Frequency-Domain Better than Time-Domain for Causal Structure Recovery in Dynamical Systems on Networks

Authors:




Mohammed Tuhin Rana[1], Mishfad Shaikh Veedu[2], James Melbourne[3], Murti Salapaka[1]

Presenter:

Mohammed Tuhin Rana, rana0082@umn.edu

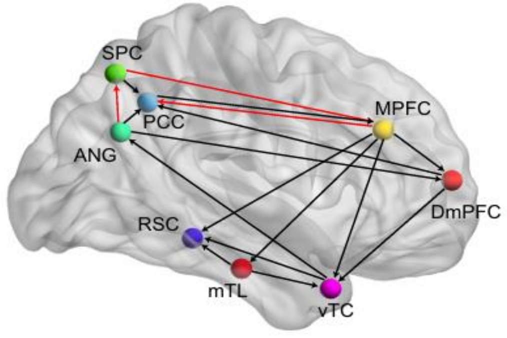
[1]*Department of Electrical and Computer Engineering, University of Minnesota -Twin Cities*

[2]*Google LLC*

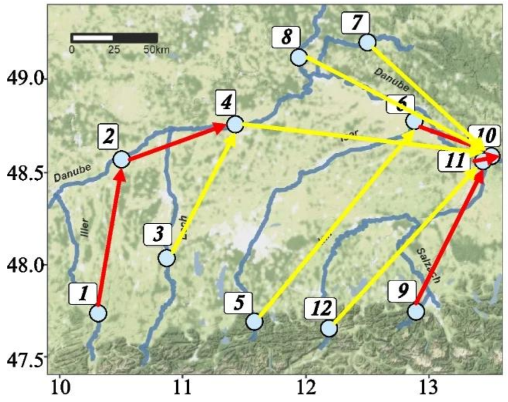
[3]*Centro de Investigación en Matemáticas*

Contents

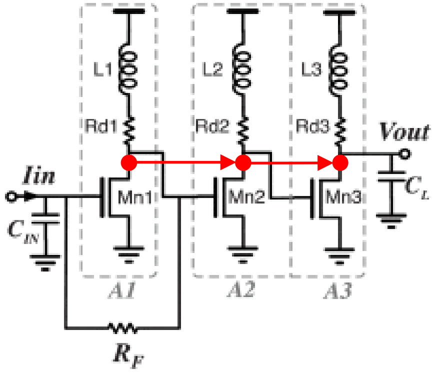
Motivation



Connectivity Structure of Brain Regions



River Flow Network



Electronic Circuits Network

- Many practical systems are described as dynamic networked systems, e.g. biological systems, electronic circuits, geographical systems, chemical reaction networks.
- Time-domain (TD) based causal discovery methods are computationally intensive.
  - TD methods often involve estimating unrolled DAGs followed by rolling back which increases computational resource requirement.
  - TD estimation of filters required for conditional independence test is computationally expensive.

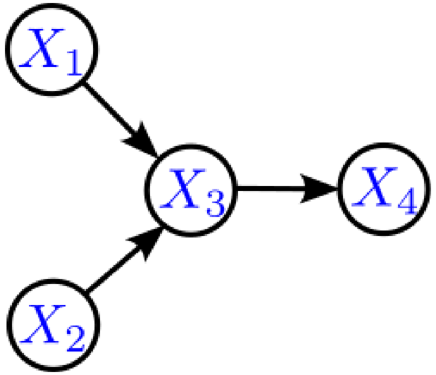
- Materassi, Donatello, and Murti V. Salapaka. "Signal selection for estimation and identification in networks of dynamic systems: A graphical model approach." IEEE Transactions on Automatic Control 65, no. 10 (2019): 4138-4153.
- Spirtes, Peter, Clark N. Glymour, and Richard Scheines. Causation, prediction, and search. MIT press, 2000.
- Wang, Xuewei, Ru Wang, Fei Li, Qiang Lin, Xiaohu Zhao, and Zhenghui Hu. "Large-scale granger causal brain network based on resting-state fMRI data." Neuroscience 425 (2020): 169-180.
- Ji, J., Wang, T., Liu, J., Wang, M., & Tang, W. (2024). River runoff causal discovery with deep reinforcement learning: J. Ji et al. Applied Intelligence, 54(4), 3547-3565.
- Li, Dan, Ming Liu, Shengwei Gao, Yongjun Shi, Yihua Zhang, Zhiyong Li, Patrick Yin Chiang, Franco Maloberti, and Li Geng. "Low-noise broadband CMOS TIA based on multi-stage stagger-tuned amplifier for high-speed high-sensitivity optical communication." IEEE Transactions on Circuits and Systems I: Regular Papers 66, no. 10 (2019): 3676-3689.
- Rana, M. T., Veedu, M. S., & Salapaka, M. V. (2025). Uncovering the Influence Flow Model of Transistor Amplifiers, Its Reconstruction and Application. arXiv preprint arXiv:2508.04977.
- Biswas, R., & Shlizerman, E. (2022). Statistical perspective on functional and causal neural connectomics: The Time-Aware PC algorithm. PLOS Computational Biology, 18(11), e1010653.

Problem Statement

Example:

$$\begin{bmatrix} X_1(\omega) \\ X_2(\omega) \\ X_3(\omega) \\ X_4(\omega) \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ H_{31} & H_{32} & 0 & 0 \\ 0 & 0 & H_{43} & 0 \end{bmatrix} \begin{bmatrix} X_1(\omega) \\ X_2(\omega) \\ X_3(\omega) \\ X_4(\omega) \end{bmatrix} + \begin{bmatrix} \mathcal{E}_1(\omega) \\ \mathcal{E}_2(\omega) \\ \mathcal{E}_3(\omega) \\ \mathcal{E}_4(\omega) \end{bmatrix}$$

$\vec{G}$:



- <u>Linear Dynamic Influence Model (LDIM, $(H, \mathcal{E})$)</u>: Network of dynamic systems expressed as $X(\omega) = H(\omega)X(\omega) + \mathcal{E}(\omega)$ with,
  - $X(\omega)$: Fourier transform of $n \times 1$ vector of measured timeseries $x(t)$,
  - $H(\omega)$: $n \times n$ transfer matrix,
  - $\mathcal{E}(\omega)$: Fourier transform of $n \times 1$ vector of wide-sense stationary inputs with **diagonal power spectral density**, $\Phi_{\mathcal{E}\mathcal{E}}(\omega)$.

- <u>Generative Graph</u>: Directed graph $\vec{G} = (V, \vec{E})$ is generative graph associated with LDIM $(H, \mathcal{E})$ if $\vec{E} = \{(X_i, X_j) | H_{ji} \neq 0\}$

- <u>Problem Statement</u>:
  - Given timeseries measurements $x(t)$ from LDIM.
  - Evaluate computational complexity of TD vs FD based causal discovery.
  - Develop efficient causal discovery algorithms.

- Materassi, Donatello, and Murti V. Salapaka. "Signal selection for estimation and identification in networks of dynamic systems: A graphical model approach." IEEE Transactions on Automatic Control 65, no. 10 (2019): 4138-4153.

Preliminaries and Wiener Filter Based Causal Discovery

- Given directed graph $\vec{G} = (V, \vec{\mathcal{E}})$. For disjoint $X, Y, Z \subset V$, $\boldsymbol{dsep(X,Z,Y)}$ holds if at least one of following hold for all paths between x and y for every $x \in X$, $y \in Y$.
  - The path has a non-collider $p \in Z$.
  - If q is a collider in the path, then neither q nor any of its descendants belong to Z.

- For JWSS processes $x(\cdot)$, and $y(\cdot) = [y_1(\cdot)\ \ y_2(\cdot)\ \ \dots y_n(\cdot)]$, the minimum mean square estimate of $x(\cdot)$ from $y(\cdot)$ is given in terms of the Wiener filter as

$$X_y(\omega) = W_{x.y}(\omega)Y(\omega) = \left[W_{x.y}[y_1](\omega) \dots W_{x.y}[y_n](\omega)\right]\begin{bmatrix} Y_1(\omega) \\ \vdots \\ Y_n(\omega) \end{bmatrix},$$

where $W_{x.y}(\omega) = \Phi_{xy}(\omega)\Phi_{yy}^{-1}(\omega)$ with $\Phi_{xy}(\omega)$ being the cross power spectral density of $x(\cdot)$ and $y(\cdot)$.

- For LDIM: $\boldsymbol{W_{y.[x,Z]}\,[x](\omega) = 0 \Leftarrow dsep(x,Z,y)}$ in corresponding generative graph.
- Faithfulness assumption: $\boldsymbol{W_{y.[x,Z]}\,[x](\omega) = 0 \Rightarrow dsep(x,Z,y)}$.

- Spirtes, Peter, Clark N. Glymour, and Richard Scheines. Causation, prediction, and search. MIT press, 2000.
- Materassi, Donatello, and Murti V. Salapaka. "Signal selection for estimation and identification in networks of dynamic systems: A graphical model approach." IEEE Transactions on Automatic Control 65, no. 10 (2019): 4138-4153.

TD vs FD Estimation of Wiener Filters

- Given $\{x(t)\}_{t=0}^{T}$, TD estimate, $\boldsymbol{W_{i.C}^{(T,L)}(\omega)}$, of $W_{i.C}(\omega)$ obtained by
  - $\beta^* := arg \min_{\beta \in \mathbb{R}^{m(2L+1)}} \frac{1}{T+1} \|\mathrm{x}_i - \mathrm{y}_C \beta\|_2^2$, where $i \in V, C = \{c_1, \dots c_m\} \subset V \backslash \{i\}, \mathrm{x}_i := [x_i(T-L) \ \dots \ x_i(L)]^\top$
  - $\boldsymbol{W_{i.C}^{(T,L)}(\omega)} = \sum_{k=-L}^{L} w_{i.C}^{(T,L)}(k) e^{-j\omega k}$, where $w_{i.C}^{(T,L)}$ is rearrangement of β^* into $(2L+1) \times m$.

  Here,

$$\mathrm{y}_C := \begin{bmatrix} x_{c1}(T) & x_{c1}(T-1) & \dots & x_{c1}(T-2L) & \dots & x_{cm}(T) & \dots & x_{cm}(T-2L) \\ x_{c1}(T-1) & x_{c1}(T-2) & \dots & x_{c1}(T-2L-1) & \dots & x_{cm}(T-1) & \dots & x_{cm}(T-2L-1) \\ & & & \vdots & & & \vdots & \\ x_{c1}(2L) & x_{c1}(2L-1) & \dots & x_{c1}(0) & \dots & x_{cm}(2L) & \dots & x_{cm}(0) \end{bmatrix}$$

- Complexity of computing **TD least square estimate is** $\boldsymbol{O(Tm^2L^2)}$.

- FD estimate, $\boldsymbol{W_{i.C}^{(f)}(\omega_k)}$, of $W_{i|C}(\omega_k)$ obtained by $W_{i.C}^{(f)}(\omega_k) := arg \min_{\beta \in \mathbb{C}^{|C|}} \frac{N}{T} \|\mathcal{X}_i(\omega_k) - \mathcal{X}_C(\omega_k)\beta\|_2^2$, where $\omega_k = \frac{2\pi k}{N}, N = 2^a, a \in \mathbb{N}$, and $\mathcal{X}_i(\omega_k)$ and $\mathcal{X}_C(\omega_k)$ are FFT of data.

- Complexity of estimating **FD estimate is** $\boldsymbol{O(Tm^2 \log N)}$.

- FD approach provides $\boldsymbol{O\left(\frac{L^2}{\log N}\right)}$ improvement over TD.

Proposed Algorithms

Wiener-PC Algorithm

Input: Data $\mathcal{X}(\omega), \omega_k \in \Omega_N, q$.
Output: $\hat{G}$.

1. Initialize the sets, $\mathcal{S} \longleftarrow$ fully connected undirected edge set, $Col \longleftarrow \{\}, DS \longleftarrow \{\}, E_{est} \longleftarrow ()$.
2. For $i = 1, \dots, n$
   a) For $j = 1, \dots, n$, and $j < i$
      i. Initialize $D \longleftarrow V \backslash \{i, j\}$
      ii. For $k = 0, \dots, q-1$
         - For z in $combinations(D, k)$
            - If $|W_{i.[j,z]}[j](\omega_k)| < \tau$
               - $DS(i,j) \longleftarrow z$
               - Delete (i,j) from $\mathcal{S}$
               - break
            - If $|z| \geq q-1$, then $DS(i,j) \longleftarrow -1$
3. For $(i,j) \in V \times V$ and $i < j$
   a) If $\{i,j\} \in \mathcal{S}$, then $E_{est}(i,j) \longleftarrow 1$ and $E_{est}(i,j) \longleftarrow -1$
4. For $(i,j) \in V \times V$ and $i < j$
   a) For $k = j+1, \dots, n$
      i. If $\{i,k\} \in \mathcal{S}$ and $\{i,j\} \in \mathcal{S}$ and $i \notin \{j,k\}$
         - If $DS(j,k) > -1$ and $i \notin DC(j,k)$
            - $Col \longleftarrow Col \cup \{i\}$
            - $E_{est}(j,i) \longleftarrow 0$
            - $E_{est}(k,i) \longleftarrow 0$
5. For $(i,j) \in V \times V$
   a) If $(i,j) \in E_{est}$ and $(j,i) \in E_{est}$
      - Fix the orientation of as many possible such that there are no directed cycles or new colliders.
6. Return $\hat{G} = (V, E_{est})$

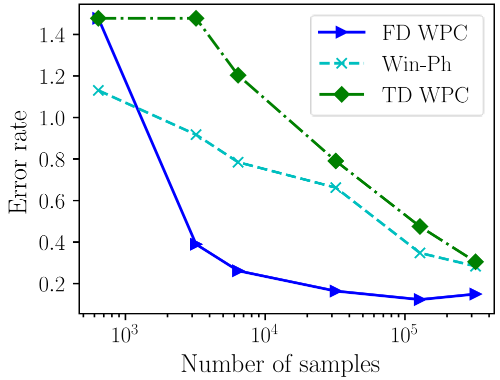
Wiener-Phase Algorithm

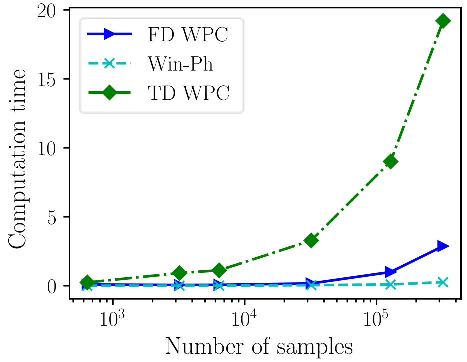
Input: Data $\mathcal{X}(\omega), \omega_k \in \Omega_N$,.
Output: $\hat{G}$.

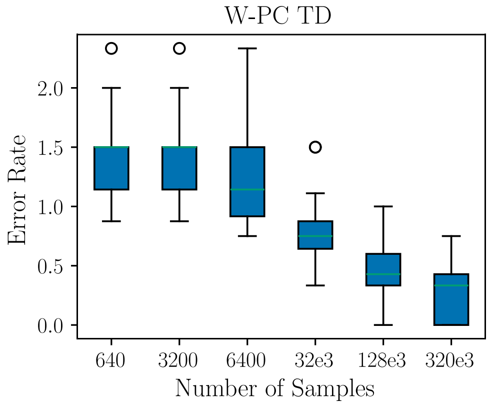
1. Initialize the ordering, $\mathcal{S} \longleftarrow ()$
2. For $i = 1, \dots, n$
   a) Compute $W_{i.\bar{\imath}}^{(f)}(\omega_k)$
   b) For $j = 1, \dots, n$
      i. If $\left|W_{i.\bar{\imath}}^{(f)}[j](\omega_k)\right| > \tau$, then $\mathcal{K} \longleftarrow \mathcal{K} \cup \{i,j\}$
      ii. If $\left|\Im\left\{W_{i.\bar{\imath}}^{(f)}[j](\omega_k)\right\}\right| > \tau$, then $\mathcal{S} \longleftarrow \mathcal{S} \cup \{i,j\}$
3. Compute $\mathcal{SP} := \mathcal{K} \backslash \mathcal{S}$
4. For $\{i,j\} \in \mathcal{SP}$
   a) For $k = 1, \dots, n$
      i. If $\{i,k\} \in \mathcal{S}$ and $\{j,j\} \in \mathcal{S}$
         - $C_{ij} \longleftarrow C_{ij} \cup \{k\}$
   b) If $|C_{ij}| = 1$, then $Col \longleftarrow Col \cup \{k\}$
   c) Else: for $c \in C_{ij}$ compute $W_{i.[j,C]}(\omega_k)$
      i. If $|W_{i.[j,z]}[j](\omega_k)| > \tau$ then $Col \longleftarrow Col \cup \{c\}$
5. Return $\hat{G} = (V, E_{est})$

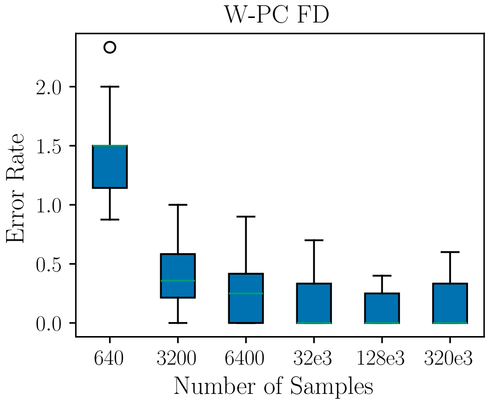
- Complexity of TD Wiener-PC is $\boldsymbol{O}\left(\boldsymbol{T}\boldsymbol{n}^{q+3}\boldsymbol{L}^2\right)$
- Complexity of FD Wiener-PC is $\boldsymbol{O}\left(\boldsymbol{T}\boldsymbol{n}^{q+3}\log \boldsymbol{N}\right)$
- Complexity of Wiener-Phase is $\boldsymbol{O}\left(\boldsymbol{n}^3\left(\boldsymbol{T}\log \boldsymbol{N} + \boldsymbol{q}^2\right)\right)$
- Wiener-Phase is most efficient followed by FD and TD Wiener-PC.
- **Wiener-Phase algorithm uses phase property of FD Wiener filters.**
  - **TD WF has no such property to be exploited.**

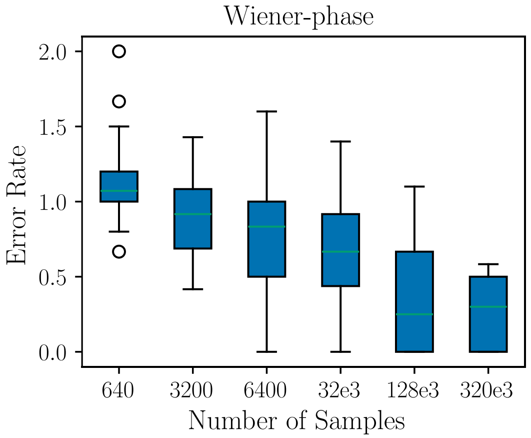
Experimental Results: TD vs FD











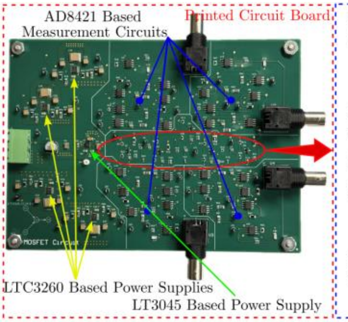
- Proposed algorithms are compared on synthetic data generated using 25 random DAGs.
- Data generative model:

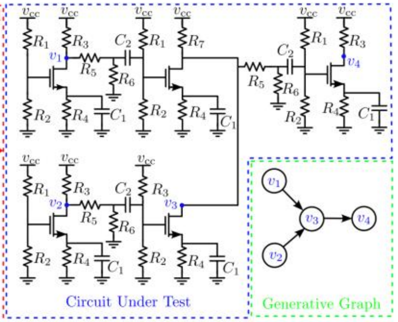
$$x_i(t) + a_i(1)x_i(t-1) + a_i(2)x_i(t-2) + a_i(3)x_i(t-3) = \sum_{j \neq i} b_{ij}x_j(t-1) + e_i(t),$$

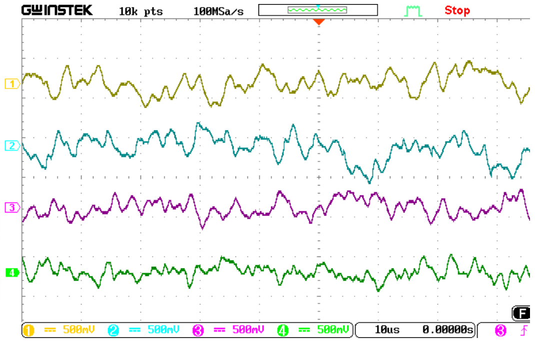
  Here, $e_i(t) \sim \mathcal{N}(0,1), b_{ij} \sim U(0.2,0.4)$.

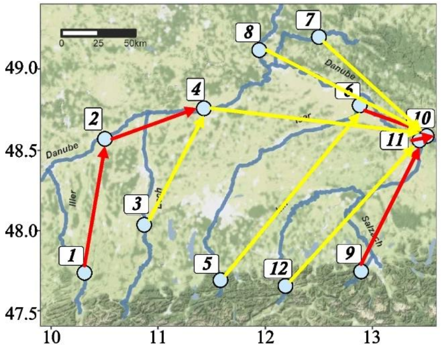
- FD Wiener-PC is best in terms of accuracy and robustness.
- Wiener-Phase is the fastest algorithm among the three.
- TD Wiener-PC is out performed by both of the FD algorithms in terms of accuracy and efficiency.

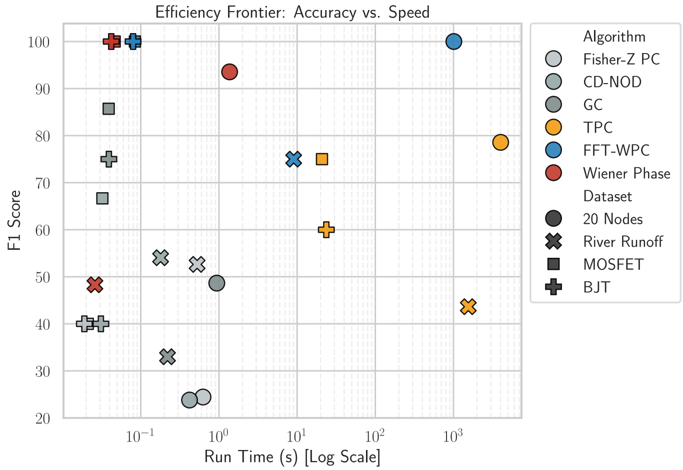
Comparison with Baselines and Application to Circuits

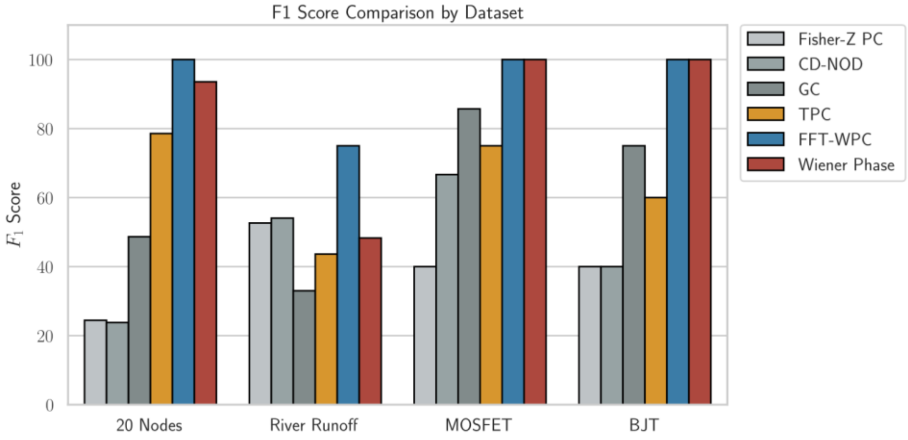




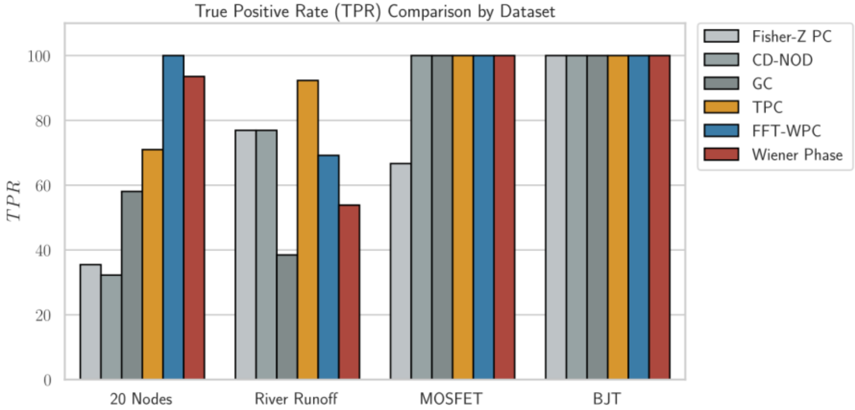


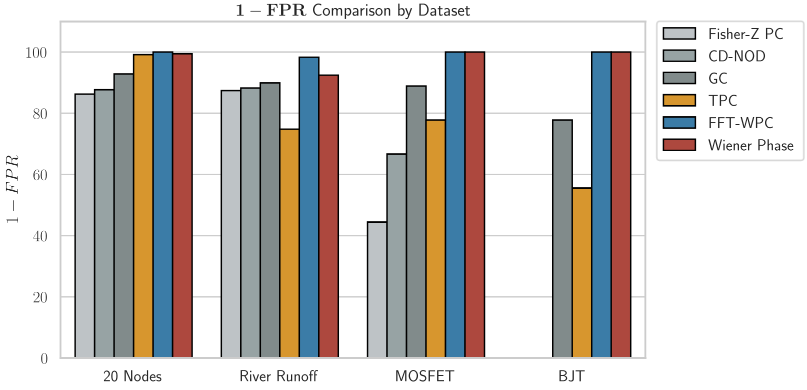






- Baselines are Fisher-Z based PC, CD-NOD, Granger Causality, and Time Aware PC.
- Comparison with baselines using a 20 nodes synthetic dataset, real-world **river-runoff data**, physical **hardware based MOSFET and BJT circuit data.**
- MOSFET and BJT hardware datasets are time-series voltage measurements (sampled at 1Ms/s) at the drain and collector nodes of the circuits.





- Proposed FD algorithms outperform most baselines in terms of accuracy.
- Most accurate baseline, TPC, is outperformed by proposed algorithms on most datasets.

Thank You!

LinkedIn, Mohammed Tuhin Rana

Salapakalab Website