ICML
International Conference
On Machine Learning

Paper Project Homepage

AdamO: A Collapse-Suppressed Optimizer for Offline RL

Nan Qiao Sheng Yue Shuning Wang Ju Ren

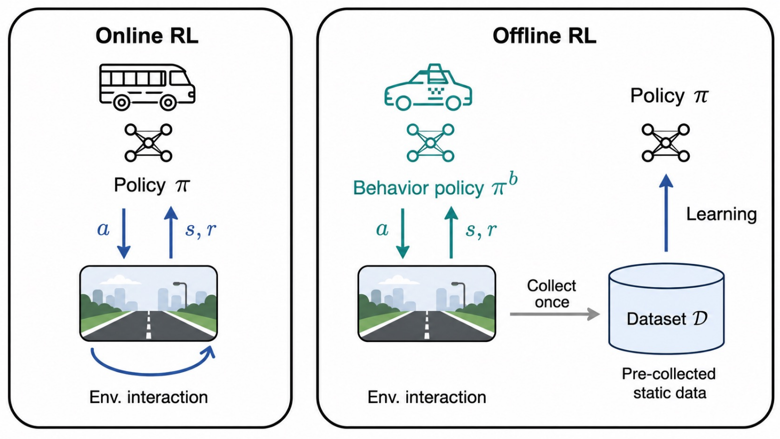
Background: Reinforcement Learning

With ever-growing intelligent decision-making applications…

Conducting such online interaction may not be feasible in many settings due to the high costs or risks involved

Background: Offline Reinforcement Learning

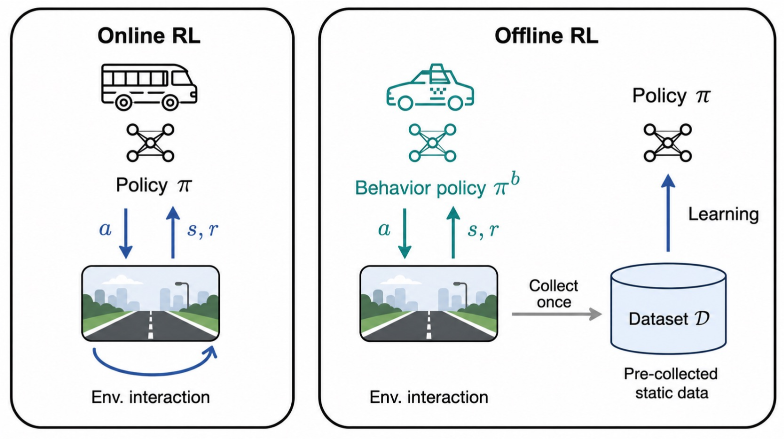
With ever-growing intelligent decision-making applications...

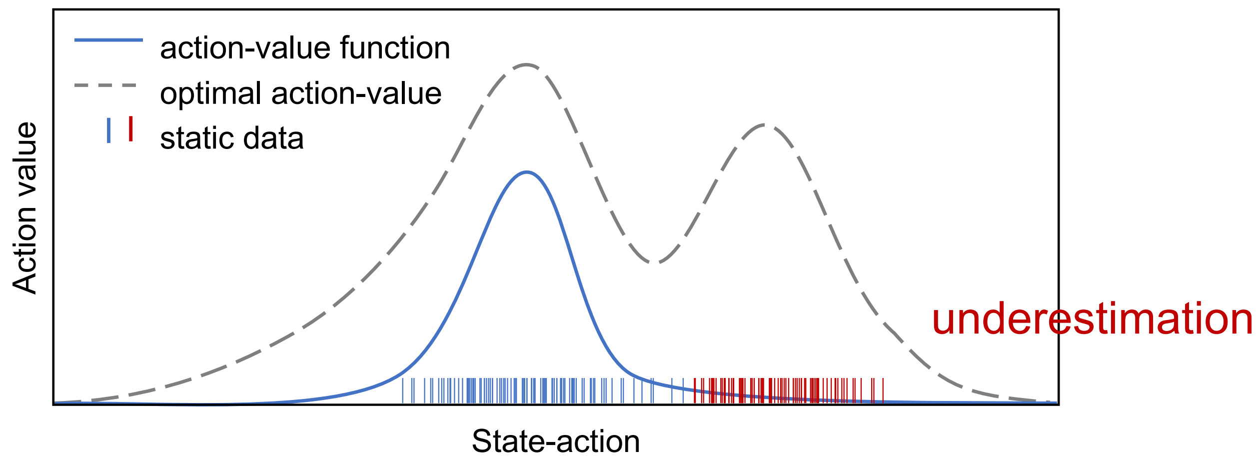


A major challenge in offline reinforcement learning is out-of-distribution shift, which is typically addressed through policy constraints or value underestimation.

Background: Offline Reinforcement Learning

With ever-growing intelligent decision-making applications…

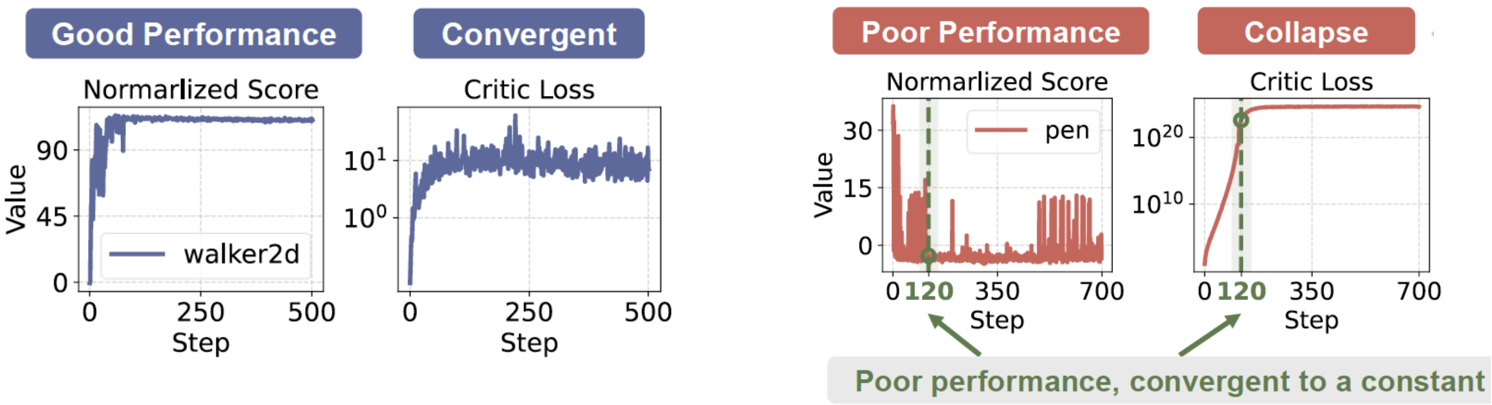




A major challenge in offline reinforcement learning is out-of-distribution shift, which is typically addressed through policy constraints or value underestimation.

Observation on offline RL

Nevertheless, we observe that training collapse frequently occurs in offline RL, primarily manifesting as unstable or diverging loss, which ultimately leads to poor performance.



Conditions for Adam-Induced Value Function Collapse

Under a first-order linearization assumption, Adam-driven TD error dynamics reduce to a second-order linear difference equation. System convergence is then governed by the Hurwitz condition on matrix S, or equivalently, by ensuring that every mode's growth factor r satisfies |r| < 1, with r defined as the per-step multiplicative scaling of the error.

A parameter of Adam

Theorem 1. *Under Assumptions 1–3, for $t \geq t_0$ the TD error satisfies*

TD error

$e_t = Q_{\theta_t}(X) - \left(r + \gamma Q_{\theta_t}(X_t^*)\right).$

$$\mathbf{e}_{t+1} = \left((1+\beta_1)I + \eta(1-\beta_1)\mathsf{S}\right)\mathbf{e}_t - \beta_1 \mathbf{e}_{t-1} + o(\eta) \quad (20)$$

learning rate

where S *is defined in Eq. (17). In the local first-order regime of Assumption 1, the linearized dynamics are divergent to zero if* S *is Hurwitz. Equivalently, in the Hurwitz case every modal growth factor of Eq. (20) satisfies* $|r| < 1$.

Here, the matrix S is a function of the parameter gradient, the discount factor, and Adam's diagonal preconditioning matrix.

Lemma 4. *Under Assumption 3, define for any finite input sets X_1 and X_2 the preconditioned Gram matrix*

$$\mathsf{K}(X_1, X_2) = Z(X_1)^\top D\, Z(X_2). \quad (16)$$

Define

$$\mathsf{S} = \gamma \mathsf{K}(\bar{X}^*, X) - \mathsf{K}(X, X). \quad (17)$$

Conditions for Adam-Induced Value Function Collapse

Under a first-order linearization assumption, Adam-driven TD error dynamics reduce to a second-order linear difference equation. System convergence is then governed by the Hurwitz condition on matrix S, or equivalently, by ensuring that every mode's growth factor r satisfies |r| < 1, with r defined as the per-step multiplicative scaling of the error.

A parameter of Adam

Theorem 1. *Under Assumptions 1–3, for $t \geq t_0$ the TD error satisfies*

TD error

$$\mathbf{e}_t = Q_{\theta_t}(X) - \Big(r + \gamma Q_{\theta_t}(X_t^*)\Big).$$

$$\mathbf{e}_{t+1} = \Big((1+\beta_1)I + \eta(1-\beta_1)\mathsf{S}\Big)\mathbf{e}_t - \beta_1 \mathbf{e}_{t-1} + o(\eta) \tag{20}$$

learning rate

where S *is defined in Eq. (17). In the local first-order regime of Assumption 1, the linearized dynamics are divergent to zero if* S *is Hurwitz. Equivalently, in the Hurwitz case every modal growth factor of Eq. (20) satisfies* $|r| < 1$.

Here, the matrix S is a function of the parameter gradient, the discount factor, and Adam' s diagonal preconditioning matrix.

Lemma 4. *Under Assumption 3, define for any finite input sets X_1 and X_2 the preconditioned Gram matrix*

$$\mathsf{K}(X_1, X_2) = Z(X_1)^\top D\, Z(X_2). \tag{16}$$

Define

TD loss to converge to zero if and only if the matrix S is Hurwitz.

Conditions for Adam-Induced Value Function Collapse

Theorem 1. *Under Assumptions 1–3, for $t \geq t_0$ the TD error satisfies*

$$\mathbf{e}_{t+1} = \Big((1+\beta_1)I + \eta(1-\beta_1)\mathsf{S}\Big)\mathbf{e}_t - \beta_1\mathbf{e}_{t-1} + o(\eta), \quad (20)$$

where S is defined in Eq. (17). In the local first-order regime of Assumption 1, the linearized dynamics are divergent to zero if S is Hurwitz. Equivalently, in the Hurwitz case every modal growth factor of Eq. (20) satisfies $|r| < 1$.

✓ A Hurwitz matrix is characterized by eigenvalues with **strictly negative real parts**, corresponding physically to a contractive system with **negative feedback that damps perturbations over time**. In contrast, a non-Hurwitz matrix exhibits **positive-feedback** behavior, **expanding errors exponentially along certain directions**.

✓ When the state-transition matrix S is Hurwitz, all error growth factors satisfy |r| < 1, ensuring exponential decay of the TD error and **convergence of the entire system**. Otherwise, the presence of any |r| ≥ 1 implies divergence."

Conditions for Adam-Induced Value Function Collapse

Theorem 1. *Under Assumptions 1–3, for $t \geq t_0$ the TD error satisfies*

$$\mathbf{e}_{t+1} = \Big((1+\beta_1)I + \eta(1-\beta_1)\mathsf{S}\Big)\mathbf{e}_t - \beta_1 \mathbf{e}_{t-1} + o(\eta), \quad (20)$$

where S *is defined in Eq. (17). In the local first-order regime of Assumption 1, the linearized dynamics are divergent to zero if* S *is Hurwitz. Equivalently, in the Hurwitz case every modal growth factor of Eq. (20) satisfies* $|r| < 1$.

Hurwitz ⟷ the real part of the largest eigenvalue of the matrix < 0 → TD error **converge to 0**

the maximum real part of the matrix eigenvalues is zero, with the corresponding imaginary part being zero → TD error **converge to a constant**

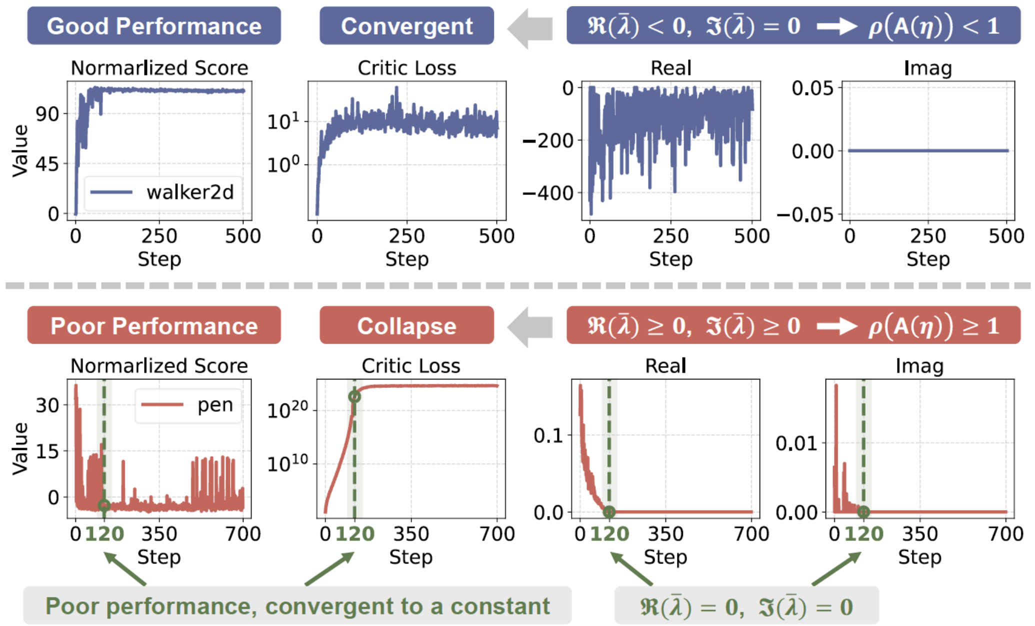
the maximum real part of the matrix eigenvalues is zero, with the corresponding imaginary part being non-zero → TD error **diverge**

the real part of the largest eigenvalue of the matrix >0 → TD error **diverge**

$\mathsf{S} = \gamma \mathsf{K}(\bar{X}^*, X) - \mathsf{K}(X, X)$ → For **supervised learning**, the associated matrix is negative semidefinite, implying that the real parts of all eigenvalues are **naturally less than or equal to zero.**

Conditions for Adam-Induced Value Function Collapse

We validate the above theoretical analysis through experiments, accurately predicting the precise points of convergence and divergence.



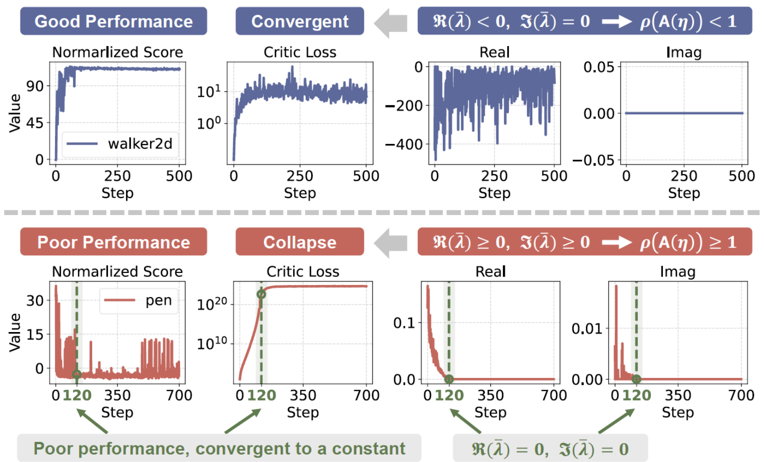
What conditions guarantee that the matrix S is Hurwitz?

What conditions guarantee that the matrix S is Hurwitz?

Key observations and derivations:

•**A sufficient condition** for the matrix **S** to be Hurwitz is that Equation (11) holds.

•**A sufficient condition** for Equation (11) to hold is to enforce approximate parameter orthogonality, together with input normalization, EMA (exponential moving average), and spectral normalization — the latter three being standard tricks widely used in reinforcement learning.



Proposition 5.1. *If*

$$\gamma\|\Phi\|_2\|\Phi_*\|_2 + \|\Phi^\top\Phi - I\|_2 < 1, \qquad (11)$$

then S is Hurwitz.

orthogonality constraint

$$R(W) = \begin{cases} \frac{1}{4}\|WW^\top - I_r\|_F^2, & r < c, \\ \frac{1}{4}\|W^\top W - I_c\|_F^2, & r \geq c, \end{cases}$$

Why impose parameter orthogonality at the optimizer level?

- ✓ The orthogonality constraint is **sufficient** to be handled by the optimizer because it only depends on the current parameters and does not require extra data or loss computation.
- ✓ It is also **necessary** to keep it separate from the loss, because TD learning and orthogonalization serve different purposes. If combined in the loss, Adam may mix their gradients and store the orthogonality effect in its moment estimates, which can interfere with later TD updates.

Algorithm 2 Adam with L_2 regularization and Adam with decoupled weight decay (AdamW)

given $\alpha = 0.001, \beta_1 = 0.9, \beta_2 = 0.999, \epsilon = 10^{-8}, \lambda \in \mathbb{R}$

initialize time step $t \leftarrow 0$, parameter vector $\boldsymbol{\theta}_{t=0} \in \mathbb{R}^n$, first moment vector $\boldsymbol{m}_{t=0} \leftarrow \boldsymbol{0}$, second moment vector $\boldsymbol{v}_{t=0} \leftarrow \boldsymbol{0}$, schedule multiplier $\eta_{t=0} \in \mathbb{R}$

repeat

    $t \leftarrow t + 1$

    $\nabla f_t(\boldsymbol{\theta}_{t-1}) \leftarrow \text{SelectBatch}(\boldsymbol{\theta}_{t-1})$ ▷ select batch and return the corresponding gradient

    $\boldsymbol{g}_t \leftarrow \nabla f_t(\boldsymbol{\theta}_{t-1}) + \lambda\boldsymbol{\theta}_{t-1}$

    $\boldsymbol{m}_t \leftarrow \beta_1 \boldsymbol{m}_{t-1} + (1-\beta_1)\boldsymbol{g}_t$ ▷ here and below all operations are element-wise

    $\boldsymbol{v}_t \leftarrow \beta_2 \boldsymbol{v}_{t-1} + (1-\beta_2)\boldsymbol{g}_t^2$

    $\hat{\boldsymbol{m}}_t \leftarrow \boldsymbol{m}_t/(1-\beta_1^t)$ ▷ β_1 is taken to the power of t

    $\hat{\boldsymbol{v}}_t \leftarrow \boldsymbol{v}_t/(1-\beta_2^t)$ ▷ β_2 is taken to the power of t

    $\eta_t \leftarrow \text{SetScheduleMultiplier}(t)$ ▷ can be fixed, decay, or also be used for warm restarts

    $\boldsymbol{\theta}_t \leftarrow \boldsymbol{\theta}_{t-1} - \eta_t \left(\alpha \hat{\boldsymbol{m}}_t / (\sqrt{\hat{\boldsymbol{v}}_t} + \epsilon) + \lambda\boldsymbol{\theta}_{t-1} \right)$

until *stopping criterion is met*

return optimized parameters $\boldsymbol{\theta}_t$

Solution：AdamO

We introduce an orthogonal projection mechanism to avoid interfering with normal gradient updates.

Algorithm 1 AdamO: Adam with Orthogonality Correction

Init $\omega_0, m_0 = v_0 = 0, \eta, \kappa, \tau$ and Adam params
for $t = 0, 1, \ldots$ **do**
    $g_t \leftarrow \nabla L_t(\omega_t)$
    Compute standard Adam update u_t
    Compute δ_t by applying (17)–(19) layer-wise
    $\omega_{t+1} \leftarrow \omega_t - \eta\,(u_t + \delta_t)$
end for

This is applied layer-wise to ensure scale invariance. To prevent the correction from hijacking task progress, we enforce a local budget constraint

$$\langle g_t, \delta_t \rangle_F \geq -\tau\big(\langle g_t, u_t \rangle_F\big)_+, \qquad \tau \in [0, 1). \tag{18}$$

Among all scalings of $\delta_{t,0}$ along its ray, AdamO takes the largest feasible one, yielding the closed form:

$$\delta_t = \begin{cases} \delta_{t,0}, & \text{if } \langle g_t, \delta_{t,0} \rangle_F \geq T_t, \\ \frac{T_t}{\langle g_t, \delta_{t,0} \rangle_F}\, \delta_{t,0}, & \text{otherwise}, \end{cases} \tag{19}$$

Theoretical guarantees

Does introducing the orthogonality term destabilize the Adam optimizer or interfere with normal task gradient updates?

Theoretical guarantees

Does introducing the orthogonality term destabilize the Adam optimizer or interfere with normal task gradient updates?

Theorem D.8. *Consider the perturbed dynamics Eq. (96) and the Adam Hamiltonian Eq. (94). Along any trajectory,*

$$\frac{d}{dt}H_{\text{Adam}}(w_t, m_t, v_t) = \frac{d}{dt}H^{(\text{Adam})}_{\text{Adam}}(w_t, m_t, v_t) - \langle g(w_t), \delta_t \rangle_F \tag{98}$$

$$\leq -\Delta_{\text{Adam}}(w_t, m_t, v_t) + \tau\big(\langle g(w_t), u_t \rangle_F\big)_+, \tag{99}$$

where Δ_{Adam} is the nonnegative dissipation term in Eq. (95). In particular:

Theorem D.5. *Assume L_t is μ-smooth. (a) In conflict-free mode ($\tau = 0$), under an explicit stepsize condition, AdamO does not increase the next-step task loss relative to Adam. (b) In budgeted mode ($\tau > 0$), we give an explicit upper bound quantifying the worst-case next-step degradation.*

- ✓ AdamO adds the orthogonality correction separately from Adam's normal gradient update, so it does not contaminate Adam's momentum or variance estimates.
- ✓ Theorems D.5 and D.8 show that this correction is controlled: it preserves task progress when conflict-free, and any possible interference is explicitly bounded.

Experiment setup

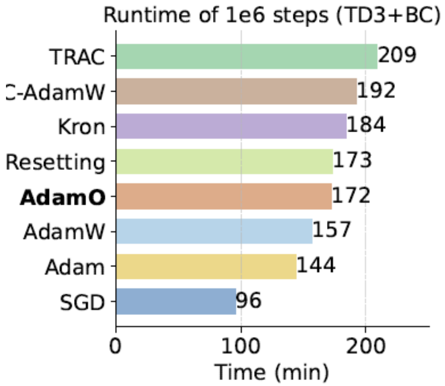
- We evaluate our algorithms using data from a standard offline RL benchmark, D4RL, based on the MuJoCo simulator

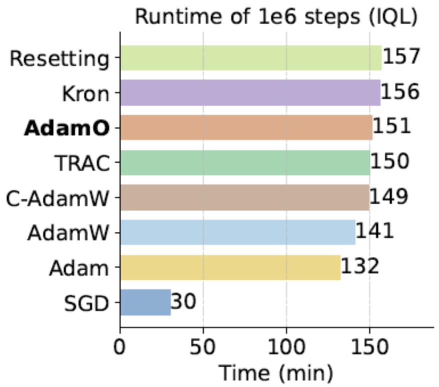
- **Baselines:** Our evaluation covers a broad range of offline RL methods, including policy-constraint and regularization approaches such as TD3+BC and ReBRAC, in-sample learning methods such as IQL and ACTIVE, and distribution-shift-robust algorithms such as PARS and SQOG. We also compare with general-purpose optimizers, including SGD, Adam, and AdamW, as well as stability-oriented methods such as periodic state resetting, TRAC, Kron, and cautious momentum variants like C-AdamW.

- **Implementation**: We implement the code using PyTorch 2.1.2 and run experiments on Ubuntu 20.04.4 LTS with 4 NVIDIA GeForce RTX 3090 GPUs

Experimental results

AdamO outperforms the baselines on all tasks

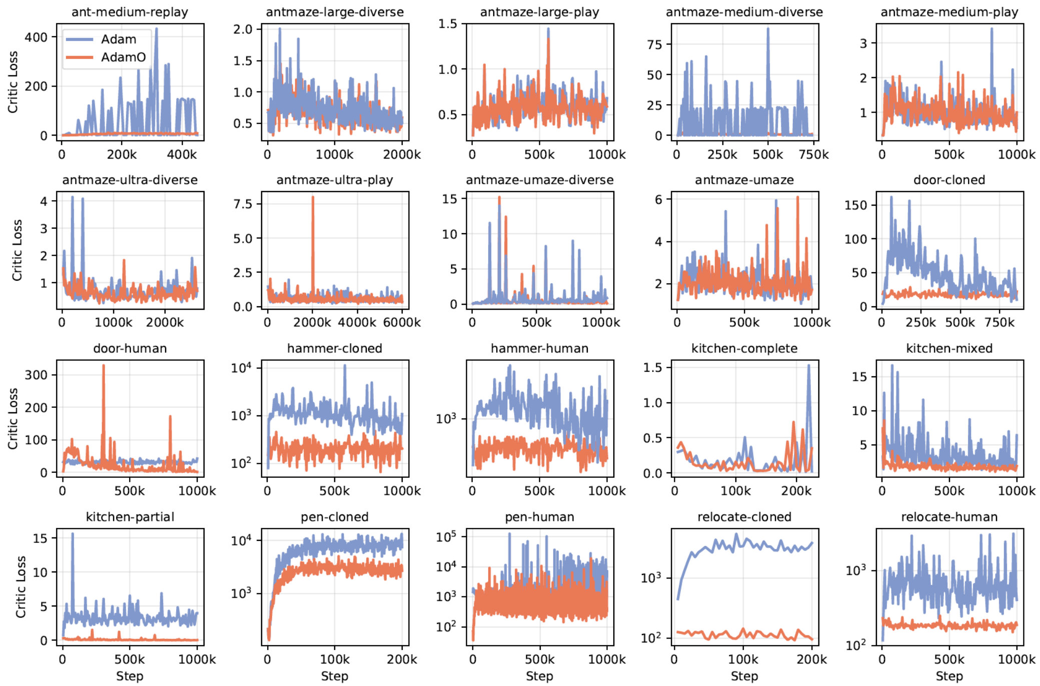




Task Name	TD3+BC		IQL		ReBRAC		ACTIVE		PARS		SQOG	
	Adam	AdamO	Adam	AdamO	Adam	AdamO	Adam	AdamO	Adam	AdamO	Adam	AdamO
AntMaze-umaze	72.2	**92.5**$_{\uparrow 28\%}$	80.5	**83.8**$_{\uparrow 4\%}$	94.3	**96.5**$_{\uparrow 2\%}$	92.4	**95.1**$_{\uparrow 3\%}$	**97.3**	93.5$_{\downarrow 4\%}$	89.6	**93.1**$_{\uparrow 4\%}$
AntMaze-umaze-div	47.0	**82.2**$_{\uparrow 75\%}$	55.8	**68.2**$_{\uparrow 22\%}$	87.2	**91.5**$_{\uparrow 5\%}$	75.9	**78.2**$_{\uparrow 3\%}$	**93.2**	92.4$_{\downarrow 1\%}$	72.8	**87.1**$_{\uparrow 20\%}$
AntMaze-med-play	0.3	**28.5**$_{\uparrow \infty}$	**70.4**	70.1$_{\downarrow 0.4\%}$	85.2	**89.4**$_{\uparrow 5\%}$	73.7	**86.5**$_{\uparrow 17\%}$	**91.5**	92.8$_{\uparrow 1\%}$	60.9	**65.4**$_{\uparrow 7\%}$
AntMaze-med-div	0.2	**16.4**$_{\uparrow \infty}$	66.9	**75.5**$_{\uparrow 13\%}$	79.8	**85.1**$_{\uparrow 7\%}$	77.8	**82.4**$_{\uparrow 6\%}$	**87.1**	90.6$_{\uparrow 4\%}$	65.8	**81.2**$_{\uparrow 23\%}$
AntMaze-large-play	0.0	**13.7**$_{\uparrow \infty}$	38.5	**41.8**$_{\uparrow 9\%}$	55.4	**61.8**$_{\uparrow 12\%}$	50.9	**56.5**$_{\uparrow 11\%}$	46.6	**52.9**$_{\uparrow 14\%}$	52.6	**57.4**$_{\uparrow 9\%}$
AntMaze-large-div	0.0	**16.5**$_{\uparrow \infty}$	32.4	**36.1**$_{\uparrow 11\%}$	50.5	**56.2**$_{\uparrow 11\%}$	46.6	**61.8**$_{\uparrow 33\%}$	50.8	**56.5**$_{\uparrow 11\%}$	47.5	**52.8**$_{\uparrow 11\%}$
AntMaze-ultra-div	0.0	**2.4**$_{\uparrow \infty}$	19.0	**31.1**$_{\uparrow 64\%}$	5.7	**9.4**$_{\uparrow 65\%}$	10.0	**19.8**$_{\uparrow 98\%}$	42.1	**48.6**$_{\uparrow 15\%}$	0.0	**4.2**$_{\uparrow \infty}$
AntMaze-ultra-play	0.0	**1.8**$_{\uparrow \infty}$	21.0	**29.9**$_{\uparrow 42\%}$	20.6	**26.8**$_{\uparrow 30\%}$	13.2	**11.5**$_{\downarrow 13\%}$	55.9	**62.4**$_{\uparrow 12\%}$	0.0	**2.1**$_{\uparrow \infty}$
AntMaze Avg.	15.0	**31.8**$_{\uparrow 112\%}$	48.1	**54.6**$_{\uparrow 14\%}$	59.8	**64.6**$_{\uparrow 8\%}$	55.1	**61.4**$_{\uparrow 11\%}$	**70.6**	73.7$_{\uparrow 4\%}$	48.7	**55.4**$_{\uparrow 14\%}$
HalfCheetah-m	35.9	**53.5**$_{\uparrow 49\%}$	29.7	**35.2**$_{\uparrow 19\%}$	42.8	**49.4**$_{\uparrow 15\%}$	42.9	**48.5**$_{\uparrow 13\%}$	44.9	**51.2**$_{\uparrow 14\%}$	36.5	**41.8**$_{\uparrow 15\%}$
HalfCheetah-mr	39.1	**46.2**$_{\uparrow 18\%}$	32.7	**38.4**$_{\uparrow 17\%}$	40.7	**47.1**$_{\uparrow 16\%}$	34.9	**40.6**$_{\uparrow 16\%}$	45.4	**50.8**$_{\uparrow 12\%}$	30.2	**35.5**$_{\uparrow 18\%}$
HalfCheetah-me	33.5	**60.1**$_{\uparrow 79\%}$	48.1	**54.6**$_{\uparrow 14\%}$	63.9	**71.5**$_{\uparrow 12\%}$	55.7	**62.3**$_{\uparrow 12\%}$	75.7	**82.9**$_{\uparrow 10\%}$	58.9	**65.4**$_{\uparrow 11\%}$
Hopper-m	40.7	**75.5**$_{\uparrow 86\%}$	38.9	**45.1**$_{\uparrow 16\%}$	78.6	**86.2**$_{\uparrow 10\%}$	26.2	**31.8**$_{\uparrow 21\%}$	73.7	**80.4**$_{\uparrow 9\%}$	45.8	**51.3**$_{\uparrow 12\%}$
Hopper-mr	21.3	**55.8**$_{\uparrow 162\%}$	46.6	**52.9**$_{\uparrow 14\%}$	64.2	**70.8**$_{\uparrow 10\%}$	62.5	**68.7**$_{\uparrow 10\%}$	67.5	**74.1**$_{\uparrow 10\%}$	61.4	**67.2**$_{\uparrow 9\%}$
Hopper-me	32.6	**84.2**$_{\uparrow 158\%}$	66.5	**73.2**$_{\uparrow 10\%}$	78.5	**85.9**$_{\uparrow 9\%}$	62.7	**69.1**$_{\uparrow 10\%}$	75.5	**81.6**$_{\uparrow 8\%}$	72.9	**78.5**$_{\uparrow 8\%}$
Walker2d-m	21.2	**68.4**$_{\uparrow 223\%}$	54.9	**61.5**$_{\uparrow 12\%}$	62.2	**69.4**$_{\uparrow 12\%}$	53.6	**59.2**$_{\uparrow 10\%}$	68.0	**74.5**$_{\uparrow 10\%}$	49.8	**55.6**$_{\uparrow 12\%}$
Walker2d-mr	19.3	**52.5**$_{\uparrow 172\%}$	51.4	**57.8**$_{\uparrow 12\%}$	69.4	**76.1**$_{\uparrow 10\%}$	45.8	**51.4**$_{\uparrow 12\%}$	63.4	**69.2**$_{\uparrow 9\%}$	58.5	**64.3**$_{\uparrow 10\%}$
Walker2d-me	22.4	**98.5**$_{\uparrow 340\%}$	57.3	**63.7**$_{\uparrow 11\%}$	74.0	**80.8**$_{\uparrow 9\%}$	58.6	**64.9**$_{\uparrow 11\%}$	81.6	**88.3**$_{\uparrow 8\%}$	68.7	**74.9**$_{\uparrow 9\%}$
Locomotion Avg.	29.6	**66.1**$_{\uparrow 123\%}$	47.3	**53.6**$_{\uparrow 13\%}$	63.8	**70.8**$_{\uparrow 11\%}$	49.2	**55.2**$_{\uparrow 12\%}$	66.2	**72.6**$_{\uparrow 10\%}$	53.6	**59.4**$_{\uparrow 11\%}$
Pen-human	−4.1	**83.1**$_{\uparrow \infty}$	79.1	**89.8**$_{\uparrow 14\%}$	**105.4**	102.5$_{\downarrow 3\%}$	106.2	**109.8**$_{\uparrow 3\%}$	88.1	**94.5**$_{\uparrow 7\%}$	**77.0**	75.9$_{\downarrow 1\%}$
Pen-cloned	5.6	**82.4**$_{\uparrow 1371\%}$	46.5	**82.2**$_{\uparrow 77\%}$	**98.5**	92.4$_{\downarrow 6\%}$	96.5	**100.2**$_{\uparrow 4\%}$	**107.1**	105.8$_{\downarrow 1\%}$	73.6	**87.4**$_{\uparrow 19\%}$
Door-human	−0.3	**0.2**$_{\uparrow \infty}$	3.5	**4.8**$_{\uparrow 37\%}$	−0.1	**0.1**$_{\uparrow \infty}$	0.0	**0.2**$_{\uparrow \infty}$	0.1	**0.3**$_{\uparrow 200\%}$	−0.1	**0.1**$_{\uparrow \infty}$
Door-cloned	−0.3	**0.1**$_{\uparrow \infty}$	3.3	**4.5**$_{\uparrow 36\%}$	0.1	**0.5**$_{\uparrow 400\%}$	0.1	**0.4**$_{\uparrow 300\%}$	0.1	**0.3**$_{\uparrow 200\%}$	0.1	**0.5**$_{\uparrow 400\%}$
Hammer-human	**1.1**	0.5$_{\downarrow 55\%}$	1.9	**3.1**$_{\uparrow 63\%}$	0.3	**0.8**$_{\uparrow 167\%}$	0.3	**0.9**$_{\uparrow 200\%}$	0.3	**0.8**$_{\uparrow 167\%}$	0.3	**0.8**$_{\uparrow 167\%}$
Hammer-cloned	0.3	**0.4**$_{\uparrow 33\%}$	1.7	**3.2**$_{\uparrow 88\%}$	5.4	**7.5**$_{\uparrow 39\%}$	5.9	**7.8**$_{\uparrow 32\%}$	4.6	**7.0**$_{\uparrow 52\%}$	5.1	**6.8**$_{\uparrow 33\%}$
Relocate-human	−0.3	**0.2**$_{\uparrow \infty}$	0.1	**0.5**$_{\uparrow 400\%}$	0.2	**0.6**$_{\uparrow 200\%}$	0.2	**0.7**$_{\uparrow 250\%}$	0.1	**0.7**$_{\uparrow 600\%}$	0.2	**0.6**$_{\uparrow 200\%}$
Relocate-cloned	**0.1**	0.0$_{\downarrow 100\%}$	0.0	0.0	0.0	**0.4**$_{\uparrow \infty}$	**0.2**	0.1$_{\downarrow 50\%}$	**0.1**	0.0$_{\downarrow 100\%}$	−0.2	**0.6**$_{\uparrow \infty}$
Adroit Avg.	0.3	**20.9**$_{\uparrow 6866\%}$	17.0	**23.5**$_{\uparrow 38\%}$	**26.2**	25.6$_{\downarrow 2\%}$	26.2	**27.5**$_{\uparrow 5\%}$	25.1	**26.2**$_{\uparrow 4\%}$	19.5	**21.6**$_{\uparrow 11\%}$

Experimental results

- AdamO effectively mitigates loss divergence.



Thank You!

Nan Qiao
nanqiao.ai@gmail.com