

Benchmarking GRN inference algorithms using topological network properties

27.04.2021

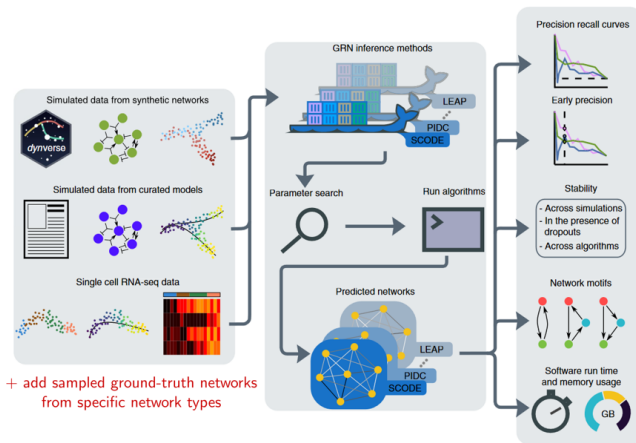
Progress report

HelmholtzZentrum münchen

Deutsches Forschungszentrum für Gesundheit und Umwelt

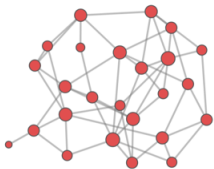
Current best practice

Pratapa et al.: *Benchmarking algorithms for gene regulatory network inference from single-cell transcriptomic data* (BEELINE)

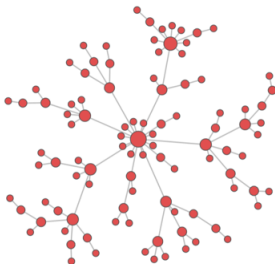


+ add analysis of
topological network
properties

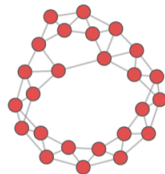
Network types



Random graph
(Erdős-Renyi)



Scale-free network
(Barabasi-Albert)



Small-world network
(Watts-Strogatz)

Network measures

Assortativity: preference for a network's nodes to attach to others that have a similar degree.

Quantified by the *assortativity coefficient*: ρ between k and k_{nn}

Degree Centrality: provide an estimate of how "centralized" a graph with a fixed number of nodes and links is

Quantified by the *Centralization*: $H = \sum_{v_i} |deg(v^*) - deg(v_i)|$

Clustering Coefficient: measure of the extend to which nodes in a graph tend to cluster together.

Quantified by the *local clustering coefficient*: $C_v = \frac{L_v}{2 * k_v * (k_v - 1)}$

and the *global clustering coefficient*: $\langle C \rangle = \frac{1}{n} \sum_{i=1}^n C_{v_i}$

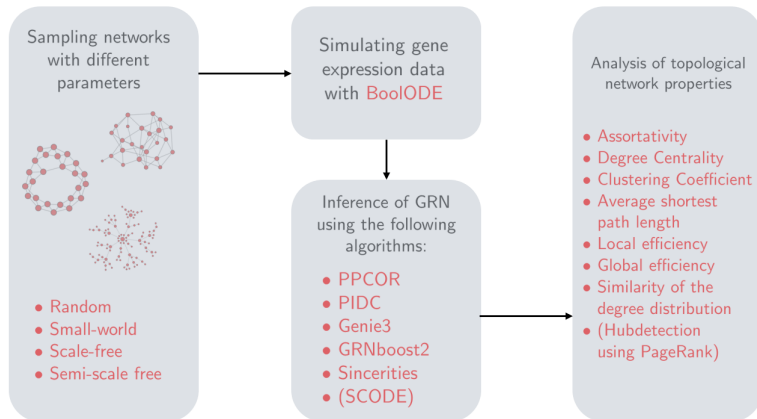
Network measures

Average shortest path length: average shortest path length between two random nodes: $a = \sum_{v,w} \frac{d(v,w)}{n \cdot (n-1)}$

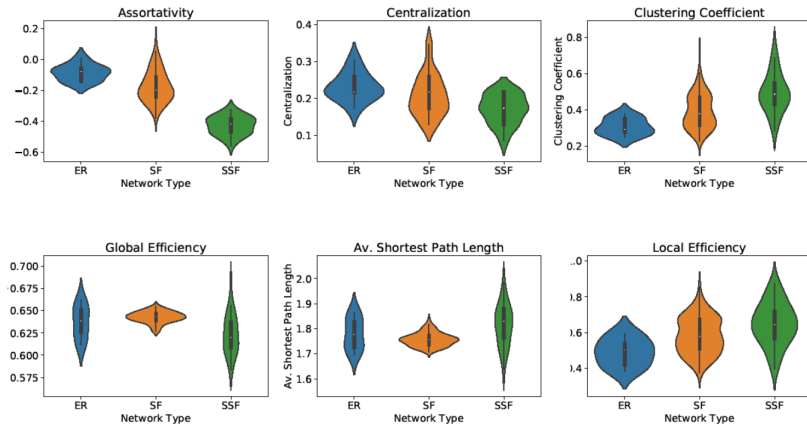
Global efficiency: efficiency of information exchange in the network
 $E_{glob}(G) = \frac{E(G)}{E_{ideal}}$ where $E(G) = \frac{1}{n \cdot (n-1)} \sum_{v \neq w} \frac{1}{d(v,w)}$

Local efficiency: resistance of the network to failure on a small scale
 $E_{loc}(G) = \frac{1}{n} \sum_v E(G_v)$

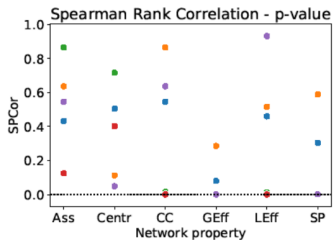
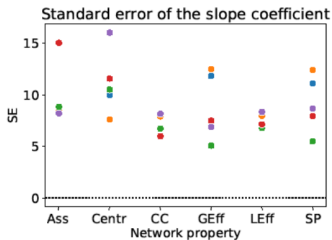
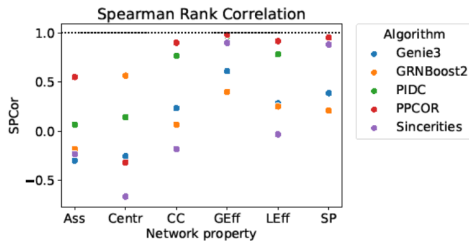
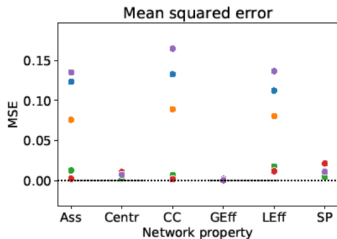
Adjusted pipeline



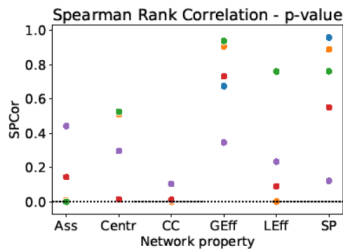
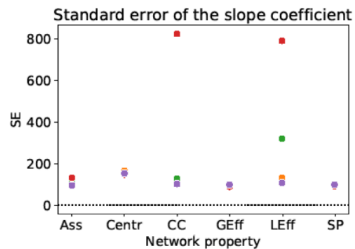
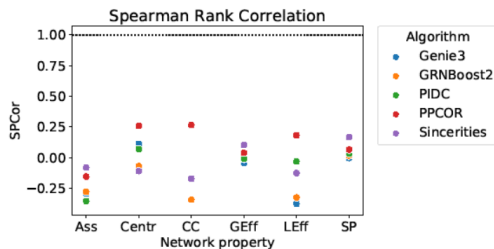
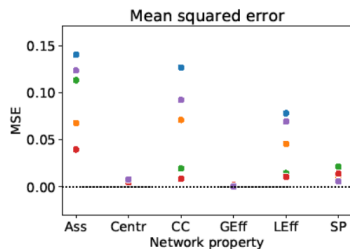
Properties of ground-truth networks



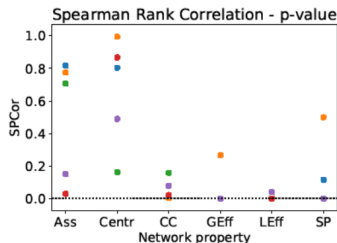
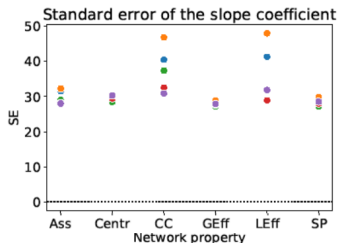
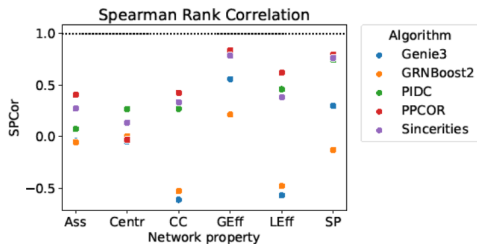
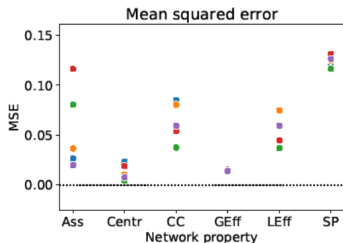
Random networks



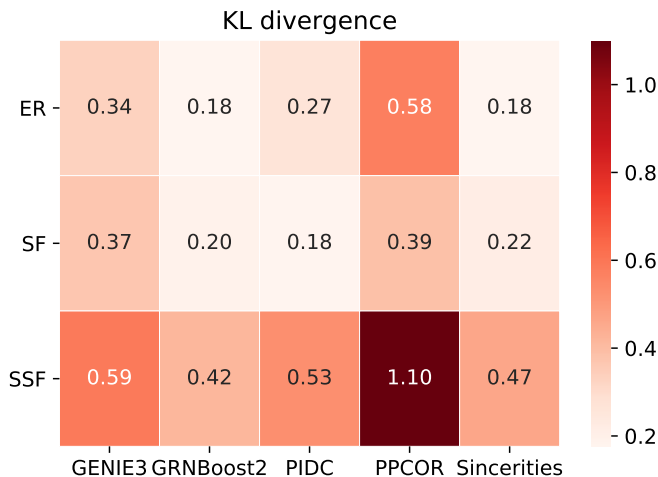
Scale-Free networks



Semi-Scale-Free networks



Similarity between the degree distributions



Next steps

- Finish analysis on synthetic data
 - Investigate the effect of dropouts and inhibiting/activating edges
 - Fix SCODE and Small-world networks
 - Analyse the effect of network parameters
 - Hub analysis with PageRank
- Transfer analysis to experimental data