Signal Processing Techniques Applied to Biomedical Diagnostics

by

Ismail Degani

Submitted to the Department of Electrical Engineering and Computer Science on January 10, 2022, in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Electrical Engineering and Computer Science

Abstract

An effective way to combat cancer and infectious disease in resource-poor settings is to implement rapid and accurate diagnostic tests that can be administered at the point of care (POC). Developing such miniaturized, portable, and low-cost systems requires innovative approaches in both assay and device design. In this thesis, we construct a novel phase-sensitive “lock-in” amplifier (LIA) based on the Fast Walsh-Hadamard Transform (FWHT), and evaluate its ability to boost the signal-to-noise ratio of optical fluorescence signals. The LIA is designed to be resilient in challenging environments containing high/unpredictable ambient noise. We then develop two rapid diagnostic systems that pair this technology with isothermal CRISPR-Cas12a-based DNA/RNA amplification to detect clinically relevant targets with high specificity. Finally, we evaluate the clinical performance of our systems in detecting target genes for (1) SARS-CoV-2, the virus responsible for the COVID-19 pandemic, and (2) Human Papilloma Virus (HPV), the causal agent of cervical cancer.

Thesis Supervisor: Hakho Lee
Title: Associate Professor of Radiology
Harvard Medical School
Acknowledgments

I am deeply grateful to the many individuals and organizations who have made this doctoral thesis possible:

First, I’d like to give thanks to my research advisor Professor Hakho Lee. Dr. Lee is an exceptional mentor who has devoted significant time and effort to support my research in the Center for Systems Biology (CSB) at the Massachusetts General Hospital. It was Dr. Lee who recommended that I explore an academic career; he was the first to encourage me to apply for a PhD and consider a new path in the field of biomedical engineering. Prior to this, I had no formal research background beyond the undergraduate level, and in particular I had no peer-reviewed publications. Despite this, Dr. Lee gave me the opportunity to participate in research projects which appeared in high impact journals. He pushed me to rapidly ramp up my academic credentials and was instrumental in my acceptance to top PhD programs at Harvard and MIT; this would have been impossible just a year earlier. Dr. Lee saw value and potential in my non-academic (engineering) skill set, and paired me with excellent collaborators to complete meaningful projects with clinical impact. His lasting influence on my career cannot be overstated.

Many other faculty at the CSB have helped me tremendously in my graduate studies and I am grateful to all of them. Professor Hyungsoon Im provided me with the opportunity to work on innovative point-of-care diagnostic systems. He was my first research supervisor while I was a master’s student, and it was his compelling work that attracted me to the CSB. Laboratory medical directors Ralph Weissleder MD/PhD and Cesar Castro MD generously provided their clinical perspective across many of my projects. Their guidance significantly strengthened the foundations of this thesis and gave me a broad understanding of the many facets of translational research. In particular, I was able to accompany them on a trip to Botswana in 2019 where I got a first-hand look at how our diagnostic systems were being tested and validated in the developing world.

The systems presented in thesis were developed in collaboration with Dr. Chang
Yeol Lee, a postdoctoral researcher at the CSB. It was a privilege to work with Chang Yeol. He enabled me to apply my software, hardware, and signal-processing expertise towards addressing fundamental challenges in molecular diagnostics. I’m grateful for his professionalism and his friendship, along with countless other members at the CSB who helped me including John Raisbeck, Matthew Allen William, Hannah Peterson, Christian Landeros, Phil McFarland, Lucas Rohrer, Hyunho Kim, Ala Jo, Jouha Min, Young Kwan Cho, HoJeong Yu, Divya Pathania and many more.

I want to thank Professor Joel Voldman at the MIT Research Laboratory of Electronics (RLE) for his support and encouragement in my first year of PhD studies. At the Voldman lab, I received hands-on lab training and was introduced to many wonderful MIT PhD students (Kruthika Kikkeri, Wei Liao, Nicha Apichitsopa, Jaemyon Lee) and postdocs (Dohyun Lee, Sarvesh Varma, Bill Chow) with whom I hope to maintain lasting relationships.

I’d like to give a special thanks to my academic advisor Professor Luca Daniel for his support and friendship throughout my academic journey. In addition to supporting my academic progress, he fostered a strong sense of community and belonging for me at MIT. Our shared enthusiasm for classical music led to weekly violin-guitar rehearsals where I had the chance to decompress and reconnect with my musical side. He invited me to his choir concerts and informal gatherings; moreover he was always available to discuss research or personal difficulties I was having. Professor Daniel served on my thesis committee along with Professor Sangeeta Bhatia, who has also been a wonderful mentor and friend. I thank them both for their support during all phases of my research, from project ideation to the defense of this thesis.

I’m indebted to the MIT Legatum Foundation for providing generous financial support for my academic studies, and for my travel to Botswana. Through the Legatum foundation I met many like-minded entrepreneurs and gained valuable insights into the realities of solving healthcare challenges in the developing world. An additional thanks is owed to the MathWorks Corporation for awarding me a generous fellowship in my final year of doctoral studies. This fellowship allowed me to focus on completing this thesis and to coordinate a smooth transition to an industry
role following graduation. I am deeply grateful for the invaluable gifts from both organizations.

I’m thankful to Professor Richard Cohen, who gave me the opportunity to serve as a teaching assistant for his Fall and Spring courses in biomedical entrepreneurship. Beyond the financial support, these courses provided valuable exposure to the Boston ecosystem of life-science startups. They equipped me with the ability to critically evaluate biomedical business models, understand clinical trial data, and appreciate the complexities of commercializing therapeutics and medical devices. I thank him for this opportunity and wish him the best on a well-deserved retirement.

I’d like to thank all of the financial administrators, graduate officers, and other administrative staff in both the EECS Department and at the Massachusetts General Hospital for facilitating my atypical PhD appointments which spanned across both institutions. Professor Leslie Kolodziejski, Janet Fischer, Alicia Duarte, Kathy McCoy, Dan Vongphayboun, Irene Huang, Bill Tilden, Serena Sullivan, and Dianne Moschella all worked very hard to ensure I had the financial support I needed each semester, while keeping administrative burdens under control.

Last but not least, I’d like to thank my family and friends for their unwavering support and kindness through this challenging, rewarding chapter of my life.
3.3.5 In-vitro RNA preparation ................................................. 37
3.3.6 SYBR green-based qRT-PCR & melting curve analysis ................. 38
3.3.7 Clinical sample collection and analyses .................................. 39

3.4 Results and Discussion ...................................................... 39
3.4.1 CODA assay ................................................................. 39
3.4.2 CODA device ............................................................... 41
3.4.3 CODA assay optimization ............................................... 45
3.4.4 Assay characterization .................................................. 47
3.4.5 Clinical sample analysis ................................................ 48

3.5 Conclusions ................................................................. 49

4 Direct Fluorescence system for rapid HPV detection 53

4.1 Introduction ................................................................. 53
4.1.1 Human Papilloma Virus ................................................ 54
4.1.2 HPV Vaccines ............................................................ 54
4.1.3 Traditional Screening Tests ............................................. 54
4.1.4 HPV DNA Tests .......................................................... 55
4.1.5 VIA/VILI Inspection ..................................................... 56
4.1.6 Overview and Summary ............................................... 56

4.2 CRISPR Isothermal Amplification ........................................ 57

4.3 Discussion: Improving the CODA Device Design ....................... 59
4.3.1 Pink Noise and Modulation Frequency ................................ 59
4.3.2 Walsh-Hadamard Modulation ......................................... 62
4.3.3 Generating Walsh-Hadamard Waveforms ............................. 66
4.3.4 Optimization of Walsh-Hadamard Lock-in Amplifier .............. 67

4.4 Materials and Methods .................................................... 69
4.4.1 Materials ................................................................. 69
4.4.2 HPV assay ............................................................... 70
4.4.3 Clinical sample collection and analyses ............................. 71
4.4.4 HPV device construction ............................................. 71
## List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-1</td>
<td>Overview of fluorescence polarization (FP) technique.</td>
<td>23</td>
</tr>
<tr>
<td>2-2</td>
<td>Bacterial detection with PAD.</td>
<td>27</td>
</tr>
<tr>
<td>3-1</td>
<td>CODA signal processing flow</td>
<td>36</td>
</tr>
<tr>
<td>3-2</td>
<td>qRT-PCR for N1 RNA</td>
<td>38</td>
</tr>
<tr>
<td>3-3</td>
<td>CODA system Assay schematic</td>
<td>40</td>
</tr>
<tr>
<td>3-4</td>
<td>CODA device configuration and optics</td>
<td>41</td>
</tr>
<tr>
<td>3-5</td>
<td>CODA device photograph and UI software</td>
<td>41</td>
</tr>
<tr>
<td>3-6</td>
<td>CODA Real-time user interface</td>
<td>42</td>
</tr>
<tr>
<td>3-7</td>
<td>Optical design of CODA device</td>
<td>43</td>
</tr>
<tr>
<td>3-8</td>
<td>Validation of CODA device</td>
<td>44</td>
</tr>
<tr>
<td>3-9</td>
<td>Optimization of the reporter probe</td>
<td>45</td>
</tr>
<tr>
<td>3-10</td>
<td>CODA assay characterization</td>
<td>46</td>
</tr>
<tr>
<td>3-11</td>
<td>CODA Specificity assessment</td>
<td>47</td>
</tr>
<tr>
<td>3-12</td>
<td>Overall workflow of COVID-19 diagnosis with CODA.</td>
<td>48</td>
</tr>
<tr>
<td>3-13</td>
<td>CODA results of clinical samples</td>
<td>49</td>
</tr>
<tr>
<td>3-14</td>
<td>Waterfall plot of CODA results</td>
<td>49</td>
</tr>
<tr>
<td>3-15</td>
<td>COVID-19 clinical sample validation tests</td>
<td>50</td>
</tr>
<tr>
<td>4-1</td>
<td>Effect of Photobleaching on Fluorescein 5(6)-isothiocyanate</td>
<td>60</td>
</tr>
<tr>
<td>4-2</td>
<td>White and 1/f ADC Noise at 37.5kHz</td>
<td>60</td>
</tr>
<tr>
<td>4-3</td>
<td>White and 1/f ADC Noise at 407.9kHz</td>
<td>62</td>
</tr>
<tr>
<td>4-4</td>
<td>Overlaid Direct Comparison of Walsh and Fourier Basis Functions</td>
<td>64</td>
</tr>
<tr>
<td>4-5</td>
<td>FWHT Sample Inputs and Outputs</td>
<td>66</td>
</tr>
</tbody>
</table>
List of Tables

3.1 Oligonucleotide sequences used in the CODA study. . . . . . . . . . . 34
3.2 Clinical sample information. . . . . . . . . . . . . . . . . . . . . . . 48
3.3 Comparison of COVID-19 molecular tests. . . . . . . . . . . . . . . . 51

4.1 HPV Diagnostics Landscape . . . . . . . . . . . . . . . . . . . . . . . 57
4.2 Cost Breakdown of HPV Fluorescence Unit . . . . . . . . . . . . . . . 72
4.3 HPV Device Validation Data . . . . . . . . . . . . . . . . . . . . . . . 75
Chapter 1

Introduction

The persistent threats of cancer and infectious disease remain top global challenges [1 2 3 4]. At the time of writing, the COVID-19 pandemic has claimed an estimated 4.3 million lives since the initial outbreak [5], and the worldwide annual death rate from cancer is more than twice this number (9.9 million [6]). Unfortunately, the burden of these diseases is acutely borne by those in low-resource settings (rural and low-income areas, developing countries) due to their limited access to adequate healthcare [7 8]. Low and middle-income countries (LMICs) suffer far worse health outcomes than their more affluent counterparts. Given the magnitude of the clinical need, there is a strong incentive to investigate new technologies that may improve such outcomes. While technology alone is unlikely to solve such disparities, it can make a meaningful impact by unburdening resource-constrained clinicians and significantly lowering certain aspects of the cost of care for patients.

1.1 Clinical Need for Point-of-Care Diagnostics

An effective way to combat infectious diseases and curable cancers in LMIC’s is to implement rapid, accurate diagnostic tests at the point of care (POC) [9 10 11]. Such tests would eliminate diagnostic delays and uncertainties, enabling the timely initiation of the most appropriate and effective drugs. Unfortunately, current diagnostic methods often fail to meet these goals, either because the assays are too lengthy or
because they are not readily available in POC settings. Gold-standard tests such as blood culture, diagnostic pathology, and immunohistochemistry are largely infeasible for on-site use; they incur long processing times (up to several days), personnel costs, and the need for specialized equipment and species-specific protocols [12]. As an alternative, nucleic-acid (NA) assays have been increasingly adopted for clinical diagnostics [13, 14, 15, 16]. These tests can produce comprehensive information for pathogen detection and cancer profiling; target sequence libraries are also rapidly expanding, aided by advances in whole-genome sequencing [17]. Technical constraints, however, still limit the use of NA tests to centralized hospital laboratories and largely exclude them from POC settings: most tests are currently priced at >$100 in order to recoup the large capital expenditure on equipment ($10,000-$150,000). Furthermore, NA systems are bulky and house sophisticated optics which require routine calibration.

1.2 Specific Aims of Thesis

Our research group has been interested in developing new biosensing modalities whose functions range from sample preprocessing [18, 19], rapid NA amplification [20], to signal detection [21, 22]. Recent efforts have focused on improving signal acquisition and processing to facilitate POC diagnostics.

We have built compact, portable devices which are potentially scalable to a large number of disease targets. We have further optimized these systems to address specific clinical needs. The aim of this thesis is to build upon the fluorescence and fluorescence polarization (FP) systems [23] that have been previously disseminated. Specifically, we will (1) address new disease targets and (2) improve the signal-to-noise ratio and overall robustness of the system’s measurements.
1.3 Thesis Overview

In this thesis, I will present new signal processing techniques and diagnostic device designs which achieve high accuracy and performance across two clinically relevant detection targets: Human Papilloma Virus (HPV) and SARS-CoV-2 (COVID-19).

In chapter 2, I will review past work done in our research lab, as well as give an overview of the working principles of fluorescence polarization (FP) and lock-in amplification (LIA), which are central to the diagnostic systems that were developed in this thesis.

The system of chapter 3 is termed CODA (CRISPR Optical Detection of Anisotropy) and was designed to rapidly detect SARS-CoV-2, the pathogen responsible for the COVID-19 pandemic. The CODA assay combines isothermal DNA amplification, CRISPR/Cas12a activation, and FP signal generation. During the amplification reaction, Cas12a gRNAs recognize target nucleic acids (NA), which lead to the cleavage of FP probes by Cas12a/gRNA complexes. CODA is able to detect SARS-CoV-2 RNA within 20 minutes, and when applied to clinical samples (n=20), the method accurately identified 100% of COVID-19 positive patients.

The system of chapter 4 evolves the CODA system to rapidly detect high-risk HPV subtypes associated with aggressive cervical cancers. Like CODA, this new assay also combines isothermal DNA amplification with CRISPR/Cas12a activation, and it delivers results within 20 minutes. However, the HPV system uses an enhanced lock-in technique which facilitates superior signal-to-noise (SNR) ratio and fast acquisition time for direct fluorescence measurement. The HPV system is still under development, but preliminary data on clinical samples (n=9) already show great promise (89% accuracy).

Chapter 3 and 4 discuss all aspects of the CODA and HPV systems, including assay design, device construction, and the results of clinical tests. Both of these chapters are essentially self-contained publications and can be read independently of the rest of the thesis.

Finally, in chapter 5, I will conclude with future directions for this line of research.
Chapter 2

Review

The material presented in this chapter has been written in collaboration with bio-
chemist Chang Yeol Lee under the direction of supervising PI Hakho Lee. A related
review article is currently pending publication.

In this chapter we will review the optical and biological principles that we employ
in the design and development of our diagnostic systems.

2.1 Direct Fluorescence Principle

The principle of fluorescence is central to biological analysis [24]. Briefly, an or-
ganic molecule called a fluorophore absorbs a photon of a particular wavelength and
subsequently emits a longer-wavelength photon after some delay $\tau$, typically in the
nanosecond range. A photodetector captures this emitted light using a set of opti-
cal filters that block out the excitation wavelength and any other exogenous light.
At a minimum, one optical filter known as an emission filter is required to block
unwanted wavelengths of excitation light from reaching the detector. In practice,
extcitation filters are also necessary because monochromatic light sources (laser, LED,
etc) do not have a sufficiently narrow bandwidth and can pollute the spectrum of the
faint fluorescent signal. Finally, dichroic mirrors add an additional layer of unwanted
wavelength rejection, thus increasing the SNR of the system even further.
In fluorescent assays, probes are designed to activate fluorophores upon binding to desired targets. This is accomplished by modifying or cleaving molecules that are prohibiting or “quenching” the fluorescence.

### 2.2 Polarization Anisotropy Principle

The principle of fluorescence polarization (FP) builds on direct fluorescence to measure sample viscosity as well as changes in molecular mass. Unlike direct fluorescence, it is largely independent of a sample’s fluorophore concentration.

Fluorescence polarization was first described in 1926\(^2\) but didn’t gain widespread adoption in the biological community until the 1980’s. In fact, even as late as 1995, FP was referred to as a “new” tool for biological analysis by industry experts\(^3\).

In FP assays, fluorescent probes are designed to alter their molecular weight upon recognizing a target. This is accomplished either by cleaving off molecules (lowering their weight) or by binding to molecules (increasing their weight). This weight change modulates the probes’ tumbling rate (Brownian motion) which leads to changes in the overall assay’s fluorescence polarization (Fig. 2-1A). Large molecules tumble slowly in solution, and so their emitted fluorescence will have a polarization angle that is very similar to that of the excitation light source. In contrast, small molecules will rotate quickly; thereby their fluorescent emissions will effectively be depolarized.

Fluorescence polarization is computed from the relative intensities of two simultaneous measurements at parallel (\(I_\parallel\)) and perpendicular (\(I_\perp\)) polarization angles; this is known as a “ratiometric” measurement. Detecting FP is advantageous over measuring raw fluorescent intensities, assuming the assay is amenable to FP. The stability requirement for the excitation light source can be less stringent, as FP is (largely) independent of the intensity of the incident light. For example, dual optical detectors, each capturing an orthogonal polarization component, can compensate one another in the event of common noise. These detectors can also be monitored to reject single-channel noise that causes the signal variance to increase asymmetrically between channels (Fig. 2-1B). Finally, since only the intensity-ratios are relevant, the
Figure 2-1: Overview of fluorescence polarization (FP) technique. (A) Upon absorbing light (blue sinusoid), a fluorophore will emit longer-wavelength photons (green sinusoid) with identical polarization angles. If the fluorophore is small, it tumbles quickly during the fluorescence lifetime and the observed emission is depolarized (top). If the fluorophore binds a larger particle, the complex would rotate slowly and the emission light largely maintains the same polarization as the excitation light (top). (B) FP signal acquisition. Intensity readings from dual optical detectors are captured continuously, and an FP value is calculated in real-time. Common noise (dotted orange arrows) present on both channels is rejected due to the ratiometric nature of the measurement (i.e. the subtraction of parallel and perpendicular signals). In addition, the measured FP value can be monitored to ensure that its variance is within a threshold (dotted black arrows), which guards against temporal or momentary noise.

The sample’s total absolute fluorescence can be any arbitrary value as long as it is strong enough to rise above the system’s noise floor, yielding more assay flexibility.
The canonical fluorescence polarization equation is given in Eq. \[2.1\]

\[
r = \frac{I_{\parallel} - I_{\perp}}{I_{\parallel} + 2 \cdot I_{\perp}}
\] (2.1)

Rearranging the terms (Eq. \[2.2\]) highlights its ratiometric quality:

\[
r = \frac{I_{\parallel}/I_{\perp} - 1}{I_{\parallel}/I_{\perp} + 2}
\] (2.2)

This is also referred to as polarization anisotropy, though “polarization” and “anisotropy” technically have different definitions. In particular, polarization is measured in mP units and does not include a weighting factor of 2 on \(I_{\perp}\) (Eq. \[2.3\], see chapter 16 of Lotze \[27\] for more information).

\[
P = \frac{I_{\parallel} - I_{\perp}}{I_{\parallel} + I_{\perp}}
\] (2.3)

Small fluorescent molecules will have near-zero \(r\) as \(I_{\parallel}\) is close to \(I_{\perp}\). Conversely, large molecules or molecules bound to stationary targets will show high \(r\) magnitudes.

Diagnostic methods employing FP detection are economically scalable, can be packaged into compact devices, and allow for high-throughput measurements. The ability of FP to detect changes in molecular weight can be used to develop bioassays that detect i) target-induced cleavage of fluorescent probes \[28, 29, 30\] and ii) binding of fluorophore-conjugated ligands to a target of interest \[31, 32\].

Accordingly, there have been previous attempts to design FP-based point-of-care (POC) systems in conjunction with molecular detection probes. One such system is the Polarization Anisotropy Diagnostic (PAD) \[23\]. PAD was tailored to the clinical application of identifying bacterial pathogens responsible for healthcare-associated infections (HAI). The system served as a pre-cursor to the devices investigated in this thesis and will be reviewed in section \[2.5.1\].
2.3 Portable and Miniaturized Systems

Fluorescence and FP are typically measured by high-throughput systems such as commercial plate readers. Unbound by portability or cost constraints, these systems can ensure pristine measurement conditions by employing a range of design options, including i) hefty transformers to provide clean power signal, ii) a thick chassis for electromagnetic and optical shielding, and iii) separate circuit boards to minimize crosstalk between digital and analog electronics. Such generous options are often inapplicable to portable and cost-effective devices: they often use compact but noisy switching power supplies, need to keep shielding within reasonable limits, and place electrical components in close proximity to minimize physical footprint. Meeting these constraints in a POC FP system thus requires alternative strategies to maximize signal integrity. In this thesis, we will explore one such strategy known as “lock-in” amplification, which is reviewed in section 2.4.

2.4 Lock-in Amplification Principle

Lock-in amplification (LIA), also known as “phase-sensitive” detection, is a principle used across many sensing modalities to increase the signal-to-noise ratio (SNR) of a detection system [33]. LIA works by perturbing or “modulating” the signal it is measuring so that it can be separated from the ambient noise of the system. In the case of fluorescence or FP measurements, the excitation light can be modulated in an on-off flicker or a sinusoidal intensity profile. The observed fluorescence will then follow the same intensity pattern, and can be extracted from the noisy background by referencing (“locking-in”) the excitation light.

Implementing a lock-in system is straightforward: microcontrollers integrated with high resolution digital-to-analog converters (DACs) can generate a reference signal that modulates an excitation source, and low-cost amplifiers can demodulate faint signals from optical detectors. Furthermore, combining optical lock-in with the continuous monitoring of system noise (i.e. signal variance) improves the system fidelity.
A simple monitoring system might take several samples of a signal and reject the observed values if their variance falls outside of a predetermined range. This technique protects against momentary temporal noise fluctuations that can corrupt a measurement.

2.5 Previous Works

Here we discuss POC tests developed in our group to enable rapid diagnosis of infectious diseases and cervical cancer. We built on the foundation of these devices to create the CODA and HPV systems.

2.5.1 PAD System

The PAD system [23] employs FP to detect clinically relevant HAI bacteria (*Escherichia coli*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa*, *Staphylococcus aureus*) and does so with high specificity. It uses aptamer sequences that are bi-specific to a DNA polymerase and bacterial NAs; the aptamer induces the polymerase to cleave FP probes when target bacterial NAs are present. This cleaving reduces the probe’s molecular weight and therefore modulates FP signal. The PAD assay demonstrated detection sensitivity down to the single bacterium level and was able to determine both drug resistance and virulence status.

The PAD prototype had three separate components: a fluidic cartridge, a mini-thermocycler, and an optical reader for FP detection. The fluidic cartridge was developed for POC NA extraction. The device was designed to be produced via injection molding for cost-effective mass production. The detection system was designed for portable operation and ease of use. It had a modular structure consisting of a base unit for signal processing and plug-in optical cube modules. Each cube contained miniaturized optics to detect sample FP. The packaged PAD system had a small form factor (8 × 8 × 8 cm³) and weighed 400 g. To further improve system portability and user-friendliness, the PAD system was equipped with a smartphone via a Bluetooth connection. A smartphone App was also developed for system operation.
Figure 2-2: **Bacterial detection with PAD.** (a) PAD assay principle. Amplified target nucleic acids (NAs) have the detection key locked into the DNA polymerase, deactivating its activity. Thus, fluorescence anisotropy signal is high. (b) Photograph of a disposable RNA extraction cartridge made in plastic. The device has an RNA extraction chamber packed with glass beads (inset). (c) Photograph of a portable system for fluorescence anisotropy detection. Four separate optical cubes can be plugged into an electronic base station. (d) PAD measurement is controlled through a custom-designed application in a smartphone. *From Park, 2016* [23]. Reprinted with permission from AAAS.

A pilot clinical study to detect pathogens in human samples, where nine patient samples were aliquoted and screened by PAD and gold-standard methods (culture followed by qPCR) in a clinical microbiology laboratory. The PAD test results, obtained in 2 hours, fully matched with conventional culture-based tests that have 3-5 day turnaround times.

The PAD system demonstrated that FP detection platforms have the potential to become powerful tools for rapid diagnostics in the infectious disease space. Such detection systems do not require highly skilled personnel or labor intensive analyses, and the assays are quick and cost effective. PAD demonstrated that it was possible to construct systems that were well-aligned with a POC workflow to aid physicians to initiate prompt and informed patient treatment. The PAD device designs served as a reference to be improved in the CODA and HPV systems discussed in this thesis.
2.5.2 HPV Diffraction System AIM

Content presented in this section has been adapted from Theranostics 2019; 9(26): 8438–8447, where Ismail Degani is a contributing author. This is an open access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/). See http://ivyspring.com/terms for full terms and conditions.

The clinical need for rapid, low-cost HPV-diagnosis in low-resource settings is discussed in more detail in Chapter 4. The AIM system [34] attempted to address this need using microholographic imaging of dimer-bead formation. HPV DNA was extracted and amplified from cervical brushings using bead-loaded disposable filters. DNA isolates were then mixed with 6-µm polystyrene (PS) and 5-µm silica beads, each coated with DNA probes complementary to the 3’ and 5’ ends of the target HPV DNA. In the presence of target DNA, the two types of beads were linked together to form detectable PS-silica beads dimers. Diffraction patterns of PS, silica, and PS-silica bead dimers were captured by a miniaturized micro-holography device and quickly analyzed by pre-trained deep-learning algorithms.

The AIM-HPV system was tested on 28 samples from patients having abnormal pap smear results. Across both brushings and biopsies, the AIM system showed full concordance when compared to a commercial DNA test (Cobas, Roche Diagnostics).

The AIM HPV device demonstrated that high-accuracy POC devices were feasible for HPV screening in resource-poor settings. The device’s per-unit manufacturing cost was estimated to be < $150 [35], and the cost of reagents and beads would scale with increased volume. The HPV system described in Chapter 4 targeted the same subtypes (HPV 16, 18) as the AIM system, and accessed the same clinical laboratory within the Massachusetts General Hospital to retrieve samples for evaluation.

2.6 Summary

Fluorescence and FP are powerful analytical tools that have been broadly used in chemistry [36], drug development [37, 38], and infectious disease diagnostics [23].
Combining these robust sensing modalities with highly specific assays and optical lock-in amplification yield powerful detection capabilities in systems that are relatively simple and cheap to design. New assay methods have the potential to lower reagent costs ($1-$2 per assay) and manual steps (washing, centrifugation, etc) while still achieving detection accuracy comparable to gold-standard methods.

These advantages make the systems discussed in this thesis well-suited to a POC workflow, facilitating rapid diagnosis in hospitals and ambulatory clinics. The pilot studies attempted in this thesis prove the concept by processing clinical samples and rigorously comparing them to gold standard methodologies. We have designed COVID-19 and HPV systems that leverage the techniques and principles described in this Chapter.

Further improvements will bring these technologies closer to clinical translation. Improved sample pre-processing (e.g., NA extraction), increased throughput, and multiplexed detection modes could allow the detection of a plurality of targets in a convenient “one-pot” format. Additionally, integrating and mechanizing all steps can further reduce the likelihood of errors from sample contamination and user intervention. A highly specific diagnosis of an infectious disease or cancer, based on a comprehensive panel of probes, will allow clinicians to effectively triage patients in order to respond optimally in resource-constrained environments.
Chapter 3

CODA: Fluorescence polarization system for COVID-19 diagnosis

The work presented in this chapter has been reproduced from the published manuscript entitled "Fluorescence polarization system for rapid COVID-19 diagnosis" by co-first authors Chang Yeol Lee and Ismail Degani, under the direction of supervising authors Hakho Lee and Jinwoo Cheon [39]. Other contributing authors include Jiyong Cheong, Jae-Hyung Lee, and Hyun-Jung Choi (2021. Biosensors and Bioelectronics, 113049). Reproduction permission is granted by Elsevier B.V.

3.1 Introduction

Prompt diagnosis, patient isolation, and contact tracing are key measures to contain the coronavirus disease 2019 (COVID-19). Molecular tests are the current gold standard for COVID-19 detection, but are carried out at central laboratories, delaying treatment and control decisions. Here we describe a portable assay system for rapid, onsite COVID-19 diagnosis. Termed CODA (CRISPR Optical Detection of Anisotropy), the method combined isothermal nucleic acid amplification, activation of CRISPR/Cas12a, and signal generation in a single assay, eliminating extra manual steps. Importantly, signal detection was based on the ratiometric measurement of fluorescent anisotropy, which allowed CODA to achieve a high signal-to-noise ratio.
For point-of-care operation, we built a compact, standalone CODA device integrating optoelectronics, an embedded heater, and a microcontroller for data processing. The developed system completed SARS-CoV-2 RNA detection within 20 min of sample loading; the limit of detection reached 3 copy/μL. When applied to clinical samples (10 confirmed COVID-19 patients; 10 controls), the rapid CODA test accurately classified COVID-19 status, in concordance with gold-standard clinical diagnostics.

### 3.2 Background

The 2019 coronavirus (COVID-19) pandemic underscores the need for rapid, point-of-care (POC) diagnostic tests [10, 11]. At about one year post-initial outbreak, the United States alone has had over 24 million COVID-19 cases with more than 400,000 deaths; other countries, who have managed to prevent the initial spread, now experience a “second wave” of COVID-19 outbreaks [40, 41]. While COVID-19 vaccines are slowly deployed over the world, the primary response to COVID-19 is still containment, i.e., widespread implementation of diagnostic testing, isolation and contact tracing [10, 12, 11]. Among many different assays reported for COVID-19 detection, nucleic acid amplification tests (NAATs), in particular reverse transcription polymerase chain reaction (RT-PCR), are the *de facto* diagnostic standard. They enable highly accurate identification of SARS-CoV-2, the causative pathogen [43, 44, 11]. Most PCR tests, however, are carried out in centralized laboratories, limited by lengthy assay time (1-2 h) and requiring bulky instrumentation [45]. This in turn causes logistic overheads (e.g., sample transfer, protection from degradation) and long turnaround times to obtain results (2-3 days).

Isothermal nucleic acid (NA) amplification is a promising technology for onsite NAATs. Using specialized DNA polymerases with the capacity of strand displacement, these methods amplify NA at a fixed temperature, thereby simplifying device design [46, 47]. Combining isothermal amplifications with clustered regularly interspaced short palindromic repeats (CRISPR) has been shown to further improve assay sensitivity and specificity [48]. Upon recognition of its target strand, a CRISPR-
associated (Cas) protein and guide RNA complex can indiscriminately cleave single-stranded NAs [19]. This property has been exploited to amplify analytical signal i) through the collateral cleavage of non-target NA probes that have a fluorescent dye and quencher pair [50, 51, 52] or ii) through the release of nanoprobe that are initially anchored by non-target NA probes [53]. CRISPR systems have demonstrated a promising potential for COVID-19 diagnostics [54, 55]. Technical challenges, however, limit the practical point-of-care (POC) use of these assays: i) most CRISPR tests proceed with sequential NA amplification and detection, requiring separate preparation and introduction of CRISPR reagents into samples; and ii) signal readout, typically performed with a dipstick-type lateral flow device, also incurs extra manual steps and produces subjective, qualitative results.

Here, we report on a rapid, quantitative, and streamlined COVID-19 test in a compact system. Termed CODA (CRISPR Optical Detection of Anisotropy), the system seamlessly combined a one-pot CRISPR test with a robust detection modality, fluorescence anisotropy (FA). Specifically, we adopted a one-pot assay scheme wherein isothermal NA-target amplification and CRISPR-based target recognition take place simultaneously. The activated CRISPR/Cas then cleaves fluorescent DNA reporters, changing the FA readout. The CODA approach offered practical advantages: i) the entire reaction, including reverse transcription, NA amplification, and signal detection, were conveniently carried out in a single tube and at constant temperature (42°C); ii) the FA measurement, unlike conventional fluorescent intensity detection, was ratiometric and robust against common noises (e.g., intensity fluctuations), which led to reliable analytical signal even at low NA concentrations [56]; and iii) FA reporters, which required only a fluorescent dye conjugation, were simpler and cheaper than common CRISPR/Cas probes. These features enabled us to advance a portable CODA system integrating low-noise optical detection, precise temperature control, and on-board data processing and display. With the integrated CODA system, we detected SARS-CoV-2 RNA within 20 min in a ‘sample-to-result’ manner and achieved a sensitivity down to 3 copies/μL. We further applied the system in a pilot clinical test, assessing 20 clinical samples; the results showed an excellent accordance
3.3 Materials and Methods

3.3.1 Materials

<table>
<thead>
<tr>
<th>Usage</th>
<th>Sequence (5' to 3')</th>
</tr>
</thead>
<tbody>
<tr>
<td>N1-Forward RPA primer</td>
<td>AGG CAG CAG TAG GGG AAC TTC TCC TGC TAG AAT</td>
</tr>
<tr>
<td>N1-Reverse RPA primer</td>
<td>TTG GCC TTT ACC AGA CAT TTT GCT CTC AAG CTG</td>
</tr>
<tr>
<td>N1-Forward Cas12a gRNA</td>
<td>UAA UUU CUA CUA AGU GUA GAA CAU CAC CGC CAU UGC CAG CC</td>
</tr>
<tr>
<td>N1-Reverse Cas12a gRNA</td>
<td>UAA UUU CUA CUA AGU GUA GAA UUG CUG CUG CUU GAC AGA UU</td>
</tr>
<tr>
<td>N2-Forward RPA primer</td>
<td>GGA ACT GAT TAC AAA CAT TGG CCG CAA ATT GCA</td>
</tr>
<tr>
<td>N2-Reverse RPA primer</td>
<td>TGC TTA TCC AGC AAA ATG ACT TGA TCT TGG AAA</td>
</tr>
<tr>
<td>N2-Forward Cas12a gRNA</td>
<td>UAA UUU CUA CUA AGU GUA GAA AAC GCU GAA GCG CUG GGG GC</td>
</tr>
<tr>
<td>N2-Reverse Cas12a gRNA</td>
<td>UAA UUU CUA CUA AGU GUA GAA GGU UGA CCC ACA CAG GUG CC</td>
</tr>
<tr>
<td>RPP30-Forward RPA primer</td>
<td>UAA UUU CUA CUA AGU GUA GAA UUG GGG AGA CAG CCG CCC AC</td>
</tr>
<tr>
<td>RPP30-Reverse RPA primer</td>
<td>UAA UUU CUA CUA AGU GUA GAA UUG GGG AGA CAG CCG CCC AC</td>
</tr>
<tr>
<td>Reporter probe (5 mer)</td>
<td>FAM-TTA TT</td>
</tr>
<tr>
<td>Reporter probe (7 mer)</td>
<td>FAM-TAT TAT T</td>
</tr>
<tr>
<td>Reporter probe (9 mer)</td>
<td>FAM-TAT TAT T</td>
</tr>
<tr>
<td>Reporter probe (11 mer)</td>
<td>FAM-ATT ATT ATT AT</td>
</tr>
<tr>
<td>Reporter probe (13 mer)</td>
<td>FAM-ATT ATT ATT A</td>
</tr>
<tr>
<td>SARS-CoV-2 N1 for IVT (b)</td>
<td>TAA TAC GAC TCA CTA TAG GCC TCA AGG AAC AAC ATT GCC A</td>
</tr>
<tr>
<td>Forward primer</td>
<td>GCT TTA GTG GCA GTA GTG T</td>
</tr>
<tr>
<td>Reverse primer</td>
<td>CAG CAG TAG GGG AAC TTC T</td>
</tr>
<tr>
<td>SARS-CoV-2 N1 for qRT-PCR</td>
<td>AGA CAT TTT GCT CTC AAG CTG</td>
</tr>
</tbody>
</table>

Table 3.1: Oligonucleotide sequences used in this study.
(a) The colors of oligonucleotide sequences correspond to those of the domains depicted in Fig. 3-3.
(b) Underlined sequence is a T7 promoter sequence.

All oligonucleotides used in this study were synthesized by Bioneer® (Daejeon, Korea). The sequences of the oligonucleotides are listed in Table 3.1. Plasmid controls of SARS-CoV-2, human RPP30, MERS-CoV, and SARS-CoV and gBlocks® gene fragments of HCoV-NL63, HCoV-229E, HCoV-OC43, and HCoV-HKU1 were purchased from Integrated DNA Technologies, Inc. (Coralville, IA, USA). TwistAmp® Basic kit and RevertAid reverse transcriptase (RTase) were purchased from TwistDx™ and ThermoFisher Scientific (Waltham, MA, USA), respectively. EnGen® Lba Cas12a
(Cas12a), RNase inhibitor (Murine) (RI), DNase I, HiScribe™ T7 High Yield T7 RNA synthesis kit, and Monarch® RNA cleanup kit were purchased from New England BioLabs (Beverly, MA, USA). KAPA HiFi HotStart ReadyMix PCR kit and TOPreal™ One-step RT qPCR kit were purchased from Roche (Basel, Switzerland) and Enzymomics (Daejeon, Korea), respectively. DEPC-DW, purchased from Bioneer®, was used in all experiments. All other chemicals were of analytical grade and used without further purification.

### 3.3.2 Device construction

The CODA device was modeled in computer-aided design software (Solidworks, 2019) and fabricated via computer numerical control machining (6061 aluminum) and 3-dimensional photopolymer resin printing (Formlabs). The main body housed a light source subassembly and two identical FA detectors. i) The light source consisted of an LED (Thorlabs M470D2), a linear polarizer (Thorlabs LPVISE100-A), and a convex focusing lens (Thorlabs LB1092-A). Components were concentrically aligned inside a standard tube mount (Thorlabs SM1A6, CP4S, SM1L03). The LED was attached to a metal heatsink through a thermal compound (Arctic Silver 5) and driven by a constant current amplifier (Meanwell LDD-1000L). ii) Each FA detector subassembly consisted of a convex lens (Thorlabs LB1092-A), a linear polarizer (Thorlabs LPVISE050-A), a band-pass filter (Thorlabs FLH05532), and a photodiode (Hamamatsu S1223). Optical components were assembled in slotted lens tubes (Thorlabs SM05L20C); the two detectors were oriented to have orthogonal polarizations. iii) FA signals were amplified using a configuration of precision amplifiers (LF356, AD54 CE9JHZ) and a lock-in amplifier (AD63 CE0JNZ). A 1-kHz carrier wave was generated by the integrated digital-analog converter (DAC) of an ARM Cortex M4 microcontroller (PJRC Teensy 3.2). iv) For sample heating, a metal holder was machined in aluminum. An integrated 10 W ceramic heater was driven by a power transistor (TIP120) attached to an 18 V supply. Temperature was monitored by a 10 kΩ thermistor (Thorlabs HT10KR) and controlled to within 0.2°C by a proportional-integral-derivative (PID) feedback loop running on the microcontroller unit (MCU). V) For the graphical user
interface (GUI), a real-time dashboard was programmed using the Qt GUI framework (Fig. S1), which communicated with the MCU over the USB serial port. The MCU firmware was written in C++ (Arduino).

### 3.3.3 CODA signal processing

**Figure 3-1:** CODA signal processing flow. In step 1, fluorescence intensity data from parallel (red) and perpendicular (blue) channels are captured as 14-bit integers every 0.1 sec. These signals may contain common noise (orange arrows). The channels are then subtracted, normalized, and scaled by factor $F$, leading to the continuous anisotropy plot of step 2. Note that the anisotropy measurement is free of the common noise fluctuations. A rolling standard deviation of the most recent 40 anisotropy samplings is monitored during data collection to ensure that the values are stable. Finally, in step 3, a measurement is taken as the average of 40 samples, provided that their standard deviation is less than 0.5%

The oscillating fluorescence signals were captured by a photodiode; amplified 10-
fold by an analog negative-feedback amplifier locked-in to the 1 kHz source signal; low-pass-filtered and amplified again by 30-fold; and converted to a digital signal by a pair of 12-bit analog-to-digital converters inside the MCU. The integrated real-time monitoring GUI displayed raw fluorescence readings from both parallel ($I_{||}$) and perpendicular ($I_{\perp}$) channels in real-time (every 0.1 s). The anisotropy ($r$) was computed according to equation 3.1, where $F$ was a scaling factor ($F = 4259$) to match CODA values with those measured by a plate reader (Sapphire 2, TECAN).

$$r = F \cdot \frac{I_{||} - I_{\perp}}{(I_{||} + 2 \cdot I_{\perp})}$$  \hspace{1cm} (3.1)

Noise fluctuations arising from both raw and computed data streams were precisely measured as the rolling standard deviation of the most recent 40 samples, to confirm that the system’s measurement was reliable (see Fig. 3-1 for the flow of signal processing).

### 3.3.4 CODA assay

The CODA master mix was prepared by combining 240 nM gene-specific RPA primers, 100 nM reporter probe, 160 nM gene-specific Cas12a gRNAs, 2 U/$\mu$L RevertAid RTase, 0.8 U/$\mu$L RI, 640 nM Cas12a, 0.2 $\times$ NEBuffer 2.1 (10 mM Tris-HCl, 50 mM NaCl, 10 mM MgCl2, 100 $\mu$g/mL BSA, pH 7.9 at 1 $\times$ concentration), and reconstituted RPA mix. Upon addition of viral RNA and 14 mM MgOAc, the CODA mix was incubated at 42°C for 20 min. FA was measured during the reaction. In the specificity test, plasmid controls and gene fragments, which are double-stranded DNAs, were used as a target and RTase was excluded.

### 3.3.5 In-vitro RNA preparation

PCR mix for in-vitro transcription (IVT) (20 $\mu$L) was prepared to contain 500 nM IVT primers, 1 $\times$ KAPA HiFi HotStart ReadyMix, and 104 copy/$\mu$L SARS-CoV-2 plasmid control. Thermocycling was performed with following steps: 95°C, 2min; 95°C, 20 s, 60°C, 39s, 35 cycles; 60°C, 5 min. After agarose gel extraction, 1 $\mu$g of IVT
PCR product was applied to HiScribe™ T7 High Yield T7 RNA synthesis mix with recommended composition of T7 RNA polymerase mix, rNTPs, and reaction buffer. IVT was then carried out at 37°C for 3 h, followed by DNase I treatment for 30 min to prevent DNA contamination. Finally, the IVT RNA product was purified with Monarch® RNA cleanup kit per manufacturer’s protocol and its concentration and purity were determined using NanoDrop™ 2000c (ThermoFisher Scientific, Waltham, MA, USA).

3.3.6 SYBR green-based qRT-PCR & melting curve analysis

For comparative analyses, qRT-PCR was performed. Viral RNA of SARS-CoV-2 was added to qRT-PCR mix composed of 250 nM primers and 1 × TOPreal™ One-step RT-qPCR reaction mix. qRT-PCR was then conducted on the ViiA 7 Real-Time PCR system (Life Technologies) with the following steps: 50°C, 30 min for RT; 95°C, 10 min; 95°C, 5 s, 60°C, 30 s, fluorescence measurement, 45 cycles. After denaturing and ramping down the PCR product to 40°C, a melting curve analysis was carried out to validate the qRT-PCR result. As the temperature is elevated (0.05°C/s), fluorescence is accordingly measured. Threshold cycle (Ct) and melting temperature were automatically determined by the system software (Fig. 3-2).

Figure 3-2: qRT-PCR for N1 RNA. (a) Cycle-dependent fluorescent signals of qRT-PCR in the presence of serially diluted N1 RNA. (b) Melting curve analysis was carried out on RT-PCR product of N1 RNA to validate the results of qRT-PCR. The observed melting temperature ($T_m = 81.1 \, ^\circ C$) matches with a predicted value (80.6 °C).
3.3.7 Clinical sample collection and analyses

The study was approved by the Institutional Review Board of Chonnam National University Hospital. Clinical samples of nasopharyngeal and oropharyngeal swabs and sputum were collected to the universal transport medium (UTM) (Asan Pharmaceutical, Seoul, Korea) and transported to Chonnam National University Hospital Laboratory which Korea Centers for Disease Control and Prevention (KCDC) approved as biosafety level-2 (BL-2) facility in accordance with institutional biosafety requirements. Total RNA was extracted from each sample, using AdvanSure™ E3 System (LG chem, Seoul, Korea) per manufacturer’s protocol. For the clinical diagnoses, extracted RNA was amplified (40 cycles) using a commercial qRT-PCR kit (PowerChek™ 2019-nCoV Real-time PCR Kit; KogeneBiotech, Seoul, Korea) and a detection system (CFX96™ Real-time PCR detection system; Bio-Rad, Hercules, CA, USA). A positive qRT-PCR result was defined as a $C_t \leq 35$.

3.4 Results and Discussion

3.4.1 CODA assay

The CODA assay (Fig. 3-3) combines isothermal NA amplification and CRISPR/Cas12a detection. In the presence of the target NA sequence, reverse transcription and recombinase polymerase amplification (RT-RPA) takes place with gene-specific RPA primers. During this process, single strand DNAs downstream to the primer binding sites (purple and blue) are exposed by the strand displacement activity of DNA polymerase. When gRNAs recognize the specific sites (green and orange) at both ends of displaced single strands, Cas12a/gRNA complexes get activated to cleave DNA [49]. Note that most cleavage actions would be limited to fluorescent reporter DNA (trans-cleavage). The protospacer adjacent motif (PAM) sequence is absent downstream to the recognition sites, which makes it unlikely that activated Cas12a invades and cuts target DNA [57]. Furthermore, RPA synthesizes complementary sequences on displaced single strands, effectively keeping Cas12a (RuvC nuclease do-
Figure 3-3: **CODA system Assay schematic.** When target viral RNA is present, RT-RPA and CRISPR/Cas detection take place together. DNA polymerase recognizes target sequences and displaces double-stranded DNA. Cas12a/gRNA complexes then bind to specific sites (green and orange) in the exposed single strand, get activated, and start to cleave nearby reporter probes. This cleaving process is amplified, as RT-RPA reaction proceeds. As a result, the fluorescence anisotropy (r) of the sample decreases (right).

As exponential NA amplification proceeds, the trans-cleaving activity is reinforced, degrading a large number of fluorescent reporters. Accordingly, the overall molecular weight of fluorescent reporters decreases, which yields a lower FA value (r) due to increased Brownian motion (Fig. 3-3 right). The isothermal, dual amplification of target NA and FA signal allows for rapid, one-step, and one-pot assay (20 min). Minimal intervention as well as a ratiometric FA readout make CODA robust to external interference. To enable onsite CODA assay, we have constructed a portable device (Fig. 3-4a) for CODA reaction and FA detection. Specifically, the device maintained a constant reaction temperature (42°C) through a feedback control, measured FA values using compact optics (Fig. 3-4b), and processed raw data to obtain r values. The prototype device (Fig. 3-5a) had a form factor of 13 × 4.5 × 5 cm³. A sample container (a PCR tube with diameter 6.2 mm) was to be inserted into a form-fitting metal sleeve for heating. The device also communicated with a computer running real-time analysis software (Fig. 3-5b, Fig. 3-6) that displayed and stored data.
3.4.2 CODA device

Fig. 3-7 shows the schematic of optoelectronics inside the CODA device. We adopted an optical lock-in method for low-noise measurements. An LED generated modulated blue light at varying intensity (f = 1 kHz) which passed through a linear polarizer onto the sample. The sample emitted green light (525 nm) oscillating in intensity at the same modulation frequency (1 kHz) with polarization components that depend on the molecular weight of FA probes. Two detectors captured both parallel...
Figure 3-6: CODA Real-time user interface. At the far left is a configuration panel where experimental parameters can be tuned, including (i) the number of measurements to collect, (ii) how long each measurement should take, and (iii) how many samplings to average in order to obtain a reliable measurement. The reference fluorescence anisotropy (FA) value \((r_0)\) can also be set here. The main dashboard is divided into four quadrants, where the user can monitor the raw fluorescent intensity for both parallel and perpendicular FA detectors (upper left), as well as a continuously computed FA value (lower left). Temperature and PID control signals are continuously monitored (upper right) and multiple patient samples can be analyzed sequentially in a batch and plotted (lower right). The dashboard is shown running on Windows 7-64 bit, though it is fully cross-platform and can run on a variety of operating systems that support the Qt GUI toolkit (MacOS/Linux)

\(I_\parallel\) and perpendicular \(I_\perp\) polarization components. Each detector consisted of a photodiode and a convex lens. An additional polarizer and a bandpass filter were used to reject the source (exciter signal) and unwanted polarization angles. Once the signal had been captured, optical lock-in amplifiers were utilized to reject common noise that may arise due to i) ambient light leaking into the device’s optical path, ii) external electrical noise due to currents flowing through ground impedances, iii)
electromagnetic interference from nearby electronic devices, and iv) other internal electrical noise.

Figure 3-7: **Optical design of CODA device.** Optical and electrical schematic. A light emitting diode (LED) illuminates a sample with linearly polarized light oscillating at 1 kHz. Fluorescence is captured by two photodetectors, each consisting of a photodiode, a 525-nm bandpass filter and a linear polarizer. The signal is processed by a sequence of integrated filtering/amplification steps: \(10 \times\) bandpass, lock-in, and \(30 \times\) lowpass. The cleaned signals are finally captured by a microcontroller. Temperature is controlled through a feedback control. AMP, amplifier; DAC, digital-to-analog converter; PID, proportional-integral-derivative.

We compared CODA’s optical performance against that of a benchtop plate reader (Sapphire 2, TECAN). We prepared standard samples of varying fluorescent anisotropy; different amounts of glycerol were added into a fluorescein solution, which changed the solution’s viscosity. For CODA, parallel \(I_{\parallel}\) and perpendicular \(I_{\perp}\)
Figure 3-8: **Validation of CODA device.** (a) The CODA system was benchmarked against a conventional plate reader. Samples were prepared in triplicate through the serial dilution of glycerol in an aqueous buffer, varying the viscosity. All samples contained the same amount of fluorescein (240 nM). An excellent correlation was observed between these two systems. (b) Sample heating curve. The system reached the target temperature (42°C) within 90 s. This temperature was maintained within ±0.2°C variations.

Polarization intensities were measured every 0.1 s, and 40 readings from each channel were averaged to reduce sampling noise. The fluorescence anisotropy (r) was then computed by the onboard MCU. The measured r values showed a good match ($R^2 = 0.97$) between the CODA device and the plate reader (Fig. 3-8a), which confirmed CODA’s accuracy in optical detection. Because the computed anisotropy (r) is essentially a normalized difference between polarization channels, we indeed observed that the anisotropy measurements were far more stable than their raw intensity components. Combining noise-monitoring with tightly controlled exposure times reduced the variability between measurements and increased the overall repeatability of experiments.

We next monitored CODA’s temperature stability. The goal was to keep the temperature of the metal sleeve holding a sample tube at 42°C ($T_{set}$), the optimal condition for the CODA assay. The MCU checked the temperature $T_{actual}$ every 0.1 s by measuring the resistance across an embedded thermistor and then used the error term ($T_{set} - T_{actual}$) for the feedback control. The system rapidly reached the target
temperature, within 1.4 min of heater activation, and maintained it with variations \( \leq 0.2 \degree C \) (Fig. 3-8b).

### 3.4.3 CODA assay optimization

We optimized the CODA assay condition for COVID-19 detection. We designed detection probes to target two regions in nucleocapsid gene (N1 and N2 genes) of SARS-CoV-2 RNA (Table 3.1). A third probe targeted human ribonuclease P subunit 30 (RPP30) gene as a sample quality check. For each target gene, Cas12a gRNAs were constructed to recognize sequences near the RPA primer binding sites. This would enable Cas12a gRNAs to bind to exposed single strand DNAs upon strand displacement, activating Cas12a’s cleavage activity.

![Figure 3-9: Optimization of the reporter probe.](image)

(a) The length of the reporter probe was varied and the resulting anisotropy signal from CRISPR cleavage was measured. The optimal length was found to be 9 bases. (b) The concentration of the 9-base probe was optimized to achieve the maximal anisotropy signal. The optimal concentration was found to be 0.1 \( \mu M \). The target concentration (SARS-CoV-2 plasmid control) was fixed at 103 copy/\( \mu L \). All experiments were performed in triplicate and the data are displayed as means ± s.e.m.

Using SARS-CoV-2 N1 RNA as a model target, we first determined the optimal FA reporter condition (Fig. 3-9). Probe lengths and concentrations were varied and the FA signal after Cas12a cleavage was measured. The maximal FA signal was observed with a nine-base reporter at 0.1 \( \mu M \). We next validated the assay feasibility by monitoring the FA signal under various reaction conditions (Fig. 3-10a). The
r values were high (i.e., high anisotropy) without effective cleavage of fluorescent reporters. The highest r was observed when RT-RPA took place but not Cas12a activation. This result was presumable due to the increase of viscosity from the RPA reaction \[58, 59\]. With all CODA components present, however, FA decreased significantly. We defined the analytical metric as \(\Delta r = r_0 - r\), where \(r_0\) (control) was measured from the CODA assay in the absence of the target gene.

Figure 3-10: CODA assay characterization. (a) Assay mechanism validation by varying the reaction condition. FA remained high when any single assay component was missing. The value significantly decreased when all key components were present. The target N1 RNA concentration was 106 copy/\(\mu\)L. RTase, reverse transcriptase. (b) Temporal evolution of FA signal among serially diluted N1 RNA samples. A sample without target RNA was used as a control. The difference \(\Delta r = r_0 - r\) was defined, where \(r_0\) was FA from the control sample. The net signal (\(\Delta r\)) increased and became commensurate with RNA concentration in 20 min. (c) Serially diluted N1 RNA samples were analyzed by CODA (20 min) and conventional RT-qPCR (80 min). CODA achieved wider dynamic range and higher sensitivity than RT-qPCR. The limit of detection was 3 copy/\(\mu\)L for CODA. \(\Delta C_t = C_{t0} - C_t\), where \(C_{t0}\) and \(C_t\) were threshold cycle values for control and RNA samples, respectively.

We assessed the CODA assay kinetics. Samples with varying N1 RNA concentrations were prepared, and their FA signals were monitored as the CODA reaction proceeded (Fig. 3-10b). Within 5 min of the reaction start, we observed that signals had already risen above the background. We set the detection time to 20 min post-reaction; at this point, the signal level was commensurate with the RNA concentration, allowing for quantitative measurements.
3.4.4 Assay characterization

Applying these assay settings, we carried out a titration experiment with serially diluted N1 RNA samples (Fig. 3-10c). Based on $3\sigma$/slope, where $\sigma$ is the standard deviation at the lowest concentration in the linear range, the limit of detection was determined to be 3 copy/µL. Notably, the CODA assay yielded a robust signal even at low RNA concentrations, achieving wider dynamic range than conventional RT-PCR. This merit can be attributed to a high signal-to-noise ratio inherent in the FA measurement: the fluorescence light entering photodetectors remained strong in the FA measurement, largely set by the initial reporter probe concentration. In RT-PCR, however, the fluorescence intensity at low RNA concentrations would be weak and competing with the detectors’ intrinsic noise.

Figure 3-11: CODA Specificity assessment. SARS-CoV-2 N1, N2, and human RPP30 probes were tested for cross reactivity with other compounding genes. Each probe generated high analytical signal only when its corresponding target was present. Gene concentration was set to 103 copy/µL in all samples. Data are displayed as means ± s. e.m. From triplicate measurements.

We further tested the specificity of prepared probes (N1, N2, RPP30) against most common human coronaviruses (HCoV-NL63, HCoV-229E, HCoV-OC43, HCoV-HKU1) and zoonotic ones (SARS-CoV, MERS-CoV) (Fig. 3-11). The N1 and N2 CODA assay was highly specific to SARS-CoV-2 with negligible cross-reactivity with SARS-CoV whose genome sequence is highly homologous (79.6%) to that of SARS-CoV-2 [60, 41].
3.4.5 Clinical sample analysis

Figure 3-12: Overall workflow of COVID-19 diagnosis with CODA. Overall workflow of COVID-19 diagnosis with CODA. By integrating the assay and real-time signal measurement, processing, and display into a compact device, CODA enabled onsite COVID-19 diagnosis from one-time sample loading.

<table>
<thead>
<tr>
<th>Patient #</th>
<th>Age</th>
<th>Gender</th>
<th>COVID-19</th>
<th>Patient #</th>
<th>Age</th>
<th>Gender</th>
<th>COVID-19</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>62</td>
<td>Male</td>
<td>Positive</td>
<td>11</td>
<td>41</td>
<td>Female</td>
<td>Negative</td>
</tr>
<tr>
<td>2</td>
<td>62</td>
<td>Female</td>
<td>Positive</td>
<td>12</td>
<td>37</td>
<td>Female</td>
<td>Negative</td>
</tr>
<tr>
<td>3</td>
<td>69</td>
<td>Female</td>
<td>Positive</td>
<td>13</td>
<td>80</td>
<td>Female</td>
<td>Negative</td>
</tr>
<tr>
<td>4</td>
<td>60</td>
<td>Male</td>
<td>Positive</td>
<td>14</td>
<td>34</td>
<td>Female</td>
<td>Negative</td>
</tr>
<tr>
<td>5</td>
<td>60</td>
<td>Male</td>
<td>Positive</td>
<td>15</td>
<td>18</td>
<td>Male</td>
<td>Negative</td>
</tr>
<tr>
<td>6</td>
<td>62</td>
<td>Male</td>
<td>Positive</td>
<td>16</td>
<td>34</td>
<td>Female</td>
<td>Negative</td>
</tr>
<tr>
<td>7</td>
<td>50</td>
<td>Male</td>
<td>Positive</td>
<td>17</td>
<td>14</td>
<td>Female</td>
<td>Negative</td>
</tr>
<tr>
<td>8</td>
<td>79</td>
<td>Female</td>
<td>Positive</td>
<td>18</td>
<td>40</td>
<td>Female</td>
<td>Negative</td>
</tr>
<tr>
<td>9</td>
<td>64</td>
<td>Female</td>
<td>Positive</td>
<td>19</td>
<td>25</td>
<td>Female</td>
<td>Negative</td>
</tr>
<tr>
<td>10</td>
<td>23</td>
<td>Male</td>
<td>Positive</td>
<td>20</td>
<td>21</td>
<td>Male</td>
<td>Negative</td>
</tr>
</tbody>
</table>

Table 3.2: Clinical sample information.

We finally applied the CODA platform to clinical COVID-19 diagnostics. Specimens (nasopharyngeal or oropharyngeal swabs, or sputum) were obtained from COVID-19 suspected individuals at Chonnam National University Hospital (CNUH, Republic of Korea). COVID-19 infection status was independently confirmed at CNUH Clinical Diagnostic Laboratory. We used ten laboratory-confirmed COVID-19 positive samples and ten confirmed negative samples (Table 3.2 for patient information). For each patient, we extracted total RNA from UTM. Three 2-µL aliquots were prepared from the RNA extract and analyzed for N1, N2, and RPP30 (Fig. 3-12).

Fig. 3-14 and 3-13 summarize the CODA assay results. For each gene (T = N1, N2, RPP30), we set the cutoff for positivity as \( \Delta r_T = 5 \cdot \sigma T \), where \( \sigma T \) is the
Figure 3-13: **CODA results of clinical samples.** Three targets (N1, N2, RPP30) were assessed in clinical samples for COVID-19 detection. Data are displayed as mean ± s.e.m from triplicate measurements.

Figure 3-14: **Waterfall plot of CODA results.** Clinically confirmed COVID-19 samples showed higher $\Delta r$ than non-COVID-19 control samples. The cutoff values (dashed lines) were equal to $5 \cdot \sigma_T$, where $\sigma_T$ was the standard deviation of signal without target gene. From triplicate measurements.

standard deviation without the target gene. Applying these criteria, we observed that all samples passed the quality check (i.e., RPP30 positive) with no statistical difference between COVID-19 patient and control samples. In contrast, N1 and N2 were positive only in COVID-19 patient samples (Fig. 3-15a), matching with clinical test results (100% concordance for the current samples). Furthermore, N1 and N2 signal levels, which would be proportional to viral loads, were positively correlated (Pearson correlation coefficient, $r_p = 0.7643$; Fig. 3-15b).

### 3.5 Conclusions

The developed CODA platform distinguishes itself from other one-pot CRISPR/Cas assays [61]. By concurrently executing isothermal amplification and CRISPR/Cas detection in a single device, CODA eliminated extra hands-on steps, completing a-
Figure 3-15: **COVID-19 clinical sample validation tests** (a) RPP30 levels were statistically identical between COVID-19 and control groups, passing the human sample quality check. N1 and N2 levels, however, were significantly different (two-tailed t-test, p < 0.001) between patient and control groups. (b) N1 and N2 signals of CODA, which are proportional to viral loads, were positively correlated (Pearson r = 0.76). Data are displayed as means ± s. e.m. From triplicate measurements.

says within a single sample loading workflow. Measuring fluorescent anisotropy (FA) also offered practical advantages: i) the signaling probe was simpler and cheaper with CODA (i.e., fluorescent DNA) than with other methods that use DNA with a pair of fluorescence dyes and a quencher; and ii) the ratiometric nature of FA measurements made the assay robust against common noise, such as fluctuations of fluorescent intensities. The current study proved the concept by adapting CODA for POC COVID-19 diagnosis. We built a compact device integrating tight temperature control and low-noise optical measurement capability, and also established probes specific to SARS-CoV-2. Using this device, we achieved high detection sensitivity (limit of detection, 3 RNA copy/μL) and completed the entire assay within 20 min of the one-time sample loading step. Overall, CODA’s assay performance was equal or superior to those of other molecular tests (see Table 3.3 for comparison).

Future improvements would accelerate CODA’s field use. Incorporating RNA extraction into the assay system would realize a true “sample-in and answer-out” test. One promising strategy is to integrate solid-phase extraction (e.g., silica beads) [69] in a detachable cartridge. Increasing the assay throughput is also necessary, particularly for COVID-19 diagnostics, to detect all three targets (N1, N2, RPP30) at the
### Table 3.3: Comparison of COVID-19 molecular tests.

<table>
<thead>
<tr>
<th>System</th>
<th>Readout</th>
<th>Assay time (min)</th>
<th>LOD (copy/µL)</th>
<th>Assay characteristics</th>
<th>Reference</th>
</tr>
</thead>
</table>
| Dual-functional plasmonic system    | LSPR             | 14               | $1.13 \times 10^5$ | • Specialized chip from microfabrication  
• Low sensitivity                                                                       | 1         |
| Sandwich electrochemical sensor     | Electrochemistry  | 180              | 0.2            | • Many assay steps including magnetic separation  
• Long assay time                                                                       | 2         |
| RT-PCR                              | Fluorescence     | 120              | 3.2            | • Requiring precise temperature control  
• Dual-labeled probe (fluorophore and quencher)  
• Long assay time                                                                      | 3         |
| RT-LAMP                             | Color            | 20               | 0.5            | • Complex primer design  
• Isothermal amplification                                                              | 4         |
| RT-RPA                              | Fluorescence     | 20               | 1.2            | • Dual-labeled probe  
• Isothermal amplification  
• 2 steps                                                                               | 5         |
| NEAR                                | Fluorescence     | 17               | 20             | • Dual-labeled probe  
• Isothermal amplification                                                              | 6         |
| NASBA                               | Fluorescence     | 120              | 2.5            | • Long assay time  
• Isothermal amplification  
• 2 steps                                                                               | 7         |
| RT-LAMP & CRISPR/Cas               | Color            | 40               | 10             | • Complex primer design  
• Isothermal amplification  
• 2 steps                                                                               | 8         |
| RT-RPA & CRISPR/Cas                | Fluorescence     | 50               | 0.07           | • Dual-labeled probe  
• Isothermal amplification  
• 2 steps                                                                               | 9         |
| CODA (current work)                | Fluorescence anisotropy | 20       | 3              | • Isothermal amplification  
• Ratiometric measurements  
• Single step                                                                           | -         |

References, in order of appearance in table: [62, 63, 64, 65, 66, 67, 68, 54, 55]

same time. Due to its simple structure (e.g., single temperature, compact optics), the CODA device would be readily scalable for such parallelization. On the clinical side, we need to test larger patient cohorts to obtain rigorous assay statistics. Also intriguing to explore is the use of saliva as a test specimen. Not only easily collectable, saliva has also been shown to contain comparable viral loads as nasopharyngeal swabs [70]. These efforts will help us to reach game-changing assay turnaround times, which are critical to mounting a prompt and effective response to emerging infectious diseases.
Chapter 4

Direct Fluorescence system for rapid HPV detection

The work presented in this chapter has been completed in collaboration with biochemist Chang Yeol Lee under the direction of supervising PI Hakho Lee. Assay protocol design, DNA probe validation, and preparation of clinical samples was led by Chang Yeol Lee, while Ismail Degani led the device design, development and validation. This research is currently pending publication.

4.1 Introduction

The worldwide number of deaths due to cervical cancer is alarming: 310,000 deaths were reported in 2018 alone [71]. Even more striking is that 87-88% of these deaths occur in low and middle income countries (LMIC’s) [71, 8]. The impact of socioeconomic status on health outcomes is clear: the age-standardized mortality rate (ASMR) of women in countries ranking ‘low’ on the Human Development Index (HDI) is more than six-fold higher than women in countries with a ‘high’ HDI rank (3.0 vs. 20.0 per 100,000). Thus, it is imperative that clinical strategies are tailored to meet the unique needs of resource-limited environments.

In this chapter, we describe a low-cost system for HPV screening. Like the CODA system of chapter 3, this system combined isothermal nucleic acid amplification, acti-
vation of CRISPR/Cas12a, and signal generation in a single assay. For point-of-care operation, we built a compact, standalone device integrating optoelectronics, an embedded heater, and a microcontroller for data processing. The developed system completed HPV detection within 20 min of sample loading; the limit of detection reached 1 copy/μL. In a pilot test with clinical samples (7 confirmed HPV+ patients; 2 controls), the rapid CODA test accurately classified HPV and HPV subtype in concordance with gold-standard clinical diagnostics.

4.1.1 Human Papilloma Virus

Human Papilloma Virus (HPV) is an important screening target because it is causally linked to 95-97% of all diagnosed cervical cancers \[72, 73\]. There are two high risk subtypes (HPV 16, 18) which alone account for 70% of case. The virus is thought to promote cancer by compromising tumor-suppressing agents within cells \[74\].

4.1.2 HPV Vaccines

The first-line defense against cervical cancer is the prevention of HPV infection. Prophylactic HPV vaccinations targeting multiple HPV subtypes\[75\] have been very effective in this regard, but infrastructural issues greatly impede widespread adoption \[76\]. Fewer than half (43%) of all countries have national HPV vaccination programs, and LMIC’s in particular report vaccination rates of only 13-20% on average \[77\]. Psychological bias and cultural influence \[78, 79\] may also play a role in low adoption rates.

Given the poor penetration of HPV vaccination programs worldwide, it is crucial to develop alternative/complimentary strategies that can together combat cervical cancer more effectively \[80, 81\].

4.1.3 Traditional Screening Tests

The longstanding gold standard for cervical cancer screening has been the Papanicolaou (Pap) smear test \[82\]. Expert pathologists or cytotechnologists manually inspect
an endocervical brushing of cells for morphological abnormalities. Widespread screening efforts have led to dramatic reductions (>50%) in both cervical cancer incidence and death rates [83]. Liquid-based cytology (LBC) is a newer technique which improves upon the Pap smear by first suspending cells in a solution before depositing them in a thin layer on a slide. This intermediate step improves the evenness and spacing of cells which improves sensitivity [84, 85]. While they have made a considerable impact since their introduction in the mid-20th century, Pap and LBC tests suffer from several drawbacks in the context of LMIC’s. First, they require evaluation by highly trained professionals; this increases cost, requires laboratory infrastructure, and introduces potential bias/subjectivity into the diagnosis. Second the sensitivity of these tests can be quite low. Pap smear sensitivity has been shown to be less than 60% [86, 87], and LBC is in the 75% range. For this reason, these tests are often paired with an HPV DNA test.

4.1.4 HPV DNA Tests

HPV DNA screening methods have far greater sensitivity (95%+) than earlier screening approaches [88]. As the clinical emphasis on HPV DNA screening has grown, commercial tests such Cobas® (Roche) and OnClarity™ (Becton Dickinson) are now in widespread use [89]. Additionally, when co-administered with a Pap smear or LBC screen, the specificity of the DNA+Cytology ensemble increases to 100% [90], giving clinicians a powerful screening tool.

However despite this excellent performance, such a screening approach requires multiple followup visits to a clinic, and can be an onerous financial burden on high-risk patients without sufficient economic means [82]. Additionally, commercial diagnostic systems such as the Cobas 8800 are “SUV-sized” machines which require government agency funding in the multi-million dollar range to purchase and operate [91]. They are clearly not designed to operate in small rural or remote environments in LMIC’s. Therefore there is still a large unmet clinical need which requires simpler workflows and more affordable technologies.
4.1.5 VIA/VILI Inspection

Due to the infrastructural and cost challenges of Pap, LBC and DNA testing, low-cost techniques known as Visual Inspection by Ascetic Acid (VIA) and Visual Inspection by Lugol’s Iodine (VILI) have emerged. As their names suggest, these involve visually examining the cervix for color changes after the application of 3-5% acetic acid (VIA) or Lugol’s Iodine (VILI). Evaluations of these assays show overall sensitivity and specificity as high as 80% and 92% respectively [88]. The method also produces results immediately, eliminating costly followup visits. However the results are highly variable due to inter-operator variability. Studies evaluating VIA’s sensitivity show results ranging from 45%-80% [92], which suggests that operator skill has an outsize effect on diagnostic quality.

4.1.6 Overview and Summary

Table 4.1 summarizes the current diagnostic landscape.

There is a clear need for a low cost and highly accurate screening method that (1) does not have the infrastructure and personnel requirements of a Pap smear/LBC test, and (2) has a far lower cost of implementation than a commercial DNA screening method [93]. Standalone VIA and VILI do not fully meet the needs of cervical cancer diagnosis due to their inter-operator variability. Therefore there is a need for new methods that produce fast quantitative results at the point-of-care with minimal cost and without expert technicians. Such a method could be paired with VIA or VILI to rival the Pap/DNA integrated test performance. There have been many next-gen diagnostic technologies proposed by ourselves and others to meet this clinical need [34, 94, 95, 96], and in this chapter we review a new isothermal CRISPR-based HPV DNA screening system that offers unique advantages over prior approaches.
<table>
<thead>
<tr>
<th>Test</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Infrastructure</th>
<th>Drawbacks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pap Smear</td>
<td>55% [86]</td>
<td>96% [86]</td>
<td>Pathologist or Cytologist; Cytology Laboratory Infrastructure Required</td>
<td>Low sensitivity (Pap), Low specificity (LBC), Subjectivity</td>
</tr>
<tr>
<td>Liquid-based Cytology (LBC)</td>
<td>75% [84]</td>
<td>76% [84]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cobas HPV DNA Test</td>
<td>93-96.3% [88]</td>
<td>69.5-70.6% [88]</td>
<td>Specimen prep and analysis equipment, trained personnel</td>
<td>Costly equipment, proprietary reagents</td>
</tr>
<tr>
<td>VILI</td>
<td>56-65% [92]</td>
<td>77% [92]</td>
<td>Minimal: Performed by nurses, midwives, paramedics</td>
<td>Low sensitivity and specificity</td>
</tr>
<tr>
<td>VIA</td>
<td>45% [92]</td>
<td>92% [88]</td>
<td>Minimal: Performed by nurses, midwives, paramedics</td>
<td>Estimates vary widely, inter-operator variation</td>
</tr>
</tbody>
</table>

Table 4.1: HPV Diagnostics Landscape

### 4.2 CRISPR Isothermal Amplification

Isothermal nucleic acid (NA) amplification is a promising technology for onsite NAATs. Using specialized DNA polymerases with the capacity of strand displacement, these methods amplify NA at a fixed temperature, thereby simplifying device design [46][47]. Combining isothermal amplifications with clustered regularly interspaced short palindromic repeats (CRISPR) has been shown to further improve assay sensitivity and specificity [48]. Upon recognition of its target strand, a CRISPR-associated (Cas)
protein and guide RNA complex can indiscriminately cleave single-stranded NAs [49]. This property has been exploited to amplify analytical signal i) through the collateral cleavage of non-target NA probes that have a fluorescent dye and quencher pair [50, 51, 52] or ii) through the release of nanoprobe that are initially anchored by non-target NA probes [53]. CRISPR systems have demonstrated a promising potential for COVID-19 diagnostics [54, 55]. Technical challenges, however, limit the practical point-of-care (POC) use of these assays: i) most CRISPR tests proceed with sequential NA amplification and detection, requiring separate preparation and introduction of CRISPR reagents into samples; and ii) signal readout, typically performed with a dipstick-type lateral flow device, also incurs extra manual steps and produces subjective, qualitative results.

Here, we report on a rapid, quantitative, and streamlined HPV test in a compact system. Like the CODA system of chapter 3, the system a one-pot CRISPR test with a robust detection modality: phase-locked fluorescence.

Specifically, we adopted a one-pot assay scheme wherein isothermal NA-target amplification and CRISPR-based target recognition take place simultaneously. The activated CRISPR/Cas then cleaves fluorescent DNA reporters, increasing the detected fluorescence. This approach offers practical advantages: i) the entire reaction, including cellular lysis, reverse transcription, NA amplification, and signal detection, were conveniently carried out in a single tube at constant temperatures (95°C for cell-lysis, 64°C for amplification); ii) the fluorescence measurement was highly robust to ambient noise due to the implementation of a digital Walsh-Hadamard lock-in amplifier, which led to reliable analytical signal even at low NA concentrations [56]. These features enabled us to advance a portable system integrating low-noise optical detection, precise temperature control, and on-board data processing and display. With the integrated CODA system, we detected HPV DNA for multiple subtypes within 20 min in a ‘sample-to-result’ manner and achieved a sensitivity down to 1 copy/μL. We further applied the system in a pilot clinical test, assessing 9 clinical samples; the results showed a very good (89%) accordance with those by a clinical laboratory.
4.3 Discussion: Improving the CODA Device Design

Developing the CODA system of chapter 3 uncovered several areas for improvement of the device design, most notably in the areas of noise-rejection and detection time.

First, the system’s phase-sensitive lock-in amplifier (LIA) was triggered by a 1-kHz sinusoidal reference wave, and the signal was retrieved using a hardware demodulator followed by a first-order (RC) low-pass filter. These parameters were chosen non-rigorously with a preference towards i) convenience of implementation, ii) minimal cost and component count, and iii) simplicity of design. As a result, the system functioned acceptably but was by no means optimal.

Another concern with the CODA system was the detection time: an average of 40 samples taken every 0.1s was required to produce an accurate and dependable value, resulting in a total exposure time of 4 seconds. During this time, the sample was continuously illuminated by a high brightness LED (irradiance rating of 21.9 $\mu$W/mm$^2$), which induced a meaningful amount of “photobleaching,” or a decrease of fluorescent signal. Photo-bleaching is especially problematic when detecting very low levels of fluorescent reporters. Prolonged exposure can destroy the desired signal in a matter of seconds if the sample only contains fluorophore concentrations in the picomolar range. Fig 4-1 shows the effect of photobleaching in the CODA system on a 240nM Fluorescein solution in 1x phosphate-buffered saline (PBS). After only 30 seconds, the mean fluorescence of the sample diminishes by about 13%.

To address these earlier drawbacks, we designed an improved system with finely-tuned parameters and it yielded far superior performance.

4.3.1 Pink Noise and Modulation Frequency

There has been considerable effort directed towards optimizing the performance of lock-in amplification systems, including waveform shaping, modulation frequency selection, and signal filtering architecture [33]. In general, higher modulation frequencies tend to yield better signal-to-noise ratios (SNR), due to the diminished influence of $1/f$ noise in the high-frequency spectrum. It has long been known that $1/f$ noise is
ubiquitous in electrical systems [97], and the CODA system is no exception: Fig. 4-2 is a frequency plot of noise sampled at 37.5kHz from the system’s analog-to-digital converter (ADC). The $1/f$ noise is nearly 40 times greater than the white noise in the low frequency range.

Figure 4-1: Effect of Photobleaching on Fluorescein 5(6)-isothiocyanate

Figure 4-2: White and $1/f$ ADC Noise at 37.5kHz
Despite this clear inverse relationship between modulation frequency and noise power, it is not necessarily optimal to simply choose the maximum possible modulation frequency. This is because hidden “noise-maxima” exist all throughout the spectrum but are averaged out when using typical FFT window sizes (512-8192 samples); they only reveal themselves at adequately high frequency resolutions. Such spikes arise due to ambient electromagnetic interference, switching power-supply noise, AC hum/harmonics, and any other idiosyncrasies of the environment (fluorescent light flicker, etc). The “true” white noise floor is far lower than what Fig. 4-2 would suggest. However, due to the low frequency resolution, these spikes “smear” across the spectrum, and contribute to the perceived white noise. To illustrate this further, Fig 4-3 shows very similar ADC noise of Fig 4-2 now sampled at 407kHz from a faster microcontroller (Kinetis ARM Cortex M7). The noise is plotted in frequency space using progressively larger FFT window sizes. Each doubling of the window size correspondingly doubles the frequency resolution, and the contours of these frequency spikes become more pronounced. It is clear that 60Hz is a very bad choice of modulation frequency because it corresponds to AC hum. However, equally poor is the higher-frequency 50kHz band, where power supply switching noise ($P_{50}$) is dominant. Modulating a signal here is 2-3 orders of magnitude noisier than the nearby 45kHz or 75kHz bands, which are much closer to the white noise power threshold ($P_{white}$). Therefore, one must tread carefully to choose a modulation frequency that doesn’t “step on any spikes.”

Furthermore, as we’ll see later in this chapter, the optimal modulation frequency is also highly dependent on the gain-bandwidth product (GBP) limitations of the system.

Aside from the frequency of the reference modulation, the waveform’s shape also has a sizeable impact on the system’s overall performance. The most commonly used modulation references are sinusoidal or square wave-based. Sinusoidal references may be optimal for downstream filtering as they can be very efficiently demodulated by RC filters or by the Fast Fourier Transform (FFT) in the case of digital LIA’s. Square wave references have advantages as well, including lower complexity: as an example,
generating an optical square wave can very easily be accomplished with a simple “chopper wheel” that periodically occludes a light beam as it rotates.

4.3.2 Walsh-Hadamard Modulation

Given the flexibility of modern digital-analog converters (DAC’s) to generate arbitrary high-resolution waveforms, experimenting with complex (spread spectrum, aperiodic, pseudo-random) waveform-shaping is relatively straightforward. The guiding principle in choosing a reference wave is to select the one that most effectively “evades” the ambient noise of the system, thus maximizing the overall SNR. This can be a non-trivial task given that the noise may be transient/non-stationary, and the system dynamics might be non-linear. Other considerations would include fast processing, and ease of signal generation, and total signal power.

In this work we examined an interesting class of functions known as the Walsh-Hadamard basis functions, and evaluated their utility as modulation references in a lock-in amplifier. These are a binary class of step functions that bear many similarities to the Fourier basis of sines and cosines, but also possess unique advantages in the
context of lock-in amplification.

\[ H_2 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} W_{(1,2)} \\ W_{(2,2)} \end{bmatrix} \]  

(4.1)

The Walsh basis functions are defined by the base 2x2 Hadamard Matrix as shown in Eq. 4.1. Each Walsh function \( W_{i,n} \) represents the \( i^{th} \) row of this \( n \times n \) matrix. For example, the first length-2 Walsh function \( W_{(1,2)} = [1, 1] \) is the first row of \( H_2 \).

To get longer sequences, we create larger Hadamard matrices by recursively stacking Hadamard matrices as shown in Eq. 4.2.

\[ H_4 = \begin{bmatrix} H_2 & H_2 \\ H_2 & -H_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} W_{(1,4)} \\ W_{(2,4)} \\ W_{(3,4)} \\ W_{(4,4)} \end{bmatrix} \]

(4.2)

This extraordinarily simple construction yields collections of length \( n \) Walsh functions with very useful properties for lock-in amplification:

- Walsh sequences of a given length \( n \) are orthogonal to all other sequences of the same length. Like the Fourier basis, they form a complete orthogonal system and the dot-product of any two sequences is zero. More formally, the cross-covariance of any two length \( n \) Walsh sequences at lag zero is always equal to zero:

\[ \text{cov}(W_{(i,n)}, W_{(j,n)})_{\tau=0} = \sum_{k=1}^{n} w_{(i,n)}[k] \cdot w_{(j,n)}[k] = 0 \]  

(4.3)

- A consequence of the orthogonality property is that all Walsh sequences sum to zero, with the exception of \( W_{(1,n)} \) (the “DC sequence”). This sequence consists solely of ones, so in order for other Walsh functions to maintain zero covariance
Figure 4-4: Overlaid Direct Comparison of Walsh and Fourier Basis Functions. This figure shows a length 16 Walsh basis (blue) overlaid with corresponding Fourier coefficients (orange). The substantial similarities are apparent including their sharing of a “DC Term” and having a natural ordering defined in terms of increasing zero-crossings. In fact, even the phase relationship between sine and cosine has an analog in Walsh space: odd sequencies $W_{(2m+1,n)}$ are phase-shifted copies of even sequencies $W_{(2m,n)}$. 
with it, they must all sum to zero:

$$\text{cov}(W_{i,n}, W_{1,n})_{\tau=0} = \sum_{k=1}^{n} w_{i,n}[k] \cdot 1 = 0 \rightarrow \sum_{k=1}^{n} w_{i,n}[k] = 0$$  \hspace{1cm} (4.4)$$

The analog of this sequency in Fourier space is the DC term: a cosine with frequency 0 and amplitude 1.

- A consequence of the zero-sum property is that all Walsh sequences (again, excluding DC) contain exactly half 1’s and half -1’s. In the context of lock-in amplification, this means they all have equal root-mean square (RMS) power when used as modulation reference functions.

- For a given $n$, all Walsh functions possess a unique number of transitions from -1 to 1, or zero-crossings. Ordering functions by the number of zero-crossings is known as “sequency ordering” because it is identical to the concept of frequency-ordering for sinusoidal functions that comprise a Fourier basis. This ordering scheme is shown side-by-side its Fourier equivalent in Fig 4-4 to illustrate its substantial similarity.

- Similar to the Fast Fourier Transform (FFT), an $O(n \log n)$ Fast Walsh-Hadamard Transform (FWHT) exists. The FWHT is especially efficient running on microcontrollers because it requires only additions and subtractions and can be optimized for 16-bit integer input via microcontroller intrinsics (i.e. the NEON instruction set, available on ARM Cortex M-Series embedded systems). The Walsh-Hadamard transform has been used across a wide range of applications including sleep-cycle analysis [98,99], quantum computing algorithms [100,101], image compression [102,103], and evolutionary tree estimation by DNA sequence analysis [104].

An excellent, in-depth overview of the many properties of Walsh functions is given in Stoffer’s review [105], but the above should provide enough background to explain the mechanics of the lock-in amplifier in the proposed HPV diagnostic system.
An important question to ask is: why use Walsh-Hadamard modulation references instead of tried-and-true sinusoids or square waves? The strongest justification for any waveform modulation scheme lies in proving that it produces the highest SNR in a given system; secondarily, it is advantageous to show that it a) is more computationally efficient, or b) reduces cost by using fewer/cheaper components. This is what we will explore in the next two sections.

### Waveform Design

![Single-Sequency (12-bit ADC Adjusted)](image1)

![Multi-Sequency (12-bit ADC Adjusted)](image2)

### Waveform Output

![ADC_MAX = 4095](image3)

Figure 4-5: **FWHT Sample Inputs and Outputs** The above illustrates how the FWHT can be used to custom-design waveforms with the precise desired spectral characteristics. These examples assume a digital-analog converter (DAC) that requires output between 0 and $2^b - 1$ where $b$ is the bit-depth specification. For strictly positive signals between 0 and ADC_MAX, the waveform must be shifted by adding $[2^{b-1} - 1/2]$ to the DC term. And when designing multi-band signals as shown in the lower plots, the total energy must be shared between each Walsh component $W_{(i,n)}$ in order to keep the waveform’s maximum $\leq$ ADC_MAX. This is why each stem is shorter in the multi-sequency case.

### 4.3.3 Generating Walsh-Hadamard Waveforms

Generating Walsh functions with specific spectral properties is very straightforward. To generate the $m^{th}$ Walsh function of length $n$ the following steps can be used: (1)
construct a length $n$ array $T_n$ of zeroes with a single delta function at location $m$ and (2) apply the WHT to this array. In a Fourier-based system, this process would be identical except the inverse-Fourier transform would be applied. A few simple variations on this technique can be used to meet the modulation requirements of most systems:

- For a system where the output waveform should be between $[-A, A]$ (where $A$ might be a hardware-dependent voltage), the delta function should be set to $A$ instead of 1.

- If the signal needs to be strictly positive $[0, A]$ (i.e. for a single power-supply configuration), a DC term equal to $A/2$ should be added to the 0th element of $T_n$.

- If a spread-spectrum signal containing $k$ sequencies is desired, one can simply add $k$ delta functions to $T_n$ wherever desired. In this case the magnitude of each delta function should be $A/k$ so that the modulated signal does not exceed the system maximum $A$.

Two examples of single-sequency and multi-sequency waveform generation used in the proposed HPV system are shown in Fig. 4-5. The modulations were used to drive a fluorescent LED light source, and so positive-only values were used. Additional details and source code for waveform construction can be found in Appendix A.

### 4.3.4 Optimization of Walsh-Hadamard Lock-in Amplifier

We first set out to determine the optimal Walsh-Hadamard waveform parameters to produce the highest signal-to-noise ratio. The sampling rate of the system (200 kHz) was pre-determined by the system hardware: the low-cost microcontroller had a maximum data acquisition speed, determined by the ADC settling times (Fig. ), the MCU’s data bus clock frequency, and the maximum supported USB serial baud rate.
Many combinations of multiple-sequency and single-sequency waveforms of varying lengths were tested. It was found that the best results were achieved by single-sequency waveforms that were in low-noise, mid-to-low frequency areas of the spectrum, as shown in Fig. 4-6.

Figure 4-6: Optimal Walsh-Sequency Determination by Exhaustive Sweep
The above shows an exhaustive sweep of the Walsh-Hadamard space of length 512 sequences (low sequencies are towards the left and high sequencies are towards the right). The mean fluorescence (green) of a sample, was measured 60 times and averaged at each sequency. The std deviation of the measurements is shown as a 95% confidence interval (blue). The noisier areas near sequency 200 and 300 exhibit wider bands and have poor SNR accordingly. The perceived fluorescence magnitude also decreases with increasing sequency due to the Gain-Bandwidth Product Limitations of the system, yielding an optimal modulation function at $W_{(107,512)} \approx 20.9$kHz.

This was a surprise because the noise profile of the ADC’s from section 4.3.1 suggested that the optimal frequency was likely located at the upper-end of the spectrum. However, the Gain-Bandwidth Product (GBP) of the system, particularly due to the AD823AN amplifier, attenuated the signal amplitude at higher frequencies. Thus even though there was overall lower noise, there was also commensurately lower signal.

Electronic circuits in general all contain parasitic capacitance sources which atten-
valuate higher frequency signal. Additionally, op-amps are often designed with specific internal junction capacitances to improve stability at the expense of high-frequency performance \[106, 107\]; and all have slew-rate limits that also determine their high-frequency performance.

Given the many factors that affect SNR, Walsh “Sequency Sweeps,” were performed to empirically determine most optimal values. These sweeps rapidly measured average SNR at every possible sequency. A length-512 sweep is shown in Fig. 4-6, which determined that the optimal sequency occurred around 20.9kHz (sequency number 107). This optimal value was set as the preferred modulation for the experiments performed in this chapter, but the reader should note that it is likely to change given the experimental conditions (indoor lab environment vs. outdoor field setting, the profile of nearby electromagnetic interference (EMI), overhead lighting, etc). Given this, the system should adaptively sample the ambient noise and generate an optimal reference modulation on the fly.

The GBP limitation also complicates the use of spread-spectral signals and is the reason why they were discarded from the system design. If a modulation sequency is constructed with two Walsh functions (i.e. \( W_{(a,n)} + W_{(b,n)} \), where \( b > a \)), the returned signal will be of the form \( W_{(a,n)} + q \cdot W_{(b,n)} \), where \( q < 1 \) is some attenuation factor imposed by the GBP characteristic. Determining this \( q \) complicates the signal acquisition and so the simpler-single frequency was opted for instead.

With an optimal digital lock-in strategy, we proceed to assay development and experimental results in the following sections.

4.4 Materials and Methods

4.4.1 Materials

All oligonucleotides used in this study were synthesized by Integrated DNA Technologies (Coralville, IA, USA), including gBlocks® gene fragments of HPV subtype targets (HPV16, 18, 31, 33, and 58) and GAPDH. TwistAmp® Basic kit was purchased from
TwistDx™. EnGen® Lba Cas12a (Cas12a) and RNase inhibitor (Murine) (RI) were purchased from New England BioLabs (Beverly, MA, USA). DEPC-DW, purchased from Thermo Fisher Scientific, was used in all experiments. All other chemicals were of analytical grade and used without further purification.

4.4.2 HPV assay

The HPV master mix was prepared by combining 240 nM gene-specific RPA primers, 200nM reporter probe, 160nM gene-specific Cas12a gRNAs, 0.8 U/μL RI, 640 nM Cas12a, 0.2 × NEBuffer 2.1 (10mM Tris-HCl, 50mM NaCl, 10mM MgCl2, 100μg/mL BSA, pH 7.9 at 1× concentration), and reconstituted RPA mix. Upon addition of viral DNA and 14 mM MgOAc, the HPV mix was incubated at 37°C for 20 min. Fluorescence was measured during the reaction.

Figure 4-7: Assay schematic: System for HPV Detection. When target viral DNA is present, RT-RPA and CRISPR/Cas detection take place together. DNA polymerase recognizes target sequences and displaces double-stranded DNA. Cas12a/gRNA complexes then bind to specific sites (green and orange) in the exposed single strand, get activated, and start to cleave nearby reporter probes. This cleaving process is amplified, as the RT-RPA reaction proceeds. As a result, the fluorescence intensity (F) of the sample increases (right).
4.4.3 Clinical sample collection and analyses

The clinical study was approved by the Partners Healthcare Institutional Review Board (Massachusetts General Hospital/Brigham and Women’s Hospital). Clinical, cervical brushing samples were collected and fixed in BD SurePath™ liquid-Pap test solution (BD Biosciences). DNA was extracted from 500μL of samples, using Quick-DNA FFPE Miniprep kit (Zymo Research) with a slight modification of manufacturer’s instructions. 2μL of extracted DNA was analyzed as described in section 4.4.2. For sample quality control and signal normalization, GAPDH, which is a housekeeping gene, was used as a control target.

4.4.4 HPV device construction

The HPV device was modeled in computer-aided design software (Solidworks, 2019) and fabricated by hand-soldering a standard perfboard (Vector 8015). The main body housed a light source sub-assembly, filter cube holder, and sample holder. i) The light source consisted of an LED (Thorlabs M470D2), and a 2 biconvex focusing lenses (Thorlabs LB1092-A). The LED was attached to a metal heat-sink through a thermal compound (Arctic Silver 5) and driven by a single-supply, low noise LED current source driver controlled by the MCU’s 16-bit digital-to-analog converter (DAC) (Analog Devices CN0370). ii) The fluorescent detector sub-assembly consisted of a FITC Filter Cube (Nikon), convex lens (Thorlabs LB1761-A) and a photodiode (Hamamatsu S1223). iii) Fluorescent readings were amplified using a single 17MHz AD823AN op amp. Lock-in amplification was performed entirely digitally in firmware, using a 20.9 kHz carrier wave generated by the integrated DAC of an ARM Cortex M4 microcontroller (MK20DX256, PJRC). iv) For sample heating, a metal sleeve was machined in aluminum. An integrated 15 W (Thorlabs HT15W) ceramic heater was driven by a power transistor (TIP120) attached to an 18 V supply. Temperature was monitored by a 10 kΩ thermistor and controlled to within 0.2 °C by a proportional-integral-derivative (PID) feedback loop running on the microcontroller unit (MCU). V) For the graphical user interface (GUI), a real-time dashboard was programmed...
using the Qt GUI framework (Fig. 4-8), which communicated with the MCU over the USB serial port. The MCU firmware was written in C++ (Arduino). Cost estimates for the system are shown in Table 4.2.

<table>
<thead>
<tr>
<th>Component</th>
<th>Specification</th>
<th>Cost ($ USD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Microcontroller</td>
<td>Cortex M4 180MHz</td>
<td>30</td>
</tr>
<tr>
<td>ADC (Built-into MCU)</td>
<td>400ksps-1Mps</td>
<td>0</td>
</tr>
<tr>
<td>DAC (Built-into MCU)</td>
<td>200ksps / 0.6us settling</td>
<td>0</td>
</tr>
<tr>
<td>LED</td>
<td>470nm, 1A</td>
<td>5</td>
</tr>
<tr>
<td>Photodiode</td>
<td>Hamamatsu S1223: 10ns rise time</td>
<td>10</td>
</tr>
<tr>
<td>Op-Amps (x2)</td>
<td>AD823AN: 17MHz, 325ns settling time</td>
<td>20</td>
</tr>
<tr>
<td>Transmission Filter</td>
<td>530nm</td>
<td>150</td>
</tr>
<tr>
<td>Excitation Filter</td>
<td>488nm</td>
<td>50</td>
</tr>
<tr>
<td>Other Optics</td>
<td>Convex Lenses</td>
<td>20</td>
</tr>
<tr>
<td>Turret Motor</td>
<td>NEMA-8</td>
<td>20</td>
</tr>
<tr>
<td>Circuit Board (PCB)</td>
<td>2 Layer, 6&quot;x4&quot;</td>
<td>25</td>
</tr>
<tr>
<td>Miscellaneous</td>
<td>Outer case, hardware, cabling</td>
<td>25</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>355</strong></td>
<td></td>
</tr>
</tbody>
</table>

Table 4.2: Cost Breakdown of HPV Fluorescence Unit

### 4.4.5 HPV device signal processing

Walsh-modulated fluorescence signals of length 512 were captured by a photodiode and directly sampled at 200 kHz by a 12-bit analog-to-digital converter built-into the MCU. The integrated real-time monitoring GUI displayed raw fluorescence readings as well as the FFT and FWHT of the signal in real-time (60 frames-per-second). The optimal (highest-SNR) modulation sequency \( W_{(\star, n)} = W_{(107, 512)} \) was identified by a sequency-sweep of the Walsh space. This sequency roughly corresponds to a frequency of 20.9kHz. The demodulated signal was obtained directly from the \( W_{(\star, n)} \) coefficient of the FWHT, which is equivalent to a cross-correlation of the modulated signal with \( W_{(\star, n)} \). The affine transformation \( \mathcal{R}(f) \) of Eq. 4.5 was applied to re-scale the device’s values to be similar to those measured by a plate reader (GENios Pro, TECAN GmbH). Raw readings are shown in table 4.3. The device’s data was...
Figure 4-8: **HPV Device Qt GUI.** A screenshot of an extended user interface, displaying FFT and FWHT transformations of the raw signal. The HPV device communicates raw samples at 200kHz to a computer to present demodulated signal data and live SNR measurements.
substantially similar to the plate reader, as shown in Fig. 4-9

\[
\mathcal{R}(f_p^i) = [f_p^i - \min(f_d)] \frac{\max(f_p) - \min(f_p)}{\max(f_d) - \min(f_d)} + \max(f_p)
\]  

(4.5)

Figure 4-9: **HPV Device Validation Fit** Thirteen samples measured for fluorescent intensity by a commercial plate reader and the HPV device.

Noise fluctuations arising from both raw and computed data streams were precisely measured as the rolling standard deviation of the most recent 60 samples, to confirm that the system’s measurement was reliable. The device’s signal-to-noise ratio was measured to be 111.26dB.

### 4.5 Results and Discussion

#### 4.5.1 HPV assay

The HPV assay combines isothermal NA amplification and CRISPR/Cas12a detection. In the presence of the target NA sequence, recombinase polymerase amplification (RPA) takes place with gene-specific RPA primers. During this process, single strand DNAs downstream to the primer binding sites (purple and blue) are
<table>
<thead>
<tr>
<th>Sample</th>
<th>Concentration</th>
<th>((\mu)_{fd})</th>
<th>((\sigma)_{fd})</th>
<th>(f_p)</th>
<th>(\mathcal{R}(f_p))</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>47.0 (\mu)M</td>
<td>31197.3</td>
<td>178.1</td>
<td>17007.0</td>
<td>31197.3</td>
</tr>
<tr>
<td>2</td>
<td>23.5 (\mu)M</td>
<td>19618.7</td>
<td>174.1</td>
<td>10070.0</td>
<td>18537.8</td>
</tr>
<tr>
<td>3</td>
<td>11.8 (\mu)M</td>
<td>11324.8</td>
<td>93.9</td>
<td>4481.0</td>
<td>8338.3</td>
</tr>
<tr>
<td>4</td>
<td>5.87 (\mu)M</td>
<td>6178.0</td>
<td>69.4</td>
<td>2149.0</td>
<td>4082.6</td>
</tr>
<tr>
<td>5</td>
<td>2.94 (\mu)M</td>
<td>3067.4</td>
<td>65.8</td>
<td>1150.0</td>
<td>2259.5</td>
</tr>
<tr>
<td>6</td>
<td>1.47 (\mu)M</td>
<td>1283.8</td>
<td>49.1</td>
<td>554.0</td>
<td>1171.8</td>
</tr>
<tr>
<td>7</td>
<td>0.73 (\mu)M</td>
<td>887.1</td>
<td>46.9</td>
<td>295.0</td>
<td>699.2</td>
</tr>
<tr>
<td>8</td>
<td>0.37 (\mu)M</td>
<td>571.9</td>
<td>45.6</td>
<td>150.0</td>
<td>434.5</td>
</tr>
<tr>
<td>9</td>
<td>0.18 (\mu)M</td>
<td>372.9</td>
<td>45.5</td>
<td>98.5</td>
<td>340.6</td>
</tr>
<tr>
<td>10</td>
<td>91.8 nM</td>
<td>272.8</td>
<td>47.6</td>
<td>47.0</td>
<td>246.6</td>
</tr>
<tr>
<td>11</td>
<td>45.9 nM</td>
<td>232.7</td>
<td>43.1</td>
<td>28.0</td>
<td>211.9</td>
</tr>
<tr>
<td>12</td>
<td>23.0 nM</td>
<td>197.3</td>
<td>43.5</td>
<td>19.0</td>
<td>195.5</td>
</tr>
<tr>
<td>13</td>
<td>11.5 nM</td>
<td>182.7</td>
<td>46.1</td>
<td>12.0</td>
<td>182.7</td>
</tr>
</tbody>
</table>

\(r^2(f_d)\) 0.99034 \hspace{1cm} 0.99034

Table 4.3: **HPV Device Validation Data.** Thirteen samples were analyzed, spanning the active range of Fluorescein where the signal is above the noise floor and below the auto-quenching concentration of the fluorophore. Columns \((\mu)_{fd}\) shows the mean fluorescence of 60 measurements gathered by the HPV device in rapid succession (2ms apart), and column \((\sigma)_{fd}\) shows the std. dev of these measurements. Column \(f_p\) shows the fluorescence intensities of the samples analyzed by a plate reader, and \(\mathcal{R}(f_p)\) applies the transform of Eq. 4.5. Adjusted \(r^2\) is computed for both raw and transformed values to confirm that \(\mathcal{R}(f_p)\) does not affect the fit.
exposed by the strand displacement activity of DNA polymerase. When gRNAs recognize the specific sites (green and orange) at both ends of displaced single strands, Cas12a/gRNA complexes get activated to cleave DNA [49]. Note that most cleavage actions would be limited to reporter DNA (trans-cleavage). The protospacer adjacent motif (PAM) sequence is absent downstream to the recognition sites, which makes it unlikely that activated Cas12a invades and cuts target DNA [57]. Furthermore, RPA synthesizes complementary sequences on displaced single strands, effectively keeping Cas12a (RuvC nuclease domain) from cleaving single strands in target DNA.

As exponential NA amplification proceeds, the trans-cleaving activity is reinforced, activating a large number of auto-quenched fluorescent reporters. Accordingly, the overall fluorescence increases (Fig. 4-7 right). The isothermal, dual amplification of target NA and the fluorescence signal allows for a rapid one-pot assay (20 min). Minimal intervention as well as a real-time quantitative readout make the HPV system robust to external interference.

A prototype HPV device (Fig. 4-10) was designed and fabricated to facilitate the HPV assay reaction and perform fluorescent detection. Specifically, the device controlled the temperature in order to lyse cells in the cervical brushing sample (64 °C, 20 min; 95 °C, 10 min) and extract viral DNA. It then maintained constant reaction temperature (37 °C) through a feedback control, rapidly measuring fluorescence values in less than 2 ms using phase-sensitive detection techniques. The prototype device had a form factor of 16 × 6 × 5 cm³. A sample container (a PCR tube with diameter 6.2 mm) was inserted into a form-fitting metal holder for heating; in the final version, a turret is planned to hold a total of 12 samples for parallel processing. The device communicated with a computer running real-time analysis software (Fig. 4-8) that displayed and stored data.

4.5.2 HPV device

Fig. 4-11 shows the schematic of optoelectronics inside the CODA device. We adopted a digital lock-in method based on the Walsh-Hadamard transform for low-noise measurements. An LED generated modulated blue light at varying intensity
The CAD design (top) and photo (bottom) of the HPV analysis device. A prototype device integrates fast-Walsh signal modulation and real-time fluorescence monitoring. A sample tube (50 μL) is inserted into a metal block whereby a photodetector detects fluorescence emission from the sample. Eventually a motorized turret will be implemented to facilitate heating and analyzing 12 samples in parallel.
\( f \approx 20.9 \text{kHz} \) which passed through a FITC transmission filter onto the sample. The sample emitted green light (525 nm) oscillating in intensity at the same modulation frequency (20.9 kHz) with an intensity profile corresponding to the fluorophore concentration. The detector captured this raw intensity data and a microcontroller filtered and processed it in real-time. The detector consisted of a photodiode, a convex lens, and an emission filter. Once the signal had been captured, a custom-designed lock-in amplifier implemented on the microcontroller adaptively rejected common noise based on the unique noise present in the environment.

![Electronic Schematic of HPV Device](image)

Figure 4-11: **Electronic Schematic of HPV Device** A Walsh Hadamard signal is configured by a user interface (GUI) and sent to a microcontroller. The buffered signal drives a high-power LED (blue section), which excites a photodiode that is read back into the MCU and transferred to the PC where it is processed. A secondary amplification stage (yellow) may be used for extremely faint signals, but only if the first stage op-amp cannot supply additional gain. The second stage will incur \( \sigma \) noise rather than the lower \( \sqrt{\sigma} \) noise of a single-stage configuration.

We compared the HPV Device’s optical performance against that of a benchtop plate reader (GENios Pro, TECAN GmbH). We prepared standard samples of varying fluorescent intensity in the active range of the fluorophore. Intensity measurements were computed every 2 ms, and 60 readings were averaged to reduce sampling noise. The fluorescence \( (f) \) was then computed by a graphical user interface on an at-
attached laptop computer. The measured $f$ values showed a good match ($R^2 = 0.9903$) between the HPV device and the plate reader [4-9], which confirmed its accuracy in optical detection. Combining noise-monitoring with tightly controlled exposure times reduced the variability between measurements and increased the overall repeatability of experiments.

4.5.3 Assay characterization

![Fluorescence graph]

Figure 4-12: HPV Assay Feasibility Study The Assay mechanism was validated by varying the reaction conditions. Fluorescence remained high when any single assay component was missing. The value significantly decreased when all key components were present. Data are displayed as means ± s. e.m. From triplicate measurements.

After successfully validating our assay parameters (Fig. 4-12), we carried out a titration experiment with serially diluted HPV16 DNA samples (Fig. 4-13, left). Based on a blank signal $+ 3\sigma$, where $\sigma$ is the standard deviation at the blank, the limit of detection was determined to be 0.04 copy/$\mu$L.

We also performed sensitivity analysis using the CaSki cell line, which is HPV 16 positive (Fig. 4-13, right). Notably, the HPV assay yielded a robust signal even at low cell concentrations and detected down to a single CaSki cell, achieving wider dynamic range than conventional RT-PCR. This merit can be attributed to a high
Figure 4-13: **HPV Assay Sensitivity** Using synthetic HPV16 DNA, the detection limit was found to be 0.04 copy/µL (left). In a cell line (CaSki), the system was able to detect HPV16 DNA with only one cell present (right). Data are displayed as means ± s. e.m. From triplicate measurements.

The signal-to-noise ratio inherent in the phase-sensitive measurement: the light entering the photodetector remained substantially robust to the detectors’ intrinsic noise.

We further tested the specificity of prepared probes for HPV16 and HPV18 against cervical cancer cell lines (CaSki, SiHa, HeLa, C33a) containing one or neither of each gene (Fig 4-14). In all cases, we obtained excellent specificity corresponding to the expected level of gene (i.e. DNA copy number) present in each cell.

### 4.5.4 Clinical sample analysis

We finally applied the HPV platform to clinical cervical cancer diagnosis in a preliminary study. Specimens (cervical brushings) were obtained from individuals at Massachusetts General Hospital (MGH, Boston, USA). HPV status and subtype was independently confirmed at MGH Clinical Diagnostic Laboratory. We used seven laboratory-confirmed HPV positive samples and two controls. For each patient, we extracted DNA and 2-µL aliquots of DNA extract were analyzed for HPV 16, 18, 31, 33 and 58.

Fig. 4-15 summarizes the HPV assay results. We first confirmed that all samples...
passed the quality check (i.e. GAPDH positive) with no statistical difference between 
HPV patient and control samples. For each gene (HPV16, 18, 31, 33 and 58), we set 
the cutoff for positivity as $F = 5 \cdot \sigma T$, where $\sigma T$ is the standard deviation without 
the target gene. Applying these criteria, we analyzed and compared the results with 
those of the hospital. The results substantially matched with clinical test results (89% 
concordance for the current samples).

81
Figure 4-15: HPV clinical sample test data Data are displayed as means ± s.e.m. From triplicate measurements.

### 4.6 Conclusions

The HPV platform represents a technological advance over the CODA system architecture, while simultaneously demonstrating that the versatile CRISPR-RPA amplification assay is applicable to a wide range of targets. Like CODA, concurrently executing isothermal amplification and CRISPR/Cas detection in a single device has the potential to streamline clinical workflows. Measuring direct fluorescence in the HPV system, rather than CODA’s anisotropy (FA) design, did have disadvantages: in particular, it reduced the common-noise rejection of the system. However, the direct
measurement also had significant advantages including a substantially lower part-count and therefore a lower cost and complexity. The HPV system only required a single detector and signal processing pipeline, and its part-count was further reduced by the digitization of the lock-in system into firmware. The HPV study proved the concept of accurate isothermal CRISPR-based HPV diagnosis by adapting and improving on the previous CODA assay designs. We established probes specific to HPV, and using a prototype device, we achieved high detection sensitivity (limit of detection, 1 RNA copy/μL). We completed the entire assay within 20 min of the one-time sample loading step. A compact HPV device is forthcoming, which will integrate tight temperature control and a 12-sample turret for parallelized sample processing. With our low-noise signal acquisition technique and improved probe design, we expect a substantially compact and accurate device.

Additionally, with our finalized device we will test larger patient cohorts to obtain rigorous assay performance statistics. The proposed HPV system may one day serve as a valuable tool that improves early cervical cancer detection rates in resource-limited areas, by equipping clinicians with better and more accurate diagnostic capabilities.
Chapter 5

Conclusions

5.1 Thesis Overview

In this thesis we have described a new technique for signal processing, and implemented two diagnostic platforms that (1) have leveraged it to detect minute quantities of DNA/RNA, (2) are very low cost and (3) have a compact form-factor. By adaptively sampling noise in the Walsh-Hadamard space and using this information to generate tailored modulation patterns, we demonstrated the ability to extract very faint signals in a computationally efficient manner. Finally, we documented the evolution of our systems from conception through to a clinical study.

The goal of this research is to create devices that can meet challenging clinical needs using new approaches, and by borrowing ideas across disciplines. Below are some future directions such research may take.

5.2 Areas for Future Development

5.2.1 Multiplexed detection

The most direct improvement on the devices discussed would be to incorporate multiplexing capability into the assay and device. Currently, each of the nanoprobe (SARS-CoV-2 N1, N2 / HPV 16, 18) is currently inspected separately for fluorescent
activity; this is a necessity because the nanoprobes are all conjugated to Fluorescein isothiocyanate (FITC) fluorophores. A more streamlined system might pair complementary fluorophores with each probe in order to simultaneously detect multiple target genes in a single step. This would require both (1) careful fluorophore selection to avoid crosstalk between excitation and emission wavelengths (2) innovative device design to minimize the number of optical filters, detectors, moving parts, and overall electronic complexity.

A promising approach to facilitate multiplexing might be to adapt earlier work [108] which improves upon the standard fluorescence detection paradigm. Instead of employing multiple emission filters, coupling multiple excitation light sources to a single quad-bandpass filter reduces the total number of optical components needed. Newer approaches described in the following sections may facilitate multiplexing in an even more scalable and cost-effective manner.

5.2.2 Gigahertz sampling for “filter-free” fluorescence

Analog optical filters have been the most expensive components (by far) of the systems we have built. These include fluorescence excitation/emission filters, as well as dichroic mirrors which further impede unwanted wavelengths from reaching the sensitive photodetectors.

These components require complex manufacturing processes, involving depositing layers of dielectrics with λ/4 spacings that tune their passband and stopband characteristics [109]. More importantly however, these filters do not benefit from the same economies-of-scale as semiconductor technologies (integrated circuits, photodiodes, microcontrollers). Most semiconductors have equally or even more complex fabrication processes, but can often be purchased for pennies due to their broad demand across multiple industries; such demand spurs production at the million or billion-unit scale. Given this, it is advantageous to reduce a system’s dependence on biomedical “specialty components” by re-purposing generic components that are already manufactured at global scale. In this spirit, investigating a “filter-free” fluorescence system might be a promising avenue to reduce cost and enable multiplexing capability. It is
well known that fluorophore species can have very different decay profiles: some may decay fully within 1ns after activation, whereas others may take hundreds of nanoseconds. Additionally, new compounds such as fluorescent nanodiamonds (FND), having superior photostability and tunable decay characteristics are actively being researched \[110\]. By integrating advances from the telecommunications industry such as field-programmable gate arrays (FPGA) and gigahertz-range ADC’s, it may be possible to directly measure fluorescent decays from a biological sample at very low-cost and with a portable form-factor. The idea is to momentarily activate the sample with a broad-spectrum light source, and sample the impulse response at the nanosecond scale to estimate the fluorophore concentration. This method could also be enhanced with a powerful multiplexing capability: multiple superimposed decay profiles may be separable via expectation maximization on a mixture model. Such a system could completely eliminate the need for multiple detectors/filters.

5.2.3 Wavelength-Division Multiplexing

Another technique to move beyond the realm of analog optical filters would be to employ a filtering technology such as Wavelength-Division Multiplexing (WDM). WDM has long been used by the telecom industry for bidirectional multiplexed communication over fiber optic lines in the infrared bands. However, it has recently been adapted to visible wavelengths specifically for biomedical applications (fluorescence and cytometry) \[111\] \[112\]. Incorporating WDM into a fluorescent or FP assay could achieve a color-multiplexing capability with an extremely small form-factor.

5.2.4 Further Miniaturization and Portability

Fig. 5-1 shows a concept rendering of a miniaturized plate reader running on battery power and having no moving parts. Each well has its own emitter/detector pair, and is excited by a unique modulation waveform. Using orthogonal phase-sensitive detection, this reader may be able to scan an entire 96-well plate in a matter of milliseconds. Measurements from all wells would take place in parallel and the system
would be agnostic to ambient light and electrical noise. The ability to scan plates so rapidly and conveniently may enable new types of research and bring fluorescence testing to unlikely places. With plate reader that is no larger than a plate, one may be able to put it inside an incubator, refrigerator, or a custom apparatus to perform periodic measurements.

5.3 Concluding Remarks

The benefits of multiplexing, miniaturization and cost reduction can be transformational. Breakthrough innovations can arise when cost/size optimization processes eventually surpass a key inflection point. For example, Moore’s law describes the steady cost reduction of transistors which lead to the personal computer revolution. Similar processes exist for megapixels per dollar in image sensors (Hendy’s Law) and the energy density of lithium batteries. In the spirit of these continuous improvement processes, this thesis disseminates new signal processing approaches, device designs and assay approaches. We hope these contributions will hasten the inexorable cost/performance improvements that will transform the field of molecular diagnostics.
In this thesis, we explored two low-cost molecular diagnostic systems having high signal-to-noise ratio and a miniaturized form-factor. We demonstrated fast assay processing times and highly accurate results that are comparable to gold-standard methods. The immediate advantages of such systems include dissemination to LMIC’s where the gold-standard approaches are too expensive or are infrastructure-limited. However, it is also interesting to speculate about second-order effects of such innovations, including (1) future displacement of gold-standards even in high-income countries such as the United States, and (2) novel, non-obvious use-cases that create entirely new opportunities for improvements in healthcare. While there are many reasons to be optimistic about the future of molecular diagnostics, there are also significant hindrances to be cognizant of. In the United States, complex reimbursement protocols and poorly-aligned stakeholder incentives might be a larger impediment to the adoption of molecular diagnostics than technological limitations. The net cost savings and improved patient outcomes of these tests are far from guaranteed, which makes payers such as private insurance and Medicare wary of reimbursing them. Preventative screening is particularly problematic for insurers because its benefit will likely accrue to a competitor as patients churn through different plans. Additionally, physicians may not embrace these diagnostics unless they result in more procedures or more costly procedures being performed. Finally, adopting companion diagnostics for patient selection in clinical trials might is a risky proposition for drug manufacturers because (1) including them generally requires a larger phase II trial population and increased per-patient cost (2) the diagnostic may significantly lower revenue and market share if the treatable population is reduced to a smaller sub-segment [113].

While reducing cost and complexity will not solve all of the problems of adoption, it is imperative to keep costs low in order to drive innovation and experimentation. Molecular diagnostics are likely to flourish in high-value areas such as oncology and autoimmune therapy at first, and later spread to other areas when their price tag drops below some cost-benefit threshold. As the cost of molecular diagnostics decreases and their value is better understood through further research and clinical study, the
promise of widespread personalized medicine via bio-marker analysis may finally come to pass.
Appendix A

ARM NEON Accelerated Fast Walsh-Hadamard Transforms

The following are two C++ implementations of the Fast Walsh-Hadamard Transform, optimized for the ARM-Cortex series of microcontrollers. The first is based on a compact Python implementation available on Wikipedia [114] (reproduced below), while the second is based on a technique described by K.G. Beauchamp [115] that automatically handles sequency re-ordering.

```python
def fwht(a) -> None:
    h = 1
    while h < len(a):
        for i in range(0, len(a), h * 2):
            for j in range(i, i + h):
                x = a[j]
                y = a[j + h]
                a[j] = x + y
                a[j + h] = x - y
        h *= 2
```

The code was designed to process packed 16-bit integer array input, which is a common format used by microcontrollers when periodically sampling ADC data. The
implementations include NEON Intrinsics for the ARM M-Series line of embedded microcontrollers. Sample inputs and outputs are shown in Fig. 4-5. A total of 7 implementations were benchmarked, and the full code is available at the following GitHub Repository:

https://github.com/deganii/fwht-arm

Generating a single sequency waveform such as $W_{(5,8)}$ is done as follows:

\[
\begin{bmatrix}
2047.5_{(DC)} \\
0 \\
0 \\
0 \\
0 \\
2047.5_{(5)} \\
0 \\
0
\end{bmatrix} \xrightarrow{\text{FWHT}} \begin{bmatrix}
4095 \\
0 \\
0 \\
0 \\
4095 \\
4095 \\
4095 \\
0
\end{bmatrix} = W_{(5,8)} \quad (A.1)
\]

Multiple sequency waveforms are also relatively trivial to generate using the FWHT. Below is an example of the $W_{([2,3,6],8)}$:

\[
\begin{bmatrix}
2047.5_{(DC)} \\
0 \\
682.5_{(2)} \\
682.5_{(3)} \\
0 \\
0 \\
682.5_{(6)} \\
0
\end{bmatrix} \xrightarrow{\text{FWHT}} \begin{bmatrix}
4095 \\
2730 \\
1365 \\
0 \\
1365 \\
2730 \\
1365 \\
2730
\end{bmatrix} = W_{(2,8)} + W_{(3,8)} + W_{(6,8)} \quad (A.2)
\]
void fwht_wiki(int16_t *dst, int16_t *src, uint16_t *seq, uint16_t n){
    uint16_t h = 1;
    // first pass: use a sum/difference SIMD
    for (uint32_t* xy = (uint32_t *)(&src[n-2]); xy >= (uint32_t *)src; xy--){
        *xy = __SSAX(*xy, *xy);
        ((int16_t *)xy)[1] = -((int16_t *)xy)[1];
    }
    h <<= 1;
    while(h < n){
        for(int16_t i = 0; i< n; i += h<<1){
            for(int16_t j = i; j< i + h; j+=2){
                uint32_t x1x2 = *((uint32_t*)(&src[j]));
                uint32_t y1y2 = *((uint32_t*)(&src[j+h]));
                *((uint32_t*)(&src[j])) = __SADD16(x1x2, y1y2);
                *((uint32_t*)(&src[j+h])) = __SSUB16(x1x2, y1y2);
            }
        }
        h <<= 1;
    }
    for(uint32_t i = 0; i < n; i++){
        dst[i] = src[seq[i]];
    }
}

inline void swap16(int16_t **a, int16_t **b){
    int16_t *temp = *a;
    *a=*b;
    *b=temp;
}

int16_t* fwht_beauchamp(int16_t *y, int16_t *x, uint16_t N){
    uint16_t S = N/2; // stage
    uint16_t G = N/2; // Number of Groups (N/2)
    uint16_t M = 2; // Number of Members in Each Group
    uint16_t iM2 = 0; // product of pass_index (i) * M * 2

    // First-pass
    for(uint16_t i = 0; i < N; i+=2){
        uint32_t x1x2 = *((uint32_t*)(&x[i]));
        *((uint32_t*)(&y[i])) = __SSAX(x1x2, x1x2);
        y[i+1] = -y[i+1];
    }
    S >>= 1;
    swap16(&x,&y);

    // Subsequent passes
    while(S > 0) {
        for(int i = 0; i < G/2; i++){  
            iM2 = i*M*2;
            for(int j = 0; j < M*2; j+=4){
                uint32_t xl = *((uint32_t*)(x[iM2 + j/2]));
                uint32_t xr = *((uint32_t*)(x[iM2+ M + j/2]));
                uint16_t *yp = (uint16_t*)(y[iM2 + j]);
                uint32_t *yp_32_1 = (uint32_t*)(y[iM2 + j + 1]);
            }
        }
        S >>= 1;
        swap16(&x,&y);
    }
    return y;
}
*yp_32_1 = __SSUB16(xl, xr);
uint32_t add16 = __SADD16(xl, xr);
yp[0] = ((int16_t*)(&add16))[0];
yp[3] = ((int16_t*)(&add16))[1];
}
}
swap16(&x,&y);
G = G/2;
M = M*2;
S >>= 1;
}
return x;
}


Dipanjan Roy Chaudhury. Usaid announces $3 million to support india’s covid-19 efforts, May 2020.


Remila Rezhake, Shang-Ying Hu, Shuang Zhao, Xiao-Qian Xu, Xue-Lian Zhao, Li Zhang, Yan Wang, Xun Zhang, Qin-Jing Pan, You-Lin Qiao, et al. Eight-type human papillomavirus e6/e7 oncoprotein detection as a novel and promis-


