**Introduction to Spectral Metrics in Biological Network Theory**

Roberto Visintainer, Giuseppe Jurman, Marco Grimaldi, Cesare Furlanello

Fondazione Bruno Kessler, Trento, Italy

---

**Introduction:**

The basic goal of network comparison is quantifying difference between two homogeneous objects in some network space. Although fruitful insights can be drawn by indirect methods (for instance comparing graph density, connectivity or communities), a distance must be employed whenever a quantitative assessment of the differences between two elements is required. A few global distances have been proposed to such extent; among them, the family of spectral measures are particularly relevant. As the names suggests, their definition is based on (functions of) the spectrum of one of the possible connectivity matrices of the network, i.e. the set of eigenvalues.

---

**Spectral Distances for Network Comparison**

<table>
<thead>
<tr>
<th>Node Correlation</th>
<th>Ipsen-Mikhailov</th>
<th>Biological Applications</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

**Ipsen-Mikhailov**

In general Ipsen & Mikhailov distance (D2) seems to be the metric more stable (in terms of varying graph density) and robust (less prone to odd behaviours). MCC has recently attracted the attention of the machine-learning community as one of the best method to summarize into a single value the confusion matrix of a binary classification task.

---

**Transcriptional Interaction Networks EC for Esherichia coli**

EC includes 519 interactions between TF and operons, with maximal eigenvalue \( \lambda_{max} = 73.0 \) and \( d = -\log(\text{AUC}(EC)) = 330.02 \).

---

**Distances in Network Time Series**

- The largest variations occur in the embryonal stage E.
- The dynamics of the networks move \( N_i \) away from \( N_0 \) until time points 20, then the following terms start getting closer again to \( N_0 \), in terms of distance \( d \).
- The same trend applies to MCC values, too, but to a quite smaller extent: the fact that MCC curve is descending from its maximum in the last 15 time points can be appreciated only by zooming in.
- This means that, after the embryonal stage, the network is getting structurally more and more similar again to \( N_0 \), but the number of links that match with those of \( N_0 \) is however very little.

---

**Synthetic Data**

Computing D1-D6 on three random synthetic benchmark datasets, averaging over \( b = 50 \) replicates.

---

**References**

- The dynamics of 519 genes related to the developmental process based on their genetic ontologies is observed, on 66 time points are selected during the developmental cycle: Embryonic (1-30 time point), Larval (31-40 time point), Pupal (41-58 time points), and Adult stages (59-66 time points). Hereafter, \( <N_i,N_j> \) and MCC \( (N_i,N_j) \) are computed for all 66 networks.
- Jakobson et al., 2002 Ipsen et al., 2002
- The dynamics of 519 genes related to the developmental process based on their genetic ontologies is observed, on 66 time points are selected during the developmental cycle: Embryonic (1-30 time point), Larval (31-40 time point), Pupal (41-58 time points), and Adult stages (59-66 time points).
- Hereafter, \( <N_i,N_j> \) and MCC \( (N_i,N_j) \) are computed for all 66 networks.

---


http://mpba.fbk.eu