

PARALLEL COMPUTATION OF DIAGONALLY DOMINANT LINEAR SYSTEMS

RABIA KAMRA



**DEPARTMENT OF MATHEMATICS
INDIAN INSTITUTE OF TECHNOLOGY DELHI
OCTOBER 2017**

©Indian Institute of Technology Delhi (IITD), New Delhi, 2017

PARALLEL COMPUTATION OF DIAGONALLY DOMINANT LINEAR SYSTEMS

by

Rabia Kamra

Department of Mathematics

Submitted

*in fulfillment of the requirements of the degree of Doctor of Philosophy
to the*



Indian Institute of Technology Delhi
October 2017

To
My Parents

Certificate

This is to certify that the thesis entitled **Parallel Computation of Diagonally Dominant Linear Systems** submitted by **Ms. Rabia Kamra** to the Indian Institute of Technology Delhi, for the award of the Degree of **Doctor of Philosophy**, is a record of the original bona fide research work carried out by her under my supervision and guidance. The thesis has reached the standards fulfilling the requirements of the regulations relating to the degree. The results contained in this thesis have not been submitted in part or full to any other university or institute for the award of any degree or diploma.

New Delhi
October 2017

Dr. S. Chandra Sekhara Rao
Professor
Department of Mathematics
Indian Institute of Technology Delhi
New Delhi-110 016

Acknowledgements

First and foremost, I thank the Almighty God for his grace and blessings on me which made me to reach this place and gave me strength to complete my work.

From the core of my heart, I express my deep sense of gratitude to my supervisor Prof. S. Chandra Sekhara Rao for his constant support, encouragement and constructive suggestions throughout in all the years of Ph.D. I am highly thankful to him for introducing me to the area of parallel computing. His wide knowledge and logical way of thinking have been of great value for this thesis. Without his support, this thesis would not have been possible.

I am greatly indebted to my S.R.C. (Student Research Committee) members: Prof. R. K. Sharma (Chairman), Prof. K. Sreenadh and Prof. H. K. Malik, for their valuable comments and suggestions during my presentations. I am very much thankful to all faculty and staff members of the Department of Mathematics, IIT Delhi for their help and support. I am grateful to IIT Delhi for providing me scholarship for my Ph.D. work. I thank IIT Delhi authorities, security personnel, hostel staff and hostel mess staff who all collectively make a peaceful environment to live and study. I also thank Prof. A. K. Saroha, Associate Dean of Hostel Management, for allowing me to stay in the New Vindhyachal apartments.

It is beyond the scope of any acknowledgement for what I have received from my parents, yet I make an effort to express my heartfelt and affectionate gratitude for their co-operation, patience and keen interest in my progress during my Ph.D. I am highly obliged to my parents Sh. Rajeev Kamra and Smt. Veena Kamra for all their hard work to provide a stable foundation for my life and future. Thanks to my sister and brothers, Priyanka,

Parav, Veeru and Divansh for their love through this long process. I owe my thanks to my husband, Dr. Suraj Goyal for always being by my side. Without his encouragement, endless patience and understanding it would have been impossible for me to complete this work. I am thankful to my son, Krishna, for being the great source of love and relief from scholarly endeavor. I acknowledge my deep gratitude to the memory of my beloved grandfather and grandmother and thank them for their everlasting blessings. My heart felt regard goes to my father in law, mother in law and sister in law for their love and moral support.

I would also like to thank all my friends Pooja Punyani, Shrashti Diwedi, Prachi Singh, Dr. Swati Sidana, Priyanka Singh, Dr. Arti Pandey, Dr. Anubha Jindal and all my juniors who generously shared time at different stages of my work, and the enjoyable moments spent with them will always be memorable. I also enjoyed the company of my senior colleagues; Sheetal Chawla and Manisha Srivastva.

New Delhi

Rabia Kamra

Abstract

In this work, we develop fast, accurate, numerically stable and scalable parallel algorithms for the solution of diagonally dominant linear systems. Solving these linear systems $Ax = f$ is central to the field of scientific and engineering computations. We give a WZ factorization for the solution of dense diagonally dominant nonsingular linear systems. The existence of the factorization is given. The backward error analysis of the WZ factorization method is given and numerical stability of the method is established.

The WZ factorization for the diagonally dominant tridiagonal coefficient matrix is given. Existence of the factorization is proved for diagonally dominant nonsingular matrices. A parallel algorithm for the solution of diagonally dominant tridiagonal linear systems based on the given WZ factorization, with a small modification, is developed using divide and conquer technique. The coefficient matrix is partitioned into r blocks along the main diagonal and all partitions are related to each other with one coupling variable. A reduced system of order $2r$ is constructed from the coupling variables of each partition. Solution of the original linear system depends on the solution of reduced system. It has been proved that the reduced system has the same properties, like nonsingularity, tridiagonal structure and diagonally dominance, as that of the original linear system which ensures the existence of solution of the reduced system. The backward error analysis of the parallel algorithm is given and numerical stability is proved. The idea used for tridiagonal linear systems is extended to the diagonally dominant narrow banded coefficient matrix of order N with semibandwidth β ($\beta \ll N$).

The block WZ factorization for the block tridiagonal toeplitz-block-toeplitz (TBT) coefficient matrix \mathcal{A} is given. The coefficient matrix is decomposed as $\mathcal{A} = A_t + S$, such that

$A_t = W_t Z_t$, where the factors W_t and Z_t are block tridiagonal and are constructed from the block W and Z factors of \mathcal{A} , and S is a block matrix with only seven non-zero block entries concentrated on the middle. A direct parallel block WZ algorithm, named DPBWZA, for the solution of block tridiagonal TBT linear system $\mathcal{A}x = f$ is constructed based on the splitting and the block WZ factorization using divide and conquer technique. Existence of the block WZ factorization of the coefficient matrix \mathcal{A} is proved for block diagonally dominant matrices. The backward error analysis of the parallel algorithm DPBWZA is presented and numerical stability is established.

A Parallel Algorithm, named PBWZA, for the solution of strictly diagonally dominant block tridiagonal linear systems is designed. The linear system is partitioned along the main diagonal and in each partition, a sequential iterative scheme is used simultaneously. Thus, an iterative scheme is developed for the parallel solution of block tridiagonal linear systems. Convergence of the iterative scheme given for parallel algorithm is assured for strictly diagonally dominant block tridiagonal linear systems.

An incomplete WZ factorization (IWZ factorization) for the diagonally dominant banded sparse matrix is given. A hybrid parallel algorithm, called Incomplete WZ Parallel Solver ($IWZPS$), for the solution of large sparse linear systems is constructed. The sparse linear system $Ax = f$ is first reordered to form a narrow banded linear system which is sparse within the band. The reordered system is then partitioned and the given IWZ factorization with some modification is used to develop the parallel algorithm. The solution obtained is refined using outer iterations of a krylov subspace method.

All the above parallel algorithms are implemented.

सार

इस काम में, हम तेज़, सटीक, संख्यात्मक रूप से स्थिर और स्केलेबल समानांतर एल्गोरिदम विकसित करते हैं। तिरछे प्रमुख रेखीय प्रणालियों के समाधान के लिए इन रेखीय प्रणालियों को हल करना $Ax=f$ वैज्ञानिक और इंजीनियरिंग कम्प्यूटेशंस के क्षेत्र में केंद्रीय है। हम देते हैं घने तिरछे नॉनसिंग्युलर रेखीय प्रणालियों के समाधान के लिए एक WZ कारक फर्किकीकरण का अस्तित्व दिया जाता है। WZ की पिछली त्रुटि विश्लेषण फैक्टरिकी पद्धति दी जाती है और विधि की संख्यात्मक स्थिरता की स्थापना की जाती है।

तिरछे त्रिनोगोनात्मक गुणांक मैट्रिक्स के लिए वाइडब्ल्यूजे का कारक दिया जाता है। फैक्टिकेशन का अस्तित्व तिरछे प्रभावशाली नॉनसिंग्युलर मैट्रिक्स के लिए सिद्ध है। तिरंगी प्रमुख त्रिदोगोल रेखिक प्रणालियों के समाधान के लिए एक समानांतर एल्गोरिथ्म दिए गए डब्ल्यूडब्ल्यू फैक्टरिकेशन के आधार पर, छोटे संशोधन के साथ, डिवाइड का उपयोग करके विकसित किया गया है और तकनीक को जीत गुणांक मैट्रिक्स को मुख्य ब्लॉकों में विभाजित किया जाता है विकर्ण और सभी विभाजन एक युग्मन वैरिएबल के साथ एक दूसरे से संबंधित हैं। एक कम ऑर्डर $2r$ की व्यवस्था प्रत्येक विभाजन के युग्मन वैरिएबल से बनाई गई है। उपाय मूल रेखीय प्रणाली का कम प्रणाली के समाधान पर निर्भर करता है। यह किया गया है। यह साबित करता है कि कम प्रणाली में एक समान गुण होते हैं, जैसे किअलंकरण, त्रिदेगोनाल संरचना और तिरछे प्रभुत्व, जो कि मूल रेखिक प्रणाली के मुताबिक सुनिश्चित करता है। कम प्रणाली के समाधान के अस्तित्व समानांतर के पिछड़े त्रुटि विश्लेषण एल्गोरिथ्म दिया जाता है और संख्यात्मक स्थिरता साबित होती है। त्रिदेगोनाल के लिए इस्तेमाल किया जाने वाला विचार रेखीय प्रणाली को तिरछे प्रभावी संकीर्ण बैंड गुणांक मैट्रिक्स तक बढ़ाया जाता है सेमबैंडविड्थ β ($\beta \ll N$).

ब्लॉक त्रिदेगोनाल टीबीटी गुणांक के लिए ब्लॉक डब्ल्यूजेड कारक कारक मैट्रिक्स ए दिया जाता है गुणांक मैट्रिक्स को $E = E + ES$ के रूप में विघटित किया जाता है; ऐसा है कि $E = WtZt$ पर; जहां कारकों Wt और Zt ब्लॉक त्रिनोगोनात्मक हैं और से निर्माण कर रहे हैं। E के ब्लॉक डब्ल्यू और जेड कारक; और ES केवल सात गैर शून्य ब्लॉक प्रविष्टियों के साथ एक ब्लॉक मैट्रिक्स है। बीच पर केंद्रित एक प्रत्यक्ष समानांतर ब्लॉक डब्ल्यूपी एल्गोरिथ्म, जिसका नाम डीपीबीडब्ल्यूएजीए है, ब्लॉक त्रिदेगोनाल टीबीटी रेखिक प्रणाली के समाधान के लिए एक्स = एफ पर आधारित है। विभाजित और तकनीक को जीतने और जीतने के लिए ब्लॉक WZ फॉक्मेचरिंग। अस्तित्व गुणांक मैट्रिक्स E के ब्लॉक डब्ल्यूजेड फॉक्साइजेशन के ब्लॉक तिरछे के लिए साबित होता है। प्रमुख मैट्रिक्स समानांतर एल्गोरिथ्म DPBWZA के पिछड़े त्रुटि विश्लेषण है।

प्रस्तुत और संख्यात्मक स्थिरता की स्थापना की है। एक समानांतर एल्गोरिथ्म, जिसका नाम पीबीडब्ल्यूएजीए (PBWZA) है, सख्ती से तिरछे प्रभावशाली के समाधान के लिए ब्लॉक त्रिदोगोल

रैखिक सिस्टम तैयार किया गया है। रेखीय प्रणाली को साथ में विभाजित किया गया है। मुख्य विकर्ण और प्रत्येक विभाजन में, अनुक्रमिक पुनर्योजी योजना का उपयोग एक साथ किया जाता है। इस प्रकार, एक टिकाऊ योजना ब्लॉक त्रिदोशल रैखिक के समानांतर समाधान के लिए विकसित की गई है। सिस्टम समानांतर एल्गोरिथम के लिए दिया जाने वाली पुनरावृत्ति योजना के अभिसरण के लिए आश्वासन दिया गया है। सख्ती से तिरका हुआ प्रमुख ब्लॉक त्रिदोशल रैखिक प्रणाली तिरछे प्रभावी बैंडिंग के लिए एक अधूरे डब्ल्यूडब्ल्यू फैक्टरिकेशन (आईडब्ल्यूजेड कारकिकी) विरल मैट्रिक्स दिया जाता है। एक हाइब्रिड समानांतर एल्गोरिथम, जिसे अपरिवेड डब्लूपी समानांतर सॉलवर कहा जाता है।

(IWZPS), बड़े विरल रेखीय प्रणालियों के समाधान के लिए बनाया गया है। विरल रैखिक प्रणाली $Ax = f$ को पहले एक संकीर्ण बैंड रेखीय प्रणाली बनाने के लिए पुनः क्रमबद्ध किया जाता है जो विरल होता है। बैंड के भीतर फिर से क्रमबद्ध सिस्टम का विभाजन किया जाता है और दिए गए IWZ गुणक कुछ संशोधनों के साथ समानांतर एल्गोरिथम को विकसित करने के लिए उपयोग किया जाता है। समाधान प्राप्त किया। एक क्रिलॉव सबस्पेस विधि के बाहरी पुनरावृत्तियों का उपयोग करके परिष्कृत किया जाता है।

सभी उपरोक्त समानांतर एल्गोरिथम कार्यान्वित हैं।

Contents

Certificate	i
Acknowledgements	iii
Abstract	v
List of Figures	xi
List of Tables	xiii
List of Symbols	xv
1 Introduction and preliminaries	1
1.1 Era of high performance computing	2
1.1.1 GPU computing and CUDA	6
1.1.2 Principles of parallel algorithm design	7
1.2 Solving linear systems is important	9
1.3 Literature survey	11
1.3.1 Dense linear systems	11
1.3.2 Tridiagonal linear systems	14
1.3.3 Banded linear systems	17
1.3.4 Toeplitz-Block-Toeplitz linear systems	19
1.3.5 Block tridiagonal linear systems	21
1.3.6 Sparse linear systems	23
1.4 Layout of the thesis	25

2	Diagonally Dominant Linear System Solver	29
2.1	Present method	30
2.2	The algorithm	32
2.3	Existence theorems	34
2.4	Error analysis	40
2.4.1	Preliminaries	40
2.4.2	Error in WZ factorization	41
2.4.3	Error in solving the subsystem $\hat{W}y = f$	48
2.4.4	Error in solving the subsystem $\hat{Z}x = \hat{y}$	52
2.5	Numerical experiments	60
2.6	Conclusions	61
3	Parallel Algorithm for Diagonally Dominant Tridiagonal Linear Systems	63
3.1	The WZ factorization	64
3.2	Parallel solver using WZ factorization	66
3.2.1	Partitioning of tridiagonal linear system	66
3.2.2	The method	68
3.3	Algorithm and its complexity	70
3.3.1	Complexity of the algorithm	72
3.4	Existence theorems	72
3.5	Error analysis	80
3.5.1	Error in WZ factorization	80
3.5.2	Error in solving the subsystem $\hat{W}y = f$	90
3.5.3	Error in solving the reduced system $Rx_{red} = \hat{y}_{red}$	94
3.5.4	Error in solving the subsystem $\hat{Z}x = \hat{y}$	95
3.6	Numerical experiments and discussion	101
3.7	Conclusions	104
4	Parallel Algorithm for Diagonally Dominant Banded Linear Systems	107
4.1	The parallel WZ factorization method	109
4.1.1	Partitioning	109
4.1.2	Factorization	111
4.1.3	The method	113

4.2	The algorithm	116
4.3	Existence theorems	119
4.4	Error analysis	126
4.4.1	Error in the WZ factorization	127
4.4.2	Error in solving the subsystem $\hat{W}y = f$	131
4.4.3	Error in solving the subsystem $\hat{Z}x = \hat{y}$	137
4.4.4	Error in solving the reduced system $Rx_{red} = y_{red}$	138
4.5	Numerical experiments and discussion	143
4.5.1	First set of experiments	144
4.5.2	Second set of experiments	145
4.6	Conclusions	149
5	Parallel Algorithm for Block Tridiagonal Toeplitz-Block-Toeplitz Linear Systems	151
5.1	Sequential case	154
5.1.1	Block WZ factorization of block tridiagonal matrix \mathcal{A}	154
5.1.2	The solution process	155
5.2	Direct parallel block WZ factorization algorithm(DPBWZA)	157
5.2.1	Partitioning of linear system $\mathcal{A}x = f$	157
5.2.2	The parallel solution process	159
5.2.3	Algorithm DPBWZA	164
5.2.4	Complexity of the parallel algorithm DPBWZA	167
5.3	Existence theorems	168
5.4	Error analysis	172
5.4.1	Error in the factorization of \mathcal{A}	173
5.4.2	Error in solving the subsystem $\hat{W}_t Y_t = F$	175
5.4.3	Error in solving the subsystem $\hat{Z}_t X_t = \hat{Y}_t$	176
5.4.4	Error in solving the reduced system $R^{(p)} X_{t_{red}}^{(p)} = \hat{Y}_{t_{red}}^{(p)}$	178
5.4.5	Error in solving the block linear system $\mathcal{A}^{(p)} X_c^{(p)} = v^{(p)}$	182
5.5	Numerical experiments and discussion	183
5.6	Conclusions	189

6	Parallel Iterative GPU Solution for Block Tridiagonal Linear Systems	191
6.1	Iterative scheme for sequential case	193
6.1.1	The splitting of matrix \mathcal{A}	194
6.1.2	The iterative scheme	195
6.1.3	Convergence analysis	196
6.2	Parallel algorithm PBWZA	199
6.2.1	The parallel splitting of matrix \mathcal{A}	201
6.2.2	The parallel iterative scheme	203
6.2.3	Convergence analysis of parallel iterative scheme	205
6.3	Numerical experiments and discussion	206
6.4	Conclusions	212
7	Parallel Hybrid Algorithm for Large Sparse Linear Systems	215
7.1	The method	217
7.1.1	Reordering of the linear system	217
7.1.2	Partitioning of the linear system	217
7.1.3	WZ factorization of the sparse matrix	218
7.1.4	Construction of IWZPS	223
7.1.5	The IWZPS algorithm	225
7.2	Numerical experiments and discussion	226
7.3	Conclusions	231
	References	233
	Bio-Data	249

List of Figures

3.1	Time in seconds for parallel WZ algorithm against different number of processors in solving tridiagonal diagonally dominant linear system for different problem sizes.	103
3.2	Time in seconds for the Truncated SPIKE algorithm against different number of processors in solving tridiagonal diagonally dominant linear system for different problem sizes.	103
3.3	Time in seconds for parallel WZ algorithm and for the Truncated SPIKE algorithm against different number of processors in solving tridiagonal diagonally dominant linear system for different problem sizes.	104
4.1	Partitioning of the matrix A, x and f for $r = 4$	110
4.2	Speedup of the parallel WZ algorithm over ScaLAPACK for diagonally dominant banded linear system of order 720,000 with varying bandwidths using 32 processors.	146
4.3	Speedup of the parallel WZ algorithm over ScaLAPACK for diagonally dominant banded linear systems of varying orders with bandwidth 121 with increasing number of processors.	149
4.4	Execution times of “solve”(on left) and “factorization”(on right) stages of the parallel WZ algorithm and ScaLAPACK for diagonally dominant banded linear systems of order $N = 1,884,960$ with bandwidth 441 with increasing number of processors.	149

5.1	Execution times of the parallel algorithm DPBWZA over cuSolverSP for block tridiagonal toeplitz-block-toeplitz linear systems of various orders N with fixed block size q	187
5.2	Execution times of the parallel algorithm DPBWZA over cuSolverSP for block tridiagonal toeplitz-block-toeplitz linear systems of fixed order N with varying block size q	187
5.3	Speedup of the parallel algorithm DPBWZA over cuSolverSP for block tridiagonal toeplitz-block-toeplitz linear systems of various orders N with fixed block size q	188
5.4	Speedup of the parallel algorithm DPBWZA over cuSolverSP for block tridiagonal toeplitz-block-toeplitz linear systems of fixed order N with varying block size q	188
6.1	Execution times of the parallel algorithm PBWZA and cuSolverSP for solving block tridiagonal linear system of fixed order N with varying block size q on GPU.	211
6.2	Execution times of the parallel algorithm PBWZA and cuSolverSP for solving block tridiagonal linear system of fixed block size q with varying order N on GPU.	211
6.3	Speedup of the parallel algorithm PBWZA over cuSolverSP for block tridiagonal linear systems of (a) fixed problem size N with varying block size q and (b) fixed block size q with varying problem size N	212
7.1	Speedup of IWZPS, DDPS and MUMPS against different number of processors for different linear systems.	231

List of Tables

2.1	Comparison of theoretical and experimental error bounds on the solution of $Ax = f$	61
3.1	Complexity of the WZ algorithm.	72
3.2	Complexity of the Truncated SPIKE algorithm.	73
3.3	Execution times of the parallel WZ algorithm and the Truncated SPIKE algorithm, SPIKE-TA, for solving nonsingular diagonally dominant tridiagonal linear systems of varying orders N with different number of processors.	102
3.4	Execution times of the parallel WZ algorithm and the Truncated SPIKE algorithm, SPIKE-TA, for solving nonsingular diagonally dominant tridiagonal linear systems of varying orders N with different number of processors.	102
4.1	Execution times and speed improvements of the parallel WZ algorithm versus ScaLAPACK for diagonally dominant banded linear system of order $N = 720,000$ on 32 processors.	145
4.2	Speed improvements of the parallel WZ algorithm over ScaLAPACK for diagonally dominant banded linear systems of varying orders N and bandwidth 121 with increasing number of processors r	148
4.3	Execution times of the parallel WZ algorithm and ScaLAPACK for diagonally dominant banded linear system of order $N = 1,884,960$ and bandwidth 441 with increasing number of processors r	148

5.1	Hardware Specifications.	185
5.2	Speed improvements of the Direct Parallel Block WZ Algorithm, DPB- WZA, over cuSolverSP for solving block tridiagonal toeplitz-block-toeplitz linear systems of varying orders N and block size q on GPU.	186
5.3	Normwise forward and backward errors for the Direct Parallel Bloack WZ Algorithm, DPBWZA, and cuSolverSP.	188
6.1	Hardware Specifications.	207
6.2	Speed improvements of the Parallel Block WZ Algorithm, PBWZA, over cuSolverSP for solving block tridiagonal linear systems of varying orders N and block size q on GPU.	209
6.3	Normwise forward and backward errors for the Parallel Block WZ Algo- rithm, PBWZA, and cuSolverSP.	210
7.1	Linear systems from the University of Florida sparse matrix collection. N and nnz denote the order and the number of nonzeros of the matrix.	229
7.2	Execution times of the Incomplete WZ Parallel Solver, IWZPS, and PAR- DISO using one core.	229
7.3	Execution times of the Domain Decomposition Parallel Solver, DDPS, and PARDISO using one core.	229
7.4	Execution times of MULTifrontal Massively Parallel sparse direct Solver, MUMPS, and PARDISO using one core.	230
7.5	Number of outer BiCGStab iterations for the Incomplete WZ Parallel Solver, IWZPS, and ILUPACK.	230
7.6	Number of outer BiCGStab iterations for the Domain Decomposition Par- allel Solver and ILUPACK.	230

List of Symbols

Symbol	Meaning
A, B, C	represent matrices
$\mathcal{A}, A_t, W_t, Z_t, S$	represent block matrices
q	size of each block
W	represents factored matrix W
Z	represents factored matrix Z
x, x_t, x_c, f	represent vectors
x_i, f_i	represent subvectors
n, N	represent order of linear system
β	semibandwidth of banded matrix
(p)	superscript represents p th partition
T	superscript represents the transpose
κ	represents condition number
g	largest element in magnitude of a matrix
TBT	Toeplitz-Block-Toeplitz matrix
\mathbf{u}	unit roundoff of the machine
b	base of the machine
t	precision
\hat{c}	represents the floating point result of the computation c