Supplementary material:

Transfer learning for predicting source terms of principal component transport in chemically reactive flow

Ki Sung Jung, Tarek Echekki, Jacqueline H. Chen, Mohammad Khalil

S1. Sensitivity of the TL methods depending on the manifold definition

Figure S1 shows the variations in the normalized root-mean squared error (NRMSE) against the test sets for the target task where $T_0 = 1300$ K and $N_\phi = 4$. Note that we trained the ANN models in the source and target tasks using each of two PCA basis, A^T, defined by either (i) the dataset with $T_0 = 1000$ K and $N_\phi = 30$ (source task) or (ii) $T_0 = 1300$ K and $N_\phi = 30$ (target test). Since the difference in A^T between the two datasets is noticeable, this additional experimental study provides insight into the effect of the PCA basis on the result. It is shown from Fig. S1 that there are noticeable differences in the error depending on the low-dimensional manifolds, especially with the case with TL1. Nevertheless, the minimum value of NMRSE is marginally affected by the choice of A^T across all clusters, and moreover, the overall efficiency of the TL is relatively low as compared to the case shown in Fig. 8 irrespective of the choice of A^T. This result indicates that the primary source of error of TL stems from the chemical source term predictions (not PCA basis), which are largely determined by the number of training samples in the target task and task similarity. Therefore, this result demonstrates that the effect of changing the PCA basis is not a primary factor in determining the optimal value of NRMSE within the current TL framework.

S2. Loss curve for training the source and target task with TL methods

Fig. S2 shows the variations in training loss as a function of the epoch for training source task (i.e., $T_0 = 1000$ K, $N_\phi = 30$) and target task (i.e., $T_0 = 1050$ K, $N_\phi = 4$) with TL3 for

Cl#2. Note that we apply the early stopping callback function based on the validation loss such that the total number of iterations is different for each case. A plateau learning rate scheduler is used to optimize the ANN model. Since the number of training samples in the target task is much smaller than that in the source task, the convergence rate of the ANN model in the target task is noticeably faster than that in the source task. As a result, the wall clock time required to train the source task at T_0 = 1050 K with N_ϕ = 30 is 3,150 s for Cl#2, while the wall clock time for training target task for T_0 = 1050 K with N_ϕ 4 is 328 s

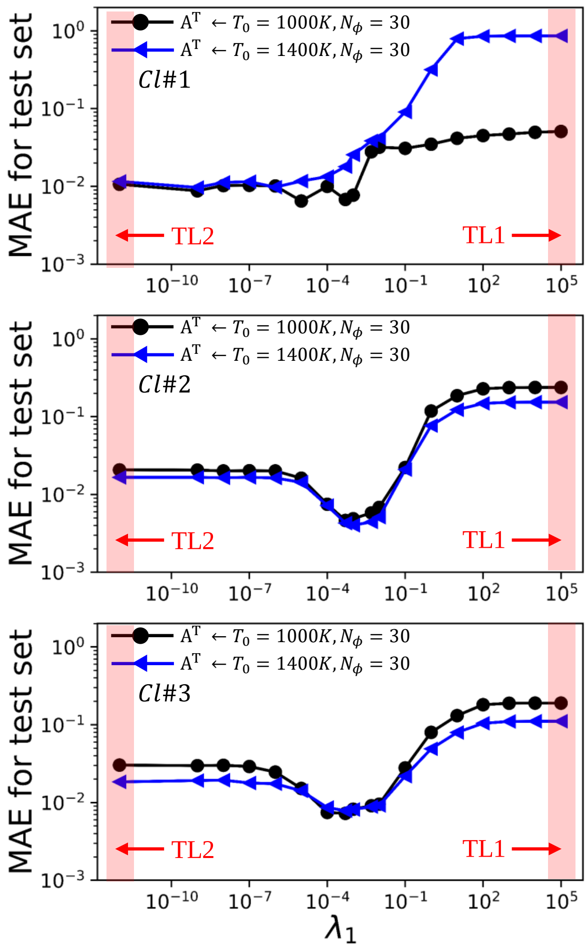

Figure S1: *A priori* evaluation of the NRMSE for the test set for the target task with T_0 = 1400 K and N_ϕ = 4, conditioned on Cluster 1, 2, and 3 (top to bottom), as a function of λ_1. The choice of PCA basis, A^T, is based on either (black) the dataset with T_0 = 1000 K and N_ϕ = 30 (source task) or (blue) T_0 = 1300 K and N_ϕ = 30 (target test).

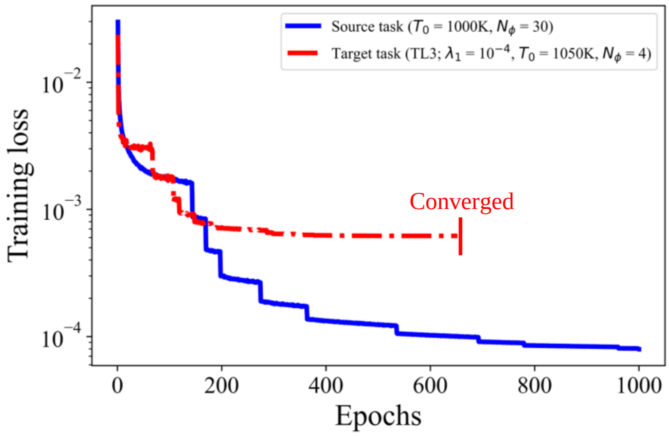

Figure S2: Variations in training loss as a function of epoch for training (blue) source task and (red) target task with $T_0 = 1050$ K, $N_\phi = 4$ for Cl#2. For the target task, TL3 is used with $\lambda_1 = 10^{-4}$. The wall clock time required to train the source task is 3,150 s, while the wall clock time for training target task is 328 s

when using TL3 with $\lambda_1 = 10^{-4}$.