Even though ICA guarantees a unique solution, given the noisy nature of the data and the fact that there are finite samples, ICA algorithms that are of iterative type will produce slightly different solutions depending on the initialization that is used.

Hence, it is important to use a scheme for selecting the “best run”, i.e., the most representative and reproducible one among a number of multiple runs for a given ICA algorithm. An important additional note is that this should not be interpreted as the desirability of using an algorithm that provides the same solution each time as most often these yield suboptimal solutions, and the more flexible ICA algorithms are the ones that are likely to yield slightly different solutions at each run.

GIFT provides three methods to select the most stable/ reliable ICA solution.

1. ICASSO

ICASSO implements a clustering approach to cluster components across different ICA runs followed by identification of qualified clusters that have a cluster size within a pre-defined range and have a quality index above a pre-defined threshold. The original implementation of ICASSO [1] selects a single centrotype for each cluster as a reliable estimate for that cluster leading to loss of information when more than one type of component is grouped into the same cluster. Hence, in [2] the authors propose a method based on ICASSO to select the most stable run. Using only the qualified clusters, the most stable run is selected as the run with highest average maximal intracluster similarity, *i.e.*, the run including the components that are close enough to all centrotypes within the qualified clusters. GIFT implements the version proposed in [2] to select the stable run.

1. Minimum spanning tree (MST) [3]

MST aligns the components across multiple ICA runs using the linear assignment problem. The minimum cost of alignment and the corresponding alignment for each pair is computed using the Hungarian algorithm followed by identifying the central run as the run that has minimum cost of alignment. The components in each run are reordered as per the central run. After alignment, a one-sample *t*-test is performed across runs in order to investigate the reliability of the estimated components. The best run is selected as the run with highest correlation between the components and the corresponding *T*-maps.

1. Cross inter-symbol interference (ISI) [4]

Cross ISI is the fastest method to find the most consistent run. Cross ISI measures the distance between a pair of ICA solutions. For each run, cross ISI is computed between that run and all the other remaining runs. The most consistent run is selected as the run with highest average cross ISI. The selected run agrees with the run selected by MST and ICASSO in a number of scenarios and provides a better solution than MST and ICASSO in other scenarios [4].

References

[1] Himberg, Johan, Aapo Hyvärinen, and Fabrizio Esposito. "Validating the independent components of neuroimaging time series via clustering and visualization." *Neuroimage* 22, no. 3 (2004): 1214-1222.

[2] Ma, Sai, Nicolle M. Correa, Xi-Lin Li, Tom Eichele, Vince D. Calhoun, and Tülay Adali. "Automatic identification of functional clusters in FMRI data using spatial dependence." *IEEE Transactions on Biomedical Engineering* 58, no. 12 (2011): 3406-3417.

[3] Du, Wei, Sai Ma, Geng-Shen Fu, Vince D. Calhoun, and Tülay Adali. "A novel approach for assessing reliability of ICA for FMRI analysis." *IEEE International Conference on* *Acoustics, Speech and Signal Processing (ICASSP)*, pp. 2084-2088, 2014.

[4] Long, Qunfang, Chunying Jia, Zois Boukouvalas, Ben Gabrielson, Darren Emge, and Tülay Adali. "Consistent run selection for independent component analysis: Application to fMRI analysis." *IEEE International Conference on* *Acoustics, Speech and Signal Processing (ICASSP),* 2018.