ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XII-4/W1-2026
21st International 3D GeoInfo Conference 2026, 28–29 September 2026, Sofia, Bulgaria

Initial Bounding of Partial Optimal Transport for Point Cloud Co-Registration

Irvin Chadraa[1], Xinru Zhu[1], Esteban Tabak[1], Debra F. Laefer[2,3]

[1] Courant Institute of Mathematical Sciences, New York University (NYU) New York, NY
(ic2184@nyu.edu, xz4694@nyu.edu, tabak@cims.nyu.edu)
[2] Dept. of Civil and Urban Engineering, Tandon School of Engineering, NYU, Brooklyn, NY
[3] Center for Urban Science and Progress, Tandon School of Engineering, NYU, Brooklyn, NY
(debra.laefer@nyu.edu)

Keywords: partial optimal transport, data co-registration, subset selection, noise reduction

Abstract

Recently published work on the use of partial optimal transport (POT) has shown promise as a robust, fully-automated means to co-register point cloud data sets where either the spatial extent does not fully overlap or where one or more elements appear within only one of the scans due to temporal-based changes. However, understanding the ultimate robustness of such an approach requires further analysis of the relationship between its performance and the sensitivity of the user-defined hyperparameters that determine which elements in the partially-overlapped data set are included in the final correspondence map. Presented herein are a set of experiments formulated to investigate this relationship. The resulting outputs show a robustness to approximately 25% outliers, beyond which even carefully tuned hyperparameters cannot reliably achieve proper mass assignment. For practitioners, exponential regression models derived from controlled experiments provide reliable starting points for hyperparameter tuning in scenes with outlier proportions below 25%. When outlier content approaches or exceeds the 25% threshold, preprocessing through rough segmentation or filtering is recommended. Spatially-aware sampling improves robustness at low outlier percentages but degrades performance at high proportions. While hyperparameter value selection has negligible computational impact, runtime exhibits exponential growth with data size. POT's robust performance independent of point cloud sparsity suggests temporary downsampling strategies can effectively reduce computational burden without compromising final registration quality.

1. Introduction

Point cloud co-registration is a fundamental task in remote sensing, with applications ranging from urban modeling (Zováthi et al., 2022) to robotic perception (Cattaneo et al., 2022). Traditional registration methods such as Iterative Closest Point (ICP) (Besl and McKay, 1992) and its modern variants assume complete or near-complete overlap between scans and are sensitive to initialization and outliers. When temporal changes occur between acquisitions (e.g., the appearance or removal of vehicles, vegetation growth, or construction activity) or when spatial coverage differs due to occlusion or sensor limitations, these methods often fail or require extensive manual preprocessing.

Optimal transport (OT) provides a non-traditional basis by which to overcome these problems. OT is a geometrically principled framework for establishing a correspondence map that matches points between data sets by minimizing the total cost of mass redistribution. However, classical OT formulations assume balanced measures with complete mass transfer, making them unsuitable for scenarios with partial overlap or temporal differences. To overcome such a limitation, partial optimal transport (POT) extends the classical framework by allowing selective matching, thereby excluding points with prohibitively high transport costs from the final correspondence map.

While Chadraa et al. (2025) demonstrated the feasibility of POT for point cloud co-registration, a critical challenge in its application to real-world registration lies in balancing the trade-off between transport cost minimization and the degree of partial matching allowed by user-defined regularization parameters. Without a principled understanding of parameter selection, practitioners must resort to time-consuming, trial-and-error approaches or risk algorithmic failure in operational settings. Additionally, the sensitivity of the method and its relationship to data characteristics such as noise or outliers, point density, and sampling strategy have to date remained unexplored.

This work addresses these gaps through systematic testing of POT across controlled experimental scenarios. Specifically, the following are investigated: (1) if regularization parameters must scale as the proportion of outlier points increases, (2) if point quantity and density affect the robustness of correspondence mapping, (3) if patterns observed in one object class generalize to others, and (4) if different data sampling strategies influence parameter sensitivity. By testing the above, this work provides practitioners with guidance for parameter selection and identifies operational limits of POT-based registration.

2. Background

Optimal transport (Monge, 1781) provides a principled way to compare two probability distributions by computing how much "effort" is needed to transform one into the other. OT is often called the Earth Mover's Problem, as it was originally designed to minimize the effort of moving physical material, such as sand or soil. The problem assumes two distributions: a source μ on set X and a target ν on set Y. Classical OT searches for a minimum joint probability distribution $\pi \in X \times Y$ that couples the source and target marginal distributions, μ and ν, respectively (Kantorovich, 1942). In the continuous setting described above, OT is expressed in the form of Equation (1),

$$\min_{\pi} \left\{ \int_{X \times Y} c(x, y)\, \mathrm{d}\pi(x, y) \,\middle|\, \pi \in \Pi(\mu, \nu) \right\} \tag{1}$$

where $\Pi(\mu, \nu)$ is the set of all couplings, in which the marginal distributions are exactly μ and ν.

In a discrete setting, OT becomes a linear program. The optimal transport map is the real-valued matrix $\Pi = (\pi_{ij}) \in \mathbb{R}^{n\times m}$ that minimizes the total cost C, a sum of pairwise costs $(c_{ij}) \in \mathbb{R}^{n\times m}$ of transport between samples $\{x_i\}_{i=1}^n$ and $\{y_j\}_{j=1}^m$. The Euclidean distance is often used to measure transport cost. Regardless of the selected distance metric, each source point x_i has an associated mass $p_i = \mu(x_i)$, each target point y_j has mass $q_j = \nu(y_j)$, and masses in each marginal sum to one.

The linear program is written in Equation (2):

$$
\begin{aligned}
\min_{\Pi\geq 0, \Pi\in\mathbb{R}^{n\times m}} \quad & \sum\nolimits_{i,j} c_{ij}\pi_{ij} \\
\text{subject to} \quad & \sum\nolimits_{j} \pi_{ij} = p_i \quad \forall i \in \{1,\ldots,n\} \\
& \sum\nolimits_{i} \pi_{ij} = q_j \quad \forall j \in \{1,\ldots,m\}
\end{aligned}
\tag{2}
$$

This formulation represents the "full" optimal transport problem in a discrete setting. While it can be solved exactly, this comes at significant computational expense. Sinkhorn's algorithm (Cuturi, 2013) can be used to obtain an approximation with less computation.

An alternative computational strategy emerged from Essid et al. (2019), who reformulated optimal transport as a minimax optimization between a transport map $T = \nabla\phi$ and an adversarial function g. Their approach minimizes the Kullback-Leibler (KL) divergence (Kullback and Leibler, 1951) by composing local transport maps across intermediate distributions. That sample-based method avoids explicit density estimation and better adapts feature representations directly from data, which makes it suited to high-dimensional registration problems. In the context of image registration, Wang et al. (2025) extended the classical OT framework by incorporating cost-free transformations that account for natural variations in data acquisition such as rotations, translations, and perspective changes.

POT builds upon classical optimal transport by allowing selective matching between point sets rather than requiring complete mass transfer. Unlike full optimal transport, which forces every point to participate in the correspondence, POT selectively excludes points with prohibitively high transport costs. This flexibility makes POT particularly useful in registration tasks where data sets exhibit partial overlap, contain temporal differences, and/or include outlier structures. Theoretical foundations for partial transport were established by Caffarelli and McCann (2010) and Figalli (2010), who proved existence and uniqueness results for continuous POT maps.

Recent algorithmic developments have focused on discrete POT formulations suitable for practical applications. Chapel et al. (2020) embedded partial transport within the standard OT framework by introducing dummy points to absorb unmatched mass, enabling application of exact solvers. Qin et al. (2022), instead, relaxed classical OT marginal constraints into a range constraint that permits zero-mass assignment for individual points, while also capping total transported mass with a global parameter. Rather than tuning a hard mass cutoff, Shen et al. (2021) regularized both marginals with KL penalties and collapsed the resulting soft transport plan into a compact per-point displacement field, enabling GPU-scalable matching for point clouds with over 100k points. Riaz et al. (2023) proposed a semi-relaxed formulation incorporating KL regularization and proximal optimization for subset selection of the source data. Bai et al. (2023) decomposed high-dimensional partial transport into sequences of one-dimensional problems that can be solved efficiently, and later applied their method to non-rigid point-cloud registration by alternating between one-dimensional partial correspondence estimation and RBF/thin-plate-spline deformation updates (Bai et al., 2025).

Most recently, Chadraa et al. (2025) introduced a method of discrete POT that treats partially-overlapped source and target point clouds as collections of weighted samples and computes a transport plan that focuses on regions with consistent overlap. Points that are distant from any reasonable counterpart receive negligible transported mass and, thus, only minimally influence the final correspondence map.

3. Methodology

This section provides an overview of the partial optimal transport formulation, on which the basis of subsequent experiments is formed. For a detailed derivation, see (Chadraa et al., 2025).

3.1 Discrete Partial Optimal Transport

The discrete POT problem extends the linear program formulation by allowing the redistribution of mass to exclude high-cost point pairs. Given source points $\{x_i\}_{i=1}^n$ with masses p_i and target points $\{y_j\}_{j=1}^m$ with masses q_j, the method computes modified masses $\tilde{p}_i$ and $\tilde{q}_j$ that can either be zero (excluded) or scaled versions of the original masses. The transport plan $\Pi = (\pi_{ij})$ then couples only the selected points.

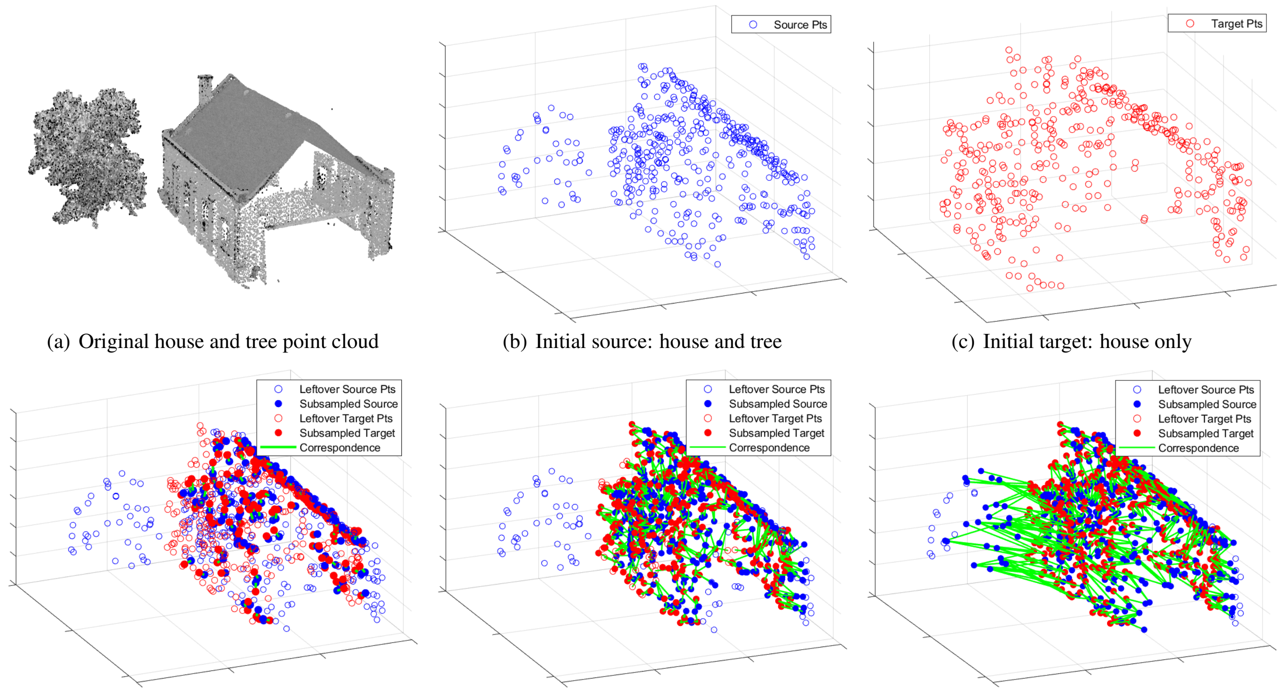
The optimization problem takes the following form.

$$
\begin{aligned}
\min_{\Pi\geq 0, \Pi\in\mathbb{R}^{n\times m}} \quad & \sum\nolimits_{i,j} c_{ij}\pi_{ij} + \lambda(\alpha+\beta) \\
\text{subject to} \quad & \alpha \geq 1, \quad \beta \geq 1, \quad \tilde{p}_i \leq \alpha p_i, \quad \tilde{q}_j \leq \beta q_j \\
& \sum\nolimits_{j} \pi_{ij} = \tilde{p}_i \quad \forall i \in \{1,\ldots,n\} \\
& \sum\nolimits_{i} \pi_{ij} = \tilde{q}_j \quad \forall j \in \{1,\ldots,m\} \\
& \sum\nolimits_{i} \tilde{p}_i = 1, \quad \sum\nolimits_{j} \tilde{q}_j = 1
\end{aligned}
\tag{3}
$$

where c_{ij} represents the pairwise Euclidean transport cost, and the scaling parameters α and β allow for redistribution of the mass of the points, while maintaining the total unit mass.

When α and β equal one, the problem reverts to the original full transport formulation. As α and β grow, the formulation allows the remaining masses at more significant points to grow, while the masses of points with the highest transport costs (i.e., the furthest points) are reduced to a mass of zero and, thus, excluded from the transport plan. Very large values of α and β lead to an "extreme" partial transport, where only the closest points between the source and target sets are matched, effectively converting this formulation into a search for the minimizer of a nearest-neighbor problem.

The challenge in applying POT to real-world scenarios lies in the sensitivity of the method to hyperparameter selection. To produce an optimal co-registration between correct elements within the data set with POT, a user-defined regularization parameter λ is introduced into the objective function. The selection of λ is critical, as it has a direct relationship with the extent of



(a) Original house and tree point cloud
(b) Initial source: house and tree
(c) Initial target: house only
(d) Partial Optimal Transport ($\lambda = 0.25$)
(e) Partial Optimal Transport ($\lambda = 5$)
(f) Partial Optimal Transport ($\lambda = 25$)

Figure 1. Qualitative results of partial optimal transport on a data set with 400 randomly sampled house points and 40 randomly sampled tree points. Effective co-registration between the house structures depends on the λ regularization parameter, with lower λ resulting in less points included in the output and higher λ including more points, but potentially more outliers as well.

subset selection. Figure 1 provides a visualization of the point cloud data used in Chadraa et al. (2025), as well as the impact of arbitrarily selected values of λ on POT. Note that smaller values of λ allow the POT algorithm to include higher α and β terms in the optimization landscape, thereby restricting the number of points in the final correspondence solution. Alternatively, larger values revert the mapping to the full transport case, where complete matching yields soft correspondences that become difficult to interpret.

3.2 Solution Procedure

The discrete POT linear program is solved using an implicit gradient descent method by converting the problem into a primal-dual saddle-point framework introduced in Essid et al. (2023). This approach reformulates the linear program as a Lagrangian with dual variables enforcing the constraints, then applies curvature-aware updates that anticipate the opponent's moves to minimize the transport matrix Π and maximize constraint multipliers. The method scales for large point clouds and avoids the diffuse matchings produced by entropy-regularized solvers like Sinkhorn, which depend upon the λ parameter.

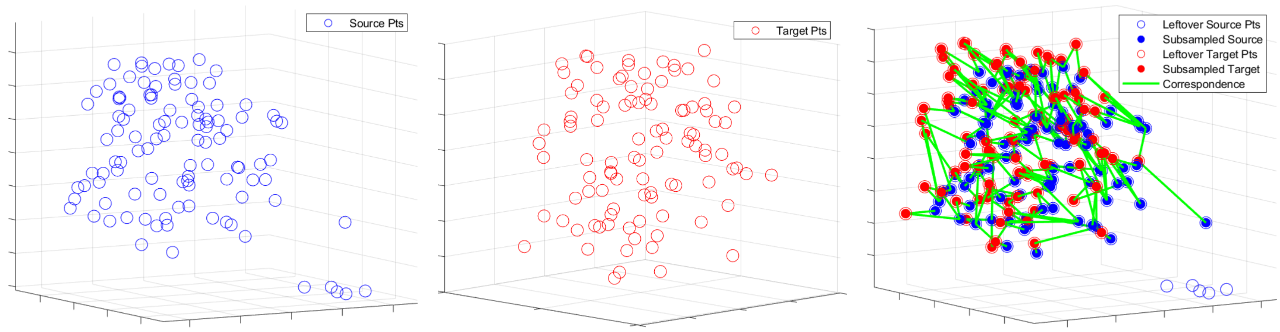
To ensure interpretable binary correspondences, two refinement steps are applied. First, a mass bounding procedure governed by tolerance ϵ snaps masses close to zero or to their maximum uniform weight, which immediately classifies most points as active or inactive. Second, a branch-and-cut algorithm resolves any remaining intermediate mass values by systematically adding constraints to the linear program. The constraints fix point masses to the maximum allowable value or to zero (Mitchell, 2002). Together, these steps produce hard, one-to-one or one-to-none correspondences suitable for point cloud registration. Additional algorithmic enhancements include robust preconditioning and adaptive learning rate selection via backtracking line search with directional conditions to ensure the saddle-point problem converges (Chadraa et al., 2025).

3.3 Data Processing

The set of experiments in Chadraa et al. (2025) used a subset of a high-density, aerial light detection and ranging (LiDAR) point cloud collected by Laefer and Vo (2019), which includes scans of distinct urban objects including houses, trees, and light poles that were manually extracted to test the proposed approach. Each object's original scan consists of over 250,000 points, which are then subsampled in CloudCompare to be easily managed with local compute power. To ensure spatial coherence across different experimental setups, additional preprocessing was conducted so that the subsampled objects were positioned appropriately near each another, and their coordinate systems were translated so that all structures would lie on the same vertical plane. This setup was created to emulate real-world, multi-object scans where irrelevant structures may be present near larger objects of interest.

4. Experiments

To provide a more principled analysis on the performance of partial optimal transport and the influence of the noise penalty parameter λ, five sets of experiments are undertaken herein. Across all experiments, a binary search over λ is performed to identify the minimum value that yields the first failure case, defined as the algorithm assigning non-zero mass to at least one outlier point. These failure λ values are then plotted against the percentage of outlier points and fit as an exponential regression curve to provide practical guidance for future applications. Figure 2 illustrates a failure case in the correspondence map where a single false correspondence is made to an outlier point.



(a) Initial source: tree and light pole

(b) Initial target: tree only

(c) Partial optimal transport ($\lambda = 20$)

Figure 2. POT applied on a sparse point cloud created from Laefer and Vo (2019) with 100 randomly sampled tree points and 6 randomly sampled lightpole points. A failure case is shown, with a single point from the light pole is included in the final correspondence map.

4.1 Experiment 1: Upper Bound with Fixed Point Density

This experiment establishes an empirically estimated upper bound of the λ parameter based on the proportion of outlier points such that all user-defined λ values below this bound yield valid correspondence maps that only match points between the true structures.

A new composite source point cloud and a target point cloud are constructed from components belonging to different data sets:

$$\text{source} = \text{true} + \text{outlier}, \qquad \text{target} = \text{true}.$$

where the true structure (a house) remains consistent across both source and target point clouds, and the smaller outlier structure (a tree) is included only within the source point cloud. This creates a controlled partial-overlap scenario where the tree represents an additional structure that should ideally be down-weighted by POT. While the house structure contains 100 randomly sampled points from one dataset, the tree structure varies from 1 to 50 points, incrementing by one to create a variety of different outlier configurations.

4.2 Experiment 2: Effect of Point Density Variation

This experiment extends the first experiment by evaluating whether point density affects the relationship between λ and the proportion of outliers. The procedure of Experiment 1 is repeated using higher-density configurations of the original data-set with sampling 200, 300, and 400 points for the house structure. The tree point count is varied proportionally to maintain comparable outlier percentages across different densities, where a range of 2 to 100 points (incrementing by 2) is sampled for the configuration with 200 points in the house structure, 3 to 150 points (incrementing by 3) are sampled for the 300-point house configuration, and 4 to 200 points (incrementing by 4) are sampled for the 400-point house configuration. By comparing the failure λ curves at these three additional density levels, this experiment assesses whether denser point clouds require different λ values to achieve robust correspondence mapping and if the decay pattern persists across sampling resolutions.

4.3 Experiment 3: Data Sampling Strategy Comparison

This experiment evaluates whether the spatial distribution of sampled points influences parameter sensitivity. Instead of random decimation, an alternative data set is created using a spatial-aware sampling method in CloudCompare. In this case, the user first defines a minimum accepted distance between two points. CloudCompare automatically samples points from the original cloud by parsing neighboring points within the input radius. This sampling method ensures more uniform coverage, avoiding clusters and holes that could potentially affect the cost matrix and the transport solution. For each data set, minimum distances are defined such that the same point counts are consistent between experiments 1, 2, and 3. The failure λ search procedure is repeated, and the resulting exponential decay curves are compared against experiments 1 and 2 to determine whether spatially-aware sampling provides greater robustness to POT. Table 1 compares point densities of the data produced using random and spatial sampling strategies.

Points	House-Tree			
	Random		Spatial	
	Source	Target	Source	Target
100	0.221 ± 0.011	0.208	0.170 ± 0.017	0.165
200	0.408 ± 0.032	0.415	0.327 ± 0.034	0.321
300	0.623 ± 0.050	0.617	0.493 ± 0.051	0.488
400	0.826 ± 0.068	0.866	0.648 ± 0.067	0.645

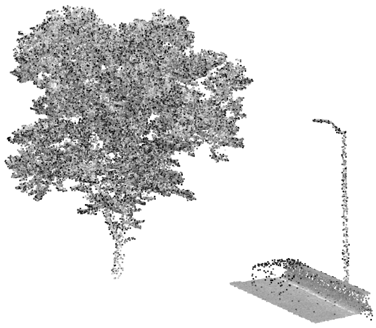
Table 1. Point cloud density comparison (points per m^2).

4.4 Experiment 4: Cross-Object Generalization

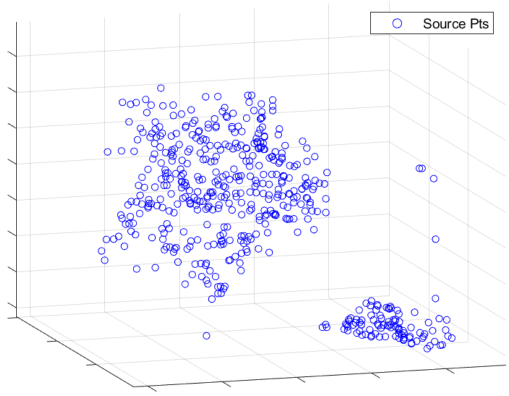
To assess whether the patterns observed in the house-tree scenario generalize to other object classes, the previous experiments are repeated using a tree as the primary structure and a light pole as the outlier. The tree structure is sampled to 100, 200, 300, and 400 points; matching the number of house points in the previous experiments. The number of points in the light pole matches the previous setup as well, by ranging from 1 to 50 for the 100 tree point setting, from 2 to 100 (incrementing by 2) for the 200 tree point setting, from 3 to 150 (incrementing by 3) for the 300 tree point setting, and from 4 to 200 (incrementing by 4) for the 400 tree point setting. This setting tests whether failure λ curves exhibit similar exponential decay patterns across structurally different objects, and if any performance thresholds observed in prior experiments represent algorithmic limits or are specific to the house-tree data geometry. Figures 2 and 3 illustrate sparse and dense versions of this alternative tree-lightpole data set, respectively. The average source point density ranges from 0.767 ± 0.067 pts/m^2 to 2.541 ± 0.268 pts/m^2.

4.5 Experiment 5: Evaluation on Alternate Urban Dataset

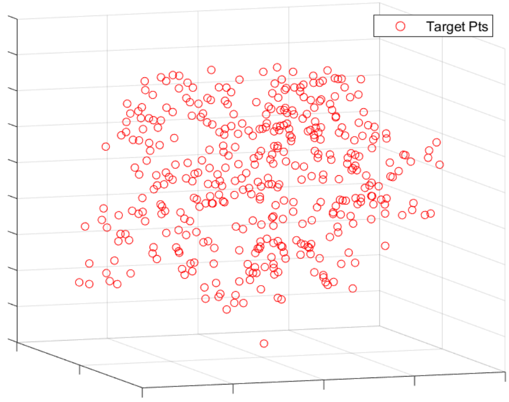
To test the real-world validity of the emulated multi-object configuration used in previous experiments, this experiment eval-



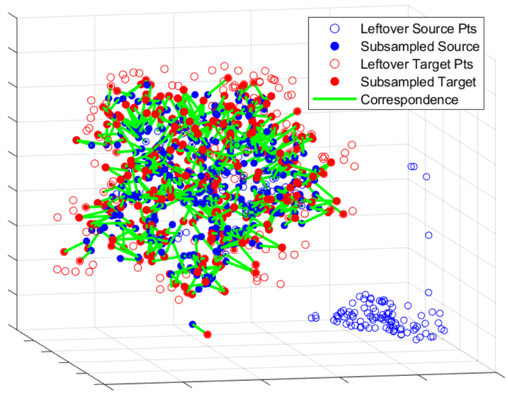
(a) Original tree and light pole point cloud



(b) Initial source: tree and light pole

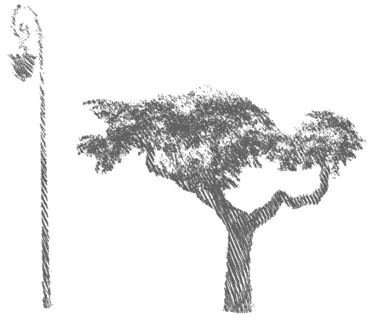


(c) Initial target: tree only

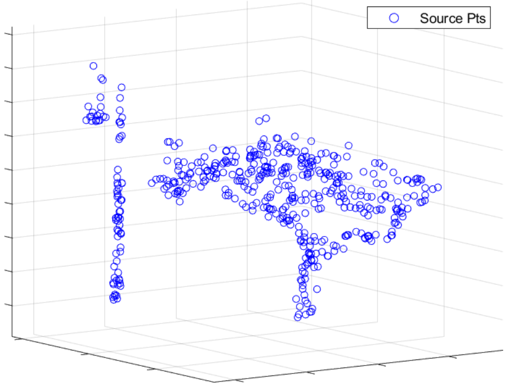


(d) Partial optimal transport ($\lambda = 3$)

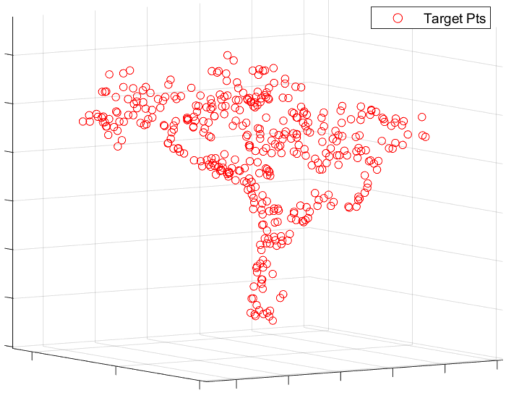
Figure 3. Results of POT on a dense point cloud with 400 randomly sampled tree points and 100 randomly sampled light pole points. The value of λ is smaller compared to the sparse data set, since the set of feasible λ shrinks, as the number of points increases.



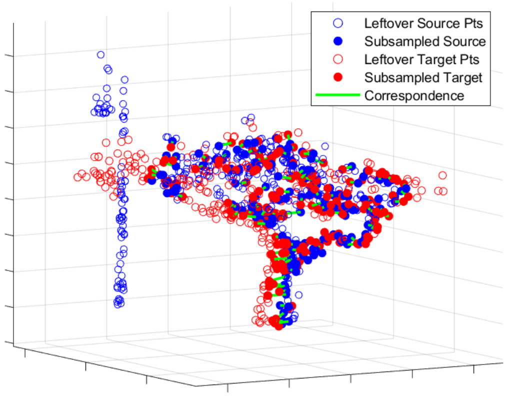
(a) Original tree and light pole point cloud



(b) Initial source: tree and light pole



(c) Initial target: tree only



(d) Partial optimal transport ($\lambda = 20$)

Figure 4. Results of POT on a real-life tree-lightpole point cloud taken from the Paris-Lille-3D dataset with 400 randomly sampled tree points and 100 randomly sampled light pole points.

uates the effects of POT on a segmented portion of Paris-Lille-3D, an open-source urban point cloud dataset collected by Roynard et al. (2018). For the source point cloud, a tree and a lightpole are captured next to each other. For the target point cloud, the light pole is segmented out, leaving only points related to the tree. The number of points in the tree and light pole is sampled exactly as in Experiment 4. Figure 4 presents this real-world data set, which is noticeably more dense. The average source point density ranges from 54.2508 ± 6.0216 pts/m^2 to 83.120 ± 23.724 pts/m^2.

5. Results

The experiments reveal consistent patterns in the relationship between λ, outlier proportion, and algorithm performance. Across all experiments, there is a recurring performance breakdown threshold, which appears to reflect an inherent limitation of POT. In cases where outlier points are numerous or when they are near the true structure, the cost of matching outliers to distant but feasible points on the true structure becomes competitive with the penalty for excluding them. This tradeoff erodes POT's ability to reliably distinguish between valid and spurious correspondences.

5.1 Runtime analysis

For all experiments, runtimes were recorded to empirically demonstrate the relation with the number of points in the data. Figure 5 displays exponential relationships between the size of the dataset and the algorithm runtime for both randomly and spatially sampled datasets. Similarly, the standard deviation increases as the dataset size increases. Further investigation of the impact of the parameter λ on the POT algorithm runtime shows no discernible influence. Figure 6 shows the runtimes on randomly sampled house-tree data set with 400 true points and 40 outliers, after sweeping a range of valid λ values. All experiments were conducted using MATLAB R2023b on a system with 13th Gen Intel Core i7-13700H CPU and 16GB RAM.

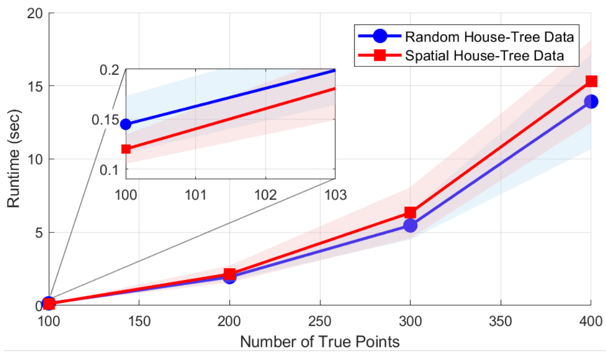


Figure 5. POT runtime on randomly and spatially-sampled house-tree data at increasing data sizes. Runtime exhibits exponential growth with variability increasing with data size.

5.2 Experiment 1

In Experiment 1, with 100 points in the house structure, λ decreases exponentially to maintain correct correspondence mapping, which is reasonable as having zero outliers would mean that any value of λ yields a valid output. The fitted exponential model takes the form $\lambda \approx Ae^{-B(x-C)}$, where x represents the outlier percentage, and constants A, B, and C are determined by least-squares regression.

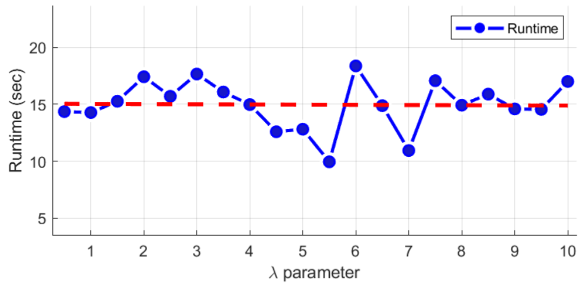


Figure 6. Randomly sampled house-tree data: runtime versus increasing λ values. No visible influence on runtime appears.

5.3 Experiment 2

Results of Experiment 2 show that low-density data configurations exhibit greater stability, tolerating slightly higher λ values before failure occurs. Conversely, the high-density configuration (400 points) shows increased volatility, with failure λ values varying more erratically, particularly at higher outlier percentages. This suggests that the randomly sampled data sets with high point counts included a number of extraneous points that made it difficult for the algorithm to distinguish between inliers and outliers. Figure 7 presents the primary findings from Experiments 1 and 2, clearly showing exponential decay patterns as the proportion of outlier points increases.

5.4 Experiment 3

Figure 8 shows that spatial sampling of the data sets yields similar exponential decay patterns compared to random sampling, but with a much wider range of λ values. For data configurations with a lower outlier percentage, a higher possible maximum λ is observed, irrespective of the total points in the data set. This implies that the spatial sampling method in these cases produces data that are more robust to partial optimal transport and that outliers are more easily identifiable by the algorithm. However, as with the random sampling case, a higher outlier proportion yields a smaller range of feasible λ values, meaning it is more difficult for the algorithm to perform well with the additional outliers, regardless of the dataset's spatial structure.

5.5 Experiment 4

In Experiment 4, similar exponential decay patterns of λ are observed when using an alternate tree and light pole dataset. Figure 9 shows that the behavior is more consistent across different density configurations, likely a result of the tree's inherent geometric structure resembling a three-dimensional Gaussian distribution. This suggests a less pressing need for large data sets when there is no preferred orientation in the data representation, as smaller and sparser data sets still capture the same POT behavior. In the case of the house, the roof is more heavily captured in the aerial LiDAR scan compared to the vertical walls; see Hinks et al. (2009) regarding native data distributions.

5.6 Experiment 5

Figure 10 shows the behavior of the failure λ values with this alternative real-world data set. Because the light pole is positioned much closer to the tree, the values in the λ plot are much lower and more concentrated together compared to the similar setup used in experiment 4, ranging from 0-1 as opposed to the more varied results in prior experiments. This is likely due to the point cloud density being more concentrated in comparison, which yields a far narrower interval of successful registration as outlier points are even closer to the primary structure.

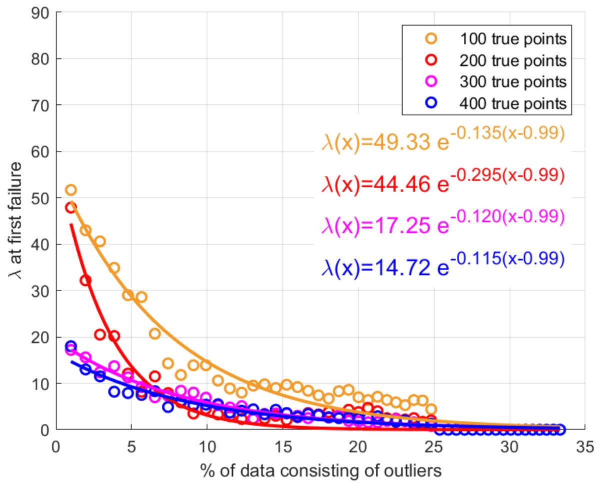


Figure 7. Experiments 1 and 2: Random sampling of house-tree data yields exponential decay in λ across point counts.

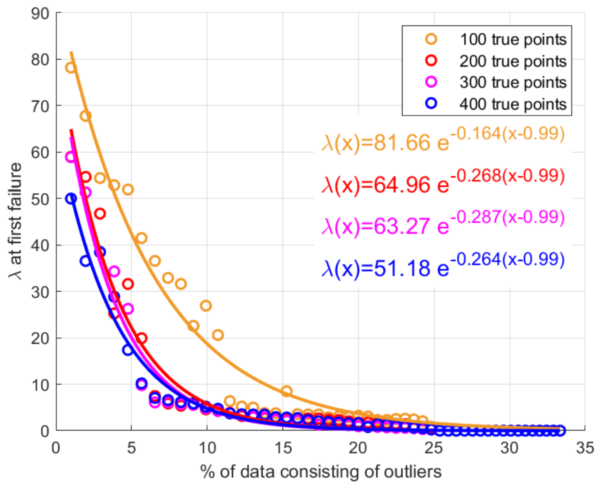


Figure 8. Experiment 3: Spatial sampling of house-tree data yields a larger range of possible λ values than those generated on randomly sampled datasets.

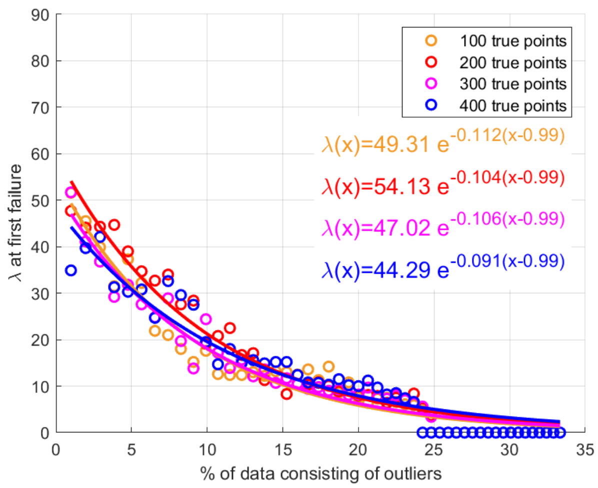


Figure 9. Experiment 4: Random sampling of the tree-light pole data results in similar behavior between different point counts.

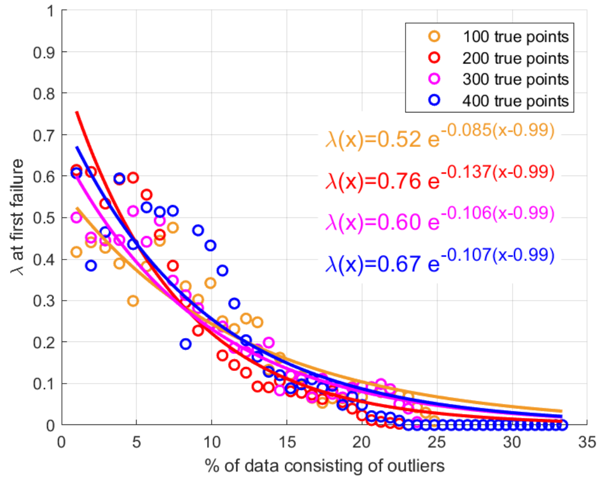


Figure 10. Experiment 5: Using an alternate dataset yields a much smaller range of λ values less than 1, as well as similar exponential behavior across different point counts.

6. Discussion

Systematic investigation of POT sensitivity reveals several important insights for practical application of the method to point cloud co-registration tasks. Through controlled experiments varying outlier proportion, point density, and sampling strategy across multiple object classes, key findings are established:

1. The relationship between minimum failure λ and outlier percentage follows a consistent exponential decay pattern, which can be characterized by simple regression models to guide parameter selection in new scenarios.
2. A fundamental breakdown threshold exists across multiple experimental and density configurations at approximately 25% outliers, beyond which reliable correspondence mapping becomes infeasible regardless of λ tuning.
3. However, both subsampling strategies (random vs spatially subsampled) and total point count impact the maximum viable λ, especially at lower percentages of outliers
4. An exponential growth in runtime is observed as the sizes of the data sets increase. However, the impact of λ on runtime is negligible.

These findings provide practitioners with practical guidance for applying POT to real-world registration tasks. For scenes with known or estimated outlier proportions below 20-25%, exponential regression models offer reasonable starting points for λ selection, which can then be refined through limited local search. For scenes approaching or exceeding the 25% threshold, preprocessing such as rough segmentation or filtering to reduce outlier content should be incorporated. Additionally, considering that POT performs well irrespective of point cloud sparsity, long runtimes can be countered by using smaller datasets without significant loss of information.

Specifically, larger data sets can be temporarily downsampled for the purpose of co-registration. Voxel-based downsampling is one method that can preserve a strong representation of the original data, while reducing compute and memory costs, although registration quality can still be compromised if too many data points are removed. Such abilities also lend themselves to matching data sets with different resolution and those from times periods. Further exploration of these topics will be the focus of future experimentation by the authors.

7. Conclusions

This work provides the first systematic characterization of hyperparameter sensitivity in partial optimal transport for point cloud co-registration, with specific focus on the critical regularization parameter λ. This principled study further advances the use of partial optimal transport to a deployable, scalable tool for real-life co-registration tasks.

To further evaluate robustness, future considerations should extend this analysis to a broader range of datasets capturing different scenarios, such as those including other objects or multiple instantiations of the same object within the same point cloud. Additional possible directions include investigating the interaction between cost function design and parameter sensitivity, or exploring adaptive λ selection strategies that incorporate geometric descriptors. Ultimately, establishing theoretical bounds that connect the empirically observed 25% outlier threshold to fundamental properties of POT is needed, as this would provide valuable insights into the ultimate limits of partial optimal transport for registration applications.

References

Bai, Y., Schmitzer, B., Thorpe, M., Kolouri, S., 2023. Sliced optimal partial transport. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 13681–13690.

Bai, Y., Tran, H., Damelin, S. B., Kolouri, S., 2025. Partial transport for point-cloud registration. *Sampling Theory, Signal Processing, and Data Analysis*, 23(1), 4.

Besl, P. J., McKay, N. D., 1992. Method for registration of 3-d shapes. *Sensor fusion IV: control paradigms and data structures*, 1611, Spie, 586–606.

Caffarelli, L. A., McCann, R. J., 2010. Free boundaries in optimal transport and Monge-Ampere obstacle problems. *Annals of Mathematics*, 673–730. doi.org/10.4007/annals.2010.171.673.

Cattaneo, D., Vaghi, M., Valada, A., 2022. LCDNet: Deep loop closure detection and point cloud registration for LiDAR SLAM. *IEEE Transactions on Robotics*, 38(4), 2074–2093. doi.org/10.1109/TRO.2022.3150683.

Chadraa, I., Tabak, E., Laefer, D. F., 2025. Partial Optimal Transport for Co-Registration of Partially-Overlapped Point Clouds. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, X-4/W6-2025, 41–48. doi.org/10.5194/isprs-annals-X-4-W6-2025-41-2025.

Chapel, L., Alaya, M. Z., Gasso, G., 2020. Partial optimal tranport with applications on positive-unlabeled learning. *Advances in Neural Information Processing Systems*, 33, 2903–2913. doi.org/10.48550/arXiv.2002.08276.

Cuturi, M., 2013. Sinkhorn distances: Lightspeed computation of optimal transport. *Advances in Neural Information Processing Systems*, 26. doi.org/10.48550/arXiv.1306.0895.

Essid, M., Laefer, D. F., Tabak, E. G., 2019. Adaptive optimal transport. *Information and Inference: A Journal of the IMA*, 8(4), 789–816. doi.org/10.1093/imaiai/iaz008.

Essid, M., Tabak, E. G., Trigila, G., 2023. An implicit gradient-descent procedure for minimax problems. *Mathematical Methods of Operations Research*, 97(1), 57–89. doi.org/10.48550/arXiv.1906.00233.

Figalli, A., 2010. The optimal partial transport problem. *Archive for Rational Mechanics and Analysis*, 195(2), 533–560. doi.org/10.1007/s00205-008-0212-7.

Hinks, T., Carr, H., Laefer, D. F., 2009. Flight optimization algorithms for aerial LiDAR capture for urban infrastructure model generation. *Journal of Computing in Civil Engineering*, 23(6), 330–339. doi.org/10.1061/(ASCE)0887-3801(2009)23:6(330).

Kantorovich, L. V., 1942. On the translocation of masses. *Dokl. Akad. Nauk. USSR (NS)*, 37, 199–201.

Kullback, S., Leibler, R. A., 1951. On information and sufficiency. *The annals of mathematical statistics*, 22(1), 79–86.

Laefer, D. F., Vo, A.-V., 2019. 2019 LiDAR Data Collection for Sunset Park, Brooklyn, NY. doi.org/10.17609/4cpx-2h33.

Mitchell, J. E., 2002. Branch-and-cut algorithms for combinatorial optimization problems. *Handbook of applied optimization*, 1(1), 65–77.

Monge, G., 1781. Mémoire sur la théorie des déblais et des remblais. *Mem. Math. Phys. Acad. Royale Sci.*, 666–704. doi.org/10.1112/plms/s1-14.1.139.

Qin, H., Zhang, Y., Liu, Z., Chen, B., 2022. Rigid registration of point clouds based on partial optimal transport. *Computer Graphics Forum*, 41(6), 365–378.

Riaz, B., Karahan, Y., Brockmeier, A. J., 2023. Partial optimal transport for support subset selection. *Transactions on Machine Learning Research*. openreview.net/forum?id=75CcopPxIr.

Roynard, X., Deschaud, J.-E., Goulette, F., 2018. Paris-Lille-3D: A large and high-quality ground-truth urban point cloud dataset for automatic segmentation and classification. *The International Journal of Robotics Research*, 37(6), 545–557.

Shen, Z., Feydy, J., Liu, P., Curiale, A. H., San Jose Estepar, R., San Jose Estepar, R., Niethammer, M., 2021. Accurate point cloud registration with robust optimal transport. *Advances in Neural Information Processing Systems*, 34, 5373–5389.

Wang, Y., Tabak, E., Laefer, D. F., 2025. Optimal transport with cost-free transformations for image co-registration. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLVIII-4/W15-2025, 179–186. doi.org/10.5194/isprs-archives-XLVIII-4-W15-2025-179-2025.

Zováthi, Ö., Nagy, B., Benedek, C., 2022. Point cloud registration and change detection in urban environment using an onboard Lidar sensor and MLS reference data. *International Journal of Applied Earth Observation and Geoinformation*, 110, 102767.