nature research
https://doi.org/10.1038/s41586-020-2713-y

Supplementary information

A giant planet candidate transiting a white dwarf

In the format provided by the authors and unedited

Peer Review File

Manuscript Title: A Giant Planet Candidate Transiting a White Dwarf

Redactions – Third Party Material

Parts of this Peer Review File have been redacted as indicated to remove third-party material.

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 white dwarf being transited by a Jupiter-sized object that the authors constrain to be of planetary mass using Spitzer 4.5 micron data. I have to say that without the Spitzer data I would be unconvinced by this result, since the optical data are perfectly consistent with a brown dwarf or even a low-mass stellar companion. However, the Spitzer result is very convincing and I believe that the authors interpretation of the planetary nature of the companion is correct. Therefore, this represents an extremely exciting result, more than worthy of being published in Nature.

The paper is well written and easy to follow, I congratulate the authors on a excellent piece of work. I have only a few minor issues to raise:

I assume the uncertainties quoted in table 1 are one sigma errors? If so this should be stated in the caption.

In the transit analysis section the authors should mention that the prior on the WD density is a Gaussian distribution.

Ordinarily it is not possible to simultaneously constrain the inclination and radii of both objects from the transit of a white dwarf (see section 3 of Parsons et al. 2017, MNRAS, 470, 4473 for example). This is because the transit only contains two pieces of information, the total width and the ingress/egress duration (made worse in this case by the partially eclipsing nature preventing a complete measure of the latter). This is significantly different to the transit of a planet across a star, where the depth of the transit yields the radius ratio. In this case the star is significantly smaller than the transiting object and so the depth of the transit is not sensitive to the radius ratio - it just tells us what fraction of the white dwarf's surface is covered (minus any dilution).

I suspect that it is the inclusion of the prior on the density of the white dwarf that stops the fit from being completely degenerate. I have absolutely no problem with the authors taking this approach (it seems to me to be the only sensible course of action with the current data), but given this I would really appreciate seeing a corner plot from the MCMC fit to see exactly how the various parameters are constrained (included as supplementary material would be sufficient). Some expansion of the text in the transit analysis section to specifically state how the fits are being constrain would also be appreciated. For example, there are obvious inclination limits above which the eclipse is total and below which there is no transit. What is physically constraining the other parameters?

Along the same line, the authors use the analytical models of Mandel & Agol 2002, but these are only valid if the planet radius is less than a 10th of the stellar radius. The authors should comment on this.

I am not entirely convinced that CE evolution can be ruled out. I note that the authors conclude that CE cannot be ruled out entirely, but I think the text comes across that it is extremely unlikely and I'm not sure this is the case. In general their analysis is sound, although note that lambda (the binding energy parameter) can be computed directly (e.g. from BSE - Hurley et al. 2002, MNRAS, 329, 897 - see also the top panel of figure 5 from Camacho et al. 2014, A&A, 566, 86, where lambda is plotted as a function of the Roche-lobe radius of the primary). As the authors hint at in the text (but made explicit in the Camacho paper) lambda can reach values of ~10 towards the end of the AGB, meaning that it would be relatively easy for the planet to survive a CE at this phase. Using these computed values of lambda, what (range of) value does that imply for alpha (which in recent years has become the parameter most quoted when discussing CE evolution)?

Given that the authors believe that the progenitor of the WD was quite a low mass star then it seems likely that it made it to the end of the AGB (i.e. this was not a higher mass star whose evolution was significantly cut short by a CE phase). The radius of giant stars in this mass range are fairly similar between the RGB and along the early AGB. This means that if the planet wasn't engulfed on the RGB (which seems like the case here, since the WD mass is >0.5Msun), then it is unlikely to be engulfed again before the end of the AGB (since the radius has not increased beyond what the star had at the tip of the RGB and any mass loss would also move the planet outward). It therefore doesn't seem like a high degree of fine tuning to get the planet to be engulfed towards the end of the AGB (its actually quite a likely phase at which the planet would be engulfed).

This being said, I am convinced by the authors arguments that the planet could have been scattered to its current location and I agree that this is probably the most likely origin, but perhaps they should lighten the language a bit when ruling out a CE origin.

Its also worth noting that recently Schreiber et al. 2019, ApJ, 887L, 4) showed that the intense UV radiation of hot WDs can evaporate gas giant planets even at relatively large distances from the WD, so it would certainly have a tough time as a post-CE object. Although the Schreiber study didn't probe planets substantially more massive than Jupiter, so perhaps objects ~10MJup are less affected by this? My guess is that this process effectively sets a lower mass limit on the planet, below which it would have been completely evaporated.

Referee #2 (Remarks to the Author):

To date several thousand exoplanets have been discovered primarily via the Doppler and transit methods. These discoveries are becoming rather routine and mundane to the point that it is rare to claim a discovery as being new and unexpected. This is not the case here. The discovery of a candidate giant planet transiting a white dwarf stars continues in the long series of surprising discoveries among exoplanets.

The authors provide strong evidence for the presence of a transiting exoplanet around a white dwarf. The discovery is sensational because the host star has evolved past its red giant phase. This means that the planet either had to survive the common envelope phase or somehow had to be "deposited" near the star. I have become a bit jaded when it comes to exoplanet discoveries, but I must say this one got me excited. This discovery will definitely be of interest to the community (and I expect it to make some headlines!)

Unfortunately, it is impossible to measure the mass of the object via the Doppler method. Because one does not have the planet mass, the authors are wise to claim this as a planet "candidate". Lacking a mass the authors must confirm the object via the process of validation, i.e. eliminating all possible scenarios such that the planet hypothesis is the only logical explanation. They have done an excellent job of removing all false positive scenarios, particularly in the use of additional data: 1) high resolution imaging date to eliminate background objects, 2) a high S/N light curve taken with a 10-m telescope, 3) spectral data, and 4) finally Spitzer data to look for thermal emission from the companion. Overall, this is a very convincing result.

With a radius of ~ 1 R_Jup the companion can still be an M dwarf star, but at a distance of 25 pcs they should have seen thermal emission with Spitzer. Whether it is really 14 MJup one has to rely on theoretical models. I generally put lower weight to these, but I am still convinced that the companion is a sub-object. Even if its true mass were a factor of a few higher it would still be an interesting discovery and still a planet in my opinion.

It is a mystery how such a planet formed. I am sure that this will keep theorists busy for a while. In this respect, much of the discussion in the "methods" serves only to provide plausible formation scenarios. I am not a theorist, so I cannot really comment on these sections. I think that the paper stands on its own merits presenting the observational evidence.

This paper is certainly appropriate for Nature. It is one of the more important discoveries in the field in the past year or so. We have always wondered whether white dwarfs can host planets - now we know, this is the most convincing evidence by far. It is a nice paper and I recommend publication.

I do have one comment. It is possible to get a photometric mass measurement via the Doppler beaming effect. This has been done with great success using Kepler. I was thus a bit surprised that the authors did not even mention the prospects of using Doppler beaming to get the mass. My guess is that the quality of the TESS light curves are not up to the task, but also because of the short time span of the measurements. Would it help to get more sector data on this target? What about using PLATO in the future? Correct me if I am wrong, but I believe that Doppler beaming will one day measure the mass of the companion.

Referee #3 (Remarks to the Author):

Dear editor,

I have read and reviewed the article "Discovery of a Giant Planet Candidate Transiting a White Dwarf" by Vanderburg et al., submitted to Nature.

This manuscript reports the discovery and follow up of a low mass companion to a white dwarf through transit. The follow up observations seem rather comprehensive and allow the authors to characterize the various stellar components of the system (The white dwarf and the binary of M dwarfs orbiting around it).

The transits themselves have also been observed several times with various instruments in the visible, and it is argued rather convincingly that they must be caused by a substellar object.

Now the main novelty of the article is that the authors have performed infrared observations of the transit and that they claim this allows them to put a stringent upper limit on the mass of the transiting companion. If these constraints are real, this object would be a true outsider compared to other transiting objects around white dwarfs.

The argument is as follows: a massive companion would emit significantly in the infrared, thus decreasing the transit depth in that band compared to the optical. The non detection of this effect thus allows the authors to put an upper limit on the flux emitted by the companion. This upper limit is then compared to models to derive upper limits on the companion's mass.

While the overall argumentation makes sense, there are two main points that need to be clarified.

1) The authors seem to imply that, in the absence of a thermal emission from the companion, the transit depth should be the same in the visible and infrared. This does not seem straightforward to me, especially in a grazing geometry, because of the limb darkening of the star that is different in both

channels. In principle, a planet hiding the same area in two different channels could lead to two different transit depths. So the equality of the transit depths does need to be corrected from this effect before any interpretation in terms of thermal flux can be made.
However, I understand that the fact that the companion hides almost exactly half the star might be saving the day here. Nevertheless, I think the following points should be elucidated:
- What is the actual fraction of the star covered by the planet in the most likely transit models? How does it differ from the percentage of flux dimming?
- The authors use the Limb darkening coefficients (LDC) of Gianninas et al. (2013) that are specifically computed for white dwarfs in the visible. Since the stellar density (and thus gravity) is evaluated during the fitting process, do the authors change the LDCs during the fit or use values for a given gravity (in the latter case, please specify which for reproducibility)?
- The authors use LDCs computed for main sequence stars in the infrared. Considering the many orders of magnitude of difference in gravity between a white dwarf and a main sequence star, can the authors comment on the validity of the LDCs used and their uncertainty?
- In any case, could the authors estimate the uncertainty on the diluting parameter (d) brought about by the uncertainties on the visible and IR LDCs used in the analysis?

2) An important part of the argumentation relies on the use of evolution models to transform fluxes constraints into mass constraints. While the authors make a rather good job at accounting for the age uncertainty, it is also notorious that various models can produce relatively different Mass-Age-Luminosity relations.
- Could the authors try to estimate the uncertainties on the mass limit due to the uncertainty on the evolution model used (for example by reproducing the analysis with another set of publicly available models)?

Another motivation could be that, during this review, I noticed this new grid published on the arxiv (https://ui.adsabs.harvard.edu/abs/2020arXiv200313717P/abstract ; but there are others) which, although producing fairly consistent results, seems to yield systematically lower masses for a given flux. This would tighten the authors argument.

The presence of clouds also tends to redden the spectrum. So clouds usually make a substellar object appear brighter at 4.5mu for a given mass, further tightening the mass constraints. The authors might want to elaborate and quantify that.

Other than these two points, I thought the paper was well written. The text is short and to the point while providing enough details in the Methods section. A few additional minor comments are listed below.

Minor points:
- Some planetary models provide outputs in magnitudes in various bands, including the Spitzer IRAC bands. To allow an easier comparison with them, could you please provide the Spitzer magnitude derived for the WD in table 1 (possibly along with the lower limits on the magnitude of the companion)

- Figure 2 shows a wide region of the parameter space where nothing happens. It might be worth restricting the figure to a smaller mass area. It could also be interesting to show contours of the brightness in the Spitzer band along with the flux upper limits.

- I myself do *not* think that an object stops being interesting when it crosses the 13Mjup mass threshold. As argued in the manuscript there are plenty of problems with the current IAU definition of a brown dwarf and I think that how we might want to label this object is actually physically irrelevant. So, while acknowledging that this is probably an editorial decision, I would strongly encourage the authors to remove the planet/brown dwarf limit in Figure 2.

Author Rebuttals to Initial Comments:

Note: Author responses in Red

We thank the editor and the referees for very useful reports that have helped us considerably strengthen the paper. We have responded to the comments below inline, and have marked the changes in the manuscript in red bold font. In addition to changes made in response to the referees' comments, we have also added some new analysis using a more general expression for the amplitude of spectral features in transmission for grazing transits and a brief discussion of another mechanism for exciting dynamical instabilities (see https://arxiv.org/abs/1508.05715).

Referees' comments:

Referee #1 (Remarks to the Author):

This paper presents the discovery of a white dwarf being transited by a Jupiter-sized object that the authors constrain to be of planetary mass using Spitzer 4.5 micron data. I have to say that without the Spitzer data I would be unconvinced by this result, since the optical data are perfectly consistent with a brown dwarf or even a low-mass stellar companion. However, the Spitzer result is very convincing and I believe that the authors interpretation of the planetary nature of the companion is correct. Therefore, this represents an extremely exciting result, more than worthy of being published in Nature.

The paper is well written and easy to follow, I congratulate the authors on a excellent piece of work. I have only a few minor issues to raise:

I assume the uncertainties quoted in table 1 are one sigma errors? If so this should be stated in the caption.

We have clarified in the table caption that unless stated otherwise, the uncertainties are one sigma errors.

In the transit analysis section the authors should mention that the prior on the WD density is a Gaussian distribution.

We have clarified that the prior is a Gaussian distribution.

Ordinarily it is not possible to simultaneously constrain the inclination and radii of both objects from the transit of a white dwarf (see section 3 of Parsons et al. 2017, MNRAS, 470, 4473 for example). This is because the transit only contains two pieces of information, the total width and the ingress/egress duration (made worse in this case by the partially eclipsing nature preventing a complete measure of the latter). This is significantly different to the transit of a planet across a star, where the depth of the transit yields the radius ratio. In this case the star is significantly smaller than the transiting object and so the depth of the transit is not sensitive to the radius ratio - it just tells us what fraction of the white dwarf's surface is covered (minus any dilution).

I suspect that it is the inclusion of the prior on the density of the white dwarf that stops the fit from being completely degenerate. I have absolutely no problem with the authors taking this approach (it seems to me to be the only sensible course of action with the current data), but given this I would really appreciate seeing a corner plot from the MCMC fit to see exactly how the various parameters are constrained (included as supplementary material would be sufficient). Some expansion of the text in the transit analysis section to specifically state how the fits are being constrain would also be appreciated. For example, there are obvious inclination limits above which the eclipse is total and below which there is no transit. What is physically constraining the other parameters?

The referee is correct that the prior on the stellar density prevents the fit from being totally degenerate. The density informs the scaled semimajor axis, and therefore the average speed of the planet throughout the orbit (normalized to the radius of the star). Knowledge of the average orbital speed links the transit duration to the size of the planet, with an added constraint on the transit impact parameter from the maximum depth of the transit. We have added text to the Methods section discussing these constraints and two new figures showing corner plots.

Along the same line, the authors use the analytical models of Mandel & Agol 2002, but these are only valid if the planet radius is less than a 10th of the stellar radius. The authors should comment on this.

We thank the referee for this comment, but as far as we can tell, the expressions we used are valid for planets of any size relative to their stars. In particular, we used the occult_quad code written by Eric Agol (and updated by Jason Eastman), which implements the equations in Mandel & Agol Section 4, which are an exact solution for transits of a star with a quadratic limb darkening law. We did not use the approximation given by Mandel & Agol in Section 5, which is indeed only valid for Rp/R* < 0.1. So as far as we can tell, the expressions we use are valid for this geometry. We have clarified in the text that we use the exact formulae for stars with quadratic limb darkening laws.

I am not entirely convinced that CE evolution can be ruled out. I note that the authors conclude that CE cannot be ruled out entirely, but I think the text comes across that it is extremely unlikely and I'm not sure this is the case. In general their analysis is sound, although note that lambda (the binding energy parameter) can be computed directly (e.g. from BSE - Hurley et al. 2002, MNRAS, 329, 897 - see also the top panel of figure 5 from Camacho et al. 2014, A&A, 566, 86, where lambda is plotted as a function of the Roche-lobe radius of the primary). As the authors hint at in the text (but made explicit in the Camacho paper) lambda can reach values of ~10 towards the end of the AGB, meaning that it would be relatively easy for the planet to survive a CE at this phase. Using these computed values of lambda, what (range of) value does that imply for alpha (which in recent years has become the parameter most quoted when discussing CE evolution)?

Given that the authors believe that the progenitor of the WD was quite a low mass star then it seems likely that it made it to the end of the AGB (i.e. this was not a higher mass star whose evolution was significantly cut short by a CE phase). The radius of giant stars in this mass range are fairly similar between the RGB and along the early AGB. This means that if the planet wasn't engulfed on the RGB (which seems like the case here, since the WD mass is >0.5Msun), then it is unlikely to be engulfed again before the end of the AGB (since the radius has not increased beyond what the star had at the tip of the RGB and any mass loss would also move the planet outward). It therefore doesn't seem like a high degree of fine tuning to get the planet to be engulfed towards the end of the AGB (its actually quite a likely phase at which the planet would be engulfed).

This being said, I am convinced by the authors arguments that the planet could have been scattered to its current location and I agree that this is probably the most likely origin, but perhaps they should lighten the language a bit when ruling out a CE origin.

This is a great point. We agree that given the uncertainties in the process of common envelope evolution, it is wise to soften our language. We have done so throughout the paper, including removing the statements about fine tuning. We have also added a reference to Camacho et al. (2014) as an empirical constraint on the most likely values of alpha and lambda. We have also highlighted the fact that WD 1856 stands out from the population of known brown dwarf/white dwarf binaries in terms of the energy available to eject the common envelope. Perhaps this empirical comparison is the best evidence that the formation of these objects is different from WD 1856.

Its also worth noting that recently Schreiber et al. 2019, ApJ, 887L, 4) showed that the intense UV radiation of hot WDs can evaporate gas giant planets even at relatively large distances from the WD, so it would certainly have a tough time as a post-CE object. Although the Schreiber study didn't probe planets substantially more massive than Jupiter, so perhaps objects ~10MJup are less affected by this? My guess is that this process effectively sets a lower mass limit on the planet, below which it would have been completely evaporated.

Evaporation from UV radiation from hot white dwarfs is indeed a strong constraint. We referenced a paper by Bear & Soker 2011 (https://ui.adsabs.harvard.edu/abs/2011MNRAS.414.1788B/abstract) which discusses mass loss for post-common-envelope objects (especially in the context of the now-refuted object claimed to be orbiting HD 149382). They conclude that while mass loss can be significant for super-Jupiters (masses 8-23 Mj), it is likely that they can survive. Lower-mass planets will likely be evaporated though. We have now also included a reference to Schreiber et al as well.

Referee #2 (Remarks to the Author):

To date several thousand exoplanets have been discovered primarily via the Doppler and transit methods. These discoveries are becoming rather routine and mundane to the point that it is rare to claim a discovery as being new and unexpected. This is not the case here. The discovery of a candidate giant planet transiting a white dwarf stars continues in the long series of surprising discoveries among exoplanets.

The authors provide strong evidence for the presence of a transiting exoplanet around a white dwarf. The discovery is sensational because the host star has evolved past its red giant phase. This means that the planet either had to survive the common envelope phase or somehow had to be "deposited" near the star. I have become a bit jaded when it comes to exoplanet discoveries, but I must say this one got me excited. This discovery will definitely be of interest to the community (and I expect it to make some headlines!)

Unfortunately, it is impossible to measure the mass of the object via the Doppler method. Because one does not have the planet mass, the authors are wise to claim this as a planet "candidate". Lacking a mass the authors must confirm the object via the process of validation, i.e. eliminating all possible scenarios such that the planet hypothesis is the only logical explanation. They have done an excellent job of removing all false positive scenarios, particularly in the use of additional data: 1) high resolution imaging date to eliminate background objects, 2) a high S/N light curve taken with a 10-m telescope, 3) spectral data, and 4) finally Spitzer data to look for thermal emission from the companion. Overall, this is a very convincing result.

With a radius of ~ 1 R_Jup the companion can still be an M dwarf star, but at a distance of 25 pcs they should have seen thermal emission with Spitzer. Whether it is really 14 MJup one has to rely on theoretical models. I generally put lower weight to these, but I am still convinced that the companion is a sub-object. Even if its true mass were a factor of a few higher it would still be an interesting discovery and still a planet in my opinion.

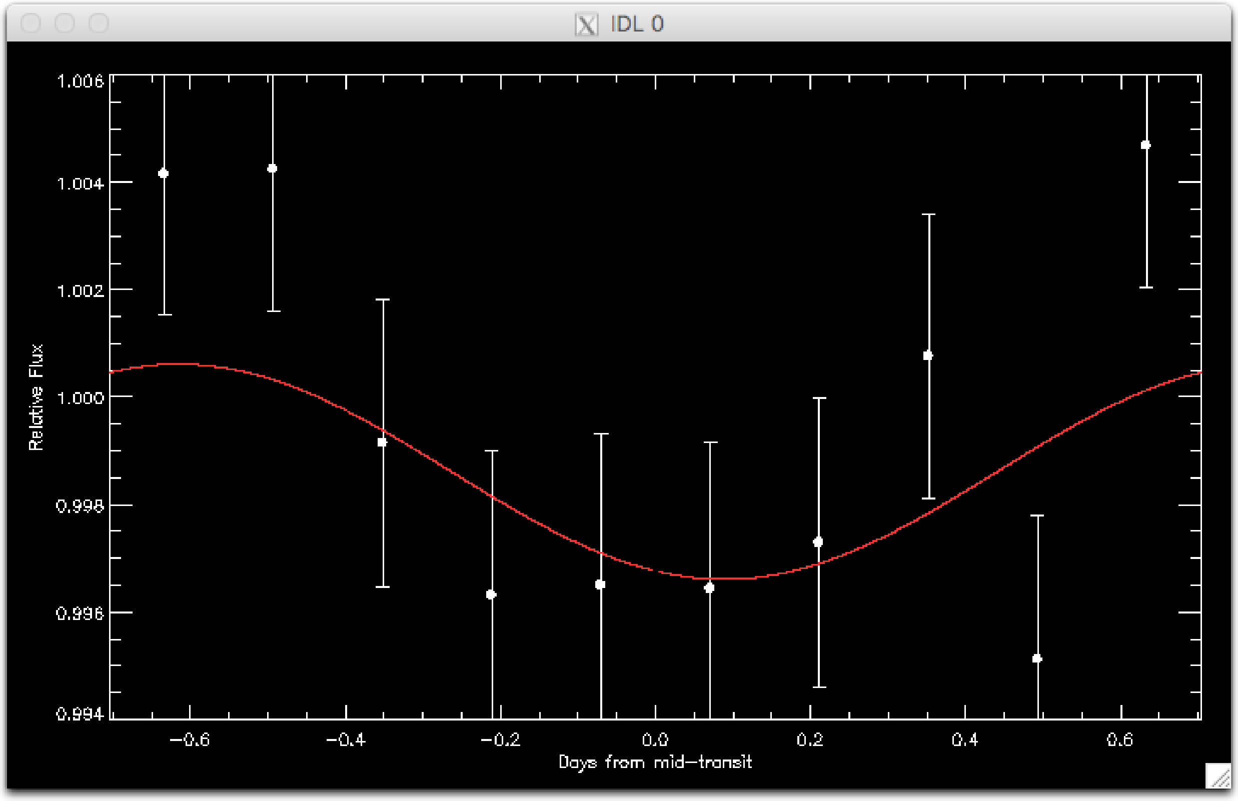
It is a mystery how such a planet formed. I am sure that this will keep theorists busy for a while. In this respect, much of the discussion in the "methods" serves only to provide plausible formation scenarios. I am not a theorist, so I cannot really comment on these sections. I think that the paper stands on its own merits presenting the observational evidence.

This paper is certainly appropriate for Nature. It is one of the more important discoveries in the field in the past year or so. We have always wondered whether white dwarfs can host planets - now we know, this is the most convincing evidence by far. It is a nice paper and I recommend publication.

I do have one comment. It is possible to get a photometric mass measurement via the Doppler beaming effect. This has been done with great success using Kepler. I was thus a bit surprised that the authors did not even mention the prospects of using Doppler beaming to get the mass. My guess is that the quality of the TESS light curves are not up to the task, but also because of the short time span of the measurements. Would it help to get more sector data on this target? What about using PLATO in the future? Correct me if I am wrong, but I believe that Doppler beaming will one day measure the mass of the companion.

This is a great point. The referee is correct that the TESS light curves are not of high enough precision to place meaningful constraints on the Doppler boosting effect. We fit the TESS light curves with a simple sine + cosine model and found the following results:
sin: -770 ppm +/- 1130 ppm (Doppler boosting)
cos:-1850 +/- 1130 (Illumination effect)

For comparison, the expected boosting amplitude is about 50 ppm for at 14 M_J planet, so TESS will not be sensitive to this effect.



It will likely be difficult for PLATO and CHEOPS with larger collecting areas, but similar angular resolution (and therefore similar challenges performing photometry near two significantly brighter stars) to improve upon this measurement as well. CHEOPS should expect photometric precision about 3x better than TESS given its larger aperture, but even after months of observations will not be close to a detection. If PLATO observes WD 1856 at the center of its field of view (with all 32 overlapping cameras), it should achieve about 7-15 times better photometric precision than TESS, but would need two full years to detect a 14 M_J planet at 1-2 sigma.

The best hope for measuring the Doppler boosting signal for WD 1856 b will likely be with either HST or JWST thanks to their much better angular resolution and collecting area, but will be quite expensive. HST has the advantage of a larger Doppler boosting signal (if observations are conducted in the blue optical), but JWST has the larger collecting area. We used HST and JWST exposure time calculators to estimate that observations of a full planet orbit with either telescope (using HST/WFC3/UVIS and the F200LP filter, or JWST/NIRCAM/SW with the F090W2 filter) could yield uncertainties on the boosting amplitude of roughly 50 ppm, and potentially detect a 14 M_j planet at 1 sigma. Stronger constraints (3 sigma) would likely require ~weeks of observations, likely prohibitively expensive.

We note that 4 years of Kepler data probably could have placed interesting constraints on the boosting amplitude.

We have added a subsection to the Methods with our constraints on the TESS

out-of-transit variations and a brief discussion of feasibility of detecting the boosting signal with other facilities.

Referee #3 (Remarks to the Author):

Dear editor,

I have read and reviewed the article "Discovery of a Giant Planet Candidate Transiting a White Dwarf" by Vanderburg et al., submitted to Nature.

This manuscript reports the discovery and follow up of a low mass companion to a white dwarf through transit. The follow up observations seem rather comprehensive and allow the authors to characterize the various stellar components of the system (The white dwarf and the binary of M dwarfs orbiting around it).

The transits themselves have also been observed several times with various instruments in the visible, and it is argued rather convincingly that they must be caused by a substellar object.

Now the main novelty of the article is that the authors have performed infrared observations of the transit and that they claim this allows them to put a stringent upper limit on the mass of the transiting companion. If these constraints are real, this object would be a true outsider compared to other transiting objects around white dwarfs.

The argument is as follows: a massive companion would emit significantly in the infrared, thus decreasing the transit depth in that band compared to the optical. The non detection of this effect thus allows the authors to put an upper limit on the flux emitted by the companion. This upper limit is then compared to models to derive upper limits on the companion's mass.

While the overall argumentation makes sense, there are two main points that need to be clarified.

1) The authors seem to imply that, in the absence of a thermal emission from the companion, the transit depth should be the same in the visible and infrared. This does not seem straightforward to me, especially in a grazing geometry, because of the limb darkening of the star that is different in both channels. In principle, a planet hiding the same area in two different channels could lead to two different transit depths. So the equality of the transit depths does *need to be corrected from this effect before any interpretation in terms of thermal flux can be made.*

The referee makes a great point which we have now explored in detail. It is true that the transit depth in the optical and infrared should not be exactly identical given the different limb darkening profiles in the two bands, and we have clarified this in the main text.

However, I understand that the fact that the companion hides almost exactly half the star might be saving the day here. Nevertheless, I think the following points should be elucidated:
- What is the actual fraction of the star covered by the planet in the most likely transit models? How does it differ from the percentage of flux dimming?

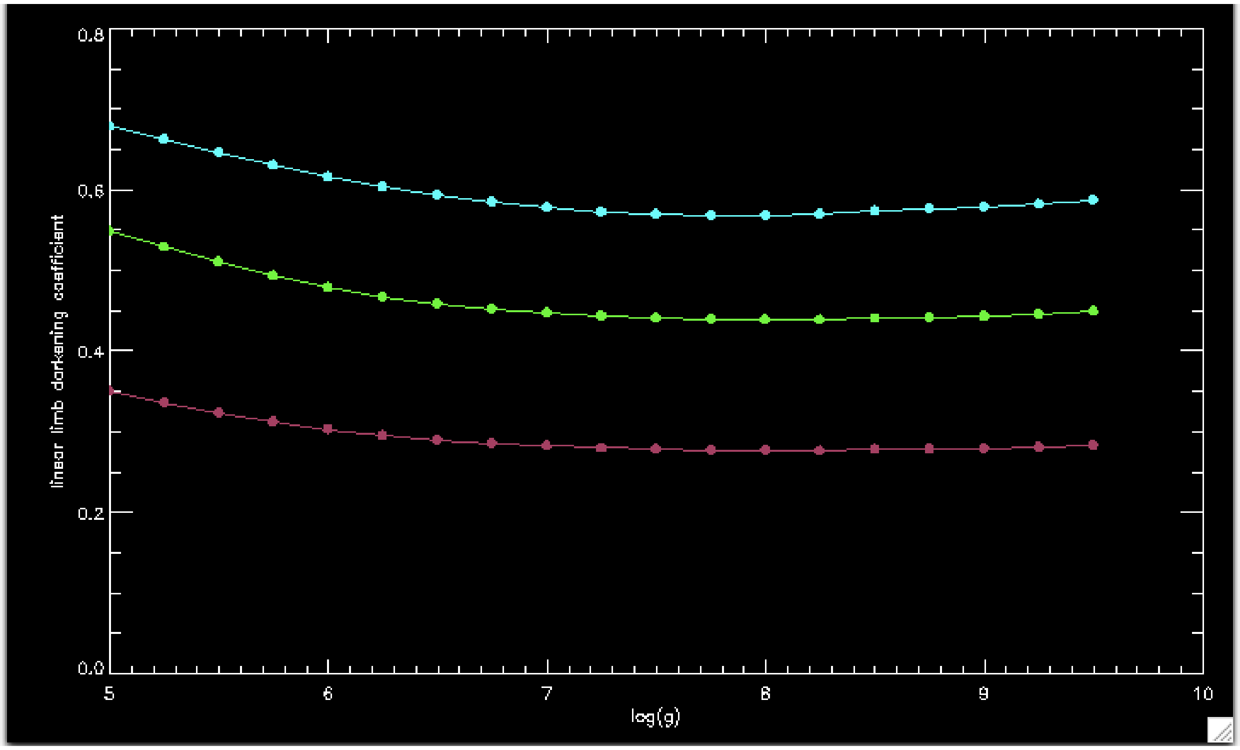
The actual fraction of the star covered by the planet is 56.21 +/- 0.17%, compared to a transit depth of 56.65% +/- 0.18 in the optical and 56.25 +/- 1.45% in the IR. As the referee suspected, the fact that the transit is close to 50% deep helps to cancel out differences due to limb darkening coefficients.

- The authors use the Limb darkening coefficients (LDC) of Gianninas et al. (2013) that are specifically computed for white dwarfs in the visible. Since the stellar density (and thus gravity) is evaluated during the fitting process, do the authors change the LDCs during the fit or use values for a given gravity (in the latter case, please specify which forreproducibility)?

We fix the GTC limb darkening coefficients to (u1 =0.0477, u2 = 0.5159), the values specified for white dwarfs with log(g) = 8 and Teff = 4750 K. For Spitzer, we used (u1 =0.00, u2 = 0.15). We have specified these in the paper.

- The authors use LDCs computed for main sequence stars in the infrared. Considering the many orders of magnitude of difference in gravity between a white dwarf and a main sequence star, can the authors comment on the validity of the LDCs used and theiruncertainty?

Surface gravity tends to have a relatively weak effect on the limb darkening coefficients. Here, we have plotted the model linear limb darkening coefficients from Gianninas et al. 2013 as a function of surface gravity for a 5000 K white dwarf in three different filters:



From top to bottom, the curves show coefficients in u-band (cyan), g-band (green), and z-band (maroon). Over more than four orders of magnitude in surface gravity, the coefficient only changes by ~10-20%.

We are also helped by the fact that limb darkening is generally weaker in the infrared than the optical ([https://ui.adsabs.harvard.edu/abs/1976ApJS...30 .... 1V/abstract](https://ui.adsabs.harvard.edu/abs/1976ApJS...30....1V/abstract)), so uncertainties in the model parameters make less of a difference to the final results (see below).

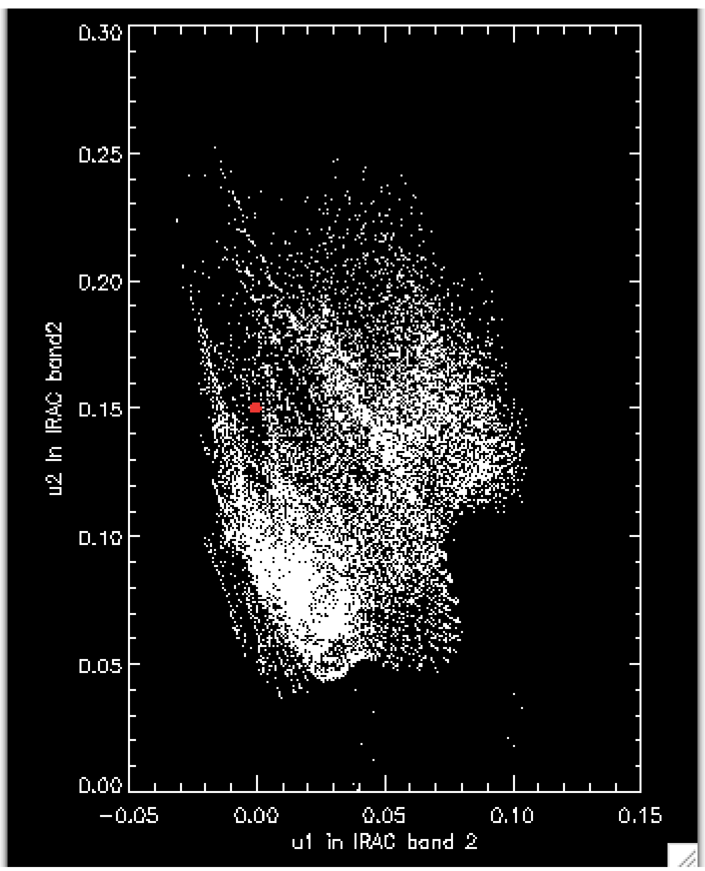
- In any case, could the authors estimate the uncertainty on the diluting parameter (d) brought about by the uncertainties on the visible and IR LDCs used in the analysis?

We performed a few experiments to quantify how our choice of limb darkening coefficients affects our final results.

The first thing we did was to assess the "worst case" scenario -- that is, we assumed we have almost no knowledge of the limb darkening profile of WD 1856. We re-ran our MCMC fit (imposing a circular orbit to speed convergence) and letting the limb darkening coefficients in both g' band and IRAC channel 2 vary freely, with only minimal physical priors. In particular, we only imposed priors from Kipping 2013 to enforce the brightness profile is always positive (the star cannot emit negative light) and decreasing towards the limb (limb darkening, not limb brightening), namely u1 >0, 0<q1<1, and 0<q2<1, where q1 and q2 are transformed limb darkening coefficients from Kipping 2013. We found that in this worst-case scenario, the value and uncertainty on the dilution parameter slightly changed from D = 0.004 +/- 0.029 (with fixed coefficients) to D = 0.010 +/- 0.029 with limb darkening free to vary (a change of 0.2σ). Allowing the limb darkening parameters to float changed the planet/star radius ratio from 7.28 +/- 0.64 to 7.04 +/- 0.67 , about 0.4σ.

The second thing we did was run a test where the limb darkening coefficients in IRAC band 2 were free to vary (with the same physical priors described before), but the GTC parameters were fixed to the model values. Here, the dilution parameter (D = 0.008 +/- 0.029) was a bit closer to the value with fixed coefficients (different by 0.13σ), but this time the planet/star radius ratio was 7.31 +/- 0.63, in very good agreement with the values with limb darkening fixed. This makes sense - information of the detailed shape of the transit is dominated by the more precise GTC light curve.

Finally, we ran a test where we fixed limb darkening parameters for the GTC light curve, and allowed Spitzer limb darkening parameters to vary only over plausible physical values. We looked at the limb darkening coefficients in IRAC band 2 for the entire suite of stellar models explored by Claret & Bloemen 2011 and found that all limb darkening parameter fell in a tight range of (u1,u2) values. The values were all close to 0, indicating stars are nearly uniform in the infrared.

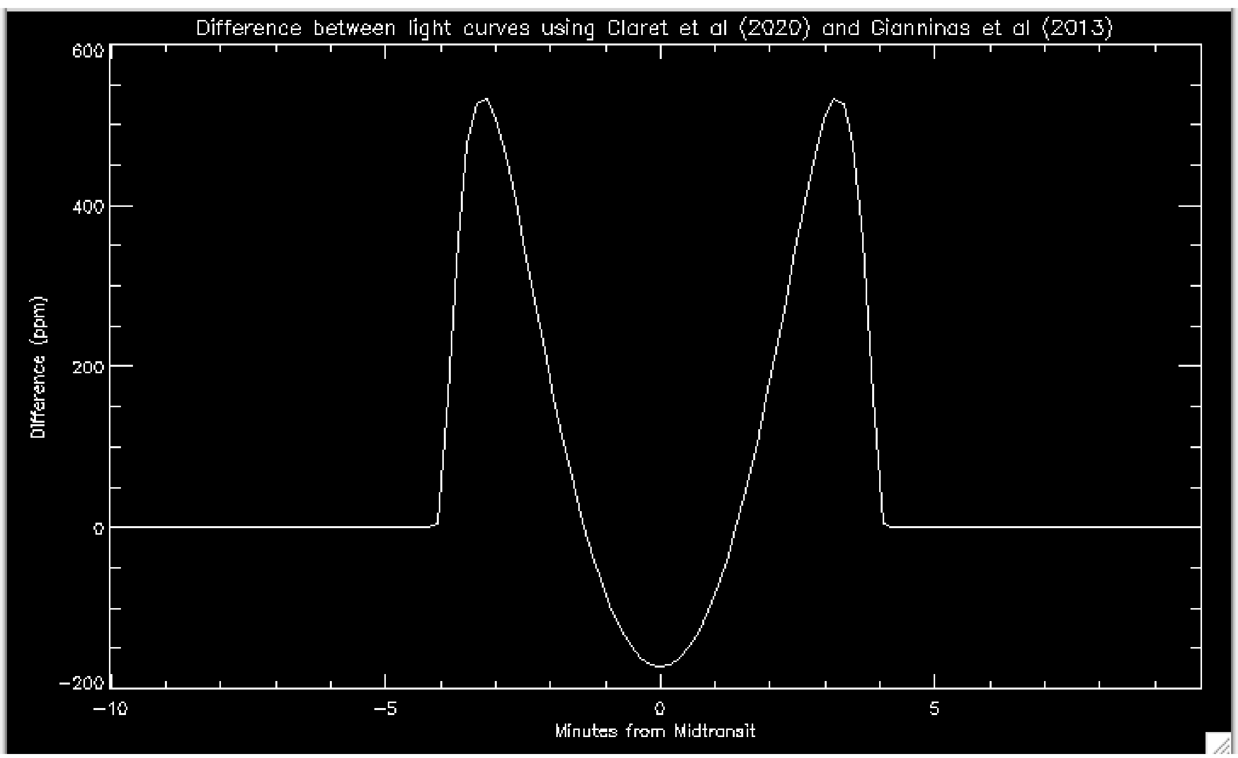


We observed that restricting our fit to allow u1 < 0.2 and u2 < 0.3 on top of the physical priors (Kipping) encompasses nearly the entire suite of Claret & Bloemen models (except for a handful of models with slightly negative u1 values, which are rejected by the

Kipping prior because they show limb brightening instead of limb darkening). When we restrict to these plausible values for Spitzer limb darkening, we find nearly identical results (d = 0.004 +/- 0.29, Rp/R* = 7.29 +/- 0.62) to our baseline fit.

The results of these tests are summarized here:

	Dilution parameter D	Rp/R*
Baseline (fixed LD coefficients for both Spitzer and GTC	0.004 +/- 0.029	7.28 +/- 0.64
Fixed GTC, free Spitzer	0.008 +/- 0.029	7.29 +/- 0.64
Free GTC, free Spitzer	0.010 +/- 0.029	7.04 +/- 0.67
Fixed GTC, Spitzer restricted to plausible values	0.004 +/- 0.029	7.29 +/- 0.62



Finally, we note that another group has independently tabulated optical limb darkening coefficients for white dwarfs (Claret et al. 2019: https://ui.adsabs.harvard.edu/abs/2020A%26A...634A..93C/abstract), and report a value for the GTC (u1 = .0699, u2 = .4749) optical nearly identical to our adopted values: (u1 =0.0477, u2 = 0.5159). The difference between light curves with these limb darkening

parameters is on the order of a few hundred ppm, much smaller than our photometric precision.

We have added a summary of these experiments to the Methods section.

2) An important part of the argumentation relies on the use of evolution models to transform fluxes constraints into mass constraints. While the authors make a rather good job at accounting for the age uncertainty, it is also notorious that various models can produce relatively different Mass-Age-Luminosity relations.
- Could the authors try to estimate the uncertainties on the mass limit due to the uncertainty on the evolution model used (for example by reproducing the analysis with another set of publicly available models)?

Another motivation could be that, during this review, I noticed this new grid published on the arxiv (https://ui.adsabs.harvard.edu/abs/2020arXiv200313717P/abstract ; but there are others) which, although producing fairly consistent results, seems to yield systematically lower masses for a given flux. This would tighten the authors argument.

We thank the referee for pointing out these new models - we have downloaded the evolutionary tracks and find fairly similar results to the Sonora models. The new ATMO 2020 models are only reported to ages of 10 Gyr, so we made a comparison between the 1,2,3 sigma limits for these models with the Sonora models at 10 Gyr.

	1 sigma	2 sigma	3 sigma
Sonora (10 Gyr)	9.4 Mj	11.9 Mj	13.6 Mj
ATMO (10 Gyr)	8.4 Mj	11.5 Mj	15.7 Mj

The ATMO models seem to show lower mass objects cooling somewhat more than higher mass objects, but the agreement between the grids is reassuring. We also tested the effects of non-equilibrium mixing on these results using the ATMO grids calculated with these different assumptions, and found minimal impact on the results.

The presence of clouds also tends to redden the spectrum. So clouds usually make a substellar object appear brighter at 4.5mu for a given mass, further tightening the mass constraints. The authors might want to elaborate and quantify that.

The effects of clouds on our mass limits are more difficult to quantify. We are not aware of any self-consistent evolutionary models that include the formation of water clouds for

cool brown dwarfs/planets below temperatures of about 375 K, so our discussion will be more qualitative.

In general, clouds introduce two competing effects: 1) they slow the cooling of the object, so an object which evolved with thick clouds will remain hotter and more luminous for longer than a cloud-free object, and 2) when present, clouds affect the spectrum of the brown dwarf. The first effect (slowing the cooling of the object) strengthens our mass limits by keeping the planet/brown dwarf more luminous for longer. But the second effect tends to decrease the flux in spectral windows (like at 4.5 microns) at temperatures where clouds are expected to form. WD 1856 b is cool enough that water clouds should be present in its spectrum, so these two effects will compete.

To roughly quantify how much of an effect clouds might have in a scenario where the changes to the spectrum dominate the effect, we looked at calculations from Morley et al. 2018, who show the effect of clouds in a 250 K brown dwarf:

[Figure 9 Redacted: Morely, C. V. et al. An L Band Spectrum of the Coldest Brown Dwarf. *The Astrophysical Journal*, **858**: 97 (2018).]

In this case, clouds decrease the brightness at 4.5 microns by about a factor of 2. In the

13.8 Gyr, 10-20 M_Jup regime, a factor of 2 decrease in the flux would correspond to an increase in our mass limits by about 25%, or about 3 M_jup. It is likely that this effect will be at least somewhat offset by slower cooling due to the presence of clouds at various points throughout the object's lifetime.

We have added a paragraph discussing these issues to the Methods section.

Other than these two points, I thought the paper was well written. The text is short and to the point while providing enough details in the Methods section. A few additional minor comments are listed below.

Minor points:
- Some planetary models provide outputs in magnitudes in various bands, including the Spitzer IRAC bands. To allow an easier comparison with them, could you please provide the Spitzer magnitude derived for the WD in table 1 (possibly along with the lower limits on the magnitude of the companion)

We calculated the IRAC band 2 magnitude for WD 1856, and lower limits on the absolute/apparent magnitude of the companion and added them to the table.

- Figure 2 shows a wide region of the parameter space where nothing happens. It might be worth restricting the figure to a smaller mass area. It could also be interesting to show contours of the brightness in the Spitzer band along with the flux upper limits.

We have added several contours of constant flux to the plot. Now that the new contours are shown, the upper region of the plot is more informative, so we left the y axis range unchanged.

*- I myself do *not* think that an object stops being interesting when it crosses the 13Mjup mass threshold. As argued in the manuscript there are plenty of problems with the current IAU definition of a brown dwarf and I think that how we might want to label this object is actually physically irrelevant. So, while acknowledging that this is probably an editorial decision, I would strongly encourage the authors to remove the planet/brown dwarf limit in Figure 2.*

We certainly agree that WD 1856 b is interesting whether its mass is 12 Mj or 14 Mj! However, since many people do consider the deuterium burning limit a meaningful dividing line, we prefer to leave the distinction in the plot.

Reviewer Reports on the First Revision:

Referees' comments:

Referee #1 (Remarks to the Author):

I am satisfied that the authors have properly addressed all the points I raised.

One final thought, I assume that the authors are able to rule out that the system is not a double white dwarf binary composed of two very similar DC white dwarfs? In this case the period would be double the quoted value (i.e. 2.8 days, with primary and secondary eclipses visible). My guess is that this can be ruled out using the SED and perhaps the eclipse duration would still be too long for two white dwarf sized objects even at this longer period. But it might be something to address in the paper, since having a double white dwarf binary with such an orbital period is not uncommon at all and you would expect very little change in eclipse depth with wavelength if this were the case.

Referee #2 (Remarks to the Author):

This is the revised version of the manuscript. For the most part I was satisfied with the original version - the Spitzer observations nailed it for me. However, the new version addresses the relatively minor issues of the referees. In particular I am satisfied with the response to my comments on Doppler beaming. I suspected this would be difficult, but it is nice to see that they at least looked into the possibility. I also see that they replied adequately (in my view) to questions from the other referees that covered aspects beyond my expertise.

This version improves on the original so I recommend publication.

Artie Hatzes

Referee #3 (Remarks to the Author):

I have read and reviewed the revised manuscript submitted by Vanderburg et al.

I am pleased to say that the authors fully answered to all of my concerns and that I can now recommend the article for publication.

I take the opportunity to thank the authors for what has been my easiest reviewing experience in years.

Author Rebuttals to First Revision:

We thank the referees for looking over the new draft and for their kind words and constructive reports throughout the process. We have responded to the one additional comment and made modifications to the manuscript highlighted in red bold font. We have also made the requested editorial changes. Finally, we fixed a few typographical errors throughout the document.

Referee #1 (Remarks to the Author):

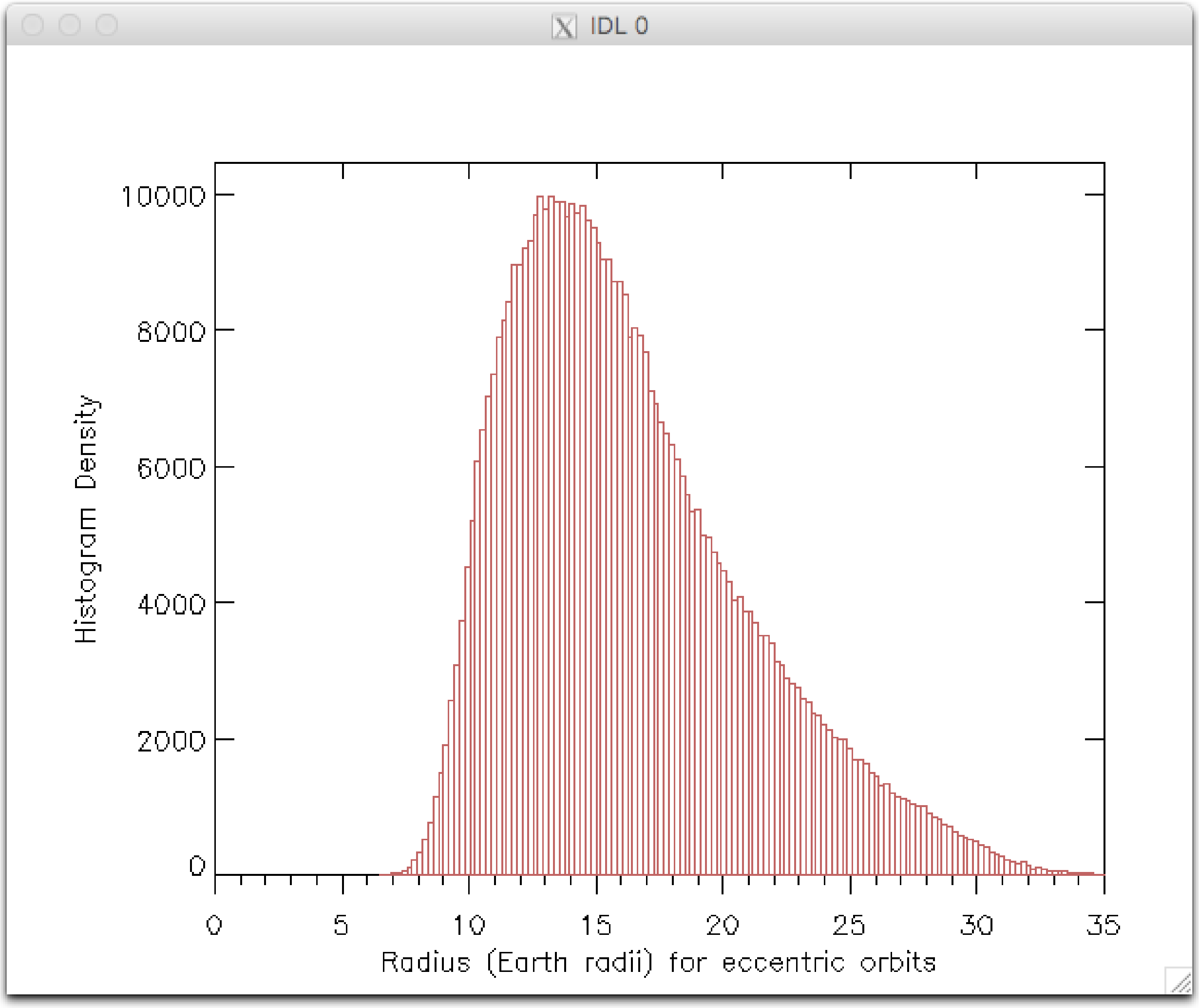
I am satisfied that the authors have properly addressed all the points I raised.

One final thought, I assume that the authors are able to rule out that the system is not a double white dwarf binary composed of two very similar DC white dwarfs? In this case the period would be double the quoted value (i.e. 2.8 days, with primary and secondary eclipses visible). My guess is that this can be ruled out using the SED and perhaps the eclipse duration would still be too long for two white dwarf sized objects even at this longer period. But it might be something to address in the paper, since having a double white dwarf binary with such an orbital period is not uncommon at all and you would expect very little change in eclipse depth with wavelength if this were the case.

Good question -- the simplest way to rule out an equal mass/equal brightness white dwarf/white dwarf binary is to note that both the odd and even transit depths are about 56% deep, while an equal brightness binary can at most produce a 50% transit depth (because the light from the occulting star will still be present during transit). Put slightly differently, for any eclipsing binary, ($D_{Primary} + D_{Secondary}$) ≤ 100%, where D is the eclipse depth -- the stars cannot block more than the total brightness of the system. If WD 1856 were a nearly equal mass EB, the sum of the eclipse depths would be about 112%, so the

true period cannot be 2.8 days.

A similar scenario is that WD 1856 is transited by a much colder (and therefore much fainter) white dwarf in a 1.4 day orbit. This is indeed ruled out by the eclipse duration, which is far too long for a ~1.5 Earth radius object to produce. For circular 1.4 day orbits (assuming the companion white dwarf is much fainter than WD 1856), as we would expect for a white dwarf/white dwarf binary, the transit duration should be at most 3 minutes for an equatorial transit, and shorter for a grazing transit. The observed transit duration is almost 8 minutes by contrast. The detailed shape of the grazing transit also seems to disfavor small companion radii - even when we relax the circular orbit assumption (decoupling the transit duration from the average orbital speed), radii smaller than about 7 Earth radii are excluded: see the following posterior probability distribution:



So unless the companion white dwarf was bloated to the size of Jupiter (which has only been seen for extremely low mass, hot, and luminous white dwarfs, which would dominate the total brightness of the system - see https://arxiv.org/abs/1502.02303), it cannot produce the transit duration we observe.

We have added a few words on these scenarios to the main paper, and some more explanation to the Methods section.