



HAL
open science

An Empirical Investigation of Statistical Backbone Filtering Techniques

Ali Yassin, Hocine Cherifi, Hamida Seba, Olivier Togni

► **To cite this version:**

Ali Yassin, Hocine Cherifi, Hamida Seba, Olivier Togni. An Empirical Investigation of Statistical Backbone Filtering Techniques. Conference on Complex Systems, Oct 2022, Palma de Mallorca, Spain. hal-04321035

HAL Id: hal-04321035

<https://hal.science/hal-04321035>

Submitted on 4 Dec 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

An Empirical Investigation of Statistical Backbone Filtering Techniques

Ali Yassin

University of Burgundy, FRANCE
ali_yassin@etu.u-bourgogne.fr

Hocine Cherifi

University of Burgundy, FRANCE
hocine.cherifi@u-bourgogne.fr

Hamida Seba

University of Lyon, FRANCE
hamida.seba@univ-lyon1.fr

Olivier Togni

University of Burgundy, FRANCE
olivier.togni@u-bourgogne.fr

What is a Network Backbone?

Filtering Noise?
Preserving Topological Properties?

No Universal Definition!

Types of Backbone Filtering Techniques

1. Structural Backbone Filtering Techniques.
2. Statistical Backbone Filtering Techniques.

Statistical Backbone Filtering Techniques

- **Noise Corrected Filter**

edge weights are drawn from a *binomial distribution*. However, it estimates the probability of observing a weight connecting two nodes using a Bayesian framework.

- **Marginal Likelihood Filter**

each unit edge randomly chooses two nodes, which results in a *binomial distribution*.

- **ECM Filter**

Employs a null model based on the *canonical maximum-entropy ensemble* of weighted networks having the same degree and strength distribution as the original network.

- **LANS Filter**

Employs the *empirical cumulative density function* to judge statistical significance.

- **Disparity Filter**

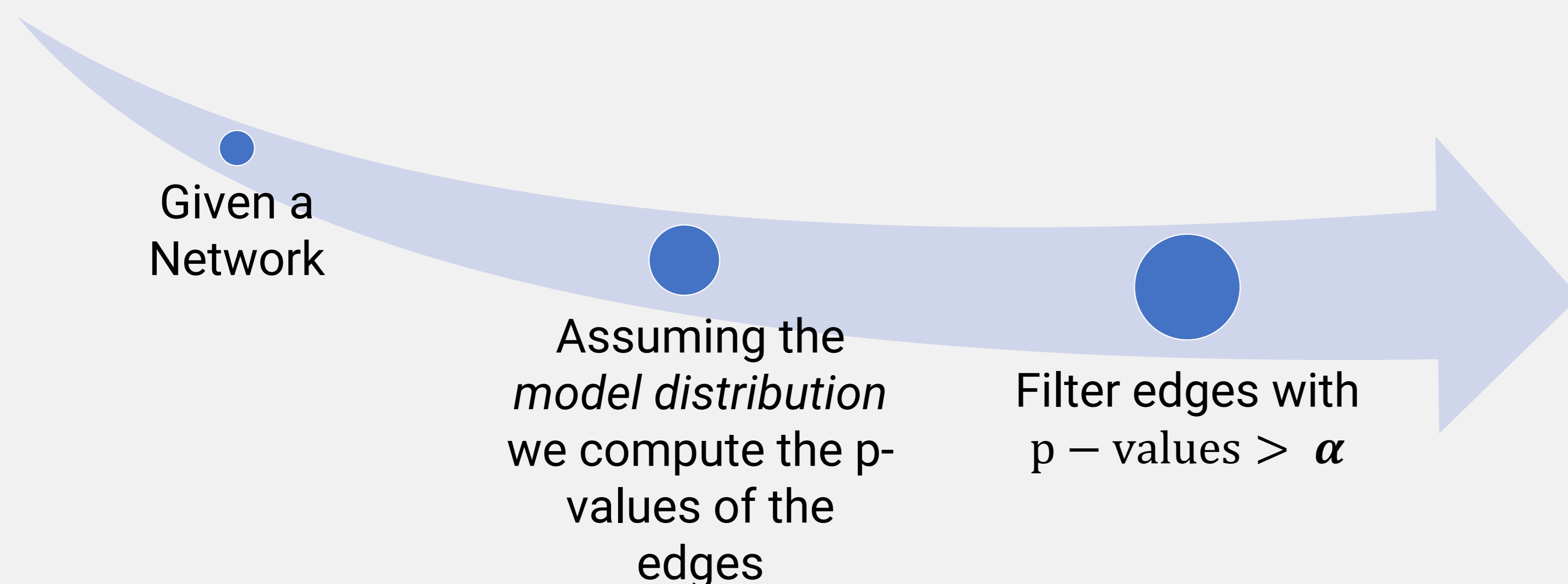
a node distributes its strength (summation of weights) *uniformly* on the edges.

- **Polya Urn Filter**

a node distributes its strength at random through a Polya process, which results in a *Beta-binomial distribution*.

- **GloSS Filter**

edge weights are randomly drawn from the *empirical weight distribution*.



Worldwide Air Transportation Network



Noise Corrected Filter



Marginal Likelihood Filter



ECM Filter



LANS Filter



Disparity Filter



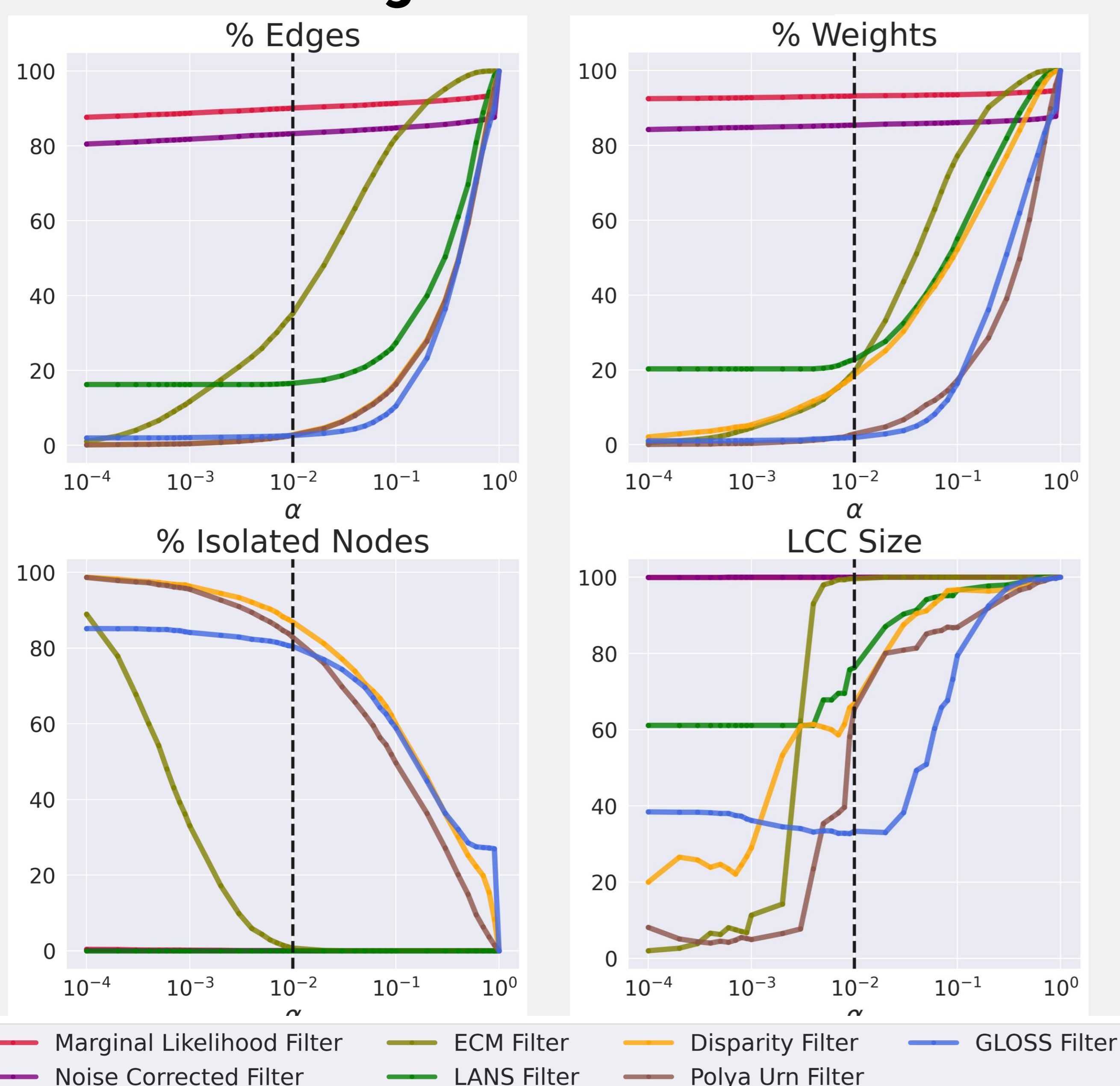
Polya Urn Filter



GloSS Filter



Evaluation Using Four Indicators



Conclusion

Filter	Technique Behavior	Giant Component	Isolated Nodes	Weights Retained
Marginal Likelihood	Conservative	Always	Low	High
Noise Corrected	Conservative	Always	Low	Small & High
ECM	Flexible	$\alpha > 10^{-2}$	$\alpha < 10^{-2}$	Small & High
LANS	Aggressive	Always	Low	High
Disparity	Aggressive	$\alpha > 10^{-2}$	High	High
Polya Urn	Aggressive	$\alpha > 10^{-2}$	High	Small & High
GloSS	Aggressive	$\alpha > 10^{-2}$	High	Small & High

What about the distributions?