nature portfolio

Peer Review File

An Adolescent and Near-Resonant Planetary System Near the End of Photoevaporation

Corresponding Author: Mr Mutian Wang

This manuscript has been previously reviewed at another journal. This document only contains information relating to versions considered at Nature Astronomy.

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)
This paper presents the TOI-2076 system, including detailed mass and radii constraints. As a three-planet system this would be interesting, but the young age (~200Myr) and opportunity to use the system as a benchmark for photoevaporation and planet evolution studies means that the work is suitable for Nature Astronomy, in my opinion. Although a single system, this has the potential to be a key piece of evidence in situating evolution processes in planetary age, and points towards early photoevaporation playing a key role. The system of well characterised planets with similar masses and very different radii presents opportunities for further understanding.

I have some questions below which I request are dealt with before publication.

1. In the abstract, it is not clear what scale of planets are being discussed. Please specify sub-Neptunes (or other terminology as the authors feel appropriate). The main issue here is distinguishing between terrestrial and gas giant planets for a general reader.

2. Line147-149. The Delta parameter and associated values for these planet pairs needs context. Could the authors provide typical Delta values for planet pairs in or near resonance to compare to the 1.5% and 0.3% given here?

3. Line214-216. Much is made of TOI-2076 being in near-resonance rather than librating, and I accept the strong and detailed analysis demonstrating this. However, it is stated that convergent disk migration establishes pristine MMRs, with librating resonance angles. Could the authors provide detail or citations for how pristine the MMRs produced by convergent disk migration should be? There must be some spread in period ratio. Given that a significant conclusion of the paper is that the TOI-2076 system must have undergone some dynamical evolution to move from a librating resonant system to a near-resonant one, some evidence should be given that the initial conditions were, or should be expected to be, a close resonance.

4. Erosion of Primordial Atmospheres. The models used to test atmospheric erosion assume identical initial gas envelopes for the planets, if I have understood the paper correctly. Is there no expectation of different envelopes due to different formation locations in the disk for these planets? One would expect that the more distant planets started further out in the disk, even potentially beyond the snow line, before migrating and being trapped in resonance. What effect on the models would such differences, in composition or in available gas mass, have?

5. Again on the evolution of H/He in the planetary atmospheres, the simulation seems to assume that the planets reached their present day location very early, and the separation stays static for the whole photoevaporation model. Is the system reaching its present day configuration before photoevaporation (i.e. on Myr timescale) realistic and can this be backed up by citation? If there is some ambiguity in arrival time at the present configuration, what effect does this have on the simulations?

6. Section2.3, Equation 4. Sigma_jit appears to be time dependent in this equation, is this correct? It would be an unusual modelling step to take if so.

7. Section2.3. “the mass constraints do not vary much..”. Please quantify this, is the variation significant or not?

(Remarks on code availability)

Reviewer #2

(Comments for the Author)
This article characterizes and contextualizes the young and compact TOI-2076 multi-planet system. The available data include time-series photometry from TESS, CHEOPS, and LCOGT and archival radial velocity measurements from HARPS, HIRES, and APF. Beyond the new transit data, the main novelties in the article come from the dynamical analysis (TTVs, near-resonant circulating state, photodynamical modeling) and -- to a lesser extent -- the atmospheric evolution models (assuming XUV-driven photoevaporation). These analyses are enabled by the mass solution derived for all four planets from a joint photodynamical TTV+RV fit. The results are that TOI-2076 appears to be in a fragile near-resonant state with an architecture that seems likely to have been sculpted by a combination of resonant migration, an unknown form of subsequent (gentle) disruption, and atmospheric loss.

The work is clear, careful, and well-written. The data collection and data analysis represent novel and significant efforts, and the models suggest a picture of early multi-planet system evolution that will interest many astronomers. The work therefore constitutes an important contribution to the field. There are nonetheless a few areas in which the article could be improved.

Comments refer to the line numbers in the original submission's converted PDF.

MAJOR

The results of the injection-recovery exercise shown in Figure 12 seem very odd. In particular, the uncertainties on the plot do not appear to correspond to the usual meaning: they are too close to the 1:1 line (especially planets c and d), and none of the planets seem to show intrinsic dispersion around a line of constant bias -- instead they each seem to have some fixed, constant offset from 1:1. Why is this happening? This issue currently hinders assessment of the injection-recovery exercise. For purposes of the review, it would be more helpful to see analogs of Figure 9 and the phase-folded RVs from Figure 1, but after having subtracted the nominal Keplerian signals and refitted the stellar activity model. This test is most relevant for cases of 0, 0.5, 1.0, 1.5, and 2.0 m/s (which are more comparable to the signal amplitudes being reported).

It seems that the dynamical modeling (and perhaps atmospheric modeling) results might depend on the masses of the planets, which in turn depend on the RV analysis. How robust are the conclusions regarding the near-resonant state of the b-c and c-d pairs to possible changes in the derived masses of the planets? Stated differently: how large of a change in M_b, M_c, and M_d would be needed for the system to *not* be in the near-resonant circulating regime? This question originates from systematic concerns about the gaussian process regression, for which tests are suggested below. However, those tests and concerns are ultimately moot if there is a factor a few flexibility in allowed planet mass for the dynamical arguments to hold.

Figure 4 vs 3: The idea of Figure 4 seems to be that we should equate the architecture of TOI-2076 and TOI-178 with that of V1298 Tau. However at t=20 Myr, Figure 3 shows the model TOI-2076 planet sizes at around two to three times smaller than the current V1298 Tau planet sizes. This suggests something may be missing from the current picture; possibilities might be addressed in discussion.

MINOR

* GPs and RV analysis:

* A more robust cross-validation test, more analogous to that from Blunt et al. 2023, would be to train only on the HARPS data and then assess whether the model predicts the contemporaneous HIRES data.

* A simple check on the FAP of the reported planetary RV signals could come from a Lomb-Scargle periodogram of the "RV - stellar activity" residuals. See for example Figure 1 of Basant et al. (2025, arXiv:2503.08095).

* Would there be any use in calculating a marginal log likelihood for the model with four Keplerian orbits against a model with no Keplerian signals? It is understood from previous work that this test can lead one astray (Suarez-Mascareno et al. 2021). Nonetheless, consideration of this test could perhaps improve the manuscript, as well as offer some statistical measure beyond "N-σ" for the systematic degree of confidence in the RV-derived masses.

* Abstract: "This trend strongly supports photoevaporation (with timescale of 10s-100s of Myr), rather than giant impacts, as the dominant mechanism driving the erosion of primordial H/He atmospheres." Suggest the more generic "atmospheric escape" over "photoevaporation", since photoevaporation was the only process that was considered in this article's modeling, and the remnant thermal energy from formation and bolometric stellar luminosity likely also play a role. The quoted timescale of 10s-100s of Myr is somewhat confusing given the few gigayear lifetimes for c and d quoted in Figure 3.

* Abstract: Does a metastable helium outflow rule out the concept of a steam-rich interior (i.e. the supercritical hydrosphere

style model from Burn et al. 2024), aka "water world"? Why can such interiors not have volatile envelopes?

* L168: "around 5 m/s": The reported semi-amplitudes seem to span 1.3 to 2.4 m/s. (A 20:1 ratio against the spot-induced variability, rather than 10:1).

* L216-217: "reveals that the system has undergone significant post-formation evolution" seems overstated.

* L555-557: "It is reasonable to assume..." The RV is more closely linked to the time derivative of the flux time-series. See Aigrain et al. MNRAS 419, 3147–3158 (2012), or Klein & Donati MNRAS 493, L92–L97 (2020).

* Figure 14: Great plot! I really appreciated the inclusion of this part of the analysis.

NITS

L243-245 "lies deep" typo, "exhibits" typo
256 Hamlitonian
263 missing "librates" around 180 degrees
348 disk "disperse"

(Remarks on code availability)

Version 1:

Reviewer comments:

Reviewer #1

(Comments for the Author)
I thank the authors for a clear and comprehensive response to the questions raised. They have all been met and I am happy to recommend the paper for publication as is.

Reviewer #2

(Comments for the Author)
Thank you to the authors for addressing my concerns.

Minor notes, no response needed:
- Idea of ~100 Myr timescales for elevated stellar activity has been debated, particular for EUV decay https://ui.adsabs.harvard.edu/abs/2021MNRAS.501L..28K/abstract
- V1298 Tau planets being caught in the boiloff phase seems quite fine-tuned given the expect timescales. See Rogers et al arXiv:2311.12295

Dear Reviewer 1,

Thank you very much for your thorough reading of our manuscript and for the constructive and insightful comments. Below we appended a detailed, point-by-point response to each of your comments. The response to your comment and the modified parts in manuscript are both highlighted in red for ease of reading. The referenced line number corresponds to the revised version of manuscript.

Best regards,
Mutian Wang (on behalf of all co-authors)

==================================

1. In the abstract, it is not clear what scale of planets are being discussed. Please specify sub-Neptunes (or other terminology as the authors feel appropriate). The main issue here is distinguishing between terrestrial and gas giant planets for a general reader.

Response:
Thank you for pointing this out. We have restricted to the discussion in sub-Neptune systems in line 99.

2. Line147-149. The Delta parameter and associated values for these planet pairs needs context. Could the authors provide typical Delta values for planet pairs in or near resonance to compare to the 1.5% and 0.3% given here?

Response:
Thanks for the suggestion. We have added context about typical Kepler near-resonant planet display Δ= 1-2% in Line 149-150.

3. Line 214-216. Much is made of TOI-2076 being in near-resonance rather than librating, and I accept the strong and detailed analysis demonstrating this. However, it is stated that convergent disk migration establishes pristine MMRs, with librating resonance angles. Could the authors provide detail or citations for how pristine the MMRs produced by convergent disk migration should be? There must be some spread in period ratio. Given that a significant conclusion of the paper is that the TOI-2076 system must have undergone some dynamical evolution to move from a librating resonant system to a near-resonant one, some evidence should be given that the initial conditions were, or should be expected to be, a close resonance.

Response:
Thanks for the suggestions on clarifying this context. The simplest model of convergent migration of sub-Neptune systems in a laminar disk (e.g., Keller et al. 2025) shows that the typical Δ from the final resonant chain is < 0.1% for 1st-order MMR, and <0.05% for 2nd-order MMR. So there appears to be some mild increase on Δ for TOI-2076 planets compared to the simplest expectation from disk migration model. Other physical processes might further modify the intial period ratio

distrbution, and is now further discussed in the main text. We now elaborate the initial conditions, potential explanations in line 275-280, 285-288.

4. Erosion of Primordial Atmospheres. The models used to test atmospheric erosion assume identical initial gas envelopes for the planets, if I have understood the paper correctly. Is there no expectation of different envelopes due to different formation locations in the disk for these planets? One would expect that the more distant planets started further out in the disk, even potentially beyond the snow line, before migrating and being trapped in resonance. What effect on the models would such differences, in composition or in available gas mass, have?

Response:
Thanks for this important question. Based on our understanding, the gas accretion history of sub-Neptunes is thought to be controlled primarily by the core mass and the cooling efficiency of the envelope, while depending only logarithmically on the local nebular gas density and temperature (Pollack 1996, Rafikov 2006). This is because the accretion process itself generates a highly opaque atmosphere: the formation of H- ions at high temperatures increases the envelope opacity, which strongly limits radiative cooling and therefore buffer the gas accretion rate. As a result, the key bottleneck is the envelope's cooling opacity rather than the absolute gas supply in the disk.

More quantitatively on the degree of gas density on gas accretion. Lee & Chiang 2015 predicts that the envelope mass scale with $\sim\Sigma^{0.12}$, with Σ denoting local gas density. In the most extreme case, where the outermost planet formed beyond snowline (>1 AU) while inner planets form more in-situ, the gas density contrast will be 30-100 times, and leads to an enhancement of 1.5-1.7 on the accreted gas of outermost planet.

For this reason, in our models we adopt comparable initial envelope mass fractions for planets because of their similar core mass, independent of their putative formation location. We opt to elaborate this point in Line 360-362, and build up this point in Section 2.5 "Initial H/He Envelope Contents" showing that variations in disk conditions are subdominant than planet core mass compared to the envelope opacity–driven regulation of gas accretion.

5. Again on the evolution of H/He in the planetary atmospheres, the simulation seems to assume that the planets reached their present day location very early, and the separation stays static for the whole photoevaporation model. Is the system reaching its present day configuration before photoevaporation (i.e. on Myr timescale) realistic and can this be backed up by citation? If there is some ambiguity in arrival time at the present configuration, what effect does this have on the simulations?

Response:
Thank you for raising this point. The time evolution of our photoevaporation model should be interpreted as starting from the epoch of disk dispersal. Before the disk dissipates, the dense nebular gas shields planets from significant atmospheric escape, and orbital migration due to tidal interaction with the disk is expected to occur during this stage.

Once the disk has dispersed, disk-driven migration ceases, and the subsequent orbital evolution is partly governed by tidal interactions with the star. However, these tidal processes operate on timescales of $\gtrsim$1 Gyr, far longer than the photoevaporation epoch. Atmosphere mass loss is also thought to induce orbital migration. However, the extent is also minimal, which is estimated to be ~1% by Hanf et al. 2025, Wang & Lin 2023.

To summarize, the planets should be largely in place where they are observed now by disk dispersal, so it would be a reasonable approximation to treat the planetary orbits as fixed during our mass-loss simulations. We have clarified this about this in Line 373-375.

6. Section2.3, Equation 4. Sigma_jit appears to be time dependent in this equation, is this correct? It would be an unusual modelling step to take if so.

Response:
Thank you for raising this. The RV jitter is time-independent. We have corrected this typo in the text.

7. Section2.3. "the mass constraints do not vary much..". Please quantify this, is the variation significant or not?

Response:
We apologize for not making this clear in the original text. The mass constraints obtained using different RV products are consistent within 1σ (tables attached). For clarity, we now include a more quantitative description of this comparison in lines 595–600.

HARPS pipelines	Planet e	Planet b	Planet c	Planet d
HARPS(DRS) +HIRES/APF	4.85±1.60	6.95±2.20	6.34±3.0	5.35±3.20
HARPS(lbl) +HIRES/APF	4.86±1.50	6.54±2.5	6.36±2.75	4.72±3.12
HARPS(serval) +HIRES/APF (adopted)	**4.91±1.34**	**7.1±1.9**	**6.9±2.4**	**6.1±2.6**

Dear Reviewer 2,

Thank you very much for your thorough reading of our manuscript and for the constructive and insightful comments. Below we appended a detailed, point-by-point response to each of your comments. The response to your comment and the modified parts in manuscript are both highlighted in red for ease of reading. The referenced line number corresponds to the revised version of manuscript.

Best regards,
Mutian Wang (on behalf of all co-authors)

====================================

1. The results of the injection-recovery exercise shown in Figure 12 seem very odd. In particular, the uncertainties on the plot do not appear to correspond to the usual meaning: they are too close to the 1:1 line (especially planets c and d), and none of the planets seem to show intrinsic dispersion around a line of constant bias -- instead they each seem to have some fixed, constant offset from 1:1. Why is this happening? This issue currently hinders assessment of the injection-recovery exercise. For purposes of the review, it would be more helpful to see analogs of Figure 9 and the phase-folded RVs from Figure 1, but after having subtracted the nominal Keplerian signals and refitted the stellar activity model. This test is most relevant for cases of 0, 0.5, 1.0, 1.5, and 2.0 m/s (which are more comparable to the signal amplitudes being reported).

Response:
We thank the referee for pointing out the confusion presented in the injection-recovery test. Planets b and e show a constant offset from the 1:1 line because we subtracted the best-fit (MAP) solution from the observed RVs. We verified that the semi-amplitudes of planets b and e in the best-fit solution differ from their posterior medians by 0.2 and 0.3 m s^{-1}, respectively. These small mismatches leave residual planetary signals, which appear as the constant offset in Figure 12 and consequently bias the injection–recovery tests.

The absence of dispersion around the 1:1 line originates from the way the original recovery test was implemented. Specifically, we subtracted the observed RVs with the nominal Keplerian model, injected synthetic Keplerian signals, and then repeated the full GP fit. In this setup, the underlying activity + white-noise model (measurement errors and instrumental jitter) used for recovery is identical to that in the fiducial fit. As a result, the GP always perfectly reproduces the activity + noise contribution, and the injected Keplerian signal is recovered without dispersion.

To address this, we performed an alternative test in which we drew different realizations of white noise for each injection. Specifically, we adopted the mean GP prediction conditioned on the original dataset at the observing epochs as the activity model, then added white noise with amplitude $(\sigma_{rv}^2 + \sigma_{jitter}^2)^{0.5}$. Each injection therefore has a distinct noise realization. We then re-fitted the data with GP included, allowing all hyperparameters to vary. The results attached below show that while the recovered semi-amplitude precisions (error bars) are comparable across injections, the median of recovered values are scattered and biased at the ~1σ level relative to the injected values.

We have updated the injection-recovery test in current draft (line 636-652, Figure 8). The phase-folded, activity-subtracted RV plots for each injection is appended at the end of this report.

2. It seems that the dynamical modeling (and perhaps atmospheric modeling) results might depend on the masses of the planets, which in turn depend on the RV analysis. How robust are the conclusions regarding the near-resonant state of the b-c and c-d pairs to possible changes in the derived masses of the planets? Stated differently: how large of a change in M_b, M_c, and M_d would be needed for the system to *not* be in the near-resonant circulating regime? This question originates from systematic concerns about the gaussian process regression, for which tests are suggested below. However, those tests and concerns are ultimately moot if there is a factor a few flexibility in allowed planet mass for the dynamical arguments to hold.

Response:

Thanks for this important question. We evaluate this by considering the resonance width, which depends on planetary mass and eccentricity (Tamayo & Hadden 2024). Briefly, the resonance width of period ratio Δ_{max} near j:(j-k) MMR depends on planet mass and eccentricity as:

$$\Delta_{max} \sim \sqrt{\mu e_{12}^{k}}$$

Where μ is the planet-to-star mass ratio, e12 is the relative eccentricity, for which we test against the eccentricity range derived from photodynamics results. In the plot below, we compare the observed period ratio offset, Δ, with the resonance width for different assumed planetary masses. The b–c pair lies well outside the first-order resonance: at the observed Δ~1.5%, the planets would need to be about ten times more massive to fall within the resonant regime and exhibit libration. The c–d pair is closer to the second-order resonance, but it would still require planet masses about two to three times larger than the current estimates to be resonant. However, the updated injection–recovery tests show that the GP-recovered masses of planets c and d are biased at fractional precision of < 50% . Therefore, the c–d pair can also be regarded as securely in a near-resonant, circulating configuration.

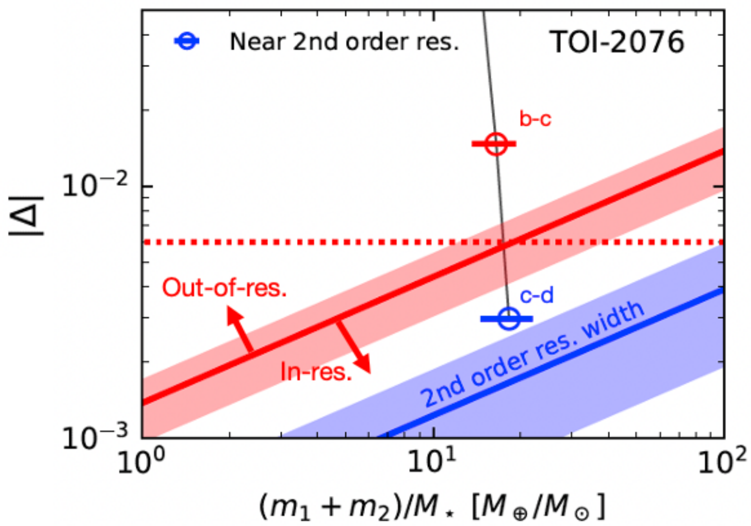


Figure 1. Comparison of observed normalized distance to MMR (Δ) and planet mass to the resonance width of first-order (red line) and second-order (blue line) MMR. The width of band reflects the resonance width's dependence on planetary eccentricity, for which we adopt the range from photodynamical result.

3. Figure 4 vs 3: The idea of Figure 4 seems to be that we should equate the architecture of TOI-2076 and TOI-178 with that of V1298 Tau. However at t=20 Myr, Figure 3 shows the model TOI-2076 planet sizes at around two to three times smaller than the current V1298 Tau planet sizes. This suggests something may be missing from the current picture; possibilities might be addressed in discussion.

Response:

We thank the referee for pointing out this difference. The difference between the radius of V1298 Tau planets and the TOI-2076 at its similar age might be attributed to the difference on the envelope mass fraction f_{env}. It is well-known that the planet radius is a proxy for the planetary envelope content mass (Lopez & Fortney 2014, Chen & Rogers 2016). According to the mass-radius relationships from Mordasini et al. 2012, the mass envelope fraction f_{env} would be 10-20% and ~40% to match the 5-6 R_earth (V1298 Tau c/d) and ~8-9.5 R_earth (V1298 Tau e/b), respectively. In our simulation, all planets start with f_{env} = 6.5%, much lower than the aforementioned f_{env} and thus is smaller than V1298 Tau. We have discussed this point in Line 401-406. Nevertheless, we have also simulated a case (attached below) where TOI-2076 bcd started with f_{env} ~ 20%, the radius started with ~6 R_earth, much more consistent with the V1298 Tau planet c/d. But initial f_{env} = 20% would be too high compared to expectations from core accretion theory (which predict f_{env} =5-10% as discussed in Section 2.5 "Initial H/He Envelope Contents"). Therefore we stay with the original version of the simulation.

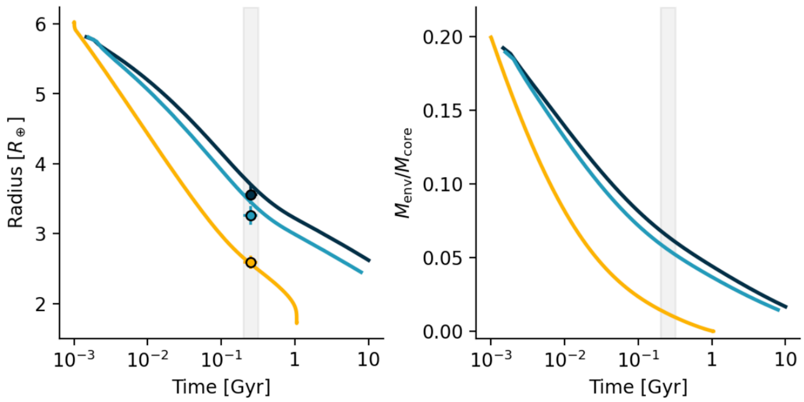


Figure 2. Same atmosphere mass loss simulation as in original paper, but start with higher initial envelope mass fraction (f_{env} = 20%).

Another possibility is that V1298 Tau planets are currently in a transient phase of rapid mass loss, often termed "boil-off" (Owen & Wu 2016, Wang & Dai 2019). Recent JWST measurements indicate that V1298 Tau b have hot (~500 K) interiors (Barat et al. 2025). Such high internal entropy can inflate planetary radii beyond standard hydrostatic models and is consistent with an ongoing boil-off phase. Our simulations do not explicitly model this process; instead, we treat it as a short "burn-in" stage that precedes the evolutionary tracks shown in the figures. Thus, the initial conditions of our models should be regarded as post–boil-off states. This scenario is now discussed in line 851-865.

4. A more robust cross-validation test, more analogous to that from Blunt et al. 2023, would be to train only on the HARPS data and then assess whether the model predicts the contemporaneous HIRES data.

Response:

We thank the referee for the suggestions of cross-validation test. We performed an additional cross-validation test following your advice. Here, we used the median values of GP hyperparameters listed in Table 3 in the draft, then we predict the HIRES and APF RV points with the GP model conditioned only on HARPS data. Below is the different between observed and predicted RV data with this exercise, in unit of the quadratic sum of RV and GP uncertainties. APF and HIRES seems to follow a broader distribution than HARPS, but is concentrated around 0 and consistent with normal distributions. Therefore we conclude that the existing GP model has reasonable predictive ability on unseen, contemporaneous data. We have updated this new cross-validation test in Line 620-633 and Figure 7.

5. A simple check on the FAP of the reported planetary RV signals could come from a Lomb-Scargle periodogram of the "RV - stellar activity" residuals. See for example Figure 1 of Basant et al. (2025, arXiv:2503.08095).

Response:

We have made Generalized Lomb-Scargle periodograms (via Pyastronomy) to (1) RV, (2) RV-stellar activity, and (3) RV - stellar activity – Keplerian data. In the second panel that shows the power spectrum of RV-activity data, planet e and b has FAP<0.1%. Planet c has FAP ~0.1%, while planet d only have FAP>10%.

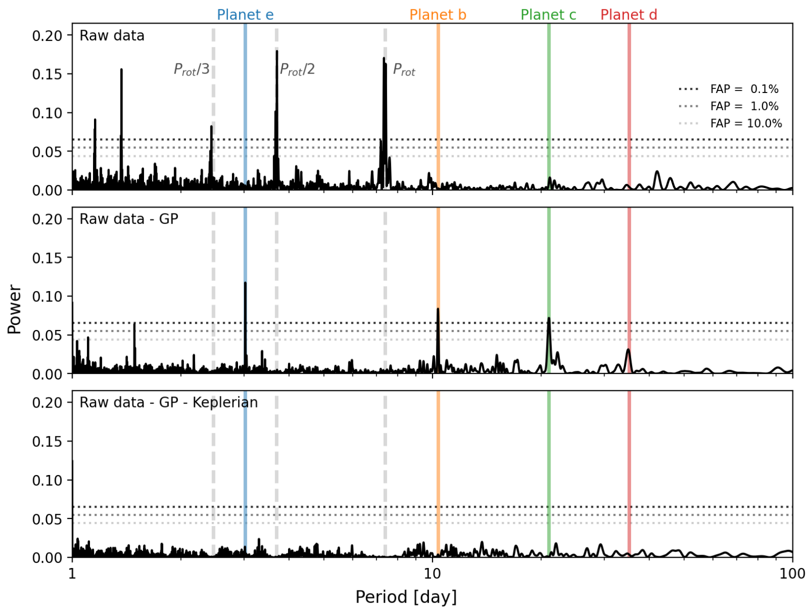


Figure 3. GLS periodogram of (1) observed RV, (2) observed RV subtracted by activity GP model, and (3) observed RV subtracted by activity GP model and all Keplerian signals. The

relevant periodicity of planets and stellar activity are marked as vertical lines.

6. Would there be any use in calculating a marginal log likelihood for the model with four Keplerian orbits against a model with no Keplerian signals? It is understood from previous work that this test can lead one astray (Suarez-Mascareno et al. 2021). Nonetheless, consideration of this test could perhaps improve the manuscript, as well as offer some statistical measure beyond "N-σ" for the systematic degree of confidence in the RV-derived masses.

Response:

We thank the referee for the suggestion. We carried out nested samplings with pure GP, GP+2pl, and GP+4pl models. The log-evidence values are listed in the table below. Compared to the pure GP model, the GP+2pl and GP+4pl models improve the evidence by ΔlogZ ~ 1.1 and 2.0, corresponding to Bayes factors of ~3 and ~7.5, respectively. This indicates positive support for including Keplerian signals. The GP+4pl model is modestly favored over the GP+2pl model (ΔlogZ ~ 1.1). We have added this into the Line 601-611.

Model	logZ
Pure GP	-1442.592
GP+planet b/e	-1441.447
GP+planet b/c/d/e	-1440.580

7. Abstract: "This trend strongly supports photoevaporation (with timescale of 10s-100s of Myr), rather than giant impacts, as the dominant mechanism driving the erosion of primordial H/He atmospheres." Suggest the more generic "atmospheric escape" over "photoevaporation", since photoevaporation was the only process that was considered in this article's modeling, and the remnant thermal energy from formation and bolometric stellar luminosity likely also play a role. The quoted timescale of 10s-100s of Myr is somewhat confusing given the few gigayear lifetimes for c and d quoted in Figure 3.

Response:

Thank you for the helpful suggestions. Our results indicate that atmosphere mass loss driven by stellar flux is the preferred explanation in this system, given its youth (200 Myr). This is consistent with expectations that photoevaporation is most effective at early times when stellar activity is elevated, and with the outcome that a 1% H/He envelope remains as a long-lived, converged result of photoevaporation. In contrast, core-powered mass loss driven by remnant heat and bolometric stellar luminosity typically operates on gigayear timescales, which appears less favored in this case. Nevertheless, we have revised the statement in the abstract to present it in a more objective way, and acknowledge that alternative atmospheric escape processes may also contribute, and complement the discussion in Lines 851-865 of the revised draft.

Regarding the quoted timescales, the "10s–100s of Myr" refers to the epoch during which most of the mass loss occurs (e.g., ~80% of the total lost mass for planet b is removed within the first 100 Myr). By comparison, the gigayear lifetimes of planets c and d indicate that their envelopes are not expected to be substantially sculpted by photoevaporation, highlighting the differential sculpting effect of photoevaporation. However, recent work on the stellar UV evolution suggests a more a more gradual decrease in XUV flux (e.g., King et al. 2020), which might lengthen the timescale of XUV-driven atmospheric mass-loss to 100 Myr - 1Gyr. Therefore we decided not to explicitly mention the “10s–100s of Myr” timescale in the text.

8. Abstract: Does a metastable helium outflow rule out the concept of a steam-rich interior (i.e. the supercritical hydrosphere style model from Burn et al. 2024), aka "water world"? Why can such interiors not have volatile envelopes?

Response:
We thank the referee for raising this point. We intended to say metastable helium rules out the pure water world (rocky core + water layer, no H/He). It is indeed true that Hycean world (rocky core + water layer + H/He) and steam world (supercritical water mixed with H/He as in Burn et al. 2024) can still produce Helium outflows. We have clarified this wording in Line 110.

9. L168: "around 5 m/s": The reported semi-amplitudes seem to span 1.3 to 2.4 m/s. (A 20:1 ratio against the spot-induced variability, rather than 10:1).

Response:
We have corrected the RV signal <2.5m/s in Lin 170.

10. L216-217: "reveals that the system has undergone significant post-formation evolution" seems overstated.

Response:
We have stated this more objectively, comparing the initial conditions expected from disk migration model and the observed architecture, and discussing possible avenues in Line 278-283.

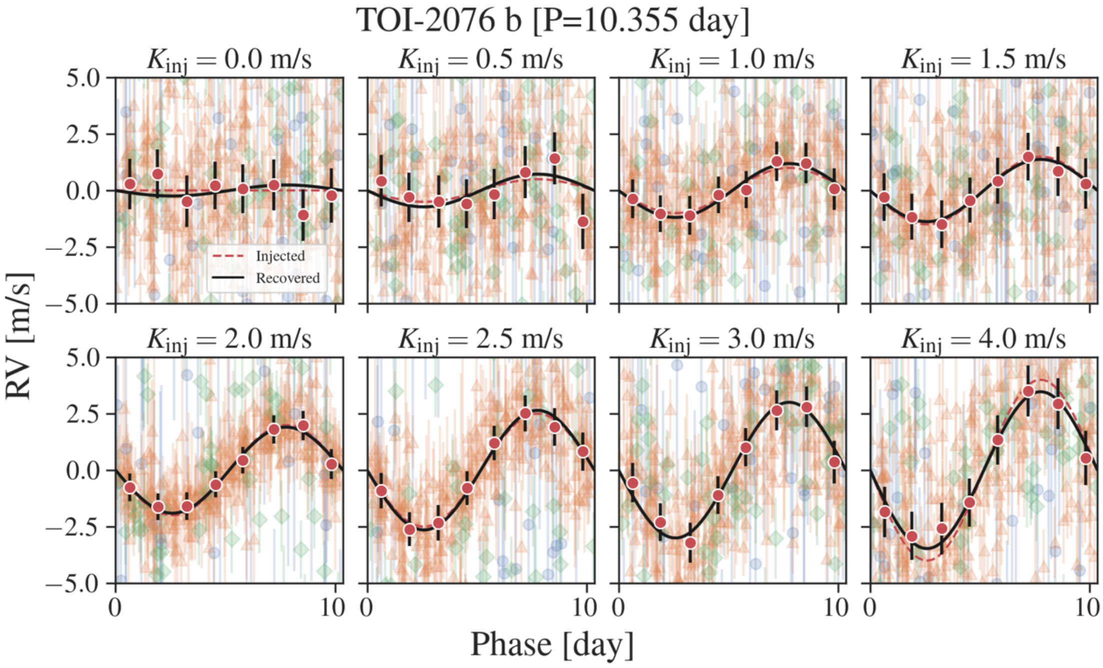
11. * L555-557: "It is reasonable to assume..." The RV is more closely linked to the time derivative of the flux time-series. See Aigrain et al. MNRAS 419, 3147–3158 (2012), or Klein & Donati MNRAS 493, L92–L97 (2020).

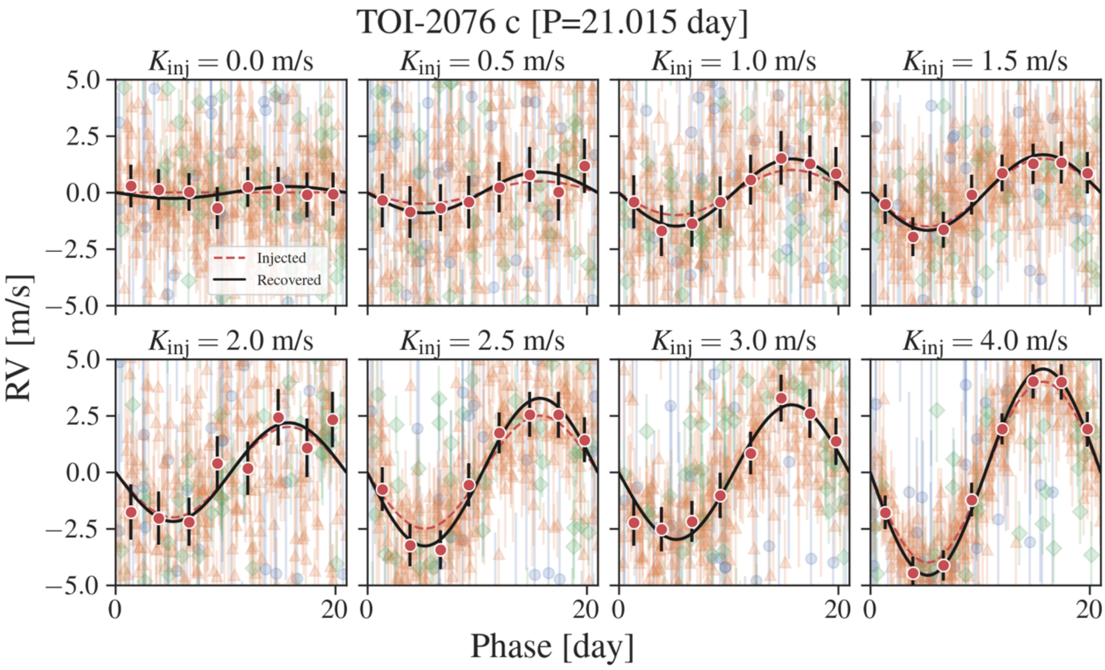
Response:
We have revised the text accordingly in Line 571-573.

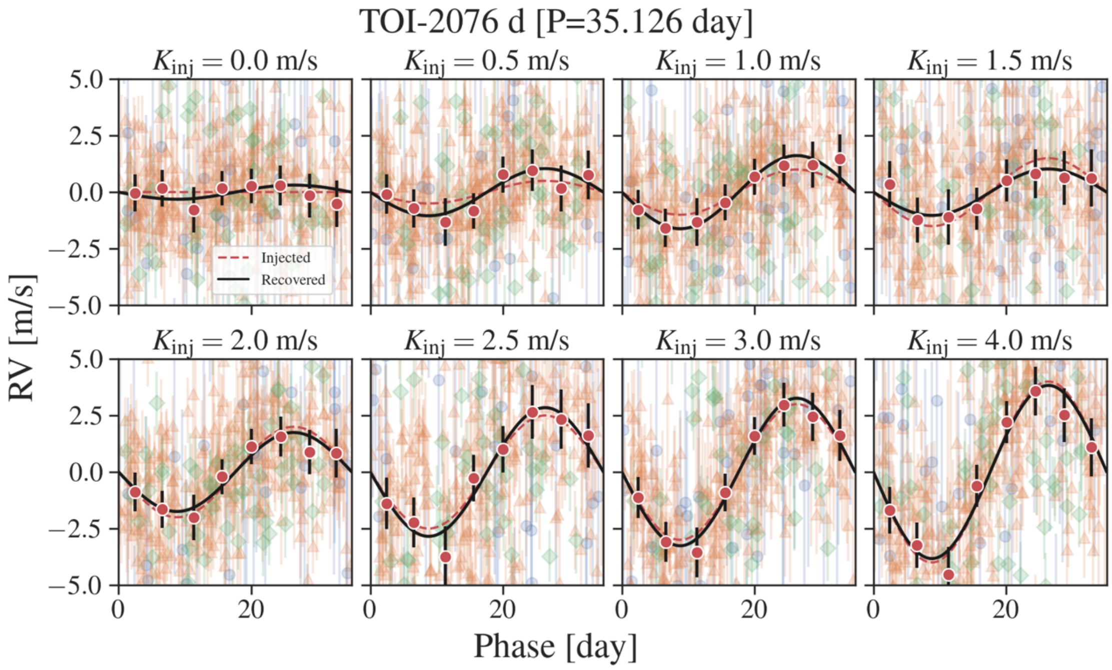
L243-245 "lies deep" typo, "exhibits" typo
256 Hamlitonian
263 missing "librates" around 180 degrees
348 disk "disperse"

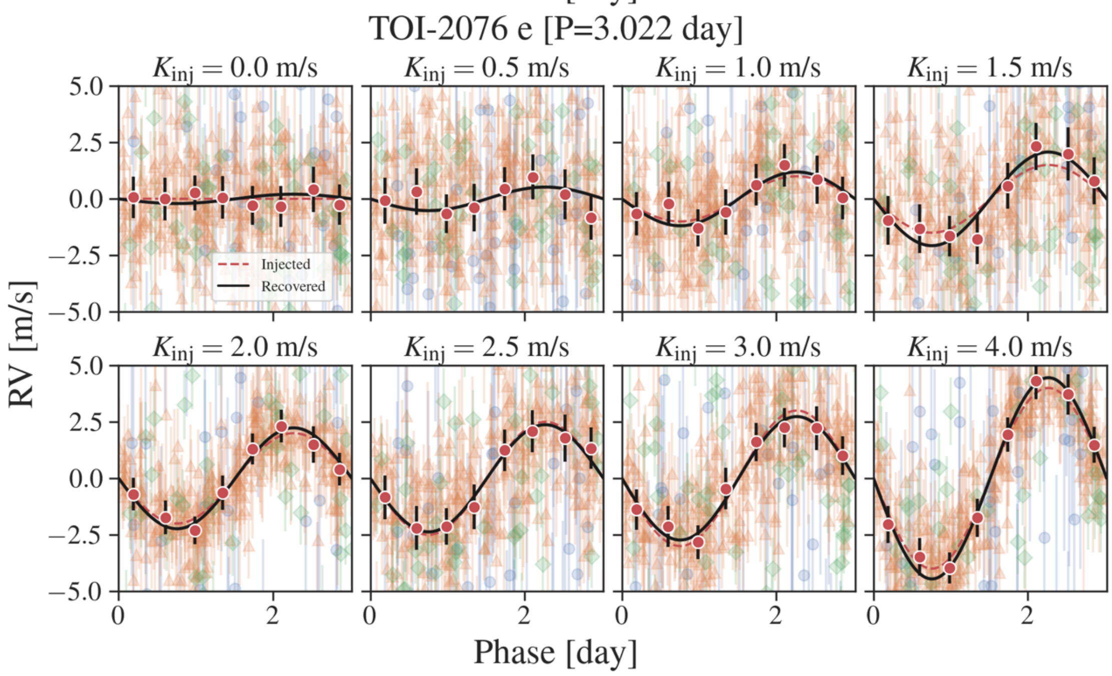
Response:
We have corrected all typos in the draft.

APPENDIX Phase-folded RV plots for Injection-Recovery Tests






TOI-2076 d [P=35.126 day]
$K_{inj} = 0.0$ m/s
$K_{inj} = 0.5$ m/s
$K_{inj} = 1.0$ m/s
$K_{inj} = 1.5$ m/s
$K_{inj} = 2.0$ m/s
$K_{inj} = 2.5$ m/s
$K_{inj} = 3.0$ m/s
$K_{inj} = 4.0$ m/s
Injected
Recovered
RV [m/s]
Phase [day]


TOI-2076 e [P=3.022 day]
$K_{inj} = 0.0$ m/s
$K_{inj} = 0.5$ m/s
$K_{inj} = 1.0$ m/s
$K_{inj} = 1.5$ m/s
$K_{inj} = 2.0$ m/s
$K_{inj} = 2.5$ m/s
$K_{inj} = 3.0$ m/s
$K_{inj} = 4.0$ m/s
Injected
Recovered
RV [m/s]
Phase [day]

Reviewer #1:

Comments for the Author:

I thank the authors for a clear and comprehensive response to the questions raised. They have all been met and I am happy to recommend the paper for publication as is.

Response: Thank you for your careful reading of our manuscript and for the insightful comments. These suggestions have helped us improve the clarity and robustness of the paper.

Reviewer #2:
Comments for the Author:

Response: Thank you for your careful reading of our manuscript and for the insightful comments. These suggestions have helped us improve the clarity and robustness of the paper.

Minor notes, no response needed:
- Idea of ~100 Myr timescales for elevated stellar activity has been debated, particular for EUV decay https://ui.adsabs.harvard.edu/abs/2021MNRAS.501L..28K/abstract

Response: Thank you for raising this point. Adopting a slower XUV decay law after the saturation phase does not substantially change the present-day atmospheric state inferred for the modeled planets, as the high-energy flux is elevated during the early evolutionary phase regardless of the decay prescription. However, it could influence the future atmospheric evolution, especially for planet b, which may lose additional envelope mass and potentially evolve toward a super-Earth on gigayear timescales under a slower EUV decay scenario. We now mention this caveat explicitly in Section 2.6 of the Methods.

- V1298 Tau planets being caught in the boiloff phase seems quite fine-tuned given the expect timescales. See Rogers et al arXiv:2311.12295

Response: Thank you for bringing this paper to our attention. Indeed, the short duration of the boil-off phase makes it an exceedingly transient snapshot compared with the 1e7-yr age of the system. This suggests that the V1298 Tau planets are larger than the initial radius of TOI-2076 more likely due to their higher initial envelope mass fractions.