Peer Review File

Manuscript Title: A 62-minute orbital period black widow binary in a wide hierarchical triple

Reviewer Comments & Author Rebuttals

Reviewer Reports on the Initial Version:

Referees' comments:

Referee #1 (Remarks to the Author):

This paper presents the discovery of a likely black widow pulsar in a very short orbit in a hierarchical triple. While some of the specific aspects of the system are not unique (low-metallicity wide-separation pulsar binaries are known, as well neutron stars in hierarchical triples), other aspects of the system are unique (the orbital period and method of discovery) and the combination of traits certainly represents an extreme member of the pulsar/spider family. The light curve data set and analysis is extensive and the identification of the system as a black widow is compelling. I have some significant comments that would need to be addressed before I could recommend the paper for acceptance, including some additional astrometric analysis to firm up the association of the likely pulsar binary with the sdK star, clarifying the constraints on the properties of the black widow companion and of the putative tertiary star and its orbit, and on the interpretation of the origin of the system.

In its present form, I feel that the paper will be of interest to people working in neutron star binaries, but the paper perhaps does not fully establish that it is of wide interest to those in broader fields, such as compact objects and supernovae. The discovery of one of these systems solely via optical data is interesting, but mostly to people in this subfield. That the system is a hierarchical triple is also interesting, but the lack of a pulsar discovery and the possible eclipsing nature of the inner binary means that e.g. interesting GR tests as done in some of the other hierarchical triples known are unlikely in this system. Other low-metallicity, wide-binary neutron star binaries are known (as mentioned in the text) and are still a bit mysterious; if a more compelling case could be made that this system was ejected from a specific globular cluster, or a case made that it definitely couldn't have come from a cluster (and hence was strong evidence of accretion-induced collapse) then *that* could be of wider interest. Some of my comments below relate to these broader issues, and it's possible that by addressing these comments the potential wider interest of the paper could be established.

More significant comments

--The analysis presented in the paper is suggestive that the spider binary and the tertiary indeed are co-moving, but the authors could do a better job really firming this up. First, the Gaia EDR3 astrometry of the tertiary has a large error (RUWE of > 1.4) suggesting the published astrometric fit is imperfect. It would be good for the authors to explore an independent measurement of the proper motion of the tertiary, using a combination of existing red imaging that maximizes the sdK star (e.g. SDSS and UKIDDS, or SDSS and their new red imaging) to verify the Gaia proper motion. Next, given how bright the system is in u in one of the SDSS epochs, it seems plausible to derive largely independent proper motion for the black widow companion from that u-band epoch and one of the newer UV/blue epochs where the system is very bright in u and hence the contribution of the sdK is minimized. This would help a lot in confirming the association of the binary and putative tertiary.

--The authors should do more to explore a possible origin in a globular cluster. First, the authors need to give additional details on the uncertainty in the sdK metallicity estimate. A very precise [Fe/H] value is listed without an associated uncertainty. Then, for the link to clusters: while the kinematics are uncertain due to the uncertain distance of the system, there are other constraints in that (a) there aren't too many clusters with metallicities around -1.17 (this is where the metallicity uncertainty is relevant) and (b) there aren't too many clusters that have a substantial population of millisecond pulsars. That immediately rules out most host clusters, and with Gaia EDR3 nearly all clusters have precise kinematics known, so a rough comparison between ZTF J1406+1222 and known clusters should still be possible to see how many origin clusters are truly plausible.

--The paper needs to be clearer about the constraints on the size of the companion to the neutron star. This is intimately tied to the model constraints on the inclination; in one place the paper says that inclinations of 30-90 degrees provide reasonable fits, but in another the authors day that fits < 30 degrees can also fit the data (pg 44 "Our heating models of the lightcurve achieve good fits for inclinations below 30 degrees"). Presumably some very face-on inclination can't work, so the authors should be more precise about the constraints, and the implications on the size of the companion. The authors should also be clearer about the constraints on the nature of the black widow companion: given the constraints on the size, what range of masses might be plausible for a white dwarf, and would such an object still be expected to have a hydrogen shell at that mass?

--The boldface paragraph emphasizes that "ZTF J1406+1222 pushes the boundaries of evolutionary models, falling close to or below the minimum orbital period thought possible for such systems.", but no additional discussion of this point appears elsewhere in the main text or methods. The authors need to add additional details and justification for this point: is it due to an extra source of angular momentum loss compared to simple models? (Simple models, in general, do not reproduce the distribution of spider companions in a mass--orbital period diagram.) Are the spectra consistent with a lower hydrogen fraction, or is this unconstrained? Is the role of the tertiary relevant? Does it favor dynamical formation?

--The paper needs to have more discussion of the inferred semi-amplitude of the emission lines: unless the system is face-on, any emission coming from near the companion should have a much, much larger semi-amplitude than the listed value of 112 +/- 15 km/s. For example, in PSR J1311--3430, while the emission lines do not follow a clean sinusoidal pattern, their semi-amplitude is relatively close to that of the secondary. Of course, there is no guarantee that in ZTF J1406+1222 the wind shows the same behavior, or that the emission is coming from the same location in the wind as in PSR J1311--3430, but it does suggest that other interpretations should be considered. One is that the binary is relatively face-on; another is that the emission lines are coming from the shock, closer to the neutron star, rather than a wind.

--The paper spends a fair amount of time using the estimated orbital properties of the tertiary to constrain the formation of the system, but as far as I can tell, doesn't list the estimated mass of that star anywhere (presumably pretty accurate from its Teff and [Fe/H]), nor the range of reasonable semi-major axes (the system could be eccentric or circular, so this is pretty big, plus the distance uncertainty). Carefully showing these values and analysis is necessary for the reader to understand the implications for the system formation.

--For data availability: my view is that "on request" is not a sufficient modern practice, since in 5 or 10 years the corresponding author may not be able to supply these for many common reasons. The complete set of photometric measurements and reduced spectra need to be made available in a data repository with a link provided in the paper. Of course my advice is only advisory to the editor, but I would not recommend acceptance of this paper until this was done.

Other, more minor comments

--The boldface paragraph states that the black widow is being ablated by the X-ray flux of an intrabinary shock. It's far from clear that this is the main source of high-energy emission, it's instead possible that the MeV/GeV gamma-ray emission from the pulsar is the most relevant aspect (which the authors of this paper indirectly confirm via their very high fit for the "heating" luminosity, much higher than the X-ray limit from the putative shock). As a number of recent papers have shown (e.g., Ginzburg & Quataert 2020), it's also unclear whether ablation is actually the main direct source of mass loss for black widows.

pg 7:

-- "We did not detect pulsations to a limit of < 0.1 mJy and < 0.2 mJy at L-band and S-band, respectively" While there is more detail in the Methods, the authors should state what sort of upper limits these are (especially since the unusual choice of 6-sigma is used).

pg 14:

--I don't understand the origin of the listed systemic velocity in Table 1; it appears nowhere else in the paper. It's not clear what spectral measurement (emission, absorption) it comes from.

pg 28:

--It would be good for the authors to estimate how the inferred heating luminosity depends on the assumed neutron star mass, and to do the same calculation for a more massive (1.8 or 2.0 Msun) neutron star, since that would imply a larger semi-major axis.

pg 43:

--The left limit of the box really ought to be extended to the left, since eta=1, while possible, is not a typical value for black widows.

Referee #2 (Remarks to the Author):

A. The results are clearly summarised in the first part of the manuscript. The discussion of key results is presented with an appreciable level of detail in the METHOD sections.

B. The manuscript reports on the discovery of a BW in a triple system, which would be the first of its class. To the best of my knowledge, only another MSP has been identified in a triple system (J1903+0327; Freire et al. 2011), with a possible candidate (J1935+1726; Mignani et al. 2014). Although in both cases the non-degenrate companion seems to be a MS star and the orbital parameters are much at variance with those of ZTF J1406+1222, it would be good to mention these cases to stress the uniqueness of this discovery.

On a different subject, the authors claim that this is the first case of a BW discovered thanks to optical observations. It might be worth to mention the cases of the BWs/RBs PSR J1653−0158, J2039-5617, J1311-3430, 2339-0533 (Nieder et al. 2020; Clark et al. 2021; Pletsch et al. 2012; Ray et al. 2014) where the discovery of optical flux modulations prompted the identification of a new BW/RB. I appreciate that ZTF J1406+1222 was identified through a blind all-sky search and not from the follow-up of gamma-ray sources but I think that it is worth to clarify this point here.

C. I have no remark on this point

D. The statistical treatment of the data is satisfactory. I would add how the discovery of ZTF J1406+1222 fits with the current BW populations and how many of those are expected to be discovered in a search such as that carried out by the authors. Is ZTF J1406+1222 the only candidate spider found in their search?

E. Conclusions are outlined quite well. Admittedly the weight is affected by the lack of non detection of the neutron star in radio, x and gamma rays. I would have loved to read more on the origin of the system. I agree with the authors that a birth in a GC is probably the most likely explanation. Since the authors master very well galactic orbit simulation codes, it would have been interesting to extrapolate back in time the orbit of the system and of Galactic GC to see whether an association was possible. THis would add more weight to the GC formation hypothesis.

F. No gamma-ray counterpart to ZTF J1406+1222 was found in the 4FGL but this could be due to selection effects in the catalogue building. The angular separation between the two optical sources in ZTF J1406+1222 is ~0.6 arcsec. Is there a chance to obtain to separate the two sources with HST and/or AO images?

G. References are properly reported. I would only add works cited in section B of this report.

H. I have no major objection on this point. I think that the manuscript will gain in clarity once the authors have accounted for my suggestions.

Referee #3 (Remarks to the Author):

I think this is a generally excellent paper about a very interesting discovery. I think that the key ideas - that this is a black widow type system, and that it is in a wide hierarchical triple, are well demonstrated. I do have a few minor comments.

The authors' claim that this could be the tip of the iceberg - a large population of similar objects could be simply awaiting discovery - is more difficult. The authors began with the GAIA all-sky survey and selected systems meeting very broad criteria. The authors do not mention whether other systems like this could be in their sample of 25,000 objects. If so, I recognize that they may not be ready for publication but this would support the claim that this is the tip of the iceberg, and perhaps the authors could at least indicate that investigation of other candidates is ongoing. If not, if this is the only plausible candidate, that suggests thst further systems like this must be below the sensitivity threshold of their survey, raising questions about just how numerous systems like this could be.

The presence of the third body is surprising. It appears to have played no role in the detection (and thus is not favoured by selection effects) but to require exotic formation scenarios. In the situation of radio millisecond pulsars, where the presence of a third body likewise does not appear in selection effects, there are between one and two hundred known, and very few appear to have a third body like this. As the authors point out, J1024 has a wide companion, PSR J0337+1715 has a relatively close third body, and one of the globular cluster millisecond pulsars appears to have a long-period Jupiter-mass companion. This rarity of higher multiplicity among discovered radio millisecond pulsars seems compatible with the exotic formation mechanisms we expect to be necessary to preserve a triple system through supernova and spin-up. Perhaps the authors could comment on whether the higher multiplicity in this system is a suprising discovery for this reason?

In the discussion of formation mechanisms, perhaps it would be good to mention the competing theories for the formation of PSR J0337+1715? Tauris and van den Heuvel present one scenario, and alternatives have also been proposed. I realize that the requirements for a wide triple are not the same as those for a relatively tight triple system, but some comments about the similarities or differences might be helpful in explaining how exotic the formation mechanisms must be. Likewise, PSR J0337+1715 has nearly circular orbits, coplanar and with their major axes nearly aligned, suggesting that some alignment mechanism acted upon the system during its formation. Is something similar expected for this system? Does this challenge the idea that the Kozai-Lidov mechhanism is relevant to its formation?

This may only need to be in the Methods, but I would have appreciated a little more detail on how convincing the association of the third body with the binary was. As I understand it, the GAIA proper motions match, and in fact the position angle of the barely-resolved image remains constant, suggesting a very high degree of matching in the motion. How convincing should this be? What is the probability of a coincidental alignment? Do the radial velocities inferred from spectral lines provide a useful compatibility check on this?

The authors measure an orbital period from the optical observations, and in fact they have quite a long span of ZTF observations. I think it would be valuable to provide an orbital ephemeris that would allow computing absolute orbital phase at arbitrary times. I think in particular that it would be valuable to indicate the orbital phase coverage of the radio pulsation search data, as many spider pulsars exhibit radio eclipses when the companion passes near the line of sight. Such an orbital ephemeris would be useful to researchers planning to follow up this discovery; in particular it would be valuable to re-observe the system with a larger radio telescope having a wider bandwidth. It would also be valuable to fold the Fermi photons at the orbital period and look for variability at this period - this might allow detection of an otherwise too-faint source, given Fermi's relatively poor photon localization. Publication of an orbital ephemeris would enable such follow-up work.

In the authors' modelling, can they say anything about how likely the companion is to be filling its Roche lobe? Some spider pulsars are observed to transition between radio millisecond pulsar and accretion-disc states. So far this has been only "redbacks" with larger companion masses, but this is an extremely close binary. Is there any evidence that there might have been a state transition during the history of observations? At a kiloparsec the source might well not have triggered all-sky monitors, but it might be worth checking. Likewise are there archival optical observations with very different magnitudes or colours?

As the authors are contrasting their method of discovery - optical variability with the more common approach of following up unassociated gamma-ray sources, I think it might be worth mentioning the gamma-ray non-detection in the main body of the paper.

I understand that the light curve modelling serves primarily to rule out a white dwarf primary, and that none of the models are able to reproduce the asymmetric colour light curve (I note that Rene Breton and hist students have done some work on reproducing asymmetric light curves from redback systems). But I am slightly concerned by the way the neutron-star modelling is described: the model seems to be an extremely hot 14-km blackbody radiator. In practice radio pulsars emit most of their spin-down power as a pair-plasma wind, whose interactions with a companion are potentially somewhat complicated (there are models for how deep in the companion the energy will be deposited). These models are complicated and beyond the scope of this paper, but I just want to check that the supposed 14-km blackbody radiator is not contributing to the model light curve. It is also worth considering the idea that the pulsar heats the companion indircetly, that the pair-plasma wind energy is reprocessed into X-rays at a shock between pulsar and companion (possibly at L1). Given the estimated irradiation luminosity, should such X-rays have been detected in the observations described in the paper?

Author Rebuttals to Initial Comments:

Referees' comments:

Referee #1 (Remarks to the Author):

This paper presents the discovery of a likely black widow pulsar in a very short orbit in a hierarchical triple. While some of the specific aspects of the system are not unique (low-metallicity wide-separation pulsar binaries are known, as well neutron stars in hierarchical triples), other aspects of the system are unique (the orbital period and method of discovery) and the combination of traits certainly represents an extreme member of the pulsar/spider family. The light curve data set and analysis is extensive and the identification of the system as a black widow is compelling. I have some significant comments that would need to be addressed before I could recommend the paper for acceptance, including some additional astrometric analysis to firm up the association of the likely pulsar binary with the sdK star, clarifying the constraints on the properties of the black widow companion and of the putative tertiary star and its orbit, and on the interpretation of the origin of the System.

Thank you for taking the time to carefully read through the manuscript, and for providing such valuable feedback. We have done our best to address all of the comments, which we believe have significantly strengthened the manuscript.

In its present form, I feel that the paper will be of interest to people working in neutron star binaries, but the paper perhaps does not fully establish that it is of wide interest to those in broader fields, such as compact objects and supernovae. The discovery of one of these systems solely via optical data is interesting, but mostly to people in this subfield. That the system is a hierarchical triple is also interesting, but the lack of a pulsar discovery and the possible eclipsing nature of the inner binary means that e.g. interesting GR tests as done in some of the other hierarchical triples known are unlikely in this system. Other low-metallicity, wide-binary neutron star binaries are known (as mentioned in the text) and are still a bit mysterious; if a more compelling case could be made that this system was ejected from a specific globular cluster, or a case made that it definitely couldn't have come from a cluster (and hence was strong evidence of accretion-induced

collapse) then *that* could be of wider interest. Some of my comments below relate to these broader issues, and it's possible that by addressing these comments the potential wider interest of the paper could be established.

More significant comments

--The analysis presented in the paper is suggestive that the spider binary and the tertiary indeed are co-moving, but the authors could do a better job really firming this up. First, the Gaia EDR3 astrometry of the tertiary has a large error (RUWE of > 1.4) suggesting the published astrometric fit is imperfect. It would be good for the authors to explore an independent measurement of the proper motion of the tertiary, using a combination of existing red imaging that maximizes the sdK star (e.g. SDSS and UKIDDS, or SDSS and their new red imaging) to verify the Gaia proper motion. Next, given how bright the system is in u in one of the SDSS epochs, it seems plausible to derive largely independent proper motion for the black widow companion from that u-band epoch and one of the newer UV/blue epochs where the system is very bright in u and hence the contribution of the sdK is minimized. This would help a lot in confirming the association of the binary and putative tertiary.

This is an excellent suggestion, and we have updated the astrometry figure to incorporate it. We included the positions in the SDSS u and z band astrometry in the figure, and the u measurement indeed appears to track the position of the BW rather than the sdK. We detect the proper motion with significance, and have confirmed that it is indeed consistent with that measured by Gaia, though the uncertainties are larger. We would also like to point out that the object is reported in the USNO-B1.0 catalog (from 2003), and has a similar proper motion reported there (-72 mas/yr pmRA and -8 mas/yr pmDec), and a similar set of values in the PPMXL catalog from 2010 (pmRA -78 mas/yr pmDec -15.5 mas/yr), and the APOP catalog from 2015 (-77.5 mas/yr pmRA and -19.1 mas/yr pmDec). These catalogs all have much larger error bars than the Gaia eDR3 value, but the similarity of their measurements further bolsters our confidence that the proper motion is reliable.

--The authors should do more to explore a possible origin in a globular cluster. First, the authors need to give additional details on the uncertainty in the sdK metallicity estimate. A very precise [Fe/H] value is listed without an associated uncertainty. Then, for the link to clusters: while the kinematics are uncertain due to the uncertain distance of the system, there are other constraints in that (a) there aren't too many clusters with metallicities around -1.17 (this is where the metallicity uncertainty is relevant) and (b) there aren't too many clusters that have a substantial population of millisecond pulsars. That immediately rules out most host clusters, and with Gaia EDR3 nearly all clusters have precise kinematics known, so a rough comparison between ZTF J1406+1222 and known clusters should still be possible to see how many origin clusters are truly plausible.

This is a good suggestion--we have quantified the uncertainties in metalicity by computing the spectroscopic indices of many individual phase-resolved spectra near minimum brightness, rather than just that of the coadded spectrum. We find a large scatter, with a large uncertainty. To help further firm this up, we obtained model spectra for low metalicity stars, and performed a fit for the metalicity and T_eff that way, and now report values from that in the table (and added an inset to the Figure illustrating the red portion of the spectrum showing this fit). The metalicity here is very close to that estimated by the spectroscopic indices, and the uncertainty is much smaller; however, we caution readers in the text that the uncertainty from the covariance matrix of the atmospheric fit is likely underestimated, as we do not know exactly how much the BW is contaminating the spectrum at minimum light.

We performed an extensive kinematic analysis, and found it quite challenging to confidently trace the system back to any particular globular cluster given the large distance uncertainty. However, the suggestion that we explore this lead us to an interesting realization. The kinematic figures in the original draft of the manuscript illustrate three cases, corresponding to our best estimated distance of 1140 pc, and the one sigma uncertainties of this value, 940 and 1340 pc. All these solutions plunge within a few kpc of the Galactic center, and the solution at 940 pc in particular passes within just 50 pc of this region, without any fine tuning on our part (by fine tuning, we can

get it even closer). We believe that this is compelling evidence in favor of a globular cluster origin (in contrast to J1903+0327, where they ruled out such an origin on the basis that the object did not pass near the Galactic Center https://ui.adsabs.harvard.edu/abs/2011MNRAS.412.2763F/abstract), and it is possible that the cluster was disrupted. We find this is compelling for a few reasons. Literature seeking to account for the Gamma ray excess in the vicinity of the Galactic center has suggested it originates from MSPs ejected from disrupted globular clusters, and, by disrupting a globular cluster, it is actually possible to eject such a wide triple from the cluster without needing to worry about a binary+binary interaction to keep it bound, since the Roche limit of the triple system is far more robust than the Roche limit of the much less dense cluster (though one could just as well eject the system with a dynamical binary+binary interaction, and disrupt the cluster later). We have updated the main text to discuss this potential origin, and believe it makes the narrative regarding the origin of the system even more compelling, broadening the appeal of the overall work to those generally interested in the formation and ejection of compact object binaries in globular clusters (something which is relevant to other communities, such as those working on gravitational waves).

--The paper needs to be clearer about the constraints on the size of the companion to the neutron star. This is intimately tied to the model constraints on the inclination; in one place the paper says that inclinations of 30-90 degrees provide reasonable fits, but in another the authors day that fits < 30 degrees can also fit the data (pg 44 "Our heating models of the lightcurve achieve good fits for inclinations below 30 degrees"). Presumably some very face-on inclination can't work, so the authors should be more precise about the constraints, and the implications on the size of the companion. The authors should also be clearer about the constraints on the nature of the black widow companion: given the constraints on the size, what range of masses might be plausible for a white dwarf, and would such an object still be expected to have a hydrogen shell at that mass?

Yes, this is a good suggestion. We omitted a key qualifying statement on the inclination constraints in the text: the lowest inclination for which we could achieve a reasonable looking fit with a completely unconstrained heating model was about 14 degrees, and we now clarify this in the methods section of the text. However, this model required m_irr/m_NS>5 and a very large distance to the system, which is obviously not physical (we believe the model was trying to inflate the irradiated object as much as it could be expanding its Roche lobe). We now explain this in the methods section of the text. The best fit LCURVE model illustrated in Figure 2 converged on a ~40 degree inclination fit (and the ICARUS models shown in the methods converged to 66 degrees, though where they converged depended on what fraction of the data we decided to include creating the model fit). One subtlety in reporting a radius for the secondary from these models is that we are unsure exactly how much flux is coming from the IBS/cometary tail at various phases, or the appropriate value to use as T_irr (since the temperature profile clearly varies as a function of phase, and the optical luminosity of the irradiated companion is quite sensitive to 10% changes in the temperature scale).

However, we can compute a fairly robust lower limit on the radius from the observables, by assuming we are seeing the object's full irradiated surface at peak flux at a temperature of T_irr=10500K across the entire surface (an idealized edge-on case) at the lower estimate of our distance estimate of 940 pc, which yields R_irr>0.03 rsun (we have added this to Table 1). Computing an upper limit to the radius is more subtle, because this requires assuming a more face on inclination, which means the fraction of irradiated surface we see could be diluted by up to a factor of 2. This would still give a fairly tight constraint (at 1340 pc, the upper limit of our distance, we estimate a 10500K irradiated surface viewed face on would correspond to a radius of 0.057 rsun), but it is still an underestimate, since the flux profile we see isn't uniformly at 10500K, but varies significantly, which means much of the illuminated face is at lower T_eff, inflating the radius needed to reproduce the peak apparent magnitudes at 1340 pc. One concrete statement we can make though is that the density of the object must be >10 g/cm^3, because of the 62 minute orbital period (by combining this with the lower limit radius estimate, we get a minimum mass of ~2 Jupiter masses, which is unphysically low, but at least a robust statement that doesn't depend on EOS). We report the minimum density in Table 1, as we believe this may be a useful constraint for people seeking to model the evolution of the system.

0.03 rsun is large for a white dwarf (a fully degenerate pure helium object would have a mass of 0.07 msun at this radius). One possibility is that the irradiated object could be an evolved main sequence star which began mass transfer near the bifurcation period (this is an important channel for producing ultracomapct AM CVn sources, and could in principle just as well be one for generating ultracompact accreting NS systems). These objects can still have substantial hydrogen envelopes, but exhibit significantly higher densities than a pure H MS star, allowing them to dip well below the period minimum while still hosting a hydrogen atmosphere. We have added some discussion in methods regarding this possible formation channel.

--The boldface paragraph emphasizes that "ZTF J1406+1222 pushes the boundaries of evolutionary models, falling close to or below the minimum orbital period thought possible for such systems.", but no additional discussion of this point appears elsewhere in the main text or methods. The authors need to add additional details and justification for this point: is it due to an extra source of angular momentum loss compared to simple models? (Simple models, in general, do not reproduce the distribution of spider companions in a mass--orbital period diagram.) Are the spectra consistent with a lower hydrogen fraction, or is this unconstrained? Is the role of the tertiary relevant? Does it favor dynamical formation?

We have added some additional discussion in the methods text. While the Kozai-Lidov mechanism may have been involved, it is not clear how that could account for the short orbital period of the system. Instead, we suggest it likely results from the donor being H-poor (He-rich), allowing it to reach periods under an hour. However, detailed modelling the evolution of this system is beyond the scope of this particular work (which is already a bit lengthy).

--The paper needs to have more discussion of the inferred semi-amplitude of the emission lines: unless the system is face-on, any emission coming from near the companion should have a much, much larger semi-amplitude than the listed value of 112 +/- 15 km/s. For example, in PSR J1311--3430, while the emission lines do not follow a clean sinusoidal pattern, their semi-amplitude is relatively close to that of the secondary. Of course, there is no guarantee that in ZTF J1406+1222 the wind shows the same behavior, or that the emission is coming from the same location in the wind as in PSR J1311--3430, but it does suggest that other interpretations should be considered. One is that the binary is relatively face-on; another is that the emission lines are coming from the shock, closer to the neutron star, rather than a wind.

This is an excellent point. While it's not obvious whether the relatively low semi-amplitude rules the wind out (since the geometry of the outflow and viewing angle can probably allow for a range of possible effective Doppler shifts of this feature), it is true that the semi-amplitude of the feature in PSR J1311--3430 is comparable to that of the donor, just quite out of phase. It is intriguing to consider that the emission features originates from the IBS, which could be at a much lower velocity than the irradiated object based on its location in the gravitational potential. We have added discussion of this possibility in the main text, and rephrased the text to no longer sound as though we assume this must originate from a cometary tail like that seen in PSR J1311-3430.

--The paper spends a fair amount of time using the estimated orbital properties of the tertiary to constrain the formation of the system, but as far as I can tell, doesn't list the estimated mass of that star anywhere (presumably pretty accurate from its Teff and [Fe/H]), nor the range of reasonable semi-major axes (the system could be eccentric or circular, so this is pretty big, plus the distance uncertainty). Carefully showing these values and analysis is necessary for the reader to understand the implications for the system formation.

We now quote the subdwarf mass estimate in the section describing its analysis (the estimated mass is 0.18-0.3 msun). We have added in some basic estimates in methods concerning the properties of the triple system (which now account for the fact that we don't know the eccentricity, etc, which is an excellent point!). The story has not changed significantly from the first manuscript (we now marginalize over the possible mass range estimates for the BW and sdK in this estimate). The main change is rather than estimating the orbital period at ~10000 years, we now instead quote a lower limit of >10000 years, since us witnessing the current separation only yields a lower limit on the possible orbital period (by assuming an orbit with eccentricity=1, and that we caught the sdK right at the maximum distance in its free fall plunge towards the inner binary).

--For data availability: my view is that "on request" is not a sufficient modern practice, since in 5 or 10 years the corresponding author may not be able to supply these for many common reasons. The complete set of photometric measurements and reduced spectra need to be made available in a data repository with a link provided in the paper. Of course my advice is only advisory to the editor, but I would not recommend acceptance of this paper until this was done.

This is a good suggestion--we have created a github repository with all of the reduced HiPERCAM and LRIS data, and provide a link to it in the data availability statement.

Other, more minor comments

--The boldface paragraph states that the black widow is being ablated by the X-ray flux of an intrabinary shock. It's far from clear that this is the main source of high-energy emission, it's instead possible that the MeV/GeV gamma-ray emission from the pulsar is the most relevant aspect (which the authors of this paper indirectly confirm via their very high fit for the "heating" luminosity, much higher than the X-ray limit from the putative shock). As a number of recent papers have shown (e.g., Ginzburg & Quataert 2020), it's also unclear whether ablation is actually the main direct source of mass loss for black widows.

This is a good point. We have accordingly modified the sentence in the bold faced paragraph to state that this emission originates from the pulsar itself and in the IBS (that particular statement is meant to be broad and reflect the state of all systems, and is not specific to this particular object).

pg 7:

-- "We did not detect pulsations to a limit of < 0.1 mJy and < 0.2 mJy at L-band and S-band, respectively" While there is more detail in the Methods, the authors should state what sort of upper limits these are (especially since the unusual choice of 6-sigma is used).

We have clarified this in methods (we actually moved most of this sentence to methods at the editor's suggestion that observation details be relocated to methods in order to shorten the main text).

pg 14:

--I don't understand the origin of the listed systemic velocity in Table 1; it appears nowhere else in the paper. It's not clear what spectral measurement (emission, absorption) it comes from.

This is a good point--we measured this from the Ca triplet lines belonging to the sdK. We have included text in the methods section discussing the sdK atmospheric analysis to clarify that this is how we measured the systematic velocity.

pg 28:

--It would be good for the authors to estimate how the inferred heating luminosity depends on the assumed neutron star mass, and to do the same calculation for a more massive (1.8 or 2.0 Msun) neutron star, since that would imply a larger semi-major axis.

Yes, this is a good suggestion. We have redone the heating luminosity calculation with a wider range of masses (now we let the total mass of the inner binary vary between 1.45-2.05 msun). This has expanded the possible range of L_H values, and we have updated the table accordingly.

pg 43:

--The left limit of the box really ought to be extended to the left, since eta=1, while possible, is not a typical value for black widows.

This is a good point--we have extended the box down to eta=0.1.

Referee #2 (Remarks to the Author):

A. The results are clearly summarised in the first part of the manuscript. The discussion of key results is presented with an appreciable level of detail in the METHOD sections.

B. The manuscript reports on the discovery of a BW in a triple system, which would be the first of its class. To the best of my knowledge, only another MSP has been identified in a triple system (J1903+0327; Freire et al. 2011), with a possible candidate (J1935+1726; Mignani et al. 2014). Although in both cases the non-degenrate companion seems to be a MS star and the orbital parameters are much at variance with those of ZTF J1406+1222, it would be good to mention these cases to stress the uniqueness of this discovery.

Yes this is a good point--one thing we wish to emphasize is that the triple aspect of the system is exciting not because this system will be an exquisite test of GR like PSR J0337+1715, but rather because from the perspective of the evolution of the system, having a wide companion in a spider binary introduces many possibilities and constraints regarding the system's evolution). We have added some text to the main text emphasizing this. We make direct comparisons to some of the known triple systems in the methods section.

On a different subject, the authors claim that this is the first case of a BW discovered thanks to optical observations. It might be worth to mention the cases of the BWs/RBs PSR J1653−0158, J2039-5617, J1311-3430, 2339-0533 (Nieder et al. 2020; Clark et al. 2021; Pletsch et al. 2012; Ray et al. 2014) where the discovery of optical flux modulations prompted the identification of a new BW/RB. I appreciate that ZTF J1406+1222 was identified through a blind all-sky search and not from the follow-up of gamma-ray sources but I think that it is worth to clarify this point here.

Yes this is a fair point--we now include a statement in the main text on this, with all of these suggested references.

C. I have no remark on this point

D. The statistical treatment of the data is satisfactory. I would add how the discovery of ZTF J1406+1222 fits with the current BW populations and how many of those are expected to be discovered in a search such as that carried out by the authors. Is ZTF J1406+1222 the only candidate spider found in their search?

This is a good point--we have indicated in the methods section that further candidates are being followed up. We have already confirmed at least one additional source, which we recovered in the optical search, and then found was associated with a Fermi Gamma-ray source and Swift x-ray source, but we think it's beyond the scope of this paper to give too many details on that one, which has a longer period, and is not as unique as ZTF J1406+1222.

E. Conclusions are outlined quite well. Admittedly the weight is affected by the lack of non detection of the neutron star in radio, x and gamma rays. I would have loved to read more on the origin of the system. I agree with the authors that a birth in a GC is probably the most likely explanation. Since the authors master very well galactic orbit simulation codes, it would have been interesting to extrapolate back in time the orbit of the system and of Galactic GC to see whether an association was possible. THis would add more weight to the GC formation hypothesis.

This is an excellent suggestion, and lead us to a very interesting discovery! While we couldn't trace the system back to any particular globular cluster, within the astrometric uncertainties, there are orbits which pass very near the Galactic center. This suggests the possibility that the system originated from a disrupted globular cluster, which we find compelling because it provies ejecting the system from a globular cluster was that one would need a fine tuned binary+binary interaction to make it work (since one would otherwise disrupt the wide triple, whose orbital velocity is much smaller than the escape velocity of a globular), but the gravitational field needed to disrupt a globular cluster is far too weak to disrupt the triple (since the average density of matter within the triple is much higher than that of a globular cluster). We have included text on this scenario, and believe it adds weight to the discovery.

F. No gamma-ray counterpart to ZTF J1406+1222 was found in the 4FGL but this could be due to selection effects in the catalogue building. The angular separation between the two optical sources in ZTF J1406+1222 is ~0.6 arcsec. Is there a chance to obtain to separate the two sources with HST and/or AO images?

There are no suitable nearby bright guide stars (and the object is just generally too faint) for AO to be feasible, but HST would certainly be able to resolve the two components. We will submit proposals to acquire these observations in future cycles (along with ultraviolet spectroscopy).

G. References are properly reported. I would only add works cited in section B of this report.

H. I have no major objection on this point. I think that the manuscript will gain in clarity once the authors have accounted for my suggestions.

Referee #3 (Remarks to the Author):

I think this is a generally excellent paper about a very interesting discovery. I think that the key ideas - that this is a black widow type system, and that it is in a wide hierarchical triple, are well demonstrated. I do have a few minor comments.

The authors' claim that this could be the tip of the iceberg - a large population of similar objects could be simply awaiting discovery - is more difficult. The authors began with the GAIA all-sky survey and selected systems meeting very broad criteria. The authors do not mention whether other systems like this could be in their sample of 25,000 objects. If so, I recognize that they may not be ready for publication but this would support the claim that this is the tip of the iceberg, and perhaps the authors could at least indicate that investigation of other candidates is ongoing. If not, if this is the only plausible candidate, that suggests thst further systems like this must be below the sensitivity threshold of their survey, raising questions about just how numerous systems like this could be.

This is a fair point--we have included a parenthetical statement in methods indicating that we are continuing to investigate other candidates. We have removed the text suggesting that this may be the tip of the iceberg of a large population.

The presence of the third body is surprising. It appears to have played no role in the detection (and thus is not favoured by selection effects) but to require exotic formation scenarios. In the situation of radio millisecond pulsars, where the presence of a third body likewise does not appear in selection effects, there are between one and two hundred known, and very few appear to have a third body like this. As the authors point out, J1024 has a wide companion, PSR J0337+1715 has a relatively close third body, and one of the globular cluster millisecond pulsars appears to have a long-period Jupiter-mass companion. This rarity of higher multiplicity among discovered radio millisecond pulsars seems compatible with the exotic formation mechanisms we expect to be necessary to preserve a triple system through supernova and spin-up. Perhaps the authors could comment on whether the higher multiplicity in this system is a suprising discovery for this reason?

This is an excellent point. We have expanded the discussion of the uniqueness of the triple nature of the system in the main text, and include additional material concerning how fine tuned/exotic the formation mechanisms must be to form the system. In addressing the referee reports, we discovered that the system actually has a Galactic orbit that within error bars, is consistent with passing very near the Galactic center, and thus this introduces a new formation scenario (the triple could have been ejected from a disrupted globular cluster, which eliminates the need for a fine tuned binary+binary interaction to keep it bound).

In the discussion of formation mechanisms, perhaps it would be good to mention the competing theories for the formation of PSR J0337+1715? Tauris and van den Heuvel present one scenario, and alternatives have also been proposed. I realize that the requirements for a wide triple are not the same as those for a relatively tight triple system, but some comments about the similarities or differences might be helpful in explaining how exotic the formation mechanisms must be. Likewise, PSR J0337+1715 has nearly circular orbits, coplanar and with their major axes nearly aligned, suggesting that some alignment mechanism acted upon the system during its formation. Is something similar expected for this system? Does this challenge the idea that the Kozai-Lidov mechanism is relevant to its formation?

We have added some additional discussion in the methods to address these points. Because of the very wide separation of the third body, it seems unlikely that any three-body orbital dynamics, mass transfer, or tidal effects involving the third body would be relevant (unlike in PSR J0337+1715). More detailed modeling in the future would be helpful but is beyond the scope of this work.

This may only need to be in the Methods, but I would have appreciated a little more detail on how convincing the association of the third body with the binary was. As I understand it, the GAIA proper motions match, and in fact the position angle of the barely-resolved image remains constant, suggesting a very high degree of matching in the motion. How convincing should this be? What is the probability of a coincidental alignment? Do the radial velocities inferred from spectral lines provide a useful compatibility check on this?

We have added to the figure demonstrating this association. We used archival SDSS epochs from 2003 (in u and z band, which should approximately track the locations of the BW and sdK, respectively), to show that both these sources appear over an arcsecond away from the corresponding Gaia J2016.0 positions, and clearly appear to co-move between the SDSS and Gaia epochs. We believe this solidifies the conclusion that the sources are associated via their common proper motion. Additionally, the systemic velocity of the sdK absorption lines 67+/-30 km/s, is consistent with the 95+/-36 km/s velocity semi-amplitude of the emission features, though we choose not to report the latter, because it’s not obvious that light originating from those features should exhibit a systemic velocity reflective of the true systemic velocity).

The authors measure an orbital period from the optical observations, and in fact they have quite a long span of ZTF observations. I think it would be valuable to provide an orbital ephemeris that would allow computing absolute orbital phase at arbitrary times. I think in particular that it would be valuable to indicate the orbital phase coverage of the radio pulsation search data, as many spider pulsars exhibit radio eclipses when the companion passes near the line of sight. Such an orbital ephemeris would be useful to researchers planning to follow up this discovery; in particular it would be valuable to re-observe the system with a larger radio telescope having a wider bandwidth. It would also be valuable to fold the Fermi photons at the orbital period and look for variability at this period - this might allow detection of an otherwise too-faint source, given Fermi's relatively poor photon localization. Publication of an orbital ephemeris would enable such follow-up work.

This is a good point--we give the u-band ephemeris in Table 1 (the relevant parameters are T_0 and P_b). T_0 is derived from the HiPERCAM data, and P_b from the ZTF observations, which when combined should give an ephemeris stable over many years.

In the authors' modelling, can they say anything about how likely the companion is to be filling its Roche lobe? Some spider pulsars are observed to transition between radio millisecond pulsar and accretion-disc states. So far this has been only "redbacks" with larger companion masses, but this is an extremely close binary. Is there any evidence that there might have been a state transition during the history of observations? At a kiloparsec the source might well not have triggered all-sky monitors, but it might be worth checking. Likewise are there archival optical observations with very different magnitudes or colours?

This is a good point--we have added a statement to the modelling section of methods, in which we mention that our models generally preferred low filling factors (<0.5). Unfortunately, the exact filling factor is poorly constrained (it depends on what fraction of the data we choose to model). The filling factor the modelling converges to is likely dominated by constraints on the distance/luminosity given T_irr, rather than being directly influenced by the lightcurve shape.

As the authors are contrasting their method of discovery - optical variability with the more common approach of following up unassociated gamma-ray sources, I think it might be worth mentioning the gamma-ray non-detection in the main body of the paper.

This is a good point--we have explicitly mentioned in the main body of the paper that the source was not detected in Gamma rays (alongside the statement on the X-ray and radio non-detection).

I understand that the light curve modelling serves primarily to rule out a white dwarf primary, and that none of the models are able to reproduce the asymmetric colour light curve (I note that Rene Breton and hist students have done some work on reproducing asymmetric light curves from redback systems). But I am slightly concerned by the way the neutron-star modelling is described: the model seems to be an extremely hot 14-km blackbody radiator. In practice radio pulsars emit most of their spin-down power as a pair-plasma wind, whose interactions with a companion are potentially somewhat complicated (there are models for how deep in the companion the energy will be deposited). These models are complicated and beyond the scope of this paper, but I just want to check that the supposed 14-km blackbody radiator is not contributing to the model light curve. It is also worth considering the idea that the pulsar heats the companion indircetly, that the pair-plasma wind energy is reprocessed into X-rays at a shock between pulsar and companion (possibly at L1). Given the estimated irradiation luminosity, should such X-rays have been detected in the observations described in the paper?

Yes this is a good point--LCURVE treats the irradiator as a simple blackbody, but we compute the flux originating from both it and the irradiated object, and have verified that the contribution directly from the 14-km is negligible in those models (the whole issue of why the WD models would not work, is that the radius of this other BB source becomes large enough that the irradiating object starts to actually contribute to the overall flux, diluting the irradiation amplitude). We also ran models in which we assumed a radius of 1 km, and verified that there was no change, other than the effective temperature on the surface of the BB increasing in order to achieve a fixed irradiation luminosity, so in this regime, LCURVE is effectively just treating the irradiator as a point source emitting some total energy flux isotropically which is intercepted by the companion, and the estimated irradiation luminosity from this is consistent with that inferred from ICARUS. Extended Data Figure 10 explores whether we expect that the system should have been detected given the estimated irradiation luminosity--the big unknown is the irradiation efficiency (which in some black widows, has even been shown to be greater than 1). The current x-ray limits would have failed to detect several known systems with a similar heating luminosity, so it is possible that this system will be detected in future X-ray observations, and we point this out in the text.

Reviewer Reports on the First Revision:

Referees' comments:

Referee #1 (Remarks to the Author):

Overall I am satisfied with most of the responses to my original comments and think the firming up of the association of the BW binary with the outer subdwarf is convincing, and also appreciate the clarifications to some of the conclusions about the properties of the inner binary and the triple.

In my mind the question of the origin of the system and its broader implications is still the most outstanding one. While the new suggestion that the triple could have originated in a globular cluster disrupted near the Galactic Center is an intriguing one, this raises additional questions that aren't addressed in the revised paper. The main one in my mind is whether the expected rate of globular cluster disruption is consistent with observing even one of these systems close (within ~ 1 kpc) of the Sun: the implied space density must be large unless an appeal to luck is made. I wonder if it's even consistent with other constraints we have on stars kicked out to the solar radius from centrally dissolved globular clusters (about half of these have weird chemical abundances and are easily identified in spectral surveys). If the authors could better establish that this was a reasonable scenario, then I do think the interest in this discovery would be broader, though I see the possible connection to the Galactic Center excess as not being too interesting. We know that globular clusters containing MSPs must have been destroyed (even if this turns out not to fully explain the gamma-ray excess) so a tenuous existence proof of one such system doesn't tell us much about the overall astrophysical problem.

In the realm of other explanations, the authors did not clearly establish in their response whether they had fully investigated whether the system could indeed have originated in any of the globular clusters with similar metallicity that have pulsars (there are 4 of these with metallicities within ~ 1 sigma of their value: M4, M62, NGC 1851, and NGC 6517, and 6 more with in ~ 2 sigma). Two of these are relatively close to the Sun (right now). It would be good to have a clearer statement about whether an association with any of these is possible or is ruled out.

Referee #2 (Remarks to the Author):

I have read the revised version of the manuscript and I appreciate the effort that the authors put in addressing my suggestions The manuscript has now been shortened, which helps gaining in clarity.

I have a comment on the revised version

l105: the authors state the variable lines cannot be originated in a nebula, like in the case of PSR J1959+2048, for which
the lines "are thought to originate primarily from the shocked interstellar medium near the system"

l110: about ZTF the authors say that the lines "do not originate from the surface of the irradiated object, but rather from an intra-binary shock"

these two parts generate some confusion and I suggest a rephrasing to avoid ambiguity, maybe marking the difference between a nebul, which can embed the entire system, and an intrabinay shock

Other than that, I agree that ZTF is a unique object but the authors ahould try to conclude their manuscript better explaining why and how this uniqueness is important from an astrophysical point of view.

Author Rebuttals to First Revision:

Referees' comments:

Referee #1 (Remarks to the Author):

Overall I am satisfied with most of the responses to my original comments and think the firming up of the association of the BW binary with the outer subdwarf is convincing, and also appreciate the clarifications to some of the conclusions about the properties of the inner binary and the triple.

In my mind the question of the origin of the system and its broader implications is still the most outstanding one. While the new suggestion that the triple could have originated in a globular cluster disrupted near the Galactic Center is an intriguing one, this raises additional questions that aren't addressed in the revised paper. The main one in my mind is whether the expected rate of globular cluster disruption is consistent with observing even one of these systems close (within ~ 1 kpc) of the Sun: the implied space density must be large unless an appeal to luck is made. I wonder if it's even consistent with other constraints we have on stars kicked out to the solar radius from centrally dissolved globular clusters (about half of these have weird chemical abundances and are easily identified in spectral surveys). If the authors could better establish that this was a reasonable scenario, then I do think the interest in this discovery would be broader, though I see the possible connection to the Galactic Center excess as not being too interesting. We know that globular clusters containing MSPs must have been destroyed (even if this turns out not to fully explain the gamma-ray excess) so a tenuous existence proof of one such system doesn't tell us much about the overall astrophysical problem.

We are hesitant to make any statements about space density or rates on the basis of a single exotic system (when the sample size is 1, there is a long tail in the posterior distribution for space density). We believe that given the kinematic difficulties of forming such a wide triple, the orbital trajectory in the Halo plunging in and out of the core of the Galaxy, and the low metaliticy of the companion, the hypothesis that this system may have originated from a disrupted cluster in the Galactic center is a reasonable one.

In the realm of other explanations, the authors did not clearly establish in their response whether they had fully investigated whether the system could indeed have originated in any of the globular clusters with similar metallicity that have pulsars (there are 4 of these with metallicities within ~ 1 sigma of their value: M4, M62, NGC 1851, and NGC 6517, and 6 more with in ~ 2 sigma). Two of these are relatively close to the Sun (right now). It would be good to have a clearer statement about whether an association with any of these is possible or is ruled out.

The main difficulty we encountered with this exercise is that we do not know the age of the system, and the astrometry is quite poorly constrained. We didn't find any compelling solutions where the system was ejected from one of the nearby clusters during its current orbit around the Galactic center (but this doesn't probe very far back, as the system's last pass by the Galactic center may have been as recently as 100 Myr ago); the uncertainties in astrometry quickly propagate as one tries to go farther back in time, especially given that this object regularly interacts with the gravitational mass near the center of the Galaxy.

Referee #2 (Remarks to the Author):

I have read the revised version of the manuscript and I appreciate the effort that the authors put in addressing my suggestions The manuscript has now been shortened, which helps gaining in clarity.

I have a comment on the revised version

l105: the authors state the variable lines cannot be originated in a nebula, like in the case of PSR J1959+2048, for which the lines "are thought to originate primarily from the shocked interstellar medium near the system"

l110: about ZTF the authors say that the lines "do not originate from the surface of the irradiated object, but rather from an intra-binary shock"

these two parts generate some confusion and I suggest a rephrasing to avoid ambiguity, maybe marking the difference between a nebul, which can embed the entire system, and an intrabinay shock

This is a good point. We have clarified in both statements. For the shocked interstellar medium, we have added an additional clause at the end of the sentence to indicate that this is a large scale structure: "which are thought to originate primarily from the shocked interstellar medium near the system\cite{Kulkarni1988B}, a large scale structure which extends far beyond the binary." . Similarly, we now explicitly indicate that the intra-binary shock is located between the pulsar and its companion "but rather from an intra-binary shock located between the pulsar and the companion"

Other than that, I agree that ZTF is a unique object but the authors should try to conclude their manuscript better explaining why and how this uniqueness is important from an astrophysical point of view.

We have added the following concluding paragraph emphasizing the uniqueness of the system. Because the last sentence of the bold faced paragraph was cut we thought to emphasize the uniqueness of the optical discovery here (this does have big implications in our opinion, because all previous work identifying these systems has carried a selection bias by relying in part on radio, X-ray, and Gamma ray emission from the system, even if an optical counterpart was eventually

identified). This does lengthen the manuscript to 1738 words–if this is an issue, we are also ok with omitting the paragraph.

“The hierarchical triple nature and short orbital period of ZTF J1406+1222 distinguish it from any known spider binary, challenging formation models of these systems. Uniquely among spider binaries, ZTF J1406+1222 was identified using only optical photons, a novel method of discovering binaries hosting neutron stars solely on the basis of their strongly irradiated companions, thus potentially eliminating strong selection effects in previous work, which identified them primarily on the basis of their emission at radio, X-ray, or gamma-ray wavelengths.”