Multiple Teacher Forcing Passes: Variance Decomposition

A. Where is the variance?

Between-prompt

Within-prompt

Base	Clone	Parallel	Premium	Unhelpful
100.0%	100.0%	99.9%	100.0%	100.0%

Variance (%)
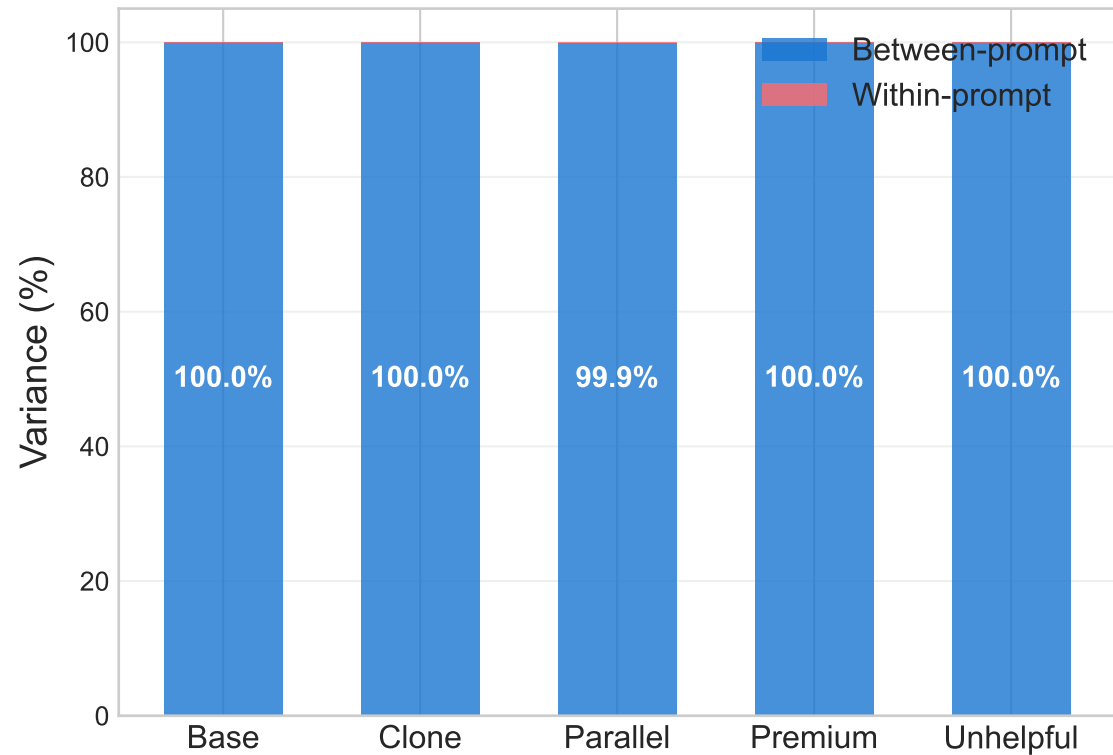
B. Five passes per prompt (Parallel policy)

Within-prompt variance: ~0.3%
Between-prompt variance: ~99.7%

Individual passes

Prompt mean

Log Probability

Different Prompts: Prompt 1, Prompt 2, Prompt 3, Prompt 4, Prompt 5