

CORONAL HOLES AND SOLAR *f*-MODE WAVE
SCATTERING OFF LINEAR BOUNDARIES

by

Shea A. Hess Webber
A Dissertation
Submitted to the
Graduate Faculty
of
George Mason University
In Partial Fulfillment of
the Requirements for the Degree
of
Doctor of Philosophy
Computational Sciences and Informatics

Committee:

| | |
|-------------|--|
| _____ | Dr. W. Dean Pesnell, Dissertation Director, Committee Member |
| _____ | Dr. Robert Weigel, Committee Chair |
| _____ | Dr. Arthur Poland, Committee Member |
| _____ | Dr. Jie Zhang, Committee Member |
| _____ | Dr. Joseph Weingartner, Committee Member |
| _____ | Dr. Kevin Curtin, Acting Director, Department of Computational and Data Sciences |
| _____ | Dr. Donna M. Fox, Associate Dean, Office of Student Affairs & Special Programs, College of Science |
| _____ | Dr. Peggy Agouris, Dean, College of Science |
| Date: _____ | Fall Semester 2016 George Mason University Fairfax, VA |

Coronal Holes and Solar f -Mode Wave Scattering Off Linear Boundaries

A dissertation submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy at George Mason University

By

Shea A. Hess Webber
Master of Science
George Mason University, 2012
Bachelor of Science
Gettysburg College, 2009

Director: Dr. W. Dean Pesnell, Affiliate Professor
Department of Physics and Astronomy

Fall Semester 2016
George Mason University
Fairfax, VA

Copyright © 2016 by Shea A. Hess Webber
All Rights Reserved

Dedication

I dedicate this dissertation to my grandfather, Richard C. Webber, from whom I inherited my analytic mind and who has always been proud of my work and my goals.

Grandpa - I've accomplished this for the both of us and I am honored to be sharing the culmination of my academic career with you.

Acknowledgments

First, thank you so much to my family for their unwavering support throughout my life, throughout my education, and throughout my science. At the start of this journey, I don't think any of us expected that I would end up here. Meechai- thank you especially for directing my path into the stars, and for keeping me centered and stable through everything life has thrown my way. You are the most amazing sister anyone could ever have. Thanks also to all my friends for understanding my long silences and absences, and for sticking with me through insanity. Alex- your support, confidence, and comfort has kept me motivated and optimistic through all the stress inherent in this process. I also want to express my appreciation for all my teachers, mentors, and professors who encouraged me and helped me achieve my academic and scientific goals. You have all played a role in my success, no matter how long ago I passed through your classrooms and offices. I must also acknowledge my old friend Sean Conover and his mother Belinda Barker, who unknowingly gave me the gift of solar physics. Thank you!

I would like to thank Dr. Tom Duvall for his advice and support in starting me down my helioseismic research path, as well as his help towards the research presented here. His knowledge of useful introductory material and related publications has been integral in my understanding of this field. He developed the time-distance helioseismology technique used in this research and has done significant research on solar f -mode waves. This dissertation would not have happened without him.

Thank you to Dr. Robert Cameron and Dr. Aaron Birch for their help and guidance on both my simulation code and grasp of time-distance theory. I would also like to thank Dr. Laurent Gizon and the Max Planck Institute for Solar System Research for allowing me to visit and collaborate, as well as hosting my stay.

I would like to thank Dr. Junwei Zhao for his help towards my understanding of the SDO/HMI data and its time-distance products, and for welcoming all my questions. Thanks also to Dr. Phil Scherrer, the Stanford University Solar group, and the JSOC team for helping me navigate the data archive and also hosting me for multiple collaborative visits.

A special thanks goes to Dr. Brad Hindman, who took time out of his own busy schedule to review and edit Section 3.2 for me. His input greatly improved the flow and clarity of the time-distance technique outlined in that highly important section.

I would also like to sincerely thank my dissertation committee for their patience and guidance, particularly Dr. W. Dean Pesnell for over a decade of advising me and molding my research skills.

This work was supported by NASA's Solar Dynamics Observatory. All of the code for this dissertation was written in IDL, with a heavy emphasis on the SolarSoft library. All SDO data was retrieved using the Joint Science Operations Center's Science Data Processing database. The SOHO/EIT images are courtesy of the SOHO/EIT consortium. The SOHO/MDI images are provided by the Solar Oscillations Investigation (SOI) team of the StanfordLockheed Institute for Space Research.

Table of Contents

| | Page |
|---|------|
| List of Tables | vii |
| List of Figures | ix |
| Abstract | xxx |
| 1 Introduction | 1 |
| 1.1 Coronal Hole Overview | 3 |
| 1.2 Wave Modes Overview | 6 |
| 1.3 Helioseismology Overview | 8 |
| 2 Coronal Hole Detection & Results | 10 |
| 2.1 Data Preparation and Detection Method | 10 |
| 2.1.1 Download Process | 11 |
| 2.1.2 Image Selection & Unpacking Routines | 16 |
| 2.1.3 Histogram Filtering & Combining Wavelength Images | 18 |
| 2.1.4 Rotation & Cropping | 22 |
| 2.1.5 Reducing to 1D | 23 |
| 2.1.6 Edge & Width Detection | 25 |
| 2.2 Results | 38 |
| 2.3 Previous Work on Coronal Holes | 46 |
| 3 Wave and Helioseismic Theory | 53 |
| 3.1 Wave and Scattering Theory | 53 |
| 3.1.1 Single Scattering Source | 54 |
| 3.1.2 Multiple Scattering Sources | 57 |
| 3.2 Time-Distance Helioseismology Theory | 69 |
| 3.2.1 Phase-Speed Filtering | 69 |
| 3.2.2 Extracting Travel Times | 74 |
| 3.2.3 Symmetry Arguments | 85 |
| 4 Wave Scattering Simulation & Results | 92 |
| 4.1 Simulation | 92 |
| 4.2 Time-Distance Analysis | 101 |

| | | |
|-------|--|-----|
| 4.2.1 | Testing Assumptions | 101 |
| 4.2.2 | Applied Comparison of Cross-Correlation and Green's Function . . . | 112 |
| 4.2.3 | Statistical Analysis | 116 |
| 4.3 | Results | 132 |
| 5 | Wave Scattering Analysis & Results | 135 |
| 5.1 | Quiet Sun Data | 136 |
| 5.1.1 | HMI Data | 136 |
| 5.1.2 | Time-Distance Results | 138 |
| 5.2 | Coronal Hole Data | 148 |
| 5.2.1 | HMI Data | 148 |
| 5.2.2 | Time-Distance Results | 149 |
| 5.3 | Statistical Comparisons | 157 |
| 6 | Conclusions | 171 |
| 6.1 | Future Work | 173 |
| A | Processing AIA Data - Details | 177 |
| A.1 | Determining the AIA Cutout Bounding Box | 177 |
| A.2 | Updating the Projected AIA Header Information | 178 |
| B | Derivation of Scattering Amplitude | 182 |
| B.1 | Cylindrical Plane Wave Expansion | 182 |
| B.2 | Cylindrical Wave Scattering | 186 |
| C | Fourier Transform Proofs | 189 |
| | Bibliography | 192 |

List of Tables

| Table | Page |
|---|------|
| 2.1 Savitzky-Golay data results to help find “the sweet spot” for smoothing and determining an error estimate in edge location and width. Code uses IDL’s SAVGOL() function where $\sigma = nleft = nright$ is the integer specifying the number of data points on either side of each point to include in the curve-fit, <i>degree</i> specifies the degree of smoothing of the polynomial, and <i>dx</i> is the sampling interval between data points. The variables <i>t1</i> and <i>t2</i> are the threshold values used to determine the inflection point and width locations, respectively. The inflection point location and width boundaries are detected using the specified parameters. The number of width detections is recorded as <i>nw</i> . Note that the widths declared here are those closest to the inflection point on either side. All locations are in AIA pixel units. | 39 |
| 2.2 Savitzky-Golay data results to help find “the sweet spot” for smoothing and determining an error estimate in edge location and width. Code uses IDL’s SAVGOL() function where $\sigma = nleft = nright$ is the integer specifying the number of data points on either side of each point to include in the curve-fit. We choose polynomial <i>degree</i> = 4, and sampling interval <i>dx</i> = 1.0 pixel. The variables <i>t1</i> and <i>t2</i> are the threshold values used to determine the inflection point and width locations, respectively. The inflection point location and width boundaries are detected using the specified parameters. The number of width detections is recorded as <i>nw</i> . Note that the widths declared here are those closest to the inflection point on either side. All locations are in AIA pixel units. | 40 |

2.3 Savitzky-Golay data results to help find “the sweet spot” for smoothing and determining an error estimate in edge location and width. Code uses IDL’s **SAVGOL()** function where $\sigma = nleft = nright$ is the integer specifying the number of data points on either side of each point to include in the curve-fit. We choose polynomial $degree = 4$, and sampling interval $dx = 1.0$ pixel. The threshold variables that determine the inflection point and width locations, respectively, are taken to be $t1 = 1E6$ and $t2 = 2E4$. The inflection point location and width boundaries are detected using the specified parameters. The number of width detections is recorded as nw . Note that the widths declared here are those closest to the inflection point on either side. Bold red values indicate where one of the listed criteria are not met, therefore ruling out that row of parameters. All locations are in AIA pixel units. 41

List of Figures

| Figure | Page |
|--|------|
| 1.1 <i>f</i> -mode waves are similar to deep-water surface waves in the ocean. The wavefronts propagate horizontally across the surface while the particle motion is circular. As mentioned in the attributed text, the circular motion decreases in radius with increasing depth until the motion disappears completely. (Figure 51-9, Feynman, Leighton, and Sands, 1963) | 7 |
| 2.1 A screenshot of the JSOC Data Export Request Generation webpage with parameter specifications for one wavelength of the SDO/AIA data set used in this work. | 12 |
| 2.2 SDO/AIA cutout images on Feb. 5 2014 at 8:00 UT (first data time step) in wavelengths 304 Å (top left), 171 Å (top right), 193 Å (bottom left), 211 Å (bottom right). The images shown are \log_{10} scale of the data. Both axes have units of pixels. Recall from Section 2.1.1, 1 AIA cutout pixel = $0.0359^\circ \simeq 0.436$ Mm. | 18 |
| 2.3 Histogram examples corresponding to the SDO/AIA images in Figure 2.2. Again, the plots are arranged by wavelength: 304 Å (top left), 171 Å (top right), 193 Å (bottom left), 211 Å (bottom right). The vertical dashed lines indicate the intensity exposure limits. The lower limit is the Gaussian center, or average intensity. The upper limit is the Gaussian center plus twice the FWHM. | 20 |
| 2.4 The same images as in Figure 2.2, with the intensity limits from Figure 2.3 applied. Both axes have units of pixels. Recall from Section 2.1.1, 1 AIA cutout pixel = $0.0359^\circ \simeq 0.436$ Mm. | 21 |

| | | |
|------|--|----|
| 2.5 | The first time step of combined AIA data, rotated 20° and cropped to remove rotation effects at the edges. The notation x', y' is used here to denote rotated coordinates. For simplicity in the remainder of this dissertation, we revert to the notation x, y for the rotated coordinates, unless otherwise noted. Both axes have units of pixels. Recall from Section 2.1.1, 1 AIA cutout pixel = $0.0359^\circ \simeq 0.436$ Mm. | 22 |
| 2.6 | The aligned data, summed over the y -dimension, shows the CH boundary through time. The time step along the y -axis has units of seconds ($dt = 45$ s), while the x -axis has units of pixels. Recall from Section 2.1.1, 1 AIA cutout pixel = $0.0359^\circ \simeq 0.436$ Mm. | 24 |
| 2.7 | Additional summing over the time dimension shows a plot through the x -dimension that displays the transition from the CH region to the quiet sun. The x -axis has units of pixels. Recall from Section 2.1.1, 1 AIA cutout pixel = $0.0359^\circ \simeq 0.436$ Mm. The summed intensity has arbitrary units. | 25 |
| 2.8 | An example region with boundary edges of width $dx = 3$, centered at $x = 26$ and $x = 72$. Intensity is indicated by color (grey transitions designate boundary regions). | 27 |
| 2.9 | Summed data from example region in Figure 2.8 with smoothed curves from two different smoothing methods: Gaussian (red dashed) and SG (blue dash-dot). The Gaussian standard deviation is $\sigma = 5$ and the number of points used on either side of the SG fit is also $\sigma = 5$ (11 total points, including center). The polynomial degree of the SG is $degree = 3$ | 28 |
| 2.10 | A comparison of varying SG smoothing parameters. The dashed lines correspond to $degree = 3$ polynomial fits, while dash-dotted lines show $degree = 4$ polynomial fits. Each different set of parameters is also represented by different colors. | 29 |
| 2.11 | Similar to Figure 2.9, but with overplotted boundary markers for each curve. The boundary locations are chosen at the points where each curve meets the limiting data values. The inner boundaries, where the curves reach $y = 100$, are shown with dash-dotted lines. The outer boundaries, where the curves reach $y = 0$, are shown with dotted lines. The summed data curve and boundaries are shown in black. The Gaussian smoothed curve and boundaries are shown in red. The SG smoothed curve and boundaries are shown in blue. | 30 |

| | | |
|------|---|----|
| 2.12 | First-order derivatives of each curve, along with width boundaries from Figure 2.11. The colors and linestyles correspond to those in Figure 2.11. . . . | 31 |
| 2.13 | Second-order derivatives of each curve, along with width boundaries from Figure 2.11. The colors and linestyles correspond to those in Figures 2.11-2.12. | 32 |
| 2.14 | Raw and smoothed example curves with detected inflection points from both smoothed cases. Both curves produce inflection points at the same locations, which are also the centers of the width regions and where both smoothed curves cross the raw data curve. The red dotted lines represent the inflection point locations determined from the Gaussian smoothed curve. The blue dashed lines represent the inflection point locations determined from the Savitky-Golay smoothed curve. | 33 |
| 2.15 | Third-order derivatives of each curve, along with width boundaries from Figure 2.11. The colors and linestyles correspond to those in Figures 2.11-2.13. The detected inflection points are represented as vertical black dashed lines. | 34 |
| 2.16 | Raw and smoothed example curves with detected inflection points and width boundaries from both smoothed cases. Both curves produce width boundaries at the same locations, which also correspond to the boundary edges from the summed data curve. The black dashed lines represent the detected inflection points. The red dashed lines represent the width boundary locations determined from the Gaussian smoothed curve. The blue dotted lines represent the width boundary locations determined from the Savitky-Golay smoothed curve. | 36 |
| 2.17 | The 2D example boundary region with the detected inflection points (solid red lines) and detected boundary widths (dashed red lines). Intensity is indicated by the background color. | 37 |
| 2.18 | 0 th -order summed AIA data (black) with the overplotted SG smoothed curve (cyan). The detected inflection point (solid red) and detected boundary widths (red dashed) are also shown. The smoothing parameters are $\sigma = 25$ and $degree = 4$. Recall from Section 2.1.1, 1 AIA cutout pixel = $0.0359^\circ \simeq 0.436$ Mm. The summed intensity has arbitrary units. | 42 |

| | | |
|------|--|----|
| 2.19 | 1 st -order summed AIA data (black) with the overplotted SG smoothed curve (cyan). The detected inflection point (solid red) and detected boundary widths (red dashed) are also shown. The smoothing parameters are $\sigma = 25$ and $degree = 4$. Recall from Section 2.1.1, 1 AIA cutout pixel = $0.0359^\circ \simeq 0.436$ Mm. The summed intensity has arbitrary units. | 43 |
| 2.20 | 2 nd -order summed AIA data (black) with the overplotted SG smoothed curve (cyan). The detected inflection point (solid red) and detected boundary widths (red dashed) are also shown. The smoothing parameters are $\sigma = 25$ and $degree = 4$. Recall from Section 2.1.1, 1 AIA cutout pixel = $0.0359^\circ \simeq 0.436$ Mm. The summed intensity has arbitrary units. | 44 |
| 2.21 | 3 rd -order summed AIA data (black) with the overplotted SG smoothed curve (cyan). The detected inflection point (solid red) and detected boundary widths (red dashed) are also shown. The smoothing parameters are $\sigma = 25$ and $degree = 4$. Recall from Section 2.1.1, 1 AIA cutout pixel = $0.0359^\circ \simeq 0.436$ Mm. The summed intensity has arbitrary units. | 45 |
| 3.1 | Scattering of an incident plane wave off of a point source scatterer; displaying only the scattered wave (left), and the total (incident + scattered) wave (right). Only the $m = 0$ mode is shown for $l = 600$. The dashed line indicates a single wavefront of the incident plane wave ψ_i . Note that the incident wave vector \mathbf{k}_i is normal to this wavefront. The resultant scattered wave ψ_s is radially symmetric about the point source in the far-field, as shown by the dotted line wavefront. The amplitudes shown here are normalized to the largest absolute value result from the total wave equation. This image was created using the <i>cyl.wav.pro</i> code, discussed in Section 4.1. | 55 |
| 3.2 | A single scatterer S_1 is shifted off the origin to a position $\mathbf{b}_1 = [b_{1x}, b_{1y}] = b_1[\cos \phi_{b_1}, \sin \phi_{b_1}]$ (the red vector). The observation point P is located at $\mathbf{r}_p = [r_{px}, r_{py}] = r_p[\cos \phi_p, \sin \phi_p]$ with respect to the origin (the blue vector). The position of P with respect to S_1 is $\mathbf{r}_1 = \mathbf{r}_p - \mathbf{b}_1 = [r_{px} - b_{1x}, r_{py} - b_{1y}]$ (the purple vector). The distance between P and S_1 is therefore $ \mathbf{r}_1 = r_1 = \sqrt{(r_{px} - b_{1x})^2 + (r_{py} - b_{1y})^2}$ | 57 |

- 3.3 A simple plane wave ψ_i scattering off of two points located at \mathbf{b}_1 and \mathbf{b}_2 . The total scattered wave is observed at point P some distance $\mathbf{r}_p = [r, \phi_p]$ from the origin O . With respect to the two scatterers, P is located at $\mathbf{r}_1 = [r_1, \phi_1]$ and $\mathbf{r}_2 = [r_2, \phi_2]$, respectively. The incident wave vector $\mathbf{k}_i = [k, \phi_i]$ and the scattered wave vectors \mathbf{k}_n have the same magnitude, k , as does the wave vector from the origin in the the direction of P , \mathbf{k}_p . The scattered waves are circular, but only the waves in the direction of P are of interest. The angles between the incident and the scattered wave vectors are $\varphi_n = \phi_i - \phi_n$ (not shown for clarity). γ_n denote the angles between \mathbf{r}_p and \mathbf{b}_n ($\mathbf{r}_p \cdot \mathbf{b}_n = r b_n \cos \gamma_n$; in this depiction, $\gamma_2 \equiv \phi_p$). Note, P is assumed to be in the same z -plane as ψ_i and the two scatterers. 60
- 3.4 Compare to Fig. 8.1, page 311 in Martin (2006). 67
- 3.5 An example of various frequency-filtering functions is shown in (a). The blue dot-dashed ($\cdot - \cdot - \cdot - \cdot -$) line is a Gaussian function, the black and green solid lines (--- , ---) are two boxcar functions of varying widths, and the red dashed curve ($- - -$) is an FTC filter. The Fourier transforms of these functions are shown in (b), with corresponding colors and linestyles. Both boxcar functions have significant side-lobes in their transforms, which disrupt the clarity of the transformed signal. The transform of the Gaussian is, as expected, another Gaussian so the transformed signal is “perfect”. However, the downside of the Gaussian filter is the abrupt loss of power about a peak frequency, which is limiting if you have a strong signal spread over a broad range of frequencies. The FTC filter is useful as a full bandpass filter of all relevant frequencies while minimizing the effect of disadvantageous side-lobes, as can be seen when comparing the transforms in (b). Note, this frequency-filtering example is a simplified illustration but the concept holds true independent of the filtering variable. 72

| | | |
|------|---|----|
| 3.6 | These phase-speed diagrams (also known as ω - k diagrams; ω is angular frequency and k is wavenumber) display the spectral results of the sample data cube shown in Figure 3.7a. The full power spectrum in (a) shows multiple waves modes, as well as a low-frequency signal (where researchers expect to find the elusive g -mode). The low-frequency signal is filtered out in both (b) and (c). Any signal below the median phase-speed distance between the f - and p_1 -modes is filtered out to show only the various p_n -modes in (b). Similarly, all phase-speeds above the f - and p_1 -mode median are filtered out of (c), leaving only the f -mode signal. | 73 |
| 3.7 | An example time step of an HMI tracked mapped doppler data cube. The tracked data is initially unfiltered (a). The various modes are filtered from the full power spectrum in the Fourier domain (as shown in Figure 3.6), and the filtered spectra are then transformed back to the spatial domain, giving just the p -modes (b) and the isolated f -mode (c). | 73 |
| 3.8 | Example cross-correlations using the f -mode filtered data from Figure 3.7c. | 78 |
| 3.9 | TD plots of three sample source-point locations, showing unshifted, unaveraged data. The time lag has units of seconds ($dt = 45$ s), while \mathbf{r} and Δ are in pixel units. As expected, the cross-correlations are strongest at $\mathbf{t} = 0$ for $\mathbf{r}_s = \mathbf{r}_r$ ($\Delta = 0$), where the source is correlating perfectly with itself. These locations are represented by the vertical dashed lines. The plots are “shifted” such that the dashed lines coincide with $\Delta = 0$ on the axis above the plots. This is representative of the TD shifting that is done before the cross-correlation results can be averaged to improve the signal-to-noise ratio. Note that some data manipulation was done to produce these images, in order to clarify the cross-correlation signals within the background noise, but this enhancement is not performed in the actual data analysis. | 79 |
| 3.10 | The same three TD plots from Figure 3.9, now shifted in terms of $\Delta = 0$. Again, time lag has units of seconds ($dt = 45$ s), while Δ is in pixel units. The cross-correlation signals are visibly propagating from $[\Delta, \mathbf{t}] = [0, 0]$. These data can now be averaged (along with the shifted TD plots from all other instances of \mathbf{r}_r), the result of which is shown in Figure 3.11. We note that other than our indication of the corresponding \mathbf{r}_s and \mathbf{r}_r , the shift loses all relevant location information about the associated source points. | 80 |

| | | |
|------|---|----|
| 3.11 | An example TD plot, averaged over the shifted data from all source points. Again, time lag has units of seconds ($dt = 45$ s), while Δ is in pixel units. The signal-to-noise ratio is much larger in this plot than in the unshifted, unaveraged data in the Figures 3.9 and 3.10 plots, even despite the signal enhancement performed in producing those images (which was not used here). | 81 |
| 3.12 | An example TD plot showing a line of constant phase (left), shifted with respect to the constant phase (middle), and then averaged over Δ (right). (Hindman <i>et al.</i> , 2004, Figure 7) | 81 |
| 3.13 | The f -mode power and cross-covariance functions. (a) A cut through of the filtered power spectrum at $k_y = 0$. (b) The spatially averaged cross-covariance function (thin line) for some point-to-point distance Δ . The thick dot-dashed line shows the window function $f(t)$ used in the travel-time measurement procedure. The thick line shows the reference cross-covariance function, C^0 . (c) A sample cross-covariance between a single pair of points that are separated by Δ (thin line). Overplotted is C^0 (thick line). Note the scale difference between (b) and (c) (Gizon and Birch, 2004, Figure 1) | 84 |
| 3.14 | We consider two source signals $\phi(\mathbf{r}_{s1}, t)$ and $\phi(\mathbf{r}_{s2}, t)$ located at $\mathbf{r}_{s1} = [x_{s1}, y_{s1}]$ and $\mathbf{r}_{s2} = [x_{s2}, y_{s2}]$. The observed signal $\phi(\mathbf{r}_r, t)$ is made at a reference point a distance $\mathbf{r}_r = [x_r, y_r]$ from the origin O . With respect to the source points, P is located at $\mathbf{\Delta}_1 = \mathbf{r}_r - \mathbf{r}_{s1}$ and $\mathbf{\Delta}_2 = \mathbf{r}_r - \mathbf{r}_{s2}$, respectively. Note that the source and receiver points are assumed to be in the same z -plane. Also note that a wavefield is not shown, for clarity, but that perturbations could be located anywhere in the domain. The scatterers do not have to be located at the considered source points. | 86 |
| 3.15 | An example of the quadrant geometry used in TD helioseismology to measure east-west flows. The black regions show pixels in the east-west quadrant arcs that are a distance Δ from the central pixel x . The gray regions cover the pixels that would be used for three separate smaller Δ . (Hindman <i>et al.</i> , 2004, Figure 6) | 90 |

- 4.1 Scattering amplitudes $f(\varphi_n)$ for scattering angle $\varphi_n = 0^\circ$ computed using the simulation code for each $0 \leq m \leq 7$. The various wavenumbers k are shown ranging in color from lowest (red) to highest (black). The color bar is plotted to the right of the image. It is apparent that for all values of k , the associated $f(\varphi_n = 0^\circ)$ converge to 0 before $m = 3$ 94
- 4.2 Scattering amplitudes $f(\varphi_n)$ for scattering angles $\varphi_n = 0^\circ, 45^\circ, 90^\circ$ computed using the simulation code for each wavenumber k . The solutions for azimuthal modes $m = 0$ (black), $m = 1$ (blue), $m = 2$ (green), and $m = 3$ (red, dashed) are labeled in each plot. The $m = 3$ case is also representative of the $f(\varphi_n)$ solutions for all higher order m since they similarly converge to $f(\varphi_n) = 0$. Note that the even-order azimuthal modes do not vary with φ_n , as expected from Equation 4.2. These plots, along with Figure 4.1, show that all orders of $m > 3$ can be neglected in the simulation scattering calculations, independent of k or φ_n 96
- 4.3 Example data produced by the scattering simulation code. Three time steps are shown at $t = 0$ (top row), $t = 100$ (middle row), and $t = 200$ (bottom row). Axes are in units of pixels. Greyscale colors indicate wave amplitude (black/white is minimum/maximum, grey is average), although the incident and scattered waves are not scaled with respect to each other since the incident wave signal is actually 2-3 orders of magnitude larger than the scattered wave signals. The incident wave (left column) is in-phase at $x = 0$ when $t = 0$, with an incident angle $\phi_i = 0^\circ$. The scattering results from a single scatterer (middle column) and 10 evenly-spaced scatterers (right column) are shown for each time step. The scatterer locations are indicated by the red dots (\bullet). For both scattering results, the scatterers are located at $x = n_x/3$. The single scatterer is placed at $y = 0$, and the multiple scatterers are spaced symmetrically around $y = 0$ with $d = \Delta y = 5$ pixels. Note that each image is scaled to itself, and that the scattered wave signals are actually much smaller than the incident signal. The total wavefields (incident + scatterer(s)) are visually indistinguishable from the incident waves shown here because the scattering signal is much weaker than the incident wave signal. 97

- 4.5 Example data produced by the scattering simulation code, displaying a change in incident angle from Figure 4.3. Three time steps are shown at $t = 0$ (top row), $t = 100$ (middle row), and $t = 200$ (bottom row). Axes are in units of pixels. Greyscale colors indicate wave amplitude (black/white is minimum/-maximum, grey is average), although the incident and scattered waves are not scaled with respect to each other since the incident wave signal is actually 2-3 orders of magnitude larger than the scattered wave signals. The incident wave (left column) is in-phase at $t = 0$, with an incident angle $\phi_i = 45^\circ$. The scattering results from a single scatterer (middle column) and 10 evenly-spaced scatterers (right column) are shown for each time step. The scatterer locations are indicated by the red dots (\bullet). For both scattering results, the scatterers are located at $x = n_x/3$. The single scatterer is placed at $y = 0$, and the multiple scatterers are spaced symmetrically around $y = 0$ with $d = \Delta y = 5$ pixels. Note that each image is scaled to itself, and that the scattered wave signals are actually much smaller than the incident signal. The total wavefields (incident + scatterer(s)) are visually indistinguishable from the incident waves shown here because the scattering signal is much weaker than the incident wave signal. 99
- 4.6 Example scattered data produced by the scattering simulation code, displaying changes in scatterer spacing, using the $t = 100$ incident conditions from Figure 4.3 (in-phase, $\phi_i = 0^\circ$). Axes are in units of pixels. Greyscale colors indicate wave amplitude (black/white is minimum/maximum, grey is average), although the three scattering cases are not scaled with respect to each other. The scatterer locations are indicated by the red dots (\bullet); however, dot size is not indicative of scatterer size. The middle plot is the same as the middle-righthand plot in Figure 4.3, for reference. The left plot shows the scattering response from 10 scatterers with $d = \Delta y = 2$ pixels; the right plot shows the 10-scatterer response for $d = \Delta y = 10$ pixels. We note that the locations in x are likewise adjustable, but not shown. 100

| | | |
|-----|--|-----|
| 4.7 | Examples of the real-part of the simulated scattered wave functions for a single scatterer (left) and a column of ten scatterers (right) at $t = 0$. Axes are in units of pixels. Greyscale colors indicate wave amplitude (black/white is minimum/maximum, grey is average), although the incident and scattered waves are not scaled with respect to each other since the incident wave signal is actually 2-3 orders of magnitude larger than the scattered wave signals. The incident wave packet (top) for these examples is initially in-phase and aligned at the $x = 0$ line. Note that these images are the same simulated data as those in the top row of Figure 4.3. | 102 |
| 4.8 | The simulated scattered data in Figure 4.7 summed over the y -dimension. The x -axis has units of pixels, while the summed amplitude has arbitrary units. Note that the summed amplitude for the 10-scatterer case is about $10\times$ the magnitude of the 1-scatterer case results. This data only shows the $t = 0$ time step. Other time steps have a similar magnitude difference in amplitude between the 1-scatterer and 10-scatterer cases. Stacking these summed amplitude data through time produces plots of the time-dependent Green's functions of the scattering. | 103 |
| 4.9 | The Green's function results for the 1- and 10-scatterer <i>scattered</i> wave simulation data are shown in the left and left-center images, respectively. The x -axis has units of pixels, while the time steps on the y -axis are in seconds ($dt = 45$ s). The vertical dashed line indicates the x -location of the scatterers. Note that if the $t = 0$ data from these two figures were plotted, the results would give Figure 4.8. These images are shown to scale with themselves, so for comparison purposes the scaled-difference (right-center) and scaled-difference overlaid with the scaled-10-scatterer case (right) are also shown. | 105 |

- 4.10 Similar to Figure 4.9, the Green's function results for the 1- and 10-scatterer *incident* wave simulation data are shown in the left and left-center images, respectively. The x -axis has units of pixels, while the time steps on the y -axis are in seconds ($dt = 45$ s). The vertical dashed line indicates the x -location of the scatterers. These images are shown to scale with themselves, so for comparison purposes the difference (right-center) and difference overlaid with the 10-scatterer case (right) are also shown. As expected, there is no difference between the incident wave linear TD results, since the incident wave packet designated in both cases is the same. Thus, the right-hand image is a mirror of the left and left-center images. 106
- 4.11 Again similar to Figure 4.9, the Green's function results for the 1- and 10-scatterer *total* wave simulation data are shown in the left and left-center images, respectively. The x -axis has units of pixels, while the time steps on the y -axis are in seconds ($dt = 45$ s). The vertical dashed line indicates the x -location of the scatterers. These images are shown to scale with themselves, so for comparison purposes the difference (right-center) and difference overlaid with the scaled-10-scatterer case (right) are also shown. Although the backscattered wave signal is too weak to be seen in both the 1-scatterer and 10-scatterer cases by themselves, the difference clearly shows both the forward and backscattered wave signals. The apparent scattering mainly comes from the the 10-scatterer case because the scattered wave magnitude is $10\times$ larger than that of the 1-scatterer case. The overlaid image on the right shows both the incident and scattered wave parts together. 107
- 4.12 Examples of the simulated real scattered wave functions for a single scatterer (left) and a column of ten scatterers (right) at $t = 0$. Axes are in units of pixels. Greyscale colors indicate wave amplitude (black/white is minimum/-maximum, grey is average), although the incident and scattered waves are not scaled with respect to each other since the incident wave signal is actually 2-3 orders of magnitude larger than the scattered wave signals. The incident wave packet (top) for these examples is initially in-phase at an angle of 45° and a radial distance of $nx/2 = ny/2$ from the domain center. Note that these images are the same simulated data as those in the top row of Figure 4.5.109

| | | |
|------|---|-----|
| 4.13 | The simulated scattered data in the 10-scatterer cases from Figures 4.7 and 4.12, summed over the y -dimension. The x -axis has units of pixels, while the summed amplitude has arbitrary units. Note that the summed amplitudes for both incident angle cases have approximately the same magnitude. This data only shows the $t = 0$ time step. Other time steps have similar amplitude magnitudes between the $\phi_i = 0^\circ$ and $\phi_i = 45^\circ$ cases. Stacking these summed amplitude data through time produces plots of the time-dependent Green's functions of the scattering. | 110 |
| 4.14 | The Green's function results for the 1- and 10-scatterer <i>scattered</i> wave simulation data, with an incident angle of 45° , are shown in the left and center images, respectively. The x -axis has units of pixels, while the time steps on the y -axis are in seconds ($dt = 45$ s). The vertical dashed line indicates the x -location of the scatterers. Note that if the $t = 0$ data from the 10-scatterer case was plotted, the results would give the left plot in Figure 4.13. For comparison purposes, the difference (right) is also shown, with no scaling applied. | 111 |
| 4.15 | The Green's function results from the 10-scatterer, in-phase, $\phi_i = 0^\circ$ simulated data, corresponding to the simulated data represented in Figure 4.3. The x -axis has units of pixels, while the time steps on the y -axis are in seconds ($dt = 45$ s). The total wave (left) and incident wave (middle) results are indistinguishable since the scattered wave is weaker than the incident wave. The scattered wave result (right) shows both a forward and backward scattered wave to either side of the line of scatterers, which is designated by the vertical dashed line. | 114 |
| 4.16 | The Green's function results from the 10-scatterer, random phased, $\phi_i = 0^\circ$ simulated data, corresponding to the simulated data represented in Figure 4.4. The x -axis has units of pixels, while the time steps on the y -axis are in seconds ($dt = 45$ s). The total wave (left) and incident wave (middle) results are indistinguishable since the scattered wave is weaker than the incident wave. The scattered wave result (right) shows both a forward and backward scattered wave to either side of the line of scatterers, designated by the vertical dashed line. The random phase results have strong signals throughout the duration of the simulation time, corresponding to small, randomly coherent incident wave packets. | 115 |

| | | |
|------|---|-----|
| 4.17 | The cross-correlated results of the multiphased Green's functions in Figure 4.16. The x -axis has units of pixels, while the time lag on the y -axis is in seconds ($dt = 45$ s). The cross-correlation between the scattered wave and the incident wave at $x = 0$ (left) produces the same result as cross-correlating the total wave with itself at $x = 0$. The backscattered wave diverges from the line of scatterers, designated by the vertical dashed line. The cross-correlation of the incident wave with itself at $x = 0$ (right) crosses the line of scatterers at the same time. | 116 |
| 4.18 | Student's t -test results between the Green's functions of in-phase, $\phi_i = 0^\circ$ scattering off of 1 and 10 scatterers. The x -axis has units of pixels. Both the 3×3 (left) and 5×5 (right) superpixel calculations are shown. | 121 |
| 4.19 | Student's t -test significance of differences results reflecting the statistical significance of the t -statistic results shown in Figure 4.18. The x -axis has units of pixels. Again, both the 3×3 (left) and 5×5 (right) superpixel calculations are shown. Note that darker shades indicate higher significance. | 122 |
| 4.20 | Histograms of the significance distributions in Figure 4.19. Both the 3×3 (top) and 5×5 (bottom) superpixel cases are shown. All significance left of the vertical dash-dotted line indicates significant differences (significance level ≤ 0.05 ; probability of significance = 95%) between the 1 and 10 scatterer cases, while the dashed line indicates <i>highly</i> significant differences (significance level ≤ 0.01 ; probability of significance = 99%). | 123 |
| 4.21 | Binary Student's t -test significance results from Figure 4.19, masked to above 95% probability of significance. The x -axis has units of pixels. Again, both the 3×3 (left) and 5×5 (right) superpixel calculations are shown. Note that black regions indicates statistically significant differences; white, statistically insignificant differences. | 124 |
| 4.22 | Student's t -test results from Figure 4.19, masked to above 95% probability of significance. The x -axis has units of pixels. Again, both the 3×3 (left) and 5×5 (right) superpixel calculations are shown. In either case, the main region of difference between the TD signals is passed through as statistically significant. | 125 |

| | | |
|------|---|-----|
| 4.23 | Student's t -test results between the Green's functions of in-phase, $\phi_i = 45^\circ$ scattering off of 1 and 10 scatterers. The x -axis has units of pixels. Both the 3×3 (left) and 5×5 (right) superpixel calculations are shown. | 127 |
| 4.24 | Student's t -test significance results reflecting the statistical significance of the t -statistic results shown in Figure 4.23. The x -axis has units of pixels. Again, Both the 3×3 (left) and 5×5 (right) superpixel calculations are shown. Note that darker shades indicate higher significance. | 128 |
| 4.25 | Histograms of the significance distributions in Figure 4.24. Both the 3×3 (top) and 5×5 (bottom) superpixel calculations are shown. All significance left of the vertical dash-dotted line indicates significant differences (significance level ≤ 0.05 ; probability of significance = 95%) between the 1 and 10 scatterer cases, while the dashed line indicates <i>highly</i> significant differences (significance level ≤ 0.01 ; probability of significance = 99%). | 129 |
| 4.26 | Binary Student's t -test significance results from Figure 4.24, masked to above 95% probability of significance. The x -axis has units of pixels. Again, both the 3×3 (left) and 5×5 (right) superpixel calculations are shown. Note that black regions indicates statistically significant differences; white, statistically insignificant differences. | 130 |
| 4.27 | Student's t -test results from Figure 4.24, masked to above 95% probability of significance. The x -axis has units of pixels. Again, both the 3×3 (left) and 5×5 (right) superpixel calculations are shown. In either case, the main region of difference between the TD signals is passed through as statistically significant with the exception of the backscattered wave signal at earlier time steps in the 5×5 superpixel case (as seen in Figure 4.26). | 131 |
| 5.1 | SDO/AIA cutout images on Nov. 13 2013 at 4:00AM (middle of observation) in wavelengths 304 Å (top left), 171 Å (top right), 193 Å (bottom left), 211 Å (bottom right). The images shown are \log_{10} scale of the data. Both axes have units of pixels. Recall from Section 2.1.1, 1 AIA cutout pixel = $0.0359^\circ \simeq 0.436$ Mm. | 137 |
| 5.2 | SDO/HMI tracked Dopplergram image on Nov. 13 2013 at 4:00AM (middle of observation). The data has already been filtered for the f -mode. Both axes have units of pixels. Recall, 1 HMI pixel = $0.06^\circ \simeq 0.729$ Mm. | 138 |

| | | |
|-----|---|-----|
| 5.3 | The f -mode filtered SDO/HMI quiet sun data, after summing over the y -dimensions. The time step along the y -axis has units of seconds ($dt = 45$ s). The x -axis has units of pixels. Recall, 1 HMI pixel = $0.06^\circ \simeq 0.729$ Mm. . . | 139 |
| 5.4 | The TD plot of the f -mode filtered SDO/HMI quiet sun data, after summing over the y -dimension. Time lag \mathfrak{t} has units of seconds ($dt = 45$ s). Δ has units of pixels. Recall, 1 HMI pixel = $0.06^\circ \simeq 0.729$ Mm. The TD results were calculated for all possible reference lines, and were subsequently shifted with respect to each reference line. The shifted results were then summed to create this result, with an improved signal-to-noise ratio. | 140 |
| 5.5 | Example reference cross-correlations, derived from the quiet sun data, used in the travel time measurements. Time lag \mathfrak{t} has units of seconds ($dt = 45$ s). Δ has units of pixels. Recall, 1 HMI pixel = $0.06^\circ \simeq 0.729$ Mm. | 143 |
| 5.6 | The $\tau_+(x_s, x_r)$ (top left), $\tau_-(x_s, x_r)$ (top right), $\tau_{\text{mean}}(x_s, x_r)$ (bottom left), and $\tau_{\text{diff}}(x_s, x_r)$ (bottom right) travel time measurements from the quiet sun data. Both axes have units of pixels. Recall, 1 HMI pixel = $0.06^\circ \simeq 0.729$ Mm. Travel times are indicated in greyscale, with black and white corresponding to shortest and longest times, respectively. We note that there is a visible pattern of horizontal and vertical lines; as of yet, we cannot conclude about the source of the pattern, but it is under investigation. | 144 |
| 5.7 | The $\tau_+^{\text{all}}(\Delta)$ (top left), $\tau_-^{\text{all}}(\Delta)$ (top right), $\tau_{\text{mean}}^{\text{all}}(\Delta)$ (bottom left), and $\tau_{\text{diff}}^{\text{all}}(\Delta)$ (bottom right) travel time measurements from the shifted TD results of the full quiet sun data. Travel times have units of seconds ($dt = 45$ s). Δ has units of pixels. Recall, 1 HMI pixel = $0.06^\circ \simeq 0.729$ Mm. These plots may seem uninteresting because there is no variation in the travel times. However, this is because we are essentially comparing our $C^0(\Delta, \mathfrak{t})$ to itself. We expect 0-variation when making this comparison, so these results act as a sanity check of our travel time calculations. Also, this leads us to understand that the variations seen in Figures 5.8 and 5.9 are indicative of the noise that is introduced when we compute travel times in sub-regions when $C(\Delta, \mathfrak{t})$ and $C^0(\Delta, \mathfrak{t})$ have different signal-to-noise ratios. | 145 |

- 5.8 The $\tau_+^{\text{in}}(\Delta)$ (top left), $\tau_-^{\text{in}}(\Delta)$ (top right), $\tau_{\text{mean}}^{\text{in}}(\Delta)$ (bottom left), and $\tau_{\text{diff}}^{\text{in}}(\Delta)$ (bottom right) travel time measurements from the shifted TD results of the quiet sun region, for a sub-region of the size of the CH region in Section 5.2. Travel times have units of seconds ($dt = 45$ s). Δ has units of pixels. Recall, 1 HMI pixel = $0.06^\circ \simeq 0.729$ Mm. These travel times are for reference and later comparison with the shifted travel times in Figure 5.13. They also give some idea of the noise difference in travel times measured using the shifted-average of the full dimensional data set, and those obtained from the sub-region. . . 146
- 5.9 The τ_+^{out} (top left), τ_-^{out} (top right), $\tau_{\text{mean}}^{\text{out}}$ (bottom left), and $\tau_{\text{diff}}^{\text{out}}$ (bottom right) travel time measurements from the shifted TD results of the quiet sun region, for a sub-region of the size of the region outside the CH in Section 5.2. Travel times have units of seconds ($dt = 45$ s). Δ has units of pixels. Recall, 1 HMI pixel = $0.06^\circ \simeq 0.729$ Mm. These travel times are for reference and later comparison with the shifted travel times in Figure 5.14. They also give some idea of the noise difference in travel times measured using the shifted-average of the full dimensional data set, and those obtained from the sub-region. . . 147
- 5.10 Averaged, shifted TD plots of the f -mode filtered SDO/HMI CH data, after summing over the y -dimension. Time lag t has units of seconds ($dt = 45$ s). Δ has units of pixels. Recall, 1 HMI pixel = $0.06^\circ \simeq 0.729$ Mm. The left-most TD result are calculated from all possible reference lines, ignoring the location of the CH boundary. The center plot shows the average shifted result from only those reference lines inside the CH region. The right-hand result is similarly that of the region outside the CH boundary. The three cases are highly similar, although it is possible that the TD signal in the outer region is attenuated faster than inside the CH. This difference will be investigate further in Section 5.3. 150

| | | |
|------|---|-----|
| 5.11 | The $\tau_+(x_s, x_r)$ (top left), $\tau_-(x_s, x_r)$ (top right), $\tau_{\text{mean}}(x_s, x_r)$ (bottom left), and $\tau_{\text{diff}}(x_s, x_r)$ (bottom right) travel time measurements from the full CH data. Both axes have units of pixels. Recall, 1 HMI pixel = $0.06^\circ \simeq 0.729$ Mm. Travel times are indicated in greyscale, with black and white corresponding to shortest and longest times, respectively. The solid lines represent the location of the center of the CH boundary, while the dashed lines indicate the width of the boundary. Where (x_s, x_r) are below/left of the CH boundary, the travel times are measured within the CH region with respect to reference lines inside the region. Likewise for (x_s, x_r) are above/right of the CH boundary, the travel times are calculated outside of the CH region with respect to reference lines outside the region. Similarly, for x_s inside (outside) the boundary and x_r outside (inside) the boundary, the travel times are measured outside (inside) the CH region with respect to signals inside (outside) the CH region. We again note that there is a visible pattern of horizontal and vertical lines; as of yet, we cannot conclude about the source of the pattern, but it is under investigation. | 153 |
| 5.12 | The $\tau_+^{\text{all}}(\Delta)$ (top left), $\tau_-^{\text{all}}(\Delta)$ (top right), $\tau_{\text{mean}}^{\text{all}}(\Delta)$ (bottom left), and $\tau_{\text{diff}}^{\text{all}}(\Delta)$ (bottom right) travel time measurements from the shifted TD results of the full CH data. Travel times have units of seconds ($dt = 45$ s). Δ has units of pixels. Recall, 1 HMI pixel = $0.06^\circ \simeq 0.729$ Mm. | 154 |
| 5.13 | The $\tau_+^{\text{in}}(\Delta)$ (top left), $\tau_-^{\text{in}}(\Delta)$ (top right), $\tau_{\text{mean}}^{\text{in}}(\Delta)$ (bottom left), and $\tau_{\text{diff}}^{\text{in}}(\Delta)$ (bottom right) travel time measurements from the shifted TD results inside the CH region. Travel times have units of seconds ($dt = 45$ s). Δ has units of pixels. Recall, 1 HMI pixel = $0.06^\circ \simeq 0.729$ Mm. | 155 |
| 5.14 | The τ_+^{out} (top left), τ_-^{out} (top right), $\tau_{\text{mean}}^{\text{out}}$ (bottom left), and $\tau_{\text{diff}}^{\text{out}}$ (bottom right) travel time measurements from the shifted TD results outside of the CH region. Travel times have units of seconds ($dt = 45$ s). Δ has units of pixels. Recall, 1 HMI pixel = $0.06^\circ \simeq 0.729$ Mm. | 156 |
| 5.15 | Basic difference plots between the four sets of compared TD results. Δ has units of pixels, and \mathfrak{t} has units of seconds ($dt = 45$ s). These plots are mainly for reference against the superpixel-average t -statistic results in Figures 5.16-5.19. | 162 |

- 5.16 The Student's t -test comparison between the shifted TD results inside and outside the CH boundary. Δ has units of pixels, and t has units of seconds ($dt = 45$ s). The top row of results employs a 3×3 superpixel; the bottom row, a 5×5 superpixel. Full maps of the t -statistic are shown on the left. The center images show the binary $\geq 95\%$ significance maps of the left-hand results. Black indicates statistically significant differences and white indicates otherwise. We threshold the t -test results on the left with the center significance masks to produce the plots on the right. These results show only the significantly different TD signals between the inner and outer regions of the CH boundary. Note that small-scale differences between the main TD branches are averaged out with this technique. 163
- 5.17 The Student's t -test comparison between the shifted TD results of the full CH data set and separate quiet sun data. Δ has units of pixels, and t has units of seconds ($dt = 45$ s). The top row of results employs a 3×3 superpixel; the bottom row, a 5×5 superpixel. Full maps of the t -statistic are shown on the left. The center images show the binary $\geq 95\%$ significance maps of the left-hand results. Black indicates statistically significant differences and white indicates otherwise. We threshold the t -test results on the left with the center significance masks to produce the plots on the right. These results show only the significantly different TD signals between the full CH data and the quiet sun data. Note that small-scale differences between the main TD branches are averaged out with this technique. 164
- 5.18 The Student's t -test comparison between the shifted TD results inside the CH region and the separate quiet sun data. Δ has units of pixels, and t has units of seconds ($dt = 45$ s). The top row of results employs a 3×3 superpixel; the bottom row, a 5×5 superpixel. Full maps of the t -statistic are shown on the left. The center images show the binary $\geq 95\%$ significance maps of the left-hand results. Black indicates statistically significant differences and white indicates otherwise. We threshold the t -test results on the left with the center significance masks to produce the plots on the right. These results show only the significantly different TD signals between the CH region and the unrelated quiet sun region. Note that small-scale differences between the main TD branches are averaged out with this technique. 165

- 5.19 The Student's t -test comparison between the shifted TD results outside of the CH boundary and the separate quiet sun data. Δ has units of pixels, and t has units of seconds ($dt = 45$ s). The top row of results employs a 3×3 superpixel; the bottom row, a 5×5 superpixel. Full maps of the t -statistic are shown on the left. The center images show the binary $\geq 95\%$ significance maps of the left-hand results. Black indicates statistically significant differences and white indicates otherwise. We threshold the t -test results on the left with the center significance masks to produce the plots on the right. These results show only the significantly different TD signals between the quiet sun region near the CH boundary and the unrelated quiet sun region. Note that small-scale differences between the main TD branches are averaged out with this technique. 166
- 5.20 The $\geq 95\%$ significance of differences between the shifted travel time measurements inside and outside the CH boundary. Shown are the significant t -statistics of $\tau_+(\Delta)$ (middle top), $\tau_-(\Delta)$ (top), $\tau_{\text{mean}}(\Delta)$ (middle bottom), and $\tau_{\text{diff}}(\Delta)$ (bottom). Travel times have units of seconds ($dt = 45$ s). Δ has units of pixels. Recall, 1 HMI pixel = $0.06^\circ \simeq 0.729$ Mm. 167
- 5.21 The $\geq 95\%$ significance of differences between the shifted travel time measurements of the entire CH data and the separate quiet sun data. Shown are the significant t -statistics of $\tau_+(\Delta)$ (middle top), $\tau_-(\Delta)$ (top), $\tau_{\text{mean}}(\Delta)$ (middle bottom), and $\tau_{\text{diff}}(\Delta)$ (bottom). Travel times have units of seconds ($dt = 45$ s). Δ has units of pixels. Recall, 1 HMI pixel = $0.06^\circ \simeq 0.729$ Mm. 168
- 5.22 The $\geq 95\%$ significance of differences between the shifted travel time measurements inside the CH and the unrelated quiet sun data. Shown are the significant t -statistics of $\tau_+(\Delta)$ (middle top), $\tau_-(\Delta)$ (top), $\tau_{\text{mean}}(\Delta)$ (middle bottom), and $\tau_{\text{diff}}(\Delta)$ (bottom). Travel times have units of seconds ($dt = 45$ s). Δ has units of pixels. Recall, 1 HMI pixel = $0.06^\circ \simeq 0.729$ Mm. 169
- 5.23 The $\geq 95\%$ significance of differences between the shifted travel time measurements of the region outside of the CH boundary and the separate quiet sun data. Shown are the significant t -statistics of $\tau_+(\Delta)$ (middle top), $\tau_-(\Delta)$ (top), $\tau_{\text{mean}}(\Delta)$ (middle bottom), and $\tau_{\text{diff}}(\Delta)$ (bottom). Travel times have units of seconds ($dt = 45$ s). Δ has units of pixels. Recall, 1 HMI pixel = $0.06^\circ \simeq 0.729$ Mm. 170

C.1 To the left are spatial plots of the Gaussian function of Equation C.2 (top) and the piecewise function of Equation C.6 (bottom). To the right are the corresponding Fourier transform functions of Equation C.5 (top) and Equation C.7 (bottom). The constant α in the equations is arbitrarily set to $\alpha = 1.5$ for these calculations. 191

Abstract

CORONAL HOLES AND SOLAR f -MODE WAVE SCATTERING OFF LINEAR BOUNDARIES

Shea A. Hess Webber, PhD

George Mason University, 2016

Dissertation Director: Dr. W. Dean Pesnell

Coronal holes (CHs) are solar atmospheric features that have reduced emission in the extreme ultraviolet (EUV) spectrum due to decreased plasma density along open magnetic field lines. CHs are the source of the fast solar wind, can influence other solar activity, and track the solar cycle. Our interest in them deals with boundary detection near the solar surface. Detecting CH boundaries is important for estimating their size and tracking their evolution through time, as well as for comparing the physical properties within and outside of the feature.

In this thesis, we (1) investigate CHs using statistical properties and image processing techniques on EUV images to detect CH boundaries in the low corona and chromosphere. SOHO/EIT data is used to locate polar CH boundaries on the solar limb, which are then tracked through two solar cycles. Additionally, we develop an edge-detection algorithm that we use on SDO/AIA data of a polar hole extension with an approximately linear boundary. These locations are used later to inform part of the helioseismic investigation; (2) develop a local time-distance (TD) helioseismology technique that can be used to detect CH boundary signatures at the photospheric level. We employ a new averaging scheme that makes use of the quasi-linear topology of elongated scattering regions, and create simulated data to

test the new technique and compare results of some associated assumptions. This method enhances the wave propagation signal in the direction perpendicular to the linear feature and reduces the computational time of the TD analysis. We also apply a new statistical analysis of the significance of differences between the TD results; and (3) apply the TD techniques to solar CH data from SDO/HMI. The data correspond to the AIA data used in the edge-detection algorithm on EUV images. We look for statistically significant differences between the TD results inside and outside the CH region.

In investigation (1), we found that the polar CH areas did not change significantly between minima, even though the magnetic field strength weakened. The results of (2) indicate that TD helioseismology techniques can be extended to make use of feature symmetry in the domain. The linear technique used here produces results that differ between a linear scattering region and a circular scattering region, shown using the simulated data algorithm. This suggests that using usual TD methods on scattering regions that are radially asymmetric may produce results with signatures of the anisotropy. The results of (1) and (3) indicate that the TD signal within our CH is statistically significantly different compared to unrelated quiet sun results. Surprisingly, the TD results in the quiet sun near the CH boundary also show significant differences compared to the separate quiet sun.

Chapter 1: Introduction

This dissertation is an amalgamation of two distinct problems in solar physics. Both stem from the overarching goal of understanding the solar dynamo, which is the mechanism that generates and evolves the Sun’s magnetic field (Charbonneau, 2014; Tobias, 2002). The two topics approach the solar dynamo from, literally, opposite ends. The first, the tracking and modeling of open magnetic field regions corresponding to coronal holes, is a “top-down” approach that studies solar atmospheric conditions and links the evolution of the open flux with internal solar dynamo processes (e.g., Fisk and Schwadron (2001); Pinto *et al.* (2011); Zirker (1977)). Coronal holes are also of interest as sources of the fast and slow solar wind, which directly affects conditions in near-Earth space and the heliosphere (Neugebauer *et al.*, 1998; Nolte *et al.*, 1976; Ohmi *et al.*, 2004; Schwenn, 2006; Wang *et al.*, 1998, 2000). The second, “bottom-up” approach uses helioseismology to probe the solar photosphere and convection zone to determine magnetic field and plasma flow properties at and below the Sun’s surface, where the dynamo is predicted to originate (Charbonneau, 2014; Tobias, 2002). The physics between the solar interior and atmosphere is still a widely open subject in solar physics, and connecting these seemingly disjointed layers could have a broad impact on our understanding of the dynamo, as well as plasma physics, magnetohydrodynamics, and space weather prediction.

In this work, we investigate coronal holes (generally studied with “top-down” methods) at the photosphere using a new “bottom-up” helioseismic technique. Boundary detection processes are combined with time-distance local helioseismology to distinguish differences in wave properties between a coronal hole region and the quiet sun, as well as potentially reveal backscattered waves. Because boundary regions are elongated features, the scattering we are considering is better represented as from a quasi-linear scattering surface than from

a point scatterer. To enhance the estimation of wave properties, we study a region of a large polar coronal hole extension with an approximately linear boundary. We also develop a simulation on which we test our linear time-distance method. In future, this method could easily be applied to other elongated solar magnetic features, such as sunspot boundaries and filaments.

The remainder of this chapter is devoted to brief synopses of the general topics necessary for understanding the scientific techniques and results of this work. In section 1.1, we review the history and current physical understanding of coronal holes. In Section 1.2, we briefly cover wave modes, focusing on f -mode oscillations, which is vital for our helioseismic discussion. Finally, we give a general overview of (specifically, local) helioseismology and its history in Section 1.3, which we will expand upon in more detail later in Chapter 3.

In Chapter 2, we investigate coronal holes using statistical properties and image processing techniques on extreme ultraviolet (EUV) images in order to detect coronal hole boundaries in the low corona and chromosphere. We develop an edge-detection algorithm used on a polar hole extension with an approximately linear boundary; the data is from the Solar Dynamics Observatory's Atmospheric Imaging Assembly instrument (SDO/AIA; Lemen *et al.*, 2012; Pesnell, Thompson, and Chamberlin, 2012). These locations are used later to inform part of the helioseismic investigation. Appendix A also includes some more specific details about this process. Additionally, data from the Solar and Heliospheric Observatory's Extreme ultraviolet Imaging Telescope (SOHO/EIT; Delaboudinière *et al.*, 1995) is used to locate polar coronal hole boundaries on the solar limb, which are then tracked through two solar cycles (Hess Webber *et al.*, 2014; Kirk *et al.*, 2009).

Chapter 3 includes two in-depth reviews of (3.1) wave and scattering theory and (3.2) time-distance local helioseismology. These topics are the basis for our new helioseismic technique, as well as the way in which we test the new approach. We also present the mathematical and geometrical symmetry arguments that, when applied to time-distance helioseismology, constitute the foundation of our technique. We employ a new averaging scheme that makes use of the linear geometry of elongated scattering regions. This method

enhances the wave propagation signal in the direction perpendicular to the linear feature and reduces the computational time of the TD analysis. Appendices B and C, respectively, provide further instruction on the fundamental mathematics needed for these two reviews.

Chapter 4 describes the analytic simulation we have developed to test the new time-distance technique, and presents results found using the technique on simulated data. We also perform a statistical analysis on these time-distance results to probe for significant differences between simulated cases. The simulation solves the scattered wave equation for an f -mode plane wave interacting with a rigid cylinder, aligned perpendicular to the axes of wave propagation. The simulation does not take any magnetic field into account, and thus is not a true representation of the solar conditions (i.e., no absorption, mode conversion, etc.); however, it is a representative model of the scattered wave interference expected from an elongated scattering region (e.g., our coronal hole boundary), and serves as proof-of-concept data for our time-distance method.

Chapter 5 presents the solar data and results found after employing our time-distance technique. We study two data sets derived from the SDO Helioseismic and Magnetic Imager (HMI) instrument (Pesnell, Thompson, and Chamberlin, 2012; Schou *et al.*, 2012). A broadband f -mode phase-speed filter is applied in both cases. The first is a quiet sun region. The second is the coronal hole region corresponding to the SDO/AIA data studied in Chapter 2. We compare the time-distance results between these two data sets using the same statistical analysis used to compare the simulated data sets.

Finally, we present and discuss our conclusions in Chapter 6. Possibilities for future work are also considered.

1.1 Coronal Hole Overview

Coronal holes (CHs) are regions of open magnetic field along which solar plasma particles flow freely away from the Sun. This outflow causes the density and electron temperature in the CHs to be lower than in the ambient corona and chromosphere (the solar “atmosphere”) (Munro and Withbroe, 1972). They are the source of the fast solar wind, dynamically change

with the solar cycle, and are key to understanding the Sun’s large-scale magnetic dynamo process.

Polar coronal holes (PCHs) are long-lived features that primarily appear at latitudes above $\pm 50^\circ$ (Waldmeier, 1981). These holes have a fixed rate of rotation distinct from the polar plasma rotation (Navarro-Peralta and Sanchez-Ibarra, 1994), which suggests that the field lines are less prone to twisting and distortion from sub-surface flows. PCHs are dominant during solar minimum. Equatorial coronal holes (ECHs) are more transient features, lasting hours up to weeks, and are more common during solar maximum. This type of hole is subject to magnetic disturbances caused by flares and coronal mass ejections. ECHs also rotate at the same rate as the solar differential rotation (Navarro-Peralta and Sanchez-Ibarra, 1994). Polar hole extensions share similar properties to the PCHs (fixed rotation rate, long-lived) but stretch down into lower latitudes where ECHs dominate.

During solar minimum, the Sun’s magnetic field is quasi-dipolar. The open magnetic field migrates toward either pole and large stable PCHs are visible in EUV and X-ray wavelengths. As the solar cycle progresses into maximum, the magnetic field gets wrapped around in the high- β plasma of the convection zone, turning the poloidal field into a toroidal one. The PCHs break apart and move down to lower latitudes as small, more transient ECHs. Zhou and Smith (2009) found that, over four solar cycles, the amount of magnetic flux from open polar magnetic field at minimum is roughly the same as the amount of magnetic flux from globally open fields at maximum.

There are two viewpoints about the structure and origin of CHs. Some in the scientific community assert that CHs are purely atmospheric features that are created and destroyed in the corona and only extend down to the top of the photosphere (the Sun’s “surface”). Others assert that CHs are driven by internal processes in the Sun and their associated open magnetic field lines propagate into the convection zone. We will briefly discuss the research supporting both viewpoints.

One widely-used theory that supports the idea that CHs are entirely driven by atmospheric processes is the quasi-steady coronal model (e.g., Antiochos *et al.*, 2007; Edmondson

et al., 2009; Schrijver and DeRosa, 2003). The quasi-steady model was originally proposed by Parker (1958). It suggests that at some distance into the corona from the photosphere, the gas pressure begins to dominate over the magnetic pressure and pulls the magnetic field lines outward. This gives rise to the idea of a “source surface,” below which all field lines are closed while field lines that cross it are defined as open (e.g., Altschuler and Newkirk, 1969; Hoeksema and Scherrer, 1986; Schatten, Wilcox, and Ness, 1969). The Potential-Field Source Surface (PFSS) model is still a common method used in global coronal magnetic field modeling. More recently, quasi-steady coronal modeling efforts have extended into the realm of magnetohydrodynamic (MHD) modeling (e.g., Linker *et al.*, 1999; Roussev *et al.*, 2003), which does not require some of the physical assumptions necessary for the PFSS model (e.g., current-free coronal magnetic field conditions). Another model that could support the idea of coronal holes being purely atmospheric features is the interchange reconnection model, which describes changes in the coronal magnetic field via interactions between open and closed field lines in the corona (Fisk, Zurbuchen, and Schwadron, 1999). However, this model has also been used to support the alternative theory that the formation and destruction of CHs may be partially driven by internal solar processes, as we discuss next.

The second theory argues that the open magnetic flux associated with CHs propagates into the convection zone and is influenced by motions therein. Bohlin (1976) suggested that CH formation is a byproduct of by the “unbalanced” flux emergence of bipolar magnetic regions. Bohlin (1977) subsequently found that CHs tend to grow and decay at the average rate of a supergranular cell. Around the same time, McIntosh (1976) proposed that CHs form within regions of unipolar magnetic cells at low latitudes, influenced by nearby PCH extensions of like polarity with which the ECHs then merge. He attributed this behavior to the evolution of global magnetic neutral lines. In 2001, Fisk and Schwadron proposed that the total amount of open magnetic flux of the Sun should remain approximately constant in time because it can only be reduced by closure with open flux of opposite polarity. This is unlikely in CHs since they have uniform polarity, particularly if the CH has a large area.

They argue that the behavior of the open flux must therefore be regulated by motions in the photosphere. Particularly, they assert that diffusion processes and convective motions govern the evolution of CHs both locally (opening and closing of flux at the boundaries) and globally (over the course of a solar cycle). Pinto *et al.* (2011) used numerical MHD simulations to study the coupling of the solar dynamo with the corona and solar wind. They look at the response of open magnetic field regions in the corona to changes in the evolving dynamo through a solar cycle. Their simulation shows a clear correspondence between their dynamo model and the formation and evolution of coronal holes. Although the dynamo model is a simplified one (e.g., magneto-rotational effects are neglected, toroidal fields and flows are not considered) and the results are somewhat idealized (e.g., the models produce regular and symmetric solar cycles), they note that the coronal features produced are consistent with the results of similar work in modeling of the coronal field evolution (Wang and Sheeley, 2003). There is also evidence that coronal holes and the fast solar wind are associated with the chromospheric network, which in turn has been found to correspond with the boundaries of supergranular cells that are driven by convection (Hassler *et al.*, 1999; Simon and Leighton, 1964).

Attempting to examine CHs from below the surface could provide an answer to this highly debated topic. Recent research in local helioseismology (see Section 1.3) has been successful at finding wave scattering signatures off of strong, closed magnetic flux below the photosphere. We suggest that similar wave scattering could occur at sub-photospheric CH boundaries, since the closed-to-open magnetic field boundary acts as a discontinuity that could reflect some incoming waves.

1.2 Wave Modes Overview

There are three kinds of waves used in helioseismology, each of which generates a different resonant wave mode. p -modes are acoustic oscillations, driven by pressure gradients; g -mode oscillations have gravity as the restoring force; f -mode oscillations are surface-gravity

waves. Each type of wave mode samples different depths of the solar interior. There are many harmonic resonance patterns within each mode category.

We focus on f -mode waves for the purpose of this dissertation. f -modes are surface waves that propagate horizontally across the photosphere, with gravity acting as the restoring force. In fluid dynamics, surface-gravity wave propagation is described by Airy (or linear) wave theory (Craig, 2004). A simple schematic is shown in Figure 1.1, from Feynman, Leighton, and Sands (1963). Surface waves are a reasonable starting point for this research because, as mentioned in Section 1.1, the research question is how deep below the photosphere do CHs make their presence known.

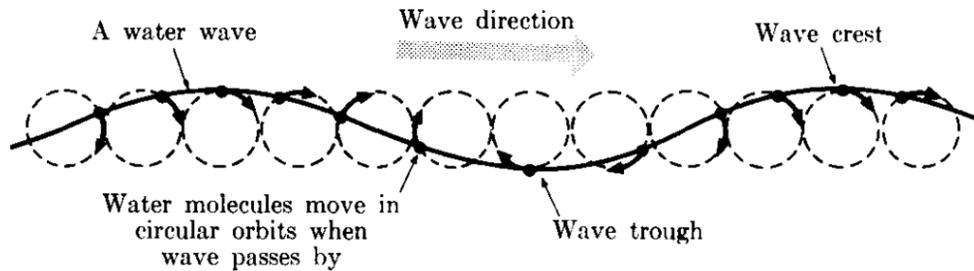


Figure 1.1: f -mode waves are similar to deep-water surface waves in the ocean. The wavefronts propagate horizontally across the surface while the particle motion is circular. As mentioned in the attributed text, the circular motion decreases in radius with increasing depth until the motion disappears completely. (Figure 51-9, Feynman, Leighton, and Sands, 1963)

The main difference between Airy wave theory and the study of waves on the Sun is the inclusion of the magnetic field. Airy wave theory is a hydrodynamic description of wave propagation (charge-neutral fluid interacting with charge-neutral perturbations), whereas the waves on the Sun are plasma waves that are influenced by and interact with the solar magnetic field; in other words, MHD. To illustrate this, consider the hydrodynamic scenario of a deep lake with an island in the center. A speed boat moving through the water creates a wake that travels as a plane wave across the lake surface, eventually intersecting the island.

The interaction with the island creates a scattered wave, the “shape” (i.e., interference pattern) of which is dependent on the “shape” (i.e., structure) of the island. In the MHD approximation, the plasma takes place of the water in the lake and the magnetic field is a barrier like the island. We can study the scattering of plasma waves off of the magnetic field in order to understand the structure of that field.

f -mode waves have a radial order of $n = 0$ (there are no nodes in the radial eigenfunction of the mode; Duvall, Jr. and Gizon (2000)). On the Sun, the peak of the f -mode frequency envelope is about 3 mHz, with a FWHM of 1 mHz. This corresponds to a spherical harmonic degree range of $600 \leq \ell \leq 1200$, with a peak at about $\ell = 880$. Note that there are no observable f -mode frequencies below 0.661 mHz ($\ell = 41$) (Schou, 2004).

The dispersion relationship for f -mode waves is similar to that of deep-water surface waves (in the limit that the wavelength is small compared to the water depth):

$$\omega^2 = gk, \tag{1.1}$$

where ω is the angular temporal frequency, g is the gravitational acceleration at the photosphere, and k_H is the horizontal wavenumber. Note that $k_H = (\ell + 1/2)/R_\odot$, where ℓ is the spherical harmonic degree and R_\odot is the solar radius (Duvall, Jr. and Gizon, 2000). f -mode oscillations are sensitive to both horizontal flow components, which allows more accuracy for scattering detections than Doppler measurements that only see along the line-of-sight (LOS). Also, the f -mode has known eigenfunctions, so the depth that is being averaged is well defined.

1.3 Helioseismology Overview

Helioseismology is the study of the solar interior from surface observations of naturally-excited internal acoustic and surface-gravity waves (Gizon and Birch, 2005). Local helioseismology, a term coined by Lindsey, Braun, and Jefferies (1993), involves the study of small regions on the solar disk and is used to create a detailed picture of the interior in

that particular region. There are multiple methods used in local helioseismic research, including ring diagram analysis, time-distance techniques, and acoustic holography (Pijpers, 2006). All current local helioseismic techniques assume wave scattering off of a point source disturbance.

The specific local technique that will be used in this dissertation is called time-distance (TD) helioseismology, first developed by Duvall, Jr. *et al.* (1993). TD helioseismology computes correlations in far-field wave patterns between two surface points. A recent review on this technique is presented by Gizon, Birch, and Spruit (2010), and we also briefly describe the method in Section 3.2. Our specific procedure varies from the general TD methodology by our exploitation of elongated feature geometry, allowing us to perform the analysis in 1D and investigate a magnetic field configuration previously unexplored. Precedence for our new technique is discussed in Section 3.2.3.2 (i.e., Cameron, Gizon, and Duvall, Jr., 2008; Duvall, Jr. *et al.*, 1996; Gizon and Birch, 2005; Hindman *et al.*, 2004).

Local helioseismology has been used to look at magnetic structure in the solar interior through flow anomalies. The initial idea for this dissertation was motivated by the work of Ilonidis, Zhao, and Kosovichev (2011). Their research employed a time-distance analysis technique and was able to predict emerging flux associated with active regions up to two days before the appearance of sunspots. We discuss implementation of time-distance helioseismology analysis in Section 3.2. Howe *et al.* (2013) used a point source analysis technique, called the ring-diagram method, to look at differences between quiet sun regions and CH regions. They found that while there was no apparent change in wave frequency and mode amplitude between quiet sun and CHs, there is evidence of less scattering from points within the CH. They attribute this to the simple magnetic field geometry and fewer emerging flux disturbances from within the hole areas. This leads to intriguing questions about what occurs at the boundary between CHs and quiet sun regions and whether we can use alternative local helioseismology techniques to investigate the subsurface properties below CHs.

Chapter 2: Coronal Hole Detection & Results

This dissertation requires knowing the location of CH boundaries. CH regions are not visible in doppler velocity data, and can even be difficult to distinguish in magnetograms. We must therefore use EUV images to detect the CHs in the chromosphere and corona, approximating their projected boundary locations in the solar atmosphere as the open magnetic field boundary location in the photosphere. In this chapter, we discuss our CH boundary detection process, from data retrieval through the detection algorithm to the boundary location results of the CH in question. Sections 2.1 and 2.2 consider the detection method and results, respectively. Additionally, we review our somewhat related previous work on PCH boundary detection using EUV images in Section 2.3.

2.1 Data Preparation and Detection Method

Prior to performing a helioseismic analysis on the SDO/HMI data, steps must be taken to align the CH boundary to the y -dimension of the domain. This allows us to take advantage of symmetry arguments for an elongated feature, as well as to approximate the boundary location and width. However, the boundary is not a visible feature in the HMI data. We therefore use the corresponding AIA data to determine the alignment and location parameters that will be applied later to the HMI images. The retrieval and post-processing methods are fairly straightforward, but many details and assumptions must be recorded for reproducibility, and to understand the parameters that will be applied to the HMI data. The steps and details of the AIA image post-processing method are discussed in this section. All of the code used for this process, as well as the TD analysis later on, is written in the Interactive Data Language (IDL; version 7.0, Unix) with heavy reliance on the SolarSoft (SSW) library.

2.1.1 Download Process

The initial step is retrieving the AIA data set that directly corresponds to the HMI data. This means not only the same time span, but also the same dimensions and location on the disk. Both data sets are archived by Stanford University’s Joint Science Operations Center (JSOC). The data can be exported via the internal JSOC export system¹. In this section, we walk through the export process for our AIA data, noting that the procedure is mainly the same for retrieving the HMI data as well.

There are multiple AIA product series available for download. The main archived data set is the AIA level-1 series (*aia.lev1*). As opposed to the raw level-0 data, the level-1 data have been processed for pixel-level calibrations including bad-pixel removal, despiking, and flat-fielding. The processing of level-1.5 data further corrects for geometric calibrations, such as plate-scaling, rotation to solar north, and shifting the image centers to match the HMI data. Further details on the differences between data processing products can be found in the SDO documentation website², run by the Lockheed Martin Solar and Astrophysics Laboratory (LMSAL). In particular, the PDF link to N. Hurlburt’s “AIA plan for producing and distributing data”³ (last update November 19, 2010, as of writing this dissertation) describes the various AIA level differences in detail. Another useful resource available via LMSAL’s SDODocs website is M. DeRosa’s “Guide to SDO Data Analysis”⁴ (last update July 27, 2016, as of writing this dissertation).

The only fully-archived AIA data series is the level-1 processed data. There are several options for converting a level-1 download to level-1.5, such as using the SSW routine *aia_prep.pro*. There are also processing options embedded into the JSOC export system that perform the same conversion, one of which will be discussed in detail later in this section.

The JSOC export system has a cutout service that allows the user to request only a

¹<http://jsoc.stanford.edu/ajax/exportdata.html>

²<https://www.lmsal.com/sdodocs/>

³https://www.lmsal.com/sdodocs/doc?cmd=dcurl&proj_num=SDOD0045&file_type=pdf

⁴<http://www.lmsal.com/sdouserguide.html>

specific region of the archived images. We use this service as our means of data retrieval. The user can define the desired coordinates of the image center, as well as the dimensions of the domain. We know what cutout parameters the desired AIA data requires, based on the location and dimensions of the HMI data cube (see Appendix A.1). Figure 2.1 shows a screenshot of the JSOC export website with parameter specifications for one wavelength of the SDO/AIA data set used in this work. We will move our discussion through the export page, starting at the top.

The screenshot displays the JSOC Data Export Request Generation webpage. The form is organized into several sections:

- RecordSet from file:** Includes a checkbox for allowing RecordSet list file uploads and a text input for the RecordSet identifier: `aia.level1[2014.02.05_8:00-2014.02.05_16:00][? WAVELENGTH=211 ?]`. A `Lookdata` button is present.
- Record Limit:** Set to `none`. An optional manual limit to number of records to export.
- Record Count:** Set to `2400`. A `Recount` button is available. A note: Limit for AIA to about 15,000 and for HMI about 30,000 in each request.
- Method:** Set to `url-tar`. A note: Choose method, url_quick: or url for now. url_quick implies protocol of "as-is".
- Filename Format:** Set to `aia.level1.{T_OBS:A}. {WAVELENGTH}.proj.fits`. A note: File name template.
- Processing:** Includes checkboxes for `no_op - none`, `hide`, `aia_scale - Scale AIA level1 to 0.6 arcsec/pixel` (checked), `HmiB2ptr - Convert HMI B_720s to Phi,Theta,R coordinates`, `resize - Resize and rotate if needed, use sub-pixel registration`, `im_patch - Extract sub-frame`, `maproj - Extract a sub-frame and remap to a chosen projection` (checked), and `rebin - Rebin with boxcar or gaussian smoothing`.
- AIA_Scale:** A note: Normalize image scale to AIA standard 0.6 arc-sec per pixel is enabled **Only for aia.level1 for now.**
- Maproj:** A detailed section for projection parameters. It includes a text box for `Projected Patch Extract` (used to export sub-frames from a sequence of images). Below it, a note: The location of the extracted sub-frame must be specified in Carrington latitude and longitude. The recordset should be specified fully. For the time being, for AIA at lower cadence use one of the time slotted series: `aia.level1_euv_12s`, `aia.level1_uv_24s`, or `aia.level1_vis_1h`. Options include `Visible grid spacing in degrees` (set to `none`), `NOAA AR number` (set to `NotSpecified`), `CLong` (set to `82`), `Carrington longitude of patch center` (checkbox `Check for disk center longitude`), `CLat` (set to `24`), `Latitude of patch center`, `MapProjection` (set to `Postel's - azimuthal equidistant projection`), `MapScale` (set to `0.0359`), `Scale in degrees per pixel after mapping. Choose <= 0.0359`, `Width` (set to `885`), `Width of extracted patch in pixels`, `Height` (set to `885`), `Height of extracted patch in pixels`, and a `Verify` button with `Check params` and `OK to submit` options.
- Protocol:** Set to `FITS`. A note: Choose protocol, "FITS", "JPEG", "MPG", "MP4", or "as-is". Note uncompressed FITS not an option.
- Notify:** A text input for email address with an `OK` button. A note: Provide your email address for notification. If Requestor is your SolarMail name you may use "solarmail" here.
- Requestor:** Set to `shessweb`. A note: Provide an identifier for you. Optional if not using SolarMail.
- Buttons:** `Check Params for Export` (OK to proceed), `Request Export Status` (Export request submitted, please wait...).
- Footer:** A note: check to show export params. `op=exp_request&ds=aia.level1[2014.02.05_8:00-2014.02.05_16:00][? WAVELENGTH=211 ?]&size=0.04668384790420532&process=n=0[aia_scale][maproj,map=Postel,clon=82,clat=24,scale=0.0359,cols=885,rows=885&requestor=shessweb¬ify=falurex25@gmail.com&method=url-tar&filenamefmt=aia.level1.{T_OBS:A}. {WAVELENGTH}.proj.fits&format=json&protocol=FITS,compress Rice`. Below this, `RequestID` is `JSOC_20160328_456` (This is the ID tag for your export request. Use the Status Request button below to retrieve the links to the data.) and `Status` is `Data Ready, size=1313MB`. `Data Location` is also indicated.

Figure 2.1: A screenshot of the JSOC Data Export Request Generation webpage with parameter specifications for one wavelength of the SDO/AIA data set used in this work.

The **RecordSet** specifies the desired data series, date and time range, and any other keyword options for the export query. It can either be entered manually, if the user knows

the query syntax and desired data set, or it can be built using the JSOC Lookdata service⁵. In accordance with the known HMI data specifications, our **RecordSet** query is for the *aia.lev1* data series on February 5th, 2014, between 8:00 (8AM) and 16:00 (4PM). The wavelength keyword in Figure 2.1 is set to 211 Å; however, we made the same query for three other wavelengths as well: 304 Å, 171 Å, and 193 Å. There is a **Record Count** of about 2400 AIA images that meet the **RecordSet** specifications for each wavelength, one image every ~12 seconds. (The **WAVELNTH=171** query actually has a count of 2401 because the first image per minute for that wavelength is taken on the 00-second, which includes the 16:00:00 time step. This is automatically corrected for in the post-processing.) We leave the **Record Limit** option set to its default “none” since we want all of the available records. We do not actually need the full cadence of this data series; the CH boundary does not change significantly over such a short period, and our analysis uses spatial and temporal averages so that small-scale variations are negligible. We use one image per minute to save on computation time and storage. However, there is currently no way to tell the JSOC to only select the first image per minute for the *aia.lev1* series (there are shortcuts for some other data series, but they are not universal). We therefore download all the data files within the time range and unpack only the desired time files (see Section 2.1.2). Note, this is one reason why we only want a cutout region and not the full resolution, full-disk data. The download and storage of a data set with 2400 images at 4096×4096 spatial resolution takes too much time and space, even in a compressed, tar-file format.

Another minor but time-saving note pertains to the retrieval format. If you do not change the delivery **Method** from url to url-tar, you will have to download each of the 2400 FITS files individually. It is much more convenient to download a single tar file containing all the data files.

The **Filename Format** is the filename template for the returned data files. Different data series have their own default templates. Changing the template is optional. Any

⁵<http://jsoc.stanford.edu/ajax/lookdata.html>

acceptable keyword of the selected data series can be used in the template, surrounded by braces. We adjust the **Filename Format** to differentiate between data by adding the **WAVELNTH** keyword. Added text without braces is included as-typed. We add “.proj.fits” to designate that these FITS files will be postel (azimuthally equidistant) projected (we make note of this and the filename designation for completeness and clarity). Note that changing the **RecordSet** query in any way automatically reverts the **Filename Format** back to its default. To ensure consistently named files from various export requests, check/reset the filename template each time the **RecordSet** is changed (e.g. changing the desired wavelength).

The **AIA_SCALE Processing** option essentially converts the level-1 data to level-1.5 data⁶. It scales and rotates the images according to the metadata information so that it matches the HMI data. Selecting this option removes the need to use *aia_prep.pro*.

The **MAPROJ**⁷ **Processing** option performs two necessary functions. First, it extracts tracked subframes from the full-disk $4K \times 4K$ resolution images. The patches are tracked at the Carrington rate. Second, the extracted patches are remapped to a specified heliographic projection from the inherent helioprojective AIA coordinate system. The available map projections include Plate Carée, cylindrical equal area, orthographic, and Postel’s projection. MAPROJ returns the data at the same processing level in which it was input (e.g. *aia.lev1* returns as level-1). The **JSOC Processing** options are done in the order they are listed on the export page⁶. Thus, using the aforementioned **AIA_SCALE** option will lead to a level-1.5 data product from MAPROJ.

Within the **MAPROJ Processing** procedure, the options for specifying the grid spacing and active region number can be left to their defaults. The **CLong**, **Clat**, **MapProjection**, **Width** and **Height** parameters are all chosen to reflect the corresponding HMI data series. All of this information is determined from the FITS headers of the corresponding HMI data. The central coordinates of the HMI data in Carrington longitude and latitude are

⁶Personal Communication with P. Scherrer, 3/25/2016

⁷http://jsoc.stanford.edu/cvs/JSOC/proj/util/apps/export_as_maproj.c?rev=HEAD&content-type=text/vnd.viewcvs-markup

($+82^\circ, +24^\circ$). These coordinates are the input values for the CLong and CLat parameters, respectively. We select the `postel` MapProjection parameter in accordance with the HMI data projection. The spatial dimensions of the HMI data are 512×512 pixels with a pixel scale of $0.06^\circ/\text{pixel}$ in Heliographic Carrington coordinates, which corresponds to a domain size of $30^\circ \times 30^\circ$. However, the inherent AIA coordinate system and pixel scale are not the same as those for HMI; this must be taken into account in the JSOC cutout parameters. The corresponding AIA domain can be approximated using the SSW World Coordinate System (WCS) commands in IDL (see Appendix A.1 for more details). The MapScale parameter is automatically set to $0.0359^\circ/\text{pixel}$. This can be adjusted but we left it set to the default since it is apparently the closest we can come to the HMI pixel scale (0.06°). Figure 2.1 shows that our desired Width and Height parameters come out to about 885×885 pixels to cover the same “square” 30° region. Note that AIA data is inherently recorded in *arcsec* but the projected cutout data is returned in degrees. We caution that the output FITS header information produced by MAPROJ is not fully converted from helioprojective coordinates to heliographic coordinates. The header information needs to be edited/corrected in order to use WCS routines (see Appendix A.2).

Before submitting an export request, the MAPROJ parameters must be verified by clicking the “Check params” button. If all the necessary parameters are specified, it will display the message “OK to submit”. Similarly, you must click the “Check Params for Export” button to verify that the necessary query information is being submitted. It will display “OK to proceed” if everything is in order for the export. Clicking “Submit Export Request” sends the query to the JSOC server. After successful submission, the **RequestID** for the current **RecordSet** will appear on the page below the submit button. Write this ID down! The form is generally *JSOC_YYYYMMDD_request#*, where the date is the date that the export request was submitted and the request number resets at the beginning of each day. This **RequestID** will also be the filename of the tar file that contains the data FITS files.

The JSOC export service will **Notify** the provided email account when the export process is complete and the data is ready for download. Note, however, that even when a notification email address is provided, experience shows that notifications are not always sent/received. As mentioned, record the **RequestID** for each export query as a precautionary measure so that the status of the request can be manually checked on the website and, when available, download the tar file. The **RequestID** is also important information to provide to the JSOC team if there is an issue that requires debugging.

2.1.2 Image Selection & Unpacking Routines

After download, the tar files are saved in wavelength-specific directories. The specific JSOC RequestID's/filenames of the tar files used in this work are as follows:

```
304 JSOC_20160328_455.tar
171 JSOC_20160328_454.tar
193 JSOC_20160328_456.tar
211 JSOC_20160328_457.tar
```

The filenames of the FITS data within the tar files have the form

aia.lev1.YYYY-MM-DDThhmmssZ.wavelnth.proj.fits

as suggested by the **Filename Format** in Figure 2.1.

The following Unix commands change to the working directory and extract only the first FITS file for every minute.

```
## change directory to location of the tar file
cd DIR

## see the names of the data files that will be extracted from the tar
file
tar -tvf JSOC_*.tar `./aia.lev1.2014-02-05T????[0]?*`

## assuming the last call looked ok, extract the data files into the
current directory
tar -xvf JSOC_*.tar `./aia.lev1.2014-02-05T????[0]?*`
```

Because the data is all taken on the same date, the files can be sorted by time. Since each wavelength only has one image between 00 and 09 seconds every minute, these few lines

effectively extract only the desired files.

The extracted files are rice-compressed and are read into IDL using the *fitsio_read_image.pro*. (The standard IDL *readfits.pro* routine should also work with **/FPACK** keyword set. This may also require download of the FPACK software⁸.) The read-in data is automatically uncompressed by the routine. The headers contain some information about the compressed data that is unnecessary moving forward. The uncompressed header information is updated and the uncompressed FITS files are saved. Recall that there is no need to run *aia_prep.pro* because the JSOC export processing option MAPROJ converts the data to level-1.5 internally.

Figure 2.2 shows the first time step of the \log_{10} data from each wavelength. The CH boundary is apparent in all four wavelengths as a change in intensity from darker (CH) to brighter (quiet sun). The intensity change is particularly noticeable in the 193 Å and 211 Å wavelengths.

⁸<http://heasarc.gsfc.nasa.gov/fitsio/fpack/>

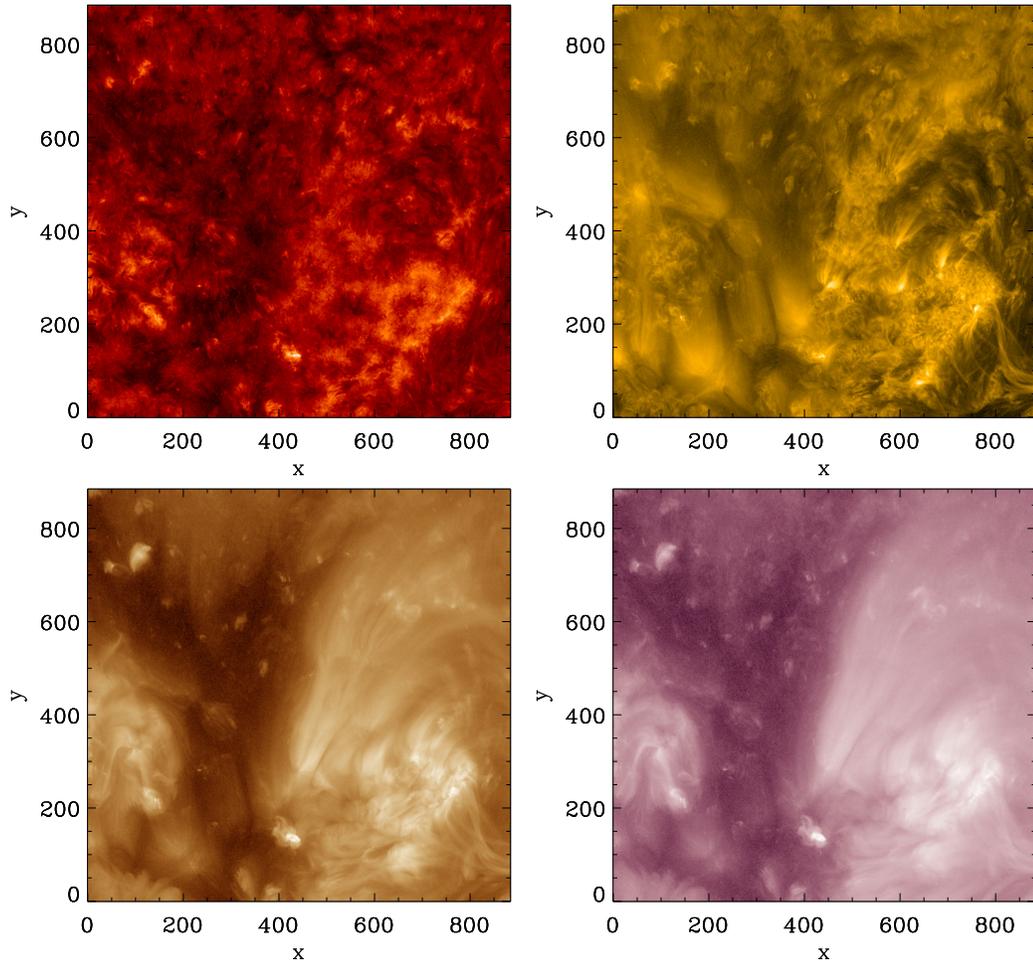


Figure 2.2: SDO/AIA cutout images on Feb. 5 2014 at 8:00 UT (first data time step) in wavelengths 304 Å (top left), 171 Å (top right), 193 Å (bottom left), 211 Å (bottom right). The images shown are \log_{10} scale of the data. Both axes have units of pixels. Recall from Section 2.1.1, 1 AIA cutout pixel = $0.0359^\circ \simeq 0.436$ Mm.

2.1.3 Histogram Filtering & Combining Wavelength Images

From the uncompressed files, we store each wavelength in separate variables. The filenames are parsed for hours and minutes. We then match the available times for each wavelength, discarding any files that do not have data at a given time in all four wavelengths. Looping through the remaining time steps, we iteratively read in the four wavelength files corresponding to each time step, checking that the data dimensions match. If not, we discard the data and continue to the next time step. Otherwise, we begin the post-processing

procedure on the set of images.

First, we perform intensity filtering to improve the contrast in the EUV images. The intensity thresholds are chosen statistically from histograms of each image's intensity. For each wavelength at every time step, we create a histogram of `ALOG10(data)`⁹ with a minimum value of 0 and a maximum value of 10. The bin sizes for each wavelength were determined by visual scrutiny of a random sample of the histograms for that wavelength. The bin sizes need not (*should* not) necessarily be the same for each wavelength. We fit a Gaussian curve to every histogram and store the coefficients. Note that the Gaussian fit is also useful for choosing a reasonable bin size for each wavelength. The coefficients are used to calculate the approximate FWHM of each histogram. From FWHM and Gaussian coefficients, we then determine the intensity limits for each wavelength, which are adjustable but should be consistent between the wavelengths. We choose a low intensity limit of `low = center` and high intensity limit of `high = center + 2*fwhm`, which effectively filters all the below-average and anomalously high intensities, leaving the CH region well defined. The histogram plots and fitted Gaussian curves for the first time step of each wavelength (corresponding to the images in Figure 2.2) are shown in Figure 2.3, along with their defined intensity boundaries.

After determining the intensity threshold values, we next create `BYTSCL()` images of each `ALOG10(data)` wavelength with max and min values set to the intensity limits defined above. The filtered data are then summed between all four wavelengths to form one time step of the combined-wavelength data set. We can check the resulting images to verify that the CH boundary is clearly defined, and that stray light effects and bright loop features are minimized. If necessary, the histogram and intensity limit parameters can be tweaked until the combined images meet those criteria. The first time step of the resulting data is shown in Figure 2.4.

⁹Inherent IDL programs will appear in this style throughout the rest of this section.

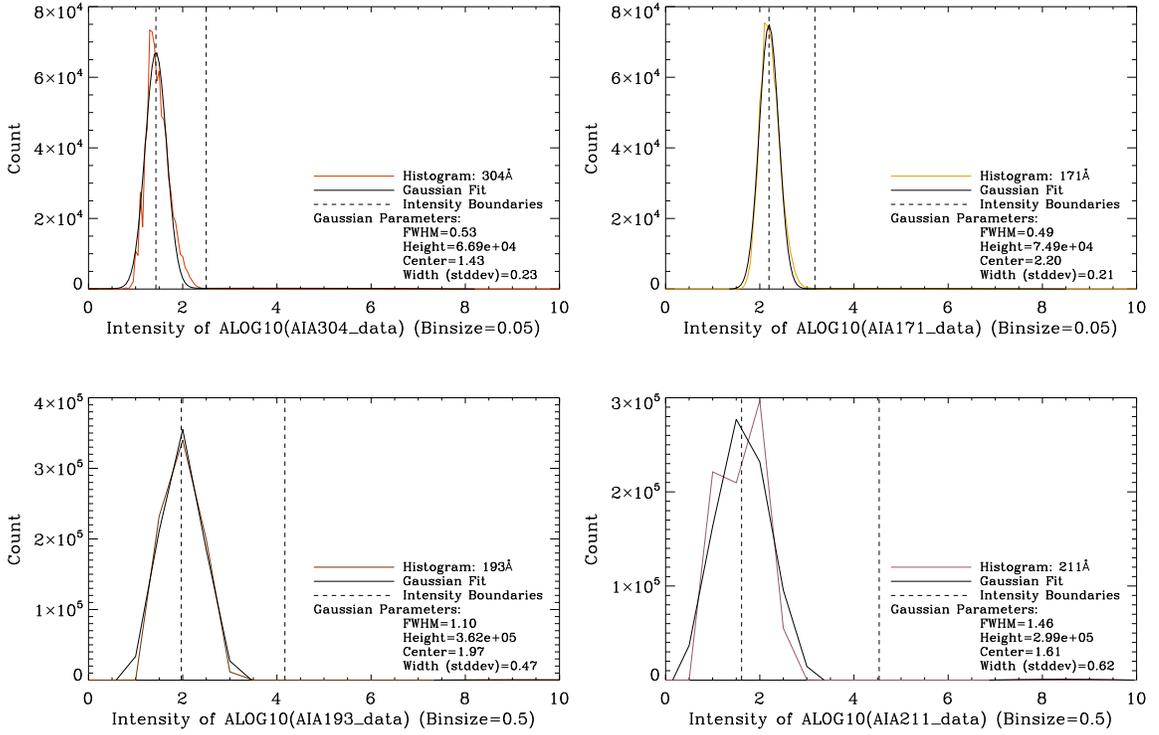


Figure 2.3: Histogram examples corresponding to the SDO/AIA images in Figure 2.2. Again, the plots are arranged by wavelength: 304 Å (top left), 171 Å (top right), 193 Å (bottom left), 211 Å (bottom right). The vertical dashed lines indicate the intensity exposure limits. The lower limit is the Gaussian center, or average intensity. The upper limit is the Gaussian center plus twice the FWHM.

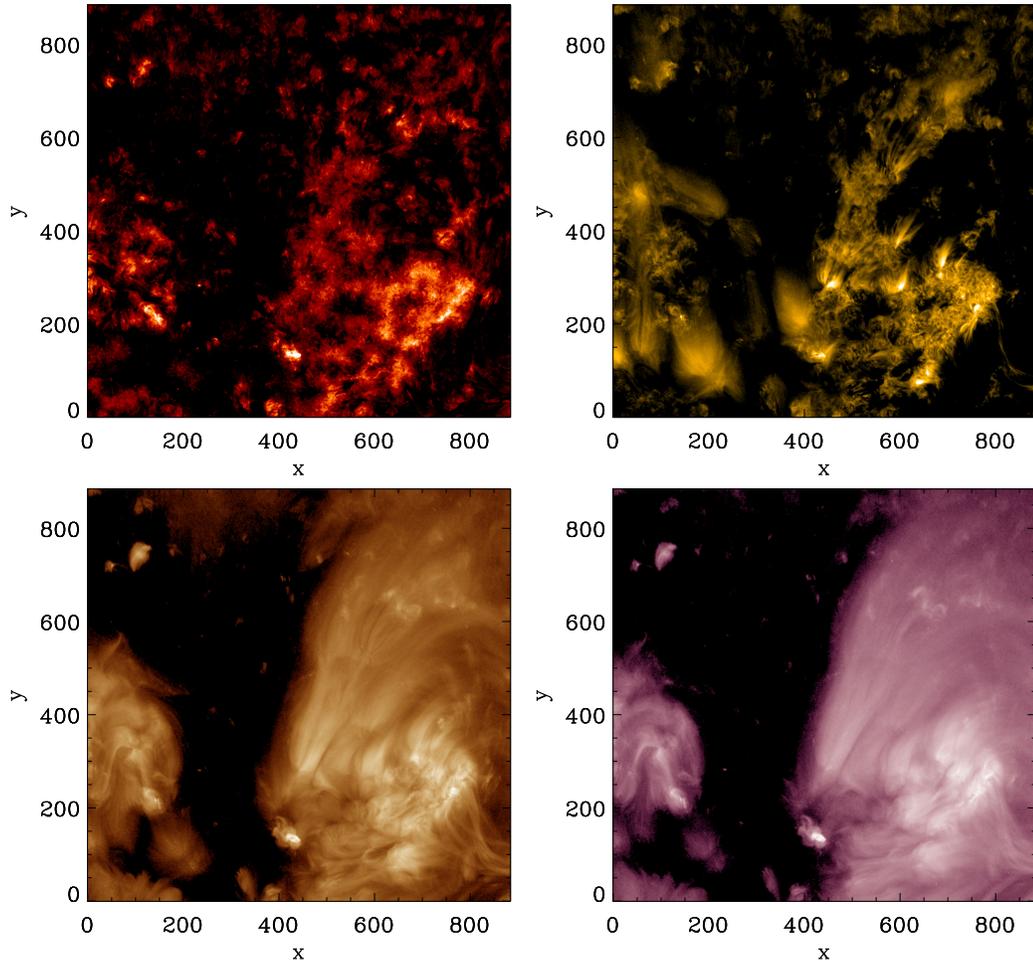


Figure 2.4: The same images as in Figure 2.2, with the intensity limits from Figure 2.3 applied. Both axes have units of pixels. Recall from Section 2.1.1, 1 AIA cutout pixel = $0.0359^\circ \simeq 0.436$ Mm.

2.1.4 Rotation & Cropping

Using IDL's `ROT()` function, rotate the combined image until the linear CH boundary is aligned to the y -axis. We have chosen a rotation of 20° about the center of the image. There is a rotational “smearing” effect at the corners of the image. This needs to be cropped out before further steps can be taken. Cropping reduces the amount of useful data so crop as little of the image as possible while removing the smearing, keeping the image dimensions square. The first time step of the rotated and cropped data is shown in Figure 2.5.

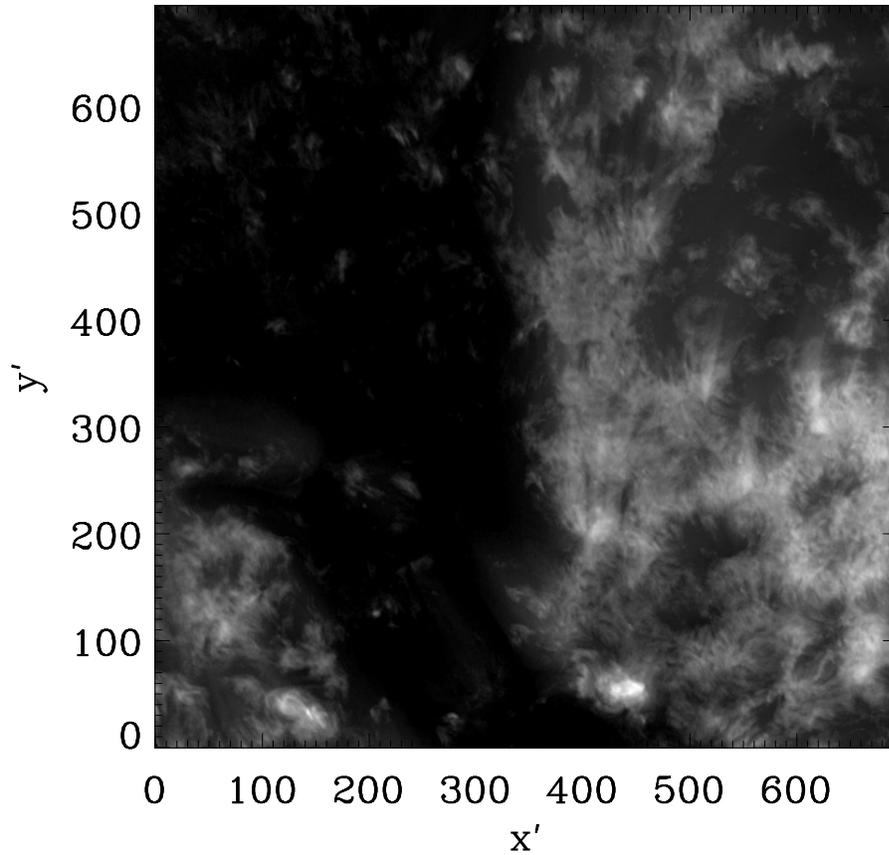


Figure 2.5: The first time step of combined AIA data, rotated 20° and cropped to remove rotation effects at the edges. The notation x', y' is used here to denote rotated coordinates. For simplicity in the remainder of this dissertation, we revert to the notation x, y for the rotated coordinates, unless otherwise noted. Both axes have units of pixels. Recall from Section 2.1.1, 1 AIA cutout pixel = $0.0359^\circ \simeq 0.436$ Mm.

The rotational step can be enhanced using the edge detection techniques discussed later in Section 2.1.6. This would allow for a more quantitative and less arbitrary definition of the rotation angle. For now, the rotation parameter visually aligns the boundary approximately with the y -axis.

We store all time steps of the combination data in a single variable, create the relevant header information (see Appendix A.2), and save the combined data and metadata in a FITS file. This data is a visual representation of the working HMI data, after the same rotational and cropping process is applied.

2.1.5 Reducing to 1D

The entire purpose of rotating and cropping the data is to be able to reduce the data set down to one dimension, taking advantage of the elongated geometry of the feature. Reducing the data not only improves the signal-to-noise ratio, but also reduces the number of cross-correlation calculations necessary to produce a time-distance result by a factor of n , the size of one dimension. This greatly decreases the amount of time it takes to compute the time-distance results.

There are two options for reducing the data to 1D. We can either average over the y -dimension, or sum over the y -dimension. These options only differ by a normalization factor. In this dissertation we use summations to reduce to 1D, both for the boundary detections using AIA data discussed here and in the linear time-distance analysis using the simulated and HMI data (see Section 3.2.3). The result is shown in Figure 2.6.

For the AIA data, a reduction in the time dimension also takes place to further improve the signal-to-noise ratio, shown in Figure 2.7. The boundary between the CH and quiet sun regions is apparent as a smooth transition from an integrated intensity of $\sim 5 \times 10^6$ up to $\sim 8 \times 10^7$. The inflection point and 3rd-order critical points of the transition region represent the center and width of the CH boundary, which we determine in Section 2.1.6 using edge detection techniques. Again, the reduction is accomplished through a summation, as opposed to averaging. The temporal reduction is not done in the corresponding HMI data

analysis.

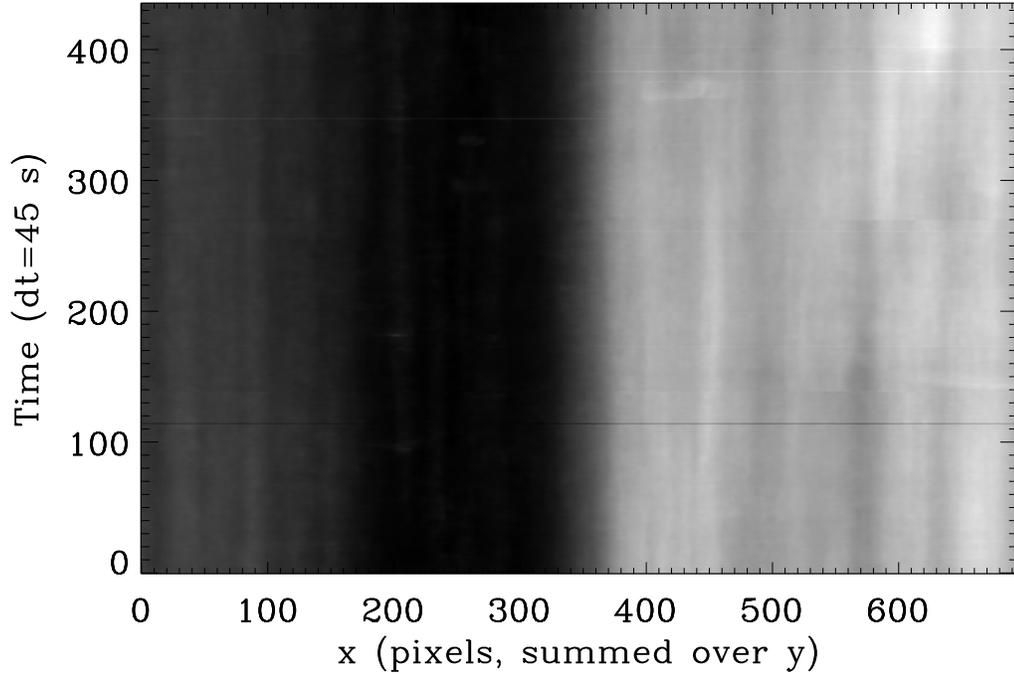


Figure 2.6: The aligned data, summed over the y -dimension, shows the CH boundary through time. The time step along the y -axis has units of seconds ($dt = 45$ s), while the x -axis has units of pixels. Recall from Section 2.1.1, 1 AIA cutout pixel = $0.0359^\circ \simeq 0.436$ Mm.

Summarizing the process so far, we 1) choose an appropriate elongated-CH data set, 2) rotate the images such that the CH boundary is parallel to y -axis (removing rotational effect around edges), 3) sum the data along the y -dimension, and 4) sum the reduced data through time. Next we will smooth the 1D data (and its derivatives) along the x -dimension in order to locate the center and width of the CH boundary.

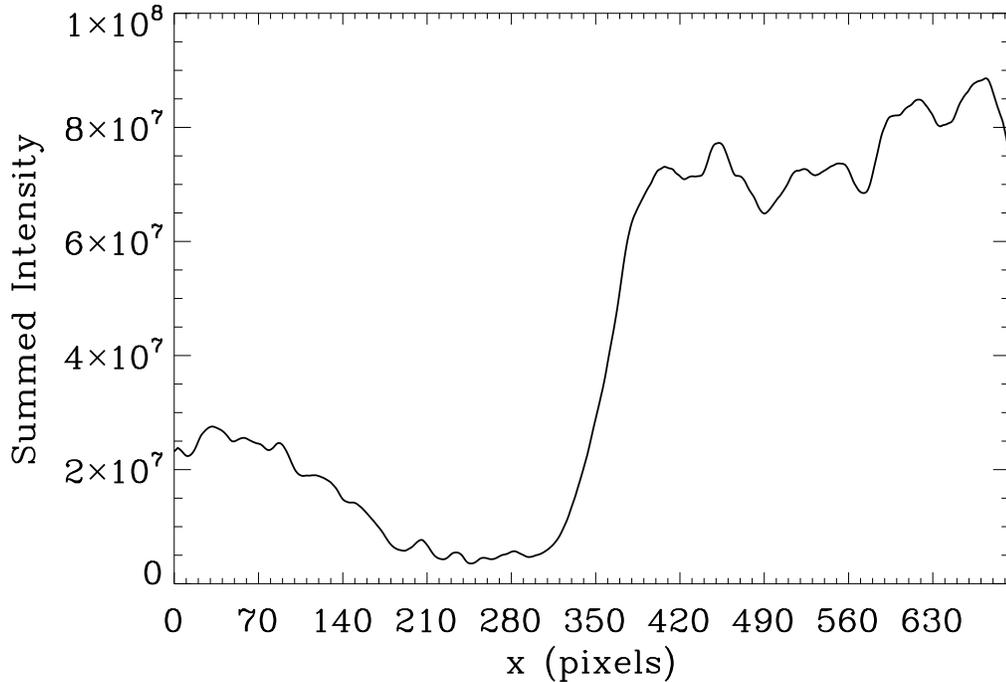


Figure 2.7: Additional summing over the time dimension shows a plot through the x -dimension that displays the transition from the CH region to the quiet sun. The x -axis has units of pixels. Recall from Section 2.1.1, 1 AIA cutout pixel = $0.0359^\circ \simeq 0.436$ Mm. The summed intensity has arbitrary units.

2.1.6 Edge & Width Detection

Savitzky-Golay (SG) smoothing is well-suited for use with data containing a lot of high-frequency noise; particularly for locating maxima, minima, and inflection points by finding of zero-point crossings in the data derivatives. This method employs curve-fitting of the data, which preserves peak heights and widths, instead of finding the maxima of smoothed data that can distort the shape and/or height of the peak, as with Gaussian fitting. This is exemplified later in Section 2.1.6.1. We note that while the 1D AIA data in Figure 2.7 does not appear noisy, the subsequent derivatives necessary for our edge-detection technique are increasingly noisy, as will be seen in Section 2.2. The SG smoothing filter is thus an important choice for this analysis.

SG smoothing is a convolution method that uses least-squares fitting to fit subsets of a data series with a low-degree polynomial (Savitzky and Golay, 1964). The convolution

coefficients are determined by least-squares parameters (Gorry, 1990; Press *et al.*, 1992). Depending on the coefficients, this method can be used to smooth noisy data or compute the smoothed derivatives (any order) of the unsmoothed data.

IDL’s `SAVGOL()` function returns the coefficients corresponding to designated parameters. The user can define the number of data points on either side of a point to be considered in the smoothing filter (left and right are defined separately for the option of asymmetric weighting; we use symmetric weighting and denote this value as σ , in comparison to the standard deviation of a Gaussian filter), the order \mathbb{N} of the desired derivative of some function f ($f^{\mathbb{N}}(x)$; $\mathbb{N} = 0$ for no differentiation), and the degree of the desired polynomial. It is noted in the `SAVGOL()` documentation that the coefficients returned for $\mathbb{N} > 0$ are not normalized. A normalization factor must be applied before convolving the coefficients with the data. This factor is expressed by

$$norm = \frac{!N}{dx^N}, \tag{2.1}$$

where dx is the sampling interval between the data.

2.1.6.1 Example

To initially develop and test an edge detection routine that also detects the width boundaries of edges, we created an example region with “wide” edges, as shown in Figure 2.8.

To imitate the process we are using on the AIA data, we summed the example data over the y -dimension. To start, we compare the technical differences between Gaussian and SG smoothing. Figure 2.9 shows the raw data along with the two smoothing methods applied to the raw data. The Gaussian smoothing employed has a standard deviation of ($\sigma = 5$). Similarly, the number of points used on either side of the SG smoothing was $\sigma = 5$, with a polynomial of $degree = 3$. No noise was added to the example data, so there are no other unwanted high-frequencies to smooth (besides the Gibbs phenomenon at the sharp boundary edges). Note that both smoothed curves intersect the center of the boundary

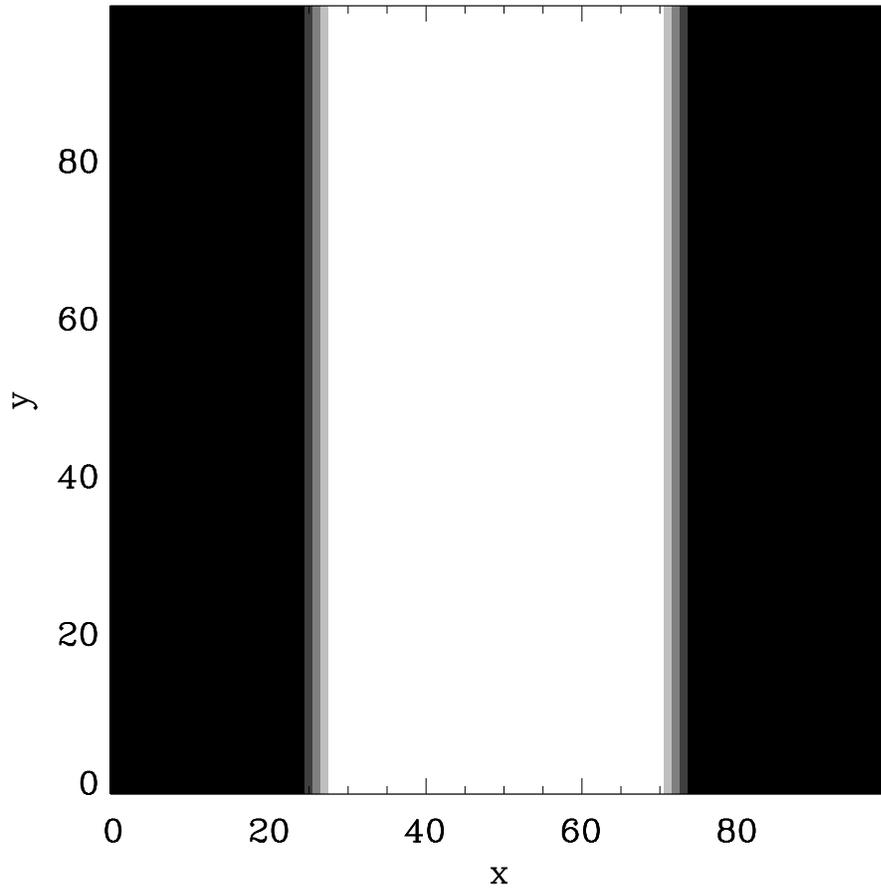


Figure 2.8: An example region with boundary edges of width $dx = 3$, centered at $x = 26$ and $x = 72$. Intensity is indicated by color (grey transitions designate boundary regions).

regions.

To better understand how the choice of parameters affect the SG smoothing, we look at how adjusting the parameters changes the smoothed curves. Figure 2.10 compares varying SG parameters with respect to the example data. A higher “sigma” (σ) broadens the smoothed width. A higher degree increases the overshoot of the smoothed width boundaries but, as mentioned, this is not highly relevant for this example. We continue with the $\sigma = 5$, *degree* = 3 case to minimize the smoothed width and for comparison with the similar Gaussian case.

For reference, we overplot boundary markers for each curve in Figure 2.11. The chosen

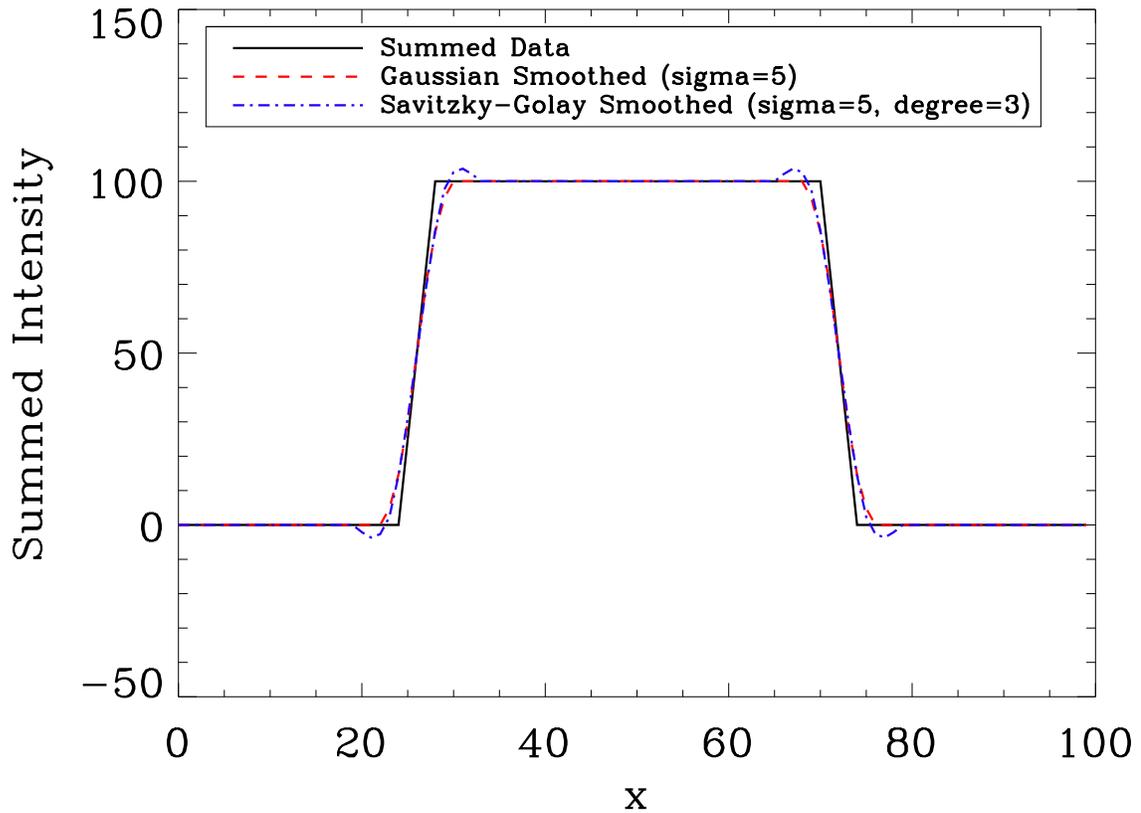


Figure 2.9: Summed data from example region in Figure 2.8 with smoothed curves from two different smoothing methods: Gaussian (red dashed) and SG (blue dash-dot). The Gaussian standard deviation is $\sigma = 5$ and the number of points used on either side of the SG fit is also $\sigma = 5$ (11 total points, including center). The polynomial degree of the SG is $degree = 3$.

boundary points are located where each curve meets the limiting data values ($min = 0$, $max = 100$).

The overplotted 1st-order derivatives of each curve, along with the 0th-order width boundaries from Figure 2.11, are shown in Figure 2.12. The signal maxima/minima are all seemingly centered between the summed data boundaries, which is desirable. The derivative curves intersect *on* the summed data boundaries, which is also anticipated. The raw data and Gaussian curves widen beyond their own 0th-order boundaries, while the SG curve

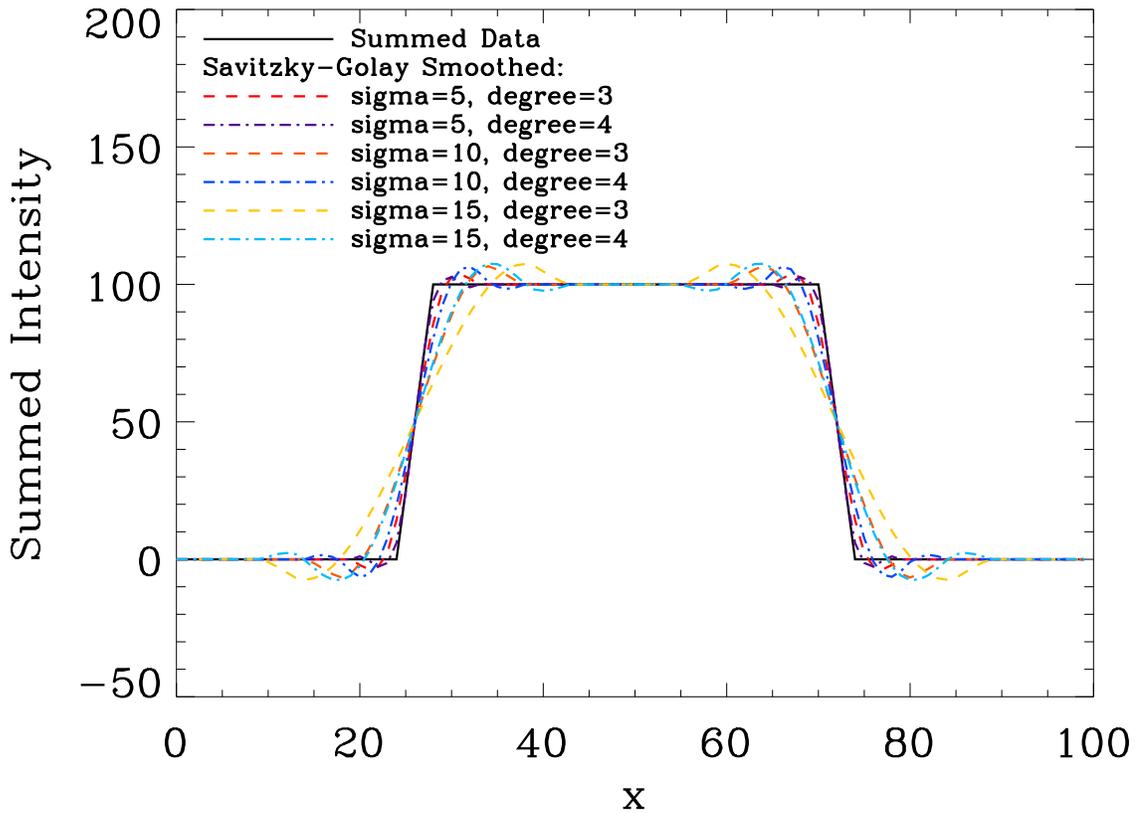


Figure 2.10: A comparison of varying SG smoothing parameters. The dashed lines correspond to $degree = 3$ polynomial fits, while dash-dotted lines show $degree = 4$ polynomial fits. Each different set of parameters is also represented by different colors.

retains its width. Also, the SG peak height is closer to raw data peak height than the Gaussian peak height. As mentioned in Section 2.1.6, the Gaussian derivative is slightly stunted and broadened as compared to the SG curve.

The overplotted 2nd-order derivatives of each curve, again with the 0th-order width boundaries, are shown in Figure 2.13. The data peaks are centered on the summed data boundaries (as expected). The smoothed peaks are widened, which is expected since the 1st-order fitted curves are wider than the data. The derivative curves all have 0-crossings centered between the raw data boundaries, which is desirable. All non-zero portions of the curves remain between the SG width boundaries.

The inflection points are determined based on the location of the 2nd-order 0-crossings,

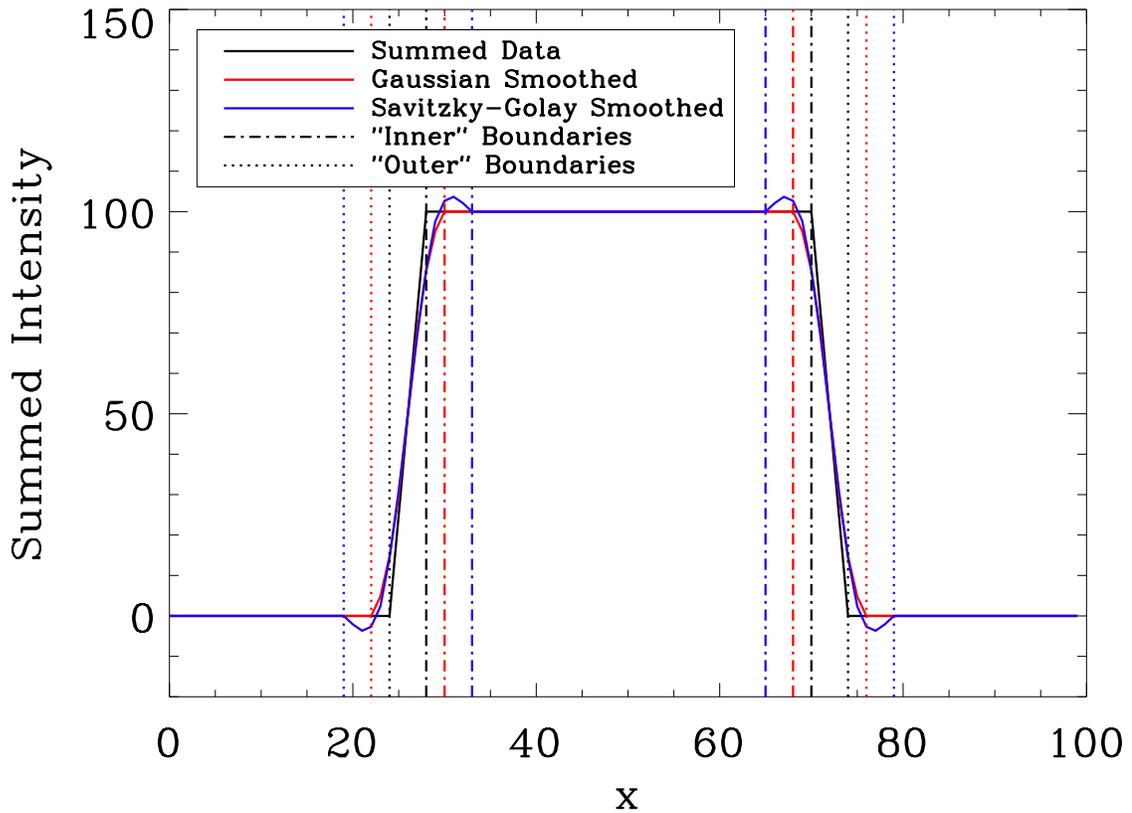


Figure 2.11: Similar to Figure 2.9, but with overplotted boundary markers for each curve. The boundary locations are chosen at the points where each curve meets the limiting data values. The inner boundaries, where the curves reach $y = 100$, are shown with dash-dotted lines. The outer boundaries, where the curves reach $y = 0$, are shown with dotted lines. The summed data curve and boundaries are shown in black. The Gaussian smoothed curve and boundaries are shown in red. The SG smoothed curve and boundaries are shown in blue.

detected near the locations of the 1st-order peak maxima and minima. Thresholding is used to locate the 1st-order peaks. The subset of code used for the detections is shown below. The Gaussian and SG smoothed curves each find the same inflection point locations, at the center of the raw data boundary, as shown in Figure 2.14.

```

threshold=10.0           ;1st-order peak threshold value
out=INTARR(nx)           ;masked output array, initialized to 0
;loop through data points
FOR ii=0,n-1 DO BEGIN
  ;test 1st-order derivative magnitude against threshold value
  IF (ABS(deriv1[ii]) GT threshold) THEN BEGIN

```

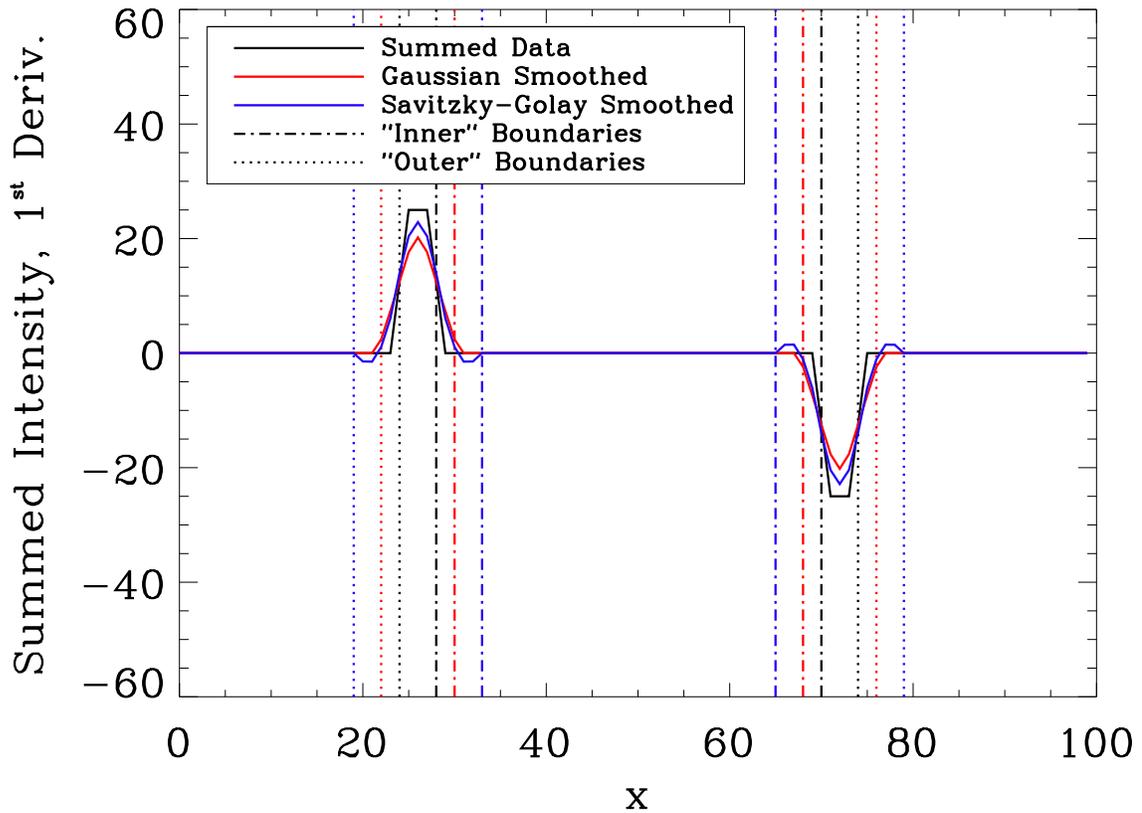


Figure 2.12: First-order derivatives of each curve, along with width boundaries from Figure 2.11. The colors and linestyles correspond to those in Figure 2.11.

```

; if above threshold, store current and prior 2nd-order derivative
values
pixel=deriv2[ii]
prevp=deriv2[ii-1]
; if 2nd-order derivative has 0-crossing, mark location in output
array
IF (pixel GT 0) AND (prevp LE 0) THEN out[ii-1]=1
IF (pixel LE 0) AND (prevp GT 0) THEN out[ii]=1
ENDIF
ENDFOR
; store count and locations of 0-crossings occurring above threshold
inflection=WHERE(out EQ 1, ni)

```

The overplotted 3rd-order derivatives of each curve with the 0th-order width boundaries are shown in Figure 2.15. The detected inflection points are also indicated on the plot. All of the 3rd-order curves appear to have 0-crossings at the raw data boundaries. The absolute maximum and minimum locations agree with the detected inflection points. All non-zero

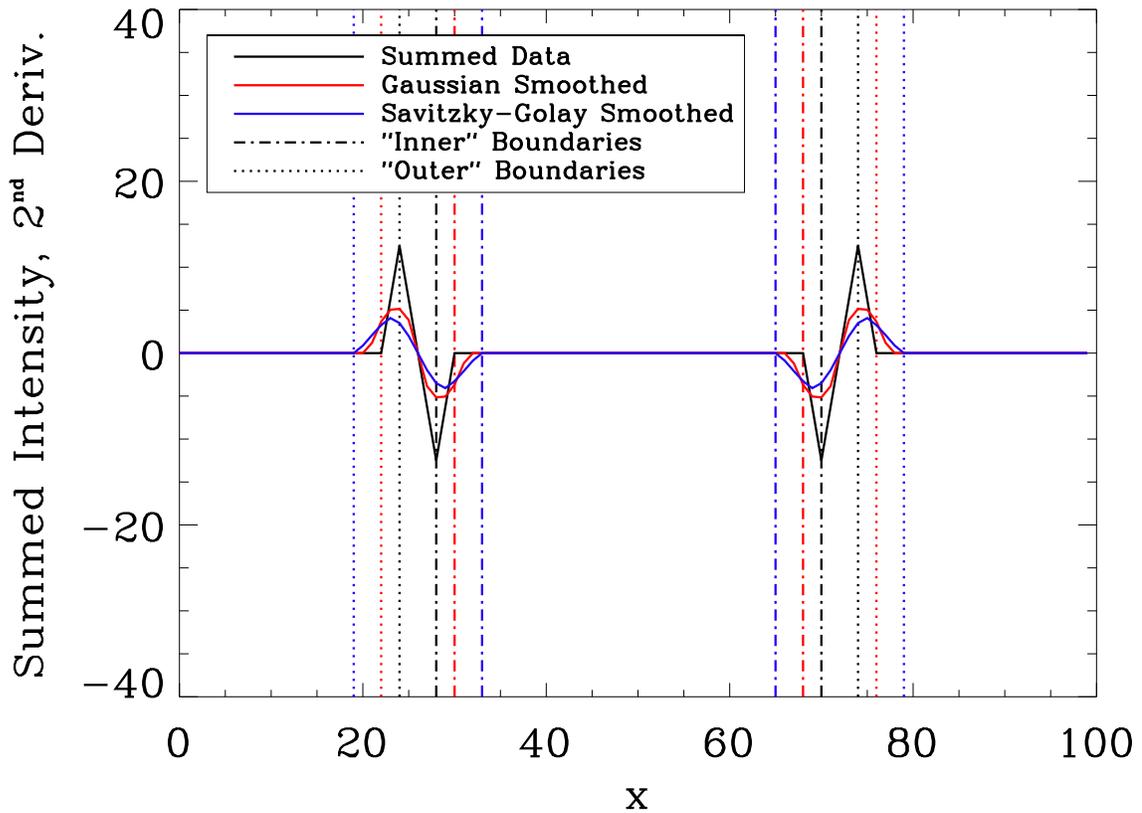


Figure 2.13: Second-order derivatives of each curve, along with width boundaries from Figure 2.11. The colors and linestyles correspond to those in Figures 2.11-2.12.

portions of the curves remain between SG width boundaries.

Similar to the inflection point detections, the width boundaries are determined based on the locations of 3rd-order 0-crossings detected near the locations of the 2nd-order peak maxima and minima. Again, thresholding was used to locate the peaks. The subset of code used for the detections is shown below. The Gaussian and SG smoothed curves find the same width boundary locations, at edges of data boundary widths, as shown in Figure 2.16.

```

threshold=1.0           ;2nd-order peak threshold value
out=INTARR(nx-1)        ;masked output array, initialized to 0
;loop through data points
FOR ii=1,n-1 DO BEGIN
    ;test 2nd-order derivative magnitude against threshold value
    IF (ABS(deriv2[ii]) GT threshold) THEN BEGIN

```

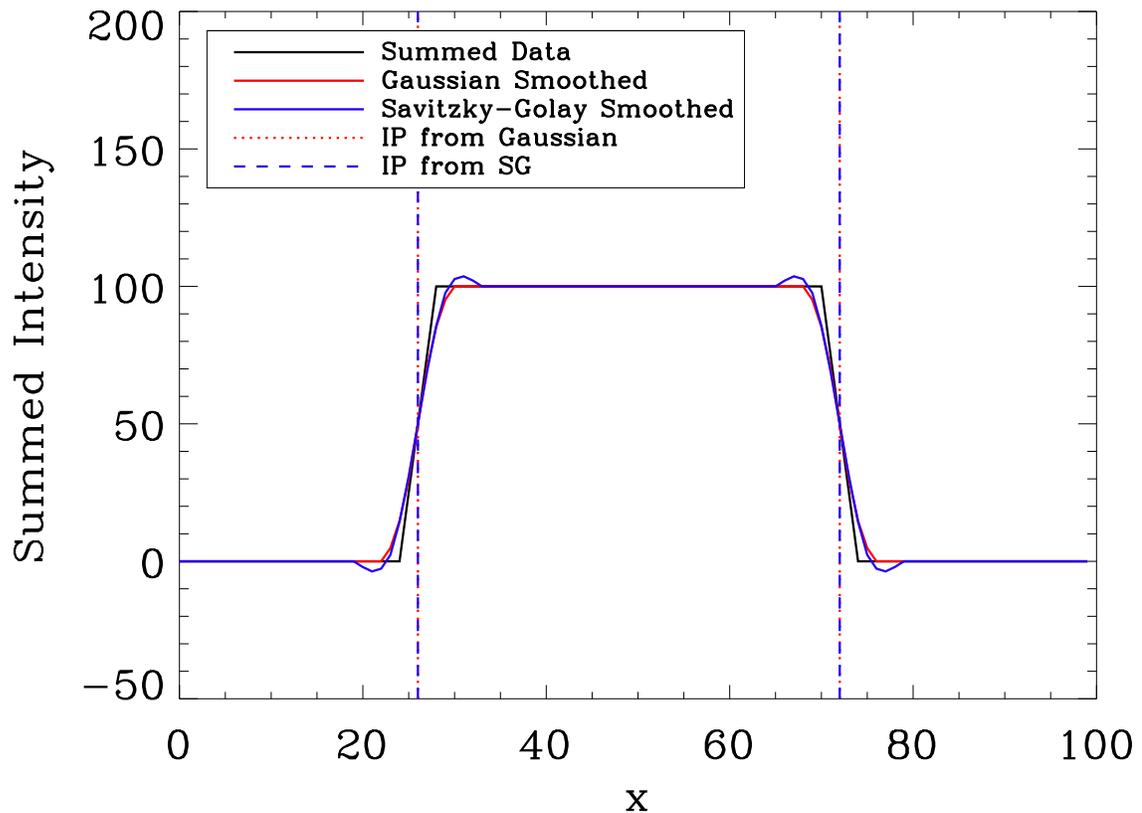


Figure 2.14: Raw and smoothed example curves with detected inflection points from both smoothed cases. Both curves produce inflection points at the same locations, which are also the centers of the width regions and where both smoothed curves cross the raw data curve. The red dotted lines represent the inflection point locations determined from the Gaussian smoothed curve. The blue dashed lines represent the inflection point locations determined from the Savitzky-Golay smoothed curve.

```

; if above threshold, store current and prior 3rd-order derivative
; values
pixel=deriv3[ii]
prevp=deriv3[ii-1]
; check sign of 1st-order derivative at current location
s=SIGN(1,deriv1[ii])
; if 3rd-order derivative has 0-crossing, mark location in output
; array, depending on corresponding 1st-order sign
IF (pixel GT 0) AND (prevp LE 0) THEN BEGIN
  IF (s GT 0) THEN out[ii-1]=1 ELSE out[ii]=1
ENDIF
IF (pixel LE 0) AND (prevp GT 0) THEN BEGIN
  IF (s GT 0) THEN out[ii]=1 ELSE out[ii-1]=1
ENDIF

```

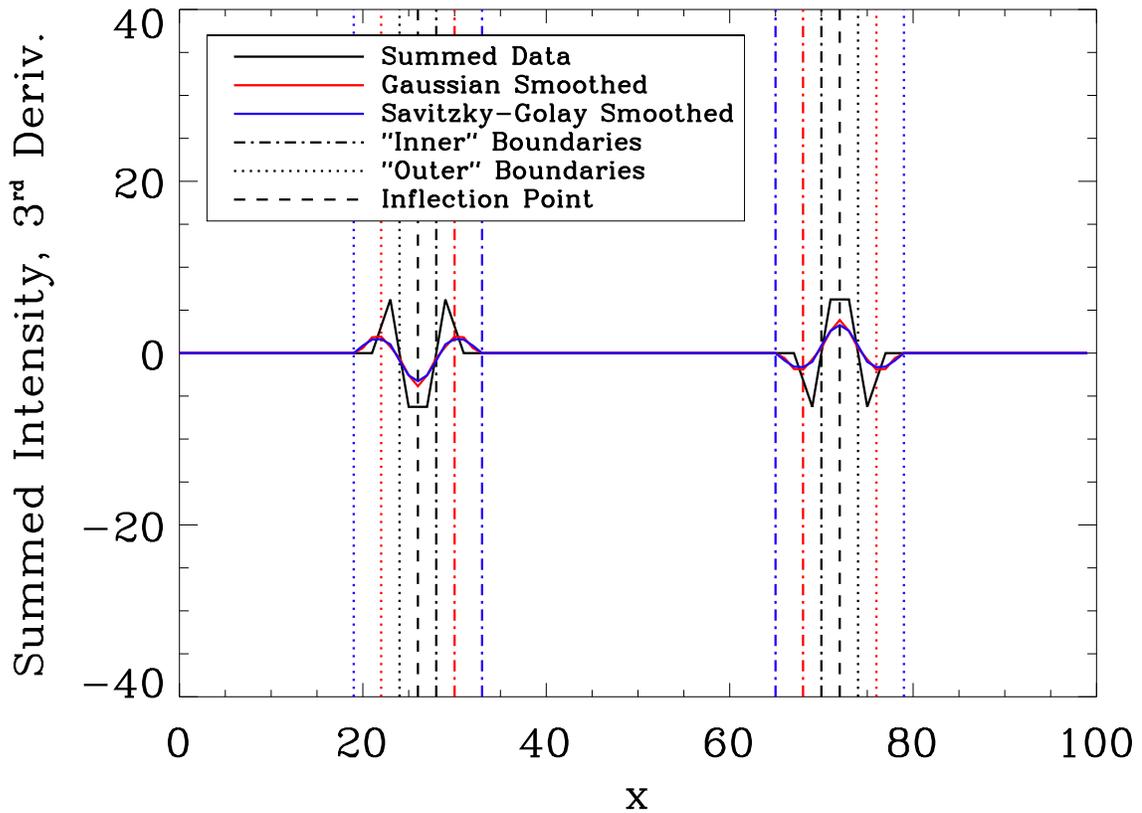


Figure 2.15: Third-order derivatives of each curve, along with width boundaries from Figure 2.11. The colors and linestyles correspond to those in Figures 2.11-2.13. The detected inflection points are represented as vertical black dashed lines.

```

ENDIF
ENDFOR
;store count and locations of 0-crossings occurring above threshold
width=WHERE(out EQ 1, nw)

```

The 2D example boundary region with detected inflection point and width boundary locations is shown in Figure 2.17. The technique developed with this simple example accurately detects the boundary center and edges. The choice of smoothing is clearly not too important here since both methods give the same edge/width results. However, as discussed in Section 2.1.6, SG is better a method for noisy data, like the AIA data derivatives (see Section 2.2). We apply the same SG method to the summed AIA data of Figure 2.7 in next section.

To briefly summarize the process outlined above, we first compute the smoothed derivatives (up to 3rd-order) of a 1D data set. The peaks of the smoothed 1st-order derivative correspond to the inflection points of the original data. We isolate the local region around a 1st-order maximum/minimum and then determine the precise location of the peak (inflection point), which corresponds to the 0-crossing point in the 2nd-order derivative. Similarly, the inflection points on either side of the 1st derivative peak are located by isolating the maxima/minima of the corresponding 2nd-order curve and the related 3rd-order 0-crossings. These secondary inflection points constitute a bounded interval within which the central peak exists.

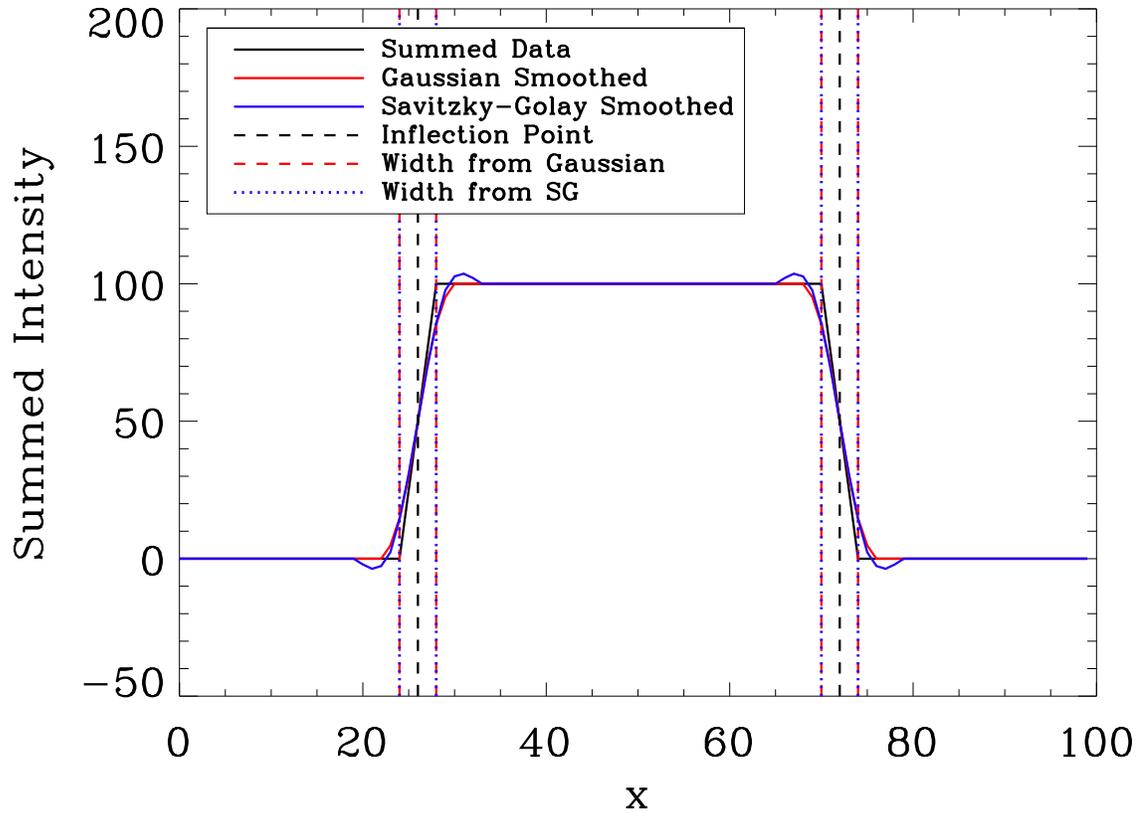


Figure 2.16: Raw and smoothed example curves with detected inflection points and width boundaries from both smoothed cases. Both curves produce width boundaries at the same locations, which also correspond to the boundary edges from the summed data curve. The black dashed lines represent the detected inflection points. The red dashed lines represent the width boundary locations determined from the Gaussian smoothed curve. The blue dotted lines represent the width boundary locations determined from the Savitzky-Golay smoothed curve.

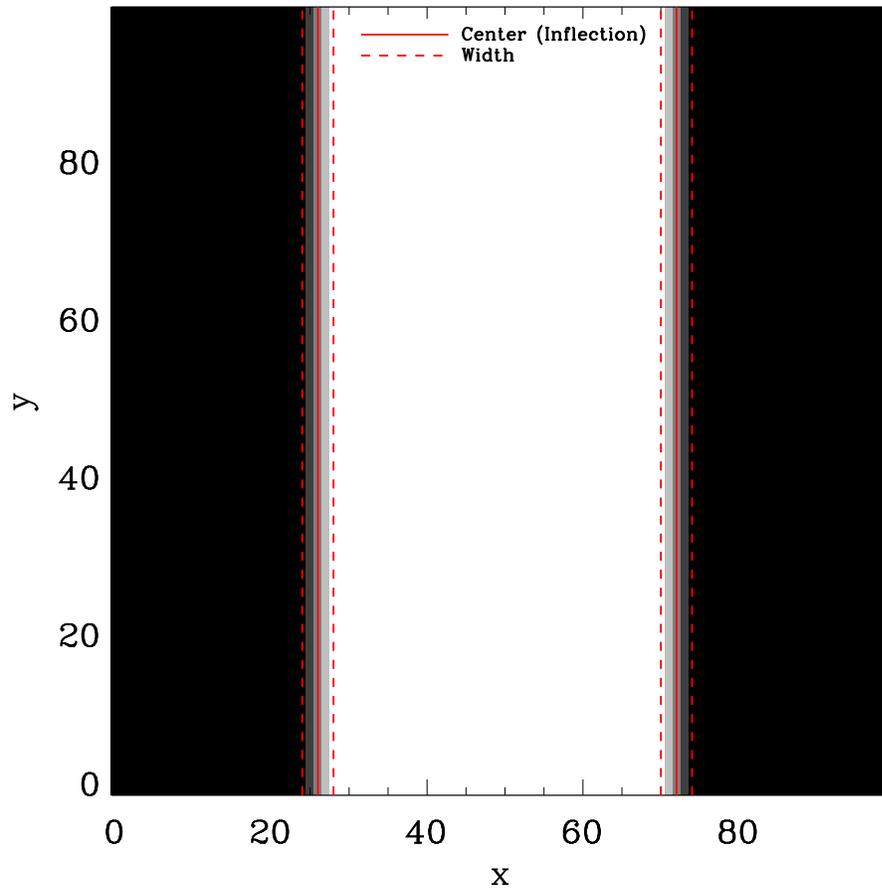


Figure 2.17: The 2D example boundary region with the detected inflection points (solid red lines) and detected boundary widths (dashed red lines). Intensity is indicated by the background color.

2.2 Results

Recall the image-processing procedure so far. Having chosen an appropriate elongated-CH data set, we rotated the multi-wavelength images such that the CH boundary is parallel to y -axis, removing the rotational effects around edges. We then summed the data along the y -dimension (axis-of-symmetry), and similarly summed the reduced data through time. Now we need to locate the center and width of the CH boundary.

Using the edge detection technique from Section 2.1.6.1, we want to determine the CH boundary's center and edges from the AIA data. These locations will be used later in the HMI data analysis. Since the AIA data has high-frequency noise, we use SG smoothing, as discussed in Section 2.1.6. To determine the best possible set of parameters for smoothing and the boundary detection algorithm, we first try various combinations of parameters and compare their detection results. Tables 2.1-2.3 show and discuss the different sets of parameters and their results when applied to the data.

For all of the calculations henceforth, we set parameter $dx = 1.0$ since the sample spacing of the data is 1 pixel. (We note that we tried setting this value to the physical data spacing in Mm, $dx = \left(0.0359 \frac{\text{deg}}{\text{pix}}\right) \left(12.15 \frac{\text{Mm}}{\text{deg}}\right) = 0.436$, but the resulting smoothed curve fit the data poorly.) It is also apparent from Table 2.1 that the Mm spacing consistently detects multiple inflection points as well. Table 2.1 also shows that $degree = 4$ is generally more sensitive to nw and incidentally, it also produces a curve that better fits the data than $degree = 3$. Therefore going forward, we select $degree = 4$ to improve the sensitivity of the width detection.

We investigate the result dependence on threshold variability in Table 2.2. Recall from Section 2.1.6.1 that the inflection point is dependent on the $t1$ parameter, while the width detections are dependent on $t2$.

For the sake of having some definable criteria moving forward, we will choose the σ and thresholding parameters that give results that best fit the following rules:

1. Only one inflection point is detected,

Table 2.1: Savitzky-Golay data results to help find “the sweet spot” for smoothing and determining an error estimate in edge location and width. Code uses IDL’s `SAVGOL()` function where $\sigma = nleft = nright$ is the integer specifying the number of data points on either side of each point to include in the curve-fit, $degree$ specifies the degree of smoothing of the polynomial, and dx is the sampling interval between data points. The variables $t1$ and $t2$ are the threshold values used to determine the inflection point and width locations, respectively. The inflection point location and width boundaries are detected using the specified parameters. The number of width detections is recorded as nw . Note that the widths declared here are those closest to the inflection point on either side. All locations are in AIA pixel units.

| σ | $degree$ | dx | $t1$ | $t2$ | inflection point | left width | right width | nw |
|----------|----------|-------|------|------|------------------|------------|-------------|------|
| 10 | 3 | 1.0 | 1E6 | 2E4 | 370 | 366 | 377 | 39 |
| | | 0.436 | | | 370^{*4} | 366 | 377 | 64 |
| 10 | 4 | 1.0 | 1E6 | 2E4 | 371 | 366 | 377 | 49 |
| | | 0.436 | | | 371^{*6} | 366 | 377 | 62 |
| 15 | 3 | 1.0 | 1E6 | 2E4 | 368 | 362 | 378 | 22 |
| | | 0.436 | | | 368^{*3} | 362 | 378 | 45 |
| 15 | 4 | 1.0 | 1E6 | 2E4 | 371 | 362 | 378 | 33 |
| | | 0.436 | | | 371^{*3} | 362 | 378 | 46 |
| 20 | 3 | 1.0 | 1E6 | 2E4 | 367 | 322 | 379 | 13 |
| | | 0.436 | | | 367^{*2} | 359 | 379 | 31 |
| 20 | 4 | 1.0 | 1E6 | 2E4 | 369 | 359 | 379 | 25 |
| | | 0.436 | | | 369^{*2} | 359 | 379 | 38 |

^{*n} Routine detected $n > 1$ inflection point edges.

2. The detected width is symmetric about the detected inflection point,
3. The detected inflection point is the also the approximate average inflection point from all tested cases,
4. If Criteria 1-3 are satisfied for multiple sets of parameters, choose the case with the widest width (it is better to overestimate the width than underestimate it).

From Table 2.2, it is apparent that $\sigma = 20, 25$ best conform to our stated criteria. We select the thresholding parameters $t1 = 1E6$ and $t2 = 2E4$. We note that our choice of the $t1$ parameter is somewhat arbitrary between the two mid-values because they both produce the same inflection point within each case of σ . We discard the lowest and highest values because they either detect too many inflection points or not enough (Criterion 1).

Table 2.2: Savitzky-Golay data results to help find “the sweet spot” for smoothing and determining an error estimate in edge location and width. Code uses IDL’s `SAVGOL()` function where $\sigma = nleft = nright$ is the integer specifying the number of data points on either side of each point to include in the curve-fit. We choose polynomial $degree = 4$, and sampling interval $dx = 1.0$ pixel. The variables $t1$ and $t2$ are the threshold values used to determine the inflection point and width locations, respectively. The inflection point location and width boundaries are detected using the specified parameters. The number of width detections is recorded as nw . Note that the widths declared here are those closest to the inflection point on either side. All locations are in AIA pixel units.

| σ | $t1$ | $t2$ | inflection point | left width | right width | nw |
|----------|-------|-------|-------------------|------------|-------------|------|
| 10 | 8E5 | 1.5E4 | 371 ^{*4} | 366 | 377 | 50 |
| | 1E6 | 2E4 | 371 | 366 | 377 | 49 |
| | 1.2E6 | 2.5E4 | 371 | 366 | 377 | 45 |
| | 1.4E6 | 3E4 | 371 | 366 | 377 | 37 |
| 15 | 8E5 | 1.5E4 | 371 ^{*2} | 362 | 378 | 38 |
| | 1E6 | 2E4 | 371 | 362 | 378 | 33 |
| | 1.2E6 | 2.5E4 | 371 | 362 | 378 | 28 |
| | 1.4E6 | 3E4 | — | 362 | 378 | 26 |
| 20 | 8E5 | 1.5E4 | 369 | 359 | 379 | 32 |
| | 1E6 | 2E4 | 369 | 359 | 379 | 25 |
| | 1.2E6 | 2.5E4 | 369 | 359 | 379 | 22 |
| | 1.4E6 | 3E4 | — | 359 | 379 | 19 |
| 25 | 8E5 | 1.5E4 | 368 | 355 | 381 | 24 |
| | 1E6 | 2E4 | 368 | 355 | 381 | 22 |
| | 1.2E6 | 2.5E4 | 368 | 355 | 381 | 19 |
| | 1.4E6 | 3E4 | — | 31 | 381 | 11 |

^{*n} Routine detected $n > 1$ inflection point edges.

Similarly, our choice of $t2$ is somewhat arbitrary between the lowest three values tested since all produce the same boundary limits within each case of $sigma$. We discard the higher value because it returns an inconsistent lower boundary limit result with higher σ .

To narrow down the “sweet spot”, we look in finer detail at values between $15 < \sigma \leq 30$ in Table 2.3. The average inflection point is approximately between $x = 368$ and 369 . The width is not as clear cut and depends mainly on the number of points considered in the fit. All the cases of σ in Table 2.3 fulfill Criterion 1. Cases $\sigma = 19, 21, 22, 23, 26, 27$ are ruled out by Criterion 2; cases $\sigma = 16, 17, 18, 28, 29, 30$ are ruled out by Criterion 3. We choose the $\sigma = 25$ case as our ultimate parameter selection because the symmetric boundary is wider than in the $\sigma = 20, 24$ cases (Criterion 4).

Table 2.3: Savitzky-Golay data results to help find “the sweet spot” for smoothing and determining an error estimate in edge location and width. Code uses IDL’s `SAVGOL()` function where $\sigma = nleft = nright$ is the integer specifying the number of data points on either side of each point to include in the curve-fit. We choose polynomial $degree = 4$, and sampling interval $dx = 1.0$ pixel. The threshold variables that determine the inflection point and width locations, respectively, are taken to be $t1 = 1E6$ and $t2 = 2E4$. The inflection point location and width boundaries are detected using the specified parameters. The number of width detections is recorded as nw . Note that the widths declared here are those closest to the inflection point on either side. Bold red values indicate where one of the listed criteria are not met, therefore ruling out that row of parameters. All locations are in AIA pixel units.

| σ | inflection point | left width | right width | nw |
|----------|------------------|------------|-------------|------|
| 16 | 370 | 361 | 378 | 29 |
| 17 | 370 | 361 | 379 | 30 |
| 18 | 370 | 360 | 379 | 29 |
| 19 | 369 | 360 | 379 | 29 |
| 20 | 369 | 359 | 379 | 25 |
| 21 | 369 | 358 | 379 | 23 |
| 22 | 369 | 357 | 379 | 22 |
| 23 | 368 | 357 | 380 | 21 |
| 24 | 368 | 356 | 380 | 21 |
| 25 | 368 | 355 | 381 | 22 |
| 26 | 368 | 354 | 381 | 20 |
| 27 | 368 | 353 | 382 | 19 |
| 28 | 367 | 352 | 382 | 18 |
| 29 | 367 | 351 | 383 | 15 |
| 30 | 367 | 350 | 384 | 15 |

* n Routine detected $n > 1$ inflection point edges.

The AIA data plots produced using the final parameter selection are shown in Figures 2.18-2.21. Each plot shows the detected inflection point and all detected width boundaries. Note again that only the width boundaries closest to either side of the inflection point are taken as the boundaries of interest. Figure 2.18 shows the data from Figure 2.7, plotted with the 0th-order SG-smoothed curve. The smoothed curve is a good representation of the data. The higher-order SG-smoothed curves are increasingly useful as the raw data derivatives become progressively noisier. The inflection line intersects the peak of the 1st-order SG derivative in Figure 2.19, the apparent 0-crossing location of the 2nd-order SG derivative

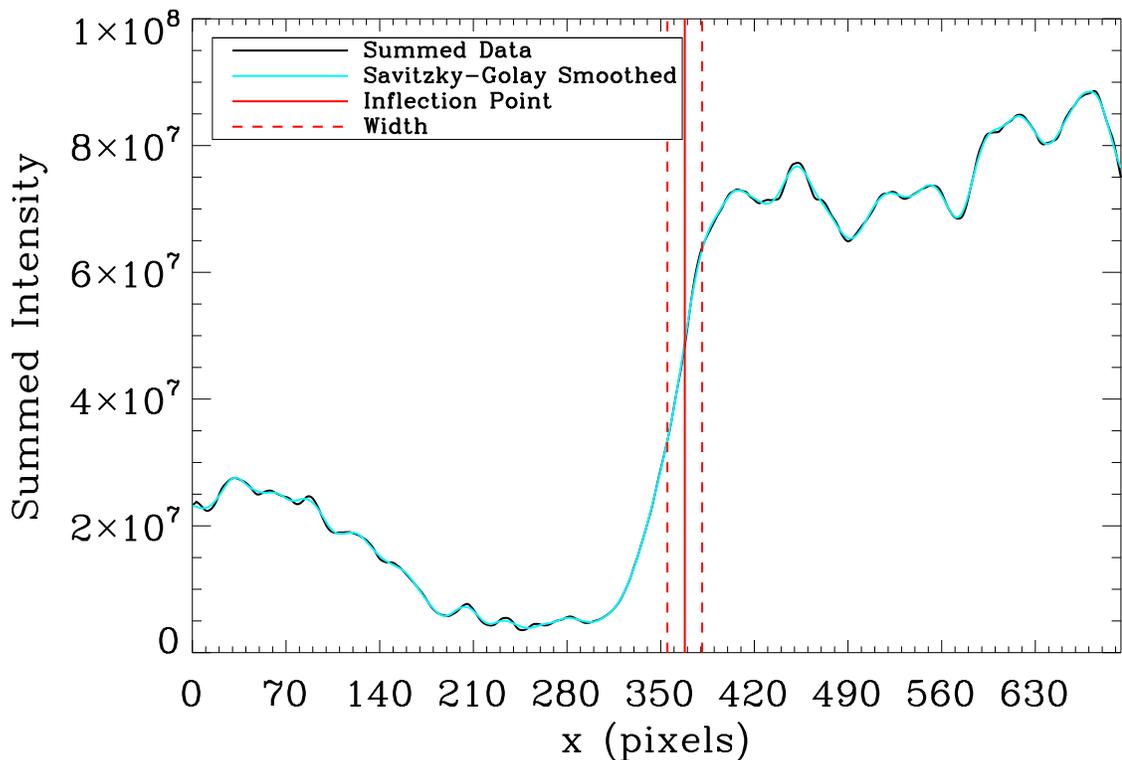


Figure 2.18: 0th-order summed AIA data (black) with the overplotted SG smoothed curve (cyan). The detected inflection point (solid red) and detected boundary widths (red dashed) are also shown. The smoothing parameters are $\sigma = 25$ and *degree* = 4. Recall from Section 2.1.1, 1 AIA cutout pixel = $0.0359^\circ \simeq 0.436$ Mm. The summed intensity has arbitrary units.

in Figure 2.20, and the local minimum of the 3rd-order SG derivative in Figure 2.21. The width lines (including those that we will neglect) intersect the local minima and maxima of the 2nd-order SG derivative in Figure 2.20, and the apparent 0-crossing locations of the 3rd-order SG derivative in Figure 2.21.

In summary, we will use a central edge location of $x = 368$, a lower width boundary location of $x = 355$, and an upper width boundary of $x = 381$. These pixel positions are converted into Heliographic Carrington coordinates using the SSW WCS routines in IDL, and then mapped onto the corresponding HMI data. This process will be discussed later when we consider the HMI analysis in Chapter 5; however, it involves WCS routines outlined in Appendix A.1.

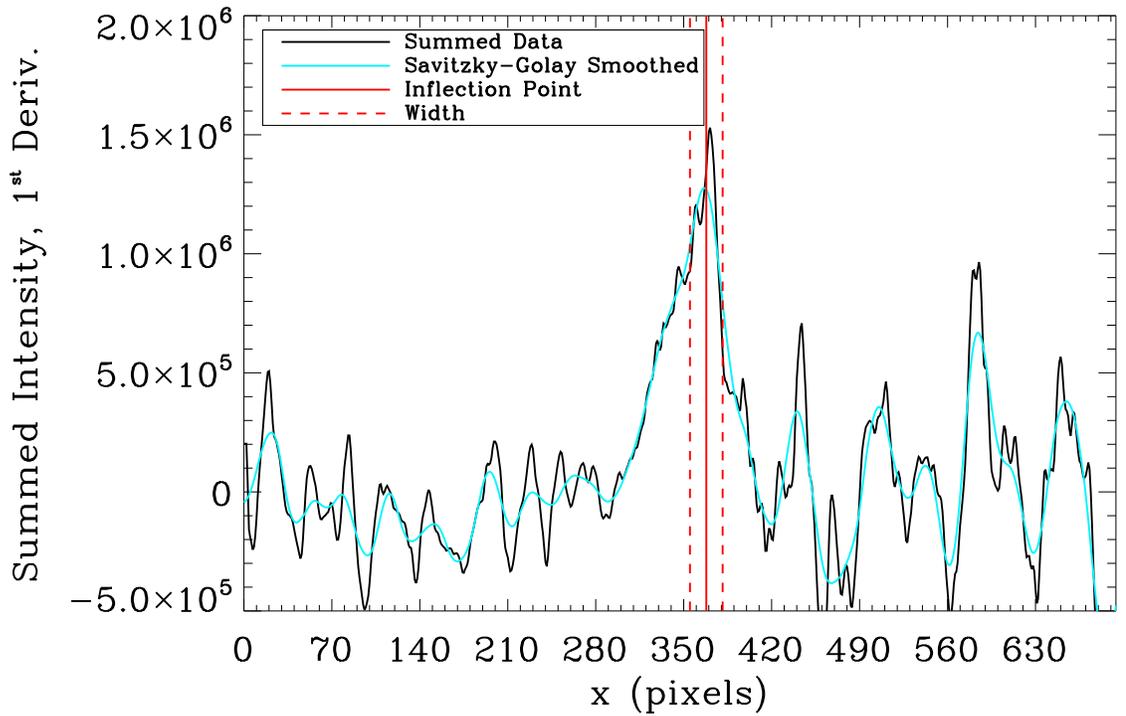


Figure 2.19: 1st-order summed AIA data (black) with the overplotted SG smoothed curve (cyan). The detected inflection point (solid red) and detected boundary widths (red dashed) are also shown. The smoothing parameters are $\sigma = 25$ and *degree* = 4. Recall from Section 2.1.1, 1 AIA cutout pixel = $0.0359^\circ \simeq 0.436$ Mm. The summed intensity has arbitrary units.

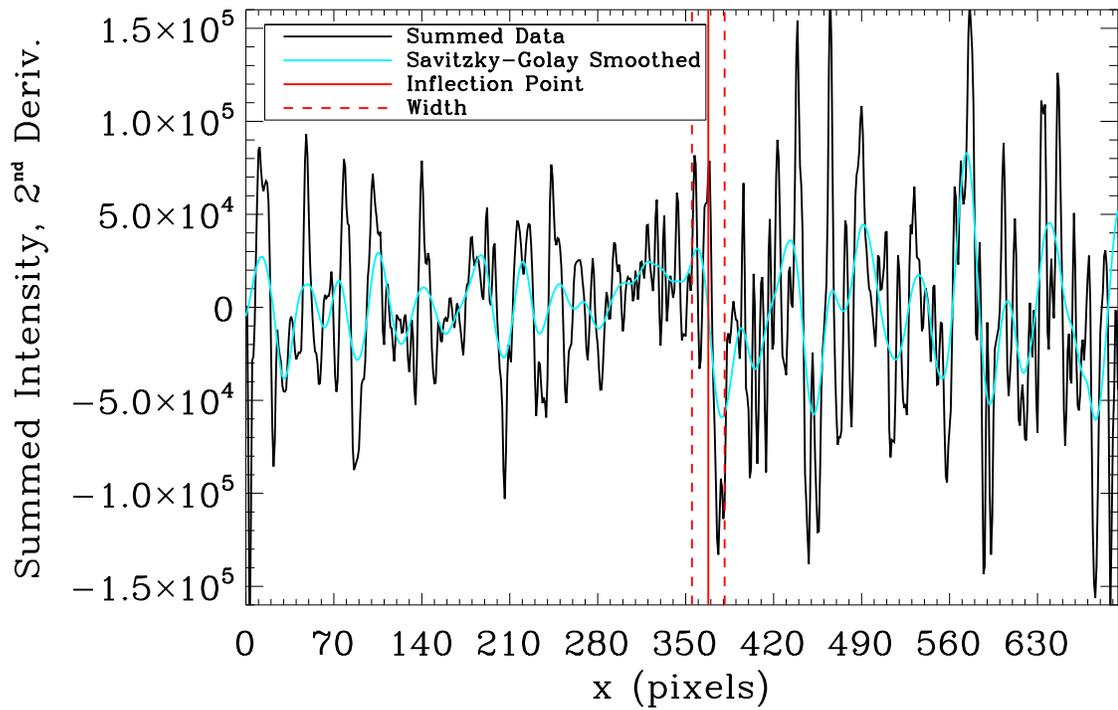


Figure 2.20: 2nd-order summed AIA data (black) with the overplotted SG smoothed curve (cyan). The detected inflection point (solid red) and detected boundary widths (red dashed) are also shown. The smoothing parameters are $\sigma = 25$ and *degree* = 4. Recall from Section 2.1.1, 1 AIA cutout pixel = $0.0359^\circ \simeq 0.436$ Mm. The summed intensity has arbitrary units.

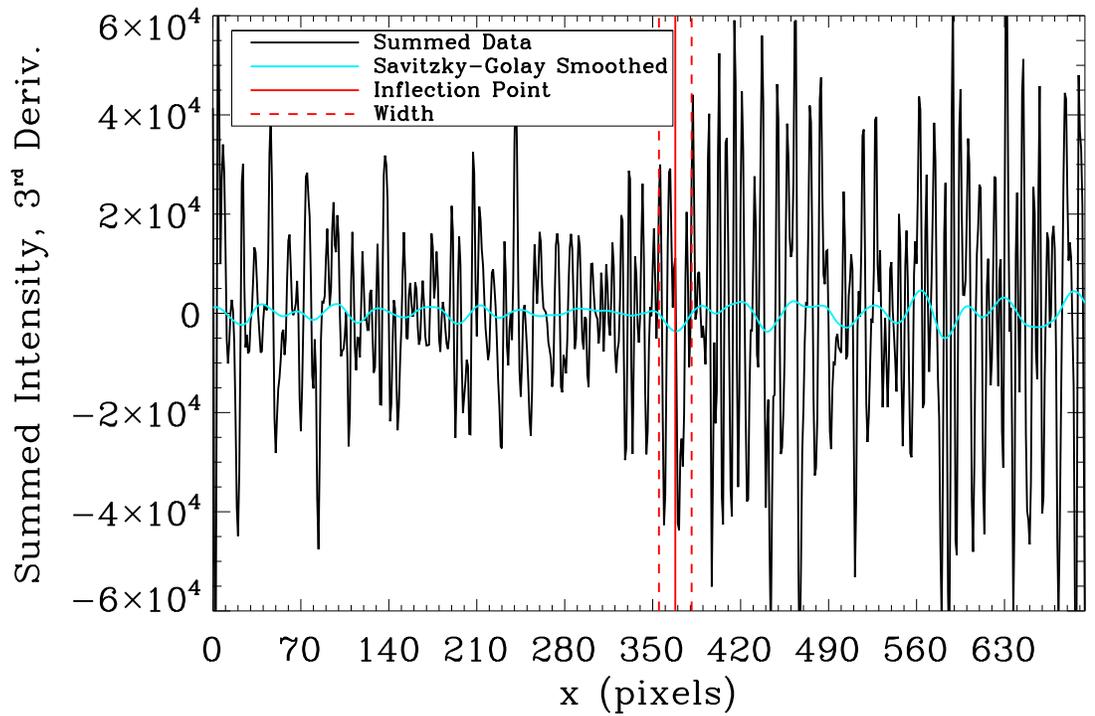


Figure 2.21: 3rd-order summed AIA data (black) with the overplotted SG smoothed curve (cyan). The detected inflection point (solid red) and detected boundary widths (red dashed) are also shown. The smoothing parameters are $\sigma = 25$ and *degree* = 4. Recall from Section 2.1.1, 1 AIA cutout pixel = $0.0359^\circ \simeq 0.436$ Mm. The summed intensity has arbitrary units.

2.3 Previous Work on Coronal Holes

Our previous research similarly studied CHs in EUV wavelengths, using SOHO/EIT data. Here we summarize the method and results that were published in Kirk *et al.* (2009) and Hess Webber *et al.* (2014). As mentioned in Chapter 1, the evolution of CHs through the solar cycle is important in terms of understanding the solar dynamo. We specifically tracked the boundaries of PCHs using a new feature tracking technique, dubbed perimeter tracing, that measures the polar hole perimeters by detecting the limb boundaries throughout a polar rotation. We then used the traced perimeter to estimate the PCH areas through time. Our goal was to determine whether the polar holes could be used as proxies for the open polar magnetic field, which then might be useful for predicting the level of solar activity in the subsequent solar cycle. In this section, we provide a brief review of this prior work and its results. We hereafter refer to Kirk *et al.* (2009) as Paper I and Hess Webber *et al.* (2014) as Paper II.

Paper I is mainly devoted to the details of the perimeter tracing method, although some initial results are also discussed. Paper II presents a few minor adjustments to the method in Paper I, but is mostly dedicated to the results of an expanded data set, as well as the development and results from a second supplemental PCH area measurement method. In both papers, we use data from three EIT wavelengths for the analysis: 304 Å, 171 Å, and 195 Å. This set of wavelengths allows us to explore the PCH boundaries at varying scale heights/temperatures, the highest of which is between ~ 60 Mm and ~ 190 Mm above the photosphere (Fe XII 195 Å); the lowest, between ~ 3 Mm and ~ 60 Mm (He II 304 Å). There are benefits and disadvantages to using any of these wavelengths. Small scale heights obscure the polar limb less than large scale heights (Figure 2 in Paper I); however, the contrast is similar between the CHs and general corrugations in the chromospheric line, which makes the boundary detections more difficult than in the high contrast coronal lines (Figure 3 in Paper I). Using the collection of all three wavelengths allows us to compare and confirm the results of any one data set. Paper II also uses synoptic-map data from both

SOHO/EIT (same three wavelengths) and SOHO/Michelson Doppler Imager (MDI).

As stated above, we detect the PCH limb boundaries through a rotation. However, the differential rotation of the Sun means that using the Carrington rotation rate at the poles would be inaccurate. Therefore, it is necessary to first define a standard rotation rate for the polar holes. We determined that the approximate rate of rotation was 33 days, and labeled this as the Harvey Rotation rate (in honor of the late Karen Harvey who was a pioneer in CH research). We also defined a Harvey Rotation number (HR) as

$$HR = \frac{JD - 2415024.5}{33}, \quad (2.2)$$

where JD is the Julian day, and HR 0 began on Jan 4, 1900.

The detection method begins with some image processing of the data. First, it is calibrated using the SSW routine *eit_prep.pro*. The images are then spatially smoothed using a combination of morphological image analysis (MIA) functions (see Paper I, Section 3.2 for more details). The effect is to (1) smooth out regions that are too small to be part of the PCH, and (2) group small, closely-spaced regions of similar intensity into the larger PCH region. These manipulations are important for removing bright points occurring within the holes, which can potentially divert the boundary detections. They are also helpful in smoothing out the corrugation and intermittent spicule structures in the 304 Å data.

Choosing intensity thresholds for the various wavelengths is the next step. Because the Sun's intensity fluctuates throughout the solar cycle, explicit intensity values for each wavelength are unreliable since the maximum intensity at one time in the cycle could be the minimum intensity at another time. Instead, we use normalized histograms of intensity for each smoothed image. Distinct intensity threshold levels are chosen for each wavelength. For example, for 171 Å images, the threshold might be the intensity bin with a value of 0.70; in other words, where 30% of the pixels in the image are darker than that bin. The threshold level for 304 Å images is lower because of their lower contrast, mentioned above. The particular threshold values are chosen so as to best reflect the visible PCH limb

boundaries throughout the solar cycle. (The threshold levels differ slightly between Paper I and Paper II, in a quest for improved detection accuracy.) We then create binary images using the intensity thresholds.

From the binary data, we isolate the solar limb by masking most of the disk and corona, so that only an annulus containing the limb remains. The annulus has a width of about 6% of the solar radius. We also mask up to latitudes of $\pm 60^\circ$ (changed to $\pm 50^\circ$ in Paper II) to isolate the polar regions. From here, the PCH boundaries are apparent around the limb. See Figure 4 in Paper I for an illustration of the steps in our detection process.

After the PCH boundaries are detected, their locations are converted into heliographic coordinates and subsequently to HR coordinates (where the central meridian of the disk is taken to be 0° longitude at the start of each rotation). We can then plot the locations on a polar map for each HR (Figure 5 in Paper I).

We also compute the “quality” of the measurements. This value is the ratio between the number of pixels within the binary limb that qualify as part of the PCH (where the binary pixels are 0) and the total number of pixels between the detected limb boundaries. If the PCH is poorly defined, the quality will be low. Any measurement with a quality under 15% is discarded to help guard against false positives.

In Paper II, we incorporate additional false positive filters. The detection-confidence (DC) factor is very similar to the quality calculation. Having already removed detections with low measurement quality, we take the ratio of the total number of PCH pixels around the limb with the total number of points above the limiting polar latitude. This is a measure of the likelihood that a detection confines a PCH. If the boundary detections are close together, then there are few potential PCH pixels in that pole and the probability that a PCH actually exists between the boundaries is low, prompting a low DC factor. DC does not indicate anything about the accuracy of the detections, but solar cycle trends are apparent in the DC factor. Particularly, it is significantly lower during maximum than through the rest of the cycle, when it ranges between about 0.4-1.0. This result lead us to implement a filter that weights the boundary fit of the detected points based on their DC factors, and to

disregard points with DC factors less than 0.4. We also reduce false positive detections do to absorption features, such as filaments, that pass through the other filters. They appear on the polar plots as smooth curves with consistently close-set latitudinal coordinates. The filter implemented to remove the absorption features exploits this smooth, clustered curvature by setting a lower-limit on the difference between the latitudinal coordinates of consecutively detected points. This filter is effective because the latitudes of detected PCH boundaries fluctuate more over time than do filaments rotating across the solar limb. Again, any data that did not meet the threshold criteria were discarded.

To calculate an area from the perimeter results, we first combine the east and west limb detections from each PCH, and then fit the combined data from one HR with a line of best fit. Combining the limb measurements increases the sample size for the fitting. This step is somewhat trivial since we are working in our defined HR coordinate system, so the coordinates for either edge of a PCH can simply be concatenated into a single array. We also apply another filter to guard against under-sampling. The filter throws out any rotation where more than half of the measured points are missing (due to the quality, DC, or absorption filters).

The array of detected points is now translated such that its “center-of-mass” is the origin of the coordinate grid. Doing so helps prevent unrealistic fits of the data. The measurements are then fitted using a non-linear least squares fitting routine with a Fourier cosine series in longitude

$$f(\phi) = \aleph_0 + \sum_{n=0}^{N_{\max}} \aleph_{2n-1} \cos(n\phi + \aleph_{2n}), \quad (2.3)$$

where \aleph are the Fourier coefficients and N_{\max} varies between two and seven to generate multiple candidate fits. Candidate fits must conform to certain criteria in order to be accepted as reasonable representations of the PCH, e.g., the fitted line cannot extend below 45° latitude. Again, for more details see Paper I. For all the remaining fits, we calculate

the area within the fitted curves by integrating over the Fourier cosine function. We then take an average of each of those areas, and accept it as the approximate area of the PCH for the HR.

Confidence measurements of the fits are determined by estimating the uncertainties of the fit parameters. We calculate the variances of each parameter from their squared partial derivatives of Equation 2.3. These variances allow us to resolve a one-sigma confidence range for each fit parameter. We then add/subtract one sigma to/from each of the parameters and use the subsequent fits to calculate a maximum/minimum area limits, using the same method described above, and again take an average of max and min areas to produce a one-sigma range about the line-of-best-fit average area.

The results of Paper I study the PCHs from the beginning of the SOHO/EIT data, up through mid-2007, spanning solar cycle 23. As expected, both PCHs show a decline into solar maximum and restoration into solar minimum. The northern PCH disappears at the end of 1999, coming back late in 2000. The southern PCH is delayed, disappearing in mid-2000 and returning around the start of 2002. The asymmetry between the northern and southern PCHs is consistent with the now widely studied hemispheric asymmetry prevailing through solar cycle 24.

The area results are also compared to those of Harvey and Recely (2002), who studied the PCH area between the middle of solar cycle 22 and the middle of solar cycle 23, using He I 10830 Å. Our results agree with theirs in terms of the time-evolution of the holes through the solar cycle. The northern hole disappears in mid-1999 and reemerges in early 2001; the southern hole disappears in mid-2000 but their study ends prior to its reappearance. The results differ in the measured magnitudes of the PCH areas. After recovering from the solar cycle 23 maximum, Harvey and Recely measure the northern PCH area to be $\sim 37\%$ larger than our measurement. Around the minimum between solar cycles 22 and 23, they measure a northern PCH to be almost twice as large as our measurement.

The Paper I areas are also compared to measurements of the polar magnetic field from the Wilcox Solar Observatory (WSO). The unsigned polar magnetic field was about 40%

weaker at the end of our study (near solar minimum) than it was in the previous cycle minimum. Because CHs are representative of regions with concentrated open magnetic field lines, one might reason that the polar holes should reflect the decrease in the polar magnetic field strength between solar minima. However, our observations show that the total PCH area in the cycle 23/24 minimum was only $\sim 15\%$ smaller than in the previous minimum. We conclude that this inconsistency suggests the PCHs are influenced more by dynamics, such as flux transport and meridional flow, than magnetic field strength.

We now discuss the results of Paper II. Recall that we use the perimeter tracing method, along with an EIT synoptic method and an MDI synoptic method. We will not go into detail about the two new synoptic methods, but encourage readers to explore those procedures in the paper. However, the results presented here are supported by all three methods, and this research has been extended using the synoptic methods on SDO/AIA and HMI data by Karna, Hess Webber, and Pesnell (2014).

Paper II expands the time-series of Paper I through the end of 2010, at which point the available SOHO data becomes too few per day to accurately measure PCH areas. Before proceeding, we note that we are explicitly not comparing PCH area values between the widely accepted times of the solar cycle minima derived from the monthly averaged sunspot numbers. Instead, we make the more fundamental comparisons of 1) the maximum PCH areas in either hemisphere between solar cycle minima, and 2) the temporal displacement of the area peaks between hemispheres. The SOHO data set does not cover the entire cycle 22/23 minimum. As a consequence of this, we cannot compare the shift in hemispheric area peaks between solar cycles, and are forced to accept the beginning of the SOHO data series as the largest PCH areas of the cycle 22/23 minimum.

The first main result of Paper II is that, for both hemispheres during both solar minima, the maximum PCH areas cover about 6% of the total solar surface. The PCH areas do not vary significantly between the observed solar cycle minima. Again, these results are confirmed by all three feature-tracking methods. They also reconfirm our conclusion from Paper I that the evolution of the PCH areas does not seem to be a good predictor of the

amount of solar cycle activity expected in the following solar maximum.

We also find that the PCHs in the northern and southern hemispheres do not reach their maximum areas at the same time, the northern hole achieving its peak almost 4 years before the southern hole (again, this comparison can only be made for the cycle 23/24 minimum). This suggests that the two poles reach solar-minimum conditions at distinct times. Similarly, the northern hole began declining toward solar-maximum conditions in 2010, while the southern hole had not begun its descent to solar maximum by the end of the available time-series. At the time, we concluded that this out-of-phase characteristic was suggestive that the subsequent polarity reversal would also be offset between the hemispheres, which has turned out to be true (Karna, Hess Webber, and Pesnell, 2014).

Lastly, we note the work done in the Appendix of Paper II, towards understanding and removing an annual periodicity prevalent in the PCH areas. Unfortunately, our efforts were unable to successfully remove the fluctuations. We finally concluded that the periodicity stemmed from inherent projection effects within the SOHO ephemeris data and cannot be fully eliminated from the calculations. However, through these efforts, we gained important insight for interpreting the area results. We determined that the upper limit of the periodic data approximates the actual PCH areas better than the average of the signal. We therefore avoided using averaging techniques to smooth the area data.

Chapter 3: Wave and Helioseismic Theory

In this chapter, we present details concerning the specific theories involved in this dissertation research. The chapter is divided into two distinct topics. In Section 3.1, we develop mathematical and conceptual ideas of wave scattering, with some associated mathematical derivations given in Appendix B. This review is vital for the development of our simulation, which will be discussed later in Chapter 4. The second theory presented in Section 3.2 concerns TD helioseismology, including the general analysis method and the specifics of our new technique. Appendix C contains some supplementary material for this section. The TD theory discussion in Section 3.2 and Appendix C is a description of the process we use to analyze both the simulated and solar data sets.

3.1 Wave and Scattering Theory

The following explanation of the plane wave and scattering equations is primarily attributed to Morse and Feschbach (1953, Ch. 9.3, Ch. 11) and Morse and Ingard (1986). A more in-depth discussion in terms of solar physics can be found in Yang *et al.* (2012).

The solution to the total, time-dependent wave equation Ψ can be written in cylindrical coordinates r, φ, z as

$$\Psi = \psi(r, \varphi, z)e^{-i\omega t} \quad (3.1)$$

where the spatial wavefunction ψ is a solution to the Helmholtz equation (see Equation B.1). The time-dependent exponential term is related to the angular wave frequency ω , which we define by the f -mode dispersion relation in Equation 1.1.

In Appendix B.1, we share an in-depth derivation of the expanded plane wave wavefunction, ψ_p , for f -modes in cylindrical coordinates

$$\psi_p(r, \varphi, z) = Ae^{ikr \cos(\varphi)} = A \sum_{m=0}^{\infty} \epsilon_m i^m J_m(kr) \cos(m\varphi). \quad (3.2)$$

We use this expanded form to further derive the scattered plane wave expansion ψ_s off a cylinder aligned along the z -axis, which is detailed in Appendix B.2.

An overview of single-source scattering is discussed below in Section 3.1.1, including the mathematical formalism for a non-scatterer-centric system. This discussion is generalized to an N -scatterer system in Section 3.1.2, including single (each scatterer is unaware of its neighbors) and multiple (scattered waves are “re-scattered” by neighbors) scattering regimes. We are specifically interested in multiple-source scattering in this dissertation because elongated scattering sources can be represented by integrating over an array of discrete scatterers (see Section 3.2.3).

3.1.1 Single Scattering Source

The total time-independent wavefunction ψ is equivalent to the sum of the incident wavefunction ψ_i and resulting scattered wavefunction ψ_s . Figure 3.1 illustrates scattering of an incident plane wave off of a cylindrical scatterer, oriented along the z -axis. Since the incident wave is a plane wave, then by Equation 3.2, the incident wavefunction at some arbitrary point is

$$\psi_i = e^{i\mathbf{k}_i \cdot \mathbf{r}} = e^{ikr \cdot \hat{\alpha}} = e^{ikr \cos \varphi}. \quad (3.3)$$

The incident wave vector is denoted as $\mathbf{k}_i = k\hat{\alpha} = k[\cos \phi_i, \sin \phi_i]$ with magnitude $|\mathbf{k}_i| = k = 2\pi/\lambda$, where λ is the wavelength. The incident angle of the incoming wave with respect to the $+x$ -axis through the origin is given by ϕ_i . The position vector of the observation with respect to the origin is $\mathbf{r} = r\hat{r} = r[\cos \phi_r, \sin \phi_r]$, where ϕ_r is the angular position with respect to the $+x$ -axis. We define the angle $\varphi = \phi_i - \phi_r$, for simplicity. Note the use of the

trigonometric angle sum identity $\cos(\alpha \pm \beta) = \cos \alpha \cos \beta \mp \sin \alpha \sin \beta$.

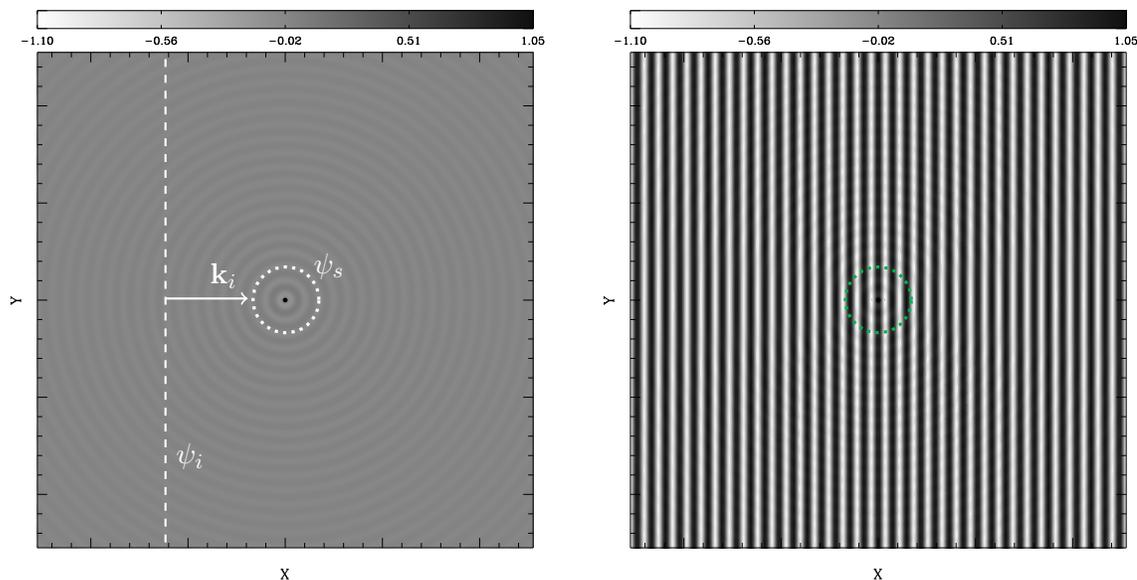


Figure 3.1: Scattering of an incident plane wave off of a point source scatterer; displaying only the scattered wave (left), and the total (incident + scattered) wave (right). Only the $m = 0$ mode is shown for $l = 600$. The dashed line indicates a single wavefront of the incident plane wave ψ_i . Note that the incident wave vector \mathbf{k}_i is normal to this wavefront. The resultant scattered wave ψ_s is radially symmetric about the point source in the far-field, as shown by the dotted line wavefront. The amplitudes shown here are normalized to the largest absolute value result from the total wave equation. This image was created using the *cyl_wav.pro* code, discussed in Section 4.1.

In the limit of $r \rightarrow \infty$, the scattered wave satisfies the boundary conditions of a wave diverging from a point source in the scattering region (see Appendix B.2). This boundary condition applies whether the scattering region is actually a point source or not. In cylindrical coordinates, this condition requires that

$$\psi_s \rightarrow f(\varphi) \frac{e^{ikr}}{\sqrt{r}}; \quad r \rightarrow \infty, \quad (3.4)$$

which is only dependent on two dimensions (Morse and Feschbach, 1953, p. 1065, Eq. 9.3.3).

The function f is called the scattering amplitude or angular distribution function. It is one dimensional since the only angular dimension is azimuthal.

In the far-field regime, the scattering amplitude is given by

$$f(\varphi) = f(\hat{\alpha}, \hat{r}) = A \sqrt{\frac{2}{\pi k}} e^{-ika} \sum_{m=0}^{\infty} \epsilon_m i^{m-1} \sin(\delta'_m) \cos(m\varphi), \quad (3.5)$$

where m is the azimuthal order, ϵ_m is the Neumann factor, a is the radius of the scattering cylinder, A is an arbitrary amplitude, and δ'_m are the derivatives of the related phase angles for the cylindrical Bessel functions (Morse and Feschbach, 1953, Ch. 11). See Appendix B.2 for an extensive derivation of this equation.

Shifting a Single Scatterer Away from the Origin

Consider an observation point P at $\mathbf{r}_p = r_p \hat{r}_p = r_p [\cos \phi_p, \sin \phi_p]$. As discussed above, in the limit as $\mathbf{r}_p \rightarrow \infty$, the scattered wave equation ψ_s of an incident plane wave from a single scatterer S_1 at the origin must look like

$$\psi_s \Rightarrow \frac{f(\hat{\alpha}, \hat{r}_p)}{\sqrt{r_p}} e^{ikr_p}. \quad (3.6)$$

If the scatterer is moved to some position $\mathbf{b}_1 = b_1 \hat{b}_1$ away from the origin, where $\hat{b}_1 = [\cos \phi_{b_1}, \sin \phi_{b_1}]$, then $\mathbf{r}_1 = \mathbf{r}_p - \mathbf{b}_1$ is the new position vector from S_1 to P (as shown in Figure 3.2).

With respect to the scattering point, the scattered wave equation becomes

$$\psi_s = \frac{f(\hat{\alpha}, \hat{r}_1)}{\sqrt{r_1}} e^{ikr_1} e^{ik\mathbf{b}_1 \cdot \hat{\alpha}}. \quad (3.7)$$

Note that as $b_1 \rightarrow 0$, $r_1 \rightarrow r_p$ and this equation reduces to Equation 3.6, as expected. Furthermore, if we shift our working coordinate space and reconsider S_1 to be the “new

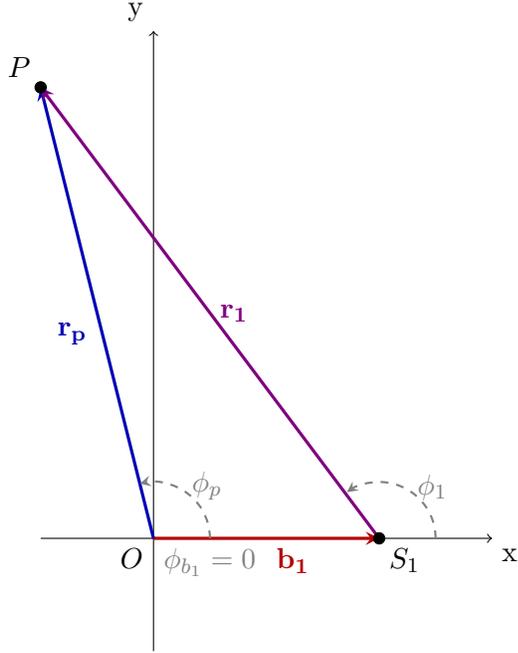


Figure 3.2: A single scatterer S_1 is shifted off the origin to a position $\mathbf{b}_1 = [b_{1x}, b_{1y}] = b_1[\cos \phi_{b_1}, \sin \phi_{b_1}]$ (the red vector). The observation point P is located at $\mathbf{r}_p = [r_{px}, r_{py}] = r_p[\cos \phi_p, \sin \phi_p]$ with respect to the origin (the blue vector). The position of P with respect to S_1 is $\mathbf{r}_1 = \mathbf{r}_p - \mathbf{b}_1 = [r_{px} - b_{1x}, r_{py} - b_{1y}]$ (the purple vector). The distance between P and S_1 is therefore $|\mathbf{r}_1| = r_1 = \sqrt{(r_{px} - b_{1x})^2 + (r_{py} - b_{1y})^2}$.

origin,” then the only difference between Equation 3.7 and Equation 3.6 is the exponential term dependent on \mathbf{b}_1 . This indicates that the only physical change made by moving the scatterer is a phase-shift!

With an understanding of the basic principles for scattering by a single point, we move on to discuss scattering when more than one scattering source is present.

3.1.2 Multiple Scattering Sources

We now consider the effects of wave scattering from more than one scatterer. Multiple scattering sources lead to wave interactions between the incident wave and all of the scattered

waves produced by individual scattering sources. A simple example of this is the double-slit experiment for light waves. The same general principle applies to acoustic and gravity waves.

If the scattering cross-section of each scatterer is smaller than the distance to its neighbors, then each scatterer behaves as though its neighbors do not exist, e.g. as a single point scatterer. This is commonly referred to as single scattering (not to be confused with scattering by a single source, discussed in Section 3.1.1). If the scattering cross-sections are large enough to influence the neighboring scattering points, each scatterer then “notices” the others. In this case, the total scattering is complicated by the interaction between the original scattered wave from each point and the “re-scattered” wave by the neighboring points. This is called multiple scattering.

Additionally, there is differentiation between dependent and independent scattering. Dependent scattering occurs when there are systematic phase relations between scatterers, which leads to strong constructive and destructive interference in the scattered waves, particularly in the direction of the incident wave. This is also referred to as coherent scattering. Independent scattering, also known as incoherent scattering, is therefore scattering without any systematic phase relationship between the scatterers. This assumption can be made if the scattering points are sufficiently far away from one another. It has been shown that a “sufficient” separation between scatterer centers is a distance of at least three times the scatterers’ radius (being identical in size) (van de Hulst, 1957). Incoherent scattering is weaker than coherent scattering. However, it is still non-zero and is important when observing the results of well-spaced single scattering, or when the observation is made off-axis to the incident wave.

In this dissertation, we focus on coherent, single scattering. In practice, our exploitation of the system geometry (see Section 3.2.3) drowns out any weaker incoherent or multiple scattering signal. However, we include these concepts in our review to better understand the overall theory, and why it is reasonable to neglect them. In Section 3.1.2.1, we discuss the basic theory of single multi-source scattering and review the mathematical formulation of

coherent versus incoherent scattering. We then briefly overview the more complex multiple scattering in Section 3.1.2.2. For further details on the differences between dependent and independent scattering, as well as single versus multiple scattering, see Martin (2006, p.1-2).

3.1.2.1 Single Scattering

Suppose an arbitrary observation point P is located in the far-field at $\mathbf{r}_p = [r, \phi_p]$ with respect to the origin. Similarly, there is a row of N cylindrical scattering points, each located at \mathbf{b}_n from the origin in the same z -plane as P , where $n = 1, 2, \dots, N$. We define γ_n to be the angle between \mathbf{r}_p and \mathbf{b}_n . Let $\mathbf{r}_n = [r_n, \phi_n]$ denote the distance vector from each scatterer to P . A plane wave, described by Equation 3.3, is incident on the scatterers with zero phase shift at the origin. Figure 3.3 shows a simple schematic of this situation for $N = 2$ scattering points.

Assume the scattering points are far enough apart that near-field interactions can be neglected. Combining Equations 3.5 and 3.7, and summing over the N points, we find the total scattered wave at P is

$$\psi_s = \sum_{n=1}^N \frac{f(\varphi_n)}{\sqrt{r_n}} e^{i(\mathbf{k}_i \cdot \mathbf{b}_n + kr_n)}. \quad (3.8)$$

For clarity of notation, we define \mathbf{k}_n to be the scattered wave vector from the n^{th} scatterer (recall, $|\mathbf{k}_i| = |\mathbf{k}_n| = k$); φ_n is the angle between \mathbf{k}_n and \mathbf{k}_i , referred to as the scattering angle. The first term in the exponent is the incident plane wave response at the n^{th} scatterer. The second exponential term describes the resultant scattered wave in the direction from \mathbf{b}_n toward P ($\mathbf{k}_n \cdot \mathbf{r}_n = kr_n$).

If the location of P is sufficiently far from the region containing the scattering points ($\mathbf{r}_p \gg \mathbf{b}_n$), then $\frac{1}{\sqrt{r_n}} \sim \frac{1}{\sqrt{r}}$, and the law of cosines, $r_n^2 = r^2 + b_n^2 - 2rb_n \cos \gamma_n$, can be

$$\begin{aligned}
&\Rightarrow \frac{f(\varphi)}{\sqrt{r}} \sum_{n=1}^N e^{i[\mathbf{k}_i \cdot \mathbf{b}_n + kr(1 - \frac{b_n}{r} \cos \gamma_n)]} \\
&\Rightarrow \frac{f(\varphi)}{\sqrt{r}} \sum_{n=1}^N e^{i[\mathbf{k}_i \cdot \mathbf{b}_n + kr - kb_n \cos \gamma_n]} \\
&\Rightarrow f(\varphi) \frac{e^{ikr}}{\sqrt{r}} \sum_{n=1}^N e^{i[\mathbf{k}_i \cdot \mathbf{b}_n - \mathbf{k}_p \cdot \mathbf{b}_n]} \\
\psi_s &\simeq f(\varphi) \frac{e^{ikr}}{\sqrt{r}} \sum_{n=1}^N e^{i\boldsymbol{\mu} \cdot \mathbf{b}_n} \tag{3.9}
\end{aligned}$$

where $\boldsymbol{\mu} = \mathbf{k}_i - \mathbf{k}_p$ and $\mu = |\mathbf{k}_i - \mathbf{k}_p| = \sqrt{\mathbf{k}_i^2 + \mathbf{k}_p^2 - 2|\mathbf{k}_i||\mathbf{k}_p| \cos(\phi_i - \phi_p)} = \sqrt{2k^2(1 - \cos \varphi)} = \sqrt{4k^2 \sin^2(\frac{\varphi}{2})} = 2k \sin(\frac{\varphi}{2})$, with $\varphi = \phi_i - \phi_p$ being the angle between \mathbf{k}_i and \mathbf{k}_p . (Note that we intentionally define all angles with respect to the x -axis as ϕ and all difference angles between vectors as φ .)

Equation 3.9 shows that in the far-field regime, the scattered wave field from multiple scatterers can be expressed simply as the far-field single-scatterer solution multiplied by a factor of the sum of the phase shifts contributed by each scatterer.

Scattering by a Row of N Equally-Spaced Scatterers

Consider a line of N scatterers, evenly distributed with spacing $d < D$ along $-D/2 \leq x \leq D/2$ on the x -axis. Assuming the x -coordinate of the first scatterer is x_1 , the x -coordinate for any of the scatterers can be written as $x_n = x_1 + (n - 1)d$ where $n = 1, 2, \dots, N$. Since the scatterers are positioned on the x -axis, $\mathbf{b}_n = x_n \hat{x}$ ($y = z = 0$). For simplicity, let us assume that $\boldsymbol{\mu} = \mu \hat{x}$ as well (in other words, $\phi_i = 180^\circ - \phi_p \Rightarrow \varphi = 2\phi_i - 180^\circ$, or ϕ_i and ϕ_p must be reflected in π). Therefore, from Equation 3.9,

$$\psi_s = f(\varphi) \frac{e^{ikr}}{\sqrt{r}} \sum_{n=1}^N e^{i\mu x_n}$$

$$\begin{aligned}
&= f(\varphi) \frac{e^{ikr}}{\sqrt{r}} \sum_{n=1}^N e^{i\mu(x_1+(n-1)d)} \\
&= f(\varphi) \frac{e^{ikr}}{\sqrt{r}} e^{i\mu x_1} \sum_{n=1}^N e^{i\mu(n-1)d}
\end{aligned} \tag{3.10}$$

Let $\alpha = \mu d$ and note that

$$\sum_{n=1}^N e^{i\alpha(n-1)} = \frac{e^{i\alpha N} - 1}{e^{i\alpha} - 1}. \tag{3.11}$$

Substituting this series into the previous equation and making use of Euler's formula ($e^{ix} = \cos x + i \sin x$), the half-angle formula ($1 - \cos \phi = 2 \sin^2[\phi/2]$), and the double-angle formula ($\sin \phi = 2 \sin[\phi/2] \cos[\phi/2]$), we simplify as follows:

$$\begin{aligned}
\psi_s &= \left(\frac{e^{ikr}}{\sqrt{r}} \right) f(\varphi) e^{i\mu x_1} \left[\frac{e^{i\alpha N} - 1}{e^{i\alpha} - 1} \right] \\
&= \left(\frac{e^{ikr}}{\sqrt{r}} \right) f(\varphi) e^{i\mu x_1} \left[\frac{\cos(\alpha N) + i \sin(\alpha N) - 1}{\cos(\alpha) + i \sin(\alpha) - 1} \right] \\
&= \left(\frac{e^{ikr}}{\sqrt{r}} \right) f(\varphi) e^{i\mu x_1} \left[\frac{-2 \sin^2(\frac{1}{2}\alpha N) + i \sin(\alpha N)}{-2 \sin^2(\frac{1}{2}\alpha) + i \sin(\alpha)} \right] \\
&= \left(\frac{e^{ikr}}{\sqrt{r}} \right) f(\varphi) e^{i\mu x_1} \left[\frac{-2 \sin^2(\frac{1}{2}\alpha N) + 2i \sin(\frac{1}{2}\alpha N) \cos(\frac{1}{2}\alpha N)}{-2 \sin^2(\frac{1}{2}\alpha) + 2i \sin(\frac{1}{2}\alpha) \cos(\frac{1}{2}\alpha)} \right] \\
&= \left(\frac{e^{ikr}}{\sqrt{r}} \right) f(\varphi) e^{i\mu x_1} \left[\frac{\sin(\frac{1}{2}\alpha N)}{\sin(\frac{1}{2}\alpha)} \right] \left[\frac{-\sin(\frac{1}{2}\alpha N) + i \cos(\frac{1}{2}\alpha N)}{-\sin(\frac{1}{2}\alpha) + i \cos(\frac{1}{2}\alpha)} \right] \\
&= \left(\frac{e^{ikr}}{\sqrt{r}} \right) f(\varphi) e^{i\mu x_1} \left[\frac{\sin(\frac{1}{2}\alpha N)}{\sin(\frac{1}{2}\alpha)} \right] \left[\frac{\cos(\frac{1}{2}\alpha N) + i \sin(\frac{1}{2}\alpha N)}{\cos(\frac{1}{2}\alpha) + i \sin(\frac{1}{2}\alpha)} \right] \\
&= \left(\frac{e^{ikr}}{\sqrt{r}} \right) f(\varphi) e^{i\mu x_1} \left[\frac{\sin(\frac{1}{2}\alpha N)}{\sin(\frac{1}{2}\alpha)} \right] \left[\frac{e^{\frac{1}{2}i\alpha N}}{e^{\frac{1}{2}i\alpha}} \right]
\end{aligned}$$

$$\begin{aligned}
&= \left(\frac{e^{ikr}}{\sqrt{r}} \right) f(\varphi) e^{i\mu x_1} \left[\frac{\sin(\frac{1}{2}\mu d N)}{\sin(\frac{1}{2}\mu d)} \right] e^{\frac{1}{2}i\alpha(N-1)} \\
&= \left(\frac{e^{ikr}}{\sqrt{r}} \right) f(\varphi) e^{i\mu(x_1 + \frac{1}{2}(N-1)d)} \left[\frac{\sin(\frac{1}{2}N\mu d)}{\sin(\frac{1}{2}\mu d)} \right] \tag{3.12}
\end{aligned}$$

We note that the $(N-1)$ term in the exponent should just be an N , according to Morse and Feschbach (1953, Eq. 11.3.84), which we were unable to resolve; however, we also point out that for large N (or small d), the effect of this discrepancy is negligible. We also observe that this term drops out for $N = 1$, leaving behind the scattering equation for a single scatterer (indeed, taking this into account may be the distinction between Equation 3.12 and that of Morse and Feschbach). Therefore, the equations are functionally the same and can be interpreted the same way.

Recall that $\mu = 2k \sin(\frac{\varphi}{2})$ is the magnitude of the difference between the incoming wave vector and the wave vector in the direction of the observation. As $\varphi \rightarrow 0$, the incident wave vector is increasingly parallel to the line of observation and thus $\mu \rightarrow 0$. In this case, $e^{i\mu(\dots)} \rightarrow 1$ and $[\sin(\frac{1}{2}N\mu d)/\sin(\frac{1}{2}\mu d)] \rightarrow N$. Thus, $\psi_s \rightarrow N f(\varphi)(e^{ikr}/\sqrt{r})$. Note that this limit also holds true for any φ while $Nkd \ll 1$ ($2\pi D \ll \lambda$), also known as the long-wavelength approximation. This behavior produces a purely coherent scattered wave.

When $\mu \rightarrow 0$ and $Nkd > 1 \gg kd$ ($2\pi D > \lambda \gg 2\pi d$), the small-angle approximation only applies to the sine term in the denominator of Equation 3.12. Therefore, the scattered wavefunction can be approximated as $\psi_s \rightarrow f(\varphi)(e^{ikr}/\sqrt{r})[\sin(\frac{1}{2}N\mu d)/(\frac{1}{2}\mu d)]$. Changes in φ thus have a strong affect on the scattering amplitude. For $\varphi \rightarrow 0 = 2\pi$, the extra amplitude term converges to a factor of N (as discussed above). However, as $\varphi \rightarrow \pi$, the scattered wave amplitude is greatly damped, considerably smaller than N times the scattering amplitude of a single scatterer. Thus, the scattered beam is weak except for an intense ‘‘shadow’’ beam on the forward-side of the scatterers (with respect to the direction of wave motion). This is a combination of both coherent and incoherent scattering.

In the small wavelength approximation, $\lambda < 2\pi d$, Equation 3.12 must be used in full

and results in a mainly incoherent wave. However, when $\mu = m\pi$, where m is an integer, the sine term in the denominator again goes to zero, producing a strong, forward-scattered beam. It may be of interest that the angles at which this occurs are the same as the Bragg angles in the diffraction of x-rays off a crystal.

Also note that we have made the assumption here that $\boldsymbol{\mu}$ is parallel to the row of scatterers, which is not always (or even generally) the case. In our study, we examine incident waves traveling perpendicular to a row of scatterers (Chapters 4-5). If the row of scatterers is perpendicular to $\boldsymbol{\mu}$, for example with $\mathbf{b}_n = y_n \hat{y}$, then

$$\begin{aligned}
 \psi_s &= f(\varphi) \frac{e^{ikr}}{\sqrt{r}} \sum_{n=1}^N e^{i\mu y_n \hat{x} \cdot \hat{y}} \\
 &= f(\varphi) \frac{e^{ikr}}{\sqrt{r}} \sum_{n=1}^N e^0 \\
 &= f(\varphi) \frac{e^{ikr}}{\sqrt{r}} \sum_{n=1}^N 1 \\
 &= N f(\varphi) \frac{e^{ikr}}{\sqrt{r}}
 \end{aligned} \tag{3.13}$$

which is the same result as for the “parallel” coherent scattering case above.

With a general understanding of single scattering from multiple scatterers, we now extend the discussion to higher order scattering terms.

3.1.2.2 Multiple Scattering

When the wave scattering off of multiple scatterers is strong, the waves around each scatterer are influenced by the presence of the other scatterers. In addition to the incoming incident wave and outgoing scattered wave from each scattering point, we must also consider the incoming re-scattered waves from all the scattering points in the surrounding near-field region (Martin, 2006; Morse and Feschbach, 1953). Because we are operating under the

assumptions of coherent, long-wavelength and far-field scattering, we can neglect the additional multiple scattering effects for the purposes of this work. However, it is instructive to understand the basics of multiple scattering, so a brief overview of the math is given here, based mainly on the discussion in Martin (2006, Ch. 8.3).

While the total scattered field from a row of scatterers is not isotropic, we will assume that each individual scatterer produces a symmetrically scattered wave. Under this assumption, the scattered wave field will act as a function $\psi_{s,j}(\mathbf{r}) = A_j G(\mathbf{r} - \mathbf{r}_j)$, where A_j is the amplitude of the scattered wave, G is the free-space Green's function, and \mathbf{r} , \mathbf{r}_j are the position vectors at the observation and j^{th} scatterer, respectively, with respect to the origin. In comparison to Equation 3.4 for a single scatterer observed in the far-field, we could relate $A_1 \implies f(\varphi)$ and $G(\mathbf{r} - \mathbf{r}_1) \implies e^{ik|\mathbf{r}-\mathbf{r}_1|}/\sqrt{|\mathbf{r} - \mathbf{r}_1|}$. For ease of notation and generality for multiple scatterers, we will continue from here using the simplified isotropic formula.

We can write the total wave field for N scatterers as

$$\begin{aligned}\psi(\mathbf{r}) &= \psi_i(\mathbf{r}) + \sum_{j=1}^N \psi_{s,j}(\mathbf{r}) \\ &= \psi_i(\mathbf{r}) + \sum_{j=1}^N A_j G(\mathbf{r} - \mathbf{r}_j).\end{aligned}\tag{3.14}$$

In and of itself, this is the same equation as used for the single scattering discussed in Section 3.1.2.1. We then further define an “effective” scattered field by

$$\psi_n(\mathbf{r}) \equiv \psi(\mathbf{r}) - A_n G(\mathbf{r} - \mathbf{r}_n) = \psi_i(\mathbf{r}) + \sum_{\substack{j=1 \\ j \neq n}}^N A_j G(\mathbf{r} - \mathbf{r}_j).\tag{3.15}$$

This expression gives the resulting incident wave field on the n^{th} scatterer due the initial

incident wave *and* all the scattered waves from the remaining $N - 1$ scatterers. The amplitude of the wave scattered by this n^{th} scatterer is proportional to the effective field incident on it, which can be expressed as

$$A_n = g_n \psi_n(\mathbf{r}_n), \quad (3.16)$$

where g_n is the scattering coefficient for the n^{th} scatterer. We will not go into details on how to determine g_n here; see Martin (2006, Ch. 8.3.1) for a thorough explanation. The definition of A_n extends to the scattered amplitudes of all j scatterers. Therefore, the effective scattered field near the n^{th} scatterer can be expressed by

$$\psi_n(\mathbf{r}_n) = \psi_i(\mathbf{r}_n) + \sum_{\substack{j=1 \\ j \neq n}}^N g_j \psi_j(\mathbf{r}_j) G(\mathbf{r}_n - \mathbf{r}_j). \quad (3.17)$$

This equation is a linear system for $\psi_j(\mathbf{r}_j)$ and can be used iteratively within the total wave equation to account for multiple scattering

$$\psi(\mathbf{r}) = \psi_i(\mathbf{r}) + \sum_{j=1}^N g_j \psi_j(\mathbf{r}_j) G(\mathbf{r} - \mathbf{r}_j). \quad (3.18)$$

The order of the multiple scattering is determined by the number of iterations of $\psi_j(\mathbf{r}_j)$. For example, if we consider 1st-order scattering to be that of single scattering (only the incident wave is scattered, there is no re-scattering of neighboring scattered waves), then the 2nd-order scattering equations is

$$\psi(\mathbf{r})_{O^2} = \psi_i(\mathbf{r}) + \sum_{j=1}^N g_j G(\mathbf{r} - \mathbf{r}_j) \left[\psi_i(\mathbf{r}_j) + \sum_{\substack{l=1 \\ l \neq j}}^N g_l \psi_l(\mathbf{r}_l) G(\mathbf{r}_j - \mathbf{r}_l) \right]. \quad (3.19)$$

Similarly, the 3rd-order multiple scattering equation is

$$\psi(\mathbf{r})_{O3} = \psi_i(\mathbf{r}) + \sum_{j=1}^N g_j G(\mathbf{r} - \mathbf{r}_j) \left[\psi_i(\mathbf{r}_j) + \sum_{\substack{l=1 \\ l \neq j}}^N g_l G(\mathbf{r}_j - \mathbf{r}_l) \right. \\ \left. \times \left[\psi_i(\mathbf{r}_l) + \sum_{\substack{m=1 \\ m \neq l}}^N g_m \psi_m(\mathbf{r}_m) G(\mathbf{r}_l - \mathbf{r}_m) \right] \right] \quad (3.20)$$

and so on. See Figure 3.4 for a conceptual schematic of this process.

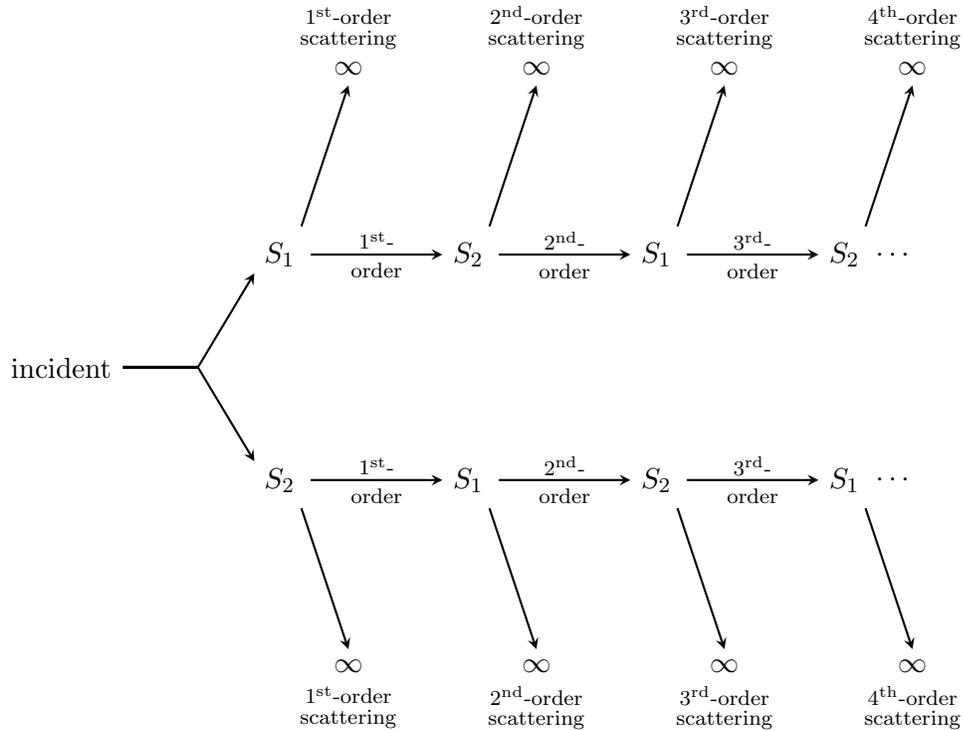


Figure 3.4: Compare to Fig. 8.1, page 311 in Martin (2006).

For a more detailed discussion, see Morse and Feschbach (1953, p. 1499-1500) and Martin

(2006, Ch. 8.3). Martin also has specific discussions for multiple scattering of acoustic, electromagnetic, elastic, and water waves.

In brief summary of Section 3.1, the cylindrical scattering equation can be derived from the interaction of plane wave on a rigid cylinder with its major axis perpendicular to the surface of the medium. The scattering from a single cylinder at the origin is produced using boundary conditions at the cylinder and at infinity. If we move the cylinder away from the origin, the change in the scattering equation equates to an additional phase-shift term, which is dependent on the new location of the scatterer. The effect of multiple scatterers is generalized simply as the sum of the scattering from each of the scatterers (think superposition). This assumes that the scatterers are spread far enough apart that they do not "notice" the scattered waves from their neighbors. If the scatterers are close together, then higher order scattering effects must be considered (e.g., the total scattered signal from any one scatterer would be the 1st order scattering from the incident plane wave, and the higher order "re-scattering" of scattered waves coming from nearby neighbors). We can implement the far-field and long-wavelength approximations to ensure that we do not need to consider these multiple-scattering effects.

3.2 Time-Distance Helioseismology Theory

In this section, we provide a detailed review of TD helioseismology concepts, as they pertain to this dissertation. We discuss the implementation of a phase-speed filter used to isolate the f -modes in the SDO/HMI data in Section 3.2.1. Section 3.2.2 describes the process used to calculate cross-covariances in the TD method, as well as the importance and application of measuring travel times. We also consider the divergence of our 1D linear method from “usual” time-distance techniques in Section 3.2.3. The mathematics behind linear correlations is reviewed, as is the connection between our summed cross-correlations and Green’s function. Finally, we provide cases of precedence that support our exploitation of elongated feature geometry in the new method.

3.2.1 Phase-Speed Filtering

In order to remove noise and unwanted wave modes, mode filters are applied to the data in the Fourier domain. These filters affect both frequency ω (temporal) and wavenumber k (spatial) dimensions, but for simplicity we will focus our discussion on frequency filtering. However, keep in mind that the Fourier and filtering details described here translate directly to wavenumber filtering, and that this research uses a full-spectral mode filter.

Depending on whether the frequencies to be removed are above or below the desired passband of frequencies, different filters are used. To remove higher frequencies, a low-pass filter is used (it allows low frequencies to pass through and blocks those that are higher than some cutoff frequency), while high-pass filters attenuate lower frequencies. In some cases, the same filtering function can be used as both a high- and low-pass filter. Combined, they create a bandpass filter, which collectively removes all undesirable frequencies, transmitting only a specific passband of signals. The shape of the filtering function dictates the amplitude and rapidity of frequency attenuation, and its width controls the bandwidth of the passband. Broad bandpass filters have wider bandwidths so the passband includes a wider range of frequencies. Similarly, narrow bandpass filters have smaller bandwidths and are more focused on specific frequencies.

A Gaussian filter is often the preferred bandpass filter function, particularly for image processing, because when a Gaussian function is transformed from the Fourier domain to the spatial domain (or vice versa), it remains a Gaussian. However, the Gaussian filter is a narrow bandpass filter with rapid signal attenuation about the peak frequency, which is limiting if we want to isolate a strong signal spread over a wide passband of frequencies. The same does not apply for broad differentiable piecewise functions such as an “Ideal” (or “Boxcar”) filter, but the sharp edges of these functions create distortions in the Fourier domain. See Appendix C for proof of these statements.

The filter function used in this work is that described by Duvall, Jr., Birch, and Gizon (2006). The filter is a flat-top function with cosine-bell edges (hereafter referred to as the FTC filter). The FTC filter is applied separately at each wavenumber k with a central frequency ω given by the f -mode dispersion relation (Equation 1.1), and a width that also depends indirectly on k , as follows. If the central f -mode frequency for some k is ω_f and the p_1 -mode frequency for k is ω_p , then the magnitude difference in frequency between the two modes is $\Delta\omega = |\omega_p - \omega_f|$. The FTC filter allows a full-power pass at frequencies between $\pm 0.5\Delta\omega$ and allows no power through at frequencies greater than $|\pm 0.75\Delta\omega|$. Half of a cosine function dictates the filtering power between $|\pm 0.5\Delta\omega| < \omega < |\pm 0.75\Delta\omega|$.

An example of the FTC filtering function is shown in Figure 3.5a, along with a Gaussian function and two Boxcar functions. The corresponding fast Fourier transforms (FFTs) for each function are shown in Figure 3.5b. Again, these functions and their transforms are only shown in the frequency/time (ω/t) domain, but the filters apply just as well to wavenumber/space (k/r) or combined full-spectral ($(\omega, k)/(r, t)$) results. The FFT of the FTC is a good approximation to the Gaussian FFT and has only minor side-lobes. For these reasons, we deem the FTC function an acceptable filter for separating out the f -mode.

The spectral results of an f -mode-filtered sample data cube with the FTC function are given in the ω - k diagrams in Figure 3.6. The unfiltered power spectrum is shown in Figure 3.6a, for comparison. An acceptable p -mode filter (Gaussian high pass, provided by J. Zhao) is applied for the results given in Figure 3.6b. Our FTC filter has been applied

for the result in Figure 3.6c. The code that applies the FTC filter to a given data cube is called *filter_yesf.pro*. For reference, the first time step of a sample HMI tracked dopplergram data cube is shown in Figure 3.7. Figure 3.7a displays the original, unfiltered data used to produce Figure 3.6a. Likewise, the resultant first time step of p -mode and f -mode filtered data, returned from transforming Figures 3.6b and 3.6c, are shown in Figures 3.7b and 3.7c, respectively.

We point out that the ω - k diagrams in Figure 3.6 exhibit the dispersion relations of the various wave modes. The spectral f -mode wave signal in Figure 3.7c is thus representative of the f -mode dispersion relation given in Equation 1.1.

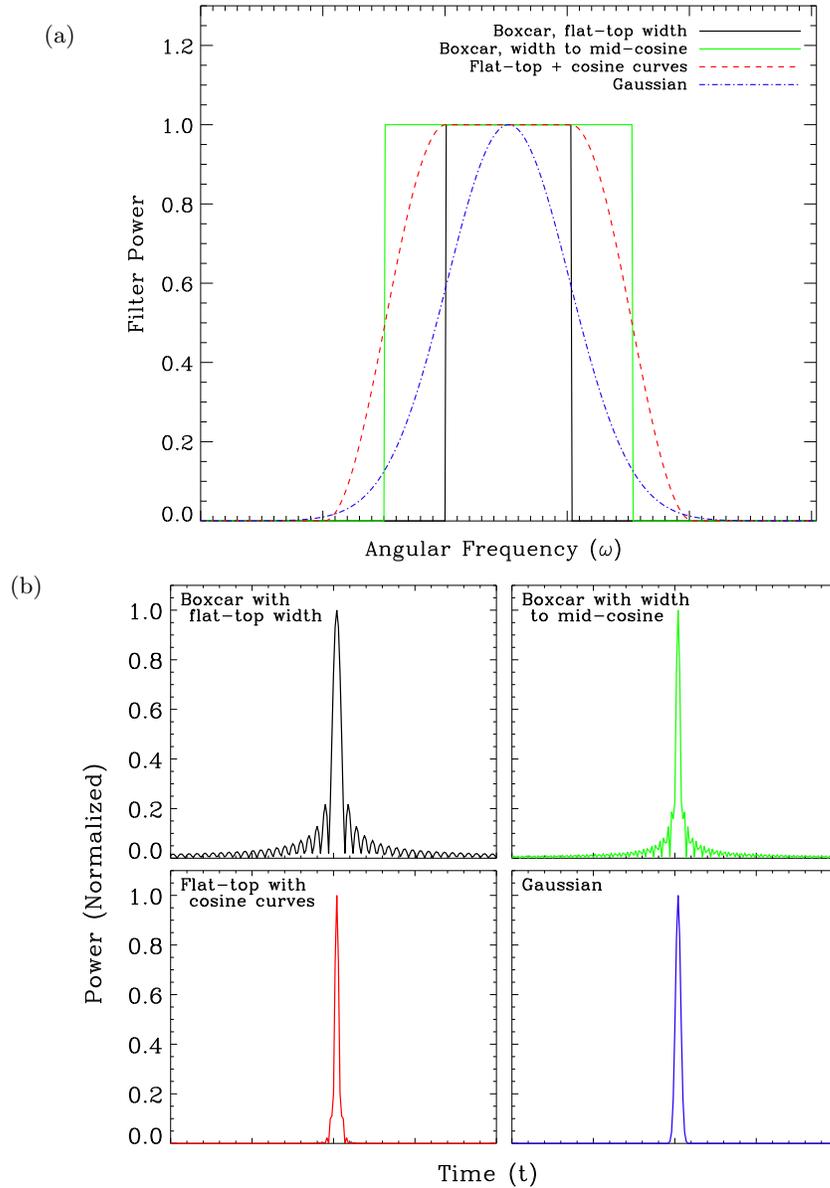


Figure 3.5: An example of various frequency-filtering functions is shown in (a). The blue dot-dashed ($\cdot - \cdot - \cdot - \cdot -$) line is a Gaussian function, the black and green solid lines (--- , ---) are two boxcar functions of varying widths, and the red dashed curve (---) is an FTC filter. The Fourier transforms of these functions are shown in (b), with corresponding colors and linestyles. Both boxcar functions have significant side-lobes in their transforms, which disrupt the clarity of the transformed signal. The transform of the Gaussian is, as expected, another Gaussian so the transformed signal is “perfect”. However, the downside of the Gaussian filter is the abrupt loss of power about a peak frequency, which is limiting if you have a strong signal spread over a broad range of frequencies. The FTC filter is useful as a full bandpass filter of all relevant frequencies while minimizing the effect of disadvantageous side-lobes, as can be seen when comparing the transforms in (b). Note, this frequency-filtering example is a simplified illustration but the concept holds true independent of the filtering variable.

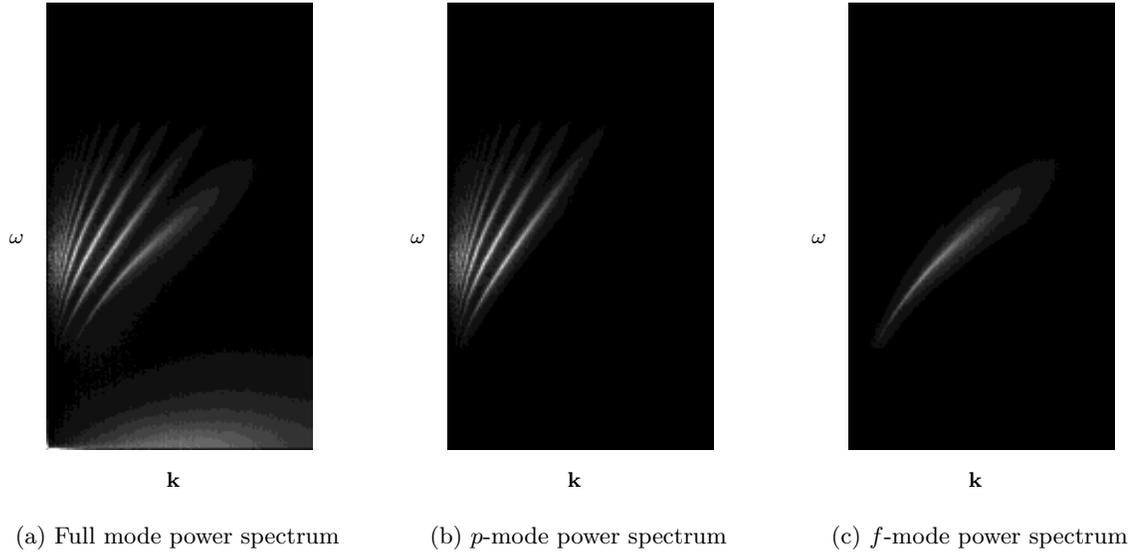


Figure 3.6: These phase-speed diagrams (also known as ω - k diagrams; ω is angular frequency and k is wavenumber) display the spectral results of the sample data cube shown in Figure 3.7a. The full power spectrum in (a) shows multiple waves modes, as well as a low-frequency signal (where researchers expect to find the elusive g -mode). The low-frequency signal is filtered out in both (b) and (c). Any signal below the median phase-speed distance between the f - and p_1 -modes is filtered out to show only the various p_n -modes in (b). Similarly, all phase-speeds above the f - and p_1 -mode median are filtered out of (c), leaving only the f -mode signal.

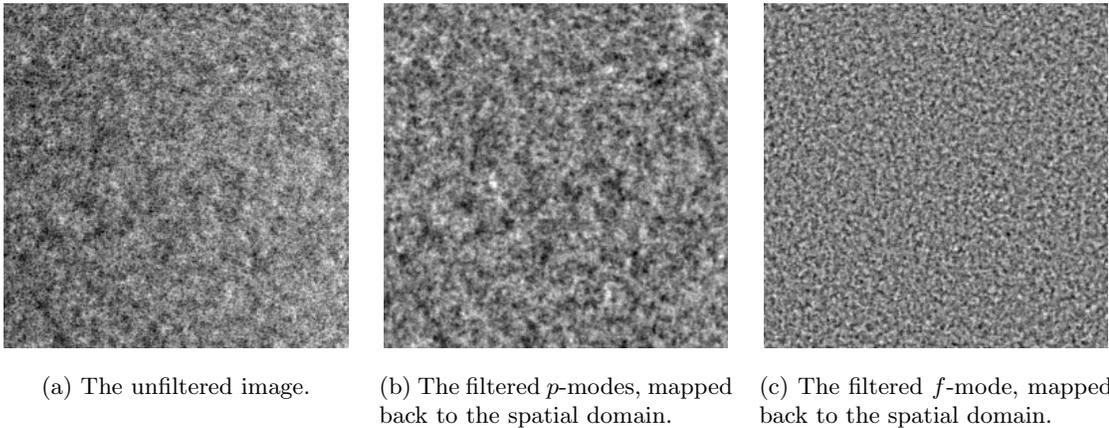


Figure 3.7: An example time step of an HMI tracked mapped doppler data cube. The tracked data is initially unfiltered (a). The various modes are filtered from the full power spectrum in the Fourier domain (as shown in Figure 3.6), and the filtered spectra are then transformed back to the spatial domain, giving just the p -modes (b) and the isolated f -mode (c).

3.2.2 Extracting Travel Times

3.2.2.1 Calculating Cross-Covariances

The cross-covariance between two complex functions $f(\mathbf{r}, t)$ and $g(\mathbf{r}, t)$ is defined as the convolution

$$C(f, g, \mathbf{t}) = f(\mathbf{r}, \mathbf{t}) * g(\mathbf{r}, \mathbf{t}) \equiv \int_{-\infty}^{\infty} f^*(\mathbf{r}, t)g(\mathbf{r}, t + \mathbf{t})dt, \quad (3.21)$$

where f^* is the complex conjugate of f , \mathbf{r} and t are position and time, respectively, and \mathbf{t} is the positive or negative time lag. If the two functions are real (as for our purposes), then $f^* = f$.

The cross-covariance between two observed points over time can be expressed as

$$C(\mathbf{r}_s, \mathbf{r}_r, \mathbf{t}) = \frac{1}{T} \int_0^T \phi(\mathbf{r}_s, t)\phi(\mathbf{r}_r, t + \mathbf{t})dt, \quad (3.22)$$

where \mathbf{r}_s and \mathbf{r}_r denote the locations of the source and receiver points, respectively, and t represents the time interval between the start of the observation at $t = 0$ and the total temporal duration of the observation T (Couvidat *et al.*, 2012; Giles, 1999; Gizon and Birch, 2005). We define $\Delta = \|\mathbf{r}_r - \mathbf{r}_s\|$ as the distance between the points. The quantity $\phi(\mathbf{r}, t)$ is the filtered LOS signal of the data cube in position and time. As discussed in Section 3.2.1, the mode filter is applied to the data cube in Fourier space. The filtered data are transformed back into spatial and temporal dimensions for the cross-covariance calculations.

Note that the difference between the cross-covariance C and the cross-correlation C_0 is simply a normalization factor. Statistically, the normalization is the inverse of the standard deviations σ , $C_0(f, g, \mathbf{t}) = C(f, g, \mathbf{t})/\sigma_f\sigma_g$. Analytically, $C(\mathbf{r}_s, \mathbf{r}_r, \mathbf{t})$ can be normalized to the cross-covariance at $\mathbf{r}_s = \mathbf{r}_r$ ($\Delta = 0$) and $\mathbf{t} = 0$, $C_0(\mathbf{r}_s, \mathbf{r}_r, \mathbf{t}) = C(\mathbf{r}_s, \mathbf{r}_r, \mathbf{t})/C(\mathbf{r}_s, \mathbf{r}_s, 0)$.

For TD helioseismology, the process of computing the cross-covariance between two point sources in an HMI data set is as follows. First, choose the source pixel, denoting the

source location $\mathbf{r}_s = [x_s, y_s]$. Choose the receiver pixel a distance Δ away at $\mathbf{r}_r = [x_r, y_r]$. Calculate the cross-covariance between the signals $\phi(\mathbf{r}_s, t)$ and $\phi(\mathbf{r}_r, t + \mathbf{t})$ for each time step (each image in the data cube), using Equation 3.22 (or the built in IDL function `C_CORRELATE()`).

This process is repeated using each non-edge pixel as \mathbf{r}_s . The “edges” are defined as all pixels less than the distance Δ away from the boundary of the image. Furthermore, the cross-covariances can be computed for multiple \mathbf{r}_r of varying distances and/or directions. The final set of cross-covariance data is a matrix with dimensions [$\#$ of \mathbf{r}_s , $\#$ of \mathbf{r}_r , $\#$ of \mathbf{t}].

This matrix (and the number of calculations required to produce it) can be reduced by considering different averaged geometries of receiver pixels in specific directions/distances. Averaging in these geometries also helps isolate flow patterns in specific directions, and improves the signal-to-noise ratio of the cross-correlations results. For example, in the common center-to-annulus geometry, \mathbf{r}_s is considered to be a central pixel surrounded an annulus of radius Δ . This averaging scheme emphasizes radial flows toward and away from the central point. The signal of each annulus is taken as the averaged signals at the receiver pixels that fall within Δ around the central source pixel. Mathematically, this averaged signal is expressed as $\bar{\phi}(\Delta, t) = \sum_{i=1}^N \phi(\mathbf{r}_{ri}, t)/N$, where N is the number of pixels that falls within the annulus radius, $\mathbf{r}_{ri} = [x_s + \Delta \cos(\theta_i), y_s + \Delta \sin(\theta_i)]$ is the location of the i^{th} receiver pixel in the annulus, and θ_i is the angle between pixel i and a line intersecting \mathbf{r}_s parallel to the $+x$ -axis. The cross-covariance in Equation 3.22 then becomes $C(\mathbf{r}_s, \Delta, \mathbf{t}) = \frac{1}{T} \int_0^T \phi(\mathbf{r}_s, t) \bar{\phi}(\Delta, t + \mathbf{t}) dt$. The calculations are repeated for multiple annuli of varying Δ . We do not use the center-to-annulus geometry in our work; however, the ability to average point-to-point cross-correlations over specific geometries is highly important for our new technique. We discuss the implications and application to our method further in Section 3.2.3.

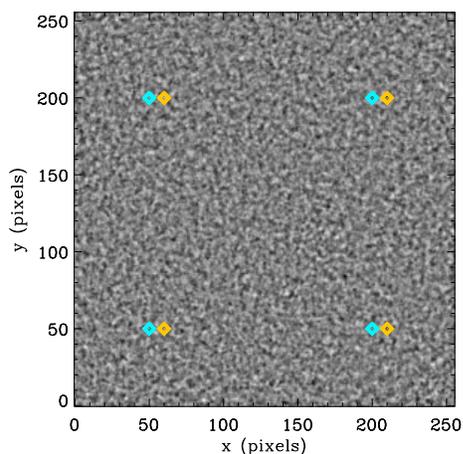
Figure 3.8 illustrates the cross-correlation process applied to the example f -mode data shown in Figure 3.7c. The data in Figure 3.8a is shown with four different source pixels,

each with a receiver point a distance $\Delta = 10$ pixels away in the x direction. Figure 3.8b displays the corresponding cross-correlation plots of the point pairs in (a). By themselves, they do not show clear signals of correlated inbound and outbound wave propagation. However, averaging the cross-correlation results of all possible source pixels with the same Δ improves the signal-to-noise ratio so that visible wave signals are evident propagating inward and outward, as shown in Figure 3.8c.

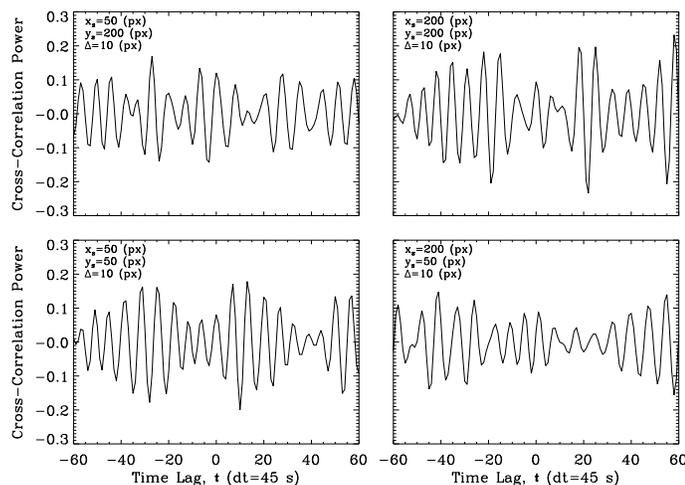
To represent the data in another way, we can “stack” the point-to-point cross-correlation results of each source point for a given receiver. As mentioned above, the cross-correlations can also be averaged over various geometries (e.g., center-to-annulus) to improve the signal-to-noise ratio and return results with specific directional information. Thus, as a clarifying example, we will briefly examine results of another geometry that emphasizes waves that propagate left-right (or east-west). Instead of averaging over annuli, we average over the north-south direction such that $\bar{y}_r = \bar{y}_s$, making r effectively one dimensional in the x -direction, and $\Delta = |r_r - r_s| = |x_r - x_s|$. The resulting averaged cross-correlation is a function of $r_s = x_s$, $r_r = x_r$, and t . We note that we will be exploring this geometry in much greater detail later in Section 3.2.3.

Stacking the cross-correlations of this east-west geometry for a particular r_r produces a TD plot: an array of t versus \mathbf{r}_s , such as those in Figure 3.9. Furthermore, we can average the cross-correlations pertaining to equivalent Δ , regardless of \mathbf{r}_s , so that we have averaged TD plots representative of equal distances between points. In practice, this requires a “shifting” of the TD results with respect to \mathbf{r}_s such that the distances to \mathbf{r}_r are “aligned” in terms of Δ in the TD plot. Visually, this equates to shifting the TD data from different sources, like in Figure 3.9 so that the vertical dashed lines are positioned at $\Delta = 0$, as in Figure 3.10. The averaged shifted TD plot, t versus Δ , produces a map with a higher signal-to-noise ratio, as in Figure 3.11, which shows a much clearer signal than the unaveraged TD plots in Figures 3.9 and 3.10. We note here that although the signal-to-noise ratio is improved, some of the spatial information is lost in taking these shifted averages. This is detrimental to studying the difference in phase correlations between explicit source points.

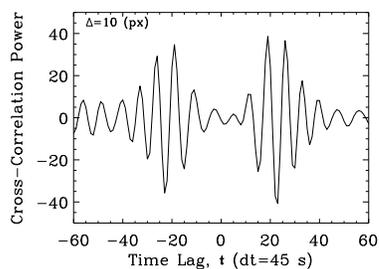
In any TD plot, the slope of the branch envelope indicates the group velocity of the wave packet. Similarly, the phase speed pertaining to each wavenumber is measured by the slope of the strong black or white signals within a branch. Airy wave theory (Section 1.2) tells us that for f -mode waves, the phase speed is equivalent to the dispersion relation divided by the wavenumber, so from Equation 1.1 the lines of constant phase in these TD plots should act as $\sqrt{g/k}$. To further improve the signal-to-noise ratio, we can make another shift in the TD plots along the lines of constant phase and then average across Δ , as illustrated in Figure 3.12 (Hindman *et al.*, 2004, Figure 7).



(a) A filtered f -mode dopplergram image indicating the location of four pairs of source (cyan) and receiver (yellow) pixels. The source and receiver points are separated by $\Delta = 10$ pixels in the x direction.



(b) The cross-correlation plots for each of the four point pairs in (a). The plot arrangement corresponds to the point locations on the image. There is no distinct cross-correlation pattern across time in these plots. Lag has units of seconds ($dt = 45$ s).



(c) The average over the cross-correlations of *all* the possible source pixels in (a), with $\Delta = 10$ pixels in the x direction. It is clear that the averaging improves signal-to-noise and reveals an ingoing and outgoing wave signal. Lag has units of seconds ($dt = 45$ s).

Figure 3.8: Example cross-correlations using the f -mode filtered data from Figure 3.7c.

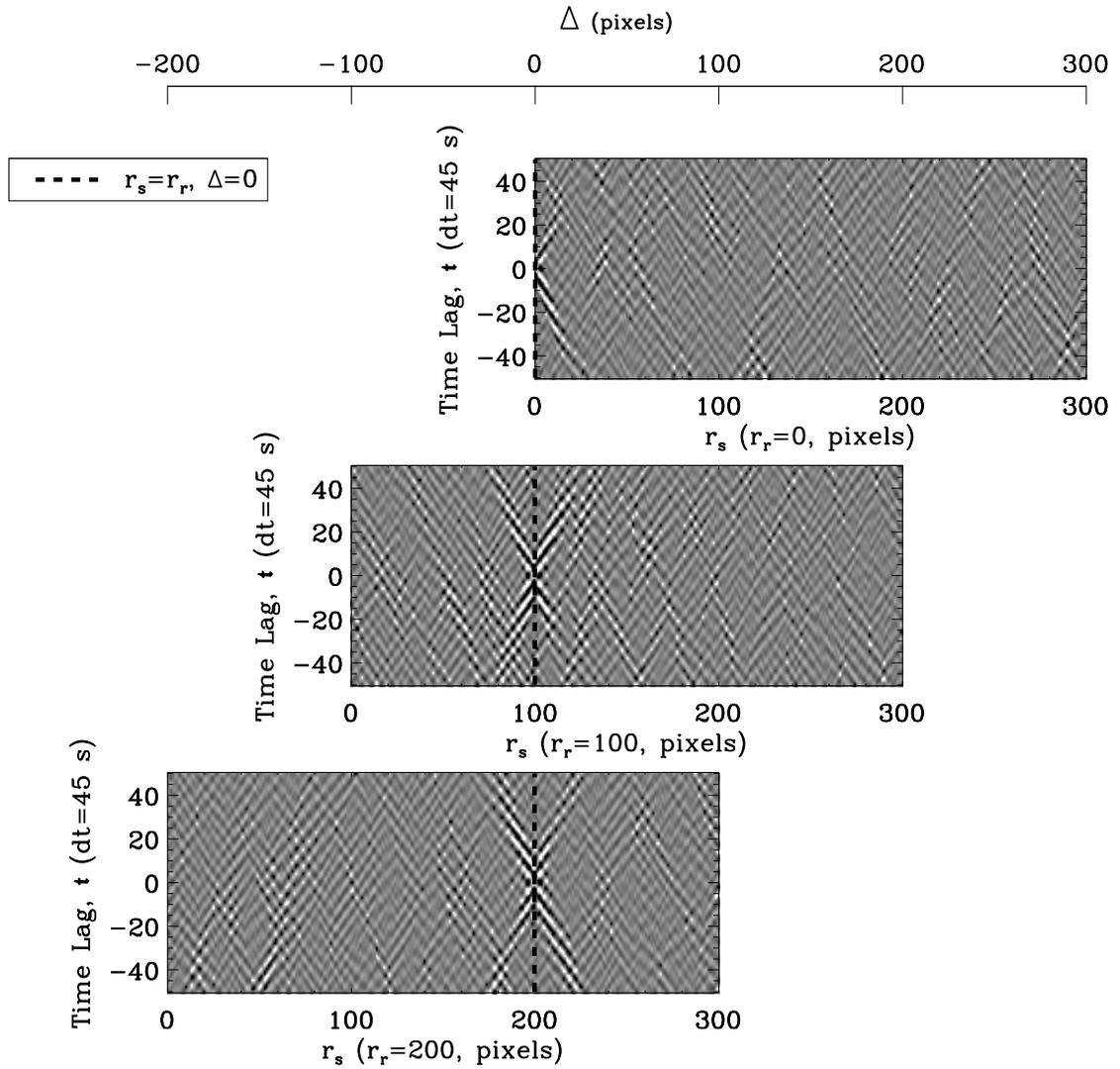


Figure 3.9: TD plots of three sample source-point locations, showing unshifted, unaveraged data. The time lag has units of seconds ($dt = 45$ s), while \mathbf{r} and Δ are in pixel units. As expected, the cross-correlations are strongest at $t = 0$ for $\mathbf{r}_s = \mathbf{r}_r$ ($\Delta = 0$), where the source is correlating perfectly with itself. These locations are represented by the vertical dashed lines. The plots are “shifted” such that the dashed lines coincide with $\Delta = 0$ on the axis above the plots. This is representative of the TD shifting that is done before the cross-correlation results can be averaged to improve the signal-to-noise ratio. Note that some data manipulation was done to produce these images, in order to clarify the cross-correlation signals within the background noise, but this enhancement is not performed in the actual data analysis.

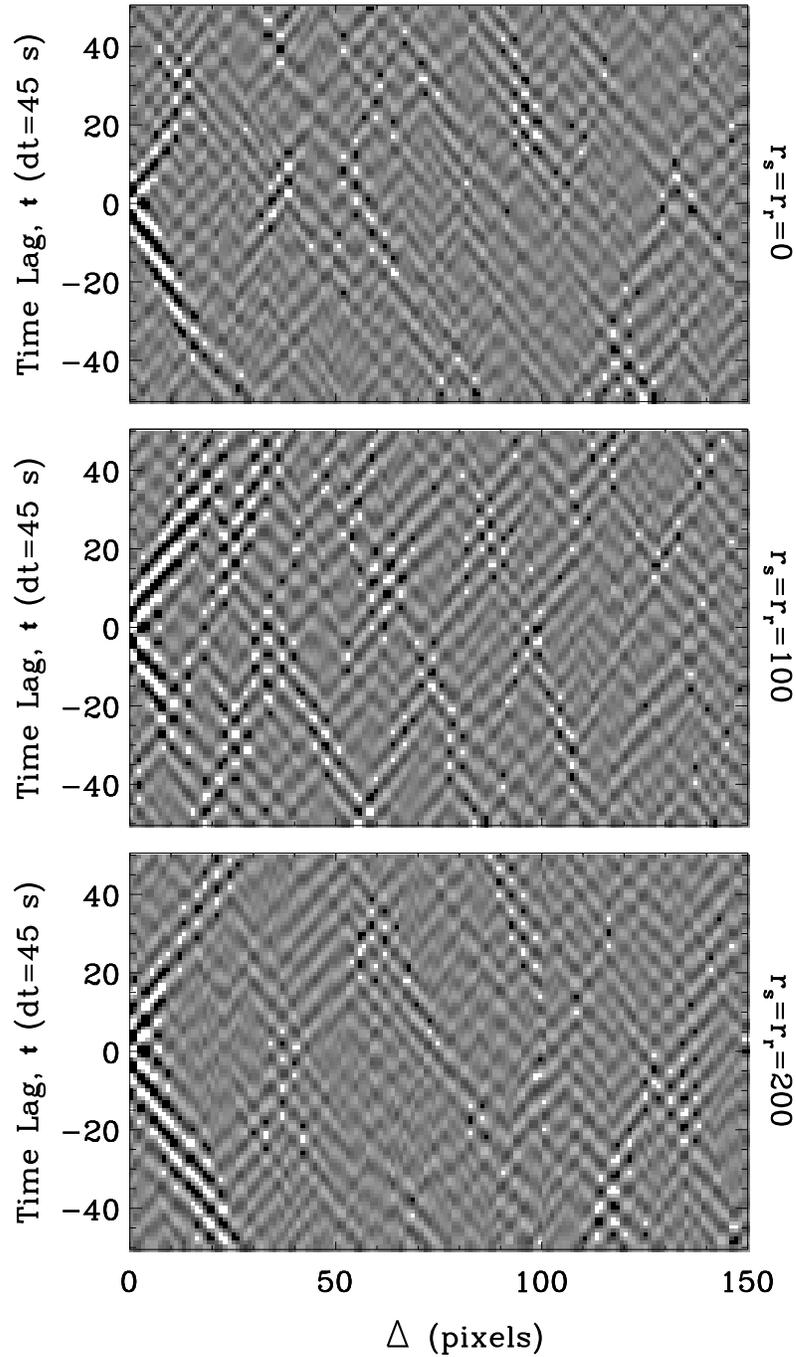


Figure 3.10: The same three TD plots from Figure 3.9, now shifted in terms of $\Delta = 0$. Again, time lag has units of seconds ($dt = 45$ s), while Δ is in pixel units. The cross-correlation signals are visibly propagating from $[\Delta, t] = [0, 0]$. These data can now be averaged (along with the shifted TD plots from all other instances of \mathbf{r}_r), the result of which is shown in Figure 3.11. We note that other than our indication of the corresponding \mathbf{r}_s and \mathbf{r}_r , the shift loses all relevant location information about the associated source points.

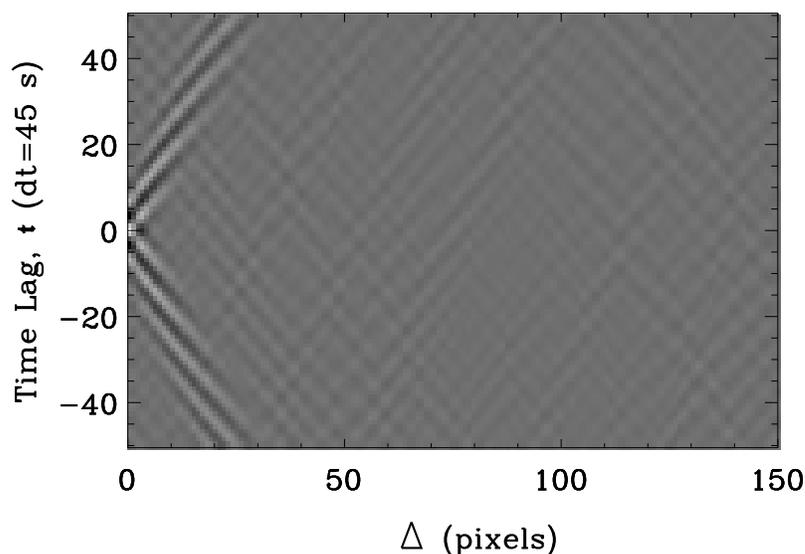


Figure 3.11: An example TD plot, averaged over the shifted data from all source points. Again, time lag has units of seconds ($dt = 45$ s), while Δ is in pixel units. The signal-to-noise ratio is much larger in this plot than in the unshifted, unaveraged data in the Figures 3.9 and 3.10 plots, even despite the signal enhancement performed in producing those images (which was not used here).

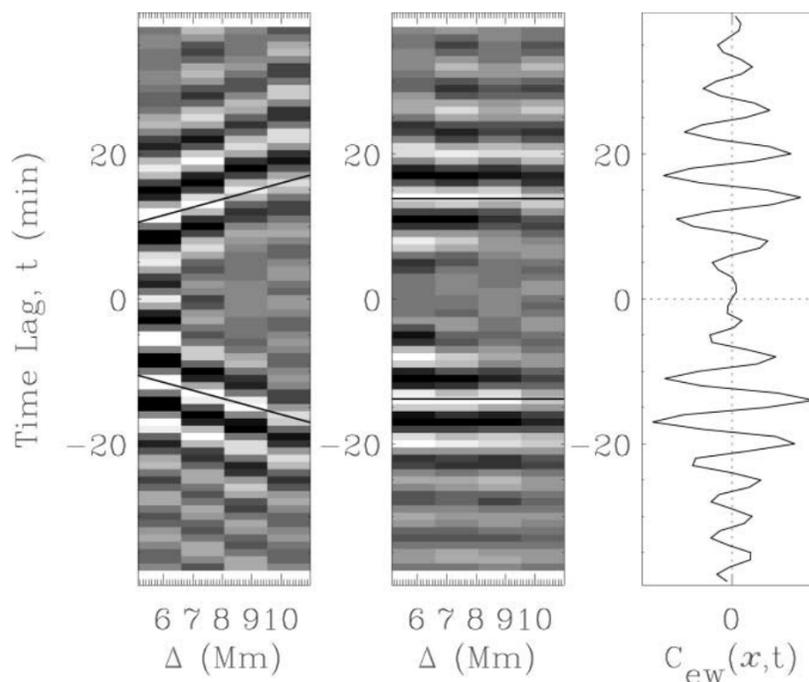


Figure 3.12: An example TD plot showing a line of constant phase (left), shifted with respect to the constant phase (middle), and then averaged over Δ (right). (Hindman *et al.*, 2004, Figure 7)

3.2.2.2 Travel Times

Perturbations alter wave travel times expected in an homogeneous domain. Measuring the travel times of waves through an inhomogeneity can tell us about the perturbation. For example, flow perturbations will cause the positive τ_+ and negative τ_- travel times to differ, since waves travel faster in the direction of the flow and are inhibited when traveling against it. Therefore, a difference between positive and negative travel times $\tau_{\text{diff}} = \tau_+ - \tau_-$ indicates that a flow perturbation is present. Similarly, the average travel time $\tau_{\text{mean}} = (\tau_+ + \tau_-)/2$ is indicative of wave-speed perturbations, such as temperature variations (Gizon and Birch, 2005).

There are multiple methods by which to calculate travel times (Gizon and Birch, 2005), including fitting Gaussian wavelets (Duvall, Jr. *et al.*, 1997; Kosovichev and Duvall Jr., 1997), and fitting model cross-covariance functions (Gizon and Birch, 2002, 2004). For this dissertation, we use the travel time definition from Gizon and Birch (2004), which minimizes the the difference between the observed cross-covariance function and a reference cross-covariance function. There are several benefits for using this method. First, it is highly robust with respect to noise. Second, the function is linear in terms of the cross-covariance, which means that averaging over different travel times is equivalent to calculating the travel time for averaged cross-covariances. Last, the probability density functions of the positive and negative travel times are unimodal Gaussian curves, which allows us to calculate errors for the travel times.

Because we will ultimately be working in one spatial dimension (see Sections 2.1.5 and 3.2.3.2), we can use a point-to-point formulation of the travel time calculations; we note that the center-to-annulus formalism is similarly defined by Gizon and Birch (2004). The mathematical definition of the point-to-point travel times is

$$\tau_{\pm}(\mathbf{x}_1, \mathbf{x}_2) = h \sum_{\mathbf{t}} W_{\pm}(\mathbf{\Delta}, \mathbf{t}) [C(\mathbf{x}_1, \mathbf{x}_2, \mathbf{t}) - C^0(\mathbf{\Delta}, \mathbf{t})], \quad (3.23)$$

where $W_{\pm}(\mathbf{\Delta}, t)$ is a weighting function given by

$$W_{\pm}(\mathbf{\Delta}, t) = \frac{\mp f(\pm t) \partial_t C^0(\mathbf{\Delta}, t)}{h \sum_{t'} f(\pm t') [\partial_{t'} C^0(\mathbf{\Delta}, t')]^2}. \quad (3.24)$$

C^0 is a smooth reference cross-covariance function, h is the temporal sampling rate, and $\mathbf{\Delta} = \mathbf{x}_2 - \mathbf{x}_1$. Figure 3.13 illustrates the procedure for determining τ_{\pm} , starting from an f -mode power spectrum, then determining $C^0(\mathbf{\Delta}, t)$, and finally minimizing the difference between C^0 and one instance of the point-to-point cross-covariance function $C(\mathbf{x}_1, \mathbf{x}_2, t)$. The function $f(\pm t)$ is a window function that isolates the positive and negative parts of the cross-covariance. The interval is selected to coincide with the arrival time of the wave packet. When measuring positive travel times, $f(t) = 0$ when $t < 0$; similarly for negative travel times when $t > 0$. For f -mode waves, the time-distance maps only have one group phase branch (as seen in Figure 3.11), so $f(\pm t)$ can be taken as a Heaviside step function, shown in the middle panel of Figure 3.13 (Gizon and Birch, 2004). The Heaviside step function for positive travel times can be represented as the discrete piecewise function with

$$f(+t) = \begin{cases} 0 & t < 0 \\ \frac{1}{2} & t = 0, \\ 1 & t > 0 \end{cases} \quad (3.25)$$

and similarly for negative travel times.

The travel times defined here are essentially minimizations of the “badness of fit” between C and C^0 . If a uniform flow rotates in the $+x$ direction, then for waves traveling from x_1 to x_2 , C will occur earlier in time than C^0 because waves travel faster with the flow than against it. Therefore, in order to minimize the difference between C and C^0 , C^0 must be shifted backwards in time so that the minimum occurs at a *negative* time τ_+ . Similarly,

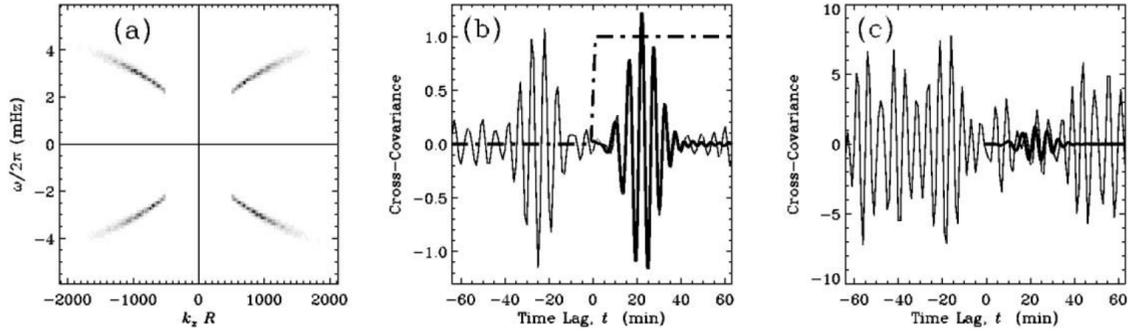


Figure 3.13: The f -mode power and cross-covariance functions. (a) A cut through of the filtered power spectrum at $k_y = 0$. (b) The spatially averaged cross-covariance function (thin line) for some point-to-point distance Δ . The thick dot-dashed line shows the window function $f(t)$ used in the travel-time measurement procedure. The thick line shows the reference cross-covariance function, C^0 . (c) A sample cross-covariance between a single pair of points that are separated by Δ (thin line). Overplotted is C^0 (thick line). Note the scale difference between (b) and (c) (Gizon and Birch, 2004, Figure 1)

we would expect τ_- to have positive values in the presence of this flow. If the flow were moving in the opposite direction, the reverse would be true with positive τ_+ and negative τ_- .

3.2.3 Symmetry Arguments

This section is the start of our divergence from “usual” TD helioseismology, described in Section 3.2.2.1. We have argued in Section 3.1 that the scattered wavefield from a linear perturbation can be approximated discretely as an integral over points along the boundary. This is analogous to the superposition of scattered wavefields from the individual points. In Section 2.1.5, we suggested that we can exploit the axis of symmetry of a 1D scattering region to (1) improve the signal-to-noise ratio of the TD results, and (2) greatly decrease the number of calculations required and therefore the computing time of the TD analysis. In this section, we present the underlying mathematical and geometrical symmetry arguments necessary for our new TD technique, as well as a case of precedence for making these assumptions.

3.2.3.1 Mathematical Symmetry

We want to measure the travel times of a wavefield from an elongated perturbation that produces a quasi-linear scattering signal. We argue that it is possible to assume a 1D system by averaging the signal along the line of symmetry, to simplify calculating the necessary cross-correlations while isolating the wave signal traveling perpendicular to the approximately-linear boundary. To do so, certain properties of the averaged cross-covariance function must be proven.

Figure 3.14 shows a simplified example considering two source point observations, $\phi(\mathbf{r}_{s1}, t)$ and $\phi(\mathbf{r}_{s2}, t)$, and the signal at a single receiver point, $\phi(\mathbf{r}_r, t)$. Ordinarily, the cross-covariances would be calculated using Equation 3.22 between each source and the receiver, separately, as

$$C_1 = C(\mathbf{\Delta}_1, \mathbf{t}) = \frac{1}{T} \int_0^T \phi(\mathbf{r}_{s1}, t) \phi(\mathbf{r}_r, t + \mathbf{t}) dt, \quad (3.26)$$

$$C_2 = C(\mathbf{\Delta}_2, \mathbf{t}) = \frac{1}{T} \int_0^T \phi(\mathbf{r}_{s2}, t) \phi(\mathbf{r}_r, t + \mathbf{t}) dt. \quad (3.27)$$

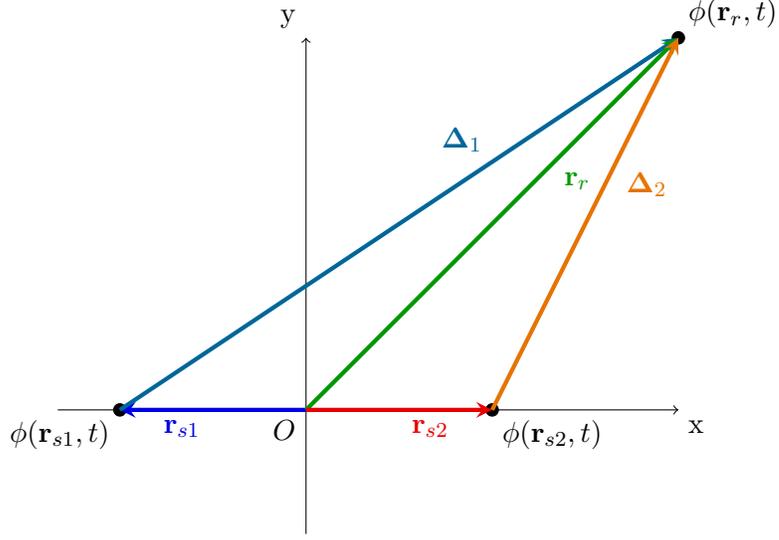


Figure 3.14: We consider two source signals $\phi(\mathbf{r}_{s1}, t)$ and $\phi(\mathbf{r}_{s2}, t)$ located at $\mathbf{r}_{s1} = [x_{s1}, y_{s1}]$ and $\mathbf{r}_{s2} = [x_{s2}, y_{s2}]$. The observed signal $\phi(\mathbf{r}_r, t)$ is made at a reference point a distance $\mathbf{r}_r = [x_r, y_r]$ from the origin O . With respect to the source points, P is located at $\Delta_1 = \mathbf{r}_r - \mathbf{r}_{s1}$ and $\Delta_2 = \mathbf{r}_r - \mathbf{r}_{s2}$, respectively. Note that the source and receiver points are assumed to be in the same z -plane. Also note that a wavefield is not shown, for clarity, but that perturbations could be located anywhere in the domain. The scatterers do not have to be located at the considered source points.

The cross-covariances C_1 and C_2 then might be averaged (along with other C_n) in order to increase the signal-to-noise ratio, as discussed in Section 3.2.2. What we aim to prove is that the cross-covariance of the averaged signals produces the same result, and how averaging over a linear geometry of signals emphasizes wave propagation in specific directions.

Firstly, we point out that the cross-covariance is a linear function. Using the two sources and receiver in Figure 3.14 and Equations 3.26-3.27, we show that

$$\begin{aligned}
\bar{C}_{1,2} &= \frac{1}{2} [C_1 + C_2] \\
&= \frac{1}{2T} \left[\int_0^T \phi(\mathbf{r}_{s2}, t) \phi(\mathbf{r}_r, t + \mathbf{t}) dt + \int_0^T \phi(\mathbf{r}_{s1}, t) \phi(\mathbf{r}_r, t + \mathbf{t}) dt \right] \\
&= \frac{1}{2T} \int_0^T [\phi(\mathbf{r}_{s2}, t) \phi(\mathbf{r}_r, t + \mathbf{t}) + \phi(\mathbf{r}_{s1}, t) \phi(\mathbf{r}_r, t + \mathbf{t})] dt
\end{aligned}$$

$$\begin{aligned}
&= \frac{1}{T} \int_0^T \frac{1}{2} [\phi(\mathbf{r}_{s2}, t) + \phi(\mathbf{r}_{s1}, t)] \phi(\mathbf{r}_r, t + \mathbf{t}) dt \\
&= \frac{1}{T} \int_0^T \bar{\phi}_{1,2}(\mathbf{r}_s, t) \phi(\mathbf{r}_r, t + \mathbf{t}) dt.
\end{aligned} \tag{3.28}$$

It follows that this is mathematically valid for any number N of the averaged cross-covariance function \bar{C}_N and averaged source signals $\bar{\phi}_N$, where $\bar{\phi}_N = \sum_{i=1}^N \phi(\mathbf{r}_{si}, t)/N$. The same logic similarly applies if we also consider multiple receiver points. As mentioned in Section 2.1.5, one could similarly choose to sum instead of average over the data. This clearly would not effect the conclusions drawn in this section, since the normalizing factor of the average is a constant which does not change the linearity of the problem.

This property of the cross-covariance can also be proven in the Fourier domain. The Fourier transform is a linear process, and the convolution between two functions in Fourier space is simply multiplicative between the transformed functions. Therefore, the transform of the sum over multiple signals is the same as the sum over the transformed signals.

Next, we investigate the effect of averaging over a line of signals in one dimension. Again starting from Figure 3.14, we look at how averaging between \mathbf{r}_{s1} and \mathbf{r}_{s2} affects Δ in the resultant averaged cross-covariance. As stated in the figure caption, we define the position vectors

$$\begin{aligned}
\mathbf{r}_{s1} &= [x_{s1}, y_{s1}], & \mathbf{r}_{s2} &= [x_{s2}, y_{s2}], & \mathbf{r}_r &= [x_r, y_r], \\
\Delta_1 &= \mathbf{r}_r - \mathbf{r}_{s1}, & \Delta_2 &= \mathbf{r}_r - \mathbf{r}_{s2}.
\end{aligned} \tag{3.29}$$

We now show that averaging over Δ_1 and Δ_2 is equivalent to the distance between the receiver's location and the averaged locations of the sources. We do so explicitly in x and y coordinates to help illustrate the directionality of the results.

$$\bar{\Delta} = \frac{1}{2} \sum_{i=1}^2 \Delta_i = \frac{1}{2} [(\mathbf{r}_r - \mathbf{r}_{s1}) + (\mathbf{r}_r - \mathbf{r}_{s2})]$$

$$\begin{aligned}
&= \frac{1}{2} [([x_r, y_r] - [x_{s1}, y_{s1}]) + ([x_r, y_r] - [x_{s2}, y_{s2}])] \\
&= \frac{1}{2} [((x_r - x_{s1}) \hat{x} + (y_r - y_{s1}) \hat{y}) + ((x_r - x_{s2}) \hat{x} + (y_r - y_{s2}) \hat{y})] \\
&= \frac{1}{2} [(x_r - x_{s1} + x_r - x_{s2}) \hat{x} + (y_r - y_{s1} + y_r - y_{s2}) \hat{y}] \\
&= \frac{1}{2} [(2x_r - (x_{s1} + x_{s2})) \hat{x} + (2y_r - (y_{s1} + y_{s2})) \hat{y}] \\
&= [x_r, y_r] - \frac{1}{2} [(x_{s1} + x_{s2}) \hat{x} + (y_{s1} + y_{s2}) \hat{y}] \tag{3.30} \\
&= [x_r, y_r] - [\bar{x}_s, \bar{y}_s] \\
&= \mathbf{r}_r - \bar{\mathbf{r}}_s
\end{aligned}$$

Again, this logic expands for any number of source points (N instead of 2), and considering multiple reference points. From Equation 3.30, we can see that if the source points are symmetrically aligned along the x - (or y -) axis, as they are in Figure 3.14, the signals in that direction cancel out ($x_{s1} = -x_{s2}$) so that only the signal perpendicular to the line of symmetry is important. Of course, the receiver point can contribute directionality in all directions; however, all but the perpendicular components of the received signal similarly cancel if multiple receiver signals are averaged along the same axis of symmetry. The problem would therefore simplify down to one spatial dimension.

We have shown that (1) the average over cross-covariances between signals is equivalent to the cross-covariance of the averaged signals and (2) averaging a line of signals preserves only the averaged signal perpendicular to the line. Together, these properties allow us to simplify the TD process for a linear perturbation down to one spatial dimension before any cross-covariance is calculated. For a data cube of size $nx \times ny \times nt$, this reduces the computational time of the point-to-point cross-covariances by a factor of ny . They also tell us that we can easily compute the unidirectional cross-correlations of a wavefield with an elongated perturbation that produces a linear signal in any direction by making use of the

axis of symmetry.

3.2.3.2 Geometric Symmetry

There is precedence for using geometric symmetry arguments in time-distance helioseismology. We briefly discussed the center-to-annulus geometry in Section 3.2.2.1. It is also a relatively common practice to find the center-to-annulus cross-correlations separately for different quadrants of the annuli (north, south, east, west), on top of averaging over each annulus as a whole. This geometry is exemplified in Figure 3.15 (Hindman *et al.*, 2004, Figure 6). Not only does this tactic give more information about directionality of flows, but it also helps to improve the signal-to-noise ratio (Duvall, Jr. *et al.*, 1996, 1997; Gizon and Birch, 2005; Hindman *et al.*, 2004). For example, eastward flows are an average of the positive lags from the eastern quadrant and the negative lags from the western quadrant, which doubles the sample size and improves the signal-to-noise ratio. Similar is true for the flows towards the other quadrants. The standard full-annulus cross-correlation represents radially inward/outward flows to/from the central pixel.

We point out that our summation along the north-south axis of symmetry of our linear feature (described in Section 2.1.5) effectively performs the same function as isolating the east-west directed waves. All wave propagation that is not parallel to the x -axis is essentially averaged out of the resultant 1D data.

Additionally, Cameron, Gizon, and Duvall, Jr. (2008) developed a time-distance method that makes use of the rotational symmetry of a nearly-circular sunspot region. They used SOHO/MDI data of AR 9787, as well as simulated sunspot data, to observe how f -mode waves respond to propagating through strong magnetic regions. They describe their cross-covariance technique (we note that the notation used below corresponds to their paper, and not necessarily the notation used in this thesis):

Let us denote by $\phi(\mathbf{r}, t)$ the filtered Doppler velocity, where \mathbf{r} is a position vector and t is time. Rather than studying the cross-covariance of ϕ between

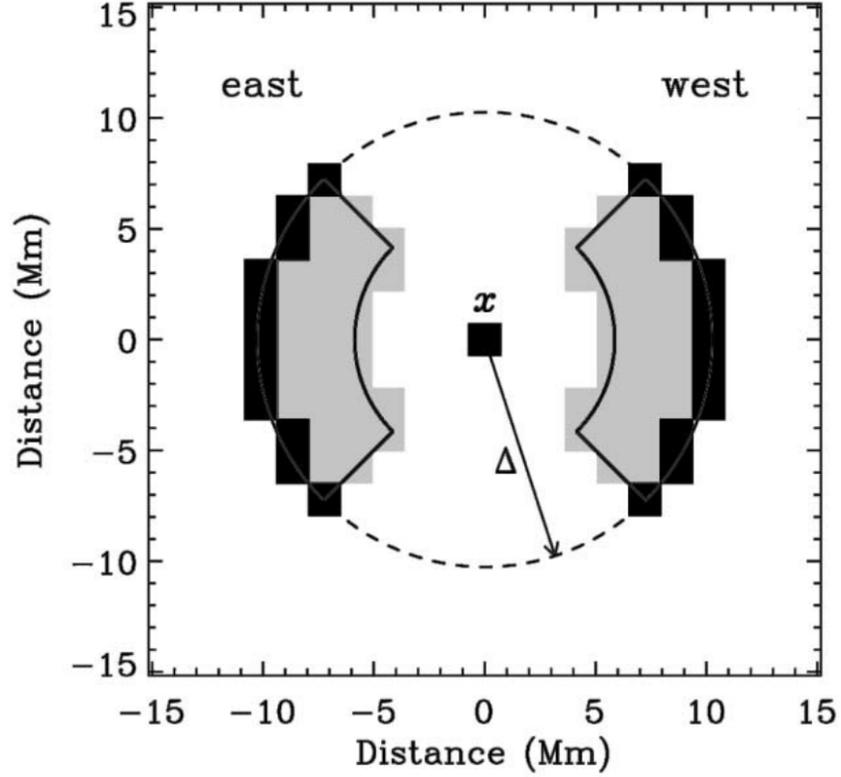


Figure 3.15: An example of the quadrant geometry used in TD helioseismology to measure east-west flows. The black regions show pixels in the east-west quadrant arcs that are a distance Δ from the central pixel x . The gray regions cover the pixels that would be used for three separate smaller Δ . (Hindman *et al.*, 2004, Figure 6)

two spatial points, we consider the cross-covariance between ϕ averaged over a great circle γ , denoted by $\bar{\phi}(\gamma, t)$, and ϕ measured at any other point \mathbf{r} :

$$C(\mathbf{r}, t) = \int_0^T \bar{\phi}(\gamma, t') \phi(\mathbf{r}, t + t') dt'$$

where the effective integration time, T , is nine days, or 36 times six hours. Since averaging ϕ over γ is equivalent to selecting only the horizontal wavevectors that are perpendicular to γ , this cross-covariance function gives us information about plane wave packets that propagate away from γ . In simple words, the

cross-covariance at time lag t tells us about the position of a wave packet, a time t after it has left γ . In the rest of the paper we fix the distance between γ and the center of the sunspot at $\Delta = 40$ Mm. Since $\Delta \gg \lambda$, where $\lambda \approx 5$ Mm is the dominant wavelength, the sunspot is in the far field of γ .

For any particular choice of orientation of γ , the computed cross-covariance is very noisy, thus explaining the need for some spatial averaging.... Because sunspot AR 9787 is almost rotationally invariant around its center, we can compute many equivalent cross-covariance functions corresponding to many different directions of the incoming wave packet, derotate these about the center of the sunspot so that they match the reference cross-covariance, and average them to reduce the noise. We have performed this averaging over all directions of the incoming wavevectors, with a fine sampling of 1° this enables us to reach a very good signal-to-noise level. (Cameron, Gizon, and Duvall, Jr., 2008, p.294-295)

The process they detail is similar to our own method, except that they are working with a rotationally invariant case whereas ours is treated as translationally invariant. Where we sum along the y -axis of symmetry to isolate east-west flows, they average the derotated cross-covariances to isolate radial inward/outward flows. By the logic presented in Section 3.2.3.1, the linearity of the cross-covariance function equates these procedures. This gives us excellent precedence for accepting our technique as reasonable.

Chapter 4: Wave Scattering Simulation & Results

4.1 Simulation

The motivation for producing simulated scattered wave data is to create a proof-of-concept model on which to test helioseismic results using known inputs. It is important to understand that we are not over-concerned with incorporating detailed solar physics, such as magnetic fields, into this simulation. We simply want to create surface wave interactions to determine how well our local helioseismology technique can retrieve the scattering information.

All of the data in this chapter were created using IDL code we wrote, called *cyl_wav.pro*. All of the calculations are performed in the complex domain, but the images show only the real part of the data, as these are the “observable” results.

Our simulation is analytic rather than numerical. The grid resolution is set to 128×128 pixels. There is no need for very-high resolution simulated data; it only serves to slow down calculations. The spatial resolution is 0.729 Mm/pixel, to mimic the HMI data. The spatial resolution ($0.729 \text{ Mm} \rightarrow \ell \simeq 6007$) and domain size ($\sim 93 \text{ Mm} \rightarrow \ell \simeq 46$) limit the need for all values of ℓ ; however, our region of interest between $600 \leq \ell \leq 1200$ ($7.29 \text{ Mm} \geq \lambda \geq 3.65 \text{ Mm}$) falls well within the maximum and minimum limiting ℓ values. We can also determine some reasonable spacing $\Delta\ell$ such that only distinct modes within the domain are used, which reduces computational time. We use $\Delta\ell = 30$, reducing the number of simulated wave modes to 21 (although the spacing could be much higher than this). The temporal cadence is set to 45 seconds, again to mimic the HMI data.

We neglect multiple scattering effects (discussed in Section 3.1.2) because the simulation is in the far-field regime and assumes the long-wavelength approximation. We can make these assumptions based on our choice of scatterer-radius a . If $a \leq 0.5$ pixels (i.e. the

diameter of each scatterer is less than a pixel), the actual size of a becomes arbitrary, in which case we can choose it to be arbitrarily small. This ensures both the long-wavelength and far-field approximations. The far-field approximation alone allows us to neglect multiple scattering. Additionally, declaring $a \ll \lambda$ means that any noticeable distance d between scatterers is large enough to ignore multiple scattering as well, regardless of the far-field approximation. In the simulation code, we set $a = \text{pixsz}/10$ to guarantee our use of the far-field and long-wavelength approximations.

We note that using long wavelengths is also desirable because the waves see the “width” of a boundary as smaller, so the boundary interactions are sharper. Sharper interactions are more likely to produce a stronger coherent backscattered signal. This point is important for “future” work that is currently underway.

From the discussions in Section 3.1 and Appendix B.2, we consider N scattering points in the dependent, single scattering limit. The total scattered wave function will look like

$$\psi_s = \sum_{n=1}^N \frac{f(\varphi_n) e^{i\mathbf{k}\mathbf{r}_n \cdot \hat{\alpha}} e^{i\mathbf{k}\mathbf{b}_n \cdot \hat{\alpha}}}{\sqrt{r_n}}. \quad (4.1)$$

which is computed in the main *cyl_wav.pro* code.

The scattering amplitude in Equation 4.1 is computed in the *cyl_scattering_amplitude.pro* function within the main code. The form of the scattering amplitude equation coded into the algorithm, for scattering angle $\varphi_n = \phi_i - \phi_n$ and arbitrary m , is given by

$$f(\varphi_n) = \sqrt{\frac{2}{\pi k}} e^{-i\delta'_m} \epsilon_m i^{m-1} \sin(\delta'_m) \cos(m\varphi_n), \quad (4.2)$$

where δ'_m are the phase angle derivatives for the Bessel functions (calculated assuming the long-wavelength approximation and Neumann “rigid” boundary conditions in the function *cyl_phase_angle.pro*). Note that because the scattering amplitude dependence on m is linear, we simply need to sum the results of Equation 4.2 over various m to incorporate more than

one azimuthal mode. Although $m \rightarrow \infty$ on the Sun, we only need the contributions up to $m = 3$. Higher than that, the scattering solution converges and the higher-order m effects are negligible. This is proven in Figures 4.1 and 4.2. Figure 4.1 shows the scattering amplitudes $f(\varphi_n = 0^\circ)$ computed for each wavenumber k with $0 \leq m \leq 7$. It is clear that the solutions have converged to 0 by $m = 3$. Figure 4.2 shows the $0 \leq m \leq 3$ solutions of $f(\varphi_n)$ with different scattering angles across varying k . Again, the $m = 3$ solution has clearly converged to 0 in each angular case.

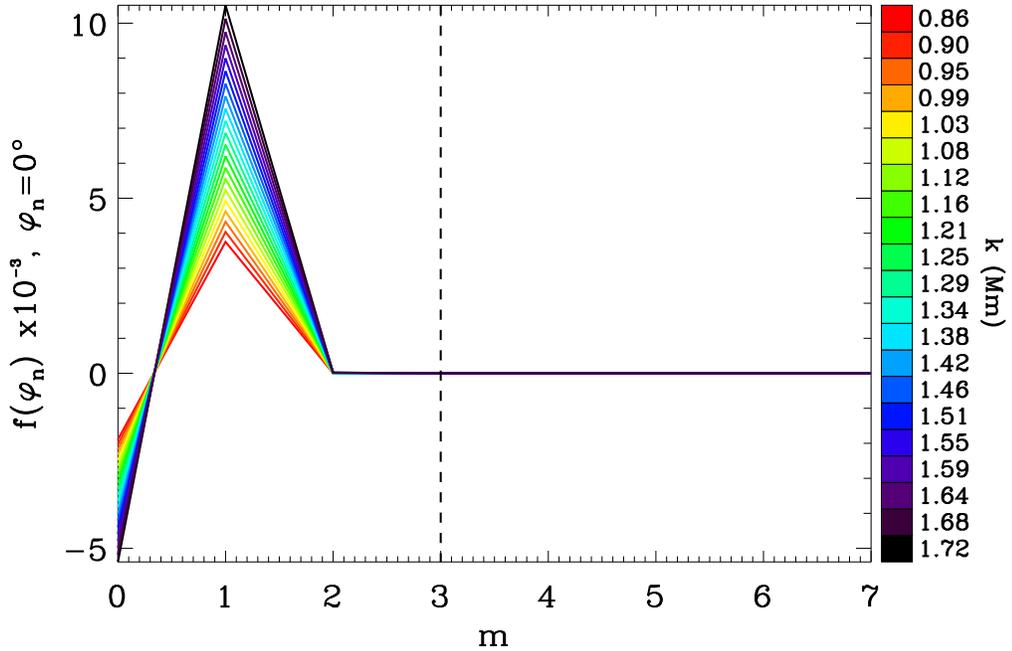


Figure 4.1: Scattering amplitudes $f(\varphi_n)$ for scattering angle $\varphi_n = 0^\circ$ computed using the simulation code for each $0 \leq m \leq 7$. The various wavenumbers k are shown ranging in color from lowest (red) to highest (black). The color bar is plotted to the right of the image. It is apparent that for all values of k , the associated $f(\varphi_n = 0^\circ)$ converge to 0 before $m = 3$.

The simulation code allows the user to change multiple parameters in Equation 4.1, including the incident angle, the number and spacing of wave modes, the phase of each wave mode, and the number and location of scattering points. We show examples of the various parameter adjustments in Figures 4.3-4.5. In each of these figures, there are three

rows corresponding to time steps $t = 0, 100, 200$ (recall, the cadence is set to 45 s). The first columns display incident plane waves at each time step. The second and third columns show the scattering responses to the incident waves at each time step for a single scatterer and 10 scatterers, respectively. At any point in time, the maximum scattered wave signal is 2-3 orders of magnitude smaller than the incident wave signal; therefore, the total wave is not shown since it is visually indistinguishable from the incident wave. Figure 4.3 displays an in-phase incident plane wave with $\phi_i = 0^\circ$ and the corresponding scattering responses. Figure 4.4 shows a change in phase with a randomly phased incident plane wave of $\phi_i = 0^\circ$ and the corresponding scattering responses. Figure 4.5 shows an adjustment to the incident angle with an in-phase incident plane wave of $\phi_i = 45^\circ$ and the corresponding scattering responses. Finally, Figure 4.6 shows similar results to the Figure 4.3 multiple scatterer results at $t = 100$, but illustrates the variable scatterer spacing in y . We note (but do not show) that the locations in x are likewise adjustable.

For the simulated data in our linear TD method, we use 10 scattering points parallel to the y -axis at $x = n_x/3$ pixels with a spacing $d = \Delta y = 5$ pixels. The incident wave packet is in-phase along the edge of the domain at $t = 0$, with an incident angle of $\phi_i = 0^\circ$, as in Figure 4.3.

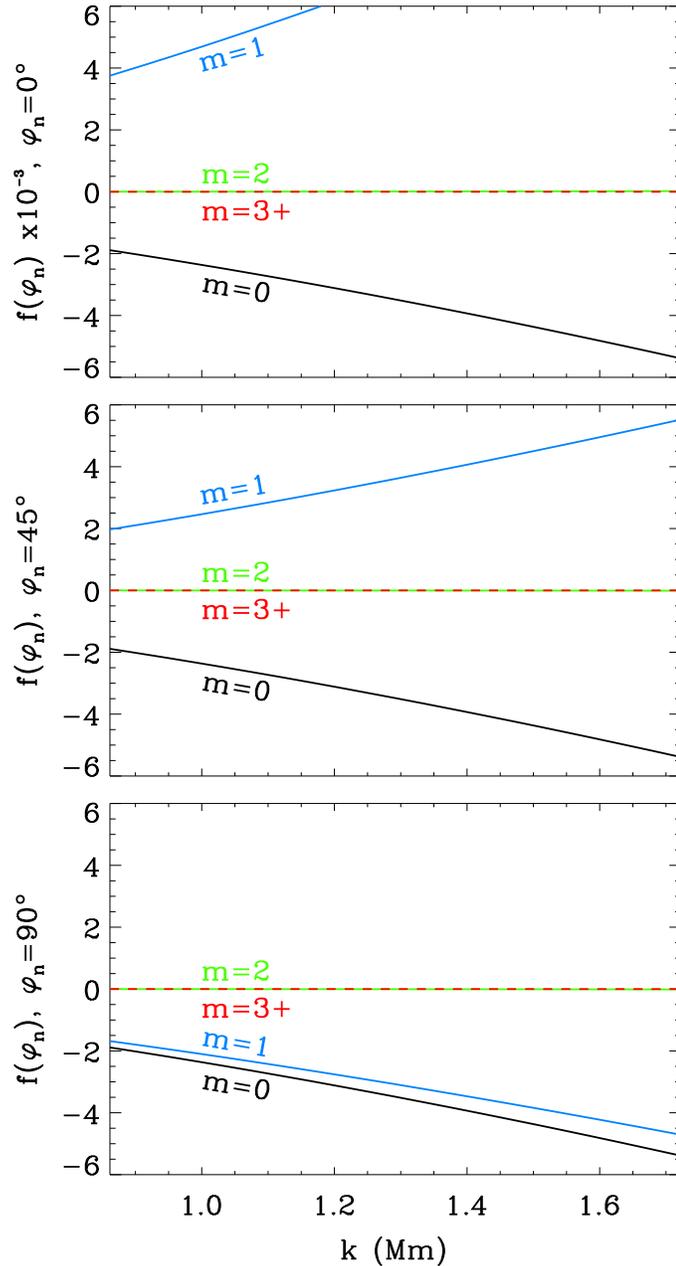


Figure 4.2: Scattering amplitudes $f(\varphi_n)$ for scattering angles $\varphi_n = 0^\circ, 45^\circ, 90^\circ$ computed using the simulation code for each wavenumber k . The solutions for azimuthal modes $m = 0$ (black), $m = 1$ (blue), $m = 2$ (green), and $m = 3$ (red, dashed) are labeled in each plot. The $m = 3$ case is also representative of the $f(\varphi_n)$ solutions for all higher order m since they similarly converge to $f(\varphi_n) = 0$. Note that the even-order azimuthal modes do not vary with φ_n , as expected from Equation 4.2. These plots, along with Figure 4.1, show that all orders of $m > 3$ can be neglected in the simulation scattering calculations, independent of k or φ_n .

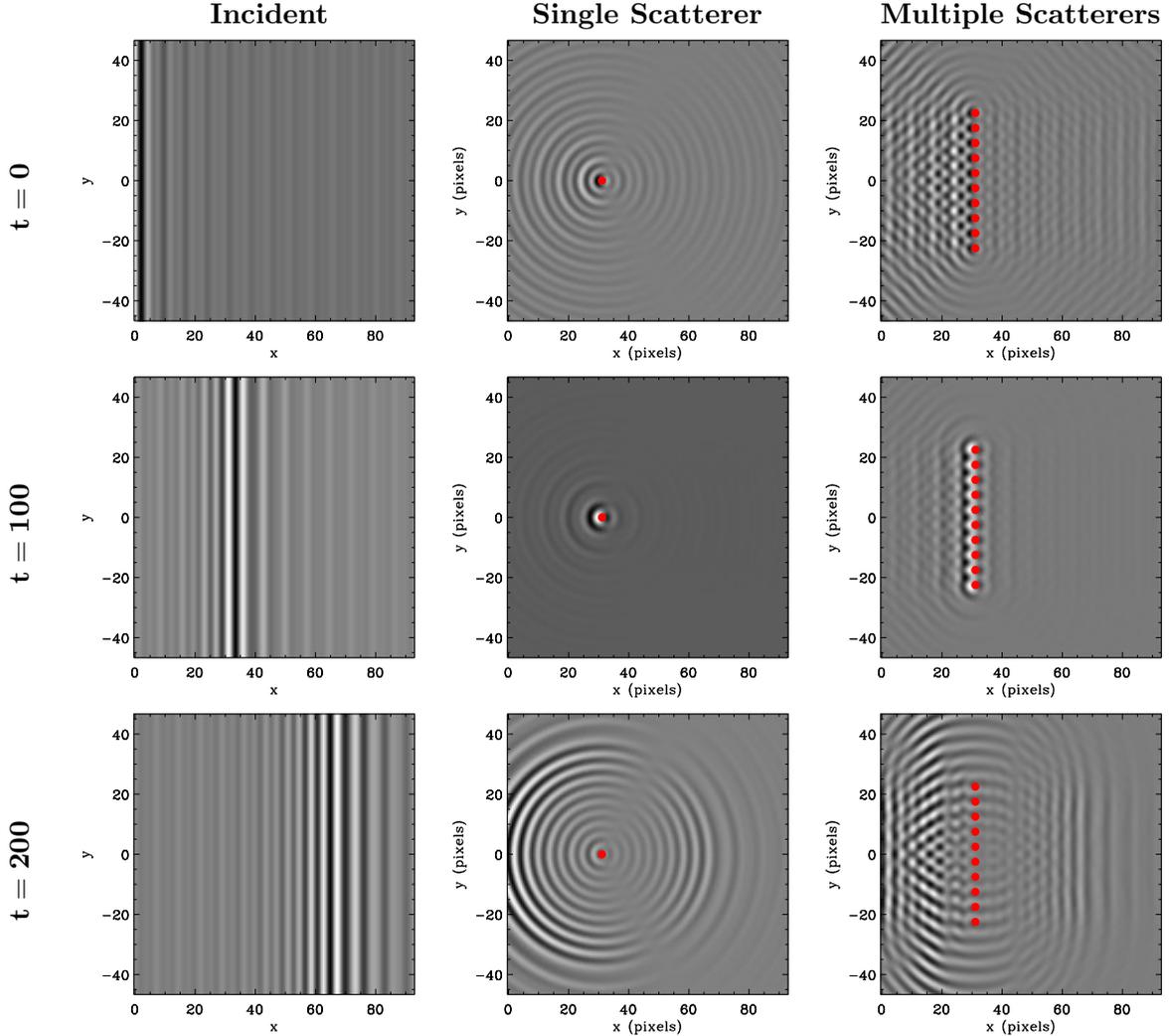


Figure 4.3: Example data produced by the scattering simulation code. Three time steps are shown at $t = 0$ (top row), $t = 100$ (middle row), and $t = 200$ (bottom row). Axes are in units of pixels. Greyscale colors indicate wave amplitude (black/white is minimum/maximum, grey is average), although the incident and scattered waves are not scaled with respect to each other since the incident wave signal is actually 2-3 orders of magnitude larger than the scattered wave signals. The incident wave (left column) is in-phase at $x = 0$ when $t = 0$, with an incident angle $\phi_i = 0^\circ$. The scattering results from a single scatterer (middle column) and 10 evenly-spaced scatterers (right column) are shown for each time step. The scatterer locations are indicated by the red dots (\bullet). For both scattering results, the scatterers are located at $x = n_x/3$. The single scatterer is placed at $y = 0$, and the multiple scatterers are spaced symmetrically around $y = 0$ with $d = \Delta y = 5$ pixels. Note that each image is scaled to itself, and that the scattered wave signals are actually much smaller than the incident signal. The total wavefields (incident + scatterer(s)) are visually indistinguishable from the incident waves shown here because the scattering signal is much weaker than the incident wave signal.

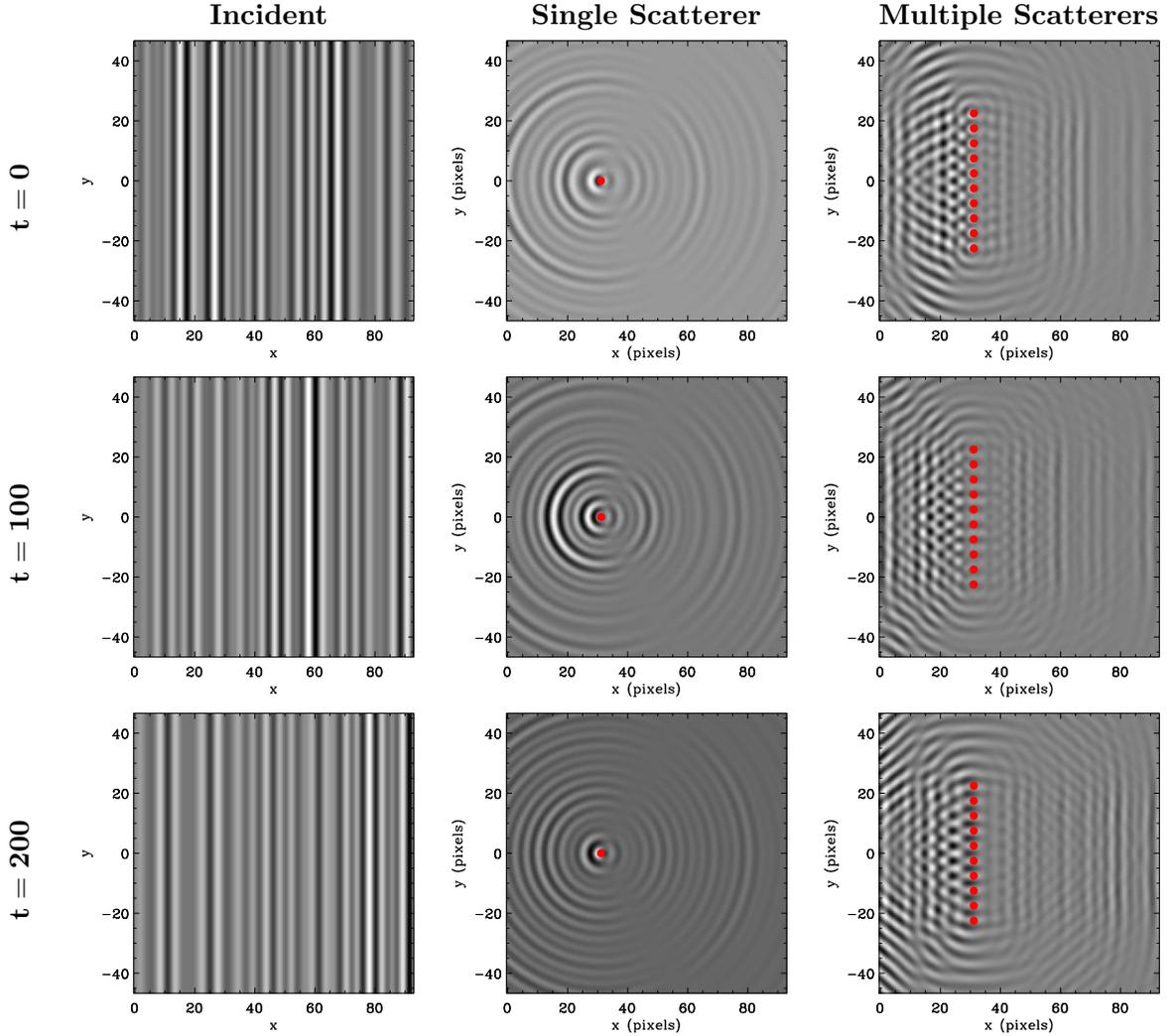


Figure 4.4: Example data produced by the scattering simulation code, displaying a change in incident phase from Figure 4.3. Three time steps are shown at $t = 0$ (top row), $t = 100$ (middle row), and $t = 200$ (bottom row). Axes are in units of pixels. Greyscale colors indicate wave amplitude (black/white is minimum/maximum, grey is average), although the incident and scattered waves are not scaled with respect to each other since the incident wave signal is actually 2-3 orders of magnitude larger than the scattered wave signals. The incident wave (left column) is randomly phased, with an incident angle $\phi_i = 0^\circ$. The scattering results from a single scatterer (middle column) and 10 evenly-spaced scatterers (right column) are shown for each time step. The scatterer locations are indicated by the red dots (\bullet). For both scattering results, the scatterers are located at $x = n_x/3$. The single scatterer is placed at $y = 0$, and the multiple scatterers are spaced symmetrically around $y = 0$ with $d = \Delta y = 5$ pixels. The total wavefields (incident + scatterer(s)) are visually indistinguishable from the incident waves shown here because the scattering signal is much weaker than the incident wave signal.

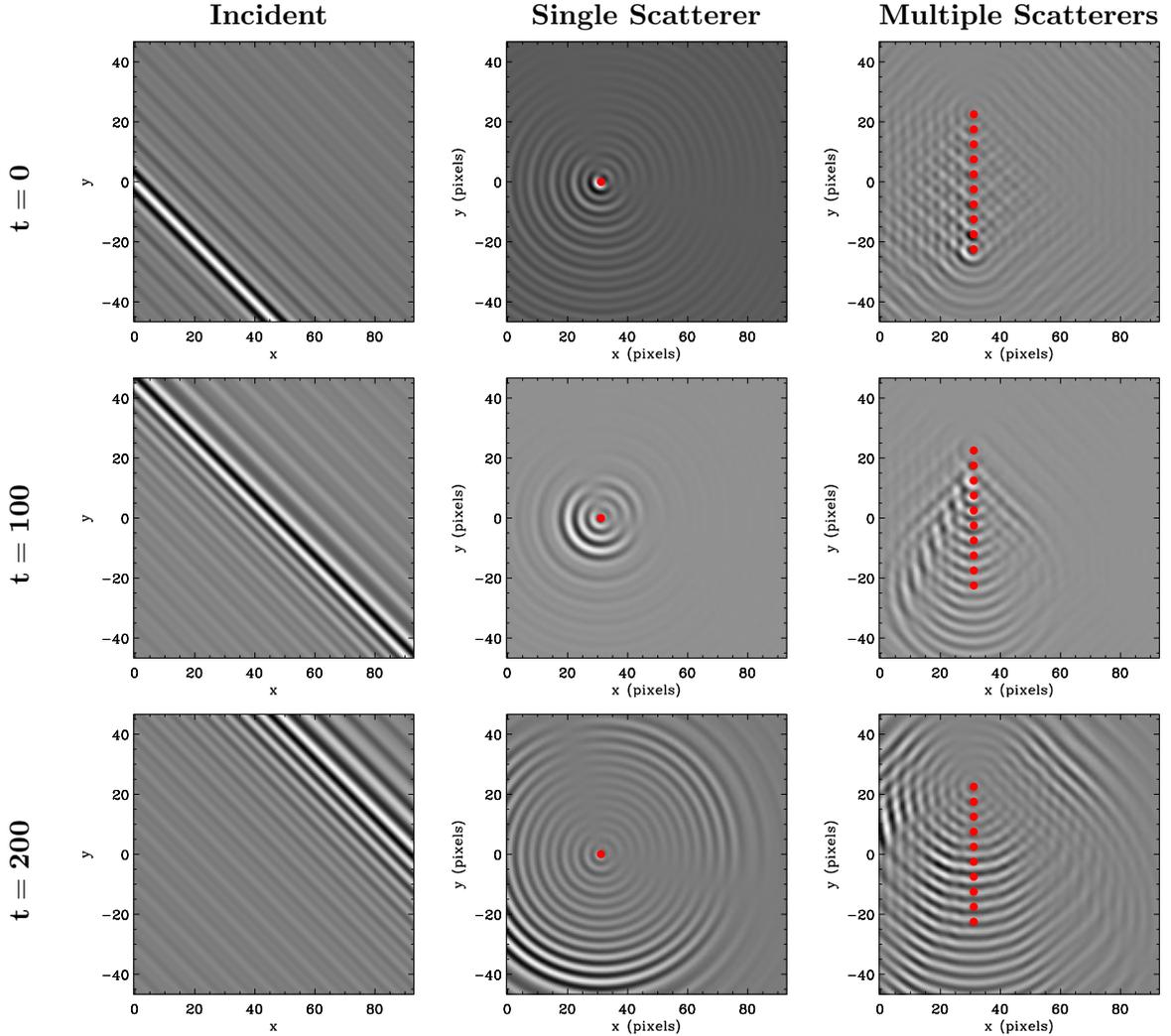


Figure 4.5: Example data produced by the scattering simulation code, displaying a change in incident angle from Figure 4.3. Three time steps are shown at $t = 0$ (top row), $t = 100$ (middle row), and $t = 200$ (bottom row). Axes are in units of pixels. Greyscale colors indicate wave amplitude (black/white is minimum/maximum, grey is average), although the incident and scattered waves are not scaled with respect to each other since the incident wave signal is actually 2-3 orders of magnitude larger than the scattered wave signals. The incident wave (left column) is in-phase at $t = 0$, with an incident angle $\phi_i = 45^\circ$. The scattering results from a single scatterer (middle column) and 10 evenly-spaced scatterers (right column) are shown for each time step. The scatterer locations are indicated by the red dots (\bullet). For both scattering results, the scatterers are located at $x = n_x/3$. The single scatterer is placed at $y = 0$, and the multiple scatterers are spaced symmetrically around $y = 0$ with $d = \Delta y = 5$ pixels. Note that each image is scaled to itself, and that the scattered wave signals are actually much smaller than the incident signal. The total wavefields (incident + scatterer(s)) are visually indistinguishable from the incident waves shown here because the scattering signal is much weaker than the incident wave signal.

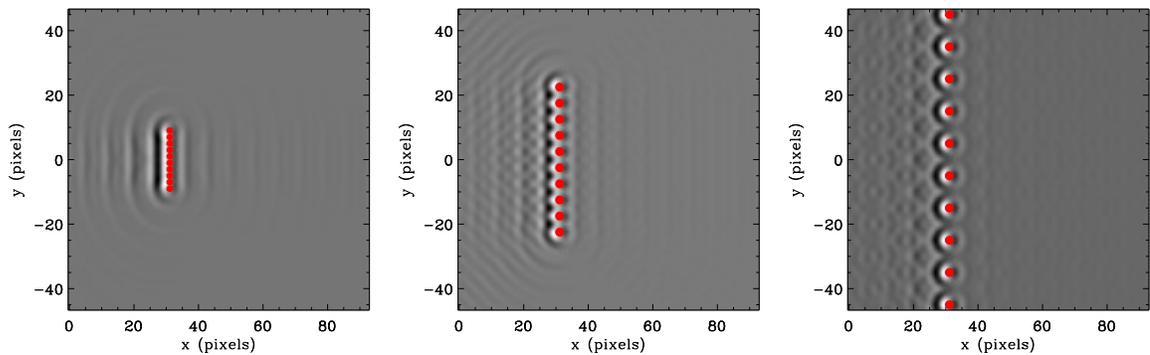


Figure 4.6: Example scattered data produced by the scattering simulation code, displaying changes in scatterer spacing, using the $t = 100$ incident conditions from Figure 4.3 (in-phase, $\phi_i = 0^\circ$). Axes are in units of pixels. Greyscale colors indicate wave amplitude (black/white is minimum/maximum, grey is average), although the three scattering cases are not scaled with respect to each other. The scatterer locations are indicated by the red dots (\bullet); however, dot size is not indicative of scatterer size. The middle plot is the same as the middle-righthand plot in Figure 4.3, for reference. The left plot shows the scattering response from 10 scatterers with $d = \Delta y = 2$ pixels; the right plot shows the 10-scatterer response for $d = \Delta y = 10$ pixels. We note that the locations in x are likewise adjustable, but not shown.

4.2 Time-Distance Analysis

4.2.1 Testing Assumptions

Before comparing simulated linear TD results with those from real data, or even before interpreting the simulated data results on their own, we need to understand whether and how various assumptions influence the analysis. If the implementation of the assumptions consequently changes the simulated results, the effects need to be recognized so that they do not confuse our interpretation of the results. In extreme cases, we may even need to change our assumptions in order to produce reasonable outcomes and conclusions. In Section 3.2.3, we mathematically proved the linear TD effects of spatially averaging data over an axis of symmetry, reducing the cross-correlation calculations to 1D. Using the simulated data, we now explore how the linear TD results differ when assuming a line of perturbations versus a single scatterer.

We compare a line of scatterers to a point scatterer when summing over the linearly symmetric dimension. Figure 4.7 shows example images for reference. As mentioned in Section 4.1, the incident wave packet is in-phase at $t = 0$ along the y -axis at $x = 0$. The waves consist of 21 evenly spaced wavelengths corresponding to $600 \leq \ell \leq 1200$. The incident waves in this example propagate in the $+x$ direction (incident angle of 0°). In the scattered wavefields, the scatterers are visibly centered at $y = ny/2$ along the $x = nx/3$ line. Multiple scattering has been ignored. The azimuthal order m terms in the expanded scattering calculation are summed over $0 \leq m \leq 3$, since any terms above $m = 3$ are negligible, as shown in Section 4.1. Across two dimensions, the 1-scatterer case shows the expected radially symmetric scattering pattern, while for the 10-scatterer case, the wave interactions from each scatterer create a quasi-linear scattering pattern.

Figure 4.8 shows plots of the scattered data in Figure 4.7, summed over the y -dimension. The magnitude ratio between the 1-scatterer and 10-scatterer cases is approximately a multiple of the number of scatterers n . This relationship is seen over all other time steps, although only the $t = 0$ time step is shown. The factor-of- n difference holds true for any

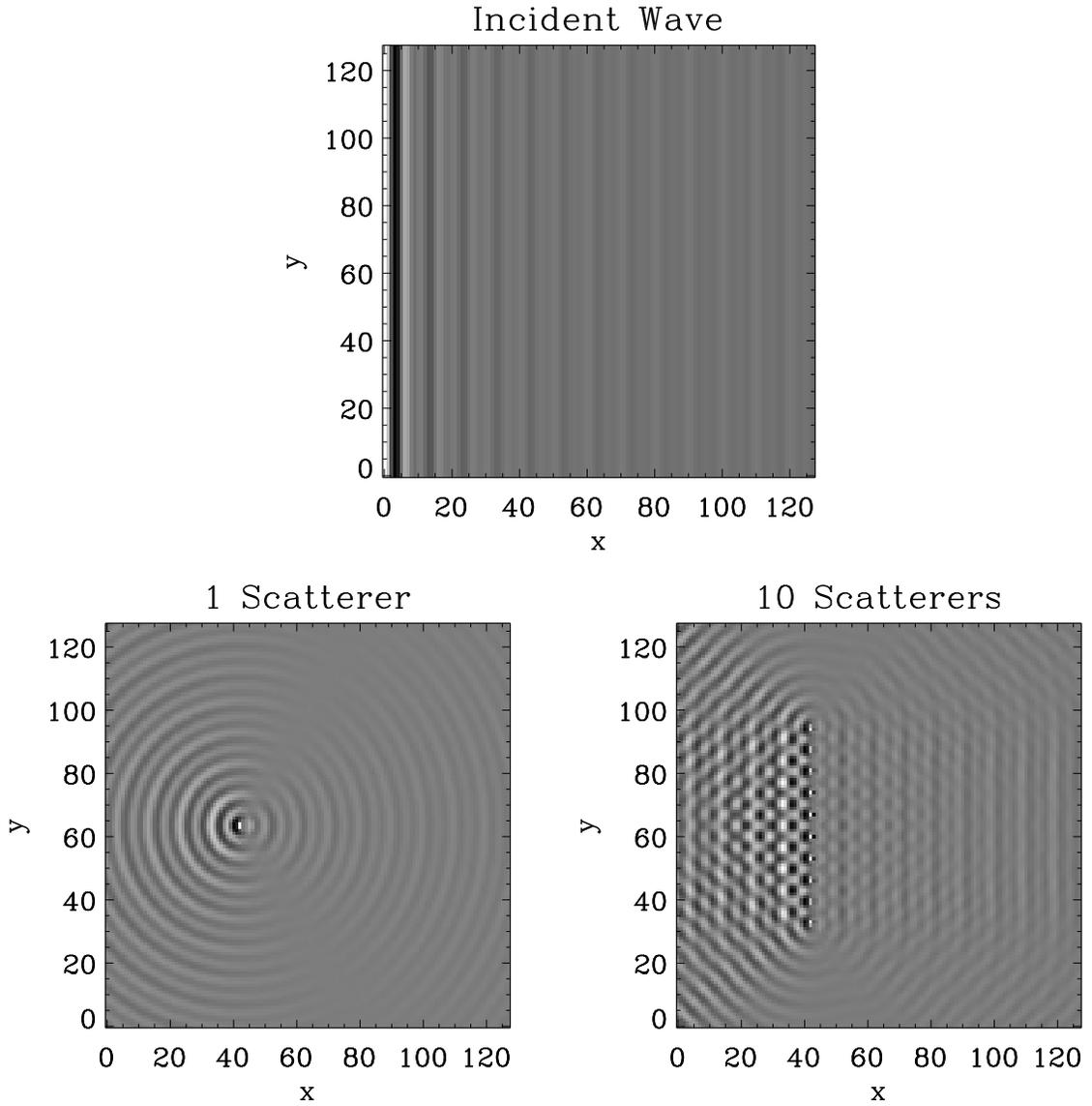


Figure 4.7: Examples of the real-part of the simulated scattered wave functions for a single scatterer (left) and a column of ten scatterers (right) at $t = 0$. Axes are in units of pixels. Greyscale colors indicate wave amplitude (black/white is minimum/maximum, grey is average), although the incident and scattered waves are not scaled with respect to each other since the incident wave signal is actually 2-3 orders of magnitude larger than the scattered wave signals. The incident wave packet (top) for these examples is initially in-phase and aligned at the $x = 0$ line. Note that these images are the same simulated data as those in the top row of Figure 4.3.

number of scatterers because of the linearity of the system. This expected result is now confirmed.

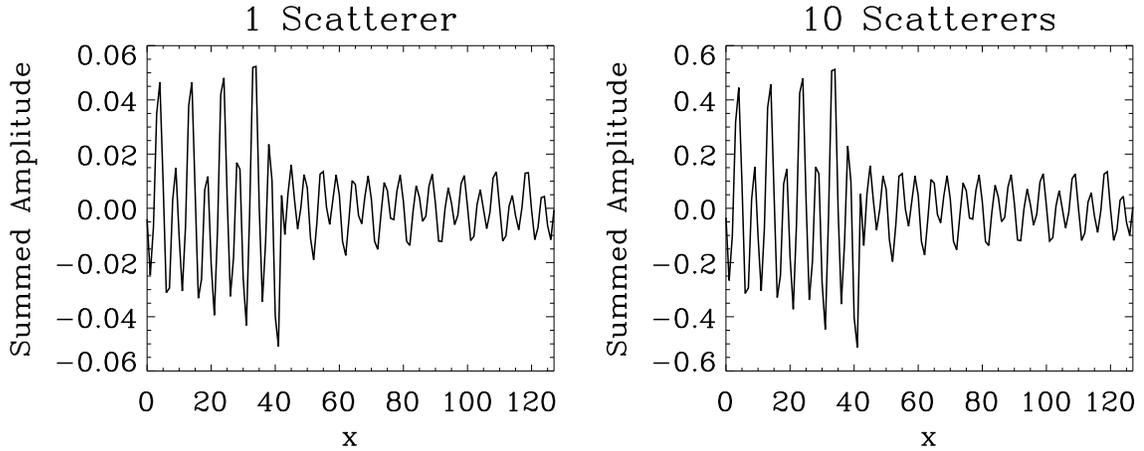


Figure 4.8: The simulated scattered data in Figure 4.7 summed over the y -dimension. The x -axis has units of pixels, while the summed amplitude has arbitrary units. Note that the summed amplitude for the 10-scatterer case is about $10\times$ the magnitude of the 1-scatterer case results. This data only shows the $t = 0$ time step. Other time steps have a similar magnitude difference in amplitude between the 1-scatterer and 10-scatterer cases. Stacking these summed amplitude data through time produces plots of the time-dependent Green’s functions of the scattering.

Stacking the summed simulation data, like Figure 4.8, through time produces plots of Green’s function $G(x_s, x_r, t)$. For in-phase incident waves with a large signal-to-noise ratio (no random noise was added to the scattering calculations), these Green’s function images are effectively TD results of cross-correlated data. (We will discuss how scattering from out-of-phase incident waves affects this relationship in Section 4.2.2.) The connection between Green’s function and the cross-correlation has been shown by many authors (e.g., Gouédard *et al.*, 2008; Lobkis and Weaver, 2001; Tsai, 2010). The general relationship is expressed by

$$\frac{d}{dt} [C(\mathbf{r}_s, \mathbf{r}_r, t)] = \frac{-A^2}{2} [G(\mathbf{r}_s, \mathbf{r}_r, t) - G(\mathbf{r}_s, \mathbf{r}_r, -t)]. \quad (4.3)$$

This equation changes somewhat depending on the restrictive conditions and assumptions of the system, including attenuation, (un)equal energy distribution between wave modes, degeneracy of modes, and low/high frequency regimes. However, in general (and for our simulation) the relationship can be approximated by Equation 4.3. For our purposes, the

Green's function has only one spatial dimension with $r = x$, since we summed over the y -dimension.

Figure 4.9 shows the Green's functions for the 1- and 10-scatterer cases. At first glance, these individual results look approximately the same; however, both plots are scaled with respect to themselves. To clearly see any differences, the 1-scatterer data must be approximately scaled to the magnitude of the 10-scatterer data before subtracting the two. We find an interesting scaled difference result. A rounded propagation feature exists in the 1-scatterer case but is diminished (if not absent) in the 10-scatterer case. This rounded feature is due to the strong circular scattering signal in the 1-scatterer case. In order to better view how this feature merges with the stronger TD feature, we also include an overlay of the scaled difference with the scaled 10-scatterer case. The approximate start time of the feature can be seen in the $t = 200$ row of Figure 4.3, as the characteristic differences between the 1 and 10 scatterer simulated data. The presence of the rounded feature could be further explored by studying the rate at which the feature diminishes with increasing n , and how large n needs to be in order to reasonably neglect the feature. However, delving into those details is beyond the scope of the current work. We simply note that we use a multiple-scatterer case moving forward to avoid this effect from the 1-scatterer assumption.

For reference, we also show the Green's function results for the incident (Figure 4.10) and total (Figure 4.11) wavefields. As expected, there is no difference between the incident wave TD results, since the incident wave packet designated in both cases is exactly the same. In the total wavefield results, the scattered wave signal is too weak to be seen compared to the incident wave signal, for both the 1-scatterer and 10-scatterer cases (the scattered wave signal being 2-3 orders of magnitude smaller than the incident wave signal). The (unscaled) difference clearly shows both the forward and backscattered waves (the incident wave contribution is effectively canceled out); however, the apparent scattering mainly comes from the 10-scatterer case since the scattered wave magnitude is $10\times$ larger than that of the 1-scatterer case. This is the same reason why the rounded propagation feature from Figure 4.9 is not visible. One might conclude that if the 1-scatterer data were

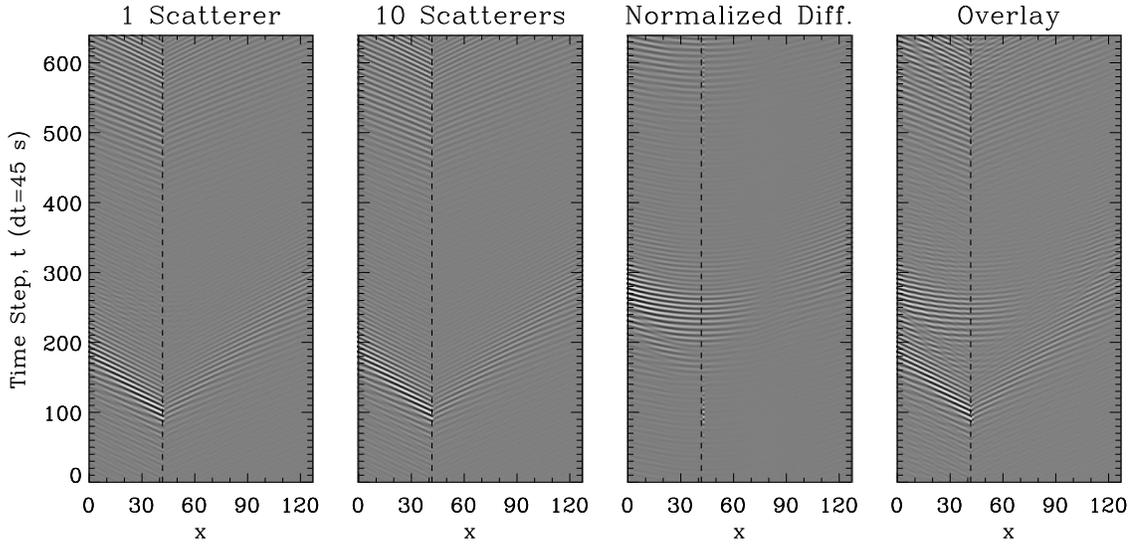


Figure 4.9: The Green's function results for the 1- and 10-scatterer *scattered* wave simulation data are shown in the left and left-center images, respectively. The x -axis has units of pixels, while the time steps on the y -axis are in seconds ($dt = 45$ s). The vertical dashed line indicates the x -location of the scatterers. Note that if the $t = 0$ data from these two figures were plotted, the results would give Figure 4.8. These images are shown to scale with themselves, so for comparison purposes the scaled-difference (right-center) and scaled-difference overlaid with the scaled-10-scatterer case (right) are also shown.

scaled in the difference between total waves (as was done in the scattering-only difference), the rounded propagation feature would appear here too. On the contrary, scaling the 1-scatterer, total-wave data would scale both the incident and scattered wave signals from the 1-scatterer case. Thus, the incident signal would still dominate and drown out *all* the scattered signals.

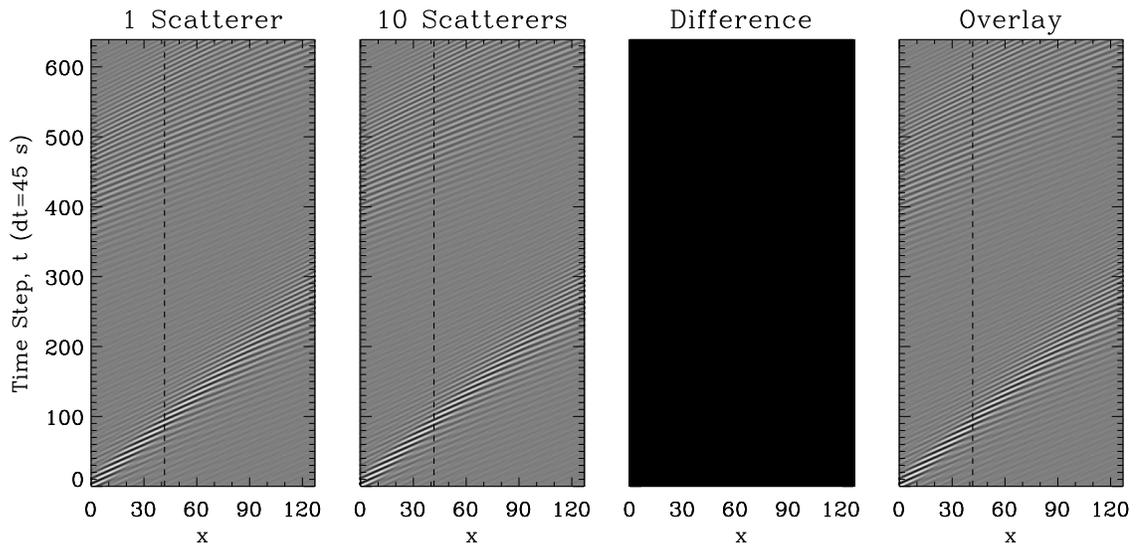


Figure 4.10: Similar to Figure 4.9, the Green's function results for the 1- and 10-scatterer *incident* wave simulation data are shown in the left and left-center images, respectively. The x -axis has units of pixels, while the time steps on the y -axis are in seconds ($dt = 45$ s). The vertical dashed line indicates the x -location of the scatterers. These images are shown to scale with themselves, so for comparison purposes the difference (right-center) and difference overlaid with the 10-scatterer case (right) are also shown. As expected, there is no difference between the incident wave linear TD results, since the incident wave packet designated in both cases is the same. Thus, the right-hand image is a mirror of the left and left-center images.

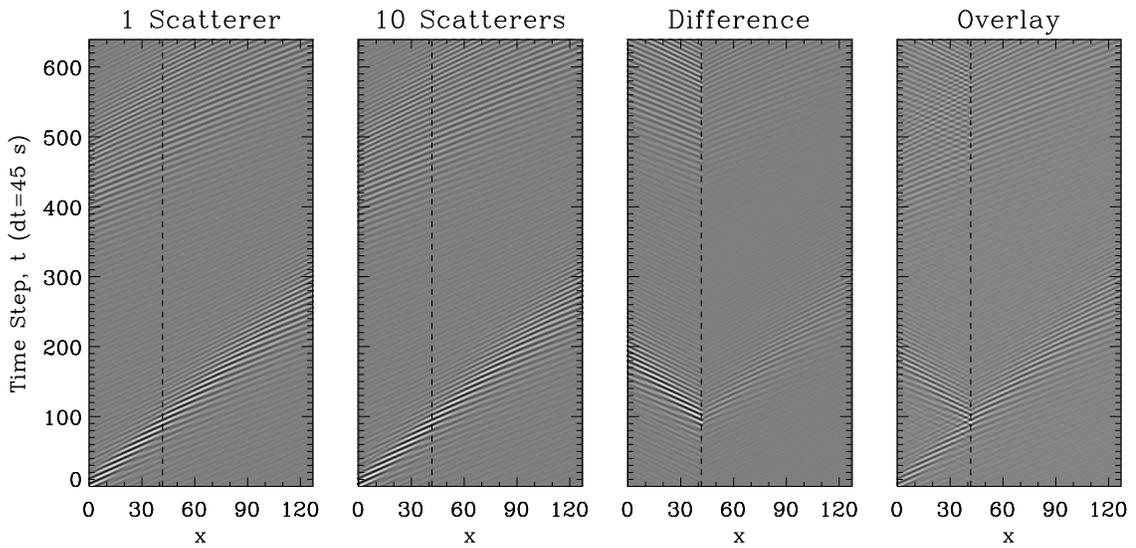


Figure 4.11: Again similar to Figure 4.9, the Green's function results for the 1- and 10-scatterer *total* wave simulation data are shown in the left and left-center images, respectively. The x -axis has units of pixels, while the time steps on the y -axis are in seconds ($dt = 45$ s). The vertical dashed line indicates the x -location of the scatterers. These images are shown to scale with themselves, so for comparison purposes the difference (right-center) and difference overlaid with the scaled-10-scatterer case (right) are also shown. Although the backscattered wave signal is too weak to be seen in both the 1-scatterer and 10-scatterer cases by themselves, the difference clearly shows both the forward and backscattered wave signals. The apparent scattering mainly comes from the the 10-scatterer case because the scattered wave magnitude is $10\times$ larger than that of the 1-scatterer case. The overlaid image on the right shows both the incident and scattered wave parts together.

What happens if we change the direction of propagation of the incident wave packet? Figures 4.12-4.14 correspond to Figures 4.7-4.9 above, but with $\phi_i = 45^\circ$. The scattering points are positioned in the same locations as in the $\phi_i = 0^\circ$ case. In Figure 4.12, notice that the effect of the non-zero azimuthal orders shows in the 1-scatterer case as a change in the direction of the strongest scattering so that it propagates parallel to the direction of the new incident wave. The same is not as obvious in the 10-scatterer case, but the strongest scattering clearly occurs at the scatterers closest to the incident wave packet (as expected). Regardless, the scatterers still produce a linear interference pattern along the y -dimension. Comparing the summed data from both 10-scatterer cases ($\phi_i = 0^\circ$ and $\phi_i = 45^\circ$) in Figure 4.13, the summed amplitude magnitudes are similar. However, the difference between the wave patterns is apparent. The Green's functions in Figure 4.14 show a somewhat surprising result. The 10-scatterer case seems to have a double TD signal, which is approximately symmetric in time about the 1-scatterer TD signal. The double signals apparently occur as the incident wave packet passes either end of the column of scatterers. While surprising, these results make sense since the column of scatterers is symmetric about $y = ny/2$, where the 1-scatterer is located. From the difference plot, it is apparent that 1) the maximum scattering amplitudes are on the same order of magnitude (no scaling required) and 2) the 1-scatterer signal indeed falls in the center between the double 10-scatterer signals. This $\phi_i = 45^\circ$ study is mainly a curiosity-driven comparison investigation and no further analysis has been pursued for this dissertation.

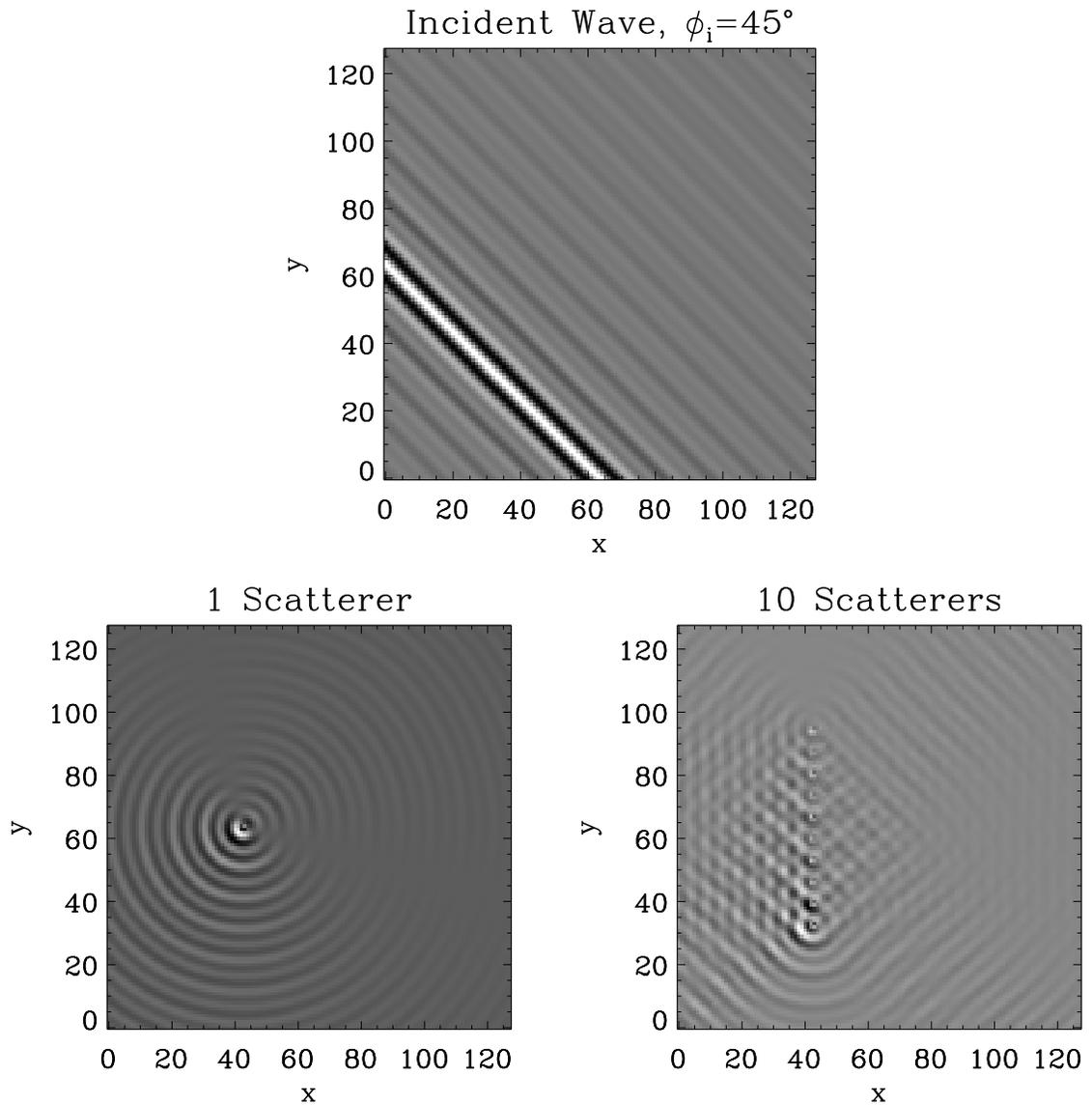


Figure 4.12: Examples of the simulated real scattered wave functions for a single scatterer (left) and a column of ten scatterers (right) at $t = 0$. Axes are in units of pixels. Greyscale colors indicate wave amplitude (black/white is minimum/maximum, grey is average), although the incident and scattered waves are not scaled with respect to each other since the incident wave signal is actually 2-3 orders of magnitude larger than the scattered wave signals. The incident wave packet (top) for these examples is initially in-phase at an angle of 45° and a radial distance of $nx/2 = ny/2$ from the domain center. Note that these images are the same simulated data as those in the top row of Figure 4.5.

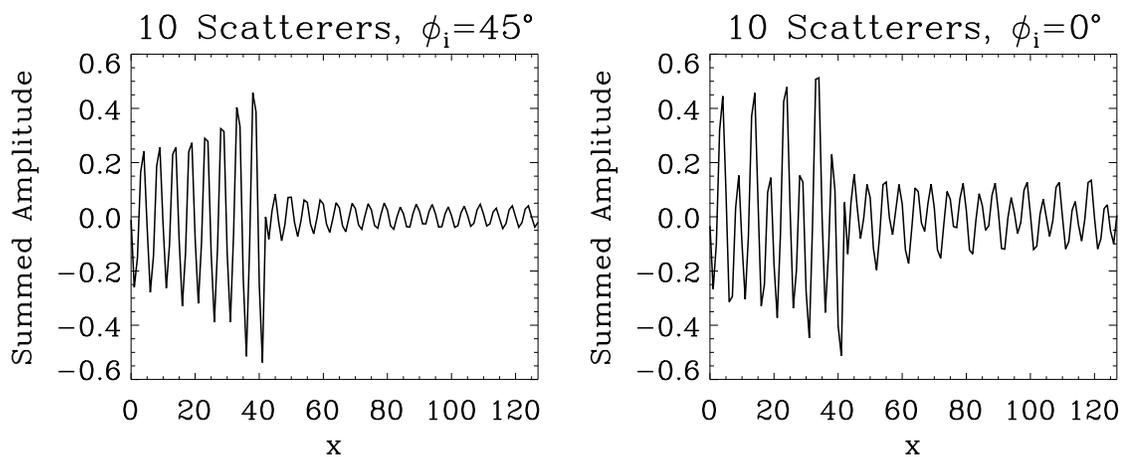


Figure 4.13: The simulated scattered data in the 10-scatterer cases from Figures 4.7 and 4.12, summed over the y -dimension. The x -axis has units of pixels, while the summed amplitude has arbitrary units. Note that the summed amplitudes for both incident angle cases have approximately the same magnitude. This data only shows the $t = 0$ time step. Other time steps have similar amplitude magnitudes between the $\phi_i = 0^\circ$ and $\phi_i = 45^\circ$ cases. Stacking these summed amplitude data through time produces plots of the time-dependent Green's functions of the scattering.

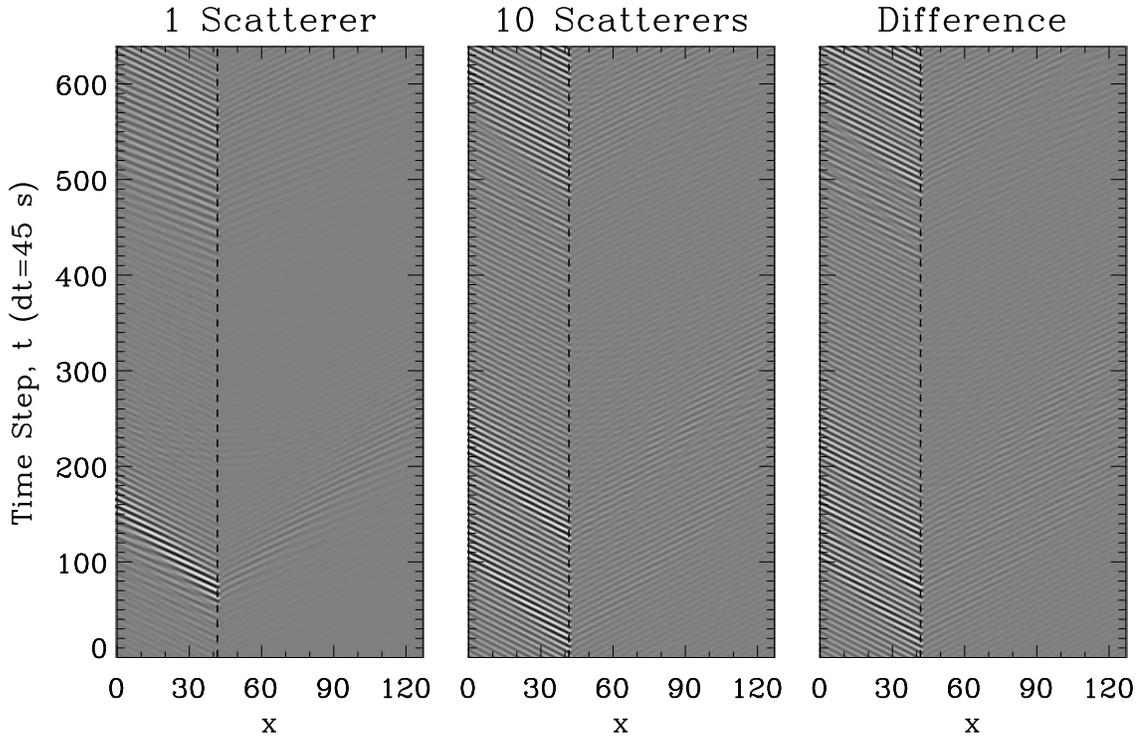


Figure 4.14: The Green's function results for the 1- and 10-scatterer *scattered* wave simulation data, with an incident angle of 45° , are shown in the left and center images, respectively. The x -axis has units of pixels, while the time steps on the y -axis are in seconds ($dt = 45$ s). The vertical dashed line indicates the x -location of the scatterers. Note that if the $t = 0$ data from the 10-scatterer case was plotted, the results would give the left plot in Figure 4.13. For comparison purposes, the difference (right) is also shown, with no scaling applied.

In summary, we have learned three things from this study, two of which are directly pertinent to the overarching research project. First, we confirmed that the intensity magnitude of the summed simulated data directly scales by a multiple of the number of scatterers n present in a line of scatterers. Second, while the main TD results for a single scatterer and column of multiple scatterers are approximately the same (when scaled by the multiplicative factor of n), the single scatterer produces an additional circular scattering effect that is not apparent in the multiple scatterer result. Finally, for the multiple scatterers case, the first two properties are partially dependent on the angle of incidence with respect to the line of scatterers. They have been proven when the incident wave propagation is perpendicular to the scattering line, but may differ as the angle becomes more acute with respect to the axis of symmetry. However, ensemble theory suggests that if multiple random waves were incident from a variety of angles, these angular effects would cancel to some extent. Furthermore, the linear summation along the axis of symmetry would give TD results similar to the purely perpendicular case, as proven in Section 3.2.3. The simulation has not been used to explore the details of this line of reasoning in this dissertation, but could be investigated in the future.

4.2.2 Applied Comparison of Cross-Correlation and Green's Function

As mentioned in the previous section, there is a close relationship between the cross-correlation and Green's function of a set of data. Assuming that the source function is white noise distributed everywhere acting at time t , Gouédard *et al.* (2008) state:

In the frequency domain, a white noise [source function] contains all the frequencies with a random phase. In the time domain, this is a random wavefield [modal sum of Green's function] such that the position and the activation time of each source are uncorrelated. (Gouédard *et al.*, 2008, p.378)

However, when the frequencies are in-phase with no added noise, the sources in the wavefield are automatically correlated. Therefore, it is not necessary to perform a cross-correlation

on in-phase simulated data; however, the out-of-phase simulation must be cross-correlated to retrieve a coherent TD signal. We can use the simulation to visualize this relationship.

Figure 4.15 shows the Green’s function results for the 10-scatterer simulated data with an in-phase incident wave packet of $\phi_i = 0^\circ$. The scatterers are located parallel to the y -axis at $x = nx/3$, with spacing $d = 5$ pixels in the y -direction. As discussed in Sections 3.2.3 and 4.2.1, these results are produced by summing over the axis of symmetry, which corresponds to the y -dimension in our case. The total wave and incident wave Green’s functions are visibly indistinguishable because the resultant scattered wave is much weaker than the incident wave. However, the scattered wave does contribute a small but significant signal, including a backscattered wave signal. The backscattered wave is apparent in the scattered wave Green’s function plot as the coherent signal to the left of the line of scatterers. Note that this reflection occurs at the same point in time as when the incident wave passes the scatterers.

Similarly, Figure 4.16 shows the Green’s function results of when the incident wave modes are randomly phased. The scatterer locations and angle of incidence are the same as for the in-phase case. In comparison to the in-phase Green’s functions, the random phase results have strong signals throughout the duration of the simulation time. The strong signals correspond to small, randomly coherent incident wave packets, as can be seen in Figure 4.4.

We can physically see the role that performing the cross-correlation plays when we apply it to the the randomly-phased Green’s functions. Although we have no choice but use the total wave in actual solar data, we are free to cross-correlate the scattered wave with the incident wave in the simulated data, which is effectively the same as the difference in TD signals between the total and incident waves. This allows us to examine the purely-scattered wave signal in detail. The scattered wave cross-correlation with respect to the $x = 0$ incident signal is shown in Figure 4.17, along with the incident wave’s cross-correlation with itself at $x = 0$ for reference. It is clear that the cross-correlation aligns the phases; the cross-correlations in Figure 4.17 resemble the coherent signal of the Green’s functions

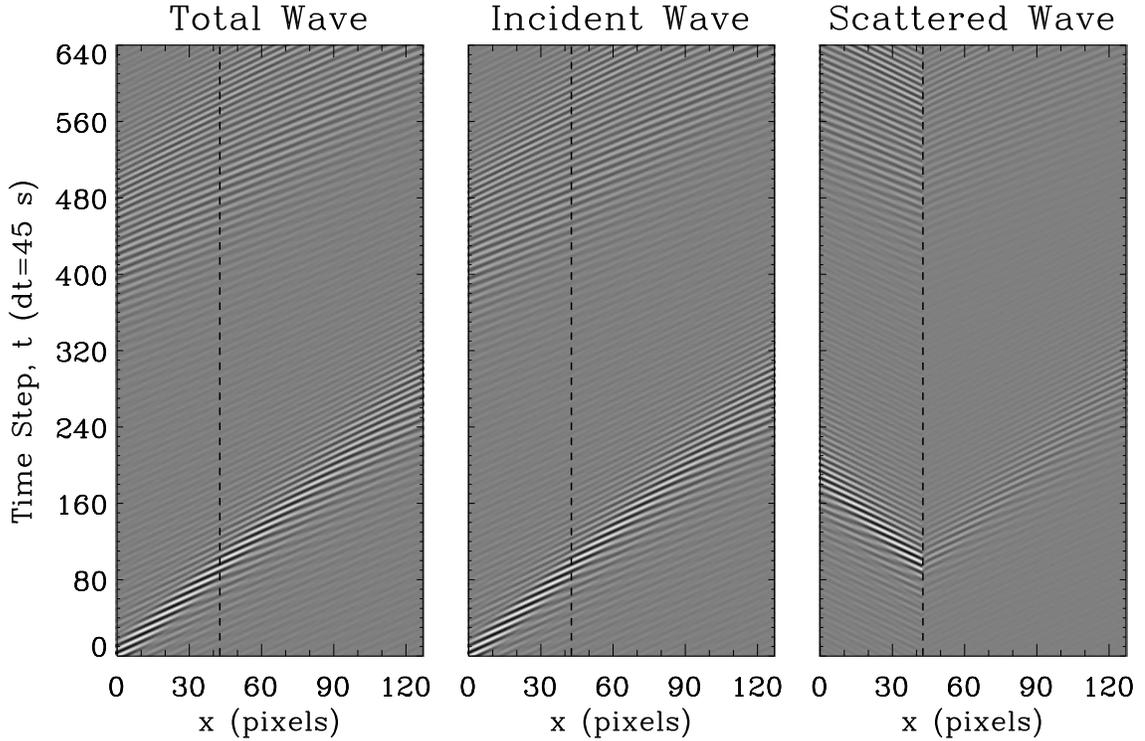


Figure 4.15: The Green's function results from the 10-scatterer, in-phase, $\phi_i = 0^\circ$ simulated data, corresponding to the simulated data represented in Figure 4.3. The x -axis has units of pixels, while the time steps on the y -axis are in seconds ($dt = 45$ s). The total wave (left) and incident wave (middle) results are indistinguishable since the scattered wave is weaker than the incident wave. The scattered wave result (right) shows both a forward and backward scattered wave to either side of the line of scatterers, which is designated by the vertical dashed line.

in Figure 4.15.

We note here that there is only the forward propagating branch of the TD signals in the simulated data, which is particularly clear in Figure 4.17. This is because we only have incident waves propagating in one direction. If the various wave modes were randomized in their incident angle, the backward propagating branch would appear in the TD results. As it is, we do not need to consider any negative time lags in the simulated cross-correlation results.

Because we focus on the in-phase simulated data in this chapter, the cross-correlation calculations are not necessary and we only use the Green's function results moving forward.

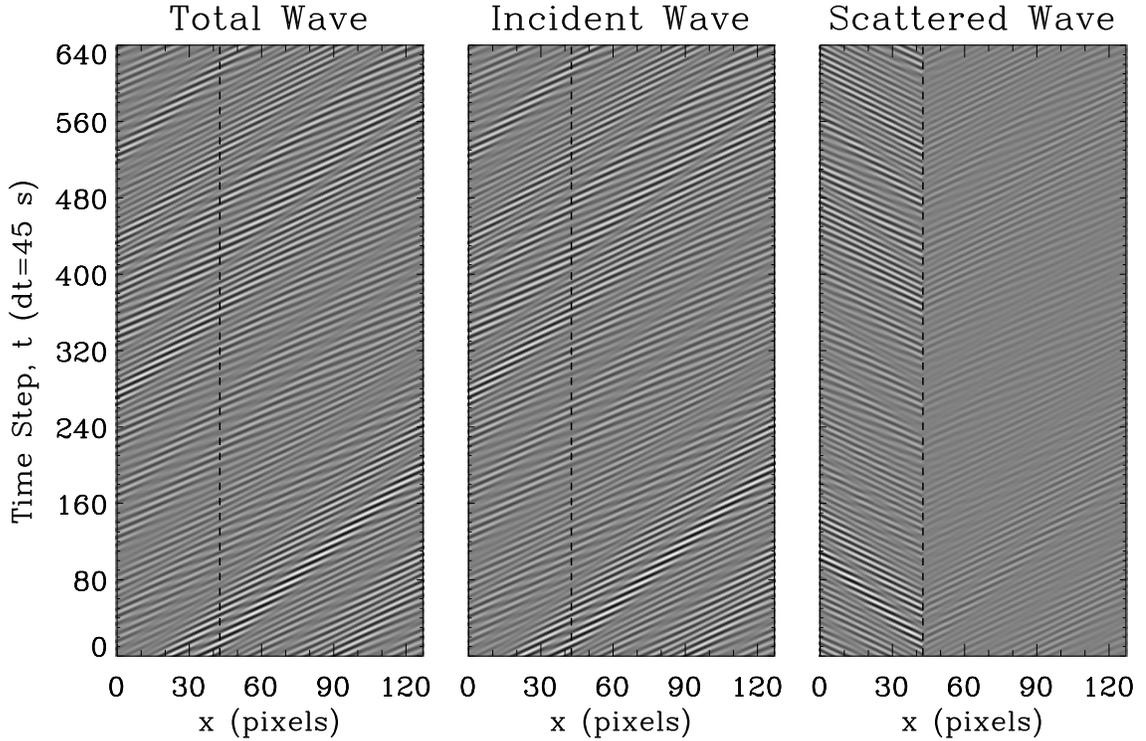


Figure 4.16: The Green's function results from the 10-scatterer, random phased, $\phi_i = 0^\circ$ simulated data, corresponding to the simulated data represented in Figure 4.4. The x -axis has units of pixels, while the time steps on the y -axis are in seconds ($dt = 45$ s). The total wave (left) and incident wave (middle) results are indistinguishable since the scattered wave is weaker than the incident wave. The scattered wave result (right) shows both a forward and backward scattered wave to either side of the line of scatterers, designated by the vertical dashed line. The random phase results have strong signals throughout the duration of the simulation time, corresponding to small, randomly coherent incident wave packets.

However, we note that the wave packet is in-phase at $t = 0$ and has shifted slightly out-of-phase by the time it reaches the line of scatterers. The Green's function results would more accurately represent true TD results if the wave packet were exactly in-phase when incident upon the scatterers. However, the phase shift is small enough that the approximation of the Green's function to TD results is reasonable.

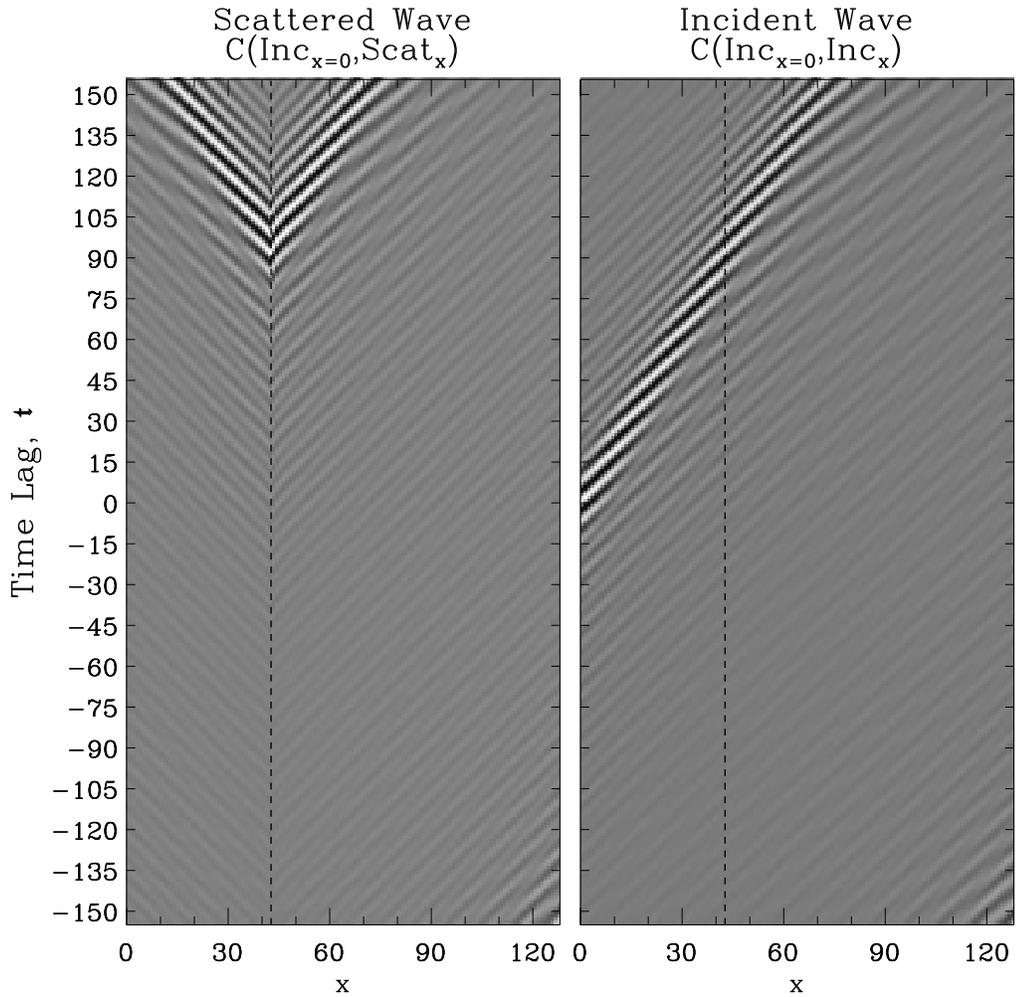


Figure 4.17: The cross-correlated results of the multiphased Green’s functions in Figure 4.16. The x -axis has units of pixels, while the time lag on the y -axis is in seconds ($dt = 45$ s). The cross-correlation between the scattered wave and the incident wave at $x = 0$ (left) produces the same result as cross-correlating the total wave with itself at $x = 0$. The backscattered wave diverges from the line of scatterers, designated by the vertical dashed line. The cross-correlation of the incident wave with itself at $x = 0$ (right) crosses the line of scatterers at the same time.

4.2.3 Statistical Analysis

When comparing noisy data, the difference between data will be even noisier. As has been discussed, one of the main challenges in TD helioseismology is improving the signal-to-noise ratio. It is difficult to perform any noise reduction techniques on the difference between TD results without affecting the already weak signal. In many cases, the signal-to-noise

ratio of the difference is so low that the signal and noise are virtually indistinguishable. The question becomes, how can we determine which difference features are significant and which are products of noise? The answer comes from performing a statistical analysis between the two TD data sets. We present a new statistical method for comparing TD results, stemming from the Student's t -test analysis. Before presenting some results, we will briefly review the equations needed to compute the t -statistic and its significance.

Assuming equal variances, the independent Student's t -test statistic between two Gaussian data distributions A and B is calculated as (Press *et al.*, 1992, p.616)

$$t = \frac{\bar{A} - \bar{B}}{s_{AB} \sqrt{\frac{1}{n_A} + \frac{1}{n_B}}}, \quad (4.4)$$

$$s_{AB} = \sqrt{\frac{(n_A - 1) s_A^2 + (n_B - 1) s_B^2}{n_A + n_B - 2}},$$

where n_* are the sample sizes, \bar{A} and \bar{B} denote the sample means, s_* are the sample standard deviations, and s_{AB}^2 is an unbiased estimator of the common variance between the two sampled data sets. The individual degrees of freedom are $d.f._* = n_* - 1$, so the total number of degrees of freedom $d.f. = d.f._A + d.f._B = n_A + n_B - 2$. In the specific case where $n_A = n_B = n$, these equations simplify such that $d.f. = 2n - 2$ and s_{AB} becomes the pooled standard deviation. Note that the denominator of t is the standard error of the difference between the means.

The significance of the t -statistic is calculated using an incomplete beta function, $I_x(a, b)$ (Press *et al.*, 1992, p.226,228,616). The probability distribution Prob_t is dependent upon the value of t and the total $d.f.$ of the two samples. It is given by

$$\text{Prob}_t = 1 - I_{\frac{d.f.}{d.f.+t^2}} \left(\frac{d.f.}{2}, \frac{1}{2} \right). \quad (4.5)$$

Statistically, Prob_t is the likelihood that, for the specified number of degrees of freedom, the

observed difference of the two sample means would be smaller than the observed difference if the means were the same. For t to indicate a significant difference at the 95% confidence level, $\text{Prob}_t \geq 0.95$ (significance level $I_x(a, b) \leq 0.05$).

If we assume unequal variances between the two distributions, the independent t -test statistic becomes (Press *et al.*, 1992, p.617)

$$t = \frac{\bar{A} - \bar{B}}{s_{A-B}}, \tag{4.6}$$

$$s_{A-B} = \sqrt{\frac{s_A^2}{n_A} + \frac{s_B^2}{n_B}}.$$

Now, s_*^2 are the unbiased estimators of the sample variances, but s_{A-B}^2 is *not* a pooled variance. The individual degrees of freedom are defined the same as in the equal-variance case, but the total number of degrees of freedom is now approximated as

$$d.f. = \frac{\left(\frac{s_A^2}{n_A} + \frac{s_B^2}{n_B}\right)^2}{\left[\left(\frac{s_A^2}{n_A}\right)^2 / (n_A - 1)\right] + \left[\left(\frac{s_B^2}{n_B}\right)^2 / (n_B - 1)\right]}. \tag{4.7}$$

This approximation is known as the Welch-Satterthwaite equation, and the unequal-variance calculation of the t -statistic is called Welch's t -test (Welch, 1947). The probability distribution is calculated in the same way as for the equal-variance Student's t -test.

We determine whether two distributions have significantly different variances by performing an F -test (Press *et al.*, 1992, p.619). The F -test tests the null hypothesis that the two sample variances are equal. The F -statistic is the ratio between the variances, so values of either $F \gg 1$ or $F \ll 1$ indicate highly significant differences. Again, the probability of producing a particular value of F if the variances are *not* equal (rejecting the null hypothesis) is calculated using an incomplete beta function of the form (Press *et al.*, 1992,

p.226,229,619)

$$\text{Prob}_F = 2I_{\frac{d.f.B}{d.f.B+d.f.AF}} \left(\frac{d.f.B}{2}, \frac{d.f.A}{2} \right). \quad (4.8)$$

The factor of 2 arises because we are generally willing to reject the null hypothesis with either very large or very small values of F , so the significance is “two-tailed” (the sum of two incomplete beta functions, one “left”-tailed for very small F and one “right”-tailed for very large F). The incomplete beta function follows the symmetry relation $I_x(a, b) = 1 - I_{1-x}(b, a)$, which indicates that the two tails are always equal so we only need to calculate one tail and double it. Also note that if the distributions are highly similar, the two tails can overlap. Such a scenario can produce probabilities greater than 1. This impossibility is resolved by subtracting the too-large probability from 2 (the sum of the upper-limit statistical probabilities (1) of both tails). With this in mind, a value of $\text{Prob}_F \leq 0.05$ implies a 95% confidence in the rejection of the null hypothesis. In other words, the variances are statistically different. We note that it is possible to have a value of $F \simeq 1$ with a statistically significant Prob_F rejecting the null hypothesis. In this case, we would assume the variances to be statistically different, despite their similar values.

Our use of the t -tests is slightly different from the standard definitions described above. We want to look at a “map” of the differences between TD results, instead of the difference of the averages. Therefore, instead of subtracting the full-data averages in the numerators of Equations 4.4 and 4.6, we use a “superpixel” method, taking the difference between a running-average of a $p \times p$ array around corresponding points in the two TD plots, and using the statistics of those superpixels. For example, a 3×3 superpixel would have $n_* = 9$, and the average and standard deviation of this array would constitute \bar{A} (\bar{B}) and s_A (s_B) in the t -statistic equations, respectively. Edges are taken into account in the statistics. The F and Prob_F values are calculated to determine whether the similar- or unequal-variance equations should be used for each superpixel. The corresponding probability maps are produced using Equation 4.5. To create a difference map of only the statistically significant differences, we threshold the subsequent probability maps to above 95% significance and

then mask the t -statistic maps by the probability threshold maps. We note that using the superpixel method loses some resolution due to the localized averaging. We are currently investigating alternative statistical comparison techniques, which are beyond the scope of this dissertation.

For the simulation, we statistically compare the in-phase scattered wave Green's function results between the 1-point and 10-point scattering cases, first for $\phi_i = 0^\circ$ (Figures 4.18-4.22) and then for $\phi_i = 45^\circ$ (Figures 4.23-4.27). In both simulation comparisons, we show the 3×3 and 5×5 superpixel results for contrast, considering that different resolutions of the spatial averaging will affect the statistical results (as mentioned above).

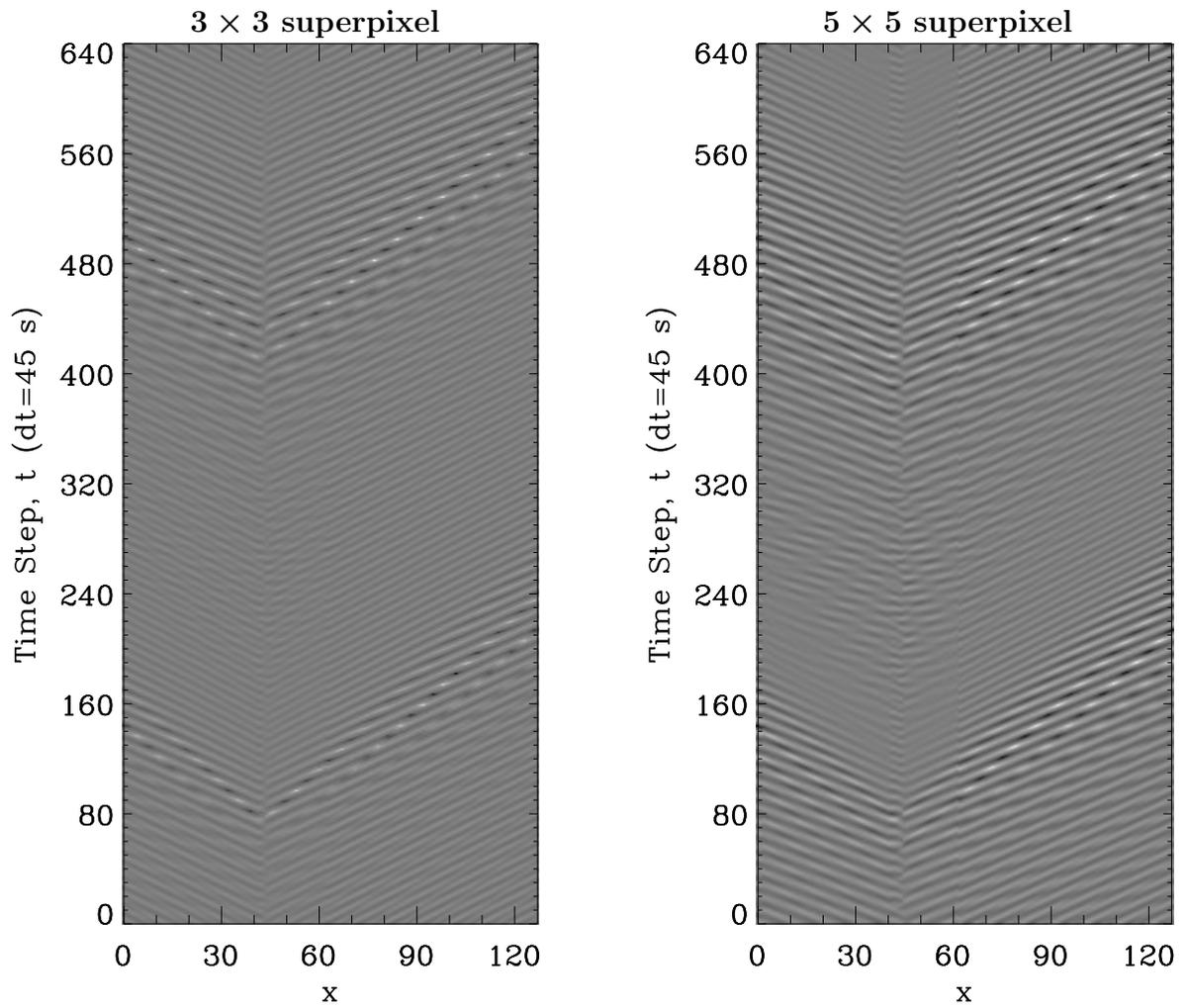


Figure 4.18: Student's t -test results between the Green's functions of in-phase, $\phi_i = 0^\circ$ scattering off of 1 and 10 scatterers. The x -axis has units of pixels. Both the 3×3 (left) and 5×5 (right) superpixel calculations are shown.

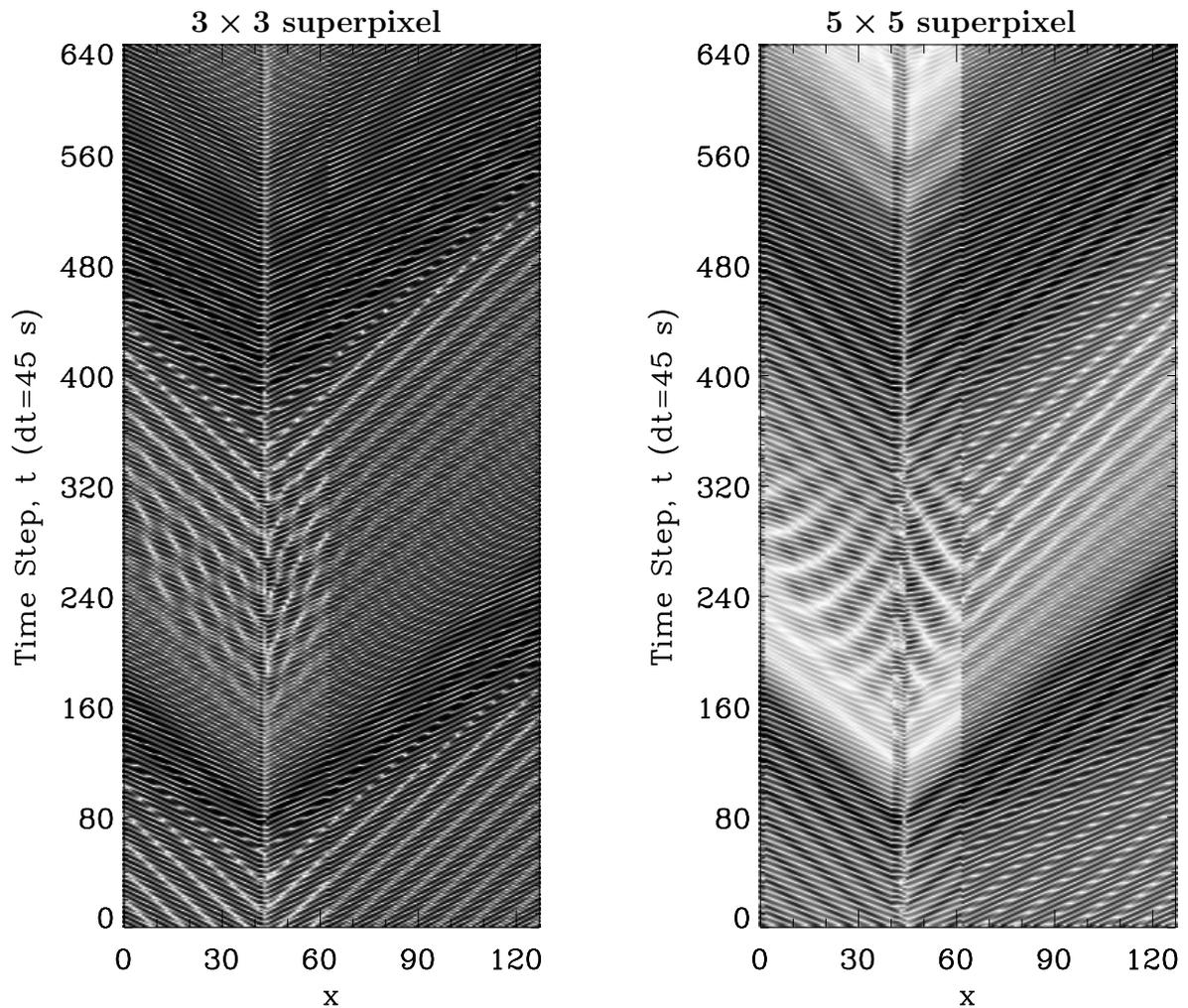


Figure 4.19: Student's t -test significance of differences results reflecting the statistical significance of the t -statistic results shown in Figure 4.18. The x -axis has units of pixels. Again, both the 3×3 (left) and 5×5 (right) superpixel calculations are shown. Note that darker shades indicate higher significance.

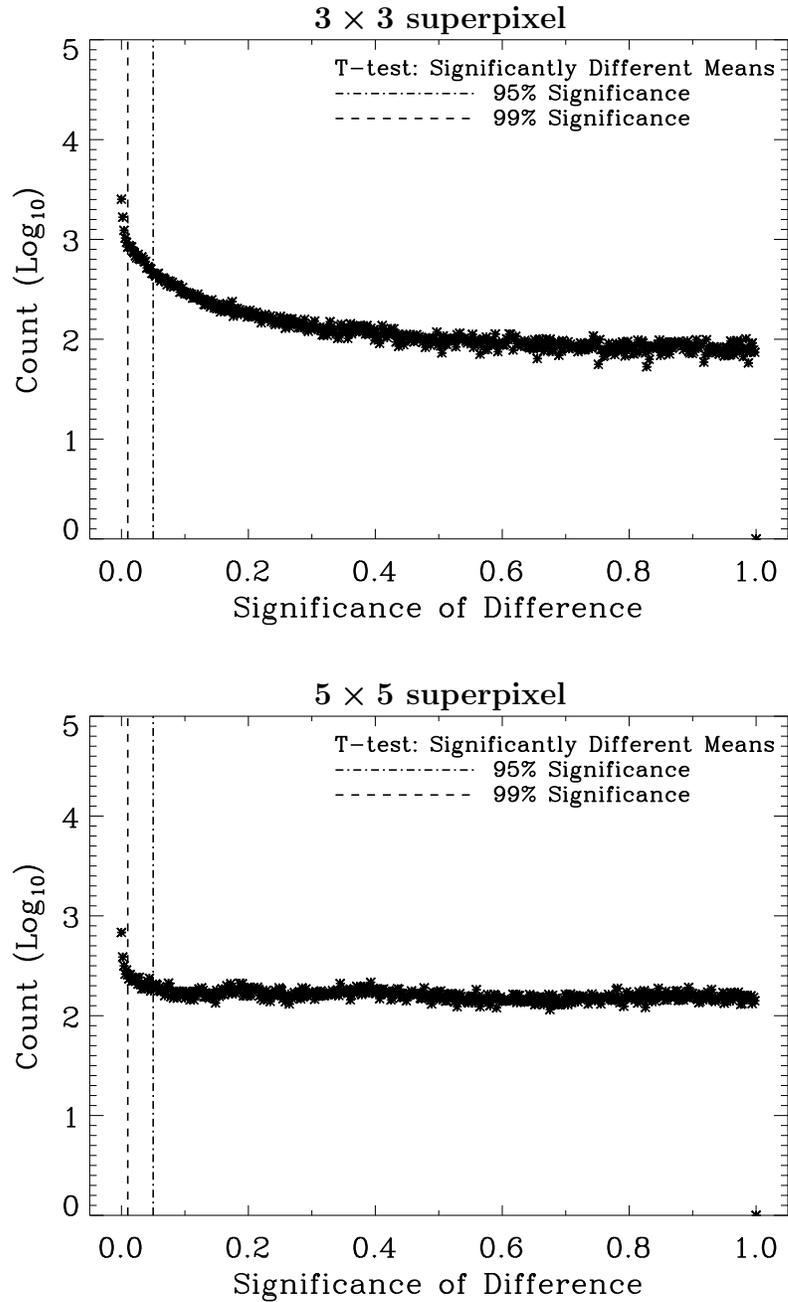


Figure 4.20: Histograms of the significance distributions in Figure 4.19. Both the 3×3 (top) and 5×5 (bottom) superpixel cases are shown. All significance left of the vertical dash-dotted line indicates significant differences (significance level ≤ 0.05 ; probability of significance = 95%) between the 1 and 10 scatterer cases, while the dashed line indicates *highly* significant differences (significance level ≤ 0.01 ; probability of significance = 99%).

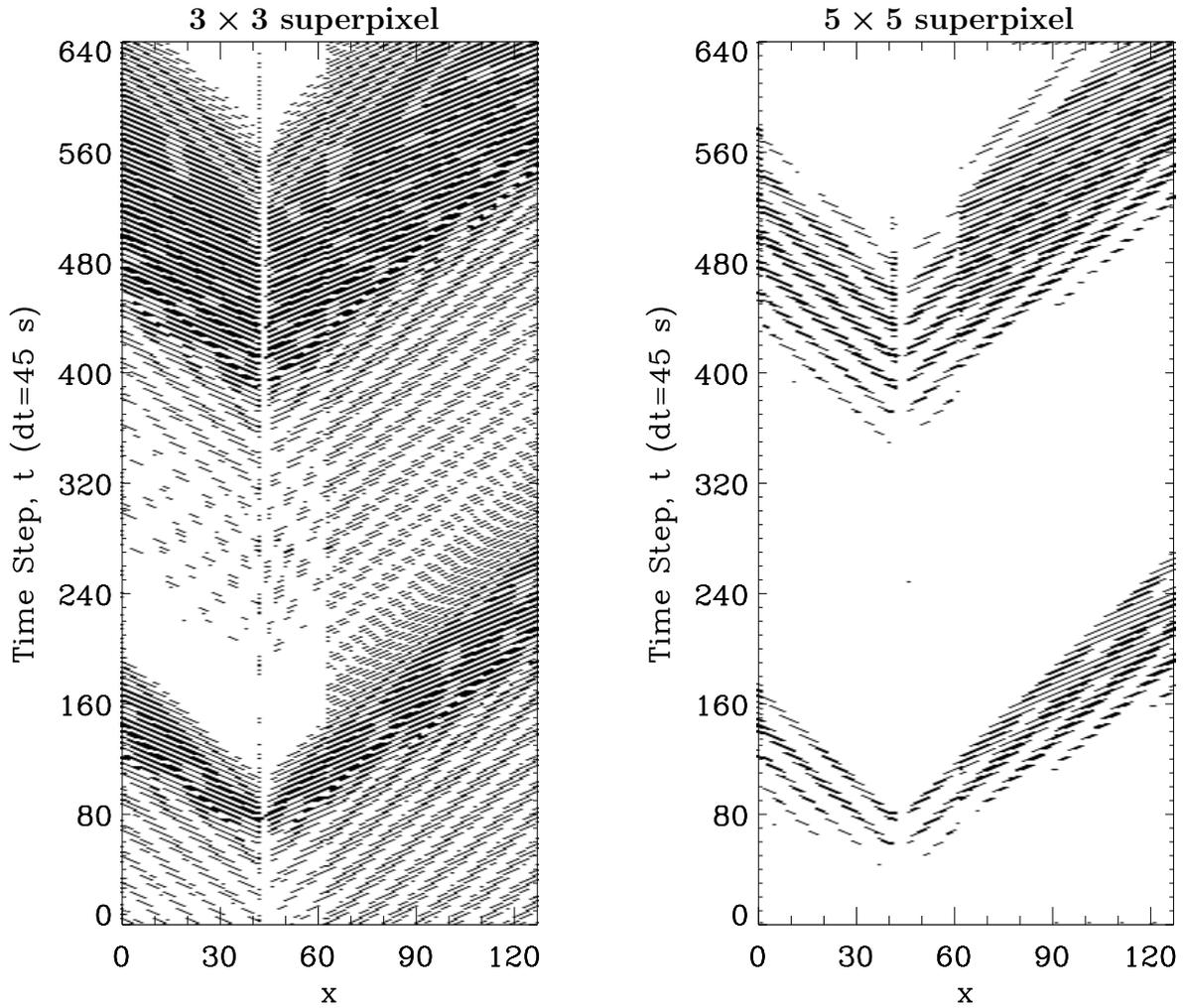


Figure 4.21: Binary Student's t -test significance results from Figure 4.19, masked to above 95% probability of significance. The x -axis has units of pixels. Again, both the 3×3 (left) and 5×5 (right) superpixel calculations are shown. Note that black regions indicates statistically significant differences; white, statistically insignificant differences.

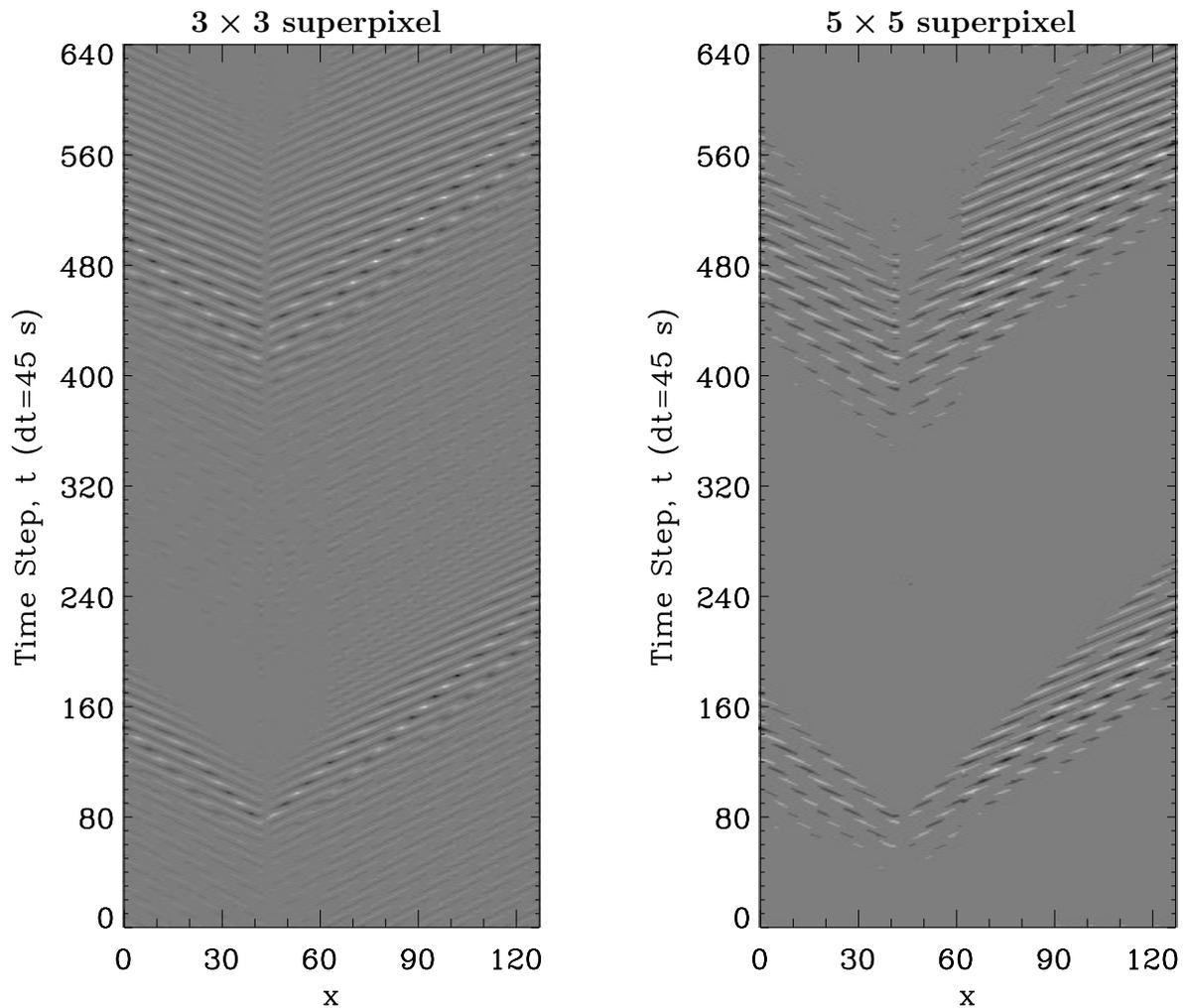


Figure 4.22: Student's t -test results from Figure 4.19, masked to above 95% probability of significance. The x -axis has units of pixels. Again, both the 3×3 (left) and 5×5 (right) superpixel calculations are shown. In either case, the main region of difference between the TD signals is passed through as statistically significant.

For the in-phase, $\phi_i = 0^\circ$ Green's function case, Figure 4.18 shows the initial t -statistic results comparing 1-scatterer versus 10-scatterers, calculated with our "superpixel" versions of Equations 4.4 and 4.6. Results from both a 3×3 and 5×5 superpixel are shown. For each running-average set of superpixels, the F and Prob_F values (Equation 4.8) are calculated and used to determine whether the similar- or unequal-variance t -test equations will be used for the statistical analysis of those localized data. Looking at the results of the significance tests in Figure 4.19, computed using Equation 4.5, the 5×5 superpixel case finds more less-significant differences than the 3×3 superpixel case, as expected due to the broader averaging scheme of the former. However, both probability calculations seem to identify a significant difference in the main wave propagation features. Figure 4.20 shows histograms of the significance values for the 3×3 and 5×5 superpixel cases. If we apply a binary threshold at the 95% significance level, displayed in Figure 4.21, we can then use the results as a mask of the t -test data so that only the significant differences are displayed, as in Figure 4.22. For both superpixel cases, the main wave propagation features of the t -test results qualify as statistically significant differences, for both the transmitted and reflected portions of the scattered wave.

Similarly, for the in-phase, $\phi_i = 45^\circ$ Green's function case, Figure 4.23 shows the initial t -statistic results comparing 1- versus 10-scatterers, using a 3×3 and 5×5 superpixel. Again, the F -statistic and Prob_F values are used to determine which superpixel locations require similar- or unequal-variance calculations. Like before, Figure 4.24 clearly shows that the 5×5 superpixel result has more instances of lower-significance of differences than the 3×3 result, but both seem to indicate that the main wave features are of higher significance. Figure 4.25 displays the histograms of the significance values for the two superpixel cases. Applying the binary threshold at the 95% significance level, we produce the significance masks in Figure 4.26. The most obvious difference between the 3×3 and 5×5 superpixel cases here is that the increased averaging (lower resolution) of the 5×5 superpixel case loses some of the main backscattered wave signal. The masked t -test results are displayed in Figure 4.27. The two results are somewhat difficult to differentiate, but the

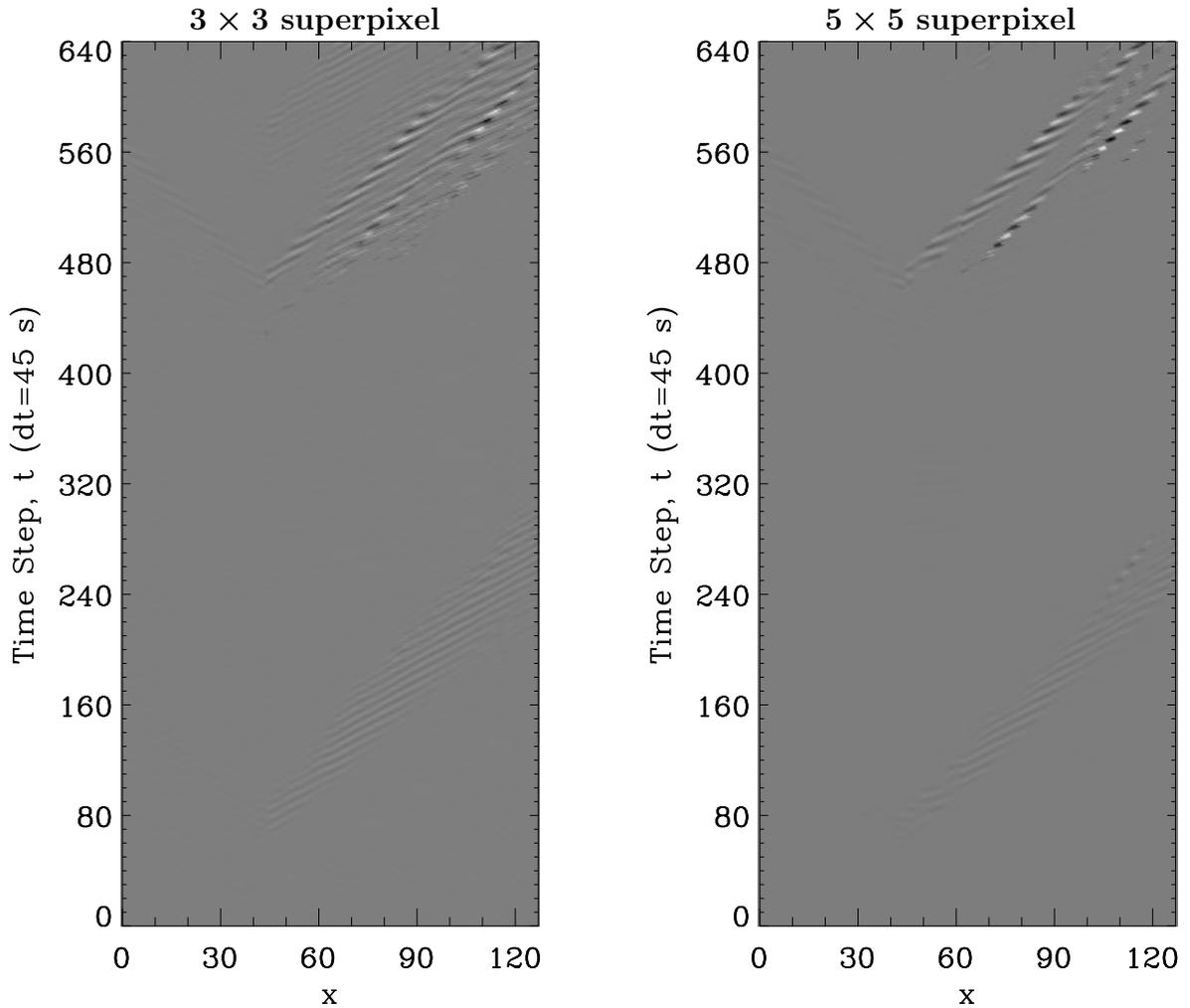


Figure 4.23: Student's t -test results between the Green's functions of in-phase, $\phi_i = 45^\circ$ scattering off of 1 and 10 scatterers. The x -axis has units of pixels. Both the 3×3 (left) and 5×5 (right) superpixel calculations are shown.

statistically significant differences in the main TD signals have already been made apparent by Figure 4.26.

Other statistical comparisons between simulated Green's function results (e.g., total waves, randomized phases, cross-correlated, etc.) were also performed but we will not present them all here. Further investigations into these distinctions, as well as the statistical analysis between other wave-parameter combinations, are areas of future work.

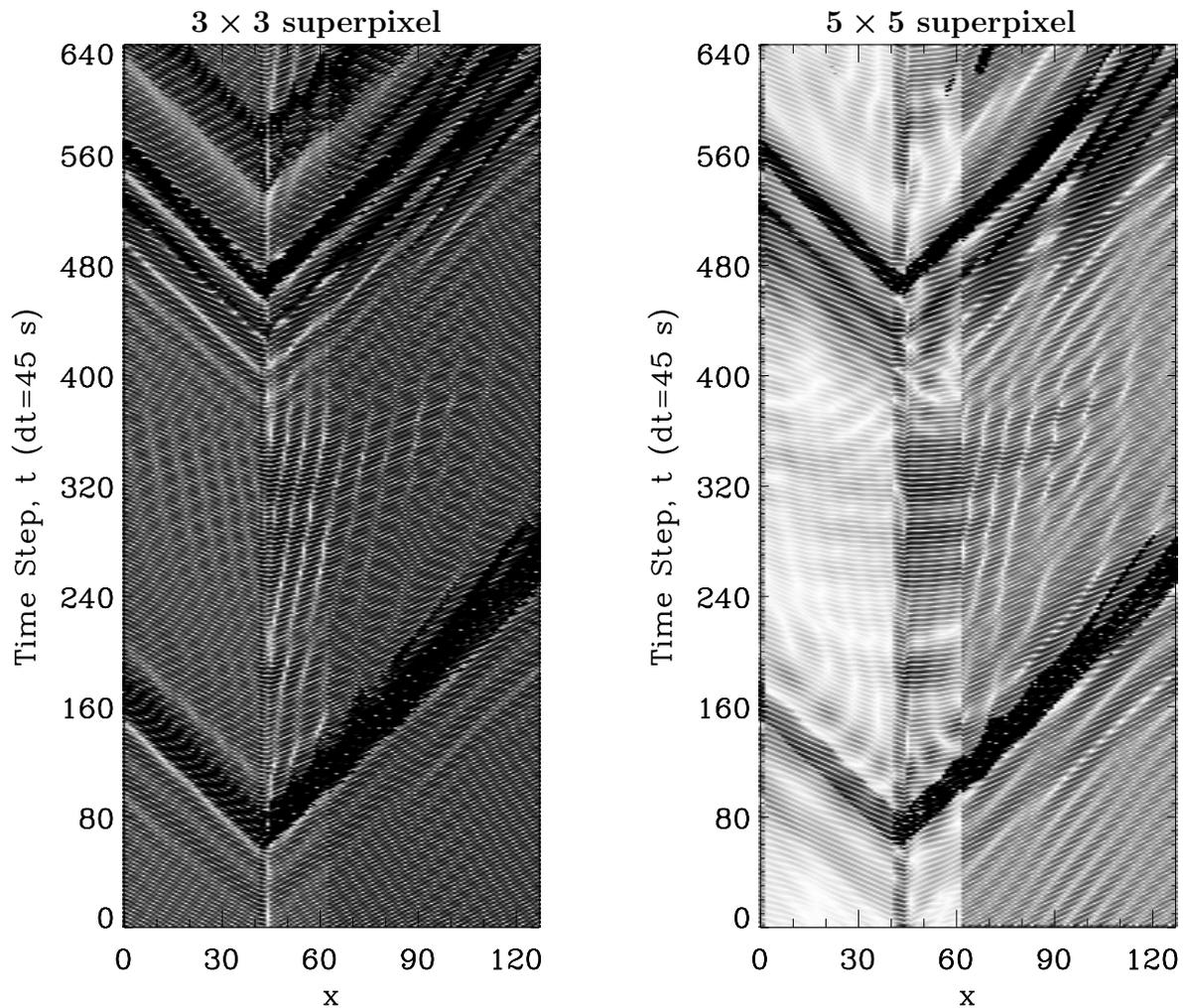


Figure 4.24: Student's t -test significance results reflecting the statistical significance of the t -statistic results shown in Figure 4.23. The x -axis has units of pixels. Again, Both the 3×3 (left) and 5×5 (right) superpixel calculations are shown. Note that darker shades indicate higher significance.

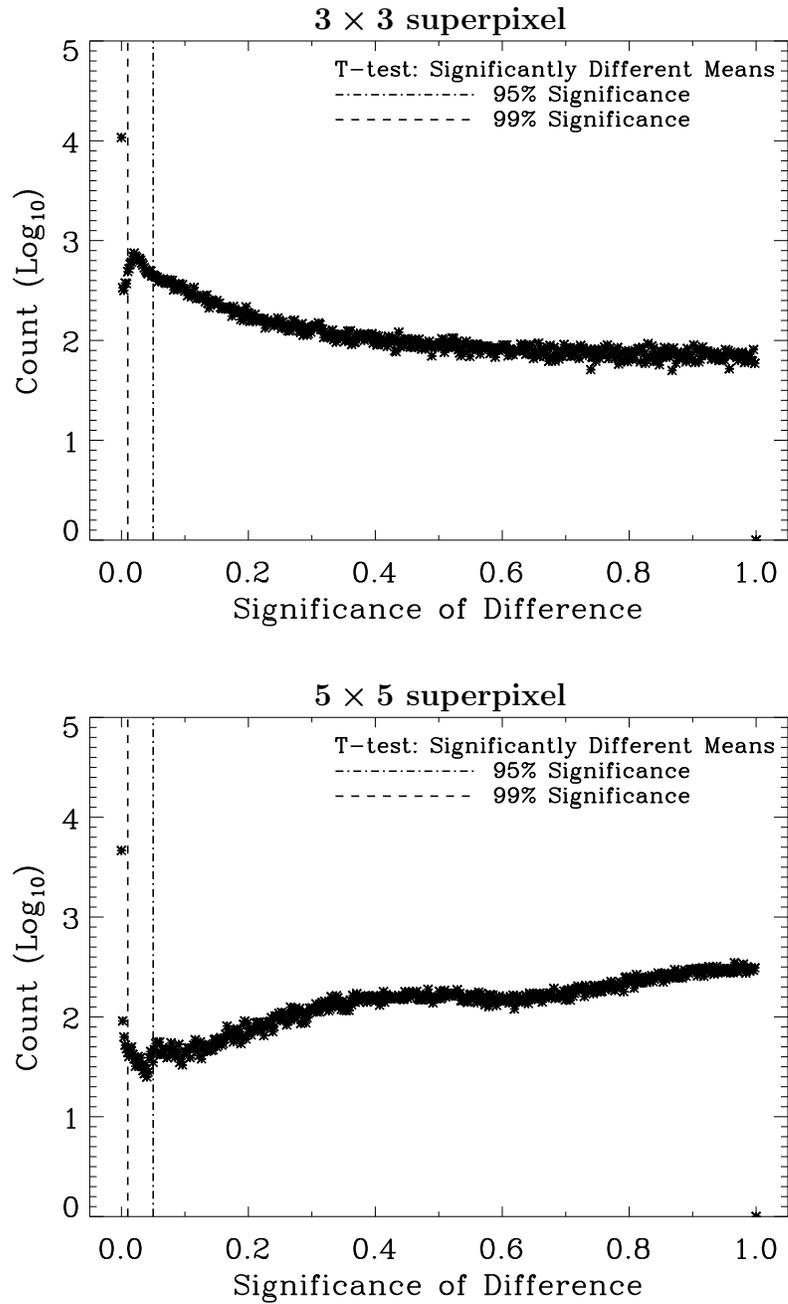


Figure 4.25: Histograms of the significance distributions in Figure 4.24. Both the 3×3 (top) and 5×5 (bottom) superpixel calculations are shown. All significance left of the vertical dash-dotted line indicates significant differences (significance level ≤ 0.05 ; probability of significance = 95%) between the 1 and 10 scatterer cases, while the dashed line indicates *highly* significant differences (significance level ≤ 0.01 ; probability of significance = 99%).

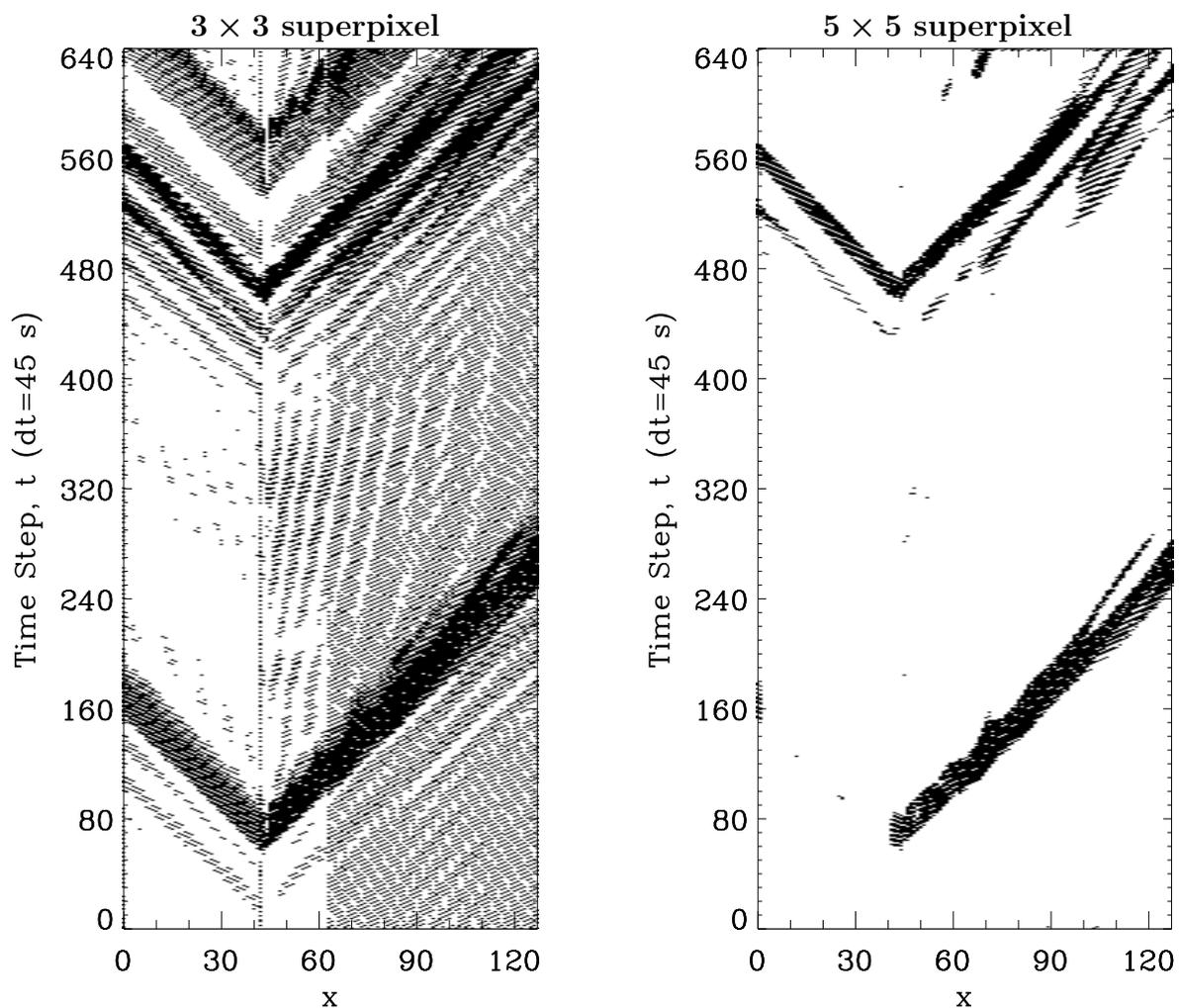


Figure 4.26: Binary Student's t -test significance results from Figure 4.24, masked to above 95% probability of significance. The x -axis has units of pixels. Again, both the 3×3 (left) and 5×5 (right) superpixel calculations are shown. Note that black regions indicates statistically significant differences; white, statistically insignificant differences.

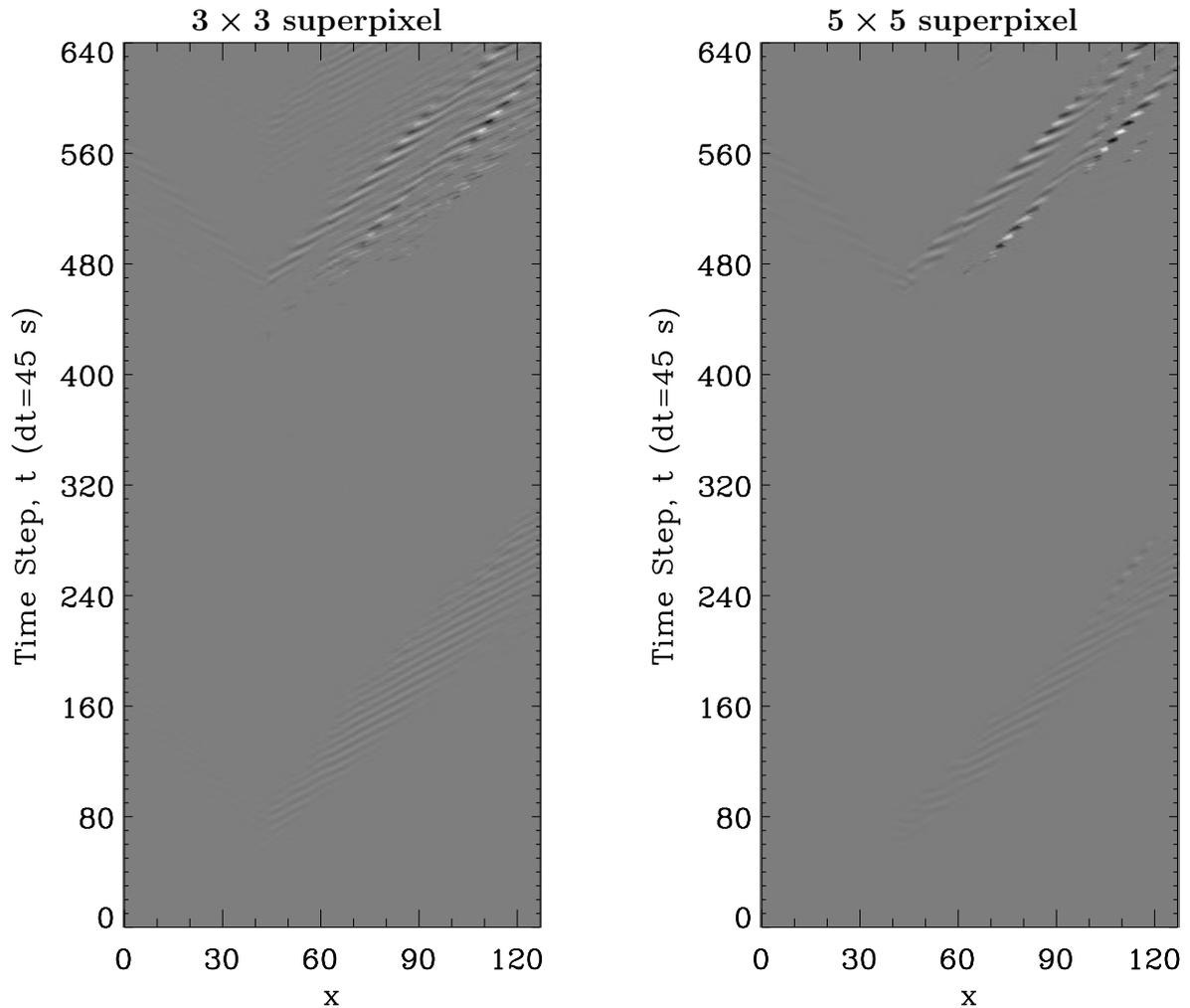


Figure 4.27: Student's t -test results from Figure 4.24, masked to above 95% probability of significance. The x -axis has units of pixels. Again, both the 3×3 (left) and 5×5 (right) superpixel calculations are shown. In either case, the main region of difference between the TD signals is passed through as statistically significant with the exception of the backscattered wave signal at earlier time steps in the 5×5 superpixel case (as seen in Figure 4.26).

4.3 Results

For the simulated data in our linear TD method, we use 10 scattering points parallel to the y -axis, centered around $y = n_y/2$, along $x = n_x/3$ with a spacing $d = \Delta y = 5$ pixels. The incident wave packet is in-phase along the edge of the domain at $t = 0$, with an incident angle of $\phi_i = 0^\circ$. As stated in Section 4.1, the spatial resolution is 0.729 Mm/pixel. The simulation is created for a total duration of 8 hours, with a 45 second cadence. The harmonic degrees range between $600 \leq \ell \leq 1200$ with a spacing $\Delta\ell = 30$; the azimuthal modes include $0 \leq m \leq 3$. Examples from this data set were presented in Figure 4.3.

Figure 4.15 shows the Green’s function results for the above case. As discussed in Sections 3.2.3 and 4.2.1, the results are produced by summing over the axis of symmetry in the simulated wavefield. The total- and incident-wave Green’s functions are visibly indistinguishable because the resultant scattered wave is much weaker than the incident wave. However, the scattered wave contributes a small but significant signal, including a backscattered wave signal. Note that the coherent scattered signal occurs at the same time as when the incident wave packet passes the line of scatterers, as expected.

Recall that the goal of creating this simulation was to test the linear TD scheme for an elongated scattering region, specifically on a data set of which we know and can control the input parameters. For this purpose, we have shown that the linear TD results produce significant differences when applied to a domain with an elongated scattering region versus a circularly symmetric scattering region. This suggests that using usual TD methods on features that are radially *asymmetric* may produce results with signatures of the anisotropy, and that the shape of the feature must be taken into account. Furthermore, our new linear technique successfully enhances the linear interference pattern from the line of scatterers, while reducing the dimensionality (and therefore computational time) of the TD analysis. Additionally, our method can easily be extended in future research to study cases in which the array of scatterers is not linearly aligned.

We also performed a modified Student’s t -test comparison of the scattered Green’s function results between the 1- and 10-scatterer cases, using a new superpixel method.

These differences are found to be statistically significant. Some work has been done toward comparing the total and incident Green’s function results for the 10-scatterer case, and the t -test results indeed return the scattered wave results (since the t -statistic is essentially a difference). However, the difference is not found to be statistically significant, despite the match to a known result. This most likely has to do with the spatial averaging of our superpixel t -test method, which loses resolution in the local differences between data. In the future, we need to determine alternate ways of calculating the localized t -test probabilities so that we can better calculate and interpret these significance results. In the mean time, this method of statistical analysis is promising for the comparison of differences in our simulated scattering results. We will also apply it to the HMI data in Chapter 5 to see if we can detect statistically significant differences between the quiet sun and our coronal hole data. There is growing interest in the local helioseismology community for the statistical error analysis of data and results, and this method is a step in that direction.

We have not presented the travel time calculations for this simulated data set. One drawback of the parameters selected is that there is no negative branch of the TD results, and therefore we cannot measure a meaningful τ_- . Without both the positive and negative travel times, we also cannot calculate meaningful mean and difference travel times. Thus, the travel times do not contribute much information about the current simulation results. Future work might include adjusting the input parameters such that the incident angle and initial mode phases are randomized, which would then produce both branches of the cross-correlation.

One of the next steps of this research that we are currently investigating is the isolation and enhancement of backscattered signals. This process involves phase-speed shifting, mentioned in Section 3.2.2.1. The phase speeds can be approximated by the calculated travel times. Since the backscattered signal is extremely weak in the solar data but is clear in the simulation, we are again using the simulation as a test case prior to applying the analysis to the HMI data. For this purpose, the measured travel times will be important.

As mentioned above, potential future work with the simulation includes adjusting various input parameters and analyzing the differences in the TD results, expanding from a linear array of scatterers to more general scattering geometries, looking at the statistical differences between various other combinations of parameters, and testing alternative statistical analysis methods for comparing total wave results. We are also currently working on isolating and enhancing the simulated backscattered signal, the method of which we will then apply to our coronal hole data. There are many other possibilities for working with the simulated data, such as incorporating wave attenuation and absorption to mimic the effect of a magnetic field, or adding various noise distributions.

Chapter 5: Wave Scattering Analysis & Results

We now discuss the SDO/HMI Dopplergrams and the application of our linear TD method to solar data. We begin by briefly discussing the data product specifications and the retrieval process. We then consider the linear TD analysis of two different data sets and their results, as well as a statistical comparison between the results.

Similar to the AIA data discussed in Chapter 2, the SDO/HMI data is downloaded from the JSOC archive. Unlike the AIA data, however, the HMI Dopplergrams are already processed and partitioned from full-disk images into smaller mapped regions that cover areas of $30^\circ \times 30^\circ$ over 8-hour intervals. These time-distance products are tracked at the differential rotation rate of the sun, and mapped into heliographic coordinates via Postel's projection. The resultant data cubes have dimensions of $512 \times 512 \times 640$, with a spatial resolution of $0.06^\circ/\text{pixel}$ ($\sim 0.729 \text{ Mm}/\text{pixel}$), and a temporal cadence of 45 s. The data is archived in the JSOC series *hmi.tdVtrack*.

Just as with the AIA data, the exported HMI data is rice-compressed. We again use the *fitsio_read_image.pro* routine to decompress the FITS file and read the data. We then take the Fourier transform of the data using IDL's `FFT(data)` function, and filter the transformed datacube to isolate the *f*-mode using the FTC filter described in Section 3.2.1. The filtered data is then transformed back into the spatial/temporal domain, again with `FFT(data,/INVERSE)`. Our code that performs this process is entitled *filter_yesf.pro*. The uncompressed header information is then updated and saved with the uncompressed, *f*-mode-filtered HMI data. This form of the data is what we use for the duration of the analysis. The routine that encompasses this whole procedure is called *process_HMI_data.pro*.

After performing the post-processing procedure on the corresponding AIA data (Sections 2.1.3-2.1.4), we apply the same rotation and cropping procedure on our HMI data.

We add the rotate and crop parameters to the header metadata and adjust other affected parameters, such as the size of the spatial dimensions. This is the final iteration of the data, on which we will perform our linear TD analysis. We also compute and store the new bounding box information, for reference and to check against the AIA bounding box. This code is called *post_process_HMI.pro*.

We perform our analysis on two different HMI data sets. First, we investigate a quiet sun region as a reference data set (Section 5.1). Next, we perform the analysis on our quasi-linear coronal hole data in Section 5.2. Finally, in Section 5.3, we discuss the statistical differences between the results.

5.1 Quiet Sun Data

5.1.1 HMI Data

The quiet sun data is from November 13, 2013, starting at 12:00 AM. It is centered around Carrington Heliographic longitude 112.9° (corresponding to a Stonyhurst longitude of 0° ; central meridian) and latitude 0° (equator). For reference, the mid-time step AIA data is shown in Figure 5.1 and the corresponding f -mode filtered HMI data is shown in Figure 5.2. We note that this quiet sun data set is observed at a different latitude than our CH data, which could affect the comparative results later in Section 5.3. However, because our method for statistical comparison is an initial attempt and will likely be improved in the future, we are not concerned about the differing latitudes within the scope of this dissertation. We will address this issue in the future.

There is technically no need for the rotation and cropping of the quiet sun data, since there is no boundary feature to align. However, we perform these image processes for consistency when we later compare results with the coronal hole data.

Implementing our linear technique, we sum over the y -dimension of the processed data. Again, this is not specifically necessary for the quiet sun data. We do so 1) to isolate the east-west signal of the random wavefield for comparability with the coronal hole data, 2) for

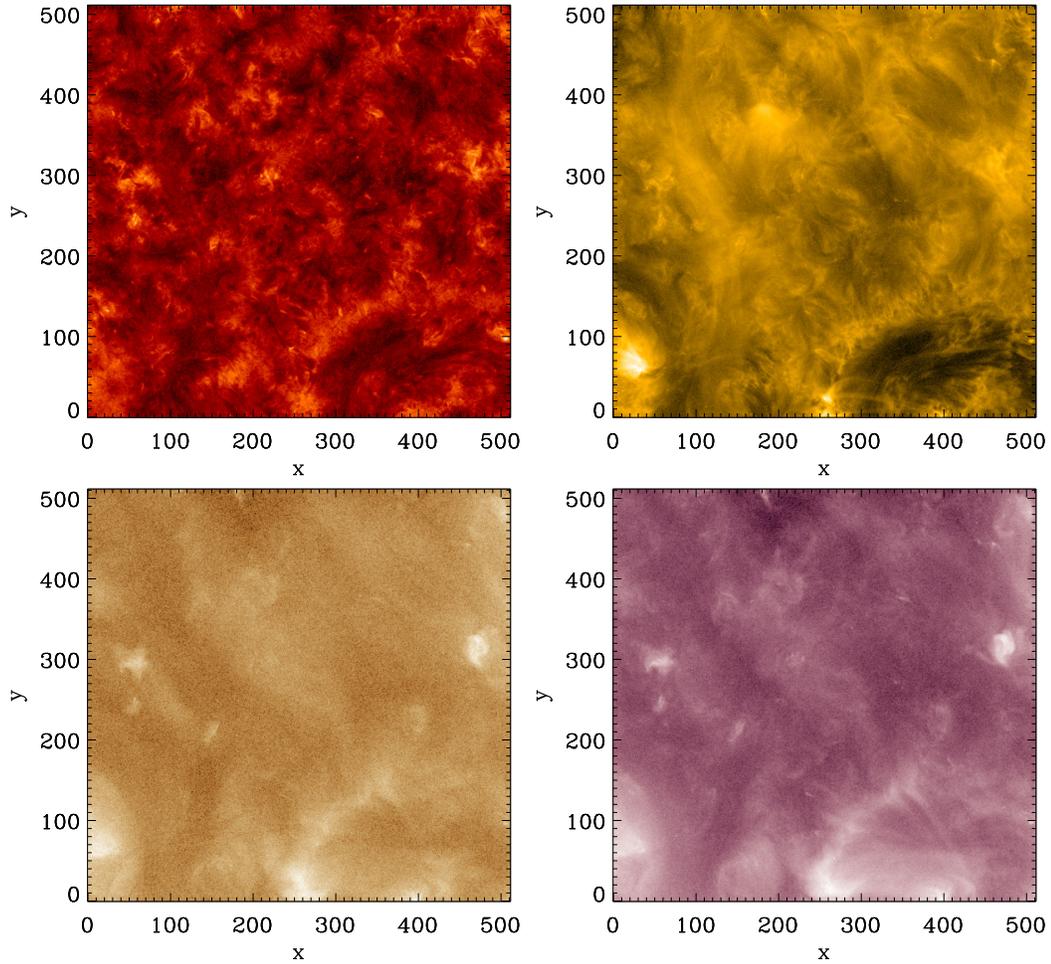


Figure 5.1: SDO/AIA cutout images on Nov. 13 2013 at 4:00AM (middle of observation) in wavelengths 304 Å (top left), 171 Å (top right), 193 Å (bottom left), 211 Å (bottom right). The images shown are \log_{10} scale of the data. Both axes have units of pixels. Recall from Section 2.1.1, 1 AIA cutout pixel = $0.0359^\circ \simeq 0.436$ Mm.

the convenience of working in only one spatial dimension, 3) to improve the signal-to-noise ratio, and 4) since a truly random wavefield should not be affected by an assumed symmetry anyway. As a matter of interest and to compare to the summed Green's function results of the simulation, we show the summed quiet sun data in Figure 5.3.

We perform the shift with respect to various reference lines, as discussed in Section 3.2.2.1. For the quiet sun, we do not care about preserving the location of a feature and we want to

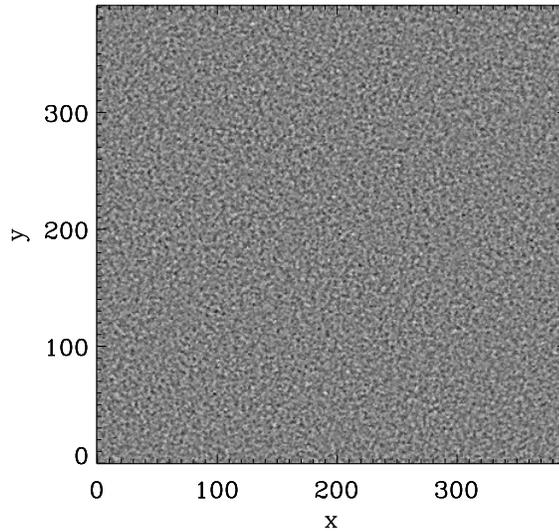


Figure 5.2: SDO/HMI tracked Dopplergram image on Nov. 13 2013 at 4:00AM (middle of observation). The data has already been filtered for the f -mode. Both axes have units of pixels. Recall, 1 HMI pixel = $0.06^\circ \simeq 0.729$ Mm.

improve the signal-to-noise ratio as much as possible. TD results are calculated for all possible reference lines of the data. Each result is then shifted by its reference line location, such that the source-receiver distance Δ becomes the dependent variable. The shifted results are then averaged to produce the final TD result, which we will discuss in the following section. We also measure and present the shifted travel times for the quiet sun data. Theoretically, the shifted travel times of the quiet sun region should be 0-valued throughout, since we are comparing C^0 with C^0 . We will see that this is true for the travel times of the full quiet sun region, but not quiet true for sub-regions of interest.

5.1.2 Time-Distance Results

The averaged, shifted TD result for our quiet sun region is shown in Figure 5.4. Examples of the resultant cross-correlations (for $\Delta = 0, 7, 15$) are shown in Figure 5.5. As expected, the strongest cross-correlation is at $(\Delta, \mathfrak{t}) = (0, 0)$, where the data is being compared to itself with no time lag.

The quiet sun travel times are measured with the method described in Section 3.2.2.2.

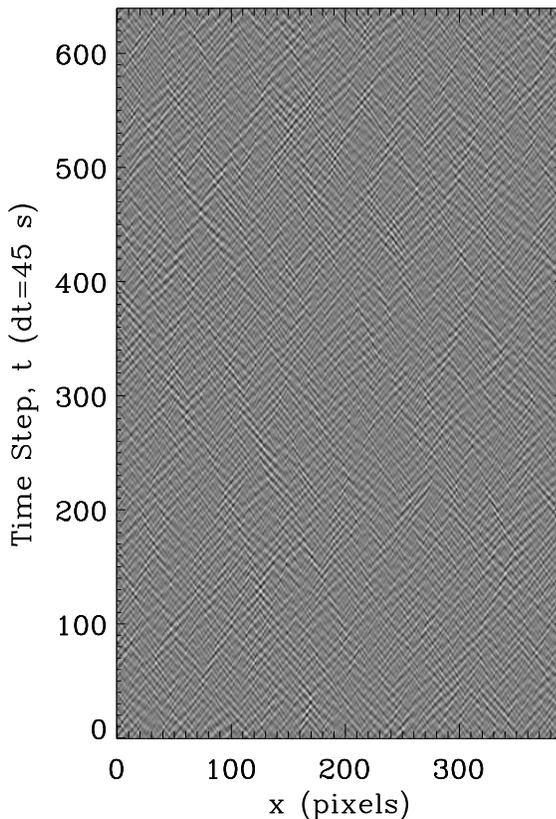


Figure 5.3: The f -mode filtered SDO/HMI quiet sun data, after summing over the y -dimensions. The time step along the y -axis has units of seconds ($dt = 45$ s). The x -axis has units of pixels. Recall, 1 HMI pixel = $0.06^\circ \simeq 0.729$ Mm.

We use a temporal sampling rate $h = 1$ because our lags change in discrete increments ($t_i = ih$), although $dt = 45$ seconds. To measure the unshifted travel times, we use the unshifted TD results as our main cross-correlation term $C(x_1, x_2, t)$ in Equation 3.23, and the shifted cross-correlations described above as $C^0(\Delta, t)$ in Equations 3.23 and 3.24. The unshifted travel time results are shown in Figure 5.6. The shifted travel time results are shown in Figures 5.7, 5.8, and 5.9, and are measured (respectively) using the shifted-average quiet sun TD results from the full data set, a sub-region the size of the CH region, and a sub-region the size of the region outside the CH. This is mainly done for later comparison with the shifted travel times from the CH sub-regions.

The unshifted travel time results are shown in Figure 5.6. The mean positive travel

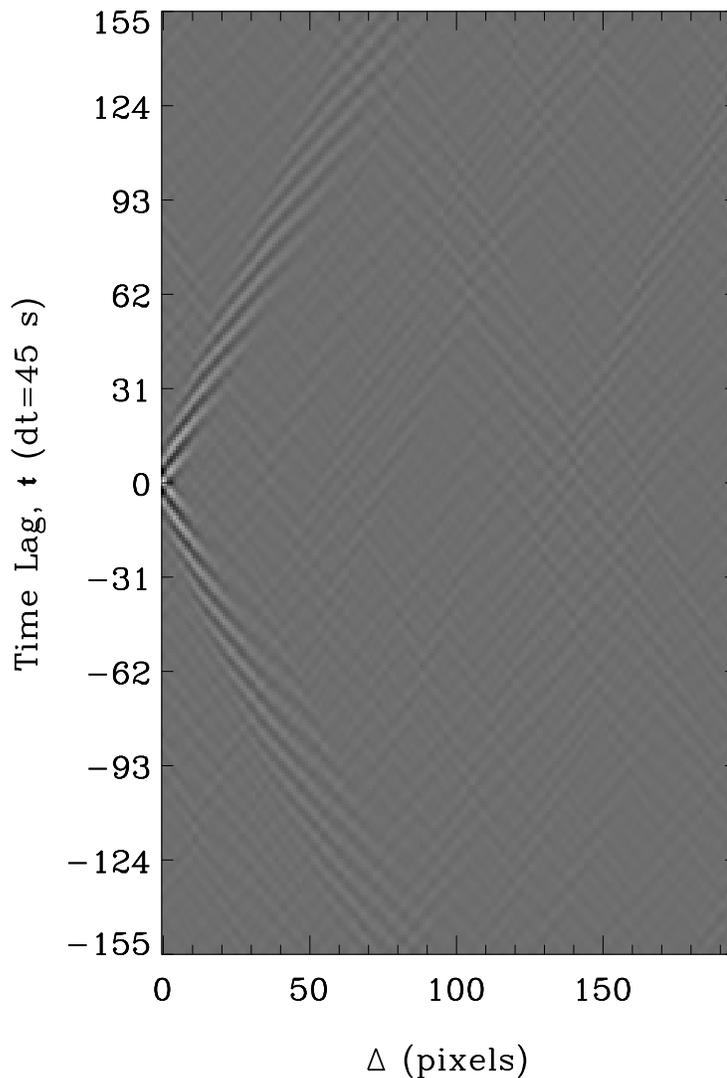


Figure 5.4: The TD plot of the f -mode filtered SDO/HMI quiet sun data, after summing over the y -dimension. Time lag t has units of seconds ($dt = 45$ s). Δ has units of pixels. Recall, 1 HMI pixel = $0.06^\circ \simeq 0.729$ Mm. The TD results were calculated for all possible reference lines, and were subsequently shifted with respect to each reference line. The shifted results were then summed to create this result, with an improved signal-to-noise ratio.

time $\bar{\tau}_+$ is a positive value, while the mean negative travel time $\bar{\tau}_-$ is negative. We expect that the *opposite* of this result might be partially due to solar rotation, as mentioned in Section 3.2.2.2. Yet because we rotated the data, our x -dimension does not directly correspond to an east-west flow, and is more along the lines of “east northeast-west southwest”

in direction. This could account for the apparent “reversal” of underlying flow in our quiet sun results. There seem to be no notable wave-speed perturbations in τ_{mean} . This is expected for a quasi-isotropic and homogeneous quiet sun region. The mean of τ_{diff} is also positive, (as expected based on its definition and the values of $\bar{\tau}_+$ and $\bar{\tau}_-$) which indicates an underlying flow in the $-x$ direction. We note that deviations in these travel time measurements are within the noise level of the results, and the average values of the positive and negative travel times are close to $\bar{\tau}_{\pm} = 0$ within one standard deviation. Also note that there is a larger signal-to-noise ratio closer to $\Delta = 0$, as we might expect from the example $C^0(\Delta, t)$ plots in Figure 5.5. This dependence of signal-to-noise on Δ should become more significant for noisier travel time measurements.

We reiterate that for a *non-rotated* data cube, the results in the $\pm x$ direction would correspond to the east-west direction. An underlying flow in the $+x$ direction might then be indicative of solar rotation, which rotates from east to west. However, when interpreting our travel time results, it is important to remember that our data is rotated. Therefore, our $\pm x$ direction is more akin to a “east northeast-west southwest” flow.

Recall that τ_{diff} is sensitive to variations in flows and τ_{mean} is sensitive to sound speed perturbations. Therefore, the physical interpretations of our unshifted quiet sun results are that 1) there are not strong differences in the plasma flows through which the f -mode waves are traveling in the quiet sun and 2) there are not cogent sound speed perturbations, such as drastic changes in temperature, in this quiet sun region. Of course, these results only consider flows and wave-speed perturbations in the $\pm x$ direction due to our use of the linear TD technique. However, when looking at elongated boundary features, these effects should be most notable in the direction perpendicular to the boundaries.

Although we mainly need the shifted travel times for comparison with the similar CH results, we can still consider some stand-alone implications of the shifted quiet sun travel time results. Figure 5.7 shows the measured travel time variations over the full quiet sun domain, which are non-existent. These plots may seem uninteresting because there is no variation in the travel times. However, this 0-result is due to the fact that we are essentially

comparing our $C^0(\Delta, \mathbf{t})$ to itself. We expect 0-variation when making this comparison, so these results act as a sanity check of our travel time calculations. Also, this leads us to the conclusion that the variations seen in Figures 5.8 and 5.9 are indicative of the noise that is introduced when we compute the travel times in sub-regions using the $C^0(\Delta, \mathbf{t})$ derived from the average of the full dimension of the TD data (in other words, when $C(\Delta, \mathbf{t})$ and $C^0(\Delta, \mathbf{t})$ have different signal-to-noise ratios).

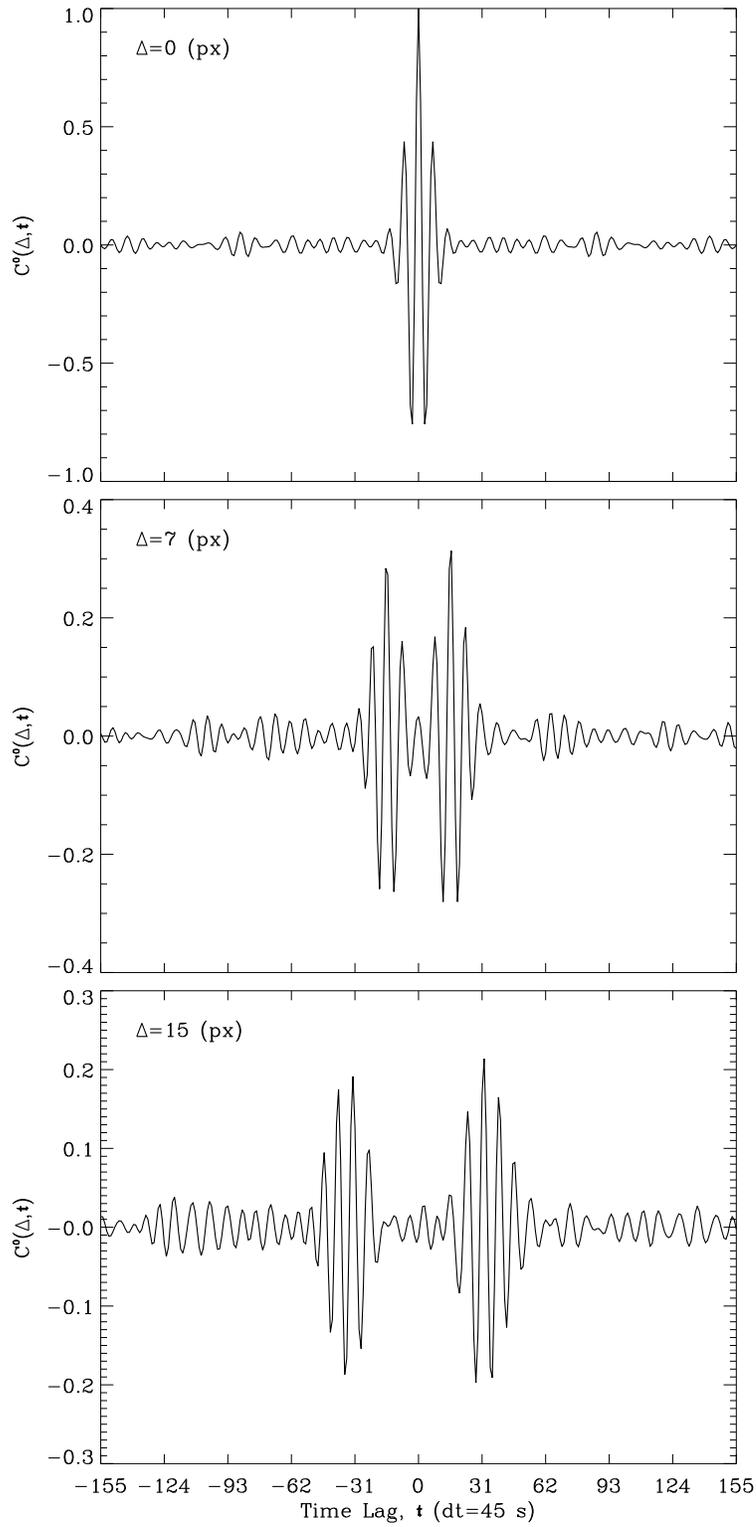


Figure 5.5: Example reference cross-correlations, derived from the quiet sun data, used in the travel time measurements. Time lag t has units of seconds ($dt = 45$ s). Δ has units of pixels. Recall, 1 HMI pixel = $0.06^\circ \simeq 0.729$ Mm.

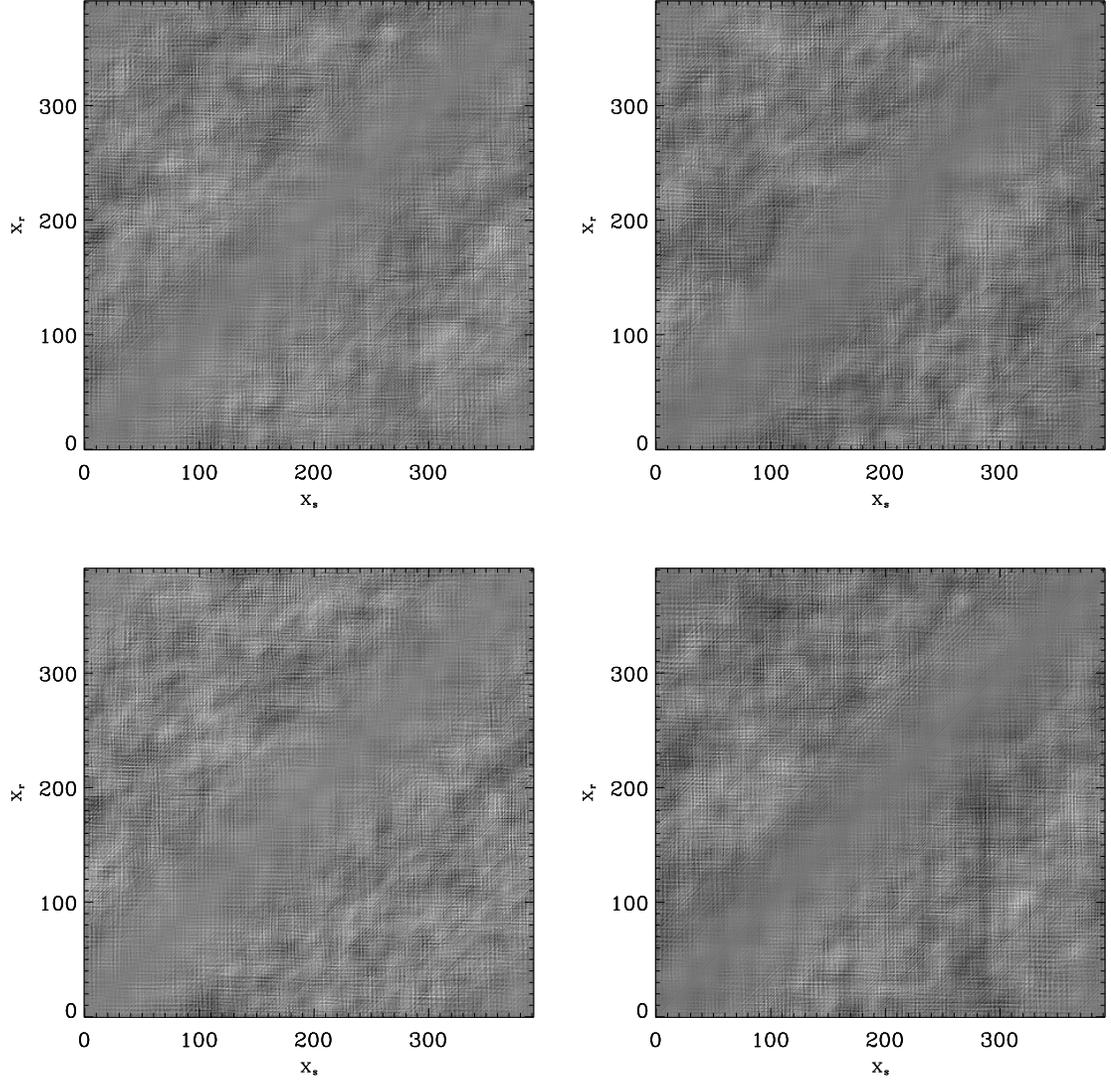


Figure 5.6: The $\tau_+(x_s, x_r)$ (top left), $\tau_-(x_s, x_r)$ (top right), $\tau_{\text{mean}}(x_s, x_r)$ (bottom left), and $\tau_{\text{diff}}(x_s, x_r)$ (bottom right) travel time measurements from the quiet sun data. Both axes have units of pixels. Recall, 1 HMI pixel = $0.06^\circ \simeq 0.729$ Mm. Travel times are indicated in greyscale, with black and white corresponding to shortest and longest times, respectively. We note that there is a visible pattern of horizontal and vertical lines; as of yet, we cannot conclude about the source of the pattern, but it is under investigation.

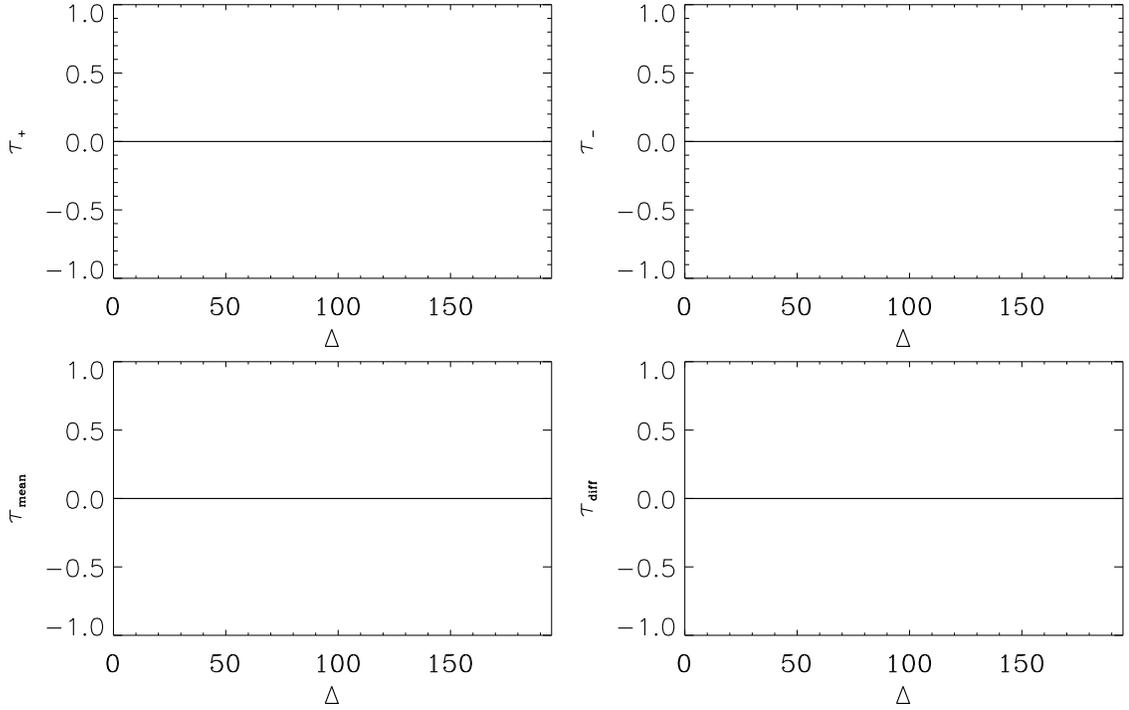


Figure 5.7: The $\tau_+^{\text{all}}(\Delta)$ (top left), $\tau_-^{\text{all}}(\Delta)$ (top right), $\tau_{\text{mean}}^{\text{all}}(\Delta)$ (bottom left), and $\tau_{\text{diff}}^{\text{all}}(\Delta)$ (bottom right) travel time measurements from the shifted TD results of the full quiet sun data. Travel times have units of seconds ($dt = 45$ s). Δ has units of pixels. Recall, 1 HMI pixel = $0.06^\circ \simeq 0.729$ Mm. These plots may seem uninteresting because there is no variation in the travel times. However, this is because we are essentially comparing our $C^0(\Delta, \mathbf{t})$ to itself. We expect 0-variation when making this comparison, so these results act as a sanity check of our travel time calculations. Also, this leads us to understand that the variations seen in Figures 5.8 and 5.9 are indicative of the noise that is introduced when we compute travel times in sub-regions when $C(\Delta, \mathbf{t})$ and $C^0(\Delta, \mathbf{t})$ have different signal-to-noise ratios.

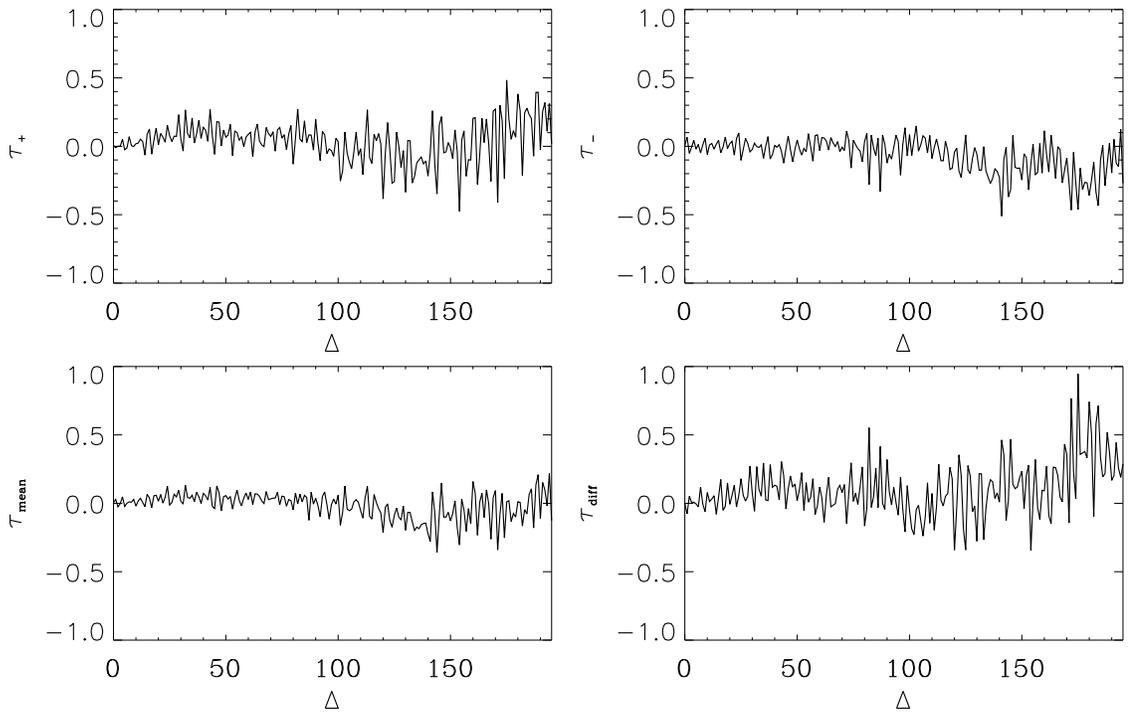


Figure 5.8: The $\tau_+^{\text{in}}(\Delta)$ (top left), $\tau_-^{\text{in}}(\Delta)$ (top right), $\tau_{\text{mean}}^{\text{in}}(\Delta)$ (bottom left), and $\tau_{\text{diff}}^{\text{in}}(\Delta)$ (bottom right) travel time measurements from the shifted TD results of the quiet sun region, for a sub-region of the size of the CH region in Section 5.2. Travel times have units of seconds ($dt = 45$ s). Δ has units of pixels. Recall, 1 HMI pixel = $0.06^\circ \simeq 0.729$ Mm. These travel times are for reference and later comparison with the shifted travel times in Figure 5.13. They also give some idea of the noise difference in travel times measured using the shifted-average of the full dimensional data set, and those obtained from the sub-region.

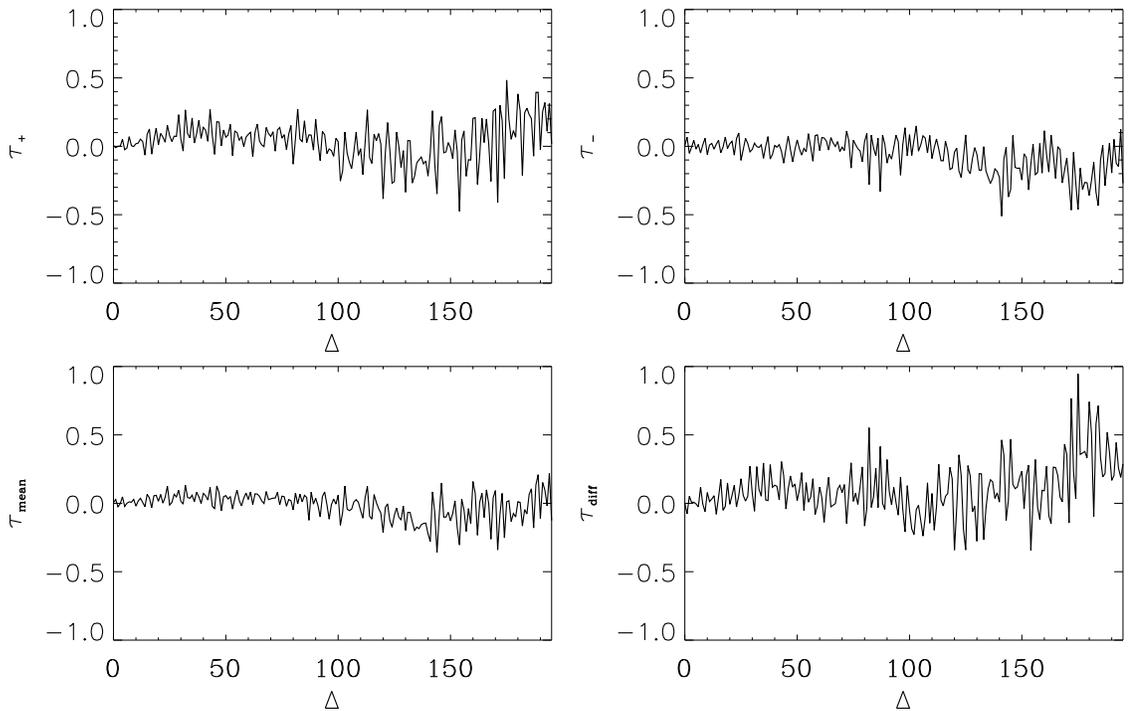


Figure 5.9: The τ_+^{out} (top left), τ_-^{out} (top right), $\tau_{\text{mean}}^{\text{out}}$ (bottom left), and $\tau_{\text{diff}}^{\text{out}}$ (bottom right) travel time measurements from the shifted TD results of the quiet sun region, for a sub-region of the size of the region outside the CH in Section 5.2. Travel times have units of seconds ($dt = 45$ s). Δ has units of pixels. Recall, 1 HMI pixel = $0.06^\circ \simeq 0.729$ Mm. These travel times are for reference and later comparison with the shifted travel times in Figure 5.14. They also give some idea of the noise difference in travel times measured using the shifted-average of the full dimensional data set, and those obtained from the sub-region.

5.2 Coronal Hole Data

5.2.1 HMI Data

The CH data is from February 5, 2014, starting at 8:00 AM. It is centered around Carrington Heliographic longitude 82° (corresponding to a Stonyhurst longitude of 0° ; central meridian) and latitude 24° (equator). The HMI data corresponds to the AIA data discussed in Sections 2.1 and 2.2. This data is visibly indistinguishable from that shown in Figure 5.2, so we will not share an example time step.

Having performed the rotation and cropping designated by the EUV results, we must convert the detected CH and width boundary locations from AIA pixels to HMI pixels. As mentioned in Section 2.1.1, the HMI data has a different pixel scale than that of the AIA data. However, both data sets, by our design, are in Postel-projected heliographic coordinates and cover the same bounding box of the solar disk. The heliographic locations of the CH inflection point and boundaries are the same regardless of the data set. It is therefore a somewhat trivial task to convert the AIA pixel locations into heliographic coordinates, and subsequently those coordinates into HMI pixel locations. To do so, we use some of the WCS routines already outlined in Appendix A.1. Namely, we use the SSW function *wcs_get_coord.pro* on the specific AIA pixels with the AIA WCS structure to retrieve the associated heliographic coordinates, followed by *wcs_get_pixel.pro* on those coordinates with the HMI WCS structure to get the associated HMI pixels. The HMI WCS structure is produced from the FITS header using the function *fitshead2wcs.pro*. We again note (as we did in Chapter 2) that the AIA WCS structure must be corrected after downloading that data series, as shown in Appendix A.2. Other than updating the AIA metadata, the procedure for converting the AIA pixels to HMI pixels is encoded into our routine *convert_aia2hmiloc.pro*. After running the conversion code, we find the necessary HMI locations as follows: a central CH location at $x = 268$, a lower width boundary location at $x = 260$, and an upper width boundary at $x = 276$. It is interesting to note that the boundary width is still symmetric about the detected center, which was one of our original

criteria for choosing width-detection parameters from the EUV data in Section 2.2.

Again, we perform a summation over the y -dimension of the CH data to exploit the assumed translational invariance of the boundary. There is still no apparent distinction between the CH data and the quiet sun data at this point. The same is true of the TD results of the summed CH data. However, as mentioned in Section 3.2.2.1, if we take the shifted average of these TD results, we lose the specific spatial information with respect to the CH boundary. In order to compare the TD results within the CH region to the nearby quiet sun or the completely unrelated quiet sun results of Section 5.1.2, we split the unshifted CH data at the boundary. We can then choose to take a shifted average inside or outside the hole for comparison, or compare the unshifted regions (full set or sub-regions) with reference cross-correlations.

We add here that while we cannot shift the CH TD results with respect to all available reference lines without losing the boundary location, we will consider alternatives in the future backscatter-detection project mentioned at the end of Chapter 4, which is outside the scope of this dissertation. We plan to shift the data with respect to the detected CH boundary in order to enhance the scattered signal in the interaction region of the CH boundary, after testing this method on the simulated backscattered signal.

5.2.2 Time-Distance Results

The shifted TD plots of the CH data are shown in Figure 5.10, for 1) an averaged shift across the entire domain, 2) that for reference lines located inside the CH region, and 3) similarly for the region outside the CH boundary. We note that the shifting of the inner and outer data is bounded by the width locations, as opposed to the detected-center location of the CH boundary. As usual, the distinctions between the three cases are difficult or impossible to see. However, upon comparison of the inside-CH and outside-CH results, it is possible that the negative branch of the TD signal is attenuated more quickly and at shorter distances in the region outside the CH than within the CH region. We will investigate this possible difference further in Section 5.3.

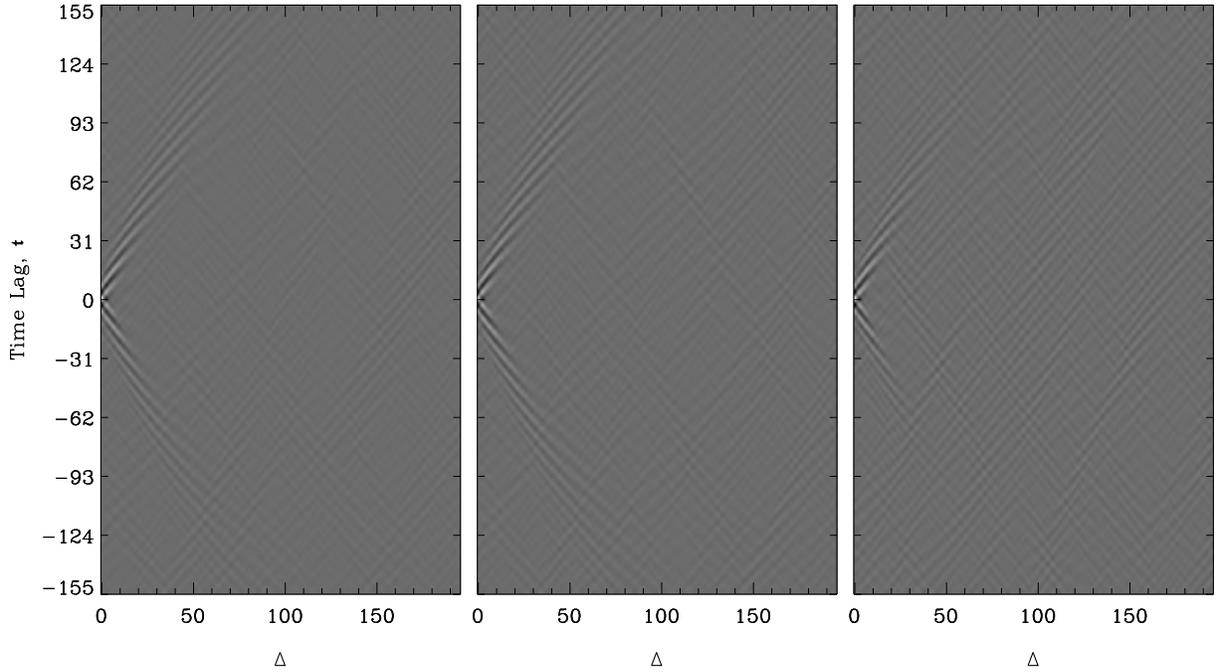


Figure 5.10: Averaged, shifted TD plots of the f -mode filtered SDO/HMI CH data, after summing over the y -dimension. Time lag t has units of seconds ($dt = 45$ s). Δ has units of pixels. Recall, 1 HMI pixel = $0.06^\circ \simeq 0.729$ Mm. The left-most TD result are calculated from all possible reference lines, ignoring the location of the CH boundary. The center plot shows the average shifted result from only those reference lines inside the CH region. The right-hand result is similarly that of the region outside the CH boundary. The three cases are highly similar, although it is possible that the TD signal in the outer region is attenuated faster than inside the CH. This difference will be investigate further in Section 5.3.

We will not show examples of the unshifted TD data, although we do use the three cases (“in”, “out”, and “all”) of the unshifted and shifted data to measure travel times. We use the same $C^0(\Delta, t)$ as was used to measure the travel times for the quiet sun data in Section 5.1.2, which was derived from that data set (see Figure 5.5).

We do not expect to see strong implications of horizontal flows or wave-speed perturbations in the travel times because CHs are not strongly magnetic or turbulent features. However, as discussed in Section 1.1, CHs are known to have outflows that cause decreased density and electron temperature in the solar atmosphere. It is uncertain whether these properties are noticeable at the surface but if so, they may manifest as differences in the measured travel times or attenuation rate between the CH region and the quiet sun.

For the unshifted travel time results, we find travel times dependent on both the source and receiver points, similar to those measured for the quiet sun data. The travel times measured across the entire domain of the CH data are shown in Figure 5.11, with indications of the CH boundary location. Compared to the quiet sun travel times in Figure 5.6, these results appear slightly noisier throughout. This is likely due to our use of reference cross-correlations, which fit much of the noise in the related quiet sun data. Using the same reference with completely unrelated data, even another quiet sun region, will propagate the uncorrelated noise variations through into the travel times. We also still see a dependence of the signal-to-noise ratio on Δ , as discussed in Section 5.1.2.

As predicted above, these unshifted results do not show any substantial flows or wave-speed perturbations. There could be differences between the travel times of the waves inside and outside the CH boundary. We are currently working on ways to test for the significance of any differences between the inner and outer unshifted travel times. We will not consider this issue in this dissertation, but it is a current topic of study. However, we can state that the average $\bar{\tau}_+$ and $\bar{\tau}_-$, as well as $\bar{\tau}_{\text{diff}}$, taken over the entire domain indicate a slight uniform flow in the $+x$ direction, as opposed to our quiet sun results in Section 5.1.2. Those averages taken specifically for the travel times within and outside of the CH region indicate the same result. We note again, however, that the standard deviations for all these results are at least one order of magnitude larger than the calculated means, so we cannot conclude anything about these flow results with certainty.

The travel times between two shifted data sets are not usually measured. However, if we want to look for very subtle differences between the travel times inside and outside the CH region, it seems that improving the signal-to-noise as much as possible is desirable. The resulting travel times will only be a function of Δ instead of combinations of sources and receivers. With our linear TD method, Δ only describes distances along the $\pm x$ direction, so the resulting travel times can be interpreted similarly to the usual “unshifted” travel times.

The results of the shifted travel time measurements are shown in Figures 5.12, 5.13, and

5.14 for the full domain, the region within the CH boundary, and the outer quiet sun region, respectively. Again, the signal-to-noise ratio is largest when $\Delta = 0$. It is difficult to see from the plots what the travel times are suggesting. However, we can use the mean values of the plots to inform ourselves about overall flows and wave-speed perturbations present in the specified regions. We will also investigate the significance of differences in the shifted results in Section 5.3.

For the shifted travel times through the full domain (Figure 5.12), the mean of the travel time difference again suggests an underlying flow propagating in the $+x$ direction. There also seems to be some wave-speed perturbation affecting travel times between distances of $150 < \Delta < 190$ in the τ_{mean} results. Similarly in the CH region (Figure 5.13), the mean of τ_{diff} also indicates a flow in the $+x$ direction. The wave-speed perturbation around $150 < \Delta < 190$ is also still present in τ_{mean} , as well as another perturbation near $\Delta = 90$. Interestingly, the mean flow outside of the CH (Figure 5.14) is in the $-x$ direction, and the wave-speed perturbation between $150 < \Delta < 190$ seems to be a much more distinctive feature. There is also a perturbation outside the CH near $\Delta = 80$, which could be affiliated with the slight perturbation inside the CH at about the same distance. We point out that the perturbation outside the CH has a negative effect on τ_{mean} , while the perturbation inside the CH has a positive effect.

We note again that the mean flow results from the shifted travel times should be considered with some skepticism because the standard deviations of all these results are large enough to cross a 0-mean value. On the other hand, the apparent wave-speed perturbation features fall outside of one standard deviation in all three cases of τ_{mean} , suggesting that the perturbations are of some significance. We investigate the significance of differences between results inside and outside the CH, and compared with the quiet sun shifted travel times, in the next section.

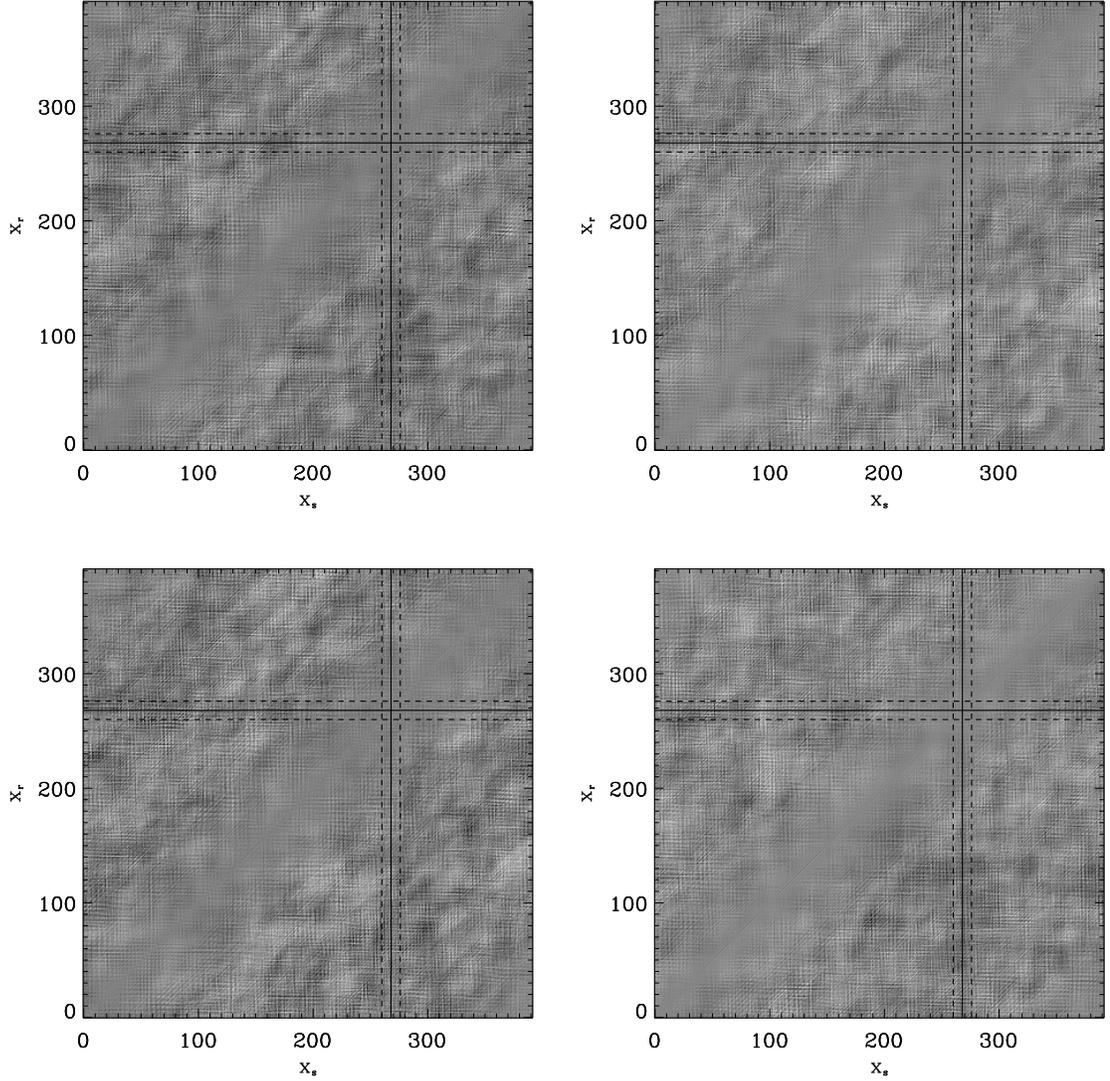


Figure 5.11: The $\tau_+(x_s, x_r)$ (top left), $\tau_-(x_s, x_r)$ (top right), $\tau_{\text{mean}}(x_s, x_r)$ (bottom left), and $\tau_{\text{diff}}(x_s, x_r)$ (bottom right) travel time measurements from the full CH data. Both axes have units of pixels. Recall, 1 HMI pixel = $0.06^\circ \simeq 0.729$ Mm. Travel times are indicated in greyscale, with black and white corresponding to shortest and longest times, respectively. The solid lines represent the location of the center of the CH boundary, while the dashed lines indicate the width of the boundary. Where (x_s, x_r) are below/left of the CH boundary, the travel times are measured within the CH region with respect to reference lines inside the region. Likewise for (x_s, x_r) are above/right of the CH boundary, the travel times are calculated outside of the CH region with respect to reference lines outside the region. Similarly, for x_s inside (outside) the boundary and x_r outside (inside) the boundary, the travel times are measured outside (inside) the CH region with respect to signals inside (outside) the CH region. We again note that there is a visible pattern of horizontal and vertical lines; as of yet, we cannot conclude about the source of the pattern, but it is under investigation.

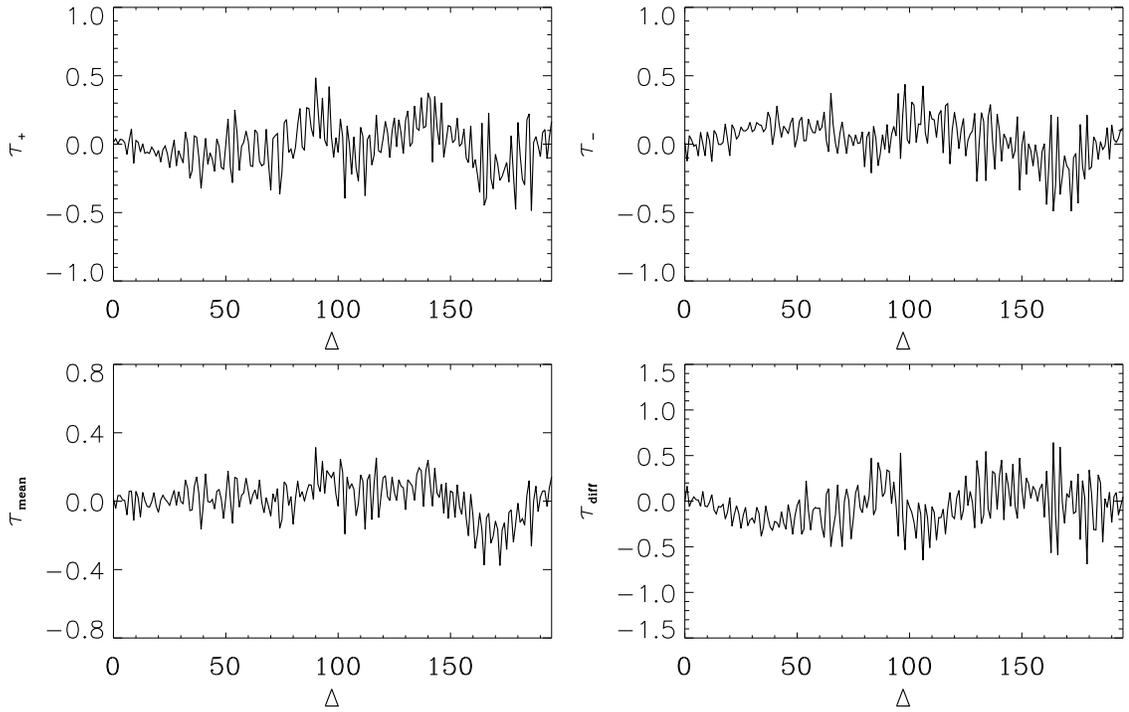


Figure 5.12: The $\tau_+^{\text{all}}(\Delta)$ (top left), $\tau_-^{\text{all}}(\Delta)$ (top right), $\tau_{\text{mean}}^{\text{all}}(\Delta)$ (bottom left), and $\tau_{\text{diff}}^{\text{all}}(\Delta)$ (bottom right) travel time measurements from the shifted TD results of the full CH data. Travel times have units of seconds ($dt = 45$ s). Δ has units of pixels. Recall, 1 HMI pixel = $0.06^\circ \simeq 0.729$ Mm.

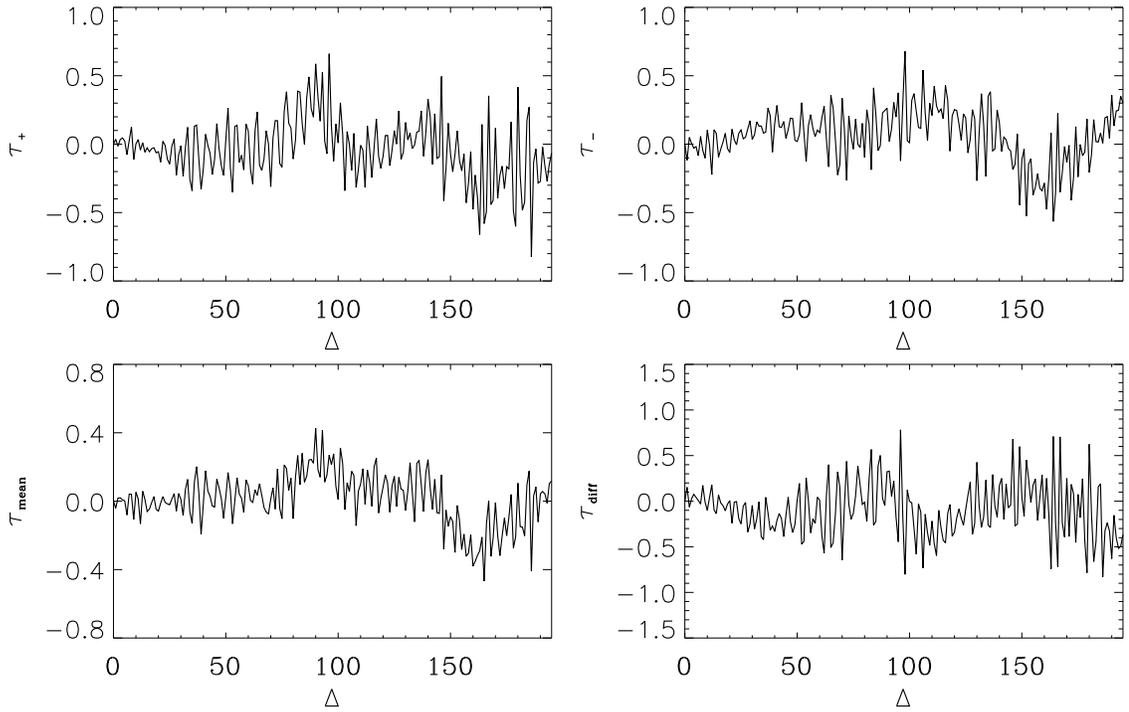


Figure 5.13: The $\tau_+^{\text{in}}(\Delta)$ (top left), $\tau_-^{\text{in}}(\Delta)$ (top right), $\tau_{\text{mean}}^{\text{in}}(\Delta)$ (bottom left), and $\tau_{\text{diff}}^{\text{in}}(\Delta)$ (bottom right) travel time measurements from the shifted TD results inside the CH region. Travel times have units of seconds ($dt = 45$ s). Δ has units of pixels. Recall, 1 HMI pixel = $0.06^\circ \simeq 0.729$ Mm.

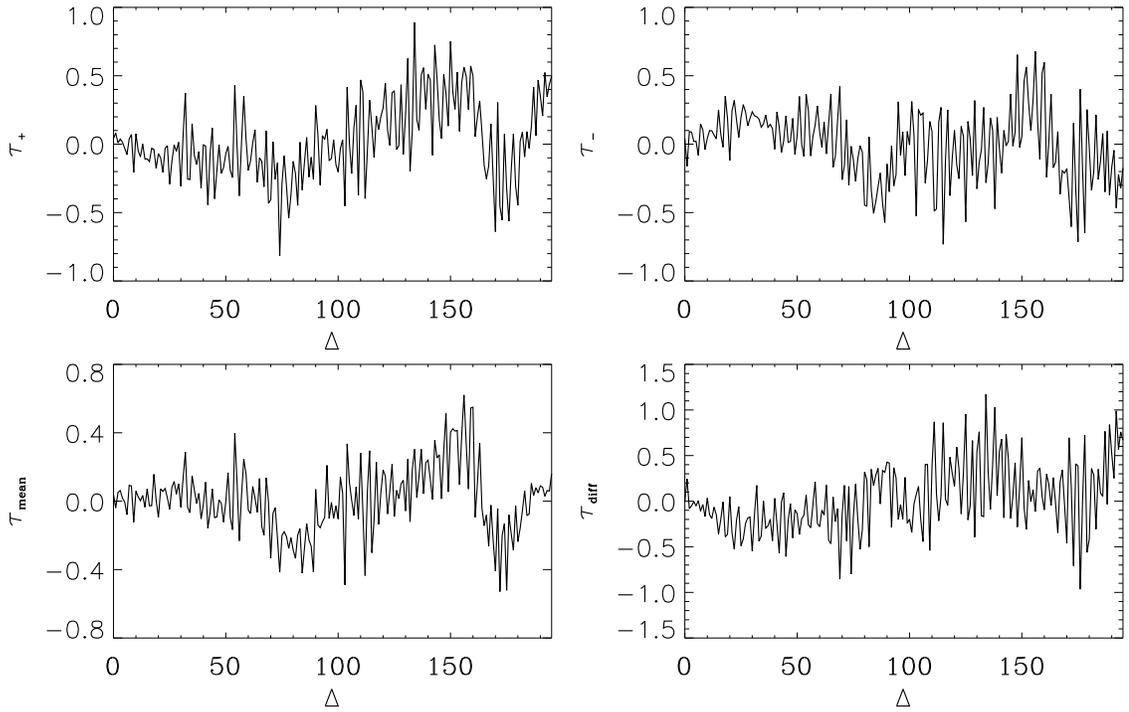


Figure 5.14: The τ_+^{out} (top left), τ_-^{out} (top right), $\tau_{\text{mean}}^{\text{out}}$ (bottom left), and $\tau_{\text{diff}}^{\text{out}}$ (bottom right) travel time measurements from the shifted TD results outside of the CH region. Travel times have units of seconds ($dt = 45$ s). Δ has units of pixels. Recall, 1 HMI pixel = $0.06^\circ \simeq 0.729$ Mm.

5.3 Statistical Comparisons

As done for the simulation analysis in Chapter 4, we want to statistically compare the TD results obtained within the CH region to those from outside the CH boundary. First, we will perform our “superpixel” Student’s t -test analysis on the shifted TD results from Figures 5.4 and 5.10. Then we will use the same method to compare the shifted travel time results from Figures 5.12, 5.13, and 5.14 with the similar quiet sun results in Figures 5.7, 5.8, and 5.9. For the shifted travel times, the superpixel method translates to a comparison of smoothed running-averages of the 1D data.

We present four statistical comparison cases between the shifted-average TD results. The basic differences between TD results in the four cases are mapped in Figure 5.15, for comparison with the related superpixel t -test results. The statistical comparison between the regions inside and outside of the CH boundary are shown in Figure 5.16; the full CH data set versus the separate quiet sun data is given in Figure 5.17; Figure 5.18 displays the significance of differences between the CH region and the unrelated quiet sun, and the comparison of the quiet sun region near the CH boundary to the unrelated quiet sun region is shown in Figure 5.19. We note that the comparisons are made between like-sized sub-regions. For each case, we show both the 3×3 superpixel results and the 5×5 superpixel results. The latter are more highly smoothed, and therefore have less spatial resolution than the former; however, the increased number of data points used in the calculations improve the statistics of the comparison.

We note that in the raw differences of the shifted TD data in Figure 5.15, there appear to be numerous strong difference attributes. However, as mentioned in Section 4.2.3, the superpixel method of comparison uses localized averaging, which loses some resolution in the data. Therefore, certain features are clearly averaged out in the t -test results, including the main TD branches, and no comparison of the small-scale differences between data sets can be made. On the other hand, this highlights regions with large-scale significant differences, and we can also compare more general properties of the main TD branch regions. Finally, while the statistical results presented below are interesting, their causes and implications

are as yet unclear. The application of hypothesis testing to TD results is a novel approach, and further analysis needs to be done in the future to understand the full implications of the statistical results. In light of this reservation, we will only point out some of the more obvious and interesting results of the comparison cases.

Between the CH region and the nearby quiet sun (Figure 5.16), there is one region that has large-scale significant differences, around $\Delta = 75$ and $t = 45$. It is difficult to interpret the meaning of this feature, since it does not appear to be a strong feature in the raw “In vs Out” difference image in Figure 5.15 and the smoothed statistical analysis interrupts the feature in the t -statistic results. However, it seems that the feature propagates to both shorter and longer distances, and into longer time lags. It is also not present in any of the other shifted TD comparison cases, implying that this significant difference is unique between the regions inside and outside of the CH.

Furthermore, we point out that between the averaged-out main TD branches in Figure 5.16, the negative branch of the main TD signal is truncated in both Δ and t , as compared to the positive branch. The difference is apparent in both the t -statistic map and the significance mask. This could indicate greater attenuation or absorption of waves propagating in the $-x$ direction versus waves traveling in the $+x$ direction. This result is not apparent in the comparisons between the full CH region and the quiet sun region (Figure 5.17) or between the region inside the CH and the quiet sun (Figure 5.18). It seems to be present in the comparison between the region outside the CH and the unrelated quiet sun (Figure 5.19), but is less distinct. This suggests that the difference in $-x$ wave propagation is unique to the quiet sun region just outside of our CH, and may be indicative of the influence on waves entering a CH region. Future case studies would be useful for validating this conclusion.

Collectively, these results suggest the presence of a wave front propagating through the quiet sun toward the CH boundary, which is somehow disrupted in the boundary region. If it were reflected by the boundary, we would “expect” to see some backscattered signal at the boundary location. However, recall that shifting the TD data obscures the location of

the boundary, so any backscattered signal is likely averaged out. It is also possible that the wave is partially absorbed by the open magnetic field when it crosses the CH boundary, so that the weakened propagating wave does not show up as a statistically significant difference between the CH region and the unrelated quiet sun.

As mentioned above, these statistical results are interesting but their causes and implications are unclear as of yet. Further analysis is needed in the future to understand the full implications of these results. Furthermore, we have not investigated the presence or potential cause of a wavefront propagating towards the CH. It is possible that some nearby solar activity produced such a wave, but we leave that determination to future work. We also note that while we used our superpixel Student’s t -test method on this data, it is also possible to use the traditional t -test considering that the shifted TD results are averages of a set of unshifted data. We have performed an initial investigation into this form of the analysis, the results of which support the conclusions presented above and possibly indicate the presence of backscattering. However, that analysis is still currently under investigation and is beyond the scope of this dissertation.

We now move on to the statistical comparison between the shifted travel time measurements from the quiet sun and the various CH sub-regions. Recall that these travel times were measured using the shifted TD results discussed above. They are therefore only dependent on Δ . Again, the signal-to-noise ratio is largest close to $\Delta = 0$. The “superpixel” t -test method used for the shifted TD comparisons translates to a smoothed running-average of the 1D shifted travel times. Multiple smoothing widths were tried, but only statistical comparisons using a width of 7 “pixels” (3 data points on either side of a central point) are shown. This smoothing value was chosen to improve the statistics, without losing too much spatial resolution. Only the final $\geq 95\%$ masked t -statistic results are presented.

In Figures 5.20-5.23, we look at the thresholded significance of difference between the measured travel times between different sub-regions. Each plot shows the significant differences between its specified travel time measurement (i.e., the τ_{mean} and τ_{diff} plots are not the average and difference of the τ_+ and τ_- significant difference plots, but the compared

significant differences between the τ_{mean} and τ_{diff} results of two sub-regions).

In Figure 5.20, the statistically significant differences between the in-CH and out-CH travel times indicate that the difference in the underlying flows is significant, and there is also a significant difference in wave-speed perturbation between $70 \leq \Delta \leq 100$. This is consistent with the results of Figures 5.13 and 5.14. Interestingly, the wave-speed perturbation around $\Delta = 150$ is similar between the two regions and the difference is relatively small (although still found to be statistically significant). This is likewise the case for the other sub-region comparisons.

In Figure 5.21, the statistically significant differences in travel times between the full CH and quiet sun regions are much smaller than the comparisons between other sub-regions. This is most likely due to the fact that the full quiet sun travel times in Figure 5.7 are exactly 0, which greatly affects the smoothed hypothesis test. If we accept the results of this comparison as real, they indicate a difference in flows effecting small distances and a possible wave-speed perturbation at $\Delta = 100$ that might correspond to the significant wave-speed perturbation in the in-CH versus out-CH comparison.

In Figure 5.22, the t -test results comparing the region inside the CH with the unrelated quiet sun region shows a significant net difference in flow, particularly for distances up to $\Delta = 50$. It also seems to indicate a strong wave-speed perturbation both around $\Delta = 90$ (similar to those seen in the other comparisons), and another strong perturbation near $\Delta = 20$ that was not apparent in the individual travel times (Figures 5.8 and 5.13). The cause of this significant perturbation is currently unidentified, and is an item for future investigation. It is possible that this feature is a signature of the wave propagation through a CH versus a quiet sun region, particularly since it is not found as a strong feature in the other travel time comparisons.

Finally, Figure 5.23 compares the quiet sun region outside of the CH to the unrelated quiet sun region. Again, there is a strong indication of a wave-speed perturbation difference near $\Delta = 80$, and a small but significant difference in flow speed between the independent quiet sun region and the near-CH quiet sun region. This is in agreement with the individual

travel time results found in Figures 5.14 and 5.9.

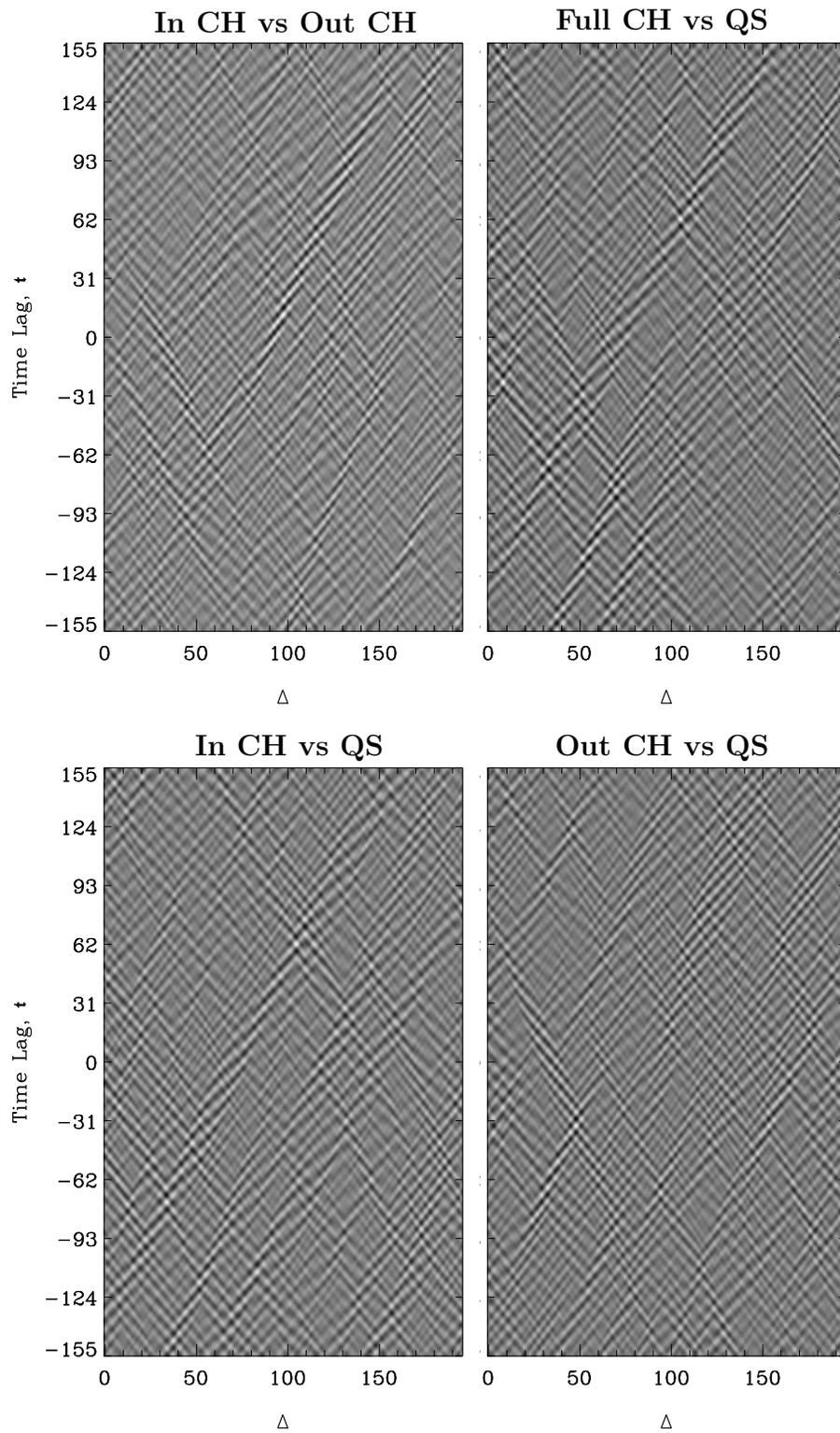


Figure 5.15: Basic difference plots between the four sets of compared TD results. Δ has units of pixels, and t has units of seconds ($dt = 45$ s). These plots are mainly for reference against the superpixel-average t -statistic results in Figures 5.16-5.19.

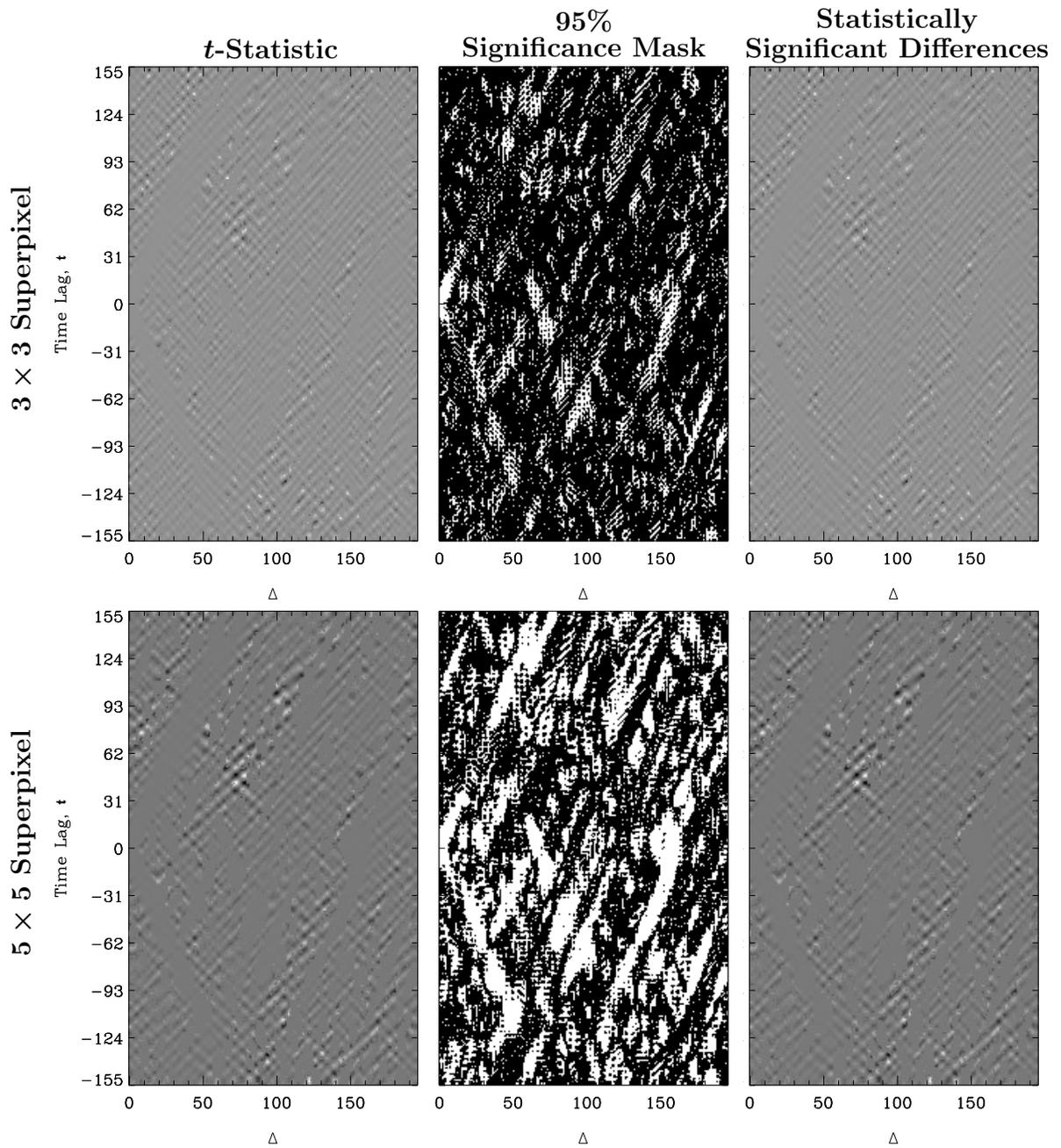


Figure 5.16: The Student’s t -test comparison between the shifted TD results inside and outside the CH boundary. Δ has units of pixels, and t has units of seconds ($dt = 45$ s). The top row of results employs a 3×3 superpixel; the bottom row, a 5×5 superpixel. Full maps of the t -statistic are shown on the left. The center images show the binary $\geq 95\%$ significance maps of the left-hand results. Black indicates statistically significant differences and white indicates otherwise. We threshold the t -test results on the left with the center significance masks to produce the plots on the right. These results show only the significantly different TD signals between the inner and outer regions of the CH boundary. Note that small-scale differences between the main TD branches are averaged out with this technique.

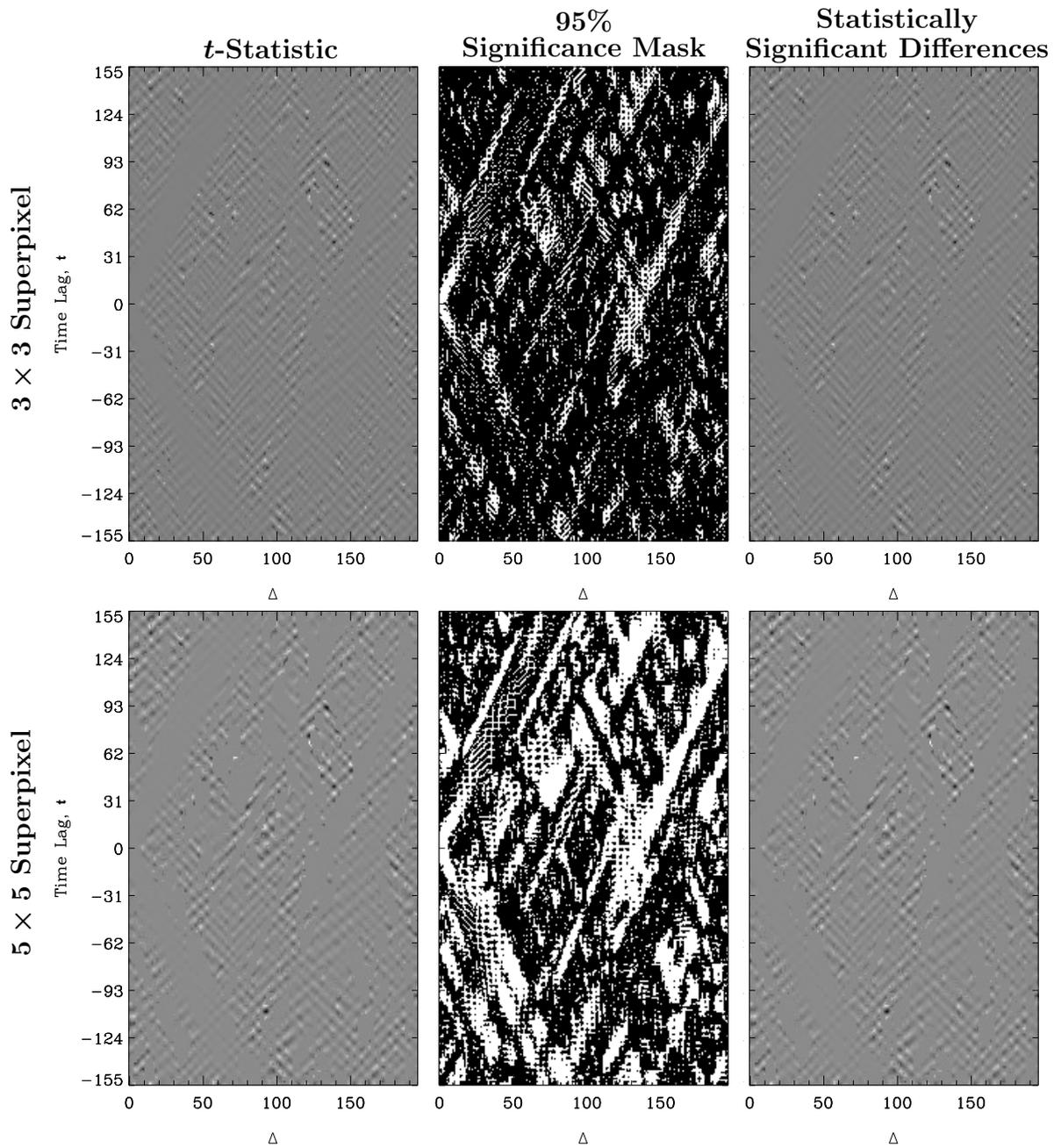


Figure 5.17: The Student’s t -test comparison between the shifted TD results of the full CH data set and separate quiet sun data. Δ has units of pixels, and t has units of seconds ($dt = 45$ s). The top row of results employs a 3×3 superpixel; the bottom row, a 5×5 superpixel. Full maps of the t -statistic are shown on the left. The center images show the binary $\geq 95\%$ significance maps of the left-hand results. Black indicates statistically significant differences and white indicates otherwise. We threshold the t -test results on the left with the center significance masks to produce the plots on the right. These results show only the significantly different TD signals between the full CH data and the quiet sun data. Note that small-scale differences between the main TD branches are averaged out with this technique.

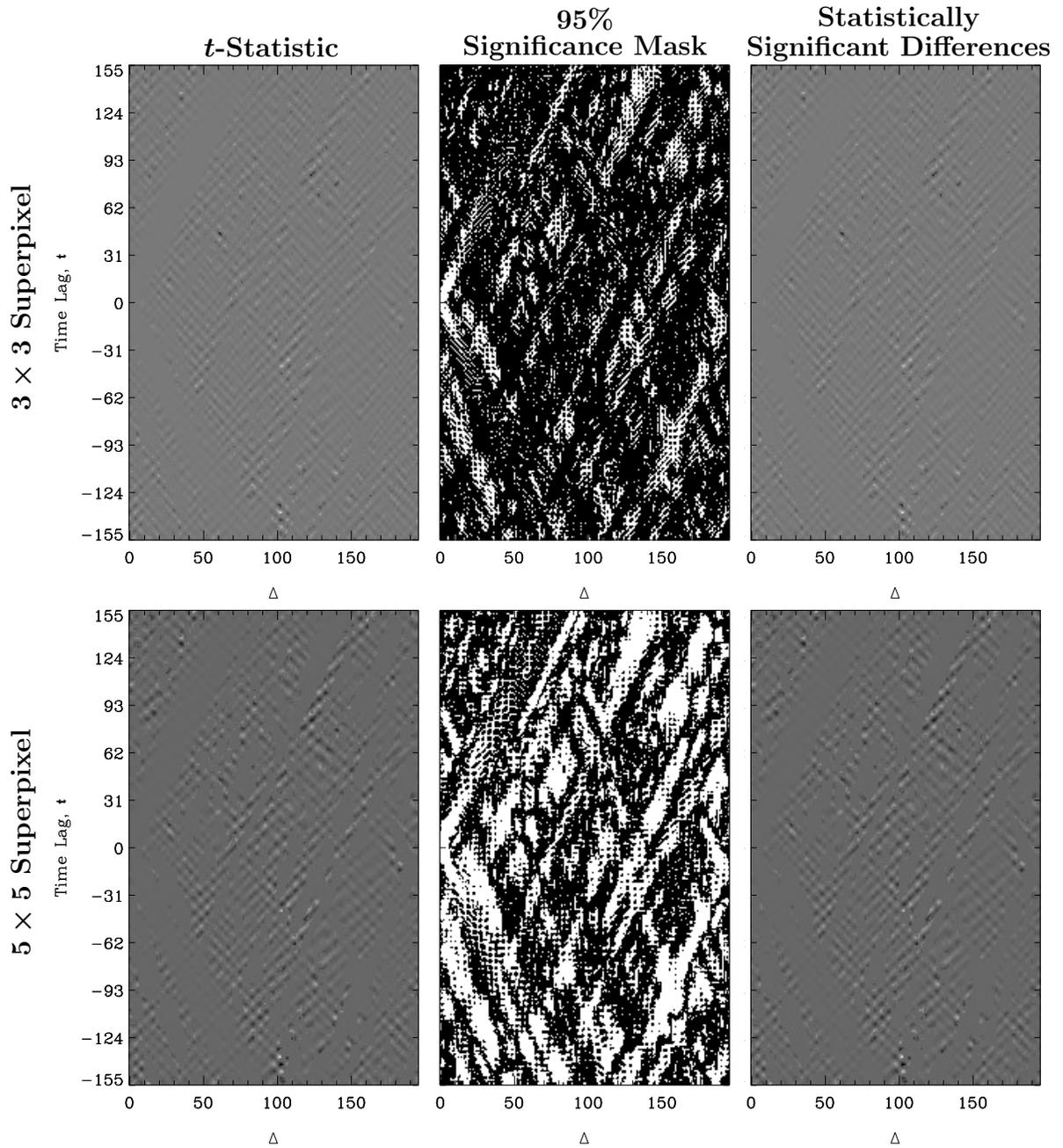


Figure 5.18: The Student’s t -test comparison between the shifted TD results inside the CH region and the separate quiet sun data. Δ has units of pixels, and t has units of seconds ($dt = 45$ s). The top row of results employs a 3×3 superpixel; the bottom row, a 5×5 superpixel. Full maps of the t -statistic are shown on the left. The center images show the binary $\geq 95\%$ significance maps of the left-hand results. Black indicates statistically significant differences and white indicates otherwise. We threshold the t -test results on the left with the center significance masks to produce the plots on the right. These results show only the significantly different TD signals between the CH region and the unrelated quiet sun region. Note that small-scale differences between the main TD branches are averaged out with this technique.

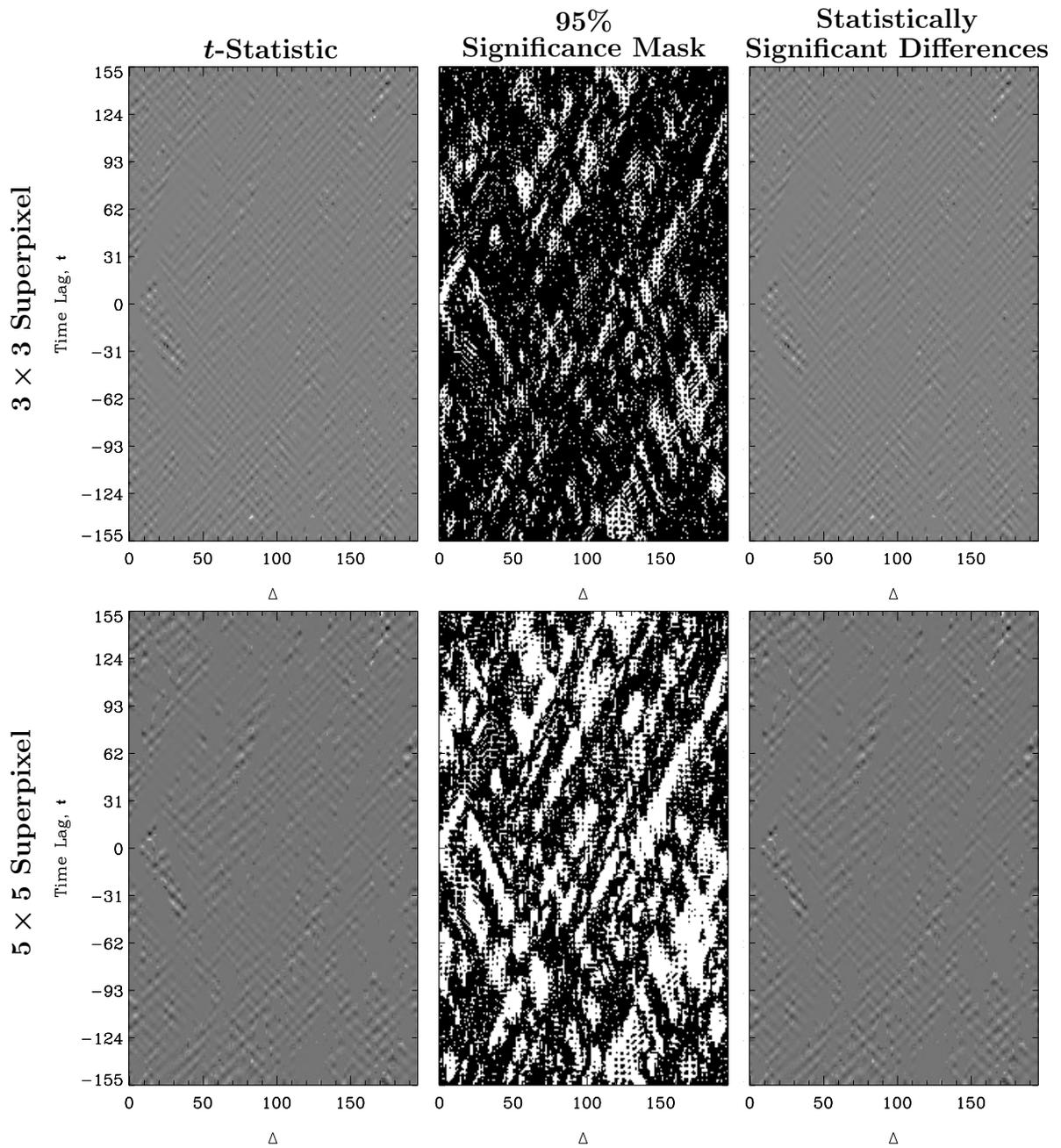


Figure 5.19: The Student’s t -test comparison between the shifted TD results outside of the CH boundary and the separate quiet sun data. Δ has units of pixels, and t has units of seconds ($dt = 45$ s). The top row of results employs a 3×3 superpixel; the bottom row, a 5×5 superpixel. Full maps of the t -statistic are shown on the left. The center images show the binary $\geq 95\%$ significance maps of the left-hand results. Black indicates statistically significant differences and white indicates otherwise. We threshold the t -test results on the left with the center significance masks to produce the plots on the right. These results show only the significantly different TD signals between the quiet sun region near the CH boundary and the unrelated quiet sun region. Note that small-scale differences between the main TD branches are averaged out with this technique.

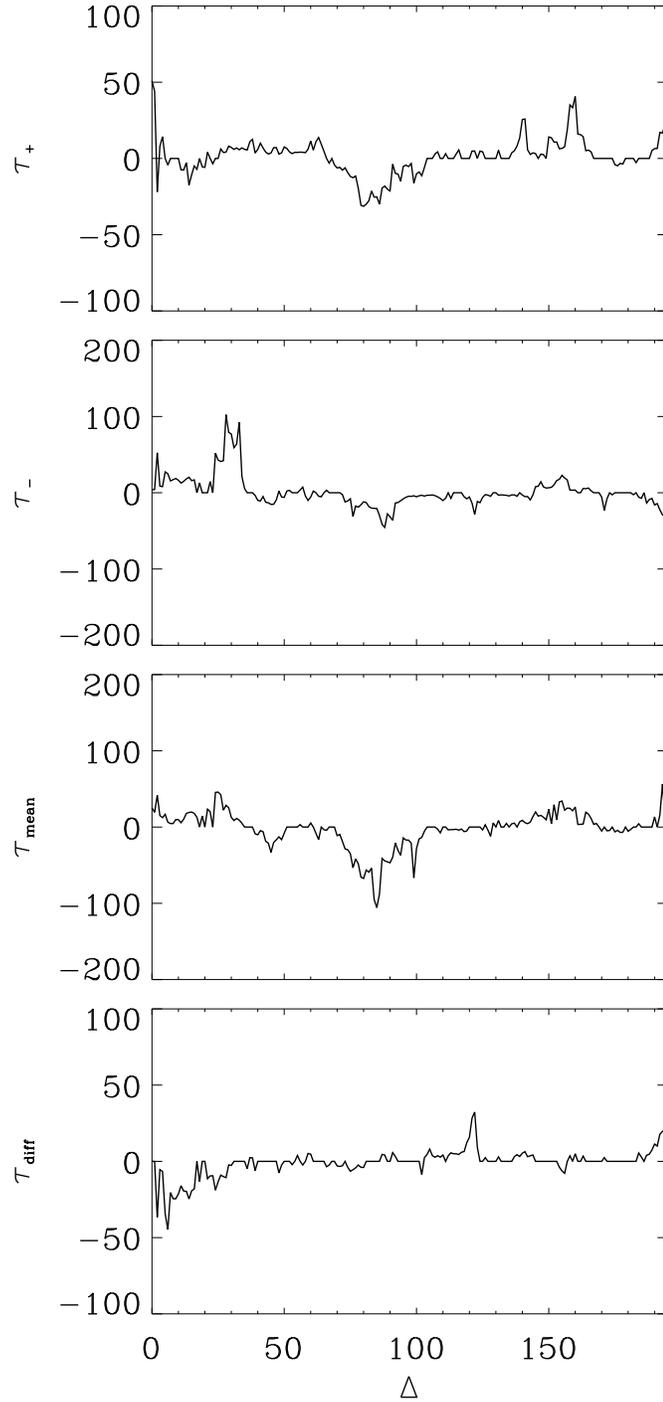


Figure 5.20: The $\geq 95\%$ significance of differences between the shifted travel time measurements inside and outside the CH boundary. Shown are the significant t -statistics of $\tau_+(\Delta)$ (middle top), $\tau_-(\Delta)$ (top), $\tau_{\text{mean}}(\Delta)$ (middle bottom), and $\tau_{\text{diff}}(\Delta)$ (bottom). Travel times have units of seconds ($dt = 45$ s). Δ has units of pixels. Recall, 1 HMI pixel = $0.06^\circ \simeq 0.729$ Mm.

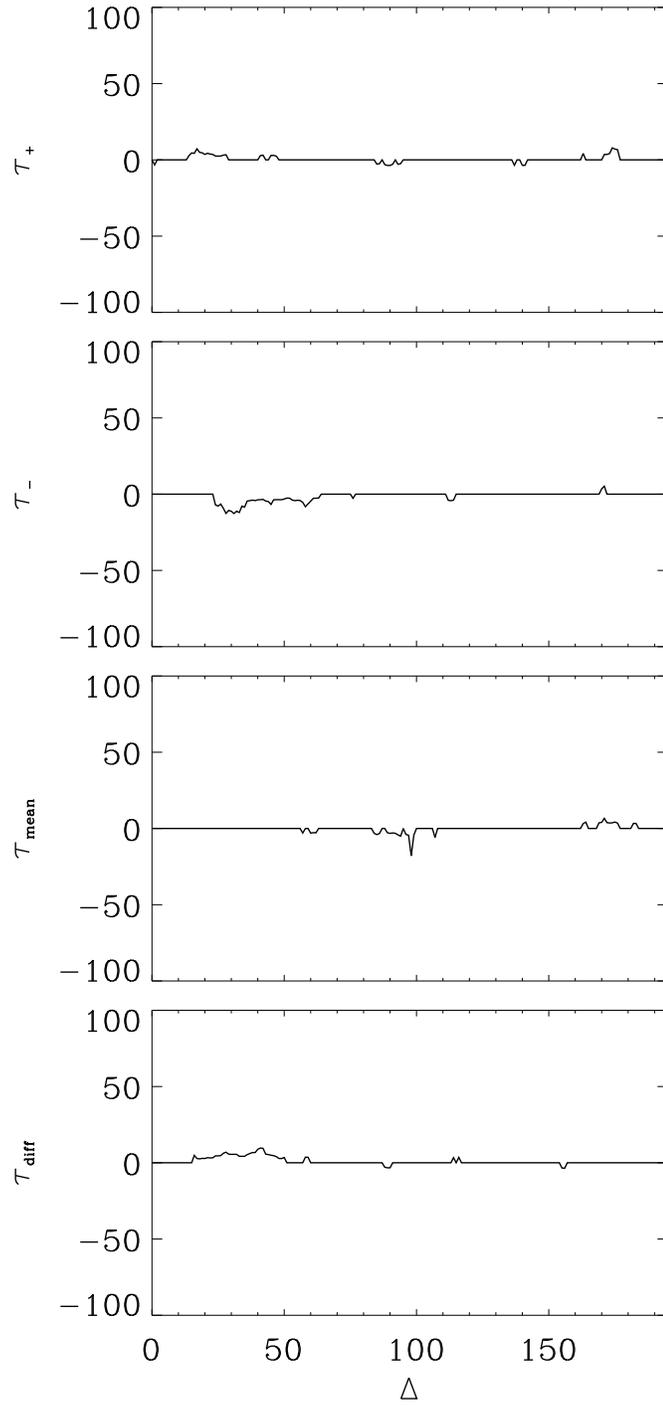


Figure 5.21: The $\geq 95\%$ significance of differences between the shifted travel time measurements of the entire CH data and the separate quiet sun data. Shown are the significant t -statistics of $\tau_+(\Delta)$ (middle top), $\tau_-(\Delta)$ (top), $\tau_{\text{mean}}(\Delta)$ (middle bottom), and $\tau_{\text{diff}}(\Delta)$ (bottom). Travel times have units of seconds ($dt = 45$ s). Δ has units of pixels. Recall, 1 HMI pixel = $0.06^\circ \simeq 0.729$ Mm.

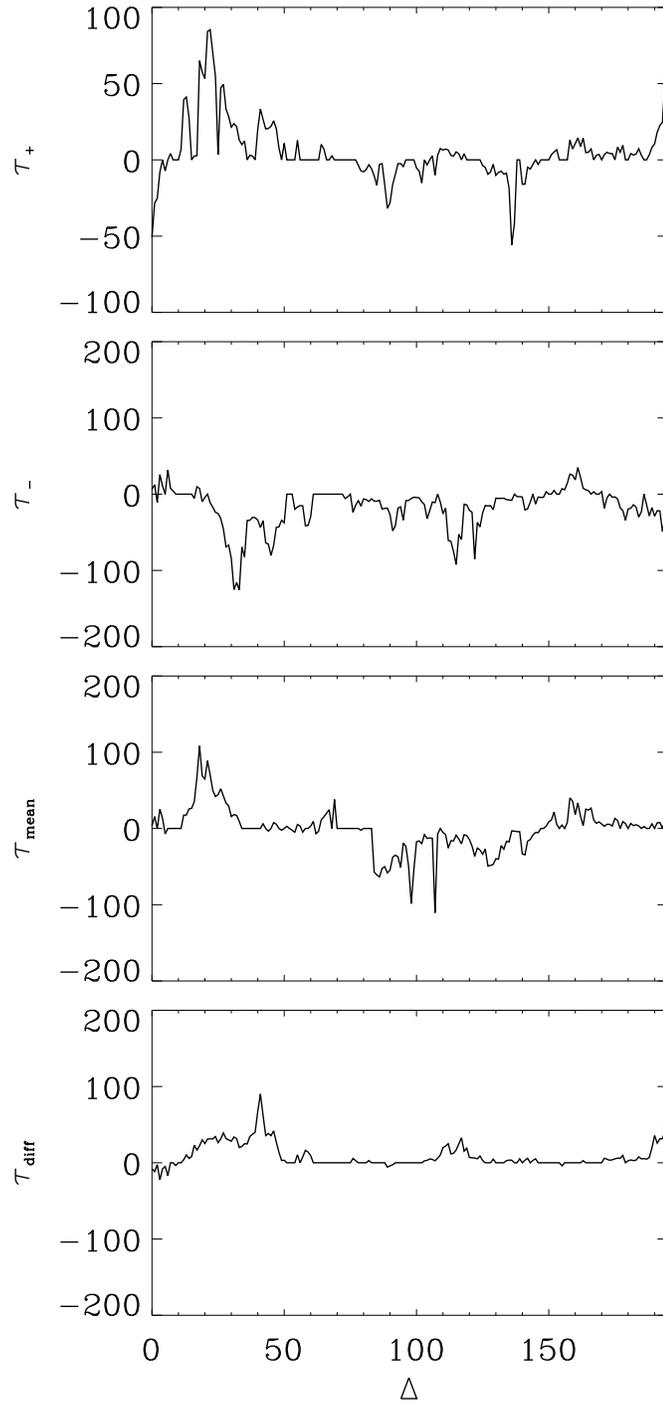


Figure 5.22: The $\geq 95\%$ significance of differences between the shifted travel time measurements inside the CH and the unrelated quiet sun data. Shown are the significant t -statistics of $\tau_+(\Delta)$ (middle top), $\tau_-(\Delta)$ (top), $\tau_{\text{mean}}(\Delta)$ (middle bottom), and $\tau_{\text{diff}}(\Delta)$ (bottom). Travel times have units of seconds ($dt = 45$ s). Δ has units of pixels. Recall, 1 HMI pixel $= 0.06^\circ \simeq 0.729$ Mm.

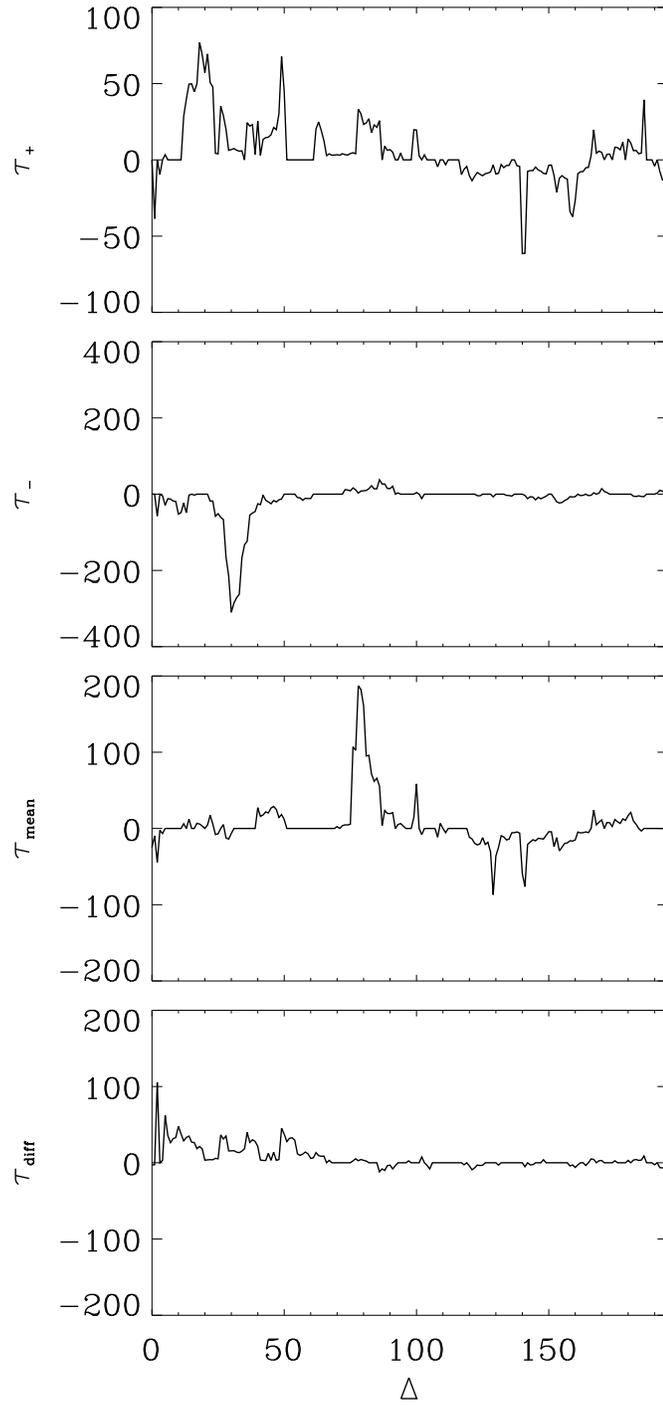


Figure 5.23: The $\geq 95\%$ significance of differences between the shifted travel time measurements of the region outside of the CH boundary and the separate quiet sun data. Shown are the significant t -statistics of $\tau_+(\Delta)$ (middle top), $\tau_-(\Delta)$ (top), $\tau_{\text{mean}}(\Delta)$ (middle bottom), and $\tau_{\text{diff}}(\Delta)$ (bottom). Travel times have units of seconds ($dt = 45$ s). Δ has units of pixels. Recall, 1 HMI pixel = $0.06^\circ \simeq 0.729$ Mm.

Chapter 6: Conclusions

This dissertation is comprised of three main topics relevant to the study of elongated f -mode wave scattering detection and its application to a solar CH boundary. In Chapter 2, we investigated CHs using statistical properties and image processing techniques on EUV images to detect CH boundaries in the low corona and chromosphere. Part of this study used SOHO/EIT data to locate polar CH boundaries on the solar limb, which were subsequently tracked through two solar cycles. Additionally, we developed an edge-detection algorithm that was used on SDO/AIA data spanning a polar CH extension that has an approximately linear boundary. These locations were also used to inform part of the helioseismic investigation. In Chapter 3, we discussed a new local TD helioseismology technique that can be used to detect elongated boundary features. The method employs a new averaging scheme that makes use of the quasi-linear topology of elongated scattering regions. It enhances the wave propagation signal in the direction perpendicular to the linear feature and reduces the computational time of the TD analysis. We also reviewed the scattering theory necessary to create a simulation that we could use to test the new technique. We discussed the simulated data and the results from the linear TD analysis in Chapter 4, as well as considering the significance of differences between the simulated TD results. Finally in Chapter 5, we applied the TD technique to solar CH data from SDO/HMI. The data correspond to the AIA data used in the edge-detection algorithm on EUV images. We looked for significant differences in the TD results and measured travel times between the CH and the quiet sun.

The investigation tracking polar CH boundaries found that the polar CH areas did not change significantly between solar minima, even though the polar magnetic field strength weakened by half.

After producing a simulated data set containing a linear array of scatterers, we subsequently tested the linear TD technique. Our new linear technique successfully enhances

the linear interference pattern from the line of scatterers, while reducing the computational time of the analysis. We found that the results clearly and significantly differ between a linear scattering region and a circularly symmetric scattering region. We conclude that using circularly-symmetric TD geometries on features that are radially asymmetric may produce results with signatures of the anisotropy. Therefore, it is important to use a geometry appropriate to the feature of interest when performing TD analysis. Additionally, our method can easily be extended to study cases in which the array of scatterers are not linearly aligned.

The results of the EUV edge-detection method provided the boundary location of the polar CH extension used in the linear TD analysis, and informed the image-processing parameters necessary for aligning the CH boundary in the Dopplergram data. We should note that our explanation of the download process for JSOC cutout data was generalized and made publicly available as guiding documentation for individuals wanting to make use of that service. Applying our linear TD analysis to the HMI data, we found that the travel time results indicate a small, uniform flow in both a quiet sun region and in the CH data, including specifically within the CH region itself. However, the uncertainties are large enough that we cannot conclude about the existence of this flow with certainty. In general, a flow in the $+x$ direction would most likely be attributed to solar rotation. Yet because of our rotation of the data, our x -dimension does not directly correspond to an east-west flow, and is more along the lines of an east northeast-west southwest flow. Considering this and the relatively high standard deviations of the travel times, we are hesitant to conclude that solar rotation is the cause of our potential underlying flow, although we recognize that we could still be detecting a slight east-west component of the rotation. Further results, averaged within the regions inside and outside of the CH, suggest that there is actually a reversed underlying flow in the region outside of the CH boundary, and that there is some wave-speed perturbation across the CH boundary. This result is further supported by the statistical comparison performed on the TD results and travel time measurements. We suggest that our findings indicate the presence of a wave front propagating through the

quiet sun toward the CH boundary, which is then disrupted and weakened in the boundary region via absorption by the open field lines or backscattering that is lost in the noise. We cannot confirm this conjecture without some of the future work that we are currently investigating.

6.1 Future Work

The possibilities for future work stemming from this dissertation are many and varied. They include everything from improving the simulation, to detecting the CH boundary more accurately, to applying the linear TD technique to other solar features, etc. We will briefly review some of the future work possibilities mentioned in this dissertation, as well as other ideas that have not been discussed.

In terms of the edge detection technique, there are several research options to try. We may be able to restrict the width of the CH boundary by performing similar image post-processing techniques on He 10830 Å data. This would allow us to determine if the boundary width in the chromosphere is similar to that in the corona. The combined AIA data is mainly coronal with large scale heights and may not represent the CH boundary width in the photosphere/low-chromosphere very well. It has also been suggested that we try using AIA 1700 Å data for similar reasons. We could also try using HMI magnetogram data to look for unipolar regions corresponding to the CH region. Unfortunately, unipolar regions may not be well defined in magnetograms, so our post-processing technique may not be useful with that data set. Additionally, it has been suggested that we apply a point-spread-function to our AIA data, in order to remove stray light effects. We did not do so because we are downloading JSOC cutout data instead of full-disk images, and the point-spread function needs to be applied to the full-disk prior to retrieving the cutout data. However, it is possible that removing stray light effects will give us higher accuracy in detecting the CH boundary. Finally, there has been a recommendation that instead of performing an approximate rotation to align the CH boundary visually, we use a 16-point edge detection of the boundary to determine the slope and thus the angle of rotation needed

for proper alignment.

There are also many potential improvements that can be made on the simulation. Most basic is to consider using the full-field Hankel function instead of the far-field approximation that we currently employ, since multiple scattering can already be neglected through our choice of scatterer radius (R. Cameron, August 2015, personal comm.). The relative importance of the far-field assumption could be an exploratory topic using the simulation. We could also incorporate more realistic physics, such as wave attenuation or absorption. There is also the possible undertaking of changing from an analytic code to a numerical simulation, which could then include more physical parameters such as density and magnetic field.

We can also perform further TD analysis on the simulated data. Future work might include adjusting the input parameters such that the incident angle and/or initial mode phases are randomized, which would then produce a more realistic cross-correlation signal. We can also change other parameters to test the TD results of, for example, an elongated scattering boundary that is not strictly linear or expand from a line of scatterers to a broader region of scatterers (either symmetrically aligned or randomly distributed). We also want to further investigate the statistical differences between varying scenarios, and consider other methods of statistical comparison. We are also currently working on isolating and enhancing the simulated backscattered signal. This requires the use of phase-speed and group-speed information to focus the backscattered region in the TD results, shift the data so that the interaction at the boundary is aligned, and then implement a phase-speed shift that allows us to average over the backscattering for better signal-to-noise. The simulated TD results will be used to test all these steps before applying them to the solar data.

In terms of the statistical analysis, we want to improve upon our superpixel t -test method and look into other possible significance-testing methods in image-processing. We might also make other statistical comparisons between data. Additionally, it is important to begin using propagated errors in our analyses, which we are not currently considering. Finally, we should compare the results of our CH region against quiet sun data observed at

the same Carrington latitude and longitude coordinates in order to account for projection, line-of-sight, and rotational effects.

To improve upon our linear TD results of the HMI data, we want to extend the data set in time. This should help improve our signal-to-noise ratio. We could also find other similar CH boundary cases to compare. We also want to expand the technique from approximately linear boundary regions to handle elongated boundaries that are piecewise-linear, as mentioned for the future testing possibilities of the simulation. Finally, we could also test our method on an approximately symmetric sunspot and compare the TD results to those found using a center-to-annulus east-west quadrant analysis. We have actually already chosen a candidate symmetric sunspot on which to perform this study.

As mentioned above, we are currently working toward the detection of backscattered signals from the CH boundary. We also plan to do more work on the statistical analysis of the TD results, as well as the travel time measurements. We are currently working on using a traditional t -test analysis on the shifted TD results, since they are averages of a set of shifted data. Finally, we have not yet investigated the presence or potential cause of a wave front propagating towards the CH, detected in our travel time measurements. We need to return to the full-disk images to determine whether any nearby solar activity produced such a wave.

There are also bigger steps that can be taken to extend this research. There is interest in the research community for applying our linear TD method on a filament region. It has also been suggested that we retry the analysis for lower ℓ (larger λ) to make boundary width “sharper”. Theoretically, this should result in a stronger backscattered signal. A λ :width ratio greater than $\sim 3:1$ in Mm is desirable (R. Cameron, August 2015, personal comm.). Similarly, we should retry the analysis for higher frequencies, which are more sensitive to atmospheric structure. We are more likely to detect backscattering from a CH boundary at higher frequencies because the main difference between a CH region and the quiet sun is the atmospheric conditions, and high frequencies are more sensitive to the coronal state. Combining this with the investigation of lower values of ℓ , it has been recommended that

we look at the at the p_4 -mode in the future, instead of the f -mode.

Appendix A: Processing AIA Data - Details

This appendix pertains to the discussion in Section 2.1. Specifically, Appendix A.1 is referenced in Section 2.1.1, and Appendix A.2 is referenced in Sections 2.1.1 and 2.1.4.

A.1 Determining the AIA Cutout Bounding Box

We share some example code for determining the pixel dimensions needed by the JSOC cutout service in order to produce AIA data with the same heliographic dimensions as the relevant HMI data. This is necessary because the AIA pixel scale is different than that of HMI, even when projecting the AIA data onto the same coordinate map as the HMI data.

A reference projected AIA cutout image is required by this code for the header information needed in the WCS routines. The dimensions can be arbitrarily chosen for this test case. We start with square pixel dimensions approximated by the desired domain size ($30^\circ \times 30^\circ$) divided by the projected AIA pixel scale (JSOC default of $0.0359^\circ/\text{pixel}$), giving $30^\circ/(0.0359^\circ/\text{pixel}) \approx 635$ pixels square.

However, the HMI file is *not* arbitrary. The HMI data should have the bounding box and center of the desired AIA output (ideally, use the actual data to be used in the analysis).

```
fnmhmi = HMI data filename
fnmaia = AIA data filename

;read in fits data and store headers
hmidata=readfits(fnmhmi,hmihead)
aiadata=readfits(fnmaia,aiahead)

;convert header information into wcs structures
wshmi=fitshead2wcs(hmihead)
wsaia=fitshead2wcs(aiahead)

;store size of hmi data axes
;note, hmi time axis doesn't matter here
nxhmi=fxpar(hmihead,'NAXIS1')
nyhmi=fxpar(hmihead,'NAXIS2')

;get hmi coordinates for lower-left and upper-right pixels
hmill=wcs_get_coord(wshmi,[0,0,0])
hmiur=wcs_get_coord(wshmi,[nxhmi-1,nyhmi-1,0])
```

```

;convert hmi hgc coordinates (deg) to projected aia hgc coordinates (deg
)
wcs_convert_to_coord , wcsaia , aiall , `hg' , hmill[0] , hmill[1] , /
  carrington
wcs_convert_to_coord , wcsaia , aiaur , `hg' , hmiur[0] , hmiur[1] , /
  carrington

;check aia coordinates compared to hmi coordinates
;these SHOULD be the same
print , aiall , hmill[0:1]
print , aiaur , hmiur [0:1]

;convert aia degrees to pixels
aiall_p=wcs_get_pixel(wcsaia , aiall)
aiaur_p=wcs_get_pixel(wcsaia , aiaur)

;report aia bounding box in pixels ,
print , 'lower-left ' , aiall_p
print , 'upper-right ' , aiaur_p
print , 'difference ' , aiaur_p-aiall_p

;the bounding box should be square ,
;the difference is approximately the size of the required aia data ,
;input these pixel dimensions into the JSOC export system!

```

Once the desired pixel dimensions are determined, we recommend that users download a sample instance of the projected AIA data from the JSOC cutout service and double check the bounding box coordinates before exporting the entire data set. This may save time in the long run if the bounding box size needs to be adjusted.

A.2 Updating the Projected AIA Header Information

Unfortunately, the MAPROJ processing option in the JSOC cutout service (discussed in Section 2.1.1) does not update enough of the relevant header information to port the metadata correctly into a WCS structure. This is a fairly straightforward thing to fix, but must be done manually (at least for now). And it *must* be done in order to move forward with determining the necessary AIA bounding box (Appendix A.1) and converting the detected coronal hole boundary location in the AIA data to the corresponding location in the HMI data (Section 2.2). In plain terms, it is a silly and tedious step but is a requirement before the “real work” can begin.

First, we create a template WCS structure in IDL:

```
template={template, $
  COORD.TYPE: '', $
  WCSNAME: '', $
  SIMPLE: BYTE(0), $
  NAXIS: LONARR(2), $
  VARIATION: '', $
  COMPLIANT: BYTE(0), $
  PROJECTION: '', $
  IX: 0L, $
  IY: 0L, $
  CRPIX: DBLARR(2), $
  CRVAL: DBLARR(2), $
  CTYPE: STRARR(2), $
  CNAME: STRARR(2), $
  CUNIT: STRARR(2), $
  CDELT: DBLARR(2), $
  ROLL_ANGLE: DOUBLE(0), $
  PC: DBLARR(2,2), $
  TIME: {TIME, $
    FITS_DATE: '', $
    OBSERV_DATE: STRARR(4), $
    OBSERV_END: STRARR(4), $
    EXPTIME: DOUBLE(0) $
  }, $
  POSITION: {POSITION, $
    SOHO: BYTE(0), $
    POS_ASSUMED: 0, $
    DSUN_OBS: DOUBLE(0), $
    SOLAR_B0: DOUBLE(0), $
    SOLAR_L0: DOUBLE(0), $
    CARR_EARTH: DOUBLE(0), $
    HGLN_OBS: DOUBLE(0), $
    HGLT_OBS: DOUBLE(0), $
    HEQ_OBS: DBLARR(3) $
  } $
}
```

Each tag in the template structure is defined according to the standard WCS structure for solar data, including the type and size of the associated values. We duplicate this template into arrays for each of our four AIA wavelengths.

We then pre-define the values that we expect for certain tags of the projected AIA data (the WCS tags from the associated HMI data can be used to identify some of these desired “default” values):

```

coordtype='Carrington-Heliographic '
wcsname='Carrington-Heliographic '
naxis=[695,695] ;after rotation and cropping
variation='CROTA'
compliant=BYTE(0)
projection='ARC'
ix=0L
iy=1L
crpix=[348,348]
ctype=['CRLN-ARC', 'CRLT-ARC']
cname=[' ', ' ']
cunit=['deg ', 'deg ']
cdelt=0.0359
rollangle=DOUBLE(-20.0)
pc=[[COS(DOUBLE(rollangle*IDTOR)),SIN(DOUBLE(rollangle*IDTOR))],[-SIN(
DOUBLE(rollangle*IDTOR)),COS(DOUBLE(rollangle*IDTOR))]]
simple=BYTE(0)
soho=BYTE(0)
posassumed=0

```

Next, we loop through the original “bad” WCS structures for the four AIA wavelengths we use. We compare the original tag values to the desired “default” tag values defined above. For those values that seem to have been updated through MAPROJ, we simply check for consistency with the default values (an error message is displayed if something goes wrong), and then set the new WCS tag value to the original value. For the values that we know were not updated by MAPROJ, we explicitly set the tag to the new default tag value.

We also use the template to create a new WCS structure in which we combine information from the four wavelengths (to be used with the combined data produced in Sections 2.1.3 and 2.1.4). The WCS string names should match, regardless of wavelength, and are easily copied into the combination WCS structure. Several of the numerical values are also simply copied (i.e. SIMPLE, SOHO, ROLLANGLE, CDELTA, etc.). The complication comes from those values that are slightly different between the four different data sets (i.e. OBSERV_DATE, OBSERV_END, EXPTIME, DSUN_OBS, SOLAR_B0, etc.). These values are slightly different because the different AIA cameras and filters take pictures at slightly different times, which translates to slightly different positions as well. However, some of these parameters are not necessary for the WCS computations, so we store an array of the

four values for reference. Those parameters that *are* important in the WCS calculations vary only slightly between the four wavelengths, so we store the average of those values for later use. We make note of this for consistency and reproducibility.

Appendix B: An In-Depth Derivation of the Scattering Amplitude and Related Quantities

B.1 Cylindrical Plane Wave Expansion

The solution to the full time-dependent wave equation in cylindrical coordinates is given by Equation 3.1, in which the spatial wavefunction ψ is a solution to the Helmholtz equation

$$(\nabla^2 + k^2) \psi = 0. \quad (\text{B.1})$$

The Laplacian operator in cylindrical coordinates is

$$\begin{aligned} \nabla^2 &\equiv \frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial}{\partial r} \right) + \frac{1}{r^2} \frac{\partial^2}{\partial \varphi^2} + \frac{\partial^2}{\partial z^2} \\ &= \frac{\partial^2}{\partial r^2} + \frac{1}{r} \frac{\partial}{\partial r} + \frac{1}{r^2} \frac{\partial^2}{\partial \varphi^2} + \frac{\partial^2}{\partial z^2}. \end{aligned} \quad (\text{B.2})$$

If we assume the wavefunction is a separable equation $\psi = R(r)\Phi(\varphi)Z(z)$, we can insert this solution into Equation B.1 and find solutions to the separate terms of ψ

$$\begin{aligned} \frac{\partial^2 \psi}{\partial r^2} + \frac{1}{r} \frac{\partial \psi}{\partial r} + \frac{1}{r^2} \frac{\partial^2 \psi}{\partial \varphi^2} + \frac{\partial^2 \psi}{\partial z^2} + k^2 \psi &= 0 \\ \Rightarrow \Phi Z \frac{\partial^2 R}{\partial r^2} + \Phi Z \frac{1}{r} \frac{\partial R}{\partial r} + R Z \frac{1}{r^2} \frac{\partial^2 \Phi}{\partial \varphi^2} + R \Phi \frac{\partial^2 Z}{\partial z^2} + k^2 R \Phi Z &= 0 \\ \Rightarrow \frac{r^2}{R} \frac{\partial^2 R}{\partial r^2} + \frac{r}{R} \frac{\partial R}{\partial r} + \frac{1}{\Phi} \frac{\partial^2 \Phi}{\partial \varphi^2} + \frac{r^2}{Z} \frac{\partial^2 Z}{\partial z^2} + k^2 r^2 &= 0 \end{aligned} \quad (\text{B.3})$$

At this point we can separate out the Φ term. Because the solution to Φ must be periodic in φ , the separation constant must be negative.

$$\frac{1}{\Phi} \frac{d^2 \Phi}{d\varphi^2} = -m^2 \longrightarrow \Phi(\varphi) = A_m \cos(m\varphi) + B_m \sin(m\varphi) \quad (\text{B.4})$$

Continuing with the remaining terms in the partially separated Helmholtz equation,

$$\begin{aligned} & \frac{r^2}{R} \frac{\partial^2 R}{\partial r^2} + \frac{r}{R} \frac{\partial R}{\partial r} + \frac{r^2}{Z} \frac{\partial^2 Z}{\partial z^2} + k^2 r^2 - m^2 = 0 \\ \Rightarrow & \frac{1}{R} \frac{\partial^2 R}{\partial r^2} + \frac{1}{rR} \frac{\partial R}{\partial r} + \frac{1}{Z} \frac{\partial^2 Z}{\partial z^2} + k^2 - \frac{m^2}{r^2} = 0 \end{aligned} \quad (\text{B.5})$$

Separating out the Z term now, the solution must not be periodic as $z \rightarrow \infty$, which necessitates a positive separation constant

$$\frac{1}{Z} \frac{d^2 Z}{dz^2} = n^2 \longrightarrow Z(z) = C_n e^{nz} + D_n e^{-nz} \quad (\text{B.6})$$

Returning to the final separation term from the Helmholtz equation, we now have

$$\begin{aligned} & \frac{1}{R} \frac{d^2 R}{dr^2} + \frac{1}{rR} \frac{dR}{dr} + k^2 - \frac{m^2}{r^2} + n^2 = 0 \\ \Rightarrow & \frac{d^2 R}{dr^2} + \frac{1}{r} \frac{dR}{dr} + \left(k^2 - \frac{m^2}{r^2} + n^2 \right) R = 0 \\ \Rightarrow & r^2 \frac{d^2 R}{dr^2} + r \frac{dR}{dr} + ((k^2 + n^2)r^2 - m^2) R = 0 \end{aligned} \quad (\text{B.7})$$

Letting $x = r\sqrt{k^2 + n^2} \longrightarrow r = x/\sqrt{k^2 + n^2}$ and $dx = \sqrt{k^2 + n^2} dr \longrightarrow dr = dx/\sqrt{k^2 + n^2}$, therefore $dx^2 = (k^2 + n^2)dr^2 \longrightarrow dr^2 = dx^2/(k^2 + n^2)$. Substituting these into the previous equation gives

$$\begin{aligned} & \frac{x^2 \cancel{(k^2 + n^2)}}{\cancel{(k^2 + n^2)}} \frac{d^2 R}{dx^2} + \frac{x \sqrt{k^2 + n^2}}{\sqrt{k^2 + n^2}} \frac{dR}{dx} + (x^2 - m^2) R = 0 \\ \Rightarrow & x^2 \frac{d^2 R}{dx^2} + x \frac{dR}{dx} + (x^2 - m^2) R = 0 \\ \Rightarrow & \frac{d^2 R}{dx^2} + \frac{1}{x} \frac{dR}{dx} + \left(1 - \frac{m^2}{x^2} \right) R = 0 \end{aligned} \quad (\text{B.8})$$

This resulting equation is the Bessel equation. Solutions are a linear combination of Bessel functions of the first and second kinds ($J_{mn}(x)$ and $N_{mn}(x)$, respectively). The general solution to the radial equation is thus

$$R(r\sqrt{k^2 + n^2}) = E_{mn}J_{mn}(r\sqrt{k^2 + n^2}) + F_{mn}N_{mn}(r\sqrt{k^2 + n^2}). \quad (\text{B.9})$$

Combining Equations B.4, B.6 and B.9, we find the general form of the spatial wavefunction to be

$$\begin{aligned} \psi(r, \varphi, z) = \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} [A_m \cos(m\varphi) + B_m \sin(m\varphi)] [C_n e^{nz} + D_n e^{-nz}] \\ \times \left[E_{mn}J_{mn}(r\sqrt{k^2 + n^2}) + F_{mn}N_{mn}(r\sqrt{k^2 + n^2}) \right]. \end{aligned} \quad (\text{B.10})$$

Note that m is the azimuthal order of the resulting wave and n is the radial order. The spherical harmonic degree l is embedded in our definition of the wavenumber k (see Section 1.2).

For our purposes, we only wish to consider f -mode waves, in which case $n = 0$. With this simplification, the previous equation becomes

$$\psi(r, \varphi, z) = C \sum_{m=0}^{\infty} [A_m \cos(m\varphi) + B_m \sin(m\varphi)] [E_m J_m(kr) + F_m N_m(kr)], \quad (\text{B.11})$$

where we have now lost any dependence on z (as might be expected for surface waves) and the only contribution from the $Z(z)$ equation is a constant factor which we call $C = C_0 + D_0$.

We can reach a more specific solution for a plane wave perpendicular to the cylindrical (z -)axis with $\psi = Ae^{ikr \cos(\varphi)}$ by starting with a fundamental property of Bessel functions (understanding from Equation B.11 that Bessel functions are part of the solution). First, note that $J_{-m}(z) = (-1)^m J_m(z)$, and we will make use of the trig identity $\cos(x) = \frac{1}{2}(e^{ix} +$

e^{-ix}). By definition,

$$\begin{aligned}
e^{z(t^2-1)/2t} &= \sum_{m=-\infty}^{\infty} t^m J_m(z) \\
&= J_0(z) + \sum_{m=-\infty}^{-1} t^m J_m(z) + \sum_{m=1}^{\infty} t^m J_m(z) \\
&= J_0(z) + \sum_{m=1}^{\infty} (-1)^m t^{-m} J_m(z) + \sum_{m=1}^{\infty} t^m J_m(z) \\
&= J_0(z) + \sum_{m=1}^{\infty} J_m(z) \left[t^m + \left(\frac{-1}{t} \right)^m \right].
\end{aligned} \tag{B.12}$$

If we set $t = ie^{i\varphi}$, we get

$$\begin{aligned}
e^{z((ie^{i\varphi})^2-1)/2ie^{i\varphi}} &= J_0(z) + \sum_{m=1}^{\infty} J_m(z) \left[(ie^{i\varphi})^m + \left(\frac{-1}{ie^{i\varphi}} \right)^m \right] \\
e^{-z(e^{i2\varphi}+1)/2ie^{i\varphi}} &= J_0(z) + \sum_{m=1}^{\infty} J_m(z) [i^m e^{im\varphi} + (-1)^m i^{-m} e^{-im\varphi}] \\
e^{iz(e^{i\varphi}+e^{-i\varphi})/2} &= J_0(z) + \sum_{m=1}^{\infty} J_m(z) [i^m e^{im\varphi} + \cancel{(-1)^m} (-1)^m i^m e^{-im\varphi}] \\
e^{iz \cos(\varphi)} &= J_0(z) + \sum_{m=1}^{\infty} i^m J_m(z) [e^{im\varphi} + e^{-im\varphi}] \\
e^{iz \cos(\varphi)} &= J_0(z) + 2 \sum_{m=1}^{\infty} i^m J_m(z) \cos(m\varphi) \\
e^{iz \cos(\varphi)} &= \sum_{m=0}^{\infty} \epsilon_m i^m J_m(z) \cos(m\varphi),
\end{aligned} \tag{B.13}$$

where ϵ_m is called the Neumann factor and is defined as $\epsilon_0 = 1$, $\epsilon_m = 2$ for $m > 0$.

Substituting in $z = kr$, we find the plane wave expansion in cylindrical coordinates to be

$$\psi = Ae^{ikr \cos(\varphi)} = A \sum_{m=0}^{\infty} \epsilon_m i^m J_m(kr) \cos(m\varphi), \quad (\text{B.14})$$

which is exactly Equation 3.2. In comparison with the general solution given in Equation B.11, this is an acceptable solution to the Helmholtz equation.

B.2 Cylindrical Wave Scattering

The following derivation of cylindrical wave scattering equations is primarily attributed to Morse and Feschbach (1953) and Morse and Ingard (1986).

The total wavefunction for an incident plane wave scattering off a cylinder of radius a , oriented along the z -axis, is $\Psi_{tot} = \psi_{tot} e^{-i\omega t} = (\psi_i + \psi_s) e^{-i\omega t}$. The incident plane wave expansion ψ_i is given by Equation 3.2, and ψ_s is the scattered wavefunction.

As $r \rightarrow \infty$, the solution to ψ_s must converge to an outgoing plane wave perpendicular to the cylindrical axis. This behavior can generally be expressed by the Hankel function of the first kind, $H_m(kr) = J_m(kr) + iN_m(kr)$, (also known as the Bessel function of the third kind), which is the full general outgoing solution to the Bessel equation. We can thus assume the solution to the scattered wavefunction is of the form

$$\psi_s = \sum_{m=0}^{\infty} B_m H_m(kr) \cos(m\varphi), \quad (\text{B.15})$$

where B_m are coefficients that could depend on azimuthal order. Comparing this assumed solution to Equation B.11, it is an appropriate solution to the Helmholtz solution. The total time-independent wavefunction is then written as

$$\psi_{tot} = \psi_i + \psi_s$$

$$\begin{aligned}
&= A \sum_{m=0}^{\infty} \epsilon_m i^m J_m(kr) \cos(m\varphi) + \sum_{m=0}^{\infty} B_m H_m(kr) \cos(m\varphi) \\
&= \sum_{m=0}^{\infty} A \epsilon_m i^m J_m(kr) \cos(m\varphi) + B_m H_m(kr) \cos(m\varphi) \tag{B.16}
\end{aligned}$$

Assuming the cylinder is a rigid scatterer, the boundary condition at $r = a$ must be $v_n = \nabla\psi_{tot} \cdot \mathbf{n} = 0$. In other words, there is no fluid flow across the rigid boundary so the normal velocity (normal gradient of the wavefunction) is non-existent at the boundary. This condition is also known as a Neumann boundary condition. Note that the normal velocity at a cylindrical surface is the simply radial velocity. Therefore, the Neumann boundary condition for our specific geometry is $v_r = \partial\psi_{tot}/\partial r = 0$. Applying the rigid boundary conditions,

$$v_r = \nabla\psi_{tot} \cdot \mathbf{r} \Big|_{r=a} = \frac{\partial\psi_{tot}}{\partial r} \Big|_{r=a} = 0 \tag{B.17}$$

$$\begin{aligned}
&\Rightarrow \sum_{m=0}^{\infty} A \epsilon_m i^m J'_m(kr) \Big|_{r=a} \cos(m\varphi) + B_m H'_m(kr) \Big|_{r=a} \cos(m\varphi) = 0 \\
&\Rightarrow A \epsilon_m i^m J'_m(kr) \Big|_{r=a} \cancel{\cos(m\varphi)} + B_m H'_m(kr) \Big|_{r=a} \cancel{\cos(m\varphi)} = 0 \\
&\Rightarrow B_m H'_m(kr) \Big|_{r=a} = -A \epsilon_m i^m J'_m(kr) \Big|_{r=a} \\
&\Rightarrow B_m = -A \epsilon_m i^m \frac{J'_m(kr)}{H'_m(kr)} \Big|_{r=a} \tag{B.18}
\end{aligned}$$

Morse and Feschbach (1953, p.1564) define the derivatives of the Bessel and Hankel functions, respectively, as $J'_m(kr) = -C'_m(kr) \sin[\delta'_m(kr)]$ and $H'_m(kr) = iC'_m(kr)e^{i\delta'_m(kr)}$, where $\delta_m(kr)$ is the phase angle and $C_m(kr)$ is the amplitude of the m^{th} -order Hankel function. Substituting $J'_m(kr)$ and $H'_m(kr)$ into the previous step, we get

$$B_m = -A \epsilon_m i^m \frac{\cancel{C'_m(kr)} \sin[\delta'_m(kr)]}{i \cancel{C'_m(kr)} e^{i\delta'_m(kr)}} \Big|_{r=a}$$

$$\begin{aligned}
&= A\epsilon_m i^{m-1} \sin [\delta'_m(kr)] e^{-i\delta'_m(kr)} \Big|_{r=a} \\
&= A\epsilon_m i^{m-1} \sin [\delta'_m(ka)] e^{-i\delta'_m(ka)}
\end{aligned} \tag{B.19}$$

Plugging B_m back into ψ_s gives the scattered wavefunction of a plane wave interacting with a rigid cylinder

$$\psi_s = A \sum_{m=0}^{\infty} \epsilon_m i^{m-1} \sin [\delta'_m(ka)] e^{-i\delta'_m(ka)} H_m(kr) \cos(m\varphi). \tag{B.20}$$

Morse and Feschbach (1953, p.1563-4) also provide us with asymptotic values of $H_m(kr)$ and $\delta'_m(ka)$. As $kr \rightarrow \infty$ (far-field), $H_m(kr) \rightarrow \sqrt{\frac{2}{\pi kr}} e^{ikr - i\pi(m+\frac{1}{2})/2}$. Assuming $ka \ll 2m + 1$, which is a valid assumption for small a and/or small k (small l , large λ , long-wavelength approximation), then $\delta'_0(ka) \rightarrow \frac{1}{4}\pi(ka)^2$ and $\delta'_{m>0}(ka) \rightarrow -\frac{\pi m}{(m!)^2} \left(\frac{ka}{2}\right)^{2m}$. Inserting these approximations into Equation B.20 gives the far-field scattering equations under the long-wavelength approximation. Notice that under these assumptions, Equation B.20 is proportional to e^{ikr}/\sqrt{r} . This is an expected condition for a scattered wave in two dimensions as $r \rightarrow \infty$ (Morse and Feschbach, 1953, 9.3). The scattering amplitude $f(\varphi)$ is equivalent to all other terms in the scattering function (see Equation 3.5).

Equation B.20 is essentially the equation in the *cyl_wav.pro* code that produces the simulated data used to test the time-distance helioseismology techniques for this research (Chapter 4). However, we do implement the long-wavelength approximation to determine $\delta'_m(ka)$ and the far-field assumption to reduce the Hankel function.

Appendix C: Fourier Transforms of Gaussian and Piecewise Functions

In 1D, the transform equation from the spatial domain to the Fourier domain is

$$\tilde{H}(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} H(x) e^{-ikx} dx \quad (\text{C.1})$$

$H(x)$ and $\tilde{H}(k)$ are the input and output functions of the transformation, where x is the spatial coordinate and k is the wavenumber. If we want to find the transform of a gaussian function, we substitute the function for $H(x)$. The transform of a piecewise function is likewise found by substituting the equation for $H(x)$.

Starting with a simple Gaussian function,

$$F(x) = e^{-x^2/\alpha^2} \quad (\text{C.2})$$

where α is an arbitrary constant. The transform equation now looks like

$$\begin{aligned} \tilde{F}(k) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-x^2/\alpha^2} e^{-ikx} dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-(\frac{x^2}{\alpha^2} + ikx)} dx \end{aligned} \quad (\text{C.3})$$

If we multiply through by a constant factor $e^{-(Ak)^2}$, we can complete the square within the integrand exponent.

$$\begin{aligned} \tilde{F}(k) &= \frac{e^{(Ak)^2}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\left(\frac{x^2}{\alpha^2} + ikx + (Ak)^2\right)} dx = \frac{e^{(Ak)^2}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\left(\frac{x}{\alpha} + Ak\right)^2} dx \\ &= \frac{e^{-\left(\frac{\alpha^2 k^2}{4}\right)}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\left(\frac{x}{\alpha} + \frac{i\alpha k}{2}\right)^2} dx \end{aligned} \quad (\text{C.4})$$

Letting $u = \frac{x}{\alpha} + \frac{i\alpha k}{2}$, and $dx = \alpha du$, the substitution yields

$$\begin{aligned}
 \tilde{F}(k) &= \frac{e^{-\left(\frac{\alpha^2 k^2}{4}\right)}}{\sqrt{2\pi}} \alpha \int_{-\infty}^{\infty} e^{-u^2} du \\
 &= \frac{e^{-\left(\frac{\alpha^2 k^2}{4}\right)}}{\sqrt{2\pi}} \alpha \sqrt{\pi} \\
 &= \frac{\alpha}{\sqrt{2}} e^{-\left(\frac{\alpha k}{2}\right)^2}
 \end{aligned} \tag{C.5}$$

Equation C.5 is the solution to the Gaussian Fourier transform and is a Gaussian function itself.

Implementing a simple piecewise function,

$$G(x) = \begin{cases} 0 & \text{for } |x| > \alpha \\ 1 & \text{for } -\alpha < x < \alpha \end{cases} \tag{C.6}$$

we can split the transform equation into three separate integrals, for each region of the piecewise function.

$$\begin{aligned}
 \tilde{G}(k) &= \frac{1}{\sqrt{2\pi}} \left[\int_{-\infty}^{-\alpha} (0) e^{-ikx} dx + \int_{-\alpha}^{\alpha} (1) e^{-ikx} dx + \int_{\alpha}^{\infty} (0) e^{-ikx} dx \right] \\
 &= \frac{1}{\sqrt{2\pi}} \left[0 + \int_{-\alpha}^{\alpha} e^{-ikx} dx + 0 \right] \\
 &= \frac{1}{\sqrt{2\pi}} \int_{-\alpha}^{\alpha} e^{-ikx} dx \\
 &= \sqrt{\frac{2}{\pi}} \frac{\sin(\alpha k)}{k}
 \end{aligned} \tag{C.7}$$

Equation C.7 is most certainly *not* a piecewise function.

The plots of functions $F(x)$ and $G(x)$ with their associated Fourier transforms $\tilde{F}(k)$ and

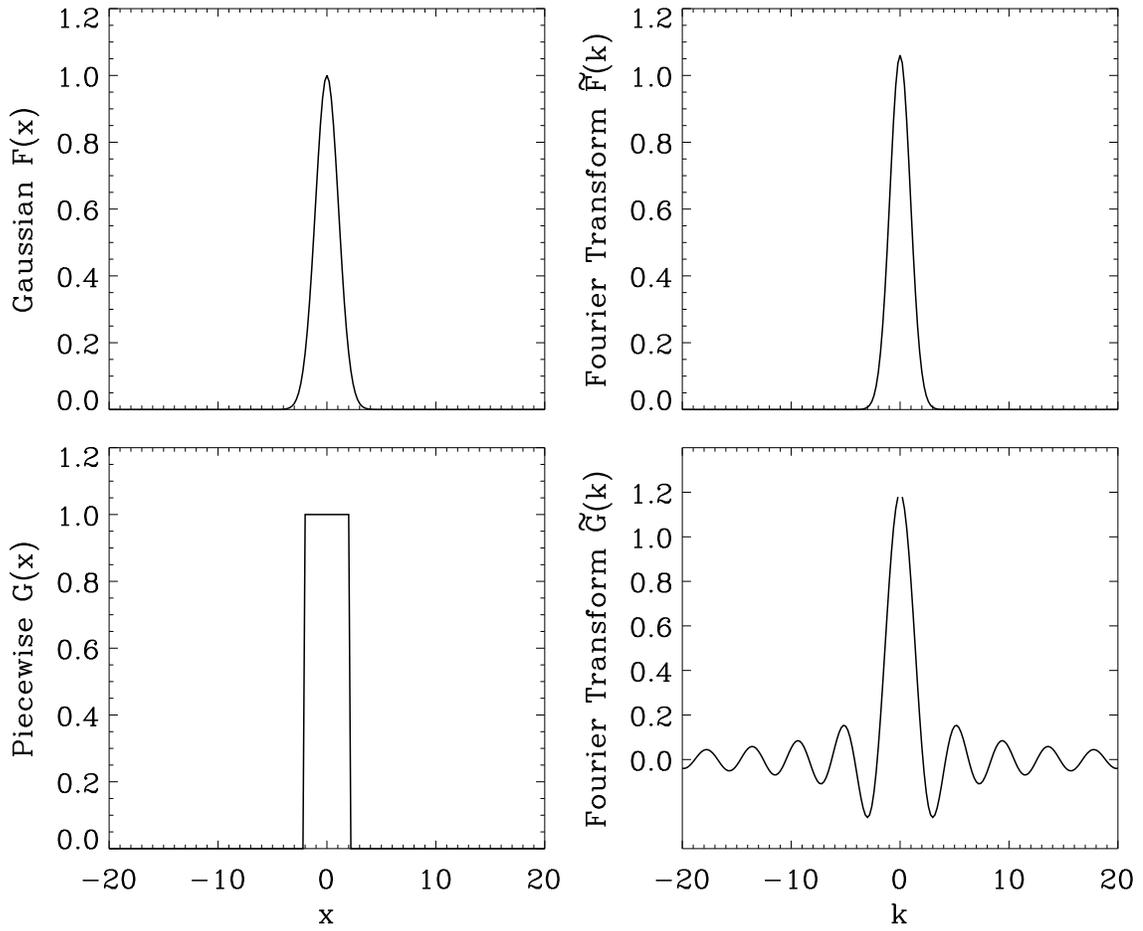


Figure C.1: To the left are spatial plots of the Gaussian function of Equation C.2 (top) and the piecewise function of Equation C.6 (bottom). To the right are the corresponding Fourier transform functions of Equation C.5 (top) and Equation C.7 (bottom). The constant α in the equations is arbitrarily set to $\alpha = 1.5$ for these calculations.

$\tilde{G}(k)$ are shown in Figure C.1.

Bibliography

- Altschuler, M.D., Newkirk, G.: 1969, Magnetic fields and the structure of the solar corona. *Solar Phys.* **9**(1), 131–149. doi:10.1007/BF00145734. <http://dx.doi.org/10.1007/BF00145734>.
- Antiochos, S.K., DeVore, C.R., Karpen, J.T., Mikić, Z.: 2007, Structure and dynamics of the sun's open magnetic field. *Astrophys. J.* **671**(1), 936. <http://stacks.iop.org/0004-637X/671/i=1/a=936>.
- Bohlin, J.D.: 1976, The physical properties of coronal holes. In: Williams, D.J. (ed.) *Physics of Solar Planetary Environments, Proceedings of the International Symposium on Solar-Terrestrial Physics* **1**, American Geophysical Union, Boulder, CO, 114–128. ISBN 9781118669044. doi:10.1029/SP007p0114. <http://dx.doi.org/10.1029/SP007p0114>.
- Bohlin, J.D.: 1977, Extreme-ultraviolet observations of coronal holes. I - Locations, sizes and evolution of coronal holes, June 1973 – January 1974. *Solar Phys.* **51**, 377–398. doi:10.1007/BF00216373.
- Cameron, R., Gizon, L., Duvall, Jr., T.L.: 2008, Helioseismology of sunspots: Confronting observations with three-dimensional MHD simulations of wave propagation. *Solar Phys.* **251**, 291–308. doi:10.1007/s11207-008-9148-1.
- Charbonneau, P.: 2014, Solar dynamo theory. *Ann. Rev. Astron. Astrophys.* **52**, 251–290. doi:10.1146/annurev-astro-081913-040012.

- Couvidat, S., Zhao, J., Birch, A.C., Kosovichev, A.G., Duvall, Jr., T.L., Parchevsky, K., Scherrer, P.H.: 2012, Implementation and comparison of acoustic travel-time measurement procedures for the Solar Dynamics Observatory/HelioSeismic and Magnetic Imager time - distance helioseismology pipeline. *Solar Phys.* **275**, 357–374. doi:10.1007/s11207-010-9652-y.
- Craik, A.D.D.: 2004, The origins of water wave theory. *Ann. Rev. Fluid Mech.* **36**(1), 1–28. doi:10.1146/annurev.fluid.36.050802.122118. <http://www.annualreviews.org/doi/abs/10.1146/annurev.fluid.36.050802.122118>.
- Delaboudinière, J.-P., Artzner, G.E., Brunaud, J., Gabriel, A.H., Hochedez, J.F., Millier, F., Song, X.Y., Au, B., Dere, K.P., Howard, R.A., Kreplin, R., Michels, D.J., Moses, J.D., Defise, J.M., Jamar, C., Rochus, P., Chauvineau, J.P., Marioge, J.P., Catura, R.C., Lemen, J.R., Shing, L., Stern, R.A., Gurman, J.B., Neupert, W.M., Maucherat, A., Clette, F., Cugnon, P., van Dessel, E.L.: 1995, EIT: Extreme-Ultraviolet Imaging Telescope for the SOHO mission. *Solar Phys.* **162**, 291–312. doi:10.1007/BF00733432. <http://link.springer.com/article/10.1007%2FBF00733432>.
- Duvall, Jr., T.L., Gizon, L.: 2000, Time-distance helioseismology with f modes as a method for measurement of near-surface flows. *Solar Phys.* **192**, 177–191.
- Duvall, Jr., T.L., Birch, A.C., Gizon, L.: 2006, Direct measurement of travel-time kernels for helioseismology. *Astrophys. J.* **646**(1), 553–559. doi:10.1086/504792. <http://stacks.iop.org/0004-637X/646/i=1/a=553>.
- Duvall, Jr., T.L., Jefferies, S.M., Harvey, J.W., Pomerantz, M.A.: 1993, Time-distance helioseismology. *Nature* **362**, 430–432. doi:10.1038/362430a0.
- Duvall, Jr., T.L., Kosovichev, A.G., Scherrer, P.H., Milford, P.N.: 1996, Detection of subsurface supergranulation structure and flows from MDI high-resolution data using time-distance techniques. In: *American Astronomical Society Meeting Abstracts #188*, *Bull. Am. Astron. Soc.* **28**, 898.

- Duvall, Jr., T.L., Kosovichev, A.G., Scherrer, P.H., Bogart, R.S., Bush, R.I., De Forest, C., Hoeksema, J.T., Schou, J., Saba, J.L.R., Tarbell, T.D., Title, A.M., Wolfson, C.J., Milford, P.N.: 1997, Time-distance helioseismology with the MDI instrument: initial results. *Solar Phys.* **170**, 63–73. http://link.springer.com/chapter/10.1007/978-94-011-5236-5_4.
- Edmondson, J.K., Lynch, B.J., Antiochos, S.K., DeVore, C.R., Zurbuchen, T.H.: 2009, Reconnection-driven dynamics of coronal-hole boundaries. *Astrophys. J.* **707**(2), 1427. <http://stacks.iop.org/0004-637X/707/i=2/a=1427>.
- Feynman, R.P., Leighton, R.B., Sands, M.: 1963, *The Feynman Lectures on Physics: Volume 1, The Feynman Lectures on Physics 1*, Addison-Wesley, Boston, 7. Chap. 51.
- Fisk, L.A., Schwadron, N.A.: 2001, The behavior of the open magnetic field of the sun. *Astrophys. J.* **560**(1), 425. <http://stacks.iop.org/0004-637X/560/i=1/a=425>.
- Fisk, L.A., Zurbuchen, T.H., Schwadron, N.A.: 1999, On the coronal magnetic field: Consequences of large-scale motions. *Astrophys. J.* **521**(2), 868. <http://stacks.iop.org/0004-637X/521/i=2/a=868>.
- Giles, P.M.M.: 1999, Time-distance measurements of large scale flows in the solar convection zone. PhD thesis, Stanford University. Chap. 2. <http://soi.stanford.edu/papers/dissertations/giles/thesis/PDF/>.
- Gizon, L., Birch, A.C., Spruit, H.C.: 2010, Local helioseismology: Three-dimensional imaging of the solar interior. *Ann. Rev. Astron. Astrophys.* **48**, 289–338. doi:10.1146/annurev-astro-082708-101722.
- Gizon, L., Birch, A.C.: 2002, Time-distance helioseismology: The forward problem for random distributed sources. *Astrophys. J.* **571**, 966–986. <http://iopscience.iop.org/0004-637X/571/2/966>.
- Gizon, L., Birch, A.C.: 2004, Time-distance helioseismology: Noise estimation. *Astrophys. J.* **614**, 472–489. doi:10.1086/423367. <http://iopscience.iop.org/0004-637X/614/1/472>.

- Gizon, L., Birch, A.C.: 2005, Local helioseismology. *Living Rev. Sol. Phys.* **2**(6). <http://www.livingreviews.org/lrsp-2005-6>.
- Gorry, P.A.: 1990, General least-squares smoothing and differentiation by the convolution (Savitzky-Golay) method. *Anal. Chem.* **62**(6), 570–573. doi:10.1021/ac00205a007. <http://dx.doi.org/10.1021/ac00205a007>.
- Gouédard, P., Stehly, L., Brenguier, F., Campillo, M., Colin de Verdière, Y., Larose, E., Margerin, L., Roux, P., Sanchez-Sesma, F.J., Shapiro, N.M., Weaver, R.L.: 2008, Cross-correlation of random fields: Mathematical approach and applications. *Geophys. Prospect.* **56**(3), 375–393. doi:10.1111/j.1365-2478.2007.00684.x. <http://dx.doi.org/10.1111/j.1365-2478.2007.00684.x>.
- Harvey, K.L., Recely, F.: 2002, Polar coronal holes during cycles 22 and 23. *Solar Phys.* **211**, 31–52. <http://link.springer.com/article/10.1023/A:1022469023581>.
- Hassler, D.M., Dammasch, I.E., Lemaire, P., Brekke, P., Curdt, W., Mason, H.E., Vial, J.-C., Wilhelm, K.: 1999, Solar wind outflow and the chromospheric magnetic network. *Science* **283**(5403), 810–813. doi:10.1126/science.283.5403.810. <http://science.sciencemag.org/content/283/5403/810>.
- Hess Webber, S.A., Karna, N., Pesnell, W.D., Kirk, M.S.: 2014, Areas of polar coronal holes from 1996 through 2010. *Solar Phys.* **289**(11), 4047–4067. doi:10.1007/s11207-014-0564-0. <http://dx.doi.org/10.1007/s11207-014-0564-0>.
- Hindman, B.W., Gizon, L., Duvall Jr, T.L., Haber, D.A., Toomre, J.: 2004, Comparison of solar subsurface flows assessed by ring and timedistance analyses. *Astrophys. J.* **613**(2), 1253–1262. doi:10.1086/423263. <http://adsabs.harvard.edu/abs/2004ApJ...613.1253H>.
- Hoeksema, J.T., Scherrer, P.H.: 1986, An atlas of photospheric magnetic field observations and computed coronal magnetic fields: 1976-1985. *Solar Phys.* **105**, 205–211. doi:10.1007/BF00156388.

- Howe, R., Haber, D.A., Bogart, R.S., Zharkov, S., Baker, D., Harra, L., van Driel-Gesztelyi, L.: 2013, Can we detect local helioseismic parameter shifts in coronal holes? *J. Phys. Conf. Ser.* **440**(1), 012019. <http://stacks.iop.org/1742-6596/440/i=1/a=012019>.
- Ilonidis, S., Zhao, J., Kosovichev, A.: 2011, Detection of emerging sunspot regions in the solar interior. *Science* **333**, 993. doi:10.1126/science.1206253.
- Karna, N., Hess Webber, S.A., Pesnell, W.D.: 2014, Using polar coronal hole area measurements to determine the solar polar magnetic field reversal in solar cycle 24. *Solar Phys.* **289**(9), 3381–3390. doi:10.1007/s11207-014-0541-7. <http://link.springer.com/10.1007/s11207-014-0541-7>.
- Kirk, M.S., Pesnell, W.D., Young, C.A., Hess Webber, S.A.: 2009, Automated detection of EUV polar coronal holes during solar cycle 23. *Solar Phys.* **257**(1), 99–112. doi:10.1007/s11207-009-9369-y. <http://link.springer.com/10.1007/s11207-009-9369-y>.
- Kosovichev, A.G., Duvall Jr., T.L.: 1997, Acoustic tomography of solar convective flows and structures. In: Pijpers, F. P. and Christensen-Dalsgaard, J. and Rosenthal, C. S. (ed.) *SCORe '96: Solar Convection and Oscillations and their Relationship*, Springer Netherlands, Dordrecht, 241–260. ISBN 978-94-011-5167-2. doi:10.1007/978-94-011-5167-2_26. http://dx.doi.org/10.1007/978-94-011-5167-2_26.
- Lemen, J.R., Title, A.M., Akin, D.J., Boerner, P.F., Chou, C., Drake, J.F., Duncan, D.W., Edwards, C.G., Friedlaender, F.M., Heyman, G.F., Hurlburt, N.E., Katz, N.L., Kushner, G.D., Levay, M., Lindgren, R.W., Mathur, D.P., McFeaters, E.L., Mitchell, S., Rehse, R.A., Schrijver, C.J., Springer, L.A., Stern, R.A., Tarbell, T.D., Wuelser, J.-P., Wolfson, C.J., Yanari, C., Bookbinder, J.A., Cheimets, P.N., Caldwell, D., Deluca, E.E., Gates, R., Golub, L., Park, S., Podgorski, W.A., Bush, R.I., Scherrer, P.H., Gummie, M.A., Smith, P., Auken, G., Jerram, P., Pool, P., Soufli, R., Windt, D.L., Beardsley, S., Clapp, M., Lang, J., Waltham, N.: 2012, The Atmospheric Imaging Assembly (AIA) on the Solar

- Dynamics Observatory (SDO). *Solar Phys.* **275**, 17–40. doi:10.1007/s11207-011-9776-8. <http://link.springer.com/article/10.1007%2Fs11207-011-9776-8>.
- Lindsey, C., Braun, D.C., Jefferies, S.M.: 1993, Local helioseismology of subsurface structure. In: Brown, T.M. (ed.) *GONG 1992: Seismic Investigation of the Sun and Stars*, *ASP Conf. Ser.* **42**, 81.
- Linker, J.A., Mikić, Z., Biesecker, D.A., Forsyth, R.J., Gibson, S.E., Lazarus, A.J., Lecinski, A., Riley, P., Szabo, A., Thompson, B.J.: 1999, Magnetohydrodynamic modeling of the solar corona during Whole Sun Month. *J. Geophys. Res. Space Phys.* **104**(A5), 9809–9830. doi:10.1029/1998JA900159. <http://dx.doi.org/10.1029/1998JA900159>.
- Lobkis, O.I., Weaver, R.L.: 2001, On the emergence of the Green's function in the correlations of a diffuse field. *J. Acoust. Soc. Am.* **110**(6), 3011–3017. doi:<http://dx.doi.org/10.1121/1.1417528>. <http://scitation.aip.org/content/asa/journal/jasa/110/6/10.1121/1.1417528>.
- Martin, P.A.: 2006, *Multiple scattering: Interaction of time-harmonic waves with n obstacles*, *Encyclopedia of Mathematics and its Applications* **107**, Cambridge University Press, New York, 437. ISBN 9780521865548. doi:10.2277/0521865549. <http://www.cambridge.org/us/academic/subjects/mathematics/differential-and-integral-equations-dynamical-systems-and-co/multiple-scattering-interaction-time-harmonic-waves-ini-obstacles?format=HB>.
- McIntosh, P.S.: 1976, An empirical model for the origin and evolution of coronal holes. In: *Bulletin of the American Astronomical Society*, *Bull. Am. Astron. Soc.* **8**, 325.
- Morse, P.M., Feschbach, H.: 1953, *Methods of theoretical physics*, McGraw-Hill Book Company, Inc., New York.
- Morse, P.M., Ingard, K.U.: 1986, 8. *Theoretical acoustics*, Reprint edn., *International Series in Pure and Applied Physics*, Princeton University Press, Princeton. ISBN 9780691024011.

- Munro, R.H., Withbroe, G.L.: 1972, Properties of a coronal “hole” derived from extreme-ultraviolet observations. *Astrophys. J.* **176**, 511. doi:10.1086/151653.
- Navarro-Peralta, P., Sanchez-Ibarra, A.: 1994, An observational study of coronal hole rotation over the sunspot cycle. *Solar Phys.* **153**, 169–178. doi:10.1007/BF00712499.
- Neugebauer, M., Forsyth, R.J., Galvin, A.B., Harvey, K.L., Hoeksema, J.T., Lazarus, A.J., Lepping, R.P., Linker, J.A., Mikic, Z., Steinberg, J.T., von Steiger, R., Wang, Y.-M., Wimmer-Schweingruber, R.F.: 1998, Spatial structure of the solar wind and comparisons with solar data and models. *J. Geophys. Res.* **103**, 14587–14600. doi:10.1029/98JA00798. <http://onlinelibrary.wiley.com/doi/10.1029/98JA00798/abstract>.
- Nolte, J.T., Krieger, A.S., Timothy, A.F., Gold, R.E., Roelof, E.C., Viana, G., Lazarus, A.J., Sullivan, J.D., McIntosh, P.S.: 1976, Coronal holes as sources of solar wind. *Solar Phys.* **46**, 303–322. doi:10.1007/BF00149859. <http://link.springer.com/article/10.1007%2FBF00149859>.
- Ohmi, T., Kojima, M., Tokumaru, M., Fujiki, K., Hakamada, K.: 2004, Origin of the slow solar wind. *Adv. Space Res.* **33**, 689–695. doi:10.1016/S0273-1177(03)00238-2. <http://www.sciencedirect.com/science/article/pii/S0273117703002382>.
- Parker, E.N.: 1958, Dynamics of the interplanetary gas and magnetic fields. *Astrophys. J.* **128**, 664. doi:10.1086/146579.
- Pesnell, W.D., Thompson, B.J., Chamberlin, P.C.: 2012, The Solar Dynamics Observatory (SDO). *Solar Phys.* **275**, 3–15. doi:10.1007/s11207-011-9841-3. <http://link.springer.com/article/10.1007%2Fs11207-011-9841-3>.
- Pijpers, F.P.: 2006, *Methods in helio- and asteroseismology*, Imperial College Press, London. ISBN 1-86094-755-7.

- Pinto, R.F., Brun, A.S., Jouve, L., Grappin, R.: 2011, Coupling the solar dynamo and the corona: Wind properties, mass, and momentum losses during an activity cycle. *Astrophys. J.* **737**(2), 72. <http://stacks.iop.org/0004-637X/737/i=2/a=72>.
- Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P.: 1992, *Numerical recipes in C: The art scientific computing*, 2nd edn. Cambridge University Press, New York. ISBN 0521431085. doi:10.2277/0521431085.
- Roussev, I.I., Gombosi, T.I., Sokolov, I.V., Velli, M., Manchester IV, W., DeZeeuw, D.L., Liewer, P., Tóth, G., Luhmann, J.: 2003, A three-dimensional model of the solar wind incorporating solar magnetogram observations. *Astrophys. J. Lett.* **595**(1), L57. <http://stacks.iop.org/1538-4357/595/i=1/a=L57>.
- Savitzky, A., Golay, M.J.E.: 1964, Smoothing and differentiation of data by simplified least squares procedures. *Anal. Chem.* **36**(8), 1627–1639. ISBN 0003-2700. doi:10.1021/ac60214a047. <http://dx.doi.org/10.1021/ac60214a047>.
- Schatten, K.H., Wilcox, J.M., Ness, N.F.: 1969, A model of interplanetary and coronal magnetic fields. *Solar Phys.* **6**(3), 442–455. doi:10.1007/BF00146478. <http://dx.doi.org/10.1007/BF00146478>.
- Schou, J.: 2004, Low frequency modes. In: Danesy, D. (ed.) *SOHO 14 Helio- and Asteroseismology: Towards a Golden Future, ESA Special Publication* **559**, 134.
- Schou, J., Scherrer, P.H., Bush, R.I., Wachter, R., Couvidat, S., Rabello-Soares, M.C., Bogart, R.S., Hoeksema, J.T., Liu, Y., Duvall, T.L., Akin, D.J., Allard, B.A., Miles, J.W., Rairden, R., Shine, R.A., Tarbell, T.D., Title, A.M., Wolfson, C.J., Elmore, D.F., Norton, A.A., Tomczyk, S.: 2012, Design and ground calibration of the Helioseismic and Magnetic Imager (HMI) instrument on the Solar Dynamics Observatory (SDO). *Solar Phys.* **275**, 229–259. doi:10.1007/s11207-011-9842-2. <http://link.springer.com/article/10.1007%2Fs11207-011-9842-2>.

- Schrijver, C.J., DeRosa, M.L.: 2003, Photospheric and heliospheric magnetic fields. *Solar Phys.* **212**(1), 165–200. doi:10.1023/A:1022908504100. <http://dx.doi.org/10.1023/A:1022908504100>.
- Schwenn, R.: 2006, Space weather: The solar perspective. *Living Rev. Sol. Phys.* **3**(2). accessed 2014-07-17. doi:10.1007/lrsp-2006-2. <http://www.livingreviews.org/lrsp-2006-2>.
- Simon, G.W., Leighton, R.B.: 1964, Velocity fields in the solar atmosphere. III. Large-scale motions, the chromospheric network, and magnetic fields. *Astrophys. J.* **140**, 1120. doi:10.1086/148010.
- Tobias, S.M.: 2002, The solar dynamo. *Philos. Trans. R. Soc. London, Ser. A* **360**(1801), 2741–2756. doi:10.1098/rsta.2002.1090. <http://rsta.royalsocietypublishing.org/content/360/1801/2741>.
- Tsai, V.C.: 2010, The relationship between noise correlation and the Green's function in the presence of degeneracy and the absence of equipartition. *Geophys. J. Int.* **182**, 1509–1514. doi:10.1111/j.1365-246X.2010.04693.x. <http://gji.oxfordjournals.org/content/182/3/1509>.
- van de Hulst, H.C.: 1957, *Light scattering by small particles*, Reprinted by Dover, New York, 1981 edn. Wiley, New York.
- Waldmeier, M.: 1981, Cyclic variations of the polar coronal hole. *Solar Phys.* **70**, 251–258. doi:10.1007/BF00151332.
- Wang, Y.-M., Sheeley, N.R. Jr.: 2003, On the topological evolution of the coronal magnetic field during the solar cycle. *Astrophys. J.* **599**(2), 1404. doi:10.1086/379348. <http://stacks.iop.org/0004-637X/599/i=2/a=1404>.

- Wang, Y.-M., Sheeley, N.R. Jr., Walters, J.H., Brueckner, G.E., Howard, R.A., Michels, D.J., Lamy, P.L., Schwenn, R., Simnett, G.M.: 1998, Origin of streamer material in the outer corona. *Astrophys. J. Lett.* **498**, 165–168. doi:10.1086/311321. <http://iopscience.iop.org/article/10.1086/311321/meta>.
- Wang, Y.-M., Sheeley, N.R., Socker, D.G., Howard, R.A., Rich, N.B.: 2000, The dynamical nature of coronal streamers. *J. Geophys. Res.* **105**, 25133–25142. doi:10.1029/2000JA000149.
- Welch, B.L.: 1947, The generalization of ‘student’s’ problem when several different population variances are involved. *Biometrika* **34**(1-2), 28–35. doi:10.1093/biomet/34.1-2.28. <http://biomet.oxfordjournals.org/content/34/1-2/28.short>.
- Yang, M.-H., Chou, D.-Y., Zhao, H., Liang, Z.-C.: 2012, Phenomenological study of interaction between solar acoustic waves and sunspots from measured scattered wavefunctions. *Astrophys. J.* **755**(1), 10. <http://stacks.iop.org/0004-637X/755/i=1/a=10>.
- Zhou, X., Smith, E.J.: 2009, Solar cycle variations of heliospheric magnetic flux. *J. Geophys. Res. Space Phys.* **114**, 3106. doi:10.1029/2008JA013421.
- Zirker, J.B.: 1977, Coronal holes and high-speed wind streams. *Rev. Geophys.* **15**(3), 257–269. doi:10.1029/RG015i003p00257. <http://dx.doi.org/10.1029/RG015i003p00257>.

Biography

Shea A. Hess Webber received her PhD in Computational Sciences & Informatics through George Mason University in 2016, while working as a graduate research assistant at NASA's Goddard Space Flight Center with Dr. W. D. Pesnell. Dr. Hess Webber received her BS in Physics in 2009 from Gettysburg College, and MS in Computational Sciences in 2012 from GMU. Dr. Hess Webber worked at GSFC as a summer intern through the Scientific and Engineering Student Internship Program from 2005 through 2008. She returned to GSFC as a full-time GRA from 2009 through 2016, through a CEPHEUS Cooperative Agreement award. During the summers of 2013 and 2014, she spent several weeks collaborating with Dr. T. Duvall and others at Stanford University in Palo Alto, CA. In 2015, Dr. Hess Webber spent a month collaborating with Dr. Duvall and others at the Max Planck Institute for Solar System Research in Göttingen, Germany.

Dr. Hess Webber's previous research includes the evolution of polar coronal holes throughout the solar cycle using EUV data from the SOHO mission. Her dissertation involves the use of time-distance helioseismology to investigate elongated magnetic boundary regions in the photosphere, using both EUV and Doppler velocity data from the SDO mission. She has presented research at multiple conferences, most recently giving a dissertation talk at the 2016 SPD meeting in Boulder, CO, and poster presentation at the SDO 2016 meeting in Burlington, VT. She also served as Executive Secretary for a NASA HSR proposal review panel. Dr. Hess Webber has been awarded numerous student awards and held the position of Student Organizing Committee Chair at the 2011 LWS/SDO Workshop in Squaw Valley, CA, as well as being a committee member and post-president adviser to the new chair at the 2013 SDO meeting in Cambridge, MD. She has worked at many Education/Public Outreach events, including the AwesomeCon NASA booth, the Philadelphia Science Festival, and Career Day lectures at her former high school. Dr. Hess Webber is an inducted member of the National Physics Honor Society $\Sigma\Pi\Sigma$, a student member of the Solar Physics Division of the American Astronomical Society, and a United States citizen.

Publications:

Karna, N., Zhang, J., Pesnell, W.D., Hess Webber, S.A., *Study of the 3D Geometric Structure and Temperature of a Coronal Cavity Using the Limb Synoptic Map Method*, ApJ, 810, 124, 2015

Hess Webber, S.A., Karna, N., Pesnell, W.D., Kirk, M.S., *Areas of Polar Coronal Holes from 1996 Through 2010*, Sol Phys, 289, 4047, 2014

Karna, N., Hess Webber, S.A., Pesnell, W.D., *Using Polar Coronal Hole Area Measurements to Determine the Solar Polar Magnetic Field Reversal in Solar Cycle 24*, Sol Phys, 289, 3381, 2014

Kirk, M.S., Pesnell, W.D., Young, C.A., Hess Webber, S.A., *Automated Detection of EUV Polar Coronal Holes during Solar Cycle 23*, Sol Phys, 257, 99, 2009