

Table 1: **Evaluation results** of PPO and SAC with the new baselines.

PPO Results								
	Lunar Lander Safe				Point Robot Navigation			
	50k		best		250k		best	
	r	c	r	c	r	c	r	c
PPO base	-126 ± 158	77 ± 29	225 ± 100	10 ± 30	15 ± 15	3 ± 3	31 ± 31	3 ± 2
Lagrange	-24 ± 146	54 ± 39	204 ± 116	12 ± 24	15 ± 16	3 ± 3	24 ± 5	3 ± 2
V1	101 ± 84	41 ± 19	205 ± 78	7 ± 16	25 ± 5	3 ± 3	30 ± 4	2 ± 1
V2	89 ± 122	44 ± 34	251 ± 28	5 ± 9	11 ± 21	3 ± 3	17 ± 15	3 ± 1
V3	144 ± 4	26 ± 22	264 ± 5	1 ± 2	27 ± 5	5 ± 2	29 ± 3	3 ± 3
CPO	-200 ± 227	100 ± 1	-42 ± 21	0 ± 0	-	-	-	-
FOCOPS	-129 ± 21	64 ± 24	117 ± 80	30 ± 19	17 ± 11	0 ± 0	33 ± 2	0 ± 0

SAC Results								
	50k				50k			
	50k		best		50k		best	
	r	c	r	c	r	c	r	c
SAC base	90 ± 108	10 ± 16	181 ± 117	3 ± 6	22 ± 19	1 ± 1	38 ± 1	1 ± 1
Lagrange	111 ± 105	17 ± 13	184 ± 128	2 ± 3	29 ± 8	0 ± 0	38 ± 1	0 ± 0
Mult	-35 ± 27	3 ± 5	-35 ± 27	3 ± 5	34 ± 3	0 ± 0	38 ± 1	0 ± 0
Mult Clipped	134 ± 94	14 ± 13	243 ± 49	8 ± 15	33 ± 4	0 ± 0	38 ± 1	0 ± 0
Mult Lagrange	125 ± 59	29 ± 15	251 ± 20	2 ± 2	37 ± 2	0 ± 0	37 ± 2	0 ± 0

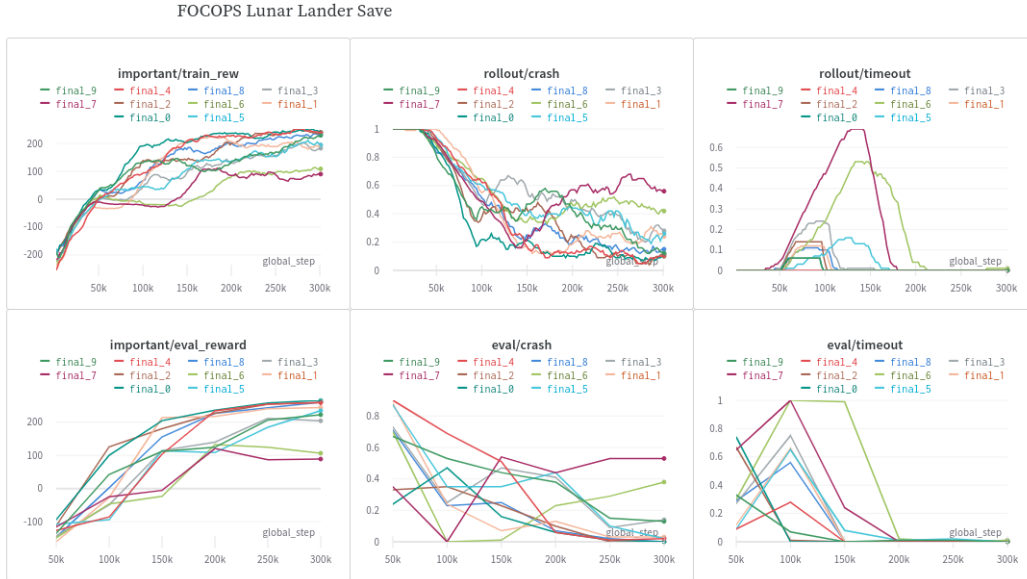


Figure 1: FOCOPS on Lunar Lander Safe. Note that for the results in the paper we trained all PPO versions for only 150k steps.

FOCOPS Point Robot Navigation

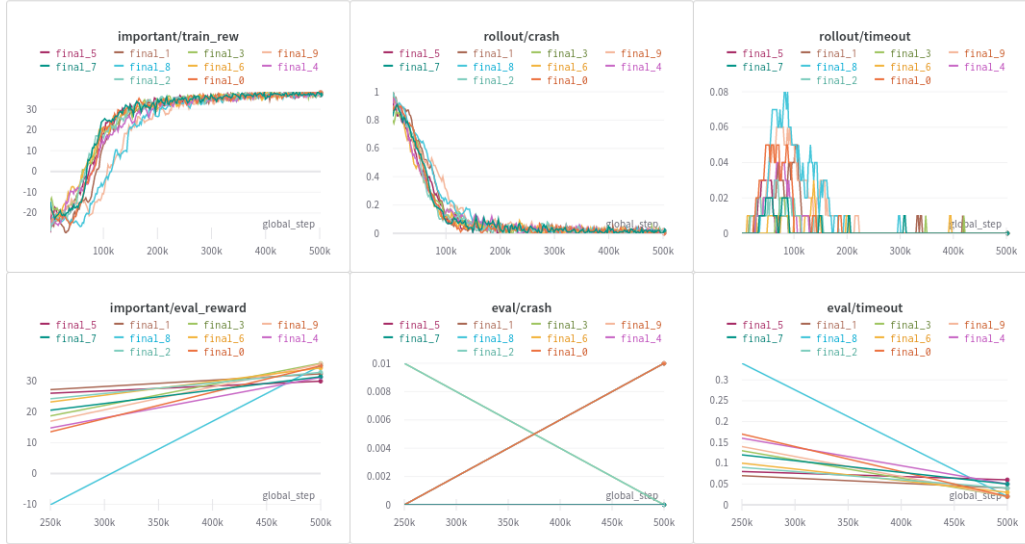


Figure 2: FOCOPS on Point Robot Navigation.

CPO Lunar Lander Safe

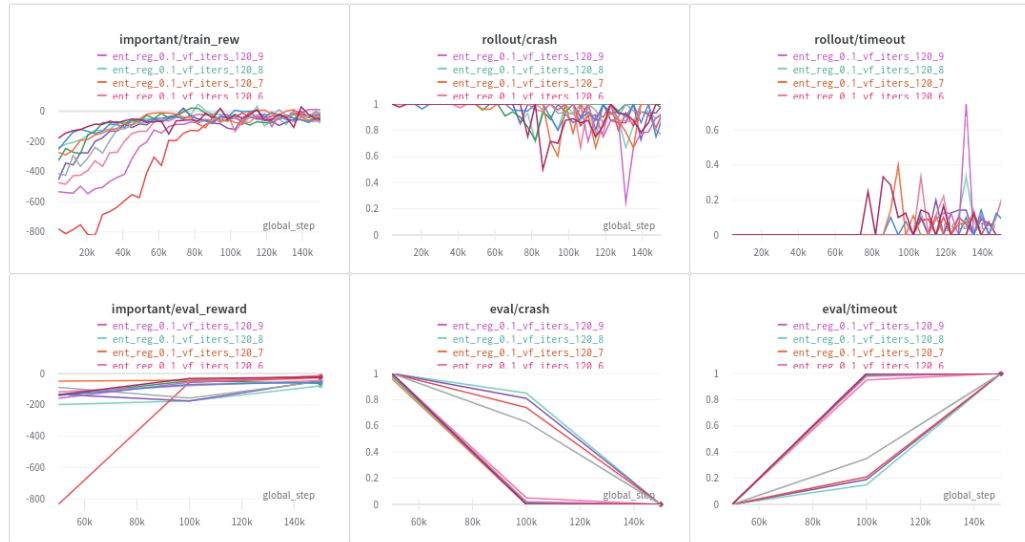


Figure 3: CPO on Lunar Lander Safe.