# INDICATE:

#### **INdex to Define Inherent Clustering And TEndencies**











#### **Talk Structure**

- Motivation
- What is INDICATE?
- Why use INDICATE?
- INDICATE in 2D
- INDICATE in 3D
- Future development & applications
- Summary









#### Traditional clustering algorithm

Assume:There are clusters in the datasetGoal:Identify cluster centroids and members

**Problem:** Do not give enough quantitative information for individual objects



Traditional clustering algorithm

Assume:There are clusters in the datasetGoal:Identify cluster centroids and members

**Problem:** Do not give enough quantitative information for individual objects

Need to quantitatively trace how "clustered" individual objects are in distribution & how this changes as the system evolves



Traditional clustering algorithm

Assume:There are clusters in the datasetGoal:Identify cluster centroids and members

**Problem:** Do not give enough quantitative information for individual objects

Need to quantitatively trace how "clustered" individual objects are in distribution & how this changes as the system evolves



#### What is INDICATE?

Statistical tool for 2D, 3D and 6D datasets

 Quantifies the degree of clustering of <u>each object</u> in a discrete distribution

Assigns a "clustering" Index to each object

 Compares number of nearest neighbours for each object in observed distribution to that expected in a evenly spaced uniform distribution





(Buckner et al., in prep)





(Buckner et al., in prep)



# Why use INDICATE?

Can be applied in any desired parameter space

Robust. No preference for data set

- size
- shape
- number of dimensions (2D, 3D, 6D)

 Directly compare Index value for objects in different distributions

Full automated





























































(Buckner et al., in prep)



#### **Future Development and Applications**

- Full testing of 3D version (shapes, simulations..etc.)
- Development & testing of 6D version
- Application to find substructures in observed and simulated data
- Application to quantify/trace mass segregation



#### Summary

#### • INDICATE is:

- A novel statistical tool
- Quantifies the degree of clustering of each object in a discrete distribution
- Advantages:
  - Full automated
  - Robust
  - Not dependent on distribution size, shape, of number of dimensions
    - Quantitatively trace the spatial evolution of each object in a distribution
- Applications:
  - Substructure
  - Mass segregation



#### **Questions?**

