Q1.

- The proposed algorithm is as follows.
  Reverse the edge direction of the graph.
  Initialize an empty list \( L_v \) for each vertex \( v \). The list \( L_v \) will maintain a set of pair \(<u, d>\) in increasing order of the hash value of \( u \).
  Denote \( v_j \) as the vertex with the \( j \)-th smallest hash value on hash function \( h_j \).
  Then we conduct a revised Dijkstra as follows. Start a Dijkstra algorithm from vertex \( v_j \) in the \( j \)-th iteration. The key difference in the Dijkstra traversal is that when a vertex \( v \) associated with distance \( d(v, v_j) \) is popped from the heap, we will check \( L_v \). If there are \( k \) pairs in \( L_v \) that have distances smaller than \( d(v, v_k) \), skip the subsequent traversal from out neighbors of \( v \) in the reverse graph. Otherwise, we add \(<h_i(v_j), d(v, v_j)\>\) at the end of \( L_v \) and then continue the Dijkstra as usual.

- The analysis of the algorithm is as follows.
  The key point is to derive the expected size for each vertex \( v \). Sort vertices by distance from \( v \) in ascending order. For a vertex that is at the \( i \)-th position to be included in \( L_v \), its hash value must be one of the \( k \)-smallest among the first \( i \) positions.
  If \( i \leq k \), \( P[\text{included}] = 1 \). Otherwise, \( P[\text{included}] = 1 - \frac{i-1}{i} \cdot \frac{i-2}{i-1} \cdots \frac{i-k}{i-k+1} = \frac{k}{i} \).
  The expected size for each vertex is \( \sum_{i=1}^{k} 1 + \sum_{i=k+1}^{n} \frac{k}{i} = O(k \log n) \).
  To achieve an \((\epsilon, \delta)\) guarantee for all vertices set \( k = \frac{1}{\epsilon^2} \) and apply the median trick with \( \log(n/\delta) \) copies of the bottom-\( k \) sketch.
  Thus, the space cost is \( O\left(\frac{n \log(n/\epsilon^2) \log n}{\epsilon^2}\right) \).
  The running cost to build is \( O\left(\frac{\log(n/\epsilon^2)}{\epsilon^2} (m \log n + n \log^2 n)\right) \).

Q2. Denote \( n \) as the number of nonzero entries.

- If \( n = 2 \), \( P[\text{False Positive}] \leq 2 \times \left(\frac{1}{2}\right)^2 = \frac{1}{2} \leq \frac{3}{4} \).
- If \( n \geq 3 \), you may apply the discussion in the slides.

Or you may consider bad events as follows.

1. All entries are hashed to the same slot and the sum is nonzero, which happens with a probability at most \( 2 \cdot \frac{1}{2n} = \frac{1}{n} \leq \frac{1}{4} \).
2. Both slots are occupied, where one has the sum equaling 0 and the other does not. Then consider an equivalent problem as follows. Given a set \( V \) of \( n \) nonzero numbers, uniformly sample a nonempty subset \( S \subseteq V \), which is the probability that either \( S \) of \( V \setminus S \) sums up to 0. If \( \text{sum}(V) = 0 \), no false positive will happen. So we consider \( \text{sum}(V) = a \neq 0 \). Define \( S_m \) as the largest \( S \) such that \( \text{sum}(S) = 0 \). Consider every \( T \subset S_m \) satisfying \( \text{sum}(T) = 0 \) and we have \( \forall e \in T, \text{sum}(T \setminus e) \neq 0 \). Thus \( \sum_{T \subset S_m} I(\text{sum}(T) = 0) < \frac{1}{2} \cdot 2|S_m| < \frac{1}{2} \cdot 2^{n-1} = 2^{n-2} \).
We have \( P[\text{either } S \text{ or } V \setminus S \text{ sums up to } 0] = \frac{2 \sum_{T \subseteq S, V \supseteq I} I(\text{sum}(T) = 0)}{2^n - 2} = \frac{2^n - 2 - 1}{2^n - 1} \leq \frac{1}{2}. \)

Therefore, \( P[\text{False Positive}] \leq \frac{1}{4} + \frac{1}{2} = \frac{3}{4}. \)

**Q3.** Set \( c_1 = 8 \) and \( c_2 = 16. \) Denote \( q_1, q_{1+}, q_2 \) and \( q_{2+} \) as the fail probability of 1-sparse recovery sketch (SRS), 1-sparse recovery sketch+ (SRS+), adaptive sample structure (ASS) and adaptive sample structure+ (ASS+) respectively.

For SRS, we have \( q_1 \leq \frac{3}{4} \) from Theorem 3.

\[
\text{For SRS+}, \quad q_{1+} \leq (q_{1+})^{c_1 \log n} \leq \left(\frac{3}{4}\right)^{c_1 \log n} = \left(\frac{3}{4}\right)^{8 \log n} \leq \left(\frac{3}{4}\right)^{\frac{2 \log n + 1}{\log(\frac{3}{4})}} = \left(\frac{3}{4}\right)^{\frac{1}{4}}(2n^2) = \frac{1}{2n^2}.
\]

For ASS, we have \( q_2 \leq \frac{7}{8} \) from Theorem 4.

\[
\text{For ASS+}, \quad q_{2+} \leq (q_{2+})^{c_2 \log n} \leq \left(\frac{7}{8}\right)^{c_2 \log n} = \left(\frac{7}{8}\right)^{16 \log n} \leq \left(\frac{7}{8}\right)^{\frac{2 \log n + 1}{\log(\frac{7}{8})}} = \left(\frac{7}{8}\right)^{\frac{1}{2}}(2n^2) = \frac{1}{2n^2}.
\]

Therefore, the total fail probability \( Q \leq \left\lceil \log_2 n \right\rceil q_{1+} + nq_{2+} \leq nq_{1+} + nq_{2+} \leq \frac{1}{2n} + \frac{1}{2n} = \frac{1}{n}. \)