In this chapter we describe the path which we took to build the in-silko simulation. First we start by re mentioning the questions that need to be answered to build a good group testing model. First, what is the best matrix that one uses to obtain best results. this question is important as the rows of the matrix represent the number of tests to be ran, so fewer rows are better, but also this should not come at a cost of the recovery. The second question which concerns the best algorithm to be used in order to get the best recovery results. Here in this work we focus on an iterative algorithms called the Hard thresholding algorithms. Where we write an implantation of the algorithms using programming language Python as well as running some performance tests.

1 The measurement matrix

Our goal is to solve the minimization problem which is sometimes refereed to as the basis pursuit,

$$\min||x||_1, \text{subject to} \quad y = Ax \tag{1}$$

Here we introduce some theoretical guarantees for finding the best measurement matrix.

1.1 Mutual coherence

The mutual coherence of a matrix A is the largest absolute correlation between the columns of the matrix A. It is important to note that if two columns of A are strongly correlated then it will hard to know their contribution to the measurement result y. In and extreme case where the the mutual coherence is very high, then it would nearly impossible to recover a sparse signal[ref here].

Theorem 1 *Assume that $||x||_0 \leq d$ for the true vector x and $\mu < \frac{1}{2d-1}$. Then, x is a solution for the l_0 and l_1 problems.*

[ref here] where μ is the mutual coherence metric, and for a matrix A it is defined by

$$\mu(A) = max_{i \neq j} \frac{|<a_i, a_j>|}{||a_i||_2||a_j||_2} \tag{2}$$

1.2 Restricted isometry property (RIP)

restricted isometry constant δ of a matrix A, $\delta(A)$ is the smallest number such that the following inequality holds:

$$1 - \delta(A)||x||_2^2 \leq ||Ax||_2^2 \leq (1 + \delta(A))||x||_2^2$$

A matrix A is said to satisfy the restricted isometry property with constant $\delta(A)$ if $\delta(A) < 1$ [ref here].

Theorem 2 *Assume that $||x||_0 \leq d$ for the true vector x and $\delta_{2d} < \sqrt{2} - 1$ then x is the unique solution of the l_0 and l_1 problems.*

Several improvements has been done the RIP condition, other types also available for instance $\delta < 0.1$ [ref here]. It is also known that stronger results can be obtained by using RIP as compared to the mutual coherence. But it can be seen that unlike the mutual coherence the RIP has a higher computation complexity.

2 Practical considerations

The measurement matrix is called deterministic if every test is obtained or determined with a probability of 1. While in contrast, the matrix is deterministic if the tests are arranged to some probabilistic distribution, or in simple words part of the matrix or the whole matrix is obtained by chance. In this work the matrix is generated randomly, particularly, the elements the matrix are generated using the **uniform random distribution**. It is known from the literature that a randomly generated matrix actually preserves the restricted isometry property [ref here].

Another consideration is the distribution of positive, i.e. the location of positives and the number of positives in the vector x, that is the vector to be recovered. There are two common distribution in the literature, the *probabilistic*, in which the locations and number of positives in the tested set is determined according to some probability. The other type is combinatorial scheme, where here the any set of up to d defectives can be positive. Here the number of defective is known in advance, where as their locations are set randomly. In our work here we considered the second case. In combinatorial case several schemes have proposed to attain a low number of test namely of $O(d^{1+o(1)}log^{1+o(1)}n)$ [ref here]. It is important to note that we are interested in the case where the number of individuals to be tested is very big i.e. $n \to \infty$, whereas the number of defectives d is constant.