Homework Assignment 1

Due Date: Sunday, December 4th, 10:30.

Instructions:

1. The homework assignment will be done individually.

2. If you use any result and/or material from books, papers, or online, you need to mention and reference it in every part of your solutions.

3. If you use any programming code in your solutions, you need to include the code you use as part of your solution.

Problem 1.
In this problem, we will calculate the expected number of writes of some rewrite codes.

1. Consider the rewrites approach from slide 17 in Lecture 2 for a cell with 7 levels (level 0, level 1,\ldots, level 6).
   
   (a) Assume a symbol with 7 values is written to this cell with the same probability for each value. What is the expected number of successful writes? (note that the worst case is 1).
   
   (b) Explain how to write to this cell a symbol with four values twice at the worst case.
   
   (c) What is the expected number of writes for the code you designed in the previous section if the four values are written with the same probability on each write.

2. Find the expected number of writes of the Rivest Shamir WOM code (slides 19-20 in Lecture 2). Assume that every message is written with the same probability (that is, the probability to write every pair of two bits is 1/4). What is then the expected number of successful writes?

Problem 2.
In this problem, we will follow up on the WOM codes construction by the coset coding scheme from slides 9-10 of Lecture 4 and the two-write WOM codes construction in slides 11-16 of Lecture 4. For shorthand, this construction will be called in this problem Construction A.

1. Prove that the WOM code given in slide 10 is a \([7, 3; 8, 8]\) WOM code.

2. Let \(H\) be the \(m \times (2^m - 1)\) parity check matrix of a Hamming code of length \(2^m - 1\), dimension \(2^m - m - 1\), and minimum distance 3. Prove that using the coset coding scheme, it is possible to construct a WOM code with \(2^m - 1\) cells such that it is possible to write \(m\) bits \(2^m - 2 + 1\) times.
3. What are the WOM code rates when choosing the matrix

\[ H = \begin{pmatrix}
1 & 1 & 1 & 0 & 0 & 0 \\
1 & 1 & 0 & 1 & 1 & 0 \\
1 & 0 & 0 & 0 & 1 & 1
\end{pmatrix}, \]

in Construction A?

4. In this part we will modify Construction A such that the construction will use two matrices \( H_1 \) and \( H_2 \), each of size \( r \times n \) and of full rank.

(a) Propose a two-write WOM code of \( n + 1 \) cells which uses the two matrices \( H_1 \) and \( H_2 \). Your construction should be as similar as possible to Construction A. Explain how the encoding and decoding maps work for each of the two writes and prove their correctness. Finally, find the individual rates of the WOM codes (in a similar way to slide 15).

(b) Demonstrate your WOM code construction with an example for the two matrices

\[ H_1 = \begin{pmatrix}
1 & 0 & 0 & 1 & 1 & 0 & 1 \\
0 & 1 & 0 & 1 & 0 & 1 & 1 \\
0 & 0 & 1 & 0 & 1 & 1 & 1
\end{pmatrix}, \quad H_2 = \begin{pmatrix}
0 & 1 & 1 & 1 & 1 & 0 & 0 \\
1 & 0 & 1 & 1 & 0 & 1 & 0 \\
1 & 1 & 0 & 1 & 0 & 0 & 1
\end{pmatrix} \]

and calculate its rate on each write.

(c) Bonus: Write a computer program that chooses the two matrices \( H_1 \) and \( H_2 \) uniformly at random and calculate the rates of the WOM code they provide. Run your program and give the WOM codes with the best sum-rate you could find.

**Problem 3.** In this problem, we will study how to build a computer program that searches for two-write WOM codes. Assume there are \( n \) cells and \( 0 \leq k \leq n \) is a given parameter. We will construct a bipartite graph consisting of two groups of vectors: \( S_L \) and \( S_R \). On its left side, the set \( S_L \) consists of all binary vectors of weight at most \( k \) and on its right side, the set \( S_R \) consists of all binary vectors of weight at least \( k \). There is an edge between a vector \( v_L \in S_L \) and a vector \( v_R \in S_R \) if and only if \( v_L \leq v_R \).

1. Draw the bipartite graph for \( n = 4 \) and \( k = 2 \).

A WOM code, according to this bipartite graph, will enable to write \( M_1 = \sum_{i=0}^{k} \binom{n}{i} \) messages on the first write, which are encoded to all possible vectors of the first group \( S_L \). On the second write, we want to design the WOM code such that \( M_2 \) messages can be written. Assume one assigns a mapping \( F : S_R \rightarrow \{1, \ldots, M_2\} \).

2. Explain how it is possible to check whether the mapping \( F \) will generate a WOM code that allows to write on the worst case \( M_2 \) messages on the second write.

3. Design an algorithm (the best you can think of) to find such a mapping \( F \) and explain how the value \( M_2 \) for the number of messages on the second write is derived from this algorithm.

4. Write a computer program that implements your algorithm and report on different WOM codes you could find using your program. In particular, provide the WOM code of largest sum-rate you could find.
Problem 4.

1. Write the capacity region of a four-write WOM codes and find the probabilities $p_1, p_2, p_3$ that maximize the sum-rate to achieve value $\log_2 5$.

2. Find the maximum sum-rate of fixed-rate four-write WOM codes.

Problem 5.
In this problem, we will prove that $\log_2 (t + 1)$ is an upper bound on the sum-rate of any $t$-write WOM code.
Assume $C$ is an $[n, t; M_1, M_2, \ldots, M_t]$ $t$-write WOM code. For any sequence of messages $(m_1, \ldots, m_t)$, where for $1 \leq i \leq t$, $m_i \in \{1, \ldots, M_i\}$, $A(m_1, \ldots, m_t)$ is an array of size $t \times n$ such that its $i$-th row, for $1 \leq i \leq t$, is the cells values after the $i$-th write. Prove the following:

1. Write the matrix $A$ for the Rivest Shamir code if the first two information bits are 10 and the second two information bits are 11.

2. Show that for any two different sequences of messages $(m_1, \ldots, m_t) \neq (m'_1, \ldots, m'_t)$, $A(m_1, \ldots, m_t) \neq A(m'_1, \ldots, m'_t)$.

3. Let $A$ be a matrix of size $t \times n$ which is equal to $A(m_1, \ldots, m_t)$ for some sequence of messages $m_1, \ldots, m_t$. How many different options does every column of $A$ have? How many different such matrices $A$ can exist?

4. Conclude that the sum-rate of $C$ is at most $\log_2 (t + 1)$, that is,

$$\sum_{i=1}^{t} \log_2 M_i \leq \log_2 (t + 1).$$