INTEGRATED WAVEFRONT CORRECTION AND BIAS ESTIMATION FOR THE HIGH-CONTRAST IMAGING OF EXOPLANETS

A J Eldorado Riggs

A DISSERTATION
Presented to the Faculty
of Princeton University
in Candidacy for the Degree
of Doctor of Philosophy

Recommended for Acceptance
by the Department of
Mechanical and Aerospace Engineering
Adviser: N. Jeremy Kasdin

May 2016
© Copyright by A J Eldorado Riggs, 2016.
All Rights Reserved
Abstract

Just over two decades ago the first planet outside our solar system was found, and thousands more have been discovered since. Nearly all these exoplanets were indirectly detected by sensing changes in their host stars’ light. However, exoplanets must be directly imaged to determine their atmospheric compositions and the orbital parameters unavailable from only indirect detections. The main challenge of direct imaging is to observe stellar companions much fainter than the star and at small angular separations. Coronagraphy is one method of suppressing stellar diffraction to provide high star-to-planet contrast, but coronagraphs are extremely sensitive to quasi-static aberrations in the optical system. Active correction of the stellar wavefront is performed with deformable mirrors to recover high-contrast regions in the image. Estimation and control of the stellar electric field is performed iteratively in the camera’s focal plane to avoid non-common path aberrations arising from a separate pupil sensor. Estimation can thus be quite time consuming because it requires several high-contrast intensity images per correction iteration.

This thesis focuses on efficient focal plane wavefront correction (FPWC) for coronagraphy. Time is a precious commodity for a space telescope, so there is a strong incentive to reduce the total exposure time required for focal plane wavefront estimation. Much of our work emphasizes faster, more robust estimation via Kalman filtering, which optimally combines prior data with new measurements. The other main contribution of this thesis is a paradigm shift in the use of estimation images. Time for FPWC has generally been considered to be lost overhead, but we demonstrate that estimation images can be used for the detection and characterization of exoplanets and disks. These science targets are incoherent with their host stars, so we developed and implemented an iterated extended Kalman filter (IEKF) for simultaneous estimation of the stellar electric field and the incoherent signal. From simulations and testbed experiments, we report the increased FPWC speed enabled by Kalman filtering and the use of the IEKF for exoplanet detection during FPWC. We discuss the relevance and future directions of this work for planned or proposed coronagraph missions.
Acknowledgements

I first thank my adviser, Jeremy Kasdin. Your continued guidance, encouragement, and enthusiasm have been essential to my professional growth at Princeton. I came into graduate school unsure how to become an engineering scientist, but you have served as a great example for me to follow. From our frequent discussions, I have learned much about estimation, astronomy, and detection. I am grateful for all the opportunities you have provided me, especially working in the High Contrast Imaging Laboratory at Princeton and collaborating with JPL. I also thank the other professors who have been involved in my exams, research, and dissertation over the years: Bob Vanderbei, Mike Littman, Rob Stengel, and David Spergel.

I cannot thank Tyler Groff enough for being my unofficial mentor, first as a fellow graduate student and later as research staff at Princeton. You have been pivotal to my overall development at Princeton and have helped me navigate every stage of graduate school. Thank you for devoting countless hours to teaching me the nuances of operating the lab and helping me truly understand optics and wavefront correction.

It has been a pleasure working on coronagraph optimization with Bob and postdocs Alexis Carlotti and Neil Zimmerman. I learned much from each of you and enjoyed working on these design problems together. To our newest postdoc, Jessica Gersh-Range, I am glad that our times at Princeton overlapped and that you will continue investigating this topic.

I am thankful for all the great graduate students who have been in the group. Hari Subedi and He Sun, I am happy that you are continuing to improve the lab, and it has been fun working with both of you. My office mates have enriched my experience at Princeton with all our random chats about life and work, and I am glad for the time I have had with Dan Sirbu, Elizabeth Young, Mary Anne Limbach, Aaron Lemmer, and Andreas Rousing.

My graduate work was made possible by several grants from NASA. My second and third years were funded by NASA Grants #NNX09AB96G and #NNX12A151G. I owe much to the NASA Space Technology Research Fellowship (NSTRF), which funded my work over the
last two years under NASA Grant #NNX14AM06H. My NSTRF-sponsored internships at the Jet Propulsion Laboratory (JPL) and Space Telescope Science Institute (STScI) greatly improved my research and introduced me to many new collaborators and friends.

I thank all my colleagues at JPL for allowing me to use their state-of-the-art testbed and making me feel welcome during my several months of internships. Stuart Shaklan, thank you for providing guidance and support as my NSTRF mentor. Eric Cady, Brian Kern, and Camilo Mejia Prada, I owe all my results from the High Contrast Imaging Testbed to your patience, training, and troubleshooting help in the last three years. I have learned much from Bala Balasubramanian and Victor White at JPL’s Microdevices Laboratory during our many iterations of coronagraph design and manufacturing. My work has benefited greatly from collaborating with the WFIRST coronagraph optical modeling group. I especially thank John Krist and Bijan Nemati for many insightful discussions about Fourier optics, coronagraphy, and wavefront correction. I am deeply grateful to Stuart, Ilya Poberezhskiy, Feng Zhao, Marissa Rubio, and Ozhen Pananyan for making all my internship stays at JPL run smoothly.

In the last few months, the NSTRF has also allowed me to collaborate with the high-contrast imaging group at STScI. I have enjoyed working with R´em´i Soummer, Mamadou N’Diaye, and Lucie Leboulleux on the wavefront correction code for HiCAT, and I hope our troubleshooting efforts pay off soon. I thank Laurent Pueyo for familiarizing me with exoplanet detection theory and helping me plan out Chapter 6. Neil, you were a gracious host during my visits, I enjoyed spending time with you in Baltimore.

I will miss my many Princeton friends from outside the lab group. I’m glad to have been close with many of my MAE class: Will Scott, Katie Fitch, Scott Dawson, Anthony Degennaro, Steve and Heidi Atkinson, Brendan Andrade, and Matt Plasek. Many thanks to the softball team for four great seasons and remembering still to have fun after our unwilling bump to the A League. Softball has been the highlight of each year and taught me many valuable life lessons, such as how not to get permanently scarred anymore while sliding into base.
I thank my family for a lifetime of love and support that has helped me get to this point in life. My parents Akiyo and Jon and my sister G.G. Comet, thank you for staying close even when I have been so far away geographically. Cleocatra, I appreciated your efforts to type the first draft of this dissertation with random letters.

I am grateful for the my new Chicago family: the Cobbs, Henehans, and many branches thereof. Thank you for having me at all your massive holiday gatherings and for giving me real professional sports teams to root for. Mary Kay, Mike, and Claire Cobb, thank you for accepting me as one of your own.

I dedicate this thesis to my best friend and amazing wife, Katie Riggs. You have been there for me every step of the way and have supported me through all the stressful times. Being with you has made me a better person. You have helped me to find a good work-life balance, which has made me happier overall and more efficient at work. Thank you for continually seeking out new adventures for us and finding all that New Jersey has to offer. I look forward to a lifetime of more adventures with you.

This dissertation carries the number T-3318 in the records of the Department of Mechanical and Aerospace Engineering.
“There can be no thought of finishing, for ‘aiming at the stars,’ both literally and figuratively, is a problem to occupy generations, so that no matter how much progress one makes, there is always the thrill of just beginning.”

-Robert H. Goddard

“To infinity... and beyond!”

-Buzz Lightyear
Contents

Abstract ................................................................. iii
Acknowledgements ................................................... iv
List of Tables .......................................................... xii
List of Figures ......................................................... xiii
List of Symbols ........................................................ xvi
Key Terminology ...................................................... xvii

1 Introduction ......................................................... 1
  1.1 Science Motivation ............................................... 1
  1.2 Overcoming Diffraction ......................................... 4
    1.2.1 Coronagraphy ............................................... 6
    1.2.2 Wavefront Correction ..................................... 9
  1.3 Proposed Coronagraphy Missions ............................. 13
  1.4 Testbeds for Coronagraphy .................................... 15
  1.5 Fourier Optics .................................................. 16
  1.6 Controllability and Observability ........................... 20
    1.6.1 Controlling the Electric Field with Deformable Mirrors ...
    1.6.2 Observability of the Image Plane Electric Field .........
  1.7 Exoplanet Signal Extraction Methods ........................ 25
  1.8 Dissertation Overview ........................................ 29
4.3 Experimental Results from Princeton’s HCIL ............................................. 85
  4.3.1 Experiments without Additional Incoherent Sources ............................. 85
  4.3.2 Experiments in the Presence of a Flat, Incoherent Background ................. 91
  4.3.3 Experiments to Recover a Faint Planet .................................................. 93
4.4 Limitations of Princeton’s HCIL and Proposed Upgrades .............................. 101
  4.4.1 DM Surface Limitations ...................................................................... 101
  4.4.2 DM Placement Limitations .................................................................. 108
  4.4.3 Other HCIL Limitations ..................................................................... 110
4.5 Summary of Results from Princeton’s HCIL ............................................... 111

5 The High Contrast Imaging Testbed at the Jet Propulsion Laboratory 113
  5.1 Layout and Hardware of JPL’s HCIT ......................................................... 114
  5.2 Design of Experiments .......................................................................... 118
  5.3 Simulations of JPL’s HCIT .................................................................... 120
    5.3.1 Full and Compact Simulation Models ................................................ 120
    5.3.2 Simulation Methodology ................................................................... 122
    5.3.3 Controller Tuning ............................................................................ 122
    5.3.4 Sources of Model Error and Measurement Noise ................................ 123
    5.3.5 Optimal Selection of Exposure Time .................................................. 126
  5.4 Experimental Results from JPL’s HCIT ...................................................... 131
    5.4.1 Correction Speed Comparison ............................................................ 131
    5.4.2 Estimate Quality Comparison .............................................................. 132
    5.4.3 Correction Runs with a Shorter Exposure Time ..................................... 137
  5.5 Current Limitations of JPL’s HCIT ............................................................. 138
  5.6 Comparison of Simulation and Lab Results ............................................... 141

6 Signal Detection Using the Bias Estimate ......................................................... 143
  6.1 Observing Scenarios .............................................................................. 144
List of Tables

2.1 Dimensions of variables for the Kalman filter .......................... 59
2.2 Dimensions of matrices for the Kalman filter ........................... 62

3.1 Dimensions of variables for the EKF ...................................... 70
3.2 Dimensions of matrices for the EKF ...................................... 71
3.3 Computational complexity of each estimator ............................... 78
List of Figures

1.1 Exoplanet characteristics and discovery methods .................................. 3
1.2 Contrast vs resolution of the telescope ................................................. 4
1.3 Examples of apodized pupils .............................................................. 6
1.4 Diagram of a Lyot coronagraph .......................................................... 8
1.5 Diagram of an adaptive optics system ................................................. 10
1.6 Coordinates for optical propagation between two planes ...................... 17

2.1 Planes used in the linearization of the electric field ............................ 33

4.1 Diagram of the HCIL optical layout ............................................... 81
4.2 The HCIL shaped pupil mask and its PSF .......................................... 82
4.3 An example laboratory PSF after creating small dark holes .................. 83
4.4 Comparison of correction speed for several estimators ....................... 86
4.5 Comparisons of the average starlight and intensity estimates ............... 89
4.6 Estimated and measured values in the HCIL dark holes ...................... 90
4.7 Final PSFs for the IEKF correction run with zodi ............................... 91
4.8 Correction curves with zodi injected ............................................... 93
4.9 Results from a 2-probe-pair IEKF correction run ................................ 94
4.10 Template PSF for a planet in the HCIL ........................................... 96
4.11 Planet signals recovered via different techniques ............................... 97
4.12 Planet signal metrics after correction ............................................. 98
4.13 Correlation between planet and template PSF after each iteration ........ 98
4.14 Estimated planet contrast after each correction iteration ................... 99
4.15 Measured surfaces for the 1st-generation HCIL Kilo-DMs ................... 102
4.16 Simulated and measured HCIL PSFs ............................................. 103
4.17 HCIL PSFs with small dark holes before and after correction ............... 104
4.18 HCIL PSFs with large dark holes before and after correction ............... 106
4.19 Measured DM surfaces for the new HCIL Kilo-DMs ........................... 107
4.20 Simulated monochromatic HCIL PSFs with DM edge clipping ............. 108
4.21 Simulated 10% broadband HCIL PSFs with DM edge clipping ............. 109

5.1 Ray trace diagram of the optical layout used in JPL’s HCIT .................. 114
5.2 AFTA pupils and their corresponding SPLC apodizer designs ............. 116
5.3 Diagram of the most recent SPLC design for the WFIRST CGI ............. 117
5.4 Images of fabricated SPLC masks used in tests at JPL ....................... 119
5.5 Diagrams of the compact and full optical models .............................. 120
5.6 Example of the empirical line search controller ........................................ 123
5.7 Correction at different NECs for 2-probe-pair estimators .......................... 127
5.8 Correction at different NECs for 1-probe-pair estimators .......................... 129
5.9 HCIT correction curves for 0.5s exposures ............................................. 132
5.10 PSFs and estimates after each 0.5s/exposure correction run ..................... 135
5.11 Average incoherent estimates for each estimator .................................... 137
5.12 HCIT correction curves for 0.3s exposures ........................................... 138
5.13 Histogram of HCIT sCMOS noise ...................................................... 139

6.1 Planet PSF used in the simulations ....................................................... 148
6.2 Correction speed curves from Monte Carlo trials ..................................... 149
6.3 \(3 \times 10^{-10}\) planet PSF correlation values .............................................. 152
6.4 Diagram of ROC curve construction ..................................................... 154
6.5 ROC and AUC curves for first RCDI test ............................................ 155
6.6 \(3 \times 10^{-10}\) planet contrast estimates .................................................. 156
6.7 ROC and AUC curves for second RCDI test ......................................... 157
6.8 \(1 \times 10^{-9}\) planet PSF correlations ..................................................... 158
6.9 \(1 \times 10^{-9}\) planet contrast estimates .................................................. 158
6.10 ROC and AUC curves for third RCDI test ........................................... 159
6.11 \(1 \times 10^{-10}\) planet PSF correlations ..................................................... 160
6.12 \(1 \times 10^{-10}\) planet contrast estimates ................................................ 160

xiv
## List of Symbols

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda$</td>
<td>Wavelength of light</td>
<td>5</td>
</tr>
<tr>
<td>$D$</td>
<td>Telescope or entrance pupil diameter</td>
<td>5</td>
</tr>
<tr>
<td>$E$</td>
<td>Complex-valued electric field</td>
<td>17</td>
</tr>
<tr>
<td>$x, y$</td>
<td>Coordinates in the collimated beam</td>
<td>17</td>
</tr>
<tr>
<td>$i$</td>
<td>Imaginary number</td>
<td>17</td>
</tr>
<tr>
<td>$\Delta z$</td>
<td>Beam propagation distance</td>
<td>17</td>
</tr>
<tr>
<td>$z$</td>
<td>Beam propagation coordinate</td>
<td>17</td>
</tr>
<tr>
<td>$\mathcal{F}{\cdot}$</td>
<td>Fourier transform operator</td>
<td>18</td>
</tr>
<tr>
<td>$\xi, \eta$</td>
<td>Focal plane coordinates</td>
<td>19</td>
</tr>
<tr>
<td>$A$</td>
<td>Nominal complex field at the DM</td>
<td>20</td>
</tr>
<tr>
<td>$\Delta \phi$</td>
<td>Change in phase from DM actuation</td>
<td>20</td>
</tr>
<tr>
<td>$\tilde{E}_0$</td>
<td>Electric field incident upon the DM</td>
<td>33</td>
</tr>
<tr>
<td>$\tilde{E}$</td>
<td>Electric field leaving the DM</td>
<td>33</td>
</tr>
<tr>
<td>$\phi$</td>
<td>Total phase contribution of the actuated DM</td>
<td>33</td>
</tr>
<tr>
<td>$k$</td>
<td>Correction iteration number of FPWC</td>
<td>33</td>
</tr>
<tr>
<td>$f(x, y)$</td>
<td>Two-dimensional influence function of a DM actuator</td>
<td>33</td>
</tr>
<tr>
<td>$\Delta u$</td>
<td>Vector of differential displacement commands to the DM actuators</td>
<td>33</td>
</tr>
<tr>
<td>$q$</td>
<td>DM actuator index</td>
<td>33</td>
</tr>
<tr>
<td>$N_{\text{act}}$</td>
<td>Number of actuators on the DM surface</td>
<td>33</td>
</tr>
<tr>
<td>$C{\cdot}$</td>
<td>Linear operator from the DM to the science camera plane</td>
<td>34</td>
</tr>
<tr>
<td>$m, n$</td>
<td>Discrete coordinates in the image plane</td>
<td>35</td>
</tr>
<tr>
<td>$N_{\text{pix}}$</td>
<td>Number of camera pixels in the dark hole</td>
<td>35</td>
</tr>
<tr>
<td>$G$</td>
<td>Control Jacobian for the dark hole pixels</td>
<td>35</td>
</tr>
<tr>
<td>$N_{DM}$</td>
<td>Number of DMs used for control</td>
<td>35</td>
</tr>
<tr>
<td>$u$</td>
<td>Vector of differential displacement commands to the DM actuators</td>
<td>35</td>
</tr>
<tr>
<td>$J$</td>
<td>Cost function</td>
<td>36</td>
</tr>
<tr>
<td>$\tilde{Q}, \tilde{R}$</td>
<td>Weighting matrices</td>
<td>36</td>
</tr>
<tr>
<td>$I$</td>
<td>Identity matrix</td>
<td>36</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>DM actuation regularization (damping) parameter</td>
<td>36</td>
</tr>
<tr>
<td>$*$</td>
<td>Complex conjugate</td>
<td>36</td>
</tr>
<tr>
<td>$\mathcal{R}{\cdot}$</td>
<td>Real part of a complex value or function</td>
<td>36</td>
</tr>
<tr>
<td>$\bar{T}$</td>
<td>Average contrast in the dark hole</td>
<td>38</td>
</tr>
<tr>
<td>$C$</td>
<td>Contrast target</td>
<td>38</td>
</tr>
<tr>
<td>$\mu$</td>
<td>Lagrange multiplier</td>
<td>39</td>
</tr>
<tr>
<td>$j$</td>
<td>Probe number index</td>
<td>41</td>
</tr>
<tr>
<td>Symbol</td>
<td>Description</td>
<td>Page</td>
</tr>
<tr>
<td>--------</td>
<td>------------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>$p$</td>
<td>Complex-valued probe electric field in the image plane</td>
<td>41</td>
</tr>
<tr>
<td>$I$</td>
<td>Measured contrast at an image plane pixel</td>
<td>41</td>
</tr>
<tr>
<td>$n$</td>
<td>Measurement noise vector</td>
<td>41</td>
</tr>
<tr>
<td>$\mathcal{I}{\cdot}$</td>
<td>Imaginary part of a complex value or function</td>
<td>42</td>
</tr>
<tr>
<td>$N_{pp}$</td>
<td>Number of probe pairs</td>
<td>42</td>
</tr>
<tr>
<td>$z$</td>
<td>Noisy measurement vector</td>
<td>42</td>
</tr>
<tr>
<td>$H$</td>
<td>Observation matrix</td>
<td>42</td>
</tr>
<tr>
<td>$x$</td>
<td>Real-valued state vector</td>
<td>42</td>
</tr>
<tr>
<td>$R$</td>
<td>Measurement noise covariance matrix</td>
<td>43</td>
</tr>
<tr>
<td>$E[\cdot]$</td>
<td>Expectation value</td>
<td>43</td>
</tr>
<tr>
<td>$\theta_j$</td>
<td>Phase of the image-plane probe</td>
<td>47</td>
</tr>
<tr>
<td>$E_R, E_I$</td>
<td>Real and imaginary parts of the E-field, respectively</td>
<td>47</td>
</tr>
<tr>
<td>$\sigma^2$</td>
<td>Variance</td>
<td>48</td>
</tr>
<tr>
<td>$I_{pk}$</td>
<td>Peak measured counts of the star in the image</td>
<td>49</td>
</tr>
<tr>
<td>$F_{pk}$</td>
<td>Peak flux of the star in the image</td>
<td>49</td>
</tr>
<tr>
<td>$t$</td>
<td>Period of time</td>
<td>49</td>
</tr>
<tr>
<td>NEC</td>
<td>Noise-equivalent contrast</td>
<td>51</td>
</tr>
<tr>
<td>$N_{exp}$</td>
<td>Number of exposures averaged</td>
<td>54</td>
</tr>
<tr>
<td>$Z$</td>
<td>Counts from zodiacal or exozodiacal dust</td>
<td>54</td>
</tr>
<tr>
<td>$D_c$</td>
<td>Counts from dark current</td>
<td>54</td>
</tr>
<tr>
<td>$\langle\cdot,\cdot\rangle$</td>
<td>Complex inner product</td>
<td>56</td>
</tr>
<tr>
<td>$\Phi$</td>
<td>State transition matrix</td>
<td>58</td>
</tr>
<tr>
<td>$\Gamma$</td>
<td>Real-valued control Jacobian</td>
<td>58</td>
</tr>
<tr>
<td>$\Lambda$</td>
<td>Real-valued disturbance Jacobian</td>
<td>58</td>
</tr>
<tr>
<td>$w$</td>
<td>Disturbance vector</td>
<td>58</td>
</tr>
<tr>
<td>$P$</td>
<td>State covariance matrix</td>
<td>60</td>
</tr>
<tr>
<td>$Q$</td>
<td>Process noise covariance matrix</td>
<td>60</td>
</tr>
<tr>
<td>$K$</td>
<td>Kalman gain matrix</td>
<td>60</td>
</tr>
<tr>
<td>$h(x)$</td>
<td>Nonlinear measurement function</td>
<td>66</td>
</tr>
<tr>
<td>$N_z$</td>
<td>Number of images in the EKF measurement vector</td>
<td>70</td>
</tr>
<tr>
<td>$N_{it}$</td>
<td>Number of EKF iterations</td>
<td>74</td>
</tr>
<tr>
<td>$C$</td>
<td>Normalized two-dimensional PSF correlation value</td>
<td>95</td>
</tr>
</tbody>
</table>
Key Terminology

- **High Contrast**: The star-to-planet flux ratio being large. By convention in the literature, the inverse ratio is reported; thus, $10^{-10}$ is higher contrast than $10^{-9}$. In this thesis we generally use the terminology “better” and “worse” contrast instead of “higher” and “lower” contrast, respectively, to reduce confusion.

- **$\lambda/D$**: Angular unit used in the image plane. One $\lambda/D$ is defined as the full-width at half maximum of the PSF’s main lobe at wavelength $\lambda$ for a circular pupil of diameter $D$.

- **Focal Plane Wavefront Correction (FPWC)**: The estimation and control of the stellar electric field in the camera’s focal plane.

- **Signal-to-Noise Ratio (SNR) vs Noise-Equivalent Contrast (NEC)**: SNR is the ratio of the mean signal being measured over the total standard deviation from measurement noise. NEC is the minimum intensity level of the starlight that the estimator can accurately sense.

- **Search Area vs Dark Hole**: The search area is the high-contrast region in the image produced in an ideal coronagraph design. The dark hole is the subset of the search area in which wavefront correction is performed.

- **Normalized Intensity (NI) vs Contrast**: NI is the measure of the starlight suppression in the dark hole. NI is the ratio of the stellar intensity in the dark hole to the unocculted peak of the stellar PSF. Contrast, as defined in this thesis, is the NI divided by the off-axis PSF suppression at an image location. Contrast equals NI for apodizer-only coronagraphs, but contrast is worse than NI for Lyot-type coronagraphs because the focal plane mask and Lyot stop block some of the off-axis light.

- **Raw Contrast vs Contrast**: “Raw contrast” is the measured contrast directly from an image (after converting from NI). “Contrast” typically refers to the effective contrast after using post-processing techniques to remove even more residual starlight. Except when specified in sections describing image processing, we mean raw, measured contrast when we use the term contrast in this thesis.
Chapter 1

Introduction

1.1 Science Motivation

Just 25 years ago the only known planets were those within the solar system. With the
first discoveries of an extrasolar planet, or exoplanet, around a pulsar in 1992[1] and around
a main-sequence star in 1995[2], the field of exoplanet science erupted into existence. As-
tronomers to date have confirmed the discovery of 1,935 exoplanets and identified another
3,701 possible candidates.[3] With thousands of exoplanets to study, scientists have be-
gun to develop theories of planetary formation, abundance, and evolution with statistical
certainty.[4]

Nearly all known exoplanets were indirectly detected from changes in their host stars’
light. The two most successful detection methods have been radial velocity (RV, also called
Doppler spectroscopy) and transit photometry. The RV method identifies the periodic
Doppler shift of a star’s spectrum due to a companion. The Doppler shift is more prominent
for planetary orbits with small inclinations to our line of sight. RV is a popular method
because it can be used on any telescope with a high-resolution spectrograph, but it is diffi-
cult to perform surveys because spectra can usually be obtained for only one star at a time.
The RV technique allows determination of a planet’s orbital period, orbital eccentricity, and
mass times the sine of the unknown orbital inclination.

Transit photometry[5], the other main detection method, first detected a new exoplanet in 2002[6] and has since found the most new planets. The transit technique measures the periodic dip in a star’s brightness from an exoplanet passing through the line of sight, so it requires a nearly edge-on orbit. Transit observations are easier for a survey mission because numerous stars can be imaged simultaneously, but the brief transit events also require continuous observation. The Kepler space telescope[7, 8], which monitored 170,000 stars in a single patch of sky for four years, yielded over half of all confirmed exoplanet detections and all the potential detections.[3] The transit method allows determination of an exoplanet’s orbital period and its relative radius compared to the host star. For a few nearby transiting exoplanets, their atmospheric spectra have been obtained from comparing the stellar spectra during and in between transit events.[9]

Besides the few cases where transit spectroscopy is possible, direct imaging of exoplanets is necessary to obtain spectra of their atmospheres. Direct imaging of exoplanets allows determination of more orbital parameters such as inclination when there is data from another detection method. Astrometry and direct imaging are the only methods capable of discovering and characterizing exoplanets with face-on orbits, but to date only one planet has been discovered via astrometric motion of its host star.[10] So far only about two dozen exoplanets have been directly imaged.[3] (Several dozen have been imaged but have large enough masses to be classified as brown dwarf stars.) Exoplanets are much fainter than their host stars and have small angular separations from them, so direct imaging requires a small field of view and long integration times. These constraints make direct imaging surveys time consuming and explain the comparatively low yield of exoplanets from this method.

Figure 1.1 shows the distribution of detected exoplanets in terms of estimated mass versus separation from the host star. Since exoplanet detection is a young field, these results represent the capabilities and selection biases of the first generation of instruments.

The RV method is biased toward finding more massive planets with smaller orbital sep-
Figure 1.1: Estimated mass vs semi-major axis for the exoplanets with confirmed detections. Not all transit detections have known masses. The planets from our solar system are included as reference points. Data were obtained in February 2015 from the exoplanets.eu catalog.

...arations, and the transit method best detects larger planets with short orbital periods. (By Kepler’s third law of planetary motion, a shorter period corresponds to a smaller orbital separation.) RV and transit surveys have been ongoing long enough to start finding exoplanets with years-long orbits. Direct imaging is most sensitive to large, bright planets with wide orbital separations. Planets are brightest when they are young and self-luminous in the infrared, and only young exoplanets have been imaged thus far. The lack of known planets with low mass or with high mass and wide separations is likely just selection bias from our current detection methods. Figure 1.1 shows that our detection methods are not yet sensitive to analogs of our solar system planets other than Jupiter.

Direct imaging is the prime candidate for detecting and characterizing the first solar system analogs of gas giants around other stars. Unlike other methods, direct imaging can detect and spectrally characterize an exoplanet from observations at a single point in time.
Monitoring transit events or tracking RV spectra for years or decades to confirm a detection is extremely challenging and time consuming. For earth and Venus-like exoplanets, the transit and RV methods are likely to make the first discoveries. Once earth-mass planets are found in similar orbits around other stars, scientists will need spectra from direct imaging to determine if these worlds are habitable, or perhaps even inhabited. To characterize the first exo-earth spectrally, we will need further technological advancements implemented on new observatories.

1.2 Overcoming Diffraction

The key to directly imaging exoplanets is overcoming the diffraction of starlight. Figure 1.2 shows a radial slice of the point-spread functions (PSFs) for a circular, Hubble-sized (i.e., 2.4m-diameter) telescope. The solid line shows the PSF of the Sun and the dotted line shows

![Graph showing radial slices of the PSFs for a circular-aperture telescope. The PSFs of companion objects are shown at a separation of 2.1 λ/D, equivalent to 1 astronomical unit as observed by a 2.4m telescope from 10 parsecs away in 550nm light. A companion star of equal brightness (dotted line) is easily resolved. An earth-like exoplanet (dash-dotted line) 10^{-10} times fainter than host star is technically resolvable, but the wings of the host star’s PSF (solid line) are 10^8 times brighter than the planet at its location.

Figure 1.2: Radial slices of the PSFs for a circular-aperture telescope. The PSFs of companion objects are shown at a separation of 2.1 \( \frac{\lambda}{D} \), equivalent to 1 astronomical unit as observed by a 2.4m telescope from 10 parsecs away in 550nm light. A companion star of equal brightness (dotted line) is easily resolved. An earth-like exoplanet (dash-dotted line) \( 10^{-10} \) times fainter than host star is technically resolvable, but the wings of the host star’s PSF (solid line) are \( 10^8 \) times brighter than the planet at its location.

the PSF of the earth if it were equally bright and viewed from 10 parsecs away in 550nm
light. The main lobes of the Sun and bright-earth are clearly separated, so the Hubble Space Telescope is technically able to resolve earth-like planets. However, the earth is actually about $10^{-10}$ times dimmer than the Sun in visible light[11] as indicated by the dash-dotted line. The wings of the solar PSF are $\approx 10^8$ times brighter than the earth at its location, so the planet signal would be completely lost in the photon shot noise of the sunlight. Thus, the main limitation with direct imaging is not telescope resolution but the high star-to-planet contrast ratio.

Several concepts have been proposed to mitigate diffraction and decrease the measured star-to-planet contrast for direct imaging. The earth-sun contrast ratio is more reasonable at $\approx 10^{-7}$ in mid-infrared light versus $10^{-10}$ in visible light,[11] but a $\approx 10$ times larger aperture would be necessary to obtain the same resolution at those longer wavelengths. A nulling interferometer[12] mission would fly several large telescopes in formation to obtain that long baseline and destructively interfere mid-infrared starlight. This concept is no longer under active consideration, and most currently proposed missions plan on visible-light imaging. An external starshade[13, 14] suppresses starlight before it ever reaches the telescope and requires formation flying of the two spacecraft. Alternatively, the light from a single telescope can be interfered with itself in a visible nulling coronagraph.[15] In this thesis, we focus on the use of an imaging coronagraph, which suppresses starlight after it enters the telescope with a series of masks, stops, apodizers, and/or specialized mirrors.

There are two main challenges to high contrast imaging with a coronagraph. The first is designing a coronagraph to suppress diffracted starlight in a region of high contrast, called the search area, without blocking off-axis light from planets and disks. The second is mitigating the optical aberrations that diffract starlight back into the search area of the coronagraph. The corrected sub-region of the search area is referred to as the dark hole.
1.2.1 Coronagraphy

Coronagraphs change the amplitude and/or phase of the starlight at one or more planes in the optical train. There are two classes of coronagraphs: those that create contrast by reshaping the PSF (apodizers), and those that create contrast by blocking the starlight (Lyot-type coronagraphs).

The reshaping of the PSF to suppress its wings is called apodization, which literally means to “remove the feet” of the PSF. Figure 1.3 shows the classical Airy pattern PSF from a circular aperture and two examples of apodized pupils and their PSFs. Apodization

![Pupil and Image Examples](image.png)

Figure 1.3: Examples of (top row) pupils and (bottom row) their PSFs to show the effects of apodization. (Left) A circular aperture contains infinite spatial frequencies and produces the Airy pattern as its PSF. (Middle) A smoothly apodized pupil concentrates most of the starlight within the core of the PSF. (Right) A shaped pupil uses binary apodization to suppress diffraction in a finite region of the PSF. The pupils are plotted on a linear scale and the PSFs are plotted on a log\(_{10}\) scale.

is generally considered to be a modification of the pupil amplitude, but the pupil phase can be modified instead of or in addition to the amplitude to alter the PSF.[16, 17, 18, 19] Since the image plane PSF is the power spectrum of spatial frequencies in the pupil plane, classical apodization works by smoothing out the sharp edges of the aperture and leaving only low
spatial frequency content in the pupil. The wings of the PSF in Fig. 1.3 (middle column) are thus suppressed out to infinity at the expense of reduced pupil transmission and lower resolution (i.e., a wider PSF core) in the image. Slepian[20] first derived that the optimal apodization to concentrate light maximally for a circular aperture is a prolate spheroidal wavefunction. Partially transmissive materials are all chromatic, so smooth apodizations have instead been manufactured with intermittently-spaced, opaque “microdots” to give the correct transmission profile on average in a region.[21, 22] Phase-induced amplitude apodization (PIAA)[23] uses two highly aspheric mirrors to remap the uniform pupil into a smoothly apodized one with nearly no throughput loss. Our group at Princeton developed the shaped pupil (SP) coronagraph, which uses binary amplitude apodization (regions are either fully transmissive or fully opaque) to generate high contrast.[24, 25, 26, 27, 28, 29, 30] Since the pupil features of the SP are sharp, the high contrast regions are generated by destructive interference instead of by smoothed pupil features like in classical apodization. The search area of the SP is thus confined to a localized region around the star, as shown with the example SP in Fig. 1.3 (right column). Freestanding SP designs have been etched out of silicon wafers as transmissive masks, and non-freestanding designs have been manufactured on silicon substrates as reflective masks by Balasubramanian et al.[31]

The first coronagraph type was named after Bernard Lyot, who invented it to block the sun and view the solar corona.[32] Figure 1.4 demonstrates the principles of a standard Lyot coronagraph. An opaque spot in the focal plane first blocks the core of the stellar PSF. In the next pupil, another mask called the Lyot stop blocks much of the remaining starlight diffracted by the focal plane mask (FPM). In the final image, the starlight is suppressed by 2-3 orders of magnitude while off-axis light sources pass through the optical system with little transmission loss.

Numerous variations of the Lyot coronagraph have been invented to generate high contrast. All use a Lyot stop and differ in their choice of FPM. Instead of a hard-edged spot, Kuchner and Traub[34] designed a band-limited amplitude profile for the FPM. Many
Figure 1.4: Diagram of a Lyot coronagraph. At the first focal plane, an opaque spot occults the core of the stellar PSF. Residual starlight diffracts around the focal plane mask and is mostly blocked by a Lyot stop at the next pupil plane. The starlight is suppressed by a few orders of magnitude in the final image plane. Intensities are plotted on a log_{10} scale. Based on a figure by Oppenheimer and Hinkley.[33]

Others[35, 36, 37, 38, 39, 40, 41, 42, 43] designed different types of phase-altering FPMs. Moody and Trauger[44] optimized both phase and amplitude for the FPM in the hybrid Lyot coronagraph (HLC). Recent efforts have combined the methods of apodization and Lyot coronagraphy for improved performance with obstructed telescope apertures.[45, 46, 47, 48, 49, 50, 51] At Princeton, Zimmerman, Riggs, et al.[50] developed the shaped pupil Lyot coronagraph (SPLC), which uses an SP to apodize the beam in a Lyot-coronagraph configuration. In this thesis, we perform simulations and lab experiments with SPs for open apertures and with SPLCs for the obstructed aperture of a proposed space mission.
1.2.2 Wavefront Correction

Every coronagraphic system requires wavefront correction to mitigate optical aberrations and recover high contrast. The wavefront is defined as the phase of the coherent beam. Wavefront correction is subdivided into three main categories based on the timescales and sources of the aberrations. First, adaptive optics (AO) compensates for atmospheric turbulence at timescales on the order of 1 millisecond. Second, low-order wavefront sensing and control (LOWFSC) cancels phase aberrations from observatory pointing jitter and mechanical deformation of the telescope on timescales of 0.1 to several seconds. Finally, focal plane wavefront correction (FPWC) mitigates quasi-static phase and amplitude aberrations from imperfections in the optics themselves at timescales of hours and longer. In this thesis, we primarily investigate FPWC and note that these different regimes are separable if the uncorrected residuals from the faster correction loops are sufficiently small.

Adaptive Optics

The purpose of AO is to cancel out atmospheric turbulence, the largest source of aberrations in ground-based astronomy. Each layer of the atmosphere introduces its own phase aberrations to the stellar wavefront, and the total phase aberrations are hundreds of nanometers peak-to-valley. This high amount of aberrations smears out the full width at half maximum (FWHM) of a point source’s image, thereby degrading the resolution of the telescope. An 8-m ground telescope observing at a wavelength of 1.2\,\mu m gives a theoretical FWHM of 0.04 arcseconds for a point source, but even the best nights at high-altitude observatories have a diffraction limit of \geq 0.4 arcseconds. AO is used to mitigate atmospheric turbulence and bring the resolution of the telescope closer to its theoretical limit.

Depending on the wind speed, the aberration timescale is on the order of a millisecond. The AO system must sense and correct the aberrations on the same short timescale to be effective. The deformable mirrors (DMs) used for AO have response times on the order of 10-100 times faster than the aberrations and are thus not a major concern. The main hurdle
is accurately sensing the stellar wavefront in such a short time. Figure 1.5 diagrams the key components in an AO system. Typically, a beamsplitter sends a portion of the starlight to a dedicated wavefront sensor (WFS). Most AO systems use a dichroic beamsplitter to re-direct visible light to the WFS and transmit the near-infrared light to the science instrument. The WFS can be fast because it uses all of the visible starlight and measures only the phase (or phase gradient).

In the past few years, several large (5-10m diameter) ground-based telescopes have implemented coronagraphs behind their AO systems to achieve moderate contrast levels. These coronagraphs are still limited to worse than about $10^{-5}$ raw (i.e., measured) contrast because of the uncorrected AO residuals. Future extremely large (20-40m diameter) telescopes plan to use extreme AO (ExAO) to cancel atmospheric turbulence even more accurately and achieve raw contrast on the order of $10^{-6}$ in coronagraphic images. The main challenges of ExAO are to sense the wavefront accurately enough, to minimize the errors in open-loop actuation of the DMs, and to minimize the time delay between the
sensing of the wavefront and the commanding of the DMs.[61]

Low-Order Wavefront Sensing and Control

The stellar electric field must be held constant at the science camera during the minutes-, hours-, or even days-long exposure times required for high-contrast imaging. After atmospheric turbulence, the other significant dynamic aberrations are from pointing jitter and mechanical deformation of the observatory. LOWFSC is the correction of these medium-timescale phase aberrations with magnitudes on the orders of 10 picometers to a few nanometers. Current AO systems have phase residuals on the order of tens of nanometers, so LOWFSC will be much more useful for future ExAO systems or space telescopes. LOWFSC has been implemented on only a few telescopes to date,[62, 63] not counting simple tip/tilt correction systems. It is being heavily investigated for future ground and space based observatories proposed to have high-contrast coronagraphs.[64, 65]

The term “low-order” refers to low-order Zernike polynomials, which correspond to many of the pure aberrations encountered in a circular pupil. In particular, the main Zernike modes sensed are tip, tilt, defocus, astigmatism, coma, trefoil, and spherical. Tip and tilt arise from pointing errors of the entire telescope at timescales on the orders of 0.1-1s. Defocus stems mainly from the movement of the structure connecting the primary and secondary mirrors. Most of the other modes arise from thermal warping of the telescope primary mirror and have timescales of approximately 1s or slower.

The schemes for low-order wavefront sensing (LOWFS) utilize differential measurements of starlight rejected by the coronagraph. Most methods use the starlight reflected off an opaque focal plane mask,[66, 67, 68] and one type uses the light reflected off the Lyot stop for coronagraphs with focal plane phase masks.[63]
Focal Plane Wavefront Correction

If ExAO and LOWFSC have stabilized the stellar wavefront, the aberrations in the optics themselves must be corrected to obtain high contrast. It is impossible to manufacture mirrors with the surface flatness and reflectivity uniformity required to achieve better than $\approx 10^{-7}$ contrast without active correction.[69] Coronagraphs and DMs, both produced in small batches, are prone to manufacturing defects and introduce even more aberrations. The main goal of quasi-static wavefront correction is thus to mitigate the effects of manufacturing errors and recover regions of high contrast in the image. The aberrations are called quasi-static instead of static because mechanical and thermal stresses on the coronagraph instrument cause slow changes to the beam path on the timescales of hours, days, or longer.

The word “correction” is somewhat of a misnomer in this case. Unlike AO which tries to conjugate atmospheric turbulence or LOWFSC which tries to drive the wavefront back to an initial setting, quasi-static wavefront correction does not in general drive the electric field to the coronagraph’s design value. Rather, it uses the DMs to minimize the electric field in the dark hole and can sometimes reach contrasts below the design level of the coronagraph. Minimizing the phase error in the pupil plane reduces the net aberrations in the focal plane but doesn’t guarantee high contrast in the dark hole, the only region where we need the starlight to be suppressed. Pupil-plane phase conjugation also does not mitigate amplitude aberrations, which typically produce aberrations in the $10^{-5}$ to $10^{-7}$ contrast range. Another important observation is that non-uniform pupil phase and amplitude are permissible as long as the starlight destructively interferes in the dark hole. For these reasons, we control the electric field directly in the focal plane instead of the phase in the pupil plane. Since we control directly in the focal plane, we perform estimation in that plane as well. The estimation and suppression of quasi-static speckles is thus called focal plane wavefront correction (FPWC). We delve into more details of FPWC in §1.6 and Chapter 2.

Several current ground instruments have an additional calibration system to measure quasi-static aberrations in the non-common path between the AO’s WFS and the corona-
These calibration systems use the light rejected by the FPM to measure the quasi-static aberrations upstream of the FPM. They ignore the aberrations after the FPM, however, so FPWC is still necessary to reach better contrast.

**Deformable Mirrors**

Wavefront correction systems generally use deformable mirrors (DMs) to control the wavefront. Spatial light modulators offer higher actuator density at lower cost, but they are not used because they are chromatic and can control only a single polarization of light. DMs are available with segmented or continuous facesheets, and the continuous facesheet ones are preferred for high-contrast imaging because they do not introduce extra diffraction off segment edges. Bimorph DMs with actuators parallel to the facesheet exist, but the DMs most commonly used for wavefront correction have actuators perpendicular to the facesheet.

Boston Micromachines Corporation (BMC) and AOA Xinetics are the two primary manufacturers of DMs for astronomical wavefront correction. BMC specializes in microelectromechanical systems (MEMS) DMs[72] made using techniques from the silicon semiconductor industry. BMC DMs are popular because they are less expensive per actuator, have zero-hysteresis capacitive actuation, and are smaller with inter-actuator pitches of 300-400 µm. Xinetics DMs have electrostrictive actuators composed of lead magnesium niobate (PMN), which compresses with an applied voltage. The PMN actuators have temperature-dependent hysteresis, which reaches its lowest value of just under 1% near room temperature.[73] Xinetics DMs are larger because of their minimum inter-actuator pitch of 1mm. They have the advantage of smoother facesheets (after polishing) and higher preparedness for use in space.

### 1.3 Proposed Coronagraphy Missions

To date no high-contrast coronagraph has flown in space. The James Webb Space Telescope (JWST), launching in 2018, will carry the first sophisticated coronagraphs into space: a
suite of band-limited Lyot and four-quadrant phase mask coronagraphs. \cite{74, 75} The JWST coronagraphs are designed to achieve moderate contrast levels because they will be used without DMs for wavefront correction.

The National Aeronautics and Space Administration’s (NASA’s) first proposed space mission to image and spectrally characterize earth-like exoplanets was the Terrestrial Planet Finder (TPF). TPF consisted of two different mission concepts, one with a single 4-8m monolithic telescope and coronagraph (TPF-C)\cite{76} and one with four telescopes acting as an interferometer (TPF-I).\cite{77} Neither was funded past the study phase, but the technology developments for the TPF-C laid the groundwork and framework for future exoplanet imaging missions. Most recently, NASA has studied two mission concepts for launch in the next decade. The three-year Exo-C mission\cite{64} would use a single-purpose, 1.5m unobscured telescope for directly imaging 20 RV-detected gas giant exoplanets and spectrally characterizing at least half of them. The Wide-Field Infrared Survey Telescope (WFIRST) is a concept using the obscured-aperture, 2.4-m Astrophysics Focused Telescope Assets (AFTA) telescope. The WFIRST Coronagraph Instrument would have one year as a secondary payload to image about 20 RV-detected gas giant exoplanets and obtain spectra from several of them.\cite{78}

The first space mission for directly imaging terrestrial exoplanets will fly in the 2030s at the earliest. Because of the large aperture required for high angular resolution and efficient light collection, all options being considered by NASA would be multi-billion dollar flagship missions. One mission concept, the Habitable Exoplanet Imaging Mission (HabEx), might have a monolithic ≥4m primary mirror, but most of the other concepts would use segmented 8-16m primary mirrors. These segmented-aperture telescopes would be multi-purpose facilities and are under study as the Advanced Technology Large Aperture Space Telescope (ATLAST)\cite{79}, High-Definition Space Telescope (HDST),\cite{80} and Large Ultraviolet Optical Infrared (LUVOIR) telescope.\cite{80}
1.4 Testbeds for Coronagraphy

For this thesis, we performed experiments in two laboratories. We first ran our proof-of-concept tests in Princeton’s High Contrast Imaging Laboratory (HCIL). We then verified our algorithms at higher contrast at the state-of-the-art High Contrast Imaging Testbed (HCIT) at the Jet Propulsion Laboratory (JPL).

High Contrast Imaging Laboratory (HCIL) at Princeton

Since its creation in the early 2000s, the HCIL at Princeton has served two main purposes. The first was to demonstrate in hardware the viability of the shaped pupil (SP) coronagraph to generate high contrast in an image. The SP has been validated after numerous laboratory demonstrations of both transmissive, free-standing SPs at Princeton[81, 82, 83, 84] and JPL [85, 86], and of reflective, non-free-standing SPs at JPL.[87, 88] The second, ongoing purpose of the HCIL is rapid testing of algorithms for FPWC such as the stroke minimization controller by Pueyo et al., [89] the pair-wise Kalman filter by Groff et al.,[83], and the pair-wise extended Kalman filter presented in this thesis.[84] The HCIL’s simple optical layout and constant availability enable fast implementation and testing of new algorithms at moderate ($10^{-6}$ to $10^{-7}$) contrast levels. After validating the algorithms in the HCIL, we often export them to the HCIT at JPL for demonstrations at high ($10^{-8}$ to $10^{-10}$) contrast.

High Contrast Imaging Testbed (HCIT) at JPL

The HCIT at JPL advances technology for the direct imaging of exoplanets. With high-quality optics in two thermally-controlled vacuum chambers, the HCIT provides the best available approximation for a space-based coronagraph instrument. The HCIT focuses on increasing the test-readiness level (TRL, defined by NASA as the preparedness level of a technology for a space mission) of coronagraph hardware, DMs, and wavefront correction algorithms. Through our continued collaboration with JPL, we have been able to validate at high contrast our SP designs and FPWC algorithms.
Over the past two decades, the HCIT has validated at high contrast the open-aperture designs for the band-limited Lyot coronagraph,[90], transmissive SP coronagraph, [85, 86] vector vortex coronagraph,[91] and PIAA coronagraph.[92] In the past two years, the HCIT has focused on testing the coronagraphs selected for the obscured aperture of WFIRST.[50, 87, 88, 93, 94, 51] Optical model validation and correction algorithm development have become higher priorities for the HCIT because of the possibility of WFIRST launching in under a decade.[95, 84]

1.5 Fourier Optics

Here we briefly review the analytical approximations and numerical tools for modeling optical propagation in a coronagraphic system. Maxwell’s equations fully describe the propagation of electromagnetic radiation but are far too cumbersome for use in end-to-end modeling of FPWC. The vectorial nature of light can be safely ignored as long as there are no wavelength-scale apertures (which act as waveguides) or interactions of the light with mask materials.[96] The next concern is polarization within the optical system, which occurs from non-normal reflections. It becomes an issue for coronagraphs when the light splits into two perpendicular polarizations that see different sets of aberrations, in which case FPWC cannot compensate well for both simultaneously. Breckenridge et al.[97] describe this issue in general, and Krist et al.[95] discuss this problem in the context of the WFIRST CGI. At present there is no way to actively control the polarization in the coronagraph instrument. Instead, the current approach is to design the optical system to minimize polarization by using suitable coatings on the mirrors and keeping angles of incidence small. If polarization is an issue for the targeted contrast level, then the only recourse is to correct and image the starlight in only one polarization at a time.

The propagation of light without material interactions is called scalar diffraction and is described by the Rayleigh-Sommerfeld diffraction integral.[98] The Rayleigh-Sommerfeld
formula is intractable for even modern computers, so the paraxial approximation is used to simplify it into the Huygens-Fresnel integral. Since the stellar wavefront is essentially flat across a meters-wide telescope after propagating for lightyears, the assumption of parallel rays is completely safe. Even so, the Huygens-Fresnel integral is still intractable except for the special, limited-use case of radial, one-dimensional propagation.[99] By approximating the optical path length to first order for each ray, one obtains the Fresnel approximation. It is only valid for narrow fields of view; fortunately, the correctable region of the coronagraphic image is typically less than an arcsecond across. Krist et al.[100] performed an extensive study and concluded that the Fresnel approximation is accurate to better than $10^{-10}$ contrast in simulation for nearly all coronagraphs. The main exception is the PIAA coronagraph, for which one of the more accurate approximations must be used to propagate between the two highly curved pupil-remapping lenses or mirrors.

The analytical form of the Fresnel approximation between the two planes in Fig. 1.6 is

$$E_2(x_2, y_2) = \frac{e^{i \frac{2\pi}{\lambda} \Delta z}}{i\lambda \Delta z} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} E_1(x_1, y_1) e^{i \frac{\pi}{\lambda\Delta z} [(x_2-x_1)^2+(y_2-y_1)^2]} dx_1 dy_1,$$

(1.5.1)

where $E$ is the electric field, $\Delta z$ is the distance between planes, and $\lambda$ is the wavelength of light.

Figure 1.6: Coordinates for the two planes in Fresnel propagation.
Two different formulations are used to compute the Fresnel integral; they are analyti-
cally identical but numerically produce different results because of discrete sampling. For
plane-to-plane propagation in a collimated beam, the angular spectrum version of the Fres-
nel approximation is used. The angular spectrum is obtained by re-writing Eq. 1.5.1 as a
convolution of the original electric field \( E_1 \) with a quadratic phase factor,

\[
E_2(x_2, y_2) = \int_{-\infty}^{+\infty} E_1(x_1, y_1) h(x_2 - x_1, y_2 - y_1)dx_1dy_1
= \mathcal{F}^{-1}\left\{ \mathcal{F}\{E_1(x_1, y_1)\}e^{-i\pi\lambda\Delta z(x_2^2+y_2^2)} \right\},
\tag{1.5.2}
\]

where \( \mathcal{F} \) is the Fourier transform (FT) operator, \( \mathcal{F}^{-1} \) is the inverse FT, and the convolution
kernel is

\[
h(x_2, y_2) = e^{i\frac{2\pi}{\lambda}\Delta z} e^{i\frac{\pi}{\lambda\Delta z}(x_2^2+y_2^2)}. \tag{1.5.3}
\]

For propagation in converging or diverging beams, Eq. 1.5.1 is re-written as a FT of the
original electric field times quadratic phase factors inside and outside the integral,

\[
E_2(x_2, y_2) = \frac{e^{i\frac{2\pi}{\lambda}\Delta z}}{i\lambda\Delta z} e^{i\frac{\pi}{\lambda\Delta z}(x_2^2+y_2^2)} \int_{-\infty}^{+\infty} E_1(x_1, y_1) e^{i\frac{\pi}{\lambda\Delta z}(x_1^2+y_1^2)}e^{-i\frac{2\pi}{\lambda\Delta z}(x_\xi y_\eta)}dx_1dy_1
= \frac{e^{i\frac{2\pi}{\lambda}\Delta z}}{i\lambda\Delta z} e^{i\frac{\pi}{\lambda\Delta z}(x_2^2+y_2^2)} \mathcal{F}\{E_1(x_1, y_1)e^{i\frac{\pi}{\lambda\Delta z}(x_1^2+y_1^2)}\} \tag{1.5.4}
\]

This formulation is typically the one referred to as Fresnel propagation. By propagating
only between parabolic reference surfaces or between a parabolic reference and the focal
plane, we can drop the quadratic phase term outside the integral in Eq. 1.5.4. Several
sources\[101, 102, 103\] detail the proper calculation of the Fresnel and angular spectrum
propagations to avoid or mitigate numerical errors. Many (but not all) of the potential
numerical pitfalls are from using fast Fourier transforms (FFTs),[104, 105] which have more
errors than brute-force discrete Fourier transforms (DFTs) but can be orders of magnitude faster to compute. The choice between Fresnel and angular spectrum propagators depends on the sampling of the quadratic phase factors in each. The phase must not change by more than $\pi$ radians in neighboring sampling points or else the wavefront aliases. This criterion leads to the choice of the angular spectrum for plane-to-plane propagation and Fresnel for propagation between a parabolic reference and the focal plane.\textsuperscript{[101, 100]}

The simplest imaging system uses a lens or collimating mirror to bring light to a focus. If an optic is placed one focal length $f$ before the collimating optic, it can be shown (see, e.g., \textsuperscript{[98, 106]}) that the relationship between that pupil plane and the focal plane is simply a FT times a scalar,

$$E_{\text{focus}}(\xi, \eta) = \frac{e^{i\pi f}}{i\lambda f} F\{E_{\text{pupil}}(x, y)\}. \quad (1.5.5)$$

This assumes that the collimating optic is unaberrated and sufficiently oversized not to clip the beam. For coronagraph design or optical modeling when the aberrations are unknown, Eq. 1.5.5 provides much faster propagation between the pupil and focal planes than using Eqs. 1.5.2 and 1.5.4 in series. Although an individual FFT is quite fast for the square matrix representations of the electric field with hundreds or thousands of sampling points per side, it is usually the slowest operation in optical propagation. FPWC simulations require thousands of FT calculations, so it is important to keep matrices small and use Eq. 1.5.5 instead of Fresnel and angular spectrum propagations whenever reasonable.
1.6 Controllability and Observability

1.6.1 Controlling the Electric Field with Deformable Mirrors

Controllability

Here we provide a summary of the electric field controllability with DMs in the context of FPWC. Other authors[107, 89, 106] derive these findings in detail. We separate the controllability into three categories: frequency coverage in the pupil, phase and amplitude in the pupil, and spectral bandwidth. Because we assume the DM corrections are generally much smaller than the wavelength, we use the first order Taylor series expansion of the phase at the DM,

\[ A(x, y)e^{i\Delta\phi(x,y)} \approx A(x, y) + iA(x, y)\Delta\phi(x, y), \]  

(1.6.1)

where \( A(x, y) \) is the nominal complex field at the DM and \( \Delta\phi(x, y) \) is the change in phase from actuating the DM.

DMs are placed at or near the pupil plane to conjugate the largest sources of phase aberrations, which are on the primary and secondary mirrors for a space telescope. We decompose the pupil-plane phase aberrations into a Fourier series of sines and cosines to determine what is controllable. The FT relationship between the pupil and image means that a sinusoid of frequency \( N \) cycles per aperture in the pupil produces symmetric delta functions (convolved with the PSF of the telescope pupil) in the image at \( \pm N/2 \lambda/D \). A DM with \( N_a \) actuators across the pupil can actuate a maximum frequency of \( N_a/2 \) cycles per aperture, thereby controlling out to \( N_a/2 \lambda/D \) from the star along each axis of the image. The DM can also suppress the low-frequency aliasing effects of beating among higher, uncontrollable spatial frequencies.

Because a DM alters phase, it can obviously conjugate phase aberrations in the pupil; control of amplitude aberrations at the DM is less intuitive. The two sources of amplitude
aberrations are reflectivity errors in the mirrors and phase-to-amplitude mixing of aberrations. The off-the-shelf mirrors for a testbed typically have reflectivity non-uniformity errors of less than 1%, and for a space telescope the errors would need to be slightly lower.\[69\] Free-space propagation of phase aberrations from the powered optics (lenses or curved mirrors) to the pupil converts some of the phase into amplitude via the Talbot effect.\[108\] In testbeds this creates amplitude aberrations of several percent in magnitude. Since the DM cannot directly control amplitude aberrations in the pupil, we decompose the amplitude aberrations into a Fourier series and analyze the effects in the image plane. Amplitude aberrations are purely real in the pupil, so they are Hermitian in the image. DM actuation is purely imaginary to first order in the pupil and is thus anti-Hermitian in the image. Therefore, a single DM can null an amplitude-induced speckle on just one side of the image; the speckle on the other side doubles in amplitude. Shaklan and Green\[69\] first derived that two DMs in series allowed correction of pupil-plane amplitude aberrations on both sides of the image. Via the Talbot effect, a sinusoid on one DM partially converts to amplitude by the time it reaches the second DM. The second DM then cancels the remaining phase component and effectively enables pure amplitude actuation with only phase-altering DMs.\[89\]

Pueyo and Kasdin\[107\] extended this analysis to incorporate spectral bandwidth. The control of amplitude aberrations via Talbot mixing is inherently chromatic. When the electric field is expanded in a Laurent series about wavelength, it becomes clear that two DMs in series can perfectly correct phase-induced amplitude aberrations at only one wavelength. Three DMs enable correction of the leading terms of the wavelength expansion, thereby providing broadband, high-contrast dark holes. For future space-based coronagraphs, the current plan is use only two DMs in series and manufacture the mirrors smooth enough to keep the uncorrectable effect of chromatic amplitude aberrations below the desired dark hole contrast level.
Control Algorithms

Trauger et al.\cite{109} demonstrated the first successful controller in the laboratory for FPWC. In this model-free scheme called “speckle nulling,” sinusoids with different phases are applied to the DM to suppress stellar speckles at the targeted spatial frequency. Speckle nulling provides only localized correction but can be made faster by targeting many speckles simultaneously at each correction iteration. Nevertheless, it requires hundreds or thousands of correction iterations to suppress the entire dark hole fully and is thus too slow for a space mission. Model-based estimation and control have enabled much faster correction.

Malbet et al.\cite{110} proposed the first model-based DM control law to create a high-contrast dark hole in the coronagraphic image. Their technique involved a nonlinear optimization that was numerically unstable and too time consuming for real-time correction.\cite{89} Borde and Traub\cite{111} proposed a linearized controller to minimize the energy in the dark hole as a quadratic cost function, but this direct minimization was numerically unstable. Give’on et al.\cite{112} applied Tikhonov regularization to stabilize the problem and called this algorithm electric field conjugation (EFC). Pueyo et al.\cite{89} derived a similar controller called stroke minimization that minimizes the DM actuation subject to a constraint on the achieved contrast. Both EFC and stroke minimization have been tested in the laboratory and proven to be much faster than speckle nulling. We discuss these linearized control laws in more detail in §2.2.

1.6.2 Observability of the Image Plane Electric Field

Phase Retrieval

In the early development of wavefront correction before dedicated phase sensors, the science camera was used as the wavefront sensor. The phase at the focus had to be estimated because cameras only measure the intensity of the incident light. With the entire electric field at the camera plane, one could perform an inverse FT to obtain the electric field at the pupil
plane. The DM would then be used to conjugate the phase.

The first technique for phase retrieval was the Gerchberg-Saxton algorithm.[113] Gerchberg-Saxton is an iterative routine that propagates back and forth between the pupil and focus and applies known constraints at each, such as the measured intensity in the focus and the known aperture at the pupil. Maximum likelihood estimation (MLE) used nonlinear optimization, often gradient based, to find the best fit of some modal basis at the pupil to the measured image.[114] Both techniques suffered from degenerate solutions when trying to determine the phase from a single image.[115] The inclusion of phase diversity greatly improved both methods.[116, 117] Phase diversity is the use of known phase modulation at the pupil to alter the intensity in the image. From the mixing of the known and unknown aberrations in the focal plane, enough data are present in several images to break any degeneracies and enable accurate estimation of the original electric field. The simplest kind of phase diversity to implement is defocus diversity, which produces pure defocus by translating the camera or using a filter wheel to switch among lenses. The DM(s) used for correction can also be used for introducing phase diversity, but there is more uncertainty in the shape of the applied phase.

Wavefront correction with phase retrieval is too slow for AO and is mostly used as a calibration tool. In a famous instance, the on-orbit phase retrieval of the Hubble Space Telescope’s flawed primary mirror enabled a later space shuttle mission to install corrective optics and save the mission.[118, 119] For coronagraphic testbeds, phase retrieval is used to measure and flatten the nominal pupil wavefront as much as possible before starting FPWC.

Phase Retrieval with a Coronagraph

In FPWC, the electric field is sensed and corrected directly in the focal plane to produce the dark hole. Therefore, the science camera must again be used as the wavefront sensor. Without a coronagraph, phase retrieval utilizes the whole PSF and both forward and backward propagation through the optical model. The coronagraphic case is much more difficult be-
cause opaque coronagraphic masks or field stops prevent backward modeling. Coronagraphic phase retrieval still relies on forms of phase diversity to estimate the phase of the light in the dark hole.

The main challenge of FPWC with a high-contrast coronagraph is to sense the electric field quickly and accurately from a set of intensity images. Because the optical model is imperfect, FPWC must be performed iteratively. As correction suppresses the starlight, longer exposures are necessary to obtain an adequate signal-to-noise ratio (SNR) in each estimation image. Assuming computational overhead is minimal on a coronagraph-equipped space observatory, FPWC will be limited by the acquisition rate of images for estimation. There is thus a strong incentive make FPWC faster by improving optical model accuracy and reducing the total exposure time required for estimation images.

Several techniques have been developed for creating diversity in the coronagraphic intensity image and then estimating the stellar electric field. The self-coherent camera [120] is an estimation technique for coronagraphs with focal plane phase masks. Pinholes outside the nominal beam radius at the Lyot stop produce an interference pattern in the image that is used to calculate the electric field. A more general, nonlinear technique is coronagraphic focal plane wavefront estimation for exoplanet detection (COFFEE),[121] which utilizes a maximum a posteriori approach to estimate pupil plane aberrations and the bias signal. COFFEE introduces large phase aberrations at the DM to create diversity in the image plane. COFFEE is currently too slow computationally to implement in real time, it requires phase diversity too large to actuate with some types of DMs, and it has not yet been used to recover high contrast in experiment. In this thesis, we generate phase diversity with small, pair-wise probes actuated on the DM.[112] We utilize only the pair-wise technique for estimation because it is linear, making it tractable and practical for real-time correction; it can be used with any coronagraph type; and it is the only model-based estimation scheme used to date to achieve better than $10^{-8}$ contrast in testbed experiments.
1.7 Exoplanet Signal Extraction Methods

The goal of image post-processing methods is to remove the starlight in the image and leave only the light from sources incoherent with the star. Within the dark hole, there are many signals to take into account. Compact bodies (such as stars and exoplanets) are far enough away that we cannot resolve their individual features. Since we observe these compact objects as points, the wavefront we receive from each of them is coherent with itself. The separately resolvable objects are then incoherent with each other. The light incoherent with the star contains signals from several sources. The science targets are exoplanets, companion stars, and circumstellar accretion disks. Within the stellar system of interest, exozodiacal dust (called exozodi) will produce an incoherent signal that increases in brightness the closer it is to the star. In the foreground, the telescope sees zodiacal dust (called zodi) in our solar system, which produces an approximately flat illumination across the miniscule field of view in the dark hole. In the background, there can be stars and galaxies. Finally, there will be some level of scattered light in the system from stray reflections.

Regardless of which post-processing method is used to separate starlight from incoherent light, the problem of disambiguating incoherent signals is common to all. In this thesis, we therefore address only the removal of the coherent stellar signal and treat the exoplanet as the only source of incoherent light during simulated detection experiments. We justify our assumption of separable incoherent sources with several possible methods. Zodi can be subtracted off as a flat background, stray light can be subtracted if it is static for long enough, background stars will have drastically different spectra from exoplanets, and exozodi might be smooth and continuous enough to model and subtract off.

Until now we focused on the reduction of diffracted starlight with hardware–coronagraphs and DMs. After the dark hole is formed, some planets will be significantly brighter than the remaining starlight and count as obvious detections. The more challenging cases are the planets hidden at or below the residual starlight level. An exoplanet will appear in the image as an unresolved PSF, which has the same spatial scale as the residual starlight.
speckles ($\approx \lambda/D$, the FWHM of the PSF). Some form of image plane diversity is therefore required to distinguish faint planets from speckles. Here we summarize state-of-the-art methods for extracting the exoplanet signal and describe how the work in this thesis enables a new detection method.

**Current Methods**

Over the past two decades, several sophisticated methods have been invented to differentiate starlight from exoplanets and disks, effectively yielding even better contrast than in the raw images. The first technique was spectral differential imaging (SDI).[122, 123] SDI utilizes the spectra obtained for planet characterization to aid in detection as well. The brightest speckles (at worse than about $10^{-5}$ contrast) are from mostly achromatic phase and amplitude aberrations in the pupil, which produce image plane speckles with fixed radii in $\lambda/D$. Therefore, these speckles appear to move radially outward from the star as the wavelength increases. Since the planet is a fixed distance on-sky from the star, the planet appears not to move with wavelength.

Angular differential imaging (ADI)[124] leverages spatial motion to differentiate speckles and incoherent sources. By rotating the telescope or sky, the quasi-static stellar speckles from the instrument remain fixed on the science camera while the planet appears to circle the star. A space telescope accomplishes ADI by rolling while pointed at the target star. The sky naturally rotates within the field of view for large ground telescopes with altitude-azimuth mounts. When the planet is farther from the star and/or the rotation rate is faster, ADI performs better since it obtains more diversity among the planet and speckles at that radius in the image. The achievable rotation rate is limited by the thermal stability for a space telescope and set by the earth’s rotation, telescope latitude, and target’s position for a ground telescope.

Reference differential imaging (RDI) is based on the principle of PSF subtraction. If a stable, static optical system images two spectrally similar stars, then subtracting those PSFs
should leave only measurement noise and the incoherent sources around each star. Current telescopes are not stable to high contrast, however, so direct PSF subtraction of two images has only marginally helped improve the effective contrast. Lafrenière et al.[125] realized that quasi-static aberrations appear only in some PSF images but not others depending on such factors as the stellar spectrum and thermal loading of the telescope. They created the locally optimized combination of images (LOCI) algorithm, which builds the reference PSF from a library of PSFs from the same instrument. Soummer et al.[126] built upon this idea with the Karhunen-Loève image projection (KLIP) algorithm. KLIP decomposes each PSF with a Karhunen-Loève basis set and uses the first few eigenmodes to build the best reference PSF for subtraction. Soummer et al.[126] used KLIP in 2012 to reveal that three of the four known HR8799 exoplanets were in 1998 Hubble Space Telescope images—a full ten years before Marois et al.[127] first imaged and discovered the planets from the ground.

These post-processing methods are being considered an integral part of the technology development for the WFIRST CGI.[128] The WFIRST CGI dark holes are expected to have raw contrast values in the range of $10^{-8}$ and $10^{-9}$,[95] whereas most of the planets to be imaged are predicted to have contrasts between $10^{-9}$ and $10^{-10}$. Researchers at JPL are investigating the use of ADI for the CGI. Ygouf et al.[129] found that SDI is useful for moderate contrast images from current ground observatories, but Soummer et al.[130] have found that SDI is less applicable in the high-contrast case of WFIRST. FPWC leaves only the highly chromatic aberrations in the pupil from the phase-to-amplitude mixing of aberrations non-conjugate to the DMs. In the image plane, this extreme chromaticity of the high-contrast stellar speckles has prevented their accurate modeling and subtraction via SDI. Soummer et al.[130] found that RDI is more promising, and Ygouf et al.[131, 132] have begun using KLIP on simulated, noiseless WFIRST CGI images to obtain approximately five times better effective contrast. Because of the one-year mission and long integration times required for each high-contrast image, RDI for the WFIRST CGI will have a smaller library of PSFs compared to previous cases in which RDI has helped. The long-integration times
and limited mission lifetime for the WFIRST CGI also mean that the ultimate achievable contrast, even with post-processing, will be set by the photon shot noise and uncorrected LOWFSC residuals at about $10^{-10}$ contrast.

**Coherence Differential Imaging**

All the post-processing techniques just described use only intensity images with no knowledge of the optical system. Techniques such as LOCI and KLIP implicitly utilize the fact that certain disturbances in the system have predictable, repeatable effects such as deformation of the primary mirror from gravity sag on earth, or position- and pointing-dependent thermal loads in space. Variations in the wavefront act on the electric field, so they are more difficult to isolate in intensity because the cross terms with the nominal stellar electric field also appear. Coherence differential imaging (CDI) uses knowledge of the stellar electric field to identify sources incoherent with the starlight such as exoplanets. So far CDI has not been utilized for detection because it requires modulation of the electric field in a known manner. Astronomical space telescopes do not yet have high-actuator-count DMs or sophisticated coronagraphs to make use of CDI. It may be feasible with current ExAO-equipped ground telescopes, but the large AO residuals and quasi-static aberrations would need to be accurately measured.

Proposed high-contrast ground- and space-based coronagraph instruments will have the high stability and accurate metrology to enable CDI for exoplanet detection. In this thesis, we explore for the first time the use of wavefront correction to provide the image diversity for CDI. A batch process CDI could be used after forming the dark hole similar to how other post-processing methods are implemented. Our approach uses the formerly wasted overhead time for FPWC to build a recursive estimate of the planet signal. The probed diversity images, originally just used for estimation, also provide the diversity for CDI. The uncorrected LOWFS residuals (and ExAO residuals for a ground telescope) will also modulate the focal plane electric field and should be included in the CDI calculation to help...
reach the fundamental detection floor from photon noise alone. As a proof of principle, we investigate only the case of static aberrations for exoplanet detection with recursive CDI (RCDI).

1.8 Dissertation Overview

The overall theme of this dissertation is efficient FPWC for the direct imaging of exoplanets. Most proposed coronagraph missions assume that the time required for wavefront correction will be negligible compared to the length of the mission. To realize that goal of having efficient FPWC, we have worked to reduce the required exposure time and utilize correction images for science. Because obtaining images for estimation will be the slowest part of FPWC during a mission, much of our work has been to implement recursive estimators to increase correction speed and robustness to noise. We present laboratory experiments and their results from Princeton’s HCIL and JPL’s HCIT to compare the performance of batch process and recursive estimators embedded in a control loop. Our implementation of recursive bias estimation also enables the use of FPWC images for exoplanet detection via RCDI. We performed simulations of FPWC for the WFIRST CGI to compare the correction speed with different estimators and to compare RCDI’s exoplanet detection probability to the fundamental shot noise limit.

Chapter 2 delves into the mathematical framework of linear focal plane wavefront correction (FPWC). We show that the two most commonly used linear control laws solve the same quadratic cost function. Because we compare the performance of several estimators in later chapters, we describe in depth the pair-wise estimation scheme conceived by Give’on et al.[112] as a batch process estimator (BPE) and incorporated into a Kalman filter (KF) by Groff and Kasdin.[83] We present our new work that calculates the theoretical accuracy limits in the starlight estimate for pair-wise estimation.[133]
Chapter 3 details our modification of the pair-wise estimation scheme for use with a non-linear estimator. We formulate an extended Kalman filter (EKF) in the context of pair-wise probing for FPWC. After showing that the EKF estimate has bias error, we describe an iterated extended Kalman filter (IEKF) for improved estimation. Although the Kalman filters have more computational complexity than the BPE, we show that the recursive estimators are no more demanding than the controller and should be viable for real-time calculation on a spacecraft.

Chapter 4 describes the lab experiments we conducted in Princeton’s HCIL. We performed correction speed comparisons for all the estimators with and without a bright uniform background to mimic light from zodiacal dust. In the first laboratory demonstration of RCDI, we injected a planet in the dark hole and compare the extracted planet signal retrieved via PSF subtraction and CDI. For consideration during a future re-design of the HCIL, we identify in simulation several sources of aberrations that limit the achievable broadband contrast in the HCIL.

Chapter 5 details simulations of the WFIRST CGI and lab experiments in JPL’s HCIT. In both cases, we compare the correction speed of the BPE, KF, and IEKF. For the first time in the laboratory, we demonstrated FPWC using extremely low SNR images to reduce the total integration time during correction.

Chapter 6 describes the first statistical analysis of RCDI. We perform Monte Carlo simulations of correction runs for a simple observation scenario with the WFIRST CGI. We then compute the probabilities of detection and probabilities of false alarms for RCDI and compare those to the fundamental limits set by photon shot noise.
Chapter 7 outlines the directions for future research in the high contrast imaging of exoplanets. We make recommendations for algorithm development to support the WFIRST CGI in the near-term and to support future ground- and space-based coronagraph missions in later decades.

Chapter 8 reviews the major conclusions from this thesis.
Chapter 2

Linear Focal Plane Wavefront Correction

This chapter focuses on the fundamentals of linear focal plane wavefront correction (FPWC) as a mathematical foundation for the later chapters. In §2.1 we define the deformable mirror (DM) surface as the superposition of individual actuator pokes and linearize the electric field about the phase at the DM. Section 2.2 describes the linear-model, quadratic-cost control laws commonly used in FPWC. In §2.3 we re-derive pair-wise estimation in its simplest form as a batch process estimator. We derive the dependence of the electric field estimate accuracy on noise sources and the probed image intensity in §2.4. Finally, in §2.5 we summarize the work of Groff and Kasdin[83] in combining pair-wise estimation with a Kalman filter. The derivations in §2.1 and §2.3 appear in a paper by Riggs et al.[84]

2.1 Linearization of the Optical Model

Here we derive the linearized electric field at the DM for use with the linear controller and linear estimator. Let $\tilde{E}_0(x, y)$ be the initial complex electric field at the DM including the incident field and the nominal complex aberrations on the DM, where $(x, y)$ are coordinates in the plane of the DM. Let $\phi_{k-1}(x, y)$ be the total phase contribution of the DM at correction
Figure 2.1: Planes used in the linearization of the electric field about the DM phase. The field at the DM, $\tilde{E}(x,y)$, propagates through the coronagraph and is focused onto the science camera plane with electric field $E(\xi, \eta)$.

Iteration $k-1$, and let $\Delta \phi_k(x,y)$ be the perturbation of the DM phase at correction iteration $k$ such that

$$\phi_{k-1}(x,y) = \sum_{j=1}^{k-1} \Delta \phi_j(x,y). \quad (2.1.1)$$

The phase at the DM is twice the surface height of the DM and scales inversely with wavelength $\lambda$ (in meters). Assuming small deformations, we approximate the DM surface as the sum of the normalized actuator influence function $f(x,y)$ times the displacement command $\Delta u_{k,q}$ (in meters) at each actuator $q$’s center location $(x_q, y_q)$. The perturbation phase at the DM is then given by

$$\Delta \phi_k(x,y) = \frac{2}{\lambda} \sum_{q}^{N_{act}} \Delta u_{k,q} f(x - x_q, y - y_q), \quad (2.1.2)$$

where $N_{act}$ is the number of DM actuators.

Again assuming small perturbation commands to the DM, we can approximate the electric field leaving the DM, $\tilde{E}_k(x,y)$, with a first order Taylor series expansion about the most
recent DM perturbation,

$$\tilde{E}_k(x, y) = \tilde{E}_0(x, y)e^{i(\phi_{k-1}(x,y) + \Delta \phi_k(x,y))}$$

$$\approx \tilde{E}_0(x, y)e^{i\phi_{k-1}(x,y)}(1 + i \Delta \phi_k(x,y)). \tag{2.1.3}$$

Nearly all coronagraphs can be modeled as a series of linear operators such as Fourier transforms, Fresnel propagations, and mask multiplications. (The notable exception is Phase Induced Amplitude Apodization (PIAA),[23] which breaks the assumptions of the Fresnel approximation with extremely large phase gradients.) Since the control and estimation methods presented here are general to all coronagraphs, we represent the propagation from the DM to the science camera by the linear operator $\mathcal{C}\{\cdot\}$ to obtain the focal plane electric field $E_k(\xi, \eta)$,

$$E_k(\xi, \eta) = \mathcal{C}\{\tilde{E}_k(x, y)\}$$

$$\approx \mathcal{C}\{\tilde{E}_0(x, y)e^{i\phi_{k-1}(x,y)}\} + \mathcal{C}\{i\tilde{E}_0(x, y)e^{i\phi_{k-1}(x,y)}\Delta \phi_k(x, y)\}$$

$$= E_{k-1}(\xi, \eta) + \sum_{q}^{N_{\text{act}}} \Delta u_{k,q} \mathcal{C}\{i\tilde{E}_0(x, y)e^{i\phi_{k-1}(x,y)}f(x - x_q, y - y_q)\}$$

$$= E_{k-1}(\xi, \eta) + \sum_{q}^{N_{\text{act}}} \Delta u_{k,q}B_{k-1,q}(\xi, \eta), \tag{2.1.4}$$

where $(\xi, \eta)$ are coordinates in the image. The aberrated focal plane electric field before the new command at correction iteration $k$ is

$$E_{k-1}(\xi, \eta) = \mathcal{C}\{\tilde{E}_0(x, y)e^{i\phi_{k-1}(x,y)}\}, \tag{2.1.5}$$

and the Jacobian of each actuator at the image is given by the function

$$B_{k-1,q}(\xi, \eta) = \mathcal{C}\{i\tilde{E}_0(x, y)e^{i\phi_{k-1}(x,y)}f(x - x_q, y - y_q)\}. \tag{2.1.6}$$
Detectors measure the intensity in finite-sized pixels, so the focal plane field is converted from the continuous coordinates $(\xi, \eta)$ to the discrete indices $(m, n)$. The detector integrates the intensity over the whole pixel whereas the model just samples discretely at each pixel. With greater than Nyquist discretization ($\geq 2$ pixels per frequency unit ($\lambda/D$)) of the PSF already required for wavefront correction, the effect of sampling the PSF at each pixel instead of integrating over that area is negligible. The discretized focal plane electric field is thus

$$E_{k,m,n} = E_{k-1,m,n} + \sum_{q}^{N_{\text{act}}} \Delta u_{k,q} B_{k-1,q,m,n},$$  \hspace{1cm} (2.1.7)

where we have implicitly defined the region as being only within the dark hole. To perform matrix operations on the discretized field, it is convenient to reshape the field into a vector of length $N_{\text{pix}}$, the number of dark hole pixels, such that

$$E_k = E_{k-1} + G_{k-1} \Delta u_k.$$  \hspace{1cm} (2.1.8)

Both $E_k$ and $E_{k-1}$ have dimensions $N_{\text{pix}} \times 1$, the complex-valued control Jacobian $G_{k-1}$ has dimension $N_{\text{pix}} \times N_{\text{act}}$, and the vector of control commands $\Delta u_k$ has dimension $N_{\text{act}} \times 1$. Since the effect of a single DM is approximated as additive, the linearized field for $N_{DM}$ DMs is

$$E_k = E_{k-1} + \sum_{l=1}^{N_{DM}} G_{k-1,l} \Delta u_{k,l},$$  \hspace{1cm} (2.1.9)

where $G_{k-1,l}$ and $\Delta u_{k,l}$ are the control Jacobian and control signal, respectively, for the $l$-th DM. After this point, we drop the differential notation of $\Delta u_k$ and write just $u_k$ for clarity. Without loss of generality, we also assume a single DM in later derivations of the controllers and estimators.
2.2 Linear Focal Plane Wavefront Control

In this section, we derive and compare the two linear control laws developed by Give’on et al.[112] and Pueyo et al.[89] We show that both solve the same quadratic cost function. Groff et al.[133] compare the two controllers in more detail.

In focal plane wavefront control, the goal is to control the system while minimizing both the deviation from the desired state and the magnitude of the control signal. For a linear system, the cost function is typically set as quadratic to have a unique minimum. We therefore define the general cost function as

\[ J = e^T \tilde{Q} e + u^T \tilde{R} u, \quad (2.2.1) \]

where \( \tilde{Q} \) and \( \tilde{R} \) are symmetric, positive-definite weighting matrices, \( e \) is a vector of the difference between the desired and actuated electric fields, and \( u \) is again the vector of control signals. In our case, the error vector of dimension \( N_{\text{pix}} \times 1 \) is the remaining electric field \( (E_{k-1} + G_{k-1} u_k) \) since we would like to zero out the energy in the dark hole. Although the pixel intensities can be weighted, we simplify the derivation by leaving them unweighted and setting \( \tilde{Q} = I \), where \( I \) is the identity matrix. We also leave the individual DM actuators unweighted for simplicity, but the actuation overall needs to be weighted with respect to the cost of the error. Thus, we set \( \tilde{R} = \alpha I \) with \( \alpha \) as the actuation weighting factor. The cost function is now

\[ J_k = (E_{k-1} + G_{k-1} u_k)^*(E_{k-1} + G_{k-1} u_k) + \alpha u^T u \quad (2.2.2) \]

\[ = E_{k-1}^* E_{k-1} + u_k^T (\alpha I + G_{k-1}^* G_{k-1}) u_k + 2 \mathcal{R}\{u_k^T G_{k-1}^* E_{k-1}\}, \quad (2.2.3) \]

where \( * \) gives the conjugate transpose. We find the minimum cost by setting \( \frac{\partial J_k}{\partial u_k} = 0 \) and
solving for $u_k$,

$$u_k = -(G_{k-1}^*G_{k-1} + \alpha I)^{-1}R\{G_{k-1}^*E_{k-1}\}. \quad (2.2.4)$$

Although Eq. 2.2.4 provides a solution to the control signal, the choice of the actuator damping factor $\alpha$ is ill-defined. A line search over $\alpha$ to minimize $J_k$ is computationally tractable. We cannot utilize the best value of $\alpha$ from an unbounded, model-based line search because the linearization and model uncertainty limit the true achievable correction step. Too little damping causes divergence and addition of energy into the dark hole, whereas too much damping yields an unnecessarily small correction step. An empirical line search (i.e., taking images at several $\alpha$ values and comparing the contrast achieved at each) is used for controller characterization in the lab but is impractical during long, photon-limited observations, such as during a space mission.

**Electric Field Conjugation**

Give’on conceived of a linear controller called electric field conjugation (EFC) that would negate the error in the estimated electric field from its design value,

$$0 = (E_{k-1} + G_{k-1}\Delta u_k) - E_{\text{design}}, \quad (2.2.5)$$

via the optimal command

$$u_k = (G_{k-1}^*G_{k-1})^{-1}R\{G_{k-1}^*(E_{\text{design}} - E_{k-1})\}. \quad (2.2.6)$$

Some coronagraph designs for open apertures theoretically cancel all starlight (i.e., $E_{\text{design}}=0$), which makes EFC equivalent to the linearized energy minimization proposed by Borde and Traub.[111] Give’on et al.[112] eventually realized that the targeted electric field should always be set to zero for any coronagraph because the design field might be unreachable and
the controller can sometimes reach contrasts below the design level of the coronagraph. Setting $\alpha$ to zero in Eq. 2.2.2 gives the cost function to minimize just the energy in the dark hole. From Eq. 2.2.4 with $\alpha=0$, the optimal command $u_k$ to minimize the dark hole electric field is

$$u_{\text{EnergyMin},k} = -(G^*_{k-1}G_{k-1})^{-1}\mathcal{R}\{G^*_{k-1}E_{k-1}\}. \quad (2.2.7)$$

Directly using Eq. 2.2.7 results in divergence because $G^*_{k-1}G_{k-1}$ is ill-conditioned and produces overly large actuator commands, so Give’on et al.[112] used Tikhonov regularization [134] to damp the actuator commands,

$$u_{\text{EFC},k} = -(G^*_{k-1}G_{k-1} + \alpha I)^{-1}\mathcal{R}\{G^*_{k-1}E_{k-1}\}. \quad (2.2.8)$$

Regularized EFC in Eq. 2.2.8 gives the same exact control command as the quadratic controller in Eq. 2.2.4 and still requires finding an acceptable value of $\alpha$.

**Stroke Minimization**

Pueyo et al.[89] solved the damping parameter selection problem by altering the cost to minimize actuator stroke alone subject to a constraint on contrast, $C_k$. In the stroke minimization algorithm, the optimization is posed as a convex program,

$$\begin{align*}
\text{minimize} & \quad u_k^T u_k \\
\text{subject to} & \quad \bar{I}_k \leq C_k
\end{align*} \quad (2.2.9)$$

where $\bar{I}_k = E_k^* E_k / N_{\text{pix}}$ is the average contrast in the dark hole. We could solve Eq. 2.2.9 with a convex optimizer, but we typically alter the problem slightly to obtain a short, closed-form solution. By instead using the equality constraint $\bar{I}_k - C_k = 0$, we can add it to the cost
function via a Lagrange multiplier $\mu$,

$$J_{SM,k} = u^T u + \mu (\bar{T}_k - C_k)$$

$$= \mu E_{k-1}^* E_{k-1} + u_k^T (\mathbb{I} + \mu G_{k-1}^* G_{k-1}) u_k + 2\mu \mathcal{R} \{ u_k^T G_{k-1}^* E_{k-1} \} - \mu N_{pix} C_k. \quad (2.2.10)$$

Setting $\partial J_{SM,k}/\partial u_k = 0$ and solving for the optimal command $u_k$, we arrive at the same expression as Eq. 2.2.4 with $\mu = 1/\alpha$. Although the choice of $C_k$ is still ad hoc, there is an advantage since the control choice has become grounded in a relevant testbed parameter. By re-defining the problem as in stroke minimization, we can use the model-based line search to try to reach a target contrast rather than arbitrarily selecting different values of $\alpha$. In either case, once $\alpha$ or $C_k$ is characterized for a coronagraphic testbed, it typically does not change without a major disturbance to the alignment.

2.2.1 Broadband Control

A major purpose of direct imaging is to obtain spectra of exoplanets and disks, necessitating broadband dark holes. A broadband control scheme is required since the contrast steeply degrades outside the controlled wavelength range. A closed-form formula of the dark hole’s electric field on wavelength has thus far been intractable, so we instead approximate the broadband dark hole as a sum of images at several discrete wavelengths. For most coronagraphs, a sub-bandpass of up to $\Delta \lambda/\lambda_0=2\%$ is the largest that can be well-approximated as a single wavelength. A 10% bandpass is therefore treated as the sum of intensities at five discrete wavelengths. The broadband control law used at JPL’s HCIT weights the each of these discrete wavelengths equally and attempts to suppress them all simultaneously.[112]
The broadband cost function is thus

\[
J_{BB,k} = \alpha u^T u + \frac{1}{N_\lambda} \sum_l w_l \left( E_{k-1}(\lambda) + G_{k-1}(\lambda) u_k \right) * \left( E_{k-1}(\lambda) + G_{k-1}(\lambda) u_k \right)
\]

\[
= \alpha u^T u + \frac{1}{N_\lambda} \sum_l w_l I_k(\lambda) + \frac{1}{N_\lambda} u_k^T \left( \sum_l w_l G_{k-1}^*(\lambda) G_{k-1}(\lambda) \right) u_k,
\]

(2.2.11)

where \( N_\lambda \) is the number of discrete wavelengths in the correction bandpass and \( E_{k-1} \) and \( G_{k-1} \) are dependent on wavelength. The weight \( w_l \) on each wavelength is set to unity since all the wavelengths are equally important for spectral analysis. Groff et al.\[133\] discuss the weighting choices and other possibilities for broadband control in more detail. With \( w_l = 1 \) for all \( l \) and minimizing \( J_{BB} \), we find that the optimal broadband command is

\[
u_{k, BB, opt} = - \left( \frac{1}{N_\lambda} \sum_l G_{k-1}^*(\lambda) G_{k-1}(\lambda) \right) + \alpha \Pi \right)^{-1} R \left\{ \frac{1}{N_\lambda} \sum_l G_{k-1}^*(\lambda) E_{k-1}(\lambda) \right\}.
\]

(2.2.12)

In practice, the control Jacobian \( G_{k-1} \) is usually not updated (\( G_0 \) is used instead) to save computation time at the expense of somewhat slower correction. From here on we will use the notation \( G \) instead of \( G_0 \) for convenience. The errors in the estimate and model, ignored nonlinearities of the DM phase contribution, and the use of regularization cause the new, corrected field to have only a slight improvement in contrast after each control step. The correction is thus iterative, with a new DM command calculated and applied after each new estimate of the electric field at each controlled wavelength.

2.3 Batch Process Pair-wise Estimation

The model-based control techniques described above require knowledge of the electric field in the dark hole. An estimation approach is thus needed to determine the field from intensity measurements in the science camera. Currently the baseline estimation method for a coronagraphic space mission is pair-wise difference imaging as developed by Give'on
et al. [112, 135], which probes the image via small DM perturbations. It is the only model-based estimation scheme that has attained better than $10^{-8}$ contrast in laboratory experiments,[85, 136, 92, 91, 86] all of which have been in JPL’s HCIT. Pair-wise estimation is characterized by several notable features: it can be used with any coronagraph, it requires no additional hardware, and it is fairly robust to model uncertainty. This method is described in detail in several papers,[112, 135, 133] but we revisit the derivation to provide the mathematical foundation for our new work. Since starlight is inherently broadband, the monochromatic measurements described are actually narrowband (typically with $\Delta \lambda / \lambda_{\text{central}} = 2–3\%$) measurements over which the electric field variation is negligible.

### 2.3.1 Constructing the Batch Process Pair-wise Estimator

In pair-wise estimation, shapes are actuated on a single DM to probe the electric field in the dark hole. Give’on et al. [135] explain one such method for choosing sets of probe shapes. To make the estimation linear, a separate image is taken for the positive and negative of each probe shape applied on the DM. At least two pairs of different probes are required to distinguish the real and imaginary parts of the electric field. In this analysis, we assume that the incoherent signal is static and that the starlight is modulated only by the DM.

Let $u_j$ be the differential control signal for the $j$-th positive probe shape. Then, the linearized change in the focal plane electric field from the probe shape is defined as $p_j = Gu_j$. For convenience we do not explicitly write the dark hole pixel index for the probe field $p$, the electric field $E$, and the intensity $I$. The focal plane intensity at each dark hole pixel for a given positive or negative probe shape is then

$$I_{k,j}^\pm = |E_k \pm p_{k,j}|^2 + I_{\text{inco},k} + n_{k,j}^\pm$$

$$= |E_k|^2 + |p_{k,j}|^2 \pm 2\Re\{E_k^* p_{k,j}\} + I_{\text{inco},k} + n_{k,j}^\pm,$$  \hspace{1cm} (2.3.1)

where $n_{k,j}^\pm$ is the zero-mean, Gaussian measurement noise and $I_{\text{inco},k}$ is any light incoherent
with the star. The difference of the positive and negative probed images is equal to twice
the cross term in the electric field plus noise,

\[ \Delta I_{k,j} = I_{k,j^+} - I_{k,j^-} = 4\Re\{E_k^* p_{k,j}\} + n_{k,j}, \tag{2.3.2} \]

where \( n_{k,j} = n_{k,j^+} - n_{k,j^-} \) is the total noise having twice the variance of a single probed
image. For a set of measurements from \( N_{pp} \) probe pairs, the measurement equation is

\[
\begin{bmatrix}
\Delta I_{k,1} \\
\vdots \\
\Delta I_{k,N_{pp}}
\end{bmatrix} = 4
\begin{bmatrix}
\Re\{p_{k,1}\} & \Im\{p_{k,1}\} \\
\vdots & \vdots \\
\Re\{p_{k,N_{pp}}\} & \Im\{p_{k,N_{pp}}\}
\end{bmatrix}
\begin{bmatrix}
\Re\{E_k\} \\
\vdots \\
\Im\{E_k\}
\end{bmatrix} + \begin{bmatrix}
n_{k,1} \\
\vdots \\
n_{k,N_{pp}}
\end{bmatrix}, \tag{2.3.3}
\]

where \( \Im\{\cdot\} \) takes the imaginary part of the complex value. We re-write Eq. 2.3.3 as

\[ z_k = H_k x_k + n_k, \tag{2.3.4} \]

where the set of measurements is

\[ z_k = \begin{bmatrix}
\Delta I_{k,1} \\
\vdots \\
\Delta I_{k,N_{pp}}
\end{bmatrix}, \tag{2.3.5} \]

the linear observation matrix is

\[ H_k = 4 \begin{bmatrix}
\Re\{p_{k,1}\} & \Im\{p_{k,1}\} \\
\vdots & \vdots \\
\Re\{p_{k,N_{pp}}\} & \Im\{p_{k,N_{pp}}\}
\end{bmatrix}, \tag{2.3.6} \]
the state vector is

$$x_k = \begin{bmatrix} R\{E_k\} \\ I\{E_k\} \end{bmatrix},$$

(2.3.7)

and the measurement noise vector is

$$n_k = \begin{bmatrix} n_{k,1} \\ \vdots \\ n_{k,N_{\text{pp}}} \end{bmatrix}.$$

(2.3.8)

The best estimate from processing a single batch of measurements is derived in Appendix A.1 and found to be

$$\hat{x}_k = (H_k^T R_k^{-1} H_k)^{-1} H_k^T R_k^{-1} z_k,$$

(2.3.9)

The weight matrix $R_k^{-1}$ is the inverse of the measurement noise covariance matrix,

$$R_k = E[n_k n_k^T],$$

(2.3.10)

where $E[\cdot]$ gives the expectation value. If the probe amplitudes are all equal, the measurement noise covariance becomes $R_k = r_k I$, where $r_k$ is a scalar. The batch process estimate of the field’s real and imaginary parts thus reduces to

$$\hat{x}_k = (H_k^T H_k)^{-1} H_k^T z_k$$

$$= H_k^L z_k,$$

(2.3.11)

where the superscript $L$ denotes the left pseudo-inverse. There must be at least two probe pairs for $H^L$ to be able to have full rank (2 in this case). In the special case of two probe pairs, $H^L$ reduces to $H^{-1}$. In §2.3.3 and §2.4, we discuss the best choices of probes to make $H^L$ full rank at each estimated image plane pixel.
In this batch process estimator (BPE), there is an implicit assumption that the wavefront is static. The estimator and controller can still create a dark hole as long as the electric field is static at the level of the contrast target over the course of a few correction iterations.

2.3.2 Utilizing the Unprobed Image

In addition to the pairs of probed images, we always take an unprobed image, \( I_k \), at each correction iteration. During a space mission, it would be wasteful to take another image just to measure the new contrast level since the next starlight estimate would provide that value. The unprobed image is required for two important calculations: measuring the amplitude of the probes and calculating the incoherent light level.

Even with the best available laboratory models, the measured and modeled probe amplitudes can have different morphologies and differ by tens of percent in magnitude. When the measurements are not dominated by noise, the empirically-based estimate of the probe amplitude,

\[
\hat{|p_{k,j}|} = \sqrt{I_{k,j}^+ + I_{k,j}^-} - I_k. \tag{2.3.12}
\]

provides a much more accurate model of the probes to include in the observation matrix. As described by Give’on et al., this technique mitigates several types of model error to enable faster, deeper correction. The phase of the probe is still calculated using the model.

Pair-wise estimation can yield a batch estimate of the incoherent light intensity at each correction iteration \( k \). The incoherent intensity estimate \( \hat{I}_{inco,k} \) at each pixel is then

\[
\hat{I}_{inco,k} = I_k - |\hat{E}_k|^2, \tag{2.3.13}
\]

where \( \hat{E}_k \) is the estimated stellar electric field. We will not derive the variance for the starlight intensity here, but from Eq. 2.3.13 we can see that the incoherent intensity batch estimate has a higher variance than a single image. Both terms in Eq. 2.3.13 are susceptible to noise sources (shot, readout, and dark current noise), and the estimated starlight intensity is
susceptible to model errors. To mitigate both model errors and measurement noise, it would be better to estimate the incoherent light recursively with a filter that can use previous data and appropriate weights on the error sources. That is the subject of Chapter 3.

2.3.3 Choice of Probes

The probes must be carefully selected for pair-wise estimation to work well. For uniform estimate quality at each pixel, the focal plane probe field $p_{k,j}$ should have uniform amplitude over the full extent of the dark hole. In addition, the phase of $p_{k,j}$ needs to be selectable for each of the probes. The commonly used method for probe selection developed by Give’on et al. [135] relies on analytical transforms, such as the Fourier transform relationship between the sinc function and the rect function. As a simple case, consider a simple optical model consisting of a DM and aperture $A$ at a pupil plane. The focal plane electric field is then a simple Fourier Transform, $\mathcal{F}\{\cdot\}$, of the pupil field. To actuate a uniform, purely imaginary probe field in a rectangle centered in the image plane, we apply sinc functions on the DM,

\[
p_k(\xi, \eta) = ic_k \mathcal{F}\{A(x, y) \text{sinc}(\pi w_\xi x) \text{sinc}(\pi w_\eta y)\}
= ic_k \mathcal{F}\{A(x, y)\} \ast \text{rect}(w_\xi \xi) \text{rect}(w_\eta \eta),
\]

(2.3.14)

where $c_k$ is the chosen probe amplitude, the imaginary value $i$ comes from the first order Taylor expansion of the DM surface, and $w_\xi$ and $w_\eta$ are the widths of the rect functions in frequency units. The convolution, denoted by $\ast$, with the nominal aperture’s Fourier transform is not an issue; it mostly has a smoothing effect, which does not adversely affect a uniform probe field. To actuate the real part of the focal plane field, the probe shape must have an antisymmetric component. Therefore, we generally multiply the sinc function by a sinusoid having a phase offset of $\theta_j$ to shift the probe field’s phase the desired amount. The
generalized probe field,

\[ p_{k,j}(\xi, \eta) = ic_k F\{A(x, y) \sin(2\pi ax + \theta_j) \sin(\pi w_\xi x) \sin(\pi w_\eta y)\} \]

\[ = ic_k F\{A(x, y)[\sin(2\pi ax) \cos(\theta_j) + \cos(2\pi ax) \sin(\theta_j)] \sin(w_\xi x) \sin(w_\eta y)\} \]

\[ = c_k F\{A(x, y)\} \ast \text{rect}(w_\xi \xi) \text{rect}(w_\eta \eta) \ast \left( \cos(\theta_j)[\delta(\xi + a) + \delta(\xi - a)] + i \sin(\theta_j)[\delta(\xi + a) - \delta(\xi - a)] \right) \]

\[ = c_k F\{A(x, y)\} \ast \text{rect}(w_\xi \xi) \text{rect}(w_\eta \eta) \ast \left[ e^{i\theta_j}[\delta(\xi + a) + e^{-i\theta_j}\delta(\xi - a)] \right] \]  \hspace{1cm} (2.3.15)

now actuates two complex rectangles offset by distance \( a \) in frequency units.

The major assumption in Eq. 2.3.15 is that the coronagraph has a negligible effect on the probe field. Because most coronagraphs do not substantially alter light off-axis from the star, the analytical selection of probes is applicable in most cases. If a numerical solution is necessary, the DM commands for a desired probe field \( p_{k,i} \) can instead be calculated by inverting the control Jacobian (similarly to Eq. 2.2.8),

\[ \Delta u_{k,j} = -(G^* G + \beta I)^{-1} \mathcal{R}\{G^* p_{k,j}\}, \]  \hspace{1cm} (2.3.16)

where \( \beta \) is an empirically found regularization term.

### 2.4 Error Analysis for Pair-wise Probing

Here we analyze the variance of the electric field estimate given bounding cases of measurement noise. The work appearing here, in §2.4, was contributed by Riggs and Kern to the paper by Groff et al.[133] We begin with the effect of stellar shot noise on the estimate, demonstrate how zodi and dark current drive us to brighter probes, and determine the order-of-magnitude limit on how bright the probes may be. We then determine what strategies
can be used to obtain the most precise and accurate electric field estimates possible. All calculations in this section are for a single correction iteration, so we temporarily drop the subscript $k$ for clarity.

### 2.4.1 Stellar Shot Noise

Here we will calculate the variance in the electric field estimate at a given pixel in the presence of only stellar shot noise. By utilizing probe shapes that actuate the entire focal plane evenly in field (such as the products of sinc and sine functions), we can safely apply these calculations to all pixels in the correction region.

We choose probes $p_j$ such that they have the same amplitude, $p = |p_j| \forall j$,

$$p_j = p e^{i \theta_j} = p \cos(\theta_j) + ip \sin(\theta_j),$$

(2.4.1)

where $\theta_j$ is an arbitrary angle. Uniformly distributing the phase of the probes, $\theta_j$, within $[0, \pi]$ to get full coverage of the real-imaginary plane, we substitute Eq. 2.4.1 into Eq. 2.3.3 to re-write the observation as

$$\begin{bmatrix}
\Delta I_1 \\
\vdots \\
\Delta I_{N_{pp}}
\end{bmatrix} = 4p \begin{bmatrix}
\cos(\alpha + \pi \frac{N_{pp} - N_{pp}}{N_{pp}}) & \sin(\alpha + \pi \frac{N_{pp} - N_{pp}}{N_{pp}}) \\
\vdots & \vdots \\
\cos(\alpha + \pi \frac{N_{pp} - 1}{N_{pp}}) & \sin(\alpha + \pi \frac{N_{pp} - 1}{N_{pp}})
\end{bmatrix}
\begin{bmatrix}
E_R \\
E_I
\end{bmatrix} + \begin{bmatrix}
n_1 \\
\vdots \\
n_{N_{pp}}
\end{bmatrix},$$

(2.4.2)

where $E_R$ and $E_I$ are shorthand for $\mathcal{R}\{E\}$ and $\mathcal{I}\{E\}$, respectively. We include the arbitrary phase offset $\alpha$ to show that only the distribution of probe field phases matters, not their absolute values. We define $H_\theta$ such that $H = 4p H_\theta$ to obtain our new state estimate equation,

$$\hat{x} = \frac{1}{4p} (H_\theta^T H_\theta)^{-1} H_\theta^T z.$$  

(2.4.3)
With some algebra, it can be shown that

\[
H_T H_\theta = \begin{bmatrix}
\sum_{j=0}^{N_{pp}-1} \cos^2 \left( \alpha + \pi \frac{N_{pp}-j}{N_{pp}} \right) & \sum_{j=0}^{N_{pp}-1} \frac{1}{2} \sin \left( 2 \left( \alpha + \pi \frac{N_{pp}-j}{N_{pp}} \right) \right) \\
\sum_{j=0}^{N_{pp}-1} \frac{1}{2} \sin \left( 2 \left( \alpha + \pi \frac{N_{pp}-j}{N_{pp}} \right) \right) & \sum_{j=0}^{N_{pp}-1} \sin^2 \left( \alpha + \pi \frac{N_{pp}-j}{N_{pp}} \right)
\end{bmatrix} = \frac{N_{pp}}{2} \mathbb{I}.
\]  

(2.4.4)

We can then write Eq. 2.4.3 as

\[
\begin{bmatrix}
\hat{E}_R \\
\hat{E}_I
\end{bmatrix} = \frac{1}{4p} \frac{2}{N_{pp}} \begin{bmatrix}
\sum_{j=0}^{N_{pp}-1} \Delta I_j \cos \left( \alpha + \pi \frac{N_{pp}-j}{N_{pp}} \right) \\
\sum_{j=0}^{N_{pp}-1} \Delta I_j \sin \left( \alpha + \pi \frac{N_{pp}-j}{N_{pp}} \right)
\end{bmatrix}.
\](2.4.5)

We now want to calculate the variance \(\sigma^2\{\cdot\}\) of the real and imaginary parts of the electric field estimate, which is dependent on the variance of the \(\Delta I_j\)'s and any uncertainty in \(p\), \(\sigma^2\{p\}\). Assuming that \(\sigma^2\{\Delta I\} = \sigma^2\{\Delta I_j\} \forall j\) since the probe amplitudes are equal, we can write

\[
\begin{bmatrix}
\sigma^2\{\hat{E}_R\} \\
\sigma^2\{\hat{E}_I\}
\end{bmatrix} = \left( \frac{1}{4p} \right)^2 \left( \frac{2}{N_{pp}} \right)^2 \sigma^2\{\Delta I\} \begin{bmatrix}
\sum_{j=0}^{N_{pp}-1} \cos^2 \left( \alpha + \pi \frac{N_{pp}-j}{N_{pp}} \right) \\
\sum_{j=0}^{N_{pp}-1} \sin^2 \left( \alpha + \pi \frac{N_{pp}-j}{N_{pp}} \right)
\end{bmatrix} + \sigma^2\{p\}/p^2 \begin{bmatrix}
\hat{E}_R^2 \\
\hat{E}_I^2
\end{bmatrix}.
\]

(2.4.6)

The second term in Eq. 2.4.6 is a fractional error in \(\hat{E}_R\) and \(\hat{E}_I\), equal to the fractional error (or uncertainty) in \(p\), \(\sigma\{p\}/p\). The most likely sources of this error are DM calibration errors and DM motions not being produced as they were commanded (e.g., hysteresis, actuator coupling, electronic drift, etc.). It is likely that these systematic issues can be better understood by consistent use of the same probes and empirical calibration. The rest of this
derivation will focus on the fundamental Poisson shot noise limit, implicitly assuming that \( \sigma\{p\} \) is sufficiently smaller than \( p \) such that the first term in Eq. 2.4.6 dominates.

The fundamental limit to measurement of \( \hat{E}_R \) and \( \hat{E}_I \) comes from Poisson statistics on the intensity measurements in Eq. 2.4.6, which is rewritten (without the \( \sigma\{p\} \) term) as

\[
\sigma^2 \{ \hat{E}_R \} = \sigma^2 \{ \hat{E}_I \} = \left( \frac{1}{4p} \right)^2 \left( \frac{2}{N_{pp}} \right) \sigma^2 \{ \Delta I \}. \tag{2.4.7}
\]

Referencing Eq. 2.3.1, we evaluate Eq. 2.4.7 in the two limiting cases of large (\( p \gg |E| \)) and small (\( p \ll |E| \)) probes, where we write \( p = |p_j| \) since we assume the probes are of equal amplitude. In the case of large probes, Eq. 2.3.1 can be approximated as

\[
I_j \pm \sim p^2. \tag{2.4.8}
\]

The Poisson statistics of any of the intensity measurements come from the number of detected photoelectrons in that measurement (photoelectrons being assumed for most silicon-based detectors), so it is necessary to relate physical units to \( I \).

In high-contrast coronagraphy, it is common to relate the intensities measured throughout the image plane to the peak of an image with no focal plane mask, \( I_{pk} \), which is determined by integration time per intensity measurement \( t \) and a flux \( F_{pk} \) (expressed in detected photoelectrons/s per pixel) such that

\[
I_{pk} = F_{pk} t. \tag{2.4.9}
\]

The intensities at each image plane locations where the probes are being applied can be assigned a contrast value \( C \), which can be used very generally to relate to any intensity.
measured, such that

\[ I = CI_{pk} \]
\[ = CF_{pk}t, \]  

(2.4.10)

where the units of \( I \) are explicitly photoelectrons. In general, we assume that \( t \) is the same for every intensity measurement in the probe sequence, and that \( I_{pk} \) is unchanged by the probe DM settings (which is accurate on the order of \( p^2/I_{pk} \)), so that only \( C \) changes through the probing sequence.

To determine the variances relevant to Eq. 2.4.7, we define \( C_p \) such that

\[ p^2 = C_p F_{pk}t, \]  

(2.4.11)

where \( C_p \) has a contrast interpretation; e.g., \( C_p = 10^{-8} \) corresponds to probes that add \( 10^{-8} \) intensity in the image plane. With this formulation, the units of \( p^2 \) are photoelectrons. For an individual intensity measurement, Poisson statistics dictate that the variance in photoelectrons\(^2\) is equal to the expected number of photoelectrons. For large probes, we have

\[ \sigma^2\{I\} \sim p^2, \]  

(2.4.12)

with the units of both sides of this equation are photoelectrons\(^2\). Since any \( \Delta I \) is the difference of two intensities, each with variance \( \sigma^2\{I\} \), the variance of a differential measurement is given by

\[ \sigma^2\{\Delta I\} \sim 2p^2. \]  

(2.4.13)

Revisiting Eq. 2.4.7 for large probes, the variance of one component of the field in units of
photoelectrons is given by

$$\sigma^2\{\hat{E}_R\} = \sigma^2\{\hat{E}_I\} \sim \left(\frac{1}{4p}\right)^2 \left(\frac{2}{N_{pp}}\right)2p^2 = \frac{1}{4N_{pp}}.$$  \hspace{1cm} (2.4.14)

The photoelectron unit is difficult for the notation to carry succinctly because photoelectrons are mathematically dimensionless, as Poisson statistics are counting statistics that apply only to the number of detected photoelectrons, not to the accumulation of electric charge. The key to following these units in Eq. 2.4.14 is that the $(1/4p)^2$ term carries $1/p^2$ with units $1/$photoelectrons, while the variance term $2p^2$ carries units of photoelectrons, so that their product yields the final answer in photoelectrons. The photoelectron normalization implicitly carries the integration time $t$ and flux $F_{pk}$ used to give units to $p^2$ for Poisson statistics in Eq. 2.4.11.

It is instructive to define a new term, a dimensionless noise-equivalent contrast (NEC), as one metric of the noise in the final electric field estimate. The electric field estimates, $\hat{E}_R$ and $\hat{E}_I$, are expressed in units common to the measurement used in the probe sequence, so the appropriate normalization for $\hat{E}_R^2$ and $\hat{E}_I^2$ is the same as for the intensity image, $I_{pk}$ from Eq. 2.4.9. Analogous to Eq. 2.4.10, NEC is defined such that

$$\text{NEC} = \frac{(\sigma^2\{\hat{E}_R\} + \sigma^2\{\hat{E}_I\})}{I_{pk}} = \frac{1}{2N_{pp}F_{pk}t},$$ \hspace{1cm} (2.4.15)

where we recall from Eq. 2.4.14 that $1/(2N_{pp})$ carries units of photoelectrons, so that NEC is truly dimensionless (i.e., a ratio).

$\hat{E}_R$ and $\hat{E}_I$ are computed from a sequence of probe images, comprising $N_{pp}$ pairs. Each pair is taken with two images using integration time $t$. We define the total time for the entire
probe sequence to be
\[ t_{\text{tot}} = 2N_{pp}t, \] (2.4.16)
allowing us to write the noise-equivalent contrast as
\[ \text{NEC} = \frac{1}{F_{pk}t_{\text{tot}}}. \] (2.4.17)

A sample calculation of this would be where \( t_{\text{tot}} \) is chosen to be long enough that the peak intensity reaches \( 10^9 \) photoelectrons. This \( t_{\text{tot}} \) time is broken into \( 2N_{pp} \) separate integrations with time \( t \), with DM modulation in between images, allowing a calculation of \( \hat{E}_R \) and \( \hat{E}_I \). The NEC is then \( 10^{-9} \). In other words, \( t_{\text{tot}} \) is the time required to reach SNR=1 at a contrast level defined by NEC.

Wavefront control in particular offers an excellent use of NEC with an interpretation in terms of contrast. Assuming a contrast of \( |E|^2/I_{pk} = 10^{-7} \), and a set of large amplitude probes that produce NEC = \( 10^{-9} \), the measured field will have real and imaginary components given by \( \hat{E}_R = \mathcal{R}\{E\} + \Delta_R \) and \( \hat{E}_I = \mathcal{I}\{E\} + \Delta_I \) with an expectation of \( \mathbb{E}[\Delta_R^2 + \Delta_I^2]/I_{pk} = \text{NEC} \). If wavefront control operates on the electric field and removes the estimated field, \( \hat{E}_R + i\hat{E}_I \), then the residual field is \( -\Delta_R - i\Delta_I \), such that the contrast after correction has an expectation value \( \mathbb{E}[(\Delta_R^2 + \Delta_I^2)/I_{pk}] = \text{NEC} \). Thus, the noise equivalent contrast is the limit to how far wavefront control can reduce the contrast for a given electric field estimate.

The form of Eq. 2.4.17 is very different than the Poisson noise associated with measurement of the contrast itself, in that the noise does not depend on \( |E| \). The relevant comparison is that the electric field variance is
\[ \text{NEC} = \mathbb{E}\left[\left|\hat{E}_R + i\hat{E}_I - (E_R + iE_I)\right|^2\right]/(F_{pk}t_{\text{tot}}), \] (2.4.18)
which is distinct from the standard deviation of measured intensity \( \sigma(\hat{E}_R^2 + \hat{E}_I^2)/I_{pk} \). The
calculation of $\sigma\{\hat{E}_R^2 + \hat{E}_I^2\}$ is algebraically laborious, but leads to an answer whose variance is twice the direct-measurement Poisson limit,

$$\sigma^2\{\hat{E}_R^2 + \hat{E}_I^2\} = 2(2N_{pp}|E|^2). \tag{2.4.19}$$

Integrating over the total time $t_{tot}$, the standard deviation is given in contrast as

$$\sigma\{\hat{E}_R^2 + \hat{E}_I^2\}/(F_{pk}t_{tot}) = \sqrt{2(2N_{pp}|E|^2)}/(F_{pk}t_{tot}). \tag{2.4.20}$$

In other words, relying on a sequence of large probe measurements to determine $\hat{E}_R$ and $\hat{E}_I$ allows accurate determination of the complex nature of $E$ independent of the brightness $|E|^2$, but poorer determination of $|E|^2$ than by simply measuring the unprobed intensity itself.

Another observation about the NEC is worth revisiting. While the NEC limit from Poisson statistics is a fundamental limit, the fractional error $\sigma\{p\}/p$ from Eq. 2.4.6 also affects the maximum improvement from a single probe estimate. For example, contrast improvements by more than a factor of 10, such as the previous example of a factor of 100 from $10^{-7}$ to $10^{-9}$, may not be possible if there are 10% errors on $p$.

### 2.4.2 Stellar Shot Noise Using Small Probe Intensities

Section 2.4.1 dealt with large probe intensities, where $p^2 \gg |E|^2$. While the “intermediate” case where $p^2 \sim |E|^2$ does not have much algebraic simplicity, the case of small probes, $p^2 \ll |E|^2$ is simple to express. For small probes, the individual measured intensities, and therefore their Poisson noise terms, are dominated by $|E|^2$. Based on Eq. 2.3.1 and analogous to Eq. 2.4.8, the probed image intensity is approximated as

$$I_{j\pm} \sim |E|^2. \tag{2.4.21}$$
Following the same logic as with the large probes, the noise equivalent contrast NEC for small probes becomes

$$\text{NEC} = \frac{1}{F_{pk} t_{tot}} \frac{|E|^2}{p^2}. \tag{2.4.22}$$

Since, by construction, small probes have $|E|^2/p^2 \gg 1$, this NEC is far larger than the corresponding large-probe NEC.

### 2.4.3 Other Noise Sources

We can expect there to be other sources of measurement noise that will degrade our estimate quality. Here we include the effects of light from zodiacal or exozodiacal dust $Z$ (in units of photoelectrons, integrated over time $t$), detector dark current $D_c$ (also in units of photoelectrons and integrated over time $t$), and readout noise variance $\sigma^2_{ron}$ (in units of photoelectrons$^2$). Large probes are assumed, as they will produce the least noise in the electric field estimates. A measurement $\Delta I$ comprised of $N_{exp}$ averaged exposures per probed image, where $N_{exp}$ exposures together accumulate $t$ integration time, will have a variance

$$\sigma^2\{\Delta I\} = 2\sigma^2\{I_{j\pm}\} + 2\sigma^2\{Z\} + 2\sigma^2\{D_c\} + 2N_{exp}\sigma^2_{ron}$$

$$= 2(p^2 + Z + D_c + N_{exp}\sigma^2_{ron}). \tag{2.4.23}$$

The factor of 2 comes from the two intensity measurements, each containing $N_{exp}$ individual exposures, that form a $\Delta I$ measurement. Normalizing by $I_{pk}$, we find

$$\text{NEC} = \frac{1}{F_{pk} t_{tot}} \left(1 + \frac{Z + D_c + N_{exp}\sigma^2_{ron}}{p^2}\right). \tag{2.4.24}$$

We can now see some new properties of the variance of the electric field estimate. Increasing the probe brightness $p$ reduces the variance contribution from incoherent background (zodi, dark current) and readout noise. If probes can be made arbitrarily large, the NEC will approach the fundamental limit from stellar shot noise alone.
2.4.4 Maximum Allowable Probe Brightness

Earlier we found that brighter probes (larger $p$) reduce the measurement noise variance of the electric field estimate. However, if we make $p$ too large, higher order terms of the Taylor series neglected in Eq. 2.1.3 will become the dominant source of error in Eq. 2.3.3. Here we will find the maximum allowable intensity of a probe before nonlinear terms become significant relative to the nominal contrast level, $|E|^2$. For simplicity, we will once again evaluate a single pixel in the image plane. Rather than assuming we already know the probe $p_j$, we will treat the additive electric field change of the $j^{th}$ probe as a general value $\Delta E_j$. In this case the measurement in Eq. 2.3.3, neglecting the additive measurement noise, is rewritten as

$$
\begin{bmatrix}
\Delta I_{k,1} \\
\vdots \\
\Delta I_{k,N_{pp}}
\end{bmatrix} = 4
\begin{bmatrix}
\mathcal{R}\{\Delta E_{k,1}\} & \mathcal{I}\{\Delta E_{k,1}\} \\
\vdots & \vdots \\
\mathcal{R}\{\Delta E_{k,N_{pp}}\} & \mathcal{I}\{\Delta E_{k,N_{pp}}\}
\end{bmatrix}
\begin{bmatrix}
\mathcal{R}\{E_k\} \\
\mathcal{I}\{E_k\}
\end{bmatrix}.
$$

(2.4.25)

We will now evaluate the effect of higher order terms appearing in the observations, $z_k$, and whether we can compensate for those nonlinear terms by constructing the observation matrix, $H_k$, with the full nonlinear model. Expanding about the DM surface as we did in Eq. 2.1.3, a higher order expansion at the DM plane is given by

$$
\tilde{E}_k \approx \tilde{E}_{k-1}[1 + i\Delta \phi_k - \frac{1}{2!}\Delta \phi_k^2 - \frac{i}{3!}\Delta \phi_k^3 + \frac{1}{4!}\Delta \phi_k^4].
$$

(2.4.26)

Applying a specific positive or negative probe to the expanded field in Eq. 2.4.26, the focal plane electric field becomes

$$
E_{k,j\pm} = \mathcal{C}\{\tilde{E}_k\}
\approx \mathcal{C}\{\tilde{E}_{k-1}\} \pm G_{k-1}\Delta \phi_{k,j} - \frac{1}{2!}\mathcal{C}\{\tilde{E}_{k-1}\Delta \phi_{k,j}^2\} \pm \frac{i}{3!}\mathcal{C}\{\tilde{E}_{k-1}\Delta \phi_{k,j}^3\} + \frac{1}{4!}\mathcal{C}\{\tilde{E}\Delta \phi_{k,j}^4\}.
$$

(2.4.27)
We expect the average change in electric field $\Delta E_k$ between the positive and negative probe to be

$$
\Delta E_k = \frac{E_{k,+} - E_{k,-}}{2} \approx G_{k-1}u_{k,j} - \frac{i}{3!}C\{\tilde{E}_{k-1}\Delta\phi^3_{k,j}\}.
$$

(2.4.28)

The higher order terms in Eq. 2.4.28 can become significant compared to $G_{k-1}u_{k,j}$ for sufficiently large probe phases, $|\Delta\phi_{k,j}|$. Since these terms will be observed in the measurement, we evaluate their contribution to our measurement, $\Delta I_{k,j}$, for a single probe. Calculating $\Delta I_{k,j}$ up to fourth order in $\Delta\phi_k$, we find

$$
\Delta I_{k,j} = |E_{k,j+}|^2 - |E_{k,j-}|^2
= 4\Re\{\langle E_{k-1}, G_{k-1}u_{k,j} \rangle \}
- 2\Re\{\langle G_{k-1}u_{k,j}, C\{\tilde{E}_{k-1}\Delta\phi^2_{k,j} \} \rangle \} - \frac{2}{3}\Re\{\langle E_{k-1}, C\{i\tilde{E}_{k-1}\Delta\phi^3_{k,j} \} \rangle \}
+ \frac{1}{3}\Re\{\langle C\{\tilde{E}_{k-1}\Delta\phi^2_{k,j} \}, C\{i\tilde{E}_{k-1}\Delta\phi^3_{k,j} \} \rangle \} + \frac{1}{6}\Re\{\langle G_{k-1}u_{k,j}, C\{\tilde{E}_{k-1}\Delta\phi^4_{k,j} \} \rangle \},
$$

(2.4.29)

where $\langle \cdot, \cdot \rangle$ denotes the complex inner product. With the observation nonlinearities in place, we now determine what nonlinear terms exist on the right hand side of Eq. 2.4.25 by expanding one row of the observer. Recalling that $\Re\{A\}\Re\{B\} + \Im\{A\}\Im\{B\} = \Re\{\langle A, B \rangle \}$, the nonlinear terms from a probe that appear in the observer are given by

$$
4\Re\{\langle \Delta E_{k,j} , E_{k-1} \rangle \} = 4\Re\{\langle E_{k}, G_{k-1}u_{k,j} \rangle \} - \frac{2}{3}\Re\{\langle E_{k-1}, C\{i\tilde{E}_{k-1}\Delta\phi^3_{k,j} \} \rangle \}.
$$

(2.4.30)

Subtracting Eq. 2.4.30 from Eq. 2.4.29, the unmodeled residuals in the measurement that are not eliminated by constructing the observation matrix with a nonlinear model are given by

$$
\Delta I_{k,j,\text{error}} = -2\Re\{\langle G_{k-1}u_{k,j}, C\{\tilde{E}_{k-1}\Delta\phi^2_{k,j} \} \rangle \} + \frac{1}{3}\Re\{\langle C\{\tilde{E}_{k-1}\Delta\phi^2_{k,j} \}, C\{i\tilde{E}_{k-1}\Delta\phi^3_{k,j} \} \rangle \}
+ \frac{1}{6}\Re\{\langle G_{k-1}u_{k,j}, C\{\tilde{E}_{k-1}\Delta\phi^4_{k,j} \} \rangle \}.
$$

(2.4.31)
By constructing the observation matrix, $H_k$, from the full nonlinear model we do indeed eliminate some of the dominant nonlinear residuals in our probe images. Had we simply used $\Delta E_{k,j} = G_{k-1}u_{k,j}$ to construct $H_k$, the estimate would have been subjected to all of the nonlinear residuals in Eq. 2.4.29. However, Eq. 2.4.31 shows us that not all higher order terms are cancelled in $\Delta I_{k,j}$ by using a nonlinear model to construct $H_k$. This is because formulating the observer fundamentally requires that we define a linear system. Thus, we have mitigated errors in our estimate from unmodeled measurement nonlinearities but we have not eliminated them entirely. It is the nonlinear residuals of Eq. 2.4.31 that define the maximum allowable probe amplitude. Recalling from our development of Eq. 2.4.27 that $|G_{k-1}u_{k,j}| = |\mathcal{C}\{\tilde{E}_{k-1}\Delta \phi_{k,j}\}|$ the first term in Eq. 2.4.31 will be of order $|\Delta \phi_{k,j}^3|$ and the other two are both of order $|\Delta \phi_{k,j}|^5$. Assuming that the fifth order terms are much smaller than the third order term, Eq. 2.4.31 can be approximated as

$$\Delta I_{k,j,\text{error}} \approx -2\mathcal{R}\{\langle G_{k-1}u_{k,j}, \mathcal{C}\{\tilde{E}_{k-1}\Delta \phi_{k,j}^2\} \rangle\}. \quad (2.4.32)$$

We now wish to determine the contrast level at which the unmodeled nonlinear residuals of Eq. 2.4.32 begin to dominate our estimate for a given probe amplitude, $p_{k,j} \approx |G_{k-1}u_{k,j}|$. Although $\mathcal{C}\{\tilde{E}_{k-1}\Delta \phi_{k,j}^2\}$ does not simplify in terms of $G_{k-1}$ and $u_{k,j}$, we approximate it as $\mathcal{C}\{\tilde{E}_{k-1}\Delta \phi_{k,j}^2\} \approx |G_{k-1}u_{k,j}|^2$ to get an order of magnitude estimate of this floor. Thus, the dominant nonlinear error in our estimate scales approximately as

$$\Delta I_{k,j,\text{error}} \approx |G_{k-1}u_{k,j}|^3 = |\mathcal{C}\{\tilde{E}_{k-1}\Delta \phi_{k,j}\}|^3. \quad (2.4.33)$$

If the peak contrast of the probes (assuming all $N_{pp}$ probes have equal amplitude) in the $k^{th}$ iteration is $10^{-4}$, $|G_{k-1}u_{k,j}| \approx 10^{-2}$ and our unmodeled nonlinear errors are at about the $10^{-6}$ contrast level. We therefore cannot expect to control below that. Solving in the other direction, if we are probing a dark hole at a contrast of $10^{-9}$, the maximum probe contrast is about $10^{-6}$. It is worth mentioning that uncertainties in the probe shape will result in
imperfect cancellation of nonlinearities in the observation. The effect of probe uncertainty is minimized by measuring the probe amplitude directly as in Eq. 2.3.12, but its phase must still be obtained from the model. One can use Eq. 2.4.33 as a starting point for wavefront correction simulations to hone in on the maximum allowable brightness of the probe.

2.5 Recursive Pair-wise Estimation with a Kalman Filter

The BPE from §2.3 is simple to implement, but it utilizes only the most recent set of images. By also including the previous state estimate and knowledge of the optical model in the estimator, one can obtain a more precise state estimate. The Kalman filter (KF)[137] is the optimal recursive estimator for linear systems and was first implemented for FPWC by Groff and Kasdin.[83] We derive the KF in Appendix A.2, and here we summarize the key parts of the KF.

2.5.1 Constructing the Pair-wise Kalman Filter

The definitions of the state vector $x_k$, measurement vector $z_k$, measurement noise vector $n_k$, and observation matrix $H_k$ are the same here as for the BPE in §2.3.1. We define the linear, dynamic model of the optical system as

$$x_k = \Phi x_{k-1} + \Gamma u_{k-1} + \Lambda w_{k-1}. \quad (2.5.1)$$

Equation 2.5.1 neglects the nonlinear terms of the DM phase Taylor expansion in Eq. 2.1.4, but we must use the linear formulation in order to use an additive control model in the focal plane. (The nonlinear model would require the state to be in the DM plane since we cannot back-propagate from the focal plane to the DM.) Nevertheless, as the contrast improves during FPWC and the amplitudes of the DM commands decrease, the ignored nonlinear terms
from the DM phase expansion in Eq. 2.5.1 become vanishingly small.

The propagation matrices in Eq. 2.5.1 are the state transition matrix $\Phi$, the real-valued control Jacobian $\Gamma$, and the disturbance Jacobian $\Lambda$. We list the variables used for the KF in Table 2.1. The vector $w_{k-1}$ is random process noise; it is included in the model to accommodate model errors and random, unknown disturbances.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Representation</th>
<th>Dimension</th>
</tr>
</thead>
<tbody>
<tr>
<td>State Estimate</td>
<td>$\hat{x}_k = \begin{bmatrix} \mathcal{R}{E_k} \ \mathcal{I}{E_k} \end{bmatrix}$</td>
<td>$2 \times 1$</td>
</tr>
<tr>
<td>Differenced Image</td>
<td>$z_k = \begin{bmatrix} (I_{k,1+} - I_{k,1-}) \ \vdots \ (I_{k,N_{pp}+} - I_{k,N_{pp}-}) \end{bmatrix}$</td>
<td>$N_{pp} \times 1$</td>
</tr>
<tr>
<td>Sensor Noise</td>
<td>$n_k$</td>
<td>$N_{pp} \times 1$</td>
</tr>
<tr>
<td>DM Commands</td>
<td>$u_k$</td>
<td>$N_{DM} N_{act} \times 1$</td>
</tr>
<tr>
<td>Process Noise</td>
<td>$w_k$</td>
<td>$N_{DM} N_{act} \times 1$</td>
</tr>
</tbody>
</table>

Table 2.1: Dimensions of variables for the Kalman filter at one image plane pixel. $N_{pp}$ is the number of probe pairs, $N_{DM}$ is the number of DMs, and $N_{act}$ is the number of actuators per DM.

In the discrete-time KF, the state vector and state covariance matrix are estimated in two stages at each sampling point in time. In the first stage, known as the time update and denoted by $(-)$, the linear, dynamic model propagates the state vector and state covariance matrix to the current time step. In the second stage, known as the measurement update and denoted by $(+)$, the estimates are updated to include the measurements from the current time step.

The state estimate time update, following the form of Eq. 2.5.1, simply propagates the previous estimated state as

$$\hat{x}_k(-) = \Phi \hat{x}_{k-1}(+) + \Gamma u_{k-1}, \quad (2.5.2)$$
where the term $\Lambda w_{k-1}$ is omitted because it is unknown and assumed to have zero mean. 

The state covariance time update is then

$$P_k(-) = E[(\hat{x}_k(-) - x_k)(\hat{x}_k(-) - x_k)^T]$$

$$= \Phi P_{k-1}(+) \Phi^T + Q_{k-1}, \quad (2.5.3)$$

where the process noise matrix is defined as

$$Q_{k-1} = \Lambda E[w_{k-1}w_{k-1}^T] \Lambda^T. \quad (2.5.4)$$

The last stage of the KF is to improve the estimates with new data in the measurement update equations,

$$\hat{x}_k(+) = \hat{x}_k(-) + K_k[z_k - H_k\hat{x}_k(-)] \quad (2.5.5)$$

$$P_k(+) = [I - K_kH_k]P_k(-), \quad (2.5.6)$$

where the Kalman gain $K_k$ optimally balances the weighting of the new measurements versus the model error and time updated estimate. The Kalman gain is defined as

$$K_k = P_k(-)H_k^T[H_kP_k(-)H_k^T + R_k]^{-1}, \quad (2.5.7)$$

where $R_k$ is the measurement noise covariance matrix as defined in Eq. 2.3.10.
For clarity, we group all five KF equations together as

\begin{align}
\hat{x}_k(-) &= \hat{x}(+)_{k-1} + \Gamma u_{k-1} \\
P_k(-) &= P_{k-1}(+) + Q_{k-1} \\
K_k &= P_k(-)H_k^T[H_kP_k(-)H_k^T + R_k]^{-1} \\
\hat{x}_k(+) &= \hat{x}_k(-) + K_k[z_k - H_k\hat{x}_k(-)] \\
P_k(+) &= [I - K_kH_k]P_k(-).
\end{align}

The definitions and dimensions of all the filter matrices are listed in Table 2.2. The Kalman filter formulation allows for a dynamic system via the state transition matrix. Nevertheless, the optical systems described in this thesis are assumed static, yielding $\Phi = I$. The only source of process noise is in our knowledge of the DM response, which sets $w_{k-1} = u_{k-1}$ and $\Lambda$ as the error in $\Gamma$. The error in the DM actuation is not necessarily zero-mean Gaussian, but since it is unknown we treat it as such in this derivation.

The time update equations and the Kalman gain do not utilize measurements, so the performance of the filter depends heavily on the accuracy of the linear model and assumed noise properties. We can determine the measurement noise covariance matrix accurately for our system via calibration of the camera since $R_k$ depends on the readout noise, dark current noise, and photon shot noise. The probed images have nearly equivalent contrast levels and are not correlated with each other, so the measurement noise covariance is given by

\[ R_k = r_kI, \]

where $r_k$ is a scalar. We discuss the calculation of $r_k$ in further detail in §3.3 for the extended Kalman filter.

Because the true state is unknown, the variance of the process noise is uncertain. We therefore start with educated guesses for the values of $P_0(\cdot)$ and $Q_{k-1}$ and tune them until
the best correction performance is obtained. Since the dark hole is suppressed approximately uniformly, we assign the same \( Q_{k-1} \) to all pixels. Furthermore, we assume that the variances of the real and imaginary parts of the electric field are the same and uncorrelated, which gives

\[
Q_{k-1} = q_{k-1} I,
\]

where \( q_{k-1} \) is a scalar. The real and imaginary estimates are correlated to some extent, but we have not observed improved performance with the pair-wise KF by adding off-diagonal elements to \( Q_{k-1} \).
2.5.2 Benefits of the Kalman Filter

From Eq. A.2.22 in the derivation in Appendix A.2, the covariance of the batch process estimate at a given iteration is

\[ P_{BPE,k} = (H_k^T R_k^{-1} H_k)^{-1}. \]  \hfill (2.5.15)

The alternate form of the KF’s state covariance measurement update in Eq. A.2.24 is

\[ P_k(+) = [P_k^{-1}(-) + H_k^T R_k^{-1} H_k]^{-1}. \]  \hfill (2.5.16)

As long as there is some confidence in the model (such that \( P_k^{-1}(-) \) is not zero), the state covariance for the KF is smaller than that of the BPE since \( P_k(-) \) and \( R_k \) are positive semi-definite. The KF is defined as optimal because it is the minimum variance estimator for a linear system with Gaussian noise.

Another benefit of the KF is that it can perform a measurement update with just a single probed image pair. The BPE calculation in Eq. 2.3.11 required at least two image pairs for \( H_k^T z_k \) to yield a full 2\times1 estimate of the state. One image pair for the BPE can provide only a 1\times1 estimate of a linear combination of both states. In the KF state measurement update, the measurement correction term \( K_k[z_k - H_k \hat{x}_k(-)] \) always has dimension 2\times1 for a positive number of probe pairs, even just one. The prior estimate of the state and knowledge of the system enable this full estimate update even with fewer new measurements than the number of states. A guaranteed estimate is not a guaranteed good estimate, however, so the single probe shape should be alternated after each correction iteration to modulate the real and imaginary parts of the field equally on average. Groff et al. [106, 83] first demonstrated in Princeton HCIL experiments that the 1-probe-pair KF enabled faster correction than the 2-pair BPE. Riggs et al. [86] later confirmed this result at high contrast in JPL’s HCIT.
Chapter 3

Recursive Bias and Starlight Estimation

The pair-wise batch process estimator (BPE) in §2.3 and pair-wise Kalman filter (KF) in §2.5 have been tested and proven to work at high contrast for suppressing coherent, on-axis light in the Jet Propulsion Laboratory’s High Contrast Imaging Testbed (HCIT).[135, 86] The ultimate goal of wavefront correction, however, is to image faint sources incoherent with a star such as exoplanets and disks. During wavefront correction the starlight speckles change as they are suppressed while the faint exoplanets and disks remain unchanged at the measured levels. A recursive estimator provides a methodology for separating the planet from speckles by exploiting the mutual incoherence of the light. This concept of modulating the starlight to distinguish the incoherent signal is known as coherence differential imaging (CDI) and is explored in §4.3.3 and Chapter 6. By then implementing Bayesian techniques [138] to locate any exoplanets or disks in the recursive incoherent estimate, we can better detect and characterize our science targets.

One possible approach to recursive incoherent estimation is to use another Kalman filter on the batch incoherent estimate in Eq. 2.3.13. While this method would let us use two linear estimators, it is inefficient because the estimates of the incoherent light and starlight
are interdependent and the incoherent light in the probed images is not utilized. To produce the best estimate of the incoherent light with all available data, we need to estimate the stellar electric field and incoherent intensity simultaneously in a nonlinear estimator that does not difference the images.

3.1 An Extended Kalman Filter for Pair-wise Estimation

As a first step into nonlinear focal plane wavefront estimators, we develop an extended Kalman filter (EKF) for focal plane wavefront correction (FPWC). We derive the EKF in Appendix A.3, and numerous texts (e.g., [139, 140]) discuss it in detail. In this thesis, the EKF utilizes the same probe selection as the KF does for a direct performance comparison. The EKF has the advantage that it can utilize the unprobed image in the measurement update as well, thus using all available information recursively. The work in this chapter first appeared in papers by Riggs et al.[141, 84]

3.1.1 Constructing the EKF

Since we would like to estimate the incoherent intensity $I_{\text{inco},k}$ at each pixel, we first augment the original state vector in Eq. 2.3.7 to have it become

$$
\begin{bmatrix}
R\{E_k\} \\
I\{E_k\} \\
I_{\text{inco},k}
\end{bmatrix}
$$

(3.1.1)

The previous measurement vector in Eq. 2.3.5 differences out the incoherent signal and thus cannot be used for estimating it. For the EKF, we use the same image set as in pair-wise estimation but do not difference the probed images. Then, $z_k$ at each dark hole pixel consists
of the unprobed image $I_k$ and the $2 \times N_{pp}$ probe images,

$$z_k = \begin{bmatrix} I_k & I_{k,1^+} & I_{k,1^-} & \cdots & I_{k,N_{pp}^+} & I_{k,N_{pp}^-} \end{bmatrix}^T$$

$$= h(x_k) + n_k,$$  \hspace{1cm} (3.1.2)

where $h(x_k)$ is the nonlinear measurement function. The additive measurement noise vector $n_k$, re-defined as

$$n_k = \begin{bmatrix} n_{k,unprobed} & n_{k,j^+} & n_{k,j^-} & \cdots & n_{k,N_{pp}^+} & n_{k,N_{pp}^-} \end{bmatrix}^T,$$  \hspace{1cm} (3.1.3)

consists of readout noise, photon shot noise, and dark current from each separate image. By substituting for the intensities in Eq. 3.1.2, we expand the nonlinear measurement function to be in terms of the state vector, control Jacobian, and DM commands as
\[ h(x_k) = \begin{pmatrix} |E_k|^2 + I_{inco,k} \\ |E_{k,1+}|^2 + I_{inco,k} \\ |E_{k,1-}|^2 + I_{inco,k} \\ \vdots \\ |E_{k,N_{pp}+}|^2 + I_{inco,k} \\ |E_{k,N_{pp}-}|^2 + I_{inco,k} \end{pmatrix} \]

\[ \approx \begin{pmatrix} |E_k|^2 + I_{inco,k} \\ |E_k + Gu_1|^2 + I_{inco,k} \\ |E_k - Gu_1|^2 + I_{inco,k} \\ \vdots \\ |E_k + Gu_{N_{pp}}|^2 + I_{inco,k} \\ |E_k - Gu_{N_{pp}}|^2 + I_{inco,k} \end{pmatrix} + O\{\Delta \phi_k^2\} \]

\[ \approx (\Re\{E_k\})^2 + (\Im\{E_k\})^2 + I_{inco,k} \]

\[ = (\Re\{E_k + Gu_1\})^2 + (\Im\{E_k + Gu_1\})^2 + I_{inco,k} \]

\[ = (\Re\{E_k - Gu_1\})^2 + (\Im\{E_k - Gu_1\})^2 + I_{inco,k} \]

\[ = (x_k[1])^2 + (x_k[2])^2 + x_k[3] \]

\[ = (x_k[1] + \Re\{G\}u_1)^2 + (x_k[2] + \Im\{G\}u_1)^2 + x_k[3] \]

\[ = (x_k[1] - \Re\{G\}u_1)^2 + (x_k[2] - \Im\{G\}u_1)^2 + x_k[3] \]

\[ = (x_k[1] + \Re\{G\}u_{N_{pp}})^2 + (x_k[2] + \Im\{G\}u_{N_{pp}})^2 + x_k[3] \]

\[ = (x_k[1] - \Re\{G\}u_{N_{pp}})^2 + (x_k[2] - \Im\{G\}u_{N_{pp}})^2 + x_k[3] \]

\[ + O\{\Delta \phi_k^2\}, \quad (3.1.4) \]

67
where \( x_k[m] \) represents the \( m \)-th element of vector \( x_k \), and \( u_j \) is the DM command for the \( j \)th probe shape. The term \( \mathcal{O}\{\Delta \phi_k^2\} \) is the model error from ignored higher-order terms of the DM-phase Taylor series expansion in Eq. 2.1.3. This term remains in the measurement now that the probed image pairs are not differenced.

With the EKF, one could use a nonlinear state equation based on the true nonlinear phase dependence of the electric field on the DM surface,

\[
E_k(\xi, \eta) = \mathcal{C}\{\tilde{E}_0(x, y)e^{i\phi_{k-1}(x, y)}e^{i\Delta \phi_k(x, y)}\},
\]

(3.1.5)

in the modeled propagation of the electric field, but we do not. Because we cannot sense the entire focal plane electric field and because opaque coronagraphic masks and field stops block light, we cannot directly back-propagate our electric field estimate from the focal plane to the DM-plane. Maximum a posteriori methods such as COFFEE[121] exist to solve this problem, but they are currently too slow computationally to implement in real time. The other reason for not utilizing the full nonlinear model and for not including more terms from the Taylor expansion in Eq. 2.1.3 is that the errors in our knowledge of \( \mathcal{C}\{\cdot\} \), \( \tilde{E}_0(x, y) \), \( \phi_{k-1}(x, y) \), and \( \Delta \phi_k(x, y) \) might outweigh the better accuracy from a higher-order model. We therefore use the same linear state dynamics of Eq. 2.5.1 from the KF for the EKF,

\[
x_k = \Phi x_{k-1} + \Gamma u_{k-1} + \Lambda w_{k-1},
\]

(3.1.6)

which becomes more accurate as the contrast improves and the ignored higher order terms in the DM phase Taylor expansion become smaller. Since the state estimate time update is the same as before, the time update equations for the state and state covariance have the same forms as Eqs. 2.5.8 and 2.5.9. The dimensions of the matrices have increased because of the augmented state vector. The new, third row of \( \Gamma \) is zeroes because the incoherent light is not modulated by the DMs. (Only the PSF core is observable for faint incoherent sources, and high-order wavefront correction primarily leaves the core of the PSF

68
undisturbed.) Nevertheless, process noise must still be added for the third state to account for other model errors and to prevent the filter from converging quickly on an incorrect value.

The EKF’s Kalman gain formula has the same form as for the KF,

\[ K_k = P_k(-)H_k^T[H_kP_k(-)H_k^T + R_k]^{-1}, \quad (3.1.7) \]

except that \( H_k \), now an approximation, is the partial derivative of the nonlinear observation evaluated at the latest state estimate \( \hat{x}_k(-) \),

\[ H_k = \frac{\partial h(\hat{x}_k)}{\partial \hat{x}_k} \bigg|_{\hat{x}_k = \hat{x}_k(-)} = \begin{bmatrix} 2\hat{x}[1] & 2\hat{x}[2] & 1 \\ 2(\hat{x}_k[1] + R\{G\}u_1) & 2(\hat{x}_k[2] + I\{G\}u_1) & 1 \\ \vdots & \vdots & \vdots \\ 2(\hat{x}_k[1] - R\{G\}u_{Npp}) & 2(\hat{x}_k[2] - I\{G\}u_{Npp}) & 1 \end{bmatrix} \bigg|_{\hat{x}_k = \hat{x}_k(-)} . \quad (3.1.8) \]

The state measurement update becomes

\[ \hat{x}_k(+) = \hat{x}_k(-) + K_k[z_k - h(\hat{x}_k(-))], \quad (3.1.9) \]

which is the same as for the KF in Eq. 2.5.11 except that the linear observation \( H_k \hat{x}_k(-) \) is replaced by the nonlinear observation \( h(\hat{x}_k(-)) \). The state covariance measurement update retains the same form as in Eq. 2.5.12 for the KF.
In summary, the five EKF equations for our formulation are

\[
\hat{x}_k(-) = \hat{x}(+)_{k-1} + \Gamma u_{k-1} \tag{3.1.10}
\]

\[
P_k(-) = P_{k-1}(+) + Q_{k-1} \tag{3.1.11}
\]

\[
K_k = P_k(-)H_k^T[H_kP_k(-)H_k^T + R_k]^{-1} \tag{3.1.12}
\]

\[
\hat{x}_k(+) = \hat{x}_k(-) + K_k[z_k - \hat{h}(\hat{x}_k(-))] \tag{3.1.13}
\]

\[
P_k(+) = [I - K_kH_k]P_k(-). \tag{3.1.14}
\]

We summarize the variables used in these equations in Table 3.1, and we list the matrices and their definitions in Table 3.2. Once again we assume that the system is static (except for control) such that \(\Phi = I\). The estimate is performed separately at each dark hole pixel to avoid the use of extremely large matrices.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Representation</th>
<th>Dimension</th>
</tr>
</thead>
<tbody>
<tr>
<td>State Estimate</td>
<td>(\hat{x}<em>k = \begin{bmatrix} R{E_k} \ I{E_k} \ I</em>{\text{inco},k} \end{bmatrix} )</td>
<td>(3 \times 1)</td>
</tr>
<tr>
<td>Intensity Measurements</td>
<td>(z_k)</td>
<td>(N_z \times 1)</td>
</tr>
<tr>
<td>Sensor Noise</td>
<td>(n_k)</td>
<td>(N_z \times 1)</td>
</tr>
<tr>
<td>DM Commands</td>
<td>(u_k)</td>
<td>((N_{DM}N_{\text{act}}) \times 1)</td>
</tr>
<tr>
<td>Process Noise</td>
<td>(w_k)</td>
<td>((N_{DM}N_{\text{act}}) \times 1)</td>
</tr>
</tbody>
</table>

Table 3.1: Dimensions of variables for the EKF at one image plane pixel. \(N_z = 1 + 2N_{pp}\).

### 3.1.2 Bias Error in the EKF Estimate

Solving the EKF cost function in Eq. A.3.3 via linearization of the observation produces sub-optimal estimates. The partial derivative of the cost in Eq. A.3.4 is approximate, so
<table>
<thead>
<tr>
<th>Matrix</th>
<th>Representation</th>
<th>Dimension</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linearized State Response</td>
<td>$\Phi = I$</td>
<td>$3 \times 3$</td>
</tr>
<tr>
<td>Nonlinear Observation</td>
<td>$h(x)$</td>
<td>$N_z \times 1$</td>
</tr>
<tr>
<td>Linearized Observation</td>
<td>$H_k = \left. \frac{\partial h(\hat{x}_k)}{\partial x_k} \right</td>
<td>_{\hat{x}_k=\hat{x}_k(-)}$</td>
</tr>
<tr>
<td>Linearized Complex Response</td>
<td>$G$</td>
<td>$1 \times N_{act}$</td>
</tr>
<tr>
<td>Response of Probing DM</td>
<td>$\Gamma = \begin{bmatrix} \mathcal{R}{G[1]} \ldots \mathcal{R}{G[N_{act}]} \ \mathcal{I}{G[1]} \ldots \mathcal{I}{G[N_{act}]} \ 0 \ldots 0 \end{bmatrix}$</td>
<td>$3 \times N_{act}$</td>
</tr>
<tr>
<td>Disturbance Response</td>
<td>$\Lambda = \Gamma$</td>
<td>$3 \times N_{act}$</td>
</tr>
<tr>
<td>State Covariance (Time Update)</td>
<td>$P_k(-) = E[(x_k - \hat{x}_k(-))(x_k - \hat{x}_k(-))^T]$</td>
<td>$3 \times 3$</td>
</tr>
<tr>
<td>State Covariance (Measurement Update)</td>
<td>$P_k(+) = E[(x_k - \hat{x}_k(+))(x_k - \hat{x}_k(+))^T]$</td>
<td>$3 \times 3$</td>
</tr>
<tr>
<td>Process Noise</td>
<td>$Q_k = \Lambda E[w_k w_k^T] \Lambda^T$</td>
<td>$3 \times 3$</td>
</tr>
<tr>
<td>Sensor Noise</td>
<td>$R_k = E[n_k n_k^T]$</td>
<td>$N_z \times N_z$</td>
</tr>
<tr>
<td>Kalman Gain</td>
<td>$K_k$ is computed</td>
<td>$3 \times N_z$</td>
</tr>
</tbody>
</table>

Table 3.2: Dimensions of matrices for the EKF at one image plane pixel.

Evaluating it at zero produces a value of $\hat{x}_k(\pm)$ that does not fully minimize the EKF’s cost function. Here we determine the dominant error in the observation from ignoring the higher order terms $O\{\Delta \phi_k^2\}$ in Eq. 3.1.4, and then we determine the effect on the estimate accuracy.

Following the notation from §2.1, the probed electric field at the DM expanded to second order about the DM probe phase is

$$\tilde{E}_{k,j}(x,y) = \tilde{E}_k(x,y) e^{\pm i \Delta \phi_{k,j}(x,y)}$$

$$= \tilde{E}_k(x,y) \left(1 \pm i \Delta \phi_{k,j}(x,y) - \frac{1}{2} \Delta \phi_{k,j}^2(x,y) \right), \quad (3.1.15)$$

71
which gives a focal plane electric field of

$$E_{k,j}(\xi, \eta) \approx C\{\tilde{E}_k(x, y)\} \pm C\{i\tilde{E}_k(x, y)\Delta\phi_{k,j}(x, y)\} - \frac{1}{2}C\{\tilde{E}_k(x, y)\Delta\phi_{k,j}^2(x, y)\}. \quad (3.1.16)$$

Defining $$\mathbf{e}_{k,j} = \frac{1}{2}C\{\tilde{E}_k(x, y)\Delta\phi_{k,j}^2(x, y)\}$$ for brevity, the measured focal plane intensity of a probed image is

$$I_{k,j} \approx |E_k| + |p_{k,j}|^2 \approx |E_k|^2 + |p_{k,j}|^2 \pm 2\Re\{E_k^*p_{k,j}\} + |\mathbf{e}_{k,j}|^2 - \Re\{E_k^*\mathbf{e}_{k,j}\} + I_{\text{inco},k} + n_{k,j}. \quad (3.1.17)$$

Comparing the measurement in Eq. 3.1.2 and the nonlinear observation in Eq. 3.1.4 to Eq. 3.1.17, we find the dominant error terms in the nonlinear measurement equation to be

$$\mathcal{O}\{\Delta\phi_{k,j}^2\} \approx |\mathbf{e}_{k,j}|^2 - \Re\{E_k^*\mathbf{e}_{k,j}\}. \quad (3.1.18)$$

The state estimate is now checked for a bias error from the ignored terms in the observation formula. The expectation value of the estimated state minus the true state would be zero for an unbiased estimate, $$E[\hat{x}_k - x_k] = 0$$. The bias of the time update is checked by using Eqs. 3.1.6 and 3.1.10 to find

$$E[\hat{x}_k(-) - x_k] = E[\hat{x}_{k-1}(-) + \Gamma u_{k-1} - (x_{k-1} + \Gamma u_{k-1} + \Lambda w_{k-1})]$$

$$= E[\hat{x}_{k-1}(+) - x_{k-1}] - \Lambda E[w_{k-1}]$$

$$= E[\hat{x}_{k-1}(+) - x_{k-1}]. \quad (3.1.19)$$

The measurement update bias error is still unknown except for the first correction iteration, which used the BPE. From Eqs. 2.3.4 and 2.3.11, the initial measurement update bias error
is found to be

\[ E[\hat{x}_1(+) - x_1] = E[(H_1^T H_1)^{-1} H_1^T z_1 - x_1] \]
\[ = E[(H_1^T H_1)^{-1} H_1^T (H_1 x_1 + n_1) - x_1] \]
\[ = (H_1^T H_1)^{-1} H_1^T E[n_1] \]
\[ = 0, \quad (3.1.20) \]

which, from Eq. 3.1.19, makes the first time update bias error

\[ E[\hat{x}_2(-) - x_2] = E[\hat{x}_1(+) - x_1] \]
\[ = 0. \quad (3.1.21) \]

The first use of the EKF measurement update then introduces an estimate bias of

\[ E[\hat{x}_2(+) - x_2] = E[\hat{x}_2(-) + K_2 (z_2 - h(\hat{x}_2(-))) - x_2] \]
\[ = E[\hat{x}_2(-) - x_2] + K_2 E[h(x_2) + O\{\Delta \phi_k^2\} + n_2 - h(\hat{x}_2(-)))] \]
\[ = K_2 E[h(\hat{x}_2(-)) - h(\hat{x}_2(-))] + K_2 E[O\{\Delta \phi_k^2\}] + K_2 E[n_2] \]
\[ = K_2 O\{\Delta \phi_k^2\}. \quad (3.1.22) \]

This estimate bias is passed unchanged to the next correction iteration’s time update via Eq. 3.1.19. The EKF measurement update bias at all following iterations is thus

\[ E[\hat{x}_k(+) - x_k] = E[\hat{x}_k(-) - x_k] + K_k O\{\Delta \phi_k^2\}. \quad (3.1.23) \]

The state measurement update in Eq. 3.1.13 could be improved by calculating and including the higher order terms from Eq. 3.1.18. The observation’s partial derivative in Eq. 3.1.8 cannot incorporate this change, however, so we seek another method of reducing the bias error in the EKF estimate.
3.2 The Iterated Extended Kalman Filter for Pair-wise Estimation

A large body of research already exists to address nonlinearities when implementing an EKF. Here we implement the simplest improvement, which is to iterate the EKF (known as an iterated EKF, or IEKF) to mitigate nonlinearities. We derive the IEKF in Appendix A.4. The main error in $H_k$ (and subsequently in $K_k$, $\hat{x}_k(\cdot)$, and $P_k(\cdot)$) comes from the linearization about the model-based time update $\hat{x}_k(\cdot)$, but after running the EKF (Eqs. 3.1.10-3.1.14) there is a more accurate estimate of the state available. Using $\hat{x}_k(\cdot)$ as the new linearization point for an updated $H_k$, the IEKF recomputes $K_k$, $\hat{x}_k(\cdot)$, and $P_k(\cdot)$. There is now an even better estimate of the state, and this process of iterating the EKF can be repeated until the state estimate converges on a solution. Defining the subscripts for the EKF iterations as $i = 0, 1, 2, ..., N_{it}$, we follow the notation of Gelb[139] and Simon[140] and write the IEKF equations as

$$H_{k,i} = \left. \frac{\partial h(x)}{\partial x} \right|_{x=\hat{x}_{k,i}(\cdot)}$$ (3.2.1)

$$K_{k,i} = P_{k(-)}H_{k,i}^T[H_{k,i}P_{k(-)}H_{k,i}^T + R_k]^{-1}$$ (3.2.2)

$$\hat{x}_{k,i+1}(\cdot) = \hat{x}_k(\cdot) + K_{k,i}(z_k - h(\hat{x}_{k,i}(\cdot)) - H_{k,i}[\hat{x}_k(\cdot) - \hat{x}_{k,i}(\cdot)])$$ (3.2.3)

$$P_{k,i+1}(\cdot) = [I - K_{k,i}H_{k,i}]P_k(-).$$ (3.2.4)

We initialize the IEKF with

$$\hat{x}_{k,i=0}(\cdot) = \hat{x}_k(\cdot)$$ (3.2.5)

$$P_{k,i=0}(\cdot) = P_k(-)$$ (3.2.6)

and then iterate the filter by updating Eqs. 3.2.1-3.2.4 to converge on a better state estimate at the $k^{th}$ time step. If $N_{it} = 0$, the IEKF simplifies back to the EKF.
The IEKF, being a Gauss-Newton minimization\cite{142} as shown in Appendix A.4, performs best when the nonlinearities are small and the initial EKF state estimate is already near optimal. For highly nonlinear systems, the linearizations may be too inaccurate and lead the IEKF to get stuck in a local minimum. The measurement function is mildly nonlinear in our case (because $h(x_k)$ is quadratic and the state values have magnitudes much less than 1), so we expect the IEKF gradient search to work well for reducing the bias error in the EKF estimate.

### 3.3 Sensor and Process Noise

In any variation of the Kalman filter, the noise covariance matrices $Q_{k-1}$ and $R_k$ are tuning parameters. The covariance of the model error is largely unknown, but the measurement noise covariance can be determined with measurements from the camera. The main sources of measurement noise are dark current noise, readout noise, and photon shot noise. The noise statistics apply to the number of photoelectrons, not the number of analog-digital units (ADU, or counts). Any measured values in ADU must first be converted into photoelectrons using the gain of the camera electronics in $e^-/ADU$. Following the notation §2.4, the total variance in photoelectrons² expected at each pixel is given by

$$\sigma_{k,\text{total}}^2 = N_{\exp}(\sigma^2\{I_k\} + \sigma^2\{D_c\} + \sigma^2_{\text{ron}})$$

$$= N_{\exp}(C_kF_{pk}t_{\exp} + F_Dt_{\exp} + \sigma^2_{\text{ron}}), \quad (3.3.1)$$

where $N_{\exp}$ is the number of averaged exposures, $I_k$ is the number of photoelectrons from the star, $D_c$ is the number of dark current photoelectrons, $\sigma^2_{\text{ron}}$ is the variance of the readout noise, $C_k$ is the average starlight contrast, $F_{pk}$ is the peak flux of the starlight, $t_{\exp}$ is the exposure time per frame, and $F_D$ is the dark current flux. The contrast across the dark hole in either the probed or unprobed images is relatively uniform, so we use the same matrix $R_k$ at each pixel. We still use separate values for probed or unprobed images since the probed
images are much brighter. The noise from image to image is uncorrelated, so $R_k$ is a diagonal matrix. This means that each diagonal entry $r_k$ in $R_k$ is simply the variance in dimensionless units of contrast$^2$,

$$r_k = \frac{\sigma^2_{k,\text{total}}}{(F_{pk}t_{\text{exp}})^2}. \quad (3.3.2)$$

The sensor noise matrix is then

$$R_k = \begin{bmatrix} r_{k,\text{unpr}} & 0 \\ 0 & r_{k,\text{pr}} \\ \vdots \\ 0 & 0 & r_{k,\text{pr}} \end{bmatrix}, \quad (3.3.3)$$

where $r_{k,\text{unpr}}$ is for the unprobed image and $r_{k,\text{pr}}$ is for the probed images.

We must include a nonzero process noise $Q_{k-1}$ in the covariance estimate extrapolation because of the uncertainty in the control step. Although we assume that the DM does not modulate the incoherent light at measurable levels, we must still include process noise to prevent the filter from converging quickly to an incorrect value. We scale the process noise for the third state with the average incoherent intensity estimate. Without location-specific information of the process noise, we assign the same $Q_{k-1}$ matrix to each image plane pixel. Similarly, we have no way of distinguishing if the real or imaginary parts of the starlight should have more or less model error, so we set those covariance values as equal. For each pixel, we thus use the process noise matrix

$$Q_{k-1} = \begin{bmatrix} q_0|\hat{E}_{k-1}|_{\text{avg}}^2 & 0 & 0 \\ 0 & q_0|\hat{E}_{k-1}|_{\text{avg}}^2 & 0 \\ 0 & 0 & q_3(\hat{I}_{\text{inco},k-1})_{\text{avg}}^2 \end{bmatrix}, \quad (3.3.4)$$

where $q_0$ and $q_3$ are constants used to tune the relative values. There should also be off-
diagonal elements in $Q_k$ since the model errors of the states are slightly cross-correlated, but we nevertheless keep $Q_{k-1}$ diagonal for simplicity. With good tuning in simulation, the values of $P_k(\cdot)$ tend to show no cross-correlation (zero values) among the electric field and incoherent states but a slight cross-correlation (about 10% of the diagonal values) between the real and imaginary parts of the field. Including this nonzero off-diagonal term in $Q_{k-1}$ does not change performance of the EKF and was therefore not included for any of the tests reported in later chapters.

### 3.4 Computational Complexity of the Pair-wise Estimators

Space-based observatories have limited computing power, so it is important to compare the computational complexity of the estimators proposed. In pair-wise estimation, the state can be estimated separately at each image plane pixel. This means that only small matrices are required, but the calculations must be repeated for the thousands of pixels. Here we derive estimates of the complexity for each estimator based on the number of matrix multiplications (and divisions) required per pixel. Because there are many possible methodological approaches (such as taking another image during calculations) or algorithmic approaches (such as using alternate forms of equations) to reduce the effective complexity of the estimators, we perform just a simple analysis as a starting point for comparison. We leave out the re-calculation of the observation matrix, $H_k$, for each estimator because calculating the new values of $\mathcal{R}\{G\} u_{k,j}$ and $\mathcal{I}\{G\} u_{k,j}$ is common to all the estimators and requires the square root calculation in Eq. 2.3.12. It should also be noted that the recursive estimators require more memory, but that is a separate issue from the processing speed considered here.

Table 3.3 shows the relative complexity of each estimator in terms of floating-point multiplications required per pixel. The BPE calculation is based on Eq. 2.3.11 and uses the minimum of two probe pairs. The total number of multiplications is 26.
<table>
<thead>
<tr>
<th>Estimator</th>
<th>Number of Floating-Point Multiplications</th>
</tr>
</thead>
<tbody>
<tr>
<td>BPE (2p)</td>
<td>26</td>
</tr>
<tr>
<td>KF (1p)</td>
<td>19</td>
</tr>
<tr>
<td>KF (2p)</td>
<td>46</td>
</tr>
<tr>
<td>EKF (1p)</td>
<td>150</td>
</tr>
<tr>
<td>EKF (2p)</td>
<td>360</td>
</tr>
<tr>
<td>IEKF (1p)</td>
<td>$150 + 159N_{it}$</td>
</tr>
<tr>
<td>IEKF (2p)</td>
<td>$360 + 375N_{it}$</td>
</tr>
</tbody>
</table>

Table 3.3: Number of scalar, floating-point multiplications (and divisions) required per image plane pixel for each estimator. The number of probe pairs used are shown in parentheses. $N_{it}$ is the number of EKF iterations in the IEKF.

We use the form of the KF given in Eqs. 2.5.8-2.5.12. The most computationally expensive calculation in the KF is the multiplication of $\Gamma_{u_{k-1}}$ in the time update of the state estimate, which requires $2N_{act}$ multiplications. For the WFIRST CGI, this would be about three thousand. The controller should already have performed this calculation to choose the optimal DM command, so we assume that it adds no complexity to the estimator. The 1-probe-pair KF requires 19 multiplications, so it is slightly less computationally expensive than the BPE. The 2-probe-pair KF, requiring 46 multiplications, needs less than twice the number of computations in the BPE.

The EKF as listed in Eqs. 3.1.10-3.1.14 requires many more calculations than the KF because of the longer state and measurement vectors. We find that the 1-pair EKF requires 150 multiplications and the 2-pair EKF requires 360. These numbers could be reduced by exploiting the sparsity of the matrices and not performing a brute force matrix inversion in the Kalman gain calculation. Iterating the filter requires several more calculations per EKF iteration because of the extra multiplication in Eq. 3.2.3. The IEKF thus requires on the order of a thousand multiplications at each of the few thousand pixels. Since the inversion in Eq. 2.2.4 of the controller can be precomputed once if the control Jacobian is not updated,
the linear controller and IEKF would each require on the order of a million calculations per correction iteration. We therefore conclude that the IEKF should still be feasible for the limited computing power available on a space observatory.
Chapter 4

The High Contrast Imaging Laboratory at Princeton

In this chapter, we present focal plane wavefront correction (FPWC) results from the High Contrast Imaging Laboratory (HCIL) at Princeton. The HCIL is a testbed for the rapid development and testing of new FPWC algorithms, and here we report the first tests with our extended Kalman filter (EKF) and iterated extended Kalman filter (IEKF) formulations that use pair-wise probed images. We first describe the HCIL optical configuration. We then compare the performance of the batch process estimator (BPE), Kalman filter (KF), EKF, and IEKF from various experiments. We compare correction speed with and without incoherent sources present, and we compare the amount of bias error for each estimator. Finally, we identify the limitations in our lab to be addressed in future work. Most of this chapter was first reported by Riggs et al.[84]

4.1 Layout of Princeton’s HCIL

In the HCIL, we exclusively use shaped pupil (SP) coronagraphs to generate high contrast. Our optical layout shown in Fig. 4.1 uses as few optics as possible to enable easy, stable alignment and to introduce few optical aberrations. We inject monochromatic, 635-nm
laser light directly from a single-mode fiber (launch 1) as the simulated stellar point source in the nominal experimental configuration. The 60-inch focal length of the first off-axis parabola (OAP) allows us to approximate the central part of the Gaussian beam as uniform. The collimated beam reflects off two Boston Micromachines Kilo-DMs and a fold mirror in series before reaching a transmissive, 10-mm diameter SP. The SP used in this paper is the freestanding Ripple3 design described by Belikov et al.[85] and Kasdin et al.[143] and shown in Fig. 4.2(a). The apodized PSF has a theoretical contrast of $3 \times 10^{-10}$ from $4 - 40 \lambda/D$ over symmetric $90^\circ$ sectors as shown in Fig. 4.2(b), and the empirical, uncorrected PSF in the HCIL is shown in Fig. 4.2(c). The second and final OAP focuses the beam onto a focal plane mask (FPM), which is used only as a field stop for better dynamic range on the detector. Two achromatic lenses then re-image the stopped-down PSF onto a CCD camera.

On-sky, coronagraphic images will contain sources incoherent with the star such as ex-
Figure 4.2: (a) The Ripple3 shaped pupil used in the HCIL along with (b) its normalized design PSF on a log scale. The ideal average contrast is $3 \times 10^{-10}$ from $4 - 40 \lambda/D$ over symmetric $90^\circ$ sectors. (c) The uncorrected PSF as measured in the HCIL is shown with the same spatial scaling but a shorter log stretch.

Ozodiacal or zodiacal dust, exoplanets, disks, or background stars or galaxies. To test our estimators in the presence of incoherent sources, we inject additional laser light at either of two locations on our bench as shown in the dashed boxes in Fig. 4.1. To create an exoplanet, we insert a beamsplitter in front of the original fiber source and position fiber launch 2 to reflect into the same beam path. To eliminate any dispersion or path difference errors for the star, launch 1 becomes the planet and launch 2 becomes the star. This is the simplest configuration to add an off-axis source, but the beamsplitter creates additional aberrations and stronger polarization dependence. We adjust the planet intensity by using a separate laser source, and we can re-position the planet and star by translating the fiber heads. To approximate a flat zodiacal dust signal (commonly referred to as zodi), we place another fiber tip (launch 3) approximately half a meter from the camera. The core of the expanding Gaussian beam is approximately uniform over the detector from this distance.

The polarization introduced by the flat beamsplitter should have a single orientation; non-uniform polarization from reflections off OAPs are more problematic because wavefront correction cannot fully control the different aberrations of each polarization channel simultaneously. In future work, we plan to compare achievable contrast in the HCIL with
and without a polarization filter installed to determine if polarization-dependent aberrations limit our achievable contrast. Refer to Breckenridge et al.[97] for a physical understanding of the problem and to Krist et al.[95] for an analysis of polarization-dependent errors for the WFIRST CGI.

We block most of the stellar PSF with a field stop and perform wavefront correction in a subset of the transmitted region. The FPM transmits light in symmetric areas from a radius of $5 - 11 \lambda/D$ over $90^\circ$ sectors as shown in Fig. 4.3. The nominal aberrations set an average starting contrast at around $6 \times 10^{-5}$ in monochromatic experiments with wavelength $\lambda = 635$ nanometers. In these experiments, we corrected only small, rectangular dark holes in the image plane with $\xi \in [-10, -7; 7, 10] \lambda/D$ and $\eta \in [-2, 2] \lambda/D$ as shown in Fig. 4.3. In §4.4, we investigate the dependence of achievable contrast on the dark hole area or chromatic bandwidth in the HCIL.

![Figure 4.3: An example laboratory PSF after creating small dark holes. A field stop blocks the bright, hourglass-shaped part of the PSF to provide better dynamic range on the CCD camera.](image)

### 4.2 Design of Experiments

The main purpose of the following experiments was to determine the performance of the new EKF and IEKF formulations relative to the previously validated BPE and KF. Because the true stellar electric field is unobtainable in experiment, we cannot measure the estimate
accuracy and covariance directly. Instead, we primarily compared the different estimators via correction speed, defined here as measured starlight contrast versus total exposure time. We assumed that a better estimate would enable a larger correction step. We also compared the relative averaged contrast estimates to determine if any estimator provided significantly different (i.e., biased) values. Our goals were thus to:

1) Compare the correction speed and achievable contrast for the BPE, KF, EKF, and IEKF...

   a) without injecting incoherent light.

   b) in the presence of bright, zodi-like, incoherent light.

2) Determine if using a single pair of probed images (instead of two pairs) per correction iteration increases the correction speed of the KF, EKF, or IEKF.

3) Determine if the EKF estimates have significant bias error.

4) Determine the minimum number of EKF iterations necessary for the IEKF state estimate to converge.

5) Determine if an injected planet signal can be accurately recovered from the incoherent light estimate.

To compare the relative performance of different estimators, we needed to distinguish testbed fluctuations from algorithmic performance. If we performed separate correction runs in our testbed on the same day, we could safely compare them. Otherwise, the optics could drift out of alignment and degrade the correction performance. Because each correction run took approximately half an hour, we had time for only one or two trials with each estimator when performing comparisons of all the estimators.

As derived in §2.4.4, brighter probes reduce the estimation error from readout noise and incoherent-light shot noise. If the probe amplitudes are too large, though, ignored non-linearities and model error degrade the estimate accuracy. We manually tuned the probe
amplitudes to be as large as possible without slowing correction or limiting the achievable contrast, which gave a probe intensity of about $10^{-6}$ for measured contrast values around $10^{-7}$. The ad hoc rule developed recently at JPL and used here is to set the probe intensity as $\sqrt{(10^{-5}\cdot I_k)}$, where $I_k$ is the current measured average dark hole contrast. This choice damps the probe signal to avoid nonlinearities in the model when the unprobed contrast is worse than $10^{-5}$, and it makes probes brighter than the nominal starlight when the unprobed contrast is better than $10^{-5}$ and measurement noise is more of a concern than model nonlinearities.

4.3 Experimental Results from Princeton’s HCIL

4.3.1 Experiments without Additional Incoherent Sources

Correction Speed Comparisons

In this experiment, we compared the correction curves for the four estimators (BPE, KF, EKF, and IEKF). We also tested the three recursive estimators with either one or two pairs of probed images, yielding seven total cases. The 2-probe-pair estimators used five images per correction iteration (1 unprobed and 2 pairs of probed images) and the 1-probe-pair estimators used three images per correction iteration (1 unprobed and 1 pair of probed images). The same probe shapes were applied every correction iteration for the 2-pair estimators. For the 1-pair estimators, the chosen probe pair alternated after every correction iteration to modulate the electric field sufficiently. The IEKF performed five EKF iterations.

We defined correction speed as the total number of images since all exposures had equal length. The exposure time was fixed for simpler dark frame subtraction and was chosen as the longest possible without saturating any detector pixels. The 0.120-second exposures gave a contrast resolution of $1.8\times10^{-8}$ per count. During a space mission, the early-iteration, low-contrast images would use shorter exposures to save time. Our experiment showed that
the correction curves for all the estimators overlapped above \( \approx 2 \times 10^{-6} \) contrast, so the fixed exposure time in the early iterations had little effect on the contrast curve comparisons later on. The fixed exposure time was more realistic for the later correction iterations when the last factor of 10 in contrast was being recovered.

The correction curves for the different estimators, as shown in Fig. 4.4, differentiated as they slowed down and approached their final achievable contrast levels. The BPE provided

![Figure 4.4: Comparison of correction speed for several estimators without an intentionally injected incoherent source. Exposure time for each image was constant. The Kalman filter variations were tested with both one and two pairs of probed images.](image)

the slowest overall correction. It achieved the worst final contrast and took more correction iterations to reach it. The 1-pair KF was second slowest but did eventually reach as high a contrast as the other recursive estimators. The 2-pair KF and EKF performed equally well, and the 2-pair IEKF performed slightly better than those two after 100 total images. The 1-pair EKF and IEKF started off slightly faster than the others and reached their best achievable contrast in less than 80% the number of images required for the 2-pair recursive estimators, thereby showing a benefit from fewer probed images per correction iteration.
Without repeated trials to average out variability, it is important not to draw too many conclusions from these results. In another run (not shown), for example, the 2-pair IEKF performed the same as the 2-pair KF and EKF. Nevertheless, we have observed that the BPE and 1-pair KF are always slower than the other methods and that the 1-pair EKF and IEKF are always slightly faster than all the 2-pair versions. The higher computational complexity of the recursive estimators is partially offset since they require fewer correction iterations to reach a given contrast level.

Most importantly, we have proven the viability of EKF and IEKF formulations that do not require image differencing. This enables planet signal extraction from the recursive incoherent estimate, which is demonstrated in §4.3.3. This benefit comes at no loss to correction speed with the 2-pair estimators, and the EKF and IEKF even provide faster correction when using just one probe pair per correction iteration.

After reaching the best achievable contrast, each correction curve diverged slightly (but did stay at or below about $2 \times 10^{-7}$ contrast). This behavior indicated that the modeled control Jacobian no longer matched the true system well at that contrast level. In a space mission, the system should be better characterized near the ultimate contrast value to prevent divergence. We will further characterize the HCIL with a newly-acquired Phasics interferometer. We plan to test a parameter-adaptive filter (PAF) to improve the system model, in particular the control Jacobian, during FPWC. We discuss several possible paths for the PAF in Chapter 7.

Although the 1-pair EKF and IEKF slightly outperformed the 2-pair recursive estimators, in later tests we used the 2-pair versions of the estimators. This method enabled us to re-run all the estimators over the same set of saved images (since the BPE cannot use 1 probe pair per correction iteration). We also found the 2-pair versions of the estimators to be more robust to noise, such as in the later experiment with bright, zodi-like background in the images.
Relative Accuracy of the Estimators

Here we compare the average contrast of the estimated starlight or incoherent light for each 2-pair estimator, including the IEKF with different number of EKF iterations. Comparing the estimated intensities can reveal a net bias but averages out the Gaussian noise of individual pixels. To eliminate variations in images between different correction runs, we ran the estimators on a set of saved images from a 2-pair IEKF correction run. There was no intentionally injected incoherent source. The only difference from using stored data instead of real-time data to feed the estimator was that the control signal was pre-determined, but this did not change the accuracy of the estimator. The true state values were unknown, but we could still observe the similarities and fluctuations of the different estimates during correction. The HCIL optical system was slightly out of alignment compared to the tests from Fig. 4.4, so the achievable contrast degraded to $1.8 \times 10^{-7}$.

All the estimators except the EKF gave almost exactly the same average starlight contrast values at each correction iteration, as shown in Fig. 4.5(a). The BPE exhibited more starlight estimate fluctuations than the other estimators during the last half of the correction because the BPE did not utilize previous estimates to average out measurement noise. The EKF exhibited a large bias error and mistook much of the starlight for incoherent light, as shown by the plot of the incoherent estimates at each correction iteration in Fig. 4.5(b). The EKF’s bias error decreased over time but still remained at the end of correction. The IEKF eliminated the starlight estimate bias error in one EKF iteration and the incoherent estimate bias error in two. We conclude that two EKF iterations are sufficient for convergence of the IEKF estimates. After this experiment, we stopped testing the un-iterated EKF in comparisons because of its heavily biased estimates.

Assuming that the incoherent signal was static, the BPE and KF gave more accurate (i.e., smaller) incoherent estimates in the first twenty iterations when the measured contrast was above $3 \times 10^{-7}$. Recall that the BPE and KF obtain a batch process estimate of the incoherent light by differencing the unprobed image and estimated starlight, as in Eq. 2.3.13.
Figure 4.5: (a) Comparison of the average starlight intensity estimate for each 2-probe-pair estimator. (b) Comparison of the average incoherent intensity estimate for each estimator. $N_{it}$ is the number of EKF iterations performed in the IEKF. All estimators in this case used the same saved set of correction run images to allow a direct comparison of their quality with the exact same data. The measured intensity is included for reference.

The bright probes in the earliest iterations made the first few incoherent estimates too large, and the IEKF’s recursive incoherent estimate needed until the twentieth correction iteration to forget those incorrect data points. (The BPE and KF, producing batch process incoherent estimates, did not suffer from this problem.) We therefore recommend that the IEKF not be turned on until the later correction iterations, such as when the contrast is below $10^{-6}$, to avoid a large initial bias error in the incoherent estimate. In later correction iterations as the contrast curve started to level off, the recursive estimate of the IEKF was much less variable than for the BPE and KF. This suggests that the IEKF adequately filters out read noise and produces a more accurate incoherent estimate near the end of correction.

The EKF estimates in Fig. 4.5 were significantly biased, and yet the correction speed in Fig. 4.4 clearly shows that the EKF was no slower or less robust at wavefront correction than the IEKF or KF. Somehow a net bias error in the estimate did not degrade performance, whereas random errors from noise on batch estimates did slow the correction and restrict the achievable contrast. This result contradicted our premise that a more accurate estimate
would yield faster wavefront correction. Nevertheless, we discontinued testing of the EKF since we would like an unbiased incoherent estimate for the exoplanet detection algorithms discussed in Chapter 6.

The starlight intensity estimates (except for the EKF’s) in Fig. 4.5(a) were approximately $1 \times 10^{-7}$ below the measured contrast at all correction iterations, and all the incoherent estimates converged to that difference of $1 \times 10^{-7}$ in Fig. 4.5(b). The final incoherent estimate’s structure for the IEKF in Fig. 4.6(b) matched neither that of the IEKF’s coherent estimate in Fig. 4.6(a) nor that of the nearly flat probe signal, so it was likely not an artifact of pairwise estimation. The incoherent signal was also not random, meaning it was not attributable to read noise. We conclude that the incoherent estimate was a true signal composed of stray light. For comparison, we show the sum of the coherent and incoherent intensity estimates in Fig. 4.6(c) and the measured intensity in Fig. 4.6(d). The IEKF produces the same structure and effectively removes the large amount of measurement noise in the true image.

![Images of intensity estimates](image)

Figure 4.6: Estimates of (a) the starlight intensity at $7.6 \times 10^{-8}$ average contrast and (b) the incoherent intensity at $9.9 \times 10^{-8}$ average contrast at the end of the correction run in Fig. 4.5. (c) The estimated total intensity obtained from summing (a) and (b). (d) The measured total intensity for the unprobed image. The image plane is cropped to just the dark holes.
4.3.2 Experiments in the Presence of a Flat, Incoherent Background

In this experiment, we compared the 2-probe-pair BPE, KF, and IEKF estimators in the presence of a bright, zodi-like background as shown in Fig. 4.7(a). In a coronagraphic space mission to image earth-like planets, the dark hole would be at approximately $10^{-10}$ contrast and the combined zodi and exozodi might be as bright as $10^{-8}$ contrast, depending on such factors as the location in the sky, the chosen star, and the location around the star. Since the HCIL can achieve contrasts of $1 - 2 \times 10^{-7}$, we set the incoherent background at $2.45 \times 10^{-5}$ to obtain the worst-case brightness factor of $\approx 100$ that might be expected during a mission. The average standard deviation at each pixel was $5.2 \times 10^{-7}$ contrast from readout noise and incoherent-light shot noise, several times above the achievable contrast without injecting an incoherent background. To verify the estimated starlight intensity, we acquired another image with the incoherent source turned off after each correction iteration, as in Fig. 4.7(b).

Figure 4.7: Final PSFs for the IEKF correction run when (a) the incoherent background is on and (b) is temporarily turned off for a contrast-verification measurement. The incoherent signal appears behind the field stop since fiber launch 3 was placed downstream of the field stop. We used the signal behind the center of the mask, $\xi \in [-1.5, 1.5] \lambda/D$, to calculate the true average incoherent intensity and the net standard deviation at each pixel.

We compared the correction speed for the batch process, KF, and IEKF in the presence of the flat background. The nominal incoherent signal at $\approx 1 \times 10^{-7}$ was still present with the zodi turned off, so the estimated starlight intensity should have been below the measured,
background-off intensity by that amount. We defined the nominal incoherent signal as the stray light when only the starlight laser was on. The KF and IEKF starlight estimates in Figs. 4.8(b) and 4.8(c), respectively, were indeed about $1 \times 10^{-7}$ below the measured, zodi-less intensities, while the BPE’s starlight intensity estimate in Fig. 4.8(a) was too large by about $1 \times 10^{-7}$. Figure 4.8(d) directly compares the measured intensities from Figs. 4.8(a)-(c). The IEKF correction run was slightly faster and less variable at times than the other two estimators, but the differences are small enough that they might be insignificant.

Compared to the correction curves without zodi in Fig. 4.4, the curves with bright background in Fig. 4.8 were much more erratic in their convergence and had worse achievable contrast. The fundamental contrast limit for the BPE is the noise-equivalent contrast (NEC) from Eq. 2.4.24, calculated as approximately $5.6 \times 10^{-8}$ in this case. Since the NEC was several times lower than the previously achievable contrast of $1-2 \times 10^{-7}$, we expected no change in the contrast floor, even with the higher measurement noise. Instead, all the estimators settled to contrasts in the range of $3-5 \times 10^{-7}$.

The unstable laser power for the bright background could explain the degraded performance. Over the course of a correction run, we observed the laser power drift by 1.5%, corresponding to about $4 \times 10^{-7}$ in contrast. Such a large drift in the allocation of intensity could explain the non-smooth correction curves and the lower achievable contrast for the bright background case. We could repeat this experiment with a dynamic model of the incoherent light in the KF, EKF, and IEKF formulations, but it does not represent a realistic behavior of an on-sky source during a space mission. The relevant dynamics to include are fluctuations in the input stellar wavefront from varying thermal and mechanical loads on a telescope. Such tests are the planned as future work in the HCIL and will require additional hardware to introduce these phase variations at the input pupil.
Figure 4.8: Correction runs with different estimators in the presence of $2.45 \times 10^{-5}$ contrast incoherent background light. The mean standard deviation per pixel from zodi shot noise and readout noise was at $5.2 \times 10^{-7}$ contrast. In all measurements, there was always also an incoherent signal from scattered light at about $1 \times 10^{-7}$ average contrast. (a) The BPE did well until reaching the noise floor. (b) The KF had slower initial correction speed but allowed correction below the noise floor. (c) The IEKF allowed the fastest correction to the noise floor and below. (d) A direct comparison of the measured intensity with the incoherent background temporarily turned off.

### 4.3.3 Experiments to Recover a Faint Planet

In this set of experiments, we injected a faint, off-axis point source to mimic an exoplanet, and then we attempted to recover its signal. We inserted the star-planet simulator (the beamsplitter and fiber launch 2) into the testbed and used the 2-probe-pair IEKF for each
correction run. We first ran wavefront correction without injecting a planet to determine the differences in nominal performance. The initial PSF is shown in Fig. 4.9(a), and the final, corrected PSF is shown in Fig. 4.9(c). The estimated starlight intensity curve in Fig. 4.9(b) was virtually the same as the cases without the star-planet simulator and without an injected background. The final starlight estimate in Fig. 4.9(d) was comparable to the one in Fig. 4.6(a). The beamsplitter introduced a much larger nominal incoherent signal at $3.6 \times 10^{-7}$ average contrast as shown in Fig. 4.9(e), probably because of ghosting of the starlight in the beamsplitter.

![Figure 4.9: Results from a 2-probe-pair IEKF correction run with the star-planet simulator installed but no planet injected. (a) The stopped-down, uncorrected initial image, shown on a log scale, had a contrast of $5.87 \times 10^{-5}$ in the correction region. (c) The final, corrected image had a measured average contrast of $4.4 \times 10^{-7}$ in the rectangular dark holes ($3.1 \times 10^{-7}$ in just the right-side dark hole). (b) The measured and estimated contrast curves. (d) The estimated coherent light after correction. (e) The final estimated incoherent signal, larger from the beamsplitter being placed in the system.](image)

Next we performed wavefront correction trials with an injected planet at four contrast levels, starting below the average incoherent background level and ending slightly above
it. The injected planet used a separate laser channel and was centered at approximately \((\xi, \eta) = (8.0, -0.6)\). We used the IEKF during real-time correction, saved those images, and re-used them for the KF trials for a more direct estimator comparison. This strategy eliminated variations between trials from noise, hardware, or the controller.

We compared three different techniques to recover the incoherent planet signal from the images. The first was simple PSF subtraction (PS). After a correction run, we took one image with the planet laser on and another with it off. Subtracting these two images yielded the PS estimate. The second technique defined the planet signal as the batch process incoherent estimate (BPIE) from Eq. 2.3.13, where the KF supplied the estimated starlight intensity, \(|E_k|^2\). We included the BPIE because it utilized the concept of coherence diversity (i.e., modulating the stellar electric field to distinguish the incoherent signal) without requiring the more complicated IEKF. The final method used the recursive incoherent estimate (RIE) from the IEKF as the planet signal. To isolate the planet signal for this analysis, we subtracted the IEKF’s best estimate of the nominal incoherent signal, as shown in Fig. 4.9(e), from the BPIEs and RIEs. Because the incoherent background is unlikely to be fully subtractable in this manner during a space mission, the analysis in this section represents only a best-case scenario. Any non-uniformities or asymmetries in the zodiacal or exozodiacal light would make the background more difficult to subtract.

We quantified the quality of the planet signal with two metrics: the accuracy of the planet’s contrast estimate and the two-dimensional (2D) correlation of the planet signal to the expected PSF. The contrast estimate was calculated by translating the normalized, on-axis PSF from Fig. 4.2(c) to the planet’s location, as shown in Fig. 4.10, and then scaling the template PSF’s contrast to fit the planet signal. The 2D correlation, \(C\), which quantitatively compares the morphology of the signals, was calculated by correlating the template PSF,
Figure 4.10: Normalized, on-axis PSF shifted to the planet location for use as a PSF template for the planet. Only the region within the full-width at half maximum (shown as the dotted line) was used to avoid fitting to noise.

\[ I_{\text{temp}} \] to the extracted planet signal, \( \hat{I}_{\text{planet}} \),

\[
C = \frac{\sum_{s=1}^{N_{\text{FWHM}}} (I_{\text{temp}}(s) \hat{I}_{\text{planet}}(s))}{\sqrt{\left(\sum_{s=1}^{N_{\text{FWHM}}} I_{\text{temp}}(s)^2\right)\left(\sum_{s=1}^{N_{\text{FWHM}}} \hat{I}_{\text{planet}}(s)^2\right)}}, \tag{4.3.1}
\]

where \( s \) is the pixel index. To avoid fitting to noise for both of these metrics, we used only the \( N_{\text{FWHM}} \) pixels located within the FWHM of the template PSF.

Figure 4.11 shows the planet signal obtained from each routine at each contrast level. The first row displays the planet-only images, which are the averages of ten exposures with the star laser off. The contrast values reported above this row yielded the least-squares fit of the normalized template PSF to the given signal. The average readout noise per pixel in these 10-exposure-averaged images was \( 2.8 \times 10^{-8} \) contrast. The second, third, and fourth rows show the planet signals obtained from the PS, BPIE, and RIE methods, respectively. Upon visual inspection, the RIE produced the least noisy planet signals and the best chance of a detection for the faintest planet setting.

For the planet signals in Fig. 4.11, we show the corresponding planet contrast estimates in Fig. 4.12(a) and correlation values in Fig. 4.12(b). There were not enough data points to determine if PS or the BPIE was more accurate than the other for either metric. The RIE
Figure 4.11: Planet signals obtained from the different techniques at the end of correction. Only the right dark hole region is shown. Each column is for a different planet contrast level. The first row is the planet PSF measured by averaging ten images with the starlight laser off.

The previous analysis used only the final estimates and images from the correction runs. That is the optimal strategy for PSF subtraction, which needs a dimmer dark hole to reduce stellar shot noise in the planet signal, but it might be unnecessary for the BPIE and RIE. Therefore, we calculated the BPIE and RIE after each correction iteration to determine how early they could estimate the planet accurately. At the time of the experiment we did not...
Figure 4.12: (a) Planet contrast estimates and (b) 2D PSF correlation values corresponding to the estimated planet signals in Fig. 4.11.

think to obtain an extra image with the planet off after each correction iteration, so PS could not be included in this comparison.

Figures 4.13(a) and 4.13(b) show the BPIE and RIE correlation values, respectively. Both

Figure 4.13: Correlation between the planet signal and the template PSF after each correction iteration for (a) the batch process incoherent estimate and (b) the recursive incoherent estimate. Each line corresponds to a different injected planet brightness. Except in the case of the faintest planet, the correlation settles near its final value by about the fifth correction iteration.
methods reached their final values by about the fifth correction iteration, which corresponded to the dark hole reaching approximately $2 \times 10^{-6}$ contrast. The exception was for the faintest planet intensity, for which the correlation values showed much higher variability. For each of the four planet settings, the average RIE correlation was significantly higher and had less variability over time. The faintest two planets merited the most attention as the most difficult ones to detect in a space mission. For the $2.0 \times 10^{-7}$ contrast planet in correction iterations 5-50, the mean correlation was 77% for the BPIE and 92% for the RIE. Similarly for the faintest planet, the mean correlation was 37% for the BPIE and 70% for the RIE. The RIE was thus a much better tool for detecting faint planets than the BPIE.

Figures 4.14(a) and 4.14(b) show the BPIE and RIE planet contrast values, respectively. The measured contrast levels are shown as dotted lines to reveal biases in the estimates.

![Figure 4.14: Estimated planet contrast after each correction iteration for (a) the batch process incoherent estimate and (b) the recursive incoherent estimate. The measured contrast of the planet is shown for comparison as a dotted line. The BPIE planet contrast values had more variability than the RIE contrasts. The BPIE values settled below the measured contrast levels, while the RIE values started above the measured values before converging to them.](image)

The values settled in about five correction iterations, and the RIE contrast values showed much less variability among correction iterations compared to the BPIE values. At each planet brightness, the BPIE contrast values started near the measured value but settled
with a slightly negative bias. The RIE contrast values started positively biased then settled at the measured contrast. The initial positive bias in the RIE estimates likely arose from starting the IEKF at poor contrast levels, where nonlinearities in the model and observation are large. We previously observed this early-iteration estimate error in Fig. 4.5(b). It may therefore be beneficial to start running the IEKF at moderate contrast levels to avoid the large initial bias in the incoherent estimate.

We have demonstrated that the incoherent light estimate can be utilized for extracting the planet signal during wavefront correction. Because the RIE utilizes the whole history of images to average out noise, it gives the best planet contrast estimate and best fit of the planet PSF compared to PS and the BPIE. These results hold when the other background light can be fully subtracted or is nonexistent, neither of which is safe to assume for a space mission. Non-uniform background light makes planet detection harder for all three of these methods, but these results indicate that the RIE is best at separating starlight from incoherent light. PS may still be the best option if the dark hole speckles are stable long enough to image two different stars. However, if the dark hole does change significantly from slewing the telescope, then coherence diversity via the RIE should be the best option for detecting a companion and estimating its contrast.

Our analysis in this section used a single correction run per planet contrast level. To compute a statistically significant planet detection probability for each of the methods described, we need to perform many wavefront correction trials. We also need to include cases where there is no planet to determine the false positive rate since we had been looking only at the true positive rate of detection. Simulations offer the easiest means of performing hundreds of correction runs and generating new initial conditions for each, so in Chapter 6 we determine the planet detection capabilities in simulations of the WFIRST CGI.
4.4 Limitations of Princeton’s HCIL and Proposed Upgrades

Here we detail the current limitations in the HCIL. A space-based, high-contrast coronagraphic instrument will achieve contrast levels between $10^{-8}$ and $10^{-10}$. Since the measured contrast floor in the HCIL is currently $10^{-7}$, we are in the worse contrast regime where nonlinearities of the DM surface expansion corrupt our estimates more than would be expected in space. A major goal is therefore to push our measured contrast floor closer to $10^{-8}$, which has been demonstrated in other in-air coronagraphic testbeds.[144, 120] The other desired improvements are to make larger dark holes in broadband light at high contrast.

4.4.1 DM Surface Limitations

Aberrations on the optical surfaces degrade the achievable contrast and chromatic bandwidth of a coronagraphic instrument,[69] even with active wavefront correction. [107] As with most coronagraphs, the SP in the HCIL was optimized for a uniform (in phase and amplitude) incident wavefront. While it is possible to improve the contrast from coronagraphic masks with careful optimization of DM shapes, random aberrations in general degrade performance from the desired level. Our MEMS DMs have the largest aberrations of all the optical surfaces in the HCIL, so here we numerically characterize the performance limitations from these aberrations.

The thin DM facesheets supported by posts cannot be polished flat like other optics, so the DM manufacturing process often leaves large nominal surface aberrations. The manufacturing process also leaves behind a corrugation of the surface, commonly called “quilting.” The HCIL has used two first-generation MEMS Kilo-DMs from Boston Micromachines Corporation (BMC). The DMs have $32 \times 32$ actuator arrays and $340 \mu$m pitch between actuators. The unpowered surface of DM1 is shown in Fig. 4.15(a) and contains large amounts of astigmatism, defocus, and spherical aberrations. The surface of DM2, shown in Fig. 4.15(b), has
smaller low-frequency aberrations but a higher lip at the edges. Our power supply has a

![Image](a)

![Image](b)

Figure 4.15: The square, 1024-actuator BMC MEMS Kilo-DMs used in the HCIL. (a) The unpowered surface measurement for DM1 in the HCIL. DM1 has large astigmatism, defocus, and spherical aberrations. (b) The unpowered surface measurement for DM2. The corrugation, called “quilting,” is a result of the manufacturing process.

voltage range of 0–50 volts, which corresponds to approximately 50nm of actuator stroke for DM1 and 160 nm of stroke for DM2. This range is insufficient for flattening the DM surfaces, so we are forced to leave the nominal aberrations uncorrected. (The custom power supply cannot be replaced or upgraded; the cables and connectors are no longer manufactured or sold.)

During lab calibration, we found empirically that smaller dark holes at smaller chromatic bandwidths enabled us to achieve higher contrast in the HCIL, but we had not tested if these limitations were from the DMs. Here we simulated focal plane wavefront control in the HCIL with a perfect model, no noise sources, full knowledge of the electric field, and the only aberrations being the DM surfaces. We subtracted off the defocus and astigmatism terms from both DM surface measurements because these were (mostly, but possibly not fully) corrected in the HCIL via optical alignment; pistoning the camera position effectively eliminated defocus, and translating one of the OAPs removed the astigmatism. In Fig. 4.16, we show (a) the simulated HCIL PSF without aberrations, (b) the simulated HCIL PSF with
only the measured DM aberrations, (c) the simulated HCIL PSF with only the measured DM aberrations minus defocus and astigmatism, and (d) the measured HCIL PSF. The measured PSF in Fig. 4.16(d) indeed appears much more like Fig. 4.16(c) than Fig. 4.16(b), but there are clearly additional aberrations distorting the actual PSF.

Figure 4.16: Measured and simulated HCIL PSFs. The simulated HCIL PSF (a) without aberrations, (b) with only the DM aberrations from Fig. 4.15(b), and (c) with only the DM aberrations minus defocus and astigmatism. (d) The measured HCIL PSF. All images are shown for the region $\xi \in [-10, 10], \eta \in [-18, 18]$.

We simulated the HCIL for the four possible combinations of a small (rectangular) or large (full opening) dark hole and monochromatic or 10% broadband light. From actual testbed experiments, we have only monochromatic results. Figure 4.17 shows the results from using the small rectangular dark holes in the image plane at $\xi \in [-10, -7; 7, 10] \lambda/D$ and $\eta \in [-2, 2] \lambda/D$. This dark hole was chosen as the minimum size necessary to image the core of an off-axis PSF. The left column contains the images before correction and the right column contains the images after correction convergence. The top row shows the monochromatic laboratory images, the second row shows the simulated monochromatic images, and the third row shows the simulated 10% broadband images. The average contrast in the dark hole region is given on each image. The starting contrast in the simulation was
Figure 4.17: Measured and simulated HCIL PSFs before and after correction for the small rectangular dark hole case. The top row shows monochromatic measurements, and the bottom two rows show monochromatic simulations with the only aberrations being from the measured DM surfaces. The average contrast of the dark hole region is written on each image.

four times worse than in the true image, possibly because of the unknown aberrations on the other optics in the true system or because of modeling error of the sharp features in the DM surface measurements.

For the small dark hole case, the simulation reached $2.8 \times 10^{-9}$ contrast in monochromatic light but only $5.9 \times 10^{-7}$ contrast in 10% broadband light. This indicates that the DM aberrations are highly chromatic, since the designed SP dark hole is achromatic without aberrations. The measured monochromatic dark hole had $1.2 \times 10^{-7}$ contrast but contained incoherent light and measurement noise; just the estimated coherent light had a final contrast
of $5.4 \times 10^{-8}$. The contrast discrepancy factor of 20 between simulation and experiment is partially attributable to other aberrations in the system, measurement noise, and the expectation that air turbulence limits coronagraphic performance to about $10^{-8}$. The main contrast limitation in the laboratory (compared to the simulation) might have been that the DM surface measurements were improperly modeled for the estimator and controller. We did not realize until long after our testbed experiments that the defocus and astigmatism terms should be subtracted from the surface measurements. This means that the linearization point for the controller was incorrect, thereby degrading the achievable contrast during our laboratory experiments. Unfortunately, we cannot repeat our HCIL experiments with the corrected DM surface model until we complete separate testing of two new DMs and return the HCIL to its original configuration. To realize the full contrast improvement in the lab, we might also require a phase retrieval estimate to identify the additional aberrations causing the discrepancies between the simulated PSF in Fig. 4.16(c) and the measured one in Fig. 4.16(d). Otherwise, the controller’s linearization point could still be significantly off and artificially limit the achievable contrast.

The case of the larger dark hole over the full field stop opening is shown in Fig. 4.18 and has the same structure as the previous figure. The simulation reached $4.8 \times 10^{-8}$ contrast in monochromatic light. The monochromatic lab result at $3.8 \times 10^{-6}$ contrast was 80 times worse than the expected limit, most likely because of the previously mentioned DM surface modeling error and uncertainties in the DM actuation and other optics’ aberrations. The 10% broadband simulation, finishing at $1.7 \times 10^{-6}$ contrast, could not correct the aberrations nearly as well as in the monochromatic case.

From these simulated results, we conclude that the large aberrations on the HCIL DMs are the dominant performance limitation for broadband wavefront correction. The only way to achieve reach $10^{-8}$ contrast in the current HCIL setup would be with small dark holes in monochromatic light. The simulated contrast floor for large, monochromatic dark holes in the HCIL is $\approx 5 \times 10^{-8}$, which is reasonable but still several times worse than our target of
Because our DM power supply is irreplaceable and does not provide the necessary voltage range to flatten the current DMs, any plan to achieve better than $10^{-6}$ contrast in large, broadband dark holes in the HCIL requires the use of new DMs. In the near future, we may be able to replace the older Kilo-DMs and power supply with new ones. We are currently testing two circular, 952-actuator Kilo-DMs for BMC, and we show the flattened DM surface map measurements in Fig. 4.19. The ability to flatten these new DMs would greatly improve our correction performance. If we continue to use the older Kilo-DMs, we may be able to improve performance by using the “woofer-tweeter” concept from conventional adaptive

---

**Figure 4.18:** Measured and simulated HCIL PSFs before and after correction of the entire transmitted region. The top row shows monochromatic measurements, and the bottom two rows show monochromatic simulations with the only aberrations being from the measured DM surfaces. The average contrast of the dark hole region is written on each image.
Figure 4.19: The circular, 952-actuator MEMS Kilo-DMs from BMC that might be used for future tests in the HCIL. (a) The surface measurement for the new DM1 after flattening the surface. (b) The surface measurement for the new DM2 after flattening the surface. The remaining surface error of these DMs after flattening is just $\approx 6$ nm RMS and primarily consists of high spatial frequencies.

optics. Instead of using the woofer DM to correct for atmospheric turbulence, we would use it to correct the low spatial frequency aberrations on the nominal Kilo-DM surfaces. We could purchase a relatively inexpensive ($\approx$ $4,000-15,000$), large-stroke, low-actuator-count MEMS or piezoelectric DM. Thorlabs sells a 10mm-diameter piezoelectric DM (model number DMP40-P01) that we could place directly upstream or downstream of the current DMs without requiring any other new optics in the layout. The $\approx 15\%$ hysteresis of this piezo DM would require re-calibration of the initial wavefront (which is not currently performed in the HCIL) with every use. The $12 \times 12$-actuator Multi-DM from BMC has no hysteresis, but it does introduce more quilting and higher-order aberrations. The Multi-DM, having fewer actuators, has a smaller active mirror surface of 4.4mm across. We would have to install another set of OAPS in the HCIL to re-image the Multi-DM onto the Kilo-DMs ($\approx 10$mm across) at the proper magnification. Even if we obtain the new Kilo-DMs, we may want to add the Multi-DM and re-imaging OAPs to the HCIL layout to enable low-order wavefront sensing and control (LOWFSC) experiments. The Multi-DM should not be used for the
HCIL if its low frequency quilting creates uncorrectable, bright speckles in the dark hole region; this study is left for a future re-design of the HCIL optical layout.

### 4.4.2 DM Placement Limitations

In addition to the DM surface aberrations, the diffraction off the DM surface edges creates another limitation on contrast. The current HCIL configuration has the DMs and SP in series. Free-space propagation of the beam after the DMs creates chromatic ringing about the edges of the beam that worsens the farther the beam propagates. The first DM is approximately 1.5 meters upstream of the SP, which for a 10.88mm-square aperture at the DMs generates intense ringing. Figure 4.20(a) shows in simulation that, with DM edge ringing as the only aberration in the HCIL, the monochromatic contrast degrades from its design level of $3.4 \times 10^{-10}$ to $6.7 \times 10^{-7}$. Wavefront correction fully compensates for the monochromatic ringing as shown in Fig. 4.20(a) and reaches a sub-design-level contrast of $6 \times 10^{-11}$. Repeating the simulation in 10% broadband light, we found that the initial average contrast of $6.5 \times 10^{-7}$ did not improve with wavefront correction, reaching only $6.3 \times 10^{-7}$ contrast. Figure 4.21(a) shows the uncorrectable broadband PSF with edge ringing, and Fig. 4.21(b) shows the PSF if the DMs were over-sized sufficiently compared to the beam,
thereby giving the designed contrast of $3.4 \times 10^{-10}$.

Figure 4.21: (a) The simulated 10% broadband PSF in the HCIL with no optical aberrations but with the beam clipped by the edges of the DMs. The average contrast is $6.5 \times 10^{-7}$ and did not improve with wavefront control. (b) The simulated 10% broadband PSF in the HCIL with no optical aberrations or beam clipping provides the design contrast of $3.4 \times 10^{-10}$.

From this finding, we conclude that the HCIL optical layout must be changed to obtain large, broadband dark holes in the desired $10^{-7}$ to $10^{-8}$ contrast range. The distance between the DMs must be carefully chosen to give adequate phase and amplitude control in the image, but the distance between the DMs and the SP should be as small as possible to mitigate edge ringing. The large DM-SP separation in the HCIL was chosen for simplicity of the layout. In several other testbeds, the first DM is set as conjugate to the apodizer or SP to have effectively zero propagation between the two planes. The second DM must still be in a different plane for double-sided image correction, so the SP may need to be undersized compared to the DM. Conjugating DM1 to the SP would help performance but would require installing a new set of re-imaging OAPS in the beam path. It may be possible simply to move the DMs closer to the SP, but the wide DM and SP mounts make a compact layout difficult without blocking the beam or requiring high incidence angles. Low angles of incidence are necessary for the DMs and SP to avoid their apertures appearing narrower to the beam via the projection effect. Because we previously concluded in §4.4.1 that one or more new DMs are required for better performance the HCIL, the analysis in this section should help dictate where they are placed in a re-designed HCIL optical layout.
4.4.3 Other HCIL Limitations

Besides the DM surfaces and placement, several other limitations need to be addressed for better performance in the HCIL. In particular, we discuss stray light control, optical model improvement through calibration, and new CCD detectors.

Because we cannot discount the possibility of the nominal incoherent signal in Fig. 4.6(b) being improperly modeled coherent light, we should suppress the ambient incoherent signal below the desired contrast floor. The first, simple fix should be to better baffle the testbed, especially since most of the 100mm-diameter beam incident upon DM1 (10.88mm across) scatters around the testbed. We also see in the top-right image of Fig. 4.17 that light from the brightest speckles spreads behind the focal plane field stop and therefore also leaks into the dark hole. This extra light must be caused by ghosting in the re-imaging lenses since they are the only optics between the field stop and camera. The $\approx 1\%$ anti-reflective (AR) coating on the lenses would suffice if we corrected over the full dark hole, but since we allow a dynamic range of about $10^4$ between the brightest speckles and the dark hole, the AR coating no longer prevents ghosting from contaminating the dark hole. Since we currently cannot correct the entire image region passing through the field stop, the alternative is to replace the lenses with spherical or off-axis parabolic mirrors.

Our final achievable contrast and the speed at which we reach it are heavily dependent on our knowledge of the model since we use model-based estimation and control. We want an accurate control Jacobian $G$, which requires us to know accurately our nominal aberrations at each DM, $\tilde{E}_0(x, y)$; the propagation from the DM to the camera, $C\{\cdot\}$; and the gains and influence function of the DMs. Without these values, our linearization point is inaccurate and wavefront correction is slower and recovers a worse contrast. We have started to implement phase retrieval algorithms to determine the net effect of the aberrations in the system at the pupil plane. We would also like to use phase retrieval or our new Phasics interferometer to characterize the voltage-displacement curve for each DM actuator.

Another limiting factor is our ability to measure contrast at the level we desire. The
detector used for the experiments in this chapter, a Starlight Xpress MXV-M9, has a high read noise of 4.9 ADU (corresponding to \( \approx 7.5 \) electrons) RMS for a 40,000-ADU linear range. Because we have speckles approximately \( 10^4 \) times brighter than the dark hole, we cannot reach a read noise standard deviation below about \( 8 \times 10^{-8} \) contrast at each pixel without saturating the detector. Averaging images is the usual method of increasing SNR in an image with a high dynamic range. We did not stack images because the higher frame rate caused the camera to overheat and leave the electronic shutter open between some frames (which overexposes the image). In future experiments we will utilize our new, actively-cooled QSI RS6.1 CCD camera. It has a lower readout noise of 3 electrons in its slow readout mode; however, the smaller pixel well depth and the higher gain (selectable as either 0.17 or 0.40 electrons/ADU) give a smaller dynamic range. Nevertheless, the QSI camera can effectively provide lower-noise images by averaging frames and/or exposing longer because of the better anti-blooming feature for saturated pixels. In the near future, we may be able to obtain an ultra-low noise electron-multiplying CCD detector to better emulate a space-based coronagraph instrument.

4.5 Summary of Results from Princeton’s HCIL

Below we list our most important findings from this chapter.

- The EKF and IEKF are viable estimators for FPWC.
- The 1-probe-pair EKF and IEKF provide the fastest correction in the HCIL, which has large amounts of model uncertainty.
- The EKF estimates (both coherent and incoherent) are heavily biased, and the IEKF sufficiently reduces the bias error.
- All the pair-wise-probing estimators worked, albeit more slowly, in the presence of bright, incoherent background light.
• The recursive incoherent estimate produced a better estimate of a planet signal than PSF subtraction or the batch process incoherent estimate when several correction iterations were required.

• The incoherent estimates during correction were already near their final values very early during correction. Digging the dimmest dark hole was unnecessary for obtaining a good planet estimate from FPWC’s incoherent estimate.

• The unflattenable DM surfaces and edge ringing from our first-generation BMC DMs currently prevent the HCIL from achieving large, broadband dark holes at better than $10^{-6}$ contrast.
Chapter 5

The High Contrast Imaging Testbed at the Jet Propulsion Laboratory

After demonstrating the viability of the extended Kalman filter (EKF) and iterated extended Kalman filter (IEKF) for focal plane wavefront correction (FPWC) at Princeton in Chapter 4, we progressed to tests with more mission-like parameters in the Jet Propulsion Laboratory’s (JPL’s) High Contrast Imaging Testbed (HCIT). The HCIT is a state-of-the-art facility for advancing technology for the direct imaging of exoplanets. Through our ongoing collaboration with JPL, we have been able to validate at high contrast our shaped pupil designs and FPWC algorithms.

In this chapter, we describe our HCIT estimator comparisons in both simulation and laboratory experiments. Riggs et al. [145] published some of the simulation work in the 2015 SPIE conference proceedings, and we will publish the HCIT lab results in the 2016 SPIE conference proceedings. We begin with a description of the layout and hardware of the HCIT. We then list the goals of these tests and the experimental design. After discussing the strategy for high-fidelity optical modeling, we describe our FPWC simulations of the current coronagraph design for the Wide-Field Infrared Survey Telescope Coronagraph Instrument (WFIRST CGI). Next we detail our laboratory HCIT experiments that used an earlier
shaped pupil Lyot coronagraph (SPLC) design for the WFIRST CGI. We conclude with a discussion of the current limitations in the HCIT and with a comparison of the results from this chapter and the previous one.

### 5.1 Layout and Hardware of JPL’s HCIT

For the HCIT laboratory experiments in this chapter, we used the optical bench diagrammed in Fig. 5.1. The powered optics were all OAPs with surface figure ratings of approximately $\lambda/20$, or 30nm, peak-to-valley. Four flat mirrors served to package the layout compactly. Broadband laserlight was filtered to the desired spectral bandwidth, spatially filtered through a pinhole, and injected into the vacuum chamber via a fiber optic cable. OAP1 collimated the beam, which then propagated to DM1 and DM2 in series. DM1 defined the entrance pupil with an iris. OAPs 2 and 3 re-imaged the beam at half magnification onto the reflective

![Figure 5.1: Ray trace diagram of the optical layout used in JPL’s HCIT. Adapted from a figure by Cady et al.[146] The most important optics are denoted by gray boxes, and only the flat fold mirrors are unlabeled.](image)
SP at the next pupil plane. The SP was mounted on a stage next to a flat mirror used for viewing the entire pupil. OAP4 focused the beam onto the FPM, and OAP5 collimated the beam before it reached the Lyot stop. Finally, OAP6 focused the light onto the science camera.

The camera sat on a linear stage to enable pre-FPWC calibration, such as pupil imaging, DM actuator registration, and phase retrieval. DM actuator registration retrieved the clocking angle and translation of each DM with respect to the SP. Phase retrieval, with the FPM and Lyot stop temporarily removed from the beam path, provided the starting electric field at the DM planes and measurement of the DM actuator gains about their starting position. The DM actuator gains could not be measured as accurately with a dedicated surface-measuring device outside the vacuum chamber because humidity in the air significantly changes the response of the actuators.

Both DMs are older Xinetics models with 1mm inter-actuator pitch. DM1 has a 64×64-actuator grid and DM2 has a 32×32-actuator grid. Only a 28×28-actuator sub-array was used on each DM to prevent clipping the edge of the beam at DM2. The fused-silica facesheet on these older Xinetics DMs exhibits print-through of the actuator array (similar to the Boston Micromachines DMs used at Princeton), but recent Xinetics DMs have been polished to reduce that aberration to a negligible level.

The simulations and testbed experiments in this chapter use the SPLC developed by Neil Zimmerman and A.J. Riggs at Princeton.[50] For the simulations, we used the most recent SPLC design with the Cycle 5 AFTA pupil shown in Fig. 5.2(b) and the matching SP apodizer shown in Fig. 5.2(d). The testbed experiments used the SP apodizer shown in Fig. 5.2(c), which was designed for the earlier Cycle 1 AFTA pupil shown in Fig. 5.2(a). The later SPLC design had slightly higher throughput for off-axis sources because of the removed circular obscurations around the secondary mirror, but otherwise the two designs provided virtually identical performance.

In Fig. 5.3 by Zimmerman et al.,[50] we diagram how the SPLC generates high contrast.
First the SP apodizes the PSF to create moderate contrast in part of the image. The FPM then transmits only the dim region of the stellar PSF. After being re-collimated, much of the remaining starlight is blocked by the Lyot stop. The residual starlight is re-focused to the camera plane and is now at high contrast. Off-axis sources, such as planets, passing through the FPM opening are only partially blocked and can still be observed in the image.

The designs shown in this chapter are nicknamed “characterization” SPLCs, meaning that the high-contrast regions have a reduced azimuthal coverage for imaging an exoplanet with a known azimuthal location. In this case, the openings are only 65° on each side of the star, from $2.5 - 9.0\lambda_0/D$. Compared to a full annular search area design (called a “discovery” design), this design enables a smaller inner working angle (IWA) of $2.8\lambda_0/D$. The IWA as defined here is the closest radius to the star at which the throughput of the planet is half of its peak value in the image plane; throughput is defined as the ratio of the flux under the

---

Figure 5.2: (a) Cycle 1 pupil of the AFTA telescope. (b) Cycle 5 AFTA pupil, with the features around the secondary mirror removed. (c) SP design for the Cycle 1 SPLC by N. Zimmerman. (d) SP design for the Cycle 5 SPLC by Zimmerman et al.[50]
trying to suppress the full correctable region. Restricting the dark hole problem thus leaves greater tolerance for ... on 01/15/2016 Terms of Use: http://spiedigitallibrary.org/ss/TermsOfUse.aspx

IVb, among the circular SPLCs described in Sec. as illustrated in Fig. matched to the optimized focal plane region in the final image, of an occulting spot, the FPM is a bowtie-shaped aperture and trials are repeated for 18 and 10% bandwidths. All optimization aberrations in the propagation model.

FWHM of the planet PSF compared to the flux of the planet light entering the telescope. This SPLC design was numerically optimized to provide an 18% spectral bandwidth for the integral field spectrograph (IFS) in the WFIRST CGI.

The SPLC masks were manufactured in the Microdevices Laboratory (MDL) at JPL by Balasubramanian et al.[31] The SP apodizers were made on polished, 2-5mm-thick silicon wafer substrates. Figure 5.4(a) shows the actual SP used for the lab experiments in this chapter. The silicon wafers were coated in aluminum, and then the SP design was etched into the silicon via a cryogenic black silicon process. The black silicon is a rough surface of black needles a few microns in size and has a specular reflectivity of \( \leq 7 \times 10^{-8} \).[31] The FPM and Lyot stop were through-hole designs etched out of computer-industry standard

Figure 5.3: Diagram of the most recent SPLC design for the WFIRST CGI from Zimmerman, Riggs, et al.[50] Masks are shown in the top row, and the intensities at the central wavelength \( \lambda_0 \) are shown in the bottom row. (a) The SP apodizes the PSF, and (b) the FPM transmits only a small region of the off-axis PSF. The core of a planet PSF would go through this region mostly unblocked. (c) The Lyot stop at the next pupil plane blocks much of the remaining starlight, producing even higher contrast at (d) the final focal plane.
300-500µm-thick silicon-on-insulator (SOI) wafers. The SOI wafers allow different etching depths on the front and back, which is used to create an optical edge of only 10-50µm on the front and a negative-bias (i.e., wider) relief etch on the back. Deep reactive-ion etching (DRIE) can etch the desired outline to better than 1µm accuracy. The etched masks are coated with a layer of metal, in this case gold, to prevent transmission of light through the silicon. Figure 5.4(b) shows a microscope image of the FPM. The Lyot stop is shown in Fig. 5.4(c), mounted (left) and under the microscope (right). The free-standing Lyot stop required struts to support the central obscuration, but the small struts had a negligible impact on the contrast and throughput of the coronagraph design.

5.2 Design of Experiments

The main purpose of these tests was to compare the pair-wise FPWC estimators in a mission-like setting at high contrast and with low model error. We used hardware similar to that being planned for the WFIRST CGI, in particular Xinetics DMs and a SPLC. We omitted the EKF from the estimator comparison because we found the EKF to have too much bias error in §4.3.1.

Without the hardware or scheduling constraints of an actual testbed, we performed numerous trials in simulation with two DMs in series. We focus on correction speed comparisons of the estimators in this chapter and save the extraction of a planet signal for Chapter 6.

We had limited time to operate the HCIT at JPL, and most of that time was spent integrating the HCIT team’s calibration data into our optical model for FPWC. In the allotted time, we could not troubleshoot all the errors in the two-DM model, so we performed the estimator comparison tests with a single DM and thus obtained a single-sided dark hole. The overhead time for performing broadband correction (by measuring 2% bandpasses sequentially) was also too great given the timeframe, so our experimental results are reported only for a single 2% bandpass centered at 550nm. We performed two estimator comparison
Figure 5.4: (a) Composite picture from microscope images of the SPLC SP design in Fig. 5.2(d). (b) Microscope image of the FPM. Both (a) and (b) are from Balasubramanian et al. [31] (c) Photograph and microscope image of the Lyot stop from Cady et al. [88]

runs with fixed exposure times; the 0.5s-exposure runs gave a noise equivalent contrast (NEC) of $9.3 \times 10^{-9}$ and the 0.3-s exposure runs had a NEC of $2.1 \times 10^{-8}$. We could not use a rolling, NI-dependent exposure time because the camera was calibrated only for a few fixed exposure times. Unlike in §4.3.1 for the Princeton experiments, the HCIT was not configured to allow an intentionally injected incoherent background or planet signal in the testbed.
5.3 Simulations of JPL’s HCIT

Before performing experiments in the HCIT, we tested the performance of the pair-wise estimators in a simple, simulated model of the WFIRST CGI. Here we present the methodology and results of simulations to compare the estimator speeds in photon-limited correction runs with the second SPLC design.

5.3.1 Full and Compact Simulation Models

Our FPWC code uses two models of the optical system: a “full” model and a “compact” model. Both are shown in Fig. 5.5 next to the optical layout being simulated. The full model, used only in simulation to generate realistic images, represents the full, aberrated system. The full model incorporates expected levels of aberrations, misalignments, model...

---

Figure 5.5: (Middle) An unfolded, simple optical layout of the WFIRST CGI. The model starts at the re-imaged pupil with the DM since everything upstream is assumed to be static. (Top) Compact model propagations used in our simulation. After the angular spectrum (AS) propagations between DMs, the rest of the propagations were simple Fourier transforms (FT) between pupil and focal planes. (Bottom) Full model propagations used in our simulation. Only AS and Fresnel (Fr) propagations were used, making the full model significantly slower than the compact model. Distances and mask sizes are not to scale.
uncertainty, and noise. To include aberrations at all the optics, only Fresnel propagations are used. The collimating optics should have surface and reflectivity maps with power-spectral densities (PSDs) representative of the actual optics used. A well-known error to include in the DM model is that of the actuators’ displacement-per-voltage gains. The full model outputs simulated images, so models of the readout noise, dark current noise, and photon shot noise are included at the camera.

The compact model is used by the estimator and controller, regardless of whether real or simulated images are being used for correction. It deviates from the perfect model only by whatever aberrations or errors are known and included (such as a starting pupil plane electric field from phase retrieval). The compact model needs to be run thousands of times (once per DM actuator) for the control Jacobian calculation, so we make the compact model as fast as possible without degrading performance. The smallest possible matrix representations of the electric field and masks are used. Because aberrations at the collimating optics are unknown, Fresnel propagations are mostly unnecessary and Fourier transforms (FTs) can be used to propagate directly between pupil planes and focal planes.

For these simulations, we used a simple model of the WFIRST CGI to simplify the simulation and avoid potential export control issues with the official model from JPL. Our modeled layout in Fig. 5.5 started at the re-imaged pupil with DM1 since we assumed a static system. The beam then propagated through several OAPs and coronagraphic masks to the camera. Above the optical layout we show all the propagations performed in the compact model: two angular spectrum propagations between the DMs followed by five FTs between pupil and focal planes. Below the optical layout, we show all the propagations used in the full model: six angular spectrum propagations and four Fresnel propagations. The Fresnel transform requires just one FT but each angular spectrum propagation requires two (because it performs a convolution). In our case, the full model requires almost twice as many FTs (16 vs 9) as the compact model. The number of FTs in each model indicates its relative speed since the FT is the slowest operation performed in each.
5.3.2 Simulation Methodology

The optical layout in the simulation, shown in Fig. 5.5, included many of the key components of the HCIT and the proposed WFIRST CGI design. As a simplification and time-saving measure, the layouts were simplified to exclude many of the extra re-imaging optics and fold mirrors. Two 48×48-actuator Xinetics DMs, under-filled to 46 actuators across, were simulated.

Several stylistic choices were made for convenience. Normalized intensity (NI) of starlight was reported instead of contrast, but the average NI was less than a factor of 1.5 better than the raw contrast for this SPLC. (Contrast accounts for the attenuation of off-axis sources passing through the FPM and LS whereas NI just divides the entire image by the peak value of the unocculted star.) The chosen correction region in simulation was over the full opening of the FPM from 2.5 to 9.0λ₀/D in symmetric 65° wedges, and the reported NI was for the sub-region from the inner working angle of 2.8λ₀/D to the outer working angle of 8.8λ₀/D. The chosen photon arrival rate, giving 10⁻⁴ photons/pixel/second at 10⁻⁹ NI in the dark hole, was an order-of-magnitude estimate for a magnitude 2.4 star being imaged in 2% narrowband light with WFIRST CGI. The value was based on one calculated by Krist et al.[95] Only one correction run trial for each setting was performed. Minor variability can occur between repeated trials with a given estimator because of stochastic noise, so only large differences in performance among the trials should be treated as significant.

5.3.3 Controller Tuning

For the fastest correction, the controller should be as aggressive as possible without diverging. Otherwise, correction is unnecessarily slow and some estimators could appear better than others solely from poor controller tuning. For instance, overly damped control steps would lead to slower correction, which favors recursive techniques that build up better estimates with more data. To tune the model-based line search, we first used the empirical line search of the controller’s Lagrange multiplier with no photon noise as shown in Fig. 5.6. This
Figure 5.6: An early control iteration performed with an empirical line search. The Lagrange multiplier $\mu$, used in stroke minimization, scales as one over the square root of the DM damping parameter. For large damping (small $\mu$), the linear model matches the observation well. For less damping (larger $\mu$), the observed normalized intensity is better than the linear model predicts for a range of values before diverging and getting worse than the linear model.

This lets us determine the fastest correction rate possible for our system. This tuning approach could be implemented in an actual space observatory with calibration runs of the wavefront correction on a very bright star. It may have to be done only once, or it may have to be repeated every few weeks or months, depending on the frequency of large disturbances to the optical system. With this information of how large a correction step can be taken at a given starting contrast, a look-up table is made for the model-based, stroke-minimizing controller (discussed in §2.2) to use in later runs. We have found that the targeted contrast improvement factor just has to be near (within a factor of about 1.5) the best value for correction still to work well, so the look-up table can be simple and still perform optimally. For instance, a target contrast correction factor of 0.5 per step might be sufficient for the first few iterations and a factor of 0.8 might work until correction has completed.

### 5.3.4 Sources of Model Error and Measurement Noise

The only source of measurement noise included was photon shot noise even though the EMCCD for the WFIRST CGI will also have low levels of readout noise and dark current noise.\(^{[147, 148]}\) Including these extra noise sources would have forced us to test with specific
parameters such as stellar flux, exposure times, and number of stacked frames, whereas with just photon noise we can test the theoretical limits of performance and ignore image stacking logistics. No incoherent light sources (such as zodi, exozodi, disks, or exoplanets) were included in these simulations as the simplest starting case.

We gave all the reflective optics (OAPs, SP, and DMs) 5nm RMS surface aberrations with a frequency power law typical for manufactured optics. This gave a starting NI worse than $10^{-4}$ and did not allow correction to high-contrast because of large low-order aberrations. To mimic the phase retrieval and flattening performed in the HCIT, we mostly flattened the phase at the pupil before the SP. We left several nm RMS of low-order aberrations on the SP itself since those are the most difficult to sense and correct in the testbed. The starting NI was then about $10^{-5}$, the same approximate starting contrast observed in well-aligned testbeds. White-noise amplitude aberrations of 0.5% RMS were included on the SP mask because the custom mirror surface has the largest amplitude aberrations of all the optics. We included DM actuator gain errors of 10% RMS to match actual testbed estimates. Sidick et al. [149] provide a detailed analysis of this topic. In this chapter, we used the same set of optical aberrations for all simulations to provide a consistent simulation model for each trial.

A ubiquitous problem in coronagraphic wavefront correction modeling is that simulations can achieve much higher contrast in many fewer correction iterations than actual testbed experiments. Most simulations, such as those reported here, that include all known sources of aberrations at measured levels reach the final achievable contrast within just five to ten correction iterations, whereas tens of iterations are necessary in testbeds. Large contrast correction steps of a factor of 100 or 1000 are easy to achieve in simulation, so the electric field linearization at the DM is not the main culprit. Most HCIT experiments to date used long exposures for high NEC and thus were not limited by measurement noise, either. Model error therefore seems to be the main reason real testbeds have difficulty taking large correction steps (factors of 10 or larger) and correcting below $10^{-8}$ to $10^{-9}$ contrast.
Numerous simulations by the WFIRST CGI modeling team at JPL have indicated that most known model errors cannot explain the slowness of the correction on the actual testbed. Some issues are specific to the type of coronagraph used. Sidick et al. [150] found that including HLC FPM manufacturing defects in the full model is one method to provide more realistic simulations of correction speeds and achievable contrast level, but the simulated errors have to be much more severe than the ones known from microscope measurements. The SPLC’s through-hole FPM is simple to manufacture accurately compared to the HLC’s FPM, but the reflective SP apodizer is difficult to manufacture flat and defect free. The DM actuator influence functions are treated as uncoupled and uniform for each actuator in the model, but this assumption is known to be incorrect at some level. A new study should be performed to determine how much of the model error these two issues introduce at high contrast. Another possible culprit for the unknown model error is the pre-FPWC calibration, which is usually assumed to be perfect in the simulations. Un-sensed low-order modes (even at the $\lambda/10$ or $\lambda/20$ level) can cause enough error in the control Jacobian linearization point to impede correction performance. Additional fine tuning of the low-order aberrations in the SPLC testbed model last fall provided significantly faster correction and better achievable contrast. Further studies are necessary to determine if these or other issues are the main causes of the remaining model error.

After including all the known sources of model error, the correction speed in our simulations was still too fast compared to in the lab. The low-order SP aberrations that we added slightly slowed down FPWC. Introducing even larger aberrations made the achievable contrast worse than the levels attainable in the lab, so we instead introduced a large vertical misalignment of $0.2\lambda_0/D$ to the FPM. We know that this level of misalignment is too high, but since the true cause of model error is unknown we used the FPM misalignment to obtain the expected FPWC behavior of a lab experiment.
5.3.5 Optimal Selection of Exposure Time

To make correction as fast as possible, the exposure time at each iteration should be chosen carefully. Since higher contrast images have fewer photons, the exposure times should scale inversely with the dark hole contrast. The calibration of exposure times is empirical and depends on the properties of a specific optical system. If the exposures are too short, the estimate will be poor and control will make the contrast worse. If the exposures are too long, correction will take just as many iterations but more overall time. Recursive estimators effectively allow averaging of noise sources and should therefore be able to utilize shorter images than the BPE. Depending on the stellar magnitude and available computing power, there might be a separate optimization to minimize the total computational and exposure time. For these initial investigations, our only concern is the total exposure time since the photon arrival rate sets the fundamental limit on performance.

We simulated the 2-pair BPE, KF, and IEKF in monochromatic light at different NECs to compare their performance. The results are shown in Fig. 5.7. Because the initial correction steps, in this case from about $10^{-5}$ to $10^{-8}$ NI, have many more available photons, they contribute little to the total exposure time. Therefore, we started each trial at $10^{-8}$ NI. Even though the design NI of the SPLC is about $2 \times 10^{-9}$, the additional amplitude control with the DMs enables several times higher achievable NI in monochromatic simulations. For each trial, we set the exposure time per image to give a specific NEC at the last measured unprobed image. We report the NEC at each iteration instead of the dark hole SNR because the NEC is a better indicator of the achievable performance in a correction step. The NEC technically only applies to the BPE, so in the context of the other estimators we use it to indicate the exposure times used. In this section, we report the NEC as the factor multiplied with the previous measured contrast (e.g., 16 instead of $16\bar{T}_{k-1}$) for simplicity. We started with an NEC of 16 and stepped down by a factor of 2 until the exposures became too short for correction to be stable for each estimator.

The left column of Fig. 5.7 shows the average NI versus total exposure time for each
Figure 5.7: Simulated correction runs of the WFIRST CGI in monochromatic light with different NECs for the 2-probe-pair BPE, KF, and IEKF. The left column contains plots of the NI vs total exposure time, and the right column shows plots of the normalized RMS starlight estimate error. The flux rate assumes correction is being performed on a 2.4 magnitude star in 2% narrowband light. The legend applies to all plots, although the BPE case does not have trials for NECs of 1/4 or 1/8.

estimator using different NECs. The KF and IEKF gave essentially the same correction speed as the BPE for NEC ≥ 4, meaning that there was a surplus of photons in these
cases. The correction for the BPE was fastest for NEC=2 and NEC=4 and became slower for NECs of 1 or 1/2. We expected the BPE to perform worse for NEC ≤ 1 because, as discussed in §2.4, the NEC should define the achievable contrast limit for the BPE. Even though the BPE should not theoretically be able improve the NI with NEC ≤ 1, we believe it still showed some improvement because the speckles brighter than the average still had enough signal to be corrected. The KF and IEKF, being able to utilize prior data, enabled much faster correction with lower NEC values. Both the KF and IEKF provided the same correction speeds for NECs down to 1/2. At the very lowest NECs of 1/4 and 1/8, the KF gave much smoother and more stable correction curves than for the IEKF. We suspect that the IEKF was less stable at the lowest NECs because of its nonlinear observation and because there was virtually no signal in the unprobed image. The KF, using only the bright probed images in its estimate, performed much better at the lowest NEC tested. By the end of the correction runs, the best-case KF gave a factor of 4 better NI than the best-case BPE run, and the best-case IEKF gave a factor of about 3. Comparing performance along the other axis, the KF and IEKF reached the best BPE run’s NI of $8 \times 10^{-8}$ in about 3 hours versus 30 hours for the BPE—a time savings factor of 10. These are promising results for making FPWC more efficient during a space mission.

Since in simulation the true (noiseless) electric field can be referenced at each correction iteration, we also plotted the average RMS error of the starlight estimate in the right column of Fig. 5.7. We did not include the mean error values because they were scatter plots and did not reveal any definitive bias errors. We saved discussion of the incoherent estimate’s accuracy for Chapter 6. Here the estimate error is calculated as the average RMS error of the electric field estimate in the dark hole and is normalized by the true average NI for easier comparisons. The high-signal cases with NECs of 16, 8, and 4 gave essentially the same error levels for each estimator, which explains why the correction curves were also the same among the estimators in those cases. For NECs below 4, the KF gave the lowest RMS error, the IEKF the second lowest, and the BPE the most. The IEKF’s error was much higher in
the initial part of correction and required time to decrease to its steady-state value. This is
typical behavior for an IEKF and could be mitigated by starting the IEKF earlier in the full
correction run starting at poor contrast.

We performed the same tests with the 1-pair KF and IEKF to determine if correction
would benefit more from fewer images or from shorter exposures at each correction iteration.
The data are shown in Fig. 5.8. The concept of NEC does not directly apply with these

![Graphs showing NI and normalized standard deviation vs total exposure time for 1-probe-pair KF and IEKF with different NECs.](Figure 5.8: Simulated correction runs of the WFIRST CGI in monochromatic light with different NECs for the 1-probe-pair KF and IEKF. The left column contains plots of the NI vs total exposure time, and the right column shows plots of the normalized RMS starlight estimate error. The legend applies to all plots.)

1-pair estimators (since the BPE requires 2 or more probe pairs), so the reported NECs here
just indicate that the same exposure time formula applies from the previous 2-pair trials.
Because each 1-pair correction iteration took 3/5 the time of the same 2-pair iteration, the NEC=n trials for the 2-pair filters corresponded closest to the NEC=n/2 trials for the 1-pair filters in terms of total signal obtained for the estimate. Indeed, the 1-pair IEKF correction started becoming erratic for NEC≤ 1 whereas the 2-pair IEKF starting becoming erratic for NEC≤1/2. The 1-pair KF correction at NEC=1/4 was slightly slower than for the 2-pair KF at NEC=1/8 and was about as erratic. The average RMS error of the KF estimates were nearly the same for the 1-pair and 2-pair estimates at the same NEC, whereas the 1-pair IEKF had significantly higher RMS error than the corresponding 2-pair trials. We conclude that correction is faster and more accurate when using two probe pairs instead of one in a given amount of exposure time per correction iteration. This disproves our notion from previous chapters that one probe pair was better; we had not used the shortest possible exposure times in those cases.

Our aims with Kalman filtering are to make FPWC faster and to use correction images for recursive coherence differential imaging (RCDI). Our results in this section indicate that the extremely low SNR estimation is fastest, most stable, and most accurate with the KF. Since the IEKF is necessary to estimate the incoherent light recursively, this presents a tradeoff. If necessary, the KF could be used in FPWC and the IEKF could be run in parallel to perform RCDI. This tradeoff may become a nonissue if the higher model uncertainty and nonzero detector noise in a real instrument make the extremely low signal cases of NEC≤ 1/4 infeasible in practice. Addition of LOWFSC or ExAO correction residuals might also favor higher-signal exposures, in which the KF and IEKF provide equal performance. We will need to explore these possibilities in simulations and testbed experiments to better predict the expected capabilities of the WFIRST CGI.
5.4 Experimental Results from JPL’s HCIT

The following results from JPL’s HCIT were obtained before the simulations of §5.3.5 were performed. Because of that timing and other constraints, there were some key differences between our simulations and testbed experiments. We did not know to attempt any of the extremely low signal trials with the KF and IEKF, so we used the same exposure times for all the estimators when making comparisons in the testbed. Ideally the exposure times would have increased with each correction step to keep constant the NEC divided by the previous NI. Unfortunately, we could not use a rolling exposure time at the HCIT because of complicated camera behavior and limited time for calibration. The dark current signal versus time for the sCMOS detector was extremely nonlinear, so we could not properly subtract the dark signal via interpolation of master dark frames. We instead created master dark frames for a few exposure times and used a single exposure time during each correction run. To help mitigate the change in performance from not increasing exposure times, we limited the contrast correction range to a factor of about five. Each run started from the same DM setting providing $\approx 3 \times 10^{-8}$ measured NI, and the achievable measured NI floor was $\approx 6 \times 10^{-9}$.

5.4.1 Correction Speed Comparison

Here we compare the BPE, KF, and IEKF for a short 0.5s exposure time, giving a NEC of $9.3 \times 10^{-9}$ NI for the BPE. (The probes were at approximately $1 \times 10^{-7}$ NI.) For the KF and IEKF, we performed separate cases with one or two probed imaged pairs. Because the SNR was so low in the short exposures, we could not use the 0.5s unprobed images to measure the current contrast level. We instead took an extra 10s unprobed image at each correction iteration solely to measure the average dark hole NI; these long images were not used by the estimators. In Fig. 5.9, we plot the measured NI values from the 10s-long images during correction with each estimator. In terms of correction iterations, the recursive estimators
Figure 5.9: HCIT correction curves for 0.5s exposures. Separate cases with one or two probed image pairs were performed for the KF and IEKF. (a) NI vs correction iteration. (b) NI versus total integration time. The recursive 1-pair estimators reach the best NIs the fastest, and the recursive 2-pair estimators achieve slightly worse NIs. In general during the correction, the BPE achieved the worst NI.

performed similarly fast but with the IEKF requiring the first twenty iterations to catch up. When compared in terms of total exposure time, the 1-pair recursive estimators performed the best, reaching measured NIs at or better than $6 \times 10^{-9}$ at the end of the allotted time. The 1-pair KF was consistently faster than the other estimators until the 1-pair IEKF caught up in the middle of correction. The 1-pair IEKF might have required time to catch up with the 1-pair KF because of worse tuning or because it needed longer to build up a good recursive estimate. The 2-pair IEKF did not exhibit this slower initial correction convergence. The BPE gave the slowest correction overall and achieved the worst NI of approximately $7 \times 10^{-9}$.

In summary, the 2-pair recursive estimators provided faster correction than the BPE, and the 1-pair recursive estimators provided even faster correction and better achievable contrast.

5.4.2 Estimate Quality Comparison

Here we analyze the quality of the final estimates. Faster correction speed is important, but the other main reason to use the recursive estimators is to provide more accurate estimates.
We first describe an additional method of estimating the starlight intensity from probed measurements.

**Model-Free Estimation**

In an unpublished memo from 2015, Dr. Eric Cady derived the following method of finding the starlight contrast without using the optical model in the calculation. We first re-write Eq. 2.3.2 to be in terms of the cosine of the difference between the probe phase and electric field phase,

\[
\Delta I_{k,j} = I_{k,j+} - I_{k,j-} = 4\mathcal{R}\{E_k^* p_{k,j}\} + n_{k,j}
\]

\[
= I_{k,j+} - I_{k,j-} = 4|E_k||p_{k,j}| \cos(\theta_{p_{k,j}} - \theta_{E_k}) + n_{k,j},
\]

(5.4.1)

where the phase angles are defined as \(\theta_{E_k} = \angle E_k\) and \(\theta_{p_{k,j}} = \angle p_{k,j}\). Utilizing Eq. 2.3.12 to compute the amplitude of the probe, we calculate the electric field amplitude times the cosine terms solely from the measured images (and ignoring the noise terms) as

\[
|E_k| \cos(\theta_{p_{k,j}} - \theta_{E_k}) = \frac{I_{k,j+} - I_{k,j-}}{4|p_{k,j}|}
\]

\[
= \frac{I_{k,j+} - I_{k,j-}}{4\sqrt{\frac{I_{k,j+} + I_{k,j-}}{2} - I_k}}.
\]

(5.4.2)

The aim is to determine \(|E_k|\) from the \(N_{pp}\) measurements of \(|E_k| \cos(\theta_{p_{k,j}} - \theta_{E_k})\) with evenly spaced phase angles in the range \(\theta_{p_{k,j}} \in [0, 2\pi)\). (As explained in §2.4.1, evenly spacing the phase of the probe’s electric field in the image provides the lowest variance pair-wise estimate.) With knowledge of the relative phase of each probe from the commanded DM shape for each, we re-order the probes in order of increasing phase. We then perform a least-squares fit of both the amplitude and phase of a single-period sinusoid to the data.

The model-free starlight estimate provides an alternative method of calculating the starlight amplitude at each pixel without requiring knowledge of the optical propagation
in the model. It is not practical for replacing the standard pair-wise estimate because it does not provide the phase. We therefore implemented the model-free starlight estimate just once after each correction run to obtain a comparison estimate for the pair-wise estimators. The model-free estimate is sensitive to errors in $|p_{k,j}|$ from low SNR images, so it is more accurate for longer exposures. The model-free estimation can provide the incoherent estimate only through differencing, the same as in Eq. 2.3.13, and is thus not calculated for comparison here.

**Estimate and Image Comparisons**

Figure 5.10 shows the measured and estimated images for each of the five correction runs with 0.5s/exposure. The left column shows the measured images after correction with a long, 10s exposure. These long exposures were used to calculate the NI values shown in Fig. 5.9. The 10s exposures were still quite noisy, so in the 2nd column we include the model-free estimate, which used one unprobed image and 20 probed images of 2.0s/exposure. The model-free starlight estimates have NIs substantially lower than in the measured images, which means that there was a relatively large incoherent signal of a few times $10^{-9}$. Other tests by the HCIT team indicated that the actual incoherent light level in the dark holes was less than $\approx 3 \times 10^{-9}$ NI, so some of the missing intensity might still be noise in the 10s exposures. The third column shows the final starlight intensity estimates from the real-time estimators. Because they used many fewer, shorter images than the model-free estimates, they are substantially noisier. The BPE produced the worst starlight estimate with numerous incorrectly bright pixels from the high noise. The recursive estimators all produced less noisy estimates with many fewer erroneously bright pixels. Compared with the final 0.5s images in the fifth column, the incoherent estimates in the fourth column were essentially the same and composed primarily of noise.

In future tests, the BPE calculation should be modified for low-flux experiments. The probe amplitude was still calculated using Eq. 2.3.12, which is extremely noisy and inaccurate
Figure 5.10: Measured PSFs and intensity estimates at the end of each 0.5s/exposure correction run from Fig. 5.9. The 2-probe-pair estimators performed 30 iterations and the 1-probe-pair estimators performed 50 iterations to use the same total exposure time. All images are shown for the region $\xi \in [0, 10] \lambda/D$, $\eta \in [-6, 6] \lambda/D$. Average NI values in the correction region are included for the first three columns but not for the last two, in which the high noise level made the values inconsistent from iteration to iteration.

at low SNR. Because the two probe pairs do not change and the probe brightness can be fixed when correction steps are small, the measured probe amplitudes can be averaged over
several iterations to reduce noise. We could alternatively revert to using the model-based probe amplitudes since the HCIT model error is low. We expect that the averaging method, and possibly the model-based method, for obtaining the probe amplitudes would improve the BPE. Nonetheless, we still expect the KF and IEKF to perform better because they utilize the entire system model and previous data to build their estimates.

We expected the IEKF to produce a less noisy, more accurate incoherent estimate, but Fig. 5.10 shows that the IEKF provided no benefit in this experiment. We attribute this mainly to the filter tuning. We did not re-tune the IEKF (or KF) from prior values tuned in simulation only. The previous tuning values provided fast correction and we did not have time to test different filter settings, so we did not discover this issue until weeks after the experiment was complete and the testbed was decommissioned. In retrospect, we should have adjusted the process noise covariance matrix $Q_k$ to make the estimates average out noise more.

The readout noise from the sCMOS camera was much higher than would be expected during an actual mission, so the IEKF would still have had difficulty reducing the noise substantially if it had been working optimally. The RMS noise at each pixel was 5.2 ADU ($2.6 \times 10^{-8}$), equivalent to $4.5 \times 10^{-8}$ NI. The Princeton HCIL’s average RMS readout noise of 4.9 ADU ($\approx 7.5 e^{-}$, or $8 \times 10^{-8}$ contrast) did not cause problems in §4.3 because we were estimating incoherent signals at or above that contrast, not below it as we did here. In our JPL HCIT tests we attempted to estimate the incoherent intensities at accuracies about two orders of magnitude below the RMS readout noise, which was too difficult for the estimators in this case. Unlike the incoherent estimate, the starlight estimate was much less sensitive to the high readout noise because of the bright probes, as explained in §2.4.3.

As shown in Fig. 5.11, all the average incoherent intensity estimates were erratic. The fluctuations can be explained by the high readout noise. Many were also negative-valued, which is not physically possible. The assumption of the photon shot noise distribution being Gaussian only applies at high flux, but at low flux the true Poisson distribution is skewed and
Figure 5.11: Average incoherent intensity estimates for the 0.5s/exposure correction runs. All the estimates fluctuate greatly between iterations because of high noise in each image. Most of the estimates are negative because the starlight intensity estimates were too large.

distinctly non-Gaussian. We therefore suspect that the Poisson noise distribution skewed the estimates of the starlight to be too high, thus causing the incoherent estimates to be too low (and even negative). This potential issue merits further investigation because it may require different derivations of the KF and IEKF for Poisson noise sources.

5.4.3 Correction Runs with a Shorter Exposure Time

We had enough time in the HCIT to perform one more round of tests, so we performed another estimator comparison with shorter, 0.3s/exposure images. This set the NEC at approximately $2.1 \times 10^{-8} \text{NI}$, twice the previously achieved NI of $6 \times 10^{-9}$. The intensity estimates and measurements were qualitatively similar to those from the 0.5s/exposure trials, so we do not include another version of Fig. 5.10 for this case. The NI correction curves, shown in Fig. 5.12, were qualitatively similar to those for the 0.5s/exposure case in Fig. 5.9 except for a one minor difference. The 1-pair IEKF was no slower than the 1-pair KF this time, so the previously observed difference was possibly just a random fluctuation. When compared in total exposure time, The 1-pair recursive filters were consistently faster than the 2-pair estimators and reached better NI values. The BPE once again had a worse NI floor of about $7 \times 10^{-9}$, versus about $6.5 \times 10^{-9}$ for the 2-pair KF and IEKF and about $6 \times 10^{-9}$ for the 1-pair KF and IEKF. All the estimators achieved NI levels better than the NEC. We
Figure 5.12: HCIT correction curves for 0.3s exposures. Separate cases with one or two probed image pairs were performed for the KF and IEKF. (a) NI vs correction iteration. (b) NI versus total integration time. The recursive 1-pair estimators reach the best NIs the fastest, but the recursive 2-pair estimators almost caught up by the end of correction. The BPE did not reach NI values better than $7 \times 10^{-9}$, whereas the recursive estimators achieved values closer to $6 \times 10^{-9}$.

expected the KF and IEKF to do so because of their recursive nature, but it was a surprise that the BPE could as well.

### 5.5 Current Limitations of JPL’s HCIT

Here we discuss the limitations of JPL’s HCIT in more detail. Most of the issues were related to the HCIT-specific hardware, in particular the sCMOS camera and Xinetics DMs. There were additional calibration issues specific to the SPLC since this was the first one ever tested in the lab.

As mentioned in §5.4.2, the read noise of the Andor Neo sCMOS camera was much higher than would be expected for a space mission. Figure 5.13 shows a histogram of the measured read noise’s standard deviation at each of the $2 \times 2$-binned pixels. The mean standard deviation was $2.58 \text{ e}^-$ and the median $2.46 \text{ e}^-$ for a zero-second exposure. The unbinned pixels had a readout noise of half that but gave too fine a sampling of 11.9 pixels
Figure 5.13: Histogram of the RMS read noise per 2×2-binned pixel on the HCIT’s Andor Neo sCMOS camera. The mean standard deviation was $2.58 \, e^{-}$ and the median was $2.46 \, e^{-}$ for a zero-second exposure. The camera gain was $0.5e^-/ADU$.

per $\lambda_0/D$ in the image plane. Unlike CCDs, sCMOS detectors cannot bin pixels noise-free on chip, so binning in software increases the RMS noise as the square root of the number of binned pixels. In the future, the HCIT should design the image plane plate scale to match the detector pixel size and/or use a low noise EMCCD detector instead of an sCMOS. The WFIRST CGI is planned to have an EMCCD with a much lower read noise of $0.2 \, e^-$,[148] so implementing one in the HCIT would provide relevant experience for the mission in the ongoing FPWC tests.

The high read noise degrades the incoherent estimates much more than the starlight estimates. As derived in §2.4.3, bright probes in pair-wise estimation nearly eliminate the effect of readout noise on the starlight estimate. The incoherent estimate does not see this benefit and can improve at best only as fast as averaging images. Even if the measurements were only of the incoherent source, the noise in the incoherent estimate could only be reduced by the square root the number of images. By the end of our correction runs, that factor was $\sqrt{150} \approx 12.2$. With other noise sources and model error, the effective reduction of noise was actually much less than that theoretical value for the recursive incoherent estimate.

We concluded that the exposure time should be rolling during correction to save time since worse-contrast images can utilize shorter exposures. This strategy would require very careful calibration of the dark current’s dependence on time for the detector. We would also
need to determine a function for the exposure time based on the contrast. A logical starting point is our strategy from the §5.3 simulations of increasing the exposure time linearly with the NI improvement factor. That keeps the dark hole’s SNR constant if the only noise source is photon noise. In the presence of readout and dark current noise as well, it should make more sense to hold constant the NEC divided by the previous contrast since bright probing can mostly circumvent those other noise sources.

The Xinetics DMs had different limitations compared to the BMC MEMS DMs used at Princeton. The main issue with the Xinetics DMs was the hysteresis of the PMN actuators. Near room temperature, the hysteresis is minimized at less than 1%. During a correction run, 1% hysteresis was not a major concern. The largest actuation was in the probe shapes, which already cannot be too large because of the ignored nonlinearity of the DM phase Taylor expansion. The hysteresis became a problem when power cycling the DMs. Prior to FPWC, phase retrieval was performed and the DMs were used to null the large low-order aberrations of the pupil phase. When the DMs were power cycled, these large actuations were not reproducible and required phase retrieval to be performed again. The HCIT team at JPL found that pre-compensation of the low-order modes to high accuracy was time-consuming but crucial for optimal performance of the SPLC. The new Xinetics DMs used on the neighboring HLC optical bench exhibited much less hysteresis after power cycling, so using new DMs may mitigate the hysteresis problems sufficiently for the WFIRST CGI mission.

As just mentioned, the low-order Zernike modes needed to be pre-compensated for the SPLC to achieve the best contrast. The SPLC in these lab experiments presented two major hurdles. First, the SP apodizer blocked much of the aperture and left only small openings in some regions. This made phase retrieval very noisy and inaccurate in much of the pupil, so a flat mirror was translated into the same plane as the SP to perform phase retrieval of the whole pupil. The low-order aberrations everywhere else in the system were then sensed and pre-nulled. Since the SP was on a substrate that was difficult to polish better than \(\approx 60\text{nm}\)
peak-to-valley, the low-order modes (mostly astigmatism) on the SP were not sensed with phase retrieval. The unknown, low-order modes on the SP made initial correction runs much slower and reach a worse NI. The HCIT team identified some of the SP’s aberrations from comparing simulations of the coronagraphic speckle pattern for each Zernike mode with the observed speckles, but this approach was not optimized and likely missed some aberrations. In the future, we will need a more automated and precise method for identifying the SP’s low-order surface aberrations.

5.6 Comparison of Simulation and Lab Results

Here we compare the results from the WFIRST CGI simulation, the JPL HCIT lab tests, and the Princeton HCIL lab tests. We primarily compare the correction speed because the settings for incoherent light estimation were quite different for each case.

In the near-perfect knowledge case of our WFIRST CGI simulation, the most notable difference was between the allowable exposure times for the BPE and the filters. The KF and IEKF could utilize images several times shorter and thereby reach given NI values several times faster. With extremely low signal estimation images, the KF provided slightly faster and more accurate correction than the IEKF. When high-signal estimation images were used, the 2-pair BPE, KF, and IEKF all provided the same correction speeds. When the total exposure time per correction iteration was as low as possible, we found that correction was faster and more stable when using two probe pairs instead of one.

The JPL HCIT lab tests verified some of the cases from the simulations. At NECs at or slightly above the achievable NI, the BPE could not average out noise as well and had a worse achievable NI than the recursive filters. Before the BPE reached its achievable NI floor, the 2-pair estimators all had similar correction speeds. The 1-pair KF and IEKF still provided faster correction at the higher NI levels but not quite the full $5/3$ times faster from using $3/5$ the number of images as the 2-pair estimators. Having performed our testbed
experiments before the simulations, we did not perform lab trials to determine the shortest possible exposure times that the KF and IEKF could use and still provide stable correction.

In the Princeton HCIL lab tests, there was much more model uncertainty. In that case, the BPE was both slower and had a worse achievable contrast. The other main difference was that the 1-pair KF needed more time than all the other recursive estimators to reach the best contrast level. Only the IEKF (and EKF) gave the fastest correction in total exposure time. The experiment with the star-planet simulator showed that the IEKF provided a much better incoherent estimate containing the planet than the KF’s batch process incoherent estimate.

Distilling all these results, we make several recommendations of the estimator choice for well-calibrated coronagraph instruments such as the WFIRST CGI. First, a recursive estimator should be used instead of the BPE. The KF and IEKF provide better achievable contrast and more robustness to measurement noise (i.e., smoother correction curves) at low flux levels. Second, the 2-pair KF is the best choice if only faster correction is desired. The KF is easier to tune and computationally less expensive than the IEKF. For an equal amount of exposure time per correction iteration at the lowest light levels, the 2-pair KF is faster and more robust than the 1-pair version. Finally, the 2-pair IEKF should be used to obtain both fast correction and the best real-time estimate of the incoherent intensity. The IEKF provides equally fast correction as the KF except at the very shortest possible exposure level, but that case has not yet been verified in the lab. We confirmed that the IEKF produces a good recursive incoherent estimate in the HCIL, but not in the HCIT with higher measurement noise and no planet signal to extract. In the next chapter, we explore the planet detection capabilities of the IEKF’s recursive bias estimate in more detail.
Chapter 6

Signal Detection Using the Bias Estimate

When an exoplanet is fainter than the residual stellar speckles in the dark hole, it is impossible to detect the planet from a single image. Some form of diversity is needed in the image to disambiguate the planet from the speckles. As discussed in §1.7, the state-of-the-art image post-processing methods for exoplanet detection utilize the movement of stellar speckles relative to the planet with wavelength (spectral differential imaging, or SDI), the movement of the planet relative to the speckles during a telescope roll (angular differential imaging, or ADI), or the subtraction of PSFs from different stars (reference differential imaging, or RDI). These post-processing methods are considered an integral part of the technology development for proposed exoplanet imaging missions.[128] For instance, the expected science yield of the Wide-Field Infrared Space Telescope’s Coronagraph Instrument (WFIRST CGI) assumes that ADI and RDI will enable detection of planets 10-30 times fainter than the dark hole contrast.[78]

A suggested but previously undeveloped detection method, coherence differential imaging (CDI), utilizes knowledge of the stellar electric field to identify incoherent sources such as exoplanets. The modulation and estimation of the starlight is already required during focal
plane wavefront correction (FPWC), so the most efficient use of telescope time would be to perform CDI with wavefront estimation images. Our formulation of the iterated extended Kalman filter (IEKF) in Chapter 3 creates a recursive estimate of the incoherent light, which we can then analyze for a planet signal. We demonstrated in §4.3.3 that this recursive coherence differential imaging (RCDI) technique worked in four laboratory trials with an injected planet. To quantify rigorously the IEKF’s ability to detect exoplanets via RCDI, many trials should be performed with and without a planet present. In this chapter, we conducted Monte Carlo simulations of FPWC with a simple model of the WFIRST CGI to prove the principle of RCDI. We begin by discussing three types of observing scenarios and the utility of RCDI in each case. We then describe the setup and goals of these first RCDI experiments in simulation. After reporting the detection results for planets at different contrast levels, we describe our plans for improving the simulation model.

6.1 Observing Scenarios

One proposed observing scheme for the WFIRST CGI would dig the dark hole on a bright star, slew to the (fainter) star of interest, and integrate until the planet is detected and/or characterized. Based on simulations of the LOWFSC and structural-thermal-optical-performance (STOP) models of the WFIRST CGI performed at JPL and Goddard Space Flight Center, the expectation is that the stellar wavefront can be held constant near the $10^{-10}$ contrast level for slews up to a few degrees. Correcting the wavefront on a “reference” star several times brighter than the planet-hosting “science” star is intended to reduce the overhead time for wavefront correction and provide a reference PSF for RDI. The exposure time on the science star would then be just long enough to confirm or discount a planet detection at a desired contrast floor or to obtain a planetary spectrum at the desired SNR. To extract the planet signal from the dark hole speckles, the current plan for the WFIRST CGI is either to use RDI or ADI.
The utility of RCDI during a space mission depends largely upon the stability of the observatory. The more stable the stellar wavefront is, the more accurate the conventional image post-processing techniques are and the less need there is for RCDI. There are three general scenarios that we might expect for a coronagraphic mission:

Scenario 1: The observatory and instrument are ultra-stable. The dark hole can be maintained to about the $10^{-10}$ contrast level or better. Thermally-induced mechanical deformations of the observatory are well counteracted by heaters and LOWFSC during slews and rolls. DM hysteresis is negligible. Coronagraphic masks are swapped in and out with repeatable positioning.

Scenario 2: The observatory is sufficiently stable during moves, but the coronagraph instrument’s actuations are not repeatable and cause significant contrast degradation. The wavefront entering the coronagraph is held constant by the thermal controls and the LOWFSC. Hysteretic DM actuation and/or mask swapping are not repeatable enough for different bandpasses, so FPWC is required after each change to a new broadband (e.g., 10-18%) bandpass.

Scenario 3: The observatory is not sufficiently stable during maneuvers, even with active heating controls and LOWFSC. The dark hole contrast becomes several times worse with each slew or roll.

In Scenario 1 with a perfectly stable dark hole, photon shot noise sets the detection limit. As shown in Eq. 2.4.20, using pair-wise probed images to estimate the stellar intensity makes the stellar shot noise’s standard deviation (at least) $\sqrt{2}$ times larger than just imaging the unprobed starlight in the same amount of time. Since this scenario is shot noise limited and RCDI introduces more shot noise from probing, the best method for planet detection is PSF subtraction or RDI. High-SNR template PSFs can be obtained on much brighter reference stars with little penalty in time.
The comparison of detection methods is more complicated in Scenario 2. The efficiency of each observing and detection scheme depends on the amount of contrast degradation, the correction speed, the telescope slew and settling times, and the degree of hysteresis in the CGI actuation. If the contrast degradation is small enough, the dark holes at each bandpass can be dug on the bright reference star sequentially and only one slew is necessary to the science star. However, a different strategy is needed if the contrast degrades substantially when switching the coronagraph between different pre-corrected bandpasses. If the combined slew and correction time is short compared to the exposure time on the planet, the telescope should be slewed back to the bright reference star for each new bandpass. Otherwise, correction should probably be performed on the science star itself, in which case RCDI might be the most time-efficient detection method.

In Scenario 3, RCDI might be the only feasible option. Without the ability to slew or roll the telescope while maintaining a stable image, ADI and RDI will have limited utility or not work at all. We assume that if the telescope stays pointed at the same target and has a constant thermal load (such as at the Earth-Sun L2 Lagrange point), the stellar wavefront entering the coronagraph would be nearly static. RCDI would then be able to dig back the dark hole and provide an estimate of the incoherent light sources.

Current models for the WFIRST observatory and CGI predict performance most closely matching Scenario 1. The repeatability of the CGI actuations when switching among bandpasses has not been investigated, however, so Scenario 2 is still possible. The LOWFSC system is critical to making Scenarios 1 and 2 possible, but LOWFSC has thus far been demonstrated in simulation only. Until there is a lab demonstration showing that LOWFSC will keep the stellar wavefront sufficiently stable for FPWC, Scenario 3 cannot be confidently ruled out. A testbed is currently being constructed in JPL’s HCIT to perform the first lab experiment that integrates LOWFSC with FPWC.

Although the most compelling cases for RCDI are Scenarios 2 and 3, we simulated Scenario 1 in this chapter. The static stellar wavefront in Scenario 1 allows us to prove the
principle of RCDI in a simple case. Future simulations of Scenarios 2 and 3 will require more sophisticated, dynamic optical models as well as advanced treatments of RDI (e.g., KLIP) beyond simple PSF subtraction.

6.2 Simulation Setup

In these simulations, we used the same shaped pupil Lyot coronagraph (SPLC) design and optical configuration as in §5.3. The central wavelength was 770 nm to match one of the bandpasses planned for this characterization SPLC in the WFIRST CGI.

For the Monte Carlo runs, we introduced several sources of variation among the trials. We included photon shot noise on the star and planet. Dark current noise and readout noise require assumptions about the detector and exposure times, so they were omitted in these initial tests. We included 10% RMS errors in the DM actuator gains. For all of the optics, we generated new, random phase aberration maps with 5nm RMS and a realistic -3.5 power law roll-off in frequency. The low-frequency components of those aberrations heavily degrade the performance of FPWC, so we mostly flattened the phase at the exit pupil with the DM to emulate pre-FPWC phase retrieval and flattening. Finally, we included 0.5% RMS amplitude errors on the shaped pupil mask to model the non-uniformity of its aluminum coating.

As discussed in §5.3.4, these aberrations alone are not enough to slow FPWC down to the levels observed in the laboratory. In §5.3 we slowed correction by displacing the FPM a fixed amount for all runs. Here we chose not to introduce any fixed model errors or aberrations because doing so would, at least partially, defeat the purpose of performing Monte Carlo trials. In future simulations, we plan to include several other sources of model error (e.g., DM registration or mask alignment errors) with values chosen at random within the expected tolerances.

To calculate the probability of detection, both the rates of true positive detections (finding a real planet) and false positive detections (finding a nonexistent planet) must be known.
Therefore, in half of our trials there was no planet, and in the other half a planet was included at one location in the image plane. This simplification minimized the number of trials to perform. As shown in Fig. 6.1(a), we placed the center of the planet at $(\xi, \eta) = (3.5, 0)\lambda_0/D$. That location is as close to the inner working angle (IWA) as possible without clipping.

![Normalized Exoplanet PSF](a) Exoplanet PSF FWHM (b)

Figure 6.1: Normalized, off-axis PSF of the exoplanet used in these simulations. (a) Log scale. (b) Linear scale showing the FWHM.

The FWHM of the planet’s PSF. The detectability closest to the star is of particular interest because the WFIRST CGI is expected to observe most exoplanets near the IWA. For planets at or below the dark hole contrast, only the peak of the PSF will be detectable. We therefore use the FWHM of the exoplanet’s PSF, shown in Fig. 6.1(b), to calculate the detection probability. The rest of the planet light is obscured by the measurement noise. We chose a reasonable plate scale sampling of 3 pixels per $\lambda_0/D$, which yielded only 11 pixels under the planet’s FWHM. Rather than injecting a static planet PSF into the final starlight images, we propagated the planet’s electric field through the entire optical model to generate a more accurate planet PSF.

The correction curves from these simulations, shown in Fig. 6.2, were unrealistically fast, but we allowed them for these initial, proof-of-principle tests. To prevent FPWC from reaching extremely high contrast in the dark hole in monochromatic light, we performed estimation and control over a 6% bandpass (three neighboring 2% sub-bandpasses). We did not perform correction over the full 18% spectral bandwidth of the coronagraph to reduce
Figure 6.2: Correction speed curves for (a) the full dark hole and (b) the planet’s FWHM region for all the trials in one experiment. There is a factor of 3 variability in achievable average normalized intensity (NI) for the full dark hole but a factor of approximately 10 for the planet region.

computational time and because we calculate the planet estimate in each 2% sub-bandpass separately anyway. Correcting in broadband sets the achievable normalized intensity (NI) near the design limit of the coronagraph. As shown in Fig. 6.2(a), the range of final average NI values in the full dark hole was within a factor of three. For the final average NI of the 11 pixels in the planet region, however, Fig. 6.2(b) shows there was much more variability among trials (about a factor of ten). It is uncertain if this large range is realistic, and we will pursue this question further after making the aforementioned improvements to our simulation model. The main difficulty with so much variation in the dark hole contrast is that the stellar shot noise varies significantly among trials. The signal-to-noise ratio (SNR) then cannot be applied to the trials as a whole, so it is difficult to determine the efficiency of RCDI.
6.3 Design of Experiments

The main purpose of these experiments was to demonstrate the ability of RCDI to detect exoplanets. Our goals were to:

1) Prove the viability of RCDI.

2) Quantify the probability of detection.

3) Compare the performance of RCDI to that of the fundamental shot noise limit.

We performed several sets of trials at different planet contrast levels. The most compelling cases are for planets fainter than the dark hole, so we tested two of those cases and one more with the planet at the dark hole contrast. In each experiment, we chose the exposure times and number of correction iterations such that tens or hundreds of photons in total should have arrived from the planet by the end of the run.

As in §4.3.3, we computed the 2D planet PSF correlation (C), given by Eq. 4.3.1, and the best-fit planet contrast from the incoherent estimate at the central wavelength. We calculated these values for each detection method at each correction iteration in each trial, even when the planet was not present. The template PSF for the planet was the ideal, model-based one shown in Fig. 6.1(b). Using the PSF correlation values, we constructed the receiver operating characteristic (ROC) curves. These curves show the tradeoff between higher probability of detection (the fraction of true planets counted as detections) and a lower false alarm rate (the fraction of spurious signals counted as detections). We describe ROC curves in more detail in the next section.

In this observing scenario, we split the correction run into two stages for the reference and science stars. The dark hole was almost entirely dug on the reference star for the first seven iterations. Since the contrast improved extremely rapidly in these simulations, the KF and IEKF were not started until the fourth iteration, when the dark hole had already reached about $10^{-8}$ average NI. This allowed the filters to avoid the larger model
nonlinearities at worse contrast levels. The IEKF was used for correction, and the KF operated in parallel on the same images without influencing the controller. At the seventh iteration, the correction run switched to the science star and the exposure time was fixed. For these initial simulations, we did not include any differences (e.g., in spectra or stellar angular size) between the reference and science star. Because our IEKF formulation builds a recursive estimate of the incoherent light, we zeroed out the previous estimate and increased the incoherent state’s covariance by 30 times when switching to the science star.

We compared the quality of the IEKF’s incoherent estimates to the fundamental shot noise limit. As a simpler alternative to the IEKF’s incoherent estimate, we also calculated the mean of all the KF’s prior incoherent estimates on the science star. (We did not do the same for the BPE because we have proven in previous chapters that the KF provides a better stellar estimate, and thus a better batch process incoherent estimate, than the BPE.) This serves as a baseline to determine if the IEKF’s extra complexity provides a better detection. We refer to both these methods as RCDI in this chapter. In these tests, the fundamental detection limit from shot noise was found using idealized PSF subtraction; the subtracted reference image was noiseless, and the only variation came from photon shot noise in the science image. It was therefore impossible for RCDI to outperform idealized PSF subtraction in this set of tests, and in fact the KF and IEKF must perform slightly worse than the fundamental limit because of their higher noise variance from the probed images. Images for idealized PSF subtraction were taken immediately after switching to the science star, and their exposure times equalled the total cumulative times used for the correction iterations that followed. That is, at each discrete time step set by the correction iterations, there was a point for each the KF average, the IEKF, and the fundamental shot noise limit.
6.4 Results from Simulation

Experiment 1: Exoplanet at $3 \times 10^{-10}$ Contrast

In this first case, we examine the detectability of an exoplanet at $3 \times 10^{-10}$ contrast. As indicated by Fig. 6.2(b), that makes the planet 1-10 times fainter than the starlight at the same location. The planet’s flux rate was set to 1 photon at the peak of the planet per image, and 5 equal-length images (4 probed and 1 unprobed) were taken per correction iteration.

We first examine the PSF correlation values for each method in Fig. 6.3. The plots include all the values from each of the 100 trials with or without the planet present, and the median values are plotted for reference as solid lines. The median is used instead of the mean because the distributions become more skewed as the average approaches the limits of -1 or 1. All three methods work well with the planet present and provide higher median PSF correlations with more exposure time. The median for the trials without a planet should be zero, which idealized PSF subtraction and the IEKF both provide. The KF averaging method instead has an unexplained negative bias of approximately 20%. The correlation distributions have more spread in the no-planet cases because the PSF is being fit to only

![Figure 6.3: Planet PSF correlation values for each method with a $3 \times 10^{-10}$ contrast exoplanet. Black points are values from the 100 trials with a planet and red points are from the 100 trials without a planet. The black points are artificially shifted right to distinguish the red and black points more easily. The solid lines are the median PSF correlation values over all trials.](image-url)
Here we explain how to quantify the probabilities of interest from the Monte Carlo trial data. A full treatment of statistical signal detection theory can be found in numerous texts (e.g., [151]). The four possible detection-choice outcomes after each observation are:

a) a true positive detection (a planet signal is correctly counted as a planet)

b) a missed detection (a planet signal is incorrectly discounted)

c) a false positive detection (a spurious signal is incorrectly counted as a planet)

d) a true negative detection (a spurious signal is correctly ignored).

The competing goals of any detection scheme are to maximize the fraction of true positive detections, also called the probability of detection ($P_D$) or completeness, while simultaneously minimizing the fraction of false positive detections, called the probability of false alarm ($P_{fa}$). The receiver operating characteristic (ROC) curve displays this tradeoff by plotting the $P_D$ against the $P_{fa}$. In Fig. 6.4, we demonstrate how to construct the ROC curve from the estimated PSF correlation values. One ROC curve represents the data from a single time step during FPWC. For easier visualization, we converted the points in the top plot of Fig. 6.4 into the histogram on the bottom left. The choice of the PSF correlation cutoff value $C_{th}$ is called the “threshold” and is parametrized in the ROC curve. The fraction of points above the threshold value from the no-planet case is the $P_{fa}$; similarly, the fraction of points above the threshold value from the planet case is the $P_D$. Several points on the ROC curve (A-E) are labeled with their corresponding threshold values in the histogram. The minimum threshold of $C_{th} = -1$ (Point A) always counts all planets and false alarms, whereas the maximum threshold of $C_{th} = 1$ (Point E) never permits any detections. In the case shown, the minimum threshold to obtain all detections ($C_{th} = -0.2$ at Point B) still allows too many false alarms (70%). The best threshold choice in this case is probably $C_{th} = 0.5$ (Point D) or slightly smaller. At the Point D threshold, the completeness is 65% with zero false alarms.
Figure 6.4: Example construction of an ROC curve from the IEKF PSF correlation values. (Top) PSF correlation values from a single time step are (bottom left) plotted in histogram form. (Bottom right) The ROC curve is constructed from the histogram. The fraction of no-planet-case PSF correlation values above a chosen threshold $C_{th}$ is the probability of false alarm ($P_{fa}$), and the fraction of planet-case PSF correlation values above the threshold is the probability of detection ($P_D$). Several points (A-E) on the histogram show the mapping of PSF correlation values along the ROC curve.

In general as the exposure time increases and more signal is collected from the planet, a desirable estimator would keep moving the ROC curves up and left toward the curve of perfect detection, going from $(P_D, P_{fa}) = (1,1)$ to $(0,1)$ to $(0,0)$ and known as the oracle.
To be useful, an estimator must provide an ROC curve above the $P_D = P_{fa}$ line, which is called the “coin-flip” line because it corresponds to unbiased, random guessing.

Figure 6.5(a) shows all the ROC curves for the IEKF. Lighter curves denote earlier correction iterations. As desired, the ROC curves generally move closer to the oracle after more correction iterations. The IEKF in this case provides a favorable, high completeness (above 80%) for low false alarm rates (below 10%) after the fifth correction iteration.

Because it can be difficult to see many ROC curves on the same plot, the area under the ROC curve (AUC) is another useful metric. As the name indicates, it gives the area under the ROC curve and more simply displays the improvement of the ROC curve with more exposure time. However, the AUC can sometimes be misleading because many different ROC curves can have the same area underneath them; a higher AUC might not actually provide a higher completeness at a given threshold. It is therefore important to observe both the ROC curves and AUC plot together for a complete analysis.

Figure 6.5(b) shows the AUC curves for each detection method. The IEKF provided performance close to the fundamental shot noise limit after five iterations. Surprisingly, the
KF average method provided essentially the same ROC curves (not shown) and AUC curve as the IEKF. We are uncertain if that means the IEKF is underperforming or that the KF averaging is inherently as good at detection as the IEKF.

The estimated planet contrast is another useful measure of detection accuracy since the PSF correlation performs a normalized fit that ignores the planet contrast. Figure 6.6 shows the planet contrast estimates for each method. The spreads in the IEKF and KF average estimates appear similar and have many more outliers than the shot noise limit. The cases with the planet should converge on the true planet contrast and the cases without the planet should converge to zero. The median contrast estimates for PSF subtraction matched the true values well. The median of the IEKF’s estimated contrast values were slightly below the true value, and the median values for the KF average were significantly below the true values. The negative bias of all the estimates for the KF average is from over-estimation of the starlight, and the negative bias in the IEKF planet estimates could be the result of sub-optimal filter tuning.

Figure 6.6: Planet contrast estimates for each method with a $3 \times 10^{-10}$ contrast exoplanet. Black points are values from the 100 trials with a planet and red points are from the 100 trials without a planet. The black points are artificially shifted right to distinguish the red and black points more easily. The solid lines are the median contrast estimates over all trials, and the dashed lines are the true planet contrast and zero.
Experiment 2: Exoplanet at $1 \times 10^{-9}$ Contrast

Here we review the results from another test with a brighter exoplanet at $1 \times 10^{-9}$ contrast. This planet should be easier to detect, so we reduced the exposure times to give an average of $1/3$ photon per image at the peak of the planet. As shown by the AUC curves in Fig. 6.7(b), RCDI provided nearly perfect detection of the planet by the seventh correction iteration after an average of 12 photons had been collected from the planet. The IEKF and KF average again provided closely matching ROC and AUC curves. The IEKF’s ROC curves in Fig. 6.7(a) show steady improvement until nearly perfect detection was achieved.

Figure 6.7: (a) Receiver operating characteristic (ROC) curves for each correction iteration with the IEKF in the $1 \times 10^{-9}$ contrast planet case. Darker curves correspond to later correction iterations. (b) Area under the ROC curves (AUCs) for all three detection methods.

In the PSF correlation plots of Fig. 6.8, the RCDI cases with a planet once again have lower median correlations than for the shot noise limit. The negative correlations in the no-planet RCDI cases are much more pronounced this time. In particular, the KF average’s no-planet correlations decrease even faster than the with-planet correlations increase. We see in Fig. 6.9 that the KF average’s negative PSF correlations correspond to a large, negative contrast estimate for a nonexistent planet. The KF average’s median planet estimate is also
Figure 6.8: Planet PSF correlations for each method with a $1 \times 10^{-9}$ contrast exoplanet. Black points are values from the 100 trials with a planet and red points are from the 100 trials without a planet. The solid lines are the median contrast estimates over all trials.

negatively biased by about $5 \times 10^{-10}$, half the brightness of the planet. The IEKF contrast estimates also exhibit negative biases, although they are less severe. The negatively biased contrast estimates might be attributable to the over-estimation of the starlight by the IEKF and KF from the shorter exposures. Further investigation is necessary to determine why the KF average has such a strong negative PSF correlation when no planet is present.

Figure 6.9: Planet contrast estimates for each method with a $1 \times 10^{-9}$ contrast exoplanet. Black points are values from the 100 trials with a planet and red points are from the 100 trials without a planet. The solid lines are the median contrast estimates over all trials, and the dashed lines are the true planet contrast and zero.
Experiment 3: Exoplanet at $1 \times 10^{-10}$ Contrast

For the last case, we tested detection with as faint a planet as possible. An attempt at $3 \times 10^{-11}$ contrast did not work, so we set the planet to $1 \times 10^{-10}$ contrast. The dark hole contrast was brighter than the planet in every trial this time. To mitigate the additional shot noise and reduce the number of iterations per trial, we chose a higher flux of 3 photons per image at the peak of the planet.

Since the dark hole in the planet region was 3-30 times brighter than the planet, the range of SNRs after collecting 80 planet photons was 1.6-4.5. It is therefore unsurprising that the RCDI did not catch up with the shot noise limit by the end of the allotted time. The IEKF’s AUC curves in Fig. 6.10(b) showed constant improvement with more exposure time and reached 0.9. The final ROC curves for the IEKF in Fig. 6.10(b) had a reasonable 70% completeness at a 10% false alarm rate. Once again the ROC and AUC curves for the KF average and IEKF were nearly identical.

Figure 6.10: (a) ROC curves for each correction iteration with the IEKF in the $1 \times 10^{-10}$ contrast planet case. Darker curves correspond to later correction iterations. (b) AUCs for all three detection methods.

Because of the higher shot noise, we observed larger spreads in the PSF correlations for
each method in Fig. 6.11. The other noticeable difference from the previous experiments is that the median PSF correlations for the IEKF and KF average had positive biases when no planet was present. The median contrast estimates without planets for the IEKF and KF average also had small positive biases in Fig. 6.12, although the median contrast estimates were unbiased when there was a planet.

Figure 6.11: Planet PSF correlations for each method with a $1 \times 10^{-10}$ contrast exoplanet. Black points are values from the 100 trials with a planet and red points are from the 100 trials without a planet. The solid lines are the median contrast estimates over all trials.

Figure 6.12: Planet contrast estimates for each method with a $1 \times 10^{-10}$ contrast exoplanet. Black points are values from the 100 trials with a planet and red points are from the 100 trials without a planet. The solid lines are the median contrast estimates over all trials, and the dashed lines are the true planet contrast and zero.


6.5 Analysis

We completed our objectives for these initial experiments and have identified several directions for future work in RCDI. Most importantly, we proved that RCDI is a viable method for the detection of exoplanets. We showed with Monte Carlo trials of different correction runs that the probability of detection with RCDI increased with more exposure time and was reasonably close to the fundamental limit set by photon shot noise. Our idealized treatment of PSF subtraction was unrealistic (being exactly at the shot noise limit), so our future work will also need more practical simulations of ADI and RDI for fairer comparisons with RCDI.

From our initial results, we have several unanswered questions. The largest mystery is whether the IEKF’s incoherent estimate is any better than the average of the KF’s batch process incoherent estimates. The accuracy of the photometry (e.g., the median estimated planet contrast) was usually more accurate with the IEKF than with the KF average, but the probability of detection based on the ROC and AUC curves was nearly identical for both in all cases tested. To improve the IEKF, we may need to implement a more sophisticated tuning law for the process noise and sensor noise covariance matrices, \( Q \) and \( R \). We currently choose the values for \( Q \) and \( R \) based on the average contrast values across the full dark hole, but we observed that the contrast values can be significantly different in the planet FWHM region. The IEKF’s estimates might therefore improve with different covariance for small, local regions within the dark hole.

Next, we need to identify the reasons the KF average’s (and to a lesser extent the IEKF’s) median contrast estimates and PSF correlations sometimes exhibited large biases. We observed that the biases worsened as the exposure time decreased (for the same stellar flux). The issue might be that the pair-wise estimation scheme and the filters assume a Gaussian noise distribution while the shot noise in low-signal images actually has a highly skewed Poisson distribution.

In our future work, we plan to simulate RCDI in more complicated and realistic scenarios where FPWC would be more useful or even required. The first priority is to slow FPWC
down to believable speeds by including more aberrations and errors in the optical model. The challenge is to allow all the parameters to change among Monte Carlo trials without having the correction speeds or achievable contrasts vary too dramatically (e.g., by factors of 10 or more). To simulate Scenarios 2 and 3, we need to generate realistic wavefront disturbances from telescope maneuvering, LOWFSC residuals, and/or hysteretic hardware motions. For all these future simulation design choices, we plan to work more closely with the WFIRST CGI optical modeling group at JPL to produce reasonable settings.
Chapter 7

Future Directions in High-Contrast Imaging

On the path to imaging earth-like exoplanets, there are still numerous technological hurdles to overcome. All parts of such an advanced observatory will need to work in tandem to achieve $10^{-10}$ contrast, so other researchers have focused on the telescope requirements (e.g., pointing stability, segmentation),\[152, 153, 154\] the capabilities of the deformable mirrors,\[155, 156\] and the coronagraph designs.\[157\] Our work has supported the development of focal plane wavefront correction (FPWC) algorithms, particularly for the planned Wide-Field Infrared Survey Telescope Coronagraph Instrument (WFIRST CGI) mission to image dozens of gas giant exoplanets. In just the past ten years, FPWC methods have progressed from model-free speckle nulling, which required thousands of correction iterations, to model-based correction that needs only a few tens of iterations and can concurrently detect science targets. Further FPWC developments during the preparation for and mission lifetime of WFIRST will prepare us for an exo-earth imaging mission in two to three decades.

By proving the viability of using a nonlinear, recursive estimator for focal plane wavefront sensing in this thesis, we have enabled several possible paths for improvement such as different probe choices and parameter estimation. We found that the iterated extended Kalman
filter (IEKF) eliminates most of the bias error of the regular extended Kalman filter. More advanced filters may provide further improvement in accuracy. Our current estimators and controllers are still limited by imperfect input parameters, so better calibration is necessary. Alternatively, an adaptive model could be used for improving itself in real-time from the measurements. To accommodate dynamic, residual aberrations from extreme adaptive optics (ExAO) or a low-order wavefront sensing and control (LOWFSC) system, our estimator formulation can be modified via sensor fusion into a multi-rate Kalman filter. In the following sections, we describe the future research directions in coronagraphy, wavefront correction, and exoplanet detection, both in the near-term for the WFIRST CGI and in the long-term for giant ground-based and space-based telescopes.

7.1 Coronagraph Design

We have focused on efficient wavefront correction, but the efficiency of the overall mission can also be increased with improved coronagraph designs. Of the numerous coronagraphs mentioned in §1.2.1, most of them are based on analytical solutions for open apertures. The complicated, obstructed apertures of WFIRST, all proposed extremely large ground telescopes, and nearly all proposed large space telescopes require numerical optimization. The numerical optimization of coronagraphs for arbitrary apertures has mostly occurred in the last decade, and the prospect of the WFIRST CGI sparked a new wave of innovation three years ago. The two numerical approaches currently validated in experiment are apodization of the pupil in a Lyot-type coronagraph[46, 50, 48] or use of a complex (phase and amplitude) focal plane phase mask.[43, 93, 51] Because the numerical optimizations are computationally expensive, have unique solutions for different apertures, and benefit heavily from prior insights, the parameter space of possible solutions has barely been explored.

Most numerically optimized designs for obstructed apertures unfortunately block much of the exoplanet light when trying to suppress the extra diffracted starlight. For the WFIRST
CGI, for instance, the SPLC and HLC designs have planet throughputs of only about 4%, and the transmission of the non-coronagraphic optics reduces the throughput by another factor of two. (We use Krist et al.’s [95] definition of throughput as the ratio of the planet light under the FWHM of the PSF over the light incident on the primary mirror of the telescope. Therefore, the throughput without the coronagraph in WFIRST is 34%, not 100%, because of the nominal diffraction pattern.) Coronagraph designs having higher planet throughput without sacrificing IWA, contrast, or spectral bandwidth could greatly improve the science yield of the WFIRST CGI and other future coronagraphic missions.

7.2 Integrated Control Loops

Our work addressed only the suppression of quasi-static speckles for achieving high contrast. An actual space-based coronagraph will need a LOWFSC system to keep the telescope stable, and a ground-based coronagraph will also require ExAO. Krist et al. [95] and Ygouf et al. [131] investigated the effects of LOWFSC residuals in the image after FPWC but not during it, and Frazin [158, 159] has begun working on a formalism to account for ExAO residuals in the image. If the LOWFSC and ExAO correction error residuals are sufficiently small then FPWC can be treated separately, but to achieve the highest detection sensitivity all the control loops should be integrated. The Kalman filter formalism is well suited for this problem because it can incorporate dynamics and model uncertainty. Since the three correction loops operate at different timescales, this technique is called a multi-rate Kalman filter. A testbed is being built in JPL’s HCIT to test FPWC and LOWFSC concurrently, and our group at Princeton plans to reconfigure the HCIL to include LOWFSC as well.

7.3 Calibration and Modeling

Better models of the coronagraphic optical system will enable faster FPWC and more accurate simulations. FPWC in the lab drastically slows down at contrast levels near $10^{-9}$
because our models do not match the testbed well at that level. Closed loop correction allows us to keep improving in contrast at the expense of much more time. Some of these model-testbed performance discrepancies have been attributed to mask manufacturing errors or un-sensed low-order modes in the pupil, but the problem still remains. As part of the preparation for the WFIRST CGI in the past two years, the high-contrast imaging group at JPL has greatly improved their simulation capabilities and testbed calibration procedures. They have been meticulously simulating and characterizing each expected source of model-testbed discrepancies, and for the testbed they have been inventing new metrics to measure the system behavior for comparison with the model. Space-based coronagraphs will not have all the metrology equipment of a ground-based instrument or laboratory, so additional work will be necessary to modify calibration techniques for the available hardware on orbit. More progress in modeling and calibration will enable faster FPWC, more realistic simulations of coronagraph missions, and higher science yields for future missions.

7.4 Estimation and Control Algorithms

Many research directions remain in the relatively young field of coronagraphic wavefront estimation and control. Our validation of the nonlinear IEKF formalism in this thesis enables the use of other nonlinear filters and other probing strategies. Nonlinear, adaptive filtering could enable a self-improving model for better estimation and control. Direct broadband estimation (instead of the currently used sequential estimation of narrow spectral bandpasses) could enable faster correction or, at the least, the need for fewer filters in an instrument. Finally, other control laws are worth investigating since electric field conjugation and stroke minimization have thus far been the only model-based controller to achieve high contrast in the lab.

Up to this point we have used the same set of one unprobed image and $N_{pp}$ probed image pairs per correction iteration because it provides sufficient diversity for wavefront correction.
The original purpose of using pairs of probed images was to yield a linear relationship between the electric field and the field change from the probes, but we showed that the nonlinear measurement in Eq. 3.1.4 along with the IEKF works at least as well if not better. We can therefore modify our measurement equation $z_k$ from Eq. 3.1.2 to a more general one comprised of any set of images, including unpaired probes. For example, the minimal set of images for an estimate update might consist of the unprobed image, a single probed image for the real part of the field, and a single probed image for the imaginary part of the field. We could also skip taking the unprobed image to save time if we are confident that the starlight estimate from only probed images is accurate. In either case of using unpaired probes or not taking the unprobed image, we would no longer be able to use Eq. 2.3.12 to measure the probe amplitudes. Our model would thus need to be even more accurate than it has been for previous lab experiments, for which the measured probe amplitude has allowed us to overcome much of the model uncertainty. We have begun to test an IEKF with the real and imaginary parts of the probe as extra states at each pixel. This strategy turns the IEKF into a parameter adaptive filter (PAF), which could help estimation by improving knowledge of the probe signal or worsen estimation by trying to estimate too many variables at once. We will discuss other possibilities for using a PAF later in this section.

In this thesis, we found the IEKF to be sufficiently accurate for FPWC. The main error in the IEKF’s estimate stemmed from the assumption of the linear state model, which has significant nonlinearity errors at poor contrast (worse than about $10^{-6}$). At high contrast, the IEKF had no issues with the quadratic measurement function since we observed no bias errors. We believe that more advanced nonlinear filters are unnecessary given the IEKF’s excellent performance at high contrast, but they should be investigated in case they can provide a better estimate or more robustness to model uncertainty, especially in the case of extremely low SNR per image. In particular, the unscented Kalman filter (UKF) is only slightly more complex than the IEKF and has the potential to propagate the state covariance better through a nonlinear system.
As we demonstrated by augmenting the state with the incoherent signal, the IEKF also allows the estimation of other system parameters by adding them to the state vector. This is known as a PAF, and it could improve the performance of both the estimator and the controller since both rely on the model. Choosing the best parameters to estimate in the PAF is difficult. In our case, we would like to estimate the control Jacobian directly to mitigate all sources of model uncertainty at once, but that would increase the number of states at each image plane pixel from three to several thousand. That formulation is intractable and unlikely to converge, so specific system parameters would have to be estimated instead. PAFs have even more settings to tune and are not guaranteed to converge, so only the most observable parameters in the image plane might be viable as states in the filter. PAFs would also require more computing power because they have more states and, in most cases, require re-computing the control Jacobian each correction iteration. A neural network or other machine learning algorithm might be better suited than an IEKF-based PAF for estimating many of the optical system’s parameters. Such a computationally expensive algorithm would only be feasible for ground-based instruments in the near future.

Some of the shortfalls in our calibration techniques and focal plane wavefront estimation methods might be solved by using the COFFEE algorithm.[121] We have not used COFFEE, however, because it is difficult to set up the nonlinear optimization, it is currently too slow for real-time correction, it currently uses very large DM stroke (too much for Xinetics DMs) for creating phase diversity, and it has not been used in the lab to create high-contrast dark holes. We propose using COFFEE in the near-term as part of the pre-FPWC system calibration. COFFEE can provide estimates of the pupil aberrations with the coronagraph in place and thus characterize the system as it is used, whereas the conventional phase retrieval method requires the temporary removal of the FPM and Lyot stop. We would need to use smaller and possibly different phase diversity probes than the ones proposed in the literature in order to use COFFEE with the current Xinetics DMs in the HCIT.

Our contribution to FPWC focused on monochromatic estimation since the current
scheme uses sequential narrowband imaging across the whole bandpass. We showed in this thesis that Kalman filtering could speed up correction by several times, and a broadband imaging scheme might further speed up correction by a factor of 3 to 10. An integral field spectrograph (IFS) generates spectra at each point in the image, and an IFS is part of the design for the WFIRST CGI. If the IFS images can be used for FPWC, then the narrowband images could be taken in parallel. As an example, the SPLC design for the WFIRST CGI’s IFS has an 18% spectral bandwidth. If 3% sub-bandpasses are used for estimation the time savings factor is 6, and if 2% sub-bandpasses are used the factor is 9. (These are only rough figures that ignore the throughput loss in the filters for sequential narrowband imaging or in the IFS’s prism for the parallel narrowband imaging.) It is currently unknown if the IFS will be suitable for FPWC because the raw images must first be processed into image data cubes, in which the third dimension is the wavelength. Processing the raw images is challenging because the camera pixels sample the spectra differently for each image location and because there is some cross-talk (i.e., overlap) of neighboring spectra at the detector. In the next year or two to prepare for the WFIRST CGI, it will be important to test the feasibility of broadband FPWC with an IFS in both simulation and on the IFS optical bench being assembled this year in JPL’s HCIT.

Besides parallel narrowband imaging, there are other possible approaches to broadband estimation. Sirbu et al.[160] have begun investigating chromatic probing to estimate the electric field at several wavelengths within a broadband image, but it is still unclear if this strategy is feasible or would provide faster estimation than sequential narrowband probing. Groff[106] attempted “windowed” estimation, in which he imaged for estimation only in a narrow, central sub-bandpass and tried to extrapolate the electric field across the rest of broad bandpass. Because back propagation was not possible within the coronagraphic optical model, windowed estimation could not adequately estimate the field at the wavelengths outside the imaged bandpass. For either the windowed or full broadband estimation to be feasible, we believe that an accurate model of the electric field’s chromaticity is necessary. As
Soummer et al.\cite{130} found while investigating post-processing on high-contrast images, there is not (yet, at least) a closed-form relationship between the wavelength and the focal plane electric field that does not require complete knowledge of the aberrations at every optic. Aberrations non-conjugate to the DMs are extremely chromatic at high contrast because they undergo phase-to-amplitude mixing before reaching the DM planes.

We focused on estimation in this thesis because we found the current controllers to be sufficient. Nevertheless, other control laws might address some of the shortfalls of the linear-model, quadratic cost controllers currently used. Instead of turning stroke minimization’s inequality constraint in Eq. 2.2.9 into an equality constraint, we could investigate the use of a convex solver on the actual inequality constraint. Unlike with the current controllers, that would enable contrast constraints at each pixel instead of on average, and it would not penalize any pixels for having a better contrast than the targeted values.

We know from simulation and the testbed that residual low-order aberrations (missed during conventional phase retrieval) limit the correction speed and achievable contrast of a coronagraph. EFC and stroke minimization as currently formulated only correct the mid-spatial-frequency aberrations that are directly observable in the dark hole. The LOWFSC systems being designed for the WFIRST CGI and other space-based coronagraphs perform only differential low-order correction about a set point, so they also cannot correct absolute low-order aberrations. If we want to completely suppress the low-order modes and achieve better contrast without having to perform extensive calibration before FPWC, we could modify the controller to actuate low-order Zernike modes as well. This might require using light from outside the dark hole to obtain more data about the low-order aberrations. With the SPLC, for instance, low-order modes cause bright speckles to appear just outside the dark hole in the area blocked by the FPM.
7.5 Detection Methods

Image post-processing is being considered an integral part of future exoplanet imaging missions. Because obstructed and/or segmented apertures are likely for future ground- and space-based observatories, the achievable raw contrast with a coronagraph is expected to be at or worse than the brightness of the exoplanets. For the WFIRST CGI, the desired contrast improvement factor is 10 to 30 based on an expected raw contrast in the range of $10^{-8}$ to $10^{-9}$.\cite{78} Of the existing post-processing techniques, reference differential imaging (RDI) and angular differential imaging (ADI) are currently the most promising for the WFIRST CGI.\cite{130, 131, 132} RDI would be limited by a small library of images or from large wavefront changes in the weeks between observations. ADI will be feasible as long as the observatory’s pointing and thermal loading can be held steady during a roll maneuver. Spectral differential imaging (SDI) will not be possible at high contrast until a closed-form expression is found for the intensity versus wavelength.\cite{130}

In this thesis, we demonstrated the first use of recursive coherence differential imaging (RCDI) to detect an injected planet in the HCIL and in simulation. RCDI, subverting the former paradigm of using only post-correction images for science, enables real-time estimation of an exoplanet during correction. In our Chapter 6 Monte Carlo simulations of a simple scenario for the WFIRST CGI, we demonstrated that RCDI can efficiently detect planets while ruling out false alarms. We have barely begun to explore the utility of CDI and RCDI, and in future work we plan to explore the capabilities of RCDI in a more sophisticated simulation of the WFIRST CGI including LOWFSC and in simulations of ground-based observatories with ExAO. We would like to perform RCDI experiments in the JPL’s HCIT, although the planet signal would have to be injected in software without a major hardware modification. If the post-ExAO wavefronts are stable enough, we would eventually like to demonstrate RCDI on current ground-based observatories with first-generation ExAO systems such as the Gemini South Telescope and Subaru Telescope.

Unlike the conventional post-processing techniques, RCDI is less accessible to astronomers.
because performing it requires more than a set of images from an observatory. To utilize knowledge of the DM response and the electric field, an optical model of the instrument is also required. Because instrument behavior is not generally released (or sometimes even recorded), most researchers are not experienced in wavefront correction, and most (or possibly all) current telescopes are not stable enough for RCDI, no one has used RCDI before for the detection of real exoplanets. We would like to advocate for the release of the control Jacobian, DM voltage maps, and electric field estimates along with the actual images to enable the average researcher to perform RCDI with minimal effort. The linearized optical model with just the control Jacobian and DM command maps is less accurate than the full nonlinear optical model, but at high contrast the error is minimal and the simple model would be easy to implement in a few lines of code. We would like to show both that RCDI is a useful detection method and that it can be easy to implement if the relevant instrument settings are provided.
Chapter 8

Conclusions

With the prospect of the Wide-Field Infrared Survey Telescope Coronagraph Instrument flying in less than a decade, efficient focal plane wavefront correction (FPWC) and exoplanet detection algorithms are critical for maximizing the mission’s science yield. The slowest part of FPWC during a space mission will be obtaining images for wavefront estimation, so our goal was to improve the estimator both by reducing the total required exposure time and by using the correction images for exoplanet detection. In this thesis, we developed and implemented an iterated extended Kalman filter (IEKF) based on an established pair-wise probing estimation scheme. Our IEKF formulation constructs recursive, near-optimal estimates of both the stellar electric field and the incoherent bias signal from all images taken during wavefront estimation. We showed that the IEKF, as one of the simplest nonlinear filters, should be within the capabilities of a space telescope’s onboard computer. In several simulations and laboratory experiments, we found that the recursive nature of the IEKF and the previously developed Kalman filter (KF) enabled significantly faster speckle suppression than with a batch process estimator. The ability of the KF and IEKF to average out read noise and photon shot noise over many correction iterations allowed much shorter exposure times. The IEKF provided the fastest correction with large model error in our lab at Princeton, but the KF and IEKF gave equally fast correction with small model error in
the well-calibrated testbed experiments at the Jet Propulsion Laboratory. In simulations with the shortest possible exposure times and little model error, the KF provided slightly faster correction than the IEKF. We therefore expect the KF and IEKF to have comparable correction speeds for a space mission. The major benefit of the IEKF is recursive estimation of the incoherent light, which includes any exoplanets. The IEKF’s bias estimation provides an alternative, coherence-based methodology for separating exoplanet light from the stellar speckles and can be performed during what is normally considered lost overhead time for wavefront correction. In Monte Carlo simulations of FPWC in a simple coronagraphic system, we proved the viability of this recursive coherence differential imaging (RCDI) technique for the first time. RCDI was reasonably efficient at detecting the planet signal and ruling out false alarms in the static cases tested, and it has the potential to outperform conventional planet detection methods in more realistic cases with dynamic wavefront disturbances. Our development of a fast, nonlinear estimator that recursively estimates incoherent light may enable the discovery and characterization of more exoplanets in upcoming coronagraphy missions.
Appendix A

Derivations of Estimators

A.1 The Batch Process Estimator

This section details the procedure for calculating the least-squares-error state estimate at a discrete time step $k$ from a single batch of measurements. The noisy measurement of the $N_{\text{state}} \times 1$-dimension state vector $x_k$ to be estimated is

$$z_k = H_k x_k + n_k,$$

(A.1.1)

where $z_k$ is the $N_{\text{meas}} \times 1$ vector of measurements; $n_k$ is the $N_{\text{meas}} \times 1$ vector of additive, Gaussian, zero-mean noise; and $H_k$ is the $N_{\text{meas}} \times N_{\text{state}}$ observation matrix. The true state is not known (or else estimation would be unnecessary), so the squared error between the observed and expected measurement is minimized instead. The least-squares state estimate minimizes the square of the error in the measurement, $(z_k - H_k \hat{x}_k)^T(z_k - H_k \hat{x}_k)$, where $\hat{x}_k$ is the estimated state vector, when measurements are all equally noisy. If individual measurements have different RMS levels of noise, the noisier measurements should be weighted less. The batch process estimate of $x_k$ minimizes the weighted square of the error,

$$J_{BPE,k} = \frac{1}{2}(z_k - H_k \hat{x}_k)^T R_k^{-1}(z_k - H_k \hat{x}_k),$$

(A.1.2)
where the weight $R_k^{-1}$ is the inverse of the measurement noise covariance matrix,

$$R_k = E[n_k n_k^T].$$

(A.1.3)

The cost function $J_{BPE,k}$ is quadratic and convex (because $R_k$ is positive semi-definite by definition), so its minimum is where the gradient equals zero. The partial derivative of the cost with respect to the state estimate is

$$\frac{\partial J_{BPE,k}}{\partial \hat{x}_k} = -H_k^T R_k^{-1} (z_k - H_k \hat{x}_k).$$

(A.1.4)

After evaluating the partial derivative at zero, the state estimate that minimizes the cost is found to be

$$\hat{x}_k = (H_k^T R_k^{-1} H_k)^{-1} H_k^T R_k^{-1} z_k.$$  

(A.1.5)

If the variance of the noise is the same for all measurements, then $R_k = r_k I$, where $r_k$ is a scalar. The batch process estimate then reduces to

$$\hat{x}_k = (H_k^T H_k)^{-1} H_k^T z_k$$

$$= H_k^L z_k,$$

where the superscript $L$ denotes the left pseudo-inverse.

A.2 The Kalman Filter

The Kalman filter (KF) is the minimum-variance state estimator for a discrete-time, linear system with additive, Gaussian noise sources.[137] The KF is used for sequential state estimation, in which it is computationally wasteful to use all past measurements in the BPE when the state was already estimated at the previous time step. That is, the KF optimally
updates the previous estimate of the state with the new measurements. The KF also takes into account state dynamics and control inputs in the system, which are used to propagate the previous state estimate forward to the current time step. Here we derive the KF based on the derivations and notations of others.[161, 162, 106]

As in the BPE of Appendix A.1, the $N_{meas} \times 1$-dimension, linear, noisy measurement vector $z_k$ is

$$z_k = H_k x_k + n_k.$$  \hfill (A.2.1)

The linear, dynamic state equation is

$$x_k = \Phi x_{k-1} + \Gamma u_{k-1} + \Lambda w_{k-1},$$  \hfill (A.2.2)

where $\Phi$ is the $N_{state} \times N_{state}$ state transition matrix, $u_{k-1}$ is the $N_{ctrl} \times 1$ control signal vector, $\Gamma$ is the $N_{state} \times N_{ctrl}$ control response matrix, and $\Lambda$ is the $N_{state} \times N_{dist}$ disturbance response matrix. The $N_{dist} \times 1$ vector $w_{k-1}$ is Gaussian, zero-mean process noise from disturbances.

The state vector and state covariance matrix are estimated in two steps at each sampling point in time. In the first step, known as the time update and denoted by $(-)$, the linear, dynamic model propagates the state and state covariance to the current time step. In the second step, known as the measurement update and denoted by $(+)$, the estimate is updated to include the measurements from the current time step.

The state estimate time update, following the form of Eq. A.2.2 but omitting the disturbances because they are unknown and zero-mean, simply propagates the previous estimated state and is written as

$$\hat{x}_k(-) = \Phi \hat{x}_{k-1}(+) + \Gamma u_{k-1}.$$  \hfill (A.2.3)
The state covariance matrix time update is then calculated as

\[ P_k(-) = E[(\hat{x}_k(-) - x_k)(\hat{x}_k(-) - x_k)^T] \]
\[ = E[(\Phi(\hat{x}_{k-1}(+) - x_{k-1}) - \Lambda w_{k-1})(\Phi(\hat{x}_{k-1}(+) - x_{k-1}) - \Lambda w_{k-1})^T] \] (A.2.4)
\[ = \Phi E[(\hat{x}_{k-1}(+) - x_{k-1})(\hat{x}_{k-1}(+) - x_{k-1})^T] \Phi^T + \Gamma E[w_{k-1}w_{k-1}^T] \Gamma^T 
+ \Phi E[(\hat{x}_{k-1}(+) - x_{k-1})w_{k-1}^T] \Gamma^T + \Gamma E[w_{k-1}(\hat{x}_{k-1}(+) - x_{k-1})^T] \Phi^T 
= \Phi E[(\hat{x}_{k-1}(+) - x_{k-1})(\hat{x}_{k-1}(+) - x_{k-1})^T] \Phi^T + \Gamma E[w_{k-1}w_{k-1}^T] \Gamma^T 
= \Phi P_{k-1}(+) \Phi^T + Q_{k-1}, \] (A.2.5)

where the process noise covariance matrix is defined as

\[ Q_{k-1} = \Lambda E[w_{k-1}w_{k-1}^T] \Lambda^T. \] (A.2.6)

The process noise and state estimate error are assumed to be uncorrelated such that

\[ E[(\hat{x}_{k-1} - x_{k-1})w_{k-1}^T] = 0. \] (A.2.7)

The measurement update of the state minimizes a new cost function,

\[ J_{KF,k} = \frac{1}{2}(\hat{x}_k - \hat{x}_k(-))^T P_k^{-1}(-)(\hat{x}_k - \hat{x}_k(-)) + \frac{1}{2}(z_k - H_k\hat{x}_k)^T R_k^{-1}(z_k - H_k\hat{x}_k). \] (A.2.8)

The second term of \( J_{KF,k} \) is the same as \( J_{BPE,k} \) from Eq. A.1.2, and the new first term normalizes the error in the state estimate with the inverse of the state covariance matrix as the weight.

Because the cost function is quadratic with respect to the state estimate and convex (because \( P_k(-) \) and \( R_k \) are positive semi-definite by definition), the minimum of \( J_{KF,k} \) is where its gradient equals zero. The partial derivative of the cost function with respect to \( \hat{x}_k \)
\[ \frac{\partial J_{KF,k}}{\partial \hat{x}_k} = P_k^{-1}(-)(\hat{x}_k - \hat{x}_k(-)) - H_k^T R_k^{-1}(z_k - H_k \hat{x}_k). \]  

(A.2.9)

After evaluating the partial derivative at zero, the measurement update of the state estimate is

\[ \hat{x}_k(+) = (P_k^{-1}(-) + H_k^T R_k^{-1} H_k)^{-1} P_k^{-1}(-) \hat{x}_k(-) \]
\[ + (P_k^{-1}(-) + H_k^T R_k^{-1} H_k)^{-1} H_k^T R_k^{-1} z_k. \]  

(A.2.10)

To simplify the form, the matrix inversion lemma is utilized on the inverted term,

\[ (P_k^{-1}(-) + H_k^T R_k^{-1} H_k)^{-1} = P_k(-) - P_k(-) H_k^T [H_k P_k(-) H_k^T + R_k]^{-1} H_k P_k(-), \]  

(A.2.11)

and substituted back into Eq. A.2.9 to obtain

\[ \hat{x}_k(+) = (P_k(-) - P_k(-) H_k^T [H_k P_k(-) H_k^T + R_k]^{-1} H_k P_k(-)) P_k^{-1}(-) \hat{x}_k(-) \]
\[ + (P_k(-) - P_k(-) H_k^T [H_k P_k(-) H_k^T + R_k]^{-1} H_k P_k(-)) H_k^T R_k^{-1} z_k \]
\[ = (\mathbb{I} - P_k(-) H_k^T [H_k P_k(-) H_k^T + R_k]^{-1} H_k) \hat{x}_k(-) \]
\[ + P_k(-) H_k^T (\mathbb{I} - [H_k P_k(-) H_k^T + R_k]^{-1} H_k P_k(-) H_k^T) R_k^{-1} z_k \]
\[ = (\mathbb{I} - K_k H_k) \hat{x}_k(-) \]
\[ + P_k(-) H_k^T (\mathbb{I} - [H_k P_k(-) H_k^T + R_k]^{-1} (H_k P_k(-) H_k^T + R_k - R_k)) R_k^{-1} z_k \]
\[ = (\mathbb{I} - K_k H_k) \hat{x}_k(-) + P_k(-) H_k^T (\mathbb{I} - \mathbb{I} + [H_k P_k(-) H_k^T + R_k]^{-1}) R_k R_k^{-1} z_k \]
\[ = (\mathbb{I} - K_k H_k) \hat{x}_k(-) + K_k z_k \]
\[ = \hat{x}_k(-) + K_k (z_k - H_k \hat{x}_k(-)). \]  

(A.2.12)

(A.2.13)
The optimal measurement update is thus the time update plus the Kalman gain, defined as

\[ K_k = P_k(-)H_k^T[H_kP_k(-)H_k^T + R_k]^{-1}, \]  

(A.2.14)
times the residual of the measurement and estimated observation.

Defining the error between the state estimate time update and the true state with a tilde as

\[ \tilde{x}_k(-) = \hat{x}_k(-) - x_k \]  

(A.2.15)
and using the definition of the Kalman gain in Eq. A.2.14, the measurement update of the state covariance matrix is found to be

\[
P_k(+) = E[(\hat{x}_k(+)-x_k)(\hat{x}_k(+)-x_k)^T] \\
= E[(\tilde{x}_k(-)+K_k(z_k-H_k\hat{x}_k(-)))(\tilde{x}_k(-)+K_k(z_k-H_k\hat{x}_k(-)))^T] \\
= E[(\tilde{x}_k(-)+K_k(H_k\tilde{x}_k(-)+n_k))(\tilde{x}_k(-)+K_k(H_k\tilde{x}_k(-)+n_k))^T] \\
= E[\tilde{x}_k(-)\tilde{x}_k(-)^T] + K_k(H_k E[\tilde{x}_k(-)\tilde{x}_k(-)^T]H_k^T + E[n_kn_k^T])K_k^T \\
- K_kH_kE[\tilde{x}_k(-)\tilde{x}_k(-)^T] - E[\tilde{x}_k(-)\tilde{x}_k(-)^T]H_k^TK_k^T \\
= P_k(-) + K_k(H_kP_k(-)H_k^T + R)K_k^T - K_kH_kP_k(-) - P_k(-)H_k^TK_k^T \\
= P_k(-) - K_kH_kP_k(-) + P_k(-)H_kK_k^T - P_k(-)H_k^TK_k^T \\
= (I - K_kH_k)P_k(-). \]  

(A.2.16)
For clarity, all five KF equations written together are

\[ \hat{x}_k(-) = \Phi \hat{x}(+)_{k-1} + \Gamma u_{k-1} \]  
(A.2.17)

\[ P_k(-) = \Phi P_{k-1}(+)\Phi^T + Q_{k-1} \]  
(A.2.18)

\[ K_k = P_k(-)H_k^T[H_k P_k(-)H_k^T + R_k]^{-1} \]  
(A.2.19)

\[ \hat{x}_k(+) = \hat{x}_k(-) + K_k[z_k - H_k \hat{x}_k(-)] \]  
(A.2.20)

\[ P_k(+) = \| - K_k H_k \| P_k(-). \]  
(A.2.21)

Without an initial estimate, the application of the KF begins on the second time step. The BPE provides the state estimate for the first correction iteration. The covariance of the batch process estimate is then calculated from Eq. A.1.5 as

\[ P_{BPE,k} = E[(\hat{x}_k - x_k)(\hat{x}_k - x_k)^T) \]

\[ = E[((H_k^T R_k^{-1} H_k)^{-1}H_k^T R_k^{-1} z_k - x_k)((H_k^T R_k^{-1} H_k)^{-1}H_k^T R_k^{-1} z_k - x_k)^T] \]

\[ = (H_k^T R_k^{-1} H_k)^{-1}H_k^T R_k^{-1} E[n_k n_k^T] R_k^{-1} H_k (H_k^T R_k^{-1} H_k)^{-1} \]

\[ = (H_k^T R_k^{-1} H_k)^{-1}, \]  
(A.2.22)

which gives

\[ P_1(+) = (H_1^T R_1^{-1} H_1)^{-1} \]  
(A.2.23)

as the initial state covariance matrix for the KF. To compare the covariance estimates of the BPE and KF directly, the Kalman gain in Eq. A.2.19 is substituted into Eq. A.2.21 and the matrix inversion lemma from Eq. A.2.11 is used to obtain

\[ P_k(+) = [P_k^{-1}(-) + H_k^T R_k^{-1} H_k]^{-1}. \]  
(A.2.24)

It is now clear that the KF provides a smaller state covariance (Eq. A.2.24) than the BPE (Eq. A.2.22) since \( P_k(-) \) is a positive semi-definite matrix, as indicated by its definition in
For future reference in the derivation of the iterated extended Kalman filter, substituting Eq. A.2.24 into Eq. A.2.10 yields an alternative form of the Kalman gain,

\[
\hat{x}_k(+) = P_k(+) P_k(+) - 1 \hat{x}_k(-) + P_k(+) H_k^T R_k - 1 z_k
\]

\[[\mathbb{I} - K_k H_k] \hat{x}_k(-) + P_k(+) H_k^T R_k - 1 z_k.\]  

(A.2.25)

Comparing Eq. A.2.25 with Eq. A.2.12, we see that the Kalman gain can also be formulated as

\[
K_k = P_k(+) H_k^T R_k - 1.
\]  

(A.2.26)

### A.3 The Extended Kalman Filter

The extended Kalman filter (EKF) is one of the simplest nonlinear estimators. When necessary, it linearizes the system so that the Kalman filter equations can be used. The EKF is thus a sub-optimal estimator because it only approximately finds the best estimate. Here we derive the EKF for linear state dynamics and a nonlinear observation, but in general the EKF can also incorporate nonlinear state dynamics.

The \(N_{\text{meas}} \times 1\)-dimension, nonlinear, noisy measurement vector \(z_k\) is

\[
z_k = h(x_k) + n_k, \tag{A.3.1}
\]

where \(n_k\) is the \(N_{\text{meas}} \times 1\) additive, Gaussian, zero-mean noise vector and \(h(x_k)\) is the nonlinear measurement function that converts the \(N_{\text{state}} \times 1\) state vector into a \(N_{\text{meas}} \times 1\) output vector. The linear, dynamic state equation is still the same as in Eq. A.2.2,

\[
x_k = \Phi x_{k-1} + \Gamma u_{k-1} + \Lambda w_{k-1}. \tag{A.3.2}
\]
where Φ is the $N_{state} \times N_{state}$ state transition matrix, $u_{k-1}$ is the $N_{ctrl} \times 1$ control signal vector, Γ is the $N_{state} \times N_{ctrl}$ control response matrix, and Λ is the $N_{state} \times N_{dist}$ disturbance response matrix. The $N_{dist} \times 1$ vector $w_{k-1}$ is Gaussian, zero-mean process noise from disturbances. Because we assume the same linear state dynamics from the KF for the EKF, the time update equations for the state and state covariance are the same as Eqs. A.2.17 and A.2.18, respectively.

The cost function for the EKF state estimate,

$$J_{EKF,k} = \frac{1}{2} (\hat{x}_k - \hat{x}_k(-))^T P_k^{-1}(-)(\hat{x}_k - \hat{x}_k(-)) + \frac{1}{2} (z_k - h(\hat{x}_k))^T R_k^{-1}(z_k - h(\hat{x}_k)), \tag{A.3.3}$$

includes the nonlinear observation function. The partial derivative of the cost function with respect to $\hat{x}_k$ is

$$\frac{\partial J_{EKF,k}}{\partial \hat{x}_k} = P_k^{-1}(-)(\hat{x}_k - \hat{x}_k(-)) - \frac{\partial h(\hat{x}_k)}{\partial \hat{x}_k} R_k^{-1}(z_k - h(\hat{x}_k))$$

$$\approx P_k^{-1}(-)(\hat{x}_k - \hat{x}_k(-)) - \frac{\partial h(\hat{x}_k)}{\partial \hat{x}_k} \bigg|_{\hat{x}_k = \hat{x}_k(-)} R_k^{-1}(z_k - h(\hat{x}_k))$$

$$= P_k^{-1}(-)(\hat{x}_k - \hat{x}_k(-)) - H_k^T R_k^{-1}(z_k - h(\hat{x}_k)), \tag{A.3.4}$$

where the partial derivative of the observation about point $\hat{x}_k(-)$ gives the approximate observation matrix of

$$H_k = \frac{\partial h(\hat{x}_k)}{\partial \hat{x}_k} \bigg|_{\hat{x}_k = \hat{x}_k(-)}. \tag{A.3.5}$$

The strategy for finding the optimal KF estimate relied on the cost function being quadratic and the partial derivative being linear with respect to the new state estimate. The partial derivative for the EKF is nonlinear, so it requires an approximation to evaluate it at zero and solve for $\hat{x}_k$. The approximation means that, in general, the EKF’s state measurement update does not fully minimize the cost function and is sub-optimal.
To solve for the new estimate, the nonlinear observation is first linearized about the state estimate time update as

$$h(\hat{x}_k) \approx h(\hat{x}_k(-)) + H_k(\hat{x}_k - \hat{x}_k(-)). \quad (A.3.6)$$

After substituting Eq. A.3.6 into Eq. A.3.4, evaluating Eq. A.3.4 at zero, and rearranging, the state estimate measurement update becomes

$$\hat{x}_k(+) = \hat{x}_k(-) + [P_k^{-1}(-) + H_k^T R_k^{-1} H_k]^{-1} H_k^T R_k^{-1} (z_k - h(\hat{x}_k(-)))$$

$$\approx \hat{x}_k(-) + [P_k^{-1}(-) + H_k^T R_k^{-1} H_k]^{-1} H_k^T R_k^{-1} (z_k - h(\hat{x}_k(-)) + H_k(\hat{x}_k(-) - \hat{x}_k(-)))$$

$$= \hat{x}_k(-) + K_k(z_k - h(\hat{x}_k(-))), \quad (A.3.7)$$

where Eq. A.3.6 is evaluated at the time update. We temporarily define the EKF’s Kalman gain as

$$K_k = [P_k^{-1}(-) + H_k^T R_k^{-1} H_k]^{-1} H_k^T R_k^{-1}. \quad (A.3.8)$$

Linearizing the measurement of the true state about the point $\hat{x}_k(-)$ as

$$h(x_k) \approx h(\hat{x}_k(-)) + H_k(x_k - \hat{x}_k(-)), \quad (A.3.9)$$
the measurement update of the state covariance matrix is found to be

\[
P_k(+) = E[(\hat{x}_k(+) - x_k)(\hat{x}_k(+) - x_k)^T]
\]

\[
= E[(\tilde{x}_k(-) + K_k(z_k - h(\hat{x}_k(-)))(\tilde{x}_k(-) + K_k(z_k - h(\hat{x}_k(-))))^T]
\]

\[
\approx E[(\tilde{x}_k(-) + K_k(h(x_k) + n_k - h(\hat{x}_k(-)))(\tilde{x}_k(-) + K_k(h(x_k)
\]

\[
+ n_k - h(\hat{x}_k(-))))^T]
\]

\[
\approx E[(\tilde{x}_k(-) + K_k(H_k\tilde{x}_k(-) + n_k)(\tilde{x}_k(-) + K_k(H_k\tilde{x}_k(-) + n_k))^T]
\]

\[
= (I - K_kH_k) P_k(-), \tag{A.3.10}
\]

where the final steps in solving for \( P_k(+) \) are exactly the same as for the KF in Eq. A.2.16. Since the EKF’s state covariance update equation is the same as the KF’s in Eq. A.2.21, the alternate definition in Eq. A.2.24 is also true for the EKF. Substituting Eq. A.2.24 into Eq. A.3.8, the EKF’s Kalman gain becomes

\[
K_k = P_k(+)H_k^T R_k^{-1}, \tag{A.3.11}
\]

the same as for the KF in Eq. A.2.26, so the Kalman gain from Eq. A.2.19 is also valid for the EKF.

In summary, the five EKF equations for a linear dynamic system and a nonlinear mea-
The equations have the same form as those in the KF except that the linear observation
$H_k \hat{x}_k(-)$ is replaced by the nonlinear observation $h(\hat{x}_k(-))$ in the state measurement update.
The observation matrix $H_k$ is an approximation in the EKF, whereas it is exact in the KF.

**A.4 The Iterated Extended Kalman Filter**

The sub-optimality of the EKF estimate arises from solving the cost function in Eq. A.3.3
as though the observation were linear. The cost can be better minimized with a nonlinear,
iterative solver. Here we implement the Gauss-Newton method, which uses the first and
second partial derivatives of the cost to perform an iterated gradient search. The Gauss-
Newton method for minimizing the EKF cost function is called the iterated extended Kalman
filter (IEKF). The IEKF is usually derived differently from the method shown below, but
Bell and Cathey[142] demonstrated that the well-known IEKF is in fact a Gauss-Newton
minimization.

The updated cost function from the EKF case is

$$J_{k,i} = \frac{1}{2} (\hat{x}_{k,i} - \hat{x}_k(-))^T P_k^{-1}(-)(\hat{x}_{k,i} - \hat{x}_k(-)) + \frac{1}{2} (z_k - h(\hat{x}_{k,i}))^T R_k^{-1}(z_k - h(\hat{x}_{k,i})),$$

(A.4.1)

which now includes the subscript $i$ to denote the Gauss-Newton iteration of the state estimate.
measurement update. We drop the (+) notation from the iterated state measurement update for clarity. Similar to the EKF case, the partial derivative of the cost with respect to \( \hat{x}_{k,i} \) is

\[
G_{k,i} = \frac{\partial J_{k,i}}{\partial \hat{x}_{k,i}}
\]

\[
= P_{k}^{-1}(-)(\hat{x}_{k,i} - \hat{x}_{k}(-)) - \frac{\partial h(\hat{x}_{k})}{\partial \hat{x}_{k}} R_{k}^{-1}(z_{k} - h(\hat{x}_{k,i}))
\]

\[
\approx P_{k}^{-1}(-)(\hat{x}_{k,i} - \hat{x}_{k}(-)) - \left. \frac{\partial h(\hat{x}_{k})}{\partial \hat{x}_{k}} \right|_{\hat{x}_{k} = \hat{x}_{k,i}} R_{k}^{-1}(z_{k} - h(\hat{x}_{k,i}))
\]

\[
= P_{k}^{-1}(-)(\hat{x}_{k,i} - \hat{x}_{k}(-)) - H_{k,i}^{T} R_{k}^{-1}(z_{k} - h(\hat{x}_{k,i})),
\] (A.4.2)

where the approximate, linearized observation matrix is

\[
H_{k,i} = \frac{\partial h(\hat{x}_{k})}{\partial \hat{x}_{k}} \bigg|_{\hat{x}_{k} = \hat{x}_{k,i}}.
\] (A.4.3)

The approximate second derivative of the cost with respect to \( \hat{x}_{k,i} \) is

\[
H_{k,i} = \frac{\partial^{2} J_{k,i}}{\partial \hat{x}_{k,i}^{2}}
\]

\[
= \frac{\partial G_{k,i}}{\partial \hat{x}_{k,i}}
\]

\[
\approx \left. \frac{\partial}{\partial \hat{x}_{k,i}} \left( P_{k}^{-1}(-)(\hat{x}_{k,i} - \hat{x}_{k}(-)) - H_{k,i}^{T} R_{k}^{-1}(z_{k} - h(\hat{x}_{k,i})) \right) \right|_{\hat{x}_{k} = \hat{x}_{k,i}}
\]

\[
= P_{k}^{-1}(-) + H_{k,i}^{T} R_{k}^{-1} \frac{\partial h(\hat{x}_{k})}{\partial \hat{x}_{k}}
\]

\[
\approx P_{k}^{-1}(-) + H_{k,i}^{T} R_{k}^{-1} \left. \frac{\partial h(\hat{x}_{k})}{\partial \hat{x}_{k}} \right|_{\hat{x}_{k} = \hat{x}_{k,i}}
\]

\[
= P_{k}^{-1}(-) + H_{k,i}^{T} R_{k}^{-1} H_{k,i}.
\] (A.4.4)

Each step of the Gauss-Newton gradient search is defined as the initial estimate minus the inverse of the Hessian times the gradient,[162]

\[
\hat{x}_{k,i+1} = \hat{x}_{k,i} - H_{k,i}^{-1} G_{k,i}.
\] (A.4.5)
It is helpful to calculate the covariance of \( \hat{x}_{k,i+1} \) next before expanding the expression for \( \hat{x}_{k,i+1} \) in Eq. A.4.5. The error in the estimate compared to the true state is given as

\[
\tilde{x}_{k,i+1} = \hat{x}_{k,i+1} - x_k \\
= \hat{x}_{k,i} - x_k - H_{k,i}^{-1} \hat{G}_{k,i} \\
= \hat{x}_{k,i} - x_k - H_{k,i}^{-1}(P_{k}^{-1}(-)(\hat{x}_{k,i} - \hat{x}_k(-)) - H^T_{k,i} R^{-1}_k(z_k - h(\hat{x}_{k,i})) \\
= \hat{x}_{k,i} - x_k - H_{k,i}^{-1} P_{k}^{-1}(-(\hat{x}_{k,i} - x_k) - (\hat{x}_k(-) - x_k))) \\
+ H_{k,i}^{-1} H^T_{k,i} R^{-1}_k(z_k - h(\hat{x}_{k,i})) \\
= (I - H_{k,i}^{-1} P_{k}^{-1}(-))\hat{x}_{k,i} + H_{k,i}^{-1} P_{k}^{-1}(-)\tilde{x}_{k,0} \\
+ H_{k,i}^{-1} H^T_{k,i} R^{-1}_k(h(x_k) + n_k - h(\hat{x}_{k,i})) \\
\approx (I - H_{k,i}^{-1} P_{k}^{-1}(-))\hat{x}_{k,i} + H_{k,i}^{-1} P_{k}^{-1}(-)\tilde{x}_{k,0} + H_{k,i}^{-1} H^T_{k,i} R^{-1}_k(h(\hat{x}_{k,i})) \\
+ H_{k,i}(x_k - \hat{x}_{k,i}) + n_k - h(\hat{x}_{k,i}) \\
= (I - H_{k,i}^{-1}(P_{k}^{-1}(-) + H^T_{k,i} R^{-1}_k H_{k,i}))\hat{x}_{k,i} + H_{k,i}^{-1} P_{k}^{-1}(-)\tilde{x}_{k,0} \\
+ H_{k,i}^{-1} H^T_{k,i} R^{-1}_k n_k \\
= H_{k,i}^{-1} P_{k}^{-1}(-)\tilde{x}_{k,0} + H_{k,i}^{-1} H^T_{k,i} R^{-1}_k n_k. \\
(A.4.6)
\]

where we use the shorthand notation of

\[
\tilde{x}_{k,i} = \hat{x}_{k,i} - x_k \\
\hat{x}_{k,0} = \hat{x}_k(-).
(A.4.7)
(A.4.8)
From Eq. A.4.6, the updated state covariance after each Gauss-Newton iteration is

\[ P_{k,i+1} = E[(\tilde{x}_{k,i+1})(\tilde{x}_{k,i+1})^T] \]

\[ = E[\mathcal{H}_{k,i}^{-1} P_k^{-1}(-)\tilde{x}_{k,0} \tilde{x}_{k,0}^T P_k^{-1}(-)\mathcal{H}_{k,i}^{-1}] + E[\mathcal{H}_{k,i}^{-1} H_k^T R_k^{-1} n_k n_k^T R_k^{-1} H_k \mathcal{H}_{k,i}^{-1}] \]

\[ = \mathcal{H}_{k,i}^{-1} P_k^{-1}(-) \mathcal{H}_{k,i}^{-1} + \mathcal{H}_{k,i}^{-1} H_k^T R_k^{-1} H_k \mathcal{H}_{k,i}^{-1} \]

\[ = H_{k,i}^{-1} \]

\[ = (P_k^{-1}(-) + H_{k,i}^T R_k^{-1} H_{k,i})^{-1}. \]

(A.4.9)

Since Eq. A.4.9 has the same form as the covariance measurement update for the KF and EKF, the alternate definition from Eq. A.3.16 is still valid and gives

\[ P_{k,i+1}(+) = [I - K_{k,i} H_{k,i}] P_k(-), \]

(A.4.10)

where the Kalman gain formula from Eq. A.2.26 also still applies as

\[ K_{k,i} = P_{k,i+1}(+) H_{k,i}^T R_k^{-1}. \]

(A.4.11)

Substituting Eqs. A.4.10 and A.4.11 into Eq. A.4.5, the Gauss-Newton state update becomes

\[ \hat{x}_{k,i+1} = \hat{x}_{k,i} - H_{k,i}^{-1} G_{k,i} \]

\[ = \hat{x}_{k,i} - P_{k,i+1}(+) (P_k^{-1}(-)(\hat{x}_{k,i} - \hat{x}_k(-)) - H_{k,i}^T R_k^{-1}(z_k - h(\hat{x}_{k,i}))) \]

\[ = \hat{x}_{k,i} - P_{k,i+1} P_k^{-1}(-)(\hat{x}_{k,i} - \hat{x}_k(-)) + P_{k,i+1} H_{k,i} R_k^{-1}(z_k - h(\hat{x}_{k,i})) \]

\[ = \hat{x}_{k,i} - (I - K_{k,i} H_{k,i})(\hat{x}_{k,i} - \hat{x}_k(-)) + K_{k,i}(z_k - h(\hat{x}_{k,i})) \]

\[ = (I - K_{k,i} H_{k,i}) \hat{x}_k(-) + K_{k,i}(z_k - h(\hat{x}_{k,i})) \]

\[ = \hat{x}_k(-) + K_{k,i} (z_k - h(\hat{x}_{k,i}(+)) + H_{k,i}[\hat{x}_{k,i} - \hat{x}_k(-)]), \]

(A.4.12)

which is the same as for the EKF except for the linear correction term \( K_{k,i} H_{k,i}[\hat{x}_{k,i} - \hat{x}_k(-)] \).
The IEKF equations are grouped together here for reference as

\begin{align*}
H_{k,i} &= \frac{\partial h(\hat{x}_k)}{\partial x_k} \bigg|_{x=\hat{x}_{k,i}(+)} \\
K_{k,i} &= P_k(-)H_{k,i}^T[H_{k,i}P_k(-)H_{k,i}^T + R_k]^{-1} \\
\hat{x}_{k,i+1} &= \hat{x}_k(-) + K_{k,i}(z_k - h(\hat{x}_{k,i}(+)) + H_{k,i}[\hat{x}_{k,i} - \hat{x}_k(-)]) \\
P_{k,i+1(+)} &= \left[ I - K_{k,i}H_{k,i} \right]P_k(-),
\end{align*}

with the initial settings of

\begin{align*}
\hat{x}_{k,i=0} &= \hat{x}_k(-) \\
P_{k,i=0(+)} &= P_k(-).
\end{align*}

If the number of EKF iterations $N_{it}$ equals zero, then the IEKF equations simplify to those of the EKF.

The Gauss-Newton gradient search assumes that the cost function is approximately quadratic, so the IEKF performs best when the nonlinearities are small and the initial EKF state estimate is already near optimal. For highly nonlinear systems, the linearizations may be too inaccurate and lead the IEKF to get stuck in a local minimum. Because of this behavior, the IEKF often produces a better estimate than the EKF but is not guaranteed to do so.
Bibliography


List of Abbreviations

ADI  angular differential imaging ................................................................. 26
ADU  analog-digital units ................................................................. 75
AFTA  Astrophysics Focused Telescope Assets ........................................... 14
AO  adaptive optics ................................................................. 9
AUC  area under the (ROC) curve .............................................................. 155
BMC  Boston Micromachines Corporation ............................................. 13
BPE  batch process estimator ................................................................. 44
BPIE  batch process incoherent estimate ............................................... 95
CCD  charge-coupled device ................................................................. 81
CDI  coherence differential imaging .............................................................. 28
COFFEE  coronagraphic focal plane wavefront estimation for exoplanet detection 24
DM  deformable mirror ................................................................. 13
DRIE  deep reactive-ion etching ................................................................. 118
EFC  electric field conjugation ................................................................. 37
EKF  extended Kalman filter ................................................................. 65
EMCCD  electron-multiplying charge-coupled device ................................ 111
ExAO  extreme AO ................................................................. 10
FPM  focal plane mask ................................................................. 7
FPWC  focal plane wavefront correction ............................................... 9
FT  Fourier transform ................................................................. 18
FWHM  full width at half maximum ......................................................... 9
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>HCIL</td>
<td>High Contrast Imaging Laboratory</td>
<td>15</td>
</tr>
<tr>
<td>HCIT</td>
<td>High Contrast Imaging Testbed</td>
<td>15</td>
</tr>
<tr>
<td>HLC</td>
<td>hybrid Lyot coronagraph</td>
<td>8</td>
</tr>
<tr>
<td>IEKF</td>
<td>iterated extended Kalman filter</td>
<td>74</td>
</tr>
<tr>
<td>IFS</td>
<td>integral field spectrograph</td>
<td>117</td>
</tr>
<tr>
<td>IWA</td>
<td>inner working angle</td>
<td>116</td>
</tr>
<tr>
<td>JPL</td>
<td>Jet Propulsion Laboratory</td>
<td>15</td>
</tr>
<tr>
<td>KF</td>
<td>Kalman filter</td>
<td>58</td>
</tr>
<tr>
<td>KLIP</td>
<td>Karhunen-Loève image projection</td>
<td>27</td>
</tr>
<tr>
<td>LOCI</td>
<td>locally optimized combination of images</td>
<td>27</td>
</tr>
<tr>
<td>LOWFSC</td>
<td>low-order wavefront sensing and control</td>
<td>9</td>
</tr>
<tr>
<td>MDL</td>
<td>Microdevices Laboratory</td>
<td>117</td>
</tr>
<tr>
<td>MEMS</td>
<td>microelectromechanical systems</td>
<td>13</td>
</tr>
<tr>
<td>NASA</td>
<td>National Aeronautics and Space Administration</td>
<td>14</td>
</tr>
<tr>
<td>NEC</td>
<td>noise-equivalent contrast</td>
<td>51</td>
</tr>
<tr>
<td>NI</td>
<td>normalized intensity</td>
<td>122</td>
</tr>
<tr>
<td>OAP</td>
<td>off-axis parabola</td>
<td>81</td>
</tr>
<tr>
<td>PAF</td>
<td>parameter-adaptive filter</td>
<td>87</td>
</tr>
<tr>
<td>PIAA</td>
<td>phase-induced amplitude apodization</td>
<td>7</td>
</tr>
<tr>
<td>PMN</td>
<td>lead magnesium niobate</td>
<td>13</td>
</tr>
<tr>
<td>PS</td>
<td>PSF subtraction</td>
<td>95</td>
</tr>
<tr>
<td>PSD</td>
<td>power-spectral density</td>
<td>121</td>
</tr>
<tr>
<td>PSF</td>
<td>point-spread function</td>
<td>4</td>
</tr>
<tr>
<td>RCDI</td>
<td>recursive coherence differential imaging</td>
<td>144</td>
</tr>
<tr>
<td>RDI</td>
<td>reference differential imaging</td>
<td>26</td>
</tr>
<tr>
<td>RIE</td>
<td>recursive incoherent estimate</td>
<td>95</td>
</tr>
<tr>
<td>RMS</td>
<td>root mean square</td>
<td>107</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
<td>Page</td>
</tr>
<tr>
<td>--------------</td>
<td>---------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>ROC</td>
<td>receiver operating characteristic</td>
<td>150</td>
</tr>
<tr>
<td>RV</td>
<td>radial velocity</td>
<td>1</td>
</tr>
<tr>
<td>sCMOS</td>
<td>scientific complementary metal-oxide semiconductor</td>
<td>131</td>
</tr>
<tr>
<td>SDI</td>
<td>spectral differential imaging</td>
<td>26</td>
</tr>
<tr>
<td>SNR</td>
<td>signal-to-noise ratio</td>
<td>24</td>
</tr>
<tr>
<td>SOI</td>
<td>silicon-on-insulator</td>
<td>118</td>
</tr>
<tr>
<td>SP</td>
<td>shaped pupil</td>
<td>7</td>
</tr>
<tr>
<td>SPLC</td>
<td>shaped pupil Lyot coronagraph</td>
<td>8</td>
</tr>
<tr>
<td>TPF</td>
<td>Terrestrial Planet Finder</td>
<td>14</td>
</tr>
<tr>
<td>WFIRST</td>
<td>Wide-Field Infrared Survey Telescope</td>
<td>14</td>
</tr>
<tr>
<td>WFIRST CGI</td>
<td>Wide-Field Infrared Survey Telescope Coronagraph Instrument</td>
<td>14</td>
</tr>
<tr>
<td>WFS</td>
<td>wavefront sensor</td>
<td>10</td>
</tr>
</tbody>
</table>