Processing Asynchronous Time-simplex Signals

Feng Ji[1], Xingchao Jian[1], Wee Peng Tay[1], Maosheng Yang[2]
[1]Nanyang Technological University, Singapore, [2]TU Delft, Netherlands
E-mails: jifeng@ntu.edu.sg, xingchao001@e.ntu.edu.sg, wptay@ntu.edu.sg, m.yang-2@tudelft.nl.

***Abstract*—Topological signal processing (TSP) over simplicial complexes typically assumes observations associated with the simplicial complexes are real scalars. In this paper, we develop a TSP framework for the case where observations on a simplicial complex belong to function spaces that are commonly used to represent time-varying signals. Our approach generalizes the Hodge decomposition and allows for signal processing tasks to be performed on these more complex observations, which can better model signals in practice. We propose a unified and flexible framework for TSP over simplicial complexes that expands its applicability to a wider range of signal processing applications. Numerical results demonstrate the effectiveness of this approach and provide a foundation for future research in this area.**

***Index Terms*—Topological signal processing, generalized signals, simplicial complex**

I. Introduction

In recent years, there has been growing interest in analyzing signals supported on different domains. For signals supported on vertices of graphs, graph signal processing (GSP) has emerged as a powerful tool [1]–[3]. For signals supported on edges of simplicial complexes such as in telecommunication traffic flows, topological signal processing (TSP) is developed [4]–[8]. The basic idea behind TSP is based on the Hodge decomposition of signals, which decomposes a signal into three orthogonal components: the irrotational, solenoidal, and harmonic components. Like GSP, there are also ideas based on the eigenspaces of the Laplacian operator and enables the consideration of concepts such as frequency, Fourier transform, wavelets, and convolution [9], [10].

These works have largely focused on the case when the signal (on each simplex) is a scalar in $\mathbb{R}$ (or $\mathbb{C}$). To deal with discrete time-varying signals, [11] has proposed an approach that embeds the signals in a product space, similar to the time-vertex framework in GSP [12]–[14]. By leveraging smoothness over both spatial and temporal domains, this method is able to effectively reconstruct time-varying flows. Using a similar strategy, [15] proposes an online imputation method for edge flows.

In practice, the signal space can be more complex. An example is given by $L^2(\Omega)$, where Ω is a continuous time interval, which allows the modeling of continuous-time signals not captured by the frameworks of [11], [15]. For a simplicial complex modeling a network like a sensor network (cf. [16, Section V.]), observations or samples collected at each simplex may not be ***time synchronous***. More specifically, signals observed at simplexes, e.g., nodes, edges, or faces, may not be sampled on the same time stamps, and signals may not be observed at uniform time intervals.

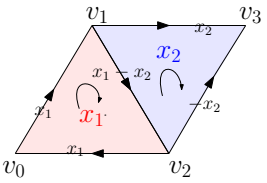
If our signal model uses only real-valued signals, then we need to consider multiple discrete timestamps and process observed signals on the entire network either separately for each timestamp [4] or jointly [11]. As observations are asynchronous, at each discrete timestamp, the available information over the whole network can be incomplete. Moreover, asynchronism may render uniform discretization difficult. On the other hand, modeling signals as functions from a space $L^2(\Omega)$ alleviates such problems by aggregating several asynchronous observations into a single signal model that captures spatial and temporal correlations for different simplexes (cf. Section IV). A similar approach is taken in generalized GSP (GGSP) [3], [16], [17]. To describe such signals, we need an algebraic entity different from $\mathbb{R}$.

The well-established setup of algebraic topology [18] provides the most general framework to describe non-scalar signals. More specifically, we may use elements in abelian groups, modules, or even local coefficient systems (sheaves) [18] to model such signals on simplicial complexes. Formally, they form chains or cochains in algebraic topology. In our paper, though in the same spirit, we shall be more concrete by focusing on normed function spaces as signal spaces, which are abelian groups. They are sufficient for us to model asynchronous time series of signals, both discrete and continuous, on simplicial complexes.

In this short paper, we formally describe the proposed framework with generalized signals on simplicial complex, to model time-simplex signals. In general, the Hodge decomposition does not always exist for such a signal (e.g., [19, Example 2]). The precise reason is beyond the scope of the paper. We highlight that essentially, a generalized signal does not necessarily belong to a *field* such as $\mathbb{R}$, in which each non-zero element has a multiplicative inverse (see [20]). To overcome the above challenge, we propose an optimization model as a substitute, which is shown to generalize the Hodge decomposition. We demonstrate how the framework can be used effectively in practice with real datasets. This work can be viewed as a simultaneous generalization of large parts of both [4] and [16]. Our work [19] is an extended version of the paper, in which a general formal framework is proposed.

II. Simplicial complexes and time-simplex signals

We first define simplicial complexes [4], [18], which are high-dimensional generalizations of graphs.



2-simplexes:
$[v_0, v_1, v_2], [v_1, v_3, v_2]$

1-simplexes (edges):
$[v_0, v_1], [v_1, v_2], [v_2, v_0]$
$[v_2, v_3], [v_1, v_3]$

Fig. 1. This exemplifies a 2-simplicial complex X with two 2-simplexes (the colored triangles) and five 1-simplexes (the edges). The oriented simplexes are listed by the side. For example, the clockwise orientation of the triangle $[v_0, v_1, v_2]$ has oriented boundary edges $[v_0, v_1], [v_1, v_2], [v_2, v_0]$ as shown. A signal in $\mathcal{A}$ assigns an element in $\mathcal{A}$ to a simplex of appropriate dimension in X. In the example, (x_1, x_2) is a 2-signal assigned to the 2-simplexes. It generates the 1-signal $(x_1, x_1 - x_2, x_1, -x_2, x_2)$ on the set of edges, via the boundary map ∂_2 (or $\mathbf{B}_2$ in the matrix form).

Definition 1. *The* standard n-simplex (or dimension n simplex) *is defined as the set*

$$\Delta_n = \big\{(r_0, \dots, r_n) \in [0,1]^{n+1} \mid \sum_{i=0}^{n} r_i = 1\big\}.$$

Each point $v_i \in [0,1]^{n+1}$ with its i-th component 1 is a vertex of Δ_n. A topological space homeomorphic to the standard n-simplex is called an n-simplex. In Δ_n, if we require $k > 0$ coordinates being 0, we get an $(n-k)$-simplex, called a face.

A simplicial complex X *is a set of simplexes such that*

- *Any face from a simplex of X is also in X.*
- *The intersection of any two simplexes $\sigma_1, \sigma_2 \in X$ is a face of both σ_1 and σ_2.*

Intuitively, a simplicial complex is obtained by gluing "triangles of different dimensions" along their boundaries.

The *geometric realization* of a simplicial complex X is obtained from gluing all its simplexes along common faces. We usually do not distinguish X from its geometric realization. For a k-simplex σ^k with vertices $\{v_0, \dots, v_k\}$, an *orientation* is an ordering of its vertices, denoted by $[v_{i_0}, \dots, v_{i_k}]$ (see Fig. 1). Two orientations are opposite if they differ by an odd number of permutations. Otherwise, they are the same.

Now we consider signals. Let $\mathcal{A}$ be a *function space* on a domain Ω that is preserved under (function) addition and subtraction, i.e., $x_1 + x_2, x_1 - x_2 \in \mathcal{A}$ provided $x_1, x_2 \in \mathcal{A}$. For **example**, $\mathcal{A}$ can be the *space of polynomials* up to a fixed degree, a *subspace of* $L^2(\Omega)$, or *splines*. We use each element x of $\mathcal{A}$ to model a time series. In many applications, it suffices to observe the value of x at a discrete or even a finite subset of Ω to fully determine x.

Definition 2. *Suppose a finite simplicial complex X has n_k oriented k-simplexes $\{\sigma_1^k, \dots, \sigma_{n_k}^k\}$ for $k = 0, \dots, K$. A k-signal $\mathbf{x} = (x_i)_{i \le n_k}$ in $\mathcal{A}$ assigns $x_i \in \mathcal{A}$ to the i-th simplex σ_i^k. The set of k-signals in $\mathcal{A}$ is denoted by $C_k(X, \mathcal{A})$.*

Intuitively, analogous to GSP, a k-signal is a collection of elements of $\mathcal{A}$, each assigned to a k-simplex of X.

In the context of our paper, a k-signal $\mathbf{x} \in C_k(X, \mathcal{A})$ associates a "time series" to each k-simplex of X. Another key player is the *boundary operator* $\partial_k : C_k(X, \mathcal{A}) \to C_{k-1}(X, \mathcal{A})$, that produces a $k-1$-signal given a k-signal.

Definition 3. *For each signal $\mathbf{x} = (x_i)_{i \le n_k}$ in $C_k(X, \mathcal{A})$, the $k-1$-signal $\partial_k(\mathbf{x})$ assigns $(-1)^j x_i$ to the oriented simplex $[v_{i_0}, \dots, v_{i_{j-1}}, v_{i_{j+1}}, \dots, v_{i_k}]$ for each $0 \le j \le k$, where $[v_{i_0}, \dots, v_{i_k}]$ is the orientation of σ_i^k (see Fig. 1).*

Intuitively, any boundary of a k-simplex σ_i^k takes the same signal value (in $\mathcal{A}$) as σ_i^k itself, up to a sign determined by the index of the missing vertex.

Equivalently, ∂_k has an $n_{k-1} \times n_k$ matrix form $\mathbf{B}_k$, whose (j,i)-th entry is 1 (resp. -1) if the σ_j^{k-1} is a face of σ_i with the same (resp. opposite) orientation and 0 otherwise. The k-th *Laplacian* $\mathbf{L}_k$ is

$$\mathbf{L}_k = \mathbf{B}_k^\intercal \mathbf{B}_k + \mathbf{B}_{k+1} \mathbf{B}_{k+1}^\intercal, \tag{1}$$

where by convention $\mathbf{B}_0 = \mathbf{B}_{K+1} = 0$. If $\mathcal{A} = \mathbb{R}$ is the space of constant functions, then the Hodge decomposition [4] claims that each k-signal $\mathbf{x}$ has a unique decomposition $\mathbf{x} = \mathbf{x}_{-1} + \mathbf{x}_0 + \mathbf{x}_1$, where $\mathbf{x}_{-1} \in \operatorname{im} \mathbf{B}_k^\intercal, \mathbf{x}_0 \in \ker \mathbf{L}_k$ and $\mathbf{x}_1 \in \operatorname{im} \mathbf{B}_{k+1}$. The decomposition does not hold for a general $\mathcal{A}$, and we provide a generalization in the next section.

III. The Fundamental Learning Model

To propose the learning model, we assume that $\mathcal{A}$ has a norm $\|\cdot\|_{\mathcal{A}}$. For **example**, the L^2-norm is a common choice if $\mathcal{A}$ is a subspace of $L^2(\Omega)$. This metrical setup can be extended to $C_k(X, \mathcal{A})$ as follows. Recall that X has n_k oriented k-simplexes $\{\sigma_1^k, \dots, \sigma_{n_k}^k\}$.

Definition 4. *For $\mathbf{x} = (x_i)_{i \le n_k} \in C_k(X, \mathcal{A})$, its* p-norm *is*

$$\|\mathbf{x}\|_{k,p}^p = \sum_{1 \le i \le n_k} \|x_i\|_{\mathcal{A}}^p.$$

As we have pointed out, given a signal $\mathbf{x} \in C_k(X, \mathcal{A})$ for a general $\mathcal{A}$, we usually do not have the Hodge decomposition as we have briefly explained in Section I. For this purpose, we propose the *fundamental learning model* as the following optimization:

$$\begin{aligned} \min_{(\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_{-1})} & \|\mathbf{x}_0\|_{k,p} + \zeta_1 \|\mathbf{x}' - \mathbf{x}\|_{k,p} + \zeta_2 R(\cdot) \\ \text{s.t. } & \mathbf{x}' = \mathbf{x}_0 + \mathbf{x}_1 + \mathbf{x}_{-1} \\ & \mathbf{x}_0 \in \ker \partial_k, \\ & \mathbf{x}_1 = \partial_{k+1}(\mathbf{y}_1),\ \mathbf{y}_1 \in C_{k+1}(X, \mathcal{A}), \\ & \mathbf{x}_{-1} = \partial_k^*(\mathbf{y}_{-1}),\ \mathbf{y}_{-1} \in C_{k-1}(X, \mathcal{A}), \end{aligned} \tag{2}$$

where the boundary operators are as defined in Definition 3.

The term $\|\mathbf{x}_0\|_{k,p}$ corresponds to the Hodge decomposition (cf. Theorem 1 below) if $\mathcal{A} = \mathbb{R}$ and $p = 2$. The term $\|\mathbf{x}' - \mathbf{x}\|_{k,p}$ is that for data fidelity. Lastly, $R(\cdot)$ is a regularizer that is problem-dependent. It usually reflects our prior knowledge of the properties of the signal such as smoothness as well as task-specific signal models. The coefficients ζ_1, ζ_2 are tunable hyperparameters. Apparently, the optimization in (2) can be applied for a general $\mathcal{A}$.

Tips to apply in practice: To solve (2), it will be convenient if we can parametrize $C_{k-1}(X, \mathcal{A})$, $C_{k+1}(X, \mathcal{A})$ and $\ker \partial_k$. For $C_{k-1}(X, A)$, an element $\mathbf{y}$ takes the form $(y_i)_{i \le n_{k-1}}$, with

$y_i \in \mathcal{A}$. Therefore, if we can parameterize $\mathcal{A}$, then we can also do so for $\mathbf{y}$. As $\ker \partial_k \subset C_k(X, \mathcal{A})$, the same approach can be applied to $C_{k+1}(X, \mathcal{A})$ and $\ker \partial_k$. In Section IV, we shall see a variant of this general idea that a parameterization is based on the eigendecomposition of the Laplacian $\mathbf{L}_k$.

We end this section by describing the relation between (2) and the classical Hodge decomposition.

Theorem 1. *For $A = \mathbb{R}$ and $p = 2$, if the parameter $\zeta_1 > 1, \zeta_2 = 0$ in (2), then the Hodge decomposition of $\mathbf{x}' = \mathbf{x}$ is the unique solution to the fundamental learning model (2).*

Proof: For $\mathbf{x} \in C_k(X, \mathbb{R})$, assume that its Hodge decomposition is $\mathbf{x} = \tilde{\mathbf{x}}_0 + \tilde{\mathbf{x}}_1 + \tilde{\mathbf{x}}_{-1}, \tilde{\mathbf{x}}_0 \in \ker \mathbf{L}_k \subset \ker \partial_k$, where $\tilde{\mathbf{x}}_1 \in \operatorname{im} \partial_{k+1}$, and $\tilde{\mathbf{x}}_{-1} \in \operatorname{im} \partial_k^*$.

Observe that to obtain an optimal solution $\mathbf{x}' = \mathbf{x}_0 + \mathbf{x}_1 + \mathbf{x}_{-1}$ to (2), we can always require $\mathbf{x}_0 \perp \mathbf{x}_1$ as the projection of $\mathbf{x}_0 + \mathbf{x}_1$ to $(\operatorname{im} \partial_{k+1})^\perp$ in $\ker \partial_k$, to ensure $\|\mathbf{x}_0\|_{k,p}$ is minimal.

We first consider the case $\mathbf{x}' = \mathbf{x}$ and (2) requires us to minimize $\|\mathbf{x}_0\|_2$. The space $C_k(X, \mathbb{R})$ has orthogonal decomposition $\operatorname{im} \partial_k^* \oplus \ker \partial_k$. Therefore, to minimize $\|\mathbf{x}_0\|_{k,2}$, we first need to project $\mathbf{x}$ to $\ker \partial_k$. Moreover, $\ker \partial_k = \operatorname{im} \partial_{k+1} \oplus \ker \mathbf{L}_k$. Since $\mathbf{x}_0 + \mathbf{x}_1 = \tilde{\mathbf{x}}_0 + \tilde{\mathbf{x}}_1$ and $\mathbf{x}_1$ is constrained to be in $\operatorname{im} \partial_{k+1}$, to minimize $\|\mathbf{x}_0\|_{k,2}$, the summand we need to remove from $\mathbf{x}$ is its projection to $\operatorname{im} \partial_{k+1}$. Therefore, the Hodge decomposition is the unique solution to (2) provided $\mathbf{x}' = \mathbf{x}$.

To conclude, it suffices to show that if $\mathbf{x}' \neq \mathbf{x}$, then

$$\|\tilde{\mathbf{x}}_0\|_{k,2} < \|\mathbf{x}_0\|_{k,2} + \zeta_1 \|\mathbf{x}' - \mathbf{x}\|_{k,2}.$$

By the orthogonality of the different components of the Hodge decomposition, we have

$$\begin{aligned}\|\mathbf{x}' - \mathbf{x}\|_{k,2}^2 &= \|\mathbf{x}_1 - \tilde{\mathbf{x}}_1\|_{k,2}^2 \\ &\quad + \|\mathbf{x}_{-1} - \tilde{\mathbf{x}}_{-1}\|_{k,2}^2 + \|\mathbf{x}_0 - \tilde{\mathbf{x}}_0\|_{k,2}^2.\end{aligned}$$

Therefore, $\|\mathbf{x}' - \mathbf{x}\|_{k,2} \geq \|\mathbf{x}_0 - \tilde{\mathbf{x}}_0\|_{k,2}$. By the triangle inequality, we obtain

$$\begin{aligned}\|\tilde{\mathbf{x}}_0\|_{k,2} &\leq \|\mathbf{x}_0\|_{k,2} + \|\mathbf{x}_0 - \tilde{\mathbf{x}}_0\|_{k,2} \leq \|\mathbf{x}_0\|_{k,2} + \|\mathbf{x}' - \mathbf{x}\|_{k,2} \\ &< \|\mathbf{x}_0\|_{k,2} + \zeta_1 \|\mathbf{x}' - \mathbf{x}\|_{k,2},\end{aligned}$$

where the last inequality follows from $\|\mathbf{x}' - \mathbf{x}\|_{k,2} > 0$ and $\zeta_1 > 1$. The result now follows. ∎

IV. Experimental results

1) Heat flow: study of bandlimited signals: In this study, we consider the heat flow dataset of the Intel Berkeley Research lab. Temperature data are collected from 53 sensors, denoted by V, placed in the lab. They form a planar graph G with an edge set E of size 87. Each edge is given an orientation such that heat flows from high temperature to low temperature. Each (directed) edge is associated with its heat flow estimated from sensor readings at every timestamp. In this way, we obtain the ground truth signals $\{\mathbf{x}_t, 1 \leq t \leq T = 864\}$, where $\mathbf{x}_t = (x_{e,t})_{e \in E}$ is the (heat flow) edge signals on G at time t. For the task, we assume that (a randomly chosen) $r \in [0.85, 0.95]$ fraction of data is missing and want to estimate all signals from the known $1 - r$ fraction.

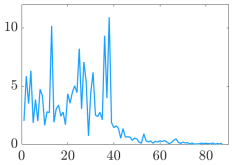


Fig. 2. The spectral plot of a typical $\mathbf{x}_t$.

We observe that a typical $\mathbf{x}_t$ is bandlimited with concentrated spectrum with respect to (w.r.t.) $\mathbf{L}_1$ (in (1)) of G (Fig. 2). Intuitively, this indicates that the heat flow follows the conservation law. Motivated by this, we propose the following "*joint*" recovery scheme, in the spirit of [16]. Specifically, let $\mathcal{A} \subset L^2([0, T/2 - 1])$ be the function space generated by short-time windowed functions [21]:

$$\begin{aligned}\psi_{\omega,t_0}(t) &= \exp(-\frac{1}{2\sigma^2}(t - t_0)^2) \cdot \sin(\omega t), \\ \phi_{\omega,t_0}(t) &= \exp(-\frac{1}{2\sigma^2}(t - t_0)^2) \cdot \cos(\omega t),\end{aligned} \tag{3}$$

where parameters $t_0 \in T_0 = \{0, 50, \ldots, 400\}$ and $\omega \in \Omega = \{0, 1/45, 2/45 \ldots, 1/5\}$. Here, σ is a tunable hyperparameter. We use these functions to model temporal patterns of the data. Therefore, a signal at an edge e is a linear combination $f_e = \sum_{t_0 \in T_0, \omega \in \Omega} \big(a_{e,\omega,t_0} \psi_{\omega,t_0} + b_{e,\omega,t_0} \phi_{\omega,t_0}\big)$. Write $\mathbf{f} = (f_e)_{e \in E}$ for a typical set of edge signals. To incorporate the bandlimited spatial information, let $\mathbf{F}$ be the low-pass filter that projects to the space spanned by eigenvectors (of $\mathbf{L}_1$) associated with the smallest 40 eigenvalues (cf. Fig. 2). Let $\mathbf{u}_i = (u_{i,e})_{e \in E}$ be the i-th eigenvector. Write $\mathbf{a}_{\omega,t_0} = (a_{e,\omega,t_0})_{e \in E}$, $\mathbf{b}_{\omega,t_0} = (b_{e,\omega,t_0})_{e \in E}$ and denote $c_{i,\omega,t_0} = \langle \mathbf{u}_i, \mathbf{a}_{\omega,t_0} \rangle$, $d_{i,\omega,t_0} = \langle \mathbf{u}_i, \mathbf{b}_{\omega,t_0} \rangle$. Similar to GSP, if $\mathbf{g} = \mathbf{F}(\mathbf{f}) = (g_e)_{e \in E}$, then g_e can be parameterized by coefficients $\{c_{i,\omega,t_0}, d_{i,\omega,t_0} \mid t_0 \in T_0, \omega \in \Omega, i \leq 40\}$ as:

$$g_e = \sum_{t_0 \in T_0, \omega \in \Omega, i \leq 40} u_{i,e} \big(c_{i,\omega,t_0} \psi_{\omega,t_0} + d_{i,\omega,t_0} \phi_{\omega,t_0}\big).$$

To solve (2), we use the fidelity term $\sum_{(e,t):\text{observed}} \|g_e(t) - x_{e,t}\|_2^2$ and regularizer $\sum_{t_0,\omega,i} \big(|c_{i,\omega,t_0}| + |d_{i,\omega,t_0}|\big)$.

The performance is evaluated using the *relative mean squared error (RMSE)* between the estimate $\hat{x}$ and ground truth x: $\text{RMSE} = \big(\sum_{(e,t):\text{test}} (\hat{x}_{e,t} - x_{e,t})^2\big) / \big(\sum_{(e,t):\text{test}} x_{e,t}^2\big)$. For comparison, we consider benchmarks in [21] and [22, (2)], called "*separate*" and "*product*" respectively. Briefly, the separate method performs an interpolation to recover the time series of each edge based on known signals using the window functions (3) as basis. However, spatial signal correlations are not used. On the other hand, the product method constructs a new simplicial complex $X = G \times P_T$, where P_T is the path graph of size T. One can then apply standard TSP techniques [4] for recovery on X with the usual signal space $\mathbb{R}$.

From Fig. 3, we see that our joint method performs significantly better than the benchmarks, particularly for large r. Moreover, RMSEs for the joint method have much smaller standard deviations. This suggests that the joint method has fully used both spatial and temporal information. On the other hand, the separate method outperforms the product method for smaller r. However, its performance deteriorates if r increases, as it does not use spatial information.

TABLE I
REAL EVENTS ASSOCIATED WITH THE CURVE PATTERNS (CHANGE OF TREND OR PROMINENT FLUCTUATIONS) IN FIG. 4

(A)	Index ≈ 600	The Fed reduced the interest rate by 50bps on 3 Mar. 2020 and by 100bps on 16 Mar. 2020 to combat COVID-19.
(B)	1340 - 1380	The Fed raised the interest rate by 25bps on 17 Mar. 2022 and 50bps on 5 May 2022 to combat inflation
(C)	1560 - 1600	The Fed raised the interest rate by 75bps and 50bps on 2 Nov. and 14 Dec. 2022, marking the end of its hawkish policies.
(D)	≈ 1305	Frequent fluctuations for EUR. It may correspond to the Russian invasion of Ukraine on 24 Feb. 2022.
(E)	≈ 540	Sudden drop in value for CNY. This may correspond to the outbreak of COVID-19 at the end of 2019.
(F)	1500 - 1550	The short premiership of Liz Truss (6 Sep. - 25 Oct. 2022). GBP hit a 37-year low on 23 Sep. [23].
(G)	105 - 450	GBP curve is unstable. It may correspond to the post-Brexit UK-EU negotiation period of the withdrawal agreement
(H)	$--$	AED, HKD, USD curves are almost identical in shape. HKD and AED have been pegged to USD since Oct. 1983, Nov. 1997.
(I)	$--$	JPY is almost always increasing. It may be due to the Bank of Japan maintaining a negative interest rate of -0.1% since 2016.

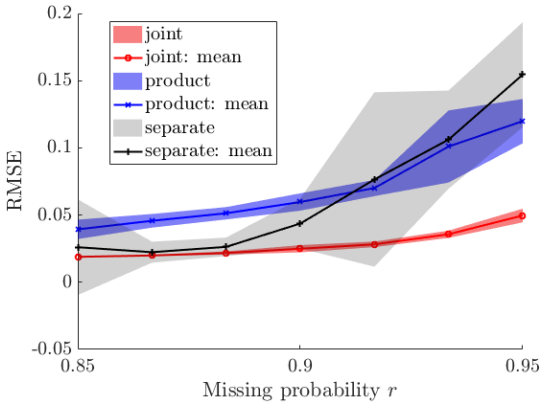


Fig. 3. Performance comparison among the joint, separate, and product methods. The shaded regions are derived from sample deviations.

2) Currency exchange: the Hodge decomposition: In this study, we consider the dataset of currency exchange rates of 9 countries from 27 Jul. 2018 to 26 Jul. 2023. For each pair of currencies, there is a single given exchange rate. For **example**, between the US dollar (USD) and Euro (EUR), only the buying value of 1 USD in EUR is provided. Therefore, if we model the currencies by nodes of a graph $G = (V, E)$, there is a canonical set of edge orientations associated with the dataset. We build a simplicial complex X with edges E and there is a 2-complex for each triple of nodes.

Suppose there is no arbitrage. Then for each triple of currencies v_1, v_2, v_3, the exchange rates $r_{1,2}, r_{2,3}$ and $r_{1,3}$ should satisfy $r_{1,2}r_{2,3} = r_{1,3}$. We may turn this into an additive identity on the associated 2-complex by applying the natural $\log$ function to the exchange rates. With such preprocessing, we observe edge signals that are $\log$ of exchange rates.

We assume that on each day, signals on $1/3$ of (uniformly) randomly chosen edges are missing. Therefore, the observed signals are both incomplete and asynchronous. The framework proposed in the paper is suitable to tackle these issues.

Specifically, for each edge $e = (v_i, v_j) \in E$ and start date t, we collect observed signals within 15 consecutive days starting from t. There are usually less than 15 signals due to missing data. However, we can still fit these signals with a polynomial $P_{e,t}$ of degree 2. Therefore, if we let $\mathcal{A}$ be the space of degree 2 polynomials (over the finite interval $[1, 15]$), $P_{e,t}$ is a signal in $\mathcal{A}$. Applying the procedure for all edges, we obtain $\mathbf{x}_t = \{P_{e,t} \mid e \in E\} \in C_1(X, \mathcal{A})$.

Endow $\mathcal{A}$ with the L^2-norm and we solve the optimization (2). As there is no non-trivial edge cycle in X (each edge cycle is the boundary of a union of triangles), it is known from algebraic topology that we do not need to consider the harmonic component for $\mathbf{x}_t$ [18]. As a result, the solution $\mathbf{x}_t \approx \mathbf{x}'_t = \partial_1^* \mathbf{y}_{-1,t} + \partial_2 \mathbf{y}_{1,t}$ with $\mathbf{y}_{-1,t} \in C_0(X, \mathcal{A})$ and $\mathbf{y}_{1,t} \in C_2(X, \mathcal{A})$.

As we have discussed, the "no arbitrage" condition enforces $\mathbf{y}_{1,t} \approx \mathbf{0}$. The signal $\mathbf{y}_{-1,t} = \{Q_{v,t} \in A \mid v \in V\}$ alone can be useful for us to probe changes and trends in the currency exchange market. Intuitively, $Q_{v,t}$ describes the divergence of the signal at node v [4], which can be used to detect change and anomalous behavior at v. As each function is over $[1, 15]$, $Q_{v,t}(1)$ indicates the divergence of the first date over the period $[t+0, t+14]$. In Fig. 4, for each $v \in V$, we visualize by plotting the time series $\{Q_{v,t}(1), t \in T\}$ where T is from 27 Jul. 2018 to 26 Jul. 2023, indexed by 1 to 1819.

For each curve in Fig. 4, we see that there are noticeable changes, patterns, and trends. We summarize some findings and compare them to major real financial and historical events in Table I. From the observations, we see that the approach indeed provides us with useful information, even with incomplete and asynchronous data.

V. CONCLUSION

In this paper, we have extended the existing framework of TSP on simplicial complexes to handle a wider range of signal spaces. We consider observations on simplexes as elements of function spaces, so that we may model asynchronous time series more accurately. We discuss the decomposition theory of such signals with a novel optimization framework. Numerical experiments show the effectiveness of the proposed framework.

REFERENCES

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

[2] G. Leus, A. G. Marques, J. M. Moura, A. Ortega, and D. I. Shuman, "Graph signal processing: History, development, impact, and outlook," *IEEE Signal Process. Mag.*, vol. 40, no. 4, pp. 49–60, Jun. 2023.

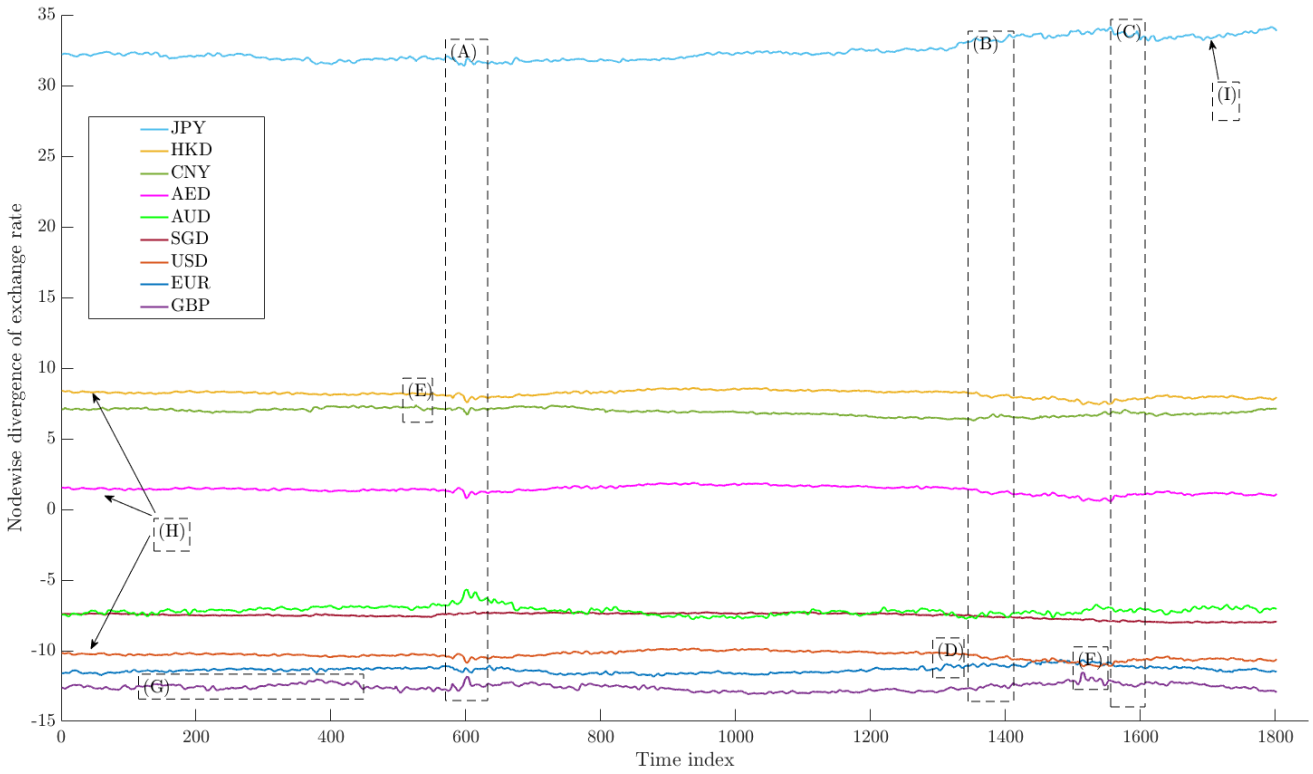


Fig. 4. Plots of $Q_{v,t}(1)$ for all nodes corresponding to the nine currencies.

[3] X. Jian, F. Ji, and W. P. Tay, "Generalizing graph signal processing: High dimensional spaces, models and structures," *Foundations and Trends® in Signal Processing*, vol. 17, no. 3, pp. 209–290, 2023.

[4] S. Barbarossa and S. Sardellitti, "Topological signal processing over simplicial complexes," *IEEE Trans. Signal Process.*, vol. 68, pp. 2992–3007, 2020.

[5] S. Sardellitti, S. Barbarossa, and L. Testa, "Topological signal processing over cell complexes," in *Proc. Asilomar Conf. on Signals, Systems and Computers*, Pacific Grove, CA, USA, Nov. 2021.

[6] S. Barbarossa and S. Sardellitti, "Topological signal processing: Making sense of data building on multiway relations," *IEEE Signal Process. Mag.*, vol. 37, no. 6, pp. 174–183, 2020.

[7] S. Sardellitti and S. Barbarossa, "Robust signal processing over simplicial complexes," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, Singapore, Singapore, May 2022.

[8] F. Ji, G. Kahn, and W. P. Tay, "Signal processing on simplicial complexes with vertex signals," *IEEE Access*, vol. 10, pp. 41 889–41 901, 2022.

[9] T. M. Roddenberry, F. Frantzen, M. T. Schaub, and S. Segarra, "Hodgelets: Localized spectral representations of flows on simplicial complexes," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, 2022, pp. 5922–5926.

[10] M. Yang, E. Isufi, M. T. Schaub, and G. Leus, "Simplicial convolutional filters," *IEEE Trans. Signal Process.*, vol. 70, pp. 4633–4648, 2022.

[11] T. M. Roddenberry, V. P. Grande, F. Frantzen, M. T. Schaub, and S. Segarra, "Signal processing on product spaces," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, Rhodes Island, Greece, 2023.

[12] N. Perraudin, A. Loukas, F. Grassi, and P. Vandergheynst, "Towards stationary time-vertex signal processing," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, New Orleans, US, Mar. 2017.

[13] F. Grassi, A. Loukas, N. Perraudin, and B. Ricaud, "A time-vertex signal processing framework: Scalable processing and meaningful representations for time-series on graphs," *IEEE Trans. Signal Process.*, vol. 66, no. 3, pp. 817–829, 2018.

[14] A. Loukas and N. Perraudin, "Stationary time-vertex signal processing," *EURASIP J. Adv. Signal Process*, vol. 2019, no. 1, pp. 1–19, Aug. 2019.

[15] R. Money, J. Krishnan, B. Beferull-Lozano, and E. Isufi, "Online edge flow imputation on networks," *IEEE Signal Process. Lett.*, vol. 30, pp. 115–119, 2023.

[16] F. Ji and W. P. Tay, "A Hilbert space theory of generalized graph signal processing," *IEEE Trans. Signal Process.*, vol. 67, no. 24, pp. 6188 – 6203, 2019.

[17] X. Jian and W. P. Tay, "Wide-Sense Stationarity in Generalized Graph Signal Processing," *IEEE Transactions on Signal Processing*, vol. 70, pp. 3414–3428, 2022.

[18] A. Hatcher, *Algebraic Topology*. Cambridge University Press, 2002.

[19] F. Ji, X. Jian, W. P. Tay, and M. Yang, "Generalized signals on simplicial complexes," *arXiv preprint arXiv:2305.06899*, 2023.

[20] S. Lang, *Algebra*, 3rd ed. New York, NY: Springer-Verlag, 2002.

[21] L. Cohen, *Time Frequency Analysis: Theory and Applications*. Prentice-Hall, 1995.

[22] T. M. Roddenberry, V. P. Grande, F. Frantzen, M. T. Schaub, and S. Segarra, "Signal processing on product spaces," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Process.*, 2023.

[23] T. Stubbington and N. Asgari, "Pound hits 37-year low against dollar on huge UK tax cut 'gamble'," *Financial Times*, Sep. 2022.