

# LLM Online Spatial-temporal Signal Reconstruction Under Noise

Yi Yan, *Student Member, IEEE*, Dayu Qin, *Student Member, IEEE*, and Ercan E. Kuruoglu, *Senior Member, IEEE*

**Abstract**—This work introduces the LLM Online Spatial-temporal Reconstruction (LLM-OSR) framework, which integrates Graph Signal Processing (GSP) and Large Language Models (LLMs) for online spatial-temporal signal reconstruction. The LLM-OSR utilizes a GSP-based spatial-temporal signal handler to enhance graph signals and employs LLMs to predict missing values based on spatiotemporal patterns. The performance of LLM-OSR is evaluated on traffic and meteorological datasets under varying Gaussian noise levels. Experimental results demonstrate that utilizing GPT-4-o mini within the LLM-OSR is accurate and robust under Gaussian noise conditions. The limitations are discussed along with future research insights, emphasizing the potential of combining GSP techniques with LLMs for solving spatiotemporal prediction tasks.

**Index Terms**—Large Language Model, Graph Signal Processing, spatial-temporal graph, online prediction.

## I. INTRODUCTION

RECENT advancements in artificial intelligence have led to breakthroughs in many fields such as healthcare diagnostics [1] and investment portfolio construction [2], culminating in the development of Large Language Models. Large Language Models (LLMs) are a type of artificial intelligence model that is designed for natural language processing (NLP) tasks for its ability to understand and generate large-scale texts [3]. BERT is a groundbreaking predecessor to modern LLMs by demonstrating the power of bidirectional transformers for natural language understanding [4]. Modern LLMs, such as GPT-3 [5], GPT-4 [6], ERNIE [7] and Kimi AI [8], are trained on datasets with billions of words and typically transformers architecture, to manipulate natural language [9]. Several applications of LLMs have been explored: LLMs enhance interactive machine translation by delivering high-quality initial translations, adapting efficiently to user feedback, and minimizing training costs [10]; LLMs enhance sentiment analysis by generating domain-specific weak labels and enabling efficient model distillation for practical applications. [11]. However, LLMs still exhibit certain limitations. For example, if the information provided to LLMs is insufficient or the prompt is misleading, inaccurate, or incomplete healthcare decisions can be made by LLMs which can lead to physical or psychological harm [12]. Despite their strength in processing text-based information, LLMs remain limited in

handling multivariate data structures, necessitating exploration into methods like Graph Signal Processing (GSP) and Graph Neural Networks (GNNs) to analyze and model such complex datasets effectively.

Graph-based methods provide a powerful framework for modeling and analyzing correlations in complex multi-variate data and have been applied in many fields, such as neurological disorders screening [13] and financial crisis prediction [14]. There are some application scenarios that are better suited for graph methods compared with computer vision (CV) and natural language processing (NLP) approaches, for instance, social network analysis [15], [16], traffic prediction [17], and quantum computing [18]. By exploiting the graph topology, interactions among multivariate data are captured along with the multi-variate data, offering task performance that outperforms non-graph algorithms. In addition, GSP with spectral approaches enables efficient representation and extraction of spectral patterns in addition to the spatial patterns seen in graphs and has a wide range of applications, including flaw detection in the wire-based directed energy deposition [19] and electroencephalography signal processing [20]. Graph-based methods are conventionally applied to static machine learning tasks such as classification, regression (on time-invariant data), and clustering, which involve data analysis without considering temporal changes [21], [22]. However, some particular tasks such as traffic prediction, climate modeling, and financial forecasting, rely heavily on capturing spatial-temporal dependencies. Spatial-temporal graph algorithms, GSP and GNNs, were designed to solve these kinds of time-varying tasks and have succeeded in dealing with such challenges by effectively modeling relationships and temporal dynamics in graph-structured data [23], [24], [25], [17].

Some recent studies have further explored the integration of GSP with Large Language Models (LLMs), revealing substantial advantages. Firstly, this integration potentially expands the application of LLMs by enabling them to process and analyze time-varying graph structure data, which allows LLMs to engage in graph-based reasoning across varied scenarios such as academic and e-commerce networks [26]. Secondly, GSP can leverage LLMs to process time-varying, multivariate data from a text-based perspective, providing a novel angle for analyzing dynamic complex networks [26]. Notably, LLMs on graphs perform well not by relying on data leakage but because they interpret graphs as languages, where the node label (signal) is deemed more crucial than the structure itself [27]. Additionally, the InstructGLM framework demonstrates how LLMs can effectively represent graph structures through natural language for node classification in citation networks. T

Yi Yan and Dayu Qin contributed equally. Yi Yan was affiliated with the Tsinghua-Berkeley Shenzhen Institute, Shenzhen International Graduate School, Tsinghua University, during the completion of this work. Dayu Qin and Ercan E. Kuruoglu are currently affiliated with the Tsinghua-Berkeley Shenzhen Institute, Shenzhen International Graduate School, Tsinghua University. Corresponding author: Ercan E. Kuruoglu; e-mail: kuruoglu@sz.tsinghua.edu.cn.

his approach eliminates the need for complex GNN pipelines and unifies graph learning with natural language processing, showcasing the potential of LLMs in graph tasks [28].

In our work, we introduced a novel method, the Large Language Model for Online Spatial-temporal Reconstruction (LLM-OSR) algorithm, which combines the strengths of the time-varying GSP method and LLMs for efficient and accurate reconstruction of missing signals in dynamic spatiotemporal complex networks. This integrated approach uses GSP to denoise or enhance signal features, ensuring high-quality input for LLM-based prediction tasks. The combination of GSP with LLMs presents noteworthy advantages. Firstly, this integration has the potential to expand the application of LLMs by enabling them to process and analyze time-varying graph structure data. Secondly, GSP can leverage LLMs to process time-varying, multivariate data from a text-based perspective, providing a novel perspective for analyzing dynamic complex networks.

There are 2 main contributions to our paper:

- We introduced the LLM-OSR algorithm, a novel approach to reconstructing spatial-temporal signals in an online manner by seamlessly combining GSP-based techniques with LLM-driven predictors. This innovative approach effectively reconstructs time-varying graph signals in the presence of noise and missing values within spatiotemporal data.
- The LLM-OSR employs a sophisticated reverse embedding approach to transform spatial-temporal signals on graphs into coherent and contextually meaningful natural language expressions, making the information readily interpretable and actionable by LLMs.

Here is the organization of this paper. The preliminary knowledge is presented in Section II. Section III provides a detailed discussion of the LLM-OSR. The experimental results and corresponding discussions are covered in Section IV. The limitations and some potential future extensions of the proposed LLM-OSR are discussed in Section V provides an in-depth discussion of the limitations of LLM-OSR and outlines potential directions for future research. Finally, Section VI concludes the paper.

## II. PRELIMINARI KNOWLEDGE

### A. GSP Preliminaries

In this paper, we consider an undirected and unweighted graph  $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ , where  $\mathcal{V} = \{v_1, \dots, v_i\}$  is the set of nodes or vertices, and  $\mathcal{E}$  is the set of edges. We can represent the topology of the graph using the adjacency matrix  $\mathbf{A}$ , where  $\mathbf{A} \in \mathbb{R}^{N \times N}$  and its elements are defined as follows:

$$A_{ij} = \begin{cases} 1, & \text{if there is an edge between } v_i \text{ and } v_j, \\ 0, & \text{otherwise.} \end{cases} \quad (1)$$

The time-varying graph signal  $\mathbf{x}[t] \in \mathbb{R}^N$  is the multi-variate numerical data recorded on the graph nodes that change over time.

In GSP, the spectral operations are defined using the graph Laplacian matrix  $\mathbf{L} \in \mathbb{R}^{N \times N}$ :

$$\mathbf{L} = \mathbf{D} - \mathbf{A}, \quad (2)$$

where  $\mathbf{D}$  is the degree matrix, defined as  $\mathbf{D} = \text{diag}(\mathbf{1}^T \mathbf{A})$  and  $\mathbf{1}$  is an all ones vector. The spectral operations are conducted through GSP by having the GFT as the analogy of the classical Fourier Transform and can be realized by the eigendecomposition of  $\mathbf{L}$ :

$$\mathbf{L} = \mathbf{U} \text{diag}(\boldsymbol{\lambda}) \mathbf{U}^T, \quad (3)$$

where  $\mathbf{U}$  is the orthonormal eigenvalue matrix and  $\boldsymbol{\lambda}$  is the vector of eigenvalues. In GSP, the Laplacian matrix  $\mathbf{L}$  of an undirected and unweighted is a symmetric semi-definite matrix. The eigenvectors serve as graph Fourier bases and eigenvalues represent graph frequencies. The eigenvalue-eigenvector pairs are sorted in increasing order: smaller eigenvalues correspond to smoother variations (low frequencies) in the graph signal and larger eigenvalues correspond to rapid variations (high frequencies) [29].

Spectral operations can be conducted through applying filters  $\sum_{f=1}^F h(\boldsymbol{\lambda})_f$  to the signal  $\mathbf{x}$  in the spectral domain through GFT:

$$\tilde{\mathbf{x}} = \mathbf{U} \sum_{f=1}^F h(\boldsymbol{\lambda})_f \mathbf{U}^T \mathbf{x}, \quad (4)$$

where  $\tilde{\mathbf{x}}$  is the processed signal. The formulation in (4) is also known as the graph convolution. By implementing various filters, such as high-pass and low-pass filters, we can perform tasks like denoising and feature enhancement [30].

### B. LLM preliminaries

The Generative Pre-trained Transformer 3 (GPT-3) is a transformer-based large-scale autoregressive language model developed by OpenAI. With 175 billion parameters, it is significantly larger than its predecessor, GPT-2, and was designed to excel in task-agnostic few-shot learning. GPT-3 can perform a wide array of NLP tasks, such as text generation, translation, summarization, and question answering, without requiring task-specific fine-tuning. Instead, it leverages in-context learning, where the model can adapt to new tasks by interpreting prompts and few-shot demonstrations directly within its input context [31]. The GPT-3 is capable perform various tasks, such as creative writing and generating code, highlighting its broad applicability [5]. Despite these strengths, GPT-3 has notable limitations, such as struggles with long-term coherence, logical consistency, and susceptibility to generating factually incorrect or biased content. Building on the GPT-3 foundation, GPT-4 introduces several key advancements that address some of these limitations of GPT-3. GPT-4 offers improved performance across a wider range of tasks and demonstrates enhanced capabilities in reasoning, inference, and contextual understanding by incorporating improvements such as multimodal processing, larger context window, and optimized scaling laws to deliver more accurate and reliable outputs [32]. These technical improvements enable GPT-4 to tackle more complex tasks, such as multimodal reasoning and handling intricate logical chains, with greater precision. For example, the GPT-4 is capable of scoring in the top 10% on simulated bar exams, compared to GPT-3.5, which could score only at the bottom 10% [33].

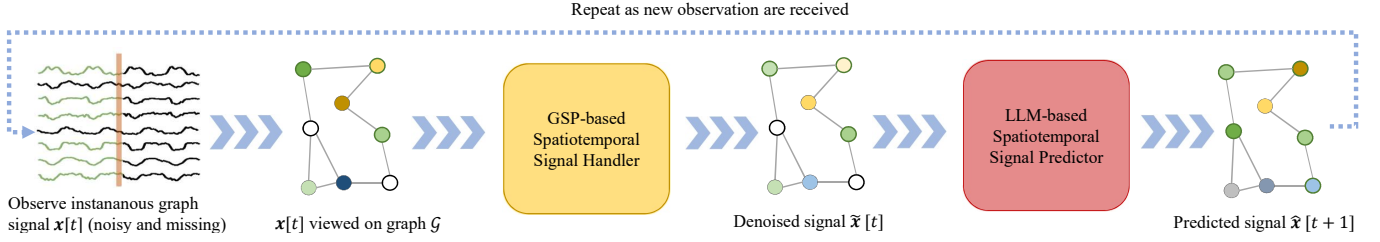


Fig. 1. An overview of the LLM-OSR workflow

### III. METHODOLOGY

#### A. Methodology overview

The LLM-OSR algorithm reconstructs missing graph signals by combining GSP-based processing and LLM prediction. To provide an intuition of the LLM-OSR, it first enhances signal features with GSP and then predicts missing values via the LLM for time-varying spatiotemporal data. The entire process operates online, meaning that the spatiotemporal reconstruction is performed in real time as new signal observations are continuously received and processed by the LLM-OSR. The signal observation model used in LLM-OSR can be expressed as

$$\mathbf{o}[t] = \mathbf{M}(\mathbf{x}_g[t] + \boldsymbol{\epsilon}[t]), \quad (5)$$

where  $\mathbf{M}$  is the observation mask,  $\mathbf{x}_g[t]$  is the ground truth graph signal, and  $\boldsymbol{\epsilon}[t]$  is the i.i.d. zero-mean additive Gaussian noise to the graph signal. Consider a graph  $\mathcal{G}$  with a total of  $N$  nodes, and a subset of these nodes  $\mathcal{O} \subseteq \mathcal{V}$ , where only  $O$  out of  $N$  nodes are observed. Using the observation set  $\mathcal{O}$ , we can construct a masking matrix  $\mathbf{M}$  to model the signal observation:

$$\mathbf{M} = \text{diag}([\mathcal{O}(v_1), \dots, \mathcal{O}(v_O)]^T), \quad (6)$$

where the membership of a node  $v_i$  in the observation set is defined as:

$$\mathcal{O}(v_i) = \begin{cases} 1, & \text{if } v_i \in \mathcal{O}, \\ 0, & \text{if } v_i \notin \mathcal{O}. \end{cases} \quad (7)$$

Here,  $\mathcal{O}(v_i)$  acts as an indicator function that determines whether  $v_i$  belongs to the observed set, and  $\mathbf{M}$  is a diagonal matrix with entries corresponding to these indicator values. This assumption is fundamental in various applications, including climate change analysis [34], skeleton-based gait recognition [35], and brain studies [36]. The workflow of one iteration of LLM-OSR is shown in Figure 1. An overview of the LLM-OSR framework is shown in Algorithm 1.

#### B. GSP-based Spatial-temporal Signal Handler

The proposed GSP-based Spatial-temporal Signal Handler aims to leverage GSP techniques to denoise and enhance spatial-temporal signals to prepare the data for the LLM-based spatiotemporal signal predictor.

During the training phase of LLM-SRO, the goal is to learn the optimal filter parameters within the GSP-based spatial-temporal signal handler from the training data. In our case, since we assume that the signal observations received by

#### Algorithm 1 LLM-OSR Overview

- 1: **Train Phase:**
- 2: Initialize and train the GSP-based spatial-temporal signal handler as seen in Algorithm 2
- 3: **Testing Phase:**
- 4: Deploy the trained GSP-based spatial-temporal signal handler
- 5: Prepare the LLM-based spatial-temporal signal predictor
- 6: **while** new observations  $\mathbf{o}[t]$  are available **do**
- 7:   Gather  $\hat{\mathbf{x}}[t-1]$ , the previous signal estimate
- 8:   Process the observations  $\mathbf{o}[t]$  with the GSP-based spatial-temporal signal handler and Collect the processed observations  $\tilde{\mathbf{o}}[t]$
- 9:   Pass  $\tilde{\mathbf{o}}[t]$  and  $\hat{\mathbf{x}}[t-1]$  into the LLM-based spatial-temporal signal predictor
- 10:   Operate the LLM-based spatial-temporal signal predictor as seen in Algorithm 3
- 11:   Collect reconstructed  $\hat{\mathbf{x}}[t]$
- 12: **end while**

the LLM-OSR contain noise and missing nodes, we would like to let the GSP-based spatial-temporal handler perform a denoising task and feed the denoised observations to the LLM-based spatial-temporal signal predictor. To give an overview, the LLM-OSR takes a universal approach to learning the filter parameters  $\sum_{f=1}^F h(\boldsymbol{\lambda})_f$  in (4) by iteratively applying the graph convolution (4) to the training data, calculating the loss of the estimation, and updating the filter parameter. We augment the training set by concatenating multiple copies to increase the number of time samples. During each iteration, the GSP-based spatial-temporal handler is trained using data from a single time instance. First, we apply the graph convolution operation in (4) to obtain the signal representation. Next, we compute the Mean Absolute Error (MAE) between the predicted output and the ground truth at each time instance  $t$ , serving as the performance metric for the current filter:

$$\text{MAE}[t] = \frac{1}{N} \sum_{n=1}^N |x_{g,n}[t] - \tilde{x}_n[t]|, \quad (8)$$

where  $\tilde{\mathbf{x}}$  is the processed graph signal,  $\hat{x}_n[t]$  it the processed signal on the  $n^{\text{th}}$  node within  $\tilde{\mathbf{x}}$ , and  $x_{g,n}[t]$  is the signal on the  $n^{\text{th}}$  node within the ground truth signal  $\mathbf{x}_g[t]$ . The actual learning and updating of the filter are done by calculating the gradient of the MAE with respect to the filter parameters

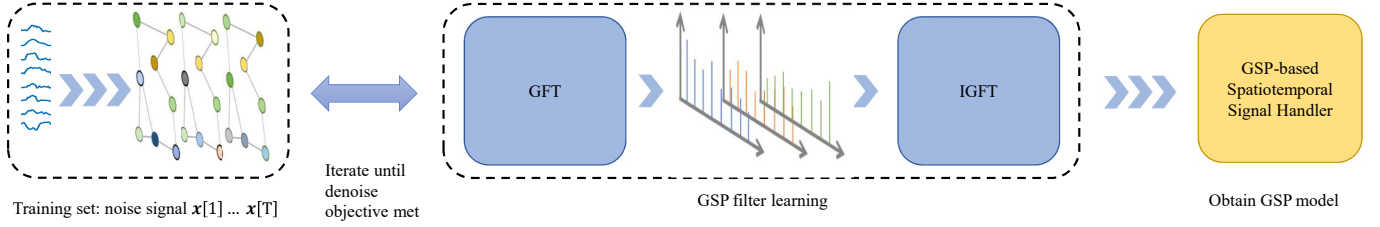


Fig. 2. The training process of the GSP-based spatial-temporal signal handler.

$\sum_{f=1}^F h(\lambda)_f$ , denoted as  $\nabla_{\mathbf{H}}\text{MAE}$  then updated the filter parameters using the gradient descent rule:

$$\sum_{f=1}^F h(\lambda)_f = \sum_{f=1}^F h(\lambda)_f - \eta \cdot \nabla_h \text{MAE}[t], \quad (9)$$

where  $\eta$  is the learning rate. We repeat this process is repeated iteratively until we run out of training samples or when the MSE converges to a steady value (early stopping). Since the graph filter parameters are applied within the graph convolution, it is essentially learned based on the knowledge of both the graph structure and associated graph signals. These parameters are optimized to minimize noise and preserve important spatial and spectral graph features in the data.

Once training is complete, the learned GSP filter parameters are fixed and stored in the GSP-based spatial-temporal signal handler for use during the test phase. We assume that the signals in the training set and the testing set have similar spectrums. During the test phase, the pre-trained GSP-based spatial-temporal signal handler is applied to unseen test samples  $\mathbf{o}[t]$  as they are received in real-time to enhance their quality by denoising. In other words, since each test sample  $\mathbf{o}[t]$  is an observation with missing values and noise, the pre-trained GSP-based spatial-temporal signal handler is applying a series of graph filters  $\sum_{f=1}^F h(\lambda)_f$  to  $\mathbf{o}[t]$  to denoise using the graph convolution (4). Since the process is completed through the graph convolution (4), the signal observations are processed with the knowledge of the underlying spatial structure of the graph  $\mathcal{G}$ . The logic of training and deploying the GSP-based spatial-temporal signal handler data can be found in Figure 2.

### C. LLM-based spatial-temporal signal predictor

The first step of our LLM-based spatial-temporal signal predictor is a reverse embedding function to process the denoised signals. In NLP tasks, the embedding process typically involves transforming words, phrases, or sentences into numerical vectors by mapping them into a designated feature space [37]. In conventional graph embedding tasks, data on the graph is typically transformed into a feature space using methods such as GCN or Node2Vec, similar to traditional NLP embeddings [38]. Here in the LLM-OSR, we take a reverse approach to the conventional embedding approaches such as the Node2Vec. Since the time-varying graph signal observations  $\mathbf{o}[t]$  are already numerical, instead of embedding data into designated space, we directly represent the local

### Algorithm 2 GSP-based spatial-temporal signal handler

- 1: **Train Phase:**
- 2: Given the training data  $\mathbf{x}[1] \dots \mathbf{x}[T]$
- 3: Initialize graph Laplacian  $\mathbf{L} = \mathbf{D} - \mathbf{A}$  (2) and GFT  $\mathbf{L} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^T$  (3)
- 4: Initialize filter parameters  $h(\Lambda)$
- 5: **while** halting condition not met **do**
- 6:   Apply  $\tilde{\mathbf{x}}[t] = \mathbf{U}\text{diag}(h(\Lambda))\mathbf{U}^T \mathbf{x}[t]$ , the graph convolution, as seen in (4)
- 7:   Compute MAE at each time instance  $t$ :  
 $\text{MAE}[t] = \frac{1}{N} \sum_{n=1}^N |x_{g,n}[t] - \tilde{x}_n[t]|$  as seen in (8)
- 8:   Calculate gradient  $\nabla_h \text{MAE}$  with respect to  $h(\Lambda)$
- 9:   Update filter parameters using gradient descent:  
 $h(\Lambda) = h(\Lambda) - \eta \cdot \nabla_h \text{MAE}$  as seen in (9)
- 10: **end while**
- 11: Export the parameters after training is complete
- 12: **Test Phase:**
- 13: **for** each received test observation sample  $\mathbf{o}[t]$  containing noise and missing node observation **do**
- 14:   Process  $\mathbf{o}[t]$  using graph convolution and the trained filter  $\text{diag}(h(\Lambda))$ :  $\tilde{\mathbf{o}}[t] = \mathbf{U}\text{diag}(h(\Lambda))\mathbf{U}^T \mathbf{o}[t]$
- 15:   Apply  $\tilde{\mathbf{o}}[t]$  to be fed into LLM for reconstruction
- 16: **end for**

topological connections of nodes on the graph in the spatial domain through LLM natural language expressions (English passages), which consists of text and numbers. Mathematically, extracting the 1-hop localized neighbors of the node  $v_i$  can be achieved by identifying the nonzero entries in the  $i^{\text{th}}$  row of the adjacency matrix  $\mathbf{A}$ . This can be expressed as:

$$\text{Neighbors}(v_i) = \mathcal{A}_{v_i} = \{v_j \mid \mathbf{A}[i, j] \neq 0, j \neq i\}, \quad (10)$$

where  $\text{Neighbors}(v_i)$  represents the set of neighboring nodes of  $v_i$ . Notice that in our problem setting, there is only a subset of the nodes being observed. So in the implementation of LLM-OSR, we only consider the processed node signals from the observed node neighbors.

These expressions are then further put into the prompt as the tasks  $\mathcal{T}(v_i)$  on each node to be solved by the LLMs, specifying the LLM to conduct a prediction based on the spatial-temporal information. In the spatial domain, the task  $\mathcal{T}(v_i)$  will consist of an expression of which nodes are neighbors to node  $v_i$  and the processed observed values of its neighbors. In the temporal domain, the task  $\mathcal{T}(v_i)$  will consist of an expression of the past estimated signal values of node  $v_i$ . Conceptually, we

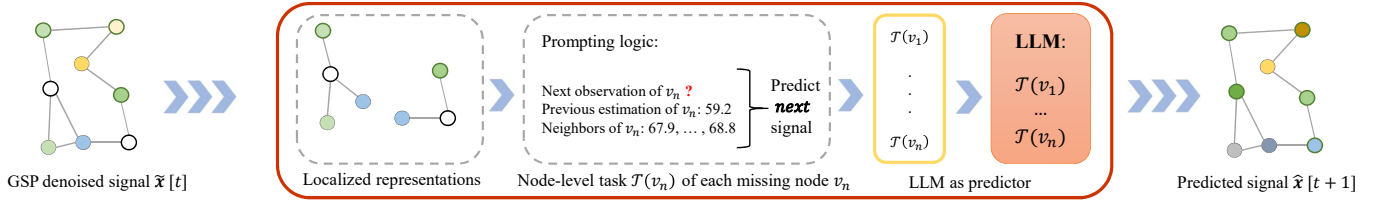


Fig. 3. The LLM-based Spatial-temporal Signal Predictor.

can understand the task of the LLM as to aggregate neighbor signals of  $v_i$  with self-aggregation; this task can be denoted as:

$$\mathcal{T}(v_i) = \text{agg}(\{(\hat{x}_i[t-1], \tilde{x}_j[t]) \mid j \in (\mathcal{A}_{v_i} \cup \mathcal{O})\}). \quad (11)$$

This approach shifts the role of the LLM from performing logical reasoning or mathematical calculations, which is a known weakness of the LLM, to generating data based on semantic understanding and context, leveraging the strengths of LLM in processing natural language descriptions. The aim of the LLM-based Spatial-temporal Signal predictor is to let the LLM leverage the smoothness assumption to infer the target values of each node based on the values of its neighboring nodes (spatial) and the past estimation (temporal).

The LLMs used within our LLM-OSR framework are GPT-3.5-turbo and GPT-4-o mini, which belong to the GPT-3 and GPT-4 families, respectively. The GPT-3 leverages in-context learning to perform various tasks—such as translation, question answering, and text completion—without requiring task-specific fine-tuning [31]. The architecture of the GPT-3 is based on transformers and was trained on a diverse dataset sourced primarily from the internet. GPT-4 further builds upon the GPT-3 foundation, incorporating key advancements such as multimodal capabilities, a larger context window, optimized scaling laws, and improved safety mechanisms [32]. These enhancements result in significant improvements in inference and logical reasoning of GPT-4, demonstrating better performance compared to GPT-3 at understanding prompts, generating more coherent and contextually appropriate responses, and handling complex tasks requiring step-by-step reasoning [33]. These improvements make GPT-4 more suitable for our tasks of online spatial-temporal signal reconstruction, which involves complex reasoning and spatial-temporal pattern recognition.

The integration of LLMs with the rest of the components within the LLM-based Spatial-temporal Signal Predictor is facilitated through the OpenAI API. To be specific, LLM-OSR directly utilizes the pre-trained model provided and deployed on the OpenAI server. That is, the LLM-OSR does not perform any additional training or fine-tuning. The API provides access to the vanilla model as-is, leveraging its existing knowledge and capabilities to generate predictions. This allows us to take full advantage of the robust, generalized understanding embedded in the model, ensuring a streamlined and efficient process for spatial-temporal signal forecasting.

We followed the approach of structuring the interaction of the reversely embedded graph signal expression with the LLM

using a dual-role setup where the LLM contains a system role and a user role. The system role serves as the supervision guide, providing the global task context and specific constraints to shape the behavior of the LLM. For our spatiotemporal task, this role defines the objective as predicting the current value of a graph node based on its previous value and the values of its neighbors, while also enforcing strict output requirements. In our LLM-OSR, the system instructs the LLM to produce only a single numeric value as output, rounded to a certain decimal place, and to avoid any extraneous text or reasoning. This ensures consistency and simplicity in the generated responses. Below is an example of system role content we used in LLM-OSR:

*The spatiotemporal task is to predict the current number on a graph based on its previous value and the value of its neighbors.*

The user role, on the other hand, dynamically generates task-specific prompts that supply the LLM with the required details for each individual prediction. Here is an example of directly expressing the neighborhood relationship along with the signals for the user role using natural language expression:

*Each indexed content is independent. Make 1 numeric prediction per indexed context. Precision round to 1 decimal point. Do not output text. Do not recall memories. Time 1439, Entity index: 322. Previous: 61.5, Neighbors: [63.9, 57.4].*

These prompts include precise temporal and spatial context, such as the time index, the node index, the previous value of the node, and a list of observed values from its neighbors. By structuring the user prompts this way, we ensure that each query to the LLM is both clear and contextually complete, reducing ambiguity in the response.

Furthermore, the interaction process is optimized by batching multiple prompts for efficiency. When a new observation is received in the LLM-OSR, the user role creates a batch of prompts corresponding in real-time. These batched prompts are sent to the LLM in a single API call, with each prompt corresponding to the task  $\mathcal{T}(v_i)$  on a single node, leveraging the prompt-response structure to efficiently handle multiple predictions.

Notice that we further included an error-checking function for invalid LLM responses. If an LLM call fails, such as when the LLM response does not include a valid numeric value, our implementation automatically retries the failed predictions. This retry mechanism regenerates the prompts for the unresolved cases and resubmits them to the LLM, with a



predefined maximum number of retries. This additional error-checking function ensures that exceptions and invalid LLM outputs are handled robustly, even in cases where the LLM struggles to produce a valid prediction.

By combining the system and user roles with these efficient handling mechanisms, this structured approach allows the LLM to focus on its core strength of generating numeric predictions while providing the necessary temporal and spatial context for accurate, context-aware results. An illustration of the prompts that provide the content and context to the LLM for solving the node-level tasks  $\mathcal{T}(v_i)|_{v=1\dots N}$  can be found in Figure 4.

---

**Algorithm 3** LLM-based spatial-temporal signal predictor

---

- 1: Initialize the LLM model and form the system and user role prompt templates
  - 2: **while** new processed observations  $\tilde{o}[t]$  are available from the GSP-based spatial-temporal signal handler **do**
  - 3:   Collect the output from the GSP-based filter spatial-temporal signal handler
  - 4:   **for** each missing node  $v_i$  in set  $\mathcal{V} - \mathcal{O}$  **do**
  - 5:     Collect  $\hat{x}_i[t-1]$ , the previous estimation of  $v_i$
  - 6:     Collect node neighborhood for  $v_i$  using (10) and the corresponding processed observed neighbor signals of node  $v_i$  in  $\tilde{o}[t]$
  - 7:     Form LLM task  $\mathcal{T}(v_i)$  from aggregation (11) and their corresponding prompts
  - 8:     Feed  $\mathcal{T}(v_i)$  and their corresponding prompts into the LLM  
      Let LLM predict  $\hat{x}_i[t]$
  - 9:   **end for** (Retry if output invalid)
  - 10:   Collect all node reconstructions and map them to  $\hat{x}[t]$
  - 11: **end while**
- 

## IV. EXPERIMENTS AND DISCUSSION

### A. Experiment Setting

Here we will provide a brief discussion about the datasets, tested algorithms, and experiment settings.

#### 1) Dataset Description:

- **Traffic Data:** We utilize the publicly available Seattle Loop Detector Dataset [39], which contains traffic flow data collected from loop detectors on the highways in the Seattle area. This dataset provides hourly traffic readings that are essential for analyzing spatiotemporal patterns. The experimental setup includes the addition of Gaussian noise with variances of 1.0 and 1.5 to evaluate the robustness of the models under varying levels of data corruption. The graph topology is constructed by mapping the physical locations of  $N = 323$  loop detectors to their corresponding positions along the actual highway path. Each loop detector is a node on the graph  $\mathcal{G}$ . The traffic speed is recorded in 5-minute intervals. We selected a sub-portion of the signal consisting of 7 days of reading, making the size of the data  $\mathbb{R}^{323 \times 2016}$ . The recordings from the first 576 will be in the training set to tune or learn the model parameters and the rest are in the

testing set. An illustration of 4 different time instances of this time-varying dataset can be found in Figure 5,

- **Meteorological Data:** Hourly wind speed and temperature data are obtained from NOAA [40]. Because of the behavior differences between wind speed and temperature, we analyze them as separate datasets in the experiments. Each node in the dataset corresponds to a geographic location defined by its latitude and longitude. We selected  $N = 197$  stations that contain no missing recordings in 3 consecutive days, giving us  $\mathbb{R}^{197 \times 96}$ . We split the first 24 time steps into the training set and the rest into the testing set. To capture the spatial dependencies among nodes, a k-nearest-neighbor (kNN) graph is constructed, where the edges' weights are computed by using a Gaussian kernel method which is described in the GNLMS framework [41].

We set the node observation ratio to be 70% for all the datasets. The missing nodes are missing throughout the entire experiment, making it challenging to infer the missing signals without the utilization of the graph topology. The goal of the experiments is that given an observation  $o[t]$  that is only partially observed, reconstruct the ground truth signal  $x_g[t]$  from the observation and past  $p$  estimations.

2) *Considered Algorithms:* We consider 2 distinct settings of the LLM-OSR. The first setting is to use GPT-3.5 turbo as the LLM within the LLM-OSR; we denote this setting as the LLM-OSR-3.5. The second setting is to use GPT-4-o mini as the LLM within the LLM-OSR, denoted as LLM-OSR-4. The 2 LLM-OSR variants will be evaluated against a variety of baseline algorithms, including graph adaptive filters, graph time-series analysis algorithms, and GNNs:

- **GLMS** [42]: An adaptive filter designed for online graph signal estimation under Gaussian noise, derived from an LMS optimization problem.
- **GNLMS** [43]: A variant of the GLMS that incorporates spectral normalization to enhance performance.
- **GNS** [44]: An adaptive filter developed for online graph signal estimation under impulsive noise, derived from an  $L_1$  optimization problem.
- **GCN** [22]: A widely recognized graph neural network (GNN) where each layer applies the graph convolution (4), incorporating spatial normalization and a non-linear activation function.
- **GVARMA** [45]: A time-series analysis method that extends the classical VARMA model into the graph signal processing (GSP) domain by defining ARMA parameters using the graph Fourier transform (GFT) in the graph spectral domain.
- **GGARCH** [46]: A time-series analysis method analogous to the classical GARCH model, adapted to the GSP domain by defining GARCH parameters using the GFT in the graph spectral domain.
- **RGDAN** [17]: A GNN architecture that combines graph diffusion-based modeling with spatial and temporal attention mechanisms to capture complex relationships in graph-structured data.

Notice that other than the adaptive filters and the LLM-OSRs,

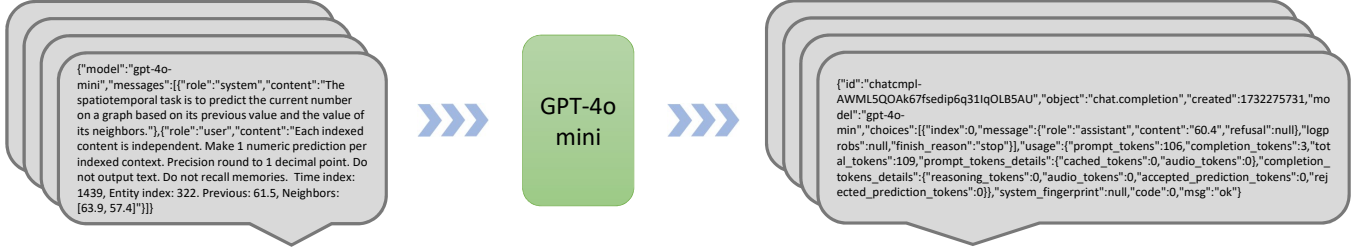


Fig. 4. The prompts prepared for LLM and the responses generated by the LLM (GPT-4o mini).

the other algorithms are offline algorithms during training.

3) *Computational Environment*: The experiments are conducted on a workstation equipped with the following hardware and software:

- **CPU**: Intel Core i9-13900K
- **GPU**: NVIDIA RTX 4090 with 24GB of G6X memory
- **Operating System**: Windows 11

These computational resources provide sufficient computational power for handling large-scale graph computations and training the LLM-OSR model.

### B. LLM-OSR on Traffic Prediction

The enhanced multilingual proficiency of GPT-4 enables it to excel in low-resource languages, and its optimized scaling laws ensure predictable improvements. LLM on graphs perform well not because of leakage, LLM understands graphs as languages instead of the topological structures, and the node label (signal) is more important than structure: [27].

The results of our experiments, as summarized in Tables I and II, demonstrate the significant performance improvements achieved by the LLM-OSR models in traffic prediction tasks on the Seattle Loop Dataset. Specifically, LLM-OSR-4 consistently outperforms all baseline models in terms of both RMSE and MAE. These include GSP-based methods like GLMS, GNLMS, and GNS, GNN-based algorithms such as GCN and RGDAN, and time-series analysis algorithms like GVARMA and GGARCH. These findings emphasize the ability of LLMs to understand graph-encoded traffic data by interpreting node-level signals as linguistic constructs instead of reading numerical data stored as graph representations or numerical embeddings. This unique approach allows LLM-OSR to extract richer, more contextually relevant patterns, thereby leading to superior predictive performance.

The experimental results further reveal a performance gap between LLM-OSR-3.5 and LLM-OSR-4. While GPT-4-based LLM-OSR-4 achieves good performance, GPT-3.5-based LLM-OSR-3.5 struggles significantly, often performing worse than many baseline graph-based methods. This performance disparity highlights the enhanced modeling capabilities of GPT-4, which can better understand the prompt and node-level signals and adapt to noisy inputs compared to GPT-3.5.

It also shows that earlier LLM versions like GPT-3.5 may lack the robustness required for spatiotemporal graph tasks.

TABLE I  
EXPERIMENT RMSE FOR SEATTLE LOOP DATASET

Model	1.0	1.5
LLM-OSR-3.5	12.23 $\pm$ 3.9e+00	13.76 $\pm$ 4.0e-00
LLM-OSR-4	<b>4.05 <math>\pm</math> 1.6e-01</b>	<b>4.69 <math>\pm</math> 6.8e-03</b>
GLMS	8.04 $\pm$ 2.1e-03	8.07 $\pm$ 2.1e-03
GNLMS	7.91 $\pm$ 1.2e-03	7.91 $\pm$ 9.0e-04
GNS	8.55 $\pm$ 1.5e-03	8.55 $\pm$ 2.3e-03
GCN	26.93 $\pm$ 7e-02	26.95 $\pm$ 4e-02
GVARMA	21.75 $\pm$ 1.3e-03	21.75 $\pm$ 1.3e-03
GGARCH	21.73 $\pm$ 4.0e-04	21.73 $\pm$ 1.3e-03
RGDAN	5.32 $\pm$ 3.2e-01	6.61 $\pm$ 2.1e+00

TABLE II  
EXPERIMENT MAE FOR SEATTLE LOOP DATASET

Model	1.0	1.5
LLM-OSR-3.5	3.52 $\pm$ 3.8e-01	4.23 $\pm$ 3.3e-01
LLM-OSR-4	<b>2.88 <math>\pm</math> 2.3e-02</b>	<b>3.62 <math>\pm</math> 7.4e-03</b>
GLMS	5.09 $\pm$ 1.8e-03	5.12 $\pm$ 2.4e-03
GNLMS	4.88 $\pm$ 9.3e-04	4.89 $\pm$ 7.6e-04
GNS	4.69 $\pm$ 1.6e-03	4.70 $\pm$ 1.2e-03
GCN	19.22 $\pm$ e+00	19.22 $\pm$ e+00
GVARMA	18.53 $\pm$ 3.7e-04	18.53 $\pm$ 1.2e-03
GGARCH	18.54 $\pm$ 3.5e-04	18.55 $\pm$ 1.2e-03
RGDAN	3.23 $\pm$ 1.3e-01	3.96 $\pm$ 1.2e+00

### C. LLM-OSR on Weather Prediction

The performance of LLM-OSR models on weather prediction tasks, as presented in Tables III, IV, V, and VI, highlights their capability in handling spatiotemporal graph data under varying noise conditions. Gaussian noise with variances of 0.2, 0.6, and 1.0 was added to simulate real-world data. The results show that while LLM-OSR-4 excels under lower noise levels, its performance degrades more significantly compared to other models as noise variance increases, which reveals a limitation of current LLM-based approaches.

For hourly wind speed prediction, LLM-OSR-4 achieves exceptional results. When the noise variance is 0.2 and 0.6 it outperforms all baselines. For noise variances of 0.6 and 1.0,

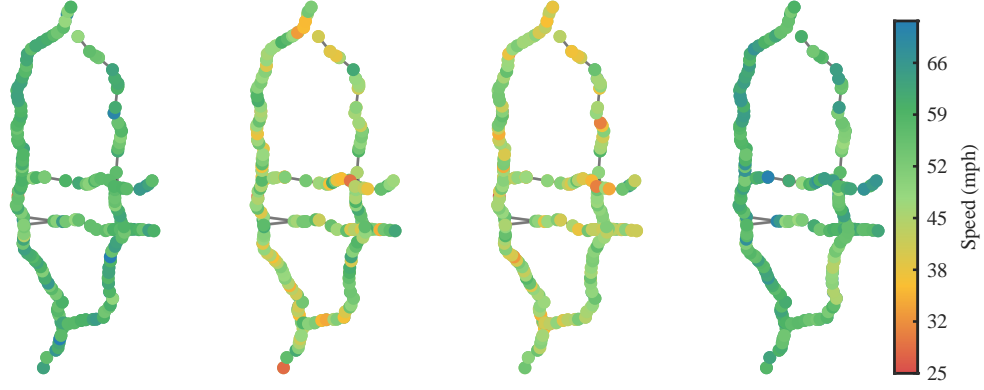


Fig. 5. The Seattle loop dataset at 4 different time instances.

LLM-OSR-4 ranks as the best-performing in terms of RMSE and second-best-performing in terms of MAE. Its second-best-performing cases are surpassed only by RGDAN, a recently proposed, sophisticated model that combines spatial and temporal embeddings. RGDAN achieves this by leveraging GNN diffusion attention mechanisms for spatial embeddings and with temporal attention for temporal embeddings. These embeddings are then integrated using transformer attention, making the RDGAN a powerful algorithm [17]. It is worth noting that both RDGAN and GPT-4-o mini are attention-based algorithms, which highlights the potential for LLM-OSR to further enhance its performance. We will discuss several potential approaches that could potentially boost the performance of LLM-OSR in the next section. In hourly temperature prediction tasks, LLM-OSR-4 maintains its leading performance under low noise conditions for a noise variance of 0.2. However, its performance declines as noise increases compared with RGDAN, but is still the second-best-performing algorithm among all the tested algorithms. Looking at the results within the LLMs, the LLM-OSR-4 outperforms LLM-OSR-3.5 again, suggesting that the GPT-4-o mini is a more appropriate choice than the GPT-3.5 turbo for LLM-OSR.

The results suggest that the performance of LLM-OSR-4 will degrade as the noise variance increases in comparison to non-LLM methods such as RGDAN. Theoretically, algorithms are expected to perform worse as noise increases, but LLM-OSR-4 appears more sensitive to this degradation. This is likely due to the inherent limitations of LLMs. While the GSP-based spatial-temporal handler is capable of noise reduction, it does not completely eliminate noise. The residual noise within the processed signals  $\tilde{o}[t]$  leads to degraded predictions of LLM-OSR. Another potential factor that leads to reduced robustness under high noise conditions is the lack of fine-tuning or retraining of the LLM in the LLM-OSR. We acknowledge the limitations and have considered them in our experimental design. A more detailed discussion of this issue will be presented in the next section. Despite these challenges, the strong performance of LLM-OSR-4 under Gaussian noise conditions demonstrates its potential for spatial-temporal sig-

nal prediction tasks with noisy real-world datasets.

TABLE III  
EXPERIMENT RMSE FOR HOURLY WIND SPEED PREDICTION

Model	0.2	0.6	1.0
LLM-OSR-3.5	$4.58 \pm 9.1e-01$	$5.11 \pm 9.8e-01$	$5.06 \pm 4.9e-01$
LLM-OSR-4	<b><math>1.39 \pm 6.1e-03</math></b>	<b><math>1.70 \pm 7.5e-03</math></b>	<b><math>1.91 \pm 1.1e-02</math></b>
GLMS	$2.19 \pm 1.8e-03$	$2.21 \pm 2.9e-03$	$2.22 \pm 5.4e-03$
GNLMS	$2.18 \pm 1.5e-03$	$2.20 \pm 3.3e-03$	$2.21 \pm 5.8e-03$
GNS	$2.38 \pm 5.8e-03$	$2.39 \pm 6.7e-03$	$2.39 \pm 8.2e-03$
GCN	$2.65 \pm 1e-02$	$2.82 \pm 4.2e-01$	$2.68 \pm 5e-02$
RGDAN	$1.83 \pm 9.0e-02$	$1.89 \pm 5.2e-02$	$1.97 \pm 5.4e-02$
GVARMA	$4.07 \pm 6.7e-03$	$4.09 \pm 1.0e-02$	$4.11 \pm 1.5e-02$
GGARCH	$3.71 \pm 1.9e-03$	$3.73 \pm 3.2e-03$	$3.76 \pm 6.7e-03$

TABLE IV  
EXPERIMENT MAE FOR HOURLY WIND SPEED PREDICTION

Model	0.2	0.6	1.0
LLM-OSR-3.5	$2.77 \pm 8.8e-01$	$3.09 \pm 8.4e-01$	$5.06 \pm 4.9e-01$
LLM-OSR-4	<b><math>1.01 \pm 6.3e-03</math></b>	<b><math>1.35 \pm 6.3e-03</math></b>	$1.54 \pm 1.0e-02$
GLMS	$1.75 \pm 1.5e-03$	$1.76 \pm 1.9e-03$	$1.77 \pm 3.7e-03$
GNLMS	$1.74 \pm 1.3e-03$	$1.75 \pm 2.1e-03$	$1.76 \pm 3.8e-03$
GNS	$1.87 \pm 3.8e-03$	$1.88 \pm 4.5e-03$	$3.53 \pm 5.9e-03$
GCN	$2.10 \pm 4.1e-02$	$2.22 \pm 1.5e-02$	$2.12 \pm 3.4e-02$
RGDAN	$1.35 \pm 6.2e-02$	$1.43 \pm 3.5e-02$	<b><math>1.51 \pm 3.6e-02</math></b>
GVARMA	$3.15 \pm 4.2e-03$	$3.17 \pm 6.5e-03$	$3.19 \pm 9.0e-03$
GGARCH	$3.10 \pm 1.9e-03$	$3.12 \pm 3.9e-03$	$3.14 \pm 7.7e-03$

## V. LIMITATIONS AND FUTURE WORK

The LLM-OSR demonstrated an impressive ability to capture complex relationships and patterns inherent in graph structures while conducting 1-step online prediction on time-varying graph signals. However, as we developed the LLM-OSR, several limitations and challenges emerged that must be addressed when leveraging LLMs on time-varying graph signals.

Let us discuss the LLM-related limitations in the LLM-OSR. First, LLMs are known to have difficulties in understanding numerical data [47]. In our experiments, there are rare occasions that the LLM will output a NaN instead of giving



TABLE V  
EXPERIMENT RMSE FOR HOURLY TEMPERATURE PREDICTION

Model	0.2	0.6	1.0
LLM-OSR-3.5	4.90 $\pm$ 9.2e-01	4.97 $\pm$ 6.8e-01	4.65 $\pm$ 7.7e-01
LLM-OSR-4	<b>1.23 <math>\pm</math> 7.5e-03</b>	1.54 $\pm$ 4.8e-03	1.72 $\pm$ 6.0e-03
GLMS	4.41 $\pm$ 5.1e-03	4.43 $\pm$ 7.1e-03	4.45 $\pm$ 9.1e-03
GNLMS	4.40 $\pm$ 5.7e-03	4.41 $\pm$ 7.0e-03	4.42 $\pm$ 9.1e-03
GNS	5.53 $\pm$ 1.6e-02	5.54 $\pm$ 2.9e-02	5.57 $\pm$ 1.8e-02
GCN	3.09 $\pm$ 5.3e-02	3.10 $\pm$ 6.3e-02	3.13 $\pm$ 7.4e-02
RGDAN	1.33 $\pm$ 1.9e-01	<b>1.33 <math>\pm</math> 9.3e-02</b>	<b>1.37 <math>\pm</math> 1.3e-01</b>
GVARMA	3.13 $\pm$ 2.7e-03	3.17 $\pm$ 4.7e-03	3.22 $\pm$ 4.4e-03
GGARCH	3.22 $\pm$ 2.7e-03	3.26 $\pm$ 4.7e-03	3.30 $\pm$ 6.1e-03

TABLE VI  
EXPERIMENT MAE FOR HOURLY TEMPERATURE PREDICTION

Model	0.2	0.6	1.0
LLM-OSR-3.5	1.78 $\pm$ 1.3e-01	1.97 $\pm$ 8.0e-02	2.02 $\pm$ 8.2e-02
LLM-OSR-4	<b>0.90 <math>\pm</math> 4.1e-03</b>	1.19 $\pm$ 3.2e-03	1.35 $\pm$ 3.6e-03
GLMS	3.03 $\pm$ 1.3e-03	3.05 $\pm$ 3.6e-03	3.07 $\pm$ 4.0e-03
GNLMS	3.02 $\pm$ 1.6e-03	3.03 $\pm$ 3.1e-03	3.04 $\pm$ 3.4e-03
GNS	3.58 $\pm$ 3.8e-03	3.55 $\pm$ 8.2e-03	3.53 $\pm$ 5.9e-03
GCN	2.32 $\pm$ 4.1e-02	2.32 $\pm$ 1.5e-02	2.29 $\pm$ 4.4e-02
RGDAN	0.98 $\pm$ 1.4e-01	<b>0.99 <math>\pm</math> 6.5e-02</b>	<b>1.04 <math>\pm</math> 1.0e-01</b>
GVARMA	2.29 $\pm$ 3.1e-03	2.34 $\pm$ 5.3e-03	2.39 $\pm$ 3.7e-03
GGARCH	2.38 $\pm$ 2.8e-03	2.43 $\pm$ 4.6e-03	2.47 $\pm$ 4.5e-03

us a numeric output. This limitation can be addressed in the future when more powerful LLMs are proposed. Diving deeper into the aspect of the intrinsic limitation of LLMs, we noticed that poorly designed prompts often fail to generate accurate numerical outputs or even any prediction at all. Prompts that are inaccurate or of low accuracy can significantly impair the capabilities of LLM. [48] To improve numerical understanding, prompts should be carefully designed to provide clear instructions and context[49]. In LLM-OSR, LLMs are used as predictors, which means that we do not train or tune the LLMs and let LLMs make zero-shot predictions. The decision not to fine-tune the LLM-OSR models stems from several considerations. However, fine-tuning large language models, such as GPT-4, requires significant computational resources and time. In addition, fine-tuning requires an appropriate balance between the size of the training dataset and the number of parameters in the model. In our case, the available datasets are relatively small compared to the parameter size of GPT-4, making fine-tuning or retraining the LLMs only marginally effective. One potential workaround to retraining and fine-tuning the LLMs would be including examples within the prompt that demonstrate how numerical outputs are expected; this can transform the LLM predictors from zero-shot learners to few-shot learners, which is expected to help guide the model for more accurate predictions. We expect there will be a performance increase if the LLMs are fine-tuned or deployed as a few-shot learning predictor instead of the current zero-shot predictor [50].

There are also spatio-temporal graph-related limitations that could be addressed in future research. In the spatial domain, our current LLM-OSR is limited to processing one node neighborhood per LLM prompt. When a single prompt is used to process multiple nodes, the outputs are frequently

incomplete or contain extraneous elements, as LLMs exhibit limited capability in handling multiple sequences of numbers in a single call. In other words, whenever we attempt to process  $N$  nodes together, LLM often returns a number of prediction outputs that are not  $N$ , and it becomes nearly impossible to align the output when the number of inputs and outputs are mismatched. In the temporal domain, LLMs struggle to comprehend long temporal behaviors and predict longer temporal sequences[51]. These limitations could potentially be solved by more advanced graph representation approaches and advanced programming techniques.

During the development of LLM-OSR, we have encountered challenges in terms of scalability and computational complexity. Currently, the speed of completing each LLM call is constrained by the speed at which LLMs process the input tokens and generate output tokens. Combined with the fact that we are processing one node neighborhood per prompt, the run speed of LLM-OSR is significantly dragged down. We also notice that in our current setting, each LLM call will have to provide content in the prompt describing the task along with the reversely embedded graph signals. Otherwise, the performance will significantly decrease as we progress. This is likely due to the fact that LLMs have limited long-term memory capabilities [52]. Similarly, when the LLM generates outputs, the numerical are often associated with contextual text even when the prompt asks it to output only numerical outputs without text. These limitations lead to a significant amount of unintended token usage that bottlenecks the I/O and bandwidth of LLM-OSR, making it challenging to scale to larger graphs and longer temporal sequences. Other than improving the LLMs themselves, an approach that we plan to take in the future is to implement LLM-OSR using distributive techniques.

Lastly, we would like to expand the fields of applications of LLM-OSR. For instance, we aim to investigate how LLM-OSR performs in applications involving impulsive and heavy-tailed noise, such as communication systems [53] and medical imaging [54]. A potential approach involves modeling the noise not with a Gaussian distribution, but with  $\alpha$ -stable distributions. These distributions are well-suited for such scenarios due to their heavy tails and impulsive characteristics. Incorporating  $\alpha$ -stable distributions into LLM-OSR may enhance its robustness by enabling it to handle extreme values more effectively, thereby improving model stability in datasets with long-tailed distributions [55], [56]. Furthermore, we aim to enhance the contextual handling capabilities of LLM-OSR to broaden its applications by leveraging its contextual inference power of LLM beyond numerical multivariate data. This expansion could pave the way for developing LLM-OSR variants tailored to CV and artificial intelligence applications in scientific domains, such as document image understanding [57] and material science [58].

## VI. CONCLUSION

The LLM-OSR algorithm shows significant potential in reconstructing spatial-temporal graph signals by combining GSP-based denoise handler with LLM-based prediction. Ex-

perimental results highlight the superior performance of LLM-OSR-4 in capturing spatial-temporal dependencies and it achieves high accuracy in signal reconstruction for traffic and weather datasets.

While the current performance of LLM-OSR is promising, significant work remains to fully unleash the capabilities of LLMs in spatial-temporal prediction to address the current limitations. The LLM-OSR could serve as a foundation to spark future studies, driving innovation and exploration in the intersection of large language models and dynamic graph signal prediction.

#### ACKNOWLEDGMENT

This work is supported by the Tsinghua Shenzhen International Graduate School Start-up fund under Grant QD2022024C, Shenzhen Science and Technology Innovation Commission under Grant JCYJ20220530143002005, and Shenzhen Ubiquitous Data Enabling Key Lab under Grant ZDSYS20220527171406015.

#### REFERENCES

- [1] J. N. Acosta, G. J. Falcone, P. Rajpurkar, and E. J. Topol, "Multimodal biomedical ai," *Nature Medicine*, vol. 28, no. 9, pp. 1773–1784, 2022.
- [2] G. Sonkavde, D. S. Dharrao, A. M. Bongale, S. T. Deokate, D. Doreswamy, and S. K. Bhat, "Forecasting stock market prices using machine learning and deep learning models: A systematic review, performance analysis and discussion of implications," *International Journal of Financial Studies*, vol. 11, no. 3, pp. 94, 2023.
- [3] A. Radford, J. Wu, D. Amodei, D. Amodei, J. Clark, M. Brundage, and I. Sutskever, "Better language models and their implications," *OpenAI blog*, vol. 1, no. 2, 2019.
- [4] J. Devlin, "Bert: Pre-training of deep bidirectional transformers for language understanding," *arXiv preprint arXiv:1810.04805*, 2018.
- [5] L. Floridi and M. Chiriatti, "Gpt-3: Its nature, scope, limits, and consequences," *Minds and Machines*, vol. 30, pp. 681–694, 2020.
- [6] E. Waisberg, J. Ong, M. Masalkhi, S. A. Kamran, N. Zaman, P. Sarker, A. G. Lee, and A. Tavakkoli, "Gpt-4: a new era of artificial intelligence in medicine," *Irish Journal of Medical Science (1971-)*, vol. 192, no. 6, pp. 3197–3200, 2023.
- [7] Y. Sun, S. Wang, S. Feng, S. Ding, C. Pang, J. Shang, J. Liu, X. Chen, Y. Zhao, Y. Lu, et al., "Ernie 3.0: Large-scale knowledge enhanced pre-training for language understanding and generation," *arXiv preprint arXiv:2107.02137*, 2021.
- [8] J. Chen, S. Li, Q. Huang, S. Yan, Z. Xie, and Y. Lu, "Application of kimi intelligent assistant in the teaching of water pollution control engineering course," *International Journal of Education and Humanities*, vol. 13, no. 3, pp. 39–43, 2024.
- [9] S. S. Sohail, F. Farhat, Y. Himeur, M. Nadeem, D. Ø. Madsen, Y. Singh, S. Atalla, and W. Mansoor, "Decoding chatgpt: a taxonomy of existing research, current challenges, and possible future directions," *Journal of King Saud University-Computer and Information Sciences*, p. 101675, 2023.
- [10] A. Navarro and F. Casacuberta, "Exploring multilingual pretrained machine translation models for interactive translation," in *Proceedings of Machine Translation Summit XIX, Vol. 2: Users Track*, 2023, pp. 132–142.
- [11] X. Deng, V. Bashlovkina, F. Han, S. Baumgartner, and M. Bendersky, "Llms to the moon? reddit market sentiment analysis with large language models," in *Companion Proceedings of the ACM Web Conference 2023*, 2023, pp. 1014–1019.
- [12] L. Tang, Z. Sun, B. Idnay, J. G. Nestor, A. Soroush, P. A. Elias, Z. Xu, Y. Ding, G. Durrett, J. F. Rousseau, et al., "Evaluating large language models on medical evidence summarization," *NPJ digital medicine*, vol. 6, no. 1, pp. 158, 2023.
- [13] F. Miraglia, F. Vecchio, C. Pappalettera, L. Nucci, M. Cotelli, E. Judica, F. Ferreri, and P. M. Rossini, "Brain connectivity and graph theory analysis in alzheimer's and parkinson's disease: the contribution of electrophysiological techniques," *Brain Sciences*, vol. 12, no. 3, pp. 402, 2022.
- [14] D. Qin and E. E. Kuruoglu, "Graph learning based financial market crash identification and prediction," in *IEEE CAI*, 2024.
- [15] M. Rostami, M. Oussalah, K. Berahmand, and V. Farrahi, "Community detection algorithms in healthcare applications: a systematic review," *IEEE Access*, vol. 11, pp. 30247–30272, 2023.
- [16] Y. Yan and E. E. Kuruoglu, "Binarized simplicial convolutional neural networks," *Neural Networks*, 2024.
- [17] J. Fan, W. Weng, H. Tian, H. Wu, F. Zhu, and J. Wu, "RGDAN: A random graph diffusion attention network for traffic prediction," *Neural Networks*, vol. 172, pp. 106093, 2024.
- [18] S. Xu, F. Wilhelm-Mauch, and W. Maass, "Quantum feature embeddings for graph neural networks," in *HICSS*, 2024, pp. 7633–7642.
- [19] B. Bevans, A. Ramalho, Z. Smoqi, A. Gaikwad, T. G. Santos, P. Rao, and J. Oliveira, "Monitoring and flaw detection during wire-based directed energy deposition using in-situ acoustic sensing and wavelet graph signal analysis," *Materials & Design*, vol. 225, pp. 111480, 2023.
- [20] R. Sharma and H. K. Meena, "Emerging trends in eeg signal processing: A systematic review," *SN Computer Science*, vol. 5, no. 4, pp. 1–14, 2024.
- [21] X. Dong, D. Thanou, L. Toni, M. Bronstein, and P. Frossard, "Graph signal processing for machine learning: A review and new perspectives," *IEEE Signal Processing Magazine*, vol. 37, no. 6, pp. 117–127, 2020.
- [22] T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," *ICLR*, 2017.
- [23] B. Yu, H. Yin, and Z. Zhu, "Spatio-temporal graph convolutional networks: A deep learning framework for traffic forecasting," *IJCAI*, 2018.
- [24] W. Wang and Q. Sun, "Robust adaptive estimation of graph signals based on welsch loss," *Symmetry*, vol. 14, no. 2, pp. 426, 2022.
- [25] Y. Yan, E. E. Kuruoglu, and M. A. Altinkaya, "Adaptive sign algorithm for graph signal processing," *Signal Processing*, vol. 200, pp. 108662, 2022.
- [26] B. Jin, G. Liu, C. Han, M. Jiang, H. Ji, and J. Han, "Large language models on graphs: A comprehensive survey," *IEEE Transactions on Knowledge and Data Engineering*, 2024.
- [27] J. Huang, X. Zhang, Q. Mei, and J. Ma, "Can llms effectively leverage graph structural information: when and why," *arXiv preprint arXiv:2309.16595*, 2023.
- [28] R. Ye, C. Zhang, R. Wang, S. Xu, Y. Zhang, et al., "Natural language is all a graph needs," *arXiv preprint arXiv:2308.07134*, vol. 4, no. 5, pp. 7, 2023.
- [29] 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.
- [30] N. Tremblay, P. Gonçalves, and P. Borgnat, "Design of graph filters and filterbanks," in *Cooperative and Graph Signal Processing*, pp. 299–324. Elsevier, 2018.
- [31] T. Brown, B. Mann, N. Ryder, M. Subbiah, J. D. Kaplan, P. Dhariwal, A. Neelakantan, P. Shyam, G. Sastry, A. Askell, S. Agarwal, A. Herbert-Voss, G. Krueger, T. Henighan, R. Child, A. Ramesh, D. Ziegler, J. Wu, C. Winter, C. Hesse, M. Chen, E. Sigler, M. Litwin, S. Gray, B. Chess, J. Clark, C. Berner, S. McCandlish, A. Radford, I. Sutskever, and D. Amodei, "Language models are few-shot learners," in *Advances in Neural Information Processing Systems*, H. Larochelle, M. Ranzato, R. Hadsell, M. Balcan, and H. Lin, Eds. 2020, vol. 33, pp. 1877–1901, Curran Associates, Inc.
- [32] J. Achiam, S. Adler, S. Agarwal, L. Ahmad, I. Akkaya, F. L. Aleman, D. Almeida, J. Altenschmidt, S. Altman, S. Anadkat, et al., "Gpt-4 technical report," *arXiv preprint arXiv:2303.08774*, 2023.
- [33] S. Bubeck, V. Chandrasekaran, R. Eldan, J. Gehrke, E. Horvitz, E. Kamar, P. Lee, Y. T. Lee, Y. Li, S. Lundberg, et al., "Sparks of artificial intelligence: Early experiments with gpt-4," *arXiv preprint arXiv:2303.12712*, 2023.
- [34] J. H. Giraldo, A. Mahmood, B. Garcia-Garcia, D. Thanou, and T. Bouwmans, "Reconstruction of time-varying graph signals via sobolev smoothness," *IEEE Transactions on Signal and Information Processing over Networks*, vol. 8, pp. 201–214, 2022.
- [35] G. Chen, X. Chen, C. Zheng, J. Wang, X. Liu, and Y. Han, "Spatiotemporal smoothing aggregation enhanced multi-scale residual deep graph convolutional networks for skeleton-based gait recognition," *Applied Intelligence*, pp. 1–21, 2024.
- [36] W. Bai, "Smoothness harmonic: A graph-based approach to reveal spatiotemporal patterns of cortical dynamics in fmri data," *Applied Sciences*, vol. 13, no. 12, 2023.

- [37] M. A. Qureshi and D. Greene, “Eve: explainable vector based embedding technique using wikipedia,” *Journal of Intelligent Information Systems*, vol. 53, pp. 137–165, 2019.
- [38] A. Grover and J. Leskovec, “node2vec: Scalable feature learning for networks,” in *SIGKDD*, 2016, pp. 855–864.
- [39] C. of Seattle, “Seattle loop detector data,” <https://github.com/zhiyongc/Seattle-Loop-Data>, 2020.
- [40] National Oceanic and Atmospheric Administration, “National oceanic and atmospheric administration (noaa) weather data,” <https://www.noaa.gov/>, 2024.
- [41] Y. Yan, R. Adel, and E. E. Kuruoglu, “Graph normalized-lmp algorithm for signal estimation under impulsive noise,” *Journal of Signal Processing Systems*, vol. 95, no. 1, pp. 25–36, 2023.
- [42] P. D. Lorenzo, S. Barbarossa, P. Banelli, and S. Sardellitti, “Adaptive least mean squares estimation of graph signals,” *IEEE Transactions on Signal and Information Processing over Networks*, vol. 2, no. 4, pp. 555 – 568, 2016.
- [43] M. J. M. Spelta and W. A. Martins, “Normalized lms algorithm and data-selective strategies for adaptive graph signal estimation,” *Signal Processing*, vol. 167, pp. 107326, 2020.
- [44] C. Peng, Y. Yan, and E. KURUOGLU, “Adaptive message passing sign algorithm,” in *Temporal Graph Learning Workshop @ NeurIPS 2023*, 2023.
- [45] E. Isufi, A. Loukas, N. Perraudin, and G. Leus, “Forecasting Time Series With VARMA Recursions on Graphs,” *IEEE Transactions on Signal Processing*, vol. 67, no. 18, pp. 4870–4885, 2019.
- [46] J. Hong, Y. Yan, E. E. Kuruoglu, and W. K. Chan, “Multivariate time series forecasting with GARCH models on graphs,” *IEEE Transactions on Signal and Information Processing over Networks*, vol. 9, pp. 557–568, 2023.
- [47] Y. Li, J. Keung, Z. Yang, X. Ma, J. Zhang, and S. Liu, “Simac: simulating agile collaboration to generate acceptance criteria in user story elaboration,” *Automated Software Engineering*, vol. 31, no. 2, pp. 55, 2024.
- [48] J. Jang, S. Ye, and M. Seo, “Can large language models truly understand prompts? a case study with negated prompts,” in *Transfer learning for natural language processing workshop*. PMLR, 2023, pp. 52–62.
- [49] F. Jia, K. Wang, Y. Zheng, D. Cao, and Y. Liu, “Gpt4mts: Prompt-based large language model for multimodal time-series forecasting,” in *Proceedings of the AAAI Conference on Artificial Intelligence*, 2024, vol. 38, pp. 23343–23351.
- [50] J. Chen, Y. Geng, Z. Chen, J. Z. Pan, Y. He, W. Zhang, I. Horrocks, and H. Chen, “Zero-shot and few-shot learning with knowledge graphs: A comprehensive survey,” *Proceedings of the IEEE*, vol. 111, no. 6, pp. 653–685, 2023.
- [51] A. Maharana, D.-H. Lee, S. Tulyakov, M. Bansal, F. Barbieri, and Y. Fang, “Evaluating very long-term conversational memory of llm agents,” in *Proceedings of the Annual Meeting of the Association for Computational Linguistics*, 2024, vol. 1, pp. 13851–13870.
- [52] S. Shahriar, B. D. Lund, N. R. Mannuru, M. A. Arshad, K. Hayawi, R. V. K. Bevara, A. Mannuru, and L. Batool, “Putting gpt-4o to the sword: A comprehensive evaluation of language, vision, speech, and multimodal proficiency,” *Applied Sciences*, vol. 14, no. 17, pp. 7782, 2024.
- [53] O. Karakus, E. E. Kuruoglu, and M. A. Altinkaya, “Modelling impulsive noise in indoor powerline communication systems,” *Signal, image and video processing*, vol. 14, no. 8, pp. 1655–1661, 2020.
- [54] W. Lee, H. S. Nam, J. Y. Seok, W.-Y. Oh, J. W. Kim, and H. Yoo, “Deep learning-based image enhancement in optical coherence tomography by exploiting interference fringe,” *Communications Biology*, vol. 6, no. 1, pp. 464, 2023.
- [55] D. Herranz, E. Kuruoğlu, and L. Toffolatti, “An alpha-stable approach to the study of the p (d) distribution of unresolved point sources in cmb sky maps,” *Astronomy & Astrophysics*, vol. 424, no. 3, pp. 1081–1096, 2004.
- [56] E. Kuruoglu, C. Molina, S. Godsill, and W. Fitzgerald, “A new analytic representation for the symmetric alpha-stable probability density function,” in *Proceedings of the 5th World Meeting of the International Society for Bayesian Analysis (ISBA)*. ASA: American Statistical Association, 1997, pp. 229–233.
- [57] E. E. Kuruoglu and A. S. Taylor, “Using annotations for summarizing a document image and itemizing the summary based on similar annotations,” May 4 2010, US Patent 7,712,028.
- [58] F. Saffarimiandoab, R. Mattesini, W. Fu, E. E. Kuruoglu, and X. Zhang, “Insights on features’ contribution to desalination dynamics and capacity of capacitive deionization through machine learning study,” *Desalination*, vol. 515, pp. 115197, 2021.