ORTHOGONAL TRANSFORMS FOR SIGNALS ON DIRECTED GRAPHS

Julia Barrufet¹, Antonio Ortega²

¹Universitat Politècnica de Catalunya, Barcelona, Spain
²University of Southern California, Los Angeles, USA

ABSTRACT

In this paper we consider the problem of defining transforms for signals on directed graphs, with a specific focus on defective graphs where the corresponding graph operator cannot be diagonalized. Our proposed method is based on the Schur decomposition and leads to a series of embedded invariant subspaces for which orthogonal basis are available. As compared to diffusion wavelets, our method is more flexible in the generation of subspaces, but these subspaces can only be approximately orthogonal.

Index Terms— Graph signal processing, directed graphs, Schur decomposition, orthogonal transforms, Diffusion wavelets.

1. INTRODUCTION

Graph signal processing (GSP) relies on the availability of a suitable frequency definition and corresponding graph Fourier transform (GFT) [1,2,3,4]. For undirected graphs, graph operators are symmetric and the GFT is orthogonal. For directed graphs, however, the GFT is not orthogonal. It can be obtained from the eigenvectors of the graph operator, if it is diagonalizable, or using the Jordan form, if not [5]. While symmetrization strategies [5] can be used to convert directed graphs into undirected graphs or slight modifications can be applied to make a defective directed graph operator diagonalizable (e.g., teleportation in PageRank [6,7]), in this paper we focus on the case where it is desirable not to modify the graph operator.

For directed graphs with an adjacency matrix that is not diagonalizable, many studies use the Jordan canonical form [8] to find a spectral decomposition [2,9,10]. More recently, in [11] a method to replace a given adjacency operator by a diagonalizable operator obtained via the Jordan-Chevalley decomposition. While this method leads to a diagonalizable graph operator, it starts by computing the Jordan form, which is well known to be numerically unstable. Indeed this is a common limitation for all solutions based on the Jordan decomposition. This situation leads to investigate alternative approaches. For example, in [12] the graph Fourier basis is built as the set of orthonormal vectors that minimizes a continuous extension of the graph cut size, known as the Lovazs extension. Other alternatives to the Jordan form are presented in [13] and [14].

Finally, an alternative approach would be to use other frequency decompositions as alternatives to the GFT. Examples include the spectral graph wavelet transforms [15]. Our proposed method is inspired by the diffusion wavelets design [16,17]. This model offers a tool to build a spectral decomposition for diffusion operators such as the adjacency matrix of a graph. The proposed method, the graph Schur transform (GST), aims to offer a valid operator to perform a spectral decomposition of a graph that can be used even in the case of defective matrices. The numerical advantages of the Schur decomposition were noted in [18] and this decomposition was used to obtain block diagonalization of matrices. This approach is used in the GraSP toolbox [19] (see also Appendix B in [4]). In this paper we introduce a novel transform design based on the Schur decomposition.

2. PRELIMINARIES

2.1. Graph signal processing and graph operators

We define graphs $G(V,E)$ with a set of nodes $V$ and edges $E$, where an edge $e_{ij}$ represents a link between node $i$ and node $j$. A graph can be weighted, if any edge $e_{ij}$ can take a real positive weight $\omega_{ij}$, or unweighted, if the weight for all its edges is 1. We focus on directed graphs where where $e_{ij}$ or $e_{ji}$ may not exist and in general $\omega_{ij} \neq \omega_{ji}$. Given a graph $G$ with $N$ nodes, we consider as the main graph operator the adjacency matrix $A$, an $N \times N$ square matrix, where the entry $a_{ij}$ will correspond to the weight $w_{ij}$. Graph signals are real vectors $x \in \mathbb{R}^N$, where the entry $x(i)$ is the real scalar corresponding to the signal associated to node $i$ and graph filters are constructed as polynomials of the graph operator $Z$ (in this paper we focus on $Z = A$).

For a diagonalizable matrix $Z$ we can construct $U$, an invertible matrix where each column is one of the eigenvectors of $Z$, and write $Z = U \Lambda U^{-1}$ where $\Lambda$ is the diagonal matrix with the eigenvalues of $Z$, and $U$ a matrix containing its eigenvectors. Now we can write any graph signal $x$ in terms of its graph frequencies:

$$x = U \tilde{x} \quad \text{where} \quad \tilde{x} = U^{-1} x \quad (1)$$

is defined as the Graph Fourier Transform (GFT) of the graph signal $x \in \mathbb{R}^N$. In this paper we focus on directed graphs (digraphs) having a defective (non-diagonalizable) adjacency matrix $A$. As shown in [20], a simple experiment where random Erdős-Rényi digraphs [21] are generated demonstrates that sparse random digraphs are very likely to be defective.

2.2. Schur Decomposition

Given $Z$ a fundamental graph operator, the Schur decomposition (SD) [22] of $Z$ can be obtained even if $Z$ is not diagonalizable as $Z = U T U^H$ where $U$ is unitary, $U^H$ is the Hermitian transpose of $U$ and $T$ is an upper triangular matrix with the eigenvalues of $Z$ along its diagonal. The Schur decomposition is not unique, a different one is obtained via the Jordan-Chevalley decomposition. While this is the case where it is desirable not to modify the graph operator. For example, in [12] the graph Fourier basis is built as the set of orthonormal vectors that minimizes a continuous extension of the graph cut size, known as the Lovazs extension. Other alternatives to the Jordan form are presented in [13] and [14].

Finally, an alternative approach would be to use other frequency decompositions as alternatives to the GFT. Examples include the spectral graph wavelet transforms [15]. Our proposed method is inspired by the diffusion wavelets design [16,17]. This model offers a tool to build a spectral decomposition for diffusion operators such as the adjacency matrix of a graph. The proposed method, the graph Schur transform (GST), aims to offer a valid operator to perform a spectral decomposition of a graph that can be used even in the case of defective matrices. The numerical advantages of the Schur decomposition were noted in [18] and this decomposition was used to obtain block diagonalization of matrices. This approach is used in the GraSP toolbox [19] (see also Appendix B in [4]). In this paper we introduce a novel transform design based on the Schur decomposition.
are subspaces of \( \mathbb{R}^N \) invariant under \( Z \). That is, if \( x \in F_k \) then \( Zx \in F_k \). Note that since \( U \) is unitary, we have that \( U^H = U^{-1} \) so that \( ZU = UT \). Thus, defining the upper triangular matrix \( T \) as the sum of a diagonal matrix of eigenvalues \( D \) and a nilpotent matrix \( N \) we have that:

\[
Z[u_1 \cdots u_N] = [u_1 \cdots u_N](D + N) = [u_1 \cdots u_N] \begin{pmatrix} \lambda_1 & n_{12} & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_N \end{pmatrix}
\]

where \( \lambda_1, \ldots, \lambda_N \) represent the eigenvalues of the \( Z \) operator and \( n_{ij} \) the corresponding element of the \( N \) matrix. Now the multiplication of each of the \( u_i \) by the operator \( Z \) can be expressed as

\[
Zu_i = \sum_{j=1}^{i-1} n_{ji}u_j + \lambda_i u_i \in F_i.
\]

(4)

Therefore, subspaces \( F_i \) are invariant but there is an overlap, since \( F_{i-1} \subset F_i \).

### 2.3. Diffusion Wavelets

Diffusion wavelets (DW) \[16\] achieve spectral and vertex domain localization, without compact support in either of these domains \[4\]. The key observation in the DW design is that successive powers of the diffusion operator \( Z \) will have increasingly lower numerical rank. This holds for a normalized operator \( Z \), i.e., whose eigenvalue magnitudes are in the range from 0 to 1, since the eigenvalues of \( Z^k \) are \( \lambda_i^k \) and can become arbitrarily small as \( k \) increases. The DW design can be summarized as follows \[4\]. First, select sets:

\[
\Phi_{Z^2} = \{ \lambda_{i,1}^2 v_1, \ldots, \lambda_{i,N}^2 v_N \}
\]

(5)

where \( \{ v_1, \ldots, v_N \} \) are the eigenvectors of \( Z \). Each of these sets corresponds to the eigenvalues of the matrices obtained from consecutive dyadic powers of the operator \( Z \) such that \( \lambda < \epsilon \). Then \( V_0 = \mathbb{R}^N \) and \( V_i = \text{span}(\Phi_i) \), where \( \Phi_i \) is an \( \epsilon \)-span of \( \Phi_{Z^i} \) \[15\]. At any stage we then find a subspace \( W_i \) such that

\[
V_i \oplus W_i = V_{i-1}.
\]

(6)

Then choosing a specific \( i \) we write

\[
V_0 = V_i \oplus W_i \oplus W_{i-1} \oplus \ldots \oplus W_1
\]

(7)

so that the basis for the \( W_i \) spaces form the orthogonal wavelets and the basis for \( V_i \) correspond to the scaling function. From this design we can finally obtain \( M + 1 \) orthogonal subspaces formed by orthogonal vectors and corresponding to the eigenvalues of \( Z \) in a decreasing order \( \{ V_1, V_2, \ldots, V_M, V_M, V_{i-1} \} \). DW has many advantages for directed graphs, including cases where \( Z \) is defective, since the basis obtained are orthogonal and an eigendecomposition of \( Z \) is not required \[16\] and \[17\]. However, this design is not flexible due to its dyadic structure (powers of two in \( Z^i \)), which may lead to difficulties when the eigenvalue distribution is highly irregular.

### 3. GRAPH SCHUR TRANSFORM (GST)

Both, DW and the SD decompose the space \( V_0 = \mathbb{C}^N \) into the direct sum of orthogonal subspaces. For SD \( \mathbb{C}^N = \bigoplus_{i=1}^N E_i \), where the subspaces \( E_i = \text{span}(u_i) \) are orthogonal, but they are not invariant. Instead, we have that the \( F_i \) spaces in \( \mathbb{C}^2 \) are invariant, but clearly they are not orthogonal to each other since \( F_{k-1} \subset F_k \).

In the DW design, for a sufficiently large \( i \), \( V_i \) contains only signals that are in the subspace corresponding to the largest eigenvalue, while \( W_1 \) corresponds to the smaller eigenvalues, and for increasing levels the subspace \( W_i \) corresponds to higher eigenvalues. Also, in DW subspaces are orthogonal by construction but they are not exactly invariant to multiplication by the graph operator. Note that if the eigenvalues are ordered in decreasing order and \( i \) is sufficiently large, then \( F_1 = E_1 = V_i \).

#### 3.1. GST Construction

The proposed GST is built on the construction of a set of invariant subspaces derived from the SD. Start with a normalized operator \( Z \), e.g., the adjacency matrix \( A \) divided by the magnitude of its largest eigenvalue \( \| A \| \). Given the embedded spaces \( F_1 \subset F_2 \subset \ldots \subset F_k \) from SD, we introduce a series of subspaces \( G_1, \ldots, G_k \), such that \( G_1 = F_1 \), \( G_1 \oplus G_2 = F_2 \) and, in general,

\[
G_{i-1} \oplus G_i = F_i.
\]

(8)

\( G_1 \) contains the basis in SD associated to eigenvalues \( \lambda_1 \) to \( \lambda_1 \), but \( G_2 \) contains a basis corresponding to the eigenvalues from \( \lambda_{i+1} \) to \( \lambda_{i+2} \) (which will be renamed \( \lambda_1^{(2)} \) to \( \lambda_2^{(2)} \)). Then to form a complete basis for \( \mathbb{R}^N \) we can represent this space as:

\[
G_1 \oplus G_2 \ldots \oplus G_k = F_k = \mathbb{R}^N
\]

(9)

where \( F_k \) corresponds to the last subspace of dimension \( N \).

From the SD \( Z = UUT^H \) we can define \( G_1 \) as:

\[
G_1 = \text{span}(u_1, \ldots, u_1),
\]

(10)

where \( u_1, u_2, \ldots \) correspond to the first columns of the matrix \( U \) and \( u_1 \) the column corresponding to the last eigenvalue included in the first subspace. Therefore, the basis for this subspace will be

\[
U_1 = [u_1, u_2, \ldots, u_1]
\]

(11)

Due to the orthogonal nature of \( U \) we have obtained an orthogonal invariant basis for the vectors in \( G_1 \).

To build the next subspace, we start by reordering the eigenvalues as \( \{ \lambda_{i+1}, \lambda_{i+2}, \ldots, \lambda_1, \lambda_2, \ldots, \lambda_1 \} \) and renaming the eigenvalues as \( \{ \lambda_1^{(2)}, \ldots, \lambda_N^{(2)} \} \). Now we can rebuild the Schur matrix in the new order:

\[
T^{(2)} = \begin{pmatrix} \lambda_1^{(2)} & n_{12}^{(2)} & n_{13}^{(2)} \\ 0 & \lambda_2^{(2)} & n_{23}^{(2)} \\ \vdots & \vdots & \ddots \\ 0 & 0 & \lambda_N^{(2)} \end{pmatrix} = U^{(2)T}ZU^{(2)}
\]

(12)

and repeating the procedure used for \( G_1 \), we can build a basis for the space \( G_2 \) taking the first \( i_2 \) columns of the matrix \( U^{(2)} \). The first column will be an eigenvector for \( \lambda_1^{(2)} \) and the other \( i_2 - 1 \) vectors will be invariant to the subspace they form. With these vectors we will have an orthogonal invariant basis, \( G_2 \), which will form the orthogonal matrix \( U_2 \):

\[
U_2 = [u_1^{(2)}, \ldots, u_i^{(2)}]
\]

At this point we have built 2 different orthogonal and invariant basis for two sets of energies of the graph: \( \lambda_1, \ldots, \lambda_1 \) and \( \lambda_1^{(2)}, \ldots, \lambda_1^{(2)} \). Repeating this procedure as many times as necessary will result in
Spectral Localization

The proposed method offers, compared to the proposed method, assume the only parameter to decide is the desired number of subspaces $M$. Then, as a criterion to group eigenvalues into different subspaces will be to find the $M − 1$ separation points where the distance between consecutive eigenvalues is greater. We test GST and DW for a random synthetic graph of 500 nodes setting the tuning parameter $M = 30$ for GST and the $\epsilon$-span parameter $\epsilon = 10^{-5}$ for DW (see Section 2.3). For this graph, the distribution of eigenvalues into subspaces is shown in Fig. 2. It is clear from these results that the GST shows a better spectral localization with a smaller variance in the magnitude of eigenvalues contained in each subspace, providing a more precise handling of the subspaces.

4. EXAMPLE GST CONSTRUCTION AND EXPERIMENTS

Dimensions of subspaces in DW and GST To illustrate the proposed method, assume the only parameter to decide is the desired number of subspaces $M$. Then, as a criterion to group eigenvalues into different subspaces will be to find the $M − 1$ separation points where the distance between consecutive eigenvalues is greater. We test GST and DW for a random synthetic graph of 500 nodes setting the tuning parameter $M = 30$ for GST and the $\epsilon$-span parameter $\epsilon = 10^{-5}$ for DW (see Section 2.3). For this graph, the distribution of eigenvalues into subspaces is shown in Fig. 2. It is clear from these results that the GST shows a better spectral localization with a smaller variance in the magnitude of eigenvalues contained in each subspace, providing a more precise handling of the subspaces.

GST for directed graphs The ultimate goal of this research was to find a method to build a basis valid to use GSP tools on directed graphs, and especially on those with a defective adjacency matrix. For a non-diagonalizable or defective graph a complete set of linearly independent eigenvectors does not exist, so a complete basis of eigenvectors for $\mathbb{R}^N$ cannot be formed.

To evaluate experimentally the proposed method, we generate $U_S$ matrices with the GST method for random synthetic graphs. We generate graphs of different sizes and select a different number of subspaces in each case. In particular, the performed verification consisted on the following tests:

- Graphs of 50, 100, 150, 200, 300 and 500 graphs were created.
- For each size, the GST was computed for different numbers of subspaces ($M$): $N/25$, $N/10$, $N/5$ and $N/2$.
- All graphs where created with a random edge probability $p \in [5/N, 1/N]$.
- For each selected $N$ and $M$, 100 graphs were created to calculate the proportion of non-diagonalizable graphs.

For each graph we construct a matrix $U_e$ with all the eigenvectors for the adjacency matrix of the graph. The adjacency matrix was defective in all cases, with a number of eigenvectors averaging around 90 (Fig. 3), while as expected the matrix $U_S$ was fullrank for all graphs and indeed provides a set of basis vectors for the space of graph signals.

![Fig. 1. Diagram showing how subspaces are created in the GST method for an example with N=10 where $\lambda_1$, $\lambda_2$, etc represent the normalized eigenvalues.](image1)

![Fig. 3. Histogram showing the rank of the Schur $U_S$ matrix (yellow) and the matrix of eigenvectors $U$ (blue) for 500 graphs with N=100. The graphic clearly shows that, while the adjacency matrix was defective for the 500 tested graphs, the GST always forms a complete basis for $\mathbb{R}^N$, with $\text{ran}(U_f) = N$.](image3)
Note that in the case of directed acyclic graphs our method will not perform correctly for various reasons. First, the adjacency matrix $A$ cannot be divided by the maximum eigenvalue since, by definition, all the eigenvalues are zero for a DAG. Second, the adjacency matrix for a DAG is a nilpotent matrix, that is, an upper triangular matrix with zeros (the eigenvalues) in its diagonal. Therefore, we have that the Schur matrix $T$ has the same form as the adjacency matrix $A$. This means that in this case the transformation matrix $U$ in the expression $A = U T U^H$ would be $U = I$. Which makes it not useful for our purpose.

Finally, the construction of our method groups eigenvalues into $M$ groups, separating them when their magnitudes are different enough. This is not possible in the case of DAGs, where all the eigenvalues have magnitude zero.

5. CONCLUSIONS

Our proposed method has several useful properties, including invariance, orthogonality within each subset and flexibility in the subset construction, including the option of constructing near orthogonal subspaces. Our future work may focus on a more thorough evaluation of the potential benefits of GSTs in applications as well further study of techniques that may be valid for directed acyclic graphs, where GST cannot be used.

6. REFERENCES

[1] 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, May 2013.

[2] Aliksei Sandryhaila and José M. F. Moura, “Discrete signal processing on graphs,” IEEE Transactions on Signal Processing, vol. 61, no. 7, pp. 1644–1656, Apr 2013.

[3] A. Ortega, P. Frossard, J. Kovačević, J. M. F. Moura, and P. Vandergheynst, “Graph signal processing: Overview, challenges, and applications,” Proceedings of the IEEE, vol. 106, no. 5, pp. 808–828, 2018.

[4] Antonio Ortega, An Introduction to Graph Signal Processing, Cambridge University Press, 2021.

[5] Libin Liu, Arpan Chattopadhyay, and Urbashi Mitra, “On solving mdps with large state space: Exploitation of policy structures and spectral properties,” IEEE Transactions on Communications, vol. 67, no. 6, pp. 4151–4165, 2019.

[6] Ryan A. Rossi and David F. Gleich, “Dynamic pagerank using evolving teleportation,” in Algorithms and Models for the Web Graph, Anthony Bonato and Jeannette Janssen, Eds. 2012, pp. 126–137. Springer Berlin Heidelberg.

[7] Lawrence Page, Sergey Brin, Rajeev Motwani, and Terry Winograd, “The pagerank citation ranking: Bringing order to the web,” in Stanford InfoLab, 1999.

[8] Richard Bronson, “Chapter 8 - Jordan canonical forms,” in Matrix Methods, pp. 176 – 220. Academic Press, 1969.
[9] Joya A. Deri, “Graph signal processing: Structure and scalability to massive data sets,” 2016.

[10] Joya Deri and Jose Moura, “Spectral projector-based graph fourier transforms,” *IEEE Journal of Selected Topics in Signal Processing*, vol. PP, Jan 2017.

[11] Panagiotis Misiakos, Chris Wendler, and Markus Püschel, “Diagonalizable shift and filters for directed graphs based on the jordan-chevalley decomposition,” Jan 2020.

[12] S. Sardellitti and S. Barbarossa, “On the graph fourier transform for directed graphs,” *IEEE Journal of Selected Topics in Signal Processing*, vol. PP, Jan 2016.

[13] Panagiotis Misiakos, Chris Wendler, and Markus Püschel, “Diagonalizable shift and filters for directed graphs based on the jordan-chevalley decomposition,” Jan 2020.

[14] S. Sardellitti and S. Barbarossa, “On the graph fourier transform for directed graphs,” *IEEE Journal of Selected Topics in Signal Processing*, vol. PP, Jan 2016.

[15] S. Sardellitti and S. Barbarossa, “On the graph fourier transform for directed graphs,” *IEEE Journal of Selected Topics in Signal Processing*, vol. PP, Jan 2016.

[16] Rasoul Shafipour, Ali Khodabakhsh, Gonzalo Mateos, and Evdokia Nikolova, “A directed graph fourier transform with spread frequency components,” *IEEE Transactions on Signal Processing*, vol. 67, no. 4, pp. 946–960, Feb 2019.

[17] Benjamin Girault, Antonio Ortega, and Shrikanth S. Narayanan, “Irregularity-aware graph fourier transforms,” *IEEE Transactions on Signal Processing*, vol. 66, no. 21, pp. 5746–5761, Nov 2018.

[18] Benjamin Girault, Shrikanth S Narayanan, Antonio Ortega, Paulo Gonçalves, and Eric Fleury, “Grasp: A matlab toolbox for graph signal processing,” in *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2017, pp. 6574–6575.

[19] Benjamin Girault, Signal processing on graphs-contributions to an emerging field, Ph.D. thesis, Ecole normale supérieure de Lyon-ENS LYON, 2015.

[20] Julia Barrufet Ribes, “A graph signal processing solution for defective directed graphs,” B.S. thesis, Universitat Politècnica de Catalunya, 2020.

[21] László Erdős, Antti Knowles, Horng-Tzer Yau, and Jun Yin, “Spectral statistics of Erdős–Rényi graphs ii: Eigenvalue spacing and the extreme eigenvalues.” *Communications in Mathematical Physics*, vol. 314, no. 3, pp. 587–640, Aug 2012.

[22] Chris Paige and Charles Van Loan, “A Schur decomposition for hamiltonian matrices,” *Linear Algebra and its Applications*, vol. 41, pp. 11 – 32, 1981.