

Figure 10: Relative Distance vs. Relative Distance. This plot shows the relative distance of the 12 heads in the attention mechanism. The x-axis represents the relative distance, ranging from 0 to 1000. The y-axis represents the relative distance, ranging from 0 to 1.0. The plot displays 12 lines, each representing a different head (Head-1 to Head-12). The lines show a general downward trend as the relative distance increases, with Head-1 (blue) and Head-2 (orange) showing the most significant decrease.

