

Order-k Markov Transformer LPE Experiments

Automated run via lpe/markov_k.transformer.py

February 13, 2026

Summary

This report summarizes order-k experiments for $k \in \{1, \dots, 7\}$ with independent Beta(1,1) priors per transition state. Training used sequence length equal to rollout length: $100 \cdot 2^k$.

Core Metrics

k	Model NLL	Bayes NLL	Gap (%)	Posterior MAE	LPE Median RelErr (%)
2	0.732413	0.546434	34.04	0.1101	81.73

Table 1: Primary metrics by order k .

Figures

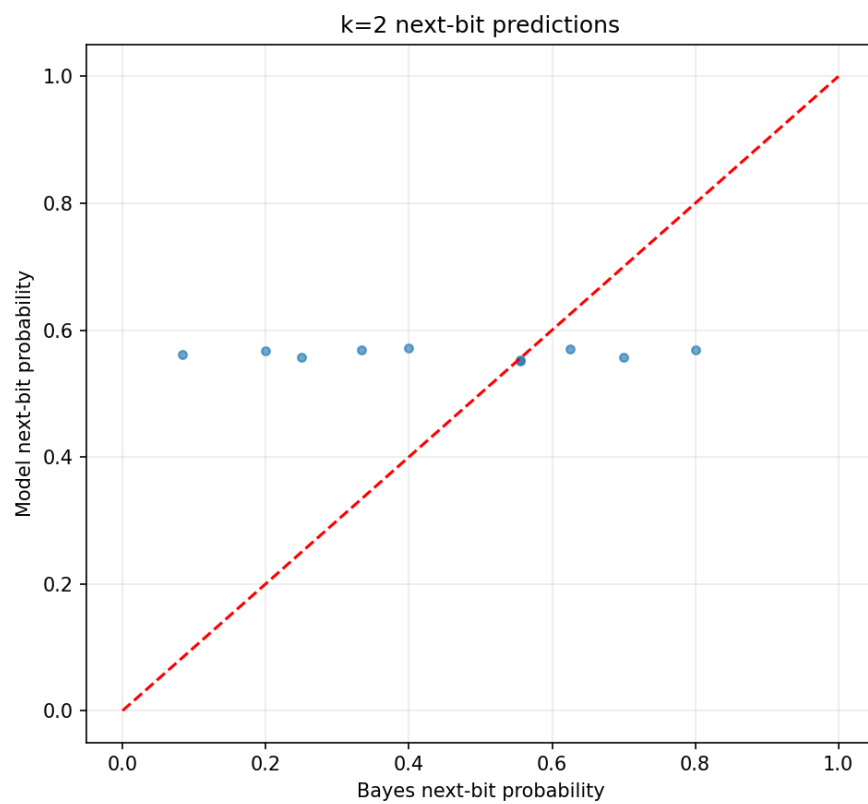


Figure 1: k=2: Step 1: next-bit prediction vs Bayes

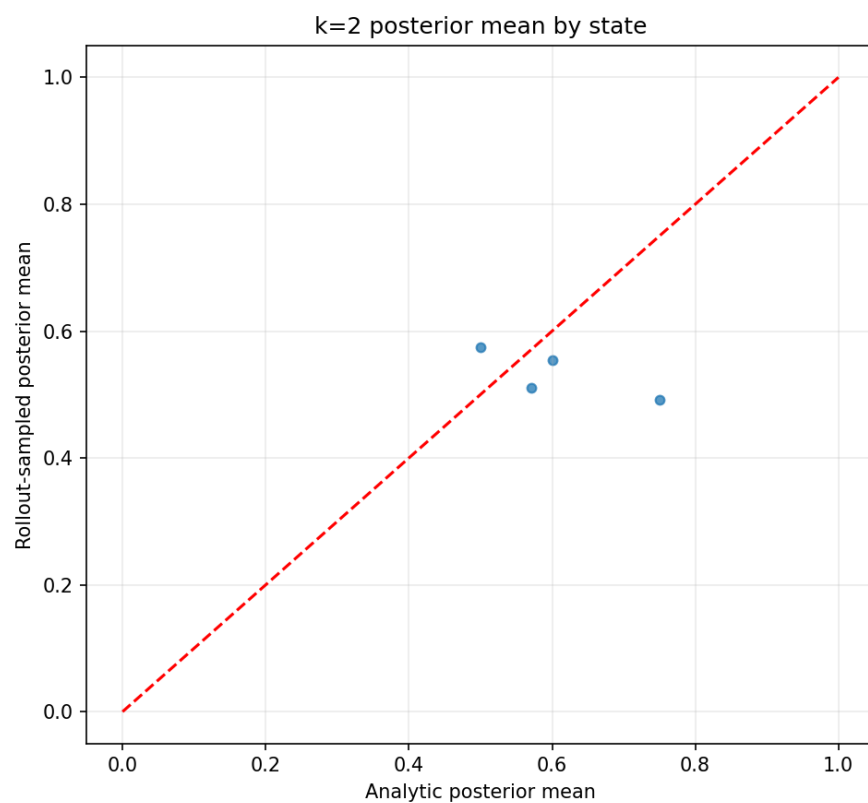


Figure 2: k=2: Step 2: posterior mean comparison

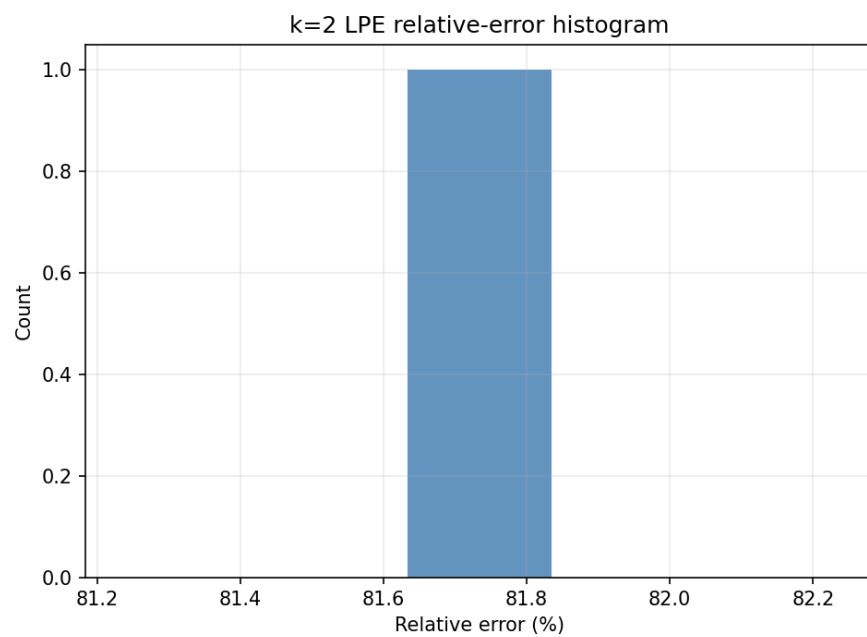


Figure 3: k=2: Step 3: LPE relative-error histogram