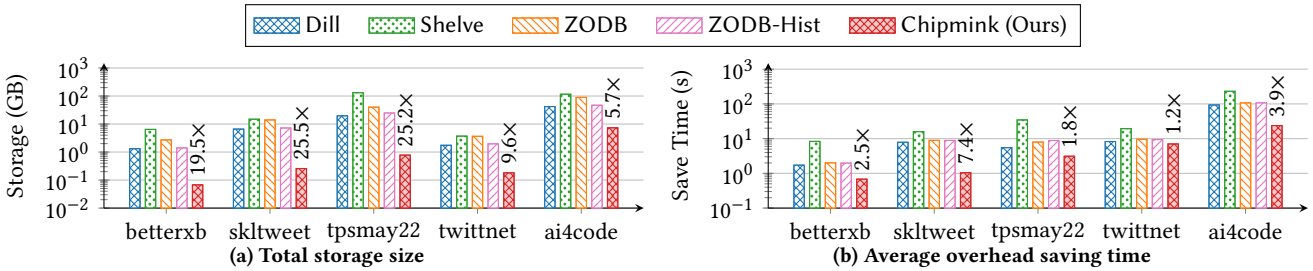


Figure 1: Chipmink compactly and quickly stores all variables, 5.7–25.5× smaller and 1.2–7.4× faster than the best baselines.

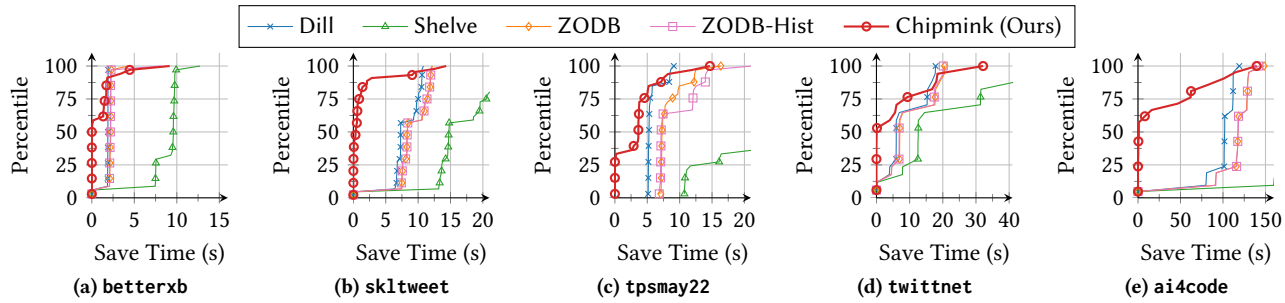


Figure 2: Empirical cumulative distributions (eCDFs) of the saving times (closer to the top-left corner are better).

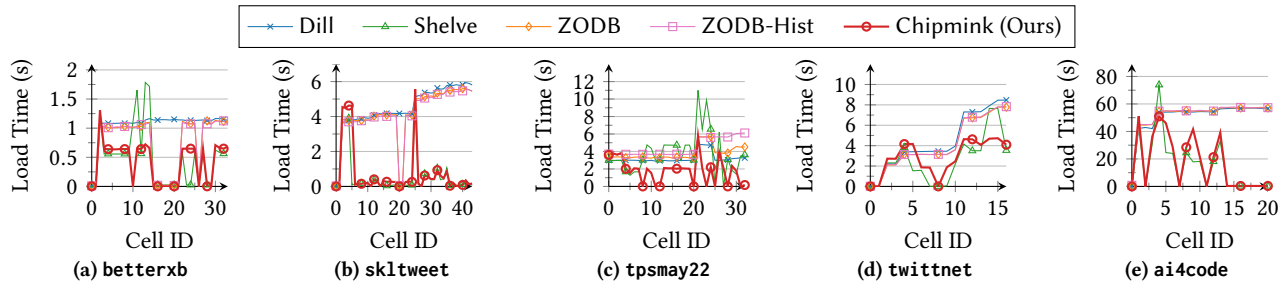


Figure 3: Partial loading time when users are interested in variables accessed at each cell. Chipmink quickly load target variables proportionally to their individual sizes, whereas some baselines' performance depend on the entire namespace size.

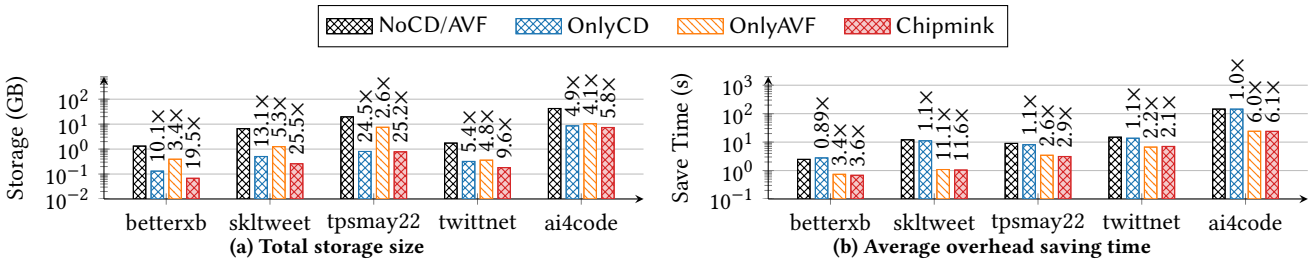


Figure 4: Ablation study: change detector (CD) and active variable filter (AVF) contribute to storage savings and speedups.

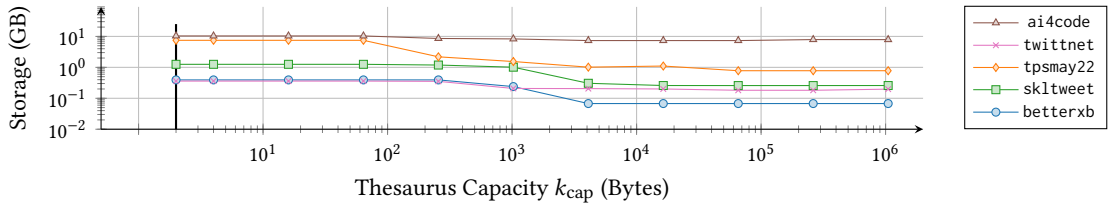
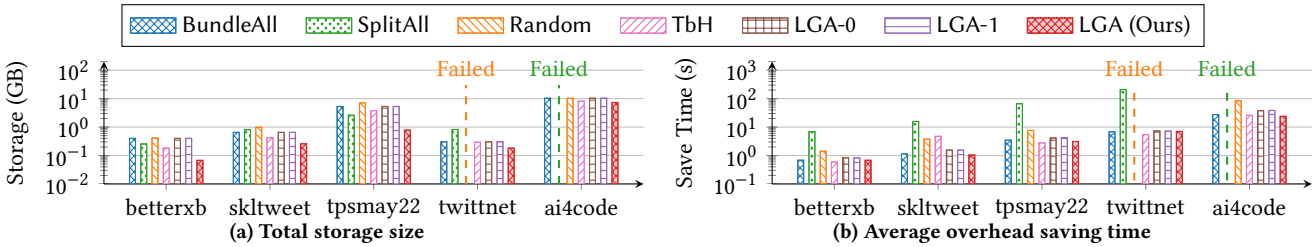


Figure 5: Higher thesaurus capacity reduces storage usage by detecting more synonymous pods. In these notebooks, storage usage converges with k_{cap} around 100 KB.



(a) Total storage size **(b) Average overhead saving time**

Figure 6: LGA is the most effective podding optimizer in discovering compact podding compared to naive methods (BundleAll, SplitAll, Random), manually derived heuristic (TbH), and LGA with inacBurate volatility models (LGA-0, LGA-1).

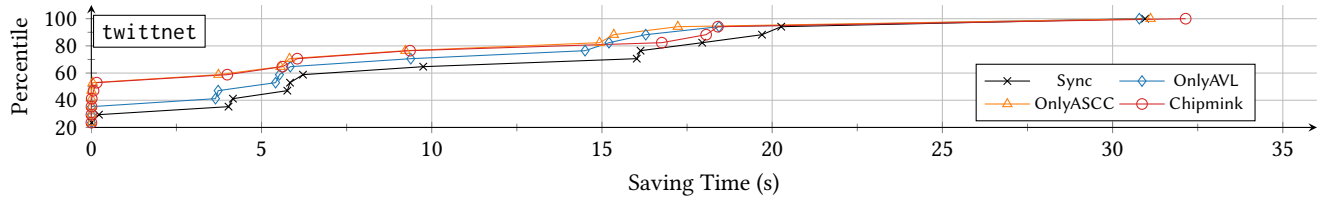


Figure 7: With parallel saving, active variable locking (AVL) and allowlist-based static code checker (ASCC) unblocks user's cell executions, improving over synchronous saving.