How to generate/replicate Figure 2:

Figure 2. Comparing PPI predictions of network shape intelligence methods.

Running the code

To run the code, execute the function *run_figure_2* with one of these options:

- Option 1: 1 to generate item with existing results. Usage: *run_figure_2(1).*
- Option 2: 2 to recreate item from original data, involving all required computations. Usage: *run_figure_2(2).*

Here is an overview of the execution of each option. The execution times reported below are measured executing the code in Windows 10 Pro with 256 GB RAM, and AMD Ryzen Threadripper PRO 3995WX 64-Cores CPU with 2.70 GHz. The software environment is MATLAB 2019a.

Option 1: Total execution time negligible

Runs *plot_figure_2* to generate item with existing results located in data folder. In data folder, there is:

- original_data: contains the raw data downloaded from their source.
- matrix: contains the adjacency matrix of each PPI network.
- sparsified_matrices: contains the sparsified networks with 10% of their links removed at random.
- network_similarities: contains CH similarity scores matrices of each sparsified network.
- pvalues_aup: contains the p-values of the permutation test for all methods in the AUP evaluation.
- script: contains *plot_figure_2* function.

Option 2: Total execution time ~**9h**

All the results of the following scripts are stored in the directory data_replicated. Below are the different steps to implement the computation:

- In order to download the data, follow the instructions reported in the file *instructions to download the data* located inside the folder "data".

- *create_[NAME PPI NET]* : creates the adjacency matrix for each 15 PPI networks. Outputs in "matrix". Total execution time: **~7 min**.

- *run_link_removal_10*: Remove at random 10% of the total links in the network. Outputs in "sparsified_matrices". Total execution time: **~50 s**.

- *run_simulation_linkrem10_CHv2_L2_L3*: For each simulation (sparsified network) of a network, computes the CH similarity scores for all the methods. Also, computes the AUP evaluation for all the methods. Outputs in "network_similarities/CH_L2_L3/results". Total execution: **~9 h**.

- *compute_pvalues_aup_top100*: computes the p-values of the permutation test for all methods in the AUP evaluation. Outputs in "pvalues_aup". Total execution time: ~**1 min**.

- *plot_figure_2*: Plot Figure 2. Total execution time: **negligible time**.

Note: running option 2 for multiple independent times might generate results slightly different because of the random link removal process performed by the function *run_link_removal_10*.