PAC-Net: A Model Pruning Approach to Inductive Transfer Learning

Sanghoon Myung[1], In Huh[1], Wonik Jang[1], Jae Myung Choe[1], Jisu Ryu[1], Dae Sin Kim[1], Kee-Eung Kim[2], and Changwook Jeong[3]

[1]Computational Science and Engineering Team, Innovation Center, Samsung Electronics

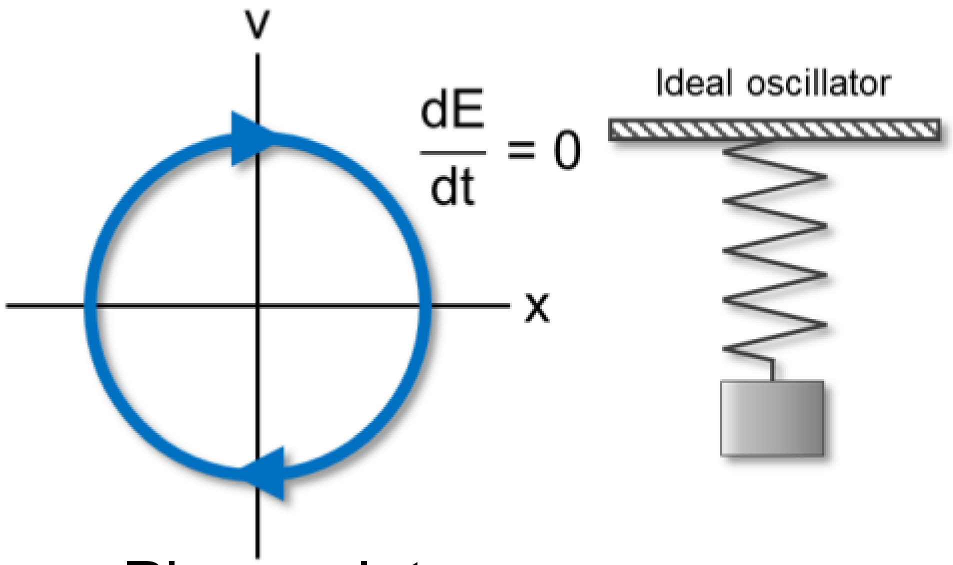
[2]Kim Jaechul Graduate School of AI, KAIST

[3]Graduate School of Semiconductor Materials and Devices Engineering, UNIST

Motivation

- Let's consider a simple problem, a spring-mass-damper system.

Source task

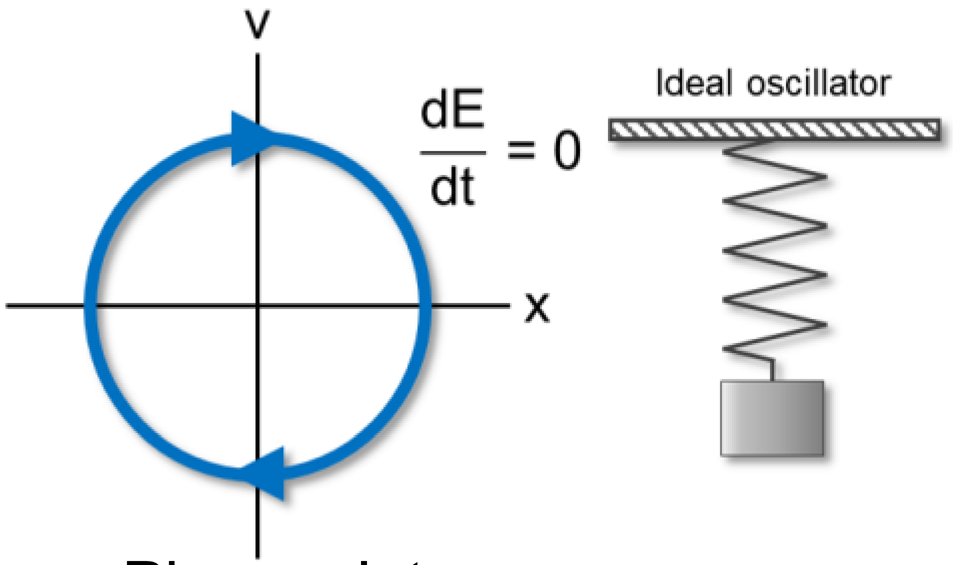


<Phase plot>

Motivation

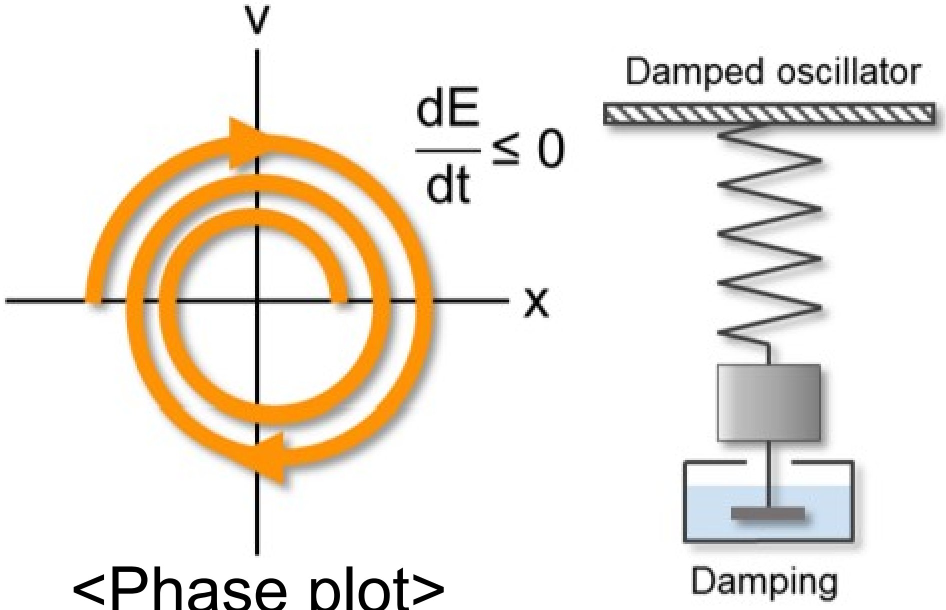
■ Let’s consider a simple problem, a spring-mass-damper system.

Source task



<Phase plot>

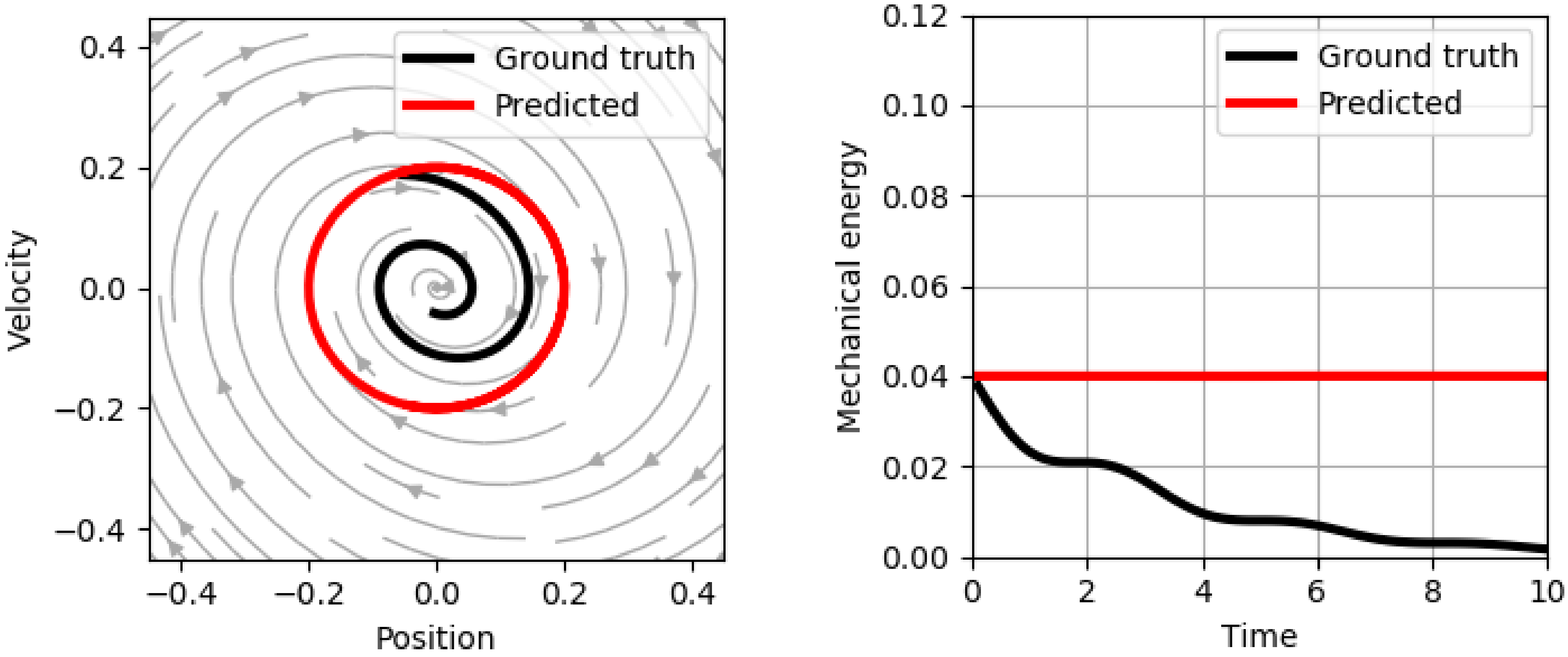
Target task



<Phase plot>

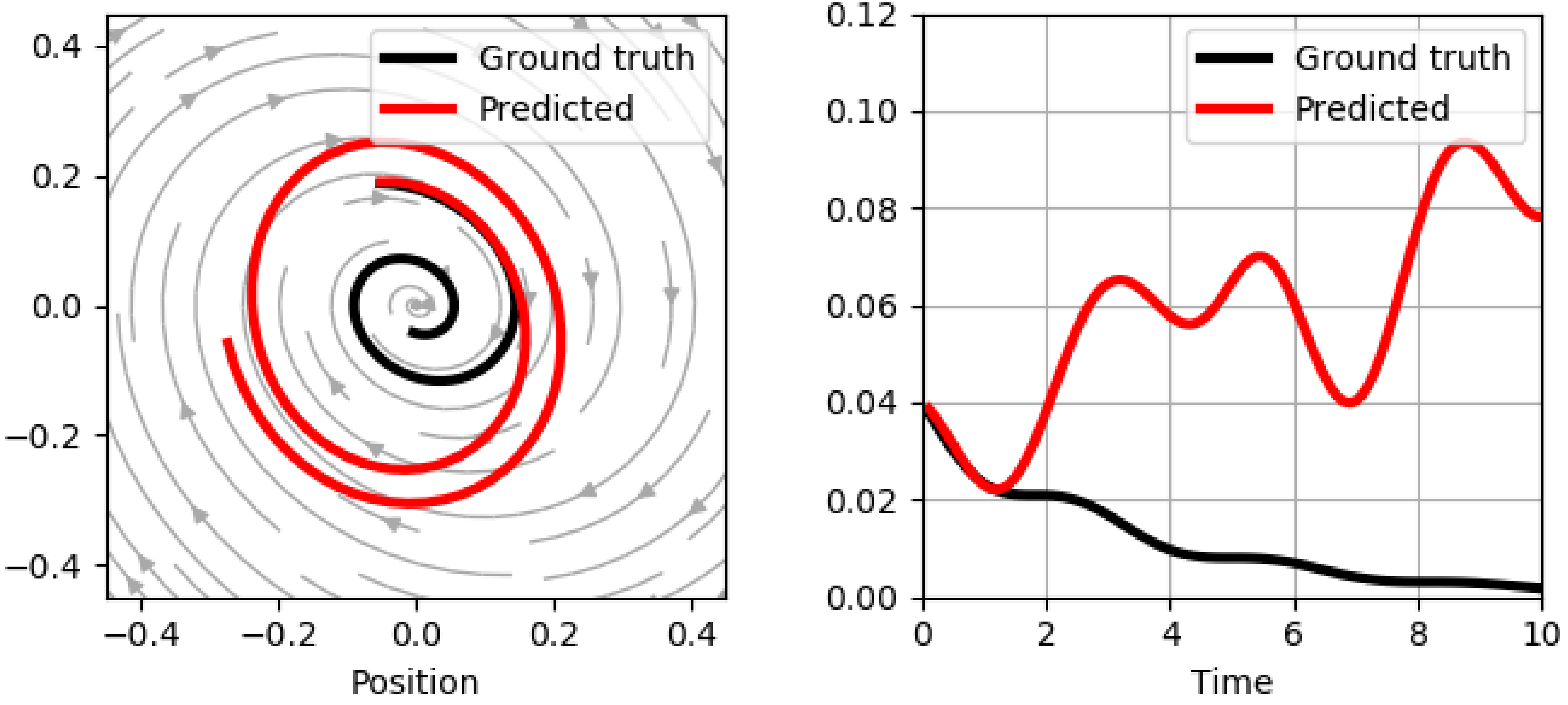
Motivation

- Let’s consider a simple problem, a spring-mass-damper system.



Motivation

- Let's consider a simple problem, a spring-mass-damper system.



Proposal

■ Hypothesis

- It is crucial to totally preserve the source knowledge for inductive transfer learning.

- If the source task is related to the target task, the weights between the source and target model are close.

■ Key idea of PAC-Net

- Deep neural network is often overparameterized to improve its performance.

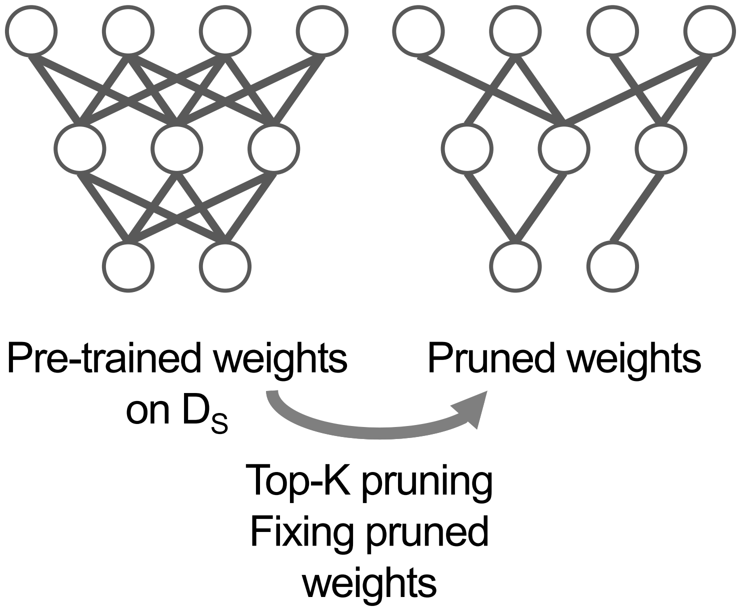
- Pruning technique can be applied to preserve the source knowledge and to reutilize the pruned weights to learn the target task.

Proposal: PAC-Net

■ Step 1: Pruning

- Prunes the weights $\mathbf{w}_S$ of the pre-trained model by applying the following binary mask ($\mathbf{m}$) that keeps the top-K large-magnitude weights:

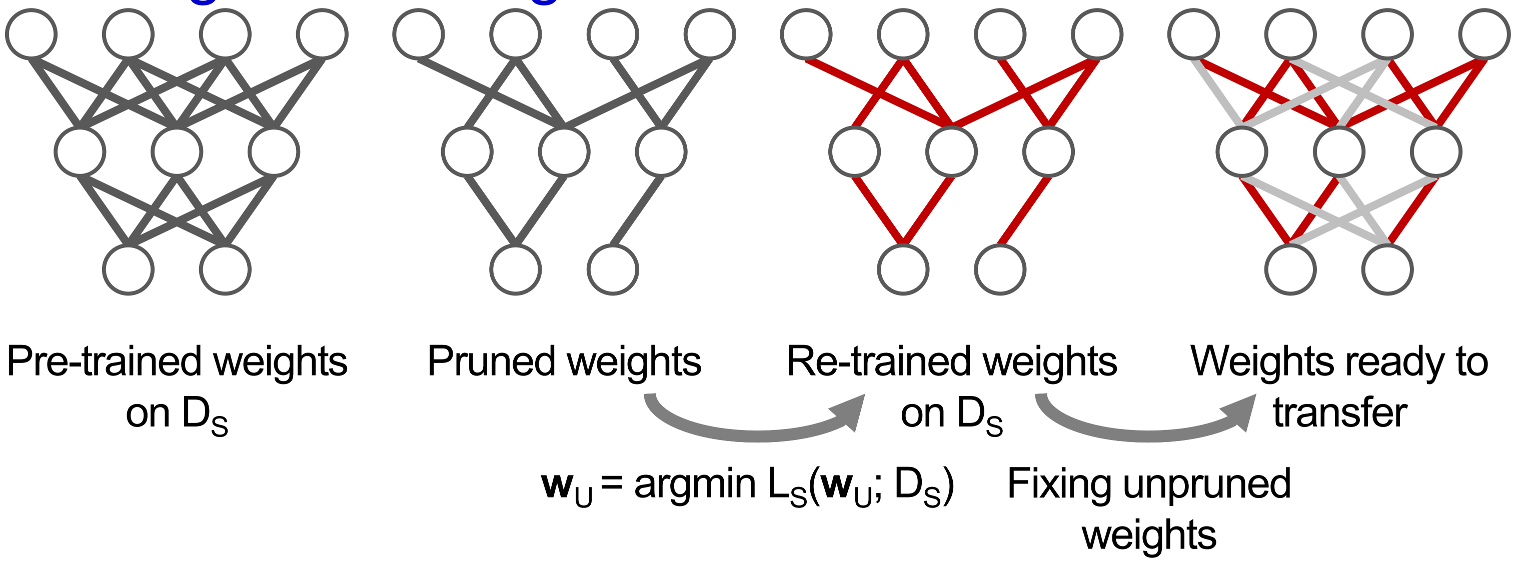
$$m^i = \begin{cases} 1, & \text{if } |w^i| > w_\kappa \\ 0, & \text{otherwise,} \end{cases}$$



Proposal: PAC-Net

■ Step 2: Allocation

- Since all the information on the source task should be embedded in $\mathbf{w}_U = \mathbf{w} \odot \mathbf{m}$, step 2 retrains the masked neural network with the source dataset.
- This procedure allocates $\mathbf{w}_U$ for the source knowledge and $\mathbf{w}_P$ that will be reutilized for the target knowledge.

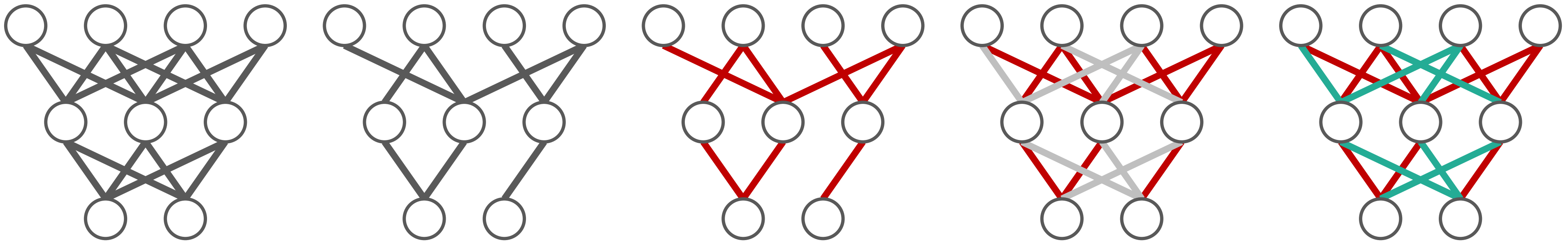


$\mathbf{w}_U$: unpruned weights
$\mathbf{w}_P$: pruned weights
D_S: source dataset
D_T: target dataset
L_S: source loss function
L_T: target loss function
Ω: regularizer

Proposal: PAC-Net

■ Step 3: Calibration

- Our objective is to calibrate $\mathbf{w}_P = \mathbf{w} \odot (1 - \mathbf{m})$ to the target task. In this procedure, only $\mathbf{w}_P$ should be updated with L^2 constraint to completely keep the source knowledge ($\mathbf{w}_U$).



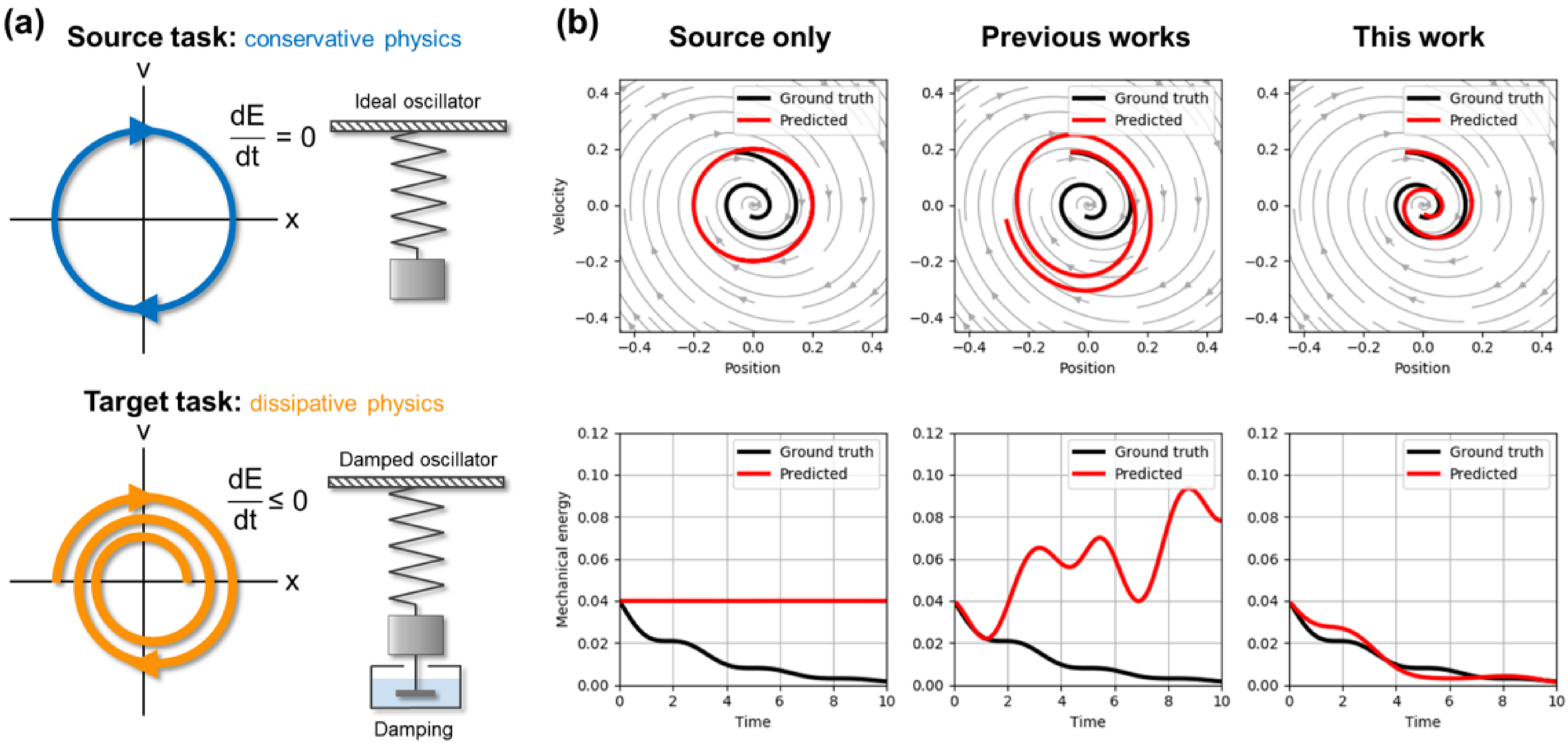
Pre-trained weights on D_S | Pruned weights | Re-trained weights on D_S | Weights ready to transfer | Trained weights on D_T

$\mathbf{w}_P = \text{argmin } L_T(\mathbf{w}_P; D_T) + \Omega(\mathbf{w}_P)$

$\mathbf{w}_U$: unpruned weights
$\mathbf{w}_P$: pruned weights
D_S: source dataset
D_T: target dataset
L_S: source loss function
L_T: target loss function
Ω: regularizer

Results

- PAC-Net can preserve the source knowledge on learning the target task.



Conclusion

■ We proposed a simple yet effective approach for inductive transfer learnin
g based on pruning.

■ Our method through pruning with regularization makes the model mitigate catastrophic forgetting, which achieves state-of-the-art performance.

■ Our method can be applicable in various tasks for classification and regres
sion as well as the impactful tasks such as ODEs and PDEs.