

Sublinear Time and Space Algorithms 2024A – Problem Set 1

Robert Krauthgamer

Due: January 3, 2024

General instructions: Please keep your answers short and easy to read. You can use results, calculations or notation seen in class without repeating them, unless asked explicitly to redo them.

1. Prove that the storage requirement of Algorithm Morris+ seen in class is, with high probability, $O(\epsilon^{-2} \log \log m)$ bits.

Hint: Use Markov's inequality.

2. Suppose we are guaranteed that some item in the stream $\sigma_1, \dots, \sigma_m$ appears more than half the time, i.e., there exists (unknown) $i \in [n]$ with frequency $x_i > m/2$. Design a streaming algorithm with space complexity $O(\log n)$ bits that finds this item i .

Hint: Store only two items.

3. Extend your algorithm in the previous question to output also a $(1 + \epsilon)$ -approximation to its frequency x_i . Make sure to clearly state the space complexity of your algorithms.