

Stable mass transfer from a substellar object onto an M dwarf

Corresponding Author: Mr Aaron Householder

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #1

(Comments for the Author)

I have reviewed the submitted paper, "Stars stably accreting from substellar objects" and find that it is an exciting and novel study of two M-dwarf systems with potential mass transfer from substellar companions. The authors have identified two systems with unusual short-period signals from ZTF and, from follow-up lightcurve observations, discovered chromaticity consistent with large hot spots. Additional reconnaissance of the systems through follow-up spectroscopy and orbital modeling present an intriguing pair of objects with complementary lines of evidence for accretion from brown dwarf companions. This would be the discovery of two unusual and rare systems, and a noteworthy result satisfying the criteria for publication in Nature Astronomy. Overall, I commend the authors for the well-written paper, which is generally quite clear. However, there are a few areas that could be expanded upon in greater detail. Given the unique and unusual nature of the object properties, I have a few questions and requests for improved clarification and supporting details for the interpretations of the data, which I list below (including line references corresponding to those in the submitted manuscript, where applicable).

In terms of high-level comments regarding the overall manuscript, the levels of constraint upon physical parameters in the two systems are quite different, with ZTF J0440 being more tightly constrained; this would be an important point to make explicitly in the paper. For example, the lines of evidence for ZTF J1444 being a mass transfer M-dwarf-BD system point to the upper limit on X-ray luminosity, transient Balmer features in the optical spectra, and the eclipse of the accretion hot spot in the light curves. However, the physical parameters of J1444 are driven in large part from the low-resolution SPHEREx spectra, which the authors note do not produce phase-resolved spectroscopy. The quality of this spectrum is rather limited and the spectroscopic distance is in tension with the (also somewhat noisy) Gaia parallax. It would be illuminating to include a discussion on how a systematic bias in effective temperature and/or different distance assumption would affect the interpretation of the other diagnostics of the system, e.g., the X-ray luminosity. The transient Balmer features are more consistent with accretion onto a stellar surface than a WD system, but perhaps are not as diagnostic for a neutron star companion (which should be otherwise precluded by the X-ray luminosity). In this case, the geometric argument would be much stronger as it is distance-independent, and the authors could consider framing it as such.

Where additional spectra/photometry data are available, it could further strengthen the results to directly test the observations for consistency with the modeled BD companion properties. For example, would one expect flux contribution in form of excess from the companion at z band or in observed spectra (LRIS, SPHEREx) given the estimated properties of 'unseen' companion? For the given age, evolutionary model, and mass, a luminosity argument could be made; e.g., for a field age BD of 30-40 Jupiter masses, one could reasonably assume such an object has a luminosity of 10^{-5} or 10^{-6} L_{sun} , based on evolutionary models (e.g., Philips et al. 2020). This is only a factor of 50-500x difference from the host accretor, so moderate contrast given instrumental sensitivity limits, and perhaps contributing non-negligibly to the observations. On such a short orbital period, irradiation would also likely be a significant factor.

Lines 82 and 515: The logical flow in the introduction does not quite follow; an M8 spectral type would have a mass closer to 0.085 M_{sun} and effective temperature of ~ 2600 K assuming a field age (e.g., Pecaut & Mamajek 2013). (Which would make the $M \sin i$ value then a bit lower, i.e., even more firmly substellar.) The accretor mass provided here seems to be determined from the MCMC framework performed later on in the paper, but this value is somewhat in tension with the empirical spectral fit, which does not seem to include a spectral type uncertainty. As such, some justification for the adopted prior mass range and discussion of the resulting posterior distribution should be provided in the "Mass and radius constraints" section of the

methods. The effective temperature/spectral type discrepancy is also noted in Fig. 6 (where the M8 spectrum clearly fits best, while a potentially corresponding BT Settl temperature of 2600K is quite different, and seemed to require subsolar metallicity; perhaps the companion is contributing to the red end of the LRIS spectrum?)

Line 107: There is at least one known black widow pulsar (PSR_J0636+5128) with an X-ray luminosity of only $\sim 8 \times 10^{29}$ erg/s (see Koljonen & Linares 2025 for an up-to-date catalog) - it would be worth remarking upon this in the discussion of the X-ray limits for both systems.

Lines 160 / 211 / 680: Brief additional details on the frequency of close m-dwarf BD binaries (e.g., Carmichael 2023 and more recent references) and current observational selection effects would provide valuable context in these areas. In particular, the empirical arguments on timescales and mass transfer lifetimes are very well-reasoned, but would be further strengthened by mentioning known separations of potential precursor systems explicitly, and possibly what characteristic inspiral times are likely under different angular momentum loss scenarios (and/or what future work is needed to inform these questions).

Line 166: The ballistic stream modeling and Coriolis prediction agreement with the observed phase offset provide compelling evidence for J0440's accretion geometry, and the authors provide a rigorous and clear case for this model describing the system well. Given that low-mass stars in this mass regime are fully convective, and likely harbor strong and complex magnetic field geometries, what effect would this have on the stream modeling/phase offset predictions? This may be small given the good agreement, but it would be interesting to understand whether this agreement places any constraints on the magnetic field of the accretor. It could also be of interest for future studies, including J1444. A brief discussion to this effect would provide valuable context, particularly in light of our current understanding of magnetospheric accretion for young low-mass stars.

Line 743: Please comment on the estimated spot filling fraction for the different hot spot size estimates, as well as their uncertainties at given adopted distances.

Line 743: From the LRIS blue-arm spectra, the presence of a Balmer jump near 340nm is often visible even for very low accretion rates for low-mass stars; is this seen in your spectra, and if not, is it because the accretion regime in the direct-impact geometry is too optically thin to produce Balmer continuum emission?

Minor comment: The aspect ratios and layout on Figures 9 and 10 make them very difficult to read, and could be significantly improved for legibility.

(Remarks on code availability)

The data and code used in the paper largely support reproducibility but could be more clearly documented with a README describing the analyses and data products in more detail. Some data products (measured radial velocities, lightcurve models, SPHEREx spectrum, best-fit BT Settl models) are not included. Some of these should likely be included directly in the repository, while others could be linked to other websites (e.g., SPHEREx archive) instead of included directly.

Reviewer #2

(Comments for the Author)

This manuscript presents two unusual ultra short period systems, and if the authors' interpretation is correct, stable mass transfer from a brown dwarf onto an M dwarf, the result would be interesting and worthy of publication. In particular, ZTF J0440+2325 appears to be a potentially important system, and I therefore find the paper promising overall.

At the same time, I do not think the current version fully supports the strength of its main interpretation. My concern is that the manuscript relies too heavily on the classical $q_{\rm crit}$ argument to support the unseen companion is a Roche-lobe filling brown dwarf donor transferring mass stably onto the visible M dwarf. More recent work has shown that mass-transfer stability generally cannot be reduced to a single fixed critical mass ratio, but depends on the donor structure, the treatment of the outer envelope, angular momentum loss, and the degree of conservativeness of the transfer.

The subsequent MCMC analysis is then carried out within that same assumed framework. As a result, the MCMC derived system parameters demonstrate internal consistency within the adopted model, but do not by themselves establish that this is the unique or securely preferred interpretation. I therefore do not think the current paper has yet shown that it has identified a brown dwarf stably transferring mass onto an M dwarf.

In addition, I would like to see a more explicit discussion of whether a second late-type photospheric component has been spectroscopically excluded. For example, could the system instead contain two late-type stars with similar spectral features?

For these reasons, I recommend substantial revision before publication.

Major comments

1. The stability argument based on a single $q_{\rm crit}$ threshold is weak to serve as a decisive BD donor and M dwarf accretor diagnostic.

More recent work has shown that mass transfer stability generally cannot be reduced to a single fixed critical mass ratio.

Woods & Ivanova (2011) argued that simple polytropic approximations can yield stability criteria that are too restrictive. Ivanova's 2014 review (https://link.springer.com/chapter/10.1007/978-3-662-44434-4_8) emphasizes that the stability of Roche-lobe overflow depends on the donor's structural response on different timescales, including adiabatic, thermal equilibrium, superadiabatic, as well as on assumptions about how conservative the transfer is.

This issue becomes even more important for donors with convective envelopes. Pavlovskii & Ivanova (2015) showed that for giant donors with well developed outer convective envelopes, the critical initial mass ratio for stable conservative mass transfer can lie in the range 1.5 to 2.2, much larger than older commonly adopted values. Ge et al. (2015) showed that the effective adiabatic stability boundary varies strongly with evolutionary state. While their results for low mass main sequence donors may be broadly consistent with a relatively low $q_{\rm crit}$ under the assumption of conservative mass transfer (<https://ui.adsabs.harvard.edu/abs/2015ApJ...812...40G/abstract>).

At minimum, the authors should discuss more explicitly whether configurations with $q \sim 1$ can be excluded.

2. Please clarify whether any spectral decomposition or explicit two-component spectral test has been performed.

For ZTF J0440+2325, the red spectrum is said to match an empirical M8 dwarf template well, and the measured RV semi-amplitude of the visible M dwarf is used to argue that the unseen companion is substellar.

I appreciate that a full spectral disentangling analysis may be challenging with the present data. However, some assessment of whether an additional cool photospheric component could be hidden in the spectrum still seems important. If such tests have not been carried out, I think this limitation should be stated more explicitly.

3. ZTF J1444+4820: the evidence is suggestive, but significantly less secure than for ZTF J0440+2325.

The manuscript acknowledges that for J1444 there are no usable red-arm spectra and no phase-resolved RVs, and that the component masses and spectral types cannot yet be dynamically constrained as directly as for J0440. I think the main text still treats the two systems somewhat too symmetrically.

I would encourage the authors to more clearly distinguish between the status of the two objects. J0440 may reasonably be presented as the anchor system. J1444 is better described as a strong candidate.

4. The authors should be clearer about which results follow directly from the data and which depend on modeling.

The major claim of this work relies on additional assumptions: Roche lobe geometry, adopted mass-radius relations, the stability framework, and the hotspot geometry. I think the manuscript should label these steps more explicitly as inference rather than direct measurement.

Minor comments

1. The main paragraphs could be phrased a bit more cautiously. I would be more comfortable with wording such as "strong evidence" for stable brown dwarf to M-dwarf mass transfer, unless the authors can further strengthen the argument discussed above.

2. line 143: I noticed at least one likely copying inconsistency: for ZTF J1444+4820, the X-ray luminosity upper limit is quoted as 7.1×10^{29} erg/s in the main text, but as 7.6×10^{29} erg/s in Table 1 and later in the discussion. Please check which value is correct.

(Remarks on code availability)

Reviewer #3

(Comments for the Author)

The manuscript presents a tantalising case for the existence of a hitherto dismissed mass transfer phase via Roche lobe overflow from a brown dwarf to a main-sequence star, involving direct-impact stream accretion. I found the case compelling and don't see any clear and obvious misguided assumptions. The analysis of the spectra made use of state-of-the-art tools following best practice, and assumptions and methods are (mostly) presented and discussed to the appropriate level. The statistical analysis deployed to consider uncertainties is appropriate, as are the conclusions (with some caveats to be added more prominently as detailed below). The references are appropriate. Overall the manuscript is well written, appropriately concise, laying out the case in a logical order, and providing an appropriate abstract, introduction and conclusions.

The lack of red-arm Keck spectra for one of the two sources leaves a level of doubt on the claimed mass and spectral type, and it might be argued that it would be prudent to await repeated observations for a more firm confirmation of the interpretation put forward. In view of the well-argued case my view is however that the paper is ready to be aired, and the unusual nature of the two systems warrants publication as a Nature Astronomy.

There are a number of minor or moderate issues I would expect the authors to address before the manuscript could be regarded as complete.

The authors talk about the transient Balmer features at the hotspot phases being consistent with a buried accretion shock,

and there certainly does seem to be obvious extra Balmer absorption in the blue lines. However the paper appears to avoid discussing H α ; a note on H α behavior and expected or unexpected phase dependence should be added. The authors should consider adding a plot of H α , beta, gamma, delta, epsilon-EW vs orbital phase.

Line 78/79 - PyHammer is referenced as Kesseli et al 2017, but the current version is detailed in Roulston et al 2020.

Line 380 - Table 1: Where the inclination is not known quote stellar mass as lower limit. Quote Gaia parallax in table.

Line 425 - Can you expand on the statistical incidence of sources such as 1444+4820 and 0440+2325, in view of the fact that they are the only two that stood out among the 1.3 billion sources you started from? How many other sources would you say are variable? Can you quantify by how much these two sources stand out? Is the fact that you found only two sources in this large sample not evidence that the sources are indeed rare?

Line 496 – Discuss the significance of and justification for adding a jitter term.

Line 509 – You state “... refined the orbital phasing by shifting the phase of the best-fit sinusoidal model and the observed radial velocities”. Why do you need to shift both the model and the observations?

Line 532 – Equation 1: Quote a mass range this approximation is valid for, and provide uncertainties.

Line 557 – The honest statement about error bars is appreciated, but please add a corresponding note to the table caption.

Line 830ff – Figures 4 and 5: is the compact Gaussian spot and elongated spot model swapped in Fig 5? It's ok in Fig 4, but the mismatch appears to be the other way round in Fig 5.

Line 585 – Make the caveat about the hot spot models more prominent in the first part of the paper.

Line 630 – $T_{\text{eff}} = 3000\text{K}$ is on the high side for an M8 star. Discuss.

Line 653 – In the absence of synchronous rotation the hot spot would presumably have a very different appearance. Discuss.

Line 691 – Justify the claimed time period with stable mass transfer (10^7 - 10^8 years).

Line 703 – Confirm that the radii obtained here are consistent with the claimed mass of the star, and what mass radius relation is used.

Line 706 – Comment on the impact of extinction, if any.

Line 807 – A slightly larger distance would bring those X-ray upper limits within range. As a note of caution do mention that the Gaia parallax if taken at face value would do this for J1444.

(Remarks on code availability)

I have not run the code, but from a quick review I see that there are two python codes with a reasonably straightforward to follow structure, albeit without any comments. The data files provided are also organised in a logical order. There is no readme file.

I recommend adding a minimum set of comments in the python code and a readme file.

Version 1:

Reviewer comments:

Reviewer #1

(Comments for the Author)

The authors have revised the manuscript following comments from the referees, and I would like to thank the authors for their careful consideration of various points raised during the refereeing process. In particular, the treatment of stellar parameters is significantly improved, and the inclusion of new LRIS spectra provides very interesting and useful context for the nature of the second candidate system. The revised and new analyses in 'Spectral Modeling and Distances' section, among other sections in the overall Methods, are thoroughly described, and the statistical analyses and treatments of uncertainties are carefully discussed. From the new data and interpretations, I have a few follow-up questions and some additional minor comments, below.

My most substantial questions involve the surprisingly complex explanation of J1444 as a spectrally-visible tertiary that is neither the donor nor accretor in the observed mass transfer system. Does the lightcurve geometry strictly enforce near edge-on inclination for the system? What are the potential degeneracies across fit solutions/impact of more complex spot prescriptions? Specifically, I am wondering whether a high latitude accretion hot spot + grazing eclipse (and thus inclined

orbit) could be consistent with the upper limit on RV variation observed in the multi-epoch LRIS spectra. To place more stringent limits on the RV variation, perhaps an instrument with higher spectral resolution than LRIS is necessary, which could be posited as future work.

In such a scenario with only two bodies, could the M4 object be the accretor with a single unseen high-density companion? Of course, this is only plausible if it can satisfy the Roche lobe geometry constraints and the lightcurve eclipse properties. From a density standpoint, it seems that this would require the donor to be a remnant substellar companion if mass transfer has already been occurring for a long period of time, wherein the companion is reduced to mostly electron degenerate material of very high density (a much older version of J0440?). The manuscript would benefit from a brief quantitative discussion of this scenario's viability (or lack thereof!), based on the available lines of observational evidence.

Minor comments:

The abstract would benefit from brief clarification on what is meant by 'candidate' for J1444 (e.g., 'strong mass transfer candidate system').

The authors may wish to consider the nomenclature for the object, as 'molar' (possibly a portmanteau of m-dwarf and polar, given that magnetically channeled accretion is not implied?) could lead to some confusion, and some additional explanation/justification would benefit the new terminology.

Line 089: Please clarify that this is Balmer line emission, as opposed to Balmer continuum.

Line 099-103: The introductory description of black widow pulsars would be beneficial to the reader earlier when first referenced.

Line 176: It would help to remind the reader that J1406+1222 is a candidate black widow pulsar system (to help distinguish its similar naming from the candidates presented in this work). On a related note, in Line 978, some clarification around the language describing J1406 as a literature comparison object would be helpful. It sounds here as though the authors are reconsidering its interpretation as a black widow pulsar candidate and instead that it may be another substellar mass transfer candidate, but the lack of optical eclipses prevent a definitive classification.

(Remarks on code availability)

The updates to the Github repository have added helpful descriptions, context, and data products that significantly improve reproducibility.

Reviewer #2

(Comments for the Author)

I appreciate the substantial revisions made by the authors. The revised manuscript is more careful in presentation. In particular:

(a) the original reliance on a q_{crit} argument has been appropriately softened and the discussion of mass transfer stability acknowledges the relevant caveats. (b) ZTF J1444 has been downgraded to a candidate. (c) the new RV anchored hot spot phase argument provides a more convincing basis for identifying the likely donor and accretor in ZTF J0440.

Overall, I believe the manuscript is substantially improved. I have only one remaining scientific reservation.

The manuscript now argues that the mass transfer is probably dynamically stable because an unstable mass transfer phase would be short that observing such a system would be unlikely. While I agree that this is a reasonable argument, I am not yet fully convinced that it is sufficient to exclude alternative evolutionary scenarios. The inference depends on the poorly known progenitor population, observational selection effects, and the duration of the unstable mass transfer phase. As a result, I view the lifetime argument as suggestive. Whether the remaining level of uncertainty is compatible with the strength of the central claim is ultimately a matter for the Editor to assess.

Minor comment:

In Figure 3, the offset is labeled as 18.3 degree, but quoted as 18.4 degree in the caption and main text. Please make these values consistent.

(Remarks on code availability)

Reviewer #3

(Comments for the Author)

I am content with the authors' responses to the issues raised in my first report. The main new addition to the paper is the interpretation of the unexpected spectral type M4 found in ZTF J1444+4820. While I don't find an obvious flaw, the new line of arguments is long enough to make the system a much less convincing example of a mass-transferring brown dwarf. The language used in the manuscript presents this in a careful and measured way, so I have no objections to the paper to remain essentially as is.

I have no substantive comments to add, just two minor items:

The consistent use of “the M4” to denote the dominant stellar component in ZTF J1444+4820 is odd. I would expect M4 to be followed by a substantive such “M4 star”, “M4 type”, “M4 tertiary” etc.

I would encourage the authors to reconsider the term “molars”. A claimed class with only one member and a second somewhat dubious member is arguably too small to yet warrant a name, and while “molar” is memorable and perhaps funny the parallel with polars is tenuous, and using m instead of p in the word is confusing at best.

(Remarks on code availability)

Version 2:

Reviewer comments:

Reviewer #1

(Comments for the Author)

I appreciate the detailed response from the authors, which satisfies all of my major and minor questions and comments. I recommend that the manuscript be accepted for publication and congratulate the authors on a very interesting discovery.

(Remarks on code availability)

Open Access This Peer Review File is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

In cases where reviewers are anonymous, credit should be given to 'Anonymous Referee' and the source.

The images or other third party material in this Peer Review File are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>

We thank the editor and the three referees for their careful reading of the manuscript and their constructive comments. Below we address each point in turn. Referee comments are shown in *blue italic text*.

Since the initial submission, we obtained new phase-resolved Keck/LRIS red-arm spectra of ZTF J1444+4820. These spectra show M4-type molecular absorption features with no radial velocity modulation at the 67-minute orbital period. Because of the lack of radial velocity modulation and because an M4 dwarf cannot fit within the Roche lobe at a 67-minute orbital period, we conclude that the M4 is a third body in a wider orbit, making ZTF J1444+4820 a hierarchical triple. The inner binary still shows all of the same accretion signatures as ZTF J0440+2325 and is most likely a low-mass mass-transferring system, but the M4 tertiary prevents us from pinning down the exact component types of the inner binary (e.g., whether it hosts a brown dwarf transferring onto an M dwarf, an M dwarf transferring onto a brown dwarf, or two brown dwarfs). Following the suggestions of multiple referees, we have downgraded ZTF J1444+4820 to a strong candidate and now present ZTF J0440+2325 as the confirmed anchor system.

Reviewer 1

In terms of high-level comments regarding the overall manuscript, the levels of constraint upon physical parameters in the two systems are quite different, with ZTF J0440 being more tightly constrained; this would be an important point to make explicitly in the paper. For example, the lines of evidence for ZTF J1444 being a mass transfer M-dwarf-BD system point to the upper limit on X-ray luminosity, transient Balmer features in the optical spectra, and the eclipse of the accretion hot spot in the light curves. However, the physical parameters of J1444 are driven in large part from the low-resolution SPHEREx spectra, which the authors note do not produce phase-resolved spectroscopy. The quality of this spectrum is rather limited and the spectroscopic distance is in tension with the (also somewhat noisy) Gaia parallax. It would be illuminating to include a discussion on how a systematic bias in effective temperature and/or different distance assumption would affect the interpretation of the other diagnostics of the system, e.g., the X-ray luminosity. The transient Balmer features are more consistent with accretion onto a stellar surface than a WD system, but perhaps are not as diagnostic for a neutron star companion (which should be otherwise precluded by the X-ray luminosity). In this case, the geometric argument would be much stronger as it is distance-independent, and the authors could consider framing it as such.

We agree that the constraints on ZTF J1444+4820 are weaker than those on ZTF J0440+2325, and we thank the referee for this helpful framing. Since submission, we obtained new Keck/LRIS red-arm spectra of ZTF J1444+4820 that show an M4 third body in a wider orbit (see above), so we have removed the SPHEREx-based characterization entirely and now adopt the geometric distance from the Gaia parallax (Bailer-Jones et al. 2021, 823 pc with a 574–1304 pc range), rather than inverting the low-significance parallax. Following this referee’s suggestion and those of Reviewer 2, we have downgraded ZTF J1444+4820 to a strong candidate and now treat ZTF J0440+2325 as the anchor system throughout. As the referee suggests, we now emphasize the distance-independent geometric arguments (the optical eclipses rule out a black widow pulsar, and the ballistic stream geometry predicts

the hot-spot phase offset) as the strongest lines of evidence. We also discuss how the X-ray upper limit for ZTF J1444+4820 at this distance ($L_X < 5.7 \times 10^{30} \text{ erg s}^{-1}$) falls within the range of some black widow pulsars, but note that the eclipses independently rule out this interpretation. The spectra also show no cyclotron humps or other features characteristic of polars or white dwarf accretors, as discussed in the revised manuscript.

Where additional spectra/photometry data are available, it could further strengthen the results to directly test the observations for consistency with the modeled BD companion properties. For example, would one expect flux contribution in form of excess from the companion at z band or in observed spectra (LRIS, SPHEREx) given the estimated properties of ‘unseen’ companion? For the given age, evolutionary model, and mass, a luminosity argument could be made; e.g., for a field age BD of 30–40 Jupiter masses, one could reasonably assume such an object has a luminosity of 10^{-5} or $10^{-6} L_\odot$, based on evolutionary models (e.g., Philips et al. 2020). This is only a factor of 50–500x difference from the host accretor, so moderate contrast given instrumental sensitivity limits, and perhaps contributing non-negligibly to the observations. On such a short orbital period, irradiation would also likely be a significant factor.

We performed this analysis for both systems. For ZTF J0440+2325, we fit BT-Settl atmosphere models to the LRIS red-arm spectrum and compared the residuals to the expected flux of the brown dwarf donor at 1000 K and 1500 K, scaled by the MCMC-derived radius at each fitted distance. The donor contributes a maximum of $\sim 0.5\%$ of the total flux at 1000 K and $\sim 2\%$ at 1500 K, both below the model-data residuals of the BT-Settl fits. For ZTF J1444+4820, we repeated the analysis with the M4 third body as the dominant source and find that the candidate inner M8 accretor contributes flux comparable to or smaller than the fit residuals. In both systems, the predicted companion flux is consistent with being present but buried beneath the model-data discrepancies. We have added a discussion of this analysis to the Methods.

Lines 82 and 515: The logical flow in the introduction does not quite follow; an M8 spectral type would have a mass closer to $0.085 M_\odot$ and effective temperature of $\sim 2600 \text{ K}$ assuming a field age (e.g., Pecaut & Mamajek 2013). (Which would make the $M \sin i$ value then a bit lower, i.e., even more firmly substellar.) The accretor mass provided here seems to be determined from the MCMC framework performed later on in the paper, but this value is somewhat in tension with the empirical spectral fit, which does not seem to include a spectral type uncertainty. As such, some justification for the adopted prior mass range and discussion of the resulting posterior distribution should be provided in the “Mass and radius constraints” section of the methods. The effective temperature/spectral type discrepancy is also noted in Fig. 6 (where the M8 spectrum clearly fits best, while a potentially corresponding BT Settl temperature of 2600 K is quite different, and seemed to require subsolar metallicity; perhaps the companion is contributing to the red end of the LRIS spectrum?)

We thank the referee for this careful observation. The original MCMC prior on the accretor mass ($0.075\text{--}0.13 M_\odot$) did not incorporate information from the spectral fits, so the posterior was pulled toward higher masses where the only constraint was that the accretor not fill its Roche lobe. We have now tightened the prior range to $0.075\text{--}0.09 M_\odot$ for ZTF J0440+2325, informed by the empirical mass range for M7–M8 dwarfs (Pecaut & Mamajek 2013), and

the posterior is now more consistent with an M8 spectral type. It is unclear why the M8 PyHammer template fits better than the corresponding BT-Settl model, but we note that these are entirely different types of templates: PyHammer provides empirical observed stellar spectra, while BT-Settl provides theoretical atmosphere models. As discussed above, the brown dwarf companion is unlikely to contribute significant flux in the red, so we do not think the companion is responsible for the discrepancy. We have also added a caveat in the Methods noting that the BT-Settl and PyHammer models and standard mass- T_{eff} calibrations were developed for isolated field dwarfs and do not account for the effects of active accretion, so how well they apply to systems like ZTF J0440+2325 remains an open question.

Line 107: There is at least one known black widow pulsar (PSR_J0636+5128) with an X-ray luminosity of only $\sim 8 \times 10^{29} \text{ erg s}^{-1}$ (see Koljonen & Linares 2025 for an up-to-date catalog) — it would be worth remarking upon this in the discussion of the X-ray limits for both systems.

We have added this reference and a note acknowledging that PSR J0636+5128 has a comparable X-ray luminosity. We also now emphasize that the X-ray limits are one of several independent lines of evidence that rule out a compact-object companion.

Lines 160 / 211 / 680: Brief additional details on the frequency of close m-dwarf BD binaries (e.g., Carmichael 2023 and more recent references) and current observational selection effects would provide valuable context in these areas. In particular, the empirical arguments on timescales and mass transfer lifetimes are very well-reasoned, but would be further strengthened by mentioning known separations of potential precursor systems explicitly, and possibly what characteristic inspiral times are likely under different angular momentum loss scenarios (and/or what future work is needed to inform these questions).

We have expanded the discussion to note that close brown dwarf companions to M dwarfs are rare (roughly a dozen known transiting systems, Carmichael 2023) and that magnetic braking remains poorly calibrated for fully convective stars (Matt et al. 2015). Nonetheless, under standard magnetic braking prescriptions, the shortest-period known precursor systems ($P \sim 2 \text{ hr}$) are expected to reach contact within a Hubble time, while many of the longer-period detached systems will not. Binary population synthesis modeling (e.g., with MESA) and a larger sample of systems will ultimately be needed to further constrain the formation channel, inspiral timescales, and expected occurrence rates.

Line 166: The ballistic stream modeling and Coriolis prediction agreement with the observed phase offset provide compelling evidence for J0440’s accretion geometry, and the authors provide a rigorous and clear case for this model describing the system well. Given that low-mass stars in this mass regime are fully convective, and likely harbor strong and complex magnetic field geometries, what effect would this have on the stream modeling/phase offset predictions? This may be small given the good agreement, but it would be interesting to understand whether this agreement places any constraints on the magnetic field of the accretor. It could also be of interest for future studies, including J1444. A brief discussion to this effect would provide valuable context, particularly in light of our current understanding of magnetospheric accretion for young low-mass stars.

The referee raises a great point. The agreement between the ballistic prediction and the

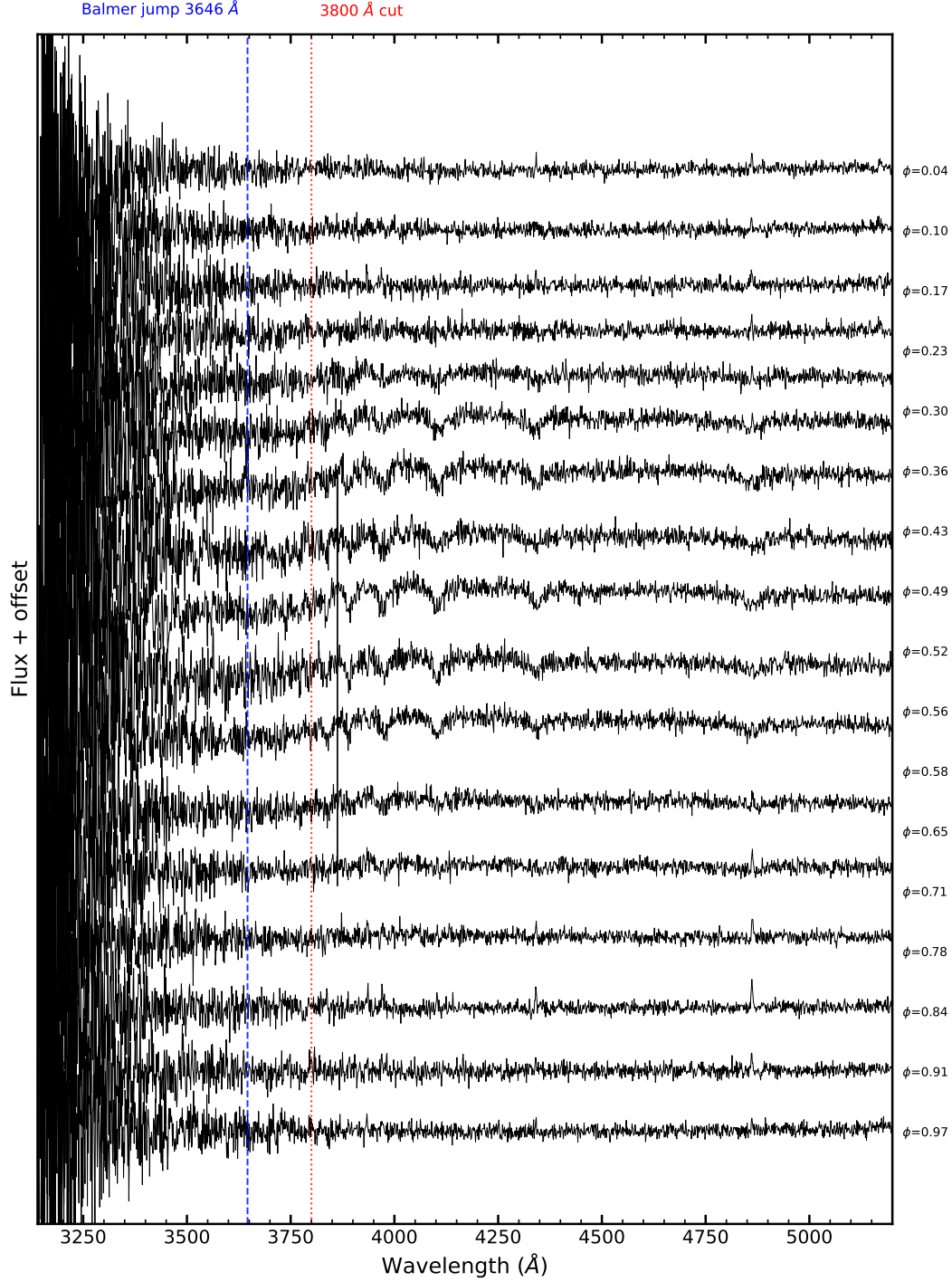
observations of ZTF J0440+2325 suggests the magnetic field does not significantly deflect the accretion stream. We also note there are a handful of unpublished results in similar systems that seem to indicate the stream is not significantly deflected by magnetic fields (i.e., the ballistic stream model seems to predict the hotspot location well). Perhaps this is not surprising, as M dwarf surface fields ($\sim 2\text{--}4$ kG, Reiners et al. 2009) are orders of magnitude weaker than the $\sim 10\text{--}100$ MG fields in polars (Cropper 1990, Ferrario et al. 2015). Ultimately, full MHD modeling would be needed to quantify this further but is beyond the scope of this work. We have added a paragraph to the Methods discussing this.

Line 743: Please comment on the estimated spot filling fraction for the different hot spot size estimates, as well as their uncertainties at given adopted distances.

We have added the hot-spot filling fractions and their uncertainties to the revised manuscript. We have also improved the hot-spot spectral fits themselves. We now subtract the scaled M dwarf templates (the M8 in ZTF J0440+2325 and the M4 tertiary in ZTF J1444+4820) from the blue-arm spectra before fitting the BT-Settl models, so that the fits are performed on the hot-spot flux alone.

Line 743: From the LRIS blue-arm spectra, the presence of a Balmer jump near 340 nm is often visible even for very low accretion rates for low-mass stars; is this seen in your spectra, and if not, is it because the accretion regime in the direct-impact geometry is too optically thin to produce Balmer continuum emission?

It is unclear whether we would expect to observe a Balmer jump in these systems. Our LRIS blue-arm spectra extend down to ~ 3140 Å, but we cut out wavelengths below 3800 Å from the analysis because of the low S/N at the blue end. We see no obvious Balmer jump in the current data (see the figure below, showing the ZTF J0440+2325 blue-arm spectra over the full wavelength coverage). We leave this as an interesting diagnostic to pursue with *HST* or ground-based UV-sensitive spectroscopy in the future.



Minor comment: The aspect ratios and layout on Figures 9 and 10 make them very difficult to read, and could be significantly improved for legibility.

We have redesigned Extended Data Figures 9 and 10 with waterfall-style layouts for improved legibility.

Remarks on code availability: The data and code used in the paper largely support repro-

ducibility but could be more clearly documented with a README describing the analyses and data products in more detail. Some data products (measured radial velocities, lightcurve models, SPHEREx spectrum, best-fit BT Settl models) are not included. Some of these should likely be included directly in the repository, while others could be linked to other websites (e.g., SPHEREx archive) instead of included directly.

We have updated the repository with a README, measured radial velocities for both systems, the LRIS red-arm spectra for ZTF J1444+4820, and comments in the analysis code. The BT-Settl models are linked to the SVO Theoretical Spectra Server and the PyHammer templates to the public PyHammer repository.

Reviewer 2

This manuscript presents two unusual ultra short period systems, and if the authors' interpretation is correct, stable mass transfer from a brown dwarf onto an M dwarf, the result would be interesting and worthy of publication. In particular, ZTF J0440+2325 appears to be a potentially important system, and I therefore find the paper promising overall.

At the same time, I do not think the current version fully supports the strength of its main interpretation. My concern is that the manuscript relies too heavily on the classical q_{crit} argument to support the unseen companion is a Roche-lobe filling brown dwarf donor transferring mass stably onto the visible M dwarf. More recent work has shown that mass-transfer stability generally cannot be reduced to a single fixed critical mass ratio, but depends on the donor structure, the treatment of the outer envelope, angular momentum loss, and the degree of conservativeness of the transfer.

The subsequent MCMC analysis is then carried out within that same assumed framework. As a result, the MCMC derived system parameters demonstrate internal consistency within the adopted model, but do not by themselves establish that this is the unique or securely preferred interpretation. I therefore do not think the current paper has yet shown that it has identified a brown dwarf stably transferring mass onto an M dwarf.

In addition, I would like to see a more explicit discussion of whether a second late-type photospheric component has been spectroscopically excluded. For example, could the system instead contain two late-type stars with similar spectral features?

For these reasons, I recommend substantial revision before publication.

Major comments

- 1. The stability argument based on a single q_{crit} threshold is weak to serve as a decisive BD donor and M dwarf accretor diagnostic.*

More recent work has shown that mass transfer stability generally cannot be reduced to a single fixed critical mass ratio. Woods & Ivanova (2011) argued that simple polytropic approximations can yield stability criteria that are too restrictive. Ivanova's 2014 review (https://link.springer.com/chapter/10.1007/978-3-662-44434-4_8) emphasizes that the stability of Roche-lobe overflow depends on the donor's structural response on different timescales, including adiabatic, thermal equilibrium, superadiabatic, as well as on assumptions about how conservative the transfer is.

This issue becomes even more important for donors with convective envelopes. Pavlovskii & Ivanova (2015) showed that for giant donors with well developed outer convective envelopes, the critical initial mass ratio for stable conservative mass transfer can lie in the range 1.5 to 2.2, much larger than older commonly adopted values. Ge et al. (2015) showed that the effective adiabatic stability boundary varies strongly with evolutionary state. While their results for low mass main sequence donors may be broadly consistent with a relatively low q_{crit} under the assumption of conservative mass transfer (<https://ui.adsabs.harvard.edu/abs/2015ApJ...812...40G/abstract>) At minimum, the authors should discuss more explicitly whether configurations with $q \sim 1$ can be excluded.

The referee raises an excellent point, and we appreciate the thoughtful critique. We have performed detailed MESA binary evolution calculations that do support stable mass transfer at mass ratios $q < 2/3$ for systems like these, and a detailed presentation of this MESA grid is currently in preparation for a separate publication. However, as the referee correctly notes, mass-transfer stability depends on assumptions about angular momentum loss, the degree of conservativeness, and the donor structure, and we cannot definitively rule out $q \sim 1$ based on those assumptions alone. We have therefore relaxed the $q_{\text{crit}} < 2/3$ claim in the revised manuscript.

In particular, ZTF J1444+4820, which previously did much of the heavy lifting for this claim, has been downgraded to a strong candidate (see above). For ZTF J0440+2325, several lines of evidence point to the M dwarf being the accretor. The mass transfer itself is clear from the extreme chromatic variability, the transient Balmer absorption near peak brightness, and the hot-spot eclipses. The M dwarf is identified from the M8 spectral match. The radial velocities, which give $K = 24.7 \pm 2.6 \text{ km s}^{-1}$ and $M \sin i = 9.7 \pm 1.2 M_{\text{Jup}}$, strongly disfavor $q \sim 1$.

The remaining question is which component is the donor and which is the accretor. The ballistic stream model predicts that the accretion stream impacts the accretor on its day side via direct impact. If the brown dwarf is the donor and the M dwarf is the accretor, the hot spot is on the M dwarf and peak optical brightness occurs near superior conjunction of the M dwarf, matching the RV-anchored phase of ZTF J0440+2325. If instead the M dwarf is the donor and the brown dwarf is the accretor, the ballistic stream impacts the brown dwarf on its day side for a grid of brown dwarf masses (0.04–0.07 $M_{\odot}$, radii from MESA evolutionary tracks), and peak optical brightness would occur near superior conjunction of the brown dwarf. This is inconsistent with the observations of ZTF J0440+2325, where peak brightness occurs at superior conjunction of the M dwarf as determined by the radial velocities. The second scenario is therefore strongly disfavored for ZTF J0440+2325 by the ballistic stream geometry combined with the RV phase anchoring. For ZTF J1444+4820, the radial velocities show no detectable modulation that could anchor the absolute orbital phase of the inner binary, so this test cannot yet be applied. We discuss both scenarios in the main text and in a new paragraph in the Methods.

2. Please clarify whether any spectral decomposition or explicit two-component spectral test has been performed.

For ZTF J0440+2325, the red spectrum is said to match an empirical M8 dwarf template

well, and the measured RV semi-amplitude of the visible M dwarf is used to argue that the unseen companion is substellar.

I appreciate that a full spectral disentangling analysis may be challenging with the present data. However, some assessment of whether an additional cool photospheric component could be hidden in the spectrum still seems important. If such tests have not been carried out, I think this limitation should be stated more explicitly.

We have performed two-component spectral tests for both systems. For ZTF J0440+2325, we fit BT-Settl atmosphere models to the red-arm LRIS spectrum and compared the residuals to the expected flux of the brown dwarf donor at 1000 K and 1500 K, scaled by the MCMC-derived radius at the fitted distance. The predicted donor flux falls below the model-data residuals of the BT-Settl fits at all wavelengths. For ZTF J1444+4820, we repeated the analysis with the M4 third body as the dominant source and again find that the predicted flux of the inner accretor falls below the fit residuals. In both systems, a second photospheric component cannot be identified in the current data, and the predicted flux contributions are small enough that they would be buried beneath the model-data residuals. Full spectral disentangling (e.g., Zucker & Mazeh 1994, Hadrava 1995, Seeburger et al. 2024) is also likely not feasible for ZTF J0440+2325: the RV semi-amplitude ($K = 24.7 \text{ km s}^{-1}$) is much smaller than the velocity resolution of both LRIS and ESI, placing the system in the regime where disentangling becomes ill-posed (Seeburger et al. 2024). Higher-resolution or near-infrared spectroscopy would allow a more sensitive search. We have added a note of this in the Methods.

3. ZTF J1444+4820: the evidence is suggestive, but significantly less secure than for ZTF J0440+2325.

The manuscript acknowledges that for J1444 there are no usable red-arm spectra and no phase-resolved RVs, and that the component masses and spectral types cannot yet be dynamically constrained as directly as for J0440. I think the main text still treats the two systems somewhat too symmetrically.

I would encourage the authors to more clearly distinguish between the status of the two objects. J0440 may reasonably be presented as the anchor system. J1444 is better described as a strong candidate.

We thank the referee for this helpful framing, which we believe has strengthened the manuscript. The revised manuscript now treats ZTF J0440+2325 as the anchor system throughout, and ZTF J1444+4820 as a strong candidate. The inner binary of ZTF J1444+4820 shows the same accretion signatures as ZTF J0440+2325 (eclipses, extreme chromatic variability, transient Balmer absorption), and although the `lcurve` fits did not converge, they yielded two families of solutions, bounded at roughly $q \approx 0.25$ and $q \approx 3$, pointing to similar component masses. These observations leave little room for interpretations other than a low-mass mass-transferring system. However, the presence of the M4 third body makes it difficult to pin down the exact component types of the inner binary (e.g., whether it is a brown dwarf transferring onto an M dwarf, an M dwarf transferring onto a brown dwarf, or two brown dwarfs) and the direction of mass transfer.

4. The authors should be clearer about which results follow directly from the data and which

depend on modeling.

The major claim of this work relies on additional assumptions: Roche lobe geometry, adopted mass-radius relations, the stability framework, and the hotspot geometry. I think the manuscript should label these steps more explicitly as inference rather than direct measurement.

We agree and have revised the manuscript to more explicitly separate what follows from the observations and what depends on modeling. For ZTF J0440+2325, the observations alone establish: (1) extreme chromatic variability and transient Balmer absorption from an accretion hot spot, (2) an M8 spectral type for the visible component, (3) a substellar companion ($M \sin i = 9.7 \pm 1.2 M_{\text{Jup}}$) from the radial velocities, (4) peak brightness near superior conjunction of the M dwarf ($\phi \approx 0.44\text{--}0.48$), and (5) no X-ray detection, disfavoring a compact object accretor. As we now state in the revised manuscript, the observations alone cannot distinguish whether the brightness peak originates from a hot spot on the day side of the M dwarf or from the night side of the brown dwarf, and the ballistic stream model is what breaks this degeneracy.

The component masses, radii, and mass ratio from the MCMC framework are model-dependent inferences that assume Roche-lobe geometry and empirical mass-radius relations. The stability of mass transfer depends on the classical q_{crit} framework, which we now caveat more explicitly.

The revised main text now presents these observational constraints first before introducing the model-dependent inferences. For ZTF J1444+4820, we have softened the language to “strong candidate” throughout and added a caveat to the Table 1 caption noting that the MCMC parameters are particularly model-dependent given the unresolved M4 third body.

Minor comments

1. The main paragraphs could be phrased a bit more cautiously. I would be more comfortable with wording such as “strong evidence” for stable brown dwarf to M-dwarf mass transfer, unless the authors can further strengthen the argument discussed above.

We have adopted the more cautious “strong evidence” wording suggested by the referee. The revised main text now states that the ballistic stream geometry, together with the Coriolis offset of peak brightness, provides strong evidence that the brown dwarf is transferring mass onto the M dwarf in ZTF J0440+2325, and we describe ZTF J1444+4820 as a “strong candidate” throughout.

2. line 143: I noticed at least one likely copying inconsistency: for ZTF J1444+4820, the X-ray luminosity upper limit is quoted as 7.1×10^{29} erg/s in the main text, but as 7.6×10^{29} erg/s in Table 1 and later in the discussion. Please check which value is correct.

We appreciate the careful reading. This was indeed a copying error introduced when updating the distance. The values have been corrected in the revised manuscript.

Reviewer 3

The authors talk about the transient Balmer features at the hotspot phases being consistent with a buried accretion shock, and there certainly does seem to be obvious extra Balmer

absorption in the blue lines. However the paper appears to avoid discussing $H\alpha$; a note on $H\alpha$ behavior and expected or unexpected phase dependence should be added. The authors should consider adding a plot of $H\alpha$, β , γ , δ , ϵ -EW vs orbital phase.

We thank the referee for this suggestion. We have added a discussion of the $H\alpha$ behavior and an equivalent width vs. orbital phase panel to Extended Data Figure 9 for ZTF J0440+2325, showing $H\alpha$ and $H\beta$ equivalent widths measured by fitting Gaussian profiles to the continuum-normalized spectra at each orbital phase. $H\alpha$ is in emission at all phases, but the emission strength weakens significantly near peak optical brightness as the hot spot rotates into view. For ZTF J1444+4820, the blue-arm spectra are too low S/N in most phases for reliable equivalent width measurements as a function of phase, but the same transient Balmer features near peak brightness are visible in the plot.

Line 78/79 — PyHammer is referenced as Kesseli et al 2017, but the current version is detailed in Roulston et al 2020.

We have added the Roulston et al. (2020) citation alongside the original Kesseli et al. (2017) reference wherever **PyHammer** is cited in the manuscript.

Line 380 — Table 1: Where the inclination is not known quote stellar mass as lower limit. Quote Gaia parallax in table.

We have added the geometric distance for ZTF J1444+4820 (Bailer-Jones et al. 2021, from the Gaia EDR3 parallax) to Table 1. We have also added the minimum companion mass $M_{\text{don}} \sin i$ for ZTF J0440+2325 from the radial velocities. We note that the component masses reported in Table 1 are inferred from the Roche-geometry MCMC framework, not from radial velocities, so they are not technically lower limits set by the unknown inclination.

Line 425 — Can you expand on the statistical incidence of sources such as 1444+4820 and 0440+2325, in view of the fact that they are the only two that stood out among the 1.3 billion sources you started from? How many other sources would you say are variable? Can you quantify by how much these two sources stand out? Is the fact that you found only two sources in this large sample not evidence that the sources are indeed rare?

These are not the only two such systems found in ZTF, with several more candidates bringing the total to roughly 8–9 systems. However, the occurrence rate is difficult to constrain given the uncertain distances and the incomplete selection function of ZTF, and will need much more work to pin down. However, the discovery of ZTF J0440+2325 at ~ 100 –150 pc suggests they are not vanishingly rare, and we expect LSST to identify many more. We leave a detailed occurrence rate calculation to future work.

Line 496 — Discuss the significance of and justification for adding a jitter term.

Without the jitter term, the fit has too many outliers relative to statistical expectations. The source of the excess RV scatter is unclear but is common in practice, and could be due to astrophysical effects such as stellar activity from the spot contamination or an unaccounted-for instrumental uncertainty. We have added a brief note of possible reasons for the additional scatter in the revised manuscript.

Line 509 — You state “...refined the orbital phasing by shifting the phase of the best-fit

sinusoidal model and the observed radial velocities”. Why do you need to shift both the model and the observations?

We apologize for the lack of clarity. The same phase offset is applied to both the model and the data. Because the ZTF photometric ephemeris is based on the optical variability and it is not known a priori which component drives the variability, the RVs are needed to anchor the absolute orbital phase of the M dwarf. We apply a single offset to both the model and the observed radial velocities so that maximum RV occurs at $\phi = 0.25$ (superior conjunction at $\phi = 0.5$). We have clarified this in the revised manuscript.

Line 532 — Equation 1: Quote a mass range this approximation is valid for, and provide uncertainties.

This relation is appropriate for zero-age main-sequence stars with $0.08 \lesssim M/M_{\odot} \lesssim 0.6$. Fitting the same functional form to both the Baraffe et al. (2015) and Nelson (2018) models over different mass ranges yields approximate uncertainties of ± 0.05 on the coefficient and ± 0.02 on the exponent, driven largely by the choice of mass range and differences between model grids. We have added this information to the revised manuscript.

Line 557 — The honest statement about error bars is appreciated, but please add a corresponding note to the table caption.

A note has been added to the Table 1 caption.

Line 830ff — Figures 4 and 5: is the compact Gaussian spot and elongated spot model swapped in Fig 5? It’s ok in Fig 4, but the mismatch appears to be the other way round in Fig 5.

Thank you for catching this. The two spot models were indeed swapped in Extended Data Figure 5. We have corrected the figure so that the compact Gaussian spot and equatorial spot models are now labeled correctly for both Extended Data Figures 4 and 5.

Line 585 — Make the caveat about the hot spot models more prominent in the first part of the paper.

We have added two notes in the main text clarifying that the light curve models qualitatively describe the data, though the simplified spot prescriptions in `lcurve` do not fully capture the detailed light curve structure.

Line 630 — $T_{\text{eff}} = 3000 \text{ K}$ is on the high side for an M8 star. Discuss.

We agree that 3000 K is on the high side for the brown dwarf donor gas temperature. We have changed this to 1500 K. As discussed in the manuscript, varying the gas temperature by factors of a few does not meaningfully change the ballistic stream trajectory or the predicted phase offset.

Line 653 — In the absence of synchronous rotation the hot spot would presumably have a very different appearance. Discuss.

We thank the referee for raising this point. The stable light curve morphology observed over many orbits is itself consistent with synchronous rotation. If the accretor were not tidally locked, the impact point of the accretion stream would likely vary significantly and the light

curve morphology would likely change from orbit to orbit. We have added a note to this effect in the Methods.

Line 691 — Justify the claimed time period with stable mass transfer (10^7 – 10^8 years).

We cannot place a definitive quantitative constraint on the mass-transfer lifetime. However, even assuming a conservative inspiral timescale of ~ 1 Gyr, a mass-transfer phase shorter than $\sim 10^7$ yr would correspond to $< 1\%$ of the inspiral timescale, making it improbable to observe even one such system. We have added this reasoning to the revised manuscript.

Line 703 — Confirm that the radii obtained here are consistent with the claimed mass of the star, and what mass radius relation is used.

We have updated the masses and radii based on Reviewer 1’s comment about the accretor mass prior (see above), and confirm they are consistent with expectations.

Line 706 — Comment on the impact of extinction, if any.

Extinction is negligible along the ZTF J1444+4820 sightline, and Gaia extinction estimates for stars along the ZTF J0440+2325 sightline show no measurable extinction in front of ~ 100 pc, though ZTF J0440+2325 lies at the near edge of the Taurus clouds. Because neglecting extinction overestimates a spectroscopic distance, correcting for any extinction would decrease the distance of ZTF J0440+2325 and strengthen the X-ray constraint rather than weaken it. The geometric distance adopted for ZTF J1444+4820 is unaffected by extinction. We have added a note to this effect in the spectral modeling and distances section of the Methods.

Line 807 — A slightly larger distance would bring those X-ray upper limits within range. As a note of caution do mention that the Gaia parallax if taken at face value would do this for J1444.

We now adopt the larger geometric distance for ZTF J1444+4820 (823 pc, with a 574–1304 pc range), at which the X-ray upper limit ($L_X < 5.7 \times 10^{30} \text{ erg s}^{-1}$, and higher toward the far end of the range) does fall within the range of some black widow pulsars. As we note in the compact-object alternatives section, the spectra and the optical eclipses independently rule out this interpretation.

Remarks on code availability: I have not run the code, but from a quick review I see that there are two python codes with a reasonably straightforward to follow structure, albeit without any comments. The data files provided are also organised in a logical order. There is no readme file. I recommend adding a minimum set of comments in the python code and a readme file.

We have added comments to the Python code and a README to the repository.

We thank the editor and the three referees for their careful reading of the revised manuscript and for their positive assessments. Below we reproduce each referee’s comments in full (in *blue italic text*) and address them in turn.

Reviewer 1 (Comments for the Author)

The authors have revised the manuscript following comments from the referees, and I would like to thank the authors for their careful consideration of various points raised during the refereeing process. In particular, the treatment of stellar parameters is significantly improved, and the inclusion of new LRIS spectra provides very interesting and useful context for the nature of the second candidate system. The revised and new analyses in ‘Spectral Modeling and Distances’ section, among other sections in the overall Methods, are thoroughly described, and the statistical analyses and treatments of uncertainties are carefully discussed. From the new data and interpretations, I have a few follow-up questions and some additional minor comments, below.

We thank the referee for the positive assessment and for the careful reading of the revised manuscript.

My most substantial questions involve the surprisingly complex explanation of J1444 as a spectrally-visible tertiary that is neither the donor nor accretor in the observed mass transfer system. Does the lightcurve geometry strictly enforce near edge-on inclination for the system? What are the potential degeneracies across fit solutions/impact of more complex spot prescriptions? Specifically, I am wondering whether a high latitude accretion hot spot + grazing eclipse (and thus inclined orbit) could be consistent with the upper limit on RV variation observed in the multi-epoch LRIS spectra. To place more stringent limits on the RV variation, perhaps an instrument with higher spectral resolution than LRIS is necessary, which could be posited as future work.

In such a scenario with only two bodies, could the M4 object be the accretor with a single unseen high-density companion? Of course, this is only plausible if it can satisfy the Roche lobe geometry constraints and the lightcurve eclipse properties. From a density standpoint, it seems that this would require the donor to be a remnant substellar companion if mass transfer has already been occurring for a long period of time, wherein the companion is reduced to mostly electron degenerate material of very high density (a much older version of J0440?). The manuscript would benefit from a brief quantitative discussion of this scenario’s viability (or lack thereof!), based on the available lines of observational evidence.

We thank the referee for these questions. Reproducing the observed radial velocity at its 3σ upper limit with an M4 star and a $0.04 M_{\odot}$ brown dwarf in the inner binary would require an inclination of $\sim 17^{\circ}$, at which the hot spot would not be eclipsed at all. While our lightcurve fits have some degeneracies in the detailed spot parameters, any inclination that still eclipses the hot spot, even a high-latitude spot with a grazing eclipse, leaves the expected radial velocity amplitude of an M4 star roughly three times the observed limit. The eclipses we observe indicate $i \gtrsim 70^{\circ}$, where the 3σ radial velocity limit would require the companion to the M4 star to be less massive than $\sim 12 M_{\text{Jup}}$, so any brown dwarf companion would be detected. Unless the brown dwarf has been stripped, this scenario is thus ruled out (and

we discuss the possibility of a stripped brown dwarf scenario below). We agree that higher-resolution spectroscopy would place a more stringent limit on any radial velocity modulation of the M4 star, and we mention this in the Methods.

The main obstacle to the scenario of an M4 accreting from an extremely stripped brown dwarf is that the M4 star is too large to fit within a 67-minute orbit. An object that does not fill its Roche lobe at this period must have a mean density $\gtrsim 90 \text{ g cm}^{-3}$, whereas a typical M4 dwarf has a mean density of only about 16 g cm^{-3} and fills its Roche lobe at $P \sim 130$ minutes. While accretion heating could potentially alter the atmosphere of a lower-mass M dwarf, for example an M8 that does not fill its Roche lobe, to mimic an M4, such a lower-mass star would only raise the expected radial velocity amplitude of the M dwarf, from $K \approx 53\text{--}57 \text{ km s}^{-1}$ for an M4 to $K \approx 83\text{--}88 \text{ km s}^{-1}$ for an M8 for inclinations of $70^\circ\text{--}90^\circ$, and further challenge the observed lack of radial velocities. We quantify this in the main text and in Figure 2, where the radial velocity limit corresponds to a 3σ companion mass upper limit of $5.7 M_{\text{Jup}}$ for a $0.08 M_\odot$ M dwarf. Explaining the data this way would thus require both a companion stripped to a few Jupiter masses and a visible star with an atmosphere strongly altered by accretion. Furthermore, based on MESA simulations in preparation, an older version of a system like ZTF J0440+2325 whose donor has been stripped below $10 M_{\text{Jup}}$ would evolve to periods longer than ~ 200 minutes. Though the exact likelihood is difficult to quantify without further modeling, we consider this scenario a possibility but less likely than the hierarchical-triple interpretation, and argue that such a triple is not particularly surprising. The black widow candidate ZTF J1406+1222, which is more likely a system similar to these (see below), is itself in a resolved triple, and a third star also provides a natural mechanism for forming these close systems, driving the inner binary to a short period through Kozai-Lidov oscillations. Better angular resolution, for example with *JWST*, would help test these alternatives, and we mention this in the text.

Minor comments:

The abstract would benefit from brief clarification on what is meant by ‘candidate’ for J1444 (e.g., ‘strong mass transfer candidate system’).

We have clarified the abstract, which now describes ZTF J1444+4820 as a “strong candidate low-mass mass-transferring system in a hierarchical triple.”

The authors may wish to consider the nomenclature for the object, as ‘molar’ (possibly a portmanteau of m-dwarf and polar, given that magnetically channeled accretion is not implied?) could lead to some confusion, and some additional explanation/justification would benefit the new terminology.

We agree and have removed the term “molars” from the manuscript.

Line 089: Please clarify that this is Balmer line emission, as opposed to Balmer continuum.

We have clarified that this refers to Balmer line emission.

Line 099-103: The introductory description of black widow pulsars would be beneficial to the reader earlier when first referenced.

We have moved the brief description to the first mention of black widow pulsars, which now

reads “black widow pulsars, compact binaries in which a neutron star irradiates and ablates a low-mass companion,” and removed the later redundant clause.

Line 176: It would help to remind the reader that J1406+1222 is a candidate black widow pulsar system (to help distinguish its similar naming from the candidates presented in this work). On a related note, in Line 978, some clarification around the language describing J1406 as a literature comparison object would be helpful. It sounds here as though the authors are reconsidering its interpretation as a black widow pulsar candidate and instead that it may be another substellar mass transfer candidate, but the lack of optical eclipses prevent a definitive classification.

We thank the referee for this comment. At the point the referee flagged, we now note in the main text that ZTF J1406+1222 was originally interpreted as a candidate black widow pulsar but is more likely a low-mass mass-transferring system. A forthcoming addendum, led by the lead author of the ZTF J1406+1222 discovery paper (Burdge et al. 2022) who is also a coauthor of the present work, will propose that it may be best explained this way rather than as a black widow pulsar. We leave most of the discussion of that system to the addendum, but note that the one signature that ZTF J1406+1222 always lacked, and that was difficult to reconcile with the black widow interpretation, was the large radial velocity modulation expected in the Balmer absorption from the irradiated face of the companion (Figure 4 of Burdge et al. 2022). The absence of these radial velocities was originally attributed to a possibly face-on inclination and the low signal-to-noise of the spectra. However, a low-mass mass-transferring interpretation now offers a much more natural explanation, given the two systems presented in this work.

Reviewer 1 (Remarks on code availability)

The updates to the Github repository have added helpful descriptions, context, and data products that significantly improve reproducibility.

Reviewer 2 (Comments for the Author)

I appreciate the substantial revisions made by the authors. The revised manuscript is more careful in presentation. In particular: (a) the original reliance on a q_{crit} argument has been appropriately softened and the discussion of mass transfer stability acknowledges the relevant caveats. (b) ZTF J1444 has been downgraded to a candidate. (c) the new RV anchored hot spot phase argument provides a more convincing basis for identifying the likely donor and accretor in ZTF J0440.

Overall, I believe the manuscript is substantially improved. I have only one remaining scientific reservation.

We thank the referee for the positive assessment and appreciate their encouragement to revisit this discussion, which we believe has significantly strengthened the work.

The manuscript now argues that the mass transfer is probably dynamically stable because an unstable mass transfer phase would be short that observing such a system would be unlikely. While I agree that this is a reasonable argument, I am not yet fully convinced that it is sufficient to exclude alternative evolutionary scenarios. The inference depends on the poorly

known progenitor population, observational selection effects, and the duration of the unstable mass transfer phase. As a result, I view the lifetime argument as suggestive. Whether the remaining level of uncertainty is compatible with the strength of the central claim is ultimately a matter for the Editor to assess.

We thank the referee for this assessment, and we agree. As the referee notes, and as we mention in the text, the progenitor population and the selection effects of surveys such as ZTF are not well understood, and we cannot definitively rule out an unstable mass-transfer phase, for example one that becomes unstable after ~ 100 Myr of mass transfer, which could still be observed if the progenitor population is large.

Much of this depends on how one interprets the word “stably” in this context. We intend it mainly to convey that these represent a new class of object, rather than as a strong statement about long-term dynamical stability. In particular, substellar companions to main-sequence stars have only been found in two categories: rapid unstable engulfment events on timescales of years and detached main-sequence–substellar binaries. These systems are neither, but instead represent a different evolutionary phase where a substellar object and a main-sequence star are in a semi-detached configuration, in which one component fills its Roche lobe and transfers mass onto the other. The exact long-term stability of that mass transfer over hundreds of Myr to Gyr timescales is of course less certain and will require a larger sample to establish. Nonetheless, the main novelty of the result (i.e., a new class of object where a star is actively accreting from a substellar object without yet engulfing it) does not depend on the word “stably,” and we would be glad to soften the title and corresponding discussion in the text, for example to “Stars accreting from substellar objects,” and are happy to work with the editor.

Minor comment:

In Figure 3, the offset is labeled as 18.3 degree, but quoted as 18.4 degree in the caption and main text. Please make these values consistent.

We have corrected the caption and the main text from 18.4° to 18.3° .

Reviewer 3 (Comments for the Author)

I am content with the authors’ responses to the issues raised in my first report. The main new addition to the paper is the interpretation of the unexpected spectral type M4 found in ZTF J1444+4820. While I don’t find an obvious flaw, the new line of arguments is long enough to make the system a much less convincing example of a mass-transferring brown dwarf. The language used in the manuscript presents this in a careful and measured way, so I have no objections to the paper to remain essentially as is.

I have no substantive comments to add, just two minor items:

We thank the referee for the positive assessment.

The consistent use of “the M4” to denote the dominant stellar component in ZTF J1444+4820 is odd. I would expect M4 to be followed by a substantive such “M4 star”, “M4 type”, “M4 tertiary” etc.

We have revised the manuscript so that “M4” is followed by a substantive throughout, using “M4 star” in the general discussion and “M4 tertiary” where its role as the tertiary is established.

I would encourage the authors to reconsider the term “molars”. A claimed class with only one member and a second somewhat dubious member is arguably too small to yet warrant a name, and while “molar” is memorable and perhaps funny the parallel with polars is tenuous, and using m instead of p in the word is confusing at best.

We have removed the term “molars” from the manuscript.