






# Quantitative Image-Based Validation Framework for Assessing Global Coronal Magnetic Field Models

CHRISTOPHER E. RURA <sup>1,2</sup> VADIM M. URITSKY <sup>1,2</sup> SHAELA I. JONES <sup>1,2</sup> COOPER DOWNS <sup>3</sup>  
CHARLES NICKOLOS ARGE <sup>2</sup> AND NATHALIA ALZATE <sup>4,2</sup>

<sup>1</sup>*The Catholic University of America  
620 Michigan Avenue NE  
Washington, DC 20064, USA*

<sup>2</sup>*NASA Goddard Space Flight Center  
Code 670, Greenbelt, MD 20771, USA*

<sup>3</sup>*Predictive Science Inc.*

*9990 Mesa Rim Road, Suite 170, San Diego, CA 92121, USA*

<sup>4</sup>*ADNET Systems, Inc.*

*Greenbelt MD 20771, USA*

## ABSTRACT

Coronagraph observations provide key information about the orientation of the Sun’s magnetic field. Previous studies used quasi-radial features detected in coronagraph images to improve coronal magnetic field models by comparing the orientation of the features to the projected orientation of the model field. Various algorithms segment these coronal features to approximate the local POS geometry of the coronal magnetic field, and their orientation can be used as input for optimizing and constraining coronal magnetic field models. We present a new framework that allows for further quantitative evaluations of image-based coronal segmentation methods against magnetic field models, and vice-versa. We compare quasi-radial features identified from QRaFT, a global coronal feature tracing algorithm, in white-light coronagraph images to outputs of MAS, an advanced MHD model. We use the FORWARD toolset to produce synthetic pB images co-aligned to real coronagraph observations, segment features in these images, and quantify the difference between the inferred and model magnetic field. This approach allows us to geometrically compare features segmented in artificial images to those segmented in white-light coronagraph observations against the POS projected MAS coronal magnetic field. We quantify QRaFT’s performance in the artificial images and observational data, and perform statistical analyses that measure the similarity of these results to compute the accuracy and uncertainty of the model output to the observational data. The results demonstrate through a quantitative evaluation that a coronal segmentation method identifies the global large-scale orientation of the coronal magnetic field within  $\sim \pm 10^\circ$  of the POS projected MAS magnetic field.

*Keywords:* Solar Corona (1483) — Coronagraphic imaging (313) — Solar magnetic fields (1503) — Cross-validation (1909) — Model selection (1912) — Solar coronal streamers (1486)

## 1. INTRODUCTION

### 1.1. Overview

The solar corona is an important region of the solar atmosphere that sets the conditions of space weather propagation into the Heliosphere. The coronal magnetic field is dominant in this region, making it the key to the physics behind most phenomena observed in this region of the sun (Solanki et al. 2006). Despite its dominance, the magnetic field in the atmospheric layers above the photosphere is notoriously difficult to measure, and therefore we must rely on extrapolations from indirect measurements of the surface magnetic field using the Zeeman effect taken from ground-

based observatories to offer the most accurate understanding of the coronal magnetic field (Solanki et al. 2006). For this reason, magnetic field models are key for accurately representing the coronal magnetic field, and being able to assess them is essential for accurate space weather forecasting.

In this paper we present a method to evaluate an image based method for assessing coronal magnetic field models. This method was developed by Jones et al. (2016, 2020), who presented a method to optimize magnetic field models by comparing the orientation of quasi-radial features segmented in polarized brightness (pB) observations to the orientation of the magnetic field model. This approach depends on the assumption that structure seen in coronal images can be treated as a proxy for the orientation of features in the coronal magnetic field. In this paper we will test the accuracy of this assumption.

Magnetic field line tracing methods use optical features observed in coronal images to segment large-scale signatures of the coronal magnetic field to measure its orientation. A major open problem in space physics is that there is little to no ground-truth precedent in which to compare the orientation of these segmented features to, which makes it difficult to determine if they are accurately representing magnetic structure. While magnetic field models such as Potential-Field Source Surface (PFSS) and Magnetohydrodynamic (MHD) models can approximate the coronal magnetic geometry, these models are challenging to constrain due to the absence of magnetic field measurements in the corona. Evaluating and validating these methods against a ground-truth precedent may illuminate to what extent we can rely on the features they produce.

### 1.2. Theoretical Basis

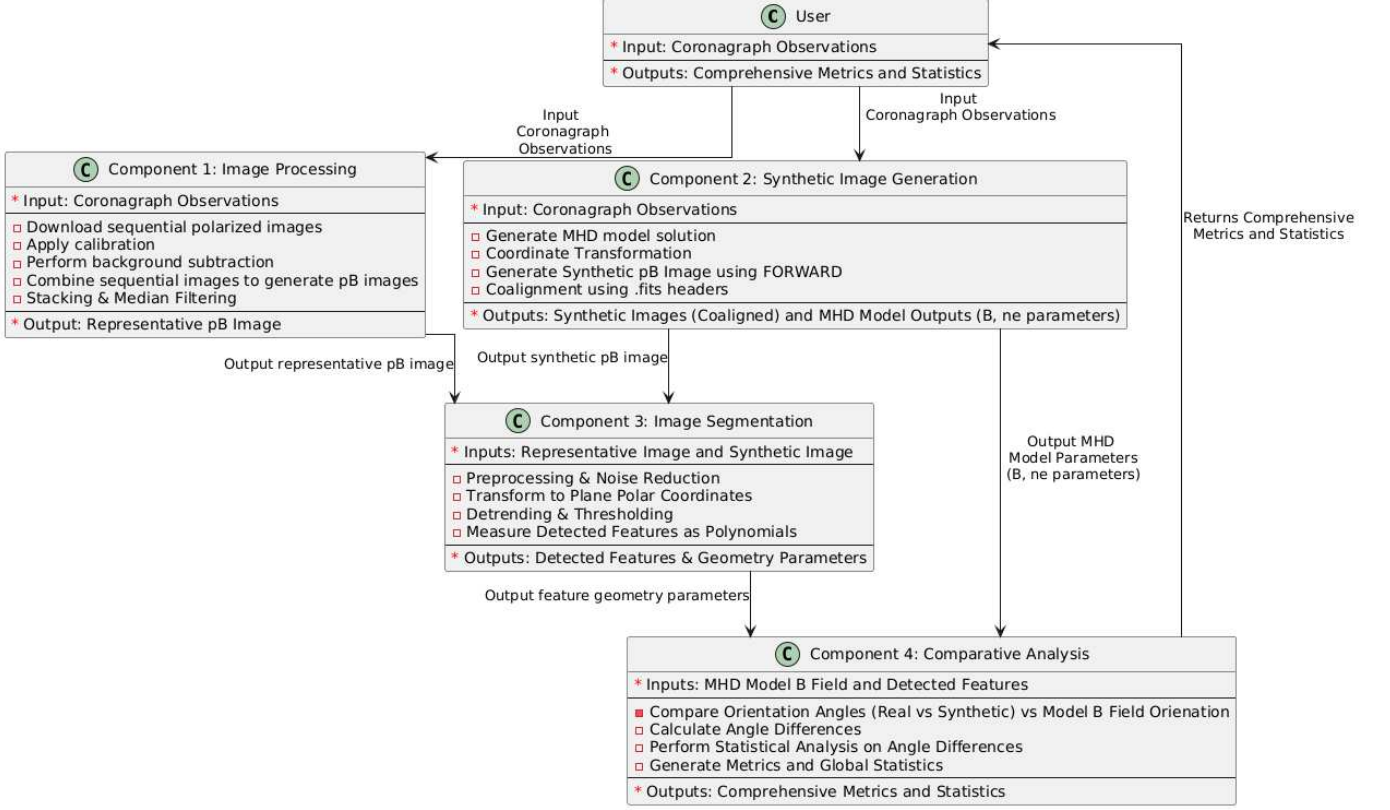
The work of Jones et al. (2020) used features segmented in pB images, which highlight the K-corona. These features are a result of intensity gradients present in the pB image, however this structure can be connected to the plasma density distribution in the K-corona. The K-corona itself is the result of photons emitted from the photosphere scattering off of coronal free electrons (Billings 1966). This process is known as Thomson scattering. Since the electric field vector is normal to the direction of propagation for electromagnetic (EM) radiation, the K-Corona is polarized (Blackwell et al. 1967). As discussed in Howard & Tappin (2009), the pB of the white-light corona can be modeled by the integration of the electron density along an observer’s line of sight. This integration must take into account several factors, such as the scattering function, the decreasing illumination of electrons with distance from the sun, and the decreasing degree of polarization with distance from the observer’s plane-of-sky (POS).

The frozen-in-flux condition states that the bulk motion of an ideal MHD plasma cannot move perpendicularly to the magnetic field ( $\vec{B}$ ) without it bringing along, or vice versa. The magnetic field therefore behaves as if it is ‘frozen-in’ with the plasma, so that the magnetic field moves with the plasma. This condition applies in highly conductive plasmas, such as in the corona and solar wind, as it accelerates into the Heliosphere (Murray 2013). Due to this, the plasma is intimately connected to the magnetic field in this region of the solar atmosphere, as the plasma evolution along individual flux tubes is isolated from neighboring flux tubes. The practical effect of this is that any density and temperature contrasts between adjacent flux tubes will naturally highlight the structure of the magnetic field itself (Aschwanden 2005). Therefore, the structure seen in coronal pB images should in theory serve as a proxy for the orientation of features in the coronal magnetic field.

### 1.3. Framework for Assessing Magnetic Field Models

In the Jones et al. (2016, 2020) studies, features are segmented in coronagraph images using the automated Quasi-Radial Feature Tracing Algorithm (Uritsky 2022) (QRaFT), which uses adaptive thresholds to enhance and segment coronal features and approximates their projected geometries using polynomials. This particular algorithm is of interest to validate, as it focuses on tracing open field line geometries, which have very weak optical signatures and are typically difficult to measure. These feature tracings are then compared to the projection of the model coronal magnetic field, assuming that the features observed in the images reflect the projected magnetic field in the image plane.

In this study, we will test this assumption using numerical outputs of the Magnetohydrodynamic Algorithm outside a Sphere Model (MAS) from PSI, which is described in section 2.4. This MHD model is useful because it incorporates a complete energy equation, which is necessary in describing both the plasma and magnetic properties of the corona (Mikić et al. 2018). We use the *SSWIDL* package FORWARD to create synthetic white-light images based on the MAS model solution for the 2017 solar eclipse (Mikić et al. 2018). These images will be as seen from the perspective of the STEREO satellite at times corresponding to the creation of real coronagraph images obtained by the STEREO COR-1 Coronagraph Instrument (COR-1).



**Figure 1.** Flowchart style overview of defined framework for quantitatively assessing image-based coronal segmentations against magnetic field models, and vice-versa. The inputs and outputs for each component are listed, as well as the general steps for each component. Each arrow represents an output from a component of the framework that is used as an input into the component the arrow is connected to.

With the **MAS** model providing a ground truth magnetic field in which to evaluate against, we segment features with the **QRaFT** algorithm, and use a series of methods to measure how well the orientations of quasi-radial plasma density features detected by **QRaFT** match the model field orientation. These methods will be discussed in more detail in section 2. By comparing the optical structure of quasi-radial features detected in coronal images to synthetic images generated by this **MHD** model as well as its expected magnetic field, we seek to provide a unique evaluation of this method’s ability to segment true magnetic structure. We present our results in this analysis in section 3, discuss potential errors in section 4, and state our conclusions in section 5.

Using these methods to produce the results shown in section 3, we have also defined a framework in which to further evaluate image-based coronal segmentation methods against model magnetic fields and vice-versa quantitatively. A high level overview of this framework is shown in figure 1, emphasizing the generality of this approach, which allows for the evaluation of segmentations from multiple data sources and segmentation algorithms against multiple magnetic field models. The potential use of this framework will be further discussed in section 4.

## 2. DATA AND METHODOLOGY

### 2.1. Synthesizing **MHD** Model Solution into Synthetic Images

As mentioned in component 2 of figure 1, we use the toolset **FORWARD** in our framework to transform the 3D **MAS** solution into **POS** view using Carrington Heliographic coordinates, where the **POS** is given by the Y and Z axes with the X-axis representing the Sun-observer line (Gibson et al. 2016). Section 1.2 noted that the K Corona’s **pB** may be computed as the integration of the electron density along an observer’s **line-of-sight** (LOS), and then applied to a scattering function that acts to vary the density along the **line-of-sight** (Howard & Tappin 2009). Schmit et al. (2009) provides an explicit equation to calculating the **pB** in a **MAS** context. **FORWARD** utilizes a scattering function that depends on both the radial distance to the Photosphere and the geometry of the observer relative to the

scattering point to incorporate illumination of the K Corona due to Thomson scattering (Gibson et al. 2016). The electron density for each pixel in the various planes of the solution is integrated along the *line-of-sight* of a hypothetical observer (in this case the X axis), and is applied to the scattering function. This process results in a synthesized *pB* image in the *POS* view from the perspective of a hypothetical observer. These synthetic *pB* images are then co-aligned to each respective observation by using *WCS* coordinates extracted from each observation’s *FITS* header as described in Thompson (2010). Additionally, we estimate the radius of the occulting disk for each synthetic *pB* image using parameters from each observation’s *FITS* header. Figure 4 shows (left) a *COR-1* polarized brightness observation used in this analysis, taken on 2017-08-29, and (right) the synthetic polarized brightness projected onto the *plane-of-sky* from the equivalent observational perspective with intensity displayed in a log scale.

## 2.2. White-Light Observations

While solar eclipses provide excellent opportunities to observe the solar corona in great detail, they do not occur frequently and are only visible from a small band of area called the path of totality. To make matters worse, this band changes with every eclipse, making it impossible to select one site that is always in this path. It is estimated that a given fixed site on Earth is able to view a total solar eclipse once every 350 years, however this estimate varies by latitude (Pasachoff 2009). Therefore, scientific instruments known as coronagraphs are needed to artificially block out the sun’s light to consistently study the solar corona. The white-light K-Corona is generated due to the electron scattering of photospheric light, meaning it is linearly polarized in a tangential direction to the solar limb (de Wijn et al. 2012). Coronagraphs take advantage of this by measuring the polarized brightness (*pB*) of the corona, which is directly related to its electron density (de Wijn et al. 2012).

The *STEREO COR-1 Coronagraph Instrument (COR-1)* is flown aboard NASA’s *STEREO A* spacecraft which, along with *STEREO B*, had the goal of observing the sun simultaneously from different perspectives. Like other coronagraphs, it can measure the *pB* of the corona. Its design is described more completely in Thompson et al. (2003), but in short, it is an occulting refractive coronagraph that utilizes an internal Lyot stop that is adapted to be used in space (Thompson et al. 2010).

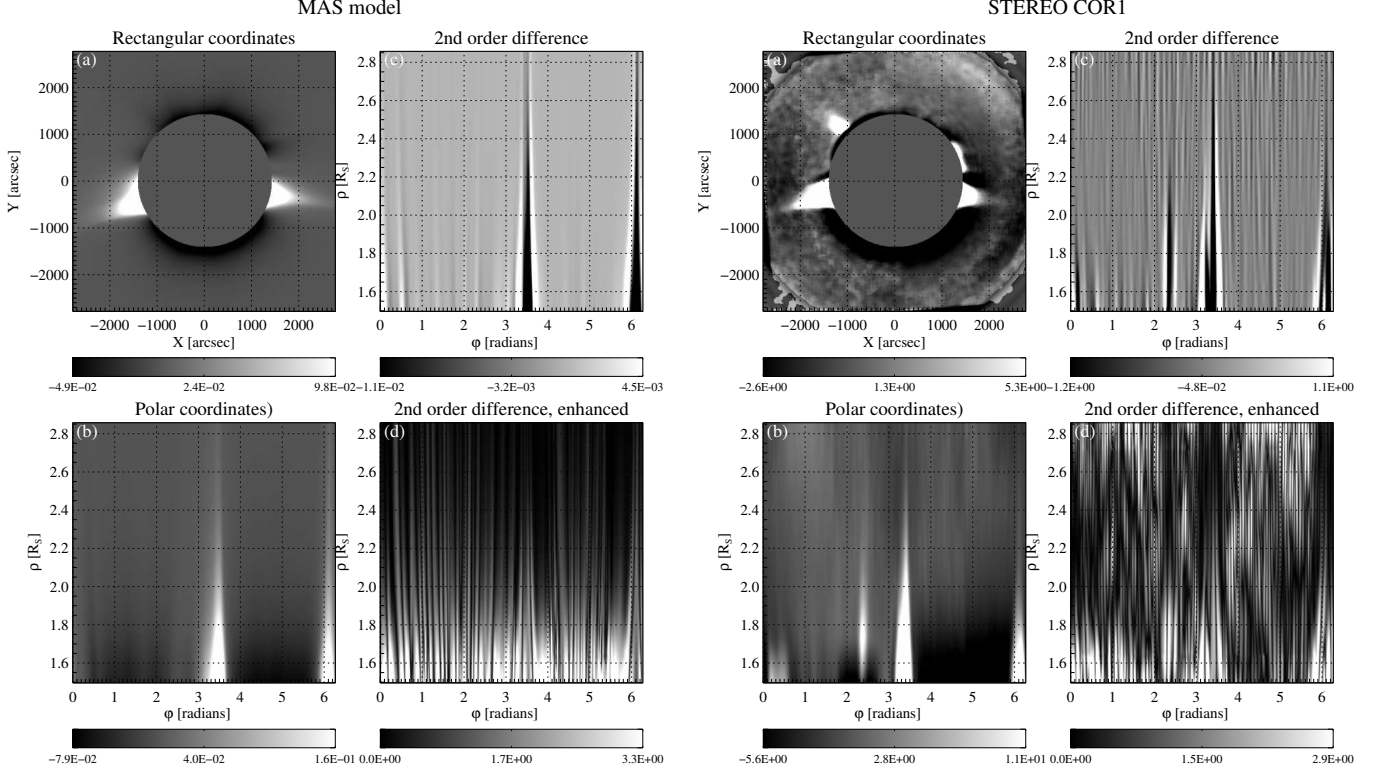
Applying this data to our framework, as mentioned in the first component of figure 1, images from this data source are downloaded at level 0 and processed to level 1 using the SolarSoft *secchi\_prep* routine. This process involves background subtraction, vignetting, flat-field correction, CCD bias correction, and other calibration processing. This routine was run with the keywords *polariz\_on*, *pB*, and *rotate\_on* to calculate the polarized brightness and rotate the image to solar north. This routine combines three polarized images, (rotated at 0°, 120°, and 240°), taken in sequence and calculates them into one *pB* image. These images were produced at a 10 minute cadence, and we aggregate images over an 8-hour timespan to retrieve 50 *pB* images. These images are then stacked and a median filter is applied to the stack to create a representative image in order to enhance signal and filter out noise. A similar process is described in Thompson et al. (2010). This process is useful for identifying quasi-radial plasma density features in *COR-1* observations as the corona’s large scale quasi-radial features remain static over a short period of time, allowing us to enhance the signal of these features and filter out imperfections such as cosmic rays and other image artifacts. Despite this enhancement, there are still artifacts present in the final output of this process, which will be discussed in further detail in section 4.

## 2.3. QRaFT Methodology

As mentioned in section 1.3, we use the *QRaFT* code as our image-based segmentation method of choice in our framework, (component three of figure 1). The *QRaFT* code enables an automatic detection of quasi-radial coronal features, which are expected to be approximately aligned with the local coronal magnetic field. The code uses a coronagraph image as an input; the *QRaFT* output is an ensemble of quasi-linear detected features approximated by polynomials, as well as a set of parameters describing the geometry of the features, including the *POS* orientation angles used in this study. A detailed description of the *QRaFT* methodology will be published elsewhere; here we present a concise outline of the main processing steps.

The studied coronal image  $I(x, y)$  first undergoes a preprocessing procedure aimed to reduce the level of the high-frequency pixel noise and to mitigate the systematic decay of the image intensity with the distance from the solar limb. The noise reduction is performed using a median boxcar filtering; the radial detrending is based on subtracting an azimuthally-averaged large-scale radial trend from the processed image.

At the next step, the coronagraph image is transformed into plane polar coordinates  $(\phi, \rho)$ , with the position angle  $\phi \in [0, 2\pi]$  measured in the counterclockwise direction relative to the western solar limb and the radial coordinate  $\rho$



**Figure 2.** QRaFT processing steps shown for (left) a synthetic polarized brightness (pB) image produced by MAS and the FORWARD code and (right) an observed pB image obtained from COR-1. In each example the upper left panel (a) is the original COR-1 pB image after radial detrending and smoothing, the bottom left panel (b) is the same image in plane polar coordinates, the top right panel (c) is the signed second-order azimuthal difference, and the bottom right panel (d) is the unsigned second order difference after azimuthal detrending revealing a fine quasi-radial structure in the example observed image.

representing the plane-of-sky (POS) distance from the solar disk center. This coordinate system is a natural choice for studying quasi-radial coronal structures in open-flux solar regions since such structures are characterized by a slowly varying position angle. Other types of structures such as closed loops require a different approach and lie outside of the QRaFT’s applicability domain.

To enhance the quasi-radial structures, we compute the unsigned second-order position-angle difference over a specified characteristic azimuthal scale  $\Delta\phi$ :

$$\Delta^2 I(\phi, \rho) = |I(\phi - \Delta\phi/2, \rho) + I(\phi + \Delta\phi/2, \rho) - 2I(\phi, \rho)| \quad (1)$$

We found that the second-order azimuthal difference to be a sensitive marker of field-aligned image structures, with the absolute value operator used in Eq. 1 enabling a consistent tracing of the structure shape on either side of the structure.

The enhanced image (1) is detrended in the position angle direction to improve the detection results for the structures that are characterized by substantially different brightness and signal-to-noise levels. The detrending is achieved by dividing the enhanced image by its low-pass filtered counterpart obtained by boxcar averaging the array  $\Delta^2 I(\phi, \rho)$  along the  $\phi$  direction.

Periodic boundary conditions were applied at the edges of the  $\phi$  interval for when performing the azimuthal differencing and detrending.

The resulting enhanced image is subject to an adaptive thresholding allowing us to identify contiguous clusters of image pixels associated with quasi-radial coronal features of interest. The method “label region” of the IDL language is used to identify and separate all clusters of pixels satisfying the chosen detection threshold. The cluster detection is performed multiple times by using different detection thresholds representing different percentile levels of the studied



image, as well as by varying the lower bound  $\rho_{min}$  of the processed range of radial positions to focus the algorithm on different coronal altitudes. We use between 10 and 20 percentile threshold choices and 10  $\rho_{min}$  values for processing each coronal image, which results in 100 - 200 individual tracing runs. Combining the outputs of these runs provides a relatively uniform detection performance at different locations around the Sun.

The clusters produced by each tracing run are approximated by polynomial functions using a nonlinear least-square fitting. These fits are used to calculate the POS orientation angles along the central line of each of the detected structures. The orientation angles produced by the QRaFT algorithm are compared with the local magnetic field direction of the MAS model as described in section 2.6. Figure 2 shows an example of each of these processing steps on a COR-1 pB observation used in this study.

#### 2.4. PSI MAS Model

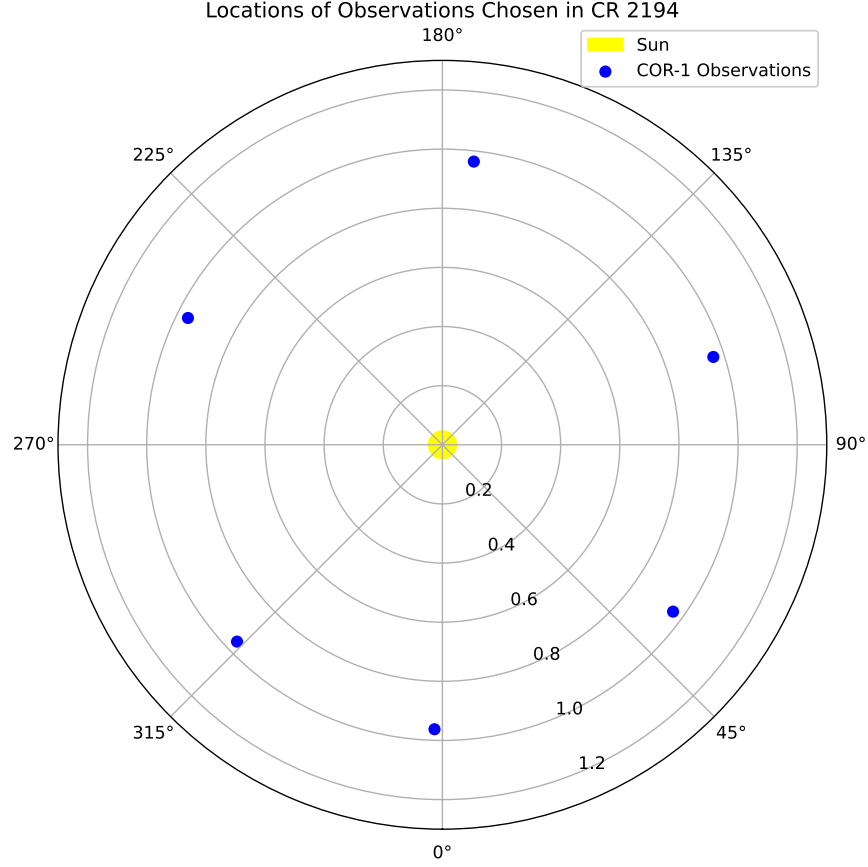
To create a ground-truth testing framework for the QRaFT technique we can leverage outputs from state-of-the-art numerical models of the solar corona. Component two of figure 1 showcases how these outputs are used in the high level overview of this framework. To this end we employ the Magnetohydrodynamic Algorithm outside a Sphere Model (MAS) model of the global solar corona (Mikić et al. 1999; Lionello et al. 2009; Mikić et al. 2018). This model (and similar codes, e.g., van der Holst et al. 2014; Réville et al. 2020) employs a ‘thermodynamic’ MHD approach, where the additional terms that describe energy flow in the corona and solar wind are included, including coronal heating, parallel thermal conduction, radiative loss, and Alfvén wave acceleration. This treatment is essential for capturing the thermal-magnetic state of the corona, including the all-important interplay between magnetic and hydrodynamic forces that open the corona and form the solar wind. Realistic magnetic configurations for targeted time-periods are obtained by using measurements of the photospheric magnetic field as the primary boundary condition. Most importantly, using such a model enables the direct computation and comparison of forward modeled observables to real observations. This technique can be used to both directly constrain the coronal model and connect coronal observables to their underlying physical state of the plasma (Lionello et al. 2009; Boe et al. 2021, 2022).

For this particular study we use the high-resolution MAS coronal prediction simulation developed for the August 21, 2017 total solar eclipse. This simulation is fully described by Mikić et al. (2018), but can be summarized as follows: The simulation utilizes a Wave-Driven-Turbulence (WTD) approach to specify coronal heating (Lionello et al. 2014; Downs et al. 2016) and an energization technique to add field aligned currents (shear/twist) over large-scale polarity inversion lines (Mikić et al. 2018; Yeates et al. 2018). To model the corona at a given time, the model must specify a full-sun map of the radial component of the magnetic field at the inner boundary. Because of the lead-time required for publishing the prediction, this simulation used a handmade splice of synoptic map data from SDO/HMI (Scherrer et al. 2012) based on what was available about 10 days prior to the eclipse. This map combined data for Carrington rotation (CR) 2192 with near-real-time data from a part of CR 2193. To capture plume-like density structures at the poles, the polar caps were also filled with a random flux distribution whose net matches observations on the net flux at high latitudes. All told, the boundary magnetic field measurements span a time-range of approximately July 16 to August 11, 2017, which placed the oldest data near the west limb during totality.

#### 2.5. Data/Model Comparison

To generate a pair of comparison images, we start with the MAS model simulation calculated at the time of the August 2017 solar eclipse. We use the toolset HelioWeb (2017) to find STEREO A’s position in Carrington Heliographic coordinates at a particular time and then query the COR-1 database in order to find observations corresponding to this time. We generate new model images from the MAS solution as described in section 2.1 from this perspective using the retrieved Carrington Heliographic latitude and longitude. We then output the appropriate model data, (which will be further described in section 2.8), for use in the segmentation comparison, which will be described in section 2.6. We then determine the next observation time corresponding to a 60° rotation of longitude and start the process again, resulting in a pair of six model slices in an approximately 360° span and accordingly six white-light images for each observatory taken over a 17 day span at the same perspective as the model images in the POS view. This is done to generate a dataset averaged over a rotation of the Sun in order to have an unbiased comparison with regards to the viewing angle of the Sun. Figure 3 shows the longitudinal distribution of these six observations across CR 2194, which corresponds to the time period of the 2017 solar eclipse and this study.

Figure 4 shows a comparison of a COR-1 observation on 2017-08-29, generated as described in section 2.2, along with a corresponding FORWARD pB model slice. Since we co-align these synthetic pB images to the COR-1 observations



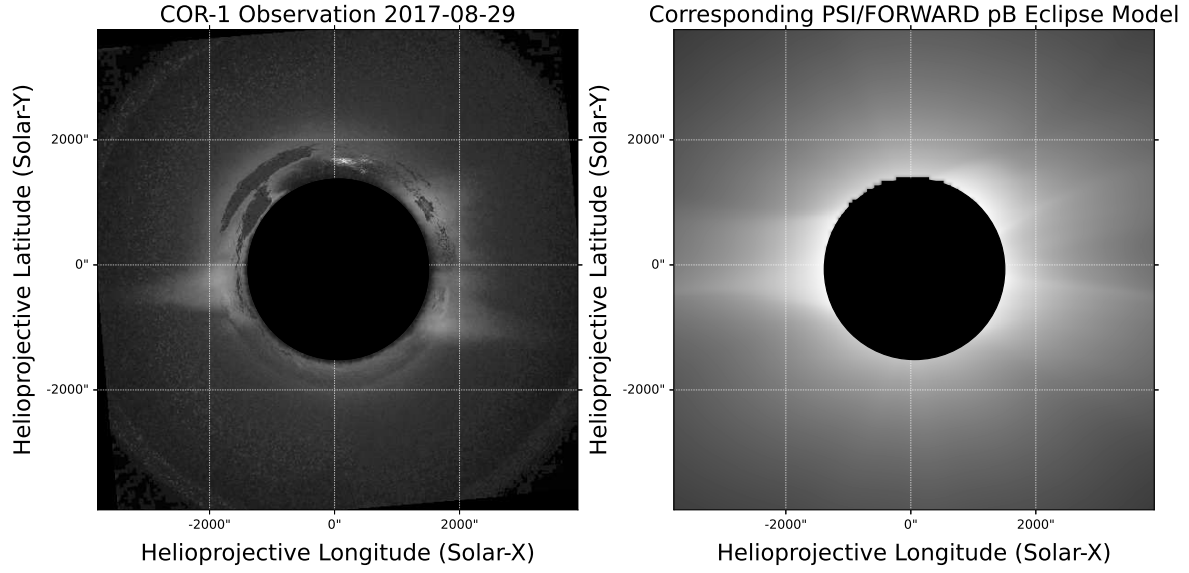
**Figure 3.** Polar plot showing the Carrington longitudes of each [COR-1](#) observation chosen in this study. These six observations correspond to [CR 2194](#) and were chosen to provide roughly  $60^\circ$  difference of Carrington longitude between each observation.

respectively, and compare each from an equivalent perspective in space, we are thus directly comparing the orientation of the observed corona to the model generated synthetic corona.

### 2.6. Calculating Angular Discrepancies

As in [Jones et al. \(2016, 2020\)](#), we utilize an angular discrepancy scheme that compares the orientation of the features segmented by [QRaFT](#) to the projected orientation of the model magnetic field for both real and synthetic [pB](#) images. This scheme is shown in [figure 6](#) and is the beginning stage of the final component of our framework, (as shown in component four of [figure 1](#)). By then comparing the results of this analysis to the results using the coronagraph observations [COR-1](#), taken from the same perspective as the synthetic [pB](#) images, we are able to evaluate how well the [QRaFT](#) algorithm performs.

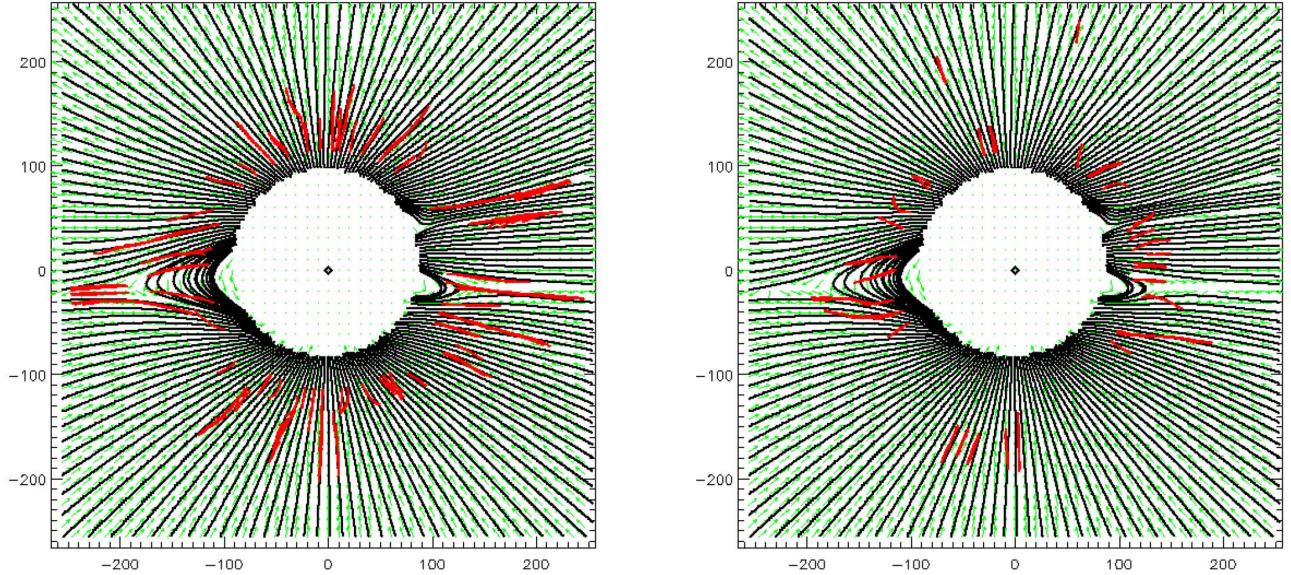
In order to generate a statistic on how well the [QRaFT](#) feature tracing method determined the orientation of each feature to the orientation of the model field, we find the angular discrepancy between the orientation of the traced feature to the orientation of the model field, as shown in [figure 6](#). For each traced feature, we compare the orientation of the slope of the polynomial that approximates the feature's orientation to the expected magnetic orientation from the model for each pixel in said feature. The expected magnetic orientation of the model field is given by the  $B_z$  vs  $B_y$  magnetic field components, given that we are using Carrington Heliographic coordinates in the plane of sky view. An angle discrepancy of  $0^\circ$  would constitute a perfect match between the orientation of the traced feature and the expected magnetic orientation from the model at a given pixel.



**Figure 4.** (Left) [COR-1](#) polarized brightness observation on 2017-08-29. (Right) Synthetic polarized brightness image generated from [FORWARD](#) co-aligned to the equivalent observational perspective as the [COR-1](#) observation.

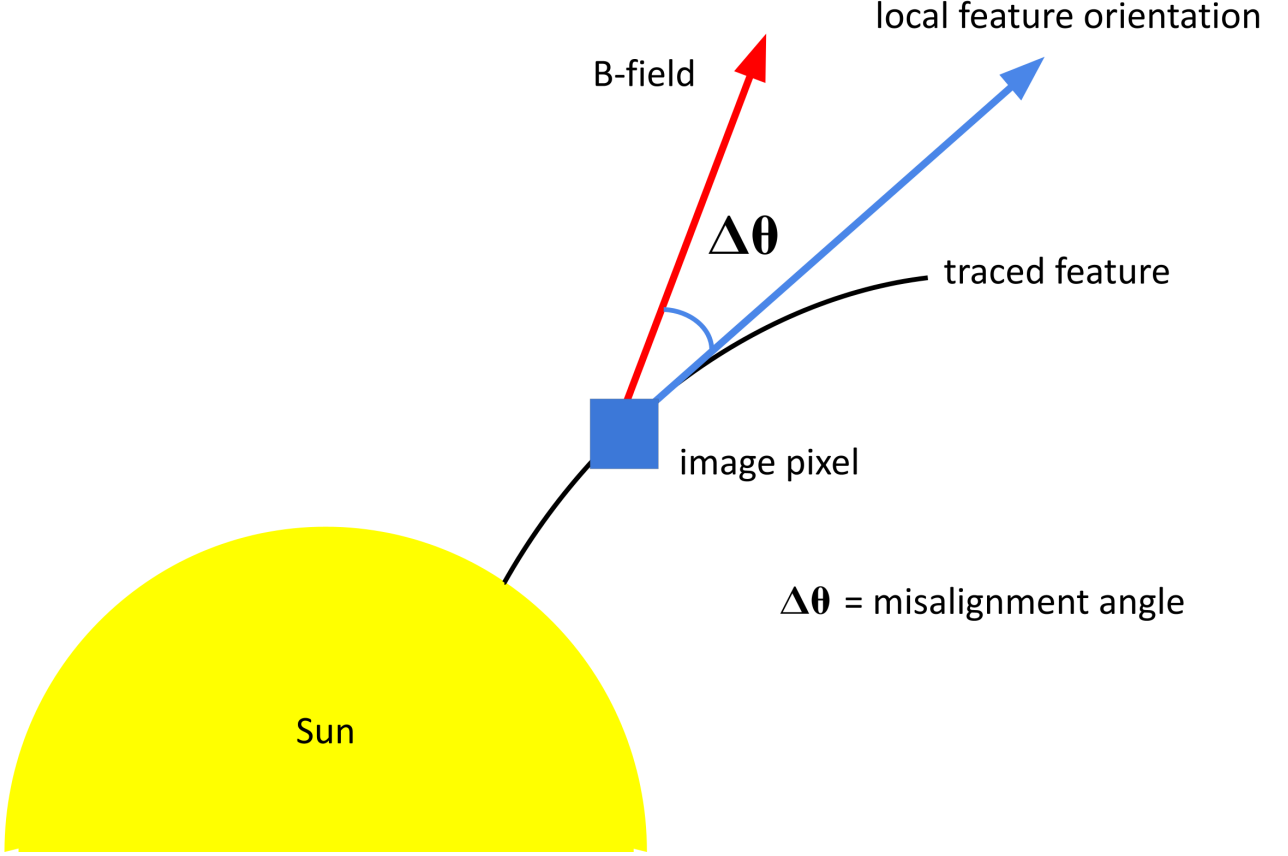
2017-08-29 PSI/FORWARD  $n_e$  Eclipse Model

2017-08-29 COR-1 Observation



**Figure 5.** (Left) Model [plane-of-sky magnetic field](#) ( $\text{POS } \vec{B}$ ) orientation with QRaFT filtered features from [MAS  \$n\_e\$](#)  POS model slice for 2017-08-29. (Right) Model  $\text{POS } \vec{B}$  orientation with QRaFT filtered features from [COR-1](#) representative median image for 2017-08-29.





**Figure 6.** Sketch of angle discrepancy calculation used to determine performance of the [QRaFT](#) feature tracing algorithm. For each pixel the expected magnetic field orientation, given by the PSI model's  $B_z$  vs  $B_y$  magnetic field components, is compared to the orientation of the segmented feature determined by the [QRaFT](#) algorithm. The angular difference is the statistic in which we determine how well the algorithm segmented each feature compared to it's known solution given by the PSI model, with an angular discrepancy of  $0^\circ$  indicating a perfect fit.

We define  $\vec{v}_1$  as the orientation of the plasma density feature traced by [QRaFT](#) in [POS](#) coordinates. We define this orientation as:

$$\vec{v}_1 = x_r \hat{x} + y_r \hat{y} \quad (2)$$

where  $x_r$  and  $y_r$  are the horizontal and vertical components of the feature segments in the [POS](#).

We define  $\vec{v}_2$  as the orientation of the of the  $\vec{B}$  from the PSI model solution translating to [POS](#) coordinates via [FORWARD](#). We define this 'B-vector' as:

$$\vec{v}_2 = \vec{B}_y \hat{x} + \vec{B}_z \hat{y} \quad (3)$$

where  $B_y$  and  $B_z$  are the  $y$  and  $z$  components of the central [POS](#) magnetic field calculated by the [MAS](#) model.

We define an angular discrepancy  $\Delta\theta$  as the angle between the orientation of [QRaFT](#)'s polynomial feature,  $v_1$ , and the orientation of the expected magnetic field generated by the model,  $v_2$ , at a given pixel. Since  $\vec{v}_1 \cdot \vec{v}_2 = |v_1| |v_2| \cos \theta$  and  $v_1^\perp \cdot v_2 = |v_1| |v_2| \sin \theta$  ([Hill 1994](#)), this angle can be calculated using the following formula.

$$\Delta\theta = \tan^{-1} \left( \frac{v_1 \hat{x} * v_2 \hat{y} - v_1 \hat{y} * v_2 \hat{x}}{v_1 \hat{x} * v_2 \hat{x} + v_1 \hat{y} * v_2 \hat{y}} \right) \quad (4)$$

This angle has a range of  $-90^\circ < \Delta\theta < 90^\circ$  and defines the correlation between the orientation of the predicted model [plane-of-sky magnetic field \(POS  \$\vec{B}\$ \)](#) and the projected orientation of plasma density features given by [QRaFT](#).

### 2.7. Statistical Analysis of Angular Discrepancies

The final step of our framework, as shown in the second half of component four in figure 1, is the generation of comprehensive statistics and metrics for this evaluation. We compute each pixel's angle discrepancy for every feature that is traced by [QRaFT](#), and calculate two sets of statistics evaluating the performance of [QRaFT](#)'s traced features. The first set of statistics we calculate are the median and average discrepancies of the orientation of the traced features when compared to the expected magnetic field predicted by the model. We do this to quantify [QRaFT](#)'s performance when compared to the model's result, as well as to signify outliers and compare the similarity of the median and average values of each dataset. Along with the average discrepancy we also calculate the standard deviation and a 95% confidence interval<sup>1</sup> of the average discrepancy.

The second set of statistics we calculate are more sophisticated metrics aimed at comparing the properties of each dataset to one another. These include the [Jensen-Shannon Divergence \(JSD\)](#), kurtosis, skew, and [Tukey's Honestly Significant Difference Test \(HSD\)](#). We use these statistics to measure the similarity of [QRaFT](#)'s performance on white light coronagraph data when compared to simulated [pB](#) images, comparing both to the expected magnetic field generated by the model. To begin, we calculate a histogram and corresponding probability density<sup>2</sup> for each dataset from the results of the angular discrepancy analysis in order to evaluate the performance of the [QRaFT](#) code on each of the images when compared to the model magnetic field orientation. This process allows us to approximate the discrete distributions of the histograms as a smooth analytical function that can be used to generate statistical comparisons of the dataset.

In order to generate a statistic for the similarities of the two distributions in each case, we use the [Jensen-Shannon Divergence \(JSD\)](#)<sup>3</sup>, which is a statistical means of determining how similar the shapes of two probability distributions are. The closer this statistic is to 0, the more similar the two distributions are. The probability distributions are normalized by amplitude when evaluating this statistics in order to evaluate the relative difference in shape of each distribution.

We use [Tukey's Honestly Significant Difference Test \(HSD\)](#)<sup>4</sup> to evaluate the relative difference in mean values between each data population described in section 2.8. Performing this analysis allows us to find the relative error contribution between each population, and since the populations vary in complexity starting from the ideal one-to-one comparison of the [MAS central plane-of-sky electron density](#), to the [MAS line-of-sight integrated electron density](#), to the [FORWARD polarized brightness](#), and finally to the white light [COR-1](#) data. We note that in our use of the [HSD](#), we evaluate this statistic at  $\alpha = 0.05$ , corresponding to a 95% confidence interval. Therefore, in each evaluation, this metric determines if the difference of means between each population is statistically significant corresponding to a 95% confidence interval<sup>5</sup>.

### 2.8. Choosing Data Populations in which to Test Framework

In order to test the data to model comparison and decompose the relative error contributions based on our assumptions, we choose four populations when comparing the spatial geometry of features segmented by [QRaFT](#) to the central [POS](#) magnetic field. First, we retain the [MAS central plane-of-sky electron density \(MAS  \$n\_e\$  POS\)](#) of each model slice to use as a one-to-one comparison of the orientation of plasma density features in the plane of sky to the central [POS](#) magnetic field. We then retain the [MAS line-of-sight integrated electron density \(MAS  \$n\_e\$  LOS\)](#) integrated along the [line-of-sight](#) of the observer as the second population. We then retain the [FORWARD polarized brightness \(FORWARD pB\)](#) calculation as the third population for each model slice, generated as previously described in section 2.1. We finally use the corresponding [COR-1](#) white-light observations as the fourth population.

These populations are selected in order to provide a varying degree of model complexity when compare the expected [POS  \$\vec{B}\$](#) . The [MAS  \$n\_e\$  POS](#) provides a base one to one comparison with the central [POS  \$\vec{B}\$](#) . The [MAS  \$n\_e\$  LOS](#) adds a degree of complexity by integrating the electron density along the [line-of-sight](#) of the observer. The [FORWARD pB](#) is a true observable and introduces the geometric scattering efficiency along the [LOS](#) into the integration. Finally, the

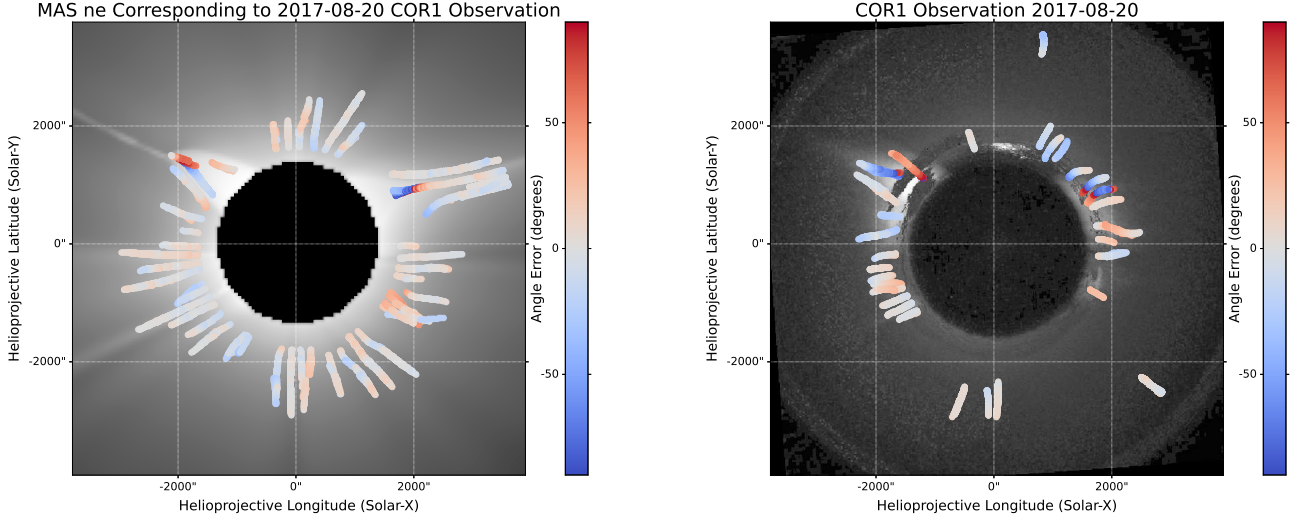
<sup>1</sup> The process for calculating a 95% confidence interval is described in appendix A.

<sup>2</sup> The process for calculating a probability density from a Gaussian kernel density estimate is described in appendix B.

<sup>3</sup> This statistic is described in appendix D.

<sup>4</sup> This statistic is described in appendix E.

<sup>5</sup> This interval is defined explicitly in appendix A



**Figure 7.** (Left) *MAS*  $n_e$  POS Model slice for 2017-08-20 with resulting *QRaFT* angular error plotted by position of pixel. (Right) *COR-1* observation for 2017-08-20 with resulting *QRaFT* angular error plotted by position of pixel. Potential sources of error are discussed in section 4.3.

white light *COR-1* observations provide the highest degree of complexity as the observational data. The consequences of the differences of resulting statistics between each of these populations will be discussed further in section 4.3. It should be noted that these potential consequences lie somewhat outside of the scope of this paper, which is defining and showcasing the actual methodology of this framework, (figure 1). Nonetheless these consequences hint at potential areas of further study utilizing this framework, which is why we choose to include them in our discussion.

### 3. RESULTS

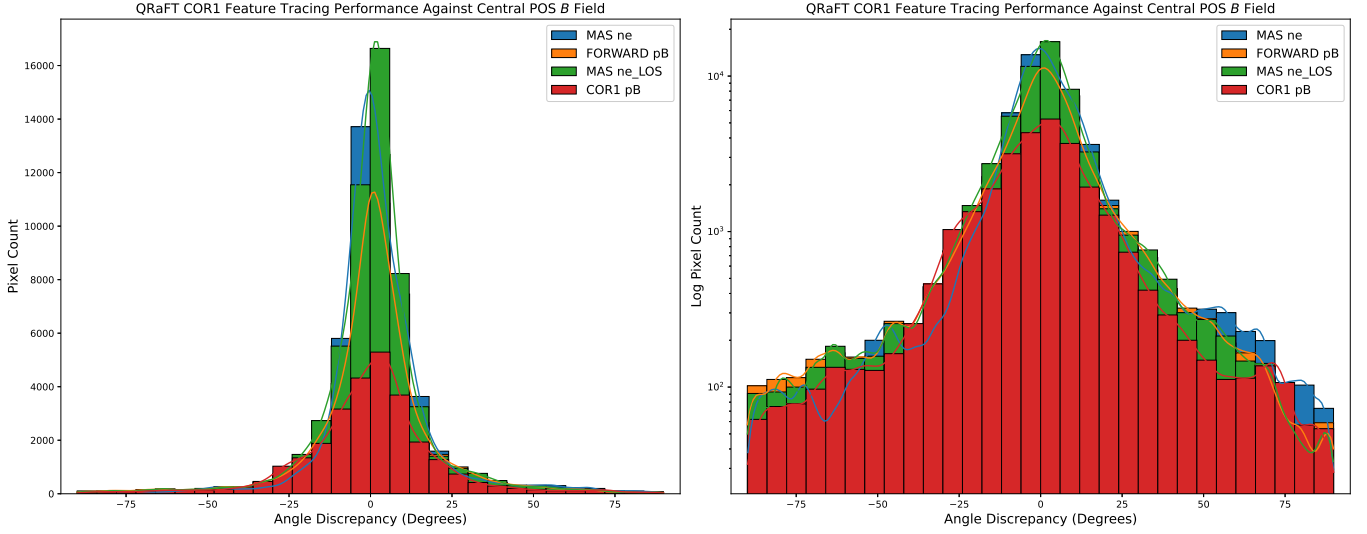
#### 3.1. Average and Median Discrepancies

Figure 7 shows an example of the angular discrepancies, as defined in section 2.6, of the *QRaFT* feature tracing method on an example slice of the *MAS* central plane-of-sky electron density and *COR-1* polarized brightness compared to the orientation of the central POS magnetic field from the *MAS* model solution. In this figure the discrepancies are mapped to their associated pixel in each feature and their magnitudes are shown using a color bar. Areas of high discrepancy are highlighted in red and blue, depending on the sign of the error.

These angular discrepancies are aggregated together across all observations and combined to produce statistics for each data population. Figure 8 shows histograms of this aggregated data divided into each respective data population defined in section 2.8. Also present in this figure are probability densities corresponding to each histogram generated using a Gaussian kernel density estimation<sup>6</sup>. The left plot show these histograms in a linear scale and the right plot show them in a log scale. The behavior of the tails of each of these histograms can be easily examined in the log scale, illuminating which data populations' tails drop off and flatten out relative to each other. It is noticeable in the left plot of figure 8 that the *COR-1* histogram has a much lower amplitude than the other three histograms, which are all based on data generated from the model. We believe this to be due to *QRaFT* detecting more features, and these features having more pixels, in the model datasets than in the *COR-1* observations.

Table 1 shows the average and median values of these angle discrepancies for each data population, as well as the standard deviation, total number of pixels, and 95% confidence interval range, organized by date of observation and

<sup>6</sup> This process is described in appendix B



**Figure 8.** (Left) linear and (Right) log-scaled histograms (with probability densities) of the combined angle discrepancies of the QRaFT filtered features on the COR-1 observations and model datasets including MAS  $n_e$  POS, MAS  $n_e$  LOS, and FORWARD pB when compared to the expected magnetic orientation from the POS  $\vec{B}$  from the PSI MAS model. The probability distributions are generated using a kernel density approximation as described in appendix B.

corresponding model datasets. In this table we also calculate the kurtosis and skewness<sup>7</sup>, of the resulting Gaussian kernel density estimate for each resulting histogram. Table 2 shows these same statistic for the combined datasets aggregated across all dates.

Figure 9 showcases box plot statistics from this aggregated data. This figure shows that when aggregated across the six observations, the MAS  $n_e$  POS population produces the lowest median discrepancy, Q1, IQR, Q3, and  $Q3+1.5*IQR$ . Each of these statistics increase when comparing each data population, with MAS  $n_e$  LOS having slightly higher statistics than MAS  $n_e$  POS, FORWARD pB having moderately higher statistics than MAS  $n_e$  LOS, and COR-1 pB having the highest statistics of the four populations. These statistics can also be seen in table 2.

We found that when compared to the central POS magnetic field of the PSI model, the average angle discrepancy of features traced in the COR-1 polarized brightness was  $14.554^\circ$ , and the average angle discrepancy of features traced in the synthetic FORWARD polarized brightness was  $12.119^\circ$ . We also found that when compared to the central POS magnetic field of the PSI model, the average discrepancy of features traced in the MAS line-of-sight integrated electron density was  $10.504^\circ$ , and the average discrepancy of features traced in the MAS central plane-of-sky electron density was  $10.479^\circ$ . The medians of each of these populations were  $9.157^\circ$  for the COR-1 dataset,  $6.883^\circ$  for the FORWARD pB dataset,  $6.035^\circ$  for the MAS  $n_e$  LOS dataset, and  $5.996^\circ$  for the MAS  $n_e$  POS dataset.

This difference between the median and average discrepancies suggests that the dataset is skewed towards lower values of angular discrepancies in each analysis, where outlying pixels that have large angular discrepancies increases the average discrepancy when compared to the median discrepancy. This shows that the QRaFT algorithm generally segmented features accurate to the magnetic field geometry within a reasonable degree of precision in each of these data populations.

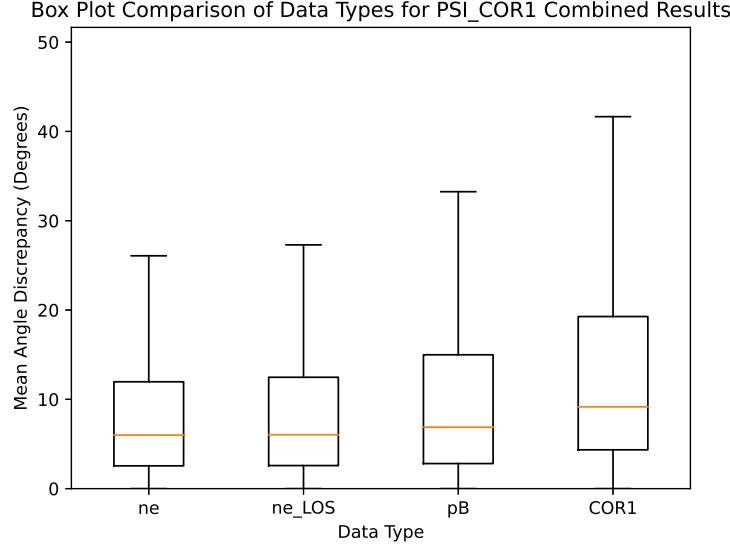
### 3.2. JSD Statistical Analysis

The Jensen-Shannon Divergence<sup>8</sup> statistic allows us to analyze the similarity of two probability distributions' shapes, which is useful for comparing the similarity of each of the data populations described in section 2.8. A probability distribution is calculated from each histogram using the method described previously in section 2.7. In figure 8, we can see from these probability distributions that there is close correlation between the shapes of the distributions of angle discrepancies of features traced in the COR-1 pB, FORWARD pB, MAS  $n_e$  LOS, and MAS  $n_e$  POS populations produced by the model when compared to the model's magnetic field orientation.

<sup>7</sup> These statistics are defined in appendix C

<sup>8</sup> This statistic, along with the Kullback-Leibler Divergence, is described in appendix D.





**Figure 9.** Boxplot of aggregated absolute angle difference for each data population described in section 2.8. Outliers above the  $Q3+1.5IQR$  range are not shown in the plot.

Since there is no ground truth precedent in which to compare the orientation of segmented coronal features to, as mentioned in section 1.1, we treat the MAS model solution as our known solution in which to base our comparison on. In order to diagnose the statistical trends between the model and data, we require an appropriate comparison framework and statistical metrics in which to quantify this correlation. We use the the model synthetic data compared to its own magnetic field orientation as the base comparison, so that the similarity between these statistics will reflect on the correlation between the data and model in regards to the statistical trend of how well quasi-radial features match the expected magnetic field orientation when segmented from coronal data vs model data.

The JSD, like the Kullback-Leibler Divergence (KLD), is defined between 0 and 1, and compares two distributions, where a metric closer to 0 indicates the two distributions are similar in shape, and a metric closer to 1 indicates the two distributions are anti-similar in shape. For transparency we offer both metrics in our results, however we consider JSD as the more important metric in evaluating the similarity of probability distributions as it is a symmetric statistic.

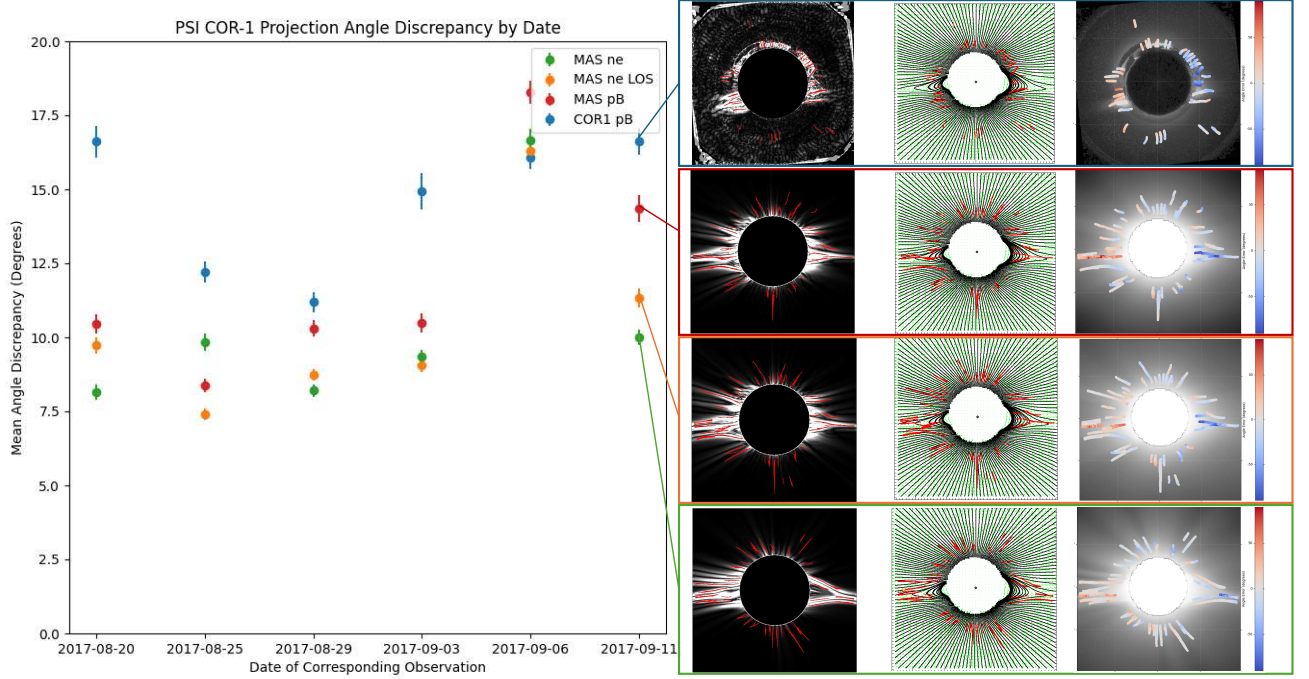
Table 4 shows the results of the JSD analysis when comparing the probability distributions of the QRaFT performance results on the aggregated data populations described in section 2.8 in relation to the orientation of the expected model magnetic orientation given by the central POS  $\vec{B}$ . The JSD and KLD statistics are combined with the HSD statistics in tables 4 and 3 for convenience, as both sets of statistics are means of the various data populations to each other.

Finally, we compare this various data populations' results from the analyses done in tables 2 and 4 on a random dataset as a control analysis. This dataset is the size of the mean of means from the four respective datasets described in section 2.8 and is randomized in a range from  $-90^\circ$  to  $90^\circ$ .

## 4. DISCUSSION

### 4.1. Overview

In this work we presented a quantitative analysis comparing white light pB observations to synthetic pB images, along with other model outputs, derived from the output of an advanced MHD model used for predicting the orientation of the solar corona during solar eclipses. Our preliminary results of this framework show close correlation between the performance of quasi-radial coronal feature tracing on both the synthetic and white-light images when compared to the expected model  $\vec{B}$ . These results are supported by the global analysis conducted in table 4, which shows a JSD metric of 0.007 between the COR-1 pB and FORWARD pB probability distributions when compared to the central POS  $\vec{B}$  and 0.007 between the MAS  $n_e$  POS and FORWARD pB probability distributions when compared to the central POS  $\vec{B}$ . These metrics signify a close similarity between both probability distributions in each case. For comparison, the



**Figure 10.** Resulting mean angle discrepancy by date with examples of each data type shown for 2017-09-11 COR-1 model projection with COR-1 observation. Error bars represent 95% confidence intervals as calculated in section A. These statistics, analyzed by date, can be found in table 1. In each set of three example plots for each data point, the left plot shows the QRaFT segmented features plotted onto the enhanced image. The middle figure is the QRaFT segmented features plotted onto the MAS POS  $\vec{B}$  orientation as shown in figure 5. The right plot is the QRaFT angular error plotted onto the original source image, as shown in figure 7.

analysis between the random probability distribution vs the FORWARD pB probability distribution results in a JSD metric of 0.231, the analysis between the MAS  $n_e$  POS and random probability distributions results in a JSD metric of 0.268, and the analysis between the COR-1 pB and random probability distributions results in a JSD metric of 0.215. These metrics are clearly higher in this case, signifying that these distributions do not show good correlation, which is what we expected the metric to show in this control case.

This analysis also showed that the segmentations by QRaFT was able to capture the global magnetic orientation between  $10.479^\circ \pm 0.118^\circ$  and  $14.554^\circ \pm 0.184^\circ$  on average for the datasets described in section 2.8 when compared to the expected magnetic orientation given by the MHD model. We can see these metrics, as well as other central tendency indicators in table 2. Comparing these results to the random population's mean discrepancy of  $45.246^\circ \pm 0.239^\circ$ , we see that the mean discrepancies of the populations chosen in section 2.8 are minimal. While these results are low in comparison to the random population, the resulting  $10.479^\circ$  average discrepancy for the MAS  $n_e$  POS population does suggest that there is baseline error in the assumptions and methodology of QRaFT. This will be further discussed in section 4.3.4, but it is worth noting as this metric is able to give us an estimate on how reliable white-light image segmentation methods are in identifying magnetic structure in the corona.

These results are significant, as the synthetic pB images are generated using MAS, which establishes a ground truth comparison between these FORWARD modeled pB images and the underlying 3D electron density and magnetic field distribution from the same calculation, allowing us to test the accuracy of observational methods that approximate the orientation of the coronal magnetic field, such as coronal segmentations. Comparing the orientation of plasma density features extrapolated by QRaFT to the expected magnetic orientation from the model establishes a quantifiable measurement of the agreement between each. Furthermore, using QRaFT to perform this comparison between the model populations described in section 2.8 vs the model's expected magnetic field orientation and then between the COR-1 polarized brightness vs the same expected magnetic field orientation also provides a quantifiable measurement of

**Table 1.** Central Tendency Statistics for each population described in section 2.8. Mean, median, standard deviation, and 95% CI values represent the absolute values in order to display magnitudes. The kurtosis and skewness are calculated on the probability distribution calculated from the signed angle histograms for each population. The table is sorted in ascending order of mean discrepancies and organized by the date of corresponding COR-1 observation.

data type	date	mean	median	std	95% CI	n	kurtosis	skewness
MAS $n_e$ LOS	2017-08-25	7.408	4.477	8.949	0.184	9059	8.127	-0.003
MAS $n_e$ POS	2017-08-20	8.150	4.182	12.429	0.251	9444	12.444	-0.342
MAS $n_e$ POS	2017-08-29	8.202	5.219	10.214	0.204	9584	8.884	1.028
FORWARD pB	2017-08-25	8.368	4.656	9.997	0.227	7473	4.945	-0.103
MAS $n_e$ LOS	2017-08-29	8.736	5.294	11.454	0.191	13878	9.642	-0.157
MAS $n_e$ LOS	2017-09-03	9.062	5.802	9.724	0.236	6512	3.402	0.592
MAS $n_e$ POS	2017-09-03	9.344	6.063	10.480	0.249	6793	6.902	-0.722
MAS $n_e$ LOS	2017-08-20	9.724	5.234	12.902	0.265	9077	6.654	0.505
MAS $n_e$ POS	2017-08-25	9.848	5.151	14.176	0.297	8739	8.983	0.404
MAS $n_e$ POS	2017-09-11	10.008	7.179	11.774	0.252	8409	8.996	0.510
FORWARD pB	2017-08-29	10.302	5.601	14.190	0.278	9999	7.222	-0.589
FORWARD pB	2017-08-20	10.448	5.787	13.403	0.318	6843	5.740	0.721
FORWARD pB	2017-09-03	10.478	6.812	11.499	0.319	5001	3.365	0.503
COR-1 pB	2017-08-29	11.193	7.858	12.718	0.334	5568	6.179	1.712
MAS $n_e$ LOS	2017-09-11	11.333	6.577	14.279	0.324	7481	5.293	-0.657
COR-1 pB	2017-08-25	12.210	8.031	12.360	0.363	4455	4.446	0.117
FORWARD pB	2017-09-11	14.350	9.227	16.636	0.456	5113	4.082	-0.461
COR-1 pB	2017-09-03	14.925	8.984	17.080	0.613	2981	3.855	1.692
COR-1 pB	2017-09-06	16.075	12.273	14.820	0.409	5047	2.141	-0.328
MAS $n_e$ LOS	2017-09-06	16.309	10.018	17.747	0.334	10821	3.165	-0.578
COR-1 pB	2017-08-20	16.615	8.021	19.529	0.542	4990	2.492	-0.079
COR-1 pB	2017-09-11	16.633	11.681	16.153	0.454	4856	1.905	-1.312
MAS $n_e$ POS	2017-09-06	16.662	9.071	18.900	0.373	9864	2.231	0.063
FORWARD pB	2017-09-06	18.281	11.925	18.689	0.390	8836	2.548	-0.671

**Table 2.** Combined central tendency statistics for each population described in section 2.8. Mean, median, standard deviation, and 95% CI values represent the absolute values in order to display magnitudes. The kurtosis and skewness are calculated on the probability distribution calculated from the signed angle histograms for each population. These populations are compared to a random sample population as a control analysis. The table is sorted in ascending order of mean discrepancies.

data type	date	mean	median	std	95% CI	n	kurtosis	skewness
MAS $n_e$ POS	combined	10.479	5.996	13.855	0.118	52833	7.108	0.270
MAS $n_e$ LOS	combined	10.504	6.035	13.348	0.110	56828	6.547	-0.256
FORWARD pB	combined	12.119	6.883	14.961	0.141	43265	5.034	-0.295
COR-1 pB	combined	14.554	9.157	15.694	0.184	27897	3.377	-0.003
random	combined	45.246	45.421	25.929	0.239	45205	-1.196	0.005

the similarity between the data and model. Having the QRaFT algorithm perform similarly on the MAS central plane-of-sky electron density, MAS line-of-sight integrated electron density, FORWARD polarized brightness, and COR-1 polarized brightness when compared to the expected orientation from the model suggests that the features segmented by QRaFT represents coronal magnetic field structure within a reasonable error, and establishes a benchmark to measure the agreement between data and model.

#### 4.2. *MAS MHD Model as a Ground-Truth Comparison*

The use of an **MHD** model, in particular **MAS**, in this statistical comparison is of special note to discuss. As mentioned in section 2.4, **MAS** includes additional energy terms that describe the energy flow in the corona and solar wind. This is necessary in order to accurately describe the interaction between the plasma and magnetic properties of the corona. Since we are particularly interested in the orientation of large scale quasi-radial plasma density features present in the solar corona as viewed in the **POS**, this makes **MAS** an ideal source of comparison for our evaluation. Using this model as a known solution allows us to compare its expected magnetic orientation to the orientation of the features segmented by **QRaFT** in order to evaluate how well they agree with these modeled magnetic structures.

Since we are using **COR-1 polarized brightness** observations to investigate segmented plasma density features in the K-Corona, the ability to model and retain the electron density is essential in reconstructing the polarization brightness of the K-Corona. **MAS**'s ability to output the electron density and magnetic field in 3D allowed for the rotation and projection of the electron density and magnetic field onto the **POS**, as described in section 2.1. The ability to work directly with the electron density when integrating along the line of sight allowed us to compare this data population to synthetic polarized brightness using the **FORWARD** code.

We purposefully chose observations that corresponded closely in time with the 2017 solar eclipse, which this particular solution was modeled to predict. While there is a similar qualitative analysis done by **PSI** to evaluate their model against eclipse observations, we target large-scale quasi-radial features as opposed to closed structures such as coronal loops. These particular features are generally difficult to detect and trace, since the optical signatures in these features are extremely weak when compared to closed features (Antonucci et al. 2020). We also apply a quantitative analysis utilizing **QRaFT** in order to verify that these features represent true magnetic structure. We use our angular discrepancy analysis described in section 2.6 to generate two types of populations, one set comparing the features of model output to itself, and the other comparing the features in the **COR-1 pB** observations to the model. This allowed us to determine the average and median angular discrepancies for each population described in section 2.8, effectively telling us the performance of **QRaFT** in each case. We then calculated probability densities using a Gaussian kernel density estimation for each population to analyze the analytical shape of each of these distributions. We can see in figure 8 that this process allows us to examine and compare the shapes of these distributions to each other, effectively comparing the distribution of discrepancies in the model comparisons to the distribution of discrepancies in the data comparison. Since each case compares the orientation of features segmented by **QRaFT** to the expected magnetic field predicted from the model, this is a direct comparison of segmentations from white-light data against this ground truth. In tables 2 and 4 we included a population mocking a randomly oriented magnetic field as a control analysis.

It is possible to qualitatively compare the shapes of these distributions, which is effectively a qualitative analysis on the similarity between the global orientation of the modeled corona when compared to the observed corona, which is similar to how **PSI** evaluated their performance in modeling broad aspects of corona in their solar eclipse comparison with eclipse images (Mikić et al. 2018). We take the additional step of analytically comparing the shapes of these probability densities to provide a quantifiable metric for how well the model represents the data. Using the **JSD**, kurtosis, and skew, we are able to provide metrics for how similar these probability densities are in shape, which gives a measurement of how well they represent each other. Combining the statistics of the average and median discrepancies of **QRaFT** as well as the kurtosis, skew, and **JSD** of the probability densities, we can describe how accurately **QRaFT** segments these features when compared to the expected result while evaluating how the distributions given by how much the probability densities represent each other. Minimizing these metrics would then in theory suggest improving the segmentation of quasi-radial features in white light **pB** observations and potentially also improve **MHD** modeling as well.

#### 4.3. *Sources of Error*

Discussing the results in tables 1 and 2, we see that the average and median discrepancies of plasma density features traced in both the synthetic **pB** images and **COR-1 pB** observations when compared to the **POS  $\vec{B}$**  orientation suggest that their orientation seem to correlate reasonably well with this expected orientation. The metrics given by the features in the synthetic **pB** images are also similar to the metrics given by the features segmented in the white-light observations. We should note that when comparing the model to itself, the performance metrics are lower as indicated in tables 1 and 2, which is to be expected. Both sets of metrics show a moderate average disagreement from the expected orientation in both the features segmented from the model populations and in the features segmented from



the [COR-1 polarized brightness](#). Due to the various models, methods, and sources of data used in this analysis, there is ample opportunity for error. We pinpoint four potential causes of error in this analysis.

#### 4.3.1. *Anomalies/Artifacts from White Light Observations*

For ground based observatories, natural imperfections relating to both atmospheric and weather effects may limit the image quality and even render certain times of observation unusable due to cloud coverage. For example, [Elmore et al. \(2003\)](#) notes that atmospheric effects such as sky transmission and sky polarization may introduce illegitimate artifacts to [K-Coronagraph Instrument at Mauna Loa Solar Observatory \(K-Cor\)](#) observations. They also note that sunlight reflected off the Earth and then by the atmosphere into the line of sight is difficult to remove from the signal of the corona ([Elmore et al. 2003](#)). Applying this framework to ground-based coronagraphs such as [K-Cor](#) would introduce the possibility of this type of error.

In the case of [COR-1](#), since this is a space-based observatory, it is not subject to the same atmospheric effects that plague ground based observatories. However even in these images artifacts from the detector can be noticed. Another potential source of error is that the [STEREO](#) spacecraft are removed from the Sun-Earth line in order to reliably detect Earth-directed CMEs ([Thompson et al. 2003](#)). While this in theory should not contribute at all to the error since the [MAS](#) model provides a 3D solution, and is thus able to be rotated to any perspective, [MAS](#) may be optimized for the viewpoint of the Sun-Earth line.

Errors in the image processing of the [COR-1 pB](#) representative observations may have also contributed to error. As discussed in section 2.2, several layers of image processing are applied to create the representative [pB](#) observations used in this study. One of these layers is image background subtraction. This is of note as [Thompson et al. \(2010\)](#) compared [COR-1 pB](#) images to [Mk4 pB](#) images to test the effectiveness of background removal, and found the overall behavior to be the same. This is noteworthy as it suggests that the background removal process for each observatory should cause similar error, if any, to the accuracy of the orientation of the quasi-radial features segmented in this study. Analysis of ground-based observatories such as [K-Cor](#) is a planned next step to further validate and analyze this framework. The similar performance of [QRaFT](#) on both the [COR-1](#) and [K-Cor](#) datasets when compared to the expected magnetic orientation would be a significant step in validating this framework between both types of observatories.

Another caveat not relating to atmospheric effects, but rather conditions on the sun is the timing of the observations chosen for this analysis. During our analysis, we focus on periods of quiet solar activity in order to preserve the large-scale magnetic structure given by the quasi-radial regions. This structure is disrupted if major space weather activity occurs, such as a coronal mass ejection (CME), since the dynamics of this event causes rapid changes to the orientation of the corona as plasma is ejected outward as a part of this process. There were time periods in which we had to select observations taken at a slightly earlier/later time than predicted using the method outlined in section 2.5. Several consequences may result from this decision. First, the images are not separated at exactly 60 degrees, but are skewed slightly more or less for each image. This is to be expected, as it is unreasonable to find every observation correlating perfectly to a 60 degree rotation of the Sun given that these observations are taken when available. The second and more consequential, is that the further the observations are taken away from the date of the solar eclipse, the less representative the corona is to the simulation. The practical effect this could have on our analysis is that the observations used at the furthest periods from the eclipse may be the least representative of the corona's overall structure in the [MHD](#) simulation, potentially skewing error towards these images. The aim of this decision was to provide a 360° view of the sun in our analysis in order to eliminate directional bias, which justifies the opportunity for potential error caused by this decision in our opinion.

#### 4.3.2. *Limitations of MHD Models*

Regardless of their level of sophistication, it is important to recognize the sources of error and inherent limitations of global coronal [MHD](#) models. In general, thermodynamic MHD models like [MAS](#) require inputs (e.g. magnetic boundary conditions) and parameterization choices (e.g. the coronal heating model), whose sources and parameterizations affect the ensuing 3D distributions of plasma properties and the vector magnetic field.

First, as mentioned in section 2.4, the primary boundary condition used in [MAS](#) is a full-sun map of the radial component of the surface magnetic field,  $B_r$ . This magnetic boundary condition is the most important input to the model, as it largely determines the overall 3D structure and morphology of the coronal magnetic field. Because measurements of the surface magnetic field are typically only available from observatories positioned along the Sun-Earth line (as is the case here), a full-sun map of  $B_r$  must be constructed from series of magnetograph observations in time, covering at least one solar rotation (e.g. a synoptic map) or produced by a data assimilative model that also

describes the evolution of the field in unobserved region (e.g. synchronic maps produced by surface flux-transport models). This implies a catch 22 for this framework and data validation as a whole, as the MHD models must inherently rely on magnetic boundary conditions that are certain to have limitations in both temporal and spatial accuracy. As mentioned in section 2.4, this particular model used a synoptic map from SDO/HMI as the input  $B_r$  at the inner boundary to predict the conditions of the corona on August 21, 2017. Since this map was generated using data from CRs 2192 and 2193, and our observational comparisons use data from CR 2194, this may be one significant source of error. This error is challenging to quantify however due to both the observing limitations and the inherent evolution of the Sun’s surface magnetic field, which together imply that there is no single map or MHD model that can be constructed that would simultaneously describe each of our observation dates. That said, the pipeline and error metrics discussed here could in-principle be used to optimize the magnetic boundary conditions for a given case (e.g.; Jones et al. 2016, 2020).

Another important element of MHD simulations is the combination of the hydrodynamic boundary conditions and coronal heating model that together determine the plasma state of the model (density, temperature, velocity). The heating model in particular is an essential piece of determining the appearance of coronal observables (Lionello et al. 2009), and there is an intrinsic feedback between the heating, plasma, and the 3D coronal magnetic field that determine where the solar wind is able to open the magnetic field and the morphological appearance of coronal structures (streamers, loops, coronal holes, etc., Downs et al. 2010). In other words, this choice may influence the misalignment angle and error metrics discussed here. In our case, the MHD model that we employ here utilizes the WTD approach to heat the corona. This approach uses the MHD equations along with auxiliary equations to describe the subgrid propagation, reflection and dissipation of Alfvén waves to determine a heating term that reasonably captures coronal heating for a broad range of morphologies (coronal holes, quiet-sun, active regions; Lionello et al. 2014; Downs et al. 2016; Mikić et al. 2018). Even so, such a model must be parameterized and may not fully capture all of the details of heating, especially in complex regions such as null points or streamer cusps. That said, unlike in the case of the magnetic boundary conditions, there is hope that our techniques could be used for heating model parameter optimization in the future, especially if done in tandem with other observational constraints (e.g. EUV and white-light intensities, in-situ measurements, etc.), and we hope to explore this in the future.

Lastly, another potentially relevant element of coronal modeling is the extent to which the presence of large-scale coronal current systems in the form of sheared or twisted filament channels (i.e. free magnetic energy) might affect the observed streamer morphology. Coronal potential field models do not have coronal currents by construction, so this is typically not a concern that is addressed. Similarly, the majority of thermodynamic global coronal MHD models are run as line-tied steady-state relaxations, where the solar wind and heliospheric current sheet(s) are formed but the low coronal domain remains potential to a large extent. There have been recent attempts to use vector magnetic field information and/or magnetofrictional methods to drive global coronal MHD models to construct coronal current systems that match observations (e.g. Hayashi et al. 2021, 2022), but between the quality limitations of full-sun vector magnetic field observations and the various pitfalls and inconsistencies of boundary driving methods (see Tarr et al. 2024), this an extremely challenging and fraught endeavor. On the other hand, as illustrated in Mikić et al. (2018), which used an ad-hoc method for building shear and magnetic flux-ropes along large-scale polarity inversion lines to construct the energized model under study here, large-scale shear and flux-ropes can influence the morphologies of streamers, particularly at their base. The extent to which the QRaFT pipeline is sensitive to such forms of energization is another interesting avenue to explore going forward, but again illustrates the challenge of the constraining the large-parameter space involved in coronal modeling, which spans choices made in model parameterizations, physical assumptions, and numerical implementations.

#### 4.3.3. Central POS Assumption

Another potential source of error is the assumption that the central POS magnetic field is completely responsible for the orientation of the large quasi-radial plasma density features that we observe in the corona. As noted in section 1.3, the features traced by QRaFT are assumed to reflect the projected coronal magnetic field in the POS. This assumption is justified by several factors. Jones et al. (2020) notes that using pB images place a much greater emphasis on coronal material near the plane of sky since the photospheric light scattered from electrons in the image plane is more polarized than the light scattered from electrons away from the plane. The consequence of this is that all coronal features appear within image plane, even those located outside of the POS (Jones et al. 2020). There are also practical limits justifying this assumption. In order to perfectly compare the 3-D magnetic solution provided by the MHD model, one would

need to know the topology of the plasma density from **pB** observations taken from Earth’s perspective. We cannot simultaneously probe the Sun for changes in its electron density in order to model the topology of both the magnetic field and the plasma density structures at different solar radii. We can therefore only measure the orientation of plasma density features in the plane of sky when analyzing **pB** observations with coronal image segmentation methods such as **QRaFT**.

While there are justifications for this assumption, it also creates ample room for error in this analysis. This assumption assumes that the central **POS** magnetic field is responsible for the orientation of the large quasi-radial plasma density features that we observe in the corona. Given that the corona is a three-dimensional dynamic structure that does not lie in a single plane and also varies along the **line-of-sight**, it is easy to see how error can be introduced from this assumption. If there are contributions to the **pB** signal from polarized structures in the corona that lie outside of the **POS**, this would introduce error into the comparison with the central **POS**  $\vec{B}$ . An example of such a signature could be the scattering of free coronal electrons that occur outside of the **POS**. The **MHD** solution of **MAS** is a 3D field, modeling the physical reality of the system. This is an assumption that can naturally lead to errors in this analysis, and would be worth exploring in future work. This could be done by using the **LOS** or emissivity weighted **LOS** integrals of  $\vec{B}$  components to produce a component in which to compare too.

It should be noted from table 4 that the difference between the **MAS**  $n_e$  **POS** and **MAS**  $n_e$  **LOS** populations was only  $0.025^\circ$ . This difference was not enough to reject the **null hypothesis** ( $H_0$ ) according to the **HSD** analysis. This suggests that the error produced from the geometric projection of the 3-D volume of the **MAS line-of-sight integrated electron density** onto the **POS** was not enough to produce significant error to this analysis. The similarity between the **pB** and **MAS**  $n_e$  **LOS** results suggests that where **QRaFT** is identifying features, the coronal material tends to be near the **plane-of-sky**. If further investigation using this framework suggests that the contributions of electrons away from the **POS** doesn’t produce significant error when comparing traced features to the **POS**  $\vec{B}$ , this would be a significant step in validating the earlier mentioned assumption. It should be cautioned that in order to truly validate this assumption, this result would need to be repeated under numerous conditions and data sources. While this method provides a means in which to do this, it should not be concluded that this lone result constitutes validation of this assumption, and more work is required to definitively come to this conclusion.

From a theoretical standpoint, one may also make the argument that the **pB** scattering should better localize structures to the **POS**, as mentioned in Jones et al. (2020), and thus should perform better than the **MAS**  $n_e$  **LOS** results. Further studies may be implemented to test these theories. For example, a cross correlation analysis between the **MAS**  $n_e$  **LOS** and **pB** signal would evaluate the similarity of the 2D arrays for each population, providing an internal check into the validity of this specific result.

It should also be noted from table 4 that there is a mean difference of  $1.616^\circ$  between the **MAS**  $n_e$  **LOS** and **FORWARD** **pB** populations in the **HSD** analysis. This difference is large enough to reject the **null hypothesis** ( $H_0$ ) and show that there is a statistically significant difference between these populations. This shows that when comparing central **POS**  $\vec{B}$  orientation to the orientation of the **MAS**  $n_e$  **LOS** and then to the **FORWARD** **pB**, there exists a statistically significant difference, and thus the attributes between these populations contributes enough error to be statistically significant. This suggests that the artificial scattering alone, along with other methods that do not involve integrating the  $n_e$  along the **LOS** in **FORWARD**, contributed enough error in this analysis to be statistically significant. Correlating this error to the error between the **MAS**  $n_e$  **POS** and **FORWARD** **pB** populations in table 4, we see that there is a mean difference of  $1.641^\circ$  between the **MAS**  $n_e$  **POS** and **FORWARD** **pB** populations. This difference is also large enough to reject the **null hypothesis** ( $H_0$ ) and show that there is a statistically significant difference between these populations, but is also roughly equal to the sum of the two previously mentioned errors. We believe this to be an example of this framework dissecting and quantifying the various contributions of error in this analysis, as this result suggests that the combined factors of the geometric projection along with the scattering of free coronal electrons contributed enough error to the analysis to be statistically significant, but that the geometric projection of the 3-D volume of the **MAS line-of-sight integrated electron density** onto the **POS** alone did not contribute enough error to the analysis to be statistically significant. As mentioned previously, if this framework is found to consistently show this result under further analysis, this would be a significant step in validating the earlier mentioned assumption and would provide a quantitative baseline in which to describe and account for this error.

Finally, it is worth noting from table 4 that the shape of the probability densities of the **MAS**  $n_e$  **POS** results and the **FORWARD** **pB** are highly similar as indicated by the **JSD** statistic of  $0.007^\circ$ . The shape of the probability densities of the **MAS**  $n_e$  **POS** results and the **MAS**  $n_e$  **LOS** results are even more highly similar as indicated by the **JSD**

statistic of  $0.006^\circ$  degrees, which is what we would have expected given pattern of the previously mentioned results when comparing these populations' mean angular differences. We believe that these quantitative results further add to the justification that the assumption that the central POS magnetic field is deterministic of the orientation of the corona is warranted. The possibility of quantitative validation of this assumption from this framework is a potentially significant milestone as several models and analyses of the solar corona rely on this assumption, but it has yet to be quantitatively justified.

#### 4.3.4. Errors Borne from QRaFT

It is possible, and in some cases likely, that QRaFT is not perfectly approximating the orientation of plasma density features in both the COR-1 pB observations and in the simulated FORWARD pB images. An example of this is QRaFT's 'blob filtering' method not being able to filter out every spurious feature detected by the algorithm. These features are normally small, and when filtered by the radial length, are able to be filtered out reasonably well.

One of the most prominent sources of error stems from QRaFT tracing radial features, (or what it perceives to be open field features), in closed field regions of the magnetic field in the POS. These regions can be especially noticed when plotting the features from QRaFT error over the expected magnetic field, such as in figure 5, and when plotting angular error over the features traced from QRaFT, such as in figure 7. These global plots of the angular error allow us to dissect the pattern of how QRaFT's features match the model magnetic field. In these plots, the regions with the most error correlate to the closed regions of the model magnetic field, as the quasi-radial features cross this field near perpendicularly in these regions. These errors have a high impact on the resulting statistics of the analysis, as seen from the discrepancy of the global means and medians in tables 1 and 2.

As mentioned in section 3, the average discrepancy of the MAS  $n_e$  POS population was  $10.479^\circ$ . As this population is a 1-1 comparison of the model POS  $\vec{B}$  solution to its corresponding MAS  $n_e$  POS, this figure is a reasonable estimate of the systematic error of QRaFT. It is noticeable from figure 10 that QRaFT segments features that do not correspond with true structure in the corona, particularly when comparing segmentations in COR-1 observations to segmentations in the MAS central plane-of-sky electron density, as shown in figure 5. In this comparison, there are clearly spurious segmentations in the COR-1 dataset that do not represent magnetic structure. In particular, features were seen to be originating from locations well beyond the solar occulting radius in several observations. While these features were able to be filtered out manually in this analysis, this still suggests that QRaFT has room for improvement in feature detection and filtration. Figure 2 shows steps of the detrending and enhancement method described in section 2.3, and in particular panel (d) of this figure shows the fine quasi-radial structure revealed from the second order detrending. There do appear to be artifacts in this enhancement that may be mistaken for features by the algorithm that are not automatically filtered when selecting the final features. Since QRaFT is in active development, the conclusion that it can be improved and from this further reduce the baseline error is reasonable.

#### 4.4. Utilizing Framework to Improve Error

Through a rigorous set of statistical tests, we believe that we were able to numerically characterize the error contributions from these assumed sources. We carefully chose four sets of populations from the data in which to measure this potential error, those being the MAS central plane-of-sky electron density (MAS  $n_e$  POS), MAS line-of-sight integrated electron density (MAS  $n_e$  LOS), FORWARD polarized brightness (FORWARD pB), and the white light COR-1 polarized brightness (COR-1 pB).

The difference between the means calculated in Tukey's Honestly Significant Difference Test<sup>9</sup> for these first two populations of the MAS  $n_e$  POS and the MAS  $n_e$  LOS should provide the error contribution from the geometric projection of the 3-D electron density volume onto the POS. Likewise the difference between the means of the next two populations given by the MAS  $n_e$  POS and model calculated FORWARD pB should provide the error contribution from the artificial scattering and other methods used in FORWARD. Finally the difference between the last two populations given by the model calculated FORWARD pB and white light COR-1 pB should provide all other error contributions associated from the comparison of the modeled POS  $\vec{B}$  to the segmented features from the observed polarized brightness. Table 3 shows the results of this error analysis for each population comparison across each date of observation chosen while table 4 shows the results of this error analysis for each population for the combined data.

<sup>9</sup> This test is described in appendix E.



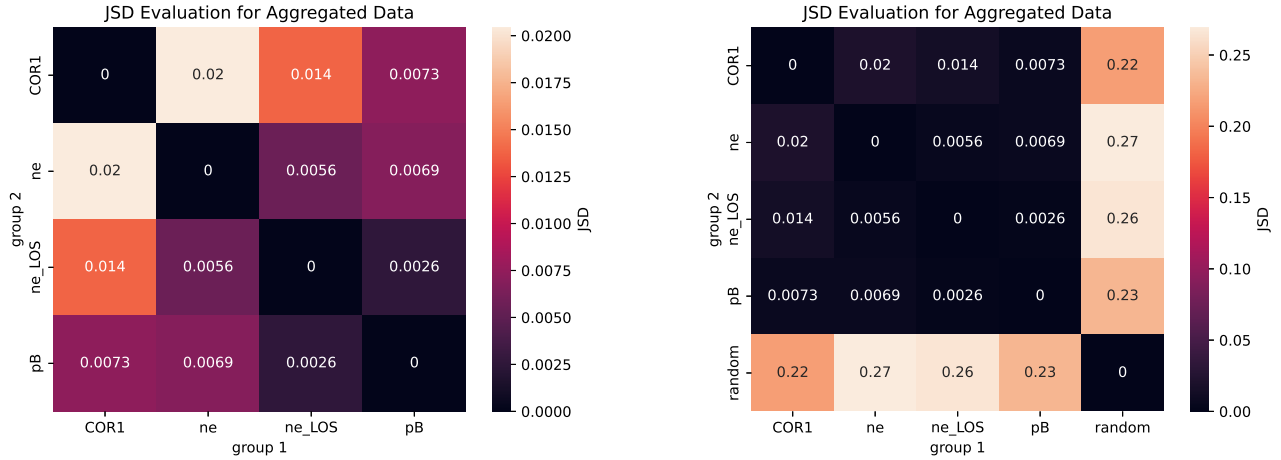
**Table 3.** Model type comparison using **HSD** and **JSD** statistics for **COR-1** dataset by date. Methods are described in appendices E and D, respectively. In the (reject  $H_0$ ?) column, a value of 1 indicates that the null hypothesis ( $H_0$ ) is rejected and a value of 0 indicates that  $H_0$  is not rejected.

group 1	group 2	date	mean diff	lower CI bound	upper CI bound	KLD	JSD	reject $H_0$ ?
COR-1 pB	MAS $n_e$ POS	2017-08-20	-8.465	-9.102	-7.827	0.179	0.048	1
COR-1 pB	MAS $n_e$ LOS	2017-08-20	-6.891	-7.533	-6.249	0.117	0.032	1
COR-1 pB	FORWARD pB	2017-08-20	-6.167	-6.845	-5.489	0.092	0.026	1
MAS $n_e$ POS	MAS $n_e$ LOS	2017-08-20	1.574	1.038	2.109	0.035	0.009	1
MAS $n_e$ POS	FORWARD pB	2017-08-20	2.297	1.719	2.876	0.118	0.027	1
MAS $n_e$ LOS	FORWARD pB	2017-08-20	0.724	0.141	1.307	0.060	0.014	1
COR-1 pB	MAS $n_e$ POS	2017-08-25	-2.363	-2.905	-1.820	0.225	0.058	1
COR-1 pB	MAS $n_e$ LOS	2017-08-25	-4.802	-5.341	-4.263	0.110	0.030	1
COR-1 pB	FORWARD pB	2017-08-25	-3.842	-4.400	-3.284	0.082	0.022	1
MAS $n_e$ POS	MAS $n_e$ LOS	2017-08-25	-2.439	-2.881	-1.998	0.196	0.043	1
MAS $n_e$ POS	FORWARD pB	2017-08-25	-1.479	-1.943	-1.015	0.239	0.056	1
MAS $n_e$ LOS	FORWARD pB	2017-08-25	0.960	0.500	1.421	0.022	0.005	1
COR-1 pB	MAS $n_e$ POS	2017-08-29	-2.991	-3.516	-2.467	1.860	0.400	1
COR-1 pB	MAS $n_e$ LOS	2017-08-29	-2.457	-2.951	-1.963	1.826	0.365	1
COR-1 pB	FORWARD pB	2017-08-29	-0.891	-1.412	-0.370	1.719	0.348	1
MAS $n_e$ POS	MAS $n_e$ LOS	2017-08-29	0.535	0.121	0.948	0.049	0.013	1
MAS $n_e$ POS	FORWARD pB	2017-08-29	2.101	1.655	2.546	0.098	0.016	1
MAS $n_e$ LOS	FORWARD pB	2017-08-29	1.566	1.157	1.974	0.016	0.004	1
COR-1 pB	MAS $n_e$ POS	2017-09-03	-5.582	-6.240	-4.923	0.269	0.070	1
COR-1 pB	MAS $n_e$ LOS	2017-09-03	-5.863	-6.525	-5.200	0.352	0.084	1
COR-1 pB	FORWARD pB	2017-09-03	-4.447	-5.141	-3.754	0.217	0.054	1
MAS $n_e$ POS	MAS $n_e$ LOS	2017-09-03	-0.281	-0.801	0.238	0.113	0.029	0
MAS $n_e$ POS	FORWARD pB	2017-09-03	1.134	0.576	1.692	0.097	0.017	1
MAS $n_e$ LOS	FORWARD pB	2017-09-03	1.415	0.852	1.978	0.048	0.012	1
COR-1 pB	MAS $n_e$ POS	2017-09-06	0.587	-0.211	1.385	0.166	0.035	0
COR-1 pB	MAS $n_e$ LOS	2017-09-06	0.235	-0.551	1.020	0.106	0.021	0
COR-1 pB	FORWARD pB	2017-09-06	2.206	1.393	3.019	0.103	0.019	1
MAS $n_e$ POS	MAS $n_e$ LOS	2017-09-06	-0.353	-0.994	0.289	0.058	0.014	0
MAS $n_e$ POS	FORWARD pB	2017-09-06	1.619	0.944	2.294	0.075	0.018	1
MAS $n_e$ LOS	FORWARD pB	2017-09-06	1.972	1.310	2.632	0.017	0.004	1
COR-1 pB	MAS $n_e$ POS	2017-09-11	-6.625	-7.292	-5.957	0.290	0.070	1
COR-1 pB	MAS $n_e$ LOS	2017-09-11	-5.300	-5.982	-4.617	0.219	0.051	1
COR-1 pB	FORWARD pB	2017-09-11	-2.283	-3.025	-1.540	0.169	0.033	1
MAS $n_e$ POS	MAS $n_e$ LOS	2017-09-11	1.325	0.737	1.914	0.052	0.015	1
MAS $n_e$ POS	FORWARD pB	2017-09-11	4.342	3.686	4.999	0.129	0.022	1
MAS $n_e$ LOS	FORWARD pB	2017-09-11	3.017	2.345	3.689	0.040	0.010	1

While each potential source of error adds uncertainty to the accuracy of our analysis, the fact that there is moderate error provides the opportunity to improve this error through further use of this framework. In section 4.2, it was noted that one of the primary methods to determine the accuracy of **PSI**'s solar eclipse analysis was by qualitative comparison with eclipse images. This analysis provides potentially the first method to quantitatively analyze the accuracy of the orientation of quasi-radial coronal features segmented from white light imaging against features segmented from a ground-truth **MHD** model when compared to their predicted magnetic orientation using this solution. This method therefore provides a benchmark in which to improve the sources of error found in this analysis. Since we now have a

**Table 4.** Model type comparison using **HSD** and **JSD** statistics for **COR-1** combined dataset. Methods are described in sections **E** and **D**, respectively. In the (reject  $H_0$ ?) column, a value of 1 indicates that the null hypothesis ( $H_0$ ) is rejected and a value of 0 indicates that  $H_0$  is not rejected.

group 1	group 2	date	mean diff	lower bound CI	upper bound CI	KLD	JSD	reject $H_0$ ?
COR-1 pB	MAS $n_e$ POS	combined	-4.076	-4.424	-3.728	0.079	0.020	1
COR-1 pB	MAS $n_e$ LOS	combined	-4.051	-4.394	-3.707	0.056	0.014	1
COR-1 pB	FORWARD pB	combined	-2.435	-2.796	-2.074	0.033	0.007	1
COR-1 pB	random	combined	30.692	30.334	31.050	0.881	0.215	1
MAS $n_e$ POS	MAS $n_e$ LOS	combined	0.025	-0.259	0.309	0.022	0.006	0
MAS $n_e$ POS	FORWARD pB	combined	1.641	1.336	1.946	0.026	0.007	1
MAS $n_e$ POS	random	combined	34.767	34.466	35.069	1.169	0.268	1
MAS $n_e$ LOS	FORWARD pB	combined	1.616	1.316	1.916	0.010	0.003	1
MAS $n_e$ LOS	random	combined	34.742	34.446	35.039	1.149	0.263	1
FORWARD pB	random	combined	33.127	32.810	33.443	0.998	0.231	1



**Figure 11.** Heatmaps of **JSD** statistics shown in table 4. The left plot displays a comparison of the **JSD** metrics between **COR-1 pB**, **MAS  $n_e$  POS**, **MAS  $n_e$  LOS**, and **FORWARD pB**. The right plot displays this same comparison, but also includes the random population in the heatmap. In each plot, each diagonal cell has a value of 0 as by definition the **JSD** between two identical probability densities is 0.

quantifiable metric for both the performance of **QRaFT** in identifying magnetic structure and for how well the **MHD** simulation represents the observational data analyzed, these results can be used as a benchmark in which to improve both respectively. For example, if a major source of error came from the assumptions made in the **MHD** model, if utilizing different assumptions produced better metrics while using this framework, then this would suggest that those different assumptions resulted in a better representation of the observational data. This would allow us to conclude that the **MHD** model performed better utilizing those different set of assumptions. An example of this could be the photospheric magnetic field map used as a constraint for generating the **MHD** model. Likewise if a major source of error came from **QRaFT**, if a modification to **QRaFT**'s code resulted in better metrics, then this would suggest that this modification improved **QRaFT**'s performance and ability to project the true orientation of quasi-radial plasma density features.

The major advantage of our framework is the quantitative metrics describing both the accuracy of feature tracing methods and how representative **MHD** models of the white-light corona are to observational white-light data. This framework could allow us to improve both feature tracing methods and **MHD** modeling, and also compare different

varieties of each. This framework has the potential to provide important insight to validating segmentations from white light data against ground-truth **MHD** models, especially since we can now provide a way to quantitatively state how representative an **MHD** model is to the large scale coronal features observed in white-light data. This may unlock insights necessary to improve both **MHD** modeling and feature tracing that would not have been obvious by utilizing only qualitative comparisons. This is important, as how accurately we are able to segment the orientation of large scale plasma density features in the corona directly impacts the accuracy and uncertainty of boundary conditions such as the source surface boundary, which is used to set the overall orientation of the **IMF** and dictate solar wind propagation by space weather models such as **WSA**. [McGregor et al. \(2008\)](#) showed that improving this condition improved the **WSA** model skill at predicting both solar wind speed and **IMF** polarity at Earth. This indicates that even a small error or uncertainty in estimating the orientation of quasi-radial magnetic field lines can significantly affect the prediction of these properties at Earth when utilizing space weather prediction models like **WSA**. The ability to quantify and improve this accuracy therefore directly translates into the ability to improve space weather forecasting.

## 5. CONCLUSIONS AND FUTURE WORK

### 5.1. *Summary of Results*

In this work we have shown several key results:

1. We have developed an innovative data analysis framework enabling a quantitative assessment of the performance of image-based coronal magnetic field tracing techniques used in observational solar physics. By studying synthetic coronal images produced by a global **MHD** model, we were able to derive a set of rigorous statistical metrics describing the consistency between the image-derived coronal magnetic geometry and the “ground truth” coronal magnetic field predicted by the model.
2. The designed methodology has been tested on the **QRaFT** field-line tracing package applied to outputs of the high-resolution **MAS** model. The results demonstrate that directions of the traced image features are accurate within approximately  $\pm 10^\circ$  of the **POS**-projected magnetic field of the **MAS** model. This result provides confidence that the large scale magnetic geometry of the **POS** projected coronal magnetic field can be measured via remote sensing techniques such as coronal segmentation.
3. In addition to standard statistical metrics such as the average and median discrepancies between the segmented features and the coronal magnetic field, we developed a set of entropy-based metrics based on an in-depth analysis of the shape of the misalignment errors characterizing the segmented features. These entropy metrics provide important additional information about how performance profiles compare across different datasets using angular errors by measuring their similarity.
4. Using the designed ensemble of performance metrics, we evaluated the contribution of several types of errors to the total misalignment error. These errors included the errors resulting from image noise, feature-tracing errors, errors introduced by **MHD** boundary conditions and first-principle physics included in the coronal model, etc.
5. The introduced methodology has universal application and can be applied to a wide variety of coronal feature tracing methodologies and **MHD** models. This research has resulted in a data product available to the space weather community available at [Rura \(2025\)](#).

### 5.2. *Quantitative Conclusions*

In this work we have shown a strong correlation between the orientation of coronal features segmented by **QRaFT**, an automated quasi-radial feature tracing method, on white light **COR-1** polarized brightness observations against corresponding model outputs including the **FORWARD** polarized brightness, **MAS** line-of-sight integrated electron density, and **MAS** central plane-of-sky electron density, when compared to the **MAS** plane-of-sky magnetic field. We conclude using the methodology and data populations described in sections 2.7 and 2.8 respectfully that the **MAS**  $n_e$  **POS** performs the best when compared to the **POS**  $\vec{B}$ , followed by the **MAS**  $n_e$  **LOS**, **FORWARD** pB, and **COR-1** pB as shown in table 2. We also conclude by utilizing metrics such as the **JSD** and **HSD** displayed in table 4 that the **MAS**  $n_e$  **POS** distribution shows the strongest correlation to the **MAS**  $n_e$  **LOS**, followed by the **FORWARD** pB and **COR-1** pB distributions. While the **MAS**  $n_e$  **POS** correlated the best with the **POS**  $\vec{B}$ , as expected, all other comparison populations also correlated reasonably well, especially when compared to the random distribution used as

a control analysis as shown in tables 2 and 4. This is also evident in figure 11, where the JSD metrics of the aggregated population datasets are shown both with and without the random population statistics for comparison.

We believe we have also created a framework that is able to diagnose error contributions quantitatively depending on the assumptions made when selecting model datasets to compare against. As mentioned in section 4.3.3, we believe we measured a  $1.616^\circ$  error contribution from the FORWARD calculated pB when compared to the MAS line-of-sight integrated electron density. We believe this is measuring the difference between this electron density integral and the theoretical calculations involved in modeling the physics of pB observables, including Thomson scattering and other methods used in FORWARD when synthesizing model pB observations. We also measured a  $0.025^\circ$  error produced from the geometric projection of the 3-D volume of the MAS  $n_e$  LOS onto the POS. The latter statistic is noteworthy as it suggests that where QRaFT is identifying features, the material tends to be near the plane-of-sky. A more comprehensive study is needed to validate the assumption mentioned in Jones et al. (2020) that all features segmented in coronal images appear within the image plane. Overall, we measured a  $2.435^\circ$  difference between the FORWARD pB and COR-1 pB populations and a  $4.076^\circ$  difference between the MAS  $n_e$  POS and COR-1 pB populations, as shown in table 4. While it’s tempting to breakdown the potential reasons between these differences in population, these lie beyond the scope of this paper. Furthermore, while we hint at the potential meaning of these differences, this is only one set of results on one comparison of one dataset. A more comprehensive study would be needed in order to confidently make any of these conclusions. It should be noted that overall, these differences are significantly smaller than the difference from the random population as shown in table 4. This gives us confidence that QRaFT is able to measure magnetic structure and that this framework has the ability to determine the accuracy of this measurement against MHD simulations.

### 5.3. Future Work

By performing a statistical analysis to determine the similarity of the QRaFT feature tracing method’s performance between model and white-light data when compared to the expected POS  $\vec{B}$  from the model, we attest that we have defined a metric to quantify the accuracy and uncertainty of the model outputs compared to the white-light observations. With this framework defined and illustrated, interesting future analyses may be performed. This framework may allow us to explore the latitudinal relationship between the performance of feature tracing and expose regions of the corona that are more poorly traced than others. We can also investigate the performance of the feature tracing analysis based on the length or location of the feature. As mentioned in section 4.3.1, we also plan to apply this framework to white-light pB data from K-Cor to analyze the similarity of QRaFT’s performance between ground-based and space-based coronagraphs. This would also allow us to confirm this framework’s reliability between both observatory types. The metrics defined in this framework give us a toolset to determine how well MHD models match white-light segmentations in specific regions of the Sun, providing opportunities for new physics to be discovered and for modeling to be improved.

It is important to note that it does not matter in the context of this framework’s statistics which data population described in section 2.8 we choose to be the “right answer” to compare against when performing the JSD analysis against the other populations since this statistic is symmetric, as can be seen in Figure 11. This helps eliminate bias in the framework and is advantageous given that there is little to no ground truth precedent in determining the photospheric magnetic field (Jones et al. 2016). Since there is no precedent in determining this magnetic field, we rely on PSI’s MAS model as our ground truth solution to compare the coronal segmentations against. Having, however an unbiased metric determining the similarity of the orientation to these coronal segmentations against this expected magnetic field without assuming which is the right answer helps to equally validate these segmentations against the model solution and vice-versa. We believe that this unbiased analysis results in a valid metric for determining both the accuracy and uncertainty of white-light coronal segmentations against MHD model outputs and vice-versa.

In conclusion, we developed a set of ready-to-use numerical tools for measuring a geometric discrepancy between a solar coronal model and real coronal structure, ranging from simple statistical indicators to sophisticated distribution comparison metrics. While we attest that this tool will have future application in quantifying the uncertainty in automated feature tracing methods and constraining boundary conditions in space weather models, we show that this tool already has immediate application by quantifying how well an MHD model of the corona fits the observed structure. We believe this is a valuable framework for space weather modeling, as it provides quantifiable metrics into how well a specific model run fits the global structure of the corona. We believe these metrics are also valuable to use for coronal segmentations, as they can be used to determine the accuracy of white-light coronal segmentations against



a ground-truth solution using an [MHD](#) model. Using this framework, we have shown that a coronal segmentation method identifies the global large-scale orientation of the coronal magnetic field in white-light [polarized brightness](#) observations within a reasonable degree of precision.

*Facilities:* [STEREO COR-1 Coronagraph Instrument \(COR-1\)](#)

*Software:* [Astropy](#) ([Astropy Collaboration et al. 2013](#)), [CoronaMETRIC](#) ([Rura 2025](#)), [FORWARD](#) ([Gibson et al. 2016](#)), [SciPy](#) ([Virtanen et al. 2020](#)), [SunPy](#) ([The SunPy Community et al. 2020](#))

We thank the [STEREO COR-1](#) team for providing coronal imaging data, and Predictive Science Inc. for their collaboration and assistance in handling their [MAS](#) model solution. C. Rura, S. Jones and V. Uritsky were funded through the Partnership for Heliophysics and Space Environment Research ([NASA](#) grant No. 80NSSC21M0180) and the Windows of the Universe Multi-Messenger Astronomy program ([NSF](#) grant AST-0946422). C.D. was supported by the [NASA](#) Living With a Star Science (80NSSC22K1021) and Living With a Star Strategic Capabilities (80NSSC22K0893) programs. N. Alzate acknowledges support from [NASA ROSES](#) through HGI grant No. 80NSSC20K1070. C.N.A. and S.I.J. were supported in part by the NASA competed Heliophysics Internal Scientist Funding Model (ISFM). C.E.R., S.I.J., and V. M.U. were supported in part by [STEREO COR-1](#) via the PHASER (80NSSC21M0180) and CEPHEUS (NNG11PL10A) awards.

#### Acronyms

$H_0$ : null hypothesis. [19](#), [21](#), [22](#), [28](#)

$\vec{B}$ : magnetic field. [2](#), [8–10](#), [12–14](#), [16](#), [19](#), [20](#), [23–25](#)

$n_e$ : electron density. [8](#), [10–15](#), [19–25](#)

**COR-1 pB**: COR-1 polarized brightness. [4–6](#), [11–17](#), [20–24](#)

**FORWARD pB**: FORWARD polarized brightness. [6](#), [10](#), [12–15](#), [19–24](#)

**MAS  $n_e$  LOS**: MAS line-of-sight integrated electron density. [10](#), [12](#), [15](#), [19–24](#)

**MAS  $n_e$  POS**: MAS central plane-of-sky electron density. [8](#), [10–15](#), [19–24](#)

**POS  $\vec{B}$** : plane-of-sky magnetic field. [8](#), [10](#), [12–14](#), [16](#), [19](#), [20](#), [23](#), [24](#)

**STEREO A**: STEREO Ahead. [4](#)

**STEREO B**: STEREO Behind. [4](#)

**CI**: confidence interval. [21](#), [22](#)

**COR-1**: [STEREO COR-1 Coronagraph Instrument](#). [2](#), [4–8](#), [10–17](#), [20–25](#)

**CoronaMETRIC**: Corona Magnetic-field Evaluation Through Real Image-model Comparison ([Rura 2025](#)). [25](#)

**CR**: Carrington rotation. [6](#), [7](#), [18](#)

**EM**: electromagnetic. [2](#)

**FITS**: Flexible Image Transport System. [4](#)

**FORWARD**: FORWARD: A Toolset for Model-Data Comparison ([Gibson et al. 2016](#)). [1–3](#), [5](#), [6](#), [8–10](#), [12–16](#), [19–25](#)

**HSD**: Tukey’s Honestly Significant Difference Test. [10](#), [13](#), [19–23](#), [28](#), [29](#)

**IMF**: Interplanetary Magnetic Field. [23](#)

**IQR:** interquartile range. [12](#)

**JSD:** Jensen-Shannon Divergence. [10](#), [12–14](#), [16](#), [19](#), [21–24](#), [27](#), [28](#)

**K-Cor:** K-Coronagraph Instrument at [Mauna Loa Solar Observatory](#). [17](#), [24](#)

**KLD:** Kullback-Leibler Divergence. [12](#), [13](#), [21](#), [22](#), [28](#)

**LOS:** line-of-sight. [3](#), [4](#), [10](#), [12](#), [15](#), [19–25](#)

**MAS:** Magnetohydrodynamic Algorithm outside a Sphere Model. [1–3](#), [5](#), [6](#), [8–17](#), [19–25](#)

**MHD:** Magnetohydrodynamic. [1–3](#), [6](#), [13–19](#), [21–24](#)

**MLSO:** Mauna Loa Solar Observatory. [17](#), [26](#)

**NASA:** National Aeronautics and Space Administration. [1](#), [4](#), [25](#)

**NSF:** National Science Foundation. [25](#)

**pB:** polarized brightness. [1–7](#), [10–25](#)

**PFSS:** Potential-Field Source Surface. [2](#)

**POS:** plane-of-sky. [1–6](#), [8–16](#), [18–25](#)

**PSI:** Predictive Science Inc.. [1](#), [2](#), [6](#), [12](#), [16](#), [20](#), [24](#)

**Q1:** lower quartile. [12](#)

**Q3:** upper quartile. [12](#)

**QRaFT:** Quasi-Radial Feature Tracing Algorithm ([Uritsky 2022](#)). [1–7](#), [9–24](#), [27](#)

**ROSES:** Research Opportunities in Space and Earth Science. [25](#)

**STEREO:** Solar Terrestrial Relations Observatory. [2](#), [4](#), [6](#), [17](#), [25](#)

**WCS:** World Coordinate System. [4](#)

**WSA:** Wang-Sheely-Argé Model. [23](#)

**WTD:** Wave-Driven-Turbulence. [6](#), [18](#)

## APPENDIX

### A. CONFIDENCE INTERVALS

A confidence interval is dependent on both the standard deviation and sample size of the dataset, and gives an estimate of the margin of error of the result. A confidence interval of  $(100 - \alpha)\%$  can be defined as  $\hat{x} - (z_{1-\alpha/2}) * SE$  to  $\hat{x} + (z_{1-\alpha/2}) * SE$ , where  $\hat{x}$  is the mean value, and  $SE$  is the standard error (Gardner 2000). Choosing  $\alpha = 0.05$ , ( $z_{0.975} = 1.96$ ) we can define a 95% confidence interval as:

$$CI = 1.96 * \frac{\sigma}{\sqrt{n}} \quad (A1)$$

where  $\sigma$  is the standard deviation and  $n$  is the sample size. A 95% confidence interval is generally considered sufficient to constitute scientific discovery filtering out background factors (Cowan et al. 2011). Determining the confidence of this is useful to evaluate its uncertainty. Because the sample sizes of histograms may differ, whether because more features were extracted in one particular image vs another, or if features traced in one particular image are longer on average than another, the confidence interval of the metric may vary. This interval is also useful determining of the performance of QRaFT as it gives a measure of the range of potential error. This confidence interval analysis is therefore beneficial to further quantifying the overall uncertainty of this method.

### B. GAUSSIAN KERNEL DENSITY ESTIMATE

For a discrete distribution  $X_N \equiv \{X_1, \dots, X_N\}$ , a Gaussian kernel density estimate can be defined as:

$$\hat{f} = \frac{1}{N} \sum_{i=1}^N \phi(x, X_i; t) \quad (B2)$$

where

$$\phi(x, X_i; t) = \frac{1}{\sqrt{2\pi t}} e^{-(x - X_i)^2 / 2t} \quad (B3)$$

note that  $N$  is the number of discrete values, and  $\phi(x, X_i; t)$  is a Gaussian kernel that has location  $X_i$  and scale  $\sqrt{t}$  (Botev et al. 2010). This allows us to approximate  $N$  different values from an unknown distribution into a continuous probability density function that describes the distribution of values analytically. We note that a bandwidth factor  $\hat{h}_i$  can influence the estimate determined by the Gaussian kernel density estimate. We use the bandwidth determined by Scott's rule, which is defined as:

$$\hat{h}_i = \hat{\sigma}_i n^{-1/(d+4)} \quad (B4)$$

where  $\hat{\sigma}_i$  is the standard deviation of the data,  $n$  is the number of data values, and  $d$  is the number of dimensions (Scott 2015).

### C. STATISTICAL METRICS OF PROBABILITY DENSITIES

Kurtosis is the measurement of whether a dataset is heavy-tailed or light-tailed compared to a normal distribution, while skewness is a measurement of symmetry (or lack thereof) of a dataset (Prins 2023). These two measurements can be extremely useful in generically describing the overall shape of a given distribution, especially when combined with other statistical measurements such as the mean and standard deviation.

The following definitions are taken from Prins (2023).

For a dataset consisting of  $Y_1, Y_2, Y_3, \dots, Y_N$ , the kurtosis can be defined as

$$kurtosis = \frac{\sum_{i=1}^N (Y_i - \bar{Y})^4 / N}{s^4} - 3 \quad (C5)$$

Likewise for a dataset consisting of  $Y_1, Y_2, Y_3, \dots, Y_N$ , the skewness can be defined as

$$g_1 = \frac{\sum_{i=1}^N (Y_i - \bar{Y})^3 / N}{s^3} \quad (\text{C6})$$

where  $\bar{Y}$  is the mean of the dataset,  $s$  is the standard deviation, and  $N$  is the sample size (or number of data points). In expression C5 subtracting the kurtosis by a factor of three is an alternative definition of kurtosis named the Fisher kurtosis. This is done so that a standard normal distribution has a kurtosis of zero. This is useful for comparing a distribution to the normal distribution, as under these definitions a normal distribution with equivalent mean and standard deviation as the comparison distribution would have zero kurtosis and zero skew.

#### D. JENSEN-SHANNON DIVERGENCE

The definition of the [Jensen-Shannon Divergence \(JSD\)](#) is:

$$D_{JS}(P||Q) = \frac{1}{2}D_{KL}(P||M) + \frac{1}{2}D_{KL}(P||M) \quad (\text{D7})$$

where  $P$  and  $Q$  are two separate probability distributions,  $D_{KL}$  is the [Kullback-Leibler Divergence \(KLD\)](#) which is defined as:

$$D_{KL}(P||Q) = \sum_x P(x) \log \left( \frac{P(x)}{Q(x)} \right) \quad (\text{D8})$$

and finally  $M$  is a mixture distribution which is defined as:

$$M = \frac{1}{2}(P + Q) \quad (\text{D9})$$

([Kashyap et al. 2020](#)).

We note from expression D8 that the [KLD](#) is defined as the sum of  $\log \left( \frac{P(x)}{Q(x)} \right)$ . This has the effect of, when comparing a distribution  $Q$  to the expected distribution  $P$  at  $x$ , that if  $Q > P$ , then  $\log \left( \frac{P(x)}{Q(x)} \right) < 0$ , and likewise if  $Q < P$ , then  $\log \left( \frac{P(x)}{Q(x)} \right) > 0$ . When we sum this expression over  $x$ , the closer  $Q$  represents  $P$ , the closer the sum approaches 0. While both [KLD](#) and [JSD](#) are metrics for the similarity of two probability distributions, the [JSD](#) is considered the more reliable metric, since the [KLD](#) is considered to be asymmetric while the [JSD](#) is a symmetric version of the [KLD](#) ([Kashyap et al. 2020](#)).

#### E. TUKEY'S HONESTLY SIGNIFICANT DIFFERENCE TEST

[Tukey's Honestly Significant Difference Test \(HSD\)](#) is a statistical test aimed at measuring the honestly significant difference between two means. [Abdi & Williams \(2010\)](#) notes that the analysis of variance, or ANOVA, test can indicate that at least one population differs from others in the analysis, but does not describe the pattern of differences between these means.

[HSD](#) aims to calculate the honest statistical difference between these means using what is called the  $q$  distribution, which gives the exact sampling distribution of the largest difference between a set of means originating from the same population ([Abdi & Williams 2010](#)).

The following notation and definitions are taken from [Abdi & Williams \(2010\)](#).

This test may analyze  $A$  groups, with each group being denoted as  $a$ . The number of observations for this group is denoted as  $S_A$ , and the total number of observations is denoted as  $N$ . The mean of group  $a$  is denoted as  $M_{a+}$ . The error source (within group) is denoted  $S(\mathcal{A})$ , and the effect (between groups) is denoted  $\mathcal{A}$ . Likewise the mean square of error is denoted  $MS_{S(\mathcal{A})}$  and the mean square of effect is denoted  $MS_{\mathcal{A}}$ .

This test aims to determine whether the [null hypothesis \( \$H\_0\$ \)](#) is true. In statistics, the [null hypothesis](#) states that there is no significant difference between two sample populations, with any observed difference being due to sampling or experimental error. [HSD](#) aims to prove or disprove the [null hypothesis](#) by evaluating the difference between groups  $a$  and  $a'$ , which can be given by the  $q$  statistics. When  $H_0$  is true, the value of  $q$  is

$$q = \sqrt{\frac{M_{a+} - M_{a'+}}{\frac{1}{2}MS_{S(\mathcal{A})} \left( \frac{1}{S_a} + \frac{1}{S_{a'}} \right)}} \quad (\text{E10})$$



and follows that the studentized range  $q$  with a range of  $A$  and  $N - A$  degrees of freedom. At a given  $\alpha$  level, the ratio  $t$  would be declared significant if  $q$  was larger than the critical value for the  $\alpha$  level obtained from the  $q$  distribution. This value is denoted  $q_{A,a}$  where  $\nu = N - A$  is the number of degrees of freedom of error, and  $A$  is the range, (which in this case is equal to the number of groups). The difference between the means of groups  $a$  and  $a'$  will be significant if

$$|M_{a+} - M_{a'+}| > \text{HSD} = q_{A,a} \sqrt{\frac{1}{2} M S_{S(A)} \left( \frac{1}{S_a} + \frac{1}{S_{a'}} \right)} \quad (\text{E11})$$

To then evaluate the difference between the means of groups  $a$  and  $a'$ , the absolute value of the difference between the means is compared to the value of **HSD**, and if

$$|M_{a+} - M_{a'+}| > \text{HSD} \quad (\text{E12})$$

then the comparison is significant at the chosen  $\alpha$  level. This procedure is then repeated for all  $\frac{A(A-1)}{2}$  comparisons.

**Prins (2023)** provides an example of the **HSD** analysis that helps to highlight its practical use in statistics. In our methods, the SciPy routine *scipy.stats.tukey\_hsd* is used to calculate the **HSD**.

## REFERENCES

- Abdi, H., & Williams, L. J. 2010, Encyclopedia of research design, 3, 1
- Antonucci, E., Harra, L., Susino, R., & Telloni, D. 2020, Space Science Reviews, 216, 117, doi: [10.1007/s11214-020-00743-1](https://doi.org/10.1007/s11214-020-00743-1)
- Aschwanden, M. J. 2005, Physics of the Solar Corona. An Introduction with Problems and Solutions (2nd edition)
- Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, A&A, 558, A33, doi: [10.1051/0004-6361/201322068](https://doi.org/10.1051/0004-6361/201322068)
- Billings, D. E. 1966, A guide to the solar corona
- Blackwell, D., Dewshirst, D., & Ingham, M. 1967, in Advances in Astronomy and Astrophysics, Vol. 5, The Zodiacal Light, ed. Z. KOPAL (Elsevier), 35, doi: <https://doi.org/10.1016/B978-1-4831-9923-8.50006-1>
- Boe, B., Habbal, S., Downs, C., & Druckmüller, M. 2021, ApJ, 912, 44
- . 2022, ApJ, 935, 173
- Botev, Z. I., Grotowski, J. F., & Kroese, D. P. 2010, arXiv e-prints, arXiv:1011.2602, <https://arxiv.org/abs/1011.2602>
- Cowan, G., Cranmer, K., Gross, E., & Vitells, O. 2011, The European Physical Journal C, 71, doi: [10.1140/epjc/s10052-011-1554-0](https://doi.org/10.1140/epjc/s10052-011-1554-0)
- de Wijn, A. G., Burkepile, J. T., Tomczyk, S., et al. 2012, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 8444, Ground-based and Airborne Telescopes IV, ed. L. M. Stepp, R. Gilmozzi, & H. J. Hall, 84443N, doi: [10.1117/12.926511](https://doi.org/10.1117/12.926511)
- Downs, C., Lionello, R., Mikić, Z., Linker, J. A., & Velli, M. 2016, ApJ, 832, 180
- Downs, C., Roussev, I. I., van der Holst, B., et al. 2010, ApJ, 712, 1219, doi: [10.1088/0004-637X/712/2/1219](https://doi.org/10.1088/0004-637X/712/2/1219)
- Elmore, D. F., Burkepile, J. T., Darnell, J. A., Lecinski, A. R., & Stanger, A. L. 2003, in Polarimetry in Astronomy, ed. S. Fineschi, Vol. 4843, International Society for Optics and Photonics (SPIE), 66 – 75, doi: [10.1117/12.459279](https://doi.org/10.1117/12.459279)
- Gardner, M. 2000, Confidence intervals rather than P values, ed. D. Altman (BMJ Books), 15–27, <https://ebookcentral.proquest.com/lib/cua/detail.action?docID=1813669>
- Gibson, S., Kucera, T., White, S., et al. 2016, Frontiers in Astronomy and Space Sciences, 3, 8, doi: [10.3389/fspas.2016.00008](https://doi.org/10.3389/fspas.2016.00008)
- Hayashi, K., Abbett, W. P., Cheung, M. C. M., & Fisher, G. H. 2021, ApJS, 254, 1, doi: [10.3847/1538-4365/abe9b5](https://doi.org/10.3847/1538-4365/abe9b5)
- Hayashi, K., Wu, C.-C., & Liou, K. 2022, ApJ, 930, 60, doi: [10.3847/1538-4357/ac6173](https://doi.org/10.3847/1538-4357/ac6173)
- HelioWeb, N. 2017, Heliocentric Trajectories for Selected Spacecraft, Planets, and Comets, <https://omniweb.gsfc.nasa.gov/coho/helios/heli.html>
- Hill, F. 1994, in Graphics Gems, ed. P. S. Heckbert (Academic Press), 138–148, doi: <https://doi.org/10.1016/B978-0-12-336156-1.50023-9>
- Howard, T. A., & Tappin, S. J. 2009, SSRv, 147, 31, doi: [10.1007/s11214-009-9542-5](https://doi.org/10.1007/s11214-009-9542-5)
- Jones, S. I., Davila, J. M., & Uritsky, V. 2016, the Astrophysical Journal, 820, 113, doi: [10.3847/0004-637X/820/2/113](https://doi.org/10.3847/0004-637X/820/2/113)
- Jones, S. I., Uritsky, V. M., Davila, J. M., & Troyan, V. N. 2020, the Astrophysical Journal, 896, 57, doi: [10.3847/1538-4357/ab8cb9](https://doi.org/10.3847/1538-4357/ab8cb9)

- Kashyap, A. R., Hazarika, D., Kan, M.-Y., & Zimmermann, R. 2020, arXiv e-prints, arXiv:2010.12198.  
<https://arxiv.org/abs/2010.12198>
- Lionello, R., Linker, J. A., & Mikić, Z. 2009, *ApJ*, 690, 902
- Lionello, R., Velli, M., Downs, C., et al. 2014, *ApJ*, 784, 120
- McGregor, S. L., Hughes, W. J., Arge, C. N., & Owens, M. J. 2008, *Journal of Geophysical Research: Space Physics*, 113, doi: <https://doi.org/10.1029/2007JA012330>
- Mikić, Z., Downs, C., et al. 2018, *Nature Astronomy*, 2, 913, doi: [10.1038/s41550-018-0562-5](https://doi.org/10.1038/s41550-018-0562-5)
- Mikić, Z., Linker, J. A., Schnack, D. D., Lionello, R., & Tarditi, A. 1999, *Physics of Plasmas*, 6, 2217
- Murray, S. A. 2013, PhD thesis, University of Dublin, Trinity College
- Pasachoff, J. M. 2009, *Nature*, 459, 789, doi: [10.1038/nature07987](https://doi.org/10.1038/nature07987)
- Prins, J. 2023, NIST/SEMATECH e-Handbook of Statistical Methods. <https://doi.org/10.18434/M32189>
- Réville, V., Velli, M., Panasenco, O., et al. 2020, *ApJS*, 246, 24
- Rura, C. 2025, crura/CoronaMETRIC: CoronaMETRIC initial release from paper 1 results., v1.0.0, Zenodo, doi: [10.5281/zenodo.15003003](https://doi.org/10.5281/zenodo.15003003)
- Scherrer, P. H., Schou, J., Bush, R. I., et al. 2012, *SoPh*, 275, 207
- Schmit, D. J., Gibson, S., de Toma, G., et al. 2009, *Journal of Geophysical Research (Space Physics)*, 114, A06101, doi: [10.1029/2008JA013732](https://doi.org/10.1029/2008JA013732)
- Scott, D. W. 2015, *Multivariate Density Estimation: Theory, Practice, and Visualization*
- Solanki, S. K., Inhester, B., & Schüssler, M. 2006, *Reports on Progress in Physics*, 69, 613, doi: [10.1088/0034-4885/69/3/r02](https://doi.org/10.1088/0034-4885/69/3/r02)
- Tarr, L. A., Kee, N. D., Linton, M. G., Schuck, P. W., & Leake, J. E. 2024, *ApJS*, 270, 30, doi: [10.3847/1538-4365/ad0e0c](https://doi.org/10.3847/1538-4365/ad0e0c)
- The SunPy Community, Barnes, W. T., Bobra, M. G., et al. 2020, *The Astrophysical Journal*, 890, 68, doi: [10.3847/1538-4357/ab4f7a](https://doi.org/10.3847/1538-4357/ab4f7a)
- Thompson, W. T. 2010, 4. [https://soho.nascom.nasa.gov/solarsoft/gen/idl/wcs/wcs\\_tutorial.pdf](https://soho.nascom.nasa.gov/solarsoft/gen/idl/wcs/wcs_tutorial.pdf)
- Thompson, W. T., Wei, K., Burkepile, J. T., Davila, J. M., & St. Cyr, O. C. 2010, *Solar Physics*, 262, 213, doi: [10.1007/s11207-010-9513-8](https://doi.org/10.1007/s11207-010-9513-8)
- Thompson, W. T., Davila, J. M., Fisher, R. R., et al. 2003, in *Innovative Telescopes and Instrumentation for Solar Astrophysics*, ed. S. L. Keil & S. V. Avakyan, Vol. 4853, International Society for Optics and Photonics (SPIE), 1 – 11, doi: [10.1117/12.460267](https://doi.org/10.1117/12.460267)
- Uritsky, V. 2022, uritsky/QRaFT: v2.0, v2.0, Zenodo, doi: [10.5281/zenodo.7410948](https://doi.org/10.5281/zenodo.7410948)
- van der Holst, B., Sokolov, I. V., Meng, X., et al. 2014, *ApJ*, 782, 81
- Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, *Nature Methods*, 17, 261, doi: [10.1038/s41592-019-0686-2](https://doi.org/10.1038/s41592-019-0686-2)
- Yeates, A. R., Amari, T., Contopoulos, I., et al. 2018, *SSRv*, 214, 99