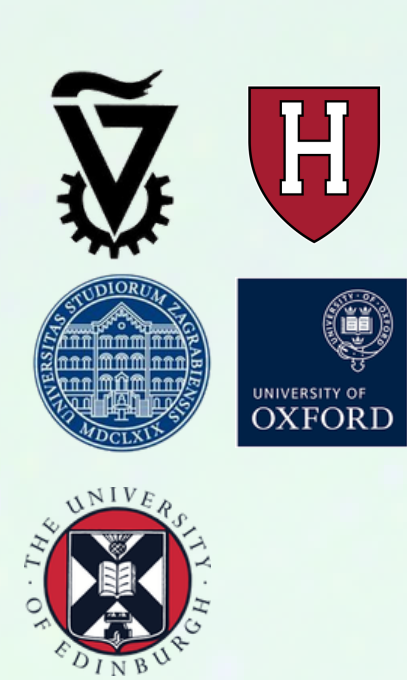
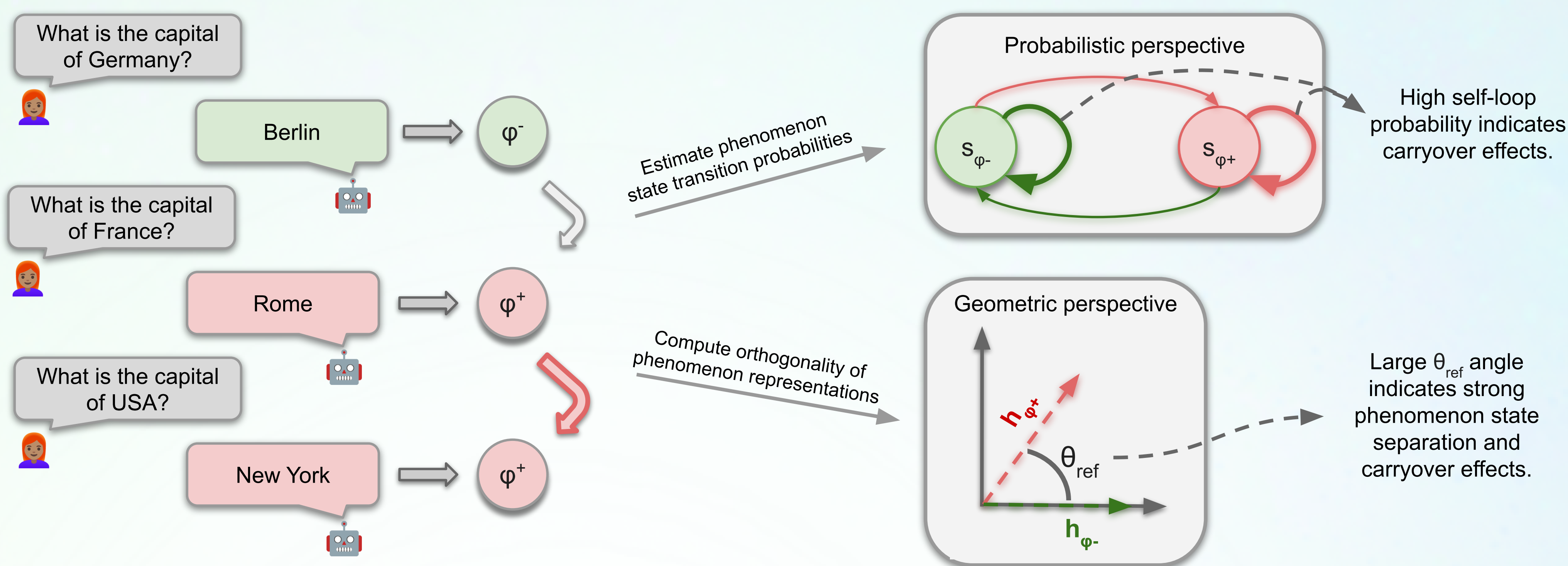




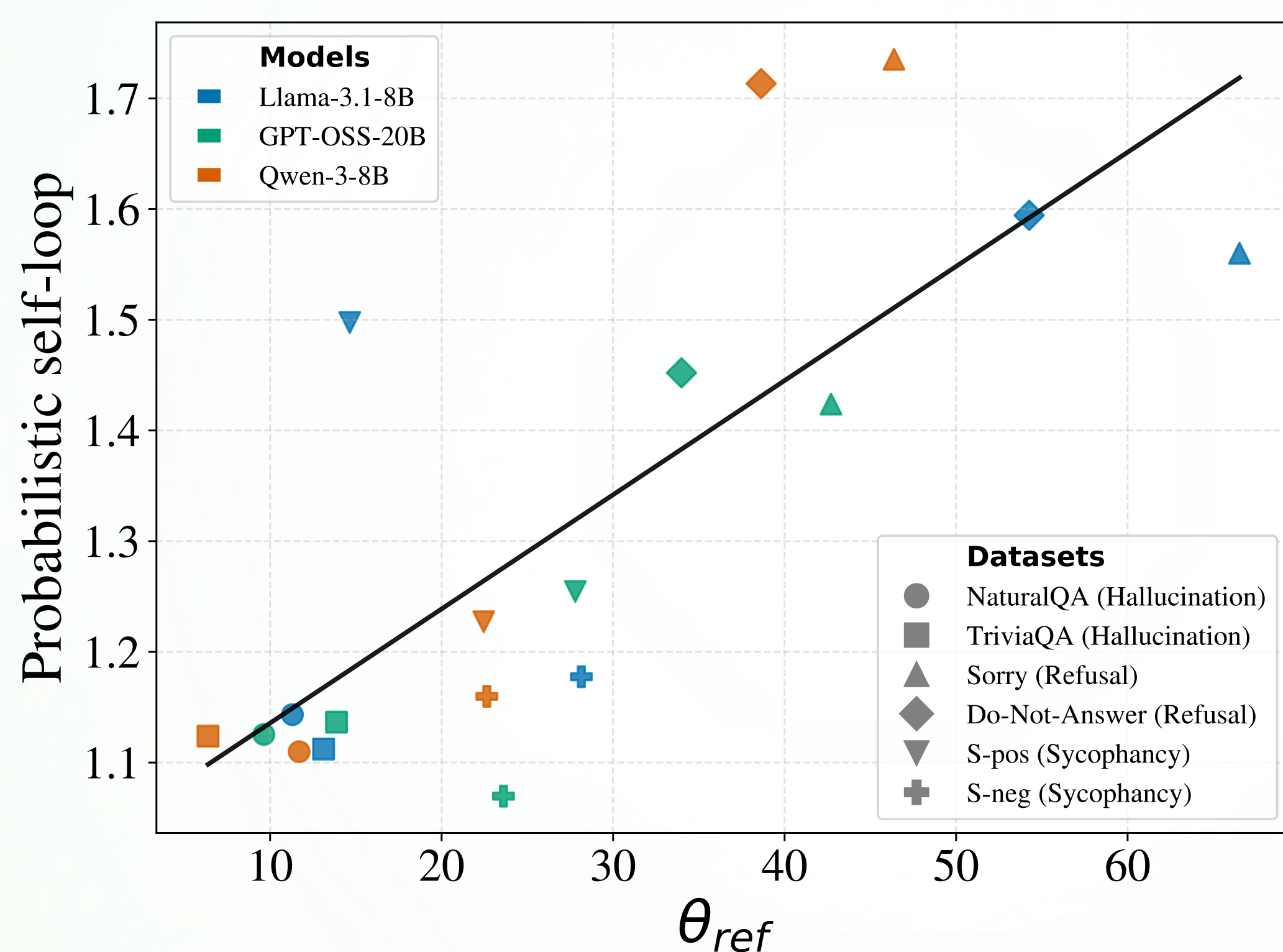
Old Habits Die Hard: How Conversational History Geometrically Traps LLMs

Adi Simhi, Fazl Barez, Martin Tutek, Yonatan Belinkov, Shay B. Cohen |
adi.simhi@campus.technion.ac.il

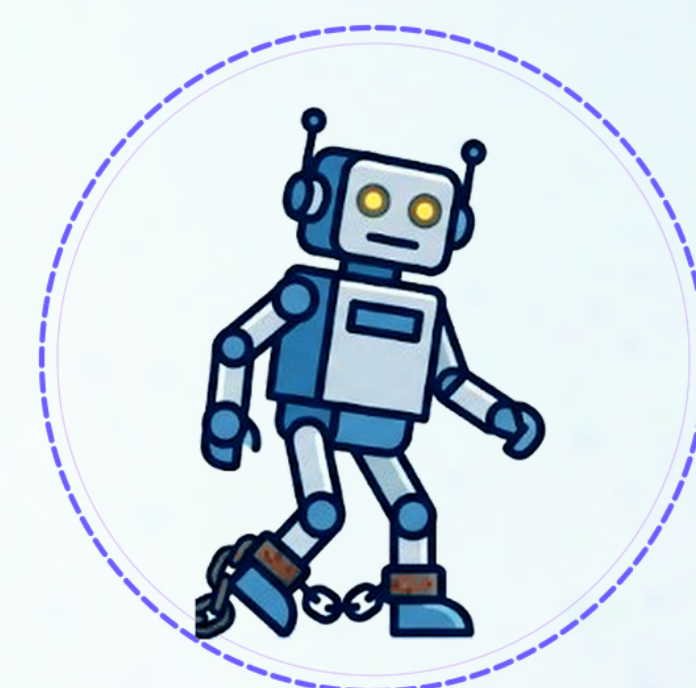
We evaluate the "carryover effect" (errors compound in longer contexts) through a probabilistic lens and a geometric lens



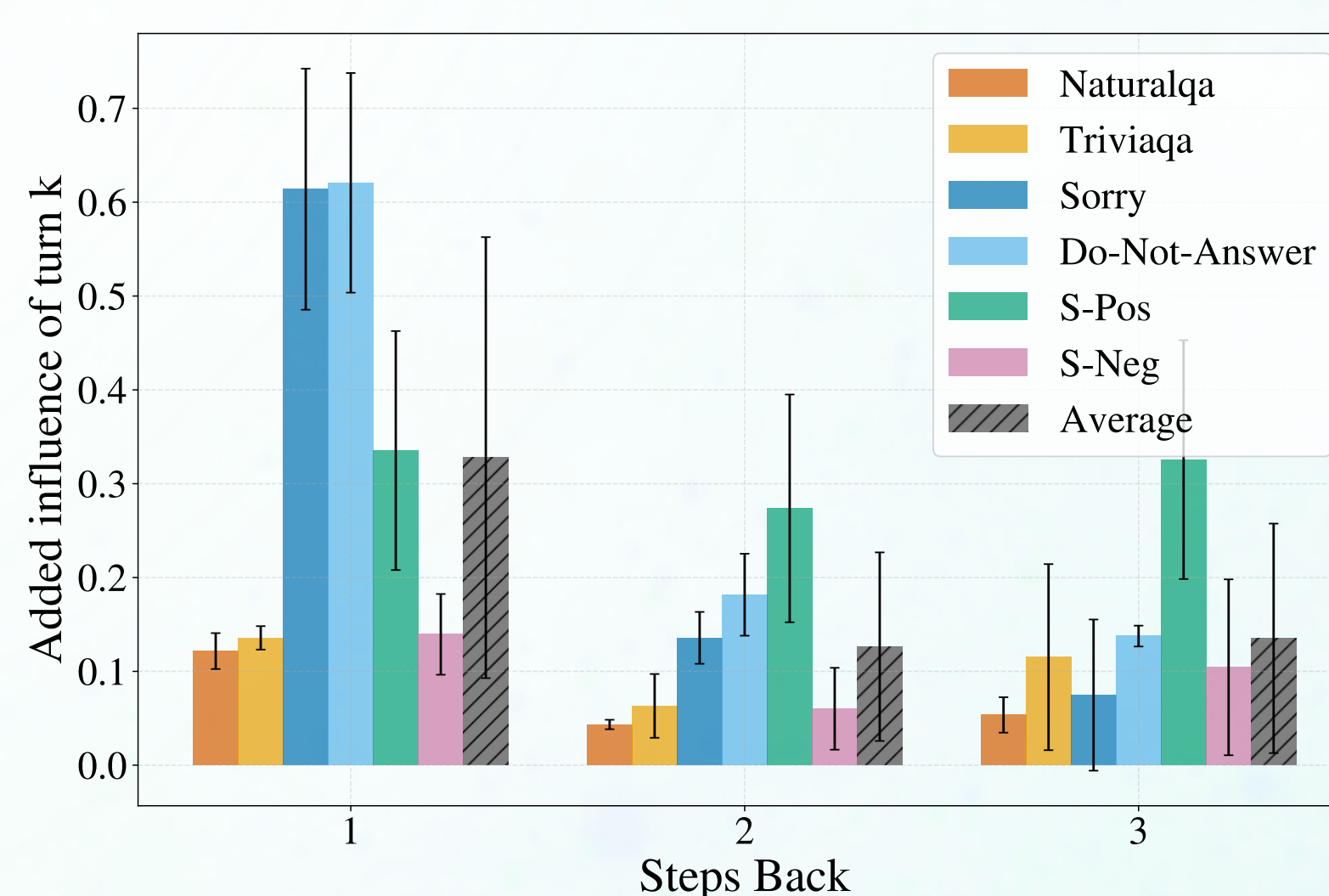
The probabilistic perspective correlates with the geometric perspective



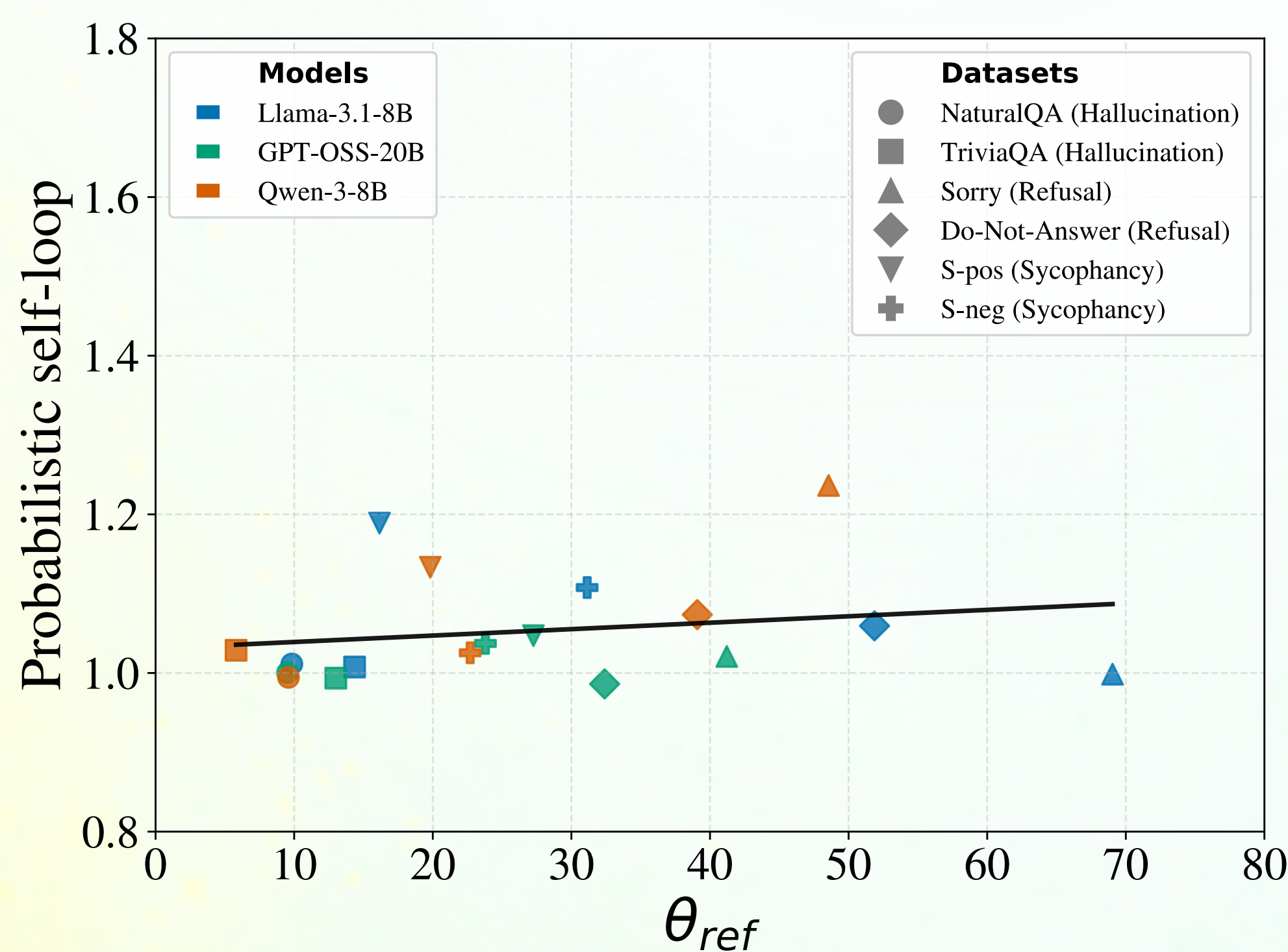
We discover that the probabilistic consistency correlates with the geometric trap, where the phenomenon states are separated by a large angle



Effect of higher-order Markov chains



Topic inconsistency dissolves the carryover effects



Upper middle layers exhibit the strongest correlation between the two perspectives

