

Name of the Participating Student: Isabella Yin

High School: Tsinghua International School

Province: Beijing

Country/Region: China

Name of Advisor(s): Laura P. Schaposnik

Affiliation of Advisor(s): Professor at the University of Illinois Chicago, Visiting
Professor at the University of Oxford

Title of Research Report: Movement Stability of D-Optimal Sensor Placement Under
Projector Perturbations

2026 S.-T. Yau High School Science Award
仅用于2026丘成桐中学科学奖论文公示

Movement Stability of D-Optimal Sensor Placement Under Projector Perturbations

Isabella Yin

Tsinghua International School

Advisor: Laura P. Schaposnik, Professor at the University of Illinois Chicago, Visiting Professor at the University of Oxford

Abstract

In time-varying D-optimal subset selection, the number of coordinates being replaced is not reflective of sensors' actual physical movement; when the optimal selection changes for a set of mobile sensors, a single one may have to move across the entire graph. Thus, we instead evaluate movement according to a bottleneck matching distance on a fixed mobility graph and define a distance margin (a minimum previous objective loss over configurations lying more than a chosen radius away from the current maximizer). Our main result shows that, if the projector drift is small and the distance margin at a chosen radius exceeds a certain threshold, then any maximizer of the new objective must lie within that radius, without requiring any tie-breaking rules or uniform conditioning assumptions over the subsets. However, additional information on competing configurations is needed to guarantee this. A weighted-path construction shows that an arbitrarily small perturbation at a fixed eigengap can move an optimal sensor from one end of the graph to the other, so any bound depending only on perturbation size and eigengap alone cannot suffice. We also present a finite certificate, conditions under which local and Voronoi restrictions preserve the optimum, and oversampled and regularized extensions, with an application to learned dictionaries. Furthermore, we end with experiments on synthetic graph families and real-world traffic datasets (PEMS-BAY and METR-LA) to test and demonstrate the certified movement bound. In our experiments, across all 26,553 real-world transitions satisfying theorem conditions, every optimizer movement lies within the certified radius, and some cases show that the certificate can be nontrivial and sharp.

Keywords: Projector perturbation, D-optimal design, Subset selection, Graph signal sampling, Bottleneck matching

Contents

1	Introduction	4
1.1	Related Work	4
1.2	Our Contributions	5
2	Setting	5
2.1	Notation	7
3	Movement Stability Under Projector Perturbations	8
4	Laplacian Perturbations and Why Projector Drift Alone Is Not Enough	10
4.1	Perturbation Size and Eigengap Alone Do Not Bound Movement	11
5	Computing the Distance Margin and Certifying Movement	13
5.1	Formulation	13
5.2	Computing the Distance Margin by Enumeration	14
6	Losslessness of P-Hop and Voronoi Restrictions	14
6.1	Intersection with the Voronoi Region	15
7	Extensions	16
7.1	Oversampled D-Optimality, $m \geq K$	16
7.2	Regularized D-Optimality	17
7.3	Dictionaries	18
8	Experiments on Synthetic Graphs	18
8.1	Protocol and Setup	18
8.2	Certifying Optimizer Movement on Synthetic Graphs	19
8.3	Distance Margin Curve	20
9	Real-World Experiments on Traffic Sensor Data	21
9.1	Datasets and Preprocessing	21
9.2	Empirical Traffic Projectors	21
9.3	Certification Protocol	22
9.4	Certification Results	22
9.5	A Nontrivial METR-LA Example	23
10	Discussion	24
11	Conclusion	25
	References	26
	Acknowledgements and Research Process	28

1. Introduction

Place a single sensor on a path of n vertices to sample a smooth signal, and assume the two endpoints are equally good. An arbitrarily small change in the edge weights can break the tie, forcing the optimal sensor from one end of the path to the other. Only one selected coordinate is replaced, but the corresponding sensor has to travel $n - 1$ hops. The number of replacements and the distance traveled are thus two different quantities, where only the second reflects a physical cost.

This difference is relevant in settings where the selected coordinates have buoys, robots, or mobile monitors attached. A newly computed global optimum might entail that a device must cross a large part of the domain, even if only one selected vertex has changed. Algorithms for dynamic placement thus often restrict attention to a local search region for each sensor. Nomura et al. update a learned dictionary, and restrict each sensor to a P -hop neighborhood intersected with a Voronoi cell determined by the current sensor positions [25]. Their experiments on synthetic signals and sea surface temperature data show improved reconstruction error relative to certain static placement baselines. Such a local restriction is practical and doable, but it raises the question that our paper aims to address: when is a local P -hop restriction guaranteed to retain a global optimum?

Answering this question requires separating four related and similar, but easily mixed, ideas and statements:

- (a) Set replacement: how many selected vertices change (measured by the symmetric-difference size $|S \Delta S'|$).
- (b) Physical movement: how far each sensor must travel, measured using a bottleneck matching distance.
- (c) Feasibility of a local search region or adjustment (whether a global optimum of the new objective belongs to the local region of the previous state).
- (d) Success of a particular algorithm: whether a greedy or distributed update finds that close optimum.

Our paper reaches results on (b) and (c), but not (d). We do not formally claim anything about whether the independent per-sensor greedy updates in Algorithm 2 of Nomura et al. [25] solve the centralized D-optimal problem; we discuss this separately in Section 6.

1.1 Related Work

Graph signal sampling: Sampling theory establishes which vertex subsets determine a band-limited graph signal in a low-frequency subspace [1, 7, 23, 27, 32] (see Shuman et al. [29] and Ortega et al. [26] for surveys). Graphical designs provide a related quadrature viewpoint for low-frequency graph functions [3, 30]. These results tell us which static sets are informative.

D-optimal design and subset selection: The D-optimal design problem in matrix sampling maximizes a log-determinant criterion, leading to the selection of informative subsets of rows, and is prevalent in problems of sensor selection [17, 19], maximum-volume submatrix selection [9], and randomized numerical linear algebra [2]. Combinatorial versions of these maximization problems are often solved using greedy and related approximation strategies. Instead of analyzing the approximation achieved by a particular algorithm, we provide theoretical guarantees on how far a maximizing configuration can move when its projector changes.

Subspace and projector perturbation: The Davis–Kahan $\sin \Theta$ theorem [10] and its further refinements [4, 18, 31, 33] describe how an invariant subspace changes under symmetric perturbations. Here, we establish a stability result under subspace perturbation for the orthogonal projector onto the subspace, rather than directly by the invariant subspace itself. We thus provide a stability result directly in terms of orthogonal projectors, which also applies to non-spectral projectors such as those arising from learned dictionaries.

Dynamic sensor placement: Analyses of adaptive and dynamic sampling strategies have also addressed real-time graph signal streams [11]. Nomura et al. recently proposed a local-update strategy [25], which directly motivates our question about local movement.

Gap in existing work: Recent work has defined margin-based certificates for top- m selection in graph-centrality problems, showing that a large enough score margin can guarantee selected sets remain unchanged under perturbation [14]. Unlike in our problem, those objectives are coordinatewise separable, whereas the D-optimal objective is a nonseparable set function. Our margin is instead defined in terms of distance in the underlying mobility graph and is used to certify a movement radius rather than only invariance of the top- m set. Replacement counts like $|S_t^* \Delta S_{t+1}^*|$ are therefore not enough for certain settings, since they record only how many selected vertices change and not how far they move.

1.2 Our Contributions

1. Movement bound under projector perturbation.

We derive a stability theorem that bounds the movement of every new maximizer using projector drift, conditioning only at the current optimizer, and a distance margin.

2. Why a distance margin is needed.

A weighted-path construction shows that no uniform movement bound over weighted graphs can depend only on perturbation size and eigengap while tending to zero as the perturbation tends to zero.

3. Computable certification.

We show how to compute the distance margin exactly on finite graphs and test the movement constraint using perfect matchings.

4. Local losslessness and extensions.

We show when P -hop restrictions intersected with Voronoi regions preserve a global optimum, and extend the stability result to oversampled and regularized settings, while showing that learned dictionary projectors fall directly within the general theorem.

5. Experiments and applications

We evaluate the certificate on synthetic graphs and real-world traffic data, illustrating its validity and sharpness under certain cases.

Section 2 provides background; Sections 3–4 establish the stability and impossibility results; Sections 5–7 present certification, local restrictions, and extensions. Section 8 presents synthetic experiments on the certificate and the distance margin curve, along with Section 9 with real-world experiments.

2. Setting

Movement is measured on a fixed graph throughout the analysis.

Assumption 2.1 (Fixed mobility graph). Let $G_{\text{mv}} = (V, E_{\text{mv}})$ be a fixed connected unweighted graph on the vertex set V with $|V| = n$, and let $d(u, v) := \min_{\pi: u \leadsto v} |\pi|$ denote its shortest-path distance, where π ranges over paths in G_{mv} and $|\pi|$ is the number of edges. All distances, P -hop neighborhoods, Voronoi cells, and matching distances in this paper refer to G_{mv} .

Separating G_{mv} from the object that models the signal is important: G_{mv} encodes where a sensor can physically travel, whereas the signal model could potentially be a weighted graph, a spectral subspace, or a learned dictionary. In the graph-signal application of Sections 4 and 6 we take G_{mv} to be the support graph of the weighted graph, and we assume that the support is fixed while the edge weights may change. The impossibility construction lies in this same setting, where only weights change.

Remark 2.2 (What the movement metric encodes versus does not). The bottleneck matching distance of Definition 2.5 treats sensors as interchangeable (only asks whether the old and new positions can be matched with small displacement). It doesn't separate sensor types or account for trajectory collisions, travel times, or edge-capacity constraints. These limitations motivate trajectory-related extensions discussed in Section 11 for future work.

Fix integers K and m with $1 \leq K < n$ and $1 \leq m \leq n$. For $s \in \{t, t+1\}$, let $\Pi_s \in \mathbb{R}^{n \times n}$ be an orthogonal projector of rank K , and choose $U_s \in \mathbb{R}^{n \times K}$ with orthonormal columns spanning its range, such that $\Pi_s = U_s U_s^\top$. Set $\rho_t := \|\Pi_t - \Pi_{t+1}\|_2$.

A configuration is a subset $S \subseteq V$ with $|S| = m$. Let $R_S \in \mathbb{R}^{m \times n}$ be the row-selector matrix and $D_S = R_S^\top R_S$ be the associated binary diagonal matrix.

Undersampled case $m \leq K$: for $s \in \{t, t+1\}$ define $C_s(S) := R_S \Pi_s R_S^\top \in \mathbb{R}^{m \times m}$, and $F_s(S) := \log \det C_s(S)$, with $F_s(S) = -\infty$ when $C_s(S)$ is singular. Maximizing F_s selects rows of the subspace basis according to the D-optimal volume criterion [7, 28]. For $m < K$, this is a row-volume criterion; when $m = K$, the compressed-projector and information-matrix determinants agree.

Oversampled case $m \geq K$: for $m > K$, the matrix $C_s(S)$ has rank at most $K < m$, so its determinant is zero and its log-determinant is $-\infty$. Thus, for $s \in \{t, t+1\}$ define instead $H_s(S) := U_s^\top D_S U_s \in \mathbb{R}^{K \times K}$, and $F_s^{\text{ov}}(S) := \log \det H_s(S)$, again with value $-\infty$ when $H_s(S)$ is singular.

At $m = K$ when the two determinant objectives agree, the undersampled perturbation estimate will be sharper and we use it in that boundary case.

Definition 2.3 (Maximizers). Write $\arg \max F_t$ for the set of configurations maximizing F_t over all m -subsets; it is nonempty because there is a finite possible family. A reference maximizer $S_t^* \in \arg \max F_t$ is any fixed element of this set, relative to which all margins are defined. No tie-breaking is needed at time $t+1$; the results hold for every $S \in \arg \max F_{t+1}$.

Lemma 2.4 (Nondegeneracy in undersampled case). *If $m \leq K$ then some m -subset S has $\det C_t(S) > 0$; consequently $d_t^* := \det C_t(S_t^*) = e^{F_t(S_t^*)} > 0$ for every reference maximizer.*

Proof. Since U_t has rank $K \geq m$, it contains m linearly independent rows, indexed by some S . Then $U_{t,S} \in \mathbb{R}^{m \times K}$ has rank m and $C_t(S) = U_{t,S} U_{t,S}^\top$ is positive definite. Any maximizer thus has objective value at least $F_t(S) > -\infty$. $\square$

Definition 2.5 (Bottleneck matching distance). For m -sets $S = \{s_1, \dots, s_m\}$ and $S' = \{s'_1, \dots, s'_m\}$, put

$$d_{\text{match}}(S, S') := \min_{\pi: \{1, \dots, m\} \rightarrow \{1, \dots, m\} \text{ bijective}} \max_{1 \leq i \leq m} d(s_i, s'_{\pi(i)}) \quad (1)$$

with d as the distance of the fixed movement graph G_{mv} . This is the smallest radius within which every old sensor can be matched bijectively to a new sensor. The value is independent of the ordering chosen for S and S' .

Proposition 2.6 (Metric property). d_{match} is a metric on the family of m -subsets of V .

Proof. The minimum in (1) is attained because there are only finitely many bijections between two m -element sets. Nonnegativity and symmetry thus follow directly from the corresponding properties of d . If $d_{\text{match}}(S, S') = 0$, then some bijection pairs each s_i with a vertex at graph distance zero (hence with s_i itself) so $S = S'$. Conversely, $d_{\text{match}}(S, S) = 0$. For the triangle inequality, let $S = \{s_i\}$, $T = \{t_j\}$, and $R = \{r_k\}$, and choose bijections π and σ attaining $d_{\text{match}}(S, T)$ and $d_{\text{match}}(T, R)$, respectively. Composing gives us

$$d(s_i, r_{\sigma(\pi(i))}) \leq d(s_i, t_{\pi(i)}) + d(t_{\pi(i)}, r_{\sigma(\pi(i))}) \leq d_{\text{match}}(S, T) + d_{\text{match}}(T, R)$$

for every i . Taking the maximum over i and then the minimum over bijections from S to R gives us the claim. $\square$

Definition 2.7 (distance margin). Fix a reference maximizer $S_t^* \in \arg \max F_t$. For an integer $P \geq 0$ define

$$M_t(P) := \min_{\substack{|S|=m \\ d_{\text{match}}(S, S_t^*) > P}} [F_t(S_t^*) - F_t(S)], \quad (2)$$

with convention $M_t(P) = +\infty$ when the minimum is over an empty family. Equivalently, $M_t(P)$ is the loss, under the old objective, of the best configuration that lies more than P hops from S_t^* .

The same empty-family convention applies to M_t^{ov} and M_t^τ introduced later, so the corresponding stability statements hold when their feasible families are empty.

2.1 Notation

The stability result depends on three main quantities.

Projector drift measures the perturbation of the underlying low-dimensional subspace, the distance margin measures the gap between the optimal configuration and distant competitors, and perturbation budget measures the objective change caused by projector drift.

Table 1: Main notation.

Symbol	Meaning
V, n	vertex set and number of vertices, $n = V $
K, m	projector rank and number of selected coordinates
S	configuration, i.e. an m -subset of V
G_{mv}, d	fixed mobility graph and its shortest-path distance
P	nonnegative integer movement radius, measured in graph hops
Π_t, U_t	rank- K orthogonal projector at time t and an orthonormal basis of its range
ρ_t	projector drift $\ \Pi_t - \Pi_{t+1}\ _2$
R_S, D_S	row-selector matrix and associated diagonal selector $D_S = R_S^\top R_S$
$C_t(S)$	compressed projector $(\Pi_t)_{SS}$, used in the undersampled case
$H_t(S)$	information matrix $U_t^\top D_S U_t$, used in the oversampled case
F_t, F_t^{ov}	undersampled and oversampled log-determinant objectives
S_t^*	fixed reference maximizer in $\arg \max F_t$
d_t^*, λ_t^*	determinant and minimum eigenvalue of $C_t(S_t^*)$
d_{match}	bottleneck matching distance between configurations
$M_t(P)$	distance margin from configurations more than P hops from S_t^*
r	certified upper bound on projector drift, $r \geq \rho_t$
$\eta_{t,1}, \eta_{t,2}, B_t$	perturbation terms appearing in Theorem 3.4, with $B_t = \eta_{t,1} + \eta_{t,2}$
ε, δ	Laplacian perturbation size $\ L_t - L_{t+1}\ _2$ and baseline eigengap $\lambda_{K+1}(L_t) - \lambda_K(L_t)$
P_{act}	actual bottleneck matching distance between consecutive optimizers
P_{cert}	smallest movement radius certified by the theorem
P_{max}	maximum possible matching distance over all possible configurations

3. Movement Stability Under Projector Perturbations

Throughout this section $m \leq K$, and Π_t, Π_{t+1} are arbitrary rank- K orthogonal projectors. The argument does not require the projectors to come from a graph; the fixed mobility graph comes in only through the matching distance of Definition 2.5. The main argument is listed below.

$$\begin{aligned}
\rho_t &= \|\Pi_t - \Pi_{t+1}\|_2 && \text{(projector drift)} \\
\implies \|C_t(S) - C_{t+1}(S)\|_2 &\leq \rho_t && \text{(Lemma 3.1)} \\
\implies |\log \det Y - \log \det X| &\leq \frac{2mr}{\lambda_{\min}(X)} && \text{(Lemma 3.2)} \\
\implies F_t(S_t^*) - F_t(S) &\leq B_t && \text{(Theorem 3.4 (a))} \\
\implies M_t(P) > B_t &\implies d_{\text{match}}(S_t^*, S) \leq P && \text{(Theorem 3.4 (b)).}
\end{aligned}$$

Lemma 3.1 (Compression). *For every m -subset S ,*

$$\|C_t(S) - C_{t+1}(S)\|_2 \leq \rho_t.$$

Proof. $C_t(S) - C_{t+1}(S) = R_S(\Pi_t - \Pi_{t+1})R_S^\top$. Since R_S has orthonormal rows, $\|R_S\|_2 = \|R_S^\top\|_2 = 1$, and submultiplicativity gives the bound. $\square$

Lemma 3.2 (Log-determinant stability). *Let $X, Y \in \mathbb{R}^{m \times m}$ be symmetric with $X \succ 0$ and $c := \lambda_{\min}(X) > 0$. If*

$$\|Y - X\|_2 \leq r \leq \frac{c}{2}, \quad (3)$$

then $Y \succ 0$ and

$$|\log \det Y - \log \det X| \leq \frac{2mr}{c}. \quad (4)$$

Proof. Write $E := Y - X$. By Weyl's inequality [15, Cor. 4.3.15], $\lambda_{\min}(Y) \geq c - r \geq c/2 > 0$, so $Y \succ 0$. Since $X \succ 0$ it has a symmetric positive-definite square root, and we may set $B := X^{-1/2}EX^{-1/2}$ so that $Y = X + E = X^{1/2}(I + B)X^{1/2}$. The matrix B is symmetric because E and $X^{-1/2}$ are, hence its eigenvalues $v_1, \dots, v_m$ are real. Taking determinants,

$$\log \det Y - \log \det X = \log \det(I + B) = \sum_{i=1}^m \log(1 + v_i).$$

By submultiplicativity, $\|B\|_2 \leq \|X^{-1/2}\|_2^2 \|E\|_2 = \|E\|_2 / c \leq r/c \leq 1/2$, so $|v_i| \leq 1/2$ for every i . Then, for $|x| \leq \frac{1}{2}$ the scalar inequality

$$|\log(1 + x)| \leq 2|x| \quad (5)$$

holds because for $x \in [0, 1/2]$, one has $\log(1 + x) \leq x \leq 2|x|$, while for $x \in [-1/2, 0]$,

$$|\log(1 + x)| = -\log(1 + x) = \int_{1+x}^1 \frac{dt}{t} \leq 2(-x) = 2|x|,$$

since $t \geq 1/2$ throughout the interval of integration. Therefore

$$|\log \det Y - \log \det X| \leq \sum_{i=1}^m |\log(1 + v_i)| \leq 2 \sum_{i=1}^m |v_i| \leq 2m \|B\|_2 \leq \frac{2mr}{c}.$$

This proves (4). $\square$

Lemma 3.2 gives the same conclusion if X and Y are exchanged, but its hypothesis must be checked for X . Thus, we may apply it in either direction between $C_t(S)$ and $C_{t+1}(S)$ if the minimum eigenvalue condition holds for chosen X . We use both directions below.

Proposition 3.3 (Determinant lower bound for the smallest eigenvalue). *Let Π be an orthogonal projector, S an m -subset, and $C = (\Pi)_{SS}$. Then $0 \preceq C \preceq I_m$, and if $\det C > 0$ then*

$$\lambda_{\min}(C) \geq \det C. \quad (6)$$

Proof. Since $0 \preceq \Pi \preceq I_n$, congruence by R_S gives $0 \preceq C \preceq I_m$, so every eigenvalue of C lies in $[0, 1]$. Writing $\det C = \prod_{i=1}^m \lambda_i(C)$ and bounding every factor other than the smallest by 1 gives us (6). $\square$

Proposition 3.3 shows that a configuration with determinant bounded away from zero automatically has its smallest eigenvalue bounded away from zero as well (and hence has controlled condition number because $\lambda_{\max} \leq 1$); the conditioning step in Theorem 3.4 ensures its positive. Thus Theorem 3.4 requires no uniform lower bound on $\lambda_{\min}(C_t(S))$ over all configurations. This argument applies only when $m \leq K$; the oversampled case uses the separate conditioning argument that will be introduced later in Section 7.1.

Theorem 3.4 (Movement bound under projector perturbation). *Assume $m \leq K$ and Assumption 2.1. Let $S_t^* \in \arg \max F_t$ be a reference maximizer and define $\lambda_t^* := \lambda_{\min}(C_t(S_t^*))$, and $d_t^* := \det C_t(S_t^*) > 0$. Let r satisfy $\rho_t \leq r$, and define*

$$\eta_{t,1} := \frac{2mr}{\lambda_t^*}, \text{ and } \eta_{t,2} := \frac{2mr e^{\eta_{t,1}}}{d_t^*}, \text{ and } B_t := \eta_{t,1} + \eta_{t,2}.$$

Suppose the two smallness conditions

$$r \leq \frac{\lambda_t^*}{2}, \quad r \leq \frac{d_t^* e^{-\eta_{t,1}}}{2}. \quad (7)$$

hold. Then:

(a) *Objective loss. For every $S \in \arg \max F_{t+1}$,*

$$F_t(S_t^*) - F_t(S) \leq B_t. \quad (8)$$

(b) *Movement. For every integer $P \geq 0$,*

$$M_t(P) > B_t \implies d_{\text{match}}(S_t^*, S) \leq P \text{ for every } S \in \arg \max F_{t+1}. \quad (9)$$

Proof. By Lemma 2.4, $d_t^* > 0$, and by Proposition 3.3, $\lambda_t^* \geq d_t^* > 0$.

First, we control the drift at the reference maximizer. By Lemma 3.1, $\|C_{t+1}(S_t^*) - C_t(S_t^*)\|_2 \leq \rho_t \leq r$, and by the first condition in (7), $r \leq \lambda_t^*/2$. Lemma 3.2 with $X = C_t(S_t^*)$, $Y = C_{t+1}(S_t^*)$, and $c = \lambda_t^*$ gives $C_{t+1}(S_t^*) \succ 0$ and

$$|F_{t+1}(S_t^*) - F_t(S_t^*)| \leq \frac{2mr}{\lambda_t^*} = \eta_{t,1}. \quad (10)$$

Next, let $S \in \arg \max F_{t+1}$. By optimality at time $t+1$ and (10),

$$F_{t+1}(S) \geq F_{t+1}(S_t^*) \geq F_t(S_t^*) - \eta_{t,1} = \log d_t^* - \eta_{t,1},$$

and so $\det C_{t+1}(S) \geq d_t^* e^{-\eta_{t,1}} > 0$. Thus $C_{t+1}(S)$ is positive definite, and Proposition 3.3 applied to Π_{t+1} gives us

$$\lambda_{\min}(C_{t+1}(S)) \geq \det C_{t+1}(S) \geq d_t^* e^{-\eta_{t,1}}. \quad (11)$$

Now compare the new maximizer back to time t . By Lemma 3.1, $\|C_t(S) - C_{t+1}(S)\|_2 \leq \rho_t \leq r$, and the second condition in (7) together with (11) gives

$$r \leq \frac{d_t^* e^{-\eta_{t,1}}}{2} \leq \frac{\lambda_{\min}(C_{t+1}(S))}{2}.$$

Lemma 3.2, applied with $X = C_{t+1}(S)$ and $Y = C_t(S)$, therefore gives

$$|F_t(S) - F_{t+1}(S)| \leq \frac{2mr}{\lambda_{\min}(C_{t+1}(S))} \leq \frac{2mr e^{\eta_{t,1}}}{d_t^*} = \eta_{t,2}. \quad (12)$$

Combining (10), optimality of S for F_{t+1} , and (12):

$$F_t(S_t^*) - F_t(S) = \underbrace{[F_t(S_t^*) - F_{t+1}(S_t^*)]}_{\leq \eta_{t,1}} + \underbrace{[F_{t+1}(S_t^*) - F_{t+1}(S)]}_{\leq 0} + \underbrace{[F_{t+1}(S) - F_t(S)]}_{\leq \eta_{t,2}},$$

which is (8).

Finally, suppose for the sake of contradiction that $M_t(P) > B_t$ and let $S \in \arg \max F_{t+1}$. If $d_{\text{match}}(S_t^*, S) > P$, then S is feasible in the minimum defining $M_t(P)$, so $F_t(S_t^*) - F_t(S) \geq M_t(P) > B_t$, contradicting (8). Therefore, $d_{\text{match}}(S_t^*, S) \leq P$ which proves the statement. $\square$

Corollary 3.5 (Determinant-based movement bound). *Under the hypotheses of Theorem 3.4 except (7), suppose instead that*

$$r \leq \frac{d_t^*}{4} \quad \text{and} \quad \frac{2mr}{d_t^*} \leq \log 2. \quad (13)$$

Then (7) holds, $B_t \leq \frac{6mr}{d_t^*}$, and so

$$M_t(P) > \frac{6mr}{d_t^*} \implies d_{\text{match}}(S_t^*, S) \leq P \text{ for every } S \in \arg \max F_{t+1}. \quad (14)$$

Proof. By Proposition 3.3, $\lambda_t^* \geq d_t^*$, so $r \leq d_t^*/4 \leq \lambda_t^*/2$, which is the first condition in (7), and $\eta_{t,1} = 2mr/\lambda_t^* \leq 2mr/d_t^* \leq \log 2$. Hence $e^{-\eta_{t,1}} \geq \frac{1}{2}$ and

$$\frac{d_t^* e^{-\eta_{t,1}}}{2} \geq \frac{d_t^*}{4} \geq r,$$

which is the second condition in (7). Finally $e^{\eta_{t,1}} \leq e^{\log 2} = 2$, so $\eta_{t,2} \leq 2 \cdot 2mr/d_t^* = 4mr/d_t^*$ and $B_t \leq 2mr/d_t^* + 4mr/d_t^* = 6mr/d_t^*$. Thus $M_t(P) > 6mr/d_t^*$ implies $M_t(P) > B_t$, and the second part of Theorem 3.4 gives the conclusion. $\square$

4. Laplacian Perturbations and Why Projector Drift Alone Is Not Enough

We now consider the case where Π_t is the low-frequency spectral projector of a graph Laplacian, and show that perturbation size and eigengap alone cannot provide a movement bound that tends to zero with the perturbation.

Let $G_t = (V, E, w_t)$ be a connected weighted undirected graph with combinatorial Laplacian $L_t = D_t - A_t$, so that $(L_t \mathbf{x})_i = \sum_{j \sim i} w_{t,ij}(x_i - x_j)$. Its eigenvalues are $0 = \lambda_1 \leq \dots \leq \lambda_n$ with orthonormal eigenvectors $\phi_1, \dots, \phi_n$, the graph Fourier modes [8, 29]. For a bandwidth $K < n$ set $U_t = [\phi_1, \dots, \phi_K]$ and $\Pi_t = U_t U_t^\top$.

Assumption 4.1 (Laplacian perturbation). The edge set E is the same at times t and $t+1$, only the weights change, and G_{mv} is the unweighted graph (V, E) . Furthermore, $\varepsilon := \|L_t - L_{t+1}\|_2$, and $\delta := \lambda_{K+1}(L_t) - \lambda_K(L_t) > 0$, and $\varepsilon < \frac{\delta}{4}$.

Lemma 4.2 (Projector drift via Davis–Kahan). *Under Assumption 4.1, the rank- K spectral projector Π_{t+1} is uniquely determined and*

$$\rho_t = \|\Pi_t - \Pi_{t+1}\|_2 \leq \frac{\varepsilon}{\delta - \varepsilon} \leq \frac{4\varepsilon}{3\delta} < \frac{4\varepsilon}{\delta}. \quad (15)$$

Proof. Write $L_{t+1} = L_t + E$ with $\|E\|_2 \leq \varepsilon$. By Weyl's inequality [15, Cor. 4.3.15],

$$\lambda_{K+1}(L_{t+1}) - \lambda_K(L_{t+1}) \geq \delta - 2\varepsilon > 0,$$

so the rank- K spectral projector of L_{t+1} is well defined. Let $U_t^\perp$ consist of eigenvectors of L_t for $\lambda_{K+1}(L_t), \dots, \lambda_n(L_t)$, with $(U_t^\perp)^\top L_t U_t^\perp = \Lambda_t^\perp$, and let U_{t+1} contain the first K eigenvectors of L_{t+1} , with $L_{t+1}U_{t+1} = U_{t+1}\Lambda_{t+1}$. Put $X := (U_t^\perp)^\top U_{t+1}$. Then

$$\Lambda_t^\perp X - X \Lambda_{t+1} = (U_t^\perp)^\top L_t U_{t+1} - (U_t^\perp)^\top L_{t+1} U_{t+1} = -(U_t^\perp)^\top E U_{t+1}.$$

Every eigenvalue of $\Lambda_t^\perp$ is at least $\lambda_{K+1}(L_t)$ and every eigenvalue of Λ_{t+1} is at most $\lambda_K(L_{t+1}) \leq \lambda_K(L_t) + \varepsilon$, so the two groups of eigenvalues are separated by at least $\delta - \varepsilon > 0$. The Davis–Kahan $\sin \Theta$ theorem [10, 31] gives

$$\|X\|_2 \leq \frac{\|(U_t^\perp)^\top E U_{t+1}\|_2}{\delta - \varepsilon} \leq \frac{\varepsilon}{\delta - \varepsilon}.$$

For orthogonal projectors of equal rank, $\|\Pi_t - \Pi_{t+1}\|_2 = \|(I - \Pi_t)\Pi_{t+1}\|_2$ [31, Thm. I.5.5], and

$$(I - \Pi_t)\Pi_{t+1} = U_t^\perp [(U_t^\perp)^\top U_{t+1}] U_{t+1}^\top = U_t^\perp X U_{t+1}^\top.$$

Multiplication by matrices with orthonormal columns preserves the nonzero singular values, so $\|\Pi_t - \Pi_{t+1}\|_2 = \|X\|_2$. And so finally, $\varepsilon < \delta/4$ gives $\delta - \varepsilon > 3\delta/4$, which gives the last two inequalities in (15). $\square$

Corollary 4.3 (Laplacian projector movement stability). *Assume $m \leq K$ and Assumption 4.1, and let $S_t^* \in \arg \max F_t$ with $d_t^* = \det C_t(S_t^*)$. If*

$$\frac{4\varepsilon}{\delta} \leq \frac{d_t^*}{4} \text{ and } \frac{8m\varepsilon}{d_t^* \delta} \leq \log 2, \quad (16)$$

then for every integer $P \geq 0$,

$$M_t(P) > \frac{24m\varepsilon}{d_t^* \delta} \implies d_{\text{match}}(S_t^*, S) \leq P \text{ for every } S \in \arg \max F_{t+1}. \quad (17)$$

Proof. Use Corollary 3.5 with $r = 4\varepsilon/\delta$, which satisfies $r \geq \rho_t$ by Lemma 4.2. Conditions (16) are just exactly (13) for this r , since $2mr/d_t^* = 8m\varepsilon/(d_t^* \delta)$, and the threshold is $6mr/d_t^* = 24m\varepsilon/(d_t^* \delta)$. $\square$

Corollary 4.3 uses the simplified bound $4\varepsilon/\delta$ to keep the constants uniform. Using $r = \varepsilon/(\delta - \varepsilon)$ instead gives a smaller threshold $6m\varepsilon/(d_t^*(\delta - \varepsilon))$ under the smallness conditions, and using $r = \rho_t$ directly, when ρ_t is computed, is still better.

4.1 Perturbation Size and Eigengap Alone Do Not Bound Movement

We now show that perturbation size and eigengap alone cannot control movement as the perturbation tends to zero. Our construction will keep the graph structure fixed and all edge weights positive, while an arbitrarily small weight change moves the optimizer across the entire path.

Lemma 4.4 (Discrete Sturm oscillation on a weighted path). *Let L be the Laplacian of a path on n vertices with strictly positive edge weights. Then the eigenvalues of L are simple, $0 = \lambda_1 < \lambda_2 < \dots < \lambda_n$, and, after deleting the zero coordinates, an eigenvector for λ_j will have $j - 1$ sign changes. Specifically, a nonconstant eigenvector with one sign change belongs to λ_2 .*

Proof. A positively weighted path Laplacian is a real symmetric tridiagonal (Jacobi) matrix with nonzero off-diagonal entries $-w_i < 0$. Such a matrix has simple eigenvalues [15, Sec. 7.9]. Fiedler’s oscillation theorem [12, Thm. 2.3] states that, after deleting zero coordinates, an eigenvector for λ_j has $j - 1$ sign changes; see also [13]. Therefore an eigenvector with one sign change corresponds to λ_2 . $\square$

Theorem 4.5 (No Movement Bound With Only Perturbation Size and Eigengap). *There is no function $g : (0, \infty)^2 \rightarrow \mathbb{R}_{\geq 0}$ satisfying both:*

(a) for every fixed $\delta > 0$, $\lim_{\varepsilon \rightarrow 0^+} g(\varepsilon, \delta) = 0$.

(b) for every $\varepsilon > 0$ and every pair of connected weighted-graph Laplacians L_t, L_{t+1} on a common vertex set and common edge set with $\|L_t - L_{t+1}\|_2 \leq \varepsilon$ and $\delta_t := \lambda_3(L_t) - \lambda_2(L_t) > 0$, every reference maximizer S_t^* and every $S \in \arg \max F_{t+1}$ of the one-sensor, $K = 2$ objective satisfy $d_{\text{match}}(S_t^*, S) \leq g(\varepsilon, \delta_t)$.

In more precise words, for every $n \geq 4$, there is a positively weighted path L_t on n vertices with fixed gap $\delta_n > 0$ and a sequence of positively weighted path Laplacians $L^{(k)}$ on the same edge set such that $\varepsilon_k := \|L^{(k)} - L_t\|_2 \rightarrow 0^+$, while the optimizer moves a matching distance of $n - 1$. The same holds under any fixed tie-breaking rule at time t .

Proof. Assume for the sake of contradiction that there exists a function $g : (0, \infty)^2 \rightarrow \mathbb{R}_{\geq 0}$ satisfying both (a) and (b). We construct a sequence of perturbations that violates this property.

Set $n \geq 4$. For $1 \leq i \leq n$, define

$$a_i := n + 1 - 2i, \quad q_i := \left(i - \frac{n+1}{2}\right)^2 - \frac{n^2 - 1}{12}.$$

Then set $u_i(\theta) := a_i + \theta q_i$. Both a and q have zero sum, the first by antisymmetry $a_{n+1-i} = -a_i$ and the second because $\sum_{i=1}^n \left(i - \frac{n+1}{2}\right)^2 = \frac{n(n^2-1)}{12}$. Hence $u(\theta) \perp \mathbf{1}$ for every θ . At $\theta = 0$, $a_i - a_{i+1} = 2$ for $1 \leq i \leq n-1$, and $A_i(0) := \sum_{j \leq i} a_j = i(n-i) > 0$ for $1 \leq i \leq n-1$, and

$$\begin{aligned} a_1 &= n-1 > 0 > 1-n = a_n, \\ |a_1| &= |a_n| = n-1 > \max_{2 \leq i \leq n-1} |a_i| = n-3. \end{aligned}$$

$u_i(\theta)$, their consecutive differences, and the partial sums $A_i(\theta)$ depend continuously on θ . Therefore, all strict inequalities that hold at $\theta = 0$ also hold for sufficiently small $|\theta|$. And thus, there exists $\theta_0 > 0$ such that, whenever $|\theta| < \theta_0$:

- (a) $u_1(\theta) > u_2(\theta) > \dots > u_n(\theta)$;
- (b) $A_i(\theta) := \sum_{j \leq i} u_j(\theta) > 0$ for $1 \leq i \leq n-1$;
- (c) $u_1(\theta) > 0 > u_n(\theta)$;
- (d) $\max_i |u_i(\theta)|$ is attained only at $i \in \{1, n\}$;
- (e) $u_1(\theta)$ and $u_n(\theta)$ have the same signs as a_1 and a_n .

We can take θ_0 to be the minimum of the finitely many positive radii by these continuity arguments.

For $|\theta| < \theta_0$, define the path edge weights...

$$w_i(\theta) := \frac{A_i(\theta)}{u_i(\theta) - u_{i+1}(\theta)}, \text{ for } 1 \leq i \leq n-1.$$

By (a) and (b), the numerator and denominator are positive, so $w_i(\theta) > 0$. Let $L(\theta)$ be the corresponding weighted-path Laplacian. Since $w_i(\theta)(u_i(\theta) - u_{i+1}(\theta)) = A_i(\theta)$,

$$\begin{aligned} (L(\theta)u(\theta))_1 &= u_1(\theta), \\ (L(\theta)u(\theta))_i &= u_i(\theta) \quad (2 \leq i \leq n-1), \\ (L(\theta)u(\theta))_n &= u_n(\theta), \end{aligned}$$

where the last equality uses $\sum_{j=1}^n u_j(\theta) = 0$. Thus $L(\theta)u(\theta) = u(\theta)$, so $u(\theta)$ is an eigenvector with eigenvalue 1. By (a) and (c), after deleting a possible zero coordinate, $u(\theta)$ has exactly one sign change. Lemma 4.4 therefore gives $\lambda_2(L(\theta)) = 1$ for all $|\theta| < \theta_0$. Since the eigenvalues are simple, $\delta_n := \lambda_3(L(0)) - 1 > 0$.

For $K = 2$ the low-frequency projector of $L(\theta)$ is

$$\Pi(\theta) = \frac{1}{n} \mathbf{1} \mathbf{1}^\top + \widehat{u}(\theta) \widehat{u}(\theta)^\top, \text{ and } \widehat{u}(\theta) = \frac{u(\theta)}{\|u(\theta)\|_2},$$

because $\lambda_1 = 0$ has eigenvector $\mathbf{1}/\sqrt{n}$ and $\lambda_2 = 1$ has eigenvector $\widehat{u}(\theta)$. With one sensor, $m = 1 \leq K$ and

$$F(\{i\}) = \log \Pi(\theta)_{ii} = \log\left(\frac{1}{n} + \widehat{u}_i(\theta)^2\right),$$

so maximizing F over single vertices is equivalent to maximizing $|u_i(\theta)|$.

At $\theta = 0$, the maximizers are $\{1\}$ and $\{n\}$, by $|a_1| = |a_n| > |a_i|$ for interior i . For $\theta \neq 0$ small, property (e) gives $|u_1(\theta)| = a_1 + \theta q_1$ and $|u_n(\theta)| = a_1 - \theta q_1$. This then gives

$$\begin{aligned} |u_1(\theta)| - |u_n(\theta)| &= 2\theta q_1, \\ q_1 &= q_n = \frac{(n-1)(n-2)}{6} > 0. \end{aligned}$$

By (d) no interior vertex can be maximal, so the unique maximizer is $\{1\}$ when $\theta > 0$ and $\{n\}$ when $\theta < 0$. Set $L_t := L(0)$ (whose eigengap is $\delta_t = \delta_n$). For sufficiently small $\theta \neq 0$, $L(\theta) \neq L(0)$: otherwise $u(0)$ and $u(\theta)$ would both lie in the eigenspace of $L(0)$ for the simple eigenvalue 1, and thus be proportional, forcing q to be a multiple of a .

This is impossible because $q \neq 0$ is symmetric, $q_{n+1-i} = q_i$, whereas $a \neq 0$ is antisymmetric, $a_{n+1-i} = -a_i$. Since each $w_i(\theta)$ is a rational function with nonvanishing denominator on $(-\theta_0, \theta_0)$, $L(\theta) \rightarrow L(0)$ in operator norm as $\theta \rightarrow 0$. Therefore we may choose a sequence $\theta_k < 0$ tending to zero such that the positive numbers

$$\varepsilon_k := \|L(\theta_k) - L(0)\|_2$$

are strictly decreasing. Then $\varepsilon_k \rightarrow 0^+$. Put $L^{(k)} := L(\theta_k)$. For every k , the perturbed objective has the unique maximizer $\{n\}$. Choose the reference maximizer $S_t^* = \{1\}$, since $\{1\} \in \arg \max F_t$. Then

$$d_{\text{match}}(S_t^*, S) = d(1, n) = n - 1 \text{ for the unique } S \in \arg \max F_{t+1},$$

because G_{mv} is the unweighted path. Property (b) of a hypothetical g would give $n - 1 \leq g(\varepsilon_k, \delta_n)$ for every k , contradicting (a) as $k \rightarrow \infty$.

Finally, if a fixed rule selects one maximizer at time t , choose the sign of the perturbation opposite to the endpoint selected by the rule; the same construction produces movement $n - 1$. $\square$

Here, we shall note why the impossibility result doesn't contradict Theorem 3.4. The long jump occurs because two distant configurations are tied at the baseline. With $L_t = L(0)$ and reference maximizer $\{1\}$, the opposite endpoint $\{n\}$ is also optimal and lies at matching distance $n - 1$. Thus $M_t(P) = 0$ for every $P < n - 1$, so the condition $M_t(P) > B_t$ fails exactly when this jump can occur. The distance margin rules out these ties.

5. Computing the Distance Margin and Certifying Movement

We now show how to compute the distance margin $M_t(P)$ and use it to certify a movement radius via Theorem 3.4.

5.1 Formulation

Proposition 5.1 (Distance margin representation). *For every integer $P \geq 0$,*

$$M_t(P) = F_t(S_t^*) - \max_{\substack{|S|=m \\ d_{\text{match}}(S, S_t^*) > P}} F_t(S),$$

where the maximum is $-\infty$ if the feasible family is empty. Moreover, $M_t(P)$ is non-decreasing in P .

Proof. The identity follows from

$$\min_S (c - F_t(S)) = c - \max_S F_t(S)$$

over the same finite family, with $c = F_t(S_t^*)$. As P increases, the constraint $d_{\text{match}}(S, S_t^*) > P$ becomes more restrictive, so the feasible family shrinks and the maximum cannot increase. This implies that $M_t(P)$ cannot decrease. $\square$

Proposition 5.2 (Perfect-matching test). *Let $S_t^* = \{q_1, \dots, q_m\}$ and $S = \{s_1, \dots, s_m\}$. Construct a bipartite graph with left vertices q_i , right vertices s_j , and an edge (q_i, s_j) whenever $d(q_i, s_j) \leq P$. Then*

$$d_{\text{match}}(S_t^*, S) \leq P$$

if and only if this bipartite graph has a perfect matching.

Proof. If the bipartite graph has a perfect matching, then there is a bijection π such that $d(q_i, s_{\pi(i)}) \leq P$ for every i . Thus, by Definition 2.5, $d_{\text{match}}(S_t^*, S) \leq P$.

Conversely, if $d_{\text{match}}(S_t^*, S) \leq P$, then since there are finitely many bijections between S_t^* and S , the minimum in Definition 2.5 is attained by some bijection π . Thus we have $d(q_i, s_{\pi(i)}) \leq P$ for every i , so the corresponding edges form a perfect matching in the bipartite graph. $\square$

This is a bipartite perfect-matching problem, which can be solved in $O(m^{2.5})$ time in the worst case [22]. The distance d_{match} is a bottleneck assignment problem [5] and can be found by binary search over the $O(m^2)$ distinct pairwise distances.

5.2 Computing the Distance Margin by Enumeration

When all $\binom{n}{m}$ configurations can be enumerated, Proposition 5.1 gives the following exact algorithm.

Algorithm 1: exact evaluation of $M_t(P)$ (undersampled case)

1. Form Π_t .
 2. For each m -subset S , evaluate $F_t(S) = \log \det(\Pi_t)_{SS}$, with $-\infty$ if singular.
 3. Record $F_t(S_t^*) = \max_S F_t(S)$ and fix a reference maximizer S_t^* .
 4. For each S , decide $d_{\text{match}}(S, S_t^*) > P$ by the perfect-matching test of Proposition 5.2.
 5. Return $M_t(P) = F_t(S_t^*) - \max\{F_t(S) : d_{\text{match}}(S, S_t^*) > P\}$, or $+\infty$ if no such S exists.
-

If Π_t is formed explicitly, preprocessing costs $O(n^2 K)$. All pairs unweighted distances can be computed by n breadth-first searches in $O(n(n + |E_{\text{mv}}|))$ time. Then, total enumeration will cost $O(\binom{n}{m}(m^3 + m^{2.5}))$. This can be seen later on starting from Section 8, where we will use enumeration of all m -subsets and bottleneck matching distances in experiments.

6. Losslessness of P -Hop and Voronoi Restrictions

Theorem 3.4 places every new maximizer in the P -hop matching ball. We now give conditions under which these restrictions preserve the optimum for the local feasible regions used in dynamic placement.

Theorem 6.1 (*P*-hop bridge). Assume the hypotheses of Theorem 3.4 and $M_t(P) > B_t$. Then every $S \in \arg \max F_{t+1}$ satisfies

$$d_{\text{match}}(S_t^*, S) \leq P,$$

and

$$\max_{\substack{|S|=m \\ d_{\text{match}}(S, S_t^*) \leq P}} F_{t+1}(S) = \max_{|S|=m} F_{t+1}(S). \quad (18)$$

Moreover,

$$\arg \max_{\substack{|S|=m \\ d_{\text{match}}(S, S_t^*) \leq P}} F_{t+1}(S) = \arg \max F_{t+1};$$

if the global maximizer is unique, then the constrained and unconstrained maximizers agree.

Proof. Theorem 3.4 (b) shows that every global maximizer of F_{t+1} satisfies $d_{\text{match}}(S, S_t^*) \leq P$. Hence the constrained family contains every global maximizer, while also being a subfamily of all m -sets. Its maximum therefore equals the global optimal value, which proves (18). A configuration in the constrained family attains this value if and only if it is a global maximizer, which proves the equality of the two argmax sets. $\square$

6.1 Intersection with the Voronoi Region

Nomura et al. [25] restrict sensor i to the strict Voronoi cell of its current position, intersected with a P -hop neighborhood. We define this region and the conditions under which it retains a global optimum. Fix an ordering $S_t^* = \{q_1, \dots, q_m\}$ and define, on the fixed movement graph G_{mv} ,

$$W_i^< := \{x \in V : d(x, q_i) < d(x, q_j) \text{ for all } j \neq i\}.$$

The strict cells are pairwise disjoint, but a vertex equidistant from two of the q_i will belong to none of them. Define the corresponding local feasible family by

$$V_P := \{\{x_1, \dots, x_m\} : x_i \in W_i^< \text{ and } d(x_i, q_i) \leq P \text{ for every } i\}.$$

Because the strict Voronoi cells are pairwise disjoint, the vertices $x_1, \dots, x_m$ in every element of V_P are distinct.

These results are relevant to the local feasible region, not the performance of a particular update algorithm. They show that (guaranteed only under the stated conditions) V_P contains a global maximizer of F_{t+1} , and not that independent per-sensor updates recover it. The regions are also centered at the reference maximizer S_t^* ; applying them to another time t configuration would require a separate analysis of tracking error (will be discussed in future works).

Corollary 6.2 (Losslessness under an assignment consistent with the Voronoi cells). Assume the hypotheses of Theorem 6.1, let $S = \{r_1, \dots, r_m\} \in \arg \max F_{t+1}$, and suppose there is a permutation π with

$$r_{\pi(i)} \in W_i^< \quad \text{and} \quad d(q_i, r_{\pi(i)}) \leq P \quad \text{for every } i. \quad (19)$$

Then $S \in V_P$, and consequently

$$\max_{S' \in V_P} F_{t+1}(S') = \max_{|S'|=m} F_{t+1}(S').$$

Proof. Condition (19) gives $r_{\pi(i)} \in W_i^<$ and $d(q_i, r_{\pi(i)}) \leq P$ for every i , so $S \in V_P$. Since S is a global maximizer and V_P is a subfamily of all m -sets, maximizing over V_P gives the same optimal value. $\square$

For $m = 1$, this condition holds automatically. For $m \geq 2$, the spacing condition below is enough.

Proposition 6.3 (Compatibility with strict Voronoi cells). *Assume $m \geq 2$. If*

$$\min_{i \neq j} d(q_i, q_j) > 2P, \quad (20)$$

then every vertex x satisfying $d(x, q_i) \leq P$ belongs to $W_i^<$. So, every bijection with bottleneck radius at most P satisfies (19).

Proof. Let $x \in V$ satisfy $d(x, q_i) \leq P$ and let $j \neq i$. By the triangle inequality and (20),

$$d(x, q_j) \geq d(q_i, q_j) - d(x, q_i) > 2P - P = P \geq d(x, q_i),$$

so q_i is strictly closer to x than any other q_j , that is, $x \in W_i^<$. If π is a bijection with $d(q_i, r_{\pi(i)}) \leq P$ for all i , then the previous argument gives $r_{\pi(i)} \in W_i^<$, which together with $d(q_i, r_{\pi(i)}) \leq P$ is (19). $\square$

Corollary 6.4 (Voronoi restriction losslessness under enough spacing). *Assume the hypotheses of Theorem 6.1. If $m = 1$, or if $m \geq 2$ and the spacing condition (20) holds, then V_P contains every $S \in \arg \max F_{t+1}$, and exact maximization of F_{t+1} over V_P is lossless in objective value.*

Proof. For $m = 1$, $W_1^< = V$. For $m \geq 2$, Theorem 6.1 supplies, for each $S \in \arg \max F_{t+1}$, a bijection from S_t^* to S with bottleneck radius at most P . Proposition 6.3 makes it compatible in the sense of (19), and Corollary 6.2 applies. $\square$

A fixed assignment of Voronoi ties does not affect Corollary 6.4, since (20) guarantees uniqueness within distance P ; furthermore, we shall highlight that it's sufficient but not necessary for Corollary 6.2.

7. Extensions

We extend the main result to oversampled and regularized D-optimality and then note how learned dictionaries fit directly into the general projector framework without further proofs.

7.1 Oversampled D-Optimality, $m \geq K$

For every m -set S one has $0 \preceq H_t(S) \preceq I_K$. Consequently, if $\det H_t(S) > 0$ then $\lambda_{\min}(H_t(S)) \geq \det H_t(S) = e^{F_t^{\text{ov}}(S)}$. Since $0 \preceq D_S \preceq I_n$ and $U_t^\top U_t = I_K$, congruence gives $0 \preceq U_t^\top D_S U_t \preceq U_t^\top U_t = I_K$, so all eigenvalues lie in $[0, 1]$ and the determinant argument of Proposition 3.3 applies.

Lemma 7.1 (Oversampled log-determinant stability). *Let $c_S := \lambda_{\min}(H_t(S)) > 0$ and let $r \geq \rho_t$ satisfy $r \leq c_S/4$. Then $H_{t+1}(S)$ is positive definite and*

$$|F_t^{\text{ov}}(S) - F_{t+1}^{\text{ov}}(S)| \leq \frac{4Kr}{c_S}. \quad (21)$$

Proof. Set $A_t := \Pi_t D_S \Pi_t = U_t H_t(S) U_t^\top$, so the nonzero eigenvalues of A_t are the same as those of $H_t(S)$, and likewise at time $t+1$. From

$$A_t - A_{t+1} = (\Pi_t - \Pi_{t+1}) D_S \Pi_t + \Pi_{t+1} D_S (\Pi_t - \Pi_{t+1})$$

and $\|D_S\|_2 \leq 1$, $\|\Pi_t\|_2 = \|\Pi_{t+1}\|_2 = 1$ we get $\|A_t - A_{t+1}\|_2 \leq 2\rho_t \leq 2r$. Order eigenvalues nonincreasingly. By Weyl's inequality, $|\lambda_j(A_{t+1}) - \lambda_j(A_t)| \leq 2r$ for all j . For $1 \leq j \leq K$ we have $\lambda_j(A_t) \geq c_S$ and $r \leq c_S/4$, so $2r \leq c_S/2$, so $\lambda_j(A_{t+1}) \geq c_S/2 > 0$. Since $\text{rank } A_{t+1} \leq K$, these are exactly the nonzero eigenvalues of A_{t+1} , hence the eigenvalues of $H_{t+1}(S)$, which is thus positive definite. Setting

$$x_j := \frac{\lambda_j(A_{t+1}) - \lambda_j(A_t)}{\lambda_j(A_t)}, \text{ and } |x_j| \leq \frac{2r}{c_S} \leq \frac{1}{2},$$

and using $\det H_t(S) = \prod_{j \leq K} \lambda_j(A_t)$, $\det H_{t+1}(S) = \prod_{j \leq K} \lambda_j(A_{t+1})$ together with (5),

$$|F_{t+1}^{\text{ov}}(S) - F_t^{\text{ov}}(S)| = \left| \sum_{j=1}^K \log(1 + x_j) \right| \leq 2 \sum_{j=1}^K |x_j| \leq \frac{4Kr}{c_S}.$$

This proves (21). $\square$

Let $S_t^* \in \arg \max F_t^{\text{ov}}$, set $\sigma_t^{*,\text{ov}} := \det H_t(S_t^*)$, and define the oversampled distance margin by

$$M_t^{\text{ov}}(P) := \min_{\substack{|S|=m \\ d_{\text{match}}(S, S_t^*) > P}} [F_t^{\text{ov}}(S_t^*) - F_t^{\text{ov}}(S)].$$

Since U_t has rank K , it has K linearly independent rows. For $m \geq K$, include those rows in an m -set S ; then $H_t(S) \succ 0$. Therefore F_t^{ov} has a finite maximum, and every reference maximizer has positive determinant.

Theorem 7.2 (Movement bound in the oversampled case). *Assume $m \geq K$ and Assumption 2.1. Let $r \geq \rho_t$ satisfy*

$$r \leq \frac{\sigma_t^{*,\text{ov}}}{8}, \text{ and } \eta_t^{*,\text{ov}} := \frac{4Kr}{\sigma_t^{*,\text{ov}}} \leq \log 2. \quad (22)$$

Then every $S \in \arg \max F_{t+1}^{\text{ov}}$ satisfies

$$F_t^{\text{ov}}(S_t^*) - F_t^{\text{ov}}(S) \leq 3\eta_t^{*,\text{ov}}.$$

Thus, consequently, for every integer $P \geq 0$,

$$M_t^{\text{ov}}(P) > 3\eta_t^{*,\text{ov}} = \frac{12Kr}{\sigma_t^{*,\text{ov}}} \implies d_{\text{match}}(S_t^*, S) \leq P \text{ for every } S \in \arg \max F_{t+1}^{\text{ov}}. \quad (23)$$

Under Assumption 4.1 one may take $r = 4\varepsilon/\delta$, in which case the hypotheses read $4\varepsilon/\delta \leq \sigma_t^{*,\text{ov}}/8$ and $16K\varepsilon/(\sigma_t^{*,\text{ov}}\delta) \leq \log 2$, and the threshold becomes $48K\varepsilon/(\sigma_t^{*,\text{ov}}\delta)$.

Proof. Assume for the sake of contradiction that (23) does not hold. Repeat Theorem 3.4's argument, using Lemma 7.1 in place of Lemma 3.2. Since $0 \preceq H_s(S) \preceq I_K$, positive determinant implies $\lambda_{\min}(H_s(S)) \geq \det H_s(S)$. Thus the reference-maximizer drift is at most $\eta_t^{*,\text{ov}}$. Optimality at time $t+1$ then gives $\det H_{t+1}(S) \geq \sigma_t^{*,\text{ov}} e^{-\eta_t^{*,\text{ov}}} \geq \sigma_t^{*,\text{ov}}/2$, so the reverse comparison contributes at most $2\eta_t^{*,\text{ov}}$. Hence $F_t^{\text{ov}}(S_t^*) - F_t^{\text{ov}}(S) \leq 3\eta_t^{*,\text{ov}}$, which contradicts the definition of the distance margin and proves (23). The Laplacian specialization follows from Lemma 4.2. $\square$

7.2 Regularized D-Optimality

For this subsection, define

$$C_s(S) := R_S \Pi_s R_S^\top, \quad s \in \{t, t+1\},$$

for arbitrary m . Lemma 3.1 still applies because its proof does not use $m \leq K$. For $\tau > 0$ set

$$F_s^\tau(S) := \log \det(C_s(S) + \tau I_m).$$

Writing $U_{s,S} := R_S U_s$, Sylvester's determinant identity gives

$$\det(\tau I_m + U_{s,S} U_{s,S}^\top) = \tau^m \det(I_K + \tau^{-1} U_{s,S}^\top U_{s,S})$$

and therefore

$$\det(C_s(S) + \tau I_m) = \tau^{m-K} \det(U_{s,S}^\top U_{s,S} + \tau I_K).$$

Thus $F_s^\tau(S)$ and the regularized $K \times K$ information log-determinant differ only by constant $(m-K) \log \tau$, so they have the same maximizers. Since $C_s(S) \succeq 0$ and $\tau > 0$, $C_s(S) + \tau I_m$ is positive definite for every S . Let $S_t^{*,\tau} \in \arg \max F_t^\tau$ be a reference maximizer and define

$$M_t^\tau(P) := \min_{\substack{|S|=m \\ d_{\text{match}}(S, S_t^{*,\tau}) > P}} [F_t^\tau(S_t^{*,\tau}) - F_t^\tau(S)].$$

Lemma 7.3 (Uniform regularized stability). *Let $r \geq \rho_t$ with $r/\tau \leq 1/2$. Then for every m -set S ,*

$$|F_t^\tau(S) - F_{t+1}^\tau(S)| \leq \frac{2mr}{\tau} =: \eta_t^\tau. \quad (24)$$

Proof. Apply Lemma 3.2 with $X = C_t(S) + \tau I_m$ and $Y = C_{t+1}(S) + \tau I_m$. Then $\lambda_{\min}(X) \geq \tau$ because $C_t(S) \succeq 0$, and $\|Y - X\|_2 = \|C_t(S) - C_{t+1}(S)\|_2 \leq \rho_t \leq r \leq \tau/2 \leq \lambda_{\min}(X)/2$ by Lemma 3.1. The bound $2mr/\tau$ follows. $\square$

Theorem 7.4 (Movement bound for the regularized objective). *Assume Assumption 2.1. Let $r \geq \rho_t$ with $r/\tau \leq 1/2$. Then for every integer $P \geq 0$, and for every $S \in \arg \max F_{t+1}^\tau$, we have*

$$M_t^\tau(P) > \frac{4mr}{\tau} \implies d_{\text{match}}(S_t^{\star, \tau}, S) \leq P. \quad (25)$$

Under Assumption 4.1 it suffices that $4\varepsilon/(\tau\delta) \leq 1/2$ and $M_t^\tau(P) > 16m\varepsilon/(\tau\delta)$.

Proof. Let $S \in \arg \max F_{t+1}^\tau$. By Lemma 7.3 applied at $S_t^{\star, \tau}$ and at S , and by optimality of S at time $t+1$,

$$\begin{aligned} F_t^\tau(S_t^{\star, \tau}) - F_t^\tau(S) &= [F_t^\tau(S_t^{\star, \tau}) - F_{t+1}^\tau(S_t^{\star, \tau})] \\ &\quad + [F_{t+1}^\tau(S_t^{\star, \tau}) - F_{t+1}^\tau(S)] \\ &\quad + [F_{t+1}^\tau(S) - F_t^\tau(S)] \\ &\leq \eta_t^\tau + 0 + \eta_t^\tau = 2\eta_t^\tau = \frac{4mr}{\tau}. \end{aligned}$$

If $d_{\text{match}}(S_t^{\star, \tau}, S) > P$ the margin definition forces the opposite strict inequality. The Laplacian statement follows from Lemma 4.2 with $r = 4\varepsilon/\delta$. $\square$

Note that the threshold here is $2\eta_t^\tau$: Lemma 7.3 is uniform over configurations, so no separate conditioning argument at the new maximizer is needed.

7.3 Dictionaries

The main theorem also applies directly when the changing signal model is given by a learned dictionary rather than a graph Laplacian. Let $A_t, A_{t+1} \in \mathbb{R}^{n \times K}$ have full column rank and define

$$Q_s := A_s(A_s^\top A_s)^{-1}A_s^\top, \quad s \in \{t, t+1\}.$$

Each Q_s is an orthogonal projector of rank K . Therefore the results of Section 3 can apply directly and smoothly with

$$\Pi_s = Q_s, \quad \rho_t = \|Q_t - Q_{t+1}\|_2.$$

In particular, for $m \leq K$ the objective is $F_s(S) = \log \det(Q_s)_{SS}$, and the movement certificate is exactly that of what we established in Theorem 3.4. No spectral eigengap or Davis–Kahan estimate is required because the projector drift can be computed directly from Q_t and Q_{t+1} .

8. Experiments on Synthetic Graphs

8.1 Protocol and Setup

The following synthetic experiments use connected weighted graphs with a fixed edge set. The corresponding unweighted graph is also the movement graph, so perturbations change only the edge weights, not the distances used by d_{match} . Together, three graph families cover different geometrical settings (regular lattice, irregular spatial graph, and one with community structure). We use $K = 4$ and $m = 2$ to stay in the undersampled setting of Theorem 3.4 while keeping the number of configurations small enough to compute all directly.

We use two baseline graphs from each of three families (with random but fixed seeds): 4×4 grids ($n = 16$; seeds 500 and 537), random geometric graphs ($n = 12$, connection radius 0.46; seeds 509 and 546), and two-block stochastic block graphs ($n = 12$, block sizes 6 and 6, intra-block connection probability 0.50, inter-block connection probability 0.08; seeds 503 and 540). If a random geometric graph is disconnected, we connect its components using the shortest Euclidean edges between components. If a stochastic block graph is disconnected, we connect the first two components using their smallest-labeled vertices and repeat until the graph is connected. Initial edge weights are drawn independently from $\text{Unif}[0.7, 1.3]$.

For each edge with weight w , we replace w by $(1 + \alpha)w$ for

$$\alpha \in \{-0.20, -0.10, -0.05, -0.02, 0.02, 0.05, 0.10, 0.20\},$$

while keeping all other edge weights fixed. These values test small and moderate single edge changes while keeping all weights positive and the edge set fixed.

For each graph Laplacian, we form the rank- K projector onto its first K eigenvectors and evaluate all $\binom{n}{m}$ configurations. Graph distances and bottleneck matching radii are integers, and log-determinants and spectral quantities will be calculated with double precision. We treat a compression as positive definite when its minimum eigenvalue exceeds 10^{-11} . We keep configurations whose objective values are within $10^{-9} + 10^{-9} \max\{1, |F_{\max}|\}$ of the maximum, set objective gaps below magnitude of 10^{-10} to zero, and test the two smallness conditions with tolerance 10^{-12} .

For a fixed reference maximizer S_t^* , we use the observed projector drift $r = \rho_t$. When both smallness conditions of Theorem 3.4 hold, define

$$P_{\text{cert}} := \min\{P : M_t(P) > B_t\},$$

$$P_{\text{act}} := \max_{S \in \arg \max F_{t+1}} d_{\text{match}}(S_t^*, S),$$

and

$$P_{\text{max}} := \max_{|S|=m} d_{\text{match}}(S_t^*, S).$$

Theorem 3.4 gives $P_{\text{act}} \leq P_{\text{cert}}$. Since $M_t(P_{\text{max}}) = +\infty$, P_{max} always certifies the entire feasible family. We thus call a certificate nontrivial when $P_{\text{cert}} < P_{\text{max}}$.

All random seeds are fixed. Calculations use Python, NumPy, SciPy, pandas, and NetworkX, and figures are generated with Matplotlib. Reproducibility code and the CSV files used for the figures are available in the GitHub repository <https://github.com/sdg-bella/continuous-graph-sampling>.

8.2 Certifying Optimizer Movement on Synthetic Graphs

The methodology produces 1,224 one-edge perturbation cases. Of these, 1,189 (97.1%) satisfy both projector perturbation smallness conditions with $r = \rho_t$ and are eligible for certification. In every eligible case, the computation gives $P_{\text{act}} \leq P_{\text{cert}}$, as required by Theorem 3.4.

Among the eligible cases, 1,035 (87.0%) have a nontrivial certificate $P_{\text{cert}} < P_{\text{max}}$. In 561 cases (47.2%), $P_{\text{cert}} = P_{\text{act}}$ with integer radius, with 547 of these being $P_{\text{cert}} = P_{\text{act}} = 0$. The optimizer moves in 33 eligible cases, and 14 of these are integer-radius sharp. The median difference $P_{\text{cert}} - P_{\text{act}}$ is two hops.

The two-hop median gap also shows that it can still be conservative. It often rules out large movements without giving the smallest possible radius. The six baselines are the same across graph families by projector drift or distance margin, so these frequencies should not be used to rank the graph families.

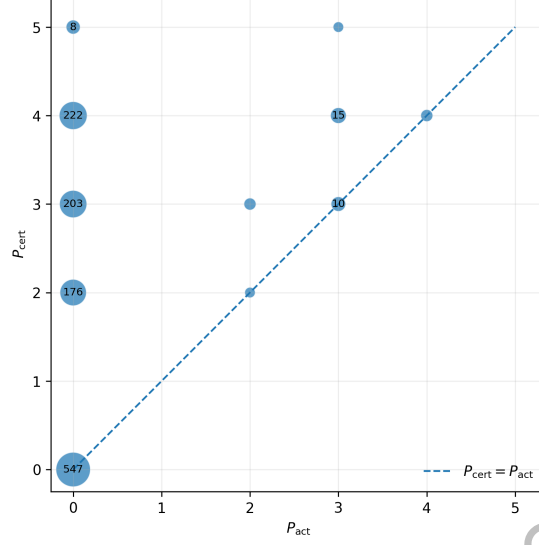


Figure 1: Actual matching movement P_{act} versus certified radius P_{cert} over the 1,189 eligible perturbation cases. Bubble area corresponds to repeated integer pairs.

8.3 Distance Margin Curve

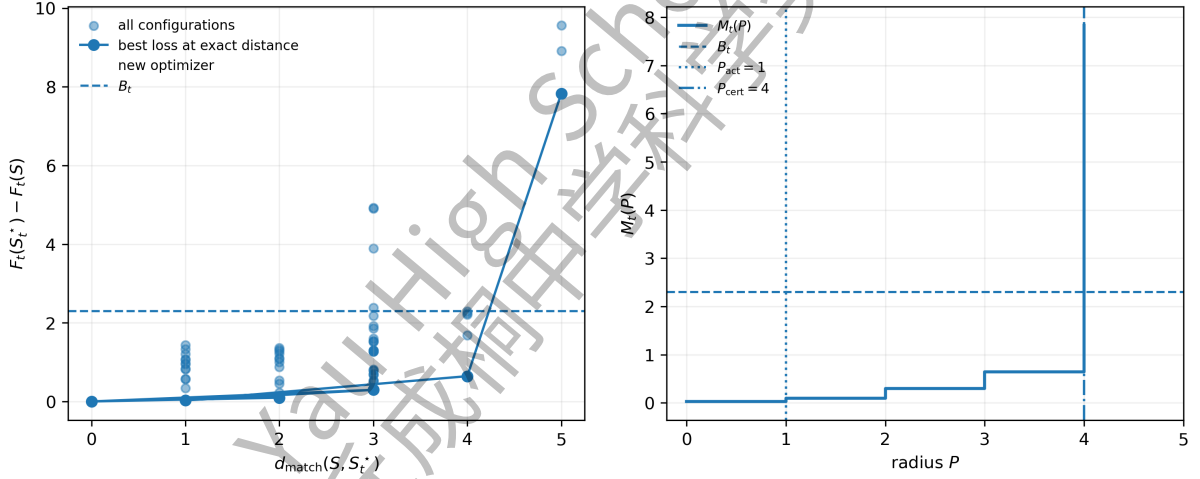


Figure 2: Objective loss by distance matching for the example random geometric graph (left) and the distance margin by radius (right). The horizontal line is B_t , and vertical are P_{act} and P_{cert}

We show the distance margin using a random geometric graph with $n = 12$, $K = 4$, and $m = 2$, selected by the fixed rule in the reproducibility code. For this perturbation,

$$P_{\text{act}} = 1, \quad P_{\text{cert}} = 4, \quad P_{\text{max}} = 5,$$

with $\rho_t = 0.05193$ and $B_t = 2.30128$.

Figure 2 plots $d_{\text{match}}(S, S_t^*)$ versus $F_t(S_t^*) - F_t(S)$ for every configuration S . The left plot therefore shows the objective loss of every configuration. For each integer radius P , $M_t(P)$ is the smallest objective loss among configurations with $d_{\text{match}}(S, S_t^*) > P$.

The right side shows how $M_t(P)$ changes with P , and the first integer radius satisfying $M_t(P) > B_t$ is P_{cert} . It's important to highlight that the left panel shows what a single second best gap misses. Configurations with nearly the same objective loss can lie at very different matching distances. The

distance margin keeps this distance information, which is what makes it useful for certifying optimizer movement.

9. Real-World Experiments on Traffic Sensor Data

The preceding sections establish the movement bound synthetically on controlled graph families. We next test the framework on real traffic data, using two traffic speed benchmarks, PEMS-BAY and METR-LA [20, 21]. The traffic measurements provide an evolving signal from which we learn an empirical low-dimensional projector, while a fixed road-based graph defines the movement metric (rather than performing traffic forecasting). The experiment therefore fits the general projector result.

9.1 Datasets and Preprocessing

PEMS-BAY contains 52,116 five minute observations from 325 sensors in the San Francisco Bay Area, while METR-LA contains 34,272 five minute observations from 207 freeway sensors in Los Angeles [6, 20, 21]. PEMS-BAY is the cleaner dataset, whereas METR-LA has more missing values and more variation.

Each dataset can be represented as a time-varying traffic matrix, where each row corresponds to one five-minute observation and each column corresponds to a fixed roadway sensor. The entries of the matrix are measured vehicle speeds at different physical locations. Because nearby sensors often experience correlated traffic patterns, the spatial variation across sensors can oftentimes be approximated by a low-dimensional subspace. In our experiments, this evolving subspace is learned from rolling traffic windows and used to construct the projector that’s core to the D-optimal selection problem.

The preprocessing protocol is fixed before checking theorem outcomes (we do not simply select configurations or intervals that produce positive results). Original timestamps and sensor identifiers are preserved, and zero or nonfinite measurements are treated as unavailable. No interpolation, filling, normalization, or temporal smoothing is performed. All 325 PEMS-BAY sensors are kept, whereas METR-LA sensors with more than 10% unavailable measurements are not, leaving 173 of the original 207 sensors. Any window that still contains an unavailable value, or that is not made up of consecutive five minute observations, is also excluded from the analysis.

Because computation of the distance margin requires complete evaluation of all possible configurations, we perform certification on a fixed 25-sensor physical subnetwork of each dataset. The movement graph is constructed without reference to projector drift, optimizer movement, margins, or whether the theorem conditions hold. Using the given road network distance metadata, each eligible sensor is connected to its three nearest sensors by road distance and the union of these edges is symmetrized. We add no new or artificial edges and keep the largest connected component, and a breadth first traversal from the first kept sensor in the original data order is performed to select 25 sensors. The resulting PEMS-BAY and METR-LA subnetworks have hop diameters 10 and 7, respectively. Thus one unit of d represents one edge in a sparse road connectivity graph rather than a unit of Euclidean / road distance.

The loop detectors in the original datasets are stationary; therefore, the matching distance does not measure physical movement of those detectors. Instead, selected sensors represent movable sensing resources placed at the same network locations, and the metric measures how far an optimal configuration would need to move when the learned traffic subspace changes.

9.2 Empirical Traffic Projectors

Let $X_t \in \mathbb{R}^{q \times n}$ denote a window of traffic-speed measurements, with rows corresponding to successive five minute observations and columns corresponding to the fixed selected sensors. The primary experiment uses $n = 25$ sensors and a window length of $q = 48$, corresponding to four hours of traffic observations. Let $V_{t,K} \in \mathbb{R}^{n \times K}$ contain the leading $K = 5$ right singular vectors of X_t , and define the empirical traffic projector

$$Q_t := V_{t,K} V_{t,K}^\top. \quad (26)$$

Because the columns of X_t are sensor coordinates, the right singular vectors describe the dominant spatial subspace across sensors, and Q_t is an orthogonal projector on $\mathbb{R}^{25}$. When the fifth and sixth singular values are distinct, Q_t is the unique projector onto the leading five dimensional right-singular subspace. In the implementation, singular-vector selection is performed by a fixed deterministic SVD routine.

Primary analysis uses raw, uncentered speed matrix in each window. Consecutive projectors are compared at the original five minute update rate, so adjacent four hour windows share 47 of their 48 observations. The experiment thus tests five minute online updates (the four hour windows are not independent samples). Invalid windows are discarded (not filled in) and transitions are evaluated only between consecutive valid windows so we never compare windows across a missing interval.

For each valid transition the projector drift is

$$\rho_t := \|Q_t - Q_{t+1}\|_2. \quad (27)$$

We don't construct a time-varying traffic Laplacian or use a Davis–Kahan bound, as the projector theorem is applied directly with $r = \rho_t$.

9.3 Certification Protocol

We use sensor budget $m = 2$ and rank $K = 5$, so the experiment is in the undersampled case of the main theorem. For each window all

$$\binom{25}{2} = 300 \quad (28)$$

two-sensor configurations are evaluated. For each configuration S we compute

$$F_t(S) := \log \det(Q_t)_{SS}, \quad (29)$$

with $-\infty$ when the compression is singular. A reference S_t^* is fixed if the old objective has multiple maximizers, while every maximizer at the new time is kept when evaluating the theorem.

At the old reference optimizer we apply Theorem 3.4 with $r = \rho_t$, computing λ_t^* , d_t^* , B_t , $M_t(P)$, P_{cert} , P_{act} , and P_{max} exactly as defined in Sections 3 and 8.1. A transition is theorem-eligible when both smallness conditions (7) hold at $r = \rho_t$.

For every eligible transition, we compute the smallest certified radius

$$P_{\text{cert}}(t) = \min\{P : M_t(P) > B_t\},$$

and compare it with the actual movement

$$P_{\text{act}}(t) = \max_{S \in \arg \max_{F_{t+1}} F_t(S)} d_{\text{match}}(S_t^*, S).$$

We call the certificate nontrivial when $P_{\text{cert}} < P_{\text{max}}$, excluding the radius that already contains every feasible configuration.

9.4 Certification Results

Table 2 summarizes the data results. The preprocessing gives us 51,824 usable consecutive transitions in PEMS-BAY and 19,724 in METR-LA. Of these, 19,259 PEMS-BAY transitions (37.2%) and 7,294 METR-LA transitions (37.0%) satisfy both smallness conditions. In total, the theorem applies to 26,553 traffic transitions.

Most five minute transitions leave the globally optimal two-sensor configuration unchanged, which is expected because adjacent rolling windows overlap heavily. Nonzero movement still occurs in 651 eligible PEMS-BAY transitions and 275 eligible METR-LA transitions. Across all 26,553 eligible transitions,

$$P_{\text{act}}(t) \leq P_{\text{cert}}(t) \quad (30)$$

with no violations. In addition, 87.2% of eligible PEMS-BAY transitions and 76.0% of eligible METR-LA transitions have $P_{\text{cert}} < P_{\text{max}}$, so most certificates are nontrivial.

Table 2: Certification results on the two real traffic datasets. “Raw/retained/used” denotes the original number of sensors, the number retained after the data-quality screen, and the size of the fixed certified subnetwork. Percentages in the $P_{\text{act}} > 0$ and nontrivial certificate columns are taken among theorem-eligible transitions. A certificate is nontrivial when $P_{\text{cert}} < P_{\text{max}}$.

Dataset	Raw/Kept/Used	Usable transitions	Eligible	$P_{\text{act}} > 0$	Nontrivial cert.	Viol.
PEMS-BAY	325/325/25	51,824	19,259 (37.2%)	651 (3.4%)	16,798 (87.2%)	0
METR-LA	207/173/25	19,724	7,294 (37.0%)	275 (3.8%)	5,543 (76.0%)	0

Figure 3 shows the actual and certified movement pairs. Each point corresponds to an integer pair $(P_{\text{act}}, P_{\text{cert}})$, with the bubble number as the number of eligible transitions there. Theorem 3.4 requires all observations to lie on or above the diagonal $P_{\text{cert}} = P_{\text{act}}$ (no point appears below it in either dataset). This also shows how much the certified radius can exceed the actual movement.

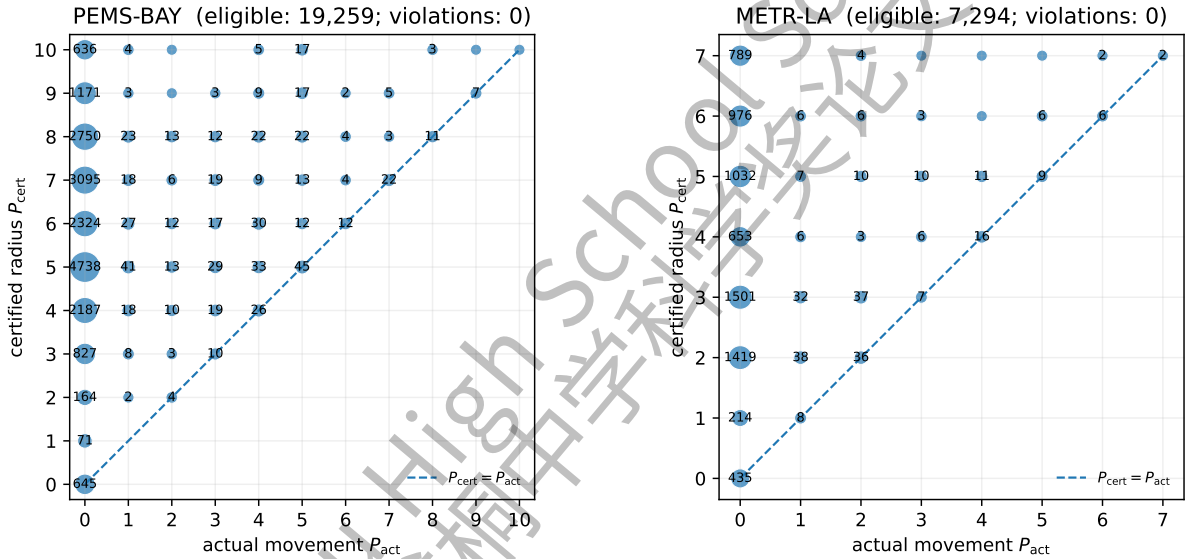


Figure 3: Actual versus certified optimizer movement on PEMS-BAY and METR-LA. Each bubble shows one $(P_{\text{act}}, P_{\text{cert}})$ pair, with bubble size and label indicating its frequency. The dashed diagonal marks equality. All eligible transitions lie on or above the diagonal, with no violations.

9.5 A Nontrivial METR-LA Example

Figure 4 shows a METR-LA transition chosen by a fixed rule. Among eligible transitions with nonzero movement, $P_{\text{cert}} = P_{\text{act}}$, and $P_{\text{cert}} < P_{\text{max}}$, we select the one with the largest P_{act} , breaking ties by earliest time. This gives the transition from May 19, 2012, 20:10 to 20:15.

For this transition the observed projector drift is approximately $\rho_t \approx 0.1087$ and the theorem gives $B_t \approx 2.5261$. The new global optimizer lies at bottleneck matching distance

$$P_{\text{act}} = 6. \quad (31)$$

For every radius $P \leq 5$, the distance margin is only

$$M_t(P) \approx 0.0544 < B_t, \quad (32)$$

so the theorem cannot exclude an optimizer beyond those radii. At $P = 6$, however, the more distant

configurations change and the margin jumps to

$$M_t(6) \approx 2.8216 > B_t. \quad (33)$$

Resultingly,

$$P_{\text{cert}} = 6 = P_{\text{act}}. \quad (34)$$

Since $P_{\text{max}} = 7$, the certificate is nontrivial: $P_{\text{cert}} = 6$ does not contain the full configuration space.

The left side of Figure 4 plots all 300 old-time objective losses against matching distance from S_t^* . The lowest point at each distance gives the best competitor at that distance. The new optimizer lies at distance 6 with a small old-time loss, so radii below 6 cannot be certified. The right plot shows the same result through $M_t(P)$, which stays below B_t through $P = 5$ and crosses it at $P = 6$.

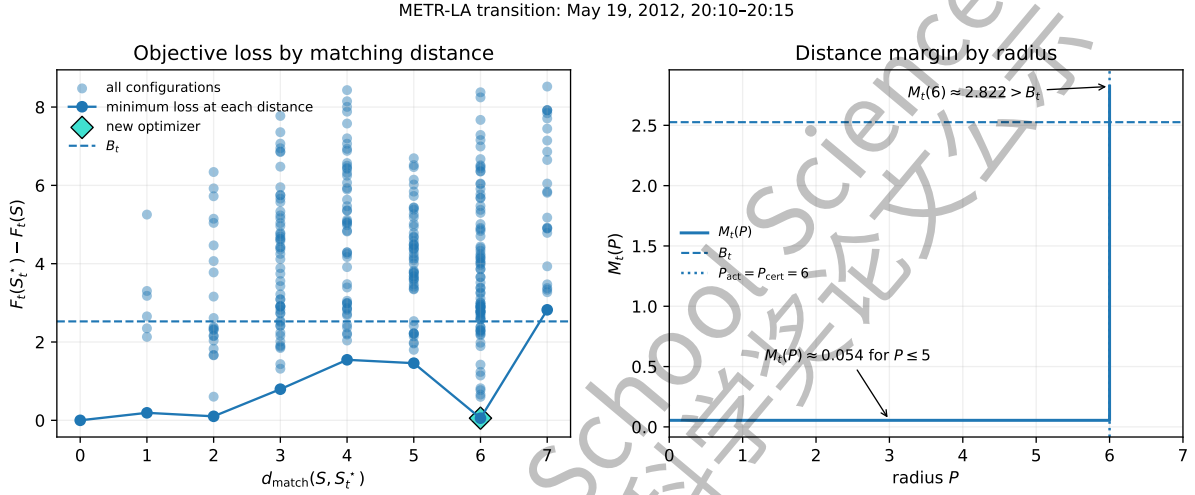


Figure 4: METR-LA transition from May 19, 2012, 20:10 to 20:15. Left: previous objective loss for all $\binom{25}{2} = 300$ configurations versus bottleneck matching distance from S_t^* ; the lowest point at each distance gives the best competitor, the diamond marks the new optimizer, and the dashed line is B_t . Right: distance margin curve $M_t(P)$. We see curve exceeds B_t at $P = 6$, giving $P_{\text{cert}} = P_{\text{act}} = 6 < P_{\text{max}} = 7$.

Together, these results show that the theorem applies to many real traffic transitions and that all eligible cases satisfy the certified movement bound. The case study shows why the distance margin matters. Projector drift alone does not determine movement; the distance margin of the old objective determines whether all farther competitors lie beyond the perturbation budget.

10. Discussion

Theorem eligibility, certificate tightness (how close P_{cert} is to P_{act}), and actual movement are different. An ineligible transition can still have $P_{\text{act}} = 0$, since the theorem gives sufficient rather than necessary conditions. An eligible transition can also have $P_{\text{cert}} > P_{\text{act}}$ when the bound is conservative. In the METR-LA case study, the certificate is both sharp and nontrivial, with $P_{\text{act}} = P_{\text{cert}} = 6 < P_{\text{max}}$. The traffic experiment also shows that the projector and movement graph can be defined separately: Q_t is learned from the traffic data, while the fixed road adjacency graph defines movement.

We also tested NOAA OISST as a real-world application [16, 24]. Under the spatial and temporal settings we tried, it was not majority of transitions that satisfied the projector perturbation conditions required for certification to make it a representative dataset for this experiment. We therefore did not adjust the setup to make the theorem apply more often. This shows that smooth physical change does not necessarily produce small projector drift, good conditioning, and a large enough distance margin (our hypothesis is that the perturbation was marginal, as temperature fluctuations are merely gradual over a long period of time; even taking monthly snapshots still had minimal perturbation).

The experiments themselves also have several limitations though. Computing $M_t(P)$ requires evaluating all configurations, so we use 25-sensor subnetworks with $m = 2$, giving $\binom{25}{2} = 300$ configurations. The traffic detectors are fixed, so movement refers to movable sensing resources rather than the installed detectors. The four hour windows are updated every five minutes and therefore overlap heavily. In the future, the hop metric can also take into account travel time, unequal edge lengths, and path constraints. Future work could study larger sensor sets and more realistically modeled movement models.

11. Conclusion

This paper studies stability using graph distance for dynamic D-optimal subset selection. Replacement count alone cannot entirely represent physical movement, as a single selected coordinate can change while the corresponding sensing resource moves across the entire graph. By measuring movement with bottleneck matching distance and measuring how much worse distant competitors are, we derive a movement bound from projector drift, conditioning, and the distance margin.

The main theorem is stated for orthogonal projectors and requires a lower eigenvalue bound only for the compressed projector at the current maximizer. Rather than imposing conditioning assumptions on all candidate subsets or on the unknown new maximizer, the proof uses optimality of the new maximizer to derive its own lower eigenvalue bound from the conditioning of the previous optimizer. The result therefore holds for every optimizer at the new time without requiring a tie-breaking rule or uniform conditioning over all candidate subsets. The Laplacian result follows from Davis–Kahan, while the weighted-path construction proves that perturbation size and eigengap alone are insufficient, so additional information about other competing configurations is necessary for a movement guarantee (the distance margin provides one such form of information). Exact finite graph computations then make the margin computable and connect the theorem to losslessness of P -hop and Voronoi-restricted optimization.

The synthetic experiments test the theorem by calculating the optimal sets, distance margins, actual matching movements, and certified radii by evaluating every sensor pair. They verify and illustrate the theorem’s implication and identify the matrix perturbation estimates and small objective gaps as the main reasons the certified radius can exceed the actual movement.

The traffic-data study extends the same analysis to projectors learned from real observations. PEMS-BAY serves as the main dataset and METR-LA as a second with more missing values and nonzero movements (more robustness and nontrivial movement to study). We do not construct time-varying road Laplacians; the experiment learns evolving empirical traffic subspaces from rolling windows of recorded speeds and applies the projector theorem directly. This lets us measure how often the theorem conditions hold, the certified radius, and the difference between actual and certified movement.

All in all, we don’t claim that every dynamic system has a small movement certificate (OISST testing already illustrates that this need not occur for certain data representation and time scale). Instead, we address: when a system’s projector drift, optimizer conditioning, and distance margin satisfy the theorem conditions, how far can a globally optimal configuration move? We provide an approach to stability questions in dynamic sensing and more general determinant-based subset-selection problems.

References

- [1] Anis, A., Gadde, A., Ortega, A.: Efficient sampling set selection for bandlimited graph signals using graph spectral proxies. *IEEE Trans. Signal Process.* 64(14), 3775–3789 (2016). <https://doi.org/10.1109/TSP.2016.2546233>
- [2] Avron, H., Boutsidis, C.: Faster subset selection for matrices and applications. *SIAM J. Matrix Anal. Appl.* 34(4), 1464–1499 (2013). <https://doi.org/10.1137/120867287>
- [3] Babecki, C.: Codes, cubes, and graphical designs. *J. Fourier Anal. Appl.* 27, 81 (2021). <https://doi.org/10.1007/s00041-021-09852-z>
- [4] Bhatia, R.: *Matrix Analysis*. Graduate Texts in Mathematics, vol. 169. Springer, New York (1997). <https://doi.org/10.1007/978-1-4612-0653-8>
- [5] Burkard, R., Dell’Amico, M., Martello, S.: *Assignment Problems*. SIAM, Philadelphia (2009)
- [6] California Department of Transportation: Performance Measurement System (PeMS). <https://pems.dot.ca.gov/>
- [7] Chen, S., Varma, R., Sandryhaila, A., Kovačević, J.: Discrete signal processing on graphs: Sampling theory. *IEEE Trans. Signal Process.* 63(24), 6510–6523 (2015). <https://doi.org/10.48550/arXiv.1503.05432>
- [8] Chung, F.R.K.: *Spectral Graph Theory*. American Mathematical Society, Providence (1997).
- [9] Çivril, A., Magdon-Ismael, M.: On selecting a maximum volume sub-matrix of a matrix and related problems. *Theor. Comput. Sci.* 410(47–49), 4801–4811 (2009). <https://doi.org/10.1016/j.tcs.2009.06.018>
- [10] Davis, C., Kahan, W.M.: The rotation of eigenvectors by a perturbation. III. *SIAM J. Numer. Anal.* 7(1), 1–46 (1970). <https://doi.org/10.1137/0707001>
- [11] Di Lorenzo, P., Banelli, P., Isufi, E., Barbarossa, S., Leus, G.: Adaptive graph signal processing: Algorithms and optimal sampling strategies. *IEEE Trans. Signal Process.* 66(13), 3584–3598 (2018). <https://doi.org/10.1109/TSP.2018.2835384>
- [12] Fiedler, M.: Eigenvectors of acyclic matrices. *Czechoslov. Math. J.* 25(4), 607–618 (1975). <https://doi.org/10.21136/CMJ.1975.101356>
- [13] Fiedler, M.: A property of eigenvectors of nonnegative symmetric matrices and its application to graph theory. *Czechoslov. Math. J.* 25(4), 619–633 (1975). <https://doi.org/10.21136/CMJ.1975.101357>
- [14] Gokavarapu, C., Gosala, S.B., Muthu, R., Pasalapudi, V., Kapakayala, T.: Finite-sample certification and operating envelopes for spectral clustering and graph centrality. *arXiv:2602.10566v2* (2026). <https://doi.org/10.48550/arXiv.2602.10566>
- [15] Horn, R.A., Johnson, C.R.: *Matrix Analysis*, 2nd edn. Cambridge University Press, Cambridge (2013). <https://doi.org/10.1017/cbo9781139020411>
- [16] Huang, B., Liu, C., Banzon, V., Freeman, E., Graham, G., Hankins, B., Smith, T.M., Zhang, H.-M.: Improvements of the Daily Optimum Interpolation Sea Surface Temperature (DOISST) Version 2.1. *J. Climate* 34(8), 2923–2939 (2021). <https://doi.org/10.1175/JCLI-D-20-0166.1>
- [17] Joshi, S., Boyd, S.: Sensor selection via convex optimization. *IEEE Trans. Signal Process.* 57(2), 451–462 (2009). <https://doi.org/10.1109/TSP.2008.2007095>

- [18] Kato, T.: Perturbation Theory for Linear Operators. Springer, Berlin (1995)
- [19] Krause, A., Singh, A., Guestrin, C.: Near-optimal sensor placements in Gaussian processes: Theory, efficient algorithms and empirical studies. *J. Mach. Learn. Res.* 9, 235–284 (2008)
- [20] Li, Y., Yu, R., Shahabi, C., Liu, Y.: Diffusion convolutional recurrent neural network: Data-driven traffic forecasting. In: International Conference on Learning Representations (ICLR) (2018). <https://arxiv.org/abs/1707.01926>
- [21] Li, Y.: DCRNN: Diffusion Convolutional Recurrent Neural Network benchmark data and code repository. GitHub. <https://github.com/liyaguang/DCRNN>
- [22] Lovász, L., Plummer, M.D.: Matching Theory. AMS Chelsea Publishing, Providence (2009).
- [23] Marques, A.G., Segarra, S., Leus, G., Ribeiro, A.: Sampling of graph signals with successive local aggregations. *IEEE Trans. Signal Process.* 64(7), 1832–1843 (2016). <https://doi.org/10.48550/arXiv.1504.04687>
- [24] NOAA National Centers for Environmental Information: NOAA 0.25-degree Daily Optimum Interpolation Sea Surface Temperature (OISST), Version 2.1. <https://www.ncei.noaa.gov/products/optimum-interpolation-sst>
- [25] Nomura, S., Hara, J., Higashi, H., Tanaka, Y.: Dynamic sensor placement based on sampling theory for graph signals. *IEEE Open J. Signal Process.* 5, 1042–1051 (2024). <https://doi.org/10.1109/OJSP.2024.3466133>
- [26] Ortega, A., Frossard, P., Kovačević, J., Moura, J.M.F., Vandergheynst, P.: Graph signal processing: Overview, challenges, and applications. *Proc. IEEE* 106(5), 808–828 (2018). <https://doi.org/10.48550/arXiv.1712.00468>
- [27] Pesenson, I.: Sampling in Paley–Wiener spaces on combinatorial graphs. *Trans. Am. Math. Soc.* 360(10), 5603–5627 (2008). <https://doi.org/10.48550/arXiv.1111.5896>
- [28] Pukelsheim, F.: Optimal Design of Experiments. Classics in Applied Mathematics. SIAM, Philadelphia (2006). <https://doi.org/10.1137/1.9780898719109>
- [29] Shuman, D.I., Narang, S.K., Frossard, P., Ortega, A., Vandergheynst, P.: The emerging field of signal processing on graphs. *IEEE Signal Process. Mag.* 30(3), 83–98 (2013). doi.org/10.1109/MSP.2012.2235192
- [30] Steinerberger, S.: Generalized designs on graphs: Sampling, spectra, symmetries. *J. Graph Theory* 93(2), 253–267 (2020). <https://doi.org/10.48550/arXiv.1803.02235>
- [31] Stewart, G.W., Sun, J.-G.: Matrix Perturbation Theory. Academic Press, Boston (1990).
- [32] Tsitsvero, M., Barbarossa, S., Di Lorenzo, P.: Signals on graphs: Uncertainty principle and sampling. *IEEE Trans. Signal Process.* 64(18), 4845–4860 (2016). <https://doi.org/10.48550/arXiv.1507.08822>
- [33] Yu, Y., Wang, T., Samworth, R.J.: A useful variant of the Davis–Kahan theorem for statisticians. *Biometrika* 102(2), 315–323 (2015). <https://doi.org/10.48550/arXiv.1405.0680>

Acknowledgements and Research Process

Research background and advisory relationship

I initiated the advisory relationship by cold emailing my advisor (November 2025), Laura Schaposnik, because I was interested in pursuing independent mathematical research. It was independent and was entirely unpaid.

At the beginning of our work together, Professor Schaposnik introduced me to the area of graphical designs, graph-based sampling, and related subset selection problems. One of the first papers I studied was Catherine Babecki’s “Codes, Cubes, and Graphical Designs,” which provided important background for our early discussions. From there, our conversations gradually moved toward questions about what happens when the underlying graph or signal structure changes over time, and I decided to explore some conjectures I had in that area.

During the early stages of the project, we met approximately once every two weeks while I learned the necessary background, read research papers, and explored possible research questions. As the project developed, we began meeting approximately once per week. In a typical meeting, I presented the work I had completed since our previous discussion, including mathematical arguments, examples, computational results, or difficulties I had encountered. Professor Schaposnik provided feedback when my results were not clear enough and suggested possible application settings in which the mathematical results could be tested.

Her role throughout the project was primarily advisory. I carried out the mathematical derivations, proof development, computational implementation, data analysis, experimental work, and manuscript writing. Professor Schaposnik provided guidance and feedback and, during the final stages, reviewed manuscript drafts and suggested areas where additional explanation was needed or where equations could be presented differently (i.e. in a separate line or with the rest of the paragraph).

Research process, contributions, division of labor, and use of AI

The project started with a question about dynamic graph sampling and developed through a failed conjecture, a counterexample, and the resulting stability theory presented in this paper.

Origin of the research question: at the beginning of the project, my advisor introduced me to graphical designs and related questions concerning how a small subset of vertices can capture information about functions on a graph. I reviewed relevant linear algebra, particularly eigenvalues, eigenvectors, and graph Laplacians, and began by studying Catherine Babecki’s work on graphical designs and broader references such as Shuman et al.’s “The Emerging Field of Signal Processing on Graphs.”

During one of our discussions, my advisor and I wondered what would happen if the underlying system evolved: specifically, whether a good sampling or sensor configuration would remain geographically close to optimal when graph weights, and consequently the relevant low dimensional subspace, changed. This became the starting question for the project.

Development of the theoretical results: our initial conjecture was that a sufficiently small perturbation, relative to an appropriate multiple of the eigengap, should prevent the optimal configuration from moving very far. I attempted to prove this statement while also testing it on synthetic graphs. When increasing the sizes of these examples, I found the conjecture was false through an $n = 100$ case experiment. My first reaction was that I had made a calculation or code error, since the result contradicted both my intuition and the pattern consistent across many runs of smaller cases. Movement was unexpectedly large, and did not remain under my conjectured bound directly proportional to eigengap. I reran the experiment with different and even larger sizes, checked the code logistics, and verified each step of the code computation independently before realizing that something was truly wrong with my intuition. This experiment changed the direction of the project.

I then started investigating why the conjecture failed rather than discarding the unsuccessful experiments. This eventually led to the weighted-path construction in the paper, which demonstrates that arbitrarily small perturbations at a fixed baseline eigengap can move an optimizer across an arbitrarily long graph, formalizing into this paper’s impossibility result. The counterexample revealed that projector drift alone is

insufficient: even when the underlying projector changes only slightly, a distant competing configuration can become optimal if it has nearly the same objective value.

Then, I reformulated the problem by asking what additional information was needed to control optimizer movement. This led me to the distance margin, which evaluates how much worse distant competing configurations are than the current optimum. Important references during this stage included the Davis–Kahan $\sin \Theta$ theorem, Stewart and Sun’s work on matrix perturbation theory and invariant subspaces, and Nomura et al.’s work on dynamic sensor placement using learned dictionaries, local neighborhoods, and Voronoi restrictions. My advisor looked over my arguments with me and identified places requiring additional justification; I constructed the examples and counterexamples and developed and checked the mathematical arguments and proofs.

Computational and experimental development: after developing the theoretical basis, I designed controlled synthetic experiments in which all configurations could be evaluated. These experiments computed actual optimal configurations, distance margins, optimizer movement, and certified movement radii, allowing the conclusions and conservatism of the theorem to be examined directly.

Finding an appropriate real-world application was more difficult. I investigated NOAA sea-surface-temperature data and several air-quality datasets after discussing possible applications with my advisor. In particular, the NOAA data didn’t satisfy some of the theorem’s earlier hypotheses when I tested. Rather than selecting the marginally few cases where my theorem could apply for that dataset, I did not use it as evidence supporting the theorem.

I later found the currently used traffic sensor data as a more suitable setting because it provides fixed physical sensor locations together with a low-dimensional traffic signal structure that changes over time. I selected the PEMS-BAY and METR-LA datasets from the public DCRNN repository. I investigated their structure, preserved timestamps and sensor identifiers, and designed the preprocessing and experimental protocol. I used ChatGPT as a coding aid for portions of the raw data parsing and asked it to evaluate some of my experiment ideas, while thoroughly reviewing the generated code, running the analyses, checking intermediate outputs, and determining the mathematical and experimental procedures used in the final study.

The final traffic experiments used fixed sensor subnetworks, consecutive traffic windows, singular-value decomposition to construct empirical low-dimensional projectors, exhaustive evaluation of sensor configurations, and direct computation of optimizer movement and certification quantities. A large experimental difficulty was avoiding making choices that aimed to simply make the theorem appear more successful than it is. Thus, I fixed certain preprocessing and baseline parameters based on purely the structure and quality of the traffic data rather than theorem outcomes and kept even the failed or ineligible transitions in the final analysis.

Writing, revision, and division of labor: I wrote the manuscript, including the literature discussion, mathematical arguments and proofs, experimental methodology, figures and tables, and interpretation of the results. Professor Schaposnik reviewed later drafts and provided feedback on clarity, organization, places requiring additional explanation, and the presentation of some equations. I revised the manuscript in response. I also used ChatGPT to help identify grammatical or stylistic issues and, as described above, as an aid during portions of coding (processing datasets); I remained responsible for checking these outputs and for the mathematical arguments, computational results, and final text.

Overall, I was responsible for the literature review, development and testing of conjectures, construction of the counterexample, formulation and proof of the theoretical results, data acquisition and preprocessing, coding and verification, experimental design and analysis, preparation of figures and tables, and manuscript writing. Professor Schaposnik served in an advisory role by introducing the initial research area, participating in the discussion that motivated the dynamic question, providing mathematical feedback, suggesting possible application directions, and commenting on manuscript drafts.

AI Usages:

Tool/model: ChatGPT, GPT-5.5 / GPT-5.6 Sol

Period: November 2025–September 2026

Frequency: regular use during active research periods, primarily for background learning, code

debugging, and manuscript proofreading; less frequent use during inactive periods.

Uses: mathematical conceptual explanation, looking for math textbooks and resources, checking mathematical reasoning already developed by the student, LaTeX/template formatting, inspection and parsing of public traffic datasets, implementation/debugging of preprocessing code based on student-specified instructions/methodology, grammar and manuscript-error checking.

Did not use AI: topic, theorem statements, proofs/core arguments, experimental questions and final methodological decisions, execution/verification of experiments, interpretation of results, and final manuscript content were independently determined, written, and verified by the student.