Learning Robot Structure and Pose Embeddings using Graph Neural Networks

J. Taery Kim [1], Jeongeun Park [2], Sungjoon Choi [2], Sehoon Ha [1]

[1] Georgia Institute of Technology [2] Korea University

ICML International Conference On Machine Learning

Georgia Tech. KOREA UNIVERSITY

1. Motivation: Can we learn embeddings to represent robotic data?

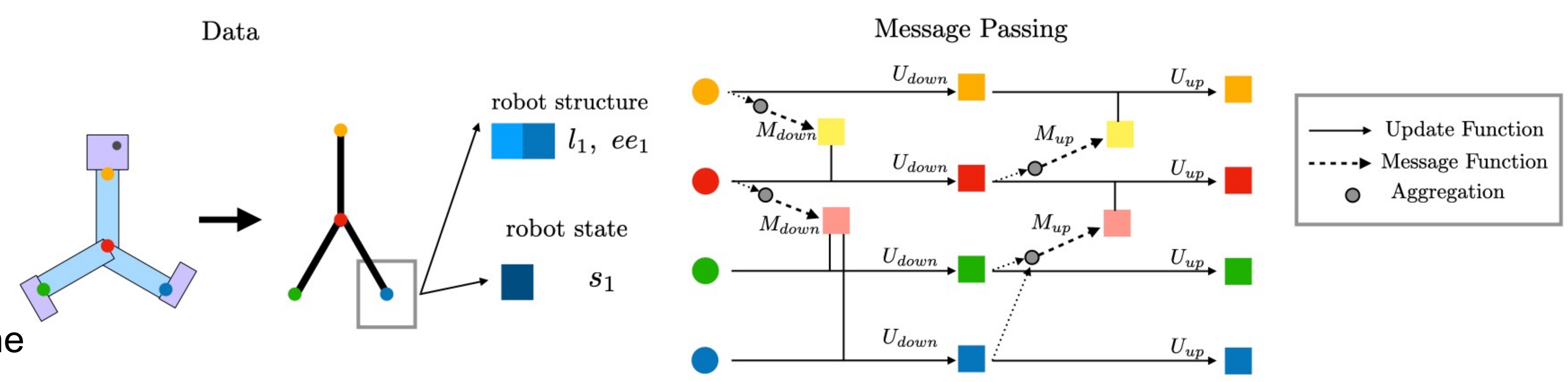
- Finding a compact and low-dimensional embedding space for complex phenomena is a key for understanding robots' behaviors.
- However, although numerous applications deal with various types of structural and motion data, the embedding of the generated data has been relatively less studied by roboticists.
- To this end, our work aims to learn embeddings for two types of robotic data: robot's design structure and pose data.

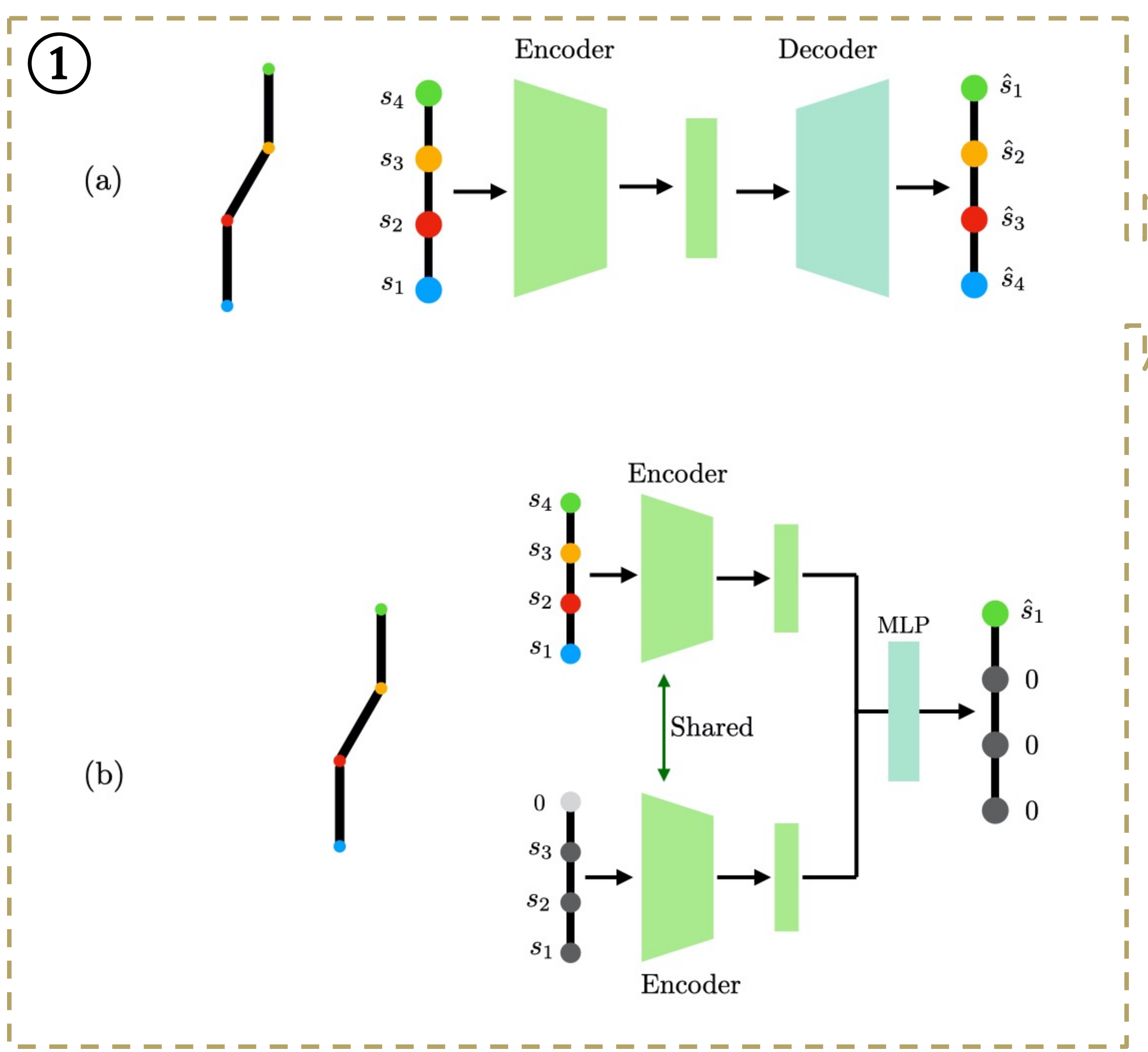
Visit our page!

2. Method: Learning Robot Embedding with GNN

❖ Tree Message Passing

- Data: Robot structure converted to tree-structured data with nodes and edges.
- Message Passing:
  - *downward passing* – to spread the information from parent to child nodes
  - Then, *upward passing* – to aggregate the information at the root node.



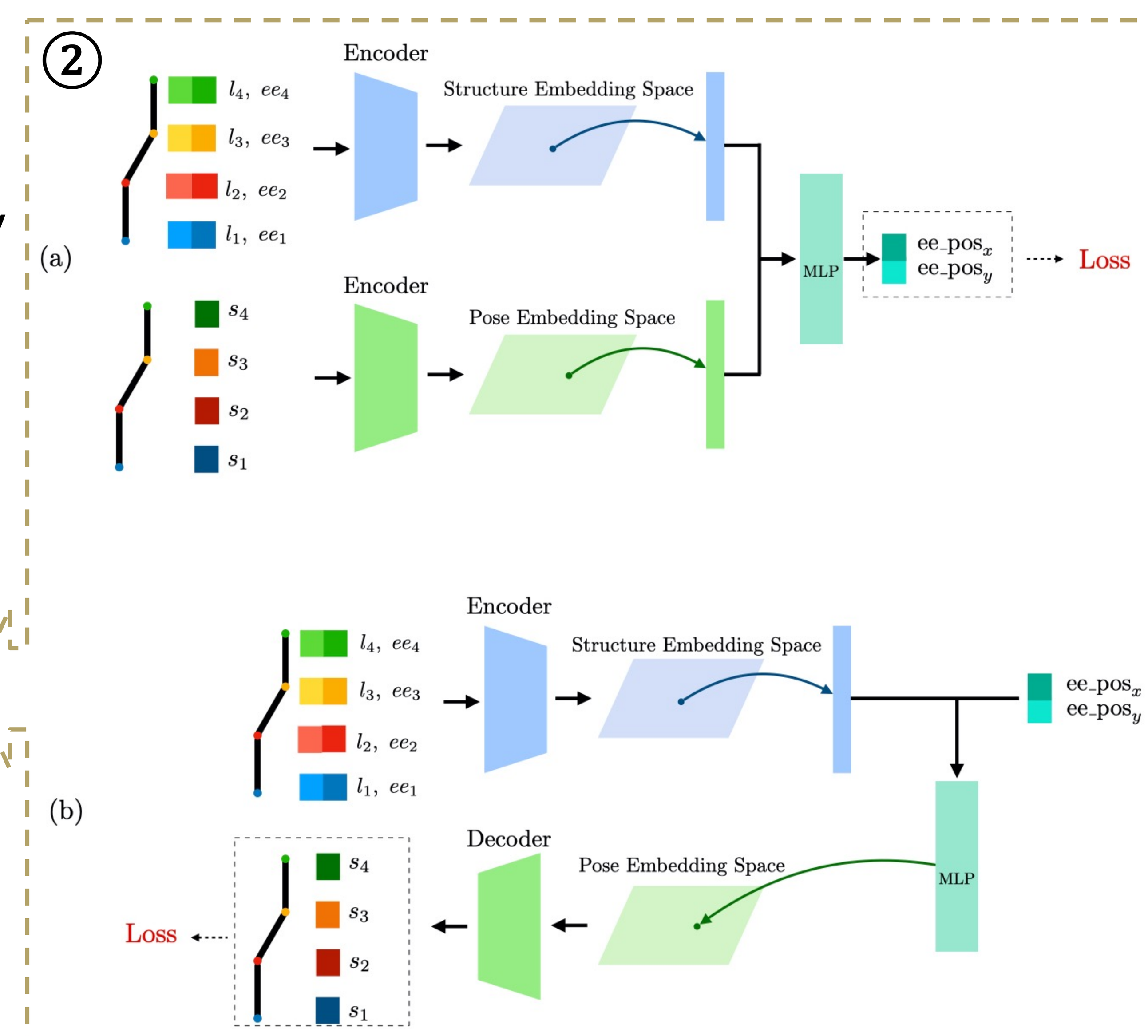


① **Pretraining of Encoders**

- To facilitate the learning process, we pretrain structural encoders by learning **self-reconstructing tasks**.
- Two strategies are compared:
  (a) encoding-decoding (ED)
  (b) fill-in-the-blank (FB)

② **Multi-task Representation Learning on Kinematics**

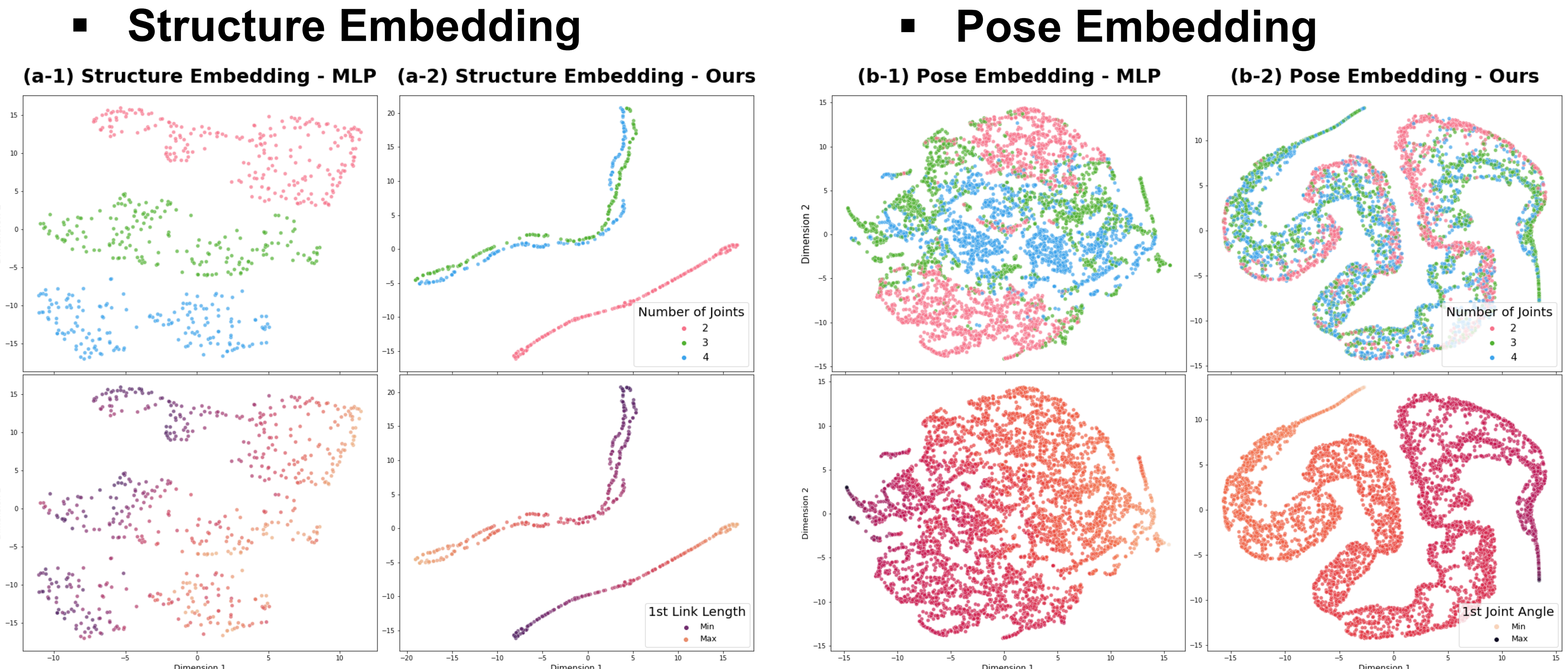
- Goal: make the **embedding space of the robot's kinematics**
- Two tasks are jointly solved:
  (a) forward kinematics (FK)
  (b) inverse kinematics (IK)



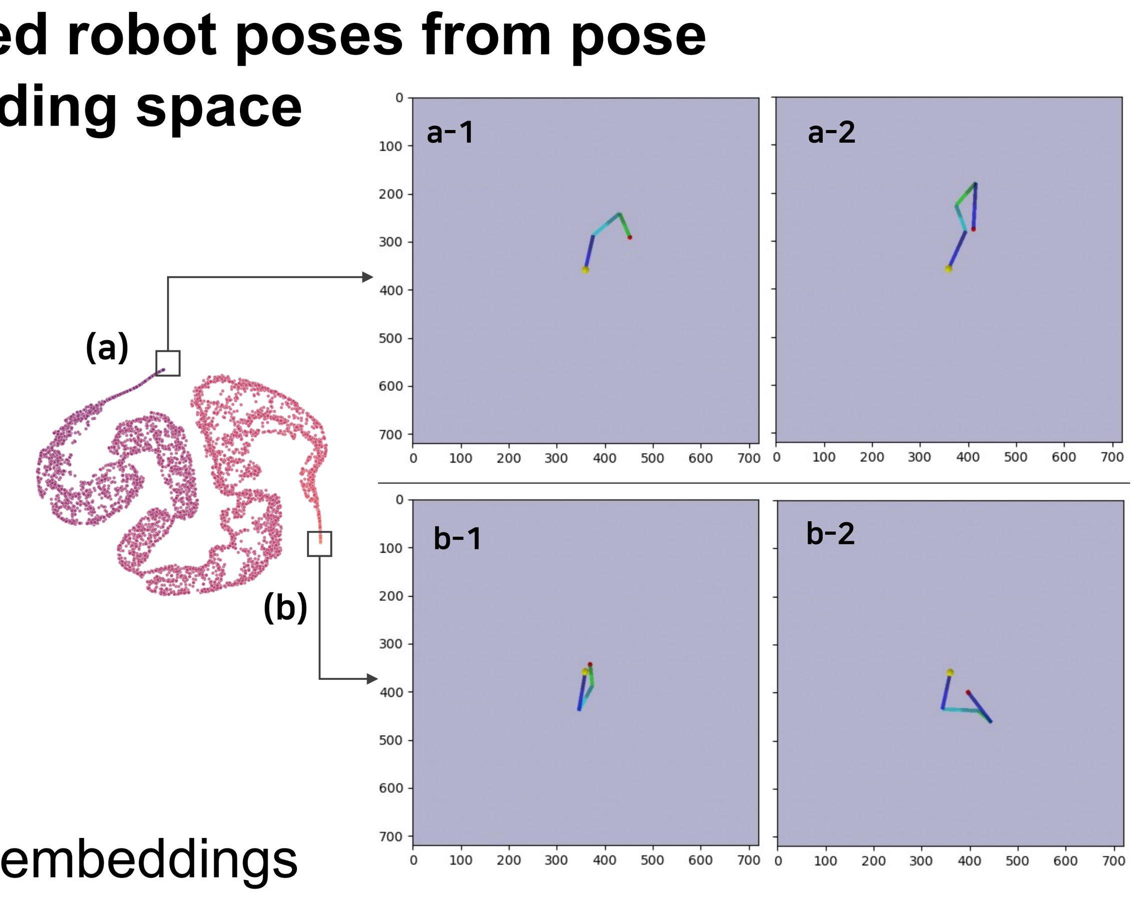
3. Experiments and Results

❖ Embedding Space Visualization

- **Structure Embedding**
- **Pose Embedding**

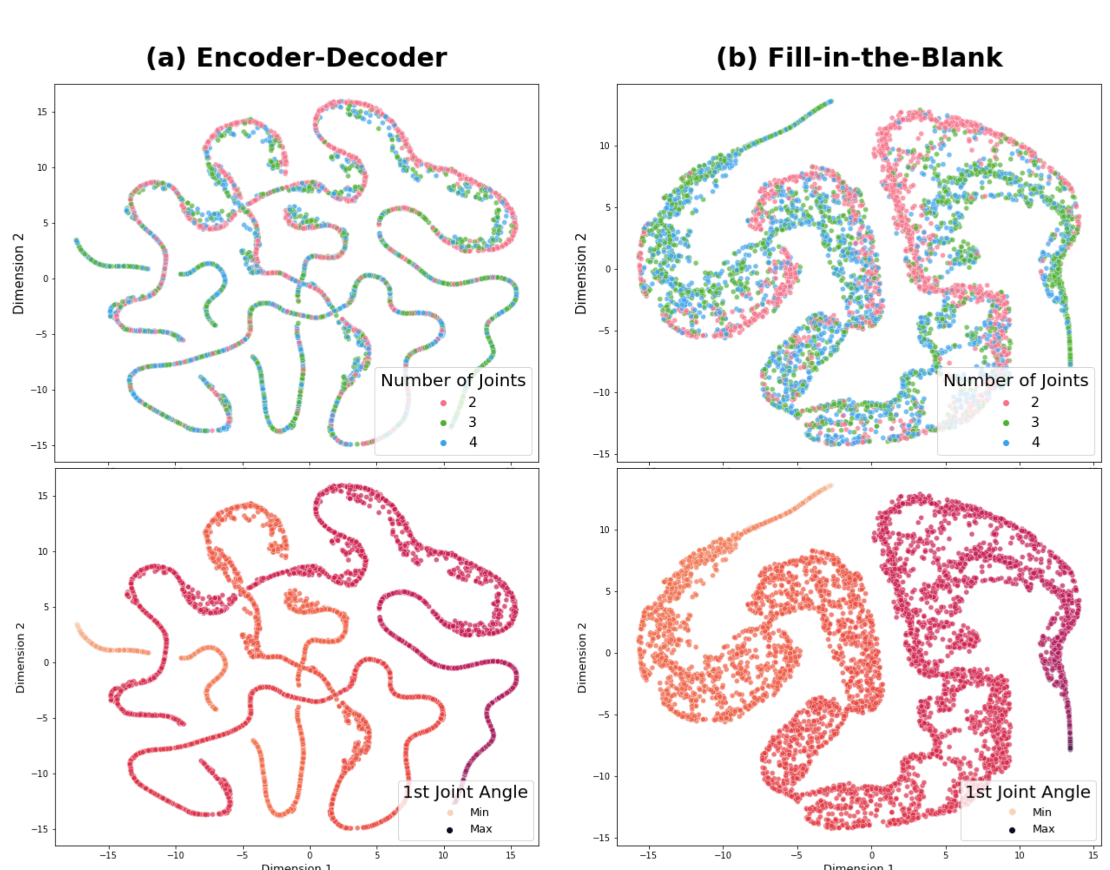


- **Sampled robot poses from pose embedding space**



- Pose embeddings mostly determined by the first joint angle

❖ Reconstruction Strategies Comparison



(a) ED reconstruction:
- No specific correlation shown with neither the number of node nor node features

(b) FB reconstruction:
- Subject to the node features showing in more stretched shapes

→ FB distinguishes each input in detail, **expected to facilitate further tuning in the embedding space.**

4. Conclusion and Future Work

- We leverage the **tree structure** existing in kinematic structures to encode the robot data into the corresponding embedding space.
- We further incorporate the **kinematic movements** of the robot through the **multi-task learning** to learn more meaningful representations.
- We test the tree message passing on a robot with a simple linear structure and visualize the embeddings.
- In the future, we hope to **add more complexity to the robot structure**, such as more nodes, various types of revolute joints, and graph structures including closed chain.