XILINX OPEN HARDWARE COMPETITION: Acceleration of LU Decomposition on FPGAs

Yichen Zhang

16th June 2022

Team number: xohw22-006
Supervisor: Dr Danial Chitnis

Abstract

A brief document to illustrate how to use LaTeX.

1 Introduction

1.1 Sparse Linear System

Solving a linear system $Ax = b$ is relatively a basic algorithm problem and is closely related to the applications of almost every field, including engineering, physics, chemistry, etc. It is at the heart of various algorithms and applications, including computational geometry, circuit simulation and data science. A variety of approaches have been developed to solve the linear system, which can mainly be classified into two types: direct methods and iterative methods.

The direct methods solve the sparse linear system mainly by computing the determinant, inversion or factorisation of a matrix. For example, Cramer's rule calculates the solution in terms of the determinants of the coefficient matrix and the right-hand-side vector. Gaussian elimination finds the solution by calculating the inverse of the system. LU decomposition solves the system by factorising the system into a lower triangular matrix L and an upper triangular matrix U.

On the other hand, the iterative method is more modern. It attempts to solve the linear system based on an initial condition and successive approximates to the final solution, such as spectral sparsificaiton of graphs [1] and division-free inversion [2]. However, the behaviour and its stability of iterative methods is heavily based on the convergence[1] of the input matrices. On the contrary, the operation cost of behaviour methods is closely related to the cost of the matrix multiplication, making it preferred to the randomised and iterative methods in real applications as the directed methods are more robust and predictable.

LU decomposition is a direct method that can solve a large sparse linear systems multiple times, with various applications in circuit simulation, structure analysis, power networks, etc. A variety of parallel sparse linear solvers have adopted LU decomposition and have been running on massively parallel supercomputers. For example, Cray XE6 [3], which is a type of distributed-memory machines, has implemented SuperLU_DIST [4] solver for LU decomposition.

---

[1]For an $n \times n$ matrix A, A is convergent if:

$$\forall i, j \in [1, n], i, j \in \mathbb{N} \Rightarrow \lim_{k\to\infty} \left(A^k\right)_{ij} = 0$$

With the continuous development of IC industry, the size of FPGAs has grown to the extent that intense and heavy floating-point operations can be now accommodated. After a decade of research, it has been proved that the accelerating algorithms on FPGAs is a promising research avenue. Modern FPGAs have made it possible to fit a large computation kernel and establish parallel computation machines. In this paper, a sparse linear solver is build based on a CPU-FPGA architecture. While the pre-processing is made on CPU, the FPGA performs the numeric factorisation and solving of the matrices.

References

[1] D. A. Spielman and S. Teng, "Spectral Sparsification of Graphs," *CoRR*, vol. abs/0808.4134, 2008, arXiv: 0808.4134, [Online]. Available: http://arxiv.org/abs/0808.4134.

[2] I. Koutis, G. L. Miller and R. Peng, "A Fast Solver for a Class of Linear Systems," *Commun. ACM*, vol. 55, no. 10, pp. 99–107, Oct. 2012, ISSN: 0001-0782, DOI: 10.1145/2347736.2347759.

[3] J. Kwack, G. Bauer and S. Koric, "Performance test of parallel linear equation solvers on Blue Waters–Cray XE6/XK7 system," English (US), 2016.

[4] X. S. Li and J. W. Demmel, "SuperLU_DIST: A Scalable Distributed-Memory Sparse Direct Solver for Unsymmetric Linear Systems," *ACM Trans. Math. Softw.*, vol. 29, no. 2, pp. 110–140, Jun. 2003, ISSN: 0098-3500, DOI: 10.1145/779359.779361.