

MÖBIUS TOTAL VARIATION FOR DIRECTED ACYCLIC GRAPHS

Vedran Mihal and Markus Püschel

Department of Computer Science
ETH Zurich, Switzerland

ABSTRACT

We propose a novel definition of total variation (TV) specifically defined for directed acyclic graphs (DAGs). It is a generalization of the classical definition for discrete-time signals (noting that discrete-time is a DAG) and obtained by inverting an integration operator on DAGs, leveraging the theory of Möbius inversion from combinatorics. We demonstrate the performance of our Möbius TV against prior TV definitions when used for denoising a set of real-world DAG signals.

Index Terms— Graph signal processing, directed acyclic graphs, Möbius inversion, total variation denoising

1. INTRODUCTION

Graph signal processing (GSP) has gained significant attention in recent years. It generalizes key concepts of classical signal processing (SP) to signals on graphs, including the notions of Fourier transform, shift, filtering, frequency and total variation [1]. Two related, but different GSP frameworks exist depending on the choice of variation or shift operator as either the graph Laplacian [2] or the adjacency matrix as in [3], which instantiates the general SP theory from [4] for graphs.

In this paper, we focus on the concept of total variation (TV), which is a measure of signal smoothness and used to order the Fourier basis from low to high frequencies. TV-based methods are a driver for many applications, including signal inpainting on graphs [5], graph clustering [6], classification, ranking and link prediction on graphs [7], denoising [8], and network flow problems [9].

Directed graphs. Most of the GSP work has focused on undirected graphs, since the symmetry of the shift guarantees the existence of an orthogonal eigen- (Fourier) basis and thus many other fundamental SP concepts. A complete generalization of GSP to directed graphs is still an open problem [1, Sec. III-A], [10] due to the general lack of a well-defined Fourier basis [11]. To overcome this problem, several solutions have been proposed by adapting the definition of shift or TV [12–19]. Particularly challenging are directed acyclic graphs (DAGs) discussed next.

Our contribution: Total variation for DAGs. DAGs constitute an important subclass of directed graphs since they

underlie the modeling of causal structures and Bayesian networks [20–22]. However, in GSP they are, in a sense, the worst case, since the spectrum collapses: all eigenvalues of the adjacency matrix are zero and thus, for example, the TV in [3] is undefined.

In this paper, we propose a novel form of TV specific for DAGs. It generalizes the classical TV in discrete-time SP (noting that the discrete-time graph is a DAG), and is obtained as the inverse of an integration operator on DAGs. Doing so leverages the theory of Möbius inversion from combinatorics [23] and yields the so-called Möbius transform as the TV operator. As a prototypical application example we perform TV denoising [8, 24] on a number of real-world DAG signals. We compare our Möbius TV to prior notions of TV, including those obtained by ignoring directions in the DAGs, and show that it yields better results.

2. TOTAL VARIATION FOR GRAPHS

We provide background on GSP and total variation.

Graphs. We define a graph as $\mathcal{G} = (\mathcal{V}, A)$ where $\mathcal{V} = \{v_1, \dots, v_n\}$ is the set of vertices and $A \in \mathbb{R}^{n \times n}$ its adjacency matrix. An element A_{ij} in A is $= 1$ if there is an edge from v_j to v_i and $= 0$ otherwise. A graph is undirected if A is symmetric, and directed otherwise.

For undirected graphs, the degree $\deg(v)$ of a vertex v is the number of its adjacent vertices. For directed graphs, we distinguish between in-degree (the number of direct predecessors) and out-degree (the number of direct successors).

For an undirected graph, the degree matrix is $D = \text{diag}(\deg(v_1), \dots, \deg(v_n))$, and the Laplacian is defined as $L = D - A$. For directed graphs, we distinguish between the in-degree and out-degree matrix (D_{in} and D_{out}). We take $L = D_{\text{in}} - A$ as the definition of directed Laplacian [13].

The undirected Laplacian can be written as $L = B^T B$, where B is the oriented incidence matrix. In B , each row corresponds to an undirected edge, and each column to a vertex. If the edge e_i connects v_j and v_k with $j < k$, then $B_{ij} = 1$, $B_{ik} = -1$, and $B_{i\ell} = 0$ otherwise.

A graph signal \mathbf{x} is a column vector $\mathbf{x} = (x_1, \dots, x_n)^T \in \mathbb{R}^n$, where x_i denotes the value at vertex v_i .

Total variation. In classical discrete-time SP total variation $\text{TV}(\mathbf{x}) = \sum_i |x_i - x_{i-1}|$ is a measure of signal smooth-

ness [25]. For a constant signal TV is zero. For graph signals this may not hold, depending on the definition used as explained next.

TV for undirected graphs. Depending on the chosen GSP framework, there are two basic definitions of graph TV. [11] defines it as

$$\text{TV}_A(\mathbf{x}) = \left\| \left(I - \frac{1}{|\lambda_{\max}|} A \right) \mathbf{x} \right\|_p^p, \quad (1)$$

where λ_{\max} is the largest magnitude eigenvalue of A . The choice for p is usually 1 or 2. For constant signals, TV_A is generally not zero. [2] uses the Laplacian:

$$\text{TV}(\mathbf{x}) = \sum_{v_i, v_j \text{ adjacent}} (x_i - x_j)^2 = \mathbf{x}^\top L \mathbf{x} = \|B\mathbf{x}\|_2^2. \quad (2)$$

TV for directed graphs. For directed graphs, [11] uses the same definition as in (1). For DAGs, $|\lambda_{\max}| = 0$ and thus it is undefined. Instead, we later use the unnormalized $\text{TV}_A(\mathbf{x}) = \|(I - A)\mathbf{x}\|_p^p$ in our experiments.

Using the directed Laplacian, [13] proposes

$$\text{TV}_L(\mathbf{x}) = \|L\mathbf{x}\|_p^p = \|(D_{\text{in}} - A)\mathbf{x}\|_p^p. \quad (3)$$

Several variants of the above have been proposed based on different scalings or by enforcing positive differences (see [14, Table I], [12], and also Table 1 later).

We note that the prior definitions of TV, including those in (1), (2), (3), and most of their variants, have the form

$$\text{TV}(\mathbf{x}) = \|\Delta \mathbf{x}\|_p^p, \quad (4)$$

where Δ is a chosen linear difference operator.

3. MÖBIUS TOTAL VARIATION FOR DAGS

In this section we introduce our novel Möbius TV for DAGs, which generalizes the classical TV for discrete-time signals. The basic idea is to define an integral operator on DAGs, which is then inverted to obtain an associated difference operator Δ for use in (4).

DAGs and partial order. Let $\mathcal{G} = (\mathcal{V}, A)$ be a DAG on n vertices. \mathcal{G} induces a partial order¹ on \mathcal{V} : $v \preceq u$ if and only if v is a predecessor of u , i.e., there is a path from v to u in \mathcal{G} . We assume \mathcal{V} to be ordered accordingly, i.e., $v_i \preceq v_j$ implies $i \leq j$. This way, A becomes lower triangular.

TV of a discrete-time signal. A discrete-time signal $\mathbf{x} = (x_i)_{i \in \mathbb{Z}}$ is defined over an infinite DAG with nodes $i \in \mathbb{Z}$ and directed edges $(i-1, i)$ for all $i \in \mathbb{Z}$. As mentioned already above, the classical TV of \mathbf{x} is defined as

$$\text{TV}(\mathbf{x}) = \sum_i |x_i - x_{i-1}| = \|\Delta \mathbf{x}\|_1, \quad (5)$$

¹A partial order on set S is a binary relation \preceq which satisfies three properties: 1. $x \preceq x$ for every $x \in S$ (reflexivity); 2. $x \preceq y$ and $y \preceq x$ implies $x = y$ (antisymmetry); 3. $x \preceq y$ and $y \preceq z$ implies $x \preceq z$ (transitivity).

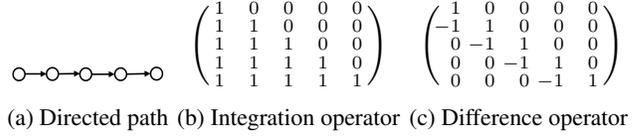


Fig. 1: Directed path on five vertices

where Δ is a linear difference operator: $(\Delta \mathbf{x})_i = x_i - x_{i-1}$. The inverse of Δ is the, also linear, integration operator Σ that sums over all predecessor nodes:

$$(\Sigma \mathbf{x})_i = \sum_{j \leq i} x_j. \quad (6)$$

TV for DAGs. We generalize (6) to signals \mathbf{x} on arbitrary DAGs $\mathcal{G} = (\mathcal{V}, A)$ by summing over all predecessors of a node in \mathcal{G} , i.e.,

$$(Z\mathbf{x})_i = \sum_{v_j \preceq v_i} x_j.$$

We write Z instead of Σ since it is known as zeta transform in combinatorics [23]. Z is lower triangular and the adjacency matrix of the so-called transitive-reflexive closure of \mathcal{G} , i.e., $Z_{ij} = 1$ if $v_j \preceq v_i$ and $= 0$ else.

Inverting Z yields the associated difference operator $\Delta = M = Z^{-1}$, which is known as Möbius transform.² A closed form for M is provided in [23]. Namely $M_{ij} = \mu(v_j, v_i)$, where μ is the Möbius function defined recursively as

$$\begin{aligned} \mu(v_i, v_i) &= 1, & \text{for all } v_i \in \mathcal{V}, \\ \mu(v_i, v_j) &= - \sum_{v_k \preceq v_i \prec v_j} \mu(v_i, v_k), & \text{for } v_i \neq v_j. \end{aligned}$$

Intuitively, $M\mathbf{x}_i$ captures the incremental change of x_i , given the values of the predecessors of v_i in \mathcal{G} , generalizing (5). As a result we obtain our

$$\text{Möbius TV} : \text{TV}_M(\mathbf{x}) = \|M\mathbf{x}\|_p^p. \quad (7)$$

For $p = 2$ this definition suggests, in analogy to (2), the eigenbasis of $M^\top M$ as associated orthogonal Fourier basis.

Möbius TV for a directed path. As a first example, we consider the simplest DAG: a finite directed path. The integration and difference operators Z and M are shown in Figure 1. In this case, $M = I - A$, the unnormalized Δ in (1). With $p = 1$ we get

$$\text{TV}_M(\mathbf{x}) = \|M\mathbf{x}\|_1 = |x_1| + \sum_{i=2}^n |x_i - x_{i-1}|, \quad (8)$$

We notice two undesirable properties: TV_M is not zero for constant signals, and, in particular, the value at the source (node with in-degree 0) v_1 occurs undifferentiated. We propose two solutions for arbitrary DAGs next.

²Interestingly, and different from this paper, M also admits an interpretation as a Fourier transform [26].

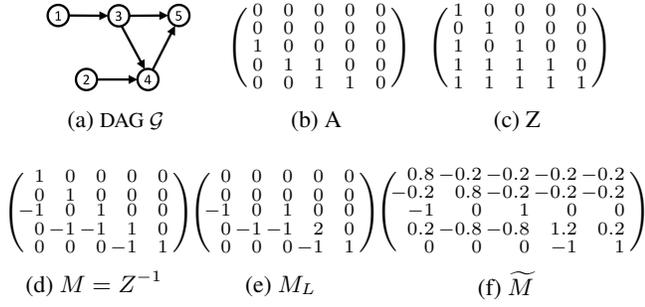


Fig. 2: DAG with five vertices, adjacency matrix, integration operator, and the proposed difference operators

Möbius Laplacian. Our first solution is similar to a Laplacian: we subtract the row sums from the diagonal elements in M :

$$M_L = M - D_M,$$

where $D_M = \text{diag}\left(\sum_{j=1}^n M_{1j}, \dots, \sum_{j=1}^n M_{nj}\right)$. The corresponding TV is then $\text{TV}_{M_L}(\mathbf{x}) = \|M_L \mathbf{x}\|_p^p$, for which constant signals have TV zero. In particular, this definition takes out sources. Indeed, for the directed path and $p = 1$ we get

$$\text{TV}_{M_L}(\mathbf{x}) = \sum_{i=2}^n |x_i - x_{i-1}|,$$

which, in this case, coincides with (3).

Modified Möbius transform. Our second proposal ensures zero-sum rows in the Möbius matrix by subtracting the corresponding row sum scaled by $1/n$ from each element in the matrix:

$$\widetilde{M} = M - \widetilde{D}_M, \quad (\widetilde{D}_M)_{ij} = \frac{1}{n} \sum_{j=1}^n M_{ij}.$$

The corresponding TV is $\text{TV}_{\widetilde{M}}(\mathbf{x}) = \|\widetilde{M} \mathbf{x}\|_p^p$ and ensures that constant signals have TV zero.

Example. Fig. 2 shows an example DAG and our three proposed difference operators. Note that none of them corresponds to a prior difference operator.

4. EXPERIMENTS

We evaluate our proposed Möbius TVs by performing classical total variation denoising [24], ported to graphs as in [8], of DAG signals. We compare against various previous notions of TV.

TV operators. We consider DAGs $\mathcal{G} = (\mathcal{V}, A)$ and overall twelve TV operators of the form $\text{TV}(\mathbf{x}) = \|\Delta \mathbf{x}\|_2^2$, with Δ shown in Table 1. The directed operators include our three versions of Möbius TV. The undirected operators are obtained

by ignoring the directions in \mathcal{G} , i.e., setting $\bar{A} = A + A^\top$, and use the TVs from (1), (2), and variants from [14, Table I].

TV denoising. Given a signal \mathbf{x} on \mathcal{G} , we add uncorrelated zero-mean Gaussian noise \mathbf{e} : $\mathbf{y} = \mathbf{x} + \mathbf{e}$, considering signal-to-noise ratios (SNRs) of 5 and 15. TV denoising is then done by solving the following minimization problem:

$$\tilde{\mathbf{x}} = \arg \min_{\mathbf{x}} (\|\mathbf{x} - \mathbf{y}\|_2^2 + \alpha \text{TV}(\mathbf{x})) \quad (9)$$

$$= (I + \alpha \cdot \Delta^\top \Delta)^{-1} \mathbf{y}, \quad (10)$$

where $\tilde{\mathbf{x}}$ is the denoised signal and the parameter α balances between the reconstruction error and the smoothness according to the chosen TV. The choice of squared 2-norm in TV ensures that we have the closed-form solution in (10).

Finding the optimal value of α is non-trivial [27, 28], so we consider a relevant range in our experiments.

Graphs and data. In synthetic experiments with smooth signals (w.r.t. the Möbius TVs) denoising with our Möbius TVs performed best by construction. Thus, we focused on real-world data to assess the potential of our TVs more realistically.

We consider three DAGs and associated data that we could find for our experiments, considering four different DAG signals in each case. The first one is the gene causal network for *Arabidopsis thaliana* plant (Arth150) [29] with 107 nodes. Nodes represent genes, and directed edges are determined by a complex statistical method described in [30], which creates a DAG. The signals assign to each gene the degree of expression in a cell.

The second one is the dependency Bayesian network for the indica rice population (Magic-irri) [29] with 64 nodes. Nodes represent single nucleotide polymorphisms and phenotypic traits, and edges represent directed stochastic dependencies between them. The network is a DAG, and the signals are allele frequencies of single nucleotide polymorphisms and phenotypic trait measurements [31].

The third one is a network of the river Thames and its major tributaries [32], where measurements are taken at 13 different sites once a week. Nodes represent pairs (site, time point); edges connect adjacent sites at the same time point directed as the water flows, and a site with itself at consecutive time points. The resulting graph is a DAG, and the signals represent mean daily flow measured once a week for the years 2013–2016 (four signals).

Since for the Thames river network only four signals were available, we decided to use the same number for the other networks for consistency. However, for Arth150 and Magic-irri network, we also considered 100 signals and obtained qualitatively similar results.

Results. We show our results in Fig. 4 with the legend provided in Fig. 3. For each DAG, there are two plots: one for the low SNR of 5, and one for the high SNR of 15. The x-axis represents the chosen parameter α , and the y-axis represents SNR after denoising. We vary the scale on the x- and

Directed operators	Δ	Undirected operators	Δ
Directed shift:	$I - A$	Undirected shift:	$I - \frac{\bar{A}}{ \lambda_{\max} }$
Directed Laplacian:	$D_{\text{in}} - A$	Undirected Laplacian:	$\bar{D} - \bar{A}$
Normalized dir. Laplacian: ¹	$I - D_{\text{in}}^{-1}A$	Normalized undirected Laplacian:	$I - \bar{D}^{-\frac{1}{2}}\bar{A}\bar{D}^{-\frac{1}{2}}$
Möbius matrix:	M	Left norm. undirected Laplacian:	$I - \bar{D}^{-1}\bar{A}$
Möbius Laplacian:	M_L	Oriented incidence matrix:	B
Modified Möbius matrix:	\widetilde{M}	Norm. oriented incidence matrix:	$B\bar{D}^{-\frac{1}{2}}$

¹ We add self-loops to all sources of the graph to avoid dividing by zero.

Table 1: Difference operators used for denoising with $\text{TV}(\mathbf{x}) = \|\Delta\mathbf{x}\|_2^2$.



Fig. 3: Legend for Fig. 4

y-axis, to focus on the range of interest for α . Undirected TVs correspond to dashed lines and our Möbius TVs are shown in different red/orange tones. The baseline SNR of \mathbf{y} is a horizontal black line. Shaded areas around lines represent the standard deviation over the four signals.

Discussion. First we note that for the chosen data and TV operators some denoising is achieved in most cases for a suitable α , with the exception of Figure 4b (right), where all methods fail. For the other five, one of our novel Möbius TVs performs best and also offers better robustness with respect to the choice of α . Further, ignoring directions (dashed lines) yields significantly inferior results in most cases.

For signals on Arth150 network (Figure 4a), the property that constant signals have zero TV appears to be important, since both Möbius Laplacian and modified Möbius yield significantly better denoising than the standard Möbius TV.

Denoising based on (9) can only work if the original signal is smooth w.r.t. the chosen TV. Indeed, we confirmed by inspecting the associated spectra that signals on the Magic-irri network are not low-frequency (but other DAG signals are; not shown due to lack of space), and it is the likely explanation for the failure on SNR = 15. For the Möbius TVs we computed the spectra as explained above (7).

Our Möbius TVs perform particularly well on the Thames network (Figure 4c). The directionality through the water flow and the measurement along time makes it an obvious DAG and gives hope that other applications that measure graph signals along discrete time steps could benefit as well. Overall, the choice of a proper TV matters but how to do so remains an open problem and is likely application-dependent.

5. CONCLUSION

DAGs are a particularly important class of directed graphs since they can model causal structures and processes as well as graph data measured along time. Prior GSP does not instantiate well for DAGs because of the lack of a well-defined Fourier basis and associated concepts. In this paper we proposed a novel form of total variation that is not based on any prior GSP variant. The key idea was to reinterpret the classical Möbius transform from combinatorics as a difference

operator, which, theoretically pleasing, generalizes the classical TV definition for the discrete-time DAG. Our experiments with TV denoising demonstrated the practical viability of our Möbius TV and invite further research to better understand its properties and to develop further applications.

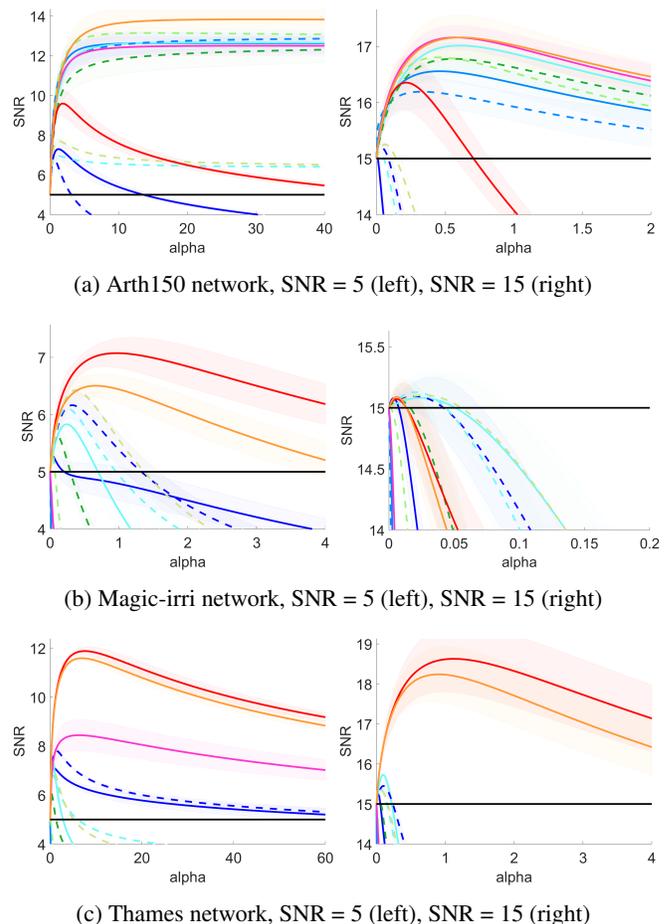


Fig. 4: TV denoising on three classes of DAG data with twelve different TV operators. The scales on the axes are different.

6. REFERENCES

- [1] A. Ortega, P. Frossard, J. Kovačević, J. M. F. Moura, and P. Vandergheynst, “Graph signal processing: Overview, challenges, and applications,” *Proc. IEEE*, vol. 106, no. 5, pp. 808–828, 2018.
- [2] D. I. Shuman, S. K. Narang, P. Frossard, A. Ortega, and P. Vandergheynst, “The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains,” *IEEE Signal Process. Mag.*, vol. 30, no. 3, pp. 83–98, 2013.
- [3] A. Sandryhaila and J. M. F. Moura, “Discrete signal processing on graphs,” *IEEE Trans. Signal Process.*, vol. 61, 10 2012.
- [4] M. Püschel and J. M. F. Moura, “Algebraic signal processing theory: Foundation and 1-D time,” *IEEE Trans. Signal Process.*, vol. 56, no. 8, pp. 3572–3585, 2008.
- [5] S. Chen, A. Sandryhaila, G. Lederman, Z. Wang, J. M. F. Moura, P. Rizzo, J. Bielak, J. H. Garrett, and J. Kovačević, “Signal inpainting on graphs via total variation minimization,” in *IEEE Int. Conf. Acoust., Speech, Signal Process.*, 2014, pp. 8267–8271.
- [6] X. Dong, P. Frossard, P. Vandergheynst, and N. Nefedov, “A regularization framework for mobile social network analysis,” in *IEEE Int. Conf. Acoust., Speech, Signal Process.*, 2011, pp. 2140–2143.
- [7] D. Zhou and B. Schölkopf, “A regularization framework for learning from graph data,” in *ICML Workshop on Statistical Relational Learning and Its Connections to Other Fields*, 2004, pp. 132–137.
- [8] S. Chen, A. Sandryhaila, J. M. F. Moura, and J. Kovačević, “Signal denoising on graphs via graph filtering,” in *IEEE Global Conf. Signal Inf. Process.*, 2014, pp. 872–876.
- [9] A. Jung, A. O., Hero III, A. C. Mara, S. Jahromi, A. Heimowitz, and Y. C. Eldar, “Semi-supervised learning in network-structured data via total variation minimization,” *IEEE Trans. Signal Process.*, vol. 67, no. 24, pp. 6256–6269, 2019.
- [10] A. G. Marques, S. Segarra, and G. Mateos, “Signal processing on directed graphs: The role of edge directionality when processing and learning from network data,” *IEEE Signal Process. Mag.*, vol. 37, no. 6, pp. 99–116, 2020.
- [11] A. Sandryhaila and J. M. F. Moura, “Discrete signal processing on graphs: Frequency analysis,” *IEEE Trans. Signal Process.*, vol. 62, no. 12, pp. 3042–3054, 2014.
- [12] R. Shafipour, A. Khodabakhsh, G. Mateos, and E. Nikolova, “A directed graph Fourier transform with spread frequency components,” *IEEE Trans. Signal Process.*, vol. 67, no. 4, pp. 946–960, 2019.
- [13] R. Singh, A. Chakraborty, and B. S. Manoj, “Graph Fourier transform based on directed Laplacian,” in *Int. Conf. Signal Process. Commun.*, 2016, pp. 1–5.
- [14] A. Anis, A. Gadde, and A. Ortega, “Efficient sampling set selection for bandlimited graph signals using graph spectral proxies,” *IEEE Trans. Signal Process.*, vol. 64, no. 14, pp. 3775–3789, 2016.
- [15] B. Seifert and M. Püschel, “Digraph signal processing with generalized boundary conditions,” *IEEE Trans. Signal Process.*, vol. 69, pp. 1422–1437, 2021.
- [16] S. Sardellitti, S. Barbarossa, and P. Di Lorenzo, “On the graph Fourier transform for directed graphs,” *IEEE J. Sel. Topics Signal Process.*, vol. 11, no. 6, pp. 796–811, 2017.
- [17] J. Domingos and J. M. F. Moura, “Graph Fourier transform: A stable approximation,” *IEEE Trans. Signal Process.*, vol. 68, pp. 4422–4437, 2020.
- [18] J. A. Deri and J. M. F. Moura, “Spectral projector-based graph Fourier transforms,” *IEEE J. Sel. Topics Signal Process.*, vol. 11, no. 6, pp. 785–795, 2017.
- [19] S. Furutani, T. Shibahara, M. Akiyama, K. Hato, and M. Aida, “Graph signal processing for directed graphs based on the Hermitian Laplacian,” in *Joint Eur. Conf. Mach. Learning and Knowledge Discovery in Databases*. Springer, 2019, pp. 447–463.
- [20] P. M. K. Illari, F. Russo, and J. Williamson, *Causality in the Sciences*, Oxford University Press, 2011.
- [21] D. Koller and N. Friedman, *Probabilistic graphical models: principles and techniques*, MIT press, 2009.
- [22] B. Schölkopf, “Causality for machine learning,” in *Probabilistic and Causal Inference: The Works of Judea Pearl*, pp. 765–804. 2022.
- [23] G.-C. Rota, “On the Foundations of Combinatorial Theory I. Theory of Möbius functions,” *Z. Wahrscheinlichkeitstheorie und Verwandte Gebiete*, vol. 2, no. 4, pp. 340–368, 1964.
- [24] L. I. Rudin, S. Osher, and E. Fatemi, “Nonlinear total variation based noise removal algorithms,” *Physica D: nonlinear phenomena*, vol. 60, no. 1–4, pp. 259–268, 1992.
- [25] S. Mallat, *A wavelet tour of signal processing (2. ed.)*, Academic Press, 1999.
- [26] B. Seifert, C. Wendler, and M. Püschel, “Causal Fourier Analysis on Directed Acyclic Graphs and Posets,” arXiv:2209.07970, 2022.
- [27] W. C. Karl, “Regularization in image restoration and reconstruction,” in *Handbook of Image and Video Process.*, pp. 183–V. Elsevier, 2005.
- [28] S. Ramani, T. Blu, and M. Unser, “Monte-Carlo SURE: A black-box optimization of regularization parameters for general denoising algorithms,” *IEEE Trans. Image Process.*, vol. 17, no. 9, pp. 1540–1554, 2008.
- [29] M. Scutari, “bnlearn - an R package for Bayesian network learning and inference, Bayesian Network Repository,” 2022, <https://www.bnlearn.com/bnrepository/>, (accessed Oct. 6, 2022).
- [30] R. Opgen-Rhein and K. Strimmer, “From correlation to causation networks: a simple approximate learning algorithm and its application to high-dimensional plant gene expression data,” *BMC systems biology*, vol. 1, no. 1, pp. 1–10, 2007.
- [31] M. Scutari, P. Howell, D. J. Balding, and I. Mackay, “Multiple quantitative trait analysis using Bayesian networks,” *Genetics*, vol. 198, no. 1, pp. 129–137, 2014.
- [32] M. J. Bowes, L. K. Armstrong, S. A. Harman, D. J. E. Nicholls, H. D. Wickham, P. M. Scarlett, and M. D. Juergens, “Weekly water quality data from the River Thames and its major tributaries (2009–2017),” 2020, <https://doi.org/10.5285/cf10ea9a-a249-4074-ac0c-e0c3079e5e45>, (accessed Oct. 6, 2022).