Score-PA: Score-based 3D Part Assembly - Supplementary

Junfeng Cheng[1]
junfeng.cheng20@imperial.ac.uk
Mingdong Wu[2]
wmingd@pku.edu.cn
Ruiyuan Zhang[3]
zhangruiyuan@zju.edu.cn
Guanqi Zhan[4]
guanqi@robots.ox.ac.uk
Chao Wu[5]
chao.wu@zju.edu.cn
Hao Dong[†2]
hao.dong@pku.edu.cn

[1] Department of Electrical and Electronic Engineering,
Imperial College London,
London, UK
[2] School of CS,
Peking University,
Beijing, China
[3] College of Computer Science and Technology,
Zhejiang University,
Hangzhou, China
[4] Department of Engineering Science,
University of Oxford,
Oxford, UK
[5] School of Public Affairs,
Zhejiang University,
Hangzhou, China

1 More Discussion about Our Score-PA

Selection of the weight function $\lambda(t)$ We discuss our training algorithm in the main text. $\lambda(t)$ is important for our training objective function. According to Song et al. [1], we need to choose a suitable $\lambda(t)$ to make our prior distribution $p(\mathbf{Q_P}(T))$ independent from the data distribution and easily sampled (We set $T = 1$). In our algorithm, the weight function is selected as $\lambda(t) = \frac{1}{2\log\sigma}(\sigma^{2t} - 1)$, which follows Song et al.'s setting [1]. We already have our SDE $d\mathbf{Q_P} = \sigma^t d\mathbf{w}$, where $t \in [0,1]$, and in this situation,

$$p_{0t}(\mathbf{Q_P}(t) \mid \mathbf{Q_P}(0), \mathbf{P}) = \mathcal{N}\left(\mathbf{Q_P}(t); \mathbf{Q_P}(0), \frac{1}{2\log\sigma}(\sigma^{2t} - 1)\mathbf{I}\right) \tag{1}$$

We have our weight function $\lambda(t) = \frac{1}{2\log\sigma}(\sigma^{2t} - 1)$, and then the prior distribution $p_{t=1}$ is,

$$\int p_0(\mathbf{y})\mathcal{N}\left(\mathbf{Q_P}; \mathbf{y}, \frac{1}{2\log\sigma}(\sigma^2 - 1)\mathbf{I}\right) d\mathbf{y} \approx \mathcal{N}\left(\mathbf{Q_P}; \mathbf{0}, \frac{1}{2\log\sigma}(\sigma^2 - 1)\mathbf{I}\right), \tag{2}$$

where σ should be large enough. The Equation 2 means that the prior distribution can be easily sampled from a normal distribution.

† Corresponding Author

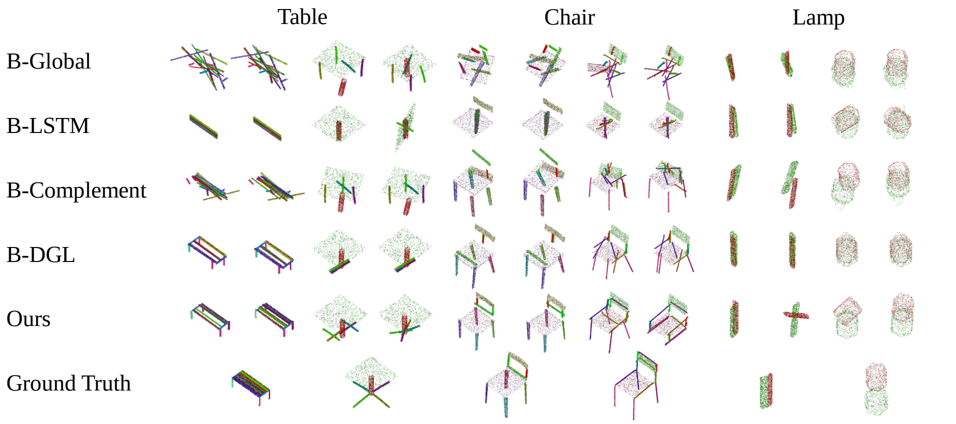

Figure 1: More qualitative comparisons between our algorithm and other baselines.

2 More Qualitative Comparisons

We present more qualitative comparisons between our algorithm and other baselines in Figure 1. Similar to the comparisons in the main text, we present two assembly results per input parts **P** for each algorithm. The comparisons show that only our algorithm is able to generate diverse results with high quality.

3 Details about Our Experiments

Training details As discussed in Section 3 of our paper, our algorithm has two important hyper-parameters T and σ in the training procedure. In our experiments on the three datasets, we set $T = 1.0$, and $\sigma = 25.0$. We train our models with 2000 Epochs on Chair and Table datasets, and 4000 Epochs on Lamp dataset. The learning rate for all datasets are set as 10^{-4}, and the Optimizer is Adam. The training batch size for all the datasets is 16.

Other details We conducted both training and testing experiments using a single RTX 3090 GPU. To ensure reproducibility, we set a fixed random seed. In our ablation study, we define the sampling steps for FPC and FPC w/o decay as $steps = N + C_F$. For the PC sampler, the number of sampling steps is simply N, as it does not involve a decay stage. In all testing experiments, including the ablation study, we set the sampling batch size for Score-PA to 4.

4 Limitation and Future Work

Currently, we achieve diverse part assembly in an ideal simulation environment. In the future, we plan to take the physical factors (*e.g.*, physical collision) into our consideration, and achieve autonomous part assembly in a real physical environment.

References

[1] Yang Song, Jascha Sohl-Dickstein, Diederik P Kingma, Abhishek Kumar, Stefano Ermon, and Ben Poole. Score-based generative modeling through stochastic differential equations. *arXiv preprint arXiv:2011.13456*, 2020.