Inverse Calculation and Regularization Process for the Solar Aspect System (SAS) of HXI Payload on ASO-S Spacecraft (Postprint)

Authors: Ji-Rui Yu, Ping Ruan, Yang Su, Ying-Hong He, Jin-You Tao, Zhe Zhang, Song Guo, Bin Xue and Jian-Feng Yang

Date: 2024-05-10T00:00:00+00:00

Abstract

For the ASO-S/HXI payload, the accuracy of the flare reconstruction is reliant on important factors such as the alignment of the dual grating and the precise measurement of observation orientation. To guarantee optimal functionality of the instrument throughout its life cycle, the Solar Aspect System (SAS) is imperative to ensure that measurements are accurate and reliable. This is achieved by capturing the target motion and utilizing a physical model-based inversion algorithm. However, the SAS optical system' s inversion model is a typical ill-posed inverse problem due to its optical parameters, which results in small target sampling errors triggering unacceptable shifts in the solution. To enhance inversion accuracy and make it more robust against observation errors, we suggest dividing the inversion operation into two stages based on the SAS spot motion model. First, the as-rigid-as-possible (ARAP) transformation algorithm calculates the relative rotations and an intermediate variable between the substrates. Second, we solve an inversion linear equation for the relative translation of the substrates, the offset of the optical axes, and the observation orientation. To address the ill-posed challenge, the Tikhonov method grounded on the discrepancy criterion and the maximum a posteriori (MAP) method founded on the Bayesian framework are utilized. The simulation results exhibit that the ARAP method achieves a solution with a rotational error of roughly $±$35 (1/2-quantile); both regularization techniques are successful in enhancing the stability of the solution, the variance of error in the MAP method is even smaller—it achieves a translational error of approximately $±$18 m (1/2-quantile) in comparison to the Tikhonov method' s error of around $±$24 m (1/2-quantile). Furthermore, the SAS practical application data indicates the method' s usability in this study. Lastly, this paper discusses the intrinsic interconnections between the regularization methods.

Full Text

Preamble

ChinaXiv Research in Astronomy and Astrophysics, 24:045003 (22pp), 2024 April © 2024. National Astronomical Observatories, CAS and IOP Publishing Ltd. Printed in China and the U.K. https://doi.org/10.1088/1674-4527/ad283b

Inverse Calculation and Regularization Process for the Solar Aspect System (SAS) of HXI Payload on ASO-S Spacecraft

Ji-Rui Yu[1,2], Ping Ruan[1], Yang Su[3,4], Ying-Hong He[1], Jin-You Tao[1], Zhe Zhang[3], Song Guo[1,2], Bin Xue[1], and Jian-Feng Yang[1]

[1] Xi' an Institute of Optics and Precision Mechanics of Chinese Academy of Sciences, Xi' an 710119, China; yujirui@opt.ac.cn

[2] University of Chinese Academy of Sciences, Beijing 100049, China

[3] Key Laboratory of Dark Matter and Space Astronomy, Purple Mountain Observatory, Chinese Academy of Sciences, Nanjing 210034, China

[4] School of Astronomy and Space Science, University of Science and Technology of China, Hefei 230026, China

Received 2023 December 10; revised 2024 January 28; accepted 2024 February 6; published 2024 March 20

Abstract

For the ASO-S/HXI payload, the accuracy of flare reconstruction depends critically on factors such as the alignment of the dual grating and precise measurement of observation orientation. To ensure optimal instrument performance throughout its lifecycle, the Solar Aspect System (SAS) is essential for providing accurate and reliable measurements by capturing target motion and employing a physical model-based inversion algorithm. However, the SAS optical system' s inversion model constitutes a typical ill-posed inverse problem due to its optical parameters, where small target sampling errors can trigger unacceptable shifts in the solution. To enhance inversion accuracy and robustness against observation errors, we propose dividing the inversion operation into two stages based on the SAS spot motion model. First, the as-rigid-as-possible (ARAP) transformation algorithm calculates the relative rotations and an intermediate variable between the substrates. Second, we solve an inverse linear equation for the relative translation of the substrates, the offset of the optical axes, and the observation orientation. To address the ill-posed challenge, we utilize the Tikhonov method based on the discrepancy criterion and the maximum a posteriori (MAP) method based on the Bayesian framework. Simulation results demonstrate that the ARAP method achieves a solution with a rotational error of approximately $±$3.5 (1/2-quantile). Both regularization techniques successfully enhance solution stability, with the MAP method exhibiting even smaller error variance—it achieves a translational error of approximately $±$18 m (1/2-quantile) compared to the Tikhonov method' s error of around $±$24 m (1/2-quantile). Furthermore, SAS practical application data confirms the method' s usability. Finally, this paper discusses the intrinsic interconnections between the regularization methods.

Key words: methods: data analysis -Sun: flares -Sun: X-rays -gamma rays

1. Introduction

With ongoing technological advancements, solar investigation has experienced renewed enthusiasm. In recent years, China' s heliophysical research has advanced significantly, yielding numerous representative outcomes and executing space-based solar expedition missions with specific objectives (Gan et al. 2019a). The ASO-S mission emerged under this opportunity as China' s inaugural comprehensive solar observatory satellite focused on detecting solar magnetic fields, flares, and coronal mass ejections (Gan et al. 2019b).

The Hard X-ray Imager (HXI) is one of the three primary payloads designed specifically for detecting solar flares in the hard X-ray spectrum. Its imaging utilizes a space-modulated dual grating collimator instrument. The principle of this imaging technique requires exact relative positioning of the grating slits (Su et al. 2019) and precise solar orientation simultaneously (Zhang et al. 2019); otherwise, the accuracy of flare intensity distribution inversion can be negatively impacted. To ensure this, an online monitoring system integrated with the instrument is necessary to measure the collimator' s status and observation direction. Solar flare detection payloads that have been launched all contain comparable systems. For instance, the Solar Aspect System (SAS) in RHESSI achieves a solar pointing accuracy of 0.4 , while the Twist Monitoring System yields a torsion measurement accuracy of a few arcseconds (Zehnder et al. 2003). The STIX payload, launched in 2020 onboard the Solar Orbiter, also featured an SAS with a solar pointing accuracy of $±$4 (Krucker et al. 2020; Warmuth et al. 2020). In HXI, for measurement convenience, we chose the mature visual measurement scheme, which has been widely applied in various fields and scenarios, such as object detection (Zou et al. 2023), parameter measurement (Hashmi et al. 2022), structure and deformation monitoring (Dong & Catbas 2021; Zhuang et al. 2022), 3D reconstruction (Ham et al. 2019), robotics (Abdulazeez & Faizi 2021), and autonomous vehicles (Joel et al. 2020). The SAS system uses images to deduce the deformation and observation orientation of the collimator by tracking light spot movement. However, significant differences in the parameters of the two SAS lenses lead to a minute eigenvalue of the coefficient matrix in the inversion function. This constitutes a common ill-posed inverse problem (Kabanikhin 2008; Engl & Groetsch 2014), where slight target sampling errors may result in oscillations and non-convergence of inversion results, making it impossible to meet accuracy requirements with conventional solution methods like Gaussian elimination, Newtonian methods, or least squares optimization.

To address this issue, we conducted two primary studies: first, utilizing the rigid body transformation model to mitigate solution instability caused by outliers; and second, implementing regularization methods to tackle the pathological inverse problem. To accomplish this objective, we commence with the physical model of SAS measurements and divide the solution process into two stages. Initially, we resolve intermediate variables from a subset of light spots in the system. Subsequently, we solve the remaining variables together with other characteristic light spots. One advantage of this approach is the utilization of the as-rigid-as-possible (ARAP) algorithm in the initial stage. This algorithm is a highly efficient 3D morphology editing tool that ensures maximum preservation of local mesh shape characteristics and produces a solution closely resembling a rigid-body transformation (Sorkine & Alexa 2007). The equations resolved during the second step following decomposition undergo formal simplification. However, the unresolved ill-posed problem presents an obstacle that limits the precision of the final solution. To address this issue, we employ regularization theory, a prevalent approach for dealing with ill-posed problems (Poggio et al. 1987; Hansen 2010). This method involves appending a penalty term to minimize the residual norm, limiting the solution' s scope so that it is no longer an "exact solution" but a harmonization of the residuals and the solution's range. Although this solution is not the precise value when the equations are solved rigorously (assuming the equations are linear and of full rank), this resolution is still more acceptable than the non-convergence of the solution resulting from the ill-posedness of the equations.

To verify the method' s effectiveness, we performed simulations of the algorithm mentioned above. The results indicate that it solved the ill-posed problem we encountered to a large extent. Additionally, the method holds significant practical value as it was implemented in the HXI instrument' s actual production process. It assists in collimator assembly and grating adjustment and also achieves in-flight monitoring and solar-pointing functions. Our research has helped SAS overcome the limitations of its principle and improve its functionality, ensuring the success of the HXI observation mission. Furthermore, our findings highlight the optimization benefits of regularization theory for ill-posed problems. This extends the potential application of visual measurement methods and provides valuable insights for the development of the next generation of solar observation equipment.

2. Measurement Principle

The SAS functions as a visual measurement system that solves the state of the HXI collimator through inverse analysis of light spot images. The system comprises two sub-optical systems: the deformation monitoring subsystem (DM) and the solar aspect subsystem (SA). The DM is a short-focus optical system with a focal length of 32.8 mm and a diagonal field of view (FOV) of 27°.8. It captures images of both the three frosted glasses affixed to the front grating substrate and the Sun, producing four spots on the detector. The SA is a longfocus optical apparatus with a focal length of approximately 1214 mm and a diagonal FOV of 45 , exclusively capable of imaging the Sun. The DM differs from the SA in that the DM lens is mounted on the SAS electronic control box, integrated with the detector, while the SA lens is suspended on the front substrate and positioned apart from the detector. The SA' s mounting configuration facilitates monitoring of the optical lens by the DM as a feature point, thus constructing an additional equation to create a more complete physical model that represents a more realistic state of the collimator. Figure 1 [Figure 1: see original paper] displays the mounting state of SAS in HXI. Figure 2 [Figure 2: see original paper] explains how SAS works.

We identify four variables that need to be monitored when tracking the collimator state: the relative rotation α of the front and rear grating substrates along the collimator main axis; the relative translation D perpendicular to the main axis between the front and rear grating substrates; the tilt offset β of the DM optical axis; and the tilt offset γ of the SA optical axis (also calibrated as the main axis of the collimator). We establish the motion pattern of the light spots to determine these four variables, of which α and D have a direct influence on the imaging effect of flares by the collimator and therefore constitute the main variables to be monitored by the SAS system. While β and γ are also state variables of HXI, they require less monitoring priority compared to α and D. This is mainly due to the fact that β represents the tilt of the SAS' s optical axis, which does not influence the imaging effect, and while γ affects flare reconstruction by representing the Sun' s pointing direction, its value can be attained through alternative methods, such as direct measurement taken by the SA in-flight.

To investigate the correlation between the motion patterns of light spots and the collimator state, we analyze three scenarios.

Case 1: The movement of the three frosted glass light spots is influenced by two factors: first, the relative rotation α and translation D that occur between the two grating substrates, and second, the axis offset β of the DM optical axis, as shown in Figure 3 Figure 3: see original paper.

Case 2: The motion of the DM solar light spot is due to two factors: the angular shift γ occurring in the SA optical axis (collimator main axis) or the offset β occurring in the DM optical axis, as shown in Figure 3(b).

Case 3: The movement of the SA solar light spot is influenced by two factors: the angular shift γ in the SA optical axis (collimator main axis) and the misalignment T in the SA optical lens concerning the detector, as shown in Figure 3(c).

To enhance our ability to create precise inversion equations, it is imperative that we calibrate the optical parameters of the DM and SA subsystems. Unlike the parameter calibration of a traditional vision measurement system, our focus is on the relationship between the motion mode of the target and its corresponding light spots to satisfy the SAS function. It is proper to directly calibrate the intermediate parameters associated with the four variables in addition to the basic internal and external parameters, principal spot position, and aberration. The primary intermediate parameters are:

1. The proportionality K of the image point' s movement resulting from the target' s movement on the front substrate (such as frosted glass or SA lens), as illustrated in Figure 4 [Figure 4: see original paper]. Considering that the SAS system contains three frosted glasses, we utilize the average of the three sets of quantities: $K = \frac{1}{3}\sum_{h=1}^{3} K_h$.
2. The relationship between the angle shift of the DM optical axis and the corresponding movement of the light spot. For the Sun target, it is denoted by L_D, and for the frosted glasses target, it is denoted by L_G.
3. The relationship L_S between the angle shift of the SA optical axis and the corresponding movement of the light spot, which can be calculated as shown in Figure 5 [Figure 5: see original paper] and Figure 6 [Figure 6: see original paper].

These four intermediate parameters constitute a crucial component in the development of SAS inversion equations. Further elaboration on these details will be provided in Section 3.3.

3. Data Processing and Regularization

Combining the physical model of the SAS, we have designed an inverse algorithm to compute the aforementioned variables. Due to computational complexity exceeding FPGA capacity, the algorithm does not operate in-flight. We simply need to record the light spots' positions in orbit as input data and correspond them to certain time points, such as flare eruption, then solve the inverse equation on the ground subsequently. Thus, we obtain the instantaneous states of the HXI collimator each time we are concerned. In the following context, the specific inverse equation set will be demonstrated and the methodology for resolving the equation will also be introduced.

In practice, the algorithm requires two types of input data: the position benchmark of the light spots' original location and their new position which occurs at a subsequent moment when changes may happen in the HXI system. Although five light spots can construct five formulas, only four unknown quantities need to be calculated, resulting in an overdetermined problem. Despite the possibility of solving it with common numerical methods, its ill-posed feature renders them all unfeasible. For this reason, the inverse process is divided into two subprocedures. The first involves computing the rotation and a so-called coupled displacement of the two substrates using the three frosted glass positions. The second involves determining the remaining variables using the outcomes from the first subprocedure and the positions of the solar light spot. Further explanation of these two subprocesses follows.

3.1. The Over-determined Function

The first subprocess that calculates the state between the two grid substrates is realized by the DM. Following the common SLAM model that projects the object coordinate to the camera coordinate, the equation can be demonstrated as follows:

$$p' = K \cdot R \cdot p + C \tag{1}$$

In the approximate case of neglecting the Z direction, $p = [x, y, 1]^T$ represents the reference position of the three frosted glasses, while p' represents the position after deformation occurs; both are input data. The K demonstrates the DM projection matrix with the form of:

$$K = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix}$$

R is the rotation matrix, and C is the translation vector. Equation (1) comprises six formulas with unknown quantities α, c_x, and c_y, constituting an overdetermined problem. Compared to the full rank equation, the overdetermined problem here cannot achieve absolute accuracy solutions but still has the advantage of obtaining a moderate solution that adapts to measurement errors, such as through least squares (LS).

However, if the LS method is used to solve Equation (1) to evaluate relative motion between grating substrates, it is necessary to determine the center point of rotation and translation first; otherwise, different definitions of the center position will yield different translations, even to the extent that at a particular center of rotation the translation is zero, which corresponds to the front grating substrate sweeping around that point by an angle.

The definition of the center of rotation on the front grating substrate is based on the theoretical center of the DM attenuator window. However, as this position cannot be measured directly, a substitute is used, namely the center of the line between the two frosted glasses, denoted as p_c in Figure 7 [Figure 7: see original paper]. Consequently, the front grating substrate corresponds to the angle of rotation of all points around p_c, and the translation corresponds to the movement of the objective function in Equation (3) at the p_c point. Specifically, to directly solve the rotation based on the location of the three frosted glasses, we have:

$$\alpha = \arg\min_{\alpha} \sum_{i=1}^{3} \| R(\alpha) \cdot (p_i - p_c) - (p'_i - p_c) \|^2 \tag{4}$$

where p_i is the initial position of the three frosted glasses, and p_i' is the new position of the three frosted glasses. Then the coupled displacement is defined as:

$$C = p_c' - p_c \tag{5}$$

where $p_c = (p_1 + p_2)/2$, and $p_c' = p_c + C$.

3.2. The Rigid Transformation Model

In actual use, we discover that light spot position measurement errors arising from the center extraction algorithm or instrument systemic errors are random quantities that cannot be predicted or corrected. This introduces non-existent deformation among the light spots' graph, which would mislead the LS solving process and sometimes cause unexpected erroneous values. The problem can be described as a non-rigid transform. To tackle this problem, we cite an ARAP iterative method that calculates the rigidity of local transformations from the surface domain. This method can also be used for comparative validation of the LS method' s calculation results.

The ARAP separates the object graph into small cells and combines the vertices of each cell to form the "energy of cell," demonstrated as Equation (6):

$$E = \sum_{i=1}^{N} \sum_{j \in N(i)} \omega_{ij} \| R_i \cdot (p_j - p_i) - (p_j' - p_i') \|^2 \tag{6}$$

where N is the sum of points in the cell, p_i is the known initial position of the ith point, p_i' is the new position after transformation, R_i denotes the rotation matrix that occurs through the ith point in the model, $j \in N(i)$ represents the jth point connecting to the ith point, and $\| \cdot \|$ denotes the general norm. We can use Equation (6) to compute each cell point rotation $R = [R_i]$ and the coordinates $p' = [p_i']$ after transformation; the smaller the energy E, the higher the rigidity of the cell shape.

Constraining the cell energy to a minimum ensures that the transformation is computed with the cell shape as rigid as possible. This would eliminate the non-rigid effect to some extent, and the solution accuracy and robustness would increase considerably. In the SAS, the frosted glasses form a triangle cell on the front grating substrate surface, as shown in Figure 8 [Figure 8: see original paper]. The frosted glasses spots'positions p_i and p_i' are known input data, which means they can be regarded as constraint points. Moreover, a virtual "controlled point" that moves following the "constraint points" passively is desired to be defined as both the rotation center and translation center of the cell. Here we choose the geometric center of frosted glasses p_1 and p_2 as the controlled point p_c, as it simply corresponds to the real center of the front substrate. Finally, we obtain a cell with four points whose variations can explicitly characterize the deviation between the two grating substrates. The deviation can be divided into two parts: first is the rotation through the controlled point, and the other is the displacement that corresponds to the movement of the controlled point.

Equation (6) computes two variables R and p' to minimize E, which is a nonlinear optimization problem that is difficult to solve for the global optimum. For simplicity, we utilize a Quasi-Newton method that fixes one variable value as known and calculates the other, then makes an exchange to realize iteration and gradually obtain global optimization. In practice, we first set an initial value for the controlled point and combine the observation data to form p', then the cell vertex vectors can be defined as Equations (7) and (8):

$$e_{ij} = p_j - p_i \tag{7}$$

$$e'_{ij} = p'_j - p'_i \tag{8}$$

By regarding the points' position values p' as known quantities, Equation (6) can be transformed into matrix form:

$$E = \sum_{i=1}^{N} \sum_{j \in N(i)} \omega_{ij} \| R_i \cdot e_{ij} - e'_{ij} \|^2 \tag{9}$$

While the unknown quantities p'_i in Equation (12) form a column vector, it can be transformed into matrix form:

$$A \cdot p' = b \tag{10}$$

where the coefficient matrix A elements are weighted by ω_{ij}:

$$A_{ij} = \begin{cases} \sum_{k \in N(i)} \omega_{ik} & \text{if } i = j \\ -\omega_{ij} & \text{if } j \in N(i) \\ 0 & \text{otherwise} \end{cases}$$

As Equation (9) is a function of variable R_i, the minimization of E only concerns the term R_i, which is:

$$R_i = \arg\max_{R_i} \mathrm{Tr}(R_i^T \cdot S_i) \tag{11}$$

where $S_i = \sum_{j \in N(i)} \omega_{ij} e_{ij} \cdot e'^T_{ij}$ builds a weighted square matrix. Using singular value decomposition (SVD), $S_i = U_i D_i V_i^T$, and according to matrix trace rules, R_i would be chosen to satisfy the maximum when $R_i^T S_i$ remains a symmetric positive semi-definite matrix. Thus, the R_i value is:

$$R_i = V_i U_i^T \tag{12}$$

This is the first iterative step we apply to obtain the rotation matrix R, and the second step is to use the known R value to deduce the new cell vertex positions p'. At this time R is fixed, we obtain p' by maximizing the E value through partial derivative calculation:

$$\frac{\partial E}{\partial p'} = 0 \tag{13}$$

Equation (11) is a linear equation set containing M formulas, where the ith formula has the form:

$$\sum_{j \in N(i)} \omega_{ij}(p'_i - p'_j) = \sum_{j \in N(i)} \omega_{ij} R_i \cdot (p_i - p_j) \tag{14}$$

Actually, we only need to compute the controlled point position p'_c, as the constrained points' positions p'_k ($k \in$ constrain) are already determined by the input data. Thus, the corresponding kth row and column of A, combined with the kth row of b, are removed, leaving only the elements concerning the controlled point. A new equation is reconstructed as:

$$A_{cc} \cdot p'_c = b_c \tag{15}$$

By solving Equation (15), the unknown controlled points p'_c are computed to satisfy the lowest energy of cells. In practice, we have three constrained points corresponding to the frosted glasses; as a result, Equation (15) has degenerated into a single element equation, and the unique variable' s solution p'_c demonstrates the new position of the rotation center.

Finally, the new rotation matrix R and new cell vertexes p' are brought back into Equations (15) and (10) separately to perform the iteration process, stopping at a suitable condition such as when the amount of iterative change is less than a certain threshold. We use the rotation center to describe the rotation and translation of the substrate; thus, $R = R_c = \alpha$ and $C = p'_c - p_c$.

3.3. The Inverse Function

As previously stated, the ARAP algorithm provides solely a relative rotation α and coupled translation C between the front and rear substrates. To solve for the remaining state variables D, β, and γ, it is essential to establish equations along with the motion pattern. Examining the first motion pattern in Figure 3(a), we see that both D and β cause coupled translation C, enabling us to obtain the first equation of motion:

$$C = K \cdot D + L_G \cdot \beta \tag{16}$$

Similarly, according to the second mode of motion in Figure 3(b), the motion of the Sun light spot in the DM is jointly caused by β and γ, yielding the second equation of motion:

$$q' - q = L_D \cdot \beta + L_S \cdot \gamma \tag{17}$$

where q is the initial position of the DM Sun light spot and q' is the position after the instrument state change. According to the third motion model in Figure 3(c), the motion of the SA Sun light spot is also synthesized by two factors, γ and T, respectively:

$$s' - s = L_S \cdot \gamma + T \tag{18}$$

where s is the initial position of the SA solar light spot, and s' is the position of the spot after a state change occurred in the collimator. T is a special variable representing the offset produced by the lens in SA placed on the front substrate relative to the detector mounted on the rear substrate. Since the SA lens is also a target spot in the DM and is itself a quantity measured by the DM, T can be expressed in terms of the rotation and translation of the front substrate:

$$T = t' - t \tag{19}$$

where t is the initial coordinate of the SA lens captured through the DM, and t' is its new coordinate after translation and rotation of the front substrate, which has $t' = R \cdot t + C$. The negative sign indicates the object-image opposite relationship. Note that the tangent calculations are simplified to arc angle calculations in Equations (16)-(18), considering that the actual deformation of the collimator is very small. Combining Equations (16)-(18), we determine the measurement equation set of the SAS:

$$B \cdot y = g \tag{20}$$

where

$$B = \begin{bmatrix} K & L_G & 0 \\ 0 & L_D & L_S \\ 0 & 0 & L_S \end{bmatrix}$$

and its elements involve the optical parameters defined in Section 2. $y = [D\ \beta\ \gamma]^T$ is the vector to be solved. The right side vector $g = [C\ q' - q + L_D \cdot \beta\ s' - s]^T$ represents the observed data. Thus, by back-calculating Equation (20), we realize the calculation of the state variables D, β, and γ.

3.3.1. Tikhonov Regularization Method Equation (20) is a full-rank linear equation that can typically be solved directly by Gaussian elimination or the Gauss-Seidel method. Unfortunately, inspecting the condition number of matrix B described as $\kappa(B) = \|B\| \cdot \|B^{-1}\|$ illustrates the illness of a linear equation. A larger condition number represents a more unstable inverse solution to the equation and a more sensitive response to noise. For example, in Equation (20), we obtain an enormous condition number exceeding 440,000. To analyze the reasons, we research the solution structure with the SVD method. Through matrix operation, the solution of the equation can be given by:

$$y = \sum_{i=1}^{n} \frac{u_i^T g}{\sigma_i} v_i \tag{21}$$

where u_i and v_i are the orthogonal column vectors of the left and right SVD matrices of B, and σ_i is the singular value. It is obvious that the coefficients $u_i^T g/\sigma_i$ are impacted extremely by σ_i, as they decay into tiny values in the ill-posed matrix. We cannot obtain an ideal input g with no noise but rather a data $g + \Delta g$ with perturbation. Consequently, the perturbations Δg are magnified and cause the theoretically exact solution y to become "unbounded." Thus, traditional methods cannot be used in ill-posed problem computation, and specialized approaches are required.

Equation (20) can be expressed as a least squares (LS) form:

$$\arg\min_y \|By - g\|^2 \tag{22}$$

To restrict the range of possible solution values, a penalization term $\|y\|^2$ is introduced in Equation (22), which is known as regularization. Then Equation (22) becomes:

$$\arg\min_y \{\|By - g\|^2 + \lambda\|y\|^2\} \tag{23}$$

Consequently, the solution of Equation (23) can be reformulated as:

$$y_\lambda = (B^T B + \lambda I)^{-1} B^T g \tag{24}$$

where λ is a positive parameter that balances the two parts: the residual norm $\|By - g\|^2$ and the solution norm $\|y\|^2$. The weight is enhanced to diminish the solution norm $\|y\|^2$ by enlarging the λ value, which stabilizes Equation (22) to some extent but may cause a larger deviation between the regularization solution and the ideal solution due to the limitation of $\|By - g\|^2$ being expanded. Conversely, if λ is reduced, the regularization effect is weakened. The choice of λ value then becomes the core of the problem. There are many well-established methods for selecting regularization parameters, such as the Discrepancy Principle (Engl 1987), the L-curve Criterion (Hansen 1992; Johnston & Gulrajani 2000), Generalized Cross Validation (Golub & von Matt 1997), and the Normalized Cumulative Periodogram (Mojabi & LoVetri 2008), among others. In this paper, to compute the appropriate λ, we utilize the discrepancy principle. According to Equation (24), we can obtain:

$$g - By_\lambda = \sum_{i=1}^{n} \frac{\lambda}{\sigma_i^2 + \lambda} (u_i^T g) u_i \tag{25}$$

Then the residual norm is:

$$\|g - By_\lambda\|^2 = \sum_{i=1}^{n} \left(\frac{\lambda}{\sigma_i^2 + \lambda} \right)^2 (u_i^T g)^2 \tag{26}$$

Taking into account $g = g_{\text{exact}} + e$, where e represents observation noise. Based on the properties of SVD decomposition, the coefficient $u_i^T g_{\text{exact}}$ follows a monotonically decreasing pattern and satisfies the discrete Picard condition. There exists a bounding value t; when $\sigma_i < t$, because of the dominance of g_{exact}, $u_i^T g_{\text{exact}}$ continues to decrease monotonically. However, the noise vector e may not align with g_{exact}, meaning that $u_i^T e$ does not necessarily satisfy the discrete Picard condition. Consequently, when $\sigma_i > t$, the noise term dominates, leading to $u_i^T e$ being larger than $u_i^T g_{\text{exact}}$, causing a platform or bounce in the value of $u_i^T g$ instead of decreasing, which would magnify the superimposed solution y_λ with tiny singular values. Hence, the regularization process aims to reduce this adverse effect, and the parameter λ works as a modifying factor.

To analyze the effect of regularization parameters, we assume the observation noise e is Gaussian distributed with standard deviation η. The main aim of regularization is to reduce the impact of noise. Then Equation (26) can be expressed as two parts:

$$\|g - By_\lambda\|^2 = \sum_{\sigma_i > \lambda\eta} \left(\frac{\lambda}{\sigma_i^2 + \lambda} \right)^2 (u_i^T g_{\text{exact}})^2 + \sum_{\sigma_i \le \lambda\eta} \left(\frac{\lambda}{\sigma_i^2 + \lambda} \right)^2 (u_i^T e)^2 \tag{27}$$

where $\lambda\eta$ is the boundary of the platform region. As Equation (27) shows, the residual norm is composed of the solution value, observation error, and exact observation value, with proportions that change with the λ value. The term $\|By_\lambda\|$ increases with λ, and e commonly has a smaller value than g_{exact}, whose proportion in the residual norm gradually diminishes along with increasing λ. We analyze two scenarios:

Case 1: There exists a critical value τ; when $\lambda < \tau$, the effect of singular values slightly increases, and the divergence of the superimposed solution $u_i^T g$ remains dominated by noise e. This indicates under-regularization (under-smoothing) and the solution $y_\lambda \approx y_{\text{exact}}$.

Case 2: When $\lambda > \tau$, σ_i increases significantly, which largely suppresses the amplitude of the superimposed solution $u_i^T g$. This demonstrates over-regularization, and the residual norm is collectively determined by both observation noise and exact observation value.

The discrepancy principle finds a critical λ such that the residual norm is exactly dominated by the observation error e, i.e., a regularization parameter that achieves equilibrium between the residual norm and the solution norm. This article advocates using the discrepancy principle when selecting the regularization parameter, owing to the ability to pre-evaluate the observation noise e (equivalent to the spot center extraction accuracy). Nevertheless, this still presents a potential risk, as an inadequate estimation of observation error could significantly impact the solution outcomes, particularly when the parameter estimation is low, resulting in under-regularization that amplifies the range of the solution norm and causes excessive deviation of y_λ from the ideal solution y_{exact}. To ensure caution, a safety factor ν around 2 is typically added to the noise for a slightly over-regularized state:

$$\|g - By_\lambda\| = \nu \cdot \|e\| \tag{28}$$

According to Equation (28), we compute λ and substitute it into Equation (24) to obtain the regularized solution.

3.3.2. In Bayesian Perspective Expanding on the Tikhonov regularization method discussed in Section 3.3.1, the Bayesian approach (Jin & Zou 2008a) offers fresh insight and provides cross-validation of Equation (23)' s solution. Assuming a prior distribution for observation noise and the target variable, one can estimate unknown information by determining the posterior probability density function (PPDF). From a probabilistic perspective, the solution aims to determine the probability value that maximizes y, taking into account the known observation g, and uses the maximum a posteriori (MAP) of y to estimate the optimal value. Based on Bayesian estimation fundamentals:

$$\pi(y|g) = \frac{\pi(g|y)\pi(y)}{\pi(g)} \tag{29}$$

where $\pi(g|y)$ is the likelihood function, $\pi(y)$ is the prior probability of y (Jaynes 1968), and $\pi(g)$ is a normalized constant usually ignored in calculations. Thus, Equation (29) can be expressed as a proportional relationship:

$$\pi(y|g) \propto \pi(g|y)\pi(y) \tag{30}$$

MAP can estimate the y value:

$$\hat{y}_{\mathrm{MAP}} = \arg\max_{y}\{\pi(y|g)\} \tag{31}$$

To resolve the issue at hand, it is essential to acquire the probability prior and likelihood as denoted in Equation (31). Initially, considering observation noise, Equation (20) can be expressed as:

$$g = By + e \tag{32}$$

The observation noise e follows a continuous probability distribution. Assuming the independence of e and y, the distribution of observation g is consistent with the noise distribution $\pi(g|y) = \pi(e)$. Usually, the noise e consists of independent identically distributed Gaussian random variables with mean zero and standard deviation σ_1:

$$\pi(g|y) \propto \exp\left(-\frac{1}{2\sigma_1^2}\|By - g\|^2\right) \tag{33}$$

Defining the a priori probability $\pi(y)$ presents difficulty as the quantity of deformation y in SAS measurements is unknown. Consequently, we should assume the HXI instrument tends to stabilize, with minimal deformation in the absence of other interferences, and a lower likelihood of larger deformations. It thus follows a Gaussian distribution with parameters $(0, \sigma_2)$, which helps establish the correlation between the Bayesian model and classical Tikhonov regularization theory. According to Bayesian theorem, the PPDF is:

$$\pi(y|g) \propto \exp\left(-\frac{1}{2\sigma_1^2}\|By - g\|^2\right) \cdot \exp\left(-\frac{1}{2\sigma_2^2}\|y\|^2\right) \tag{34}$$

As observation g is not affected by hyper-parameter λ, the first term on the right-hand side has $\pi(g|y, \lambda) = \pi(g|y)$. If the coefficient matrix B is considered, there exist interrelationships between observations. In this article, g is composed of the computational centers of three types of light spots, and observation noise reflects the extraction accuracy of light spots. Since the components are independent, the noise has covariance matrices:

$$S = \begin{bmatrix} \eta_c^2 & 0 & 0 \\ 0 & \eta_d^2 & 0 \\ 0 & 0 & \eta_s^2 \end{bmatrix} \tag{35}$$

where η_c, η_d, and η_s are the error variances of the three vectors of observation data. Due to the regularization parameter being a penalty factor applied to the solution norm $\|y\|^2$, it manifests as a correction of the prior probability $\pi(y|\lambda)$ within the Bayesian framework. As y follows a Gaussian distribution and its components satisfy the covariance matrix W (Calvetti & Somersalo 2018), λ serves as a scaling factor ultimately modifying the solution norm of y. In this scenario, y satisfies the Gaussian prior $\pi(y|\lambda) \sim \mathcal{N}(0, \lambda W)$ with hyper-parameter λ:

$$\pi(y|\lambda) \propto \exp\left(-\frac{\lambda}{2} y^T W^{-1} y\right) \tag{36}$$

Combining Equations (31), (33), and (36), we obtain:

$$\pi(y, \lambda|g) \propto \exp\left(-\frac{1}{2}\|By - g\|^2_{S^{-1}}\right) \cdot \exp\left(-\frac{\lambda}{2} y^T W^{-1} y\right) \cdot \pi(\lambda) \tag{37}$$

Then the objective function of MAP estimation is:

$$\arg\max_{y,\lambda} \pi(y, \lambda|g) \iff \arg\min_{y,\lambda} \left\{\|By - g\|^2_{S^{-1}} + \lambda y^T W^{-1} y - \log \pi(\lambda)\right\} \tag{38}$$

Considering that the components of y are independent of each other, W must be a diagonal matrix. This paper analyzes the different deformations represented by D, β, and γ according to SAS design specifications, combined with certain empirical assumptions and simulation data, arguing they should conform to different Gaussian priors. This involves estimating the variance of deformation distributions based on simulation input setup. It is stated that:

$$W = \begin{bmatrix} \sigma_d^2 & 0 & 0 \\ 0 & \sigma_\beta^2 & 0 \\ 0 & 0 & \sigma_\gamma^2 \end{bmatrix} \tag{39}$$

where σ_d, σ_β, and σ_γ represent the standard deviations of the three preset deformations, respectively. $\pi(\lambda)$ represents the prior distribution of hyper-parameters and can be described by the conjugate prior (Gelman et al. 1995). For this study, we use the gamma distribution with parameters α and β:

$$\pi(\lambda) = \frac{\beta^\alpha}{\Gamma(\alpha)} \lambda^{\alpha-1} e^{-\beta\lambda} \tag{40}$$

where Γ is the gamma function. Substituting Equations (35), (39), and (40) into Equation (38), we obtain the joint PPDF of y and λ:

$$\pi(y, \lambda|g) \propto \lambda^{\alpha-1} \exp\left(-\frac{1}{2}\|By - g\|^2_{S^{-1}} - \frac{\lambda}{2} y^T W^{-1} y - \beta\lambda\right) \tag{41}$$

Then the objective function of MAP estimation can be easily determined:

$$\arg\min_{y,\lambda} \{\|By - g\|^2_{S^{-1}} + \lambda y^T W^{-1} y - 2(\alpha - 1)\log\lambda + 2\beta\lambda\} \tag{42}$$

Equation (42) comprises two optimization variables, y and λ, which can be resolved via an iterative approach (Jiang et al. 2020).

Step 1: Define the initial value of parameter λ_0, and subsequently, y_k for the kth iteration is evaluated according to Equation (24):

$$y_k = (B^T S^{-1} B + \lambda_k W^{-1})^{-1} B^T S^{-1} g \tag{43}$$

Step 2: Update the parameter λ_{k+1} by differential operation $\partial\pi(y, \lambda|g)/\partial\lambda = 0$, then λ_{k+1} satisfies the function:

$$\lambda_{k+1} = \frac{2(\alpha - 1)}{y_k^T W^{-1} y_k + 2\beta} \tag{44}$$

This iterative process continues until convergence, providing a solution that resembles Tikhonov regularization and serves as a cross-check of the solution’ s reliability.

4. Numerical Simulation

For the procedures discussed previously, it is necessary to conduct simulations to assess their effectiveness in regularizing the ill-posed inverse problem in SAS. For this purpose, simulated data will be used to validate inversion accuracy, and real test results will be presented later. However, before this, we first calculate the light spot extraction accuracy of SAS, which serves as the observation error.

4.1. Accuracy of Light Spot Position

During both ground testing and in-flight operations, SAS detects various types of light spots due to different illumination sources, as demonstrated in Figures 9 [Figure 9: see original paper] and 10 [Figure 10: see original paper]. Different center coordinate extraction algorithms are implemented for various light spot forms based on their distinguishing features:

1. **Circular light spots with symmetric shape and uniform energy:** Binarized center of mass algorithm
2. **Bessel-shaped circular light spot:** Ellipse fitting center method (Fitzgibbon et al. 1999; Zhou et al. 2018)
3. **Full-disk large circular light spot:** Four-quadrant area fitting algorithm or 14-row edge data fitting algorithm

The principles behind these algorithms and extraction accuracy will be thoroughly described in our forthcoming publication. Due to space constraints, we cite the findings and present the results in Table 1 . It is evident that the extraction precision equals the noise level in observed data g, which also serves as crucial input for subsequent simulation calculations.

4.2. Operation Result

The iteration ends when the abort condition $\|y_n - y_{n-1}\| < \epsilon$ is met. We achieve the estimated solution y with the Bayesian model, which differs from the Tikhonov approach. The numerical simulations conducted in this paper were executed on a computer equipped with an Intel Core i7 CPU and 16 GB of RAM, utilizing Python 3.8.5. For all subsequent simulations, we established a consistent initial position g_0 for the spots and generated 1000 sets of random deformations as simulation inputs. The simulation sequence employed is as follows:

Step 1: Utilizing the imaging principle of SAS, we first derive the altered position g_1 of the light spot caused by simulated deformation. The resulting value $g = g_1 - g_0$ represents an ideal observation free from noise. It should be noted that the rotation and translation center of the front grating substrate is defined as the center of the line joining frosted glass 1 and 2 during forward derivation of the changing coordinate g_1.

Step 2: To the ideal observation, we add observation error: $g_{\text{noisy}} = g + e$, where e represents the extraction accuracy of the light spot center.

Step 3: Solve the rotation angle and coupled displacement with LS or ARAP method.

Step 4: Solve the inverse problem Equation (20) utilizing the Tikhonov and Bayesian methods and subsequently evaluate inversion algorithm accuracy.

It should be noted that we assumed any deformation in the HXI collimator during its full life cycle will not surpass the permissible threshold that would affect flare imaging quality when conducting numerical simulations. This constrains us to set the peak of simulated deformation equal to the maximum allowable deformation of the collimator: $\alpha = \pm 10''$ (the threshold that could impact imaging details), $D = \pm 18\mu$m (matching the period of the thinnest grating slit), $\beta = \pm 10''$, and $\gamma = \pm 10''$.

For comparison, assuming no noise (i.e., $e = 0$), the results obtained from directly solving Equation (1) with the LS and ARAP methods, and Equation (20) using Gaussian Elimination (GE), are presented in Figures 11 [Figure 11: see original paper] and 12 [Figure 12: see original paper]. Under this condition, the results demonstrate high accuracy reaching almost zero inversion error. The LS method has higher inversion accuracy compared to the ARAP method because ARAP does not solve equations exactly but instead calculates them through iterative optimization. These high accuracies show the correctness of the inverse process modeling.

Figures 13 [Figure 13: see original paper] and 14 [Figure 14: see original paper] display the outcomes upon adding observation noise e to observation data g. The two methods, LS and ARAP, yield very similar arithmetic results. The peak error of α is about $\pm 15''$, and the coupled displacement is about ± 0.1 pixels (as C is an intermediate process variable, we demonstrate it in "pixel" units). Similarly, we analyze error values at the 1/4, 1/2, and 3/4 quantiles, indicated in Figures 11–19. At the 1/2 quantile, the error of α is about $\pm 3.5''$, and the error of c_x and c_y is about 0.03 pixels. In subsequent numerical simulations, we use only the ARAP method to calculate the rotation angle for simplicity.

As a consequence, due to the ill-posed nature of inverse function (20), an unregularized computation of variables D, β, and γ (such as GE) results in a highly deviated and unacceptable solution far from ideal, as shown in Figure 15 [Figure 15: see original paper].

When using the Tikhonov method to solve Equation (20), results are displayed in Figure 16 [Figure 16: see original paper]. The α calculation is unaffected and retains the same accuracy of about $3.5''$, and regularization significantly optimizes solution outcomes for D, β, and γ, whose peak errors are approximately $\pm 55\mu$m, $\pm 10''$, and $\pm 10''$, respectively, with errors displaying uniform distribution within their peak range. At the 1/2 quantile, the solution error is estimated to be around $D = \pm 24\mu$m, $\beta = \pm 5''$, and $\gamma = \pm 5''$.

Solution results utilizing the Bayesian framework exhibit distinct characteristics. The PPDF, which accounts for Gaussian noise and Gaussian prior distribution, generates inverse error with a Gaussian distribution structure of mean value 0. This suggests that computation based on the Bayesian framework is more likely to achieve smaller inversion errors, as illustrated in Figure 17 [Figure 17: see original paper]. Based on findings, MAP produces slightly larger peak errors than the Tikhonov approach. Specifically, the peak value for D is about $\pm 65\mu$m, while β and γ both slightly exceed $\pm 10''$. However, error values at the 1/2 quantile are lower: $D = \pm 20\mu$m, $\beta = \pm 3''$, and $\gamma = \pm 3''$.

We observed that the maximum inversion error changes with preset simulated deformations. Specifically, when preset deformations are large, the final inversion error increases, whereas it decreases when deformations are small. Besides, the distributional characteristics of error persist without alteration. The conclusions presented in Figures 16 and 17 are outcomes acquired from predetermined deformations of α within $\pm 10''$, D within $\pm 18\mu$m, β within $\pm 10''$, and γ within $\pm 10''$. Applying another set of predetermined values (α within $\pm 5''$, D within $\pm 5\mu$m, β within $\pm 3''$, γ within $\pm 3''$) yields inversion outcomes exhibited in Figures 18 [Figure 18: see original paper] and 19 [Figure 19: see original paper]. Compared to larger preset deformations, rotation computation does not present an ill-posed problem, so solution accuracy remains unchanged. The peak error of D, β, and γ solutions is lower with the Tikhonov method, achieving peak errors of about $D = \pm 20\mu$m, $\beta = \pm 3''$, and $\gamma = \pm 3''$, and 1/2-quantile errors of $D = \pm 8\mu$m and $\beta = \pm 1.5''$, $\gamma = \pm 1.5''$. The MAP method achieves peak errors of about $D = \pm 25\mu$m, $\beta = \pm 4''$, $\gamma = \pm 4''$, and at the 1/2 quantile, error values are $D = \pm 6\mu$m, $\beta = \pm 1''$, $\gamma = \pm 1''$. Error distributions of both methods are unaffected by the magnitude of preset deformation. The reason is that regularization incurs a penalty factor in the form of $\|y\|^2$, and for the identical inverse equation, the ideal solution norm increases as preset deformation increases. At this time, to meet the optimization objective of Equation (23), the solution norm contour expands and intersects the smaller residual norm contour, resulting in smaller $\|By_\lambda\|$. Such regularization leads the inverse problem being closer to an unbiased solution. Hence, SAS monitoring, which necessitates the assumption that the instrument would not exhibit substantial deviation, is valid as indicated at the beginning.

Furthermore, if the collimator indeed undergoes an unacceptably large deformation that would compromise HXI imaging quality, prioritizing SAS measurement accuracy may become unnecessary if HXI usability cannot be ensured. Fortunately, HXI underwent precise calibration using an X-ray beam; additionally, installation and integration processes occurred in relatively stable circumstances, and the instrument was accurately temperature-controlled in-flight to ensure an unchanged state. Based on these factors, it is reasonable to assume that significant deformation would not occur in the collimator. These assumptions and conclusions were subsequently confirmed through in-flight testing.

In summary, simulation results are presented in tabulated form for clarity. Table 2 compares inversion results using both Tikhonov regularization and Bayesian frameworks for large preset deformations. The MAP approach offers some advantages due to the Gaussian distribution, which guarantees a higher proportion of small errors; results from both methods will be jointly considered. Table 3 presents inversion outcomes obtained with smaller preset simulated deformations; some corresponding values are smaller compared to Table 2. Since simulation data utilizes 1000 groups of randomly preset values, each simulation could produce slightly varied outcomes, but results remain at the same level. It is worth noting that the discrepancy principle can reduce peak error when over-regularized by increasing the safety factor ν. However, arbitrarily raising ν risks constraining the inversion solution to a range close to 0, potentially causing inversion error to always match actual deformation and rendering assessment of inversion error meaningless.

Additionally, numerous practical measurements were conducted using SAS throughout HXI assembly, launch, and in-orbit operation. Before satellite integration, third-party measuring equipment such as the Coordinate Measuring Machine (CMM) could validate SAS measurements comparatively. Figure 20 [Figure 20: see original paper] exhibits status monitoring outcomes of HXI at 14 crucial time points (illustrated in Table 4). The relatively flat curves in measurement results suggest the HXI collimator is in a stable state. The SAS solution deviates from CMM measurements with α less than $5''$, D_x less than 10μm, and D_y less than 13μm. Moreover, the two regularization methods show some deviations in results, with D_x less than 6μm and D_y less than 5μm, demonstrating SAS' s effective usability.

4.3. Discussion

The Tikhonov regularization and MAP methods differ in principle; however, their solution results are similar, indicating a close connection (Antoni et al. 2023). Equation (23) shows the effective solution equation is of the form $\|By - g\|^2 + \lambda\|y\|^2$, which is usually effective but not the only choice. Considering the general form where the penalty term takes the form $\|Ly\|^2$, Tikhonov regularization can be expressed as:

$$\arg\min_{y}\{\|By - g\|^2 + \lambda\|Ly\|^2\} \quad (45)$$

(Hansen 1989), where L is a discrete approximation of some operation operator on variable y in the first kind of Fredholm integral equations (Hansen 1989, 2010), describing interrelationships among variables to be solved. Tikhonov regularization is a particular case where $L = I$.

Examining Equation (42), it is evident that in the first term, the residual norm is modified with observation noise matrix S, while the second term outlines covariance matrix W to depict correlation between observed variables y. Subsequently, if $W^{-1} = L^T L$, the form coheres with Equation (45). The form of L relates to the definition of Markov Random Field used in heat conduction problems (Jin & Zou 2008b), which defines a priori probability of similar hyperparameters in this study, while y components are independent.

In actual measurement, D varies mainly due to the flatness of the satellite mounting surface and collimator deformation caused by gravity unloading after launch, set according to standard deviation $\sigma_d = 18\mu$m. The variation of β is primarily due to possible temperature changes in the satellite mounting plane, which cause the front and rear grating substrates of the collimator to tilt, set according to $\sigma_\beta = 10''$. The variation of γ is mainly due to changes in the observation optical axis caused by uncertainty in satellite attitude control, set according to $\sigma_\gamma = 10''$. This implies a constraint on y components in different proportions, and although this fits the prediction of a priori probability, it does not necessarily achieve optimal regularization in terms of Equation (45)' s form, as larger a priori variance makes L smaller, potentially resulting in under-regularization. This assertion is evident from numerical simulation and actual measurement results. Given these considerations, this research employs two regularization methods sharing the common trait of being based on generalized SVD but differing in their defining processes: the Tikhonov method builds upon LS optimization by incorporating solution range constraints, while the MAP method constructs the PPDF to build the optimization function through observation noise distribution and solution prior distribution. Nevertheless, both optimization functions strive to balance residuals and solution ranges, evidencing a close association between the two methods.

5. Conclusion

This paper addresses the problem that SAS is susceptible to noise from observation data during HXI collimator monitoring, resulting in significant deviation of the inversion solution. We propose methods to decrease SAS inversion error by utilizing ARAP and regularization techniques following the SAS physical model. Simulation data demonstrate the validity of the SAS inversion model and regularization process efficacy. Additionally, test data provide evidence that these methods can realize online monitoring functions. In summary, we make the following conclusions:

1. Based on data simulation, the LS and ARAP methods yield measurement accuracy of about $\pm 15''$ in peak and $\pm 3.5''$ in 1/2-quantile for rotational deformation. The Tikhonov regularization approach results in peak measurement $D = \pm 55\mu$m for translation, and peak measurement errors for the two optical axes are $\beta = \pm 10''$ and $\gamma = \pm 10''$, respectively. Its accuracy at the 1/2-quantile is $D = \pm 24\mu$m, $\beta = \pm 5''$, and $\gamma = \pm 5''$. The MAP method' s peak errors are slightly larger than Tikhonov' s, but 1/2-quantile accuracy is superior: $D = \pm 20\mu$m, $\beta = \pm 3''$, and $\gamma = \pm 3''$. All simulated data have considered light spot center extraction error.
2. The accuracy of the SAS algorithm' s inversion is linked to deformation scale. If deformation is insignificant, inversion accuracy is high; conversely, if deformation itself is significant, inversion accuracy is reduced.
3. Analyzing several practical test data instances, disparities exist between SAS and CMM test outcomes: specifically, α is less than $5''$, D_x is less than 10μm, and D_y is less than 13μm. Furthermore, disparities exist between Tikhonov and MAP methods, with D_x less than 6μm and D_y less than 5μm.

In summary, this paper demonstrates that regularization methods significantly improve SAS inversion calculation stability and enhance robustness against observation noise pollution. Furthermore, SAS' s monitoring feature provides vital assistance in completing the HXI mission, aiding in instrument assembly and tuning during production and monitoring state throughout HXI's entire lifecycle. Additionally, it enables solar orientation pointing during in-orbit operations.

Unfortunately, inherent limitations exist in visual measurement systems; for instance, improving light spot extraction accuracy indefinitely is impossible. Also, solution accuracy for inversion equations is predominantly determined by optical parameters, making it difficult and inefficient to increase accuracy solely through data processing methods. Therefore, future research should focus on designing more reasonable optical systems to improve inversion equation robustness and anti-interference capability for SAS-like monitoring systems, assisting data stability algorithms to enhance accuracy. This study significantly improves SAS usability, allowing it to successfully perform intended functions and make further progress toward achieving HXI mission goals. Ultimately, we anticipate such investigations will yield theories and techniques that can assist in designing forthcoming flare observation instruments.

Acknowledgments

This research received support from the Strategic Priority Research Program on Space Science of the Chinese Academy of Sciences, grant No. XDA15320104, with additional contributions from the Purple Mountain Observatory (PMO) of the Chinese Academy of Sciences and the National Space Science Center (NSSC).

ORCID iDs Ji-Rui Yu https://orcid.org/0000-0001-8903-2874

References

Abdulazeez, A. M., & Faizi, F. S. 2021, Turkish Journal of Computer and Mathematics Education (TURCOMAT), 12, 1563 Antoni, J., Idier, J., & Bourguignon, S. 2023, InvPr, 39, 065016 Calvetti, D., & Somersalo, E. 2018, WIREs Computational Statistics, 10, e1427 Dong, C.-Z., & Catbas, F. N. 2021, Structural Health Monitoring, 20, 692 Engl, H. W. 1987, JOTA, 52, 209 Engl, H. W., & Groetsch, C. W. 2014, Inverse and Ill-posed Problems (Amsterdam: Elsevier) Fitzgibbon, A., Pilu, M., & Fisher, R. B. 1999, ITPAM, 21, 476 Gan, W.-Q., Yan, Y.-H., & Huang, Y. 2019a, SSPMA, 49, 059602 Gan, W.-Q., Zhu, C., Deng, Y.-Y., et al. 2019b, RAA, 19, 156 Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. 1995, Bayesian Data Analysis (New York: Chapman and Hall/CRC) Golub, G. H., & von Matt, U. 1997, J. Computational and Graphical Statistics, 6, 1 Ham, H., Wesley, J., & Hendra, H. 2019, Int. J. Electrical and Computer Eng., 9, 2394 Hansen, P. C. 1989, BIT Numerical Mathematics, 29, 491 Hansen, P. C. 1992, SIAMR, 34, 561 Hansen, P. C. 2010, Discrete Inverse Problems: Insight and Algorithms (Philadelphia, PA: SIAM) Hashmi, A. W., Mali, H. S., Meena, A., et al. 2022, Materials Today: Proceedings, 56, 1939 Jaynes, E. T. 1968, IEEE Transactions on Systems Science and Cybernetics, 4, 227 Jiang, J.-H., Tang, H.-Z., Mohamed, M. S., Luo, S.-Y., & Chen, J.-D. 2020, ApSci, 10, 6348 Jin, B.-T., & Zou, J. 2008a, InvPr, 25, 025001 Jin, B.-T., & Zou, J. 2008b, IJNME, 76, 521 Joel, J., Fatma, G., Aseem, B., & Andreas, G. 2020, Foundations and Trends® in Computer Graphics and Vision, 12, 1 Johnston, P. R., & Gulrajani, R. M. 2000, ITBE, 47, 1293 Kabanikhin, S. I. 2008, Journal of Inverse Ill-Posed Problems, 16, 317 Krucker, S., Hurford, G. J., Grimm, O., et al. 2020, A&A, 642, A15 Mojabi, P., & LoVetri, J. 2008, Progress in Electromagnetics Research M, 1, 111 Poggio, T., Torre, V., & Koch, C. 1987, in Readings in Computer Vision, ed. M. A. Fischler & O. Firschein (San Francisco, CA: Morgan Kaufmann), 638 Sorkine, O., & Alexa, M. 2007, in Fifth Eurographics Symposium on Geometry Processing, ed. A. Belyaev & M. Garland (Goslar: Eurographics Association), 109 Su, Y., Liu, W., Li, Y.-P., et al. 2019, RAA, 19, 163 Warmuth, A., Önel, H., Mann, G., et al. 2020, SoPh, 295, 1 Zehnder, A., Bialkowski, J., Burri, F., et al. 2003, Innovative Telescopes and Instrumentation for Solar Astrophysics, 4853, 41 Zhang, Z., Chen, D.-Y., Wu, J., et al. 2019, RAA, 19, 160 Zhou, P., Wang, X.-Q., Huang, Q.-Y., & Ma, C. 2018, in 2018 2nd IEEE Advanced Information Management, Communicates, Electronic Automation Control Conference (IMCEC) (New Jersey: IEEE), 316 Zhuang, Y.-Z., Chen, W.-M., Jin, T., et al. 2022, Senso, 22, 3789 Zou, Z.-X., Chen, K.-Y., Shi, Z.-W., et al. 2023, Proc. IEEE, 111, 257

Note: Figure translations are in progress. See original paper for figures.