# DESIGN AND ANALYSIS OF HIGH PERFORMANCE AND LOW POWER MATRIX FILLING FOR DNA SEQUENCE ALIGNMENT ACCELERATOR USING ASIC DESIGN FLOW

**Project report presented in the partial fulfillment for the award of the Bachelor of Electrical Engineering (Hons)**

**UNIVERSITI TEKNOLOGI MARA**


**NURUL FARHANA BINTI ABD RAZAK**

**2007270476**

**FACULTY OF ELECTRICAL ENGINEERING**
**UNIVERSITI TEKNOLOGI MARA**
**40450 SHAH ALAM, SELANGOR**

# ACKNOWLEDGEMENT

In the name of Allah, The Most Gracious, The Most Merciful and The Most Beneficant. Praise in only Allah S.W.T for his bounty and blessing upon us. It is with deepest sense of gratitude to Allah who has given me the strength and ability to complete this project as it is today.

This paper could not be a possible without Puan Norhazlin bt Khairudin. I do pay great homage and concede Puan Norhazlin bt Khairudin for her pioneering in designing the matrix filling and her immense technical support to accomplish this goal. I've spent couple of hours with her to share vision and thoughts which are most memorable for me and I look forward to see her again in near future.

I would also like to express my deepest thanks to my friends, as my course mates in UiTM who had shares with me their valuable ideas in this project.

Also special thank to Mr Abdul Karimi and Mr Syed who are my co-supervisor for their guidance and knowledge in Xilinx and ASIC flow.

Special thanks also to all lecturers in Faculty of Electrical Engineering, UiTM and to those who have devoted their time either directly or indirectly, for their ideas, support and a lot of contribution towards the success of this project.

Last but not least, to both of my parents and family for their understanding and support throughout the years. You all are the source of my strength and inspiration.

Nurul Farhana bt Abd Razak
Faculty of Electrical Engineering
University Teknologi MARA (UiTM)

# ABSTRACT

Efficient sequence alignment is one of the most important and challenging activities in bioinformatics. Many algorithms have been proposed to perform and accelerate sequence alignment activities. Among them Smith-Waterman (S-W) is the most sensitive (accurate) algorithm. This paper presents a novel approach and analysis of High Performance and Low Power Matrix Filling for DNA Sequence Alignment Accelerator by using ASIC design flow. The objective of this paper is to improve the performance of the DNA sequence alignment and to optimize power reduction of the existing technique by using Smith Waterman (SW) algorithm. The scope of study is by using the matrix filling method which is in parallel implementation of the Smith-Waterman algorithm. This method provides more efficient speedup compared to the traditional sequential implementation but at the same time maintaining the level of sensitivity. The methodologies of this paper are using FPGA and Synopsis. This technique is used to implement the massive parallelism. The design was developed in Verilog HDL coding and synthesized by using LINUX tools. Matrix Cells with a design area 8808.307$mm^2$ at 40ns clock period is the best design. Thus the power required at this clock period also smaller; dynamic power 111.1415uW and leakage power 212.9538Nw. This is a large improvement over existing designs and improves data throughput by using a ASIC design flow.

## TABLE OF CONTENT

# CHAPTER 1

# INTRODUCTION

## 1.1 BACKGROUND OF STUDY

In the past decade there has been an explosive growth of biological data the rapid expansion in digitization of patient biological data [1]. This has led to the emergence of a new area of research: bioinformatics, where powerful computational techniques are used to store, analyze, simulate, and predict biological information. Bioinformatics refers to the storage, analysis, and simulation of biological information and the prediction of experimental outcomes.

The general problem statement was the genomic data at GenBank (the NIH genetic sequence database, an annotated collection of all publicly available DNA sequences) is doubling every six months. Presently, the problem is the proteomic and cellular imaging data is expected to grow even faster. Post-genomic-era bioinformatics will require high-performance computing power of the order of several hundred of teraflops or more. In recent years, FPGAs have emerged as high-performance computing accelerators capable of implementing fine-grained, massively parallelized versions of computationally intensive algorithms. The reprogrammability of FPGAs enables algorithm specific computing architectures to be implemented using the same hardware resource across a range of algorithms. In this thesis, it will describe the design flow, test results, and system optimization for implementing a Smith Waterman algorithm using a ASIC design flow. The significance of the study is to improve the performance and optimize the power consumption used. Besides, the bioinformatics