Network Analysis of Stars and Galaxies

Francesco Xotta and Francesco Montagno Bozzone

Alma Mater Studiorum - University of Bologna
francesco.xotta@studio.unibo.it francesco.montagno@studio.unibo.it

October 14, 2022

Abstract

The aim of this paper is to apply complex network techniques for the study of galaxy and star distributions. Starting from a sample catalog (the HYG catalog), we build the correspondent network from the Adjacency matrix and measure some network centralities: Betweennes Centrality (BC), Closeness Centrality (CL), Degree Centrality (DC) and Page Rank Centrality (PC). We define several star and galaxy populations basing on the value of the centralities measure and then try to argue links and correlations with some astrphysical observables such as magnitude. We also study the topology of the network, illustrating its properties and main feature.

I. Introduction

During last decades it has been established that the evolution of a galaxy is strongly correlated with its environment. Many numerical simulations have shown that the Universe has some topological properties whose origin is related to the matter distribution. Then, many studies have tried to relate astrophysical observables with measures of environmental density [Hong and Dey, 2015]; however, due to the previous topological properties, density is not the only important parameter, but other local geometrical effect could play a primary role. Among all these methods, the measure of 2-point correlation function has been widely used since 1980 [Peebles, P. J. E. (1980)]: this gives a good measure of the clustering coefficient of a population. Unfortunately, high order corrections require a huge number of samples and great computing power. In this paper we try to use a new method which highlight the topological properties of the cosmic structure using Network Science. During last decades it has been applied in many areas, from social networks to biology. Taking advantage of this huge background, we apply some measures to a database of observed stars trying to correlate Network Sciences parameters with astrophysical observables and trying to understand the role played by the topology in a galaxy.

The paper is structured as follows: in section II, we introduce some Network definitions and establish the terminology used along the paper. In this section we define the adjacency matrix and the topological measures used for the analysis.

In section III we show how to construct a network from an astrophysical dataset and then we visualize the network by using a technique called Multidimensional Scaling in section IV (an explanation on the dataset used in this work can be found in section VI).

We then compare our network to a random network in section VII, obtaining that our network exhibits a non random behaviour that is shown more explicitly in section VIII and then compared to an astrophysical quantity (in this case the magnitude) in section i.

We conclude the paper by illustrating our results and suggesting new possibilities for this area of research.

II. Complex Network: properties and definitions

A network is defined as a set of vertices and links. An useful representation of a network is given by the Adjacency Matrix.
It is defined as follows:

$$A_{ij} = \begin{cases} 1 \text{ if there is a link between i and j} \\ 0 \text{ if there is no link between i and j} \end{cases}. \tag{1}$$

We are interested in undirected networks, where $A_{ij} = A_{ji}$ and $A_{ii} = 0$.
The topology of a network can be studied through the concept of centralities: we will use the degree centrality, given as

$$\sum_j A_{ij} = k_i \tag{2}$$

which represents the amount of neighbors for the i-th node.
The betweenness centrality measures the amount of shortest path (geodesics) passing through a certain vertex

$$x_i = \sum_{jk} \frac{n^i_{jk}}{g_{jk}} \tag{3}$$

where n^i_{jk} is the number of shortest paths passing through the node i, connecting j and k nodes while g_{jk} is the total number of shortest paths connecting j and k nodes.
The closeness centrality is a measure of topological center defined as the inverse of the average of the shortest distance from a given vertex to all the others

$$C_i = \left(\frac{1}{N-1} \sum_{j \neq i} d_{ij} \right)^{-1} \tag{4}$$

where N is the total number of nodes and d_{ij} is the shortest path linking i and j.
For completeness, we have also used Page Rank Centrality; after defining briefly the Eigenvalue Centrality as

$$x_i = \frac{1}{\lambda} \sum_{j \in V} a_{ij} x_j \tag{5}$$

where G=(V,E) is the graph having $A = (a_{ij})$ as its adjacency matrix (with $a_{ij} = 1$ if vertex i is linked to vertex j, otherwise $a_{ij} = 0$); we can rewrite it in vector notation as

$$\mathbf{A}\mathbf{x} = \lambda \mathbf{x} \tag{6}$$

The desired centrality measure will be, due to Perron-Frobenius Theorem, the eigenvector with the highest eigenvalue [M. E. J. Newman (2006)].
PageRank Centrality is a variation of the Eigenvalue Centrality given as:

$$PR[A] = \frac{1-d}{N} + d \left(\sum_{k=1}^{n} \frac{PR[P_k]}{C[P_k]} \right) \tag{7}$$

where d is the damping factor (decided by Google), PR[A] is the pagerank of node (which would be a page) A, n is the number of pages that are linked to A, N is the total number of pages, PR[P_k] are the pageranks of pages P_k and $C[P_k]$ is the total number of links contained in page P_k.

III. How to construct a network from astrophysical datas

The approach we have followed for the building of the network is to define a certain linking length (following physical intuition and trying to keep the network as sparse as possible [de Regt et al. (2018)]) so that if the physical distance between the node i and j is lower than the linking length we can assume that those nodes are linked together. This allows to construct the Adjacency Matrix and, then, to evaluate the centralites.
Given that the average stellar density in the Milky Way is approximately $0.14 \frac{stars}{pc^3}$ [Gregersen, Erik (2010)], we can compute the linking length by assuming all stars are distributed uniformally in space, so that, if we consider a cubical lattice containing a star on each vertex and we assume that stars of opposite vertexes must interact with each other, we estimate a linking length of approximately 4 pc.

IV. Visualization of the Network

Once the network has been built, it is possible to visualize it. Network visualization is important to identify patterns and structures inside the network itself. The easiest way to visualize it is to represent stars using their location in the space. We have chosen a reference frame where the Sun is placed at the origin. The result can be visualized in fig. 1; as expected, the Sun is at the center of the scatter plot and only one portion of the sky is analysed.
A more enlightening way to obtain informations from the visualization of the network is to use the Multi-Dimensional Scaling (MDS).
MDS is a technique that permits to visualize the level of similarity of an ensemble of points using some criteria. In particular, chosen a dimension N and a square matrix containing the similarity parameter for each couple of points (for example the distance between two nodes), an MDS-algorithm places the nodes into a N-dimensional Cartesian space, preserving the similarity parameter as well as possible [Mead, A (1992)].
In our case, the network is scaled in a three-dimensional space, using as square matrix the Adjacency matrix.
The result of the MDS can be visualized in fig. 2. However, further analysis suggest that this property is not an intrinsic one, but it is related with the chosen sampling for the analysis. In fact, if we select only a group of stars belonging to a cubic region, the final result is different while spirals structures seems to be related with the spherical surface of the considered volume: a random network of points generated over a spherical surface returns the same spiral structure.

It can be observed that the betweenness is high close to the origin and small far away from the centre of the MDS space. Moreover, the correlation coefficient between MDS-distance (which is the norm of the vector of cartesian coordinates in the MDS space) and degree centrality is $r = 0.75$, confirming the existence of a strong correlation between the parameters.
After investigating fig. 4 one can easily notice a logarithmic behaviour, which is confirmed upon fitting, observing an R-Squared value of 0.7125.

As stated above, the MDS structure obtained seems to be dependent on the data sampling; by generating points randomly on a small spherical shell, we obtain a MDS structure as shown in figure 5.

By taking only stars in a certain cubical volume we instead obtain instead figure 6.

V. Clusters of Stars

Due to the process of condensation of prestellar clouds stars are formed in clusters (the catalogue generally used today for this kind of object is the NGC [Dreyer, J.L.E. (1953)]); this means that clusters contain stars formed during the same time period which share the same chemical properties.
Stars are indeed divided into pop.1 (or metal rich stars, generally formed later with a starting metallicity, just like the Sun) and pop.2 (or metal poor, generally formed earlier; they are more frequent in globular clusters [van Albada, T. S.; Baker, N. (1973)]).
Clusters are divided in two main types: open and globular clusters. Open clusters contain from 10 to 1000 stars and have a global positive energy, which means that they are unbounded systems due to the rotation and gravitational perturbations of their galaxy. It is important to notice that open clusters are confined on the galaxy plane and can be found within spiral arms.
Globular Clusters instead are more rich of stars and can have up to 10^5 stars, distributed spherically with a density approximately 10 times the density of an open cluster; from the HR diagram we can see that those clusters contain mainly red stars of low masses with a rich red giant branch due to the stars with higher masses have already exited the Main Sequence (MS), becoming red giants (this indicates that the age of the cluster is very high); we will prove later that globular clusters have, for the aforementioned reason, lower magnitudes.

Young stars can also form associations with

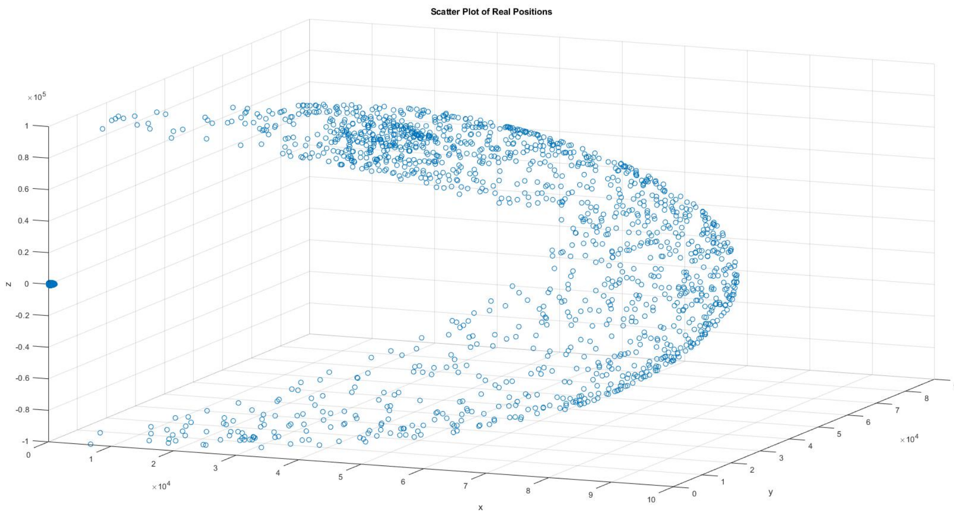

Figure 1: *Scatter plot of the network in a reference frame where the Sun is located at the origin.*

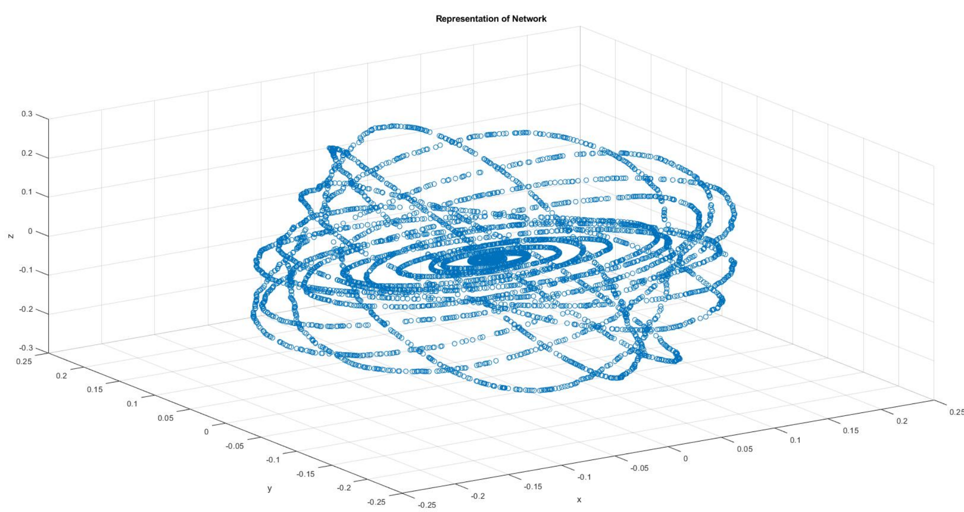

Figure 2: *Multi-dimensional scaling of the network using N=3 and the Adjacency matrix.*

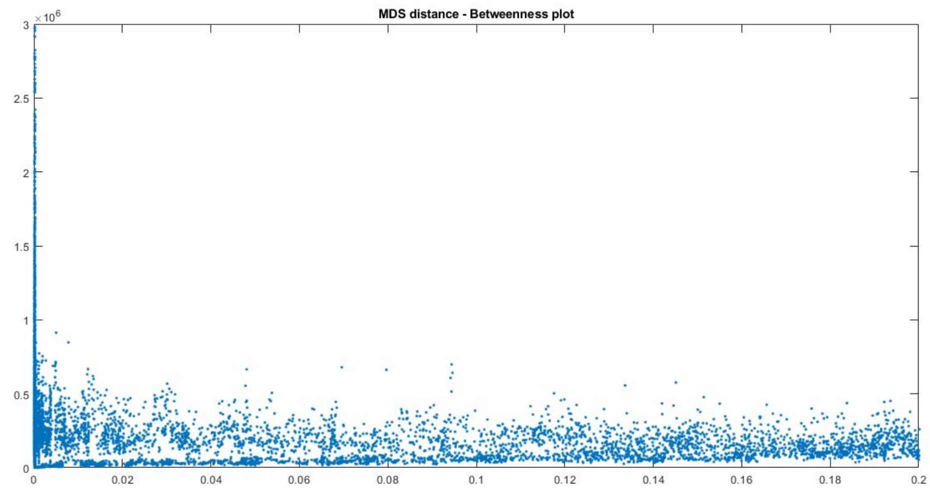

Figure 3: *On the x axis the distance from the centre of the MDS-origin; on the y axis the betweenness centrality .*

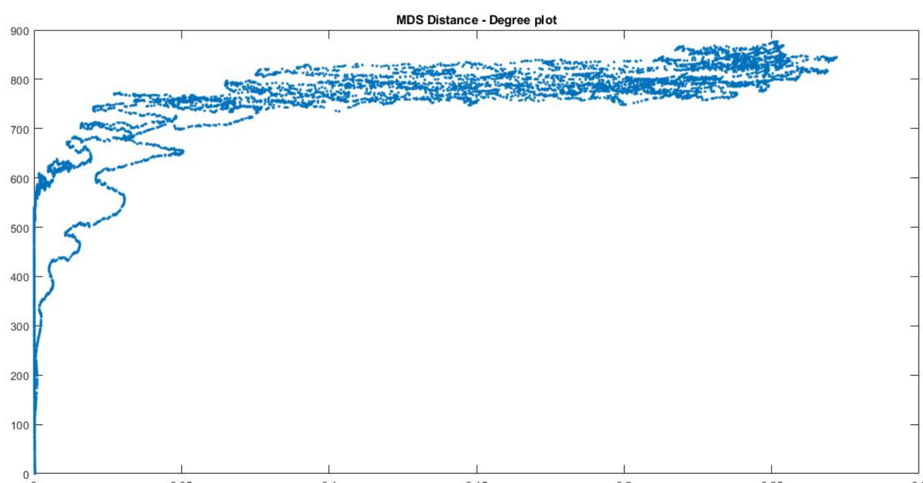

Figure 4: *On the x axis the distance from the centre of the MDS-origin; on the y axis the degree centrality.*

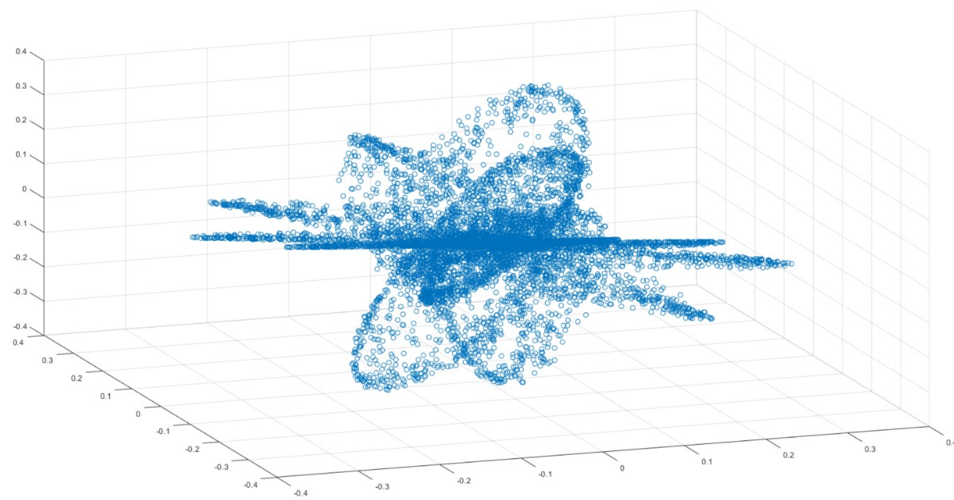

Figure 5: *Multidimensional scaling of a set of points randomly generated on a spherical shell*

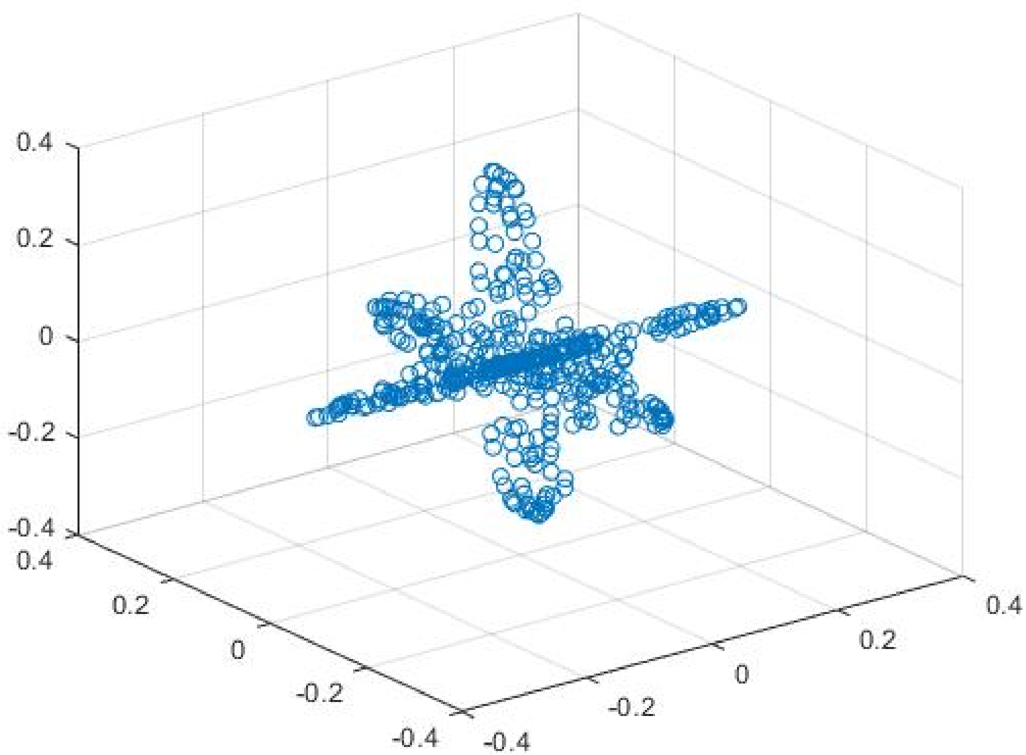

Figure 6: *Multidimensional Scaling of a selection of stars taken in a small cubical volume on the scatter plot.*

usually very few stars (around 10); associations can contain very bright stars and, just like open clusters, they have positive global energy and are thus unbounded.

VI. The HYG Catalog

In order to test the methods described this far we used the HYG catalogue, a database containing all stars in the Hipparcos, Yale Bright Stars and Gliese catalogs.

The Hipparcos Catalogue [Perryman+, 1997] is a materialisation of the ICRS reference system, containing 118218 stars.
The Yale Catalogue [Bahcall+, 1987] contains 9110 stars with their respective magnitudes, positions and spectral types.
The Gliese Catalogue [Gliese, 1957] contains 915 stars within 20 parsecs of Earth which was later extended in 1969 to 1529 stars [Gliese, 1969].

The HYG database contains the absolute magnitudes of almost 120000 stars which we are going to compare with the aforementioned centrality measures; it also contains the spectral type of the star when known, the color index and the proper name (the common name for the star), which we are not going to use in this work.

The HYG catalogue does not contain stars with a distance from the Sun below 160 light years as can be seen from figure 1.

VII. Is a stellar network random?

Once we have built the network as described in (III), it is possible to start our analysis. The first question to address is whether the network is random or not (see Appendix B for a brief description of random networks and definitions used in this section). Since the number of nodes n (i.e., the number of stars in the database) is known and the number of links m can be obtained counting the connections in the network, we can build a random network as follows: we extract randomly two nodes and we link them, repeating this process m times.
Then, we can verify that the obtained network is random plotting, for example, the frequency histogram of the degree k.
As we can see in figure 1, the behaviour is gaussian, as expected for the continuum limit (N large) of a binomial distribution (see Appendix B).

Moreover, we can also graph the betweenness as function of the degree. From figure 2, it can be observed a linear behaviour, as expected: increasing the degree in a random network we also increase the betweenness due to the random nature of the graph.

Finally, we can also graph the same results for our real network. For the first one (figure 3) no binomial distribution is observed, showing that the star network is not random. This result is also confirmed by the betweenness vs degree plot (figure 10), where no linear relation is observed

VIII. Results for the Stellar Network

As expected from the theory, the network presents a structure of different (and separated) clusters; as a difference with the expectation, we observe the presence of "bridges" (low degree centrality but high betweenness centrality) that will be discussed below more in detail.

As we have already seen in figure 9 from the distribution of degrees in the stellar network we can see that stars are more likely found in the high degree region, meaning they tend to be in clusters and groups (this can be related to how stars are formed and to the structures defined in section V); we can observe also a peak for very low degree centrality values given by the fact that, due to the low number of points used in this study (20000), lots of stars don't have all their neighbours included and are for this reason not part of any groups or cluster (while, as we have seen, this is physically not possible). It can also be noticed that the average Degree for our network is approximately 499 while for the random network it's 501: in reference to figure 9 it is obvious that the average is lowered by the high amount of stars that

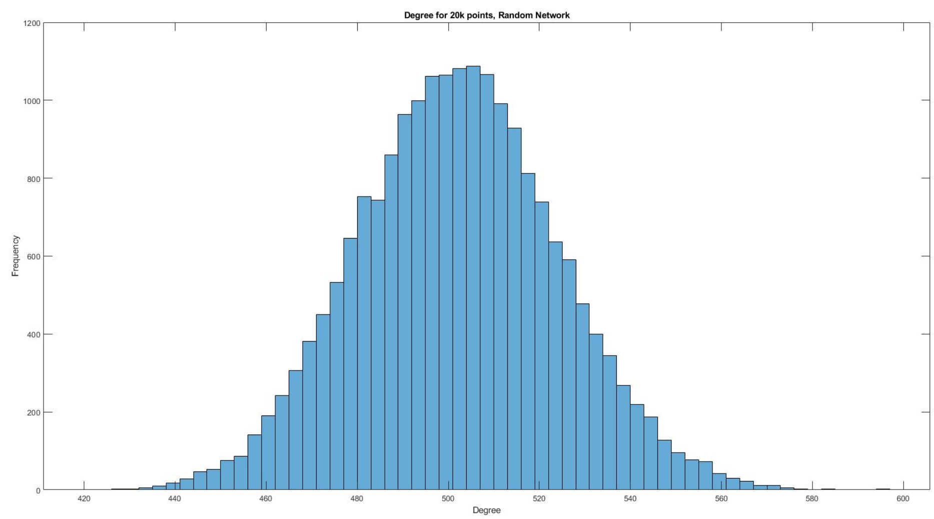

Figure 7: *Frequency histogram for the degree k. As expected, the behaviour is approximatively gaussian (for $N \to \infty$ the binomial distribution tends to a gaussian).*

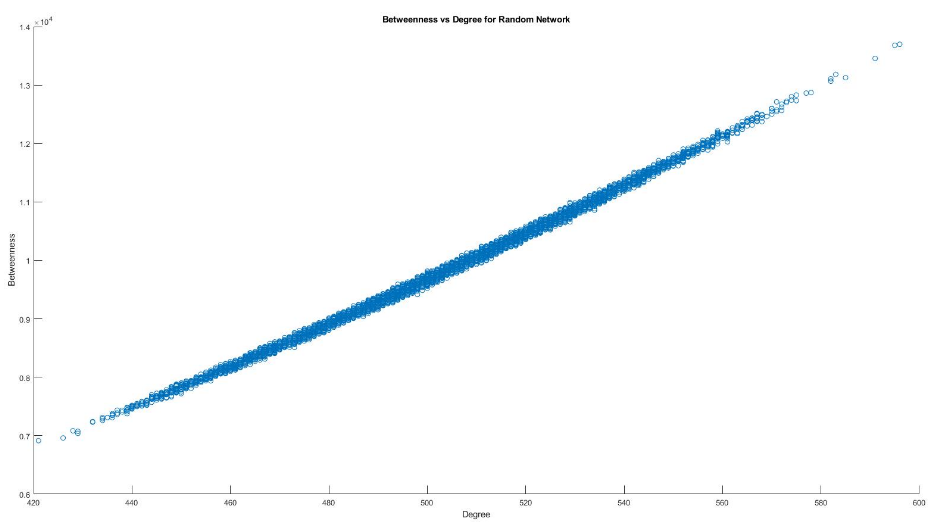

Figure 8: *Degree centrality vs Betweenness centrality for our random network. As expected, we can observe a linear relation.*

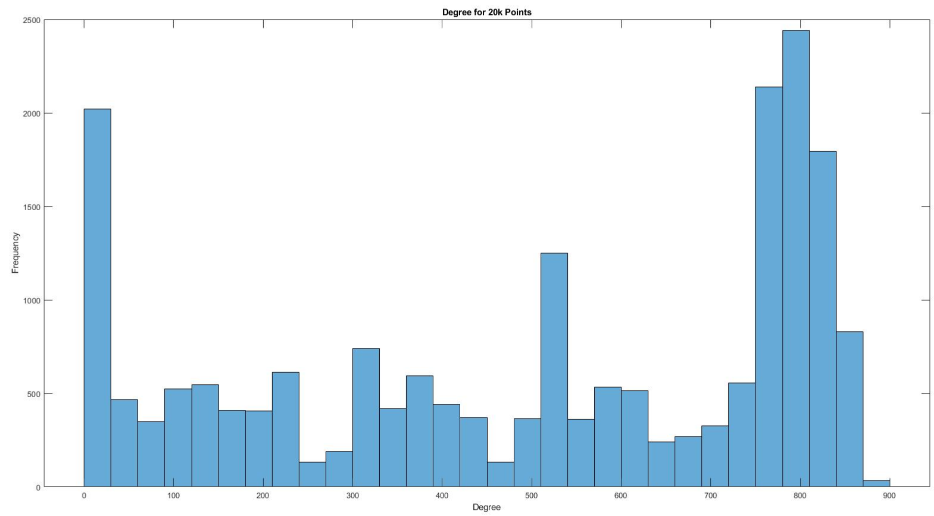

Figure 9: *Frequency histogram for the degree of the star network. The peak around the zero is a consequence of the not high enough number of nodes. The distribution does not follow the binomial one (as for the random graphs)*

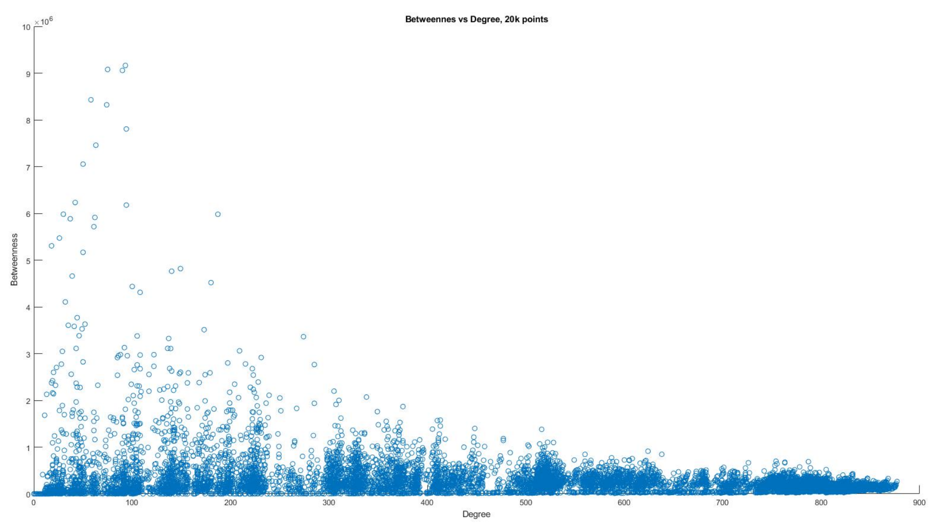

Figure 10: *The figure shows no linear relation between degree and betwennes, confirming again that the stellar network is not random.*

are incorrectly isolated due to the low number of points used in this analysis.

We can also see from figure 10 that generally the Betweenness Centrality of stars is constant, something that can be explained by remembering that stars form in groups from the same clouds and, for this reason, they generally don't have stars that belongs to different clusters at the same time. It can however be noted that there are some stars with very high Betweenness Centrality compared to the constant value and, at the same time, low Degree Centrality: this can be explained by reminding of the low sample of stars considered, which may have caused the division of clusters that would have been otherwise united leading to some of the stars in said clusters becoming part of two different clusters and having for this reason a very high Betweenness Centrality while keeping a very low Degree Centrality.

We can then look at the Closeness Centrality in 11 which we plot observing a behaviour that is, again, different than the expected random gaussian; closeness centrality tends to recall Degree Centrality, showing the same behaviour meaning that the most central nodes are the one with more neighbours.

We are however more interested now in testing another property, called assortativity (also said "Assortative Mixing") which is a preference for nodes to connect to similar nodes.
We define now a quantity e_{ij} to characterize the mixing which will be the fraction of edges in a network that connect a vertex of type i to one of type j such that $\sum_{ij} e_{ij} = 1$, $\sum_i e_{ij} = b_j$, $\sum_j e_{ij} = a_i$ and, only for undirected networks, $e_{ij} = e_{ji}$.
It is then possible to quantify the level of assortative mixing in a network by defining an "assortative coefficient" given as

$$r = \frac{\sum_i e_{ii} - \sum_i a_i b_i}{1 - \sum_i a_i b_i} \tag{8}$$

which is, while physically different, mathematically similar to the Pearson correlation coefficients.[M. E. J. Newman (2003)]
Examples of assortative network are social networks, while E-R random networks are generally non assortative.
For our stellar network we can compute the Pearson correlation coefficient between the degree centrality of each node and the average degree centrality of every adjacent node, obtaining a value of r=0.9994 (a nearly perfect assortative pattern) which means every star will be adjacent mostly to stars which share a similar number of neighbours. This supports the idea that stars may form together in gravitational collapses of the same clouds, meaning that different clusters with different average degrees are separated from each others.
We can then plot the degree centrality of each node compared to the average degree centrality of every adjacent node, obtaining figure 12.

i. Magnitude and Centralities

We can now compare the absolute magnitude (illustrated more in detail in Appendix i) to the aforementioned centrality measures using mainly Pearson correlation coefficient, and we can see that:

- Betweenness Centrality and Absolute Magnitude are not related at all (correlation coefficient of 0.0510), something that is obvious considering that betweenness centrality is constant nearly everywhere
- Degree Centrality and Absolute Magnitude are moderately correlated (correlation coefficient of 0.6278), which can be explained by remembering that very bright stars in globular clusters (where stars have higher degree centralities and closeness centralities) have already evolved past the Main Sequence stage
- Closeness Centrality and Magnitude are also moderately correlated (correlation coefficient of 0.6996) for the same reason that we have just stated
- Pagerank Centrality and Magnitude are strongly correlated (correlation coefficient of 0.7589), which means that highly linked nodes are also less bright (having higher magnitudes). However, also this property -as the MDS Scaling- seems to be related

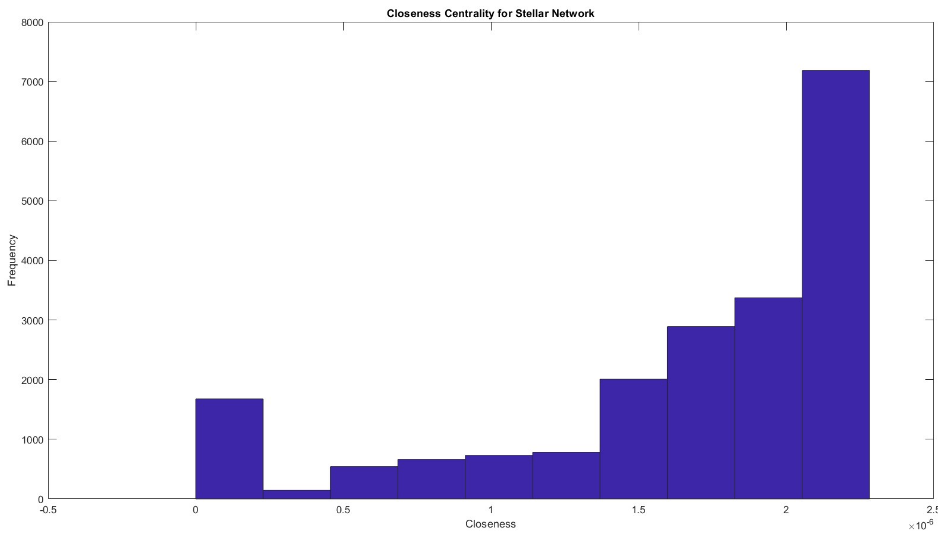

Figure 11: *Closeness Centrality for the Stellar Network*

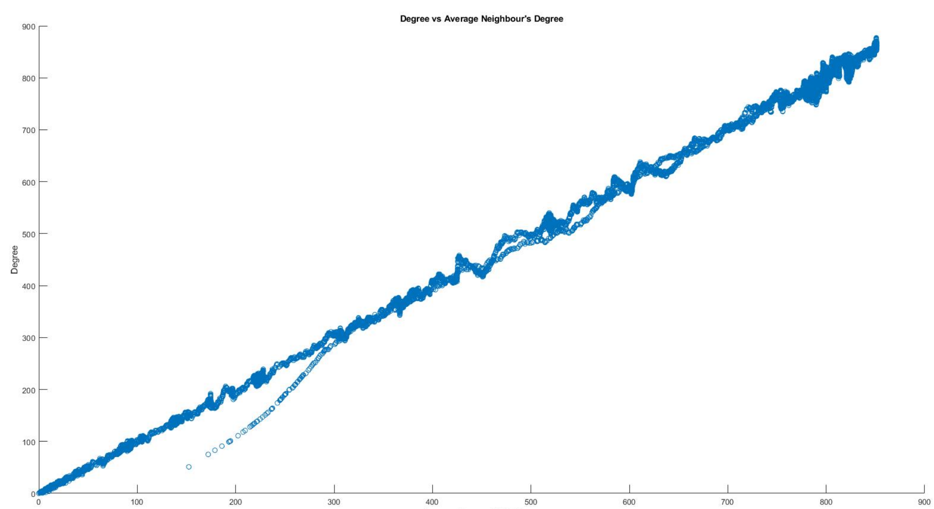

Figure 12: *average degree of each neighbour compared to each node degree*

with the chosen sampling. Restricting our analysis to a cubic region, the correlation coefficient is reduced.

Finally, we have calculated the same values using only a cubic region of our network. The correlation coefficients obtained are the following:

- Degree Centrality vs Magnitude: 0.4592
- Betweenness Centrality vs Magnitude: 0.0082
- Closeness Centrality vs Magnitude: 0.0.3104
- Pagerank Centrality vs Magnitude: 0.0657

The correlations shown seem to all agree with the fact that globular clusters' stars have higher magnitudes, while the less grouped the stars the lower the magnitude will be; we can also compare our results to random results, observing no correlation at all to any of the random centralities (correlation coefficient of 0.0084).
In figure 13 we can see that there is a range of magnitudes (from -5 to 15) for which there is a clear well defined correlation between Page-Rank and absolute magnitude, while for the values between -20 and -5 pagerank centrality does not seem related to magnitude; this can be related to the fact that some of those magnitudes probably don't correspond to stars (as, indeed, the brightest star known has a bolometric absolute magnitude of -12.41 [Brands et al.(7 April 2022)], while some of those stars have absolute magnitudes below -15) or may have incorrect magnitudes. The correlation is also slightly stronger (correlation coefficient of 0.7719) between the logarithm of the pagerank centrality and the magnitude, as suggested by the apparent "flatness" of the plot in figure 13. The same behaviour can be seen for degree centrality and closeness centrality.

IX. Conclusions

As we have seen there is a clear difference between the scatter plot of real positions and the multidimensional scaling; in figure 2 it is possible to visualize the cluster structure of the network in the shape of spirals whose centres contains stars with the highest betweenness centrality. This kind of visualization can be useful as every star in the same spiral share the same properties (i.e. magnitude, age, metallicity, position, ...); the presence of clusters was expected as a consequence of the degree distribution (figure 9) which is strongly correlated to clustering coefficient for this network ($r \simeq 1$).

As proven in section VII, the network is not random, meaning that we can expect an evolution of galaxies and their stellar population following certain criterias specified by classic astrophysics (as explained in V), allowing us to look for correlations between physical parameters (like magnitude) and topological network properties (degree centrality, betweenness centrality, ...); the obtained results are based on the HYG Catalogue (VI) but we expect them to be general, with similar results using different datasets.

In section VIII we can see that more points are needed for this analysis to be precise (as we can see from figure 9 and figure 10).
The degree centrality is correlated to closeness centrality and clustering coefficient while it has no correlation to betweenness centrality, suggesting that clusters may not be linked together. The network also shows an assortative behaviour (as seen in figure 12), which means that stars with high degree link to other similar stars, showing the existence of clusters of different dimensions.

In section i we have shown that the values of degree centrality, closeness centrality and pagerank centrality are correlated to magnitude, meaning that stars in globular clusters are less bright which is related to the advanced age of a globular cluster; blue and more brilliant stars (i.e. blue giants) remains in the main sequence for less time with respect to less massive and red stars (not to be confused with red giants, which are instead very massive due to their bigger dimension), which makes the vast majority of stars of a globular clusters.
In figure 13 it can be viewed a comparison between pagerank centrality and magnitude, showing a relationship between the two param-

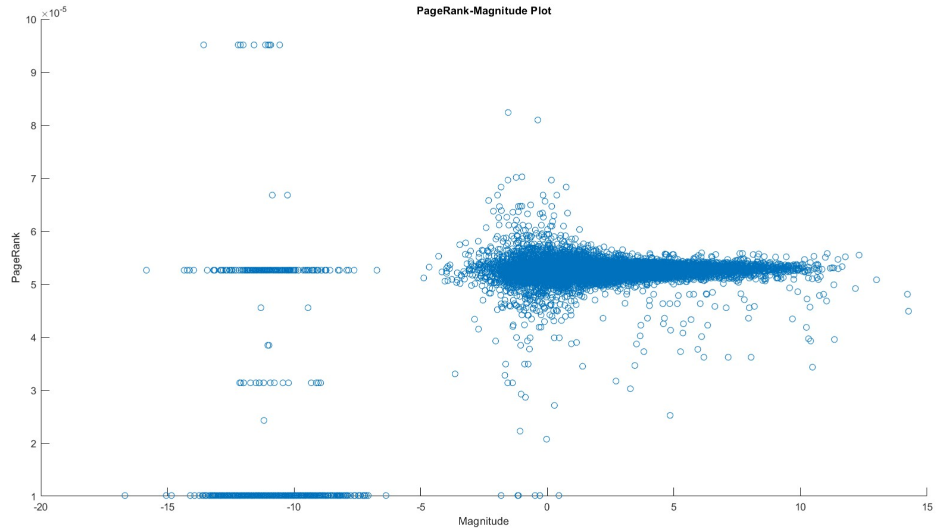

Figure 13: *PageRank-Magnitude plot for the Stellar Network*

eters but also the presence of some incorrect absolute magnitudes in the dataset; the precision of this analysis could be increased by using a different, and more recent, dataset.

i. Future Perspectives

As we have seen some of the limits of this analysis were related to the low amount of points used (due to the computing power available), but also to the dataset used (since the HYG catalogue is a bit outdated, even if it works well as a didactic catalogue).
A suggested example of a more modern catalogue may be Gaia EDR3 ([Babusiaux et al. (2022)]), even if it is way more complex to handle with.
From the computational point of view it may be better to use at least 10^5 points, required as globular clusters can have up to this amount of stars (see section V), which is hardly possible with a personal computer.

The obtained results are highly promising and may permit to obtain further informations on galaxies' star compositions (young MS stars, red giants, ...) and their topology.

Finally this analysis can be extended to galaxies as shown by [Hong and Dey, 2015] and in the future, thanks to the James Webb Space Telescope, we could extend this analysis to primordial galaxies, which requires to look at them in the far infrared, trying to discover something more about the beginning of our universe.

X. Appendix

i. Appendix A - Magnitude

The magnitude is a measure of the brightness of an astronomical object in a specific passband, usually the visible or the IR. The origin of this concept must be attributed to Hipparchus (190-120 BC) which produced a catalogue of the apparent brightness of stars, giving values from 1 (for the brightest) to 6 (the less bright ones). Nowadays the scale used is logarithmic and follows this definition: a magnitude 1 star is exactly 100 times brighter than a magnitude 6 star. It means that the lower the value of the magnitude is, the brighter the object appears. However, we must distinguish between two types of magnitude: apparent and absolute. The apparent magnitude (m) is defined as the

brightness of an object as it appears in the sky: naturally, this definition depends on the luminosity of the object but also on the distance. The absolute magnitude (M) is instead a measure of intrinsic luminosity of an object, and it is therefore independent from the distance. More formally, it is defined as the apparent magnitude (m) of that object if it were placed at 10 parsecs from the Earth.
Fixed an object as a reference, the apparent magnitude of another object 1 is given by $m_1 - m_{ref} = -2.5\log_{10}(\frac{I_1}{I_{ref}})$ where I is the power per unit area (intensity) measured from Earth. Naturally, with this formula we can go beyond the 1-6 Hipparchus scale.
Following the previous definition, the absolute magnitude can be calculated from the apparent one (knowing the distance of the object): $m - M = 2.5\log_{10}(\frac{d}{10})^2$ where d is the distance in parsecs and m is the apparent magnitude.
Some numbers: the sun has an apparent magnitude of -27, Sirius (the brightest one in the sky) -1.46, the ISS (international space station) arrives to -6.

ii. Appendix B - Random Networks

During the analysis, an important role is played by the topology of the graph. For this reason, it is important to understand whether if our graphs are random or not. Usually, by random graph we mean a network where some parameters are fixed while others are chosen randomly. For example, it is possible to fix the number of node n and link m. At this point we can choose m times 2 nodes in a random way and connect them. This is the original idea by Erdős and Rényi. However, it is also possible to follow the Gilbert model fixing the number of nodes n and a probability p: each pair of nodes is connected following the fixed probability. A network such this can be constructed starting with n nodes, choosing two of these nodes and generating a number between 0 and 1: if this number exceeds p, a link between the nodes is created. Then, we can repeat this operation $N(N-1)/2$ times. It is possible to define the degree distribution p_k as the probability that a random node has degree k. Moreover, it is not difficult to see that this distribution must be binomial: the probability that the chosen node has exactly k links is given by the product of three terms: the probability that k links are present (p^k), the probability that the remaining n-1-k are not present ($(1-p)^{n-1-k}$) and the number of all possible way to select k links from N-1 available. Therefore, we have $p_k = \binom{n-1}{k} p^k (1-p)^{n-1-k}$ and then we can obtain the network's average degree $<k>= \sum_{k=0}^{n} kp_k = Np$ and the variance. Notice that for sparse networks, where we have $<k> \ll n$ the binomial distribution is well-approximated by the Poisson distribution $p_k = e^{-<k>}\frac{<k>^k}{k!}$, giving the advantage that average values are easier to calculate. Finally, since there is no n dependece in the Poisson distribution, if we consider two networks with different sizes but same $<k>$ these will be undistinguishable.

References

[Hong and Dey, 2015] Sungryong Hong and Arjun Dey. (2015). Network Analysis of Cosmic Structures : Network Centrality and Topological Environment. *Royal Astronomical Society*, 450:1999–2015.

[Anders+, 2022] Anders et al. (2022). StarHorse2, Gaia EDR3 photo-astrometric distances. *VizieR Online Data Catalog*.

[Perryman+, 1997] Perryman, M. A. C. and Lindegren, L. and Kovalevsky, J. and Hoeg, E. and Bastian, U. and Bernacca, P. L. and Crézé, M. and Donati, F. and Grenon, M. and Grewing, M. and van Leeuwen, F. and van der Marel, H. and Mignard, F. and Murray, C. A. and Le Poole, R. S. and Schrijver, H. and Turon, C. and Arenou, F. and Froeschlé, M. and Petersen, C. S.. (1997). The HIPPARCOS Catalogue *Astronomy and Astrophysics, Vol. 323, p.L49-L52*.

[Bahcall+, 1987] Bahcall, John N. and Casertano, Stefano and Ratnatunga, Kavan U..

(1987). An Analysis of the Yale Bright Star Catalog *Astrophysical Journal v.320, p.515* .

[Gliese, 1957] Gliese, W.. (1957). Katalog der Sterne näher ALS 20 Parsek für 1950.0 *Mitt. Astron. Rechen-Inst. Heidelberg (1957)* .

[Gliese, 1969] Gliese, W.. (1957). Catalogue of Nearby Stars. Edition 1969 *Veroeff. Astron. Rechen-Inst., Vol. 22, p. 1 (1969)* .

[Gregersen, Erik (2010)] Gregersen E.. (2010). The Milky Way and beyond *The Rosen Publishing Group. pp. 35–36* .

[Dreyer, J.L.E. (1953)] New General Catalogue of Nebulae and Clusters of Stars (1888), Index Catalogue (1895), Second Index Catalogue (1908). *London: Royal Astronomical Society, 1953. (28,2157)* .

[van Albada, T. S.; Baker, N. (1973)] On the two Oosterhoff groups of globular clusters. *Astrophysical Journal. 185: 477–498.* .

[M. E. J. Newman (2003)] Mixing patterns in networks. *Department of Physics, University of Michigan, Ann Arbor, MI 48109–1120 and Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM 87501* .

[M. E. J. Newman (2006)] The Mathematics of Networks.

[Brands et al.(7 April 2022)] The R136 star cluster dissected with Hubble Space Telescope/STIS. III. The most massive stars and their clumped winds. *Astronomy and Astrophysics* .

[de Regt et al. (2018)] Network analysis of the COSMOS galaxy field. *Monthly Notices of the Royal Astronomical Society, Volume 477, Issue 4, July 2018, Pages 4738–4748,* .

[Mead, A (1992)] Review of the Development of Multidimensional Scaling Methods. *Journal of the Royal Statistical Society, Series D (The Statistician). 41 (1): 27–39* .

[Peebles, P. J. E. (1980)] The large-scale structure of the universe. *Princeton University Press, 1980. ISBN: 978-0-691-08240-0* .

[Babusiaux et al. (2022)] Gaia Data Release 3: Catalogue Validation. *A and A Gaia Data Release 3 special issue* .