Passive and Active Sampling for Piecewise-Smooth Graph Signals

Rohan Varma
Dept. of Electrical and Computer Engineering
Carnegie Mellon University
Pittsburgh, U.S.A
rohanv@andrew.cmu.edu

Jelena Kovačević
Tandon School of Engineering
New York University
New York City, U.S.A
jelenak@nyu.edu

***Abstract*—In this work, we study the sampling of piecewise smooth-graph signals that exhibit an inhomogeneous level of smoothness over the graph and are characterized by have abrupt, localized discontinuities between smooth regions of the graph. We propose an extension to the graph trend filtering framework under the sampling setting and present an ADMM-algorithm to efficiently reconstruct piecewise-smooth graph signals. Further, to alleviate the limitations of passive sampling in this setting, we develop an active sampling strategy that incorporates feedback to focus the sampling procedure near the boundary or discontinuities. We then conduct experiments that exhibit the performance of our algorithm on large complex graphs and validate the efficacy of our sampling strategies.**

***Index Terms*—sampling, piecewise smooth, graph signal processing, graph trend filtering**

I. Introduction

With the explosive growth of information and communication, data is being generated at an unprecedented rate from various sources, including multimedia, sensor networks, biological systems, social networks, and physical infrastructure [1]. Research in graph signal processing (GSP) aims to develop tools for processing such data by providing a framework for the analysis of high-dimensional data signals we refer to as *graph signals* defined on irregular graph domains [2], [3], [4]. Research on the sampling and recovery of graph signals has been prevalent in recent years [5], [6], [7], [8], [9].

The assumption that graph signals vary slowly or are smooth over the graph is a natural one to make. However, in social networks, within a given community or social circle, users' profiles tend to be homogeneous, while within a different social circle they will be different, yet still homogeneous. Such signals are characterized by large variation between regions or pieces and slow variation within pieces. In this work, we study the sampling and reconstruction of such piecewise-smooth graph signals that exhibit a spatially inhomogeneous level of smoothness over regions of the graph and have abrupt, localized discontinuities. This class of piecewise-smooth signals is complementary to the class of smooth graph signals that exhibit spatially homogeneous levels of smoothness over the graph. The sampling of such smooth signals has been well studied in previous work both within the field of graph signal processing as well as in the context of Laplacian regularization.

In the context of semi-supervised classification on graphs, each vertex represents one data point to which a label is associated and a graph can be formed by connecting vertices with weights corresponding to the affinity or distance between the data points in some feature space. It is then natural to assume that the *label signal* is piecewise-smooth on the graph. Since samples are often sparse or expensive, designing efficient sampling and reconstruction tools for semi-supervised classification and active learning is notably valuable. The graph trend filtering (GTF) framework [10], which applies total variation denoising on graphs [11], is a particularly flexible and attractive approach to process piecewise-smooth graph signals that is based on minimizing the ℓ_1 norm of discrete graph differences. In this work, we present an extension to the GTF framework under the sampling setting, that is, where we only partially observe the signal.

Most sampling strategies fall under the umbrellas of either (1) passive sampling where there is no feedback and we simply sample the space without any knowledge of key signal characteristics, or (2) active sampling where we can incorporate feedback in a sequential process. Unlike sampling smooth signals that have no discontinuities, the localized nature of the discontinuities in piecewise-smooth signals make the detection of these discontinuities inherently decoupled from the global or neighborhood features of the graph signal. It then follows that the passive sampling of piecewise-smooth graph signals is a significantly harder or even futile task than the same for globally smooth signals. For the latter, it is often sufficient to sample such that we uniformly cover the space. Consequently, we propose studying the active sampling of piecewise-smooth signals by designing algorithms and strategies that incorporate feedback. Particularly, we develop active sampling methods that can capitalize on the localized nature of the boundary by focusing the sampling process in the estimated vicinity of the boundary.

II. Graph Signal Processing and piecewise-Smooth Graph Signals

A. Graphs

We consider a weighted undirected graph $G = (\mathcal{V}, \mathcal{E}, \mathbf{A})$, where $\mathcal{V} = \{v_1, \ldots, v_N\}$ is the set of nodes, $\mathcal{E} = \{e_1, \ldots, e_m\}$ is the set of edges, and $\mathbf{A} = [A_{j,k}] \in \mathbb{R}^{N\times N}$ is the *graph shift operator* [12], or the weighted adjacency matrix. The edge set $\mathcal{E}$ represents the connections of the undirected graph G, and the positive edge weight $A_{j,k}$ between nodes v_j and v_k measures the underlying relation between the jth and the kth node, such as a similarity, a dependency, or a communication pattern. Let a *graph signal* be defined as

$$\boldsymbol{\beta} \ = \ \left[\beta_1, \beta_2, \ldots, \beta_N\right]^T \in \mathbb{R}^n,$$

where β_i denotes the signal coefficient at the ith node.

Let $\boldsymbol{\Delta} \in \mathbb{R}^{m\times N}$ be the oriented incidence matrix of G, where each row corresponds to an edge. That is, if the edge $e_i = (j,k) \in \mathcal{E}$ connects the jth node to the kth node ($j < k$), the entries in the ith row of $\boldsymbol{\Delta}$ is then given as

$$\Delta_{i,\ell} = \begin{cases} -\sqrt{A_{j,k}}, & \ell = j; \\ \sqrt{A_{j,k}}, & \ell = k; \\ 0, & \text{otherwise} \end{cases}.$$

This work is supported in part by NSF under grants CCF-1563918.

The entries of the signal $\mathbf{\Delta}\boldsymbol{\beta} = [\sqrt{A_{j,k}}(\beta_k - \beta_j)]_{(j,k)\in\mathcal{E}}$ specifies the weighted pairwise differences of the graph signal over each edge. As a result, $\mathbf{\Delta}$ can be interpreted as a *graph difference operator*. In graph signal processing, a signal is called *globally smooth over a graph* G if $\|\mathbf{\Delta}\boldsymbol{\beta}\|_2^2 = \sum_{(j,k)\in\mathcal{E}} A_{j,k}(\beta_k - \beta_j)^2$ is small.

B. Piecewise-Smooth Graph Signals

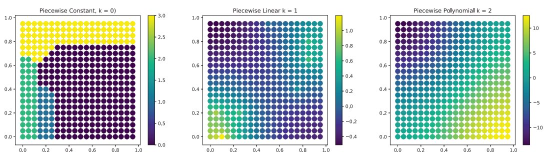

Fig. 1: From left to right, piecewise-constant (k=0), linear (k=1) and quadratic signals (k=2) on a 20×20 *2-dimensional grid graph (N = 400)*

In practice, the graph signal may not be necessarily smooth over the entire graph, but only locally within different pieces of the graph. To model inhomogeneous levels of smoothness over a graph, we say that a graph signal $\boldsymbol{\beta}$ is piecewise-constant over a graph G if many of the differences $\beta_k - \beta_j$ are zero for $(j,k) \in \mathcal{E}$. Consequently, the difference signal $\mathbf{\Delta}\boldsymbol{\beta}$ is sparse and $\|\mathbf{\Delta}\boldsymbol{\beta}\|_0$ is small.

We can generalize this notion to characterize *piecewise-kth order polynomial* signals on a graph, where the piecewise-constant case corresponds to $k = 0$, by generalizing the notion of graph difference operators. Specifically, we use the following recursive definition of the kth order graph difference operator $\mathbf{\Delta}^{(k+1)}$ [10]. Let $\mathbf{\Delta}^{(1)} = \mathbf{\Delta}$ for $k = 0$. For $k \geq 1$, let

$$\mathbf{\Delta}^{(k+1)} = \begin{cases} \mathbf{\Delta}^{(1)T}\mathbf{\Delta}^{(k)} \in \mathbb{R}^{N\times N}, & \text{odd } k \\ \mathbf{\Delta}^{(1)}\mathbf{\Delta}^{(k)} \in \mathbb{R}^{m\times N}, & \text{even } k \end{cases}.$$

The signal $\boldsymbol{\beta}$ is said to be a piecewise-kth order polynomial graph signal if $\|\mathbf{\Delta}^{(k+1)}\boldsymbol{\beta}\|_0$ is small. To further illustrate, let us consider the piecewise-linear graph signal, corresponding to $k = 1$, as a signal whose value at a node can be linearly interpolated from the weighted average of the values at neighboring nodes. It is easy to see that this is the same as requiring the second-order differences $\mathbf{\Delta}^T\mathbf{\Delta}\boldsymbol{\beta}$ to be sparse. Similarly, we say that a signal has a piecewise-quadratic structure if the differences between the second-order differences defined for piecewise-linear signals are mostly zero, that is, if $\mathbf{\Delta}\mathbf{\Delta}^T\mathbf{\Delta}\boldsymbol{\beta}$ is sparse. We illustrate examples of piecewise-kth order polynomial graph signals for $k = 0, 1, 2$ on a 2-dimensional grid graph in Fig 1.

C. Sampling

We consider the procedure of sampling and recovery as follows: we sample M coefficients in a graph signal $\boldsymbol{\beta} \in \mathbb{R}^N$ with Gaussian noise to produce a noisy sampled signal $\boldsymbol{y} \in \mathbb{R}^M (M < N)$, that is,

$$\boldsymbol{y} \quad = \quad \mathbf{\Psi}\boldsymbol{\beta} + \epsilon \;\equiv\; \boldsymbol{\beta}_{\mathcal{M}} + \epsilon, \tag{1}$$

where $\epsilon \sim \mathcal{N}(0, \sigma^2 \, \mathrm{I}_{M\times M})$, and $\mathcal{M} = (\mathcal{M}_1, \cdots, \mathcal{M}_M)$ denotes the sampling set where $\mathcal{M}_i \in \{1, \cdots, N\}$. The sampling operator $\mathbf{\Psi}$ is a linear mapping from $\mathbb{R}^N$ to $\mathbb{R}^M$, defined as

$$\mathbf{\Psi}_{i,j} = \begin{cases} 1, & j = \mathcal{M}_i; \\ 0, & \text{otherwise.} \end{cases} \tag{2}$$

We then reconstruct $\boldsymbol{\beta}$ from $\boldsymbol{y}$ to get $\widehat{\boldsymbol{\beta}} \in \mathbb{R}^N$.

Passive sampling refers to the setting where we are constrained to strategies that are blind to any samples of the signal. That is, we design strategies by considering only the underlying graph structure and any modeling assumptions we have made. In contrast, active or adaptive sampling strategies are able to choose samples in an online fashion by allowing feedback: the decision of where to sample next depends on all the observations made previously. While it's obvious that good active sampling strategies should never perform worse than passive sampling strategies, we aim to develop active sampling strategies that are able to achieve substantial gains in performance.

Under passive sampling, we can consider two different sampling settings: *random sampling* where the sample indices are chosen from from $\{1, \cdots, N\}$ independently and uniformly randomly; and *experimentally designed sampling* where the sample indices can be chosen beforehand based on the graph structure. Since we do not a-priori make any assumptions or have any information on the location of the boundary or discontinuities of the piecewise-smooth graph signal, we can show that experimentally designed sampling does not outperform random sampling and in fact, can often be detrimental. In other words, these discontinuities are fundamentally dissociated from samples outside their locations on the graph unlike globally smooth signals where key characteristics of the signal are *spread* out over local neighborhoods and consequently some nodes can be more informative than others. Consequently, we only consider uniform random sampling for passive sampling. We note that previous work that has studied the fundamental limits of passive and active sampling on graphs for globally smooth signals, has shown that active sampling *does not* fundamentally outperform passive sampling. However, experimentally designed sampling outperforms random sampling for irregular graphs where some nodes can be more informative than others.

III. Sampling and Recovery via GTF

Graph trend filtering (GTF) [10] is a flexible framework for estimation on graphs that is adaptive to inhomogeneity in the level of smoothness and localized characteristics of an observed signal across nodes. The kth order GTF estimate is defined as:

$$\hat{\boldsymbol{\beta}} = \mathbf{argmin}_{\boldsymbol{\beta}\in\mathbb{R}^n} \frac{1}{2}\|\boldsymbol{y} - \boldsymbol{\beta}\|_2^2 + \lambda\|\mathbf{\Delta}^{(k+1)}\boldsymbol{\beta}\|_1, \tag{3}$$

which can be regarded as applying total variation or fused lasso with the graph difference operator $\mathbf{\Delta}^{(k+1)}$ [11], [13]. The sparsity-promoting properties of the ℓ_1 norm have been well-studied [14]. Consequently, applying the ℓ_1 penalty in GTF sets many of the graph differences to zero while keeping a small fraction of nonzero values. GTF is then *adaptive* over the graph; its estimate at a node adapts to the smoothness in its localized neighborhood.

Under the sampling and recovery framework, we propose solving following modified version of the GTF formulation GTF-S:

$$\hat{\boldsymbol{\beta}} = \mathbf{argmin}_{\boldsymbol{\beta}\in\mathbb{R}^n} \frac{1}{2}\|\boldsymbol{y} - \mathbf{\Psi}\boldsymbol{\beta}\|_2^2 + \lambda\|\mathbf{\Delta}^{(k+1)}\boldsymbol{\beta}\|_1, \tag{4}$$

Remark 1. Note that we can use mixed piecewise penalties to encourage different kinds of piecewise polynomial behavior by stacking the graph difference matrices since we can transform $\lambda\|\mathbf{\Delta}^{(l+1)}\|_1 + \gamma\|\mathbf{\Delta}^{(m+1)}\|_1$ as $\|\mathbf{\Delta}\|_1$ where

$$\mathbf{\Delta} = \begin{bmatrix} \lambda\mathbf{\Delta}^{(l+1)} \\ \hline \gamma\mathbf{\Delta}^{(m+1)} \end{bmatrix}$$

In the following exposition however, we only consider the basic graph difference operator for a given k.

We solve this GTF-S formulation in (3) via the alternating direction method of multipliers (ADMM) framework for solving separable

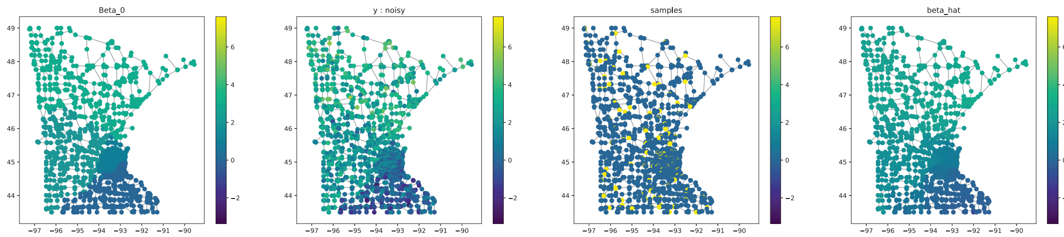

Fig. 2: Example of sampling a piecewise-constant (k=0) graph signal with 4 pieces on the Minnesota road graph with a random 5% of samples. From left to right, we have the true signal, the noisy signal, the location of the samples, and the reconstructed signal. Noisy input signal SNR = 5dB, Reconstructed signal SNR = 12.8dB

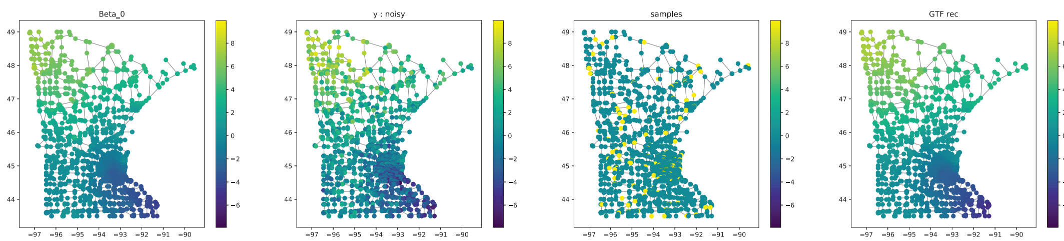

Fig. 3: Example of sampling a piecewise-linear (k=1) graph signal on the Minnesota road graph with a random 5% of samples. From left to right, we have the true signal, the noisy signal, the location of the samples, and the reconstructed signal. Noisy input signal SNR = 5dB, Reconstructed signal SNR = 14.5dB

optimization problems [15]. Via a change of variable defining $\boldsymbol{\eta} = \boldsymbol{\Delta}^{(k+1)}\boldsymbol{\beta}$, we can write the transformed problem:

$$\hat{\boldsymbol{\beta}} = \mathbf{argmin}_{\boldsymbol{\beta}\in\mathbb{R}^n} \frac{1}{2}\|\boldsymbol{y} - \boldsymbol{\Psi}\boldsymbol{\beta}\|_2^2 + \lambda\|\boldsymbol{\eta}\|_1 \quad \text{s.t.} \quad \boldsymbol{\eta} = \boldsymbol{\Delta}^{(k+1)}\boldsymbol{\beta}$$

and its corresponding Lagrangian as:

$$L(\boldsymbol{\beta}, \boldsymbol{\eta}, \boldsymbol{u}) = \frac{1}{2}\|\boldsymbol{y} - \boldsymbol{\Psi}\boldsymbol{\beta}\|_2^2 + \lambda\|\boldsymbol{\eta}\|_1 + \frac{\tau}{2}\|\boldsymbol{\Delta}^{(k+1)}\boldsymbol{\beta} - \boldsymbol{\eta} + \boldsymbol{u}\|_2^2 - \frac{\tau}{2}\|\boldsymbol{u}\|_2^2 \tag{5}$$

where $\boldsymbol{u}$ is the Lagrangian multiplier, and τ the parameter. Algorithm 1 shows the ADMM updates based on the Lagrangian in (5). For an appropriately chosen τ, the algorithm converges in a fixed number of iterations. In Fig. 2 and Fig. 3, we illustrate with

Algorithm 1 ADMM Optimization for GTF-S

Inputs: $\boldsymbol{y}$, $\boldsymbol{\Psi}$, $\boldsymbol{\Delta}^{(k+1)}$, and parameters λ, τ
Initialize:
  $\boldsymbol{D} \leftarrow \boldsymbol{\Delta}^{(k+1)}$, $\boldsymbol{\eta} \leftarrow \boldsymbol{D}\boldsymbol{\beta}$, $\boldsymbol{u} \leftarrow \boldsymbol{D}\boldsymbol{\beta} - \boldsymbol{\eta}$,
  $\boldsymbol{\beta} \leftarrow \boldsymbol{y}$ or $\boldsymbol{\beta}_{init}$ if given.
repeat
  $\boldsymbol{\beta} \leftarrow (\boldsymbol{\Psi}^T\boldsymbol{\Psi} + \tau\boldsymbol{D}^T\boldsymbol{D})^{-1}(\tau\boldsymbol{D}^T(\boldsymbol{\eta} - \boldsymbol{u}) + \boldsymbol{\Psi}^T\boldsymbol{y})$
  for `i ← 1 to length(`$\boldsymbol{D}\boldsymbol{\beta}$`)` **do**
    $\eta_i \leftarrow \text{prox}_\rho([\boldsymbol{D}\boldsymbol{\beta}]_i + u_i; \lambda/\tau)$
    ▷ $\text{prox}_\rho(t;\alpha)$ = soft-thresholding operator on t with $\alpha\rho$
  end for
  $\boldsymbol{u} \leftarrow \boldsymbol{u} + \boldsymbol{D}\boldsymbol{\beta} - \boldsymbol{\eta}$
until termination

an example the sampling and recovery of a piecewise-constant and piecewise-linear graph signal on the Minnesota road graph [16] with the GTF-S framework.

A. Theoretical Analysis

We now present bounds on the error rate of the GTF-S fit $\|\boldsymbol{\Psi}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}^\star)\|_2$ that help elucidate the relationship between the sample complexity (number of samples M needed for accurate reconstruction) with respect to structural properties of the graph and complexity of the boundary $\|\boldsymbol{\Delta}^{(k+1)}\beta\|_1$. For simplicity, let us assume the graph is fully connected, that is there is only 1 connected component, the dimension of the null space of $\boldsymbol{\Delta}^{(k+1)}$. Note that if there were multiple connected components, the problem becomes fully separable over each connected component.

Proposition 1. On a fully connected graph, we have $\boldsymbol{\Delta}^{(k+1)\dagger}\boldsymbol{\Delta}^{(k+1)} = \boldsymbol{I} - \frac{1}{N}\boldsymbol{J}$

Theorem 1 (Weak Consistency Error bound of the GTF-S minimizer)**.** Let $\widehat{\boldsymbol{\beta}}$ to be the minimizer of (4), r be the number of rows of $\boldsymbol{\Delta}^{(k+1)}$, ζ be the maximum ℓ_2 norm of the columns of $\boldsymbol{\Psi}\boldsymbol{\Delta}^{(k+1)\dagger}$. Set $\lambda = \sigma\zeta\sqrt{2\log(\frac{r}{\delta})}$, then with probability at least $1-2\delta$, we have:

$$\|\boldsymbol{\Psi}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}^\star)\|_2^2 \leq \sigma^2\left(1 + 2\sqrt{2\log(\frac{1}{\delta})}\right) + 4\sigma\zeta\sqrt{2\log(\frac{r}{\delta})}\|\boldsymbol{\Delta}^{(k+1)}\boldsymbol{\beta}^\star\|_1$$

Proof. Proof omitted due to lack of space □

Hence, the GTF-S fit is consistent if $\|\boldsymbol{\Delta}^{(k+1)}\boldsymbol{\beta}^\star\|_1$ grows at a rate slower than $\frac{1}{\sigma\zeta\sqrt{\log(\frac{r}{\delta})}}$. From [17], we can show that $\zeta \leq 1/\lambda_{\min}(\boldsymbol{\Delta}^{(2)})^{\frac{k+1}{2}}$, where $\lambda_{\min}(\boldsymbol{\Delta}^{(2)})$ is the smallest *nonzero* eigenvalue of the graph Laplacian matrix $\boldsymbol{\Delta}^{(2)}$ and quantifies the algebraic connectivity of the graph. We note that this result is consistent with basic error rates for graph trend filtering [10] as $M \to N$.

IV. Active Sampling

We expect localized behavior that may be hard to detect to hamper the performance of passive sampling strategies. Consequently, in this section, we seek to employ active sampling strategies when the signal exhibits inhomogenous behavior over the graph and contains discontinuities as in the case of piecewise-smooth signals. This gain in performance can be measured both in terms of the error rates and the sample complexity required to achieve a particular guarantee on the error. In spirit, our work follows previous work that studied the capabilities of passive and active sampling for recovering non-smooth functions from samples; the difference is that we consider a discrete setting and deal with irregular structures. For a smooth function, it has been shown that active sampling, experimentally designed sampling and uniform sampling have the same performance from a statistical perspective. For brevity, we only consider the piecewise-constant ($k = 0$) setting here, however the ideas and strategies presented here can easily be extended to general k. Given a sampling budget of M samples (assume for simplicity that M is even), we employ a two-step approach based in part on the active sampling procedures discussed in [18].

- In the first step, called the preview step, we randomly sample $M/2$ samples uniformly distributed over the graph with $\tilde{\Psi}$ and use the GTF-S estimator (4) of the signal to get the rough estimate $\tilde{\boldsymbol{\beta}}$.
- In the second step, called the refinement step, we select the remaining half of our budget $M/2$ samples, near the perceived locations of the boundaries estimated in the preview step. Particularly, in this step we define a probability distribution over the nodes such that $\pi_i \propto \sum_{j \in \mathcal{N}(i)} |\tilde{\boldsymbol{\beta}}_i - \tilde{\boldsymbol{\beta}}_j|$ where $\mathcal{N}(i)$ denotes the neighborhood of i, the nodes it shares an edge with. We sample $M/2$ nodes with replacement such that in each of the $M/2$ rounds, the probability of the i-th node being selected is proportional to π_i. Consequently, at the end of this process we can construct a randomized sampling set represented by $\widehat{\Psi}$ such that the samples are largely concentrated in the vicinity of the boundary or discontinuities. We then use the GTF-S estimator (4) with the full set of M samples with sampling operator $\Psi^T = \left[\tilde{\Psi}^T | \widehat{\Psi}^T\right]$ to get our final estimate $\widehat{\boldsymbol{\beta}}$.

This prescribed strategy is a natural way to take advantage of the idea that estimating the signal near the boundary is key to obtaining better reconstruction performance. We consider the even split of the sampling budget between the preview and refinement step only for simplicity. In addition, instead of a two-step procedure, one can reprise this idea, performing multiple refinement steps where in each step we acquire a new estimate of the boundary. However, for simplicity, we only consider the two-step procedure here.

V. Numerical Experiments

In this section, we perform numerical experiments on the synthetic piecewise-constant and piecewise-linear graph signals on the Minnesota road network with $N = 2642$ nodes and $m = 3304$ edges illustrated in Figures 2 and 3. We construct the piecewise-constant graph signal with 4 pieces by randomly choosing the location of 4 seed nodes and connecting every node to the closest seed by shortest path distance. We construct the piecewise-linear signal by randomly choosing the location of 50 discontinuities and solving the Poisson equation $\boldsymbol{\Delta}^{(2)}\boldsymbol{\beta} = \boldsymbol{b}$ where the non-zero entries in sparse vector $\boldsymbol{b}$ correspond to the discontinuities. We tune the hyperparameters ρ and τ in Algorithm 1 by grid-search for the below experiments.

A. Passive Sampling

In this section, we study the performance of our proposed algorithm for different sampling densities and noise settings. We inject white gaussian noise such that the noisy signal has a specified SNR (5dB,10dB,15dB) before uniformly randomly sampling the signal and recovering it with Algorithm 1. The results are illustrated in Figure 4.

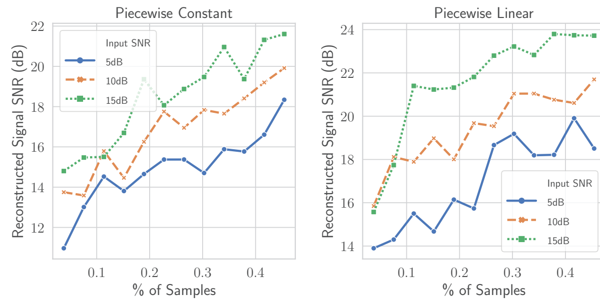

Fig. 4: Reconstructed signal SNR versus sampling density for different input SNR settings for both a piecewise-constant and a piecewise-linear graph signal

We see that we can only accurately reconstruct at moderately higher sampling densities. Note that we can reconstruct a piecewise-linear signal with better accuracy than a piecewise-constant signal with the same sample budget since the piecewise-linear graph signal is more homogenous and its key characteristics are less localized.

B. Active Sampling

We repeat the same experiment as that in Section V-A for the 5dB input SNR setting but additionally employ the active sampling strategy described in Section IV. The results are illustrated in Figure 5. Note that for piecewise-linear signals, in the refinement step, we define $\pi_i \propto |\tilde{\boldsymbol{\beta}}_i - \frac{1}{|\mathcal{N}_i|}\sum_{j \in \mathcal{N}_i} \tilde{\boldsymbol{\beta}}_j|$.

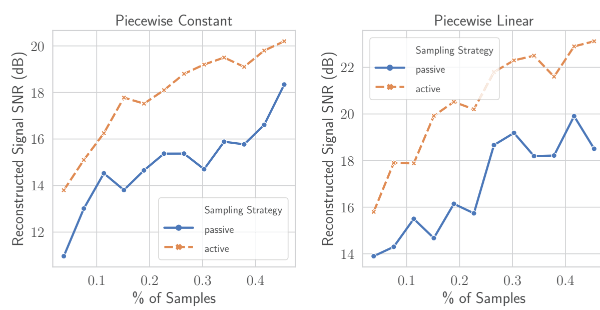

Fig. 5: Reconstructed signal SNR versus sampling density for passive and active sampling settings for both a piecewise-constant and a piecewise-linear graph signals

We see that for both piecewise-constant and piecewise-linear graph signals, active sampling consistently and significantly outperforms passive sampling for the same sampling budget. This performance gain is particularly substantial at lower sampling densities.

VI. Conclusion

In this work, we studied the sampling and recovery of piecewise smooth graph signals. Via an extension of the graph trend filtering framework, we presented an algorithm that allows us to efficiently sample and reconstruct such signals . Further, we developed an active sampling strategy that helps to better detect the location of the boundaries and discontinuities. We then presented numerical experiments on a real-world graph to test the performance of our sampling strategies and recovery algorithms.

REFERENCES

[1] M. Newman, *Networks*. Oxford University Press, 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 Processing Magazine*, vol. 30, no. 3, pp. 83–98, 2013.

[3] 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.

[4] A. Sandryhaila and J. M. Moura, "Big data analysis with signal processing on graphs: Representation and processing of massive data sets with irregular structure," *IEEE Signal Processing Magazine*, vol. 31, no. 5, pp. 80–90, 2014.

[5] S. Chen, R. Varma, A. Sandryhaila, and J. Kovačević, "Discrete Signal Processing on Graphs: Sampling Theory," *IEEE transactions on signal processing*, vol. 63, no. 24, pp. 6510–6523, 2015.

[6] R. Varma and J. Kovačević, "Sampling Theory for Graph Signals on Product Graphs," *arXiv preprint arXiv:1809.10049*, 2018.

[7] M. Tsitsvero, S. Barbarossa, and P. Di Lorenzo, "Signals on graphs: Uncertainty principle and sampling," *IEEE Transactions on Signal Processing*, vol. 64, no. 18, pp. 4845–4860, 2016.

[8] G. Puy, N. Tremblay, R. Gribonval, and P. Vandergheynst, "Random sampling of bandlimited signals on graphs," *Applied and Computational Harmonic Analysis*, vol. 44, no. 2, pp. 446–475, 2018.

[9] S. Chen, R. Varma, A. Singh, and J. Kovačević, "Signal recovery on graphs: Fundamental limits of sampling strategies," *IEEE Transactions on Signal and Information Processing over Networks*, vol. 2, no. 4, pp. 539–554, 2016.

[10] Y.-X. Wang, J. Sharpnack, A. J. Smola, and R. J. Tibshirani, "Trend filtering on graphs," *The Journal of Machine Learning Research*, vol. 17, no. 1, pp. 3651–3691, 2016.

[11] S.-J. Kim, K. Koh, S. Boyd, and D. Gorinevsky, "\ell_1 Trend Filtering," *SIAM Review*, vol. 51, no. 2, pp. 339–360, 2009.

[12] A. Sandryhaila and J. M. Moura, "Discrete signal processing on graphs," *IEEE Transactions on Signal Processing*, vol. 61, no. 7, pp. 1644–1656, 2013.

[13] R. J. Tibshirani, *The solution path of the generalized lasso*. Stanford University, 2011.

[14] R. Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 267–288, 1996.

[15] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends in Machine learning*, vol. 3, no. 1, pp. 1–122, 2011.

[16] N. Perraudin, J. Paratte, D. Shuman, L. Martin, V. Kalofolias, P. Vandergheynst, and D. K. Hammond, "GSPBOX: A toolbox for signal processing on graphs," *arXiv preprint arXiv:1408.5781*, 2014.

[17] F. R. Chung and F. C. Graham, *Spectral graph theory*. American Mathematical Soc., 1997.

[18] R. M. Castro and R. D. Nowak, "Minimax bounds for active learning," *IEEE Transactions on Information Theory*, vol. 54, no. 5, pp. 2339–2353, 2008.