

Figure S1. Normalized performance comparison between zero-shot inference and few-shot finetuning of UniSO-T across three tasks (RobotPush, Rover, and LunarLander). The light blue bars represent zero-shot performance, while the dark blue bars show the improved performance after few-shot finetuning, with error bars indicating performance variance. The red dashed line ($\mathcal{D}(\text{best})$) denotes the best score in the few-shot offline dataset.

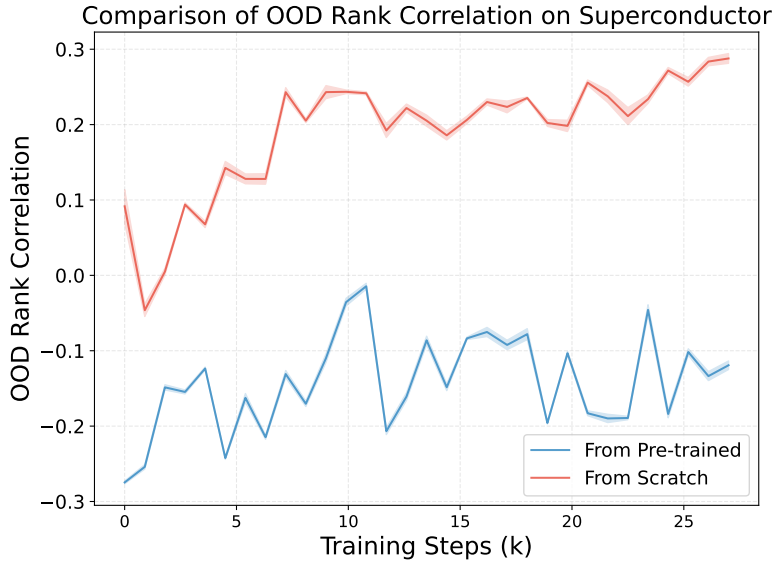


Figure S2. Comparison of OOD Spearman rank correlation on Superconductor task between UniSO-N from pre-trained embedder (blue) and from scratch (red). The OOD dataset is constructed following [Chen et al. \(2023\)](#). Higher correlation indicates better surrogate model for offline BBO ([Tan et al., 2025](#)).

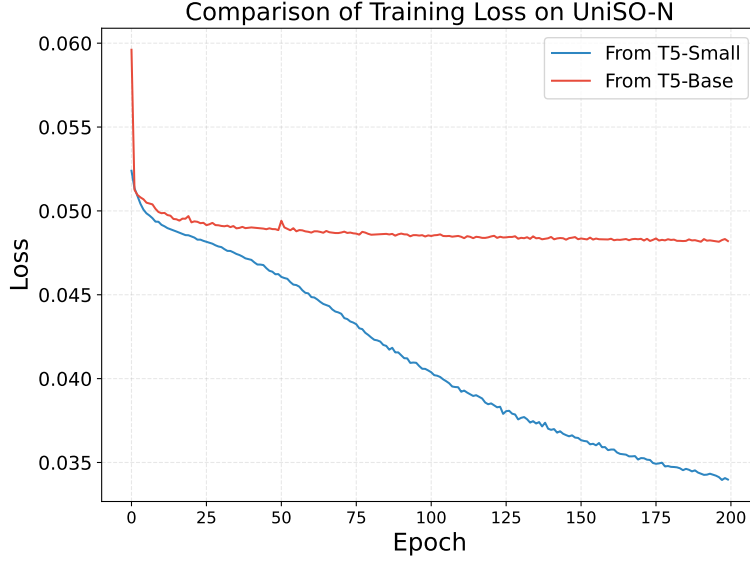


Figure S3. Training loss curves comparing UniSO-N with T5-Small versus T5-Base embedders during training.

Table S1. Performance comparison between UniSO-N with pre-trained and randomly initialized embedders on Design-Bench and SOO-Bench unconstrained tasks, where the better one is **Bold**. $\mathcal{D}(\text{best})$ denotes the best score in the offline dataset.

Task	$\mathcal{D}(\text{best})$	UniSO-N from scratch	UniSO-N from pre-trained
Ant	165.326	305.576 \pm 120.465	269.691 \pm 77.425
D’Kitty	199.363	223.104 \pm 23.838	173.911 \pm 46.662
Superconductor	74.000	101.754 \pm 6.972	67.333 \pm 10.838
TF Bind 8	0.439	0.883 \pm 0.048	0.833 \pm 0.005
TF Bind 10	0.005	0.578 \pm 0.108	0.959 \pm 0.115
GTOPX 2	-195.586	-102.632 \pm 22.635	-124.995 \pm 56.170
GTOPX 3	-151.190	-55.743 \pm 7.490	-62.622 \pm 22.261
GTOPX 4	-215.716	-102.684 \pm 19.227	-110.284 \pm 17.559
GTOPX 6	-112.599	-56.469 \pm 10.337	-57.435 \pm 18.832
Avg. Rank	/	1.111 \pm 0.314	1.889 \pm 0.314

Table S2. Comparison of UniSO-N equipped with embedders of different scales (e.g., T5-Small and T5-Base) on unconstrained tasks from Design-Bench and SOO-Bench, where the better one is **Bold**. $\mathcal{D}(\text{best})$ denotes the best score in the offline dataset.

Task	$\mathcal{D}(\text{best})$	UniSO-N + T5-Small	UniSO-N + T5-Base
Ant	165.326	269.691 \pm 77.425	255.336 \pm 209.169
D’Kitty	199.363	173.911 \pm 46.662	199.186 \pm 0.000
Superconductor	74.000	67.333 \pm 10.838	93.816 \pm 11.763
TF Bind 8	0.439	0.833 \pm 0.005	0.597 \pm 0.081
TF Bind 10	0.005	0.959 \pm 0.115	0.394 \pm 0.140
GTOPX 2	-195.586	-124.995 \pm 56.170	-163.866 \pm 74.375
GTOPX 3	-151.190	-62.622 \pm 22.261	-71.990 \pm 11.158
GTOPX 4	-215.716	-110.284 \pm 17.559	-96.400 \pm 39.728
GTOPX 6	-112.599	-57.435 \pm 18.832	-88.123 \pm 20.223
Avg. Rank	/	1.333 \pm 0.471	1.667 \pm 0.471