CURRENT STRUCTURES AND RECONNECTION EVENTS ANALYSIS IN HYBRID-KINETIC TURBULENCE SIMULATIONS USING UNSUPERVISED MACHINE LEARNING

MANUELA SISTI1,2, GIORGIO PEDRAZZI3, FRANCESCO FINELLI1, MATTEO FAGANELLO1, FRANCESCO CALIFANO1, FRANCESCA DELLI PONTI1, ANDONG HU4, and JANNIS TEUNISSEN4

Corresponding authors: manuela.sisti@univ-amu.fr
1 Aix-Marseille University, CNRS, P2M UMR 7345 - Marseille, France, 2 Physics Department, University of Pisa - Pisa, Italy, 3 HPC Business Unit, Cineca, Bologna, Italy, EU, 4 Centrum Wiskunde & Informatica (CWI), Amsterdam, The Netherlands, EU

INTRODUCTION
In the context of the study of magnetized turbulent space plasmas, a key point concerns the formation of coherent structures and their disruption through reconnection. Up to now no well verified technique to automatically detect such structures has been developed, and reconnection can be only identified by human analysis looking at possible sites one by one. Lots of data are produced by simulations (MHDs, Hall-MHDs, hybrid ones, PICs) and by satellite measurements. Thus there is a need to find a way to rapidly and efficiently identify reconnection events among these big data. Our goal is to set-up an algorithm using unsupervised Machine Learning aimed at automatically detecting the presence of magnetic structures where reconnection is occurring (2D). The final objective will be to adapt these algorithms to satellite data.

This research is developed in the framework of the european project AIDA (Artificial Intelligence Data Analysis).

MODEL AND SIMULATION SETUP

Hybrid Vlasov-Maxwell code

Normalized equations used:

\[
\begin{align*}
\frac{\partial f}{\partial t} + \mathbf{v} \cdot \nabla f + \mathbf{E} \cdot \nabla \times \mathbf{v} + \mathbf{B} \cdot \nabla \times \mathbf{B} = 0
\end{align*}
\]

(1)

The Ohm’s law for the electric field:

\[
\mathbf{E} = -\nabla \times \mathbf{B} - \frac{1}{\Omega} \left( J \times \mathbf{B} \right)
\]

(2)

Faraday’s equation:

\[
\frac{\partial \mathbf{B}}{\partial t} = -\nabla \times \mathbf{E} + J = \nabla \times \mathbf{B}
\]

(3)

\[m_i \approx m_e\]

Box: 2DxLx = Lx = 2π x 50d, 302\times302 grid points, resolution ~ 0.1d. Initial set-up:

\[B_0 = 1, m_i/m_e = 100, \beta_i = 8\pi nT_0/B_0^2 = 1\]

Initial distribution function: Maxwellian (T_0 = T_e, Te)

Turbulence is initialized with random, isotropic magnetic-field perturbations, with k in [0.02, 0.12] and dBmax ~ 0.2B.

RESULTS 1: OVERVIEW

Variables used as reconnection proxies (normalized):

- current density |J|
- in-plane electron velocity |V_e|
- electron vorticity \Omega_e = \nabla \times \mathbf{V}_e
- in-plane magnetic field \mathbf{B}
- electron decoupling: E’ = E + V_e \times B, z-component
- J - E’

Algorithm’s steps:

1. Parallel tuning k for the k-means model:
   - to avoid overfitting or underfitting
   - k = 11 for t 247 (tuning with Davis-Bouldin index [1])

2. K-means’ Lloyd’s algorithm:
   - find clusters in the variables space
   - we choose the cluster in variables space where the mean value of |J| is the highest

3. DBscan algorithm
   - the cluster chosen using K-means is made up by different structures in physical space, we use DBscan to distinguish them
   - e = 50 grid points ~ 5d (search radius for DBscan)
   - Minimum points number to form a cluster (DBscan): 100

RESULTS 1: PERFORMANCE

Comparison between (1) method = linear + DBscan + AR th, (2) method = threshold over J + threshold over AR

Precision and precision non-mr (mean over three central times)

\[\text{Precision}\]

\[\text{Precision non-mr}\]

RESULTS 2: PERFORMANCE

\[1, 35, 22, 0.49, 0.87\]

\[247, 494, 80\%\]

\[30, 10, 0.80, 0.80\]

\[1, 0.87, 0.97, 0.95\]

\[49, 41, 0.71, 0.66\]

\[30, 41, 0.73, 0.73\]

\[1\]

CONCLUSION

In particular:

- We are working to obtain a quite accurate method to automatically find reconnecting current sheets in turbulence simulations

In general:

- We are creating an utility linked to the reconnection rate \(R\), in particular \(R \sim 10\) for fast reconnection model [2], which gives

\[AR = \frac{\text{length}}{\text{time}} \sim 10\]

Quantities to evaluate performance:

- \(\text{Precision} = \frac{\text{# reconnection sites among sites selected}}{\text{# sites selected}}\)
- \(\text{Precision non-mr} = \frac{\text{# non-reconnections among excluded sites}}{\text{# sites excluded}}\)

In Figure we show the final clusters (in different colors) which we obtain by applying the first three steps. We compute the aspect ratio of each of these structures.

REFERENCES


K-means and DBscan are techniques of unsupervised ML with the aim to learn a grouping structure in a dataset (clustering techniques) [1].

*Average over three times (230, 247, 282)

*Average over four times (230, 247, 282, 494)