Order-k Markov Transformer LPE Experiments (k=1,2,3)

Automated run via lpe/markov_k_transformer.py

February 15, 2026

Summary

This report summarizes order-k experiments for $k \in \{1, \ldots, 7\}$ 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 (%)
1	0.491818	0.487222	0.94	0.0160	100.00
2	0.504101	0.493396	2.17	0.1642	2970363860660420608.00
3	0.498595	0.489566	1.84	0.0647	128.16

Table 1: Primary metrics by order k.
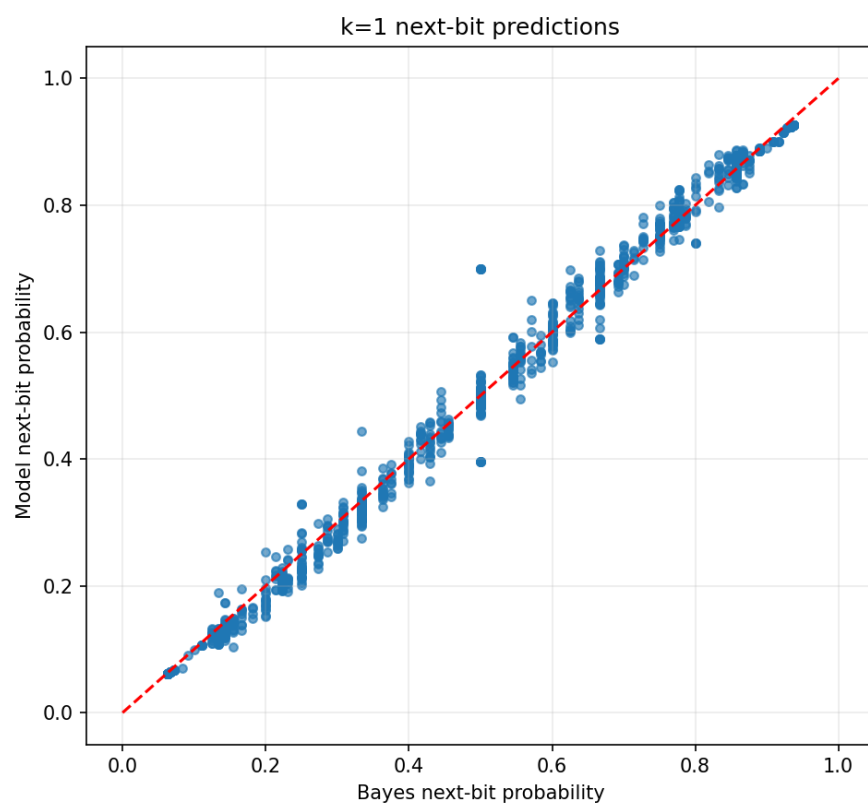
Figures

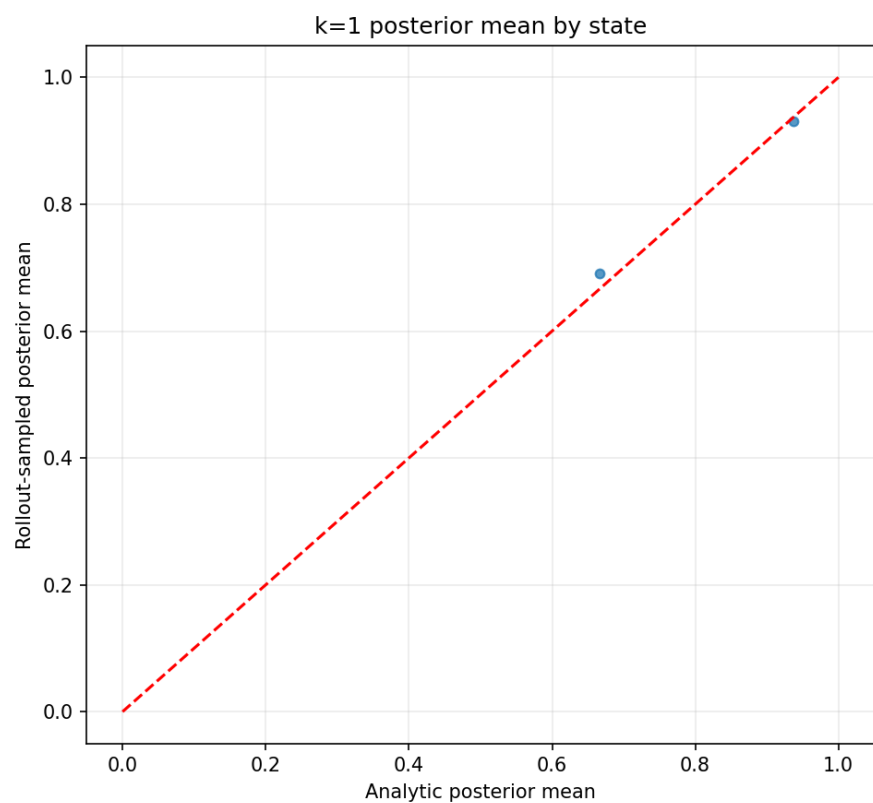Figure 1: k=1: Step 1: next-bit prediction vs Bayes

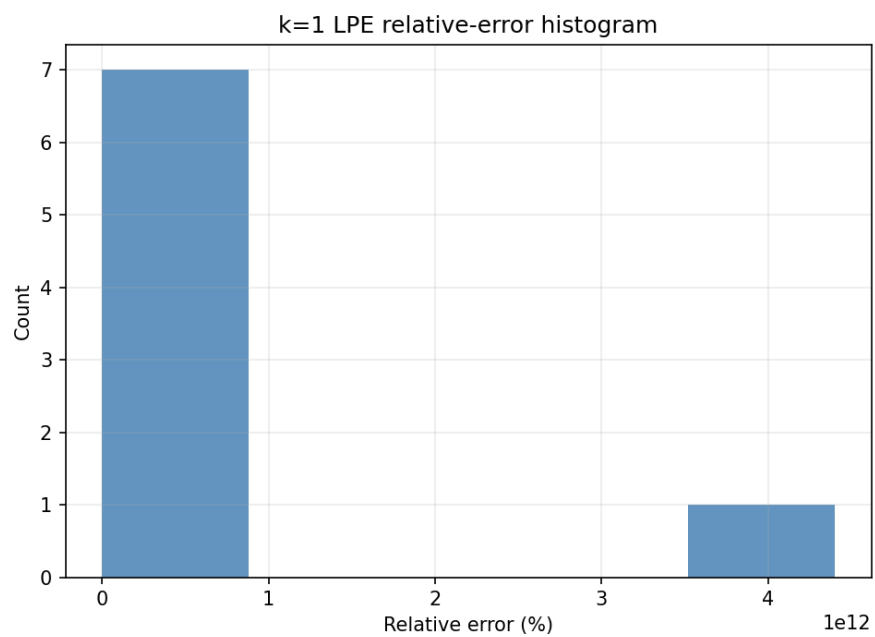Figure 2: k=1: Step 2: posterior mean comparison

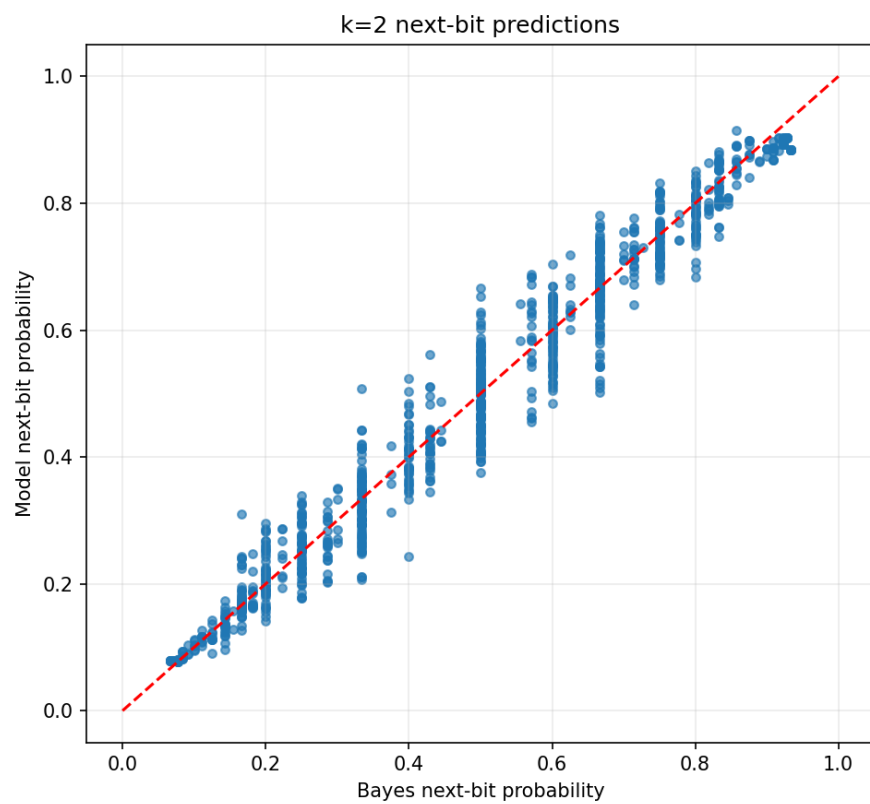Figure 3: k=1: Step 3: LPE relative-error histogram

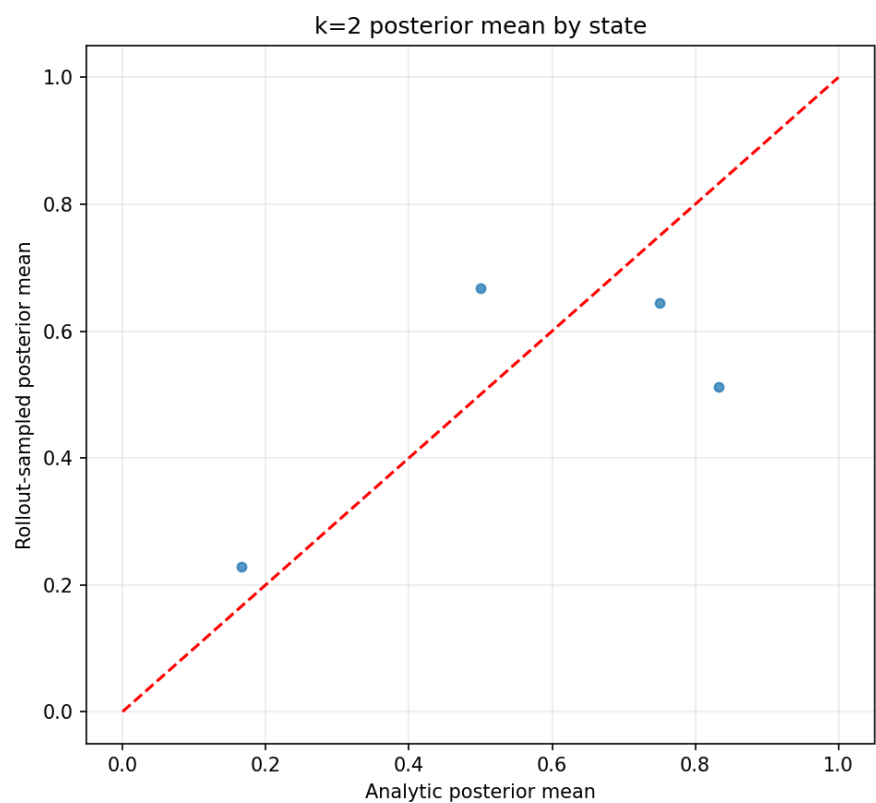Figure 4: k=2: Step 1: next-bit prediction vs Bayes

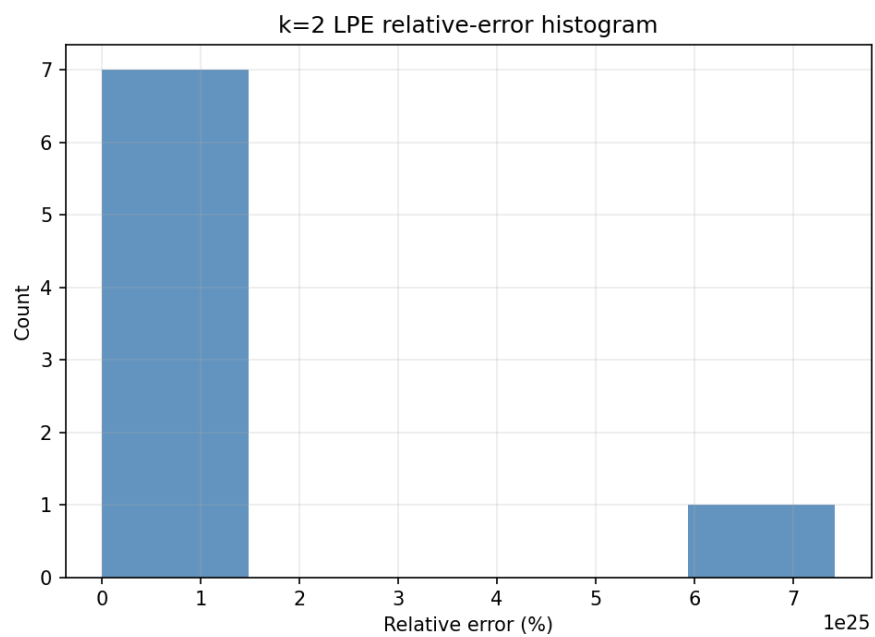Figure 5: k=2: Step 2: posterior mean comparison

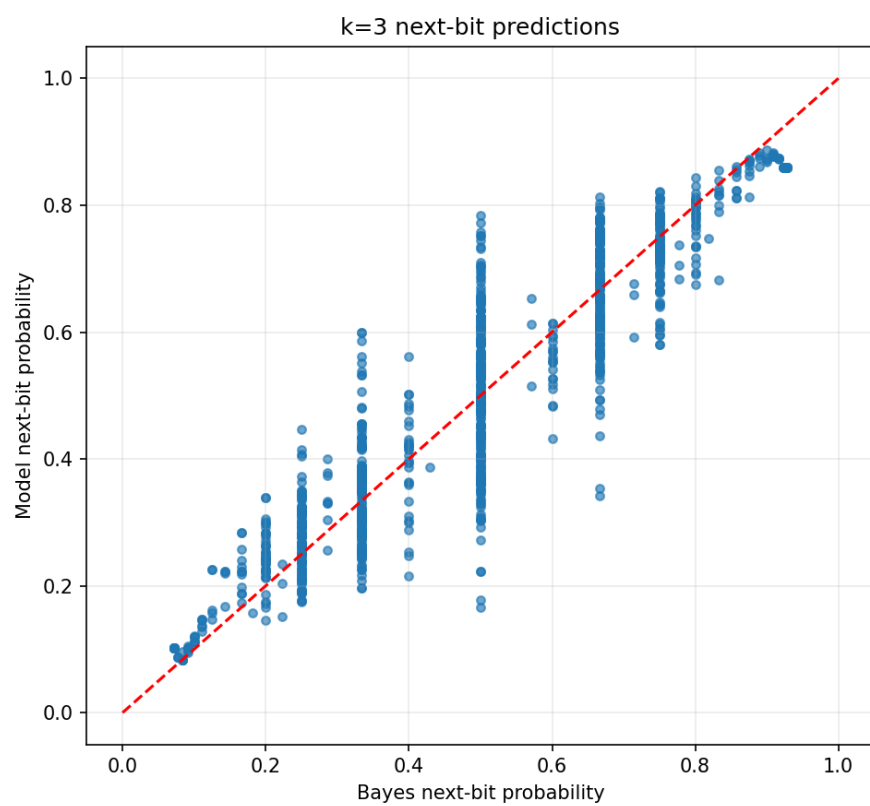Figure 6: k=2: Step 3: LPE relative-error histogram

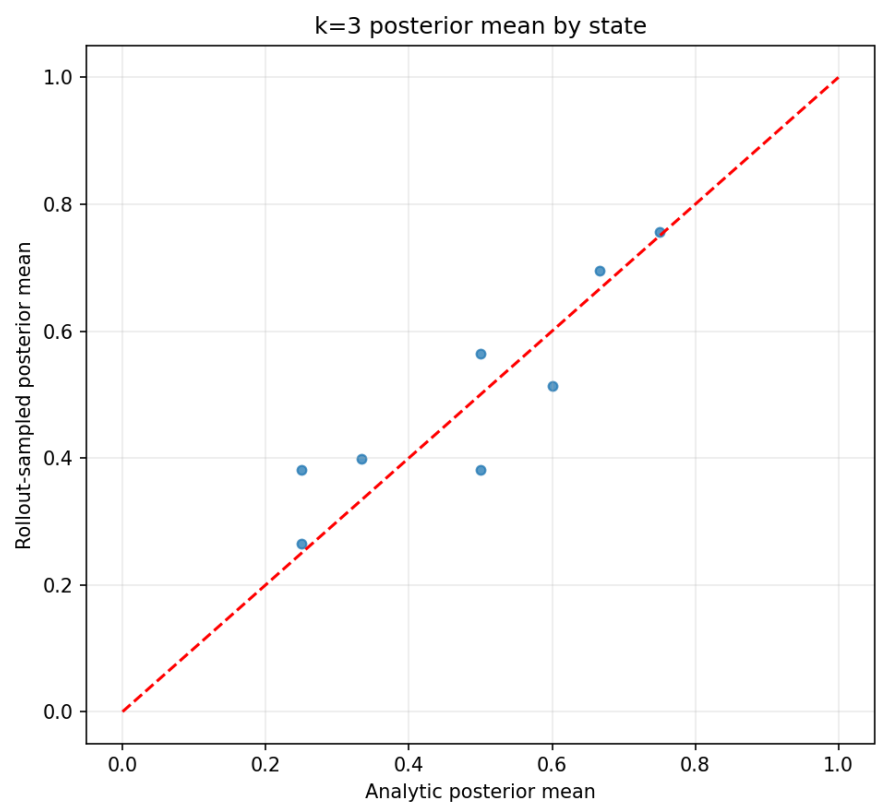Figure 7: k=3: Step 1: next-bit prediction vs Bayes
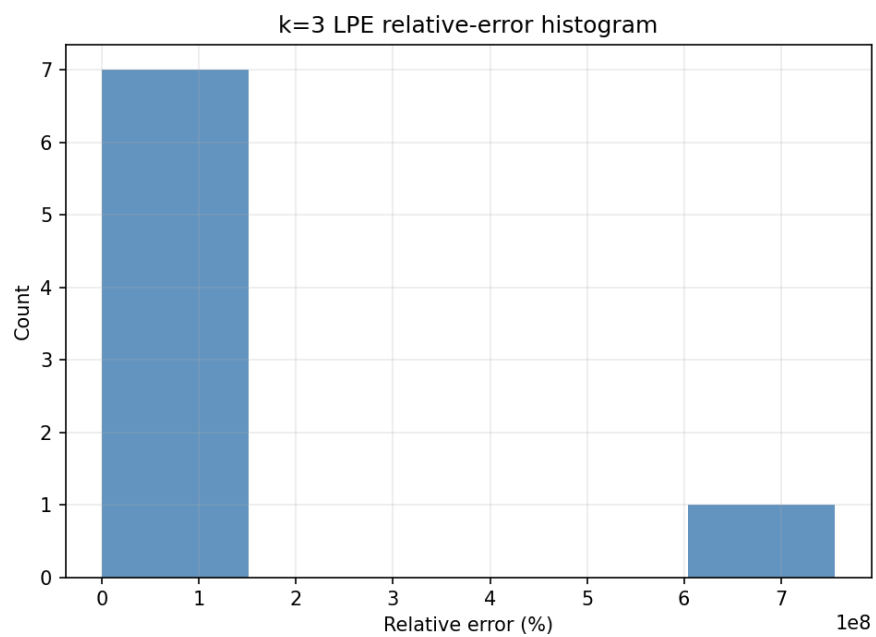
Figure 8: k=3: Step 2: posterior mean comparison

Figure 9: k=3: Step 3: LPE relative-error histogram