aida-space.eu

HORIZON 2020

This poster has received funding from the European Unions Horizon 2020 research and innovation programme under grant agreement No 776262 (AIDA, www.aida-space.eu).

Characterizing magnetic reconnection using Gaussian Mixture Models on electron particle velocity distributions

R. Dupuis[1], M. V. Goldman[2], D. L. Newman[2], J. Amaya[1], and G. Lapenta[1]
[1]Center for mathematical Plasma Astrophysics, KU Leuven, Belgium
[2]Center for Integrated Plasma Studies, Colorado University, USA
E-mail: romain.dupuis@kuleuven.be

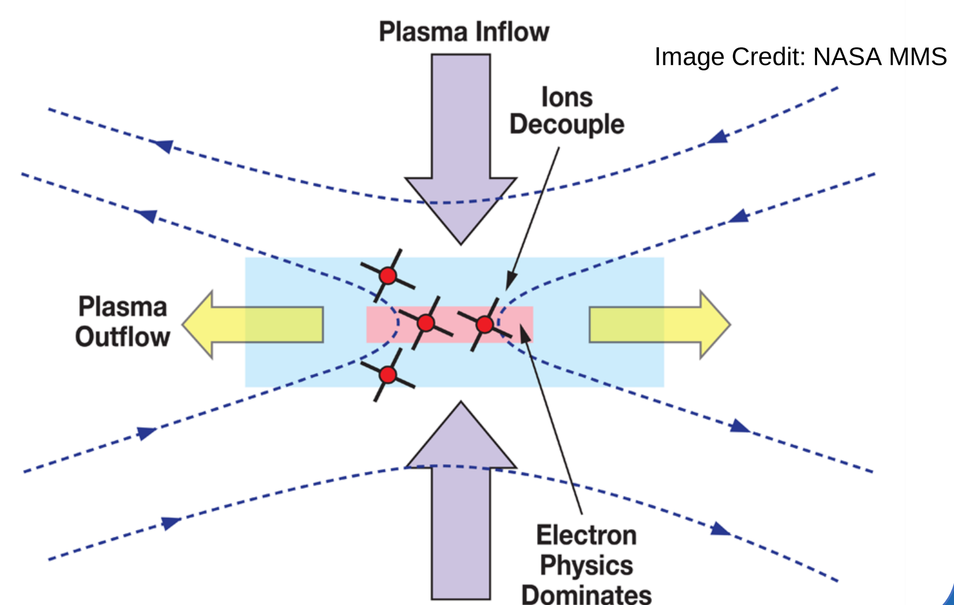
We present a method based on unsupervised machine learning **to automatically identify and characterize regions of interest using particle velocity distributions** as a signature pattern. **A density estimation technique** is applied to particle distributions provided by PIC simulations **to study magnetic reconnection regions**. The key components of the method involve: i) **a Gaussian Mixture Model** determining the presence of a given number of subpopulations within an overall population, and ii) a model selection technique with **Bayesian Information Criterion** to estimate the appropriate number of subpopulations. Thus, this method identifies automatically the presence of complex distributions, such as beams or other non-Maxwellian features, and can be used as a detection algorithm able to identify reconnection.

Magnetic reconnection

Reconnection plays **a crucial role in collisionless plasmas**. By breaking down the frozen-in magnetic fields, the magnetic energy is converted into kinetic, energy, and particle acceleration energy. Reconnection occurs at **various locations** such as laboratory plasma, solar atmosphere, or the magnetotail.

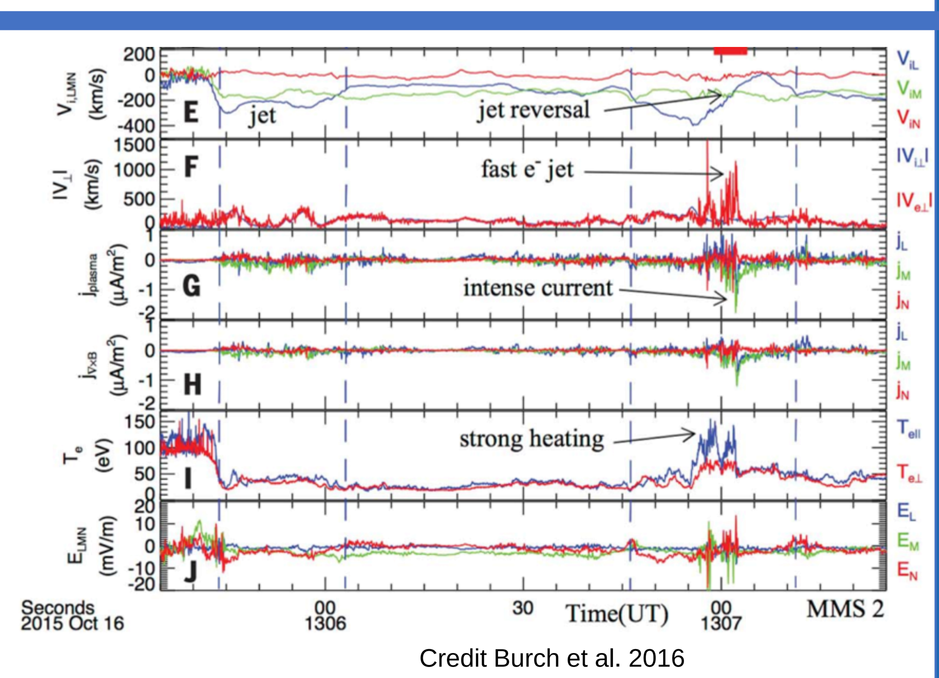
This process represents **one of the most important sources of particle acceleration in space.**

Magnetic reconnection is studied with **in-situ measurements** (MMS mission [1]) and **simulations** [2].



Challenging detection from data

- Simulations: thousands of particles per cell with a total data size nowadays reaching trillions of particles defined as **6D data**.
- Observations: **distributions** are available at specific time ranges defined manually by an operator or basic rules using to **fields data**.



Credit Burch et al. 2016

Metrics have been developed from statistical moments of the distributions. One can cite **the measure of gyrotropy** [3] and the agyrotropy.

Our approach: using directly the full distribution

Gaussian Mixture Model

The probability density function is approximated **by a sum of Gaussians** $\mathcal{N}$:

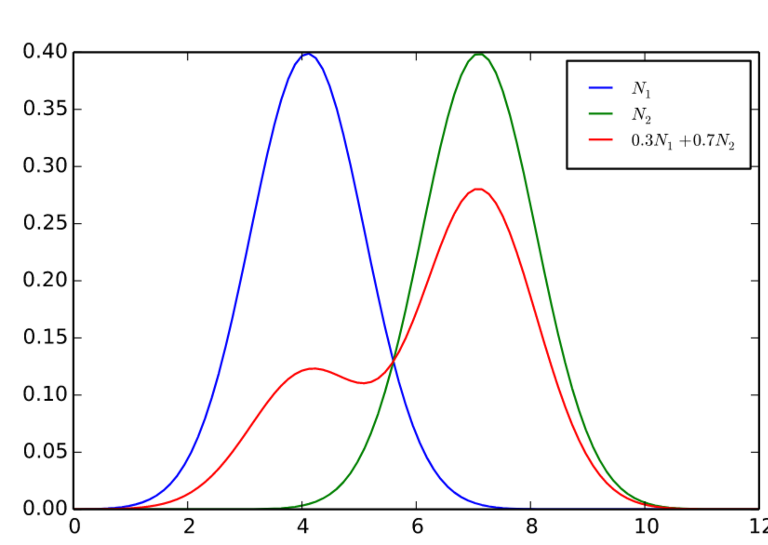
$$p(\boldsymbol{x}|\boldsymbol{\Phi}) = \sum_{k=1}^{K} w_k \,\mathcal{N}(\boldsymbol{x}|\boldsymbol{\theta_k})$$

- $\boldsymbol{\theta_k}$ k-th mean and covariance
- w_k weight of the k-th Gaussian
- K the number of Gaussians

$$\boldsymbol{\Phi} = [w_1, \cdots, w_K, \boldsymbol{\theta_1}, \cdots, \boldsymbol{\theta_K}]$$

The vector $\boldsymbol{\Phi}$ is computed as follow:

1) Discretize density by a list of particles
2) The likelihood function **L** is maximized
3) Expectation-Maximization algorithm
4) No closed form, only local maximums



Model selection problem: How to determine K in advance?

Bayesian Information Criterion (BIC) is maximized among a set of models

$$BIC = \underbrace{\ln(n)k}_{\text{complexity}} - \underbrace{2\ln(L)}_{\text{goodness of fit}}$$

- k number of free parameters
- n number of particles

The value K **does not necessary correspond to the number of subpopulations**:

- It may be beams or electron subpopulation
- It can also reflects complex distributions (flat-top, Kappa, crescent shape, etc.)

Diagnostic tool

Thermal energy as the second moment may be inappropriate if applied on several subpopulations :

$$E_{thermal} = \frac{1}{N_p}\sum_{i=1}^{3}\left[\sum_{p}\left(\boldsymbol{V_p} - \langle \boldsymbol{V_p}\rangle\right)^2\right]_i$$

The second statistic moment on the mixture has additional terms:

$$(\sigma^2)^{(K)} = \sum_{i=1}^{3}\left[\underbrace{\sum_{k=1}^{K} w_k^2 (\boldsymbol{\sigma_k})^2}_{E_{thermal}^{(K)}} + \underbrace{\sum_{k=1}^{K} w_k (\boldsymbol{\mu_k})^2 - \left(\sum_{k=1}^{K} w_k (\boldsymbol{\mu_k})\right)^2}_{E_{dev}^{(K)}}\right]_i$$

$E_{thermal}^{(K)}$ relates to the thermal energy of the mixture. The ratio with the energy shows the reduction in thermal speed (heating vs acceleration)

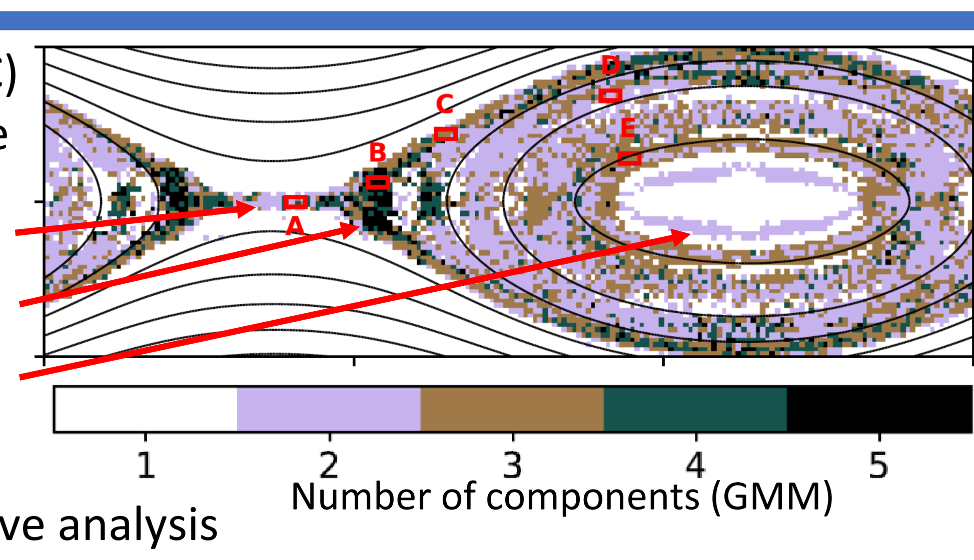
$$E_{drop} = \frac{E_{thermal}^{(K)}}{E_{thermal}}$$

$E_{dev}^{(K)}$ is the weighted variance of the mixture means measuring. The ratio with the energy highlights the complexity of the distribution.

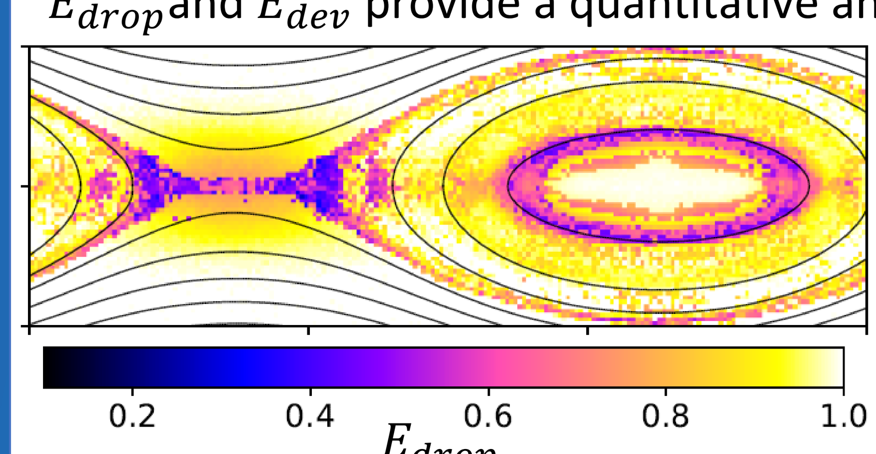
$$E_{dev} = \frac{E_{dev}^{(K)}}{E_{thermal}}$$

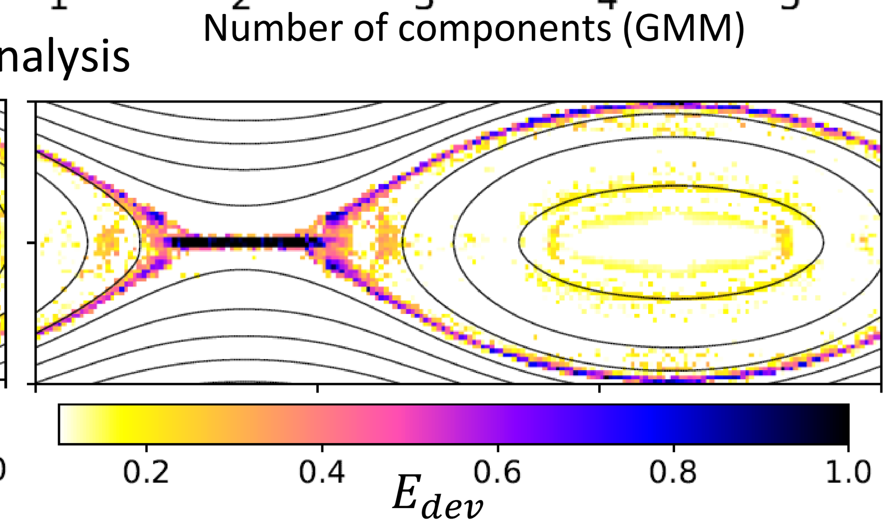
Results

2.5D **Particle-In-Cell** (PIC) simulation, double Harris sheet case



E_{drop} and E_{dev} provide a quantitative analysis

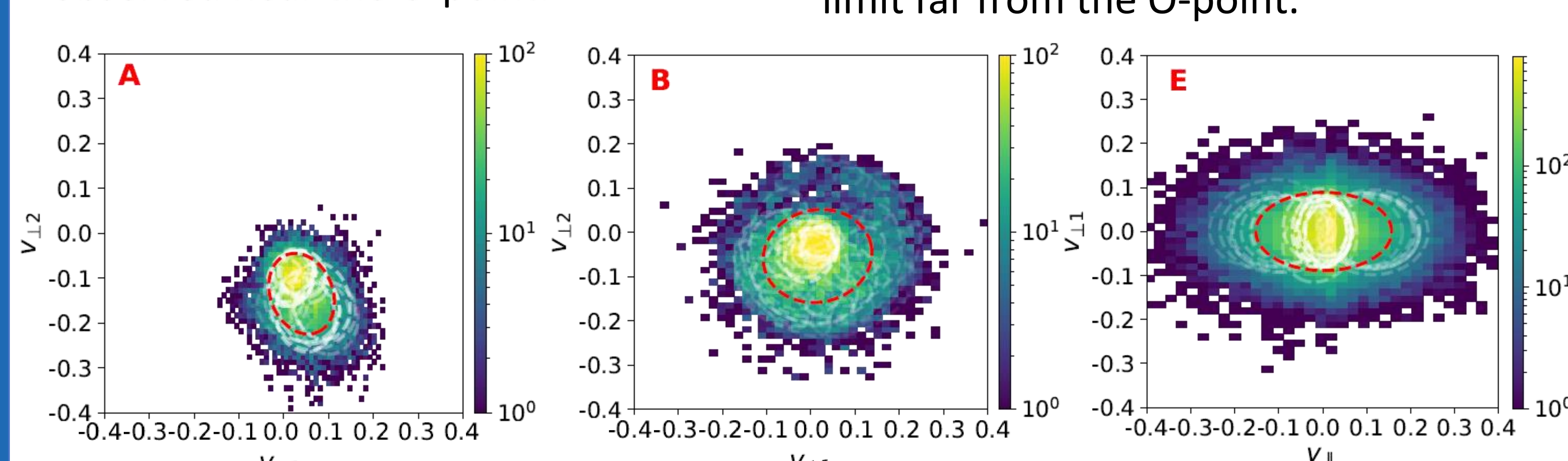




Several regions are identified. The inflow is characterized by a large region of high value. The EDR and the outflow show low values, meaning the mixture extend is reduced compared to the second moment. A ring is also observed near the O-point.

A very narrow E_{dev} peak is observed around the EDR, suggesting complex distributions are present.

Typical topological boundaries are highlighted by E_{dev}, in particular at the outflow, near the separatrix and at the limit far from the O-point.



The above distributions refer to the labels in the first figure of the section. They illustrate various behaviors, such as a Gaussian core associated with a surrounding halo (A and B) or an anisotropic heavy-tail distribution (E).

https://arxiv.org/abs/1910.10012

[1] Burch et al., "Magnetospheric multiscale overview and science objectives", Space Science Reviews, 2016
[2] Goldman et al. 2016, "What can we learn about magnetotail reconnection from 2D PIC Harris-sheet simulations?" Space Science Reviews, 199, 651
[3] Swisdak, "Quantifying gyrotropy in magnetic reconnection", Geophysical Research Letters