GRAPH FREQUENCY ANALYSIS OF COVID-19 INCIDENCE TO IDENTIFY COUNTY-LEVEL CONTAGION PATTERNS IN THE UNITED STATES

Yang Li and Gonzalo Mateos

Dept. of Electrical and Computer Engineering, University of Rochester, Rochester, NY, USA

ABSTRACT

The COVID-19 pandemic severely changed the way of life in the United States (US). From early scattered regional outbreaks to current country-wide spread, and from rural areas to highly populated cities, the contagion exhibits diverse patterns at various timescales and locations. We thus conduct a graph frequency analysis to investigate the spread patterns of COVID-19 in different US counties. The commute flows between all 3142 US counties were used to construct a graph capturing the population mobility. The numbers of daily confirmed COVID-19 cases per county were collected and represented as graph signals, which were then mapped into the frequency domain via the graph Fourier transform. The concept of graph frequency in Graph Signal Processing (GSP) enables the decomposition of graph signals (i.e., daily confirmed cases) into modes with smooth or rapid variations with respect to the underlying mobility graph. These different modes of variability are shown to relate to COVID-19 spread patterns within and across counties. Changes in the nature of spread within geographical regions are also revealed by graph frequency analysis at finer temporal scales. Overall, our GSP-based approach leverages case count and mobility data to unveil spatio-temporal contagion patterns of COVID-19 incidence for each US county. Results here support the promising prospect of using GSP tools for epidemiology knowledge discovery on graphs.

Index Terms— Graph signal processing, graph frequency analysis, COVID-19 spread patterns, network data, flow graph.

1. INTRODUCTION

Networks are ubiquitous and their graph representations offer an ideal tool to record and analyze massive amounts of data from almost every aspect of human life [1]: social networks [2, 3], traffic networks [4, 5] and biological networks [6, 7], just to name a few. Network data usually reside on irregular structures, requiring graph algorithms for analysis of emergent complex behavior [8].

Graphs enable modeling complex interactions within data by defining nodes as data entities and edges as relations between nodes. It is often beneficial to also consider nodal attributes that represent certain features of the elements of interest. Such attributes are often conceptualized as signals defined on graphs [1]. Unlike in classical signal processing (SP), the underlying graph topology provides a fair amount of prior information about the said graph signals, while the graph signals themselves can also determine and update pairwise node relationships embedded within graph edges [9, 10]. Accordingly, the field of Graph Signal Processing (GSP) [11, 12] emerged to fruitfully leverage the relational structure encoded in the graph when carrying out information processing tasks. Fundamental concepts in classical SP were generalized to accommodate graph data, notably the graph Fourier transform (GFT) to enable characteristic operations such as filtering and sampling. Noteworthy GSP ad-

vances include inference and generation of graph signals from network structures [13–15], network topology inference from graph signals [10, 16], and integration of both graph signals and topology for knowledge discovery in various timely applications [17–19]. The required GSP background is briefly introduced in Section 2.

As COVID-19 spreads on United States (US) soil and severely impacts a multitude of counties at different scales, there has been a great amount of interest in understanding the spread patterns of the epidemic. Most current work has focused on pathology analysis from biological perspectives [20, 21], studies of contagion within specific locations [22, 23], or case prediction through forecasting [24], just to cite a few noteworthy research directions. In this work, we bring to bear recent GSP advances to investigate the spread pattern of COVID-19 across all US counties, providing a comprehensive spatio-temporal analysis of the contagion via:

- **Spatio-temporal study**. Case data from 3142 US counties was collected, offering a macroscopic nationwide view of the contagion. The number of daily confirmed cases studied ranges from January 22, 2020 to August 31, 2020. Such extended period facilitates analysis of the evolution of contagion patterns across time and offers flexibility for temporal analysis at different time scales.

- **Graph frequency analysis**. The commute flows between all US counties were used to construct a graph capturing the population mobility. A GFT-based analysis was conducted to extract information from the graph spectral domain, leveraging the mobility graph and going beyond traditional vertex or time-domain analyses. Specifically, we established the correspondence between graph frequency components (via low/high-pass graph filtering) and spatial contagion patterns (within/across counties, respectively) in the network. GFT coefficients reveal fundamentally different contagion patterns across locations, and help identify counties at risk of major outbreaks in a data-driven manner.

- **Study of transformation between contagion patterns**. Beyond frequency component binning for each county, we also investigated changes in the nature of contagion patterns within counties. We considered counties near New York city as prototypical examples that morphed from localized outbreaks to cross-county spread. Through the GFT lens, such transformations are revelead in the form of (temporal) transitions between frequency band occupancies.

2. GRAPH-THEORETIC PRELIMINARIES

As the Data Science revolution keeps gaining momentum, it is only natural that complex signals with irregular structure become increasingly of interest. While there are many possible sources and models of added complexity, a general proximity relationship between signal elements is not only a plausible but a ubiquitous model across science and engineering. In this section, we briefly review needed graph-theoretic fundamentals and introduce the concepts of GSP, GFT and filtering operations in the graph frequency domain.

2.1. Graph signal fundamentals

Consider signals whose values are associated with nodes of a weighted, undirected, and connected graph. Formally, we consider the signal $\mathbf{x} = [x_1, \ldots, x_N]^T \in \mathbb{R}^N$ and the weighted graph $\mathcal{G}(\mathcal{V}, \mathcal{E}, \mathbf{W})$, where $\mathcal{V} = \{1, \ldots, N\}$ is a set of N vertices or nodes and $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ is the set of edges. Scalar x_i denotes the signal value at node $i \in \mathcal{V}$. The map $\mathbf{W} : \mathcal{V} \times \mathcal{V} \to \mathbb{R}_+$ from the set of unordered pairs of vertices to the nonnegative reals associates a weight $W_{ij} \geq 0$ with the edge $(i, j) \in \mathcal{E}$, while $W_{ij} \equiv 0$ for $(i, j) \neq \mathcal{E}$. The symmetric coefficients $W_{ij} = W_{ji}$ represent the strength of the connection (i.e., the similarity or influence) between nodes i and j. Henceforth the graph nodes will be $N = 3142$ US counties and edge weights correspond to the average amount of population commuting between counties; see also Section 3.2 for details on the graph construction process. In terms of the signal $\mathbf{x}$ defined by the daily number of confirmed cases across all counties, this means that when the weight W_{ij} is large the signal values x_i and x_j tend to be similar, based on the assumption that the spread of virus highly depends on the population commuting flows. Conversely, when the weight W_{ij} is small, the signal values x_i and x_j are not directly related except for what is implied by their weak connections to other nodes. Such an interpretation of the edge weights establishes a link between the signal values and the graph topology, motivating the fresh GSP approach advocated in this paper.

2.2. Graph Fourier transform and signal smoothness

An instrumental GSP tool is the GFT, which decomposes a graph signal into orthonormal components describing different modes of variation with respect to the graph topology. The GFT allows to equivalently represent a graph signal in two different domains – the vertex domain consisting of the nodes in $\mathcal{V}$, and the graph frequency domain spanned by the spectral basis of $\mathcal{G}$. Therefore, signals can be manipulated in the frequency domain to induce different levels of interactions between neighbors in the network; see Section 2.3 for more on graph filters for frequency decomposition.

To elaborate on this concept, consider the eigenvector decomposition of the combinatorial graph Laplacian $\mathbf{L} := \mathrm{diag}(\mathbf{W1}) - \mathbf{W}$ to define the GFT and the associated notion of graph frequencies. With $\mathbf{\Lambda} := \mathrm{diag}(\lambda_1, \ldots, \lambda_N)$ denoting the diagonal matrix of non-negative Laplacian eigenvalues and $\mathbf{V} := [\mathbf{v}_1, \ldots, \mathbf{v}_N]$ the orthonormal matrix of eigenvectors, one can always decompose the symmetric graph Laplacian as $\mathbf{L} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^T$. The GFT of $\mathbf{x}$ with respect to the Laplacian $\mathbf{L}$ is the signal $\tilde{\mathbf{x}} = [\tilde{x}_1, \ldots, \tilde{x}_N]^T$ defined as $\tilde{\mathbf{x}} = \mathbf{V}^T\mathbf{x}$. The iGFT (inverse GFT) of $\tilde{\mathbf{x}}$ is given by $\mathbf{x} = \mathbf{V}\tilde{\mathbf{x}}$, which is a proper inverse by the orthonormality of $\mathbf{V}$.

The iGFT formula $\mathbf{x} = \mathbf{V}\tilde{\mathbf{x}} = \sum_{k=1}^{N} \tilde{x}_k \mathbf{v}_k$ allows one to synthesize $\mathbf{x}$ as a sum of orthogonal frequency components $\mathbf{v}_k$. The contribution of $\mathbf{v}_k$ to the signal $\mathbf{x}$ is the GFT coefficient $\tilde{x}_k$. The GFT encodes a notion of signal variability over the graph akin to the notion of frequency in Fourier analysis of temporal signals. To understand this analogy, define the total variation of the graph signal $\mathbf{x}$ with respect to the Laplacian $\mathbf{L}$ (also known as Dirichlet energy) as the following quadratic form

$$\mathrm{TV}(\mathbf{x}) := \mathbf{x}^T\mathbf{L}\mathbf{x} = \sum_{i \neq j} W_{ij}(x_i - x_j)^2. \qquad (1)$$

The total variation $\mathrm{TV}(\mathbf{x})$ is a smoothness measure, quantifying how much the signal $\mathbf{x}$ changes with respect to the presumption on variability that is encoded by the weights $\mathbf{W}$ [1, 11].

Back to the GFT, consider the total variation of the eigenvectors $\mathbf{v}_k$, which is given by $\mathrm{TV}(\mathbf{v}_k) = \mathbf{v}_k^T\mathbf{L}\mathbf{v}_k = \lambda_k$. It follows that the eigenvalues $0 = \lambda_1 < \lambda_2 \leq \ldots \leq \lambda_N$ can be viewed as graph frequencies, indicating how the eigenvectors (i.e., frequency components) vary over the graph $\mathcal{G}$. Accordingly, the GFT and iGFT offer a decomposition of the graph signal $\mathbf{x}$ into spectral components that characterize different levels of variability.

2.3. Graph filtering

For graph signal $\mathbf{x}$ with GFT coefficients $\tilde{\mathbf{x}}$, filtering can be done in the frequency domain in a way akin to classical SP of time-varying signals. As discussed in Section 2.2, eigenvalues of the Laplacian correspond to graph frequencies and eigenvectors serve as frequency basis. For instance, a low-pass filter can be designed by isolating the lowest N_L eigenvalues and their corresponding eigenvectors [17]. Define a spectral operation $\tilde{\mathbf{x}}_L = \tilde{\mathbf{H}}_L\tilde{\mathbf{x}}$, where $\tilde{\mathbf{H}}_L = \mathrm{diag}(\tilde{\mathbf{h}}_L)$ and $\tilde{h}_{L,n} = \mathbb{I}\{n < N_L\}$ ($\mathbb{I}\{\cdot\}$ is an indicator function). This is equivalent to the following convolution operation in the vertex domain

$$\mathbf{x}_L = \mathbf{V}\tilde{\mathbf{x}}_L = \mathbf{V}\tilde{\mathbf{H}}_L\tilde{\mathbf{x}} = \mathbf{V}\tilde{\mathbf{H}}_L\mathbf{V}^T\mathbf{x} = \mathbf{H}_L\mathbf{x}, \qquad (2)$$

where $\mathbf{H}_L = \mathbf{V}\tilde{\mathbf{H}}_L\mathbf{V}^T$ is the low-pass graph filter. In addition to $\mathbf{H}_L$, a graph band-pass filter $\mathbf{H}_M$ and high-pass filter $\mathbf{H}_H$ can also be defined analogously. In this way, all graph frequencies are decomposed and assigned to each graph filter where $\mathbf{H}_L$ takes the lowest N_L frequencies, $\mathbf{H}_M$ takes the middle N_M frequencies and $\mathbf{H}_H$ takes the highest N_H frequencies, with $N_L + N_M + N_H = N$. As these filters are mutually exclusive and span all graph frequencies, we can map signals to the spectral domain via the GFT, filter them and use the iGFT to map each frequency component back to the vertex domain. This decomposes the original graph signal into $\mathbf{x} = \mathbf{x}_L + \mathbf{x}_M + \mathbf{x}_H$, which increases the resolution of the signal by partitioning it into components $\mathbf{x}_L, \mathbf{x}_M, \mathbf{x}_H$ that exhibit low, medium and high variability with respect to $\mathcal{G}$. In Section 3.3, we perform this graph frequency decomposition of COVID-19 case data to investigate various contagion patterns across US counties.

3. GSP ANALYSIS OF COVID-19 DATA

In this section, a graph frequency analysis is carried out on COVID-19 prevalence data in the US. First, we define the graph signals as well as the graph constructed for the study. Then a thorough frequency analysis is conducted to identify contagion patterns across US counties.

3.1. COVID-19 data as graph signals

The raw data[1] is the cumulative number of confirmed COVID-19 cases per 100k residents for each of the $N = 3142$ counties in US from Jan 22 to August 31 (223 days in total); see Fig. 3a. From the cumulative data, we compute the number of daily confirmed cases per 100k residents of each county. The daily graph signals can be stacked as columns of the matrix $\mathbf{X} \in \mathbb{R}^{3142 \times 223}$, where row i is a time series of length 223 recording the daily confirmed cases in county i normalized by its population size. Selected time series are depicted in Fig. 1, for a few of them two "waves" are apparent.

From Fig. 3a and Fig. 1 we may get a rough picture of which counties suffer the most from COVID-19 infections. However, both the snapshot in Fig. 3a and the trends in Fig. 1 offer limited amount of information. There are hidden relationships between the signals of each county that can offer further insights via a network-analytic study. As the spread of epidemic diseases highly depends on the population mobility patters, it is natural to take into account the commute flow between counties in building a network graph. In this way,

---

[1]COVID-19 Data Repository by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University, https://systems.jhu.edu/research/public-health/ncov/
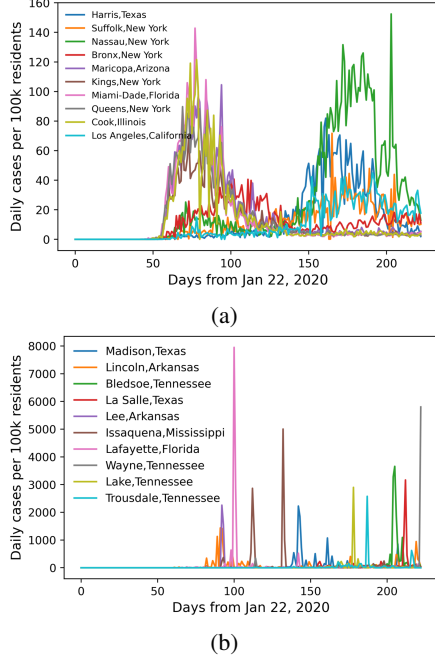
(a)

(b)

Fig. 1: Daily confirmed number of cases for counties ranked in top 10 of (a) average of cumulative number of cases; (b) variance of daily cases. Notice that counties exhibit different contagion patterns in the forms of relatively smooth lines and sudden spikes.

we may find different features of graph signals corresponding to the commute flow graph and evidences of different patterns of the virus contagion in different locations within the country.

3.2. Network graph construction

The US Census Bureau commute flow data from year 2011 to 2015 was used to construct the underlying graph that captures the population mobility between US counties[2]. This is the most recent publicly available dataset at the country level, and while imperfect in many ways it still captures nominal flows trends. A weighted undirected graph $\mathcal{G}$ was constructed with $N = 3142$ counties as nodes and $\binom{3142}{2}$ edges, where edge weights represent the average population mobility flow between two counties after being normalized by the intra-flows (mobility within each county). Fig. 2 shows the topology $\mathbf{W}$ of the resulting graph. As the counties are grouped by the state they belong to, the dense diagonal blocks were observed as expected as large portion of the commute flows happen within states. Meanwhile, we can still observe relatively significant amount of connections between counties in different states, suggesting long distance commutes widely exist, especially between large cities or traffic network hubs. The pairwise similarity patterns conveyed by this flow network are much richer than those in the distance-based graph adopted in our earlier (non-archival) work [25].

3.3. Frequency decomposition of graph signals

With the graph signals $\mathbf{X}$ and graph $\mathcal{G}$, we can follow the procedure in Section 2.2 and Section 2.3 and carry out frequency decomposition of the daily case signals. The Laplacian matrix is formed using $\mathbf{W}$ and its eigenvalues and eigenvectors were computed. A low/high-pass filter was constructed by taking the lowest/highest

---

[2]https://www.census.gov/data/tables/2015/demo/metro-micro/commuting-flows-2015.html
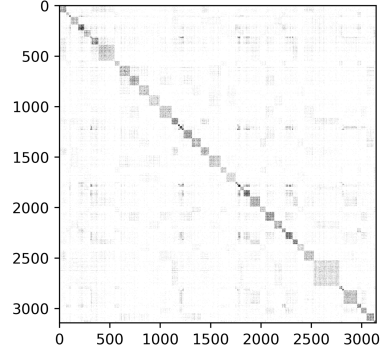

Fig. 2: Adjacency matrix of the commute graph of 3142 US counties. Each entry is the normalized population mobility flow between two counties.

one fifth of the eigenvalues, respectively. After graph filtering as in Section 2.3, the original graph signal $\mathbf{X}$ is now decomposed into $\{\mathbf{X}_L, \mathbf{X}_M, \mathbf{X}_H\} \in \mathbb{R}^{3142 \times 223}$, which represents signal components that change slowly/mildly/rapidly with respect to the underlying commute graph $\mathcal{G}$ [26]. In this work, we focus on low-pass and high-pass signals as they correlate with distinct contagion patterns. Following the same procedure in [17, 26], we take the row-wise average of the absolute values in $\mathbf{X}_L, \mathbf{X}_H$ and obtain two vectors of length 3142 that quantify the signal magnitude per county with respect to their energy occupancy in low-pass and high-pass spectral band. For simplicity, in the following discussion, we use **LP regions** to represent counties with high magnitude of low-pass signals, and **HP regions** for counties with high magnitude of high-pass signals.

We thus decompose the signal of each county into low-pass and high-pass frequency contributions and assess whether its signal concentrates more in either frequency band. Fig. 3 overlays the magnitudes of low-pass signals and high-pass signals of all counties on a US map. As low-pass signals exhibit low variability with respect to the underlying graph, LP regions have similar number of daily confirmed cases normalized by population size with neighboring counties, aligned with the commute flows in between. On the other hand, high-pass signals are related with high variability regardless of the graph structure. Thus HP regions shall present markedly distinct and abnormal case counts relative to nearby counties, despite of the population mobility between them. Analysis in the next section establishes a neat link between graph frequencies (low/high-pass) and contagion spread patterns (i.e. due to cross-county spread or within county outbreak) based on descriptive analysis of the graph signals.

3.4. Frequency analysis w.r.t. contagion patterns

Here, we build a correspondence between low-pass (i.e., smooth, diffusive) signals and cross-county contagion, as one possible consequence of cross-county spread is that nearby counties will tend to exhibit similar number of confirmed cases, leading to smooth signals on $\mathcal{G}$. Meanwhile, high-pass signals shall relate to within-county outbreak, which makes the signal of the current county very dissimilar to nearby counties, resulting in spiky signals. Intuitively, LP regions are representative of cross-county spread and HP regions capture highly localized within-county outbreaks.

3.4.1. Regions in each frequency component

Fig. 3a shows a snapshot of the cumulative confirmed cases for each county. From the original data we can only see which county has the most severe situation. Using the GFT and graph filtering, information is gleaned from the frequency domain. It is clear from Fig. 3b that counties labeled as LP regions form various spatial clusters, and low-pass signals are smoothly spread on the map. This
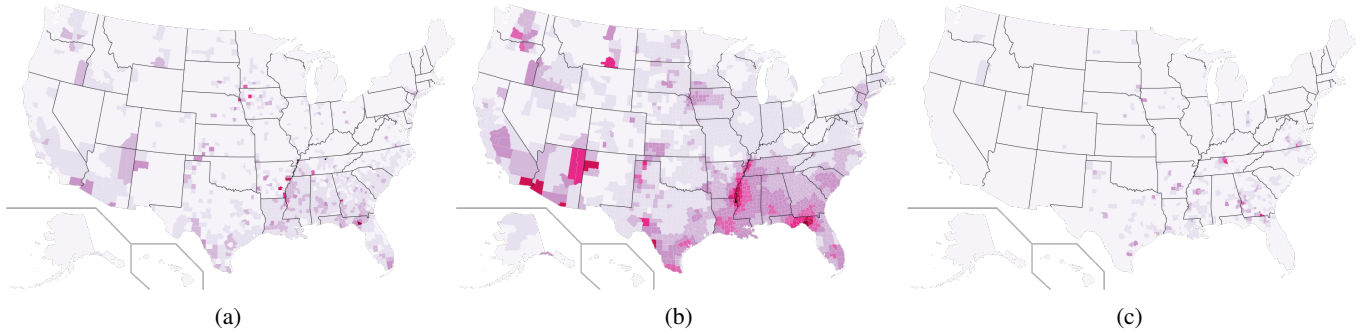
(a)   (b)   (c)

Fig. 3: (a) Cumulative number of confirmed COVID-19 cases per 100k residents for each county by Aug 31; Magnitude of (b) low-pass (b) high-pass signals of each county. Higher frequency components tend to be more localized in the vertex domain, whereas the signal energy distribution in the low-pass signal is more spatially diffused. Due to the highly skewed distribution of case counts, each one of the 3142 values was assigned to one of nine severity levels via range partitioning. Darker color (corresponding to higher severity level) represents higher number of cases.
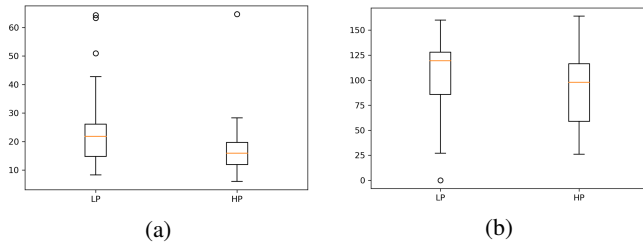

(a)   (b)

Fig. 4: Distribution of (a) average daily cases and (b) days to reach peak for LP and HP regions. T-tests with $p = 0.05$ were carried out in both comparison to prove the statistically difference between LP and HP regions.

suggests that most US counties are going through a wide-ranging cross-county spread. On the other hand, HP regions are localized in several isolated areas in Fig. 3c. Further investigations of local news consistently reveal that all these HP regions have concentrated outbreaks at migrant workers clinics, prisons, food plants, just to list a few [27–29]. This makes their signals very different from neighboring counties, thus being picked up by the high-pass graph filter.

3.4.2. *Average daily cases per frequency band*

One key metric in a pandemic is the average number of daily cases, as it measures the severity of the contagion. For the top-50 ranked LP and HP regions by signal magnitude, their average daily cases were computed and shown in Fig. 4a. As observed, LP regions have significant more daily confirmed cases than HP regions.

3.4.3. *Number of days to reach peaks*

Another key indicator is the number of days from the date of the first case to the date with the most daily cases. This metric gives insights on how fast the spread is. For the same top-50 LP and HP regions, the number of days to reach peak was computed as depicted in Fig. 4b. Apparently, HP regions take fewer days to reach peaks, suggesting relatively faster spread of the virus.

The findings of this section linked the graph frequency components with descriptive findings from raw case data such as average daily cases and number of days to reach peaks. The fact that LP regions have typically more daily cases and take longer to reach peaks (compounded with the fact these regions are more densely gathered as shown in Fig. 3b), leads to the conclusion that LP regions suffer from cross-county spread with a continuous increasing rate and a longer contagion period. Meanwhile, HP regions are spread-out in isolated areas. These regions suffer from local outbreaks in highly-concentrated places, which explains why they have very fast case increase rates. They concentrate in the high-pass band as these outbreaks make their signals very dissimilar to nearby regions.
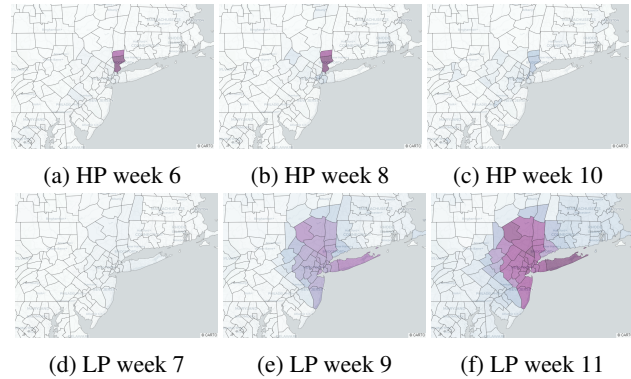

(a) HP week 6 (b) HP week 8 (c) HP week 10

(d) LP week 7 (e) LP week 9 (f) LP week 11

Fig. 5: Magnitudes of high-pass and low-pass signals in different weeks. Intermediate weeks of the same trends were not shown due to limited space.

3.5. Temporal analysis of case counts

In the previous static analysis, a scalar value was assigned to each county in each frequency component by taking a temporal (column-wise) average of $\mathbf{X}_L, \mathbf{X}_H$. This approach was usefuk to link low/high graph frequencies with contagion pattern of cross/within-county spread. However, it ignores rich temporal information by averaging the filtered signals. Instead of taking the average, here we partitioned all 223 days into 7-day windows representing each week from Jan. 22. We then took the column-wise average of each weekly period. For concreteness, we focus on the New York city area as it is represents a nice example of how a localized outbreak transformed into widely regional spread. As shown in Fig. 5, a (hihg-pass) localized outbreak started in the Westchester county at week 6. After that, magnitudes of high-pass signals decreased while low-pass signals started gaining dominance in the region. From the relation between frequency bands and contagion patterns argued in previous sections, we can see the transformation of a localized outbreak to a wider spread across NYC regions, and how it is manifests via transitions on frequency-band occupancies.

4. CONCLUSIONS AND FUTURE WORK

In this work, we applied GSP tools for the analysis of COVID-19 contagion patterns. Novel information extracted from graph frequency domain helps establish the link between graph low/high frequency and the across/within-county contagion spread for US counties. We show that GSP can provide novel insights for mining and learning vital information from complex data residing on irregular domains. As more (currently proprietary) data are released, future work consider dynamic mobility graphs with an emphasis on joint time-vertex analyses at finer scales.

5. REFERENCES

[1] A. Ortega, P. Frossard, J. Kovačević, J. M. Moura, and P. Vandergheynst, "Graph signal processing: Overview, challenges, and applications," *Proc. IEEE*, vol. 106, no. 5, pp. 808–828, 2018.

[2] B. Perozzi, R. Al-Rfou, and S. Skiena, "Deepwalk: Online learning of social representations," in *ACM SIGKDD Intl. Conf. on Knowledge Discovery and Data Mining (KDD)*, 2014, pp. 701–710.

[3] J. Zhang and J. M. Moura, "Diffusion in social networks as SIS epidemics: Beyond full mixing and complete graphs," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 4, pp. 537–551, 2014.

[4] X. Dong, A. Ortega, P. Frossard, and P. Vandergheynst, "Inference of mobility patterns via spectral graph wavelets," in *IEEE Intl. Conf. Acoust., Speech and Signal Process. (ICASSP)*. IEEE, 2013, pp. 3118–3122.

[5] M. Crovella and E. Kolaczyk, "Graph wavelets for spatial traffic analysis," in *INFOCOM*, vol. 3. IEEE, 2003, pp. 1848–1857.

[6] Y. Li, R. Shafipour, G. Mateos, and Z. Zhang, "Supervised graph representation learning for modeling the relationship between structural and functional brain connectivity," in *IEEE Intl. Conf. Acoust., Speech and Signal Process. (ICASSP)*. IEEE, 2020, pp. 9065–9069.

[7] L. Wang, F. V. Lin, M. Cole, and Z. Zhang, "Learning clique subgraphs in structural brain network classification with application to crystallized cognition," *BioRxiv*, 2020.

[8] M. Newman, *Networks*. Oxford university press, 2018.

[9] S. Segarra, A. G. Marques, G. Mateos, and A. Ribeiro, "Network topology inference from spectral templates," *IEEE Trans. Signal Inf. Process. Netw.*, vol. 3, no. 3, pp. 467–483, 2017.

[10] G. Mateos, S. Segarra, A. G. Marques, and A. Ribeiro, "Connecting the dots: Identifying network structure via graph signal processing," *IEEE Signal Process. Mag.*, vol. 36, no. 3, pp. 16–43, 2019.

[11] 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 Process. Mag.*, vol. 30, no. 3, pp. 83–98, 2013.

[12] A. Sandryhaila and J. M. Moura, "Discrete signal processing on graphs: Frequency analysis," *IEEE Trans. Signal Process.*, vol. 62, no. 12, pp. 3042–3054, 2014.

[13] C. J. Honey, O. Sporns, L. Cammoun, X. Gigandet, J.-P. Thiran, R. Meuli, and P. Hagmann, "Predicting human resting-state functional connectivity from structural connectivity," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 106, no. 6, pp. 2035–2040, 2009.

[14] Y. Li, R. Shafipour, G. Mateos, and Z. Zhang, "Mapping brain structural connectivities to functional networks via graph encoder-decoder with interpretable latent embeddings," in *IEEE Global Conf. Signal and Info. Process. (GlobalSIP)*. IEEE, 2019, pp. 1–5.

[15] S. Segarra, G. Mateos, A. G. Marques, and A. Ribeiro, "Blind identification of graph filters with sparse inputs," in *IEEE Intl. Wrksp. Computat. Advances Multi-Sensor Adaptive Process. (CAMSAP)*, 2015, pp. 449–452.

[16] X. Dong, D. Thanou, P. Frossard, and P. Vandergheynst, "Learning laplacian matrix in smooth graph signal representations," *IEEE Trans. Signal Process.*, vol. 64, no. 23, pp. 6160–6173, 2016.

[17] L. Goldsberry, W. Huang, N. F. Wymbs, S. T. Grafton, D. S. Bassett, and A. Ribeiro, "Brain signal analytics from graph signal processing perspective," in *IEEE Intl. Conf. Acoust., Speech and Signal Process. (ICASSP)*. IEEE, 2017, pp. 851–855.

[18] J. Wang, V. D. Calhoun, J. M. Stephen, T. W. Wilson, and Y. Wang, "Integration of network topological features and graph Fourier transform for fMRI data analysis," in *Intl. Symp. Biomed. Imaging*. IEEE, 2018, pp. 92–96.

[19] J. Richiardi, S. Achard, H. Bunke, and D. Van De Ville, "Machine learning with brain graphs: predictive modeling approaches for functional imaging in systems neuroscience," *IEEE Signal Process. Mag.*, vol. 30, no. 3, pp. 58–70, 2013.

[20] M. R. Mehra, S. S. Desai, S. Kuy, T. D. Henry, and A. N. Patel, "Cardiovascular disease, drug therapy, and mortality in COVID-19," *N. Engl. J. Med.*, 2020.

[21] Y. Zhang, M. Xiao, S. Zhang, P. Xia, W. Cao, W. Jiang *et al.*, "Coagulopathy and antiphospholipid antibodies in patients with COVID-19," *N. Engl. J. Med.*, vol. 382, no. 17, p. e38, 2020.

[22] A. Remuzzi and G. Remuzzi, "COVID-19 and Italy: what next?" *The Lancet*, 2020.

[23] W. Liu, Q. Zhang, J. Chen, R. Xiang, H. Song, S. Shu, L. Chen, L. Liang, J. Zhou, L. You *et al.*, "Detection of COVID-19 in children in early January 2020 in Wuhan, China," *N. Engl. J. Med.*, vol. 382, no. 14, pp. 1370–1371, 2020.

[24] A. Kapoor, X. Ben, L. Liu, B. Perozzi, M. Barnes, M. Blais, and S. O'Banion, "Examining COVID-19 forecasting using spatio-temporal graph neural networks," *arXiv preprint arXiv:2007.03113*, 2020.

[25] Y. Li and G. Mateos, "Graph frequency analysis of COVID-19 prevalence in the United States," in *KDD Workshop on Mining and Learning with Graphs*. San Diego, CA, Aug. 24, 2020.

[26] W. Huang, T. A. Bolton, J. D. Medaglia, D. S. Bassett, A. Ribeiro, and D. Van De Ville, "A graph signal processing perspective on functional brain imaging," *Proc. IEEE*, vol. 106, no. 5, pp. 868–885, 2018.

[27] C. Herbert, "Echols adds 81 cases in 10 days," 2020. [Online]. Available: https://www.valdostadailytimes.com/news/echols-adds-81-cases-in-10-days/article_401df1be-a01f-11ea-b340-b3c9f358f3e1.html

[28] T. Loller, "Tennessee county leads US in coronavirus cases per capita," 2020. [Online]. Available: https://www.timesfreepress.com/news/breakingnews/story/2020/may/12/tennessee-county-leads-us-coronavirus-cases-capita/522779/

[29] W. Joy, "Titus County: A look at one of the worst COVID-19 hotspots in Texas," 2020. [Online]. Available: https://www.wfaa.com/article/news/health/coronavirus/titus-county-a-look-at-one-of-texass-worst-covid-19-hotspots/287-2ac2d1ce-98ec-460d-bb6b-2df5b14e8827