Feature Histogram

- **Histogram Construction**
- **From 1D to 2D histograms**
- **Preserve the order of the attribute values by dividing streamlines into segments**
- **Distance between histograms**

\[
\text{Distance} = \sum_{i} \sum_{j} \text{EMD}(X_i, Y_j)
\]

- **DTW**: minimum mapping cost

\[
\text{DTW}(X, Y) = \min_{v} \left( \sum_{i=1}^{n} d(v_i, y_i) \right)
\]

**Application: Similar Streamline Query**

- The user can specify some sample streamlines, and uses the distribution-based distance to find all similar streamlines.

**Parallel Stream Surface**

- **Overview of the pipeline**
- **Computation**

- **Parallelization Strategies**
  - Cut seeding curves into segments, each process computes part of the stream surface
  - I/O strategies:
    - Data loading on demand from other processes over network instead of disk
    - Use cache to store the recently received data blocks
  - Dynamic load balancing strategies:
    - Work stealing
    - Runtime seeding curve segment subdivision

**Case Study: Solar Plume**

- Solar Plume: Triage of complex segments extracted based on box counting ratio. Top: Extracted feature segments, Bottom: clusters on the feature space correspond to various clustered streamline segments for this dataset.