



# **Machine learning quantification of cluster structure**

**Oliver Lomax**  
**Matthew Bates & Anthony Whitworth**

**30th August 2018**  
**Star Form Mapper Workshop**

# Outline

# Outline

- Q parameter, used to measure substructure in star clusters, is unreliable.

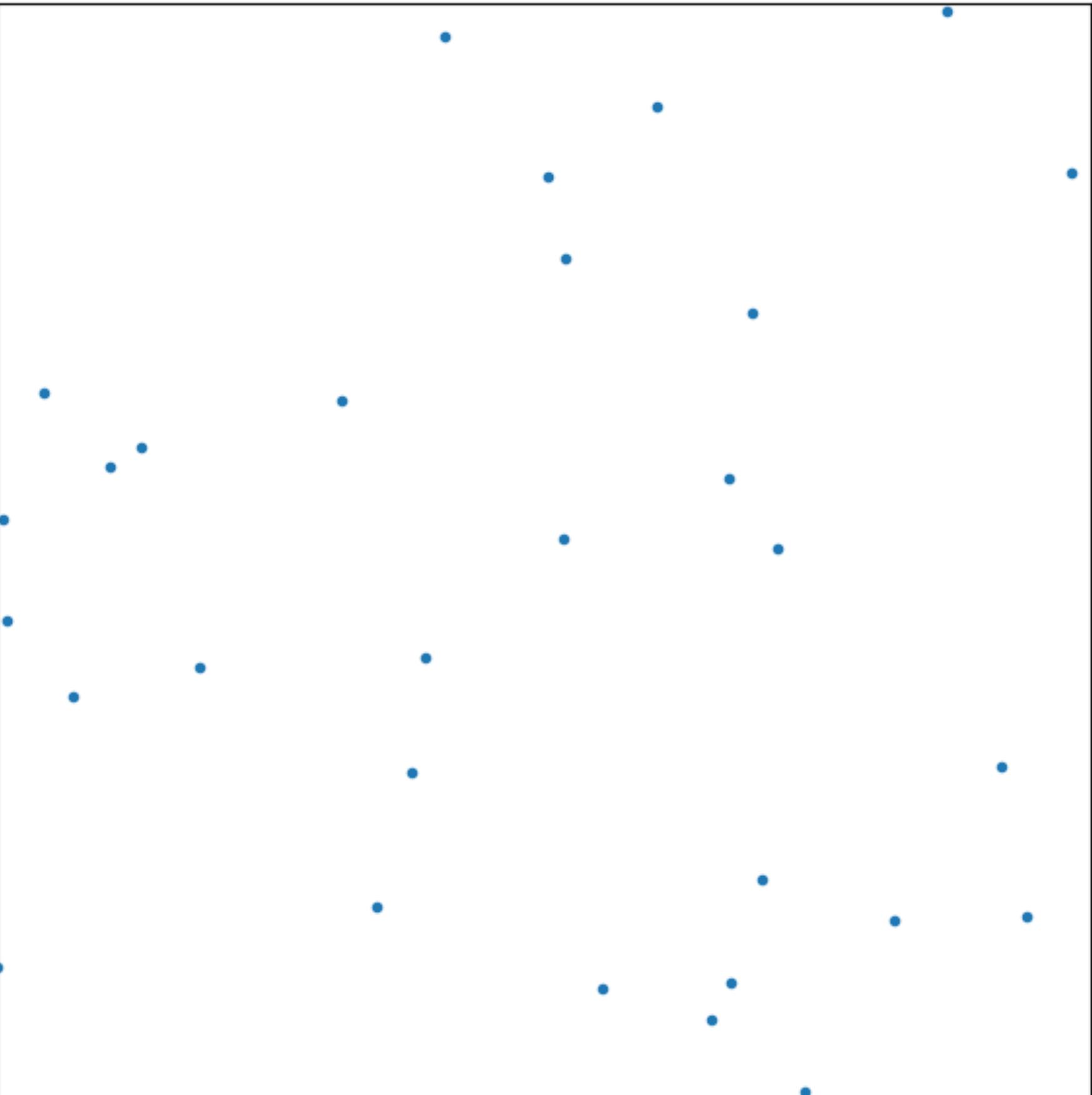
# Outline

- Q parameter, used to measure substructure in star clusters, is unreliable.
- We have tools available to do better!

# Outline

- Q parameter, used to measure substructure in star clusters, is unreliable.
- We have tools available to do better!
- Disclaimer: I am neither an observer or an  $N$ -body numericist!

# Q parameter



$$Q = \frac{\bar{m}}{\bar{s}}$$

$$\bar{m} = \frac{\sum_{i=1}^{N_m} m_i}{(N_m + 1)^{\frac{E-1}{E}} R}$$

$$\bar{s} = \frac{\sum_{i=1}^{N_s} s_i}{N_s R}$$

# Q parameter



Minimum Spanning Tree

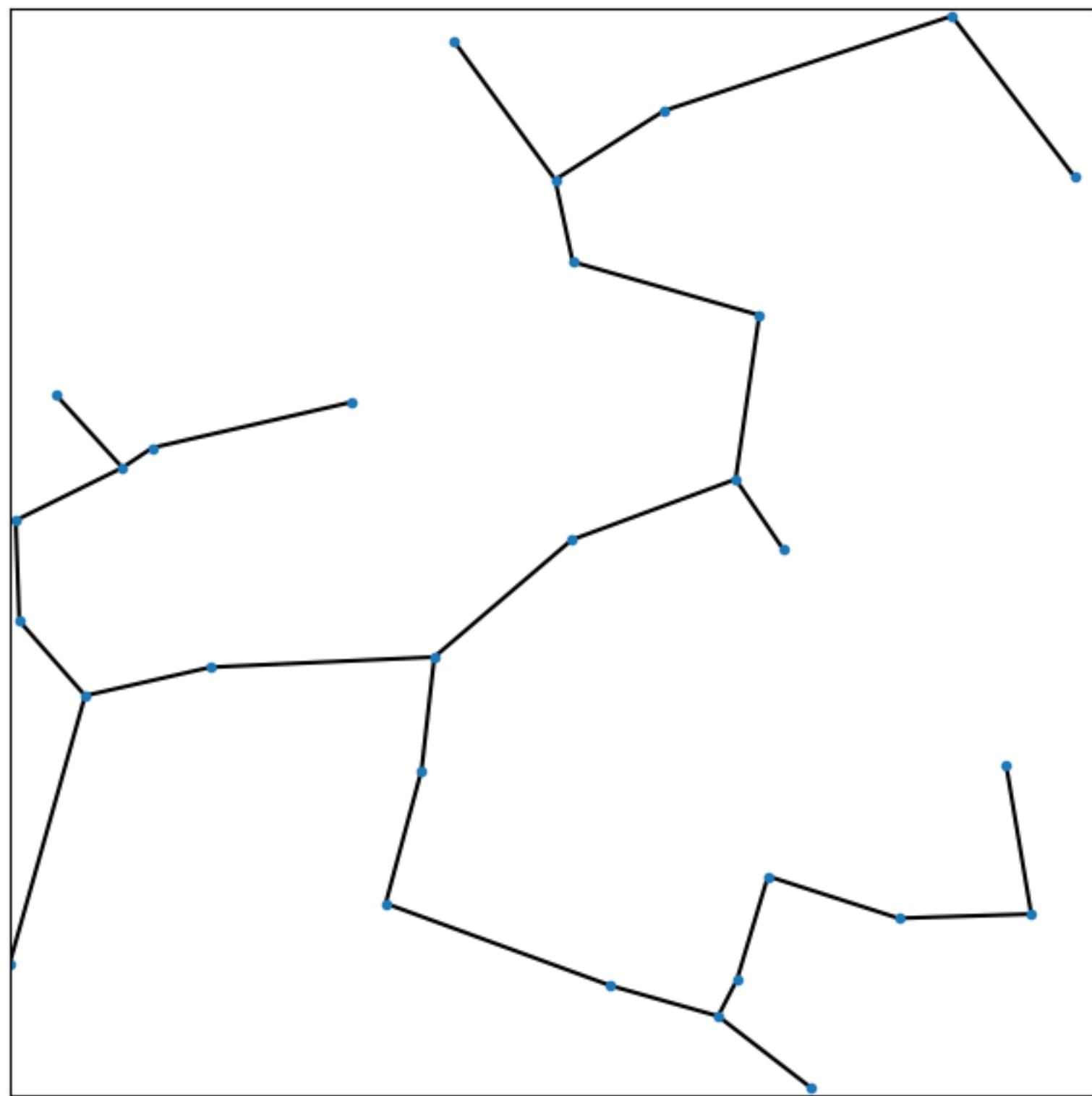
$$Q = \frac{\bar{m}}{\bar{s}}$$

$$\bar{m} = \frac{\sum_{i=1}^{N_m} m_i}{(N_m + 1)^{\frac{E-1}{E}} R}$$

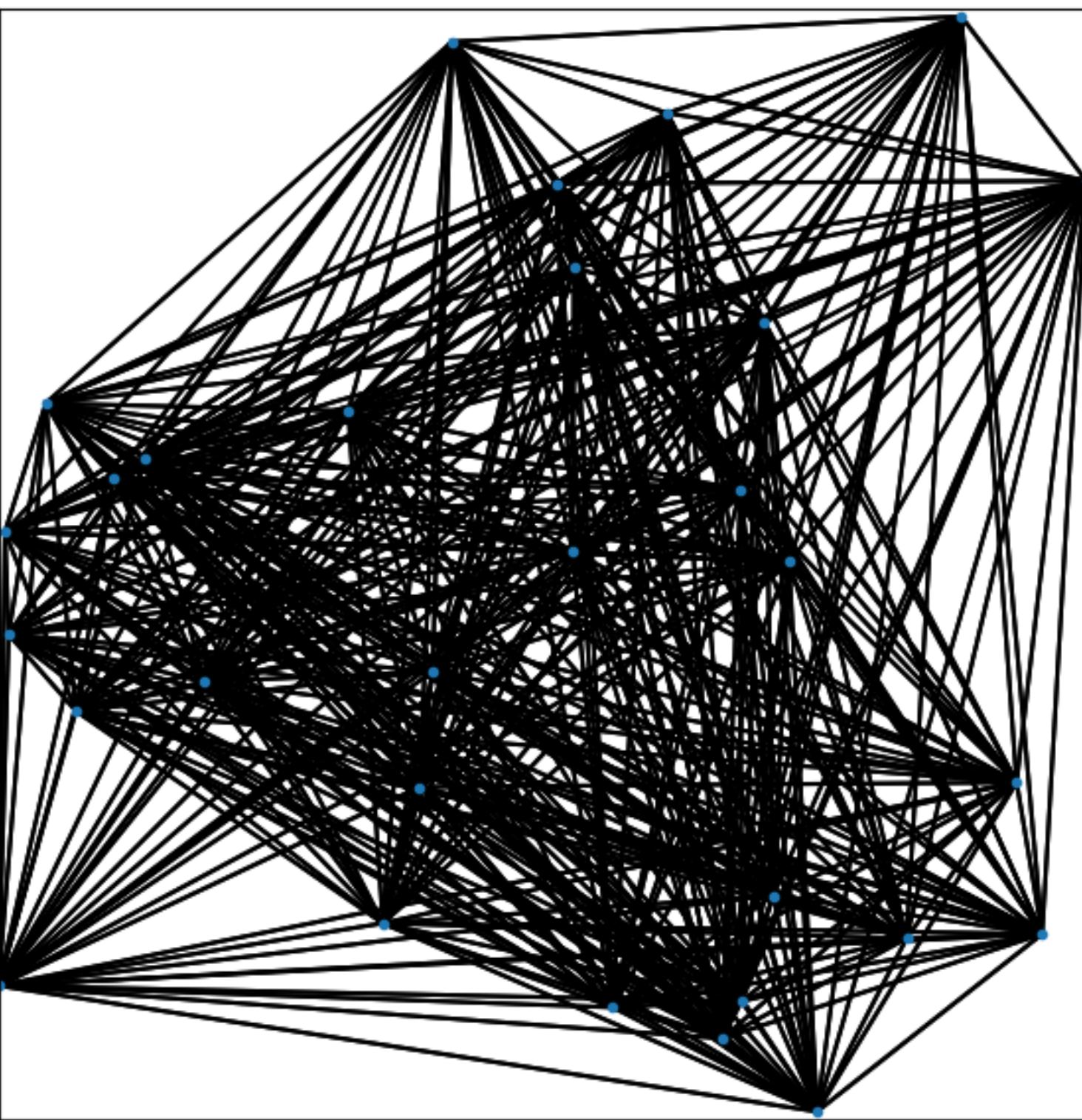
$$\bar{s} = \frac{\sum_{i=1}^{N_s} s_i}{N_s R}$$

Cartwright and Whitworth, 2004,  
MNRAS, 348, 589

# Q parameter



Minimum Spanning Tree



Complete Graph

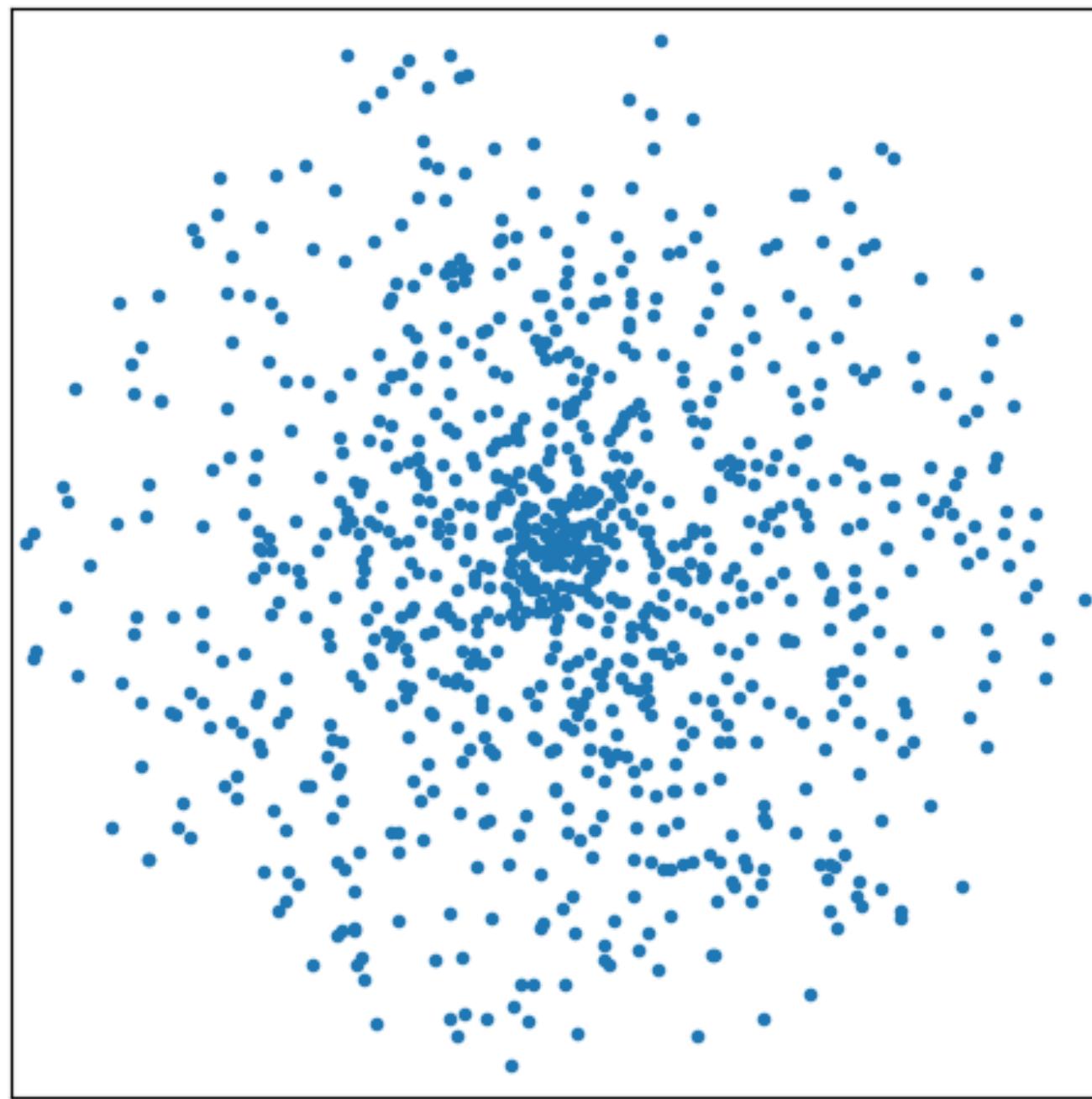
$$Q = \frac{\bar{m}}{\bar{s}}$$

$$\bar{m} = \frac{\sum_{i=1}^{N_m} m_i}{(N_m + 1)^{\frac{E-1}{E}} R}$$

$$\bar{s} = \frac{\sum_{i=1}^{N_s} s_i}{N_s R}$$

Cartwright and Whitworth, 2004,  
MNRAS, 348, 589

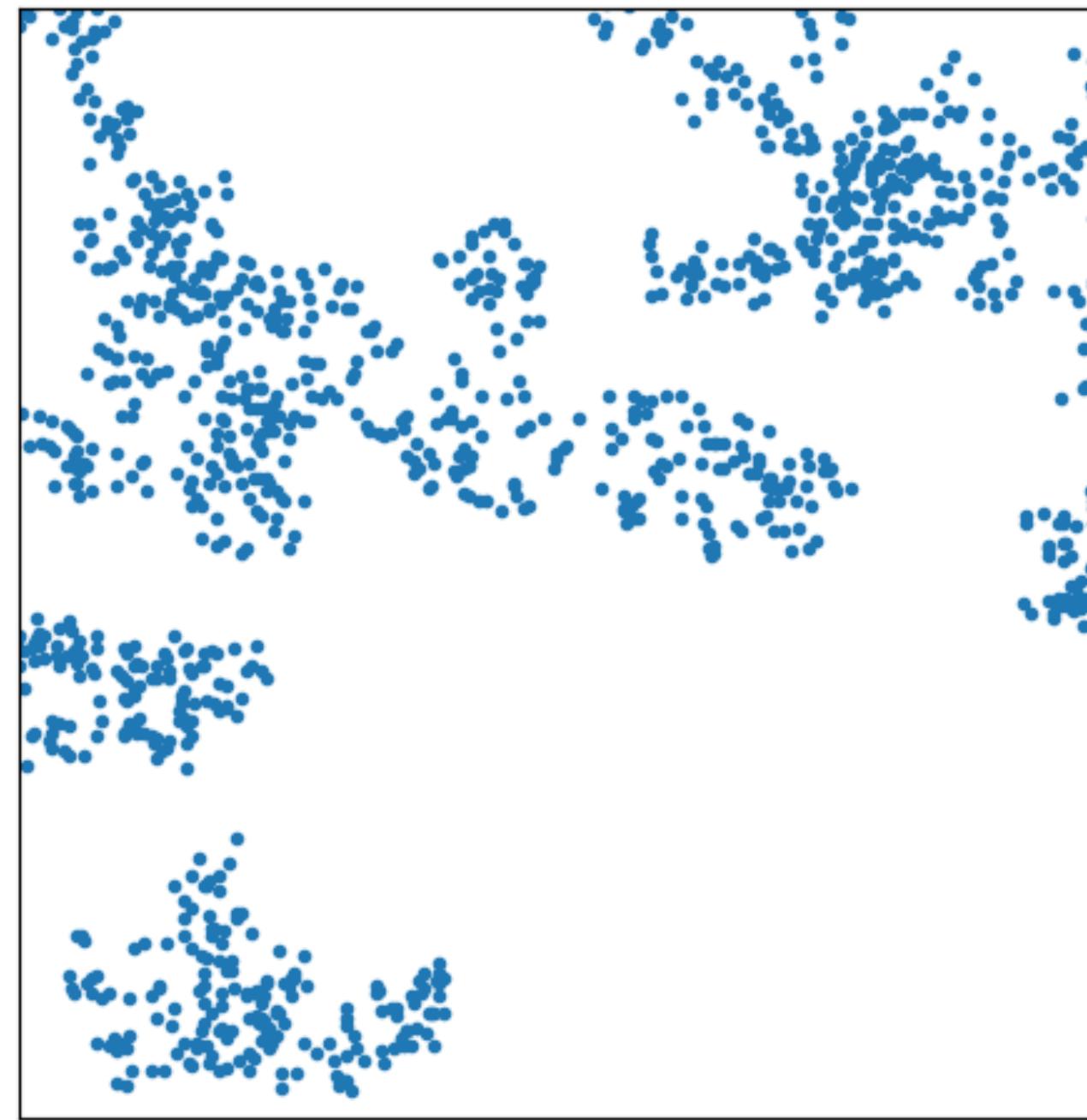
# Q parameter



**Radial Profile**

$$\rho(r) \propto r^{-\alpha}$$

$$Q \gtrsim 0.8$$



**Box Fractal**

$$P(\text{sub-cube}) = \frac{2^D}{2^E}$$

$$Q \lesssim 0.8$$

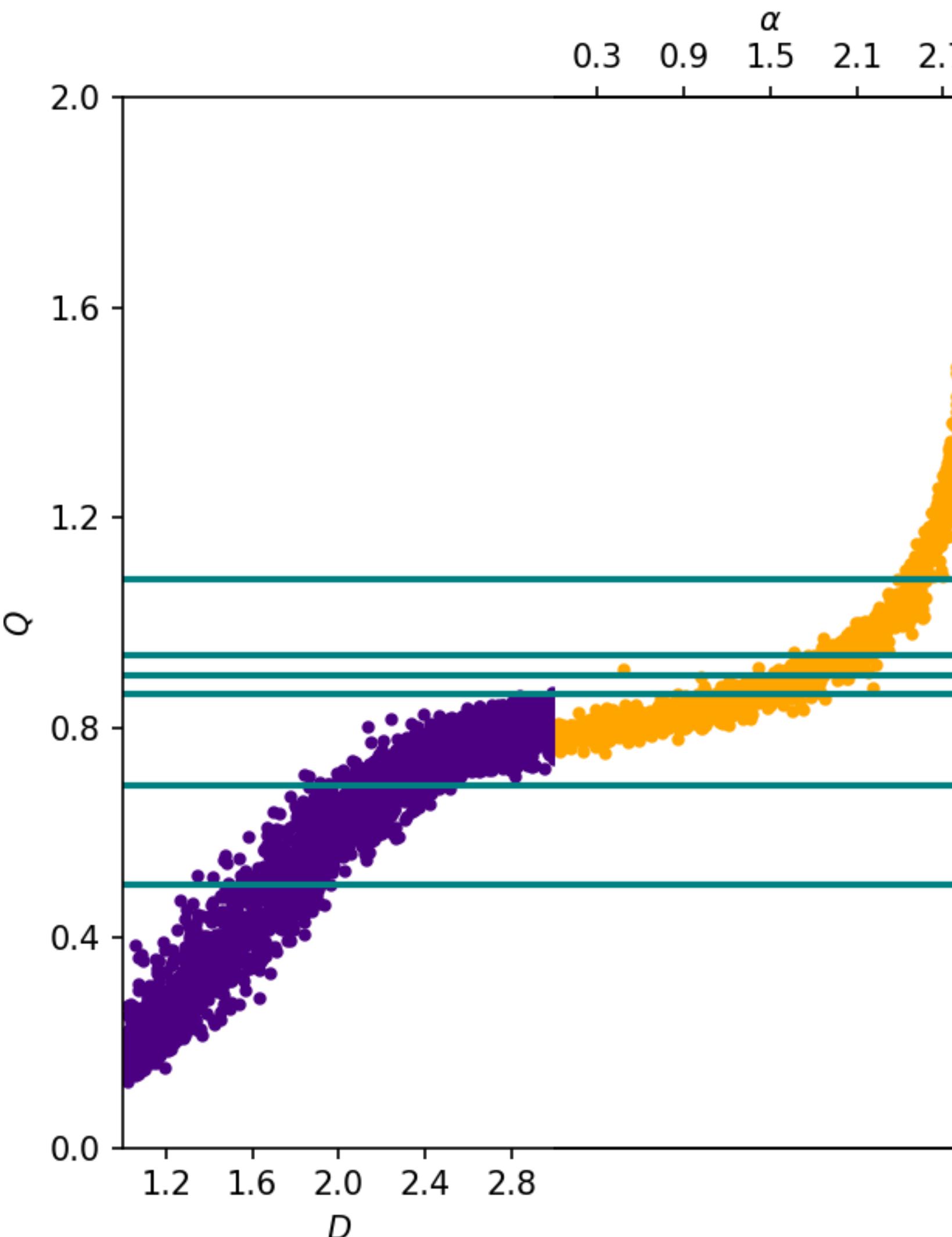
$$Q = \frac{\bar{m}}{\bar{s}}$$

$$\bar{m} = \frac{\sum_{i=1}^{N_m} m_i}{(N_m + 1)^{\frac{E-1}{E}} R}$$

$$\bar{s} = \frac{\sum_{i=1}^{N_s} s_i}{N_s R}$$

Cartwright and Whitworth, 2004,  
MNRAS, 348, 589

# Q parameter



| Cluster    | # | N   |
|------------|---|-----|
| Lupus 3    | 1 | 67  |
| IC 348     | 2 | 350 |
| $\rho$ Oph | 3 | 198 |
| IC 2391    | 4 | 200 |
| Cha I      | 5 | 234 |
| Taurus     | 6 | 335 |

Comerón (2008)  
 Lada et al. (2006); Muench et al. (2007)  
 Bontemps et al. (2001)  
 Barrado y Navascués et al. (2001)  
 Luhman (2007)  
 Luhman et al. (2010)

$$Q = \frac{\bar{m}}{\bar{s}}$$

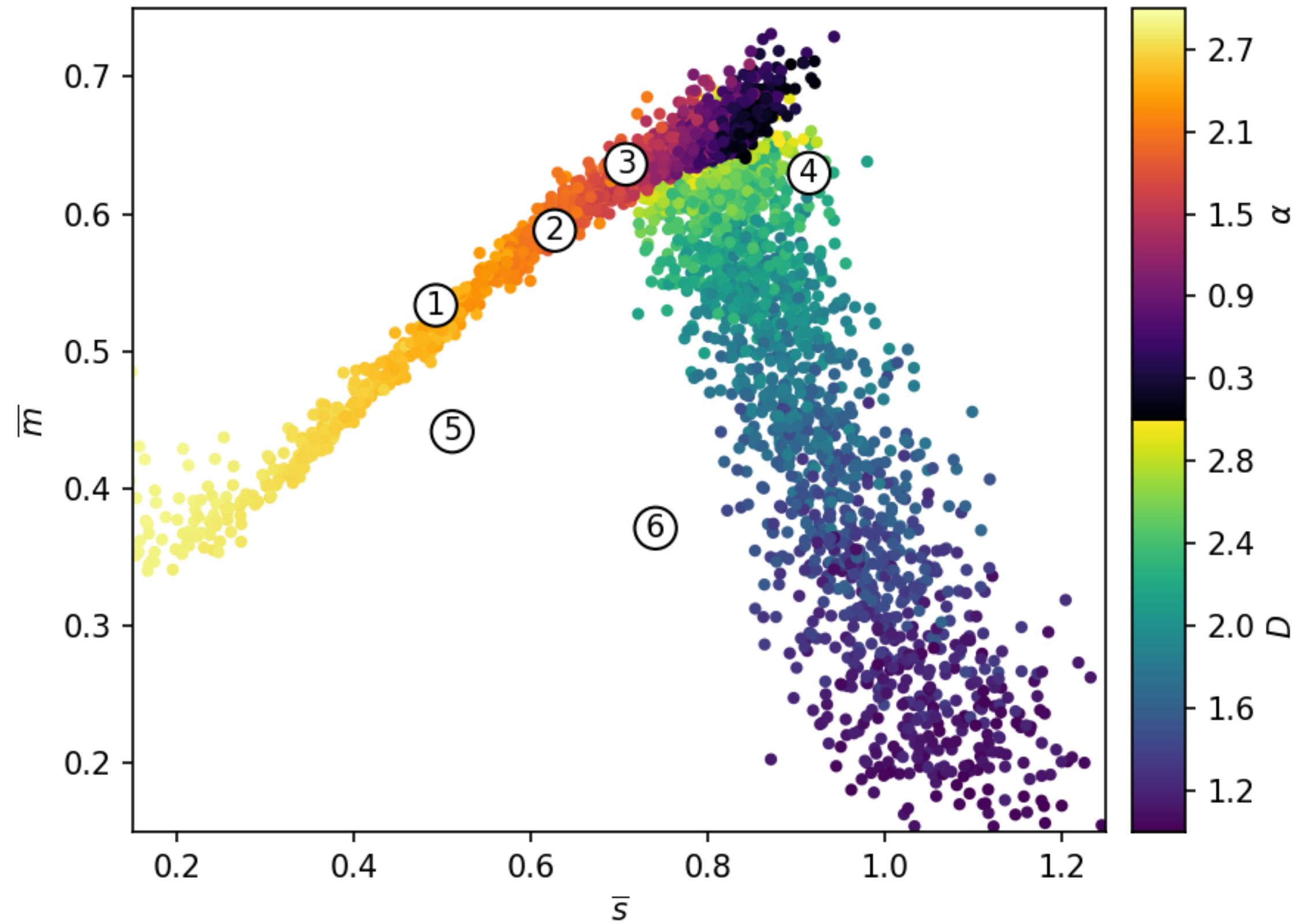
$$\bar{m} = \frac{\sum_{i=1}^{N_m} m_i}{(N_m + 1)^{\frac{E-1}{E}} R}$$

$$\bar{s} = \frac{\sum_{i=1}^{N_s} s_i}{N_s R}$$

Cartwright and Whitworth, 2004,  
 MNRAS, 348, 589

# $\bar{m}$ - $\bar{s}$ plots

| Cluster    | # | N   |
|------------|---|-----|
| Lupus 3    | 1 | 67  |
| IC 348     | 2 | 350 |
| $\rho$ Oph | 3 | 198 |
| IC 2391    | 4 | 200 |
| Cha I      | 5 | 234 |
| Taurus     | 6 | 335 |

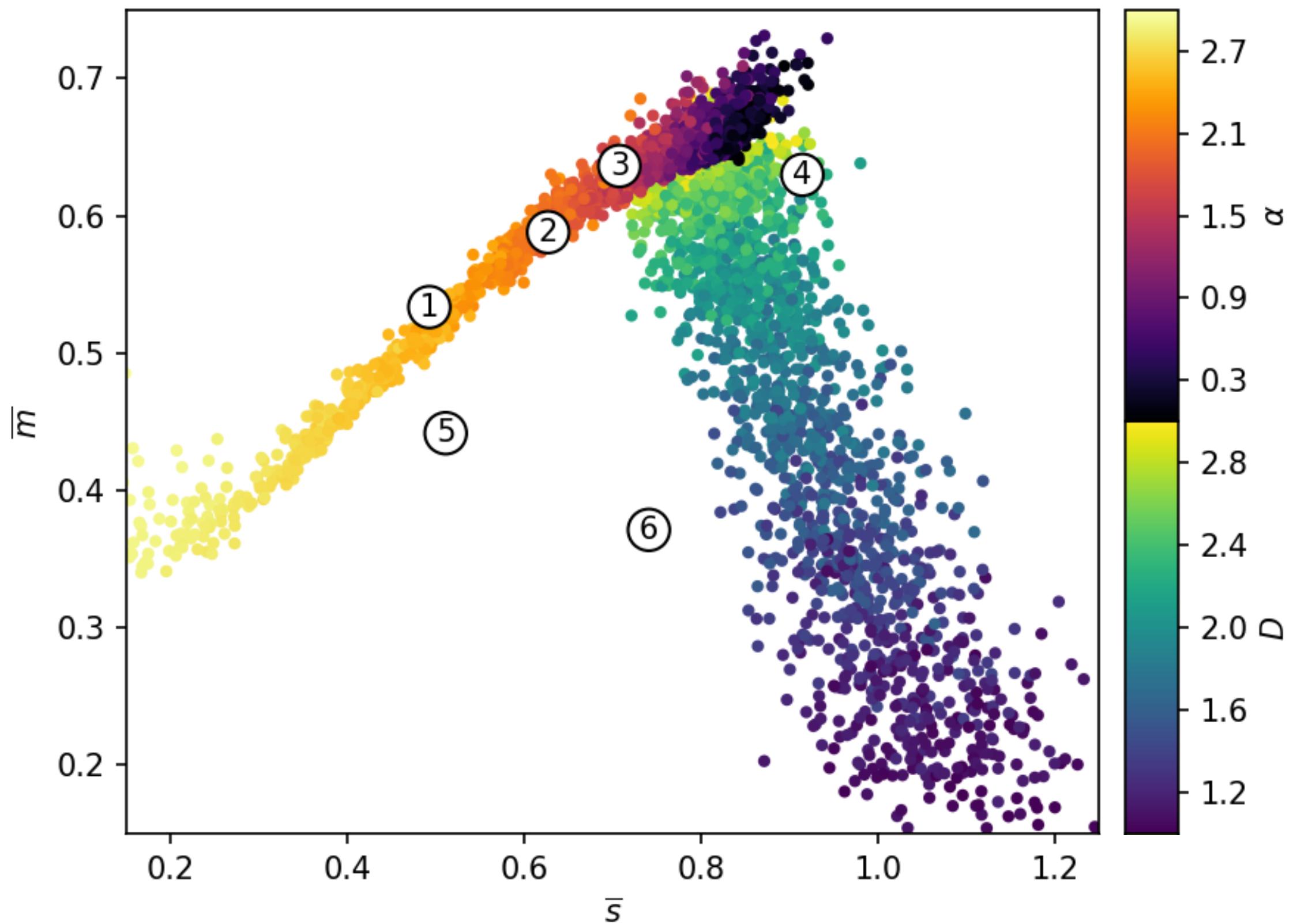


Comerón (2008)  
Lada et al. (2006); Muench et al. (2007)  
Bontemps et al. (2001)  
Barrado y Navascués et al. (2001)  
Luhman (2007)  
Luhman et al. (2010)

Cartwright, 2009, MNRAS, 400, 1427

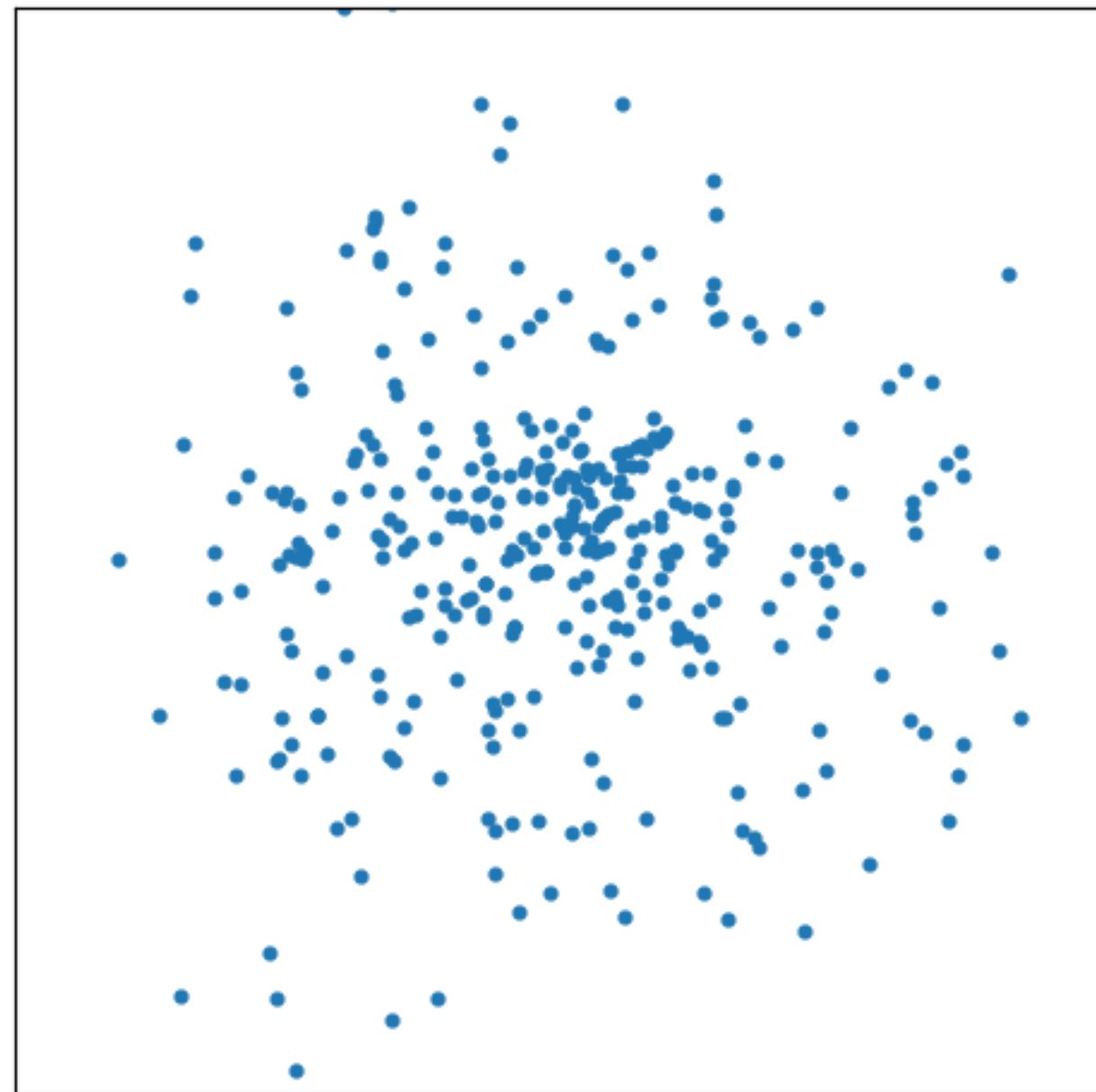
# $\bar{m}$ - $\bar{s}$ plots

| Cluster    | # | N   |
|------------|---|-----|
| Lupus 3    | 1 | 67  |
| IC 348     | 2 | 350 |
| $\rho$ Oph | 3 | 198 |
| IC 2391    | 4 | 200 |
| Cha I      | 5 | 234 |
| Taurus     | 6 | 335 |



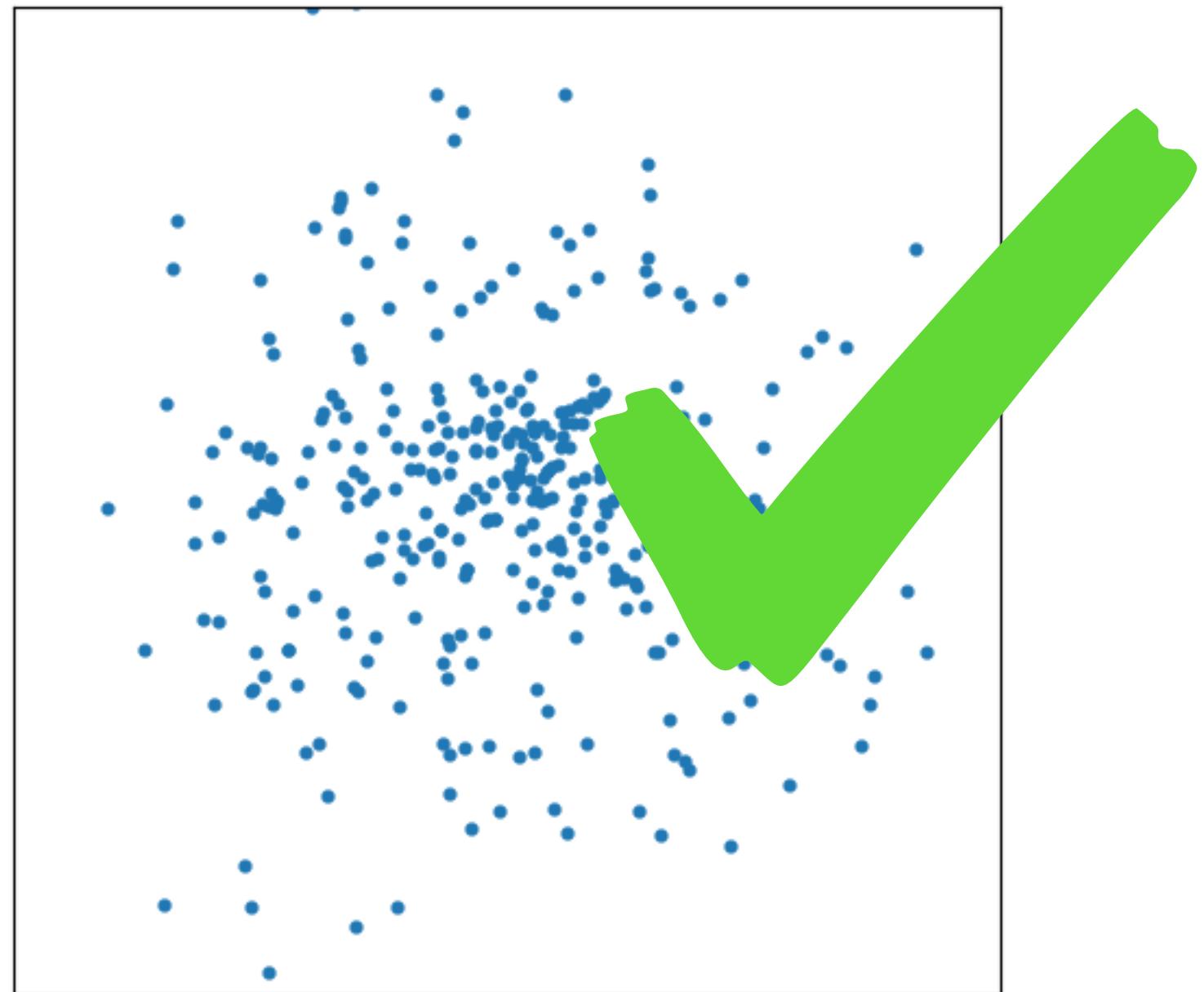
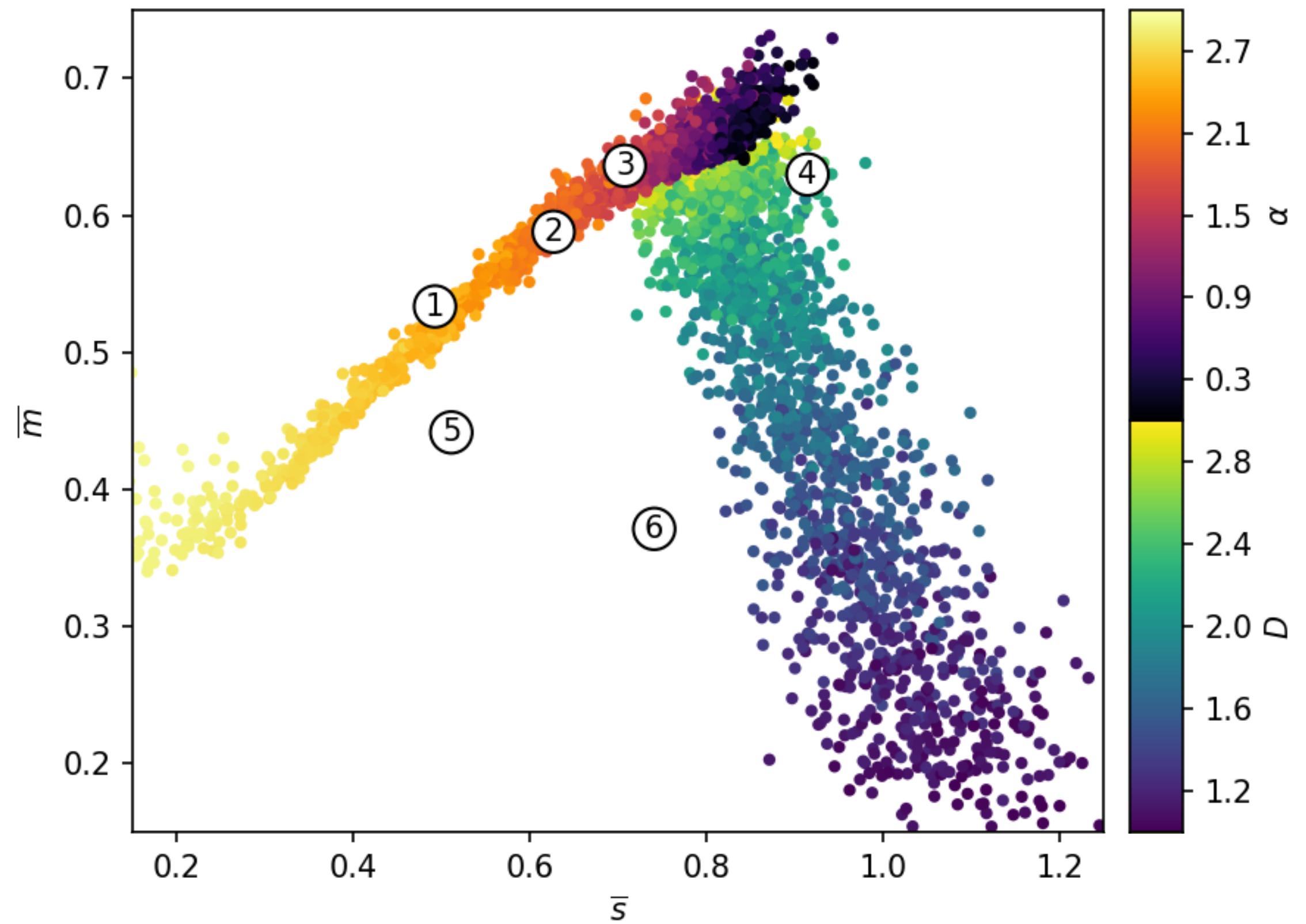
Comerón (2008)  
Lada et al. (2006); Muench et al. (2007)  
Bontemps et al. (2001)  
Barrado y Navascués et al. (2001)  
Luhman (2007)  
Luhman et al. (2010)

Cartwright, 2009, MNRAS, 400, 1427



# $\bar{m}$ - $\bar{s}$ plots

| Cluster    | # | N   |
|------------|---|-----|
| Lupus 3    | 1 | 67  |
| IC 348     | 2 | 350 |
| $\rho$ Oph | 3 | 198 |
| IC 2391    | 4 | 200 |
| Cha I      | 5 | 234 |
| Taurus     | 6 | 335 |

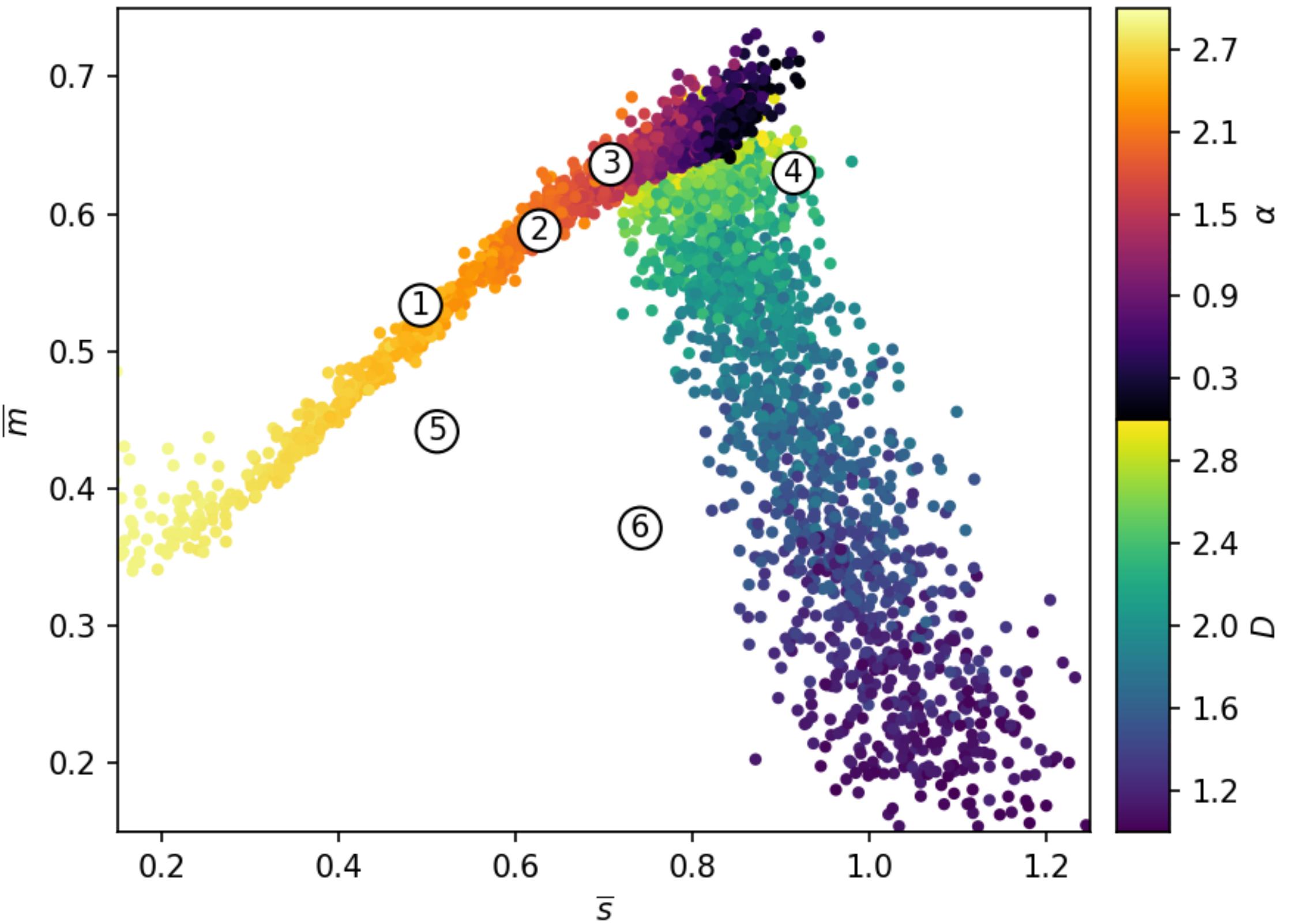


Comerón (2008)  
 Lada et al. (2006); Muench et al. (2007)  
 Bontemps et al. (2001)  
 Barrado y Navascués et al. (2001)  
 Luhman (2007)  
 Luhman et al. (2010)

Cartwright, 2009, MNRAS, 400, 1427

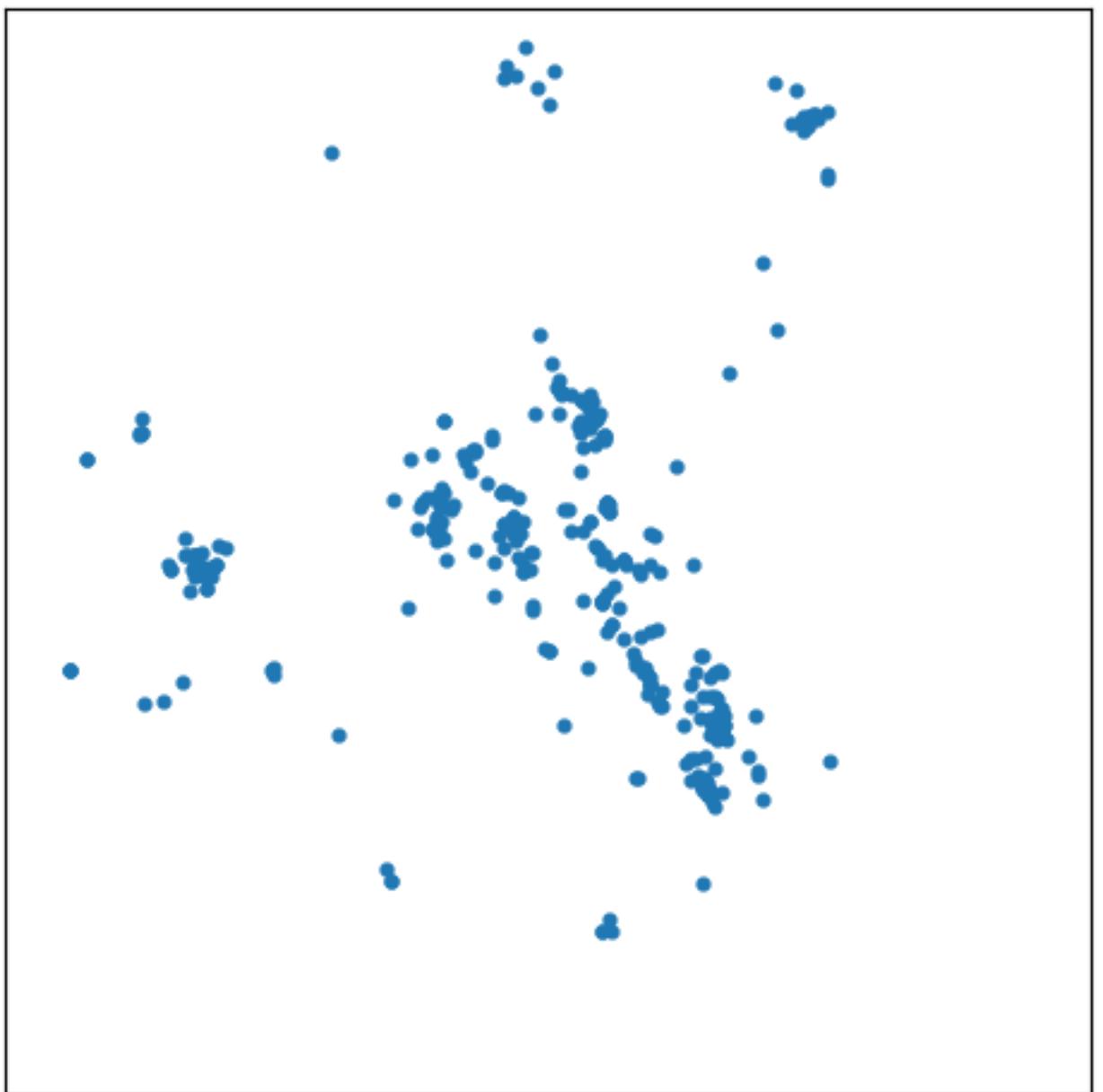
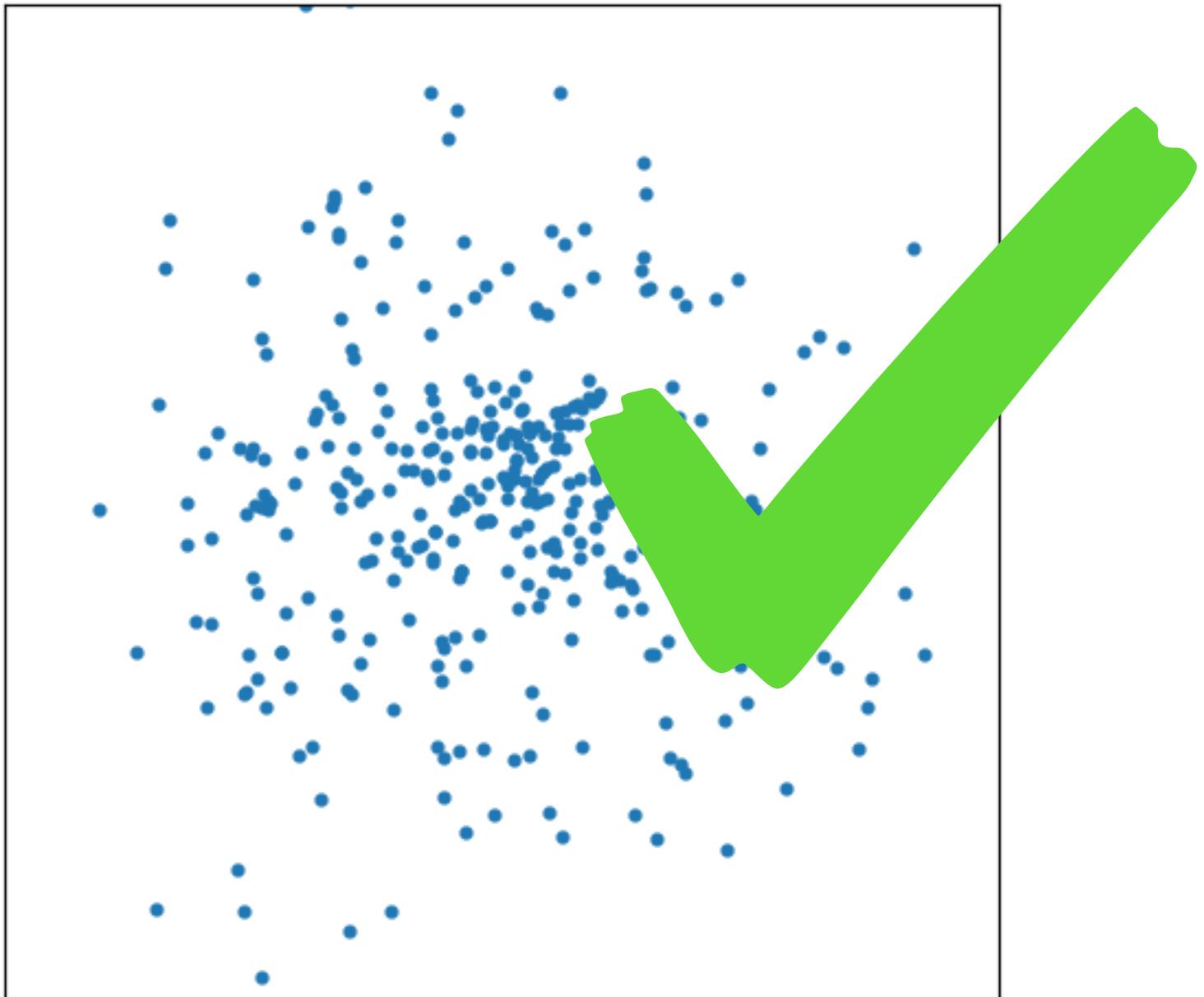
# $\bar{m}$ - $\bar{s}$ plots

| Cluster    | # | N   |
|------------|---|-----|
| Lupus 3    | 1 | 67  |
| IC 348     | 2 | 350 |
| $\rho$ Oph | 3 | 198 |
| IC 2391    | 4 | 200 |
| Cha I      | 5 | 234 |
| Taurus     | 6 | 335 |



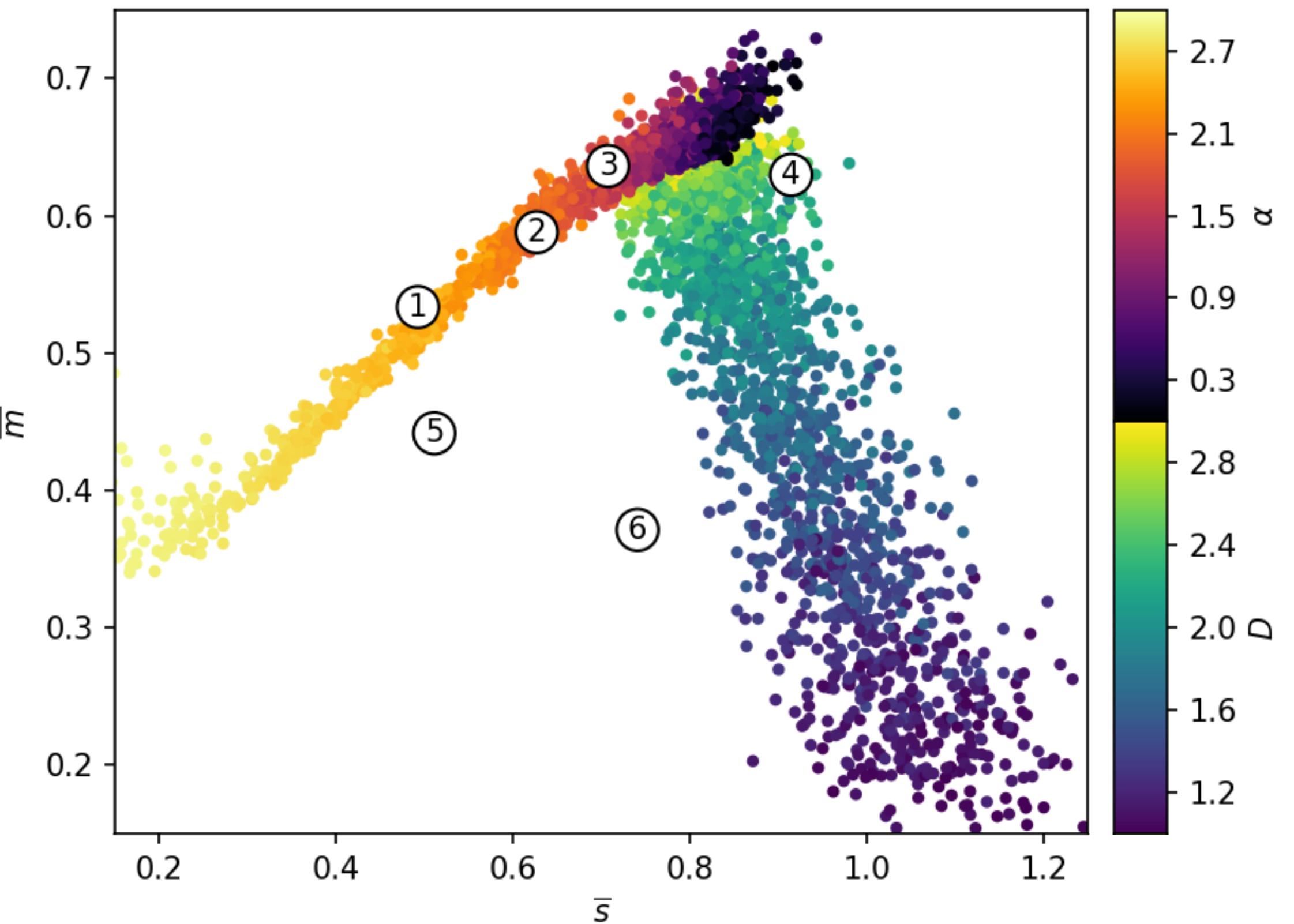
Comerón (2008)  
 Lada et al. (2006); Muench et al. (2007)  
 Bontemps et al. (2001)  
 Barrado y Navascués et al. (2001)  
 Luhman (2007)  
 Luhman et al. (2010)

Cartwright, 2009, MNRAS, 400, 1427



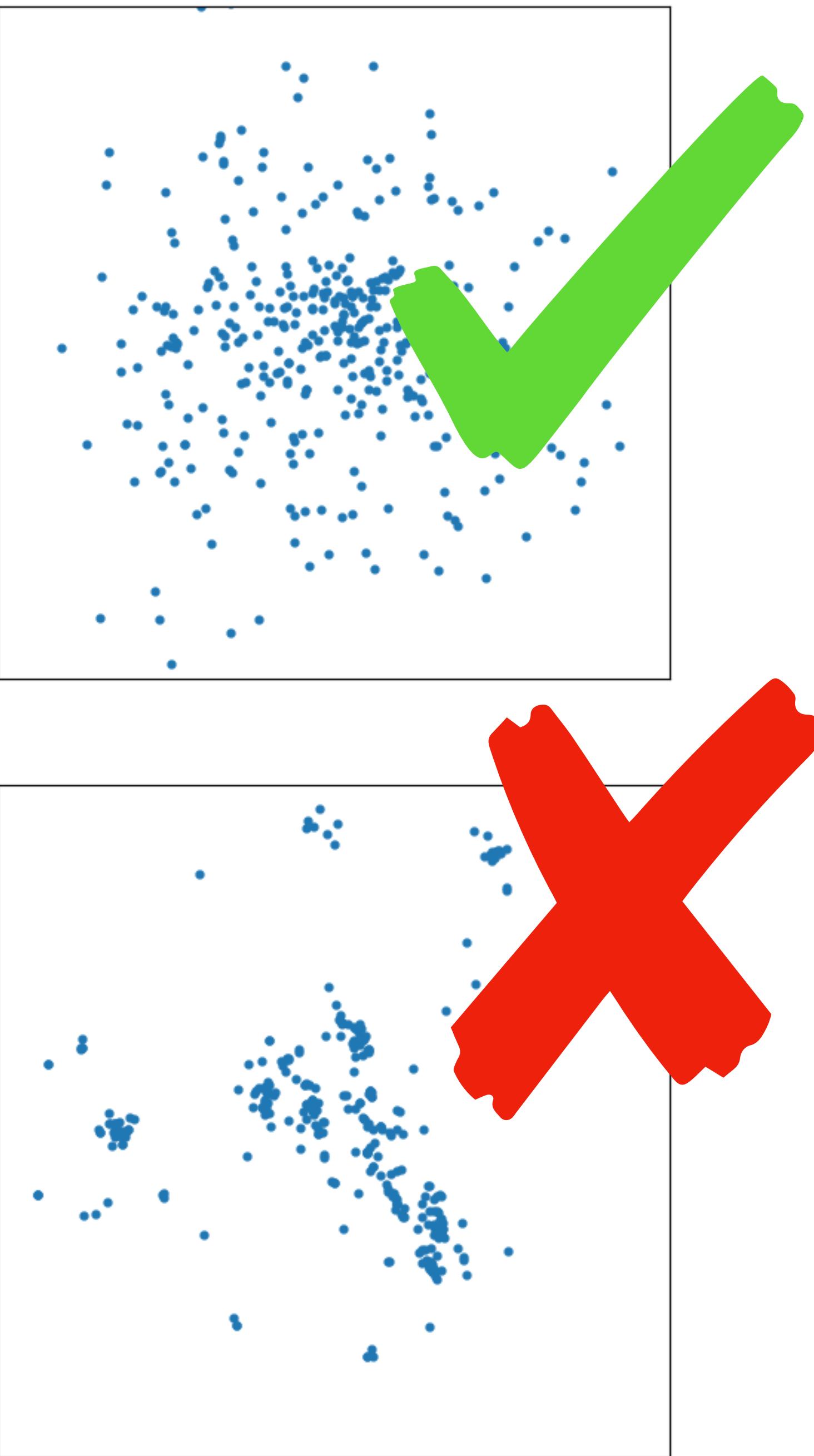
# $\bar{m}$ - $\bar{s}$ plots

| Cluster    | # | N   |
|------------|---|-----|
| Lupus 3    | 1 | 67  |
| IC 348     | 2 | 350 |
| $\rho$ Oph | 3 | 198 |
| IC 2391    | 4 | 200 |
| Cha I      | 5 | 234 |
| Taurus     | 6 | 335 |



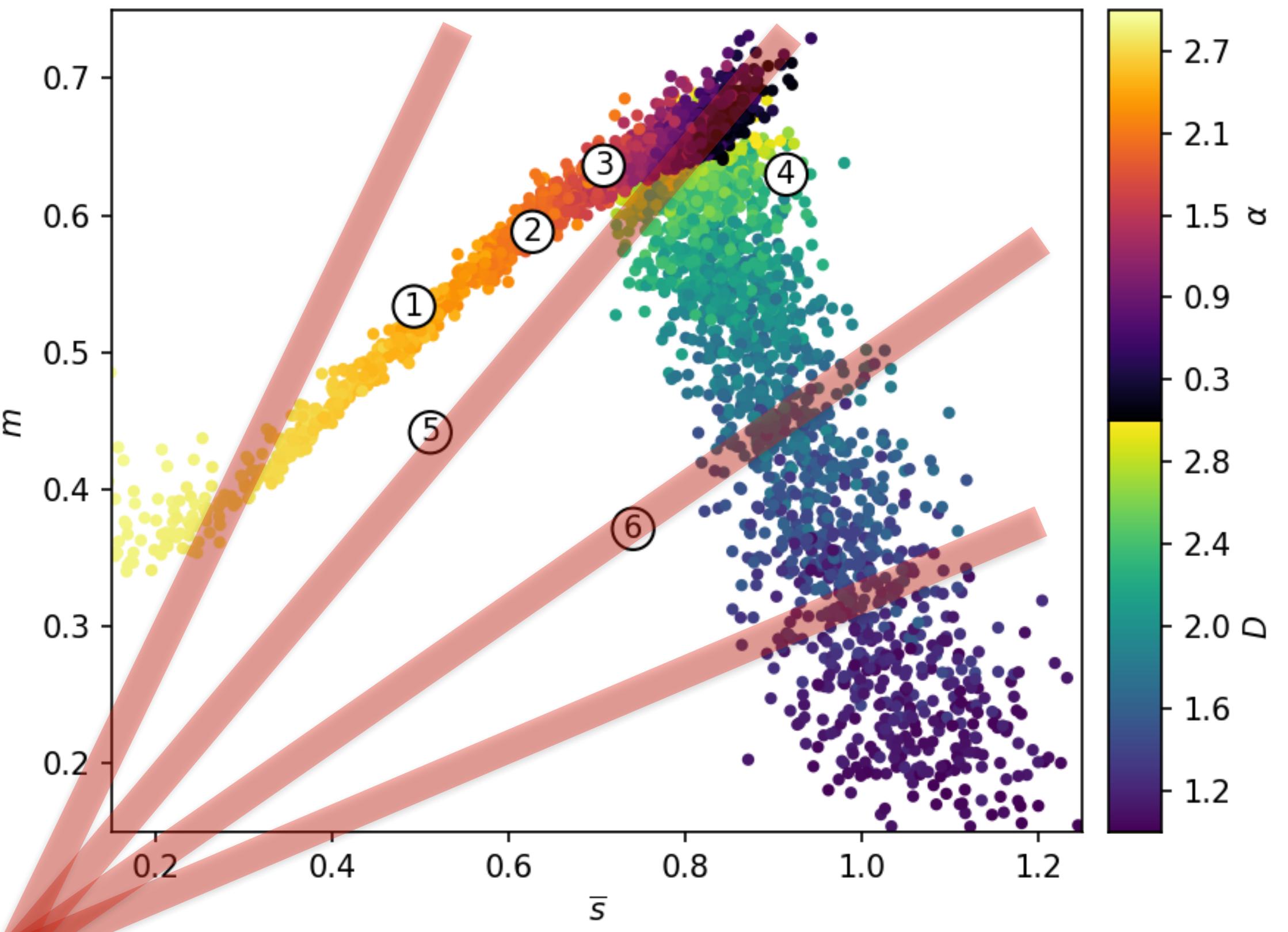
Comerón (2008)  
Lada et al. (2006); Muench et al. (2007)  
Bontemps et al. (2001)  
Barrado y Navascués et al. (2001)  
Luhman (2007)  
Luhman et al. (2010)

Cartwright, 2009, MNRAS, 400, 1427



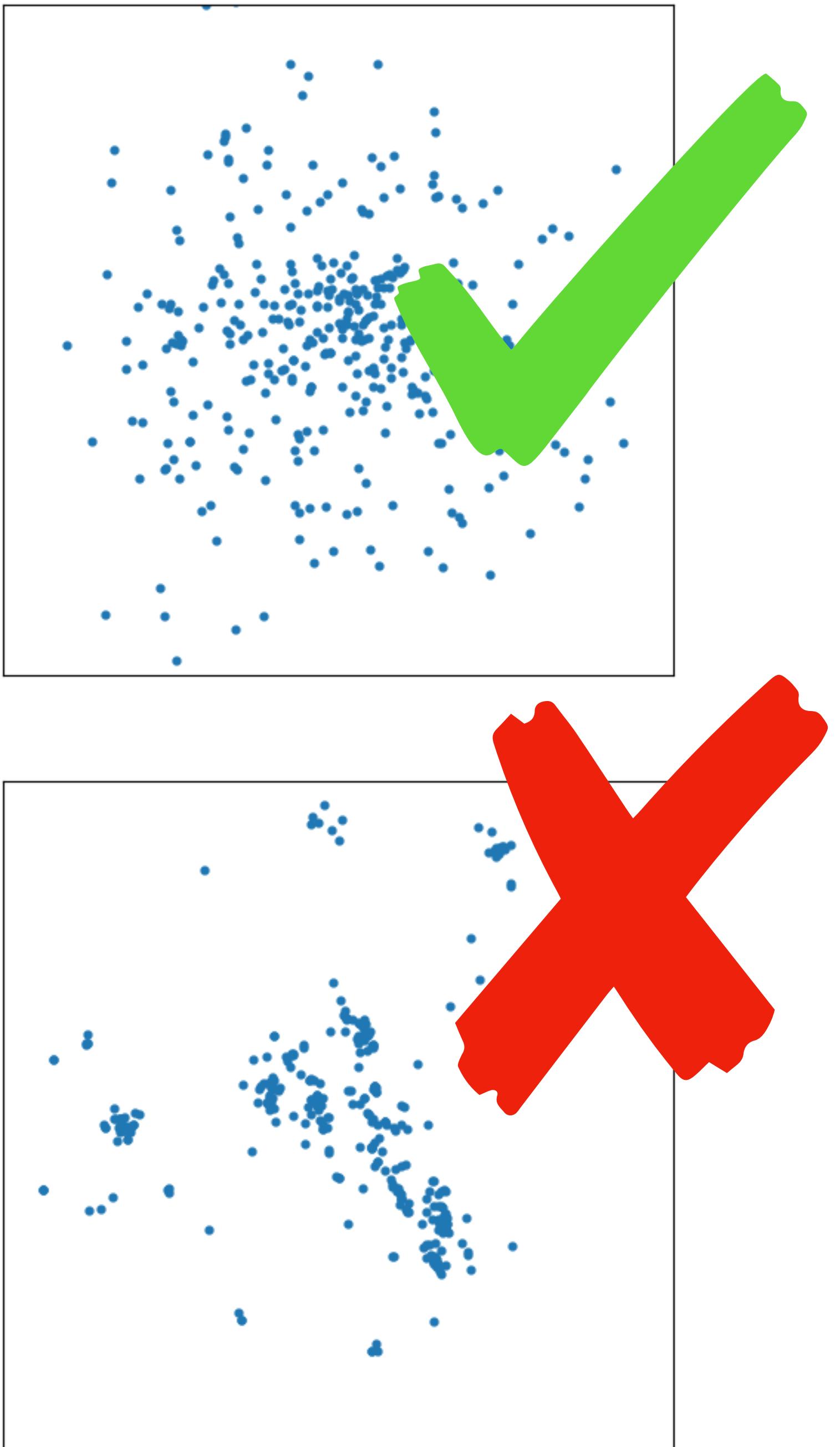
# $\bar{m}$ - $\bar{s}$ plots

| Cluster    | # | N   |
|------------|---|-----|
| Lupus 3    | 1 | 67  |
| IC 348     | 2 | 350 |
| $\rho$ Oph | 3 | 198 |
| IC 2391    | 4 | 200 |
| Cha I      | 5 | 234 |
| Taurus     | 6 | 335 |



Comerón (2008)  
 Lada et al. (2006); Muench et al. (2007)  
 Bontemps et al. (2001)  
 Barrado y Navascués et al. (2001)  
 Luhman (2007)  
 Luhman et al. (2010)

Cartwright, 2009, MNRAS, 400, 1427



# Stop using “square” models...



# Stop using “square” models...

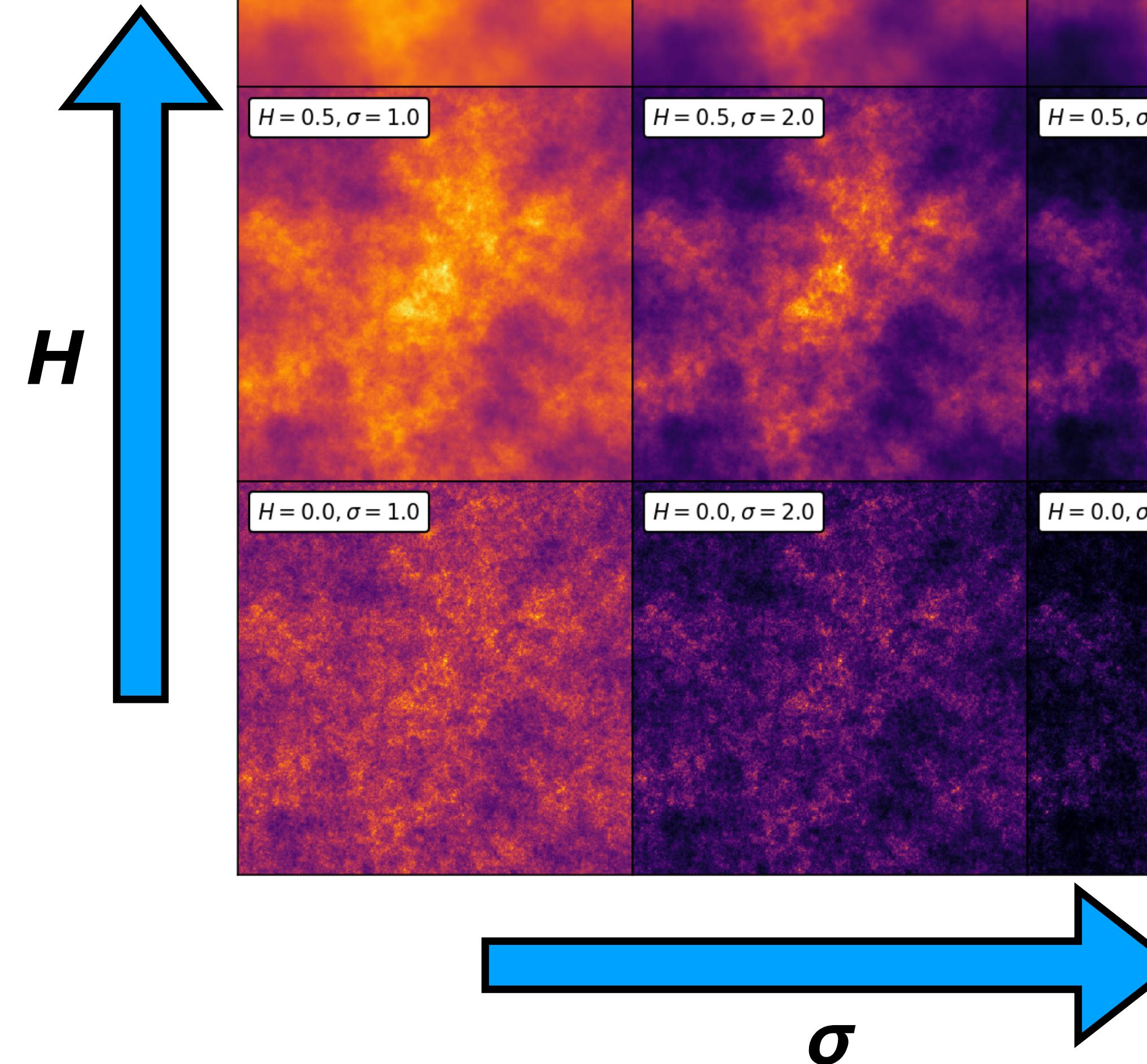


# Fractional Brownian Motion

- A generalisation of classical Brownian motion.
- Used to interpret properties of molecular clouds (e.g. Stutzki et al., 1998, Elia et al., 2014).
- Parameterised by the “Hurst index”  $H$  and a standard deviation  $\sigma$ .
- Relatively simple to generate in Fourier space:

$$P(k) \propto k^{-\beta}$$

$$\beta = E + 2H$$

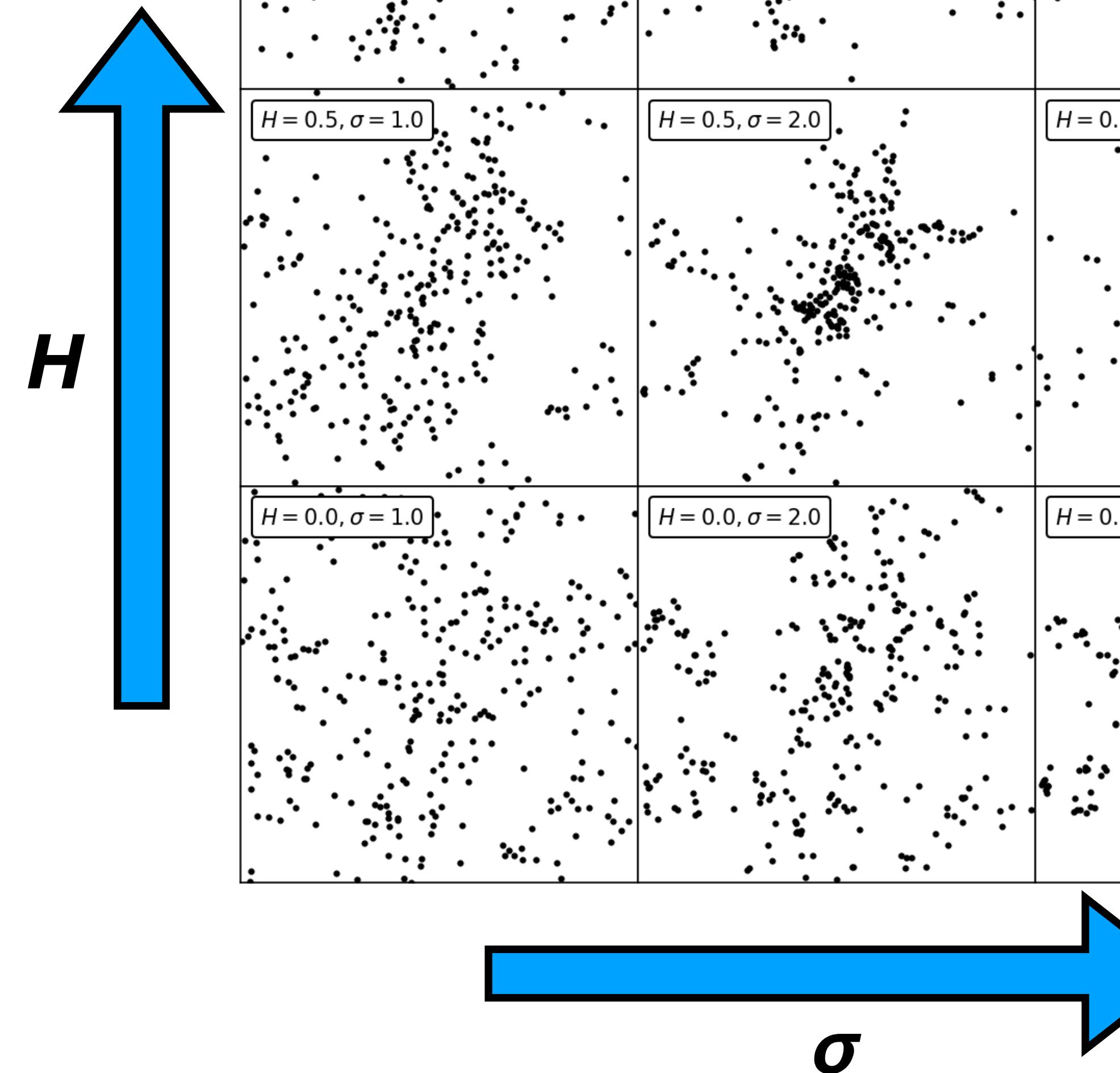


# Fractional Brownian Motion

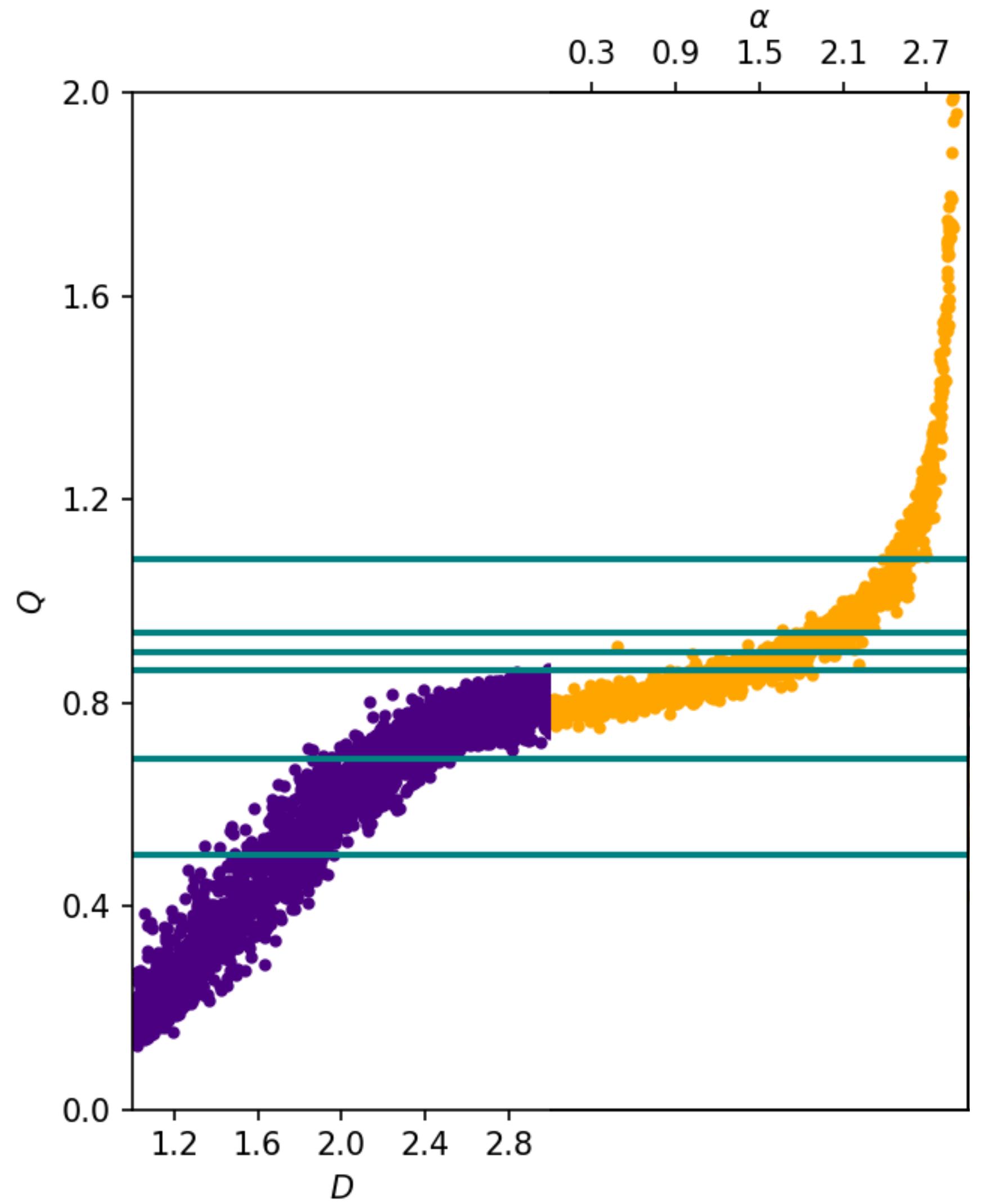
- A generalisation of classical Brownian motion.
- Used to interpret properties of molecular clouds (e.g. Stutzki et al., 1998, Elia et al., 2014).
- Parameterised by the “Hurst index”  $H$  and a standard deviation  $\sigma$ .
- Relatively simple to generate in Fourier space:

$$P(k) \propto k^{-\beta}$$

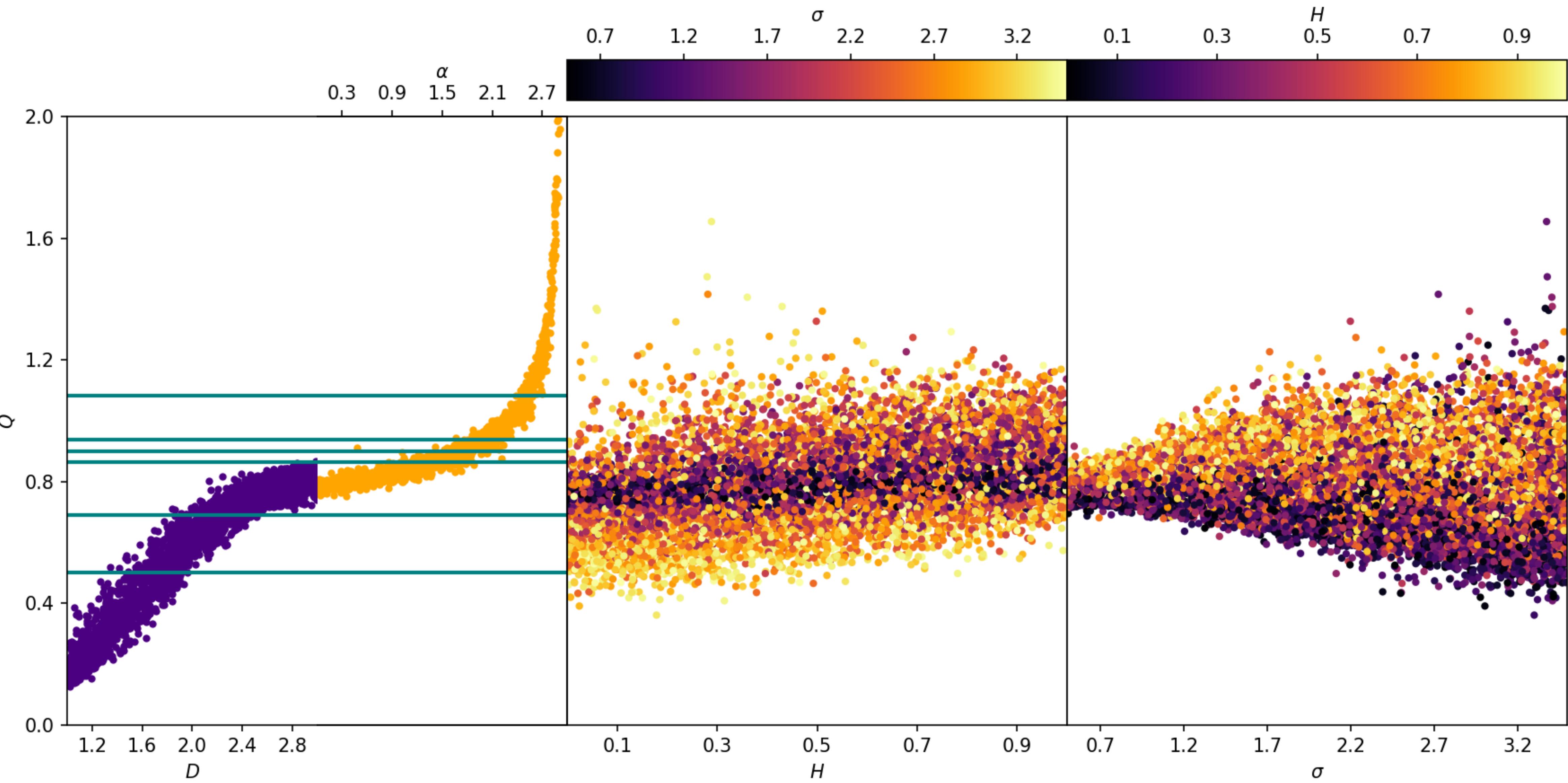
$$\beta = E + 2H$$



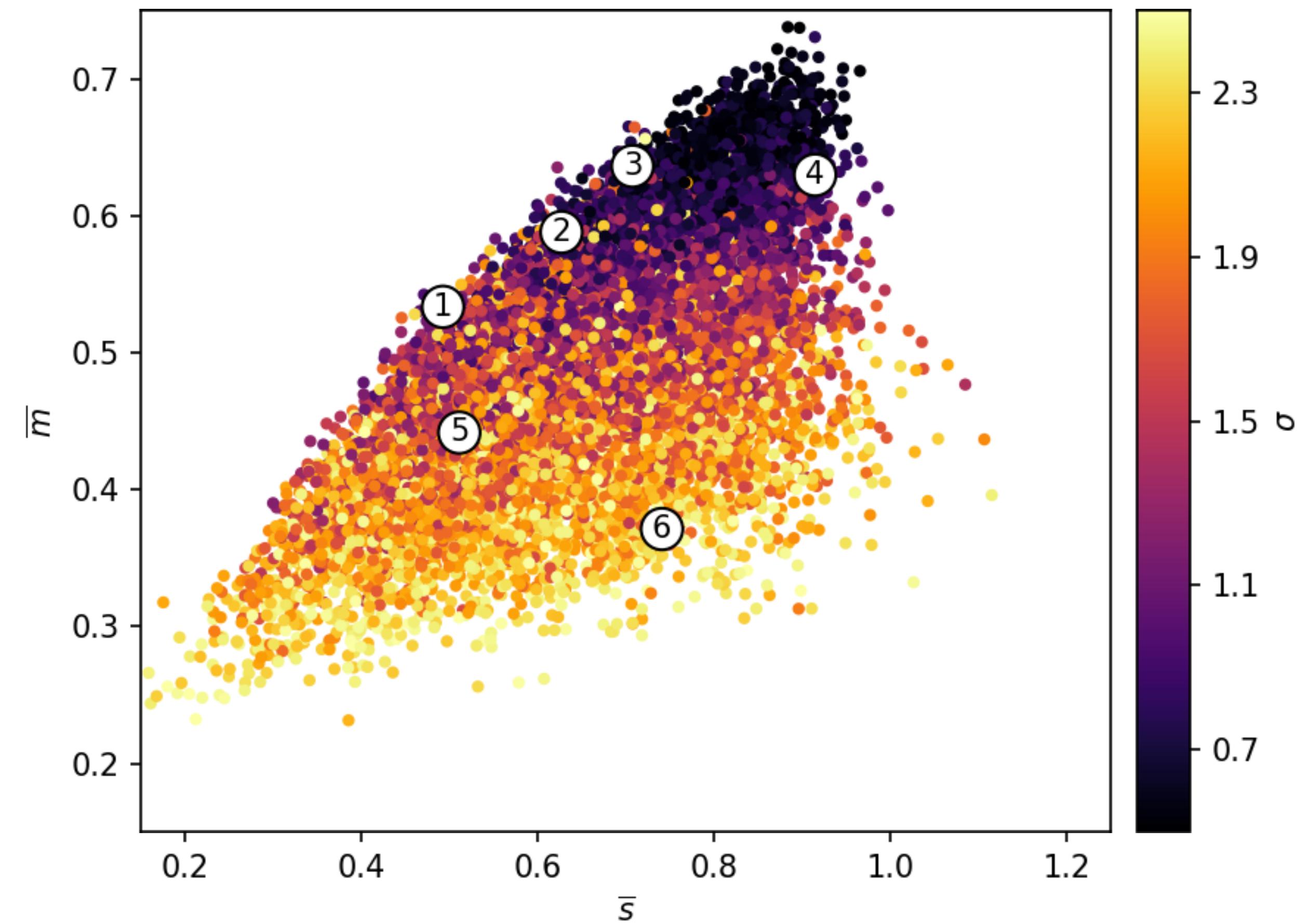
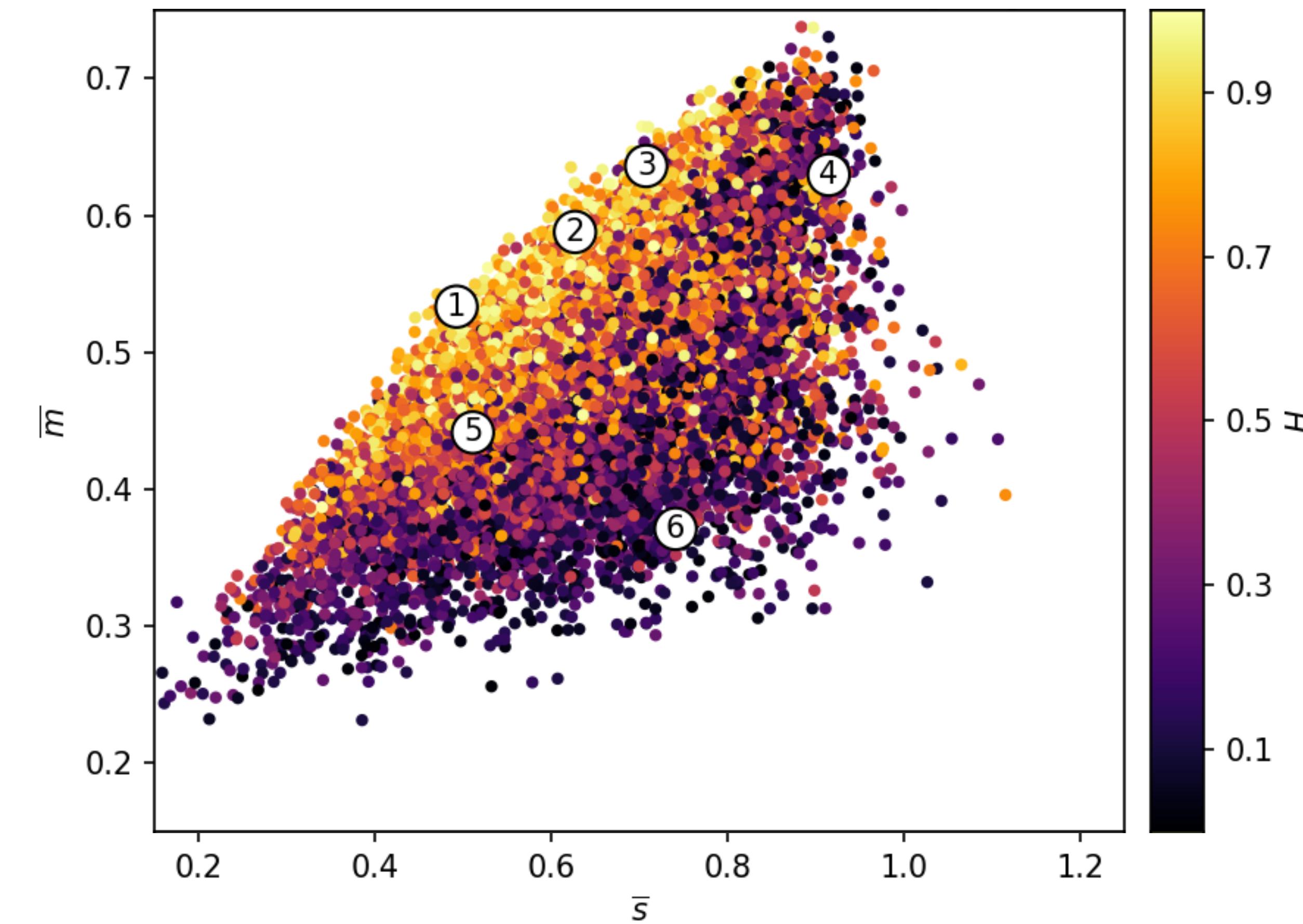
# Q parameter



# Q parameter is rubbish



# $\bar{m}$ - $\bar{s}$ plots

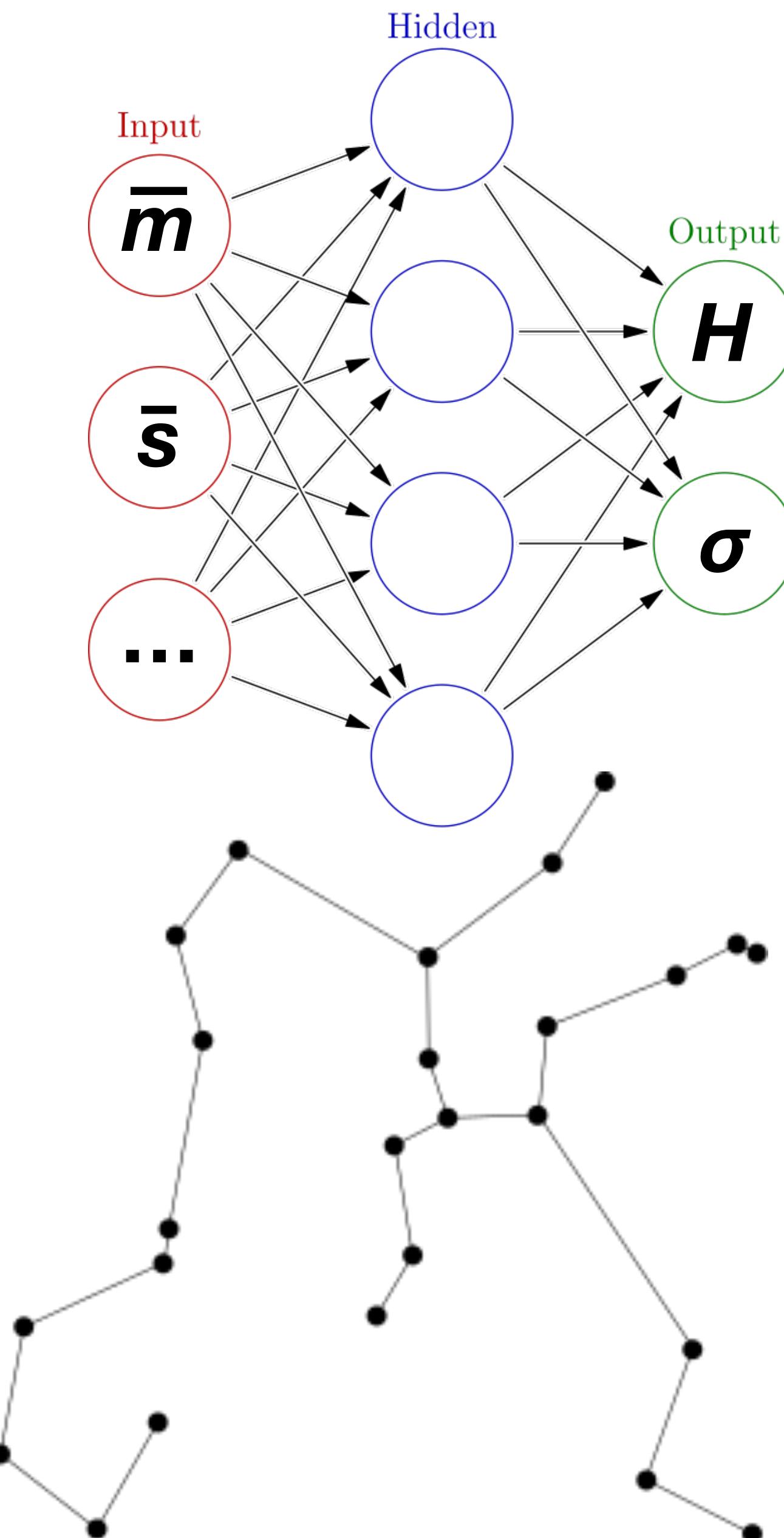


# Machine learning

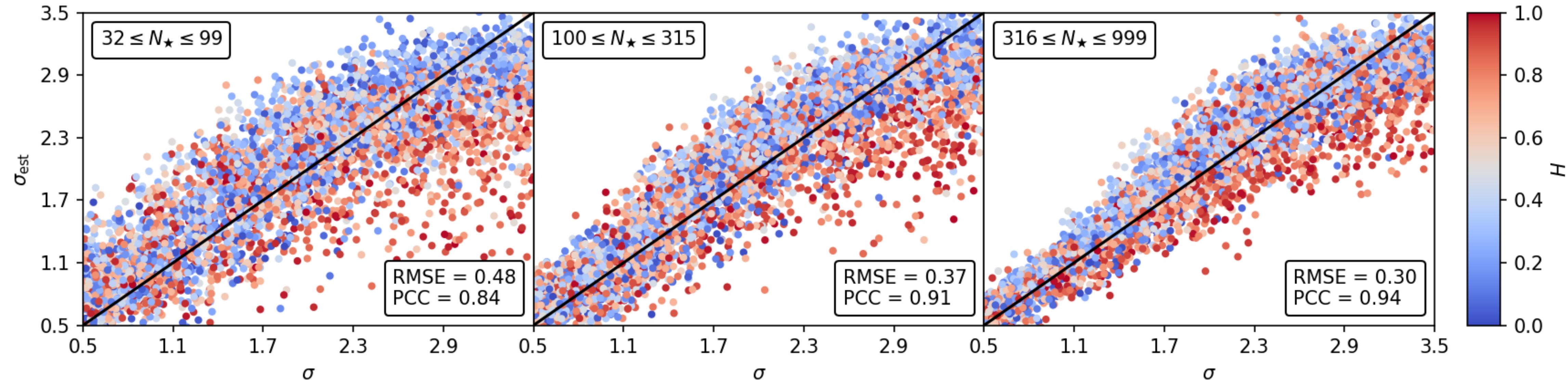
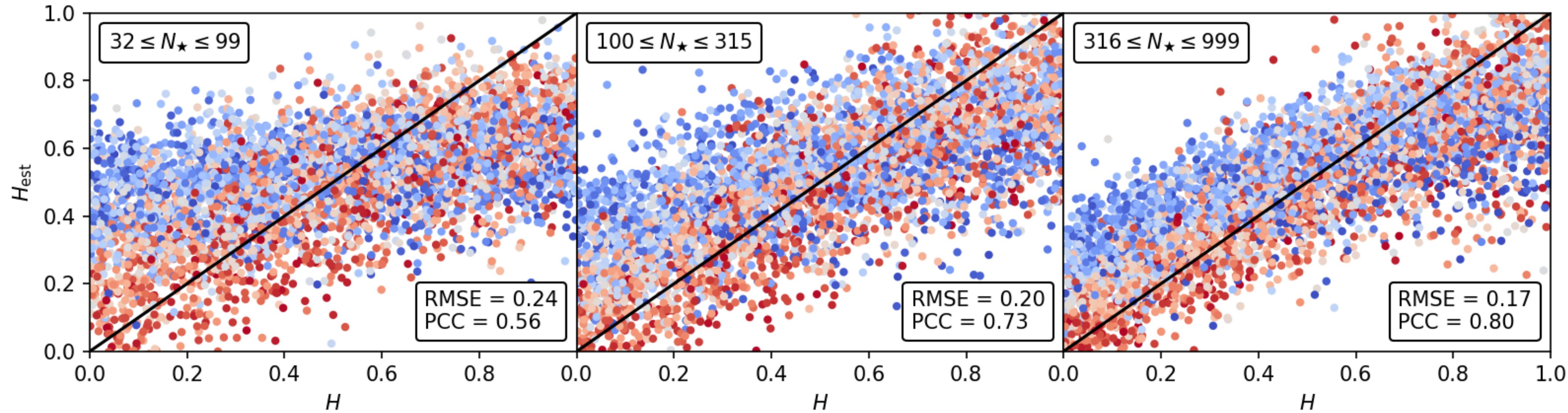
- Cluster distribution is “whitened”, i.e.  $\text{Cov}(\mathbf{x})=\mathbf{I}$
- Parameters estimated using minimum spanning tree and complete graph.
- Neural net method used to map edge length moments of graphs to underlying parameters.

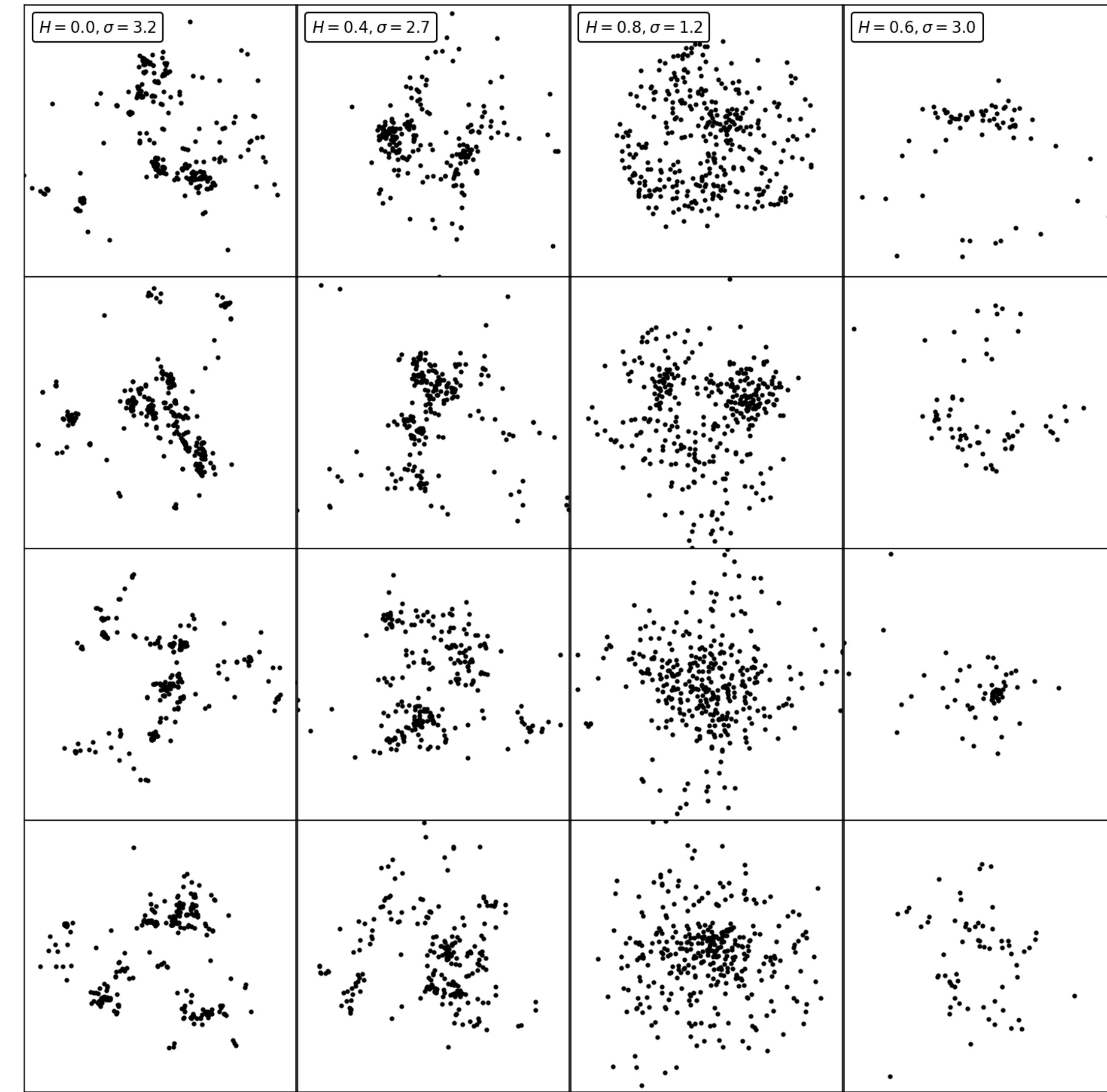
$$M_n(\mathbf{x}) = \left( \frac{1}{N} \sum_{i=1}^N (x_i - \mu_{\mathbf{x}})^n \right)^{1/n}$$

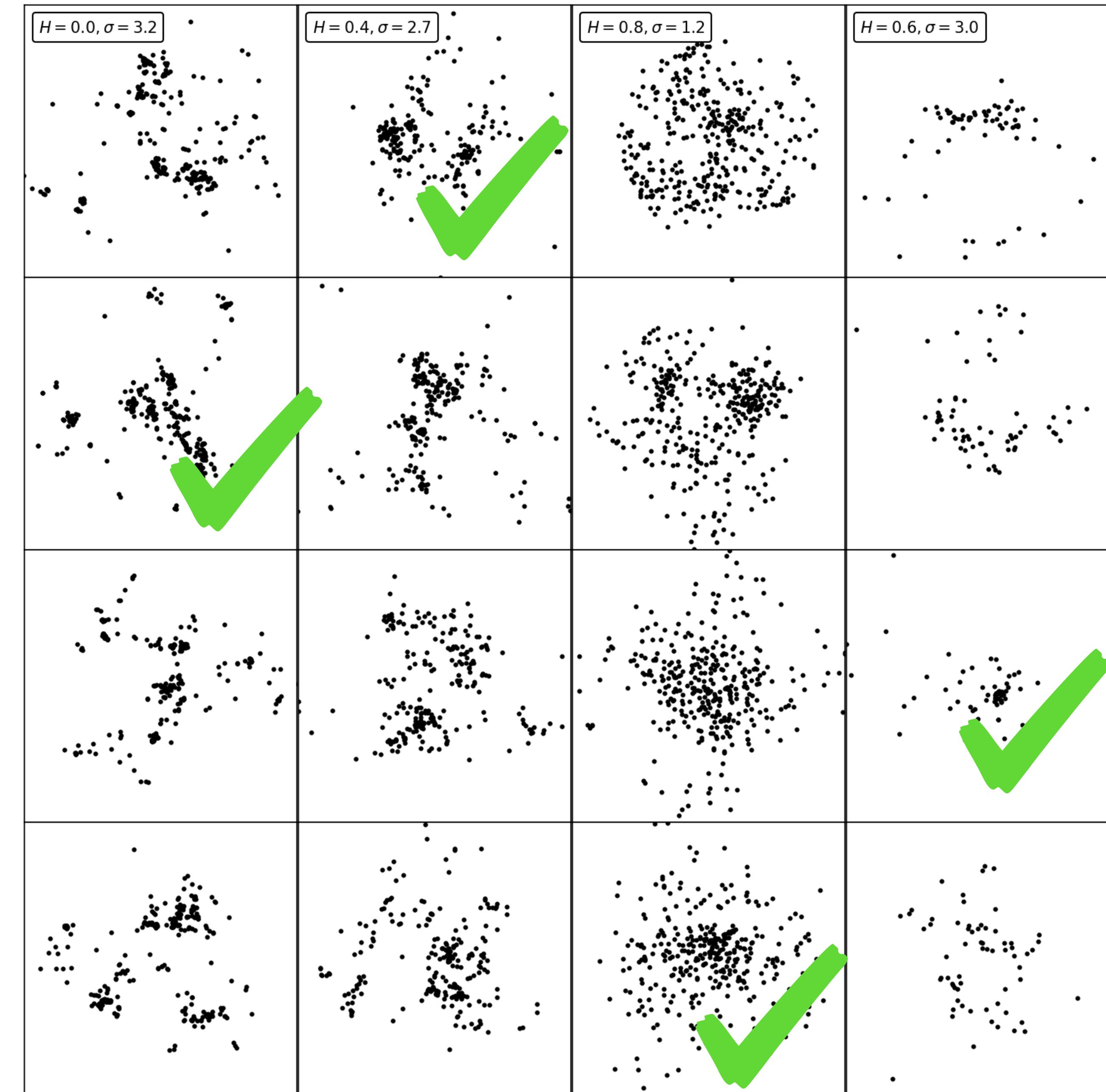
GitHub: [github.com/odlomax/clusterfrac](https://github.com/odlomax/clusterfrac)



# Parameter Estimation

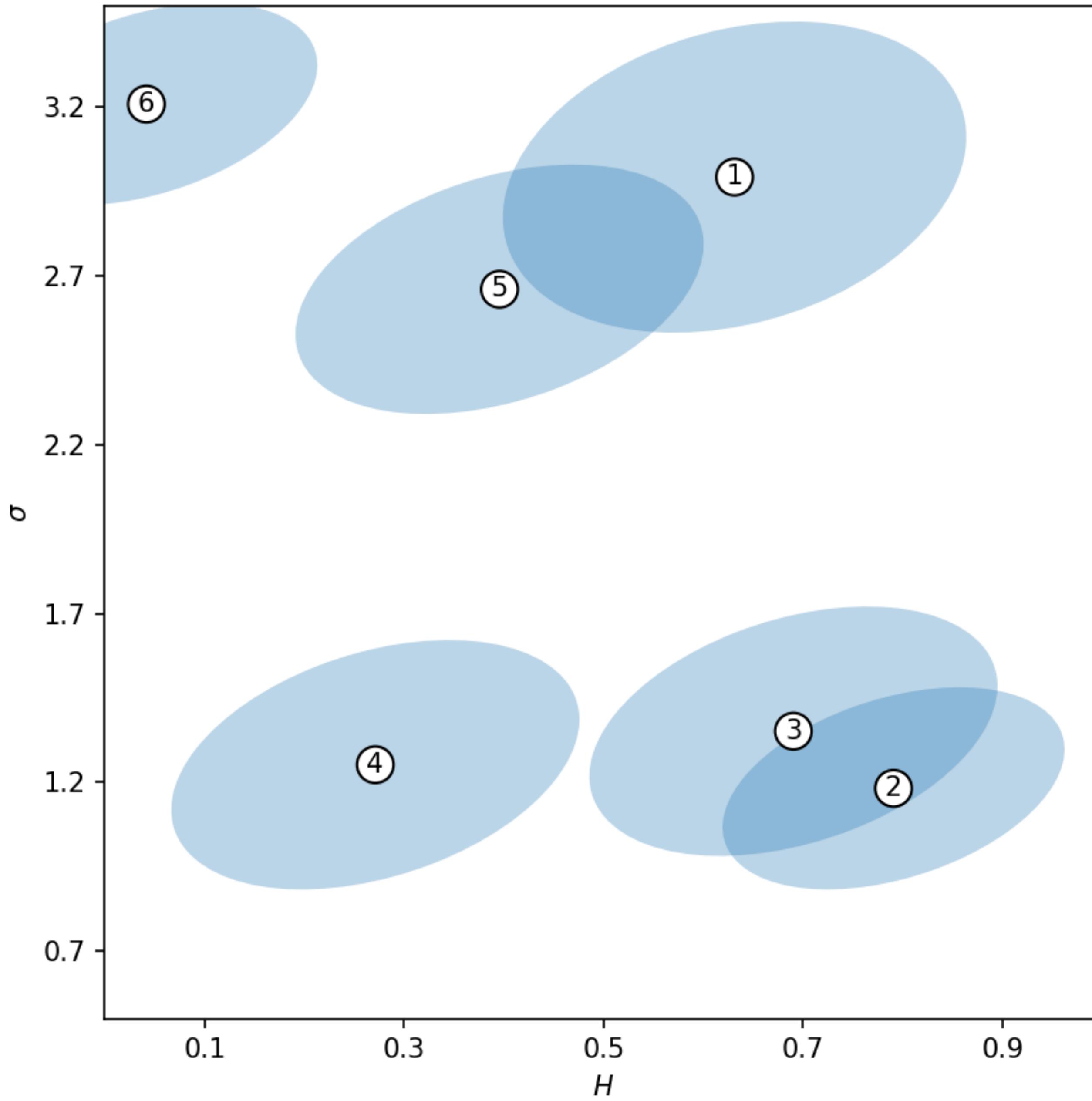






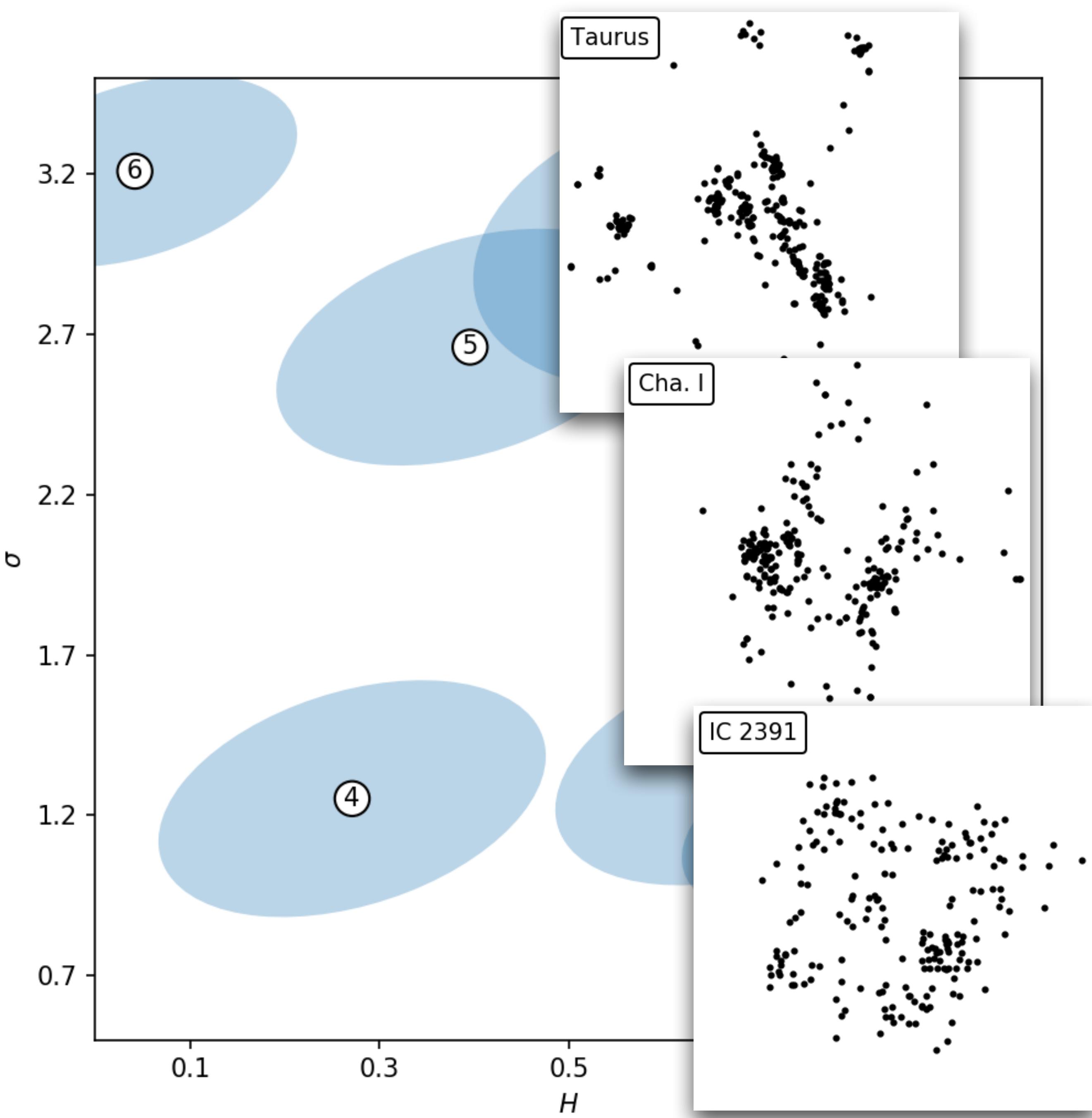
| Cluster    | # | <i>N</i> |
|------------|---|----------|
| Lupus 3    | 1 | 67       |
| IC 348     | 2 | 350      |
| $\rho$ Oph | 3 | 198      |
| IC 2391    | 4 | 200      |
| Cha I      | 5 | 234      |
| Taurus     | 6 | 335      |

Comerón (2008)  
 Lada et al. (2006); Muench et al. (2007)  
 Bontemps et al. (2001)  
 Barrado y Navascués et al. (2001)  
 Luhman (2007)  
 Luhman et al. (2010)



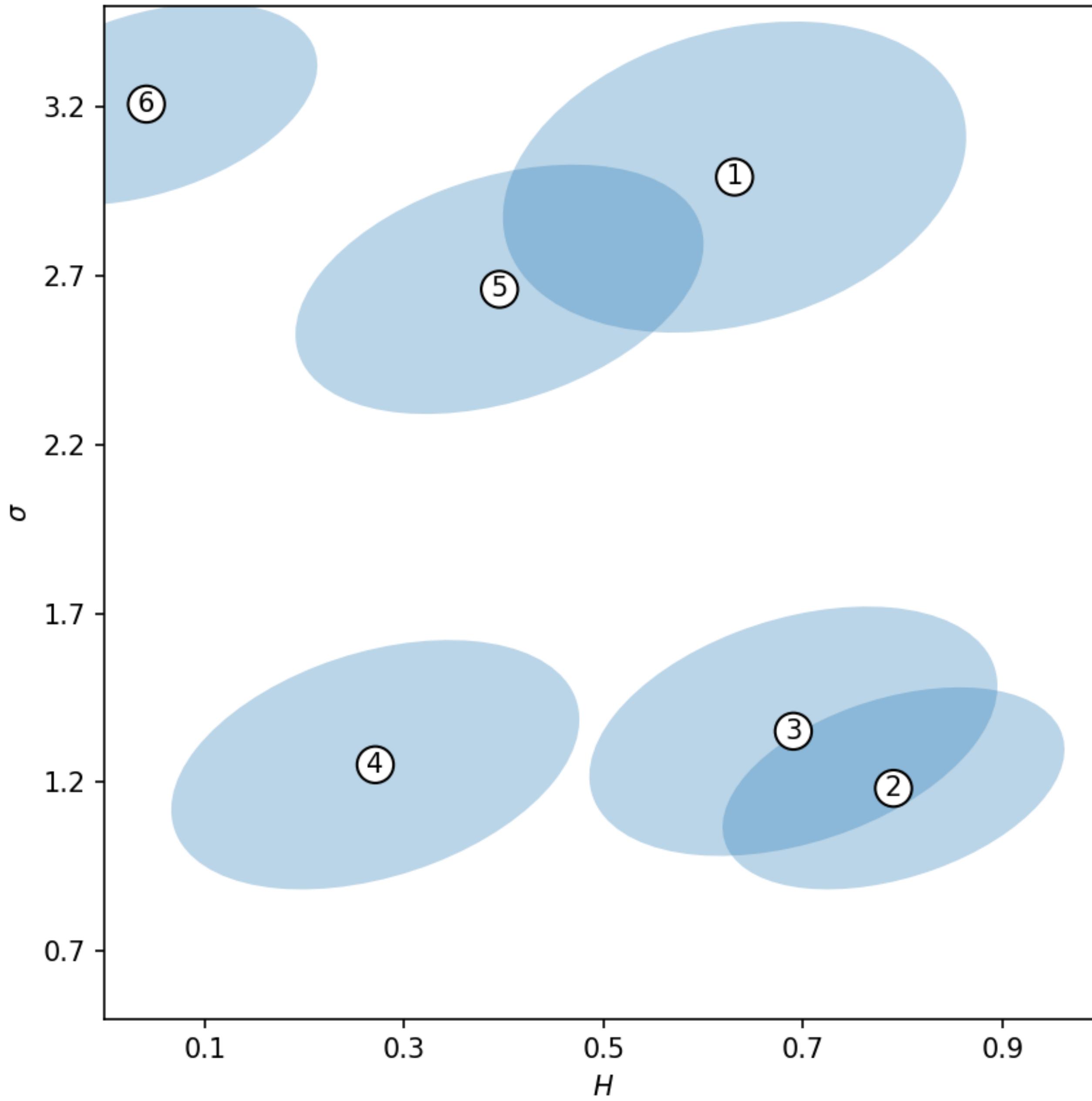
| Cluster    | # | <i>N</i> |
|------------|---|----------|
| Lupus 3    | 1 | 67       |
| IC 348     | 2 | 350      |
| $\rho$ Oph | 3 | 198      |
| IC 2391    | 4 | 200      |
| Cha I      | 5 | 234      |
| Taurus     | 6 | 335      |

Comerón (2008)  
 Lada et al. (2006); Muench et al. (2007)  
 Bontemps et al. (2001)  
 Barrado y Navascués et al. (2001)  
 Luhman (2007)  
 Luhman et al. (2010)



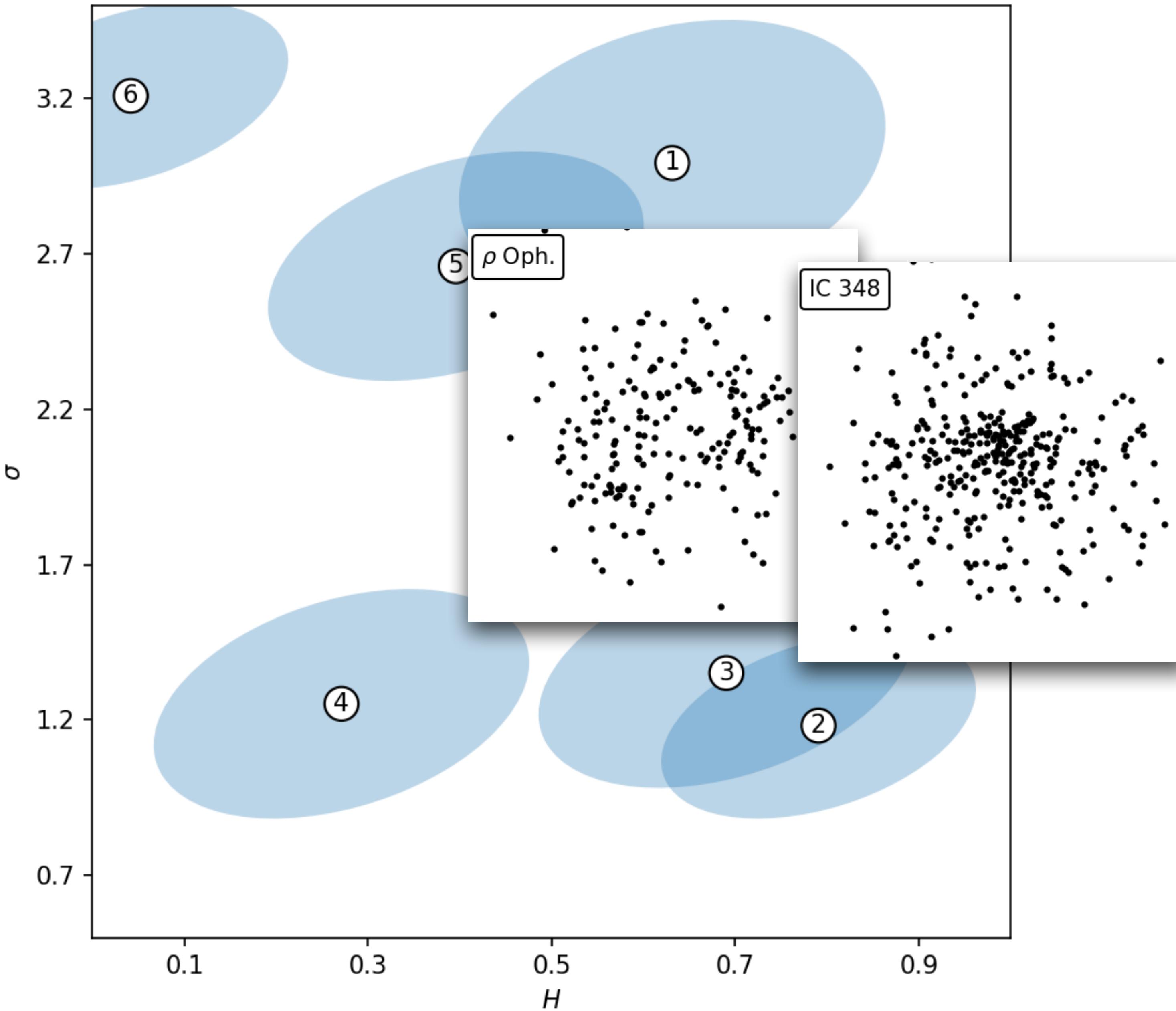
| Cluster    | # | <i>N</i> |
|------------|---|----------|
| Lupus 3    | 1 | 67       |
| IC 348     | 2 | 350      |
| $\rho$ Oph | 3 | 198      |
| IC 2391    | 4 | 200      |
| Cha I      | 5 | 234      |
| Taurus     | 6 | 335      |

Comerón (2008)  
 Lada et al. (2006); Muench et al. (2007)  
 Bontemps et al. (2001)  
 Barrado y Navascués et al. (2001)  
 Luhman (2007)  
 Luhman et al. (2010)



| Cluster    | # | <i>N</i> |
|------------|---|----------|
| Lupus 3    | 1 | 67       |
| IC 348     | 2 | 350      |
| $\rho$ Oph | 3 | 198      |
| IC 2391    | 4 | 200      |
| Cha I      | 5 | 234      |
| Taurus     | 6 | 335      |

Comerón (2008)  
 Lada et al. (2006); Muench et al. (2007)  
 Bontemps et al. (2001)  
 Barrado y Navascués et al. (2001)  
 Luhman (2007)  
 Luhman et al. (2010)



# Conclusions

- Q analysis of stellar structures has some serious shortcomings, mainly due to models.
- Fractional Brownian motion (FBM) provides a convenient model of star cluster morphology, related to molecular clouds.
- Machine learning techniques can be used to modernise Q analysis.
- So far, we can classify clusters, and generate synthetic analogues. Simulations needed to better understand physical meaning of parameters.

**Article:** [Lomax et al., 2018, MNRAS, 480, 371](#)

**arXiv:** [1804.06844](#)

**GitHub:** [github.com/odlomax/clusterfrac](#)