n-steps non-overlapping, total epochs: 181 1.00 -**---** Source: 0 : *t*₀ --- Target: $t_0 + 1: t_0 + 100$ 0.75 -Prediction: $t_0 + 1 : t_0 + 100$ 0.50 Attention point: $t_0 + 1$ 0.25 0.00 -?0.25 -?0.50 -?0.75 -?1.00 -Attention pattens for $t_0 + 1$ attn_layer_1 0.012 attn_layer_2 attn_layer_3 0.010 0.008 0.006 0.004 0.002 -

200

100

150

250

300