

**STUDIES ON EFFICIENT ALGORITHMS FOR
CONVEX AND NON-CONVEX PROBLEMS
WITH APPLICATIONS IN SIGNAL
PROCESSING, COMMUNICATION, AND
MACHINE LEARNING**

R. Jyothi



**Centre for Applied Research in Electronics
Indian Institute of Technology Delhi
March, 2022**

© Indian Institute of Technology Delhi (IITD), New Delhi, 2022

**STUDIES ON EFFICIENT ALGORITHMS FOR
CONVEX AND NON-CONVEX PROBLEMS
WITH APPLICATIONS IN SIGNAL
PROCESSING, COMMUNICATION, AND
MACHINE LEARNING**

by

R. Jyothi

Centre for Applied Research in Electronics

Submitted

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

to the



Indian Institute of Technology Delhi

March, 2022

Dedicated to

My Parents

Certificate

This is to certify that the thesis entitled “**Studies on efficient algorithms for convex and non-convex problems with applications in signal processing, communication, and machine learning**” being submitted by **Ms. R. Jyothi** to the Centre for Applied Research in Electronics (CARE), Indian Institute of Technology Delhi, for the award of the degree of **Doctor of Philosophy** is the record of the bonafide research work carried out by her under our supervision. In our opinion, 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 either in part or in full to any other university or institute for the award of any degree or diploma.



(Prof. Prabhu Babu)

CARE

Indian Institute of Technology Delhi

Hauz Khas, New Delhi 110016

India



(Prof. Rajendar Bahl)

CARE

Indian Institute of Technology Delhi

Hauz Khas, New Delhi 110016

India

Acknowledgements

First and foremost, I would like to thank my guide, Prof. Prabhu Babu, for having faith in me and giving me an opportunity to work with him. I am extremely grateful to him for introducing me to the field of optimization. His thoughtful advice has helped me a lot over the last five years of study. I am also thankful to Prof. Rajendar Bahl, Prof. Arun Kumar, Prof. Ananjan Basu, and Prof. Saif Khan Mohammed for their suggestions during my work. I would like to thank all my friends at Signal Processing, CARE for supporting me during my study. I am also thankful to all my other IIT Delhi friends for making my stay at Delhi enjoyable. Especially, I would like to thank my best friend Rijul, who has always offered to help at any time. Finally, I would like to thank my mom, dad, and brother for their unconditional love and moral support.

R. Jyothi
R. Jyothi

Abstract

The efficiency of an optimization algorithm plays a crucial role in the performance of several engineering applications. In this thesis, the focus is on utilizing the Majorization Minimization (MM) framework to develop efficient iterative optimization algorithms. In particular, MM-based algorithms are proposed to solve convex, non-convex, and non-smooth problems that have applications in the field of signal processing, communication, and machine learning. Also, for some problems, the combination of the MM framework with other commonly used optimization algorithms, such as the Alternating Direction Method of Multipliers (ADMM), is explored. One of the motivations behind studying the MM framework is its flexibility in solving a broad spectrum of problems. For instance, apart from solving convex problems, it can solve problems whose cost functions are non-convex and non-smooth. Also, it can handle problems with non-convex constraints. Another key advantage of MM is that it helps in exploiting the problem structure to develop problem-specific optimization algorithms.

In the first part of the thesis, the focus is on developing efficient optimization algorithms for the problems encountered in the field of signal processing. In this regard, four different problems related to signal processing applications are considered - 1) Source localization, 2) Non-negative matrix factorization, 3) Atomic Norm Minimization and, 4) Toeplitz covariance matrix estimation problems. For the first problem, which is a non-convex and non-smooth problem, an optimization algorithm named SOLVIT is proposed. SOLVIT is an MM-based iterative algorithm which, unlike the state-of-the-art algorithms, tackles the problem exactly without any approximations. Next, the non-negative matrix factorization problem, which is a non-convex problem

with multiple optimization variables, is studied. By inspecting the structure of this problem, two different iterative optimization algorithms named INOM and PARINOM are presented. While the former optimization algorithm is based on the block MM framework - an extension of the MM technique employed to solve problems with several block variables, the latter optimization algorithm is based on the principle of the MM procedure. Subsequently, for the convex Atomic Norm Minimization and the non-convex Toeplitz covariance matrix estimation problems, iterative optimization algorithms named DYANOM and ATOM are proposed. To develop both the optimization algorithms, a hidden structure of the underlying optimization problem is exploited. Moreover, DYANOM is based on the integration of the MM framework with Dykstra's projection. Also, two versions of ATOM, based on the combination of MM with ADMM and MM with Dykstra's projection, are proposed. For each problem, numerical simulations are conducted, which indicate that the proposed optimization algorithms have superior performance when compared to the state-of-the-art algorithms.

The second part of the thesis deals with tackling two different optimization problems that have applications in the field of communication. The first problem is concerned with the construction of unimodular sequences with good autocorrelation properties. To attain such sequences, an iterative optimization algorithm named SLOPE is proposed that minimizes the maximum peak sidelobe level of a unimodular sequence. Therefore, SLOPE optimizes a non-convex minimax problem. SLOPE arrives at a stationary point of the problem by leveraging the adaptation of the MM framework for minimax problems coupled with the utilization of the Mirror Descent Algorithm (MDA). Later, SLOPE is extended to handle additional constraints such as the energy, peak-to-average-power ratio, and spectral constraints. Numerical simulations confirm that SLOPE can generate sequences of considerably longer lengths with lower peak sidelobe levels. Next, a different non-convex minimax problem is optimized to construct optimal frames. Such frames comprises of a set of unit norm vectors such that the maximum correlation among them is minimal. To solve the minimax problem, two iterative optimization algorithms named TELET and FLIP with different updating

schemes are proposed. While the former MM-based algorithm updates all the columns comprising the frame vectors at once, the latter block MM-based algorithm updates the frame in a column-wise fashion. In addition to utilizing the MM framework, TELET and FLIP borrow MDA and state-of-the-art linear programming optimization algorithms, respectively, to construct optimal frames. Numerical simulations reveal that TELET and FLIP can construct frames of longer lengths with lower coherence values.

In the third part of the thesis, the focus is on solving optimizations problems that have applications in the field of machine learning. Under this field, the maximum likelihood parameter estimation for the classical multinomial logistic regression classifier is considered. Although the problem is convex and differentiable, it does not enjoy a closed-form solution. To solve this problem, an MM-based iterative algorithm named PIANO is proposed, which can parallelly update each element of the optimization variable. Moreover, unlike the state-of-the-art algorithms, PIANO can also be easily modified to handle the parameter estimation problem for the sparse multinomial logistic regression classifier. Simulations are conducted to compare PIANO with the existing methods, and it is found that PIANO performs better than the existing methods in terms of speed of convergence. Then, the ℓ_p norm linear regression optimization problem, which has applications in semi-supervised learning, is studied. To tackle this problem, an iterative optimization algorithm named PROMPT is proposed based on the MM technique. Unlike the state-of-the-art algorithms, PROMPT can handle the underlying cost function for any value of p . Moreover, PROMPT updates each element of the optimization variable parallelly. Numerical simulations highlight that PROMPT performs better than the state-of-the-art algorithms in terms of speed of convergence.

सारांश

एक अनुकूलन एल्गोरिथ्म की दक्षता कई इंजीनियरिंग अनुप्रयोगों के प्रदर्शन में महत्वपूर्ण भूमिका निभाती है। इस थीसिस में, कुशल पुनरावृत्त अनुकूलन एल्गोरिदम विकसित करने के लिए मेजराइजेशन मिनिमाइजेशन (एमएम) ढांचे का उपयोग करने पर ध्यान केंद्रित किया गया है। विशेष रूप से, एमएम-आधारित एल्गोरिदम उत्तल, गैर-उत्तल, और . को हल करने के लिए प्रस्तावित हैं गैर-चिकनी समस्याएं जिनमें सिग्नल प्रोसेसिंग, संचार और मशीन सीखने के क्षेत्र में अनुप्रयोग हैं। इसके अलावा, कुछ समस्याओं के लिए, एमएम फ्रेमवर्क का संयोजन अन्य सामान्य रूप से उपयोग किए जाने वाले अनुकूलन एल्गोरिदम के साथ, जैसे कि अल्टरनेटिंग गुणक की दिशा विधि (ADMM) की खोज की गई है। एमएम ढांचे का अध्ययन करने के पीछे की प्रेरणाओं में से एक समस्याओं के व्यापक स्पेक्ट्रम को हल करने में इसकी लचीलापन है। उदाहरण के लिए, उत्तल समस्याओं को हल करने के अलावा, यह उन समस्याओं को भी हल कर सकता है जिनकी लागत कार्य गैर-उत्तल और गैर-चिकनी हैं। साथ ही, यह गैर-उत्तल बाधाओं के साथ समस्याओं को संभाल सकता है। MM का एक अन्य प्रमुख लाभ यह है कि यह समस्या-विशिष्ट अनुकूलन एल्गोरिदम विकसित करने के लिए समस्या संरचना का दोहन करने में मदद करता है।

थीसिस के पहले भाग में, सिग्नल प्रोसेसिंग के क्षेत्र में आने वाली समस्याओं के लिए कुशल अनुकूलन एल्गोरिदम विकसित करने पर ध्यान केंद्रित किया गया है। इस संबंध में, सिग्नल प्रोसेसिंग अनुप्रयोगों से संबंधित चार अलग-अलग समस्याओं पर विचार किया जाता है - 1) स्रोत स्थानीयकरण, 2) गैर-नकारात्मक मैट्रिक्स कारककरण, 3) परमाणु मानदंड न्यूनतमकरण, और, 4) Toeplitz सहप्रसरण मैट्रिक्स अनुमान समस्याएं। बाकी समस्या के लिए, जो एक गैर-उत्तल और गैर-चिकनी समस्या है, SOLVIT नामक एक अनुकूलन एल्गोरिथ्म प्रस्तावित है। SOLVIT एक MM-आधारित पुनरावृत्त एल्गोरिथ्म है, जो अत्याधुनिक एल्गोरिदम के विपरीत, बिना किसी अनुमान के समस्या से निपटता है। इसके बाद, गैर-ऋणात्मक मैट्रिक्स गुणनखंडन समस्या का अध्ययन किया जाता है, जो कई अनुकूलन चरों के साथ एक गैर-उत्तल समस्या है। इस समस्या की संरचना का निरीक्षण करके, INOM और PARINOM नामक दो अलग-अलग पुनरावृत्ति अनुकूलन एल्गोरिदम प्रस्तुत किए जाते हैं। जबकि पूर्व अनुकूलन एल्गोरिथ्म ब्लॉक MM ढांचे पर आधारित है - कई ब्लॉक चर के साथ समस्याओं को हल करने के लिए नियोजित MM तकनीक का एक विस्तार, बाद वाला अनुकूलन एल्गोरिथ्म MM प्रक्रिया के सिद्धांत पर आधारित है। इसके बाद, उत्तल परमाणु मानक न्यूनीकरण और गैर-उत्तल टोप्लिट्ज सहप्रसरण मैट्रिक्स आकलन समस्याओं के लिए, पुनरावृत्त अनुकूलन एल्गोरिदम DYANOM और ATOM नाम प्रस्तावित हैं। दोनों अनुकूलन एल्गोरिदम विकसित करने के लिए, अंतर्निहित अनुकूलन समस्या की एक छिपी संरचना का शोषण किया जाता है। इसके अलावा, DYANOM डायक्स्ट्रा के प्रक्षेपण के साथ MM ढांचे के एकीकरण पर आधारित है। साथ ही, ADMM के साथ MM के संयोजन के आधार पर ATOM के दो संस्करण और डायक्स्ट्रा के प्रक्षेपण के साथ एमएम प्रस्तावित हैं। प्रत्येक समस्या के लिए, संख्यात्मक सिमुलेशन आयोजित किए जाते हैं, जो इंगित करते हैं कि प्रस्तावित अनुकूलन एल्गोरिदम का अत्याधुनिक एल्गोरिदम की तुलना में बेहतर प्रदर्शन है।

थीसिस का दूसरा भाग दो अलग-अलग अनुकूलन समस्याओं से निपटने से संबंधित है जिनका संचार के क्षेत्र में अनुप्रयोग है। पहली समस्या अच्छे ऑटोसहसंबंध गुणों के साथ अनिमोड्यूलर अनुक्रमों के निर्माण से संबंधित है। ऐसे अनुक्रमों को प्राप्त करने के लिए, SLOPE नामक एक पुनरावृत्त अनुकूलन एल्गोरिथ्म प्रस्तावित है जो एक अनिमोड्यूलर अनुक्रम के अधिकतम शिखर साइडलॉब स्तर को कम करता है। इसलिए, SLOPE एक गैर-उत्तल न्यूनतम अधिकतम समस्या का अनुकूलन करता है। मिरर डिसेंट एल्गोरिथ्म के उपयोग के साथ मिलकर मिनिमैक्स समस्याओं के लिए एमएम फ्रेमवर्क के अनुकूलन का लाभ उठाकर स्लोप समस्या के एक स्थिर बिंदु पर आता है (एमडीए)। बाद में, SLOPE को अतिरिक्त बाधाओं जैसे कि ऊर्जा, शिखर-से-औसत-शक्ति अनुपात, और वर्णक्रमीय बाधाओं को संभालने के लिए बढ़ाया गया है। संख्यात्मक सिमुलेशन इस बात की पुष्टि करते हैं कि SLOPE निचली

चोटी के साथ काफी लंबी लंबाई के अनुक्रम उत्पन्न कर सकता है साइडलॉब स्तर। अगला, एक अलग गैर-उत्तल मिनीमैक्स समस्या को इष्टतम फ्रेम बनाने के लिए अनुकूलित किया गया है। इस तरह के फ्रेम में यूनिट नॉर्म वेक्टर का एक सेट होता है, जिसमें उनके बीच अधिकतम सहसंबंध न्यूनतम होता है। मिनीमैक्स समस्या को हल करने के लिए, अलग-अलग अद्यतन योजनाओं के साथ TELET और FLIP नामक दो पुनरावृत्त अनुकूलन एल्गोरिदम प्रस्तावित हैं। जबकि पूर्व MM-आधारित एल्गोरिथम फ्रेम वेक्टर वाले सभी कॉलम को एक साथ अपडेट करता है, बाद वाला ब्लॉक MM-आधारित एल्गोरिथम कॉलम-वार फैशन में फ्रेम को अपडेट करता है। MM ढांचे का उपयोग करने के अलावा, TELET और FLIP इष्टतम फ्रेम बनाने के लिए क्रमशः एमडीए और अत्याधुनिक रैखिक प्रोग्रामिंग अनुकूलन एल्गोरिदम उधार लेते हैं। संख्यात्मक सिमुलेशन से पता चलता है कि TELET और FLIP कम सुसंगतता मूल्यों के साथ लंबी लंबाई के फ्रेम का निर्माण कर सकते हैं।

थीसिस के तीसरे भाग में, उन अनुकूलन समस्याओं को हल करने पर ध्यान केंद्रित किया गया है जिनमें मशीन सीखने के पुराने समय में अनुप्रयोग हैं। इस क्षेत्र के तहत, क्लासिकल मल्टीनोमियल लॉजिस्टिक रिग्रेशन क्लासिफायरियर के लिए अधिकतम संभावना पैरामीटर अनुमान पर विचार किया जाता है। हालांकि समस्या उत्तल और अलग-अलग है, यह नहीं है एक बंद फॉर्म समाधान का आनंद लें। इस समस्या को हल करने के लिए, पियानो नामक एक एमएम-आधारित पुनरावृत्त एल्गोरिदम प्रस्तावित है, जो अनुकूलन चर के प्रत्येक तत्व को समानांतर रूप से अद्यतन कर सकता है। इसके अलावा, अत्याधुनिक एल्गोरिदम के विपरीत, विरल बहुपद के लिए पैरामीटर अनुमान समस्या को संभालने के लिए पियानो को आसानी से संशोधित किया जा सकता है लॉजिस्टिक रिग्रेशन क्लासी एर। मौजूदा तरीकों के साथ पियानो की तुलना करने के लिए सिमुलेशन आयोजित किए जाते हैं, और यह पाया जाता है कि पियानो अभिसरण की गति के मामले में मौजूदा तरीकों से बेहतर प्रदर्शन करता है। फिर, एलपी मानदंड रैखिक प्रतिगमन अनुकूलन समस्या, जिसमें अर्ध-पर्यवेक्षित शिक्षण में अनुप्रयोग हैं, का अध्ययन किया जाता है। इससे निपटने के लिए समस्या, MM तकनीक के आधार पर PROMPT नामक एक पुनरावृत्त अनुकूलन एल्गोरिथम प्रस्तावित है। अत्याधुनिक एल्गोरिदम के विपरीत, PROMPT p के किसी भी मूल्य के लिए अंतर्निहित लागत फ्रंक्शन को संभाल सकता है। इसके अलावा, PROMPT अनुकूलन चर के प्रत्येक तत्व को समानांतर रूप से अद्यतन करता है। संख्यात्मक सिमुलेशन इस बात पर प्रकाश डालते हैं कि PROMPT अभिसरण की गति के मामले में अत्याधुनिक एल्गोरिदम से बेहतर प्रदर्शन करता है।

Contents

Certificate	i
Acknowledgements	ii
Abstract	iii
List of Figures	xi
List of Tables	xvi
Abbreviations	xviii
Notations	xxi
1 Introduction	1
1.1 Motivation	1
1.2 Research contributions	22
1.3 Organization of the thesis	29
2 SOLVIT: Source Localization via an Iterative Technique	36
2.1 Introduction	36
2.2 Problem formulation	38
2.3 Algorithms for source localization problems based on MM	40
2.3.1 SOLVIT: SOurce Localization Via an ITERative technique	41
2.3.2 Comparison of SOLVIT with [97]	44
2.3.3 Proof of Convergence	47

2.3.4	Initialization of SOLVIT	48
2.3.5	Per-iteration computational complexity of SOLVIT.	48
2.3.6	Derivation of Standard Fixed Point algorithm using Majorization Minimization	49
2.4	Numerical simulations and experiments	52
2.5	Conclusion	60
3	Novel iterative algorithms for Nonnegative Matrix Factorization	62
3.1	Introduction	62
3.2	Problem Formulation	64
3.3	Iterative algorithms for NMF problem	65
3.3.1	INOM: Iterative NONnegative Matrix Factorization	65
3.3.2	PARINOM: Parallel Iterative NONnegative Matrix Factorization	70
3.3.3	Convergence Analysis	74
3.3.4	Accelerated PARINOM	76
3.4	Iterative algorithm for KL-NMF problem	77
3.5	Numerical Simulations	80
3.6	Conclusion	88
4	Dykstra’s projection based atomic norm solver	90
4.1	Introduction	90
4.2	Problem Formulation	92
4.3	DYANOM: DYkstra’s projection based Atomic NORm solver	95
4.4	Numerical simulations	100
4.5	Conclusions	103
5	New Methods for Maximum Likelihood Estimation of Toeplitz Struc- tured Covariance Matrices	105
5.1	Introduction	105
5.2	Problem Formulation	107
5.3	Algorithms for Toeplitz covariance matrix estimation	108

5.3.1	ATOM1	110
5.3.2	Computational complexity of ATOM1	113
5.3.3	ATOM 2	114
5.3.4	Computational complexity of ATOM2	118
5.3.5	Proof of convergence	118
5.3.6	Extensions of ATOM2	119
5.4	Cramer-Rao Lower bound calculation	125
5.4.1	Toeplitz matrix	125
5.4.2	Banded Toeplitz matrix	126
5.4.3	Toeplitz-block-Toeplitz matrix	127
5.5	Numerical Simulations	128
5.6	Conclusion	137
5.7	Appendix	138
6	SLOPE: A monotonic algorithm to design sequences with good auto-correlation properties by minimizing the peak sidelobe level.	142
6.1	Introduction	142
6.2	Problem Formulation	146
6.3	Monotonic minimizer for PSL metric	147
6.3.1	SLOPE: Sequence with LOW Peak sidelobe level	147
6.3.2	Mirror Descent Algorithm for problem in (6.33)	154
6.3.3	Convergence proof and computational complexity of SLOPE	155
6.3.4	Energy, PAPR and spectral constraints	157
6.3.5	Accelerated SLOPE	159
6.4	Numerical simulations and channel estimation applications	160
6.5	Conclusion	169
7	Iterative algorithms for the design of high-dimensional incoherent frames	170
7.1	Introduction	170
7.2	Problem formulation	173

7.3	TELET: MM based algorithm for constructing large dimensional incoherent frames	174
7.3.1	Majorization Function construction	175
7.3.2	Solution to the Minimization Problem	177
7.3.3	Proof of convergence	181
7.3.4	SQUAREM acceleration scheme	182
7.4	FLIP: Block Minorization Maximization based algorithm for constructing large dimensional incoherent frames	183
7.5	Numerical Simulations of TELET and CS Application	188
7.6	Numerical Simulations of FLIP	204
7.7	Conclusion	207
7.8	Appendix	208
8	PIANO: A Fast Parallel Iterative Algorithm for Multinomial and Sparse Multinomial Logistic Regression	211
8.1	Introduction	211
8.2	Problem Formulation	214
8.3	Proposed Algorithm for Multinomial Logistic Regression	216
8.3.1	PIANO: Parallel Iterative Algorithm for MultiNomial LOGistic Regression	216
8.3.2	Sparse Multinomial Logistic Regression with ℓ_1 regularization	221
8.3.3	Sparse Multinomial Logistic Regression with ℓ_0 regularization	225
8.3.4	Proof of Convergence for PIANO	226
8.4	Performance study: simulations and real-life data set	227
8.5	Conclusion	235
9	PROMPT: Parallel Iterative Algorithm for ℓ_p norm linear regression	236
9.1	Introduction	236
9.2	Problem Formulation	239
9.3	Proposed Algorithm for ℓ_p norm linear regression problem	240

9.3.1	PROMPT: Parallel Iterative Algorithm for ℓ_p norm regression via Majorization Minimization	240
9.3.2	Special Cases	244
9.3.3	Proof of convergence	247
9.3.4	Graph Based semi-supervised learning	248
9.4	Numerical simulation	250
9.5	Conclusion	257
10	Conclusions and Future scope	258
10.1	Conclusions	258
10.2	Future work	260

List of Figures

1.1	MM procedure	9
1.2	Block MM procedure	24
2.1	Comparison of the surrogate functions of SOLVIT and R-Means algorithm for problem dimension $n = 2$ and $m = 4$	47
2.2	Comparison of convergence rate of SOLVIT, SOLVIT-Acc, R-Means and R-Means-Acc for problem dimension $n = 2$ and $m = 10$	54
2.3	RMSE vs SNR for random array geometry for problem dimension $n = 2$ and $m = 6$	56
2.4	Representative figure of the test setup taken inside the anechoic chamber. The corner of the room was taken as the origin. A speaker emitting a sinusoidal signal was considered as the sound source. A linear array with four microphones (with a spacing of 20 cm between each sensor) was used to receive the source signal	57
2.5	RMSE vs reverberation time for problem dimension $n = 2$ and $m = 4$ for three different acoustic source signals.	60
2.6	RMSE vs SNR for problem dimension $n = 2$ and $m = 4$ for three different reverberation times for speech source signal	61
3.1	Comparison of run time and convergence of proposed algorithm with existing algorithms when $n = 100$, $m = 200$ and $r = 1$	81
3.2	Comparison of proposed algorithm with existing algorithms by varying r , $m = 50000$ and $n = 10000$ and when \mathbf{V} is a dense matrix.	82

3.3	Comparison of proposed algorithm with existing algorithms by varying r , $m = 50000$ and $n = 10000$ and when \mathbf{V} is a 70% sparse matrix . . .	83
3.4	Comparison of proposed algorithm with existing algorithms for varying m , $r = 100$ and $n = 1000$ when \mathbf{V} is dense matrix.	84
3.5	Comparison of proposed algorithm with existing algorithms for varying m , $r = 100$ and $n = 1000$ when \mathbf{V} is 70% sparse matrix.	85
3.6	(a)-(e) represents the noisy source signals; (f)-(h) represents some of the observed signals	86
3.7	Reconstructed signal by HALS algorithm	87
3.8	Reconstructed signal by INOM algorithm	87
3.9	Objective vs. run time of the proposed algorithm and MU algorithm proposed for the KL-NMF problem	88
4.1	Objective vs. iteration	101
4.2	MSE vs. SNR	101
4.3	MSE vs. N.	102
4.4	Execution time vs. N.	102
4.5	Comparison of spatial spectrum of DYANOM and IVDST for $K = 2$ sources.	104
5.1	On-grid frequencies scenario for $m = 6$ and $n = 460$. a) Negative log-likelihood $f(\mathbf{R})$ vs. outer iterations; b) $\log \mathbf{X} $ vs. outer iterations . . .	130
5.2	Off-grid frequencies scenario for $m = 6$ and $n = 460$. a) Negative log-likelihood $f(\mathbf{R})$ vs. outer iterations; b) $\log \mathbf{X} $ vs. outer iterations . . .	131
5.3	MSE vs. number of samples n for Toeplitz covariance matrix. a) on-grid frequencies; b) off-grid frequencies	133
5.4	MSE(\mathbf{r}) vs. number of samples n for banded Toeplitz covariance matrix	134
5.5	MSE(\mathbf{r}) vs. number of samples n for TBT covariance matrix	134
5.6	MSE(\mathbf{r}) vs. number of samples n for Toeplitz covariance matrix with condition number constraint. a) on-grid frequencies; b) off-grid frequencies	135

5.7	Average SINR vs θ in the presence of multiple jammers. a) $m = n$ b) $n = 2m$, and c) $n = 3m$	137
6.1	PSL value vs. Iteration of SLOPE, MM-PSL-adaptive algorithm and the CPM algorithm for sequence lengths 30 and 300.	162
6.2	PSL vs N for SLOPE, CPM and MM-PSL	163
6.3	Comparison of SLOPE, MM-PSL-adaptive, and CPM algorithm with respect to three different metrics - PSL value, CPU time, and the number of iterations for different values of sequence length N	164
6.4	Comparison of Correlation level of the sequence for length $N = 49$ and length $N = 100$ when the proposed algorithm, MM-PSL-adaptive and the CPM algorithm are initialized with a random unimodular sequence.	165
6.5	Comparison of Correlation level of the sequence for length $N = 49$ and length $N = 100$ when SLOPE algorithm, MM-PSL-adaptive and the CPM algorithm are initialized with Frank sequence.	166
6.6	Comparison of Correlation level of the sequence for length $N = 49$ and length $N = 100$ when the SLOPE algorithm, MM-PSL-adaptive and the CPM algorithm are initialized with Golomb sequence.	166
6.7	Autocorrelation and normalized power spectral density plots of sequences generated with and without spectral constraints. The spectral constraints are applied over the band $[f_l = 0.3142, f_u = 1.2566]$ (rads/sec).	167
6.8	(a) MMSE vs N and (b) MMSE vs Noise power for channel estimation using sequences generated via SLOPE, CPM and MM-PSL	169
7.1	Convergence plot for $d = 3$ and $N = 9$	189
7.2	Convergence plot for different values of d and N for real and complex ETF.	192
7.3	Convergence plot: $\max \mathbf{x}_i^H \mathbf{x}_j $ vs. iteration of complex frames of large dimensions	197
7.4	Convergence plot: $\max \mathbf{x}_i^H \mathbf{x}_j $ vs. iteration of real frames of large dimensions	202

7.5	MSE ($\hat{\mathbf{U}}$) vs. Sparsity K and measurement dimension d for Haar wavelet dictionary using BP reconstruction algorithm	203
7.6	(a) Original image (b) Reconstructed image by the proposed algorithm (c) Reconstructed image by the algorithm in [116]	204
7.7	(a) Mutual coherence value vs. Iteration (b) Percentage of elements within 2% of the attained mutual coherence value vs. Iteration for frame dimensions (1) 6×10 , (2) 19×26 and (2) 60×81	206
8.1	$f'_1(w)$ vs w	218
8.2	$f'_2(w)$ vs w	219
8.3	$f'_5(w)$ vs w	223
8.4	Comparison of convergence speeds of the proposed algorithm with the state-of-the-art algorithms	228
8.5	Comparison of average run time for different values of feature dimension d , class dimension n and number of samples n for the MLR case.	229
8.6	Comparison of convergence rate of the proposed algorithm with the state-of-the-art algorithms	230
8.7	Comparison of average run time for different values of feature dimension d , class dimension n and number of samples n for the sparse MLR case.	232
8.8	Monotonic convergence of PIANO for Sparse MLR with ℓ_0 regularization	232
9.1	Comparison of convergence of the proposed algorithm with the modified IRLS and the IRLS algorithm with \mathbf{x}^0 equal to the optimal point of the ℓ_2 norm regression problem.	251
9.2	Comparison of convergence of the proposed algorithm with the modified IRLS and the IRLS algorithm with \mathbf{x}^0 generated randomly from Normal distribution	252
9.3	Objective value vs run time of the proposed algorithm for different random initialization.	253
9.4	Comparison of average run time of the proposed algorithm with the modified IRLS algorithm for varying size of the observation vector \mathbf{y}	254

9.5	Objective value vs run time of the proposed and the IRLS algorithm for the ℓ_1 norm regression problem	255
9.6	Objective value vs run time of the proposed algorithm for the ℓ_∞ norm regression problem	255
9.7	Comparison of average run time of the proposed algorithm with the modified IRLS algorithm for simulated data.	256

List of Tables

1.1	Summary of commonly used optimization algorithms	8
1.2	Summary of the research contribution	28
2.1	Source position and the Estimation error by SOLVIT and Accelerated R-Means algorithm for the real-life experiment.	58
5.1	Comparison of the average run time (in seconds) of the iterative algorithms	131
6.1	Comparison of PSL value of SLOPE with the other state-of-the-art iterative algorithms.	163
6.2	Comparison of average run time (in seconds) of SLOPE with the other state-of-the-art iterative algorithms.	164
6.3	Comparison of average number of iterations of SLOPE with the other state-of-the-art iterative algorithms.	164
7.1	Comparison of Mutual coherence of complex frames of small and medium dimensions.	193
7.2	Comparison of Mutual coherence of large dimension complex frames constructed by different algorithms	194
7.3	Comparison of run time (seconds) of the algorithms to construct complex frames of small and medium dimensions.	195
7.4	Comparison of run time (seconds) of the proposed and state-of-the-art algorithms to construct large dimensional complex frames.	196
7.5	Comparison of Mutual coherence of real frames constructed by different algorithms for small and medium dimensions	198

7.6	Comparison of Mutual coherence of large dimension real frames constructed by different algorithms	199
7.7	Comparison of run time (seconds) of the proposed and state-of-the-art algorithms to construct real frames of small, medium and large dimensions	200
7.8	Comparison of run time (seconds) of the proposed and state-of-the-art algorithms to construct large dimension real frames.	201
7.9	Performance comparison of frames of different dimensions constructed by FLIP and ISPM.	207
7.10	Run time (in seconds) of FLIP and ISPM.	207
8.1	Parameter values used for comparison for the different experiments conducted for the MLR case.	228
8.2	Parameter values used for comparison for the different experiments conducted for the sparse MLR case.	231
8.3	Summary of the Data Set Statistics	233
8.4	Comparison of the algorithms for the motion capture data set	234
8.5	Comparison of the algorithms for the amazon commerce review data set	234
8.6	Comparison of the algorithms for the shall bidding data set	234
8.7	Comparison of the algorithms for DrivFace data set	235
9.1	Comparison of run time of algorithms in seconds for IRIS data set . . .	257
9.2	Comparison of run time of algorithms in seconds for SPECT data set .	257

Abbreviations

ADMM	Alternating Direction Method of Multipliers
ANM	Atomic Norm Minimization
APG	Accelerated Proximal Gradient
ATOM	Alternating projection based TOeplitz covariance Matrix
BCD	Block Coordinate Descent
BSS	Blind Source Separation
CAN	Cyclic Algorithm New
CDMA	Code Division Multiple Access
CE	Circulant Embedding
CRLB	Cramer-Rao Lower bound
CS	Compressed Sensing
DoA	Direction of Arrival
DYANOM	DYkstra's projection based Atomic NOrM solver
EM	Expectation Maximization
ETF	Equiangular Tight Frames
FA	Factor Analysis
FFT	Fast Fourier Transform
FLIP	Frame design using LInear Programming
GF	Grassmannian Frame
HALS	Hierarchical Alternating Least Square
INOM	Iterative Nonnegative Matrix Factorization
IRLS	Iterative Reweighted Least Squares
ISL	Integrated Sidelobe Level

ISPM	Incoherence via Shifted Power Method
KKT	Karush Kuhn Tucker
KL	Kullback-Leibler
LP	Linear Programming
MDA	Mirror Descent Algorithm
MIMO	Multiple Input Multiple Output
MM	Majorization Minimization
MLE	Maximum Likelihood Estimation
MLR	Multinomial Logistic Regression
MU	Multiplicative Update
NMF	Non-negative Matrix Factorization
PAPR	Peak-to-Average-Power Ratio
PARINOM	Parallel Iterative Nonnegative Matrix Factorization
PCA	Principal Component Analysis
PIANO	Parallel Iterative Algorithm for MultiNomial LOGistic Regression classifier
PROMPT	Parallel iterative algorithm for ℓ_P norm regression via Majorization Minimization
PSD	Positive Semi-Definite
PSL	Peak Sidelobe Level
POCS	Projection onto Convex Sets
RD-LS	Range-Differences by the Least-Square
R-LS	Range Least Squares
RMSE	Root Mean Square Error
RIR	Room Impulse Response
SCM	Sample Covariance Matrix
SDP	SemiDefinite Programming
SIDCO	Sequential Iterative Decorrelation by Convex Optimization
SFP	Standard Fixed Point
SLOPE	Sequence with LOW Peak sidelobe
SNR	Signal to Noise ratio
SOLVIT	Source Localization Via an Iterative technique

SVD	Singular Value Decomposition
SQUAREM	Squared Iterative method
TDOA	Time Difference of Arrival
TELET	TEchnique to devise Large dimensional Equiangular Tight frames
TOA	Time of Arrival
TSC	Toeplitz Structured Covariance
UNTF	Unit Norm Tight Frame

Notations

a	Scalar
\mathbf{a}	Vector
\mathbf{A}	Matrix
\mathbb{R}, \mathbb{C}	Set of real and complex numbers, respectively
$\mathbb{R}^{m \times 1}, \mathbb{C}^{m \times 1}$	Set of m dimensional vectors of real and complex numbers, respectively
$\mathbb{R}^{m \times m}, \mathbb{C}^{m \times m}$	Set of $m \times m$ matrices of real numbers and complex numbers, respectively
$\mathbb{H}^{m \times m}$	Set of $m \times m$ Hermitian matrices
\mathbf{R}^+	Set of non-negative real numbers.
$(\cdot)^T$	Transpose
$(\cdot)^*$	Complex conjugate
$(\cdot)^H$	Complex conjugate transpose
$(\cdot)^{-1}$	Inverse
\odot	Hadamard product
\otimes	Kronecker product
$ \mathbf{A} $	Determinant of matrix \mathbf{A}
$\text{Re}(\cdot)$	Real part
$\text{Tr}(\mathbf{A})$	Trace of matrix \mathbf{A}
$\text{Rank}(\mathbf{A})$	Rank of matrix \mathbf{A} .
$\text{Vec}(\mathbf{A})$	Vector obtained by stacking the columns of \mathbf{A}
\mathbf{I}_N	Identity matrix of size N
∇f	Gradient of a function f
$\nabla^2 f$	Hessian of a function f .
$\mathbf{A} \succeq 0$	\mathbf{A} is a PSD matrix

$\mathbf{A} \succ 0$	\mathbf{A} is a positive definite matrix
$\mathbf{A} - \mathbf{B} \succeq 0$	Matrix $\mathbf{A} - \mathbf{B}$ is PSD
$\mathbf{A} - \mathbf{B} \succ 0$	Matrix $\mathbf{A} - \mathbf{B}$ is positive definite
$\lambda_{max}(\mathbf{A})$	Maximum eigenvalue of matrix \mathbf{A}
$\lambda_{min}(\mathbf{A})$	Minimum eigenvalue of matrix \mathbf{A}
$ \mathbf{a} $	Element-wise absolute of the vector \mathbf{a}
$\ \mathbf{a}\ _2$	Euclidean norm of the vector \mathbf{a}
$\ \mathbf{a}\ _1$	ℓ_1 norm of the vector \mathbf{a}
$\ \mathbf{a}\ _p$	ℓ_p norm of the vector \mathbf{a}
$\ \mathbf{A}\ _F$	Frobenious norm of matrix \mathbf{A}
$\mathbf{E}[\cdot]$	Statistical expectation
$\max(\mathbf{A}, \mathbf{B})$	Element wise maximum between \mathbf{A} and \mathbf{B} .
\circ	Element wise multiplication.
\oslash	Element wise division.