

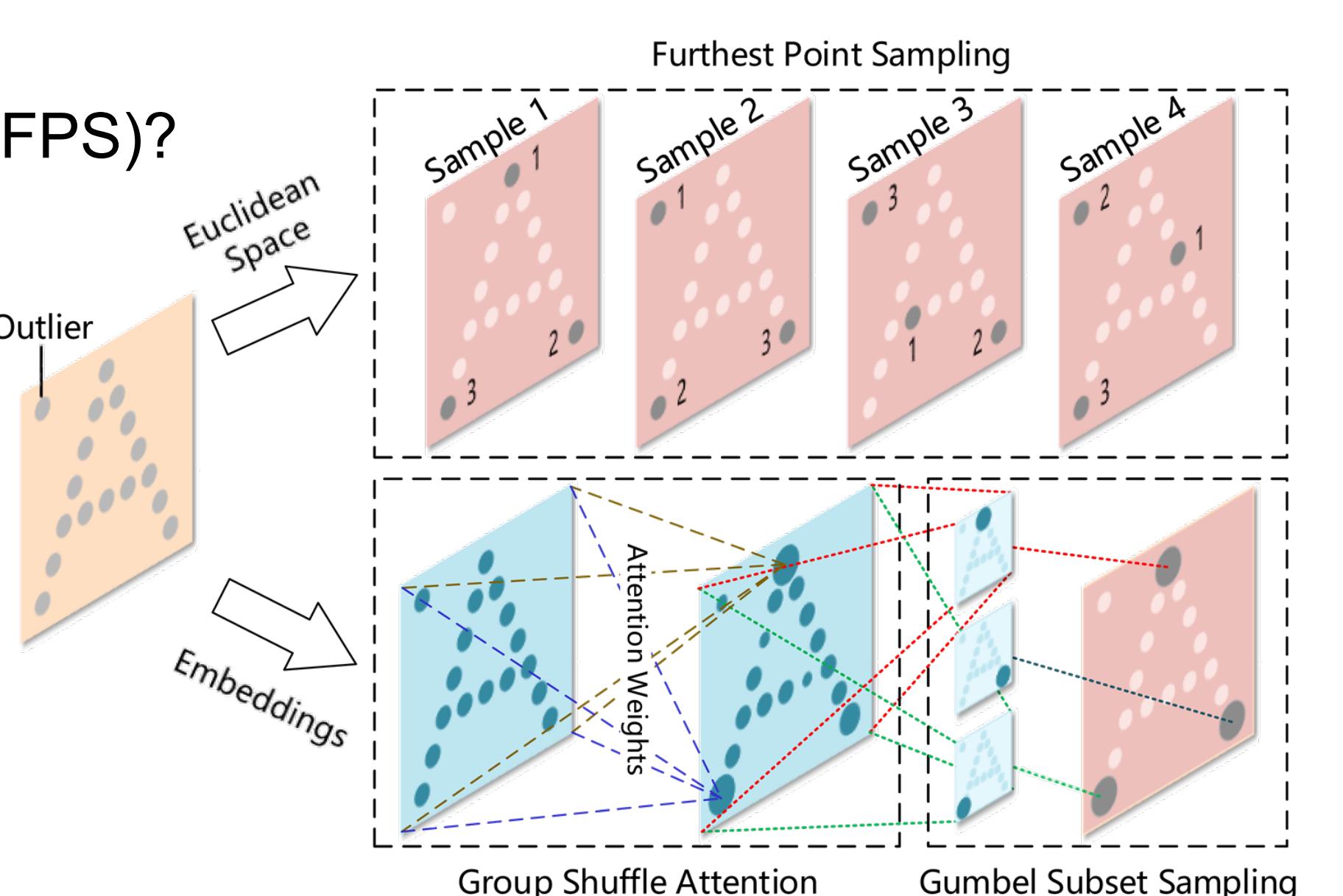
Introduction

➤ Research questions on modeling point clouds

- Relation learning between points?
- Efficiency in self-attention?
- Improve down-sampling over FPS?
- End-to-end point cloud down-sampling?

➤ What is wrong with Furthest Point Sampling (FPS)?

- Permutation-variant
- Sensitive to outliers
- Not learnable



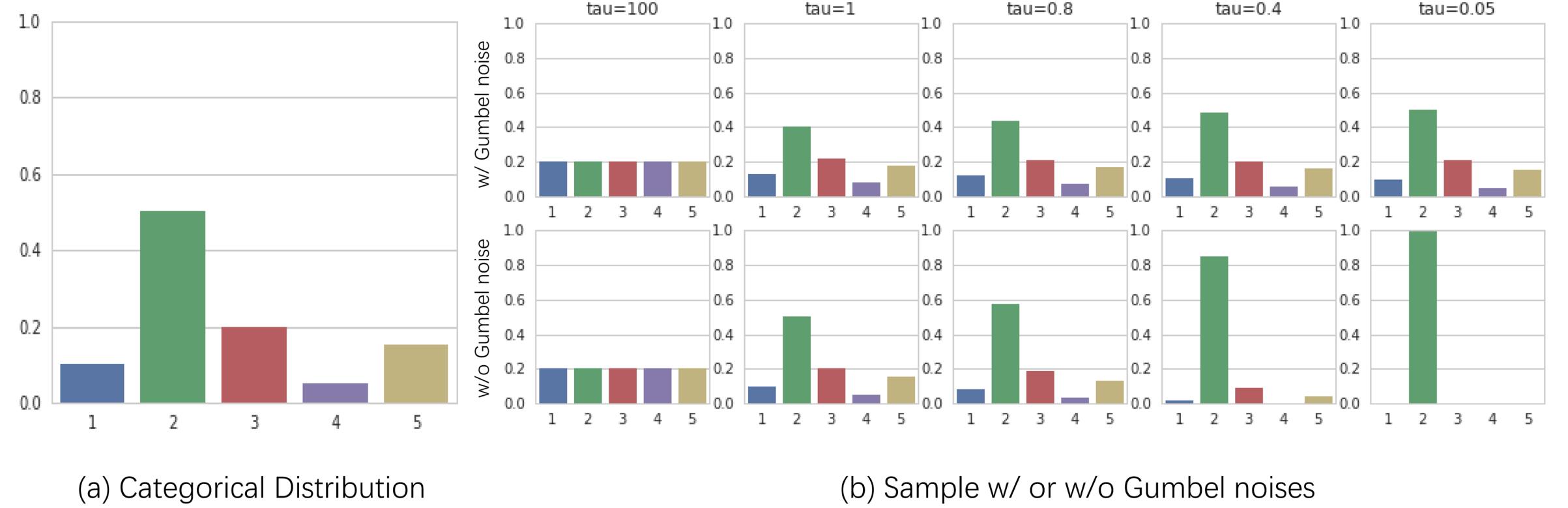
Preliminaries

➤ Self-attention

- Multi-Head Attention from "Attention Is All You Need".

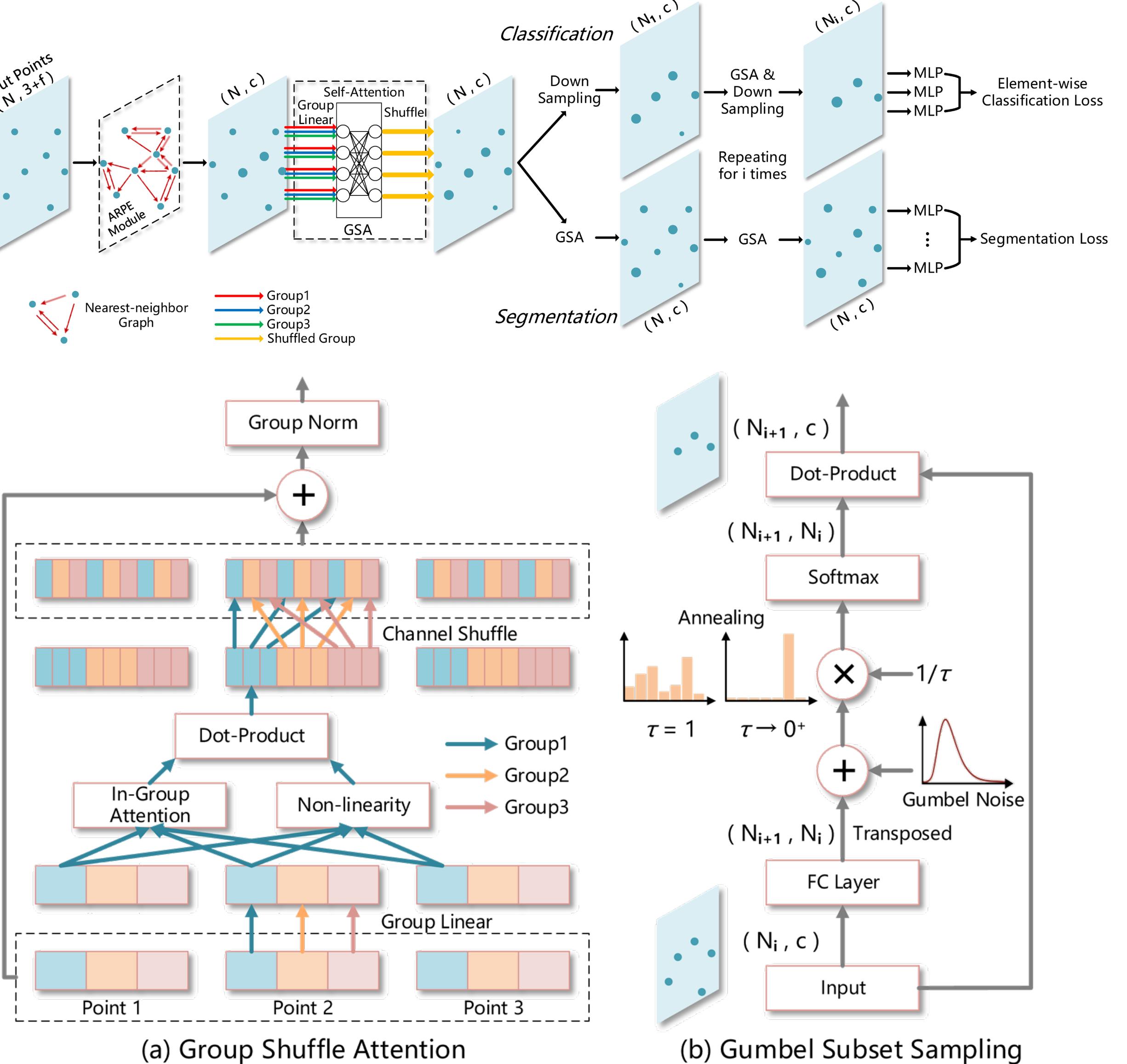
➤ Gumbel reparameterization

- Un-biased discrete reparameterization for categorical distribution.
- Low variance (compared with REINFORCE / Straight-Through Estimators).



Methodology

➤ Point Attention Transformer architecture



- Group Shuffle Attention: parameter-efficient self-attention
 - Self-attention to consume **size-varying** and **order-invariant** point sets.
 - Replace Multi-Head Attention by **channel grouping** and **shuffling**.
 - Proven **permutation-equivariance**: symmetric function for sets.
- Gumbel Subset Sampling: learnable point cloud down-sampling
 - Proven **permutation-invariance**: consistent down-sampling results.
 - **Sampling from embedding space**: less sensitive to outliers.
 - **Differentiable**: end-to-end learnable.

Applications & Results

➤ Benchmarks

Classification: ModelNet40

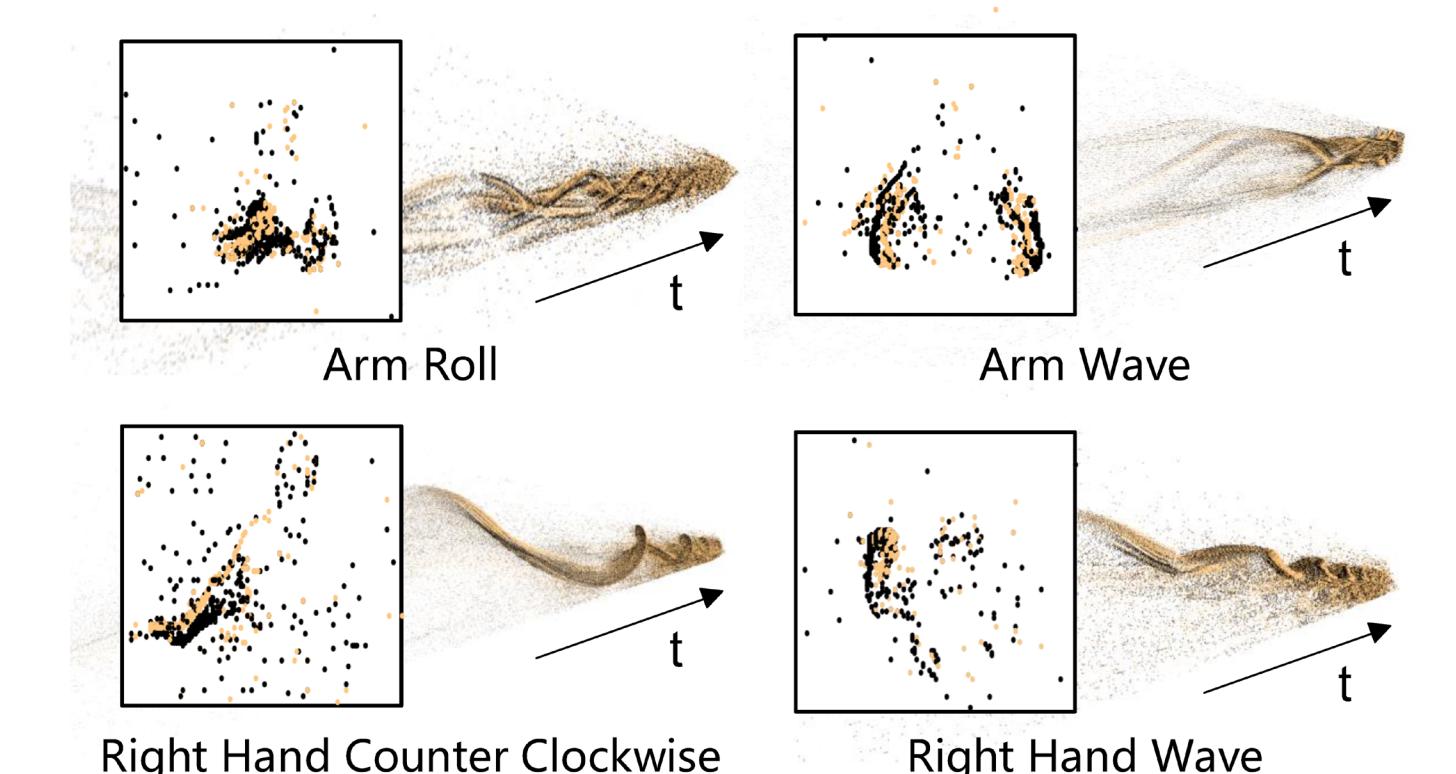
Method	Points	Accuracy (%)
DeepSets	5000	90.0
PointNet	1024	89.2
Kd-Net	1024	90.6
PointNet++	1024	90.7
KCNet	1024	91.0
DGCNN	1024	92.2
PointCNN	1024	92.2
PAT (GSA only)	1024	91.3
PAT (GSA only)	256	90.9
PAT (FPS)	1024	91.4
PAT (FPS + GSS)	1024	91.7

Segmentation: S3DIS

Method	mIoU	mIoU Area 5	Size (MB)
RSNet	56.47	-	-
SPGraph	62.1	58.04	-
PointNet	47.71	47.6	4.7
DGCNN	56.1	-	6.9
PointCNN	65.39	57.26	46.2
PAT	64.28	60.07	6.1

➤ Event camera stream as spatio-temporal point clouds: DVS128 Gesture Recognition

Event Camera Stream



Method

Method	10-CLASS	11-CLASS	Time (ms)	Size (MB)
CNN ¹	96.5	94.4	-	-
PointNet	89.5	88.8	2.8	6.5
PointNet++	95.6	95.2	18.2	12
PAT (GSA only)	96.9	95.6	16.9	5
PAT (GSA only, N256)	96.9	95.6	7.5	5
PAT (FPS)	96.5	95.2	12.7	5
PAT (FPS + GSS)	97.4	96.0	13.1	5.8

¹Running on a low-power processor,

Sliding Window Pre-processing

