

Revealing asymmetry on midplane of proto-planetary disc through modelling of axisymmetric emission: methodology

Masataka Aizawa,^{1,2,3*} Takayuki Muto,^{4,5,6} Munetake Momose³

¹*Tsung-Dao Lee Institute, Shanghai Jiao Tong University, Shengrong Road 520, 201210 Shanghai, P. R. China*

²*RIKEN Cluster for Pioneering Research, 2-1 Hirosawa, Wako, Saitama 351-0198, Japan*

³*College of Science, Ibaraki University, 2-1-1 Bunkyo, Mito, Ibaraki 310-8512, Japan*

⁴*Division of Liberal Arts, Kogakuin University, 1-24-2 Nishi-Shinjyuku, Shinjyuku-ku, Tokyo 163-8677, Japan*

⁵*Leiden Observatory, Leiden University, P.O. Box 9513, NL-2300 RA Leiden, The Netherlands*

⁶*Department of Earth and Planetary Sciences, Tokyo Institute of Technology, 2-12-1 Oh-okayama, Meguro-ku, Tokyo 152-8551, Japan*

Accepted XXX. Received YYY; in original form ZZZ

ABSTRACT

This study proposes an analytical framework for deriving the surface brightness profile and geometry of a geometrically-thin axisymmetric disc from interferometric observation of continuum emission. Such precise modelling facilitates the exploration of faint non-axisymmetric structures, such as spirals and circumplanetary discs. As a demonstration, we simulate interferometric observations of geometrically-thin axisymmetric discs. The proposed method can reasonably recover the injected axisymmetric structures, whereas Gaussian fitting of the same data yielded larger errors in disc orientation estimation. To further test the applicability of the method, it was applied to the mock data for $m = 1, 2$ spirals and a point source, which are embedded in a bright axisymmetric structure. The injected non-axisymmetric structures were reasonably recovered except for the innermost parts, and the disc geometric parameter estimations were better than Gaussian fitting. The method was then applied to the real data of Elias 20 and AS 209, and it adequately subtracted the axisymmetric component, notably in Elias 20, where substantial residuals remained without our method. We also applied our method to continuum data of PDS 70 to demonstrate the effectiveness of the method. We successfully recovered emission from PDS 70 c consistently with previous studies, and also tentatively discovered new substructures. The current formulation can be applied to any data for disc continuum emission, and aids in the search of spirals and circumplanetary discs, whose detection is still limited.

Key words: radio continuum: planetary systems – methods: data analysis – protoplanetary discs

1 INTRODUCTION

High-resolution imaging of proto-planetary discs by the Atacama Large Millimeter/submillimeter Array (ALMA) has revolutionized our view of planet formation. The spatial resolutions, down to $\sim 0.01''$, have enabled the identification of rings and gap structures in the disc continuum emission (ALMA Partnership et al. 2015; Andrews et al. 2018; Huang et al. 2018a). The mechanisms for invoking such annular structures have been under debate, and multiple explanations have been provided, such as a disc-planet interaction (Lin & Papaloizou 1979; Goldreich & Tremaine 1980; Dong et al. 2018), snowline (Zhang et al. 2015), sintering (Okuzumi et al. 2016), secular gravitational instability (Takahashi & Inutsuka 2014), and non-ideal MHD effects (Flock et al. 2015). Nevertheless, the disc-planet interaction is independently supported by recent discoveries of kinematic signals of embedded planets (Pinte et al. 2018, 2020, 2023). Here, the signals, usually referred to as velocity “kinks”, appear as localized deviations from the Keplerian motion of discs in velocity maps, and there have been dozens of such detections.

In contrast to multiple detections of rings, gaps, and kinks, the

direct detection of embedded planets and the circumplanetary discs within gaps remains limited. Currently, there are a few cases presenting the robust detection of embedded planets in gaps, such as, PDS 70 (Keppler et al. 2018; Wagner et al. 2018; Benisty et al. 2021), AB Aur (Currie et al. 2022), HD 169142 (Reggiani et al. 2014; Hammond et al. 2023), and MWC 758 (Wagner et al. 2023). These embedded planets have been somewhat selectively identified around (pre-)transitional discs with the ages greater than 4 Myr. In addition, Andrews et al. (2021) searched for circumplanetary discs in gaps of DSHARP discs (Andrews et al. 2018), whose typical age is 1 Myr; however, no compelling case remains.

Another supportive evidence for the planetary hypothesis can be obtained from a planetary spiral, which is excited by planetary gravity. However, despite the discovery of many spirals in discs (Muto et al. 2012; Grady et al. 2013; Huang et al. 2018b), only the spirals in systems with detected embedded planets are highly likely to be produced by the planetary gravity. The other spirals can also originate from non-planetary mechanisms, for example, gravity from stellar companions, gravitational instability or stellar flyby. Indeed, the spiral in HD 100453 A is a notable example, where its M-dwarf companion is thought to be the driver for the structure (Wagner et al. 2015; Dong et al. 2016).

* E-mail: masataka.aizawa.sw35@vc.ibaraki.ac.jp

Recently, [Speedie & Dong \(2022\)](#) searched for planetary spirals in 10 discs with kinematic signatures of embedded planets; however, they did not find any conclusive case. The limited detection of circumplanetary discs and spirals might indicate that the current sensitivities are still inadequate for most of the cases, or certain annular structures may have been caused via non-planetary mechanisms.

Nevertheless, an observational effort to search for circumplanetary discs and spirals is still essential for testing the planetary hypothesis. However, one obstacle to their detection is that these signals appear as small perturbations to the bright disc emission. This is similar to the situation of searching for a needle in a haystack. Thus, the precise modelling of the background disc emission is necessary to subtract its contribution.

One reasonable approximation for the background emission is an axisymmetric disc. However, the modelling of a vertically extended axisymmetric disc is difficult because of the geometric effect (e.g., shadows) and additional parameters to be solved (e.g., radial disc heights). Such geometric effects are significant at optical to infrared wavelengths, because of their sensitivity to vertically extended small dust grains. On the other hand, the disc emission at the radio band is well approximated by a geometrically thin disc model with no vertical structure owing to its sensitivity to large grains, which are more settled to the disc midplane. This approximation is to a certain extent justified by the observations of many clear concentric structures of discs in radio continuum emissions (e.g., [ALMA Partnership et al. 2015](#); [Andrews et al. 2018](#)).

Recently, assuming a geometrically thin and axisymmetric disc, [Jennings et al. \(2020\)](#) proposed an algorithm for deriving a radial surface brightness profile of continuum emission in the radio band. Their method can resolve annular structures finer than the standard imaging technique, CLEAN, using simulated and real data ([Jennings et al. 2020, 2022a,b](#)). Using their method, [Andrews et al. \(2021\)](#) also searched for circumplanetary disc emission from DSHARP data by subtracting the axisymmetric structures of discs.

Although the method proposed in [Jennings et al. \(2020\)](#) can reasonably recover the brightness profile of the disc, there is still room for improvement. Specifically, in their modelling, the geometric parameters of the axisymmetric model, such as an orientation and a central position of a disc, and hyperparameters for the Gaussian Process model that is used to prevent overfitting must be fixed. This limitation can be problematic, because a slight change in geometric parameters of the disc model, for example, several mas in a central position, can introduce the apparent non-axisymmetric structures in the image with their axisymmetric structure being subtracted ([Andrews et al. 2021](#)). Such false structures can be degenerate with the real structures, and may disrupt the detection of spirals and circumplanetary discs. Till date, in the frameworks proposed for axisymmetric disc models, these geometric parameters or hyperparameters for regularization have been manually tuned ([Jennings et al. 2020](#); [Andrews et al. 2021](#)), or estimated through parameterized models ([Zhang et al. 2015](#); [Tazzari et al. 2017](#); [Jennings et al. 2020](#); [Kanagawa et al. 2021](#)) and image-based analysis ([Huang et al. 2018a](#)).

Therefore, this study proposed an analytical framework to derive all of the parameters for a geometrically-thin axisymmetric disc and hyperparameters for the Gaussian Process kernel from observations assuming a geometrically thin disc. This study employed the formulation for the inverse modelling presented in [Kawahara & Masuda \(2020\)](#), which recovered a planetary surface map and its spin and orbital geometry from planetary reflection light. Interestingly, the problem considered in that study is mathematically similar to the current problem. Thus, we can apply their methodology for the current problem. We demonstrate the feasibility of the current method

via its application to mock and real data. In addition, we perform injection and recovery experiments for circumplanetary discs and spirals, and discuss the applicability and limitations of the method.

The remainder of this paper is organized as follows. In [Sec 2](#), the method for estimating an axisymmetric structure from visibilities is formulated. In [Sec 3](#), the methodology for studying non-axisymmetric features in a residual image is discussed. In [Sec 4](#), the feasibility of the proposed method for mock observations is investigated. In [Sec 5](#), circumplanetary discs and spirals are injected to mock data, and the ability to recover the structures is examined. In [Sec 6](#), our method was also applied to the real data for Elias 20 and AS 209. [Sec 7](#) presents an analysis of the PDS 70 data, demonstrating the recovery of circumplanetary emission. Finally, the conclusions and future improvements are presented in [Sec 8](#).

2 ANALYTICAL FORMULATION FOR PARAMETERS FOR A GEOMETRICALLY THIN AXISYMMETRIC DISC

2.1 Model

We model a geometrically thin axisymmetric disc, whose parameters are composed of the brightness profile \mathbf{a} and its geometry \mathbf{g} . Here, $\mathbf{a} = \{I(r_k)\}$ for $k = 1, 2, \dots, N$ is a vector obtained by discretizing the radial surface brightness profile $I(r)$, and r_k is the k -th collocation point of the Fourier-Bessel series:

$$r_k = R_{\text{out}} \frac{j_{0k}}{j_{0N}}, \quad (1)$$

where R_{out} is the outer boundary of $I(r)$ with $I(r > R_{\text{out}}) = 0$, and j_{0k} is the k -th root of the zero-order Bessel function of the first kind $J_0(r)$; $J_0(j_{0k}) = 0$.

The disc geometry is specified by the disc orientation and the central position: $\mathbf{g} = (\Delta x_{\text{cen}}, \Delta y_{\text{cen}}, \cos i, \text{PA})$. Here, the disc centre $(\Delta x_{\text{cen}}, \Delta y_{\text{cen}})$ is assumed to be shifted from a phase centre of the observation, and the disc orientation is specified by the position angle PA and inclination i . PA is defined as the angle of the major axis measured counter-clockwise from the north direction, and i is defined as the angle between the line of sight and the axis normal to the disc plane. Further, the formulation adopts $\cos i$, which represents the aspect ratio for the disc ellipse.

2.2 Visibilities for axisymmetric disc

2.2.1 Face-on disc

As the simplest problem, we consider the face-on disc with no positional offset from the phase centre; $\mathbf{g} = (\Delta x_{\text{cen}} = 0'', \Delta y_{\text{cen}} = 0'', \cos i = 1, \text{PA} = 0)$. Visibilities for an azimuthally symmetric geometrically thin disc with a brightness profile $I(r)$ are expressed as the Hankel transformation ([Thompson et al. 2017](#); [Jennings et al. 2020](#)):

$$\begin{aligned} V(u, v) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} I(x, y) \exp(-2\pi j(ux + vy)) \mathbf{d}\mathbf{x} \quad (2) \\ &= 2\pi \int_0^{\infty} J_0(2\pi r q_{\cos i=1}) I(r) r dr, \quad (3) \end{aligned}$$

where $V(u, v)$ is the complex visibility at a spatial frequency (u, v) , $q_{\cos i=1} = \sqrt{u^2 + v^2}$ is the deprojected spatial frequency for the face-on disc, and J_0 is the zero-order Bessel function of the first kind.

We assume M observational spatial frequencies $\{\mathbf{u}, \mathbf{v}\}_j = (u_j, v_j)$, where $j = 1, 2, \dots, M$. At these spatial frequencies, we compute model visibilities, denoted by $(\mathbf{V} = \{V_j(u_j, v_j)\})$. Using

equation (3) and assuming \mathbf{a} , the model visibilities \mathbf{V} are expressed as follows (e.g., Jennings et al. 2020):

$$\mathbf{V} = \mathbf{H}\mathbf{a}, \quad (4)$$

where the matrix for the Hankel transformation \mathbf{H} is expressed as follows:

$$\{\mathbf{H}\}_{j,k} = \frac{4\pi R_{\text{out}}^2}{j_0^2(N+1)J_1^2(j_0k)} J_0(2\pi q_{j,\cos i=1} r_k), \quad (5)$$

$$q_{j,\cos i=1} = \sqrt{u_j^2 + v_j^2}, \quad (6)$$

where J_1 is the first-order Bessel function of the first kind.

2.2.2 Inclined disc

The computations of model visibilities can be simply extended to the case of the inclined disc. For the calculation, we prepare three different coordinates (x, y) , (x', y') , and (x'', y'') ; Fig. 1 shows a schematic of the coordinates and disc. Here, (x, y) is the observational coordinate system with the phase centre being the origin, (x', y') is the shifted coordinate system with the disc centre being the origin, and (x'', y'') is the deprojected coordinate system according to (PA, i) . (x, y) and (x', y') are related as $(x, y) = (x' + \Delta x_{\text{cen}}, y' + \Delta y_{\text{cen}})$, and (x', y') and (x'', y'') are converted as follows:

$$\begin{pmatrix} x'' \\ y'' \end{pmatrix} = \begin{pmatrix} 1/\cos i & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \cos(\text{PA}) & -\sin(\text{PA}) \\ \sin(\text{PA}) & \cos(\text{PA}) \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix} \quad (7)$$

The brightness profiles for the coordinates $I(x, y)$, $I'(x', y')$, and $I''(x'', y'')$ are assumed to be invariant to coordinate transformations:

$$I(x, y) = I'(x', y') = I''(x'', y''). \quad (8)$$

This assumption is a mathematical hypothesis for modeling purposes, and it is not necessarily consistent with the physical picture of what would actually occur if the disc were observed from the face-on view¹.

Assuming the brightness profile in a deprojected frame $I''(r'')$ with $r'' \equiv \sqrt{x''^2 + y''^2}$, we aimed to compute the model visibilities. First, we transform equation (2) as follows:

$$\begin{aligned} V(u, v) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} I(x, y) \exp(-2\pi j(ux + vy)) \mathbf{d}\mathbf{x} \\ &= \exp(-2\pi j(u\Delta x_{\text{cen}} + v\Delta y_{\text{cen}})) \\ &\quad \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} I'(x', y') \exp(-2\pi j(ux' + vy')) \mathbf{d}\mathbf{x}', \\ &= \exp(-2\pi j(u\Delta x_{\text{cen}} + v\Delta y_{\text{cen}})) \\ &\quad \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} I''(x'', y'') \exp(-2\pi j(u'x'' + v'y'')) \\ &\quad |\cos i| \mathbf{d}\mathbf{x}'', \end{aligned} \quad (9)$$

where we define (u', v') as follows:

$$\begin{pmatrix} u' \\ v' \end{pmatrix} = \begin{pmatrix} \cos i & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \cos(\text{PA}) & -\sin(\text{PA}) \\ \sin(\text{PA}) & \cos(\text{PA}) \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix}. \quad (10)$$

Using equation (3), equation (9) is reduced as follows:

$$\begin{aligned} V(u, v) &= 2\pi |\cos i| \exp(-2\pi j(u\Delta x_{\text{cen}} + v\Delta y_{\text{cen}})) \\ &\quad \times \int_0^{\infty} J_0(2\pi r'' q) I''(r'') r'' dr'', \end{aligned} \quad (11)$$

where we define a deprojected radial spatial frequency q as follows:

$$q = \sqrt{u'^2 + v'^2}. \quad (12)$$

Note that $q_{\cos i=1}$ is the special case of q with $(\cos i, \text{PA}) = (1, 0)$. In the calculation, we assume an axisymmetric disc, whose brightness profile with $I''(x'', y'') = I''(r'')$, and we discretize $I''(r'')$ using \mathbf{a} in the same manner as that presented in Sec 2.2.1.

For \mathbf{V}_{real} and \mathbf{V}_{imag} , which are defined as real and imaginary parts of \mathbf{V} , respectively, we derive a linear equation as follows:

$$\begin{pmatrix} \mathbf{V}_{\text{real}} \\ \mathbf{V}_{\text{imag}} \end{pmatrix} = \begin{pmatrix} \mathbf{C}_{\text{real}} \mathbf{H}' \\ \mathbf{C}_{\text{imag}} \mathbf{H}' \end{pmatrix} \mathbf{a}, \quad (13)$$

where \mathbf{H}' , \mathbf{C}_{real} , and \mathbf{C}_{imag} are defined as follows:

$$\{\mathbf{H}'\}_{j,k} = \frac{4\pi R_{\text{out}}^2}{j_0^2(N+1)J_1^2(j_0k)} J_0(2\pi q_j r_k), \quad (14)$$

$$\{\mathbf{C}_{\text{real}}\}_{j,k} = |\cos i| \cos(-2\pi(\Delta x_{\text{cen}} u_j + \Delta y_{\text{cen}} v_j)) \delta_{j,k}, \quad (15)$$

$$\{\mathbf{C}_{\text{imag}}\}_{j,k} = |\cos i| \sin(-2\pi(\Delta x_{\text{cen}} u_j + \Delta y_{\text{cen}} v_j)) \delta_{j,k}. \quad (16)$$

2.3 Inverse modelling

On the basis of the forward modelling in Sec 2.2, we consider inverse modelling for brightness profile \mathbf{a} and geometric parameters $\mathbf{g} = (\Delta x_{\text{cen}}, \Delta y_{\text{cen}}, \cos i, \text{PA})$ from visibilities.

2.3.1 Data with Gaussian noise

Let us assume that there are M visibilities $\mathbf{d}_{\text{obs}} = \{d_{\text{obs},i}(u_i, v_i)\}$ ($i = 1, 2, \dots, M$), with the noise of each visibility data obtained via the standard deviation of $\sigma_{\text{obs}} = \{\sigma_{\text{obs},i}\}$. We separate the real and imaginary parts of the visibilities as $\mathbf{d}_{\text{real}} \equiv \Re(\mathbf{d}_{\text{obs}})$ and $\mathbf{d}_{\text{imag}} \equiv \Im(\mathbf{d}_{\text{obs}})$, respectively. We assume that observational noises for \mathbf{d}_{real} and \mathbf{d}_{imag} obey the multivariate normal distribution. Here, the multivariate normal distribution for a N -dimensional random vector \mathbf{x} is expressed as follows:

$$\mathcal{N}(\mathbf{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{(2\pi)^{N/2} |\det \boldsymbol{\Sigma}|^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu})\right), \quad (17)$$

where $\boldsymbol{\mu}$ is the mean vector and $\boldsymbol{\Sigma}$ is the covariance matrix. For \mathbf{d}_{real} and \mathbf{d}_{imag} , we assume the covariance matrix to be $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_{\text{obs}} \equiv \{\delta_{i,j}/\sigma_i^2\}$.

2.3.2 Gaussian prior on brightness profile

In interferometric observations, a brightness profile \mathbf{a} is generally under-constrained. If it is directly obtained by solving a linear equation, for example, equation (4), the solution can diverge, leading to overfitting. This can be avoided by introducing regularization in optimization. Specifically, we assume the Gaussian Process kernel with a set of hyperparameters $\boldsymbol{\theta} = (\alpha, \gamma)$ on \mathbf{a} as follows:

$$p(\mathbf{a}|\boldsymbol{\theta}) = \mathcal{N}(\mathbf{a}|\mathbf{0}, \boldsymbol{\Sigma}_a(\boldsymbol{\theta})), \quad (18)$$

where we adopt the radial basis function (RBF) kernel $\boldsymbol{\Sigma}_a(\boldsymbol{\theta})$, which is expressed as;

$$(\boldsymbol{\Sigma}_a(\boldsymbol{\theta}))_{j,k} = \alpha k_{\text{RBF}}(|r_j - r_k|; \gamma) + \alpha \epsilon \delta_{jk}, \quad (19)$$

$$k_{\text{RBF}}(|r_j - r_k|; \gamma) = \exp\left(-\frac{\|r_j - r_k\|^2}{2\gamma^2}\right), \quad (20)$$

¹ Nevertheless, the assumption is physically consistent when the disc is a geometrically-thin, optically thick disc.

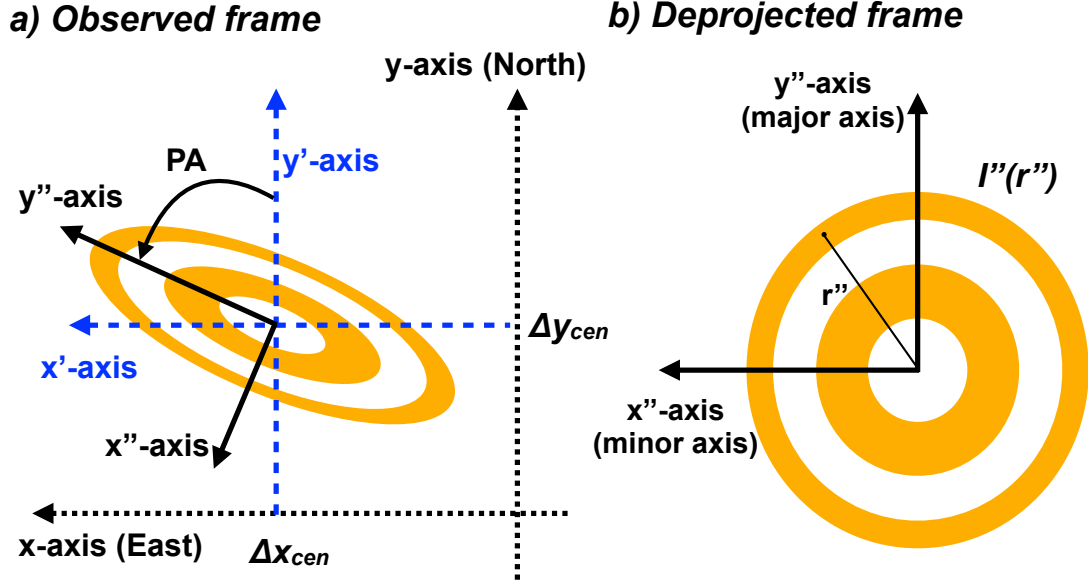


Figure 1. Schematic of a disc in observed (a) and deprojected frames (b). The disc is inclined with $\cos i$ and oriented by PA with respect to the x -axis in the projected frame. The origin of (x, y) is set to be the phase centre, and (x', y') is shifted from (x, y) by $(\Delta x_{cen}, \Delta y_{cen})$. (x'', y'') is the coordinate for the deprojected frame, and the brightness profile is assumed to be only radially dependent on $I''(r'')$.

where α determines the relative weight of the prior information, and γ determines the spatial scale for regularization. The second term in Eq (19) is the small identity matrix, which stabilizes the computation of the inverse matrix $\Sigma_a(\theta)$ (Kawahara et al. 2022). We adopted $\epsilon = 10^{-8}$ in this paper.

This prior tends to promote smooth solutions, and can aid in preventing overfitting. A previous study by Jennings et al. (2020) adopted parameters for controlling the Gaussian Process prior on powers of the brightness profiles in the frequency domain. However, in this study, we adopted the parameters for regulating the profile in the spatial domain to render the discussion simpler.

2.3.3 Likelihood function for geometry and hyperparameters

Following Kawahara & Masuda (2020), we derive the likelihood function $p(\mathbf{d}|\mathbf{g}, \theta)$. We briefly overview their formulation and apply it to the current problem. For further details, see equations (14)-(21) and Appendix A.1-A.3 in the previous study.

We introduce the data vector \mathbf{d} as follows:

$$\mathbf{d} = \begin{pmatrix} \mathbf{d}_{\text{real}} \\ \mathbf{d}_{\text{imag}} \end{pmatrix} \quad (21)$$

Recalling Bayes' theorem, we obtain

$$p(\mathbf{a}|\mathbf{d}, \mathbf{g}, \theta)p(\mathbf{d}|\mathbf{g}, \theta) = p(\mathbf{d}|\mathbf{a}, \mathbf{g}, \theta)p(\mathbf{a}|\mathbf{g}, \theta). \quad (22)$$

This can be rewritten as follows:

$$p(\mathbf{d}|\mathbf{g}, \theta) = \frac{p(\mathbf{d}|\mathbf{a}, \mathbf{g}, \theta)p(\mathbf{a}|\mathbf{g}, \theta)}{p(\mathbf{a}|\mathbf{d}, \mathbf{g}, \theta)} = \frac{p(\mathbf{d}|\mathbf{a}, \mathbf{g})p(\mathbf{a}|\theta)}{p(\mathbf{a}|\mathbf{d}, \mathbf{g}, \theta)}, \quad (23)$$

where we use $p(\mathbf{d}|\mathbf{a}, \mathbf{g}, \theta) = p(\mathbf{d}|\mathbf{a}, \mathbf{g})$ and $p(\mathbf{a}|\mathbf{g}, \theta) = p(\mathbf{a}|\theta)$.

In the equation, $p(\mathbf{d}|\mathbf{a}, \mathbf{g})$ is calculated using equation (13):

$$\begin{aligned} p(\mathbf{d}|\mathbf{a}, \mathbf{g}) &= p(\mathbf{d}_{\text{real}}|\mathbf{a}, \mathbf{g})p(\mathbf{d}_{\text{imag}}|\mathbf{a}, \mathbf{g}) \\ &= \mathcal{N}(\mathbf{d}_{\text{real}}|\mathbf{C}_{\text{real}}\mathbf{H}'\mathbf{a}, \Sigma_{\text{obs}})\mathcal{N}(\mathbf{d}_{\text{imag}}|\mathbf{C}_{\text{imag}}\mathbf{H}'\mathbf{a}, \Sigma_{\text{obs}}) \\ &= \mathcal{N}(\mathbf{d}|\bar{\mathbf{H}}\mathbf{a}, \bar{\Sigma}_{\mathbf{d}}), \end{aligned} \quad (24)$$

where $\bar{\Sigma}$ and $\bar{\mathbf{H}}$ are defined as follows:

$$\bar{\mathbf{H}} = \begin{pmatrix} \mathbf{C}_{\text{real}}\mathbf{H}' \\ \mathbf{C}_{\text{imag}}\mathbf{H}' \end{pmatrix}, \quad (25)$$

$$\bar{\Sigma}_{\mathbf{d}} = \begin{pmatrix} \Sigma_{\text{obs}} & \mathbf{O} \\ \mathbf{O} & \Sigma_{\text{obs}} \end{pmatrix}. \quad (26)$$

Further, using equations (18) and (24), we obtain

$$p(\mathbf{a}|\mathbf{d}, \mathbf{g}, \theta) \propto p(\mathbf{d}|\mathbf{a}, \mathbf{g})p(\mathbf{a}|\theta) = \mathcal{N}(\mathbf{d}|\bar{\mathbf{H}}\mathbf{a}, \bar{\Sigma}_{\mathbf{d}})\mathcal{N}(\mathbf{a}|\mathbf{0}, \Sigma_a(\theta)). \quad (27)$$

As \mathbf{a} appears as a quadratic form in $p(\mathbf{a}|\mathbf{d}, \mathbf{g}, \theta)$ in the exponent, we can obtain $p(\mathbf{a}|\mathbf{d}, \mathbf{g}, \theta)$ as follows:

$$p(\mathbf{a}|\mathbf{d}, \mathbf{g}, \theta) = \mathcal{N}(\mathbf{a}|\bar{\mu}_{\mathbf{a}|\mathbf{d}}, \bar{\Sigma}_{\mathbf{a}|\mathbf{d}}), \quad (28)$$

where we define

$$\bar{\mu}_{\mathbf{a}|\mathbf{d}} = (\bar{\mathbf{H}}^T \bar{\Sigma}_{\mathbf{d}}^{-1} \bar{\mathbf{H}} + \Sigma_a(\theta)^{-1})^{-1} \bar{\mathbf{H}}^T \bar{\Sigma}_{\mathbf{d}}^{-1} \mathbf{d}^T, \quad (29)$$

$$\bar{\Sigma}_{\mathbf{a}|\mathbf{d}} = (\bar{\mathbf{H}}^T \bar{\Sigma}_{\mathbf{d}}^{-1} \bar{\mathbf{H}} + \Sigma_a(\theta)^{-1})^{-1}. \quad (30)$$

Using equations (23), (24), and (28), we obtain

$$p(\mathbf{d}|\mathbf{g}, \theta) \propto \mathcal{N}(\mathbf{d}|\bar{\mathbf{H}}\bar{\mu}_{\mathbf{a}|\mathbf{d}}, \bar{\Sigma}_{\mathbf{d}})/\mathcal{N}(\mathbf{a}|\bar{\mu}_{\mathbf{a}|\mathbf{d}}, \bar{\Sigma}_{\mathbf{a}|\mathbf{d}}). \quad (31)$$

Again, \mathbf{d} appears as a quadratic form in $p(\mathbf{d}|\mathbf{g}, \theta)$ in its exponent. Thus, we can obtain the likelihood function $p(\mathbf{d}|\mathbf{g}, \theta)$ as follows:

$$p(\mathbf{d}|\mathbf{g}, \theta) = \mathcal{N}(\mathbf{d}|\mathbf{0}, \bar{\Sigma}_{\mathbf{d}} + \bar{\mathbf{H}}\Sigma_a\bar{\mathbf{H}}^T). \quad (32)$$

2.3.4 Posterior distribution for geometry and hyperparameters

The posterior distribution for $p(\mathbf{g}, \theta|\mathbf{d})$ is given by the Bayes' theorem as follows:

$$p(\mathbf{g}, \theta|\mathbf{d}) \propto p(\mathbf{d}|\mathbf{g}, \theta)p(\mathbf{g}, \theta) = \mathcal{N}(\mathbf{d}|\mathbf{0}, \bar{\Sigma}_{\mathbf{d}} + \bar{\mathbf{H}}\Sigma_a\bar{\mathbf{H}}^T)p(\mathbf{g}, \theta), \quad (33)$$

Based on $p(\mathbf{g}, \theta|\mathbf{d})$, we draw samples for (\mathbf{g}, θ) using an MCMC sampler. The prior distribution $p(\mathbf{g}, \theta)$ is assumed as follows. We assume uniform distributions $\mathcal{U}(0, 1)$ and $\mathcal{U}(0, \pi)$ for $\cos i$ and PA,

respectively. Further, we consider $\mathcal{U}(-1'', 1'')$ for Δx_{cen} and Δy_{cen} . In addition, we assume the log-uniform prior from 10^{-4} to 10^4 for α and a uniform prior $\mathcal{U}(0.01'', 0.15'')$ for γ . We checked that the likelihood is exceedingly small when α and γ fall outside their respective prior ranges, thus confirming the sufficient broadness of the prior. As the computation for $p(\mathbf{d}|\mathbf{g}, \boldsymbol{\theta})$ can be time-consuming, we transform the equation to a more efficient form (further details in Appendix A).

Moreover, the computation also relies on the number of data points. In Appendix B, we introduce and discuss the use of data binning to reduce the computation. Specifically, we prepare two-dimensional linear and log grids, apply them to the simulated visibilities, and compare the binning errors. The binning with the log grid was found to yield overall low errors. A more detailed discussion is presented in Appendix B. In the following analysis, we simply adopt the log grid for the analysis.

2.3.5 Posterior distribution for all of parameters

We can draw samples for $\{\mathbf{a}, \mathbf{g}, \boldsymbol{\theta}\}$ by using the conditional formula for the joint probability: $p(\mathbf{a}, \mathbf{g}, \boldsymbol{\theta}|\mathbf{d}) = p(\mathbf{a}|\mathbf{d}, \mathbf{g}, \boldsymbol{\theta})p(\mathbf{g}, \boldsymbol{\theta}|\mathbf{d})$. Specifically, we can draw a sample $(\mathbf{g}^\dagger, \boldsymbol{\theta}^\dagger)$ from $p(\mathbf{g}, \boldsymbol{\theta}|\mathbf{d})$ as done in Sec 2.3.4, and subsequently take a sample for \mathbf{a} from $p(\mathbf{a}|\mathbf{d}, \mathbf{g}^\dagger, \boldsymbol{\theta}^\dagger)$ using equation (28). We can iterate this procedure to make samples, and this sampling is indeed equivalent to drawing sample from the joint distribution $p(\mathbf{a}, \mathbf{g}, \boldsymbol{\theta}|\mathbf{d})$. Using the samples for $\{\mathbf{a}, \mathbf{g}, \boldsymbol{\theta}\}$, we can also compute statistics for parameters (e.g., mean and standard deviation for \mathbf{a}).

2.3.6 Difference in formulation between current study and frank

We here highlight the key differences between our approach and that of `frank`. Note that the direct comparison of the recovered profiles is also given in Section 6. While both methods vary in their regularization strategies, the distinct difference is that our approach solve all parameters, including brightness profiles, geometric parameters, and hyperparameters, whereas `frank` is limited to solving brightness profiles. In this sense, our formulation can be seen as a natural extension of that of `frank`.

One advantage of our method is that it can optimize geometric parameters and hyperparameters directly from the data, and thus it is less susceptible to human biases caused by manual tuning. Moreover, while the current paper is limited to the simplest model with single frequency and single source, the methodology itself can be easily extended to more complex problems with multi-frequency data or multiple sources. In such cases, manual tuning of optimal parameters can be both challenging and less reliable, making our approach more effective.

On the other hand, the disadvantage of our approach is that it is more computationally demanding than `frank`. This is because we attempted to solve all parameters, including non-linear parameters, rather than deriving only brightness profiles like `frank`. Furthermore, our current formulation cannot support imposing a non-negative condition on brightness profiles, a feature available in `frank`. This is because the analytical expression for the marginalization over \mathbf{a} becomes inapplicable under the non-negative condition. Although we can implement the condition by considering the same problem as `frank`, the current study prioritizes simultaneous fitting of all parameters, so we do not implement the function.

3 EXTRACT ASYMMETRIC FEATURES

This section summarizes the process of extracting and quantifying the non-axisymmetric components from observations using the derived parameters in Sec 2.3. The method is applied to data analyses in Sec 4 and 5.

3.1 Making residual images in observed and deprojected frames

3.1.1 Making residual image

After subtracting the axisymmetric model from visibilities, we expect the residual visibilities to contain only non-axisymmetric components. Therefore, imaging using these residual visibilities can reveal the non-axisymmetric component of the disc (Jennings et al. 2020; Andrews et al. 2021). Based on this principle, we created images to study the non-axisymmetric components.

Drawing samples from a posterior distribution as described in Sec 2, we selected the maximum a posteriori probability (MAP) estimate for \mathbf{g} and $\boldsymbol{\theta}$. Subsequently, we randomly drew a brightness profile using equation (28). For the chosen parameters and the brightness profile, we computed the model visibilities of an axisymmetric disc, and subtracted them from observed visibilities. Then, the phase centres of measurement sets were aligned with the disc centres using `fixvis` in CASA.

For the shifted and subtracted measurement set, we created images using the CLEAN algorithm, which is implemented as `tclean` in CASA. In this study, the pixel scale for the image was set as 6 mas, and the image size was 1000×1000 pixels. In the imaging process, Although it is typical to adopt 0 iteration for CLEAN in making the residual map (e.g., Jennings et al. 2020), we found that employing non-zero iterations for CLEAN result in slight or moderate improvement in the quality of the residual image, especially when there remain substantial residuals. Thus, in this paper, we implemented thresholds with `nsigma= 3.5`, which stops the iterations if the maximum residual on the image is below 3.5 times the root-mean-square error. Further, we adopted the Briggs weighting with a robust parameter of 0.5 and without UV-taper. Appendix C presents a discussion on the effects of changing robust parameters on the recovered residual images, and 0.5 was determined to be the balanced choice. The brightness of the output image from CASA was measured in Jy beam^{-1} , where ‘beam’ is the nominal area where the brightness is defined². The standard deviation of the residual image was also measured outside the disc emission.

3.1.2 Deprojection of residual image

In addition to the residual image in an observational frame, we created the deprojected residual image using the disc geometry. Spatial frequencies of the shifted measurement set were converted using the geometric parameter, $(u, v) \rightarrow (u', v')$ in equation (10). Consequently, the image was created in the same manner as Sec 3.1.1. This operation aligns the major axis of the image with the y -axis (north direction), and stretches the image along the x axis (east direction). This corresponds to the conversion in Figure 1 from a deprojected to a projected frame.

² In CASA, the synthesized beam is approximated using the Gaussian function, and its area is expressed as $\frac{\pi}{4\ln 2} b_{\text{maj}} b_{\text{min}}$, where $(b_{\text{maj}} b_{\text{min}})$ are the full widths at half maximum of the Gaussian in major and minor axes.

In the imaging process, we should be cautious that the transformation $(u, v) \rightarrow (u', v')$ apparently decreases the brightness in Jy sr^{-1} by a factor of $\cos i$, while brightness in Jy beam^{-1} remains unchanged before and after deprojection. This is because the transformation increases the disc area (and beam) while the total flux is unchanged. Throughout the paper, we consistently presented the brightness of images in Jy beam^{-1} , which remains constant, and we did not apply any correction by a factor of $\cos i$.

Although widely used, the current simple deprojection methods provide the true image of the disc viewed face-on only in the case of an infinitesimally thin disc. The other geometric effects such as the vertical structures or shadows cannot be incorporated with our simple operations.

3.2 Mode decomposition of residual image in polar direction

Spirals are often characterized by dominant modes in a polar direction. We can decompose $I(r, \phi)$ in a polar coordinate (r, ϕ) as follows (e.g., [Binney & Tremaine 2008](#)):

$$I(r, \phi) = \sum_{m=0} I_m(r) \cos(m(\phi - \phi_m(r))). \quad (34)$$

where $I_m(r)$ is the amplitude at the m -th mode, and $\phi_m(r)$ is a phase shift for the m -th component. Note that $I_0(r)$ corresponds to the brightness profile. The amplitudes and phases can be derived from the residual images $I(x, y)$ as follows:

$$I_m(r) = 2\sqrt{C_m^2 + S_m^2}, \quad (35)$$

$$\phi_m(r) = \text{atan2}(S_m, C_m), \quad (36)$$

where atan2 is the 2-argument arctangent, and (C_m, S_m) is defined as follows:

$$C_m \equiv \frac{1}{2\pi} \int_0^{2\pi} I(r, \phi) \cos(m\phi) d\phi, \quad (37)$$

$$S_m \equiv \frac{1}{2\pi} \int_0^{2\pi} I(r, \phi) \sin(m\phi) d\phi. \quad (38)$$

The values of (C_m, S_m) can be computed for the real data. In this study, we computed them by interpolating a residual image.

3.3 Extraction of odd and even symmetric components of images

Odd-symmetric and even-symmetric components of an image can be decomposed by using either of each imaginary or real part of the data. We start with decomposing an image into even- and odd-symmetric components.

$$I(x, y) = I_{\text{even}}(x, y) + I_{\text{odd}}(x, y), \quad (39)$$

where we define

$$I_{\text{even}}(x, y) = \frac{1}{2}(I(x, y) + I(-x, -y)) \quad (40)$$

$$= \sum_{m=\text{even}} I_m(r) \cos(m(\phi - \phi_m(r))), \quad (41)$$

$$I_{\text{odd}}(x, y) = \frac{1}{2}(I(x, y) - I(-x, -y)) \quad (42)$$

$$= \sum_{m=\text{odd}} I_m(r) \cos(m(\phi - \phi_m(r))), \quad (43)$$

which satisfy $I_{\text{even}}(x, y) = I_{\text{even}}(-x, -y)$, and $I_{\text{odd}}(x, y) = -I_{\text{odd}}(-x, -y)$. However, using equation (2), the real and imaginary

parts are connected to $I_{\text{even}}(x, y)$ and $I_{\text{odd}}(x, y)$ as follows:

$$\Re(V(u, v)) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} I_{\text{even}}(x, y) \exp(-2\pi j(ux + vy)) dx, \quad (44)$$

$$\Im(V(u, v)) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} I_{\text{odd}}(x, y) \exp(-2\pi j(ux + vy)) dx. \quad (45)$$

Thus, we can create $I_{\text{even}}(x, y)$ and $I_{\text{odd}}(x, y)$ using either the real or imaginary part of the data. Notably, previous studies have already used the imaginary part of the data to extract an odd-symmetric component ([Hashimoto et al. 2021](#); [Kanagawa et al. 2021](#)).

4 APPLICATION TO SIMULATED DATA FOR GEOMETRICALLY THIN AXISYMMETRIC DISC

4.1 Injection and recovery test

As a test case, we applied our method to simulated data for an axisymmetric disc to examine the ability to recover input parameters. We used two radial profiles for our test. The continuum brightness profiles for AS 209 and WaOph 6 of DSHARP observations³ ([Andrews et al. 2018](#)) were used. AS 209 is chosen as the most structured disc in a radial direction, and WaOph 6 is among $m = 2$ spiral discs in the DSHARP sample. Subsequently, we assume the model discs have the geometric parameters of $(\Delta x_{\text{cen}}, \Delta y_{\text{cen}}, \cos i, \text{PA},) = (0'', 0'', 0.75, 45^\circ)$. We normalize the brightness profiles to render the total fluxes of the AS 209 and WaOph 6 models as 0.288 and 0.161 Jy, respectively.

To produce mock data with realistic UV-coverage and noise, we incorporated them using the actual data of DSHARP observations. We downloaded the self-calibrated continuum measurement set for AS 209 and WaOph 6. The data were averaged at each spectral window to produce a single channel data. Thereafter, time averaging was applied for 30 s. A new measurement set was created using the averaged data to obtain the UV-coverage and the weight at each spatial frequency. The signals of the model disc at each spatial frequency were calculated via Fourier transforms of the model surface brightness distribution. Consequently, the Gaussian noises were added to the model visibility according to the weights. During the analysis, we used `tb.getcol` in CASA to export the measurement sets to handle them numerically in `python`⁴.

The data weights recorded in the measurement sets were overestimated for the entire DSHARP data. Following Appendix E in [Hashimoto et al. \(2021\)](#), we deprojected the real and imaginary parts of visibilities, and compared the recorded standard deviations and the deviations of observed visibilities from the binned visibilities in the deprojected spatial frequencies. Consequently, we confirmed that the overestimation existed in both the real and imaginary parts of the visibilities, and the degree of overestimation was within $\sim 10\%$ for both. Therefore, we manually reduced the recorded weights by 3.44 and 3.66 for AS 209 and WaOph 6, respectively.

To reduce the computational time, we binned the data using a logarithmically spaced grid with $N_{\text{bin}} = 500$, where the number of grid points was $(2N_{\text{bin}} + 3)^2$. The boundaries of the grids are defined

³ <https://bulk.cv.nrao.edu/alldata/lp/DSHARP>

⁴ We used `tb.getcol` in CASA to export (u, v) from measurement sets, but we found that signs of the outputted arrays for (u, v) were reversed. Specifically, they have opposite signs with respect to those obtained with `plotms`. In the analysis, we simply flipped signs of (u, v) output from `tb.getcol`.

by the parameters (x_{\min}, x_{\max}) , which are introduced in Appendix B. We set (x_{\min}, x_{\max}) to $(10^2\lambda, 10^7\lambda)$ in our analyses. This binning decreased the number of data points by an approximate factor of 10. Appendix B presents the details of data gridding using a log grid, and we show that the choice of $N_{\text{bin}} = 500$ is acceptable.

Using equation (33), we drew samples from the posterior distribution for $(\gamma, \log_{10} \alpha, \Delta x_{\text{cen}}, \Delta y_{\text{cen}}, \cos i, \text{PA})$. In sampling, we used emcee, which implemented the affine-invariant ensemble sampler for Markov Chain Monte Carlo (Foreman-Mackey et al. 2013). The initial parameters for emcee were set to be close to the input parameters. Here, 16 walkers were prepared and then evolved for at least 8,000 steps. For the drawn samples, we discarded the initial 1,250 steps during the burn-in phase. The convergence is predominantly achieved within 2,000-4,000 steps according to the auto-correlation time analysis. Specifically, we have verified that all samples meet the condition $N > 50\tau$, where N represents the step size and τ represents the integrated autocorrelation time for the chain⁵

The total computational time for the single disc was approximately 12 hours using 8 cores. This represents a significantly higher computational demand in contrast to frank's approach. This stems from our strategy to retrieve all parameters in a simultaneous manner, as opposed to frank's approach of fixing parameters except for the brightness profiles.

Figs. 2 and 3 show the recovery and injection test results for the simulated data. The upper-left panels shows the posterior distribution of the parameters for the two simulated cases. All the geometric parameters were reasonably recovered by the current method.

The analyses resulted in two different length parameters; $\gamma \approx 0.04''$ for the simulated case of AS 209, and $\gamma \approx 0.1''$ for the case of WaOph 6. The differences in γ were largely due to the differences in the injected brightness profiles; the injected profile for WaOph 6 was smoother than that of AS 209. The UV-coverage would also affect the length scale, but it remains unclear in the current two cases. The length scale parameter for WaOph 6 was larger than the beam size $(0.091'', 0.037'')$, whereas it was smaller than the beam size $(0.076'', 0.040'')$ for AS 209. Note that the beam sizes reflect the UV-coverage, but they vary with assumed parameters in CLEAN; for example, robust parameter and UV-taper. In Appendix D, we further study the variations in γ by stretching the injected brightness profiles and UV-coverage. In conclusion, the optimal length scales γ appear to be determined by multiple factors, including the brightness profiles and UV-coverage (beam size).

The upper-right panels in Figs. 2 and 3 compare the injected brightness profiles with 10 samples for α using the method presented in Sec 2.3.5. The radial profiles were also adequately recovered by the current method. However, the residuals for the brightness profiles were not perfectly consistent with the zero line for both cases, although an oscillating feature was observed in the radial direction. The oscillating length scales for the observed residuals indeed corresponded to the derived length scales γ , which determined the characteristic scales that could be resolved. Moreover, we also identified the large residuals at the very innermost parts owing to the steep changes in the fluxes in the radial direction. The oscillating features have been also reported in Jennings et al. (2020), which adopted regularization in the frequency domain, and such residuals might be inevitable in the Gaussian Process.

The lower-left panels in Figs. 2 and 3 show the simulated and model visibilities. The model visibilities were computed by choosing the MAP solution for the geometry and hyperparameters, and the

brightness profile was randomly sampled. The residual visibilities are shown in the lower panels. The models were well consistent with the simulated visibilities from the low to high spatial frequencies. However, we observed that the powers of the model visibilities at high frequencies were forcibly suppressed. This cut-off scale was determined by the optimized length scale γ , which determined the scales below which the oscillating features appeared in the brightness profiles. At the low spatial frequencies, particularly at $q < 0.3 \text{ M}\lambda$, we also found oscillating residuals with a length scale of $\sim 0.05 \text{ M}\lambda$. This length scale corresponded to the radius for the outer boundary $R_{\text{out}} = 2''$, which limited the model's applicability to the large-scale structure.

In the lower-right panels in Figs. 2 and 3, we show the residual images, which are produced by the method presented in Sec 3.1. The residual images did not exhibit noticeable features, demonstrating that we reasonably estimated the injected parameters.

4.2 Residual images constructed from shifted geometric parameters

To demonstrate how incorrect determination of geometric parameters affects the resultant residual images, we shifted each geometric parameter before subtracting axisymmetric models from the visibilities. For the simulated case of AS 209, we shifted each geometric parameter from the MAP estimate as follows: $\Delta x_{\text{cen}} = 5 \text{ mas}$, $\Delta y_{\text{cen}} = 5 \text{ mas}$, $\Delta \cos i = 0.01$, $\Delta \text{PA} = 2 \text{ deg}$. For the simulated case of WaOph 6, we shifted the parameters as follows: $\Delta x_{\text{cen}} = 2 \text{ mas}$, $\Delta y_{\text{cen}} = 2 \text{ mas}$, $\Delta \cos i = 0.03$, $\Delta \text{PA} = 3 \text{ deg}$. To simulate the realistic situation as much as possible, we assume these incorrect geometric parameters with the hyperparameters from the MAP solution, and draw a brightness profile from the posterior distribution, which nevertheless gives the profile similar to our best-fit model. With the subtracted visibilities made with the new models, we created the residual images using the CLEAN following the method presented in Sec 3.1. We also created images from unsubtracted visibilities for comparison. For that, we adopted 0.04 mJy as the CLEAN threshold and used the multiscale CLEAN algorithm with scale parameters of $[0, 30, 120, 360, 720, 1440] \text{ mas}$ (Rau & Cornwell 2011).

Fig. 4 and 5 show the unsubtracted and residual images with their geometric parameters being shifted. They are all shown in observed frames. The residual images based on the best-fitting models did not exhibit any noticeable structure. This demonstrates the feasibility of the current method. However, the residual images with shifted geometric parameters exhibited coherent patterns. These residual images obtained from the current study were overall consistent with those presented in Andrews et al. (2021), wherein the dependence of residual images was studied by shifting the geometric parameters. As shown in the figures, the difference in the injected brightness profiles changed the residual images. We highlight two outer rings for AS 209 by black dashed ellipses in Fig. 4, and we find that the ring locations indeed corresponded to the boundaries, where the signs of the residuals were flipped. However, the brightness profile for WaOph 6 was more continuous than AS 209, and the residual images were less structured.

Fig. 6 shows $I_{m=1,2}(r)$ for the residual images in the simulated case for WaOph6. The shifts in the central positions introduced residuals with the $m = 1$ component, whereas those in $\cos i$ and PA introduced the residuals with $m = 2$ component. As discussed in the later sections, these residuals can be degenerate with real signals (e.g., spirals); thus they can introduce biases for estimating such non-axisymmetric structures in discs.

As a comparison with the proposed method, we also applied the

⁵ <https://emcee.readthedocs.io/en/stable/tutorials/autocorr/>.

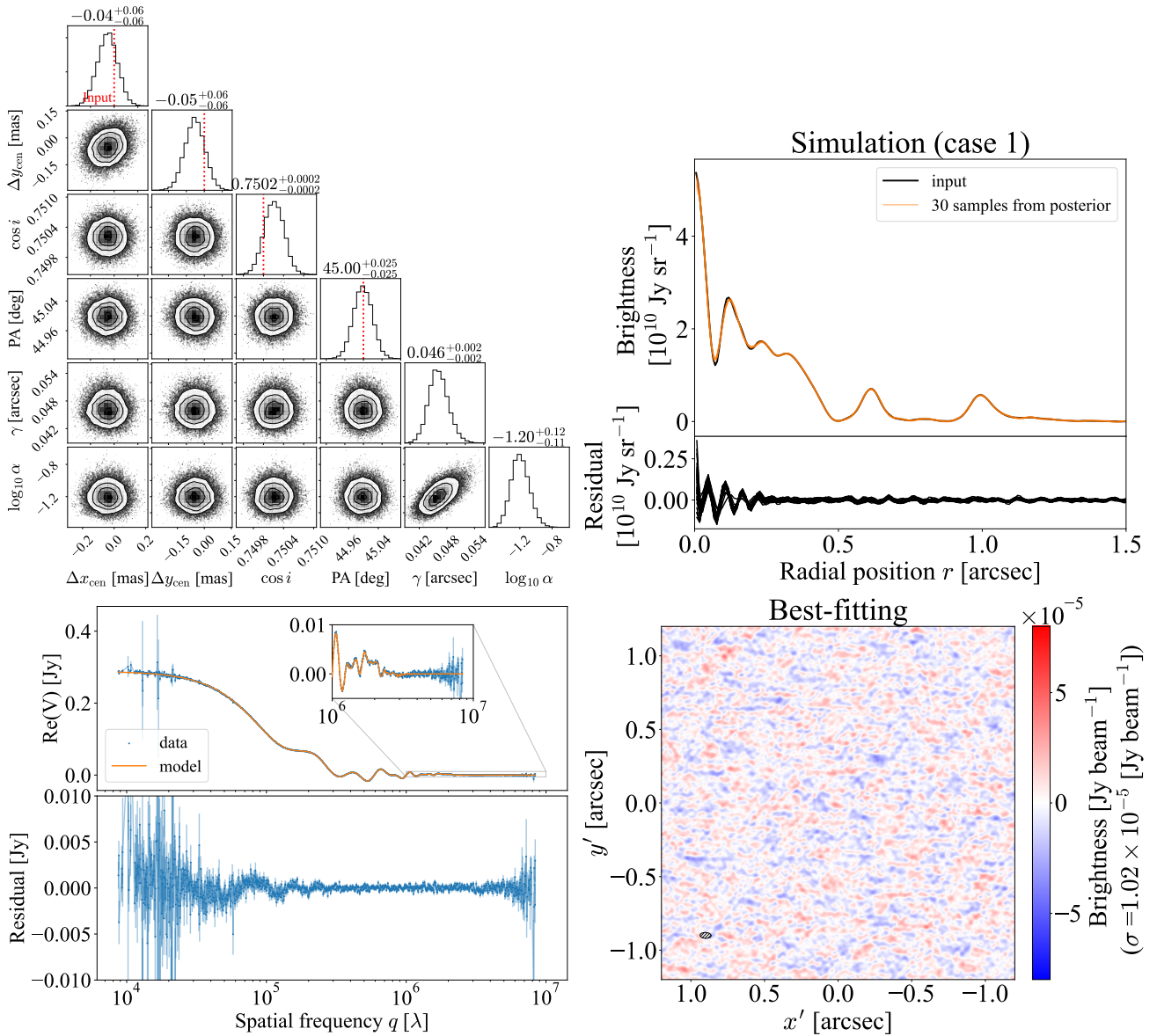


Figure 2. Recovery and injection test for mock data of AS 209. (upper-left) Posterior distribution for $(\Delta x_{\text{cen}}, \Delta y_{\text{cen}}, \cos i, \text{PA}, \alpha, \gamma)$. Dotted lines indicate the input values. (upper-right) Injected brightness profile, denoted by the black line, and 10 samples drawn from the posterior distribution of brightness profiles, denoted by red lines. At the bottom, the residuals between injected and recovered models are shown. (lower-left) Simulated visibilities denoted by blue points and model visibilities denoted by an orange line. The visibilities are binned with a logarithmic grid with $N = 2000$. In the lower panels, the residuals indicating the difference between the simulated and model visibilities are shown. (lower-right) Residual image produced with the MAP estimate in the observed frame. The synthesized beam size $(0.076'', 0.040'')$ is shown at the bottom left.

Gaussian fitting to the visibilities to estimate the geometric parameters. Here, the Gaussian fitting of visibilities corresponds to the fitting of the Gaussian function on the image plane, and it is frequently used for estimating a geometry of a disc. The model was specified by six parameters, where four parameters, $(\Delta x_{\text{cen}}, \Delta y_{\text{cen}}, \cos i, \text{PA})$ are common to our model. We implemented a posterior sampling for the Gaussian model using *emcee* (Foreman-Mackey et al. 2013).

The first and second data from the left in each panel of Fig. 7 show the injected and recovered geometric parameters obtained from the current method and the Gaussian fitting. The central positions were well recovered in both of cases. However, the parameters $\cos i$ and PA obtained from the Gaussian fitting deviated from the injected values; $\Delta \cos i \approx 0.01 - 0.05$ and $\Delta \text{PA} \approx 1 - 5$ deg. These deviations can be

problematic when searching for faint signals (e.g., circumplanetary discs and spirals) because they introduce fake non-axisymmetric structures in the residual images as shown in Figs. 4 and 5.

5 INJECTION AND RECOVERY TEST FOR SPIRALS AND CIRCUMPLANETARY DISC EMBEDDED INTO AXISYMMETRIC EMISSION

As a simple extension to Sec 4, we considered an additional non-axisymmetric perturbation to the symmetric disc. We injected spirals and a point source into an axisymmetric model and simulated the visibilities incorporating noises. Subsequently, we applied our

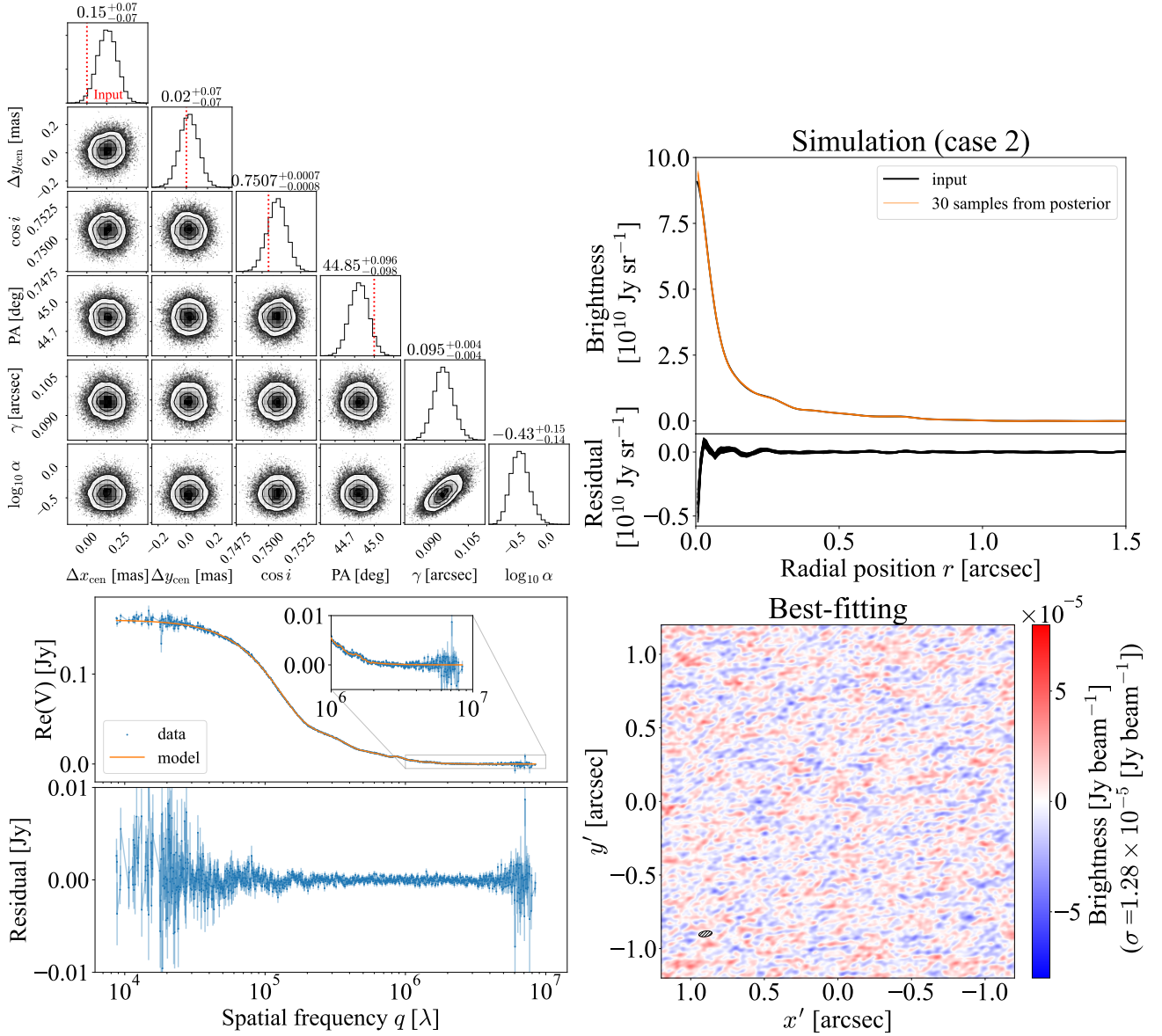


Figure 3. Recovery and injection test for mock data of WaOph 6 in Sec 4. The format of the figure is same as that of Figure 2. The synthesized beam size (0.091″, 0.037″) for the residual image is shown at the bottom left.

method to extract the axisymmetric model and construct residual images, which were then compared against the injected structures.

5.1 Axisymmetric structure + spirals

5.1.1 Injection of spirals

We injected spiral structures into an axisymmetric structure assuming the same observational setup as the simulated case for WaOph6. The brightness profile was assumed to be the same as the case for WaOph6. In this simulation, we assume zero brightness outside $r = 1.11''$, beyond which the brightness in the literature can become negative. Three different models were considered; an odd-symmetric spiral with $m = 1$ component, an even-symmetric spiral with $m = 2$ component, and the combined image for both spirals. Specifically, we assumed perturbations in the form of $\Delta I(m, r, \phi) = \Delta I_m(r) \cos(m(\phi - \phi_m(r)))$, where we adopted an Archimedean

spiral $\phi_m(r) = ar + b$. We assumed $(m, a, b) = (1, 20, 1)$ for the odd-symmetric spiral and $(m, a, b) = (2, -10, -\pi/6)$ for the even-symmetric mode. The amplitudes $I_m(r)$ were assumed as follows:

$$I_1(r) = \begin{cases} 5 \times 10^{-5} \exp(-((r - 0.4'')/0.1'')^2) \\ \text{[Jy beam}^{-1}] \end{cases} \quad (46)$$

$$I_2(r) = \begin{cases} 5 \times 10^{-5} \text{ [Jy beam}^{-1}] & (0 < r < 0.6'') \\ 0 & (0.6'' < r) \end{cases} \quad (47)$$

where the beam area was assumed to be $0.00354 \text{ [arcsec]}^2$. For the calculation of model visibilities, we used functions for the non-uniform fast Fourier transform from PyNUFFT (Lin & Chung 2017; Lin 2018). We adopted the size of the image grid $N_d = 1024$, the size of the oversampled Fourier grid $K_d = 2048$, and the size of the interpolator $J_d = 6$.

Fig. 8 shows the amplitudes of the injected brightness profile and

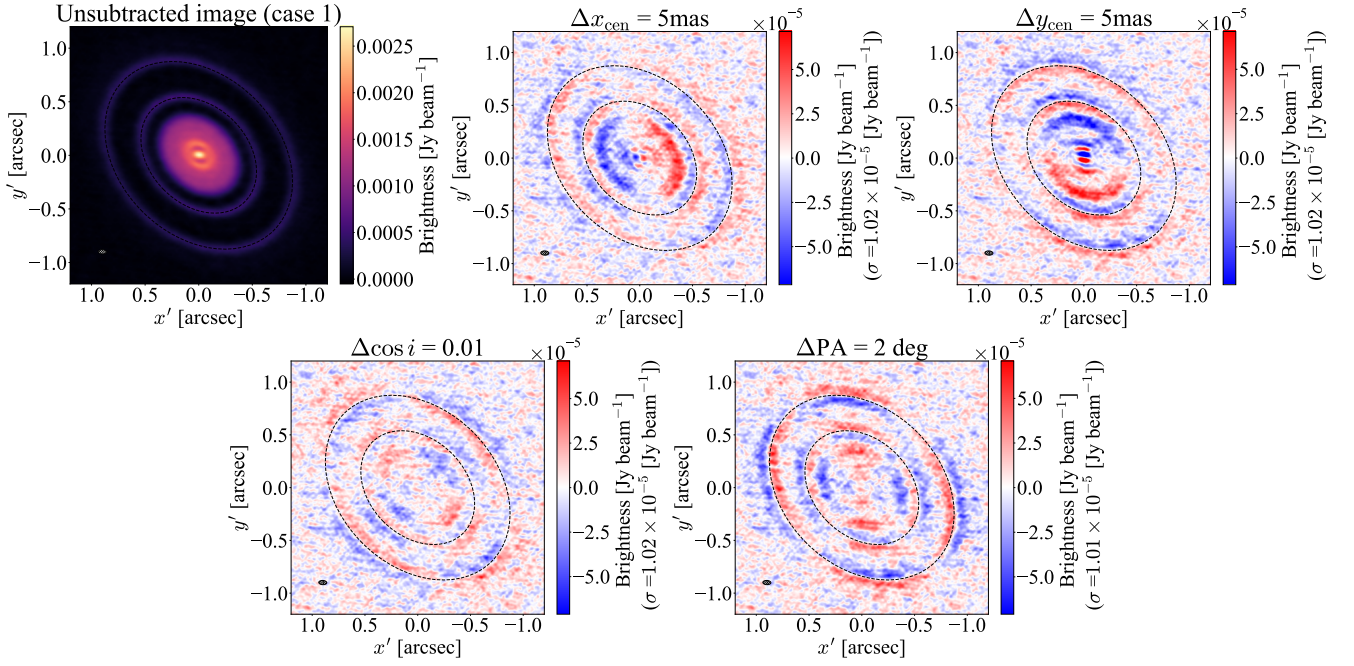


Figure 4. Unsubtracted image and residual images produced with incorrect geometric parameters in simulated case for AS 209. The standard deviation for the brightness σ is computed in the region outside the disc. (upper-left panel) The image produced with CLEAN for unsubtracted data. (upper-middle panel) The residual image produced with residual visibilities produced from the MAP estimate. (other panels) The residual images with each of the geometric parameters being shifted. The synthesized beam size ($0.076''$, $0.040''$) is shown at the bottom left.

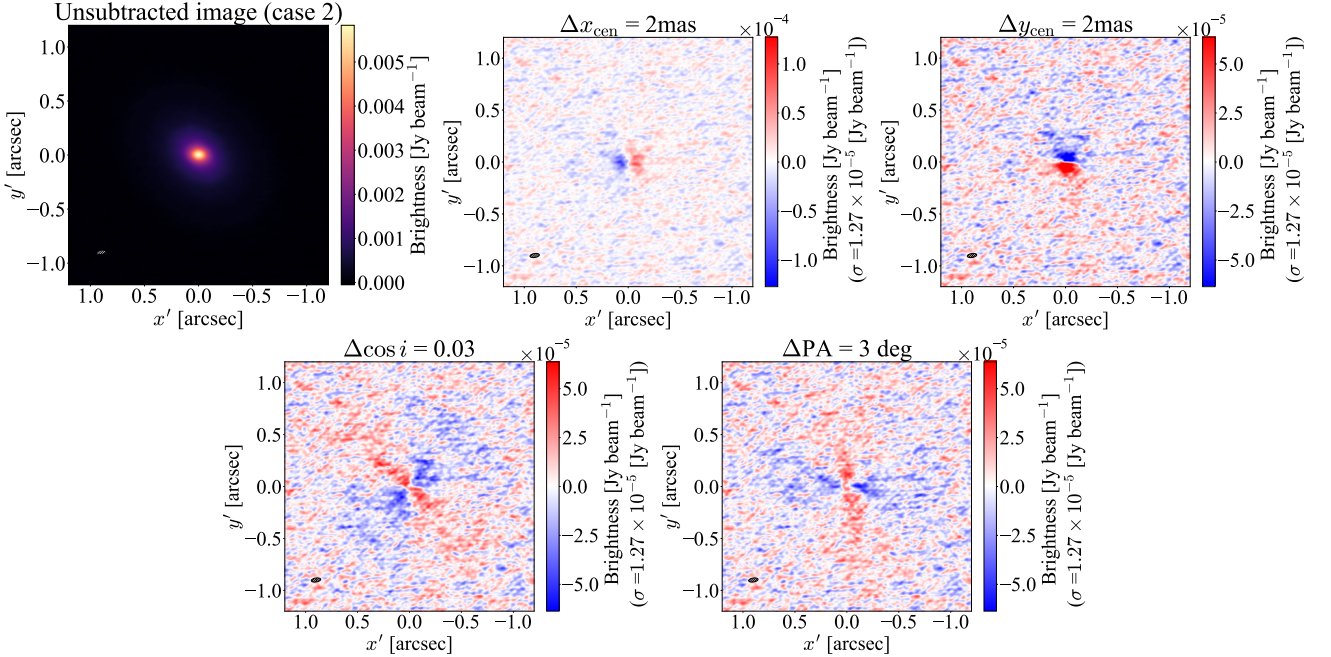


Figure 5. Same as Fig. 4 albeit for simulated data of WaOph 6. The synthesized beam size ($0.091''$, $0.037''$) is shown at the bottom left.

spirals. The amplitudes of the injected spirals were at most ~ 30 per cent of that of the axisymmetric part of the surface brightness, and they were comparable to the amplitudes of the observed spirals with $m = 2$ mode (e.g., WaOph6 and IM Lup) such that the current simulation indeed considered the realistic situation. The left panels in Fig. 9 show the injected images for the three cases.

5.1.2 Comparison of injected and recovered models

After simulating the visibilities, we applied our method to estimate the axisymmetric structures. We selected the MAP estimate on geometric parameters, drew a brightness profile, and subtracted their

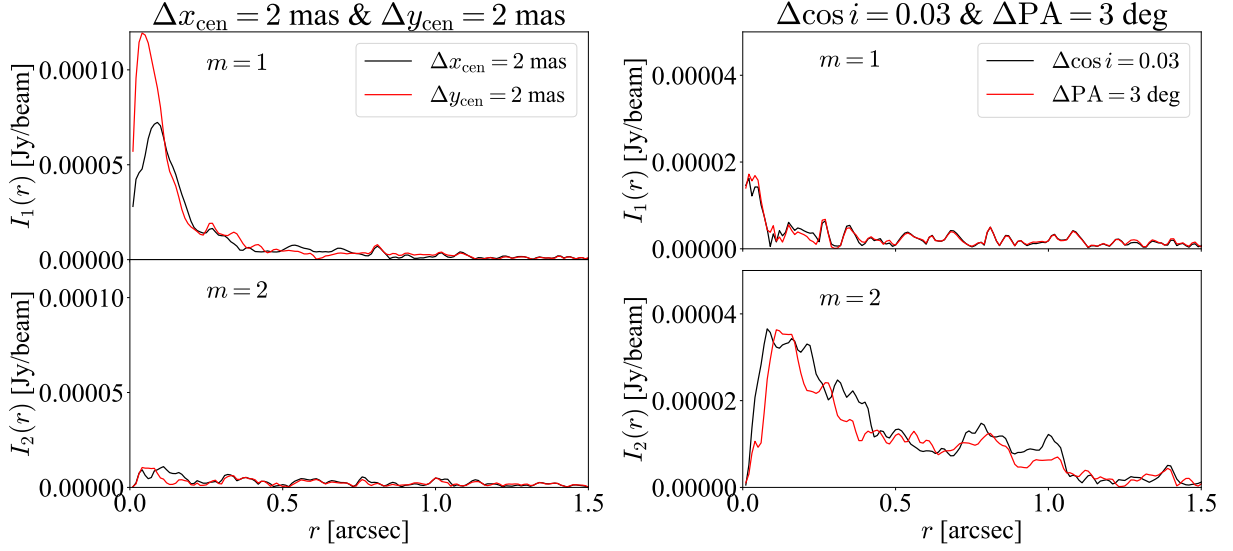


Figure 6. $I_{m=1,2}(r)$ computed for residual images in Fig. 5. The left and right panels show the results for the shifted central position and the shifted disc orientation, respectively.

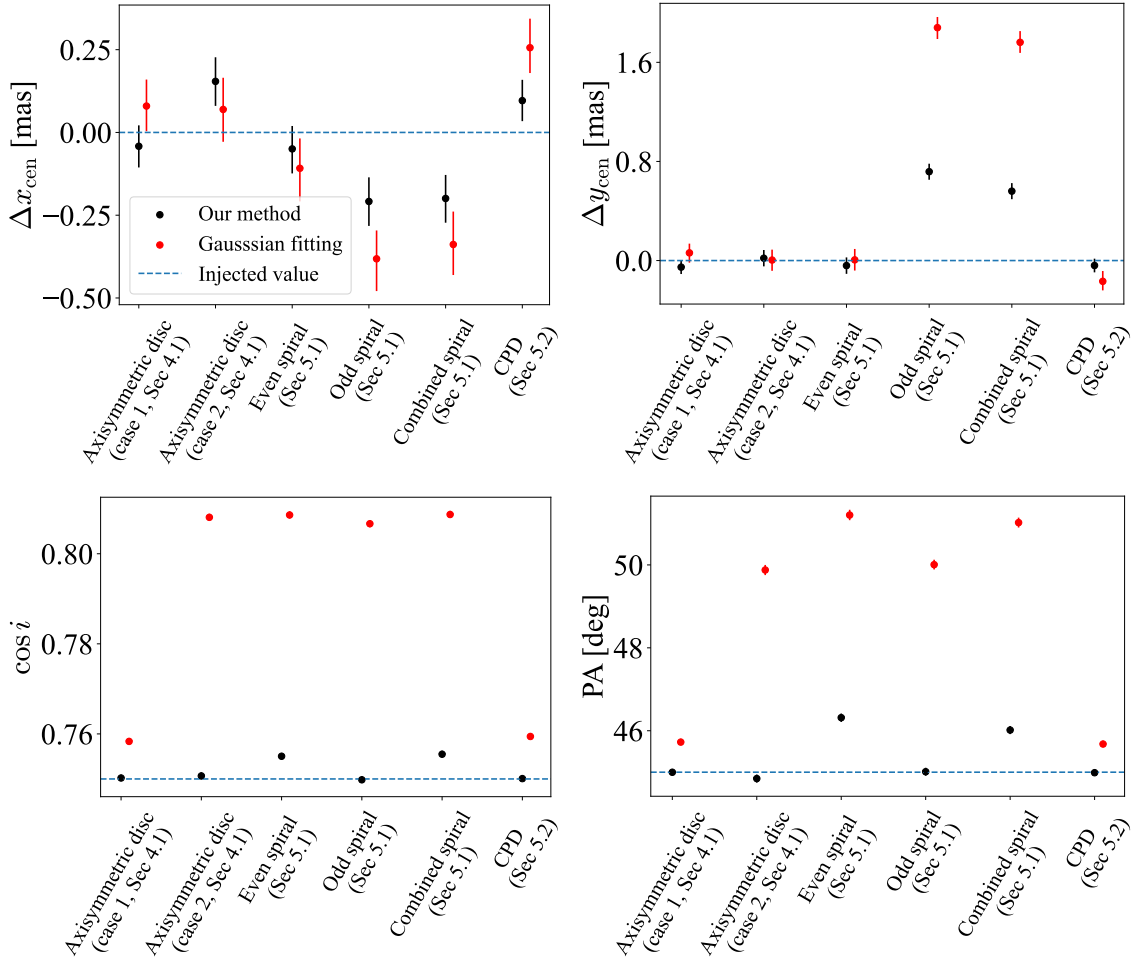


Figure 7. Injected and recovered geometric parameters in simulations in Sec 4 and 5. Each panel compares the geometric parameter: the central position and the disc orientation. The geometry is estimated using the proposed method (black) and the Gaussian fitting (red). The injected parameters are indicated by the dashed blue lines.

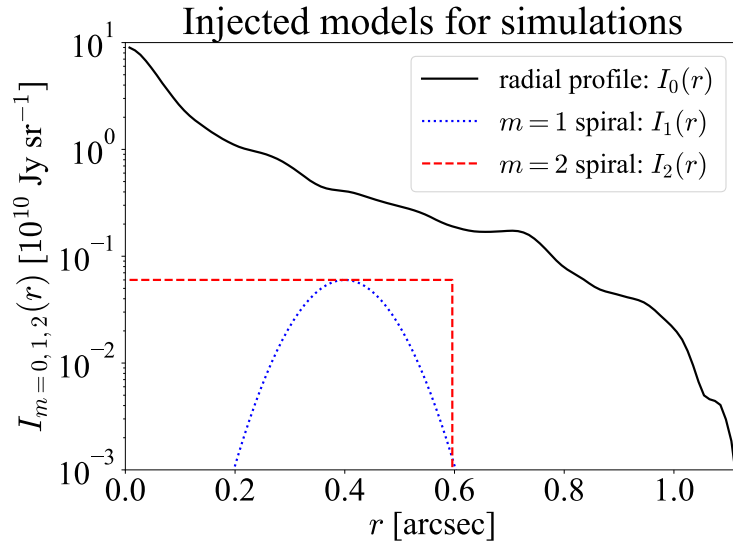


Figure 8. Amplitudes of injected brightness profile and spirals in simulations in Sec 5.

contributions from the observations. The residual visibilities were then used to construct the residual images using CLEAN.

For the comparison, we also created the residual images by adopting geometric parameters estimated from the Gaussian fitting to the visibilities. Specifically, we assumed geometry with Gaussian fitting and hyperparameters from our method, and we drew a brightness profile. Subsequently, the brightness model was subtracted from the visibilities, and the residual visibilities were imaged with CLEAN.

The residual images are shown in Fig. 9. In the second column, we show the recovered residual images obtained via the current method. The last column shows the residual images created from the geometry given by the Gaussian fitting. The residual images from our method were reasonably consistent with the injected maps, whereas those based on the Gaussian fitting exhibited larger residuals. However, the innermost structures of the discs at $r < 0.2 - 0.3''$ were not perfectly recovered even with our method. Specifically, the excess in the amplitude of the spiral was observed at $r < 0.3''$ for the odd-symmetric spiral. Moreover, no clear spiral structure was observed at $r < 0.2''$ for the even-symmetric spiral. These inconsistencies arise from the biases in the estimated geometric parameters, as discussed next.

The comparison of geometric parameters for three cases is shown in Fig. 7. For every parameter, larger deviations from the injected parameters were observed in the case of the Gaussian modelling than that using the proposed method. However, our method was still affected with biases; $(\Delta x_{\text{cen}}, \Delta y_{\text{cen}})$ for the odd-symmetric spiral, and $(\text{PA}, \cos i)$ for the even-symmetric spiral. In the case of the combined spiral, the shifts in all the geometric parameters roughly corresponded to the summation of shifts in the odd and even-symmetric spirals. These shifts are reasonable because, as discussed in Sec 4, the shifts in $(\Delta x_{\text{cen}}, \Delta y_{\text{cen}})$ introduce the residuals with $m = 1$ mode, which can be degenerate with the injected odd-symmetric spiral. Similarly, the shifts in $(\cos i, \text{PA})$ introduce the residuals with $m = 2$ mode, which can be degenerate with the even-symmetric spiral. These degeneracies produce the biases in estimating geometric parameters. Consequently, these biases impede recovery of injected non-axisymmetric structures, as shown in Fig. 9.

The left panels in Fig. 10 compare the residual image with the injected odd-symmetric spiral. The phase for the spiral was reasonably

recovered in the residual image. The upper right and lower panels show the recovered radial phases $\phi_1(r)$ and the amplitudes $I_1(r)$. The amplitudes and phases were reasonably recovered in the range of $0.3'' < r < 0.5''$, where the spiral amplitude was large. However, the amplitudes were clearly overestimated at inner radii $r < 0.3''$. This overestimation was due to the biased central position, and is consistent with the failure of the recovery of the spiral at the innermost part in Fig. 9.

The right panels in Fig. 10 show the comparison in the case of the even-symmetric spiral. The phases were reasonably recovered in most of the radial range; however, the deviations were large at the innermost part $r < 0.1''$. The amplitudes were well recovered at $0.3'' < r < 0.6''$, whereas they were underestimated in the inner regions. The deviations in the phases and amplitudes originated from the biases in the estimation of $(\cos i, \text{PA})$, which are associated with $m = 2$ feature.

Finally, Fig. 11 shows the residual images obtained with either real or imaginary parts of images. As evident, the imaging with only the real part successfully extracted the even-symmetric spiral, whereas that with only the imaginary part extracted the odd-symmetric spiral. This is consistent with the discussion in Sec 3.3. Notably, we successfully extracted odd- and even-symmetric spirals for the case of the combined image. Thus, in a realistic situation, where odd- and even-symmetric components are mixed, the imaging with either the real or imaginary part is indeed useful for separating them. In addition, the noise level decreased by a factor of $\sim \sqrt{2}$, or the signal-to-noise ratio increased by a factor of $\sim \sqrt{2}$ because of the assumed symmetry $I(x, y) = I(-x, -y)$ or $I(x, y) = -I(-x, -y)$. Thus, this method would be helpful in the search for faint signals.

5.2 Axisymmetric structure + circumplanetary disc emission

We injected a point source to the simulated data assuming circumplanetary disc emission and tested the recovering ability. The flux for the point source was assumed to be 0.1 mJy, which is approximately 10 times more significant than the noise level but 10 times lower than the ambient emission at the inner disc. The planetary location was set to be at $0.5''$ in the deprojected frame $(x'', y'') = (0.5'', 0)$, such that the source resided in a large gap of a disc. We considered the

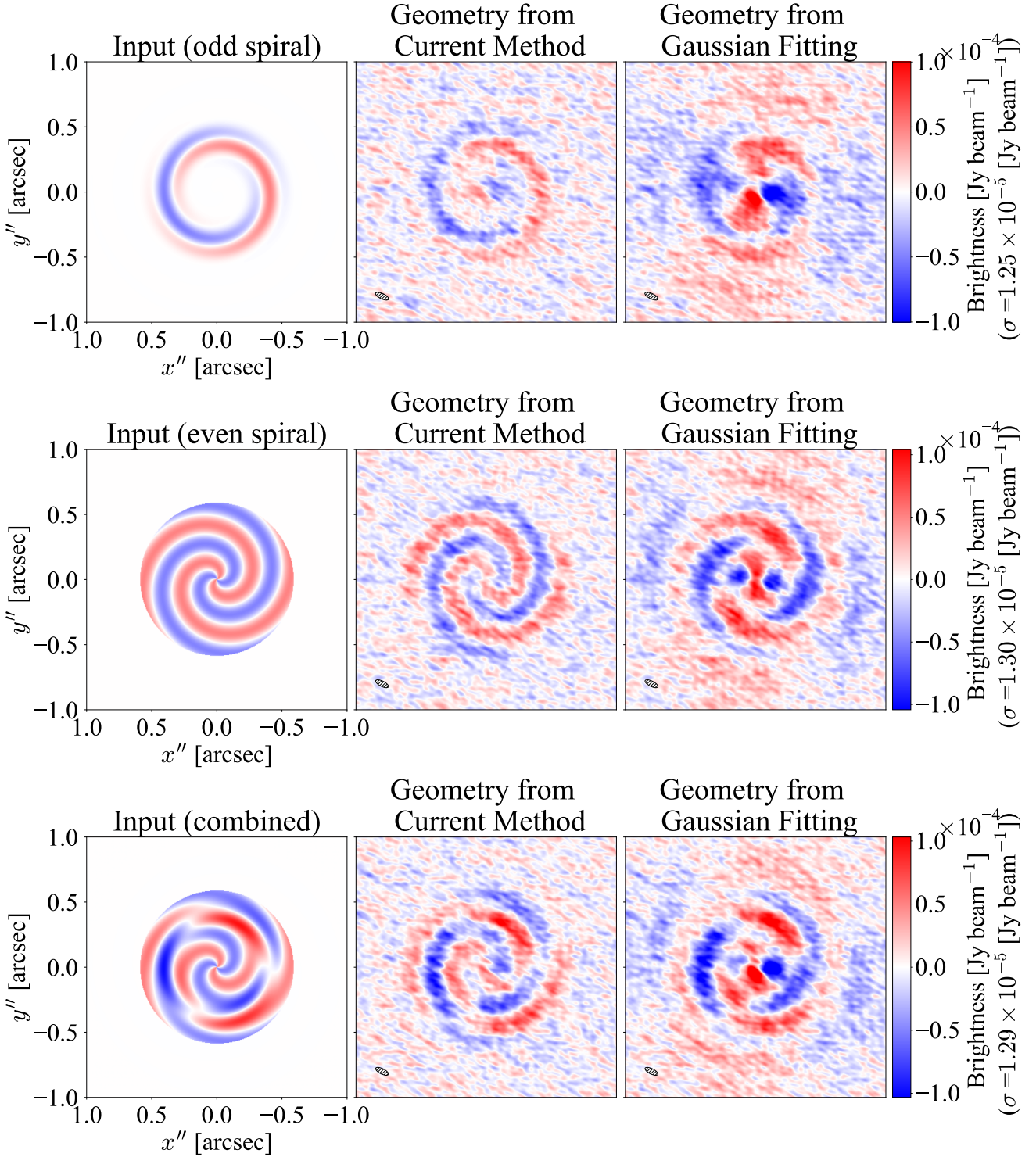


Figure 9. Injected and recovered residual images for spirals in deprojected frames; (top rows) Even-symmetric spiral (middle rows) Odd-symmetric spiral (bottom rows) Combined spirals. (left columns) Injected perturbations (middle columns) Recovered residual images based on the proposed method (right columns) Recovered residual images based on the geometry obtained with Gaussian fitting. The synthesized beam size for the deprojected image ($0.11'', 0.041''$) is shown at the bottom left, and the beam size for the image before deprojection is ($0.091'', 0.037''$).

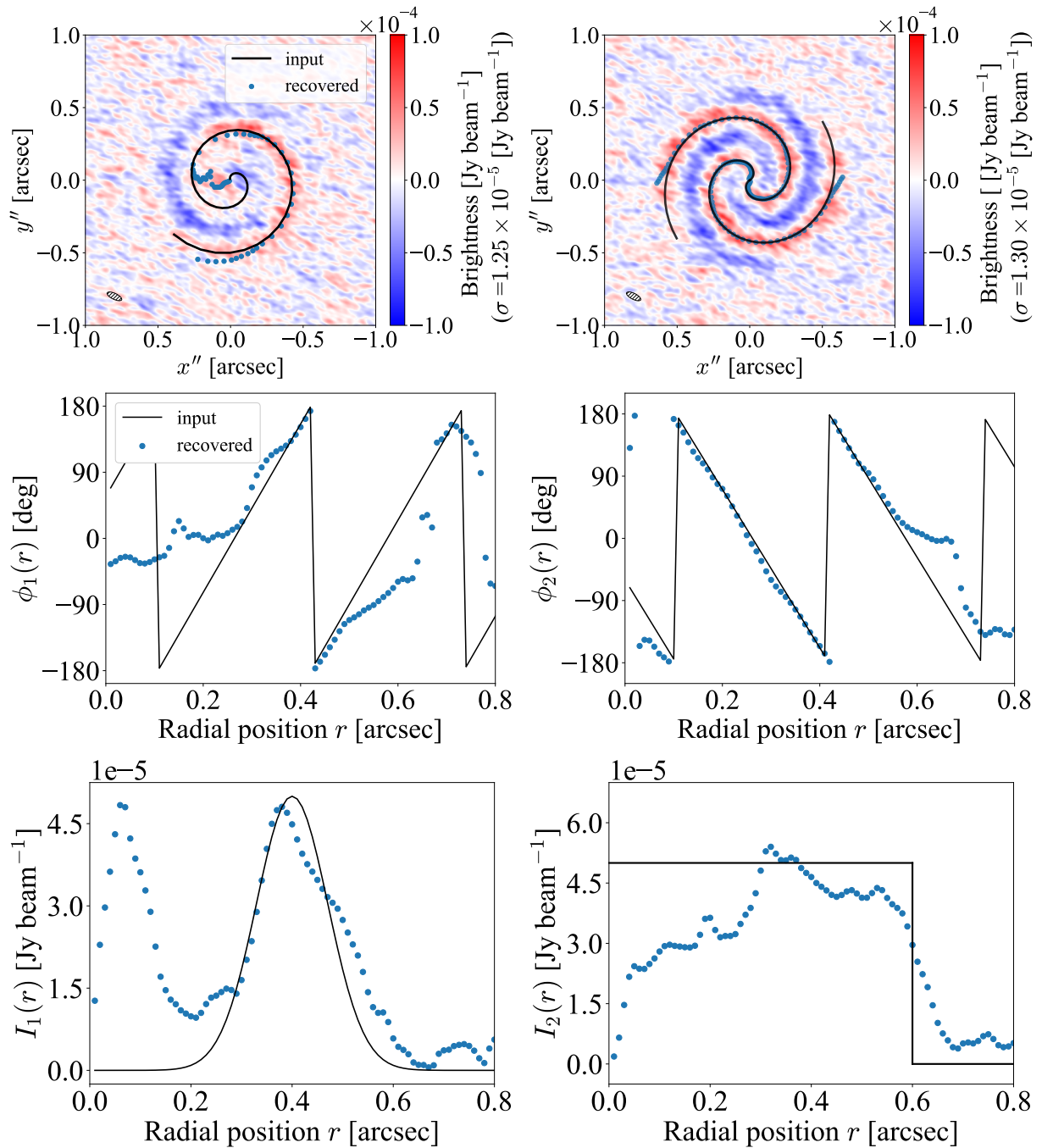


Figure 10. Phases and amplitudes for injected and recovered odd- (left) and even- (right) symmetric spirals in Sec 5. (upper panels) The residual images in deprojected frame overplotted with the locations of injected (black) and recovered phases (blue). (middle panels) Input and recovered phases in the radial direction. (lower panels) Input and recovered amplitudes of spirals.

same brightness profile and geometric configuration as the simulated case for AS209 in Sec 4 and computed the visibilities and added noises to the data assuming the same observational setup.

Using our method, we similarly recovered the brightness profile and geometric parameters from the simulated data. With the MAP estimate of geometric parameters, we randomly drew one brightness profile and subtracted the model from the simulated visibilities. Fig. 12 shows the residual images, which are generated from the residual

visibilities. The position of the injected point source was reasonably recovered, and the flux density was well recovered as well.

Moreover, the central position of the disc was slightly biased, suggesting that the point source might introduce another bias. To check this possibility, we injected the brighter sources with fluxes of 1 and 2 mJy, and attempted to estimate the central positions of the discs by repeating the same analyses. Resultantly, the central positions were estimated to be $(\Delta x_{\text{cen}}, \Delta y_{\text{cen}}) = (0.58^{+0.07}_{-0.07} \text{ mas}, -0.50^{+0.07}_{-0.07} \text{ mas})$ for the 1 mJy

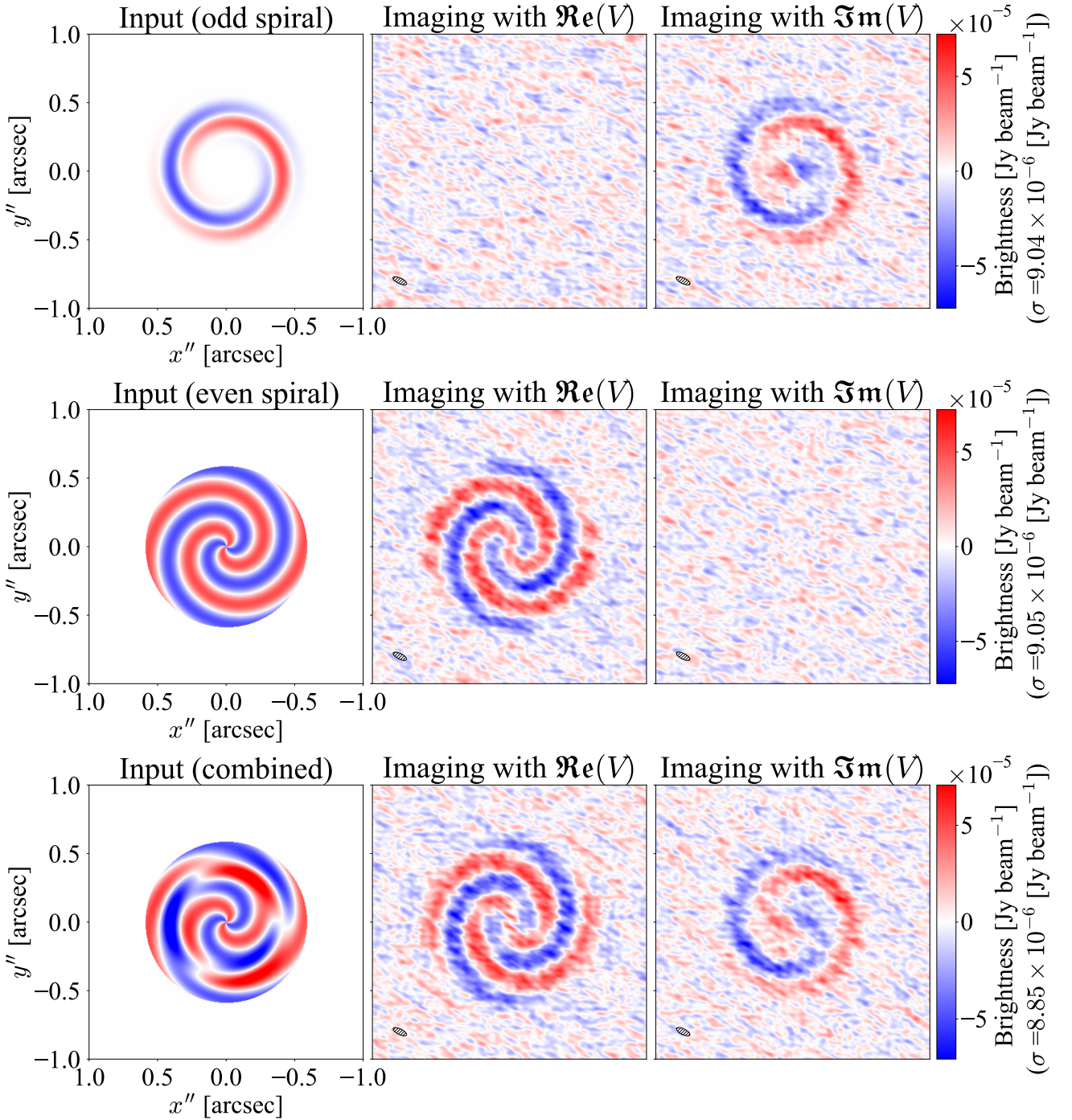


Figure 11. Demonstration of extraction of odd- and even-symmetric spirals with real and imaginary parts of the visibilities; (top rows) Residual images for even-symmetric spiral in deprojected frames (middle rows) Residual images for odd-symmetric spiral (bottom rows) Residual images for combined spirals (left columns) Injected spirals (middle column) Image produced with only the real part of the data (right column) Image produced with only the imaginary part of the data.

source and $(\Delta x_{\text{cen}}, \Delta y_{\text{cen}}) = (0.97_{-0.07}^{+0.07} \text{mas}, -0.82_{-0.07}^{+0.07} \text{mas})$ for the 2 mJy source. However, the positions of the flux centre with respect to the disc centre were calculated as $(0.9 \text{mas}, -0.9 \text{mas})$ and $(1.8 \text{mas}, -1.8 \text{mas})$ for the 1 and 2 mJy sources, respectively, and they were comparable to the biases in the estimation. Thus, even with our method, the central positions of the disc can be biased by approximately half of the positional difference between the disc and

flux centres. This holds true not only for the point source but also for localized emission; for example, a crescent-shaped emission. In the data analysis, in case of bright localized emission, it is recommended that such an additional emission be modelled or removed because it affects the residual image creation process. Specifically, in the case of the point source, we can directly include the point source model in the axisymmetric model. However, if the emission is more complicated,

we can model them on the image plane and remove their visibilities from the observations, as presented in [Andrews et al. \(2021\)](#).

6 APPLICATION TO REAL DATA: ELIAS 20 AND AS 209

We applied the proposed method to the real DSHARP data of two discs, Elias 20 and AS 209, to demonstrate its feasibility. AS 209 is one of most structured discs in the DSHARP sample, and Elias 20 shows a non-axisymmetric feature in the residual map as shown later. More systematic studies on the DSHARP discs will be presented in future studies.

We downloaded the measurement sets of the DSHARP data for the two discs, and applied time and spectral averaging in the same manner as that of Sec 4.1. We manually reduced the recorded weights by 3.50 and 3.44 for Elias 20 and AS 209, respectively, to match the observations. The data were then binned with a log grid with $N_{\text{bin}} = 500$, and they were analyzed with the current method. After sampling the posterior distribution for the parameters, we created the residual images following the method presented in Sec 3.1.1. To create residual images, we used two different geometries: one is obtained from our method, and the other obtained from [Huang et al. \(2018a\)](#), who estimated the geometric parameters through ellipse fitting of annular substructures on image planes.

Table 1 presents the derived geometries for Elias 20 and AS 209. For a reference, we also show the geometric parameters from [Huang et al. \(2018a\)](#). The length scale parameters γ for Elias 20 and AS 209 are comparable to their beam sizes (0.076'', 0.040'') and (0.048'', 0.028''), respectively. This result is consistent with the discussion in Sec 4.1 and Appendix D; γ was determined by the combination of the intrinsic brightness profile and the UV-coverage.

The derived geometries from our method were mostly consistent with the previous estimates from [Huang et al. \(2018a\)](#), although there are slight or moderate discrepancies. The central positional estimates for AS209 are offset by approximately 1 mas for both directions. Similarly, in the case of Elias 20, the positions also show a difference of 1-2 mas, and notably, there is a 0.07 difference in $\cos i$.

These small discrepancies are important for creating the residual images. Figs. 13 and 14 show the brightness profiles, visibilities, and residual images. For comparison, we also show the brightness profiles and model visibilities obtained by [Jennings et al. \(2022a\)](#), who systematically analysed the DSHARP data. Here, they used `frank` ([Jennings et al. 2020](#)), which reconstructs the radial brightness profile by fitting the real part of the deprojected visibilities. Note that [Jennings et al. \(2022a\)](#) adopted the non-negativity condition on the brightness profile in case of AS 209, and assumed additional point-source emission with flux of 0.66 mJy at the disc centre to suppress the artificial oscillating features for Elias 20 (detailed discussion on point-source correction is presented in Appendix A in their paper). Their model visibilities, as shown in Fig. 13, for Elias 20 are thus converged to 0.66 mJy, which corresponds to the flux of this point-source emission.

The brightness profiles and model visibilities of the proposed method and that of `frank` were mostly consistent; however, there existed slight or moderate differences.

In the case of Elias 20, the model visibilities are horizontally offset, and this is due to the difference in the $\cos i$. In addition, the models using the proposed method gradually converged to zero at high spatial frequencies, whereas those from `frank` sharply converged to the flux of the point source around 5 $M\lambda$. The locations of the tipping points for the convergence were determined by γ in our method or hyperparameters in `frank`, and the differences in the way

of convergence were due to the different choices for the regularization.

The notable difference was observed in the peak brightness values near $r = 0''$. The peak brightness is generally hard to estimate accurately because of the small flux at the small radii. This is mainly due to the point-source correction with 0.66 mJy adopted in [Jennings et al. \(2022a\)](#), which was however not included in their brightness profile. The innermost brightness $I(r_1)$ in our model corresponding to the flux 0.66 mJy is $\sim 6 \times 10^{10} \text{ Jy sr}^{-1}$, which can largely explain the observed difference of about $10 \times 10^{10} \text{ Jy sr}^{-1}$ between this study and [Jennings et al. \(2022a\)](#). Another potential explanation could be the difference in the strength of the regularization for smoothing, although identifying the superior model remains challenging at this stage.

In the case of AS 209, the brightness profiles and model visibilities were mostly consistent for our method and that of `frank`. The peak brightness values, on the other hand, showed the difference of about $2.5 \times 10^{10} \text{ Jy sr}^{-1}$, which is yet smaller than that of Elias 20. The discrepancy potentially arises from the difference in regularization strengths or the non-negative condition adopted in [Jennings et al. \(2020\)](#). Overall, our method exhibited the same capability at recovering the brightness profiles as that of `frank`.

The estimation of the residual images was also improved using our updated geometry. In case of Elias 20, the residual image derived with the geometry from [Huang et al. \(2018a\)](#) exhibits the $m = 2$ pattern, which is mainly attributed to the inclination offset, as shown in Figs. 4 and 5. This coherent pattern, also seen in [Jennings et al. \(2020\)](#), mostly disappears in our update image, suggesting that our method suppresses the artificial pattern owing to wrong geometric parameters. In case of AS 209, the previous literature identified significant residual patterns ([Guzmán et al. 2018](#); [Jennings et al. 2020](#)). The residual image with the updated geometry is less noisy than the previous estimate; however, structured patterns were still present. We thus conclude that the residual pattern for AS 209 cannot be removed by solely optimizing the geometric parameters for a geometrically-thin axisymmetric model.

There can be multiple possible explanations for the residual pattern in AS 209. It may be due to the real non-axisymmetry in the physical parameters, or because of the geometric effect. The individual rings may have different geometric parameters, that is, misalignment or positional offsets, where the latter case was reported for HL Tau ([ALMA Partnership et al. 2015](#)). Further, the vertical structure may also render the residual pattern complicated, although the observed residual image is not perfectly consistent with this hypothesis, which predicts the symmetric pattern with respect to the x axis (minor axis). To unveil the origin of the residual pattern for AS 209, we require a more complicated model with multiple geometric parameters or vertical structures; however this is beyond the scope of this paper.

7 APPLICATION TO REAL DATA: PDS 70

7.1 Analysis with axisymmetric disc model

As a practical application to recovering a circumplanetary emission, we applied our method to the data of PDS 70 in [Benisty et al. \(2021\)](#). The same data were already analyzed by `frank` in [Benisty et al. \(2021\)](#). We used the combined dataset, including long, medium, and short baseline data, for continuum emission in Band 7 as used in [Benisty et al. \(2021\)](#). The data were then averaged in a similar manner to Sec 4 to reduce the data size.

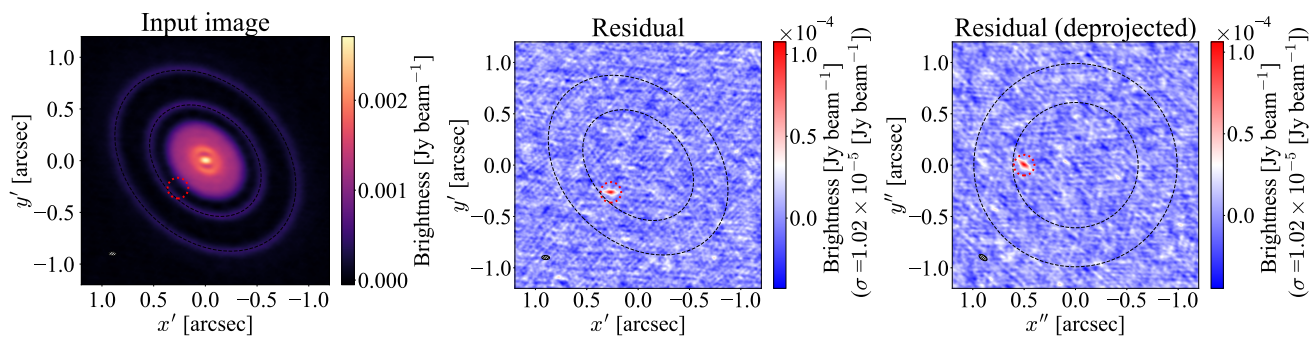


Figure 12. Residual images for the simulated data that include a simulated circumplanetary disc emission, whose position is indicated by a red dashed circle. The central and right panels show the maps in an observed and deprojected frames, respectively. The synthesized beam sizes before and after deprojection are $(0.076'', 0.040'')$ and $(0.089'', 0.045'')$ are shown at the bottom left, respectively.

Table 1. Hyperparameters for Gaussian Process and geometric parameters of Elias 20, AS 209, and PDS 70. Geometric parameters assumed in previous studies are also shown for comparison.

Name	γ ["]	$\log_{10} \alpha$	Δx_{cen} [mas]	Δy_{cen} [mas]	$\cos i$	PA [deg]
Elias 20 (this study)	$0.038^{+0.001}_{-0.001}$	$0.800^{+0.161}_{-0.150}$	$-53.919^{+0.064}_{-0.065}$	$-488.849^{+0.071}_{-0.073}$	$0.584^{+0.001}_{-0.001}$	$153.930^{+0.060}_{-0.060}$
Elias 20 (Huang et al. (2018a))	-54.5	-491.0	0.656	153.2
AS 209 (this study)	$0.028^{+0.001}_{-0.001}$	$-1.202^{+0.092}_{-0.087}$	$0.767^{+0.068}_{-0.066}$	$-1.225^{+0.048}_{-0.050}$	$0.821^{<0.001}_{>-0.001}$	$85.766^{+0.036}_{-0.038}$
AS 209 (Huang et al. (2018a))	1.9	-2.5	0.819	85.8
PDS 70 (this study)	$0.064^{+0.003}_{-0.003}$	$-2.396^{+0.112}_{-0.110}$	$10.827^{+0.079}_{-0.078}$	$14.953^{+0.101}_{-0.105}$	$0.6427^{+0.0002}_{-0.0002}$	$160.003^{+0.020}_{-0.022}$
PDS 70 Benisty et al. (2021)	12	15	0.6494	161

In the continuum emission, there is a notable crescent feature in the North-West direction.

citbenisty2021 removed the asymmetric feature by following the method described in Andrews et al. (2021). Specifically, using the CLEAN model image, they defined the asymmetry model by isolating the emission of the crescent feature, and subtracted the mean radial profile outside the area from the model, leaving only the asymmetric contribution. The constructed model was then Fourier transformed, and the model visibilities are subtracted from the observed visibilities, which were analyzed using frank. As demonstrated in Andrews et al. (2021), the method is effective for extracting strong asymmetric features that hinder the detection of weak signals, such as emission from CPD (see Appendix B of their paper for details).

In this paper, we applied our method to the original data without any such prior subtraction, aiming to minimize the manual adjustment. A major concern with this simpler approach was that the localized crescent feature might bias the estimates of the geometric parameters, especially the position estimate $(\Delta x_{\text{cen}}, \Delta y_{\text{cen}})$, similar to that in Sec. 5.2. However, as will be shown later, the derived parameters agree well with those estimated by Benisty et al. (2021), with slight differences, such as 1 mas in Δx_{cen} and 1 deg in PA. The effect of the positional difference in the residual image is negligible in the current analysis (see Fig. F1 in Appendix F). Therefore, we simply present the result of the analysis with our method, without implementing the prior subtraction of the asymmetric features.

We employed a logarithmically spaced grid with $N_{\text{bin}} = 500$ and $(x_{\text{min}}, x_{\text{max}}) = (10^2 \lambda, 1.1 \times 10^7 \lambda)$ to bin the data. We drew samples from the posterior of parameters using emcee with 16 walkers and 10,000 samples, and we ensured the convergence of MCMC. Table 1 lists the parameters from this study and Benisty et al. (2021) (specifi-

cally, Appendix B in their paper). The two studies yielded consistent values, while there are slight differences: approximately 1 mas in Δx , 0.007 in $\cos i$, and 1° in PA.

We constructed a visibility model using the MAP solution of parameters, and subtracted it from the observed visibilities. Subsequently, the residuals were processed with CLEAN, with a threshold of $n_{\text{sigma}} = 3.5$. Note that we did not adopt JvM correction adopted in the previous literature (Czekala et al. 2021; Benisty et al. 2021; Balsalobre-Ruza et al. 2023) to prevent potential exaggeration of signal-to-noise ratios in an image (Casassus & Cárcamo 2022). In the imaging process, we experimented with various robust parameters, specifically setting them to be 0.0, 0.5, 1.0, and 1.5. We found that a setting of 1.0 offers a small standard deviation image while maintaining relatively high angular resolution. For comparison, we also generated a residual image using the geometric parameters in Benisty et al. (2021) while adopting the hyperparameters in the MAP solution from our modeling.

Figure 15 illustrates the brightness profiles and visibilities derived from our method. For the brightness profiles, we show 30 random samples, indicated by the light orange lines. The brightness profile was characterized by the presence of the inner disc as well as the outer disc with two local maxima, consistently with previous studies (Keppler et al. 2019; Benisty et al. 2021). At the outer disc, there is another shoulder at $0.85''$, as also reported in Benisty et al. (2021). A slight positive excess was also observed in the cavity region, between the inner and outer discs. The flux would arise from emission around PDS 70 b&c and their Lagrange points (Benisty et al. 2021; Balsalobre-Ruza et al. 2023).

Figure 15 shows the residual images for PDS 70; the left figure in the lower panels shows the residual map based on our study in

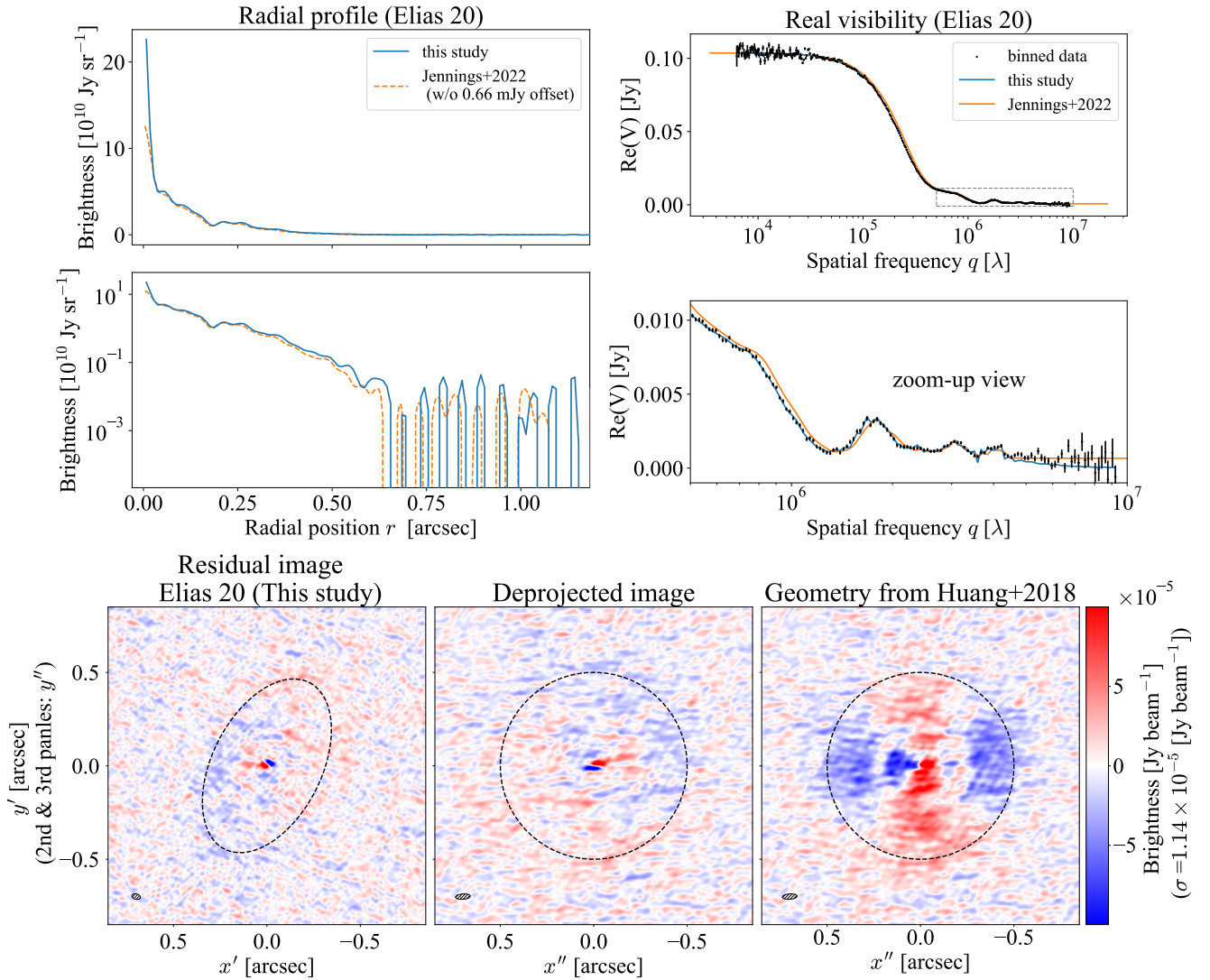


Figure 13. Analysis of DSHARP data of Elias 20. (upper-left panels) Brightness profiles recovered by this study and Jennings et al. (2022a). The profiles are shown in linear and log scales in the upper and lower panels, respectively. (upper-right panels) Model and observed real visibilities and deprojected spatial frequencies. The model by Jennings et al. (2022a) is also shown for comparison. (lower panels) Residual image made with the geometry from our method in the observational frame, that in the deprojected frame, and the image made with that of Huang et al. (2018a) in the deprojected frame. In the residual images, the reference lines at $r = 0.5''$ are also shown. As reference, the ellipse and circles are shown. The synthesized beam sizes before and after deprojection are $(0.048'', 0.028'')$ and $(0.079'', 0.029'')$ are shown at the bottom left, respectively.

the observational frame, the middle one is shown in the deprojected frame, and the right one is based on the geometry in Benisty et al. (2021). The residual images from both geometries are consistent, but there is a slight discrepancy due to difference in geometry. We investigated which parameter made the large difference, and found that the difference in PA largely accounts for the discrepancy (also see Appendix F).

The upper panels in Figure 16 shows the residual images in deprojected coordinate and the polar coordinate. Our methods successfully recovered the circumplanetary emission around PDS 70 c, in addition to the crescent feature. The emission of the circumplanetary disc remains unresolved, consistently with the analysis in Benisty et al. (2021). We measured the brightness of the circumplanetary emission around PDS 70 c. The estimated peak brightness of the emission were $95.2 \pm 17.7 \mu\text{Jy beam}^{-1}$, $89.8 \pm 13.1 \mu\text{Jy beam}^{-1}$, and $91.2 \pm 11.9 \mu\text{Jy beam}^{-1}$ for robust = 0, 0.5, 1.0, respectively. Our estimates for

robust = 0 and 0.5 were consistent with the result from Benisty et al. (2021), which reported $86 \pm 10 \mu\text{Jy beam}^{-1}$ and $79 \pm 7 \mu\text{Jy beam}^{-1}$. It should be noted that in our study, the minor/major beam sizes exhibit a slight deviation from those reported in Benisty et al. (2021) (see Table 4) with discrepancies ranging between 0.01 and $0.03''$. This minor variation is likely due to the time and spectral averaging processes applied to our measurement sets, but the discrepancies give negligible impacts on the measurements of intensities for PDS 70 c. On the other hand, our estimate for robust = 1.0 significantly differed from $170 \pm 4 \mu\text{Jy beam}^{-1}$ reported in Benisty et al. (2021). The discrepancy likely arises from the contamination from the main disc emission, which is however mostly mitigated by our method.

An extended emission around PDS 70 b was reported in Benisty et al. (2021). Our analyses similarly identified a positive excess around it, and the enclosed flux for each robust parameter was around $50 \mu\text{Jy}$, comparable to $38 \mu\text{Jy}$ reported by Benisty et al. (2021), al-

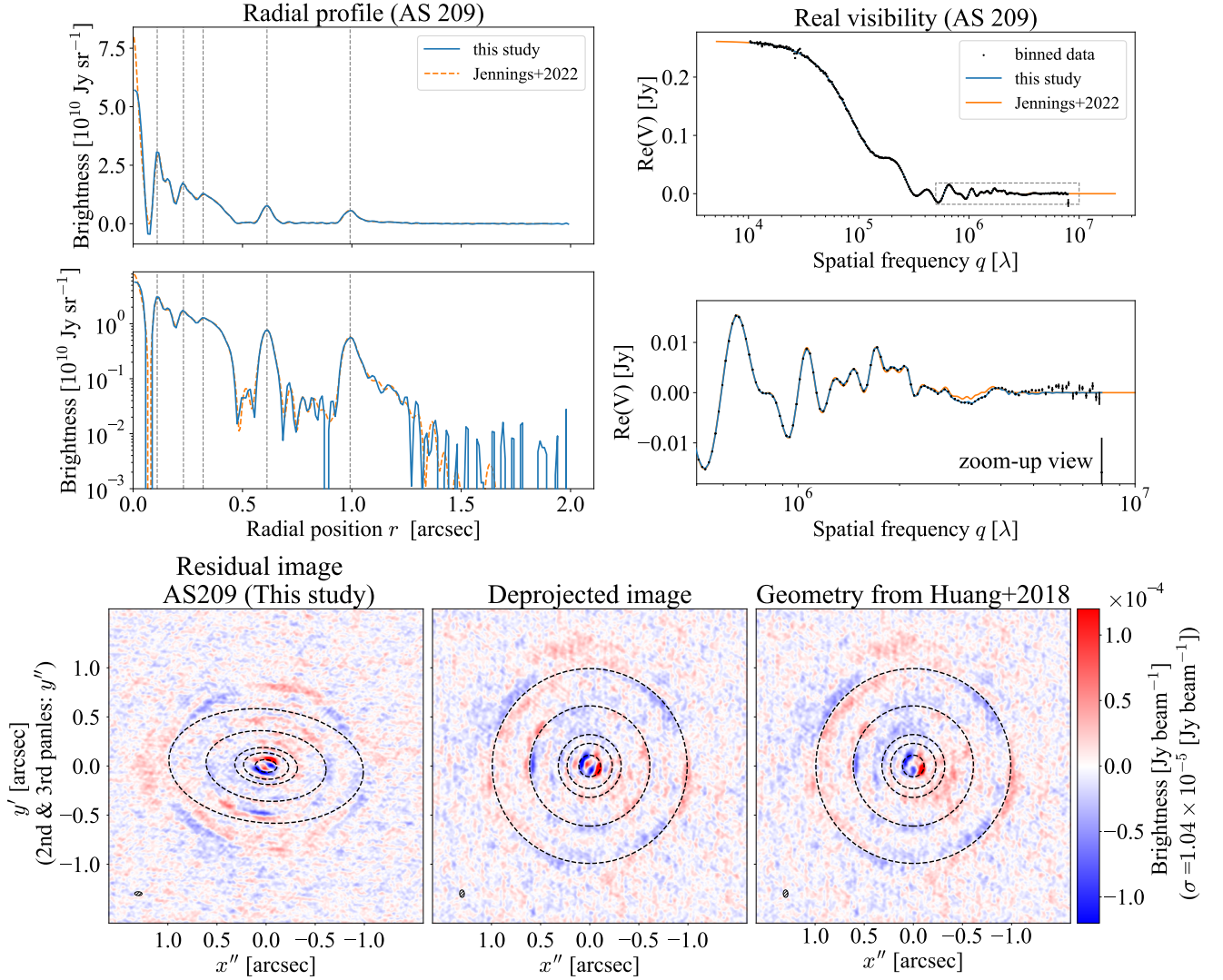


Figure 14. Analysis of DSHARP data of AS 209. The format of the figure is the same as that of Fig. 13. The reference lines corresponding to gaps and rings are shown in brightness profiles and the residual images. The synthesized beam sizes before and after deprojection are $(0.076'', 0.040'')$ and $(0.076'', 0.049'')$, respectively.

though our measured flux was influenced by the cumulative noise within the region. Additionally, there was a claim on the tentative co-orbital emission near PDS 70 b’s L5 Lagrangian point (Balsalobre-Ruza et al. 2023). We found the signal to be visually insignificant in the images. Indeed, the peak emission in this region showed significance of $2.0\text{--}2.7\sigma$ for robust = $(0, 0.5, 1)$, slightly lower than the significance around $3.3\text{--}3.4\sigma$ reported in Balsalobre-Ruza et al. (2023) for the robust parameters that we considered. The slight discrepancy might be due to the subtraction of the annular brightness in our study. On the other hand, Balsalobre-Ruza et al. (2023) observed higher significance at $5\text{--}6\sigma$ if they employed robust ≥ 1.5 and JvM correction, which were not investigated in this study. The further observations and analyses of these marginal signals will be needed.

7.2 Additional analysis on residual images by subtracting crescent model

Apart from the crescent feature and circumplanetary emission, residual features still persist in the image. Benisty et al. (2021) also iden-

tified the residual features in the image after subtracting the axisymmetric component and asymmetric feature including the crescent feature (rightmost end of Fig. 8 in their paper), although the detailed discussion was not presented. Here, we revisited the residual feature in more detail using our updated residual image.

For the further investigation, the high contrast of the bright crescent feature hinders the search of faint features in the images. In addition, the bright crescent unnecessarily introduces the negative flux, because the residual image is supposed to have zero flux when averaged over the polar direction. To avoid the latter problem, Benisty et al. (2021) removed the asymmetric feature in the image-based analysis before the visibility analysis. To mitigate these effects, we instead attempted to subtract the bright crescent feature by employing a provisional analytical model. Specifically, we considered a model comprising of the super Gaussian function in the radial direction and the von Mises distribution in the polar direction as follows:

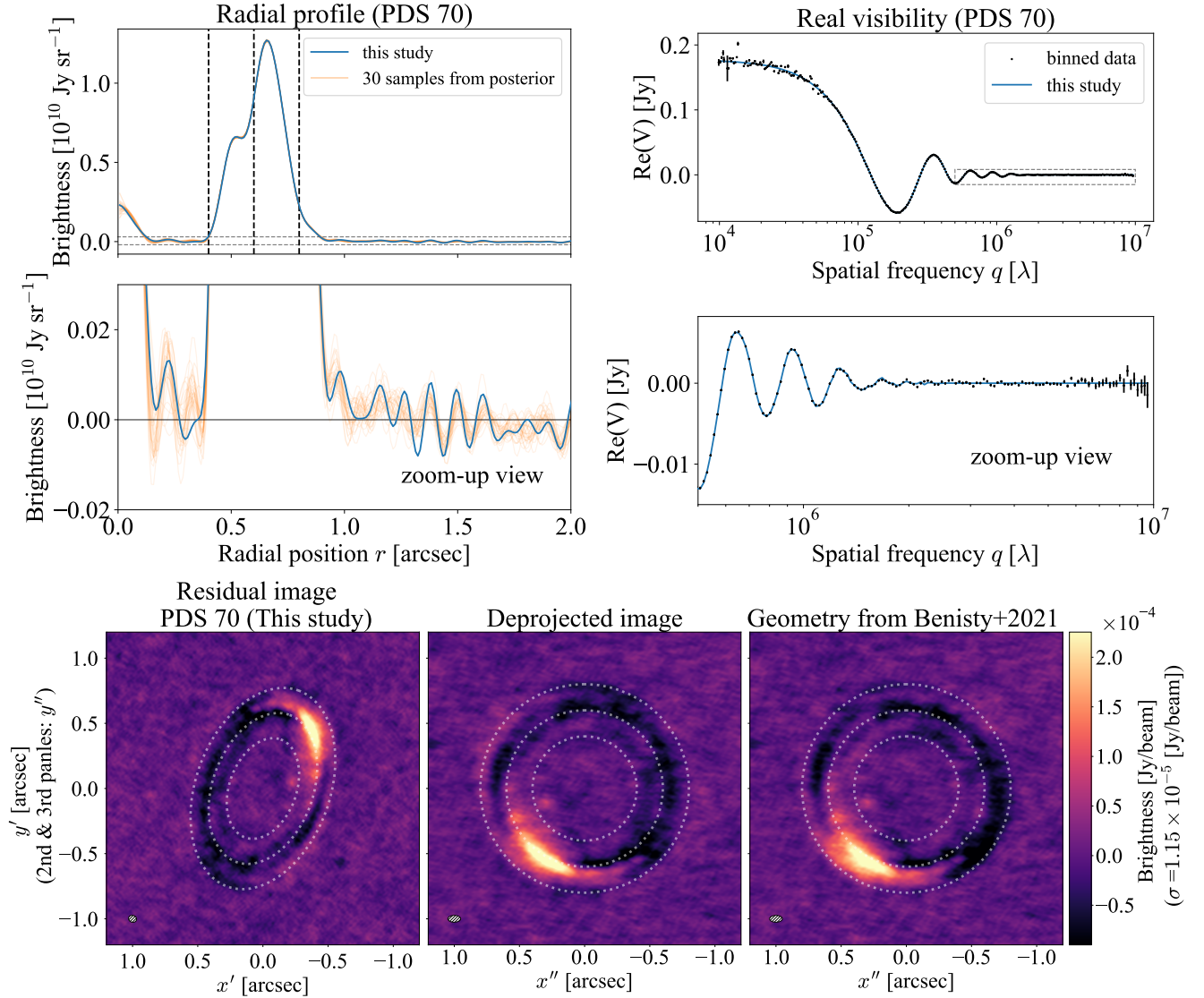


Figure 15. Analysis for PDS 70. The format of the figure is the same as that of Fig. 13 except that the zoom-up view of the radial profile is shown in the linear scale. The reference lines are shown at $r = (0.4'', 0.6'', 0.8'')$ in the panels of the brightness profile and the residual maps. The synthesized beam sizes before and after deprojection are $(0.063'', 0.053'')$ and $(0.098'', 0.053'')$, respectively.

$$I_{\text{crescent}}(r, \phi) = I_{\text{crescent,amp}} \exp\left(-\left(\frac{(r-r_0)^2}{2\sigma_r^2}\right)^\alpha\right) \times \left(\frac{\exp(\kappa \cos(\phi - \mu))}{2\pi I_{n=0, \text{modified}}(\beta)} - \frac{1}{2\pi}\right), \quad (48)$$

where $I_{n, \text{modified}}(\beta)$ is the modified Bessel function of the first kind, $I_{\text{crescent,amp}}$ determines the amplitude of the crescent model, α is the exponent for the super Gaussian function, r_0 denotes the radial position for the crescent peak, σ_r indicates the radial width, κ specifies the azimuthal concentration, and μ determines the azimuthal location for the peak. The integral of the equation (48) along ϕ is zero, consistent with the construction of the residual image. The adopted model does not perfectly explain the crescent feature, but it is still useful for removing the bright component of the crescent. We fitted the model to the residual image in Sec 7.1, and the optimization was performed us-

ing `curve_fit` in `scipy`, yielding $(I_{\text{crescent,amp}}, \alpha, r_0, \sigma_r, \kappa, \mu) = (0.339 \text{ mJy/beam}, 0.793, 0.623'', 0.0723'', 5.88, -0.960 \text{ rad})$.

The lower panels in Figure 16 shows the residual images with the crescent models being subtracted in deprojected coordinate and the polar coordinate. We confirmed the presence of emission from PDS 70 c, while no significant detection for point-source emission was observed from PDS 70 b and its L5 point.

The subtraction process facilitated the identification of faint features in the residual images. We here briefly summarize their characteristics in the following:

(a) *Crescent is trailing*: As evident in the residual images with and without the subtraction in a polar coordinate, the observed crescent feature appears to exhibit a trailing pattern rather than circular shape. Here the disc rotation is clockwise, as we deduced from the velocity gradient map and the emission surface of the gas (Keppler et al. 2019). One interesting possibility is that a planetary gravity perturbs the crescent, resulting in the trailing shape.

(b) *Extended emission from crescent to PDS 70 c?*: We also observed that the flux excess at the crescent appears to extend toward the vicinity of PDS 70 c. One interesting possibility is that this feature implies a dust accretion flow to PDS 70 c from the outer disc.

(c) *Faint arm stemming from inner end of crescent*: A faint leading arm is observed to stem from the inner end of the crescent. This feature is also faintly discernible in the residual images without the subtraction.

(d) *Arm/Vortex-like structure around $0.75''$ stemming from crescent*: Another arc or vortex stemming from the crescent was observed, and it is radially concentrated around $r = 0.75''$. This feature is more pronounced in our updated residual image than the residual image based on the geometry from Benisty et al. (2021), due to the difference in the adopted geometries, especially in PA (also see Appendix F). The base of the arm is overlapped with the crescent model in the lower panels, potentially resulting in over-subtraction of the flux of the arm near the region.

(e) *Depletion near crescent*: We observed a strong negative feature near the crescent feature. It is visible on both residual maps, with and without subtraction of crescent models. This might be attributed to the counter effect of the dust concentration at the crescent.

(f) *Blobs in crescent*: We observed blobs inside the crescent, suggesting that the crescent is not unimodal but rather multimodal. There are at least two visible blobs in the crescent feature, located at $(x'', y'') = (0.53'', -0.24'')$ and $(x'', y'') = (0.58'', 0.02'')$ in the deprojected frames. The blobs exhibited an excess of $3\text{--}5\sigma$ with respect to the background level of the the crescent feature.

All of the features appear to be associated with the crescent feature. In Appendix F, we presented the difference in the residual images derived with geometries from our study and Benisty et al. (2021). Apart from the possible arm (d), all of the features were similarly identified in both of the residual images, reinforcing the coherence of the signals within the current data set.

Substructures in the PDS 70 disc have also been reported in different bands. The arm-like structure was found in the northwest direction in the near-infrared bands, near the crescent observed in the radio band (Müller et al. 2018; Juillard et al. 2022; Christiaens et al. 2024). Juillard et al. (2022) investigated the motion of the arm over six years. However, they did not find the anticipated rotation expected if the arm were comoving with PDS 70 c. This suggests that the arm might be a vortex rather than a spiral. Indeed, as demonstrated in Fig 4 of Marr & Dong (2022), such a circular arc/vortex may be misunderstood as one-armed spiral in the near-infrared band, where the disc thickness becomes significant. On the other hand, the crescent observed in the radio band in this study displays a trailing pattern, which is unlikely to result from a purely geometric effect. One alternative explanation for the arm might be the presence of an undetected companion in the outer disc. Christiaens et al. (2024) set an upper limit on the mass of such a potential planet, obtaining limits 1–4 Jupiter masses from the JWST observation in the near-infrared band.

Additionally, Christiaens et al. (2024) identified a spiral accretion stream to the vicinity of PDS 70 c in the near infrared band by the JWST observation. It might be related to the extended emission excess near PDS 70 in the radio band (feature b in this study), but the connection between the features is not clear at this point.

The present findings rely on the assumption that the disc structure is well approximated by the thin axisymmetric disc with a single set of geometric parameters. This underlying assumption may be inadequate if the inner and outer discs are misaligned, or the disc thickness is substantial. In addition, there is still room for improvement for the modeling of the crescent feature, and the model might be better to

be incorporated into the framework of the axisymmetric modelling. A comprehensive analysis of these points is essential for the robust detection of the features, and we leave this issues for the future study.

8 SUMMARY

This study proposed a scheme for estimating the geometry, hyper-parameters, central position, and brightness profiles assuming a geometrically thin disc in radio interferometry. Our approach is less susceptible to human biases due to manual tuning of parameters in contrast to frank, where the non-linear parameters need to be determined a priori.

Simulating observations for an axisymmetric disc, we demonstrated that the proposed method can successfully retrieve geometric parameters in a more precise manner than that using Gaussian fitting. Additionally, we performed injection and recovery tests for the non-axisymmetric perturbations to the simulated data, and showed that the proposed method can reasonably recover the injected structures. However, the estimated geometric parameters were slightly shifted from the assumed values. This is attributed to the degeneracy between the non-axisymmetric perturbations and the residuals caused by biases in the geometric parameters.

The model was then applied to the real data for Elias 20 and AS 209, and the ability of the method to determine the disc geometry and brightness profile was demonstrated especially for Elias 20. Moreover, the data for the continuum emission of the PDS 70 were reanalyzed with our method. Our methods successfully identified the circumplanetary emission from PDS 70 c and the crescent feature. Furthermore, we tentatively identified several new features, including trailing nature of crescent, extended emission near PDS 70 c, arm-like structures, dust depletion near crescent, and blobs. The origins of these features are unclear, and a further modeling is needed.

The current methodology can be applied to any type of continuum data in radio interferometric observations, and future studies will analyse the DSHARP data (Andrews et al. 2018) to explore their non-axisymmetric structures.

One of the notable strengths of this research is its capacity to adapt to more complex problems involving a multitude of geometric or hyper parameters, which are difficult to adjust manually. Below, we present the potential directions for extending our study:

- The model can be extended to cover multiple rings/gaps with different central positions, inclinations, or position angles as observed in GW Ori (Bi et al. 2020). This can be realized by considering multiple sets of geometric parameters, each of which is applied to one of the separate ranges.
- The current model can be extended to incorporate frequency-dependent radial profiles for multi-frequency data. We can straightforwardly expand our formulation to include the frequency dependence in a linearized formulation by Taylor expansions in a manner similar to multi-frequency CLEAN (Rau & Cornwell 2011):

$$I(r; \nu) = \sum_t \left(\frac{\nu - \nu_0}{\nu_0} \right)^t I_{t, \nu_0}(r), \quad (49)$$

where $I(r; \nu)$ is the radial brightness at the frequency ν , ν_0 is the reference frequency, and $I_{t, \nu_0}(r)$ is the coefficient of t -th Taylor expansions. With this expression, we can analytically derive the model parameters in the same manner as that presented in the current study.

- The current study focuses on the axisymmetric component $I(r)$; however, we may be able to include non-axisymmetric components

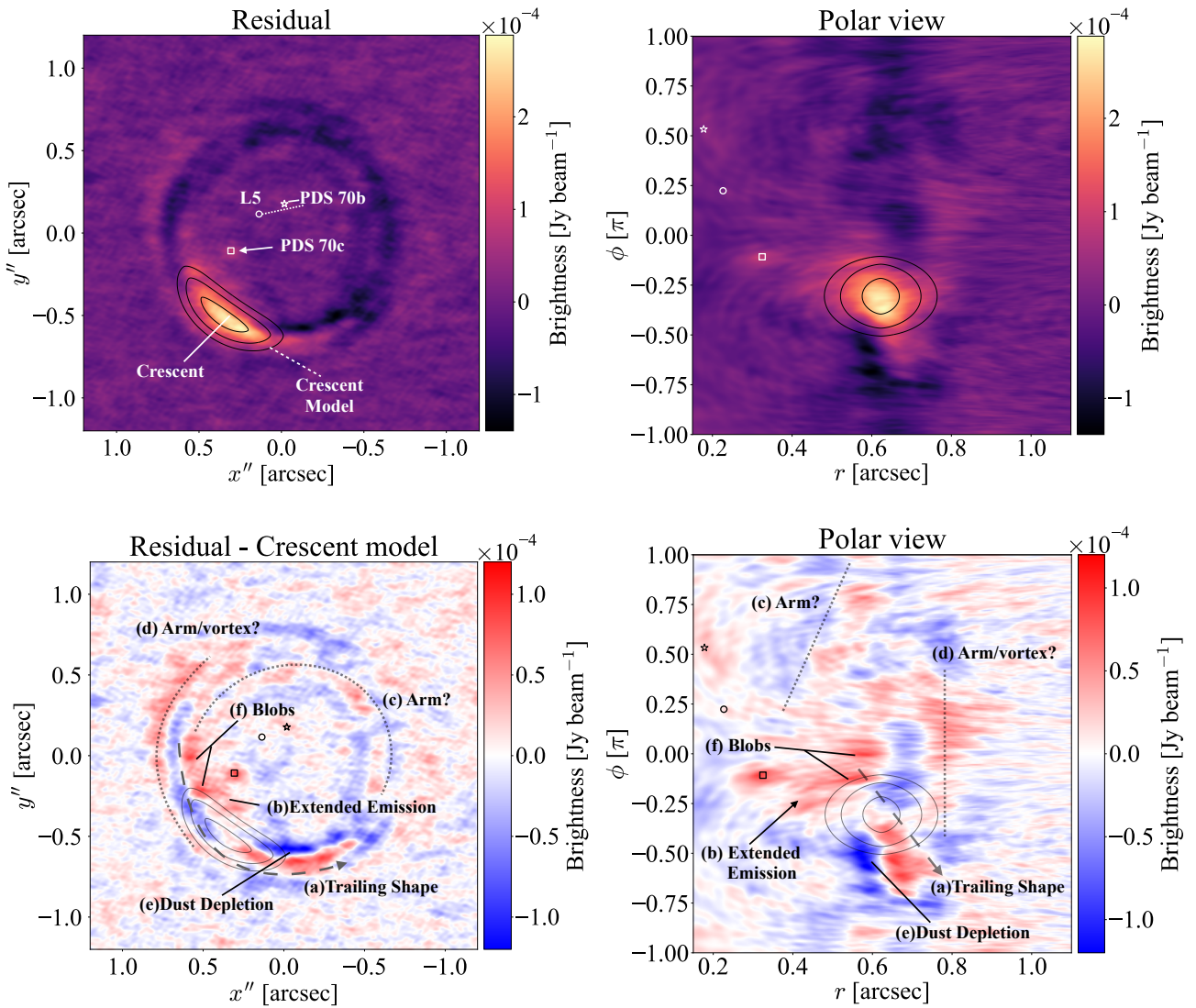


Figure 16. The residual images for PDS 70 in the deprojected coordinate (left panels) and in the polar coordinate (right panels). The upper and lower panels show the residual images without and with the subtraction of the crescent models, as discussed in Sec 7.2. Contours of the crescent models are also overlaid as black lines with levels of (50, 100, 200) $\mu\text{Jy beam}^{-1}$. Annotations in the upper panels show the locations of the protoplanets and the L₅ point of PDS 70 b, as provided by Balsalobre-Ruza et al. (2023). Annotations in the lower panels specify the possible structures identified in this system, and the further explanations for the features (a)–(f) are provided in Sec 7.2.

($I_m(r)$, $\phi_m(r)$) in the model, and directly solve them. Such modelling efforts might ease the degeneracy between the geometric parameters and non-axisymmetric structures.

- Although we created residual images with CLEAN, we can also use other imaging techniques, for example, sparse modelling, to produce images alternatively (Honma et al. 2014; Nakazato & Ikeda 2020; Yamaguchi et al. 2020; Aizawa et al. 2020). In particular, sparse modelling favours solutions with many zeros, and it can achieve better angular resolutions; thus it will be helpful for resolving or identifying the new substructures, including spirals and circumplanetary emission.

- A more comprehensive analysis for PDS 70 will be essential for the reliable detection of the discovered structures. Incorporating the crescent model into our model will allow us to separate the residual signal accurately. Exploring more complicated models that consider

a disc thickness or misalignment between inner and outer disc planes will be rewarding. The multi-wavelength data will be also useful for assessing the consistency of the signals across different wavelengths.

These will be considered in future studies.

ACKNOWLEDGMENTS

We thank Kazuhiro Kanagawa, Hongping Deng, and Ruobing Dong for their insightful discussions. We also thank the anonymous referee for constructive comments, which significantly improved the quality of the paper. Furthermore, we would like to extend our gratitude to Myriam Benisty for sharing the data for PDS 70 used in this research.

This work was supported by JSPS KAKENHI grant numbers 17H01103, 18H05441, and 23K03463. M.A. is supported by Special

Postdoctoral Researcher Program at RIKEN. T.M. is supported by Yamada Science Foundation Overseas Research Support Program.

Data analysis was in part carried out on the Multi-wavelength Data Analysis System operated by the Astronomy Data Center (ADC), National Astronomical Observatory of Japan. This research is based on the following ALMA data #2013.1.00226.S, #2015.1.00486.S, and #2016.1.00484.L. ALMA is a partnership of ESO (representing its member states), NSF (USA) and NINS (Japan), together with NRC (Canada), MOST and ASIAA (Taiwan), and KASI (Republic of Korea), in cooperation with the Republic of Chile. The Joint ALMA Observatory is operated by ESO, AUI/NRAO and NAOJ.

Software: *Astropy* (Astropy Collaboration et al. 2013), *CASA* (McMullin et al. 2007), *corner* (Foreman-Mackey 2016), *emcee* (Foreman-Mackey et al. 2013), *Jupyter Notebook* (Kluyver et al. 2016), *Matplotlib* (Hunter 2007), *NumPy* (van der Walt et al. 2011), *Pandas* (McKinney 2010), *PyNUFFT* (Lin 2018; Lin & Chung 2017), *SciPy* (Virtanen et al. 2020).

DATA AVAILABILITY

The DSHARP data used in this article are available in the DSHARP Data Release at <https://bulk.cv.nrao.edu/alldata/lp/DSHARP>. The data underlying this article will be shared on reasonable request to the corresponding author.

REFERENCES

- ALMA Partnership et al., 2015, *ApJ*, 808, L3
- Aizawa M., Suto Y., Oya Y., Ikeda S., Nakazato T., 2020, *ApJ*, 899, 55
- Andrews S. M., et al., 2018, *ApJ*, 869, L41
- Andrews S. M., et al., 2021, *ApJ*, 916, 51
- Astropy Collaboration et al., 2013, *A&A*, 558, A33
- Balsalobre-Ruza O., et al., 2023, *A&A*, 675, A172
- Benisty M., et al., 2021, *ApJ*, 916, L2
- Bi J., et al., 2020, *ApJ*, 895, L18
- Binney J., Tremaine S., 2008, *Galactic Dynamics: Second Edition*
- Casassus S., Cárcamo M., 2022, *MNRAS*, 513, 5790
- Christiaens V., et al., 2024, *arXiv e-prints*, p. arXiv:2403.04855
- Currie T., et al., 2022, *Nature Astronomy*, 6, 751
- Czekala I., et al., 2021, *ApJS*, 257, 2
- Dong R., Zhu Z., Fung J., Rafikov R., Chiang E., Wagner K., 2016, *ApJ*, 816, L12
- Dong R., Li S., Chiang E., Li H., 2018, *ApJ*, 866, 110
- Flock M., Ruge J. P., Dzyurkevich N., Henning T., Klahr H., Wolf S., 2015, *A&A*, 574, A68
- Foreman-Mackey D., 2016, *The Journal of Open Source Software*, 1, 24
- Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, *PASP*, 125, 306
- Goldreich P., Tremaine S., 1980, *ApJ*, 241, 425
- Grady C. A., et al., 2013, *ApJ*, 762, 48
- Guzmán V. V., et al., 2018, *ApJ*, 869, L48
- Hammond I., Christiaens V., Price D. J., Toci C., Pinte C., Juillard S., Garg H., 2023, *MNRAS*, 522, L51
- Hashimoto J., Muto T., Dong R., Liu H. B., van der Marel N., Francis L., Hasegawa Y., Tsukagoshi T., 2021, *ApJ*, 911, 5
- Honma M., Akiyama K., Uemura M., Ikeda S., 2014, *PASJ*, 66, 95
- Huang J., et al., 2018a, *ApJ*, 869, L42
- Huang J., et al., 2018b, *ApJ*, 869, L43
- Hunter J. D., 2007, *Computing in Science and Engineering*, 9, 90
- Jennings J., Booth R. A., Tazzari M., Rosotti G. P., Clarke C. J., 2020, *MNRAS*, 495, 3209
- Jennings J., Booth R. A., Tazzari M., Clarke C. J., Rosotti G. P., 2022a, *MNRAS*, 509, 2780
- Jennings J., Tazzari M., Clarke C. J., Booth R. A., Rosotti G. P., 2022b, *MNRAS*, 514, 6053
- Juillard S., Christiaens V., Absil O., 2022, *A&A*, 668, A125
- Kanagawa K. D., et al., 2021, *ApJ*, 909, 212
- Kawahara H., Masuda K., 2020, *ApJ*, 900, 48
- Kawahara H., Kawashima Y., Masuda K., Crossfield I. J. M., Pannier E., van den Bekerom D., 2022, *ApJS*, 258, 31
- Keppler M., et al., 2018, *A&A*, 617, A44
- Keppler M., et al., 2019, *A&A*, 625, A118
- Kluyver T., et al., 2016, in , IOS Press. pp 87–90, doi:10.3233/978-1-61499-649-1-87
- Lin J.-M., 2018, *Journal of Imaging*, 4
- Lin J.-M., Chung H.-W., 2017, *Building Bridges in Medical Sciences*
- Lin D. N. C., Papaloizou J., 1979, *MNRAS*, 186, 799
- Marr M., Dong R., 2022, *ApJ*, 930, 80
- McKinney W., 2010, in van der Walt S., Millman J., eds, *Proceedings of the 9th Python in Science Conference*. pp 51 – 56
- McMullin J. P., Waters B., Schiebel D., Young W., Golap K., 2007, in Shaw R. A., Hill F., Bell D. J., eds, *Astronomical Society of the Pacific Conference Series Vol. 376, Astronomical Data Analysis Software and Systems XVI*. p. 127
- Müller A., et al., 2018, *A&A*, 617, L2
- Muto T., et al., 2012, *ApJ*, 748, L22
- Nakazato T., Ikeda S., 2020, *PRIISM: Python module for Radio Interferometry Imaging with Sparse Modeling*, *Astrophysics Source Code Library*, record ascl:2006.002 (ascl:2006.002)
- Okuzumi S., Momose M., Sirono S.-i., Kobayashi H., Tanaka H., 2016, *ApJ*, 821, 82
- Pinte C., et al., 2018, *ApJ*, 860, L13
- Pinte C., et al., 2020, *ApJ*, 890, L9
- Pinte C., et al., 2023, *arXiv e-prints*, p. arXiv:2301.08759
- Rau U., Cornwell T. J., 2011, *A&A*, 532, A71
- Reggiani M., et al., 2014, *ApJ*, 792, L23
- Speedie J., Dong R., 2022, *arXiv e-prints*, p. arXiv:2211.05757
- Takahashi S. Z., Inutsuka S.-i., 2014, *ApJ*, 794, 55
- Tazzari M., et al., 2017, *A&A*, 606, A88
- Thompson A. R., Moran J. M., Swenson George W. J., 2017, *Interferometry and Synthesis in Radio Astronomy*, 3rd Edition, doi:10.1007/978-3-319-44431-4.
- Virtanen P., et al., 2020, *Nature Methods*, 17, 261
- Wagner K., Apai D., Kasper M., Robberto M., 2015, *ApJ*, 813, L2
- Wagner K., et al., 2018, *ApJ*, 863, L8
- Wagner K., et al., 2023, *Nature Astronomy*, 7, 1208
- Yamaguchi M., et al., 2020, *ApJ*, 895, 84
- Zhang K., Blake G. A., Bergin E. A., 2015, *ApJ*, 806, L7
- van der Walt S., Colbert S. C., Varoquaux G., 2011, *Computing in Science and Engineering*, 13, 22

APPENDIX A: EFFICIENT COMPUTATION OF EVIDENCE GIVEN GEOMETRY

The direct computation of $\mathcal{N}(d|0, \bar{\Sigma}_d + \bar{H}\bar{\Sigma}_a\bar{H}^T)$ in equation (32) requires the inverse matrix of $\bar{\Sigma}_d + \bar{H}\bar{\Sigma}_a\bar{H}^T$, whose computational cost can be $O(M^3)$ with M the number of data points. As $O(M^3)$ can be very large for the usual case in interferometric observations $M > 10^5-6$, we aimed at reducing this computation through the formula transformation.

Using the Woodbury identity $(A+UWV)^{-1} = A^{-1} - A^{-1}U(W^{-1} + VA^{-1}U)^{-1}VA^{-1}$, we obtain

$$(\bar{\Sigma}_d + \bar{H}\bar{\Sigma}_a\bar{H}^T)^{-1} = \bar{\Sigma}_d^{-1} - \bar{\Sigma}_d^{-1}\bar{H}\bar{\Sigma}_a^{-1}\bar{H}^T\bar{\Sigma}_d^{-1}. \quad (A1)$$

With this equation, we can compute the log probability for $p(d|g, \theta)$

as follows:

$$\begin{aligned} -2 \log(p(\mathbf{d}|\mathbf{g}, \boldsymbol{\theta})) &= \log |\det((\bar{\boldsymbol{\Sigma}}_d + \bar{\mathbf{H}}\bar{\boldsymbol{\Sigma}}_a\bar{\mathbf{H}}^T))| + \mathbf{d}^T \bar{\boldsymbol{\Sigma}}_d^{-1} \mathbf{d} \\ &\quad - \mathbf{d}^T \bar{\boldsymbol{\Sigma}}_d^{-1} \bar{\mathbf{H}}\bar{\boldsymbol{\Sigma}}_a \bar{\mathbf{H}}^T \bar{\boldsymbol{\Sigma}}_d^{-1} \mathbf{d}. \end{aligned} \quad (\text{A2})$$

Using another identity $\det(A + UWV) = \det(W^{-1} + VA^{-1}U)\det(W)\det(A)$, we can compute the first term as follows:

$$\begin{aligned} \log |\det((\bar{\boldsymbol{\Sigma}}_d + \bar{\mathbf{H}}\bar{\boldsymbol{\Sigma}}_a\bar{\mathbf{H}}^T))| &= \log |\det(\bar{\boldsymbol{\Sigma}}_a^{-1} \mathbf{d})| \\ &\quad + \log |\det(\bar{\boldsymbol{\Sigma}}_d)| + \log |\det(\bar{\boldsymbol{\Sigma}}_a)|. \end{aligned} \quad (\text{A3})$$

Consequently, we find the following expression:

$$\begin{aligned} -2 \log(p(\mathbf{d}|\mathbf{g}, \boldsymbol{\theta})) &= \log |\det(\bar{\boldsymbol{\Sigma}}_a^{-1} \mathbf{d})| + \log |\det(\bar{\boldsymbol{\Sigma}}_d)| \\ &\quad - \mathbf{d}^T \bar{\boldsymbol{\Sigma}}_d^{-1} \bar{\mathbf{H}}\bar{\boldsymbol{\Sigma}}_a \bar{\mathbf{H}}^T \bar{\boldsymbol{\Sigma}}_d^{-1} \mathbf{d} + c, \end{aligned} \quad (\text{A4})$$

where c corresponds to the terms independent on $(\mathbf{g}, \boldsymbol{\theta})$. In the case of $M \gg N$, this incurs a computational cost that is much lesser than that using equation (32).

APPENDIX B: REVISITING VISIBILITY BINNING WITH UNIFORM AND LOG GRIDDING

The computational cost strongly depends on the number of visibilities. Here, we discuss the data binning in an interferometric observation assuming linear and log grids.

B1 Formulation

We consider the data \mathbf{d}_{obs} with the data weights \mathbf{w}_{obs} . We binned them on a grid specified by bin edges $\{(E_i, E_j)\}$ with $i = -(N_{\text{bin}} + 1), -N_{\text{bin}}, \dots, N_{\text{bin}}, N_{\text{bin}} + 1$, wherein N_{bin} determines the number of grids. For a cell with $E_{i-1} < u_k < E_i$ and $E_{j-1} < v_k < E_j$, we define the summing operation for a vector \mathbf{x} with the same length as \mathbf{d} as follows:

$$\text{Bin}(\mathbf{x}, i, j) = \sum_{\substack{E_{i-1} < u_k < E_i \\ E_{j-1} < v_k < E_j}} x_k, \quad (\text{B1})$$

where E_i denotes either $E_{\text{uniform}, i}$ or $E_{\text{log}, i}$. Bin edges for a uniform gridding are defined as follows:

$$E_{\text{uniform}, i} = \begin{cases} x_{\min} + (i-1) \left(\frac{x_{\max} - x_{\min}}{N_{\text{bin}}} \right) & i = 1, 2, \dots, N_{\text{bin}} + 1 \\ 0 & i = 0 \\ -x_{\min} + (i+1) \left(\frac{x_{\max} - x_{\min}}{N_{\text{bin}}} \right) & i = -1, -2, \dots, -(N_{\text{bin}} + 1), \end{cases} \quad (\text{B2})$$

where (x_{\min}, x_{\max}) are values that determine the limits of the grid.

Similarly, we define bin edges for a log grid as follows:

$$E_{\text{log}, i} = \begin{cases} x_{\min} \times 10^{(i-1)\Delta_{\text{log}}} & i = 1, 2, \dots, N_{\text{bin}} + 1 \\ 0 & i = 0 \\ -x_{\min} \times 10^{-(i+1)\Delta_{\text{log}}} & i = -1, -2, \dots, -(N_{\text{bin}} + 1) \end{cases} \quad (\text{B3})$$

where we define a spacing Δ_{log} for the log grid as follows:

$$\Delta_{\text{log}} = \frac{\log_{10}(x_{\max}) - \log_{10}(x_{\min})}{N_{\text{bin}}}. \quad (\text{B4})$$

In each cell, we computed a weighted average \mathbf{d}_{bin} , a total sum of weights \mathbf{w}_{bin} , standard deviations for the noise $\sigma_{\text{bin}, i, j}$, and weighted average of spatial frequencies $(\mathbf{u}_{\text{bin}}, \mathbf{v}_{\text{bin}})$ as follows:

$$w_{\text{bin}, i, j} = \text{Bin}(\mathbf{w}, i, j) \quad (\text{B5})$$

$$\sigma_{\text{bin}, i, j} = \frac{1}{\sqrt{w_{\text{bin}, i, j}}} \quad (\text{B6})$$

$$d_{\text{bin}, i, j} = \frac{\text{Bin}(\mathbf{d} \circ \mathbf{w}, i, j)}{w_{\text{bin}, i, j}} \quad (\text{B7})$$

$$u_{\text{bin}, i, j} = \frac{\text{Bin}(\mathbf{u} \circ \mathbf{w}, i, j)}{w_{\text{bin}, i, j}} \quad (\text{B8})$$

$$v_{\text{bin}, i, j} = \frac{\text{Bin}(\mathbf{v} \circ \mathbf{w}, i, j)}{w_{\text{bin}, i, j}}, \quad (\text{B9})$$

where \circ denotes the Hadamard product for vectors \mathbf{x} and \mathbf{y} : $(\mathbf{x} \circ \mathbf{y})_i = x_i y_i$.

B2 Quantifying binning errors with uniform and log grids

We quantified binning errors by simulating an observation of a protoplanetary disc. The simulation setup was same as simulated case of AS 209 in Sec 4.

The data were binned with $N_{\text{bin}} = (125, 250, 500, 1000, 2000, 4000)$, $x_{\min} = 10^2 \lambda$, and $x_{\max} = 10^7 \lambda$. At the binned spatial frequencies $(\mathbf{u}_{\text{bin}}, \mathbf{v}_{\text{bin}})$, we computed model visibilities $\mathbf{d}_{\text{bin}, \text{model}}$, and quantified a binning error relative to an observational error as follows:

$$\chi_{\text{bin}, i, j} = \frac{d_{\text{bin}, i, j} - d_{\text{model}, i, j}}{\sigma_{\text{bin}, i, j}}, \quad (\text{B10})$$

which is the ratio of the binning error with respect to the noise amplitude. Additionally, we also computed the mean squared binning errors $\chi_{\text{bin}, \text{error}, \text{mean}}^2$:

$$\chi_{\text{bin}, \text{mean}}^2 = \frac{1}{N_{\text{d}, \text{bin}}} \sum \chi_{\text{bin}, i, j}^2, \quad (\text{B11})$$

where $M_{\text{d}, \text{bin}}$ is the number of data after binning.

We computed the binning errors for the linear and log grids by adopting $N_{\text{bin}} = (125, 250, 500, 1000, 2000, 4000)$. The left panel in Fig. B1 shows the relation between the mean squared binning errors $\chi_{\text{bin}, \text{mean}}^2$ and the number of binned data $M_{\text{d}, \text{bin}}$, which determines the computational time. As N_{bin} increases, the binning errors became small, and $M_{\text{d}, \text{bin}}$ became large as expected, thereby increasing the computational cost. On average, uniform and log grids yielded similar binning errors with $M_{\text{d}, \text{bin}}$ being fixed as shown in Fig. B1.

The right panel in Fig. B1 shows the binning errors $|\chi_{\text{bin}}|$ at deprojected spatial frequencies q . To obtain a similar number of data points $M_{\text{d}, \text{bin}} \approx 10^5$, we assumed $N_{\text{bin}} = 2000$ for the uniform grid and $N_{\text{bin}} = 500$ for the log grid (see left panel in Fig. B1). The binning errors in the uniform grid increased at low spatial frequencies, and they reduced at higher frequencies. With $N_{\text{bin}} = 2000$ for the uniform grid, the binning error can reach 40 per cent at most, and such errors might be problematic to estimations on large-scale emissions. The binning errors increase at small spatial frequencies because the binning errors in the uniform grid are proportional to the first derivative of the radial visibility profile, which tends to be large at smaller scales. This is supported by the small binning errors at large spatial frequencies. In the case of the log grid, the binning errors are suppressed at small spatial frequencies, in contrast to the uniform grid. This accords with the narrow bin widths at small spatial frequencies in the log grid; the grid is sufficiently fine to resolve the visibility profile. At the middle to high spatial frequencies, the binning errors increased and decreased, with a peak at $10^6 \text{ M}\lambda$. At

the higher spatial frequencies, the grids become coarse, while the first derivative of the radial visibility profile becomes small. These two different trends result in this dependency.

Comparing the two grids, the log grid tends to yield moderate binning errors, whereas the uniform grid yields errors in a broad range. Considering the robustness, we adopted the log grid in the analysis of the main text, because the large binning errors at small spatial frequencies for the uniform grid can degrade the estimated accuracy of the flux on a large scale.

B3 Dependence of geometric parameters and hyperparameters on number of grids

In the main text, we adopted the log grid with $N_{\text{bin}} = 500$. Here, we show that the choice of N_{bin} does not affect the estimation on non-linear parameters. Simulating the same observational setup and the brightness profile as in the simulated case for AS 209 in Sec 4, we derived the geometric parameters by varying the number of grids, $N_{\text{bin}} = 125, 250, 500, \text{ and } 1000$. We injected $(\Delta x_{\text{cen}}, \Delta y_{\text{cen}}, \cos i, \text{PA}) = (0'', 0'', 0.75, 45^\circ)$. Fig. B2 shows the results for recovered geometric parameters and hyperparameters. The results appear to be unaffected by the choice of N_{bin} .

APPENDIX C: DEPENDENCE OF RESIDUAL IMAGES ON ROBUST PARAMETERS

We varied the robust parameters as $[0, 0.5, 1]$ in Briggs gridding, and investigated the dependence of the resultant residual images. Fig. C1 shows the deprojected residual images adopting three different robust parameters. They were constructed from the same residual visibilities as those used in the case of the even-symmetric spiral in Fig. 10. Visual inspections indicated that the robust parameter of 0.5 yielded the most balanced image in terms of both sensitivity and resolution.

APPENDIX D: RECOVERED SPATIAL SCALES FOR DIFFERENT ANGULAR RESOLUTIONS AND BRIGHTNESS PROFILES

We changed the length scales for the input profile $I(r)$ and observational spatial frequencies $\{u_j, v_j\}$ to investigate the variations of recovered length parameters γ . Specifically, we considered a modified brightness profile $I_a(r) = I(ar)$ and modified spatial frequencies $\{u_{b,j}, v_{b,j}\} = \{bu_j, bv_j\}$, where we introduced scaling parameters (a, b) . We considered the same observational setup and brightness profile as that of AS 209 in Sec 4, simulated the data, and derived the posterior distribution for parameters including γ . Fig. D1 shows the result with varying scale of brightness profile $a = (0.5, 1, 1.5)$ and fixed observed spatial frequency $b = 1$. The optimized length scale γ positively scaled with a . Beyond the optimized length scale, the power of model visibilities was suppressed and exhibited damped oscillations. Fig. D2 shows the result with varying scale for spatial frequency for the data $b = (0.5, 1, 1.5)$ and the same brightness profile $a = 1$. The optimized length scale was unchanged for $b = 1, 1.5$; however, it reduced for $b = 0.5$. This indicates that if the structure is already well resolved, the improvement in the angular resolution does not change the optimized length scale γ . However, if we adopt the worse angular resolution, the resolution cannot be sufficient to resolve the structure. This results in the larger value of γ . From this

discussion, we can conclude that the optimized length scale γ is determined by both the brightness profile and the UV-coverage.

APPENDIX E: MEAN AND VARIANCE OF RESIDUAL IMAGES

Assuming the simulated data for AS 209 in Sec 4, we produced 10 different residual images by drawing samples from posterior distribution of $(\mathbf{g}, \theta, \mathbf{a})$. We computed the mean and the standard deviation for 10 residual images, and calculated the differences of two random residual images from the mean image. Fig. E1 shows the result. The mean residual image was consistent with that shown in Fig. 4. The image for the standard deviation was nearly axisymmetric, and the amplitude was $10^{-5} \text{ Jy beam}^{-1}$ at most in the innermost part; whereas, it was $2-4 \times 10^{-6} \text{ Jy beam}^{-1}$, which is a few times smaller than the observational error in the residual image, for most of the disc plane. In the lower panels, we show the differences from the mean image, and they are nearly axisymmetric as well. The axisymmetry would arise from the larger variances of the brightness profile than those of geometry and hyperparameters. Considering the small amplitudes and axisymmetry of the differential images, it is reasonable to conclude that their effects would be negligible in searching for non-axisymmetric structures, at least in the current observational setup.

APPENDIX F: COMPARISON OF RESIDUAL IMAGES FROM TWO GEOMETRIES

The upper left panel in Fig F1 illustrates the difference in the deprojected residual images derived from geometries in our study and Benisty et al. (2021) (see Fig 15). We observed a noticeable $m = 2$ pattern in the difference, primarily attributed to the discrepancy in PA by 1° . Additionally, adopting the geometry in Benisty et al. (2021), we also subtracted the crescent model from the residual image following the same procedure described in Sec 7.2, while the model was re-optimized. In the upper right and lower panels in Fig F1 present the subtracted residual images in a deprojected coordinate and a polar coordinate.

The estimate of PA appears to be influenced by large-scale faint residuals rather than localized emission (also see Fig F1). Whether to adopt the removal of visibility of localized emission in a specific region as in Benisty et al. (2021) might affect the estimation. However, both the removal and non-removal can equally introduce bias, because the removal can result in the loss of information in the data. Thus, although our estimate is obtained by optimization of the data, we have not conclusively determined which estimation is more accurate.

A major difference is that the residual image derived from geometry in Benisty et al. (2021) exhibits a wide leading arm extended from the crescent to the point around $(x'', y'') = (-0.75'', 0.4'')$. In addition, the significance of an arm (feature (d) in Sec 7.2) appears to be diminished compared to the image in Fig 16. On the other hand, all of the other features (a-c,e,f) described in Sec 7.2 are consistently identified in the two residual images with different geometries.

This paper has been typeset from a \LaTeX file prepared by the author.

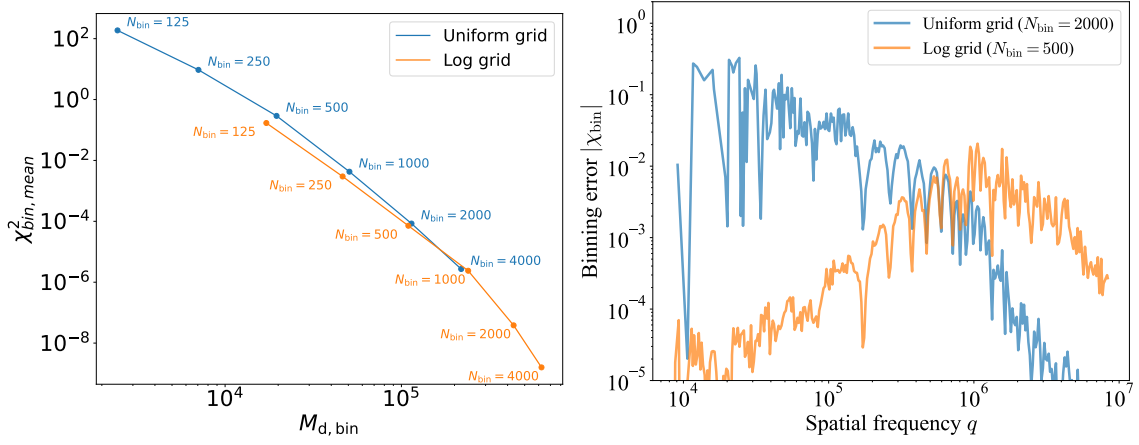


Figure B1. Comparison of binning errors for log (orange line) and linear (blue line) grids. (left) The average binning error $\chi_{bin, mean}^2$ and the number of binned data $M_{d, bin}$. We vary N_{bin} in the range of 125-4000. (right) Individual binning error $|\chi_{bin}|$ and deprojected spatial frequency q . We assume $N_{bin} = 2000$ for the uniform grid and $N_{bin} = 500$ for the log grid.

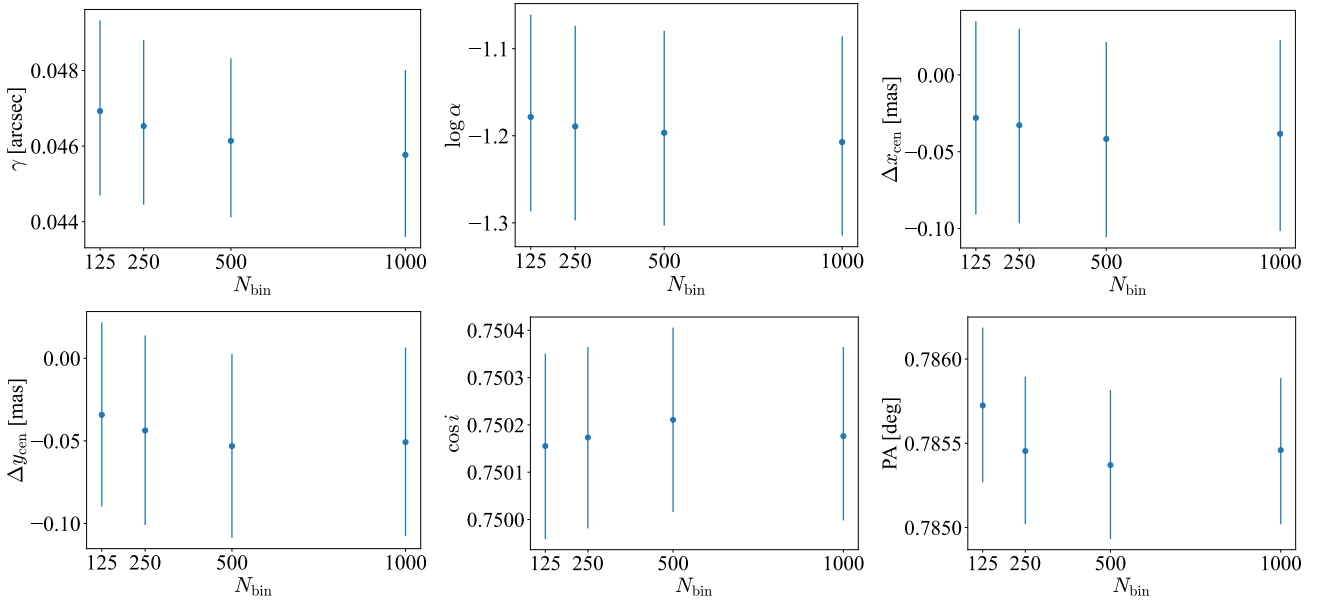


Figure B2. Recovered geometric and hyperparameters with different numbers of bins N_{bin} .

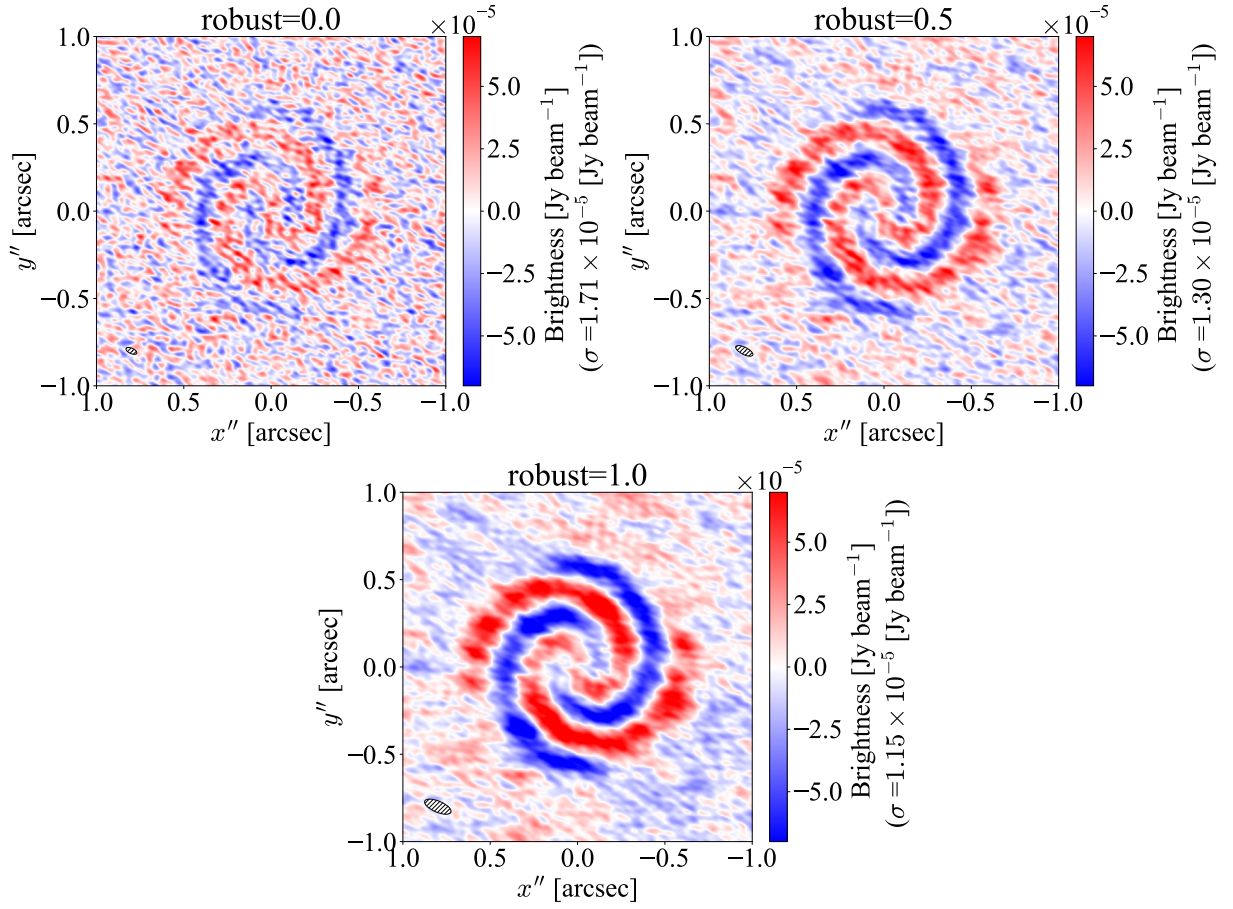


Figure C1. Dependence of residual images on robust parameters [0, 0.5, 1.0]. The synthesized beam sizes (0.067'', 0.031''), (0.11'', 0.041''), and (0.16'', 0.063'') are shown at the bottom left of panels.

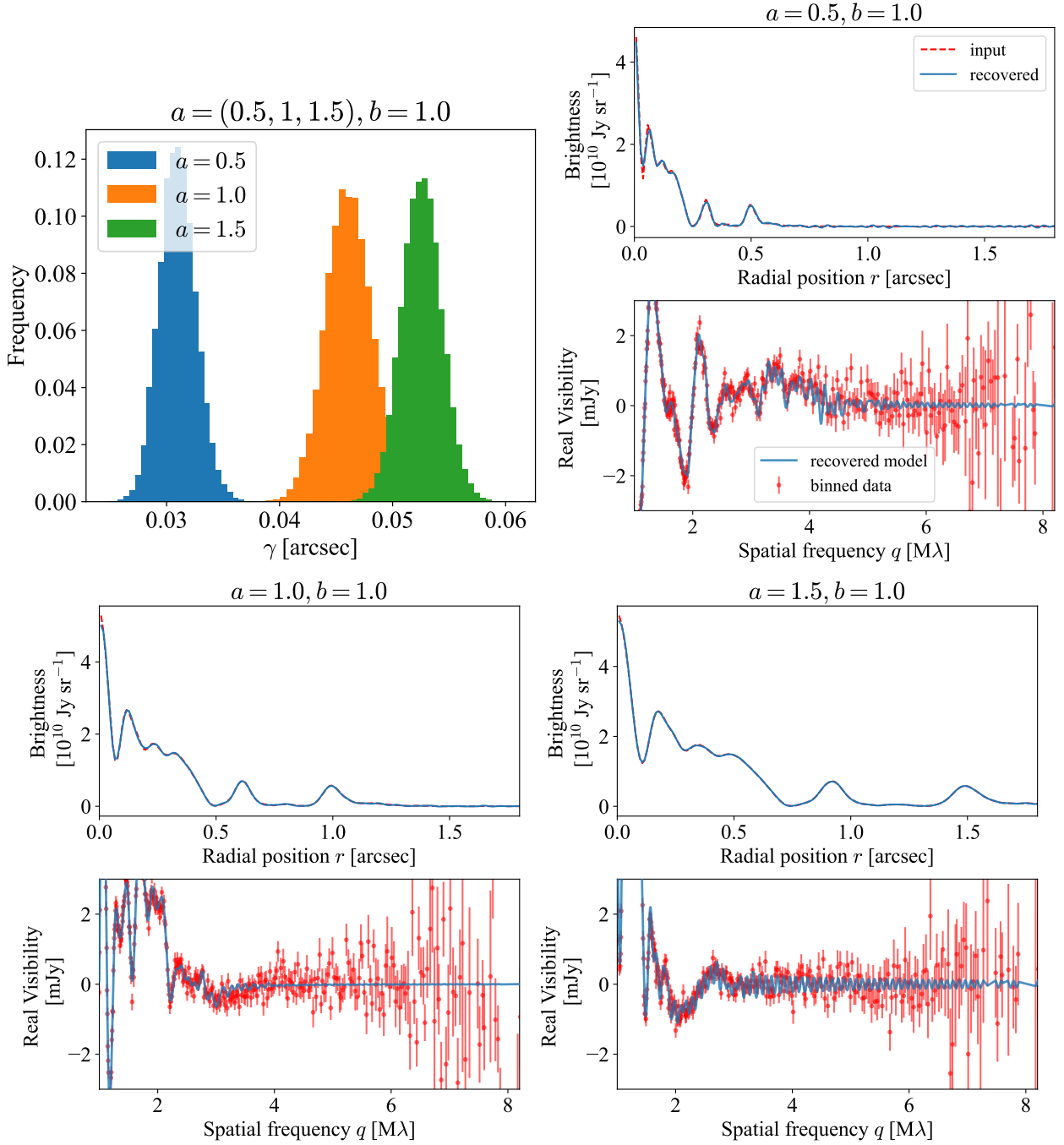


Figure D1. Simulation result with varying length scales of injected brightness profile $f_a(r) = f(ar)$ with $a = (0.5, 1, 1.5)$. (left upper) Posterior distribution of recovered length scale γ for $a = (0.5, 1, 1.5)$. (right upper) Result with $a = 0.5$. Injected and recovered brightness profile in the upper panel, and model and observed visibilities in the lower panel. The observed visibilities are binned with a log grid with $N_{\text{bin}} = 2000$. (lower panels) Result with $a = (1, 1.5)$.

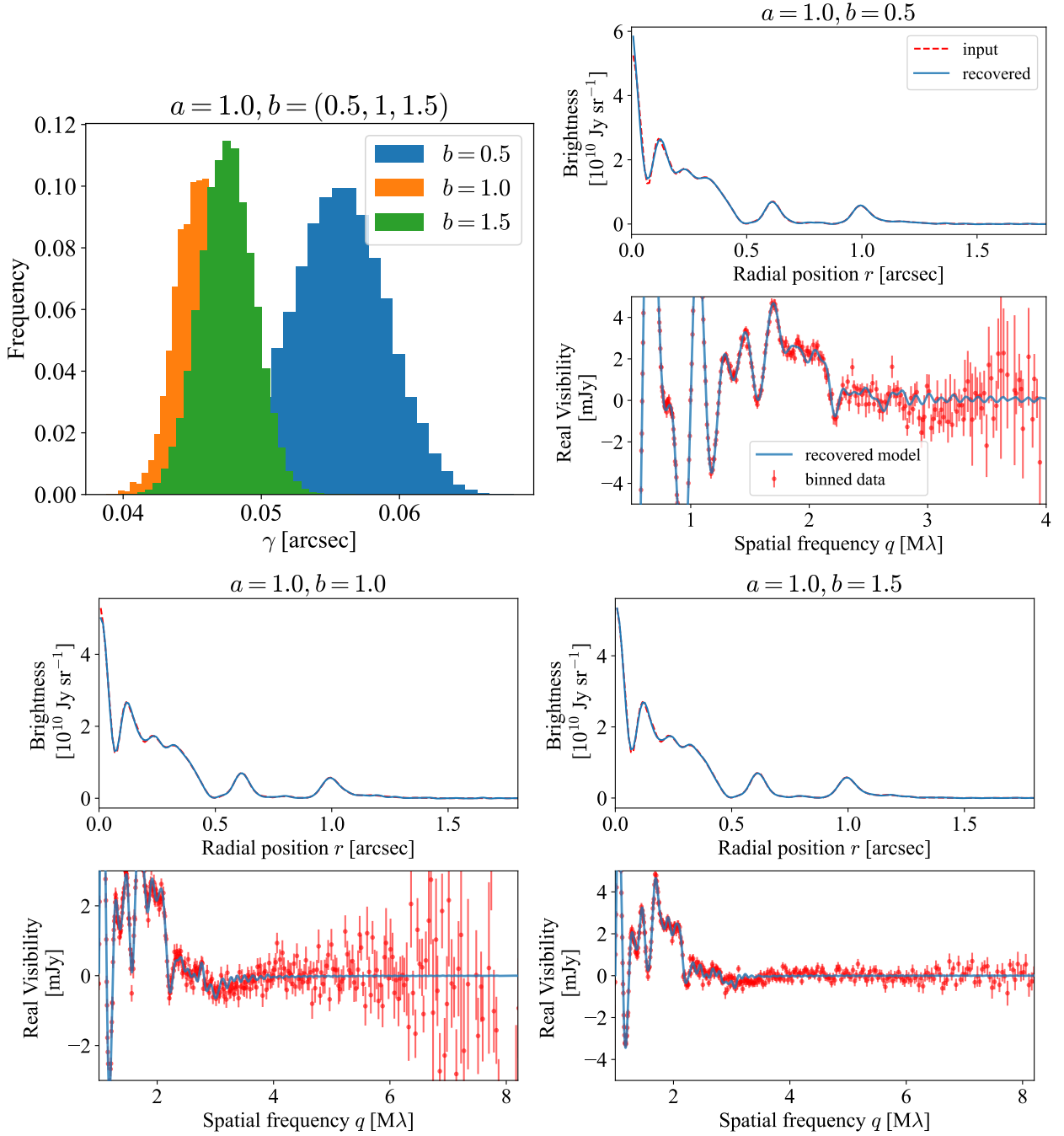


Figure D2. Same figure as Fig. D1 except that we vary $b = (0.5, 1, 1.5)$ and fix $a = 1$.

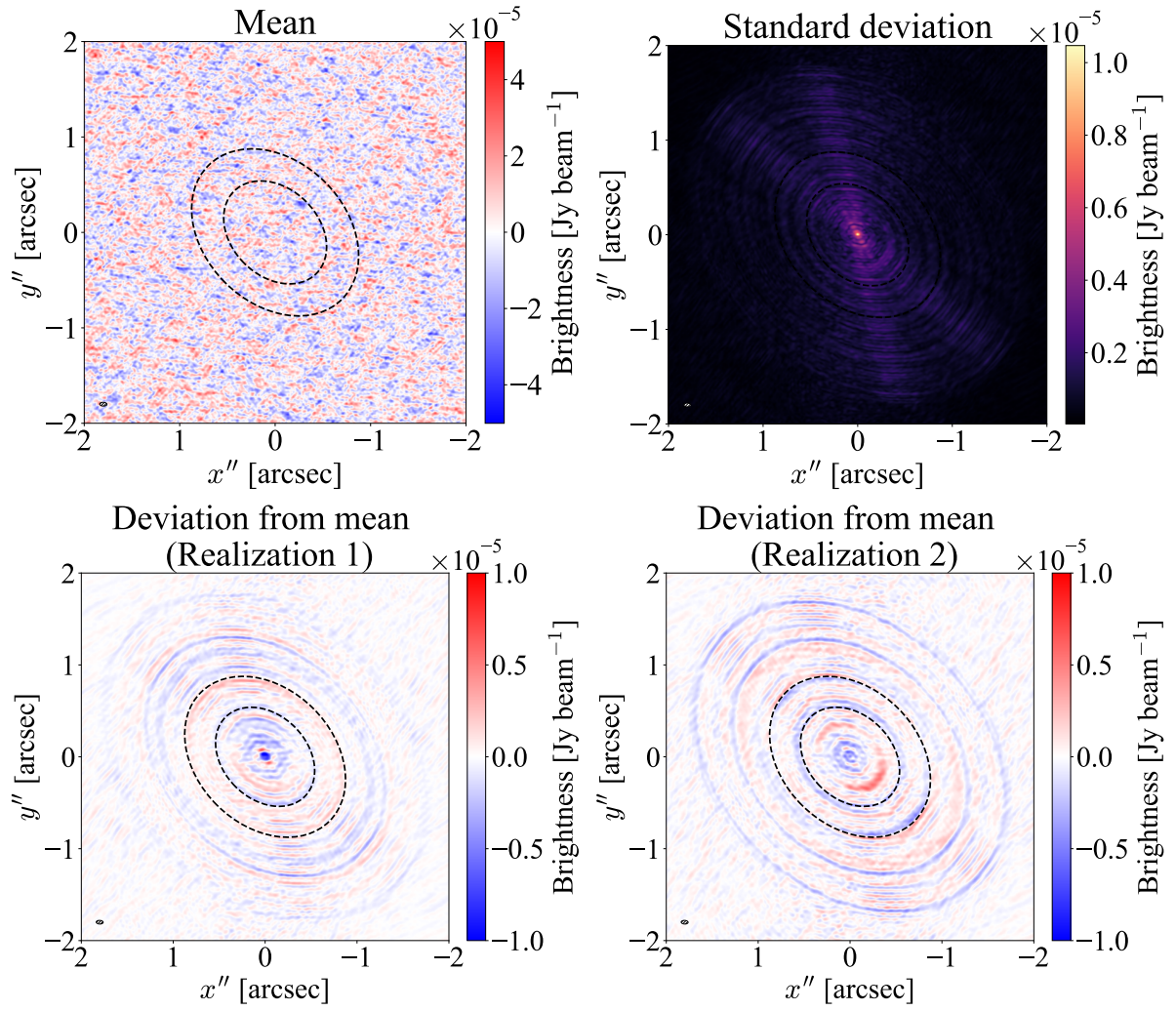


Figure E1. Mean and variance of residual images produced with 10 samples of parameters derived from simulated data of AS 209 in Sec 4. Upper and left panels show the mean and standard deviation of the residual images. The lower panels show the difference maps of two random residual images from the mean residual image.

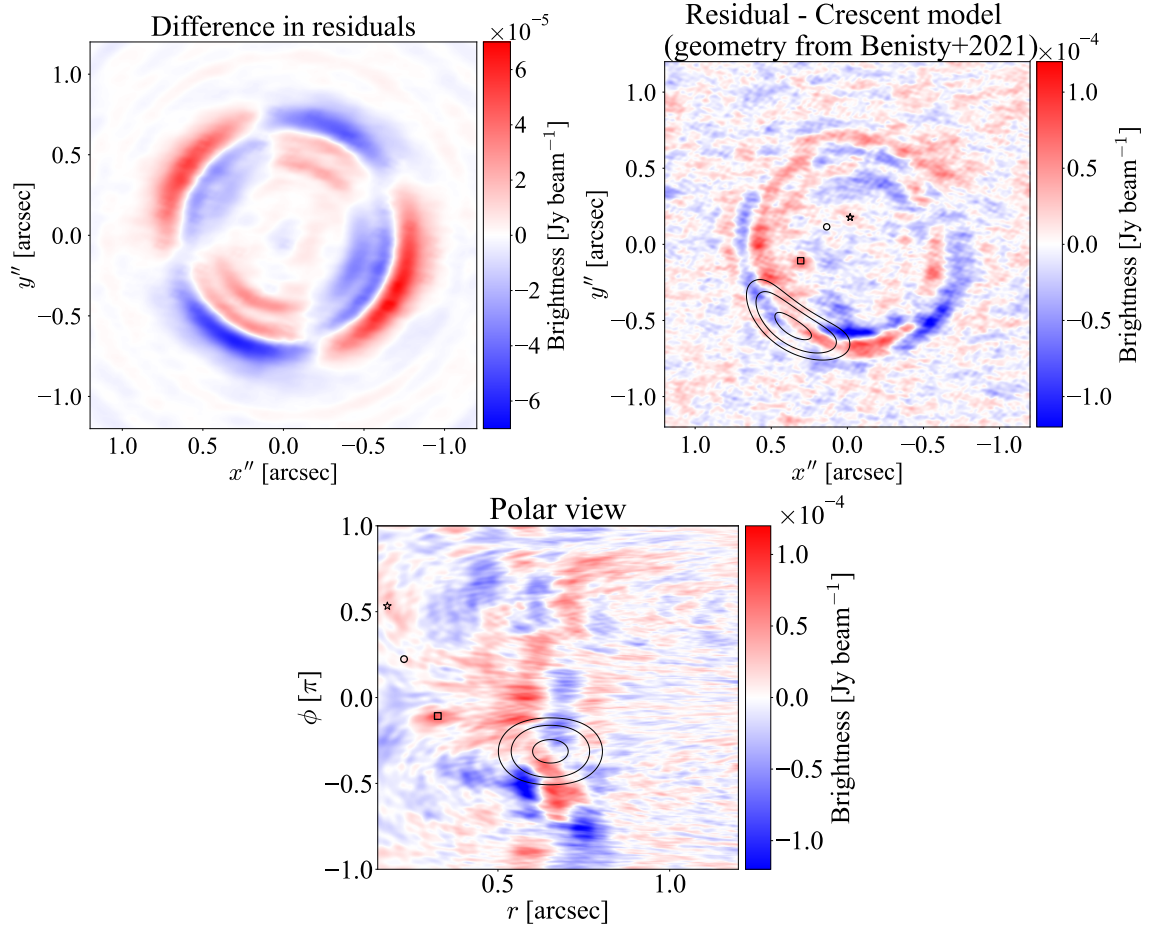


Figure F1. Difference in residual images derived from geometry in our study and Benisty et al. (2021), and residual images based on the geometry from Benisty et al. (2021) with the crescent model being subtracted. (upper left) Difference in the residual images based on geometries from two studies in Fig 15. (upper right) The residual image with the subtraction of the crescent model based on geometry from Benisty et al. (2021). The format is same as that in Fig 16. (lower) The image with the crescent mode being subtracted in polar coordinate.