MIA threshold Results (Hybrid High-Risk) 0.8 -**Generation Model** 0.7 minGPT minGPT DP ( $\epsilon$ =1) 0.6 -0.5 -0.3 -0.2 -0.1 -0.0 Attack Advantage ΑÚC F1 Score Recall Accuracy Precision Metric