nature research

https://doi.org/10.1038/s41586-020-2779-6

Supplementary information

Four annular structures in a protostellar disk less than 500,000 years old

In the format provided by the authors and unedited

Peer Review File

Manuscript Title: Four annular structures in a protostellar disk with an age <500,000 years

Reviewer Comments & Author Rebuttals

Reviewer Reports on the Initial Version:

Referees' comments:

Referee #1 (Remarks to the Author):

The authors present new, high spatial resolution observations with ALMA of the disk around the Class I protostar IRS 63, and find that the disk exhibits deviations from the expected smooth disk profile, concluding that the disk exhibits substructures. They consider this source compared with other particularly young sources with substructures and conclude that it is the youngest source to-date found to have such features. They analyze the structure of these features, discuss their possible origins, and ultimately conclude that regardless of the origin, we are observing features related to planet formation in the youngest source known to-date.

I believe the observations are quite nice, the existence of these features is interesting, that we are probably seeing some form of "substructures" in a young disk, and that the protostar would then probably be the youngest source published thus far to have substructures. However, I do have some concerns regarding their analysis that I believe should be addressed, and in particular I think in many cases the conclusions drawn are too strong. I detail those concerns below:

MAJOR CONCERNS:

- Page 1, Line 29: I think the quoted age of ~135,000 years is remarkably precise given the considerable uncertainties inherent in how these sorts of calculations are done, and is probably too strong of a statement. Personally, I think to be as conservative as possible this is probably best quoted as something like <500,000 years, which more or less covers the breadth of possibilities. Suggesting we know the age of this source to much better than this is probably an overstatement… Some of the other studies that have looked at this (e.g. Evans et al. 2009, Dunham et al. 2014) should probably also be cited.

- Page 2, lines 11-13 + Page 13, lines18-22: Was the centroid of the radiative transfer model fit explored? It seems to me like this asymmetry could be produced simply by not having the model properly centered with the data. Alternatively (or possibly also), the authors assume a relatively flared disk (beta=1.25), but previous studies have shown that HL Tau's dust disk is very flat (Pinte et al. 2016, I believe?). Could this asymmetry be removed with a flatter disk? I guess overall what I'm trying to get at is that interpreting this as a true asymmetry in the the dust emission given the relatively simple model seems a little strong.

- Page 2, Lines 42-43 and page 14, lines 23-24: Given how shallow and narrow these features are, these seem to be quite massive putative planets, if indeed there are planets at all. One possibility is the assumed model h/r, which appears to be quite high (~0.35 at r_0) vs. a maximum value of 0.1 in the Kanagawa et al. 2016 paper. This may tie in to the comment above that the dust disk at least may not be as flared as the model assumes. Alternatively, I seem to remember that Dong & Fung 2017 suggested that the gap width is actually relatively insensitive to planet mass, and that their results are consistent with Kanagawa et al. 2016 if you assume a different definition of gap width. I'm not certain how those definitions compare to what is used here, but I think this is all to say that there's quite a bit of uncertainty in these numbers and my impression is that the numbers quoted here would be fairly extreme upper limits. The authors do say that these are likely upper limits, but at the very least this could probably be emphasized more strongly to not give readers the wrong impression (especially considering the very small errors bars shown in the Methods section).

- Page 14, lines 2-3, lines 10-11, and lines 29-30: I strongly suspect that these measured values and uncertainties do not account for uncertainty in the radiative transfer model (i.e. the underlying smooth disk profile) that has been subtracted off, and so they are likely a good bit more uncertain than is estimated here. I would suggest the authors try to account for that more rigorously, though I understand that radiative transfer modeling is time consuming.

- Page 3, lines 16-18: This seems quite speculative to me considering the lack of gas observations, and is also heavily dependent on the assumed underlying smooth disk model. You need some amount of gas density enhancement (presumably, though I suppose it could be a temperature enhancement as well...) here to drive the pressure maximum that leads to the dust pile up. I don't know the numbers well, they might be mentioned in Pinilla et al. 2012 or other studies like that one and perhaps that would be something to mention for comparison, but could imagine that the 10-20% enhancement in dust is offset by a 10-20% enhancement in the gas, hence no increase in the dust-to-gas ratio.

NOT QUITE MAJOR CONCERNS:

- Throughout: I would suggest that the authors consider adjusting the references to "rings and gaps" though much of the text to "annular substructures", and referring to them as rings or gaps when appropriate - e.g. when discussing putative planets, it would be appropriate to discuss these in the context of "gaps". I don't believe there's been a community-wide definition of this yet, so this may be personal preference, but in my mind "rings and gaps" refer to local extrema in the surface density profile. DSHARP referred to similar features in their data as "plateau"-like, but still considered them "annular substructures", though they did note that there's a good chance they may be local extrema in the density profile (Huang et al. 2018). What sort of local extrema these features are, and indeed whether the features are local extrema, depends fairly significantly on the smooth disk model that is subtracted off (both in the case of the unsharp masking and the radiative transfer modeling), particularly given how shallow the features are. And until these features are observed at higher resolution or more optically thin wavelengths, it remains a possibility that this is indeed a change in the profile without an extrema. (I could potentially imagine a few other ways to try and tease this out in the current data as well, though - e.g. imaging with different visibility weights, or fitting some form of model in the visibility plane where there might be some baselines sensitive to smaller scale structures than what is seen in the image).

That said, I don't think the above suggestion has any bearing on the interpretation of these features as indications of early planet formation, as even in the perhaps most conservative case of a change in surface density power law index, this might e.g. indicate the beginnings of a pile up of dusty material at particular radii in the disk.

MINOR CONCERNS:

- Page 1, Line 43: The citations here should probably also include de Valon et al. 2020, which I believe was the first to show the substructures in DG Tau B, at least on the arXiv, if not in terms of which was published first. Though either way this should probably be included.

- Page 2, Line ~25: It may be worthwhile here to explicitly mention the typical ages (e.g. 1 - 10 Myr) for Class II disks for context.

- Page 2, Line 32: The citations here should probably also include Teague et al. 2018, though I admit I'm skeptical of whether that particular result definitively demonstrates the presence of planets, or whether it simply shows the presence of a gap through the kinematics.

- Page 11, line 25: Should "(Figure 2, right)" be "(Figure 2, left)"?

- Page 11, lines 25-26: Were additional values of the robust parameter considered, and/or superuniform weighting? If so, could those bring out the ring structure more, or if not, why not? Given the high S/N and the desire to look for small scale features, it seems at least worthwhile to check.

- Page 11, line 34: I'm not tremendously familiar with the details of unsharp masking, but in this case of a fairly elongated beam, would it make sense to double both major and minor axes? The 0.06" smoothing is just barely larger than the major axis of the beam so there is very little smoothing happening in that direction.

- Page 11, line 45: I think the "clumpy structures" referenced here are probably just noise, I'm not sure there's enough evidence to suggest otherwise…

- Page 12, lines 11-12: I don't necessarily disagree with WL 17 not being a direct analog of IRS 63… but it might be prudent to comment on the effectiveness of photoevaporation at dispersing the disk in sources this young, as my understanding is that that is a longer timescale process. On the other hand, there was a paper by Takahashi et al. 2018 that could explain this structure with a magnetized disk wind that might be prudent to mention here.

- Page 14, line 5: The 6 au feature looks fairly clearly like a model that isn't totally perfect to me, which is understandable considering how high S/N these data are. But I don't know as I'd go as far as to suggest this could be a ring…

Referee #2 (Remarks to the Author):

A. The authors present high resolution observations of the thermal dust continuum of a circumstellar disk around an embedded, young protostar. They find that, on top of a smooth distribution, the disk shows two low amplitude rings at 27au and 51au radii and discuss the implications for planet formation. With an estimated age of only 135kyr, the possibility that protoplanets have grown to Saturn mass levels at these large radii would provide a severe challenge to models of core accretion. However, as the authors note, even if planets are not the cause of the disk structure, the data suggest that planet formation has already begun during this very early phase as the disk itself is still growing.

B. These are exciting new data. The observatory used here, ALMA, has already shown rings and other substructures to be a common, arguably ubiquitous, feature in optically visible protostars with ages > 1 Myr. Here, we see the likely precursor of those features in an order of magnitude younger disk around an optically obscured, protostar surrounded by a dusty envelope. There are a handful of similar cases to date -- this is an area of active research -- and the authors compare their observations and results with the literature in the Methods section. They argue that this source, IRS63. is younger than other ringed disks based on its SED and disk-to-envelope mass ratio. This is reasonable and possibly as good as can be done but inherently imprecise. For instance the intrinsic protostellar temperature and envelope geometry also affect the SED, whereas disks are known to have a wide range of masses at any given age. Nevertheless, even though this is not necessarily unique or known with certainty to be the absolute youngest, the number of detailed observations and analysis of disks at these ~0.1Myr timescales is very small so this work is indeed significant.

C/D. The data are of high quality given the challenge of observing a deeply embedded, relatively low mass disk through its surrounding envelope at ~5au resolution. The analysis is good and both clearly described and presented. I make some suggestions to augment this, including the treatment of uncertainties, below.

E. The conclusions regarding the location of the rings and gaps are robust. The interpretation regarding planet masses and the nature of the rings are necessarily less so because these data alone cannot distinguish between dust surface density and opacity. I agree with the authors' general premise that planet formation has probably already advanced significantly at this early age. However, this is not ironclad as snowlines themselves might be able to account for the observed features and cannot be categorically ruled out from the data at least as presented here.

F. I pose my suggestions for improvements through a series of questions ordered by location in the manuscript:
- Note the challenge that a 0.5 M_Jupiter = M_Saturn planet at >20au creates for core-accretion models of planet formation, both in terms of building up such a mass so quickly but then also stopping it form undergoing exponentially fast runaway growth.
- Plot a temperature profile in the methods section and state what assumptions are used for the stellar properties here. Note that young stars can undergo sporadic luminosity bursts that can push snowlines out with longer recovery timescales than that of the star. The subtle features here remind me of the V883 Ori disk (Cieza et al. 2016) where the change in intensity was attributed to the water snowline based on a change in opacity. Consider looking at the intra-band spectral index to see if you find a similar effect.
- Explain the "back reaction" of dust on the gas since that theory states that this is only important for dust/gas ratios near unity which is two orders of magnitude greater than the values used/derived here.
- Figure 2 showing the difference between the observations and model is azimuthally asymmetric and suggests that the model is slightly offset. How was the center determined? If the asymmetry is real, this would seem to be a profound result that should be commented on.
- The radial features (rings and gaps) are extremely subtle, i.e., very broad and low amplitude. Discuss how they compare with the DSHARP features. Compare with the temperature profile to determine the opacity variation. If the rings reach tau > 0.6, then scattering may be important and the surface density variations

could be much higher than the often-assumed purely absorption based radiative transfer modeling used here.
- The Brinch & Jorgensen (2013) paper cited here for the stellar mass shows that this is actually quite poorly determined. How might a 50% change in M_star affect the temperature and consequent modeling of the results? Is there any evidence or constraints on accretion? Again, what effect might that have on the T, tau profile? How is the 135kyr age used in the abstract determined?
- The modeling used a grid-based approach which is reasonable given the computational complexity of modeling radiative transfer. However, this makes it hard to determine errors accurately. Discuss the error budget both from input parameters and fitting uncertainty.
- If the change in intensity in the rings were not due to a difference in dust-to-gas ratio but a change in dust opacity, what would be the effect (e.g., composition, a_max) and interpretation? Could this be explained as snowlines or dust pileup and agglomeration?
- Note that the moderate dust mass here shows that this young disk is already well into the (globally) gravitationally stable regime.

G. The references are fine.

H. The paper was well written: clear and lucid.

Author Rebuttals to Initial Comments:

NOTE: Author responses in RED

Referee #1 (Remarks to the Author):

The authors present new, high spatial resolution observations with ALMA of the disk around the Class I protostar IRS 63, and find that the disk exhibits deviations from the expected smooth disk profile, concluding that the disk exhibits substructures. They consider this source compared with other particularly young sources with substructures and conclude that it is the youngest source to-date found to have such features. They analyze the structure of these features, discuss their possible origins, and ultimately conclude that regardless of the origin, we are observing features related to planet formation in the youngest source known to-date.

I believe the observations are quite nice, the existence of these features is interesting, that we are probably seeing some form of "substructures" in a young disk, and that the protostar would then probably be the youngest source published thus far to have substructures. However, I do have some concerns regarding their analysis that I believe should be addressed, and in particular I think in many cases the conclusions drawn are too strong. I detail those concerns below:

MAJOR CONCERNS:

- Page 1, Line 29: I think the quoted age of ~135,000 years is remarkably precise given the considerable uncertainties inherent in how these sorts of calculations are done, and is probably too strong of a statement. Personally, I think to be as conservative as possible this is probably best quoted as something like <500,000 years, which more or less covers the breadth of possibilities. Suggesting we know the age of this source to much better than this is probably an overstatement… Some of the other studies that have looked at this (e.g. Evans et al. 2009, Dunham et al. 2014) should probably also be cited.

Response:
We agree with the referee that the most conservative age estimate is likely <500,000 years. We have changed the reference to ~135,000 years in the main text to the most conservative <500,000 years. We have also added a larger discussion of the ages in Methods, since an age of 130,000 to 140,000 years is indicated from the bolometric temperature based on the modeling of a 1 M_{sol} protostar through time from the onset of collapse (Young & Evans 2005). That study indicated that a lower mass protostar (which could be the case here, given that the 1 L_{sol} luminosity used in the model is most certainly partially from accretion) would imply a younger age for the same T_{bol}. 135,000 years is also the half-life of the Class I phase (Kristensen et al. 2018), indicating that IRS 63 is likely a quite typical Class I source in terms of age. We have also added a citation to Evans et al. 2009, as suggested.

- Page 2, lines 11-13 + Page 13, lines18-22: Was the centroid of the radiative transfer model fit explored? It seems to me like this asymmetry could be produced simply by not having the model properly centered with the data. Alternatively (or possibly also), the authors assume a relatively flared disk (beta=1.25), but previous studies have shown that HL Tau's dust disk is very flat (Pinte et al. 2016, I believe?). Could this asymmetry be removed with a flatter disk? I guess overall what I'm trying to get at is that interpreting this as a true asymmetry in the the dust emission given the relatively simple model seems a little strong.

Response:
We have changed the text in Methods to clarify some points of the radiative transfer modeling. We thank the referee for the questions, and we feel the changed text discusses the presence of the asymmetry better. We explored values of beta from 1.0 to 1.5, while fixing beta - q = 2.25. Pinte et al. (2016) found beta = 1.15 for HL Tau, which is covered in our parameter space. We point out that asymmetries purely from flaring angle geometry should occur along the minor axis of the disk, which is not the case for IRS 63. We also clarify in the text that the center of the model is fixed to the Gaussian-measured peak of the central protostar from the data. Early tests with the centering of the data revealed to us that the asymmetries were always present and moving from the data-fit center increased the residuals in the central region. The asymmetries are also not model-dependent since they are subtly detectable in the raw data. Taking radial-intensity cuts through the data (before deprojecting or azimuthal averaging) also showed excess emission in the north-west side of the disk compared to the south-west side. The asymmetry is simply more obvious when subtracting the smooth disk model due to the dynamic range of the data.

- Page 2, Lines 42-43 and page 14, lines 23-24: Given how shallow and narrow these features are, these seem to be quite massive putative planets, if indeed there are planets at all. One possibility is the assumed model h/r, which appears to be quite high (~0.35 at r_0) vs. a maximum value of 0.1 in the Kanagawa et al. 2016 paper. This may tie in to the comment above that the dust disk at least may not be as flared as the model assumes. Alternatively, I seem to remember that Dong & Fung 2017 suggested that the gap width is actually relatively insensitive to planet mass, and that their results are consistent with Kanagawa et al. 2016 if you assume a different definition of gap width. I'm not certain how those definitions compare to what is used here, but I think this is all to say that there's quite a bit of uncertainty in these numbers and my impression is that the numbers quoted here would be fairly extreme upper limits. The authors do say that these are likely upper limits, but at the very least this could probably be emphasized more strongly to not give readers the wrong impression (especially considering the very small errors bars shown in the Methods section).

Response:
Thank you. We agree that the values given are extreme upper limits. It is true that the disks modeled in Kanagawa et al. 2016 do not cover exactly the parameter space as indicated for the IRS 63 disk and the exact method of determining gap width can also vary the planet mass. As the referees suggest, we emphasize more in the text how extreme the upper limits are, and we point out in Methods that the uncertainties are likely larger than the quoted values due to the difference in gap measurement and the disk flaring.

- Page 14, lines 2-3, lines 10-11, and lines 29-30: I strongly suspect that these measured values and uncertainties do not account for uncertainty in the radiative transfer model (i.e. the underlying smooth disk profile) that has been subtracted off, and so they are likely a good bit more uncertain than is estimated here. I would suggest the authors try to account for that more rigorously, though I understand that radiative transfer modeling is time consuming.

Response:
The referee is right that these uncertainties do not reflect the inherent uncertainties in the model, which are not easily quantified. We have added text in this section to clearly point that out, and to also emphasize that the annular substructure locations and widths are not dependent on the smooth disk model (see also discussion below).

- Page 3, lines 16-18: This seems quite speculative to me considering the lack of gas observations, and is also heavily dependent on the assumed underlying smooth disk model. You need some amount of gas density enhancement (presumably, though I suppose it could be a temperature enhancement as well...) here to drive the pressure maximum that leads to the dust pile up. I don't know the numbers well, they might be mentioned in Pinilla et al. 2012 or other studies like that one and perhaps that would be something to mention for comparison, but could imagine that the 10-20% enhancement in dust is offset by a 10-20% enhancement in the gas, hence no increase in the dust-to-gas ratio.

Response:
The referee makes a good point. We had already in the Methods of the previous version re-framed the variations in intensity as variations in the dust opacity coefficients, assuming a constant dust to gas ratio in the disk. We added wording in the main text to point out that an alternative explanation is already given in the Methods section.

NOT QUITE MAJOR CONCERNS:

- Throughout: I would suggest that the authors consider adjusting the references to “rings and gaps” though much of the text to “annular substructures”, and referring to them as rings or gaps when appropriate - e.g. when discussing putative planets, it would be appropriate to discuss these in the context of “gaps”. I don’t believe there’s been a community-wide definition of this yet, so this may be personal preference, but in my mind “rings and gaps” refer to local extrema in the surface density profile. DSHARP referred to similar features in their data as “plateau”-like, but still considered them “annular substructures”, though they did note that there’s a good chance they may be local extrema in the density profile (Huang et al. 2018). What sort of local extrema these features are, and indeed whether the features are local extrema, depends fairly significantly on the smooth disk model that is subtracted off (both in the case of the unsharp masking and the radiative transfer modeling), particularly given how shallow the features are. And until these features are observed at higher resolution or more optically thin wavelengths, it remains a possibility that this is indeed a change in the profile without an extrema. (I could potentially imagine a few other ways to try and tease this out in the current data as well, though - e.g. imaging with different visibility weights, or fitting some form of model in the visibility plane where there might be some baselines sensitive to smaller scale structures than what is seen in the image).

Response:
As the referee suggested, we have changed most of the mentions of "rings and gaps" to "annular substructures", except where the interpretation required either "rings" or "gaps". It is true that the extrema here are very shallow and the amplitude of the extrema for any given annular structure (R1, R2, G1, G2) may change depending on the smooth model being subtracted off. The locations, Gaussian widths, and (as the referee rightly pointed out) regions of change in the intrinsic data radial profile are not affected by the exact smooth disk model subtracted off, and the phrase "annular substructure" is certainly applicable to all 4 features in the disk. Please see our answer below for our comments on visibility weights and working with the visibilities directly.

- That said, I don’t think the above suggestion has any bearing on the interpretation of these features as indications of early planet formation, as even in the perhaps most conservative case of a change in surface density power law index, this might e.g. indicate the beginnings of a pile up of dusty material at particular radii in the disk.

Response:
We agree. To make this point more clear, we added a sentence in the final paragraph of the main text. Thank you.

MINOR CONCERNS:

- Page 1, Line 43: The citations here should probably also include de Valon et al. 2020, which I believe was the first to show the substructures in DG Tau B, at least on the arXiv, if not in terms of which was published first. Though either way this should probably be included.

Response:
Thank you. As suggested, we have added the reference.

- Page 2, Line ~25: It may be worthwhile here to explicitly mention the typical ages (e.g. 1 - 10 Myr) for Class II disks for context.

Response:
We have noted the typical age of the known Class II disks with annular substructures (~1 Myr).

- Page 2, Line 32: The citations here should probably also include Teague et al. 2018, though I admit I'm skeptical of whether that particular result definitively demonstrates the presence of planets, or whether it simply shows the presence of a gap through the kinematics.

Response:
We agree with the referee that Teague et al. 2018 is important work; however, where Pinte et al. 2020 shows both radial and azimuthal indications of possible planets, Teague et al. 2018 shows only radial indications. Due to the limitation on citation count, we choose to cite only Pinte et al. 2020 with the more-suggestive evidence of planets.

- Page 11, line 25: Should "(Figure 2, right)" be "(Figure 2, left)"?

Response:
Yes, thank you. We have corrected the typographical error.

- Page 11, lines 25-26: Were additional values of the robust parameter considered, and/or superuniform weighting? If so, could those bring out the ring structure more, or if not, why not? Given the high S/N and the desire to look for small scale features, it seems at least worthwhile to check.

Response:
The referee is right that these parameters are worth checking. We did experiment early in the analysis with different visibility weightings for imaging, including uniform and super uniform weighting. While the overall signal to noise of the data is high, because these extrema are shallow/low contrast compared to the disk emission, the sacrifice in noise was too expensive in the trade off for better resolution. While the substructures were still present, the uncertainties inherent in the data became closer to the order of the amplitude of the extrema once the smooth disk model was subtracted off. Ultimately we chose a weighting of robust = 1.0 as a compromise between spatial resolution and noise level in the final image.

We also tried to cut out the lower-resolution (short-baseline) uv-data in an attempt to filter out the larger-scale disk emission even before imaging to highlight the rings, experimenting with multiple uv-cuts. Once enough short baselines were cut to more clearly reveal the annular substructures, the so-called "basket-weave" artifact (a checker-board pattern associated with missing flux on larger scales) appeared and significantly increased the noise of the substructures. This also introduced a negative bowl of emission around the disk which could naively be interpreted as another annular substructure to a reader unfamiliar with working with visibilities. While the annular substructures were again still present with this method, the uncertainties in the data increased enough to make analysis with this method much less straightforward, especially to a reader unfamiliar with interferometric techniques.

- Page 11, line 34: I'm not tremendously familiar with the details of unsharp masking, but in this case of a fairly elongated beam, would it make sense to double both major and minor axes? The 0.06" smoothing is just barely larger than the major axis of the beam so there is very little smoothing happening in that direction.

Response:
Because the dark annular structures (G1 and G2) are already similar to the size of the beam, we needed to smooth with the minimum smoothing kernel possible to avoid introducing azimuthal asymmetries in these structures. For instance, using an elliptical smoothing kernel, because of the orientation of the long-axis of beam and the minor axis of the disk (parallel to each other), will smooth over a substantially larger annular width along the minor axis compared to the major axis.

By keeping the smoothing kernel circular and small, on the size scale of the smallest observable substructures, this naturally minimizes such possible artifacts. The technique was used to only visually highlight the annular substructures independent of an underlying radiative transfer model, and all analysis was done with either unsmoothed data or the residuals from the radiative transfer modeling.

- Page 11, line 45: I think the “clumpy structures” referenced here are probably just noise, I’m not sure there’s enough evidence to suggest otherwise…

Response:
We disagree on this small point. GWYL 91 was also observed at lower resolution at the same wavelength in the ODISEA survey (Cieza et al. 2019, figure 3, target 047). The lower resolution Cieza et al. (2019) data also show azimuthal asymmetries/clumpy structures with a similar distribution as those seen in the higher resolution Sheehan et al. (2018) data. The other well-resolved disks in the Cieza et al. (2019) data do not show such a clumpy morphology. Therefore, we do think the clumpy structures presented in Sheehan et al. (2018) are tracing real structures and not noise.

- Page 12, lines 11-12: I don't necessarily disagree with WL 17 not being a direct analog of IRS 63… but it might be prudent to comment on the effectiveness of photoevaporation at dispersing the disk in sources this young, as my understanding is that that is a longer timescale process. On the other hand, there was a paper by Takahashi et al. 2018 that could explain this structure with a magnetized disk wind that might be prudent to mention here.

Response:
Thank you for pointing out the Takahashi et al. (2018) reference. The referee is right that typically photoevaporations happens on generally longer time scales than the length of the Class I phase (see also the discussion of WL 17 in Sheehan et al. 2018). Unfortunately due to length constraints on the text, some of the discussion regarding this had to be cut to make space for other analysis in the Methods.

- Page 14, line 5: The 6 au feature looks fairly clearly like a model that isn’t totally perfect to me, which is understandable considering how high S/N these data are. But I don’t know as I’d go as far as to suggest this could be a ring…

Response:
We agree. To avoid confusion for readers, we have removed discussion of the 6 au feature since we do not consider it in any analysis.

Referee #2 (Remarks to the Author):

A. The authors present high resolution observations of the thermal dust continuum of a circumstellar disk around an embedded, young protostar. They find that, on top of a smooth distribution, the disk shows two low amplitude rings at 27au and 51au radii and discuss the implications for planet formation. With an estimated age of only 135kyr, the possibility that protoplanets have grown to Saturn mass levels at these large radii would provide a severe challenge to models of core accretion. However, as the authors note, even if planets are not the cause of the disk structure, the data suggest that planet formation has already begun during this very early phase as the disk itself is still growing.

B. These are exciting new data. The observatory used here, ALMA, has already shown rings and other substructures to be a common, arguably ubiquitous, feature in optically visible protostars with ages > 1 Myr. Here, we see the likely precursor of those features in an order of magnitude younger disk around an optically obscured, protostar surrounded by a dusty envelope. There are a handful of similar cases to date -- this is an area of active research -- and the authors compare their observations and results with the literature in the Methods section. They argue that this source, IRS63. is younger than other ringed disks based on its SED and disk-to-envelope mass ratio. This is reasonable and possibly as good as can be done but inherently imprecise. For instance the intrinsic protostellar temperature and envelope geometry also affect the SED, whereas disks are known to have a wide range of masses at any given age. Nevertheless, even though this is not necessarily unique or known with certainty to be the absolute youngest, the number of detailed observations and analysis of disks at these ~0.1Myr timescales is very small so this work is indeed significant.

C/D. The data are of high quality given the challenge of observing a deeply embedded, relatively low mass disk through its surrounding envelope at ~5au resolution. The analysis is good and both clearly described and presented. I make some suggestions to augment this, including the treatment of uncertainties, below.

E. The conclusions regarding the location of the rings and gaps are robust. The interpretation regarding planet masses and the nature of the rings are necessarily less so because these data alone cannot distinguish between dust surface density and opacity. I agree with the authors' general premise that planet formation has probably already advanced significantly at this early age. However, this is not ironclad as snowlines themselves might be able to account for the observed features and cannot be categorically ruled out from the data at least as presented here.

F. I pose my suggestions for improvements through a series of questions ordered by location in the manuscript:

- Note the challenge that a 0.5 M_Jupiter = M_Saturn planet at >20au creates for core-accretion models of planet formation, both in terms of building up such a mass so quickly but then also stopping it form undergoing exponentially fast runaway growth.

Response:
Thank you. We have added text to point out these challenges.

- Plot a temperature profile in the methods section and state what assumptions are used for the stellar properties here. Note that young stars can undergo sporadic luminosity bursts that can push snowlines out with longer recovery timescales than that of the star. The subtle features here remind me of the V883 Ori disk (Cieza et al. 2016) where the change in intensity was attributed to the water snowline based on a change in opacity. Consider looking at the intra-band spectral index to see if you find a similar effect.

Response:
We have plotted a temperature profile, as the referee suggested (Supplemental Figure 1). Given the concerns both referees had about the model uncertainties, we have removed the model temperature from our analysis in relation to snowlines. Instead, we followed the method of Cieza et al. (2016) to split the data into two frequencies, and used those to observationally derive the spectral index, optical depth, and temperature profiles. Please see the snowline section of Methods for more details. From this more reliable observed temperature profile, we do see an increase in the spectral index and sharp decrease in optical depth at the G2 annular substructure (coincident with the CO snowline). This is comparable to the effect Cieza et al. (2016) saw for the change in intensity at the

water snowline. Thank you for the suggestion. We feel this greatly improved the analysis of our work.

- Explain the "back reaction" of dust on the gas since that theory states that this is only important for dust/gas ratios near unity which is two orders of magnitude greater than the values used/ derived here.

Response:
We point out that in the simulations of Gonzalez et al. (2017), self induced dust traps can form on 100,000 year time scales even with an initial dust to gas ratio of 0.01. Thus, there remains the possibility that the rings of IRS 63 may arise from self-induced dust traps in the process of formation. As suggested, we have added text to further explain these points.

- Figure 2 showing the difference between the observations and model is azimuthally asymmetric and suggests that the model is slightly offset. How was the center determined? If the asymmetry is real, this would seem to be a profound result that should be commented on.

Response:
Please see the discussion above in response to a similar question about the asymmetry from the first referee. At the referee's suggestion, we add a comment in the main text that this asymmetry could be one or more young vortices forming, which are possible to occur on timescales less than the Class I lifetime (Bae et al. 2015) and could further assist with piling up dusty material required for planetesimal formation.

- The radial features (rings and gaps) are extremely subtle, i.e., very broad and low amplitude. Discuss how they compare with the DSHARP features. Compare with the temperature profile to determine the opacity variation. If the rings reach tau > 0.6, then scattering may be important and the surface density variations could be much higher than the often-assumed purely absorption based radiative transfer modeling used here.

Response:
Thank you for the suggestions. We have added a sentence in the main text directly comparing the narrow and high-contrast rings in older sources to our broad and low-contrast substructures. Please see the answer above for an explanation of the derivation of the opacity variation with radius. Indeed, we find that tau > 1 in most of the disk and self-scattering may play a factor in masking emission in the more optically thick ring regions (Zhu et al. 2019). Self-scattering can also explain the values of alpha < 2 at the disk center (Liu 2019). We comment on this in the Methods section.

- The Brinch & Jorgensen (2013) paper cited here for the stellar mass shows that this is actually quite poorly determined. How might a 50% change in M_star affect the temperature and consequent modeling of the results? Is there any evidence or constraints on accretion? Again, what effect might that have on the T, tau profile? How is the 135kyr age used in the abstract determined?

Response:
We agree with the referee that the protostellar mass from Brinch & Jorgensen (2013) is relatively uncertain; however it is currently the best estimate of central protostellar mass available. We use the M_{star} value of 0.8 M_{Sun} only for the planet mass-width determination. We make a note in the Methods about the planet mass determination that this could be a key contributor to realistic uncertainties.

The model is based on a provided central point source luminosity (see Schmiedeke et al. 2016 for technical setup details), which in principle accounts for both intrinsic luminosity and any accretion luminosity from the disk onto the protostar. The 1.0 L_{Sun} used in the modeling is from Brinch & Jorgensen (2013) and hence implies an M_{star} of 1.0 M_{Sun} or less. We had also experimented with a family of models with 1.5 L_{Sun}, which would imply either a larger mass or a higher accretion rate. While the optimal disk geometry such as flaring power and surface density radial exponent remained similar between the two luminosity families, all profiles for the 1.5 L_{Sun} family of models remained far too centrally peaked (even with far lower central density values) and with residuals much higher than for the 1.0 L_{Sun} case.

The effect of the larger luminosity value on the T and tau profile is unclear because other parameters like central density must also be adjusted to better match the data and exact values for these are highly uncertain. We have also changed the snowline portion of the analysis and in Methods to reflect the T and tau profiles based on the observations rather than using the modeled T profile as we had done previously. Thus, in the current version of the draft the only result of the modeling used for analysis is the intensity profile, used to create residual map in Figure 2 and the residual 1D profiles in Figures 3 and 4. As we pointed out to the referee above, while uncertainties in the intensity from the model could change the amplitude of the annular substructures after subtracting data from the model, the locations, widths, and regions of change in the intrinsic data radial profile are not affected by the exact smooth disk model.

Please see the discussion in response to referee above for the age and new discussion we added to Methods.

- The modeling used a grid-based approach which is reasonable given the computational complexity of modeling radiative transfer. However, this makes it hard to determine errors accurately. Discuss the error budget both from input parameters and fitting uncertainty.

Response:
As the referee points out, it is difficult to determine the errors accurately with a grid based approach. As we discussed in multiple areas above, the exact values are unconstrained for any given individual parameter because many of the parameters are degenerate and we lack multi-wavelength high resolution data to break these degeneracies. For reproducibility, however, it is necessary to report the values of the individual parameters used for the best model shown. We have emphasized more in the text that the exact values of the parameters are not to be used at face value.

- If the change in intensity in the rings were not due to a difference in dust-to-gas ratio but a change in dust opacity, what would be the effect (e.g., composition, a_max) and interpretation? Could this be explained as snowlines or dust pileup and agglomeration?

Response:
While we did provide a change in dust opacity as a possible explanation for the radial variation in intensity, without data at a different wavelength, the degeneracies between composition and maximum grain size (which both impact the opacity) are highly uncertain. We feel any discussion on this would be too speculative for the data presented here.

- Note that the moderate dust mass here shows that this young disk is already well into the (globally) gravitationally stable regime.

Response:

The referee is right that the disk is likely indeed gravitationally stable. Due to length restrictions on the text, we lack space to discuss this in detail, and instead remove the citation to secular gravitation instabilities as a possible ring formation mechanism in the main text to avoid confusion for readers.

G. The references are fine.

H. The paper was well written: clear and lucid.

Reviewer Reports on the First Revision:

Referee #1 (Remarks to the Author):

I thank the authors for considering my suggestions, and am largely satisfied with the changes that have been made, in particular softening the statements regarding the age of the source. I have a few remaining thoughts and comments that I would encourage the authors to consider.

SUGGESTIONS FOR IMPROVEMENT:

- The authors answer regarding their choice of robust parameter was tremendously helpful, and may be worth including in the additional material. Though I understand that if there is a word limit that would make this difficult.

- In a few places the authors refer to the "older rings and gaps" and "young azimuthal asymmetries". I would suggest changing these to reflect that the *disks* are younger or older (as compared with the features themselves).

- I still find myself skeptical that the asymmetry in the residual image is real and am worried about over-emphasizing the feature considering the implications. The authors say that they have explored varying the center, and that the asymmetry is visible in the raw, un-model subtracted image as well. However I think further demonstration of these tests in the manuscript is warranted given the degree of certainty with which the feature is discussed. Related to this - where does the center determined by a Gaussian fit actually fall on the image, and how does that, say, compare with the location of the peak (and do these choices affect the appearance/presence of the asymmetry). Also on a that note, the paper by Pinte et al. 2016 may have assumed beta = 1.15, but they also included a prescription for dust settling and found that the large grains were very flat (a scale height of 1 au at 100 au), which is *much* flatter and less flared than the geometry the authors find. I don't necessarily disagree that asymmetries from this geometry should be along the minor axis, but this along with centering could all play into the asymmetry that is seen, and I still think should be explored more robustly, or discussed more circumspectly.

ADDITIONAL COMMENTS:

- I don't necessarily think any additional action is needed on this comment, as I think the authors have mostly already dealt with it. But I did want to re-emphasize that the models the authors refer to regarding the age of 130k - 140k years are tremendously uncertain, and perhaps more importantly, totally uncalibrated (to the best of my knowledge). Being conservative here is the best course of action.

- I'm not quite certain how to approach this, but there was a recent paper claiming the detection of substructures in the disk of L1527, which is also quite a young (Class 0/I) source. I am skeptical, at best, about that result, but it is also a published paper and so perhaps it should not be ignored?

Author Rebuttals to First Revision:

Referee #1 (Remarks to the Author):

I thank the authors for considering my suggestions, and am largely satisfied with the changes that have been made, in particular softening the statements regarding the age of the source. I have a few remaining thoughts and comments that I would encourage the authors to consider.

SUGGESTIONS FOR IMPROVEMENT:

- The authors answer regarding their choice of robust parameter was tremendously helpful, and may be worth including in the additional material. Though I understand that if there is a word limit that would make this difficult.
Response: Thank you for the suggestion. We have added information to the Methods text describing the reasoning of the choice of robust parameter.

- In a few places the authors refer to the "older rings and gaps" and "young azimuthal asymmetries". I would suggest changing these to reflect that the *disks* are younger or older (as compared with the features themselves).
Response: Thank you. This is a good point, and we have changed the wording through the text to reflect the fact that the disks or protostars are younger, not the substructures.

- I still find myself skeptical that the asymmetry in the residual image is real and am worried about over-emphasizing the feature considering the implications. The authors say that they have explored varying the center, and that the asymmetry is visible in the raw, un-model subtracted image as well. However I think further demonstration of these tests in the manuscript is warranted given the degree of certainty with which the feature is discussed. Related to this - where does the center determined by a Gaussian fit actually fall on the image, and how does that, say, compare with the location of the peak (and do these choices affect the appearance/presence of the asymmetry). Also on a that note, the paper by Pinte et al. 2016 may have assumed beta = 1.15, but they also included a prescription for dust settling and found that the large grains were very flat (a scale height of 1 au at 100 au), which is *much* flatter and less flared than the geometry the authors find. I don't necessarily disagree that asymmetries from this geometry should be along the minor axis, but this along with centering could all play into the asymmetry that is seen, and I still think should be explored more robustly, or discussed more circumspectly.

Referee 1's additional requests:

There are a number of tests that the authors have described having done (in their response, and to some degree in the manuscript), and I think having some plots to *show* those tests would be helpful, e.g.:
- changing the centering in different directions and showing how that changes the residuals
- they say the asymmetry shows up in radial intensity cuts, so showing plots of this (perhaps including how the choice of centroid affects it) could be a good demonstration of the asymmetry. If I were doing this, I would probably try to fit some sort of smooth, axisymmetric power-law model (e.g. I ~ r^-gamma) that includes the centroid as a free parameter as a part of the fit, without radiative transfer modeling so that it goes fast enough that everything can be fit together in as rigorous a way as possible (e.g. with MCMC?).
Response: As suggested by the referee, we have taken care to add clear discussion to the text regarding the asymmetry with a dedicated "Asymmetry in the disk" subsection added to Methods. We have added a number of demonstrations and figures showing that the asymmetry is real, which we summarize here:

1) We have recentered the best radiative transfer model by shifting the centroid of the model to minimize the rms of the residuals. The result is shown in Extended Data Fig. 1. The asymmetry persists with the new centroid and we can rule out a radiative transfer centering issue as the cause of the asymmetry.

2) As the referee suggested, to address the issue of potential flaring effects, we have fit a smooth,

axisymmetric power law analytic profile with a flat disk geometry to the data with MCMC (which includes leaving the centroid as well as the disk position and inclination angles as free parameters). The results, shown in Extended Data Figs. 2 and 3, also show asymmetry with excess emission to the north-west and not along the major or minor axes. This both confirms that centering or assumed position and inclination angles are not an issue and demonstrates that even a model with no flaring cannot explain the asymmetry. Hence, the MCMC fit confirms that the asymmetry is real.

3) The referee also suggested we show the radial intensity cuts of the data and show how the choice of centering impacts the cuts. In Extended Data Figs. 4, 5, and 6 we show radial intensity cuts along the major axis, minor axis, and a cut from the south-east to north-west for center positions of the peak intensity position, the Gaussian fit position, and the MCMC fit position. We note that all three center positions are within 2 pixel radii of each other (approximately 2 au or less than half the minor axis beam size). In all cases, there is excess emission in the north-west side of the disk again confirming the asymmetry, this time without the use of any modeling.
   - Note: Extended Data Figs. 4, 5, 6 are rather similar. If the editorial staff prefers to streamline the presentation of this information, we can show only one of the three figures and emphasize in the text that the trends are similar for the other two centroid positions. Please inform us if editorial changes should be made regarding this.

Thank you very much for these suggestions for how to investigate the asymmetry. We feel the new Methods subsection and new figures robustly demonstrate that the asymmetry is real and the paper is made stronger by the additions.

ADDITIONAL COMMENTS:

- I don't necessarily think any additional action is needed on this comment, as I think the authors have mostly already dealt with it. But I did want to re-emphasize that the models the authors refer to regarding the age of 130k - 140k years are tremendously uncertain, and perhaps more importantly, totally uncalibrated (to the best of my knowledge). Being conservative here is the best course of action.
Response: The referee is right that these are uncertain, yet using the models to convert from bolometric temperature to age in a relative sense still places IRS 63 as younger than the other similar sources with rings discussed. We have added a short mention of this in the Methods text for clarity.

- I'm not quite certain how to approach this, but there was a recent paper claiming the detection of substructures in the disk of L1527, which is also quite a young (Class 0/I) source. I am skeptical, at best, about that result, but it is also a published paper and so perhaps it should not be ignored?
Response: We agree that this work (Nakatani et al. 2020), which came out after we had submitted this manuscript, should not be ignored and is suggestive of possible substructures in a disk but not conclusive. Further, due to the edge-on orientation of the L1527 disk, the true nature of these possible substructures (rings, spirals, or another morphology) cannot be definitively determined. We have added text in the Methods section discussing the Nakatani et al. (2020) result.

Reviewer Reports on the Second Revision

Referee #1:

I have reviewed the updated manuscript, and am happy that the authors followed my suggestions for further analysis to explore the existence of the asymmetry in the disk. The authors have done everything I asked, and the additional analysis, text, and figures are very helpful for assessing this feature. At this point I am satisfied with the changes that the authors have made, and do not need to see any further changes.

Author Rebuttals to the Second Revision:

N/A