Energy Landscape Analysis Toolbox (ELAT) User's Guide (ver. 3.2)

Takahiro Ezaki (https://sites.google.com/site/ezakitakahiro/cv)

Naoki Masuda (http://www.naokimasuda.net/index.html)

Latest update: 27 Dec 2022

Table of Contents

Introduction: What is the energy landscape analysis?

Thank you for considering using the energy landscape analysis toolbox (ELAT). The energy landscape analysis is a computational method that enables intuitive interpretation of multivariate time series. This analysis comprises four steps (Fig. 1.): (1) Binarization of the data, (2) estimation of the maximum entropy model (Boltzmann distribution), (3) construction of a disconnectivity graph and basin of energy local minimums, and (4) computation of dynamics measures on the energy landscape. This method was originally designed for analyzing fMRI data, but it is in principle applicable to other types of data. In our experience, the energy landscape analysis works nicely when the number of variables is roughly 6 to15. For more variables, the computational cost becomes large and interpretation of the results becomes difficult. In such cases, we recommend reducing the number of variables (consider using ICA, merging variables, etc.).

For detailed algorithms of the energy landscape analysis, please refer to Ref [1]. Examples of applications of this analysis are found in Refs. [1-7]. If you have any questions or find an error, please contact ezaki0705@gmail.com.

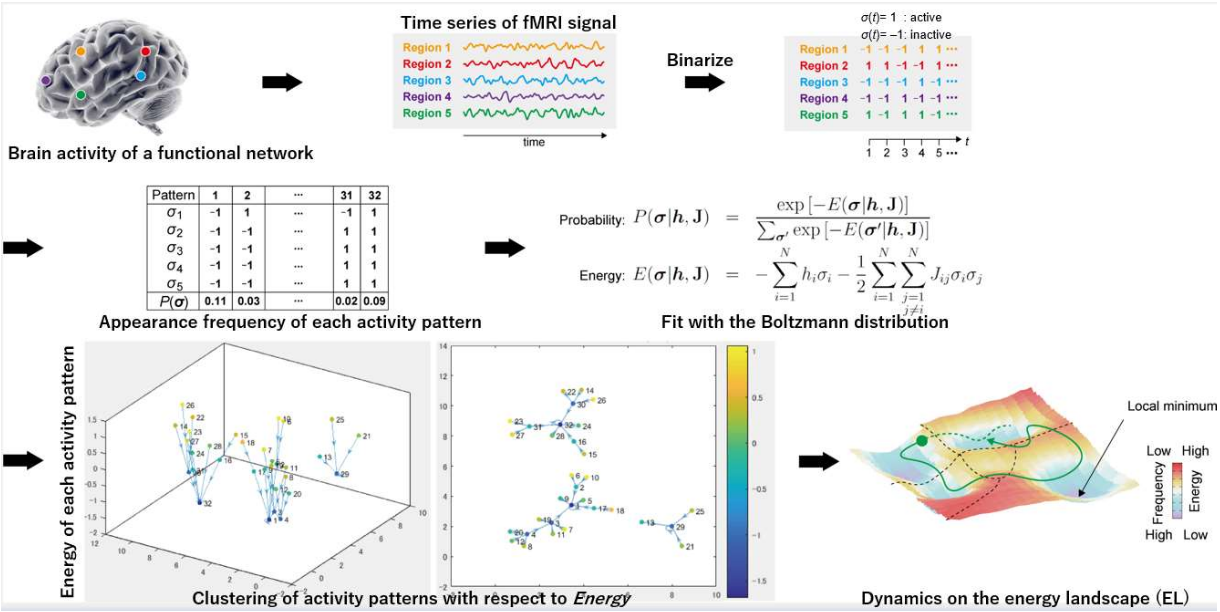

Fig. 1. Overview of the energy landscape analysis.

This toolbox performs the entire computations necessary for the energy landscape analysis (Fig. 2). The Full Analysis option computes everything including the dynamics measures of individuals. The Energy Landscape Construction option computes the energy landscape and does not perform further analysis.

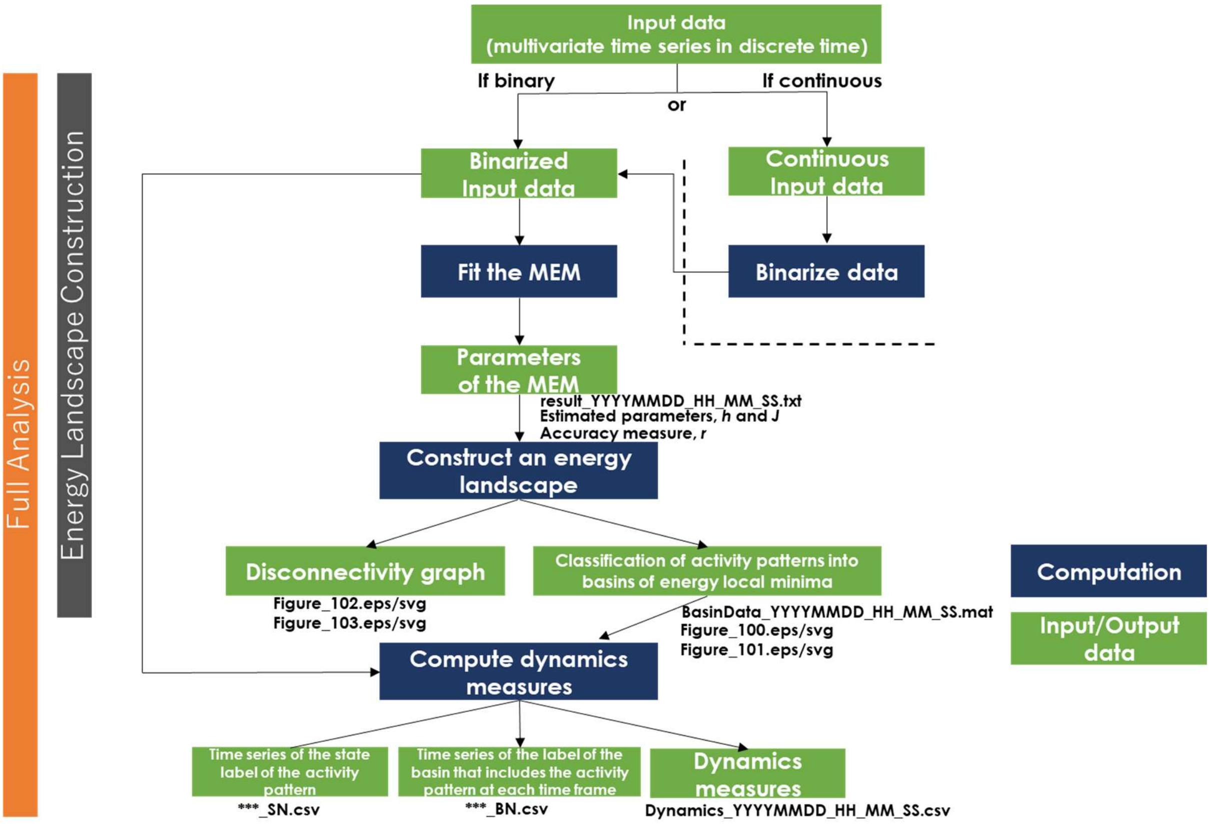

Fig. 2. Overview of ELAT.

Technical details of the energy landscape analysis

We denote the input data by $x_i(t)$, where i ($i = 1, \ldots, N$) labels a variable and t ($t = 1, \ldots, t_{max}$) represents time.

1. Binarization

The input data for the energy landscape analysis must be binary (i.e., +1 / -1). Thus, if your data $\{x_i(t)\}$ take continuous values, you have to binarize them with an appropriate threshold. In our previous studies [1,7], we binarized our fMRI data by setting a threshold to the average of the signal for each variable, i.e., $Y_i(t) = 1$ (if $x_i(t) > \langle x_i \rangle$); $Y_i(t) = -1$ (otherwise), where $\langle A \rangle$ denotes the time average of A. Because the baseline of the signal might be different across individuals, we recommend that you separately perform binarization for each individual and each variable. This procedure is implemented in this toolbox and will be explained later.

2. Labeling the states

The state of the system at time t is represented by a binary vector $(Y_1(t), Y_2(t), \ldots, Y_N(t))$. This state vector takes one of the 2^N states. For convenience, we enumerate the state with a simple conversion from binary to decimal:
$s = 1 + \sum 2^i \ (Y_i + 1) / 2$. The function $(Y_i + 1) / 2$ maps –1 and1 to 0 and 1, respectively. This mapping transforms, for example,

(–1, –1, –1, –1) to 1

(1, 1, 1, 1) to 16

(1, –1, 1, –1) to 6.

Thus, the binarized states are now labeled from 1 to 2^N.

3. Accuracy of fitting

The accuracy of the fitting is measured by r, which is output on the console. The definition of this measure is found in Ref. [1]. When r is close to 1, it means the MEM fitting is successful. We recommend reporting this value in your work. Note that this value inevitably becomes small when the data is short [1].

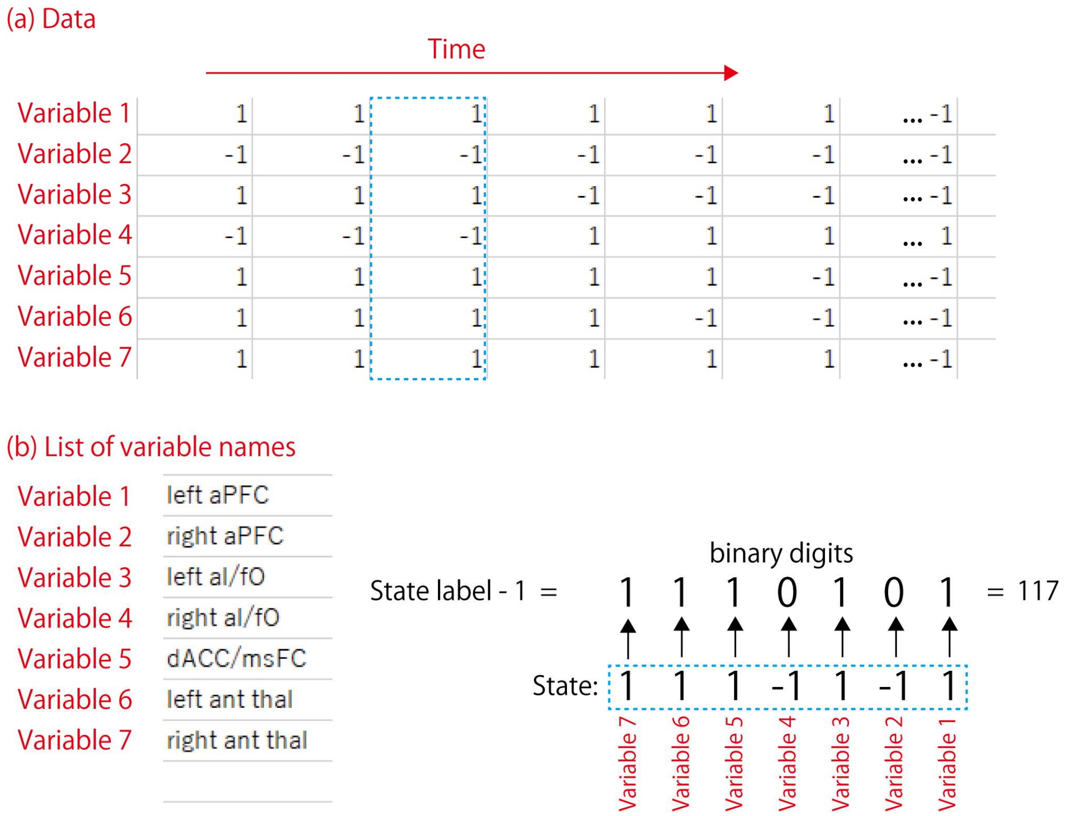

Fig. 3. Input data and the definition of the numerical label of the state.

How-to 1: A tutorial

Here we show how the toolbox works. (See Fig. 4.)

1. Prepare data

First, prepare the data in a file with extension ".dat". The file must have *N* (number of variables) rows and t_{max} columns (see Fig. 3). Each value must be separated by a tab. See "testdata.dat" contained in the toolbox folder for an example. The values in the input data can be either continuous or binarized. If the data is continuous, ELAT binarizes the data by thresholding based on the average of each variable (see 5.). If you do not have the data and want to see how it works, please use "testdata_1.dat", …, "testdata_4.dat" in the folder.

2. Launch the toolbox

Unzip the toolbox and place it somewhere convenient for you. Open "StartProgram.m" with MATLAB, and run int. Older versions of MATLAB might not support some of the functions used, so please update your MATLAB to the latest version (in many cases, updating is available for free).

3. Select type of analysis

Here, simply select "Full Analysis" (default).

4. Select input data

Select the data you prepared in 1. Multiple files can be selected. If more than one file is selected, a single energy landscape is computed by concatenating the files, but dynamics measures are computed for each file based on the estimated energy landscape. If you want to use the test data, select "testdata_1.dat", …, "testdata_4.dat" in the folder.

5. Select binarization option

Select "Binarized data" or "Continuous data." If you select "Continuous data" and wish to use the average of ***x*** as the threshold for binarization, please leave the "threshold" 0 (default). If you change this value to ***x***, the threshold will be set to **average +** ***x***. If you selected the test data, please select "Binarized data."

6. Select Basin Data Generation option

Here, just go with "Construct energy landscape from input files." We will explain the "Read basin data" option later (p. 12).

7. Select variable name (ROI name)

If you have a list of the names of the variables, please select it here. The format of the file must be a ".dat" file with *N* rows and 1 column. This is optional. For an example, see "roiname.dat" in the folder.

8. Select output folder

Select the output folder. We recommend that you create a new folder dedicated to this analysis. If you want to see the list of the energy landscape basin, please also check "Save Basin List."

9. Execute

Execute!

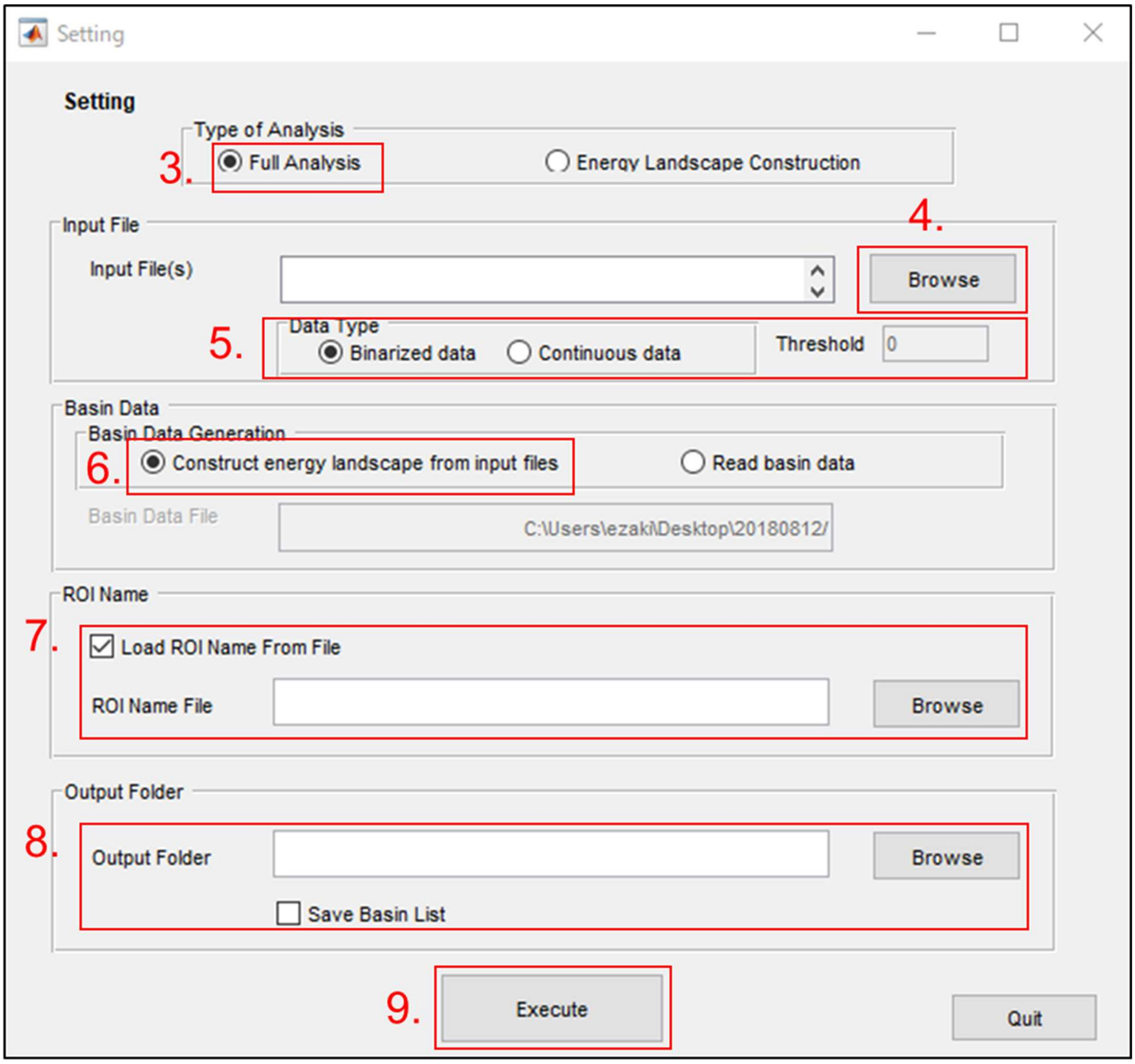

Fig. 4. Configuration window.

,

How-to 2: Results

1. Energy landscape

As a result, first, you will get four figures. The first two figures visualize the energy landscape (Fig. 5). Each node represents a state specified by its numerical label. Each edge visualizes the edge between neighboring states which have the largest energy difference (i.e., the edge on the steepest path from an arbitrary state to the local minimum of the energy). The absolute positions of the nodes are not important. Each isolated cluster of nodes is a “**basin**” of each local minimum of the energy. The energy value of each state is shown with a color, which is also visualized by the height in the 3D version of the figure (shown to the right in Fig. 5). The states with a small energy value are considered to appear frequently, and thus they should be important states in many cases.

The two figures are saved in the output folder.

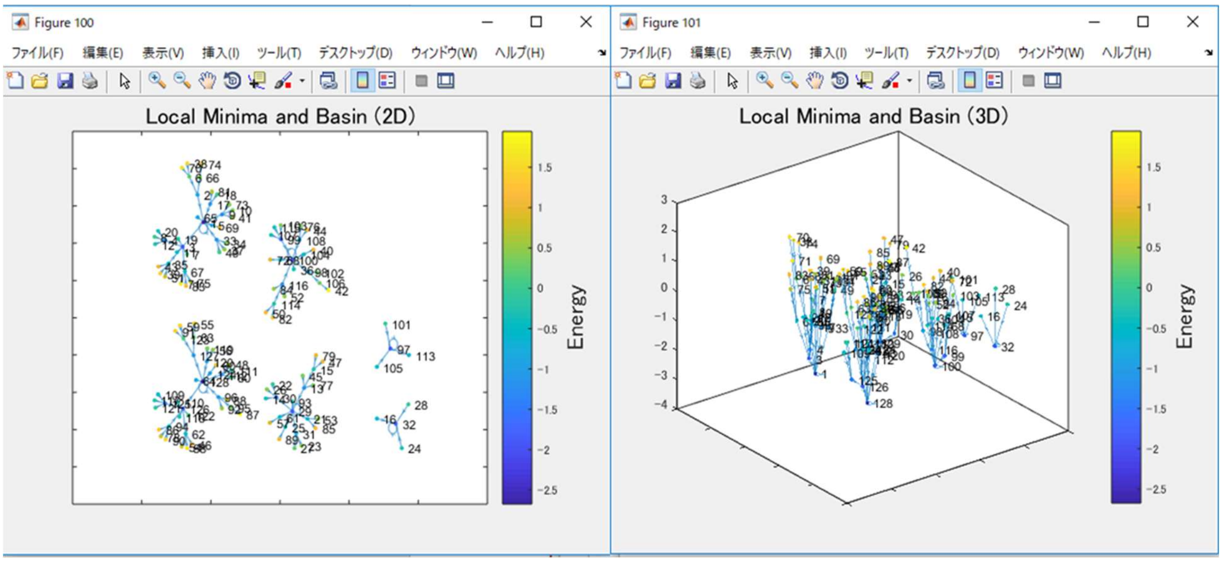

Fig. 5. Output figures 1 and 2.

The other two figures give a more abstract representation of the energy landscape. The disconnectivity graph (Fig. 6 left) shows the energy values of the local minimum states (specified in Fig. 6 right) and the height of energy barriers between them. If you perform this analysis for more than one group, separately for each group, this disconnectivity graph is useful for comparing the groups. Note that the information about the basin is dropped here.

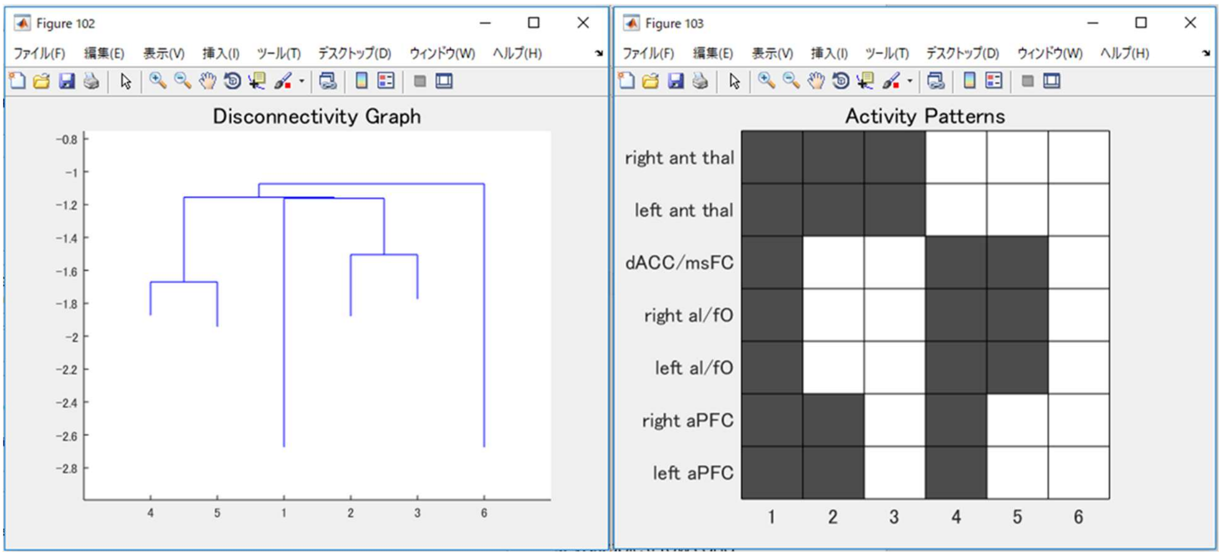

Fig. 6. Disconnectivity graph (left) and activity patterns (states) of each energy local minimum (right). The numbers on the x-axis labeling the energy local minimum states are consistently used in both panels. White and black cells in the right panel represent that variables (ROIs) are active (+1) and inactive (–1), respectively. Note that the numerical labels of the energy local minimum states in these figures are NOT the same as those used in Figs. 3 and 5.

2. Time series

In the output folder, csv files are created for each input file.

"*_SN.csv" (***: input file name):**

The time series of the numerical label of the state between 1 and 2^N, ***s(t)***, is saved.

"*_BN.csv":**

The time series of the numerical label of the basin to which ***s(t)*** belongs, denoted by, ***b(t)***, is saved.

3. Dynamics measures

In the output folder, a csv file, "Dynamics_yyyymmdd_HH_MM_SS.csv" is created. The dynamics measures are computed for each input file.

InputFile:

The name of the data file.

Frequency of *Bj*:

The fraction of time in which ***b(t)*** = ***j***.

Direct transition from *Bj* to *Bk*:

The number of direct transitions form basin ***j*** to basin ***k***, divided by t_{max}. Indirect transitions, e.g.,

$j \rightarrow m \rightarrow k$ $(m \neq j, k)$

are excluded.

Transition from *Bj* to *Bk*:

The number of transitions from basin ***j*** to basin ***k***, divided by t_{max}. Indirect transitions, e.g.,

$j \rightarrow m \rightarrow k$ $(m \neq j, k)$

are included.

Other options

1. Compute dynamics measures on other energy landscapes

In the tutorial, we constructed an energy landscape from the input data and computed dynamics measures on it. By choosing the “Read basin data” option, you can compute the dynamics measures based on a different energy landscape constructed from other input data. When you run the Full Analysis, the basin data is saved as “BasinData_YYYYMMDD_HHMMSS.mat” in the output folder.

If you perform a two-group study and separately run Full Analysis for each group, the difference in the dynamics measures strongly reflects the difference in the energy landscapes. Thus, if you want to compare dynamics of two groups, we recommend that you confirm that the number of states in each basin is similar between the two energy landscapes constructed separately for each group, or compute dynamics measures on the same energy landscape that is, for example, constructed by using all the participants in both groups.

2. Basin data

“BasinData_YYYYMMDD_HHMMSS.mat” contains a variable called “BasinGraph.” The first column is the numerical label of the state. The second column shows the neighboring state that has the smallest energy value. The connection between the state in the first column and that in the second column is shown as edges in Fig. 5. The third column is the local minimum state that the state belongs to.

References

[1] T. Ezaki, T. Watanabe, M. Ohzeki, and N. Masuda, "Energy landscape analysis of neuroimaging data," Phil. Trans. R. Soc. A 375, 20160287 (2017).

[2] T. Watanabe, N. Masuda, F. Megumi, R. Kanai, G. Rees "Energy landscape and dynamics of brain activity during human bistable perception," Nat. Commun. 5:4765 (2014).

[3] T. Watanabe, et al. "Energy landscape analysis of resting-state brain networks," Front. Neuroinform. 8:12 (2014).

[4] A. Ashourvan, S. Gu, M.G. Mattar, J.M. Vettel, D.S. Bassett "The energy landscape underpinning module dynamics in the human brain connectome," Neuroimage 157:364–380 (2017).

[5] J. Kang, C. Pae, H.J. Park "Energy landscape analysis of the subcortical brain network unravels system properties beneath resting state dynamics," Neuroimage 149:153–164 (2017).

[6] Watanabe T, Rees G "Brain network dynamics in high-functioning individuals with autism," Nat. Commun. 8:16048 (2017).

[7] T. Ezaki, M. Sakaki, T. Watanabe, N. Masuda "Age-related changes in the ease of dynamical transitions in human brain activity," Hum. Brain Map. 39, 2673–2688 (2018).