NoPain: No-box Point Cloud Attack via Optimal Transport Singular Boundary

Zezeng Li[1,2], Xiaoyu Du[1], Na Lei[1*], Liming Chen[2], Weimin Wang[1*]
[1]School of Software, Dalian University of Technology, China [2]École Centrale de Lyon, France

This document provides additional experimental results on the impact of the parameter λ and includes further quantitative and qualitative comparisons. We also discuss the limitations of our approach and outline directions for future work.

1. Additional Results

In this section, we designed experiments to evaluate the impact of the parameter λ in Eq. (8) of the manuscript on the effectiveness of the attack. Additionally, we present further quantitative comparisons on the ScanObjectNN [8] dataset, along with qualitative comparison results on ModelNet40 [9] and ShapeNetPart [1].

As stated in our manuscript, the adversarial feature $\widehat{\boldsymbol{y}}$ can be obtained through the following equation (Eq. (8) of the manuscript):

$$\widehat{\boldsymbol{y}} = \tilde{T}(\boldsymbol{x}) = \lambda_i T(\boldsymbol{c}_i) + \lambda_{i_k} T(\boldsymbol{c}_{i_k}) = \lambda_i \boldsymbol{y}_i + \lambda_{i_k} \boldsymbol{y}_{i_k}. \quad (1)$$

Where the μ-mass center $\boldsymbol{c}_i$ is approximated by the mean value of all the Monte-Carlo samples inside W_i, $\boldsymbol{x}$ is a random $\boldsymbol{x}$ from W_i. The coefficient λ_i and λ_{i_k} satisfies $\lambda_i + \lambda_{i_k} = 1$. To simplify notation, we denote λ_{i_k} as λ. Consequently, $\lambda_i = 1 - \lambda$.

To evaluate the impact of λ on the attack, we conducted experiments on the ModelNet40 and ShapeNetPart using different λ values. We recorded the attack success rate (ASR) on PointNet [7] along with the corresponding Chamfer Distance (CD). The results are presented in Fig. 1, where the left graph shows the results on the ModelNet40, and the right graph shows the results on ShapeNetPart.

As shown in Fig. 1, as λ increases, the attack success rate improves, while the Chamfer distance between the adversarial and original point clouds also increases. This result is expected since higher values of λ lead to greater divergence of the generated adversarial features from the original features, resulting in decoded point clouds that progressively differ from the originals. The objective of our proposed algorithm is to identify a critical boundary where the generated adversarial samples can successfully execute attacks while ensuring the perturbations remain sufficiently small.

*Corresponding authors.

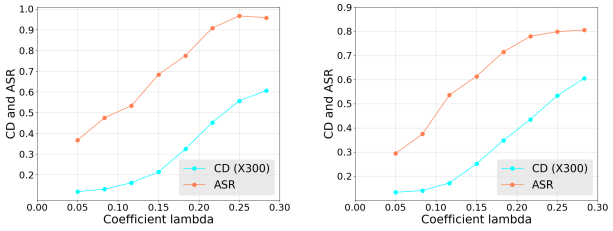

Figure 1. Effects of λ to ASR and CD on ModelNet40 (left) and ShapeNetPart (right). To present these two metrics in a single graph, we scaled the CD values by a factor of 300.

We perform additional quantitative comparison experiments on the ScanObjectNN dataset. The Attack Success Rate (ASR) and Chamfer Distance (CD) are presented in **??**, where all baselines are optimized using PointNet++ and then used to attack PointNet, DGCNN, and PointMamba [4]. The results demonstrate that our method (NoPain-PD) achieves the highest ASR for transferable attacks while maintaining a comparable perturbation intensity.

The results in Fig. 2 demonstrate that our method successfully attacks five classification models simultaneously, causing them to misclassify. In contrast, baseline methods, while effective on the surrogate model PointNet++, often fail when targeting other unknown classification models, highlighting their limited transferability. This issue is especially pronounced on the ShapeNetPart dataset. Furthermore, adversarial point clouds generated by baseline methods such as AdvPC [2], AOF [5], SI-Adv [3], and SS-attack [10] exhibit significant, perceptible noise. Although HiT-ADV [6] produces fewer outliers, it still introduces distortions, as seen in the keyboard example in the first column. By contrast, our method not only achieves transferable attacks but also maintains smaller, imperceptible perturbations.

2. Limitation and Discussion

Due to the absence of reconstruction error constraints, our method cannot guarantee imperceptible perturbations in the generated adversarial samples. Although the overall re-

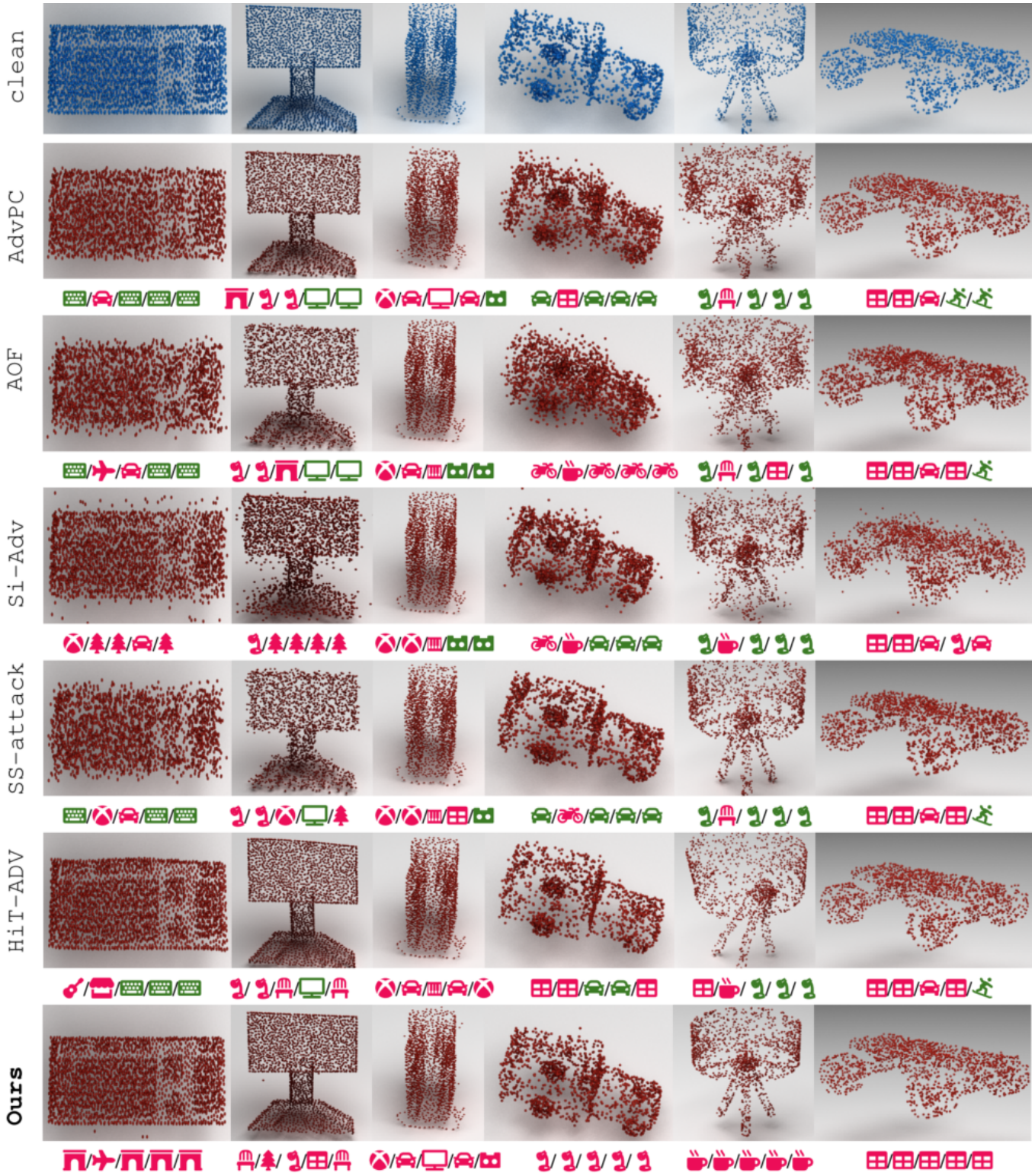

Figure 2. Visualizations of adversarial samples on data from ModelNet40 (left three columns) and ShapeNetPart (right three columns). The icons below point clouds indicate their category prediction by PointNet, PointNet++, PointConv, DGCNN and PCT, where red and green indicate successful and failed attacks.

sults produced by our approach show low reconstruction errors, some outputs may still include noticeable outliers. In future work, we will explore optimal transport mapping and point cloud attacks that incorporate reconstruction error constraints. Additionally, our observations reveal that the hyperparameter threshold τ in Algorithm 2 of our manuscript varies significantly across different datasets. To address this, we intend to develop an adaptive algorithm for automatically adjusting the threshold τ.

References

[1] Angel X Chang, Thomas Funkhouser, Leonidas Guibas, Pat Hanrahan, Qixing Huang, Zimo Li, Silvio Savarese, Manolis Savva, Shuran Song, Hao Su, et al. Shapenet: An information-rich 3d model repository. *arXiv preprint arXiv:1512.03012*, 2015. 1

[2] Abdullah Hamdi, Sara Rojas, Ali Thabet, and Bernard Ghanem. Advpc: Transferable adversarial perturbations on 3d point clouds. In *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XII 16*, pages 241–257. Springer, 2020. 1

[3] Qidong Huang, Xiaoyi Dong, Dongdong Chen, Hang Zhou, Weiming Zhang, and Nenghai Yu. Shape-invariant 3d adversarial point clouds. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 15335–15344, 2022. 1

[4] Dingkang Liang, Xin Zhou, Wei Xu, Xingkui Zhu, Zhikang Zou, Xiaoqing Ye, Xiao Tan, and Xiang Bai. Pointmamba: A simple state space model for point cloud analysis. In *Advances in Neural Information Processing Systems*, 2024. 1

[5] Binbin Liu, Jinlai Zhang, and Jihong Zhu. Boosting 3d adversarial attacks with attacking on frequency. *IEEE Access*, 10:50974–50984, 2022. 1

[6] Tianrui Lou, Xiaojun Jia, Jindong Gu, Li Liu, Siyuan Liang, Bangyan He, and Xiaochun Cao. Hide in thicket: Generating imperceptible and rational adversarial perturbations on 3d point clouds. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 24326–24335, 2024. 1

[7] Charles R Qi, Hao Su, Kaichun Mo, and Leonidas J Guibas. Pointnet: Deep learning on point sets for 3d classification and segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 652–660, 2017. 1

[8] Mikaela Angelina Uy, Quang-Hieu Pham, Binh-Son Hua, Duc Thanh Nguyen, and Sai-Kit Yeung. Revisiting point cloud classification: A new benchmark dataset and classification model on real-world data. In *International Conference on Computer Vision (ICCV)*, 2019. 1

[9] Zhirong Wu, Shuran Song, Aditya Khosla, Fisher Yu, Linguang Zhang, Xiaoou Tang, and Jianxiong Xiao. 3d shapenets: A deep representation for volumetric shapes. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1912–1920, 2015. 1

[10] Jinlai Zhang, Yinpeng Dong, Jun Zhu, Jihong Zhu, Minchi Kuang, and Xiaming Yuan. Improving transferability of 3d adversarial attacks with scale and shear transformations. *Information Sciences*, 662:120245, 2024. 1