## H. Training with Relaxations

We also explored to idea of incorporating the SPH relaxation during training, hoping that the learned model can be regularized toward predicting better particle distributions, which could make the SPH relaxation during inference unnecessary. We explored two degrees of freedom when training a GNS-10-128 model on the 2D LDC dataset: (a) dependence on the relaxation parameter $\alpha$, and (b) performance when trained with relaxation but evaluated with or without it.

**Basic setup**. We remind the reader that according to Table 2, the optimal relaxation parameters on 2D LDC are $\alpha = 0.03$ and 5 relaxation steps, but from the ablation in Fig. 13, we see that even one relaxation step significantly improves the dynamics. Thus, for simplicity, we use $\alpha = 0.03$ with 1 relaxation step for our training with relaxation. We implemented this training scheme by adding the relaxation to every forward call of the model, i.e. when pushforward is applied, the relaxation is executed at every pushforward step.

**Training with ”negative” relaxation**. One highly appealing idea is to train the model with what we call ”negative” relaxation, i.e. flipping the sign of the relaxation term by setting $\alpha$ to a negative value, by which the model would learn to over-correct unfavorable distributions. However, the results for $\alpha < 0$ in Fig. 29 are rather discouraging.

**Training and inference with relaxation**. Similar to subtracting the external force from the learning target, which we discussed in length and seems very useful, we investigated how the model would perform when it can predict an even worse particle distribution, which is then corrected through a relaxation both during training and inference, see $\alpha > 0$ in Fig. 30. But also here, we get worse results than only applying relaxation during inference. In addition, training with relaxation requires separate retraining until $\alpha$ is tuned, which is not the case with our inference time relaxation.

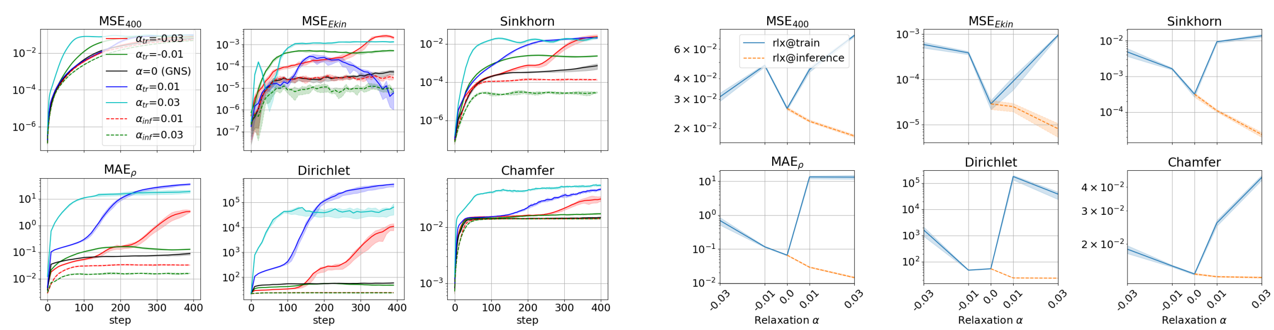


*Figure 29.* GNS-10-128 trained on 2D LDC with relaxation, and but evaluated **without** relaxation. We denote with $\alpha_{\text{tr}}$ that the model has experienced relaxation only during training and with $\alpha_{\text{inf}}$ only during inference. Metrics over the simulation length (left) and the average thereof (right).

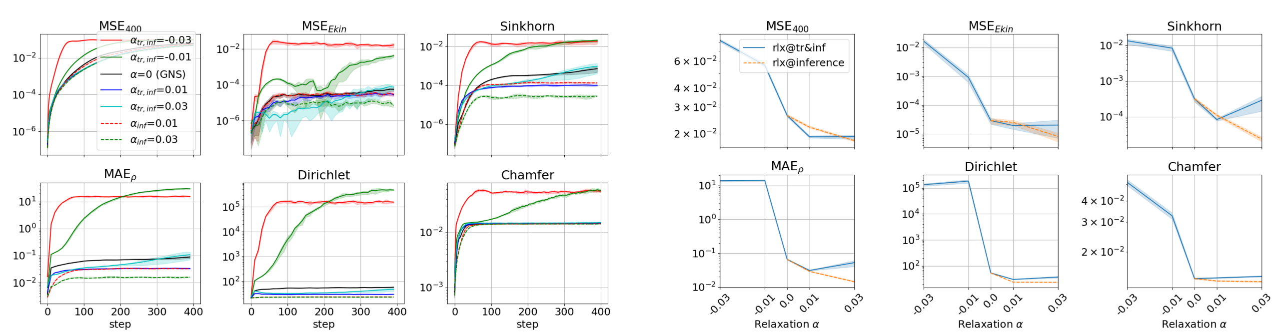


*Figure 30.* GNS-10-128 trained on 2D LDC with relaxation, and also evaluated **with** relaxation. We denote with $\alpha_{\text{tr,inf}}$ that the model has experienced relaxation both during training and inference and with $\alpha_{\text{inf}}$ only during inference. Metrics over the simulation length (left) and the average thereof (right).