Institute of Space Sciences
EXCELENCIA MARÍA DE MAEZTU
ICREA

Visualizing the pulsar population using graph theory

Diego F. Torres

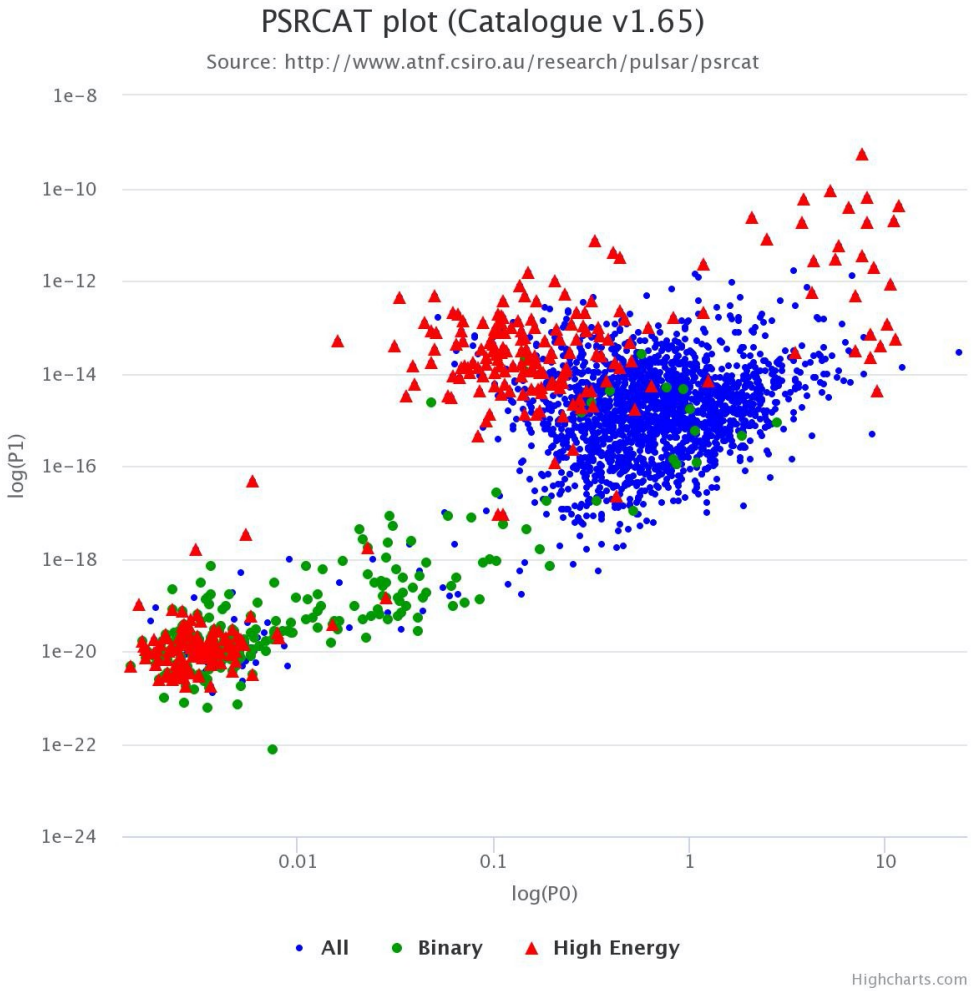
with C. Rodriguez and A. Patruno

The pulsar population

2509 ATNF pulsars with period and period deriv.

Institute of Space Sciences

EXCELENCIA MARÍA DE MAEZTU



The P-Pdot diagram.

Looked a million times in pulsar research…

version 1.67 (2022) of the catalog:
3282 pulsars listed
2509 pulsars with P and Pdot

Manchester et al. 2005, ATNF web: https://www.atnf.csiro.au/research/pulsar/psrcat/

Dipole model as proxy for physical conditions: Intrinsic properties

- Spin period:

P [s],

- Spin period derivative:

$\dot{P}$ [s s^{-1}],

- Surface magnetic flux density (equator):

$$B_s = (3c^3 I)^{1/2}/(8\pi^2 R^6 \sin^2\alpha)^{1/2}\sqrt{P\dot{P}} \simeq 3.2\times 10^{19} P^{1/2}\dot{P}^{1/2}\mathrm{G},$$

- Magnetic field at the light cylinder:

$$B_{lc} = B_s(\Omega R)^3/c^3 \simeq 3\times 10^8 P^{-5/2}\dot{P}^{1/2}\mathrm{G},$$

- Spin-down energy loss rate:

$$\dot{E}_{sd} = 4\pi^2 I\dot{P}P^{-3} \simeq 3.95\times 10^{46} P^{-3}\dot{P} \quad \mathrm{erg\ s^{-1}},$$

- Characteristic age:

$$\tau_c = P/2\dot{P} \simeq 15.8\times P\dot{P}^{-1}\mathrm{Myr},$$

- Surface electric voltage:

$$\Delta\Phi = (B_s 4\pi^2 R^3)/(2cP^2) \simeq 6.3\times 10^5 P^{-3/2}\dot{P}^{1/2}\mathrm{V},$$

- Goldreich-Julian charge density:

$$\eta_{GJ} = (\Omega B_s)/(2\pi ce) \simeq 7\times 10^{10} P^{-1/2}\dot{P}^{1/2}\mathrm{cm^{-3}}.$$

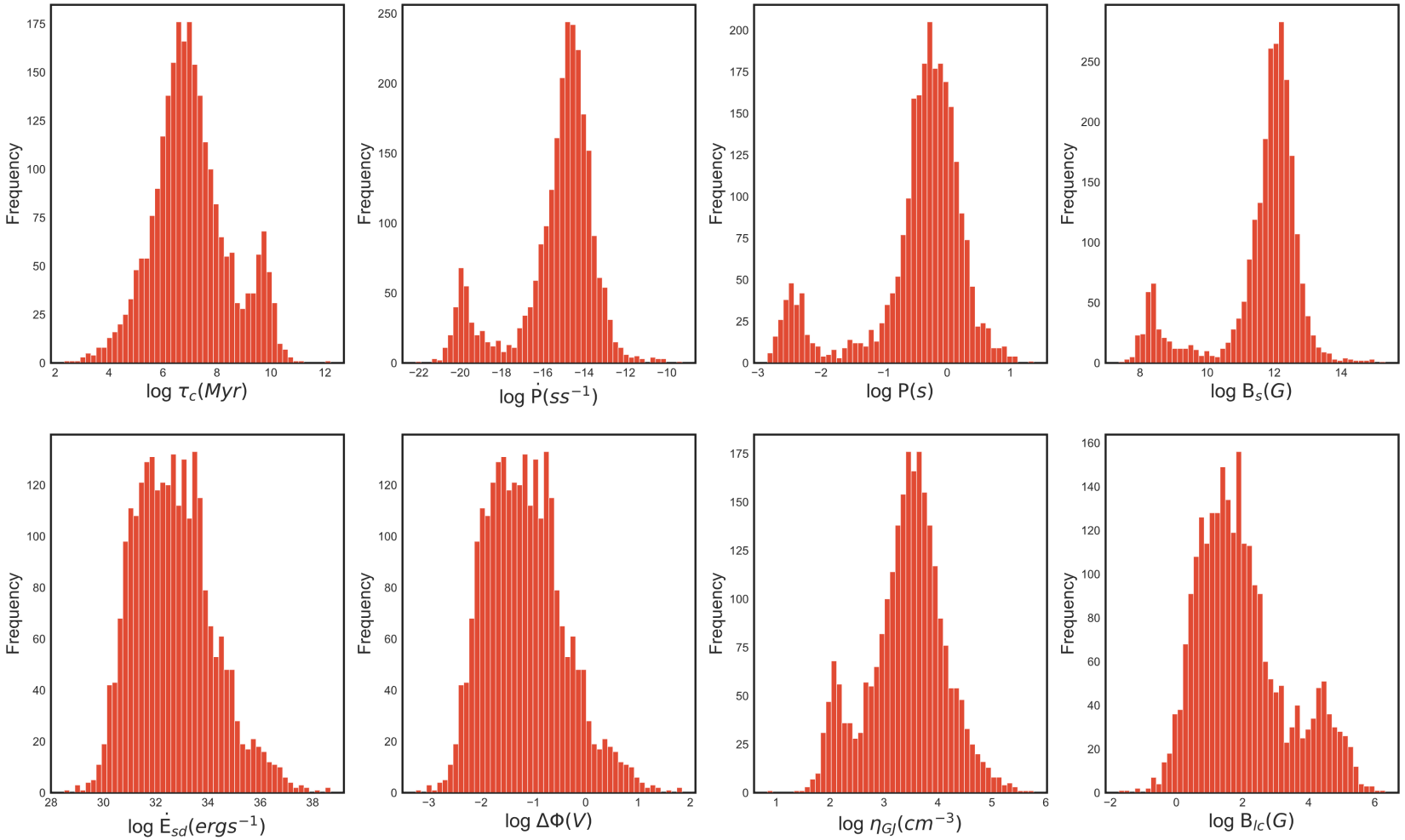
- The measurable quantities P and Pdot are the leading magnitudes in this set, from which all others are calculated using the rotating dipole model, as is usual for pulsar estimations

e.g. Lorimer & Kramer 2012

Many orders of magnitude... Log is better

Institute of Space Sciences

EXCELENCIA MARÍA DE MAEZTU



C. R. García, DFT, A. Patruno, MNRAS 2022

Treating variables

Normalizing

- Given that the distributions are not normal, we use the robust scaler for normalizing the log variables

$$x_i^{\dagger} = \frac{x_i - Q_2}{IQR}$$

- Q1, Q2, Q3 are the 1st quartile, median, and 3rd quartile of the distribution, respectively, and IQR is the interquartile range, (Q3-Q1)

- The distributions of the variables after being normalized have a median equal to zero and an IQR equal to one. Note that if we take the logarithm of the variables and then normalize it, relations between variables are more clearly uncovered. For instance, $\log \dot{E}_{sd}^{\dagger} = \log \Delta\Phi^{\dagger}$

- Considering both of them at once in defining the nearness of two given pulsars relates to the fact that the physical meaning represented by the two original magnitudes is different.

Principal components analysis

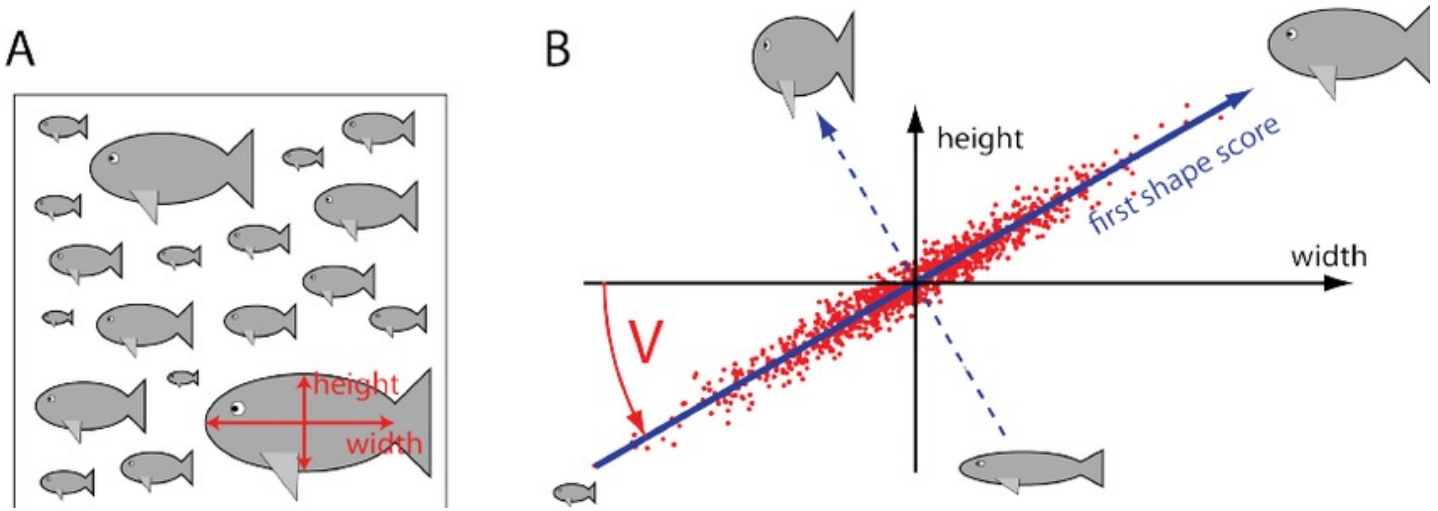
Concept

- It is a technique to reduce the dimensionality of a problem, capturing as much as the original variance of the population in as few variables as possible
  - Said otherwise:
  - *which linear combinations of variables capture most of the information about the dataset?*

- In our case, we know from the start that we have only 2 independent variables on which all others linearly (because of the log) depend, thus the dimension is 2.

- But are these the principal components?

Concept



We can describe the shape of a fish with two variables: height and width. However, these two variables are not independent of each other. In fact, they have a strong correlation. Given the height, we can probably estimate the width; and viceversa. Thus, we may say that **the shape of a fish can be described with a single component.**

P and Pdot are not the principal components. The population variance is not contained by them.

$$
\begin{aligned}
PC_1 \;=\;& 0.21 B^{\dagger}_{lc_l} - 0.29 \eta^{\dagger}_{GJ_l} + 0.05 \Delta\Phi^{\dagger}_{l} + 0.05 \dot{E}^{\dagger}_{sd_l} - 0.46 \dot{P}^{\dagger}_{l} \\
& -0.59 B^{\dagger}_{s_l} - 0.47 P^{\dagger}_{l} + 0.29 \tau^{\dagger}_{c_l}, \\
& (71.6\% \;\text{of the explained variance}), \\
PC_2 \;=\;& 0.43 B^{\dagger}_{lc_l} + 0.32 \eta^{\dagger}_{GJ_l} + 0.47 \Delta\Phi^{\dagger}_{l} + 0.47 \dot{E}^{\dagger}_{sd_l} + 0.19 \dot{P}^{\dagger}_{l} \\
& +0.05 B^{\dagger}_{s_l} - 0.36 P^{\dagger}_{l} - 0.32 \tau^{\dagger}_{c_l}, \\
& (28.4\% \;\text{of the explained variance}),
\end{aligned}
$$

where the †-quantities refer to the normalized ones
the sub-index l stands to note that the normalization is applied to the logarithm of the variable.

P and Pdot are not the principal components. The population variance is not contained by them.

$$PC_1 = -8.471 - 1.178 \log P - 0.832 \log \dot{P},$$
$$PC_2 = 14.182 - 2.931 \log P + 1.105 \log \dot{P}.$$

Note that no dag marking is herein needed, as we have absorbed the corresponding IQR and median of each variable into the coefficients and that the units of P and $\dot{P}$ are s and s/s as before

P and Pdot are not the principal components.
The population variance is not contained by them.

Institute of Space Sciences

EXCELENCIA MARÍA DE MAEZTU

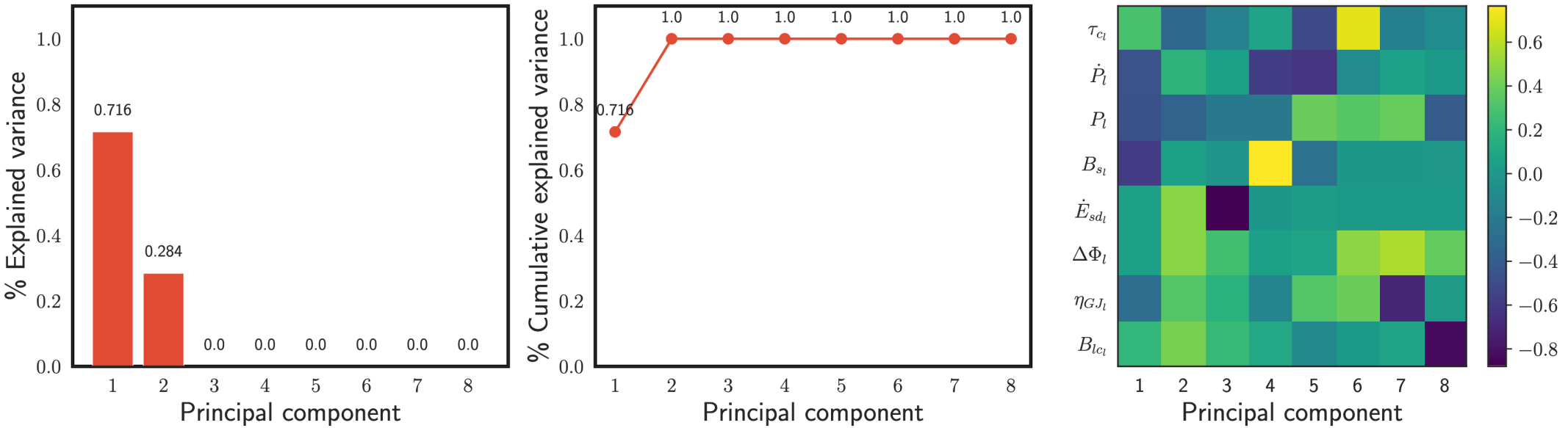
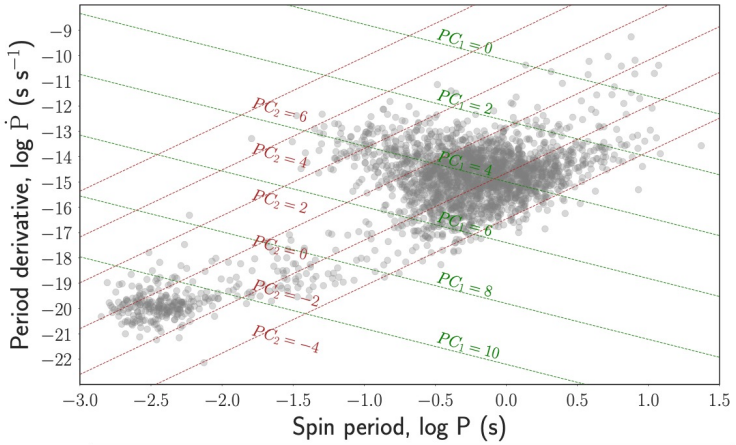
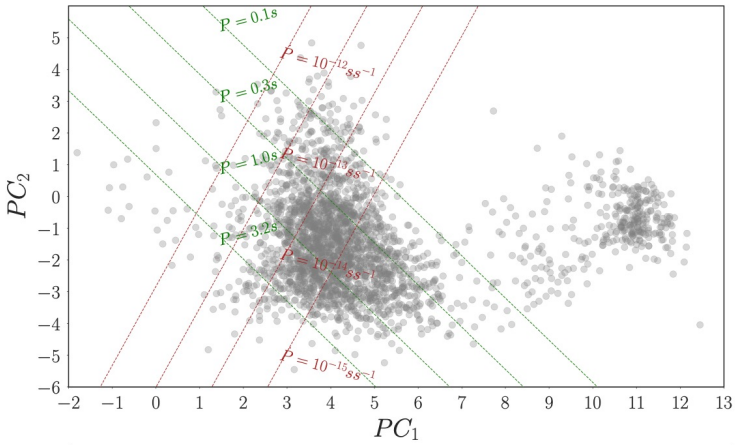


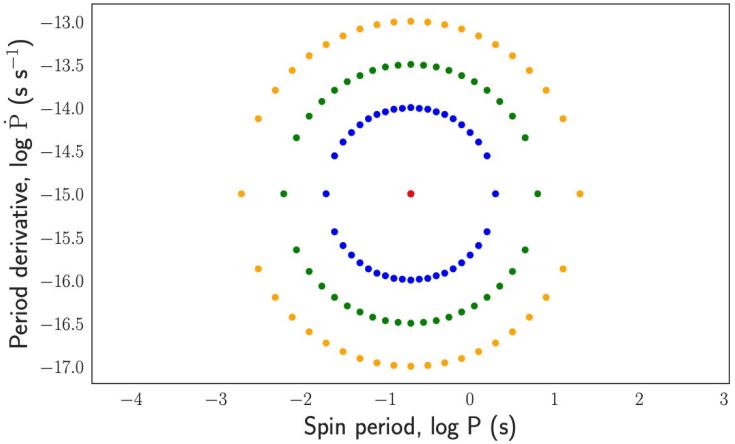
Figure 2. PCA results using the logarithm of the set of variables for the whole population of 2509 pulsars. The left panel shows the explained variance contributed by each of the PCs according to the eigenvalues of the covariance matrix. It represents the amount of information contained in each PC. The central panel shows the cumulative variance explained by the new set of variables which has been defined through the PCA analysis. The right panel shows the 'weight' that each variable has with respect to each PC. This value is the coefficient that is held in each eigenvector. Negative values imply that the variable and the PC in question are negatively correlated. Conversely, a positive value shows a positive correlation between the PC and the variable.

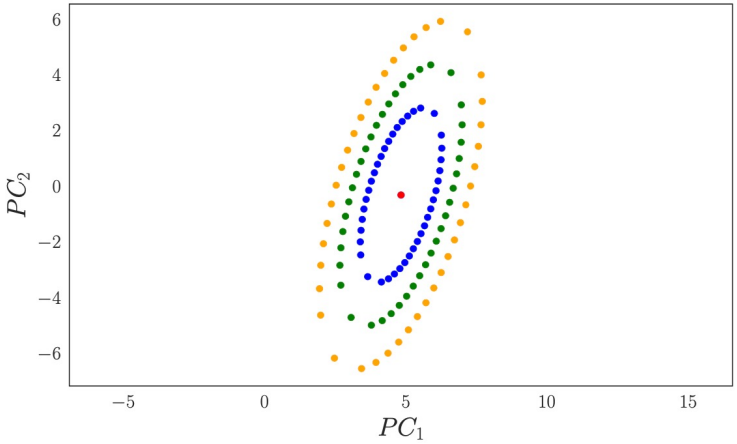
P and Pdot are not principal components.
The population variance is not contained by them.

Institute of Space Sciences
EXCELENCIA MARÍA DE MAEZTU

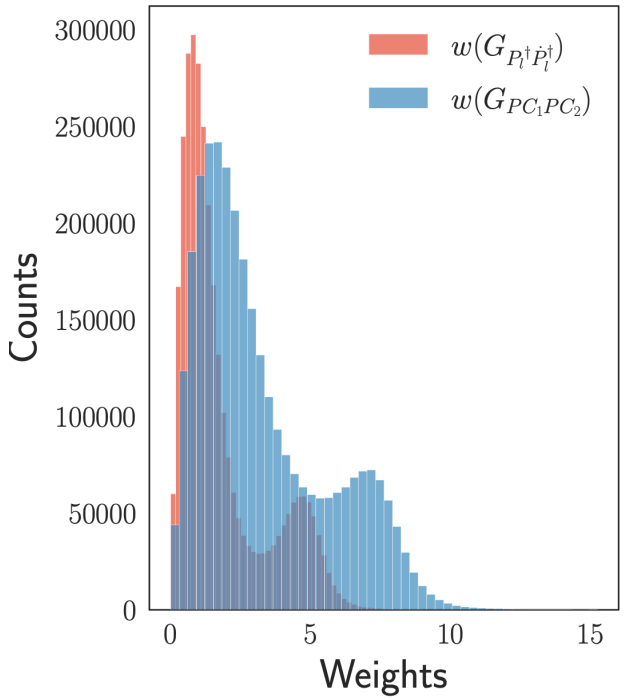








The distribution of distances using only P-Pdot and all the variables of interest differ.



Take home message:

a distance ranking or visualization based only on P and Pdot is potentially misleading.

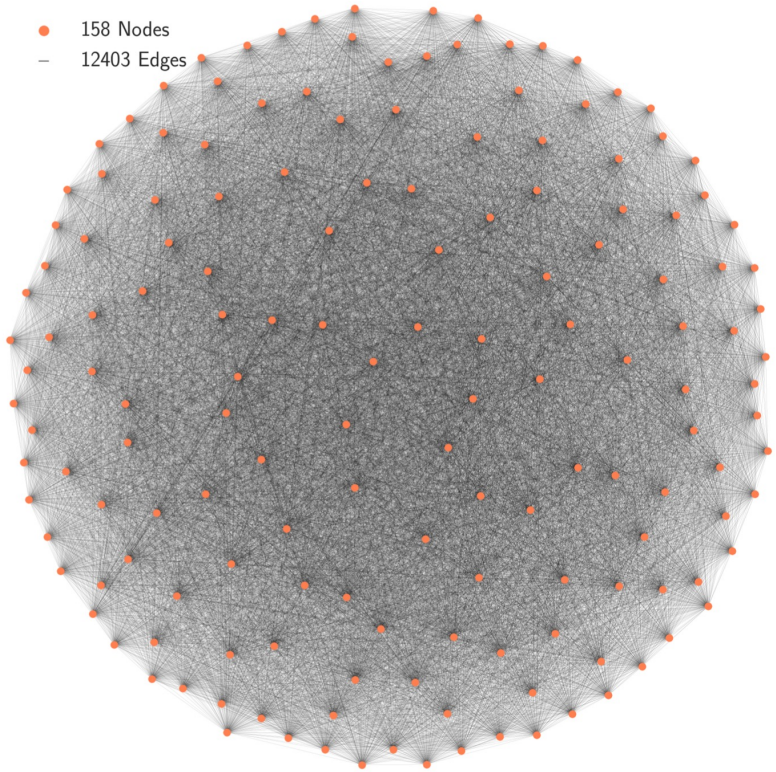
Graph theory over variables representing the full variance of the population

Basics of graph theory I: Distance and weight

- Aim: establish relationships among objects based on the connections they have, representing them in a graph.
- G (V, E) will denote a graph of a set V, of nodes v; and a set E, of edges e joining these nodes.
- The edges are assigned with a weight w (e.g., a distance value) representing the relationship between the nodes
- To assign the value of w for a given edge, the simplest possibility is to assume the weight to be the Euclidean distance between the nodes in an N−dimensional space of interest

$$d_{nm} = \sqrt{\sum_{j=1}^{N} \sum_{n=1}^{V} \sum_{m>n}^{V} (v_{jn} - v_{jm})^2}$$

Basics of graph theory II:
An example of a complete graph, with 158 nodes



Complete, undirected and weighted graph $G(158, 12403)$.

The total weight of a graph G will be obtained from the sum of each specific weight on each edge

$$w(G) = \sum_{e \in E(G)} w(e).$$

A path is a sequence of nodes through G in which no node is encountered more than once.

Then G is a connected graph where any pair of nodes can communicate through a path.

If a path is closed it is called a cycle. I

The partition of V into two sets (S, V − S) is called a cut.

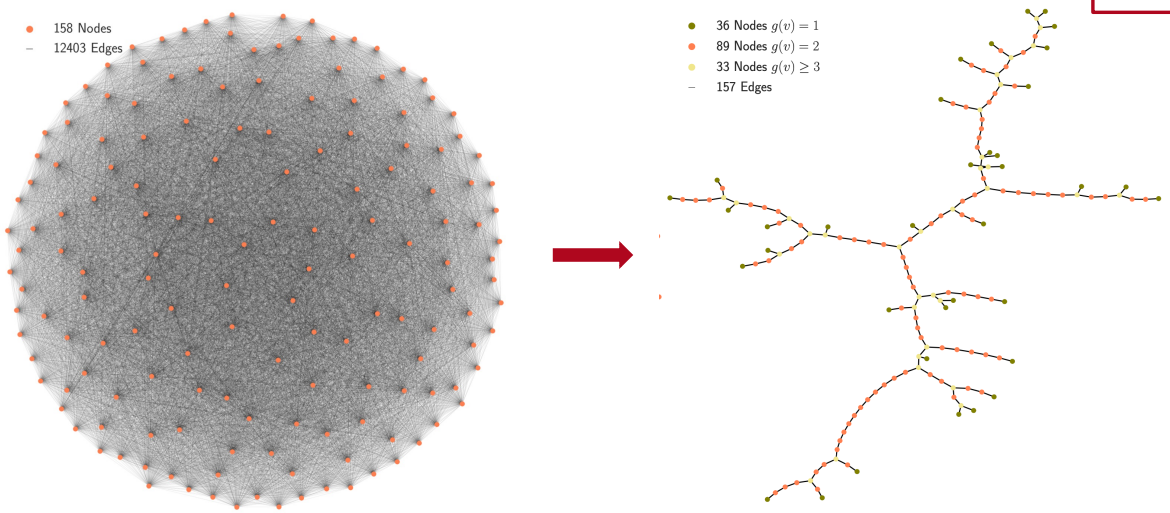
The minimal spanning tree (MST)

The MST is a sub-graph of G such that the sum of all weights is the smallest possible.

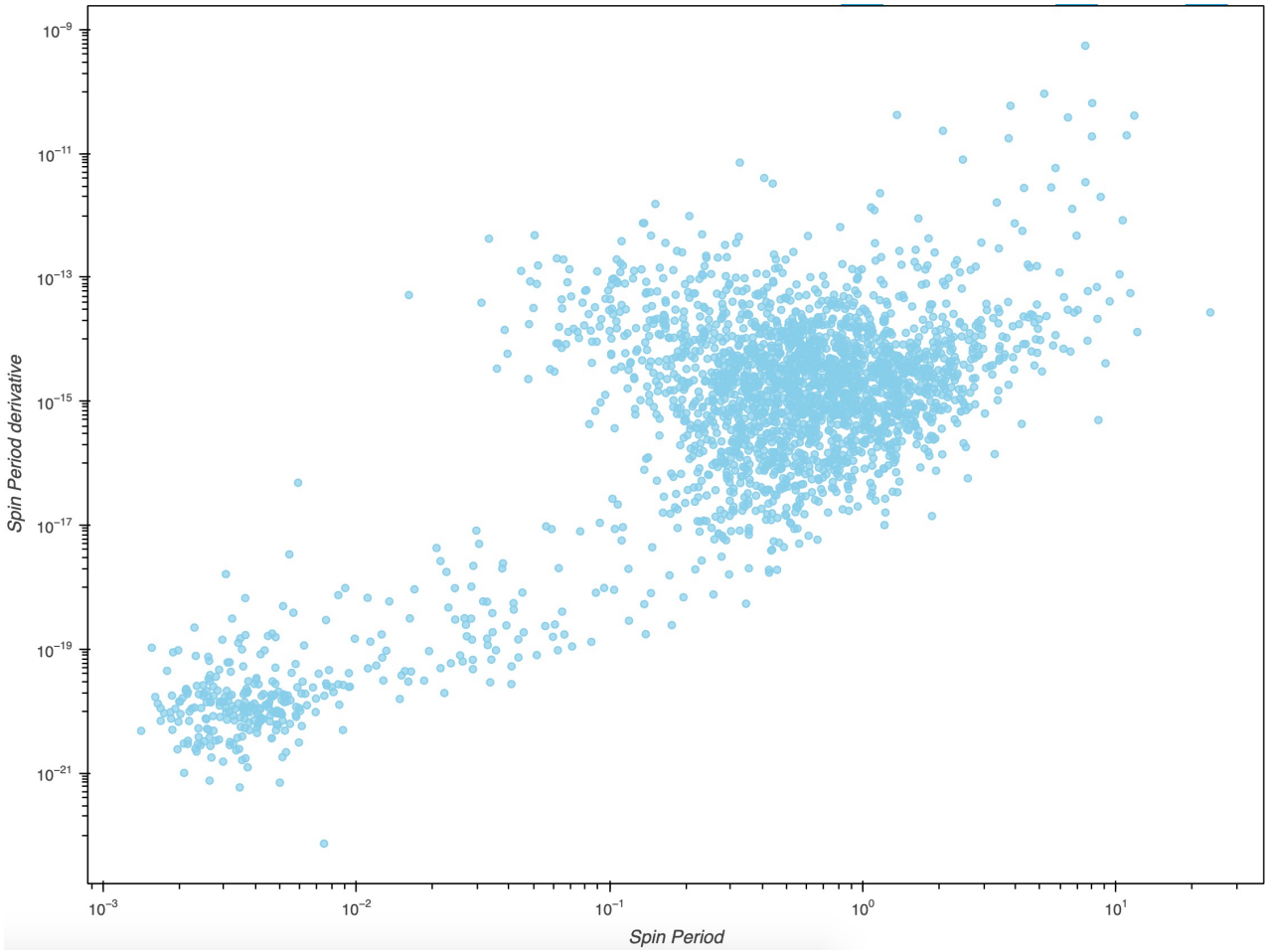
The MST connects all the nodes taking into account the minimum distance between them at a local level, but under the condition that there is a global minimization for connecting the whole population of nodes.

There are several equivalent ways to compute MSTs, one such is the Kruskal algorithm:

- Order the edges by increasing weight.
- Choose the edge with the smallest $w(e)$ and add it to E'.
- Make sure that the chosen e does not produce any cycle in the structure of T.
- The process finishes when all nodes V are connected, thus resulting in a graph $T(V, E', w')$, where $|E'| = |V| - 1$ and $w'(T)$ is the weight of T according to Eq. (12)

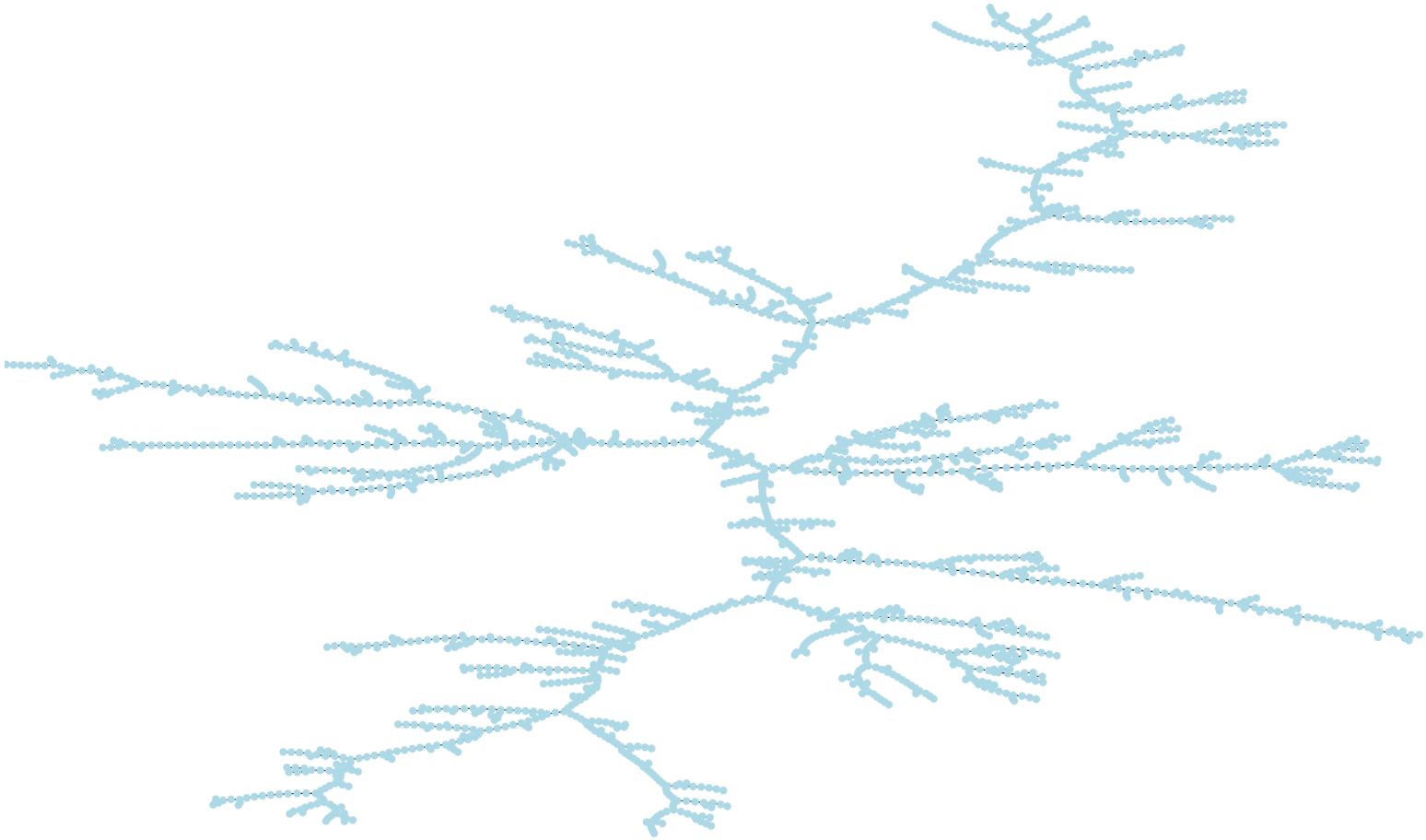


So we move from this...



C. R. García, DFT, A. Patruno, MNRAS 2022

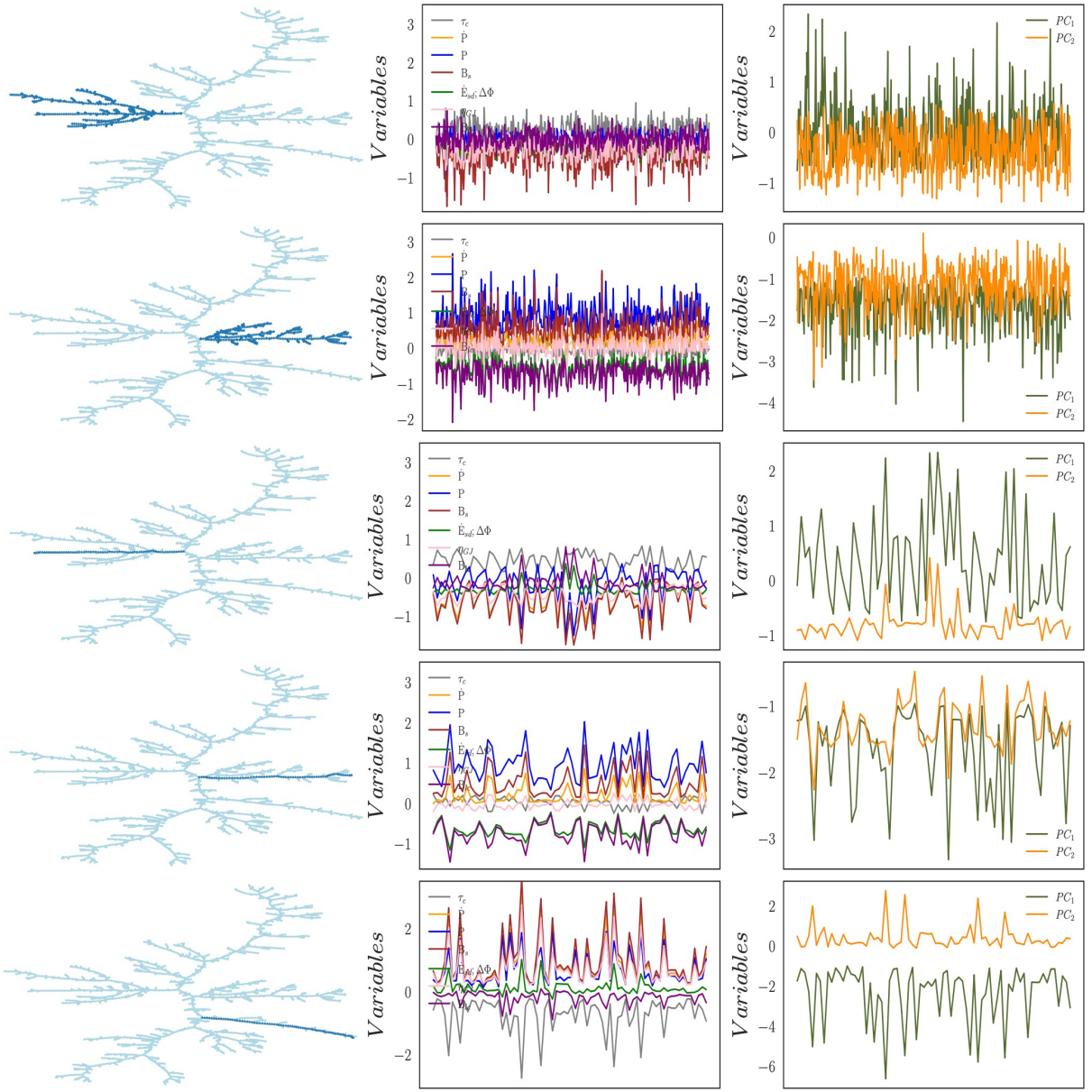
to this.. MST of the pulsar population



C. R. García, DFT, A. Patruno, MNRAS 2022

Basic branch analysis

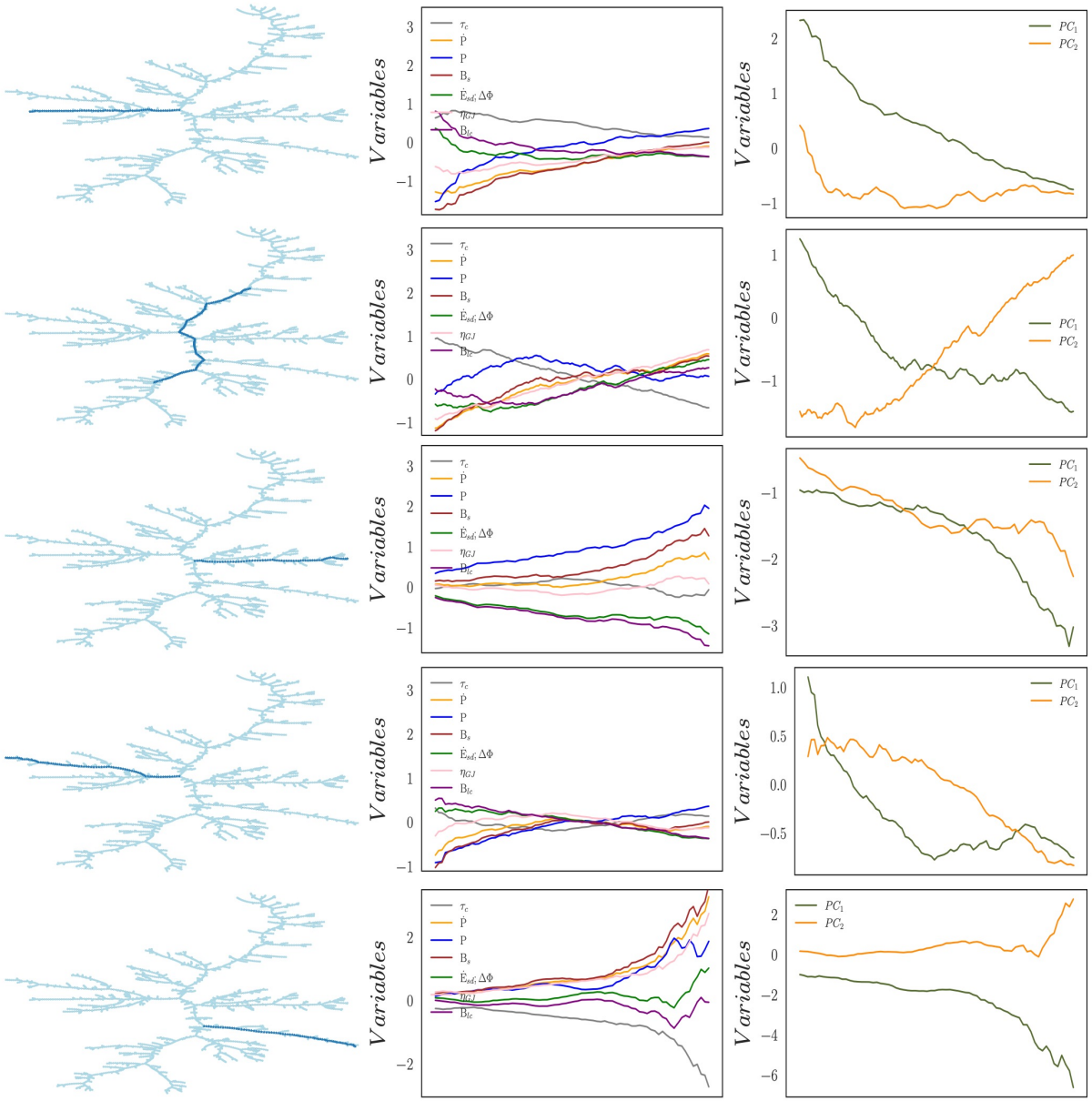
- Mixing nodes from different branches of the MST produces a scattered distribution of variables.
- This is a generic behavior that happens for any mixing of the branches in the MST, and for any mixing of the nodes even within a single branch.
- If we read the MST in a disordered manner, nothing is learned from it.



C. R. García, DFT, A. Patruno, MNRAS 2022

Branch analysis

- Instead, if we choose one of the branches at a time and run along with the nodes in it in an orderly manner, a smooth behavior of the variables naturally appears.

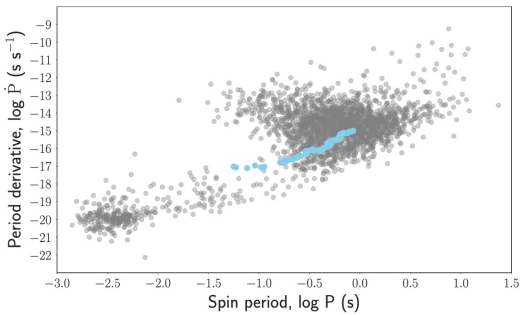


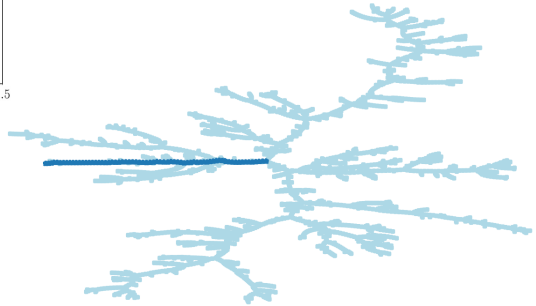
C. R. García, DFT, A. Patruno, MNRAS 2022

Branch analysis: read branches in order

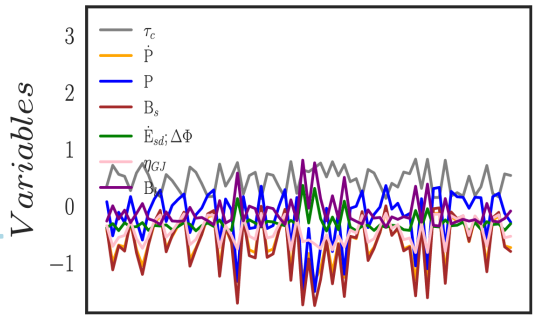
Institute of Space Sciences

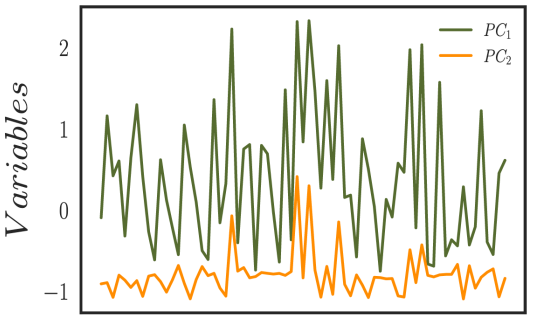
EXCELENCIA MARÍA DE MAEZTU



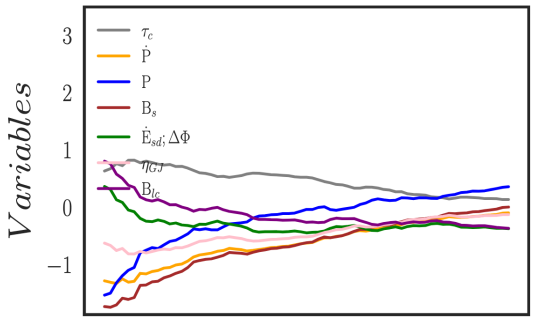


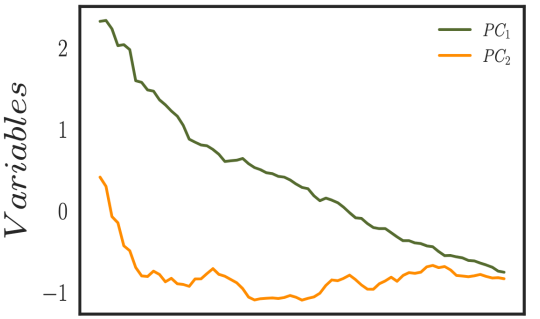
disordered





in order of the connections

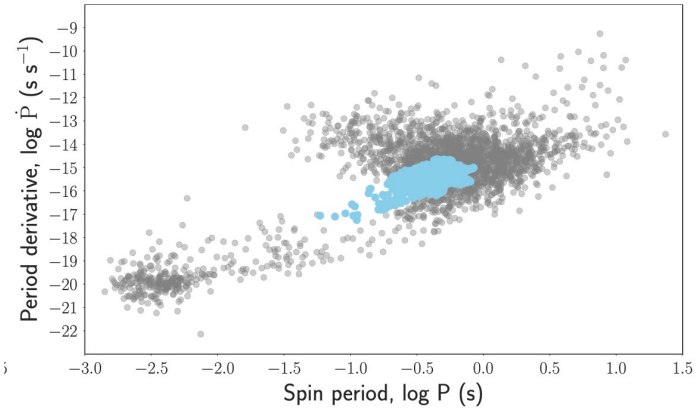


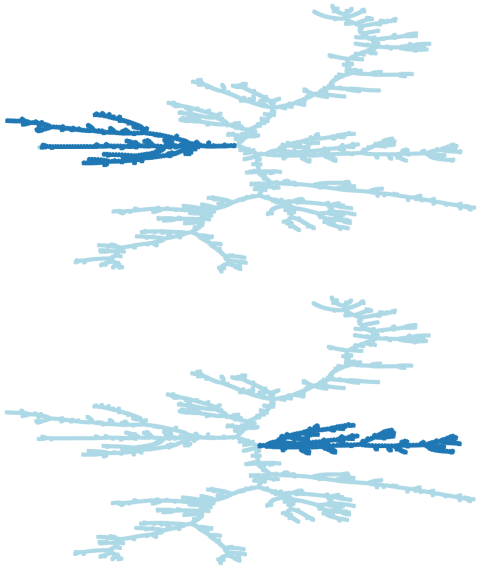


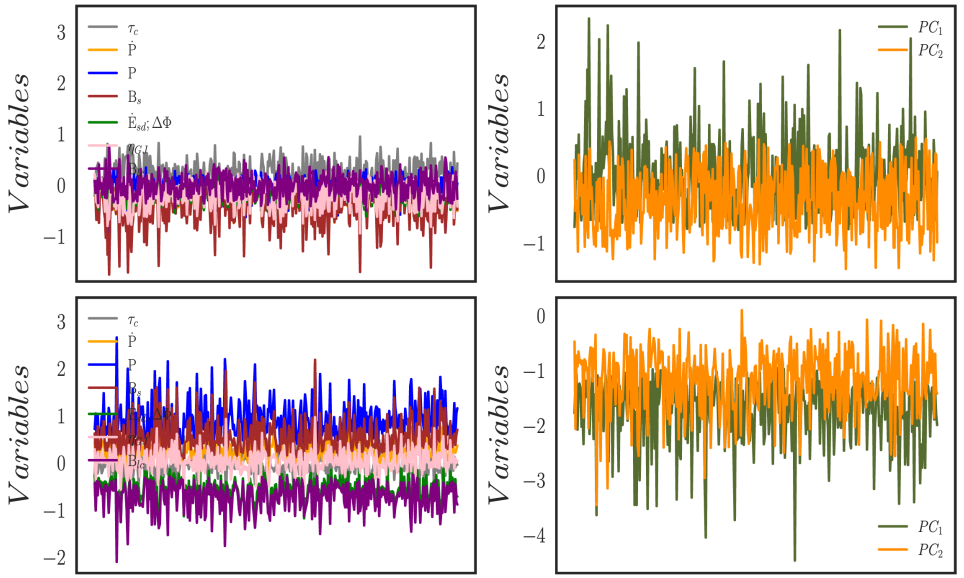
Branch analysis: PPdot nearness is like reading in disorder a branch of the MST

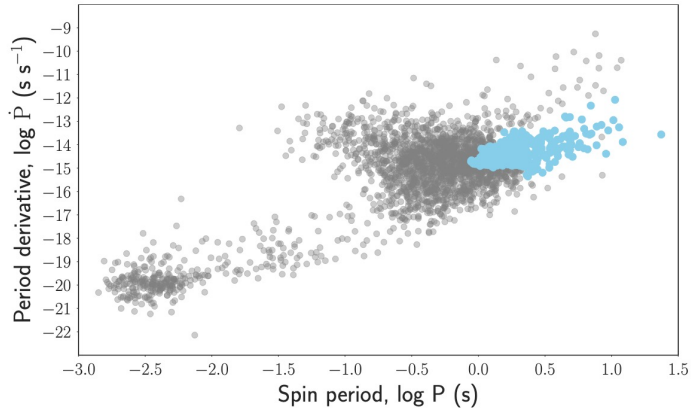
Institute of Space Sciences

EXCELENCIA MARÍA DE MAEZTU



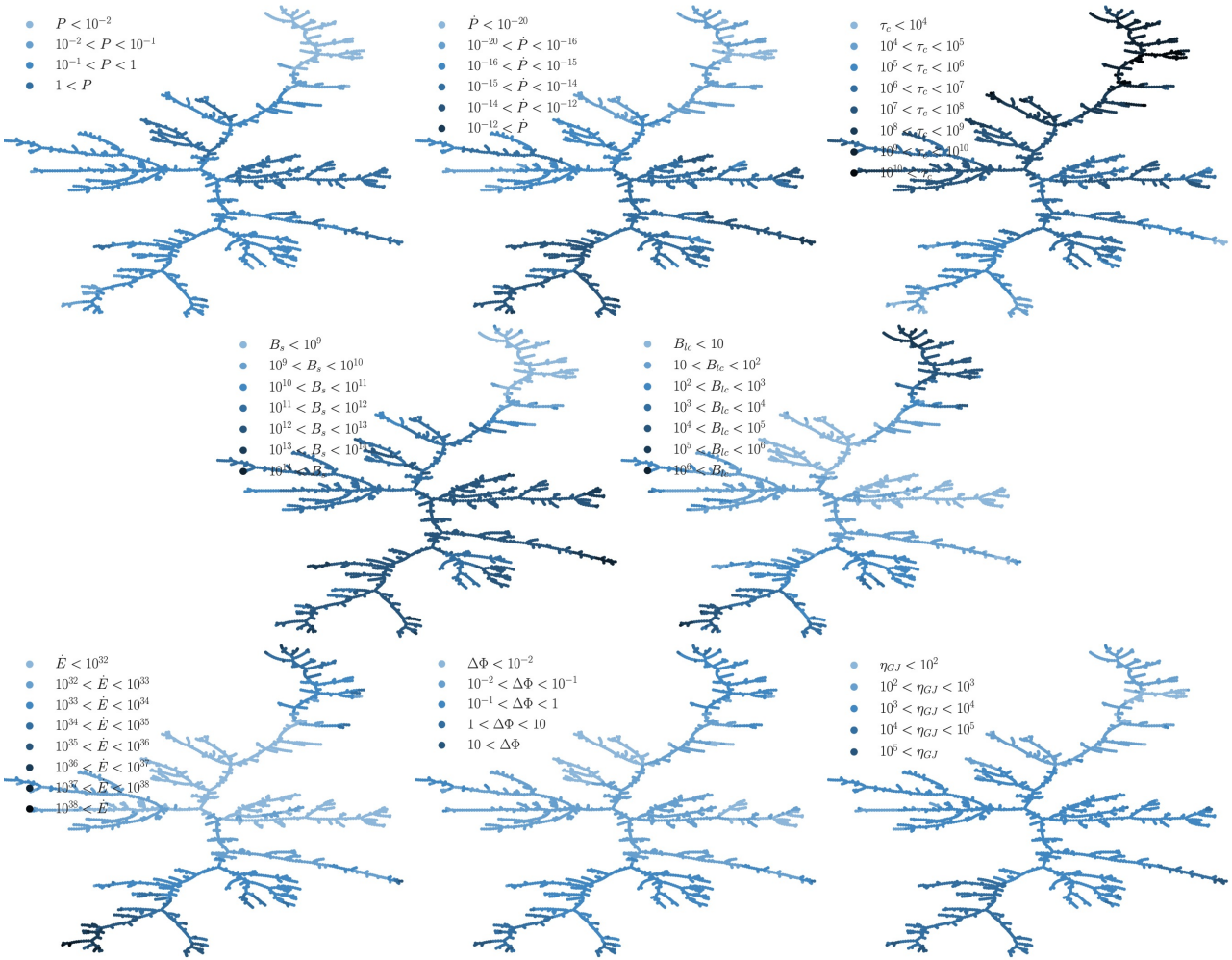






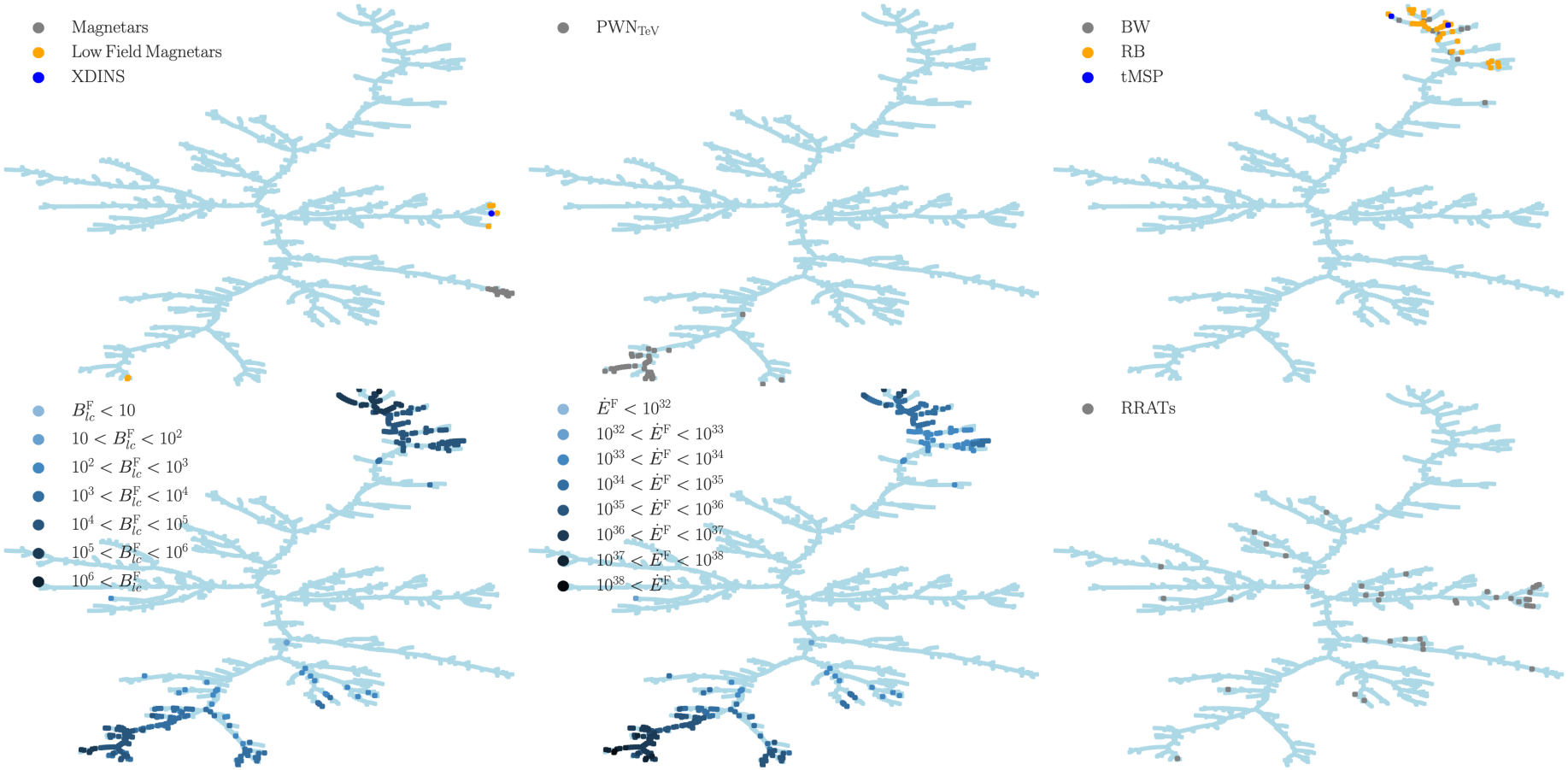
C. R. García, DFT, A. Patruno, MNRAS 2022

The pulsar tree as a descriptive tool

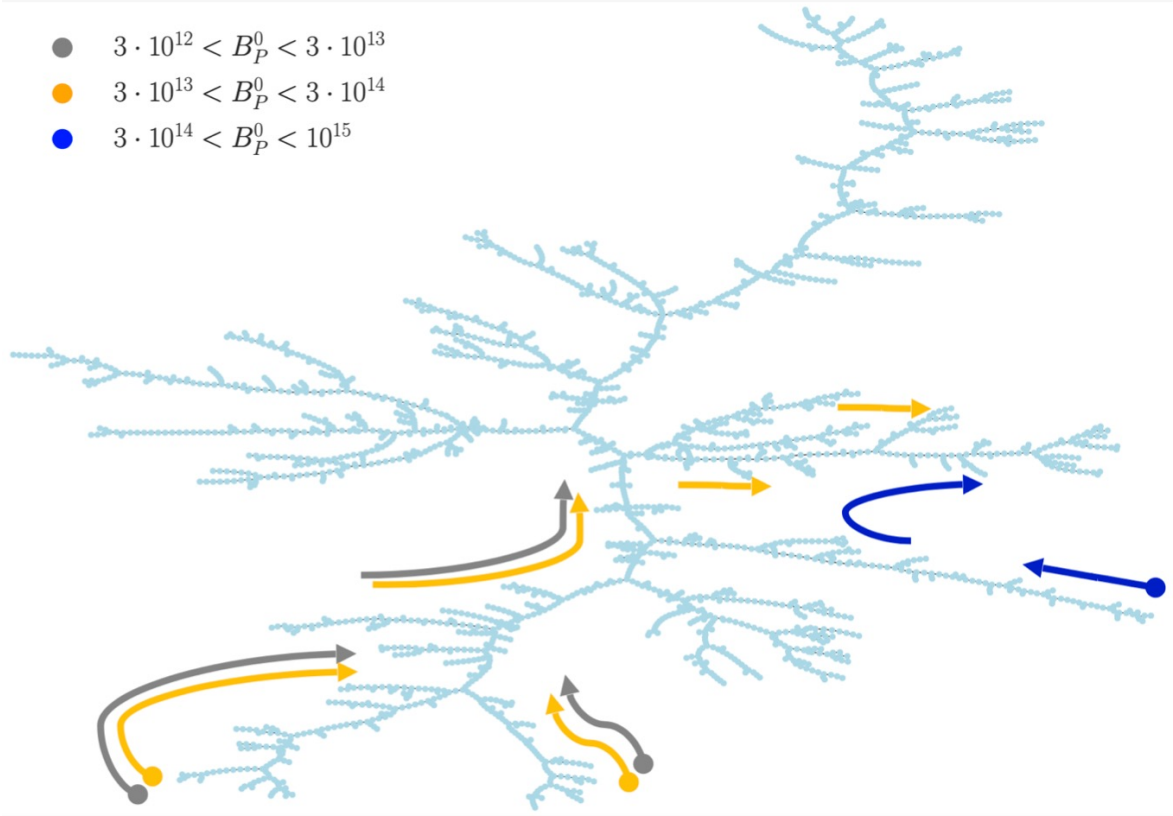


C. R. García, DFT, A. Patruno, MNRAS 2022

The pulsar tree as a descriptive tool



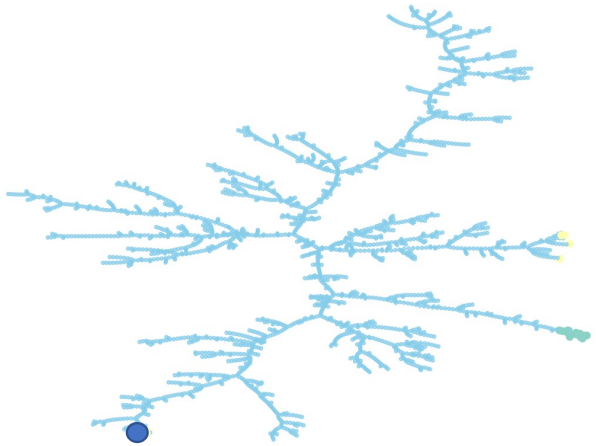
Evolutionary tracks in the MST



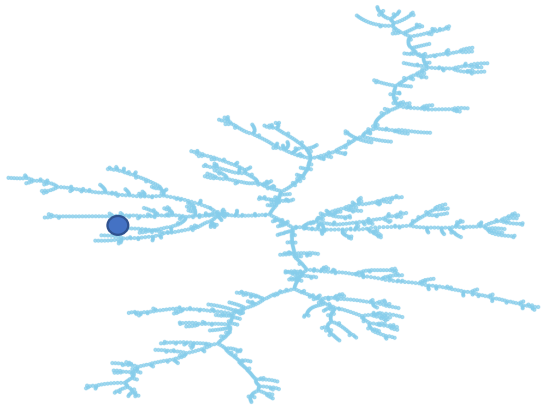
C. R. García, DFT, A. Patruno, MNRAS 2022

MST as an alerting tool

- **Based on the location of the energetic low-field magnetars J1846-0258, J1119-6127** at the bottom part of the MST, and due as well to its nearness ranking, **other pulsars with essentially the same characteristics are noted**, in particular, **J1208-6238**. It has been suggested as a possible low-field magnetar in the literature (Clark et al. (2016)) and is second (first) in the distance ranking of J1846-0258 (J1119-6127) after J1119-6127 (followed by J1846-0258). PSRs **J1513-5908** (in the composite SNR MSH 15-52**), J1640-4631**, and **J1930+1852** follow in the ranking of J1846-0258; and J1640+4631, **J1614-5048**, and J1513-5908 do so in the ranking of J1119-6127.



- **few locations of the MST show $Blc \sim$ 100 G or beyond and no detected gamma-ray pulsars yet. These regions become of special interest for future searches.** In particular, those near J1231-5113 (which is already detected) in the MST appear to be promising potential targets. A few neighboring pulsars to the outlier J1231-5113, at the end of this branch, also show a relatively large Blc with a similar range of spin-down power, and other variables, in comparison to *Fermi*-LAT pulsars. Likewise, **PSR J1915+1616** and **J2129+1210B** at the end of the nearby branch are of interest. Again, note that both have Blc > 103 G, and E > 1033 erg s−1

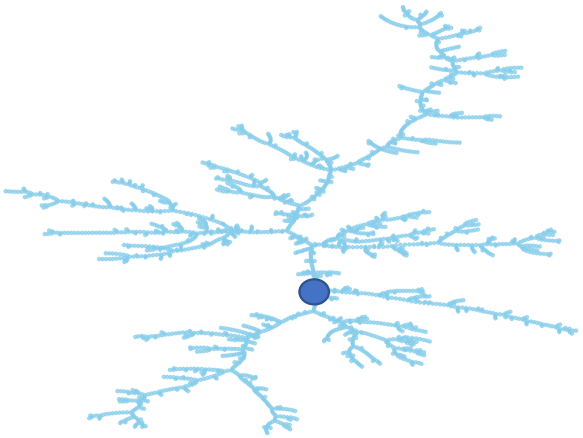


MST as an alerting tool

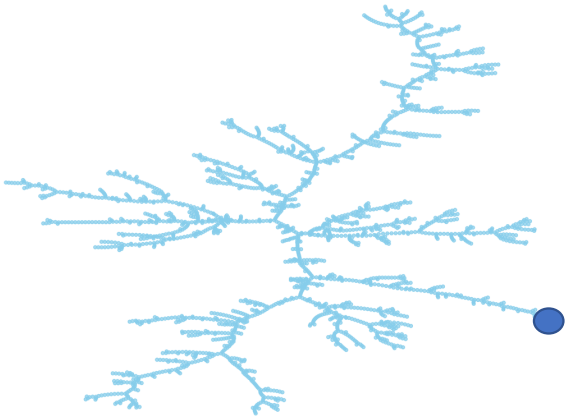
Institute of Space Sciences

EXCELENCIA MARÍA DE MAEZTU

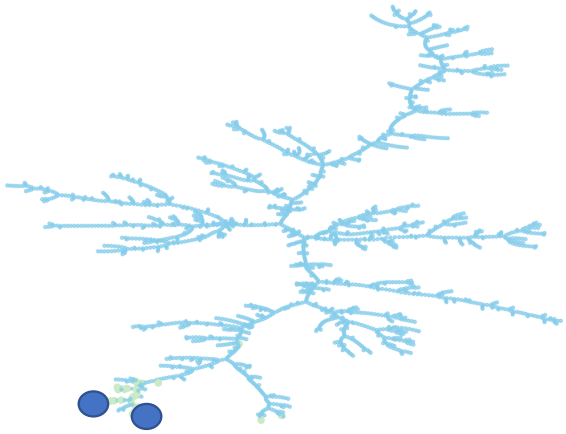
- **The detection of the radio pulsar PSR J2208+4056** by the Fermi LAT in spite of its low spin-down power has been ascribed by Smith et al. (2019) to a possible case of favorable geometry. If this is the case, it may remain indeed isolated in the MST where appears close to the main trunk. Its closest neighbours (J0532-6639, J0502+4654, and J1848-0123) call for attention in order to test this.



- Others pulsars-of-interest regarding their possible detection in gamma-rays maybe J1818-1607 and J1550-5418. These lie in the magnetar branch, where no other detected *Fermi*-LAT pulsar is located However, they **have similar properties to other pulsars already detected by *Fermi*-LAT.** The distance ranking of these two magnetars is uncommon to others: other magnetars in the branch have neither a *Fermi*-LAT pulsar nor other pulsars located in high-density areas of *Fermi*-LAT detections in the first positions of their distance ranking.



MST as an alerting tool



- The pulsars J1842-0905 and J1457-5902, and J1413-6141 and J1907+0631 are the closest neighbours to the pulsar wind nebulae J1745-3040/PWN B1742-30(1) and J0007+7303/PWN CTA 1, respectively. The latter are somewhat outliers of the pulsar wind nebulae population and thus their **neighbours are of interest to test whether this region of the pulsar parameter space is prone to producing observable nebulae**.

- Finally, the **higher-degree nodes** –particularly those connected with the main trunk, the nodes that are extremes of significant branches, and possibly other topologically selected nodes, are suggested for individual study, as they may have undiscovered relevance.

Play with it yourself:

http://www.pulsartree.ice.csic.es

Play with the pulsar tree web

http://www.pulsartree.ice.csic.es/

Institute of Space Sciences

EXCELENCIA MARÍA DE MAEZTU

THE PULSAR TREE

Visualizing the pulsar population using graph theory

Labels Left-panel

Labels Right-panel

X-Axis

Spin Period

Y-Axis

Spin Period derivative

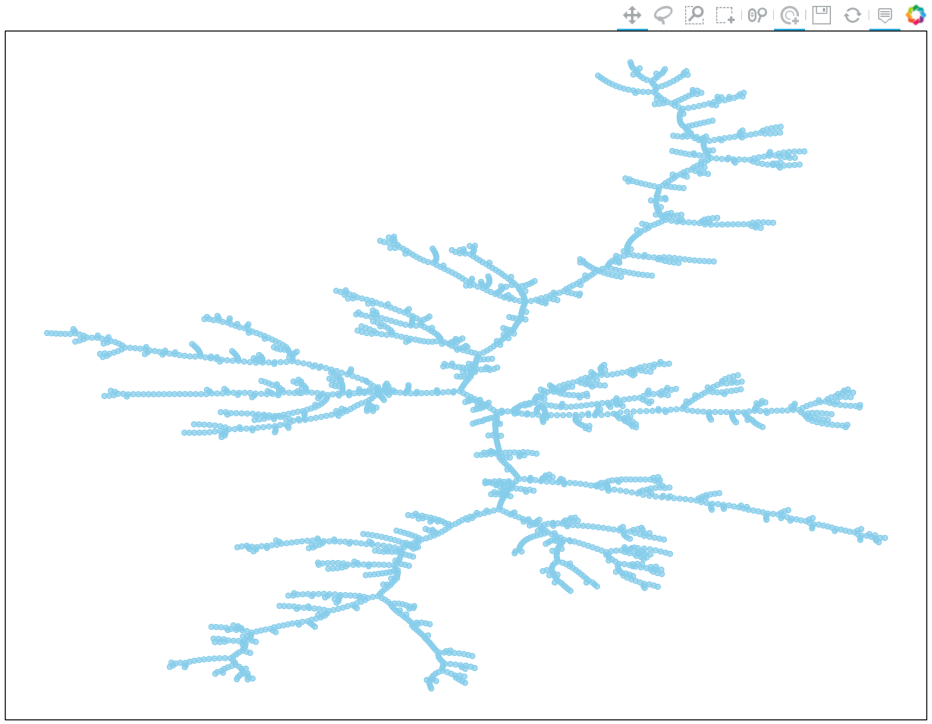
Characteristics

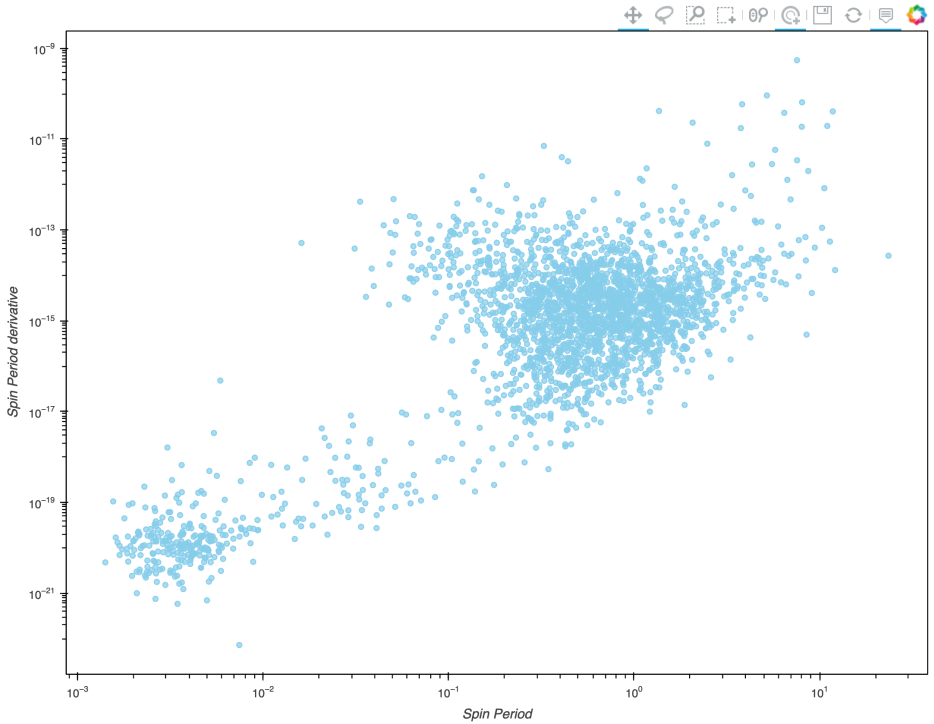
None

- Magnetars
- Low Field Magnetars
- XDINs
- RRATs
- Binaries
- Fermi
- Black widows
- Red backs
- tMSPs
- Nulling and intermittent pulsars
- PWN
- Glitches

Pulsar finder

Range finder





Selected pulsars

Download

Ranking for a specific pulsar

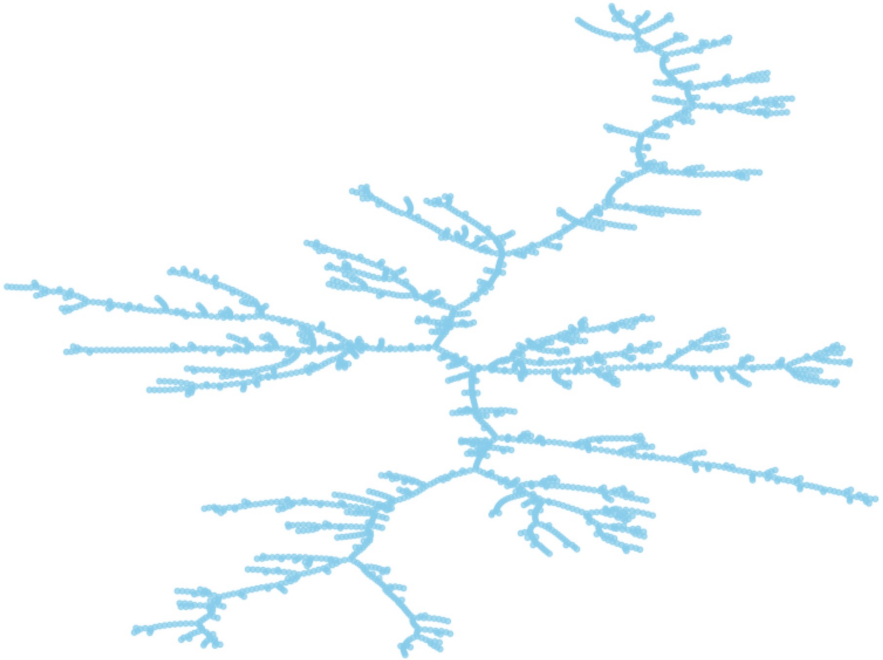
Download

Conclusions

A fresh look may change perspectives

Institute of Space Sciences

EXCELENCIA MARÍA DE MAEZTU



Credit: C. R. García, D F. Torres, and A. Patruno (Institute of Space Sciences)

The Pulsar Tree

Studying small objects at large distances is an incredible challenge, but that's what faces scientists who study neutron stars. Neutron stars are only just about 10 kilometers (6 miles) across, located at distances of quadrillions of kilometers or more. How can they even be found? Fortunately, neutron stars have ways of making their presence known. Many neutron stars have strong magnetic fields, and spin rapidly, so that the radiation generated near the magnetic poles pulses on and off as the neutron star spin. These pulses can reveal the neutron star as a compact, high density, fast-spinning, radio or X-ray pulsar. The pulse period makes neutron stars highly precise celestial clocks, useful for deep-space navigation and detection of gravitational waves from merging supermassive black holes and other massive objects. But the pulses are not precisely stable, since forces near the neutron star can slow down (or speed up) the star's rotation. These small changes in period actually provide scientists with important clues about the age of the neutron star, and can be used to classify pulsars, determine how they age and evolve, and ultimately provide a deeper understanding of the weird matter inside them. Researchers are continuously searching for news ways to understand the connections between the known neutron star pulsars and their fundamental properties. The image above is the "pulsar tree", an innovative way to explore pulsars based on a set of 8 intrinsic pulsar properties. Each branch of the tree separates pulsars into distinct groups with given properties, emphasizing the evolutionary connections in the known pulsar population. By climbing the pulsar tree, astrophysicists gain a clearer view of the connections between pulsar spin periods, period changes, pulsar ages, magnetic fields, compositions and other important properties.

Published: July 25, 2022

A fresh look may change perspectives

- The MST approach offers a different visualization of the pulsar population and may lead to new applications
  - analysis of the MST regarding clustering, centralization, betweenness, and closeness can illuminate physical connections and link pulsars among themselves.
  - The MST can be used to tell –due to the appearance of new branches- when new sub-classes of pulsars appear.
  - Similarly, a quantitative comparison of MSTs constructed for synthetic populations of pulsars offers a way to qualify the goodness of the population synthesis model that created them.
- The method used here can also be generalized to consider other variables in the distance definition. For instance, for binary pulsars, environmental and orbital parameters (DM, orbital period, orbital size, nature of the companion, etc.) can be used.
- Another possibility would be to consider distance definitions containing spectral parameters in a given energy range, or light curve properties. The forthcoming third *Fermi*-LAT pulsar catalog may be especially appropriate for working in this direction, we are testing it already with 2PC

Thank you

https://sites.google.com/view/dft-research

@dft_research

GOBIERNO DE ESPAÑA
MINISTERIO DE CIENCIA E INNOVACIÓN

CSIC
Institute of Space Sciences

EXCELENCIA MARÍA DE MAEZTU

ICREA