Table 1. FLOPs of different models.

	Ant-P	Ant-V	Hopper-P	Hopper-V	Walker-P	Walker-V
GFLOPs reward denoise	3.3354	3.3352	3.335	3.3349	3.3352	3.3351
GFLOPs observation denoise	3.3365	3.3359	3.3355	3.3352	3.336	3.3356
FLOPs causal	278528	180224	147456	98304	212992	147456

Table 2. Runtime per update compared with baselines (ms).

	Ant-P	Ant-V	Hopper-P	Hopper-V	Walker-P	Walker-V
CSR-ADM	208.95	219.06	221.69	209.09	211.93	219.99
SAC	25.03	25.06	24.35	24.6	24.74	24.17
DMBP	86.64	87.67	87.11	87.04	88.26	87.56
DBC	40.29	41.03	41.27	41.95	41.41	41.09

(a) Ant-P   (b) Hopper-P   (c) Walker-P

(d) Ant-V   (e) Hopper-V   (f) Walker-V
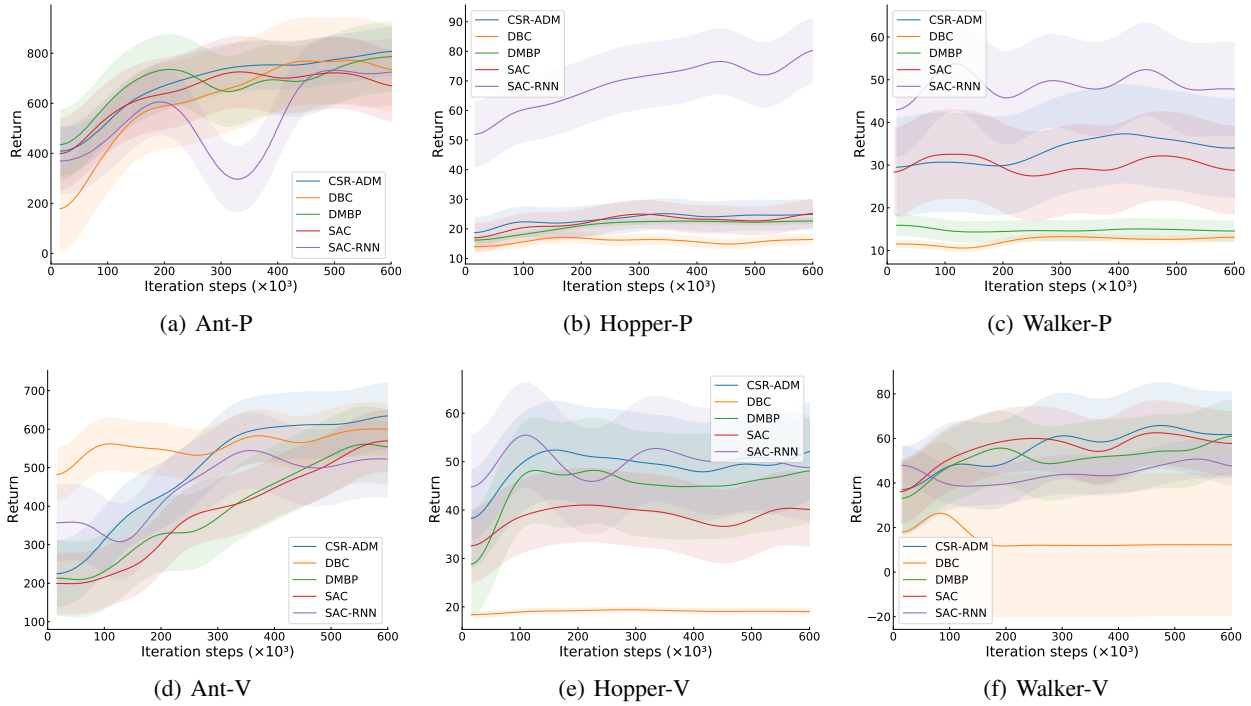
Figure 1. Comparison of CSR-ADM and baselines on six environments.

1

(a) Ant-P
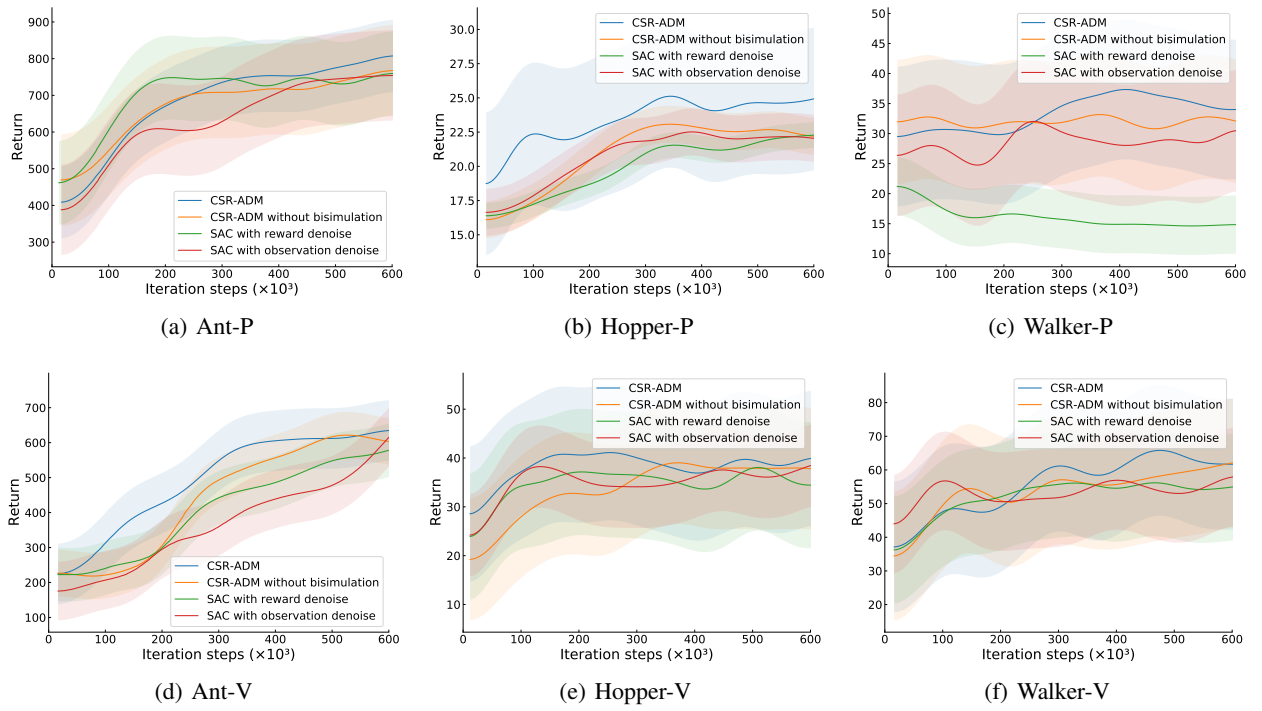
(b) Hopper-P

(c) Walker-P

(d) Ant-V

(e) Hopper-V

(f) Walker-V

Figure 2. Ablation studies of CSR-ADM on six environments.