Custom printed-circuit boards (PCBs) were designed for testing the chip functionality and for eventual use in experiments. The system primarily consists of two boards - a “daughterboard” that houses sensitive analog electronics and a “motherboard” that contains the digital data acquisition backend that interfaces with a host computer. Figure 3.8 shows the block diagram for this setup and the components contained within each of the two boards.

The 25 pairs of differential outputs from the chip are multiplexed down to five, with the row number controlling the select bits for the multiplexer. Each of these five column output pairs then has a set of boosting filters, attenuators, level-shifters, antialiasing filters, ADCs and logic buffers. The outputs of the attenuators which are required for completing the feedback loop are then demultiplexed back to the 25 corresponding pins on the chip with the row number determining the select bits.
Figure 3.8: Nanopore amplifier board-level block diagram

for the demultiplexer as well. Most of the resistors used in the design are thin-film variants to reduce their flicker noise contribution [146].

3.4.2 Power domains

The daughterboard is powered by an external 5 V DC power supply and uses voltage regulators to generate two local voltages - 1.8 V and 3.3 V. The 1.8 V supply is used for powering the chip and the bias currents, while the 3.3 V is used for powering the remaining blocks on the daughterboard. The voltage supplies are filtered using a 1 Ω resistor to further suppress power supply noise. Although there might be some concern as to noise on the 5 V supply coupling into the measurement, testing showed that noise levels were identical whether the daughterboard was powered using batteries or using a bench-top DC power supply. The reference and mid-level voltages required for proper operation of the TIA are derived from the 1.8 V supply by using a programmable resistor as a digital-to-analog converter (DAC) and are low-pass filtered before being used. The low-pass filter is crucial since noise on the reference voltage has the same effect as the $v_n$ of the TIA itself.

The motherboard is powered using the same 5 V DC power supply and also generates two local voltages that are both 3.3 V but are used for powering either analog ($AV_{DD}$, $AV_{SS}$) or digital ($DV_{DD}$, $DV_{SS}$) circuitry on the motherboard. $AV_{DD}$ powers the DC level-shifter, the anti-aliasing filters and the ADCs, and $AV_{SS}$ is tied to the
ground of the daughterboard. $DV_{DD}$ powers the logic buffers and digital isolators, whereas $DV_{SS}$ is connected to $AV_{SS}$ at a single point (star ground).

### 3.4.3 Boosting filters

The on-chip TIA has a 3-dB bandwidth set by $1/(2\pi R_F(C_F/8))$. The 20-dB/decade roll-off in the gain beyond this frequency is compensated by using active filters on the PCB. Figure 3.9 shows a schematic of the chip followed by two stages of “boosting filters” that restore flat frequency response up until 10 MHz. $R_F$ and $C_F$ are adjusted such that the low-pass filter always has its cutoff frequency at approximately 200 kHz. Thus, in the “high-gain” setting, $R_F = 45$ MΩ and $C_F = 0.15$ pF while in the “low-gain” setting, the values are 7.5 MΩ and 0.9 pF respectively.

![TIA schematic with boosting filters to increase the measurement bandwidth](image)

The boosting then starts at 200 kHz and ends at 10 MHz. Therefore, the net gain provided by the boosting filter at 10 MHz relative to that at 200 kHz is $10$ MHz/200 kHz = 50. Realizing a gain of 50 at 10 MHz using an opamp on a PCB would require a minimum UGB of 500 MHz, which is challenging to implement. Therefore, two stages
are employed; the first stage boosts from 200 kHz to 1.4 MHz while the second boosts from 1.4 MHz to 10 MHz. The net UGB requirement for the opamp in each stage is then reduced to approximately 70 MHz, which is more practical.

### 3.4.4 Data acquisition

The output of the boosting filter is connected to a level-shifter that changes the DC level of the signal from 0.9 V (half the supply voltage of the TIA chip) to 1.65 V (half the supply voltage on the PCB). This then feeds into a fourth-order Bessel filter with a cutoff frequency of 10 MHz. Bessel filters are popularly used as anti-aliasing filters because of their uniform group delay characteristics. The Bessel filters are implemented using the Sallen-Key topology as shown in Figure 3.10. This also provides a small gain to compensate for the attenuation in the signal from the boosting filters such that the net gain from the output of the chip to the output of the anti-aliasing filters is approximately unity.

![Fourth-order 10 MHz Bessel filter schematic implemented using the Sallen-Key topology](image)

**Figure 3.10:** Fourth-order 10 MHz Bessel filter schematic implemented using the Sallen-Key topology

The equation that governs the fourth-order Bessel filter frequency response (nor-
malized for phase) as a function of the cutoff frequency $f_c$ is given by [147]

$$H(jf') = H\left(\frac{j\pi f}{f_c}\right) = \frac{105}{f'^4 - 10jf'^3 - 45f'^2 + 105jf' + 105}$$

(3.3)

$$\therefore |H(jf')| = \frac{105}{\sqrt{f'^8 + 10f'^6 + 135f'^4 + 1575f'^2 + 11025}}$$

(3.4)

In this form, the gain at DC is unity and the attenuation at $f_c$ is approximately $-7$ dB. The delay-normalized version of the Bessel filter scales down $f'$ by a factor of $\pi$.

The output of the anti-aliasing filter feeds into an ADC operated with a sampling frequency of 40 million samples per second (MSPS). The ADC has 12 bits of resolution and is capable of generating both unsigned and two’s complement outputs. The 12-bit resolution implies that for a full-scale voltage of 3.3 V, the quantization error power given by $V_{LSB}^2/12$, where $V_{LSB} = V_{FS}/2^{12}$ is the value of the least-significant bit (LSB), is $5.41 \times 10^{-8}$ V$^2$. This power then yields a density of $4.81 \times 10^{-29}$ A$^2$/Hz spread across the 20 MHz measurement bandwidth when referred to the input in the worst-case, where the TIA is operated in the low-gain setting of 7.5 MΩ. This noise power is equivalent to the thermal noise of a 330 MΩ resistor and is significantly lower than the other noise sources discussed in Section 2.3.4.

The ADC outputs are low-pass filtered and buffered [148]. The buffered outputs are connected to a field-programmable gate array (FPGA, Opal Kelly XEM6310). The FPGA communicates with a host PC over USB 3.0 and transfers the ADC data to the PC where it is visualized in real-time. The 5 ADCs generate data at a rate of $5 \times 40 \times 10^6 \times 12 = 2.4$ Gbit s$^{-1} = 300$ MB s$^{-1}$. The maximum real-world transfer rate over a single USB 3.0 link is approximately 300 MB s$^{-1}$. For this reason and to increase the number of available programmable pins on the FPGA, the design incorporates a “master” FPGA that handles programmability and data transfer for two channels and a “slave” FPGA that handles data transfer for the remaining three channels. The data is buffered using the available 128 MB DRAM on the FPGAs in order to avoid data loss while waiting for the PC to request new data.
3.4.5 Shielding

The TIA’s input is extremely sensitive to EMI. A Faraday cage is thus necessary to shield the TIA’s input from external sources of EMI whenever a long wire or a volume of electrolyte is connected to it. In the system described here, the daughterboard is housed inside an aluminum box that is tied to the daughterboard’s ground and acts as a Faraday cage. The motherboard is placed outside this Faraday cage since it contains fast switching digital circuits that may generate EMI of their own. Digital signals that need to be transferred to the daughterboard first pass through RF isolators that serve to separate the possibly noisy supply voltage on the motherboard from that on the daughterboard.

Since analog signals flow from the daughterboard to the motherboard (Figure 3.8), a low impedance return path is provided between the grounds of these two boards in order to minimize any inductance on this connection that could set the two grounds at different RF potentials with respect to each other.

Finally, the experiments are performed on a vibration-isolated air table to reduce mechanical vibrations that may introduce noise through physical movement of the electrolyte connected to the TIA’s input. Figure 3.11 shows a photograph of the measurement setup.

3.5 Software

3.5.1 FPGA

The Opal Kelly XEM6310 contains a Spartan 6 series FPGA. Custom Verilog code was written in order to control the operation of the FPGA. ADC data coming into the FPGA is first stored in a 32-bit wide 4k-deep first in first out (FIFO) buffer. A custom-designed DRAM controller monitors the input buffer and transfers this data to the DRAM in bursts once sufficient data has accumulated in the buffer. The
controller and the DRAM both operate on a 312.5 MHz clock. Depending on the ADC, different data packing schemes are employed in order to maintain real-time operation. The master FPGA receives two sets of 12-bit data at 40 MHz which are padded with eight zeros to form a 32-bit entry. The input FIFO buffer used here is an asymmetric 32-in-128-out buffer to maximize the speed at which the data are written to the RAM. The slave FPGA receives three sets of 12-bit data at the same clock frequency. Here, three samples per channel are collected and then padded with 20 zeros in order to form a 128-bit entry.

Similar to the input buffers, output buffers are created in order to have some data immediately available as soon as the host PC requests it. Both FPGAs have an asymmetric 128-in-32-out output buffer. “Block-throttled” pipes are used in order to free up the USB interface for other data transfers without significantly compromising transfer speeds [149].

The master FPGA also handles programming the DACs on the daughterboard over a standard serial peripheral interface (SPI). Programming the TIAs on the chip through the scan-chains is also handled by the master FPGA.
3.5.2 PC

Custom Python code running on the PC controls the entire system. For the experiments described here, single-channel operation was sufficient and so the software was designed to prioritize display of data from a user-defined row and column. Specifically, the code converts the raw binary data coming in from the FPGA first into an appropriate voltage and then to an equivalent input-referred current.

The PyQt design toolkit is used to create the GUI with the data plotted using pyqtgraph. Pyqtgraph is significantly better than the Python standard matplotlib at displaying large datasets. 100 ms of data are nominally shown which is further downsampled by a factor of 10 such that $400 \times 10^3$ points are rendered per frame. The refresh rate and frame length are user customizable.

The data are saved as-is to disk in binary format since the overhead associated with converting to floating point numbers first and then saving those has too large a performance overhead and hinders real-time operation. Data are nominally saved in 1 s chunks although this can be changed by the user. Since no filtering is performed on the data obtained from the FPGA, information about $R_F$ and which row and column amplifier the data corresponded to is also saved to facilitate easy reconstruction when loaded from disk.

Filtering during display is handled using scipy’s implementation of a digital approximation of a fourth-order Bessel filter. This mimics the effect of recording with a system with the same bandwidth as the filter cutoff. The filter cutoff is programmable and use of the filter does not compromise real-time operation during recording. Separate tabs show the PSD and histogram of the time trace currently visible. Statistics such as the mean and standard deviation of the time trace are constantly calculated and updated. Further, noise values at various bandwidths are calculated by taking the square root of the integral of the PSD until that bandwidth mimicking the effect of a brick-wall filter.
3.6 Measurement results

Determining the frequency response of a low-noise TIA can be challenging. In this work, the frequency response is determined by coupling in a square wave of voltage through a small capacitor into the TIA input. The small capacitor is realized by connecting the ground of the square wave generator to the ground of the daughterboard and holding the signal wire of the square wave generator near the TIA’s input. Although it is hard to determine exactly how much this capacitance is, it can be varied by varying the proximity of the wire to the TIA input. The square wave of voltage induces an impulse train of currents in the TIA. By ensuring that the impulse amplitude is small enough to not saturate the TIA and by choosing an appropriate square wave frequency, the Fourier transform of this impulse train is also an impulse train. This is used to extract the shape of the frequency response as shown in Figure 3.12.

![Normalized PSD of an impulse train injected at the input of the TIA.](image)

**Figure 3.12:** Normalized PSD of an impulse train injected at the input of the TIA. The peak at each frequency represents the AC gain at that particular frequency. The envelope of the curve thus yields the AC response of the TIA.

The DC gain is measured separately in order to determine the overall tran-
simpedance gain. A known voltage is applied across a known resistance (50 MΩ in this case) connected to the input of the TIA. Thus, the injected current is known and the value of $R_F$ entered into the software is adjusted until the measured current becomes equal to the injected current as shown in Figure 3.13. These two measurements combined help determine the frequency response of the TIA.

![Figure 3.13: Measured DC response of the TIA as a function of the injected current. The TIA has a linear output across a wide range of input current amplitudes.](image)

Figure 3.14 compares the measured and simulated input-referred noise PSD of the TIA in the low-gain setting with $R_F = 7.5$ MΩ. The transistor-level simulations agree very well with the actual measurement. The roll-off at high frequencies is due to the hardware anti-aliasing filter at 10 MHz.

Figure 3.15 shows the same 10 ms long open-headstage noise measurement filtered down to 200 kHz, 1 MHz and 5 MHz using a digital approximation of a fourth-order Bessel filter. The 10 MHz trace already includes the effect of the hardware anti-aliasing filter and is not filtered further in software. The RMS noise at 100 kHz and 200 kHz is 6.1 pA$_{RMS}$ and 8.1 pA$_{RMS}$ respectively. This flicker noise is dominated
by contributions from opamps used in the boosting filter and not the TIA itself - confirmed by replacing the opamps with lower-noise, lower-bandwidth equivalents. Since this work focuses primarily on high-bandwidth recordings, this added low-frequency noise did not significantly affect high frequency performance.

![Graph](image)

Figure 3.14: Measured and simulated open-headstage input-referred noise power spectral density of the nanopore amplifier

### 3.7 Summary

This chapter discussed the design considerations and implementation specifics of a CMOS TIA array optimized for low-noise nanopore recordings. The chip was fabricated in a standard 0.18 µm CMOS process and contains a $5 \times 5$ array of amplifiers in a $5 \text{mm} \times 5 \text{mm}$ die. Each channel can be independently operated and recorded from and supports up to 10 MHz of recording bandwidth with an input-referred voltage noise of $2.6 \text{nV}/\sqrt{\text{Hz}}$. The TIA design was electrically characterized for its frequency response and noise performance.
Figure 3.15: Concatenated time trace of a 10 ms long open-headstage measurement. Each section corresponds to the same trace filtered using a digital four-pole Bessel filter to cutoff frequencies of 200 kHz, 1 MHz and 5 MHz, respectively. The 10 MHz trace already includes the effect of an analog four-pole Bessel filter and is not filtered further. $I_{RMS}$ values are indicated.
Chapter 4

High-Bandwidth Solid-State Nanopore Recordings

4.1 Introduction

For several years after the introduction of solid-state nanopores, research effort was directed primarily at achieving geometrical features comparable to biological nanopores. An interesting research direction is that of integration of nanopores with the measurement electronics. By virtue of the material properties of their membranes, solid-state nanopores are ideally suited for tight integration with well-designed electronics. The bandwidth improvements that this combination offered were first shown in 2012 [59]. Since then, although several new amplifier designs have been introduced, they have not been able to improve on the μs-resolution set by the then state of the art. Since
DNA translocates through solid-state nanopores at rates of several bases per µs, breaking this barrier is key to getting closer to the target of being able to sequence free-running DNA using a solid-state nanopore.

4.2 High-bandwidth nanopore measurement platform

As discussed in Section 2.3.4, reducing the wiring capacitance $C_W$ is one of the requirements for achieving high-bandwidth nanopore recordings. Tight integration of the nanopore with the measurement electronics can help eliminate the need for long shielded cables between them and thus help significantly decrease $C_W$. The solution adopted here is based on a platform proposed previously [59, 144]. Briefly, the on-chip electrode acts as one of the two electrodes required for the measurement, while a conventional Ag/AgCl pellet acts as the second electrode. A fluid chamber is constructed directly above the amplifier chip and acts as the “trans” side. The nanopore is mounted on this fluid chamber and the aforementioned pellet connects to the other side of the nanopore (the “cis” side). This design reduces $C_W$ to a few pF and enables high-bandwidth DNA translocation studies.

4.2.1 Amplifier packaging and fluid chamber construction

The amplifier chip is packaged as discussed in Section 3.3.7. The doughnut encapsulation technique allows access to the electrodes on the chip surface while mechanically and chemically protecting the sensitive gold wirebonds from the fluids that are to be introduced. A watertight fluid chamber is created above the chip surface by attaching a short segment from a polypropylene tube using KWIK-CAST silicone elastomer. The chip, with and without the fluid chamber, is shown in Figure 4.1.

The amplifier chip with the fluid chamber attached is placed on a compression-
Figure 4.1: Packaged nanopore amplifier chip with and without the fluid chamber mount socket that is soldered onto the daughterboard. The opening in the top piece of the compression mount was widened in order to allow the fluid chamber to pass through it (Figure 3.11).

4.2.2 Ag/AgCl microelectrode formation

Chips encapsulated in this manner are then subjected to aluminum etchant (Al etchant Type A, Transene) in order to remove the top-metal aluminum from the exposed pads connected to the amplifiers. This removes the aluminum and exposes the underlying layer, which, in the technology used for fabricating this chip, is an electrically conductive titanium nitride adhesion and diffusion barrier layer.

The next step is to deposit Ag on the titanium nitride layer. This is achieved through electrodeposition. The chip is mounted in the socket on the daughterboard and the system is powered on. Without turning on the bias for any of the amplifiers, the amplifier whose electrode needs the Ag deposition has its on-chip electrode connected to the external input on the I/O ring by programming the appropriate switches. Silver electroplating solution containing silver cyanide (Transene) is then added to the fluid chamber above the chip and a silver wire connected to a voltage
source is placed in it. The other terminal of the voltage source is connected to the external input on the I/O ring (which is connected to header pins on the daughterboard).

Voltage of appropriate polarity is then applied in order to drive electrodeposition of Ag on the on-chip electrode. Since no seed layer is present, it was found that lower current densities at the start of the process yielded more uniform coverage of the electrode surface area. Thus the voltage is adjusted so as to maintain $0.2 \mu$A for 10 minutes, followed by $0.5 \mu$A for 10 minutes and finally, $1 \mu$A for 20 minutes. This nominally resulted in an Ag layer that was a few µm thick. While the majority of the electrodeposition occurs in the final phase, skipping the initial phases was found to cause isolated regions of deposition, where an initial deposition of silver would then create a low impedance path and aid further deposition on the same spot. Electrodes that formed in this way were not used for recordings. While the process described here only refers to a single electrode being plated at a time, it is possible to plate multiple electrodes in parallel. If the amplifiers are turned on and the plating current is read out through the amplifier, information about which electrodes are being plated successfully can also be obtained. In the experiments described in this work, single-channel operation was sufficient and so parallel electroplating was not explored.

The silver microelectrodes thus formed are chlorided either by applying a drop of concentrated bleach or 50 mM FeCl$_3$ for 30 s [150]. In most of the nanopore experiments described here, the direction of the current is such that the microelectrode is at a positive polarity with respect to the pellet electrode. Thus, over the course of the experiment, the microelectrode continues to get electrically rechlorided. However, over the course of several hours of experiments, the microelectrode would lose its AgCl layer likely through dissolution in water. At this point, the chemical chloriding step would be repeated and the electrode would be used again.

Figure 4.2 shows a photograph of the on-chip electrode before and after the pro-
The Ag/AgCl interface formed in this way is not completely ideal. In particular, it has a charge-transfer resistance associated with it that is a source of noise in the measurement [109]. The electrodes were characterized for their charge-transfer resistance through electrochemical impedance spectroscopy (EIS) measurements. The values for $R_{ct}$ are nominally in the range of 100 $\Omega$ - 200 $\Omega$. At 200 $\Omega$, this contributes an additional $v_{n,ct}^2 = 4kTR_{ct} = 1.8 \text{nV}/\sqrt{\text{Hz}}$ which increases the total input-referred voltage noise to $v_n = \sqrt{v_{n,ct}^2 + v_{n,amp}^2} = 3.15 \text{nV}/\sqrt{\text{Hz}}$.

The access resistance of a 100 $\mu$m diameter electrode in 3 M KCl is approximately 75 $\Omega$ [82]. Reducing the amplifier’s noise below that added by this resistance will yield diminishing returns. The charge-transfer resistance could be reduced either by increasing the dimensions of the microelectrode or by increasing the exposed surface area for charge-transfer by patterning the electrode to create fingers.

4.2.3 Additional chip-surface passivation

Modern CMOS processes typically include a final passivation layer consisting of silicon nitride, silicon dioxide and polyimide above the top metal layer. In this technology,
the net thickness of this passivation layer is nominally $4.3 \mu m$. The fluid in the chamber above the chip can potentially form a parasitic parallel-plate capacitance with the top-level routing on the chip, with the passivation acting as the intermediate dielectric. The upper bound for this capacitance is approximately $\epsilon A/d = 180 \text{pF}$ assuming $\epsilon_r = 3.5$ and $A = 25 \text{mm}^2$. This value can be decreased by applying additional layers of passivation on the surface of the chip.

In this work, the KWIK-CAST silicone used for adhering the fluid chamber to the chip is also used for passivating the chip surface. An initial layer of silicone is first painted by hand such that the silicone touches the edges of the electrode to be used. Since single-channel operation is sufficient for the experiments described here, the other electrodes are covered with silicone. Once this initial layer of silicone has been painted, a fine pipette tip is placed above the opening in the silicone (above the electrode) and additional silicone is added around the pipette tip. The tip of the pipette prevents the silicone from flowing in and covering the exposed electrode. This process is repeated until the silicone is a few mm thick. Figure 4.3 shows one such chip with silicone passivation.

With the amplifier mounted in its socket on the daughterboard, the parasitic capacitance from the solution to the daughterboard’s ground was measured to be approximately $2 \text{pF}$.

4.2.4 Ultra-thin nanopore fabrication

As discussed in Section 2.3.2, noise is just one half of the SNR equation. High conductance nanopores capable of generating large signals are necessary in order to fully take advantage of the high-bandwidth capabilities of the measurement platform described in this work. The nanopores used in this study were fabricated by collaborators in the Drndić lab at the University of Pennsylvania. Compared to conventional silicon nitride nanopores, these were thinned down to near the theoretical limit ($\approx 1 \text{nm}$)
The thinning process is briefly described below.

A scanning transmission electron microscope (STEM) based ablation technique is used to achieve nanometer scale membrane thickness. STEM thinning uses electron irradiation with rastering of the electron probe of a JEOL 2010F S/TEM over a defined area of silicon nitride. This causes sputtering of silicon and nitrogen atoms with the final membrane consisting of amorphous silicon due to the higher rate of sputtering of nitrogen. A two-step process is used with an initial thinning of a 65 nm × 65 nm region of 50 nm thick freestanding silicon nitride membrane to 10 nm amorphous silicon by using a 2.5 nm probe size with membrane thickness controlled by quantifying the mass loss using electron-energy loss spectroscopy (EELS) (Figure 4.4). A second thinning in a smaller 25 nm × 25 nm region is made using a 0.5 nm spot size, bringing the membrane thickness down from 10 nm to less than 4 nm. The same spot size is then used for actual drilling of the nanopore.
Figure 4.4: (a) Bright-field TEM image of nanopores made in STEM-thinned membranes. Circles indicating diameters of 1.7 nm, 2.0 nm, and 2.6 nm are shown in overlay with corresponding nanopores. (b) High Annular Dark Field image of STEM-thinned regions of silicon nitride. Regions are labeled according to the steps in the STEM-thinning process, with Region 1 unthinned, Region 2 thinned to 10 nm, and Region 3 thinned to 3.5 nm. (c) Electron-energy loss spectra (EELS) taken of silicon nitride before (Region 1) and after (Region 2) initial thinning. Red markers indicate the Si peak, which drops to 20% of its original value; this corresponds to a thinning from 50 nm to 10 nm. (d) Electron Energy Loss Spectra (EELS) taken of silicon nitride before (Region 2) and after (Region 3) the second thinning process. Red markers indicate the Si peak, which drops to 35%, corresponding to a thinning from 10 nm to 3.5 nm. All images were taken using a JEOL 2010F STEM instrument.

4.2.5 Nanopore assembly

Before use in experiments, the nanopore chip is cleaned in piranha for between 5 and 10 minutes using an approach described previously [108]. This treatment helps reduce the flicker noise of the pore and increases its chances of wetting. The pore is then mounted on a custom polydimethylsiloxane (PDMS) cell using the same silicone elastomer used for attaching the fluid chamber to the chip. This provides a watertight seal as well as serves as an additional passivation layer to decrease $C_{pore}$ (Figure 4.5). Since this silicone is painted by hand, there is some variability in terms
of the membrane area that is not covered by the silicone. $C_{pore}$ is typically about 10 pF which is an order of magnitude lower than the 300 pF that was observed in earlier studies [103]. Reduction in $C_{pore}$ can yield further improvements in SNR and bandwidth.

Figure 4.5: Schematic of the cross-section of the nanopore chip including the silicone passivation

### 4.3 Open-pore noise measurements

As part of the experimental setup, the nanopore is mounted on a PDMS cell and placed on top of the fluid chamber above the nanopore amplifier chip as described in the previous section. Of the several nanopores with different diameters tested in this manner, Figure 4.6 shows the input-referred noise PSDs of the amplifier chip with two different nanopores - Pore 1 ($C_{pore} = 10\,\text{pF}, d = 1.3\,\text{nm}, t_{eff} = 1.4\,\text{nm}$) and Pore 2 ($C_{pore} = 13\,\text{pF}, d = 1.7\,\text{nm}, t_{eff} = 1.2\,\text{nm}$). As a representative example, Figure 4.7 shows the concatenated time trace of a 10 ms long noise recording with Pore 1. The input-referred noise for Pore 1 at 200 kHz is $23.2\,\text{pA}_{\text{RMS}}$, at 1 MHz is $125.7\,\text{pA}_{\text{RMS}}$, at 5 MHz is $1.43\,\text{nA}_{\text{RMS}}$ and at the full bandwidth of 10 MHz is $4.19\,\text{nA}_{\text{RMS}}$.

It is worth noting that the flicker noise discussed in Section 3.6 does not significantly affect high-frequency noise levels (Table 4.1). Since the focus in this work was
Figure 4.6: Input-referred noise power spectral density of the nanopore amplifier chip with two different nanopores. Pore 1 has $C_{pore} \approx 10\,\text{pF}$ while Pore 2 has $C_{pore} \approx 13\,\text{pF}$.

In maximizing bandwidth for a given SNR, high-bandwidth opamps were preferred in the boosting filter over their lower-noise lower-bandwidth counterparts.

<table>
<thead>
<tr>
<th></th>
<th>Noise up to 100 kHz</th>
<th>Noise up to 1 MHz</th>
<th>Noise up to 5 MHz</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Absolute (pA$_{RMS}$)</td>
<td>Contribution of noise at 100 kHz bandwidth to total noise</td>
<td>Absolute (pA$_{RMS}$)</td>
</tr>
<tr>
<td><strong>Open headstage</strong></td>
<td>6.1</td>
<td>100%</td>
<td>47.8</td>
</tr>
<tr>
<td><strong>Pore 1</strong></td>
<td>14.9</td>
<td>100%</td>
<td>125.7</td>
</tr>
</tbody>
</table>

Table 4.1: Contributions of low frequency noise to integrated noise at different bandwidths
4.4 Single-stranded DNA translocation

The experiments were performed with short ssDNA samples that were either 40 or 100 nucleotides (nt) long and prepared in aliquots of 200 nM in 3 M KCl. Since $\Delta I$ plays a significant role in improving the measurement bandwidth, 3 M KCl is used instead of the commonly used 1 M KCl. Increasing molarity increases $\Delta I$ but does not affect the relative blockade $\Delta I/I_{baseline}$. In order to accommodate larger baseline currents (and thereby maximize $\Delta I$), the nanopore amplifier is operated in its low-gain setting.

Once the nanopore has been determined to be open and have a stable baseline current, the DNA is added to the cis side and a negative bias voltage is applied to electrophoretically drive the DNA through the nanopore. The bias voltage is varied from 300 mV to 900 mV. Given the small diameters of the pores tested here, voltage biases lower than 300 mV frequently caused blockage of the pore.

Figure 4.8 shows a concatenated time trace of 100 nt ssDNA translocation recordings through Pore 3 ($C_{pore} = 15 \text{ pF}$, $d = 1.9 \text{ nm}$, $t_{eff} = 3.3 \text{ nm}$) at different applied
bias voltages. Figure 4.9 shows the same 0.2 s long time trace of ssDNA translocations through Pore 1 at 900 mV bias filtered using different four-pole Bessel filters. The translocation events are clearly visible even at filtering bandwidths as high as 5 MHz whereas at the conventional measurement bandwidth of 10 kHz, events are either severely attenuated or lost altogether.

![Figure 4.8: Concatenated time trace of 100 nt ssDNA translocation recordings through Pore 3 at biases of 0, 300, 600 and 900 mV. Each trace is 0.2 s long and filtered using a four-pole Bessel filter with a cutoff frequency of 1 MHz.](image)

4.5 Event calling

Several efforts have been focused towards creating a suite of tools for analyzing nanopore data [151–154]. However, the algorithms used in these works are computationally more expensive than the simple thresholding approach used here. Specifically, for high-bandwidth data, thresholding is extremely fast since it can be performed in $O(n)$.

Custom code written in Python is used to analyze the recorded nanopore traces. First, the baseline $I_{baseline}$ and the standard deviation $I_{RMS}$ are calculated for the
Figure 4.9: Concatenated time trace of a 0.2 s long recording of 100 nt ssDNA translocation through Pore 1 at 900 mV bias. The traces are filtered using a four-pole Bessel filter to 10 kHz, 200 kHz, 1 MHz and 5 MHz bandwidths. Low cutoff frequencies show severe degradation of signal amplitudes.

data in a known event-free region. Next, a threshold is set $\text{SNR}_{\text{min}} \times I_{RMS}$ below $I_{\text{baseline}}$. Whenever a data point in the derivative of the original time trace exceeds the threshold, it signals an edge of an event. Depending on the sign of the derivative, it indicates either the start or the end of the event. From this event point, a local search finds the nearest data point that is $3I_{RMS}$ below the baseline and this point is treated as the actual start or end of the event, respectively.

In this work, $\text{SNR}_{\text{min}}$ is set to either six or eight. This results in a negligibly low false event-rate as discussed in Section 2.3.2. Further, minimum and maximum event widths are set in order to reject events that correspond to either bumping events or to pore clogging.

4.6 Short-lived intra-event features

The need for a high-bandwidth platform is best demonstrated by its ability to detect and resolve events and features that were undetectable previously. Figure 4.10 shows
examples of 100 nt ssDNA translocation events through Pore 1 filtered to 1 MHz and 5 MHz bandwidths. The four-pole Bessel filter used for filtering has a rise time of $\approx 0.5 \mu s$ and $0.1 \mu s$ at 1 MHz and 5 MHz cutoff frequencies respectively. Consequently, events with durations less than twice these times will have their amplitudes severely attenuated [5]. Figure 4.10 shows examples of features that are visible only at bandwidths made possible by this work, some with durations of less than 200 ns; all labeled features exceed the $3I_{RMS}$ noise level below the baseline. Some of these features reveal extremely brief interactions of the molecule with the pore during its entry and exit. Some of the features, however, correspond to a deeper event within a relatively shallow level. These features are likely to be the real translocation events with the shallow level corresponding to some intermediate state where the molecule is not completely in the pore. In these cases, events are likely to be fast even when measured at a bandwidth of 1 MHz, leading to distortion of event depth and duration.

Figure 4.10: Example events from 100 nt ssDNA translocation through Pore 1 at 900 mV bias showing features visible at 5 MHz bandwidth that are invisible at 1 MHz bandwidth. Feature durations are indicated.
4.7 Analysis of two-state events

Several previous studies have reported two-level translocation current waveforms as are observed in Figure 4.11 [50, 59, 155]. Because the diameters of the nanopores used in these experiments are smaller than in any of these previous studies, the access resistance $R_a$ now starts to play a more important role in determining the overall ionic current (Section 2.3.3). In particular, modulation of $R_a$ by a molecule in the vicinity of the nanopore can be significant. The two-level behavior observed is attributed to a molecule that gets trapped as it enters or exits the pore. Such an explanation in our case is further bolstered by the fact that the standard deviation of the shallow levels observed ($I_{RMS} = 834\, \text{pA}_{RMS}$ to $1.81\, \text{nA}_{RMS}$ for the data of Figure 4.11) is significantly higher than that of the baseline current itself ($I_{RMS} = 281\, \text{pA}_{RMS}$). If a molecule in the vicinity of the pore is indeed the cause of the shallow level, then Brownian motion of the molecule could explain the increased standard deviation in the current.

However, the deep level is not present only at the end of a translocation event. Figure 4.12 shows examples of several events where the deep level is either at the beginning or the middle of an event. The explanation that the shallow level is due to the molecule blocking the pore entrance before translocating through [59] thus does not hold. Our data are in agreement with those reported in other studies [50].

4.8 DNA translocation statistics

Figure 4.13a shows a scatter plot of average current blockade values as a function of dwell time for a 4s trace recorded for 100 nt ssDNA at 900 mV bias. Event detection is determined by setting a threshold that is $6I_{RMS}$ away from the baseline. The large spread in the average $\Delta I$ values is due to the variance in the dwell time in the access region during a translocation event. A long dwell in the access region corresponds
Figure 4.11: Concatenated events from 100 nt ssDNA translocation through Pore 3 at 900 mV bias padded with baseline points for reference. The dashed black lines indicate the shallow level corresponding to the molecule being in the access region and the deep level corresponding to the actual translocation. The standard deviation of the current in the shallow region is significantly higher than even that of the baseline.

To a relatively shallow average $\Delta I$. Larger average $\Delta I$ values are associated with short dwell times in the access region. As the filtering cutoff frequency is reduced, $I_{RMS}$ decreases, but so does the amplitude of short events which means that some events fail to get detected at lower bandwidths. Conversely, shallow and long events are more likely to be seen at lower bandwidths. Figure 4.13b shows fits of the dwell times to $A_1 e^{-t/\tau_1} + A_2 e^{-t/\tau_2}$ where $\tau_1 < \tau_2$ [61]. More aggressive filtering increases $\tau_1$ indicating that the increased bandwidth results presented here are more accurate in capturing the average translocation rate.
Figure 4.12: Concatenated events from 100 nt ssDNA translocation through Pore 3 at 900 mV bias padded with baseline points for reference. The upper three events exhibit a deep level followed by a shallow level while the lower three events exhibit a deep level flanked by shallow level dwells before a return to the baseline.

4.9 Glass-passivated nanopores

In the work presented thus far, $C_{pore}$ is the largest contributor to $\Sigma C_i$. $C_{pore}$ can be reduced even further through additional and reproducible passivation of the membrane dielectric. Several approaches have been proposed for reducing $C_{pore}$ [37, 38, 108, 156–160]. In particular, the glass-based passivation approach was used here in order to reduce $C_{pore}$. The processing flow developed by the collaborators at the University of Pennsylvania is described in detail elsewhere, and is briefly reproduced here [157]. 300 µm thick glass chips are pre-deposited with silicon nitride layers, patterned by photolithography, and etched by hydrofluoric acid to created suspended silicon nitride membranes for drilling nanopores. Compared to the conventional silicon substrate
Figure 4.13: (a) Mean current blockage vs dwell time scatter plot for \( n = 2008 \) 100 nt ssDNA translocation events through Pore 1 at 5 MHz, 1 MHz, and 200 kHz filtering bandwidths (\( n \) indicated for the 5 MHz filtering bandwidth). Filtering to lower frequencies clearly indicates increased attenuation especially for events close to the inverse of the filter’s cutoff frequency. (b) Histogram plot of counts vs dwell time for the data presented in (a). Each of the plots are fitted to \( A_1e^{-t/\tau_1} + A_2e^{-t/\tau_2} \) and \( \tau_1 \) is indicated for each of the fits where \( \tau_1 < \tau_2 \). Increasing the filtering bandwidth indicates a reduction in the characteristic dwell time for the event, suggesting that even at bandwidths as high as 1 MHz translocation events were distorted.

with silicone passivation as illustrated in Figure 4.5, fused-silica (glass) based substrates are fabricated with cross-sections as depicted in Figure 4.14. An additional layer of silicone is added on the top for extra passivation. The resulting capacitance of the fused-silica device is consistently of the order of a few pF and the results
are highly repeatable. This passivation process is compatible with the membrane thinning steps described in Section 4.2.4.

![Diagram of a nanopore chip](image)

Figure 4.14: Schematic of the cross-section of the glass-passivated nanopore chip

Of the several nanopores tested in this manner, the input-referred noise PSD for Pore 4 \( (C_{\text{pore}} \approx 3 \, \text{pF}, d = 4 \, \text{nm}, t_{\text{eff}} = 5 \, \text{nm}) \), as measured with the amplifier in high-gain mode, is compared against that of Pore 1 as well as the amplifier’s open-headstage noise floor in Figure 4.15. Figure 4.16 shows an open-pore noise measurement with a glass-passivated nanopore similar to that shown in Figure 4.7. The glass passivation reduces noise levels significantly compared to the regular silicone passivation.

Figure 4.17 shows 200 nt ssDNA translocation through Pore 5 \( (C_{\text{pore}} \approx 3 \, \text{pF}, d = 1.2 \, \text{nm}, t_{\text{eff}} = 3 \, \text{nm}) \) filtered to several different bandwidths. Compared to the measurements through Pore 1 shown in Figure 4.9, there is significant improvement in SNR due to the reduced membrane capacitance.

### 4.9.1 Inverting the amplifier’s low-pass response

Assuming that the current-divider loop shown in Figure 3.1 is fast enough and the integrator opamp has an open-loop transfer function given by \( A(s) \), it can be shown that the open-loop gain of the system in open-headstage configuration \( (R_{\text{pore}} = \infty, \)}
Figure 4.15: Input-referred noise power spectral density of the nanopore amplifier chip with two different nanopores. Pore 1 has $C_{\text{pore}} \approx 10 \text{ pF}$ while the glass-passivated Pore 4 has $C_{\text{pore}} \approx 3 \text{ pF}$.

When $C_{\text{pore}} = 0$, the closed-loop transimpedance gain is given by

$$F(s) = \frac{A(s)}{1 + A(s) sC_F}$$

Note that this expression also neglects the effect of the amplifier’s input capacitance. With this $L(s)$, the closed-loop transimpedance gain is then given by

$$F(s) = \frac{A(s) \frac{8}{sC_F}}{1 + A(s) sC_F} \frac{8}{1 + A(s) sC_F R_F}$$

$$= \frac{A(s) \frac{8}{sC_F}}{1 + A(s) \left(1 + \frac{8}{sC_F R_F}\right)}$$

During actual experiments, $C_{\text{pore}} \neq 0$ implying that $\sum C_i > C_F$ and its effect on the closed-loop transfer function needs to be accounted for. In general, the open-loop
Figure 4.16: Concatenated time trace of a 10 ms long noise measurement with Pore 4 ($C_{\text{pore}} \approx 3 \text{ pF}$). Each section corresponds to the same trace filtered using a digital four-pole Bessel filter to cutoff frequencies of 200 kHz, 1 MHz, and 5 MHz, respectively. The 10 MHz trace already includes the effect of an analog four-pole Bessel filter and is not filtered further. $I_{\text{RMS}}$ values are indicated.

Gain of the system is modified to

$$L(s) = \frac{A(s) \frac{C_F}{\sum C_i} \frac{1}{sC_F}}{1 + A(s) \frac{C_F}{\sum C_i} \frac{1}{R_F sC_F}}$$  \hspace{1cm} (4.3)
Figure 4.17: Concatenated time trace of a 0.2 s long recording of 200 nt ssDNA translocation through Pore 5 at 900 mV bias. The traces are filtered using a four-pole Bessel filter to 200 kHz, 1 MHz, 5 MHz and 10 MHz bandwidths. Compared to silicone-passivated nanopores, there is significant improvement in SNR.

and the closed-loop transfer function changes accordingly to

\[
F(s) = \frac{A(s) \frac{C_F}{\sum C_i} \frac{8}{1 + A(s) \frac{C_F}{\sum C_i} \frac{sC_F}{1 + A(s) \frac{C_F}{\sum C_i} \frac{8}{sC_FR_F}}}}{1 + A(s) \frac{C_F}{\sum C_i} \frac{sC_F}{\sum C_i} \left(1 + \frac{8}{sC_FR_F}\right)}
\]

(4.4)

If \(\sum C_i > C_F\) and \(A(s) = \omega_{UGB}/s\) with \(\omega_{UGB}\) being the UGB of the opamp, an additional pole is introduced at \(\omega_{UGB}C_F/\sum C_i\). In the experiments with glass-passivated pores, \(C_F = 0.15\) pF and \(\sum C_i \approx 6\) pF placing an additional pole at approximately 2.5 MHz. The effect of this additional low-pass filter needs to be inverted.
Consider a signal $y(t)$ that is the result of a signal $x(t)$ passed through a first-order low pass filter with 3-dB frequency $\omega_0$. Their Fourier transforms are related as

$$Y(j\omega) = \frac{X(j\omega)}{1 + \frac{j\omega}{\omega_0}} \quad (4.5)$$

$$\therefore X(j\omega) = Y(j\omega) \left(1 + \frac{j\omega}{\omega_0}\right) \quad (4.6)$$

Thus, the signals are related in the time domain as

$$x(t) = y(t) + \frac{1}{\omega_0} \frac{dy(t)}{dt} \quad (4.7)$$

And, equivalently, in the discrete time domain as (assuming sufficient oversampling)

$$x[n] = y[n] + \frac{1}{\omega_0} (y[n] - y[n - 1]) \quad (4.8)$$

where $\omega_0 \in [0, \pi]$ now is the angular frequency of the low-pass filter in the discrete domain. In the system described here with a sampling frequency of 40 MHz and with a nominal $f_0 = 2.5$ MHz, $\omega_0 = \pi/8$. The glass-passivated pore recordings are subjected to this inverse filtering in order to restore their full 10 MHz bandwidth.

### 4.10 Comparison to state of the art

Table 4.2 compares the performance of the system described in this chapter to other high-bandwidth translocation studies employing similar platforms. The work described in this chapter currently represents the fastest reported recording of DNA translocations through nanopores.

### 4.11 Summary

This chapter discussed the design and construction of a low-noise high-bandwidth nanopore recording platform. The platform was experimentally validated by translocating 40 - 200 nt long ssDNA molecules through ultra-thin nanopores with silicone
<table>
<thead>
<tr>
<th>Metric</th>
<th>Rosenstein et al [59]</th>
<th>Balan et al [157]</th>
<th>Fraccari et al [161]</th>
<th>This work</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pore type</td>
<td>Silicon nitride</td>
<td>Silicon nitride</td>
<td>Quartz nanopipette</td>
<td>a-Silicon</td>
</tr>
<tr>
<td>Pore conductance (nS)</td>
<td>10</td>
<td>4.5</td>
<td>27–43</td>
<td>20–30</td>
</tr>
<tr>
<td>Pore diameter (nm)</td>
<td>3.5</td>
<td>5</td>
<td>20</td>
<td>&lt; 2</td>
</tr>
<tr>
<td>Bias (mV)</td>
<td>600</td>
<td>1000</td>
<td>800</td>
<td>900</td>
</tr>
<tr>
<td>$\Delta I$ (nA)</td>
<td>1.2</td>
<td>1</td>
<td>0.22</td>
<td>15</td>
</tr>
<tr>
<td>$C_{pore}$ (pF)</td>
<td>6</td>
<td>&lt; 1</td>
<td>1 - 10</td>
<td>3</td>
</tr>
<tr>
<td>$C_{headstage}$ (pF)</td>
<td>2</td>
<td>20</td>
<td>3.5</td>
<td>3.5</td>
</tr>
<tr>
<td>$v_n$ (nV/√Hz)</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>3.15</td>
</tr>
<tr>
<td>Max. bandwidth (MHz)</td>
<td>1</td>
<td>1</td>
<td>0.2</td>
<td>10</td>
</tr>
<tr>
<td>Input-referred noise at target bandwidth (pA RMS)</td>
<td>155</td>
<td>110</td>
<td>29.3</td>
<td>2230</td>
</tr>
</tbody>
</table>

Table 4.2: Comparison of this work with other high-bandwidth DNA translocation studies through solid-state nanopores

and glass passivation to enable recording bandwidths as high as 10 MHz. A comparison to the current state of the art reveals that this platform currently enables the highest reported recording bandwidth for DNA translocation through nanopores.
Chapter 5

Wavelet Denoising of Nanopore and Ion Channel Recordings

5.1 Introduction

Solid-state nanopores can be considered to be the biomimetic analogs of leaky ion channels. While not generally gateable, recordings from translocation experiments in solid-state nanopores show striking similarities in terms of the signal shape to those from single ion channels. The instrumentation amplifier used for both these types of recordings is thus, largely identical. Advancements in the design and integration of these amplifiers have led to significant improvements in the measurement bandwidth
of these recordings [59, 60, 162, 163].

This chapter explores the use of wavelet denoising for further improving the SNR in high-bandwidth nanopore and ion channel measurements. Having optimized the hardware as discussed in Chapter 4, optimizing the signal processing forms the next logical step in developing a low-noise recording platform.

5.2 Traditional approaches to denoising

A standard technique for improving the SNR after measurement is to low-pass filter the data. The term low-pass here refers to the frequency domain in that low-frequency signals are allowed to pass through with their magnitudes unaltered while the magnitudes of high-frequency signals are suppressed. Most classical filters are designed in the frequency domain. Of the variety of filters that can be thus designed, Bessel and Gaussian filters have found popular use in nanopore and ion channel recordings because of their uniform group delay characteristics [5].

Besides Bessel filters, a number of approaches based on Hidden Markov Models (HMMs) have been applied to denoise nanopore [164–166] as well as ion channel [167–170] recordings. Markov models are used to analyze systems that consist of states that transition to and from each other and can be directly observed. HMMs add an extra layer of depth in that the states are not observable directly. While HMMs are an extremely powerful technique, the assumption of an underlying model is, by definition, necessary. Further, HMM-based denoising techniques can be computationally expensive. Most such implementations also assume that the additive noise is spectrally white implying that the correlation in the noise between measurement points is 0. This does not hold in the context of high-bandwidth nanopore and ion channel recordings where the corrupting noise has an $f^2$ input-referred PSD. Workarounds for this have been proposed that involve the use of “meta-states” but these further increase the computational complexity [167].
In comparison, frequency-domain filters make no assumptions about the underlying model itself and are relatively computationally cheap. However, they are not particularly suited for denoising temporally localized signals such as a pulse. Figure 5.1 shows a simulated pulse and its corresponding Fourier transform. The simulation replicates the sampling of a continuous-time pulse with a width of 1 µs and an amplitude of $-10 \text{nA}$ at a rate of 40 MSPS. The plot reveals that information is spread across a range of frequencies with many non-zero coefficients. This is because the Fourier transform uses sinusoids that exist for all time as its basis functions and a large number of such sinusoids are necessary for accurately representing a time-limited signal like a pulse. Low-pass filters serve to reduce the amplitude of the high frequency components and consequently smooth the signal.

![Figure 5.1: Simulated (a) 1 µs pulse at 40 MSPS and (b) its corresponding Fourier transform](image)

5.3 Wavelet transform

While the Fourier transform uses sinusoids of different frequencies as its basis functions, the wavelet transform uses scaled and shifted versions of a “mother wavelet”
$\psi(t)$. The scaled and shifted version is related to its mother wavelet as

$$
\psi_{a,b}(t) = \frac{1}{\sqrt{a}} \psi \left( \frac{t - b}{a} \right)
$$

(5.1)

where $a$ is the scaling factor and $b$ is the translation factor. The inner product of this window with the signal of interest then gives the wavelet transform as

$$
X_W(a, b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} \psi^* \left( \frac{t - b}{a} \right) x(t) dt
$$

(5.2)

where the * denotes complex conjugation. This approach is similar to the windowing approach used in the short term Fourier transform (STFT). The key difference is that the STFT utilizes a window of a fixed width and hence provides fixed temporal and frequency resolution across frequencies. In comparison, the wavelet transform uses windows of varying widths which provide fine temporal resolution for high frequency components and fine frequency resolution for low frequency components.

While the continuous wavelet transform is a useful tool, it is the discrete wavelet transform (DWT) that is of particular relevance in practical scenarios. DWTs have found popular use in image compression algorithms [171, 172] since it can be efficiently computed using a pyramidal algorithm [171]. The DWT computes the wavelet transform coefficients at a particular decomposition level $j$. At every $j$, the wavelet function $\psi_j(x)$ has a counterpart $\phi_j(x)$, termed the “scaling” function, such that the two are orthogonal to each other.

For computational efficiency equivalent to that of the fast Fourier transform (FFT), the DWT employs dyadic scaling which implies that $a = 2^j$. Further, $b = k2^j$ such that translations are only by integer amounts. A consequence of this choice of $b$ is that low frequency components (large $j$) are shifted by larger amounts than high frequency components (small $j$). This trade-off is acceptable since low frequency components, by the sampling theorem, do not need to be sampled as finely as high frequency components. Further, the scaling and wavelet functions are related by the
equations

\[ \phi_j(x) = \sum_n g_n \phi_{j-1}(x - n) \]  

\[ \psi_j(x) = \sum_n h_n \phi_{j-1}(x - n) \]  

where \( g_n \) and \( h_n \) are discrete-time filters that relate how functions at a lower level of decomposition relate to those at a higher level. Thus, since \( \phi_j(x) \) can be analyzed exactly by using \( \phi_{j-1}(x) \), it is evident that any function that can be analyzed using \( \phi_j(x) \) can also be analyzed using \( \phi_{j-1}(x) \). The DWT coefficients can now be computed using \( g_n \) and \( h_n \) as

\[ d_j[n] = \sum_k h_k \ast a_{j-1}[2n - k] \]  

\[ a_j[n] = \sum_k g_k \ast a_{j-1}[2n - k] \]  

where \( d_j[n] \) and \( a_j[n] \) are termed the “detail” and “approximation” coefficients at level \( j \) and represent the inner product of the signal \( x[n] \) with the wavelet and scaling functions respectively. Note that \( a_0[n] = x[n] \) and the subsequent detail and approximation coefficients can be computed once \( g_n \) and \( h_n \) have been determined. With DWT, the approximation and detail coefficients at every level are downsampled by a factor of two to remove redundancy of information. This scheme is depicted in Figure 5.2 for decomposition up to \( J \) levels.

Figure 5.2: Block diagram of the wavelet transform. In the DWT, the approximation and detail coefficients are downsampled by a factor of two at each stage (not shown in figure).

In general, the detail coefficients represent the high frequency components at a particular level whereas the approximation coefficients represent the low frequency
components. The DWT results in a non-redundant representation of the signal such that a signal of length \( N \) yields \( N \) total wavelet coefficients. The stationary wavelet transform (SWT) is a variant of the DWT that does not downsample the detail and approximation coefficients at each level. Thus, a signal of length \( N \) yields a total of \( N(J + 1) \) wavelet coefficients for \( J \) levels of decomposition (\( N \) detail coefficients for each of the \( J \) levels and \( N \) approximation coefficients at the \( J^{th} \) level).

Figure 5.3 shows the approximation and detail coefficients at the first level of decomposition for the pulse shown in Figure 5.1a using the biorthogonal 1.5 wavelet as the wavelet basis. Compared to the Fourier Transform shown in Figure 5.1b, a lot fewer coefficients in the SWT are non-zero. In principle, this should allow for denoising schemes that reduce noise without smoothing the signal as aggressively.

![ SWT of a simulated 1 μs pulse](image)

**Figure 5.3: SWT of a simulated 1 μs pulse**

### 5.3.1 Denoising using the wavelet transform

The use of wavelet transforms for denoising has been explored extensively over the last few decades [173]. While several variations are possible, the basic framework for all the denoising schemes is more or less identical. The following steps are performed in a typical wavelet based denoising scheme:

1. Choose a wavelet basis
2. Choose the maximum decomposition level $J$

3. Compute the DWT/SWT for the signal $x[n]$ to obtain detail coefficients $d_j[k]$ and approximation coefficients $a_j[k]$ where $j = 1, \cdots, J$ and the range of $k$ depends on the choice of DWT or SWT.

4. Calculate the noise threshold $\lambda_j$ for $d_j[k]$.

5. Choose a thresholding scheme $T$ and threshold $d_j[k]$ to obtain $d'_j[k]$.

6. Perform the inverse DWT/SWT using $a_j[k]$ and $d'_j[k]$.

While the wavelet denoising procedure described here is for one-dimensional signals, it can be easily extended to higher dimensional signals. The following parameters must be considered carefully for achieving good denoising:

**Wavelet transform** The SWT yields a highly redundant representation of the input signal and is computationally more expensive than the DWT. However, the DWT is not translation-invariant which can cause artifacts during reconstruction after denoising [174].

**Wavelet basis** The wavelet basis should be chosen so as to concentrate the signal information into a few high-valued coefficients while the noise is spread uniformly across coefficients. This usually occurs when the scaling function resembles the signal of interest.

**Levels of decomposition** Higher levels of decomposition afford greater denoising but also smooth the signal to a greater extent and vice versa.

**Threshold** The threshold seeks to demarcate coefficients that are due to noise from those that are due to the signal. Large threshold values can suppress noise to a greater extent but also risk suppressing actual signal information.
Thresholding scheme Once a threshold has been chosen, the scheme determines how coefficients with values larger than the threshold are handled. An aggressive scheme would leave these coefficients unaltered and risk generating reconstruction artifacts while a conservative scheme would reduce the magnitude of these coefficients and risk reducing signal amplitudes.

Wavelet transform and denoising of time-series data has been explored in a wide variety of contexts [175–179]. Specifically, wavelet denoising has been used to improve SNR of data from ensemble ion channels [180], Coulter counter [181] and nanopore data [182]. However, the intrinsic SNR available at high bandwidths in these studies was relatively low due to the absence of electronics optimized for high-bandwidth measurements.

5.4 Wavelet denoising of simulated pulses

The pulses here are simulated assuming that they were sampled at a rate of 40 MSPS in order to mimic the data acquisition system described in Section 3.4.4. The pulses are nominally generated with a width of 1 µs unless described otherwise. They are padded with 5000 points (125 µs) on either side. Finally, the amplitudes are ascribed units of nA (magnitude nominally 10 nA) although wavelet denoising itself is amplitude-agnostic and should work with other signals, such as voltages, too [183].

Of the various possible noise sources, the two most likely to dominate high frequency noise performance - white and $f^2$ - are added to the pulses in order to model the actual measurement setup. White noise samples with a normal amplitude distribution are easily generated in Python. To simulate $f^2$ noise, white noise of standard deviation $\sigma/\sqrt{2}$ is first generated. The derivative of this then creates noise with an $f^2$ PSD and a standard deviation of $\sigma$ as desired.

For the reasons described in Section 5.3.1, the SWT is preferred over the DWT.
Previous work has suggested ways to automate determining the optimal number of levels of decomposition [184]. However, this scheme does not work well for the $f^2$ noise that is of interest here and is not used. Instead, the levels of decomposition are adjusted on a case-by-case basis in order to offer good denoising while retaining sharp transitions.

Wavelet processing is performed in Python using the PyWavelets module [185].

### 5.4.1 Choice of wavelet basis

The choice of wavelet basis can greatly impact the performance of wavelet denoising. In this work, the optimal basis is determined in the sense of minimizing the mean squared error (MSE) when compared to the ground truth of a noiseless pulse. A 1 µs long pulse sampled at 40 MSPS is corrupted with $f^2$ noise with $\sigma = 8 \text{nA}_{\text{RMS}}$. Note that the noise is spread across the entire available bandwidth of 20 MHz. With a pulse amplitude of 10 nA, this results in an initial SNR of 2 dB.

For each wavelet basis, five levels of SWT decomposition and garrote thresholding are used. The threshold is determined using the level-dependent threshold (LDT). The MSE is computed on a 5 µs long window around the actual pulse. All the wavelets available within the PyWavelets module are included in this test. This includes wavelets belonging to the following families:

- Haar (haar)
- Daubechies (db) [186]
- Symlets (sym)
- Coiflets (coif)
- Biorthogonal (bior) [187, 188]
- Reverse biorthogonal (rbio)
Discrete approximation of the Meyer wavelet (dmey) totaling 106 different wavelet bases.

Figure 5.4 plots the ten bases with the lowest MSE. The three wavelet bases with the lowest MSEs are consistently rbio3.1, rbio3.3 and bior1.5. Visual inspection of the rbio-filtered time traces showed that while these denoised pulses have extremely fast rising and falling edges, they suffer from significant ringing near transitions. As a result, the bior1.5 wavelet is chosen as the basis for the denoising discussed here.

![Figure 5.4: Effect of wavelet basis on the MSE of a wavelet denoised pulse](image)

The numbers in the bior and rbio naming schemes refer to the number of vanishing moments in the analysis and synthesis filters. Biorthogonal (and rbio) wavelets are different from the classical wavelets because they use separate scaling and wavelet functions for analysis and synthesis. Unlike traditional wavelet families, \( \phi_j(x) \) and \( \psi_j(x) \) at a given \( j \) are not orthogonal complements of each other in the space spanned by \( \phi_{j-1} \). However, the separation of the functions used for analysis and synthesis allows for the corresponding filters to be symmetric and linear phase.
For biorthogonal wavelets, the analysis scaling and wavelet functions are denoted by $\phi(x)$ and $\psi(x)$ respectively while the synthesis scaling and wavelet functions are denoted by $\tilde{\phi}(x)$ and $\tilde{\psi}(x)$ respectively. Figure 5.5 plots these functions for bior1.5.

![Analysis scaling function](image)

![Analysis wavelet function](image)

![Synthesis scaling function](image)

![Synthesis wavelet function](image)

Figure 5.5: Analysis and synthesis scaling and wavelet functions for bior1.5

As is the case with the traditional wavelets, corresponding filter banks $g_n, h_n, \tilde{g}_n$ and $\tilde{h}_n$ can be defined for the biorthogonal analysis and synthesis scaling and wavelet functions. These are related to the various functions by the following equations

$$\phi_j(x) = \sum_n g_n \phi_{j-1}(x - n)$$

$$\psi_j(x) = \sum_n h_n \phi_{j-1}(x - n)$$

$$\tilde{\phi}_j(x) = \sum_n \tilde{g}_n \tilde{\phi}_{j-1}(x - n)$$

$$\tilde{\psi}_j(x) = \sum_n \tilde{h}_n \tilde{\phi}_{j-1}(x - n)$$

and are shown in Figure 5.6.
5.4.2 Choice of wavelet threshold and thresholding scheme

Choosing the appropriate threshold and thresholding scheme is imperative when denoising using wavelets. The threshold value needs to be chosen high enough such that most of the noise-related detail coefficients have magnitudes less than the threshold but low enough such that the signal-related detail coefficients do not fall below the threshold. It is common to use a scaled version of the standard deviation of the noise in the detail coefficients at a particular level ($\sigma_j$) as the threshold ($\lambda_j$). If pure-noise data is available, that can be used to compute $\sigma_j$ directly. In many instances, it is hard to obtain such signal-free recordings and in such cases, the median absolute deviation (MAD) is a robust estimator for $\sigma_j$ and is given by

$$\sigma_j = \frac{\text{median}(|d_j[k] - \text{mean}(d_j[k])|)}{0.6745}$$

(5.7)

where the scaling factor of 0.6745 is for noise with a Gaussian amplitude distribution which is a valid assumption for the noise types discussed here.

Once $\sigma_j$ is known, a variety of techniques have been proposed for determining
\[ \lambda_j = \sigma_j \sqrt{2 \ln(N_j)} \]  

where \( N_j \) is the number of detail coefficients at level \( j \) and is the same as the length of the input signal itself in the case of the SWT.

After \( \lambda_j \) has been established, a variety of choices also exist for the thresholding scheme \( T \). These include the classical hard and soft \([189]\), firm \([192]\), non-negative garrote \([193]\), and others \([184, 194]\). The mathematical expressions for some of these thresholding schemes are shown below

\[
T_{\text{hard}}(x, \lambda) = \begin{cases} 
0, & \text{if } |x| \leq \lambda \\
x, & \text{if } |x| > \lambda 
\end{cases} \tag{5.9}
\]

\[
T_{\text{soft}}(x, \lambda) = \begin{cases} 
0, & \text{if } |x| \leq \lambda \\
\text{sign}(x)(|x| - \lambda), & \text{if } |x| > \lambda 
\end{cases} \tag{5.10}
\]

\[
T_{\text{garrote}}(x, \lambda) = \begin{cases} 
0, & \text{if } |x| \leq \lambda \\
x - \frac{\lambda^2}{x}, & \text{if } |x| > \lambda 
\end{cases} \tag{5.11}
\]

\[
T_{\text{firm}}(x, \lambda) = \begin{cases} 
0, & \text{if } |x| \leq \lambda_1 \\
\text{sign}(x) \frac{\lambda_2(|x| - \lambda_1)}{\lambda_2 - \lambda_1}, & \text{if } \lambda_1 < |x| \leq \lambda_2 \\
x, & \text{if } |x| > \lambda_2 
\end{cases} \tag{5.12}
\]

Of the thresholding schemes described above, only firm thresholding requires an additional threshold to be set. Figure 5.7 visualizes the effect of the schemes listed above with \( \lambda = 2 \) and the input \( x \) varying from -10 to 10. For firm thresholding, \( \lambda_1 = 2 \) and \( \lambda_2 = 2\lambda_1 = 4 \) are used. Hard thresholding leaves all coefficients larger
than the threshold unchanged. As a result, there often are discontinuities when
denoising using hard thresholding due to some noise coefficients that are just larger
than the threshold passing through unaltered. On the other hand, soft thresholding
suppresses all coefficients that are larger than the threshold, including those with
amplitudes much larger than the threshold that are likely to be from the signal. As a
result, while the denoised output is free of visual artifacts, the signal is significantly
smoother than in the case when hard thresholding is used. Garrote thresholding
offers a reasonable compromise between the two by gracefully transitioning from soft
thresholding for values near the threshold to hard thresholding for values that are far
from the threshold.

Figure 5.7: Input-output relationship for various thresholding schemes. $\lambda = 2$ for all
the schemes except firm where $\lambda_1 = 2$ and $\lambda_2 = 2\lambda_1 = 4$.

5.4.3 Comparison to Bessel filters

Bessel filters are commonly used to denoise recordings from nanopore and ion chan-
nels. Thus, the performance of wavelet denoising is directly compared to that of
fourth-order Bessel filters at different cutoff frequencies for simulated pulses. Figure 5.8a shows an example of a 1 µs pulse corrupted by white noise to an initial SNR of ≈ 2 dB. The noisy pulse is then filtered using fourth-order Bessel filters with cutoff frequencies of 1 MHz, 2.5 MHz, 5 MHz and 10 MHz respectively to mimic the processing described in Section 4.9. The same pulse is also denoised using five-level SWT with LDT and garrote thresholding. Figure 5.8b implements the same parameters but for $f^2$ noise instead of white noise.

Figure 5.8: Simulated 1 µs long pulse corrupted with (a) white and (b) $f^2$ noise to an initial SNR of approximately 2 dB. The pulse is filtered using fourth-order Bessel filters as well as wavelet denoising. Wavelet denoising offers good noise suppression while retaining edge sharpness.

Wavelet denoising performs especially well in the context of $f^2$ noise offering significantly lower noise than the Bessel filters while still retaining edge sharpness.

5.4.4 Effect of pulse width and initial SNR

It is important to also evaluate the effect of the pulse width on the performance of wavelet denoising. Figure 5.9a shows that over 500 iterations of the simulation, wavelet denoising offers good signal-shape retention when compared with the same set of fourth-order Bessel filters. The initial pulse width is varied from 100 ns to 1 µs and the initial SNR is set to be 14 dB with white noise PSD. The line $y = x$ represents the
output of the ideal filter that removes noise without distorting the signal whatsoever. Wavelet denoising retains pulse-width information better than both the 1 MHz and 2.5 MHz Bessel filters. Figure 5.9b repeats the simulation again but with $f^2$ noise instead of white noise. Here, wavelet denoising offers pulse-width retention similar to that offered by the 2.5 MHz Bessel filter.

![Graph](image)

Figure 5.9: 500 pulses are simulated with widths ranging from 100 ns to 1 µs and corrupted with (a) white and (b) $f^2$ noise to an initial SNR of approximately 14 dB. The pulse is filtered using fourth-order Bessel filters as well as wavelet denoising. The $y = x$ line represents the output of an ideal filter that does not distort the signal.

Besides the temporal response, it is also necessary to test the effect of the initial SNR on the performance of wavelet denoising. This is because, unlike Bessel filters, wavelet denoising is a non-linear operation. While the forward and inverse wavelet transforms themselves are linear operations, none of the thresholding schemes explored here are. Figure 5.10a shows that when averaged over 100 simulations, wavelet denoising offers SNR improvements compared to all of the Bessel filters evaluated here. For these simulations, a 1 µs pulse is corrupted with white noise to the initial SNR indicated. Wavelet denoising is performed using a five-level SWT with LDT and garroete thresholding. The actual calculation is for the reduction in the baseline noise level. However, since the amplitude of the signal remains unchanged in all of these filtering schemes, reduction in baseline noise directly translates to an improvement in SNR. For white noise, wavelet denoising offers an SNR improvement of over 5 dB com-
pared to the 2.5 MHz Bessel filter while exceeding its shape-retention characteristics (Figure 5.9a).

![Figure 5.10](image-url)

(a) 100 µs long pulses are simulated and corrupted with (a) white and (b) $f^2$ noise to a range of initial SNR values. The pulse is filtered using fourth-order Bessel filters as well as wavelet denoising. When averaged over 100 simulations, wavelet denoising offers consistent SNR improvements compared to all the Bessel filters evaluated here.

The SNR improvement is even higher when the noise is $f^2$ instead of white (Figure 5.10b). Here, wavelet denoising achieves $\approx 17$ dB improvement in SNR while simultaneously offering edge sharpness like that offered by the 2.5 MHz Bessel filter (Figure 5.9b).

### 5.5 Wavelet denoising of nanopore recordings

The simulations presented in the previous section were designed to mimic the nanopore recordings described in Section 4.9. Given the improvements that can be obtained in a simulated context, wavelet denoising was then applied to high-bandwidth nanopore recordings. The data analysed here are obtained using Pore 6 ($C_{pore} = 3$ pF, $d = 1.2 \text{ nm} \times 1.7 \text{ nm}$, $t_{eff} = 8 \text{ nm}$) in 3 M KCl (Figure 5.11). The pore is passivated with glass as described in Section 4.9 in order to reduce its membrane capacitance. 90 nt ssDNA is then added to the solution and its translocation through the nanopore is
recorded at a bias voltage of 700 mV. The SNR at 10 MHz bandwidth is greater than six.

Figure 5.11: TEM image of the 1.2 nm × 1.7 nm solid-state nanopore

Figure 5.12 shows a 0.2 s long time trace of the current through this nanopore. The raw data are filtered to bandwidths of 500 kHz, 1 MHz, 2.5 MHz and 5 MHz using fourth-order Bessel filters. The same data are also denoised using wavelets with eight-level SWT using LDT and garrotes thresholding. The numbers in the legend indicate the baseline noise level in units of pA_{RMS}, as calculated from the standard deviation of the time trace over an event-free 1 ms window. In these measured data where both white and f^2 noise contaminate the signal of interest, wavelet denoising performs remarkably well, offering lower noise than that achieved even by the 500 kHz Bessel filter.

The zoomed-in plot shown in Figure 5.12 shows that the wavelet denoised event has better shape retention than the 1 MHz filtered event and comparable to that of the 2.5 MHz filtered event. The baseline noise is reduced by approximately 12 dB compared to the 2.5 MHz Bessel filter. The similarities in the signal-shape preservation properties of the 2.5 MHz filter and wavelet denoising also hold in a statistical context. Figure 5.13 shows the amplitude vs dwell-time histogram of over 1000 translocation events. The dwell-time here is calculated as described in Section 4.5. The event count
Figure 5.12: Time trace of 90 nt ssDNA translocation through a nanopore in 3 M KCl at 700 mV bias. The same signal segment is filtered to different bandwidths using Bessel filters vs eight-level SWT with LDT and garrote thresholding. The numbers in the legend indicate the noise level (in pA_{RMS}) calculated over a 1 ms window. The expanded time-trace of a single event shows that wavelet denoising offers shape retention better than the 1 MHz filter and comparable to that of the 2.5 MHz filter.
(indicated in the legend) as well as the spread of events are extremely similar for the 2.5 MHz filtered data and the wavelet denoised data. Events were detected here with a simple thresholding algorithm (Section 4.5) with $\text{SNR}_{\text{min}} = 8$.

Figure 5.13: Amplitude vs dwell-time histogram showing that the wavelet denoised trace statistically offers nearly identical performance compared to the 2.5 MHz filter. Event count is indicated in the legend. Events were detected with an $8\sigma$ threshold from the baseline.

5.6 Wavelet denoising of ion channel recordings

Given the many similarities between ion channel and nanopore recordings, it is also possible to use the system designed in Chapter 3 for studying single ion channel activity. The ion channel used in this study [163] is the Type 1 ryanodine receptor (RyR1). RyR1 is a $\text{Ca}^{2+}$-induced $\text{Ca}^{2+}$ release channel that is found on the sarcoplasmic reticulum in skeletal muscle and has been implicated in major human diseases including heart failure [195], sudden cardiac death [196] and muscular dystrophy [197].

The packaging process for measuring ion channels is different from that described in Chapter 4 for solid-state nanopores. This process was developed by collaborators...
in the lab at Columbia University and consists of forming a microwell on top of the on-chip electrode, creating a lipid bilayer on this well, and incorporating a single ion channel into this bilayer. The packaging process is briefly described below.

Photoresist (Shipley 1813, Microchem) is patterned on the CMOS chip surface with 150 µm × 150 µm openings over the electrodes on a Suss MicroTec MA6 mask aligner. These openings are then used to etch the aluminum electrodes (Transene Al Etchant Type A), leaving the titanium nitride diffusion barrier underneath exposed. This step is similar to the aluminum etching described previously in Section 4.2.2. Next, 5 nm of titanium and 400–500 nm of silver are electron-beam evaporated (Angstrom EvoVac) on the surface covering the photoresist as well as the exposed titanium nitride electrode. The photoresist is lifted off in an acetone bath overnight. To form the SU-8 microwell, the chip surface is first primed with hexamethyldisilazane (HMDS-100%) (Transene). Next, a 5–7 µm thick layer of SU-8 3005 (Microchem) is spun onto the die and 30 µm diameter circular openings are patterned using an EVG 620 mask aligner with a dose of 300 mJ/cm².

The CMOS chips are then assembled onto the same package described in Section 4.2.1 and the Ag microelectrode is chlorided chemically as described in Section 4.2.2.

The lipids are prepared by mixing together 1,2-diphytanoyl-sn-glycero-3-phosphocholine (DPhPC) and DOPE at 5:3 (% mol) (both from Avanti Polar Lipids). The lipid mixture is dried under a nitrogen stream to remove the chloroform and then reconstituted in n-decane (Sigma-Aldrich). To form a membrane over the SU-8 well, a pipette tip is dipped into the lipid solution and then transferred into the buffer solution, where an air bubble is formed and brushed over the SU-8 opening, forming a bilayer. The formation of the bilayer is ascertained by applying a fixed potential difference between the two electrodes and monitoring the current flow. Once the bilayer is formed, it introduces a large series resistance between the two electrodes and consequently, the current drops to 0.
Proteoliposomes containing RyR1 are then prepared from rabbit skeletal muscle. Glycerol is added to the proteoliposome lumen to promote fusion with the bilayers using ultrasound treatment. Approximately 1–2 µL of the proteoliposome solution is added to the cis chamber which has a volume of approximately 1mL. The ion channel recordings are performed at room temperature in a buffer solution (7.3 pH) containing 20 mM Hepes, 1 M KCl, 1 mM EGTA, 1 mM ATP, and 40 µM free Ca^{2+}.

Figure 5.14 shows a schematic of this experimental setup. The data are recorded using the same acquisition system described in Section 3.4.4 with a sampling rate of 40 MSPS. However, since the signals here are not as large in amplitude as in the case of the solid-state nanopores, the data are first filtered with a fourth-order Bessel filter at 1 MHz and then downsampled to 4 MSPS.

Figure 5.14: Schematic of the high-bandwidth ion channel measurement setup. A single RyR1 channel is incorporated into a suspended lipid bilayer created on an SU-8 well fabricated directly on top of the CMOS amplifier chip.

Figure 5.15 shows an example of a 1 s long time trace showing single RyR1 channel activity at an applied bias voltage of 200 mV. The data are further filtered using
fourth-order Bessel filters with cutoff frequencies of 25 kHz, 100 kHz and 500 kHz. The 1 MHz filtered data are also denoised using a seven-level SWT with hard thresholding. The thresholds used here are extracted from the wavelet transform of pure-noise data recorded before the start of the experiment. This is necessary since the ion channel data show a relatively larger number of events compared to the nanopore data, which can lead to inaccurate estimation of the thresholds.

The expanded time traces shown in Figure 5.15 indicate that the wavelet denoised data retain temporal features better than those seen in the 100 kHz filtered data. The 25 kHz filter, on the other hand, significantly attenuates the amplitude of short-lived events.

The use of the hard thresholding scheme results in the upward spikes visible in Figure 5.15. Figure 5.16 plots a zoomed-in version of one such spike. The spike resembles the synthesis high-pass filter shown in Figure 5.6 and can be easily discarded from analysis because of its unphysical nature.

Figure 5.17 plots the all-points amplitude histogram for the time traces shown in Figure 5.15. Two key findings are evident. First, wavelet denoising offers noise performance comparable to that of the 25 kHz filter as shown by the width of the peak near the open channel current. Second, wavelet denoising preserves intermediate states in the histogram that are rendered invisible by the surrounding noise in the 100 kHz filtered data. The fact that the same peaks show up in the histogram for the 25 kHz filtered data confirms that these features are real and not an artifact of the wavelet denoising process.

5.7 Summary

This chapter discussed the use of wavelet transforms for denoising high-bandwidth nanopore and ion channel signals. Compared to Bessel filters, wavelet denoising offers a superior combination of shape retention and SNR in both simulations and
Figure 5.15: 1 s long time-trace of single ion channel recording filtered to different bandwidths using Bessel filters and denoised using a seven-level SWT with hard thresholding and LDT thresholds extracted from noisy data. The upward spikes in the wavelet denoised trace are artifacts created by the hard thresholding scheme. The expanded trace of a single event shows that wavelet denoising offers shape retention comparable to that of the 100 kHz Bessel filter and significantly better than the 25 kHz filter.
Figure 5.16: Hard-thresholding reconstruction artifact

Figure 5.17: All-points amplitude histogram of the time trace shown in Figure 5.15 with the Y-axis showing the bin count on a logarithmic scale. Wavelet denoising has noise performance comparable to that of the 25 kHz filter as shown by the width of the peak near the open channel current.
measurements. Combined with recently achieved high-resolution cryogenic electron microscopic structures of RyR1 [198, 199], improved SNR recordings hold the potential to provide important insights concerning ion channel regulation and mechanisms of dysfunction linked to human diseases that may lead to new therapies. Greater edge sharpness in nanopore recordings will particularly help in analyte detection experiments where accurate estimation of the dwell time is necessary [57].
Chapter 6

Design of an Integrated Multi-Clamp Amplifier

6.1 Introduction

Intracellular electrophysiological recordings from neurons is a high-fidelity neuroscience technique that enables the fundamental understanding of neuronal computation and function [5]. These recordings are typically performed using glass pipettes with micro or nanometer scale diameters mounted on a micromanipulator. The pipette and the bath are filled with solutions containing electrolytes and the pipette is lowered into the bath. An electrode placed in the pipette is then connected to an amplifier headstage. The amplifier is operated in either current- or voltage-clamp mode, depending on the study and parameters of interest (Figure 6.1).
Recording these µV-to-mV scale voltages in current-clamp and pA-to-nA scale currents in voltage-clamp necessitates the use of precision low-noise instrumentation amplifiers similar to the one described in Chapter 3. Benchtop amplifiers perform these recordings with high SNR [122]. However, they use discrete components in their design, increasing the cost and weight of these systems while limiting their bandwidth and scalability. IC-based solutions can address these problems but have been difficult to realize owing to the large resistance values required in these designs and the resultant limits in dynamic range. While several efforts have demonstrated either current-clamp [200], or voltage-clamp separately [125–127, 134, 201, 202], only a few have combined both and the requisite compensation circuitry onto a single chip [128, 129].

Here we explore the design and testing of a CMOS IC that includes both current- and voltage-clamp capabilities with a shared input and the ability to switch between the two modes. The IC was designed in a commercial CMOS 0.18µm process and occupies an area of 2.725 mm × 3.225 mm. The die photograph of the chip is shown in Figure 6.2.

6.2 Current-clamp

6.2.1 Design considerations

A current-clamp amplifier is required to measure intracellular or extracellular voltages with respect to a reference electrode located in the bath that the cell is placed in. Ideally, the amplifier should have a high-impedance input so as to not load the voltage source that is being measured. A voltage buffer with a high input impedance is commonly used as the first stage of a current-clamp amplifier. In the special case where the injected current $I_{inj} = 0$, the only additional circuitry that is needed is that required to compensate for $C_p$ (Figure 2.7). $C_p$ is typically of the order of a few pF.
Figure 6.1: Schematic depicting a typical intracellular recording from a neuron using a pipette. The amplifier can be operated in either current- or voltage-clamp modes and has dedicated compensation circuitry to account for the non-idealities associated with the pipette.

Figure 6.2: Die photograph of the multi-clamp amplifier chip
The exact magnitude of $C_p$ is governed by the specifics of the pipette geometry and the insertion depth [123].

In many cases, it is desirable to inject controlled amounts of current into the cell in order to elicit activity. If the cell is patched in the whole-cell configuration, a holding current may need to be applied in order to maintain the cell at its RMP. The magnitude of the current injected is usually in the pA–nA range and must be programmable. If a current is being injected into the cell through the pipette, a voltage drop of $I_{inj}R_S$ develops across $R_S$ and causes the measured voltage to be different from the membrane voltage. While circuit solutions to this problem are possible, if $R_S$ is determined before the start of the recording and is assumed to remain unchanged, the offset voltage can be exactly canceled in software.

### 6.2.2 Voltage buffer

The voltage buffer in the current-clamp circuit consists of an OTA in unity-gain feedback. Figure 6.3 shows a transistor-level schematic of the two-stage OTA used in this design. The first stage consists of a dual n- and p-input differential pair in a folded cascode configuration. The dual inputs allow for rail-to-rail input swing. The common-source second-stage allows for rail-to-rail output swing and provides further amplification. The amplifier is stabilized using Miller-compensation, with a series resistance to cancel out the right-half plane zero.

The input transistors are each sized $160\mu$m/$1\mu$m. The large sizes add to the total capacitance at the input of the OTA but help reduce flicker noise and voltage offsets due to random mismatch. Systematic offsets are minimized through the use of inter-digitation and by surrounding the input pair transistors with dummy devices. Thick-oxide variants are used for all the input-pair transistors in order to reduce gate leakage to sub-fA levels. The transistors are biased in extreme weak inversion in order to maximize their $g_m/I_d$. 

123
The bias circuitry adjusts the currents through the input-pair transistors in order to ensure a relatively flat $g_m$ across the input common-mode range. This is done by sensing the common-mode voltage at the differential inputs $(V_+ + V_-)/2$. Replica devices in the bias circuit are then used to steer current away from the bias generation block for the tail current sources when both sets of input pairs are on. When the common-mode voltage approaches $V_{DD}$, the PMOS input pair turns off. The bias generation block then increases the amount of current flowing through the NMOS input pair in order to maintain the $g_m$ approximately constant [203].

Under nominal biasing in the typical corner, simulations for the OTA showed that it is unity-gain stable and has a UGB of approximately 10 MHz. This is far in excess of the < 10 kHz bandwidth that is typically required for intracellular voltage measurements. However, this same OTA design is used across the chip and the bandwidth requirement is dictated by other blocks. The OTA occupies an area of 140 µm × 90 µm including the three units of the 20 µm × 20 µm compensation capacitor.

### 6.2.3 Capacitance compensation

The capacitance compensation circuit implemented here is based on the classical idea illustrated in Figure 2.8. The technique utilizes positive feedback. The filtered membrane voltage at the pipette ($V_p$) is sensed and buffered through the voltage buffer as $V_{buf}$. The buffered voltage is then multiplied by a scaling factor $\alpha$ with magnitude usually between 1 and 2. Lower values of $\alpha$ help reduce the noise injected by the capacitance compensation circuitry. $\alpha V_{buf}$ is then applied to one terminal of the injection capacitor with the other terminal connected to the input of the voltage buffer and the pipette establishing a potential difference of $(\alpha - 1)V_p$ across the capacitor if $V_{buf} \approx V_p$. Thus, the effect of $C_p$ can be reduced if $(\alpha - 1)C_{inj} \rightarrow C_p$. The stability of the capacitance compensation loop worsens as the value of $(\alpha - 1)C_{inj}$ approaches that of $C_p$ from below.
Figure 6.3: Circuit schematic of the dual-input OTA

Figure 6.4: Schematic of the capacitance compensation circuit. $A$ is a ten-bit programmable value that controls the gain while $B$ is a two-bit programmable value that controls the magnitude of the injection capacitor.
Figure 6.4 shows the block diagram of the on-chip implementation of the capacitance compensation circuitry. Here, $\alpha = A$ is a ten-bit digitally programmable value. Each of the opamps in non-inverting configuration provide a gain between 1 and 2 and proportional to the value of the appropriate bits of $A$. The upper path can be considered to provide the “fine” control over the amount of capacitance compensation while the lower one provides the “coarse” control. Splitting the amplification into two blocks increases the overall power consumption but decreases area and layout complexity and reduces the effect of parasitic capacitances in the programmable resistors.

The injection capacitor is made two-bit programmable ($B$ in Figure 6.4) with the unit having a magnitude of 5 pF. The possible values for $C_{inj}$ are thus 5 pF, 10 pF and 15 pF. The smallest change that can be made in $C_{inj}$ depends on the magnitude of $C_{inj}$ itself since the minimum change in $A$ is fixed at $1/1024$. Thus, $\Delta C = C_{inj}/1024$ and is approximately 5 fF, 10 fF and 15 fF respectively, for the three configurations possible here.

Matching between the programmable and fixed resistors of the non-inverting amplifier is critical to achieving accurate amplification since the absolute value of the resistance can vary due to process and temperature variations. Appropriate layout techniques are followed to ensure good matching. The capacitance compensation circuit occupies a total area of 980 $\mu$m $\times$ 440 $\mu$m.

### 6.2.4 Current injection

Figure 6.5 shows a simplified implementation of the current injection circuitry [122]. The pipette voltage is measured and then added to a command voltage $V_{command}$. This sum is then applied to one terminal of a large resistor while the other end of the resistor is connected to the input of the voltage buffer and the pipette. The effective voltage drop across the resistor is then $V_{command}$ (assuming $V_p \approx V_{buf}$) which causes
a proportional current to be injected into the pipette. Typically, the largest $V_{\text{command}}$ that can be applied is fixed and related to the supply voltage. Hence, the magnitude of the resistance used sets an upper limit on the largest current that can be injected.

![Simplified block diagram of the current-clamp current injection circuit](image)

Figure 6.5: Simplified block diagram of the current-clamp current injection circuit

Despite the limit they place on the maximum current that can be injected, large values of $R_{\text{inj}}$ are desirable to reduce the noise contributed by $R_{\text{inj}}$. However, a parasitic capacitance across $R_{\text{inj}}$ can act as a low impedance path for high frequency currents. Thus, a step change in $V_{\text{command}}$ will result in a current pulse with slow rise and fall times. Further, large resistances have traditionally been difficult to realize on a CMOS chip for reasons described in Section 3.3.4. The current injection circuit implemented here uses a resistor followed by an active current reducer similar to the feedback network described in Section 3.3.4 [132].

Figure 6.6 shows the circuit schematic of the on-chip current injection block. The summing block shown in Figure 6.5 is implemented using an opamp. The output of the opamp is $V_{\text{command}} - V_{\text{CM}} + V_{\text{buf}}$ which is applied across a resistor whose other end is held at $V_{\text{buf}}$ by the following opamp. The net current injected into the resistor is thus $(V_{\text{command}} - V_{\text{CM}})/R$. The common-mode voltage $V_{\text{CM}}$ is set to $V_{\text{DD}}/2$ and is connected to the bath electrode. $V_{\text{command}}$ is specified around this DC level. The injected current then goes through two stages of current division ($N = 32$) to yield a net current reduction of 1024. The value of injection resistance realized by
this structure is thus $R_{inj} = 1024 \times R$. Additional ratioed capacitors (not shown in figure) in parallel with the transistors in the current divider extend the current division to high frequencies.

![Circuit schematic of the current-clamp current injection block](image)

Figure 6.6: Circuit schematic of the current-clamp current injection block

In this design, $R = 100 \, \text{k}\Omega$ in the typical corner so that $R_{inj} \approx 100 \, \text{M}\Omega$. The current injection block occupies a total area of $950 \, \mu\text{m} \times 560 \, \mu\text{m}$ on the chip.

### 6.3 Voltage-clamp

#### 6.3.1 Design considerations

A voltage-clamp amplifier is required to measure ion channel currents in either cell-attached or whole-cell modes. The current flow is from the electrode in the pipette, through the ion channel in the cell membrane and back through the reference electrode in the bath. Ideally, the amplifier should have a low-impedance input so as to clamp the voltage irrespective of the magnitude of the current flow. A low-impedance input is also desirable for sinking in all of the signal currents that may be generated by ion channel activity. A TIA satisfies these requirements and is commonly used as the first stage of a voltage-clamp amplifier. Indeed, the simplest voltage clamp requires only the use of a TIA.
In many studies, it is desirable to change the clamp voltage instantaneously to observe the response of different ion channels to the voltage step. The instantaneous voltage step leads to charging currents in the parasitic capacitance at the amplifier’s input. In this work, the capacitance compensation circuit in the voltage-clamp is identical to the one described in Section 6.2.3. Resistance compensation is also desirable when the impedance of the pipette exceeds a few MΩ, in order to ensure accurate clamping.

Commercial voltage-clamp amplifiers routinely have feedback resistances in the TIA with values exceeding 100 MΩ. Larger gains in the TIA decrease the noise contribution from the feedback resistor but also place limits on the maximum amplitude of the signals that can be measured. As with the current-clamp, the input to the amplifier is preferred to have low levels of leakage current.

6.3.2 Transimpedance amplifier

The TIA implemented here is based on the principle of current amplification using negative feedback and transistors operating in the subthreshold regime [204]. Figure 6.7 shows the circuit schematic of the TIA. The first two stages provide current amplification by a factor of N each. Accurate current amplification relies on good matching between the devices connected to the output of the opamps. If the current through the pipette flows in the direction indicated in the figure, the output of the first opamp goes high, passing the current through the PMOS in the feedback path. However, since the magnitudes of these currents are small, the PMOS is in deep subthreshold and consequently, the output voltage depends logarithmically on the input current. It is possible to measure the logarithmic output voltage directly and convert it into a corresponding current [105, 129]. However, here the same output voltage is applied across N transistors in parallel. Since the N transistors experience the same potentials at their gate, source and drain as the feedback transistor, the current is
amplified \( N \) times. Similarly ratioed capacitors (not shown in figure) in parallel with these transistors extend the amplification to high frequencies.

\[
V_{\text{out}} = V_p + N^2 R I_p
\]

Figure 6.7: Circuit schematic of the TIA in the voltage-clamp

The amplified current is then passed through a final transimpedance stage with a four-bit programmable feedback resistance. The resistance value can be changed uniformly from 0 to 225 k\( \Omega \) in steps of 15 k\( \Omega \). The effective feedback resistance value is then \( R_F = N^2 R \) and can be as large as 231 M\( \Omega \) in this design with \( N = 32 \). The diodes in parallel with the feedback resistor limit the range of output voltages that are linear with respect to the input current but allow for the operation of the TIA in the same manner as explored previously once appropriate calibrations have been performed [129].

A key difference between the TIA designed here and that available in commercial voltage-clamps is the 3 dB bandwidth and the thermal noise from the feedback resistance. The 3 dB bandwidth is determined by \( R_F \) and \( C_F \) in conventional designs. However, in this design, \( C_F \) appears in parallel with \( R_F/N^2 \). Thus, the worst-case 3 dB bandwidth set by \( R_F \) and \( C_F \) for this TIA is 700 kHz. However, the current amplification stage sets a cutoff frequency of approximately 350 kHz which, consequently, limits the overall 3 dB bandwidth. Note that this is significantly greater than the \( 1/(2\pi R_F C_F) = 700 \) Hz limit that would have been set if the TIA had been
implemented using passive components with the same values of \( R_F \) and \( C_F \).

The input-referred noise contribution of the feedback resistance in a TIA with \( R_F = 231 \, \text{M}\Omega \) is \( 4kT/R_F = 6.93 \, \text{pA}^2/\text{Hz} \). Since the physical resistor used here is \( N^2 \) times smaller, the noise it generates is \( N^2 \) times larger. However, when this noise is referred back to the pipette, it sees a division by a factor of \((N^2)^2\) such that the input-referred noise from the feedback resistance for the TIA designed here is \( 4kT/N^4 \times (N^2/R_F) \) which is \( N^2 = 1024 \) times smaller than the noise generated by a passive resistor with value \( R_F \).

These advantages come at the expense of increased power consumption and design complexity in the current amplification stage. The TIA designed here occupies a total area of \( 910 \, \mu\text{m} \times 410 \, \mu\text{m} \).

### 6.3.3 Resistance compensation

The classical resistance compensation scheme shown in Figure 2.9 utilizes positive feedback. However, this technique is susceptible to ringing when the percentage of the resistance compensated exceeds 80\%. In this work, the resistance compensation instead is based on the state-estimator theory proposed previously [205]. For a current \( I_p \) flowing through the pipette, the membrane potential can be expressed as

\[
V_M = V_p - I_p R_S
\]

where \( V_p \) is the voltage applied to the positive terminal of the TIA and appears on the pipette through the clamping action of the TIA and \( R_S \) is the series resistance of the pipette. In the ideal case where \( R_S = 0 \), the membrane voltage is exactly equal to \( V_p \).

Figure 6.8 shows the block diagram for the feedback loop implemented in this work. The current flowing through the pipette is passed through an on-chip estimate of the membrane resistance \( R_{S,\text{est}} \) which generates a proportional voltage \( I_p R_{S,\text{est}} \).
This gives a local estimate of the membrane voltage as

\[ V_{M,est} = V_p - I_p R_{S,est} \]  \hspace{1cm} (6.2)

This is then compared to the command voltage and the error is fed to an integrator. The output of the integrator generates \( V_p \). If the sign of the overall feedback around the integrator is negative and the loop is acting as intended, the action of the integrator will be to drive its input to zero. This implies the following relation

\[ V_{\text{command}} = V_{M,est} \]  \hspace{1cm} (6.3)

The right-hand side of this equation equals \( V_M \) when \( R_{S,est} = R_S \) and indicates full \( R_S \) compensation if the loop is stable.

\[ \text{Figure 6.8: Block diagram of the state-estimator resistance compensation circuit} \]

If the loop is broken at the input to the integrator, the loop gain can be written as (ignoring the negative sign)

\[ L(s) = \frac{\omega_u}{s} \left( 1 - \frac{R_{S,est}}{Z_{\text{cell}}} \right) \]  \hspace{1cm} (6.4)

where \( Z_{\text{cell}} \) represents the net looking-in impedance into the pipette and can be written for the whole-cell configuration (neglecting the effect of \( R_M \)) as

\[ Z_{\text{cell}} = \frac{\left( R_S + \frac{1}{sC_M} \right) \frac{1}{sC_p}}{R_S + \frac{1}{sC_M} + \frac{1}{sC_p}} \]

\[ \approx \frac{1}{sC_M} \frac{1 + s\tau_M}{1 + s\tau_p} \quad \text{if} \quad C_M >> C_p \quad (6.5) \]
where $\tau_M = R_SC_M$ and $\tau_p = R_SC_p$.

From Equations (6.4) and (6.5), we have

$$L(s) = \frac{\omega_u}{s} \left( 1 - \frac{s R_{S,est} C_M (1 + s \tau_p)}{1 + s \tau_M} \right)$$

$$= \frac{\omega_u}{s} \left( \frac{1 + s \tau_M (1 - \alpha) - \alpha s^2 \tau_M \tau_p}{1 + s \tau_M} \right)$$

where $\alpha = R_{S,est}/R_S$ and represents the fractional compensation. If the time constant attributed to $\tau_p$ can be made small enough and the left-half plane zero in the numerator appears before the unity-gain crossing, then the loop will be stable with very good phase margin. Additional poles can be introduced after the unity-gain crossing to account for the high-frequency zero added by $\tau_p$. Note that for stable operation, $C_p$ must be reduced either physically or by electronic compensation. The need for the reduction of $C_p$’s impact can be understood by the fact that the current that flows into $C_p$ does not flow through the pipette’s $R_S$ and does not affect $V_M$. Thus, any such component in the measured current leads to an inaccurate local estimate of the membrane voltage.

Figure 6.9 shows the circuit-level implementation of the resistance compensation block. Compared to discrete systems where it is difficult to record $I_p$ at high-bandwidths, the current amplification circuit used to pre-amplify $I_p$ is based on the same structure used in the TIA and has extremely high measurement bandwidth. This preamplification also removes the need for a physical large-valued resistor. The current is first preamplified by a factor of $N \times N_1 \times N_2 = 32 \times 4 \times 8 = 1024$, and passed through a TIA with a ten-bit programmable feedback resistor. The feedback resistance can be changed uniformly from 0 to $256 \text{k} \Omega$ in steps of $250 \Omega$ allowing for compensation of a resistance as large as $262 \text{M} \Omega$.

The output of this TIA is then $V_{M,est}$, which is compared to $V_{\text{command}}$ as depicted in the block diagram (Figure 6.8). The difference must then be fed to an integrator. The differencing is realized by connecting the voltages to the terminals of a five-bit programmable $g_m$ block, implemented using a differential amplifier with the bits
controlling the amount of bias current in the input transistors of the differential pair. The output of the $g_m$ block is connected to an OTA with capacitive feedback ($C = 64\,\text{pF}$) which implements the integration as well as keeps the output voltage of the $g_m$ block constant for improved linearity. $\omega_u$ in Equation (6.6) is then equal to $g_m/C$.

The block occupies a total area of $1230\,\mu\text{m} \times 900\,\mu\text{m}$ on-chip.

6.4 Board design

6.4.1 System overview

Custom PCBs are designed for validating the chip’s functionality both electrically and in vitro. The design is split across two boards - a small daughterboard housing the chip and the sensitive analog circuitry, with a larger motherboard containing the data acquisition and the digital circuitry. Figure 6.10 shows a block diagram of this setup.

The design goal was to make the daughterboard as small and light as possible for potential use in experiments with moving mice. As a result, the daughterboard is created on a two-layer 3.56 mm × 5.08 mm PCB. Ground and power planes are avoided so as to limit the overall weight. The weight of the unassembled PCB is 2.02 g while the weight of the fully-assembled PCB is approximately 4.5 g. This weight can be further reduced as more board-level components are integrated onto the chip.

The daughterboard is connected to the motherboard by means of two flex-cables. Power and analog signals are transferred over one cable while digital signals are transferred over the other. The net impedance of the flex-cables is high enough so as to lead to high-frequency noise spikes in measurements. These spikes are beyond the bandwidth of interest and do not affect circuit performance.
Figure 6.9: Circuit-level implementation of the resistance compensation circuit

Figure 6.10: Multi-clamp amplifier board-level block diagram
6.4.2 Chip packaging and assembly

In order to reduce the weight of the daughterboard and parasitics on the input nodes, the multi-clamp amplifier chip is wirebonded directly to the PCB. The landing pads on the PCB are first cleaned by soaking in BPS-106 (Versum Materials) to clean the landing pads and improve yield during bonding. The chip is attached to the PCB using Epo-Tek H20E (Epoxy Technology). The chip has 70 wirebonds in total. After the wirebonding is complete, the chip is encapsulated using Epo-Tek OG116-31 (Epoxy Technology) to protect the wirebonds from damage during handling. The weight of the PCB after chip assembly and encapsulation is 2.10 g. Figure 6.11 shows a photograph of the daughterboard PCB with the chip assembled.

Figure 6.11: Photograph of the multi-clamp daughterboard with the chip attached

6.4.3 Power domains

The daughterboard is powered by a single 6 V DC power supply that is connected via the motherboard using the flex-cable. A 3.3 V regulator is used to power all the components on the daughterboard, including the chip. Since power planes are absent on the daughterboard, the power and ground routing is done by hand. While
connecting the DC power supply directly to the daughterboard instead of via the flex-
cables reduced some high-frequency spikes in the noise spectrum, this is not favored
since it would require additional wiring to the daughterboard.

The motherboard is powered using two power supplies. A 6 V supply is used
to generate a 3.3 V supply for powering the anti-aliasing filters, part of the digital
isolators and the digital pins on the ADC as well as for generating a 5 V supply for
powering the analog portion of the ADC. A separate 5 V supply is used exclusively
to power the Opal Kelly FPGA.

The motherboard employs a split ground-plane design. The “analog” ground
plane is connected to the ground of the daughterboard, the analog components on
the motherboard and the ADC, and is completely isolated from the “digital” ground
of the FPGA. Digital isolators are used for all data exchange across ground planes.
The split-plane design helps reduce noise that might couple in capacitively to the
voltage-clamp amplifier’s input node.

6.4.4 Data acquisition

The chip has three primary outputs of interest - the buffered pipette voltage $V_{buf}$
in current-clamp, and the pipette voltage $V_p$ and the output of the TIA $V_{out,TIA}$ in
voltage-clamp. The chip outputs are passed through an analog buffer and fourth-
order Bessel anti-aliasing filters implemented using a topology similar to that shown
in Figure 3.10 but with a cutoff frequency of 100 kHz. The outputs of the anti-aliasing
filter are then digitized at 200 kSPS each using a six-channel time-interleaved ADC
with 16-bit resolution.

The ADC outputs are then passed through digital isolators that transmit the data
to an Opal Kelly FPGA (XEM6010). Data is then transferred from the FPGA to a
host PC over a USB 2.0 link. The data rate is $16 \times 6 \times 200 \times 10^3 = 19.2 \text{ Mbit s}^{-1} = 2.4 \text{ MB s}^{-1}$
and is well within the 30 MB s$^{-1}$ real-world limits of the USB 2.0 link.


6.4.5 Shielding and assembly

Many of the shielding considerations here are the same as those described in Section 3.4.5, especially for the voltage-clamp configuration. A small aluminum box is used as a Faraday cage. The daughterboard PCB is affixed to the box via screws. The pipette holder and the pipette, however, necessarily need to extend outside the box. Adding a ground shield around the pipette risks increasing $C_p$. Instead, the optical table that the setup is placed on is connected to the ground of the daughterboard using a low-impedance connection. In our experiments, this generally provided sufficient shielding from EMI.

Appropriate holes are drilled in the aluminum box for the flex-cables and the pipette holder to pass through. The hole for the pipette holder is made just large enough for it to fit so that the holder remains firmly secured to the box. A separate aluminum piece with a dovetail cut is attached to the box via screws. This dovetail piece is then attached to the micromanipulator during biological experiments. The ground of the micromanipulator is connected to the Faraday cage.

Commercial systems generally have the notion of a “signal ground” that is connected to all the connections mentioned previously as well as the bath. This is possible since these systems employ bipolar supplies. In the system described here, a unipolar supply is used. The bath potential is thus distinct from the ground of the system. In fact, the bath is tied to a potential that is nominally $V_{DD}/2 = 1.65\,\text{V}$.

Figure 6.12 shows a photograph of the daughterboard inside the Faraday cage with the pipette holder and flex-cables attached.

6.5 Electrical characterization

The amplifier’s electrical functionality was first evaluated using a model cell built using electrical components as shown in Figure 2.6. The pipette was modeled using a
single $R_S$ mimicking the whole-cell configuration. In the measurements described below, $R_S = 20 \, \text{M}\Omega$, $R_M = 100 \, \text{M}\Omega$ and $C_M = 0$ in the current-clamp testing and $20 \, \text{pF}$ in the voltage-clamp testing, unless indicated otherwise. $C_p$ was largely determined by the parasitic trace capacitance to ground on the PCB.

### 6.5.1 Current-clamp

The noise performance of the voltage buffer is first evaluated. This is done by connecting a low-pass filtered $1.65 \, \text{V} \, \text{DC}$ voltage source directly to the input of the voltage buffer bypassing the model cell completely. Figure 6.13a shows an example output time trace of the voltage buffer filtered in software to different bandwidths using digital approximations of fourth-order Bessel filters. In a $10 \, \text{kHz}$ bandwidth, the RMS value of the input-referred voltage noise is $20 \, \mu\text{V}_{\text{RMS}}$. However, this noise is found to be dominated by components on the PCB rather than the chip itself and can potentially be reduced further by using lower-noise equivalents. This is not pursued here since the noise levels were already low enough for the intended application. Figure 6.13b shows the PSD for a $3 \, \text{s}$ long noise measurement.
Figure 6.14 shows the output of the voltage buffer with and without capacitance compensation. A function generator is used to inject a 200 mVpp 1 kHz square wave with a DC offset of 1.65 V at \( V_M \). The voltage is thus applied to the input of the voltage buffer through \( R_S \). The effect of \( R_S \) and \( C_p \) is to low-pass filter the input and slow down the rise and fall times of the square wave. The value of the parasitic \( C_p \) is estimated to be approximately 3.5 pF and the resulting 10%–90% rise time is approximately 350 \( \mu \)s. With capacitance compensation, however, the rise time drops to less than 20 \( \mu \)s and approaches the rise time constraints imposed by the 100 kHz Bessel anti-aliasing filter implemented in hardware.

Figure 6.15a plots the linearity of the injected current as a function of the command voltage. For this measurement, \( R_S = 100 \text{ M}\Omega \) and \( R_M = 0 \). \( V_{\text{command}} \) is swept from \(-100 \text{ mV} \) to \( 100 \text{ mV} \) and is changed in steps of 1 mV. The injected current should thus change from \(-1 \text{ nA} \) to \( 1 \text{ nA} \) in steps of 10 pA. The figure also plots a least-squares straight line fit to the measured data. The deviation of the measured data from the fit is due to offset voltages in the opamps that form the current divider [200]. It is worth noting that the effective transconductance is slightly higher for negative currents and lower for positive currents than the desired value. If needed, this could be eliminated by appropriate calibration.

Figure 6.15b plots the buffer’s output voltage as \( V_{\text{command}} \) is stepped from \(-5 \text{ mV} \) to \( 5 \text{ mV} \) about \( V_{CM} \) at a rate of 2 Hz. This corresponds to an injected current square wave of amplitude 100 pApp. In the absence of capacitance compensation, the rise and fall times of the measured voltage are slow while with capacitance compensation, they are sped up considerably.

### 6.5.2 Voltage-clamp

The diodes in the feedback path of the TIA allow for a linear range of operation of approximately \( \pm 0.8 \text{ V} \). The output voltage of the TIA in the voltage clamp as shown
Figure 6.13: (a) 200 ms long time trace of the output of the voltage buffer connected to a fixed DC voltage source filtered using 1 kHz, 5 kHz and 10 kHz fourth-order Bessel filters. RMS values of the noise at the filtered time-traces are indicated. (b) Input-referred noise PSD of the current-clamp voltage buffer.

Figure 6.14: For a 200 mV<sub>pp</sub> square wave at 1 kHz injected through a 20 MΩ resistor, capacitance compensation improves the rise and fall times from approximately 350 µs to less than 20 µs.
(a) Linearity of the injected current as a function of the command voltage.

(b) Step response of the injected current with and without capacitance compensation.

Figure 6.15: Linearity and step response of the current injection circuit.
in Figure 6.7 can be mapped to the input current through the relation

\[ I_p = \frac{V_{out} - V_p}{R_F} \]  \hspace{1cm} (6.7)

Thus, the range of input currents for which the output voltage is linearly related to the current depends on the gain setting.

Figure 6.16a shows the linearity of the measured voltage as the input current is swept from \(-500\) pA to \(500\) pA in steps of \(10\) pA with \(R_F \approx 120\) M\(\Omega\). The current is injected through \(R_S + R_M = 120\) M\(\Omega\) by stepping the voltage applied at the other end of \(R_M\). A linear least-squares fit is then generated from the measured data. Since the current amplifier in the TIA employs a similar topology as the divider in the current injection circuit of the current-clamp, there is a similar deviation from the linear fit (Figure 6.15a). By scaling positive and negative valued currents by different but fixed amounts, the true current can be obtained from the measured data. This is shown in Figure 6.16b where all currents \(> 0\) are divided by 1.042 while currents \(< 0\) are divided by 1.133. The line shown in the figure represents the ideal \(y = x\) input-output characteristic and the adjusted data is remarkably similar. The implication of this result is that the gains are slightly different for positive and negative currents.

Figure 6.17a shows a 0.2 s long time-trace of the output of the TIA in the open-headstage configuration with \(R_F \approx 225\) M\(\Omega\) and adjusted for linearity as described above. The same data are filtered to different bandwidths in software using digital approximations of fourth-order Bessel filters. At a bandwidth of 5 kHz, the input-referred RMS noise is 225 fA\(_{\text{RMS}}\), despite the relatively low value of the feedback resistance. Figure 6.17b shows the input-referred noise PSD of the TIA in the same configuration as calculated over a 3 s long recording. When compared to the Axopatch 200B (\(R_F = 500\) M\(\Omega\)), a commercial system used for patch-clamp recordings, the noise performance achieved here is significantly better.

The \(C_p\) compensation circuitry in the voltage-clamp is identical to that used in the current-clamp. Its functionality is verified by connecting a purely capacitive load to
(a) Measured linearity of the measured current as a function of the input current amplitude. The line provides a least-squares fit to the measured data.

(b) Adjusted measured current as a function of the input current. The measured data is multiplied by different factors depending on whether the value is positive or negative. The line shows the ideal $y = x$ input-output characteristic.

Figure 6.16: Linearity of the voltage-clamp TIA as a function of input current amplitude
the input of the TIA and stepping $V_{command}$ as shown in Figure 6.18 (data filtered to 10 kHz bandwidth). In the absence of any capacitance at the TIA’s input, the current recorded by the TIA must be constant at 0. Steps in $V_{command}$ inject charging currents into the capacitor which are recorded by the TIA. After tuning the $C_p$ compensation circuitry to cancel out $\approx 2 \text{ pF}$ of capacitance, the recorded current does not show any spikes, indicating the $C_p$ compensation circuitry works as expected.

Finally, the functionality of the $R_S$ compensation circuit is evaluated. Figure 6.19 shows the result of these measurements filtered to a 4 kHz bandwidth using a fourth-order Bessel filter. Here, $C_M$ is not populated on the PCB and the parasitic $C_M$ is estimated to be a few tens of fF. This value of $C_M$ is appropriate for small structures such as spines, where $C_M \approx 30 \text{ fF}$. The resistances $R_S$ and $R_M$ are both set to 100 M$\Omega$. $V_{command}$ is stepped from $-30 \text{ mV}$ to $30 \text{ mV}$ in steps of 5 mV with each step being held for 500 ms. Before enabling $R_S$ compensation, the parasitic capacitance $C_p$ is first compensated. In the absence of $R_S$ compensation, $V_{command}$ appears across $R_S+R_M = 200 \text{ M}\Omega$ and generates a maximum current amplitude of 150 pA. When the $R_S$ compensation is tuned to 50 $\%$, the net resistance is $R_M + (1 - \alpha)R_S = 150 \text{ M}\Omega$ and generates a maximum current of 200 pA. When the $R_S$ is almost completely compensated out, all of the applied voltage falls across $R_M$ and generates a maximum current of approximately 300 pA.

A similar test is performed with a value of $C_M = 20 \text{ pF}$ with $R_S = R_M = 100 \text{ M}\Omega$. This situation is similar to what would be seen if the soma were being clamped. In this case, full compensation would lead to loop instability, but high compensation levels are still possible. Figure 6.20 shows the result of $V_{command}$ being stepped from $-50 \text{ mV}$ to $50 \text{ mV}$ in steps of 5 mV with no $R_S$ compensation and $\alpha = 0.83$ (filtered to 1 kHz bandwidth). With no compensation, the current varies between $\pm 250 \text{ pA}$ while with compensation, the current varies between $\pm 425 \text{ pA}$. The larger spikes during transition with compensation enabled are the result of increased charging transients through $C_M$. As the effective value of $R_S$ is reduced, the voltage across
Figure 6.17: (a) 200 ms long time trace of the output of the TIA in open-headstage configuration with $R_F \approx 225 \, \text{M}\Omega$ filtered using 1 kHz, 5 kHz and 10 kHz fourth-order Bessel filters. RMS values of the noise at the filtered time-traces are indicated. (b) Input-referred noise PSD of the voltage-clamp TIA and the Axopatch 200B.

Figure 6.18: For a 10 mV$_{pp}$ square wave at 1 Hz applied at $V_{\text{command}}$ for a TIA with only a capacitance connected to its input, $C_p$ compensation removes charging transients from the recorded current.
more closely resembles the desired staircase waveform and consequently creates larger transient currents.

### 6.5.3 Comparison to state of the art

Table 6.1 compares the performance of the multi-clamp amplifier chip to existing commercial discrete as well as integrated solutions. The amplifier designed in this work compares favorably with all the other systems and improves the noise performance by over a factor of 3 in voltage-clamp.

<table>
<thead>
<tr>
<th></th>
<th>MultiClamp 700B [122]</th>
<th>Goldstein et al. [128]</th>
<th>Harrison et al. [129]</th>
<th>This work</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Type</strong></td>
<td>Discrete</td>
<td>IC</td>
<td>IC</td>
<td>IC</td>
</tr>
<tr>
<td><strong>Technology</strong></td>
<td>-</td>
<td>0.5 µm SoS</td>
<td>0.35 µm CMOS</td>
<td>0.18 µm CMOS</td>
</tr>
<tr>
<td><strong>Die size</strong></td>
<td>-</td>
<td>4 mm × 8 mm</td>
<td>4.7 mm × 3.0 mm</td>
<td>3.23 mm × 2.73 mm</td>
</tr>
<tr>
<td><strong>Supply voltage</strong></td>
<td>-</td>
<td>3.3 V</td>
<td>-</td>
<td>3.3 V</td>
</tr>
<tr>
<td><strong>Power consumption</strong></td>
<td>30 W</td>
<td>30 mW</td>
<td>-</td>
<td>7 mW</td>
</tr>
<tr>
<td><strong>$C_p$ compensation range</strong></td>
<td>0–36 pF (VC)</td>
<td>0–10 pF</td>
<td>0–10 pF</td>
<td>0–15 pF</td>
</tr>
<tr>
<td><strong>Input-referred voltage noise in CC</strong></td>
<td>-</td>
<td>150 µV$_{RMS}$ (5 kHz)</td>
<td>8.2 µV$_{RMS}$ (10 kHz)</td>
<td>20 µV$_{RMS}$ (10 kHz)</td>
</tr>
<tr>
<td><strong>VC TIA gain</strong></td>
<td>50 MΩ–50 GΩ</td>
<td>49 kΩ–100 MΩ</td>
<td>Non-linear</td>
<td>0–225 MΩ</td>
</tr>
<tr>
<td><strong>Input-referred current noise in VC (5 kHz)</strong></td>
<td>0.8 pA$_{RMS}$</td>
<td>3.3 pA$_{RMS}$</td>
<td>1.1 pA$_{RMS}$</td>
<td>0.23 pA$_{RMS}$</td>
</tr>
<tr>
<td><strong>Rs compensation range</strong></td>
<td>0.4–744.7 MΩ</td>
<td>0–100 MΩ</td>
<td>0–32 MΩ</td>
<td>0–250 MΩ</td>
</tr>
</tbody>
</table>

Table 6.1: Comparison of the integrated multi-clamp amplifier to existing state of the art. CC indicates current-clamp and VC indicates voltage-clamp.

### 6.6 Experimental results

Experiments involving cells were performed in collaboration with the Yuste lab at Columbia University. The amplifier and its Faraday cage are mounted on a microma-
Figure 6.19: For $R_S = R_M = 100 \, \text{M} \Omega$ and $V_{\text{command}}$ varying from $-30 \, \text{mV}$ to $30 \, \text{mV}$ in steps of $5 \, \text{mV}$, the amount of resistance compensated is programmable and can compensate $R_S$ nearly completely.

Figure 6.20: For $R_S = R_M = 100 \, \text{M} \Omega$ with $C_M = 20 \, \text{pF}$ and $V_{\text{command}}$ varying from $-50 \, \text{mV}$ to $50 \, \text{mV}$ in steps of $5 \, \text{mV}$, resistance compensation increases the steady-state current. The spikes with compensation enabled are larger due to the increased charging transients through $C_M$. 

148
nipulator housed with a custom-slice microscope setup. The first set of experiments involved the use of a 100 nm diameter high-impedance sharp microelectrode filled with 3 M KCl. These microelectrodes are particularly useful for targeting small structures such as spines [206]. The amplifier is configured in current-clamp mode and a 100 pA_{pp} square-wave current is injected into the pipette. Figure 6.21 shows the measured voltage response of the pipette immersed in a bath containing artificial cerebrospinal fluid (ACSF) comprising (in mM) 126 NaCl, 26 NaHCO_3, 1.145 NaH_2PO_4, 10 glucose, 3 KCl, 2 MgSO_4, and 2 CaCl_2. The capacitance compensation circuitry is tuned to cancel out approximately 8 pF of parasitic capacitance. The response indicates a measured resistance of 90 MΩ with slight overcompensation of the capacitance.

![Current injection through a sharp microelectrode in the bath](image)

**Figure 6.21:** Current injection through a sharp microelectrode in the bath

The pipette is then used to perform intra- as well as extracellular recordings from cortical layer-5 pyramidal neurons from day-40 wild-type mouse slices in ACSF. The temperature of the bath is set to 37°C. The measured resting membrane potential is −58 mV. Prior to cell entry, distinct extracellular neuronal action potentials are ob-
served. After entering the cell, intracellular action potentials are observed with high SNR, millisecond time-scales and approximately 50 mV amplitudes. These recordings are shown in Figure 6.22 and compare favorably to a spike-triggered average of recordings made using the MultiClamp 700B (Molecular Devices) in terms of SNR, timescales and signal fidelity.

![Current-clamp recordings from neurons showing intra- as well as extracellular action potentials. The intracellular signals compare favorably to measurements made using the MultiClamp 700B in terms of SNR, timescales and signal fidelity.](image)

Next, experiments are performed in neuronal cultures using patch pipettes (Figure 6.23). The internal solution for the patch pipettes is comprised of (in mM) 130 K-gluconate, 5 NaCl, 2 MgSO$_4$, 10 HEPES, 5 EGTA, 4 MgATP, 0.4 Na$_2$GTP, 7 Na$_2$-phosphocreatine, 2 pyruvic acid, pH adjusted to $\approx 7.2$. With the pipette in the bath, the bath potential is first adjusted so as to bring the pipette current to zero. Next, a pulse voltage is applied at $V_{\text{command}}$ with amplitude of 5 mV and frequency of 1 Hz. Figure 6.24a shows the current recording (filtered to 1 kHz bandwidth) through one such pipette. This allows for the characterization of $R_S$ which is typically between 7 and 14 M$\Omega$ for the pipettes used in this study. As the pipette approaches the cell membrane and suction is applied, the access resistance increases, as indicated by the decrease in the amplitude of the current pulse shown in Figure 6.24b. Finally, when the giga-seal is formed, the amplitude of the pulse waveform becomes negligibly
small as seen in Figure 6.24c. The spikes are due to the remaining uncompensated $C_p$.

![Image](image.png)

Figure 6.23: Photograph of a patch-clamp measurement from a neuron. The pipette tip is approximately 1 µm in diameter.

![Image](image.png)

Figure 6.24: Voltage-clamp current recordings of a patch pipette (a) in the bath, (b) as the pipette approaches the cell membrane and suction is applied, and (c) when the gigaseal is formed. The spikes visible after gigaseal formation are due to residual uncompensated $C_p$.

Measurements with patch pipettes are also performed on 3-D cultured hippocampal neurons grown as spheres [207, 208]. Figure 6.25a shows a recording of the current through the pipette as measured by the amplifier in voltage-clamp mode, filtered down to a bandwidth of 2 kHz in software. Here, although a gigaseal is not formed, the seal resistance is measured to be a few hundreds of MΩ. The patch is held at $−70$ mV.
with respect to the bath. The cell showed spontaneous activity and the voltage-clamp measured several action potentials, one of which is shown in Figure 6.25b.

Figure 6.25: (a) Voltage-clamp recording through a patch pipette in loose-seal configuration. The cell showed spontaneous activity when the patch was held at a potential of $-70 \text{ mV}$ with respect to the bath. (b) A zoom-in of the event near the 40 s mark reveals high SNR and millisecond time scales for the action potential.

Current-clamp measurements are also performed from a neuron in tight-seal configuration. These results are shown in Figure 6.26a after being filtered down to 10 kHz bandwidth. Over the course of several seconds, the cell showed spontaneous activity as current is injected to maintain approximately $-50 \text{ mV}$ at the patch with respect to the bath. The zoomed-in events shown in Figure 6.26b reveal amplitudes of several mV indicating that these were likely tightly-coupled extracellular action potentials [209]. Figure 6.27 (data filtered to 2 kHz bandwidth) shows two excitatory postsynaptic potentials (EPSP) and an inhibitory postsynaptic potential (IPSP) which are not as large as the action potentials that are observed.

6.7 Summary

This chapter discussed the design of an integrated voltage- and current-clamp amplifier with current injection capability and compensation circuitry to account for
Figure 6.26: (a) Current-clamp recording through a patch pipette in tight-seal configuration. The cell showed spontaneous activity when the patch was held at a potential of $-50 \text{ mV}$ with respect to the bath. (b) A zoom-in of the event near the 2.1 s mark reveals high SNR and that the events were tightly-coupled extracellular action potentials.

Figure 6.27: Current-clamp recordings of EPSP and IPSP
the pipette’s resistive and capacitive non-idealities. The system is experimentally validated \textit{in vitro} with signals showing high SNR both in intracellular as well as extracellular configurations. A comparison reveals that the amplifier performs well when compared to the current state of the art while improving on important metrics such as the noise performance and power consumption.
Chapter 7

Conclusions

7.1 Summary of contributions

This dissertation describes a body of work related to the design of amplifiers for nanoscale sensors and the subsequent signal conditioning when measuring biochemical signals. The underlying theme was to miniaturize these amplifiers while exploiting the resultant decrease in parasitics for improvement in performance. The first such work is a CMOS amplifier optimized for high-bandwidth nanopore recordings. By designing a low-capacitance platform and integrating passivated solid-state nanopores, noise at high frequencies is reduced considerably compared to the state of the art. This allows for the single-molecule measurement of DNA translocation through the nanopore at heretofore unprecedented timescales revealing dynamics that would otherwise have been rendered invisible. The platform is also used to study ion channels at high bandwidths. A wavelet transform-based denoising technique is then applied to these
recordings for further improvements in SNR.

The second work is a CMOS amplifier designed for intracellular recordings. By exploiting options that are only available in an IC context, the amplifier is able to achieve performance comparable to or better than that of commercial counterparts while consuming power that is orders of magnitude lower. In particular, the noise is among the lowest-reported among all integrated multi-clamp efforts. The amplifier is validated electrically and in vitro in neuronal cultures and slices.

These combined works have made several original contributions to single-molecule nanopore sensing and intracellular recordings:

- The highest-bandwidth nanopore recordings demonstrated to date (SNR > 6 at 10 MHz)
- The first reported recordings of DNA translocations through ultra-thin sub 2 nm-diameter silicon nitride nanopores
- The lowest reported voltage-clamp noise level to date of any integrated multi-clamp system
- An implementation of a novel resistance compensation scheme utilizing current preamplification
- The first reported application of wavelet denoising to high-bandwidth nanopore and ion channel recordings

These contributions have resulted in the following peer-reviewed publications:

• Shekar S., Chien C.-C., Hartel A., Ong P., Clarke O. B., Marks A., Drndić M., Shepard K. L. Wavelet denoising of high-bandwidth nanopore and ion channel signals. *Nano Letters, In press*


• Shekar S.*, Chien C.-C.*, Niedzwiecki D. J., Shepard K. L., Drndić M. Low-noise measurements of high-bandwidth DNA translocation in sub 2 nm a-Si/SiO$_x$ nanopores. *In preparation*


• Hartel A. J. W., Shekar S., Ong P., Schroeder I., Thiel G., Shepard K. L. High-bandwidth approaches in nanopore and ion channel recordings – A tutorial review. *In review*

### 7.2 Future work

**Nanopore amplifier noise reduction**

Commercial amplifiers used for nanopore recordings employ opamps with thermal voltage noise floors of $1\text{nV/}\sqrt{\text{Hz}}$ that is significantly lower than the $2.6\text{nV/}\sqrt{\text{Hz}}$ achieved in this work. It should be possible to design integrated opamps in order
to meet this goal. Combining this amplifier with an extremely low-capacitance measurement technique could yield further improvements in bandwidth [37]. A relevant consideration once this has been achieved will be to appropriately decrease the charge-transfer resistance of the on-chip Ag/AgCl electrode. This could be achieved using one of the solutions discussed in Section 4.2.2.

**Parallelized nanopore recordings**

The chip designed in Chapter 3 contains 25 individually accessible amplifiers on the same die. However, this work primarily explored the use of a single amplifier for both nanopore and ion channel recordings. While the ion channel setup is easily extended to multiple channels, the setup for nanopore measurements will need to be adapted in order to enable parallelized recordings.

A current limitation for the number of amplifiers that can be recorded from simultaneously is the requirement of a dedicated signal processing chain containing the boosting filters, the feedback attenuator, the anti-aliasing filters and ADCs per channel. If these could be integrated on-chip, the digital output could be multiplexed onto a few pads allowing for potentially larger scale multiplexing.

**Integrated circuitry to support nanopore fabrication**

One of the difficulties in achieving parallelized nanopore recordings is ensuring proper alignment and wetting of the membrane in relation to the amplifiers. A possible solution to this might be to first form wells above the amplifier electrodes, fill the wells with appropriate buffer solutions and then seal a pristine membrane above the wells. If the amplifier possesses the circuitry to fabricate a nanopore via CDB, this would ensure isolated operation as well as guarantee wetting of the nanopores.
Custom wavelet design

The wavelet denoising explored in this work used standard wavelets that were available as part of the PyWavelets module. However, it is possible to design wavelets tailored to the application of interest [210]. Further improvements to the denoising performance may be possible through the use of these custom-designed wavelets.

Wavelet denoising for detecting intra-event features

The denoising explored in this work primarily dealt with the detection of translocation events and enhancing the associated SNR. A natural question to ask would be how this could be modified in order to retain intra-event steps. Standard wavelet denoising operates purely on the amplitude of the detail coefficients. Since the intra-event steps will likely have smaller amplitudes than the event itself, the corresponding detail coefficients will also be smaller in amplitude and might get rejected by the thresholding scheme. But once the event edges have been determined, the threshold scheme could be modified to not be as aggressive for coefficients with indices between the start and end of the event.

On-chip ADC integration for multi-clamp

The patch-clamp amplifier chip designed in this work does not have the anti-aliasing filters and ADCs on the chip. Incorporating these on-chip will further reduce the footprint of the system and open up the possibility of performing moving-animal experiments with this amplifier. Currently, these experiments require the use of a compact headstage unit with long cables with sufficient slack to ensure freedom of movement for the animal. A fully-functional integrated amplifier will greatly improve the ease of performing such experiments.
**Improved robustness of $R_S$ compensation**

The $R_S$ compensation circuitry designed in this work relies on accurate compensation of the parasitic $C_p$ in order to remain stable. While this is also a requirement for the compensation implemented in commercial systems, it would be worth exploring the formal control theoretic requirements for such a loop to remain stable. Reducing the constraints on $C_p$ compensation can make the system more usable in biological experiments where great care must be taken in order to ensure that $C_p$ does not vary much once the pipette is in the bath.

**Multi-channel multi-clamp**

This avenue is similar to that discussed for the nanopore amplifier. One of the greatest strengths of CMOS integration is its ability to scale. Multiple multi-clamp systems have been demonstrated on the same die before but their performance did not compare favorably to commercial systems. Multiple channels on the same die were achieved by sharing components between the current- and voltage-clamp circuits. This approach can be adapted to the patch-clamp amplifier designed in this work and combined with the ADC integration described above to achieve true multi-channel multi-clamp operation.

**7.3 Final thoughts**

As Moore’s Law comes to an end, an often cited paradigm is the so-called “More than Moore”, where technology enabled by Moore’s Law is applied to problems beyond the realm of what might be considered traditional electrical engineering [211]. In that sense, I hope this thesis gave a taste of the fascinating challenges that lie at the intersection of different fields and the benefits afforded by modern semiconductor technology. The goal of technology has always been to enable man to achieve what
was not possible before, and I hope my work was one small example of that.
References


M. Zhang, C. Ngampeerapong, D. Redin, A. Ahmadian, I. Sychugov, and J. Lin-
nros, “Thermophoresis-Controlled Size-Dependent DNA Translocation through

D. Fologea, J. Uplinger, B. Thomas, D. S. McNabb, and J. Li, “Slowing DNA
translocation in a solid-state nanopore,” Nano Letters, vol. 5, no. 9, pp. 1734–
1737, 2005.

C. Plesa, N. Van Loo, and C. Dekker, “DNA nanopore translocation in glutama-

R. M. M. Smeets, U. F. Keyser, D. Krapf, M. Y. Wu, N. H. Dekker, and
C. Dekker, “Salt dependence of ion transport and DNA translocation through

DNA translocation through a nanopore in lithium chloride,” Nano Letters,

J. K. Rosenstein and K. L. Shepard, “Temporal resolution of nanopore sensor
recordings,” Proceedings of the Annual International Conference of the IEEE

“Generalized Noise Study of Solid-State Nanopores at Low Frequencies,” ACS
Sensors, p. acssensors.6b00826, 2017.

R. M. M. Smeets, U. F. Keyser, N. H. Dekker, and C. Dekker, “Noise in solid-
state nanopores,” Proceedings of the National Academy of Sciences, vol. 105,

R. M. M. Smeets, N. H. Dekker, and C. Dekker, “Low-frequency noise in solid-

R. Sarpeshkar, Ultra Low Power Bioelectronics: Fundamentals, Biomedical Ap-


R. M. M. Smeets, U. F. Keyser, M. Y. Wu, N. H. Dekker, and C. Dekker,
“Nanobubbles in solid-state nanopores,” Physical Review Letters, vol. 97, no. 8,

V. Tabard-Cossa, D. Trivedi, M. Wiggin, N. N. Jetha, and A. Marziali, “Noise
analysis and reduction in solid-state nanopores,” Nanotechnology, vol. 18,

noise processes in electrode electrolyte interfaces,” Journal of Applied Physics,


[168] L. Venkataramanan, R. Kuc, and F. Sigworth, “Identification of hidden Markov models for ion channel currents


