Attention Heads Effect (Noising) gpt2 - Next-Item Sequencing Task 0 - 0.015 1 -2 -0.010 3 -- 0.005 4 · Effect on Logit Difference 5 Head Index 0.000 6 7 --0.005 8 -0.010 9 10 -- -0.015 11 -Ó i 2 3 5 7 9 10 11 4 6 8 Layer