

Graph Signal Processing for Traffic Sensor Placement and Network-Wide Traffic Recovery: A Survey and Case Study On METR-LA

Sristi Yadav¹, Kumar Gaurav²

^{1,2}*Department of Electronics and Communication Engineering, Harcourt Butler Technical University (HBTU), Kanpur, India*

doi.org/10.64643/IJIRT209798

Abstract—Deciding where to place a limited number of traffic sensors, and how to rebuild speeds on unobserved roads from what those sensors report, is a long-standing problem in intelligent transportation systems. Classical sensor location models focus on flow observability and origin–destination estimation, while recent data-driven models forecast traffic with deep neural networks on road graphs. Graph signal processing (GSP) sits between these views: it treats traffic speeds as a signal on the road graph and offers sampling theorems, sampling-set selection algorithms and reconstruction methods with explicit guarantees. This paper surveys GSP-based sensor placement and network-wide traffic recovery. We review the sensing and data background, the foundations of graph sampling theory, experimental-design and submodular approaches to sensor selection, and the main families of recovery methods: bandlimited projection, Laplacian-regularized, kernel-based, sparse and learned. We organize the literature in a taxonomy and compare the families by assumptions, guarantees, computational cost and robustness. A case study on the 207-sensor METR-LA highway graph, using greedy E-optimal placement on synthetic bandlimited signals, illustrates how placement, bandlimit and recovery method interact, and shows that the best recovery method depends on graph topology. We close with open problems, namely validation on real non-stationary traffic, dynamic graphs, graph learning, sensor failures, heterogeneous sensing and standard benchmarks, and give practical guidelines for agencies.

Index Terms—E-optimal design, Graph sampling, Graph signal processing, Intelligent transportation systems, Laplacian regularization, METR-LA, Sensor placement, Survey, Traffic state recovery

I. INTRODUCTION

Road agencies see what is happening on their network through a limited number of fixed sensors. Inductive

loop detectors, radar units, cameras and Bluetooth readers feed speed advisory boards, incident detection, ramp metering and travel-time services [1], [2]. Every sensor has to be bought, installed, powered, calibrated and maintained, and on freeways installation can require lane closures. Almost no agency can instrument every road segment, so almost every agency has to answer two practical questions: where should the next sensor go, and how should the speeds on unmeasured segments be estimated from the measured ones?

Both questions have been studied in transportation engineering for more than two decades. The classical sensor location literature asks which links must be observed so that all link flows, or an origin–destination (OD) matrix, can be inferred [3]– [5]. These models are exact and rest on flow conservation, but they say little about how well the *value* of an unmeasured link can be predicted from its neighbors, because they do not use the spatial correlation of the traffic state itself. A second line of work, driven by public data sets such as METR-LA, learns this correlation with graph neural networks and forecasts traffic on road graphs [6]– [9]. Those models are accurate when they are given all sensors and long histories. They are not designed to decide which sensors to buy, and they give no guarantee for nodes that were never observed.

Graph signal processing (GSP) offers a third view [10]– [12]. The road network is a graph, the sensors are nodes, and the speeds measured at one time instant form a signal on the nodes. Neighboring road segments tend to have similar speeds, so the signal is smooth on the graph: its energy sits mainly in the low-frequency eigenvectors of the graph Laplacian. A signal with this property is called approximately

bandlimited, and a bandlimited signal can be recovered from a small, well-chosen subset of nodes [13]– [15]. In one framework this gives a data model, a criterion for choosing sensor locations, a family of reconstruction algorithms and, under stated assumptions, error bounds.

The ingredients are well developed, but they are scattered. Sampling theory appears in signal processing journals, sensor placement in transportation and control venues, and traffic recovery in intelligent transportation systems (ITS) and machine learning venues. A transportation engineer who wants to know whether GSP is worth trying, or a signal processing researcher who wants to know what traffic data demand from a sampling algorithm, has no single place to look. This survey is meant to be that place.

The contributions of this paper are as follows.

- 1) A unified formulation of traffic sensor placement and network-wide recovery as a graph sampling and reconstruction problem, with notation that lets methods from different communities be compared.
- 2) A taxonomy (Fig. 2) that organizes the literature into classical sensor location, experimental-design and submodular selection, graph sampling-set selection, and signal recovery, with a comparison of assumptions, guarantees and cost.
- 3) A worked case study on the 207-sensor METR-LA graph with greedy E-optimal placement, taken from the authors' companion study [16], used to show how the recovery method, the bandlimit and the sensor count interact.
- 4) A critical discussion of open problems and a set of practical guidelines for agencies.

Scope. We concentrate on fixed-location sensors that measure a scalar traffic state, mostly speed, on a road graph, and on methods that either choose the sensing set or reconstruct the unobserved nodes. We do not review traffic forecasting in depth, for which good surveys exist [9], [17], and we discuss flow-conservation and OD models only where they contrast with the graph-signal view. The coverage is representative rather than exhaustive.

Organization. Section II gives the sensing, data and GSP background. Section III formulates the problem and presents the taxonomy. Sections IV to VII review classical sensor location, experimental-design selection, graph sampling theory and recovery methods. Section VIII discusses what is specific to traffic. Section IX presents the case study, Section X compares the families, Section XI lists open problems, Section XII gives practical guidelines, and Section XIII concludes.

II. BACKGROUND

A. Traffic sensing

Fixed roadside sensors remain the main source of continuous traffic measurements on freeways. Table I summarizes the common modalities. Loop detectors are the long-standing reference for point speed and occupancy, but they require pavement cutting and are exposed to wear and failure [1]. Non-intrusive radar and video units reduce the installation burden, and Bluetooth readers or GPS probes provide travel times without dedicated roadside hardware at each location [2]. Different modalities have different noise, latency and cost, so a placement method that treats all sensors as identical is making a simplification that should be stated.

Table I. Common traffic sensing modalities (qualitative summary).

Modality	Typical measurement	Strengths	Limitations
Inductive loop	Count, occupancy, point speed at a fixed location	Mature technology; high time resolution	Pavement cutting and lane closure for installation; wear and failures
Radar / microwave	Speed and count, often several lanes	Non-intrusive; side-fire mounting	Occlusion in dense traffic; calibration needs
Video camera	Count, speed, classification, incident verification	Rich information per site	Lighting and weather sensitivity; processing cost; privacy
Bluetooth / Wi-Fi reader	Travel time between reader pairs	Low unit cost; direct travel time	Only a sample of vehicles; pairing errors

Modality	Typical measurement	Strengths	Limitations
Probe / floating car (GPS)	Trajectories and route speeds	No roadside hardware; wide coverage	Variable penetration; latency; privacy

Fixed sensors and probes are complementary. Fixed sensors give dense time series at a few places, whereas probes give sparse samples along many routes. State-estimation work has shown how to combine the two inside a flow model [2], [18]. In this survey we treat each sensor as returning one scalar speed per time step, which is the setting of the public benchmarks.

B. Data sets

Two data sets dominate the GSP and deep-learning traffic literature, and both come from the Caltrans Performance Measurement System (PeMS) [19]. METR-LA contains speeds from 207 loop detectors on Los Angeles County highways at 5-minute intervals over four months in 2012, and PEMS-BAY contains 325 sensors in the San Francisco Bay Area [6] (Table II). Both ship with an adjacency matrix computed from road-network distances between sensors, using a thresholded Gaussian kernel (see (5) below).

Table II. Widely used highway speed benchmarks.

Data set	Source	Sensors	Interval	Period
METR-LA	Loop detectors, Los Angeles County highways (PeMS)	207	5 min	March–June 2012 (about four months)
PEMS-BAY	Loop detectors, San Francisco Bay Area (PeMS)	325	5 min	January–May 2017 (about five months)

These data sets are convenient but imperfect. Raw series contain missing values coded as zeros, sensor drift and outages. The distance-based graph is a modeling choice, not a measurement: it ignores direction of travel, ramps and capacity. Any conclusion about sensor placement drawn from these graphs therefore inherits the quality of the graph, an issue we return to in Section VIII-B.

C. Graph signal processing fundamentals

Let $G = (V, E, W)$ be a weighted undirected graph with $N = |V|$ nodes and symmetric weight matrix W . Let D

$= \text{diag}(W1)$ be the degree matrix. The combinatorial Laplacian and its eigendecomposition are $L = D - W = U \Lambda U^T$, $0 = \lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_N$, (1) where the columns of U are the graph Fourier basis. The second eigenvalue λ_2 is the algebraic connectivity, and it is small for long, thin graphs such as a freeway corridor [20]–[22].

A graph signal $x \in \mathbb{R}^N$ assigns one value, here a speed, to each node. Its graph Fourier transform is

$$\hat{x} = U^T x, \quad (2)$$

and its smoothness is measured by the quadratic form $x^T L x = \frac{1}{2} \sum_{ij} W_{ij} (x_i - x_j)^2 = \sum_k \lambda_k \hat{x}_k^2$. (3)

Equation (3) shows why smooth signals are low-pass: a signal whose neighboring values are close has a small quadratic form, so its energy sits at small λ_k [10], [11]. A signal is K -bandlimited if $\hat{x}_k = 0$ for $k > K$, that is, if

$$x = U_{:K} \alpha, \quad U_{:K} \in \mathbb{R}^{N \times K}, \quad (4)$$

where $U_{:K}$ holds the K eigenvectors with smallest eigenvalues.

For road networks the weights are usually built from road distance d_{ij} with the thresholded Gaussian kernel introduced for DCRNN [6],

$$W_{ij} = \exp(-d_{ij}^2 / \sigma^2) \text{ if this value } \geq \kappa, \text{ and } 0 \text{ otherwise,} \quad (5)$$

where σ^2 is the variance of the pairwise distances and κ is a sparsification threshold. Measurements are taken at a sampling set $S \subseteq V$ with $|S| = m$, giving the model

$$y = C x + \varepsilon, \quad C \in \{0,1\}^{m \times N}, \quad \varepsilon \sim \mathcal{N}(0, \sigma_\varepsilon^2 I), \quad (6)$$

where C selects the rows indexed by S . For a K -bandlimited signal, x is determined by α , and the least-squares estimate is $\hat{x} = U_{:K} (C U_{:K})^\dagger y$, where † denotes the pseudoinverse.

This is well posed only if $C U_{:K}$ has full column rank, which requires $m \geq K$ and a suitable S [13], [14]. The estimation error covariance is $\sigma_\varepsilon^2 (U_{:K}^T C^T C U_{:K})^{-1}$, so the choice of S directly controls how much noise is amplified.

This observation turns sensor placement into an experimental-design problem, which we develop in Section V.

III. PROBLEM FORMULATION AND TAXONOMY

A. Problem statement

Given a graph G , a signal model and a budget of m sensors, the placement problem is to select the set S that minimizes the expected reconstruction error over the unobserved nodes,

$$S^* = \arg \min_{|S| = m} \mathbb{E} \|x - \hat{x}(C, y)\|^2, \quad (7)$$

where the expectation is over the noise and over the distribution of signals. Practical variants add a cost for each candidate location, require that existing sensors

be retained, or ask for a placement that performs well across several traffic regimes. The combinatorial nature of (7), with $C(N, m)$ candidate sets, is why almost all methods in this survey are greedy or relaxed. Placement and recovery are coupled. A placement is only optimal with respect to a recovery method and a signal model, and a recovery method is only reliable at placements that make its linear system well conditioned. Surveying them together, as in Fig. 1, is therefore more informative than surveying them separately.

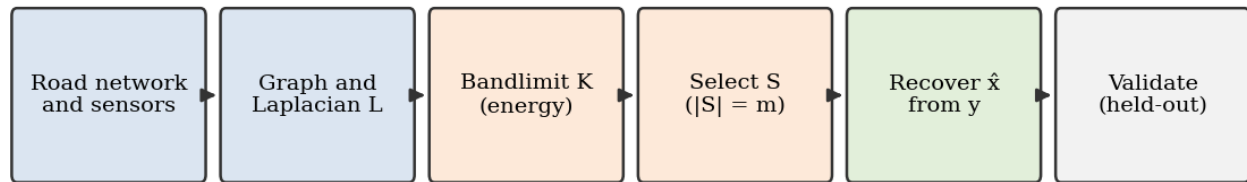


Fig. 1. Generic GSP pipeline for sensor placement and recovery.

B. Taxonomy

We organize the literature along two axes. The first is *what is selected*: the sensing set S . The second is *what is estimated*: the full signal x from y . Within each axis

we separate approaches by the assumption they place on the signal and the guarantee they offer. Fig. 2 shows the resulting taxonomy, and Sections IV to VII follow its branches.

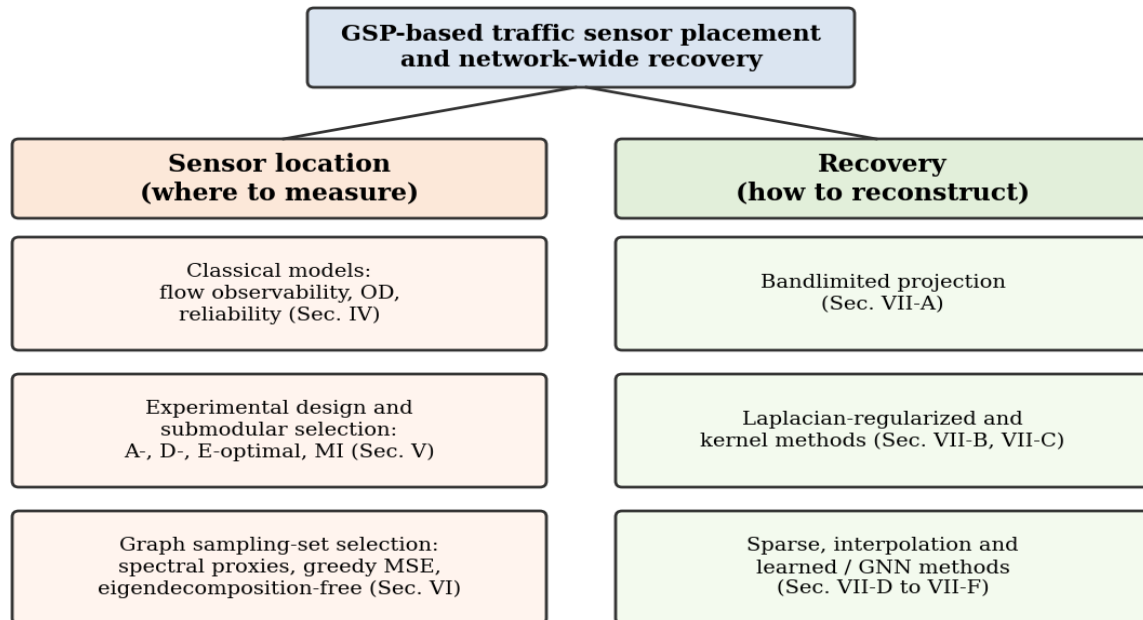


Fig. 2. Taxonomy of the literature surveyed, with section pointers

IV. CLASSICAL TRAFFIC SENSOR LOCATION

A. Flow observability and OD estimation

The first generation of sensor location models was driven by OD estimation. Yang and Zhou proposed

rules for placing traffic counters so that an OD matrix can be estimated well, based on how much of the network flow and how many OD pairs each location captures [3]. Later work reformulated the problem in terms of link-flow inference: choose a set of observed

links so that the flows on all other links are determined by flow conservation and a small set of additional assumptions [4]. Reviews of the field catalog the problem variants, such as counting versus identification sensors, link versus path observability, and static versus dynamic settings [5]. Mixed-integer programming is the standard solution tool, and reliability-oriented variants account for sensor failures [23]. These models answer a well-defined question: *which links must be measured so that flows elsewhere are logically implied?* The answers are exact, they are grounded in the physics of conservation, and they are useful when counts or OD flows are the target quantity.

B. Traffic state and travel-time estimation

A second group of works asks how sensors should be placed for estimating the traffic *state* (density or speed), usually with a dynamic model. The cell transmission model [24] and the underlying kinematic wave theory [25] describe how density evolves along a freeway, and Kalman-type estimators combine this model with measurements [26]. Studies that incorporate probe measurements into such estimators show that the value of a sensor depends on where it is relative to bottlenecks and ramps [18]. A survey of highway state estimation covers these methods in detail [2].

C. Limits of the classical view

Three limits motivate the GSP approach. First, observability models guarantee that flows can be *inferred*, not how accurately speeds can be *predicted*: they carry no notion of reconstruction error and no use of the spatial smoothness that real data show. Second, model-based estimators need a calibrated dynamic model for each section. Third, the combinatorial programs can be heavy at metropolitan scale. GSP does not replace the classical view, because it does not enforce conservation laws, but it gives a complementary, data-agnostic answer when the target is the speed field.

V. EXPERIMENTAL DESIGN AND SUBMODULAR SELECTION

A. Optimal design criteria

For the bandlimited model, Section II-C showed that the estimator covariance is $\sigma^2 \varepsilon^2 M(S)^{-1}$, where the information matrix is

$$M(S) = U_K^T C^T C U_K = \sum_{v \in S} u_{\cdot v} u_{\cdot v}^T, \quad (8)$$

and $u_{\cdot v}$ is the v -th row of U_K written as a column vector. Classical optimal design theory defines a family of scalar criteria on $M(S)$ [27]. A-optimality minimizes the trace of M^{-1} (average error variance), D-optimality maximizes $\log \det M$ (volume of the confidence ellipsoid), and E-optimality maximizes $\lambda_{\min}(M)$, which minimizes the largest eigenvalue of M^{-1} and thus the worst-case direction of error. For a noise vector ε , the coefficient error satisfies $\|\hat{a} - a\| \leq \|\varepsilon\| / \sqrt{\lambda_{\min}(M)}$, so E-optimality is the criterion that directly controls this worst-case bound. Each criterion has an interpretation in traffic terms: A controls the average speed error, D the joint uncertainty, and E the least well observed spatial pattern.

B. Greedy selection and submodularity

Exact optimization of any of these criteria is combinatorial, so the standard approach is greedy: starting from an empty set, add the node that most improves the criterion. If the objective is monotone and submodular, greedy selection is guaranteed to reach at least $(1 - 1/e)$ of the optimum [28]. The log-determinant criterion is submodular, which makes greedy D-optimal selection attractive and has been exploited for sensor selection [29]. The mean-square error and $\lambda_{\min}$ criteria are not submodular in general. For the mean-square error of graph signals, Chamon and Ribeiro showed that greedy selection still comes with a near-optimality guarantee that depends on how far the cost is from submodular [30]. For E-optimal greedy selection we are not aware of a comparable general guarantee, so its quality is supported mainly by empirical evidence and by the minimax interpretation above. This is a point where the literature and some claims in applied papers can diverge, and it should be stated carefully.

C. Convex relaxation

A second route is to relax the binary selection vector to the interval $[0, 1]$, solve a convex program, and round. Joshi and Boyd formulated sensor selection for linear measurements in this way, with log-determinant and related objectives [31]. Ranieri et al. studied sensor placement for linear inverse problems and showed that the frame potential, a surrogate for the mean-square error, is submodular and that greedy minimization of it is near optimal [32]. Convex relaxations are attractive when side constraints are

present, for example budget limits per corridor, because constraints can be added directly.

D. Gaussian-process and information-theoretic criteria

If the signal is modeled as a Gaussian process with covariance K , placement can maximize the mutual information between selected and unselected locations. Krause et al. showed that this objective is submodular under mild conditions, so greedy placement is near optimal, and that it is more robust than variance-only criteria [33]. There is a direct connection to GSP: a covariance of the form $K = f(L)$, a function of the Laplacian, is a graph kernel, and the associated posterior mean is the kernel regression estimator used for graph signal recovery [34], [35]. In other words, Gaussian-process placement and bandlimited placement are two points on a continuum of priors.

E. Data-driven bases

The bandlimited model uses a basis U_K obtained from the graph. When historical data are available, a basis can instead be learned from them, for example by proper orthogonal decomposition, and sensors selected by pivoted QR factorization [36]. The formal problem is the same, with a data-driven basis in place of U_K . The trade-off is that a data-driven basis needs training data and is tied to the regimes in those data, whereas the graph basis needs only the network. For newly built roads, or for a network with little history, the graph basis is the only option, and for mature networks the two can be compared.

VI. GRAPH SAMPLING THEORY AND SAMPLING-SET SELECTION

A. Sampling theorems

Pesenson introduced Paley–Wiener spaces on combinatorial graphs and gave conditions under which a set of nodes uniquely determines a bandlimited function [13]. Anis et al. extended these ideas to

practical sampling sets. They defined a spectral proxy that bounds the cut-off frequency of the largest bandlimited space a given set can uniquely identify, namely

$$\Omega_k(S) = \min_{\phi \neq 0 \text{ supported on } V \setminus S} \left(\frac{\|L^k \phi\|}{\|\phi\|} \right)^{1/k}, \quad (9)$$

and proposed to select S to maximize $\Omega_k(S)$ without computing the full eigendecomposition [14], [37]. Chen et al. developed a sampling theory for graph signals under a general shift operator, including the construction of interpolation operators and conditions for perfect recovery [15], and derived fundamental limits of sampling strategies in the presence of noise [38]. Tsitsvero et al. connected sampling to an uncertainty principle on graphs, using localization operators to characterize which sets of nodes and frequencies can be simultaneously concentrated [39]. A review gives a unified view of these results together with sampling of graph signals in the vertex and spectral domains [40].

Unlike classical sampling, graph sampling has no shift invariance and therefore no single Nyquist rate: the minimum number of samples depends on the specific graph and on the chosen bandlimit. One consequence is that results obtained on one network do not transfer automatically to another, a point we return to in Section IX.

B. Selection algorithms

Table III lists representative selection algorithms. They differ in the criterion they optimize, in whether they need the eigendecomposition of the Laplacian, and in the strength of their guarantees. Marques et al. proposed sampling with successive local aggregations, where a single node collects values from neighbors over several steps, which is relevant if a communication network among sensors exists [41]. Segarra et al. considered reconstruction through percolation from seed nodes [42]. For active learning on graphs, Gadde et al. used sampling theory to choose which nodes to label [43], a problem mathematically equivalent to choosing where to place a sensor.

Table III. Representative sampling-set selection approaches.

Approach	Criterion	Eigen-decomposition	Remarks
Spectral proxies [14]	Maximize $\Omega_k(S)$ (9)	Not required	Fast; bound on identifiable bandwidth
Greedy MSE [30]	Mean-square reconstruction error	Used for basis	Near-optimality guarantee tied to submodularity gap

Approach	Criterion	Eigen-decomposition	Remarks
Fundamental limits [38]	Noise-aware recovery bounds	Used for analysis	Characterizes best possible performance
Uncertainty principle [39]	Localization operators	Used	Links sampling with vertex–frequency concentration
Eigendecomposition-free [44]	Polynomial approximation of the band	Not required	Scales to large graphs
Gershgorin disc alignment [45]	Lower bound on smallest eigenvalue	Not required	Fast; targets the E-type worst-case bound
Greedy A/D/E-optimal (Sec. V)	Functions of $M(S)$	Used for U_K	Simple; D has $(1 - 1/e)$ guarantee, A and E empirical

C. Scalability

A dense eigendecomposition costs $O(N^3)$, which is manageable for a few hundred sensors but not for metropolitan networks with tens of thousands of nodes.

Three remedies are used. First, only the K lowest eigenpairs are needed and can be obtained iteratively, for example with Lanczos methods. Second, polynomial approximations of the Laplacian replace explicit eigenvectors, which is the idea behind eigendecomposition-free selection [44].

Third, bounds based on Gershgorin discs avoid eigenvalue computation entirely [45]. The choice depends on N and K , and for the 207-node benchmark the explicit approach is entirely adequate.

D. Noise and model mismatch

Real traffic speeds are not exactly bandlimited. Write $x = U_K \alpha + r$, where r collects the out-of-band content. Then with $y = C x + \varepsilon$ and the projection estimate $\hat{x} = U_K (C U_K)^\dagger y$, the error satisfies $\|x - \hat{x}\| \leq \|r\| + (\|r\| + \|\varepsilon\|) / \sigma_{\min}(C U_K)$. (10) Equation (10) shows three effects.

The first term is model bias, which no placement can remove. The second is amplification of both the out-of-band residual and the noise by the inverse of the smallest singular value of $C U_K$, which is exactly what E-optimal placement tries to make large. Raising K reduces the residual but makes $C U_K$ harder to condition and needs more samples, so there is a bias–variance trade-off governed by K . This trade-off is central to the sensitivity results in Section IX.

VII. SIGNAL RECOVERY METHODS

A. Bandlimited projection

Projection onto the K -bandlimited subspace, $\hat{x} = U_K (C U_K)^\dagger y$, is the most direct approach when the signal model holds and $C U_K$ is well conditioned [13], [15]. It has the lowest computational cost of the methods surveyed and exact recovery in the noiseless bandlimited case. Its weakness, shown in (10), is sensitivity to ill-conditioning and out-of-band content.

B. Laplacian-regularized recovery

A smoothness prior leads to the regularized estimator $\hat{x} = \arg \min_{\text{over } x} \|y - C x\|^2 + \mu x^T L x = (C^T C + \mu L)^{-1} C^T y$. (11)

As $\mu \rightarrow 0$ with noiseless data, (11) becomes harmonic interpolation, which minimizes $x^T L x$ subject to $x_S = y$, a formulation known from semi-supervised learning [46], [47] and graph interpolation [48], [49]. An important practical property is that $C^T C + \mu L$ is positive definite whenever the graph is connected and at least one node is sampled, because the null space of L is the constant vector. The estimator therefore does not require $C U_K$ to be well conditioned or the bandlimit K to be chosen, and the penalty μ trades fidelity to the data against smoothness. On a sparse, long-corridor graph with a small Fiedler value this robustness can matter a great deal, as the case study shows.

C. Kernel-based recovery

Kernel regression generalizes (11) by using the posterior mean of a Gaussian process with a graph kernel $K = f(L)$,

$$\hat{x} = K C^T (C K C^T + \sigma^2 I)^{-1} y, \quad (12)$$

where f is chosen so that high graph frequencies are attenuated [34], [35]. The Laplacian-regularized estimator is the limiting case in which the prior precision is proportional to L , so the two are closely related. Kernel methods add flexibility, since f can be tuned, and a Bayesian reading, since the posterior covariance gives uncertainty at unobserved nodes. Their quality depends on the kernel hyperparameters, and poor tuning can give results far from the Laplacian-regularized baseline even though the two share the same prior form.

D. Sparse and filter-based models

A different prior writes the signal as the output of a low-pass graph filter H applied to a sparse activation vector δ , so that $x = H\delta$, and recovers δ by sparse regression. This models traffic as the diffusion of a few localized events, such as incidents or bottlenecks, and is related to compressed sensing [50]. It can represent non-smooth structure better than a purely bandlimited model but is sensitive to the choice of filter and regularization.

E. Local interpolation

The simplest baseline assigns each unsensed node the weighted average of its sensed neighbors. It needs no

spectral information and costs very little. It fails when a node has no sensed neighbor, which is common at low sampling ratios on sparse corridors, and a faithful implementation must define a fallback, for example iterating the averaging to propagate information [42], [48].

F. Learned recovery

Learning-based methods do not assume a spectral model. Matrix and tensor completion exploit low-rank structure in traffic data and are widely used for imputation [51], [52]. Graph neural networks learn spatial dependence from data: spatio-temporal architectures such as DCRNN, STGCN and Graph WaveNet were designed for forecasting but are also used for imputation [6]–[8], [53]–[55]. Inductive graph neural networks for kriging estimate the signal at sensor locations that were never seen during training [56], which is exactly the task of reconstructing unobserved nodes. The key caveat is that graph convolutions of the type introduced in [57], [58] are themselves low-pass graph filters, so these models also rely on the smoothness assumption, but they learn the filter from data instead of fixing it. Table IV compares the recovery families.

Table IV. Recovery families compared.

Family	Prior on signal	Training data	Main failure mode	Cost
Bandlimited projection	Exactly K-bandlimited	Not needed	Ill-conditioned $C U_K$; out-of-band content	Low
Laplacian-regularized	Smooth on graph	Not needed	Over-smoothing at bottlenecks; choice of μ	Linear solve, sparse
Kernel regression	Gaussian process, $f(L)$	Tuning only	Poor kernel hyperparameters	Solve of size m
Sparse / filter-based	Filtered sparse events	Tuning only	Sensitive to filter and λ	Optimization
Local interpolation	Local averaging	Not needed	No sensed neighbor; slow propagation	Very low
Learned (GNN, completion)	Learned from data	Required	Needs history; weak guarantees; domain shift	High to train

VIII. TRAFFIC-SPECIFIC CONSIDERATIONS

A. Non-stationarity and congestion

A bandlimited model is most plausible in free-flow conditions, when speeds vary slowly along a corridor. Congestion breaks this: queues form at bottlenecks and spread upstream as kinematic waves [25], creating

sharp speed gradients that carry energy at high graph frequencies. A placement tuned to the low-frequency subspace may then do poorly exactly when the operator needs the estimate most. This is not a defect specific to one method; it is a mismatch between a global spectral prior and a regime-switching process. Remedies include graph wavelets that localize in

vertex and frequency [59], regime-specific bases, and placements optimized across an ensemble of regimes. Direction is a second issue. Freeway speeds in opposite directions are different signals, and information in congested traffic travels against the flow. Undirected Laplacians discard this asymmetry, while GSP on directed graphs [11], [60] retains it at the cost of a less convenient spectral theory.

B. Graph construction

The benchmark graph is built from road distance with the kernel in (5) [6]. Placement conclusions depend on it. If the graph links segments that are physically close but separated by a barrier or a ramp that decouples their speeds, the smoothness assumption is wrong and the selection will be misled. Graph learning methods estimate the graph from observed signals under a smoothness prior [61], [62], and several traffic models learn an adaptive adjacency during training [8], [54], [55].

Using a learned graph for placement is attractive but circular, because the learning needs data from the sensors that are yet to be placed. A sensible procedure is to learn the graph from the existing sensors, place new ones, and re-learn.

C. The temporal dimension

Sensors are fixed but traffic changes by the minute and by the season. Joint time-vertex models treat the data as a signal on the product of the road graph and a time axis [63], and stationarity models allow spectral statistics to be estimated from samples [64]. A snapshot placement ignores temporal correlation, which could reduce the number of sensors needed. The converse risk is that a placement based on one regime generalizes badly to another, so evaluation should span peak, off-peak and incident periods.

D. Missing data and failures

Real deployments lose data. Imputation methods fill gaps using temporal and spatial neighbors [51], [52], while reliability-oriented sensor location models plan for failures [23]. In the GSP setting a sensor failure is the removal of a row from C , and a placement is robust if the remaining set still gives a well-conditioned information matrix. Choosing S to maximize $\lambda_{\min}(M)$ is, in effect, a way to maximize the margin against such losses, but explicit robust criteria remain an open topic.

E. GSP versus deep learning

Deep spatio-temporal models achieve strong forecasting accuracy on METR-LA [6], [8], [9]. They need long histories, do not by themselves decide where to place sensors and give limited theoretical guarantees for unobserved nodes. GSP placement uses only the graph, can be applied to a network with no history, and gives interpretable criteria. The two are complementary: a GSP-selected set can serve as the observed node set for a learned model, and the learned model can in turn supply the realistic signal ensembles that GSP placement lacks. Evaluating this combination is an open research question.

IX. CASE STUDY: E-OPTIMAL PLACEMENT ON THE METR-LA GRAPH

To show how the pieces of Sections V to VII interact on a real network graph, we summarize the authors' companion study [16]. All numbers in this section are quoted from that study. The aim here is to illustrate the behavior of the methods, not to claim a new benchmark.

A. Setup

The graph is the DCRNN adjacency of METR-LA with $N = 207$ nodes, built with the kernel of (5) and $\kappa = 0.1$, which the study reports as having 1,313 undirected edges and an algebraic connectivity of $\lambda_2 = 0.0023$. The bandlimit $K = 15$ was chosen from the cumulative spectral energy of a set of signal profiles (95.3% in the first 15 modes) and from a gap between the 15th and 16th Laplacian eigenvalues. The measurement noise is Gaussian with $\sigma_\varepsilon = 2$ km/h. Sensors are selected with the greedy E-optimal rule of Section V-A, that is, at each step the node that maximizes $\lambda_{\min}$ of the information matrix is added. Evaluation uses 100 signal realizations, with a random-placement baseline averaged over 30 trials. Five recovery methods are compared: Laplacian-regularized least squares (LS), kernel ridge regression (KRR), a low-pass-filtered sparse delta model (LPFD-A), bandlimited sample projection (SP-L), and unconstrained local interpolation (UI).

An important caveat is that the signals in this study are *synthetic* $K = 15$ bandlimited profiles generated from the model of (4) plus noise, not recorded METR-LA time series. The study states this as its main limitation.

The results therefore show what the methods do when the model assumption holds, and they should not be read as performance on real traffic.

B. Results

With $m = 66$ sensors, 31.9% of the nodes, LS recovery gives an RMSE of 1.667 km/h and $R^2 = 0.9513$ over all 207 nodes. The random-placement baseline gives 1.985 ± 0.229 km/h, so E-optimal placement reduces RMSE by about 16% and the maximum node error from 12.333 to 5.572 km/h (Table V).

Table V. Reconstruction at $m = 66$ sensors (values from the companion study [16]).

Placement method	and	RMSE (km/h)	R^2	Max node error (km/h)
E-optimal, recovery	LS	1.667	0.9513	5.572
Random (30-trial mean), LS recovery		1.985 ± 0.229	0.9221	12.333

Fig. 3 compares the five recovery methods at the same E-optimal set. LS is lower in RMSE than the best alternative (SP-L, 33.478 km/h) by a factor above 20, and the other methods give errors between 33 and 100 km/h, with negative R^2 . The study attributes this to conditioning: it reports that the condition number of C_{U_K} is about 180 at $m = 66$ even under E-optimal placement, and much larger under random placement, so methods that invert it amplify noise. This agrees with (10) and with the positive-definiteness argument of Section VII-B.

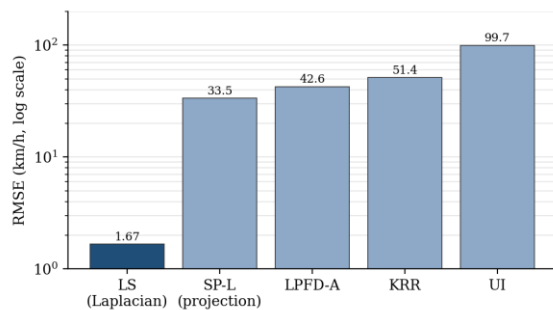


Fig. 3. RMSE of five recovery methods at $m = 66$ E-optimal sensors (log scale), plotted from the values reported in [16].

Two points deserve care when reading Fig. 3. First, errors of 33 to 100 km/h on a signal whose standard deviation is about 12 km/h are larger than a constant

guess would give, which points to an implementation or tuning issue in the weaker baselines (for example nodes without sensed neighbors in UI, or an untuned kernel in KRR) rather than a property of the method families. Section VII-C indicates that a well-tuned graph kernel should approach the Laplacian-regularized estimator. Second, the large gap is specific to this sparse long-corridor graph at this sampling ratio, as the authors note, and a denser graph can give a different ranking.

Table VI and Fig. 4 show the effect of the bandlimit. Lowering K to 10 makes RMSE worse, 2.431 km/h, because the model under-represents the signal, while raising K to 20 or 25 improves RMSE modestly but requires 84 or 103 sensors. This is the bias-variance trade-off of (10), seen in terms of sensor cost.

Table VI. Sensitivity to the assumed bandlimit K (values from [16]).

K	Sensors m^*	RMSE (km/h)	R^2
10	42	2.431	0.8912
15	66	1.667	0.9513
20	84	1.421	0.9620
25	103	1.298	0.9688

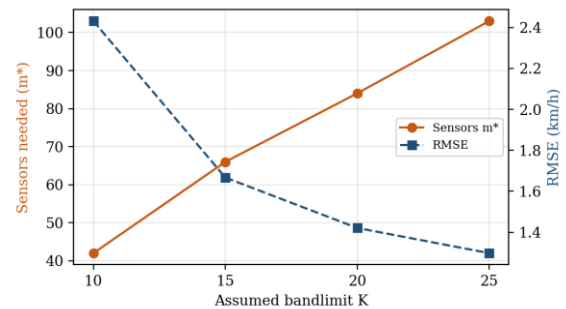


Fig. 4. Required sensors and RMSE as a function of the assumed bandlimit K , plotted from [16].

Finally, the study reports results by freeway corridor. E-optimal placement is not uniform: of the 66 selected sensors, 31 (47%) fall on the I-405/I-5 corridor that holds 35% of the nodes, and only 4 falls on I-10, which holds 16% of the nodes (Fig. 5). The corridor with the largest gain from optimization over random placement is I-405/I-5, at 24.4% lower RMSE. The pattern is consistent with the idea that selection follows the structure of the graph, concentrating sensors where the low-frequency eigenvectors have most of their local variation.

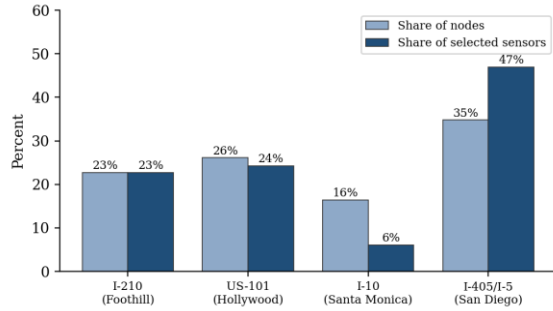


Fig. 5. Share of nodes and share of selected sensors by corridor, computed from the counts in [16].

C. What the case study shows, and what it does not

The case study demonstrates that the pipeline of Fig. 1 can be run end to end on a real benchmark graph in a fraction of a second, that greedy E-optimal selection beats random selection under the bandlimited model, and that the choice of recovery method and bandlimit changes the conclusion more than the placement does. It also illustrates a general point: the method that wins depends on conditioning and therefore on topology, which is why reports should include graph statistics such as mean degree, clustering coefficient and diameter, and not only error metrics.

Several limits should be kept in mind. The test signals lie in the same subspace for which the placement was optimized, which favors the method by construction. The criteria used to argue that $m^* = 66$ is a practical

optimum (spectral gap, RMSE knee, conditioning, cost per unit accuracy, marginal information gain and coverage at ± 2 km/h) are largely different views of the same RMSE curve and the same conditioning of $C_{U,K}$, so they are better read as mutually consistent indicators than as independent confirmations. Cost figures depend on assumed unit prices that are agency specific. The most useful next step, which the study itself names, is validation on recorded METR-LA series across peak, off-peak and incident conditions. Related work on a dense city-road graph has reported different method rankings [65], which supports the topology-dependence argument.

X. COMPARATIVE DISCUSSION

Table VII summarizes the paradigms by signal model, guarantee, data needs, scalability and main weakness. No paradigm dominates. Classical observability models give strong logical guarantees about flows but none about the accuracy of speeds. Gaussian-process criteria give submodular guarantees but need a covariance, which in practice comes from historical data or a graph kernel. Bandlimited selection needs only the graph and gives interpretable conditioning-based criteria, but its guarantees rely on the bandlimit assumption, which real congested traffic violates. Learned approaches fit real data best but need history and offer weak guarantees at unobserved locations.

Table VII. Comparison of sensor placement and recovery paradigms.

Paradigm	Signal model	Guarantee	Data need	Scalability	Main weakness
Flow observability / OD [3], [4]	Flow conservation	Logical observability of flows	Network only	MIP; heavy at scale	No speed-accuracy guarantee
GP / mutual information [33]	Gaussian process	$(1 - 1/e)$ under conditions	Covariance	Moderate	Needs a covariance model
Bandlimited A/D/E greedy [27], [30]	K-bandlimited	Bound via $\sigma_{\min}$; D submodular	Graph only	$O(N^3)$ eigen	Real traffic not bandlimited
Spectral-proxy / eigen-free [14], [44], [45]	Bandlimited	Identifiable-bandwidth bounds	Graph only	Near-linear	Looser than exact criteria
Learned / kriging [56]	Learned	Mostly empirical	Long history	Training cost	Domain shift; no guarantee

Choosing a pipeline is largely a question of what is known. If only the road network is available, bandlimited selection with Laplacian-regularized

recovery is a sound first choice. If history is available, a data-driven basis [36] or a graph learned from data [61] should be compared against the graph basis. If the

network is very large, eigendecomposition-free selection is needed. In all cases, a held-out evaluation on recorded data is the only reliable test.

XI. OPEN PROBLEMS

Validation on real, non-stationary data.

Most GSP placement studies, including the case study above, use synthetic bandlimited signals or static snapshots. The central open task is to evaluate placements on recorded series covering peak, off-peak, incident and adverse-weather conditions, and to report how the optimal sensor count changes with regime.

Beyond bandlimited models.

Congestion produces piecewise-smooth speed fields with sharp transitions. Models based on graph wavelets [59], regime switching, or sparse-plus-smooth decompositions may capture these better, but there are few sampling guarantees for them.

Dynamic and directed graphs.

Closures, incidents and changing demand alter the effective graph, and information in traffic flows in a direction. Placement methods that are robust to edge deletions and weight perturbations, and sampling theory for directed graphs [60], are needed.

Graph learning for placement.

The quality of the graph limits the quality of the placement. Learning the graph from the existing sensors [61], [62], quantifying the uncertainty in the graph, and propagating that uncertainty into the placement criterion are largely unexplored.

Robustness to failures and drift.

Sensor outages and miscalibration are routine. Criteria that explicitly maximize performance under random sensor loss [23] would connect the experimental-design view with reliability engineering.

Heterogeneous and mobile sensing.

Real systems combine loops, radar, cameras and probe vehicles with different noise levels and costs [2]. Selection with heterogeneous noise and per-sensor cost, and joint design of fixed and mobile sensing, remain open.

Scalability and online operation.

Eigendecomposition-free methods exist [44], [45], but evaluating them on metropolitan-scale road graphs with realistic constraints is still rare. Online updating of the spectral basis as the graph changes is a related need.

Costs and fairness of coverage. Cost models in the literature use assumed unit prices, while actual cost depends on site conditions. Because optimal placement concentrates sensors where the signal varies most, an agency may also need constraints that guarantee minimum coverage in every area, so that the benefit of monitoring is not distributed unevenly.

Benchmarks and reporting. Studies differ in graph, signal model, noise level and metrics, which makes comparison hard. We recommend that future papers report topology statistics, state whether signals are real or synthetic, give the noise level, use multiple random seeds, include a random-placement baseline and at least one tuned graph-kernel baseline, and release code and graphs.

XII. PRACTICAL GUIDELINES

The following sequence summarizes how an agency or researcher can apply the methods in this survey.

- 1) Build the graph from road distances or connectivity, check that it is connected, and compute basic statistics such as mean degree, clustering coefficient, diameter and algebraic connectivity.
- 2) Test the smoothness assumption on historical data from existing sensors: compute the fraction of signal energy in the lowest graph frequencies for several regimes before choosing a bandlimit K .
- 3) Choose K using held-out data rather than only the energy criterion, and examine how the required sensor count changes around the chosen value, as in Fig. 4.
- 4) Select the sensing set with greedy A-, D- or E-optimal selection when N is in the hundreds or low thousands, and with eigendecomposition-free methods for larger networks. Always compare with random placement.
- 5) Recover the unobserved nodes first with a Laplacian-regularized estimator, then compare kernel and learned methods. Use projection only if the conditioning of $C U_K$ is verified.

- 6) Validate by hiding real sensors, not only synthetic nodes, and report errors by corridor and by traffic regime.
- 7) Plan redundancy for sensor failures, and re-run the placement when the network or the graph changes.
- 8) Use the agency's own cost data, not generic unit prices, when comparing deployment tiers.

XIII. CONCLUSION

Graph signal processing links three questions that the transportation literature has mostly treated separately: how smooth traffic is over a road network, which sensors capture that smoothness best, and how to reconstruct unobserved segments from them. This survey organized the work on these questions into a taxonomy that runs from classical flow-observability models through optimal-design and submodular selection to graph sampling theory and reconstruction methods, and compared the families by assumptions, guarantees, data needs and cost.

Three messages stand out. First, placement and recovery cannot be judged separately, because the conditioning of the sampled spectral subspace couples them. Second, the best recovery method depends on graph topology and sampling ratio, so studies should report graph statistics along with error metrics. Third, the main weakness of current evidence is that bandlimited models are rarely tested on recorded, non-stationary traffic. The case study on METR-LA illustrates all three points and also shows where the evidence is thin. Closing the gap with real-data validation, dynamic graphs, robust criteria and standard benchmarks is the most valuable direction for future work.

REFERENCES

- [1] L. A. Klein, M. K. Mills, and D. R. P. Gibson, *Traffic Detector Handbook*, 3rd ed., vol. I, FHWA-HRT-06-108. Federal Highway Administration, 2006.
- [2] T. Seo, A. M. Bayen, T. Kusakabe, and Y. Asakura, "Traffic state estimation on highway: A comprehensive survey," *Annual Reviews in Control*, vol. 43, pp. 128–151, 2017.
- [3] H. Yang and J. Zhou, "Optimal traffic counting locations for origin–destination matrix estimation," *Transportation Research Part B*, vol. 32, no. 2, pp. 109–126, 1998.
- [4] S.-R. Hu, S. Peeta, and C.-H. Chu, "Identification of vehicle sensor locations for link-based network traffic applications," *Transportation Research Part B*, vol. 43, nos. 8–9, pp. 873–894, 2009.
- [5] M. Gentili and P. B. Mirchandani, "Locating sensors on traffic networks: Models, challenges and research opportunities," *Transportation Research Part C*, vol. 24, pp. 227–255, 2012.
- [6] Y. Li, R. Yu, C. Shahabi, and Y. Liu, "Diffusion convolutional recurrent neural network: Data-driven traffic forecasting," in *Proc. Int. Conf. Learning Representations (ICLR)*, 2018.
- [7] B. Yu, H. Yin, and Z. Zhu, "Spatio-temporal graph convolutional networks: A deep learning framework for traffic forecasting," in *Proc. Int. Joint Conf. Artificial Intelligence (IJCAI)*, 2018, pp. 3634–3640.
- [8] Z. Wu, S. Pan, G. Long, J. Jiang, and C. Zhang, "Graph WaveNet for deep spatial-temporal graph modeling," in *Proc. Int. Joint Conf. Artificial Intelligence (IJCAI)*, 2019, pp. 1907–1913.
- [9] W. Jiang and J. Luo, "Graph neural network for traffic forecasting: A survey," *Expert Systems with Applications*, vol. 207, Art. no. 117921, 2022.
- [10] D. I. Shuman, S. K. Narang, P. Frossard, A. Ortega, and P. Vandergheynst, "The emerging field of signal processing on graphs," *IEEE Signal Processing Magazine*, vol. 30, no. 3, pp. 83–98, 2013.
- [11] A. Sandryhaila and J. M. F. Moura, "Discrete signal processing on graphs," *IEEE Transactions on Signal Processing*, vol. 61, no. 7, pp. 1644–1656, 2013.
- [12] 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.
- [13] I. Pesenson, "Sampling in Paley–Wiener spaces on combinatorial graphs," *Transactions of the American Mathematical Society*, vol. 360, no. 10, pp. 5603–5627, 2008.
- [14] A. Anis, A. Gadde, and A. Ortega, "Efficient sampling set selection for bandlimited graph signals using graph spectral proxies," *IEEE*

- Transactions on Signal Processing, vol. 64, no. 14, pp. 3775–3789, 2016.
- [15] S. Chen, R. Varma, A. Sandryhaila, and J. Kovačević, “Discrete signal processing on graphs: Sampling theory,” IEEE Transactions on Signal Processing, vol. 63, no. 24, pp. 6510–6523, 2015.
- [16] [Authors], “Optimal sensor placement in urban traffic networks using graph signal processing,” ICTACT Journal, manuscript, 2026.
- [17] E. I. Vlahogianni, M. G. Karlaftis, and J. C. Golias, “Short-term traffic forecasting: Where we are and where we’re going,” Transportation Research Part C, vol. 43, pp. 3–19, 2014.
- [18] J. C. Herrera and A. M. Bayen, “Incorporation of Lagrangian measurements in freeway traffic state estimation,” Transportation Research Part B, vol. 44, no. 4, pp. 460–481, 2010.
- [19] C. Chen, K. Petty, A. Skabardonis, P. Varaiya, and Z. Jia, “Freeway performance measurement system: Mining loop detector data,” Transportation Research Record, no. 1748, pp. 96–102, 2001.
- [20] M. Fiedler, “Algebraic connectivity of graphs,” Czechoslovak Mathematical Journal, vol. 23, no. 2, pp. 298–305, 1973.
- [21] F. R. K. Chung, Spectral Graph Theory. Providence, RI, USA: American Mathematical Society, 1997.
- [22] U. von Luxburg, “A tutorial on spectral clustering,” Statistics and Computing, vol. 17, no. 4, pp. 395–416, 2007.
- [23] X. Li and Y. Ouyang, “Reliable sensor deployment for network traffic surveillance,” Transportation Research Part B, vol. 45, no. 1, pp. 218–231, 2011.
- [24] C. F. Daganzo, “The cell transmission model: A dynamic representation of highway traffic consistent with the hydrodynamic theory,” Transportation Research Part B, vol. 28, no. 4, pp. 269–287, 1994.
- [25] M. J. Lighthill and G. B. Whitham, “On kinematic waves II: A theory of traffic flow on long crowded roads,” Proceedings of the Royal Society of London A, vol. 229, no. 1178, pp. 317–345, 1955.
- [26] Y. Wang and M. Papageorgiou, “Real-time freeway traffic state estimation based on extended Kalman filter: A general approach,” Transportation Research Part B, vol. 39, no. 2, pp. 141–167, 2005.
- [27] F. Pukelsheim, Optimal Design of Experiments. New York, NY, USA: Wiley, 1993. Reprinted by SIAM, 2006.
- [28] G. L. Nemhauser, L. A. Wolsey, and M. L. Fisher, “An analysis of approximations for maximizing submodular set functions—I,” Mathematical Programming, vol. 14, pp. 265–294, 1978.
- [29] M. Shamaiah, S. Banerjee, and H. Vikalo, “Greedy sensor selection: Leveraging submodularity,” in Proc. IEEE Conf. Decision and Control, 2010, pp. 2572–2577.
- [30] L. F. O. Chamon and A. Ribeiro, “Greedy sampling of graph signals,” IEEE Transactions on Signal Processing, vol. 66, no. 1, pp. 34–47, 2018.
- [31] S. Joshi and S. Boyd, “Sensor selection via convex optimization,” IEEE Transactions on Signal Processing, vol. 57, no. 2, pp. 451–462, 2009.
- [32] J. Ranieri, A. Chebira, and M. Vetterli, “Near-optimal sensor placement for linear inverse problems,” IEEE Transactions on Signal Processing, vol. 62, no. 5, pp. 1135–1146, 2014.
- [33] A. Krause, A. Singh, and C. Guestrin, “Near-optimal sensor placements in Gaussian processes: Theory, efficient algorithms and empirical studies,” Journal of Machine Learning Research, vol. 9, pp. 235–284, 2008.
- [34] A. J. Smola and R. Kondor, “Kernels and regularization on graphs,” in Proc. Conf. Learning Theory (COLT), 2003, pp. 144–158.
- [35] D. Romero, M. Ma, and G. B. Giannakis, “Kernel-based reconstruction of graph signals,” IEEE Transactions on Signal Processing, vol. 65, no. 3, pp. 764–778, 2017.
- [36] K. Manohar, B. W. Brunton, J. N. Kutz, and S. L. Brunton, “Data-driven sparse sensor placement for reconstruction,” IEEE Control Systems Magazine, vol. 38, no. 3, pp. 63–86, 2018.
- [37] A. Anis, A. Gadde, and A. Ortega, “Towards a sampling theorem for signals on arbitrary graphs,” in Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP), 2014, pp. 3864–3868.
- [38] S. Chen, R. Varma, A. Singh, and J. Kovačević, “Signal recovery on graphs: Fundamental limits of sampling strategies,” IEEE Transactions on

- Signal and Information Processing over Networks, vol. 2, no. 4, pp. 539–554, 2016.
- [39] D. Tsitsvero, S. Barbarossa, and P. di Lorenzo, “Signals on graphs: Uncertainty principle and sampling,” *IEEE Transactions on Signal Processing*, vol. 64, no. 18, pp. 4845–4860, 2016.
- [40] Y. Tanaka, Y. C. Eldar, A. Ortega, and G. Cheung, “Sampling signals on graphs: From theory to applications,” *IEEE Signal Processing Magazine*, vol. 37, no. 6, pp. 14–30, 2020.
- [41] A. G. Marques, S. Segarra, G. Leus, and A. Ribeiro, “Sampling of graph signals with successive local aggregations,” *IEEE Transactions on Signal Processing*, vol. 64, no. 7, pp. 1832–1843, 2016.
- [42] S. Segarra, A. G. Marques, G. Leus, and A. Ribeiro, “Reconstruction of graph signals through percolation from seeding nodes,” *IEEE Transactions on Signal Processing*, vol. 64, no. 16, pp. 4363–4378, 2016.
- [43] A. Gadde, A. Anis, and A. Ortega, “Active semi-supervised learning using sampling theory for graph signals,” in *Proc. ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining*, 2014, pp. 492–501.
- [44] A. Sakiyama, Y. Tanaka, T. Tanaka, and A. Ortega, “Eigendecomposition-free sampling set selection for graph signals,” *IEEE Transactions on Signal Processing*, vol. 67, no. 10, pp. 2679–2692, 2019.
- [45] Y. Bai, F. Wang, G. Cheung, Y. Nakatsukasa, and W. Gao, “Fast graph sampling set selection using Gershgorin disc alignment,” *IEEE Transactions on Signal Processing*, vol. 68, pp. 2419–2434, 2020.
- [46] X. Zhu, Z. Ghahramani, and J. Lafferty, “Semi-supervised learning using Gaussian fields and harmonic functions,” in *Proc. Int. Conf. Machine Learning (ICML)*, 2003, pp. 912–919.
- [47] M. Belkin, P. Niyogi, and V. Sindhwani, “Manifold regularization: A geometric framework for learning from labeled and unlabeled examples,” *Journal of Machine Learning Research*, vol. 7, pp. 2399–2434, 2006.
- [48] S. K. Narang, A. Gadde, and A. Ortega, “Signal processing techniques for interpolation in graph structured data,” in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, 2013, pp. 5445–5449.
- [49] S. Chen, A. Sandryhaila, J. M. F. Moura, and J. Kovačević, “Signal recovery on graphs: Variation minimization,” *IEEE Transactions on Signal Processing*, vol. 63, no. 17, pp. 4609–4624, 2015.
- [50] D. L. Donoho, “Compressed sensing,” *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.
- [51] L. Li, Y. Li, and Z. Li, “Efficient missing data imputing for traffic flow by considering temporal and spatial dependence,” *Transportation Research Part C*, vol. 34, pp. 108–120, 2013.
- [52] M. T. Asif, N. Mitrovic, L. Garg, J. Dauwels, and P. Jaillet, “Low-dimensional models for missing data imputation in road networks,” in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, 2013, pp. 3527–3531.
- [53] L. Zhao et al., “T-GCN: A temporal graph convolutional network for traffic prediction,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 9, pp. 3848–3858, 2020.
- [54] L. Bai, L. Yao, C. Li, X. Wang, and C. Wang, “Adaptive graph convolutional recurrent network for traffic forecasting,” in *Proc. Advances in Neural Information Processing Systems (NeurIPS)*, 2020, pp. 17804–17815.
- [55] Z. Wu, S. Pan, G. Long, J. Jiang, X. Chang, and C. Zhang, “Connecting the dots: Multivariate time series forecasting with graph neural networks,” in *Proc. ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining*, 2020, pp. 753–763.
- [56] Y. Wu, D. Zhuang, A. Labbe, and L. Sun, “Inductive graph neural networks for spatiotemporal kriging,” in *Proc. AAAI Conf. Artificial Intelligence*, vol. 35, no. 5, 2021, pp. 4478–4485.
- [57] M. Defferrard, X. Bresson, and P. Vandergheynst, “Convolutional neural networks on graphs with fast localized spectral filtering,” in *Proc. Advances in Neural Information Processing Systems (NeurIPS)*, 2016, pp. 3844–3852.
- [58] T. N. Kipf and M. Welling, “Semi-supervised classification with graph convolutional networks,” in *Proc. Int. Conf. Learning Representations (ICLR)*, 2017.
- [59] D. K. Hammond, P. Vandergheynst, and R. Gribonval, “Wavelets on graphs via spectral graph theory,” *Applied and Computational Harmonic Analysis*, vol. 30, no. 2, pp. 129–150, 2011.

- [60] A. Sandryhaila and J. M. F. Moura, "Discrete signal processing on graphs: Frequency analysis," *IEEE Transactions on Signal Processing*, vol. 62, no. 12, pp. 3042–3054, 2014.
- [61] X. Dong, D. Thanou, P. Frossard, and P. Vandergheynst, "Learning Laplacian matrix in smooth graph signal representations," *IEEE Transactions on Signal Processing*, vol. 64, no. 23, pp. 6160–6173, 2016.
- [62] V. Kalofolias, "How to learn a graph from smooth signals," in *Proc. Int. Conf. Artificial Intelligence and Statistics (AISTATS)*, 2016, pp. 920–929.
- [63] 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 Transactions on Signal Processing*, vol. 66, no. 3, pp. 817–829, 2018.
- [64] N. Perraudin and P. Vandergheynst, "Stationary signal processing on graphs," *IEEE Transactions on Signal Processing*, vol. 65, no. 13, pp. 3462–3477, 2017.
- [65] R. A. Martínez Márquez and G. Patané, "Recovery of traffic information through graph signal processing," *EURASIP Journal on Advances in Signal Processing*, vol. 2025, Art. no. 11, 2025.