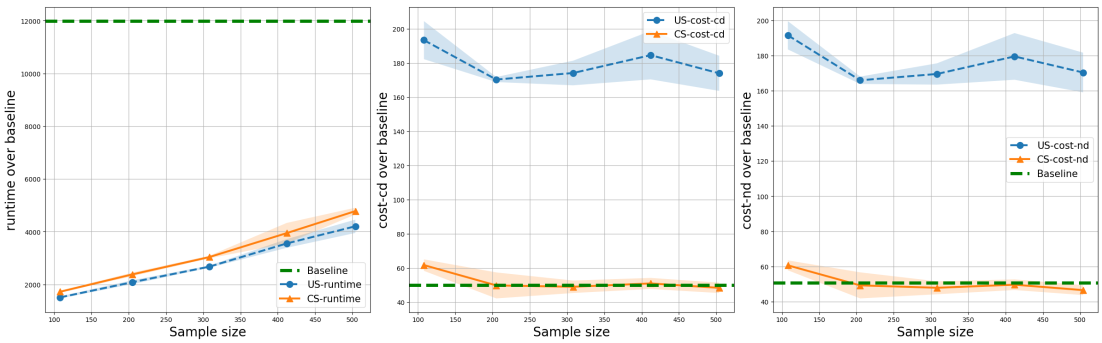
C.1. Experiments on S3DIS for Review

The *S3DIS* (Armeni et al., 2016) is a large-scale indoor 3D point cloud dataset. It consists of 13 semantic categories, including ceiling, floor, wall, beam, column, window, door, table, chair, sofa, bookcase, board, and clutter. We first extract individual point cloud shapes from the dataset and then perform sampling on each shape, representing it with 300 points to form our point cloud dataset.

Efficiency of our coreset method: The green dashed line represents the results on the full dataset. From Figure 10, it can be seen that our CS method significantly outperforms the US method.



(a) Comparison of runtime (b) Comparison of cost-cd (c) Comparison of cost-nd

Figure 10: Comparison of the US method and our CS method across varying sample sizes on S3DIS dataset.

Ablation experiments on k: From Table 7, our CS method consistently has an advantage over the US method for different k.

Table 7: Comparison of the US method and our CS method with varying values of k on S3DIS dataset. We fix the sample size as 504, $\sigma = 1$, and $\zeta = 0.1$.

k	SM	cost-nd(↓)	cost-cd(↓)	Runtime(↓)
10	US	$170.37_{\pm 11.29}$	$173.89_{\pm 10.28}$	$4205.60_{\pm 252.75}$
	CS	$\mathbf{46.53}_{\pm 2.83}$	$\mathbf{48.33}_{\pm 3.04}$	$4225.40_{\pm 129.21}$
20	US	$160.31_{\pm 0.38}$	$166.94_{\pm 0.07}$	$5160.80_{\pm 145.69}$
	CS	$\mathbf{27.77}_{\pm 2.80}$	$\mathbf{28.77}_{\pm 2.30}$	$5244.40_{\pm 86.73}$
30	US	$160.12_{\pm 0.43}$	$166.12_{\pm 1.71}$	$5958.20_{\pm 322.21}$
	CS	$\mathbf{24.49}_{\pm 5.45}$	$\mathbf{25.33}_{\pm 5.66}$	$5897.80_{\pm 209.25}$

Ablation experiments on τ: From Table 8, a smaller τ leads to faster computation and lower cost. Thus, we recommend using a relatively small τ.

Table 8: Comparison of our CS method using varying parameter τ of noise. We fix the sample size as 504, $\sigma = 1$, and $\zeta = 0.1$.

τ	SM	SS	cost-nd(↓)	cost-cd(↓)	Runtime (↓)
5	CS	504	**43.64**	**44.89**	4032.00
10	CS	504	46.53	48.33	4225.40
20	CS	504	49.00	51.02	5145.33
50	CS	504	57.75	58.37	11173.33
100	CS	504	53.51	57.49	31920.00

Selection of ζ: Each of our data items contains 0.1 mass of noise, but we run experiments with varying ζ. The results in Table 6 confirm that slightly overestimating ζ has only a minor impact, while underestimating it severely degrades the solution quality. Therefore, when the actual noise mass is unknown, we recommend setting ζ slightly larger than the expected noise mass.

Table 9: Comparison of our CS method with varying values of ζ on S3DIS dataset with 0.1 mass of noise. We fix the sample size as 504 and $\sigma = 1$.

ζ	SM	cost-nd($\downarrow$)	cost-cd($\downarrow$)	Runtime($\downarrow$)
0.05	CS	$104.96_{\pm 11.80}$	$206.60_{\pm 6.81}$	$3991.60_{\pm 278.49}$
0.1	CS	$46.53_{\pm 2.83}$	$48.33_{\pm 3.04}$	$4225.40_{\pm 129.21}$
0.2	CS	$43.52_{\pm 52.93}$	$52.93_{\pm 7.41}$	$4148.60_{\pm 156.73}$
0.3	CS	$45.62_{\pm 3.32}$	$59.94_{\pm 3.96}$	$4237.60_{\pm 40.49}$