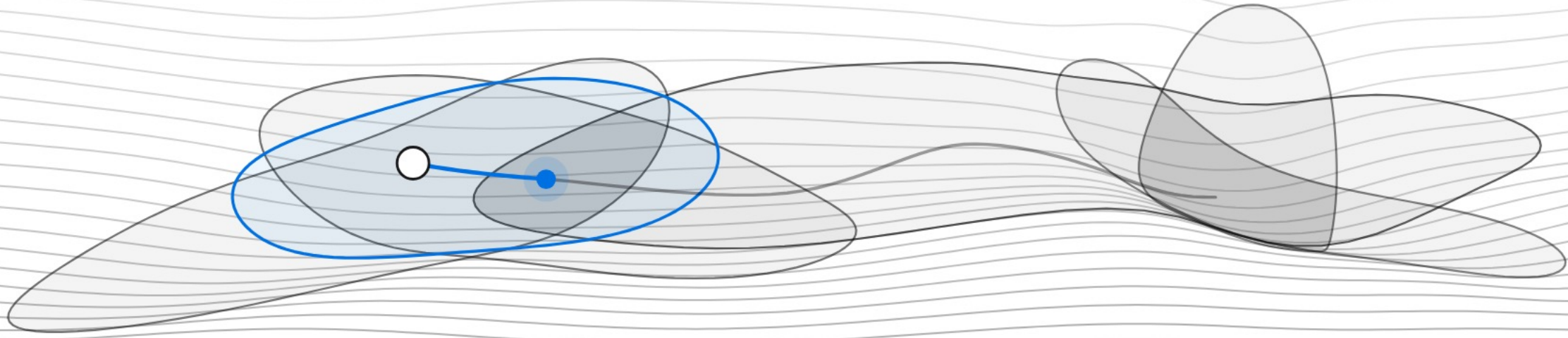
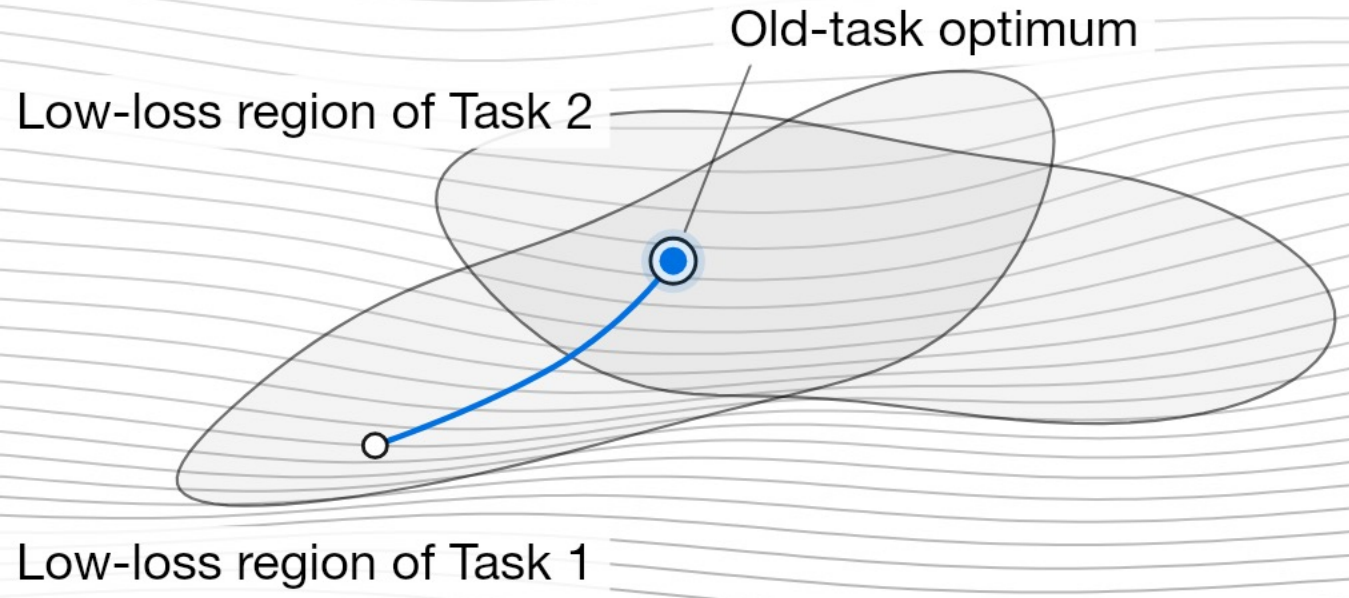
NeurIPS 2026

Trust Region Continual Learning as an Implicit Meta-Learner

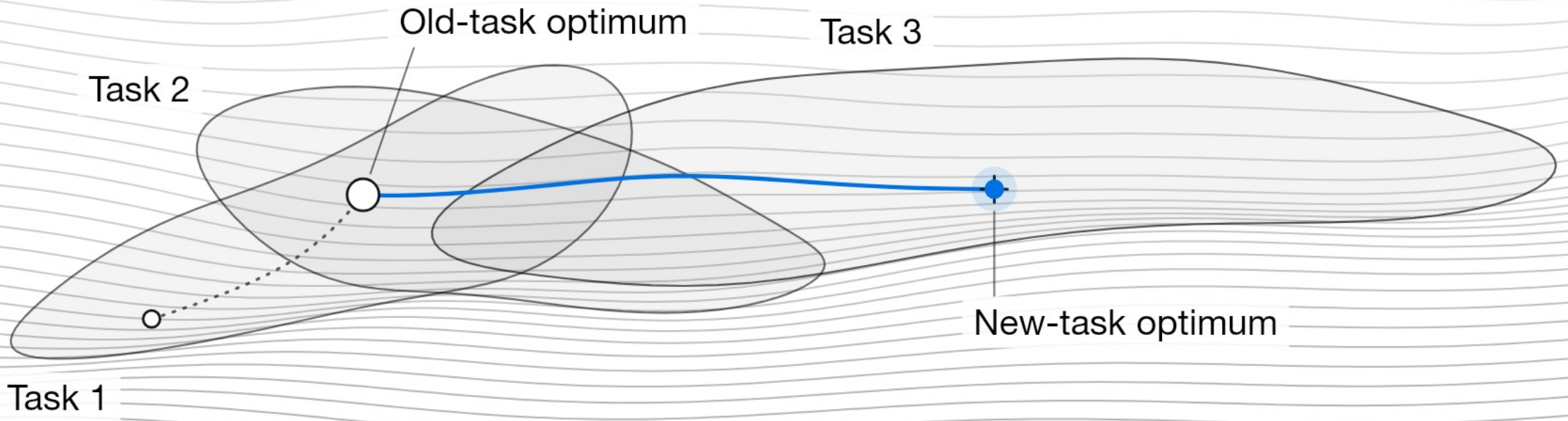
Zekun Wang* Anant Gupta* Christopher J. MacLellan
Teachable AI Lab, Georgia Institute of Technology * Equal contribution



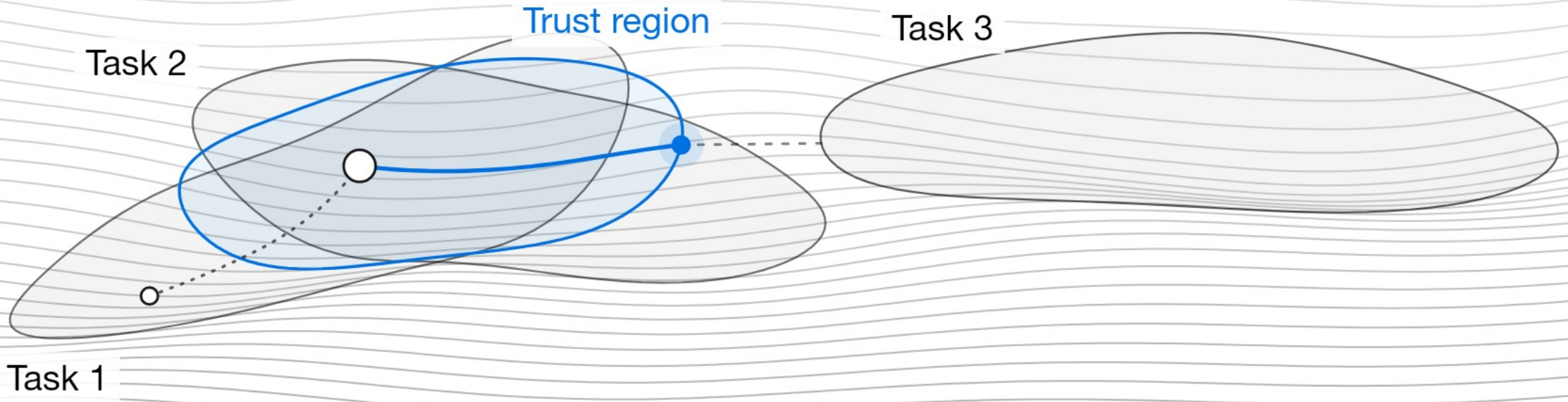
A model learns tasks one after another



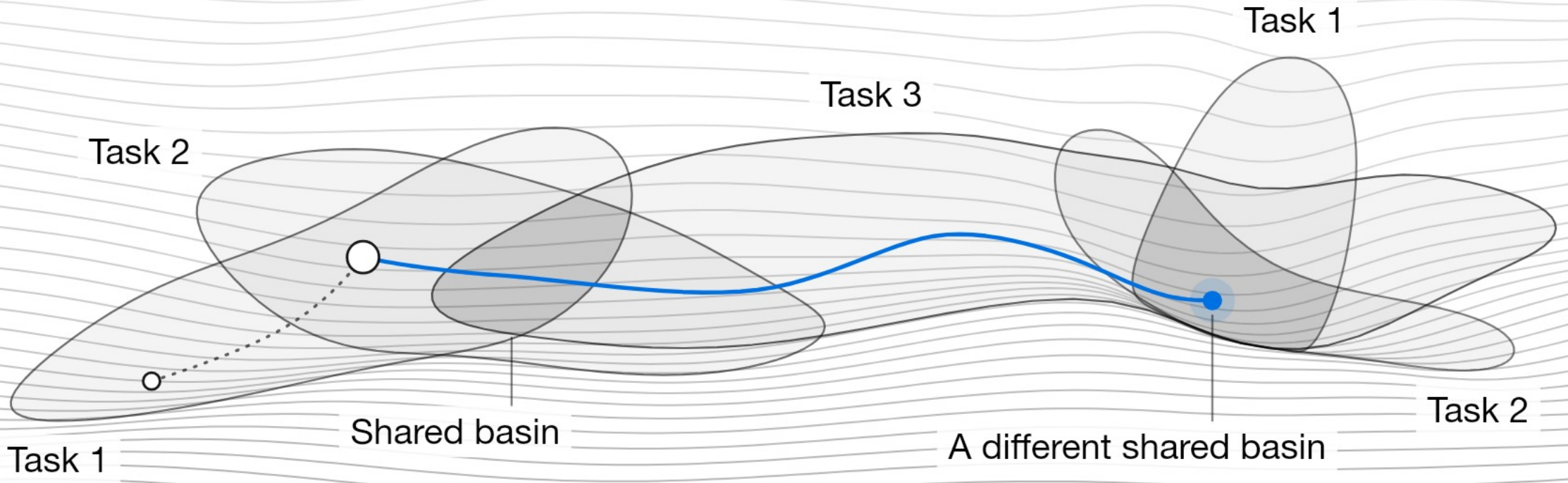
Finetuning on a new task forgets the old tasks



EWC stops at the boundary of the trust region

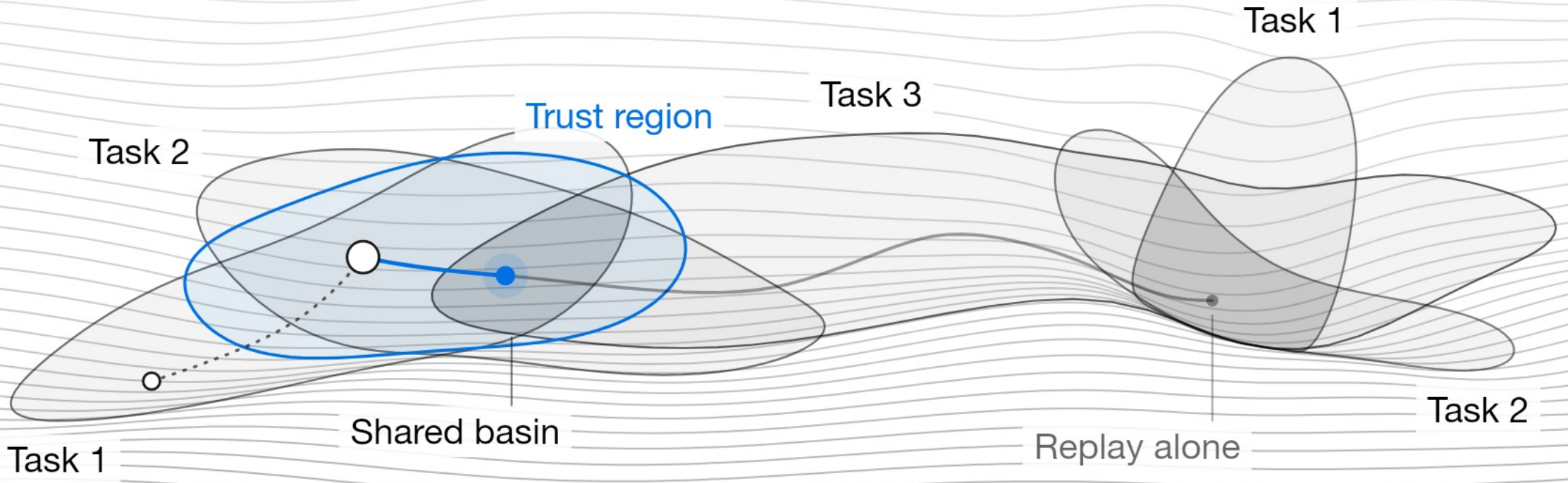


Replay drifts to a different shared basin



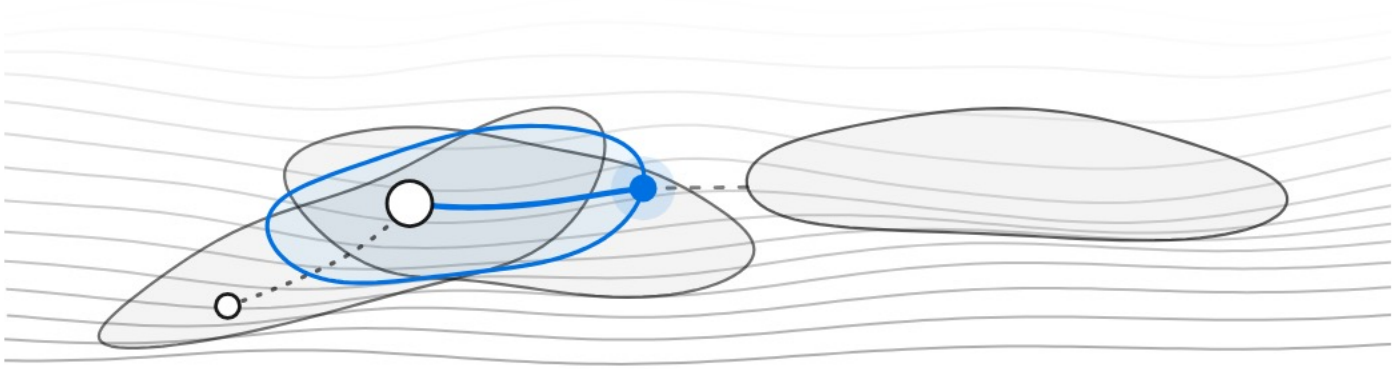
Trust Region = replay + Fisher penalty

$$\min_{\theta} \underbrace{\mathcal{L}_{\mathcal{T}_t}(\theta; \mathcal{D}_t)}_{\text{new-task loss}} + \underbrace{\beta \sum_{i<t} \mathcal{L}_{\mathcal{T}_i}(\theta; \tilde{\mathcal{D}}_i)}_{\text{replay loss}} + \underbrace{\frac{\lambda}{2} \sum_{i<t} (\theta - \theta_i^\star)^\top F_i \, (\theta - \theta_i^\star)}_{\text{Fisher penalty}}$$



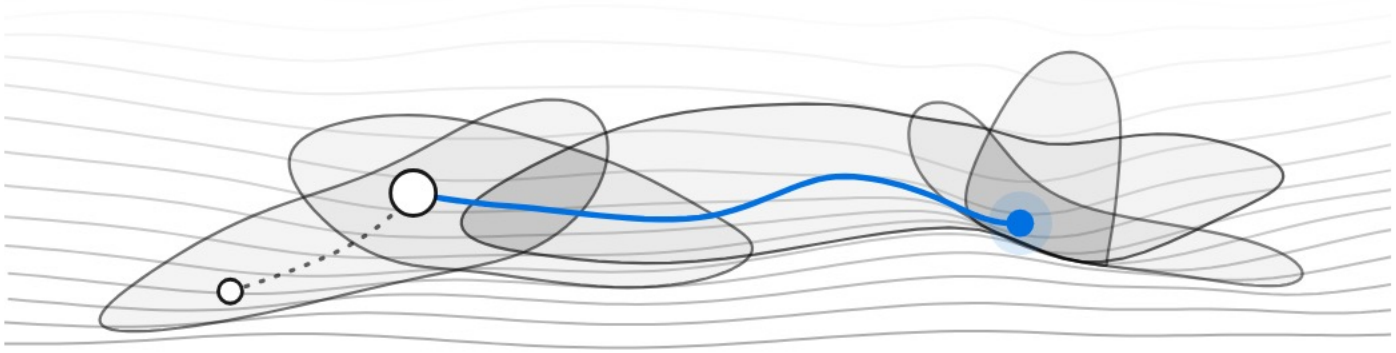
Parameter updates for a new task

EWC



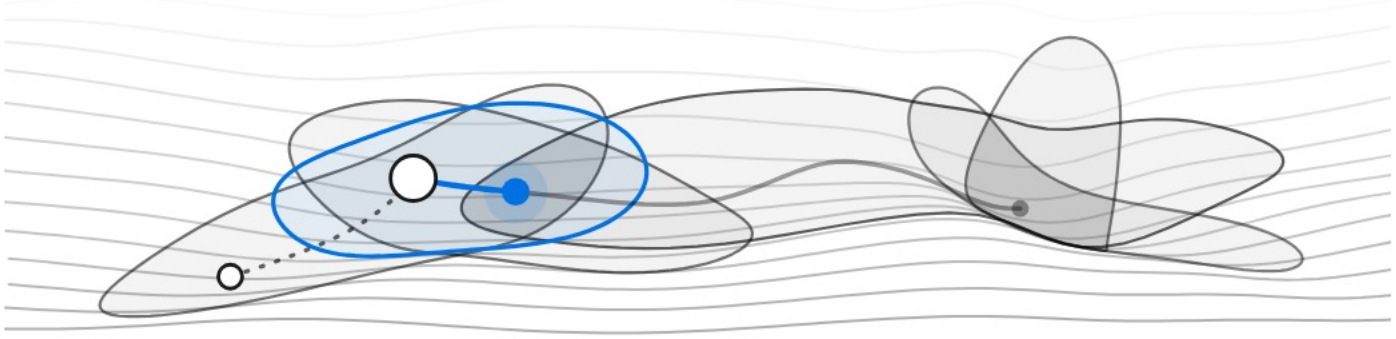
Stops at the boundary of the trust region

Replay



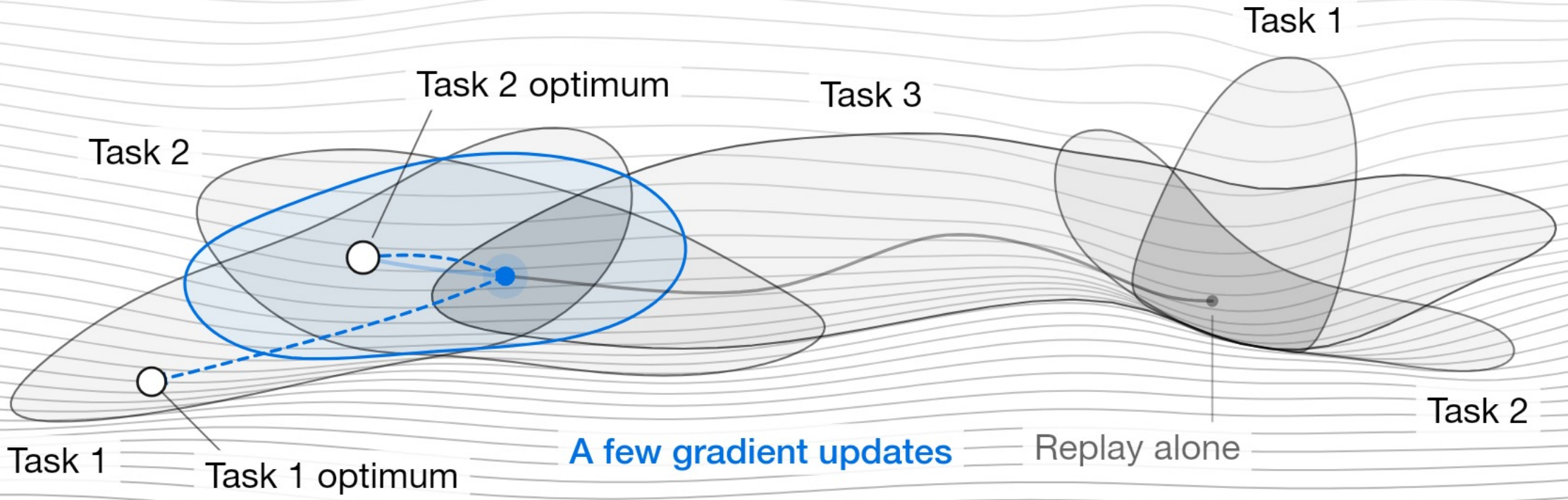
Drifts to a different shared basin

Trust Region



Stays in the shared basin inside the trust region

What does the model learn inside the trust region?



One gradient step on the Trust Region objective

$$\theta \leftarrow \theta - \eta \Big(\underbrace{\nabla_\theta \mathcal{L}_{\mathcal{T}_t}(\theta; \mathcal{D}_t)}_{\textbf{(A)}\ \text{new-task gradient}} + \sum_{i<t} \underbrace{\beta\, \nabla_\theta \mathcal{L}_{\mathcal{T}_i}(\theta; \tilde{\mathcal{D}}_i)}_{\textbf{(B)}\ \text{replay gradient}} + \sum_{i<t} \underbrace{\lambda\, F_i\, (\theta - \theta_i^\star)}_{\textbf{(C)}\ \text{Fisher penalty}} \Big)$$

One step of MAML on the old tasks

$$\theta \leftarrow \theta - \eta \Big(\underbrace{\nabla_\theta \mathcal{L}_{\mathcal{T}_t}(\theta; \mathcal{D}_t)}_{\textbf{(A)}\ \text{new-task gradient}} + \sum_{i<t} \underbrace{\beta\, \nabla_\theta \mathcal{L}_{\mathcal{T}_i}(\theta; \tilde{\mathcal{D}}_i)}_{\textbf{(B)}\ \text{replay gradient}} + \sum_{i<t} \underbrace{\lambda\, F_i\, (\theta - \theta_i^\star)}_{\textbf{(C)}\ \text{Fisher penalty}} \Big)$$

$$\theta \leftarrow \theta - \eta \sum_{i<t} \Big(\underbrace{\nabla_{\theta'} \mathcal{L}_{\mathcal{T}_i}\big(\theta_i'; \mathcal{D}_{\mathcal{T}_i}^{\text{te}}\big)}_{\textbf{(I)}\ \text{query gradient}} - \underbrace{\alpha\, H_{\theta_i}^{\text{tr}}\, \nabla_{\theta'} \mathcal{L}_{\mathcal{T}_i}\big(\theta_i'; \mathcal{D}_{\mathcal{T}_i}^{\text{te}}\big)}_{\textbf{(II)}\ \text{support curvature correction}} \Big)$$

$$\theta_i' = \theta - \alpha\, \nabla_\theta \mathcal{L}_{\mathcal{T}_i}\big(\theta; \mathcal{D}_{\mathcal{T}_i}^{\text{tr}}\big) \qquad H_{\theta_i}^{\text{tr}} = \nabla_\theta^2 \mathcal{L}_{\mathcal{T}_i}\big(\theta; \mathcal{D}_{\mathcal{T}_i}^{\text{tr}}\big)$$

Both updates point along the same direction

$$\theta \leftarrow \theta - \eta \Big(\underbrace{\nabla_\theta \mathcal{L}_{\mathcal{T}_t}(\theta; \mathcal{D}_t)}_{\textbf{(A)}\ \text{new-task gradient}} + \sum_{i<t} \underbrace{\beta\, \nabla_\theta \mathcal{L}_{\mathcal{T}_i}(\theta; \tilde{\mathcal{D}}_i)}_{\textbf{(B)}\ \text{replay gradient}} + \sum_{i<t} \underbrace{\lambda\, F_i\, (\theta - \theta_i^\star)}_{\textbf{(C)}\ \text{Fisher penalty}} \Big)$$

$$\theta \leftarrow \theta - \eta \sum_{i<t} \Big(\underbrace{\nabla_{\theta'} \mathcal{L}_{\mathcal{T}_i}\big(\theta_i'; \mathcal{D}^{\text{te}}_{\mathcal{T}_i}\big)}_{\textbf{(I)}\ \text{query gradient}} - \underbrace{\alpha\, H^{\text{tr}}_{\theta_i}\, \nabla_{\theta'} \mathcal{L}_{\mathcal{T}_i}\big(\theta_i'; \mathcal{D}^{\text{te}}_{\mathcal{T}_i}\big)}_{\textbf{(II)}\ \text{support curvature correction}} \Big)$$

Trust Region	**(B)** + **(C)**	$\approx\ (\beta + \lambda)\ \boldsymbol{F_i\, \delta_i}$
One-step MAML	**(I)** − **(II)**	$\approx\ (1 - \alpha\rho_i)\ \boldsymbol{F_i\, \delta_i}$

$$\delta_i = \theta - \theta_i^\star \qquad F_i \approx \rho_i\, u_i u_i^\top$$

Replay supplies the query gradient. The Fisher penalty approximates the support curvature correction, with no second-order backpropagation.

Task 1 re-converges within 55 gradient updates

	ImageNet-500: FID threshold +30%									**CW10**: success-rate threshold 90%								
Method	T2	T3	T4	T5	T6	T7	T8	T9	T10	T2	T3	T4	T5	T6	T7	T8	T9	T10
Finetuning	2	25	–	–	–	25	50	–	–	–	–	–	–	–	–	–	–	–
EWC	8	30	55	–	–	55	45	–	–	–	–	–	–	–	–	–	–	–
Replay	2	2	5	–	–	–	–	–	–	10	50	2	20	30	2	35	2	8000
FTML	5	55	–	–	–	–	–	–	–	550	20000	–	3000	–	30	–	–	–
VR-MCL	4	15	–	–	–	–	–	–	–	750	650	500	250	–	–	–	–	–
Trust Region (ours)	**20**	**15**	**2**	**35**	**35**	**25**	**55**	**40**	**2**	**10**	**45**	**2**	**30**	**30**	**2**	**25**	**2**	**2**

Gradient updates for Task 1 to re-converge after each task transition. Dash: not reached during that task.

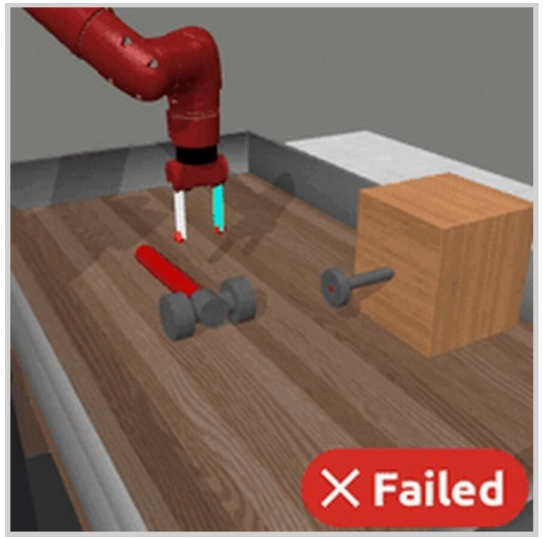
Best final performance, lowest forgetting

	ImageNet-500		CW10	
Method	$\mathcal{A}$FID ↓	$\mathcal{F}$ ↓	SR (%) ↑	$\mathcal{F}$ ↓
Finetuning	$86.2_{\pm 7.3}$	$50.4_{\pm 7.2}$	$17.9_{\pm 2.4}$	$73.0_{\pm 2.0}$
EWC	$77.2_{\pm 1.2}$	$41.2_{\pm 2.3}$	$17.6_{\pm 1.8}$	$73.7_{\pm 1.3}$
Replay	$53.4_{\pm 6.0}$	$18.2_{\pm 4.6}$	$85.3_{\pm 2.0}$	$8.2_{\pm 2.0}$
FTML	$172.5_{\pm 9.1}$	$128.5_{\pm 9.0}$	$78.5_{\pm 3.8}$	$13.3_{\pm 3.3}$
VR-MCL	$142.2_{\pm 8.5}$	$96.6_{\pm 8.4}$	$77.9_{\pm 1.5}$	$11.7_{\pm 2.0}$
Trust Region (ours)	$\mathbf{44.5}_{\pm 2.3}$	$\mathbf{10.6}_{\pm 3.0}$	$\mathbf{88.3}_{\pm 0.4}$	$\mathbf{4.4}_{\pm 0.9}$

Mean ± standard error after 10 tasks. AFID: average FID. SR: success rate. F: forgetting.

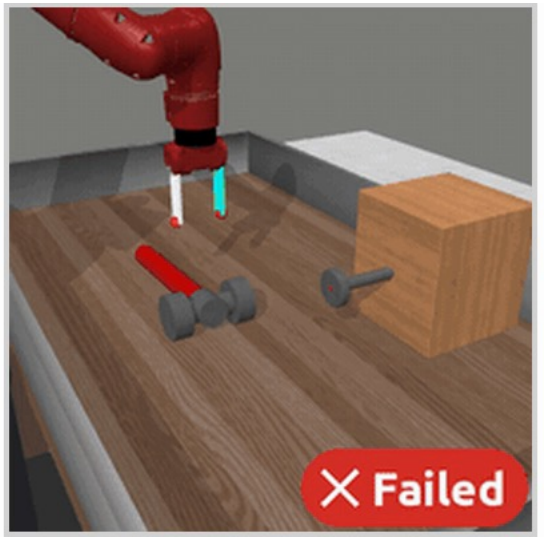
Task 1 (hammer) after training on all 10 tasks

Finetuning



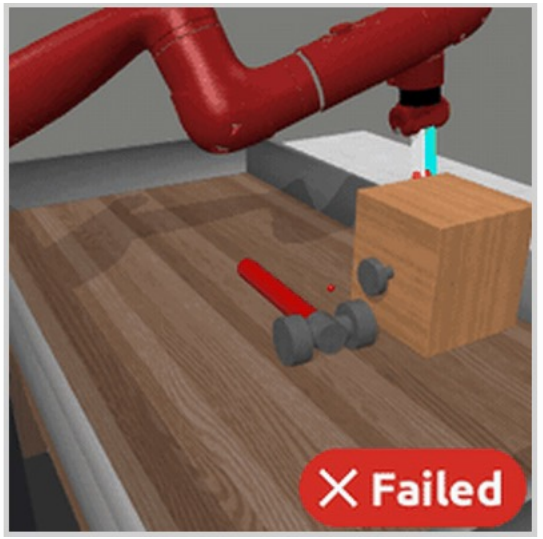
Never lifts the hammer

EWC



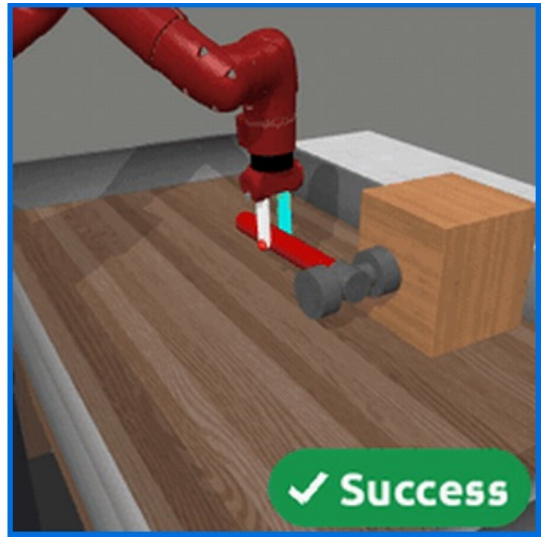
Never lifts the hammer

Replay



Lifts the hammer, then drops it

Trust Region



Drives the nail in

Diffusion Policy on CW10. Right after training on Task 1, all four methods succeed.

Trust Region is an implicit meta-learner

Trust Region continual learning trains on new-task and replay data inside a Fisher-metric trust region around the old-task optima.

Under local approximations, its old-task update takes the form of one-step MAML, without bilevel optimization.

It re-converges on old tasks in fewer gradient updates, with the best final performance and the lowest forgetting on ImageNet-500 and CW10.

Paper

Blog post

Thank you {zekun, agupta886, cmaclell}@gatech.edu