

# FLOW SMOOTHING AND DENOISING: GRAPH SIGNAL PROCESSING IN THE EDGE-SPACE

Michael T. Schaub<sup>a,b</sup> and Santiago Segarra<sup>a</sup>

<sup>a</sup>Institute for Data, Systems, and Society, Massachusetts Institute of Technology, Cambridge, MA, USA

<sup>b</sup>Department of Engineering Science, University of Oxford, Oxford, UK

## ABSTRACT

This paper focuses on devising graph signal processing tools for the treatment of data defined on the edges of a graph. We first show that conventional tools from graph signal processing may not be suitable for the analysis of such signals. More specifically, we discuss how the underlying notion of a ‘smooth signal’ inherited from (the typically considered variants of) the graph Laplacian are not suitable when dealing with edge signals that encode a notion of flow. To overcome this limitation we introduce a class of filters based on the Edge-Laplacian, a special case of the Hodge-Laplacian for simplicial complexes of order one. We demonstrate how this Edge-Laplacian leads to low-pass filters that enforce (approximate) flow-conservation in the processed signals. Moreover, we show how these new filters can be combined with more classical Laplacian-based processing methods on the line-graph. Finally, we illustrate the developed tools by denoising synthetic traffic flows on the London street network.

**Index Terms**— Graph Signal Processing, Hodge-Laplacian, Simplicial Complexes, Flow Denoising.

## I. INTRODUCTION

As the availability of relational data continues to increase, graph-based techniques to model, filter, and process such data have become a mainstay in the current literature, cutting across disciplines. Consequentially, graph signal processing (GSP), which aims to provide a theoretical grounding for the processing of signals defined on graphs, has become a rapidly growing area of research [1], [2]. However, while significant research activity has been devoted to the development of foundational tools for signal processing on graphs, including algorithms for sampling [3]–[5], reconstruction [6]–[8], and filter design [9]–[11], most research to date has focused on signals defined on the nodes of the graph.

However, in many problems modeled using graphs, the data of interest is located on the edges (as opposed to the nodes) of these graphs. A typical scenario of practical interest is a flow on the edges – signal, mass, energy, information – of a graph that is measured and has to be analyzed further, such as traffic flow associated with the edges of a traffic network [12]. Further prototypical examples include data emerging from spatially embedded networks such as supply networks (e.g., power grids) [13], mobility and migration data [14], or information flows in brain networks or other biological tissue [15]–[19]. While spatially embedded networks provide many of the most intuitive examples, edge-flows are also the natural

objects of interest for web-traffic, click-stream data, and various other communication patterns that are not embedded in any type of physical space.

Despite this large space of applications, there has been comparably little activity related to edge-based analysis of signals on graphs. One reason for this lack of research activity may be that such situations appear to be transformable into a vertex-based problem by using a line-graph transformation [20]. Line-graphs, which record adjacency relationships between edges, provide indeed some potential modeling strategy to consider edge data. For instance, line-graphs have been considered for the task of detecting overlapping communities, by clustering edges [21], [22].

However, as we discuss in this paper, line-graphs are not sufficient as modeling tools for edge data that comes in the form of flows. As we will see, one deeper underlying reason for this fact is that certain notions from ‘classical’, vertex-based GSP do not seamlessly extend from vertex to edge data: unlike vertex data, flows typically carry an *orientation*, and accounting appropriately for such orientations can be crucial to reach the desired results. Interestingly, the edge-space processing perspective that we advocate here may be interpreted in the sense of processing ‘higher-order’ data [23]–[25], and can be generalized further by calling upon tools from algebraic topology [26] and discrete calculus [27], thereby opening up a range of interesting avenues for future research. However, to motivate our work and provide a concrete problem context, in this paper we focus on the denoising of edge-flow data as a specific example, leaving data defined in higher-order simplicial complexes as a matter of future research.

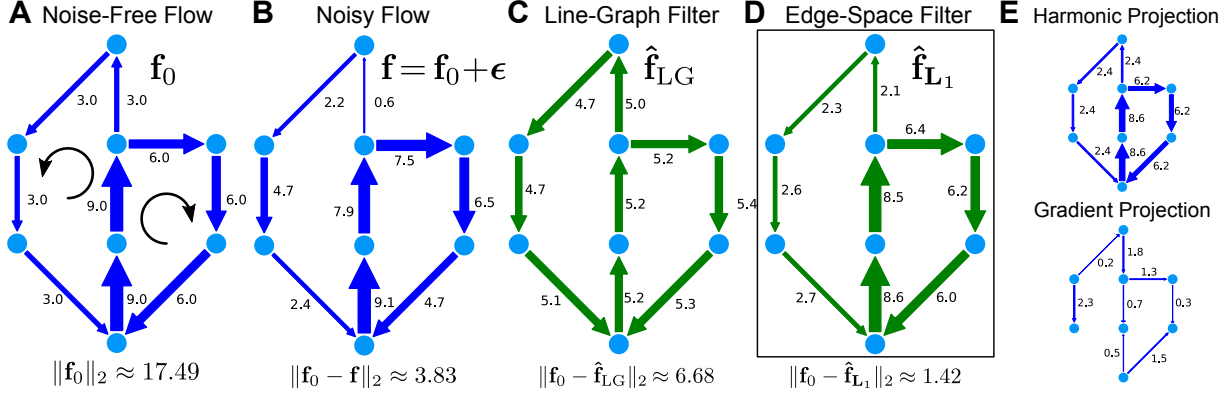
**Contributions and paper outline.** We introduce the problem of flow denoising and smoothing for edge-space data. Specifically, we discuss how the space of edge-flows decomposes into cyclic (harmonic) and gradient flows, and how this leads to a novel notion of low-pass filter for oriented edge-data. Our proposed filters are based on the Edge-Laplacian  $\mathbf{L}_1$ , a particular form of the more general Hodge-Laplacian encountered in algebraic topology [28]. We show how the Edge-Laplacian  $\mathbf{L}_1$  solves an optimal filtering problem that aims to filter out gradient flows in the edge-space, and contrast it with the more typically encountered graph Laplacian  $\mathbf{L}$ .

We start in Section II by recalling fundamental notions from GSP. In Section III we first provide an example that motivates the need for edge-space filtering, before discussing some of its various properties more formally in Section IV. We finally demonstrate the utility of the developed methodology by denoising traffic flows in the London street network (Section V), before concluding with a brief discussion on avenues for future research.

## II. BACKGROUND

**Graphs, incidence matrices, and the graph Laplacian.** An undirected graph  $\mathcal{G}$  consists of a node set  $\mathcal{N}$  of cardinality  $N$  and an

MTS received funding from the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 702410. SS received an IDSS MIT seed grant. The funders had no role in the design of this study; the results presented here reflect solely the authors’ views. Emails: {mschaub, segarra}@mit.edu



**Fig. 1: Flow smoothing on a graph – illustrative example.** **A** An undirected graph with a pre-defined, oriented flow pattern  $\mathbf{f}_0$  on the edges. **B** The observation of the flow  $\mathbf{f}_0$  is distorted by a Gaussian white noise vector  $\epsilon$ . **C-D** We may want to denoise the flow by applying a Laplacian filter based on the line-graph of this topology (**C**). However, this does not yields a satisfactory performance, in particular compared to the edge-space filter approach (**D**) introduced in this paper. **E** The originally observed flow  $\mathbf{f}$  can be decomposed into its harmonic (cyclic) component and its gradient part. Notice that the edge-space filter acts by reducing the gradient flow in the input edge-data.

edge set  $\mathcal{E}$  of unordered pairs of elements in  $\mathcal{N}$  of cardinality  $E$ . The edges can be conveniently collected as entries of the symmetric adjacency matrix  $\mathbf{A}$ , such that  $A_{ij} = A_{ji} = 1$  for all  $(i, j) \in \mathcal{E}$ , and  $A_{ij} = 0$  otherwise. Defining the diagonal degree matrix  $\mathbf{D} := \text{diag}(\mathbf{A}\mathbf{1})$ , the (combinatorial) Laplacian matrix associated with  $\mathcal{G}$  is given by  $\mathbf{L} := \mathbf{D} - \mathbf{A}$ , and has eigendecomposition  $\mathbf{L} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^\top$  where the eigenvalues in the diagonal matrix  $\mathbf{\Lambda}$  are sorted in increasing order.

An alternative way to encode a graph is by means of a node-to-edge incidence matrix  $\mathbf{B} \in \mathbb{R}^{N \times E}$ . To this end, we endow each edge  $e$  with an arbitrary reference orientation from its tail node  $t(e)$  to its head node  $h(e)$ . The incidence matrix is then defined as  $\mathbf{B}_{e t(e)} = -1$ ,  $\mathbf{B}_{e h(e)} = 1$  and  $\mathbf{B}_{ek} = 0$  otherwise. Based on this definition, it can be shown that the graph Laplacian can be alternatively written as  $\mathbf{L} = \mathbf{B}\mathbf{B}^\top$ .

**Laplacian smoothing and denoising on graphs.** Consider the observation of a noisy signal vector  $\mathbf{y} = \mathbf{y}_0 + \epsilon \in \mathbb{R}^N$  defined on the nodes of a connected graph  $\mathcal{G}$ , where  $\mathbf{y}_0$  is a certain signal of interest and  $\epsilon$  represents zero-mean additive noise. A canonical task in GSP is to provide a filtered signal  $\hat{\mathbf{y}} \approx \mathbf{y}_0$ , based on the observation of the noisy  $\mathbf{y}$ .

A typical assumption made in this context is that the signal should be smooth with respect to the underlying graph. To leverage this information we can define the following denoising optimization problem

$$\min_{\hat{\mathbf{y}}} \{ \|\hat{\mathbf{y}} - \mathbf{y}\|_2^2 + \alpha \hat{\mathbf{y}}^\top \mathbf{L} \hat{\mathbf{y}} \}, \quad (1)$$

where  $\alpha > 0$  can be understood as a regularization parameter that regulates the influence of the smoothness promoting regularizer  $\hat{\mathbf{y}}^\top \mathbf{L} \hat{\mathbf{y}} = \sum_{i,j} A_{ij} (\hat{y}_i - \hat{y}_j)^2$ . Notice that this quadratic form can be interpreted as a measure of the variation of the signal on the graph (measured along the edges) and is minimized by a constant vector  $\hat{\mathbf{y}} \propto \mathbf{1}$ . The optimal solution of the above optimization problem is given by

$$\hat{\mathbf{y}} = (\mathbf{I} + \alpha \mathbf{L})^{-1} \mathbf{y} \quad (\text{node denoising}). \quad (2)$$

Instead of solving the above problem, another strategy to obtain a filtered signal  $\hat{\mathbf{y}}$  is to use a simpler iterative smoothing operation

for a suitably chosen update parameter  $\mu$  and a certain fixed number of rounds  $k$

$$\hat{\mathbf{y}} = (\mathbf{I} - \mu \mathbf{L})^k \mathbf{y} \quad (\text{node smoothing}). \quad (3)$$

Both the denoising and smoothing procedures respectively defined in (2) and (3) are specific instances of *low-pass graph filters*. A graph filter  $\mathbf{H}$  is a linear map between graph signals that can be expressed as a matrix polynomial of  $\mathbf{L}$ . A graph filter  $\mathbf{H}$  is low-pass if its frequency response  $\hat{\mathbf{h}} = \text{diag}(\mathbf{V}^\top \mathbf{H} \mathbf{V})$  is a vector of decreasing values. It is immediate to corroborate that the filters in (2) and (3) are indeed low-pass filters.

**Line-graphs and their algebraic representation.** The line-graph of a graph  $\mathcal{G}$  is the graph  $\mathcal{G}_{\text{LG}}$  whose nodes correspond to the edges of  $\mathcal{G}$ . Two nodes in  $\mathcal{G}_{\text{LG}}$  are connected if the corresponding edges in the original graph  $\mathcal{G}$  share an incident node. Given the incidence matrix  $\mathbf{B}$  of the original graph, the adjacency matrix of the line-graph can be represented as  $\mathbf{A}_{\text{LG}} := |\mathbf{B}^\top \mathbf{B} - 2\mathbf{I}|$ , where the absolute value is applied elementwise. The graph Laplacian of the line-graph is correspondingly defined as  $\mathbf{L}_{\text{LG}} = \text{diag}(\mathbf{A}_{\text{LG}}\mathbf{1}) - \mathbf{A}_{\text{LG}} = \mathbf{B}_{\text{LG}}\mathbf{B}_{\text{LG}}^\top$ , where  $\mathbf{B}_{\text{LG}}$  is the incidence matrix of the line-graph.

### III. EDGE-SPACE FILTERING – AN ILLUSTRATIVE EXAMPLE

To understand the need for a theory of GSP for signals defined on the edges of a graph, let us consider the example graph in Fig. 1. We assume that the (a priori undirected) graph has been endowed with oriented flows  $\mathbf{f}_0$  on the edges as displayed in Fig. 1A. However, we only get to observe a perturbed version  $\mathbf{f} = \mathbf{f}_0 + \epsilon$  that has been distorted by white Gaussian noise  $\epsilon$  (Fig. 1B). Our goal is now to *smooth* this observed signal  $\mathbf{f}$  to get a better estimate of the original flows  $\mathbf{f}_0$ .

Translating the techniques previewed in (2) and (3) into edge-space analysis, we may apply these procedures by treating the observed edge-flows  $\mathbf{f}$  as node-data on the line-graph  $\mathcal{G}_{\text{LG}}$  associated with the original network. Fig. 1C shows the result of applying the line-graph filtering as in (3) substituting  $\mathbf{L}_{\text{LG}}$  for  $\mathbf{L}$ , and setting  $k = 10$  and  $\mu = 1/5$ . The differences between adjacent

edge signals are smoothed-out, and the output signal  $\hat{\mathbf{f}}_{LG}$  is driven towards the global average of the input flows  $\mathbf{f}$ . Indeed, we may interpret each smoothing operation  $(\mathbf{I} - \mu \mathbf{L}_{LG})$  as a gradient descent step for the objective  $\min_{\mathbf{f}} \|\mathbf{B}_{LG}^\top \mathbf{f}\|^2$ , where each iteration drives down the difference between adjacent edge flows. For a connected line-graph, this (local) averaging behavior leads ultimately towards a global averaging of the flows, akin to a consensus protocol.

Although the above discussion provides a good understanding of the line-graph filtering operation, the filtering results are clearly not satisfactory for our considered example: the structure of the initial flows  $\mathbf{f}_0$  has been distorted, and the error between the noise-free and the filtered signal  $\|\mathbf{f}_0 - \hat{\mathbf{f}}_{LG}\|_2 \approx 6.68$  has in fact *increased* when compared to the unfiltered signal  $\|\mathbf{f}_0 - \mathbf{f}\|_2 \approx 3.83$ . The problem here stems from the notion of smoothness that we inherited from the node-based filter, where low-pass signals are associated with the small eigenvalues of the graph Laplacian. This assumption is well matched to applications in which we expect small variation of the signal across connected nodes, e.g., if the graph corresponds to a network of sensors that aims to monitor a smoothly varying scalar field such as a temperature. However, in the case of flow data considered here this notion of a smooth signal is not desirable. Arguably, a notion of smoothness for flow data should capture the idea of flow conservation, i.e., input and output at every node should (approximately) balance out. Based on this view, cyclic flows such as depicted in Fig. 1A would correspond to smooth flows.

Interestingly, precisely this notion of smooth flows can be captured by the so called Edge-Laplacian defined as  $\mathbf{L}_1 := \mathbf{B}^\top \mathbf{B}$ . The Edge-Laplacian, which may be interpreted as a special case of the more general Hodge-Laplacian, has many properties that render this operator especially useful for flow-filtering, as we discuss in the next section.

#### IV. EDGE-SPACE FILTERING – THEORY

The  $\mathbf{L}_1$  Edge-Laplacian as an object on its own interest has been studied so far only sparsely in the literature, e.g., in the context of consensus protocols [29], or as a special case of the more general Hodge-Laplacian (see e.g. [28], [30], [31]). In the context of GSP, it has been recently considered in an independent work [24]. We now present a number of important properties of this linear operator.

**Proposition 1** *Given a graph with Edge-Laplacian  $\mathbf{L}_1 = \mathbf{B}^\top \mathbf{B}$ , the following properties hold:*

- 1) *The null space of  $\mathbf{L}_1$  is equivalent to the cycle space  $\mathcal{F}_C = \ker(\mathbf{B})$ , also known as the space of harmonic or cyclic flows.*
- 2) *The image of  $\mathbf{L}_1$  is equivalent to the gradient space  $\mathcal{F}_G = \text{im}(\mathbf{B}^\top)$ , also known as the cut space or space of potential flows.*

**Proof:** To prove (1), note that if  $\mathbf{z} \in \ker(\mathbf{B})$  then  $\mathbf{z} \in \ker(\mathbf{L}_1)$ , from the definition of  $\mathbf{L}_1$ . For the converse, assume that  $\mathbf{L}_1 \mathbf{z} = 0$ , then  $\mathbf{z}^\top \mathbf{L}_1 \mathbf{z} = \|\mathbf{Bz}\|_2^2 = 0$ , so that  $\mathbf{z} \in \ker(\mathbf{B})$  concluding the proof.

For (2), we proceed similarly. Clearly,  $\text{im}(\mathbf{L}_1) \subset \text{im}(\mathbf{B}^\top)$ . Conversely, for every non-zero  $\mathbf{z} \in \text{im}(\mathbf{B}^\top)$  we can find some  $\mathbf{y} \perp \ker(\mathbf{B}^\top)$  such that  $\mathbf{z} = \mathbf{B}^\top \mathbf{y} \neq 0$ . Since this implies that  $\mathbf{y} \in \text{im}(\mathbf{B})$ , there exists some  $\mathbf{x}$  such that  $\mathbf{z} = \mathbf{B}^\top \mathbf{Bx} = \mathbf{L}_1 \mathbf{x}$ , and hence  $\text{im}(\mathbf{B}^\top) \subset \text{im}(\mathbf{L}_1)$ . ■

Note that the name cycle space for  $\mathcal{F}_C$  derives from the fact that  $\ker(\mathbf{B})$  is the set of edge-vectors corresponding to (oriented) cycles

on the graph, i.e. edge-vectors that fulfill a flow conservation constraint at every node. It can further be shown that the dimension of  $\mathcal{F}_C$  is equal to the number of independent cycles in the graph [20]. The gradient space  $\mathcal{F}_G$  derives its name from the fact that it corresponds exactly to all those flow vectors  $\mathbf{f}_\phi = \mathbf{B}^\top \phi \in \mathbb{R}^E$  that can be induced as the difference of a scalar potential function  $\phi \in \mathbb{R}^N$  defined on the nodes.

Proposition 1 hints at the fact that an ideal low-pass filter  $\mathbf{H}_{LP}$  defined on  $\mathbf{L}_1$  should cancel out all non-cyclic flows. This intuition can be formalized as follows. Denote by  $\lambda_i(\mathbf{L}_1)$  the  $i$ -th eigenvalue of the matrix  $\mathbf{L}_1$ , then defining the frequency response  $\tilde{\mathbf{h}}_{LP}$  of an ideal low-pass filter (cf. Section II) as

$$[\tilde{\mathbf{h}}_{LP}]_i = \begin{cases} 1 & \text{if } \lambda_i(\mathbf{L}_1) = 0, \\ 0 & \text{otherwise,} \end{cases} \quad (4)$$

the following result holds.

**Proposition 2** *Given any edge-signal  $\mathbf{f}$ , it holds that*

$$\mathbf{H}_{LP} \mathbf{f} = \underset{\hat{\mathbf{f}}}{\text{argmin}} \|\hat{\mathbf{f}} - \mathbf{f}\|_2^2 \quad \text{s.t. } \mathbf{B}\hat{\mathbf{f}} = 0. \quad (5)$$

**Proof:** Consider the Lagrangian  $\mathcal{L}(\hat{\mathbf{f}}, \boldsymbol{\nu}) = \|\hat{\mathbf{f}} - \mathbf{f}\|_2^2 + \boldsymbol{\nu}^\top \mathbf{B}\hat{\mathbf{f}}$  associated with problem (5), and denote by  $\hat{\mathbf{f}}^*$  and  $\boldsymbol{\nu}^*$  the optimal primal and dual solutions of the problem. From the conditions of optimality it follows that

$$\frac{\partial \mathcal{L}}{\partial \hat{\mathbf{f}}} = 2(\hat{\mathbf{f}}^* - \mathbf{f}) + \mathbf{B}^\top \boldsymbol{\nu}^* = 0, \quad (6)$$

$$\frac{\partial \mathcal{L}}{\partial \boldsymbol{\nu}} = \mathbf{B}\hat{\mathbf{f}}^* = 0. \quad (7)$$

Left-multiplying (6) by  $\mathbf{B}$ , and substituting (7) into the obtained expression, it follows that  $\mathbf{B}\mathbf{B}^\top \boldsymbol{\nu}^* = 2\mathbf{B}\mathbf{f}$ . From here we see that an optimal dual solution for  $\boldsymbol{\nu}$  is given by the node potentials  $\boldsymbol{\nu}^* = 2\mathbf{L}^\dagger \mathbf{B}\mathbf{f}$ . Plugging this equality back into (6), and using the properties of the pseudoinverse it immediately follows that  $\hat{\mathbf{f}}^* = (\mathbf{I} - \mathbf{B}^\top \mathbf{L}^\dagger \mathbf{B})\mathbf{f} = (\mathbf{I} - \mathbf{B}^\dagger \mathbf{B})\mathbf{f}$ , indicating that  $\hat{\mathbf{f}}^*$  is the orthogonal projection of  $\mathbf{f}$  onto the kernel of  $\mathbf{B}$ . Now (5) follows from Proposition 1. ■

From the frequency response in (4) it follows that the filter  $\mathbf{H}_{LP}$  outputs a projection of the input onto the kernel space of  $\mathbf{L}_1$ . Proposition 2 formalizes the action of this filter from a denoising viewpoint. More precisely, the constraint  $\mathbf{B}\hat{\mathbf{f}} = 0$  guarantees that  $\hat{\mathbf{f}}$  represents a flow-preserving edge signal. Thus, in (5) we are seeking for the closest signal to  $\mathbf{f}$  that is flow preserving. Proposition 2 shows that such a signal is obtained by filtering out all the non-cyclic components present in  $\mathbf{f}$ .

**Low-pass filtering in the edge-space.** Proposition 2 notwithstanding, in many scenarios ideal low-pass filters are either undesirable or untenable in practice. Thus, leveraging the result in Proposition 1, we can define the analogues to (2) and (3) for filtering flow signals  $\mathbf{f}$  in the edges. More precisely, we have

$$\hat{\mathbf{f}} = (\mathbf{I} + \alpha \mathbf{L}_1)^{-1} \mathbf{f} \quad (\text{flow denoising}), \quad (8)$$

where (8) can be deemed as the solution of a regularized least-squares recovery akin to (1), and

$$\hat{\mathbf{f}} = (\mathbf{I} - \mu \mathbf{L}_1)^k \mathbf{f} \quad (\text{flow smoothing}), \quad (9)$$

which can be applied iteratively via  $k$  successive applications of a single smoothing operator  $(\mathbf{I} - \mu \mathbf{L}_1)$ .

In order to illustrate the above filtering operations, let us apply (9) with  $\mu = 1/5$  and  $k = 10$  to the noisy flow in Fig. 1B, leading to the result in Fig. 1D. As we can see, in this case the filtering leads to a much more desirable outcome: the cyclic structure of the original flow is essentially recovered, and the signal error  $\|\mathbf{f}_0 - \hat{\mathbf{f}}_{L_1}\|_2$  is reduced when compared to the original signal. Formally, from Proposition 1 we can conclude that the filtering operations (8) and (9) drive the input towards the cycle space  $\mathcal{F}_C$ . To illustrate this, we further plot the decomposition of the observed flows  $\mathbf{f}$  onto the cycle (harmonic) and the gradient spaces in Fig. 1E. Crucially, the cycle space is here of dimension 2, meaning that the filter accounts for the two ‘degrees of freedom’ in the original flow profiles (given by the two cycles), which explains the success of the filter in approximating the original flows. In contrast, the null-space of the line-graph Laplacian  $\mathbf{L}_{LG}$  is only one-dimensional, meaning that the filtered signal is driven towards a more restricted space, which is here completely unrelated to the cycle space.

The flow denoising procedure in (8) can be deemed as the solution of a specific instance of a regularized least-squares problem where there is no potential function  $\phi$  defined on the nodes, i.e., where the regularizer promotes flow conservation. In general, one might know the location of sources and sinks within the network, which can then be incorporated into the constraints in a formulation like (5), or included as a modified regularizer. This motivates the following optimization problem given a noisy flow signal  $\mathbf{f}$   $\min_{\hat{\mathbf{f}}} \{\|\hat{\mathbf{f}} - \mathbf{f}\|_2^2 + \alpha\|\mathbf{B}\hat{\mathbf{f}} - \phi\|_2^2\}$ , where the denoised signal  $\hat{\mathbf{f}}$  is sought to be close to the observed  $\mathbf{f}$  while approximately satisfying the flow balance equation. The solution to this augmented problem is  $\hat{\mathbf{f}} = (\mathbf{I} + \alpha\mathbf{L}_1)^{-1}(\mathbf{f} + \alpha\mathbf{B}^\top\phi)$ . Notice that, as expected, this reduces to (8) in the absence of sources and sinks ( $\phi = \mathbf{0}$ ).

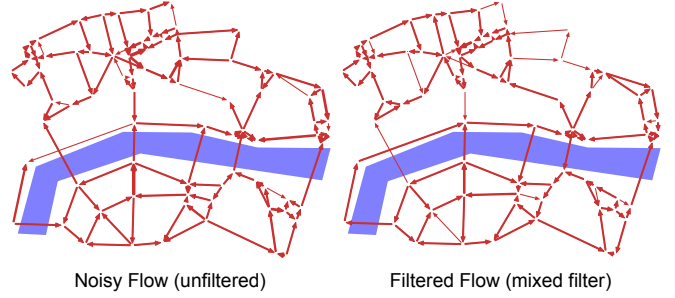
**Mixed edge filters.** We have thus far unveiled that the processing of signals on the edges of a graph is substantially different when considering the Laplacian of the line-graph  $\mathbf{L}_{LG}$  as opposed to the Edge-Laplacian  $\mathbf{L}_1$ . More precisely, the latter is preferable when considering flow signals that are subject to (approximate) conservation laws. However, promoting similar signal values among incident edges – as done by low-pass filters based on  $\mathbf{L}_{LG}$  – can be useful in different settings. For example, consider a bipartite graph representing students and homework problems, where the edges encode the perceived difficulty of each problem. Most probably some problems will be deemed as hard (or easy) by most of the students, thus presenting a smooth variation in the line graph. In general, *mixed edge filters* can combine the processing capabilities of  $\mathbf{L}_{LG}$  and  $\mathbf{L}_1$ . More precisely, given a noise signal  $\mathbf{f}$  on the edges of a graph, we define the filtered output as

$$\hat{\mathbf{f}}_{\text{mixed}} = (\mathbf{I} + \alpha\mathbf{L}_1 + \beta\mathbf{L}_{LG})^{-1}\mathbf{f}, \quad (10)$$

which is the optimal solution of the doubly-regularized problem  $\min_{\hat{\mathbf{f}}} \{\|\hat{\mathbf{f}} - \mathbf{f}\|_2^2 + \alpha\hat{\mathbf{f}}^\top\mathbf{L}_1\hat{\mathbf{f}} + \beta\hat{\mathbf{f}}^\top\mathbf{L}_{LG}\hat{\mathbf{f}}\}$ . Mixed edge filters as in (10) can be especially relevant for the *interpolation* of flow signals where large portions of the edges are unobserved, thus, flow conservation alone is not a good-enough regularizer. However, a detailed treatment of the problem of interpolation of edge-signals is left as future work.

## V. NUMERICAL EXPERIMENTS

We consider a (subset) of the street network of London – as considered in [17], [32] – that we illustrate in Fig. 2 along with the



**Fig. 2: Flow denoising on the London street network.** Left: unfiltered noisy flow on the London street network. Right: filtered flows using the mixed flow-denoising discussed method in (10) with parameters  $\alpha = 28, \beta = 0.06$ , chosen via a grid-search. Note how some flows have switched direction after filtering.

River Thames for geographical orientation. The network comprises 82 crossings (nodes) and 130 streets (edges) on which we place a generic initial flow  $\mathbf{f}_0$  distorted by white Gaussian noise (see Fig. 2-left). Since most real edge data is neither perfectly flow-preserving nor exactly constant, the initial flow  $\mathbf{f}_0$  was generated as a random mixture of a harmonic flow and a potential flow based on the line-graph Laplacian, thus containing low-pass components according to both filter classes.

On this data we apply a line-graph denoising (with parameter  $\alpha = 0.16$ ), the flow denoising procedure (8) ( $\alpha = 37$ ), and the mixed denoising filter (10) ( $\alpha = 28, \beta = 0.06$ ), where all the parameters have been chosen through a grid search for best performance. In this scenario, the initial error  $\|\mathbf{f}_0 - \mathbf{f}\|_2 \approx 8.45$  is reduced by all the filters. The line-graph filter leads to an error of  $\|\mathbf{f}_0 - \hat{\mathbf{f}}_{LG}\|_2 \approx 7.08$ . The Edge-Laplacian filter provides a better result with  $\|\mathbf{f}_0 - \hat{\mathbf{f}}_{L_1}\|_2 \approx 5.65$ . However, the best results are achieved by the mixed edge filter with an error  $\|\mathbf{f}_0 - \hat{\mathbf{f}}_{\text{mixed}}\|_2 \approx 5.24$ , depicted in Fig. 2-right. The superior performance of the mixed edge filter indicates that both the conventional line-graph filters and the proposed Edge-Laplacian filters are indeed well suited to eliminate different types of noise. A more detailed examination of real flow data, and a characterization of the settings where mixed filters achieve marked performance improvements – in terms of the noiseless signals, the type of noise present, and the underlying graph topology – are left as future work.

## VI. CONCLUSIONS

We presented a GSP framework for the treatment of signals defined on the edges of a graph. Instead of considering the line-graph and replicating the tools of node-centric GSP, we proposed a different approach based on the Edge-Laplacian. We showed that the proposed approach is indeed better suited for the treatment of signals representing flows in networks, and presented a series of filters for the denoising and smoothing of flow signals. Finally, we illustrated that the line-graph and Edge-Laplacian approaches can be combined, obtaining better denoising results in practice.

A number of exciting future directions follow naturally from the work here presented including: i) Edge-Laplacian methods for interpolation and treatment of missing data; ii) Non-linear filtering in the edge-space; and iii) Extensions to signals defined on higher-order simplicial complexes.

## VII. REFERENCES

- [1] A. Sandryhaila and J.M.F. Moura, "Discrete signal processing on graphs," *IEEE Trans. Signal Process.*, vol. 61, no. 7, pp. 1644–1656, Apr. 2013.
- [2] D. Shuman, S. 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. G. Marques, S. Segarra, G. Leus, and A. Ribeiro, "Sampling of graph signals with successive local aggregations," *IEEE Trans. Signal Process.*, vol. 64, no. 7, pp. 1832–1843, April 2016.
- [4] S. Chen, R. Varma, A. Sandryhaila, and J. Kovacevic, "Discrete signal processing on graphs: Sampling theory," *IEEE Trans. Signal Process.*, vol. 63, no. 24, pp. 6510–6523, Dec 2015.
- [5] M. Tsitsvero, S. Barbarossa, and P. Di Lorenzo, "Signals on graphs: Uncertainty principle and sampling," *IEEE Trans. Signal Process.*, vol. 64, no. 18, pp. 4845–4860, Sept 2016.
- [6] S. Segarra, A. G. Marques, G. Leus, and A. Ribeiro, "Reconstruction of graph signals through percolation from seeding nodes," *IEEE Trans. Signal Process.*, vol. 64, no. 16, pp. 4363–4378, Aug 2016.
- [7] X. Wang, M. Wang, and Y. Gu, "A distributed tracking algorithm for reconstruction of graph signals," *IEEE J. Sel. Topics Signal Process.*, vol. 9, no. 4, pp. 728–740, June 2015.
- [8] D. Romero, M. Ma, and G. B. Giannakis, "Kernel-based reconstruction of graph signals," *IEEE Trans. Signal Process.*, vol. 65, no. 3, pp. 764–778, Feb 2017.
- [9] S. Segarra, A. G. Marques, and A. Ribeiro, "Optimal graph-filter design and applications to distributed linear network operators," *IEEE Trans. Signal Process.*, vol. 65, no. 15, pp. 4117–4131, Aug 2017.
- [10] E. Isufi, A. Loukas, A. Simonetto, and G. Leus, "Autoregressive moving average graph filtering," *IEEE Trans. Signal Process.*, vol. 65, no. 2, pp. 274–288, Jan 2017.
- [11] O. Teke and P. P. Vaidyanathan, "Extending classical multirate signal processing theory to graphs – part i: Fundamentals," *IEEE Trans. Signal Process.*, vol. 65, no. 2, pp. 409–422, Jan 2017.
- [12] J. A. Deri and J. M. F. Moura, "Taxi data in New York city: A network perspective," in *Asilomar Conf. on Signals, Systems, and Computers*, Nov 2015, pp. 1829–1833.
- [13] I. Jablonski, "Graph signal processing in applications to sensor networks, smart grids, and smart cities," *IEEE Sensors Journal*, vol. 17, no. 23, pp. 7659–7666, Dec 2017.
- [14] G. Carlsson, F. Mémoli, A. Ribeiro, and S. Segarra, "Hierarchical clustering of asymmetric networks," *Advances in Data Analysis and Classification*, vol. 12, no. 1, pp. 65–105, Mar 2018.
- [15] W. Huang, T. A. W. Bolton, J. D. Medaglia, D. S. Bassett, A. Ribeiro, and D. Van De Ville, "A graph signal processing perspective on functional brain imaging," *Proc. IEEE*, vol. 106, no. 5, pp. 868–885, May 2018.
- [16] J. D. Medaglia, W. Huang, S. Segarra, C. Olm, J. Gee, M. Grossman, A. Ribeiro, C. T. McMillan, and D. S. Bassett, "Brain network efficiency is influenced by the pathologic source of corticobasal syndrome," *Neurology*, vol. 89, no. 13, pp. 1373–1381, 2017.
- [17] M. T. Schaub, J. Lehmann, S. N. Yaliraki, and M. Barahona, "Structure of complex networks: Quantifying edge-to-edge relations by failure-induced flow redistribution," *Network Science*, vol. 2, no. 1, pp. 66–89, 2014.
- [18] K. A. Bacik, M. T. Schaub, M. Beguerisse-Díaz, Y. N. Billeh, and M. Barahona, "Flow-based network analysis of the *Caenorhabditis elegans* connectome," *PLoS Computational Biology*, vol. 12, no. 8, pp. e1005055, 2016.
- [19] B. R. C. Amor, M. T. Schaub, S. N. Yaliraki, and M. Barahona, "Prediction of allosteric sites and mediating interactions through bond-to-bond propensities," *Nature Communications*, vol. 7, pp. 12477, 2016.
- [20] C. Godsil and G. Royle, *Algebraic graph theory*, vol. 207, Springer-Verlag, Graduate Texts in Mathematics, 2001.
- [21] T. S. Evans and R. Lambiotte, "Line graphs, link partitions, and overlapping communities," *Phys. Rev. E*, vol. 80, pp. 016105, Jul 2009.
- [22] Y.-Y. Ahn, J. P. Bagrow, and S. Lehmann, "Link communities reveal multiscale complexity in networks," *Nature*, vol. 466, no. 7307, pp. 761, 2010.
- [23] S. Barbarossa and M. Tsitsvero, "An introduction to hypergraph signal processing," in *IEEE Intl. Conf. Acoust., Speech and Signal Process. (ICASSP)*, March 2016, pp. 6425–6429.
- [24] S. Barbarossa, S. Sardellitti, and E. Ceci, "Learning from signals defined over simplicial complexes," in *Data Science Workshop*, March 2018, pp. 51–55.
- [25] A. R. Benson, R. Abebe, M. T. Schaub, A. Jadbabaie, and J. Kleinberg, "Simplicial closure and higher-order link prediction," *arXiv preprint arXiv:1802.06916*, 2018.
- [26] A. Hatcher, *Algebraic topology*, Cambridge University Press, 2002.
- [27] L. J. Grady and J. R. Polimeni, *Discrete calculus: Applied analysis on graphs for computational science*, Springer Science & Business Media, 2010.
- [28] Lek-Heng Lim, "Hodge Laplacians on graphs," in *Proceedings of Symposia in Applied Mathematics, Geometry and Topology in Statistical Inference*. 2015, vol. 73, AMS.
- [29] D. Zelazo and M. Mesbahi, "Edge agreement: Graph-theoretic performance bounds and passivity analysis," *IEEE Trans. Auto. Control*, vol. 56, no. 3, pp. 544–555, March 2011.
- [30] D. Horak and J. Jost, "Spectra of combinatorial laplace operators on simplicial complexes," *Advances in Mathematics*, vol. 244, pp. 303 – 336, 2013.
- [31] M. T. Schaub, A. R. Benson, P. Horn, G. Lippner, and A. Jadbabaie, "Random Walks on Simplicial Complexes and the Normalized Hodge Laplacian," *arxiv preprint arxiv:1807.05044*, July 2018.
- [32] H. Youn, M. T. Gastner, and H. Jeong, "Price of anarchy in transportation networks: efficiency and optimality control," *Physical Review Letters*, vol. 101, no. 12, pp. 128701, 2008.