Distributed Denoising over Simplicial Complexes using Chebyshev Polynomial Approximation

Sai Kiran Kadambari, Robin Francis, and Sundeep Prabhakar Chepuri
Indian Institute of Science, Bangalore, India

***Abstract*—In this work, we focus on denoising smooth signals supported on simplicial complexes in a distributed manner. We assume that the simplicial signals are dominantly smooth on either the lower or upper Laplacian matrices, which are used to compose the so-called Hodge Laplacian matrix. This corresponds to denoising non-harmonic signals on simplicial complexes. We pose the denoising problem as a convex optimization problem, where we assign different weights to the quadratic regularizers related to the upper and lower Hodge Laplacian matrices and express the optimal solution as a sum of simplicial complex operators related to the two Laplacian matrices. We then use the recursive relation of the Chebyshev polynomial to implement these operators in a distributed manner. We demonstrate the efficacy of the developed framework on synthetic and real-world datasets.**

***Index Terms*—Chebyshev polynomials, denoising, distributed signal processing, simplicial complex.**

I. Introduction

Graph signal processing (GSP) has well-documented merits for processing data defined on irregular domains [1], [2]. GSP methods capture complex pairwise relations between different entities as explained by the structure of the underlying graph and process signals supported on its nodes. However, in many real-world applications, interactions between entities are not always limited to pairwise relations, but they could be of higher-order. Some example networks with higher-order interactions are co-authorship, social, and biological networks, to name a few. For instance, people communicate or work in groups in a social network. In a biological network, more than two proteins interact simultaneously. Higher-order interactions in these networks cannot be modeled using simple graphs.

Simplicial complexes are mathematical objects that can naturally represent higher-order interactions between more than two entities [3]–[6]. A simplicial complex is a collection of simplices of different orders, e.g., node is a zero-order simplex, an edge is first-order simplex, a triangle is a second-order simplex, and so on. Signals defined over simplices of different order are referred to as higher-order simplicial signals. For example, in an email communication network, the zero-order simplices can model the network users and the interactions between a subset of 3 users form signals on a second-order simplex.

Traditional GSP methods have been extended to process signals defined on higher-order networks; see [3], [4], [7]–[9]. For processing signals on simplicial complexes, [7]–[11] represent a simplicial complex using the so-called Hodge Laplacian matrix [12]. A Hodge Laplacian matrix is the sum of upper and lower Laplacian matrices that encodes the interactions between the adjacent upper and lower simplices, respectively. Some examples of signal processing tasks over simplicial complexes are filtering [7], denoising [3], [10], and sampling [4], [9]. The denoising problem considered in [3], [10] focuses on signals that are equally smooth on both lower and upper Laplacian matrices. In many real-world datasets, the signals are dominantly smooth on either the lower or upper Laplacian matrices (see Section IV-B for illustration). For denoising and processing such signals over higher-order networks, it is of interest to have a distributed implementation.

In this work, we focus on distributed denoising of higher-order signals that are dominantly smooth on either lower or upper Hodge Laplacian matrices. In other words, we focus on denoising non-harmonic signals. Specifically, the contributions of this paper are as follows. We propose a weighted total variation (`WTV`) norm regularized denoising problem, where we use different weights for the smoothness promoting quadratic terms related to the lower and upper Hodge Laplacian matrices. We leverage the properties of the Hodge Laplacian matrices and Chebyshev polynomials to implement the closed-form solution of the `WTV` denoising problem in a distributed setting. The computation complexity of the proposed method scales linearly in the number of higher-order simplices. Finally, we demonstrate the performance of the proposed method on synthetic and real-world datasets.

Throughout the paper, we use calligraphic letters $\mathcal{V}$ to represent sets and $|\mathcal{V}|$ denotes its cardinality. We denote matrices and vectors with boldface upper and lower case letters as $\mathbf{A}$ and $\mathbf{a}$, respectively. For a vector $\mathbf{a}$, $[\mathbf{a}]_n$ represents the nth element of the vector. Given a sparse matrix $\mathbf{A}$, $\mathrm{nnz}(\mathbf{A})$ represents the number of non-zero entries in $\mathbf{A}$. We use $\mathbf{A}^T$ and $\mathbf{A}^{-1}$ to represent the trace and inverse of the matrix $\mathbf{A}$, respectively.

II. Preliminaries

In this section, we give a brief introduction to simplicial complexes and signals defined on simplicial complexes. We then introduce the Chebyshev polynomial approximation technique to efficiently implement the product of a function of a matrix and a vector.

A. Simplicial complex

Given a finite set of vertices $\mathcal{V}$, a collection of $k+1$ vertices is called a $k-$simplex $\mathcal{S}_k$, if $|\mathcal{S}_k| = k+1$. A simplicial

This work was supported in part by the Pratiksha Trust Fellowship and IFCAM project 19/56.

complex $\mathcal{S}$ is a collection of simplices of different order such that for any simplex $\mathcal{S}_k \in \mathcal{S}$, if $\sigma \subseteq \mathcal{S}_k$ then $\sigma \in \mathcal{S}$. We denote the number of $k-$simplices $\mathcal{S}_k$ present in $\mathcal{S}$ by N_k. The dimension of the simplicial complex $\mathcal{S}$ is the highest order simplex present in it. We denote the dimension of $\mathcal{S}$ as K.

The structure of a simplicial complex is completely captured by the higher-order Hodge Laplacian matrices, $\mathbf{L}_k \in \mathbb{R}^{N_k \times N_k}$, defined as

$$\begin{aligned} \mathbf{L}_k &= \mathbf{B}_k^T \mathbf{B}_k + \mathbf{B}_{k+1} \mathbf{B}_{k+1}^T \\ &= \mathbf{L}_{kl} + \mathbf{L}_{ku}, \quad k = 1, 2, \ldots, K, \end{aligned} \tag{1}$$

where $\{\mathbf{B}_k\}_{k=1}^K$ are the higher-order incidence matrices. Specifically, $\mathbf{B}_k \in \mathbb{R}^{N_{k-1} \times N_k}$ captures the interactions between $(k-1)-$simplices and $k-$simplices. The incidence matrix takes values $\{-1, 1\}$ if a $(k-1)-$simplex is incident on a $k-$simplex and it has zeros, elsewhere. For a $k-$th order Hodge Laplacian matrix $\mathbf{L}_k$, we define $\mathbf{L}_{kl} = \mathbf{B}_k^T \mathbf{B}_k$ and $\mathbf{L}_{ku} = \mathbf{B}_{k+1} \mathbf{B}_{k+1}^T$ as the lower and upper Hodge Laplacian matrices. The matrices $\mathbf{L}_k$, $\mathbf{L}_{kl}$, and $\mathbf{L}_{ku}$ are symmetric and positive semidefinite by construction. Thus, the Hodge Laplacian matrix $\mathbf{L}_k$ admits the eigenvalue decomposition

$$\mathbf{L}_k = \mathbf{U}_k \mathbf{\Lambda}_k \mathbf{U}_k^T, \tag{2}$$

where we collect the eigenvalues of $\mathbf{L}_k$ in the diagonal matrix $\mathbf{\Lambda}_k \in \mathbb{R}^{N_k \times N_k}$ and the corresponding eigenvectors in the columns of $\mathbf{U}_k \in \mathbb{R}^{N_k \times N_k}$. The eigenvalues of $\mathbf{L}_k$ are non-negative and has an interpretation of higher-order frequencies related to the simplicial complex. Further, we have $\mathbf{B}_k \mathbf{B}_{k+1} = \mathbf{0}$ [4]. Thus, the lower and upper Laplacian matrices satisfy the relation $\mathbf{L}_{kl} \mathbf{L}_{ku} = \mathbf{L}_{ku} \mathbf{L}_{kl} = \mathbf{0}$.

B. Signals over a simplicial complex

Given a simplicial complex $\mathcal{S}$, the signal $\mathbf{x}_k \in \mathbb{R}^{N_k}$ on a $k-$simplex is a real-valued map from the $k-$simplex to $\mathbb{R}^{N_k}$. For instance, the vectors $\mathbf{x}_0 \in \mathbb{R}^{N_0}$, $\mathbf{x}_1 \in \mathbb{R}^{N_1}$ and $\mathbf{x}_2 \in \mathbb{R}^{N_2}$ are, respectively, signals indexed by the nodes, edges, and triangles of a simplicial complex of dimension 2.

The amount of smoothness of $\mathbf{x}_k \in \mathbb{R}^{N_k}$ with respect to (w.r.t.) the kth order Hodge Laplacian matrix $\mathbf{L}_k \in \mathbb{R}^{N_k \times N_k}$ is captured by the quadratic total variation form

$$\begin{aligned} \mathbf{x}_k^T \mathbf{L}_k \mathbf{x}_k &= \mathbf{x}_k^T \left(\mathbf{L}_{kl} + \mathbf{L}_{ku} \right) \mathbf{x}_k \\ &= \mathbf{x}_k^T \mathbf{L}_{kl} \mathbf{x}_k + \mathbf{x}_k^T \mathbf{L}_{ku} \mathbf{x}_k. \end{aligned} \tag{3}$$

Smaller the quadratic form, smoother the $k-$simplicial signal is. Further, the quadratic forms $\mathbf{x}_k^T \mathbf{L}_{kl} \mathbf{x}_k$ and $\mathbf{x}_k^T \mathbf{L}_{ku} \mathbf{x}_k$ measure the total variation of the signal $\mathbf{x}_k$ w.r.t. the lower and upper Laplacian matrices, respectively. Harmonic signals $\mathbf{x}_k^{\mathrm{Har}}$ are the smoothest possible signals with the smallest total variation $(\mathbf{x}_k^{\mathrm{Har}})^T \mathbf{L}_k \mathbf{x}_k^{\mathrm{Har}} = 0$ and are spanned by the eigenvectors corresponding to the zero eigenvalues of $\mathbf{L}_k$. This also means that $\mathbf{x}_k^{\mathrm{Har}}$ is jointly and uniformly smooth on both the upper and lower Laplacian matrices $\mathbf{L}_{kl}$ and $\mathbf{L}_{ku}$.

In many cases, the variation of a smooth $k-$simplicial signal is dominant w.r.t. either $\mathbf{L}_{kl}$ or $\mathbf{L}_{ku}$. This implies that, the signal is more smooth on either the lower or upper Laplacian matrices. For example, consider a smooth flow signal $\mathbf{x}_1$ on a simplicial complex as shown in Fig. $1(a)$. This signal is more smooth on the lower Laplacian matrix $\mathbf{L}_{kl}$ compared to the upper Laplacian matrix $\mathbf{L}_{ku}$.

C. Chebyshev polynomial approximation

For a real symmetric matrix $\mathbf{A} \in \mathbb{R}^{N \times N}$ with eigenvalue decomposition $\mathbf{A} = \mathbf{U} \mathbf{\Lambda} \mathbf{U}^T$, the matrix function $g(\mathbf{A})$ satisfies $g(\mathbf{A}) = \mathbf{U} g(\mathbf{\Lambda}) \mathbf{U}^T$, where $g(\mathbf{\Lambda})$ is a function defined on eigenvalues of $\mathbf{A}$. Given a matrix $\mathbf{A} \in \mathbb{R}^{N \times N}$ and a vector $\mathbf{y} \in \mathbb{R}^N$, computing the matrix vector product $g(\mathbf{A})\mathbf{y}$ involves an eigenvalue decomposition that costs about $\mathcal{O}(N^3)$ flops. This computation complexity can be reduced by approximating $g(\cdot)$ with a truncated and shifted Chebyshev polynomials.

A function g defined on $[\lambda_{\text{min}}, \lambda_{\text{max}}]$ can be expressed using the shifted Chebyshev polynomials

$$g(\lambda) = \frac{1}{2} c_0 + \sum_{p=1}^{\infty} c_p \bar{T}_p(\lambda), \;\text{ for all }\; \lambda \in [\lambda_{\text{min}}, \lambda_{\text{max}}], \tag{4}$$

where $\left\{\bar{T}_p(\lambda)\right\}_{p=0}^{\infty}$ are the shifted Chebyshev polynomials with coefficients $\{c_p\}_{p=1}^{\infty}$ given by

$$c_k := \frac{2}{\pi} \int_0^{\pi} \cos(k\phi) g\left(\beta \cos(\phi) + \alpha\right) d\phi.$$

Here, $\alpha = \left(\lambda_{\text{max}} + \lambda_{\text{min}}\right)/2$ and $\beta = \left(\lambda_{\text{max}} - \lambda_{\text{min}}\right)/2$. The shifted Chebyshev polynomials satisfy the following recurrence relation

$$\bar{T}_p(\lambda) := \begin{cases} 1, & \text{if, } p = 0 \\ \frac{\lambda - \alpha}{\beta}, & \text{if, } p = 1 \\ \frac{2}{\beta}(\lambda - \alpha)\bar{T}_{p-1}(\lambda) - \bar{T}_{p-2}(\lambda), & \text{if, } p \geq 2. \end{cases}$$

Given a vector $\mathbf{y} \in \mathbb{R}^N$ and symmetric matrix $\mathbf{A} \in \mathbb{R}^{N \times N}$, we have the recurrence relation

$$\bar{T}_p(\mathbf{A})\mathbf{y} = \frac{2}{\beta} \left(\mathbf{A} - \alpha \mathbf{I} \right) \bar{T}_{p-1}(\mathbf{A})\mathbf{y} - \bar{T}_{p-2}(\mathbf{A})\mathbf{y}, \tag{5}$$

which means that $\bar{T}_p(\mathbf{A})\mathbf{y}$ can be computed from $\bar{T}_{p-1}(\mathbf{A})\mathbf{y}$ and $\bar{T}_{p-2}(\mathbf{A})\mathbf{y}$.

We can then approximate the matrix vector product $g(\mathbf{A})\mathbf{y} = \sum_{i=1}^N g(\lambda_i) \mathbf{u}_i \mathbf{u}_i^T \mathbf{y}$ using the shifted Chebyshev approximation as

$$\begin{aligned} g(\mathbf{A})\mathbf{y} &\overset{(a)}{\approx} \sum_{i=1}^N \left(\frac{1}{2} c_0 + \sum_{p=1}^P c_p \bar{T}_p(\lambda_i) \right) \mathbf{u}_i \mathbf{u}_i^T \mathbf{y} \\ &= \left(\frac{1}{2} c_0 + \sum_{p=1}^P c_p \bar{T}_p(\mathbf{A}) \right) \mathbf{y}, \end{aligned} \tag{6}$$

where (a) is the Pth order approximation of $g(\lambda)$ in (4). The approximation error is smaller for larger P, which depends on the operator $g(\cdot)$. For a sparse matrix $\mathbf{A}$, computing $g(\mathbf{A})\mathbf{y}$ using the Pth order Chebyshev polynomial approximation requires $\mathcal{O}(\text{nnz}(\mathbf{A})P)$ flops, which is very small compared to implementing $g(\mathbf{A})\mathbf{y}$ in its original form.

III. Denoising simplicial complex signals

In this section, we state the problem of denoising simplicial complex signals with varying levels of smoothness w.r.t. the upper and lower Laplacian matrices. We then present its closed-form solution. To implement this solution in a distributed manner, we express the solution in terms of linear operators related to the simplicial complex and approximate it using Chebyshev polynomials. Finally, we show how the recurrence relation of the Chebyshev polynomials enables the distributed implementation.

A. Denoising with weighted total variation regularizer

Given a noisy $k-$simplicial signal $\mathbf{y}_k = \mathbf{x}_k + \mathbf{v}_k$ on a simplicial complex $\mathcal{S}$, where $\mathbf{x}_k \in \mathbb{R}^{N_k}$ is the true signal and $\mathbf{v}_k \in \mathbb{R}^{N_k}$ is the observation noise, we aim to denoise the observed signal $\mathbf{y}_k$ to recover $\mathbf{x}_k$. We assume that, the true signal $\mathbf{x}_k$ is dominantly smooth on either the lower or upper Laplacian matrix. Mathematically, we solve the following optimization problem

$$\hat{\mathbf{x}}_k = \arg\min_{\mathbf{x}_k} \|\mathbf{x}_k - \mathbf{y}_k\|_2^2 + \alpha_l \mathbf{x}_k^T \mathbf{L}_{kl} \mathbf{x}_k + \alpha_u \mathbf{x}_k^T \mathbf{L}_{ku} \mathbf{x}_k, \quad (7)$$

where $\hat{\mathbf{x}}_k \in \mathbb{R}^{N_k}$ is the estimate of the true signal $\mathbf{x}_k$. The second and third quadratic terms in the objective function measure the smoothness of $\mathbf{x}_k$ w.r.t. the lower and upper Laplacian matrices $\mathbf{L}_{kl}$ and $\mathbf{L}_{ku}$. The hyperparameters $\alpha_l > 0$ and $\alpha_u > 0$ control the amount of smoothness of $\mathbf{x}_k$ on $\mathbf{L}_{kl}$ and $\mathbf{L}_{ku}$, respectively. The optimization problem (7) is an unconstrained convex optimization problem. From the first-order optimality conditions, the optimal solution can be obtained in closed form as

$$\hat{\mathbf{x}}_k = \left(\mathbf{I} + \alpha_l \mathbf{L}_{kl} + \alpha_u \mathbf{L}_{ku}\right)^{-1} \mathbf{y}_k := \mathbf{H}_k \mathbf{y}_k, \quad (8)$$

where $\mathbf{H}_k$ is a simplicial complex denoising filter.

Let us define $\mathbf{A} = \alpha_l \mathbf{L}_{kl} + \alpha_u \mathbf{L}_{ku}$ with the eigenvalue decomposition $\mathbf{A} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^T$. Then the simplicial filter $\mathbf{H}_k$ can be expressed as a function of $\mathbf{A}$ as $\mathbf{H}_k = \mathbf{U} g_k(\boldsymbol{\Lambda})\mathbf{U}^T$, where $g_k(\lambda) = 1/(1+\lambda)$. From the recurrence relation of the Chebyshev polynomials, (8) can be implemented using a Pth order Chebyshev polynomial approximation which costs about $\mathcal{O}\left(P(\mathrm{nnz}(\mathbf{L}_{kl}) + \mathrm{nnz}(\mathbf{L}_{ku}))\right)$ flops. To achieve a desired approximation error, P should be chosen appropriately based on the smoothness of the simplicial filter $g_k(\lambda)$. Further, to implement this filter response in a distributed setting using the recurrence relation of Chebyshev polynomials, each $k-$simplex should scale its signal with α_l (respectively, α_u), while communicating with the lower (respectively, upper) adjacent neighbours. In a simplicial complex, a pair of $k-$simplices may appear as both the lower and upper adjacent neighbours. Thus a $k-$simplex should be able to choose the scaling factor appropriately while communicating with its neighbors and not just aggregate data as in a traditional distributed setting.

Since we focus on signals that are dominantly smooth on either the upper or lower Laplacian matrices, the communication costs can be reduced further. In what follows, we leverage properties of the Hodge Laplacian matrix to express the optimal solution (8) in terms of lower and upper simplicial complex multiplier operators. We then approximate them using the truncated Chebyshev polynomials of different orders to arrive at a communication-efficient distributed implementation. Before doing so, we make the following remark.

Remark 1. *When $\alpha_l = \alpha_u$, problem (7) boils down to the denoising problem in [3]. Henceforth, we refer to this case as unweighted total variation (`UTV`) based denoising. `UTV` penalizes both the smoothness promoting quadratic forms related to the lower and upper Laplacian matrices with the same weight. Thus, the `UTV` denoising approach is useful for denoising harmonic $k-$simplicial signals, where the signals are equally smooth on both $\mathbf{L}_{kl}$ and $\mathbf{L}_{ku}$. Furthermore, with $\alpha_l = \alpha_u = \alpha$, the simplicial filter $\mathbf{H}_k = \left(\mathbf{I} + \alpha\left(\mathbf{L}_{kl} + \mathbf{L}_{ku}\right)^{-1}\right) = \mathbf{U}_k \left(\mathbf{I} + \alpha \boldsymbol{\Lambda}_k\right)^{-1} \mathbf{U}_k^T$ can be approximated using Chebyshev polynomials and implemented in a distributed manner by scaling the signals while aggregating.*

B. Simplicial complex multiplier operator

For a square matrix $\mathbf{A} \in \mathbb{R}^{N\times N}$, we have $\left(\mathbf{I} + \mathbf{A}\right)^{-1} = \sum_{i=0}^{\infty}(-1)^i \mathbf{A}^i$. Using this property, we rewrite the matrix inverse in (8), which is a simplicial denoising filter $\mathbf{H}_k = \left(\mathbf{I} + \alpha_1 \mathbf{L}_{kl} + \alpha_2 \mathbf{L}_{ku}\right)^{-1}$ as

$$\begin{aligned}
\mathbf{H}_k &= \sum_{n=0}^{\infty}(-1)^n \left(\alpha_l \mathbf{L}_{kl} + \alpha_u \mathbf{L}_{ku}\right)^n \\
&\overset{(a)}{=} \mathbf{I} + \sum_{n=1}^{\infty}(-1)^n \alpha_l^n \mathbf{L}_{kl}^n + \sum_{n=1}^{\infty}(-1)^n \alpha_u^n \mathbf{L}_{ku}^n \\
&= \sum_{n=0}^{\infty}(-1)^n \alpha_l^n \mathbf{L}_{kl}^n + \sum_{n=0}^{\infty}(-1)^n \alpha_u^n \mathbf{L}_{ku}^n - \mathbf{I} \\
&\overset{(b)}{=} \underbrace{\left(\mathbf{I} + \alpha_l \mathbf{L}_{kl}\right)^{-1}}_{\mathbf{H}_{kl}} + \underbrace{\left(\mathbf{I} + \alpha_u \mathbf{L}_{ku}\right)^{-1} - \mathbf{I}}_{\mathbf{H}_{ku}}, \qquad (9)
\end{aligned}$$

where we use the fact that $\mathbf{L}_{kl}\mathbf{L}_{ku} = \mathbf{0}$ to arrive at (a) and the identity $\left(\mathbf{I} + \mathbf{A}\right)^{-1} = \sum_{i=0}^{\infty}(-1)^i \mathbf{A}^i$ in (b). Thus, we have $\mathbf{H}_k = g_l(\mathbf{L}_{kl}) + g_u(\mathbf{L}_{ku})$, where $g_l(\cdot)$ and $g_u(\cdot)$ are the functions defined on the eigenvalues of the matrices $\mathbf{L}_{kl}$ and $\mathbf{L}_{ku}$, given by $g_l(\lambda) = 1/(1+\alpha_l\lambda)$ and $g_u(\lambda) = -\alpha_u\lambda/(1+\alpha_u\lambda)$, respectively. The matrices $\mathbf{H}_{kl}$ and $\mathbf{H}_{ku}$ are the different simplicial complex filters related to the lower and upper Laplacian matrices.

C. Distributed implementation

In this section, we focus on implementing the `WTV` simplicial denoising filter in a distributed setting. The recurrence relation of the Chebyshev polynomials (5) allows us to implement the filter response $\mathbf{H}_k\mathbf{y}$ in a distributed setting. In other words, the aim is to compute $\left[\mathbf{H}_k\mathbf{y}_k\right]_n$ at each $k-$simplex by communicating its signal values with its lower and upper adjacent $k-$simplices.

To do so, in a simplicial complex, we assume that each $k-$simplex has the knowledge of its local neighbours, namely, its lower and upper adjacent $k-$simplices and can transport its signal $[\mathbf{y}_k]_n$ to them. From (9), we have $[\mathbf{H}_k\mathbf{y}_k]_n =$

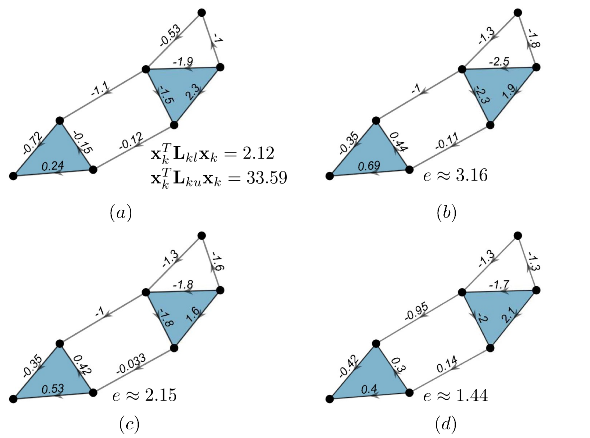

Fig. 1: Flow denoising on a simplicial complex. (a) A signal that is dominantly smooth on $\mathbf{L}_{kl}$ compared to $\mathbf{L}_{ku}$. (b) Observed noisy flow. (c) Denoised signal using `UTV` [10]. (d) Denoised signal using the proposed `WTV` method.

$[\mathbf{H}_{kl}\mathbf{y}_k]_n + [\mathbf{H}_{ku}\mathbf{y}_k]_n$. Since $\mathbf{H}_{kl}$ and $\mathbf{H}_{ku}$ are different simplicial complex filters, we use different approximation orders based on their smoothness, i.e., we approximate the filter responses $\mathbf{H}_{kl}\mathbf{y}$ and $\mathbf{H}_{ku}\mathbf{y}$ using P_l and P_u order Chebyshev polynomial approximation, respectively. We also assume that, each $k-$simplex can compute the Chebyshev polynomial coefficients $\{c_{p,l}\}_{p=0}^{P_l}$ and $\{c_{p,u}\}_{p=0}^{P_u}$ related to the lower and upper simplicial filters. Thus to compute $[\mathbf{H}_k\mathbf{y}_k]_n$, each $k-$simplex computes $\left\{\left[\bar{T}_p(\mathbf{L}_{kl})\mathbf{y}_k\right]_n\right\}_{p=1}^{P_l}$ and $\left\{\left[\bar{T}_p(\mathbf{L}_{ku})\mathbf{y}_k\right]_n\right\}_{p=1}^{P_u}$ using

$$\left[\bar{T}_p(\mathbf{L}_t)\mathbf{y}_k\right]_n = \frac{2}{\beta}\left[\mathbf{L}_t\bar{T}_{p-1}(\mathbf{L}_t)\mathbf{y}_k\right]_n - \frac{2\alpha}{\beta}\left[\bar{T}_{p-1}(\mathbf{L}_t)\mathbf{y}_k\right]_n - \left[\bar{T}_{p-2}(\mathbf{L}_t)\mathbf{y}_k\right]_n, \quad t \in \{kl, ku\},$$

and computes the sum as in (6). For $p \geq 1$ and $t \in \{kl, ku\}$, computing $\left[\bar{T}_p(\mathbf{L}_t)\mathbf{y}_k\right]_n$ at each simplex, requires communicating $\left[\bar{T}_{p-1}(\mathbf{L}_t)\mathbf{y}_k\right]_n$ with its local neighbours. This corresponds to one round of communication. For P_l and P_u communication rounds related to upper and lower simplicial complex filters, implementing $\mathbf{H}_k\mathbf{y}_k$ costs about $\mathcal{O}\left(P_l\mathrm{nnz}(\mathbf{L}_{kl}) + P_u\mathrm{nnz}(\mathbf{L}_{ku})\right)$.

IV. Numerical Experiments

In this section, we demonstrate the performance of the proposed weighted total variation (`WTV`) based denoising method and its distributed implementation using the truncated shifted Chebyshev polynomials on synthetic and real datasets. We compare the performance in terms of approximation error with state-of-the-art iterative methods. For the $k-$simplicial signal $\mathbf{x}_k \in \mathbb{R}^{N_k}$, the approximation error is defined as $e_q = \|\mathbf{x}_k^q - \mathbf{x}_k\|_2^2$, where $\mathbf{x}_k^q$ is the solution at the qth order Chebyshev polynomial approximation or the solution at the qth step of the iterative algorithm.

A. Synthetic dataset

To evaluate the performance of the `WTV` denoising method, we consider the simplicial complex shown in Fig. 1. The simplicial complex has 7 nodes, 10 edges, and 2 triangles. We generate a synthetic smooth flow signal on the simplicial complex by taking random linear combination of the eigenvectors $\mathbf{U}_1 \in \mathbb{R}^{10\times 10}$ of the Hodge Laplacian matrix $\mathbf{L}_1 \in \mathbb{R}^{10\times 10}$. The smooth flow signal is given by $\mathbf{x}_1 = \mathbf{U}_1\mathbf{z}$, where $\mathbf{z} \sim \mathcal{N}(0, \mathbf{I})$. The true flow is dominantly smooth on the lower Laplacian matrix $\mathbf{L}_{1l}$ as compared to the upper Laplacian matrix $\mathbf{L}_{1u}$ as shown in Fig. $1(a)$. To demonstrate the denoising performance, we add Gaussian noise of variance 0.5 to the true flow signal and it is as shown in Fig. $1(b)$. Given the noisy flow signal, assuming that the true flows are smooth on the Hodge Laplacian matrix, we estimate the true signal $\mathbf{x}_k$ from the observed noisy flows using `WTV` and `UTV` [3] methods from the closed-form solution in (8). For `WTV`, we choose $\alpha_1 = 0.5$ and $\alpha_2 = 0.05$, whereas for `UTV`, we choose $\alpha_1 = \alpha_2 = 0.1$. These chosen parameters yield the lowest error for the signal realization. The `UTV` method penalizes both the smoothness promoting quadratic cost related to the lower and upper Laplacian matrices equally. Thus, the approximation error for `UTV` is more compared to the proposed `WTV` method as can be see in in Figs. $1(c)$ and $1(d)$. Now, we compare the proposed distributed implementation of the filter $\mathbf{H}_k\mathbf{y}$ using the truncated Chebyshev polynomial approximation (6) with Jacobi method [13] and Jacobi with the Chebyshev acceleration method [14]. Instead of implementing a denoising filter response $\mathbf{H}_k\mathbf{y}_k$ directly, we can also use the Jacobi method solves a system of linear equations $\mathbf{Q}\mathbf{x}_k = \mathbf{y}_k$, where $\mathbf{Q} = \mathbf{H}_k^{-1}$. Starting with a random initial value $\mathbf{x}_k^0$, the Jacobi method iteratively updates its solution as

$$\mathbf{x}_k^{(t+1)} = \mathbf{Q}_D^{-1}\mathbf{Q}_O\mathbf{x}_k^{(t)} + \mathbf{Q}_D^{-1}\mathbf{y}, \;\; t = 0, 1, \ldots, T-1,$$

where the matrices $\mathbf{Q}_D$ and $\mathbf{Q}_O$ contain the diagonal and off-diagonal entries of $\mathbf{Q}$ and satisfy $\mathbf{Q} = \mathbf{Q}_D - \mathbf{Q}_O$. The Jacobi method converges, if and only if, the spectral radius of the matrix $\mathbf{Q}_D^{-1}\mathbf{Q}_O$ is less than one. Convergence of the Jacobi method can be accelerated by adding a momentum term to the Jacobi iterates, henceforth referred to as Jacobi with Chebyshev acceleration [14], [15].

In Fig. $2(a)$, we compare the approximation error $\|\mathbf{x}_k^q - \hat{\mathbf{x}}_k\|_2^2$ for different distributed implementations discussed above as a function of iteration index. Recall that, $\hat{\mathbf{x}}_k$ is the closed-form solution (8) for the `WTV` denoising problem. It is clear from Fig. $2(a)$ that the Chebyshev polynomial approximation converges faster to the true solution compared to the iterative counterparts.

In Fig. $2(b)$, we plot the resulting approximation errors $e_t = \|\hat{\mathbf{x}}_t - \tilde{\mathbf{x}}_t\|_2$, $t \in \{kl, ku\}$ for lower and upper simplicial complex filtering operations $\mathbf{H}_{kl}\mathbf{y}_k$ and $\mathbf{H}_{ku}\mathbf{y}_k$, respectively. As the lower simplicial filter response is more smooth in this experiment, it converges faster and requires fewer communication rounds as compared to the upper simplicial complex filter.

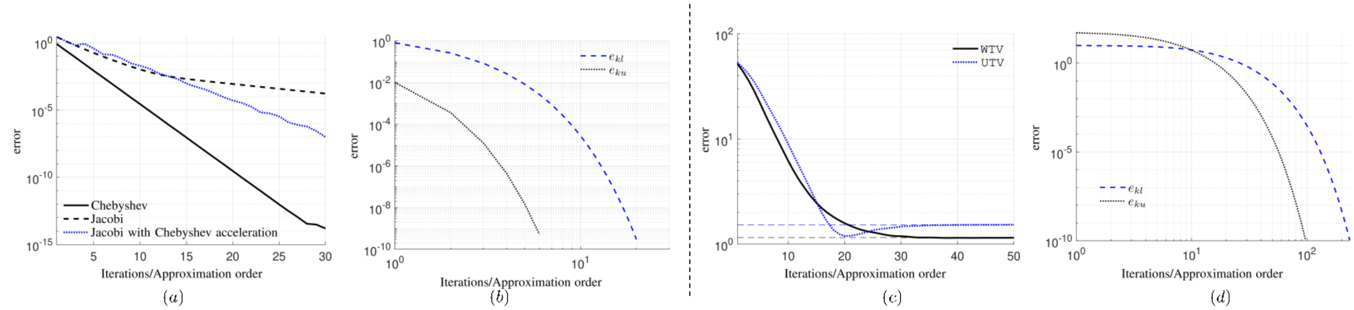

Fig. 2: Denoising on synthetic and real-world datasets. (a) `WTV` on the synthetic dataset. (b) Chebyshev polynomial approximation errors for lower and upper simplicial complex filters responses to indicate that we can use different approximation orders based on smoothness. (c) `UTV` and `WTV` on the primary school contact dataset. The dashed line indicates the errors $\|\hat{\mathbf{x}}_k - \mathbf{x}_k\|_2$ achieved from the closed-form solution using `UTV` and `WTV` methods. (d) Chebyshev polynomial approximation errors for lower and upper simplicial complex filters responses on the primary school contact dataset.

B. Real dataset

In this section, we evaluate the denoising performance of the `WTV` method on the primary school contact dataset [16]. This dataset is a collection of higher order simplices encoding the interactions between students in a primary school. For illustration, we limit our analysis to the 4th order simplices. The dataset has $N_0 = 242$, $N_1 = 8317$, $N_2 = 5139$, $N_3 = 381$, and $N_4 = 9$ higher-order simplices. The signal on each $k-$simplex is the number of interactions between $k+1$ students.

For illustration, we consider denoising signals over the 4th order simplices. We assume that the true signal is $\mathbf{x}_4 \in \mathbb{R}^{381}$ is smooth on the 4th order Hodge Laplacian matrix $\mathbf{L}_4 \in \mathbb{R}^{381\times 381}$. The amount of smoothness of $\mathbf{x}_4$ on the lower and upper Laplacian matrices are 4.12 and 0.0204, respectively. We add Gaussian noise of variance 5 to the true signal.

We solve the simplicial denoising problem in (7) using `WTV` and `UTV` in a distributed manner by implementing (6). For denoising using `WTV`, we choose $\alpha_1 = 30$, $\alpha_2 = 10$ and for `UTV`, we choose $\alpha_1 = \alpha_2 = 15$. The iterative methods Jacobi and Jacobi with Chebyshev acceleration do not converge as the spectral norm of the matrix $\mathbf{Q}_D^{-1}\mathbf{Q}_O$ for this dataset is greater than one. Hence, we do not report their performance. In Fig. 2(b), we plot the approximation error $\|\mathbf{x}_k^q - \mathbf{x}_k\|_2$ as a function of the approximation order P. It is clear from Fig. 2(c) that, for $P > 40$, both the `WTV` and `UTV` denoising methods converge to their corresponding closed-form solutions. Furthermore, `WTV` achieves a smaller denoising error than `UTV`. Figure 2(d) shows that to achieve an approximation error of 10^{-8} with $P_l = P_u = 131$ it costs about 174754 flops, whereas by choosing $P_l = 131$ and $P_u = 56$, we can achieve same error with about 157879 flops.

V. Conclusions

We developed a framework for denoising signals defined over simplicial complexes in a distributed setting. To denoise signals that are dominantly smooth w.r.t. either the upper or lower Hodge Laplacian matrices, we presented a weighted total variation regularized denoising solution and a computationally efficient solution based on the Chebyshev polynomial approximation. The performance of the proposed `WTV` denoising method is better than the `UTV` denoising method. We demonstrated the effectiveness of the proposed method on synthetic and real-world datasets. The computational complexity of the proposed distributed solution scales linearly with the number of higher-order simplices.

References

[1] A. Ortega, P. Frossard, J. Kovačević, J. M. Moura, and P. Vandergheynst, "Graph signal processing: Overview, challenges, and applications," *Proceedings of the 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] M. T. Schaub, Y. Zhu, J.-B. Seby, T. M. Roddenberry, and S. Segarra, "Signal processing on higher-order networks: Livin'on the edge... and beyond," *Signal Processing*, vol. 187, p. 108149, 2021.

[4] S. Barbarossa and S. Sardellitti, "Topological signal processing over simplicial complexes," *IEEE Trans. Signal Process.*, vol. 68, pp. 2992–3007, 2020.

[5] J. R. Munkres, *Elements of algebraic topology*. CRC press, 2018.

[6] A. R. Benson, D. F. Gleich, and J. Leskovec, "Higher-order organization of complex networks," *Science*, vol. 353, no. 6295, pp. 163–166, 2016.

[7] M. Yang, E. Isufi, M. T. Schaub, and G. Leus, "Finite impulse response filters for simplicial complexes," in *Proc. of the European Signal Process. Conf. (EUSIPCO)*, Dublin, Ireland, May 2021.

[8] S. Barbarossa, S. Sardellitti, and E. Ceci, "Learning from signals defined over simplicial complexes," in *Proc. of the IEEE Data Sci. Work.*, Lausanne, Switzerland, Jun. 2018.

[9] S. Barbarossa and S. Sardellitti, "Topological signal processing: Making sense of data building on multiway relations," *IEEE Signal Process. Mag.*, vol. 37, no. 6, pp. 174–183, 2020.

[10] M. T. Schaub and S. Segarra, "Flow smoothing and denoising: Graph signal processing in the edge-space," in *Proc. of the Global Conf. Signal and Info. Process. (GlobalSIP)*, California, USA, Nov. 2018.

[11] J. Jia, M. T. Schaub, S. Segarra, and A. R. Benson, "Graph-based semi-supervised & active learning for edge flows," in *Proc. of the ACM SIGKDD Int. Conf. on Knowl. Disc.Data Mini. (KDD)*, Anchorage, USA, 2019.

[12] L.-H. Lim, "Hodge Laplacians on graphs," *SIAM Review*, vol. 62, no. 3, pp. 685–715, 2020.

[13] Y. Saad, *Iterative methods for sparse linear systems*. SIAM, 2003.

[14] J. W. Demmel, *Applied numerical linear algebra*. SIAM, 1997.

[15] D. I. Shuman, P. Vandergheynst, D. Kressner, and P. Frossard, "Distributed signal processing via Chebyshev polynomial approximation," *IEEE Trans. Signal Inf. Process. Netw.*, vol. 4, no. 4, pp. 736–751, 2018.

[16] J. Stehlé, N. Voirin, A. Barrat, C. Cattuto, L. Isella, J.-F. Pinton, M. Quaggiotto, W. Van den Broeck, C. Régis, B. Lina *et al.*, "High-resolution measurements of face-to-face contact patterns in a primary school," *PloS one*, vol. 6, no. 8, p. e23176, 2011.