We would like to thank the reviewer for their careful consideration of the manuscript and their valuable comments. Our responses are provided below.

Comment 1: *Lines 585-645: My main concern is how well the curlometer current J and the JxB force can be used to describe this dynamical situation. Because of the large inter-spacecraft separation of the spacecraft, the estimates of these quantities are averages over a large volume.*

Response 1: We fully acknowledge that the separation of the spacecraft has implications for the interpretation of the curlometer analysis. As such, we very consciously draw only qualitative conclusions regarding the $\boldsymbol{J} \times \boldsymbol{B}$ control of the flows, in relation to how the measured estimates compare to modelled values derived from the expected background field. We are confident that the extent to which we interpret these estimates is reasonable. We are reassured through discussions with Malcolm Dunlop (Dunlop et al., curlometer technique papers, 1988, 2002; and as acknowledged in the manuscript) regarding our application of the curlometer technique in case of large inter-spacecraft separation. He confirmed that although large, the s/c configuration looks acceptable. The estimate should be stable, but as the reviewer rightly says, only of the average of a large volume.

Comment 2: *The flapping of the current sheet is observed only in one part of this volume. Should one compute these quantities specifically for C1 if that would be possible? If one assumes that the computed current J is stable and represents the current over the region covered by the Cluster tetrahedron, would it be reasonable to compute the JxB force using that J and then the B field measured only by C1? That would be a more local estimate for the JxB force at the C1 position. The authors could compute that and compare to the present estimate.*

Response 2: We agree that the flapping of the current sheet is only observed in one part of the tetrahedron volume, given that C1 is the only of the Cluster spacecraft to observe this. The reviewer puts forward an interesting and helpful suggestion to compute the $\boldsymbol{J} \times \boldsymbol{B}$ force using $\boldsymbol{J}$ (as calculated from all 4 SC), but using $\boldsymbol{B}$ only measured by C1 (instead of averaging $\boldsymbol{B}$ across the four SC). The results of this analysis are presented in Figure R1, where in panel vi) one may observe the results of calculating $\boldsymbol{J} \times \boldsymbol{B}$ using the measured magnetic field of C1 (solid line) and the modelled magnetic field of C1 (dashed line).

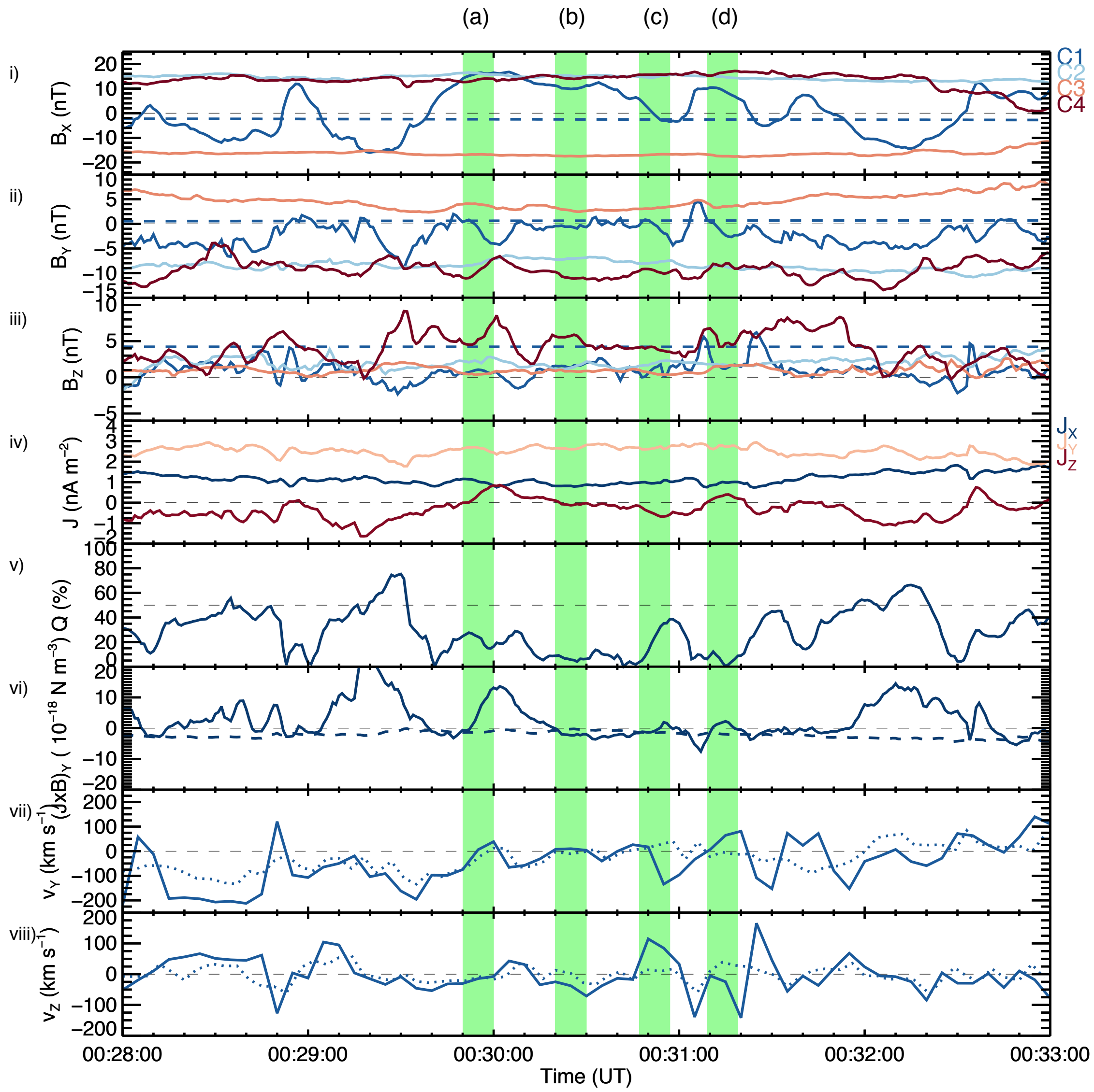

Figure R1: As in Figure 6 in the manuscript, but with $(\boldsymbol{J} \times \boldsymbol{B})_Y$ calculated using $\boldsymbol{B}$ from C1 only.

Firstly, consistent with the original Figure 6, panel vi), as perhaps to be expected, the 'model' $\boldsymbol{J} \times \boldsymbol{B}$ force (dashed line, where $\boldsymbol{J}$ has been calculated using the model C1 field and the true fields from C2-C4) is still weakly dawnward, consistent with the background 'curvature' of the magnetic field at the pre-midnight location as suggested in the original manuscript (note the different y-scale compared to the original panel vi). Particularly different, however, is the magnitude of the newly calculated $\boldsymbol{J} \times \boldsymbol{B}$ force. In the older analysis, the calculated $\boldsymbol{J} \times \boldsymbol{B}$ force was always positive with respect to the model $\boldsymbol{J} \times \boldsymbol{B}$ force, but still net negative for most of the interval. In this newer analysis, the $\boldsymbol{J} \times \boldsymbol{B}$ force is still mostly positive with respect to the model $\boldsymbol{J} \times \boldsymbol{B}$ force but is now also mostly net positive. If one takes the original panel vi) and subtracts the model $(\boldsymbol{J} \times \boldsymbol{B})_Y$ from $(\boldsymbol{J} \times \boldsymbol{B})_Y$, the result would be similar to $(\boldsymbol{J} \times \boldsymbol{B})_Y$ in the new analysis. This suggests that just using $\boldsymbol{B}$

from C1 in calculating $(\boldsymbol{J} \times \boldsymbol{B})_y$, rather than averaging across the four SC, has reduced the effects of the larger-scale background field curvature (incorporated by including the other SC). Compared with Fig. 6, the perturbations to $(\boldsymbol{J} \times \boldsymbol{B})_y$ are now also much larger in magnitude (previously they ranged between +/- 4 Nm^{-3}, but now they range between around -8 and 25 Nm^{-3}). We suggest that this is related to the fact that we are now using the (highly variable) magnetic field of C1 to provide the $\boldsymbol{B}$ in our computation of $\boldsymbol{J} \times \boldsymbol{B}$, as opposed to an averaged $\boldsymbol{B}$ – which, previously, was much steadier and closer to 0 in magnitude.

Previously, we drew attention to two key features of this figure. Firstly, we argued that the perturbations to $(\boldsymbol{J} \times \boldsymbol{B})_y$ were mostly associated with the magnetic field perturbations observed by C1. Inherently, this is even more apparent now. Secondly, we argued that 'the dawnward flow bursts tend to occur when $(\boldsymbol{J} \times \boldsymbol{B})_y$ is more negative, with the weak duskward flow bursts occurring when $(\boldsymbol{J} \times \boldsymbol{B})_y$ is less negative'. The new analysis suggests a slight adjustment to this interpretation is necessary. The dynamics evident in panels (vii) and (viii) now appear to be almost always associated with positive (duskward) enhancements in $(\boldsymbol{J} \times \boldsymbol{B})_y$, in contrast to the background (model) dawnward sense of $(\boldsymbol{J} \times \boldsymbol{B})_y$. This suggests that there is less variability in the direction of the $(\boldsymbol{J} \times \boldsymbol{B})_y$ perturbations, and that instead, the dynamic behaviour of $(\boldsymbol{J} \times \boldsymbol{B})_y$ is simply consistent with localised kinks in the magnetic field that are associated with the transient perturbations to the dawn-dusk flow. In fact, this is now more consistent with our "cartoon" interpretation presented in Fig. 7. Overall, if we were to proceed with this new version of the analysis, it will not alter our fundamental conclusions, in which we already acknowledge the uncertainly in trying to make detailed one-to-one association between the $(\boldsymbol{J} \times \boldsymbol{B})_y$ and flow perturbations. We consider that both approaches involve assumptions that limit the extent of the interpretation. Perhaps a wise solution would be to present both approximations in a new Fig. 6, and to highlight that the conclusions we have drawn are supported by both?

Comment 3: *Lines 525-532, Summary and Abstract, lines 106-107: Second, the typical extents of the IMF* B_y *penetration that is overriding the tail field line flaring (and causing tail magnetic field line twisting) in the case of clearly nonzero IMF* B_y *(IMF* $|B_y| > 3$ *nT) can be seen in Figure 2 of Pitkänen et al. (2019, GRL). Their Figure 2a and 2b show that under clearly positive IMF* B_y *conditions, the (slow) earthward convection is expected to be on average duskward both above and below the neutral sheet at the position of the Cluster spacecraft of the present manuscript. The tail magnetic field in this position is expected to be governed by the flaring. In the case of the present manuscript, the magnitude of positive IMF* B_y *was mostly less than +3 nT. Therefore, the global flow pattern in the magnetotail could be assumed to be even less asymmetric and the tail field line twisting occur at smaller extents than in Figure 2 of Pitkänen et al. (2019). The Cluster magnetic field data (C2-C4 data) clearly demonstrate the appearance of the field line flaring in the case of the present manuscript and not the twisting of the field lines due to IMF* B_y *influence. Furthermore, I think that while model results, Figure 5b in the present manuscript nicely demonstrates the spatial limits of the IMF* B_y *penetration to twist the tail magnetic field lines. The authors* *could modify the Summary section and add there that in this event, the IMF B_y influence in the position of Cluster was not strong enough to twist the magnetic field lines and the measured flows were associated with the localized magnetic field perturbation. So, the current sheet flapping was not overriding the IMF B_y control, because the control did not exist at the location of Cluster. Also then the end of the abstract (and the text elsewhere where IMF B_y overriding is discussed) would need to be modified.*

Response 3: We agree with the reviewer that Fig. 2 of Pitkänen et al. (2019) nicely illustrates the extents of the IMF B_y penetration, and that at the location of the Cluster spacecraft, the convection is expected to have a duskward component both above and below the neutral sheet. Of course, in our study, this is not what we see; the flow observed by Cluster during the flapping interval tended to have a clear dawnward component in the southern hemisphere, in disagreement with the spatially-averaged picture of slower (< 200 kms^{-1}) flow presented by Pitkänen et al. (2019). The local B_y that C1 observed during the flapping interval was also mostly negative (irrespective of hemisphere), also inconsistent with the average picture of Pitkänen et al. (2019).

We also agree with the reviewer that the C2-C4 data demonstrate field line flaring. However, we think that the suggestion that these observations do not indicate the presence of field line 'twisting' due to IMF B_y influence is a bit ambiguous. If, by 'twisting', the reviewer means that this IMF $B_y > 0$ perturbation was unable to change the sign of the (expected) $B_y < 0$ field in the pre-midnight northern hemisphere (e.g. at the location of C2 and C4), then the reviewer is absolutely correct. However, as we noted in the manuscript (lines 225-227, 520-525), and in relation to the discussion of Fig. 5, the spacecraft observing $B_x = 0$ and $B_y = 0$ just prior to the flapping interval was, in itself, indicative of IMF $B_y > 0$ penetration.

In relation to the convection: the flows that Cluster observed (locally) could have been expected in the case of a situation where we had IMF $B_y < 0$ (lines 351-355). However, the IMF and SuperDARN data allowed us to confirm an absence of any IMF $B_y < 0$ (and in-fact, the large-scale picture was one which seemed consistent with IMF $B_y > 0$). As the reviewer therefore suggests, the observed flows must have been associated with the localized magnetic field perturbations in B_y (lines 280-281) and the current sheet flapping, and could not be explained by IMF B_y control. Clearly, any IMF $B_y > 0$ associated perturbation at the location of Cluster was not significant enough to control (or perhaps override the flapping-related control) the dusk-dawn flow. We therefore agree with the reviewer on this point, and have reworded where applicable in the manuscript: rather than the 'flapping overriding the IMF $B_y > 0$ control of the flow', we now suggest that the 'IMF $B_y > 0$ penetration at the Cluster location was unable to override the variable dusk-dawn flow associated with the flapping'.

Comment 4: *Line 401: Maybe write here "the Harang reversal" instead of "the Harang discontinuity", because the authors are investigating flows.*

Response 4: We will amend this to ‘Harang reversal’ in the revised manuscript.

Comment 5: *Line 701: Which plasma sheet magnetic field observations the authors do mean here? The TA15 model results?*

Response 5: Here, we are referring to how e.g. prior to the flapping interval the SC tended to observe $B_y = 0$ at $B_x = 0$, which through our discussion related to Fig. 5 (the model results) we used to show was an effect of IMF $B_y > 0$ penetration. We will reword this to be clearer: “The IMF, ionospheric convection, and comparison of the plasma sheet magnetic field observations to the TA15 model field, all lead to…”