B) H during training A) RNN prediction + Match / Non-match (M/NM) task prediction $\hat{\mathbf{v}}_{\scriptscriptstyle{+}}$ input x_t neural activity r_t 000000 1,00 Cross entropy *H* 0.75 0.50 0.25 M/NM 0 weights V D A_{α} 0 0 B_{α} ŏ $C_{\beta 1}$ 0 ŏ o' М 0.00 ιŏ weights **U** D Ō NM 5 10 15 20 weights T weights W Epoch t+1 C) Distribution of cross-temporal α decoding accuracy D) Histogram of red regime Variance α accuracy Average α accuracy 1.0 0.10 10.0 0.8 0.08 Training time Training time 7.5 Frequency 0.6 0.06 5.0 0.4 0.04 2.5 0.2 0.02 0.0 00 Δ 0.0 0.2 0.4 0.6 0.8 1.0 Average accuracy 0.0 0.00 A A ВВ C C D D A A ВВ C C D D Testing time $t \rightarrow$ Testing time $t \rightarrow$ E) Number of switch neurons F) Number of stable neurons G) Match vs Non-match 1.00 1.00 projecting neurons Prediction only Stable P = 3e-05Fraction of runs Fraction of runs 5 Prediction + Switch 0.75 0.75 M/NM 0.50 0.50 3 P = 0.020.25 Prediction only 0.25 Prediction + M/NM 0.00 0.00 0.0 2.5 5.0 7.5 10.0 10 15 Match Non-match # Switch neurons # Stable neurons