

AFIT/DS/ENG/98-06

Linear Reconstruction
of Non-Stationary Image Ensembles
Incorporating Blur and Noise Models

DISSERTATION
Stephen D. Ford
Captain, USAF

AFIT/DS/ENG/98-06

DTIC QUALITY INSPECTED 2

19980519 089

Approved for public release; distribution unlimited

| REPORT DOCUMENTATION PAGE | | | Form Approved OMB No. 0704-0188 | |
|---|--|---|---|--|
| Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503. | | | | |
| 1. AGENCY USE ONLY (Leave blank) | 2. REPORT DATE March 1998 | 3. REPORT TYPE AND DATES COVERED Doctoral Dissertation | | |
| 4. TITLE AND SUBTITLE LINEAR RECONSTRUCTION OF NON-STATIONARY IMAGE ENSEMBLES INCORPORATING BLUR AND NOISE MODELS | | | 5. FUNDING NUMBERS | |
| 6. AUTHOR(S) Stephen D. Ford, Capt, USAF | | | | |
| 7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Air Force Institute of Technology 2750 P Street WPAFB OH 45433-7765 | | | 8. PERFORMING ORGANIZATION REPORT NUMBER AFIT/DS/ENG/98-06 | |
| 9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Capt Bruce Stribling AFRL/DEBI 535 Lipoa Pkwy, Suite 200 Kihei HI 96753 | | | 10. SPONSORING/MONITORING AGENCY REPORT NUMBER | |
| 11. SUPPLEMENTARY NOTES | | | | |
| 12a. DISTRIBUTION AVAILABILITY STATEMENT Approved for Public Release; Distribution Unlimited | | | 12b. DISTRIBUTION CODE | |
| 13. ABSTRACT (Maximum 200 words) Two new linear reconstruction techniques are developed to improve the resolution of images collected by ground-based telescopes imaging through atmospheric turbulence. The classical approach involves the application of constrained least squares (CLS) to the deconvolution from wavefront sensing (DWFS) technique. The new algorithm incorporates blur and noise models to select the appropriate regularization constant automatically. In all cases examined, the Newton-Raphson minimization converged to a solution in less than 10 iterations. The non-iterative Bayesian approach involves the development of a new vector Wiener filter which is optimal with respect to mean square error (MSE) for a non-stationary object class degraded by atmospheric turbulence and measurement noise. This research involves the first extension of the Wiener filter to account properly for shot noise and an unknown, random optical transfer function (OTF). The vector Wiener filter provides superior reconstructions when compared to the traditional scalar Wiener filter for a non-stationary object class. In addition, the new filter can provide a superresolution capability when the object's Fourier domain statistics are known for spatial frequencies beyond the OTF cutoff. A generalized performance and robustness study of the vector Wiener filter showed that MSE performance is fundamentally limited by object signal-to-noise ratio (SNR) and correlation between object pixels. | | | | |
| 14. SUBJECT TERMS image reconstruction; image restoration; atmospheric turbulence; constrained least squares; Wiener filter | | | 15. NUMBER OF PAGES 158 | |
| | | | 16. PRICE CODE | |
| 17. SECURITY CLASSIFICATION OF REPORT Unclassified | 18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified | 19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified | 20. LIMITATION OF ABSTRACT UL | |

GENERAL INSTRUCTIONS FOR COMPLETING SF 298

The Report Documentation Page (RDP) is used in announcing and cataloging reports. It is important that this information be consistent with the rest of the report, particularly the cover and title page. Instructions for filling in each block of the form follow. It is important to *stay within the lines* to meet *optical scanning requirements*.

Block 1. Agency Use Only (Leave blank).

Block 2. Report Date. Full publication date including day, month, and year, if available (e.g. 1 Jan 88). Must cite at least the year.

Block 3. Type of Report and Dates Covered. State whether report is interim, final, etc. If applicable, enter inclusive report dates (e.g. 10 Jun 87 - 30 Jun 88).

Block 4. Title and Subtitle. A title is taken from the part of the report that provides the most meaningful and complete information. When a report is prepared in more than one volume, repeat the primary title, add volume number, and include subtitle for the specific volume. On classified documents enter the title classification in parentheses.

Block 5. Funding Numbers. To include contract and grant numbers; may include program element number(s), project number(s), task number(s), and work unit number(s). Use the following labels:

C - Contract
G - Grant
PE - Program
Element

PR - Project
TA - Task
WU - Work Unit
Accession No.

Block 6. Author(s). Name(s) of person(s) responsible for writing the report, performing the research, or credited with the content of the report. If editor or compiler, this should follow the name(s).

Block 7. Performing Organization Name(s) and Address(es). Self-explanatory.

Block 8. Performing Organization Report Number. Enter the unique alphanumeric report number(s) assigned by the organization performing the report.

Block 9. Sponsoring/Monitoring Agency Name(s) and Address(es). Self-explanatory.

Block 10. Sponsoring/Monitoring Agency Report Number. (If known)

Block 11. Supplementary Notes. Enter information not included elsewhere such as: Prepared in cooperation with....; Trans. of....; To be published in.... When a report is revised, include a statement whether the new report supersedes or supplements the older report.

Block 12a. Distribution/Availability Statement. Denotes public availability or limitations. Cite any availability to the public. Enter additional limitations or special markings in all capitals (e.g. NOFORN, REL, ITAR).

DOD - See DoDD 5230.24, "Distribution Statements on Technical Documents."

DOE - See authorities.

NASA - See Handbook NHB 2200.2.

NTIS - Leave blank.

Block 12b. Distribution Code.

DOD - Leave blank.

DOE - Enter DOE distribution categories from the Standard Distribution for Unclassified Scientific and Technical Reports.
Leave blank.

NASA - Leave blank.

NTIS -

Block 13. Abstract. Include a brief (*Maximum 200 words*) factual summary of the most significant information contained in the report.

Block 14. Subject Terms. Keywords or phrases identifying major subjects in the report.

Block 15. Number of Pages. Enter the total number of pages.

Block 16. Price Code. Enter appropriate price code (*NTIS only*).

Blocks 17. - 19. Security Classifications. Self-explanatory. Enter U.S. Security Classification in accordance with U.S. Security Regulations (i.e., UNCLASSIFIED). If form contains classified information, stamp classification on the top and bottom of the page.

Block 20. Limitation of Abstract. This block must be completed to assign a limitation to the abstract. Enter either UL (unlimited) or SAR (same as report). An entry in this block is necessary if the abstract is to be limited. If blank, the abstract is assumed to be unlimited.

The views expressed in this dissertation are those of the author and do not reflect the official policy or position of the Department of Defense or the United States Government.

AFIT/DS/ENG/98-06

Linear Reconstruction
of Non-Stationary Image Ensembles
Incorporating Blur and Noise Models

DISSERTATION

Presented to the Faculty of the School of Engineering
of the Air Force Institute of Technology
Air University
In Partial Fulfillment of the
Requirements for the Degree of
Doctor of Philosophy

Stephen D. Ford, B.S.E.E., M.S.E.E.
Captain, USAF

March, 1998

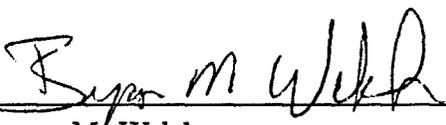
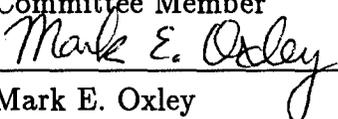
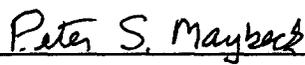
Approved for public release; distribution unlimited

Linear Reconstruction
of Non-Stationary Image Ensembles
Incorporating Blur and Noise Models

Stephen D. Ford, B.S.E.E., M.S.E.E.

Captain, USAF

Approved:

| | |
|--|---------------------------|
|  Bryon M. Welsh Research Advisor | <u>6 March 98</u> Date |
|  Michael C. Roggemann Committee Member | <u>6 March 98</u> Date |
|  Mark E. Oxley Committee Member | <u>6 March 98</u> Date |
|  Peter S. Maybeck Dean's Representative | <u>6 March 98</u> Date |


Robert A. Calico, Jr
Dean

Acknowledgements

I owe a debt of thanks to many people for their help and encouragement during my research phase. First, I am very thankful to have had a great advisor, Dr. Byron Welsh. He always found time to answer my questions no matter how trivial. More importantly, he never stopped asking me questions and challenging my assumptions. I would also like to thank my other committee members, Dr. Michael Roggemann and Dr. Mark Oxley for their ideas and observations. Dr. Roggemann's ideas were especially important in getting this work started. The Air Force Research Laboratory (AFRL) sponsored this research and discussions with AFRL scientists were very helpful. In particular, I would like to thank Dr. Chuck Matson for his insightful comments on my first paper. Thanks also to Mr. Dan Zambon and Mr. Dave Doak for keeping the computers running and providing much needed help in processing my memory intensive algorithms. On a personal level, I would like to thank all the members of Cornerstone Baptist Church for their prayers and friendship. Finally, and most of all, thank you Michele, Travis, Jacob, and Laura for your love and encouragement during these years at AFIT.

Stephen D. Ford

Table of Contents

| | Page |
|--|------|
| Acknowledgements | iii |
| List of Figures | viii |
| List of Tables | xiv |
| Abstract | xv |
| I. Introduction | 1-1 |
| 1.1 Problem Statement | 1-1 |
| 1.2 Justification for Conducting the Proposed Research | 1-5 |
| 1.3 Approach | 1-9 |
| 1.4 Scope and Assumptions | 1-11 |
| 1.5 Significant Contributions and Results | 1-12 |
| 1.6 Summary | 1-13 |
| II. Background | 2-1 |
| 2.1 Introduction | 2-1 |
| 2.2 Image Degradation Model | 2-1 |
| 2.3 Atmospheric Turbulence | 2-5 |
| 2.3.1 Turbulence Statistics | 2-5 |
| 2.3.2 Optical Transfer Function Statistics | 2-9 |
| 2.4 Classical Estimation | 2-12 |
| 2.4.1 Least Squares | 2-12 |
| 2.4.2 Gerchberg-Saxton Algorithms | 2-16 |
| 2.4.3 Maximum Likelihood | 2-17 |
| 2.5 Bayesian Estimation | 2-20 |

| | Page |
|-------|---|
| 2.5.1 | Wiener Filter 2-20 |
| 2.5.2 | Kalman Filter 2-22 |
| 2.5.3 | Maximum A Posteriori 2-22 |
| 2.6 | Summary 2-23 |
| III. | Constrained Least Squares Incorporating Wavefront Sensing 3-1 |
| 3.1 | Introduction 3-1 |
| 3.2 | Traditional Estimator 3-1 |
| 3.3 | Modified Constrained Least Squares Formulation 3-2 |
| 3.3.1 | Problem Statement 3-2 |
| 3.3.2 | Closed Form Solution 3-3 |
| 3.3.3 | Newton-Raphson Iterative Solution 3-4 |
| 3.4 | Sample Results 3-9 |
| 3.4.1 | Assumptions 3-9 |
| 3.4.2 | Uncompensated Images 3-10 |
| 3.4.3 | Adaptive Optics Compensated Images 3-15 |
| 3.5 | Summary 3-17 |
| IV. | Vector Wiener Filter Analysis 4-1 |
| 4.1 | Introduction 4-1 |
| 4.2 | Fourier Domain Filter Derivation 4-1 |
| 4.2.1 | Wiener-Hopf Expression 4-2 |
| 4.2.2 | Object-Detected Image Cross-correlation 4-3 |
| 4.2.3 | Detected Image Autocorrelation 4-5 |
| 4.2.4 | Final Result 4-7 |
| 4.3 | Signal-to-Noise Ratio Interpretation 4-8 |
| 4.4 | Mean Square Error Expressions 4-9 |
| 4.5 | Alternate Filter Expressions 4-11 |
| 4.6 | Comments on Filter Optimality 4-13 |

| | Page |
|-------|---|
| 4.6.1 | Background 4-15 |
| 4.6.2 | Signal-Independent Noise 4-16 |
| 4.6.3 | Semi-Classical Model 4-16 |
| 4.7 | Summary 4-17 |
| V. | Vector Wiener Filter Processing of Binary Star Pairs 5-1 |
| 5.1 | Introduction 5-1 |
| 5.2 | Simulation 5-1 |
| 5.2.1 | Random Object 5-2 |
| 5.2.2 | Measurement Noise 5-3 |
| 5.2.3 | Performance Metrics 5-4 |
| 5.3 | Deterministic OTF 5-6 |
| 5.4 | Random OTF Due to Atmospheric Turbulence 5-18 |
| 5.5 | Summary 5-26 |
| VI. | Vector Wiener Filter Performance and Robustness Study 6-1 |
| 6.1 | Introduction 6-1 |
| 6.2 | Generalized Models 6-1 |
| 6.2.1 | Object 6-2 |
| 6.2.2 | Optical Transfer Function 6-4 |
| 6.3 | Filter Expressions and Mean Square Error Metrics 6-6 |
| 6.4 | Performance Study 6-8 |
| 6.4.1 | Assumptions 6-9 |
| 6.4.2 | Data 6-9 |
| 6.5 | Robustness Study 6-14 |
| 6.5.1 | Assumptions 6-18 |
| 6.5.2 | Data 6-19 |
| 6.6 | Summary 6-19 |

| | Page |
|---|--------|
| VII. Conclusions and Recommendations | 7-1 |
| 7.1 Major Results | 7-1 |
| 7.2 Discussion | 7-2 |
| 7.3 Recommendations for Future Work | 7-2 |
| 7.3.1 Alternate Constraint Functions | 7-3 |
| 7.3.2 Sparse Matrix Tools | 7-3 |
| Appendix A. Derivation of Key Conditional Expectations | A-1 |
| A.1 Equation (4.10) | A-1 |
| A.2 Equation (4.20) | A-3 |
| A.2.1 $n = m$ Term | A-4 |
| A.2.2 $n \neq m$ Term | A-4 |
| Appendix B. Detected Image Probability Density Function | B-1 |
| Bibliography | BIB-1 |
| Vita | VITA-1 |

List of Figures

| Figure | | Page |
|--------|--|------|
| 1.1. | Diagram of a typical AO system as part of a large telescope. | 1-6 |
| 1.2. | Block diagram of the speckle imaging process. | 1-7 |
| 1.3. | Block diagram of DWFS. | 1-8 |
| 2.1. | Turbulent eddies and their effects on an incoming unperturbed plane wave. | 2-6 |
| 2.2. | Geometrical interpretation of ULS estimation. (a) Subspace spanned by the linearly independent column vectors of the full rank matrix H . (b) The object estimate $\hat{\mathbf{o}}$ is the orthogonal projection of the detected image \mathbf{d} onto the subspace shown in (a). The quantity ϵ denotes the error between data and object estimate. | 2-13 |
| 2.3. | Geometrical interpretation of CLS estimation with a rigid linear constraint. The constrained solution $\hat{\mathbf{o}}_c$ is a projection of the unconstrained solution $\hat{\mathbf{o}}$ onto the constraint subspace. | 2-14 |
| 2.4. | Block diagram associated with the error reduction algorithm. | 2-17 |
| 2.5. | Block diagram associated with Ayers-Dainty iterative blind deconvolution. | 2-18 |
| 3.1. | Satellite computer rendering used to test CLS algorithm performance. Negative image shown for clarity. | 3-10 |
| 3.2. | CLS algorithm comparison with manual parameter selection, $r_o = 20\text{cm}$, $m_\nu = +4$, $\sigma_r = 15$ electrons per pixel. (a) Single short exposure image. (b) CLS algorithm estimate, $\gamma = 0.0015$, 4 iterations. (c), (d), (e), and (f) manual parameter selection, $\epsilon = 0.1, 0.01, 0.0001$, and 0.00001 , respectively. | 3-11 |
| 3.3. | CLS algorithm output versus atmospheric turbulence strength, $m_\nu = +2$, $\sigma_r = 15$ electrons per pixel. (a) Detected image, $r_o = 10\text{cm}$. (b) CLS estimate, $r_o = 10\text{cm}$, $\gamma = 0.000012$, 4 iterations. (c) Detected image, $r_o = 20\text{cm}$. (d) CLS estimate, $r_o = 20\text{cm}$, $\gamma = 0.000067$, 4 iterations. | 3-13 |

| Figure | | Page |
|--------|---|------|
| 3.4. | CLS algorithm output versus object brightness, $r_o = 10\text{cm}$, $\sigma_r = 15$ electrons per pixel. (a) Detected image, $m_\nu = +4$. (b) CLS estimate, $m_\nu = +4$, $\gamma = 0.00027$, 4 iterations. (c) Detected image, $m_\nu = +6$. (d) CLS estimate, $m_\nu = +6$, $\gamma = 0.0093$, 5 iterations. (e) Detected image, $m_\nu = +8$. (f) CLS estimate, $m_\nu = +8$, $\gamma = 0.8180$, 7 iterations. | 3-14 |
| 3.5. | CLS algorithm output versus detector read noise strength, $r_o = 10\text{cm}$, $m_\nu = +4$. (a) Detected image, $\sigma_r = 10$ electrons per pixel. (b) CLS estimate, $\sigma_r = 10$ electrons per pixel, $\gamma = 0.00014$, 4 iterations. (c) Detected image, $\sigma_r = 30$ electrons per pixel. (d) CLS estimate, $\sigma_r = 30$ electrons per pixel, $\gamma = 0.00097$, 4 iterations. | 3-15 |
| 3.6. | CLS algorithm output for AO compensated images, $m_\nu = +4$, $\sigma_r = 15$ electrons per pixel. (a) Detected image, $r_o = 10\text{cm}$. (b) CLS estimate, $r_o = 10\text{cm}$, $\gamma = 0.0048$, 4 iterations. (c) Detected image, $r_o = 20\text{cm}$. (d) CLS estimate, $r_o = 20\text{cm}$, $\gamma = 0.0039$, 4 iterations. | 3-16 |
| 4.1. | The vector Wiener filter is optimal with respect to minimum error variance only among the class of linear filters represented by the small circle marked "Linear". Further statements with respect to optimality require assumptions regarding object and data distributions. | 4-14 |
| 5.1. | Sample binary star object realizations ($h_p = 1$, $h_s = 0.5$ pixels), (a) 16 x 16 detector ($w_p = w_s = 0.5$ pixels), (b) 32 x 32 detector ($w_p = w_s = 1$ pixel). | 5-3 |
| 5.2. | Normalized average MSE, $\overline{\varepsilon_V^2}/\overline{\varepsilon_D^2}$ and $\overline{\varepsilon_S^2}/\overline{\varepsilon_D^2}$, versus support constraint dimension W . Pupil size $D_p = 6$ pixels and $\overline{K} = 10000$ photoevents ($\text{SNR}_k = 100$, $\text{SNR}_r = 278$). | 5-9 |
| 5.3. | Normalized mesh plots showing the improved performance characteristics of the vector filter on binary Gaussian function objects. (a) True object realization, object randomness parameter $W = 8$ pixels, (b) mean object, (c) PSF, pupil size $D_p = 6$ pixels, (d) detected image $\overline{K} = 10000$ photoevents ($\text{SNR}_k = 100$, $\text{SNR}_r = 278$), (e) scalar filter estimate, (f) vector filter estimate. | 5-10 |

| Figure | | Page |
|--------|--|------|
| 5.4. | Radially averaged $ \gamma_{O\hat{O}} $ as a function of normalized spatial frequency ρ for the scalar and vector filters. The spatial frequency at unity corresponds to the OTF cutoff frequency for a $D_p = 6$ pixel pupil function. Support constraint dimension $W = 8$ pixels and $\bar{K} = 1000$ photoevents ($\text{SNR}_k = 32, \text{SNR}_r = 28$). | 5-11 |
| 5.5. | Radially averaged MSPE, φ^2 , as a function of normalized spatial frequency ρ for the scalar and vector filters. The spatial frequency at unity corresponds to the OTF cutoff frequency for a $D_p = 6$ pixel pupil function. Support constraint dimension $W = 8$ pixels and $\bar{K} = 1000$ photoevents ($\text{SNR}_k = 32, \text{SNR}_r = 28$). | 5-12 |
| 5.6. | Normalized average MSE, $\overline{\varepsilon_V^2}/\overline{\varepsilon_D^2}$ and $\overline{\varepsilon_S^2}/\overline{\varepsilon_D^2}$, versus pupil size D_p . Support constraint dimension $W = 8$ pixels and $\bar{K} = 1000$ photoevents ($\text{SNR}_k = 32, \text{SNR}_r = 28$). | 5-13 |
| 5.7. | Normalized average MSE, $\overline{\varepsilon_V^2}/\overline{\varepsilon_D^2}$ and $\overline{\varepsilon_S^2}/\overline{\varepsilon_D^2}$, versus \bar{K} . Support constraint dimension $W = 8$ pixels and pupil size $D_p = 6$ pixels. | 5-14 |
| 5.8. | Radially averaged $ \gamma_{O\hat{O}} $ as a function of normalized spatial frequency ρ for the scalar and vector filters. The spatial frequency at unity corresponds to the OTF cutoff frequency for a $D_p = 6$ pixel pupil function and support constraint dimension $W = 8$ pixels. The v and s designators differentiate between vector and scalar filter traces. | 5-15 |
| 5.9. | Radially averaged MSPE, φ^2 , as a function of normalized spatial frequency ρ for the scalar and vector filters. The spatial frequency at unity corresponds to the OTF cutoff frequency for a $D_p = 6$ pixel pupil function and support constraint dimension $W = 8$ pixels. The v and s designators differentiate between vector and scalar filter traces. | 5-16 |
| 5.10. | Normalized mesh plots showing the effect of measurement noise on filter performance. (a) True object realization, object support constraint dimension $W = 8$ pixels, (b) mean object, (c) PSF, pupil size $D_p = 6$ pixels, (d) detected image, $\bar{K} = 500$ photoevents ($\text{SNR}_k = 22, \text{SNR}_r = 14$), (e) scalar filter estimate, (f) vector filter estimate. | 5-17 |
| 5.11. | Mean object used to generate the random OTF data ($W = 10$ pixels). | 5-18 |
| 5.12. | Tilt-removed OTF statistics due to atmospheric turbulence versus radially averaged spatial frequency ρ (von Karman turbulence statistics $L_o/r_o = 100, D/r_o = 1, 2,$ and 4). (a) Mean, (b) variance. The OTF due to diffraction is provided as a reference to the reader in (a). | 5-20 |

| Figure | Page |
|--|------|
| 5.13. Tilt-removed OTF SNR, $\text{SNR}_{\mathcal{H}}(\rho)$, due to atmospheric turbulence versus radially averaged spatial frequency ρ (von Karman turbulence statistics $L_o/r_o = 100$, $D/r_o = 1, 2$, and 4). Note that $\text{SNR}_{\mathcal{H}}(\rho) < 1$ when $\rho > 0.3$ in the $D/r_o = 4$ case. | 5-21 |
| 5.14. Normalized average MSE, $\overline{\varepsilon}_V^2/\overline{\varepsilon}_D^2$ and $\overline{\varepsilon}_S^2/\overline{\varepsilon}_D^2$, versus \overline{K} for the diffraction-limited OTF and $D/r_o = 1, 2$, and 4 . The v and s designators differentiate between vector and scalar Wiener filter traces. | 5-22 |
| 5.15. Normalized MSE, $\varepsilon_V^2(\rho)/\varepsilon_D^2(\rho)$ and $\varepsilon_S^2(\rho)/\varepsilon_D^2(\rho)$, for the diffraction-limited OTF and $D/r_o = 1, 2$, and 4 ($\overline{K} = 10000$ photoevents). The v and s designators differentiate between vector and scalar Wiener filter traces. Filter MSE performance is severely degraded at radial frequencies ρ where $\text{SNR}_{\mathcal{H}}(\rho) < 1$ in the $D/r_o = 4$ case. | 5-23 |
| 5.16. Normalized MSE, $\varepsilon_V^2(\rho)/\varepsilon_D^2(\rho)$ and $\varepsilon_S^2(\rho)/\varepsilon_D^2(\rho)$, for the diffraction-limited OTF and $D/r_o = 1, 2$, and 4 ($\overline{K} = 1000$ photoevents). The v and s designators differentiate between vector and scalar Wiener filter traces. | 5-24 |
| 5.17. Normalized MSE, $\varepsilon_V^2(\rho)/\varepsilon_D^2(\rho)$ and $\varepsilon_S^2(\rho)/\varepsilon_D^2(\rho)$, for the diffraction-limited OTF and $D/r_o = 1, 2$, and 4 ($\overline{K} = 500$ photoevents). The v and s designators differentiate between vector and scalar Wiener filter traces. | 5-24 |
| 5.18. MSPE data $\varphi^2(\rho)$. Turbulence strength traces ($\overline{K} = 10000$ photoevents). The v and s designators differentiate between vector and scalar Wiener filter traces. | 5-25 |
| 5.19. MSPE data $\varphi^2(\rho)$. Light level traces ($D/r_o = 2$). The v and s designators differentiate between vector and scalar Wiener filter traces. | 5-26 |
| 5.20. Normalized mesh plots showing the improved performance characteristics of the vector Wiener filter on photon-limited binary star objects degraded by atmospheric turbulence. (a) True object realization, (b) mean PSF ($D/r_o = 4$), (c) random PSF, (d) detected image ($\overline{K} = 1000$ photoevents), (e) scalar filter estimate, (f) vector filter estimate. | 5-27 |
| 6.1. Mesh plots showing mean object, $\overline{o}(x, y)$, for mean object width parameter, $w_o = 0.5$. Support constraint dimension (a) $W = 10$ pixels, (b) $W = 16$ pixels. | 6-3 |

| Figure | | Page |
|--------|---|------|
| 6.2. | Mesh plot showing the mean OTF, $\overline{\mathcal{H}}(u, v)$, for mean OTF width parameter, $w_{\mathcal{H}} = 0.5$ | 6-5 |
| 6.3. | OTF SNR, $\text{SNR}_{\mathcal{H}}$, versus radial spatial frequency, ρ . The OTF SNR roll-off parameter, $\mu_{\mathcal{H}}$, controls the spatial frequency at which the $\text{SNR}_{\mathcal{H}}$ function falls below unity. | 6-7 |
| 6.4. | Normalized average MSE, $\overline{\varepsilon_V^2}/\overline{\varepsilon_S^2}$, versus support constraint dimension, W , $\rho_o = 0.5$. The OTF is non-random ($\text{SNR}_{\mathcal{H}} = \infty$ for all spatial frequencies (u, v)). | 6-10 |
| 6.5. | Normalized average MSE, (a) $\overline{\varepsilon_V^2}/\overline{\varepsilon_S^2}$, and (b) $\overline{\varepsilon_V^2}/\overline{\varepsilon_D^2}$, versus the object SNR parameter, SNR_o . The object support constraint dimension is fixed at $W = 10$ pixels. The OTF is non-random ($\text{SNR}_{\mathcal{H}} = \infty$ for all spatial frequencies (u, v)). | 6-11 |
| 6.6. | Normalized average MSE, (a) $\overline{\varepsilon_V^2}/\overline{\varepsilon_S^2}$, and (b) $\overline{\varepsilon_V^2}/\overline{\varepsilon_D^2}$, versus the object correlation coefficient, ρ_o . The object support constraint dimension is fixed at $W = 10$ pixels. The OTF is non-random ($\text{SNR}_{\mathcal{H}} = \infty$ for all spatial frequencies (u, v)). | 6-13 |
| 6.7. | (a) Normalized average MSE, $\overline{\varepsilon_V^2}/\overline{\varepsilon_S^2}$, and (b) raw average MSE data, $\overline{\varepsilon_V^2}$, versus the OTF SNR roll-off parameter, $\mu_{\mathcal{H}}$. Object model parameters: $W = 10$ pixels, $\text{SNR}_o = 2$, $\rho_o = 0.5$ | 6-15 |
| 6.8. | (a) Normalized average MSE, $\overline{\varepsilon_V^2}/\overline{\varepsilon_S^2}$, and (b) raw average MSE data, $\overline{\varepsilon_V^2}$, versus the OTF SNR roll-off parameter, $\mu_{\mathcal{H}}$. Object model parameters: $W = 16$ pixels, $\text{SNR}_o = 0.5$, $\rho_o = 0.5$ | 6-16 |
| 6.9. | Normalized average MSE, (a) $\overline{\varepsilon_V^2}/\overline{\varepsilon_S^2}$, and (b) $\overline{\varepsilon_V^2}/\overline{\varepsilon_D^2}$, versus the OTF correlation coefficient, $\rho_{\mathcal{H}}$. Object model parameters: $W = 10$ pixels, $\text{SNR}_o = 2$, $\rho_o = 0.5$ | 6-17 |
| 6.10. | Normalized average MSE, (a) $\overline{\varepsilon_V^2}/\overline{\varepsilon_S^2}$, and (b) $\overline{\varepsilon_V^2}/\overline{\varepsilon_D^2}$, versus percent error in the object SNR parameter, SNR_o , and the object correlation coefficient, ρ_o . OTF model parameters: $w_{\mathcal{H}} = 0.5$, $\rho_{\mathcal{H}} = 0.5$. True parameter values: $\text{SNR}_o = 2$, $\rho_o = 0.5$ | 6-20 |
| 6.11. | Normalized average MSE, (a) $\overline{\varepsilon_V^2}/\overline{\varepsilon_S^2}$, and (b) $\overline{\varepsilon_V^2}/\overline{\varepsilon_D^2}$, versus percent error in the object SNR parameter, SNR_o . OTF model parameters: $w_{\mathcal{H}} = 0.5$, $\rho_{\mathcal{H}} = 0.5$. True parameter values: $\text{SNR}_o = 2$, $\rho_o = 0.5$ | 6-21 |

- 6.12. Normalized average MSE, (a) $\overline{\varepsilon_V^2}/\overline{\varepsilon_S^2}$, and (b) $\overline{\varepsilon_V^2}/\overline{\varepsilon_D^2}$, versus percent error in the object correlation coefficient, ρ_o . OTF model parameters: $w_{\mathcal{H}} = 0.5$, $\rho_{\mathcal{H}} = 0.5$. True parameter values: $\text{SNR}_o = 2$, $\rho_o = 0.5$ 6-22

List of Tables

| Table | | Page |
|-------|--|------|
| 3.1. | Average Photoevents per Integration Time. | 3-10 |
| 5.1. | Measurement Noise Cases | 5-7 |
| 6.1. | Object Model Parameters | 6-4 |
| 6.2. | Optical Transfer Function Model Parameters | 6-7 |

Abstract

Two new linear reconstruction techniques are developed to improve the resolution of images collected by ground-based telescopes imaging through atmospheric turbulence. The classical approach involves the application of constrained least squares (CLS) to the deconvolution from wavefront sensing (DWFS) technique. The new algorithm incorporates blur and noise models to select the appropriate regularization constant automatically. In all cases examined, the Newton-Raphson minimization converged to a solution in less than 10 iterations. The non-iterative Bayesian approach involves the development of a new vector Wiener filter which is optimal with respect to mean square error (MSE) for a non-stationary object class degraded by atmospheric turbulence and measurement noise. This research involves the first extension of the Wiener filter to account properly for shot noise and an unknown, random optical transfer function (OTF). The vector Wiener filter provides superior reconstructions when compared to the traditional scalar Wiener filter for a non-stationary object class. In addition, the new filter can provide a superresolution capability when the object's Fourier domain statistics are known for spatial frequencies beyond the OTF cutoff. A generalized performance and robustness study of the vector Wiener filter showed that MSE performance is fundamentally limited by object signal-to-noise ratio (SNR) and correlation between object pixels.

Linear Reconstruction of Non-Stationary Image Ensembles Incorporating Blur and Noise Models

I. Introduction

1.1 Problem Statement

Since before the time of Galileo and Newton, man has used optical devices to form images of distant objects. The term *image* refers to the two dimensional picture associated with the light or irradiance collected by an imaging system such as the eye, a camera, or telescope. The term *object* denotes the light or radiant exitance that caused the image to be formed. A perfect optical system will produce an image that is identical to the object within the limits of diffraction. In reality, a loss of resolution may occur due to blur associated with distortions in the optical device or randomness in the imaging medium. An image may also be distorted by noise due to low light level or limitations in the recording device. The imaging scenario describes the degradations affecting an optical system in a given application. For example, images collected using a ground-based astronomical telescope are degraded by the turbulent atmosphere and film-grain or electronic detector noise.

With the widespread availability of computers, the concept of a digital image has become important. A digital image is an array of real or complex numbers which represents a sampled version of the two dimensional continuous image described above. The elements of an image array are known as *pixels*. In this mathematical form, a distorted image can be manipulated by a computer using a variety of techniques [41]. Image reconstruction refers to digital image processing techniques that attempt to recover an accurate object estimate based on a priori knowledge of the imaging scenario. Statistical estimation theory plays an important role in many modern reconstruction algorithms. This theory can be divided into two main approaches: classical and Bayesian [43]. In a classical approach, the

parameter of interest is assumed to be a deterministic but unknown constant. In contrast, the unknown parameter is assumed to be a random variable in a Bayesian approach. Here, the random parameter is described by a known prior probability density function (PDF), and the goal is to estimate a particular realization. Common optimization criteria for Bayesian estimators include minimizing mean square error (MSE) between parameter and estimate as well as maximizing a posteriori probability. Both classical and Bayesian estimators can sometimes be difficult to implement, requiring multidimensional integration or intensive iterative optimization. In many cases, constraining the estimator to be linear allows for substantial simplification and ease of analysis. Two well-known linear methods are least squares and linear minimum MSE estimation, also known as the Wiener filter [43].

Least squares is a classical estimation method first used by Gauss to study planetary motions in 1795 [43]. The goal is to find the object estimate that minimizes the squared error between the given distorted image and some deterministic image model. No probabilistic assumptions are made about the data [43]. Actual performance is dependent on two factors: the noise properties of the distorted image and the image model accuracy. For example, low light images degraded by atmospheric turbulence are not good candidates for least squares processing when the image model or blur cannot be estimated accurately. In this case, a wavefront sensor (WFS) can be used to provide an accurate estimate of the pupil plane phase aberration. The phase measurement can then be used to estimate the blur function. This technique is known as deconvolution from wavefront sensing (DWFS) [10, 17, 66] and is based on the least squares paradigm. A statistical noise model is not incorporated in the traditional DWFS estimator to suppress noise effects. The standard solution is to use a regularization constant in the estimator denominator [20]. The regularization constant is adjusted by the user based on perceived image quality.

The Wiener filter minimizes MSE between the true object and object estimate [43]. It was first derived for two dimensional images by Helstrom [29] and Slepian [80]. Their research followed the paradigm established in the seminal work by Norbert Wiener with stationary time series [94]. Here, the object and noise are assumed to be wide sense stationary random processes. In this dissertation, the term stationary will refer to a wide sense stationary

random process. A two dimensional stationary random process has a constant mean and autocorrelation that is only dependent on the distance between pixels, not the individual pixel locations [62]. In the Fourier domain, a stationary random process has uncorrelated spatial frequency components [62]. Since the Fourier components are uncorrelated, the Wiener filter reduces to a simplified *scalar* form which weights each distorted image spatial frequency independently to produce the estimated object spectrum. However, many practical image ensembles are non-stationary [39] and have correlated Fourier components [62]. For example, consider the image of a satellite in low-earth orbit as collected by a ground-based telescope. Multiple images of the satellite are collected from different perspectives as it passes over the observation site. Thus, the mean object is a blurred version of the true satellite against the black background of space. Clearly, this object random process is not constant mean. In addition, it is not uncommon to use a support constraint when processing these images which generates a non-stationary image domain covariance [14]. The scalar Wiener filter is not capable of incorporating complete object and noise Fourier domain correlations. Thus, it is sub-optimal with respect to MSE for non-stationary image ensembles.

As noted above, nonlinear classical and Bayesian estimation techniques often involve multidimensional integration and intensive iterative optimization. Linear techniques can offer advantages related to computational savings and analysis, often at the expense of performance. Thus, the desire to enhance the performance of linear reconstruction techniques is the key motivation for solving the problem addressed in this dissertation:

Develop enhanced linear reconstruction filters for non-stationary image ensembles incorporating a priori blur and noise models. Investigate performance limitations associated with imaging through atmospheric turbulence.

Both a classical and a Bayesian approach are addressed in this dissertation. The classical approach involves the application of constrained least squares (CLS) to DWFS. CLS incorporates a priori knowledge of the imaging scenario to constrain the set of possible object estimates. The new algorithm incorporates blur and noise models to select the appropriate regularization constant *automatically* [11]. No ad hoc regularization adjustment is required. CLS processing of DWFS data is demonstrated using simulated satellite objects degraded by atmospheric turbulence and measurement noise. Measurement noise is defined as the com-

bined contribution of shot noise and signal-independent detector noise [14]. The Bayesian approach involves the development of a new *vector* Wiener filter which is optimal with respect to MSE for a non-stationary object class degraded by atmospheric turbulence and measurement noise. Here, the term vector alludes to the dependence of each estimated Fourier component on other distorted image Fourier components. It should be noted that the general Wiener filter is the steady-state constant-gain version of the Kalman filter for shift-invariant image models and stationary noise models [43]. Thus, it is valid to refer to the new vector Wiener filter as a Kalman filter when the noise is non-stationary. However, previous image processing research has included reference to a Wiener filter in this application. Pratt proposed a generalized vector Wiener filter that is optimal with respect to MSE for non-stationary object ensembles in signal-independent noise [63]. This theory was extended to images degraded by both known blur and signal-independent noise [65, 76]. The research presented in this dissertation will use the term Wiener filter when referring to this image reconstruction application. This work involves the first extension of this theory to account properly for shot noise [14]. Vector Wiener filter performance is also investigated when blur statistics are substituted for exact knowledge of the blur function [12, 13].

This dissertation is organized into seven chapters. Chapter I presents the justification for pursuing the suggested study, details the problem to be solved, and outlines significant results. Chapter II contains background material associated with atmospheric turbulence and image reconstruction. Chapter III outlines a new CLS processing technique for DWFS. Chapter IV contains the complete derivation of a new vector Wiener filter which incorporates model-based statistical knowledge of object, blur, and noise. Chapter V illustrates vector Wiener filter performance on astronomical images degraded by atmospheric turbulence and noise. Chapter VI presents vector Wiener filter performance and robustness data for generalized object and blur models. Conclusions and recommendations for further research are found in Chapter VII. Mathematical details not included in the main text are compiled in the Appendices.

1.2 Justification for Conducting the Proposed Research

Today, the United States faces an ever-growing number of potential adversaries with satellite launch capability. Clear, resolvable images of space objects from ground-based telescopes are an absolute requirement to determine an opponent's intentions in space. Thus, the Air Force has a requirement for high resolution imagery of earth-orbiting objects as part of its space surveillance mission. In general, two broad classes of techniques are used to increase Fourier domain signal-to-noise (SNR) ratio in astronomical images: (1) pre-detection processing via adaptive optics (AO) [27] and (2) post-detection processing such as speckle imaging [44, 47, 50].

AO compensates for atmospheric turbulence-induced wavefront aberrations in real time before the light is detected at the image plane. The important components of an AO system are the deformable mirror (DM), wavefront sensor (WFS), and actuator control computer [74]. Voltages applied to the DM actuators allow its figure to be changed in real time. The WFS senses the aberrations in the incoming wave by measuring gradients in small subapertures of the telescope pupil [91]. This information is then sent to the actuator control computer which adjusts the DM to apply an estimate of the conjugate of the wavefront aberration. The correction imposed by the DM cancels out the aberration, leading to a narrower blur or point spread function (PSF) and an improved image. This process must occur at speeds on the order of the rate of change of the wavefront aberration to be effective [91]. Typically, these speeds range from approximately tens of Hertz to a few hundred Hertz [91]. Figure 1.1 illustrates the installation of an AO system on a ground-based telescope.

The first work to address the post-detection processing of images degraded by atmospheric turbulence was speckle imaging [44, 47, 50]. The term speckle refers to the data, which consist of a set of short exposure, speckled images. In this context, short exposure refers to a sufficiently short integration time to freeze an individual realization of the atmospheric turbulence-induced aberration in the image measurement. In this technique, the object is usually estimated by first estimating the modulus and phase of its Fourier transform. Labeyrie showed that the squared modulus of the object Fourier transform could be estimated from a large set of short exposure images [47]. The method requires an ensemble

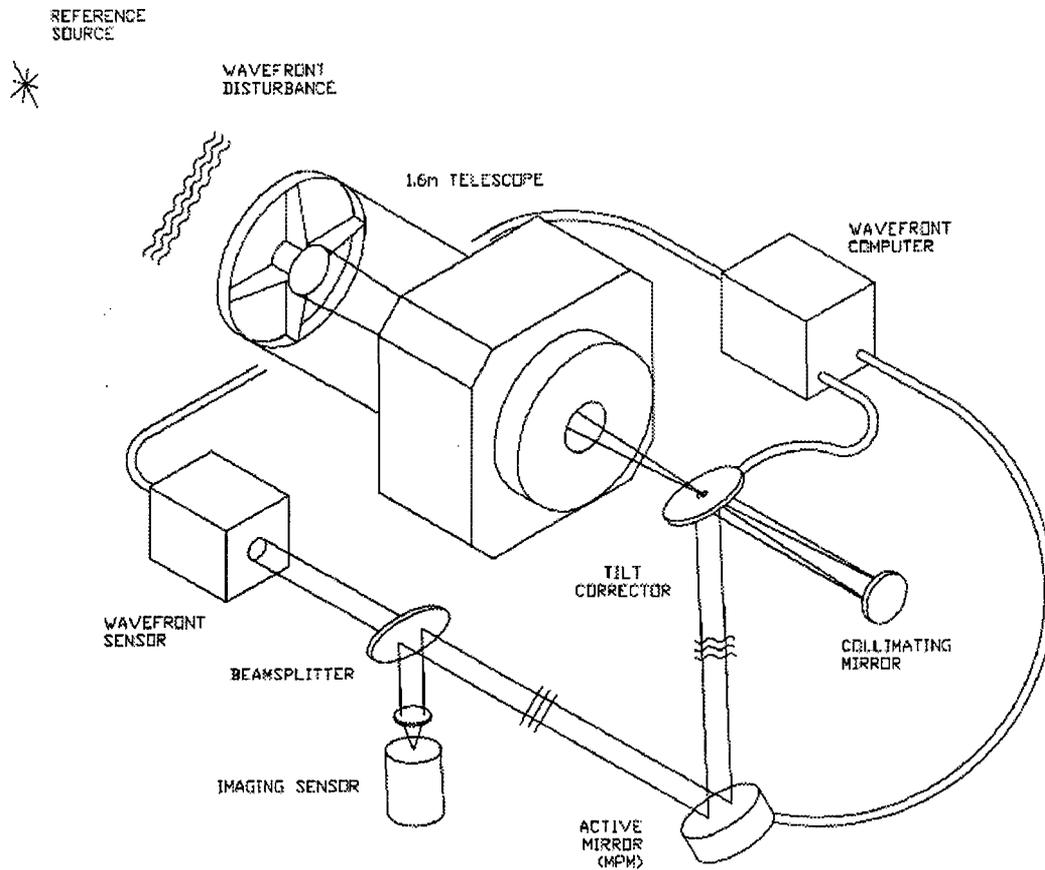


Figure 1.1 Diagram of a typical AO system as part of a large telescope.

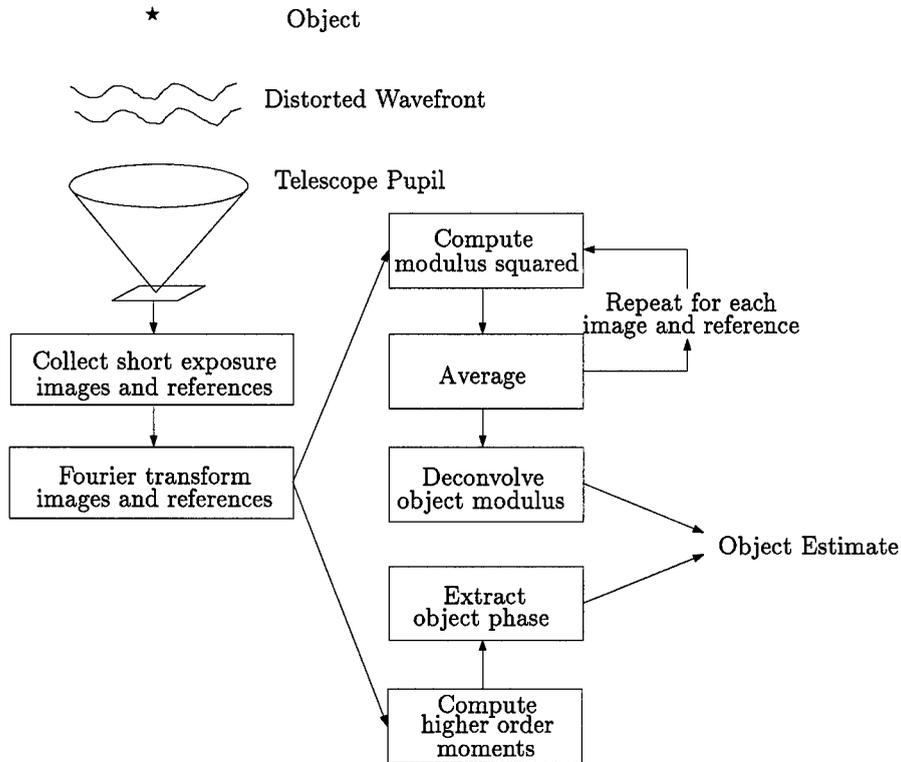


Figure 1.2 Block diagram of the speckle imaging process.

of reference point source images along with the data set. The reference images are used to estimate the Fourier transform of the atmospheric-optical system PSF or optical transfer function (OTF). The squared modulus of both data and OTF estimates are averaged to reduce noise. Then, the average squared modulus of the OTF estimates is used in a deconvolution procedure to estimate the squared modulus of the object. This estimation scheme is possible because the squared modulus of the OTF is non-zero out to the diffraction-limited cutoff of the optical system [74]. While Fourier modulus information can be extracted using the above technique, phase information is usually required to form a usable image [74]. Two methods are commonly used to extract the Fourier phase from the data ensemble: the Knox-Thompson [44] and bispectrum [50] methods. Both methods are based on the fact that certain higher order moments of the complex Fourier transform of speckled images contain encoded information about the object phase. Figure 1.2 gives a block diagram of the speckle imaging process.

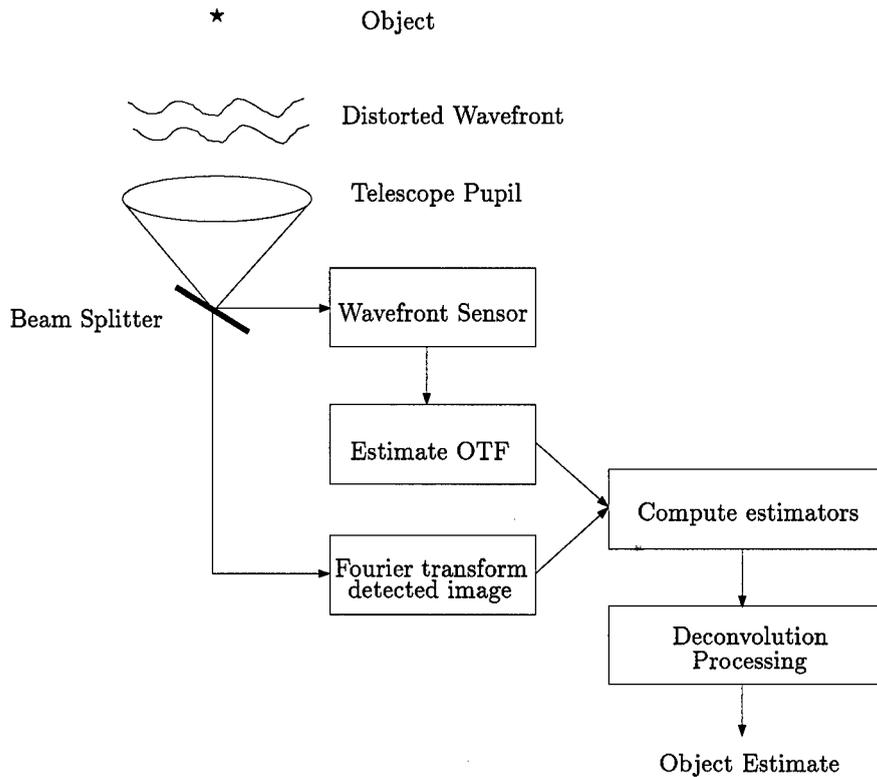


Figure 1.3 Block diagram of DWFS.

A third class of imaging techniques involves the combination of both AO and post-detection processing and is known as hybrid imaging [74]. A notable hybrid technique is DWFS [10, 17, 66]. In DWFS, a WFS is used to measure the pupil plane phase aberration associated with each short exposure image. The phase measurement is used to estimate the OTF. The short exposure image and OTF estimate can then be used to estimate the object via a deconvolution filter. The DWFS technique was first proposed by Fontanella [10], extended by Fried [17], and further developed by Primot *et al.* [66]. Primot *et al.* also conducted the first performance analysis of DWFS [66]. A variant of their estimator was later validated on astronomical data [20]. Welsh and Von Niederhausern further investigated DWFS performance by incorporating an optimal wavefront phase estimate [93]. Roggemann *et al.* [75] showed that the Primot estimator was biased and suggested a related unbiased estimator. Roggemann and Welsh also derived an SNR expression for DWFS [73] and conducted further comparison with speckle imaging and traditional linear deconvolution [92]. Figure 1.3 gives a block diagram of the DWFS process.

Speckle imaging techniques generally incorporate a deconvolution filter to estimate the object modulus [74]. AO compensated images also benefit from linear reconstruction to deconvolve blurring due to the attenuation of high spatial frequencies in the compensated images [70,72]. Typical linear reconstruction methods such as the inverse and pseudo-Wiener filters [21] require an explicit estimate of the OTF. A priori knowledge of the object class and noise is typically not used to deconvolve compensated images. CLS processing of DWFS data and the vector Wiener filter offer the potential to use statistical model-based information about the OTF and noise to improve images from Air Force ground-based surveillance sites. As of this writing, no sources have been found in the literature which document the use of CLS estimation to process DWFS data. However, Primot *et al.* note the potential of such a scheme in their key paper [66]. Similarly, no sources are available which document the use of a Fourier domain linear minimum MSE estimator to process non-stationary image ensembles degraded by random blur and shot noise. Several researchers have addressed the problem of image reconstruction in the presence of random blur but only in the image domain [4,25,26,87]. The potential benefits of statistical model-based PSF information has been mentioned by prominent researchers in the area of blind deconvolution. Schulz [77] has recommended the use of PSF statistical models in his algorithms when this information is available. However, his work and the work of other researchers in this field have not considered the degrading PSF as a random quantity [31,33,48,77,81]. Thus, the research outlined in this dissertation makes a unique contribution to a critical Air Force mission and the field of image reconstruction.

1.3 Approach

The problem statement is addressed in two ways. This two-prong approach is consistent with the natural classical-Bayesian division prevalent in statistical estimation theory [43]. First, the application of CLS to DWFS is investigated. CLS represents a classical estimation approach since the object is assumed to be a deterministic, yet unknown quantity. Only the constraint incorporates model-based statistical information. Second, a new linear Bayesian estimator, referred to as the vector Wiener filter, is derived. In this case, model-based statistical knowledge of object, blur, and noise is assumed.

The traditional DWFS deconvolution filter given by Primot *et al.* [66] provides unacceptable noise amplification when processing low light image ensembles. The standard solution is to add a parametric regularization constant [20] or SNR term [75] to the filter denominator. These approaches are analogous to CLS [37] and the parametric Wiener filter [21], respectively. The regularization is adjusted by the user based on the perceived image quality. Thus, the resultant object estimate does not satisfy any mathematical optimality criterion. In this dissertation, a modified CLS estimation scheme is developed which provides optimal processing of noisy DWFS data [11]. Here, optimal refers to the object estimate which minimizes a CLS objective function incorporating DWFS data. Unlike previous CLS algorithms [37,71], this approach incorporates ensemble average data directly to reduce noise effects. The solution uses the Lagrange multiplier technique [21,37]. A closed form solution for the object estimate is obtained which is analogous to the traditional DWFS deconvolution filter [20,75] with the Lagrange multiplier serving as a regularization constant. An iterative approach based on Newton-Raphson minimization [40,71] is used to find the appropriate regularization constant. The iteration incorporates the statistics of both the OTF and noise.

CLS processing of noisy DWFS data relies on WFS hardware and iterative processing to deconvolve turbulence effects. A non-iterative Bayesian approach to the reconstruction of astronomical images is also developed, known as the vector Wiener filter. As noted previously, a scalar filter weights each Fourier component of the distorted data independently. In contrast, a vector filter incorporates many Fourier components of the distorted image to estimate a given Fourier component of the object. The appropriate weighting of each component is determined by the object, OTF, and noise correlations. First, a vector Wiener filter is derived that properly accounts for a random OTF and shot noise effects [14]. This new linear filter has the advantage of incorporating object, OTF, and shot noise correlations between *different* spatial frequency components. When a scalar Wiener filter is applied to a non-stationary image ensemble, this correlation information is not used. The result is a sub-optimal solution with respect to MSE when compared to the vector Wiener filter. Next, vector Wiener filter performance is investigated for both a fixed and random OTF [12–14].

Finally, filter performance and robustness are examined for generalized object and blur models.

This research effort relies on theoretical analysis and Monte Carlo simulation. The Monte Carlo simulation for the atmospheric turbulence-degraded images is based on a Fourier-series based phase screen generator developed by Dr. Byron Welsh [89]. This new phase screen generator properly models the spatial and temporal correlations between wavefront phase screens based on von Karman statistics [89]. However, this study is only concerned with the spatial correlation of the OTF. Thus, the temporal capability of the phase screen generator is not used. Performance comparison between the scalar and vector Wiener filters is an important part of this study. Here, estimator performance is based on visual image comparison, MSE, and mean square phase error (MSPE). MSE at a given image pixel is defined as the expected value of the squared difference between the true object and the object estimate at that pixel. Similarly, MSPE for a given Fourier component is the expected value of the squared difference between the true object phase and the estimated object phase for that Fourier component.

1.4 Scope and Assumptions

Two new applications of model-based statistical knowledge to linear filter theory are presented in this dissertation. The primary study variables are the object class, light level, detector read noise variance, and turbulence strength. The emphasis here is on analysis and simulation.

In the discussion presented in Section 1.3, the object irradiance distribution was assumed to be a random process with known spatial frequency statistics. The concept of a random object in image reconstruction is not new [29, 63, 80] and is critical to a Bayesian development. In Chapters IV and V, perfect a priori knowledge of the object class statistics is assumed. For example, one common class of astronomical objects is the binary star pair. A priori knowledge could include the number of components (two), ratio between primary and secondary component irradiance, and object support. In this case, exact knowledge of the true object irradiance distribution is unavailable since the filter has no knowledge of the

component separation or orientation. In general, the statistical object model can be viewed as a constraint on the filter output [43]. Clearly, the object statistics can never be known perfectly in a real application. Thus, filter robustness is studied in Chapter VI.

1.5 *Significant Contributions and Results*

This section highlights overall contributions and results associated with this dissertation research. Here, the logical questions “What is new?” and “What is important?” are answered. The respective chapters are listed to aid the reader in finding topics of interest.

- Chapter III outlines a *new* application of CLS to the DWFS processing of low light images. The technique is practical and computationally inexpensive. In all cases examined, the Newton-Raphson iteration converged to a solution in less than 10 iterations.
- Chapter IV presents the derivation of a *new* vector Wiener filter incorporating the semi-classical model of photoelectric light detection. The filter uses complete OTF and shot noise statistical models.
- Chapter V presents the *first* application of second order OTF statistics between different spatial frequencies in a Wiener filter. These OTF statistics are associated with imaging through atmospheric turbulence.
- Chapter VI contains the *first* performance study to establish quantitative limits on the vector Wiener filter associated with object and OTF statistical models. The object spatial SNR and correlation coefficient provide a fundamental limit on vector Wiener filter MSE performance. For some object classes, the OTF SNR also limits MSE performance.
- Chapter VI contains the *first* robustness study of the vector Wiener filter with respect to object model error. Error in the object spatial SNR produces a substantial increase in MSE.

1.6 *Summary*

In this chapter, the dissertation problem has been presented, which is to develop linear filters for non-stationary image ensembles incorporating blur and noise statistical models. This research is justified based on the Air Force requirement for clear, resolvable images of space-based objects. Linear reconstruction provides an important function by deconvolving images processed via more sophisticated techniques and hardware. The approach outlined in this dissertation is to investigate both CLS processing of DWFS data and a new vector Wiener filter. Both techniques incorporate blur and noise statistics. In addition, the vector Wiener filter relies on model-based statistical knowledge about the object class. Before presenting these techniques in Chapters III and IV, the next chapter provides important background regarding atmospheric turbulence and key image reconstruction methods.

II. Background

2.1 Introduction

The primary objective of this chapter is to review background material related to general image reconstruction concepts and, more specifically, the reconstruction of astronomical images. The image degradation model associated with a noisy, turbulence-degraded image is introduced as well as information about atmospheric turbulence. Unless otherwise noted, this image model will be used throughout the dissertation. The rest of the chapter will briefly review some important classical and Bayesian techniques in image reconstruction. In each case, the strengths and weaknesses of past research will be highlighted.

The rest of this chapter is organized as follows. Section 2.2 presents the image degradation model. Atmospheric turbulence theory is reviewed in Section 2.3 with emphasis on expressions for the optical transfer function (OTF) statistics. Classical estimation schemes are addressed in Section 2.4 to include least squares, maximum likelihood (ML), and Gerchberg-Saxton iteration. Section 2.5 introduces Bayesian estimation concepts to include the Wiener filter, Kalman filter, and maximum a posteriori (MAP) estimation. Section 2.6 provides a brief summary of the chapter and relates this past research to the work outlined in this dissertation.

2.2 Image Degradation Model

The standard model for a noise-free linear shift-invariant imaging system can be written as [22]

$$i(x, y) = h(x, y) * o(x, y), \quad (2.1)$$

where $i(x, y)$ is the noiseless image, $h(x, y)$ is the impulse response or point spread function (PSF) associated with the blur, $o(x, y)$ is the object, (x, y) is a discrete point in the image domain, and $*$ denotes convolution. Taking the Fourier transform of both sides of Eq. (2.1) yields

$$I(u, v) = \mathcal{H}(u, v)O(u, v), \quad (2.2)$$

where a capital letter denotes the discrete spatial Fourier transform of an associated lower case quantity and $\mathcal{H}(u, v)$ is the Fourier transform of $h(x, y)$ also known as the OTF. The point (u, v) denotes a discrete spatial frequency.

The OTF for a diffraction-limited full aperture is a deterministic, tapered low-pass filter. Random aberrations in the optical system pupil due to atmospheric turbulence further degrade performance, resulting in a more attenuated OTF, especially at high spatial frequencies [22]. To complicate matters, the noiseless image $i(x, y)$ is typically not available. An image model must account for additional degradation due to the detection process. Photon-matter interactions in light detectors are random and require a statistical description. The semi-classical model is based on three assumptions about photon statistics [74]:

1. The probability of occurrence of a single photoevent in a small area dA during a time interval dt is

$$P(1, dt, dA) = \eta dt dA i(x, y, t), \quad (2.3)$$

where $i(x, y, t) = h(x, y, t) * o(x, y, t)$, dA is small compared to the coherence area of the light, dt is short compared to the coherence time of the light, and η is the quantum efficiency of the detector.

2. The probability of more than one photoevent occurring in the area dA during time interval dt is very small compared to the probability associated with either one or zero photoevents.
3. The number of photoevents K occurring in non-overlapping space or time intervals is statistically independent.

Based on these assumptions, the random variable K is governed by Poisson statistics [23, 74].

Poisson random process sample functions consist of sets of Dirac delta functions [62]. Thus, a photon-limited image is defined as [74]

$$d(x, y) = \sum_{n=1}^K \delta(x - x_n, y - y_n), \quad (2.4)$$

where (x_n, y_n) is the location of the n^{th} photoevent in the image plane and K is the total number of photoevents making up the image. Randomness is associated with the number and location of the photoevents. The randomness considered by the semi-classical model is referred to as photon or shot noise and is a form of signal-dependent noise. Signal-dependent refers to the situation in which the strength of the noise depends on the number and distribution of photoevents [74]. Photon noise typically imposes more severe limitations than diffraction, especially at low light levels. The Dirac delta functions of Eq. (2.4) are discontinuous in the image domain, which presents difficulties when conducting statistical analysis. Thus, we are motivated to analyze imaging performance in the Fourier domain via the Fourier transform of Eq. (2.4) which can be written as

$$D(u, v) = \sum_{n=1}^K \exp \{-j2\pi(ux_n + vy_n)\}. \quad (2.5)$$

Signal-independent noise is also present in many detectors used for image collection. For example, charge-coupled device (CCD) detector output is degraded by signal-independent additive noise known as read noise [74]. In this dissertation, signal-independent noise will be represented by the random variable n_p having the following statistical properties:

1. n_p is zero mean.
2. n_p is spatially uncorrelated with uniform variance σ_r^2 .
3. n_p is statistically independent of K and (x_n, y_n) .

An image model which properly accounts for both signal-dependent and signal-independent noise can now be written as [74]

$$d(x, y) = \sum_{n=1}^K \delta(x - x_n, y - y_n) + \sum_{p=1}^P n_p \delta(x - x_p, y - y_p), \quad (2.6)$$

where (x_p, y_p) is the location of the p^{th} image pixel, P is the total number of pixels in the detector array, and $d(x, y)$ now represents a detected image as collected by a CCD detector.

The corresponding Fourier domain expression can be written as

$$D(u, v) = \sum_{n=1}^K \exp \{-j2\pi(ux_n + vy_n)\} + \sum_{p=1}^P n_p \exp \{-j2\pi(ux_p + vy_p)\}. \quad (2.7)$$

Equation (2.7) is a key expression used to model noise effects in this dissertation.

To investigate both constrained least squares (CLS) processing of deconvolution from wavefront sensing (DWFS) data and the vector Wiener filter, a vector-matrix expression is needed for Eq. (2.6). The new expression includes an additive noise vector \mathbf{n} which incorporates both shot and read noise effects based on writing the total noise as the difference between the detected image and the image degraded by the PSF only, denoted $n(x, y) = d(x, y) - i(x, y)$ [14, 37]. Thus, Eq. (2.6) can now be written as [14]

$$\mathbf{d} = H\mathbf{o} + \mathbf{n}, \quad (2.8)$$

where \mathbf{d} and \mathbf{o} are P -length vector versions of the functions $d(x, y)$ and $o(x, y)$, respectively. The matrix H is a $P \times P$ block-circulant matrix representing the shift-invariant PSF. H and \mathbf{o} are properly ordered to perform the discrete convolution of Eq. (2.1) [21, 37]. Equation (2.7) can also be written using this vector-matrix notation such that

$$\mathbf{D} = \mathcal{H} \odot \mathbf{O} + \mathbf{N}, \quad (2.9)$$

where \mathbf{D} , \mathbf{O} , and \mathbf{N} are P -length vector versions of the Fourier domain functions $D(u, v)$, $O(u, v)$, and $N(u, v)$, respectively. The notation \odot represents an entrywise or Hadamard product [35] and \mathcal{H} is a P -length vector containing the OTF. Equations (2.8) and (2.9) are key expressions in the development of CLS processing of DWFS data and the vector Wiener filter in Chapters III and IV, respectively.

The next section reviews atmospheric turbulence theory. Expressions for the mean and spatial correlation of the turbulence-degraded OTF are presented. These OTF statistics will play an important role in both CLS processing of DWFS data and the vector Wiener filter.

2.3 Atmospheric Turbulence

An undergraduate Physics textbook provides the necessary tools to predict the resolution of an imaging system. Theoretically, the minimum resolvable angle seen by a telescope is limited by diffraction and can be expressed as

$$\theta = \frac{1.22\bar{\lambda}}{D}, \quad (2.10)$$

where θ is the minimum resolvable angle in degrees, $\bar{\lambda}$ is the mean wavelength of the incident light, and D is the diameter of the aperture. The overline notation denotes the expectation operator applied to the designated quantity. As noted in the previous section, ground-based imaging systems rarely achieve diffraction-limited performance. Instead, the atmosphere imposes a fundamental limit on spatial resolution. Atmospheric turbulence affects imaging systems by causing both spatial and temporal random fluctuations in the index of refraction of the atmosphere. These index of refraction fluctuations impose random phase aberrations on the incoming light [23, 69]. The primary consequence of these random phase aberrations is a general broadening of the PSF which manifests itself as blurring and lowered resolution when compared to the system predicted by Eq. (2.10).

Atmospheric turbulence is caused by turbulent air motion. The source of this air motion is the heating and cooling of the Earth by the sun. Large air masses gain heat directly from the sun during the day. At night, heat is also coupled to these air masses as the Earth cools. As a result, large-scale temperature variations are produced. These temperature variations lead to pressure differences which result in large scale air motion. Initial large scale air motions break down into smaller and smaller scale motions until the atmosphere is distributed into randomly sized pockets of air, each with its own temperature, as shown in Fig. 2.1. These pockets of air are called turbulent eddies [69]. Since the index of refraction of air is dependent on temperature, the atmosphere has a non-uniform index of refraction.

2.3.1 Turbulence Statistics. Atmospheric turbulence creates a medium which has a non-uniform or random index of refraction associated with the distribution of the turbulent

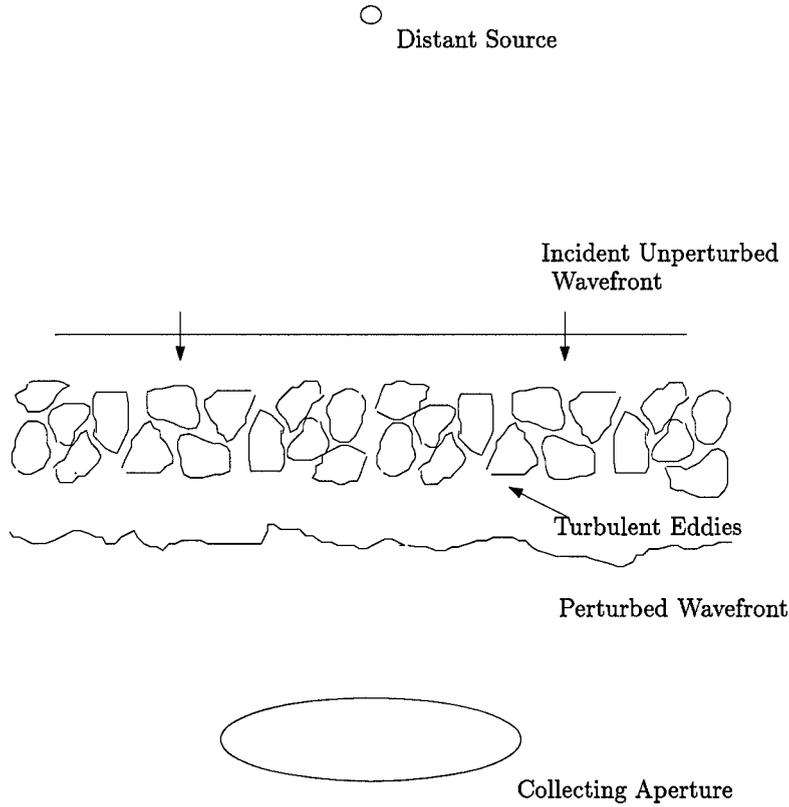


Figure 2.1 Turbulent eddies and their effects on an incoming unperturbed plane wave.

eddies. Thus, statistical models are required to understand turbulence effects fully. Turbulence modeling has received much attention in the literature. However, most published research flows from the key results by Kolmogorov [45], Tatarskii [83], and Fried [15,16]. Kolmogorov provided a statistical model related to spatial structure in turbulent air flows [45]. Tatarskii used Kolmogorov's results to model wave propagations through random index of refraction distributions [83]. Fried applied and extended Tatarskii's work to optical propagation problems [15,16].

Kolmogorov theory gives a mathematical description for index of refraction fluctuations [45]. The index of refraction spatial power spectral density (PSD) provides a frequency domain statistical model for the number and size of the turbulent eddies and can be written as [74]

$$\Phi_n^K(\kappa) = 0.033C_n^2(z)\kappa^{-\frac{11}{3}}, \quad \frac{2\pi}{L_o} \leq \kappa \leq \frac{2\pi}{l_o}, \quad (2.11)$$

where the superscript K denotes the Kolmogorov spectrum, κ is the scalar wavenumber, C_n^2 is a measure of turbulence strength called the structure constant, and the subscript n denotes the atmospheric index of refraction [23]. The constants L_o and l_o denote the outer and inner scale, respectively. These quantities represent the characteristic dimension of the largest and smallest turbulent eddies [74]. In many practical systems, the turbulence strength is a function of the distance from the imaging aperture, denoted by z . Equation (2.11) is not a useful model for the index of refraction PSD when $\kappa \rightarrow 0$ because of the non-integrable singularity at $\kappa = 0$. An alternate form known as the von Karman spectrum is finite for all $\kappa \geq 0$ and can be expressed as [74]

$$\Phi_n^V(\kappa) = \frac{0.033C_n^2(z)}{(\kappa^2 + \kappa_o^2)^{11/6}}, \quad (2.12)$$

where $\kappa_o = 2\pi/L_o$ and the superscript V represents the von Karman spectrum. The von Karman statistical model will be used in theoretical development and simulation throughout this dissertation.

Equation (2.12) is important in deriving statistical models for the turbulence-induced perturbations on a propagating wavefront. Two limiting cases are commonly studied: near field and far field. In the near field case, only the perturbations affecting the wavefront phase are considered. Under the far field assumption, both amplitude and phase effects are modeled. Here, it is assumed that the wavefront amplitude perturbations are negligible compared to the phase perturbations. This assumption is commonly used in a standard geometrical optics model [74]. Thus, only the near field case is considered in this dissertation. Using a layered turbulence model and assuming the index of refraction fluctuations are a Gaussian random process [15, 23, 74], it is possible to derive an expression for the spatial correlation function of the phase perturbations. Consider an incident wavefront and let $\phi(x, y)$ indicate the wavefront phase perturbations in the optical system pupil. Then the spatial correlation of $\phi(x, y)$, denoted $R_{\phi\phi}^V(\Delta x, \Delta y)$, is [74]

$$\begin{aligned}
R_{\phi\phi}(x, y; x', y') &= E[\phi(x, y)\phi(x', y')] \\
R_{\phi\phi}^V(\Delta x, \Delta y) &= \frac{3.089(2\pi)^{-5/6}}{2^{5/6}\Gamma[(11/6)]} \left[\left(\frac{L_o\sqrt{(\Delta x)^2 + (\Delta y)^2}}{(r_o)^2} \right) \right]^{5/6} \\
&\quad \times K_{5/6} \left[\left(\frac{2\pi\sqrt{(\Delta x)^2 + (\Delta y)^2}}{L_o} \right) \right],
\end{aligned} \tag{2.13}$$

where $E[\bullet]$ denotes the expectation operator, $\Delta x = x - x'$, $\Delta y = y - y'$, (x, y) and (x', y') are discrete points in the pupil plane, $\Gamma[(\bullet)]$ is the Gamma function, $K_{5/6}[(\bullet)]$ is the Bessel function of the second kind of order $5/6$, and r_o is the Fried's parameter defined as [74]

$$r_o = 0.185 \left[\frac{\bar{\lambda}^2}{\int_0^\infty C_n^2(z) dz} \right]^{3/5}. \tag{2.14}$$

The Fried parameter can be interpreted as the seeing cell or aperture size beyond which further increases in optical system diameter result in no further increase in resolution [74]. The spatial correlation function given in Eq. (2.13) is the key statistical quantity used to model the effect of atmospheric turbulence on imaging system performance.

While knowledge of the phase correlation function greatly enhances our understanding of turbulence effects, a related quantity known as the phase structure function is of interest in many applications. The phase structure function, denoted $D_{\phi\phi}(\Delta x, \Delta y)$, is defined as [23]

$$D_{\phi\phi}(\Delta x, \Delta y) = 2R_{\phi\phi}(0, 0) - 2R_{\phi\phi}(\Delta x, \Delta y). \tag{2.15}$$

Substituting the von Karman phase correlation function of Eq. (2.13) into Eq. (2.15) yields [74]

$$\begin{aligned}
D_{\phi\phi}^V(\Delta x, \Delta y) = & 3.089(2\pi)^{-5/3} \frac{6}{5} \left(\frac{L_o}{r_o}\right)^{5/3} \left[\left[1 - \frac{\Gamma[(1/6)]}{\pi^{1/6}} \left(\frac{\sqrt{(\Delta x)^2 + (\Delta y)^2}}{L_o}\right)^{5/6} \right. \right. \\
& \left. \left. \times K_{5/6} \left[\left(\frac{2\pi\sqrt{(\Delta x)^2 + (\Delta y)^2}}{L_o}\right) \right] \right] \right]. \tag{2.16}
\end{aligned}$$

As shown below, Eq. (2.16) is an important quantity in the derivation of first and second order OTF statistics.

2.3.2 Optical Transfer Function Statistics. In Section 2.2, the semi-classical model for light detection was introduced, as shown in Eq. (2.4). This model incorporates Dirac delta functions which are discontinuous in the image domain. Therefore, statistical analysis is more straightforward in the Fourier domain. To incorporate model-based information about the blur or optical system PSF in this alternate domain, OTF statistical expressions are needed. The random OTF $\mathcal{H}(u, v)$ is defined as [74]

$$\mathcal{H}(u, v) = \frac{W(u\bar{\lambda}f, v\bar{\lambda}f) \star W(u\bar{\lambda}f, v\bar{\lambda}f)}{N_F}, \tag{2.17}$$

where f is the optical system focal length, $u = x/(\bar{\lambda}f)$, $v = y/(\bar{\lambda}f)$, $N_F = W(0, 0) \star W(0, 0)$, \star denotes the two dimensional correlation operation, and $W(x, y)$ is the generalized pupil function which incorporates the phase aberrations such that [74]

$$W(x, y) = W_p(x, y) \exp\{j\phi(x, y)\}. \tag{2.18}$$

$W_p(x, y)$ is a real-valued function describing the unaberrated pupil.

The mean OTF can be derived using two distinct approaches [74]. The first approach relies on an interferometric view of imaging and the Van-Cittert-Zernike Theorem [23]. The second uses a thin screen model and proceeds directly from Eq. (2.17). Regardless of the

approach taken, the result is the straightforward expression [74]

$$\overline{\mathcal{H}}(u, v) = \mathcal{H}_a(u, v)\mathcal{H}_o(u, v), \quad (2.19)$$

where $\mathcal{H}_a(u, v)$ is the transfer function due to the atmospheric turbulence and $\mathcal{H}_o(u, v)$ is the transfer function associated with the aberration-free optics. Fried derived both long exposure and short exposure expressions for $\mathcal{H}_a(u, v)$ [15]. The term “long exposure” refers to the situation in which the imaging system has been exposed to many independent realizations of the atmospheric turbulence. Here, the term “short exposure” refers to the case in which wavefront tilt has been removed. Fried’s long exposure derivation relies on the assumption that the phase perturbation $\phi(x, y)$ is a stationary Gaussian random process [74]. After a change of variables and simplification, the final expression for the long exposure OTF is [74]

$$\mathcal{H}_{a_{LE}}(u, v) = \exp \left\{ -\frac{1}{2} D_{\phi\phi}(u\bar{\lambda}f, v\bar{\lambda}f) \right\}. \quad (2.20)$$

The short exposure or wavefront tilt-removed transfer function $\mathcal{H}_{a_{SE}}(u, v)$ is also of interest. Here, the residual phase after tilt removal $\phi_r(x, y)$ can be written as [74]

$$\phi_r(x, y) = \phi(x, y) - (a_x x + a_y y), \quad (2.21)$$

where a_x and a_y are coefficients describing the tilt of the wavefront phase over the pupil. In his derivation of $\mathcal{H}_{a_{SE}}(u, v)$, Fried assumed that $\phi_r(x, y)$ was uncorrelated with tilt coefficients a_x and a_y [15]. While this assumption is not strictly valid from a mathematical viewpoint, the correlation is small when D/r_o is large [74]. Thus, the short exposure OTF can be written as

$$\mathcal{H}_{a_{SE}}(u, v) = \exp \left\{ -\frac{1}{2} \left(D_{\phi\phi}(u\bar{\lambda}f, v\bar{\lambda}f) - \frac{1}{2} \bar{\lambda}^2 f^2 (a_x^2 + a_y^2) (u^2 + v^2) \right) \right\}. \quad (2.22)$$

Equations (2.20) and (2.22) provide theoretical expressions for the first order OTF statistics. The mean OTF was applied to the reconstruction of atmospheric turbulence-degraded images as early as 1967 [80]. Second order OTF statistics have also been applied to such

problems as speckle imaging [23, 47, 74]. In this application, the speckle transfer function $E\{|\mathcal{H}(u, v)|^2\}$ is an important theoretical quantity. However, the value of a priori knowledge of the correlations between *different* OTF spatial frequencies has yet to be explored. In this case, we require the complete second order statistical quantity

$$R_{\mathcal{H}\mathcal{H}}(u, v; u', v') = E[\mathcal{H}(u, v)\mathcal{H}^*(u', v')], \quad (2.23)$$

where the superscript * denotes a complex conjugate. Following the same technique used to derive an expression for the speckle transfer function [74], the OTF correlation function is

$$\begin{aligned} R_{\mathcal{H}\mathcal{H}}(u, v; u', v') = & \frac{\exp\left\{-\frac{1}{2}\left(D_{\phi\phi}(\bar{\lambda}f\sqrt{u^2+v^2}) + D_{\phi\phi}(\bar{\lambda}f\sqrt{u'^2+v'^2})\right)\right\}}{N_F^2} \\ & \times \int \int \int \int W_p(x, y)W_p(x - u\bar{\lambda}f, y - v\bar{\lambda}f) \\ & \times W_p(x', y')W_p(x' - u'\bar{\lambda}f, y' - v'\bar{\lambda}f) \\ & \times \exp\left\{-\frac{1}{2}\left(D_{\phi\phi}\left(\sqrt{(x-x')^2+(y-y')^2}\right)\right. \right. \\ & \quad \left. \left. + D_{\phi\phi}\left(\sqrt{(x-x'-u\bar{\lambda}f+u'\bar{\lambda}f)^2+(y-y'-v\bar{\lambda}f+v'\bar{\lambda}f)^2}\right)\right. \right. \\ & \quad \left. \left. - D_{\phi\phi}\left(\sqrt{(x-x'-u\bar{\lambda}f)^2+(y-y'-v\bar{\lambda}f)^2}\right)\right. \right. \\ & \quad \left. \left. - D_{\phi\phi}\left(\sqrt{(x-x'+u'\bar{\lambda}f)^2+(y-y'+v'\bar{\lambda}f)^2}\right)\right)\right\} \\ & \times dx dy dx' dy'. \end{aligned} \quad (2.24)$$

Equation (2.24) can be evaluated numerically to give theoretical values for the OTF correlations between any two arbitrary spatial frequencies (u, v) and (u', v') . However, simulation via random phase screen generator and optical system models offers a significant decrease in computational complexity over numerical evaluation. A phase screen generator and Monte Carlo simulation are used to generate the required OTF statistics in this dissertation.

The next section reviews classical estimation techniques important in processing turbulence degraded images. These techniques include least squares estimation, Gerchberg-Saxton algorithms, and ML estimation. In each case, basic theory and past research related to astronomical imaging are reviewed.

2.4 Classical Estimation

2.4.1 Least Squares. Consider an inverse problem based on the image model given in Eqs. (2.8) and (2.9). A simple approach involves ignoring the noise and assuming no a priori knowledge about the object. This method, known as unconstrained least squares (ULS), has an objective function [43]

$$J(\hat{\mathbf{o}}) = \|\mathbf{d} - H\hat{\mathbf{o}}\|^2, \quad (2.25)$$

where \mathbf{d} and \mathbf{o} are vector versions of the detected image and object, respectively. The matrix H is a block-circulant matrix representing the shift-invariant PSF. The notation $\|\bullet\|^2$ represents the Euclidean norm of a column vector, i.e. $\|\mathbf{a}\|^2 = \mathbf{a}^T \mathbf{a}$, and the superscript T denotes the matrix transpose operator. The estimate that minimizes Eq. (2.25) is the well-known expression [43]

$$\hat{\mathbf{o}} = (H^T H)^{-1} H^T \mathbf{d} \quad (2.26)$$

when the matrix H is full rank, thus the invertibility of $H^T H$ is guaranteed. Geometrically, $\hat{\mathbf{o}}$ is the orthogonal projection of \mathbf{d} onto the subspace spanned by the columns of H , as shown in Fig. 2.2. If H is a square matrix and has full rank, Eq. (2.26) reduces to the intuitive form

$$\hat{\mathbf{o}} = H^{-1} \mathbf{d}. \quad (2.27)$$

Recall that H is a block-circulant matrix and is diagonalized by the discrete Fourier transform. Thus, the Fourier domain equivalent to Eq. (2.27) is the scalar inverse filter which can be expressed as [41]

$$\hat{O}(u, v) = \frac{D(u, v)}{\mathcal{H}(u, v)}. \quad (2.28)$$

Consistent with the original objective function given at Eq. (2.25), the inverse filter incorporates no prior knowledge about the object or noise. In addition, Eq. (2.28) is not valid at spatial frequencies (u, v) where $\mathcal{H}(u, v) = 0$. Even when the OTF is non-zero, the filter amplifies noise at high spatial frequencies [41]. Despite these drawbacks, a modified inverse filter is commonly applied to deconvolve AO compensated images [72]. An iterative imple-

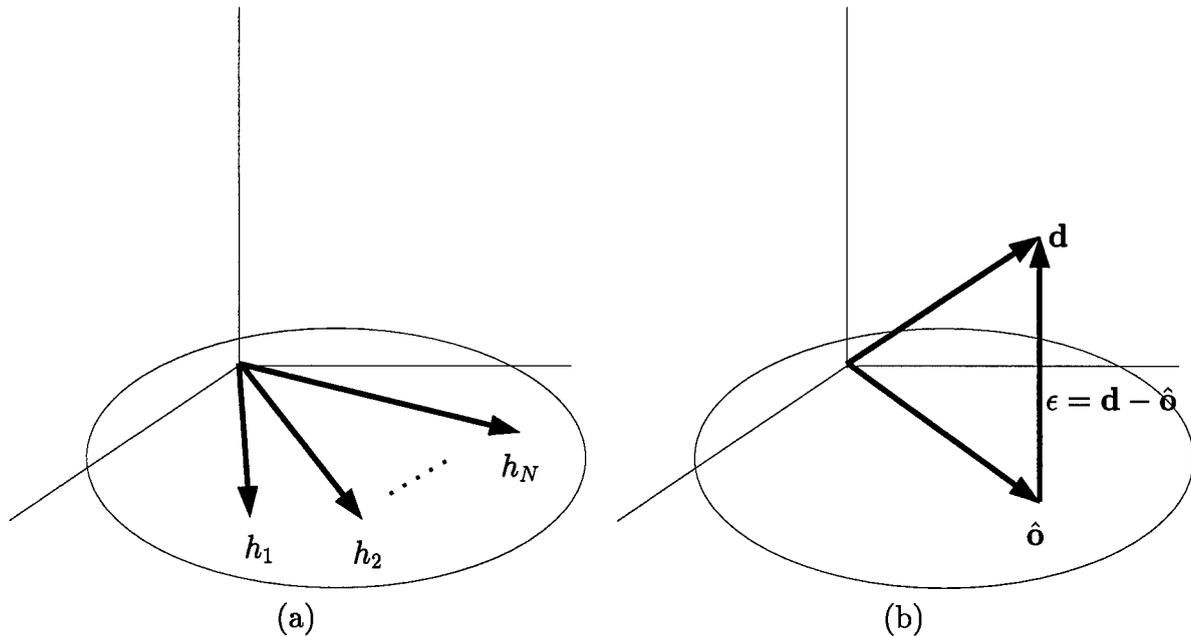


Figure 2.2 Geometrical interpretation of ULS estimation. (a) Subspace spanned by the linearly independent column vectors of the full rank matrix H . (b) The object estimate $\hat{\mathbf{o}}$ is the orthogonal projection of the detected image \mathbf{d} onto the subspace shown in (a). The quantity ϵ denotes the error between data and object estimate.

mentation, known as the Van Cittert method [2], is also available. The iterative approach can be advantageous for two reasons: (1) the iterations can be stopped before convergence and excessive noise amplification, and (2) no matrix inversions are required [2].

Due to the ill-conditioned nature of the previous inversion problem, the ULS solution is corrupted by high spatial frequency noise. Constraints can be used to improve estimator performance by incorporating a priori knowledge. Some signal processing applications use rigid linear constraints which reduce the size of the solution subspace. A typical linear constraint can be written as [43]

$$A\mathbf{o} = \mathbf{b}, \quad (2.29)$$

where A is a known full rank matrix and \mathbf{b} is a known vector. Figure 2.3 illustrates the geometrical interpretation associated with CLS. Note that the constrained solution $\hat{\mathbf{o}}_c$ can be viewed as a projection of the unconstrained solution $\hat{\mathbf{o}}$ onto the constraint subspace [43].

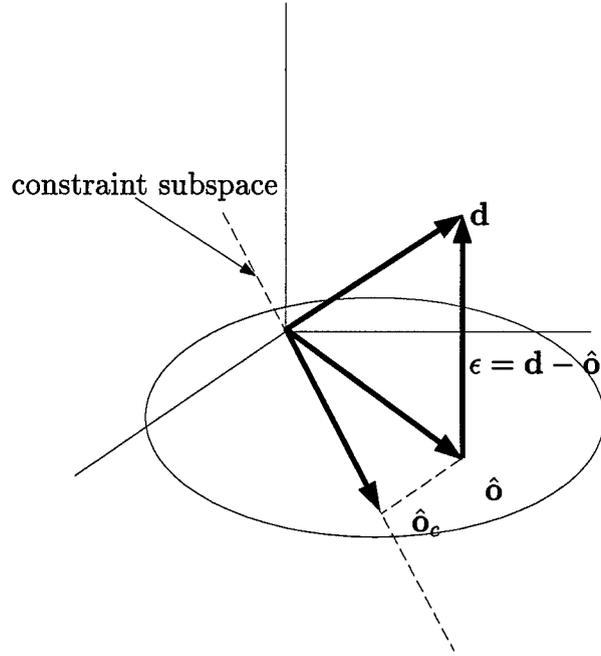


Figure 2.3 Geometrical interpretation of CLS estimation with a rigid linear constraint. The constrained solution $\hat{\mathbf{o}}_c$ is a projection of the unconstrained solution $\hat{\mathbf{o}}$ onto the constraint subspace.

Linear constraints provide a useful illustration of CLS estimation. However, rigid prior knowledge of the true object may not be available when processing astronomical images. Instead, a flexible constraint such as smoothness can be used with Lagrangian minimization. Hunt [37] incorporated a quantitative smoothness measure by minimizing

$$\|C\hat{\mathbf{o}}\|^2, \quad (2.30)$$

subject to

$$\|\mathbf{d} - H\hat{\mathbf{o}}\|^2 = E\{\|\mathbf{n}\|^2\}, \quad (2.31)$$

where \mathbf{n} is the measurement noise and C is a matrix incorporating a smoothness measure such as the two dimensional Laplacian [21]. A straightforward Lagrangian minimization yields the solution [37]

$$\hat{\mathbf{o}} = (H^T H + \gamma C^T C)^{-1} H^T \mathbf{d}, \quad (2.32)$$

where $\gamma = 1/\lambda$ and λ is a Lagrange multiplier. Equation (2.32) is identical to Tikhonov-Miller regularization [59, 86]. Thus, the CLS estimator is valid for a space-variant PSF and corresponding iterative techniques similar to the Van Cittert method are available [2]. When H and C are block circulant, the Fourier domain version of Eq. (2.32) takes the convenient scalar form [37]

$$\hat{O}(u, v) = \frac{\mathcal{H}^*(u, v)D(u, v)}{|\mathcal{H}(u, v)|^2 + \gamma|\mathcal{C}(u, v)|^2}, \quad (2.33)$$

where $\mathcal{C}(u, v)$ is the Fourier transform of the smoothness measure. The smoothness measure provides regularization which is controlled by γ . The optimal Lagrange multiplier λ can be found using a Newton-Raphson minimization based on the statistics of the noise process \mathbf{n} in Eq. (2.31) [37]. Thus, a practical version of the CLS filter in Eq. (2.33) is iterative. The filter provides fidelity to the “rough” inverse filter solution while satisfying the “smooth” constraint.

There are many variants on CLS estimation that have been used in image reconstruction to include weighted least squares [87], constrained total least squares [36], and regularized constrained total least squares [58]. Recently, a new CLS estimation algorithm addressed optimal use of object model information [71]. In practical applications, the object model could be a low resolution image of the object. This work, also known as model-based CLS, is important because it provides a simple method for incorporating object model information using a CLS-based algorithm. CLS estimation has also been applied to images degraded by random blur [3]. As of this writing, CLS estimation has not been applied to noisy DWFS data.

Least squares estimation provides a set of straightforward techniques which can incorporate some a priori knowledge about the imaging scenario. In general, no probabilistic assumptions are made about the data. In many practical applications, some sort of iterative technique is needed to find a useful solution. Another class of iterative techniques which does not require a priori probabilistic information about the data will now be discussed. These techniques are based on the Gerchberg-Saxton phase recovery algorithm [19].

2.4.2 Gerchberg-Saxton Algorithms. In 1972, Gerchberg and Saxton introduced a simple iterative algorithm for the recovery of a complex wave function from only modulus measurements [19]. The method depends on the Fourier transform relationship given by the Van Cittert-Zernike Theorem [23]. The input data are the modulus measurements from source and pupil planes as might be available in electron microscopy [19]. The algorithm begins with an initial guess of the wave function phase. This phase guess is combined with the measured source plane modulus data and then Fourier transformed to the pupil plane domain. Here, the measured pupil plane modulus data is imposed on the new function and the result inverse Fourier transformed back to the source domain. The measured amplitude data is imposed again in the source domain and the process repeated until a suitable convergence criterion is met. The algorithm is based on the fact that a change in amplitude alone in one domain of the Fourier transform results in a change in both amplitude and phase distributions in the other domain [19]. Fienup [7] modified the Gerchberg-Saxton algorithm to address the phase retrieval problem in speckle imaging [47]. The problem here is to find an object consistent with measured Fourier modulus data. In finding an estimate of the object, the Fourier domain phase is “retrieved” in the process. The Fienup solution, known as the error reduction algorithm, involves Fourier transforming back and forth between domains, satisfying the constraints in one before returning to the other. Figure 2.4 shows a block diagram of this simple technique. The technique and related extensions, such as the input-output algorithm [7], have been applied to turbulence-degraded images [8].

The general Gerchberg-Saxton approach can be applied to any problem in which partial constraints (data or a priori information) are available in each of two domains [9]. Thus, problems such as blind deconvolution can be addressed via this paradigm. The term blind deconvolution refers to an image reconstruction problem in which the degrading PSF is unknown. Figure 2.5 gives the block diagram of a blind deconvolution algorithm first proposed by Ayers and Dainty [1]. In this application, no knowledge is available about the PSF or object except the degraded data and image domain constraints. Two modified inverse filters are used to find an object and PSF simultaneously that are consistent with the data and constraints. A similar algorithm has been investigated which replaces the modified inverse

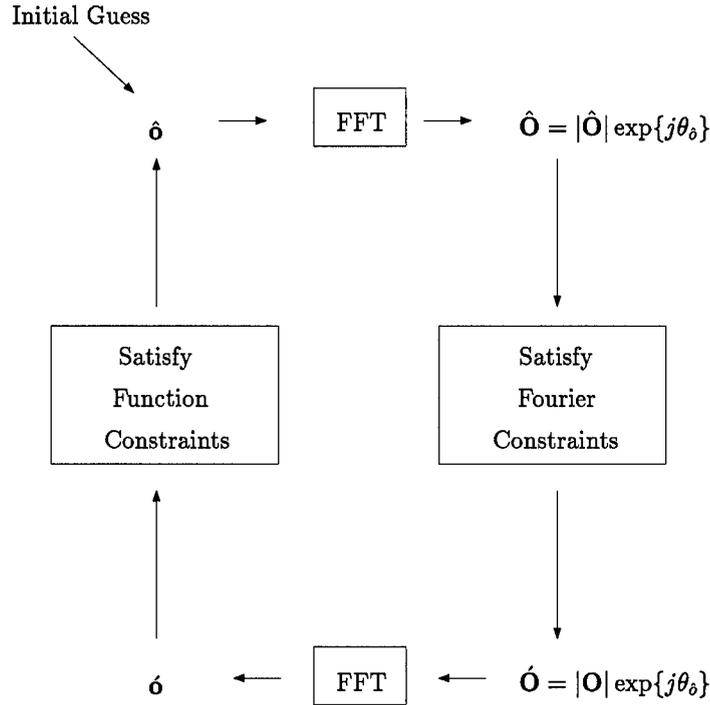


Figure 2.4 Block diagram associated with the error reduction algorithm.

filters with pseudo-Wiener filters [5]. This algorithm has also been successfully extended to incorporate multiple data frames [60].

Gerchberg-Saxton algorithms provide a powerful alternative to least squares processing in many imaging applications. However, as with most iterative algorithms involving non-convex objective functions, convergence to a global “best” solution is not guaranteed. The error reduction algorithm, as applied to the phase retrieval problem, is known to stagnate after a few iterations [9]. Solution uniqueness is also a concern [78]. Iterative blind deconvolution based on the Gerchberg-Saxton approach is also known to suffer from convergence problems and sensitivity to the choice of an initial estimate [46]. A classical method which introduces statistical assumptions about the data is now discussed. Here, statistical analysis will be important to understanding algorithm performance.

2.4.3 Maximum Likelihood. Let us now consider an image reconstruction problem in which the probability density function (PDF) of the detected data \mathbf{d} , denoted $p_{\mathbf{d}}(\mathbf{d}; \mathbf{o}, H)$,

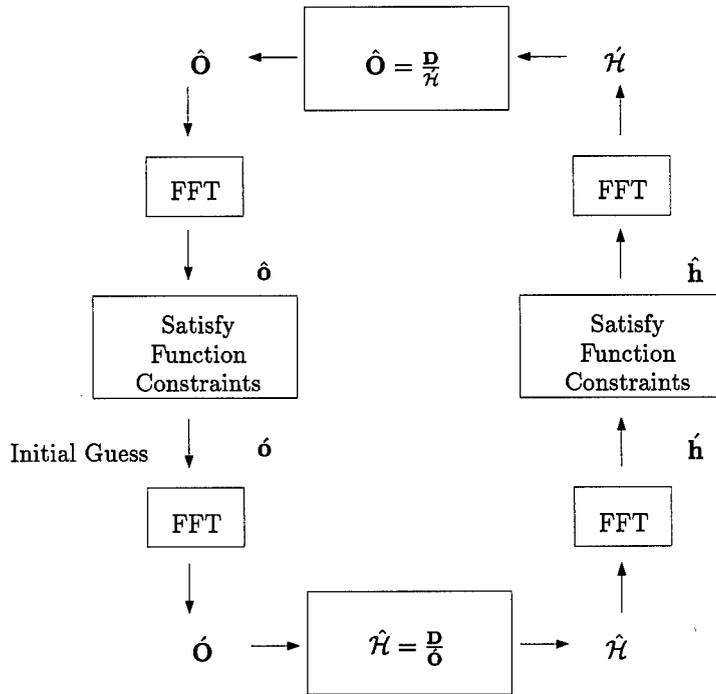


Figure 2.5 Block diagram associated with Ayers-Dainty iterative blind deconvolution.

is known. Here, the semi-colon indicates that the PDF is parameterized by \mathbf{o} and H . The ML principle can be stated as [43]:

Given a realization of \mathbf{d} and the PSF matrix H , find $\hat{\mathbf{o}}$ which maximizes $p_{\mathbf{d}}(\mathbf{d}; \mathbf{o}, H)$.

In general, the ML estimate is obtained by evaluating the likelihood function using the data realization \mathbf{d} and then searching for the appropriate $\hat{\mathbf{o}}$. When the likelihood function is continuously differentiable in \mathbf{o} and convex, the ML estimate may be determined by differentiating $p_{\mathbf{d}}(\mathbf{d}; \mathbf{o}, H)$ with respect to \mathbf{o} , setting equal to the zero vector, and solving for \mathbf{o} . ML estimation is a popular technique because it is a “turn-the-crank” procedure [43] for complicated estimation problems. ML estimation is also closely related to least squares when the likelihood function is Gaussian distributed. In this case, the ML and weighted least squares estimators are identical when the weighting matrix is chosen as the inverse of the data covariance matrix [43].

In some cases, the ML estimator cannot be found by taking the analytic derivative of the likelihood function. Another advantage of ML estimation is that $\hat{\mathbf{o}}$ can always be found by maximizing the likelihood function numerically. A straightforward grid search can be used

if \mathbf{o} is known to exist on a finite interval. However, when processing imagery, the unknown parameters represent irradiance values, which do not always exist on the required finite interval. Thus, iterative maximization methods must be used. Some commonly used iterative maximization techniques are Newton-Raphson [40], gradient descent [41], and expectation-maximization (EM) [6] algorithms. The EM algorithm is of particular interest in processing astronomical images because it is well suited to vector parameters. The EM algorithm is built on the hypothesis that some data sets allow easier determination of the ML estimate than the data \mathbf{d} . This new data set \mathbf{y} is known as the complete data, while \mathbf{d} is known as the incomplete data. The standard procedure is to assume that there is a many-to-one complete to incomplete data transformation. Thus, the EM strategy is to trade the difficult problem of maximizing $p_{\mathbf{d}}(\mathbf{d}; \mathbf{o}, H)$ for the easier problem of maximizing $p_{\mathbf{y}}(\mathbf{y}; \mathbf{o}, H)$. Since \mathbf{y} does not really exist, the algorithm incorporates the following iterative expectation-maximization procedure:

1. **Expectation (E)** Evaluate $E_{\mathbf{y}|\mathbf{d}}[p_{\mathbf{y}}(\mathbf{y}; \mathbf{o}_k, H)]$ using the k^{th} object estimate. The notation $E_{A|B}$ denotes the conditional expected value of the random variable A given B.
2. **Maximization (M)** Use the k^{th} conditional expectation from the previous step to generate \mathbf{o}_{k+1} .
3. Repeat the expectation step using \mathbf{o}_{k+1} .

The EM algorithm is guaranteed to converge to at least a local minimum under some mild mathematical conditions [6, 43].

ML estimation in all its various forms has been widely applied in image reconstruction to include astronomy [51, 68, 81], tomography [79], fluorescence microscopy [32, 34], and a variety of blind deconvolution problems [33, 48, 77, 85]. As noted above, iterative maximization techniques, such as the EM algorithm, are widely used to generate the ML estimate. These iterative techniques are not guaranteed to converge to a global maximum [43]. Instead, ML algorithms may suffer convergence problems due to numerical inaccuracies, sensitivity to local minima, and the choice of an initial estimate. Convergence problems are a special

concern in the blind deconvolution application [46]. Finally, ML estimation requires the maximization of a random likelihood function. Thus, predictive performance analysis can be difficult.

ML estimation is a powerful statistical technique which relies on knowledge of the data PDF. In the section below, Bayesian estimation, which relies on statistical assumptions about data and *object*, is discussed. The Bayesian techniques reviewed in this chapter include the Wiener filter, Kalman filter, and MAP estimation. Once again, basic theory and past research related to astronomical imaging are reviewed.

2.5 Bayesian Estimation

2.5.1 Wiener Filter. The main drawback to ULS processing and the inverse filter is sensitivity to noise. This deficiency is no surprise since ULS does not take into account noise effects. In contrast, the Wiener filter incorporates a statistical description of object and noise such that the object estimate $\hat{\mathbf{o}}$ can be written as [64]

$$\hat{\mathbf{o}} = R_{oo}H^T(HR_{oo}H^T + R_{nn})^{-1}\mathbf{d}, \quad (2.34)$$

where $R_{oo} = E[\mathbf{o}\mathbf{o}^T]$ is the object correlation matrix and $R_{nn} = E[\mathbf{n}\mathbf{n}^T]$ is the noise correlation matrix. In general image reconstruction applications, the mean object, denoted $\bar{\mathbf{o}}$, is non-zero. Equation (2.34) is derived based on minimizing mean square error (MSE) between the true object irradiance \mathbf{o} and $\hat{\mathbf{o}}$. The object estimate $\hat{\mathbf{o}}$ is constrained to be linear to the related detected data \mathbf{d} , thus producing a mathematically tractable non-iterative solution. In this application, the object statistical model is static. A more general form could allow for a dynamic object model [43]. When object and noise are stationary, R_{oo} and R_{nn} are block Toeplitz matrices [62]. These correlation matrices can be made to approximate block circulant matrices and, therefore, are diagonalized by the discrete Fourier transform [21]. The result is a Fourier domain scalar Wiener filter given as [21]

$$\hat{O}(u, v) = \frac{\mathcal{H}^*(u, v)D(u, v)}{|\mathcal{H}(u, v)|^2 + \frac{\mathcal{G}_n(u, v)}{\mathcal{G}_o(u, v)}}, \quad (2.35)$$

where $\mathcal{G}_n(u, v) = E[|N(u, v)|^2]$ and $\mathcal{G}_o(u, v) = E[|O(u, v)|^2]$ are the noise and object PSDs, respectively. This approach has been used in many imaging applications to include enhancement of scanning electron micrographs [49], multichannel processing [18], image recognition [52], and deconvolution of AO compensated images [72]. The scalar Wiener filter is suboptimal with respect to MSE for non-stationary object ensembles. Since many object ensembles are non-stationary in the mean and may be non-stationary with regard to covariance, the filter may not incorporate valuable a priori information about a given object class.

Pratt [63] proposed a generalized vector Wiener filter that is optimal with respect to MSE for non-stationary object ensembles in signal-independent noise. Pratt's Wiener filter theory has been extended to images degraded by both blur and signal-independent noise. The resultant Fourier domain filter expression can be written as [76]

$$\hat{\mathbf{O}} = R_{OO}\mathcal{H}_d^*\{\mathcal{H}_dR_{OO}\mathcal{H}_d^* + R_{NN}\}^{-1}\mathbf{D}, \quad (2.36)$$

where $R_{OO} = E[\mathbf{OO}^H]$ is the spatial frequency correlation matrix of the object, $R_{NN} = E[\mathbf{NN}^H]$ is the spatial frequency correlation matrix of the signal-independent noise, \mathcal{H}_d is a diagonal matrix of known OTF elements, and the superscript $*$ denotes the complex conjugate of the matrix elements. The role of the object correlations between *different* spatial frequencies has not been extensively studied. In addition, Eq. (2.36) does not consider a random OTF due to atmospheric turbulence.

Imaging through atmospheric turbulence is a severe manifestation of the broader problem of randomness in the pupil function of an optical system. A random pupil function can be caused by a number of factors such as camera movement relative to an object, dust particles on optical surfaces, or turbulence in water. Wiener filter theory has been applied to images degraded by random blur. Slepian [80] extended the Helstrom scalar Wiener filter to incorporate the first order OTF statistic $E[\mathcal{H}(u, v)]$ and the second order OTF statistic $E[|\mathcal{H}(u, v)|^2]$ associated with atmospheric turbulence. Ward and Saleh [87, 88] investigated an iterative Wiener filter for one dimensional data. Guan and Ward modified this iterative Wiener filter to process two dimensional images [25] and investigated a closely related

geometrical mean filter [26]. This previous work does not investigate the role of OTF correlations between *different* spatial frequencies. In addition, shot noise effects are not considered. The Wiener filters given in Eqs. (2.34)-(2.36) incorporate static statistical models. Both the Wiener and Kalman filters can also incorporate dynamic models. In the next section, the Kalman filter in adaptive image reconstruction is briefly discussed.

2.5.2 Kalman Filter. The Wiener filter given in Eq. (2.34) is based totally on knowledge of the data autocorrelation and their cross-correlation with the object. Also, changes in the object model with respect to time are not considered. General Wiener-Kalman filter theory can provide an adaptive image reconstruction capability when dynamic statistical models are available. The seminal work by Woods and Radewan [96] led to a two dimensional reduced update scalar Kalman filter (RUKF). Here, only the pixels in a small neighborhood of the current pixel are updated. It is assumed that a pixel is uncorrelated with other pixels outside this neighborhood called the update region [42]. The RUKF has been applied to deconvolution-type problems [95]. More recent advances in Kalman filter processing of two dimensional data involve a reduced order model (ROM) representation [42] and the ROM Kalman filter [97]. A two dimensional Kalman filter has also been investigated for images degraded by random blur [67]. However, the necessary dynamic object state model is not available in many applications.

The next section introduces a Bayesian technique which is closely related to the ML estimator. However, the MAP estimator incorporates a prior PDF associated with the object.

2.5.3 Maximum A Posteriori. The MAP estimation principle can be stated as follows [43]:

Given a realization of \mathbf{d} and the PSF matrix H , find $\hat{\mathbf{o}}$ which maximizes the posterior PDF $p_{\mathbf{o}}(\mathbf{o}|\mathbf{d}; H)$.

The posterior PDF is a conditional distribution. Maximizing $p_{\mathbf{o}}(\mathbf{o}|\mathbf{d}; H)$ has been shown to minimize a “hit-or-miss” cost function which assigns no penalty for small errors and maximum penalty for all errors in excess of a threshold δ [43]. This cost function can be

expressed mathematically as [43]

$$\text{Cost}(\epsilon) = \begin{cases} 0 & |\epsilon| < \delta \\ 1 & |\epsilon| > \delta \end{cases}, \quad (2.37)$$

where the variable ϵ represents error and $\delta > 0$. If δ is made very small, this cost function assigns the same penalty for all errors and no penalty for no error.

The posterior PDF can be written in a more intuitive form. Using Bayes rule [62], $p_{\mathbf{o}}(\mathbf{o}|\mathbf{d}; H)$ becomes

$$p_{\mathbf{o}}(\mathbf{o}|\mathbf{d}; H) = \frac{p_{\mathbf{d}}(\mathbf{d}; H|\mathbf{o})p_{\mathbf{o}}(\mathbf{o})}{p_{\mathbf{d}}(\mathbf{d}; \mathbf{o}, H)}. \quad (2.38)$$

Thus, maximizing the numerator of Eq. (2.38) is equivalent to maximizing the posterior PDF. Notice that $p_{\mathbf{d}}(\mathbf{d}; H|\mathbf{o})p_{\mathbf{o}}(\mathbf{o})$ is very similar to the ML likelihood function given in Section 2.4.3 except that the object is now considered a random process with known prior PDF $p_{\mathbf{o}}(\mathbf{o})$. As in the ML case, a candidate MAP estimate can be found by evaluating $p_{\mathbf{d}}(\mathbf{d}; H|\mathbf{o})p_{\mathbf{o}}(\mathbf{o})$ at the given data realization, differentiating with respect to \mathbf{o} , setting equal to the zero vector, and solving for \mathbf{o} . When a closed form solution is not practical, the MAP estimate can be found via iterative maximization methods such as gradient ascent algorithms [24, 38] and the EM algorithm [28]. Iterative MAP estimation can suffer from the same numerical convergence problems noted for ML estimation in Section 2.4.3.

2.6 Summary

This chapter provided background material related to image reconstruction. The emphasis here is on the deconvolution of astronomical images. Thus, the image degradation model associated with atmospheric turbulence, shot noise, and detector read noise was reviewed as well as statistical theory related to atmospheric turbulence. The rest of the chapter presented important classical and Bayesian estimation techniques, most of which involve iterative optimization. While these techniques are powerful, drawbacks related to convergence and ease of analysis do exist. Current linear reconstruction methods, such as the scalar Wiener filter, do not improve Fourier domain signal-to-noise ratio (SNR). In addition, the application of object, OTF, and noise correlations between *different* spatial frequencies has

not been investigated. In the rest of this dissertation, two techniques which bridge the gap between linear processing and nonlinear iterative optimization are presented: CLS processing of DWFS data and the vector Wiener filter. The vector Wiener filter provides a useful, non-iterative complement to nonlinear iterative optimization. The complete vector Wiener filter is derived in Chapter IV with performance data given in Chapters V and VI. First, Chapter III presents a novel application of CLS post-detection image processing.

III. Constrained Least Squares Incorporating Wavefront Sensing

3.1 Introduction

In this chapter, a technique for processing noisy deconvolution from wavefront sensing (DWFS) data based on constrained least squares (CLS) estimation is presented. The new algorithm selects a value for the regularization constant which is consistent with the ensemble-averaged data and a constraint [11]. This problem is solved using the Lagrange multiplier technique [21,37]. A closed form solution for the object estimate is obtained which is analogous to the traditional DWFS estimator [20,75] with the Lagrange multiplier serving as a regularization constant. An iterative approach based on a Newton-Raphson minimization [40,71] is used to find the optimal Lagrange multiplier. The iteration incorporates the statistics of both the optical transfer function (OTF) and noise. The sample results given here demonstrate that CLS estimation provides high quality processing of noisy DWFS data *automatically*. No ad hoc tuning of the regularization constant is necessary. CLS object estimates are compared with those processed via manual parameter selection. In all cases, the new technique provides images with comparable resolution. In addition, the algorithm is computationally inexpensive, converging to a solution in less than 10 iterations [11].

The rest of this chapter is organized as follows. Section 3.2 reviews the traditional DWFS estimator. In Section 3.3, a CLS algorithm is derived which incorporates a new constraint based on the ensemble-averaged DWFS data. Section 3.4 gives sample results associated with various imaging conditions, while Section 3.5 provides a brief summary.

3.2 Traditional Estimator

Raw DWFS data consists of an ensemble of short exposure images and corresponding estimates of the wavefront phase from a wavefront sensor (WFS). The phase estimates $\tilde{\phi}_i(x, y)$ can be used to generate an estimate of the OTF $\tilde{\mathcal{H}}_i(u, v)$ via a normalized autocorrelation of the generalized pupil function as shown in Eq. (2.17). The subscript i refers to the i^{th} realization, while the tilde denotes a quantity estimated directly from information provided

by WFS hardware. The DWFS estimator suggested by Primot *et al.* [66] can be written as

$$\hat{O}(u, v) = \frac{\langle \tilde{\mathcal{H}}^*(u, v)I(u, v) \rangle}{\langle |\tilde{\mathcal{H}}(u, v)|^2 \rangle}, \quad (3.1)$$

where $\langle X(u, v) \rangle$ is defined as

$$\langle X(u, v) \rangle = \frac{1}{L} \sum_{i=1}^L X_i(u, v), \quad (3.2)$$

$I(u, v)$ is the noiseless image Fourier spectrum and L is the number of images in the ensemble. Equation (3.1) does not provide acceptable reconstructions due to residual noise in the OTF estimation process. When detector noise is present, $I(u, v)$ is not available and the detected image Fourier spectrum $D(u, v)$ must be substituted in Eq. (3.1). The standard solution is to incorporate a regularization constant ϵ in the filter denominator which gives the new estimator [20, 93]

$$\hat{O}(u, v) = \frac{\langle \tilde{\mathcal{H}}^*(u, v)D(u, v) \rangle}{\langle |\tilde{\mathcal{H}}(u, v)|^2 \rangle + \epsilon}. \quad (3.3)$$

The regularization constant ϵ is adjusted in a completely ad hoc manner by the user based on the perceived image quality.

In the next section, a CLS algorithm is derived which takes advantage of noise reduction through ensemble averaging by directly incorporating $\langle \tilde{\mathcal{H}}^*(u, v)D(u, v) \rangle$ and $\langle |\tilde{\mathcal{H}}(u, v)|^2 \rangle$. Thus, the objective function used to derive this estimator takes on an unfamiliar form when compared to more traditional CLS applications in image processing [21, 37, 71].

3.3 Modified Constrained Least Squares Formulation

3.3.1 Problem Statement. This CLS estimation problem can be stated as follows: “Given the DWFS-derived estimate $\langle \tilde{\mathcal{H}}^*(u, v)D(u, v) \rangle$, the ensemble-averaged magnitude-squared OTF estimate $\langle |\tilde{\mathcal{H}}(u, v)|^2 \rangle$, and statistical models for the OTF and noise, find the CLS optimal estimate of the object that caused the detected image ensemble.” To accomplish

this task, consider finding an object estimate $\hat{\mathbf{o}}$ that minimizes

$$\langle \|\tilde{H}C\hat{\mathbf{o}}\|^2 \rangle, \quad (3.4)$$

subject to the constraint

$$\|\langle \tilde{H}^T \mathbf{d} \rangle - \langle \tilde{H}^T \tilde{H} \rangle \hat{\mathbf{o}}\|^2 = \text{E}\{\|H^T \mathbf{n}\|^2\}, \quad (3.5)$$

where C is a block-circulant constraint matrix which enforces prior knowledge of the true object such as smoothness or support and \tilde{H}_i is the block-circulant estimated PSF matrix for the i^{th} realization. Here, the $\langle \tilde{H}^*(u, v)D(u, v) \rangle$ and $\langle |\tilde{H}(u, v)|^2 \rangle$ quantities are incorporated in the constraint function directly. Also, the MSE associated with the measurement is now $\text{E}\{\|H^T \mathbf{n}\|^2\}$ instead of the norm-squared of the noise $\text{E}\{\|\mathbf{n}\|^2\}$ used previously [21, 37]. Even though Eqs. (3.4) and (3.5) are unfamiliar with respect to traditional CLS applications in image processing [21, 37, 71], it will be shown that the associated Fourier domain filter has the same general form as Eq. (3.3).

3.3.2 Closed Form Solution. The appropriate objective function $J(\hat{\mathbf{o}}, \lambda)$ for the Lagrange minimization can be written as

$$J(\hat{\mathbf{o}}, \lambda) = \hat{\mathbf{o}}^T C^T \langle \tilde{H}^T \tilde{H} \rangle C \hat{\mathbf{o}} + \lambda \left\{ \|\langle \tilde{H}^T \mathbf{d} \rangle - \langle \tilde{H}^T \tilde{H} \rangle \hat{\mathbf{o}}\|^2 - \text{E}\{\|H^T \mathbf{n}\|^2\} \right\}, \quad (3.6)$$

where λ is a Lagrange multiplier [21, 37]. Since the Lagrangian in Eq. (3.6) is quadratic, it can be minimized by differentiating $J(\hat{\mathbf{o}}, \lambda)$ with respect to $\hat{\mathbf{o}}$, setting the derivative equal to zero, and solving for $\hat{\mathbf{o}}$. Applying the appropriate vector-matrix identities [56] to take the derivatives in Eq. (3.6) yields

$$\hat{\mathbf{o}} = \left(\langle \tilde{H}^T \tilde{H} \rangle \langle \tilde{H}^T \tilde{H} \rangle + \gamma C^T \langle \tilde{H}^T \tilde{H} \rangle C \right)^{-1} \langle \tilde{H}^T \tilde{H} \rangle \langle \tilde{H}^T \mathbf{d} \rangle, \quad (3.7)$$

where $\gamma = 1/\lambda$. It is not computationally efficient to evaluate Eq. (3.7) directly since C and \tilde{H} are large matrices for standard-sized image arrays. However, these matrices are

block circulant. Block circulant matrices are diagonalized by the discrete Fourier transform allowing the transform domain equivalent of Eq. (3.7) to provide the simpler scalar form [21, 37]

$$\hat{O}(u, v) = \frac{\langle \tilde{\mathcal{H}}^*(u, v) D(u, v) \rangle}{\langle |\tilde{\mathcal{H}}(u, v)|^2 \rangle + \gamma |\mathcal{C}(u, v)|^2}, \quad (3.8)$$

where $\mathcal{C}(u, v)$ is the Fourier transform of the constraint $c(x, y)$ associated with the matrix C . When $\mathcal{C}(u, v) = 1$ for all spatial frequencies (u, v) , Eq. (3.8) is identical to Eq. (3.3). However, Eq. (3.8) will accommodate a general Fourier domain constraint function which can be tailored to prior knowledge or a specific application.

3.3.3 Newton-Raphson Iterative Solution. Instead of tuning γ manually, a Newton-Raphson technique is used to find the parameter value which minimizes the objective function $J(\hat{\mathbf{o}}, \lambda)$ given in Eq. (3.6). To implement the Newton-Raphson technique, the derivative of $J(\hat{\mathbf{o}}, \lambda)$ with respect to λ is required. This derivative is straightforward to derive and given by

$$\frac{\partial J(\hat{\mathbf{o}}, \lambda)}{\partial \lambda} = \|\langle \tilde{H}^T \mathbf{d} \rangle - \langle \tilde{H}^T \tilde{H} \rangle \hat{\mathbf{o}}\|^2 - \mathbb{E}\{\|\mathbf{H}^T \mathbf{n}\|^2\}. \quad (3.9)$$

This iteration can also be accomplished with respect to the γ parameter [37]. However, the required derivative is not as straightforward as Eq. (3.9) [37, 71]. The first norm-squared term in Eq. (3.9) is a function of the DWFS data and the object estimate. It can be computed conveniently using the Fourier domain quantities $\langle \tilde{\mathcal{H}}^*(u, v) D(u, v) \rangle$ and $\langle |\tilde{\mathcal{H}}(u, v)|^2 \rangle$. However, the second norm-squared term $\mathbb{E}\{\|\mathbf{H}^T \mathbf{n}\|^2\}$ is not a function of the DWFS data and must be derived.

Rewriting $\mathbb{E}\{\|\mathbf{H}^T \mathbf{n}\|^2\}$ in the Fourier domain using Parseval's Theorem [21] yields

$$\begin{aligned} \mathbb{E}\{\|\mathbf{H}^T \mathbf{n}\|^2\} &= \mathbb{E}\left\{\sum_{u,v} |\mathcal{H}^*(u, v) N(u, v)|^2\right\} \\ &= \mathbb{E}\left\{\sum_{u,v} |\mathcal{H}^*(u, v) (I(u, v) - D(u, v))|^2\right\} \\ &= \mathbb{E}\left\{\sum_{u,v} |\mathcal{H}^*(u, v) (\mathcal{H}(u, v) O(u, v) - D(u, v))|^2\right\}. \end{aligned} \quad (3.10)$$

Substituting the image model given in Eq. (2.7) into Eq. (3.10) introduces the noise statistics such that

$$\begin{aligned} \mathbb{E}\{|H^T \mathbf{n}|^2\} = \mathbb{E} \left\{ \sum_{u,v} |\mathcal{H}^*(u,v) (\overline{K} \mathcal{H}(u,v) O_n(u,v) \right. \\ \left. - \sum_{n=1}^K \exp[-j2\pi(ux_n + vy_n)] \right. \\ \left. - \sum_{p=1}^P n_p \exp\{-j2\pi(ux_p + vy_p)\}) \right|^2 \Bigg\}, \end{aligned} \quad (3.11)$$

where (x_n, y_n) is the location of the n^{th} photoevent in the image plane, (x_p, y_p) is the location of the p^{th} image pixel, and \overline{K} is the average number of photoevents per image. The normalized object spectrum $O_n(u, v) = O(u, v)/\overline{K}$ has also been introduced to the above expression. Now the linearity property can be used to move the expectation inside the summation over spatial frequencies (u, v) in order to evaluate one term of Eq. (3.11). At this point, the right side of Eq. (3.11) can be simplified using standard techniques for evaluating expectations of doubly stochastic Poisson random processes [74]. These techniques use nested conditional expectations over the random quantities (x_n, y_n) , K , and \mathcal{H} [74]. Nine sub-terms of Eq. (3.11) are evaluated below:

1. $\mathbb{E}\{(\overline{K})^2 |O_n(u, v)|^2 |\mathcal{H}(u, v)|^4\}$. Here, the only random quantity is $|\mathcal{H}(u, v)|^4$. Thus, the \overline{K} and $O_n(u, v)$ terms are brought outside the expectation yielding the final result

$$\text{Term 1} = (\overline{K})^2 |O_n(u, v)|^2 \mathbb{E}\{|\mathcal{H}(u, v)|^4\}. \quad (3.12)$$

2. $-\mathbb{E}\left\{\overline{K} O_n(u, v) |\mathcal{H}(u, v)|^2 \mathcal{H}(u, v) \sum_{n=1}^K \exp[j2\pi(ux_n + vy_n)]\right\}$. The random quantities underlying this expectation are (x_n, y_n) , K , and \mathcal{H} . Bayes rule [62] can be used to rewrite the joint probability density function (PDF) in the first term using conditional PDFs. These conditional PDFs translate to conditional expectations such that the second term can be written as

$$\begin{aligned} \text{Term 2} &= -\bar{K}O_n(u, v) \\ &\times E_{\mathcal{H}} \left(|\mathcal{H}(u, v)|^2 \mathcal{H}(u, v) E_{K|\mathcal{H}} \left\{ E_{x_n, y_n|K, \mathcal{H}} \left[\sum_{n=1}^K \exp[j2\pi(ux_n + vy_n)] \right] \right\} \right) \end{aligned} \quad (3.13)$$

The innermost expectation $E_{x_n, y_n|K, \mathcal{H}}[\bullet]$ has been evaluated previously [74]. The result is repeated below as

$$E_{x_n, y_n|K, \mathcal{H}}[\bullet] = K\mathcal{H}^*(u, v)O_n^*(u, v), \quad (3.14)$$

where the \bullet notation represents the argument of the innermost expectation in Eq. (3.13). Substituting Eq. (3.14) into Eq. (3.13) and evaluating the remaining nested conditional expectations gives the final result for the second term

$$\text{Term 2} = -(\bar{K})^2 |O_n(u, v)|^2 E\{|\mathcal{H}(u, v)|^4\}. \quad (3.15)$$

3. $-E\{\bar{K}O_n(u, v)|\mathcal{H}(u, v)|^2 \mathcal{H}(u, v) \sum_{p=1}^P n_p \exp[j2\pi(ux_p + vy_p)]\}$. In Chapter II, the detector read noise random variable n_p was assumed to be zero mean and statistically independent of K and \mathcal{H} . Thus, the third term is zero.

$$\text{Term 3} = 0. \quad (3.16)$$

4. $-E\{\bar{K}O_n^*(u, v)|\mathcal{H}(u, v)|^2 \mathcal{H}^*(u, v) \sum_{n=1}^K \exp[-j2\pi(ux_n + vy_n)]\}$. Term 4 is the complex conjugate of Term 2 above. Thus, the final result for the fourth term is

$$\text{Term 4} = -(\bar{K})^2 |O_n(u, v)|^2 E\{|\mathcal{H}(u, v)|^4\}. \quad (3.17)$$

5. $E\{|\mathcal{H}(u, v)|^2 \sum_{n=1}^K \sum_{m=1}^K \exp[-j2\pi(u(x_n - x_m) + v(y_n - y_m))]\}$. The random quantities underlying this expectation are (x_n, y_n) , (x_m, y_m) , K , and \mathcal{H} . As with the second

term above, conditional expectations can be used to rewrite the fifth term such that

$$\text{Term 5} = E_{\mathcal{H}} \left(|\mathcal{H}(u, v)|^2 E_{K|\mathcal{H}} \left\{ E_{x_n, y_n, x_m, y_m | K, \mathcal{H}} [\bullet] \right\} \right). \quad (3.18)$$

Here, \bullet represents the double summation shown in the first line of 5 above. The innermost expectation $E_{x_n, y_n, x_m, y_m | K, \mathcal{H}} [\bullet]$ has been evaluated previously [74]. The result is repeated below as

$$E_{x_n, y_n, x_m, y_m | K, \mathcal{H}} [\bullet] = K + (K^2 - K) |O_n(u, v)|^2 |\mathcal{H}(u, v)|^2. \quad (3.19)$$

To evaluate the expectation over K , recall that the image photon count is Poisson distributed which implies that $\overline{K^2} = (\overline{K})^2 + \overline{K}$ [62, 74]. Thus, the final result for Term 5 is

$$\text{Term 5} = \overline{K} E\{|\mathcal{H}(u, v)|^2\} + (\overline{K})^2 |O_n(u, v)|^2 E\{|\mathcal{H}(u, v)|^4\}. \quad (3.20)$$

6. $E \left\{ |\mathcal{H}(u, v)|^2 \left(\sum_{n=1}^K \exp[-j2\pi(ux_n + vy_n)] \right) \left(\sum_{p=1}^P n_p \exp[j2\pi(ux_p + vy_p)] \right) \right\}$. The sixth term is zero since n_p is zero mean and statistically independent of (x_n, y_n) , K , and \mathcal{H} .

$$\text{Term 6} = 0. \quad (3.21)$$

7. $-E \left\{ \overline{K} O_n^*(u, v) |\mathcal{H}(u, v)|^2 \mathcal{H}^*(u, v) \sum_{p=1}^P n_p \exp[-j2\pi(ux_p + vy_p)] \right\}$. The seventh term is zero since n_p is zero mean and statistically independent of K and \mathcal{H} .

$$\text{Term 7} = 0. \quad (3.22)$$

8. $E \left\{ |\mathcal{H}(u, v)|^2 \left(\sum_{p=1}^P n_p \exp[-j2\pi(ux_p + vy_p)] \right) \left(\sum_{n=1}^K \exp[j2\pi(ux_n + vy_n)] \right) \right\}$. The eighth term is the complex conjugate of the sixth term. Thus,

$$\text{Term 8} = 0. \quad (3.23)$$

9. $E \left\{ |\mathcal{H}(u, v)|^2 \sum_{p=1}^P \sum_{q=1}^P n_p n_q \exp[-j2\pi(u(x_p - x_q) + v(y_p - y_q))] \right\}$. First, recall that \mathcal{H} and n_p are statistically independent. Therefore, the ninth term can be written as

$$\text{Term 9} = E \left\{ |\mathcal{H}(u, v)|^2 \right\} E \left\{ \sum_{p=1}^P \sum_{q=1}^P n_p n_q \exp[-j2\pi(u(x_p - x_q) + v(y_p - y_q))] \right\} \quad (3.24)$$

The second expectation in Eq. (3.21) has been evaluated previously [74] for the assumptions given in Chapter II such that the final result for Term 9 can be expressed as

$$\text{Term 9} = P\sigma_r^2 E \{ |\mathcal{H}(u, v)|^2 \}. \quad (3.25)$$

Adding the nine sub-terms presented above gives the result for a single spatial frequency of Eq. (3.11):

$$E \left\{ |\mathcal{H}^*(u, v) N(u, v)|^2 \right\} = (\bar{K} + P\sigma_r^2) E \left\{ |\mathcal{H}(u, v)|^2 \right\}. \quad (3.26)$$

Summing the result in Eq. (3.26) over all spatial frequencies gives the final result

$$E \{ \|H^T \mathbf{n}\|^2 \} = (\bar{K} + P\sigma_r^2) \sum_{u,v} E \left\{ |\mathcal{H}(u, v)|^2 \right\}. \quad (3.27)$$

If the OTF statistics are unavailable in Eq. (3.27), the OTF estimate data can be substituted, which yields

$$E \{ \|H^T \mathbf{n}\|^2 \} \approx (\bar{K} + P\sigma_r^2) \sum_{u,v} \langle |\tilde{\mathcal{H}}(u, v)|^2 \rangle. \quad (3.28)$$

The result given in Eq. (3.28) is used in Eq. (3.9) to find the current derivative of $J(\hat{\mathbf{o}}, \lambda)$ with respect to the Lagrange multiplier λ . This derivative is then used in the Newton-Raphson iteration to update the value of λ and generate an object estimate using Eq. (3.7) or its Fourier domain equivalent Eq. (3.8). The iteration continues until the object estimate meets a pre-determined stopping criterion. A variety of criteria can be used to stop the Newton-Raphson iteration [40]. Since the goal is to minimize the Lagrangian given in Eq. (3.6), the iteration is stopped when $J(\hat{\mathbf{o}}, \lambda)$ is sufficiently small. All CLS images shown in this chapter were processed until $J(\hat{\mathbf{o}}, \lambda) \leq 0.001$. For applications in which the appropriate value for the Lagrangian is unknown, the algorithm can terminate when the change in λ has stabilized

from iteration to iteration. In the next section, sample results are presented to illustrate the technique.

3.4 Sample Results

In this section, sample results are presented which illustrate CLS processing of DWFS data. A variety of imaging conditions are investigated by varying turbulence strength, light level, and detector read noise variance. Data is presented for uncompensated and adaptive optics (AO) compensated images of a representative satellite object.

3.4.1 Assumptions. The satellite object shown in Fig. 3.1 was used to generate all data in this chapter. The detector array was 256×256 elements for a total of $P = 65536$ pixels. The satellite was 10m in length and in low earth orbit at a range of 500km. The diameter of the telescope was 1m with both imaging and wavefront sensor wavelengths set at $\bar{\lambda} = 500\text{nm}$. The simulated spectral bandwidth was $\pm 5\%$ of $\bar{\lambda}$, with the object assumed to have the same spectral signature as the sun. The WFS subaperture size was 10cm for a total of 60 subapertures within the entrance pupil. The AO system model had 1.2 actuators per r_o .

Atmospheric turbulence and detector noise effects were modeled using an existing computer simulation [74]. For each data realization, the simulation created a random phase screen with the appropriate turbulence statistics [89], calculated the true OTF, and formed a detected image realization $d_i(x, y)$. At the same time, a WFS model was used to generate a phase estimate $\tilde{\phi}_i(x, y)$, which was then used to compute the estimated OTF. Finally, the required quantities $\tilde{\mathcal{H}}_i^*(u, v)D_i(u, v)$ and $|\tilde{\mathcal{H}}_i(u, v)|^2$ were accumulated and the process repeated 150 times to generate the ensemble averages. The simulation incorporates an intensity splitter which sends 40% of the photons to the image plane and 60% to the WFS. Integration times of 10ms were used for both WFS and imaging system. A range of object brightness was modeled using four visual magnitude levels $m_v = +2, +4, +6,$ and $+8$. The visual magnitude allows astronomers to compare object brightness in the night sky [57]. A step in visual magnitude indicates a corresponding factor 2.5 change in brightness, with smaller values of m_v indicating brighter objects [57]. The resulting photoevents per inte-

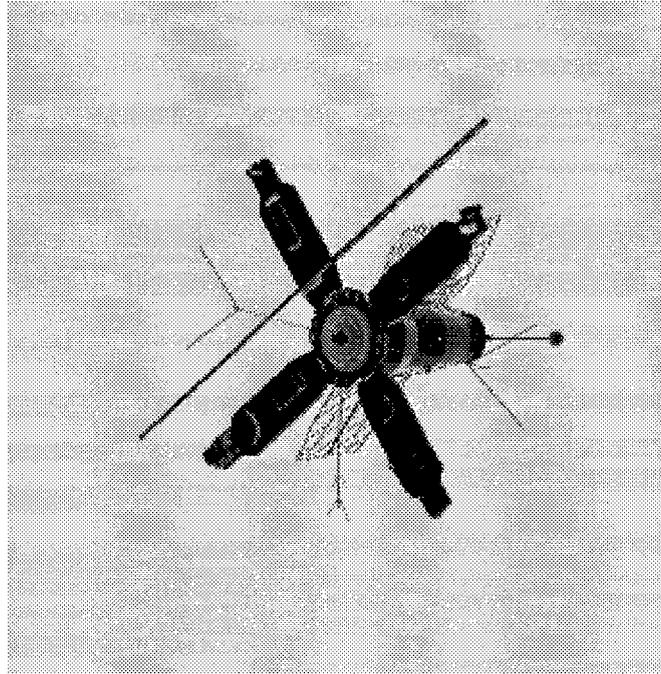


Figure 3.1 Satellite computer rendering used to test CLS algorithm performance. Negative image shown for clarity.

Table 3.1 Average Photoevents per Integration Time.

| m_ν | \overline{K}_W | \overline{K} |
|---------|------------------|----------------|
| 2 | 20,215 | 1,040,430 |
| 4 | 3,234 | 166,469 |
| 6 | 518 | 26,635 |
| 8 | 83 | 4,262 |

gration time for these cases are presented in Table 3.1, where \overline{K}_W is the average number of photoevents across an individual WFS subaperture.

3.4.2 Uncompensated Images. Figure 3.2 gives the detected short exposure image and CLS algorithm output associated with excellent seeing conditions ($r_o = 20\text{cm}$), a moderate light level satellite object ($m_\nu = +4$), and detector read noise representing a high quality CCD detector ($\sigma_r = 15$ electrons per pixel). A single short exposure image is given in (a) to illustrate shot noise and detector read noise effects. Clearly, this image is degraded by the detection process. After CLS processing of the DWFS data as shown in (b), image resolution is greatly enhanced. In this case, the CLS algorithm automatically selected the

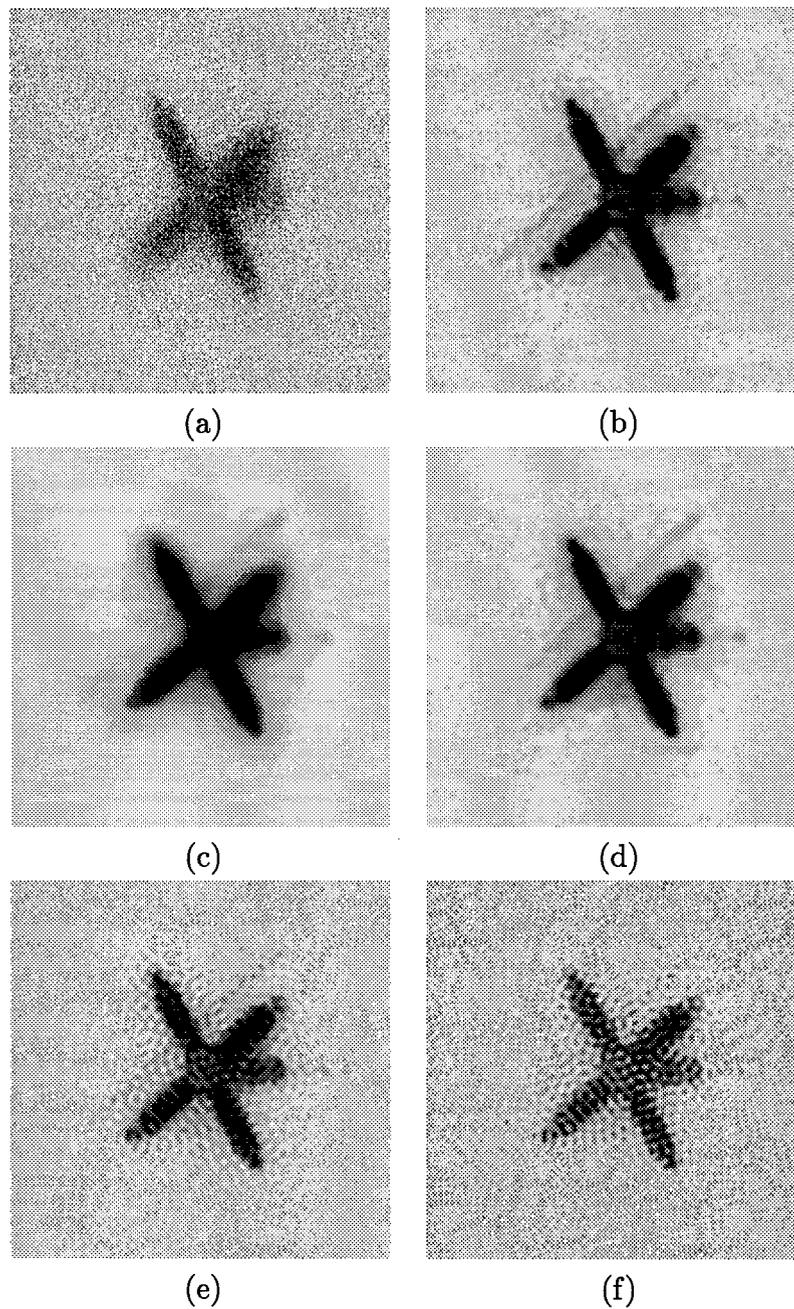


Figure 3.2 CLS algorithm comparison with manual parameter selection, $r_o = 20\text{cm}$, $m_\nu = +4$, $\sigma_r = 15$ electrons per pixel. (a) Single short exposure image. (b) CLS algorithm estimate, $\gamma = 0.0015$, 4 iterations. (c), (d), (e), and (f) manual parameter selection, $\epsilon = 0.1, 0.01, 0.0001$, and 0.00001 , respectively.

regularization parameter $\gamma = 0.0015$ in 4 iterations. In (c), (d), (e), and (f) the regularization parameter ϵ associated with the traditional DWFS estimator given in Eq. (3.3) was selected manually. For images (c) and (d), $\epsilon > \gamma$. For images (e) and (f), $\epsilon < \gamma$. Notice that the quality of the CLS image in (b) is as good or better than the manual images. Thus, Fig. 3.2 illustrates that the CLS algorithm selected a reasonable value for γ in this case.

Now consider a brighter object and different seeing conditions. Figure 3.3 gives the detected image and CLS algorithm output associated with changing turbulence strength and a brighter satellite object ($m_\nu = +2$). The CCD detector read noise remains unchanged from the previous case. Images (a) and (c) provide the detected image data for the $r_o = 10\text{cm}$ and 20cm cases, respectively. Images (b) and (d) give the corresponding CLS algorithm output. Here, $\gamma = 0.000012$ for the $r_o = 10\text{cm}$ case and $\gamma = 0.000067$ for the $r_o = 20\text{cm}$ case. The CLS algorithm provides more regularization or “smoothing” as r_o increases. This observation is consistent with the form of Eq. (3.28), where better seeing conditions lead to larger values for the quantity $\sum_{u,v} \langle |\tilde{\mathcal{H}}(u,v)|^2 \rangle$. In general, larger values for the quantity $\sum_{u,v} \langle |\tilde{\mathcal{H}}(u,v)|^2 \rangle$ lead to larger γ values and more noise smoothing.

While r_o influences CLS algorithm performance, object brightness provides a more severe limit. Not only does shot noise degrade the short exposure data, it also restricts WFS accuracy. Without a sufficiently accurate OTF estimate, DWFS performance is severely degraded. To illustrate these limitations, consider Fig. 3.4. Here, $r_o = 10\text{cm}$ and detector read noise remains unchanged from the previous cases. Images (a), (c), and (e) show the short exposure data for object brightness cases $m_\nu = +4, +6, \text{ and } +8$, respectively. In (c) and (e), noise dominates the data realization to the point that no satellite image is visible. Images (b), (d), and (f) give the corresponding CLS estimates where $\gamma = 0.00027, 0.0093, \text{ and } 0.8180$, respectively. As object brightness decreases, the output images are more blurred. This effect is consistent with a relatively large γ value and deconvolution using a poor quality estimate of the OTF.

To emphasize the limitations imposed by shot noise further, consider Fig. 3.5. Here, $r_o = 10\text{cm}$ and $m_\nu = +4$. Detector read noise variance is adjusted such that image (a) was collected with a low noise array ($\sigma_r = 10$ electrons per pixel) and image (b) with a high

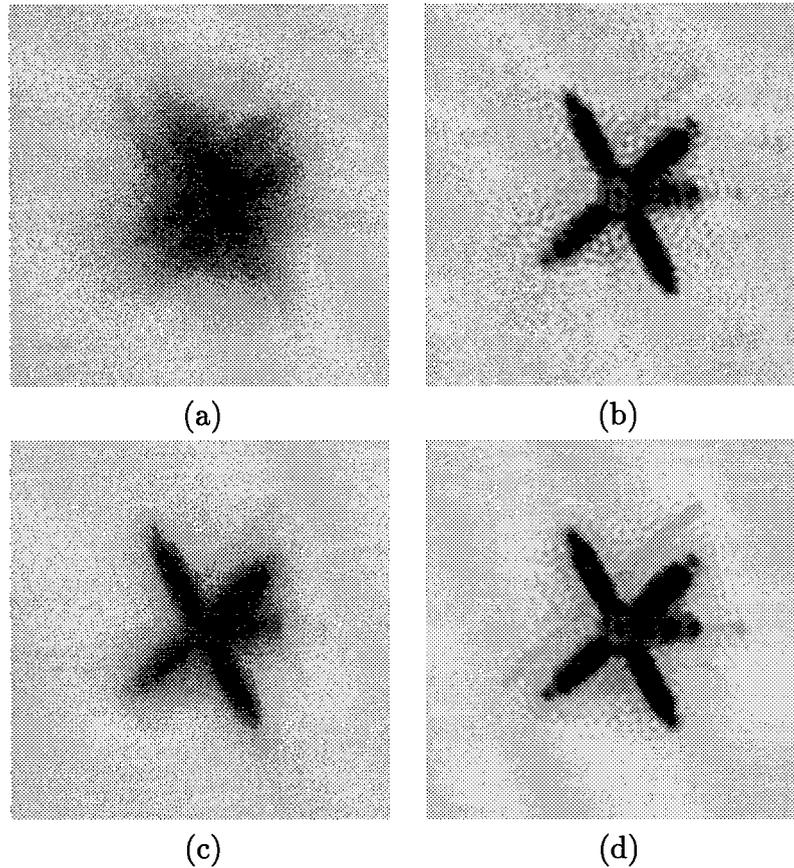


Figure 3.3 CLS algorithm output versus atmospheric turbulence strength, $m_\nu = +2$, $\sigma_r = 15$ electrons per pixel. (a) Detected image, $r_o = 10\text{cm}$. (b) CLS estimate, $r_o = 10\text{cm}$, $\gamma = 0.000012$, 4 iterations. (c) Detected image, $r_o = 20\text{cm}$. (d) CLS estimate, $r_o = 20\text{cm}$, $\gamma = 0.000067$, 4 iterations.

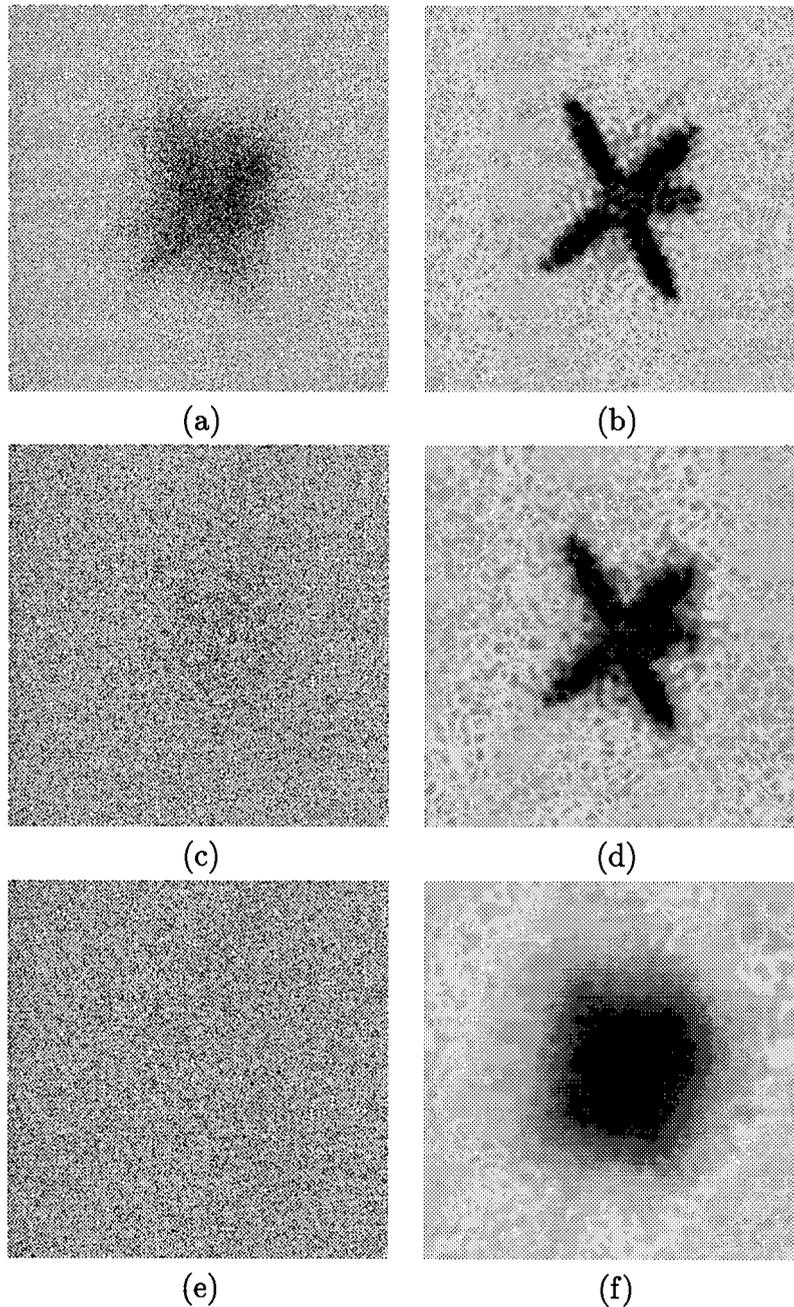


Figure 3.4 CLS algorithm output versus object brightness, $r_o = 10cm$, $\sigma_r = 15$ electrons per pixel. (a) Detected image, $m_\nu = +4$. (b) CLS estimate, $m_\nu = +4$, $\gamma = 0.00027$, 4 iterations. (c) Detected image, $m_\nu = +6$. (d) CLS estimate, $m_\nu = +6$, $\gamma = 0.0093$, 5 iterations. (e) Detected image, $m_\nu = +8$. (f) CLS estimate, $m_\nu = +8$, $\gamma = 0.8180$, 7 iterations.

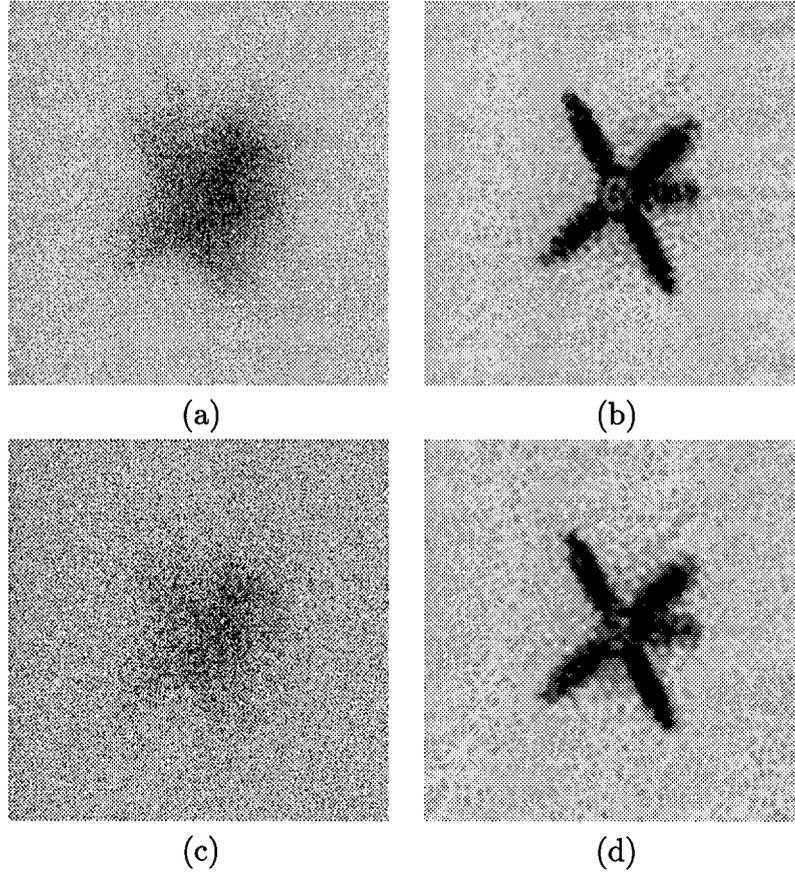


Figure 3.5 CLS algorithm output versus detector read noise strength, $r_o = 10\text{cm}$, $m_\nu = +4$. (a) Detected image, $\sigma_r = 10$ electrons per pixel. (b) CLS estimate, $\sigma_r = 10$ electrons per pixel, $\gamma = 0.00014$, 4 iterations. (c) Detected image, $\sigma_r = 30$ electrons per pixel. (d) CLS estimate, $\sigma_r = 30$ electrons per pixel, $\gamma = 0.00097$, 4 iterations.

noise array ($\sigma_r = 30$ electrons per pixel). Images (b) and (d) give the corresponding CLS estimates where $\gamma = 0.00014$ and 0.00097 , respectively. Notice that an increase in read noise strength does not have the drastic effect on image quality observed in Fig. 3.4. Detector read noise does not affect the accuracy of the OTF estimate provided by the WFS. Thus, the effect on algorithm output is less pronounced.

3.4.3 Adaptive Optics Compensated Images. DWFS processing can also be applied to AO compensated images. Here, the algorithm deconvolves atmospheric turbulence effects associated with the residual error between the true phase perturbation and the phase imposed by the deformable mirror (DM). As noted earlier, the ratio between DM actuator number

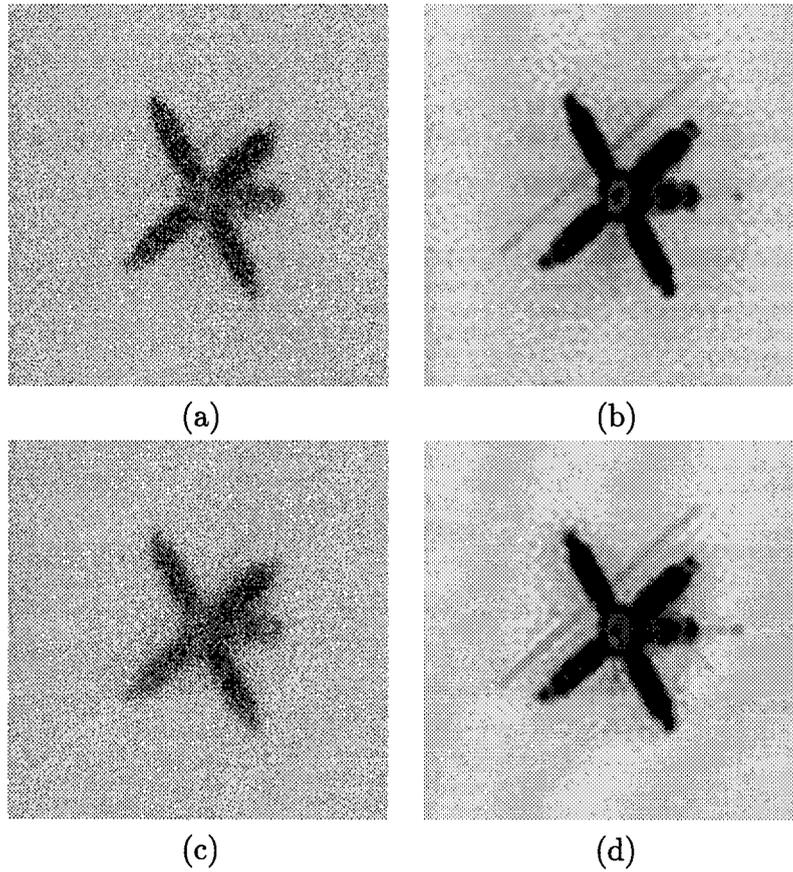


Figure 3.6 CLS algorithm output for AO compensated images, $m_\nu = +4$, $\sigma_r = 15$ electrons per pixel. (a) Detected image, $r_o = 10\text{cm}$. (b) CLS estimate, $r_o = 10\text{cm}$, $\gamma = 0.0048$, 4 iterations. (c) Detected image, $r_o = 20\text{cm}$. (d) CLS estimate, $r_o = 20\text{cm}$, $\gamma = 0.0039$, 4 iterations.

and r_o was fixed at 1.2. Thus, the number of actuators will change as r_o is adjusted. Two AO systems associated with $r_o = 10\text{cm}$ (57 DM actuators) and $r_o = 20\text{cm}$ (13 DM actuators) are modeled. Figure 3.6 illustrates CLS algorithm performance on AO compensated images when $m_\nu = +4$ and detector read noise is $\sigma_r = 15$ electrons per pixel. Images (a) and (c) give examples of raw AO compensated data when $r_o = 10\text{cm}$ and 20cm , respectively. Images (b) and (d) show the corresponding CLS estimates, where $\gamma = 0.0048$ in (b) and $\gamma = 0.0039$ in (d). The CLS algorithm does a good job of noise suppression for AO compensated images for the same reason noted above in the discussion associated with Fig. 3.3. Here, AO compensation leads to larger values for the quantity $\sum_{u,v} \langle |\tilde{\mathcal{H}}(u,v)|^2 \rangle$ in Eq. (3.28) and larger γ values.

3.5 Summary

A CLS estimator that incorporates noisy DWFS data, noise statistics, and OTF statistics was investigated. For a particular choice of Fourier domain constraint, the estimator selects the regularization parameter automatically. No ad hoc tuning is necessary to reduce high spatial frequency noise effects in the DWFS image. The CLS estimator uses a Newton-Raphson iteration to select a Lagrange multiplier which minimizes an objective function. The objective function uses ensemble-averaged data directly, which aids in noise suppression. The sample results show that the new algorithm produces DWFS images comparable in quality to manual regularization with minimal computational expense. While turbulence and detector read noise strength impact algorithm performance, shot noise imposes a fundamental limit on the deconvolution process. Finally, the CLS estimator was derived to incorporate a general Fourier domain constraint. Thus, other constraint functions can be used based on the specific application.

In the next chapter, a non-iterative Bayesian deconvolution filter is derived. Unlike this CLS approach, the vector Wiener filter incorporates object, OTF, and noise correlations between *different* spatial frequencies. Chapters V and VI will further explore the value of this model-based information for several applications related to astronomical imaging.

IV. Vector Wiener Filter Analysis

4.1 Introduction

In this chapter, a Fourier domain vector Wiener filter is derived which incorporates complete object, blur, and noise correlation statistics. The derivation extends the original work by Pratt [63] to account properly for a random optical transfer function (OTF) and measurement noise. This analysis is consistent with related research which showed that shot noise is correlated with respect to spatial frequency [54]. The amount of correlation depends on the product of the mean OTF and the mean object spectrum at a difference frequency. This linear filter can provide a useful alternative to nonlinear iterative techniques when appropriate statistical models are available.

Chapter IV is organized as follows. Section 4.2 outlines the complete derivation of the new filter for both random and deterministic OTF cases. Section 4.3 gives a signal-to-noise ratio (SNR) interpretation to the estimator. In Section 4.4, theoretical filter mean square error (MSE) expressions are derived for the vector Wiener filter, scalar Wiener filter, and the unfiltered data. These expressions are important for demonstration and analysis in Chapters V and VI. Section 4.5 provides alternate filter expressions which incorporate the mean and covariance of the random quantities. Finally, some comments are made in Section 4.6 regarding the optimality of the new filter expressions. The chapter ends with a brief summary in Section 4.7.

4.2 Fourier Domain Filter Derivation

In Chapter II, the potential performance advantages associated with the vector Wiener filter were noted for non-stationary image ensembles. However, short exposure images collected through atmospheric turbulence are blurred by an unknown random OTF. In addition, measurement noise further degrades optical system resolution. Equation (2.36) does not account for these factors and must be extended for this imaging application.

4.2.1 Wiener-Hopf Expression. Let us consider an $N \times N$ image array. Based on the Fourier domain vector-matrix image model expression given in Eq. (2.9), the objective is to find an $N^2 \times N^2$ Fourier domain filter matrix M_R such that

$$\hat{\mathbf{O}} = M_R \mathbf{D}, \quad (4.1)$$

where the MSE

$$\varepsilon^2 = \text{E} [(\mathbf{O} - \hat{\mathbf{O}})^H (\mathbf{O} - \hat{\mathbf{O}})] \quad (4.2)$$

is minimized. Using the matrix trace operator $\text{Tr}\{\bullet\}$ [35], Eq. (4.2) can be written as

$$\varepsilon^2 = \text{E} [\text{Tr} \{(\mathbf{O} - \hat{\mathbf{O}})(\mathbf{O} - \hat{\mathbf{O}})^H\}]. \quad (4.3)$$

Substituting Eq. (4.1) into Eq. (4.3), bringing the expectation operator inside the trace operator, and expanding the expression yields

$$\varepsilon^2 = \text{Tr} \{R_{OO} - M_R R_{OD}^H - R_{OD} M_R^H + M_R R_{DD} M_R^H\}, \quad (4.4)$$

where $R_{OO} = \text{E}[\mathbf{O}\mathbf{O}^H]$ is the object Fourier domain autocorrelation matrix, $R_{OD} = \text{E}[\mathbf{O}\mathbf{D}^H]$ is the object-detected image Fourier domain cross-correlation matrix, and $R_{DD} = \text{E}[\mathbf{D}\mathbf{D}^H]$ is the detected image Fourier domain autocorrelation matrix. To find the filter transformation matrix M_R that minimizes ε^2 , take the derivative of Eq. (4.4) with respect to M_R , set this derivative equal to the zero matrix, and solve for M_R . The resultant derivative is [35]

$$\frac{\partial \varepsilon^2}{\partial M_R} = -2R_{OD} + 2R_{DD} M_R^H = 0. \quad (4.5)$$

Thus, the linear minimum MSE Fourier spectrum estimate is

$$\hat{\mathbf{O}} = R_{OD}(R_{DD})^{-1} \mathbf{D}, \quad (4.6)$$

where the transformation $M_R = R_{OD}(R_{DD})^{-1}$ satisfies a Wiener-Hopf equation [43].

4.2.2 *Object-Detected Image Cross-correlation.* To write the new vector Wiener filter expression, the Fourier domain correlation matrices R_{OD} and R_{DD} must be derived using the image model given in Eq. (2.7). First, consider the cross-correlation between the object spatial frequency (u, v) and detected image spatial frequency (u', v') denoted

$$R_{OD}(u, v; u', v') = \mathbb{E}[O(u, v)D^*(u', v')]. \quad (4.7)$$

Substituting Eq. (2.7) into Eq. (4.7) and writing as the sum of two terms yields

$$\begin{aligned} R_{OD}(u, v; u', v') &= \mathbb{E} \left[O(u, v) \sum_{n=1}^K \exp \{j2\pi(u'x_n + v'y_n)\} \right] \\ &+ \mathbb{E} \left[O(u, v) \sum_{p=1}^P n_p \exp \{j2\pi(u'x_p + v'y_p)\} \right]. \end{aligned} \quad (4.8)$$

The second term in Eq. (4.8) is zero because n_p is both independent of O and zero mean. The first term can be evaluated using nested conditional expectations following the technique presented in Refs. [23] and [74]. The random quantities are the object spectrum, the OTF, the total number of photoevents K , and the photon arrival location (x_n, y_n) . Bayes rule [62] can be used to rewrite the joint probability density function (PDF) in the first term using conditional PDFs. These conditional PDFs translate to conditional expectations such that

$$\begin{aligned} R_{OD}(u, v; u', v') &= \mathbb{E}_O \left[\mathbb{E}_{\mathcal{H}|O} \left(\mathbb{E}_{K|\mathcal{H},O} \left\{ \mathbb{E}_{x_n, y_n|K, \mathcal{H}, O} \left[O(u, v) \right. \right. \right. \right. \\ &\quad \left. \left. \left. \times \sum_{n=1}^K \exp \{j2\pi(u'x_n + v'y_n)\} \right] \right\} \right) \right], \end{aligned} \quad (4.9)$$

where the notation $\mathbb{E}_{A|B}$ denotes the conditional expected value of the random event A given B . An expression similar to the innermost conditional expectation of Eq. (4.9) has been evaluated previously [74]. The derivation details in this case are given in Appendix A.1 with the final result written as

$$\mathbb{E}_{x_n, y_n|K, \mathcal{H}, O} \left[O(u, v) \sum_{n=1}^K \exp \{j2\pi(u'x_n + v'y_n)\} \right] = K\mathcal{H}^*(u', v')O(u, v)O_n^*(u', v'), \quad (4.10)$$

where the normalized object Fourier spectrum $O_n(u, v)$ is defined as

$$O_n(u, v) = \frac{O(u, v)}{O(0, 0)} = \frac{O(u, v)}{\bar{K}}. \quad (4.11)$$

Evaluation of the remaining nested expectations is trivial since K , \mathcal{H} , and O are all mutually independent. This mutual independence converts all the nested quantities into an uncoupled product of expectations, which gives the following expression for Eq. (4.7)

$$R_{OD}(u, v; u', v') = \bar{K} \bar{\mathcal{H}}^*(u', v') \text{E} [O(u, v) O_n^*(u', v')], \quad (4.12)$$

where $\bar{\mathcal{H}}(u, v)$ is the mean OTF. Clearly, the expectation above is the object Fourier domain autocorrelation, except for a normalization factor associated with $O(u, v)$. Thus, Eq. (4.11) can be used to write Eq. (4.12) as

$$\bar{K} \bar{\mathcal{H}}^*(u', v') \text{E} [O(u, v) O_n^*(u', v')] = (\bar{K})^2 \bar{\mathcal{H}}^*(u', v') \text{E} [O_n(u, v) O_n^*(u', v')]. \quad (4.13)$$

Finally, Eq. (4.7) becomes

$$R_{OD}(u, v; u', v') = (\bar{K})^2 \bar{\mathcal{H}}^*(u', v') R_{O_n O_n}(u, v; u', v'), \quad (4.14)$$

where $R_{O_n O_n}(u, v; u', v')$ is the autocorrelation between the (u, v) and the (u', v') spatial frequencies of the normalized object Fourier spectrum. Equation (4.14) can be expressed using the vector-matrix notation introduced in Chapter II. Note that the functional dependence of the OTF on the spatial frequency (u', v') is equivalent to multiplying a particular column of the matrix $R_{O_n O_n}$ by $\bar{\mathcal{H}}^*(u', v')$. This process is equivalent to multiplying by the diagonal matrix $\bar{\mathcal{H}}_d^*$ such that the final result for the object-detected image Fourier domain correlation matrix is

$$R_{OD} = (\bar{K})^2 R_{O_n O_n} \bar{\mathcal{H}}_d^*. \quad (4.15)$$

4.2.3 *Detected Image Autocorrelation.* To complete the derivation, the correlation between detected image spatial frequencies (u, v) and (u', v') denoted

$$R_{DD}(u, v; u', v') = E[D(u, v)D^*(u', v')], \quad (4.16)$$

must be derived. Substituting the image model given in Eq. (2.7) into Eq. (4.16) and expanding the expression within the expectation yields

$$\begin{aligned} R_{DD}(u, v; u', v') &= E \left[\sum_{n=1}^K \sum_{m=1}^K \exp \{-j2\pi(ux_n - u'x_m + vy_n - v'y_m)\} \right. \\ &\quad \left. + \sum_{p=1}^P \sum_{q=1}^P n_p n_q \exp \{-j2\pi(ux_p - u'x_q + vy_p - v'y_q)\} \right] \\ &= R_{DD}^{(1)}(u, v; u', v') + R_{DD}^{(2)}(u, v; u', v'), \end{aligned} \quad (4.17)$$

where the cross terms are zero because n_p is zero mean and independent of K , \mathcal{H} , and \mathcal{O} . Consider the first term, $R_{DD}^{(1)}(u, v; u', v')$, and write using nested conditional expectations [23, 74] as before which yields

$$\begin{aligned} R_{DD}^{(1)}(u, v; u', v') &= E_{\mathcal{O}} \left[E_{\mathcal{H}|\mathcal{O}} \left(E_{K|\mathcal{H},\mathcal{O}} \left\{ E_{x_n, y_n, x_m, y_m | K, \mathcal{H}, \mathcal{O}} \right. \right. \right. \\ &\quad \left. \left. \left. \times \left[\sum_{n=1}^K \sum_{m=1}^K \exp \{-j2\pi(ux_n - u'x_m + vy_n - v'y_m)\} \right] \right\} \right) \right]. \end{aligned} \quad (4.18)$$

The double summation in $E_{x_n, y_n, x_m, y_m | K, \mathcal{H}, \mathcal{O}}$ has two types of terms: K terms in which $n = m$ and $K^2 - K$ terms in which $n \neq m$. Thus, $E_{x_n, y_n, x_m, y_m | K, \mathcal{H}, \mathcal{O}}$ in Eq. (4.18) can be rewritten in a different form:

$$\begin{aligned} &E_{x_n, y_n, x_m, y_m | K, \mathcal{H}, \mathcal{O}} \left[\sum_{n=1}^K \sum_{m=1}^K \exp \{-j2\pi(ux_n - u'x_m + vy_n - v'y_m)\} \right] = \\ &E_{x_n, y_n | K, \mathcal{H}, \mathcal{O}} \left[\sum_{n=1}^K \exp \{-j2\pi((u - u')x_n + (v - v')y_n)\} \right]_{n=m} \\ &+ E_{x_n, y_n, x_m, y_m | K, \mathcal{H}, \mathcal{O}} \left[\sum_{n=1}^K \sum_{m=1}^K \exp \{-j2\pi(ux_n - u'x_m + vy_n - v'y_m)\} \right]_{n \neq m}. \end{aligned} \quad (4.19)$$

The derivation details associated with Eq. (4.19) are given in Appendix A.2 with the final result written as

$$\begin{aligned} E_{x_n, y_n, x_m, y_m | K, \mathcal{H}, O} [\bullet] &= (K^2 - K) O_n(u, v) O_n^*(u', v') \mathcal{H}(u, v) \mathcal{H}^*(u', v') \\ &+ K \mathcal{H}(u - u', v - v') O_n(u - u', v - v'), \end{aligned} \quad (4.20)$$

where the \bullet notation represents the bracketed quantity on the left side of Eq. (4.19). With Eq. (4.20) in hand, the remaining conditional expectations can be evaluated as before by noting the mutual independence of K , \mathcal{H} , and O , which leads to

$$\begin{aligned} R_{DD}^{(1)}(u, v; u', v') &= (\overline{K^2} - \overline{K}) R_{O_n O_n}(u, v; u', v') R_{\mathcal{H}\mathcal{H}}(u, v; u', v') \\ &+ \overline{K} \overline{\mathcal{H}}(u - u', v - v') \overline{O}_n(u - u', v - v'), \end{aligned} \quad (4.21)$$

where $R_{\mathcal{H}\mathcal{H}}(u, v; u', v')$ is the autocorrelation between (u, v) and (u', v') spatial frequencies of the OTF and $\overline{O}_n(u, v)$ denotes the normalized mean object spectrum. The random variable K is conditionally Poisson distributed, given \mathcal{H} and O [74]. Therefore, the second moment of K can be written as [62, 74]

$$E[K^2] = \overline{K^2} = \overline{K} + (\overline{K})^2. \quad (4.22)$$

Using Eq. (4.22), Eq. (4.21) can be written in its final form:

$$\begin{aligned} R_{DD}^{(1)}(u, v; u', v') &= (\overline{K})^2 R_{O_n O_n}(u, v; u', v') R_{\mathcal{H}\mathcal{H}}(u, v; u', v') \\ &+ \overline{K} \overline{\mathcal{H}}(u - u', v - v') \overline{O}_n(u - u', v - v'). \end{aligned} \quad (4.23)$$

The second term in Eq. (4.17) was evaluated previously [74] for the assumptions given in Chapter II such that

$$R_{DD}^{(2)}(u, v; u', v') = P \sigma_r^2 \delta(u - u', v - v'), \quad (4.24)$$

where $P = N^2$ is the number of pixels in the detector. Combining the results from Eqs. (4.23) and (4.24) gives the final expression for $R_{DD}(u, v; u', v')$:

$$\begin{aligned} R_{DD}(u, v; u', v') &= (\bar{K})^2 R_{O_n O_n}(u, v; u', v') R_{\mathcal{H}\mathcal{H}}(u, v; u', v') \\ &\quad + \bar{K} \bar{\mathcal{H}}(u - u', v - v') \bar{O}_n(u - u', v - v') \\ &\quad + P \sigma_r^2 \delta(u - u', v - v'). \end{aligned} \quad (4.25)$$

Equation (4.25) can be written conveniently using vector-matrix notation as before, which yields a compact expression for the matrix R_{DD}

$$R_{DD} = (\bar{K})^2 R_{O_n O_n} \odot R_{\mathcal{H}\mathcal{H}} + R_{NN}, \quad (4.26)$$

where the measurement noise Fourier domain autocorrelation between (u, v) and (u', v') spatial frequencies is defined as

$$R_{NN}(u, v; u', v') = \bar{K} \bar{\mathcal{H}}(u - u', v - v') \bar{O}_n(u - u', v - v') + P \sigma_r^2 \delta(u - u', v - v'). \quad (4.27)$$

4.2.4 Final Result. Returning to Eq. (4.6), inserting the results given in Eqs. (4.15) and (4.26), and dividing through by the scalar $(\bar{K})^2$ give the final expression for the extended vector Wiener filter [14]:

$$\hat{\mathbf{O}} = R_{O_n O_n} \bar{\mathcal{H}}_d^* \{ R_{O_n O_n} \odot R_{\mathcal{H}\mathcal{H}} + \bar{K}^{-2} R_{NN} \}^{-1} \mathbf{D}, \quad (4.28)$$

where the elements of the matrix R_{NN} are defined in Eq. (4.27). Equation (4.28) is the main result of this section and the subject of discussion and experimentation in Chapters V and VI. Equation (4.28) is new and has only recently been applied to images degraded by a random OTF and measurement noise [12–14]. If the OTF is deterministic and known, Eq. (4.28) can be written in the following form:

$$\hat{\mathbf{O}} = R_{O_n O_n} \mathcal{H}_d^* \{ \mathcal{H}_d R_{O_n O_n} \mathcal{H}_d^* + \bar{K}^{-2} R_{NN} \}^{-1} \mathbf{D}, \quad (4.29)$$

where the deterministic OTF is substituted for the mean OTF when defining the Fourier domain noise autocorrelation matrix R_{NN} .

4.3 Signal-to-Noise Ratio Interpretation

It is convenient to express the measurement noise Fourier domain correlation in terms of SNR quantities. Multiplying R_{NN} by the scalar \bar{K}^{-2} as required by Eq. (4.28) yields

$$\begin{aligned} \bar{K}^{-2} R_{NN}(u, v; u', v') &= \bar{K}^{-1} \bar{\mathcal{H}}(u - u', v - v') \bar{O}_n(u - u', v - v') \\ &+ [P\sigma_r^2 / (\bar{K})^2] \delta(u - u', v - v'). \end{aligned} \quad (4.30)$$

The single frame image spectrum SNR is defined as [74]

$$\text{SNR}_1(u, v) = \frac{\text{E}[|D(u, v)|]}{\{\text{Var}[D(u, v)]\}^{1/2}}, \quad (4.31)$$

and was calculated previously for an image degraded by measurement noise, yielding [74]

$$\text{SNR}_1(u, v) = \frac{|\bar{\mathcal{H}}(u, v)| |O_n(u, v)|}{\{|O_n(u, v)|^2 \text{Var}[\bar{\mathcal{H}}(u, v)] + \bar{K}^{-1} + P\sigma_r^2 / (\bar{K})^2\}^{1/2}}, \quad (4.32)$$

where $\text{Var}[\bullet]$ denotes the variance of the bracketed expression. The last two quantities in the denominator of Eq. (4.32) are related to the shot noise and the signal-independent detector read noise, respectively. Consider the following SNR expressions [90]

$$\text{SNR}_k = \sqrt{\bar{K}}, \quad (4.33)$$

$$\text{SNR}_r = \frac{\bar{K}}{\sqrt{P\sigma_r^2}}, \quad (4.34)$$

where SNR_k is associated with the shot noise and SNR_r is associated with the detector read noise. Equations (4.33) and (4.34) were used previously to rewrite the single frame image spectrum SNR [90]. In the same way, the measurement noise Fourier domain correlation in

Eq. (4.28) can be written as

$$\begin{aligned} \overline{K}^{-2} R_{NN}(u, v; u', v') &= \text{SNR}_k^{-2} \overline{\mathcal{H}}(u - u', v - v') \overline{O}_n(u - u', v - v') \\ &+ \text{SNR}_r^{-2} \delta(u - u', v - v'). \end{aligned} \quad (4.35)$$

These SNR quantities control the amount of regularization provided to each spatial frequency correlation component by the vector Wiener filter. When the measurement noise is low (i.e., high light level and minimal detector read noise), negligible regularization occurs and the filter is faithful to the data. When the measurement noise is high (i.e., low light level or significant detector read noise), regularization is applied based on the spatial frequency correlation of the shot noise and the strength of the uncorrelated detector read noise. Under extremely high noise conditions, the vector Wiener filter will smooth the data, but always consistent with the known object Fourier domain statistical model. In the next section, MSE expressions are introduced which will be important for filter performance comparison in Chapters V and VI.

4.4 Mean Square Error Expressions

The vector Wiener filter given in Eq. (4.28) minimizes the scalar MSE as given in Eq. (4.2). Thus, filter performance can be analyzed by examining the statistics of the zero mean error between the true and estimated Fourier spectra $\epsilon = \mathbf{O} - \hat{\mathbf{O}}$. The error correlation matrix $R_{\epsilon\epsilon}$ associated with a Fourier domain linear minimum MSE filter can be written as [43]

$$R_{\epsilon\epsilon} = \text{E} [(\mathbf{O} - \hat{\mathbf{O}})(\mathbf{O} - \hat{\mathbf{O}})^H]. \quad (4.36)$$

In Eq. (4.1), a linear estimator model was considered which incorporated a specific vector Wiener filter transformation matrix M_R . Here, the same linear model will be used but with an arbitrary transformation matrix M_X such that Eq. (4.1) becomes

$$\hat{\mathbf{O}} = M_X \mathbf{D}. \quad (4.37)$$

Substituting Eq. (4.37) for $\hat{\mathbf{O}}$, expanding terms, and evaluating the expectation yields [14,43]

$$R_{\epsilon\epsilon} = R_{OO} - M_X \bar{\mathcal{H}}_d R_{OO} - R_{OO} \bar{\mathcal{H}}_d^* M_X^H + M_X (R_{OO} \odot R_{\mathcal{H}\mathcal{H}} + R_{NN}) M_X^H. \quad (4.38)$$

For the vector Wiener filter transformation M_R given in Eq. (4.28), Eq. (4.38) reduces to

$$R_{\epsilon\epsilon} = R_{OO} - M_R \bar{\mathcal{H}}_d R_{OO} \quad (\text{Vector Wiener Filter}). \quad (4.39)$$

A similar expression is also available for the analogous scalar Wiener filter whose diagonal transformation matrix M_S has non-zero elements defined as

$$M_S(u, v) = \frac{\bar{\mathcal{H}}^*(u, v)}{\text{E}[|\mathcal{H}(u, v)|^2] + (\bar{K} + P\sigma_r^2) / \mathcal{G}_o(u, v)}, \quad (4.40)$$

where $\mathcal{G}_o(u, v)$ is the PSD of the object as used in Eq. (2.35). The error correlation matrix associated with the scalar Wiener filter is

$$R_{\epsilon\epsilon} = R_{OO} - M_S \bar{\mathcal{H}}_d R_{OO} - R_{OO} \bar{\mathcal{H}}_d^* M_S^* + M_S (R_{OO} \odot R_{\mathcal{H}\mathcal{H}} + R_{NN}) M_S^* \quad (\text{Scalar Wiener Filter}). \quad (4.41)$$

Finally, it is often of interest to know the baseline error statistics associated with the unfiltered detected image. When $M_X = I$, where I is the identity matrix, Eq. (4.37) gives $\hat{\mathbf{O}} = \mathbf{D}$ as required and the error correlation matrix becomes

$$R_{\epsilon\epsilon} = R_{OO} - \bar{\mathcal{H}}_d R_{OO} - R_{OO} \bar{\mathcal{H}}_d^* + (R_{OO} \odot R_{\mathcal{H}\mathcal{H}} + R_{NN}) \quad (\text{Detected Image}). \quad (4.42)$$

Equations (4.39), (4.41), and (4.42) are only dependent on the statistics of the object class, OTF, and noise. Thus, a theoretical performance study based on these error statistics is possible without running a Monte Carlo simulation. In the next section, two alternate vector Wiener filter expressions are derived which incorporate mean and covariance statistics.

4.5 Alternate Filter Expressions

The vector Wiener filter given in Eqs. (4.28) and (4.29) is a function of correlation matrices. In many statistical signal processing and image processing texts, Bayesian linear minimum MSE filters are expressed in terms of mean and covariance [41, 43]. Thus, two complementary vector Wiener filter equations are given below in the interest of completeness.

Let us now reconsider the form of the random object realization \mathbf{O} . Such a random vector can always be written as a function of its mean $\bar{\mathbf{O}}$ and a random, zero mean component $\Delta\mathbf{O}$ yielding

$$\mathbf{O} = \bar{\mathbf{O}} + \Delta\mathbf{O}, \quad (4.43)$$

where $\Delta\mathbf{O}$ is zero mean with covariance matrix C_{OO} . In general, a complex covariance matrix is defined as [62]

$$C_{XX} = E [(\mathbf{X} - \bar{\mathbf{X}})(\mathbf{X} - \bar{\mathbf{X}})^H], \quad (4.44)$$

where \mathbf{X} is a complex random vector with mean $\bar{\mathbf{X}}$. The object estimate vector can now be written as

$$\begin{aligned} \hat{\mathbf{O}} &= \bar{\mathbf{O}} + \Delta\hat{\mathbf{O}} \\ &= \bar{\mathbf{O}} + M_C \Delta\mathbf{D}, \end{aligned} \quad (4.45)$$

where $\Delta\mathbf{D} = \mathbf{D} - \bar{\mathbf{D}}$, $\bar{\mathbf{D}}$ is the mean detected image vector, and M_C is a new filter transformation matrix. Following analysis similar to that given in Eqs. (4.1)-(4.5) produces a Wiener-Hopf expression incorporating covariance matrices

$$\Delta\hat{\mathbf{O}} = C_{OD} (C_{DD})^{-1} \Delta\mathbf{D}, \quad (4.46)$$

where C_{OD} is the cross-covariance between object and detected image and C_{DD} is the autocovariance of the detected image. The analysis in Section 4.2 will now be used to find expressions for C_{OD} and C_{DD} .

By definition, the matrix C_{OD} can be written as [62]

$$C_{OD} = R_{OD} - \bar{\mathbf{O}} \bar{\mathbf{D}}^H. \quad (4.47)$$

The matrix R_{OD} was derived above and is given in Eq. (4.15). Previous analysis has shown that the mean detected image vector is $\bar{\mathbf{D}} = \bar{K} \bar{\mathcal{H}}_d \bar{\mathbf{O}}_n$ [74]. Substituting Eq. (4.15) and the mean detected image Fourier spectrum into Eq. (4.47) yields

$$\begin{aligned} C_{OD} &= (\bar{K})^2 R_{O_n O_n} \bar{\mathcal{H}}_d^* - (\bar{K})^2 \bar{\mathbf{O}}_n \bar{\mathbf{O}}_n^H \bar{\mathcal{H}}_d^* \\ &= (\bar{K})^2 (R_{O_n O_n} - \bar{\mathbf{O}}_n \bar{\mathbf{O}}_n^H) \bar{\mathcal{H}}_d^* \\ &= (\bar{K})^2 C_{O_n O_n} \bar{\mathcal{H}}_d^*. \end{aligned} \quad (4.48)$$

The matrix C_{DD} can also be written in a form which incorporates R_{DD} and $\bar{\mathbf{D}}$ such that [62]

$$C_{DD} = R_{DD} - \bar{\mathbf{D}} \bar{\mathbf{D}}^H. \quad (4.49)$$

R_{DD} was derived above and is given in Eq. (4.26). Substituting Eq. (4.26) and the mean detected image Fourier spectrum into Eq. (4.49) yields

$$C_{DD} = (\bar{K})^2 R_{O_n O_n} \odot R_{\mathcal{H}\mathcal{H}} + R_{NN} - (\bar{K})^2 (\bar{\mathbf{O}}_n \bar{\mathbf{O}}_n^H) \odot (\bar{\mathcal{H}} \bar{\mathcal{H}}^H), \quad (4.50)$$

where $\bar{\mathcal{H}}$ denotes a P -length vector containing the mean OTF elements as noted in Chapter II. Equation (4.50) can be rewritten in terms of covariance matrices using the general form shown in Eqs. (4.47) and (4.49). Combining Eq. (4.48) with the modified form of Eq. (4.50) gives an alternate filter expression incorporating covariance matrices

$$\begin{aligned} \hat{\mathbf{O}} &= \bar{K} \bar{\mathbf{O}}_n + C_{O_n O_n} \bar{\mathcal{H}}_d^* \\ &\times (C_{O_n O_n} \odot C_{\mathcal{H}\mathcal{H}} + C_{\mathcal{H}\mathcal{H}} \odot (\bar{\mathbf{O}}_n \bar{\mathbf{O}}_n^H) + C_{O_n O_n} \odot (\bar{\mathcal{H}} \bar{\mathcal{H}}^H) + \bar{K}^{-2} C_{NN})^{-1} \\ &\times (\mathbf{D} - \bar{K} \bar{\mathcal{H}}_d \bar{\mathbf{O}}_n), \end{aligned} \quad (4.51)$$

where $C_{NN} = R_{NN}$ since the measurement noise is zero mean and R_{NN} is defined as given in Eq. (4.27). When the OTF is deterministic and known, Eq. (4.51) becomes

$$\hat{\mathbf{O}} = \bar{K} \bar{\mathbf{O}}_n + C_{O_n O_n} \mathcal{H}_d^* \left(\mathcal{H}_d C_{O_n O_n} \mathcal{H}_d^* + \bar{K}^{-2} C_{NN} \right)^{-1} \left(\mathbf{D} - \bar{K} \mathcal{H}_d \bar{\mathbf{O}}_n \right). \quad (4.52)$$

The next section provides some brief comments related to the optimality of the new vector Wiener filter. Here, the linear constraint will be investigated with respect to the underlying PDFs associated with object and noise.

4.6 Comments on Filter Optimality

The vector Wiener filter is optimal with respect to MSE or error variance, regardless of the stationarity of the underlying random processes. The scalar Wiener filter is only optimal in this sense when object and noise are stationary. Here, an “optimal” filter is one that provides the minimum error variance for a particular imaging scenario and filter class. The filter class referred to here is the class of Bayesian *linear* filters only. A nonlinear estimator may exist that provides an ensemble of solutions with less error variance than the vector Wiener filter as illustrated in Fig. 4.1. In the context of this figure, the vector Wiener filter is guaranteed to provide minimum error variance among the class of linear Bayesian filters represented by the small circle marked “Linear”, not the larger circle marked “General”.

As noted in Chapter II, linear minimum MSE estimators have been extensively studied and applied to a wide variety of signal processing problems [43]. This body of knowledge reveals one important case in which the linear estimator is optimal with respect to MSE for the class of all Bayesian filters. This situation exists when the noise and object are Gaussian distributed and statistically independent [43]. However, this case does not apply to the image model given in Eq. (2.6) since the shot noise is not independent of the signal. The rest of this section will investigate the effect of the semi-classical model on the optimality of the vector Wiener filter. Section 4.6.1 will review key background details related to the linear Bayesian estimator. Section 4.6.2 presents the signal-independent noise case. Finally, Section 4.6.3 outlines optimality issues related to the semi-classical image model.

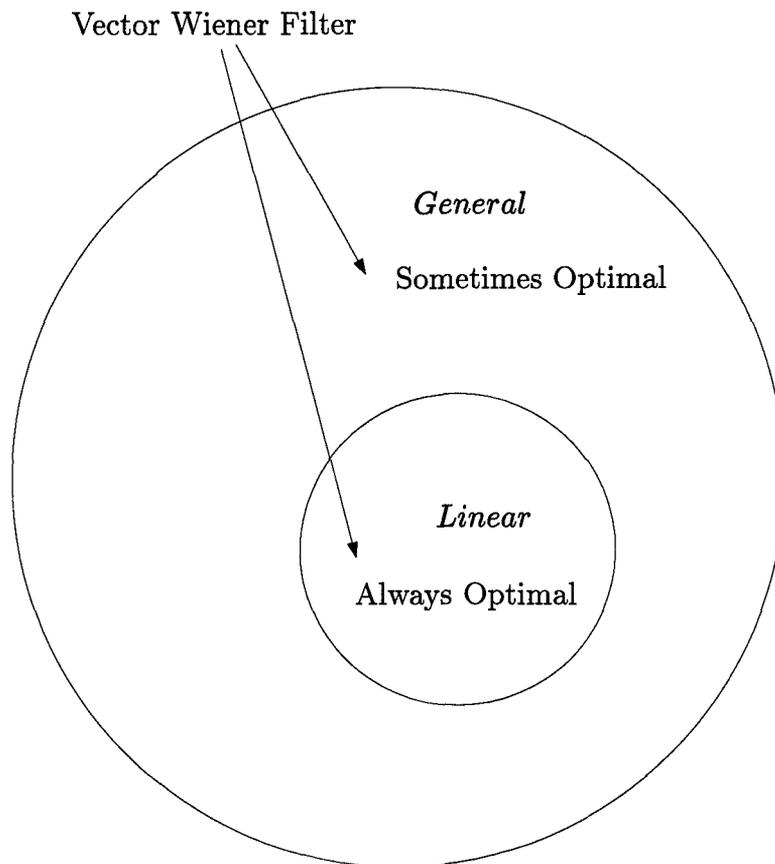


Figure 4.1 The vector Wiener filter is optimal with respect to minimum error variance only among the class of linear filters represented by the small circle marked "Linear". Further statements with respect to optimality require assumptions regarding object and data distributions.

4.6.1 Background. By definition, the Bayesian minimum MSE estimator is the mean of the posterior PDF, represented by $E_{\mathbf{O}|\mathbf{D}}[\mathbf{O}|\mathbf{D}]$ [43]. If $E_{\mathbf{O}|\mathbf{D}}[\mathbf{O}|\mathbf{D}]$ can be shown to be linearly related to the data, then the vector Wiener filter must be the minimum error variance filter over the class of *all* Fourier domain Bayesian minimum MSE filters. This special case is analogous to constraining the “best” estimator to exist within the small circle marked “Linear” in Fig. 4.1. Thus, to understand these optimality issues fully, the conditions under which $E_{\mathbf{O}|\mathbf{D}}[\mathbf{O}|\mathbf{D}]$ is linearly related to the data must be investigated. To lay the groundwork for this investigation, consider two jointly complex Gaussian random vectors \mathbf{X} and \mathbf{Y} whose vector $[\mathbf{X}^T \mathbf{Y}^T]^T$ can be described as

$$\begin{bmatrix} \mathbf{X} \\ \mathbf{Y} \end{bmatrix} \sim \mathcal{N} \left(\begin{bmatrix} \bar{\mathbf{X}} \\ \bar{\mathbf{Y}} \end{bmatrix}, \begin{bmatrix} C_{XX} & C_{XY} \\ C_{YX} & C_{YY} \end{bmatrix} \right), \quad (4.53)$$

where the notation \mathcal{N} denotes a random quantity with Gaussian distribution, given mean vector, and given covariance matrix. An equivalent representation of Eq. (4.53) is that the joint PDF $p_{\mathbf{X},\mathbf{Y}}(\mathbf{X}, \mathbf{Y})$ is Gaussian. Since a Gaussian PDF, real or complex, is completely described by its mean and covariance, Eq. (4.53) provides a complete statistical description. Further, it has been shown that if $p_{\mathbf{X},\mathbf{Y}}(\mathbf{X}, \mathbf{Y})$ is complex Gaussian as described in Eq. (4.53), the conditional PDF $p_{\mathbf{Y}|\mathbf{X}}(\mathbf{Y}|\mathbf{X})$ is also complex Gaussian with mean [43]

$$E_{\mathbf{Y}|\mathbf{X}}[\mathbf{Y}|\mathbf{X}] = \bar{\mathbf{Y}} + C_{YX} (C_{XX})^{-1} (\mathbf{X} - \bar{\mathbf{X}}). \quad (4.54)$$

Equation (4.54) is the mean of the posterior PDF and also the Bayesian minimum MSE estimator of \mathbf{Y} . Thus, if \mathbf{X} and \mathbf{Y} are jointly complex Gaussian, $E_{\mathbf{Y}|\mathbf{X}}[\mathbf{Y}|\mathbf{X}]$ is linearly related to \mathbf{X} . When the vector \mathbf{D} is substituted for \mathbf{X} and the vector \mathbf{O} is substituted for \mathbf{Y} , Eq. (4.54) can be viewed in terms of the image model given in Eq. (2.9), repeated as

$$\mathbf{D} = \mathcal{H} \odot \mathbf{O} + \mathbf{N}, \quad (4.55)$$

and the vector Wiener filter given in Eq. (4.52). Therefore, if \mathbf{D} and \mathbf{O} are jointly complex Gaussian, the general Bayesian minimum MSE estimate is linear to the data and the vector Wiener filter is the general minimum error variance estimator.

Under what conditions are \mathbf{D} and \mathbf{O} jointly complex Gaussian? The following discussion will answer this question based on statistical assumptions about the general linear model given in Eq. (4.55).

4.6.2 Signal-Independent Noise. Let us take a step back from the image degradation model first given in Chapter II and assume that the OTF is deterministic. In addition, let us assume that the complex noise vector \mathbf{N} is independent of \mathbf{O} and Gaussian distributed. This assumption is used in many image processing applications [41] but is not consistent with an image degraded by shot noise or the image model in Eq. (2.6). The joint PDF $p_{\mathbf{D},\mathbf{O}}(\mathbf{D}, \mathbf{O})$ can be rewritten in terms of conditional and marginal PDFs such that [62]

$$p_{\mathbf{D},\mathbf{O}}(\mathbf{D}, \mathbf{O}) = p_{\mathbf{D}|\mathbf{O}}(\mathbf{D}|\mathbf{O})p_{\mathbf{O}}(\mathbf{O}). \quad (4.56)$$

Since the noise has a Gaussian PDF and \mathcal{H} is deterministic, $p_{\mathbf{D}|\mathbf{O}}(\mathbf{D}|\mathbf{O})$ is also Gaussian distributed. Thus, if $p_{\mathbf{O}}(\mathbf{O})$ is Gaussian distributed, the joint PDF is Gaussian [62] and the vector Wiener filter is the general Bayesian minimum error variance filter.

4.6.3 Semi-Classical Model. While the Gaussian distributed signal-independent noise case is valid in many applications, it does not adequately represent low light imaging of atmospheric turbulence-degraded images. Here, shot noise and the randomness of the OTF introduce dependence on the signal. To consider this case, it is helpful to rewrite the OTF in the following form

$$\mathcal{H}_d = \bar{\mathcal{H}}_d + \Delta\mathcal{H}_d, \quad (4.57)$$

where $\Delta\mathcal{H}_d$ is a diagonal matrix representing a realization of the random OTF component. Now Eq. (4.55) can be rewritten using Eq. (4.57) such that

$$\mathbf{D} = \bar{\mathcal{H}}_d\mathbf{O} + \mathbf{N}_T, \quad (4.58)$$

where the total noise is defined as $\mathbf{N}_T = \Delta\mathcal{H}_d\mathbf{O} + \mathbf{N}$ and is dependent on the object via the random OTF and shot noise components. Even if the object PDF $p_{\mathbf{O}}(\mathbf{O})$ is Gaussian distributed, the joint PDF $p_{\mathbf{D},\mathbf{O}}(\mathbf{D}, \mathbf{O})$ will not be Gaussian since \mathbf{N}_T is not strictly Gaussian and is dependent on the object. Further, the signal dependence of the noise precludes the existence of any case where $p_{\mathbf{D},\mathbf{O}}(\mathbf{D}, \mathbf{O})$ is Gaussian distributed regardless of the choice of noise or object PDF. Thus, in a strict sense, no statements about the optimality of the vector Wiener filter can be made outside of the class of Bayesian linear filters.

However, in a limiting sense, the vector Wiener filter approaches optimality across the class of all Bayesian filters when two conditions are present: the average photon count \bar{K} grows large and the OTF covariance grows small. In Appendix B, it is shown that $p_{\mathbf{D}|\mathbf{O}}(\mathbf{D}|\mathbf{O})$ approaches a complex Gaussian distribution when the total photon count K is sufficiently large, via a Central Limit Theorem argument [23, 62]. For most practical images, K is very large, on the order of 1,000 to 1,000,000 photoevents. This value is more than adequate for valid application of the Central Limit Theorem [62]. The additional assumption of small OTF covariance supports the argument that \mathbf{N}_T approaches a Gaussian distribution. In addition, these two conditions yield the signal-independent noise case. Thus, when $p_{\mathbf{O}}(\mathbf{O})$ is Gaussian, the joint PDF approaches a Gaussian distribution, and the vector Wiener filter can be viewed as essentially optimal for all Bayesian minimum error variance filters. The two conditions noted above represent imaging under high light level conditions where an accurate estimate of the turbulence-induced OTF is available.

4.7 Summary

In this chapter, a Fourier domain filter was derived which incorporates object, blur, and noise statistical models. This analysis extends the vector Wiener filter to account properly for both shot noise and detector read noise as modeled in Eq. (2.7). The shot noise correlation depends on the product of the mean OTF and the mean object spectrum at a difference frequency [14, 54]. The theoretical optimality of the vector Wiener filter was also investigated with respect to assumptions about the underlying distributions of the data and object. When the detected image is degraded by signal-dependent shot noise and a random OTF, the vector

Wiener filter is only guaranteed to be the minimum error variance estimator among the class of linear Bayesian filters.

In the next chapter, simulated binary star data is processed with the vector Wiener filter. The binary star represents an important astronomical imaging application. These results include data degraded by both a fixed and a random OTF.

V. Vector Wiener Filter Processing of Binary Star Pairs

5.1 Introduction

In the previous chapter, a Fourier domain vector Wiener filter was derived which incorporates complete object, blur, and noise correlation statistics. Equation (2.7) was used in the derivation to model all degradation effects in the detected image. This model is consistent with an image degraded by atmospheric turbulence, shot noise, and detector read noise. In this chapter, the vector Wiener filter is used to process an important non-stationary astronomical object class, the binary star pair. The goal is to study vector Wiener filter performance when the object and optical transfer function (OTF) statistics are known exactly. Here, performance is compared against the scalar Wiener filter using mean square error (MSE), a correlation coefficient, mean square phase error (MSPE), and image realizations for both a deterministic and random OTF. In all cases, the vector Wiener filter provides reconstructions that are superior to those of the scalar Wiener filter. The results also illustrate the superresolution capability of the vector Wiener filter. In this dissertation, superresolution is defined as the extension of the detected image Fourier spectrum to regions where no data was measured [55].

Chapter V is organized as follows. Section 5.2 provides general details of the Monte Carlo simulation used to generate the data to include information about the random object, measurement noise, and performance metrics. Sections 5.3 and 5.4 present the data and conclusions associated with the deterministic and random OTF cases, respectively. The chapter ends with a summary in Section 5.5.

5.2 Simulation

In this section, details of a simulation used to study vector Wiener filter performance are provided. The term *simulation* is used to refer to Monte Carlo experiments involving the repeated application of the filter to random draws of object, OTF, and measurement noise. The discussion here includes details associated with random object and measurement noise.

5.2.1 *Random Object.* To compare vector and scalar Wiener filter performance, a class of simple non-stationary random objects is needed with Fourier domain statistics that can be derived analytically. The data generated for this study are based on a pair of Gaussian functions with locations that are random. Figure 5.1 shows (a) 16×16 and (b) 32×32 detector array object realizations that can be expressed mathematically as [22]

$$o(x, y) = h_p \exp \left\{ -\pi \left(\frac{(x - x_p)^2 + (y - y_p)^2}{w_p} \right) \right\} + h_s \exp \left\{ -\pi \left(\frac{(x - x_s)^2 + (y - y_s)^2}{w_s} \right) \right\}, \quad (5.1)$$

where (x_p, y_p) is the location of the primary function, (x_s, y_s) is the location of the secondary function, h_p is the peak irradiance of the primary, h_s is the peak irradiance of the secondary, w_p is the primary width parameter, and w_s is the secondary width parameter. The function locations (x_p, y_p) and (x_s, y_s) are independent, uniformly distributed random vectors restricted to a W by W pixel region in the center of the image plane. For example, $W = 1$ requires both functions to be located at the center of the image plane, and the object realization is deterministic. Since the function locations have a uniform distribution, the parameter W can also be viewed as the dimension of a spatial domain support constraint. As shown in Fig. 5.1, the normalized peak irradiance of the primary is $h_p = 1$ and the secondary is $h_s = 0.5$. Both primary and secondary functions have width parameters $w_p = w_s = 0.5$ pixels for the 16×16 detector array in (a) and $w_p = w_s = 1$ pixel for the 32×32 detector array in (b). In both cases, the width parameters were chosen to simulate unresolved point sources.

The object Fourier domain statistics are straightforward to derive for the Gaussian binary star pairs. The mean object spectrum is [14]

$$\bar{O}(u, v) = W^2 \text{sinc}(Wu, Wv) \left[h_p w_p \exp \{ -\pi (w_p (u^2 + v^2)) \} + h_s w_s \exp \{ -\pi (w_s (u^2 + v^2)) \} \right], \quad (5.2)$$

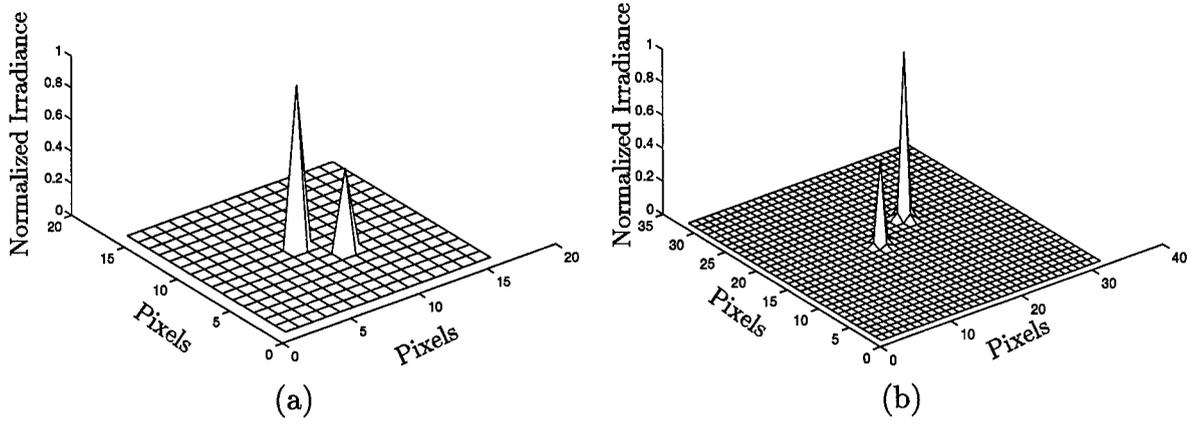


Figure 5.1 Sample binary star object realizations ($h_p = 1$, $h_s = 0.5$ pixels), (a) 16 x 16 detector ($w_p = w_s = 0.5$ pixels), (b) 32 x 32 detector ($w_p = w_s = 1$ pixel).

where the sinc function [22] is the Fourier transform of the uniform probability density function associated with a rectangular support constraint of dimension W . The correlation of (u, v) and (u', v') object Fourier components can be expressed as [14]

$$\begin{aligned}
 R_{OO}(u, v; u', v') &= W^2 \text{sinc}(W(u - u'), W(v - v')) \\
 &\times \left[h_p^2 w_p^2 \exp\{-\pi w_p(u^2 + u'^2 + v^2 + v'^2)\} \right. \\
 &\quad \left. + h_s^2 w_s^2 \exp\{-\pi w_s(u^2 + u'^2 + v^2 + v'^2)\} \right] \\
 &+ W^4 \text{sinc}(Wu, Wv) \text{sinc}^*(Wu', Wv') h_p h_s w_p w_s \\
 &\times \left[\exp\{-\pi(w_p(u^2 + v^2) + w_s(u'^2 + v'^2))\} \right. \\
 &\quad \left. + \exp\{-\pi(w_s(u^2 + v^2) + w_p(u'^2 + v'^2))\} \right]. \quad (5.3)
 \end{aligned}$$

Equations (5.2) and (5.3) provide the object Fourier domain statistics for all experiments. Each experiment trial begins by generating a new random object realization which is a member of the ensemble with these statistics.

5.2.2 Measurement Noise. Equation (4.35) suggests a measurement noise Fourier domain correlation matrix which is the sum of two components. The shot noise component is the product of the mean OTF and the mean object spectrum at a difference frequency with a scaling factor SNR_k^{-2} . The component associated with the uncorrelated detector read noise increases the diagonal elements by a scaling factor SNR_r^{-2} . Thus, the measurement noise is

correlated with respect to spatial frequency. Since the scalar Wiener filter does not exploit this information, the vector Wiener filter should have an advantage in noise suppression.

Each experiment trial generates a detected image which is corrupted by both shot noise and detector read noise. Both noise sources can be varied independently using the \bar{K} and σ_r parameters which correspond directly to light level and detector read noise, respectively. An individual trial begins by generating a normalized object realization and scaling to the selected light level as given by \bar{K} . A noiseless image is then created by multiplying the object spectrum and OTF. A Poisson random number generator is used to corrupt the noiseless image by using its irradiance values as mean parameters in the Poisson distribution. Finally, zero-mean Gaussian random numbers with standard deviation σ_r are added to model signal-independent detector read noise. This process is repeated many times to constitute a complete experiment.

5.2.3 Performance Metrics. To compare filter performance properly, metrics must be used that provide a realistic performance measure. MSE, a correlation coefficient, and MSPE are used to compare vector and scalar Wiener filter performance in this chapter. Each of these performance metrics provides a different method of comparison between the true and estimated object spectra.

5.2.3.1 Mean Square Error. Consider the definition of the error correlation matrix $R_{\epsilon\epsilon}$ presented in Eqs. (4.39)-(4.42). The diagonal elements of $R_{\epsilon\epsilon}$ represent the theoretical MSE of the respective filter. This data is presented in two forms in this chapter. First, consider the average MSE defined as

$$\bar{\epsilon}^2 = \frac{1}{P} \text{Tr} \{R_{\epsilon\epsilon}\}, \quad (5.4)$$

where P is the number of pixels in the detector array as before. Second, consider a MSE metric given as a function of radial spatial frequency ρ via radial averaging. This quantity is denoted $\epsilon^2(\rho)$. Radial averaging is the process of averaging a two dimensional function along concentric circles to produce a one dimensional plot. All MSE data in this chapter is

normalized by the error associated with the detected image. This normalization forces the MSE data values to be between 0 and 1 when the filters provide improved performance with respect to the detected image. Here, the subscripts V , S , and D denote MSE associated with the vector Wiener filter, scalar Wiener filter, and detected image, respectively.

The previous paragraph presented theoretical MSE metrics. Experimental MSE data can be collected via Monte Carlo simulation. Here, the average squared error per realization is computed and then accumulated over a large ensemble of images created under identical statistical conditions. The average squared error as calculated for the i^{th} experiment trial is

$$\varepsilon_i^2 = \frac{1}{P}(\mathbf{O}_i - \hat{\mathbf{O}}_i)^H(\mathbf{O}_i - \hat{\mathbf{O}}_i). \quad (5.5)$$

The pertinent sample statistics are the mean of Eq. (5.5) and variance of the mean defined as [84]

$$\sigma_{\varepsilon^2}^2 = \frac{1}{L(L-1)} \sum_{i=1}^L (\varepsilon_i^2)^2 - (\overline{\varepsilon^2})^2, \quad (5.6)$$

where L is the number of images in the ensemble. The theoretical average MSE metric represented by Eq. (5.4) and the experimental average MSE of Eq. (5.5) were within $\pm\sigma_{\varepsilon^2}$ in all cases. Thus, the Monte Carlo simulation used to generate other performance metrics was shown to perform accurately with respect to MSE [14].

5.2.3.2 Magnitude and Phase Error. In some cases, MSE may be a misleading quality indicator with respect to the human observer [82]. Clearly, filter estimates can be compared visually. However, visual comparison is only valid for individual realizations. Additional metrics are needed which reflect performance over a complete ensemble. In the Fourier domain, magnitude and phase play very different roles. In fact, phase plays the more important role in many situations [61]. The scalar Wiener filter cannot compensate for phase distortions due to noise which leads to less deblurring as noise increases [41]. In contrast, the vector Wiener filter incorporates some a priori knowledge of the true object phase, which can help compensate for noise without excessive smoothing. Thus, a separate comparison of magnitude and phase error performance is needed to understand the value of additional a

priori object information. Also, the metrics given below will be important in demonstrating filter performance beyond the OTF cutoff frequency.

A correlation coefficient between true and estimated Fourier spectra has been used previously to compare estimation performance [92]. In this chapter, the following correlation coefficient between $O(u, v)$ and $\hat{O}(u, v)$ is used and can be written as

$$\gamma_{O\hat{O}}(u, v) = \frac{\langle O(u, v)\hat{O}^*(u, v) \rangle}{\sqrt{\langle |O(u, v)|^2 \rangle \langle |\hat{O}(u, v)|^2 \rangle}}. \quad (5.7)$$

$|\gamma_{O\hat{O}}(u, v)|$ takes on values between 0 and 1. A value close to one implies the filter is doing a good job of reconstructing that Fourier component. Now consider the phase of the reconstructed object spectrum. MSPE is defined as

$$\varphi^2(u, v) = \langle (\phi_o(u, v) - \phi_{\hat{o}}(u, v))^2 \rangle, \quad (5.8)$$

where $\phi_o(u, v)$ is the phase of the true object spectrum and $\phi_{\hat{o}}(u, v)$ is the phase of the estimated object spectrum. Both $\gamma_{O\hat{O}}$ and φ^2 are two dimensional functions. Direct comparison of two dimensional functions is difficult. However, Fourier spectra exhibit a high degree of radial symmetry. This symmetry allows for radial averaging of $\gamma_{O\hat{O}}$ and φ^2 . The data in Sections 5.3 and 5.4 associated with these two metrics will be presented in one dimensional form using radial averaging.

5.3 Deterministic OTF

The vector Wiener filter derivation given in Eq. (4.28) assumed a random OTF and incorporated a statistical description of this quantity. However, much valuable information can be gained from scenarios in which the OTF is deterministic and known as shown in Eq. (4.29). All results presented in this section incorporate a known OTF. In all cases, a square pupil function is used to compute the fixed OTF via Eq. (2.17). The point spread function (PSF) associated with the square pupil is a $\text{sinc}^2(x, y)$ pattern with the width of the first zero crossing given by [22]

$$\Delta l = \frac{2\lambda z}{D}, \quad (5.9)$$

Table 5.1 Measurement Noise Cases

| \bar{K} | SNR_k | SNR_r |
|-----------|---------|---------|
| 10000 | 100 | 278 |
| 1000 | 32 | 28 |
| 750 | 27 | 21 |
| 500 | 22 | 14 |
| 250 | 16 | 7 |
| 100 | 10 | 3 |

where Δl is the width of the first zero crossing, λ is the imaging wavelength, z is the observation distance, and D is the square pupil dimension. To study the effect of the OTF on filter performance, the OTF cutoff frequency is changed by adjusting D . Since it is more convenient to express these quantities in a normalized pixel space, Eq. (5.9) becomes

$$\Delta l_p = \frac{2N}{D_p}, \quad (5.10)$$

where Δl_p is the width of the first zero crossing in pixels, N is the length of one side of the detector in pixels, and D_p is the square pupil dimension in pixels. In all cases, D_p will be used to identify the pupil size and corresponding PSF-OTF for each case.

Table 5.1 provides corresponding \bar{K} , SNR_k , and SNR_r values for cases of interest in this section. The detector size here is 16×16 pixels with detector read noise fixed at $\sigma_r = \sqrt{5}$ electrons per pixel in all cases. $\bar{K} = 10000$ photoevents represents high light level and low noise. Note that SNR_k and SNR_r provide an indication of the relative contribution of the two noise effects in each case. For instance, when $\bar{K} = 10000$ photoevents, SNR_k is much lower than SNR_r , which implies detector read noise will not have a large impact on performance. When $\bar{K} = 100$ photoevents, SNR_r is lower, indicating detector read noise will play a more significant role. All MSE plots in this section were generated via Eq. (5.4). The $|\gamma_{\delta\delta}(\rho)|$ and $\varphi^2(\rho)$ plots were generated via Monte Carlo simulation where $L = 100,000$ images.

As presented in Section 5.2, W represents the square dimension in the center of the image plane where the random binary Gaussian object components are allowed to exist. Thus, W represents the amount of randomness associated with the object ensemble. Figure

5.2 compares scalar and vector Wiener filter normalized average MSE performance as a function of W when the pupil size is $D_p = 6$ pixels and $\bar{K} = 10000$ photoevents ($\text{SNR}_k = 100$, $\text{SNR}_r = 278$). First, note that the vector filter provides lower MSE over a large range of W values. This improved performance is due in large part to the a priori object Fourier domain correlation information used by the new filter. The scalar filter has access only to the object power spectral density (PSD), which is equivalent to the diagonal values of the R_{OO} matrix. Since the object ensemble is non-stationary, R_{OO} has non-zero off-diagonal elements. The scalar filter cannot incorporate this valuable information and, therefore, is no longer optimal with respect to MSE as evidenced by Fig. 5.2.

It should also be noted that the MSE associated with the vector Wiener filter increases as the support constraint dimension W increases. Matson noted similar performance for iterative algorithms [54]. His analysis showed that applying support constraints can provide both a superresolution effect and variance reduction in the noisy Fourier data [54]. Support constraints provide variance reduction by maintaining Fourier domain correlations in the data which provide weighted interpixel averaging [54]. The weights are associated with the Fourier transform of the support function. Thus, as support increases, the Fourier transform of the support function narrows, providing less averaging and degraded performance. Similarly, the vector Wiener filter also provides interpixel averaging based on enforcing degraded data Fourier domain correlations. As W increases, support size increases and the Fourier transform of the support function narrows. The off-diagonal elements of R_{OO} are reduced, which results in a filter transformation matrix with less off-diagonal structure. Less off-diagonal structure in the filter transformation matrix is analogous to less interpixel averaging as each Fourier component is estimated. In the limit, when no support constraint is used, the filter transformation matrix is diagonal and no interpixel averaging occurs. Hence, the MSE performance of the vector and scalar Wiener filters is the same for $W = 16$ pixels in Fig. 5.2.

Figure 5.3 shows a single detected image realization with the associated filter outputs for $W = 8$ pixels, pupil size $D_p = 6$ pixels, and $\bar{K} = 10000$ photoevents ($\text{SNR}_k = 100$, $\text{SNR}_r = 278$). Each mesh plot is normalized to a peak value of unity for visual comparison.

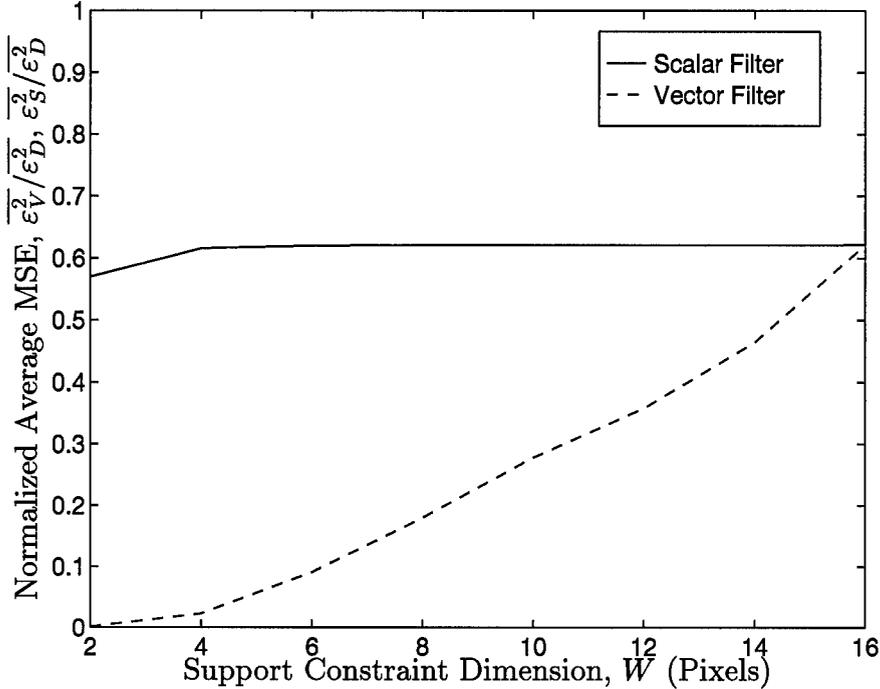


Figure 5.2 Normalized average MSE, $\overline{\varepsilon_V^2}/\overline{\varepsilon_D^2}$ and $\overline{\varepsilon_S^2}/\overline{\varepsilon_D^2}$, versus support constraint dimension W . Pupil size $D_p = 6$ pixels and $\bar{K} = 10000$ photoevents ($\text{SNR}_k = 100$, $\text{SNR}_r = 278$).

Clearly, the vector filter produces a sharper output which more closely resembles the true object realization, as suggested by the MSE data. The mean object in (b) and the PSF in (c) are provided for the convenience of the reader and later reference.

As noted in the discussion associated with Fig. 5.2, the vector Wiener filter can provide superresolution. Figures 5.4 and 5.5 illustrate this idea by comparing vector filter $|\gamma_{O\hat{O}}(\rho)|$ and $\varphi^2(\rho)$ with those associated with the scalar filter. In both plots, the pupil size is $D_p = 6$ pixels, the support constraint dimension is $W = 8$ pixels and $\bar{K} = 1000$ photoevents ($\text{SNR}_k = 32$, $\text{SNR}_r = 28$). Note that the light level has been reduced, compared to the results depicted in Figs. 5.2 and 5.3. In addition, the spatial frequency unity represents the OTF cutoff frequency. The scalar filter cannot provide superresolution since it sets spatial frequencies beyond the OTF cutoff to zero [41]. Thus, $|\gamma_{O\hat{O}}(\rho)|$ for the scalar filter drops sharply beyond the cutoff frequency. However, the vector Wiener filter is able to maintain a high correlation coefficient at these same frequencies. Figure 5.5 supports the idea of a superresolution effect since $\varphi^2(\rho)$ associated with the vector filter is much lower beyond

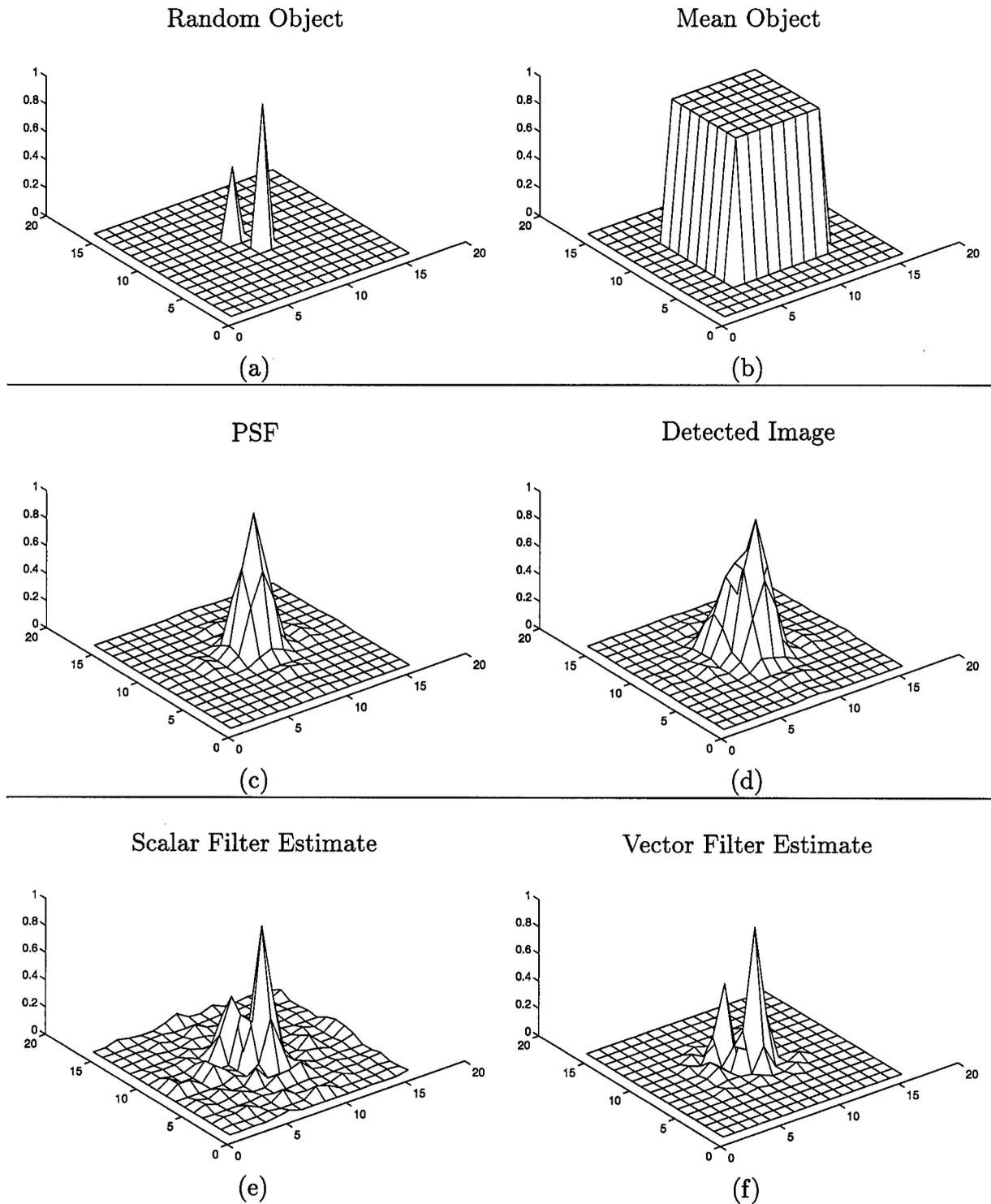


Figure 5.3 Normalized mesh plots showing the improved performance characteristics of the vector filter on binary Gaussian function objects. (a) True object realization, object randomness parameter $W = 8$ pixels, (b) mean object, (c) PSF, pupil size $D_p = 6$ pixels, (d) detected image $\bar{K} = 10000$ photoevents ($\text{SNR}_k = 100$, $\text{SNR}_r = 278$), (e) scalar filter estimate, (f) vector filter estimate.

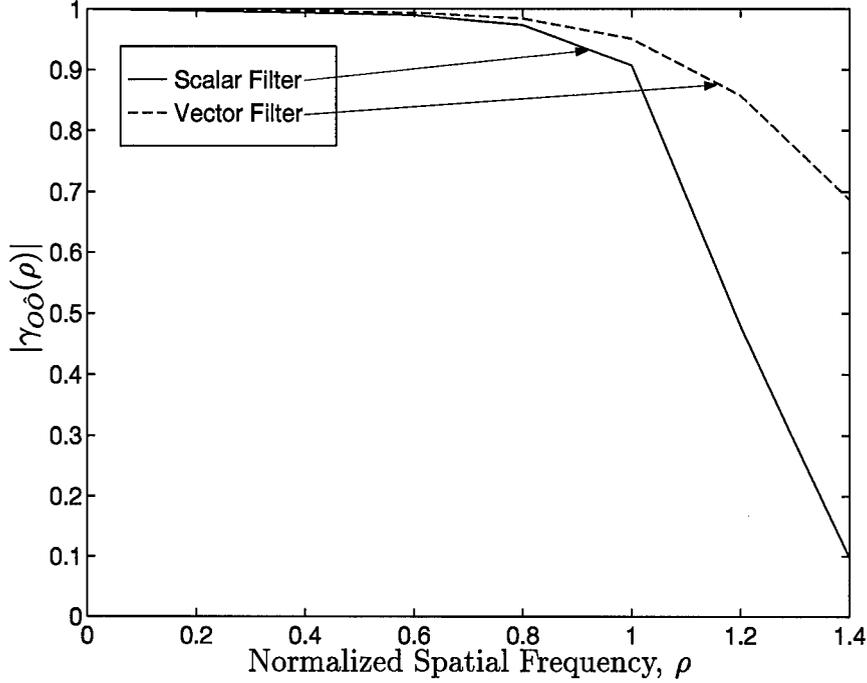


Figure 5.4 Radially averaged $|\gamma_{OO}|$ as a function of normalized spatial frequency ρ for the scalar and vector filters. The spatial frequency at unity corresponds to the OTF cutoff frequency for a $D_p = 6$ pixel pupil function. Support constraint dimension $W = 8$ pixels and $\bar{K} = 1000$ photoevents ($\text{SNR}_k = 32$, $\text{SNR}_r = 28$).

unity spatial frequency. As noted above, the a priori object knowledge incorporated in the vector Wiener filter provides superresolution, since W is associated with a spatial domain support constraint. This support constraint is incorporated directly in R_{OO} and \bar{O} . Support constraints have a long research history as a means to superresolve data in the Fourier domain [54, 55].

Now reconsider the form of the vector Wiener filter given by Eq. (4.28) and the scalar Wiener filter given by Eq. (4.40). It is obvious that neither filter incorporates detected image information beyond the OTF cutoff frequency. At those high spatial frequencies, the vector filter relies exclusively on object statistical information. In contrast, the scalar filter is not capable of incorporating this knowledge. Instead, the frequency components beyond the OTF cutoff are set to zero. Thus, the vector filter should continue to perform better than the scalar filter as the OTF cutoff frequency is adjusted lower. Figure 5.6 shows filter MSE performance as a function of the pupil size D_p when $W = 8$ pixels and $\bar{K} = 1000$ photoevents ($\text{SNR}_k = 32$, $\text{SNR}_r = 28$). Figure 5.6 shows that the vector filter provides lower MSE than

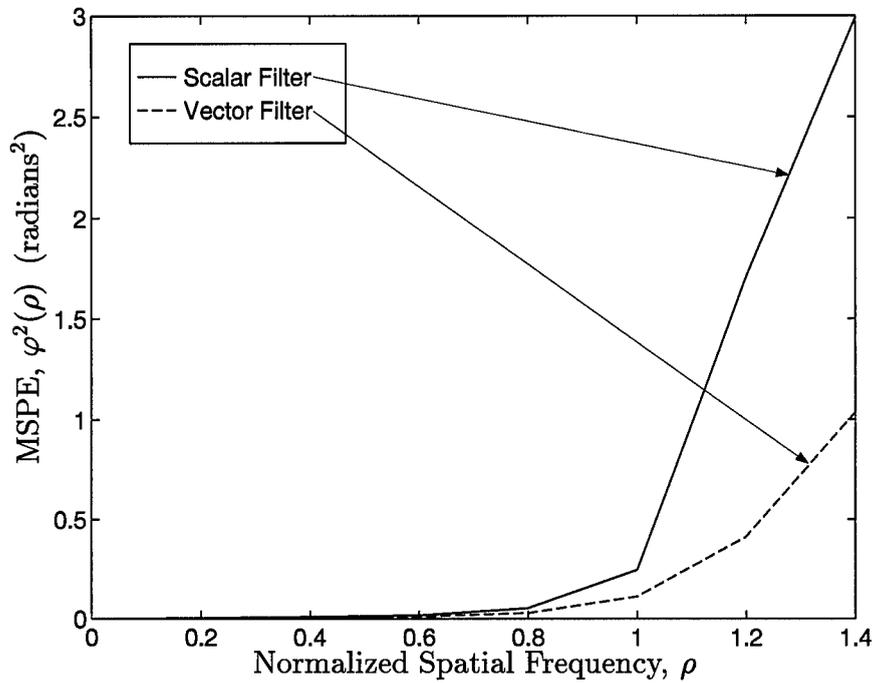


Figure 5.5 Radially averaged MSPE, φ^2 , as a function of normalized spatial frequency ρ for the scalar and vector filters. The spatial frequency at unity corresponds to the OTF cutoff frequency for a $D_p = 6$ pixel pupil function. Support constraint dimension $W = 8$ pixels and $\bar{K} = 1000$ photoevents ($\text{SNR}_k = 32$, $\text{SNR}_r = 28$).

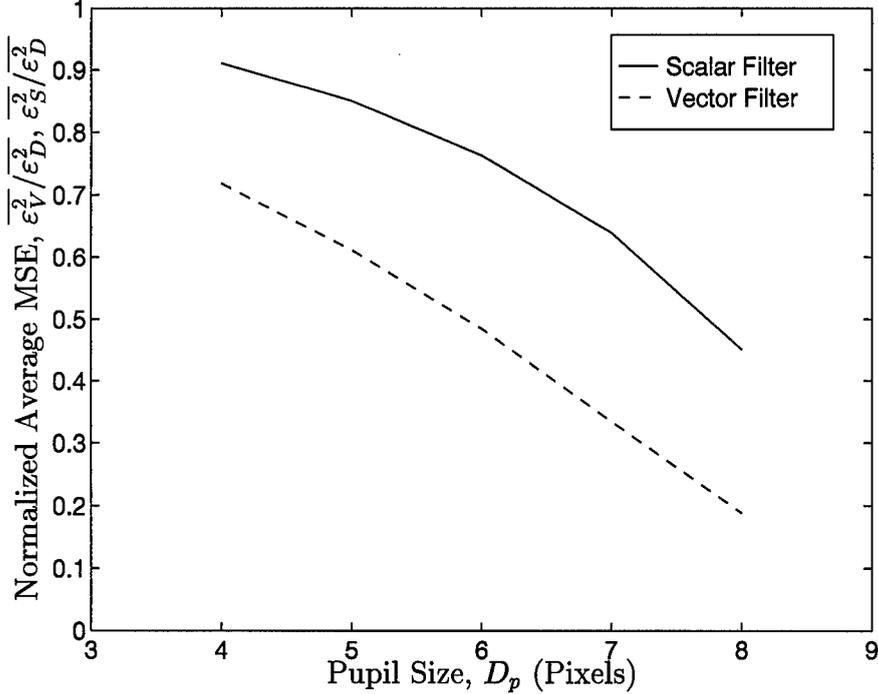


Figure 5.6 Normalized average MSE, $\overline{\varepsilon_V^2}/\overline{\varepsilon_D^2}$ and $\overline{\varepsilon_S^2}/\overline{\varepsilon_D^2}$, versus pupil size D_p . Support constraint dimension $W = 8$ pixels and $\overline{K} = 1000$ photoevents ($\text{SNR}_k = 32$, $\text{SNR}_r = 28$).

the scalar filter in all cases. As in Fig. 5.2, this MSE performance is the combined result of both a superresolution effect and enforcing detected data Fourier domain correlations.

In Chapter II, the vector Wiener filter was derived based on the image model given in Eq. (2.7) with the idea of properly modeling all noise effects. Thus, we are clearly interested in the performance of this new vector filter as measurement noise becomes more dominant. Figure 5.7 provides MSE performance as a function of \overline{K} for $W = 8$ pixels and pupil size $D_p = 6$ pixels. Clearly, the performance of both filters improves as light level increases. However, it must also be noted that the vector filter reduces MSE by a wider margin as light level increases. At $\overline{K} = 100$ photoevents the vector Wiener filter provides an additional 15% decrease in MSE below the baseline established by the scalar filter. At $\overline{K} = 1000$ photoevents the decrease in MSE is 37%. This trend is expected since the off-diagonal elements of the measurement noise Fourier domain correlation matrix R_{NN} are due to the shot noise. As light level is reduced, shot noise is less dominant and R_{NN} becomes more diagonal, which minimizes the importance of the off-diagonal elements. As noted before, it is access to the

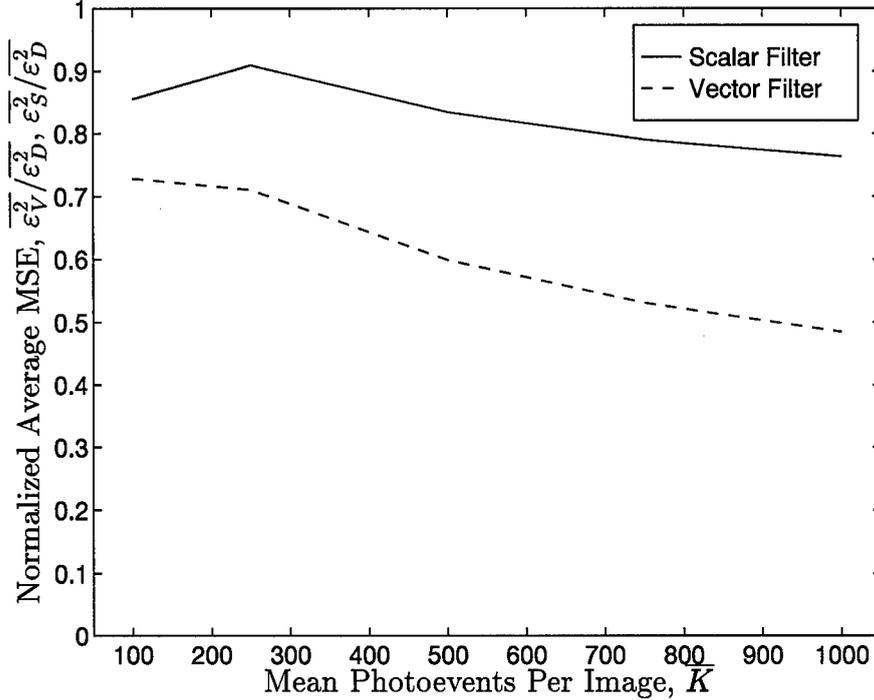


Figure 5.7 Normalized average MSE, $\overline{\epsilon_V^2}/\overline{\epsilon_D^2}$ and $\overline{\epsilon_S^2}/\overline{\epsilon_D^2}$, versus \overline{K} . Support constraint dimension $W = 8$ pixels and pupil size $D_p = 6$ pixels.

off-diagonal components of the correlation matrices which provides the difference between the vector and scalar Wiener filters. Figures 5.8 and 5.9 also illustrate the superior noise suppression of the vector Wiener filter for which the support constraint dimension W and pupil size D_p are unchanged from Fig. 5.7. The $|\gamma_{\hat{O}\hat{O}}(\rho)|$ and $\varphi^2(\rho)$ metrics are superior across all noise cases with a superresolution effect still visible beyond the OTF cutoff frequency.

Figure 5.10 provides detected image and filter outputs when $\overline{K} = 500$ photoevents ($\text{SNR}_k = 22, \text{SNR}_r = 14$). The scalar filter simply smoothes the noisy data and, therefore, is not able to resolve the two object components in (e). In contrast, the vector Wiener filter resolves the components in (f). Note how the object Fourier domain statistics enforce the support constraint imposed by the parameter W . The next section illustrates vector Wiener filter performance when the OTF is random and associated with atmospheric turbulence.

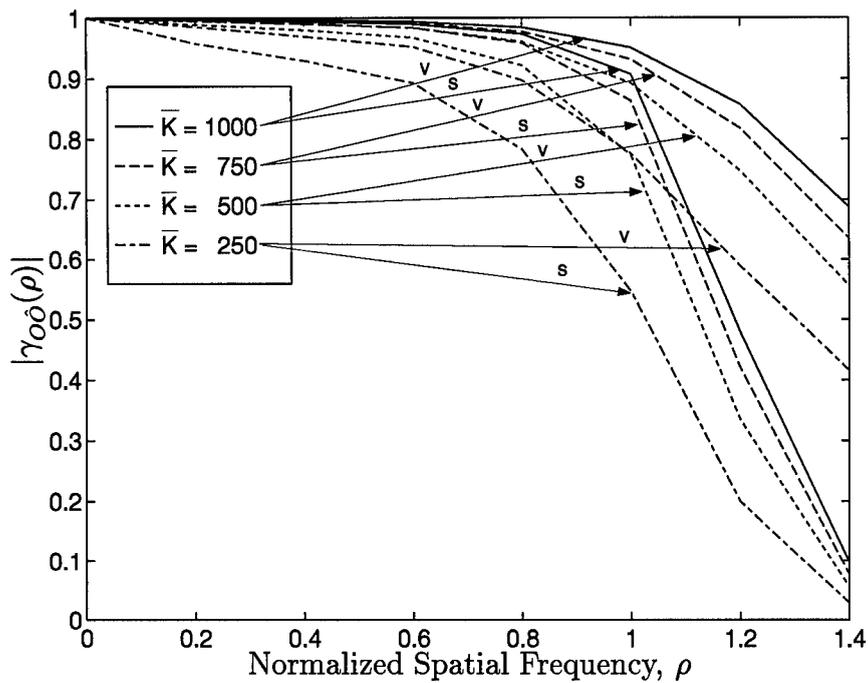


Figure 5.8 Radially averaged $|\gamma_{00}|$ as a function of normalized spatial frequency ρ for the scalar and vector filters. The spatial frequency at unity corresponds to the OTF cutoff frequency for a $D_p = 6$ pixel pupil function and support constraint dimension $W = 8$ pixels. The v and s designators differentiate between vector and scalar filter traces.

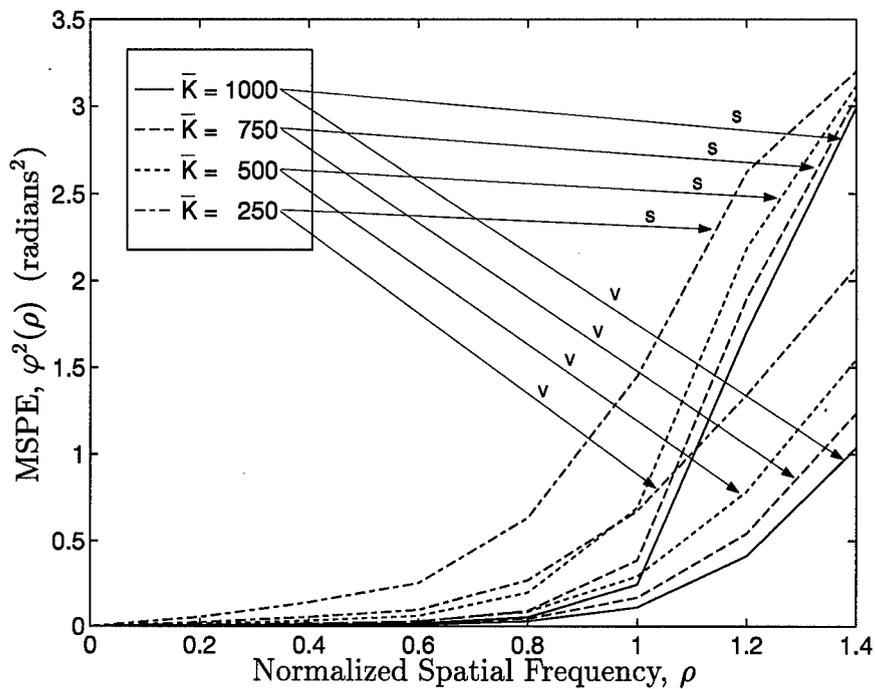


Figure 5.9 Radially averaged MSPE, φ^2 , as a function of normalized spatial frequency ρ for the scalar and vector filters. The spatial frequency at unity corresponds to the OTF cutoff frequency for a $D_p = 6$ pixel pupil function and support constraint dimension $W = 8$ pixels. The v and s designators differentiate between vector and scalar filter traces.

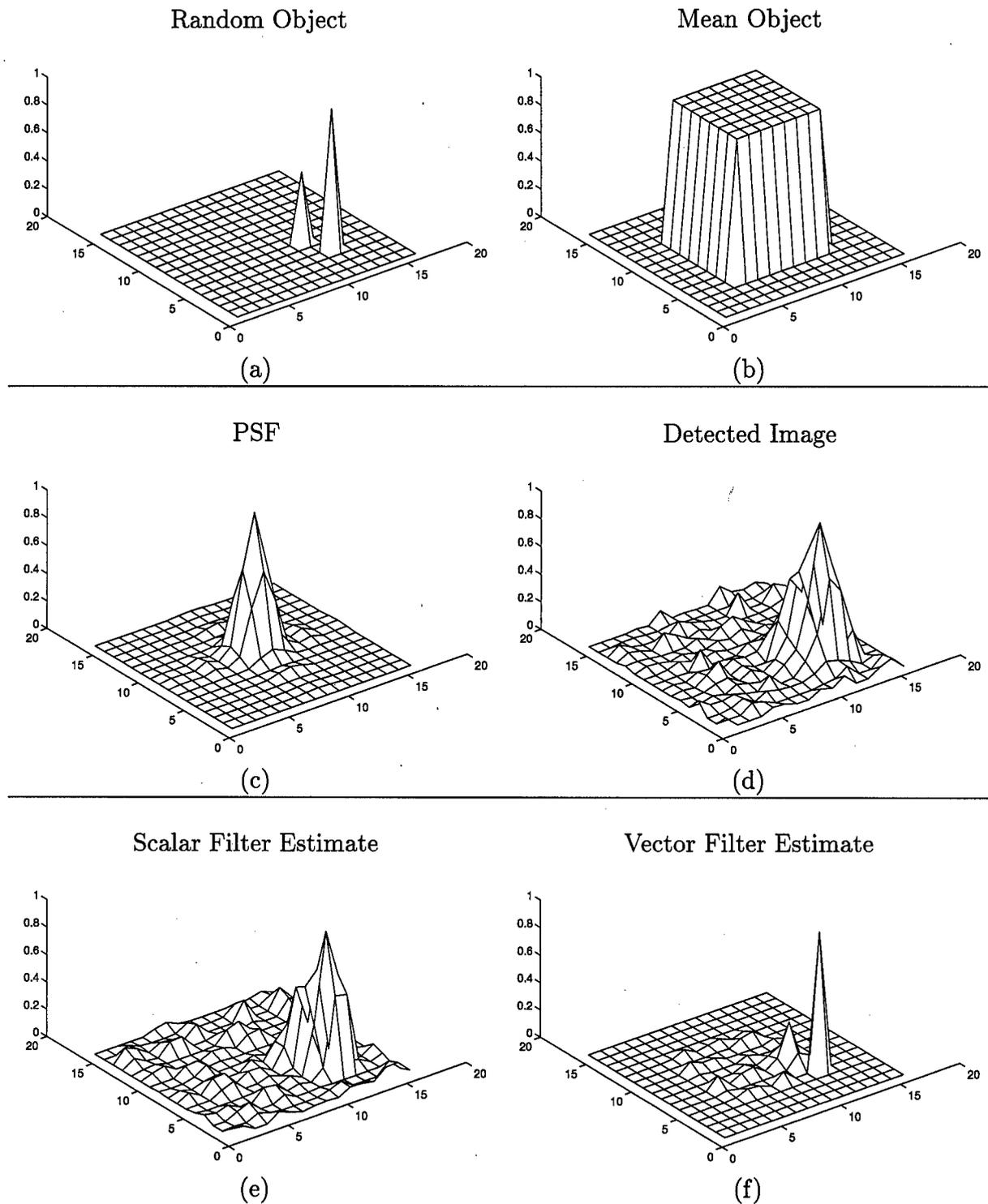


Figure 5.10 Normalized mesh plots showing the effect of measurement noise on filter performance. (a) True object realization, object support constraint dimension $W = 8$ pixels, (b) mean object, (c) PSF, pupil size $D_p = 6$ pixels, (d) detected image, $\bar{K} = 500$ photoevents ($\text{SNR}_k = 22$, $\text{SNR}_r = 14$), (e) scalar filter estimate, (f) vector filter estimate.

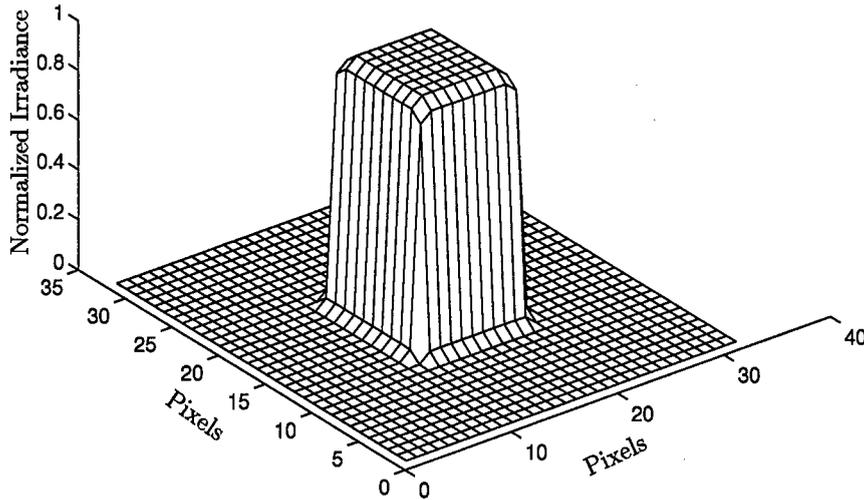


Figure 5.11 Mean object used to generate the random OTF data ($W = 10$ pixels).

5.4 Random OTF Due to Atmospheric Turbulence

In this section, the OTF is random and associated with atmospheric turbulence. Since our objective here is to study the effect of turbulence on the vector Wiener filter, the object support constraint is fixed at $W = 10$ pixels and the images are photon-limited ($\sigma_r = 0$ electrons per pixel). Figure 5.11 gives the mean object associated with this constraint. All data in this section is based on a 32×32 image array. A priori knowledge of the imaging scenario is assumed in all cases. Therefore, perfect knowledge of the OTF statistics $\overline{\mathcal{H}}_d$ and $R_{\mathcal{H}\mathcal{H}}$ is available. However, no knowledge is assumed about the individual OTF realizations.

Within the Monte Carlo simulation, a random OTF is generated via Eq. (2.17) using the von Karman statistics and a Fourier series-based phase screen generator [89]. The pupil plane residual phase aberration $\phi_r(x, y)$ is due to tilt-removed distortions. Thus, a rudimentary, first order AO system is simulated. The OTF statistics were collected by repeatedly creating independent, random OTF realizations using the phase screen generator, removing tilt, and then averaging over an ensemble of size 10000. The optical system pupil is square with dimension D . Fried's parameter r_o is used to indicate turbulence strength [23]. Four turbulence strength cases are examined using $D/r_o = 1, 2, 4$, and a diffraction-limited OTF. The ratio between the turbulence outer scale L_o and r_o is fixed at $L_o/r_o = 100$. Thus, the largest turbulent eddies are two orders of magnitude larger than the seeing cell size. Figure

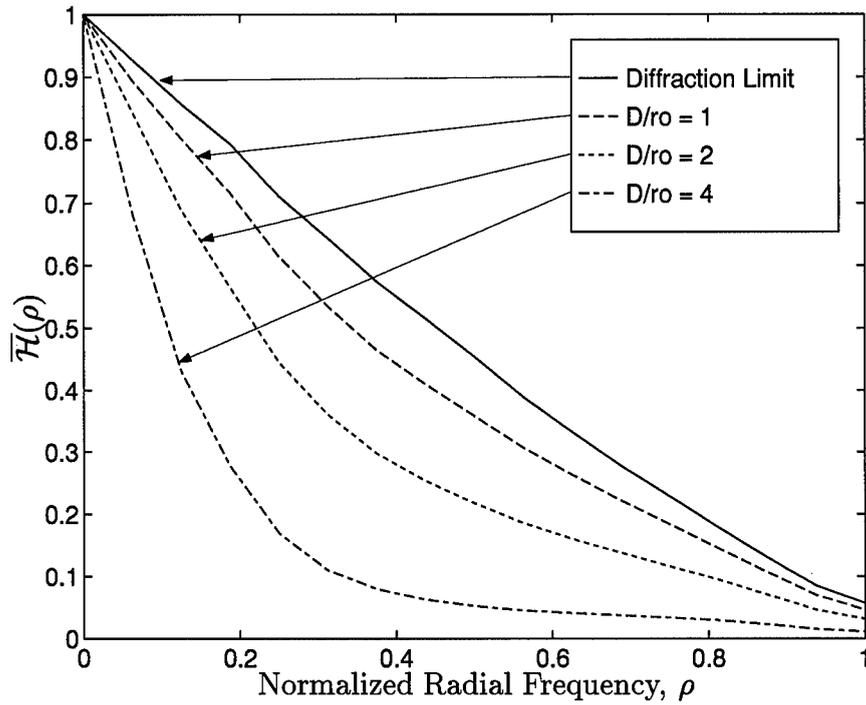
5.12 gives tilt-removed OTF statistics as a function of the radially averaged spatial frequency variable ρ for the cases noted above. The normalized spatial frequency $\rho = 1$ is associated with the diffraction-limited cutoff of the optical system as before.

Recall the vector and scalar Wiener filter expressions given in Eqs. (4.28) and (4.40). Here, it can be seen that the mean OTF plays a key role in selecting the amount of data information in the final reconstructions. As the OTF is attenuated, performance will be degraded as the filters rely on the object model to recover spatial frequencies [12, 13]. Increasing OTF variance should also degrade filter performance. A metric is needed to combine information about the OTF mean and variance into a single quantity. OTF SNR is defined as

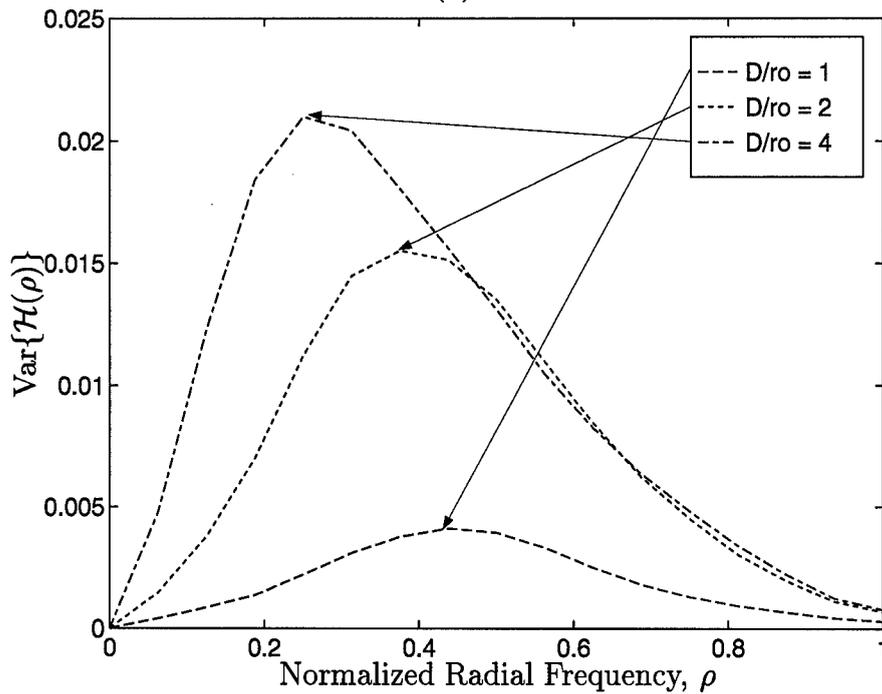
$$\text{SNR}_{\mathcal{H}}(u, v) = \frac{|\overline{\mathcal{H}}(u, v)|}{\sqrt{\text{Var}[\mathcal{H}(u, v)]}}. \quad (5.11)$$

Figure 5.13 gives $\text{SNR}_{\mathcal{H}}$ as a function of the radially averaged spatial frequency variable ρ . Notice when $D/r_o = 4$, $\text{SNR}_{\mathcal{H}}(\rho) < 1$ across a broad band of spatial frequencies. In contrast, $\text{SNR}_{\mathcal{H}}(\rho) > 1$ out to the diffraction-limited cutoff of the optical system in the other cases. This observation will be useful in predicting filter performance with regard to the results given below.

As noted earlier, our primary objective is to investigate the performance of the vector Wiener filter with regard to atmospheric turbulence. However, some performance trends associated with light level should be noted. Figure 5.14 gives the vector Wiener filter normalized average MSE versus \overline{K} for the four turbulence strength cases. Both vector and scalar Wiener filter data are presented for comparison. As expected, MSE increases as light level decreases [14]. In cases where more light is available, the randomness associated with atmospheric turbulence is the dominant effect. Also note the performance degradation between $D/r_o = 2$ and 4. This trend is also expected based on the increased attenuation of the mean OTF and increased OTF variance as shown in Fig. 5.12. However, the relative size of the increase in MSE shown in Fig. 5.14 seems to separate $D/r_o = 4$ from the other cases. Two questions naturally arise from this observation:



(a)



(b)

Figure 5.12 Tilt-removed OTF statistics due to atmospheric turbulence versus radially averaged spatial frequency ρ (von Karman turbulence statistics $L_o/r_o = 100$, $D/r_o = 1, 2$, and 4). (a) Mean, (b) variance. The OTF due to diffraction is provided as a reference to the reader in (a).

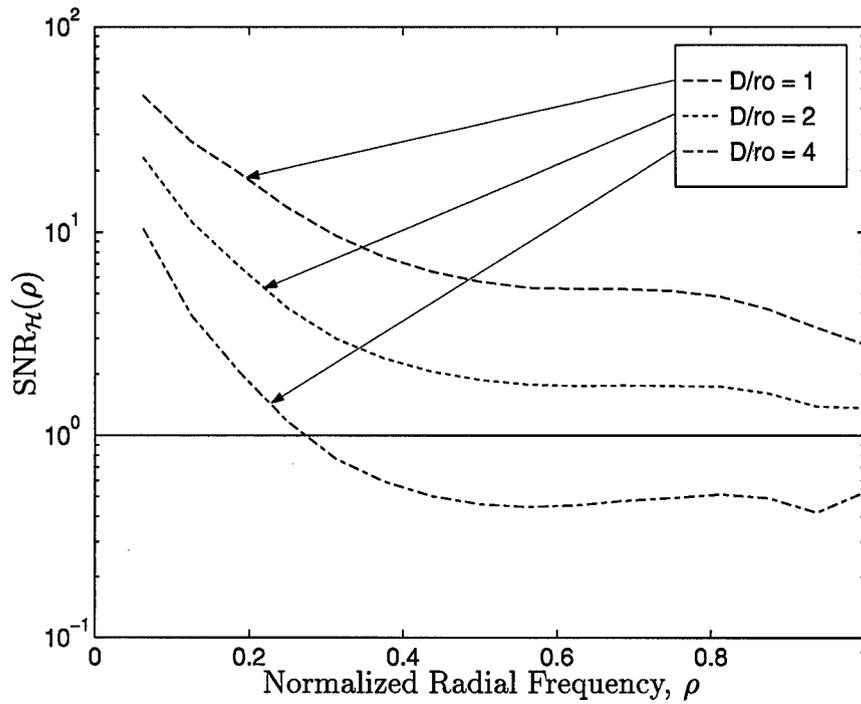


Figure 5.13 Tilt-removed OTF SNR, $SNR_{\mathcal{H}}(\rho)$, due to atmospheric turbulence versus radially averaged spatial frequency ρ (von Karman turbulence statistics $L_o/r_o = 100$, $D/r_o = 1, 2$, and 4). Note that $SNR_{\mathcal{H}}(\rho) < 1$ when $\rho > 0.3$ in the $D/r_o = 4$ case.

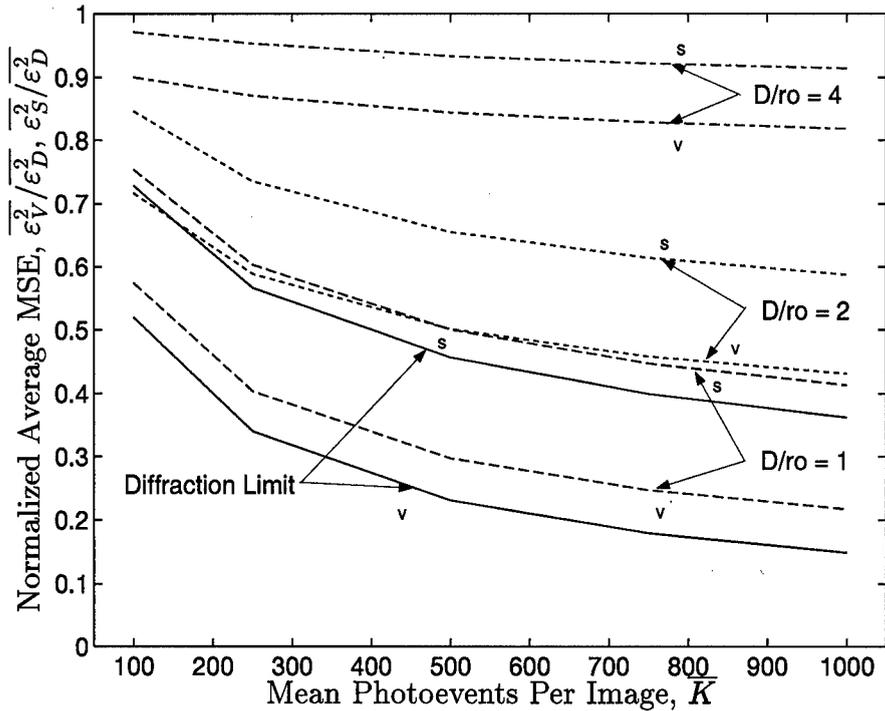


Figure 5.14 Normalized average MSE, $\overline{\varepsilon_V^2/\varepsilon_D^2}$ and $\overline{\varepsilon_S^2/\varepsilon_D^2}$, versus \overline{K} for the diffraction-limited OTF and $D/r_o = 1, 2,$ and 4 . The v and s designators differentiate between vector and scalar Wiener filter traces.

1. Is there a limit to how much OTF randomness the vector Wiener filter can handle before performance is severely degraded?
2. Can $\text{SNR}_{\mathcal{H}}$ be used to predict vector Wiener filter performance?

To address the questions posed above, consider the normalized MSE at individual spatial frequencies. Figure 5.15 shows radially averaged normalized MSE for the four turbulence strength cases under high light level conditions ($\overline{K} = 10000$ photoevents). In general, the normalized MSE associated with both filters increases as ρ increases. This trend is due to the attenuation of the OTF and increased noise effects at high spatial frequencies. However, note the jump in MSE that occurs when $\rho \approx 0.2$ to 0.4 in the $D/r_o = 4$ case. Returning to Fig. 5.13, we observe that $\text{SNR}_{\mathcal{H}}(\rho)$ falls below one when $\rho \approx 0.3$ for this case. In contrast, the other three turbulence strength cases exhibit a much more gradual increase in MSE with ρ . Thus, the vector Wiener filter seems to perform differently with respect to MSE when $\text{SNR}_{\mathcal{H}}(\rho) < 1$. Also, note that the scalar Wiener filter seems to follow the same trends except that its performance is more degraded. The vector Wiener filter has a

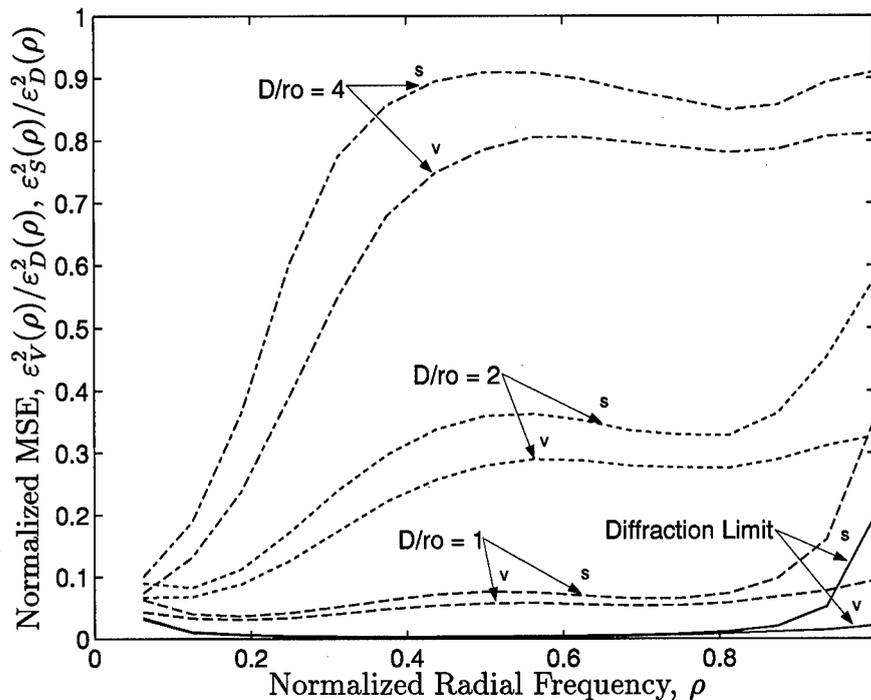


Figure 5.15 Normalized MSE, $\epsilon_v^2(\rho)/\epsilon_D^2(\rho)$ and $\epsilon_s^2(\rho)/\epsilon_D^2(\rho)$, for the diffraction-limited OTF and $D/r_o = 1, 2, \text{ and } 4$ ($\bar{K} = 10000$ photoevents). The v and s designators differentiate between vector and scalar Wiener filter traces. Filter MSE performance is severely degraded at radial frequencies ρ where $\text{SNR}_{\mathcal{H}}(\rho) < 1$ in the $D/r_o = 4$ case.

particular advantage at high spatial frequencies due to interpixel averaging provided by the object statistical model.

Now consider a change in light level. Figures 5.16 and 5.17 provide the same data as Fig. 5.15 except light level has been drastically reduced ($\bar{K} = 1000$ and 500 photoevents, respectively). Thus, MSE has increased across most spatial frequencies. However, the general trends with respect to $\text{SNR}_{\mathcal{H}}(\rho)$ remain. $\text{SNR}_{\mathcal{H}}(\rho) = 1$ is a relevant performance indicator for this object class, even at low light level.

Let us now examine the MSPE performance of the filters with respect to atmospheric turbulence and shot noise. Figures 5.18 and 5.19 provide φ^2 traces plotted versus radially averaged spatial frequency ρ . The vector Wiener filter produces reconstructions with less MSPE due to superresolution and Fourier data variance reduction associated with the a priori object statistics. As with the MSE above, note the substantial jump in MSPE in Fig. 5.18

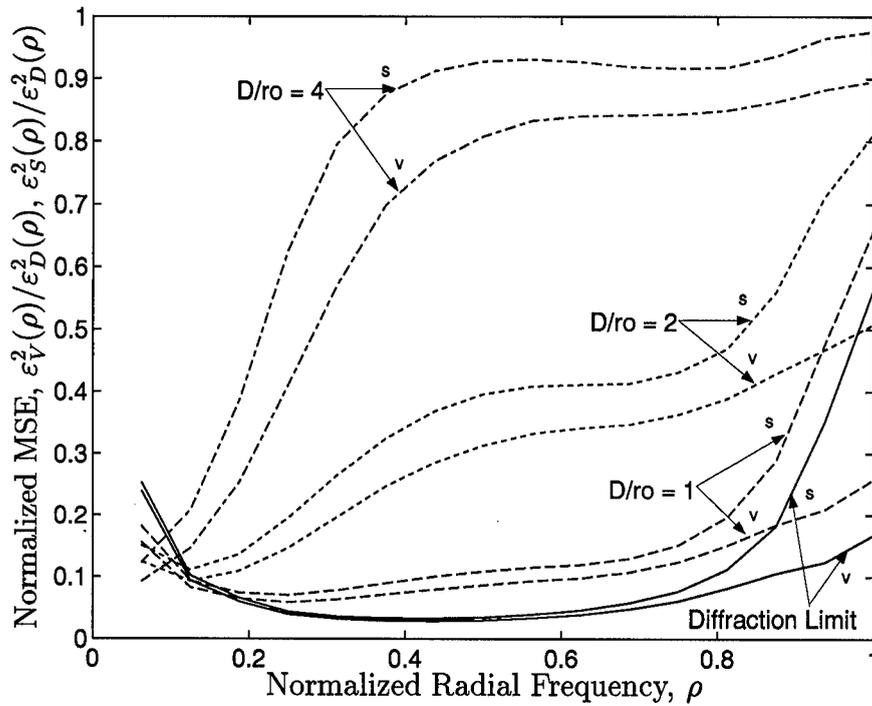


Figure 5.16 Normalized MSE, $\epsilon_V^2(\rho)/\epsilon_D^2(\rho)$ and $\epsilon_S^2(\rho)/\epsilon_D^2(\rho)$, for the diffraction-limited OTF and $D/r_o = 1, 2,$ and 4 ($\bar{K} = 1000$ photoevents). The v and s designators differentiate between vector and scalar Wiener filter traces.

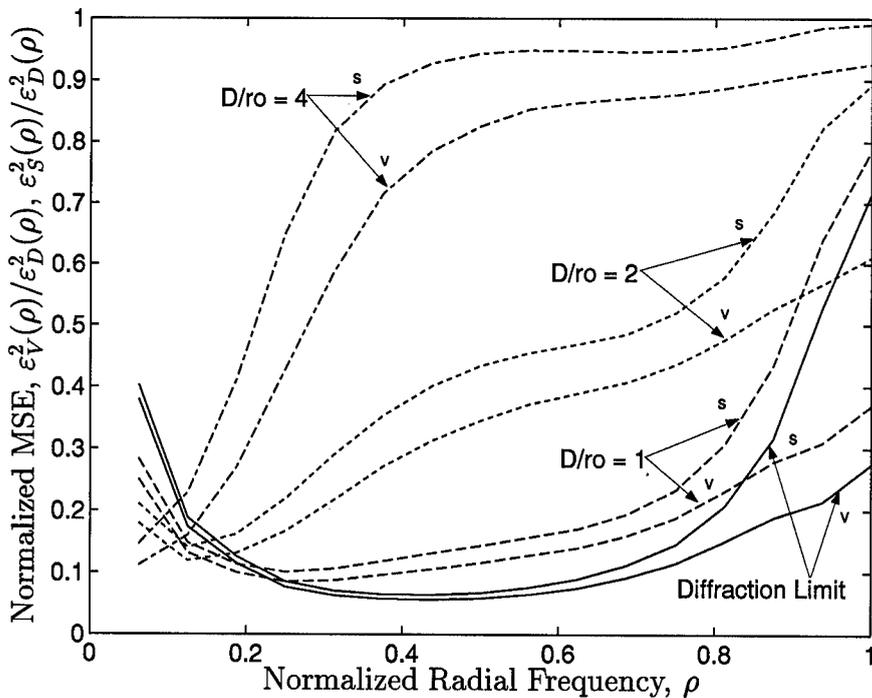


Figure 5.17 Normalized MSE, $\epsilon_V^2(\rho)/\epsilon_D^2(\rho)$ and $\epsilon_S^2(\rho)/\epsilon_D^2(\rho)$, for the diffraction-limited OTF and $D/r_o = 1, 2,$ and 4 ($\bar{K} = 500$ photoevents). The v and s designators differentiate between vector and scalar Wiener filter traces.

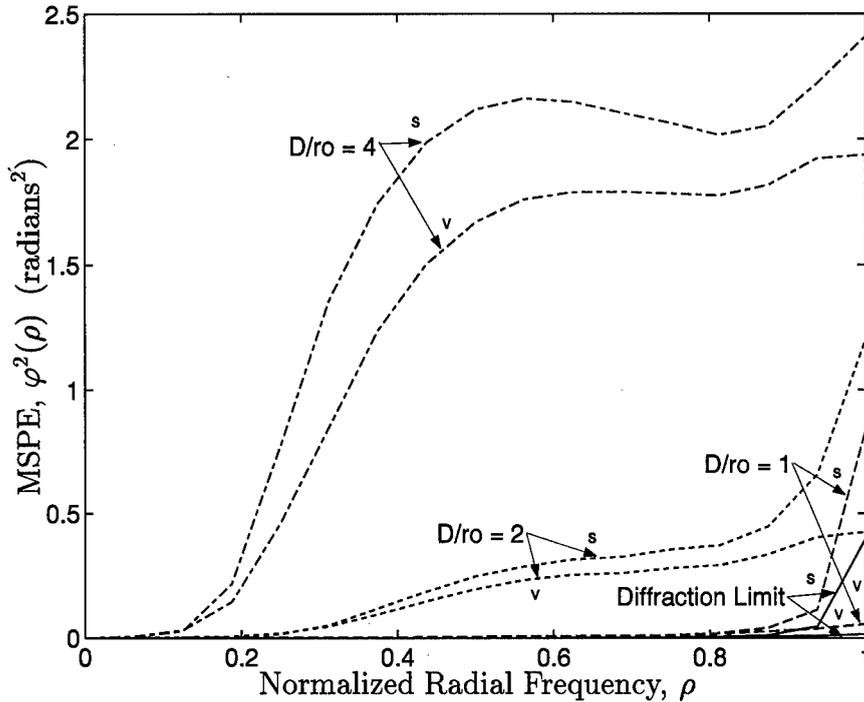


Figure 5.18 MSPE data $\varphi^2(\rho)$. Turbulence strength traces ($\bar{K} = 10000$ photoevents). The v and s designators differentiate between vector and scalar Wiener filter traces.

between turbulence strength cases $D/r_o = 2$ and 4 . In this plot, $\bar{K} = 10000$ photoevents, which represents a high light level. This observation lends additional support to the use of $\text{SNR}_{\mathcal{H}}(\rho)$ as a performance indicator for the vector Wiener filter. Figure 5.19 shows the same type of data except light level is varied and $D/r_o = 2$. Light level has a minimal effect on $\varphi^2(\rho)$ when $\bar{K} = 10000$ and $\bar{K} = 5000$ photoevents. Once again, the vector Wiener filter provides superior performance across all cases. However, the greatest performance gains are associated with spatial frequencies where $\text{SNR}_{\mathcal{H}}(\rho) > 1$.

Figure 5.20 shows detected image and filter outputs when $D/r_o = 4$ and $\bar{K} = 1000$ photoevents. The scalar filter simply smoothes the atmospheric turbulence-degraded data in (e). In contrast, the vector Wiener filter produces a sharper and better resolved reconstruction in (f). These image realizations illustrate a superior reconstruction even when filter performance has been degraded due to the strength of the atmospheric turbulence.

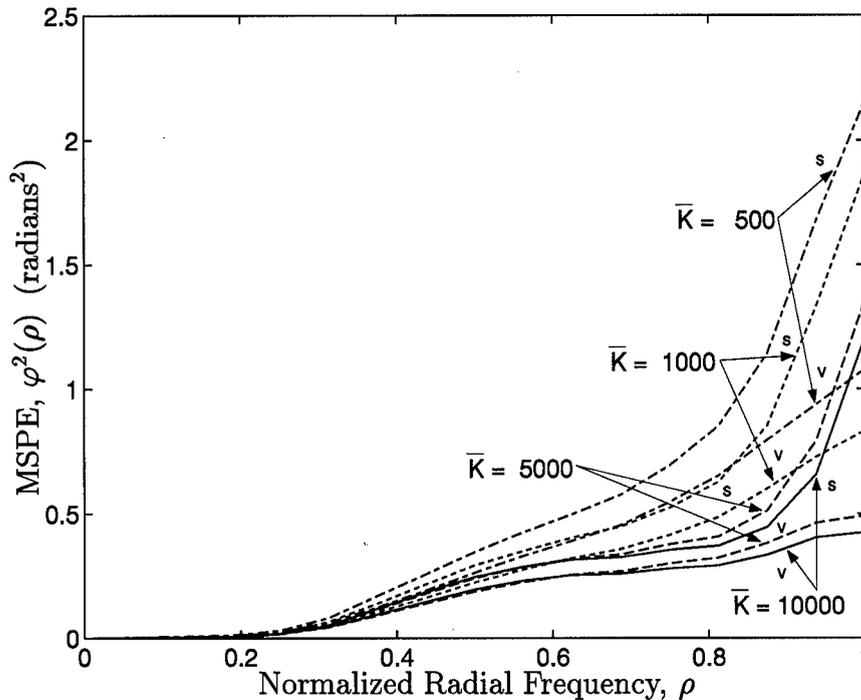


Figure 5.19 MSPE data $\varphi^2(\rho)$. Light level traces ($D/r_o = 2$). The v and s designators differentiate between vector and scalar Wiener filter traces.

5.5 Summary

Binary star simulation results show that the vector Wiener filter provides superior reconstructions when compared to the scalar Wiener filter for non-stationary object ensembles. Comparisons were conducted while varying the object support constraint dimension W , the support of the OTF, and the measurement noise level. In addition, a superresolution capability of the vector filter was illustrated by examining performance beyond the OTF cutoff frequency. Vector Wiener filter performance was also studied for photon-limited images degraded by atmospheric turbulence. Random OTFs associated with tilt-removed atmospheric turbulence cases were generated using a Fourier series-based phase screen generator [89]. This experiment was the first application of complete OTF correlation statistics to the reconstruction of turbulence-degraded images [12, 13]. Comparisons were conducted while varying the turbulence strength and light level. The vector Wiener filter was superior to the scalar Wiener filter with respect to normalized MSE and MSPE across all cases examined. However, substantial performance degradation was noted at spatial frequencies where

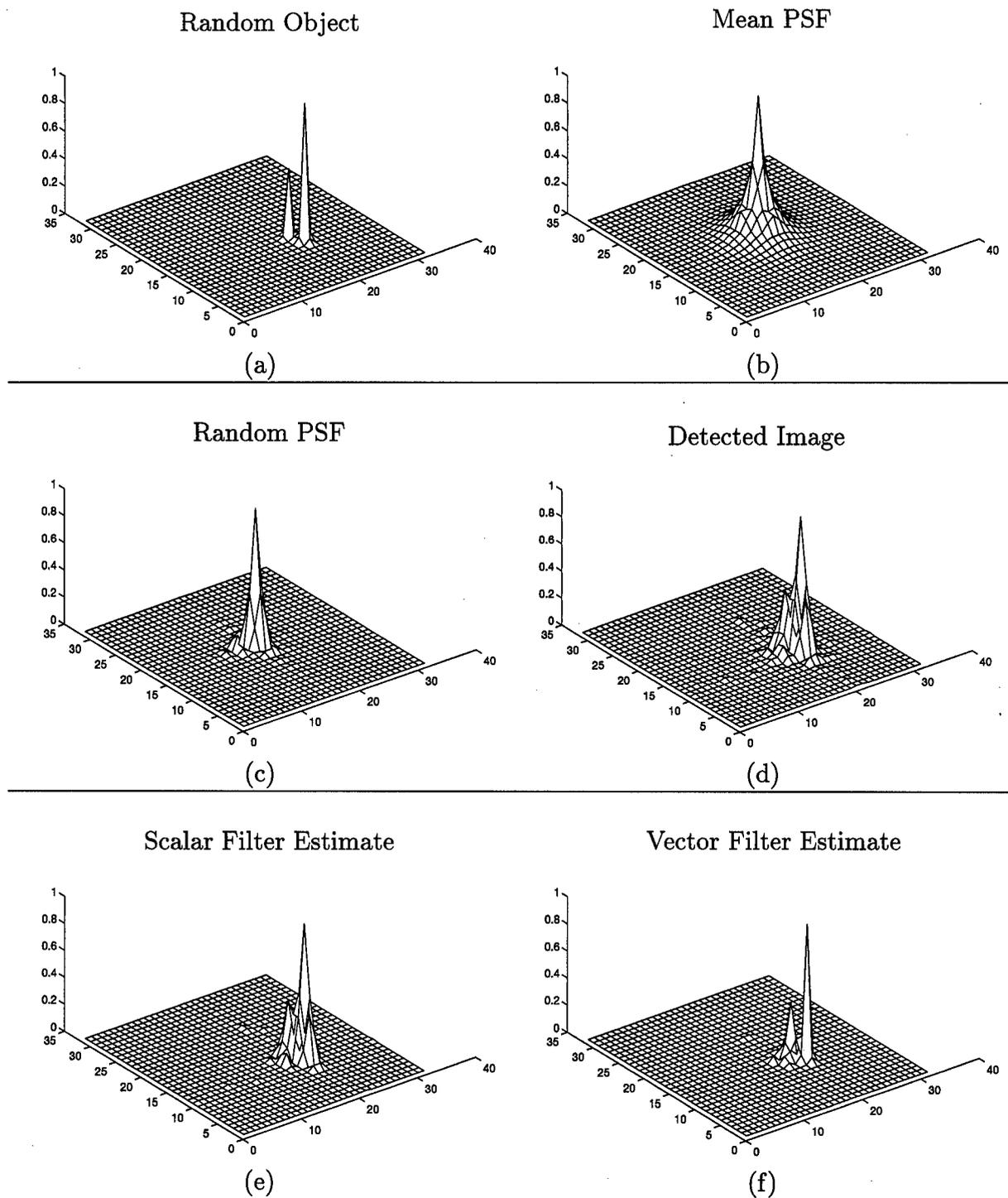


Figure 5.20 Normalized mesh plots showing the improved performance characteristics of the vector Wiener filter on photon-limited binary star objects degraded by atmospheric turbulence. (a) True object realization, (b) mean PSF ($D/r_o = 4$), (c) random PSF, (d) detected image ($\bar{K} = 1000$ photoevents), (e) scalar filter estimate, (f) vector filter estimate.

$\text{SNR}_{\mathcal{H}}(\rho) < 1$. Finally, individual filter output realizations were presented to demonstrate vector Wiener filter capabilities graphically.

The object irradiance distribution was assumed to be a random process with known first and second order Fourier domain statistics. Clearly, the vector Wiener filter is not applicable in situations in which no information is available about the type of objects to be imaged. The next chapter investigates the performance and robustness of the vector Wiener filter for generalized object and OTF statistical models. The objective is to study conditions under which the new filter is most useful.

VI. Vector Wiener Filter Performance and Robustness Study

6.1 Introduction

In Chapter V, vector Wiener filter performance was investigated for an important astronomical object class, the binary star pair. In that case, the new filter outperformed the scalar Wiener filter over a wide range of imaging conditions. However, these conclusions are limited since only binary stars and optical transfer functions representing atmospheric turbulence were examined. In this chapter, performance and filter robustness will be investigated for generalized object and optical transfer function (OTF) models. The objective is to draw conclusions about the application of the vector Wiener filter in general imaging scenarios. The following questions are of primary interest:

1. How do the object statistics limit filter performance?
2. How do the OTF statistics limit filter performance?
3. How robust is the vector Wiener filter to errors in the object statistical model?

The statistical models and resultant mean square error (MSE) data are based on a 16×16 pixel image array in all cases.

This chapter is organized as follows. Section 6.2 presents the object and OTF generalized statistical models. Section 6.3 reviews the filter expressions and relevant MSE metrics. Section 6.4 gives performance data as key object and OTF model parameters are varied. Here, all statistical models are assumed to be accurate and known. Section 6.5 gives robustness data when error is introduced to key object model parameters. Finally, Section 6.5 summarizes the chapter and highlights the major conclusions.

6.2 Generalized Models

To pursue this study, object and OTF statistical models are used which can represent a variety of imaging scenarios. These models are based on parameters that control the mean and covariance of the respective random process. The discussion below defines object and OTF generalized statistical models, key parameters, and supporting assumptions.

6.2.1 *Object.* An object statistical model for use in the vector Wiener filter is defined by the Fourier domain mean and covariance, denoted $\bar{\mathbf{O}}(u, v)$ and $C_{OO}(u, v; u', v')$, respectively. However, Fourier domain quantities can be difficult to visualize. Thus, the generalized object model used in this study is defined via the image domain statistics $\bar{o}(x, y)$ and $C_{oo}(x, y; x', y')$. These quantities are then transformed to the Fourier domain using the vector-matrix expressions [64]

$$\bar{\mathbf{O}} = F\bar{\mathbf{o}}, \quad (6.1)$$

and

$$C_{OO} = FC_{oo}F^H, \quad (6.2)$$

where $\bar{\mathbf{O}}$ and $\bar{\mathbf{o}}$ are $P \times 1$ ordered column vectors representing Fourier and image domain mean objects, respectively; C_{OO} and C_{oo} are $P \times P$ Fourier and image domain object covariance matrices, respectively; F is the $P \times P$ Fourier transformation matrix; H denotes a matrix Hermitian transpose; and P is the total number of pixels in the image array. An arbitrary element of the matrix F is defined as [64]

$$F(x, y; u, v) = \exp \left\{ -\frac{j2\pi}{N}(ux + vy) \right\}, \quad (6.3)$$

where an $N \times N$ square detector array is assumed and $N = \sqrt{P}$.

With the relationship between image and Fourier domain statistics defined, the following important question remains: "What is a representative object model?" In general, the mean object provides low spatial frequency information about the object class. A Gaussian function, centered at the origin, can represent this low pass characteristic. As noted in the previous chapter, support constraints have been widely used in image reconstruction algorithms [41, 54]. A mean object based on a Gaussian function and incorporating a support constraint can be written as

$$\bar{o}(x, y) = \begin{cases} \exp \{ -\pi ((x^2 + y^2)/\alpha_o^2 w_o^2) \} & (x, y) \in \text{support region} \\ 0 & \text{else} \end{cases}, \quad (6.4)$$

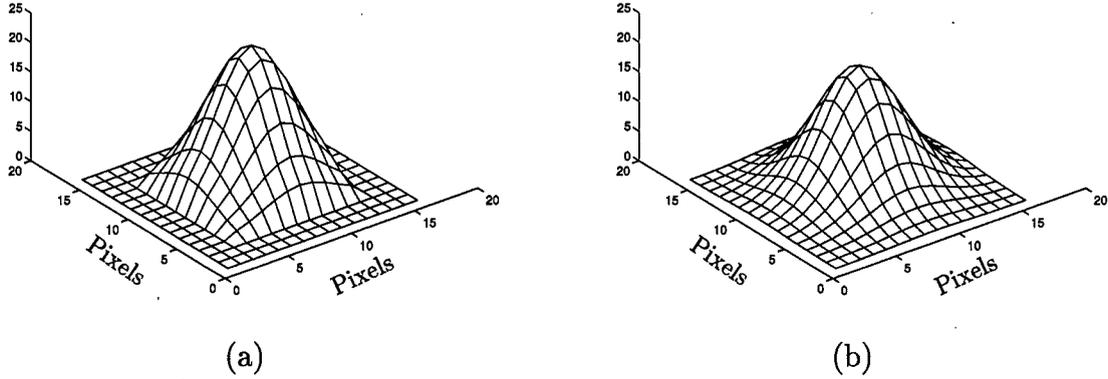


Figure 6.1 Mesh plots showing mean object, $\bar{o}(x, y)$, for mean object width parameter, $w_o = 0.5$. Support constraint dimension (a) $W = 10$ pixels, (b) $W = 16$ pixels.

where the support region is a $W \times W$ pixel rectangular area in the center of the image array, α_o is a mean object normalization constant, and w_o is the mean object width parameter. The mean object normalization constant was chosen such that the e^{-1} width of the Gaussian function is $N/2$ pixels when $w_o = 0.5$. In addition, the mean object is normalized such that

$$\sum_{x,y} \bar{o}(x, y) = 1000, \quad (6.5)$$

which is associated with an average photon count $\bar{K} = 1000$ photoevents and a moderate light level as noted in Chapter V. Figure 6.1 illustrates this mean object model for $w_o = 0.5$. The support constraint dimension is $W = 10$ pixels in (a) and $W = 16$ pixels in (b).

Clearly, the object covariance will vary widely depending on the particular imaging problem. However, standard image covariance models do exist [41]. The object covariance used in this study is based on a two dimensional, first order Markov model and is defined as

$$C_{oo}(x, y; x', y') = \begin{cases} \sigma_o(x, y)\sigma_o(x', y')\rho_o^d & (x, y) \text{ and } (x', y') \in \text{support region} \\ 0 & \text{else} \end{cases}, \quad (6.6)$$

where $\sigma_o(x, y)$ is the object standard deviation at the pixel (x, y) , ρ_o is the object correlation coefficient, and $d = \sqrt{(x - x')^2 + (y - y')^2}$ is the Euclidean distance between arbitrary pixels.

Table 6.1 Object Model Parameters

| <i>Parameter</i> | <i>Description</i> | <i>Default</i> | <i>High</i> | <i>Low</i> |
|------------------|-------------------------|----------------|-------------|------------|
| W | Support Dimension | 10 | 16 | N/A |
| w_o | Mean Width | 0.5 | 0.9 | 0.1 |
| SNR_o | Image Domain SNR | 2 or 5 | 10 | 0.5 or 2 |
| ρ_o | Correlation Coefficient | 0.5 | 0.9 | 0.1 |

The object standard deviation $\sigma_o(x, y)$ is determined based on an image domain signal-to-noise ratio (SNR) metric such that

$$\sigma_o(x, y) = \begin{cases} \bar{o}(x, y)/\text{SNR}_o & (x, y) \in \text{support region} \\ 0 & \text{else} \end{cases}, \quad (6.7)$$

where SNR_o is assumed to be uniform across all object pixels. The incorporation of a support constraint in the model ensures that the object covariance is non-stationary. The object covariance is also controlled by the SNR_o and ρ_o parameters. As SNR_o increases, the model provides more precise information about the object realizations. As $\rho_o \rightarrow 1$, the object pixels are more highly correlated. Table 6.1 lists the key object model parameters including values which represent high, low, and default cases in the performance and robustness studies.

6.2.2 Optical Transfer Function. The object statistics described in the previous section are directly dependent on model parameters as listed in Table 6.1. Once the object model parameters are selected, the mean and covariance are transformed to the Fourier domain as described in Eqs. (6.1) and (6.2). A similar model is also required for the OTF. In this case, the statistics will be defined directly in the Fourier domain. Both the mean and covariance elements are assumed to be real numbers. This assumption is consistent with an OTF ensemble representing imaging through atmospheric turbulence [74].

The mean OTF is a low pass filter which attenuates high spatial frequencies, yielding a blurred image. As noted above, this characteristic can be represented using a Gaussian

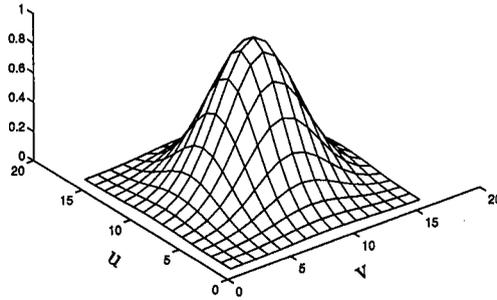


Figure 6.2 Mesh plot showing the mean OTF, $\bar{\mathcal{H}}(u, v)$, for mean OTF width parameter, $w_{\mathcal{H}} = 0.5$.

function. Thus, the mean OTF is defined as

$$\bar{\mathcal{H}}(u, v) = \exp \left\{ -\pi \left((u^2 + v^2) / \alpha_{\mathcal{H}}^2 w_{\mathcal{H}}^2 \right) \right\}, \quad (6.8)$$

where $\alpha_{\mathcal{H}}$ is a mean OTF normalization constant and $w_{\mathcal{H}}$ is the mean OTF width parameter. As with α_o above, $\alpha_{\mathcal{H}}$ was chosen such that the e^{-1} width of the Gaussian function is $N/2$ pixels when $w_{\mathcal{H}} = 0.5$. Figure 6.2 illustrates the mean OTF when $w_{\mathcal{H}} = 0.5$.

The OTF covariance model is the same as the object covariance model without a support constraint. Thus, $C_{\mathcal{H}\mathcal{H}}(u, v; u', v')$ is defined as

$$C_{\mathcal{H}\mathcal{H}}(u, v; u', v') = \sigma_{\mathcal{H}}(u, v) \sigma_{\mathcal{H}}(u', v') \rho_{\mathcal{H}}^d, \quad (6.9)$$

where $\sigma_{\mathcal{H}}(u, v)$ is the standard deviation of the OTF at the spatial frequency (u, v) , $\rho_{\mathcal{H}}$ is the OTF correlation coefficient, and $d = \sqrt{(u - u')^2 + (v - v')^2}$ measures the relative distance between arbitrary spatial frequencies. In the previous chapter, the OTF SNR, $\text{SNR}_{\mathcal{H}}$, was used to predict performance limits associated with processing binary star images collected through atmospheric turbulence. One objective of this chapter is to study performance further with respect to the $\text{SNR}_{\mathcal{H}}$ metric. Thus, $\text{SNR}_{\mathcal{H}}$ will be used to define the OTF

standard deviation in Eq. (6.9) via a simple manipulation of Eq. (5.11) which yields

$$\sigma_{\mathcal{H}}(u, v) = \frac{|\overline{\mathcal{H}}(u, v)|}{\text{SNR}_{\mathcal{H}}(u, v)}. \quad (6.10)$$

As illustrated in Fig. 5.13, a realistic $\text{SNR}_{\mathcal{H}}$ function is infinite at $(u, v) = (0, 0)$ which reflects a non-random DC value. The function then rolls off to a smaller value at a higher spatial frequency depending on the attenuation of $\overline{\mathcal{H}}(u, v)$ and the relative value of $\sigma_{\mathcal{H}}(u, v)$. For the purposes of this study, it is assumed that $\sigma_{\mathcal{H}}(u, v) < 1$ for all spatial frequencies (u, v) . This assumption is consistent with a realistic random OTF realization as defined in Eq. (2.17). Based on these characteristics and assumptions, $\text{SNR}_{\mathcal{H}}$ is modeled via the function

$$\text{SNR}_{\mathcal{H}}(u, v) = \frac{\beta_{\mathcal{H}}\mu_{\mathcal{H}}}{u^2 + v^2} + \overline{\mathcal{H}}(u, v), \quad (6.11)$$

where $\beta_{\mathcal{H}}$ is an OTF SNR normalization constant and $\mu_{\mathcal{H}}$ is an OTF SNR roll-off parameter. The OTF SNR normalization constant $\beta_{\mathcal{H}}$ was chosen such that Eq. (6.11) is unity at one-half the maximum radial frequency when $\mu_{\mathcal{H}} = 0.5$. Thus, $\mu_{\mathcal{H}}$ controls the spatial frequency at which $\text{SNR}_{\mathcal{H}}$ falls below unity as illustrated in Fig. 6.3.

Table 6.2 lists the key OTF model parameters including values which represent high, low, and default cases in the performance and robustness studies. The default value $w_{\mathcal{H}} = 0.5$ represents moderate attenuation of the object high spatial frequencies by the mean OTF. OTF covariance associated with the default values $\mu_{\mathcal{H}} = 1.0$ and $\rho_{\mathcal{H}} = 0.5$ represent moderate randomness and correlation. The term moderate refers to blur imposed by imaging through the turbulent atmosphere under good seeing conditions. This information, as well as the object parameters in Table 6.1, are an important reference to the reader with respect to the performance data given in Section 6.4. The next section reviews the filter expressions and relevant MSE metrics used in Sections 6.4 and 6.5.

6.3 Filter Expressions and Mean Square Error Metrics

As noted in the previous section, the mean object photon count is fixed at $\overline{K} = 1000$ photoevents. The impact of changing light level is well understood, based on the data and

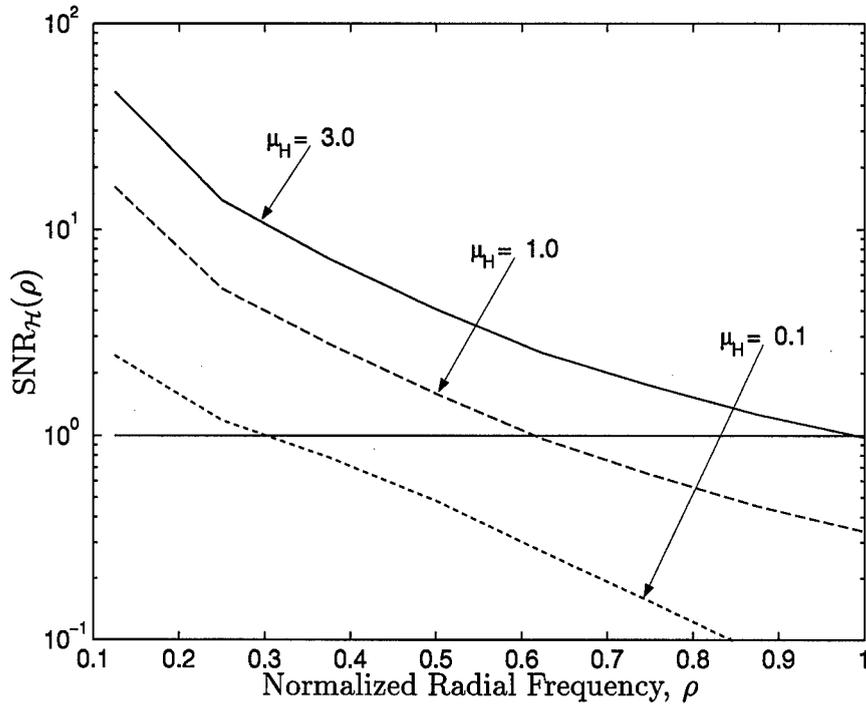


Figure 6.3 OTF SNR, $SNR_{\mathcal{H}}$, versus radial spatial frequency, ρ . The OTF SNR roll-off parameter, $\mu_{\mathcal{H}}$, controls the spatial frequency at which the $SNR_{\mathcal{H}}$ function falls below unity.

Table 6.2 Optical Transfer Function Model Parameters

| <i>Parameter</i> | <i>Description</i> | <i>Default</i> | <i>High</i> | <i>Low</i> |
|----------------------|-------------------------|----------------|-------------|------------|
| $w_{\mathcal{H}}$ | Mean Width | 0.5 | 0.9 | 0.1 |
| $\mu_{\mathcal{H}}$ | SNR Roll-Off | 1.0 | 3.0 | 0.1 |
| $\rho_{\mathcal{H}}$ | Correlation Coefficient | 0.5 | 0.9 | 0.1 |

discussion in Chapter V. Thus, light level is not a variable in these studies and a general form of the vector Wiener filter, defined as

$$\hat{\mathbf{O}} = R_{OO} \bar{\mathcal{H}}_d^* \{R_{OO} \odot R_{\mathcal{H}\mathcal{H}} + R_{NN}\}^{-1} \mathbf{D}, \quad (6.12)$$

is applicable. Equation (6.12) is equivalent to Eq. (4.28) since $R_{O_n O_n} = R_{OO}/(\bar{K})^2$. All image realizations are assumed to be photon-limited for these studies. Thus, $\sigma_r = 0$ electrons per pixel and an arbitrary element of the Fourier domain noise autocorrelation matrix is

$$R_{NN}(u, v; u', v') = \bar{\mathcal{H}}(u - u', v - v') \bar{\mathcal{O}}(u - u', v - v'). \quad (6.13)$$

The analogous scalar Wiener filter is defined via Eq. (4.40) as before.

In Chapter V, two MSE metrics were used to compare the performance of the vector and scalar Wiener filters. In this chapter, only average MSE is used. This approach supports efficient presentation of large amounts of data. Thus, in this chapter, the term MSE refers exclusively to average MSE as defined in Eq. (5.4). The required theoretical error correlation matrices are generated using Eqs. (4.39), (4.41), and (4.42). In addition, vector Wiener filter MSE is normalized with respect to the scalar Wiener filter and detected image MSE. As in Chapter V, the subscripts V , S , and D denote MSE associated with the vector Wiener filter, scalar Wiener filter, and detected image, respectively.

6.4 Performance Study

The objective of this study is to establish quantitative limits on vector Wiener filter performance as key object and OTF model parameters are varied to represent a variety of imaging conditions. Here, all statistical models are assumed accurate and known without error as the parameters change. In other words, the filters have knowledge of the true imaging scenario and are adjusted accordingly. Section 6.4.1 outlines the important assumptions associated with this work while Section 6.4.2 presents both object and OTF model data.

6.4.1 Assumptions. There are two additional assumptions associated with the object and OTF statistical models which pertain to the performance study. First, the mean object width parameter is fixed at $w_o = 0.5$. This assumption is valid since a change in mean object size in the image domain can be viewed as altering the optical system magnification. Second, the mean OTF width parameter is fixed at $w_{\mathcal{H}} = 0.5$ for all performance study data. The role of the mean OTF in filter performance is relatively straightforward as established in Chapter V. Thus, the second order statistical quantities are of primary concern in this study.

6.4.2 Data. Based on the assumptions noted above, this study will concentrate on object and OTF statistical model parameters which impact the covariance. The important object parameters are the support constraint dimension, W , the SNR parameter, SNR_o , and the correlation coefficient, ρ_o . The following plots will illustrate typical vector Wiener filter performance for a variety of imaging conditions.

Figure 6.4 shows vector Wiener filter MSE, normalized with respect to the scalar Wiener filter MSE, as a function of the support constraint dimension, W . Here, $\rho_o = 0.5$ and the OTF is non-random ($\text{SNR}_{\mathcal{H}} = \infty$ for all spatial frequencies (u, v)). The data shows that vector filter normalized MSE increases as W increases and SNR_o decreases. In Chapter V, it was shown that a rectangular support constraint has an important impact on vector filter performance for binary star pairs. Support constraints provide variance reduction by maintaining Fourier domain correlations in the data which provide weighted interpixel averaging [54]. This data clearly illustrates the value of interpixel averaging in improving vector filter MSE when object SNR is low. In the $\text{SNR}_o = 2$ case, performance improvement over the scalar filter is less than 10% when no support constraint is used. When $W = 8$ pixels, this factor improves to more than 20%. Also note the importance of SNR_o on vector filter MSE. Figure 6.4 shows that the vector filter provides over 40% less MSE than the scalar filter when $\text{SNR}_o = 10$ and no support constraint is used.

Figure 6.4 illustrated the importance of object support on filter MSE performance, especially when W is small. Many objects cannot be confined to a small support region and are better represented by a larger W value. Thus, the support constraint dimension is fixed

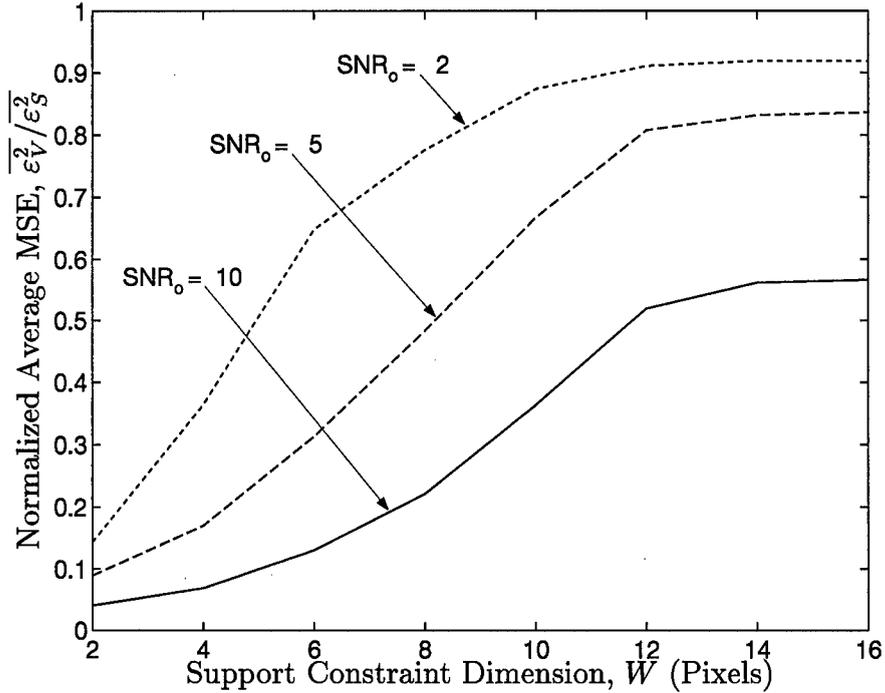
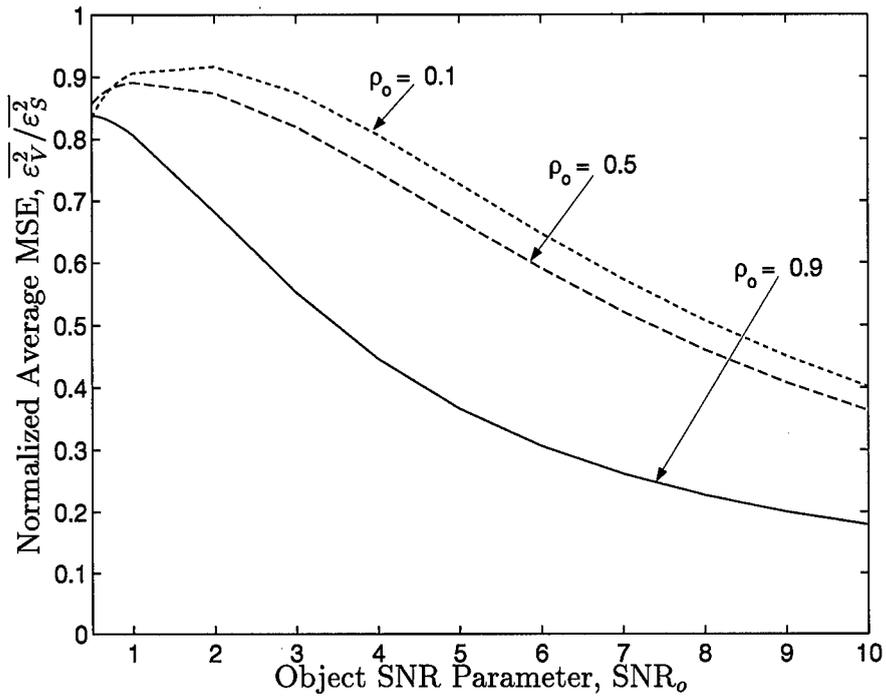
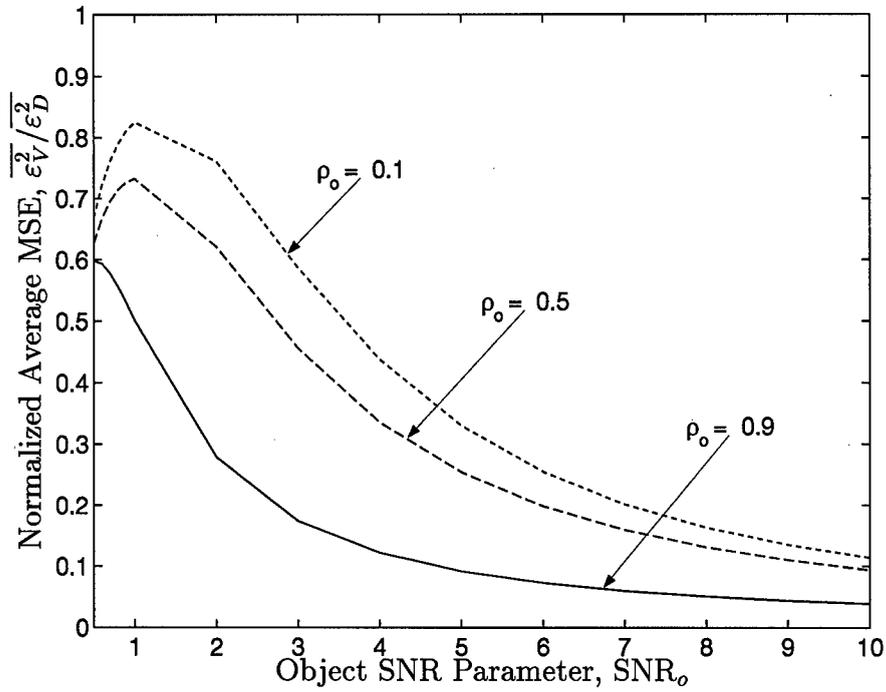


Figure 6.4 Normalized average MSE, $\overline{\varepsilon_V^2}/\overline{\varepsilon_S^2}$, versus support constraint dimension, W , $\rho_o = 0.5$. The OTF is non-random ($\text{SNR}_{\mathcal{H}} = \infty$ for all spatial frequencies (u, v)).

at $W = 10$ pixels for the remaining data plots to represent a typical object class. Figure 6.5 gives normalized MSE as a function of the object SNR parameter, SNR_o . Once again, the OTF is non-random ($\text{SNR}_{\mathcal{H}} = \infty$ for all spatial frequencies (u, v)). First, consider the MSE normalized with respect to the scalar filter in (a). Here, vector filter MSE approaches to within 10% of the scalar filter when $\rho_o = 0.1$. Note that the normalized MSE begins to decrease slightly when $\text{SNR}_o < 2$ and $\rho_o = 0.1$ or 0.5 . When $\text{SNR}_o > 2$, both filters perform well, with the vector filter providing greater MSE improvement as SNR_o increases. When $\text{SNR}_o < 2$, MSE increases significantly for both filters, but the scalar filter performance degrades faster than the vector filter. Thus, the object spatial SNR provides a natural performance threshold in these cases. The data in (a) also shows that a high object correlation coefficient can have a substantial effect on MSE. When $\rho_o = 0.9$, the vector filter offers approximately 20% more MSE improvement over the scalar filter than in the $\rho_o = 0.1$ or 0.5 cases. In (b), the MSE is normalized with respect to the detected image. In almost all cases, the vector filter provides at least 20% improvement. As in (a), the MSE decreases when $\text{SNR}_o < 2$ and $\rho_o = 0.1$ or 0.5 . Here, the detected image MSE is increasing much



(a)



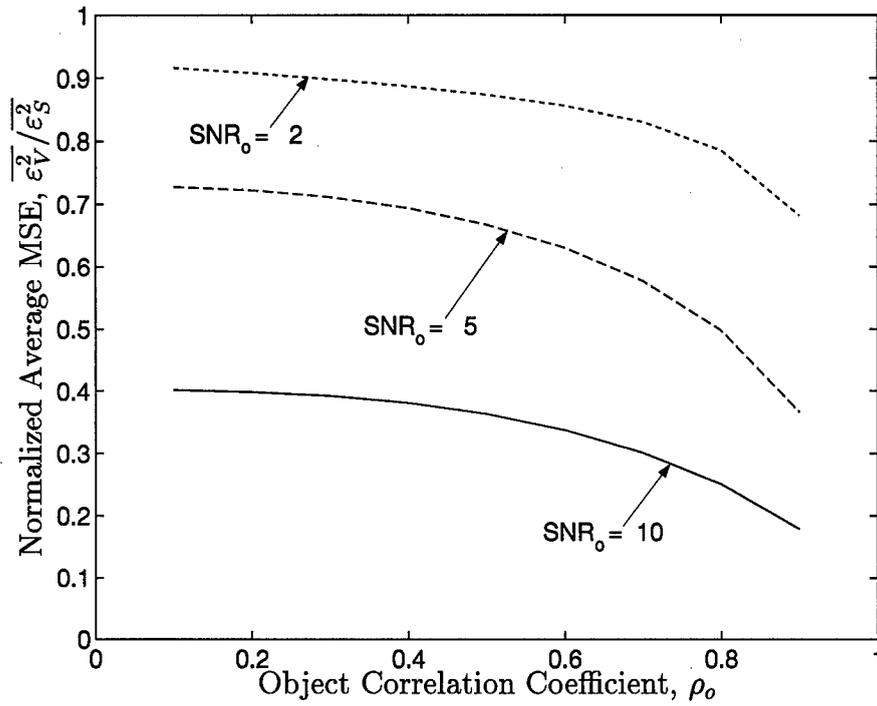
(b)

Figure 6.5 Normalized average MSE, (a) $\overline{\epsilon_V^2}/\overline{\epsilon_S^2}$, and (b) $\overline{\epsilon_V^2}/\overline{\epsilon_D^2}$, versus the object SNR parameter, SNR_o . The object support constraint dimension is fixed at $W = 10$ pixels. The OTF is non-random ($\text{SNR}_{\mathcal{H}} = \infty$ for all spatial frequencies (u, v)).

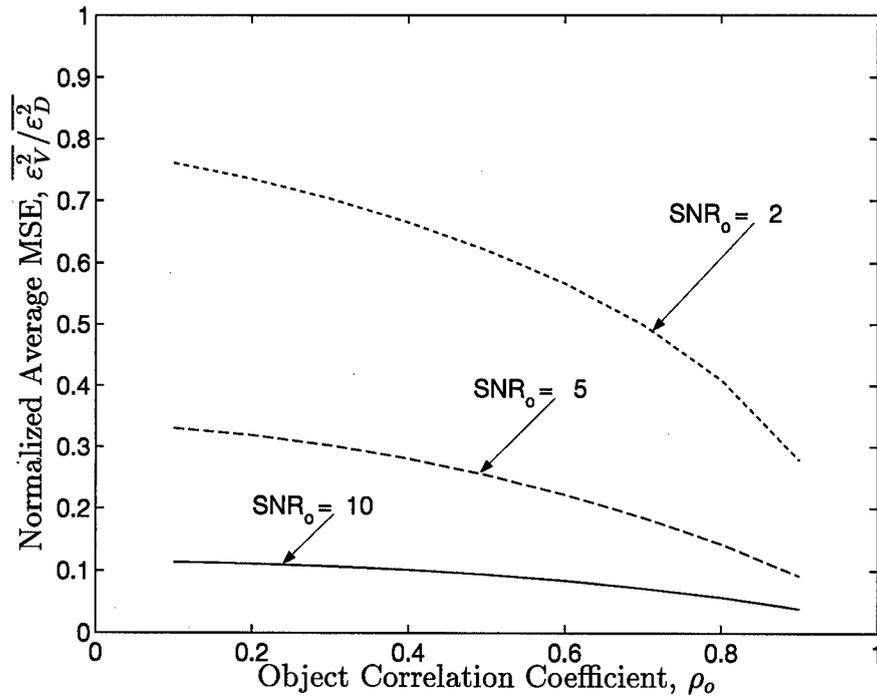
faster than the vector filter MSE. In general, the vector Wiener filter provides the best MSE performance when $\text{SNR}_o > 2$ and ρ_o is high. This MSE improvement can be as much as 80% over the scalar filter and 90% over the detected image when $\text{SNR}_o = 10$.

Now consider Fig. 6.6 in which normalized MSE is presented specifically as a function of the object correlation coefficient, ρ_o . As before, the OTF is non-random ($\text{SNR}_H = \infty$ for all spatial frequencies (u, v)). The data in (a) is normalized with respect to the scalar filter. Note that $\rho_o = 0.9$ provides 20-30% improvement over $\rho_o = 0.1$ regardless of the SNR_o case. Also, the object correlation coefficient seems to provide the greatest MSE performance boost over the scalar filter when $\rho_o > 0.5$. In (b), the vector filter MSE data is normalized with respect to the detected image. The plot shows that ρ_o has the most impact on MSE performance when SNR_o is low. High object SNR reduces the relative contribution of the object covariance function, $C_{oo}(x, y; x', y')$, to the object correlation function, $R_{oo}(x, y; x', y')$. Hence, ρ_o has little effect on vector filter MSE when $\text{SNR}_o = 10$ in (b). In general, low object SNR and $\rho_o > 0.5$ are conditions under which the object correlation coefficient has the greatest effect on filter MSE performance.

For the object model parameters investigated above, the OTF was non-random ($\text{SNR}_H = \infty$ for all spatial frequencies (u, v)). Now consider a random OTF and investigate filter MSE performance with respect to the OTF model parameters. The key parameters in this case are the OTF SNR roll-off parameter, μ_H , and the OTF correlation coefficient, ρ_H . Figure 6.7 presents MSE as a function of the OTF SNR roll-off parameter, μ_H . In (a), the vector Wiener filter MSE data is normalized with respect to the scalar filter with the object model parameters fixed at the default values ($W = 10$ pixels, $\text{SNR}_o = 2$, $\rho_o = 0.5$). Note that vector filter MSE performance improves greatly when μ_H decreases below 0.5. In fact, the vector filter MSE drops by approximately 50% in this case. At first glance, the data in (a) would seem to indicate that low OTF SNR improves filter MSE performance. In reality, both vector and scalar filters experience an increase in MSE as μ_H falls below this threshold. However, the scalar filter MSE increases more dramatically than the vector filter MSE as shown via the raw MSE data in (b). Here, only the $\rho_H = 0.5$ case is shown. The vector



(a)



(b)

Figure 6.6 Normalized average MSE, (a) $\overline{\varepsilon_V^2}/\overline{\varepsilon_S^2}$, and (b) $\overline{\varepsilon_V^2}/\overline{\varepsilon_D^2}$, versus the object correlation coefficient, ρ_o . The object support constraint dimension is fixed at $W = 10$ pixels. The OTF is non-random ($\text{SNR}_{\mathcal{H}} = \infty$ for all spatial frequencies (u, v)).

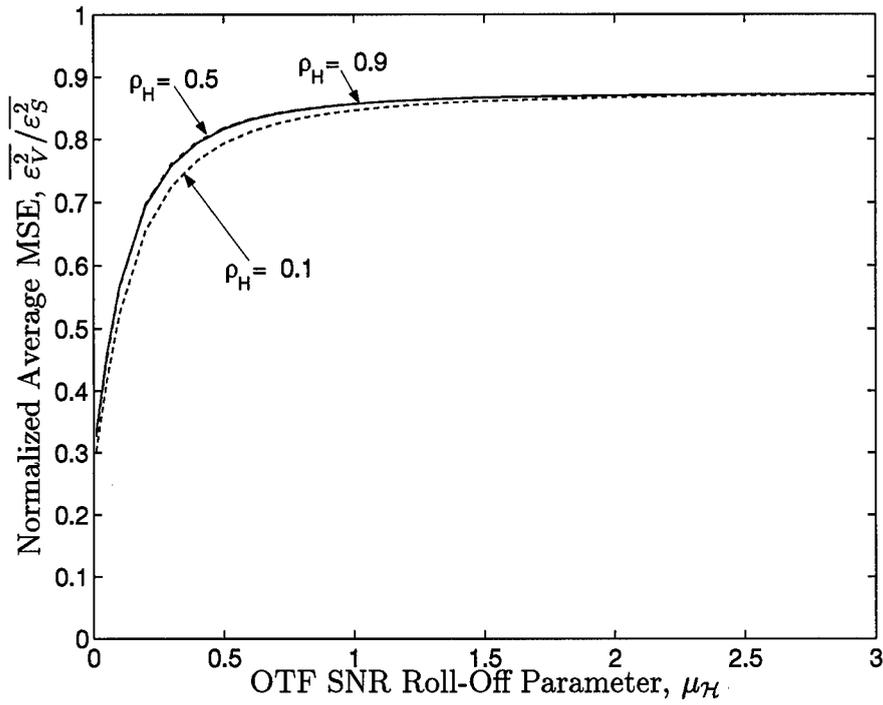
Wiener filter is able to maintain much better performance for low OTF SNR due to the a priori object information provided by the statistical model.

Now consider the same type of data in Fig. 6.8 in which the object model parameters are now less favorable to vector filter processing ($W = 16$ pixels, $\text{SNR}_o = 0.5$, $\rho_o = 0.5$) and $\rho_{\mathcal{H}} = 0.5$. As before, the vector filter MSE data is normalized with respect to the scalar filter in (a). Note that the MSE still decreases for low $\mu_{\mathcal{H}}$ values. However, the effect is minimal compared to the previous plots. As shown via the raw data in (b), the vector filter MSE increases just as fast as the scalar filter with decreasing $\mu_{\mathcal{H}}$. In this case, the object spatial SNR is below the performance threshold established earlier and the support constraint has been removed. Thus, the object model cannot compensate for the increased randomness in the OTF. In general, the vector Wiener filter provides the best MSE performance when $\mu_{\mathcal{H}} > 0.5$. This performance threshold corresponds with the OTF SNR function falling below unity at normalized radial frequency $\rho = 0.5$. Thus, this data seems to be in general agreement with the binary star data in Chapter V. For a typical object class, vector Wiener filter MSE increases dramatically when $\text{SNR}_{\mathcal{H}}$ falls below unity at the mid and high spatial frequencies.

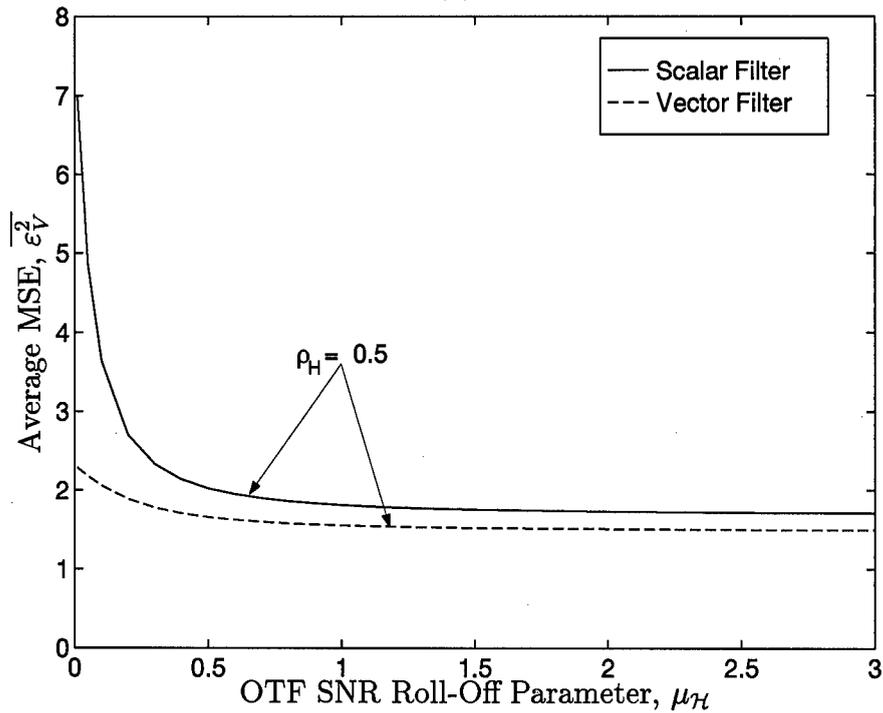
In Fig. 6.7 (a), changing the OTF correlation coefficient seemed to have little impact on filter MSE. This performance trend is confirmed in Fig. 6.9 in which normalized MSE is given versus the OTF correlation coefficient, $\rho_{\mathcal{H}}$. In (a) and (b), the vector Wiener filter MSE is normalized with respect to the scalar filter and detected image, respectively. Regardless of the $\mu_{\mathcal{H}}$ value, the plot traces are almost flat. This performance is relatively consistent across a variety of object classes. Thus, $\rho_{\mathcal{H}}$ has minimal impact on filter performance.

6.5 Robustness Study

In some applications, the statistical models may only approximate the true imaging scenario. This situation is usually caused by a lack of a priori knowledge about the problem, especially with regard to the object model. The objective of this study is to investigate the robustness of the vector Wiener filter to error in key object model parameters. The required

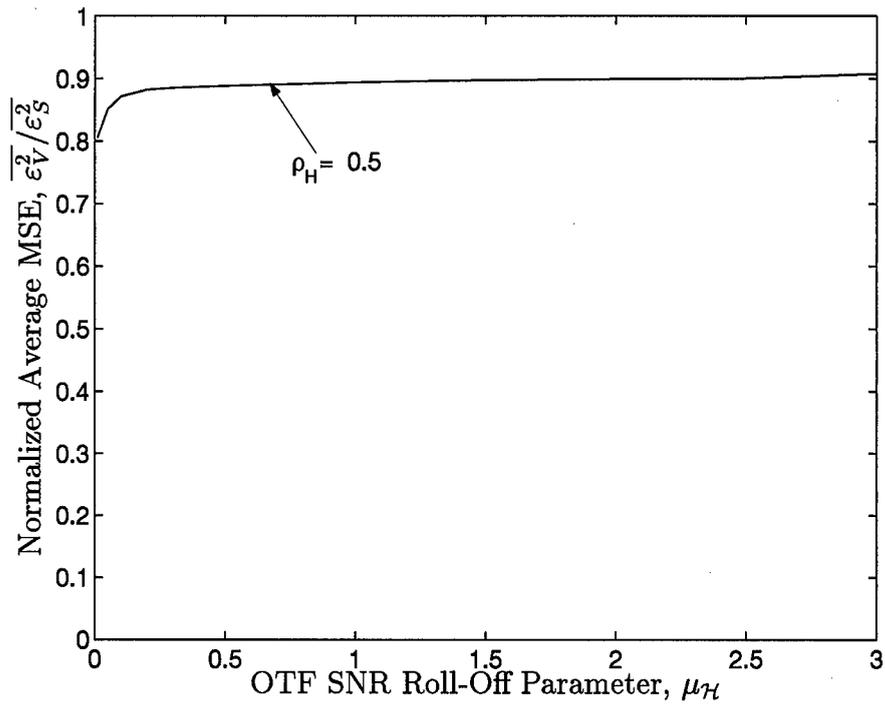


(a)

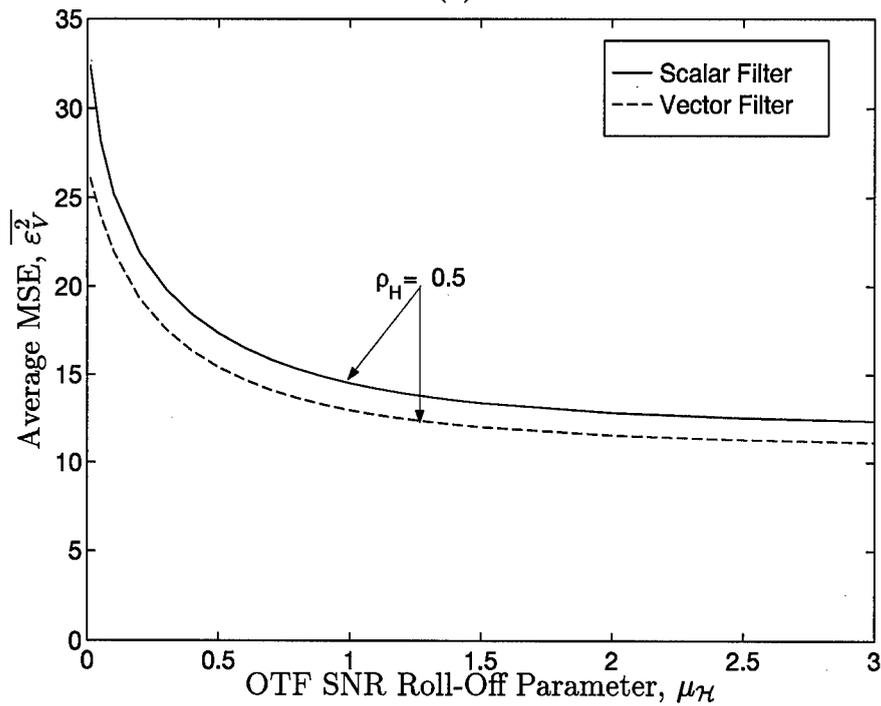


(b)

Figure 6.7 (a) Normalized average MSE, $\overline{\varepsilon_V^2}/\varepsilon_S^2$, and (b) raw average MSE data, $\overline{\varepsilon_V^2}$, versus the OTF SNR roll-off parameter, $\mu_{\mathcal{H}}$. Object model parameters: $W = 10$ pixels, $\text{SNR}_o = 2$, $\rho_o = 0.5$.

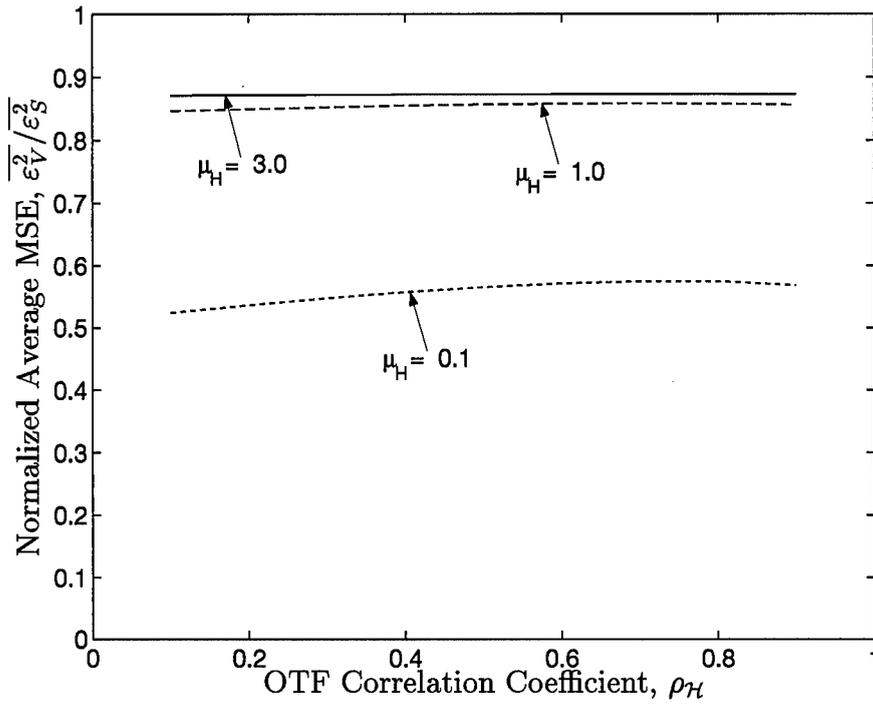


(a)

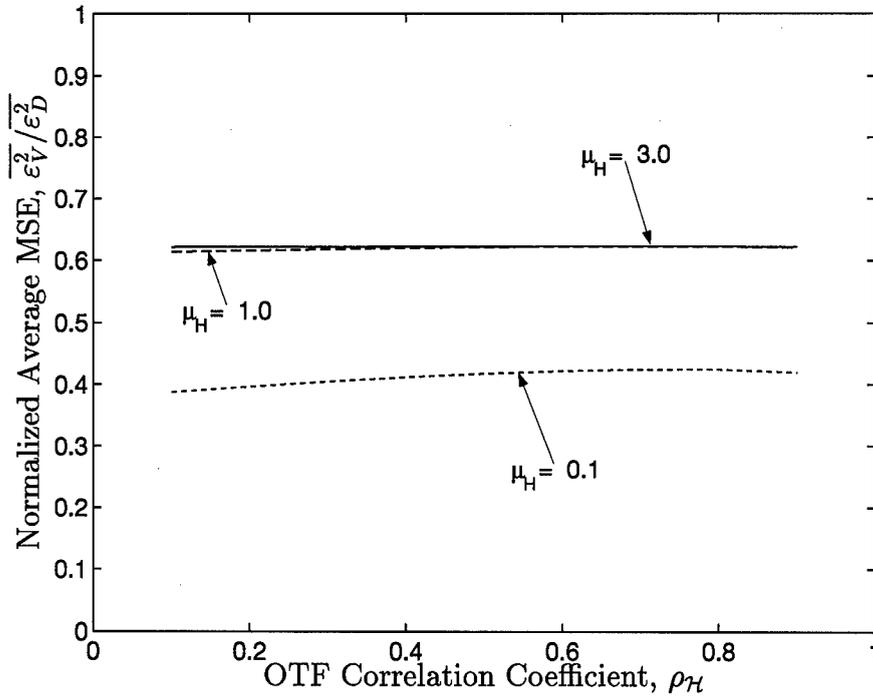


(b)

Figure 6.8 (a) Normalized average MSE, $\overline{\epsilon_V^2}/\overline{\epsilon_S^2}$, and (b) raw average MSE data, $\overline{\epsilon_V^2}$, versus the OTF SNR roll-off parameter, μ_H . Object model parameters: $W = 16$ pixels, $\text{SNR}_o = 0.5$, $\rho_o = 0.5$.



(a)



(b)

Figure 6.9 Normalized average MSE, (a) $\overline{\epsilon_V^2}/\overline{\epsilon_S^2}$, and (b) $\overline{\epsilon_V^2}/\overline{\epsilon_D^2}$, versus the OTF correlation coefficient, ρ_H . Object model parameters: $W = 10$ pixels, $\text{SNR}_o = 2$, $\rho_o = 0.5$.

theoretical error correlation matrix can be written as

$$R_{\epsilon\epsilon} = R_{OO} - M_w R_{OD}^H - R_{OD} M_w^H + M_w R_{DD} M_w^H, \quad (6.14)$$

where R_{OO} , R_{OD} , and R_{DD} are based on the true model parameters. The filter transformation matrix M_w is dependent on the corrupted model parameters. The rest of the section is outlined as follows. Section 6.5.1 gives the important assumptions associated with this work while Section 6.5.2 presents the data.

6.5.1 Assumptions. There are several additional assumptions associated with the object and OTF statistical models which pertain to the robustness study. First, it is assumed that the OTF model is known with reasonable accuracy. Thus, only the object model parameters are studied. This assumption is valid since an ensemble of bright star images is often used to obtain an accurate estimate of the atmospheric-optical system OTF for use in linear deconvolution [72,74]. These same point source images could also be used to estimate the OTF SNR. Therefore, the mean OTF width parameter and OTF correlation coefficient are fixed at $w_{\mathcal{H}} = 0.5$ and $\rho_{\mathcal{H}} = 0.5$, respectively. Second, the object support constraint dimension and the mean object width parameter are fixed at $W = 10$ pixels and $w_o = 0.5$, respectively. Here, it is assumed that these parameters are also known with reasonable accuracy. Thus, this robustness study will concentrate on the object SNR parameter, SNR_o , and the object correlation coefficient, ρ_o . These parameters are associated with the object covariance. It is this part of the object statistical model that may not be readily available in some imaging applications. Finally, percent error is used as an independent variable in the study. This metric is defined as

$$\text{Percent Error} = \left(\frac{p_w - p_t}{p_t} \right) \times 100, \quad (6.15)$$

where p_w is the corrupted model parameter and p_t is the true model parameter. In all cases shown below, the true parameter values are $\text{SNR}_o = 2$ and $\rho_o = 0.5$. Percent error is displayed from -100% to +100% in all cases. When the normalized MSE is unity, the vector

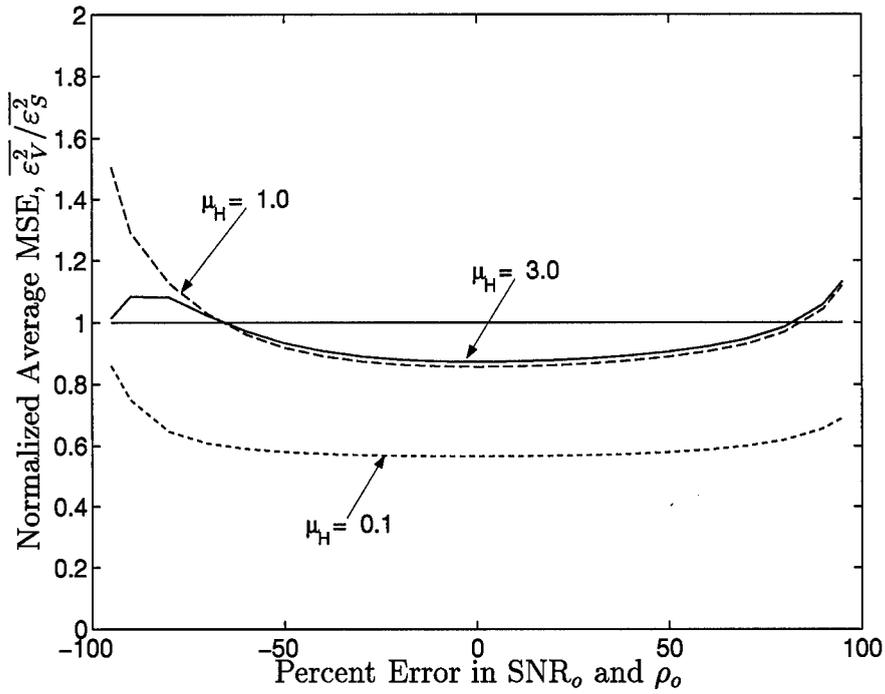
Wiener filter provides the same MSE as the scalar filter or detected image for that error magnitude.

6.5.2 Data. First, consider error in both the object SNR parameter, SNR_o , and the object correlation coefficient, ρ_o . Figure 6.10 gives normalized MSE as a function of percent error in both SNR_o and ρ_o . As in the previous study, the vector filter MSE is normalized with respect to the scalar filter in (a). Notice the substantial difference in performance between the $\mu_{\mathcal{H}} = 0.1$ and $\mu_{\mathcal{H}} = 1.0$ and 3.0 traces. This difference is associated with the OTF SNR performance threshold first established in Fig. 6.7. Clearly, vector filter performance degrades as error is introduced into these parameters. In fact, the MSE performance ratio between vector and scalar filters is greater than 90% when the error magnitude is greater than 60% in the $\mu_{\mathcal{H}} = 1.0$ and 3.0 cases. Now consider the vector filter MSE normalized with respect to the detected image in (b). The performance trends here are similar to (a) except the filter MSE is much larger when the parameters are underestimated in the $\mu_{\mathcal{H}} = 1.0$ and 3.0 cases.

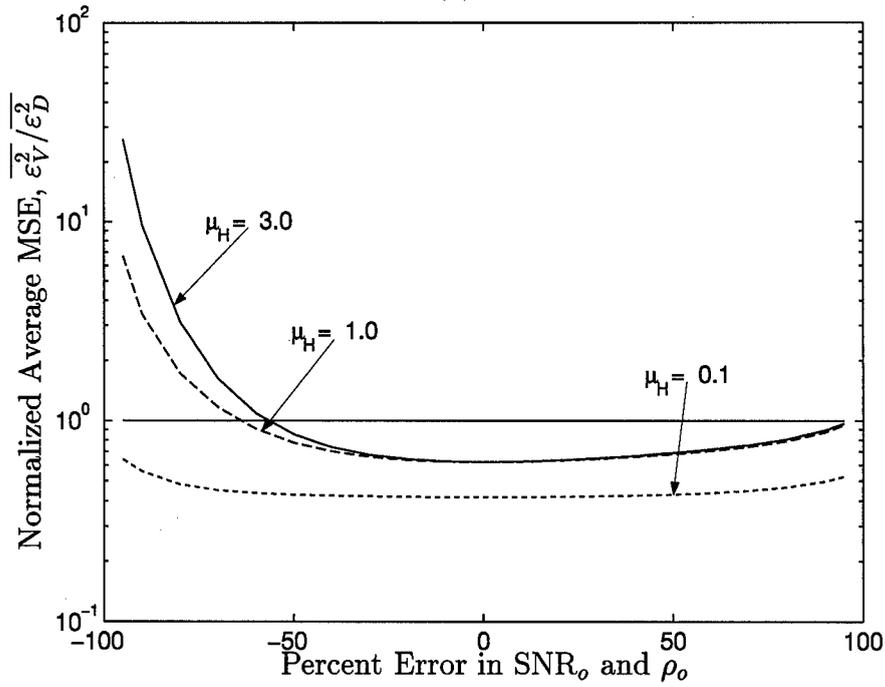
The previous plots revealed the sensitivity of the vector Wiener filter to simultaneous error in the two key parameters which control object covariance. Figures 6.11 and 6.12 show vector filter normalized MSE for error in only one parameter. In both figures, the plot at (a) is normalized with respect to the scalar filter and the plot at (b) is normalized with respect to the detected image. In each plot, only the $\mu_{\mathcal{H}} = 1.0$ case is shown. Clearly, the SNR_o parameter generates the largest proportion of the MSE shown in Fig. 6.10. Thus, SNR_o is more sensitive to error than ρ_o . In fact, Fig. 6.12 shows that error in the ρ_o parameter never degrades vector filter performance such that the normalized MSE is greater than unity.

6.6 Summary

In this chapter, filter performance and robustness were investigated for generalized object and OTF models. The objective was to draw quantitative conclusions about the application of the vector Wiener filter in general imaging scenarios. The performance study involved examining vector and scalar filter MSE performance as key statistical model parameters were varied. In all cases, the model parameters were assumed known without error.

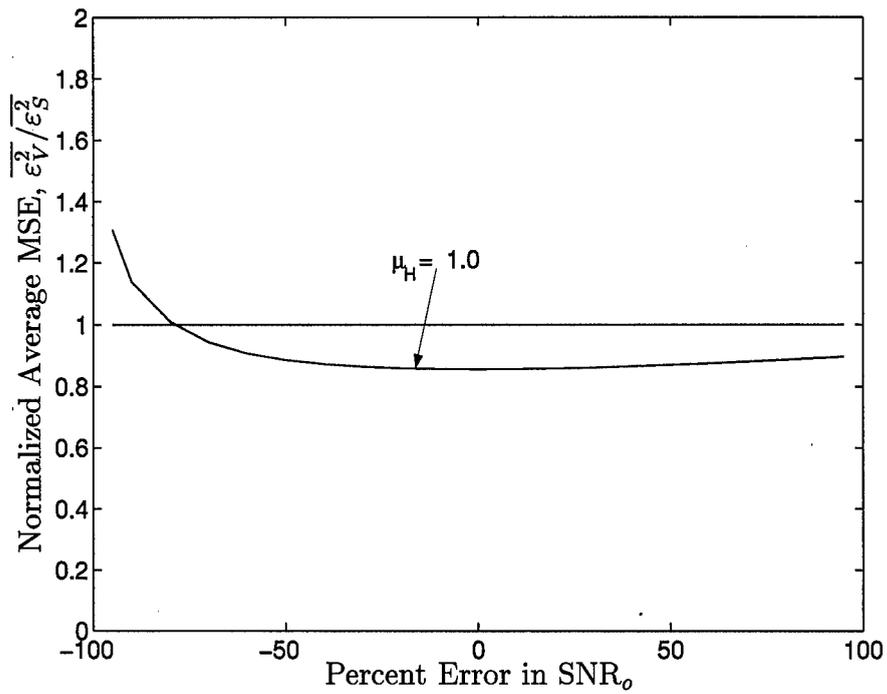


(a)

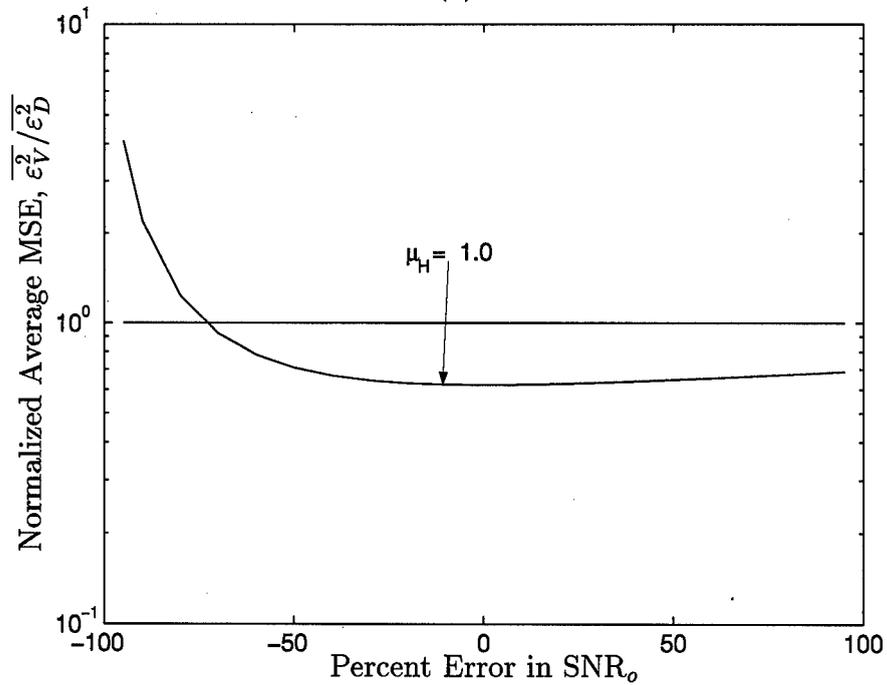


(b)

Figure 6.10 Normalized average MSE, (a) $\overline{\epsilon_V^2}/\overline{\epsilon_S^2}$, and (b) $\overline{\epsilon_V^2}/\overline{\epsilon_D^2}$, versus percent error in the object SNR parameter, SNR_o , and the object correlation coefficient, ρ_o . OTF model parameters: $w_H = 0.5$, $\rho_H = 0.5$. True parameter values: $\text{SNR}_o = 2$, $\rho_o = 0.5$.

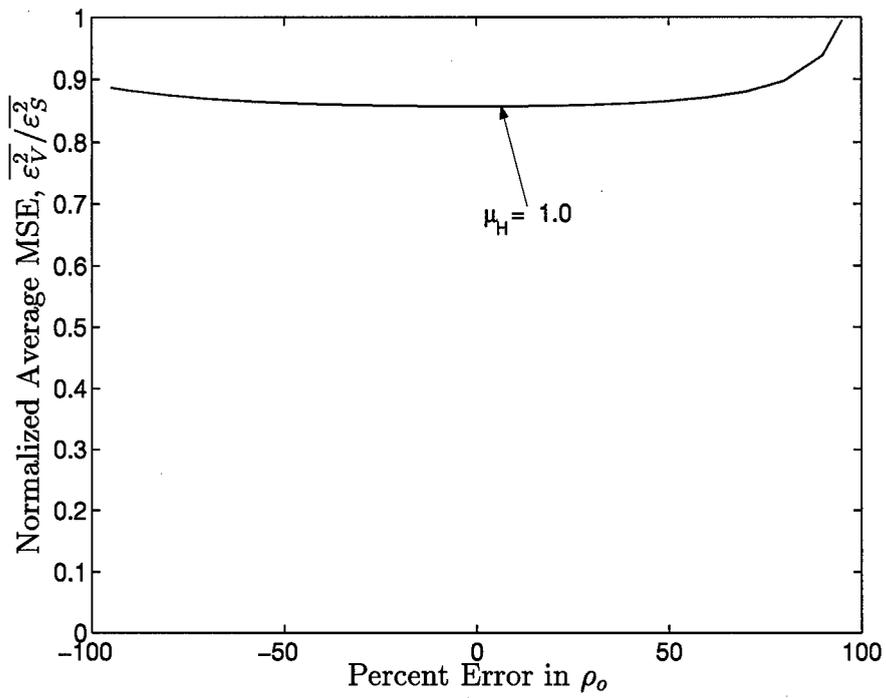


(a)

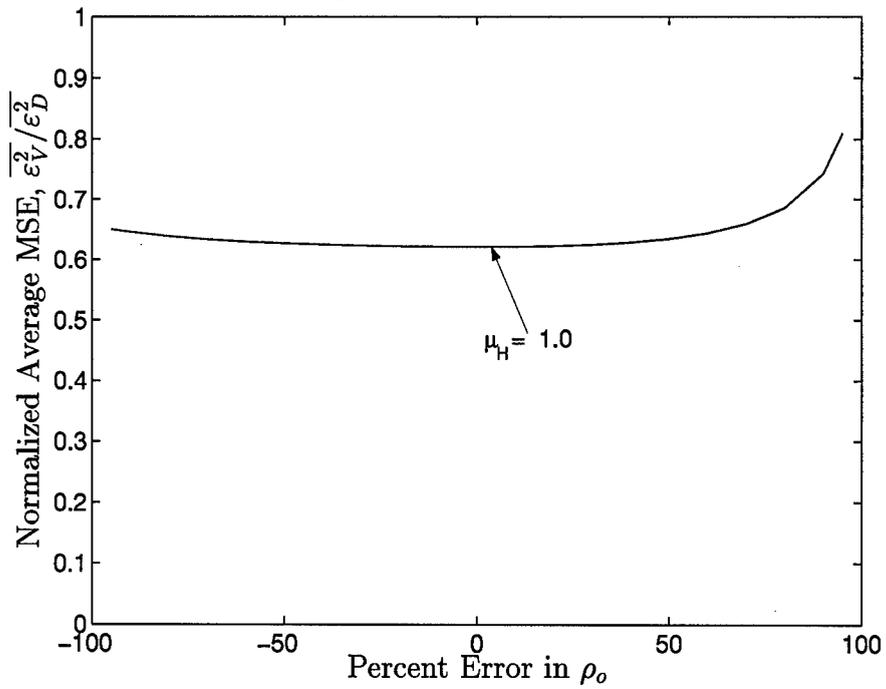


(b)

Figure 6.11 Normalized average MSE, (a) $\overline{\varepsilon_V^2}/\overline{\varepsilon_S^2}$, and (b) $\overline{\varepsilon_V^2}/\overline{\varepsilon_D^2}$, versus percent error in the object SNR parameter, SNR_o . OTF model parameters: $w_H = 0.5$, $\rho_H = 0.5$. True parameter values: $\text{SNR}_o = 2$, $\rho_o = 0.5$.



(a)



(b)

Figure 6.12 Normalized average MSE, (a) $\overline{\varepsilon_V^2}/\overline{\varepsilon_S^2}$, and (b) $\overline{\varepsilon_V^2}/\overline{\varepsilon_D^2}$, versus percent error in the object correlation coefficient, ρ_o . OTF model parameters: $w_{\mathcal{H}} = 0.5$, $\rho_{\mathcal{H}} = 0.5$. True parameter values: $\text{SNR}_o = 2$, $\rho_o = 0.5$.

In contrast, the robustness study involved introducing error to the filter-assumed object statistical model.

The results of the performance study showed that vector Wiener filter MSE performance can be limited by both object and OTF statistical models. An object support constraint is extremely advantageous to vector filter processing since interpixel averaging is enhanced. The object SNR parameter, SNR_o , and the object correlation coefficient, ρ_o , provide a fundamental limit on filter MSE. When $\text{SNR}_o < 2$ and $\rho_o < 0.5$, the vector filter provides only marginal improvement in MSE over that of the scalar filter. However, $\rho_o > 0.5$ can help compensate for low object SNR in many cases. The OTF SNR can also provide a limit on filter MSE performance for some object classes. In these cases, the OTF SNR roll-off parameter, $\mu_{\mathcal{H}}$, must be large enough to boost the OTF SNR above unity at the mid spatial frequencies. For the data shown here, the vector Wiener filter provides the best MSE performance when $\mu_{\mathcal{H}} > 0.5$. However, the vector Wiener filter can continue to perform well below this OTF SNR threshold if the object SNR is high. Finally, it was shown that the OTF correlation coefficient, $\rho_{\mathcal{H}}$, has minimal impact on filter normalized MSE.

The robustness study investigated the effect of error in the filter-assumed object model parameters. These results showed that the vector Wiener filter is less robust than the scalar Wiener filter with respect to these errors. This effect was anticipated, since the vector filter better exploits the real world information about the imaging scenario provided by more detailed statistical models. In general, simultaneous error in the SNR_o and ρ_o parameters resulted in marginal performance improvement over the scalar filter when the error magnitude was greater than 60%. The greatest impact on performance was associated with error in the object spatial SNR.

VII. *Conclusions and Recommendations*

7.1 *Major Results*

The major results of this research effort are the following:

1. Development of a new constrained least squares (CLS) algorithm for deconvolution from wavefront sensing (DWFS) processing of low light images.
 - Computationally inexpensive. In all cases examined, the Newton-Raphson iteration converged to a solution in less than 10 iterations.
 - Shot noise imposes the fundamental performance limit.
2. Derivation of a new vector Wiener filter incorporating the semi-classical model of photoelectric light detection.
 - Yields superior reconstructions with respect to mean square error (MSE) and mean square phase error (MSPE) when compared to the scalar Wiener filter for binary star objects.
 - Provides superresolution when the object's Fourier domain statistics are known for spatial frequencies beyond the optical transfer function (OTF) cutoff.
3. Quantitative results showing the performance and limitations of the vector Wiener filter when applied to binary star images degraded by atmospheric turbulence.
 - First application of second order OTF statistics between different spatial frequencies in a Wiener filter.
 - Filter MSE performance degraded for spatial frequencies at which the OTF signal-to-noise ratio (SNR) is less than unity.
4. Quantitative results showing performance limits on the vector Wiener filter associated with generalized object and OTF models.
 - The object SNR parameter, SNR_o , and the object correlation coefficient, ρ_o , provide a fundamental limit on filter MSE. When $\text{SNR}_o < 2$ and $\rho_o < 0.5$, the vector filter provides only marginal improvement in MSE over the scalar filter.

- For a typical object class, the OTF SNR must be above unity for normalized radial frequencies $\rho > 0.5$.
5. Quantitative results showing the robustness of the vector Wiener filter with respect to object model error.
- Simultaneous error in the object SNR parameter, SNR_o , and the object correlation coefficient, ρ_o , resulted in marginal performance improvement over the scalar filter when the error magnitude was greater than 60%.
 - The greatest impact on vector Wiener filter performance was associated with error in the object spatial SNR.

7.2 Discussion

The previous chapters introduced two new linear reconstruction techniques which complement existing linear filters and more intensive iterative optimization schemes. The first, CLS incorporating wavefront sensing, is practical for large image arrays and easy to apply when wavefront sensor (WFS) hardware is available. The algorithm is fundamentally limited by shot noise effects in the phase estimates. In addition, the CLS algorithm tends to underestimate the regularization constant for small data ensembles (< 50 images). The second technique, a new vector Wiener filter, offers superior performance over the existing scalar Wiener filter for non-stationary image ensembles. However, computational complexity severely limits the practical application of this filter, since reconstruction of an $N \times N$ array involves the inversion of an $N^2 \times N^2$ matrix. For a 256×256 image array, this means inversion and storage of a 65536×65536 matrix! For the vector Wiener filter to be widely applicable, methods must be found to speed the computational process and reduce memory requirements. The next section offers some ideas for future work related to these limitations.

7.3 Recommendations for Future Work

Two primary areas remain to be explored with regard to these linear filter schemes.

7.3.1 Alternate Constraint Functions. In Chapter III, the Fourier domain constraint function $\mathcal{C}(u, v) = 1$ was used in the CLS algorithm such that the manual regularization constant ϵ and the inverse of the Lagrange multiplier γ were equivalent. Alternate constraint functions, such as the two dimensional Laplacian or a support constraint, should be investigated. A different constraint function may provide better algorithm performance for smaller ensemble sizes.

7.3.2 Sparse Matrix Tools. In many cases, the object, OTF, and noise correlation arrays may be well approximated by relatively sparse matrices. A sparse matrix is a special class of matrix that contains a significant number of zero-valued elements [53]. This important property leads to:

1. Reduced memory requirements since only the non-zero entries and their locations in the original matrix need be stored.
2. Reduced computation time by eliminating operations on zero elements.

MATLAB supports sparse matrix computations, including a number of iterative methods for solving simultaneous linear equations [53]. These techniques could be applied to the vector Wiener filter solution to expand the practical computational limits of the filter.

Appendix A. Derivation of Key Conditional Expectations

This appendix presents the derivations of two important conditional expectations in Chapter IV. As an aid to the reader, the following additional information about the semi-classical model of photoelectric light detection is provided.

1. The random variable K , the number of detected photoevents in an image realization, obeys Poisson statistics and is described by the probability density function (PDF) [74]

$$P_K(K; A) = \frac{1}{K!} \left(\int \int_A \lambda(x, y) dx dy \right)^K \exp \left\{ - \int \int_A \lambda(x, y) dx dy \right\}, \quad (\text{A.1})$$

where A represents the area associated with an individual detector element and $\lambda(x, y)$ denotes the rate function. The rate function is proportional to the noiseless image irradiance $i(x, y)$.

2. The random arrival location of the n^{th} photoevent (x_n, y_n) has a PDF related to $\lambda(x, y)$ which can be written as [74]

$$p_{x_n, y_n}(x_n, y_n) = \frac{\lambda(x_n, y_n)}{\int \int_A \lambda(x, y) dx dy}. \quad (\text{A.2})$$

The mean number of photoevents occurring in the differential area $dx dy$ is $\lambda(x, y) dx dy$ [74].

A.1 Equation (4.10)

In this section, Eq. (4.10) is derived and repeated below

$$\mathbb{E}_{x_n, y_n | K, \mathcal{H}, O} \left[O(u, v) \sum_{n=1}^K \exp \{ -j2\pi(u'x_n + v'y_n) \} \right] = K \mathcal{H}^*(u', v') O(u, v) O_n^*(u', v'). \quad (\text{A.3})$$

The reader should recognize that the left side of Eq. (A.3) is the expected value of a function of the random variable (x_n, y_n) . Recall, that the definition of such an expectation is [62]

$$\mathbb{E}_x[g(x)] = \int g(\alpha) p_x(\alpha) d\alpha, \quad (\text{A.4})$$

where α is a dummy variable of integration. Rewriting the left side of Eq. (A.3), using Eq. (A.4), yields

$$\begin{aligned} E_{x_n, y_n | K, \mathcal{H}, O}[\bullet] &= O(u, v) \int \int \left(\sum_{n=1}^K \exp \{-j2\pi(u'x_n + v'y_n)\} \right) \\ &\times p_{x_n, y_n}(x_n, y_n | K, \mathcal{H}, O) dx_n dy_n, \end{aligned} \quad (\text{A.5})$$

where the \bullet notation represents the bracketed expression in Eq. (A.3). Since both integration and summation are linear operations, the order of Eq. (A.5) can be rearranged such that

$$\begin{aligned} E_{x_n, y_n | K, \mathcal{H}, O}[\bullet] &= O(u, v) \sum_{n=1}^K \left(\int \int p_{x_n, y_n}(x_n, y_n | K, \mathcal{H}, O) \right. \\ &\times \exp \{-j2\pi(u'x_n + v'y_n)\} dx_n dy_n \left. \right). \end{aligned} \quad (\text{A.6})$$

The PDF associated with the photoevent arrival location was given in Eq. (A.2) above. Substituting this expression for $p_{x_n, y_n}(x_n, y_n | K, \mathcal{H}, O)$ into Eq. (A.6) yields

$$\begin{aligned} E_{x_n, y_n | K, \mathcal{H}, O}[\bullet] &= \frac{O(u, v)}{\int \int \lambda(x_n, y_n) dx_n dy_n} \sum_{n=1}^K \left(\int \int \lambda(x_n, y_n) \right. \\ &\times \exp \{-j2\pi(u'x_n + v'y_n)\} dx_n dy_n \left. \right). \end{aligned} \quad (\text{A.7})$$

The integral in the parenthesis is the Fourier transform of the rate function. If $\Lambda(u, v)$ denotes the Fourier transform of $\lambda(x, y)$, Eq. (A.7) can be written as

$$E_{x_n, y_n | K, \mathcal{H}, O}[\bullet] = \frac{O(u, v)}{\int \int \lambda(x_n, y_n) dx_n dy_n} \sum_{n=1}^K \Lambda^*(u', v'), \quad (\text{A.8})$$

where the complex conjugate of $\Lambda(u', v')$ is introduced based on the positive sign of the complex exponential kernel. Recall that the mean number of photoevents occurring in the differential area $dx dy$ is $\lambda(x, y) dx dy$. Thus, the integration over this quantity in Eq. (A.8) is equal to the average number of photoevents per image \bar{K} . Also note that $\Lambda(u, v)$ can be normalized such that

$$\Lambda_n(u, v) = \frac{\Lambda(u, v)}{\bar{K}}. \quad (\text{A.9})$$

Thus, Eq. (A.8) now becomes

$$\mathbb{E}_{x_n, y_n | K, \mathcal{H}, O}[\bullet] = \sum_{n=1}^K O(u, v) \Lambda_n^*(u', v'). \quad (\text{A.10})$$

Finally, recall that the rate function $\lambda(x, y)$ is proportional to the noiseless image $i(x, y)$. Thus, in the Fourier domain we can write [74]

$$\Lambda(u, v) = \mathcal{H}(u, v) O(u, v). \quad (\text{A.11})$$

Making this substitution in Eq. (A.10) and noting that $\sum_{n=1}^K$ can be replaced with the variable K yields the final result as stated in Chapter IV

$$\mathbb{E}_{x_n, y_n | K, \mathcal{H}, O}[\bullet] = K \mathcal{H}^*(u', v') O(u, v) O_n^*(u', v'). \quad (\text{A.12})$$

A.2 Equation (4.20)

In this section, Eq. (4.20) is derived and repeated below

$$\begin{aligned} \mathbb{E}_{x_n, y_n, x_m, y_m | K, \mathcal{H}, O}[\bullet] &= (K^2 - K) O_n(u, v) O_n^*(u', v') \mathcal{H}(u, v) \mathcal{H}^*(u', v') \\ &\quad + K \mathcal{H}(u - u', v - v') O_n(u - u', v - v'), \end{aligned} \quad (\text{A.13})$$

where the \bullet notation in Eq. (A.13) represents the bracketed expression on the left side of Eq. (A.14) above. Let us begin with Eq. (4.19) repeated here as

$$\begin{aligned} \mathbb{E}_{x_n, y_n, x_m, y_m | K, \mathcal{H}, O} \left[\sum_{n=1}^K \sum_{m=1}^K \exp \{-j2\pi(u x_n - u' x_m + v y_n - v' y_m)\} \right] &= \\ \mathbb{E}_{x_n, y_n | K, \mathcal{H}, O} \left[\sum_{n=1}^K \exp \{-j2\pi((u - u') x_n + (v - v') y_n)\} \right]_{n=m} & \\ + \mathbb{E}_{x_n, y_n, x_m, y_m | K, \mathcal{H}, O} \left[\sum_{n=1}^K \sum_{m=1}^K \exp \{-j2\pi(u x_n - u' x_m + v y_n - v' y_m)\} \right]_{n \neq m} & \end{aligned} \quad (\text{A.14})$$

The derivation can be divided into two parts: the $n = m$ and $n \neq m$ terms.

A.2.1 n = m Term. First, rewrite the $n = m$ term using the general expectation definition given in Eq. (A.4) such that

$$\begin{aligned} E_{x_n, y_n | K, \mathcal{H}, O_{n=m}} &= \int \int \left(\sum_{n=1}^K \exp \{ -j2\pi((u - u')x_n + (v - v')y_n) \} \right) \\ &\times p_{x_n, y_n}(x_n, y_n | K, \mathcal{H}, O) dx_n dy_n. \end{aligned} \quad (\text{A.15})$$

As in A.1 above, let us rearrange terms, substitute the rate function $\lambda(x, y)$ for the PDF, and recognize that the resultant integral is the Fourier transform of $\lambda(x, y)$, such that Eq. (A.15) becomes

$$E_{x_n, y_n | K, \mathcal{H}, O_{n=m}} = \sum_{n=1}^K \frac{\Lambda(u - u', v - v')}{K}. \quad (\text{A.16})$$

Now writing Eq. (A.16) in terms of the normalized Fourier domain rate function and replacing the summation with the variable K yields

$$E_{x_n, y_n | K, \mathcal{H}, O_{n=m}} = K \Lambda_n(u - u', v - v'). \quad (\text{A.17})$$

Finally, replacing the rate function with the OTF and object spectrum quantities gives the $n = m$ term

$$E_{x_n, y_n | K, \mathcal{H}, O_{n=m}} = K \mathcal{H}(u - u', v - v') O_n(u - u', v - v'). \quad (\text{A.18})$$

A.2.2 n ≠ m Term. As noted in Chapter II, one of the key assumptions associated with the semi-classical model is that the number of photoevents occurring in non-overlapping intervals are statistically independent [74]. Thus, we can split the joint PDF $p_{x_n, y_n, x_m, y_m}(x_n, y_n, x_m, y_m)$ into two marginal PDFs such that

$$p_{x_n, y_n, x_m, y_m}(x_n, y_n, x_m, y_m) = p_{x_n, y_n}(x_n, y_n) p_{x_m, y_m}(x_m, y_m). \quad (\text{A.19})$$

Using this product of marginal PDFs and writing the $n \neq m$ term using Eq. (A.4) gives

$$\begin{aligned}
\mathbb{E}_{x_n, y_n, x_m, y_m | K, \mathcal{H}, O_{n \neq m}} &= (K^2 - K) \\
&\times \left(\int \int \exp \{-j2\pi(ux_n + vy_n)\} \right. \\
&\times p_{x_n, y_n}(x_n, y_n) dx_n dy_n \\
&\times \left(\int \int \exp \{j2\pi(u'x_m + v'y_m)\} \right. \\
&\times p_{x_m, y_m}(x_m, y_m) dx_m dy_m \Big). \tag{A.20}
\end{aligned}$$

As before, substitute the rate function $\lambda(x, y)$ for the PDFs and note that the integrals are Fourier transforms which yields

$$\mathbb{E}_{x_n, y_n, x_m, y_m | K, \mathcal{H}, O_{n \neq m}} = (K^2 - K) \frac{\Lambda(u, v) \Lambda^*(u', v')}{(\overline{K})^2}. \tag{A.21}$$

Finally, normalizing and replacing the rate function with the OTF and object spectrum quantities gives the $n \neq m$ term

$$\mathbb{E}_{x_n, y_n, x_m, y_m | K, \mathcal{H}, O_{n \neq m}} = (K^2 - K) O_n(u, v) O_n^*(u', v') \mathcal{H}(u, v) \mathcal{H}^*(u', v'). \tag{A.22}$$

Now combining Eqs. (A.18) and (A.22) gives the Chapter IV result

$$\begin{aligned}
\mathbb{E}_{x_n, y_n, x_m, y_m | K, \mathcal{H}, O} [\bullet] &= (K^2 - K) O_n(u, v) O_n^*(u', v') \mathcal{H}(u, v) \mathcal{H}^*(u', v') \\
&+ K \mathcal{H}(u - u', v - v') O_n(u - u', v - v'), \tag{A.23}
\end{aligned}$$

Appendix B. Detected Image Probability Density Function

In this appendix, it is shown that $p_{\mathbf{D}|\mathbf{O}}(\mathbf{D}|\mathbf{O})$ approaches a Gaussian distribution for large photoevent counts K based on the Fourier domain image degradation model given in Eq. (2.7) and repeated here as

$$D(u, v) = \sum_{n=1}^K \exp \{-j2\pi(ux_n + vy_n)\} + \sum_{p=1}^P n_p \exp \{-j2\pi(ux_p + vy_p)\}. \quad (\text{B.1})$$

In general, \mathbf{O} is random. Therefore, the conditional probability density function (PDF) given a specific realization of the random object is considered. A given realization for the underlying optical transfer function (OTF) \mathcal{H} is also assumed.

Let us consider the second term first and rewrite using Euler's identity [23] as

$$\begin{aligned} \sum_{p=1}^P n_p \exp \{-j2\pi(ux_p + vy_p)\} &= \left(\sum_{p=1}^P n_p \text{Re} [\exp \{-j2\pi(ux_p + vy_p)\}] \right) \\ &\quad + j \left(\sum_{p=1}^P n_p \text{Im} [\exp \{-j2\pi(ux_p + vy_p)\}] \right) \\ &= \left(\sum_{p=1}^P n_p \text{Cos} [-2\pi(ux_p + vy_p)] \right) \\ &\quad + j \left(\sum_{p=1}^P n_p \text{Sin} [-2\pi(ux_p + vy_p)] \right). \end{aligned} \quad (\text{B.2})$$

Now recall that the detector read noise n_p was assumed to be a zero-mean, uncorrelated, Gaussian random variable with uniform variance σ_r^2 . Thus, the real and imaginary parts of Eq. (B.2) are sums of scaled Gaussian random variables of the form

$$\text{Re} \left[\sum_{p=1}^P n_p \exp \{-j2\pi(ux_p + vy_p)\} \right] = k_1 n_1 + k_2 n_2 + \dots + k_P n_P, \quad (\text{B.3})$$

and

$$\text{Im} \left[\sum_{p=1}^P n_p \exp \{-j2\pi(ux_p + vy_p)\} \right] = l_1 n_1 + l_2 n_2 + \dots + l_P n_P, \quad (\text{B.4})$$

where the real constants k_p and l_p are cosine and sine functions, respectively. Written in this form, both the real and imaginary parts of the second term of Eq. (B.1) are Gaussian distributed, since scaled sums of statistically independent Gaussian random variables are also Gaussian distributed [30]. Thus, the complex random variable

$$\sum_{p=1}^P n_p \exp \{-j2\pi(ux_p + vy_p)\}, \quad (\text{B.5})$$

is Gaussian distributed.

Now consider the first term of Eq. (B.1). Here, the Gaussian nature is less obvious since the underlying PDF associated with the random arrival location (x_n, y_n) is not Gaussian [74]. As before, let us rewrite in terms of real and imaginary parts using Euler's identity such that the first term of Eq. (B.1) becomes

$$\begin{aligned} \sum_{n=1}^K \exp \{-j2\pi(ux_n + vy_n)\} &= \left(\sum_{n=1}^K \text{Re} [\exp \{-j2\pi(ux_n + vy_n)\}] \right) \\ &\quad + j \left(\sum_{n=1}^K \text{Im} [\exp \{-j2\pi(ux_n + vy_n)\}] \right) \\ &= \left(\sum_{n=1}^K \text{Cos} [-2\pi(ux_n + vy_n)] \right) \\ &\quad + j \left(\sum_{n=1}^K \text{Sin} [-2\pi(ux_n + vy_n)] \right). \end{aligned} \quad (\text{B.6})$$

In Chapter II, it was noted that non-overlapping photoevent arrival locations are statistically independent based on the semi-classical model [74]. Thus, the real and imaginary parts in Eq. (B.6) are sums of independent random variables. In fact, these quantities are the result of *large* sums of independent random variables since K is on the order of 1,000 to 1,000,000 for typical astronomical imaging applications. We argue that the first term of Eq. (B.1) has an approximate Gaussian PDF based on the Central Limit theorem which can be stated as follows [23, 62]:

Given N independent random variables x_i with arbitrary PDFs (not necessarily the same), we form their sum

$$x = x_1 + x_2 + \dots + x_n. \quad (\text{B.7})$$

The random variable x has mean $\eta = \eta_1 + \eta_2 + \dots + \eta_n$ and variance $\sigma^2 = \sigma_1^2 + \sigma_2^2 + \dots + \sigma_n^2$. The Central Limit theorem states that under certain general conditions, x approaches a Gaussian distribution, with the same mean and variance, as N increases.

When the x_i are identically distributed, the general sufficient condition for application of the theorem is that the means and variances of the random variables must be finite [23]. The sums of sines and cosines shown in Eq. (B.6) fit the requirements of the Central Limit theorem in that the means and variances are guaranteed to be finite and K is very large for typical applications. Ref. [62] states that $n = 30$ is sufficient in most applications in which the random variables are independent and identically distributed. Thus, the first term of Eq. (B.1)

$$\sum_{n=1}^K \exp \{-j2\pi(ux_n + vy_n)\}, \quad (\text{B.8})$$

approaches a Gaussian distribution for this imaging application.

Now combining the conclusions regarding the first and second terms of Eq. (B.1), it can be seen that $D(u, v)$ is the sum of two independent Gaussian random variables. Thus, $p_{\mathbf{D}|\mathbf{O}}(\mathbf{D}|\mathbf{O})$ is also Gaussian, since the sum of two arbitrary Gaussian random variables is Gaussian distributed [30].

Bibliography

1. Ayers, G. R. and J. C. Dainty. "Iterative blind deconvolution method and its applications," *Opt. Lett.*, *13*:547-549 (1988).
2. Biemond, J., et al. "Iterative Methods for Image Deblurring," *Proc. IEEE*, *78*:856-883 (1990).
3. Bilgen, M. and H. S. Hung. "Constrained least-squares filtering for noisy images blurred by random point spread function," *Opt. Eng.*, *33*:2020-2023 (1994).
4. Combettes, P. L. and H. J. Trussell. "Methods for Digital Restoration of Signals Degraded by a Stochastic Impulse Response," *IEEE Trans. Acoust. Speech, Sig. Proc.*, *37*:393-401 (1989).
5. Davey, B. L. K., et al. "Blind Deconvolution of Noisy Complex-Valued Image," *Opt. Comm.*, *69*:353-356 (1989).
6. Dempster, A. P., et al. "Maximum Likelihood from Incomplete Data via the EM Algorithm," *J. R. Stat. Soc. B*, *39*:1-37 (1977).
7. Fienup, J. R. "Reconstruction of an object from the modulus of its Fourier transform," *Opt. Lett.*, *3*:27-29 (1978).
8. Fienup, J. R. "Space object imaging through the turbulent atmosphere," *Opt. Eng.*, *18*:529-534 (1979).
9. Fienup, J. R. "Phase retrieval algorithms: a comparison," *Appl. Opt.*, *21*:2758-2769 (1982).
10. Fontanella, J. "Analyse de surface d'onde, deconvolution et optique active," *J. Opt. (France)*, *16*:257-268 (1985).
11. Ford, S. D., et al. "Constrained Least Squares Estimation in Deconvolution from Wavefront Sensing." Accepted for publication in *Opt. Comm.* February 1998.
12. Ford, S. D., et al. "Reconstruction of turbulence-degraded images using the vector Wiener filter." Submitted to *Opt. Eng.* November 1997.
13. Ford, S. D., et al. "Reconstruction of turbulence-degraded images using the vector Wiener filter." *SPIE Conference Proceedings on Image Reconstruction and Restoration II, San Diego, California:3170*. 21-32. 1997.
14. Ford, S. D., et al. "Reconstruction of low-light images by use of the vector Wiener filter," *J. Opt. Soc. Am. A*, *14*:2678-2691 (1997).
15. Fried, D. L. "Optical resolution through a randomly inhomogeneous medium for very long and very short exposures," *J Opt. Soc. Am.*, *56*:1372-1379 (1966).
16. Fried, D. L. "Anisoplanatism in adaptive optics," *J Opt. Soc. Am.*, *72*:52-61 (1982).

17. Fried, D. L. "Postdetection wave-front compensation." *SPIE Conference Proceedings on Digital Image Recovery and Synthesis, San Diego, California:828*. 127-133. 1987.
18. Galatsanos, N. P. and R. T. Chin. "Digital Restoration of Multichannel Images," *IEEE Trans. Acoust. Speech, Signal Processing*, *37*:415-421 (1989).
19. Gerchberg, R. W. and W. O. Saxton. "A Practical Algorithm for the Determination of Phase from Image and Diffraction Plane Pictures," *Optik*, *35*:237-246 (1972).
20. Gonglewski, J. D., et al. "First astronomical application of postdetection turbulence compensation," *Appl. Opt.*, *29*:4527-4529 (1990).
21. Gonzales, R. C. and R. E. Woods. *Digital Image Processing*. Addison-Wesley: Reading, Massachusetts, 1992.
22. Goodman, J. W. *Introduction to Fourier Optics*. McGraw-Hill: New York, 1968.
23. Goodman, J. W. *Statistical Optics*. John Wiley and Sons: New York, 1985.
24. Guan, L. and R. K. Ward. "A Maximum A Posteriori Approach to the Restoration of Randomly Distorted Signals." *Proc. IEEE Conference on Acoust. Speech, Signal Processing, New York, New York*. 1770-1773. 1988.
25. Guan, L. and R. K. Ward. "Restoration of Randomly Blurred Images by the Wiener Filter," *IEEE Trans. Acoust. Speech, Signal Processing*, *37*:589-592 (1989).
26. Guan, L. and R. K. Ward. "Restoration of stochastically blurred images by the geometrical mean filter," *Opt. Eng.*, *29*:289-295 (1990).
27. Hardy, J. W. "Active Optics: A New Technology for the Control of Light," *Proc. IEEE*, *66*:651-696 (1978).
28. Hebert, T. J. and K. Lu. "Expectation-Maximization Algorithms, Null Spaces, and MAP Image Restoration," *IEEE Trans. Image Processing*, *4*:1084-1095 (1995).
29. Helstrom, C. W. "Image restoration by the method of least squares," *J. Opt. Soc. Am.*, *57*:297-303 (1967).
30. Hogg, R. V. and A. T. Craig. *Introduction to Mathematical Statistics*. Prentice-Hall: Englewood Cliffs, New Jersey, 5th ed., 1995.
31. Holmes, T. J. "Maximum-likelihood image restoration adapted for noncoherent optical imaging," *J. Opt. Soc. Am. A*, *5*:666-673 (1988).
32. Holmes, T. J. "Expectation-maximization restoration of band-limited, truncated point-process intensities with application in microscopy," *J. Opt. Soc. Am. A*, *6*:1006-1014 (1989).
33. Holmes, T. J. "Blind deconvolution of quantum-limited incoherent imagery: maximum-likelihood approach," *J. Opt. Soc. Am. A*, *9*:1052-1061 (1992).
34. Holmes, T. J. and Y. Liu. "Richardson-Lucy/maximum likelihood image restoration algorithm for fluorescence microscopy: further testing," *Appl. Opt.*, *28*:4930-4938 (1989).

35. Horn, R. A. and C. R. Johnson. *Topics in Matrix Analysis*. Cambridge University Press: New York, 1991.
36. Hung, H. S. "New restoration filter for signals distorted by random impulse response," *Opt. Eng.*, *36*:1452-1457 (1997).
37. Hunt, B. R. "The Application of Constrained Least Squares Estimation to Image Restoration by Digital Computer," *IEEE Trans. Comput.*, *C-22*:805-812 (1973).
38. Hunt, B. R. "Bayesian Methods in Nonlinear Digital Image Restoration," *IEEE Trans. Comput.*, *C-26*:219-229 (1977).
39. Hunt, B. R. and T. M. Cannon. "Nonstationary Assumptions for Gaussian Models of Images," *IEEE Trans. Sys. Man Cyber.*, *SMC-6*:876-882 (1976).
40. Isaacson, E. and H. B. Keller. *Analysis of Numerical Methods*. Dover Publications: New York, 1994.
41. Jain, A. K. *Fundamentals of Digital Image Processing*. Prentice-Hall: Englewood Cliffs, New Jersey, 1989.
42. Kaufman, H. and A. M. Tekalp. "Survey of Estimation Techniques in Image Restoration," *IEEE Contr. Syst. Mag.*, *11*:16-24 (1991).
43. Kay, S. M. *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice-Hall: Englewood Cliffs, New Jersey, 1993.
44. Knox, K. T. and B. J. Thompson. "Recovery of images from atmospherically degraded short exposure images," *Astrophys. J.*, *193*:45-48 (1974).
45. Kolmogorov, A. N. "The local structure of turbulence in incompressible viscous fluids for very large reynolds' numbers." *Turbulence, Classic Papers on Statistical Theory* edited by S. K. Friedlander and L. Topper, 151-155, Wiley-Interscience: New York, 1961.
46. Kundur, D. and D. Hatzinakos. "Blind Image Deconvolution," *IEEE Signal Process. Mag.*, *13*:43-64 (1996).
47. Labeyrie, A. "Attainment of diffraction limited resolution in large telescopes by Fourier analyzing speckle patterns in star images," *Astron. Astrophys.*, *6*:85-87 (1970).
48. Lagendijk, R. L., et al. "Identification and Restoration of Noisy Blurred Images Using the Expectation-Maximization Algorithm," *IEEE Trans. Acoust. Speech Sig. Proc.*, *38*:1180-1191 (1990).
49. Lewis, B. L. and D. J. Sakrison. "Computer Enhancement of Scanning Electron Micrographs," *IEEE Trans. Circuits Syst.*, *CAS-22*:267-278 (1975).
50. Lohman, A., et al. "Speckle masking in astronomy: triple correlation theory and applications," *Appl. Opt.*, *22*:4028-4037 (1983).
51. Lucy, L. B. "An iterative technique for the rectification of observed distributions," *Astron. J.*, *79*:745-754 (1974).

52. Marom, E. and H. Inbar. "New interpretations of Wiener filters for image recognition," *J. Opt. Soc. Am. A*, *13*:1325-1330 (1996).
53. The Math Works, Inc., Natick, Massachusetts. *MATLAB: The Language of Technical Computing* (Version 5 Edition), 1996.
54. Matson, C. L. "Fourier Spectrum Extrapolation and Enhancement using Support Constraints," *IEEE Trans. Sig. Processing*, *42*:156-163 (1994).
55. Matson, C. L. and M. C. Roggemann. "Noise reduction in adaptive-optics imagery with the use of support constraints," *Appl. Opt.*, *34*:767-780 (1995).
56. Melsa, J. L. and D. L. Cohn. *Decision and Estimation Theory*. McGraw-Hill: New York, 1978.
57. Menzel, D. H. *A Field Guide to the Stars and Planets*. Houghton-Mifflin: Boston, 1964.
58. Mesarovic, V., et al. "Regularized Constrained Total Least Squares Image Restoration," *IEEE Trans. Image Processing*, *4*:1096-1107 (1995).
59. Miller, K. "Least Squares methods for ill-posed problems with a prescribed bound," *SIAM J. Math. Anal.*, *1*:52-74 (1970).
60. Miura, N., et al. "Parallel scheme of the iterative blind deconvolution method for stellar object reconstruction," *Appl. Opt.*, *32*:6514-6520 (1993).
61. Oppenheim, A. V. and J. S. Lim. "The importance of phase in signals," *Proc. IEEE*, *69*:529-541 (1981).
62. Papoulis, A. *Probability, Random Variables, and Stochastic Processes* (3rd Edition). McGraw-Hill: New York, 1991.
63. Pratt, W. K. "Generalized Wiener Filter Computation Techniques," *IEEE Trans. Comput.*, *C-21*:636-641 (1972).
64. Pratt, W. K. *Digital Image Processing*. Wiley and Sons: New York, 2nd ed., 1991.
65. Pratt, W. K. and F. Davarian. "Fast Computational Techniques for Pseudoinverse and Wiener Image Restoration," *IEEE Trans. Comput.*, *C-26*:571-580 (1977).
66. Primot, J., et al. "Deconvolution from wave-front sensing: a new technique for compensating turbulence-degraded images," *J. Opt. Soc. Am. A*, *7*:1598-1608 (1990).
67. Qureshi, A. G. and M. M. Fahmy. "2-D Kalman Filtering for the Restoration of Stochastically Blurred Images." *Proc. IEEE Conference on Acoust. Speech, Signal Processing, New York, New York*. 1024-1027. 1988.
68. Richardson, W. H. "Bayesian-Based Iterative Method of Image Restoration," *J. Opt. Soc. Am.*, *62*:55-59 (1972).
69. Roddier, F. "The Effects of Atmospheric Turbulence in Optical Astronomy." *Progress in Optics XIX*, edited by E. Wolf, North-Holland: New York, 1981.
70. Roggemann, M. C., et al. "Compensated speckle imaging: theory and experimental results," *Appl. Opt.*, *33*:7429-7441 (1994).

71. Roggemann, M. C. and D. W. Tyler. "Model-based image reconstruction by means of a constrained least-squares solution," *Appl. Opt.*, *36*:2360–2369 (1997).
72. Roggemann, M. C., et al. "Linear reconstruction of compensated images: theory and experimental results," *Appl. Opt.*, *31*:7429–7441 (1992).
73. Roggemann, M. C. and B. M. Welsh. "Signal-to-noise ratio for astronomical imaging using deconvolution from wave-front sensing," *Appl. Opt.*, *33*:5400–5414 (1994).
74. Roggemann, M. C. and B. M. Welsh. *Imaging Through Turbulence*. CRC Press: Boca Raton, Florida, 1996.
75. Roggemann, M. C., et al. "Biased estimators and object-spectrum estimation in the method of deconvolution from wave-front sensing," *Appl. Opt.*, *33*:5754–5763 (1994).
76. Rosenfeld, A. and A. C. Kak. *Digital Picture Processing, 1*. Academic Press: New York, 2nd ed., 1982.
77. Schulz, T. J. "Multiframe blind deconvolution of astronomical images," *J. Opt. Soc. Am. A*, *10*:1064–1073 (1993).
78. Seldin, J. H. and J. R. Fienup. "Iterative blind deconvolution algorithm applied to phase retrieval," *J. Opt. Soc. Am. A*, *7*:428–433 (1990).
79. Shepp, L. A. and Y. Vardi. "Maximum Likelihood Reconstruction for Emission Tomography," *IEEE Trans. Med. Imaging*, *MI-1*:113–122 (1982).
80. Slepian, D. "Linear least-squares filtering of distorted images," *J. Opt. Soc. Am.*, *57*:918–922 (1967).
81. Snyder, D. L. and T. J. Schulz. "High-resolution imaging at low-light levels through weak turbulence," *J. Opt. Soc. Am. A*, *7*:1251–1265 (1990).
82. Sondhi, M. M. "Image restoration: the removal of spatially invariant degradations," *Proc. IEEE*, *60*:842–853 (1972).
83. Tatarskii, V. I. *Wave Propagation in a Turbulent Medium*. Dover Publications: New York, 1967.
84. Taylor, J. R. *An Introduction to Error Analysis*. University Science Books: Mill Valley, California, 1982.
85. Thiebaut, E. and J.-M. Conan. "Strict a priori constraints for maximum-likelihood blind deconvolution," *J. Opt. Soc. Am. A*, *12*:485–492 (1995).
86. Tikhonov, A. N. and V. Y. Arsenin. *Solutions of Ill-Posed Problems*. Wiley and Sons: New York, 1977.
87. Ward, R. K. and B. E. A. Saleh. "Restoration of images distorted by systems of random impulse response," *J. Opt. Soc. Am. A*, *2*:1254–1259 (1985).
88. Ward, R. K. and B. E. A. Saleh. "Restoration of images distorted by systems of random time-varying impulse response," *J. Opt. Soc. Am. A*, *3*:800–807 (1986).

89. Welsh, B. M. "A Fourier series based atmospheric phase screen generator for simulating nonisoplanatic geometries and temporal evolution." Submitted to *Appl. Opt.* February 1997.
90. Welsh, B. M. "Speckle imaging signal-to-noise ratio performance as a function of frame integration time," *J. Opt. Soc. Am. A*, 12:1364-1374 (1995).
91. Welsh, B. M. and C. S. Gardner. "Performance analysis of adaptive-optics systems using laser guide stars and slope sensors," *J Opt. Soc. Am. A*, 6:1913-1923 (1989).
92. Welsh, B. M. and M. C. Roggemann. "Signal-to-noise comparison of deconvolution from wave-front sensing with traditional linear and speckle image reconstruction," *Appl. Opt.*, 34:2111-2119 (1995).
93. Welsh, B. M. and R. N. VonNiederhausern. "Performance analysis of the self-referenced speckle-holography image-reconstruction technique," *Appl. Opt.*, 32:5071-5077 (1993).
94. Wiener, N. *Extrapolation, Interpolation, and Smoothing of Stationary Time Series*. John Wiley and Sons: New York, 1949.
95. Woods, J. W. and V. K. Ingle. "Kalman Filtering in Two Dimensions: Further Results," *IEEE Trans. Acoust. Speech, Signal Processing*, 29:188-197 (1981).
96. Woods, J. W. and C. H. Radewan. "Kalman Filtering in Two Dimensions," *IEEE Trans. on Info. Theory*, IT-23:473-482 (1977).
97. Wu, Z. "A Multidimensional State-Space Model Kalman Filter with Applications to Image Restoration," *IEEE Trans. Acoust. Speech, Signal Processing*, 33:1576-1592 (1985).

Vita

Captain Stephen D. Ford [REDACTED]

[REDACTED] he graduated from Lowell Senior High School in Lowell, Michigan, and then attended Michigan State University in East Lansing, Michigan from 1985 to 1989. While attending college, he enrolled in the Air Force Reserve Officer Training Corps leading to an Air Force commission in June of 1989. Second Lieutenant Ford then served as an electronics engineer at Wright Laboratory's Armament Directorate, Eglin Air Force Base, Florida. His duties included program management responsibility for the Ballistic Sight Technology Improving Night/Day Gunnery program which sought to improve the first burst accuracy of small crew-served weapons. In December 1994, Captain Ford graduated from the Air Force Institute of Technology (AFIT) with a Master of Science Degree in Electrical Engineering. He entered the Ph.D. program at AFIT in 1995.

[REDACTED]
[REDACTED]

VITA-1