ABSTRACT

Linear graph filters are a class of graph-signal operators that combine mathematical tractability with practical relevance. However, a number of meaningful problems cannot be satisfactorily addressed within the linear domain. Motivated by this, in this paper we introduce the concept of nonlinear weighted median graph filters (WMGF). The goal is to generalize the definition of classical weighted median filters to operate over graph signals not defined in regular domains such as time or a rectangular lattice. Under the proposed definition the output of a WMGF can be viewed as a nonlinear (median) combination of shifted versions of the input. Additionally, the value of each of these shifted versions at a particular node is found iteratively by computing the weighted median of the values of the previous shifted input within the one-hop neighborhood of the node. Within the class of WMGFs we pay especial attention to center-WMGF – where each node can select its own weight but must weight all its neighbors evenly –, and analyze some of their properties, including the characterization of their roots. Simulations illustrating how our definitions can be used to model and analyze nonlinear network diffusion processes close the paper.

Index Terms— Nonlinear graph signal processing, Nonlinear graph filters, Weighted median graph filters, Distributed nonlinear operators, Nonlinear diffusion.

1. INTRODUCTION

Graph Signal Processing addresses the problem of modeling, analyzing and processing information associated with the nodes (links) of a graph. The tacit assumption is that the network structure defines a notion of proximity or dependence among the nodes of the graph [1,2], which must be leveraged when designing algorithms tailored to graph signals. A plethora of graph-supported signals exist, with examples ranging from gene-expression patterns defined on top of gene networks to the spread of a rumor over a social network. Transversal to the particular application, we must address the question of how to redesign traditional tools originally conceived to study and process signals defined on regular domains and extend them to the more complex graph domain. The problem that we investigate in this paper is the definition of weighted median graph filters (WMGFs), which can be understood as a generalization of classical median filters operating with signals defined in irregular domains represented by a generic weighted graph.

Linear time-invariant filters are a centerpiece in classical signal processing (SP); hence, several works have investigated the problem of designing linear graph filters (LGFs). LGFs can be defined either in the frequency domain [1,3]; or in the original graph domain [2,4], using the concept of graph-shift operator (GSO). Except for pathological cases, those definitions are equivalent and have been successfully applied to the problems such us denoising, signal reconstruction, or classification [5,8], to name a few. LGFs have also been used to model linear diffusion and percolation processes with the GSO accounting for the local structure of the network where the process takes place [8,10]. However, many relevant problems in classical SP cannot be satisfactorily addressed within the linear domain. A subclass of nonlinear operators that has been shown to be effective in dealing with a number of practical problems while retaining some mathematical tractability are median filters, as well as their (negative) weighted generalizations [11,12]. Weighted median filters yield robust estimates, can deal effectively with impulsive noise and outliers, and do not suffer from edge effects [11]. Motivated by these attractive features, our goal is to generalize the definition of a classical weighted median filter to operate over graph signals guaranteeing that the processing steps that transform the input into the output account for the (local) structure of the graph where the signals are defined.

With that purpose in mind, we start by reviewing the classical definition of weighted median filters (Sec. 1.1) and LGFs (Sec. 1.2), emphasizing their interpretation as local linear operators. We then introduce the definition of WMGF (Sec. 2) and discuss its relevance to model nonlinear local network dynamics. Sec. 3 particularizes this definition for center-WMGF, which are a more tractable class of WMGF where each node is allowed to select its own weighting coefficient, but must use the same weight for all its one-hop neighbors. We first provide two definitions for center-WMGF and then analyze some of their properties, including the so-called saturation coefficient, and their filter roots, which are graph signals that remain invariant after the application of the filter (Sec. 3.2). Numerical experiments illustrating the behavior and relevance of the proposed definitions (Sec. 3) together with concluding remarks (Sec. 5) wrap-up the manuscript.

1.1. Weighted median filters

Consider in this section that \( x \) represents a time signal of finite length. In its simplest definition, a median filter of memory (length) \( L \) is a nonlinear operator that given an input \( x \) generates an output \( y \) with the same length whose \( t \)-th element is

\[
y_t = \text{Med}(x_{t:L}), \quad \text{with} \quad x_{t:L} := \{x_t, x_{t-1}, \ldots, x_{t-L'+1}\} \quad \text{and} \quad L' := \min(t, L).
\]

The median operator \( \text{Med}(\cdot) \) sorts the input samples in ascending order and picks as output the middle one. We adopt the conven-
tion that, when \( L \) is even, among the two feasible values, the median is defined as the maximum one. Since the median filter in (1) weights all the \( L \) samples evenly, a natural generalization is to consider non-homogeneous scenarios where some samples (e.g., the most recent ones) are more important than the others. One way to accomplished this is by considering a set of non-negative integer weights \( h = [h_0, ..., h_{L-1}]^T \), so that higher values imply that the sample is more relevant. To this end, given a non-negative integer \( c \) and a scalar input \( x \) we introduce the replication operator \( \circ \) so that \( c \circ x \) is a sequence (vector) of length \( c \) with all the entries equal to \( x \). Moreover, using \( \delta \) to denote the weighted median operator, the output of a weighted median filter with non-negative integer coefficients \( h \) is found as follows.

**Definition 1** Given the coefficients \( h \in \mathbb{N}_L^d \) and the input \( x \in \mathbb{R}^N \), the output \( y \in \mathbb{R}^N \) of a weighted median filter is generated as

\[
y_t := h \tilde{x}_{t|L}(h), \quad \text{with} \quad \tilde{x}_{t|L}(h) := [h_0 \circ x_1, h_1 \circ x_{t-1}, ..., h_{L-1} \circ x_{t-L+1}].
\]

Note that sequence \( \tilde{x}_{t|L}(h) \) has length \( M_{t|L} = \sum_{l=0}^{L-1} h_l \) and recall that \( L' = \min(t, L) \). Alternatively, one can also view (2) as the solution to \( y_t = \arg\min_{x_t} \sum_{l=0}^{L'} h_l |x_t - x| \), where we recall that when the minimum is not unique, we adopted the convention that the output is the largest feasible \( x \) [12]. This optimization problem is useful to generalize further the definition of a weighted median filter by considering negative and non-integer weights as follows:

\[
y_t := h \tilde{x}_{t|L}(cL, h), \quad \text{where} \quad \tilde{x}_{t|L}(cL, h) := [h_0 \circ x_1, h_1 \circ x_{t-1}, ..., h_{L-1} \circ x_{t-L+1}].
\]

### 1.2. Graph signals and LGFs

Let \( G \) denote a directed graph with a set of \( N \) nodes or vertices \( \mathcal{N} \) and a set of links \( \mathcal{E} \), such that if node \( i \) is connected to \( j \), then \((i, j) \in \mathcal{E}\). The (incoming) neighborhood of \( i \) is defined as the set of nodes \( \mathcal{N}_i = \{j | (j, i) \in \mathcal{E}\} \) connected to \( i \). The adjacency and degree matrices of \( G \) are denoted as \( A \) and \( D \), with \( D = \text{diag}(A1) \).

**Graph signals and shift operator.** The focus of this paper is not on analyzing \( G \), but graph signals defined on the set of nodes \( \mathcal{N} \). Formally, each of these signals can be represented as a vector \( x = [x_1, ..., x_N]^T \in \mathbb{R}^N \) where the \( i \)-th element represents the value of the signal at node \( i \). Since the vectorial representation of a graph signal does not convey explicitly the structure in \( G \), the graph is endowed with a GSO \( S \) [2][13]. The operator \( S \) is an \( N \times N \) matrix whose entry \( S_{ij} \) can be non-zero only if \( i = j \) or if \((i, j) \in \mathcal{E}\). The sparsity pattern of the matrix \( S \) captures the local structure of \( G \), but we make no specific assumptions on the values of the non-zero entries of \( S \). Choices for \( S \) are \( A \) [2][13], the graph Laplacian \( L = D - A \), and their respective generalizations [14]. The intuition behind \( S \) is to represent a transformation that can be computed locally at the nodes of the graph. To be rigorous, define \( s_i \in \mathbb{R}^N \) as the \( i \)-th row of the GSO \( S_i := [S_{i1}, ..., S_{iN}]^T \) and consider the signal \( y \) obtained as \( y = Sx \), so that the value of \( y \) at node \( i \) is \( y_i = S_{ii}x_i \). Since \( s_i \) is non-zero only for the nodes within the one-hop neighborhood of \( i \), node \( i \) can compute \( y_i \) provided that it has access to the value of \( x_j \) at \( j \in \mathcal{N}_i \).

**LGFs.** A graph-signal operator is a transformation \( T : \mathbb{R}^N \rightarrow \mathbb{R}^N \) between graph signals. LGFs are a particular class of linear graph-signal operators that can be represented as matrix polynomials of the shift operator \( S \), i.e., \( N \times N \) matrices of the form \( H = \sum_{l=0}^{L-1} h_l S^l \), where \( h = [h_0, ..., h_{L-1}]^T \) are the filter coefficients. The formal definition stating the input-output relation of a LGF is given next.

**Definition 2** Given the coefficients \( h \in \mathbb{R}^L \) and the input graph signal \( x \in \mathbb{R}^N \), the output \( y \in \mathbb{R}^N \) of a LGF is generated as

\[
y_t := h^T z_t, \quad \text{where} \quad z_t := \begin{bmatrix} z_t(0) & \ldots & z_t(L-1) \end{bmatrix}^T,
\]

\[
z_t(l) := s_t^L z(l-1) \quad \text{and} \quad z_t(0) := x_t.
\]

Note that the previous is equivalent to saying that \( y := \sum_{l=0}^{L-1} h_l z(l) \) with \( z(t) := S z(l-1) \) and \( z(0) = x \).

Hence, as is the case for classical time signals, the output of a linear graph filter can be viewed as a linear combination of shifted versions of the input. Definition 2 tries to emphasize the fact that LGFs can be implemented locally, with \( L-1 \) exchanges of information among neighbors [3]. This is true, because: i) node \( i \) can obtain \( z_t(0) \) locally based on the values of \( z_j(l-1) \) at \( j \in \mathcal{N}_i \), as \( z_t(l) = \sum_{j \in \mathcal{N}_i} S_{ij} z_j(l-1) \); and ii) signals \( z(t) := [z(0), z(1), \ldots, z(L-1)]^T \) can be computed sequentially in \( L \) steps using (4).

### 2. WEIGHTED MEDIAN GRAPH FILTERS

For classical time signals, linear time-invariant filters find the value of the output at a given time instant as a linear combination of values of the input. Hence, the idea behind the median filters in Sec. 1 is to replace such a linear combination with a weighted median. When signals are supported on irregular graph domains, Definition 2 revealed that LGFs use two different linear combinations. The first one is a local linear combination dictated by the rows of \( S \) [cf. (3)]. The second one, which is the counterpart of the one for time-invariant filters, dictates how to find the output by combining linearly the different shifted inputs according to the coefficients \( h \) [cf. (3)].

Our proposal in this paper is to generalize the definition of a weighted median filter to the graph domain by replacing the linear combinations in (3) and (4) with weighted medians. This gives rise to the following definition of WMGF.

**Definition 3** Given the coefficients \( h \in \mathbb{R}^L \) and the input graph signal \( x \in \mathbb{R}^N \), the output \( y \in \mathbb{R}^N \) of a WMGF is generated as

\[
y_t := h^T z_t, \quad \text{where} \quad z_t := \begin{bmatrix} z_t(0) & \ldots & z_t(L-1) \end{bmatrix}^T,
\]

\[
z_t(l) := s_t^L z(l-1) \quad \text{and} \quad z_t(0) := x_t.
\]

To emphasize the dependence on \( h \) and \( S \), we will denote the WMGF as \( H(h, S) \). Moreover, we will use \( y = H(h, S) \hat{x} \) to denote that \( y \) is the output obtained after applying the filter to the input \( x \). When the above definition is particularized for the adjacency matrix associated with the directed cycle, which is the support for classical time signals [1]. Definition 3 reduces to the classical weighted median filter in Definition 1. To understand the operation of a WMGF, consider first the local nonlinear dynamics in (6), which establishes that the shifted inputs \( z(l) \) are found using a nonlinear weighted median, with coefficients given by the entries of \( S \). Since \( s_i \) is only non-zero within the one-hop neighborhood, the values of \( z_t(l) \) in (6) can be obtained locally by \( i \). For example suppose that \( i = 2, N_2 = \{1, 2, 3\} \) and \( s_2 = [2, 3, 2, 0, ..., 0]^T \). Then, \( z_2(l) := s_2^L z(l) \) can be found locally as \( z_2(l) := \text{Med}([x_1, x_2, x_2, x_2, x_3, x_3]) \). Similarly, when all the shifted signals have been generated, the value of the output signal \( y \) at every node is found by using the weighted median operation in (5). For example, one can choose the output \( y \) to be \( z_{(L-1)} \) i.e., the result of applying the nonlinear operator \( S \) to the input \( L-1 \) consecutive times– by setting \( h = [0, 0, 0, 1]^T \).
3. CENTER-WEIGHTED MEDIAN GRAPH FILTERS

Center-weighted median filters are a particular class of traditional weighted median filters that set all the weights to one except the one corresponding to value of the input at the current instant (center of the window) \[15\]. In this section we propose two different generalizations for filters operating in graph domains.

**Definition 4** Given an unweighted and undirected graph \( G \) with no self-loops, adjacency matrix \( A \), and degree matrix \( D \):

i) A uniform center-WMGF (uCMGF) is a WMGF where \( S \) is of the form

\[
S = C_u(\rho) := \rho I + A, \quad \text{with} \quad \rho > 0. \tag{7}
\]

ii) A degree-based center-WMGF (dCMGF) is a WMGF where \( S \) is of the form

\[
S = C_d(\alpha) := \alpha D + A, \quad \text{with} \quad \alpha > 0. \tag{8}
\]

Both (7) and (8) are natural extensions of classical center-weighted median filters. To see this, consider the classical definition in time with (center) weight \( W \) and window size \( 2K + 1 \). Then, we build a graph \( G \) whose nodes represent discrete time instants and where each node \( i \) is connected to all the nodes (time instants) included in the window centered at \( i \). In this case, the application of the classical filter is equivalent to the application of either \( C_u(\rho) \) with \( \rho = W \), or \( C_d(\alpha) \) with \( \alpha = W/2K \). Notice that in regular graphs \( D \) is a multiple of \( I \), so that (7) and (6) describe the same class of shifts.

Since the graph \( G \) constructed from the discrete time points is regular, the distinction between uniform and degree-based filters is irrelevant in classical SP. However, for general graphs, they lead to different behaviors. For simplicity, Secs. 3.1 and 3.2 focus on analyzing uCMGFs, but similar results can be stated for dCMGFs.

3.1. Saturation property

A central feature of classical center-weighted median filters is that they attenuate extreme values of the input signal while leaving intermediate values unperturbed \[15\]. This saturation property is preserved by their generalizations to graphs.

Given the graph signal \( x \in \mathbb{R}^N \), we denote by \( x^1 \) the signal of length \( |N| + 1 \) collecting the values of \( x \) at node \( i \) and its neighborhood \( N_i \). Furthermore, given any signal \( x \in \mathbb{R}^N \) and integer \( k \), we denote by \( x^{(k)}(i) \) the \( k \)-th element of \( x \) when sorted in increasing order.

\[ x^{(k)}(i) := x(i) \quad \text{for} \quad k < 1 \quad \text{and} \quad x^{(k)}(i) := x(i + N) \quad \text{for} \quad k > N. \]

For consistency, we set \( x^{(k)} := x(1) \) for \( k < 1 \) and \( x^{(k)} := x(N) \) for \( k > N \). The following saturation property is satisfied by uCMGFs.

**Proposition 1** The output \( y \) of applying \( C_u(\rho) \) to the input \( x \) is

\[
y_i = \text{Med}([x^{(1)}(i), x_i, x^{(1)}(u_i)]) \tag{9}
\]

where the lower and upper saturation limits are given by

\[
l_\rho = \lfloor |N_i| - \rho/2 \rfloor + 2, \quad u_\rho = \lceil |N_i| + \rho/2 \rceil + 1. \tag{10}
\]

**Proof:** See appendix A. \[\blacksquare\]

From (9) it follows that the shifted output at node \( i \) is equal to the input as long as \( x_i \) is contained in the interval \([x^{(1)}(i), x^{(1)}(u_i)]\). Otherwise, the shifted output takes the value of the corresponding limit that is violated. This characteristic makes them specially suitable for removing outliers and can be used for the denoising of signals with impulsive noise, in the same way that classical center-weighted median filters can be used to treat ‘salt and pepper’ noise in images \[16\].

Lastly, it is not hard to check that for the graph \( G \) representing a center-weighted median filter for time signals, the classical saturation property \[15\] (Chapter 5.2) is recovered by Prop. 1 for odd and integer \( \rho \) and even \( |N_i| = 2K \).

3.2. Root analysis

Since complex exponentials are not eigenfunctions of classical median filters, frequency analysis is not a suitable tool to study WMGFs. When analyzing median filters the basic descriptors of their deterministic properties are their roots, which are signals invariant to filtering \[11\].

Define a graph signal structure \( \mathcal{X} \) on a given graph \( G \) as the set of all graph signals that share a common constant piecewise definition, regardless of the actual values taken by the signal. For example, all constant signals \( x \) are contained in the structure \( \mathcal{X}_0 := \{ x \mid \exists \alpha \in \mathbb{R} \text{ s.t. } x = \alpha \cdot a, \ldots, a \} \) regardless of their value. Notice that each node clustering \( C \) (i.e., each partition of the elements in \( N \)) gives rise to a different signal structure \( \mathcal{X}_C \), with the pieces where the signals are constant given by the clusters in \( C \). With these definitions, we say that \( \mathcal{X}_C \) is a weight-independent root structure of a median graph filter \( H(h, S) \) if

\[
H(h, S)x = x, \tag{11}
\]

for all \( h \) and all graph signals \( x \in \mathcal{X}_C \). Furthermore, given an undirected graph \( G \) and a clustering \( C \), define \( e_i^{\text{out}} \) and \( e_i^{\text{in}} \) as the number of edges connecting node \( i \) to nodes outside and inside its cluster, respectively. We then say that the clustering \( C \) is \( \rho \)-stable if

\[
e_i^{\text{out}} - e_i^{\text{in}} < \rho, \quad \text{for all} \ i. \tag{12}
\]

This allows us to characterize next the root structures of uCMGFs.

**Proposition 2** The signal structure \( \mathcal{X}_C \) is a weight-independent root structure of \( H(h, C_u(\rho)) \) if and only if \( C \) is \( \rho \)-stable.

**Proof (sketch):** First show that \( \mathcal{X} \) is a weight-independent root structure of \( H(h, C_u(\rho)) \) if and only if it is a root structure of \( C_u(\rho) \). One direction follows from the definition of root while the other follows from choosing \( h = [0, 1] \). Then show the statement of the proposition for the specific filter \( C_u(\rho) \). Given a \( \rho \)-stable clustering \( C \), that every signal \( x \in \mathcal{X}_C \) is a root follows from

\[
[C_u(\rho)x]_i := \text{Med}([x_i, \ldots, x_i, \rho \circ x_i, x_{i+1}, \ldots, x_{i+N}]) = x_i, \tag{13}
\]
for $K = e^\text{out}_i$ and where the unweighted $x_i$ is repeated $e^\text{in}_i$ times. The last equality in (13) follows from (12). For the opposite direction, we may write (13) but with $e^\text{out}_i$ and $e^\text{in}_i$ not satisfying (12) and build a counterexample by selecting, e.g., $x_i > x_{jk}$ for all $k$.

The above result states that structures representing signal segregation in the graph are impervious to local median dynamics. Moreover, the only signal structures unperturbed by these interactions are that of constant-signal neighborhoods. Notice that, Prop. \[5\] notwithstanding, other root signals exist for specific choices of signal values, however, no other signal structure $X$ exists such that all signals in that structure are roots. Finally, from classical nonlinear SP [17], it is known that the only root structures of a median graph filter of window width $2K + 1$ are constant neighborhoods of size at least $K$. This result can be obtained as a particular case of Prop. \[5\] for $\rho = 1$ where the minimum size of $K$ is necessary and sufficient to guarantee 1-stability of the constant neighborhoods.

4. NUMERICAL EXPERIMENTS

The first experiment illustrates the results in Prop. 1. We consider a connected symmetric Erdös-Rényi (ER) random graph with $N = 20$ nodes and edge probability $p = 0.3$ [18]. A smooth input $x$ (generated as a bandlimited signal in the eigenbasis of the Laplacian of the graph $G$) is filtered using a uCMGF with weight $\rho = 5$, to generate the output $y$. Both $x$ and $y$ as well as the thresholds in (10) are shown in Fig. 1b. The plots confirm that for most of the nodes the value of the input and the output are the same. Only for nodes where the value of the input signal is very large (node 10) or very small (nodes 12 and 15), the output is a clipped version of the input.

The second experiment illustrates how WMGFS can act as barriers to prevent the spreading of anomalies. A graph with $N = 12$ nodes generated using a stochastic block model (SBM) with 2 communities of equal size, intra-community link probability of $p = 0.90$, and inter-community link probability of $q = 0.10$ is considered [18]. The values of the input signal $x$ are drawn independently across nodes from a real uniform distribution between 0 and 1. Then value of $x$ at three nodes, two of them belonging to the cluster formed by nodes 1-6 and the third one from the cluster formed by nodes 7-12, is corrupted with an anomaly, as shown in the leftmost column of Fig. 1c. This corrupted signal is then filtered using three uCMGFs with different values of $\rho$ and each of them of degree two, so that the GSO is applied twice. The output generated by each

These filters, as well as the intermediate signal obtained after the first application of the shift, are shown in Fig. 1b. The results confirm that lower values of $\rho$ and higher number of degrees are more effective filtering the anomalies. However, if too aggressive, those configurations can also cause information loss, eliminating the details of the non-corrupted information. For intermediate values of $\rho$ and filter degree ($\rho = 2$ and $L = 1$), we also observe that when anomalies affect several nodes within the same community, they are more likely to survive.

The last experiment illustrates the results in Prop. 2. Two types of graphs are considered. The first one is the ER graph in our first experiment; the second one is a SBM-graph with $N = 20$ nodes, 4 communities of size 5, and intra and inter community link probability of $p = 0.90$ and $q = 0.05$, respectively. For each graph, we first generate a random input, which is then iteratively filtered, so that the output generated at iteration (time) $t$ serves as input for iteration $t + 1$. The original input, as well as the filtered outputs, are shown in Fig. 1c both for the SBM graph (top) and for the ER graph (bottom). The filters implemented correspond to a uCMGF with $\rho = 1$. The figures show that in both cases, the output of the last stage converges to a piecewise constant signal that is a root structure of the graph [cf. (11)]. For the SBM, after 3 stages the output converges to a signal that is constant within each of the four communities (clusters) in the graph. For the ER graph, 6 stages are required to convergence, and the root signal is constant across the entire graph. Note that since the ER model does not impose a clustered structure and the value of $p$ is very small, it is not likely to have a root structure other than the constant signal satisfying the stability condition in (12). While the results shown in Fig. 1c correspond to a single realization, the experiment was repeated for 100 signals. For the SBM graph the output converged to a signal constant within the communities more than 90% of the times, while for the ER graph the output converged to a constant signal around 50% of the realizations.

5. CONCLUSIONS

This paper introduced the definitions of WMGF and center-WMGF, which are two classes of nonlinear graph-signal operators. Differences relative to classical median filters and linear graph filters were discussed, and several properties of center-WMGF were analyzed in detail. Since the output of a WMGF is obtained based on local nonlinear operators that depend on the structure of the graph, the proposed definitions were shown to be relevant in applications involving nonlinear network dynamics.
6. REFERENCES


