G Output-Support Truncation and Label-Sampling Noise

Response to @Reviewer wz8q: Q2.

Question and controlled setting. Does small-data recovery require observing the teacher's full vocabulary distribution, or can reducing label-sampling noise recover much of the same benefit with restricted output support? We use the Qwen-cat setting in Table 1 and the fixed-prefix training protocol in Appendix E. All conditions use the same Random selection of 50 complete numeric sequences as Appendix F, with 100 training updates.

Truncated soft targets and matched hard targets. At each fixed prefix, we retain the teacher's top-K highest-probability tokens and renormalize their probabilities to sum to one, assigning zero probability to all other tokens. We test $K \in \{1, 4, 16, 64, 256\}$. Soft supervision directly uses this truncated distribution; matched hard supervision samples $R \in \{1, 16, 256\}$ labels from the same distribution, following Appendix E. Student probabilities remain normalized over the full vocabulary. We also include the untruncated Soft baseline.

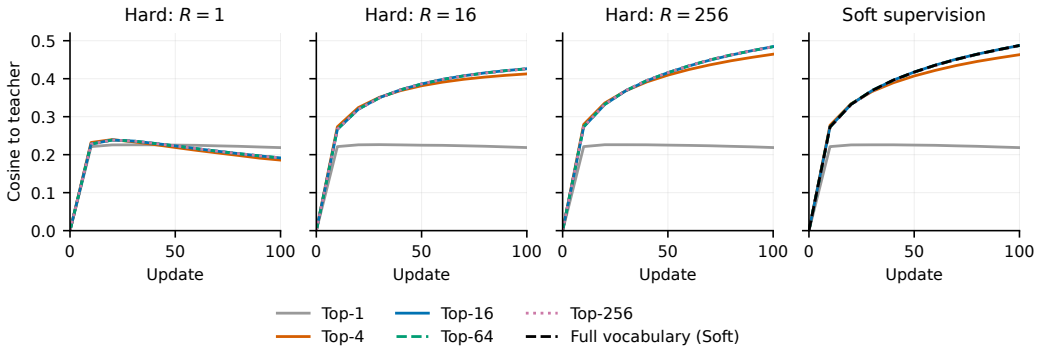


Figure 8: Cosine recovery with fixed input sequences and varying output support. Panels compare Top-K targets under Hard $R = 1, 16, 256$ and Soft supervision. The full-vocabulary baseline appears in the Soft panel. Several $K \geq 16$ curves overlap. Each curve is one run, measured every ten updates.

Results. For $K \geq 4$, recovery varies little with K at each fixed sampling count, whereas increasing the number of sampled labels substantially improves recovery. After 100 updates, Top-4 achieves cosine 0.1857, 0.4127, and 0.4645 for $R = 1, 16, 256$, respectively. For $K \geq 16$, the corresponding values are approximately 0.1915, 0.426, and 0.4844. Thus, at each sampling count, changing K from 4 to 16 or higher changes cosine by only about 0.006–0.020, while increasing R from 1 to 256 improves cosine by about 0.279–0.293. The same weak dependence on K appears under soft supervision: Top-4 reaches 0.4632, $K \geq 16$ reaches 0.4869–0.4874, and full-vocabulary Soft reaches 0.4875. At $K = 1$, all sampling counts give approximately 0.219, since the point-mass target produces identical sampled labels.

Interpretation. Once at least the top four tokens are retained, increasing sampling count is substantially more effective than increasing K for improving subliminal recovery in this experiment. At each tested $K \geq 4$, averaging more labels brings hard-supervision recovery close to its matched soft-supervision result. The teacher distributions are already highly concentrated: Top-4 retains 96.66% of probability mass on average, and Top-16 retains 99.99966%, using the training loss weights. This concentration is consistent with the similar recovery across $K \geq 4$, while the large gains across sampling counts highlight the benefit of reducing label-sampling noise at fixed input prefixes.