

Searching for planet-driven dust spirals in ALMA visibilities

Edward T. Stevenson¹,^{*} Álvaro Ribas,¹ Jessica Speedie², Richard A. Booth³ and Cathie J. Clarke¹

¹*Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK*

²*Department of Physics and Astronomy, University of Victoria, Victoria, BC V8P 1A1, Canada*

³*School of Physics and Astronomy, University of Leeds, Leeds LS2 9JT, UK*

Accepted 2024 April 18. Received 2024 March 25; in original form 2023 December 8

ABSTRACT

Atacama Large Millimetre/submillimetre Array (ALMA) observations of the thermal emission from protoplanetary disc dust have revealed a wealth of substructures that could evidence embedded planets, but planet-driven spirals, one of the more compelling lines of evidence, remain relatively rare. Existing works have focused on detecting these spirals using methods that operate in image space. Here, we explore the planet detection capabilities of fitting planet-driven spirals to disc observations directly in visibility space. We test our method on synthetic ALMA observations of planet-containing model discs for a range of disc/observational parameters, finding it significantly outperforms image residuals in identifying spirals in these observations and is able to identify spirals in regions of the parameter space in which no gaps are detected. These tests suggest that a visibility-space fitting approach warrants further investigation and may be able to find planet-driven spirals in observations that have not yet been found with existing approaches. We also test our method on six discs in the Taurus molecular cloud observed with ALMA at 1.33 mm, but find no evidence for planet-driven spirals. We find that the minimum planet masses necessary to drive detectable spirals range from ≈ 0.03 to $0.5 M_{\text{Jup}}$ over orbital radii of 10–100 au, with planet masses below these thresholds potentially hiding in such disc observations. Conversely, we suggest that planets $\gtrsim 0.5$ – $1 M_{\text{Jup}}$ can likely be ruled out over orbital radii of ≈ 20 – 60 au on the grounds that we would have detected them if they were present.

Key words: accretion, accretion discs – methods: data analysis – methods: statistical – techniques: interferometric – planet–disc interactions – protoplanetary discs – submillimetre: planetary systems.

1 INTRODUCTION

During its formation, a planet gravitationally interacts with its natal protoplanetary disc and generates substructure that is potentially observable. For instance, every planet drives a wake in its disc, which is sheared into spiral arms by the disc’s differential rotation (Kley & Nelson 2012). For larger planets, these spiral arms steepen into shocks as they propagate and can eventually lead to the opening of annular gaps in the disc (Lin & Papaloizou 1986; Goodman & Rafikov 2001). To what extent these substructures can be identified in observations has been explored across various regimes, including near-IR scattered light images (e.g. Dong et al. 2015; Juhász et al. 2015; Dong & Fung 2017), (sub-)mm dust thermal emission (Weber et al. 2019; Speedie, Booth & Dong 2022; Binkert, Szulágyi & Birnstiel 2023), and CO emission lines (Teague et al. 2018a, b; Bae, Teague & Zhu 2021). Through detection of these substructures, there is the possibility of inferring the presence of young planets and constraining their demographics, which, in turn, will serve to test planet formation theories.

Exoplanet surveys have made it clear that planets are abundant (Zhu & Dong 2021), and spirals are a robust prediction of planet–disc interactions (Goldreich & Tremaine 1980). Therefore, one would

expect a large number of planet-driven spirals in protoplanetary discs. This theoretical expectation has not, however, translated into a large number of clear detections. The majority of spiral detections to date have come from scattered light observations of sub- μm grains (e.g. Fukagawa et al. 2004; Muto et al. 2012). Spiral features have also more recently been observed in CO emission lines that trace gas kinematics (e.g. Christiaens et al. 2014; Tang et al. 2017; Teague et al. 2019). Continuum dust emission, which has revealed plenty of annular structure, i.e. gaps and rings (e.g. ALMA Partnership et al. 2015; Andrews et al. 2016, 2018; Clarke et al. 2018; Kudo et al. 2018; Long et al. 2018; Huang et al. 2018a; Keppler et al. 2019; Huang et al. 2020; Macías et al. 2021), has additionally delivered a small number of spiral detections. Among these, the discs around single stars that display spiral patterns in dust emission include Elias 2–27, IM Lupi, WaOph 6, and HD 143006, which have been found to contain large-scale dust spirals (Pérez et al. 2016; Huang et al. 2018b; Andrews et al. 2021), and a smaller scale dust spiral found in MWC 758 (Dong et al. 2018a). These spirals have not been definitively attributed to planets, with Mawet et al. (2012) finding no evidence for an embedded companion in their direct imaging analysis of IM Lupi. In the case of Elias 2–27, Meru et al. (2017) suggested that the spirals could be consistent with gravitational instability or a 10–13 M_{Jup} companion located at a large orbital radius external to the spirals (between ≈ 300 – 700 au), but considered the spirals unlikely to be due to a planet internal to the spirals. Similarly, Brown-Sevilla

* E-mail: edward.stevenson100@gmail.com

et al. (2021) found the spirals in WaOph 6 to require a planet of minimum mass $10 M_{\text{Jup}}$ exterior to the spirals (at orbital radii > 100 au), and Dong, Najita & Brittain (2018b) found the spirals in MWC 758 could be consistent with a planet of a few M_{Jup} exterior to the spirals at ≈ 100 au. The faint spiral in HD 143 006 reported by Andrews et al. (2021) lends support for the 10–20 M_{Jup} embedded companion already suggested by other forms of evidence such as CO kinematics (Pérez et al. 2018) and asymmetries in scattered light (Benisty et al. 2018). Therefore, the few spirals found could only have a planetary origin for extremes of mass or orbital radius, and no spirals attributable to more moderate planet masses have yet been found in the mm continuum.

The extent to which this scarcity of planet-driven dust spirals is genuine, or a result of observational limitations, is currently poorly understood. In addition, it is unclear how reliably dust spirals could be detected across the disc parameter space and interferometric set-ups. Speedie et al. (2022, henceforth S22) shed light on these uncertainties by using image residuals and contouring to analyse disc asymmetries, applied to synthetic Atacama Large Millimetre/submillimetre Array (ALMA) observations of planet-containing model discs.

In this work, we build on S22 by exploring a different approach to analysing disc asymmetries that specifically targets planet-driven spirals, and aims to test the limits of planet recovery with spirals in ALMA observations. Existing works with efforts to highlight or detect spiral structures in ALMA continuum images have used a variety of approaches: unsharp masking (e.g. Pérez et al. 2016; Meru et al. 2017), high-pass filtering (e.g. Rosotti et al. 2020; Norfolk et al. 2022), image contouring (e.g. Jennings et al. 2022), and axisymmetric brightness subtraction (e.g. Andrews et al. 2021; Jennings et al. 2022; S22; Speedie & Dong 2022). While some of these approaches perform the suppression of axisymmetric emission in image space and others in visibility space, all of them lead to the process of identifying spiral structure being done in image space. These image-space methods therefore require transformation from the observed visibilities into an image, generally performed with the CLEAN deconvolution algorithm (Högbom 1974; Clark 1980; Cornwell 2008). Although CLEAN is the standard and clearly effective technique used across radio astronomy, it suffers the well-known drawbacks of information loss, primarily from the convolution of the CLEAN model with the CLEAN beam, and correlated noise with ill-defined uncertainties.

Fitting models via comparison between synthesized and observed visibilities is therefore preferable to model fitting in the image plane, since observational data are acquired in the visibility plane and its uncertainties are well understood. Here, we develop a method that implements this approach and serves as a proof of concept, providing insights into its potential uses and challenges.

We test our new method on (an extended version of) the set of synthetic observations of planet-containing discs from S22, and compare its planet recovery ability to looking for spirals in CLEAN image residuals (the principal spiral detection method tested in S22), and looking for gaps in CLEAN image profiles. We also test the method on a sample of six discs (Table 2) from the ALMA survey of the 1.33 mm dust thermal emission of 32 discs in the Taurus star-forming region (~ 0.12 arcsec resolution, $50 \mu\text{Jy beam}^{-1}$ mean sensitivity) presented in Long et al. (2019). These discs were chosen for their ‘smoothness’ (relative lack of annular substructure) and for easy comparison with the analysis of this sample undertaken in Jennings et al. (2022), who used both image contouring and axisymmetric brightness subtraction (performed in visibility space) to analyse the disc asymmetries.

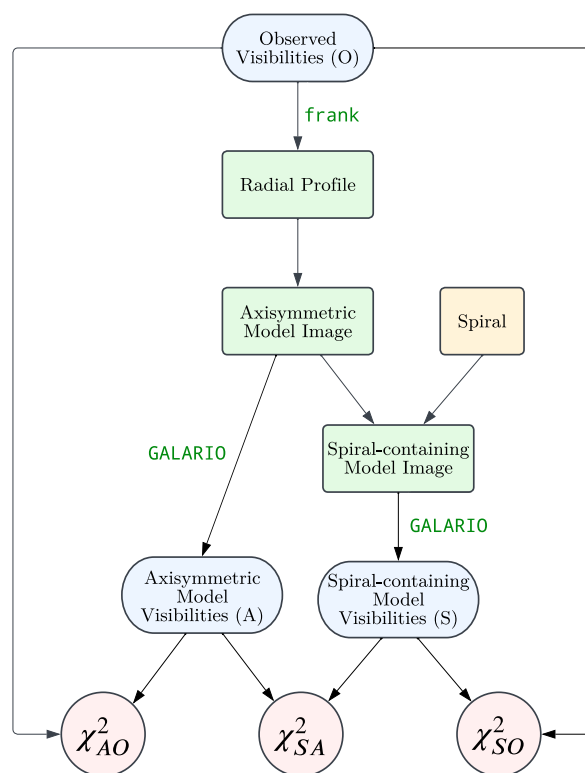


Figure 1. Flow chart showing the key steps of the method, including how the χ^2 metrics are generated. The steps going from the observed visibilities to the axisymmetric model image are described in Section 2.1. How the spiral is combined with the axisymmetric model image to create the spiral-containing model image is described in Section 2.2. The transformation to visibility space and χ^2 metrics are explained in Section 2.3.

Section 2 describes our method. Sections 3 and 4 present and analyse the main results of our tests of the method on the synthetic observations and Taurus disc sample, respectively. Section 5 discusses the significance of these results and suggests avenues for further work. We summarize and conclude in Section 6.

2 METHOD

We require two components to apply our method to a given disc observation (whether real or synthetic): an axisymmetric model image of the disc (Section 2.1) and a spiral perturbation driven by a planet (Section 2.2). Combining these elements, we construct a model disc image that incorporates the planet’s influence. We then transform to visibility space and use χ^2 statistics to compare the axisymmetric and spiral-containing model to the observed data. Repeating this for different planet positions and spiral simulation parameters, we use the χ^2 metrics to determine which parameters give the best fits to the observed visibilities and interpret the results using the Bayesian Information Criterion (BIC) and more qualitative methods (Section 2.3). The steps of the method, including how the χ^2 metrics are generated, are summarized visually in Fig. 1.

2.1 Creating the axisymmetric model image

First, we use FRANK (Jennings et al. 2020) to reconstruct the disc’s 1D radial profile. FRANK fits the observed visibilities directly to construct a non-parametric, axisymmetric model of the disc’s brightness.

We then generate an axisymmetric model image I_A from this radial profile by interpolating it and sweeping over 2π radians, making sure to use a large enough image plane to avoid aliasing, and small enough pixels to Nyquist sample the uv -points.

Next, we partially reintroduce the disc's on-sky projection, inclining the model but forgoing its rotation or offset. The justification for this is that the inclination of the disc affects its total observable flux; here, we take the optically thick limit, in which the total flux is reduced by a factor $\cos i$ for inclination i . By contrast, neither rotation nor offset affects total flux, and both can be implemented in visibility space, rather than image space, with better performance and accuracy (Briggs, Schwab & Sramek 1999). We thus defer their inclusion to the final step of generating the model visibilities.

2.2 Adding the spiral

To introduce the spiral into the model, we employ an extended version of the set of two-dimensional (2D) gas + dust hydrodynamic simulations generated in S22. In brief, the simulations were performed with FARGO3D (Benítez-Llambay & Masset 2016) modified to compute dust dynamics (Booth, Sijacki & Clarke 2015; Rosotti et al. 2016), with a planet placed at $r_p = 50$ au. The 2D cylindrical (r, ϕ) domain spans 2π in azimuth and $0.1 r_p$ to $3.0 r_p$ in radius, with a resolution of $N_r \times N_\phi = 1100 \times 2048$ cells (spaced logarithmically and linearly in the r and ϕ directions, respectively). Wave damping zones (de Val-Borro et al. 2006) were employed to minimize wave reflections. The dust layer was given open/inflow radial boundary conditions. The equation of state of the gas was varied to be locally isothermal or approximately adiabatic; in the latter case, the energy equation was solved using a β -cooling prescription (see below). The simulations were evolved to 1500 orbits, after which the disc was radially truncated to $0.2 r_p$ to $2.2 r_p$ in order to remove the damping zones, and radiative transfer was performed on the outputs analytically. In the approximately adiabatic case, this was done directly with the dust surface density and temperature distributions, assuming $T_{\text{dust}} = T_{\text{gas}}$; in the isothermal case, an axisymmetric temperature distribution was generated. We refer the reader to sections 2.1 and 2.2 of S22 for details.

The result of these steps is a continuum brightness map of a disc containing a planet (the left-hand panel of Fig. 2). To isolate the planet's asymmetric influence, we subtract the map's azimuthally averaged brightness and normalize by it, giving us a fractional residual brightness map (see the right-hand panel of Fig. 2).¹ We then interpolate this map on to the same Cartesian grid as the axisymmetric model image. This allows the spiral perturbation and axisymmetric model to be superimposed, giving us our spiral-containing model disc image. Mathematically,

$$I_S = I_A \left(1 + \frac{I_{\text{sim}} - \langle I_{\text{sim}} \rangle_\phi}{\langle I_{\text{sim}} \rangle_\phi} \right) = I_{\text{sim}} \times \frac{I_A}{\langle I_{\text{sim}} \rangle_\phi} \quad (1)$$

where I_A is the axisymmetric disc image from FRANK, $(I_{\text{sim}} - \langle I_{\text{sim}} \rangle_\phi) / \langle I_{\text{sim}} \rangle_\phi$ is the fractional residual map of a spiral simulation, and I_S is our resulting spiral-containing model disc image. This procedure is equivalent to rescaling the brightness map by our disc's axisymmetric profile, as indicated by the second equality.

¹Throughout this work, we refer to the planet's asymmetric influence simply as a 'spiral', though note it can include additional, non-linear structure, particularly for simulations of planets above a thermal mass.

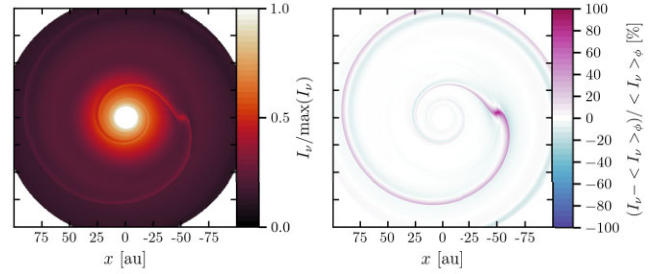


Figure 2. Simulation of a $3 M_{\text{th}}$ planet in an approximately adiabatic ($\beta = 10$), marginally optically thick ($\tau_0 = 1$) disc. Left: Brightness map. Right: Corresponding residual brightness map in percentage units of the azimuthally averaged brightness.

2.2.1 Simulation parameters

(i) **Planet mass.** The planet mass M_p can take the values 0.1, 0.3, 1, or $3 M_{\text{th}}$, where the thermal mass² the local scale height of the gas disc. $M_{\text{th}} = (H/r_p)^3 M_*$ with r_p the planet orbital radius, H the local scale height, and M_* the stellar mass. For these simulations, we fix $r_p = 50$ au, with $(H/r_p) = 0.07$ and $M_* = 0.8 M_\odot$, giving a range of $0.1 M_{\text{th}} \approx 0.5 M_{\text{Nep}}$ to $3 M_{\text{th}} \approx 0.9 M_{\text{Jup}}$ at this radius. The disc aspect ratio H/r affects the spiral morphology (Bae & Zhu 2018), and therefore, we investigate its effects on our method in Section 3.1. The planet mass parameter has the strongest effect on spiral morphology in these simulations.

(ii) **Optical depth.** This is parametrized by the initial (i.e. after zero orbits) optical depth at the planet's orbital radius (50 au) τ_0 . For these simulations, $\tau_0 = 0.1, 0.3, 1, \text{ or } 3$, thus ranging from optically thin through to optically thick. This optical depth is coupled to the observational wavelength, for which we consider two values in this work: 0.87 and 1.33 mm. We scale the brightness of the simulations to ensure the optical depths are correct for both wavelengths.

(iii) **Cooling time.** The gas spiral morphology and amplitude depends on the cooling time-scale of the disc gas, which then affects the dust spiral via dust–gas coupling (Miranda & Rafikov 2020; Zhang & Zhu 2020). We parametrize the cooling time-scale by β , where $t_{\text{cool}} = \beta/\Omega$ and Ω is the local angular speed. We use two β values to explore the extremes: $\beta = 0$ for isothermal, and $\beta = 10$ for approximately adiabatic. As a shorthand, we often refer to approximately adiabatic simply as adiabatic (this terminology is consistent with S22).

(iv) **Spiral handedness.** Because it is not generally possible to tell in which direction the disc is rotating – which determines the winding sense of the spiral – from our observations, it is necessary to introduce a handedness parameter for the spiral sense. We define the spiral handedness such that the spirals shown in Figs 2 and 3 are right-handed.

In total, this gives $4 \times 4 \times 2 \times 2 = 64$ possible spiral simulations. We then need two parameters to specify the planet's location in the disc: its orbital radius r_p and its azimuthal angle ϕ_p , which is defined anticlockwise from the right arm of the disc's minor axis following deprojection (assuming conventional sky coordinates). Hence, for any given spiral, we have six parameters: four simulation parameters (M_p , τ_0 , β , and handedness) and two positional parameters (r_p and ϕ_p). In the rest of this work, we refer to a given set of these six parameters as a spiral model instance. We use the term 'spiral

²Physically, the thermal mass corresponds to the mass at which the planet's Hill radius is equal to $3^{-1/3}$ times.

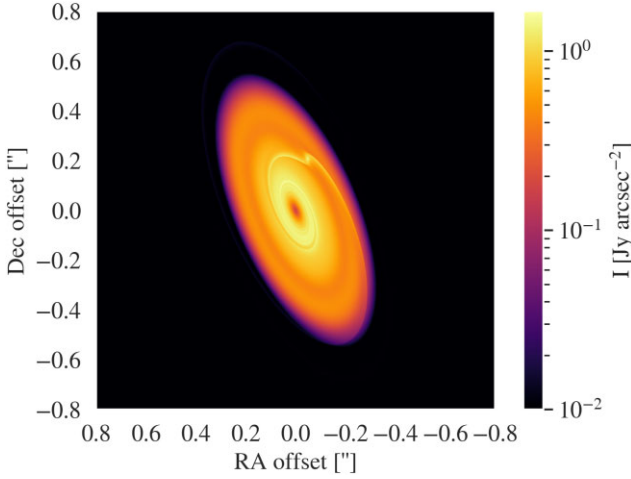


Figure 3. Projected disc model of RY Tau containing a $1 M_{\text{th}}$ planet at $r_p = 0.34$ arcsec, $\phi_p = 30^\circ$. Inclination = 65.0° , position angle = 23.1° .

templates’ to refer to spiral model instances in the context of their use fitting disc observations in our method (via their residual brightness maps).

Note that although we consider different r_p for our spiral templates, we only use spiral *simulations* taken at a fixed r_p ($= 50$ au). We assume the simulations are scale-free, meaning that the brightness perturbation due to the spiral depends only on the ratio r/r_p . This allows us to scale r_p for a template without running a new spiral simulation. However, because the magnitude of the spiral brightness perturbation scales approximately with the ratio M_p/M_{th} (Zhu et al. 2015), and we keep the brightness perturbation amplitude constant with r_p in our templates, we must also keep M_p/M_{th} constant with r_p . Therefore, since the simulations assume a disc flaring of $H/r \propto r^{1/4}$, implying $M_{\text{th}} \propto r_p^{3/4}$, this means $M_p \propto r_p^{3/4}$. That is, the absolute planet mass of a given template changes with r_p .

As a checkpoint, by the end of this step we have generated our model image, consisting of a spiral template superimposed with an axisymmetric model of the disc (the disc is inclined but not yet rotated or offset to match the real disc observation, which is performed in visibility space). Fig. 3 shows a representation of one of these model discs (displayed at a higher resolution and with an on-sky geometry matching the real disc observation).

2.3 Detecting a spiral

The visibilities of a given model disc image are generated via the 2D Fourier transform, incorporating the necessary rotation and offset to match the real disc observation. We achieve this using GALARIO (Tazzari, Beaujean & Testi 2018). For clarity, we introduce a shorthand for the three visibilities we are working with: O for the Observed data, A for those of the Axisymmetric model, and S for those of the Spiral model instance.

We use χ^2 metrics to assess the goodness of fit, where the χ^2 taken between a pair of visibilities P and Q is defined as

$$\chi_{PQ}^2 = \sum_k w_k |V_{P,k} - V_{Q,k}|^2, \quad (2)$$

where $w_k = 1/\sigma_k^2$ is the weight of the k th visibility point and σ_k^2 is the variance of its real and imaginary components. Since we can take a χ^2 between any pair of visibilities, we have three relevant

χ^2 metrics: χ_{AO}^2 , χ_{SA}^2 , and χ_{SO}^2 .³ The weights used for all of these metrics are those of the observed visibilities.⁴

For a spiral model instance that is a better fit to the observed visibilities than the axisymmetric model, we expect $\chi_{SO}^2 < \chi_{AO}^2$. For brevity, we introduce the quantity $\Delta\chi^2 = \chi_{AO}^2 - \chi_{SO}^2$, meaning that a positive (negative) $\Delta\chi^2$ corresponds to a model instance that is a better (worse) fit than the axisymmetric model.

To confidently claim a spiral detection in an observation using our method, there are two tests the data must pass:

(i) The spiral model must be a better model than the axisymmetric model for the data. We evaluate this using the Bayesian Information Criterion (BIC; Schwarz 1978) defined as

$$\text{BIC}(M) = 2 \ln L_M - |M| \ln N,$$

where L_M is the maximized value of the likelihood of model M , $|M|$ is the number of parameters estimated by the model, and N is the number of data points.⁵ The change in BIC between the spiral and axisymmetric model is given by

$$\Delta\text{BIC} = \text{BIC}(S) - \text{BIC}(A) \approx \Delta\chi_{\text{max}}^2 - 6 \ln N, \quad (3)$$

which follows because the noise in visibility data is approximately Gaussian and we estimate an additional six parameters from the data for the spiral model. $\Delta\text{BIC} > 0$ provides evidence for the spiral model, with $\Delta\text{BIC} > 10$ typically being taken as very strong evidence (Kass & Raftery 1995). For our data (Sections 3 and 4), $\Delta\chi^2 \gtrsim 70$ would generally constitute strong evidence by this criterion.⁶

(ii) $\Delta\chi^2$ must have a spatial (r_p , ϕ_p) structure that indicates unambiguously that the improvement is caused by a spiral, not some other asymmetry. We cannot discern this with statistics alone; we must also rely on more qualitative methods. For this purpose, we introduce $\Delta\chi^2$ heatmaps (e.g. Figs 4–7) which show the $\Delta\chi^2$ for different planet positions in the deprojected disc. This test is needed because the spiral model is capable of fitting certain non-spiral asymmetries to the extent that it is favoured over the axisymmetric model by the BIC. This ‘false fitting’ can therefore pass our first test and, if unchecked, would lead to incorrectly inferring a spiral. In Section 3, we will see examples of this false fitting and how to discern it from true fitting.

3 RESULTS: SYNTHETIC OBSERVATIONS

Here, we generate synthetic observations of model discs and then treat these as we would a real disc observation in which we are seeking evidence of spiral structure (i.e. we evaluate the difference in fits to the observed visibilities between an axisymmetric model and models including various spiral structures). We investigate five different observational set-ups, which are described in Table 1.

³ χ_{SA}^2 is not used directly in the method; however, it is useful for exploring the magnitude of effect on the visibilities of adding a spiral to an axisymmetric model disc (Appendix B).

⁴ For the synthetic observations, we use the weights computed from the median variance of the deprojected and binned visibilities via the FRANK package (Jennings et al. 2020), as the weights returned by CASA (McMullin et al. 2007) are not statistical.

⁵ The BIC is well-suited to large N (as is the case for visibility data), whereas p -value-based model selection is less reliable for such data sets (Raftery 1995).

⁶ ΔBIC approximates the Bayes factor between the models B_{SA} via $\Delta\text{BIC} \approx 2 \ln B_{SA}$.

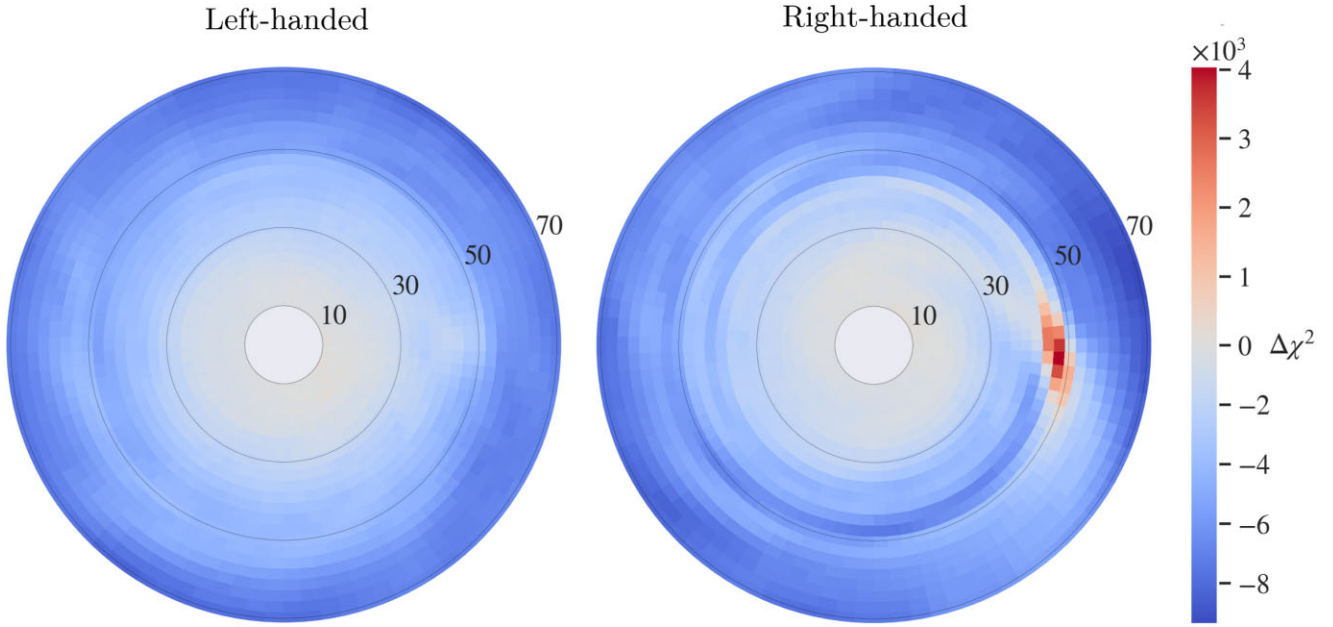


Figure 4. $\Delta\chi^2$ heatmaps of a synthetic observation from our fiducial set of a model disc ($\tau_0 = 3$, $\beta = 10$, $M_p = 1 M_{\text{th}}$, right-handed spiral) for its matching right-handed spiral template (right) and a left-handed, but otherwise matching, spiral template (left). The disc's planet is located at $r_p = 50$ au, $\phi_p = 0$ (along the right horizontal axis). The heatmap spatially represents the disc, with each point on the heatmap corresponding to a spiral template with its planet placed at that location. Recall $\Delta\chi^2 = \chi_{AO}^2 - \chi_{SO}^2$, i.e. the difference in goodness of fit between the axisymmetric model and spiral model instance. Positive $\Delta\chi^2$ values (red) mean the spiral model instance is a better fit, and negative values (blue) mean the axisymmetric model is a better fit. The maximum $\Delta\chi^2$ value for the right-handed template is 4.1×10^3 , occurring in the region around the planet position. The maximum $\Delta\chi^2$ value for the left-handed template in the same region is -430 , indicating a strong preference for the right-handed model. The heatmaps are truncated at the disc's inner boundary of 10 au, and an outer boundary of 70 au (the 70–110 au region of these discs offers little interest). The radial ticks are in au and the colourbar applies to both heatmaps to aid comparison. Note that the colourbar diverges asymmetrically about zero.

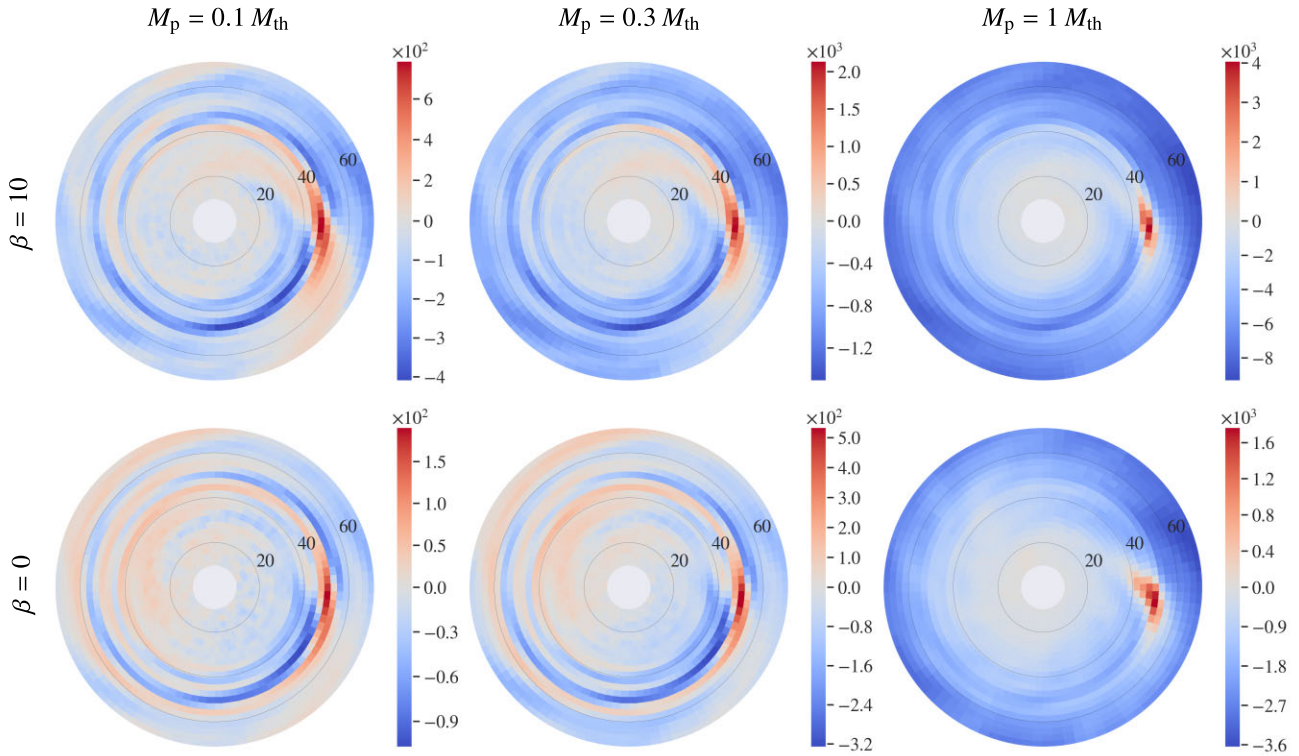


Figure 5. $\Delta\chi^2$ heatmaps of a synthetic observation from our fiducial set of a model disc ($\tau_0 = 3$, $\beta = 10$, $M_p = 1 M_{\text{th}}$, right-handed spiral) for $\tau_0 = 3$ spiral templates. $M_p/M_{\text{th}} = 0.1, 0.3, 1$ from left to right, and $\beta = 0, 10$ from bottom to top. The heatmap for the matching spiral template is therefore in the top right. These heatmaps demonstrate that the method can recover the planet with non-matching spirals.

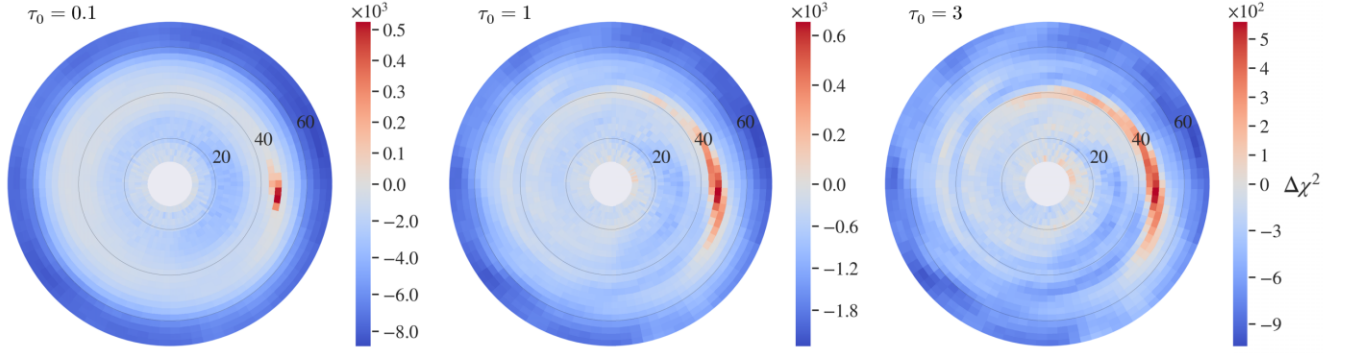


Figure 6. $\Delta\chi^2$ heatmaps of a synthetic observation from our fiducial set of a model disc ($\tau_0 = 0.3$, $\beta = 10$, $M_p = 1 M_{\text{th}}$, right-handed spiral) for wrong optical depth, but otherwise matching, spiral templates. $\tau_0 = 0.1, 1, 3$ from left to right (the heatmap for the matching $\tau_0 = 0.3$ template appears almost identical to the $\tau_0 = 1$ one). These heatmaps demonstrate that the method can recover the planet without having to accurately predict the disc’s optical thickness.

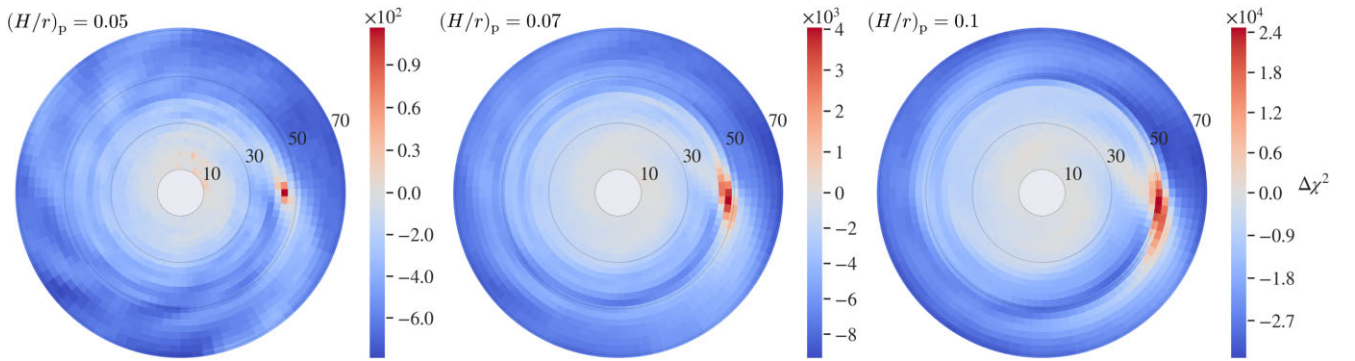


Figure 7. $\Delta\chi^2$ heatmaps of model discs with $(H/r)_p = 0.05, 0.07, 0.1$ (left, centre, right), and otherwise the same parameters: $\tau_0 = 3$, $\beta = 10$, $M_p = 1 M_{\text{th}}$, and a right-handed spiral. The discs are synthetically observed with our fiducial set-up. The $\Delta\chi^2$ heatmaps are generated with the $(H/r)_p = 0.07$, and otherwise matching, spiral template in each case. This demonstrates that the method can recover the planet without having to accurately predict the disc aspect ratio.

Table 1. ALMA set-ups and observational parameters for the synthetic observations in this work. The sensitivities given refer to *requested* sensitivities with the ALMA Cycle 8 Observing Tool. The combined (i.e. summed compact and extended configuration) on-source times given are those needed to achieve these requested sensitivities. After CLEANING, the measured rms noise in all our images is $\sim 80\text{--}95$ percent of the requested sensitivity. The penultimate column gives the $\Delta\chi^2$ value for which $\Delta\text{BIC} = 0$, meaning the spiral and axisymmetric models are equally likely. From equation (3), this relates to the number of data points N by $\Delta\chi^2(\Delta\text{BIC} = 0) = 6\ln N$. Hence, a $\Delta\chi^2 > \Delta\chi^2(\Delta\text{BIC} = 0) + 10$ indicates a very strong preference for the spiral model. The visibilities and images for all of our sets are available at doi.org/10.6084/m9.figshare.25464109. The final column indicates if the set was also analysed in S22. For comparison, the bottom two rows give the corresponding information for the Taurus disc sample analysed in Section 4 and the DSHARP survey.

Descriptor	Angular resolution (arcsec)	Sensitivity ($\mu\text{Jy beam}^{-1}$)	Combined on-source time (min)	ALMA configuration	Observational wavelength (mm)	$\Delta\chi^2$ ($\Delta\text{BIC} = 0$)	In S22
Synthetic observations							
Fiducial set	0.061	35	40	C-4 + C-7	0.87	63	Yes
Higher sensitivity set	0.061	10	480	C-4 + C-7	0.87	78	Yes
Lower sensitivity set	0.061	50	19	C-4 + C-7	0.87	59	No
Higher resolution set	0.028	35	40	C-5 + C-8	0.87	63	Yes
Taurus comparison set	0.13	50	6	C-6	1.33	52	No
Real observations							
Taurus survey (Long et al. 2019)	~ 0.12	~ 50	~ 6	C-6	1.33	~ 60	–
DSHARP survey (Andrews et al. 2018)	~ 0.035	~ 20	~ 90	C-5 + C-8/9	1.25	–	–

The synthetically observed discs are based on the same set of spiral simulations as the spiral templates which we use in our method. Therefore, these observations clearly present an idealized situation for analysis with our method. In light of this, we first explore whether an improvement in fit is only manifest for the spiral template with exactly the same parameters as those used in the model disc we observe (which we call the ‘matching’ spiral template) or whether the spiral can be recovered, and the location of the planet constrained, for a range of spiral templates (Section 3.1). We then compare the method’s ability to identify planets to image residuals (Section 3.2), and a simple gap-based method (Section 3.3). The $\Delta\chi^2$ values and heatmaps from these synthetic observations will also serve as comparison points for the analysis of the Taurus disc sample in Section 4. The effects of errors in the fitted disc geometry are considered in Appendix A.

The synthetically observed discs we examine in this section are all face-on, with zero rotation or offset. Their inner and outer boundaries are at 10 and 110 au, with the planet positioned at $r_p = 50$ au, $\phi_p = 0^\circ$, and producing a right-handed spiral (orbiting anticlockwise). They are computed with a distance to the disc of 140 pc. The host stars are given a mass of $0.8 M_\odot$ and luminosity of $1.5 L_\odot$, which are the approximate medians of the sample in the DSHARP survey (table 1 of Andrews et al. 2018). The chosen planet location of 50 au aligns roughly with the orbital radius where many of the DSHARP survey’s gaps and rings are observed (fig. 7 of Huang et al. 2018a).

To generate the radial brightness profiles for the synthetic discs we use the following FRANK hyperparameters: $N = 200$, $R_{\text{out}} = 1.2$ arcsec, $\alpha = 1.05$, $p_0 = 10^{-15} \text{ Jy}^2$, and $w_{\text{smooth}} = 0.1$. These are within the ranges suggested by Jennings et al. (2020) and our resulting profiles are quite insensitive to variations within those ranges.

3.1 Effect of spiral model parameters

To test the method, we first examine the goodness of fit when a synthetic observation of a model disc is fitted with spiral templates of different parameters to that disc. For these tests, we consider discs in our fiducial set described in Table 1. For spiral detection in this set, our second test tends to be more exacting than our first (Section 2.3), meaning our main obstacle is discerning true spiral signals from false ones. In this section, the maximum $\Delta\chi^2$ in all of the heatmaps shown well-exceed the threshold for $\Delta\text{BIC} > 10$.

One would expect the greatest improvement in fit (i.e. the largest $\Delta\chi^2 > 0$) to occur for the matching spiral template, and indeed this is the case. A useful result is that the right-handed spiral (the correct one) shows a much larger maximum $\Delta\chi^2$ than the left-handed spiral (Fig. 4). We can use this as a preliminary test for whether a spiral template is actually fitting a spiral or some other asymmetry that is approximately azimuthally symmetric, such as an arc; if it is fitting such an asymmetry, we would expect the fit improvement to be similar for a right- or left-handed spiral.

When the wrong planet mass is used in the spiral template, we see that the lower planet mass models still show improved fits compared with the axisymmetric model when the planets are placed at a similar location to the disc’s true planet location, although the improvement in χ^2 is inferior to that achieved with the correct planet mass (Fig. 5). In contrast, templates with too large a planet mass ($3 M_{\text{th}}$) do not show any improved fits. Using the wrong cooling time-scale decreases the improvement, but still recovers the planet, even for the smaller masses (Fig. 5). Similarly, a poorly estimated optical thickness does not prevent the planet from being recovered in most cases (Fig. 6; a strongly underestimated optical thickness

can sometimes limit recovery). Decreasing the spiral template’s optical thickness increases its fractional brightness amplitude with a relatively modest impact on the spiral morphology, as indicated by a corresponding increase in $\Delta\chi^2$ amplitudes in its associated heatmap with little change in its structure. This means that for templates with amplitudes that are too low (e.g. due to a planet mass that is too low), decreasing their optical thickness can sometimes improve recovery, and vice-versa for templates with amplitudes that are too high. This also indicates a degree of degeneracy between the optical thickness and planet mass (this may be resolvable through alternative avenues to estimate the optical thickness, e.g. from the disc’s brightness temperature). A general theme is that when a template’s amplitude is much too high, such as when the planet mass is too large, the $\Delta\chi^2$ heatmaps lose their ability to recover the planet. Whereas, when a template’s amplitude is too low, there is often still improvement near the planet’s location and the planet can be recovered.

The disc aspect ratio H/r affects the spiral morphology, with spiral arms tending to be less tightly wound for higher H/r (Rafikov 2002; Bae & Zhu 2018). In our method, we take a fixed $(H/r)_p = 0.07$ for all of our spiral templates. Therefore, to investigate the effect of disc aspect ratio on our method’s ability to recover spirals, we generate two discs with alternative ratios, $(H/r)_p = 0.05$ and 0.1 , and compare the fits derived using our $(H/r)_p = 0.07$ spiral templates to the fit for the matching $(H/r)_p = 0.07$ disc. As can be seen in Fig. 7, the $\Delta\chi^2$ heatmaps recover the planet in each case. In fact, the $(H/r)_p = 0.1$ disc displays a larger peak $\Delta\chi^2$ than the matching $(H/r)_p = 0.07$ disc. This is because, by increasing the aspect ratio, we increase the disc temperature (these parameters are coupled in the simulations), and therefore the disc brightness. This effectively increases the signal-to-noise ratio of the observation. The same effect compounds the decrease in peak $\Delta\chi^2$ for the $(H/r)_p = 0.05$ disc. Therefore, the method is clearly sensitive to H/r , but does not rely on using an accurate value. Each disc fit also retains a strong handedness preference (like in Fig. 4), confirming that the method is not merely sensitive to emission near the planet position. This result justifies our decision to model the spirals as being scale-free. That is, given the method’s relative insensitivity to $(H/r)_p$ and that $H(r)/r$ only varies weakly with r , we adopt a single value for $(H/r)_p$ ($= 0.07$) and then apply an overall scaling with r_p .

In complement to this section, in Appendix B, we explore the effect on the χ^2 amplitude of spiral templates with different parameters. Instead of using the discs considered in this section (which are face-on and would have a simple power-law surface density and temperature profile in the absence of their planets), we consider discs at a range of inclinations and with more complex axisymmetric brightness profiles, informed by our Taurus disc sample analysed in Section 4. The outcomes are consistent with the findings of this section.

3.2 Visibility-space fitting versus image residuals

Here, we compare the spiral detection ability of CLEAN image residuals, generated by subtraction of the azimuthally averaged disc brightness, to our visibility-space fitting method. We examine our fiducial and higher resolution sets, both of which have a requested sensitivity of $35 \mu\text{Jy beam}^{-1}$ (the lowest considered in S22, meaning the easiest observations to make and the hardest spirals to recover) and resolutions of 0.061 and 0.028 arcsec, respectively. For comparison, the Taurus sample discs analysed in Section 4 were observed at an angular resolution ~ 0.12 arcsec and mean sensitivity $\sim 50 \mu\text{Jy beam}^{-1}$ (Long et al. 2019). Therefore, the synthetic observations with the lowest sensitivity and resolution considered in this

section are still more sensitive and higher resolution than the Taurus sample analysed in Section 4, but exhibit lower sensitivities and resolutions than many ALMA observations (e.g. DSHARP survey, Andrews et al. 2018; and several case studies of individual systems, such as: TW Hya, Andrews et al. 2016; HD 169142, Pérez et al. 2019; HL Tau, ALMA Partnership et al. 2015; and MP Mus, Ribas et al. 2023).

A series of comparisons between image residuals and $\Delta\chi^2$ heatmaps for the two resolutions are shown in Fig. 8. In the image residuals, we see lower spiral signal-to-noise ratio for the higher resolution observations. This result occurs because the spiral signal is larger for a larger beam (due to the beam’s larger area). This means that, if noise (in Jybeam^{-1}) is independent of beam size, a lower angular resolution (larger beam) observation will have a higher spiral signal-to-noise ratio, provided it can resolve the distance between spirals.

This is reflected in the $\Delta\chi^2$ values, for which the higher resolution observations give somewhat lower amplitudes, indicating that the spiral perturbations generally exhibit larger amplitudes at shorter baselines (larger spatial scales). However, the higher angular resolution observations do result in a tighter fit around the planet position and reduced ‘false fitting’ ($\Delta\text{BIC} > 0$ in disc regions far from the planet’s true position).

Both images and visibilities, therefore, exhibit a similar trade-off between spiral ‘strength’ (signal-to-noise/spiral model support) and spiral ‘clarity’ (how unambiguously spiral-like the asymmetry is), with higher resolution (at fixed sensitivity) yielding a weaker but clearer spiral signal.

Of the synthetic observations considered in Fig. 8, only the 0.061 arcsec, $1 M_{\text{th}}$ one ((a), bottom row) shows a noticeable spiral in the image residuals, and even here it is an incomplete spiral. We see a much stronger detection in the $\Delta\chi^2$ heatmaps. For the same observations, they recover spirals that are completely washed out by noise in the residuals, and clearly locate the planet.

3.3 Gaps versus spirals

Here, we compare the detectability of gaps and spirals as signals of the embedded planets. For the gap method, we use the radial brightness profile extracted from the CLEAN image (generated with the `radial_profile` function of GOFISH; Teague 2019). We use CLEAN image profiles instead of FRANK fits for two reasons: first, because this aligns with conventional approaches, and second, out of necessity – the model discs have sharp edges at 10 and 110 au, which cause problems for FRANK fits (Jennings et al. 2020) (the FRANK fits are still closer to the true disc profiles than the CLEAN image profiles in many cases, but the oscillatory artefacts [due to Gibbs phenomenon] can produce gap false positives). In these radial profiles, we use gap depth as a measure of detection strength, which we define as

$$\delta = I_{\text{ring}}/I_{\text{gap}}, \quad (4)$$

where I_{gap} is the brightness minimum in the gap and I_{ring} is the brightness maximum in the ring at the gap edge (as in Zhang et al. 2018).

For spirals, we use our visibility-space fitting method and consider our two tests: (i) the spiral model support, indicated by the maximum $\Delta\chi^2$ achieved, and (ii) how clearly spiral-like the asymmetry is, inferred by visual inspection of the $\Delta\chi^2$ heatmaps. Our qualitative criteria for such inferences, along with heatmaps for our fiducial set, are given in Appendix C.

Based on these considerations, we divide the strength of planet recovery into three levels: no recovery, marginal recovery, and clear recovery. Marginal recovery expresses that there is a detectable signal, but that potentially small decreases in observation quality, or slightly less favourable model parameters, could be enough to render the signal undetectable. Since, for a given set of observation/model parameters, the synthetic observations present best case scenarios for detection – known planet location, disc models matching our spiral templates, and no interference from other disc substructure – the observations with marginal planet recovery represent optimistic detection limits for the respective methods. For gaps, we take marginal recovery to be $1 < \delta < 1.5$, with values below and above this range indicating no recovery and clear recovery, respectively. For spirals, we take it to mean clearly passing one of our tests, but only marginally passing the other, i.e. $0 < \Delta\text{BIC} < 10$ or a suggestive, but not unambiguous, structure in the $\Delta\chi^2$ heatmap.

The recovery levels of the two methods are shown in Fig. 9 for various sets of observations. In terms of spiral simulation parameters, we see that

- (i) As expected, the recovery of both gaps and spirals improves with increasing planet mass.
- (ii) For greater optical depths, spirals become easier to recover, whereas gaps become harder to recover.
- (iii) For the longer cooling time, gaps are harder to recover, whereas spirals are much easier to recover.
- (iv) There are clear regions of the parameter space where gaps are not recovered, but spirals are.

Similar trends for planet mass, optical depth, and cooling time are seen for spiral recovery in image residuals (see e.g. fig. 7 of S22). The main reason for the spiral recovery trend for optical depth is likely that decreasing optical depth leads to a fainter disc and therefore a lower absolute amplitude of the spirals. At some point, further increases in optical depth should lead to a decreased spiral signal, as the spiral perturbation becomes small relative to the background emission (especially for the isothermal case, where there is no temperature perturbation). For image residuals, this limit appears to be reached between $\tau_0 = 1\text{--}3$ in some cases (e.g. $\beta = 0$, $M_p = 3 M_{\text{th}}$); however, for visibility-space fitting, this limit is not reached in any of the cases ($\tau_0 \leq 3$) we consider. The main reason for spiral recovery declining for shorter cooling times is likely that the spiral temperature perturbation decreases (disappearing for our isothermal case).

The opposite gap and spiral recovery trends for optical thickness and cooling time may also partly indicate that deeper/wider gaps interfere with the spiral signal. This is certainly true for images, for which deeper/wider gaps reduce the brightness/amount of disc area over which the spiral can be traced (see fig. 10 of S22). This effect appears replicated in the visibilities, although it is somewhat diminished, meaning our visibility-space approach may be less vulnerable to interference from gaps.

While we have only sparsely explored the parameter space here, and some additional caveats should be kept in mind (Section 5), our inferences suggest that looking for spirals to evidence embedded planets would be most useful in optically thicker discs with longer cooling times, and that detectable spirals may be present in regions of the parameter space with no detectable gaps in the image.⁷

⁷Some of these gaps may be detectable with super-resolution fits to the radial profile such as with GALARIO (Tazzari et al. 2018) and FRANK (Jennings et al. 2020).

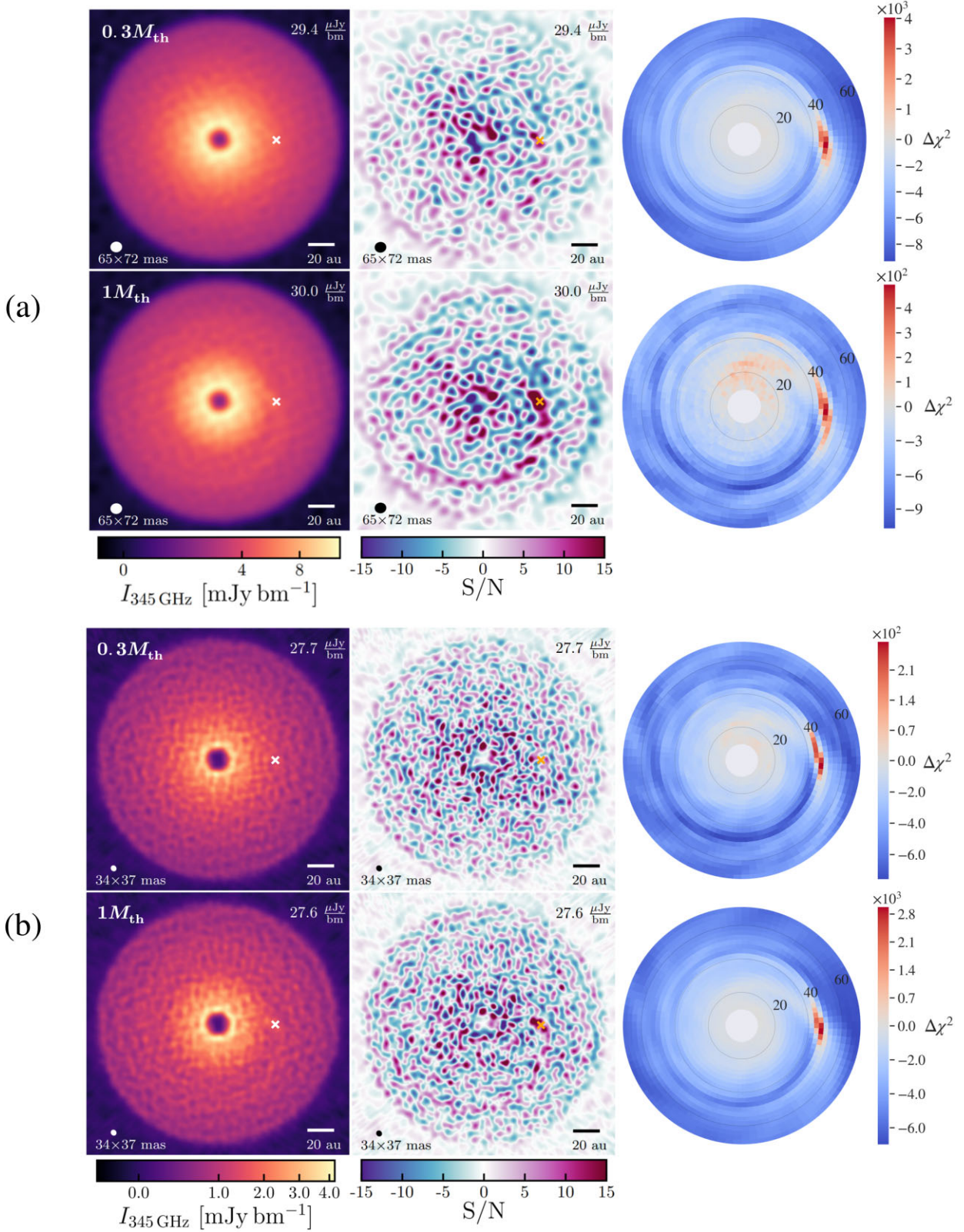


Figure 8. Planet recovery for $\beta = 10$, $\tau_0 = 3$ discs from our (a) (top two rows) fiducial set (0.061 arcsec resolution) and (b) (bottom two rows) higher resolution set (0.028 arcsec resolution). Each set has a requested sensitivity of $35 \mu\text{Jybeam}^{-1}$ (40 min of on-source time), and contains a planet of mass $0.3 M_{\text{th}}$ (top) or $1 M_{\text{th}}$ (bottom). Left: CLEAN image. Centre: Corresponding image residuals (generated by subtraction of the azimuthally averaged disc brightness). Right: Corresponding $\Delta\chi^2$ heatmaps. Despite the spiral only being noticeable in one of the image residuals, the $\Delta\chi^2$ heatmaps clearly recover the planet in all four cases. In the synthetic observations and image residuals: the cross indicates the planet position (50 au), the measured rms noise in each observation after CLEANing is written in the top right corner of each panel (and in each case is ~ 10 – 20 per cent less than the requested sensitivity), and the synthesized beam is shown in the bottom left corner. The radial ticks in the heatmaps are in au.

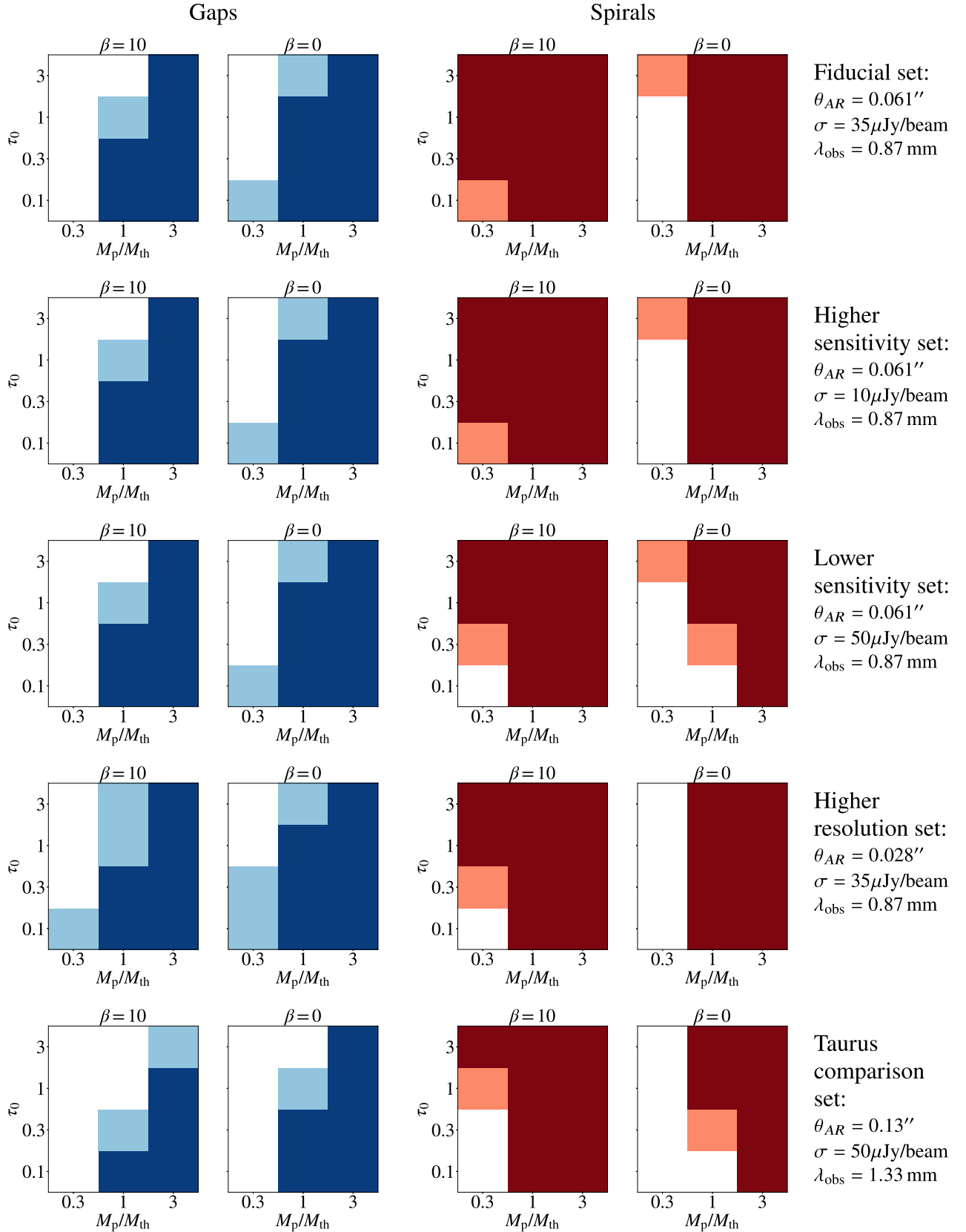


Figure 9. Planet recovery across the spiral simulation parameters (cooling time β , optical thickness τ_0 , and planet mass M_p) for gaps (left) and spirals (right), using clean image profiles and visibility-space fitting, respectively. White squares indicate no detection, light blue/red indicates a marginal gap/spiral detection, and a dark blue/red indicates a clear detection. The rows correspond to our various observation sets, labelled on the right by their descriptor from Table 1 along with their angular resolution, sensitivity, and observational wavelength. The radial profiles and $\Delta\chi^2$ heatmaps for our fiducial set are given in Appendix C.

Table 2. Parameters of our Taurus sample discs. d is the distance to the source (using *Gaia* DR2 measurements from Bailer-Jones et al. 2018). Total specific flux $F_{1.33\text{mm}}$ and $R_{\text{eff},95\%}$ (the disc radius containing 95 per cent of the flux) are derived from the FRANK profiles for each disc. The inclination (Inc.) and position angle (PA) are from Long et al. (2019).

Disc	d (pc)	$F_{1.33\text{mm}}$ (Jy)	$R_{\text{eff},95\%}$ (arcsec)	Inc. ($^\circ$)	PA ($^\circ$)
Compact discs					
BP Tau	129	0.045	0.321	$38.2^{+0.5}_{-0.5}$	$151.1^{+1.0}_{-1.0}$
DO Tau	139	0.12	0.263	$27.6^{+0.3}_{-0.3}$	$170.0^{+0.9}_{-0.9}$
DR Tau	195	0.13	0.276	$5.4^{+2.1}_{-2.6}$	$3.4^{+8.2}_{-8.0}$
FT Tau	127	0.091	0.357	$35.5^{+0.4}_{-0.4}$	$121.8^{+0.7}_{-0.7}$
Extended discs					
RY Tau	128	0.21	0.509	$65.0^{+0.1}_{-0.1}$	$23.1^{+0.1}_{-0.1}$
UZ Tau E	131	0.13	0.667	$56.1^{+0.4}_{-0.4}$	$90.4^{+0.4}_{-0.4}$

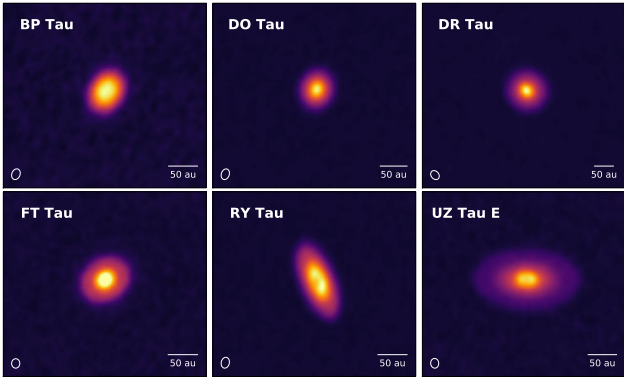


Figure 10. The 1.33 mm CLEAN images for our disc sample. Each panel is $2.4 \text{ arcsec} \times 2.4 \text{ arcsec}$, with the synthesized beam shown in the bottom left corner of each panel.

To explore the effect of observational parameters (angular resolution and sensitivity), we tested five different ALMA set-ups, which are described in Table 1. The recovery levels across the spiral simulation parameters for each are shown in Fig. 9. The complete set of radial profiles and $\Delta\chi^2$ heatmaps used in determining the recovery levels of our fiducial set are shown in Figs C1–C5.

(i) **Sensitivity.** For gaps, we see that sensitivity has a negligible effect on the CLEAN image profiles, with no change in recovery levels from the lower sensitivity to fiducial to higher sensitivity sets. This is unsurprising since, once one achieves a reasonable sensitivity, the estimate of mean intensity in a given radial bin will already well-approximate the true mean, and further increases in sensitivity will have little effect.

For spirals, we find that increasing sensitivity increases spiral model support (evidenced by larger ΔBIC values), but has negligible effect on the spatial structure of $\Delta\chi^2$ seen in the heatmaps. This is because the noise at each visibility point is largely independent of observing time, and the $\Delta\chi^2$ at each visibility point characterizes the strength of the spiral relative to the axisymmetric background. Therefore, for a given ALMA configuration and spiral model instance, $\Delta\chi^2$ will scale linearly with the number of visibility points (or equivalently, the observing time), increasing model support but leaving the spatial structure of $\Delta\chi^2$ unchanged. For the lower sensitivity set, spiral

model support is the limiting factor for spiral detection. Hence, going from the lower sensitivity to fiducial set results in improved recovery levels. For the fiducial set, however, the limiting factor for spiral detection becomes identifying the asymmetry (that is providing that spiral model support) as clearly spiral like. Since we rely on the spatial structure of $\Delta\chi^2$ for this, going from the fiducial to higher sensitivity set brings no improvement in recovery levels.

This is in stark contrast to the results for images in S22, which show a strong increase in spiral recovery for sensitivities beyond (our fiducial set's) $35 \mu\text{Jy beam}^{-1}$ (see e.g. their fig. 7). However, only for our higher sensitivity set (which offer the best spiral recovery in images of the observations considered here or in S22) does the spiral recovery in images approach that of visibility-space fitting.

(ii) **Resolution.** For gaps, our higher resolution set does slightly better than our fiducial set, with a few additional marginal recoveries, due to better resolving the gap.

For spirals, our higher resolution set gives $\Delta\chi^2$ heatmaps with spatial structures that indicate spirals slightly more clearly than our fiducial set. However, for this improvement they trade a decrease in $\Delta\chi^2$ magnitude, giving weaker spiral model support. This demonstrates the trade-off between spiral strength and clarity with resolution at fixed sensitivity mentioned in Section 3.2. In this case, as for image residuals (S22), the balance slightly favours a lower resolution (fiducial set).

Our Taurus comparison set's observational parameters are chosen to align closely with the Taurus disc sample analysed in Section 4. They are observed at a slightly longer wavelength of 1.33 mm, and are our lowest resolution and lowest sensitivity set (joint with our lower sensitivity set). As for the fiducial and higher sensitivity sets, the limiting factor for spiral detection in this set is the strength of the spiral signal. Comparing the detections in this set with the lower sensitivity set (which has the same sensitivity but higher resolution) suggests that, in this case, the trade-off between spiral strength and clarity favours a higher resolution, in contrast to the fiducial set versus higher resolution set comparison.

For these Taurus comparison observations, we also note a peculiar result: the spirals in the $M_p = 3 M_{\text{th}}$, $\beta = 0$ discs are not clearly recoverable with their *matching* spiral templates, with only the $\tau_0 = 0.1$ disc showing marginal spiral recovery, and the rest showing no recovery. All the planets can be clearly recovered, however, using *non-matching* templates with lower amplitudes, achieved by increasing τ_0 or decreasing M_p . For example, the planet in the $\tau_0 = 1$ disc can be recovered using the otherwise matching $M_p = 1 M_{\text{th}}$ template (as shown in the bottom right of Fig. 11). The reason for these template amplitudes being too high for their matching model discs may be explained by errors in the FRANK profiles. The FRANK fits produce profiles that have slightly brighter rings and darker gaps than the true profiles of the discs, an error likely due to their sharp inner and outer edges (Jennings et al. 2020). The bright ring results in the template inducing a spiral amplitude that is too large in the ring's vicinity. As a result, templates using a lower amplitude reproduce the data more accurately. While this issue could arise when fitting real observations, real discs are unlikely to have sharp edges and thus the effect should be weaker.

Our key conclusions from this section are as follows: planets are likely easier to detect in discs with shorter cooling times and lower optical depths via their gap signature. Many such gaps have already been detected in the field and are often interpreted as evidencing a planet. In contrast, in discs with longer cooling times and higher optical depths, planets may be easier to detect via their spiral signature. However, with typical image-space approaches,

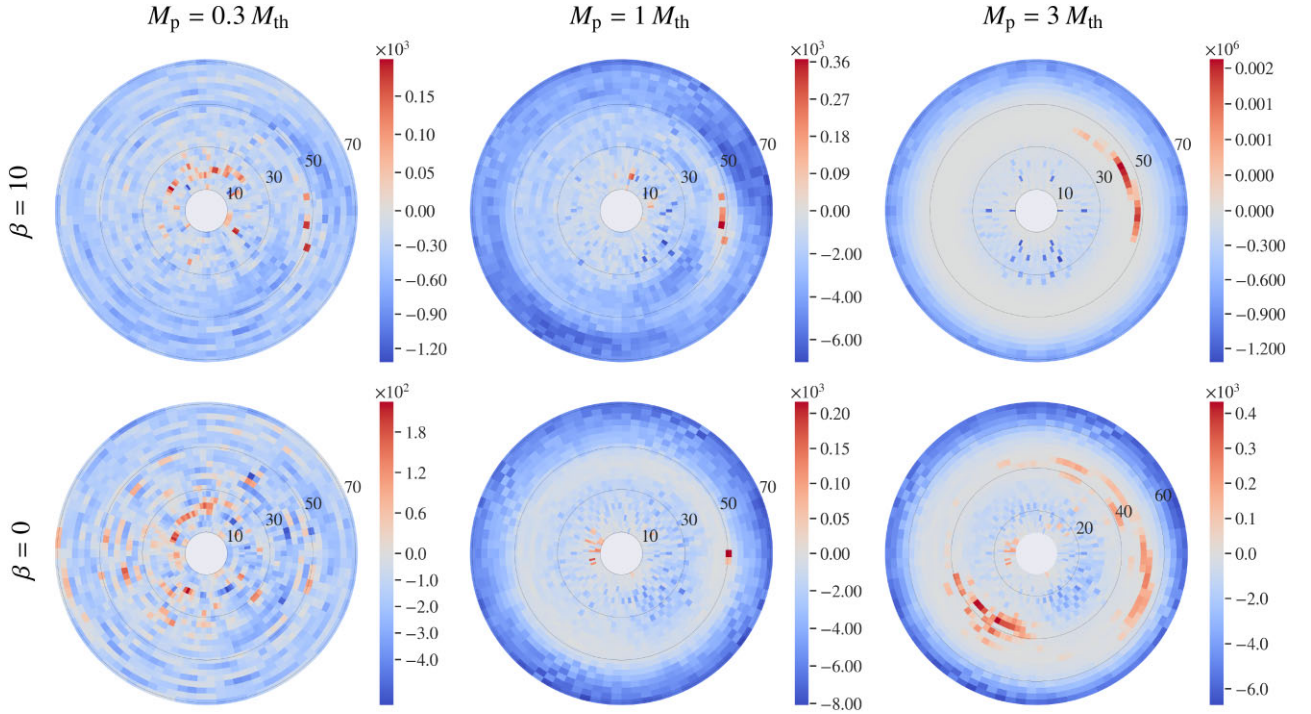


Figure 11. $\Delta\chi^2$ heatmaps of $\tau_0 = 1$, right-handed model discs with their matching spiral templates (except for the $\beta = 0$, $M_p = 3 M_{\text{th}}$ disc, whose planet is not recovered with its matching template, but is with the otherwise matching $M_p = 1 M_{\text{th}}$ template (among others)). $M_p/M_{\text{th}} = 0.3, 1, 3$ from left to right, and $\beta = 0, 10$ from bottom to top. The discs are observed at 1.33 mm with an angular resolution of 0.13 arcsec and sensitivity of $50 \mu\text{Jybeam}^{-1}$ (ALMA C-6 configuration with 6 min of on-source time).

these spiral detections could require very high sensitivities (e.g. the $10 \mu\text{Jybeam}^{-1}$ of our higher sensitivity set, requiring 8 h of on-source time), which could partly explain the paucity of detections. A visibility-space fitting approach may be able to achieve sufficient detection capabilities with much more moderate sensitivities (e.g. the $35 \mu\text{Jybeam}^{-1}$ of our fiducial set, requiring only 40 min of on-source time).

In complement to this section, in Appendix C we provide the full set of radial profiles and $\Delta\chi^2$ heatmaps for our fiducial set of observations, and provide interpretations of the heatmaps that we made in arriving at our spiral recovery levels (the middle row of Fig. 9).

4 RESULTS: TAURUS DISC SAMPLE

In this section, we apply our method to six ‘smooth’ (relative lack of annular substructure) discs from the ALMA Taurus survey presented in Long et al. (2019) (Table 2, Fig. 10). The Taurus survey observations have an angular resolution ~ 0.12 arcsec and a mean sensitivity $\sim 50 \mu\text{Jybeam}^{-1}$ at 1.33 mm (Band 6), resulting from on-source times between 4 and 10 min in the C-6 configuration. To generate the Taurus discs’ radial brightness profiles in our method, we use the same `FRANK` hyperparameters as for the synthetic observations. Again, the resulting profiles are quite insensitive to variations within the recommended ranges.

Our Taurus sample’s mean observational parameters (wavelength, resolution, and sensitivity) align closely with our Taurus comparison set of synthetic observations (Fig. 1, the bottom row of Fig. 9). Their total fluxes, derived from the `FRANK` profiles, range from ≈ 0.05 to 0.2 Jy (BP Tau to RY Tau). Those of our Taurus comparison set are generally somewhat brighter, and range from ≈ 0.2 to 1.3 Jy

($\tau_0 = 0.1$ to 3, with some small variations with planet mass and cooling time). The synthetic observations are therefore effectively somewhat more sensitive than our Taurus observations. The effective radii of our Taurus sample discs (mean $R_{\text{eff},95 \text{ per cent}} \approx 55$ au) are also somewhat smaller than that of our model discs (≈ 95 au). Overall, the Taurus sample and comparison synthetic observations are broadly comparable, and we can use the successful spiral recoveries achieved in those synthetic observations to inform what a spiral-like signal should look like in the $\Delta\chi^2$ heatmaps of the Taurus sample. The heatmaps of an illustrative set of these synthetic observations is shown in Fig. 11. Recall that $\Delta\chi^2 = \chi_{AO}^2 - \chi_{SO}^2$, i.e. the difference in goodness of fit between the axisymmetric model and spiral model instance. Positive (negative) $\Delta\chi^2$ values represent an improvement (decline) in fit with the spiral model instance. Since, of the simulation parameters, optical thickness has the smallest effect on spiral morphology (mostly just acting to increase or decrease the amplitude of the spiral perturbation), we present the heatmaps in this section at a fixed optical thickness ($\tau_0 = 1$) for simplicity.

4.1 Spiral fits

Of the six discs in our Taurus sample (Table 2 and Fig. 10), BP Tau, DO Tau, FT Tau, and UZ Tau E show little asymmetry, each giving $\Delta\text{BIC} < -10$, and therefore strongly favouring the axisymmetric model. In contrast, DR Tau and RY Tau exhibit consistent localized regions of $\Delta\chi^2 \sim 100$ in the inner disc (e.g. Fig. 13), strongly favouring the spiral model and passing our first test (Section 2.3). They do not, however, pass our second test. Both fail to clearly distinguish handedness, with left- and right-handed templates achieving similar peak $\Delta\chi^2$. Hence, we do not find evidence for planet-driven spirals in these discs. Nevertheless, the $\Delta\chi^2$ heatmaps

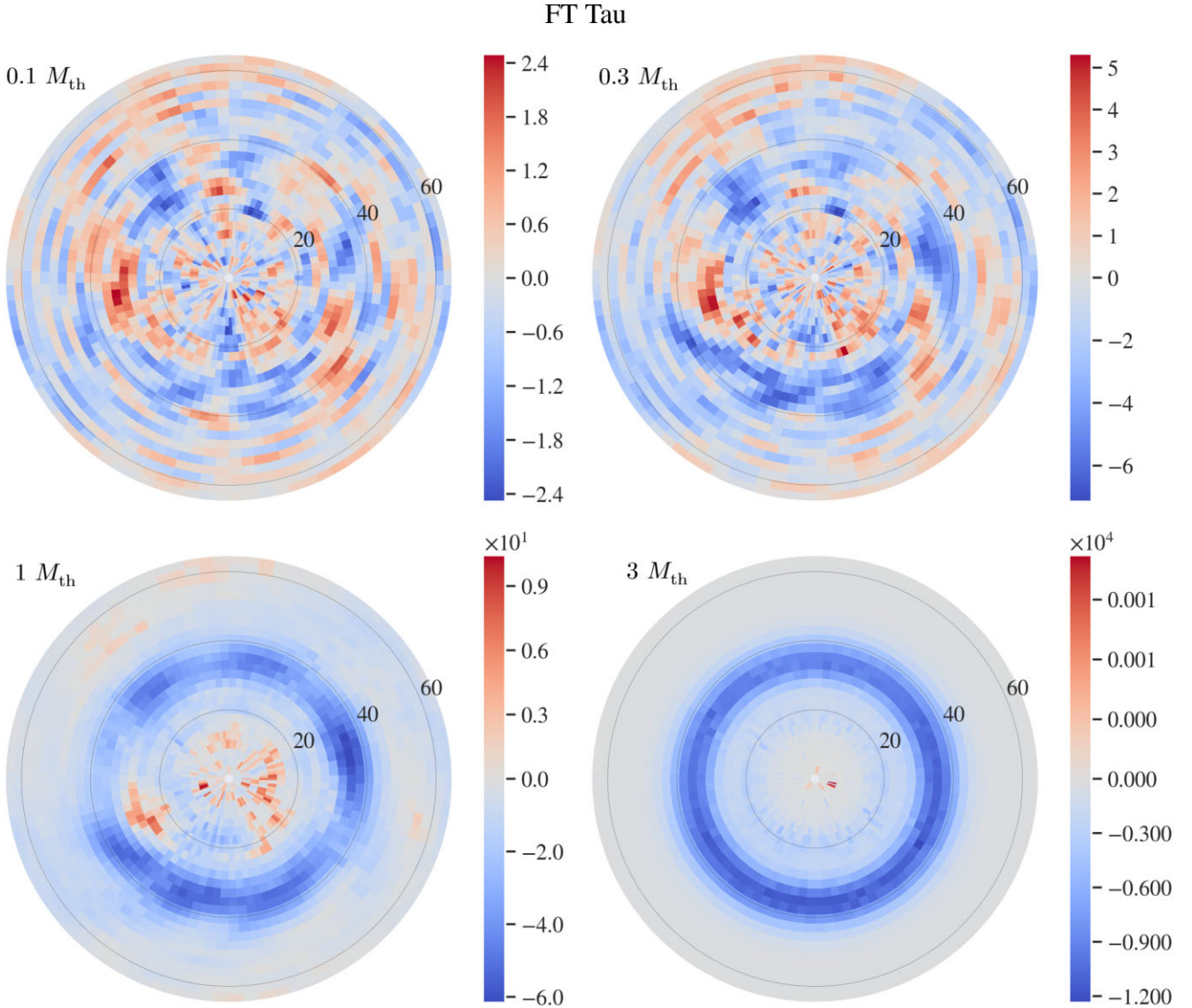


Figure 12. $\Delta\chi^2$ heatmaps of FT Tau for $\beta = 10$, $\tau_0 = 1$, and right-handed spiral templates. The heatmap spatially represents the deprojected disc, with the disc’s minor axis projected onto the horizontal axis ($\phi_p = 0$). The radial ticks are in au.

provide insights into the nature of the observations and the method, and the possibility of planets hiding in the discs.

Using FT Tau as an example of the low-asymmetry discs, the $\Delta\chi^2$ heatmaps in Fig. 12 for the high-mass (1 and $3 M_{\text{th}}$) planets exhibit large negative $\Delta\chi^2$ (~ -50 for the $1 M_{\text{th}}$ case and $\sim -10^4$ for the $3 M_{\text{th}}$ case) over significant regions of the disc. These planets therefore clearly have a significant effect on the disc structure and would be detectable if they were present, as suggested by the Taurus comparison set. Considering this, one can rule out the presence of these higher-mass planets over the majority of the radial extent of FT Tau. For example, it is evident from the $\Delta\chi^2$ in the $3 M_{\text{th}}$ heatmap (the bottom right-hand panel of Fig. 12) that a $3 M_{\text{th}}$ planet residing between 10 and 60 au is unlikely. It is more difficult to draw conclusions about planets in the inner disc, where there are some $\Delta\chi^2 > 0$ planet positions and results are inherently unreliable due to the spiral contributing on scales smaller than the angular resolution of the observations. Similarly, conclusions for large r_p are limited by low disc brightness. We apply this reasoning more fully in Section 4.2.

The two blue rings (at ≈ 18 and 34 au) visible in the $3 M_{\text{th}}$ heatmap of Fig. 12 coincide with radii of elevated brightness (and are consistent with the radial χ^2_{SA} plots; see Fig. B2).

For the low-mass (0.1 and $0.3 M_{\text{th}}$) planets, the large negative $\Delta\chi^2$ disappear. Instead, the heatmaps appear ‘noisy’ and have small $\Delta\chi^2$ values ($|\Delta\chi^2| < 10$), roughly symmetrically distributed about zero. This is suggestive of a detectability limit in the observations; if a low-mass planet were present in the disc, its spiral would likely be indistinguishable from observational noise. The equivalent heatmaps of RY Tau and UZ Tau E (the extended discs) generally show slightly more structure than the compact discs, possibly due to their larger size and brightness allowing for effectively a greater resolution and sensitivity. Both sets of discs, however, suggest a minimum planet mass for detectability of between 0.3 and $1 M_{\text{th}}$ across the disc radii considered, which is broadly similar to the range suggested by the Taurus comparison set (as indicated in the bottom right of Fig. 9).

The $\Delta\chi^2$ analysis for the two more asymmetric discs (DR Tau and RY Tau) shares many of the same features as for the low-asymmetry

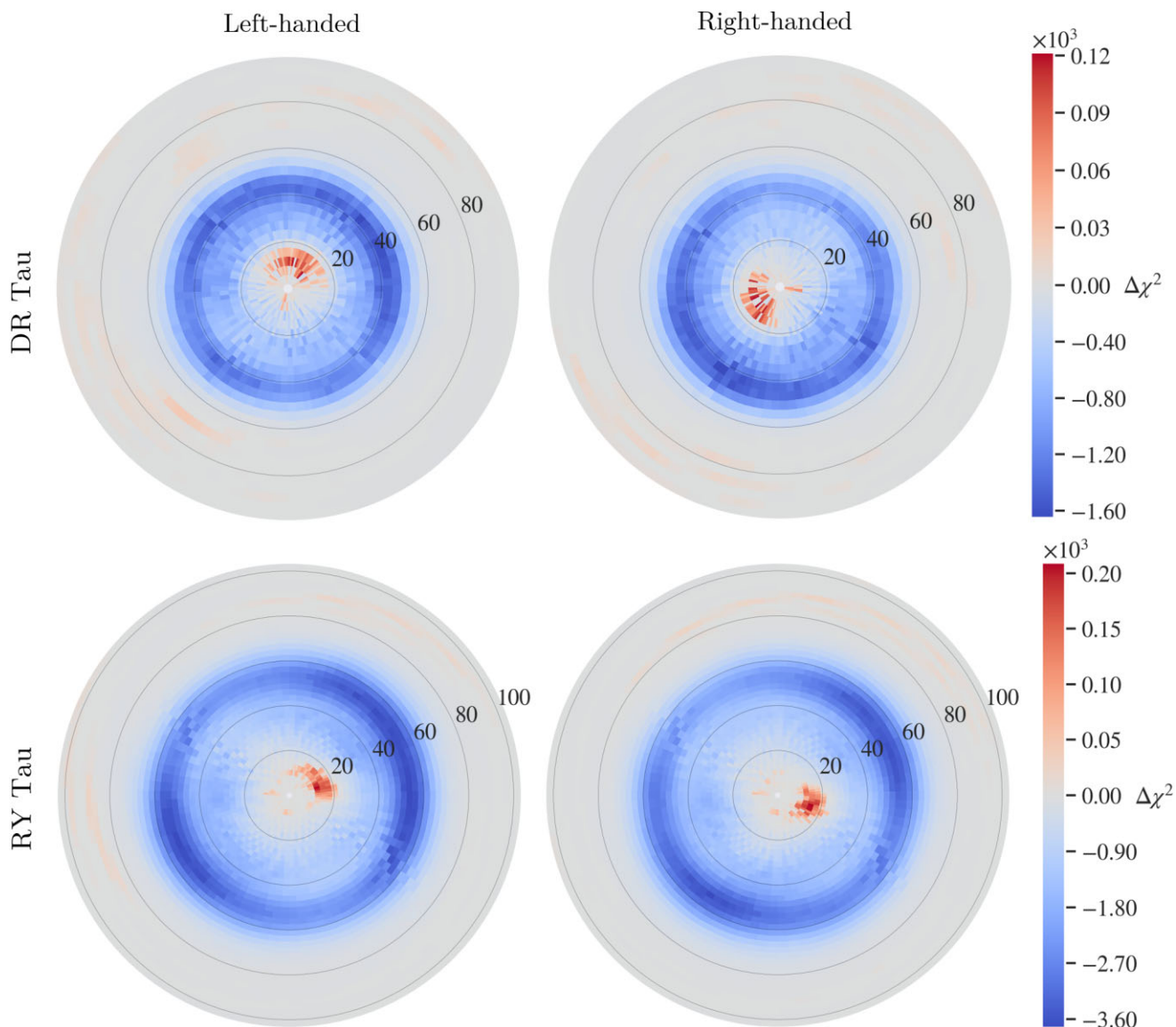


Figure 13. $\Delta\chi^2$ heatmaps of DR Tau (top row) and RY Tau (bottom row) for $\beta = 10$, $\tau_0 = 3$, $M_p = 3 M_{\text{th}}$ left- and right-handed spiral templates. DR Tau and RY Tau show maximum $\Delta\chi^2 \approx 100$ and 200, respectively, with similar results for both hands. The colourbars apply to each row of heatmaps.

discs: noisy heatmaps suggesting the low detectability of low-mass planets, and blue rings and consistent $\Delta\chi^2 < 0$ suggesting a lack of high-mass planets across most of the disc. However, both of these discs exhibit much larger $\Delta\chi^2 > 0$ at inner radii ($0.04 \text{ arcmin} \leq r_p \leq 0.08 \text{ arcmin}$ for DR Tau and $0.08 \text{ arcmin} \leq r_p \leq 0.18 \text{ arcmin}$ for RY Tau, which has an inner cavity) for the high-mass planets, suggesting a much stronger asymmetry in these discs. These regions of improvement are localized and consistent across different spiral templates.⁸

However, as can be seen in Fig. 13, spiral templates of either handedness result in similarly large peak $\Delta\chi^2$ values. This indicates that the templates are probably not fitting genuine spirals, but some other asymmetry. The small position offsets between the peak left- and right-handed $\Delta\chi^2$ for RY Tau suggest that these high-mass

planet templates could be fitting some arc-like asymmetry in the disc. In fact, higher resolution scattered light observations for RY Tau (Francis & van der Marel 2020) show such an arc, extending over $\sim 180^\circ$ in azimuth, at the same location as indicated by the $\Delta\chi^2$ heatmaps. Since RY Tau has a significant inner cavity and is highly inclined (65.0°), this could even be from the viewing angle of the hotter disc wall at the edge of the cavity (Ribas et al., in preparation), rather than any asymmetry in the disc density.

The fitting in the equivalent heatmaps of DR Tau show a slightly more ‘spotty’ structure that broadly resembles the false fitting seen for errors in the disc phase centre (Appendix A). Jennings et al. (2022) find the same asymmetry in their imaged FRANK residuals and show that it is not due to an incorrect determination of the disc phase centre (see their appendix A), instead concluding that there is genuine unresolved inner disc structure. DR Tau is also only slightly inclined (5.4°), hence the viewing angle of a disc wall is not a likely explanation. Other possible explanations include an azimuthal dust trap (such as in Oph-IRS 48; Bruderer et al. 2014), or a vortex (such

⁸They are also in reasonable agreement with where one might expect an asymmetry to be from the corresponding imaged FRANK residuals in figs 6 and 11 of Jennings et al. (2022).

as in HD 34282; Marr & Dong 2022). Such structures could of course point to the presence of a planet in the disc, but the evidence would be separate to the spirals considered in this work.

4.2 Which planets can we rule out?

Extending the reasoning in the previous section, for a given planet mass and location within a disc, we infer that there is unlikely to be a planet of that mass or greater at that location if all the plausible spiral templates which account for such a planet yield $\Delta\chi^2$ values which give $\Delta\text{BIC} < -10$.⁹ We also require that these planet masses are *capable* of inducing a significant $\Delta\chi^2$, which is indicated by a large negative $\Delta\chi^2$ at other comparable locations in the disc. The first condition is a requirement for there being no evidence for the planet, and the second is a requirement for the planet to be detectable.

Hence, using the $\Delta\chi^2$ heatmaps, we can suggest planets to reject over ranges of disc radii. Table 3 provides suggestions for the r_p ranges over which one can rule out planets greater than 1 and $3 M_{\text{th}}$ (≈ 0.3 and $0.9 M_{\text{Jup}}$ at 50 au) based on these considerations.

From Table 3, we can place an upper mass limit of approximately $3 M_{\text{th}}$ for planets at intermediate radii of the discs, with conclusions at the innermost and outermost radii being limited by insufficient resolution and low disc brightness, respectively. In each case, the outermost radii are notably beyond the effective ‘edges’ of the discs in 1.33 mm. We can drop this limit to approximately $1 M_{\text{th}}$ over most of DO Tau, and smaller annuli (centred on bright rings) in FT Tau and UZ Tau E. Recalling that our models have $M_{\text{th}} \propto r_p^{3/4}$, we note that these thermal mass limits correspond to different absolute masses at different orbital radii. The mapping between thermal masses and Jupiter masses is shown in Fig. 14, and the resulting M_p ranges are included in Table 3.

One can compare our locus of upper mass limits to the masses and orbital radii Lodato et al. (2019) inferred from the Taurus disc sample in Long et al. (2018) (Fig. 14). Their focus is on radially structured discs, whereas our sample consists of relatively smooth discs. Only FT Tau, RY Tau, and UZ Tau E from our sample are included in their analysis, with their inferred embedded planet masses sitting well below the masses that we rule out. RY Tau and UZ Tau E’s inferred planets also sit notably below our suggested minimum mass locus for driving a detectable spiral ($0.3\text{--}1 M_{\text{th}}$ across the disc radii considered), with FT Tau’s sitting in the zone between detectable and undetectable. Therefore, if such planets are present, they are unlikely to be detectable via their spiral signature in these observations. It is worth noting that Lodato et al. (2019) did not impose a minimum gap depth to make these inferences, and the gaps appear quite shallow in the discs’ parametric profiles (UZ Tau E’s in particular) that they used (from Long et al. 2018).¹⁰

A number of the gaps considered in Lodato et al. (2019) (from various sources) do yield inferred planet masses sitting above the lower limit for detectability. This limit is also likely lower for discs observed at higher sensitivities and resolutions than the sample from Long et al. (2018), such as the DSHARP discs whose gap widths were measured in Zhang et al. (2018). The wider gaps in these disc

observations will give us a better chance to support or oppose planets as their progenitors.

5 DISCUSSION

The method described in this work is the first implementation of a visibility-space spiral fitting approach, and it faces a number of issues and uncertainties. Rather than offering a ready-to-go visibility-space method for finding planet-driven spirals, this paper is intended as a proof-of-concept for this approach and a foundation for a more sophisticated method based upon it. Here, we give a couple key issues that the approach faces, and offer suggestions for how they could be addressed in future work.

5.1 Misleading signals

As we saw for DR Tau and RY Tau in Section 4.1, non-spiral asymmetries such as arcs can be fit by the spiral model (to a statistically significant degree). In order to reliably recover a planet-driven spiral, it is necessary to distinguish between these misleading signals and true signals. The handedness test (Fig. 4) is an example of an initial attempt at this. This can be readily improved upon, for instance, by introducing models of non-spiral asymmetries (e.g. bright spots or arcs) and comparing their signal to that of the spiral model. Distinguishing between the morphology of spirals driven by planets, and those driven by other mechanisms (e.g. stellar fly-bys or gravitational instability) may also be worthwhile.

5.2 Alternative spiral morphologies

The model discs present more favourable observations for our method than would be the case for real discs because of the perfect correspondence between their spiral and the method’s spiral templates. The multitude of assumptions made in generating these spiral simulations will undoubtedly result in a deviation from the planet-driven spiral structure in real discs. Our tests in Section 3.1 suggest that this may not be a big problem for planet recovery as the method can often pick out the presence of a spiral when using templates that do not exactly match the model disc’s (e.g. Figs 5 and 7). It is possible, however, that certain parameters assumed constant here, or neglected physics, could have a stronger effect on planet recovery. Examples include: the assumption of the planet being on a circular, co-planar, and non-migrating orbit (Quillen et al. 2005; Duffell & Chiang 2015), viscosity (Fung, Shi & Chiang 2014), dust scattering (Sierra & Lizano 2020), vertical disc structure (Krapp, Kratter & Youdin 2022), disc inclination (S22), MHD winds (Wafflard-Fernandez & Lesur 2023), and interfering disc substructures (e.g. spirals/gaps from other planets). Related to this, the degeneracy in fit across cooling time, optical depth, and planet mass (particularly optical depth and planet mass), as well as simulation assumptions and neglected physics, make it difficult to infer planet mass from a spiral signal (a similar degeneracy occurs for images; S22), although with constraints on the cooling time-scale and optical thickness from other data this would improve. Such constraints would lead to a similar improvement in our ability to rule out planets (as done in Section 4.2).

The extent of these limitations depends on the coverage of possible spiral morphologies provided by the spiral simulations we use in the fitting approach. The method which we developed here employs the pre-generated hydrodynamical spiral simulations of S22. Whilst these ‘first-principles’ simulations have many advantages, they are computationally expensive and slow. A shift towards faster simulation methodologies would enable a larger and finer exploration of the

⁹The BIC is strictly only valid for the maximum $\Delta\chi^2$ achieved by the spiral model, but it can still serve as a guide to which $\Delta\chi^2 > 0$ values are plainly insignificant.

¹⁰They remain shallow in the discs’ FRANK profiles and do not appear in their CLEAN image profiles. See figs 6 and 11 of Jennings et al. (2022) for comparisons of these three profiles for each of the discs.

Table 3. Approximate orbital radii ranges for which we can rule out planets of masses greater than 3 and $1 M_{\text{th}}$ for our Taurus sample discs. The corresponding planet mass range for the 3 and $1 M_{\text{th}}$ lower limits are given in the adjacent columns to the right. $R_{\text{eff},95\%}$ is the disc radius containing 95 per cent of the total flux. The final two columns give the planet orbital radius and mass inferred from the gap present in these discs (table 1 of Lodato et al. 2019).

Disc	$R_{\text{eff},95\%}$ (au)	r_p exclusion range for $M_p \geq 3M_{\text{th}}$ (au)	$M_p = 3M_{\text{th}}$ mass range (M_{Jup})	From Lodato et al. (2019)			
				r_p exclusion range for $M_p \geq 1M_{\text{th}}$ (au)	$M_p = 1M_{\text{th}}$ mass range (M_{Jup})	Inferred r_p (au)	Inferred M_p (M_{Jup})
BP Tau	41	(10, 54)	(0.26, 0.91)	–	–	–	–
DO Tau	37	(8, 61)	(0.22, 1.00)	(11, 61)	(0.09, 0.33)	–	–
DR Tau	54	(23, 62)	(0.48, 1.01)	–	–	–	–
FT Tau	45	(10, 64)	(0.26, 1.04)	(33, 43)	(0.21, 0.26)	25	0.15
RY Tau	65	(23, 82)	(0.48, 1.25)	–	–	43	0.077
UZ Tau E	87	(13, 97)	(0.31, 1.42)	(55, 63)	(0.31, 0.34)	69	0.023

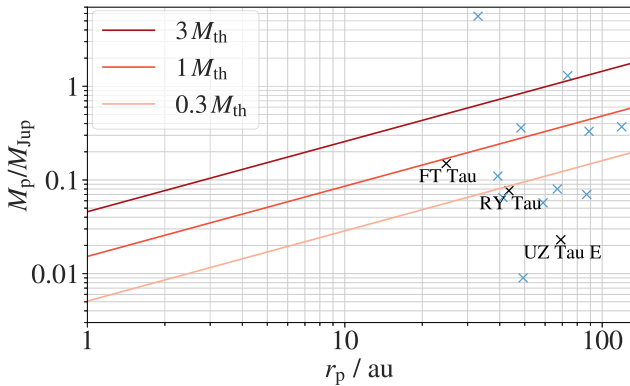


Figure 14. Planet mass against orbital radius for 0.3, 1, and 3 thermal masses ($M_{\text{th}} \propto r_p^{3/4}$). The minimum mass for driving a detectable spiral is between 0.3 and $1 M_{\text{th}}$ for disc observations like those presented in Long et al. (2019). The crosses indicate the planet masses and orbital radii inferred by Lodato et al. (2019) for the disc sample in Long et al. (2018) (excluding the $15.7 M_{\text{Jup}}$ planet inferred for CI Tau). The labelled black crosses indicate discs that we also consider in this work.

parameter space, allowing for a greater coverage of possible spiral morphologies (including possibly other non-spiral asymmetries) and tighter constraints on planet masses.

One possibility for achieving this is machine learning models. For instance, PPDONET (Mao et al. 2023) can produce 2D gas surface density perturbation maps for arbitrary combinations of planet mass, disc viscosity, and disc aspect ratio in ‘less than a second on a laptop’. Alternatively, one could take the (semi-)analytic route; WAKEFLOW (Bollati et al. 2021; Hilder et al. 2023) can swiftly generate 2D gas surface density perturbation heatmaps for a range of planet and disc properties, and has the added advantage of interfacing with MCFOST (Pinte et al. 2006, 2009) to perform the necessary radiative transfer.

6 SUMMARY AND CONCLUSIONS

In this paper, we have explored the planet detection capabilities of fitting planet-driven dust spirals to protoplanetary disc observations in visibility space. We devised a method for this approach in Section 2 which combines an axisymmetric model of the given disc with a series of planet-driven spiral simulations to create a series of spiral-containing model discs. The model disc images are transformed to visibility space and fit to the observed visibilities using χ^2 statistics. Through our tests on synthetic disc observations from an extended version of the set generated in Speedie et al. (2022, hereafter S22),

we explored how different model parameters affect the fitting results, and compared its ability to identify planets with looking for spirals in CLEAN image residuals and gaps in CLEAN image profiles. We also applied our method to six smooth discs from the ALMA Taurus survey presented in Long et al. (2019) (Section 4). Our main conclusions can be summarized as follows:

(i) The method developed in this work significantly outperforms image residuals in recovering spiral-driving planets from the synthetic observations, showing clear signals for spirals well beyond the detection limits of image residuals (Fig. 8). The planet recovery ability of the method may also be less vulnerable to interference from annular gaps than image-space methods, and could potentially be used to support or oppose a planet origin for such observed gaps. Taken together, the discovery of a spiral co-located with a gap would constitute strong evidence for a planet.

(ii) The trends in the detectability of gaps and spirals with cooling time-scale and optical depth are opposite, with spirals being easier to detect for longer cooling times and higher optical depths, while gaps are more easily opened for shorter cooling times and lower optical depths. In the synthetic observations, spirals were recovered in significant regions of the disc/observational parameter space where gaps were not (Fig. 9).

(iii) Our sample of Taurus disc observations and the synthetic Taurus comparison set suggest planets with masses > 0.3 – $1 M_{\text{th}}$ can drive detectable spirals in such observations. This corresponds to an absolute mass range of ≈ 0.03 – $0.5 M_{\text{Jup}}$ over orbital radii of 10–100 au.

(iv) In our sample of Taurus discs, we find evidence of inner disc asymmetries in DR Tau and RY Tau, but no planet-driven spirals (Section 4.1). Higher resolution/sensitivity disc observations (of which many already exist), and more structured discs, may offer a better chance of detection.

(v) For our sample of Taurus discs, we can reasonably rule out the existence of any planets with masses $\gtrsim 3 M_{\text{th}}$ over orbital radii of 20–60 au, corresponding to a mass range of ≈ 0.5 – $1 M_{\text{Jup}}$. The exclusion of $\gtrsim 1 M_{\text{th}}$ planets is possible over more limited radii ranges for some of the discs (Table 3).

(vi) The method developed in this work is the first implementation of a visibility-space spiral fitting approach, and it has a large scope for improvement. Despite this, it demonstrated a promising ability to recover planet-driven spirals in the synthetic observations. This suggests that such an approach warrants further investigation and may be able to find planet-driven spirals in disc observations that have not yet been found with existing methods.

ACKNOWLEDGEMENTS

ETS thanks the Leverhulme Centre for Life in the Universe for their support via an internship. AR has been supported by the UK Science and Technology Facilities Council (STFC) via the consolidated grant ST/W000997/1 and by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement no. 823823 (RISE DUSTBUSTERS project). RAB thanks the Royal Society for their support through a University Research Fellowship. This research was enabled in part by support provided by the Digital Research Alliance of Canada (alliancecan.ca).

DATA AVAILABILITY

The data underlying this article are available at doi.org/10.6084/m9.figshare.25464109.

REFERENCES

- ALMA Partnership et al., 2015, *ApJ*, 808, L3
 Andrews S. M. et al., 2016, *ApJ*, 820, L40
 Andrews S. M. et al., 2018, *ApJ*, 869, L41
 Andrews S. M. et al., 2021, *ApJ*, 916, 51
 Bae J., Zhu Z., 2018, *ApJ*, 859, 118
 Bae J., Teague R., Zhu Z., 2021, *ApJ*, 912, 56
 Bailer-Jones C. A. L., Rybizki J., Fousneau M., Mantelet G., Andrae R., 2018, *AJ*, 156, 58
 Benisty M. et al., 2018, *A&A*, 619, A171
 Benítez-Llambay P., Masset F. S., 2016, *ApJS*, 223, 11
 Binkert F., Szulágyi J., Birnstiel T., 2023, *MNRAS*, 523, 55
 Bollati F., Lodato G., Price D. J., Pinte C., 2021, *MNRAS*, 504, 5444
 Booth R. A., Sijacki D., Clarke C. J., 2015, *MNRAS*, 452, 3932
 Briggs D. S., Schwab F. R., Sramek R. A., 1999, in Taylor G. B., Carilli C. L., Perley R. A., eds, ASP Conf. Ser. Vol. 180, Synthesis Imaging in Radio Astronomy II. Astron. Soc. Pac., San Francisco, p. 127
 Brown-Sevilla S. B. et al., 2021, *A&A*, 654, A35
 Bruderer S., van der Marel N., van Dishoeck E. F., van Kempen T. A., 2014, *A&A*, 562, A26
 Christiaens V., Casassus S., Perez S., van der Plas G., Ménard F., 2014, *ApJ*, 785, L12
 Clark B. G., 1980, *A&A*, 89, 377
 Clarke C. J. et al., 2018, *ApJ*, 866, L6
 Cornwell T. J., 2008, *IEEE J. Sel. Top. Signal Proc.*, 2, 793
 de Val-Borro M. et al., 2006, *MNRAS*, 370, 529
 Dong R., Fung J., 2017, *ApJ*, 835, 38
 Dong R., Zhu Z., Rafikov R. R., Stone J. M., 2015, *ApJ*, 809, L5
 Dong R. et al., 2018a, *ApJ*, 860, 124
 Dong R., Najita J. R., Brittain S., 2018b, *ApJ*, 862, 103
 Duffell P. C., Chiang E., 2015, *ApJ*, 812, 94
 Francis L., van der Marel N., 2020, *ApJ*, 892, 111
 Fukagawa M. et al., 2004, *ApJ*, 605, L53
 Fung J., Shi J.-M., Chiang E., 2014, *ApJ*, 782, 88
 Goldreich P., Tremaine S., 1980, *ApJ*, 241, 425
 Goodman J., Rafikov R. R., 2001, *ApJ*, 552, 793
 Hilder T., Fasano D., Bollati F., Vandenberg J., 2023, *J. Open Source Softw.*, 8, 4863
 Högbom J. A., 1974, *A&AS*, 15, 417
 Huang J. et al., 2018a, *ApJ*, 869, L42
 Huang J. et al., 2018b, *ApJ*, 869, L43
 Huang J. et al., 2020, *ApJ*, 891, 48
 Jennings J., Booth R. A., Tazzari M., Rosotti G. P., Clarke C. J., 2020, *MNRAS*, 495, 3209
 Jennings J., Tazzari M., Clarke C. J., Booth R. A., Rosotti G. P., 2022, *MNRAS*, 514, 6053
 Juhász A., Benisty M., Pohl A., Dullemond C. P., Dominik C., Paardekooper S.-J., 2015, *MNRAS*, 451, 1147
 Kass R. E., Raftery A. E., 1995, *J. Am. Stat. Assoc.*, 90, 773
 Keppler M. et al., 2019, *A&A*, 625, A118
 Kley W., Nelson R., 2012, *ARA&A*, 50, 211
 Krapp L., Kratter K. M., Youdin A. N., 2022, *ApJ*, 928, 156
 Kudo T., Hashimoto J., Muto T., Liu H. B., Dong R., Hasegawa Y., Tsukagoshi T., Konishi M., 2018, *ApJ*, 868, L5
 Lin D. N. C., Papaloizou J., 1986, *ApJ*, 309, 846
 Lodato G. et al., 2019, *MNRAS*, 486, 453
 Long F. et al., 2018, *ApJ*, 869, 17
 Long F. et al., 2019, *ApJ*, 882, 49
 Macías E., Guerra-Alvarado O., Carrasco-González C., Ribas Á., Espaillat C. C., Huang J., Andrews S. M., 2021, *A&A*, 648, A33
 Mao S., Dong R., Lu L., Yi K. M., Wang S., Perdikaris P., 2023, *ApJ*, 950, L12
 Marr M., Dong R., 2022, *ApJ*, 930, 80
 Mawet D. et al., 2012, *A&A*, 544, A131
 McMullin J. P., Waters B., Schiebel D., Young W., Golap K., 2007, in Shaw R. A., Hill F., Bell D. J., eds, ASP Conf. Ser. Vol. 376, Astronomical Data Analysis Software and Systems XVI. Astron. Soc. Pac., San Francisco, p. 127
 Meru F., Juhász A., Ilee J. D., Clarke C. J., Rosotti G. P., Booth R. A., 2017, *ApJ*, 839, L24
 Miranda R., Rafikov R. R., 2020, *ApJ*, 904, 121
 Muto T. et al., 2012, *ApJ*, 748, L22
 Norfolk B. J. et al., 2022, *ApJ*, 936, L4
 Pérez L. M. et al., 2016, *Science*, 353, 1519
 Pérez L. M. et al., 2018, *ApJ*, 869, L50
 Pérez S., Casassus S., Baruteau C., Dong R., Hales A., Cieza L., 2019, *AJ*, 158, 15
 Pinte C., Ménard F., Duchêne G., Bastien P., 2006, *A&A*, 459, 797
 Pinte C., Harries T. J., Min M., Watson A. M., Dullemond C. P., Woitke P., Ménard F., Durán-Rojas M. C., 2009, *A&A*, 498, 967
 Quillen A. C., Varnière P., Minchev I., Frank A., 2005, *AJ*, 129, 2481
 Rafikov R. R., 2002, *ApJ*, 569, 997
 Raftery A. E., 1995, *Sociol. Methodol.*, 25, 111
 Ribas Á. et al., 2023, *A&A*, 673, A77
 Rosotti G. P., Juhász A., Booth R. A., Clarke C. J., 2016, *MNRAS*, 459, 2790
 Rosotti G. P. et al., 2020, *MNRAS*, 491, 1335
 Schwarz G., 1978, *Ann. Stat.*, 6, 461
 Sierra A., Lizano S., 2020, *ApJ*, 892, 136
 Speedie J., Dong R., 2022, *ApJ*, 940, L43
 Speedie J., Booth R. A., Dong R., 2022, *ApJ*, 930, 40(S22)
 Tang Y.-W. et al., 2017, *ApJ*, 840, 32
 Tazzari M., Beaujean F., Testi L., 2018, *MNRAS*, 476, 4527
 Teague R., 2019, *J. Open Source Softw.*, 4, 1632
 Teague R., Bae J., Bergin E. A., Birnstiel T., Foreman-Mackey D., 2018a, *ApJ*, 860, L12
 Teague R., Bae J., Birnstiel T., Bergin E. A., 2018b, *ApJ*, 868, 113
 Teague R., Bae J., Huang J., Bergin E. A., 2019, *ApJ*, 884, L56
 Wafflard-Fernandez G., Lesur G., 2023, *A&A*, 677, A70
 Weber P., Pérez S., Benítez-Llambay P., Gressel O., Casassus S., Krapp L., 2019, *ApJ*, 884, 178
 Zhang S., Zhu Z., 2020, *MNRAS*, 493, 2287
 Zhang S. et al., 2018, *ApJ*, 869, L47
 Zhu W., Dong S., 2021, *ARA&A*, 59, 291
 Zhu Z., Dong R., Stone J. M., Rafikov R. R., 2015, *ApJ*, 813, 88

APPENDIX A: EFFECTS OF ERRORS IN THE FITTED DISC GEOMETRY

In our analysis of the synthetic observations, we assume that the disc phase centre, inclination and position angle are known. However, incorrect determination of these parameters can produce asymmetry artefacts that could potentially disrupt spiral recovery. Therefore, here we briefly examine how such errors manifest in the $\Delta\chi^2$ heatmaps [see appendix A of Andrews et al. (2021) for a similar discussion for image residuals] and how they affect spiral recovery.

We consider two example synthetic observations of a disc with geometric parameters $(\Delta\text{Dec}, \Delta\text{RA}) = (0 \text{ mas}, 0 \text{ mas})$, $\text{inc.} = 30^\circ$, $\text{PA} = 90^\circ$. Both discs have $\beta = 10$, $\tau_0 = 1$ and are observed with our fiducial set-up. One contains no planet, and the other contains a $0.3 M_{\text{th}}$ planet (at $r_p = 50 \text{ au}$, $\phi_p = 0^\circ$ [along the projected disc's minor axis]), and producing a right-handed spiral). We attempt to fit both with a $\beta = 10$, $\tau_0 = 1$, $M_p = 0.3 M_{\text{th}}$ spiral template (i.e. the template matching the parameters of the planet-containing disc). The $\Delta\chi^2$ heatmaps for the planet-less disc observation allow us to see the false fitting ($\Delta\text{BIC} > 0$ planet positions) that errors in disc geometry can introduce (Fig. A1), and those of the planet-containing disc allow us to see how this impacts recovery for our smallest planet case (Fig. A2). We expect the recovery of any larger mass planets to be less affected.

First we note that the effect on spiral recovery of the disc's inclination is small: compared to the face-on version, (Fig. C2) the peak $\Delta\chi^2$ is reduced by ~ 10 per cent, and the $\Delta\chi^2$ heatmaps appear structurally similar. The effect is larger for higher inclinations, and is reduced if the planet is closer to the major axis of the deprojected disc. For errors in disc geometry, we find that

(i) False fitting due to phase centre errors is most prominent in the inner disc. We see more significant false fitting for errors in ΔDec than ΔRA . This is likely due to the projected disc's minor axis being aligned with the Declination axis in this example, meaning that errors in ΔDec represent a larger fraction of the projected disc radius compared to errors in ΔRA , which are aligned with the major axis.

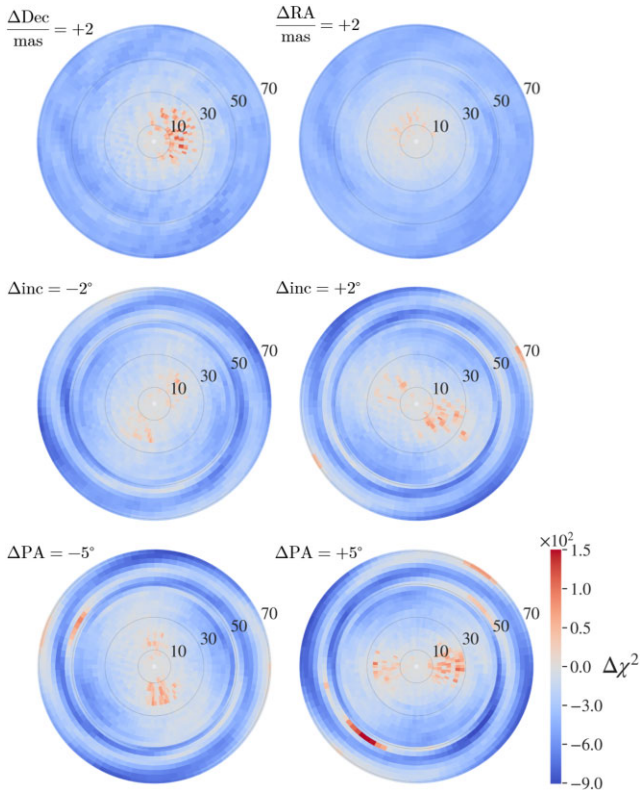


Figure A1. $\Delta\chi^2$ heatmaps for a ($\beta = 10$, $\tau_0 = 1$, $M_p = 0.3 M_{\text{th}}$) spiral template applied to a ($\beta = 10$, $\tau_0 = 1$) planet-less disc. The disc is fitted with deviations from its correct geometric parameters (in the text) noted in the upper left corner of each panel. The colourbar in the bottom right applies to all the heatmaps.

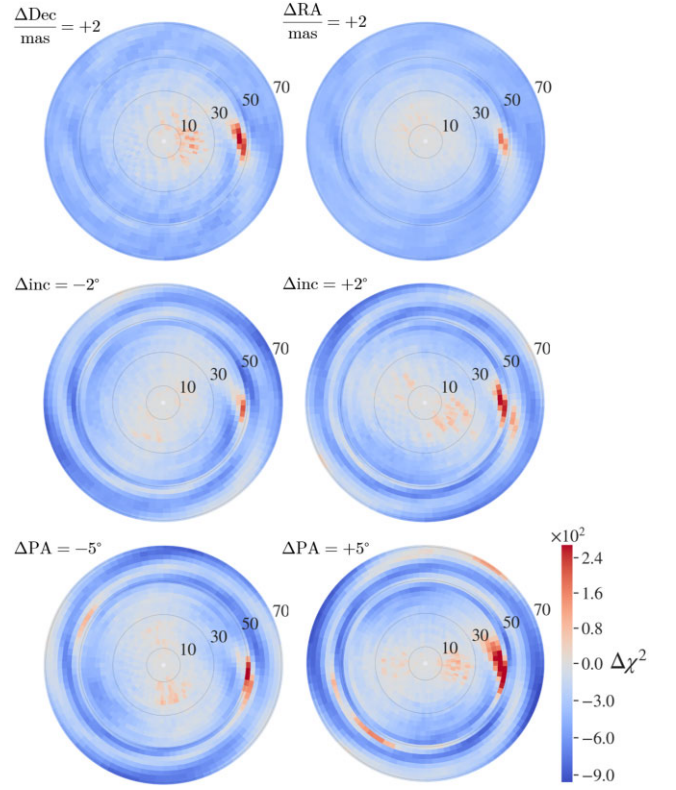


Figure A2. $\Delta\chi^2$ heatmaps as in Fig. A1, except the disc contains a $0.3 M_{\text{th}}$ planet.

(ii) False fitting due to errors in inclination and position angle also appear prominently in the inner disc, though can additionally appear as streaks in the rest of the disc. These streaks are perhaps the most problematic for spiral recovery as they somewhat resemble a spiral signal. However, the other characteristics of these errors (multiple streaks, two clusters of inner disc fits) should help distinguish them from a true spiral signal. We generally see more significant false fitting for errors in inclination than position angle.

(iii) Phase centre errors give one cluster of false fits on one side of the assumed disc centre, whereas orientation errors give two clusters on opposite sides of the disc centre, which can help distinguish between them.

(iv) For the given errors in disc geometry, the $0.3 M_{\text{th}}$ planet is still clearly recoverable by its spiral signature. They serve as a cautious (and usually readily achievable) upper limits for attempting spiral detection with our approach.

APPENDIX B: THE VISIBILITY-SPACE EFFECT OF ADDING A SPIRAL TO AN AXISYMMETRIC DISC

To explore how strong of an effect planet-driven spirals can have on visibilities, we try adding them to discs that are perfectly axisymmetric, but projected and with realistic radial structure. We obtain such discs from the visibilities of the axisymmetric disc model (A) and the spiral model instance (S) and quantify their difference using $\chi_{SA}^2 = \sum_k w_k |V_{P,k} - V_{Q,k}|^2$. χ_{SA}^2 is always a positive quantity, with its magnitude representing the strength of effect on the visibilities the spiral is having.

We calculate χ_{SA}^2 across four of the spiral template parameters, neglecting to vary ϕ_p or spiral handedness as we expect these to have little impact on the results due to the axisymmetry of the disc. To justify this neglect, we must consider that the disc orientation will affect the strength of expression of a spiral from a planet placed at different azimuthal positions in the disc. To demonstrate the magnitude of this effect, we include an example χ_{SA}^2 heatmap for RY Tau (Fig. B1), which is the most inclined disc in our Taurus sample.

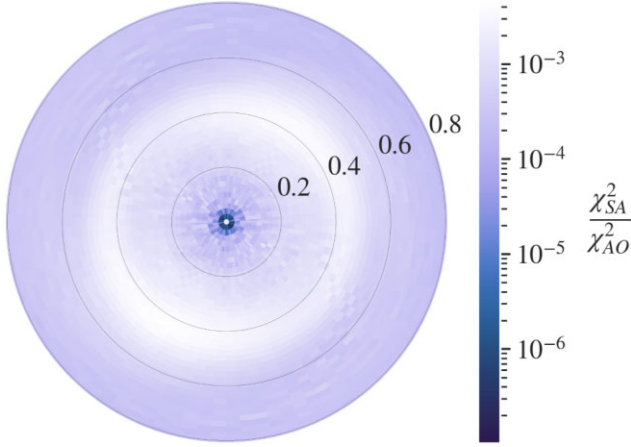


Figure B1. χ_{SA}^2 heatmap, in units of χ_{AO}^2 , of RY Tau (inclination = 65.0°) for a $\beta = 10$, $\tau_0 = 3$, $M_p = 1 M_{th}$, right-handed spiral template. The heatmap spatially represents the deprojected disc, with the disc's minor axis projected onto the horizontal axis ($\phi_p = 0$). Radial ticks are in arcseconds.

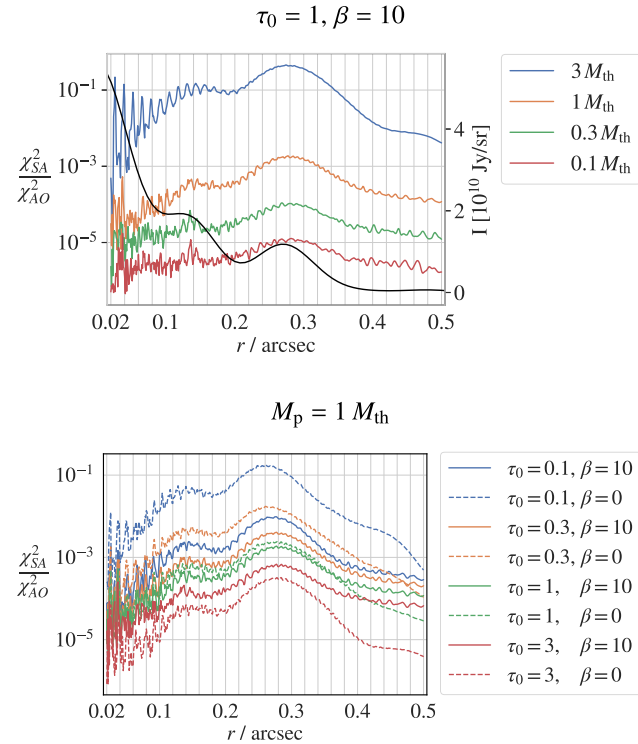


Figure B2. χ_{SA}^2 plots, in units of $\chi_{AO}^2 = 2.48 \times 10^5$, of FT Tau for right-handed, $\phi_p = 0$ spiral templates. Top: Variation with planet mass for adiabatic, $\tau_0 = 1$ templates. The FRANK brightness profile also is shown for comparison. Bottom: Variation with cooling time and optical depth for $M_p = 1 M_{th}$ templates.

Even in this highly inclined case, the effect is small, with the heatmap appearing very close to axisymmetric. Therefore, in order to facilitate comparison between spiral templates, we let all spiral templates take $\phi_p = 0$. We also let all spiral templates be right-handed as the effect of handedness should be negligible for the same reason. We can then compare the effects of the remaining spiral model parameters: orbital radius, planet mass, optical depth, and cooling time.

We explore this in Fig. B2 for FT Tau, a typical compact disc from our Taurus sample. We see significant radial variations in χ_{SA}^2 , with intermediate radii having the greatest effect across the discs in our Taurus sample. This is unsurprising as large r_p place the strongest parts of the spiral in faint reaches of the disc, and small r_p limit the strongest parts of the spiral to a small area of the disc. We also see peaks and troughs that coincide with radii of elevated and reduced brightness, respectively. This is to be expected because the spiral templates represent *fractional* residual brightness maps; thus, placing the planet in a brighter region will lead to a stronger effect. At low r_p , χ_{SA}^2 becomes increasingly noisy. This is likely due to the fact that the spiral is contributing on scales smaller than the angular resolution of the observations. This high χ^2 variability at inner disc radii also appears in many of our $\Delta\chi^2$ heatmaps.

Of the simulation parameters, planet mass has the largest effect, with higher mass planets unsurprisingly having greater effects on the disc visibilities. Optically thinner spiral templates also show larger magnitude effects, which is opposite to what is seen in the corresponding image residuals (fig. 6 of S22). The reason for this is that, in images, decreasing optical depth leads to a decrease in overall disc brightness, whereas, since our method incorporates the spiral as a fractional residual brightness perturbation, there is no change in overall disc brightness (the overall disc brightness is instead set by the axisymmetric FRANK model). Hence, only fractional change matters, and the fractional change is larger for optically thinner models. The radial structure of χ_{SA}^2 changes little between different optical depths, consistent with its modest impact on spiral morphology.

The isothermal templates tend to exhibit a slightly greater dependence on orbital radius than the (approximately) adiabatic ones. Additionally, the isothermal templates appear more sensitive to optical depth variations than the adiabatic ones, with the strongest spiral template being the $\tau_0 = 0.1$ isothermal one across most radii, and the weakest usually being the $\tau_0 = 3$ isothermal one. This is somewhat counterintuitive since the temperature contrast of the adiabatic templates generally results in a greater spiral brightness contrast than in the corresponding isothermal templates. However, the isothermal spiral simulations drive a deeper/wider gap in their discs and this manifests as a higher amplitude brightness perturbations around the planet's position in the spiral templates ($I_S \propto 1 / \langle I_{sim} \rangle_\phi$ [equation (1)], and for a deep gap $\langle I_{sim} \rangle_\phi$ is small). This appears to have a greater effect than the temperature perturbation of the adiabatic templates for the optically thin cases, although this reverses as optical depth increases.

APPENDIX C: SUPPLEMENTARY RADIAL PROFILES AND $\Delta\chi^2$ HEATMAPS

For our fiducial set of synthetic observations (0.061 arcsec resolution and $35 \mu\text{Jybeam}^{-1}$ sensitivity), we include here the full set of CLEAN image profiles (Fig. C1) used in determining the gap recovery grids, and $\Delta\chi^2$ heatmaps (Figs C2–C5) used in determining the spiral recovery grids (Fig. 9).

All the model discs contain right-handed spirals; therefore, for the right-handed spiral templates (Figs C2 and C4), a positive $\Delta\chi^2$ near $r_p = 50$ au, $\phi_p = 0$ is interpreted as a true signal for the planet.

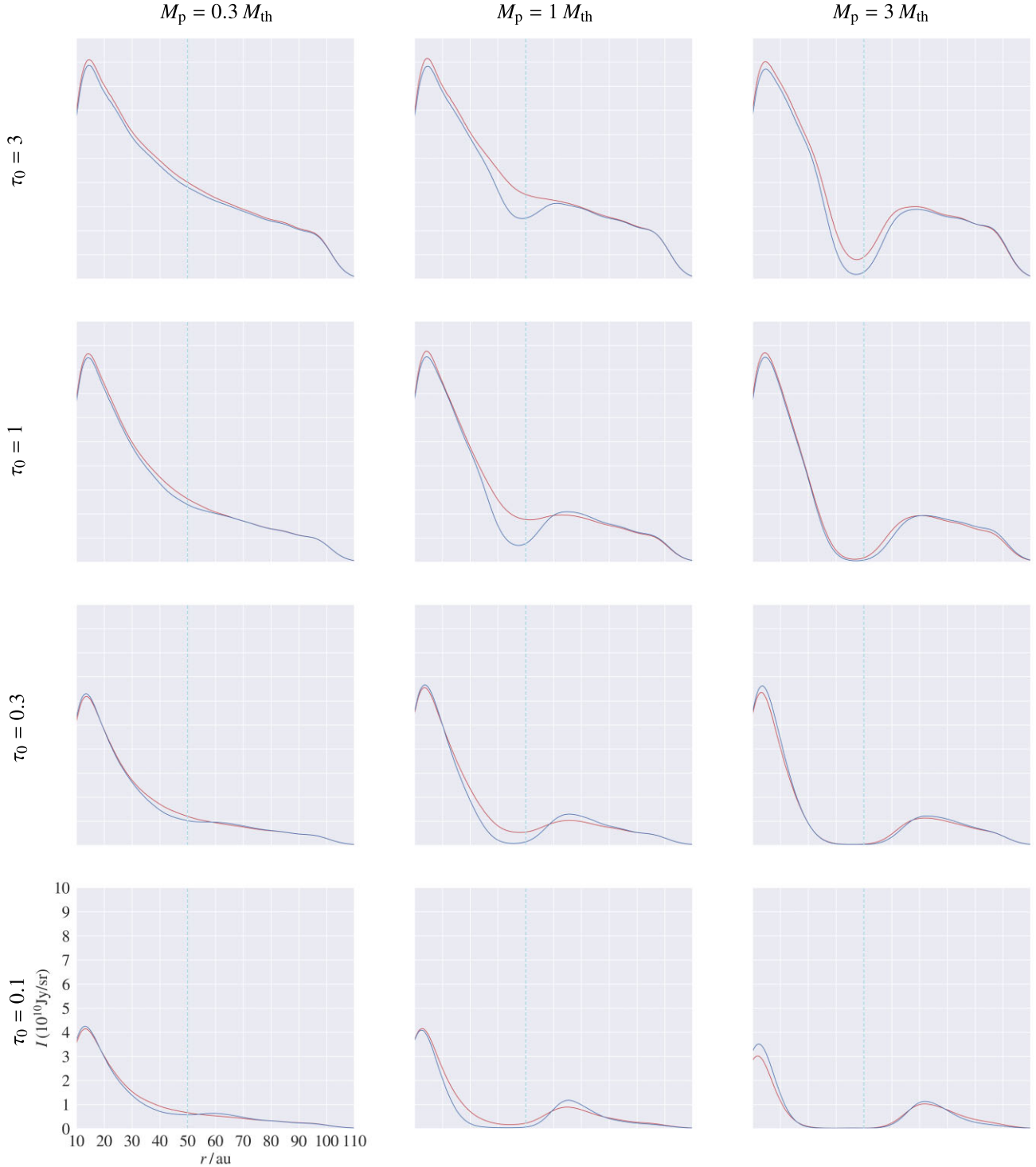


Figure C1. Radial profiles for adiabatic ($\beta = 10$) (red lines) and isothermal ($\beta = 0$) (blue lines) model discs. The axes shown in the bottom left profile apply to all the profiles. The planet radial position (50 au) is indicated by a blue dashed line. The gaps in the isothermal profiles are generally deeper than in the adiabatic profiles. $M_p/M_{\text{th}} = 0.3, 1, 3$ from left to right, and $\tau_0 = 0.1, 0.3, 1, 3$ from bottom to top. The discs are observed with our fiducial set-up (0.061 arcsec angular resolution and $35 \mu\text{Jy beam}^{-1}$ sensitivity, from ALMA configuration pair C-4 + C-7 with 40 min of on-source time).

By contrast, the left-handed templates (Figs C3 and C5) serve as handedness tests for spiral recovery, and low $\Delta\chi^2$ values (compared to the corresponding right-handed template) are interpreted as giving a clearer spiral recovery, particularly for regions near the peak $\Delta\chi^2$ in the right-handed templates.

A ‘perfect’ spiral signal is a ‘spot’ around the planet position showing a significantly higher $\Delta\chi^2$ than anywhere else. However, since it is possible to partially fit the true spiral with an incorrect spiral, we tolerate some degree of false fitting ($\Delta\text{BIC} > 0$) with the left-handed templates and incorrectly positioned right-

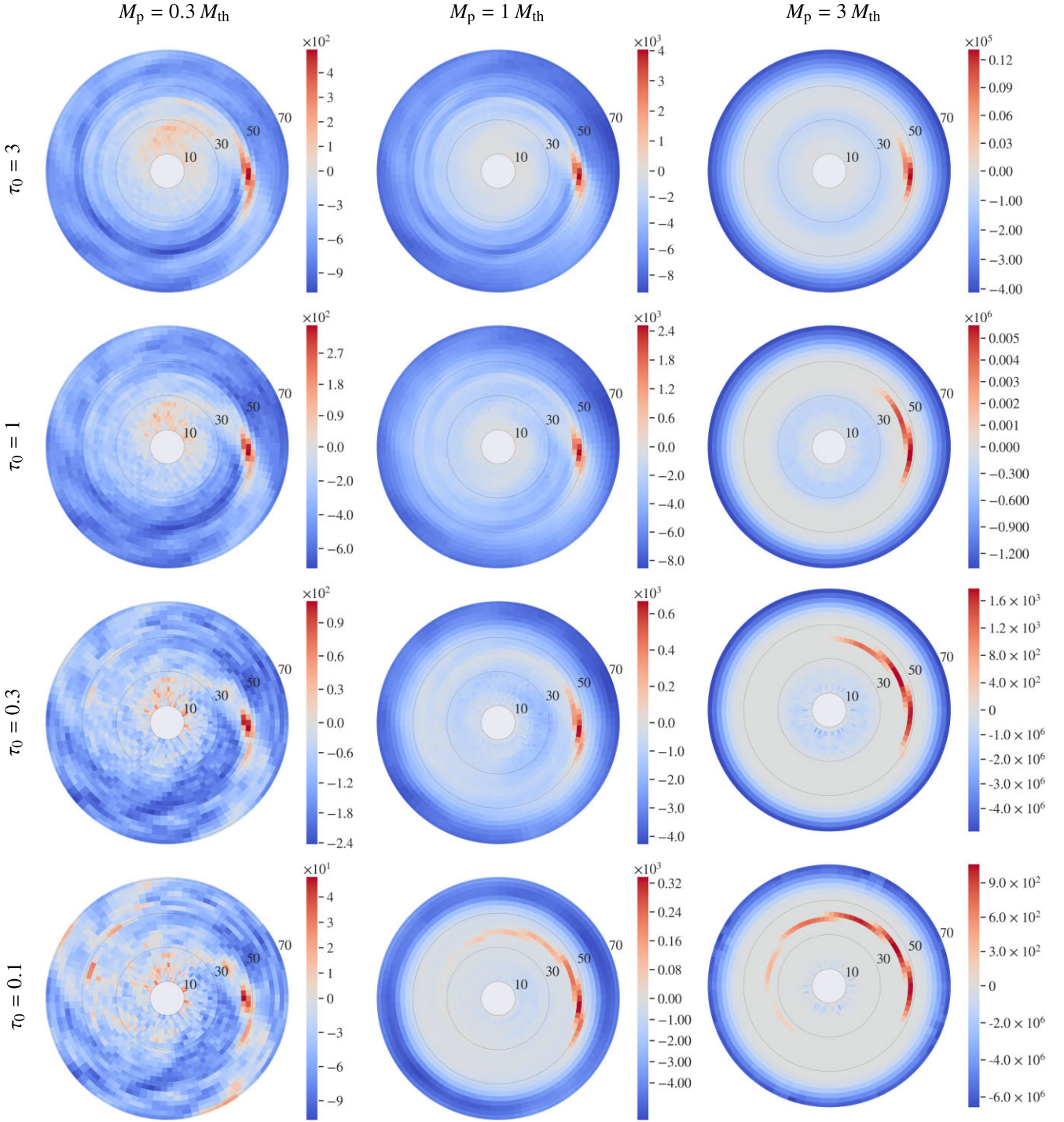


Figure C2. $\Delta\chi^2$ heatmaps of adiabatic ($\beta = 10$), right-handed model discs with their matching spiral templates. The discs host their planet at 50 au along the right horizontal axis, and are observed with our fiducial set-up (0.061 arcsec angular resolution and $35 \mu\text{Jybeam}^{-1}$ sensitivity). For these observations, $\Delta\chi^2(\Delta\text{BIC} = 0) = 63$, as determined by the number of visibility points. All the adiabatic discs, except the $\tau_0 = 0.1$, $M_p = M_{\text{th}}$ case, achieve a maximum $\Delta\chi^2$ significantly larger than this, indicating a very strong preference for the spiral model.

handed templates. We interpret this false fitting as not significantly impeding spiral recovery if its spatial structure is clearly not noise-like and not consistent with other asymmetries. For instance, some heatmaps have false fitting that traces out part of a spiral arm, such as for the optically thin, high planet mass templates towards the bottom right of Figs C2 and C4. Whereas, for the

$M_p = 0.3 M_{\text{th}}$ isothermal templates in the left column of Fig. C4, although there is clearly some spatial structure, it is not clearly differentiable from noise or non-spiral asymmetries (the $\tau_0 = 3$ case is interpreted as a marginal recovery, owing to the right-handed template's notably higher peak $\Delta\chi^2$ than the left-handed one).

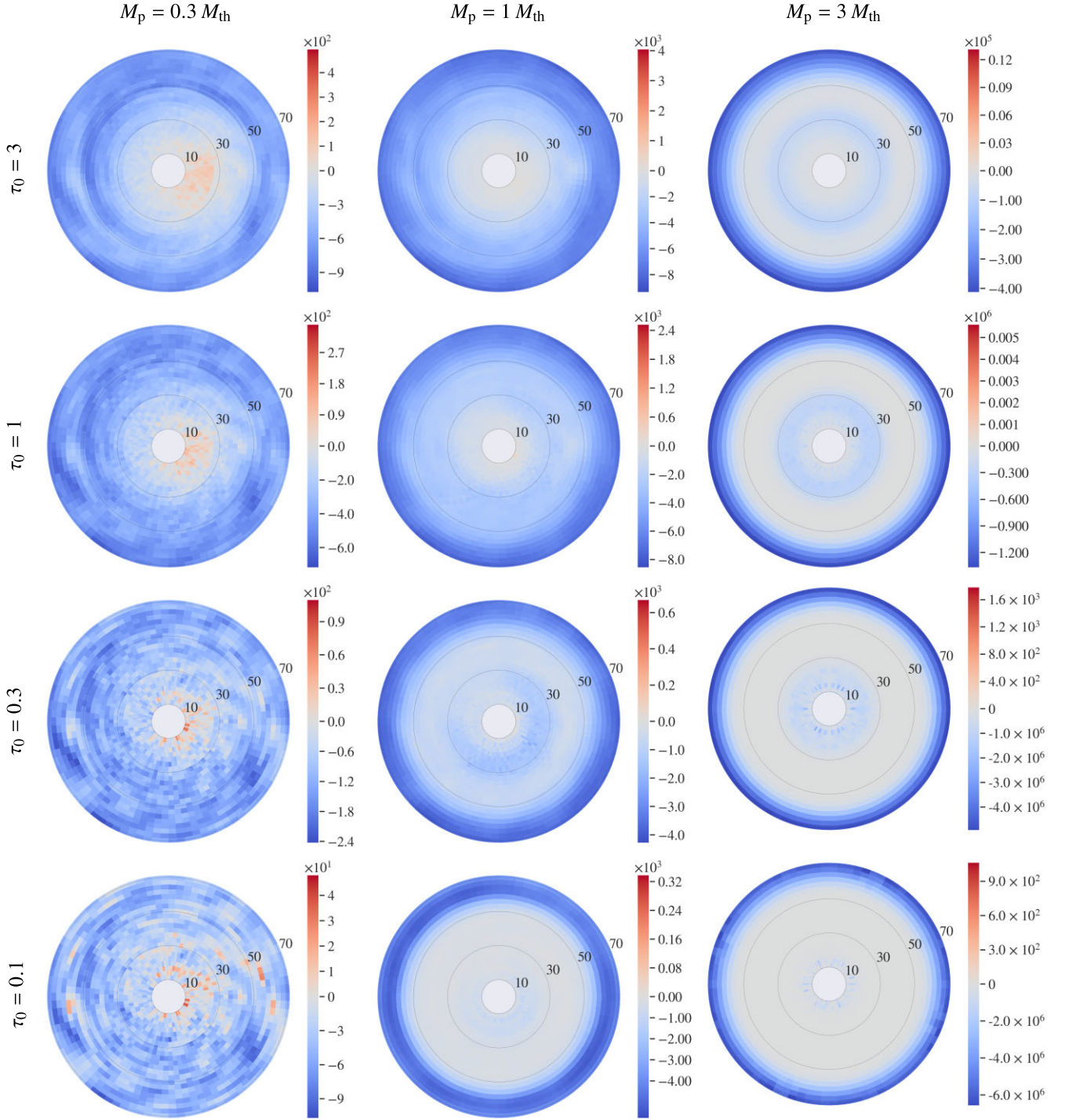


Figure C3. $\Delta\chi^2$ heatmaps of adiabatic ($\beta = 10$), right-handed model discs with their left-handed, but otherwise matching, spiral templates. Because these spirals are of the wrong handedness, low $\Delta\chi^2 > 0$ values (compared to the corresponding right-handed template in Fig. C2) are interpreted as giving a clearer spiral recovery. The colourbars for each heatmap are the same as for the corresponding right-handed template in Fig. C2 to aid comparison.

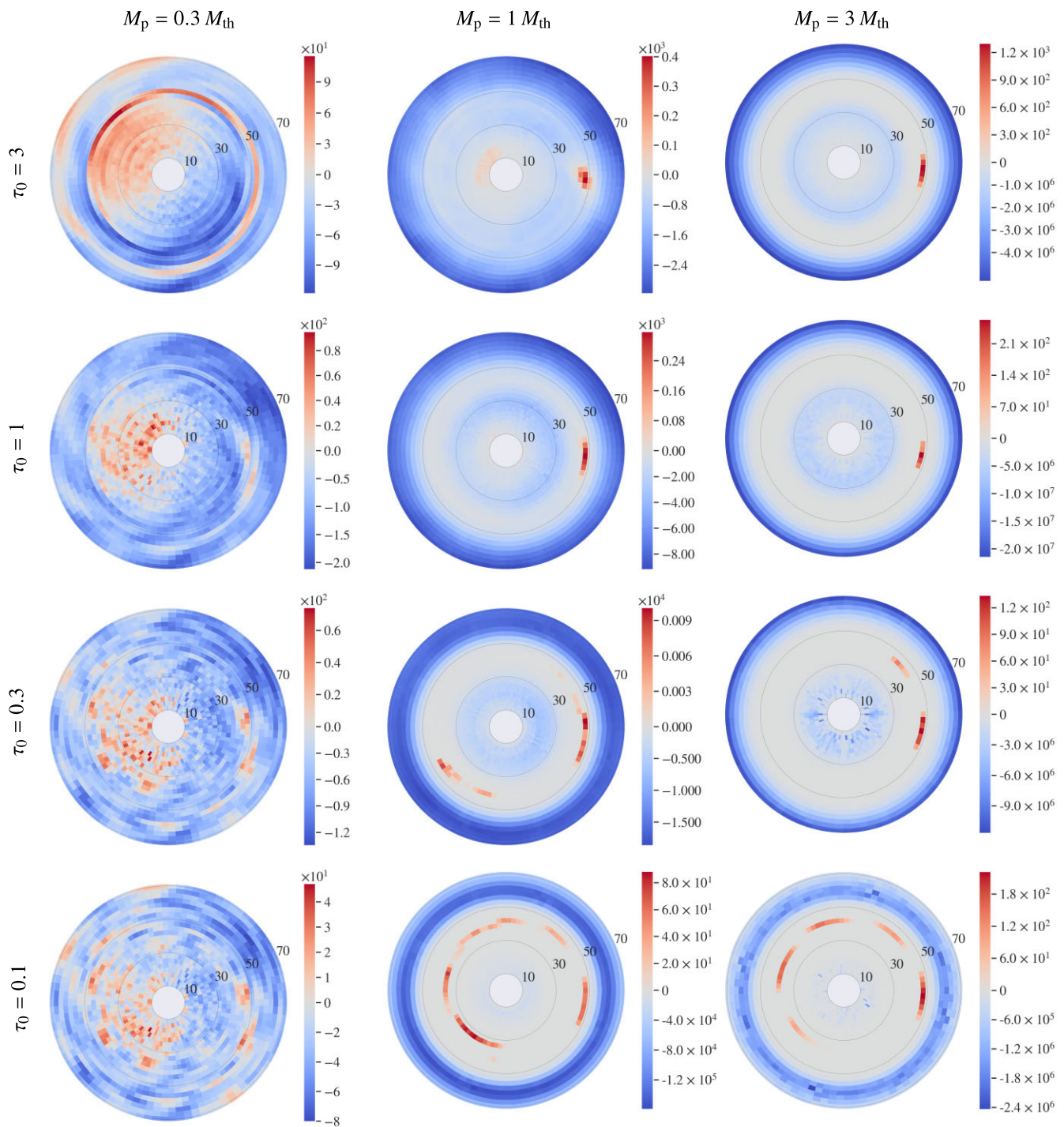


Figure C4. $\Delta\chi^2$ heatmaps of isothermal ($\beta = 0$), right-handed model discs with their matching spiral templates. Like in Fig. C2, the discs host their planet at 50 au along the right horizontal axis, and are observed with our fiducial set-up (0.061 arcsec angular resolution and $35 \mu\text{Jybeam}^{-1}$ sensitivity).

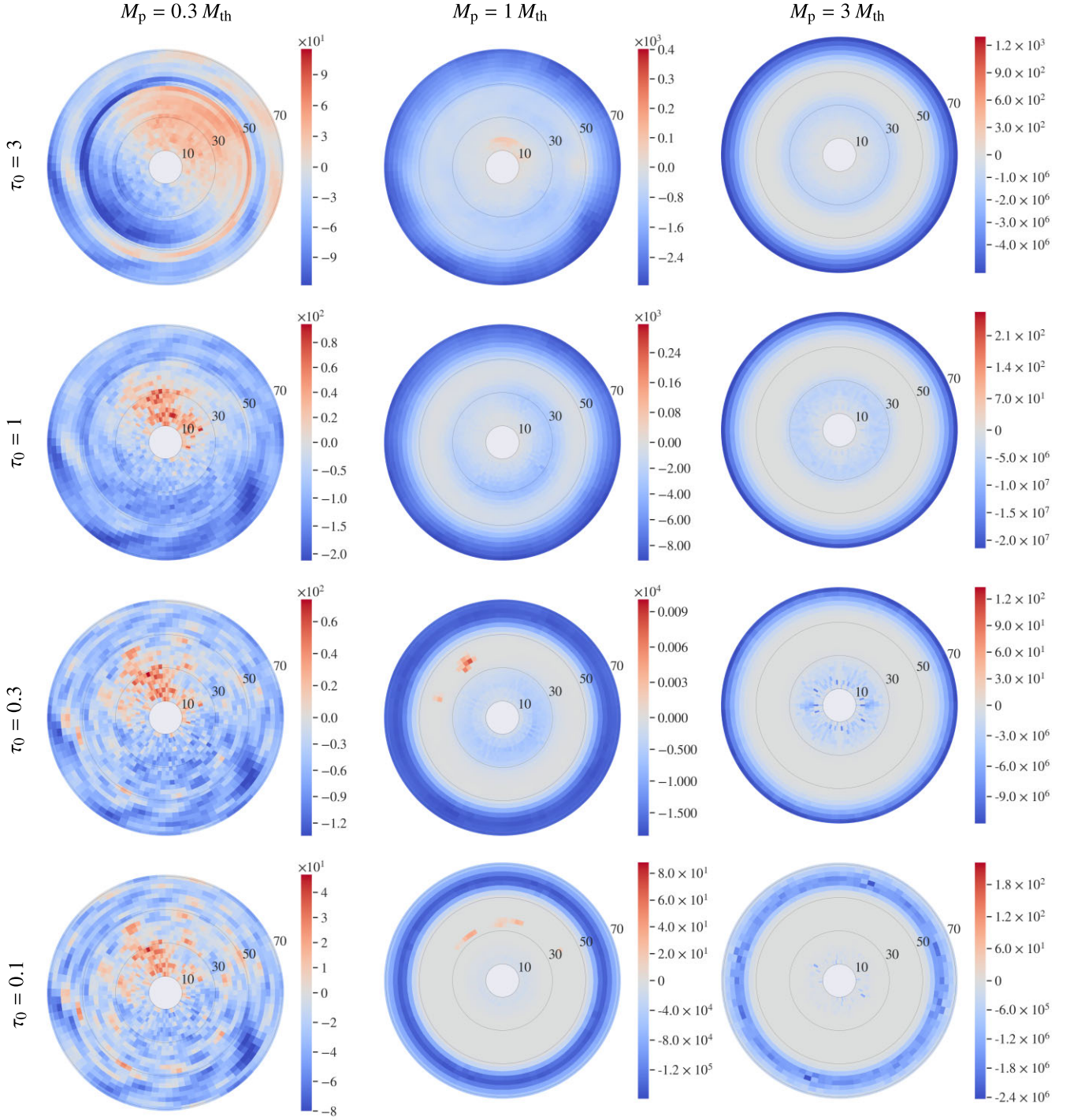


Figure C5. $\Delta\chi^2$ heatmaps of isothermal ($\beta = 0$), right-handed model discs with their left-handed, but otherwise matching, spiral templates. Because these spirals are of the wrong handedness, low $\Delta\chi^2 > 0$ values (compared to the corresponding right-handed template in Fig. C4) are interpreted as giving a clearer spiral recovery. The colourbars for each heatmap are the same as for the corresponding right-handed template in Fig. C4 to aid comparison.

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