Exchange96Transformer Bottom 30% performance Top 30% performance 10 Number of trials 8 6 2 0 0.51 0.58 0.66 0.73 0.81 0.28 0.36 0.43 active layer ratio