Consistent Soundscape Connectomes via Stability-Refined Graph Learning

Maria J. Guerrero*§, Aref Einizade†, Jhony H. Giraldo‡, César A. Uribe§
* SISTEMIC, Engineering Faculty, Universidad de Antioquia, Medellín, Colombia
† SAMOVAR, Télécom SudParis, Institut Polytechnique de Paris, Evry, France
‡ LTCI, Télécom Paris, Institut Polytechnique de Paris, Palaiseau, France
§ Department of Electrical and Computer Engineering, Rice University, Houston, TX, USA

***Abstract*—Passive Acoustic Monitoring (PAM) enables large-scale biodiversity assessment using networks of autonomous recorders. A natural way to summarize relationships across sites is to learn a weighted graph whose nodes are recorders and whose edges encode acoustic similarity. However, standard Graph Learning (GL) methods can be structurally unstable: removing a node (e.g., due to sensor failure) can trigger global rewiring, changing the graph's connectivity and undermining ecological interpretation and longitudinal comparisons.**

We propose a prior-guided refinement approach that learns (or updates) the graph on the surviving nodes while (i) promoting structural similarity to a reference/prior graph and (ii) enforcing signal smoothness so that acoustically similar sites remain consistently connected. We evaluate the method on synthetic benchmarks and real PAM data and show that it yields more consistent graphs under node dropouts, preserving prior structural patterns while maintaining smoothness, thereby improving robustness and ecological interpretability.

***Index Terms*—Graph signal processing, bioacoustic monitoring, passive acoustic monitoring, learning graphs from data, topological consistency.**

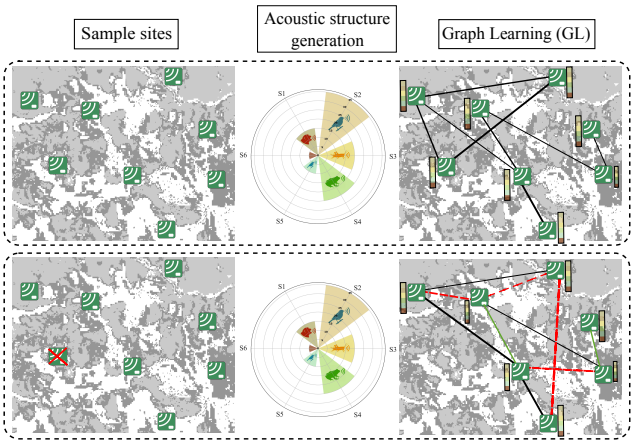

Fig. 1. Inconsistencies in the soundscape connectomes generated via graph learning in the presence of node dropouts. Site-level acoustic structures (from sonotype composition [8]) are used to learn a graph; re-learning after sensor loss may change edges (black: unchanged, red dashed: new, green: missing).

I. Introduction

Passive Acoustic Monitoring (PAM) [1], [2] enables scalable, long-term, and non-invasive biodiversity assessment by deploying autonomous recorders across various landscapes to record local biophony (the collective vocal activity of animal communities). These acoustic signals reflect ecological processes and variations at specific sites, making them a valuable tool for monitoring ecosystem dynamics and guiding conservation planning [3], [4]. Recent studies have shown that different land-cover types have unique acoustic signatures [5], [6], and that sound-based features can help discriminate among them. This approach offers a promising complement to remote sensing and traditional field surveys [7], [8].

Graph Learning (GL)[1] [9], [10] offers a suitable framework for modeling the relationships among ecological sampling sites. In this setting, *soundscape connectomes* represent sites as nodes and connect them through edges that encode acoustic similarity computed from site-level sound descriptors (e.g., acoustic structures derived from sonotype composition [8], [11]). These representations may facilitate ecological interpretation and help reveal spatial structure in vocal communities [12], [13].

As illustrated in Fig. 1, the graph is learned from the available site descriptors. However, when a single recorder becomes unavailable (e.g., due to device failure or corrupted data), many standard approaches for learning graphs from data, including sparse inverse covariance estimation [14] and Laplacian-constrained smoothness models [15], can be sensitive to the specific set of observed nodes. Since these models often optimize global signal properties [10], removing a node can induce structural rebalancing, leading to an unstable topology. Consequently, edges between the remaining sites may inconsistently appear, disappear, or both.

In fields like neuroscience, changes in graph structure after node removal may be acceptable, as virtual lesions or ablation studies explicitly investigate how the absence of nodes affects brain network function [16], [17]. In contrast, ecological networks derived from PAM data aim to preserve consistent inter-site relationships even in the presence of occasional sensor dropout, since biological similarities among locations are driven by underlying site conditions rather than the set of

This research was supported by Universidad de Antioquia–CODI and RICE University (code project: 2024-73410), the ANR project DeSNAP (ANR-24-CE23-1895-01), and the National Science Foundation under Grants #2213568 and #2443064.

[1]In this paper, we use *graph learning* to mean *learning graphs from data*, such as graph topology/structure inference as commonly studied in graph signal processing.

available recorders. Accordingly, the inferred topology should remain as consistent as possible when nodes are removed, especially in long-term monitoring scenarios where sensor loss (failure) is common [18], [19].

To address topology changes induced by missing nodes, we introduce a GL Refinement (GL-Ref) objective to enhance topological consistency when nodes are removed. Instead of reapplying existing GL methods, our approach refines their outputs by augmenting the objective function to encourage similarity to a prior graph structure while maintaining signal smoothness. Evaluations on both synthetic and real PAM datasets show consistent improvements in preserving graph structure, supporting more reliable ecological interpretation.

Our contributions are as follows: *i)* We propose a convex GL-Ref objective function that combines signal smoothness with a similarity-preserving regularizer. This approach improves the topological consistency of existing GL methods when nodes are removed. *ii)* We incorporate GL-Ref into a real PAM pipeline, using sonotype-based acoustic features extracted from real field deployments. *iii)* Our results demonstrate that this refinement leads to more stable topological structures and improved edge consistency compared to GL baselines such as Graphical Lasso [14] and Laplacian-constrained smoothness models [15], [20].

II. Graph Learning Refinement (GL-Ref)

Preliminaries. We represent a graph by $G = (\mathcal{V}, \mathcal{E}, \mathbf{A})$, where $\mathcal{V}$ is the node set with $|\mathcal{V}| = N$, $\mathcal{E} \subseteq \{(i,j) \mid i,j \in \mathcal{V}, i \neq j\}$ is the edge set, and $\mathbf{A} \in \mathbb{R}^{N\times N}$ is a (symmetric) weighted adjacency matrix. We focus on connected, undirected, weighted graphs. We also define the Laplacian as $\mathbf{L} = \mathbf{D} - \mathbf{A}$, with $\mathbf{D} = \mathrm{diag}(\mathbf{A}\mathbf{1})$ as a diagonal degree matrix, where $\mathbf{1}$ denotes the all-ones vector (dimension implied by context).

For a removed node $s \in \{1, \dots, N\}$, let $\mathbf{P}_s \in \{0,1\}^{(N-1)\times N}$ denote the selector that deletes the s-th row from $\mathbf{A}$ to model the effect of removing this node. Then, by collecting node feature vectors on G as $\mathbf{X} \in \mathbb{R}^{N\times F}$ with F being the number of node features, the node removal gets the reduced node feature matrix $\tilde{\mathbf{X}} \in \mathbb{R}^{(N-1)\times F}$ as $\tilde{\mathbf{X}} = \mathbf{P}_s\mathbf{X}$, and use $\mathbf{A}_{\text{proj}}$ to denote the projected prior adjacency ($\mathbf{A}_{\text{prior}}$) on the observed nodes as $\mathbf{A}_{\text{proj}} = \mathbf{P}_s\,\mathbf{A}_{\text{prior}}\,\mathbf{P}_s^\top \in \mathbb{R}^{(N-1)\times(N-1)}$.

A. Problem Statement

Let $\mathbf{X} \in \mathbb{R}^{N\times F}$ be the feature matrix that collects F graph signals described by a generative process f on the prior graph G with adjacency $\mathbf{A}_{\text{prior}}$ as $\mathbf{X} = f(G)$. Given a new (reduced) version of $\mathbf{X} \in \mathbb{R}^{N\times F}$ as $\tilde{\mathbf{X}} \in \mathbb{R}^{(N-1)\times F}$, obtained from removing (disconnecting) node s from the prior graph G with graph signals collected in $\mathbf{X}$, we aim to learn a graph G_{r} with adjacency $\tilde{\mathbf{A}}$ such that $\mathbf{A}_{\text{proj}} \approx \tilde{\mathbf{A}}$ and it can represent the graph signals in $\tilde{\mathbf{X}}$ as $\tilde{\mathbf{X}} = f'(G_{\text{r}})$. We assume the standard graph-signal smoothness prior used in GL [15], i.e., the signals generated by both f and f' exhibit small local variations with respect to the underlying Laplacian (equivalently, they have low Laplacian quadratic form), as in [15], [20].

Our goal is to reduce topology inconsistencies in GL, characterized by the introduction or loss of edges after node dropout in the new graph. Inferring soundscape connectomes (in terms of structural similarity to the prior graph) can serve as a practical proof of concept of their applicability.

B. Optimization Problem

Figure 2 shows an overview of GL-Ref. We learn the adjacency of the reduced graph $\tilde{\mathbf{A}} \in \mathbb{R}^{(N-1)\times(N-1)}$ on the observed nodes by balancing two objectives: *i) signal smoothness* on the observed features imposed by f' and *ii) structural similarity* to a projected prior $\mathbf{A}_{\text{proj}}$. We define the pairwise squared-distance matrix $\mathbf{Z} \in \mathbb{R}^{(N-1)\times(N-1)}$, where $\mathbf{Z}_{ij} = \|\tilde{\mathbf{x}}_i - \tilde{\mathbf{x}}_j\|_2^2$, for $i,j = 1,\dots,N-1$. The smoothness of signals in $\tilde{\mathbf{X}}$ on a graph with Laplacian $\tilde{\mathbf{L}}$ can be written as [15]:

$$\mathrm{tr}\big(\tilde{\mathbf{X}}^\top\tilde{\mathbf{L}}\tilde{\mathbf{X}}\big) \;=\; \tfrac{1}{2}\sum_{i,j}\tilde{\mathbf{A}}_{ij}\,\|\tilde{\mathbf{x}}_i - \tilde{\mathbf{x}}_j\|_2^2 \;=\; \tfrac{1}{2}\langle \mathbf{Z}, \tilde{\mathbf{A}}\rangle, \qquad (1)$$

where $\langle\cdot,\cdot\rangle$ denotes the Frobenius inner product. This identity allows expressing the smoothness term linearly in $\tilde{\mathbf{A}}$ via $\mathbf{Z}$. By taking both signal smoothness and structural similarity into account, we propose the following objective to optimize:

$$\begin{aligned} \min_{\tilde{\mathbf{A}},\tilde{\mathbf{L}}} \quad & \mathrm{tr}\big(\tilde{\mathbf{X}}^\top\tilde{\mathbf{L}}\tilde{\mathbf{X}}\big) \;+\; \lambda\,\|\tilde{\mathbf{A}} - \mathbf{A}_{\text{proj}}\|_F \\ \text{s.t.} \quad & \tilde{\mathbf{A}} = \tilde{\mathbf{A}}^\top, \quad \tilde{\mathbf{A}} \geq 0, \quad \mathrm{diag}(\tilde{\mathbf{A}}) = \mathbf{0}, \\ & \tilde{\mathbf{L}} = \mathrm{diag}(\tilde{\mathbf{A}}\mathbf{1}) - \tilde{\mathbf{A}},\ \tilde{\mathbf{L}}\mathbf{1} = \mathbf{0}, \\ & \tilde{\mathbf{L}} \succeq 0,\ \tilde{\mathbf{L}}_{ij} \leq 0 \text{ for } i \neq j, \end{aligned} \qquad (2)$$

where $\lambda > 0$ is a regularization parameter, and the constraints ensure the validity of $\tilde{\mathbf{L}}$ and $\tilde{\mathbf{A}}$ as a Laplacian and adjacency matrix, respectively. The problem described in the objective (2) is convex, because the smoothness term is linear in $\tilde{\mathbf{A}}$, as shown in equation (1), the Frobenius norm is convex, and the feasible set is also convex.

We solve (2) with a first-order operator-Splitting Conic Solver (*SCS*) [21]. The optimization involves $\mathcal{O}((N-1)^2)$ decision variables (the entries of $\tilde{\mathbf{A}} \in \mathbb{R}^{(N-1)\times(N-1)}$) and a semidefinite constraint $\tilde{\mathbf{L}} \succeq 0$, which makes the overall computational cost grow with N; accordingly, the runtime is expected to scale roughly as $\mathcal{O}(N^3)$.

Finally, to obtain an unweighted graph for binary/topological evaluations, we prune near-zero weights introduced by solver tolerance using a fixed τ, i.e., we set $\tilde{A}_{ij} = 0$ if $\tilde{A}_{ij} < \tau$ (a common practice in GL; *e.g.*, [15]).

III. Experiments and Results

We evaluate our method on both synthetic (Erdős–Rényi (ER) graphs) and real (PAM) datasets, as follows:

A. Datasets

Synthetic dataset. We generate ground truth ER graphs with the adjacency $\mathbf{A}_{\text{gt}} \in \mathbb{R}^{N\times N}$ and Laplacian $\mathbf{L}$, using a connection probability $p = 2\frac{\log(N)}{N-1}$, which yields connected graphs without isolated vertices [22]. Smooth graph signals are sampled from the Laplacian spectrum following the approach

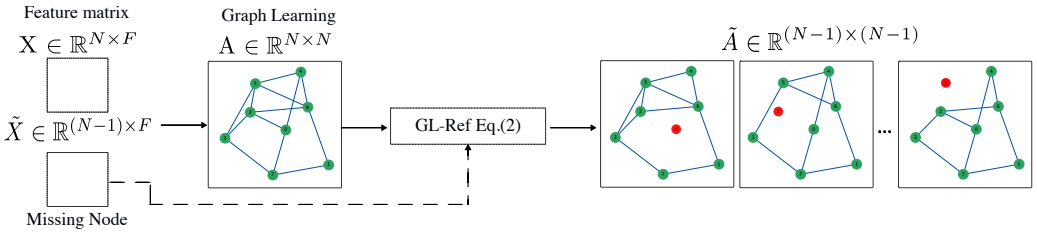

Fig. 2. Overview of the proposed refinement under node removal. Given node features $\mathbf{X}$, a graph $\mathbf{A}$ is first inferred. Upon node removal, the reduced feature matrix $\tilde{\mathbf{X}}$ is used to estimate $\tilde{\mathbf{A}}$ via convex optimization, enforcing signal smoothness and structural consistency with the subgraph $P_s A_{\text{prior}} P_s^\top$.

in [20]. Precisely, using the Eigenvalue Decomposition (EVD) of the Laplacian as $\mathbf{L} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^\top$, the smooth signals are generated as $\mathbf{X} = \mathbf{V}\mathbf{H} + \boldsymbol{\varepsilon}$, where $\mathbf{H} \sim \mathcal{N}(\mathbf{0}, \mathbf{\Lambda}^\dagger)$, and $\boldsymbol{\varepsilon}$ is additive Gaussian noise that is scaled to achieve target Signal-to-Noise Ratio (SNR) levels.

Real-world PAM dataset. We consider a network of N passive acoustic sensors deployed across a region, where each site (node) i is represented by an acoustic feature vector $\mathbf{x}_i \in \mathbb{R}^F$ summarizing the site's local biophony. Following the unsupervised soundscape-to-graph pipeline [6], we extract *sonotypes*, which are acoustic entities containing information in both time and frequency that may correspond to animal vocalizations, and use them to construct per-site *acoustic structures* that act as ecosystem fingerprints. Stacking all site vectors yields the node feature matrix $\mathbf{X} \in \mathbb{R}^{N\times F}$ [8].

The acoustic data used in this study come from two distinct PAM deployments conducted in different regions of Colombia. The first dataset was acquired in Puerto Wilches–Santander, where 17 Song Meter Mini recorders were deployed. Devices recorded 1-minute clips every 10 minutes at 48 kHz, producing about 19,598 recordings over 10 days. The second dataset was collected in a rural region of Alejandría–Antioquia, encompassing ecosystems along the slopes of the Andes. This deployment used 59 AudioMoth recorders configured to capture 1-minute recordings every 15 minutes at a sampling rate of 192 kHz, yielding approximately 91,993 recordings collected over 18 days. Both datasets were preprocessed using filtering and activity-detection steps to emphasize biophonic content before sonotype extraction [8].

B. Evaluation Metrics

Our primary consistency metric is the per-removal Normalized Mean Squared Error (NMSE) between the inferred adjacency in the reduced space and the projected reference.

Let $\tilde{\mathbf{A}}$ be the adjacency learned on the $N-1$ observed nodes. With the projection operator $\mathbf{P}_s$, we define NMSE_s as:

$$\text{NMSE}_s \;=\; \frac{\left\|\tilde{\mathbf{A}} - \mathbf{P}_s\mathbf{A}\mathbf{P}_s^\top\right\|_F}{\left\|\mathbf{P}_s\mathbf{A}\mathbf{P}_s^\top\right\|_F}.$$

For the synthetic dataset, we also report edge-recovery ROC-AUC and topological deltas, including the absolute difference in clustering coefficient, average shortest-path length, betweenness centrality, and degree distribution distance [23], [24]. For averaging over multiple realizations, we generated 50 random trials. For the PAM experiment, we report NMSE and complement the numeric results with a topological interpretation of inferred connectomes.

C. Implementation Details

All experiments follow a cross-validation protocol to select λ. We perform a grid search and choose the value that minimizes the mean per-removal NMSE across trials (20–80% dropout); once selected, $\lambda^\star$ is fixed for the final evaluation (metrics are computed per removed node and reported as means and dispersion). We used the CVXPY [25], [26] package to optimize (2) with the SCS solver. The implementation codes are available at https://github.com/mariajguerrerom/GL-REF.

D. Synthetic Experiment

In our study, we consider SNR values of $\{\infty, 20, 0, -5\}$ dB and vary N by selecting values from the set $\{20, 40, 60, 80, 100\}$, which is representative of the range of sensor counts typically encountered in passive acoustic monitoring (PAM) datasets, from small to relatively large field deployments. Additionally, we evaluate different node dropout regimes (20%, 40%, 60%, and 80%) by performing node-by-node removals on sampled subsets.

For each removed node s, we (i) form $\tilde{\mathbf{X}} = \mathbf{P}_s\mathbf{X}$, (ii) set the projected ground truth $\mathbf{A}_{\text{proj}} = \mathbf{P}_s\mathbf{A}_{\text{gt}}\mathbf{P}_s^\top$, and (iii) solve the optimization (2) to obtain $\tilde{\mathbf{A}}$. Then, it is remapped via $\mathcal{R}_s(\tilde{\mathbf{A}})$ and compared to $\text{Mask}_s(\mathbf{A}_{\text{gt}})$ (prior graph with isolated target node) to compute the evaluation metrics.

Table I–(a) summarizes the topological consistency across dropout percentages for $N = 100$, SNR $= \infty$. NMSE remains moderate, and topological deltas (absolute difference between the ground-truth and the inferred graph) remain small across the tested removal rates (20%–80%), indicating that the learned graphs preserve their global structure despite extensive node removals. Table I–(b) isolates the effect of measurement noise for the 60% dropout case (and $N = 100$). As expected, reconstruction quality degrades with increasing noise: NMSE increases and AUC for edge recovery decreases as SNR falls from ∞ to -5 dB. Moreover, the proposed method still produces reasonable topological fidelity at moderate noise levels (20 dB, 0 dB), and only under the most adverse noise (-5 dB) do both edge-wise and global statistics deteriorate markedly. This behavior was expected and shows that consistency is achievable while the SNR changes.

TABLE I

CONSISTENCY METRICS (IN TERMS OF NMSE, AUC, AND THE DIFFERENCE BETWEEN CLUSTERING COEFFICIENT ($\delta_{\mathcal{C}}$), BETWEENNESS CENTRALITY ($\delta_{\mathcal{B}}$), PATH LENGTH ($\delta_{\mathcal{P}}$), DEGREE DISTRIBUTION ($\delta_{\mathcal{D}}$) OF THE LEARNED AND PRIOR GRAPHS) IN TWO SETTINGS (ST): (A) VARYING DROPOUT LEVELS WITH INFINITE SNR; (B) VARYING NOISE LEVELS WITH 60% DROPOUT.

St	NMSE ↓	AUC ↑	$\delta_{\mathcal{C}}$ ↓	$\delta_{\mathcal{B}}$ ↓	$\delta_{\mathcal{P}}$ ↓	$\delta_{\mathcal{D}}$ ↓
(a) Dropout levels @ Infinite SNR						
20%	0.314	0.862	0.038	0.003	0.425	0.097
40%	0.256	0.881	0.018	0.003	0.438	0.381
60%	0.207	0.909	0.025	0.002	0.246	0.008
80%	0.226	0.895	0.025	0.002	0.292	0.013
(b) Noise levels (dB) @ 60% Dropout						
∞	0.207	0.909	0.025	0.002	0.246	0.008
20	0.273	0.864	0.028	0.003	0.356	0.198
0	0.501	0.742	0.051	0.005	0.820	0.906
-5	0.544	0.704	0.056	0.005	0.761	1.523

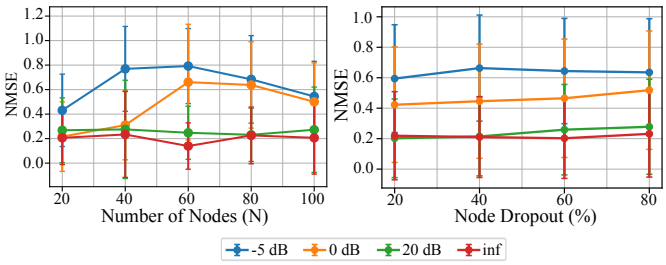

Fig. 3. Synthetic NMSE results. The left panel shows the NMSE as a function of the number of nodes N under a fixed node dropout of 60%, while the right panel reports the NMSE as a function of the node dropout percentage for a fixed graph size $N = 100$. Colors indicate the SNR values, with SNR $= \infty$ denoting the noise-free case.

Figure 3 illustrates two complementary views of the reconstruction error for a fixed graph configuration. The left panel reports the NMSE as a function of the number of nodes N, while the right panel shows the NMSE as a function of the node dropout percentage. Each curve corresponds to a different noise level introduced in the synthetic graph signals, with colors indicating the SNR values $\{-5, 0, 20, \infty\}$ dB, where SNR $= \infty$ denotes the noise-free case. As N increases, the NMSE remains bounded across all SNR regimes, indicating that the proposed approach scales stably with graph size. Similarly, increasing the node dropout fraction results in gradual degradation in reconstruction performance, particularly under severe noise conditions, while the noise-free and high-SNR cases remain relatively stable. Error bars reflect variability across multiple graph realizations and node-removal samplings. Together, the numerical and graphical results support the conclusion that the proposed strategy promotes structural consistency under node dropout across a wide range of practical settings.

E. Real PAM Experiment

Given the real site-by-sonotype feature matrix $\mathbf{X}$, we first compute full-data priors $\mathbf{A}_{\text{prior}}$ using the following selected graph learners: (i) Sparse statistical (Glasso) [14]: estimates a sparse precision matrix (the inverse covariance); (ii) Smooth graph learner (GL-SigRep) [15], [20]: learns the graph so that site signals vary little across connected nodes. (iii) k–Nearest Neighbors (kNN): builds a local graph where each site connects to its k closest Euclidean neighbors. For each node s, we run two procedures: (a) baseline leave-one-out re-inference by reapplying the same graph learner to $\tilde{\mathbf{X}}$ (this yields the baseline adjacency for the reduced set), and (b) our proposed GL-Ref method, where we set $\mathbf{A}_{\text{proj}} = \mathbf{P}_s \mathbf{A}_{\text{prior}} \mathbf{P}_s^\top$ and solve (2) to obtain $\tilde{\mathbf{A}}$. All outputs are remapped with $\mathcal{R}_s(\cdot)$ and evaluated using NMSE and complementary topology-based analysis. We select λ from the search grid by minimizing the mean NMSE across removals and report comparisons between the baseline re-inference and the stability-refined results. Finally, we use the learned connectomes and sonotype composition to support qualitative ecological interpretation.

We report results on two PAM datasets (Santander and Antioquia) using the same analysis framework; differences arise from the underlying soundscape composition and spatial configuration.

Puerto Wilches–Santander dataset: Figure 4 illustrates a topology comparison for the PAM deployment, where panel **(A)** shows the baseline graph inferred using Graphical Lasso on full $\mathbf{X}$, panel **(B)** shows the baseline re-inference after removing a single node (the empirical instability of the baseline method), and panel **(C)** shows the stability-refined result produced by our optimization (2). The refinement tends to preserve the neighborhood structure of the observed nodes while avoiding spurious rewiring. Node colors indicate habitat types, positioned according to their exact geographical locations; the red node denotes the removed site, blue edges correspond to expected connections, red edges to unexpected connections, and dashed light-blue edges to missing links.

Table II summarizes the mean NMSE across the 17 leave-

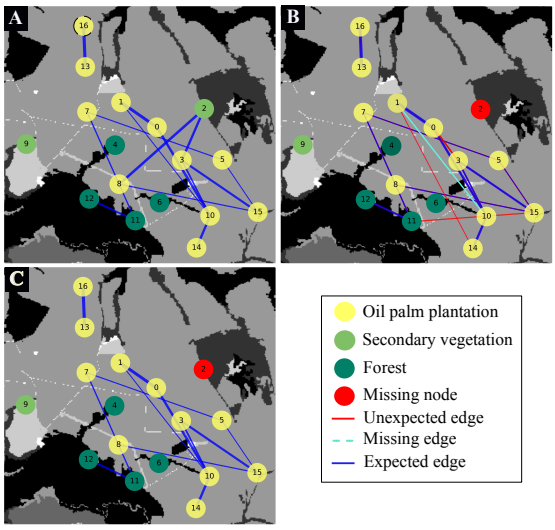

Fig. 4. Topology comparison for the Puerto Wilches–Santander PAM deployment. A) learned graph on full data, B) baseline re-inference after node removal, C) graph after stability-refinement.

TABLE II
MEAN NMSE ACROSS 17 NODE-REMOVAL EXPERIMENTS FOR EACH METHOD, BEFORE AND AFTER APPLYING THE REFINEMENT APPLIED TO THE PUERTO WILCHES-SANTANDER DATASET. LOWER IS BETTER.

Method	GL (Baseline)	→	GL-Ref (Proposed)
Glasso [8], [14]	0.22	→	**0.02**
GL-SigRep [15], [20]	0.58	→	**0.37**
kNN	0.16	→	**0.12**

one-out experiments, comparing baselines with and without the proposed refinement. The results indicate that our method yields clear reconstruction improvements, particularly when the prior graph provides a consistent, smooth structure [20].

By stabilizing the topology inference (GL) under node removal, our refinement helps to preserve significant inter-site connections that reflect consistent sonotype dynamics. This approach ensures that the resulting connectomes remain ecologically interpretable, as acoustically linked sites share biophonic patterns. This relationship was previously demonstrated in [8].

Alejandría–Antioquia dataset: We report results on a second PAM dataset collected in the Andean region of Alejandría (Antioquia), which exhibits higher landscape heterogeneity and was acquired with a different recorder model and sampling rate. Figure 5 summarizes the qualitative effect of node removal on the inferred connectome. Panel **(A)** shows the graph learned from the full dataset, revealing spatially coherent acoustic relationships across land-cover types. Panel **(B)** illustrates the baseline re-inference after removing one node, where the learned topology changes with respect to the full-data graph, introducing new links. Panel **(C)** shows the refined graph produced by GL-Ref, which reduces these edge changes and largely preserves the dominant inter-site connectivity patterns despite the missing node.

TABLE III
MEAN NMSE ACROSS 59 NODE-REMOVAL EXPERIMENTS FOR EACH METHOD, BEFORE AND AFTER APPLYING THE REFINEMENT APPLIED TO THE ALEJANDRÍA-ANTIOQUIA DATASET. LOWER IS BETTER.

Method	GL (Baseline)	→	GL-Ref (Proposed)
Glasso [8], [14]	0.18	→	**0.16**
GL-SigRep [15], [20]	0.16	→	**0.03**
kNN	0.12	→	**0.07**

Table III reports the mean NMSE across the 59 node-removal experiments for the Alejandría dataset. Compared to the baseline methods, the proposed refinement consistently reduces reconstruction error, with particularly pronounced improvements for the three methods. These results indicate that stabilizing the inferred topology under node removal yields more reliable signal reconstruction, especially when the initial graph captures smooth, ecologically meaningful acoustic relationships.

Overall, the Alejandría results (both qualitative topology comparisons and mean NMSE) support that GL-Ref improves topological consistency under node removal while preserving an interpretable acoustic connectome. This setting is particularly challenging in heterogeneous Andean landscapes, where soundscape composition can change over short distances. Moreover, as deployments scale to denser sensor networks, the likelihood of missing recordings due to device malfunction, corrupted/lost files, or incomplete retrieval increases, making stable inference under node loss practically important.

F. Sensitivity analysis and running time.

Regularization parameter selection and sensitivity. We select the regularization parameter λ by grid search as is described in section III-C, Table IV reports the mean NMSE obtained for each candidate λ (for $N = 100$ and SNR $= \infty$) across dropout regimes. The best-performing λ is consistent across dropouts (here, $\lambda = 2$), indicating low sensitivity within the tested grid.

TABLE IV
SYNTHETIC EXPERIMENT ($N = 100$, SNR $= \infty$): MEAN NMSE FOR EACH CANDIDATE λ ACROSS DROPOUT REGIMES. LOWER IS BETTER.

Dropout (%)	$\lambda = 0.01$	0.1	0.2	1	2	5
20	0.781	1.000	0.466	0.397	**0.070**	1.514
40	1.000	0.893	0.566	0.372	**0.048**	1.000
60	0.949	0.645	0.544	0.242	**0.019**	0.989
80	0.961	0.980	0.777	0.325	**0.038**	1.296

Running time and scalability. We report runtime measurements for the *synthetic* experiment, using the same leave-one-out node-removal protocol described in Sec. III-C. The observed runtime increases with the graph size N, consistent with the polynomial scaling expected for the conic optimization problem in (2) (Sec. II-B). For dropout $= 80\%$, Table V summarizes the mean runtime for $N \in \{20, 40, 60, 80, 100\}$ and includes an extrapolation to $N = 1000$ (denoted by $\approx$). In practice, PAM deployments for this application typically involve networks with tens to ~ 100 recorders. Larger-scale validation is therefore left for future work, as public PAM datasets with more than 100 sensors remain uncommon due to the cost of large deployments and the still-growing availability of large-scale releases.

TABLE V
RUNTIME VS. GRAPH SIZE (DROPOUT = 80%).

N	20	40	60	80	100	1000
Time [s]	1.523	7.230	21.936	55.726	128.769	$\approx 5.54 \times 10^4$

IV. CONCLUSIONS

Our findings show that incorporating a similarity-preserving regularizer into smoothness-based Graph Learning (GL) improves the topological consistency of inferred structures under node removal. This is particularly important for ecological applications that use Passive Acoustic Monitoring (PAM),

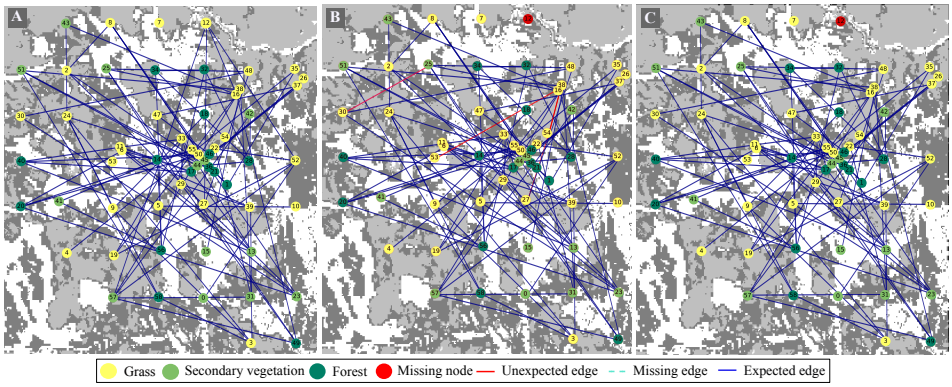

Fig. 5. Topology comparison for the Alejandría–Antioquia PAM deployment. A) learned graph on full data, B) baseline re-inference after node removal, C) graph after stability-refinement.

where occasional sensor dropout can otherwise trigger global reconfiguration of the learned topology.

Across controlled synthetic experiments, GL-Ref yields more stable connectomes under node-removal tests, improves reconstruction accuracy, and preserves key topological statistics. We also observe low sensitivity of the regularization parameter within the tested grid, supporting practical deployment without extensive re-tuning across dropout rates.

On real PAM deployments, GL-Ref reduces the structural drift induced by node removal. Rather than re-learning a substantially different topology, it preserves the connectivity patterns implied by the chosen prior while keeping the smoothness fit to the observed signals. This consistency makes the resulting connectomes a more reliable representation of site-to-site acoustic relationships for downstream ecological analysis, highlighting (i) recurrent links among sites with similar land-cover and (ii) cross-land-cover links that may reflect shared biophonic activity or transitional soundscape dynamics.

Future work will focus on extending the framework to time-varying graph models to capture seasonal and intra-day dynamics in soundscapes.

References

[1] R. Gibb et al., "Emerging opportunities and challenges for passive acoustics in ecological assessment and monitoring," *Methods in Ecology and Evolution*, vol. 10, pp. 169–185, 2019.

[2] B. Pijanowski et al., "What is soundscape ecology? an introduction and overview of an emerging new science," *Landscape Ecol*, vol. 26, pp. 1213–1232, 2011.

[3] B. Pijanowski et al., "Soundscape Ecology: The Science of Sound in the Landscape," *BioScience*, vol. 61, pp. 203–216, 2011.

[4] D. Stowell and J. Sueur, "Ecoacoustics: acoustic sensing for biodiversity monitoring at scale," *Remote Sensing in Ecology and Conservation*, vol. 6, pp. 217–219, 2020.

[5] A. Castro-Ospina et al., "Graph-based audio classification using pre-trained models and graph neural networks," *Sensors*, vol. 24, pp. 2106, 2024.

[6] M. J. Guerrero et al., "Acoustic animal identification using unsupervised learning.," *Methods in Ecology and Evolution*, vol. 14, pp. 1500–1514, 2023.

[7] D. Bormpoudakis et al., "Spatial heterogeneity of ambient sound at the habitat type level: Ecological implications and applications," *Landscape Ecology*, vol. 28, pp. 495–506, 2013.

[8] M. J. Guerrero et al., "Graphical representation of landscape heterogeneity identification through unsupervised acoustic analysis," *Methods in Ecology and Evolution*, vol. 16, pp. 1255–1272, 2025.

[9] X. Dong et al., "Learning graphs from data: A signal representation perspective," *IEEE Signal Processing Magazine*, vol. 36, no. 3, pp. 44–63, 2019.

[10] G. Mateos et al., "Connecting the dots: Identifying network structure via graph signal processing," *IEEE Signal Processing Magazine*, vol. 36, no. 3, pp. 16–43, 2019.

[11] I. Brugere et al., "Network structure inference, a survey: Motivations, methods, and applications," *ACM Computing Surveys*, vol. 51, pp. 1–39, 2018.

[12] D.A. Nieto-Mora et al., "Systematic review of machine learning methods applied to ecoacoustics and soundscape monitoring," *Heliyon*, vol. 9, pp. e20275, 2023.

[13] N. Rendon et al., "Letting ecosystems speak for themselves: An unsupervised methodology for mapping landscape acoustic heterogeneity," *Environmental Modelling & Software*, vol. 187, pp. 106373, 2025.

[14] J. Friedman et al., "Sparse inverse covariance estimation with the graphical lasso," *Biostatistics*, vol. 9, pp. 432–41, 2008.

[15] V. Kalofolias, "How to learn a graph from smooth signals," in *International Conference on Artificial Intelligence and Statistics*, 2016, vol. 51, pp. 920–929.

[16] B. J. Williamson et al., "Virtual lesions in meg reveal increasing vulnerability of the language network from early childhood through adolescence," *Nature Communications*, vol. 14, pp. 1–10, 2023.

[17] C. Chen et al., "Inpainting-driven graph learning via explainable neural networks," in *AAAI Conference on Artificial Intelligence*, 2022.

[18] J. Shonfield and E. Bayne, "Autonomous recording units in avian ecological research: Current use and future applications," *Avian Conservation and Ecology*, vol. 12, pp. 14, 2017.

[19] J. W. Ribeiro et al., "Passive acoustic monitoring as a tool to investigate the spatial distribution of invasive alien species," *Remote Sensing*, vol. 14, no. 18, pp. 4565, 2022.

[20] X. Dong et al., "Learning laplacian matrix in smooth graph signal representations," *IEEE Transactions on Signal Processing*, vol. 64, pp. 6160–6173, 2016.

[21] B. O'Donoghue et al., "Conic optimization via operator splitting and homogeneous self-dual embedding," *Journal of Optimization Theory and Applications*, vol. 169, no. 3, pp. 1042–1068, 2016.

[22] A. Frieze et al., *Introduction to Random Graphs*, Cambridge University Press, 2016.

[23] M. Rubinov and O. Sporns, "Complex network measures of brain connectivity: Uses and interpretations," *NeuroImage*, vol. 52, pp. 1059–1069, 2010.

[24] S. Sosa-Orozco et al., "Network measures in animal social network analysis: Their strengths, limits, interpretations and uses," *Methods in Ecology and Evolution*, pp. 10–21, 2020.

[25] S. Diamond et al., "CVXPY: A Python-embedded modeling language for convex optimization," *Journal of Machine Learning Research*, vol. 17, no. 83, pp. 1–5, 2016.

[26] A. Agrawal et al., "A rewriting system for convex optimization problems," *Journal of Control and Decision*, vol. 5, no. 1, pp. 42–60, 2018.