HAL open science

Localization bounds for the graph translation

Benjamin Girault, Paulo Gonçalves, Shrikanth Narayanan, Antonio Ortega

To cite this version:

Benjamin Girault, Paulo Gonçalves, Shrikanth Narayanan, Antonio Ortega. Localization bounds for the graph translation. IEEE Global Conference on Signal and Information Processing, Dec 2016, Washington DC, United States. pp.331-335, 10.1109/GlobalSIP.2016.7905858 . hal-01368817

HAL Id: hal-01368817

https://inria.hal.science/hal-01368817v1

Submitted on 9 Dec 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

LOCALIZATION BOUNDS FOR THE GRAPH TRANSLATION

Benjamin Girault[⋆], *Paulo Gonçalves*[◇], *Shrikanth S. Narayanan*[⋆], *Antonio Ortega*[⋆]

[⋆] University of Southern California, Los Angeles, CA 90089, USA
[◇] Univ Lyon, Inria, ENS de Lyon, CNRS, UCB Lyon 1, 69342, Lyon, FRANCE

ABSTRACT

The graph translation operator has been defined with good spectral properties in mind, and in particular with the end goal of being an isometric operator. Unfortunately, the resulting definitions do not provide good intuitions on a vertex-domain interpretation. In this paper, we show that this operator does have a vertex-domain interpretation as a diffusion operator using a polynomial approximation. We show that its impulse response exhibit an exponential decay of the energy way from the impulse, demonstrating localization preservation. Additionally, we formalize several techniques that can be used to study other graph signal operators.

***Index Terms*—** graph signal processing, graph filters approximation, graph translation

1. INTRODUCTION

The field of graph signal processing aims at extending the tools of classical signal processing to irregular domains, and more precisely to finite and discrete irregular structures. In the past few years, the field has seen many successes, such as the graph Fourier transform [1, 2], shift invariant filters [2], ARMA models [3], several graph wavelet transforms [4–7], or vertex-frequency decompositions [7,8]. This is by no means an exhaustive list of GSP tools introduced so far, but an illustration to how tools involving the time shift operator have been extended to the graph setting. In particular, several operators equivalent to the time shift have been devised for graph domains. Chronologically, the first of these is the *graph shift* [2] defined in the vertex domain as an operator *shifting* the energy from one vertex to its neighbors and according to the edges weights. The second is the *generalized translation* [8] defined as a generalized convolution by a delta signal and having the property of localizing smooth signals about said delta signal. The operator we are interested in is the *graph translation* [9] defined in the Fourier domain as an isometric operator shifting the Fourier components by a complex phase and according to the graph frequencies. Finally, in [10], the authors identified the need of an isometric operator after [9], but the phase shifts of the Fourier components are not related to the graph frequencies. Of these shift operators only the last two are isometric [9], *i.e.* the only ones preserving the energy of a signal, and only the graph translation has meaningful (*w.r.t.* the graph) complex phase shifts.

Numerical experiments with this operator have shown a localization preservation property [9, 11], which has not been shown analytically. The localization property can be formalized in three different ways from the impulse response. The strictest is to verify that the impulse response is itself an impulse, as verified by the time shift operator. We showed in [11] that it is impossible to have both the isometry and this strict property. The second definition is to have an impulse response compactly supported. This property is shown by the graph shift. We believe that this definition is also incompatible with isometry, but this has not yet been proven. We are focusing here on the third definition involving an exponential decay of the energy as vertices get farther away from the center of the impulse. The localization property of [8] is of this kind. We use a similar method and draw a general framework for showing such a property. More precisely, given an operator, we approximate it using a polynomial of the adjacency matrix. Such a polynomial operator has itself an impulse response compactly supported since the kth power of the adjacency matrix corresponds to the k-hop paths in the graph. Showing that the error made by the approximation decays exponentially as the order of the polynomial increases is then enough to show that the operator is preserving localization according to the third definition. We show that the graph translation verifies this third definition of localization property.

We start by setting the definitions and notations of GSP we use. Then we give several general results and techniques that can be used to bound the response of any operator written as a function of a given matrix (typically the Laplacian matrix or the adjacency matrix). Finally, we recall the definition of the graph translation, and give its associated localization preservation result.

2. GRAPH SIGNAL PROCESSING

Let $\mathcal{G} = (V, E, w)$ be a weighted undirected graph, with V the set of vertices with $|V| = N$, E the set of edges such that if ij is an edge, ji is also an edge, and $w : E \to \mathbb{R}$ the weight of the edges. We denote $\mathbf{A}$ the adjacency matrix with $A_{ij} = w(i, j)$ the weight of the corresponding edge, $A_{ij} = 0$ if no such edge exists. Let $\mathbf{D} = \operatorname{diag}(d_1, \ldots, d_N)$ be the diagonal matrix of vertex degrees $d_i = \sum_j A_{ij}$. Let $\mathbf{L} = \mathbf{D} - \mathbf{A}$ be the Laplacian matrix of the graph, and $\mathcal{L} = \mathbf{D}^{-1/2}\mathbf{L}\mathbf{D}^{-1/2}$ the normalized Laplacian matrix.

The three matrices $\mathbf{A}$, $\mathbf{L}$ and $\mathcal{L}$ are real symmetric, and as such diagonalizable in an orthonormal basis. In particular, we have $\mathbf{L} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^*$ with $\mathbf{\Lambda}$ a diagonal matrix, and $\mathbf{U}$ the unitary[1] eigenvector matrix. The Graph Fourier Transform (GFT) is then defined as the projection on those eigenvectors and denoted $\widehat{\mathbf{x}} = \mathbf{U}^*\mathbf{x}$. Similar approaches can be derived using $\mathbf{A}$ and $\mathcal{L}$. The orthonormality allows then for Parseval's identity to be verified: $\sum_i |x_i|^2 = \|\mathbf{x}\|_2^2 = \|\widehat{\mathbf{x}}\|_2^2 = \sum_l |\widehat{x}(l)|^2$, where the signal energy in both the vertex and spectral domains is given by the ℓ_2-norm.

Let $\rho_{\mathcal{G}}^2 = \max_i 2d_i(d_i + \bar{d}_i)$, with $\bar{d}_i = \sum_j A_{ij}d_j/d_i$. The reduced[2] graph frequencies associated to the Laplacian-based GFT

This work was supported in part by NSF under grants CCF-1410009, CCF-1527874, CCF-1029373.

[1] U being real, it is orthogonal, but since we are dealing with complex operators and signal, the unitary property is more convenient.

[2] In the temporal domain, reduced frequencies lie in the interval $[-1/2, 1/2]$, with low frequencies close to zero.

are then $\nu_l = \frac{1}{2}\sqrt{\lambda_l/\rho_{\mathcal{G}}} \in [0, \frac{1}{2}]$, where λ_l is the lth eigenvalue[3] of $\mathbf{L}$. Similarly, for the GFT based on $\mathcal{L}$, we have $\nu_l = \frac{1}{2}\sqrt{\mu_l/2} \in [0, \frac{1}{2}]$, and for $\mathbf{A}$, we have $\nu_l = 1 - \gamma_l/\gamma_{\max} \in [0, 2]$ (with μ_l and γ_l the eigenvalues of $\mathcal{L}$ and A respectively). We refer the interested reader to [9, 11, 12] for justifications of these definitions.

Finally, we call a convolutive operator a graph filter $\mathbf{H}$ such that there exists a graph signal $\mathbf{h}$ verifying $\mathbf{H}\mathbf{x} = \mathbf{h} * \mathbf{x}$ where $*$ is the convolution operation defined as $\widehat{\mathbf{h} * \mathbf{x}}(l) = \widehat{h}(l)\widehat{x}(l)$. We denote $\widehat{\mathbf{H}}$ the dual operator in the spectral domain such that $\widehat{\mathbf{H}\mathbf{x}} = \widehat{\mathbf{H}}\widehat{\mathbf{x}}$. Note that equivalently, $\mathbf{H}$ and $\mathbf{L}$ (resp. $\mathcal{L}$, $\mathbf{A}$) are jointly diagonalizable or $\widehat{\mathbf{H}}$ is a diagonal matrix.

3. PRELIMINARY: POLYNOMIAL APPROXIMATION BOUNDS

We state the main fundamental result we rely on:

Theorem 1. *Let $f(x)$ be an analytical function and $p_K^{(f)}(x)$ a polynomial approximation of degree K of f such that $|f(x) - p_K^{(f)}(x)| \leq \kappa_f(K)$ on $X \subseteq \mathbb{R}$. Let $\mathbf{M}$ be an Hermitian[4] matrix with eigenvalues in X. We have then:*

$$\left\| f(\mathbf{M})\mathbf{x} - p_K^{(f)}(\mathbf{M})\mathbf{x} \right\|_2 \leq \kappa_f(K)\|\mathbf{x}\|_2.$$

Proof. First of all, $\mathbf{M}$ being Hermitian, it can be diagonalizable into $\mathbf{M} = \mathbf{V}\mathbf{\Theta}\mathbf{V}^*$, with $\mathbf{\Theta}$ diagonal and $\mathbf{V}$ unitary. It follows that $f(\mathbf{M}) = \mathbf{V}f(\mathbf{\Theta})\mathbf{V}^*$ and $p_K(\mathbf{M}) = \mathbf{V}p_K(\mathbf{\Theta})\mathbf{V}^*$. We have then:

$$\begin{aligned}
\|f(\mathbf{M})\mathbf{x} - p_K(\mathbf{M})\mathbf{x}\|_2^2 &= \|[f(\mathbf{\Theta}) - p_K(\mathbf{\Theta})]\,\mathbf{V}^*\mathbf{x}\|_2^2 \\
&= \sum_l |f(\theta_l) - p_K(\theta_l)|^2 |\widehat{\mathbf{x}}(l)|^2 \\
&\leq \kappa_f(K)^2 \sum_l |\widehat{\mathbf{x}}(l)|^2 \\
&= \kappa_f(K)^2 \|\widehat{\mathbf{x}}\|_2^2 = \kappa_f(K)^2 \|\mathbf{x}\|_2^2.
\end{aligned}$$

where the first and last equality follow from $\mathbf{V}$ being unitary, and the inequality follows from $\theta_l \in X$. □

We now state three lemmas giving the bound $\kappa_f(K)$.

Lemma 1. *Let f be an analytical function and p_K its polynomial approximation verifying:*

$$f(x) = \sum_{k=0}^{\infty} f_k\,(x-a)^k \qquad p_K(x) = \sum_{k=0}^{K} f_k\,(x-a)^k,$$

such that f is well defined[5] on the convex set $X \subseteq \mathbb{R}$. We have then:

$$\kappa_f(K) = \frac{1}{(K+1)!} \max_X \left\{ |f^{(K+1)}| \right\} \max_{x \in X} \left\{ |x-a|^{K+1} \right\}. \quad (1)$$

Proof. We first remark that using the Taylor series decomposition about a, we have $f_k = f^{(k)}(a)/k!$. Using Taylor's theorem, for all $x \in X$, there exits $\zeta \in [a, x]$ (or $\zeta \in [x, a]$ if $x < a$) such that:

$$f(x) - p_K(x) = \frac{f^{(K+1)}(\zeta)}{(K+1)!}(x-a)^{K+1}.$$

Maximizing the absolute value of the quantity above on the set X yields (1). □

Lemma 2. *Let f be an analytical function verifying:*

$$f(x) = \sum_{k=0}^{\infty} (-1)^k f_k\,(x-a)^k$$

with $f_k(x-a)^k$ of constant sign for any $x \in X \subseteq \mathbb{R}$. We have then:

$$\kappa_f(K) = \left| f_{K+1}(x-a)^{K+1} \right|. \quad (2)$$

Proof. The sum in f is an alternating series. □

Note that we need $X \subseteq [a, \infty)$ or $X \subseteq (\infty, a]$ for Lem. 2.

Lemma 3. *Let $f(x) = g(x)h(x)$ be the product of two analytical functions well defined on $X \subseteq \mathbb{R}$, and $p_{P,Q}^{(f)}(x) = p_Q^{(g)}(x)p_P^{(h)}(x)$ be its polynomial approximation. Let $\kappa_g(Q)$ and $\kappa_h(P)$ be the associated upper bounds. We have then:*

$$\kappa_f(P, Q) = \kappa_g(Q) \max_X |h| + \kappa_h(P)(\max_X |g| + \kappa_g(Q)).$$

Proof. We split the difference in two parts (we removed the arguments of the functions for conciseness):

$$\begin{aligned}
\left| f - p^{(f)} \right| &\leq \left| gh - p_Q^{(g)}h \right| + \left| p_Q^{(g)}h - p_Q^{(g)}p_P^{(h)} \right| \\
&= \left| g - p_Q^{(g)} \right| |h| + \left| p_Q^{(g)} \right| \left| h - p_P^{(h)} \right| \\
&\leq \kappa_g(Q) \max_X |h| + (\max_X |g| + \kappa_g(Q))\kappa_h(P),
\end{aligned}$$

where we used the property $|p_Q^{(g)}| \leq |g| + |g - p_Q^{(g)}|$ in the second inequality. □

We will be using Thm. 1 with $\mathbf{M} = \mathbf{L}/\rho_{\mathcal{G}}$, $\mathbf{M} = \mathcal{L}/2$ and $\mathbf{M} = \mathbf{I} - \mathbf{A}/\gamma_{\max}$ in section 5.

4. GRAPH TRANSLATION

The goal of this paper is to show that the graph translation of [9] preserves localization in the third sense: The impulse response has a decaying energy from the impulse vertex to its neighbors and beyond. Given a GFT and a set of graph frequencies $\{\nu_l\}_l$, we define the *Graph Translation* operator as the isometric convolutive operator verifying $\widehat{\mathbf{T}} = \exp\left(-\imath 2\pi \operatorname{diag}(\nu_0, \cdots, \nu_{N-1})\right)$ [11]. Isometry is understood here as the property $\|\mathbf{T}x\|_2 = \|x\|_2$. This leads to the following equivalent definitions:

Definition 1 (Graph Translation with L)**.**

$$\mathbf{T}_{\mathcal{G}} = \exp\left(-\imath\pi\sqrt{\frac{\mathbf{L}}{\rho_{\mathcal{G}}}}\right). \quad (3)$$

Note that $\mathbf{T}_{\mathcal{G}}$ is not just a normalized version of $\mathbf{L}$, but a completely new operator built from $\mathbf{L}$. A similar operator can be defined using the normalized Laplacian matrix:

Definition 2 (Graph Translation with $\mathcal{L}$)**.**

$$\mathcal{T}_{\mathcal{G}} = \exp\left(-\imath\pi\sqrt{\frac{\mathcal{L}}{2}}\right). \quad (4)$$

[3] The rationale behind the square root is the fact that the eigenvalues of the continuous Laplacian are squared frequencies (hence the loss of the sign), and the discrete graph Laplacian is similar to the continuous Laplacian [1].

[4] $\mathbf{M}$ and its conjugate transpose $\mathbf{M}^*$ are equal.

[5] *i.e.* the sum converges.

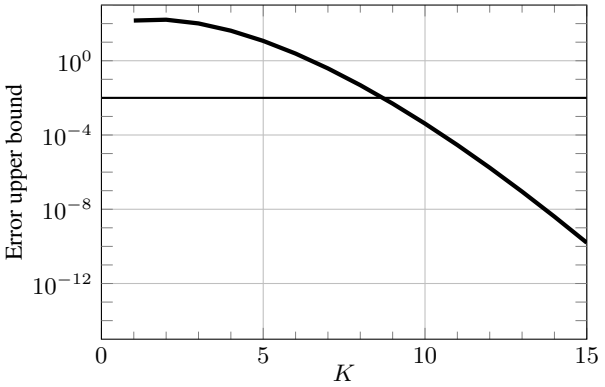

Fig. 1: Approximation curve associated to (7) with $\alpha = 1$.

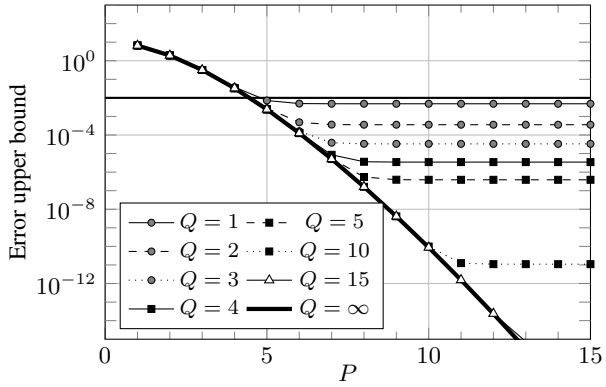

Fig. 2: Approximation curves associated to (10) for different values of Q with $\alpha = 1$ and $\varrho = 0.1$

We refer the interested reader to [11] for illustrations on the comparison between these two operators. A similar definition can also be devised with the adjacency matrix-based GSP framework. We denote it $\mathcal{A}_\mathcal{G}$ to stress the fact that this is more than just a normalization of the adjacency matrix.

Definition 3 (Graph Translation with A)**.**

$$\mathcal{A}_\mathcal{G} = \exp\left(-\imath\pi\left(\mathbf{I} - \frac{\mathbf{A}}{\gamma_{max}}\right)\right). \tag{5}$$

Def. 3 rescales first the adjacency matrix to have eigenvalues of $\mathbf{I} - \frac{\mathbf{A}}{\gamma_{\max}}$ within $[0, 2]$, *i.e.* the associated graph frequencies lie within $[0, 2\pi]$. The issue of the graph frequency sign does not appear here since the eigenvalues of the adjacency matrix are considered as linear representations of the graph frequencies according to [12].

Note that the operator $\mathcal{A}_\mathcal{G}$ is not guaranteed to be isometric when the graph is directed since the eigenvalues of $\mathbf{A}$ can be complex: if $z = x + \imath y$ is a complex eigenvalue of $\mathbf{A}/\gamma_{\max}$, we have $\exp(-\imath\pi(1 - z)) = e^{-\pi y}\exp(-\imath\pi(1 - x))$ as an eigenvalue of $\mathcal{A}_\mathcal{G}$, and y being non-zero, the operator is not isometric. As specified in section 2, we focus on undirected graphs which do not present such a difficulty.

5. GRAPH TRANSLATION LOCALIZATION PROPERTY

As shown in [11], the graph translation does not preserve the strictest localization preservation property in general. The main contribution of this paper is then to show that all three graph translations still have localization bounds in the vertex domain. To that end, we give polynomial approximations of these operators that verify an exponential decay of the error as the degree of the polynomial increases. As shown in the introduction, this guarantees a polynomial decay of the impulse response from the impulse vertex to its neighbors.

Since the square root in Def. 1 and 2 introduces an additional degree of complexity, we begin with $\mathcal{A}_\mathcal{G}$. All bounds are shown for the graph translation to the power α, where α plays the role of a *vertex-diffusion* factor. The error bounds we obtain show that localization decreases as the vertex-diffusion increases.

5.1. Adjacency-based Graph Translation

In this section, we have $\mathbf{M} = \mathbf{I} - \mathbf{A}/\gamma_{\max}$, such that:

$$\mathcal{A}_\mathcal{G}^\alpha = \cos(\alpha\pi\mathbf{M}) - \imath\sin(\alpha\pi\mathbf{M}).$$

We consider then the following analytic formulas on $\mathbb{R}_+$:

$$\cos(\alpha\pi x) = \sum_{k=0}^{\infty}(-1)^k\frac{(\alpha\pi)^{2k}}{(2k)!}x^{2k}$$

$$\sin(\alpha\pi x) = \sum_{k=0}^{\infty}(-1)^k\frac{(\alpha\pi)^{2k+1}}{(2k+1)!}x^{2k+1},$$

and the following truncated sum approximation of $\mathcal{A}_\mathcal{G}^\alpha$:

$$\mathcal{A}_{\mathcal{G},K}^\alpha = \sum_{k=0}^{K}(-1)^k\left[\frac{(\alpha\pi\mathbf{M})^{2k}}{(2k)!} - \imath\frac{(\alpha\pi\mathbf{M})^{2k+1}}{(2k+1)!}\right]. \tag{6}$$

Using Thm. 1 and Lem. 2 with $a = 0$ and $X = [0, 1]$, we have then:

$$\frac{\left\|\left[\mathcal{A}_\mathcal{G}^\alpha - \mathcal{A}_{\mathcal{G},K}^\alpha\right]\mathbf{x}\right\|_2}{\|\mathbf{x}\|_2} \quad\leq\quad \frac{(\alpha 2\pi)^{2K+2}}{(2K+2)!}\left(1 + \frac{\alpha 2\pi}{2K+3}\right) \tag{7}$$

5.2. Laplacian-based Graph Translation

We now have $\mathbf{M} = \mathbf{L}/\rho_\mathcal{G}$, such that:

$$\mathbf{T}_\mathcal{G}^\alpha = \cos(\alpha\pi\sqrt{\mathbf{M}}) - \imath\sqrt{\mathbf{M}}\frac{\sin(\alpha\pi\sqrt{\mathbf{M}})}{\sqrt{\mathbf{M}}} \tag{8}$$

Let $\varrho = \lambda_1/\rho_\mathcal{G}$ be the spectral gap of $\mathbf{M}$ and $\epsilon = (1-\varrho)/(1+\varrho) < 1$. Note that the case $\epsilon = 1$ corresponds to $\varrho = 0$, *i.e.* the graph is disconnected. However graphs are always assumed connected for GSP since we can perform GSP separately on each connected component for the same results. We consider then the following sums:

$$C^{(K)}(x) = \sum_{k=0}^{K}(-1)^k\frac{(\alpha\pi)^{2k}}{(2k)!}x^k$$

$$S^{(K)}(x) = \sum_{k=0}^{K}(-1)^k\frac{(\alpha\pi)^{2k+1}}{(2k+1)!}x^k$$

$$R^{(K)}(x) = \sqrt{\tfrac{1}{1+\epsilon}}\sum_{k=0}^{K}\frac{(-1)^k(2k)!}{(1-2k)(k!)^2 4^k}\left((1+\epsilon)x - 1\right)^k,$$

such that $\cos(\alpha\pi\sqrt{x}) = C^{(\infty)}(x)$, $\sin(\alpha\pi\sqrt{x})/\sqrt{x} = S^{(\infty)}(x)$ for $x \in \mathbb{R}_+$, and $\sqrt{x} = R^{(\infty)}(x)$ for $x \in [\varrho, 1]$. This last sum follows from the Taylor expansion of $\sqrt{1+y}$ about $y = 0$ for $y \in [-\epsilon, \epsilon]$ and $y = (1+\epsilon)x - 1$. We use the following approximation sum of

the graph translation:

$$T^{\alpha}_{\mathcal{G},P,Q} = C^{(P)}(\mathbf{M}) - \imath R^{(Q)}(\mathbf{M}) S^{(P)}(\mathbf{M}), \tag{9}$$

Where P may be different than Q to account for the slower convergence speed of the sum in $R^{(K)}$. Indeed, Stirling's formula for the factorial shows that the leading fraction in the sum in $R^{(K)}$ is equivalent to $(-1)^k(\sqrt{\pi k}(1-2k))^{-1}$, such that the convergence speed of $S^{(K)}$ and $C^{(K)}$ are much faster than that of $R^{(K)}$. We have then the following bounds:

$$\kappa_C(P) = \frac{(\alpha\pi)^{2P+2}}{(2P+2)!} \qquad \kappa_S(P) = \frac{(\alpha\pi)^{2P+3}}{(2P+3)!}$$

$$\kappa_R(Q-1) = \sqrt{\tfrac{1}{1-\epsilon^2}} \frac{(2Q)!}{(2Q-1)(Q!)^2 4^Q} \left(\epsilon(1-\epsilon)\right)^Q,$$

using Lem. 2 on $[0,1]$ for C and S, and Lem. 1 on $[\varrho,1]$ for R with $f(y) = \sqrt{1+y}$ on $[-\epsilon,\epsilon]$. Notice also that $S^{(P)}(0) = 0 = S(0)$ for all P, such that:

$$\begin{aligned} R^{(Q)}(\mathbf{M}) S^{(P)}(\mathbf{M})\mathbf{x} \\ &= \sum_{l=0}^{N-1} R^{(Q)}\left(\frac{\lambda_l}{\rho_{\mathcal{G}}}\right) S^{(P)}\left(\frac{\lambda_l}{\rho_{\mathcal{G}}}\right) \widehat{x}(l)\mathbf{u}^{(l)} \\ &= \sum_{l=1}^{N-1} R^{(Q)}\left(\frac{\lambda_l}{\rho_{\mathcal{G}}}\right) S^{(P)}\left(\frac{\lambda_l}{\rho_{\mathcal{G}}}\right) \widehat{x}(l)\mathbf{u}^{(l)}, \end{aligned}$$

for all P and Q, with $\mathbf{u}^{(l)}$ the lth Fourier mode. It is then enough to have κ_R and κ_{RS} for the interval $[\lambda_1/\rho_{\mathcal{G}}, 1] = [\varrho, 1]$ to obtain the result of Thm. 1. This also shows:

$$\mathbf{T}_{\mathcal{G}} = \mathbf{T}_{\mathcal{G},\infty,\infty}.$$

Using Lem. 3, we obtain on $[\varrho, 1]$:

$$\kappa_{RS}(P,Q) = \kappa_R(Q) + \kappa_S(P)\left(1 + \kappa_R(Q)\right).$$

and finally:

$$\begin{aligned} &\left\| \left[\mathbf{T}^{\alpha}_{\mathcal{G}} - \mathbf{T}^{\alpha}_{\mathcal{G},P,Q}\right] \mathbf{x} \right\|_2 \\ &\quad \leq \left[\kappa_C(P) + \kappa_S(P) + \kappa_R(Q)(1 + 2\kappa_S(P))\right] \|\mathbf{x}\|_2. \end{aligned} \tag{10}$$

Overall, this bound is dominated by the term $\kappa_R(Q)$ since $\kappa_C(P)$ and $\kappa_S(P)$ decrease very quickly with P due to the factorial. Fig. 2 shows the evolution of $\|[\mathbf{T}_{\mathcal{G}} - \mathbf{T}_{\mathcal{G},P,Q}]\mathbf{x}\|_2/\|\mathbf{x}\|_2$ according to P and Q. We see that the error is mainly explained by Q: Larger values of P do not decrease the error, with a plateau on the error explained by Q. The approximation error is about 1% when $P = 5$ and $Q = 1$, and each increment of one of Q leads to a an approximation error divided by about $1/\epsilon$ (for a large P). With these values of P and Q, we have approximated $\mathbf{T}_{\mathcal{G}}$ with a local operator acting in a 6-hops neighborhood of a given vertex.

Note that the faster convergence compared to Fig. 1 is explained by two factors. First, the graph frequencies of $\mathbf{L}$ lie in $[0, 1/2]$ while those of $\mathbf{A}$ lie in $[0,1]$ such that the error curve is slightly shifted towards bigger values of K on Fig. 1. Second, the use of the square root of $\mathbf{L}$ for $\mathbf{L}_{\mathcal{G}}$ compared to the plain matrix $\mathbf{A}$ for $\mathcal{A}_{\mathcal{G}}$ leads to a steeper slope on Fig. 2 (for $Q = \infty$), and a better approximation.

Fig. 3 shows the minimal value of $P + Q$ such that the approximation in (10) yields an error smaller than ξ from 0.5 to 10^{-4} and for different values of the vertex-diffusion factor α. We observe that the more the graph translation is applied, the more diffused the signal can be with a looser bound. For $\alpha = 1$, we obtain the result that 50% of the energy is within a 3-hops radius, 90% of the energy is within

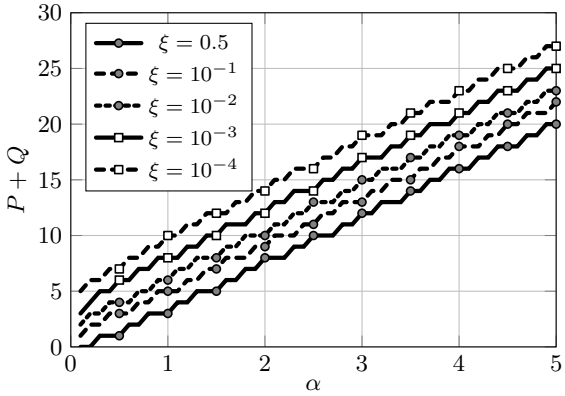

Fig. 3: Minimal value of $P + Q$ to have a maximum error of ξ in (10) for $\varrho = 0.1$.

a 5-hops radius and 99% is within a 6-hops radius. Approximating the graph translation through a polynomial operator is therefore a trade-off between accuracy and the size of the neighborhood. Also, the slopes on Fig. 3 are steep and the localization decreases much faster as α increases compared to the graph shift operator (having a unitary slope), illustrating another trade-off between localization and isometry.

Finally, note that this is an upper bound, and the output may remain highly localized in the vertex domain no matter the value of α. This is especially the case if one Fourier mode is highly localized: If $\mathbf{u}^{(l)} \approx \delta_i$, then $\mathbf{T}^{\alpha}_{\mathcal{G}}\delta_i \approx \mathbf{T}^{\alpha}_{\mathcal{G}} u^{(l)} = e^{-\imath\alpha\nu_l}\mathbf{u}^{(l)} \approx e^{-\imath\alpha\nu_l}\delta_i$, *i.e.* $\mathbf{T}^{\alpha}_{\mathcal{G}}\delta_i$ is localized about vertex i. Our bound corresponds then to the worst case scenario of delocalized Fourier modes.

5.3. Normalized-Laplacian-based Graph Translation

The technique is the same as before using $\mathbf{M} = \mathcal{L}/2$ instead of $\mathbf{M} = \mathbf{L}/\rho_{\mathcal{G}}$, and leading to the exact same bound with $\varrho = \mu_1/2$:

$$\begin{aligned} &\left\| \left[\mathcal{T}^{\alpha}_{\mathcal{G}} - \mathcal{T}^{\alpha}_{\mathcal{G},P,Q}\right] \mathbf{x} \right\|_2 \\ &\quad \leq \left[\kappa_C(P) + \kappa_S(P) + \kappa_R(Q)(1 + 2\kappa_S(P))\right] \|\mathbf{x}\|_2. \end{aligned} \tag{11}$$

6. CONCLUSION

We have shown in this paper that the graph translation defined in [9] indeed preserves the localization of a graph signal in the sense of an exponential decay, as suggested by the numerical experiments. Equivalently, we showed that this operator is acting as a diffusion operator in the sense that the energy spreads from one vertex to its neighbors. Also, when iterated, the bounds are less and less tight such that the energy may spread across farther and farther vertices of the graph just like a diffusion process would. Note that the bounds we showed only depend on the spectral gap of the Laplacian, and as such are very general bounds. We refer the interested reader to [13] for a review on the spectral gap of the Laplacian (called the algebraic connectivity of the graph). In addition, we showed several techniques applicable to other operators that can be used to give them polynomial approximation and to bound the error and the spread of the impulse response.

One last noticeable remark is that the impulse response of the graph translation can be seen as a generalized translation of a particular graph signal: $T_{\mathcal{G}}\delta_i = h * \delta_i$ where $\widehat{h}(l) = \exp(-\imath 2\pi\nu_l)$. This property allows to further study the impulse response of the graph translation using the results dedicated to the generalized translation, and will be used in a future paper.

REFERENCES

[1] David I. Shuman, Sunil K. Narang, Pascal Frossard, Antonio Ortega, and Pierre 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.

[2] Aliaksei Sandryhaila and José M. F. Moura, "Discrete Signal Processing on Graphs," *IEEE Transactions on Signal Processing*, vol. 61, no. 7, pp. 1644–1656, 2013.

[3] Andreas Loukas, Andrea Simonetto, and Geert Leus, "Distributed Autoregressive Moving Average Graph Filters," *IEEE Signal Processing Letters*, vol. 22, no. 11, pp. 1931–1935, Nov 2015.

[4] Ronald R. Coifman and Mauro Maggioni, "Diffusion wavelets," *Applied and Computational Harmonic Analysis*, vol. 21, no. 1, pp. 53–94, 2006.

[5] David K. Hammond, Pierre Vandergheynst, and Rémi Gribonval, "Wavelets on graphs via spectral graph theory," *Applied and Computational Harmonic Analysis*, vol. 30, no. 2, pp. 129–150, 2011.

[6] Nora Leonardi and Dimitri Van De Ville, "Tight Wavelet Frames on Multislice Graphs," *IEEE Transactions on Signal Processing*, vol. 61, no. 13, pp. 3357–3367, July 2013.

[7] David I. Shuman, Christoph Wiesmeyr, Nicki Holighaus, and Pierre Vandergheynst, "Spectrum-Adapted Tight Graph Wavelet and Vertex-Frequency Frames," *IEEE Trans. Signal Processing*, vol. 63, no. 16, pp. 4223–4235, 2015.

[8] David I. Shuman, Benjamin Ricaud, and Pierre Vandergheynst, "Vertex-frequency analysis on graphs," *Applied and Computational Harmonic Analysis*, vol. 40, no. 2, pp. 260–291, 2016.

[9] Benjamin Girault, Paulo Gonçalves, and Éric Fleury, "Translation on Graphs: An Isometric Shift Operator," *Signal Processing Letters, IEEE*, vol. 22, no. 12, pp. 2416–2420, Dec 2015.

[10] Adnan Gavili and Xiao-Ping Zhang, "On the Shift Operator, Graph Frequency and Optimal Filtering in Graph Signal Processing," *ArXiv e-prints*, Nov. 2015.

[11] Benjamin Girault, *Signal Processing on Graphs - Contributions to an Emerging Field*, Phd thesis, Ecole normale supérieure de lyon - ENS LYON, Dec. 2015.

[12] Aliaksei Sandryhaila and José M. F. Moura, "Discrete Signal Processing on Graphs: Frequency Analysis," *Signal Processing, IEEE Transactions on*, vol. 62, no. 12, pp. 3042–3054, June 2014.

[13] Nair Maria Maia de Abreu, "Old and new results on algebraic connectivity of graphs," *Linear Algebra and its Applications*, vol. 423, no. 1, pp. 53–73, 2007.