Naive 0.99 0.61 0.52 0.18 0.89 0.78 0.46 0.64 0.67 0.92 0 0.98 0.11 0.47 0.57 0.18 0.73 0.83 0.33 0.86 0.6 \vdash 0.97 0.55 0.77 0.54 0.13 0.66 0.57 0.83 0.14 0.31 \sim 0.67 0.94 0.57 0.95 0.23 0.75 0.86 0.35 0.55 0.76 $_{\odot}$ Training task 0.08 0.21 0.53 0.02 0.97 0.44 0.48 0.88 0.38 0.75 0.03 0.88 0.76 0.22 0.26 0.97 0.86 0.36 0.8 0.54 0.04 0.82 0.78 0.33 0.41 0.79 0.97 0.33 0.51 0.66 9 0.99 0.42 0.87 0.09 0.57 0.04 0.46 0.41 0.64 0.78 / 0.89 0.74 0.35 0.29 0.58 0.45 0.54 0.98 0.22 0.31 ∞ 0.76 0.86 0.01 0.36 0.13 0.56 0.83 0.87 0.46 0.96 0 0 5 1 2 3 4 6 7 8 9 **Evaluation task**

- 0.8

- 0.6

- 0.4

- 0.2