Answer to reviewers

Paper egusphere-2022-645: “Statistical distribution of mirror mode-like structures in the magnetosheaths of unmagnetised planets: 1. Mars as observed by the MAVEN spacecraft” by Simon Wedlund et al.

2023-02-05: Point-by-point response to reviewers.
2023-03-03: Table 1 caption changed.

We would like to thank the two anonymous referees for their constructive comments, questions and suggestions. Please find our answers below, with all the points raised during the review addressed in the following.
Throughout, our responses are marked as “AC” (author comment AC1 or AC2 for reviewer 1 or 2) and in italic, whereas all other instances (non-italic text) are the original reviewer comments and questions.

Reviewer 1, Aug. 2022

The manuscript focuses on the identification of mirror mode like waves in the magnetosheath of Mars, using in-situ observations made by NASAs MAVEN spacecraft. Mirror mode waves are common in planetary magnetospheres and are generated by temperature anisotropies in the plasma; they thus provide insight into the physical processes active there. The manuscript provides a good summary of this background information and the motivation for the presented work.

The manuscript describes an automated detection algorithm to identify mirror mode like waves in the magnetosheath of Mars, via in-situ magnetic field measurements that are made by MAVEN. The algorithm is initially trained using supplementary ion observations, but only utilizes magnetic field observations. This enables the results to be compared to other spacecraft missions, on which magnetometers are commonly carried. This intentional design of the algorithm is a well thought out strength, which should result in new and interesting comparisons being made in the future (there is a companion manuscript to this one, investigating results at Venus, for example). This manuscript presents the first set of statistical results obtained via the application of this algorithm to the MAVEN magnetometer dataset at Mars.

The method is well described and the manuscript presents a thorough discussion of the results that includes appropriate comparison to previous related works. The authors find that mirror mode like waves are most commonly observed just behind the shock, and close to the induced magnetosphere boundary, in agreement with theory and previous works. One interesting discovery is that the occurrence rate of mirror mode like waves is lower during higher solar EUV conditions, contrary to initial expectations. The proposed explanation for this seems reasonable.

The manuscript is well written with very few typos, and the figures are of high quality. I have some minor comments and questions noted below, but these should not affect the overall conclusions drawn from this work, and I expect that this manuscript will be suitable for publication once addressed.

AC1: *We warmly thank the reviewer for an in-depth evaluation of our manuscript. We strived to make all necessary corrections as per the reviewer’s suggestions, which we feel have substantially improved the manuscript’s clarity and impact. Please find below our point-by-point answers.*

Minor questions:

- Criterion 6 in the automated algorithm restricts event detection to the magnetosheath. The manuscript provides justification for this, and this question is more of a curiosity. Do you know how this “pre-selecting” affects your results? Clearly, you are unlikely to identify mirror mode like structures outside of the magnetosheath, but do you have an idea of how many “real” events you might be removing, that lie outside of the magnetosheath? Given the instability criterion noted in the manuscript, is the magnetosheath where you most expect mirror mode waves to be observed anyway?

AC1: *We thank the reviewer for their remark, which made us think more about those numbers. Event detection is indeed restricted to the magnetosheath and magnetosphere, because this is the place where mirror modes are most expected to occur, as explained in the introduction. This is due to the temperature anisotropy becoming large in the wake of the quasi-perpendicular shock or due to other local mechanisms in the magnetosheath. It is true that by discarding all events in the solar wind, we might also discard true mirror mode events (for example created by the pickup ion process in very specific magnetic field configurations), but the probability of this latter case to occur is indeed empirically low, when examining the data over sufficiently long times. In effect, mirror modes are mostly present downstream of the shock, whereas magnetic holes, isolated events in pressure balance that have similar but not decisively the same characteristics as mirror modes, are mostly observed in the solar wind. These isolated events and solar wind events were excluded from our study, as the nature of magnetic holes is still highly debated and not (always) necessarily linked to the mirror instability.*

Following Madanian et al. (2020, https://doi.org/10.1029/2019JA027198*), magnetic holes in the solar wind represent about 2 events per day, lasting about 20 s, that is, about 40 s of detection per day. If we count all events including those in the solar wind meeting our detection Criteria 1–4 (that is, not applying Criteria 5–6), we end up with about 120 s of events per day on average. Removing the 40 s found by Madanian et al. (2020), we end up with 80 s, a number quite comparable to our lower estimate of 70 s/day of MM-like detections in the magnetosheath. Moreover, as suggested later by the reviewer, and due to MAVEN spending about 30% of its time in the solar wind per day on average, our estimate of 68 and 120 s/day represent in fact an accumulation of detections over only 70% of a day. Correcting for this would increase the detection rate to ~100 s/day and 180 s/day on average.*

For clarity of exposure, we decided to apply this previous reasoning so that the paragraph starting with "Finally, if we assume that a typical structure lasts about 5-10 s on average" is changed to:

"Following Madanian et al. (2020) who used MAVEN measurements over a period of 3 months, magnetic holes in the solar wind represent about 2.1 LMHs/day, lasting about 20 s each, that is, a total of 40–50 s of LMH detection per day. This number can be tentatively compared to our results if we consider that applying Criteria 1–4 only on MAVEN's magnetometer data does not filter out solar wind LMHs from our MM-like event database. In that case, we obtain ~ 180 s of events per day. Removing the 50 s/day of LMHs, we end up with 130 s/day, a number marginally larger than our corrected (lower) estimate of 68 × 1.5 ~ 100 s of MM-like detections in the magnetosheath (Criteria 1–6). This would be statistically consistent with the hypothesis that the majority of the events captured in the solar wind are isolated events reminiscent of LMHs, as discussed when attempting to remove false positive detections (see Sect. 2.2.2). Conversely, if we assume that a typical MM-like structure lasts about 10 s on average (as in Simon Wedlund et al., 2022c), we end up with 100 s divided by 10 s, i.e., ≈ 10 MMs/day in Mars' magnetosheath."

- Line 173: the algorithm only accepts magnetic field structures that rotate by less than 10-20 degrees. Table 1 notes that this value is 10 degrees – so is this the value actually used? In addition, will this criterion still allow fast magnetosonic waves to be included in the identification (I think criteria 2-4 remove fast magnetosonic wave modes, which are typically circularly polarized)? The manuscript notes that variations in plasma density and magnetic field strength are anti-correlated for mirror mode waves (in contrast to correlated for fast magnetosonic waves) – but this characteristics cannot be taken advantage of when using solely magnetometer measurements.

  AC1: *Yes, we use the value of 10 degrees to constrain the rotation, in a manner similar to Lucek et al. Ann. Geophys. (1999). As remarked by the reviewer, fast mode-type (magnetosonic) waves are right-hand circularly polarised (Gary, JGR 97, 3103-3111, 1992). By limiting the rotation of the field over the structure, we are discarding most of the magnetosonic waves except for those that we encounter at the right "wave phase" in which they perform a slight rotation of the field over half a wavelength.*

- Line 180: the manuscript rejects isolated single events – do you require a certain number of oscillations to occur one after the other (a wave train), for example? I think lines 180-185 describe this process, but it's not quite clear to me how this is achieved. Can you explain this section or rephrase it?

  AC1: *We thank the reviewer for pointing this out. In our mitigation strategy, an isolated event is defined as a magnetic field dip or peak lasting no more than 1 s within a 30-s period of detections.*

This is because our detection criteria are applied to every 1-s magnetic field measurement, and so the whole of a MM structure is usually not detected (only parts of it), leading to an underestimate of the true duration of a MM-like structure. That said, if, for a long magnetic field dip lasting say 15 s, we only get one detected event of 1 s within a period of consecutive 30 s, we consider this detection as an isolated event, which we discard. If multiple detections occur during one 30-s period, we keep the events in question, which can be in effect either a large MM-like structure (less probable after a cursory examination of the datasets) or a train of shorter MM-like structures (more probable), depending on the length of the period they belong to. We reformulate this part to make it clearer to the reader, as well as change "region" to "detection period" to reflect the temporal nature of the dataset:

"First, within our original database of candidate MM-like detections, we define so-called `detection periods' containing all events consecutively detected, with two periods separated by a minimum of 30 s between one another. We discard periods of isolated singular events, defined as a detected event lasting no more than 1 s within a period. In this way, some of these discarded isolated events (DIE) are not part of the usual quasi-periodic train of MM structures, which would result in more than one detection within a period. They are instead more reminiscent of the so-called 'linear magnetic holes' (LMH), in the original definition of Turner et al. (1977), structures that we want to filter out from the database. In contrast, if multiple detections occur consecutively within a whole period, these events may represent in reality either a large MM-like structure or a train of shorter MM-like structures, depending on the length of the period they belong to."

- Lines 185-190: when you calculate the standard deviation of the change in magnetic azimuth and elevation angles, do you calculate this based on the full cadence (32 Hz) magnetic field dataset, or a filtered version of this? Essentially – do you calculate the standard deviation of the underlying average background field, so that you ignore any smaller scale, higher frequency rotations / features that may also be present, on top of the lower frequency mirror mode like structures?

  AC1: *We calculate the standard deviation of the azimuth/elevation angles based on the 1-s downsampled (averaged over 1 s) **B**-field measurements, thereby ignoring higher frequency features which are not at the same scale as the ion-scale mirror mode structures. Furthermore, we do not perform this standard deviation directly on the 1-s background field (Butterworth-filtered as explained in the paper), because the background field has a much larger scale size than the mirror mode structures we seek to characterise, but rather we calculate first the angles on the 1-s downsampled measurements and then calculate their moving standard deviation over a 2-min interval. Granted, if the running averages are performed on the exact same overlapping intervals, both calculations should give approximately the same result. However, we prefer to calculate the angles based on the non-background 1-s components and then calculate their standard deviation, in order to prevent an accidental downsampling of the final results. This will be clarified when introducing the expression for the angles. We suggest to alter the text with the following:*

  "First, within our original database of candidate MM-like detections, we define so-called `detection periods' containing all events consecutively detected, with two periods separated by a minimum of 30 s between one another. We discard periods of isolated singular events, defined as a detected event lasting no more than 1 s within a period. In this way, some of these discarded isolated events (DIE) are not part of the usual quasi-periodic train of MM structures, which would result in more than one detection within a period. They are instead more reminiscent of the so-called 'linear magnetic holes' (LMH), in the original definition of Turner et al. (1977), structures that we want to filter out from the database. In contrast, if multiple detections occur consecutively within a whole period, these events may represent either a large MM-like structure (less probable after a cursory examination of the dataset) or a train of shorter MM-like structures (more probable), depending on the length of the period they belong to."

- Line 224: In Figure 1 (and perhaps 2), there are some green lines that lie over data where the values of delta(B)/B and delta(N)/N are both positive (or both negative), which I believe means these quantities vary in phase (as opposed to anti-phase). Can you comment on this and why these events are still classified as mirror mode like? There's reference to this at line 229, but I'm curious why these events are still included, if these quantities are varying in phase?

  AC1: *The reviewer is correct, and indeed several detected structures appear to be in phase rather than antiphase at the resolution of the ion instrument shown. The purpose of the two figures is to show the success and caveats of the method, as explained in this section. Generally speaking, determining the correlation between B and N is not a straightforward issue and ideally requires a*

more careful case study as in Simon Wedlund et al. (2022c, JGR 127, e2021JA029811). One reason a discrepancy may exist here is that the detection is performed on the 1-s magnetic field datasets, and not on the 4- to 8-s downsampled dB/B and dN/N fluctuations that are shown in in the last panels of Figs. 1 and 2. Many of the finer B-field structures are hence smoothed out, including dips and peaks of a duration of the order of the resolution adopted. That said, an antiphase behaviour at that resolution is, in combination with the magnetic field criteria, a good sign that the structures are indeed mirror modes (see the first few oscillations around 00:19:30 on Fig. 1), whereas a phase behaviour does not rule this conclusion out entirely (see earlier oscillations before and later oscillations around 00:20 which show an apparent mix of phase and antiphase behaviours). This has to do with the scales probed by each instrument. A visual and cursory examination of the litigious cases using the electron analyser SWEA onboard densities (used as a first approximation of the plasma densities), which have a higher resolution than SWIA, shows a mostly in-phase behaviour with magnetic field strength (hence more reminiscent of fast mode-type 'magnetosonic' waves), with a few events appearing in antiphase (mirror modes proper). Later in the time series, some oscillations closer to 00:30 UT have a clearer B-N antiphase while not being captured by our algorithm due to a slight nonlinear polarisation. This further proves that B-field criteria only are not enough to unambiguously confirm these events as mirror modes, as discussed in f.ex. Simon Wedlund et al. (2022c), hence our taxonomy as "mirror-mode"-like. This is a well known consequence of this method of detection, as previously shown at Earth (see Lucek et al. Ann. Geophys., 1999; Genot et al. Ann. Geophys., 2009, etc.).
In the end, our approach is to count even the litigious detections as MM-like in this particular case, as they are difficult to filter out entirely in our statistical database, even with our quite stringent mitigation strategies.
Since the manuscript was submitted in July 2022, a complementary study has been started using calibrated ion and electron density moments to overcome these difficulties, in a manner similar as the recent study of Jin et al. (2022, https://iopscience.iop.org/article/10.3847/1538-4357/ac5f00*), which makes use of both ion and magnetometer instruments. This is the subject of a new paper to be submitted, which will show how both methods (B-field only, and B-field & B-N anti correlation) compare in a more quantitative way.*
For the moment, we add some clarification and caveat about those specific occurrences in the text and the ambiguities that remain:

*"The events in green display a mix of B-N in-phase and antiphase behaviours (bottom panel), with some of the shorter detected structures appearing in phase, a characteristic more reminiscent of fast mode-type (e.g., magnetosonic) modes; this is however difficult to unambiguously ascertain owing to SWIA's 8-s resolution which cannot capture these shorter magnetic field structures. Around 00:19:30 UT, some oscillations are clearly in antiphase, whereas before and after, they become in phase again. Closer to 00:30 UT, B-N fluctuations appear again in antiphase but are not captured by the detection algorithm, due to constraints on the linearity of the structures. This points further to the necessity of a more in-depth case study as in Simon Wedlund et al. (2022c) to ascertain the nature of the fluctuations captured by our algorithm and shows the limitations of a **B**-field only algorithm to detect MM structures. From a general point of view, the temperature anisotropies responsible for the generation of MMs are expected to take place in the wake of a quasi-perpendicular shock (see Hoilijoki et al. 2016). Consequently, we expect such a time series, together with that discussed in Simon Wedlund et al. (2022c) (24 December 2014 around 11:25 UT, also at high solar activity) or Ruhunusiri et al. (2015) (26 December 2014 around 15:00 UT in their Fig. 1), to harbour `textbook' examples of MM occurrences."*

- Lines 260: This is mostly a question of curiosity: if at Earth most mirror mode like waves are elliptically polarized, why does this study focus on linearly polarized ones?

  AC1: *We thank the reviewer for pointing this out. Génot et al. (2001) used a small set of mirror mode events (13 cases) from the AMPTE spacecraft to conclude that elliptical polarisation may occur more often than linear polarisation in certain conditions, a statement we oversimplified in our paper. Theoretically, single mirror modes are purely growing modes, which an elliptical polarisation would prevent, hence our choice of linear polarisation here. We modify the text to read:*

  "Theoretically, single MMs are purely growing modes, which an elliptical polarisation would prevent. However, as shown in studies made in the Earth's magnetosheath (Génot et al. 2001, 2009), MMs can also be elliptically polarised in certain conditions. [...] Therefore, our method may miss a certain (but difficult to estimate) amount of structures which are MMs but are non-linearly polarised in nature."

- Lines 314-315: typo: the second square parentheses in each pair need reversing direction. Which side of the dates are the inclusive ones here? For example, MY32 = [31 Jan 2013 – 18 Jun 2015], MY33 = [18 Jun 2015 – 05 May 2017]. Does 18 June 2015 fall in MY32 or MY33?

  AC1: *The inverse brackets were intended to show that the 18 June 2015 was partially not included in MY32. However, we realise this is an oversimplification. We decided to add the limiting hours when necessary. For example the limit between MY32 and MY33 is 18 June 2015 12:34 UT, which is accurate within 3 min, according to Allison & McEwen (2000): this is the limit we used in our calculations. Similarly, MY33→34 occurs on 5 May 2017 11:45 UT, MY34→35 on 23 March 2019 11:32 UT and the end of MY35 is on 7 February 2021 11:12 UT. This information will be added.*

- (Line 350: I look forward to reading Wedlund+ 2026, once MAVEN has sampled more MYs.)

  AC1: *We thank the reviewer for their humour and will certainly try to indulge them.*

- Line 393: Discussing possible reasons why the detection rates are smaller in the presented study than in Ruhunusiri + 2015. You mention that your data set covers a much longer time span than Ruhunusiri+2015. Could you do a quick chop of your dataset and plot results for just the time range covered by them? This isn't necessary, but it would be interesting to see if the results are consistent with this hypothesis.

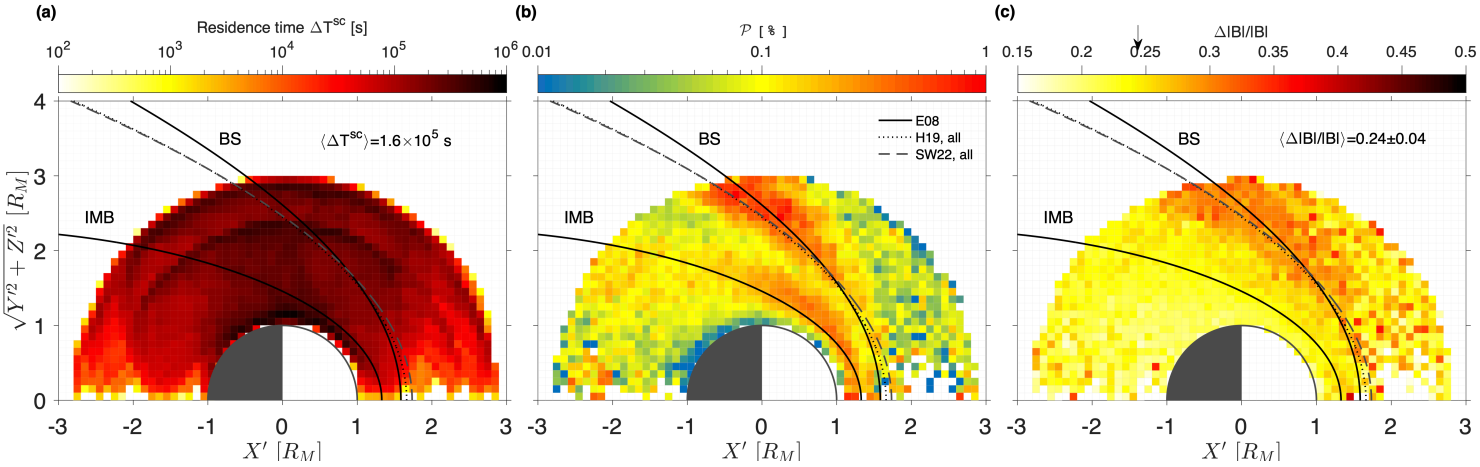
  AC1: *This is a good point and so we performed a new smaller analysis on the early dataset of Ruhunusiri et al. (2015), spanning 28 Nov. 2014 to 20 March 2015. We found 10,292 events in total, before applying Criteria 5-6, and obtaining 4,732 events. Unsurprisingly, the results are essentially the same as those presented in our MY32 discriminated dataset (Fig. 8a), since MY32 ends on 18 June 2015 and so we refer to that figure instead. The average value in the box is still around 0.1% (mostly yellow values) with maxima at 0.5% during that reduced dataset, which is a factor 10-20 less than Ruhunusiri's results, which were based on transport ratios downsampled at 4 s.*
  In this paragraph, we have also added a mention to another recent paper. Since the initial submission of our manuscript, a new study on the subject by Jin et al. (2022, ApJ) has indeed come to our attention (see Reviewer 2 comment for ll. 392-396). They use a similar technique as ours but with B-N antiphase behaviour included on 4 years of MAVEN data: their results are very similar to ours, with probabilities of less than 2% on average over the full volume of space. We change the text as follows to reflect this discussion:

  "Although our main detection regions are similar to those of Ruhunusiri et al. (2015), both in position and shape, we report here much lower absolute detection probabilities of MM-like structures (maximum of 0.8%). If we take into account the length of the datasets considered in each study, their results included about four months of observation during MY32, and thus are most comparable to our Fig. 8a. However, a quantitative comparison with the values of Ruhunusiri et al. (2015) appears extremely challenging at this stage. One reason is that the two detection methods are fundamentally different, ours using a B-field-only 1-s resolution at the expense of an ambiguity in the nature of the detected structure with a clear underestimate of the total duration of the found events, theirs using wave analysis techniques based on transport ratios with a cruder time resolution (4–8 s with a Fourier transform on consecutive 128 s intervals) looking for the mode producing the maximum of B-field power in each 128 s window. In that way, our quantitative results are more comparable to those of Jin et al. (2022), who recently found, with similar techniques as ours (but using the additional B–N antiphase behaviour), an occurrence rate of less than 2–5% on average over the first fours years of MAVEN data. The strategies we applied to help remove possible false positive detections may, to a certain extent, have filtered out legitimate events. Moreover, as explained in Section 2.3.1, the total duration of MM-like structures ΔT^{struct} is underestimated in our approach by more than 50% because of inherent limitations in the detection method. All points combined, this implies that our detection probability should be seen as a lower estimate (see Section 2.2.2)."

- Line 393: "trains of short events" – I'm not sure this has been explained fully, and refers back to my comment about line 180 above. In addressing that, you will probably address this.

  AC1: *We agree with the reviewer and clarify this part as our answer in l. 180 above. Otherwise, no other changes.*

- Line 396: the manuscript mentions that when false positives are removed, the automated algorithm may also remove some number of real events. If you skip this step (and keep all data), what do the results look like – are the detection rates consistent with the earlier studies, and the differences can indeed be explained by the proposed reasons? This again is not necessary (if difficult to do for example), but would be a first step to test this hypothesis.



AC1: *We thank the reviewer for this idea. When keeping all data unchanged, the spatial distribution changes only little, with two main loci of MM-like structures as before. This figure is reproduced below, in the same format as Fig. 3. The maximum occurrence probability is closer to 5%, with a significant number of events present in the solar wind and around the statistical average position of the shock, where the probability peaks. The detection probability in space and magnitude remains almost unchanged near the sunward IMB with respect to our final results with false positives removed, which suggests that single events do not occur often in this region. This conclusion is also consistent with considerations on the overall statistics when taking all events into account and correcting for the daily rates of magnetic holes in the solar wind, see our response to earlier AC1 "Criterion 6…". In contrast, the tail of the magnetosphere is having significantly higher probabilities. We decide not to include this figure in the manuscript to avoid cluttering it.*

- Line 399: Does the pre-selection of events in the magnetosheath play a role here? Did Ruhunusiri+2015 search for events throughout the entire magnetosphere, with no limits on spatial location?

  AC1: *The preselection of events within the downstream part of the shock (including magnetosheath and magnetosphere) should, to our knowledge, not affect the detection probabilities as we normalise the detection times to the residence time of the spacecraft in the spatial grid. Differently than us, Ruhunusiri et al. (2015) have no restrictions on the spatial location of the events, and include solar wind (where Alfvén waves are largely dominant), magnetosheath (mix of Alfvén and quasi-|| slow modes dominant) and magnetosphere (fast magnetosonic waves dominant). Our text here should stay unchanged.*

- Line 422: Is the 68s of structures detected per day averaged over the entire MAVEN orbit, or just when MAVEN is sampling the magnetosheath? Can you comment on how only allowing events to be counted in the magnetosheath may skew this number, based on how you normalize it "per orbit"? For example, if MAVEN spends 20 minutes of every 4.5 hour orbit sampling the magnetosheath, it can only ever detect mirror mode like events for 20 minutes per orbit at most, given the current constraint that events are only counted when MAVEN is in the magnetosheath.

  AC1: *This is a very good point which we had originally missed, and we are grateful to the reviewer for it. The histogram of Fig. 6 assumes a full day of observation, but because we have removed the events in the solar wind and MAVEN spends about 30% of its time on average in the solar wind (see, for example, the Figure 1 of Simon Wedlund et al. 2022c), all numbers should be compensated for this lack of temporal coverage by at least an equal number (multiplied by a factor 10/7 ~ 1.5 at least). To clarify this, we first add some statistics in Table 3 (see below).*

  Table 3 now gives the percentage of MM-like detections with respect to this time, whereas we also give the actual percentage time spent by MAVEN in the magnetosheath and magnetosphere (about 70% of the total time spent around Mars). It is now introduced as:

  "Table 3 presents the general statistics of our MM-like structure database, with a total of 176,041 events detected between 1 Nov. 2014 and 7 Feb. 2021 (last day of MY35), with a total residence time of the spacecraft in the magnetosheath and magnetosphere of about 4.4 Earth years, compared to a total orbiting time of 6.1 Earth years. For each controlling parameter we calculate the residence time of MAVEN inside the bow shock of Mars and the global observation ratio of MM-like events. By

Table 3. Statistics of MM-like event in the magnetosheath of Mars found from 1 Nov. 2014 to 7 February 2021 with the MAVEN/MAG instrument, following the detections performed in Section 2.2 and for different cases. N_{MM} represents the total number of MM-like events found (equivalent to a duration in s because of the magnetometer resolution of 1 s). ΔT^{sc}_{in} is the total duration that the spacecraft spent inside the bow shock of Mars during that time (excluding thus the time spent in the solar wind), whereas ΔT^{sc}_{tot} is the total time spent by MAVEN in the whole volume of space, given here as a comparison. Observation ratio $R_{MM} = N_{MM}/\Delta T^{sc}_{in}$ is the percentage of MM-like detections in the magnetosheath, and $O_{msh} = \Delta T^{sc}_{in}/\Delta T^{sc}_{tot}$, the proportion of the full orbit coverage that MAVEN effectively spends in the magnetosheath, are given in percentages.

Case	N_{MM}	ΔT^{sc}_{in} [s]	R_{MM} [%]	ΔT^{sc}_{tot}	O_{msh} [%]
MY32	14,285	12,519,185	0.114	18,005,770	69.5
MY33	58,584	41,047,441	0.143	59,120,895	69.4
MY34	50,315	39,389,107	0.128	59,188,581	66.5
MY35	52,857	46,748,279	0.113	56,662,768	82.5
EUV flux $\geq 2.77\,mW\,m^{-2}$	88,688	74,105,086	0.120	105,764,081	70.1
EUV flux $< 2.77\,mW\,m^{-2}$	87,353	65,598,926	0.133	87,213,933	75.2
Ls1 = [315° – 45°]	49,497	39,072,449	0.127	52,184,340	74.9
Ls2 = [45° – 135°]	56,591	38,690,683	0.146	51,082,098	75.7
Ls3 = [135° – 225°]	36,928	30,274,297	0.122	41,449,456	73.0
Ls4 = [225° – 315°]	33,025	31,666,583	0.104	48,262,120	65.6
All	176,041	139,704,012	0.126	192,978,014	72.4

contrast, we also calculate the total residence time of MAVEN and the proportion in percentage that MAVEN spends inside the bow shock during its orbiting time around Mars, using the fast bow shock detection of Simon Wedlund et al. (2022b) to pinpoint where the magnetosheath finishes and the solar wind starts in the individual orbits. During the time span covered, MAVEN remains about 70% of its orbiting time downstream of the bow shock. […]"

Then, we keep Figure 6 as it is, because this figure shows the actual number of seconds captured by our algorithm in MAVEN's data — that said, we make modifications to the text and the caption to explain this 30% underestimate "per day".

"Figure 6 displays the daily detection rate of MM-like detections inside the bow shock during the entire mission. The numbers quoted here represent an accumulation of the detected events in the magnetosheath over a full 24h of observation by MAVEN. However, during the time span considered here, MAVEN spent at most 30% of its time in the solar wind per orbit (see O_{msh} in Table 3), and so all magnetosheath detection rates quoted here should be multiplied by about a factor at least 10/7 ~ 1.5 to compensate for the absence of temporal coverage when in the solar wind. For simplicity, we will quote the numbers below as they are, and apply a corrective factor when generalising and comparing to other works."

- Around line 425: when discussing occurrence rates as a function of MY, did MAVEN sample the spatial extent of the magnetosheath evenly and equally in each MY? For example, if MAVEN sampled just the nose of the magnetosheath in one MY, and then just the tail magnetosheath in another, you would likely obtain different occurrence rates. I imagine it might be difficult to fully address given that you can only slice and dice the data so many ways while retaining good statistics, but can you comment on this at least?

  AC1: *We thank the reviewer for a nice suggestion and fully agree that these global numbers have a lot of conditions underlying them, one of which being that we assume an even distribution of the orbit throughout the Mars Years (MY). As the reviewer remarks, this was not entirely the case as can be seen from Figure 8, left column. That is why these numbers can only be a first indicator pending the more in-depth analysis later presented in the manuscript (Section 3.2). For example, MAVEN's orbit during MY32 does not cover the subsolar regions of the magnetosheath well, in contrast to the other MYs. In that way, the two MYs most comparable with respect to MAVEN orbits are MY33 and MY34, whereas MY35 MAVEN orbits were more compact, which may have altered the final result. We add the reviewer's suggestion that the spatial coverage is also adding to the reason why MY32 felt like an outlier (we stated only that the coverage was much less extended in time with respect to the other years). The new text is the following:*

"To further comparisons, we look at the evolution of this detection rate with respect to MY, assuming that MAVEN's orbit coverage of the magnetosheath was similar between MYs. The latter assumption is mostly fulfilled for MY33 and MY34 (as can be seen later in Fig. 8, left column), with similar orbits and a similar amount of time spent in Mars' environment, whereas MY32, and to a lesser extent MY33, have quite different spatio-temporal coverages. The mean daily detection rate over each MY changes little (green dotted line on the figure), with MY33 having more detections (85 detections/day in mostly high EUV flux) than any other year, and MY34 (73 detections/day in mostly low EUV flux) having less detections. As expected, MY32 seems to be a clear outlier due to a looser coverage around the subsolar magnetosheath and MAVEN probing only the later portion of the full MY. This suggests that we cannot compare absolute detection numbers between MYs without first normalising to the spacecraft's residence time during that period. Such a normalisation is performed and discussed in Sect. 3.2."

- Line 553 and Figure 10: bottom panel in particular: by eye I would have placed the orange fit further to the right. It seems like the extreme values on the left (large and negative) are skewing the fit to the left. If I integrate the black area by eye, there is significantly greater area to the right of the vertical dashed orange line, than to the left. Is this consistent with the negative reported skewness, which seems to be calculated in a way such that extreme outliers can bias it heavily?

  AC1: *We thank the reviewer for pointing this out. With the new PDFs calculated for accumulated residence times above 2 hrs for MAVEN in any grid cell (as discussed in the next question), the distributions have changed. We will rework the fits (exploring distributions other than the Gaussian distribution with skewness = 0 by definition) and recalculate the skewness for each revised distribution, as well as add the position of the mean/median/mode value to illustrate the value of the skewness better. We will show this in the revised manuscript and make the corresponding corrections in the text.*

- Figure 3, and all figures with statistical results: do you require a minimum number of data points to lie within each grid cell, to be displayed?

  AC1: *In our first submission, we did not have any constraints on the number of detections per grid cell to show the results. However, there was a small caveat due to the log10 scale for the deep blue colour which corresponds in effect to probabilities below 0.01%, with many of those coloured grid cells being in effect strictly zero. We have now removed the zero values from the representation here, which now are the same white colour as the 'no-data' background. Moreover, the reviewer's remark alerted us to the more general issue of reliable statistics in each grid cell, and we decided to constrain the probabilities of detecting MM-structures assuming MAVEN spends a minimum of 2 hours of cumulated duration in any grid cell. This is equivalent to half a full orbit of MAVEN, 7200 s, that is, about 250 times larger than the longest single structures at Mars (~30 s). The non-constrained full spacecraft coverage is, as before, shown in panels (a) of Fig.3 and first column of subsequent figures. The text is modified whenever needed and we add the following description when introducing Fig. 3 in Section 3.1:*

  "The spacecraft residence time ΔT_{sc} is then used as a normalising factor to calculate the probability P of detecting MM- like events (Fig. 3b), expressed here in percentage and in logarithmic scale. In this representation, we ignore all grid cells for which $\Delta T_{sc} < 2$ hr to ensure a good statistics throughout the grid: MAVEN thus stays a minimum cumulated time in each cell equivalent to about 250 times the duration of the longest single MM-like structure at Mars (~30 s). MM-like structures are mostly confined to two main regions, one in the immediate vicinity of the predicted shock (at SZA > 45 degrees), and one in the magnetosheath pressed against the IMB."

  And in the caption of Fig. 3:

  "[…] (b) Percentage occurrence of detecting MM-like structures P, for $\Delta T^{sc} \geq 2$ hr in any given grid cell. […] In panel (a), the average residence time in a grid cell $0.1 \times 0.1 R_M$, noted $\langle \Delta T^{sc} \rangle$ is given for reference. In panel (b), the average probability $\langle P \rangle$ in a grid cell, ignoring all grid cells for which $\Delta T^{sc} < 2$ hr (vertical arrow in the colour bar of panel a), is also given. The average positions of the 'nominal' bow shock (BS) and of the induced magnetospheric boundary (IMB) are given for reference as black continuous lines (Edberg et al., 2008, noted 'E08'), based on MGS data. Other bow shock positions representative of Mars Express and MAVEN datasets are in dotted lines (Hall et al., 2019, all points, noted 'H19') and as dashed lines (Simon Wedlund et al., 2022b, all points, noted 'SW22'). The average IMB and the BS lines roughly delimit the magnetosheath region; of

note, detections seemingly outside of this average bow shock on the picture are in reality always inside the shock for individual events. […]"

Mentions to this limitation in the representation are added when describing Figs. 8, 9, 11.

- Figure 7b: Should the red and blue lines overlap at all? The definition appears to split EUV conditions greater or less than 2.77mWm^-2, so I don't believe there should be any overlap?

  AC1: *This apparent overlap is an artifice of the histogram binning and bin size chosen since 2.77 mW/m2 is a two digit precision number, whereas bins of 0.1 mW/m2 were used here, which we wish to keep for an easy reading of the figure. Consequently, EUV values between 2.7 and 2.77 mW/m2 (low EUV conditions) will fall in the same bin as EUV values between 2.77 and 2.8 mW/m2 (high EUV conditions), that is, bin 2.7-2.8 mW/m2. No changes are thus needed here, although we will add a mention of this effect in the discussion, for clarity:*

  "Note that in Fig. 7b and f, the finite width of EUV (single precision compared to the double precision I_{EUV} threshold) and Ls bins results in expected overlaps at the borders between blue, red, orange and purple lines."

- Figures 8, 9, 11: for the right hand panels that show % difference, with values ranging from negative through zero to positive, I suggest using a colorbar where, for example, negative numbers are darker shades of blue, zero is white, and more positive numbers are red. This would allow the reader to more easily identify the different regions in those plots.

  AC1: *We thank the reviewer for their suggestion. We follow their advice and adopt the blue-to-red colour scale for the third column of panels in all relevant images, modifying the main text whenever needed.*

Answer to reviewers

Paper egusphere-2022-645: “Statistical distribution of mirror mode-like structures in the magnetosheaths of unmagnetised planets: 1. Mars as observed by the MAVEN spacecraft” by Simon Wedlund et al.

We would like to thank the two anonymous referees for their constructive comments, questions and suggestions. Please find our answers below, with all the points raised during the review addressed in the following.
Throughout, our responses are marked as “AC” (author comment AC1 or AC2 for reviewer 1 or 2) and in italic, whereas all other instances (non-italic text) are the original reviewer comments and questions.

Reviewer 2, Dec. 2022

Mirror Modes (MMs) are often observed in the magnetosheath of many solar system objects, like Earth, Venus, Mars and comets. This paper presents the statistical maps of mirror mode-like (MM) structures in the magnetosheath of Mars based on the magnetometer data only of MAVEN from November 2014 to February 2021 (MY32-MY35). Based on the magnetic field-only criteria and two mitigation strategies, 176,041 MM-like structures are detected in their final database. Then, they analyze the characteristics and calculate the detection probabilities of MM-like structures with respect to several controlling parameters (including Mars Years, EUV flux and Mars season, Ls) and map them in the magnetosheath of Mars. The results indicate MM-like structures appear in two main regions: behind the shock and close to the induced magnetospheric boundary. And the detection probabilities are higher with low solar EUV flux, which contradicts the prospect, explained by two combining effects: plasma beta effects and non-gyrotropy of pickup ion distributions. The detection method and mitigation strategies are well described and the properties of MMs are well shown by plenty of figures. Before I recommend accepting to publication in AG, below comments/ suggests needs to address.
AC2: *We thank the reviewer for their thorough review of our manuscript. Please find below our point-by-point answers.*

As shown in the introduction, MMs likely share a common ancestor with magnetic holes (MHs); statistical characteristics of magnetic holes have already been analyzed in the Martain magnetosheath. For instance, Huang et al. (2021, ApJ, 922:107) analyzed the distribution and parameter characteristics of kinetic-size magnetic holes (KSMHs) in the Martain magnetosheath using MAVEN’s data and tries to explain the mechanism of KSMHs as electron vortex. These results may also shed light on the mechanism of MMs. The authors should try to compare the properties between MMs and MHs to figure out the relationship between these structures and their mechanisms and their differences.
AC2: *We thank the reviewer for their suggestion. Indeed, several theories connect mirror modes with magnetic holes, although this link is still highly debated. While KSMHs have very different scales than the structures detected in our paper, investigating the connection between ion-scale MHs and MMs would be indeed interesting. It is however outside of the scope of our current paper, and we leave this excellent idea for a coming study on their relationship.*

- Line 185-186: “moreover this ensures that rotations could be calculated for trains of MM-like structures for which the 2-min windowed background magnetic field values would be representative.” This sentence is complicated and vague.

  > **AC2:** *We decided to remove this sentence altogether, as it is implicitly contained in the whole paragraph discussion. The point was that the 30 s duration between detection periods (formerly “regions” in this paragraph) ensured that trains of mirror modes (as opposed to isolated structures) were collected into one detection period per se, so that the 2-min averaged field was more representative of their background environment as a whole.*

- Line 245-250: “Conversely, we also expect the method to keep structures that are likely not MMs but situated in the magnetosheath (with B and N in phase).” What do the authors want to express by this

sentence? Besides, the authors try to keep the isolated structures out of the final database to make the results accurate. Now that, what does this sentence "As a consequence, on this criterion only (isolated structure), the frequency of MM-like detections in our final database could be underestimated by at least 10%." mean? Should the isolated structures be included in the database or not?

> **AC2:** *We have clarified this statement, which was intended to recall how our algorithm may discard some real events because they are seen as isolated by the algorithm and keep some others which are more dubious, as emphasised in the discussion of Fig. 1. The last part of the paragraph means exactly what is written: in the final database, we decide to remove all isolated events indiscriminately, although some are expected to be real mirror modes, just not captured accurately by the algorithm. The point of the next sentence is only to evaluate the maximum uncertainty for the final numbers presented here, in case ALL isolated events are real (which they are not). After some further tests, we estimate that we might wrongly discard about 5-10% of the original database. The text now reads:*
>
> *"This shows that although the detection method using Criteria 1--6 appears quite apt at detecting regions where MM-like structures are present and removing structures that are clearly not MMs, it may also ignore promising candidates (especially around 18:30--18:34~UT). Conversely, as already mentioned in Fig. 2, we expect the method to also keep structures that are likely not MMs although situated in the magnetosheath but with B and N in phase. […] As a consequence, on this criterion only (isolated event), we estimate that the frequency of true MM-like detections in our final database could be underestimated by about 10%."*

- Line 277-278: "…, which we expect to be about 25%, as evaluated from visual comparisons in a subset of events (see Figs. 1 and 2)." How do the authors evaluate the extent of the underestimation and from which subsets of events?

  > **AC2:** *We have revised this number by adding new cases. We looked at all events on Fig. 1 and especially Fig. 2 containing the clearest structures, which are clearly mirror mode-like (compressional, linearly polarised and B-N antiphase). Following the methodology presented in Simon Wedlund et al. (2022c, JGR 127, e2021JA029811), we first identified their start and beginning (start time of dip and end time of dip) with respect to background magnetic field fluctuations. We then compared the total number duration of the structures in the time span covered in the figures (± 15 minutes around the detections) with the total number of seconds captured by our algorithm. The difference was about 25-30% on average for these two figures. This is within the ballpark of the numbers quoted in Simon Wedlund et al. (2022c), with 33 s out of 78 s of visually picked structures captured by the algorithm (57%), when considering that the algorithm can capture simultaneously the ascending and descending parts of a full sinusoidal oscillation, even though we would count that visually as one structure. Since the time the paper was originally submitted, we have looked at several more cases to define this number better and found that we underestimate the total duration of confirmed MM structures by about 50% on average. This sentence now reads:*
  >
  > *"Following Simon Wedlund et al. (2022c), we evaluate such underestimate to at least 50%, based on visual comparisons in the subset of events presented in Fig. 1, Fig. 2 and in Simon Wedlund et al. (2022c), who, with the same detection algorithm, captured 33 out of a total of 77 s of visually identified MM structures."*

- Line 392-396: "Secondly, the two detection methods differ significantly: our B-field-only criteria detection permits us to capture trains of short events with 1-s resolution at the expense of an ambiguity in the nature of the detected structure, whereas Ruhunusiri et al. (2015) used wave analysis techniques based on transport ratios with a cruder time resolution (4–8 s with a Fourier transform on consecutive 128 s intervals)." The authors try to compare the detection probabilities between this work and Ruhunusiri et al. (2015) and explain the lower estimate due to the detection method. However, the lower resolution in Ruhunusiri et al. (2015) could also lead to the omission of smaller structures, resulting in a lower estimate for their results. It's hard to make a comparison and draw a conclusion in this way.

  > **AC2:** *We thank the reviewer for their remark and agree that it is a difficult task to quantitatively compare both approaches, but since only a few other studies of mirror modes are available at Mars, our intention was to contrast our results, at least qualitatively, with first those of Ruhunusiri et al. (2015), who use a complementary technique to detect MM waves, but with lower time resolution. Both works may look at different aspects of MMs statistically, as the maximum of B-field power in any given mode is taken over 128 s in total out of the number of observations in that time (also a multiple of 128 s). Hence counting the occurrences of each wave mode can result in significant differences overall. We have simplified the discussion in the text only stating the obvious difference*

of different detection techniques, as indeed, it is difficult to justify the increase or decrease of occurrence rates, one way or the other. We also quote a new study that has since come to our attention (Jin et al. ApJ 2022), who used a similar technique as ours and obtained very commensurable results to ours. The text now is hopefully clearer and to the point, and reads (see also our response to Reviewer 1, l. 393):

"Although our main detection regions are similar to those of Ruhunusiri et al. (2015), both in position and shape, we report here much lower absolute detection probabilities of MM-like structures (maximum of 0.8%). If we take into account the length of the datasets considered in each study, their results included about four months of observation during MY32, and thus are most comparable to our Fig. 8a. However, a quantitative comparison with the values of Ruhunusiri et al. (2015) appears extremely challenging at this stage. One reason is that the two detection methods are fundamentally different: our use a B-field-only 1-s resolution at the expense of an ambiguity in the nature of the detected structure with a clear underestimate of the total duration of the found events, theirs use wave analysis techniques based on transport ratios with a cruder time resolution (4–8 s with a Fourier transform on consecutive 128 s intervals) looking for the mode producing the maximum of B-field power in each 128 s window. In that way, our quantitative results are more comparable to those of Jin et al. (2022), who recently found, with similar techniques as ours (but using the additional B–N antiphase behaviour), an occurrence rate of less than 2% on average over the first fours years of MAVEN data. The strategies we applied to help remove possible false positive detections may, to a certain extent, have filtered out legitimate events. Moreover, as explained in Section 2.3.1, the total duration of MM-like structures ΔT^{struct} is underestimated in our approach by more than 50% because of inherent limitations in the detection method. All points combined, this implies that our detection probability should be seen as a lower estimate (see Section 2.2.2)."

- Line 427-428: "…, we cannot compare absolute detection numbers with the other MYs without first normalising to the spacecraft's residence time during that period." The authors could try to calculate the detection probabilities of each Mars year, so far as the absolute detection numbers and the spacecraft's residence time can be obtained, as shown in Fig. 8.

  AC2: *In Table 3, we calculated the detection probability for every Mars Year, throughout the entire volume of space, which we think is what the reviewer is asking here. However, following also comments from Reviewer 1, we decided to calculate the ratios of Table 3 by normalising the detection durations to the time spent by MAVEN in the magnetosheath and magnetosphere (about 30% of the total orbiting time), at the exclusion of the solar wind. To easily compare, we also included the total time spent in the volume of space by MAVEN. That said, our point in this paragraph could have been made clearer and so we decided, also following Reviewer 1 comment l. 425, to change that particular text to:*

  ***"**As expected, MY32 seems to be a clear outlier due to a looser coverage around the subsolar magnetosheath and MAVEN probing only the later portion of the full MY. This suggests that we cannot compare absolute detection numbers between MYs without first normalising to the spacecraft's residence time during that period. Such a normalisation is performed and discussed in Sect. 3.2."*

  We refer to our responses to Reviewer 1 for added clarifications.

- Table3: "NMM represents the total number of MM-like events found (equivalent to a duration in s because of the magnetometer resolution of 1 s)".
  Line 419-421: "On average throughout MY32 to MY35 with MAVEN, we find $\langle N \rangle = 68 \pm 43$ structures per day (ignoring single isolated 1-s events, see Sect. 2.2.2) fulfilling the criteria of Table 1."
  Line 429-430: "Finally, if we assume that most structures last 5–10s on average, we end up with 68/10–68/5 ≈ 7–14 MMs/day in Mars' magnetosheath." What's the definition of an MM-like event, a structure and a MM? Are they the same or not? What is the relationship between them?

  AC2: *This is a very good point and we thank the reviewer for allowing us to clarify our definitions. We have strived to homogenise the vocabulary throughout the entire manuscript, with "events/ detected events" = actual detections (not necessarily imaging the full extent of a dip or peak MM-like structure, as already explained in the text), and "structures" referring to the whole MM-like fluctuations (that is, a dip or a peak, or category "other", as defined in previous works such as Joy et al., 2006 at Jupiter). "Event" is thus used exclusively as the result of the detections using the B-field-only criteria, whereas "structure" is left for observations and theoretical concepts, in a more statistical sense (for example the maps showing the probability of detected MM-like structures, once*

the individual events have been accumulated into a grid cell). We have added this mention in the introduction of Sect. 2.2:

"Moreover, we refer in the following to the detections as `events' as we go through each 1 s of magnetic field data, whereas `structures' refer to a MM-like fluctuation as a whole (a dip or a peak, or a mix of them, as in Joy et al. 2006), which may contain several detected events. When accumulating detection events in a statistical spatial grid, the detection probability will simply be referred to as 'probability of MM-like structures'."

- Line 445: "… , with the threshold <I> just above the irradiance local peaks during MY34 and MY35." The authors should define the threshold <I> when introducing the EUV flux levels in Line 316-322.

  AC2: *The first occurrence is indeed on line 317 in the text and was only defined in the caption of Fig. 4. We add the mention:*

  "This limit, noted <I> in the following, is the median of the EUV flux in the 2014-2021 period […]."

  as well as in the caption of Fig. 7:

  "[…]; In panels (b) and (e), the median value of the EUV flux $\langle I \rangle = 2.77\ mW\ m^{-2}$; […]."

- Line 537-540: "As the exosphere expands with increasing EUV flux, the obstacle to the solar wind flow grows in size, with the bow shock and IMB both swelling up. This is illustrated in Fig. 9 by the dashed bow shock curves of Simon Wedlund et al. (2022b) and how they compare to the fixed curves of Hall et al. (2019)." The bow shock and IMB will both swell up with increased EUV flux. However, in Fig. 9, the dashed lines only show the expanding bow shock, but the modeled boundary of IMB is fixed. The authors should describe it more clearer.

  AC2: *We thank the reviewer for this point. The following text modification will remedy this:*

  "As the exosphere expands with increasing EUV flux, the obstacle to the solar wind flow grows in size, with the bow shock and IMB both swelling up. The swelling of the bow shock is illustrated in Fig. 9 by comparing the dashed shock curves of Simon Wedlund et al. (2022b) in panels a and b, representing 'EUV low' and 'EUV high' conditions, respectively, with the fixed curves of Hall et al. (2019)."

  To make sure that this is understood in the figures as well, we have modified the captions accordingly, adding that the IMB positions of Edberg et al. 2008 were measured for high EUV fluxes, when relevant.

- Figs. 8, 9 & 11: As we can see from those figures, the detection probability P may reach artificially low values, especially in the region in front of the modeled bow shock and in the tail. However, when calculating the difference from the total detection probability âP/Ptot, the data coverage is quite different among these figures. Do the authors remove the small value when plotting the âP panels? What is the criteria the authors chose to remove these data points? Why not discard the data points with low P since they make little sense?

  AC2: *We thank the reviewer for their comment. In our original manuscript, there was a small caveat due to the log10 scale for the deep blue colour which corresponds in effect to probabilities below 0.01%, with many of those coloured grid cells being in effect strictly zero. We have now removed the zero values from the representation here, which now are the same white colour as the no-data background. The full spacecraft coverage is, as before, shown in panel (a) of Fig.3 and first panels of subsequent figures. Following a comment by reviewer 1, we now only show in the second panels the probability of detecting MM-like structures in grid cells where the spatial coverage is at least 2 hours of cumulated observations from MAVEN (that is, 250 times larger than the largest MM-like structure observed at Mars). This effectively removes the more uncertain calculations from the representation.*

Minor Issues:

- Line 8-9: I suppose that the time range of the dataset should be corrected from "February 2020" to "February 2021".

  AC2: *We thank the reviewer for noticing this typo. It is now corrected.*

- Line 314-315: "For reference, MY32 = [] … ." There is an error in the representation of parentheses.

AC2: *We have added the exact times for each beginning and end of MY for clarity, as also suggested by reviewer 1.*

- Line 602: "… (see also Fig. 7g and Table 3)." There is no Fig. 7g. Please check it.

  AC2: *Again, we thank the reviewer for noticing this typo. We meant Fig. 7f.*

- Line 628-629: Add "are used" at the end of the sentence.

  AC2: *We do not think it necessary to add anything there.*