AGU PUBLICATIONS

Journal of Geophysical Research Space Physics

Supporting Information for

Classifying Magnetosheath Jets using MMS - Statistical Properties

Savvas Raptis[1], Tomas Karlsson[1], Ferdinand Plaschke[2], Anita Kullen[1], Per-Arne L. Lindqvist[1]

[1]Space and Plasma Physics, School of Electrical Engineering and Computer Science, KTH Royal Institute of Technology, Stockholm, Sweden

[2]Space Research Institute, Austrian Academy of Sciences, Graz, Austria

Contents of this file

Additional supporting information (Files uploaded separately)

Introduction

The supporting information consists of:

- (Text S1): A detailed description of the algorithm used for the classification of the jets used in the analysis of the main paper. The purpose of this text is to inform the reader of the details of the procedure not given in the appendix.
- (Figure S2): A detailed flowchart to be read along with text S1 for a detailed step-by-step guide through the algorithm used for the classification of jets.
- (Dataset S3): A full table of the dataset that was primarily used for the analysis (See Table 3 on the main paper) is included.

Text S1.

As described in the main article in subsection 3.2, we first identified 8499 jets from MMS1 measurements between May 2015 and May 2019 according to the criteria shown in Equations (1) and (3) in the main article.

These are then filtered to remove 'bad events' and sorted into the different classes (Qpar, Qperp, Boundary, and Encapsulated jets) according to the algorithm, described here and in the flow chart (Figure S2).

Data Pre-process:

This initial stage consists of finding cases of "Missing data" (Class 8) and "Border" (Class 7) jets from the 8499 unclassified cases. Class 7 jets are those found close to a magnetopause or a bow shock crossing.

As shown in Eq. 3 of the main article, the initial necessary information for the classification of a jet contains the pre-jet, jet and post-jet periods. Therefore, the first step is to find jets containing unreliable measurements within these periods, to remove them from the classification process. These jets correspond to the Class 8.

The second class removed in the initial pre-process is class 7 ("Border jets") which corresponds to jets found very close to a magnetopause or a bow shock crossing. These jets are found by checking whether a crossing was observed up to 5 minutes before or after the jet. If so, these jets are removed from the dataset. All the crossings were found from an automatic procedure that is also used to find where MMS resides in magnetosheath measurements (See subsection 3.1 on the main article).

These procedures remove 45 (Class 8) and 1346 (Class 7) jets. The rest of the database is filtered with the help of the following stages to determine the different jet classes and provide a sufficiently large sample to conduct statistical analysis.

Each of the remaining jets is moved to the next stages of the algorithm until is classified into one of the main classes. The main classes are the Qpar, Qperp, boundary and encapsulated jets (Table 3). If a jet is not classified in these 5 stages it is automatically considered "Unclassified" (Class 0)

Stage 1 – Initial Classification:

The first stage corresponds to a non-iterative algorithm that tries to directly classify jets to one of the main classes.

This is done by applying the thresholds described in Table 3 while using the pre/post jet time shown in Eq.6 of the main article. In particular, the code assigns a characterization for the three periods (Pre, jet, post) and then depending on these three values determines the class of the jet.

Firstly, the rules N.1 are applied. If the jet is not classified then, by using N.2, the algorithm determines whether there is a good indication that the jet can be classified in a future stage. These rules are used to determine if at least 1 period for possible boundary jets or 2 periods for possible encapsulated jets are not characterized as unknown (class 0). If so, these jets are moved to the next stages for further process.

If a jet is found to have all its corresponding periods (pre, jet, post) classified as "unknown" (class 0) then the whole is moved to the unclassified category and is not analyzed furtherly. This stage is the most robust and works very well for Quasi-parallel (Class 1) and Quasi-perpendicular (Class 2) jets. However, while allowing some cases of boundary and encapsulated jets to be classified, the majority of these jets were moved temporarily to classes 4 and 6 to be further processed in later stages and get possibly classified.

Stage 2 – Adjusting pre/post time:

In the second stage, the pre/post time of each jet that was not classified previously is changed.

The adjustment that takes place is of two different variations. The first one that is applied is to move the pre and post time period by 1/2 of its value backward and forward in time respectively. After doing that, we try to classify the jets once more.

At first, the algorithm determines if 4/5 of the total measurements of the whole period (Including pre-time, jet time and post time) correspond to either quasi-parallel or quasi-perpendicular plasma (Rules N.3). If so, we classify the jet to its corresponding class of Qpar or Qperp jet. This addition compared to the previous stage was done to avoid misclassification cases that could result from the variance of the pre-jet and post-jet periods.

The same rules as stage are then applied to determine if a jet belongs to one of the main classes. The only difference originates from the adjustment of the pre and post jet time periods.

The above variation is repeated 6 times, with each iteration adjusting the pre-jet and post-jet time further away from the jet by 1 measurement (4.5 seconds).

If a jet fails to be classified with the above variation, another one is used. Specifically, the algorithm takes up to a 30% increase of the initial time and up to 30% decrease to account for individual variations per jet that were possibly not accurately captured in Eq. 6.

Once more, the procedure follows the method described in Stage 1. Therefore, in total 6 tries for variation A of time adjustment and 6 more tries of Variation B are applied. If a jet remains in classes 4 and 6 it is moved to the next stage.

Stage 3 – Changing average time window:

In the third stage, the same adjustments of the pre and post jet time periods are used, while changing the thresholds that are required to be satisfied.

In all previous stages, we have used a 60-second average window for the magnetic field and a 30 second one for the rest of the quantities. However, doing so, we filtered out small time scale changed that are useful to determine boundary and encapsulates cases. As a result, as shown in Table A1 of the main article (second row), a new set of thresholds is used corresponding to different smoothing of the quantities. In particular, a 30-second window is now used for the magnetic field while the rest of the quantities remain as originally obtained from the MMS.

This stage was effective in finding a few more cases of Boundary (Class 3) and Encapsulated (Class 5) jets. Most importantly, it finalizes the dataset for the Qpar and the Qperp jets.

At this point, it was found that both Qpar and Qperp jets that fit the necessary and the extra criteria (Table A1 and discussion in appendix) have a large enough sample to treat them statistically. As a result, to avoid any false-positive cases, we stop searching for classes 1 and 2 and we keep the jets that reached stage 3 as our final sample for these two classes.

It is important to mention that at this point only a very few cases of boundary and encapsulated jets (Tables B1, B2) are found. This shows that the complexity of these jets is difficult to be captured by the techniques used so far.
To increase the sampling of the underrepresented classes (boundary/encapsulated), the algorithm uses only possibly candidates (Classes 4 and 6) to pass through the next stages.

Stage 4 – Normalizing each quantity:

In this stage, a normalization technique is applied to the measurements creating relative thresholds for each case (Table A1).

This procedure increases the number of cases that were initially not classified due to the strict thresholds imposed in the previous stages. On the other hand, it could also increase the number of false positives, making manual verification in a later stage necessary.

At this point, the code introduces a normalization to the quantities (Table A1, last row) and utilizes only one variation of pre/post jet time adjustment (variation B).

Jets that still did not get classified to either category are moved to the final stage.

Stage 5 – Removing a necessary criterion:

In stage five, the exact same procedure as in stage 4 is applied but with removing one necessary criterion. The criterion removed from necessary criteria is the one corresponding to high energy flux F_H (Table A1)

By doing so, more samples were allowed to be classified, enlarging significantly the database. Every jet that fails to be classified at this stage is automatically named "Unclassified" (class 0).

Manual Verification:

As described above, while Qpar and Qperp jets contained a few to no false positives, this is not the case for the boundary and encapsulated ones. Stages 4 and 5 allowed us to significantly increase the size of the database but at the cost of allowing several false-positive cases.

As a result, the first and the second author of the article did the following procedure to ensure that the database accurately reflects the intended classes:

At first, we removed the very few cases of Qpar and Qperp jets that appeared to be close to partial crossing of bow shock or magnetopause. From that procedure, we also found very few cases that contained rapid changes of the magnetosheath (from Qpar to Qperp or vice versa). It was decided that these jets should be moved to "Unclassified" as part of the manual verification procedure.

Finally, plenty of boundary cases were removed since they were considered false positives. These cases originated from stages 4 and 5 which due to the relative thresholds applied (Table A1) classified many jets but were prone to false positives. The same procedure was done for the encapsulated jets, which reduced slightly their final number (Table 3).

This final process provides the "final" cases that are highlighted in Table 3 of the main articles. These cases are also given in the accompanying supplementary material (Data set S3).

Figure S2. Flowchart of classification algorithm along with basic information of the algorithm.

Data Set S3. Class, starting time, and ending time of all the jets used in the analysis of the main article ("final" cases in Table. 3).