Periodic curves dataset, n-steps non-overlapping, total epochs: 43 175 -**---** Source:  $0:t_0$ --- Target:  $t_0 + 1: t_0 + 100$ 150 -Prediction:  $t_0 + 1 : t_0 + 100$ 125 -Attention point:  $t_0 + 1$ 100 75 · 50 -25 0 ?25 Attention pattens for  $t_0 + 1$ 0.175 attn\_layer\_1 attn\_layer\_2 0.150 attn\_layer\_3 0.125 0.100 0.075 -0.050 -0.025 0.000 -

200

100

150

250

300