**Full Proof**

Here are proofs about space complexity to achieve -approximation with probability.

The target is to guarantee

With Chernoff bound: , for sub-Gaussian ERR with a variance of .

# Disk-Bounded Chunk Sketches

The total variance is

To achieve , there should be . Thus

When is small enough, there will be satisfying above. Thus , i.e.,

Thus, the total size of chunk sketches

# Disk-Bounded SSTable Sketches

When ，the variance of error of SSTable sketches is at most times of concatenated chunk sketches.

Now the target is to satisfy

At level L, the size of top sketch is . Then the size of largest top sketch is .

Let . Then ,，

The space complexity is

# Memory-Constrained Chunk Sketches

To make the merge-and-compressed chunk sketches more accurate than streaming KLL, in other words, .

That requires

For any top capacity in the streaming KLL,

Note that

I/O cost is

When , that is

# Memory-Constrained SSTable Sketches

The height of the top sketch in level L SSTable is

The target is to make the merge-and-compacted SSTable sketches more accurate than streaming KLL, invoke the lemmas bounding SSTable sketch error with Chunk sketch error and we have:

Let . Then ,

The target is to satisfy

Recall that

Again, the I/O cost is . However, L is determined by N and we need further analyze.

Now the target is to satisfy

Now Hc is bounded and the I/O cost is

When , T=10, Ts=5, b=0.3979, it is