Unsupervised Sampling Promoting for Stochastic Human Trajectory Prediction Supplementary Material

1. The algorithm of BOsampler

We give the algorithm to give a short but clear summarization of **BOsampler** . The detailed algorithm is shown in Algorithm 1.

Algorithm 1 Sampling procedure of **BOsampler**

Require: generator G, observed trajectory X, pseudo score evaluation function f.
  for n in $\{1,...,\mathrm{w}\}$ **do** ▷ Warm-up
    Sample $\mathrm{z}_n \sim p_{\mathrm{z}}$
    $\hat{Y}_n \leftarrow G(X, \mathrm{z}_n)$
  end for
  for n in $\{\mathrm{w}+1,...,\mathrm{N}\}$ **do** ▷ **BOsampler**
    Fit the Gaussian Process $\mathcal{GP}$ by $\mathrm{z}_{1:n-1}$ and $f(\mathrm{z}_{1:n-1})$
    Use the posterior of $\mathcal{GP}$ to build the acquisition function $\phi_n(\mathrm{z})$
    $\mathrm{z}_n \leftarrow \underset{\mathrm{z}}{\mathrm{argmax}} \phi_n(\mathrm{z})$
    $\hat{Y}_n \leftarrow G(X, \mathrm{z}_n)$
  end for
  Return $\hat{Y} = \{\hat{Y}_n | n \in [1, ..., N]\}$

2. Complete quantitative experimental results on ETH/UCY

We present the complete experiment results on five scenes of ETH/UCY datasets in Table 1. For all baseline methods, **BOsampler** consistently outperforms the MC sampling method, which shows the effectiveness of the proposed method, though not much. **BOsampler** also shows an improvement over the QMC method on most baselines. **BOsampler** doesn't have the same level of performance gain in the complete set of ETH/UCY compared to the exception subset. It is because the uncommon trajectories only comprise a small portion of the dataset. **BOsampler** achieves more performance gain on the ETH dataset. The reason is probably the same: uncommon trajectories show up more frequently in ETH dataset than in the other four datasets in ETH/UCY. Although **BOsampler** does not achieve a huge improvement in the complete set, considering **BOsampler** has a significant gain in the exception subset, it is clear that **BOsampler** balances exploitation of the baseline model's distribution and exploration of the edge cases. So **BOsampler** helps promote the robustness of the sampling process of baseline models. We highlight that **BOsampler** focus on trajectories with low probability. Although these cases are the minority, it is meaningful and crucial to consider them in intelligent transportation and autonomous driving.

3. Experiments on exception subsets with different ratios

In Section 4.2 of the main manuscript, we set the ratio of the exception set as 4%. In order to verify the validity of the chosen subset, we further provide a 12% variation, which demonstrates a consistent performance trend with the original subset, thus confirming the representativeness and appropriateness of the ratio setting of our experiments. The experiment results are presented in Table 2. When the ratio goes to 12%, **BOsampler** still achieves considerably better performance than baselines.

4. Visualization of the failure cases

We provide the visualization of the failure cases to better understand the method. From the visualization in Figure 1, we found **BOsampler** may lose the ground truth trajectory when the most-likely prediction is far away from the ground truth. It is not surprising since our method is based on the assumption of the good prior distribution. Solving this problem without accessing more data is not trivial. A potential solution is to modify the prior distribution using the testing data during the inference [6].

5. Demo

We present a demonstration video to illustrate how **BOsampler** iteratively samples trajectories. The video is named "demo.mp4" which can be found in the attachment.

References

[1] Agrim Gupta, Justin Johnson, Li Fei-Fei, Silvio Savarese, and Alexandre Alahi. Social gan: Socially acceptable trajectories

Table 1. Quantitative results on the ETH/UCY dataset with Best-of-20 strategy in ADE/FDE metric. Lower is better. * updated version of Trajectron++

Baseline Model	Sampling	ETH		HOTEL		UNIV		ZARA1		ZARA2		AVG	
		ADE	FDE	ADE	FDE	ADE	FDE	ADE	FDE	ADE	FDE	ADE	FDE
Social-GAN [1]	MC	0.77	1.40	0.43	0.88	0.75	1.50	0.35	0.69	0.36	0.72	0.53	1.05
	QMC	0.76	1.38	0.43	0.87	0.75	1.50	0.34	0.69	0.35	0.72	0.53	1.03
	BOsampler	0.73	1.28	0.43	0.87	0.75	1.50	0.34	0.69	0.35	0.71	0.52	1.01
	BOsampler + QMC	0.72	1.26	0.43	0.87	0.74	1.49	0.34	0.69	0.35	0.71	0.52	1.00
Trajectron++ [3]	MC	0.43	0.86	0.12	0.19	0.22	0.43	0.17	0.32	0.12	0.25	0.21	0.41
	QMC	0.43	0.84	0.12	0.19	0.22	0.42	0.17	0.31	0.12	0.24	0.21	0.40
	BOsampler	0.34	0.64	0.12	0.21	0.18	0.40	0.14	0.30	0.11	0.24	0.18	0.36
	BOsampler + QMC	0.34	0.64	0.12	0.21	0.18	0.40	0.14	0.30	0.11	0.24	0.18	0.36
Trajectron++ [3]*	MC	0.57	1.06	0.16	0.26	0.30	0.61	0.22	0.42	0.16	0.33	0.28	0.54
	QMC	0.57	1.05	0.16	0.26	0.30	0.61	0.22	0.42	0.16	0.33	0.28	0.54
	BOsampler	0.49	0.82	0.15	0.23	0.27	0.53	0.20	0.38	0.15	0.29	0.25	0.45
	BOsampler + QMC	0.49	0.82	0.15	0.23	0.27	0.53	0.20	0.38	0.15	0.29	0.25	0.45
PECNet [4]	MC	0.61	1.07	0.22	0.39	0.34	0.56	0.25	0.45	0.19	0.33	0.32	0.56
	QMC	0.60	1.04	0.21	0.38	0.33	0.53	0.24	0.43	0.18	0.31	0.31	0.54
	BOsampler	0.56	0.92	0.21	0.38	0.32	0.52	0.24	0.42	0.18	0.31	0.30	0.51
	BOsampler + QMC	0.56	0.91	0.21	0.37	0.31	0.51	0.24	0.41	0.18	0.31	0.30	0.50
Social-STGCNN [5]	MC	0.65	1.10	0.50	0.86	0.44	0.80	0.34	0.53	0.31	0.48	0.45	0.75
	QMC	0.62	1.03	0.38	0.57	0.36	0.63	0.32	0.52	0.29	0.50	0.39	0.65
	BOsampler	0.57	0.90	0.44	0.82	0.43	0.76	0.34	0.54	0.26	0.45	0.41	0.69
	BOsampler + QMC	0.49	0.74	0.39	0.73	0.41	0.72	0.32	0.52	0.26	0.40	0.37	0.62
STGAT [2]	MC	0.74	1.34	0.35	0.68	0.56	1.20	0.34	0.68	0.29	0.59	0.46	0.90
	QMC	0.73	1.32	0.35	0.67	0.56	1.20	0.34	0.68	0.29	0.59	0.45	0.89
	BOsampler	0.70	1.15	0.35	0.67	0.55	1.17	0.34	0.68	0.29	0.59	0.44	0.85
	BOsampler + QMC	0.68	1.11	0.35	0.67	0.55	1.17	0.33	0.67	0.30	0.59	0.44	0.84

	top 4%	top 12%	full
MC	1.21/2.33	0.88/1.76	0.32/0.56
QMC	1.20/2.33	0.86/1.70	0.31/0.54
`BOsampler`	0.97/1.75	0.74/1.38	0.30/0.51

Table 2. The results of PECNet on ETH-UCY with different exception ratios.

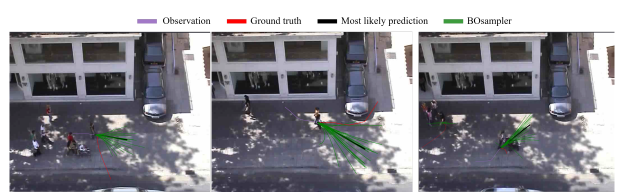

Figure 1. Failure cases: In the first and second graph, we show that when a pedestrian changes direction suddenly, our method may not predict such a change. In the third graph, when a pedestrian stops suddenly (e.g. a pram gets stuck in the ground), or when a pedestrian suddenly starts walking, our method may also fail.

with generative adversarial networks. In *CVPR*, pages 2255–2264, 2018. 2

[2] Yingfan Huang, HuiKun Bi, Zhaoxin Li, Tianlu Mao, and Zhaoqi Wang. Stgat: Modeling spatial-temporal interactions for human trajectory prediction. In *ICCV*, pages 6272–6281, 2019. 2

[3] Boris Ivanovic and Marco Pavone. The trajectron: Probabilistic multi-agent trajectory modeling with dynamic spatiotemporal graphs. In *ICCV*, pages 2375–2384, 2019. 2

[4] Karttikeya Mangalam, Harshayu Girase, Shreyas Agarwal, Kuan-Hui Lee, Ehsan Adeli, Jitendra Malik, and Adrien Gaidon. It is not the journey but the destination: Endpoint conditioned trajectory prediction. In *ECCV*, 2020. 2

[5] Abduallah Mohamed, Kun Qian, Mohamed Elhoseiny, and Christian Claudel. Social-stgcnn: A social spatio-temporal graph convolutional neural network for human trajectory prediction. In *CVPR*, 2020. 2

[6] Yu Sun, Xiaolong Wang, Zhuang Liu, John Miller, Alexei Efros, and Moritz Hardt. Test-time training with self-

supervision for generalization under distribution shifts. In *ICML*, pages 9229–9248, 2020. 1