# An uncertainty principle for lowband graph signals

Antoine Mazarguil, Laurent Oudre, and Nicolas Vayatis

***Abstract*—In this article, we introduce a novel lower bound on the support size of lowband graph signals. This result allows the deduction of an optimality criterion for the lowband and sparse decomposition of any graph signal, establishing the uniqueness of well behaving solutions. A comparison of the new bound with previously introduced results is performed, showing the improvements brought by the present work. An illustration on a practical denoising usecase on a real graph is also provided.**

***Index Terms*—graph signal processing, uncertainty principle, lowband graph signals**

## I. Introduction

In the literature, uncertainty principles encompass a wide variety of inequalities that jointly limit the occupancy of a signal in different representation spaces. The most famous uncertainty principle is the widely known Heisenberg inequality, which states that the "spread" of a signal and its Fourier transform cannot be arbitrarily small simultaneously (see [1], [2] for review and applications).
With the development of digital signal processing, a tremendous line of search has been to derive similar inequalities for discrete signals. Numerous approaches have been investigated, each requiring a new definition of signal uncertainty. However, a first challenge that appeared is that with the most natural notion of spread (i.e. standard deviation in the studied space), there exist signals that are perfectly localized in a space (i.e. the spread is null) while their representation in an other space is well defined (i.e. $< \infty$). Therefore, a discrete equivalent to the Heisenberg's inequality cannot be derived. The search for uncertainty principles then requires defining a more suitable notion of signal spread or localization, and each definition gives rise to a specific uncertainty principle. For an extensive review on discrete uncertainty principles and an overview of the main approaches and results, see [1]–[5].

In the context of discrete signals, an important uncertainty principle involving sparsity has been introduced in [6], [7]. In this framework, the 0-norm is presented as a relevant notion of signal spread, as it quantifies the size of the signal support. The associated uncertainty principle is stated as follows. Provided two vector bases $\Phi = (\phi^1, ..., \phi^N)$ and $\Psi = (\psi^1, ..., \psi^N)$ and a signal $\mathbf{x} \in \mathbb{R}^N$, we denote by $\mathbf{x}^\Phi$ and $\mathbf{x}^\Psi$ the respective signal expression in each basis. Then, the following inequality holds:

$$||\mathbf{x}^\Phi||_0 \, ||\mathbf{x}^\Psi||_0 \geq \frac{1}{\mu(\Phi, \Psi)^2} \tag{1}$$

with $\mu(\Phi, \Psi)$ the mutual coherence of the two bases, defined as follows :

$$\mu(\Phi, \Psi) = \max_{1 \leq i,j \leq N} |\langle \phi^i, \psi^j \rangle|. \tag{2}$$

All authors are with Université Paris-Saclay, ENS Paris-Saclay, CNRS, Centre Borelli, F-91190, Gif-sur-Yvette, France

Manuscript received April 19, 2005; revised August 26, 2015.

For pair of bases such that $\mu(\Phi, \Psi)$ is small enough, inequality (1) prevents the representations of the signal in the two bases from being both sparse simultaneously. This theorem leads to theoretical results which have concrete applications in compressed sensing for instance [8].

Since the emergence of the Graph Signal Processing (GSP) field, the search for uncertainty principles adapted to this framework and to the Graph Fourier Transform has been an active line of research [9], [10]. As for discrete signals, different notions of signal localization or spread lead to different uncertainty principles. The sparsity uncertainty principle (1) has been adapted to the graph framework in [11], deriving a first sparse graph uncertainty principle. We now present two other major approaches.

The first graph spectral uncertainty principle has been introduced and studied in [12]. Based on a generalization of the notion of spread, they introduce $\Delta_g(\mathbf{x})$ the graph spread of a signal $\mathbf{x}$ and $\Delta_s(\mathbf{x})$ the graph spectral spread of $\mathbf{x}$. Their work derives an additive uncertainty principle involving the two spreads, and present a extensive analysis of the region of feasible pairs $(\Delta_g(\mathbf{x}), \Delta_s(\mathbf{x}))$.

Another major approach to deriving a graph uncertainty principle is to use the theory presented in [4], [13] for the graph framework. This study, conducted in [14], focuses on the following quantities for a signal $\mathbf{x}$:

$$\alpha^2 = \frac{||\mathbf{D}_\mathcal{S}\mathbf{x}||_2^2}{||\mathbf{x}||_2^2} \qquad \beta^2 = \frac{||\mathbf{B}_\mathcal{F}\mathbf{x}||_2^2}{||\mathbf{x}||_2^2} \tag{3}$$

where $\alpha^2$ and $\beta^2$ represent the amount of energy confined in the set of vertices $\mathcal{S}$ and the set of frequencies $\mathcal{F}$ (with $\mathbf{D}_\mathcal{S}$ and $\mathbf{B}_\mathcal{F}$ the corresponding projection matrices). An uncertainty principle involving these quantities is presented along with theoretical results on graph signal sampling. Other approaches for the elaboration of an uncertainty principle for graph signal or related work can be found in [15]–[17].

The main contribution of this paper is a novel uncertainty principle for lowband graph signals, which provides theoretical guarantees for instance on the graph signal decomposition problem. From this principle, we derive an optimality criterion for the problem of sparse-lowband decomposition of a graph signal, which has practical consequences in the graph signal denoising task.

## II. Uncertainty principle for lowband graph signals

The present work focuses on inequalities of the form (1) for a graph signal and its Graph Fourier Transform (GFT). The aim of this section is to introduce our novel bound for lowband graph signal.

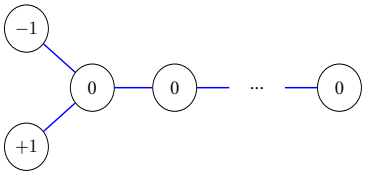


Fig. 1. A graph signal achieving sparsity in both nodal and spectral domains.

### A. Notation and Background

We consider a symmetric, connected and undirected graph $\mathcal{G} = (V, E)$ consisting of a set of $N$ nodes $V = \{1, 2, ..., N\}$, along with a set of edges $E = \{(i, j)\} \in V \times V$. The graph is supposed to be connected, i.e. there exist a path in $\mathcal{G}$ between each pair of nodes. A signal $\mathbf{x}$ over a graph $\mathcal{G}$ is defined as a mapping from the vertex set to the set of complex numbers, i.e. $x : V \longrightarrow \mathbb{C}$. We denote by $||\cdot||_p$ the p-norm of a signal, for p $\in \{0, 1, 2\}$. The combinatorial Laplacian matrix of the graph is defined as $\mathbf{L} = \mathbf{D} - \mathbf{A}$, with $\mathbf{D}$ the graph degree matrix and $\mathbf{A}$ the adjacency matrix. The Laplacian matrix is the characteristic operator that quantifies the smoothness of a graph signal, i.e. how strongly the signal values vary along the graph edges. The smoothness $S_{\mathcal{G}}(\mathbf{x})$ of a graph signal $\mathbf{x}$ is defined as $S_{\mathcal{G}}(\mathbf{x}) = \sum_{(i,j)\in E}(x_i - x_j)^2 = \mathbf{x}^T\mathbf{L}\mathbf{x}$ In order to control their smoothness, graph signals can be studied in the eigen-basis of the Laplacian operator. We introduce the eigen-decomposition of the Laplacian matrix $\mathbf{L} = \mathbf{U}\Lambda\mathbf{U}^T$ where $\Lambda$ is a diagonal matrix with non-negative real eigenvalues $0 = \lambda_1 < \lambda_2 \leq ... \leq \lambda_N$ and $\mathbf{U}$ is an orthogonal matrix. The Graph Fourier Transform (GFT) $\hat{\mathbf{x}}$ of a graph signal $\mathbf{x}$ is defined as $\hat{\mathbf{x}} = \mathbf{U}^T\mathbf{x}$. A signal $x$ is said to be $k$-lowband if only the $k$ first frequencies are active in its Fourier transform, i.e. $\hat{\mathbf{x}} = (\hat{x}_1, ..., \hat{x}_k, 0, ..., 0)^T$. The space of $k$-lowband signals is denoted by $\text{PW}_{\mathcal{G}}(\lambda_k)$, and by extension we denote by $\text{PW}_{\mathcal{G}}(\lambda)$ the space of signals whose active frequencies are restricted to the band $[0, \lambda]$. The symbol $|\mathcal{S}|$ denotes the cardinality of set $\mathcal{S}$, i.e. the number of elements of $\mathcal{S}$, and $\mathbf{x}_{\mathcal{S}}$ denotes the values of the vector $\mathbf{x}$ restricted to the indexes in $\mathcal{S}$.

### B. Demotivating example : sparsity in both node and frequency spaces.

Let us consider a graph $\mathcal{G}$ defined as a path of size $N$, to which two nodes have been added and connected to the first node of the path. Figure 1 displays the constructed graph. It is straightforward that the vector $\mathbf{u} = (1, -1, 0, ..., 0)^T$ is an eigenvector of the combinatorial laplacian of $\mathcal{G}$, associated to the eigenvalue 1. Disregarding other possible multiplicities of the eigenvalue 1, the signal $\mathbf{u}$ satisfies $||\mathbf{u}||_0 = 2$ and $||\hat{\mathbf{u}}||_0 = 1$, therefore achieving sparsity in both domain simultaneously. This example demonstrates that, without any hypothesis on the graph $\mathcal{G}$ or the signal $\mathbf{u}$, a relevant inequality similar to (1) cannot be derived (see [11]). In fact, the appearance of specific sub-structures of size $n << N$ in the graph $\mathcal{G}$ can lead to the existence of a sparse eigenvector in the decomposition of the Laplacian, inducing a $n + 1$ sparse representation over the two domains. It is also relevant to note that, apart from these sub-structures, the rest of the graph can be of any size or organization, implying no existence of upper-bounds depending on the graph size $N$.

### C. Uncertainty principle

As stated previously, there exist sparse signals in both node and GFT space for specific graph structures. However, the node-sparsity of these signals seems to induce a relatively high variation on the graph. An intuitive line of search is then to study the link between signal variation and node sparsity. We state the first result as follows.

***Theorem** 1 (Lowband graph uncertainty principle):* Given a graph $\mathcal{G}$ of size $N > 2$ and a non-null signal $\mathbf{x} \in \text{PW}_{\mathcal{G}}(\lambda)$ with $\lambda < 2$, the following inequality holds:

$$||\mathbf{x}||_0 \geq \frac{1}{2}\left(\frac{\pi}{\cos^{-1}(1 - \frac{\lambda}{2})} - 1\right) \tag{4}$$

*Sketch of proof:* The theorem is based on two major lemmas. The first lemma lower-bounds the smoothness of a signal by the squared gaps between its ordered values. If $\mathbf{x}$ is a signal on $\mathcal{G}$ and $\mathbf{y}$ is a vector containing the ordered values of $\mathbf{x}$, the following inequality holds:

$$S_{\mathcal{G}}(\mathbf{x}) = \sum_{(i,j)\in E}(x_i - x_j)^2 \geq \sum_{i=1}^{N-1}(y_i - y_{i+1})^2 \tag{5}$$

This lemma is proven by an algorithmic procedure that transforms the first sum into the second one, while the process decreases the overall quantity at each step. Sparsity appears in the second sum as the gaps between the multiple occurrences of the value 0 do not contribute. The second lemma establishes the minimal value of the second sum for a signal of unit norm assuming that it contains a null value, yielding a second inequality that involves $||\mathbf{x}||_0$. The composition of the two lemmas along with the property $S_{\mathcal{G}}(\mathbf{x}) \geq \lambda||\mathbf{x}||^2$ for $\mathbf{x} \in \text{PW}_{\mathcal{G}}(\lambda)$ yields the theorem. For a more detailed version of the proof, see Appendix A.

In the rest of the article, we will denote by $C(\lambda)$ this lower-bound.

The inequality presented in Theorem 1 lower-bounds the size of the support of lowband signal, which are extensively used in the GSP literature [9], [14]. It is relevant to note that, even though the bound is well defined for any $\lambda \in ]0, 4[$, this inequality does not provide interesting result for $\lambda > (3 - \sqrt{5})/2 \approx 0.38$. In this case, the bound $C(\lambda)$ is lower than 2 and the theorem yields $||\mathbf{x}||_0 > 1$, which is always respected by any signal $\mathbf{x} \neq 0$. The bound $C(\lambda)$ is a decreasing function of the maximum active frequency, and we emphasize that the theorem is as useful as $\lambda$ is low.

We inform the reader that the as the bound of Theorem 1 is a function of $\lambda$, its efficiency is intrinsically linked to the graph spectrum distribution. Relations between the graph eigenvalue distribution and topological properties (such as appearance of clusters, degree distribution or appearance of hubs in the graph) are a complex question that are the subject of a wide corpus [18]–[20]. Therefore, for a study of the impact of such topological properties on bound of Theorem 1, we redirect the reader to relevant literature. In the present article, this analysis is limited to experiments on 4 different graphs.

## III. Optimality criterion for sparse-lowband signal decomposition

In this section, we present a consequence of the uncertainty principle of Theorem 1. First, we introduce the notion of **Sparse-Lowband Decomposition (SLD)** of a graph signal. Then we derive an optimality condition for this decomposition problem.

### A. Sparse-lowband decomposition of a signal

In several usecases of signals gathered on a sensor network, transmission error or sensor malfunction can lead to corruption in the collected data, i.e. a subset of the values can be altered. The task of removing this impulsive noise from the corrupted signal is then of utmost importance, as well as identifying the corrupted nodes. Assuming that the studied process is lowband, this recovery takes the form of a **Sparse-Lowband Decomposition (SLD)** of the measurements.

***Definition** 1 (Sparse-Lowband Decomposition (SLD)):* Provided a graph signal $\mathbf{x}$, a $\lambda$-**SLD** of $\mathbf{x}$ is a pair $(\mathbf{t}, \mathbf{s}) \in (\mathbb{R}^N)^2$ such that:

$$\begin{cases} \mathbf{x} = \mathbf{t} + \mathbf{s} \\ \mathbf{s} \in \text{PW}_{\mathcal{G}}(\lambda) \\ \mathbf{t} \text{ is sparse} \end{cases} \tag{6}$$

A $\lambda$-SLD for a given signal $\mathbf{x}$ is optimal if there is no other $\lambda$-SLD that achieves a better sparsity on the $\mathbf{t}$ component. An optimal decomposition is therefore a solution of the following optimization problem:

$$\begin{cases} \mathbf{t}^* \in \underset{\mathbf{t}}{\text{argmin}} \ ||\mathbf{t}||_0 \quad \text{s.c.} \quad \mathbf{x} - \mathbf{t} \in \text{PW}_{\mathcal{G}}(\lambda) \\ \mathbf{s}^* = \mathbf{x} - \mathbf{t}^* \end{cases} \tag{7}$$

By construction, the problem (7) decomposes a signal $\mathbf{x}$ such that one component is as sparse as possible ($\mathbf{t}^*$) while the other component is lowband ($\mathbf{s}^* \in \text{PW}_{\mathcal{G}}(\lambda)$). However, the non-convexity of the $L_0$ term in (7) makes the problem difficult to solve. Classical optimization methods would either include combinatorial exploration or could lead to local minimums. As in many $L_0$ optimization problems, a practical way to search for a minimum is to introduce the $L_1$ counterpart of the problem (7). The $L_1$ equivalent of the previous problem is convex and can be handled by linear programming. Yet, in many practical use-case, the optimal solution happens to be sparse, thus revealing a candidate solution for problem (7) [8].

### B. Uniqueness of the Sparse-Lowband decomposition

In this subsection, we present a sufficient condition for an $\lambda$-**SLD** to be optimal, which is a consequence of Theorem 1.

***Theorem** 2 (SLD unicity criterion):* Let us consider a graph $\mathcal{G}$, a non-null signal $\mathbf{x}$ and a cutoff frequency $\lambda < 2$. Let us also consider $(\mathbf{t}, \mathbf{s})$ a $\lambda$-**SLD** of $\mathbf{x}$ such that:

$$||\mathbf{t}||_0 < \frac{C(\lambda)}{2} \tag{8}$$

Then, the pair $(\mathbf{t}, \mathbf{s})$ is the unique optimal $\lambda$-**SLD** of $\mathbf{x}$.

*Sketch of proof:* Consider $(\mathbf{t}_1, \mathbf{s}_1)$ and $(\mathbf{t}_2, \mathbf{s}_2)$ two different optimal $\lambda$-**SLD** of a vector $\mathbf{x}$ satisfying the conditions of the theorem. Observing that $||\mathbf{t}_1||_0 + ||\mathbf{t}_2||_0 < ||\mathbf{t}_1 - \mathbf{t}_2||_0 = ||\mathbf{s}_1 - \mathbf{s}_2||_0$ and applying Theorem 1 yields a contradiction, thus proving the theorem.

The previous theorem provides a methodology to check if a $\lambda$-SLD is the only optimal, even if the decomposition problem is not convex. In a deeper sense, this theorem induces a data corruption limit below which the exact reconstruction of lowband signals is optimal in the sense of sparsity. We emphasize that this application is restricted to lowband signals, and that the theorem is stronger when the frequency band is low.

## IV. Experiments

In this section, we compare the lower bound obtained in Theorem 1 to other literature graph uncertainty principles and display some illustrations of the consequences of Theorem 2 to the classical denoising problem.

### A. Comparison of sparse graph uncertainty principles

In this first experiment, we compare the uncertainty bound of Theorem 1 to another sparse uncertainty bound that is derived from [7]. Applying exactly the same demonstration to a signal $\mathbf{x} \in \text{PW}_{\mathcal{G}}(\lambda_k)$, we obtain the following result:

$$||\mathbf{x}||_0 \, ||\hat{\mathbf{x}}||_0 \geq \frac{1}{\mu_k^2} \tag{9}$$

with $\mu_k$ the mutual coherence of the canonical basis and $(\mathbf{u}^1, ..., \mathbf{u}^k)$ the $k$ first eigenvectors of the Laplacian matrix. This leads to the following expression for $\mu_k$:

$$\mu_k = \max_{\substack{1 \leq i \leq N \\ 1 \leq j \leq k}} |u_i^j| \tag{10}$$

Moreover, $\mathbf{x} \in \text{PW}_{\mathcal{G}}(\lambda_k)$ implies that $||\hat{\mathbf{x}}||_0 \leq k$, which leads to a lower bound for the sparsity of $\mathbf{x}$ :

$$||\mathbf{x}||_0 \geq \frac{1}{k \, \mu_k^2} \tag{11}$$

We denote this lower bound by $B(k)$, and compare it to $C(\lambda_k)$ introduced in Theorem 1 for 4 different non-weighted symmetric graphs:

- $G_1$ : **Uniformly sampled graph**.
  A set of $500$ positions are uniformly drawn in the unit square. The graph is generated through a 4-nearest-neighbor procedure and symmetrized.
- $G_2$ : **4-blog graph**.
  A set of $500$ positions are drawn according to gaussian distributions of variance 1 centered on the corners of the unit square. The graph is generated through a 4-nearest-neighbor procedure and symmetrized.
- $G_3$ : **Minnesota graph**.
  The Minnesota road network from the MatlabBGL library [21], composed of $2642$ nodes and $3304$ vertices.
- $G_4$ : **Bunny graph**.
  The Stanford Bunny is a set of 3D points representing the surface of a bunny [22]. The graph is generated through a 3-nearest-neighbor procedure and symmetrized.

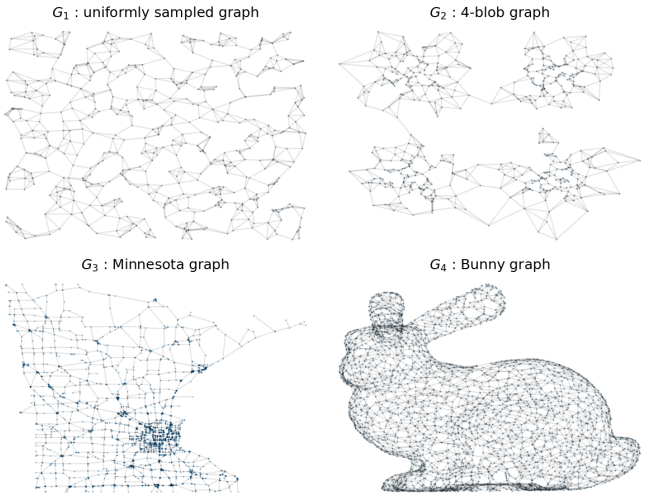


Fig. 2. Graphs $G_1$ to $G_4$ represented in their native spaces.

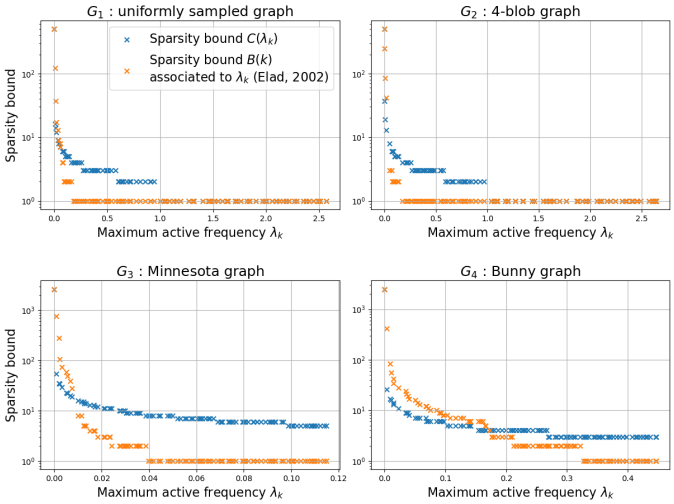


Fig. 3. Evolution of the bound of **Theorem 1** and $B(k)$ as a function of the maximum active frequency $\lambda_k$. Only the $k = 100$ first eigenvalues are presented.

Fig. 2 displays the studied graphs, and Fig. 3 displays the sparsity bounds on these graphs as a function of the maximum active frequency $\lambda_k$. As we can see, for each studied graph, the lower bound $C(\lambda_k)$ is higher that $B(k)$ for a significant range of $\lambda_k$, therefore ensuring an improved uncertainty inequality for the any signal in $\mathrm{PW}_{\mathcal{G}}(\lambda_k)$. This improvement can be especially seen for larger values of $\lambda_k$ for $G_3$ and $G_4$, which corresponds to $k > 100$ for these graphs. The poor quality of the bound $B(k)$ for large values of $k$ is a consequence of the definition of the mutual coherence $\mu_k$ that becomes inefficient quite fast. Indeed, as soon as an eigenvector $\mathbf{u}$ is included in $\mathrm{PW}_{\mathcal{G}}(\lambda_k)$ such that $||\mathbf{u}||_\infty \approx 1/k$, the bound $B(k)$ becomes close to 1. On the other hand, for lower values of $k$, the bound derived from [7] is usually better, especially for $k < 5$ for $G_1$ to $G_3$, and $k < 20$ for $G_4$ in the presented experiments. We emphasize that in the presented results, both bounds are correct, and therefore each can be used in the ranges where it is most effective. As a remark,

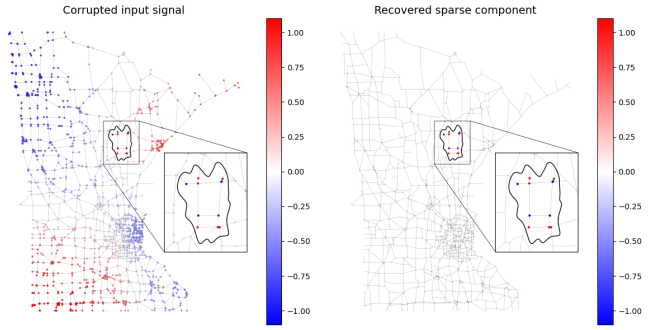


Fig. 4. The corrupted lowband signal (left) and the recovered sparse component (right). The set of corrupted nodes is highlighted by the black curve.

in the case of duplicate eigenvalues ($\lambda_k = \lambda_{k+1}$) which occurs naturally when working with non-weighted graphs, the bound $C(\lambda)$ remains unchanged whether the space $\mathrm{PW}_{\mathcal{G}}(\lambda_k)$ or $\mathrm{PW}_{\mathcal{G}}(\lambda_{k+1})$ is considered, whereas the bound $B(k)$ is usually less efficient for $\mathrm{PW}_{\mathcal{G}}(\lambda_{k+1})$.

### B. Sparse-Lowband signal decomposition

As a second illustration, we perform a **SLD** of a corrupted graph signal on the **Minnesota** graph ($G_3$). The "true" lowband signal $\mathbf{s}$ is generated with $k = 5$ active frequencies, with activities drawn according to a centered normal distribution of variance 100. A sparse corruption noise $\mathbf{t}$ is generated on a local set of 11 nodes, with values drawn according to a Rademacher distribution ($\mathbf{t}_i = 1$ or $-1$ with probability 0.5). The corrupted signal is given by $\mathbf{x} = \mathbf{s} + \mathbf{t}$, and a $\lambda_5$-**SLD** is sought through the resolution of the $L_1$ counterpart of problem (7). Figure 4 displays the corrupted signal and the recovered sparse component. In the presented experiment, the true decomposition of $\mathbf{x}$ is recovered, hereby identifying the corrupted nodes. Moreover, observing that the value $C(\lambda_5)/2 \approx 11.05$ is above the sparsity of the recovered component $||\mathbf{t}||_0 = 11$, the condition for the application of Theorem 2 are met, ensuring that the computed solution is also the unique optimal $\lambda_5$-**SLD**.

## V. Conclusion

In this paper we have presented a novel uncertainty principle for lowband graph signals, that lower-bounds the size of their supports. The main asset of this novel principle resides in its generality: the bound can be applied for any graph, independently of the studied structure. The problem of corrupted graph signal denoising is introduced through the formalism of optimal sparse-lowband decomposition, and an optimality criterion for this optimization problem is derived from the proposed uncertainty principle. Comprehensive illustrations of the proposed uncertainty principle are performed, and an instance of the task of sparse-lowband decomposition problem is presented. Interesting further developments include the refinement of the uncertainty principle under stronger assumptions and implications of the presented principle regarding other graph-related problems (more specifically, the presented uncertainty principle yields interesting results regarding interpolation).

## REFERENCES


[1] D. L. Donoho and P. B. Stark, "Uncertainty principles and signal recovery," *SIAM Journal on Applied Mathematics*, vol. 49, no. 3, pp. 906–931, 1989.

[2] G. B. Folland and A. Sitaram, "The uncertainty principle: A mathematical survey," *Journal of Fourier analysis and applications*, vol. 3, no. 3, pp. 207–238, 1997.

[3] T. Tao, "An uncertainty principle for cyclic groups of prime order," *Math. Res. Lett.*, vol. 12, Sep. 2003. DOI: 10.4310/MRL.2005.v12.n1.a11.

[4] R. Somaraju and L. W. Hanlen, "Uncertainty principles for signal concentrations," in *2006 Australian Communications Theory Workshop*, IEEE, 2006, pp. 38–42.

[5] S. Nam, "An uncertainty principle for discrete signals," *arXiv preprint arXiv:1307.6321*, 2013.

[6] D. L. Donoho, X. Huo, *et al.*, "Uncertainty principles and ideal atomic decomposition," *IEEE transactions on information theory*, vol. 47, no. 7, pp. 2845–2862, 2001.

[7] M. Elad and A. M. Bruckstein, "A generalized uncertainty principle and sparse representation in pairs of bases," *IEEE Transactions on Information Theory*, vol. 48, no. 9, pp. 2558–2567, 2002.

[8] S. G. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Transactions on signal processing*, vol. 41, no. 12, pp. 3397–3415, 1993.

[9] 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.

[10] B. Pasdeloup, R. Alami, V. Gripon, and M. Rabbat, "Toward an uncertainty principle for weighted graphs," in *2015 23rd European Signal Processing Conference (EUSIPCO)*, IEEE, 2015, pp. 1496–1500.

[11] O. Teke and P. P. Vaidyanathan, "Uncertainty principles and sparse eigenvectors of graphs," *IEEE Transactions on Signal Processing*, vol. 65, no. 20, pp. 5406–5420, 2017.

[12] A. Agaskar and Y. M. Lu, "A spectral graph uncertainty principle," *IEEE Transactions on Information Theory*, vol. 59, no. 7, pp. 4338–4356, 2013.

[13] H. J. Landau and H. O. Pollak, "Prolate spheroidal wave functions, fourier analysis and uncertainty—ii," *Bell System Technical Journal*, vol. 40, no. 1, pp. 65–84, 1961.

[14] 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.

[15] N. Perraudin, B. Ricaud, D. I. Shuman, and P. Vandergheynst, "Global and local uncertainty principles for signals on graphs," *APSIPA Transactions on Signal and Information Processing*, vol. 7, 2018.

[16] B. Pasdeloup, V. Gripon, R. Alami, and M. G. Rabbat, "Uncertainty principle on graphs," in *Vertex-Frequency Analysis of Graph Signals*, Springer, 2019, pp. 317–340.

[17] W. Erb, "Shapes of uncertainty in spectral graph theory," *IEEE Transactions on Information Theory*, vol. 67, no. 2, pp. 1291–1307, 2020.

[18] D. Cao and H. Yuan, "The distribution of eigenvalues of graphs," *Linear Algebra and Its Applications*, vol. 216, pp. 211–224, 1995.

[19] G. J. Rodgers, K. Austin, B. Kahng, and D. Kim, "Eigenvalue spectra of complex networks," *Journal of Physics A: Mathematical and General*, vol. 38, no. 43, p. 9431, 2005.

[20] I. J. Farkas, I. Derényi, A.-L. Barabási, and T. Vicsek, "Spectra of" real-world" graphs: Beyond the semicircle law," in *The Structure and Dynamics of Networks*, Princeton University Press, 2011, pp. 372–383.

[21] *Matlabbgl library*, http://www.cs.purdue.edu/homes/dgleich/packages/matlab_bgl/index.html.

[22] G. Turk and M. Levoy, "Zippered polygon meshes from range images," in *Proceedings of the 21st annual conference on Computer graphics and interactive techniques*, 1994, pp. 311–318.

## APPENDIX A
## PROOF OF THEOREM 1

In order to prove the Theorem 1, we need to introduce several objects. First, let $\mathcal{G} = (V, E)$ be a connected graph of size $N > 2$, and let $\mathbf{x}$ be a non-null graph signal on $\mathcal{G}$. We denote by $S_{\mathcal{G}}(\mathbf{x})$ the smoothness of $\mathbf{x}$ on $\mathcal{G}$.

**Smoothness of the sorted values.** We introduce an ordered sequence of values of $\mathbf{x}$, i.e. a sequence $x_{\sigma(1)} \leq ... \leq x_{\sigma(N)}$. Then, we define the smoothness of the ordered values as follows :

$$S^{sorted}(\mathbf{x}) \;=\; \sum_{i=1}^{N-1}(x_{\sigma(i)} - x_{\sigma(i+1)})^2 \tag{12}$$

We are now ready to introduce the main lemmas:

***Lemma** 1:* From graph smoothness to sorted smoothness. The following inequality holds:

$$S_{\mathcal{G}}(\mathbf{x}) \;\geq\; S^{sorted}(\mathbf{x}) \tag{13}$$

*Sketch of proof:*

Let us consider $\mathcal{T} = (V, E_{\mathcal{T}})$ a spanning tree of the graph $\mathcal{G}$. Such a tree exists because the graph $\mathcal{G}$ is connected, and we have $E_{\mathcal{T}} \subset E$, which yields:

$$S_{\mathcal{G}}(\mathbf{x}) \;\geq\; S_{\mathcal{T}}(\mathbf{x}) \tag{14}$$

Then, starting from the tree $\mathcal{T}_1 = \mathcal{T}$, we will construct a sequence of trees $(\mathcal{T}_k)_{1\leq k\leq N}$ such that:

$$\begin{cases} S_{\mathcal{T}_k}(\mathbf{x}) \;\geq\; S_{\mathcal{T}_{k+1}}(\mathbf{x}) & \forall\; k < N \\ S_{\mathcal{T}_N}(\mathbf{x}) = S^{sorted}(\mathbf{x}) \end{cases} \tag{15}$$

In order to generate such trees, we construct a sequence of path $(\mathcal{P}_k)_{1\leq k\leq N}$ on which the values of $\mathbf{x}$ are sorted and such that each $\mathcal{P}_k$ is a subgraph of $\mathcal{T}_k$. Please note that $\mathcal{T}_k$ and $\mathcal{P}_k$ are not subgraphs of the initial graph $\mathcal{G}$. These paths are constructed along the sequence of trees $(\mathcal{T}_k)_{1\leq k\leq N}$ as described by the following procedure :

- We initialize $\mathcal{T}_1$ to $\mathcal{T}$, and $\mathcal{P}_1$ to a single random node of the tree $\mathcal{T}_1$.
- Provided $\mathcal{T}_k$ and $\mathcal{P}_k$, we select a random node $n$ that is connected by an edge $e$ to $\mathcal{P}_k$ but that is not contained by $\mathcal{P}_k$. Then, we construct the path $\mathcal{P}_{k+1}$ by inserting the node $n$ in $\mathcal{P}_k$ such that the values of $\mathbf{x}$ are sorted along $\mathcal{P}_{k+1}$. The tree $\mathcal{T}_{k+1}$ is obtained by performing the same insertion and removing the edge $e$.

Figure 5 displays an example of the presented procedure. By construction, the following results hold for $k \leq N$:

$$\begin{cases} \mathcal{T}_k \text{ is a tree} \\ \mathcal{P}_k \text{ is a subgraph of } \mathcal{T}_k \\ \mathcal{P}_k \text{ is a path of size } k \\ \text{The restriction of } \mathbf{x} \text{ to } \mathcal{P}_k \text{ is ordered along } \mathcal{P}_k \end{cases} \tag{16}$$

Moreover, a simple enumeration of the possible cases yields for any $k < N$:

$$S_{\mathcal{T}_k}(\mathbf{x}) \;\geq\; S_{\mathcal{T}_{k+1}}(\mathbf{x}) \tag{17}$$

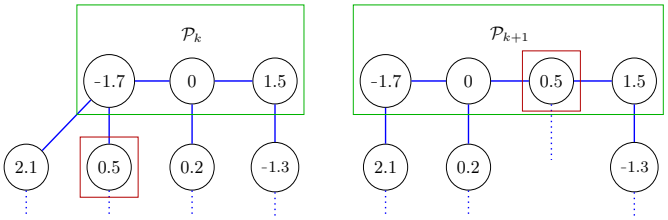


Fig. 5. An iteration of the algorithm transforming the tree $\mathcal{T}_k$ (left) into the tree $\mathcal{T}_{k+1}$ (right). The values of the signal $\mathbf{x}$ are displayed on the nodes. The paths $\mathcal{P}_k$ and $\mathcal{P}_{k+1}$ are delimited by the green boxes. The node $n$ selected randomly and added to $\mathcal{P}_{k+1}$ is highlighted in red. The insertion is performed such that the values of $\mathbf{X}$ are ordered in $\mathcal{P}_{k+1}$. All edges of $\mathcal{T}_k$ not represented on the figure remain unchanged in $\mathcal{T}_{k+1}$.

Observing that $\mathcal{T}_N = \mathcal{P}_N$, we have proven the assertions (15). Along with (14), the proof of lemma 1 is complete.

***Lemma** 2:* Sorted smoothness lower-bound.

Let us consider a $\mathbf{x} \in \mathbb{R}^N$ with exactly $M < N$ non-zero values and such that $||\mathbf{x}||_2 = 1$. Then we have:

$$S^{sorted}(\mathbf{x}) \;\geq\; 2\left(1 - \cos\left(\frac{\pi}{2M+1}\right)\right) \;=\; f(M) \tag{18}$$

*Sketch of proof:* As the count of zeros in a vector $\mathbf{x}$ does not change the value $S^{sorted}(\mathbf{x})$, the demonstration can be reduced to the case $M = N - 1$. Let us consider an integer $N > 1$ and a vector $\mathbf{y} \in \mathbb{R}^N$. For any index $k$ in $1..N$, we introduce the optimization problem $(\mathcal{P}_k)$:

$$\min_{\mathbf{y}} \sum_{i=1}^{N-1}(y_i - y_{i+1})^2 \quad \textbf{s.c.} \quad \begin{cases} ||\mathbf{y}||_2^2 & = 1 \\ y_k & = 0 \end{cases} \tag{19}$$

In this sketch of proof, we will only consider the case $k < N/2$, but it can be handled symmetrically. The optimal conditions for the Lagrangian yields solutions of the form:

$$y_j = \begin{cases} 0 & \text{if } j \leq k \\ \alpha \sin(\frac{(j-k)\pi}{2p+1}) & \text{if } j > k \end{cases} \tag{20}$$

Using $||\mathbf{y}||_2 = 1$, we can compute the minimum (19):

$$2\left(1 - \cos\left(\frac{\pi}{2(M-k)+1}\right)\right) \;=\; f(M,k) \tag{21}$$

Taking the minimum of (21) over $k$ yield the minimal value of $S^{sorted}(\mathbf{y})$, thus proving lemma 2.

**Proof of Theorem 1.**

Les us consider a graph signal $\mathbf{x} \in PW_{\mathcal{G}}(\lambda)$. Without loss of generality, we will assume that $||\mathbf{x}||_2 = 1$. From, lemma 1 and lemma 2 successively, he have:

$$S_{\mathcal{G}}(\mathbf{x}) \;\geq\; S^{sorted}(\mathbf{x}) \;\geq\; f(||\mathbf{x}||_0) \tag{22}$$

For a signal $\mathbf{x} \in PW_{\mathcal{G}}(\lambda)$, we have

$$S_{\mathcal{G}}(\mathbf{x}) \;\leq\; \lambda||\mathbf{x}||_2^2 \;=\; \lambda \tag{23}$$

Then, combining (22) and (23) yield

$$\lambda \;\geq\; 2\left(1 - \cos\left(\frac{\pi}{2||\mathbf{x}||_0 + 1}\right)\right) \tag{24}$$

A step by step inversion of this inequality leads to the result presented in Theorem 1.