Attention Heads Effect (Denoising) gpt2 - Next-Item Sequencing Task 0 1 -- 0.10 2 -3 -- 0.05 4 · Effect on Logit Difference 5 · Head Index 0.00 6 7 --0.05 8 -9 -0.10 10 -11 -Ö i 2 3 5 7 9 10 11 4 6 8 Layer