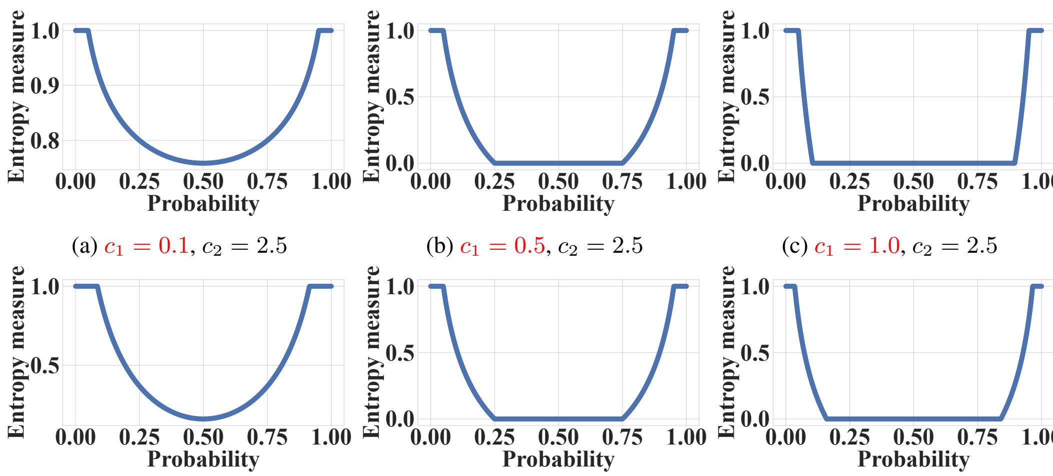
Table A1: Generalization evaluation of NeuOpt (10 runs) on real-world TSPLIB instances.

AM -mix	POMO -mix	AMDKD (POMO)	DACT (sol.10k) Avg.	DACT (sol.10k) Best	Ours (sol.10k) Avg.	Ours (sol.10k) Best	DACT (sol.40k) Avg.	DACT (sol.40k) Best	Ours (sol.25k) Avg.	Ours (sol.25k) Best
19.59%	0.92%	1.18%	3.29%	1.59%	**0.85%**	**0.50%**	2.13%	1.09%	**0.58%**	**0.35%**

Table A2: Generalization evaluation of NeuOpt-GIRE (10 runs) on real-world CVRPLIB instances.

AM -mix	POMO -mix	AMDKD (POMO)	DACT (sol.10k) Avg.	DACT (sol.10k) Best	Ours (sol.10k) Avg.	Ours (sol.10k) Best	DACT (sol.60k) Avg.	DACT (sol.60k) Best	Ours (sol.60k) Avg.	Ours (sol.60k) Best
15.87%	8.05%	5.77%	5.21%	3.68%	**4.80%**	**3.27%**	3.74%	2.85%	**3.51%**	**2.62%**



(a) $c_1 = 0.1$, $c_2 = 2.5$ (b) $c_1 = 0.5$, $c_2 = 2.5$ (c) $c_1 = 1.0$, $c_2 = 2.5$

(d) $c_1 = 0.5$, $c_2 = 1.5$ (e) $c_1 = 0.5$, $c_2 = 2.5$ (f) $c_1 = 0.5$, $c_2 = 3.5$

Figure A1: Effects of c_1 and c_2 on $\mathbb{H}[P]$ pattern: (a)-(c) fix $c_2 = 2.5$ and vary c_1; (d)-(f) fix $c_1 = 0.5$ and vary c_2. Compared to the pattern (b) and (e) used in the original paper, varying c_1 and c_2 either **constricts** the penalty, shown in (c) and (f), or **expands** the penalty, shown in (a) and (d).

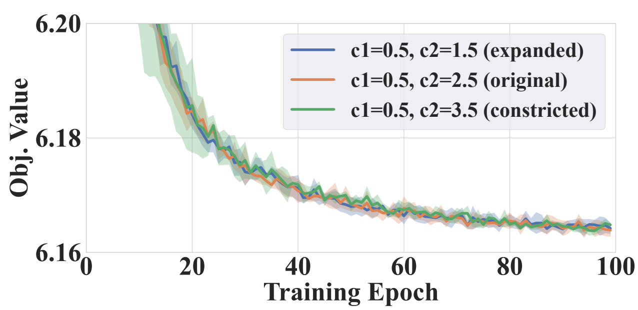


Figure A2: No much influence of **constricted** and **expanded** patterns of $\mathbb{H}[P]$ on the training stability.

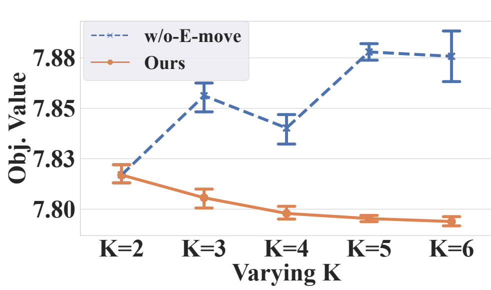


Figure A3: Influence of K and E-move.

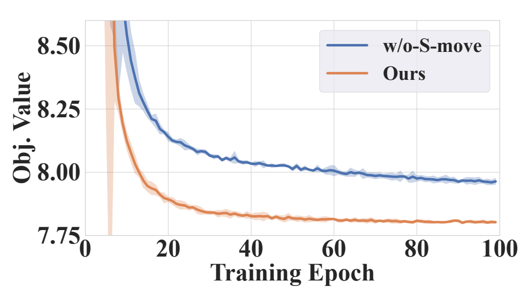


Figure A4: Influence of S-move ($K = 4$).