

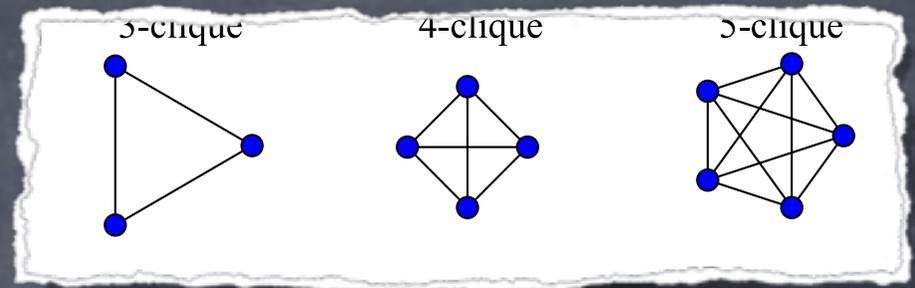
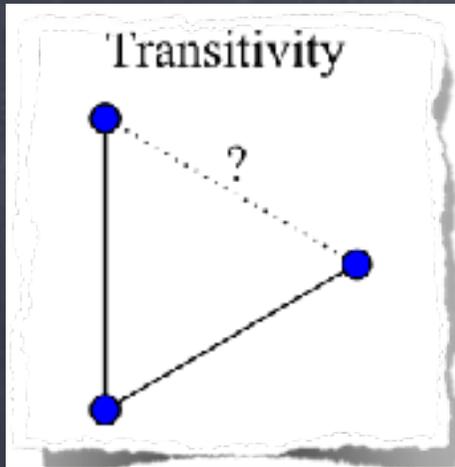
A Brief Introduction of Network Cosmology: What I have learned from galaxy Facebook



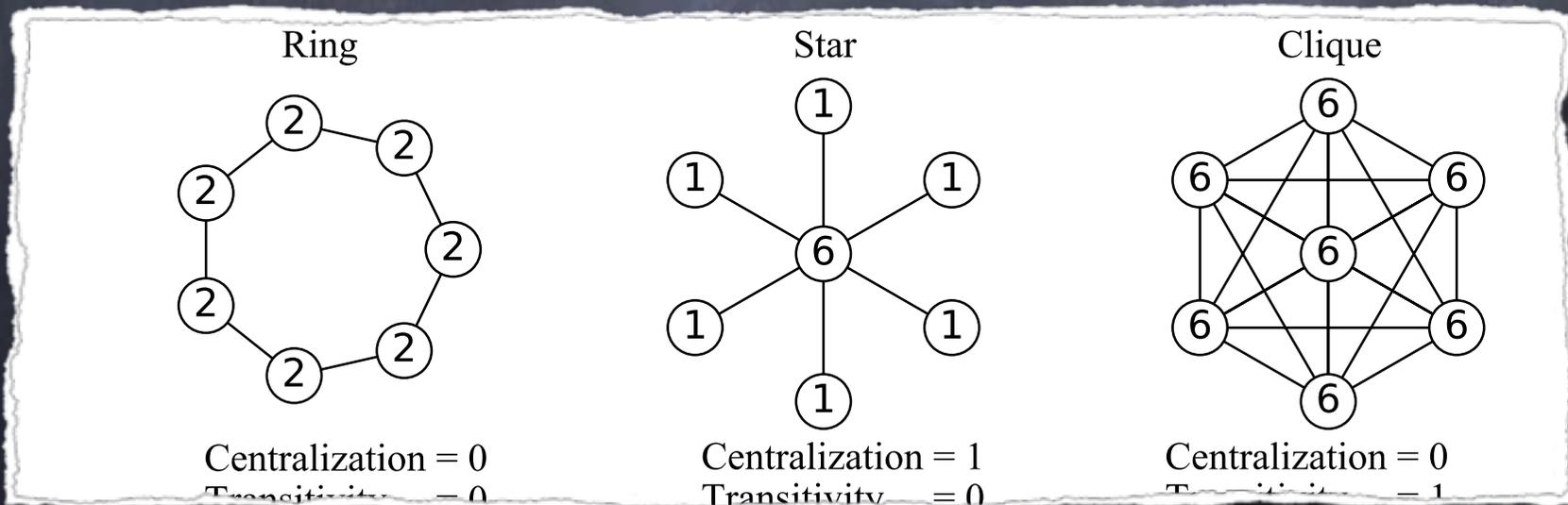
Sungryong Hong
1/17/2018

SSG

facebook



Graph Topology of Galaxy Point Distribution



Outlines

- Introduction:
 - Topology, Graph Theory, and Network Algorithms
 - Topology 101 : Metric Space vs. Topological Space
- Previous Applications in Astronomy
 - Network Centrality and Topological Environment
 - Illustris Galaxies vs. Levy Flights
 - Statistics of 2pt correlation and network topology for Lyman alpha emitters at $z=2.66$
- Current and Future Applications

Practical Applications in Astronomy

Galaxy Point Distributions

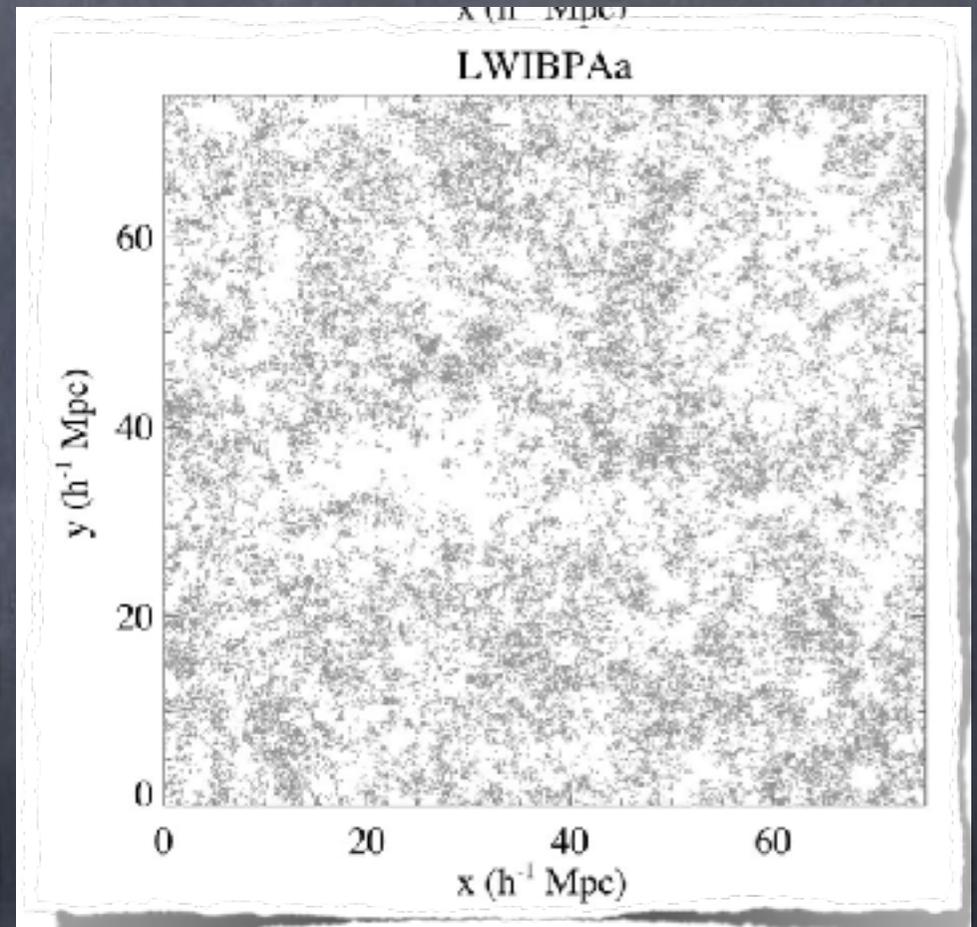
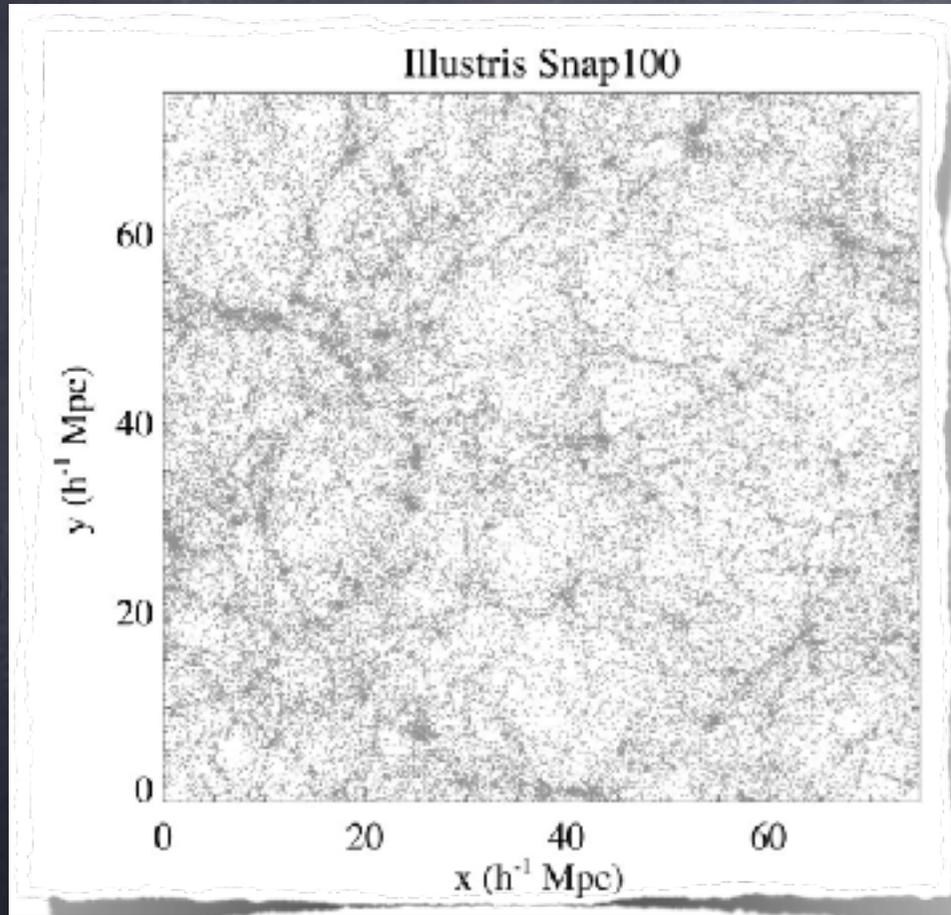
-> Generating Networks

(e.g., friends-of-friends networks)

-> Measuring Graph Topology

Example (2) : Illustris vs. Levy walks

How long does it take for you to recognize any difference between the two spatial distributions below?

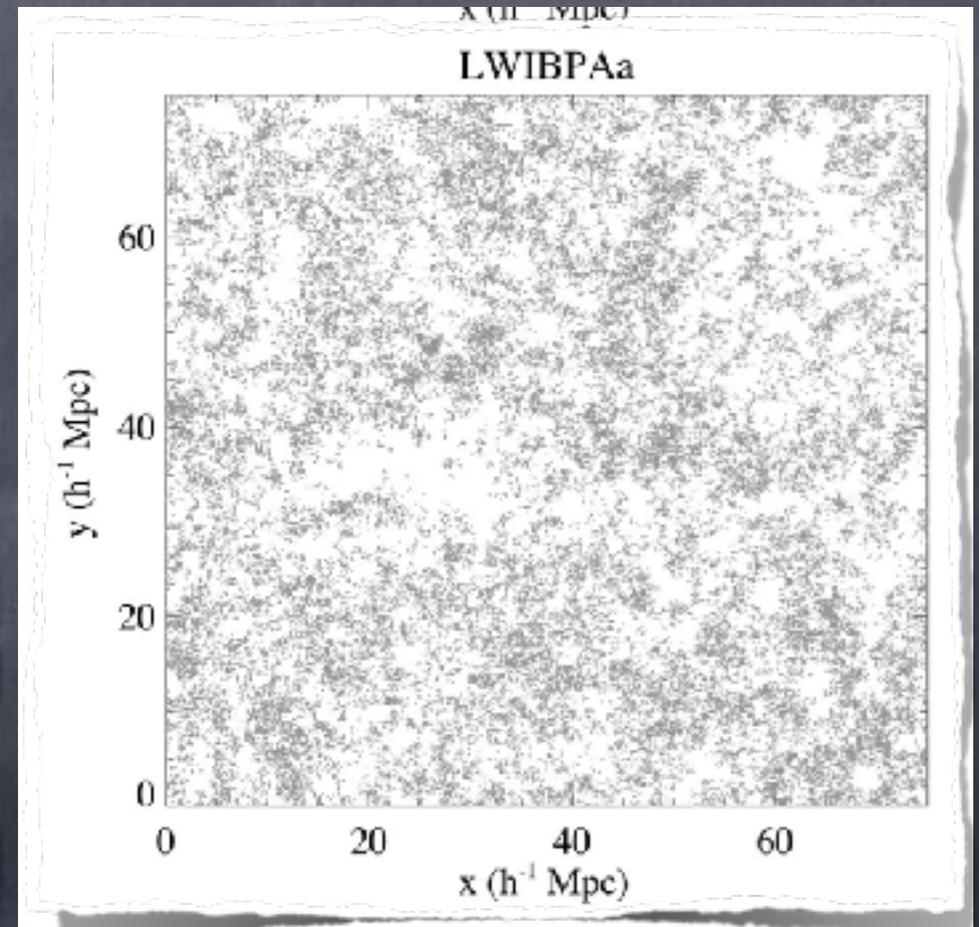
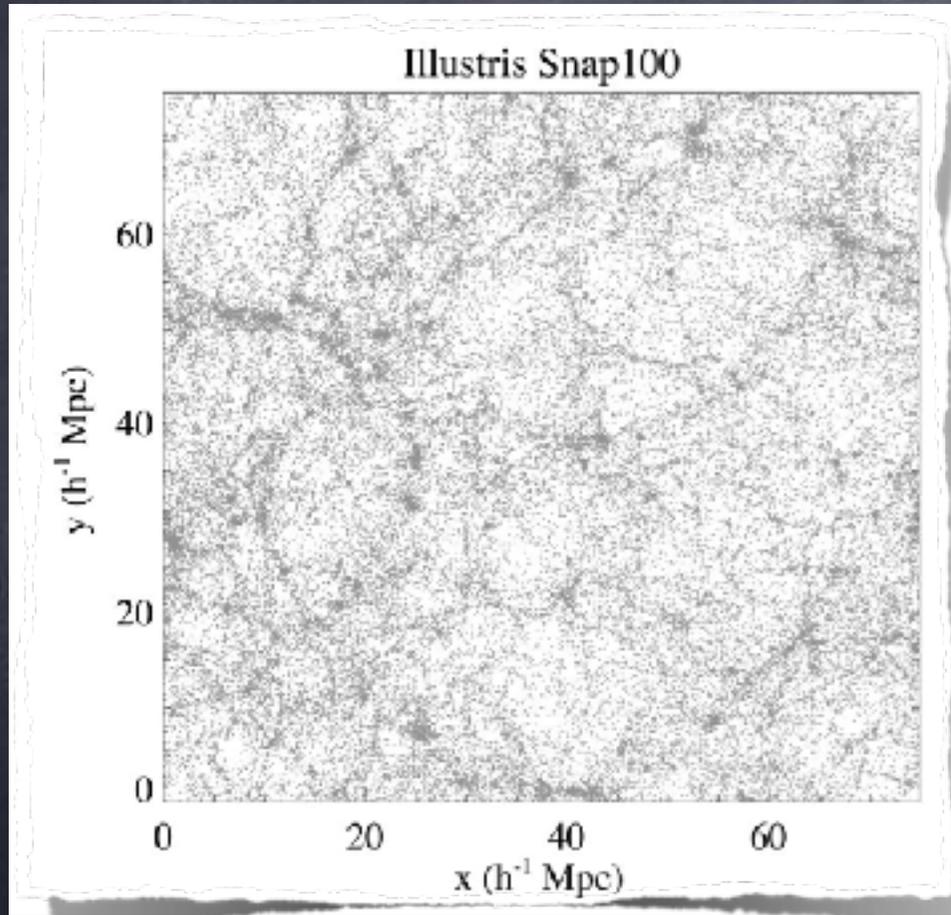


Illustris Galaxies : simulated galaxies with stellar masses greater than 10^8 Msol
Levy Fractal : random walks reproducing a single power-law 2pt function

(Hong+2016)

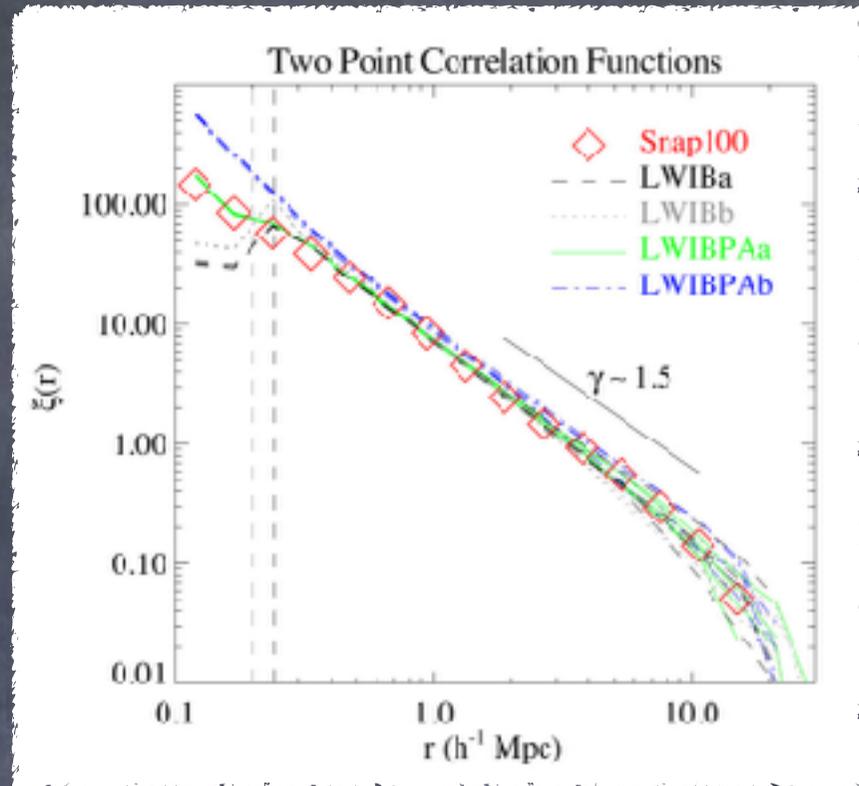
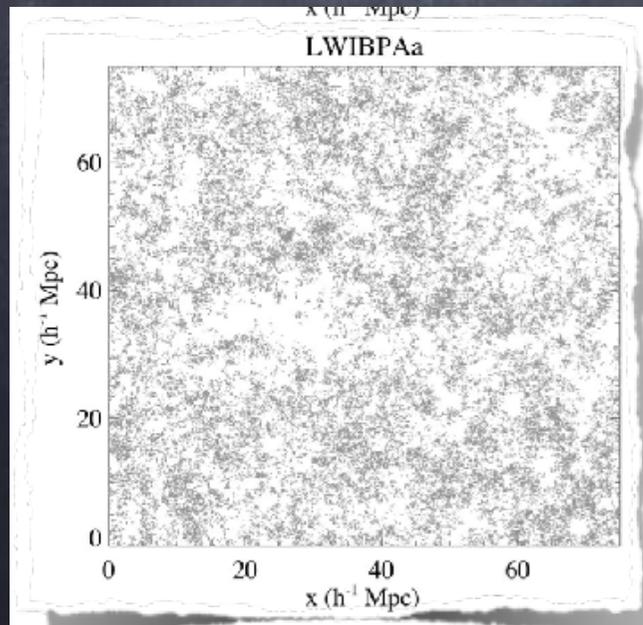
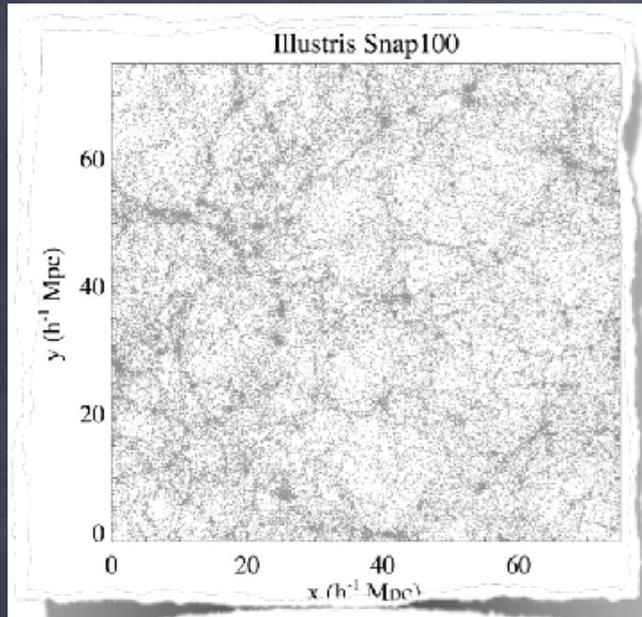
Example (2) : Illustris vs. Levy walks

How long does it take for you to recognize any difference between the two spatial distributions below?



Intuitively, we can recognize topological differences in a second, but n -point functions can't. (Hong+2016)

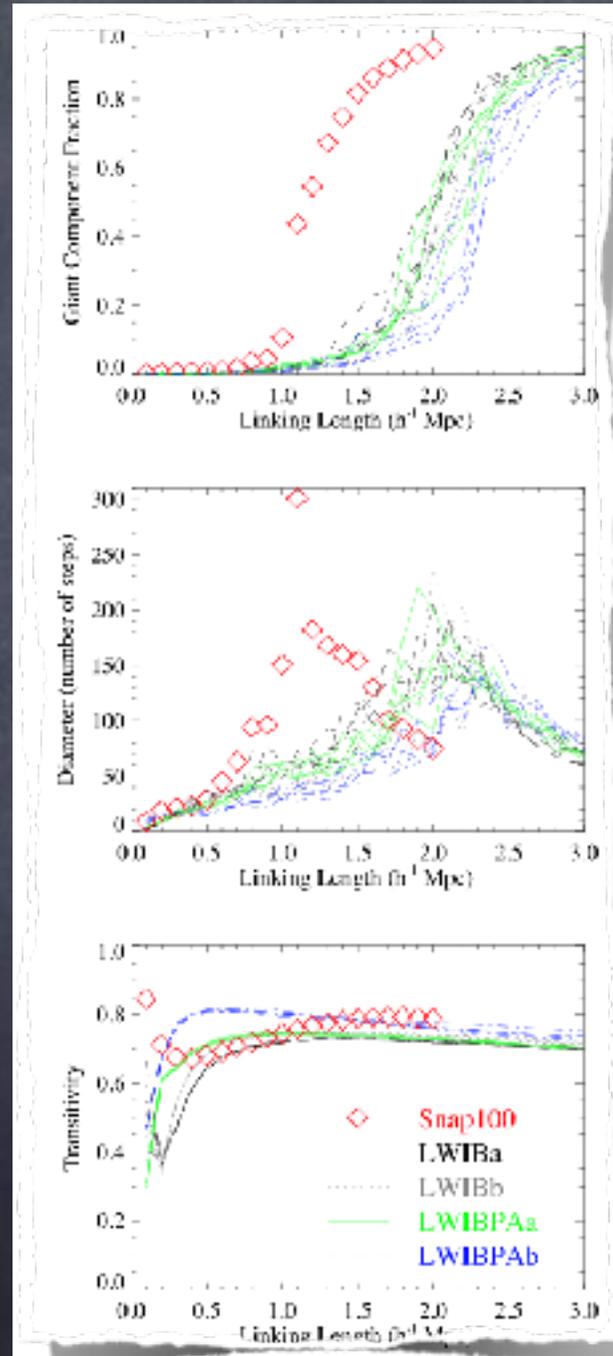
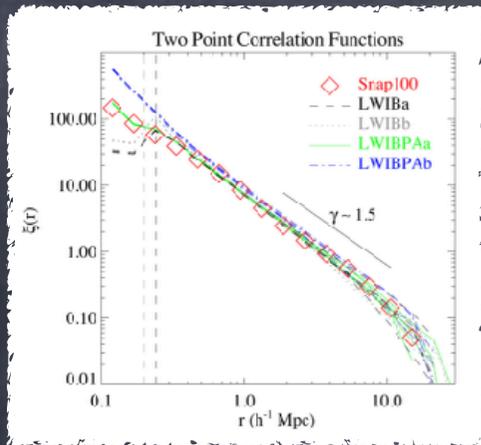
Example (2) : Illustris vs. Levy walks



If we follow the conventional way,
we have this problem :
the two models have the same
abundance and two-point function,
but they look different!

Degeneracy of 2pt function !

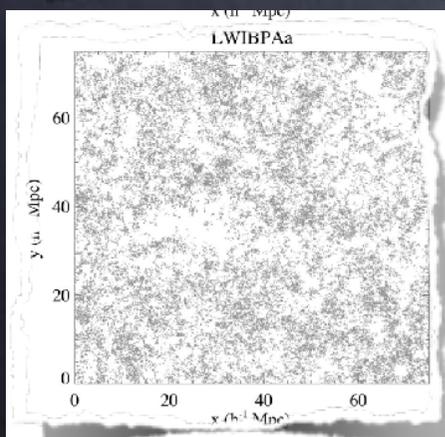
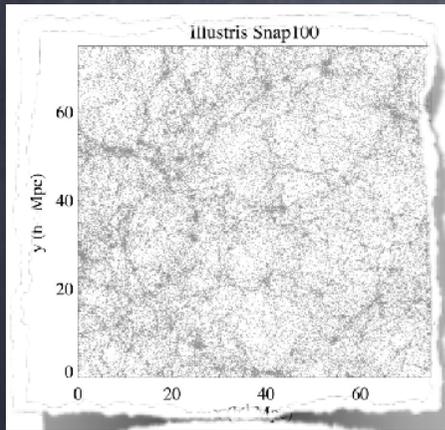
Example (2) : Illustris vs. Levy walks



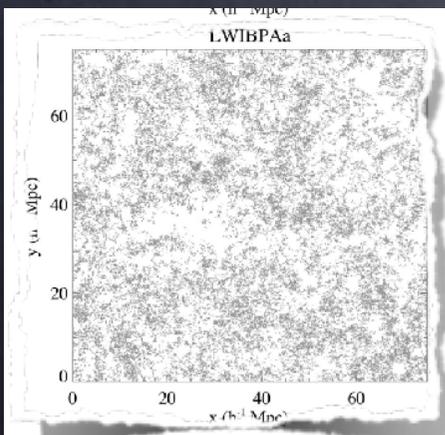
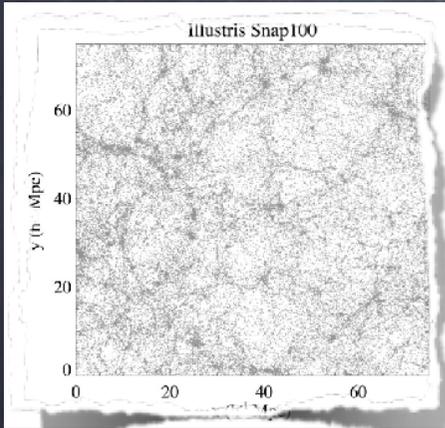
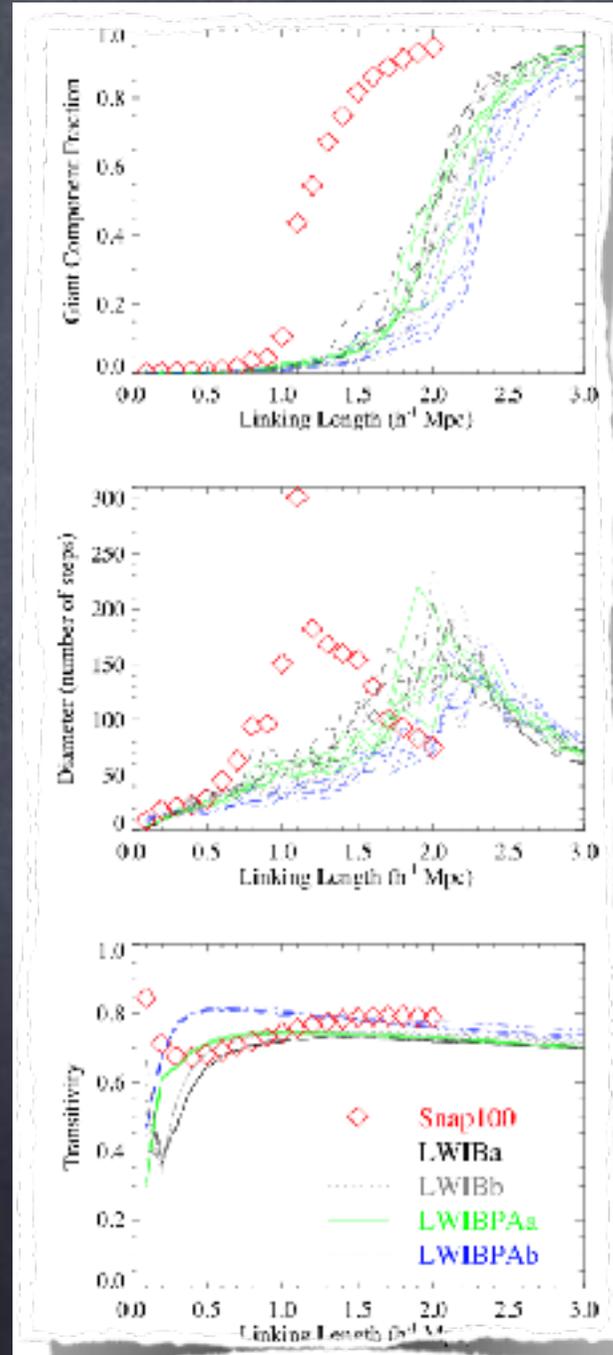
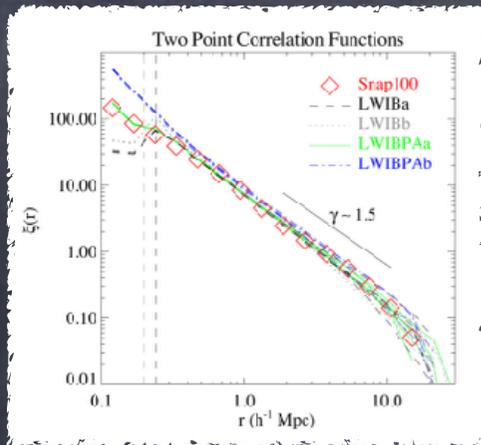
Giant component :
the largest connected
subgraph

Diameter :
the longest geodesic path
length

Transitivity :
Roughly, triangle density



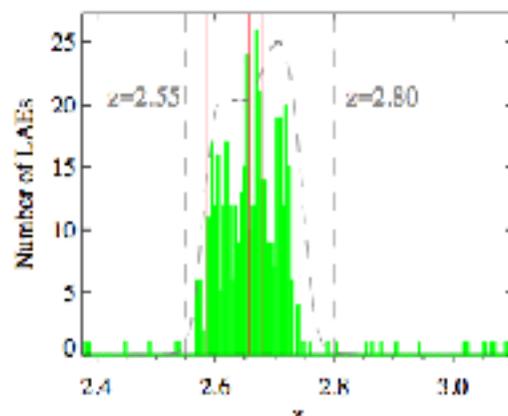
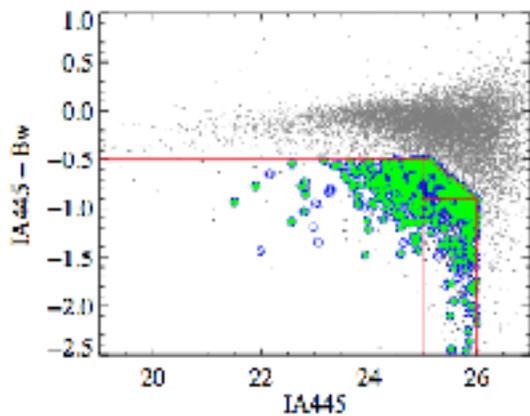
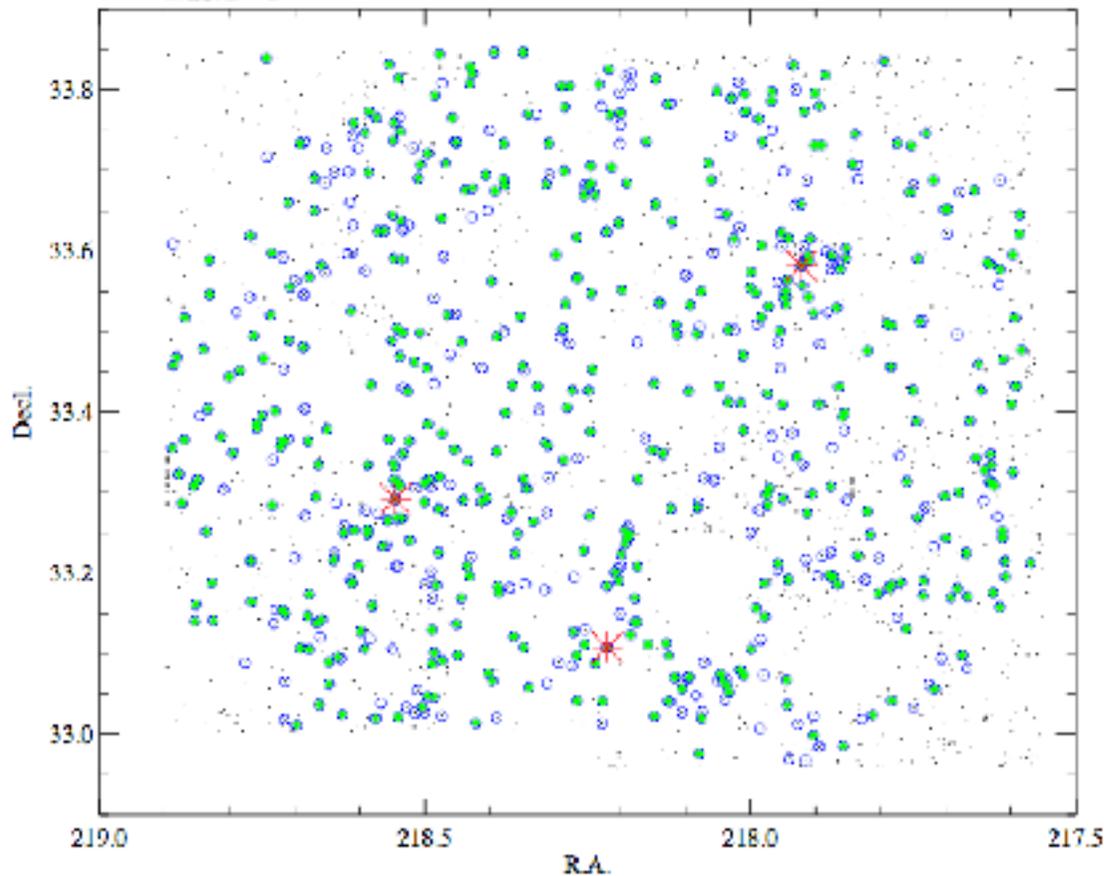
Example (2) : Illustris vs. Levy walks



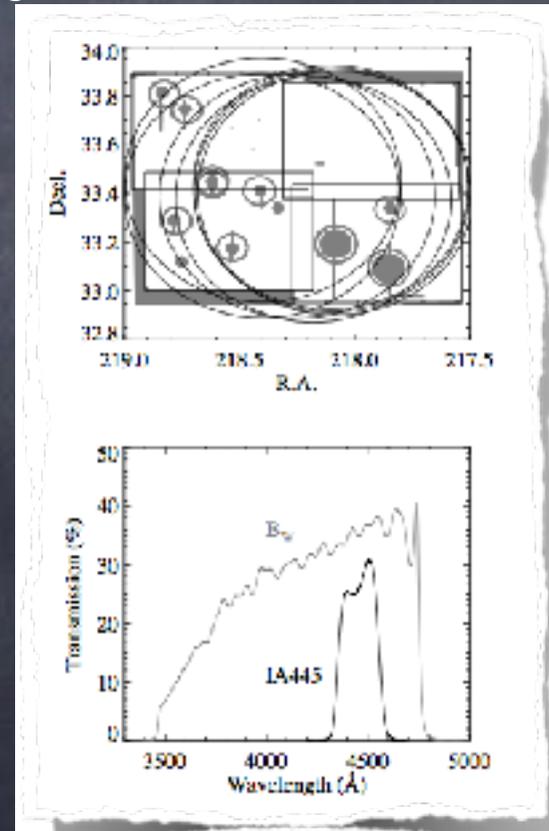
1. Quantitatively we show that the distribution of Illustris galaxies is not a Levy fractal.

2. Network (or graph) topology has great potential as topological diagnostics of galaxy point distributions, complementary to n-point statistics!

Example (3) : Bootes LAEs at $z \approx 2.66$



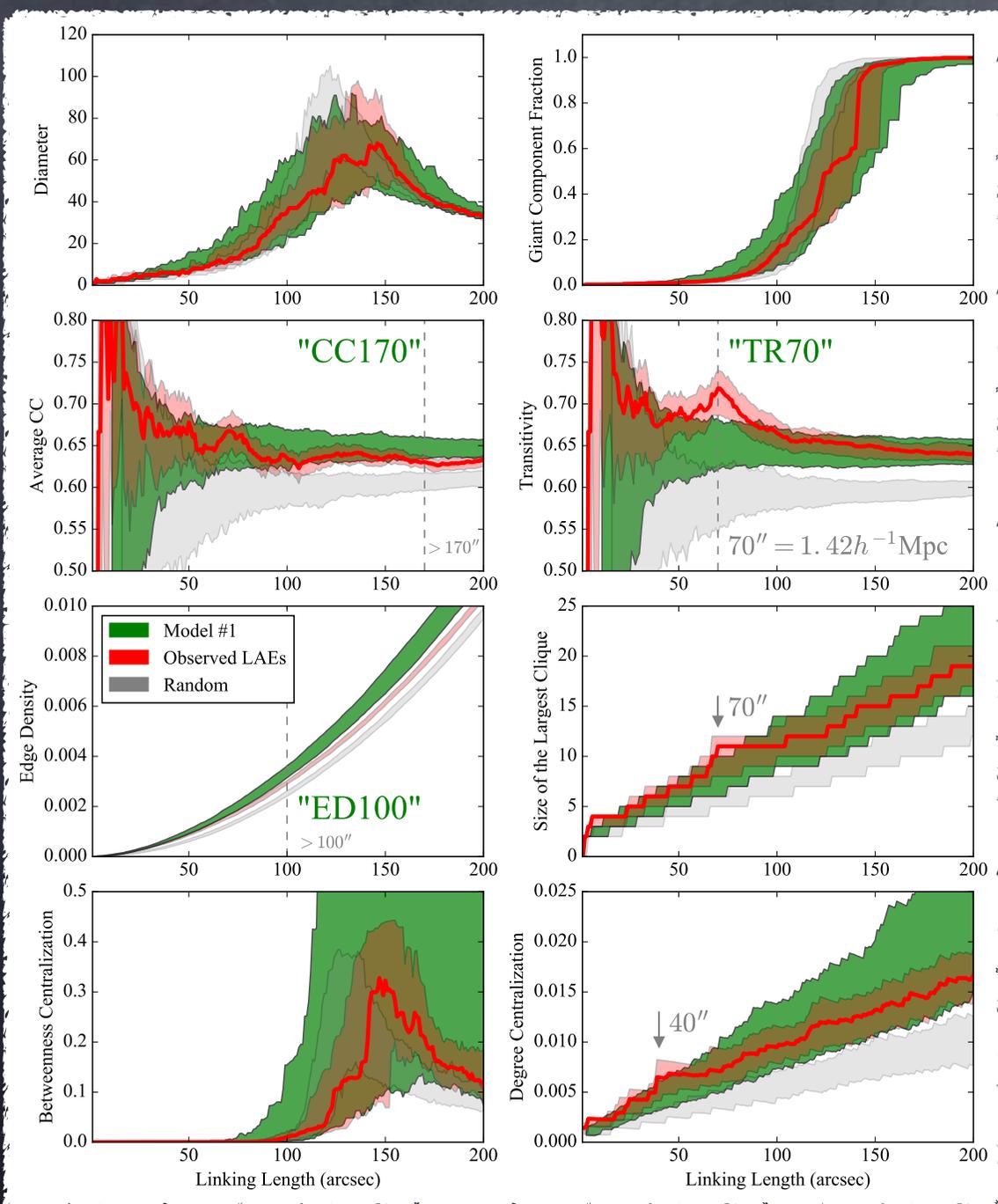
- Color Selection :
1956 LAE candidates (black dots)
- Spectroscopic Followup :
 - 635 observed LAE candidates (open blue circles)
 - 415 confirmed LAEs (green solid diamond)



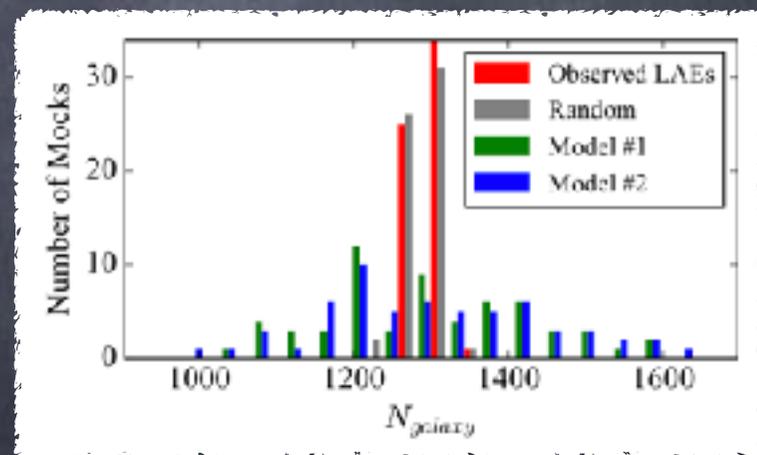
Statistics of Network Topology for LAEs at $z \approx 2.66$

- Diameter
- Giant Component Fraction
- Average Clustering Coefficient (Average CC)
- Transitivity
- Edge Density
- Size of the Largest Clique
- Betweenness Centralization
- Degree Centralization

Statistics of Network Topology for LAEs at $z \approx 2.66$

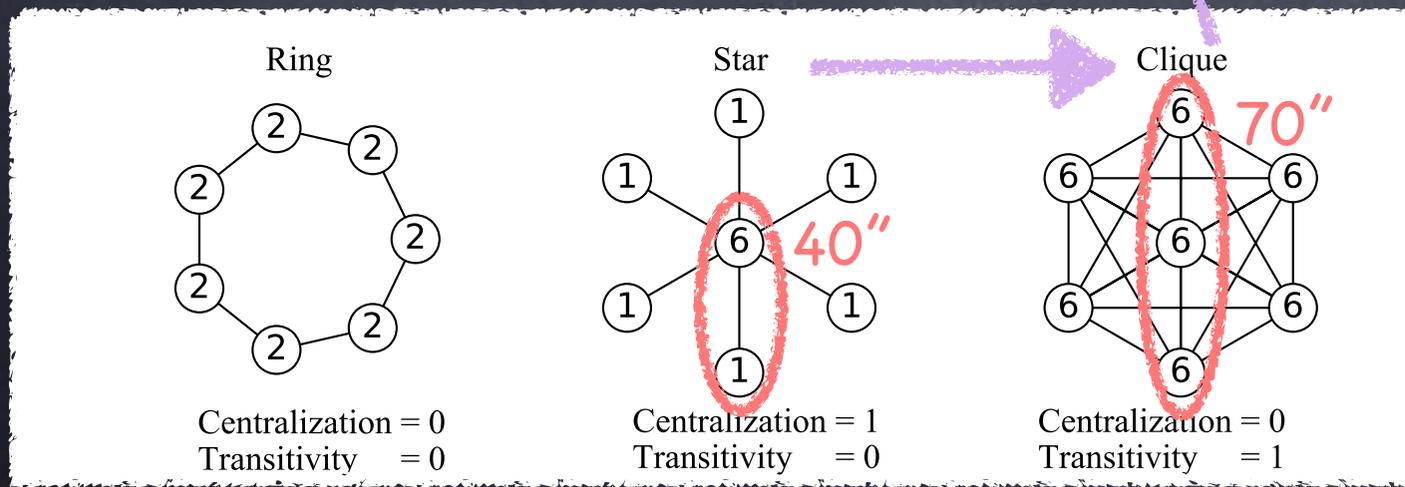
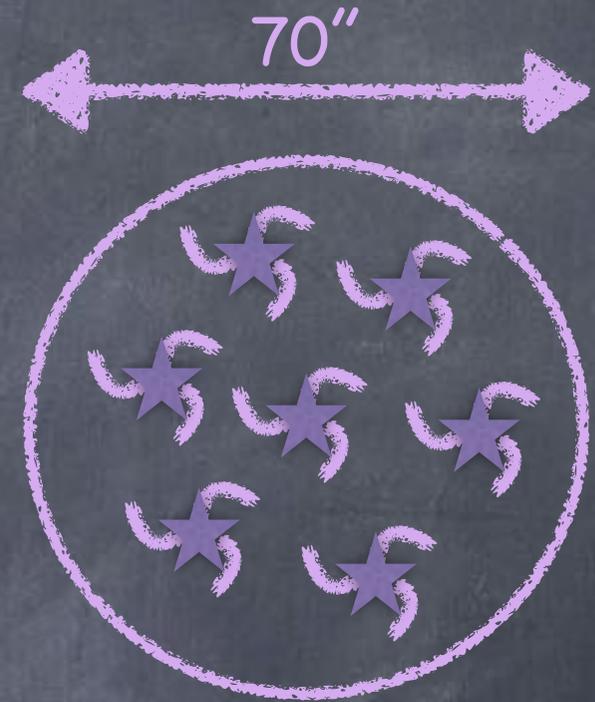
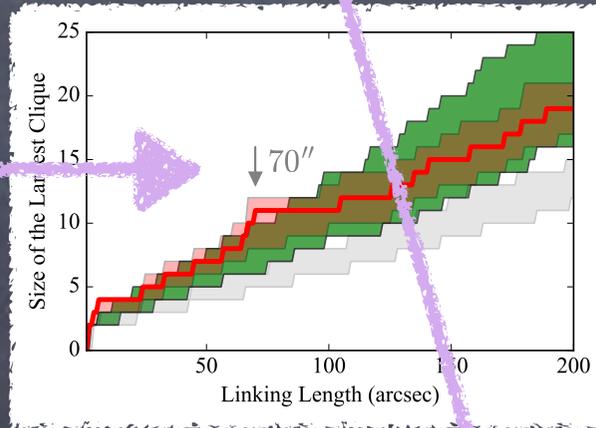
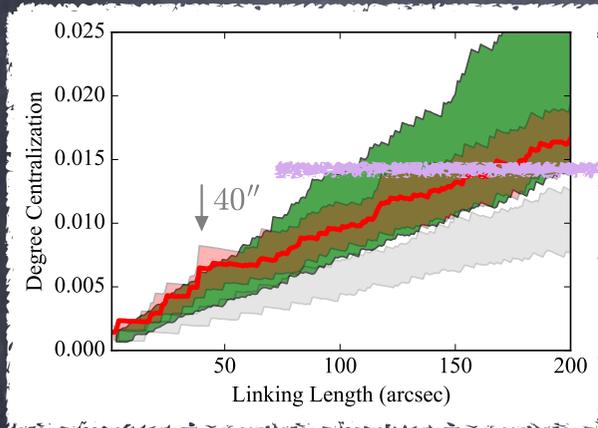
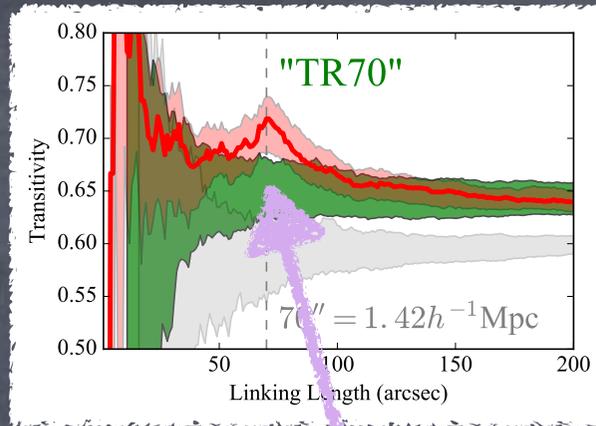
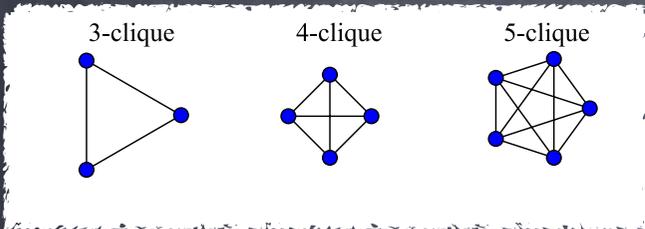
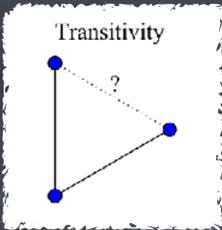


- Diameter
- Giant Component Fraction
- Average Clustering Coefficient (Average CC)
- Transitivity
- Edge Density
- Size of the Largest Clique
- Betweenness Centralization
- Degree Centralization



60 mocks for each model;
 Observed LAEs, Random Model,
 Model#1, and Model#2

Diameter of LAE Clique (Group) = 70"



2.1 Mpc in Comov.
600 kpc in Phy. at z=2.7

Current and Future Works

- Let's "Spark" the Universe

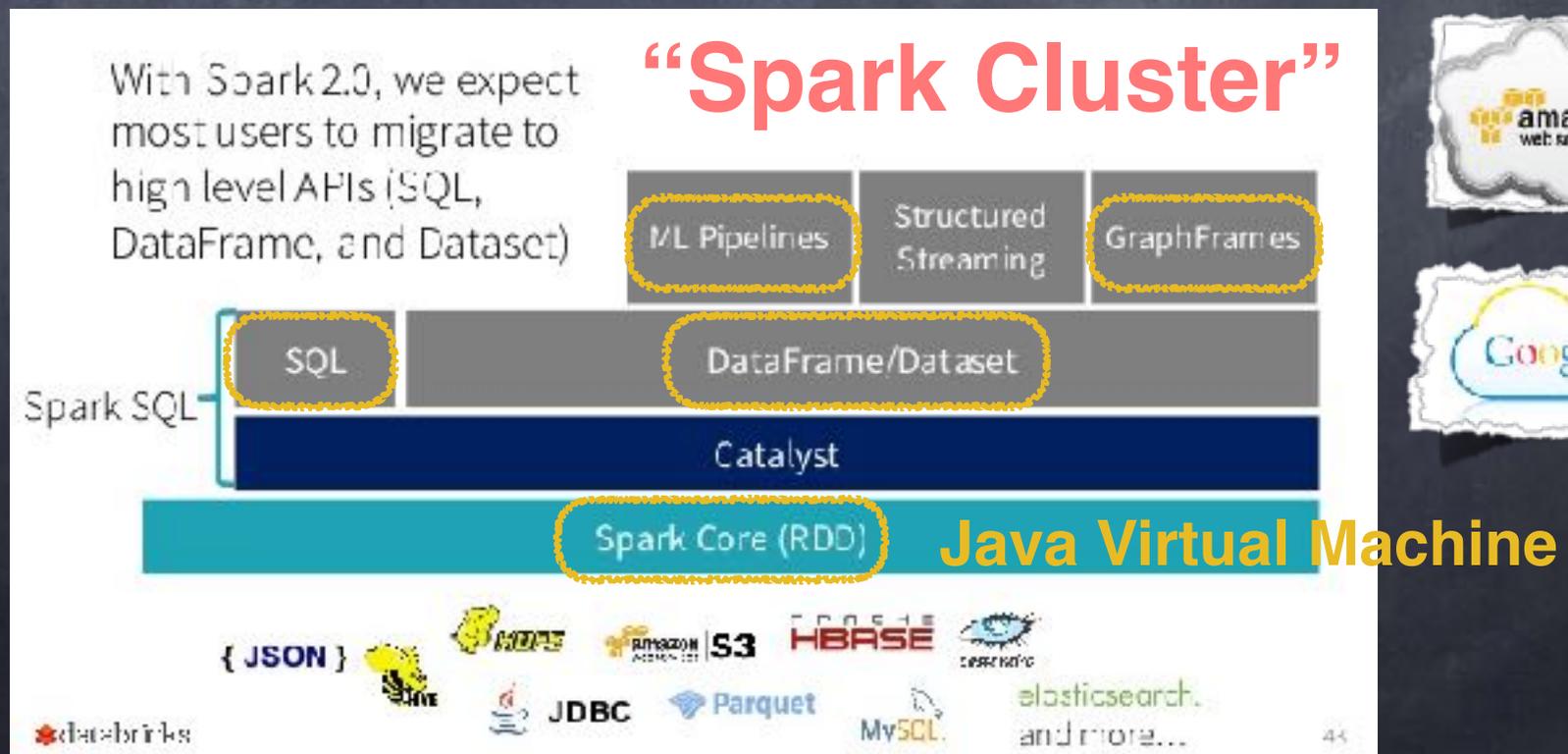
We need a tool for "Large-scale" Graph Calculations.

HETDEX : Millions of LAEs, DESI : 10 Millions of Galaxies

Fortunately, we have Apache Giraph and Spark GraphX

which can handle **billions vertices** and **trillions edges**

for analyzing **Facebook** and **Twitter** networks.



Specific Current Works

- Measuring network statistics of **SDSS Skeleton galaxies**
 - The Great Wall in the SDSS seems special?!
- Measuring network statistics of **5 Multiverses**

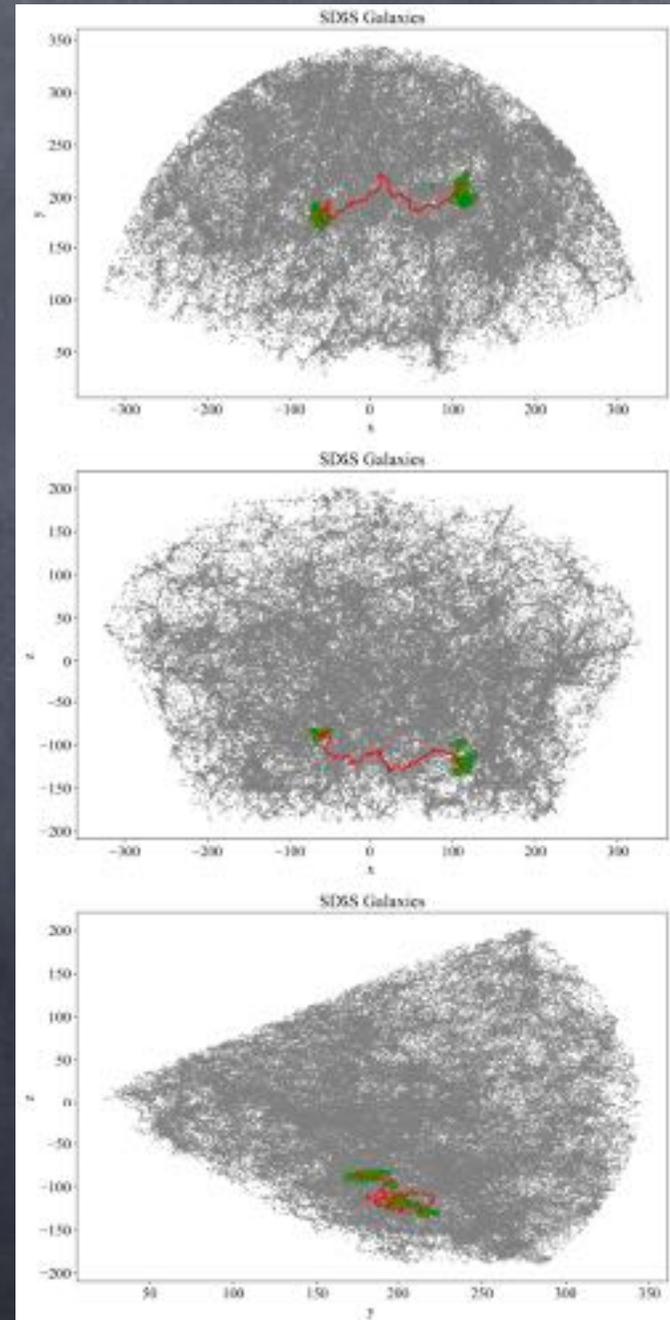
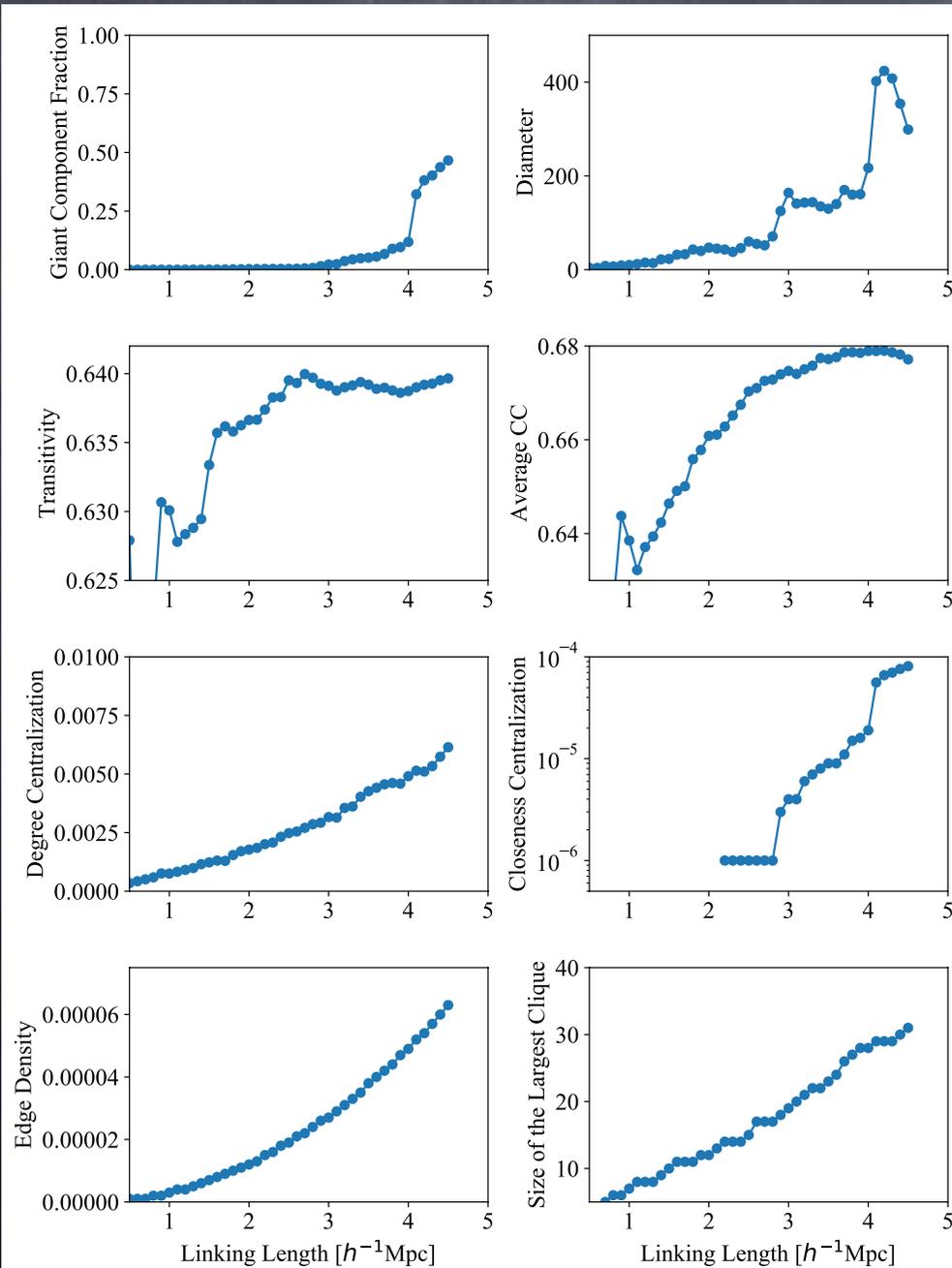
iGraph on Desktop

vs.

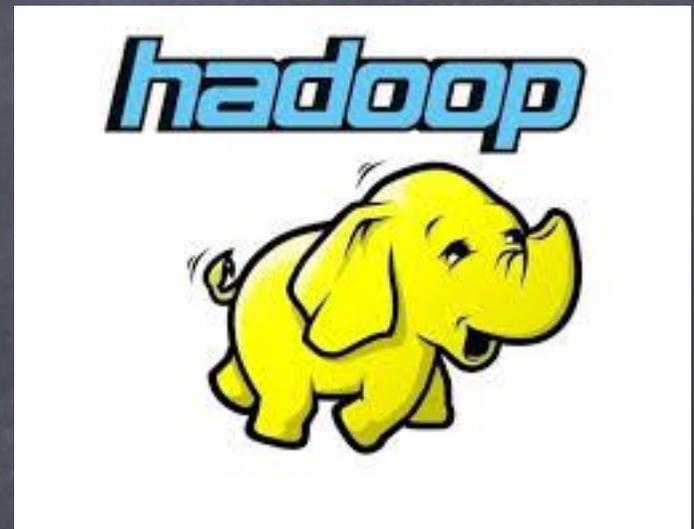
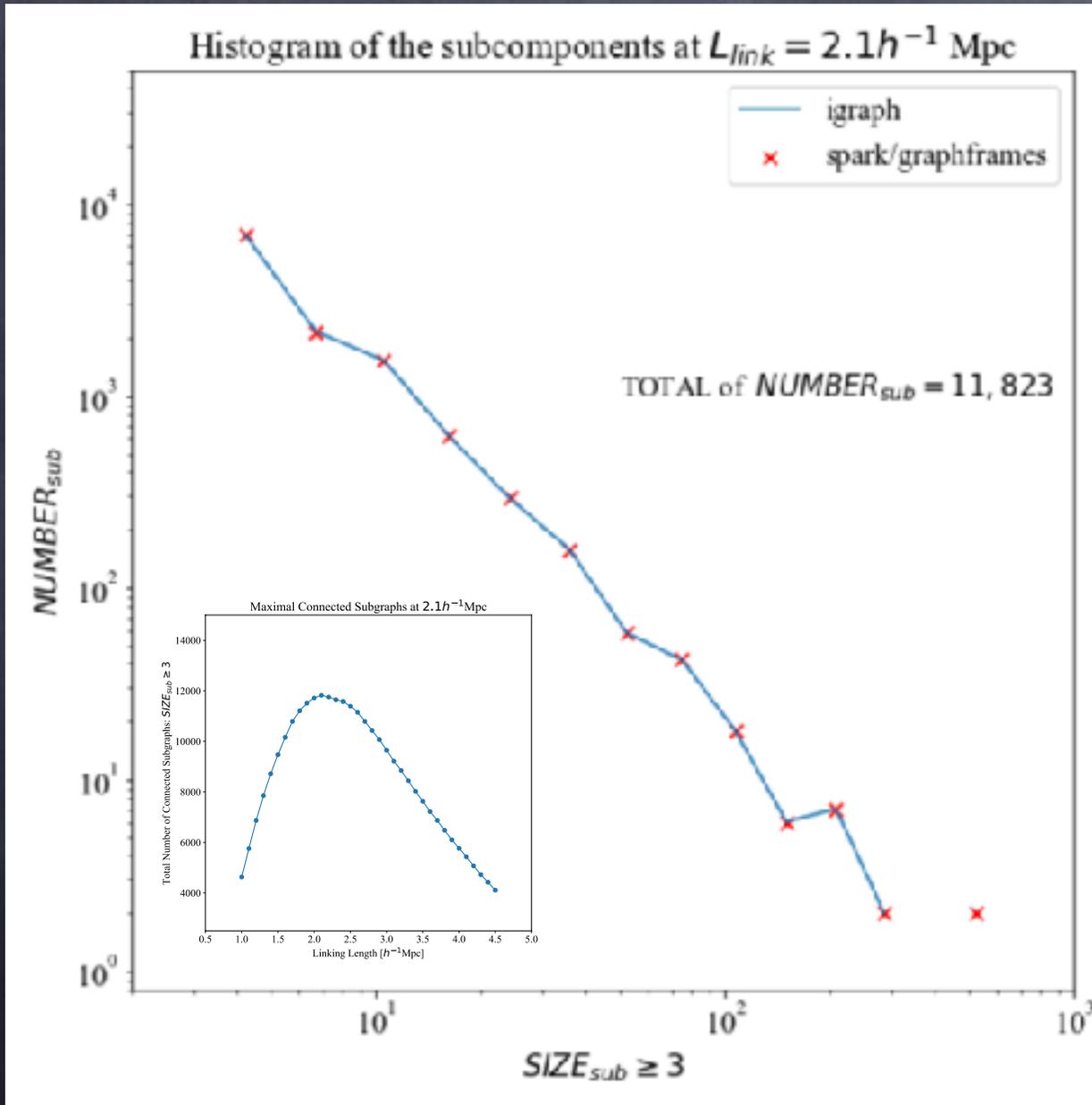
Spark GraphFrames on Spark Cluster

- **CMB hot and cold spots**
 - Identifying extrema points from CMB maps
 - Then, measuring graph statistics of CMB extrema!
- Network Epidemic Model in the Cosmological Context!
 - We may answer **the origin of life on Earth.**

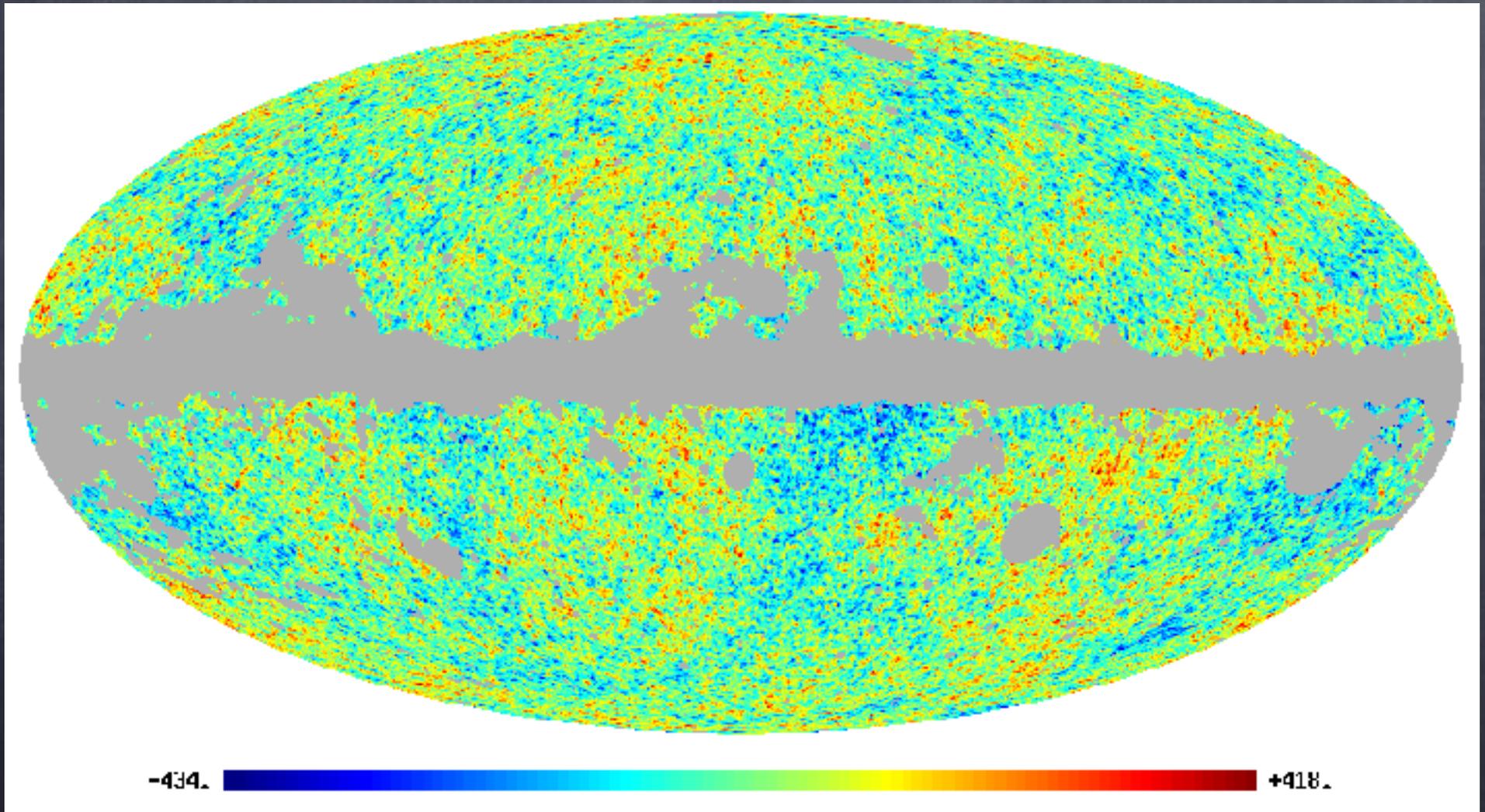
SDSS Skeleton Galaxies



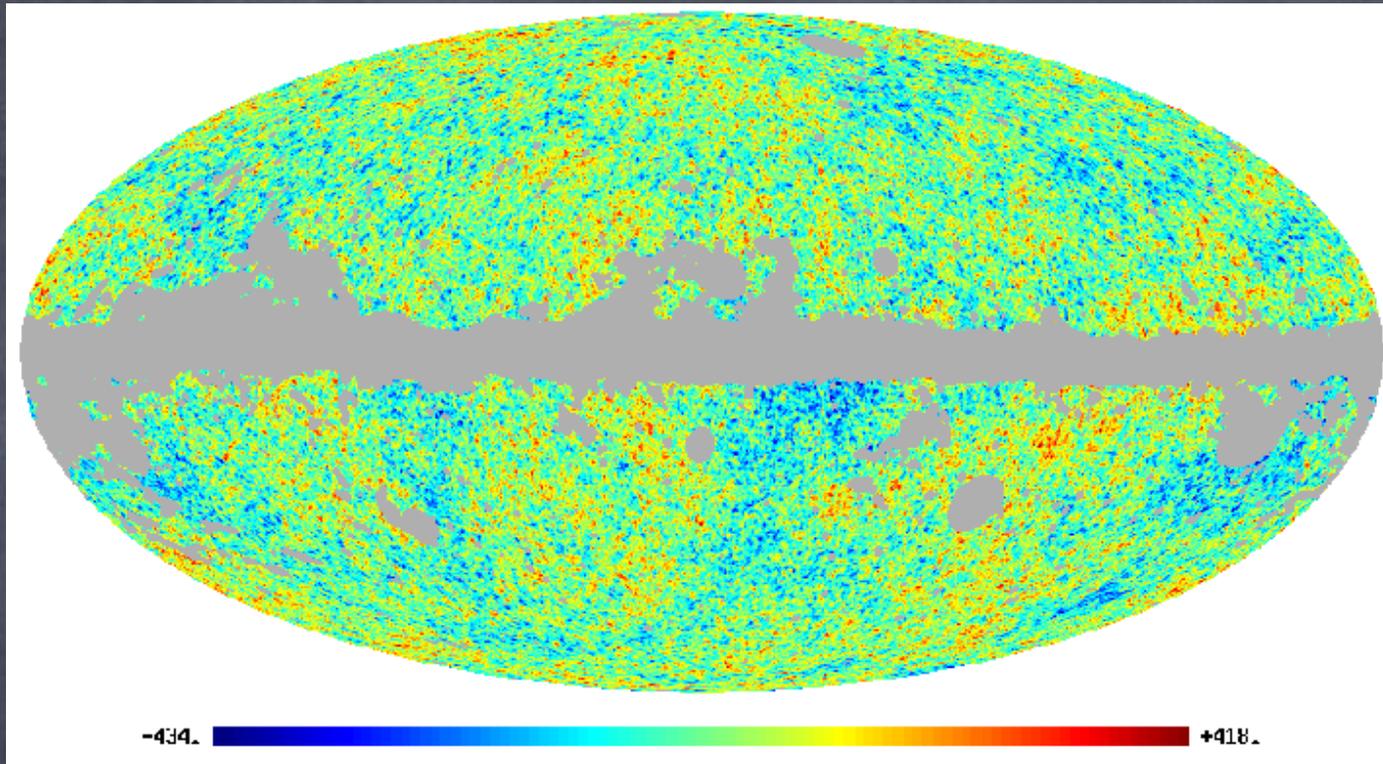
Graph Statistics of Multiverses



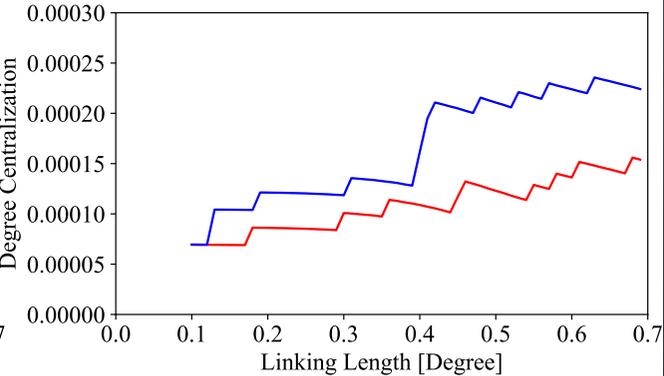
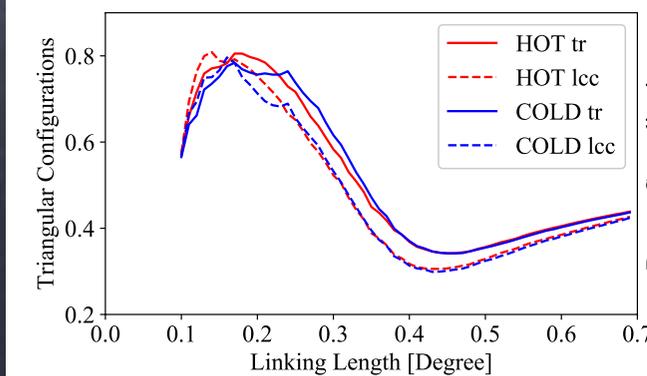
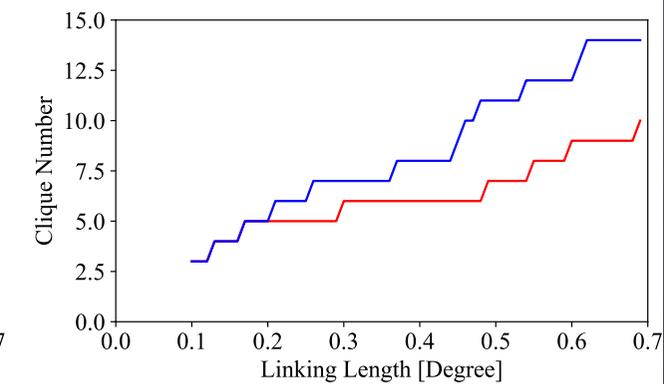
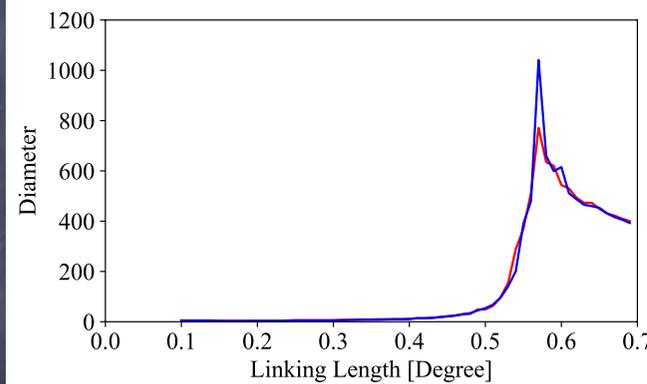
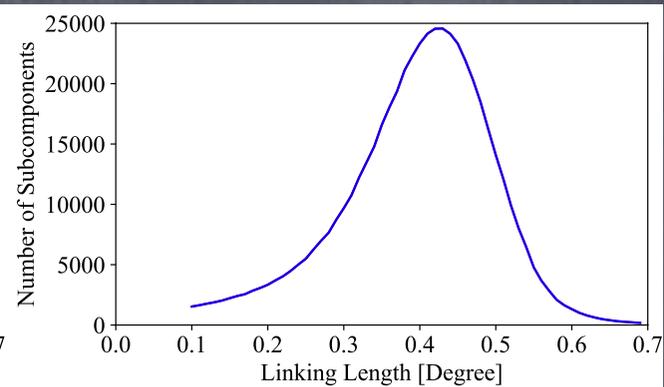
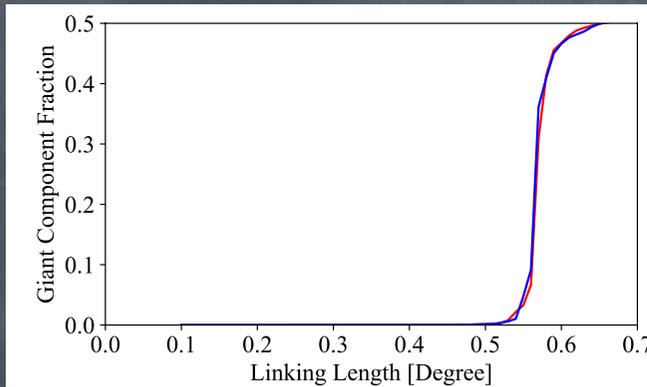
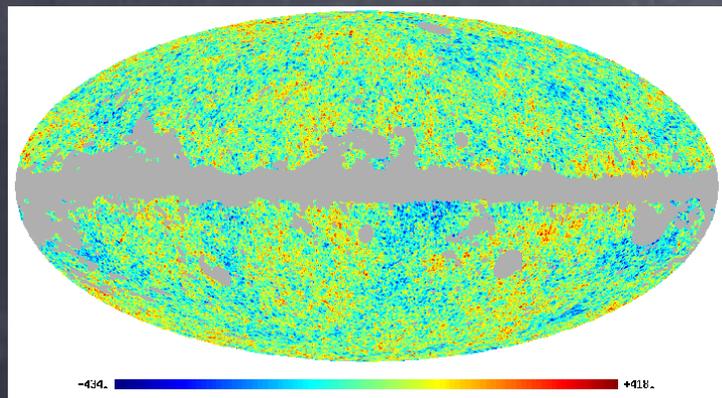
Non-Gaussianity in CMB?



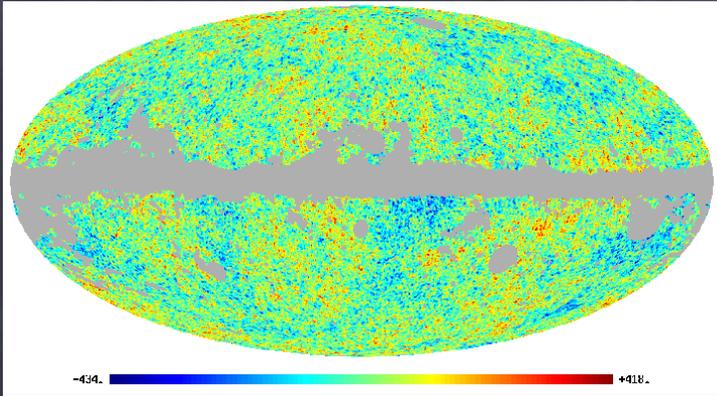
Non-Gaussianity in CMB?



Non-Gaussianity in CMB?



Non-Gaussianity in CMB?

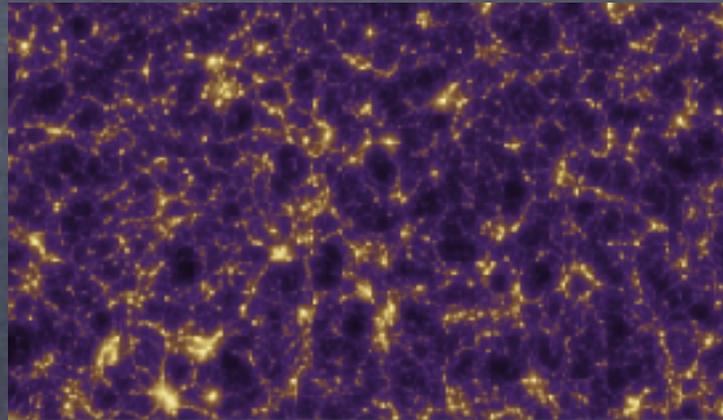


We will see ...

One More Future Thing ...

- What if we mix :

Cosmological Simulation



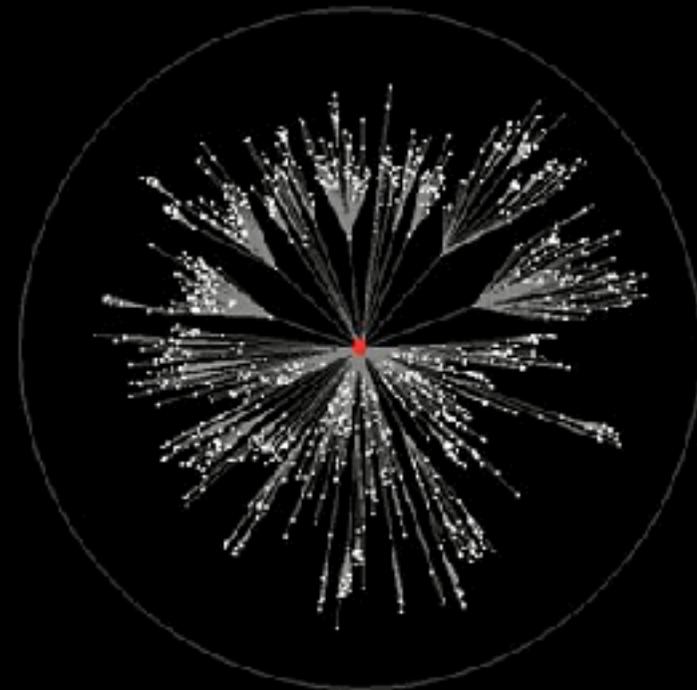
+

Network Epidemic Model



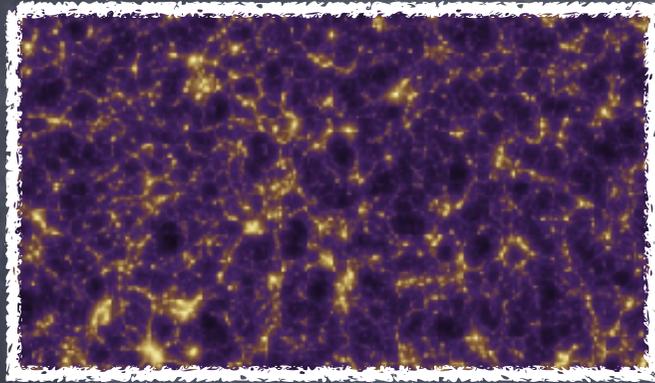
One More Future Thing ...

- Epidemic Model on Networks

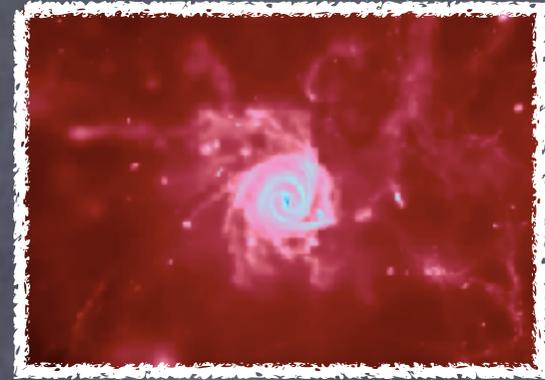


One More Future Thing ...

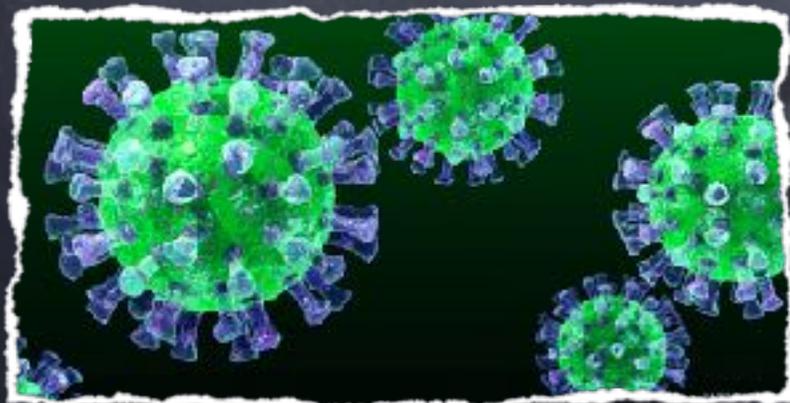
- Zoom-in Simulation for a MilkyWay-like galaxy



KIAS Horizon Run



Illustris Simulation



Interstellar Infection of Microbial Life



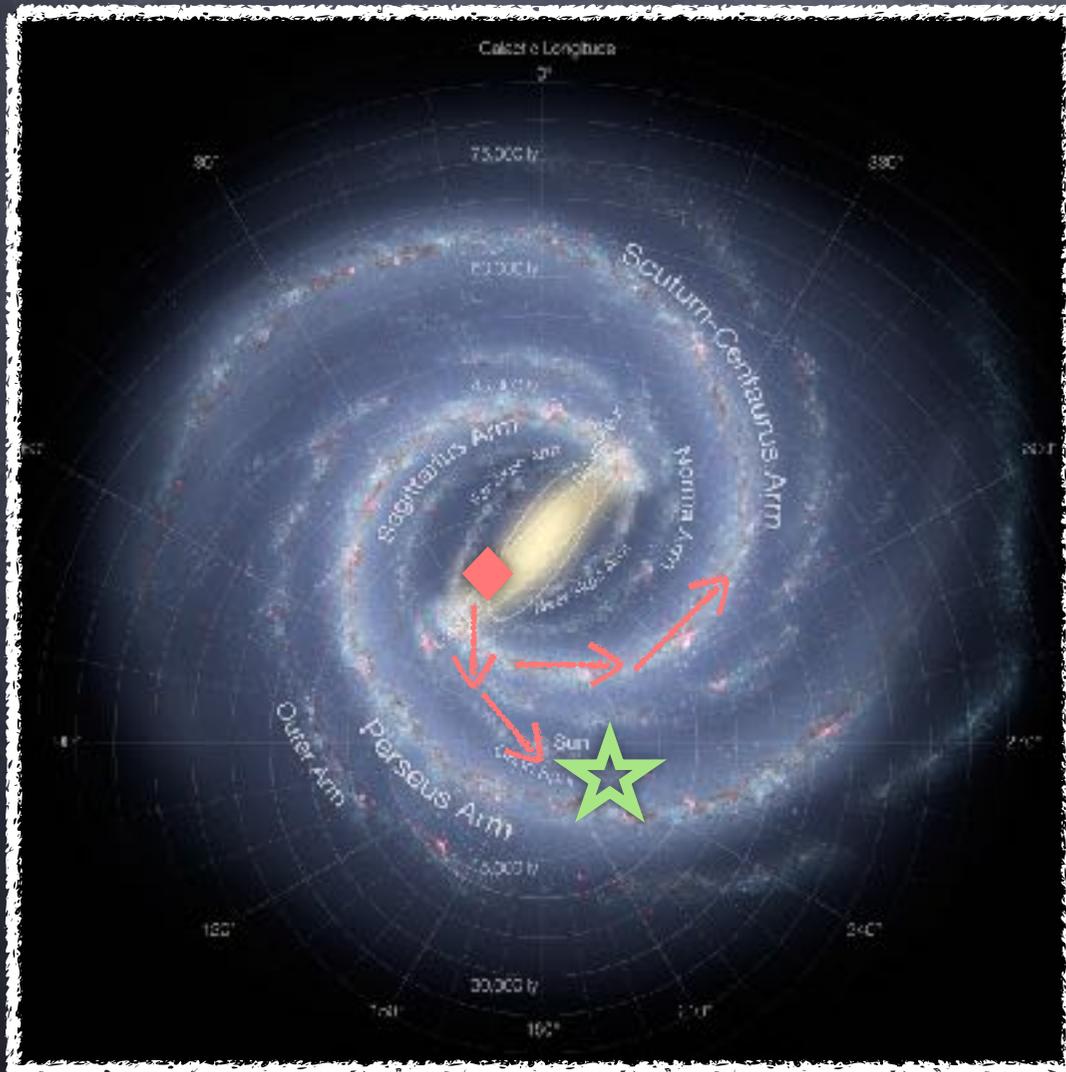
Stellar Network (from Caleb Jones' blog)

One More Future Thing ...

- What kind of implications can we get from this work?

One More Future Thing ...

- What kind of implications can we get from this work?



The origin of life on Earth :

in-situ genesis?

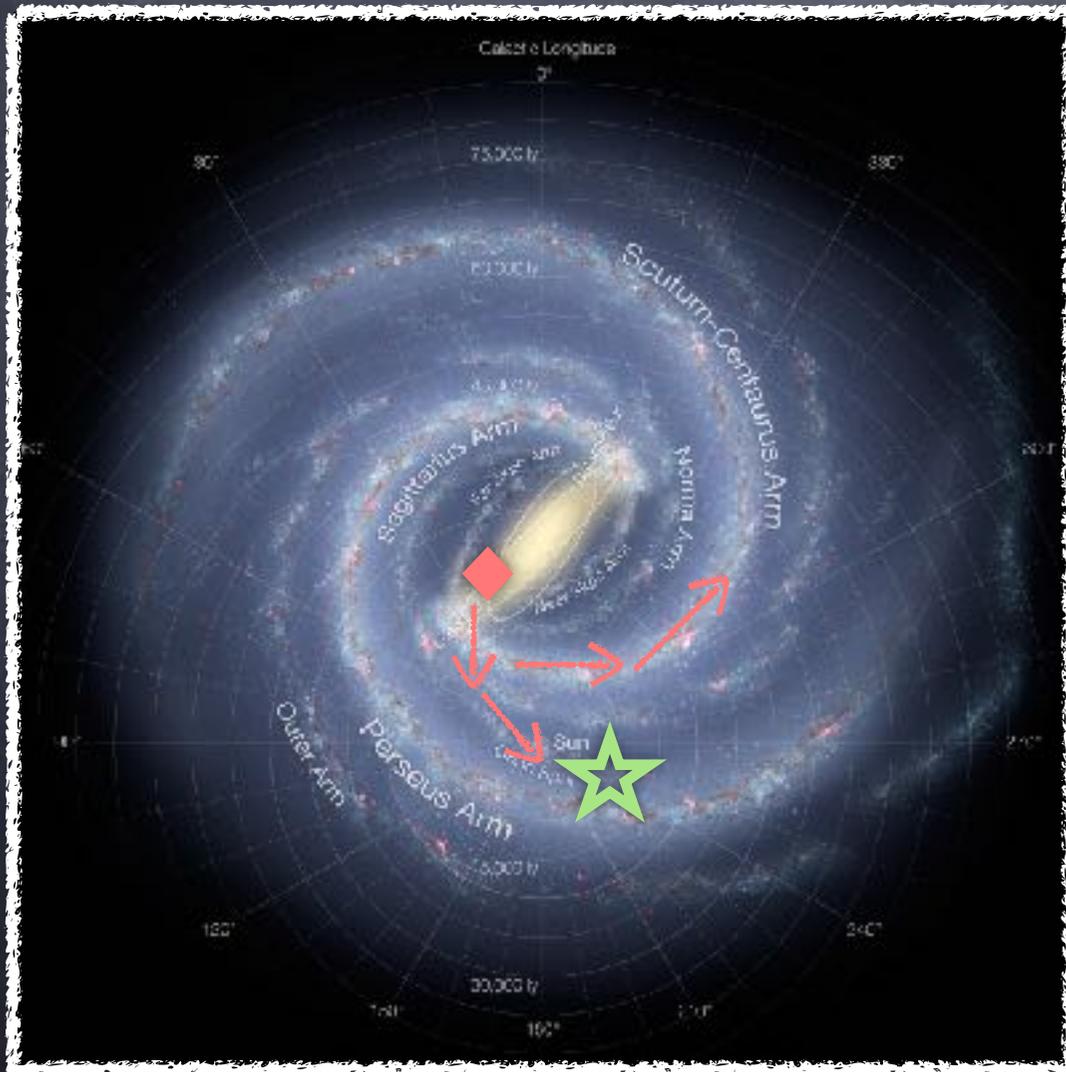
or

exo-genesis
by interstellar infection?

- Low Contagion Probability:
"Isolated Drake Equation"
- High Contagion Probability:
"Galactic Ecosystem"

One More Future Thing ...

- What kind of implications can we get from this work?



The origin of life on Earth :

in-situ genesis?

or

exo-genesis
by interstellar infection?

We may be able to answer which scenario is more reliable for explaining the origin of life on Earth.

Thank you

Are you Pro-ML or Anti-ML?

IRAF, IDL

Python



PySpark

