Design of sparse systems based on optimization methods

Jurišić Bellotti, Maja

University of Zagreb

FACULTY OF ELECTRICAL ENGINEERING AND COMPUTING

Maja Jurišić Bellotti

DESIGN OF SPARSE SYSTEMS BASED ON OPTIMIZATION METHODS

DOCTORAL THESIS

Zagreb, 2023

University of Zagreb
FACULTY OF ELECTRICAL ENGINEERING AND COMPUTING

Maja Jurišić Bellotti

DESIGN OF SPARSE SYSTEMS BASED ON OPTIMIZATION METHODS

DOCTORAL THESIS

Supervisor:
Professor Mladen Vučić, PhD

Zagreb, 2023

Sveučilište u Zagrebu
FAKULTET ELEKTROTEHNIKE I RAČUNARSTVA

Maja Jurišić Bellotti

DIZAJN RIJETKIH SUSTAVA TEMELJEN NA OPTIMIZACIJSKIM POSTUPCIMA

DOKTORSKI RAD

Mentor:
Prof. dr. sc. Mladen Vučić

Zagreb, 2023.

Doctoral thesis was written at the University of Zagreb Faculty of Electrical Engineering and Computing, Department of Electronic Systems and Information Processing.

The thesis was supported by the Croatian Science Foundation under the Project IP-2019-04-4189 Efficient Signal Processing Systems for Software Defined Radio.

Supervisor: Professor Mladen Vučić, PhD

PhD thesis contains 112 pages.
Thesis no.: _____

About the supervisor

Mladen Vučić was born in Karlovac in 1965. He received BSc, MSc and PhD degrees in electrical engineering from the University of Zagreb, Faculty of Electrical Engineering and Computing (FER), Zagreb, Croatia, in 1989, 1993 and 1999, respectively.

From March 1989 he has been working at the Department of Electronic Systems and Information processing at FER. In 2001 he was promoted to an Assistant Professor, in 2006 to an Associate Professor, in 2011 to a Professor and in 2016 to a Full Professor. He led three scientific and one research project, and participated in four other projects funded by the Ministry of Science and Technology of the Republic of Croatia. In addition, he participated in one EU FP7 project. Currently, he leads the project *Efficient Signal Processing Systems for Software Defined Radio* funded by Croatian Science Foundation and participates as a researcher in the project *DATACROSS - Advanced Methods and Technologies for Data Science and Cooperative Systems* funded by EU Structural and Investment Funds. He published more than 60 papers in journals and conference proceedings in the area of circuit theory, analog and digital signal processing, optimization theory and applications, digital system design, and embedded systems.

Prof. Vučić is a member of IEEE and KoREMA. From 2013 to 2016 he was serving as the chair of *IEEE Circuits and Systems Chapter Croatia*. From 2016 to 2018 he was Head of the Department of Electronic Systems and Information Processing at FER. In 1997, he was awarded from the Ministry of Defence and the Ministry of Science and Technology of the Republic of Croatia by the *Annual award for scientific contribution to the development and strengthening of the defence system of the Republic of Croatia*. In 2019, he received the *Fran Bošnjaković award* from the University of Zagreb.

O mentoru

Mladen Vučić rođen je u Karlovcu 1965. godine. Diplomirao je, magistrirao i doktorirao u polju elektrotehnike na Sveučilištu u Zagrebu, Fakultetu elektrotehnike i računarstva (FER), 1989., 1993. odnosno 1999. godine.

Od ožujka 1989. godine radi na Zavodu za elektroničke sustave i obradbu informacija FER-a. Godine 2001. izabran je u znanstveno-nastavno zvanje docenta, 2006. u zvanje izvanrednog profesora, 2011. u zvanje redovitog profesora, a 2016. u zvanje redovitog profesora u trajnom zvanju. Dosad je vodio tri znanstvenoistraživačka i jedan tehnologijski istraživačko razvojni projekt, te je sudjelovao na još četiri znanstvenoistraživačka projekta Ministarstva znanosti i tehnologije Republike Hrvatske. Također, bio je istraživač na jednom EU FP7 projektu. Trenutno vodi projekt *Učinkoviti sustavi za obradu signala namijenjeni programski definiranom radiju* financiran od Hrvatske zaklade za znanost, a istraživač je na projektu *DATACROSS - Napredne metode i tehnologije u znanosti o podacima i kooperativnim sustavima* financiranom iz Europskih strukturnih i investicijskih fondova. Objavio je više od 60 radova u časopisima i zbornicima konferencija u području teorije električnih krugova, analogne i digitalne obrade signala, teorije i primjene optimizacijskih postupaka, dizajna digitalnih sustava te ugradbenih računalnih sustava.

Prof. Vučić član je udruga IEEE i KoREMA. Od 2013. do 2016. godine bio je predsjednik Odjela za električne krugove i sustave Hrvatske sekcije IEEE. Od 2016. do 2018. godine bio je predstojnik Zavoda za elektroničke sustave i obradbu informacija FER-a. 1997. godine dobio je *Godišnju nagradu za sveukupne znanstveno-istraživačke doprinose razvoju i jačanju sustava obrane Republike Hrvatske*, koju su zajednički dodijelili Ministarstvo obrane i Ministarstvo znanosti i tehnologije Republike Hrvatske. 2019. godine dobio je *Nagradu Fran Bošnjaković* koju je dodijelilo Sveučilište u Zagrebu.

Acknowledgement

I want to express my greatest gratitude to supervisor Mladen whose constant energy, knowledge and creativity kept me pushing towards and gave me ideas and motivation to overcome all obstacles on the way.

As all stories have their start, mine starts with parents Ankica and Velimir, without whose support and belief in me I would not be where I am today, and for that, I am forever grateful.

It was not easy to quit a prosperous job to become a doctoral student and pursue my dreams. However, my husband, Davor gave me the courage and strength to make this jump. I am deeply indebted to his love and support that kept me going even in the hardest times.

Special thank goes to my son Filip who didn't make things easier but made them more challenging and fun.

In the end, thanks to the neighbor's rooster whose aperiodic crowing kept me awake when Filip was sleeping so that I could have time to finish this PhD.

Abstract

Efficient implementation of electric filters is important because it reduces hardware complexity and decreases its power consumption. The most demanding component from both of these aspects is multiplier, which is used to implement coefficients of filter's transfer function. Consequently, efficient systems are often designed to minimize number of multipliers or to simplify their structures. The former approach leads to systems with sparse coefficients, which are also known as sparse systems. The latter is common in the design of spatial filters, in which high operating frequencies impose additional requirements such as the constraint for low dynamic range ratio of excitation coefficients. Numeric optimization is a common tool used in such designs. However, optimization problems appearing in these cases are nonconvex and therefore difficult to solve.

In this dissertation, methods for the design of both − sparse systems and systems with constrained dynamic range ratio − are considered. In particular, a method for the design of sparse FIR filters constrained in peak-error sense is presented. The method is based on signomial programming which utilizes the l_p norm with $0 < p < 1$. It is applied in the design of linear phase filters and filters without phase specifications. Furthermore, a method for the design of sparse linear phase FIR filters based on global optimization is described. This method utilizes a branch and bound algorithm with efficient tree pruning. It is suitable for the design of filters constrained in peak- or quadratic-error sense. In the area of spatial filter design, two methods are developed. The first of them utilizes branch and bound algorithm to obtain pencil beams with constrained dynamic range ratio of excitation coefficients. The second extends this design by allowing the coefficients to take zero values. Such an approach leads to sparse design, which enables more design freedom and, consequently, improves the obtained radiation patterns.

Keywords: branch and bound, compressed sensing, convex optimization, dynamic range ratio, global optimization, l_0-norm, l_p-norm, FIR filter, minimax, peak-error, pencil beam, quadratic-error, signomial programming, sparsity, spatial filter.

Dizajn rijetkih sustava temeljen na optimizacijskim postupcima

Električki filtri predstavljeni su prijenosnim funkcijama. Ove funkcije opisane su koeficijentima koji obuhvaćaju određen, najčešće širok, dinamički raspon. Prilikom realizacije prijenosnih funkcija, spomenuti koeficijenti predstavljeni su množilima. Implementacija množila zahtijeva znatne sklopovske resurse. Stoga se učinkoviti sustavi često dizajniraju tako da minimiziraju broj množila ili da pojednostavne njihovu strukturu. Prvi pristup vodi na sustave s rijetkim koeficijentima, poznate i pod nazivom rijetki sustavi. Drugi pristup koristi se u dizajnu prostornih filtara, kod kojih rad na visokim frekvencijama nameće dodatne zahtjeve, kao što je zahtjev za malim dinamičkim rasponom pobudnih koeficijenata.

Jedan pravac u dizajnu učinkovitih sustava temelji se na numeričkoj optimizaciji. No, pripadajući optimizacijski problemi često su nekonveksni i stoga teško rješivi. Dosad su razvijene mnoge iterativne metode koje se učinkovito nose s ovim problemima. Međutim, one kao rezultat daju lokala rješenja. S druge strane, razmatrane su i metode za globalnu optimizaciju. Iako one rezultiraju globalnim optimumom, mnogo su kompleksnije i zahtijevaju dulje vrijeme izvršavanja.

U ovoj disertaciji razmatrane su metode dobivanja rijetkih filtara te filtara s ograničenim dinamičkim rasponom pobudnih koeficijenata. Istraživanje je usmjereno na dizajn rijetkih filtara s konačnim impulsnim odzivom (FIR) te na dizajn učinkovitih prostornih filtara s uskim snopom zračenja.

Izvorno, rijetki sustavi dobivaju se minimizacijom l_0-norme. No zbog kombinatorne složenosti optimizacijskog problema koji sadrži spomenutu normu, često se uvode njene relaksacije. U disertaciji je razvijena metoda za dizajn rijetkih FIR filtara koja koristi relaksaciju l_0-norme l_p-normom gdje je $0 < p < 1$. Za rješavanje tako dobivenog optimizacijskog problema predloženo je signomijalno programiranje. Metoda minimizira maksimalno odstupanje amplitudne karakteristike filtra od specifikacija danih u području propuštanja i području gušenja, a primjenjiva je na filtre s linearnom fazom te na filtre bez zahtjeva na faznu karakteristiku.

U disertaciji su razmatrane i metode za globalnu optimizaciju rijetkih FIR filtara s linearnom fazom. Predložena je metoda koja pretražuje stablo sa svim položajima nultih koeficijenata filtra, a koristi grananje i ograničavanje s učinkovitim odsijecanjem određenih grana. Pretraživanje stabla izvedeno je kretanjem po okomitim čvorovima, kao i kombinacijom kretanja po okomitim i vodoravnim čvorovima. Predložena metoda pogodna je

za dizajn filtara čija je amplitudna karakteristika ograničena u smislu maksimalnog ili kvadratnog odstupanja.

U domeni prostornih filtara, predložena je metoda za dizajn antenskih nizova s uskim snopom zračenja i ograničenim dinamičkim rasponom pobudnih koeficijenata. Ovi nizovi su optimirani u smislu najvećeg gušenja bočnih latica. Metoda je temeljena na algoritmu grananja i ograničavanja. Konačno, ova metoda proširena je dopuštanjem određenim koeficijentima da poprime vrijednost nula čime se dobivaju bolji dijagrami zračenja za isti dinamički raspon.

Ova doktorska disertacija je podijeljena u osam poglavlja.

Prvo poglavlje opisuje problematiku istraživanja i daje širi kontekst područja u koje je ono smješteno. Na temelju toga, dana je motivacija za razvoj metoda predloženih u disertaciji. Naglasak je stavljen na dizajn rijetkih digitalnih filtara te učinkovitih prostornih filtara. Na kraju, opisana je struktura disertacije kroz pregled narednih poglavlja i kratak opis građe koje ta poglavlja sadrže.

U drugom poglavlju, dan je opis sustava koji će biti opisani u disertaciji. Pritom su posebno razmatrani FIR filtri te prostorni filtri. Navedeni su optimizacijski problemi koji se koriste u klasičnom dizajnu ovih sustava. Također, razmatrane su tehnike za povećanje njihove učinkovitosti. Pritom je kod FIR filtara razmatrano maskiranje frekvencija, opisivanje koeficijenata filtara pomoću sume potencija broja dva te dobivanje rijetkih koeficijenata. U području dizajna prostornih filtara, kao elementi učinkovitosti prepoznati su ograničen dinamički raspon koeficijenata i rijetkost. Mnoge metode za dizajn rijetkih sustava inspiraciju su pronašle u teoriji sažetog očitavanja, gdje se rijetkost postiže korištenjem l_1-norme. Stoga je jedan dio teksta posvećen pregledu ovog područja. Pokazano je kakvo značenje l_1-norma ima u rekonstrukciji rijetkih signala te zašto nije izravno primjenjiva u dizajnu rijetkih filtara. Na kraju, dan je pregled postojećih metoda za dizajn rijetkih FIR filtara te metoda za dizajn učinkovitih antenskih polja s ograničenim dinamičkim rasponom pobudnih koeficijenata.

S obzirom da metode predložene u disertaciji koriste optimizaciju, u trećem poglavlju dana je terminologija vezana uz ovo područje, Nadalje, opisani su postupci za konveksnu optimizaciju i to linearno programiranje, kvadratno programiranje, optimizacija konveksne funkcije nad prostorom omeđenim stošcima drugog reda te geometrijsko programiranje. Od metoda koje ne provode konveksnu optimizaciju opisano je signomijalno programiranje. Na kraju, opisana je metoda grananja i ograničavanja, koja pripada globalnoj optimizaciji. Kod ove metode razmatrani su razni pristupi pretraživanju stabla.

U četvrtom poglavlju opisana je metoda za dizajn rijetkih FIR filtara koja koristi aproksimaciju l_0-norme l_p-normom gdje je $0 < p < 1$. Optimizacijski problem koji opisuje ovakav dizajn nije konveksan. Stoga je za njegovo rješavanje potrebno odabrati neku od iterativnih metoda koja nije osjetljiva na početnu točku. U disertaciji je za rješavanje spomenutog problema predloženo signomijalno programiranje. Iako ovo programiranje ne osigurava globalnost dobivenog rješenja, moderni postupci za rješavanje signomijalnih programa u mnogim slučajevima konvergiraju ka njemu. U disertaciji je odabran postupak koji pretvara originalni problem u slijed geometrijskih problema koji se rješavaju u svakom koraku dok se ne postigne željena točnost. Opisana metoda je primijenjena na FIR filtre s linearnom fazom te na filtre bez zahtjeva na faznu karakteristiku. Amplitudna karakteristika optimirana je u smislu najmanjeg maksimalnog odstupanja u području propuštanja i području gušenja. Na većem broju primjera pokazano je da predložena metoda daje filtre čija je rijetkost bolja ili jednaka onoj koju daju druge suvremene metode.

Većina postojećih metoda za dizajn rijetkih filtara temelje se na lokalnoj optimizaciji ili na heurističkim pristupima. Stoga one ne mogu garantirati globalnost predloženog rješenja. Poznavanje globalnog rješenja važno je s aspekta samog dizajna, ali i s aspekta procjene kvalitete postojećih postupaka. U petom poglavlju razmatrana je globalna optimizacija rijetkih FIR filtara s linearnom fazom. Položaj nultih koeficijenata u impulsom odzivu takvih filtara nije unaprijed poznat. Stoga je pripadajući optimizacijski problem kombinatoran. Jedan način za njegovo rješavanje je iscrpno pretraživanje svih kombinacija koeficijenata koji su jednaki nuli. Međutim, broj kombinacija koje je potrebno pretražiti raste eksponencijalno s redom filtra. Stoga je iscrpno pretraživanje moguće provesti samo za filtre niskog reda. Jedan od načina za ubrzavanje pretraživanja je korištenje algoritama s grananjem i ograničavanjem. Kod takvih algoritama formira se stablo koje se zatim pretražuje. Ubrzanje u odnosu na iscrpno pretraživanje postiže se odsijecanjem pojedinih grana, koje se provodi na temelju dodatnih kriterija. U ovom poglavlju, predložena je metoda koja za pretraživanje stabla koristi kombinaciju ispitivanja vodoravnih i okomitih čvorova. Odsijecanje se provodi na temelju analize izvedivosti optimizacijskog problema. Metoda koristi i predznanje o očekivanoj rijetkosti, pri čemu je moguće krenuti od broja koeficijenata različitih od nule poznatog filtra koji zadovoljava dane specifikacije. Metoda je oblikovana za dizajn FIR filtara s linearnom fazom kod kojih je amplitudna karakteristika ograničena u smislu najmanjeg maksimalnog ili najmanjeg kvadratnog odstupanja. U disertaciji je dan veći broj primjera dizajna FIR filtara niskog i srednjeg reda.

Dizajn FIR filtara ima mnogo dodirnih točaka s dizajnom linearnih antenskih polja koja imaju jednak razmak između elemenata. Stoga su mnoge metode za dizajn FIR filtara inspiraciju pronašle u dizajnu antenskih polja i obrnuto. Međutim, kod korištenja ove sličnosti, potrebno je uvažiti specifičnosti koje diktira promatrani dizajn. U šestom poglavlju, metoda grananja i ograničavanja primijenjena je u dizajnu antenskih polja s uskim snopom zračenja i ograničenim dinamičkim rasponom pobudnih koeficijenata. Uvođenjem ograničenja na dinamički raspon koeficijenata problem postaje nekonveksan. U disertaciji je pokazano da je problem moguće napisati u konveksnom obliku ako su poznati predznaci svih koeficijenata. No, s obzirom da ti predznaci u nekom dizajnu nisu poznati unaprijed, cjelovita analiza zahtijeva iscrpno pretraživanje svih njihovih kombinacija. Ovaj problem riješen je pomoću metode grananja i ograničavanja koja odluku o odsijecanju grana donosi na temelju rješavanja konveksnog problema dobivenog relaksacijom polaznog problema. Ova metoda primijenjena je na dizajn antenskih polja s uskim snopom zračenja koja osiguravaju maksimalno gušenje bočnih latica. Metoda omogućuje dizajn nizova s pozitivnim i negativnim koeficijentima. Također, kao poseban slučaj, omogućava i globalnu optimizaciju nizova čiji su koeficijenti jednaki 1 i –1.

U dizajnu koji ne ograničava dinamički raspon pobudnih koeficijenata, taj raspon može poprimiti vrlo velike vrijednosti. Posebno je to slučaj kod nizova sa strmim bokom glavne latice. Jasno je da velik dinamički raspon nužno znači da neki koeficijenti poprimaju male apsolutne vrijednosti. S druge strane, ograničavanje raspona potiskuje te elemente u područje većih vrijednosti, te tako pogoršava dijagram zračenja. Ovo pogoršanje može se umanjiti ako se koeficijentima dozvoli da poprime vrijednost nula. S obzirom da takvi koeficijenti ne sudjeluju u određivanju dinamičkog raspona, njihova prisutnost smanjuje broj ograničenja u optimizacijskom problemu i time otvara prostor za poboljšanje dijagrama zračenja. Takav pristup vodi na rijedak dizajn. U sedmom poglavlju opisana je metoda za dizajn rijetkih antenskih polja s ograničenim dinamičkim rasponom koeficijenata i maksimalnim gušenjem u području bočnih latica. Spomenuta metoda koristi algoritam s grananjem i ograničavanjem koji pretražuje stablo sačinjeno od pozitivnih, negativnih i nultih koeficijenata. Opravdanost kombiniranja rijetkosti i ograničavanja dinamičkog raspona koeficijenata pokazana je na većem broju primjera.

Osmo poglavlje sažima značajke predloženih metoda i doprinos disertacije.

U okviru ove disertacije ostvaren je znanstveni doprinos koji se očituje u tri dijela. Prvi dio sačinjava metoda za dizajn rijetkih FIR filtara temeljena na signomijalnom programiranju. Ovdje je optimizacijski problem aproksimiran relaksiranjem originalnog l_0 problema l_p

problemom gdje je $0 < p < 1$. Problem je riješen formiranjem niza konveksnih podproblema koji se iterativno rješavaju. Iako je metoda lokalna i iterativna, daje dobre rezultate i primjenjiva je na filtre niskih i srednjih redova. Dani su primjeri dizajna filtara s linearnom fazom i bez zahtjeva na faznu karakteristiku, čija je amplitudna karakteristika optimalna u minimax smislu.

Drugi dio znanstvenog doprinosa predstavlja metoda za globalnu optimizaciju rijetkih FIR filtara s linearnom fazom. S obzirom da iterativne metode ne mogu garantirati globalnost dobivenog rješenja, razvijena je metoda koja koristi grananje i ograničavanje s učinkovitim odsijecanjem pojedinih grana stabla. Metoda je oblikovana za dizajn filtara optimalnih u smislu najmanje maksimalne ili najmanje kvadratne pogreške. Metoda je pogodna za optimizaciju filtara niskog i srednjeg reda u prihvatljivom vremenu. Nadalje, pogodna je za procjenu kvalitete postojećih postupaka jer može pokazati koliko su njihova rješenja blizu globalnim optimumima.

Primjena globalne optimizacije u dizajnu prostornih filtara rezultirala je trećim dijelom doprinosa disertacije. Razvijena je metoda za dizajn rijetkih i ne-rijetkih antenskih polja s uskim snopom zračenja, ograničenim dinamičkim rasponom koeficijenata i maksimalnim gušenjem bočnih latica. Metoda se temelji na algoritmu grananja i ograničavanja. Nadalje, pokazano je da se omogućavanjem rijetkosti u navedenim sustavima postižu bolji dijagrami zračenja.

Ključni pojmovi: uska zraka, dinamički raspon koeficijenata, FIR filtar, globalna optimizacija, konveksna optimizacija, l_0-norma, l_p-norma, grananje i ograničavanje, maksimalna pogreška, minimax, kvadratna pogreška, prostorni filtar, rijetkost, sažeto očitavanje, signomijalno programiranje.

Contents

1 Introduction 1
2 Efficient Systems 3
2.1 Electric Filters 3
2.1.1 Digital Filters 3
2.1.2 Antenna Arrays 8
2.2 Efficient Electric Filters 13
2.3 Sparsity in Filter Design 14
2.3.1 Compressed Sensing 14
2.3.2 Reconstruction of Sparse Signals 14
2.3.3 Application of L_1-Norm in Compressed Sensing 16
2.3.4 Application of L_1-Norm in Filter Design 17
2.4 State-of-the-Art in Sparse FIR Filter Design 18
2.4.1 Design Based on Approximate Solving of Greedy Algorithms 18
2.4.2 Design Based on Minimization of L_p-Norm 19
2.4.3 Design Based on Branch and Bound Algorithm 20
2.4.4 Design Based on Heuristic Algorithms 21
2.4.5 Joint Optimization of Sparsity and Filter Order 21
2.5 Efficient Spatial Filters 22
2.5.1 Pencil Beams With Low Dynamic Range Ratio 23
2.5.2 Sparse Pencil Beams 24
2.5.3 Simultaneous Application of Dynamic Range Ratio and Sparsity 25
3 Optimization Methods 26
3.1 Linear Programming 27
3.2 Quadratic Programming 27
3.3 Second-Order Cone Programming 28
3.4 Geometric and Signomial Programming 28
3.5 Branch and Bound 29
4 Design of Sparse FIR Filters Based on Signomial Programming 33
4.1 Arithmetic-Geometric Mean Approximation 33
4.2 Design of Sparse FIR Filters With Linear Phase 34
4.2.1 Design of Filters With Peak-Error Constraints 34
4.2.2 Design Examples 41
4.3 Design of Sparse FIR Filters Without Phase Specifications 49
4.3.1 Design of Filters With Peak-Error Constraints 49
4.3.2 Design Examples 53

5 Design of Sparse FIR Filters Based on Global Optimization 59
5.1 Problem Formulation 59
5.2 Feasibility Test 61
5.2.1 Feasibility Test for FIR Filters With Peak-Error Constraints 61
5.2.2 Feasibility Test for FIR Filters With Quadratic Constraints 62
5.2.3 Feasibility Test for FIR Filters With Peak- and Quadratic- Error Constraints 64
5.3 Branch and Bound Method for Sparse Filter Design 65
5.3.1 Initial Sparsity 65
5.3.2 Choosing Search Strategy 66
5.4 Additional Refinement of Approximation Error 69
5.5 Properties of Proposed Methods 70
5.5.1 Subproblem Analysis and Execution Time Comparison 70
5.5.2 Comparison With Exhaustive Search 72
5.6 Design Examples 73
5.6.1 Lowpass Filters With Peak-Error Constraints 73
5.6.2 Lowpass Filters With Quadratic Constraints 79
5.6.3 Lowpass Filters With Peak-Error and Quadratic Constraints 82
6 Global Optimization of Spatial Filters With Constrained Coefficients' Dynamic Range Ratio 85
6.1 Optimization-Based Design of Pencil Beams with Minimax Sidelobes 85
6.2 DRR Constrained Design 85
6.3 Design Examples 88
6.3.1 Example 1 88
6.3.2 Example 2 89
6.3.3 Example 3 90
6.3.4 Example 4 91
6.3.5 Example 5 93
7 Global Optimization of Sparse Spatial Filters With Constrained Coefficients' Dynamic Range Ratio 95
7.1 DRR Constrained Design With Sparsity 95
7.2 Design Examples 97
7.2.1 Example 1 97
7.2.2 Example 2 98
7.2.3 Example 3 98
7.2.4 Example 4 100
8 Conclusion 103
9 References 104
Biography 110
Životopis 112

1 Introduction

Electric filters are uniquely represented with their transfer functions. These functions are described with coefficients which cover a certain, usually large, dynamic range. When realizing transfer functions, these coefficients are represented by multipliers. The implementation of multipliers requires significant hardware resources. Therefore, efficient systems are often designed to minimize the number of multipliers or to simplify their structures. The former approach leads to systems with sparse coefficients, also known as sparse systems. The latter is used in the design of spatial filters where constraining the dynamic range ratio of excitation coefficients is beneficial. One stream in the design of efficient systems is based on numerical optimization. However, the corresponding optimization problems are often nonconvex and therefore difficult to solve. Many iterative, local methods have been developed which efficiently cope with these problems. The methods for global optimization have also been considered. However, these methods are more complex and require more time for the design.

In the scope of this dissertation, methods for the design of both – sparse systems and systems with constrained dynamic range ratio – will be considered. In the field of sparse systems, the original problem leads to minimization of the l_0-norm. However, due to combinatorial nature of problems addressed in such a way, relaxations have been introduced. Here, a method for the design of sparse FIR filters with peak-error constraints is proposed. The method is based on signomial programming and utilizes the relaxation of the l_0-norm with an l_p-norm where $0 < p < 1$. It is applied in the design of filters with linear phase and the filters with no phase specifications. Global methods for the design of sparse FIR filters with linear phase have also been addressed. In particular, a method based on branch and bound algorithm which constrains the filter's magnitude response in peak-, quadratic-error, as well as in both senses is proposed. The method utilizes the depth-first as well as a combination of the depth-first and breadth-first search strategy.

In the domain of spatial filters, a method for the design of pencil beams with constrained dynamic range ratio of excitation coefficients is proposed. These pencil beams are optimized to have minimax sidelobes. The method is based on the branch and bound algorithm which incorporates efficient pruning. Furthermore, it is shown that by allowing

some coefficients to take zero values, a better radiation patterns can be achieved for the same dynamic range ratio.

The thesis is organized as follows. The second chapter provides an introductory overview of digital filters and antenna arrays, with focus on efficiency. The concept of sparsity in these areas has its origin in compressed sensing. Therefore, its basic concepts will also be given. In a compressed sensing, the l_1-norm is utilized to obtain sparse signals. However, the differences between the compressed sensing and filter design are discussed and it is shown why the use of l_1-norm is not so effective in filter design. At the end, state of the art methods in both sparse FIR filter design and antenna array design are presented.

Design methods used in this research are based on optimization techniques, so the third chapter will provide the introduction to optimization as well as an overview of all optimization techniques used later in the dissertation.

Chapters four to seven present new methods in the design of efficient systems. In chapter four, signomial programming is applied to sparse filter design. The design examples are shown on FIR filters with linear phase and FIR filters without phase specifications, both constrained in peak-error sense.

The fifth chapter presents a method for global optimization of sparse FIR filters based on efficient branch and bound algorithm. This method is applied to linear phase FIR filters with peak- and quadratic-error constraints, as well as to filters with both of these constraints.

In the sixth chapter, a global method based on branch and bound algorithm is applied in the design of pencil beams with minimax sidelobes and constrained coefficients' dynamic range ratio.

In the seventh chapter, the concept of sparsity is introduced to improve the radiation patterns of pencil beams with constrained coefficients' dynamic range ratio.

2 Efficient Systems

2.1 Electric Filters

Filtering is the most widely used signal processing technique [1], [2]. It is implemented by the devices called filters. Filter design can be performed in the time, frequency or spatial domain. In a time-domain synthesis, filters are designed to ensure certain parameters of the processed waveforms. Frequency-domain synthesis usually focuses on the selectivity. The design of spatial filters is performed to shape the structure of radiated wave.

Frequency-selective filters can be categorized as lowpass, highpass, bandpass, bandstop or, as a special case, allpass filters. The lowpass filter passes the signal components whose frequencies are lower than a specified cutoff frequency, f_c, while rejecting the other components. Its complement is the highpass filter, which passes the components whose frequencies are higher than f_c. The bandpass filter passes the components between cutoff frequencies f_{c1} and f_{c2}. Its complement is the bandstop filter, which rejects these components. There are also special filter types, such as notch filters, which block single frequency, multiband filters, which pass more than one frequency band, comb filters, which reject periodically spaced frequency regions, etc.

Spatial filters are used in a wide area of applications ranging from optical imaging, acoustics, communications, etc. The most common applications of these filters are antenna arrays, which direct the radiated energy into a spatial region of interest. In this sense, they generate a narrow beam, wide beam, multiple beams or other form of beam required by the application at hand. The antenna arrays appear in linear, planar, circular, cylindrical and spherical forms.

In this research, methods for efficient design of frequency- and spatially-selective filters are investigated with focus set on the design of sparse digital filters and linear antenna arrays.

2.1.1 Digital Filters

Digital filter is a system that performs specific mathematical operation on a discrete-time signal. Filter's properties are uniquely determined by its impulse response, $h(n)$. Filter's response to an arbitrary signal $x(n)$ is given by [1]

$$y(n) = \sum_{k=-\infty}^{\infty} x(k)h(n-k) \tag{2.1}$$

Considering impulse response length, there are two types of digital filters – finite impulse response (FIR) and infinite impulse response (IIR) filters. FIR filters are more common because they allow ideal linear phase. Furthermore, they are inherently stable and their implementation is simple. IIR filters have more complex implementation. However, they exhibit better performances than the FIR filters of the same orders. Unfortunately, they lack the phase linearity and special attention must be paid to achieve their stability.

2.1.1.1 FIR Filters

For a causal FIR filter of order *N*, the impulse response is given by

$$h(n) = h(0)\delta(n) + h(1)\delta(n-1) + \ldots + h(N)\delta(n-N) = \sum_{k=0}^{N} h(k)\delta(n-k) \tag{2.2}$$

where $\delta(n)$ is the Kronecker delta function. Filter's transfer function is obtained as

$$H(z) = Z\{h(n)\} = h(0) + h(1)z^{-1} + \cdots + h(N)z^{-N} = \sum_{k=0}^{N} h(k)z^{-k} \tag{2.3}$$

where $Z\{\cdot\}$ denotes the *z*-transform. The frequency response is obtained by evaluating $H(z)$ along the unit circle, as in

$$H(e^{j\omega}) = \sum_{k=0}^{N} h(k)e^{-j\omega k} \tag{2.4}$$

where $\omega = 2\pi f$ is angular frequency.

FIR filters can be designed to have ideal linear phase which is achieved if the impulse response is symmetric or antisymmetric. Such a response also enables reducing the number of multipliers in the implementation by the factor of two. FIR filters designed without requirements set on the phase response allow arbitrary distribution of their coefficients. This introduces additional degrees of freedom, which can be spent on increasing the selectivity.

FIR filters with linear phase are categorized into four different types. Type 1 and 2 FIR filters have symmetric, whereas Type 3 and 4 have antisymmetric impulse responses. The

examples of symmetric impulse responses are shown in Figure 2.1 for an even and an odd filter order.

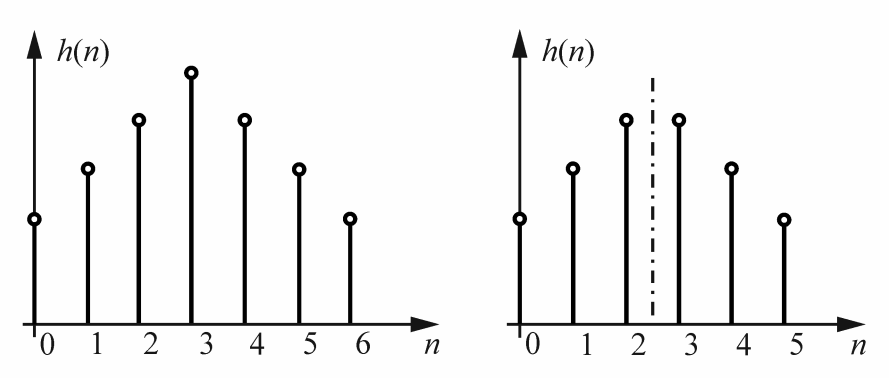

Figure 2.1 Impulse response of Type 1 (left) and Type 2 (right) FIR filter.

The frequency response of Type 1 filter is obtained by introducing $h(k) = h(N-k)$ into (2.4) and by rearranging, which results in

$$H\left(e^{j\omega}\right) = \sum_{k=0}^{N/2-1} h(k)\left(e^{-j\omega k} + e^{-j\omega(N-k)}\right) + h\left(N/2\right)e^{-j\omega N/2} \tag{2.5}$$

By further rearranging, the expression in (2.5) takes the form

$$H\left(e^{j\omega}\right) = e^{-j\omega N/2}\left\{\sum_{k=0}^{N/2-1} h(k)\left(e^{j\omega(N/2-k)} + e^{-j\omega(N/2-k)}\right) + h\left(N/2\right)\right\} \tag{2.6}$$

By using trigonometric identities, the frequency response is obtained as

$$H\left(e^{j\omega}\right) = e^{-j\omega N/2}\left\{2\sum_{k=0}^{N/2-1} h(k)\cos\left[\omega\left(N/2-k\right)\right] + h\left(N/2\right)\right\} \tag{2.7}$$

Component $e^{-j\omega N/2}$ is filter's phase shift whereas the remaining part is the zero-phase amplitude response.

Similar to Type 1, Type 2 frequency response is obtained by using $h(k) = h(N-k)$, as in

$$H\left(e^{j\omega}\right) = e^{-j\omega N/2}\cdot 2\sum_{k=0}^{(N-1)/2} h(k)\cos\left[\omega\left(N/2-k\right)\right] \tag{2.8}$$

The examples of antisymmetric impulse responses are shown in Figure 2.2 for an even and an odd filter order. By using $h(k) = -h(N-k)$ and $h(N/2) = 0$, the frequency response of Type 3 FIR filter is obtained as

$$H\left(e^{j\omega}\right) = e^{-j\omega N/2} e^{-j\pi/2} \cdot (-2) \sum_{k=0}^{N/2-1} h(k) \sin\left[\omega\left(N/2-k\right)\right] \tag{2.9}$$

Similarly, by using $h(k) = -h(N-k)$ the frequency response of Type 4 filter is obtained in a form

$$H\left(e^{j\omega}\right) = e^{-j\omega N/2} e^{-j\pi/2} \cdot (-2) \sum_{k=0}^{(N-1)/2} h(k) \sin\left[\omega\left(N/2-k\right)\right] \tag{2.10}$$

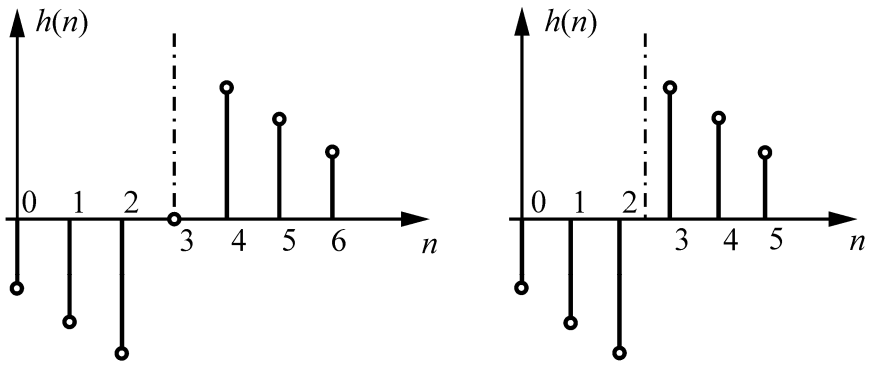

Figure 2.2 Impulse response of Type 3 (left) and Type 4 (right) FIR filter.

2.1.1.2 Optimization-Based FIR Filter Design

When designing FIR filters, the coefficients are often obtained by using numerical optimization. The optimization-based design of an FIR filter with real-valued coefficients can be described as

$$\underset{\mathbf{h}}{\text{minimize}} \quad \varepsilon\left[\mathbf{h}, H_d\left(\omega\right)\right] \tag{2.11}$$

where ε is approximation error, $\mathbf{h} = [h_0, h_1, \ldots, h_N]^T$ is filter's impulse response and $H_d(\omega)$ is desired frequency response. The error in (2.11) can be expressed as the integral of p-powered deviation of the frequency response, as in

$$\varepsilon_p(\mathbf{h}) = \left(\frac{1}{\pi} \int_0^{\pi} W(\omega) \left| H(\mathbf{h},\omega) - H_d(\omega) \right|^p d\omega \right)^{\frac{1}{p}} \tag{2.12}$$

where $W(\omega) \geq 0$ is a weighting function and $H(\mathbf{h},\omega)$ is filter's amplitude response. In this theses, two most common choices of p will be further investigated, $p \to \infty$ and $p = 2$. For $p \to \infty$, the error in (2.11) can be formulated as

$$\varepsilon_\infty(\mathbf{h}) = \max_{0 \leq \omega \leq \pi} W(\omega) \left| H(\mathbf{h},\omega) - H_d(\omega) \right| \tag{2.13}$$

The minimization of $\varepsilon_\infty(\mathbf{h})$ results in filter's frequency response which is optimum in a minimax sense. The coefficients of such a filter can be obtained by solving the problem

$$\begin{aligned} &\underset{\mathbf{h},\delta}{\text{minimize}} \quad \delta \\ &\text{subject to} \quad W(\omega) \left| H(\mathbf{h},\omega) - H_d(\omega) \right| \leq \delta \end{aligned} \tag{2.14}$$

where δ is approximation error. One popular method for minimax FIR filter design was proposed by Parks and McClellan [3].

Minimax criterion does not take into account the energy obtained in certain filter's bands. However, the energy criterion can be easily incorporated into an optimization-based design by applying $p = 2$ in (2.12), which results in

$$\varepsilon_2(\mathbf{h}) = \sqrt{\frac{1}{\pi} \int_0^{\pi} W(\omega) \left| H(\mathbf{h},\omega) - H_d(\omega) \right|^2 d\omega} \tag{2.15}$$

Since $\varepsilon_2(\mathbf{h}) \geq 0$,

$$\arg\min_{\mathbf{h}} \varepsilon_2^2(\mathbf{h}) = \arg\min_{\mathbf{h}} \varepsilon_2(\mathbf{h}) \tag{2.16}$$

Consequently, the filter's coefficients can be obtained by the optimization of squared error, rather than the error itself. Therefore, the coefficients can be obtained by solving the problem

$$\begin{aligned} &\underset{\mathbf{h},\gamma}{\text{minimize}} \quad \gamma \\ &\text{subject to} \quad \frac{1}{\pi}\int_0^{\pi} W(\omega)\left[H(\mathbf{h},\omega)-H_d(\omega)\right]^2 d\omega \le \gamma \end{aligned} \tag{2.17}$$

where γ is approximation error. The resulting filters are optimum in the least-squares sense.

In many applications, the total stopband energy and maximum stopband gain are both important [4]. Furthermore, narrow band filters require constraining the energy in the stopbands while limiting the maximum gain deviation in all care bands. Therefore, it is useful to combine both approximation errors. The examples of such designs will be analyzed more thoroughly in the following chapters.

2.1.2 Antenna Arrays

Antenna arrays are composed of several radiation sources called antenna elements. Their radiation is analyzed in two different regions, the near and the far field. In the near field, nonradiative behavior dominates. The near field is followed by the transition zone in which nonradiative as well as radiative behavior is encountered. In the far field, nonradiative effects can be neglected. This region is considered to begin at the distance of several wavelengths from the antenna. In this dissertation, only antenna arrays in the far field region are considered.

The antenna arrays are characterized by their gain or radiation pattern, which is sometimes called beampattern. Since this pattern is obtained by spatial filtering, its generation is also called beamforming. The simplest radiation pattern consists of one main and several side lobes. The basic beamforming task is to design the main lobe pointing at a specified direction and to keep the side lobes within the specified boundaries. In addition, some applications require changing the direction of the main lobe, which is performed by the beam steering.

Based on the shape of the radiation patterns, antenna arrays form wide, narrow, shaped, pencil or focused, flat top, multiple and other beam patterns. These patterns are obtained by various excitations and arrangements of antenna elements, utilizing linear, planar, circular, cylindrical, spherical or other structures. In these structures, the elements can be positioned with equal or unequal interelement spacing, thus forming uniform or nonuniform arrays. Nonuniform arrays, as well as the uniform arrays with missing certain elements, are treated as sparse arrays.

2.1.2.1 Single Antenna Response

Figure 2.3 shows single antenna element placed at the point [x_0, 0] of the x-z coordinate system. If this element is excited with the signal $U_0\cos(\omega_0 t + \varphi)$ where U_0 is the excitation's magnitude, ω_0 is angular frequency and φ is feed's angle, the electric field in element's vicinity is given by [5]

$$E_0(t) = CF(\theta)U_0 \cos(\omega_0 t + \varphi) \tag{2.18}$$

where C is a constant, $F(\theta)$ is antenna's radiation pattern and θ is azimuth direction angle in range $-\pi/2 \le \theta \le \pi/2$.

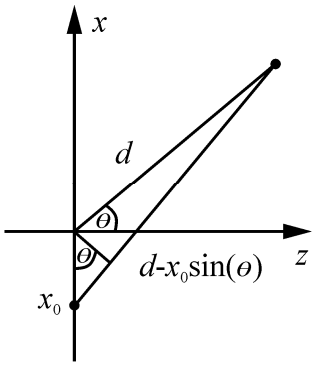

Figure 2.3 One antenna element positioned at coordinate x_0 on x-axis.

At the point distanced d from the origin, at an angle θ, electric field takes the form

$$E_d(t) = \frac{C}{d^2} F(\theta)U_0 \cos\left[\omega_0 (t-\tau) + \varphi\right] \tag{2.19}$$

By using analytical continuation of (2.19), the electric field can be expressed as a phasor

$$E_d(t) = \frac{C}{d^2} F(\theta)U_0 e^{j(\omega_0 t + \varphi - \omega_0 \tau)} \tag{2.20}$$

where

$$\tau = 2\pi \frac{d - x_0 \sin(\theta)}{\lambda \omega_0} \tag{2.21}$$

and λ is the wavelength of the transmitted signal. By substituting (2.21) into (2.20) and rearranging, the electric field is obtained as

$$E_d(t) = F(\theta) e^{j\omega_0 t} \frac{C}{d^2} e^{-\frac{2\pi d}{\lambda}} U_0 e^{j\varphi} e^{j\frac{2\pi x_0 \sin(\theta)}{\lambda}} \tag{2.22}$$

In (2.22), the expression

$$F_d(\theta) = F(\theta) U_0 e^{j\varphi} e^{j\frac{2\pi x_0 \sin(\theta)}{\lambda}} \tag{2.23}$$

is recognized as the complex radiation pattern obtained at the observed angle. If isotropic antenna is used, $F(\theta) = 1$. Assuming complex excitation

$$a_0 = U_0 e^{j\varphi} \tag{2.24}$$

the radiation pattern takes the form

$$F_d(\theta) = a_0 e^{j\frac{2\pi x_0 \sin(\theta)}{\lambda}} \tag{2.25}$$

2.1.2.2 Multiple Antenna Response

The response of multiple antennas can be easily obtained by using superposition. Figure 2.4 shows an example of linear antenna array with elements placed along the x axis at the positions x_k, $k = 1, 2, \ldots, N$. If all elements have equal radiation patterns $F(\theta)$ and excitation coefficients a_k, $k = 1, 2, \ldots, N$, the array's radiation pattern is obtained as

$$F_a(\mathbf{a}, \theta) = F(\theta) \sum_{k=1}^{N} a_k e^{j\frac{2\pi x_k \sin(\theta)}{\lambda}} \tag{2.26}$$

where $\mathbf{a} = [a_1, \ldots, a_N]^T$. The sum in (2.26) is called array factor [5], [6]

$$A(\mathbf{a}, \theta) = \sum_{k=1}^{N} a_k e^{j\frac{2\pi x_k \sin(\theta)}{\lambda}} \tag{2.27}$$

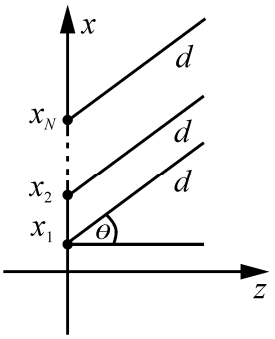

Figure 2.4 Multiple antenna elements positioned at coordinates x_k, $k = 1, 2,..., N$ on x-axis.

2.1.2.3 Relationship Between Antenna Arrays and FIR Filters

If an antenna array utilizes $N + 1$ equidistant elements placed at $x_k = k\lambda/2$, $k = 0, 1, 2, \ldots, N$, the array factor is given by

$$A(\mathbf{a}, \theta) = \sum_{k=0}^{N} a_k e^{jk\pi \sin(\theta)} \tag{2.28}$$

An FIR filter with $N + 1$ coefficients has a frequency response

$$H(e^{j\omega}) = \sum_{k=0}^{N} h_k e^{-jk\omega} \tag{2.29}$$

Clearly, the array factor can be obtained from the frequency response of an FIR filter assuming

$$\omega = -\pi \sin(\theta) \tag{2.30}$$

Apparently, the methods used in FIR filter design can be applied in the design of antenna arrays with equal interelement spacing. However, it should be noted that the frequency scale is warped with the factor $\sin(\theta)$. Similar expressions can be obtained for interelement spacing of λ.

2.1.2.4 Optimization-Based Pencil Beam Design

Antenna arrays can be designed to direct energy into a narrow spatial angle. This type of radiation pattern is called focused or pencil beam.

Optimization-based pencil beam design aims to minimize the sidelobe error, while keeping the maximum power density at angle $\theta = 0$ rad. The latter is usually achieved by forcing $A(\mathbf{a}, 0°) = 1$. For isotropic antenna elements, the optimization problem takes the form

$$\begin{aligned} &\underset{\mathbf{a}}{\text{minimize}} \quad \varepsilon(\mathbf{a}) \\ &\text{subject to} \quad \sum_{k=1}^{N} a_k = 1 \end{aligned} \tag{2.31}$$

where $\varepsilon(\mathbf{a})$ is the sidelobe error and a_k, $k = 1, 2, \ldots, N$ are excitation coefficients. In this problem $A(\mathbf{a}, 0°) = 1$ is ensured by the constraint, whereas the sidelobe region is shaped by an appropriate choice of the objective function. Similarly to the design of FIR filters, $\varepsilon(\mathbf{a})$ can be expressed as the integral of p-powered radiation pattern. In this case, it takes the form

$$\varepsilon_p(\mathbf{a}, \theta_s) = \left(4\pi \int_{\theta_s}^{\pi/2} \left| F_a(\mathbf{a}, \theta) \right|^p \cos(\theta)\, d\theta \right)^{\frac{1}{p}} \tag{2.32}$$

where $0 < \theta_s \leq \pi/2$ is the start of the sidelobe region. In a classic design, $p \to \infty$ or $p = 2$ is chosen, ensuring the sidelobe region is optimum in a l_∞- or l_2-sense. For $p \to \infty$ the error in (2.32) can be expressed as

$$\varepsilon_\infty(\mathbf{a}, \theta_s) = \max_{\theta_s \leq \theta \leq \pi/2} \left| F_a(\mathbf{a}, \theta) \right| \tag{2.33}$$

The corresponding optimization problem is given by

$$\begin{aligned} &\underset{\mathbf{a}, \delta}{\text{minimize}} \quad \delta \\ &\text{subject to} \quad \sum_{k=1}^{N} a_k = 1 \\ &\qquad\qquad \left| F_a(\mathbf{a}, \theta) \right| \leq \delta, \quad \theta_s \leq \theta \leq \frac{\pi}{2} \end{aligned} \tag{2.34}$$

where δ is approximation error. For $p = 2$, the error takes the form

$$\varepsilon_2(\mathbf{a}, \theta_s) = \sqrt{4\pi \int_{\theta_s}^{\pi/2} \left| F_a(\mathbf{a}, \theta) \right|^2 \cos(\theta)\, d\theta} \tag{2.35}$$

The corresponding optimization problem is given by

$$\begin{aligned} \underset{\mathbf{a},\gamma}{\text{minimize}} \quad & \gamma \\ \text{subject to} \quad & \sum_{k=1}^{N} a_k = 1 \\ & 4\pi \int_{\theta_s}^{\pi/2} \left|F_a(\mathbf{a},\theta)\right|^2 \cos(\theta)\, d\theta \le \gamma \end{aligned} \tag{2.36}$$

where γ is approximation error.

2.2 Efficient Electric Filters

As mentioned in previous section, FIR filters can easily achieve linear phase and they are inherently stable. These features make FIR filters very convenient and popular. However, their relatively high implementation complexity might be a drawback in some applications. Many techniques have been developed for lowering the complexity, most of which try to reduce the number of general purpose multipliers required for filter's implementation. Popular techniques include frequency response masking, using the sums of signed-powers-of-two and sparse filter design.

Frequency response masking is a set of procedures used to design very sharp linear-phase FIR filters with arbitrary passbands [7]. These filters achieve efficient implementation at the cost of increased group-delay. In addition, frequency masking cannot be used for obtaining non-frequency-selective filters, such as digital integrators. A low implementation complexity can be obtained if all filter coefficients are represented as sums of signed power-of-two terms. This leads to hardware design with no general-purpose multipliers, which are here implemented by using adders and shifts. The design of sparse filters is focused on minimization of the number of nonzero coefficients in filter's impulse response, regardless of the technique used for its implementation.

The above techniques reduce the amount of hardware required for filter's implementation. However, in some applications it is mandatory to meet the requirements set on particular components. Such is the case in the design of antenna arrays, in which the coefficients of spatial filters are incorporated in the feeding networks. Since these networks operate at high frequencies, it is difficult to implement the coefficients which have a large dynamic range ratio (DRR). Consequently, the design with constrained DRR is preferable. Additional relaxation is achieved if some coefficients are missing, which is obtained in a sparse design.

In this dissertation, efficient filters will be considered in the scope of sparse FIR filters, classic antenna arrays with constrained DRR as well as sparse antenna arrays with constrained DRR.

The concept of sparsity has its origin in the early stages of filter design but has regained much popularity in the last few decades because of the intense development of compressed sensing. Therefore, a short introduction to compressed sensing will be given in the next section. Furthermore, state-of-the-art methods in the sparse FIR filter design will be presented. Finally, efficient spatial filters with recent methods for their design will also be described.

2.3 Sparsity in Filter Design

2.3.1 Compressed Sensing

Compressed sensing (CS) is a technique for data acquisition which enables complete reconstruction of signals from incomplete sets of measurements [8], [9], [10], [11], [12], [13]. Good surveys on this topic can be found in [14] and [15].

One of the pioneer compressed sensing applications is single pixel camera [16]. This camera utilizes a digital micromirror device which consists of many small mirrors randomly positioned to stop or let the light onto the optical detector. Different combinations of opened and closed mirrors capture information about the object. Each mirror setting brings one measurement. This principle is also used for 3D scene imaging [17].

One of first industry-important CS applications utilizes compressed sensing techniques in processing of magnetic resonance images for 3D angiography, heart and brain imaging [18], [19], etc. The use of CS in medical imaging has reduced the acquisition time significantly. This is crucial because majority of errors occur when imaging is slow.

Recently, a compressed sensing application for reconstructing a scene from a defocused image was introduced [20]. Classical imaging system projects the surface of a 3D object onto a 2D camera. These surface points are sparse with respect to all points present in 3D space. If an object consists of multiple light sources, it is possible to reconstruct their positions by using compressed sensing.

2.3.2 Reconstruction of Sparse Signals

The main task of the compressed sensing is to reconstruct the unknown sparse vector $\mathbf{x} \in \mathcal{R}^N$ from a measured vector $\mathbf{y} \in \mathcal{R}^M$ assuming $M < N$. The measurement is taken as

$$\mathbf{y} = \mathbf{A}\mathbf{x} \tag{2.37}$$

where $\mathbf{A} \in \mathcal{R}^{M \times N}$ is measurement matrix. If the solution of (2.37) exists, then it surely isn't unique. Moreover, since $M < N$, the number of solutions is infinite. However, unique reconstruction of $\mathbf{x}$ is possible if certain properties are met. The key requirement for successful reconstruction is that $\mathbf{x}$ is sparse in a specific domain. Furthermore, $\mathbf{A}$ should exhibit a small mutual coherence [9] or satisfy restricted isometry property [13].

If sparsity reconstruction is formed as an optimization problem, the sparsity requirement is expressed with the objective function whereas linear measurements in (2.37) represent the constraints. The sparsity is usually expressed with specific norm. In mathematics, the l_p-norm is defined as

$$\left\|\mathbf{x}\right\|_p = \left(\sum_{k=1}^{N} \left|x_k\right|^p \right)^{\frac{1}{p}}, \quad p \geq 1 \tag{2.38}$$

where $\mathbf{x} = [x_0, x_1, \ldots, x_N]^T$ and $p \in \mathcal{R}$. Frequently used choices of p are 1, 2, and ∞, leading to l_1-norm, l_2-norm or Euclidian norm and l_∞-norm or minimax norm. Strict definition of norm is limited to $p \geq 1$, but the definition in (2.38) can be extended to $0 \leq p < 1$. For such values of p, the norm is considered improper because the triangular inequality is violated. However, useful properties are obtained for l_p-norm with $0 \leq p < 1$.

The sparsity of $\mathbf{x}$ can be directly expressed with l_0-norm [14]. The reconstruction problem is than expressed as

$$\begin{aligned} &\underset{\mathbf{x}}{\text{minimize}} \quad \left\|\mathbf{x}\right\|_0 \\ &\text{subject to} \quad \mathbf{y} = \mathbf{A}\mathbf{x} \end{aligned} \tag{2.39}$$

The problem in (2.39) is computationally demanding. Since the positions of nonzero elements in $\mathbf{x}$ are not known in advance, all subsets of $\mathbf{x}$ containing nonzero elements need to be examined. Such an examination can be performed by using an exhaustive search. Apparently, this problem cannot be solved in polynomial time. To avoid the need for the exhaustive search, two approaches are used. First of them solves the problem in (2.39) approximately, whereas the second approximates the problem in (2.39) and then exactly solves the approximate problem. The former approach is based on several variants of greedy algorithm. The most popular of them can be found in [14]. The latter approach approximates the l_0-norm

with some proper norm, resulting in convex optimization problem. This approximation is described in the following sections.

2.3.3 Application of L_1-Norm in Compressed Sensing

If l_0-norm is replaced with general l_p-norm, the problem in (2.39) becomes

$$\begin{aligned} \underset{\mathbf{x}}{\text{minimize}} \quad & \|\mathbf{x}\|_p^p \\ \text{subject to} \quad & \mathbf{y} = \mathbf{A}\mathbf{x} \end{aligned} \tag{2.40}$$

where $p \geq 0$. The solution of problem in (2.40) needs to be identical to the solution of the problem in (2.39). Additionally, the problem in (2.40) needs to be globally solvable which is possible if the problem is convex. The solutions of

$$\|\mathbf{x}\|_p^p \leq r^p \tag{2.41}$$

lie in an l_p-ball with radius r. Points in a ball for which

$$\|\mathbf{x}\|_p^p = r^p \tag{2.42}$$

have the same value of the objective function. By minimizing the objective function, the minimum r is obtained. Therefore, the solution to the problem in (2.40) is a point which satisfies (2.37) and has the minimum r. Solving the problem can be visualized by inflating the l_p-ball until it touches the hyperplane $\mathbf{y} = \mathbf{A}\mathbf{x}$. Figure 2.5 illustrates this example for a two-dimensional $\mathbf{x}$ and for various values of p. For $0 \leq p \leq 1$, the touching point of a ball and the hyperplane result in a sparse solution, whereas for $p > 1$ the results are nonsparse with high probability. On the other hand, objective function is convex for $p \geq 1$. Given that $p \leq 1$ ensures sparsity, and $p \geq 1$ ensures the convexity, these two opposing requirements are met only for $p = 1$, resulting in the problem

$$\begin{aligned} \underset{\mathbf{x}}{\text{minimize}} \quad & \|\mathbf{x}\|_1 \\ \text{subject to} \quad & \mathbf{y} = \mathbf{A}\mathbf{x} \end{aligned} \tag{2.43}$$

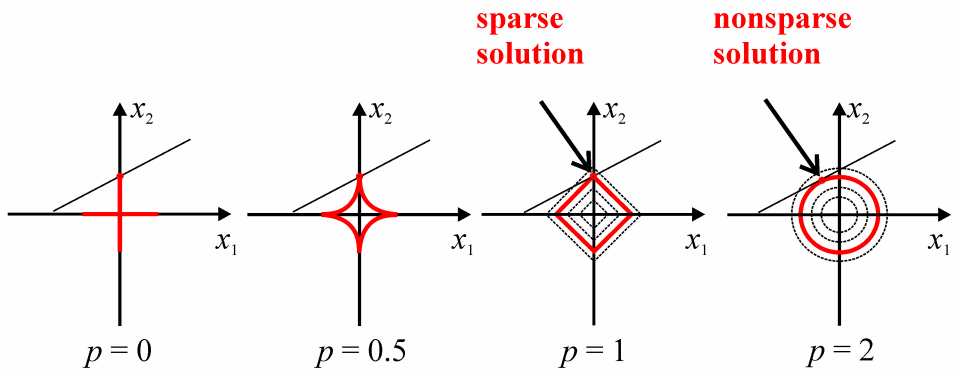

Figure 2.5 Visualization of CS reconstruction by inflating l_p-ball for various values of p.

2.3.4 Application of L_1-Norm in Filter Design

Using the l_1-norm in the compressed sensing works well (see [14] and the references therein). Therefore, a logical step is to apply the same principles in sparse filter design. Unfortunately, this is not a straightforward process as the basic prerequisites necessary for the unique solution are not met here. In Figure 2.6 it is shown that feasible region in a sparse filter design is polytope obtained by intersections of inequality constraints, rather than a hyperplane [21]. Consequently, it is more likely that minimization of l_1-norm results in a nonsparse solution than it does in a CS reconstruction. On the other hand, minimization of l_p-norm with $0 < p < 1$ provides sparse solution in majority of cases [21].

Although sparse filter design based on the l_1-norm does not result in sparse coefficients, it provides a certain number of coefficients with small values. This feature is used in many iterative methods where these small coefficients are set to zero and the optimization process is then restarted. Some of these methods are described in the following sections.

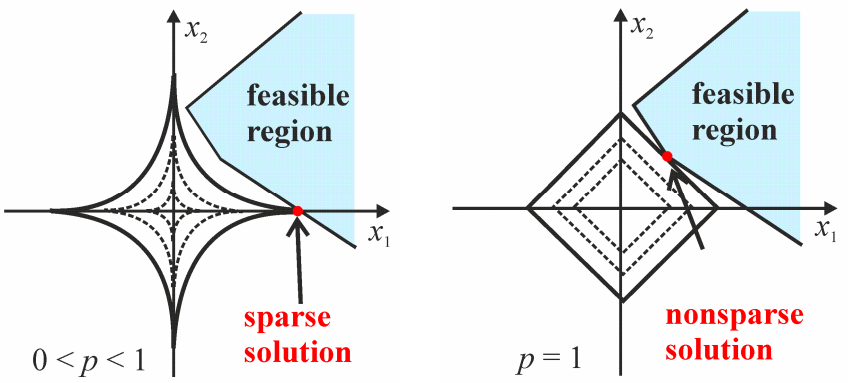

Figure 2.6 Illustration of sparse filter design based on l_p-norm with $0 < p < 1$ and $p = 1$.

2.4 State-of-the-Art in Sparse FIR Filter Design

Sparse filters attract researchers' attention for more than three decades, resulting in many design methods [21]–[53]. Most of these methods deal filters with peak-error constraints [21], [22], [23], [24], [25], [27], [28], [29], [30], [34], [36], [37], [38], [39], [40], [41], [43], [44], [47], [48], [49], [50], [51], [52], [53] and filters with quadratic-error constraints [24], [28], [31], [32], [33], [34], [37], [42], [45], [46], [50]. In further sections some of these methods are briefly described. However, they are grouped with respect to design methodology rather than used approximation techniques.

2.4.1 Design Based on Approximate Solving of Greedy Algorithms

As in the compressed sensing, l_0 problem in FIR filter design can also be solved approximately. In this approach, different greedy algorithms have been used, such as [25], [38]. Greedy algorithms solve the optimization problem iteratively, dividing the original problem to several stages. In each stage, the best local solution based on some criteria is found. For example, in [38] authors successively add zero samples to the impulse response. In each step, feasible region of every coefficient is calculated. The coefficient with middle point of feasible range closest to zero is forced to zero, and the remaining coefficients are recalculated.

Generally, solution's globality when using methods based of greedy algorithms cannot be guaranteed. However, good approximate solutions can be found in a reasonable time.

2.4.2 Design Based on Minimization of L_p-Norm

Approximating the objective function in (2.39) with a norm other than the zeroth norm brings several benefits. If a proper norm is used, the optimization problem becomes convex and can be solved globally even for large number of unknowns. However, the application of an l_p-norm with $p \geq 1$ in sparse filter design is not straightforward, as elaborated in Section 2.3.4. Instead, iterative procedure is usually needed. On the other hand, an improper norm with $0 < p < 1$ allows formulating the filter design as a single optimization problem. However, due to nonconvexity of the improper norm, solving of this problem is also realized by iterative procedure.

Sometimes, greedy algorithm is combined with convex optimization. Such an approach is used in [27] where the nonzero set is found by using linear programming. The method referred to utilizes successive thinning and CS inspired l_1-norm design. The former finds the set of zero coefficients by following two rules – minimum-increase and smallest coefficient rule. The minimum-increase rule chooses a new zero-coefficient from the nonzero set as the one that minimally increases the filter approximation error. On the other hand, the smallest coefficient rule finds the coefficient with the smallest absolute value and sets it to zero. Remaining coefficients are then reoptimized by using the linear programming. The latter utilizes l_1-norm design which results in many small coefficients. The ones with the smallest absolute values are zeroed and remaining coefficients are recalculated. Similar technique is presented in [30] where iterative-shrinkage-thresholding inspired algorithm is used to determine the nonzero coefficient set. This algorithm divides the original problem to simpler subproblems and solves them by using second order cone programming. In [30], the algorithm is applied to sparse filter design with peak-error constraints. Similar approach is used in [31] but applied to filters which have minimum weighted least-squares approximation error.

Relaxation of l_0-norm by using the l_1-norm was used in [28]. In the paper referred to, peak- or quadratic-error constraints in problems (2.14) and (2.17) are moved to the same objective function, whose minimization indicates which coefficients should be zeroed. The optimization problem is expressed as

$$\underset{\mathbf{h}}{\text{minimize}} \left\| H\left(\mathbf{h},\omega\right) - H_d\left(\omega\right) \right\|_{2,\infty} + \chi \left\| \mathbf{h} \right\|_1 \tag{2.44}$$

where $\chi > 0$ is a scalar weight and $\|\cdot\|_{2,\infty}$ denotes the l_2-norm for quadratic and l_∞-norm for peak-error constraints, respectively. More advanced search of zero coefficients by using the l_1-norm is given in [29] where larger weights are given to the coefficients that are more likely to be zero. These weights are then updated in each iteration. Similar idea with specific algorithm enhancements is given in [39] and [40]. Additionally, methods presented in [44] and [52] showed that coefficients at specific positions cannot be zero. These coefficients are taken out of the optimization search before solving, thus decreasing algorithm's complexity. All methods referred to are iterative and follow a two-step procedure. In the first step, positions of nonzero coefficients are found based on some criteria, and in the second step some of these coefficients are set to zero and the optimization is rerun to calculate the remaining coefficients.

As l_1-norm doesn't ensure sparsity in the filter design (see Section 2.3.4), attention has been paid to l_p-norm with $0 < p < 1$. In [21], such a norm is utilized with local search algorithm that resembles a simplex program for linear programming problems. Furthermore, this dissertation presents a method that utilizes signomial programming to solve an optimization problem containing the l_p-norm with $0 < p < 1$. The examples will be provided in two filter applications – sparse linear-phase FIR filters [49] and FIR filters which have no phase specifications [48].

2.4.3 Design Based on Branch and Bound Algorithm

The methods described in Sections 2.4.1 and 2.4.2 solve either the original problem of sparse FIR filter design approximately or approximate problem exactly. All of them provide useful results but none of them provides the global solution of the original problem. However, significant effort has been made in developing global methods for sparse FIR filter design. Promising but computationally expensive approach to global optimization of sparse FIR filters is the branch and bound method, which goes through all possible sets of sparse coefficients which form a tree [24], [33], [44]. If the pruning is done efficiently, branch and bound method can decrease convergence time compared to an exhaustive search approach for several orders of magnitude.

In [24], a branch and bound method has been proposed for filters with peak-error constraints in care bands and energy constraint in the stopband. The method utilizes depth-first algorithm and pruning based on problem's feasibility.

In [33], a fast branch and bound method for sparse quadratically-constrained FIR filters has been proposed. This method is capable of designing a 100th order filter in a few hours. In [44], the branch and bound algorithm has also been applied in the design of quadratically-constrained filters. The complexity of the algorithm has been reduced by reducing the search space. In this context, the method referred to solves the approximate problem rather than the original one.

In [26], a method for the design of sparse half-band filters with peak-error constraints has been presented. The design is given in the form of mixed integer programming problem. For a small number of variables, this problem can also be solved by utilizing branch and bound technique.

2.4.4 Design Based on Heuristic Algorithms

Combinatorial problems for sparse FIR filter design can also be solved by using heuristic algorithms. Such an approach sacrifices optimality in favor of speed. However, in some applications it gives satisfactory results. Typical example of heuristic approach is the method in [34], which performs the optimization by using the genetic algorithm. The method in [42] employs the cuckoo search algorithm in the design of asymmetric FIR filters. This method utilizes joint optimization of weighted least-squares approximation error and l_1-norm of filter's coefficients.

2.4.5 Joint Optimization of Sparsity and Filter Order

Most of the methods for sparse FIR filter design do not minimize its order. However, high-order filters have high group delays, which are not desirable in some applications. To cope with this issue, methods that jointly optimize filter's order and sparsity have been developed [36], [41], [47].

In [36], researchers have designed a multiobjective function which simultaneously takes into account sparsity of the coefficients and the filter order. The optimization starts with a high filter order and successively decreases it until the specifications are met. The corresponding optimization problem is given by

$$\begin{aligned} \underset{\mathbf{h}}{\text{minimize}} \quad & (1-\eta)\|\mathbf{h}\|_0 + \eta\rho(\mathbf{h}) \\ \text{subject to} \quad & \left|H(\mathbf{h},\omega) - H_d(\omega)\right| \le \delta(\omega) \end{aligned} \tag{2.45}$$

where $\eta \in [0, 1]$ is balancing weight for sparsity and filter order terms, $\delta(\omega)$ is desired approximation error, and $\rho(\mathbf{h})$ is defined as

$$\rho(\mathbf{h}) = \left\| \begin{bmatrix} |h_{N/2}| \\ |h_{N/2}| + |h_{N/2-1}| \\ \vdots \\ |h_{N/2}| + \cdots + |h_1| + |h_0| \end{bmatrix} \right\|_0 \tag{2.46}$$

It is clear that $\rho(\mathbf{h})$ gives higher weights to coefficients near the tails of the impulse response. Consequently, the optimization tends to minimize impulse response's length. When $\eta = 0$ the problem in (2.45) considers only sparsity and it is equivalent to classic sparse FIR filter design with peak-error constraints. On the other hand, $\eta = 1$ leads to a problem whose optimum is the filter with minimum length. Clearly, in all other cases, the solution is a compromise between sparsity and filter's length.

The problem in (2.45) is nonconvex. Therefore, its global solving is a challenging task. In [36], two algorithms have been proposed which form a relaxed version of the problem in (2.45) and then solve it by using the method of iteratively reweighted least squares. Unfortunately, these algorithms cannot guarantee the globality of the optimum obtained. Moreover, different starting filter orders result in different solutions.

The method in [41] tackles the same problem but without relaxation of the objective function. Instead, it solves the problem in (2.45) by using alternating direction method of multipliers. These enhancements enable solving of high order problems. The resulting filters have many zero coefficients and low lengths.

In [47], the same problem is solved by using the 0-1 exchange algorithm. This approach provides betters results than encountered in [36]. Furthermore, it is insensitive to the starting filter order.

2.5 Efficient Spatial Filters

In antenna array design, a lot of research has been dedicated to the synthesis of pencil beam patterns. In addition, effort is made to increase their efficiency. One approach which ensures efficiency, lowers coefficients' dynamic range ratio [54]–[63]. Another approach minimizes the number of nonzero antenna coefficients [23], [27], [64]–[70] leading to sparse design. Finally, some methods combine constrained dynamic range ratio and sparsity [71], [72]. The representatives of the techniques referred to will be discussed in further text.

2.5.1 Pencil Beams With Low Dynamic Range Ratio

When implementing antenna arrays, each coefficient is represented with one feeding element. Unfortunately, the feeding elements can only be manufactured to cover narrow dynamic range, especially when operating at high frequencies. Consequently, the antenna coefficients should be bounded in the design of the beampattern. One early method in this area utilizes iterative inverse Fourier technique to obtain a desired array factor [54]. The paper referred to constrains the minimum values of antenna coefficients thus achieving lower DRR. However, the maximum coefficients' values are not explicitly bounded. Therefore, this method cannot guarantee the desired DRR value.

Analytical methods for the antenna array synthesis are fast and robust. However, these methods ensure a low DRR indirectly [55], [56], [57]. Consequently, they cannot achieve the exact value of specified DRR. Instead, a low DRR is achieved with trial-and-error procedure. On the other hand, optimization based methods ensure low DRR either by constraining its value [58], [59] or by optimizing it directly [61], [62], [63].

In an optimization-based design, the DRR is kept below a specified value D as in

$$\frac{\max\limits_{1\le k\le N}\left\{\left|a_k\right|\right\}}{\min\limits_{1\le k\le N}\left\{\left|a_k\right|\right\}} \le D \tag{2.47}$$

These constraints can be easily imported into (2.31) resulting in

$$\begin{aligned} \underset{\mathbf{a},t}{\text{minimize}} \quad & \varepsilon(\mathbf{a}) \\ \text{subject to} \quad & \sum_{k=1}^{N} a_k = 1 \\ & \left|a_k\right| \le Dt, \quad k = 1,2,...,N \\ & \left|a_k\right| \ge t, \quad k = 1,2,...,N \\ & t \ge 0 \end{aligned} \tag{2.48}$$

where t is a slack variable. The constraints $|a_k| \ge t,\ k = 1, 2, \ldots, N$ are nonconvex, making solving the problem in (2.48) difficult. It can be addressed by using relaxations of the original problem or by applying iterative [58], evolutional [59], and other types of approximate algorithms. In [58], projection based iterative algorithm is used to find a desired radiation pattern with constrained DRR. On the other hand, in [59] genetic algorithm is used in designing a dual beam with phase only control and specified maximum value of the DRR.

In this dissertation, a global method based on the branch-and-bound algorithm is proposed for design of pencil beams with constrained DRR [60].

Several interesting works are focused on optimizing DRR directly but like in the case when the DRR is constrained, the corresponding problem is not convex. In [61], several different problems in the antenna array design have been addressed. One part of the research handles the DRR constraints by minimizing the difference between maximum and minimum antenna coefficient, assuming all coefficients are positive. The design problem is given by

$$\begin{aligned} \underset{\mathbf{a}}{\text{minimize}} \quad & \max(\mathbf{a}) - \min(\mathbf{a}) \\ \text{subject to} \quad & \sum_{k=1}^{N} a_k = 1 \\ & \left|F_a(\mathbf{a},\theta)\right| \le \delta(\theta), \quad \theta \in \left[\theta_s, \frac{\pi}{2}\right] \\ & \mathbf{a} \ge 0 \end{aligned} \tag{2.49}$$

where $\delta(\theta)$ is upper bound of the sidelobe error defined in finite number of frequency points. When coefficients are allowed to take both positive and negative values, the complexity of the problem rises. Moreover, the optimization problem becomes nonconvex. In [63] the design of complex antenna coefficients with minimum DRR has been presented. It optimizes both the DRR and the radiation pattern. Since this optimization problem is not convex, it is translated to several convex subproblems thus forming an iterative procedure.

2.5.2 Sparse Pencil Beams

Sparse antenna arrays contain a lower number of elements than those which are normally expected looking at arrays' structures or their sizes. The latter are sometimes called aperiodic or nonequidistant arrays. In further text, the term sparse arrays will be used.

The research of sparse arrays has been active for many decades. One of the early works in this area uses method based on linear programming to design beamformers and sparse FIR filters [23]. They show that for a special class of filters, the sidelobe level is decreased up to 20 dB by increasing the number of zero coefficients. Furthermore, they show that their method can be used for cases when some elements become faulty during operation. In [64] researchers use the l_p-norm with $0 < p < 1$ to obtain both linear and planar sparse arrays. They use a simplex search algorithm.

The research of sparse arrays has regained popularity with the emergence of compressed sensing. Several works apply CS algorithms in the design of sparse arrays [27],

[65], [66], [67]. In [27], compressed sensing technique has been applied in the design of sparse FIR filters as well as equidistant beamformers, or more precise symmetric arrays with real coefficients. In [65] and [66], iterative weighted l_1 algorithm for the design of pencil beams has been formed assuming arbitrary coefficients. In addition, in [66] the original problem has been expanded by introducing the near-field constraints. In [67], reweighted l_1 minimization is utilized to obtain sparse equidistant antenna arrays. Then, the obtained positions are refined by using clustering strategy to determine final nonequidistant positions.

An interesting work is presented in [68]. It utilizes convex optimization to synthesize both linear and planar sparse arrays with interelement spacing larger than half-wavelength. The method referred to minimizes neither the dynamic range ratio nor sparsity. However, by eliminating elements with small coefficient values, DRR as well as sparsity is promoted. In [69], nonequidistant antenna positions are obtained by using alternating direction method of multipliers. It is interesting that this method takes into account mutual coupling between antenna elements. Nature-inspired optimization methods are also applied in the design of sparse antenna arrays [70].

2.5.3 Simultaneous Application of Dynamic Range Ratio and Sparsity

Dynamic range ratio and sparsity both simplify the design of antenna arrays. However, not many papers consider combining these two approaches. This topic is addressed in the synthesis of two-dimensional arrays forming shaped beam [71]. The design starts with array's elements placed on rectangular grid. An algorithm which iteratively chooses elements with zero coefficients is then applied and antenna coefficients are optimized. The method is efficient. However, it approximates the original optimization problem by using convex relaxations. Furthermore, it does not ensure an explicit control of DRR.

Chapter 7 of this dissertation discusses optimization problems which simultaneously consider DRR and sparsity. In addition, this chapter presents a method for global design of sparse antenna arrays with constrained dynamic range ratio [72].

3 Optimization Methods

Optimization is a set of techniques for finding the best solution under the set of rules and among many other possible solutions. From mathematical point of view, optimization aims to minimize or maximize an objective function over specified domain formed by a set of constraints [73], [74]. If $\mathbf{x} = [x_1, x_2, ..., x_N]^T \in \mathcal{R}^N$ is a vector of optimization variables defined on a set Ω, $f_0(\mathbf{x})$ is the objective function, $f_i(\mathbf{x})$, $i = 1, 2, ..., I$ and $f_k(\mathbf{x})$, $k = 1, 2, ..., K$ are the set of equality and inequality constraints that $\mathbf{x}$ must satisfy, general minimization problem can be described as

$$\begin{aligned} \underset{\mathbf{x}\in\Omega}{\text{minimize}} \quad & f_0(\mathbf{x}) \\ \text{subject to} \quad & f_i(\mathbf{x}) = 0, \quad i = 1, 2, \cdots, I \\ & f_k(\mathbf{x}) \geq 0, \quad k = 1, 2, \cdots, K \end{aligned} \tag{3.1}$$

Maximization problem is obtained if $-f_0(\mathbf{x})$ is used as the objective function. Therefore, in further text, only minimization problems will be analyzed. If $I = 0$ and $K = 0$, (3.1) becomes an unconstrained optimization problem. If the objective function $f_0(\mathbf{x})$ as well as the constraints $f_i(\mathbf{x})$, $i = 1, 2, ..., I$ and $f_k(\mathbf{x})$, $k = 1, 2, ..., K$ are convex, the problem in (3.1) becomes a convex optimization problem [75]. Convex functions satisfy

$$f\left(\alpha x_1 + (1-\alpha) x_2\right) \leq \alpha f\left(x_1\right) + (1-\alpha) f\left(x_2\right), \forall \alpha \in [0,1] \tag{3.2}$$

where x_1 and x_2 are feasible points.

Point $\mathbf{x}_{min}$ is a local solution of the optimization problem if $f_0(\mathbf{x}_{min}) \leq f_0(\mathbf{x})$ in some neighborhood of $\mathbf{x}_{min}$. On the other hand, $\mathbf{x}_{min}$ is a global solution if $f_0(\mathbf{x}_{min}) \leq f_0(\mathbf{x})$ for all $\mathbf{x} \in \Omega$. If the optimization problem is convex, a local solution is also the global solution.

When solving an engineering problem, knowing the best solution is an ultimate goal. If the problem is solved by using optimization, convex programming is preferable. However, recognizing the problem at hand as a convex problem requires special skills. On the other hand, some problems cannot be expressed in a convex form. Still, the optimization can be used to obtain an acceptable solution. For such a solution, globality cannot be proved. However, if the solution satisfies the application's requirements, it is exploited. If such is not the case, an attempt can be made to solve the problem by using global optimization.

Global optimization seeks the global solution of an optimization problem. If the problem is nonlinear and nonconvex, finding the global solution is a challenging task. This task can be rarely performed in polynomial time, especially if the problem has combinatorial nature. Global optimization includes exact and heuristic techniques. The former incorporates search over the entire domain or feasible set. Examples of this approach are the exhaustive search and branch and bound. The latter include evolutionary methods, simulated annealing, etc.

In the following sections the optimization techniques used in this dissertation will be introduced.

3.1 Linear Programming

Linear programming (LP) is a convex optimization in which the objective function is linear whereas the constraints are affine functions [75]. Such optimization problem can be expressed in a form

$$\begin{aligned} \underset{\mathbf{x}}{\text{minimize}} \quad & \mathbf{c}^T\mathbf{x} \\ \text{subject to} \quad & \mathbf{G}\mathbf{x} \leq \mathbf{f} \\ & \mathbf{A}\mathbf{x} = \mathbf{b} \end{aligned} \tag{3.3}$$

where $\mathbf{c} \in \mathcal{R}^N$, $\mathbf{G} \in \mathcal{R}^{M\text{x}N}$, $\mathbf{A} \in \mathcal{R}^{P\times N}$, $\mathbf{f} \in \mathcal{R}^M$, $\mathbf{b} \in \mathcal{R}^P$, N is the number of variables and M and P are the numbers of inequality and equality constraints, respectively. In LP problem, the minimization is performed over a convex polytope, which is defined as an intersection of subspaces formed by inequality and equality constraints. Linear programs are solved by Dantzig's simplex method or by interior-point methods. The simplex method searches for the optimum by moving systematically along the polytope's vertexes. For large scale problems, such a search is time consuming. The interior point methods converge to the optimum following the path placed inside the polytope. Such convergence is fast, what makes interior point methods suitable for solving large problems.

3.2 Quadratic Programming

Special category of convex optimization problems are quadratic programming (QP) problems, whose objective function is a convex quadratic function and constraint set is affine [75]. The QP problem is expressed as

$$\begin{aligned} \underset{\mathbf{x}}{\text{minimize}} \quad & \frac{1}{2}\mathbf{x}^T\mathbf{B}\mathbf{x} + \mathbf{c}^T\mathbf{x} \\ \text{subject to} \quad & \mathbf{G}\mathbf{x} \le \mathbf{f} \\ & \mathbf{A}\mathbf{x} = \mathbf{b} \end{aligned} \tag{3.4}$$

where $\mathbf{B} \in \mathcal{R}^{N\text{x}N}$ is positive semidefinite matrix, $\mathbf{c} \in \mathcal{R}^N$, $\mathbf{G} \in \mathcal{R}^{M\text{x}N}$, $\mathbf{A} \in \mathcal{R}^{P\times N}$, $\mathbf{f} \in \mathcal{R}^M$, and $\mathbf{b} \in \mathcal{R}^P$. Linear programs can be considered as a special case of quadratic problems when $\mathbf{B} = \mathbf{0}$.

3.3 Second-Order Cone Programming

Second-order cone programming (SOCP) is also part of convex programming. A SOCP problem is given by [75] [76]

$$\begin{aligned} \underset{\mathbf{x}}{\text{minimize}} \quad & \mathbf{f}^T\mathbf{x} \\ \text{subject to} \quad & \left\|\mathbf{A}_i\mathbf{x} + \mathbf{b}_i\right\|_2 \le \mathbf{c}_i^T\mathbf{x} + d_i, \quad i = 1, 2..., M \end{aligned} \tag{3.5}$$

where $\mathbf{f} \in \mathcal{R}^N$, $\mathbf{A}_i \in \mathcal{R}^{(P_i-1)\times N}$, $\mathbf{b}_i \in \mathcal{R}^{(P_i-1)}$, $\mathbf{c}_i \in \mathcal{R}^N$, $d_i \in \mathcal{R}$, and M is the number of constraints The constraints in (3.5) are called second-order cone constraints. They are equivalent to

$$\begin{bmatrix} \mathbf{A}_i \\ \mathbf{c}_i^T \end{bmatrix}\mathbf{x} + \begin{bmatrix} \mathbf{b}_i \\ d_i \end{bmatrix} \in C_i \tag{3.6}$$

where C_i represents a second-order (Lorentz or ice-cream) cone of dimension P_i. It is clear from (3.5) that SOCP includes LP and QP. However, it covers a broader class of problems than classic linear and quadratic programming.

3.4 Geometric and Signomial Programming

For understanding of geometric and signomial programming, it is necessary to introduce the terms of monomial, posynomial and signomial. A real valued function

$$f(\mathbf{x}) = cx_1^{q_1}x_2^{q_2}\cdots x_N^{q_N} \tag{3.7}$$

where $\mathbf{x} = [x_1, x_2, ..., x_N]^T$, $c > 0$, and $q_1, q_2, ..., q_N \in \mathcal{R}$ is called a monomial [77]. Sum of monomials

$$f(\mathbf{x}) = \sum_{k=1}^{K} c_k x_1^{q_{1k}} x_2^{q_{2k}} \cdots x_N^{q_{Nk}} \tag{3.8}$$

is called a posynomial. The name posynomial originates from the term positive polynomial.

Multiplication and division of monomials results in another monomial function. Furthermore, addition, multiplication, and positive scaling of posynomials result in a posynomial function. Finally, division of posynomial with a monomial provides posynomial.

Signomial is a function expressed with (3.8) where $c_k \in \mathcal{R}$. In this context, signomial can be considered as a generalization of posynomial.

Geometric programming (GP) is a convex programming technique expressed with

$$\begin{aligned} &\text{minimize} && f_0(\mathbf{x}) \\ &\text{subject to} && f_i(\mathbf{x}) \le 1, \quad i = 1, 2, \ldots, I \\ &&& f_k(\mathbf{x}) = 1, \quad k = 1, 2, \ldots, K \\ &&& \mathbf{x} > 0 \end{aligned} \tag{3.9}$$

where $f_0(\mathbf{x}), f_i(\mathbf{x})$, $i = 1, 2, \ldots, I$ are posynomials and $f_k(\mathbf{x})$, $k = 1, 2, \ldots, K$ are monomials [77].

Signomial programming, or more precisely signomial geometric programming (SGP), is an optimization technique that solves the problem in (3.9) where the objective function as well as inequality and equality constraints can be signomials [75], [78]. SGP problems are very difficult to solve and generally cannot be solved globally. However, efficient techniques for solving SGP problems have been developed (see [78] and the references therein). Many of them transform the SGP problem into series of GP problems thus forming iterative procedure. One such technique will be more thoroughly analyzed in this dissertation.

Techniques for convex optimization are intensively studied, resulting in several commercial and free tools. Popular tools for solving LP, QP and SOCP problems are MOSEK [81], SeDuMi [82], CVX [83], [84]. Special solvers have been developed for geometric programming, such as GPPLAB [85]. However, efficient GP solver is also included in MOSEK [81].

3.5 Branch and Bound

Optimization problems involving discrete variables usually have combinatorial complexity. Minimization of the l_0-norm can also be transformed into such problem class. Generally, combinatorial problems are difficult to solve globally because many solutions should be examined. The simplest method for examining all solutions is the exhaustive

search. However, this search cannot be performed in polynomial time, so it is suitable only for solving simple problems. To deal with real world problems, the entire procedure should be sped up. The solutions explored by an exhaustive search are often organized in a rooted tree. The time required for exploring the entire tree can be reduced by an appropriate pruning. Such an approach leads to the branch and bound (BnB) algorithm, which was first introduced in [79]. In BnB algorithm, analysis is made in each node to remove the solutions that cannot yield an optimum. Their removal also removes all solutions below the node under inspection.

Basic branch and bound algorithm is described by the pseudocode shown in Algorithm 3.1 [80]. The goal is to find $\mathbf{x}$ which minimizes the objective function, $f_0(\mathbf{x})$, assuming $\mathbf{x} \in X$, where X is a set of valid solutions. The set X is called the search space and it is represented by a rooted tree. The exploration of X is performed by analyzing the sets of subproblems $S_k \subseteq X$, $k = 0, 1, 2, ..., K$, which are represented by the subtrees of the initial tree. At the root, $S_0 = X$. The exploration starts with the Step 1 in which a list of unexplored solutions is initialized as $L = \{S_0\}$, a feasible solution $\mathbf{x}_0$ is found, and $\mathbf{x}_{opt} = \mathbf{x}_0$ is set. Step 2 is performed until all subproblems from L have been explored. In this step, a subproblem S_k is taken from L. Then, the attempt is made to find $\mathbf{x}_k \in S_k$ for which $f(\mathbf{x}_k) \leq f(\mathbf{x}_{opt})$. If such $\mathbf{x}_k$ is found, an update $\mathbf{x}_{opt} = \mathbf{x}_k$ is made. In addition, the subproblem S_k is analyzed. If S_k can be proven not to contain solution better than $\mathbf{x}_{opt}$, S_k can be pruned, that is, removed from L. Otherwise, a new set of subproblems S_r, $r = 1, 2, ..., R$ should be generated from S_k and inserted into L. After all subproblems from L have been explored, $\mathbf{x}_{opt}$ is returned as optimum. Note that index k is added here for better visualization of the procedure. However, the timeline k, that is, the iteration count is not necessary for exploring the tree.

```
 1: Initialize L = {S₀}
 2: Set x_opt = x₀
 3: While L ≠ ∅
 4:    Choose an S_k from L
 5:    If x_k ∈ S_k can be found such that f(x_k) ≤ f(x_opt)
 6:       Set x_opt = x_k
 7:    End if
 8:    If S_k cannot be pruned
 9:       Make S₁, S₂, ..., S_R from S_k
10:       Insert S₁, S₂, ..., S_R into L
11:    End if
12:    Remove S from L
13: End while
14: Return x_opt as optimum
```

Algorithm 3.1 General branch and bound algorithm described with pseudocode.

The order in which subproblems S_k are selected from the list L in step 4 of Algorithm 3.1 is called the search strategy. Common strategies include the depth-first, breadth-first, best-first and cyclic best-first approach [80].

Depth-first strategy starts at the root of the tree and explores the tree in the depth before returning in the direction of the root node again. After reaching the leaf, it continues the search from the last feasible solution higher in the tree. An example of depth-first BnB tree is given in Figure 3.1 where red numbers denote the exploration order.

Breadth-first approach starts at the root and explores all nodes at the present depth prior to moving to the next tree level. An example of breadth-first BnB tree is given in the Figure 3.2.

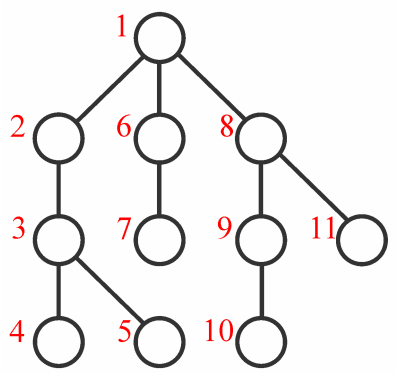

Figure 3.1 Example of depth-first branch and bound tree. Red numbers denote exploration order.

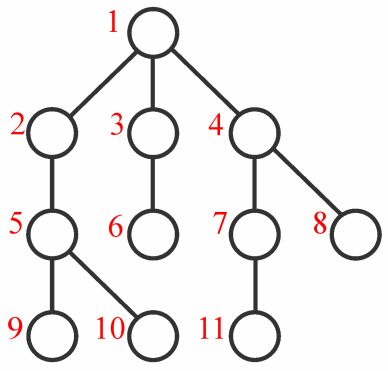

Figure 3.2 Example of breadth-first branch and bound tree. Red numbers denote exploration order.

An example of best-first search strategy is given in Figure 3.3. In this figure, red numbers indicate the exploration order whereas the blue numbers show lower bounds on the objective function achieved in particular nodes. Best-first strategy stores in memory all results of objective function's lower bound. These results are chosen to determine moving along the tree from their lower to higher values. Apparently, the best-first strategy is memory consuming and might be impractical for exploration of large trees, especially if the problem exhibits small probability of pruning.

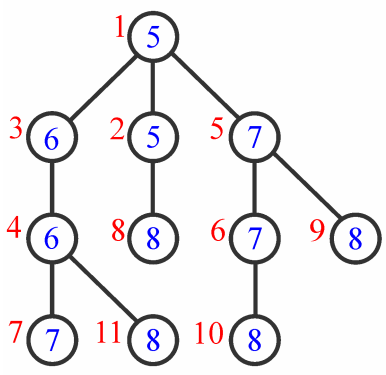

Figure 3.3 Example of best-first branch and bound tree. Red numbers denote exploration order. Blue numbers show lower bounds on objective function.

Recently, a so called cyclic-best search has been proposed, mainly for task scheduling. This strategy generalizes all the above strategies. It relies on the assumption that the problem can be solved by using parallel computing, thus being solved more efficiently.

Choosing the right approach for specific optimization problem leads to more efficient pruning and lowers computation time. Therefore, an appropriate choice of the search strategy can significantly influence the amount of computational time and memory required by the BnB procedure. Generally, this choice depends on the problem at hand.

4 Design of Sparse FIR Filters Based on Signomial Programming

As previously stated, finding optimum nonzero coefficients and their positions in sparse FIR filter design is a difficult task. Various algorithms have been proposed for its solving, most of which offer approximate or local solutions. Therefore, the research is still conducted to find a design method which provides filters whose sparsity is close to the maximum possible.

In this chapter, the method for the design of sparse FIR filters based on signomial geometric programming will be presented. The proposed SGP method will be tailored for the design of FIR filters with linear phase [49] and the filters without phase specifications [48]. Both approaches utilize filters which are constrained in a peak-error sense.

4.1 Arithmetic-Geometric Mean Approximation

Many methods for solving SGP problems are based on iterative procedures. Typically, such procedures transform the original SGP problem into sequences of GP problems. It is convenient because GP problems can be solved with available solvers. The method in [78] presents one such approach which has a global convergence. It solves general SGP problem where both objective function and constraints are non-posynomial functions. In the first step, these functions are transformed to posynomials. Afterwards, additional transformations are performed to fit them into GP framework.

In sparse FIR filter design, the transformation of the objective function to a posynomial is a straightforward process. In addition, FIR filter design does not contain equality constraints. Therefore, the method in [78] can be easily tailored to solve sparse FIR filter design problem.

Inequality constraints from (3.9) in which $f_i(\mathbf{x})$ are signomials are transformed into the following form

$$f_i^+(\mathbf{x}) - f_i^-(\mathbf{x}) \le 1, \quad i = 1, 2, \ldots, I \tag{4.1}$$

where $f_i^+(\mathbf{x})$ and $f_i^-(\mathbf{x})$ are posynomial functions that collect the terms defined in (3.8) with $c_k > 0$ and $c_k < 0$, respectively. The constraints in (4.1) can be expressed as

$$\frac{f_i^+(\mathbf{x})}{f_i^-(\mathbf{x})+1} \le 1, \quad i = 1,2,\ldots,I \tag{4.2}$$

Division of posynomial function by a monomial results in posynomial function. Therefore, to obtain a posynomial in the left-hand side of (4.2), its denominator needs to be approximated with a monomial. This is performed by introducing arithmetic-geometric mean inequality [77]

$$g(\mathbf{x}) \ge \hat{g}(\mathbf{x}) = \prod_v \left(\frac{u_v(\mathbf{x})}{\alpha_v(\mathbf{y})} \right)^{\alpha_v(\mathbf{y})} \tag{4.3}$$

where $g(\mathbf{x}) = f_i^-(\mathbf{x}) + 1$ for those $i = 1, 2,\ldots, I$ which are not a monomial, $\hat{g}(\mathbf{x})$ is best monomial approximation of $g(\mathbf{x})$,

$$\alpha_v(\mathbf{y}) = \frac{u_v(\mathbf{y})}{g(\mathbf{y})}, \forall v \tag{4.4}$$

and $\mathbf{y} > \mathbf{0}$.

4.2 Design of Sparse FIR Filters With Linear Phase

4.2.1 Design of Filters With Peak-Error Constraints

Optimization problem presenting design of sparse linear-phase FIR filters constrained in a peak-error sense is given by

$$\begin{aligned} &\underset{\mathbf{h}}{\text{minimize}} && \|\mathbf{h}\|_0 \\ &\text{subject to} && \left| H(\mathbf{h},\omega) - H_d(\omega) \right| \le \delta(\omega), \quad \omega \in \Omega \end{aligned} \tag{4.5}$$

where $\|\cdot\|_0$ denotes the l_0-norm, $\mathbf{h}$ is filter's impulse response, $H(\mathbf{h},\omega)$ is filter's amplitude response, $H_d(\omega)$ is desired amplitude response, $\delta(\omega)$ is acceptable upper bound of approximation error, and Ω is the union of frequency bands of interest. The problem in (4.5) is nonconvex and highly nonlinear. To solve it, it is here approximated with an l_p-norm, $0 < p < 1$, which is given by

$$\left\|\mathbf{h}\right\|_p = \left(\sum_{k=1}^{N} |h_k|^p\right)^{\frac{1}{p}} \tag{4.6}$$

The constraints in (4.5) are evaluated on a finite set of Q frequency points ω_q, $q = 1, 2, ..., Q$, $\omega_q \in \Omega$. Consequently, $d_q = H_d(\omega_q)$ and $\delta_q = \delta(\omega_q)$. Furthermore, by assuming

$$\arg\min_{\mathbf{h}} \left\|\mathbf{h}\right\|_p = \arg\min_{\mathbf{h}} \left\|\mathbf{h}\right\|_p^p \tag{4.7}$$

the problem in (4.5) can be expressed as

$$\begin{aligned} &\underset{\mathbf{h}}{\text{minimize}} && \left\|\mathbf{h}\right\|_p^p \\ &\text{subject to} && \left|H(\mathbf{h},\omega_q) - d_q\right| \le \delta_q, \quad q = 1,2,\ldots,Q \end{aligned} \tag{4.8}$$

The problem in (4.8) is suitable for the design of all types of linear phase FIR filters. However, due to simplicity, in the following text, detailed expressions will be provided only for Type I filters. Such filters are uniquely described by $K = N/2+1$ impulse response samples. Therefore, a new optimization variable is introduced as

$$\mathbf{z} = \left[z_1, z_2, \ldots, z_K\right]^T = \left[h_{K-1}, h_K, \ldots, h_N\right] \tag{4.9}$$

where $\mathbf{z}$ represents K right-hand side samples of the impulse response. By using $\mathbf{z}$, the problem in (4.8) becomes

$$\begin{aligned} &\underset{\mathbf{z}}{\text{minimize}} && |z_1|^p + 2|z_2|^p + \cdots + 2|z_K|^p \\ &\text{subject to} && \left|H\left(\mathbf{z},\omega_q\right) - d_q\right| \le \delta_q, \qquad q = 1,2,...,Q \end{aligned} \tag{4.10}$$

The variables z_k with $k = 2, 3, \ldots, K$ represent the right-hand side, but also the left-hand side impulse response samples. Therefore, they contribute to sparsity by factor 2. Amplitude response can be expressed in a matrix form as in

$$H\left(\mathbf{z},\omega_q\right) = \mathbf{A}_q\mathbf{z} \tag{4.11}$$

where $\mathbf{A}_q$ is given by

$$\mathbf{A}_q = \begin{bmatrix} 1 & 2\cos(\omega_q) & 2\cos(2\omega_q) & \dots & 2\cos((K-1)\omega_q) \end{bmatrix} \tag{4.12}$$

Some transformations of the problem in (4.10) are necessary to write it in a form suitable for SGP method as presented in Section 4.1. The first step is to ensure that all elements in **z** are positive.

4.2.1.1 Conditioning of SGP Problem

To make objective function a posynomial, the following substitution is introduced

$$\mathbf{y} = \mathbf{z} + \boldsymbol{\rho} \tag{4.13}$$

where $\boldsymbol{\rho} = [\rho, \rho, ..., \rho]^T$ is a vector with K equal positive constants which ensure that all elements in **y** are positive. By applying (4.13), the problem in (4.10) becomes

$$\begin{aligned} \underset{\mathbf{y}}{\text{minimize}} \quad & |y_1 - \rho|^p + 2|y_2 - \rho|^p + \cdots + 2|y_K - \rho|^p \\ \text{subject to} \quad & \left|H(\mathbf{y}, \omega_q) - d_q\right| \le \delta_q, \quad q = 1, 2, ..., Q \\ & \mathbf{y} > \mathbf{0} \end{aligned} \tag{4.14}$$

where

$$H(\mathbf{y}, \omega_q) = \mathbf{A}_q \mathbf{y} - \mathbf{A}_q \boldsymbol{\rho} \tag{4.15}$$

To avoid operation of the absolute value in the objective function, a vector of slack variables $\mathbf{t} = [t_1, t_2, ..., t_K]^T$ is introduced, which transforms the problem in (4.14) into

$$\begin{aligned} \underset{\mathbf{t},\mathbf{y}}{\text{minimize}} \quad & t_1^p + 2t_2^p + \cdots + 2t_K^p \\ \text{subject to} \quad & y_k - \rho \le t_k && k = 1, 2, ..., K \\ & -y_k + \rho \le t_k && k = 1, 2, ..., K \\ & \mathbf{A}_q \mathbf{y} - (\mathbf{A}_q \boldsymbol{\rho} + d_q) \le \delta_q, && q = 1, 2, ..., Q \\ & -\mathbf{A}_q \mathbf{y} + (\mathbf{A}_q \boldsymbol{\rho} + d_q) \le \delta_q, && q = 1, 2, ..., Q \\ & \mathbf{y} > \mathbf{0}, \ \mathbf{t} > \mathbf{0} \end{aligned} \tag{4.16}$$

The problem in (4.16) is SGP problem of standard form with $2K$ variables and $2(K+Q)$ inequality constraints. Constraints $\mathbf{t} > \mathbf{0}$ and $\mathbf{y} > \mathbf{0}$ are implicitly satisfied in signomial programming.

4.2.1.2 Solving of SGP Problem

To solve the problem by using the procedure described in Section 4.1, the problem in (4.16) is expressed in a compact form

$$\begin{aligned} \underset{\mathbf{x}}{\text{minimize}} \quad & x_1^p + 2x_2^p + \cdots + 2x_K^p \\ \text{subject to} \quad & f_i(\mathbf{x}) = a_{0i} + \sum_{k=1}^{2K} a_{ki} x_k \le 0 \quad i = 1, 2, ..., 2(K+Q) \\ & \mathbf{x} > \mathbf{0} \end{aligned} \tag{4.17}$$

where

$$\mathbf{x} = \left[x_1, x_2, ..., x_{2K}\right]^T = \begin{bmatrix} \mathbf{t} \\ \mathbf{y} \end{bmatrix} \tag{4.18}$$

It should be noted that constraints are left in the form $f_i(\mathbf{x}) \le 0$ rather than $f_i(\mathbf{x}) \le 1$, which is required by the signomial program. However, such notation will simplify further procedure. The procedure from Section 4.1 applied to the problem in (4.17) can be described with Algorithm 4.1, where r is algorithm iteration counter.

```
0: Choose feasible point x^(0) and solution accuracy η > 0
1: Set r = 0
2: Approximate original SGP problem with a GP problem around x^(r)
3: Solve the GP problem to obtain x^(r+1)
4: If ||x^(r+1)−x^(r)|| > η
      Set r = r+1 and go to step 2
5: End if
6: Return x_min = x^(r+1)
```

Algorithm 4.1 Pseudocode of SGP method for sparse FIR filter design.

The Step 2 of the algorithm contains the approximation of SGP problem in (4.17) with a GP problem. This transformation involves only modifications of the problem constraints, because the objective function in (4.17) is already a posynomial. Each constraint $f_i(\mathbf{x}) \le 0$ is first written in a form

$$f_i^+(\mathbf{x}) \le f_i^-(\mathbf{x}) \tag{4.19}$$

Note that a_{0i} is also a part of $f_i^+(\mathbf{x})$ or $f_i^-(\mathbf{x})$, depending on its sign. Now, $f_i^-(\mathbf{x})$ is approximated with

$$\hat{g}_i(\mathbf{x}) = \prod_{k \in P_i} \left(\frac{u_{ki}(\mathbf{x})}{\alpha_{ki}\left(\mathbf{x}^{(r)}\right)} \right)^{\alpha_{ki}\left(\mathbf{x}^{(r)}\right)} \tag{4.20}$$

where $u_{ki}(\mathbf{x})$ is an kth term in the ith constraint, $P_i \subset \{0, 1, \ldots, 2K\}$ is set of indices k which are present in $f_i^-(\mathbf{x})$ and

$$\alpha_{ki}\left(\mathbf{x}^{(r)}\right) = \frac{u_{ki}\left(\mathbf{x}^{(r)}\right)}{f_i^-\left(\mathbf{x}^{(r)}\right)} \tag{4.21}$$

By using (4.20) and (4.21), inequality in (4.19) is obtained in a form

$$f_i^+(\mathbf{x}) \le \hat{g}_i(\mathbf{x}) \tag{4.22}$$

Now, the optimization problem takes the form

$$\begin{aligned} \underset{\mathbf{x}}{\text{minimize}} \quad & x_1^p + 2x_2^p + \cdots + 2x_K^p \\ \text{subject to} \quad & f_i^+(\mathbf{x}) \cdot \hat{g}_i^{-1}(\mathbf{x}) \le 1, \qquad i = 1, 2, \ldots, 2(K+Q) \\ & \mathbf{x} > \mathbf{0} \end{aligned} \tag{4.23}$$

The objective function and the constraints in (4.23) are posynomials. Consequently, this problem is a standard GP problem. It can be easily solved with commercial solvers such as MOSEK [81] and GGPlab [85].

GP solvers handle only positive optimization variables, but l_p-norm pushes some of the variables in $\mathbf{t}$ to zero. Commercial solvers can cope with this irregularity correctly. However, optimization can be sped up if the constraints $\mathbf{t} \ge \boldsymbol{\mu}$, where $\boldsymbol{\mu} = [\mu, \mu, \ldots, \mu]^T$, are introduced into optimization problem (4.16) instead of $\mathbf{t} \ge \mathbf{0}$. A sufficiently small μ also promotes sparsity. Final problem is expressed as

$$\begin{aligned} \underset{\mathbf{x}}{\text{minimize}} \quad & x_1^p + 2x_2^p + \cdots + 2x_K^p \\ \text{subject to} \quad & f_i^+(\mathbf{x}) \cdot \hat{g}_i^{-1}(\mathbf{x}) \le 1 \qquad i = 1, 2, \ldots, 2(K+Q) \\ & \mu x_k^{-1} \le 1 \qquad k = 1, 2, \ldots, K \\ & x_k > 0 \qquad k = K+1, \ldots, 2K \end{aligned} \tag{4.24}$$

The procedure in Algorithm 4.1 requires feasible initial point. An initial point $\mathbf{x}_0$ is made of two parts. First is $\mathbf{t}_0$ which measures sparsity. Second is $\mathbf{y}_0$ which describes the initial filter coefficients. The initial filter, with coefficients $\mathbf{h}_0$, which is optimum in minimax sense, can be obtained by a conventional design method, such as Parks-McClellan. An initial $\mathbf{z}$, denoted with $\mathbf{z}_0$, is obtained from $\mathbf{h}_0$ by using (4.9). If $\mathbf{z}_0$ contains positive and negative coefficients, the value of ρ is obtained as

$$\rho = -2 \cdot \min\left(\mathbf{z}_0\right) \tag{4.25}$$

Initial $\mathbf{t}_0$ is obtained as

$$\mathbf{t}_0 = 1.1 \cdot \left|\mathbf{z}_0\right| + \boldsymbol{\mu} \tag{4.26}$$

Finally, initial $\mathbf{y}_0$ is calculated by using (4.13).

4.2.1.3 Choice of L_p-Norm

Proposed SGP method utilizes the l_p-norm with $0 < p < 1$. The values closer to zero ensure higher sparsity but decrease the convergence rate. On the other hand, the values closer to one increase the convergence rate but do not enforce sparsity well. Here, the value of p is chosen experimentally. Figure 4.1 shows three different lowpass filters obtained for various values of p. All filters are obtained with the order $N = 70$, passband and stopband edge frequencies $\omega_p = 0.3\pi$ and $\omega_s = 0.5\pi$, and maximum passband error $\delta_p = 0.001$ dB. The maximum stopband errors δ_s of -60 dB, -70 dB, and -80 dB are chosen. The solution accuracy $\eta = 1\text{e}{-8}$ is used. It is evident from the figure that the number of nonzero coefficients decreases with a decrease in p. Clearly, $p = 0.1$ is sufficient for ensuring minimum number of nonzero coefficients in all cases. Similar behavior has been encountered in a large number of optimization runs. Therefore, in further examples the value of $p = 0.1$ will be used unless otherwise noted.

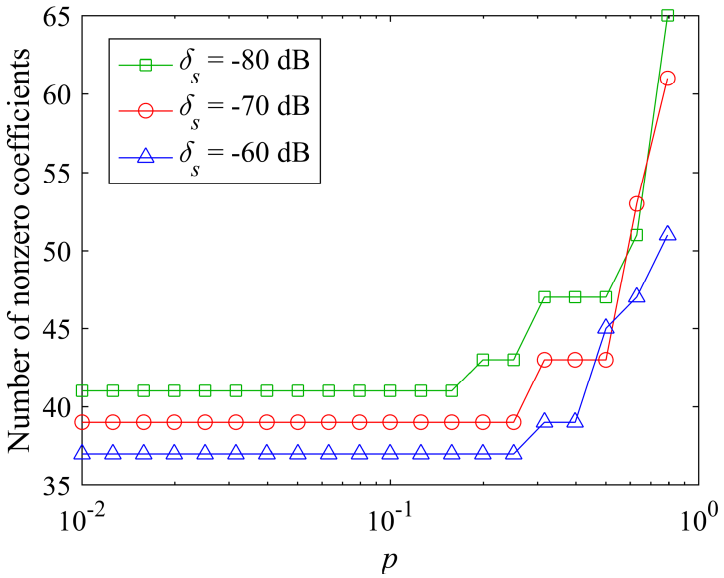

Figure 4.1 Number of nonzero coefficients obtained for three different lowpass filters and various values of p.

4.2.1.4 Refinement of Magnitude Response

The presented method minimizes sparsity, provided the obtained filter satisfies given specifications. However, after the positions of nonzero elements are obtained, additional refinement of magnitude response can be made. Positions of zero and nonzero coefficients are obtained from $\mathbf{z}$ as $S^0 = \{k : |z_k| < 1.01\mu,\ k \in \{1, 2, ..., K\}\}$ and $S^1 = \{1, 2, ..., K\} \backslash S^0$. The refinement of the magnitude response is obtained by solving the problem

$$\begin{aligned} \underset{\xi,\, z_k,\, k\in S^1}{\text{minimize}} \quad & \xi \\ \text{subject to} \quad & \frac{1}{\delta_q}\left|\mathbf{A}_q\mathbf{z} - d_q\right| \le \xi, \quad q = 1,2,...,Q \\ & z_k = 0, \qquad\qquad k \in S^0 \end{aligned} \tag{4.27}$$

where ξ is new optimization variable describing maximum approximation error and the reciprocal of δ_q is used as a weighting function.

4.2.2 Design Examples

In the following examples $Q = 10N$ and $\eta = 1\text{e}{-8}$ are used. The value for μ is set to 1e−8.

4.2.2.1 Example 1

In the first example, lowpass filters of order 60, 70 and 80 are designed with passband and stopband edge frequencies $\omega_p = 0.3\pi$ and $\omega_s = 0.5\pi$, maximum passband error $\delta_p = 0.001$ dB, and maximum stopband errors δ_s of −60 dB, −65 dB, −70 dB, −75 dB, and −80 dB. Initial filters are designed with Matlab function *firpm* which implements Parks-McClellan algorithm, resulting in filters optimum in a minimax sense. In this example, filters with equal weights in the passband and stopband are used. Table 4.1 shows the zero (marked with 0) and nonzero (marked with 1) coefficient positions of the obtained impulse responses. Due to the symmetry, only the right-hand sides of the impulse responses are shown. Furthermore, the length of the impulse response, W_{IR}, and the total number of nonzero coefficients, L_{NZ}, are given and compared with the total number of nonzero coefficients in filters presented in [39], which are obtained with the same specifications. As it is clear from the table, the proposed method achieves higher sparsities in six cases, which are marked in bold. In other cases, the proposed method provides sparsities which are equal to that obtained by the method in [39]. It is interesting to observe that lengths of the impulse responses are low and that by increasing of filter order they remain the same.

Table 4.1 Optimum zero and nonzero coefficient positions, length of impulse response, W_{IR}, and number of nonzero coefficients, L_{NZ}, of proposed filters and the filters in [39], obtained for various filter orders N, maximum passband error $\delta_p = 0.001$ dB, and various stopband approximation errors δ_s.

N	δ_s, dB	Optimum positions of zero and nonzero coefficients	W_{IR}	L_{NZ}	L_{NZ} [39]
60	–60	11111011110111101111101000010000	53	37	37
70	–60	11111011110111101111010000100000000	53	37	37
80	–60	11111011110111101111010000100000000000000	53	37	37
60	–65	11111011110111101111101000010000	53	**37**	39
70	–65	11111011110111101111010000100000000	53	**37**	39
80	–65	11111011110111101111010000100000000000000	53	**37**	39
60	–70	11111011110111101111101100010000	53	39	39
70	–70	11111011110111101111011000100000000	53	39	39
80	–70	11111011110111101111011000100000000000000	53	39	39
60	–75	11111011110111101111101100010000	53	**39**	41
70	–75	11111011110111101111011000100000000	53	**39**	41
80	–75	11111011110111101111011000100000000000000	53	**39**	41
60	–80	11111011110111101111101111000000	49	41	41
70	–80	11111011110111101111011110000000000	49	41	41
80	–80	11111011110111101111011110000000000000000	49	41	41

Figure 4.2 shows the absolute values of impulse response samples obtained after solving SGP problem for the filter from Table 4.1 with $N = 60$ and $\delta_s = –65$ dB. It is clear that the $l_{0.1}$-norm forces sparsity well by pushing certain coefficients towards μ. Clearly, the coefficients which are near μ can be neglected. After neglecting them and refining the magnitude response, the coefficients of the final filter are obtained. They are shown in Figure 4.3.

Figure 4.4 shows the corresponding magnitude response. It is shown in comparison with the response of the filter designed with Matlab function *firpm*, which has the same number of nonzero coefficients as the proposed filter. Figure 4.5 shows the enlarged passband. The proposed sparse filter has a smaller approximation error than does its nonsparse counterpart. Sparse filter has the length of the impulse response of 53 samples among which only 37 are nonzero, whereas the nonsparse filter has 37 coefficients in total. Clearly, the improvement in magnitude response is paid in longer filter's impulse response length.

Numerical values of filter's coefficients are given in Table 4.2, for convenience.

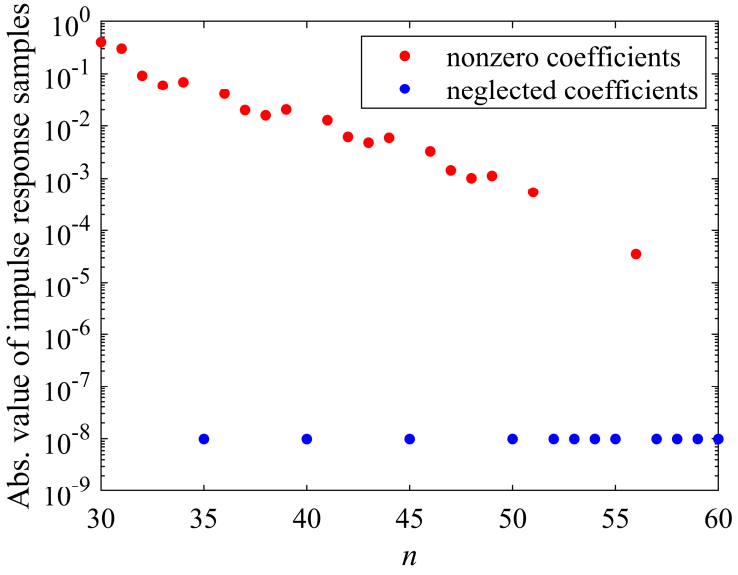

Figure 4.2 Absolute values of impulse response samples obtained after solving SGP problem for filter from Table 4.1 with $N = 60$ and $\delta_s = -65$ dB. Only right-hand side is shown.

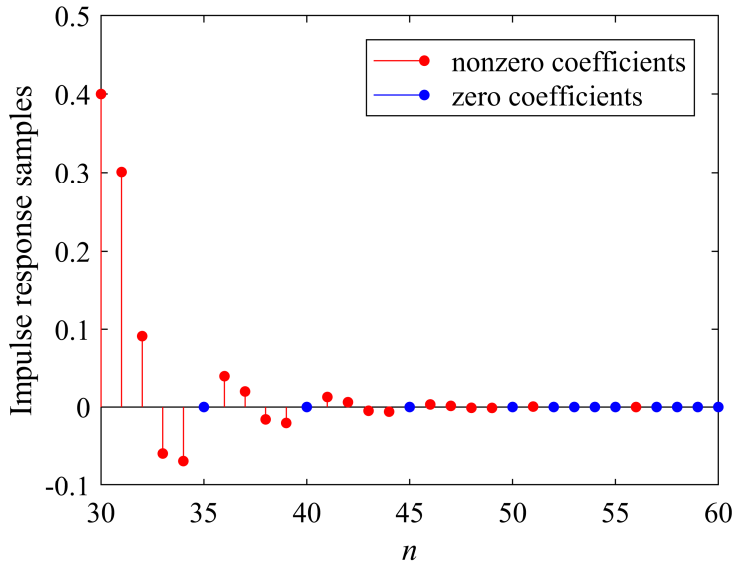

Figure 4.3 Impulse response samples of filter from Table 4.1 with $N = 60$ and $\delta_s = -65$ dB. Only right-hand side is shown.

Table 4.2 Coefficients of proposed sparse filter from Table 4.1 with $N = 60$ and $\delta_s = -65$ dB.

n	$h(n) = h(N-n)$	n	$h(n) = h(N-n)$	n	$h(n) = h(N-n)$
30	0.40000	41	0.01270	52	0
31	0.30090	42	0.00614	53	0
32	0.09130	43	–0.00477	54	0
33	–0.05905	44	–0.00589	55	0
34	–0.06865	45	0	56	–0.00004
35	0	46	0.00327	57	0
36	0.04043	47	0.00143	58	0
37	0.01974	48	–0.00101	59	0
38	–0.01571	49	–0.00112	60	0
39	–0.02025	50	0		
40	0	51	0.00054		

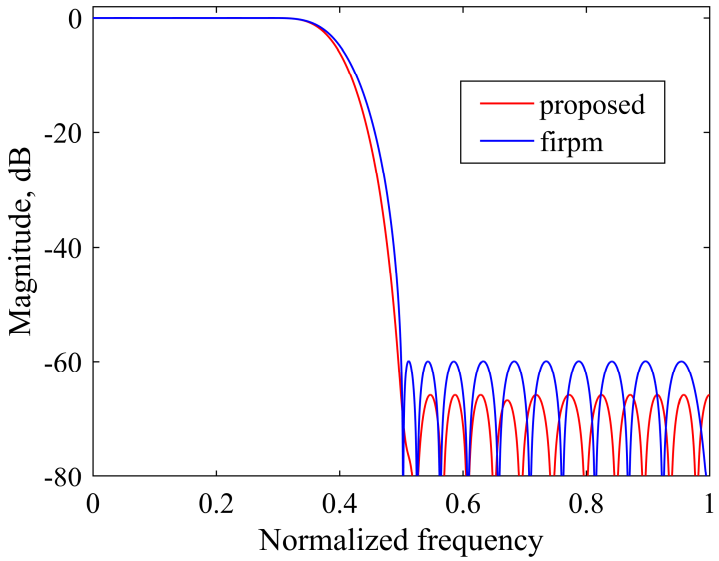

Figure 4.4 Magnitude response of sparse lowpass filter from Table 4.1 with $N = 60$ and $\delta_s = -65$ dB, compared to nonsparse minimax filter obtained with function *firpm* with same number of nonzero coefficients.

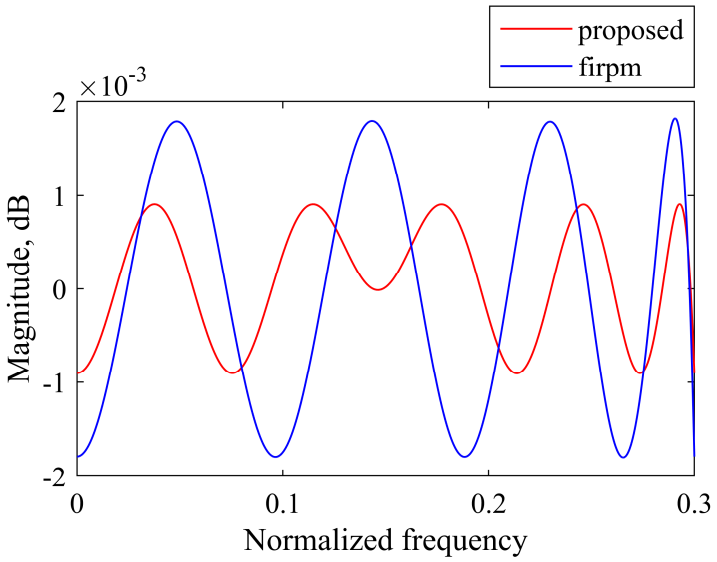

Figure 4.5 Enlarged passband of filters shown in Figure 4.4.

4.2.2.2 Example 2

In this example, passband and stopband edge frequencies as well as the maximum stopband error are kept the same as in Example 1. However, maximum passband error is increased to $\delta_p = 0.5$ dB. Initial filters are designed with function *firpm*, but here the weighting factors for the passband and stopband are set to 1 and δ_p/δ_s, respectively.

Table 4.3 shows the optimum zero and nonzero coefficient positions of the obtained impulse responses. In addition, the length of the impulse response, W_{IR}, as well as the total number of nonzero coefficients, L_{NZ}, are given. The latter parameter is compared to that obtained by the filters in [40]. In 13 out of 15 cases, which are marked in bold, the proposed method results in sparser filters then does the method in [40]. In the remaining two cases, the proposed method provides equal sparsities.

The obtained results show that in most cases the proposed SGP method is insensitive to an increase in filter order. However, some filters with higher orders result in somewhat lower sparsity. In these cases, better results could be achieved with different initial points or by adjusting input parameters. Nonetheless, the proposed method still gives much better results than the other method used in comparison.

Table 4.3 Optimum zero and nonzero coefficient positions, length of impulse response, W_{IR}, and number of nonzero coefficients, L_{NZ}, of proposed filters and the filters in [40] for various filter orders N, maximum passband error $\delta_p = 0.5$ dB, and various stopband approximation errors δ_s.

N	δ_s, dB	Optimum positions of zero and nonzero coefficients	W_{IR}	L_{NZ}	L_{NZ} [40]
60	–60	1111111101011010100000000000000	33	**25**	29
70	–60	11111111010110101000000000000000000	33	**25**	29
80	–60	111111110101101010000000000000000000000000	33	25	25
60	–65	1111111101101011000000000000000	31	**25**	29
70	–65	11111111010110110000000000000000000	31	**25**	29
80	–65	111111110101101010011000000000000000000000	41	29	29
60	–70	1111111101101011000000000000000	31	**25**	33
70	–70	11111111011010110000000000000000000	31	**25**	33
80	–70	111111110110101100000000000000000000000000	31	**25**	31
60	–75	1111111101101011010000000000000	35	**27**	33
70	–75	11111111011010110100000000000000000	35	**27**	33
80	–75	111111110110101101000000000000000000000000	35	**27**	33
60	–80	1111111101101011011000000000000	37	**29**	37
70	–80	11111111011010110101000000000000000	39	**29**	37
80	–80	111111110110101101010000000000000000000000	39	**29**	35

4.2.2.3 Example 3

This example presents the design of various lowpass filters with specifications given in [41]. The filters presented in the reffered paper are obtained by iterative method which jointly minimizes the number of nonzero coefficients and the length of the impulse response. Although the proposed SGP method does not minimize the length of the impulse response, the filters referred to are interesting because they have high filter orders, steep transition bands and low approximation errors.

In this example, the value of $p = 0.05$ is used. The initial filters are obtained by the function *firpm* assuming equally weighted passband and stopband. The obtained results are shown in Table 4.4. In this table, the first five columns contain filter number and its specifications. The total number of nonzero coefficients obtained with the proposed method, L_{NZ}, and with the method presented in [41] is shown, along with the length of the impulse response, W_{IR}. Clearly, higher sparsities are obtained for one filter, worse for two and equal for four filters.

Table 4.4 Length of impulse response, W_{IR}, and number of nonzero coefficients, L_{NZ}, of proposed filters compared to results presented in [41] for various filter orders N, passband and stopband edge frequencies, ω_p and ω_s, and passband and stopband approximation errors, $\delta_p = \delta_s$.

Filter no.	N	ω_p	ω_s	$\delta_p = \delta_s$	W_{IR}	L_{NZ}	W_{IR} [41]	L_{NZ} [41]
1	160	0.12	0.18	0.0010	115	105	112	103
2	184	0.12	0.18	0.0008	117	107	114	107
3	184	0.22	0.28	0.0010	141	85	110	85
4	200	0.22	0.28	0.0010	141	85	110	85
5	200	0.325	0.385	0.0010	139	**105**	110	107
6	200	0.325	0.385	0.0005	129	121	122	121
7	200	0.0436	0.0872	0.00023	201	187	194	183

Since the proposed SGP method does not optimize impulse response length, the obtained filters have longer responses then the filters presented in [41]. The filter no. 5 has higher sparsity then the corresponding filter in [41]. This design has been verified with initial filter of the order of 200, as well as of the order of 110, which is optimum length in [41]. In both cases, the proposed method gives 105 nonzero coefficients. The optimum coefficients obtained in the latter case are shown in Table 4.5.

Table 4.5 Coefficients of right-hand side of proposed sparse filter with $N = 110$, $\omega_p = 0.325$, $\omega_s = 0.385$ and $\delta_p = \delta_s = 0.0010$.

n	$h(n) = h(N–n)$	n	$h(n) = h(N–n)$	n	$h(n) = h(N–n)$	n	$h(n) = h(N–n)$
55	0.35504	69	0.00182	83	–0.00113	97	0.00044
56	0.28569	70	–0.01541	84	0.00467	98	–0.00113
57	0.12544	71	–0.01394	85	0.00476	99	–0.00136
58	–0.02139	72	0.00165	86	0	100	0
59	–0.07624	73	0.01316	87	–0.00413	101	0.00104
60	–0.04062	74	0.00928	88	–0.00325	102	0.00086
61	0.02049	75	–0.00365	89	0.00088	103	0
62	0.04386	76	–0.01080	90	0.00339	104	–0.00083
63	0.01835	77	–0.00566	91	0.00209	105	–0.00041
64	–0.01919	78	0.00464	92	–0.00119	106	0.00003
65	–0.02919	79	0.00862	93	–0.00263	107	0.00054
66	–0.00774	80	0.00306	94	–0.00112	108	0.00025
67	0.01745	81	–0.00487	95	0.00129	109	–0.00029
68	0.02030	82	–0.00661	96	0.00195	110	–0.00046

4.2.2.4 Example 4

Here, the bandpass linear phase FIR filter of order 160 is designed. Filter specifications are taken from [41]. Such filter has stopband edge frequencies $\omega_{s1} = 0.425\pi$ and $\omega_{s2} = 0.575\pi$, passband edge frequencies $\omega_{p1} = 0.485\pi$ and $\omega_{p2} = 0.515\pi$, and maximum passband and stopband errors $\delta_p = \delta_s = 0.01$. Since passband and stopband errors are equal, *firpm* function with equal weights is used for the initial point. The value of p is set to 0.05.

The optimum filter contains only 35 nonzero coefficients, which is also obtained in [41]. The impulse response is shown in Figure 4.6. It is clear that initial filter order can be decreased since most of zeros are positioned at the impulse response's tail. Magnitude response of the obtained filter is shown in Figure 4.7.

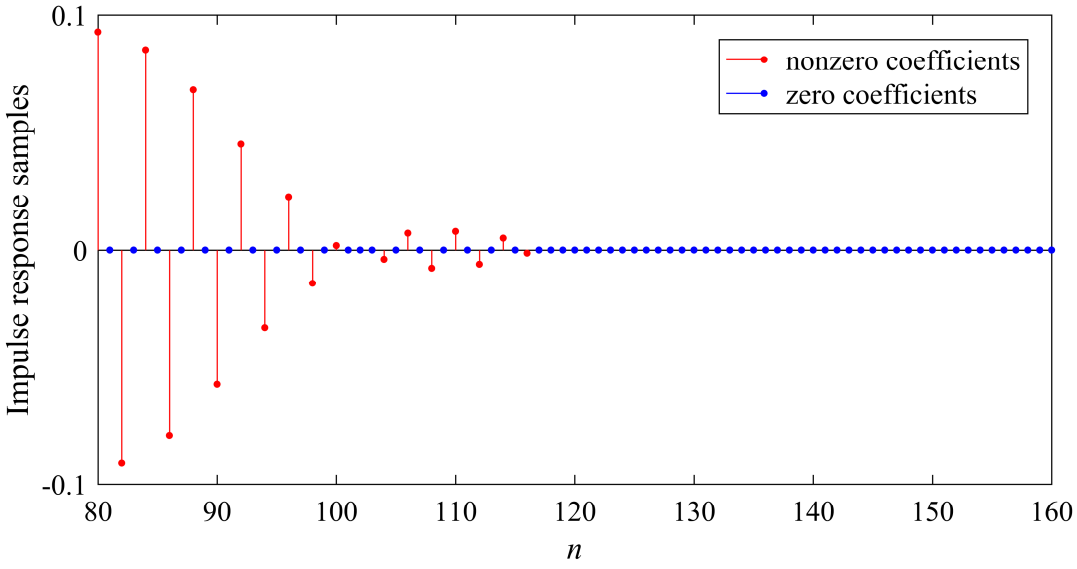

Figure 4.6 Impulse response samples of sparse bandpass filter with $N = 160$, stopband edge frequencies, $\omega_{s1} = 0.425\pi$ and $\omega_{s2} = 0.575\pi$, passband edge frequencies, $\omega_{p1} = 0.485\pi$ and $\omega_{p2} = 0.515\pi$, and passband and stopband approximation errors, $\delta_p = \delta_s = 0.01$. Only right-hand side is shown.

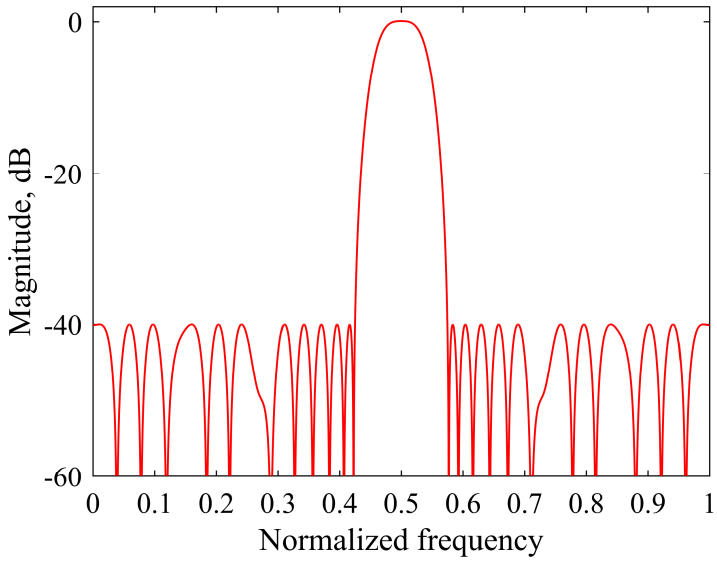

Figure 4.7 Magnitude response of sparse bandpass filter with $N = 160$, stopband edge frequencies, $\omega_{s1} = 0.425\pi$ and $\omega_{s2} = 0.575\pi$, passband edge frequencies, $\omega_{p1} = 0.485\pi$ and $\omega_{p2} = 0.515\pi$, and passband and stopband approximation errors, $\delta_p = \delta_s = 0.01$.

4.3 Design of Sparse FIR Filters Without Phase Specifications

To achieve specified selectivity, FIR filters without phase specifications require lower order than do their linear-phase counterparts. However, their realization complexity is not necessarily lower because they require one general purpose multiplier per coefficient, whereas linear phase FIR filters require one multiplier per coefficient pair. However, lowering filter's order decreases the delay of the filtered signal, what is desirable in many applications.

4.3.1 Design of Filters With Peak-Error Constraints

4.3.1.1 Optimization Problem

The design of FIR filters with magnitude response constrained in peak-error sense in which sparsity is promoted by the l_p-norm is given with (4.8). Here, this problem is generalized to enable design of filters without phase specifications. The problem is given with

$$\begin{aligned} \underset{\mathbf{h}}{\text{minimize}} \quad & \|\mathbf{h}\|_p^p \\ \text{subject to} \quad & \left|\left|H(\mathbf{h},\omega_q)\right| - d_q\right| \le \delta_q, \quad q = 1,2,\ldots,Q \end{aligned} \tag{4.28}$$

where $H(\mathbf{h},\omega_q)$, $q = 1, 2,\ldots, Q$, is filter's frequency response, d_q is desired frequency response and δ_q is required peak-error.

Each constraint in (4.28) can be rewritten by using two constraints, as

$$\left|H\left(\mathbf{h},\omega_q\right)\right| \le d_q + \delta_q \tag{4.29}$$

$$\left|H\left(\mathbf{h},\omega_q\right)\right| \ge d_q - \delta_q \tag{4.30}$$

The desired amplitude for a lowpass filter is given by

$$d_q = \begin{cases} 1 & \text{for} \quad q = 1,2,...,Q_p \\ 0 & \text{for} \quad q = Q_p+1, Q_p+2,...,Q \end{cases} \tag{4.31}$$

where Q_p is the number of frequency points in the passband. Clearly, since $d_q = 0$ in the stopband, the constraint (4.30) can be omitted while specifying this region. Now, the design of sparse lowpass FIR filters without phase specifications can be expressed by the problem

$$\begin{aligned} \underset{\mathbf{h}}{\text{minimize}} \quad & |h_0|^p + |h_1|^p + \cdots + |h_N|^p \\ \text{subject to} \quad & \left|H\left(\mathbf{h},\omega_q\right)\right| \le 1+\delta_p, \qquad q = 1,2,...,Q_p \\ & \left|H\left(\mathbf{h},\omega_q\right)\right| \ge 1-\delta_p, \qquad q = 1,2,...,Q_p \\ & \left|H(\mathbf{h},\omega_q)\right| \le \delta_s, \qquad q = Q_p+1,...,Q \end{aligned} \tag{4.32}$$

Filter's frequency response is given by

$$H\left(\mathbf{h},\omega\right) = \sum_{k=0}^{N} h_k e^{-jk\omega} = \sum_{k=0}^{N} h_k \cos\left(k\omega\right) - j\sum_{k=0}^{N} h_k \sin\left(k\omega\right) \tag{4.33}$$

Its squared magnitude response can be obtained as

$$\left|H\left(\mathbf{h},\omega\right)\right|^2 = H\left(\mathbf{h},\omega\right)H\left(\mathbf{h},-\omega\right) \tag{4.34}$$

Substituting (4.33) into (4.34) and arranging results in

$$\left|H\left(\mathbf{h},\omega\right)\right|^2 = \sum_{k=0}^{N}\sum_{i=0}^{N} h_k h_i \cos\left(k\omega\right)\cos\left(i\omega\right) + \sum_{k=0}^{N}\sum_{i=0}^{N} h_k h_i \sin\left(k\omega\right)\sin\left(i\omega\right) \tag{4.35}$$

By using simple trigonometric identities, the above expression is recognized in the form

$$\left|H\left(\mathbf{h},\omega\right)\right|^2 = \sum_{k=0}^{N}\sum_{i=0}^{N} h_k h_i \cos\left[\left(k-i\right)\omega\right] \tag{4.36}$$

The expression (4.36) is a signomial and can be easily included in a standard SGP problem. To avoid square root in calculating $H(\mathbf{h},\omega)$, the constraints in (4.32) are also squared. Optimization problem in (4.32) then takes the form

$$\begin{aligned}
\underset{\mathbf{h}}{\text{minimize}} \quad & \left|h_0\right|^p + \left|h_1\right|^p + \cdots + \left|h_N\right|^p \\
\text{subject to} \quad & \left|H\left(\mathbf{h},\omega_q\right)\right|^2 \le \left(1+\delta_p\right)^2, \quad q = 1,2,...,Q_p \\
& \left|H\left(\mathbf{h},\omega_q\right)\right|^2 \ge \left(1-\delta_p\right)^2, \quad q = 1,2,...,Q_p \\
& \left|H\left(\mathbf{h},\omega_q\right)\right|^2 \le \delta_s^2, \qquad\quad q = Q_p+1,...,Q
\end{aligned} \tag{4.37}$$

To ensure all variables included in the SGP problem are positive, the transformation

$$\mathbf{y} = \mathbf{h} + \boldsymbol{\rho} \tag{4.38}$$

is used, where $\boldsymbol{\rho} = [\rho, \rho, ..., \rho]$ is vector with $N + 1$ equal positive elements. Using (4.38), the problem (4.37) takes the form

$$\begin{aligned}
\underset{\mathbf{y}}{\text{minimize}} \quad & \left|y_0-\rho\right|^p + \left|y_1-\rho\right|^p + \cdots + \left|y_N-\rho\right|^p \\
\text{subject to} \quad & \left|H\left(\mathbf{y},\omega_q\right)\right|^2 \le \left(1+\delta_p\right)^2, \quad q = 1,2,...,Q_p \\
& \left|H\left(\mathbf{y},\omega_q\right)\right|^2 \ge \left(1-\delta_p\right)^2, \quad q = 1,2,...,Q_p \\
& \left|H\left(\mathbf{y},\omega_q\right)\right|^2 \le \delta_s^2, \qquad\quad q = Q_p+1,...,Q
\end{aligned} \tag{4.39}$$

where $|H(\mathbf{y}, \omega_q)|^2$ is calculated by using (4.36) as

$$\left|H\left(\mathbf{y},\omega_q\right)\right|^2=\sum_{k=0}^{N}\sum_{i=0}^{N}\left(y_k-\rho\right)\left(y_i-\rho\right)\cos\left[\left(k-i\right)\omega_q\right] \tag{4.40}$$

Like in the Section 4.2, here the problem in (4.39) is equivalent to

$$\begin{aligned}
\underset{\mathbf{y},\mathbf{t}}{\text{minimize}}\quad & t_0^p+t_1^p+\cdots+t_N^p & \\
\text{subject to}\quad & y_k-\rho\le t_k, & k=0,1,...,N\\
& -y_k+\rho\le t_k, & k=0,1,...,N\\
& \left|H\left(\mathbf{y},\omega_q\right)\right|^2\le\left(1+\delta_p\right)^2, & q=1,2,...,Q_p\\
& \left|H\left(\mathbf{y},\omega_q\right)\right|^2\ge\left(1-\delta_p\right)^2, & q=1,2,...,Q_p\\
& \left|H\left(\mathbf{y},\omega_q\right)\right|^2\le\delta_s^2, & q=Q_p+1,...,Q\\
& \mu t_k^{-1}\le 1, & k=0,1,...,N\\
& y_k>0, & k=0,1,...,N
\end{aligned} \tag{4.41}$$

The problem in (4.41) contains 2(N+1) optimization variables given by

$$\mathbf{x}=\left[x_1,x_2,...,x_{2(N+1)}\right]^T=\begin{bmatrix}\mathbf{t}\\ \mathbf{y}\end{bmatrix} \tag{4.42}$$

and 3(N+1) + Q_p + Q inequality constraints. The value of ρ is obtained by using

$$\rho=-2\cdot\min\left(\mathbf{h}_0\right) \tag{4.43}$$

where $\mathbf{h}_0$ are initial point coefficients obtained with Matlab function for nonsparse general filter design, *firlpnorm*. Initial $\mathbf{t}_0$ is obtained as

$$\mathbf{t}_0=1.1\cdot\left|\mathbf{h}_0\right|+\boldsymbol{\mu} \tag{4.44}$$

and initial $\mathbf{y}_0$ is calculated from (4.38) by using $\mathbf{h}_0$ and $\boldsymbol{\rho}$. The design problem in (4.41) utilizes filter's magnitude rather than the amplitude response samples. Therefore, filters with positive and negative DC gain can be obtained. To avoid the solutions with negative gain, additional constraints might be added. However, changing the coefficient signs after the optimization is more practical than adding additional constraints into the original problem.

4.3.1.2 Flipping Impulse Response in Time

The optimization problem (4.41) does not consider the phase response. Therefore, the same amplitude response corresponds to an impulse response and its time reversal. However, one of these responses introduces larger delay. Which of them is obtained as the result of optimization is not known in advance. Therefore, after the optimization, the delay of the resulting as well as of the reversed response should be checked and an appropriate choice should be made.

4.3.1.3 Complexity of Design

Generally, an SGP optimization problem for the design of FIR filters without phase specifications is more complex than the problem for the design of linear phase filters of the same order. It is a consequence of larger number of free variables, as well as the more complex expressions for magnitude response. Therefore, the maximum order which is still possible to design in an acceptable time is lower for the filters without phase specifications. However, such filters generally require lower orders than their linear phase counterparts.

4.3.2 Design Examples

In the following examples $Q = 10N$, $\eta = 1e-5 \cdot \sqrt{N}$, $p = 0.1$ and $\mu = 1e{-}5$ are used.

4.3.2.1 Example 1

In the first example, lowpass filter of initial order of 34 with passband and stopband edge frequencies, $\omega_p = 0.6\pi$ and $\omega_s = 0.7\pi$, and maximum passband and stopband errors $\delta_p = \delta_s = 0.01$, is designed. For the starting point, initial filter designed with Matlab function *firlpnorm* is used. This function returns a filter whose ripple is 0.00712, which is smaller than required δ_p and δ_s. It makes this filter a good candidate for generation of starting point for the optimization procedure.

Figure 4.8 shows the impulse response of the obtained sparse filter, which contains 29 nonzero and 6 zero coefficients. This filter is compared to nonsparse FIR filters obtained with Matlab functions *firlpnorm* and *firpm* which have the same number of nonzero coefficients. Their impulse responses are shown in Figures 4.9 and 4.10, respectively. Evidently, the impulse responses of nonsparse filters are shorter than the response of the proposed sparse filter. However, the advantage can be noticed by observing the magnitude responses. Figure

4.11 shows magnitude responses of both sparse and nonsparse filters. Furthermore, Figure 4.12 shows the enlarged passband of these filters. The sparse filter exhibits 40 dB of stopband attenuation, as required in the design process, whereas the nonsparse filter obtained with *firlpnorm* has the attenuation of 37.5 dB. Furthermore, nonsparse filter obtained with *firpm* has the worst attenuation, which equals 30.2 dB. In the passband, the sparse filter achieves the ripple of 0.010, as required, whereas the nonsparse filters with nonlinear and linear phase achieve 0.0132 and 0.031, respectively. Table 4.6 shows numeric values of the coefficients of the proposed sparse filter.

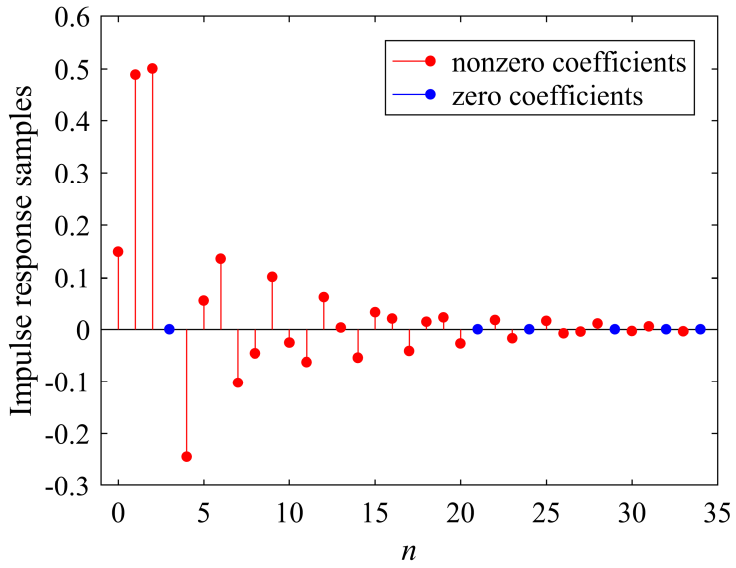

Figure 4.8 Impulse response samples of sparse lowpass filter with $N = 34$, passband and stopband edge frequencies, $\omega_p = 0.6\pi$ and $\omega_s = 0.7\pi$, and passband and stopband approximation errors, $\delta_p = \delta_s = 0.01$.

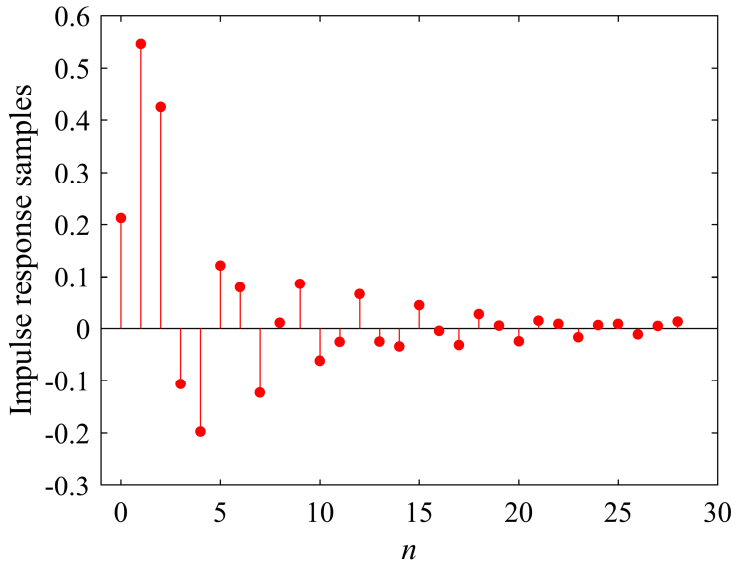

Figure 4.9 Impulse response of nonsparse lowpass filter with 29 nonzero samples, passband and stopband edge frequencies, $\omega_p = 0.6\pi$ and $\omega_s = 0.7\pi$, and passband and stopband approximation errors, $\delta_p = \delta_s = 0.01$, obtained with Matlab function *firlpnorm*.

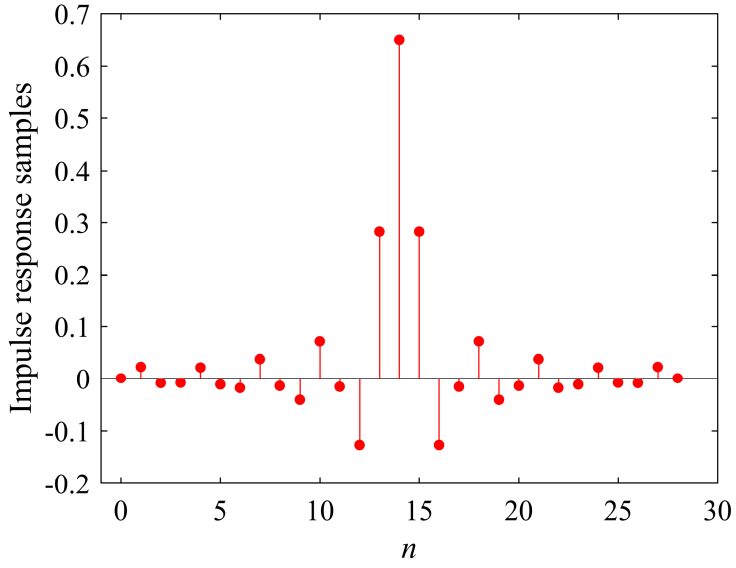

Figure 4.10 Impulse response of nonsparse lowpass filter with 29 nonzero samples, passband and stopband edge frequencies, $\omega_p = 0.6\pi$ and $\omega_s = 0.7\pi$, and passband and stopband approximation errors, $\delta_p = \delta_s = 0.01$, obtained with Matlab function *firpm*.

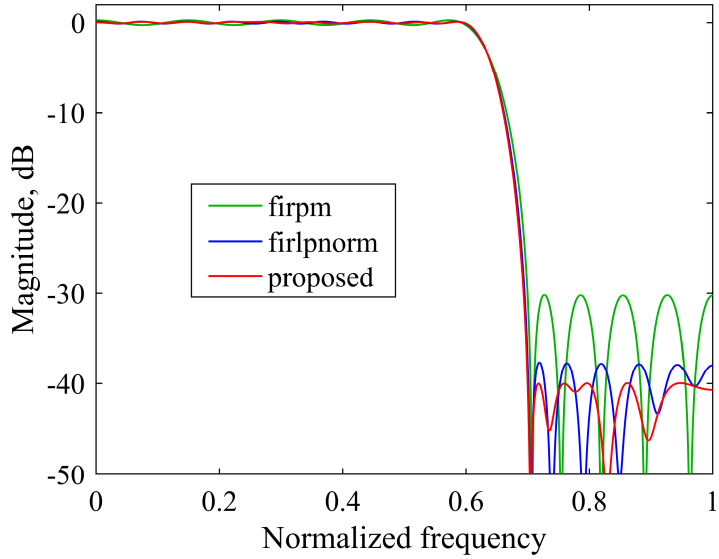

Figure 4.11 Magnitude response of optimum sparse FIR filter compared with nonsparse filters obtained by using Matlab functions *firlpnorm* and *firpm*.

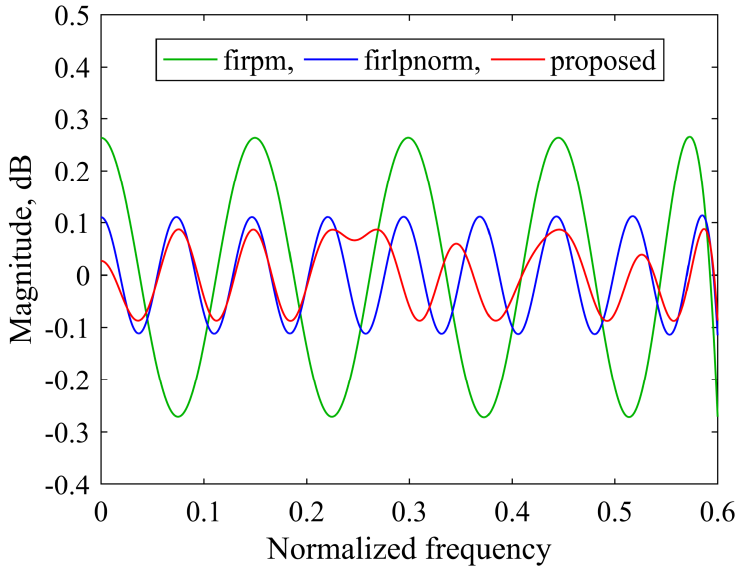

Figure 4.12 Passband magnitude of sparse FIR filter compared with nonsparse filters obtained by using Matlab functions *firlpnorm* and *firpm*.

Table 4.6 Coefficients of proposed sparse filter with $N = 34$, passband and stopband edge frequencies, $\omega_p = 0.6\pi$ and $\omega_s = 0.7\pi$, and passband and stopband approximation errors, $\delta_p = \delta_s = 0.01$.

n	$h(n)$	n	$h(n)$	n	$h(n)$
0	0.14963	12	0.06180	24	0
1	0.48947	13	0.00340	25	0.01599
2	0.50081	14	–0.05399	26	–0.00768
3	0	15	0.03301	27	–0.00447
4	–0.24424	16	0.02036	28	0.01136
5	0.05514	17	–0.04117	29	0
6	0.13643	18	0.01477	30	–0.00299
7	–0.10214	19	0.02270	31	0.00585
8	–0.04572	20	–0.02697	32	0
9	0.10187	21	0	33	–0.00346
10	–0.02530	22	0.01788	34	0
11	–0.06217	23	–0.01707		

4.3.2.2 Example 2

In this example, eight filters with initial order of 34 with passband edge frequencies $\omega_p \in \{0.1\pi, 0.2\pi, ..., 0.8\pi\}$, transition band of 0.1π, and maximum approximation errors, $\delta_p = \delta_s = 0.01$, are designed. The filter number 6 was already included into the analysis in Example 1.

Optimum positions of zero and nonzero filter coefficients together with total number of nonzero coefficients are given in Table 4.7. It is interesting to observe that most of the zero coefficients are placed at the end of the impulse response, having in mind that the filters with shorter impulse response of two possible solutions are chosen. Also, some filters have zero as the last coefficient thus indicating that initial filter order can be decreased.

Comparison of maximum stopband magnitude of the proposed filters and nonsparse filters with the same number of nonzero coefficients obtained with Matlab functions *firlpnorm* and *firpm* is given in Figure 4.13. Clearly, the highest stopband attenuation is achieved with sparse design in all cases.

Table 4.7 Optimum positions of zero and nonzero coefficients and total number of nonzero coefficients, L_{NZ}, of proposed sparse filters with $N = 34$, passband edge frequencies $\omega_p \in \{0.1\pi, 0.2\pi, ..., 0.8\pi\}$ and transition band of 0.1π.

Filter no.	ω_p	Optimum positions of zero and nonzero coefficients	L_{NZ}
1	0.1	11111111011111111111111111011111111	33
2	0.2	11111111111111111111101110111100110	30
3	0.3	11111101111111111111111111111101110	32
4	0.4	11111111001111011011111111010111101	28
5	0.5	11111111111111111111010111111101010	30
6	0.6	11101111111111111111011011111011010	29
7	0.7	11111111111111101111111111110001011	30
8	0.8	11111111111111111111111100111100001	29

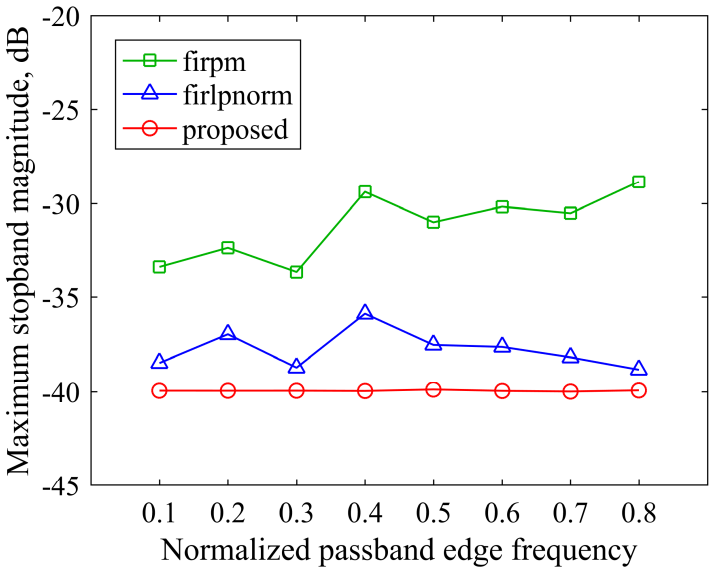

Figure 4.13 Maximum stopband magnitude of filters with $N = 34$, passband edge frequencies $\omega_p \in \{0.1\pi, 0.2\pi, ..., 0.8\pi\}$ and transition band of 0.1π, compared to nonsparse filters obtained with Matlab functions *firlpnorm* and *firpm*.

5 Design of Sparse FIR Filters Based on Global Optimization

In Chapter 4, the method based on signomial programming was proposed for the design of sparse FIR filters constrained in peak-error sense. Although this method is efficient, it cannot guarantee solutions globality. However, knowing the global solution is important not only from application point of view but also for the evaluation of nonglobal methods. So far, the only technique which can provide global optimum in sparse filter design incorporates tree search. Unfortunately, basic exhaustive search enables only the design of low-order filters. A real-word design of medium-order filters is possible with the assistance of branch and bound algorithm.

The application of branch and bound algorithm in sparse FIR filter design was proposed in [24]. The method referred to utilizes depth-first search with pruning based on the feasibility check. This method was improved in [21] where efficient estimation of lower bounds was presented. However, only the filters with quadratic constraints were considered.

In this chapter, a method for global optimization of sparse linear-phase FIR filters is proposed. The method is based on branch and bound algorithm which utilizes the depth-first and the combination of depth-first and breadth-first search. Furthermore, it incorporates two types of improved pruning. Finally, it covers the design of sparse filters constrained in peak-, quadratic-error, as well as in both senses. Proposed global optimization method works fast for small to medium filter orders, while for higher filter orders methods such as SGP presented in the Chapter 4 are more convenient. Design examples will show comparison of both methods.

5.1 Problem Formulation

Sparse linear-phase FIR-filter design which approximates the frequency response in a peak-error sense is described with the problem in (4.5). This problem is repeated here for convenience as

$$\begin{aligned} &\underset{\mathbf{h}}{\text{minimize}} && \left\|\mathbf{h}\right\|_0 \\ &\text{subject to} && \left|H\left(\mathbf{h},\omega\right)-H_d\left(\omega\right)\right| \le \delta\left(\omega\right), \quad \omega \in \Omega \end{aligned} \tag{5.1}$$

where $\|\cdot\|_0$ is the l_0-norm, $\mathbf{h}$ is filter's impulse response, $H(\mathbf{h},\omega)$ is filter's amplitude response, $H_d(\omega)$ is desired amplitude response, $\delta(\omega)$ is acceptable upper bound of approximation error, and Ω is the union of frequency bands of interest.

Linear-phase FIR filters constrained in peak-error sense are the most frequently used digital filters. However, some applications prefer energy criterion, in which the frequency response is optimized in l_2 sense. Sparse design of such filters is given by the optimization problem

$$\begin{aligned} &\underset{\mathbf{h}}{\text{minimize}} && \|\mathbf{h}\|_0 \\ &\text{subject to} && \frac{1}{\pi}\int_0^{\pi} W(\omega)\left[H(\mathbf{h},\omega)-H_d(\omega)\right]^2 d\omega \le \gamma \end{aligned} \tag{5.2}$$

where $W(\omega)$ is a positive weighting function and γ is maximum acceptable l_2 approximation error.

Filters that constrain the total energy in care bands exhibit nonequal ripples which increase near the cutoffs. To avoid such behavior, maximum peak approximation error is often incorporated. Sparse design of such filters is described by an optimization problem which combines the constraints from (5.1) and (5.2), as in

$$\begin{aligned} &\underset{\mathbf{h}}{\text{minimize}} && \|\mathbf{h}\|_0 \\ &\text{subject to} && \left|H(\mathbf{h},\omega)-H_d(\omega)\right| \le \delta(\omega) \\ & && \frac{1}{\pi}\int_0^{\pi} W(\omega)\left[H(\mathbf{h},\omega)-H_d(\omega)\right]^2 d\omega \le \gamma \end{aligned} \tag{5.3}$$

The problems in (5.1), (5.2) and (5.3) cover all types of FIR filters. However, in further text, only the method for the design of Type 1 FIR filters will be described. The expressions necessary for the design of other filter types can be obtained similarly.

Type 1 FIR filter is uniquely described with $K = N/2 + 1$ coefficients. To determine which of these coefficients can be set to zero, a tree is built in which each leaf represents one possible combination of zero and nonzero coefficients. To avoid the examination of a huge number of tree nodes, a branch and bound method is introduced. The branch and bound algorithms rely on pruning, which cuts tree branches that cannot yield the optimum. In the proposed method, the pruning is based on so called feasibility test. In the next section, this test will be evaluated for filters with peak as well as quadratically constrained magnitude

response. Then, the branch and bound procedure will be considered with the search strategy and presumed sparsity points of view.

5.2 Feasibility Test

In an optimization problem, the region above which the search for the optimum is performed is defined by a set of constraints, as in

$$f_i(\mathbf{x}) \le 0, \quad i = 1, 2, ..., I \tag{5.4}$$

If this region is empty, the problem is called infeasible. The feasibility can be tested by solving the problem [75]

$$\begin{aligned} &\underset{\mathbf{x},s}{\text{minimize}} \quad s \\ &\text{subject to} \quad f_i(\mathbf{x}) \le s, \quad i = 1, 2, ..., I \end{aligned} \tag{5.5}$$

The problem (5.5) is known as the *phase I* optimization problem because it is usually used to find initial solution in many interior point methods for convex optimization [75]. Here, it is used only as a test of feasibility. The problem is feasible if and only if $s \le 0$. If $s = 0$, the problem is considered feasible, but not strictly feasible.

5.2.1 Feasibility Test for FIR Filters With Peak-Error Constraints

Since the filter has a linear phase, its coefficients are symmetrical. Therefore, a new variable **z** is introduced which collects only the right-hand side coefficients, as shown in (4.9). Furthermore, the constraints in (5.1) are evaluated on a finite set of Q equidistant frequency points ω_q, $q = 1, 2, ..., Q$, $\omega_q \in \Omega$. The optimization problem then takes the form

$$\begin{aligned} &\underset{\mathbf{z}}{\text{minimize}} \quad \|\mathbf{z}\|_0 \\ &\text{subject to} \quad \left| H(\mathbf{z}, \omega_q) - d_q \right| \le \delta_q, \quad q = 1, 2, ..., Q \end{aligned} \tag{5.6}$$

where $d_q = H_d(\omega_q)$ and $\delta_q = \delta(\omega_q)$.

The problem (5.6) is tested for feasibility in several nodes of the tree. By applying (5.5) and matrix form of amplitude response defined in (4.11) the feasibility test takes the form

$$\begin{aligned} &\underset{\mathbf{z},s}{\text{minimize}} \quad s \\ &\text{subject to} \quad \left|\mathbf{A}_q\mathbf{z} - d_q\right| \le \delta_q + s, \quad q = 1,2,...,Q \end{aligned} \tag{5.7}$$

where $\mathbf{A}_q$ is defined in (4.12). Furthermore, in certain nodes of the tree, some coefficients can take zero values. To collect such coefficients, a vector $\mathbf{n}$ is extracted from $\mathbf{z}$ by taking only its nonzero coefficients. Then, the response $H(\mathbf{n},\omega_q)$ can be written in a matrix form as

$$H\left(\mathbf{n},\omega_q\right) = \mathbf{U}_q\mathbf{n} \tag{5.8}$$

where $\mathbf{U}_q$ is matrix composed of the columns of $\mathbf{A}_q$ that correspond to nonzero coefficients in $\mathbf{z}$. Finally, the optimization problem for feasibility testing takes the form

$$\begin{aligned} &\underset{\mathbf{n},s}{\text{minimize}} \quad s \\ &\text{subject to} \quad \left|\mathbf{U}_q\mathbf{n} - d_q\right| \le \delta_q + s, \quad q = 1,2,...,Q \end{aligned} \tag{5.9}$$

5.2.2 Feasibility Test for FIR Filters With Quadratic Constraints

After introducing the variable $\mathbf{z}$ given in (4.9), the problem in (5.2) takes the form

$$\begin{aligned} &\underset{\mathbf{z}}{\text{minimize}} \quad \left\|\mathbf{z}\right\|_0 \\ &\text{subject to} \quad \frac{1}{\pi}\int_0^{\pi} W\left(\omega\right)\left[H\left(\mathbf{z},\omega\right) - H_d\left(\omega\right)\right]^2 d\omega \le \gamma \end{aligned} \tag{5.10}$$

The constraint in (5.10) can be exactly expressed by filter's coefficients. If $W(\omega) = w$ and $H_d(\omega) = d$, where w and d are constants, the approximation error in the interval $\omega \in [\omega_1, \omega_2]$ can be expressed as

$$E = \frac{w}{\pi}\int_{\omega_1}^{\omega_2}\left[H\left(\mathbf{z},\omega\right) - d\right]^2 d\omega \tag{5.11}$$

Assuming the filter is of Type I, its frequency response is given by

$$H(\mathbf{z},\omega) = z_1 + 2\sum_{k=1}^{K-1} z_{k+1}\cos\left(k\omega\right) \tag{5.12}$$

By introducing (5.12) into (5.11) and squaring, the integral takes the form

$$\begin{aligned} E \;=\; & \frac{1}{\pi}\int_{\omega_1}^{\omega_2} z_1^2 d\omega + \frac{2}{\pi}\int_{\omega_1}^{\omega_2} z_1 \sum_{k=1}^{K-1} z_{k+1}\cos(k\omega) d\omega + \frac{2}{\pi}\int_{\omega_1}^{\omega_2} z_1 \sum_{q=1}^{K-1} z_{q+1}\cos(q\omega) d\omega \\ & -\frac{2}{\pi}\int_{\omega_1}^{\omega_2} z_1 d d\omega + \frac{4}{\pi}\int_{\omega_1}^{\omega_2} \sum_{k=1}^{K-1}\sum_{q=1}^{K-1} z_{k+1} z_{q+1}\cos(k\omega)\cos(q\omega) d\omega + \frac{1}{\pi}\int_{\omega_1}^{\omega_2} d^2 d\omega \\ & -\frac{2}{\pi}\int_{\omega_1}^{\omega_2} d\sum_{k=1}^{K-1} z_{k+1}\cos(k\omega) d\omega - \frac{2}{\pi}\int_{\omega_1}^{\omega_2} d\sum_{q=1}^{K-1} z_{q+1}\cos(q\omega) d\omega \end{aligned} \tag{5.13}$$

After integrating the addends in (5.13), and rearranging the results obtained, the integral can be written in a form

$$E = \mathbf{z}^T \hat{\mathbf{B}} \mathbf{z} - 2\mathbf{z}^T \hat{\mathbf{c}} + \hat{e} \tag{5.14}$$

where one element of matrix $\hat{\mathbf{B}}$ and vector $\hat{\mathbf{c}}$ is given as

$$\hat{B}_{kq} = \begin{cases} \frac{1}{\pi}(\omega_2 - \omega_1), & k = 0,\ q = 0 \\ \frac{2}{\pi q}\left[\sin(q\omega_2) - \sin(q\omega_1)\right], & k = 0,\ q = 1,2,...,K-1 \\ \frac{2}{\pi k}\left[\sin(k\omega_2) - \sin(k\omega_1)\right], & k = 1,2,...,K-1,\ q = 0 \\ \frac{1}{\pi k}\left[\sin(2k\omega_2) - \sin(2k\omega_1)\right] + \frac{2}{\pi}(\omega_2 - \omega_1), & k = q = 1,2,...,K-1 \\ \left.\begin{array}{l} \frac{2}{\pi(k-q)}\sin\left((k-q)\omega_2\right) - \sin\left((k-q)\omega_1\right) \\ + \frac{2}{\pi(k+q)}\sin\left((k+q)\omega_2\right) - \sin\left((k+q)\omega_1\right) \end{array}\right\}, & \begin{cases} k = 1,2,...,K-1 \\ q = 1,2,...,K-1 \\ k \neq q \end{cases} \end{cases} \tag{5.15}$$

$$\hat{c}_k = \begin{cases} \frac{d}{\pi}(\omega_2 - \omega_1), & k = 0 \\ \frac{2d}{\pi}\frac{\sin(k\omega_2) - \sin(k\omega_1)}{k}, & k = 1,2,...,K-1 \end{cases} \tag{5.16}$$

and constant $\hat{e}$ is calculated as

$$\hat{e} = \frac{1}{\pi} d^2 (\omega_2 - \omega_1) \tag{5.17}$$

Assuming filter's bands of interest are the union of R subbands within which $W(\omega) = [w_1, w_2, ..., w_R]$ and $H_d(\omega) = [d_1, d_2, ..., d_R]$, the total approximation error is given by

$$E_t = \mathbf{z}^T \mathbf{B} \mathbf{z} - 2\mathbf{z}^T \mathbf{c} + e \tag{5.18}$$

where

$$\mathbf{B} = \sum_{r=1}^{R} w_r \hat{\mathbf{B}}_r \tag{5.19}$$

$$\mathbf{c} = \sum_{r=1}^{R} w_r \hat{\mathbf{c}}_r \tag{5.20}$$

$$e = \sum_{r=1}^{R} w_r \hat{e}_r \tag{5.21}$$

Finally, by using (5.18), (5.19), (5.20), and (5.21) the problem for feasibility test takes the form

$$\begin{aligned} &\underset{\mathbf{z},s}{\text{minimize}} \quad s \\ &\text{subject to} \quad \mathbf{z}^T \mathbf{B} \mathbf{z} - 2\mathbf{z}^T \mathbf{c} + e \le \gamma + s \end{aligned} \tag{5.22}$$

To form the feasibility test which operates with the coefficients among which some are assigned zero values, vector **n** is extracted from **z**, as elaborated in the previous section. Accordingly, new matrix **V** is obtained by taking the rows and columns from **B** which correspond to nonzero coefficients in **z**. New vector **g** is obtained from **c** by taking the elements corresponding to nonzero coefficients in **z**. The feasibility test of sparse filter is then obtained as

$$\begin{aligned} &\underset{\mathbf{n},s}{\text{minimize}} \quad s \\ &\text{subject to} \quad \mathbf{n}^T \mathbf{V} \mathbf{n} - 2\mathbf{n}^T \mathbf{g} + e \le \gamma + s \end{aligned} \tag{5.23}$$

5.2.3 Feasibility Test for FIR Filters With Peak- and Quadratic- Error Constraints

If a filter is constrained by both – peak and quadratic constraints – the feasibility test is performed by solving the problem

$$\begin{aligned} \underset{\mathbf{z},s}{\text{minimize}} \quad & s \\ \text{subject to} \quad & \left|\mathbf{A}_q\mathbf{z} - d_q\right| \le \delta_q + s, \qquad q = 1,2,...,Q \\ & \mathbf{z}^T\mathbf{B}\mathbf{z} - 2\mathbf{z}^T\mathbf{c} + e \le \gamma + s \end{aligned} \tag{5.24}$$

The feasibility test which operates with the coefficients among which some are assigned zero values then takes the form

$$\begin{aligned} \underset{\mathbf{n},s}{\text{minimize}} \quad & s \\ \text{subject to} \quad & \left|\mathbf{U}_q\mathbf{n} - d_q\right| \le \delta_q + s, \qquad q = 1,2,...,Q \\ & \mathbf{n}^T\mathbf{V}\mathbf{n} - 2\mathbf{n}^T\mathbf{g} + e \le \gamma + s \end{aligned} \tag{5.25}$$

which is obtained by applying the constraints from (5.9) and (5.23).

5.3 Branch and Bound Method for Sparse Filter Design

5.3.1 Initial Sparsity

The design starts by finding any filter which satisfies specifications. Note that such a filter passes the feasibility test in (5.9), (5.23), or (5.25), depending on the approximation used. This filter can be designed by using a known sparse method. The number of its nonzero coefficients is here denoted by C. The value C is used as an initial sparsity. The optimization will improve this sparsity or prove its globality. If such a filter is not available, C can be estimated by using the Algorithm 5.1. In this algorithm, a nonsparse filter of minimum order is found, which satisfies the requirements.

```
 0: Specify dq, δq, γ, N, Q
 1: K = N/2+1
 2: Calculate ωq, q=1,...,Q
 3: For k = 1 to K
 4:    Calculate Aq by using (4.12) if minimax constraints are present
 5:    Calculate B, c, and e by using (5.19), (5.20), and (5.21) if
       quadratic constraints are present
 6:    Calculate s by using (5.7), (5.22), or (5.24)
 7:    If s <= 0
 8:       C=2k-1
 9:       Return
10:    End if
11: End for
12: Error: N is too small to satisfy design requirements
```

Algorithm 5.1 Algorithm for estimation of initial sparsity based on nonsparse filter design.

5.3.2 Choosing Search Strategy

After the initial sparsity has been estimated, a strategy for tree search must be chosen. In scope of this dissertation, two search strategies have been considered – the depth-first as well as the combination of the depth-first and breadth-first search.

Example of a tree with $K = 4$ elements is given in Figure 5.1. In each node of the tree, one of 2^K possible combinations of zero and nonzero coefficients is explored. A vector $\mathbf{P} = [P_1, P_2, ..., P_K]$, which describes such a combination, consists of elements

$$P_k = \begin{cases} 0, & k \in S^0 \\ 1, & k \in S^1 \end{cases} \tag{5.26}$$

where S^0 and S^1 are sets of indices corresponding to zero and nonzero coefficients in $\mathbf{z}$.

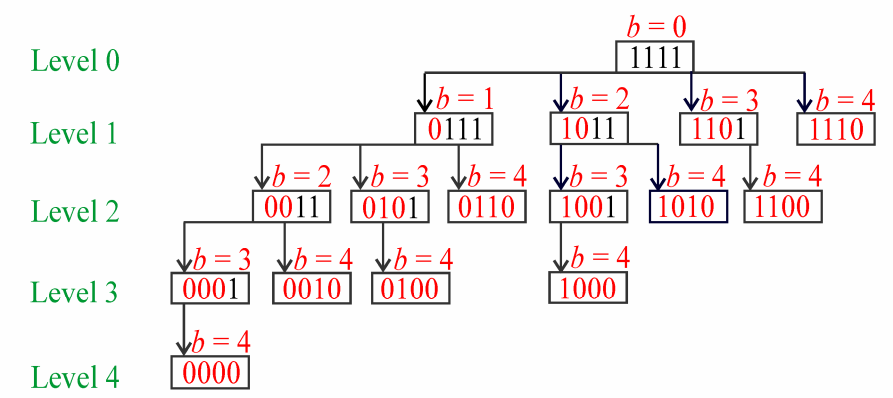

Figure 5.1 Example of tree with $K = 4$ elements. Position of last added zero is labeled as b.

5.3.2.1 Depth-First Search

At the beginning, all coefficients are assumed nonzero, that is, S^1 is a full set with K indices. Then, a zero is placed at the positions $b = 1, 2, 3,$ and 4. Note that below the node being processed, first b bits are fixed. Before the feasibility test in a node is performed, for each of the combinations obtained, simple justification is conducted. The number of nonzero coefficients in the fixed part of the node is calculated as

$$F = P_1 + 2\sum_{k=2}^{b} P_k \tag{5.27}$$

In an optimistic scenario, all of the remaining coefficients, z_k, $k = b+1, b+2, \ldots, K$, will take the value 0. Consequently, if $F > C$, the node cannot result in a higher sparsity than C and the corresponding branch should be pruned.

If $F \leq C$, the feasibility test is performed by using (5.9), (5.23), or (5.25). If the problem is feasible, the tree is further examined by going to the next level. If the problem is not feasible, the corresponding branch is pruned. The pseudo code of the described search strategy is given in Algorithm 5.2.

```
 1: (C,L) = search (r,C,P,L)
 2:    r = r + 1
 3:  For k = r to K
 4:     R = P, R_k = 0
 5:     Calculate F by using (5.27)
 6:     If F <= C
 7:        Calculate s by solving (5.9), (5.23), or (5.25)
 8:        If s <= 0
 9:           Calculate C_1 as the number of nonzero elements in R
10:           If C_1 = C
11:              Insert R in the list L
12:           Else if C_1 < C
13:              C = C_1
14:              Clear list L
15:              Insert R in the list L
16:           Else
17:               // do nothing
18:           End if
19:           (C,L) = search (k,C,R,L)
20:        Else
21:            // do nothing
22:        End if
23:     End if
24:  End for
```

Algorithm 5.2 Pseudo code for depth-first search strategy.

5.3.2.2 Combination of Depth-First and Breadth-First Search

In a depth-first search, examining each subtree improves current optimum. This optimum is then used in processing the next subtree that is rooted at the same level. However, if a zero coefficient cannot be placed at certain position regardless of the values of the remaining coefficients, this information can be utilized in node justification. Fortunately, this information can be acquired by examining the entire level before moving to the next level. This approach is named here the combination of the depth-first and the breadth-first search.

The information about positions that cannot contain zero coefficients are described by a mask vector $\mathbf{M} = [M_1, M_2, ..., M_K]$, where

$$M_k = \begin{cases} 0, & \text{zero-valued coefficient is not allowed} \\ 1, & \text{zero-valued coefficient is allowed} \end{cases} \tag{5.28}$$

The value of M_k is calculated by using the feasibility test, which is performed during breadth (horizontal) part of the search at each level. Note that the mask is not a global variable. Instead, it is updated at each level and then proceeded deeper. Clearly, the pruning at the deeper level can be performed immediately if the mask for the branch being processed is zero.

The mask can also be incorporated into justification. Assuming some of the coefficients z_k, $k = b + 1, b + 2, \ \ldots, K$, cannot take zero values, the expression (5.27) becomes

$$F = P_1 + 2\sum_{k=2}^{b} P_k + 2\sum_{k=b+1}^{K} \overline{M_k} \tag{5.29}$$

where $\overline{M_k}$ is a complement of M_k.

The pseudo code of the described search strategy is given in Algorithm 5.3.

In the presented search, improvement in pruning and consequently algorithm's convergence is achieved by using two enhancements. First is the mask, which indicates the positions at which zero coefficients cannot be placed. Second is using the prior knowledge of the number of nonzero coefficients, *C.* The features of these improvements will be evaluated in Section 5.5.

```
 1: (C,L) = search (M,r,C,P,L)
 2:    r = r + 1
 3:    t = [0,0,...,0]     // with K elements
 4:    For k = r to K
 5:       If M_k = 0
 6:          Continue
 7:       End if
 8:       R = P, R_k = 0
 9:       Calculate F by using (5.29)
10:       If F <= C
11:          Calculate s by solving (5.9), (5.23), or (5.25)
12:          If s <= 0
13:             t_k = 1
14:             Calculate C_1 as the number of nonzero elements in R
15:             If C_1 = C
16:                Insert R in the list L
17:             Else if C_1 < C
18:                C = C_1
19:                Clear list L
20:                Insert R in the list L
21:             Else
22:                // do nothing
23:             End if
24:          Else
25:             M_k = 0
26:          End if
27:       End if
28:    End for
29:    For k = r to K
30:       If t_k = 0
31:          Continue
32:       End if
33:       R = P
34:       R_k = 0
35:       (C,L) = search (M,k,C,R,L)
36:    End for
```

Algorithm 5.3 Pseudo code for proposed combination of depth-first and breadth-first search strategy.

5.4 Additional Refinement of Approximation Error

The branch and bound search provides list, **L**, which contains all combinations of zero coefficients corresponding to maximum sparsity. Although these combinations satisfy design parameters with equally sparse sets of coefficients, they provide different approximation errors. To find which of them ensures the smallest error, an additional refinement should be

performed. In a peak-, quadratic-error as well as a combined peak- and quadratic-error constrained design, the refinement is performed by using the following optimization problems

$$\begin{aligned} &\underset{\mathbf{n},\xi}{\text{minimize}} \quad \xi \\ &\text{subject to} \quad \frac{1}{\delta_q}\left|\mathbf{U}_q\mathbf{n} - d_q\right| \le \xi, \quad q = 1,2,...,Q \end{aligned} \tag{5.30}$$

$$\begin{aligned} &\underset{\mathbf{n},\zeta}{\text{minimize}} \quad \zeta \\ &\text{subject to} \quad \mathbf{n}^T\mathbf{V}\mathbf{n} - 2\mathbf{n}^T\mathbf{g} + e \le \zeta \end{aligned} \tag{5.31}$$

$$\begin{aligned} &\underset{\mathbf{n},\zeta}{\text{minimize}} \quad \zeta \\ &\text{subject to} \quad \left|\mathbf{U}_q\mathbf{n} - d_q\right| \le \delta_q, \qquad q = 1,2,...,Q \\ &\qquad\qquad \mathbf{n}^T\mathbf{V}\mathbf{n} - 2\mathbf{n}^T\mathbf{g} + e \le \zeta \end{aligned} \tag{5.32}$$

The optimum values of ξ and ζ are the approximation errors of the refined filters. The best filter from **L** is selected as the one with the smallest approximation error.

5.5 Properties of Proposed Methods

5.5.1 Subproblem Analysis and Execution Time Comparison

In this section, the design methods based on the depth-first as well as the combination of depth-first and breadth-first search strategy are analyzed. Their performances are compared on the design of filters constrained in peak-error sense. In particular, filters with orders $N = 60$, 70, and 80, passband and stopband edge frequencies $\omega_p = 0.3\pi$ and $\omega_s = 0.5\pi$, passband approximation error $\delta_p = 0.001$ dB, and stopband approximation errors $\delta_s = -60$ dB, -65 dB, -70 dB, -75 dB, -80 dB, are designed by using both search strategies. The initial sparsity was obtained by using Algorithm 5.1.

Figures 5.2 and 5.3 show the number of subproblems and the execution time required for the design of all filters. It is clear from figures that the combination of depth-first and breadth-first strategy exhibits better properties in both parameters than does the depth-first strategy. Consequently, the combination approach is used in further examples.

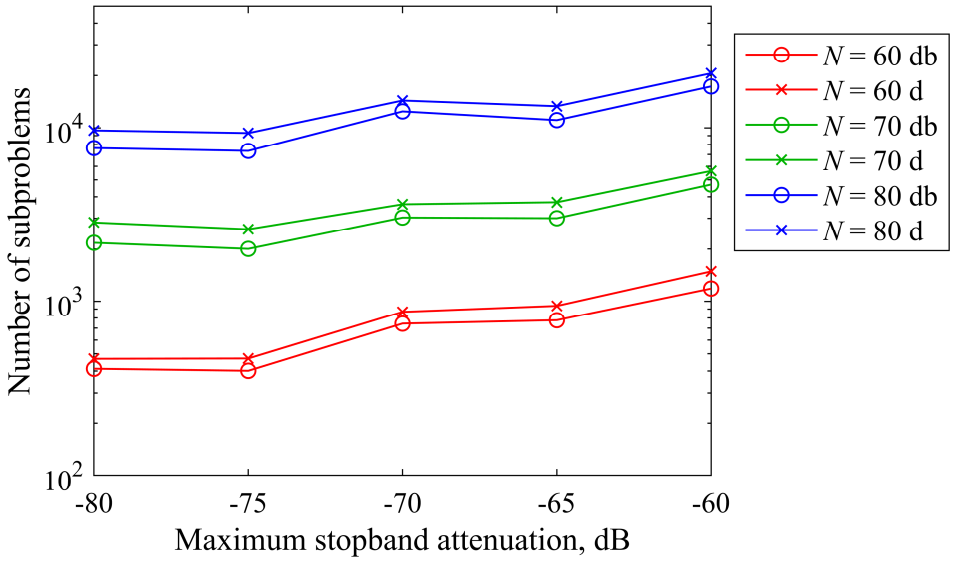

Figure 5.2 Number of subproblems solved in design of various filters obtained by depth-first (d) and combination of depth- and breadth-first (db) search strategy.

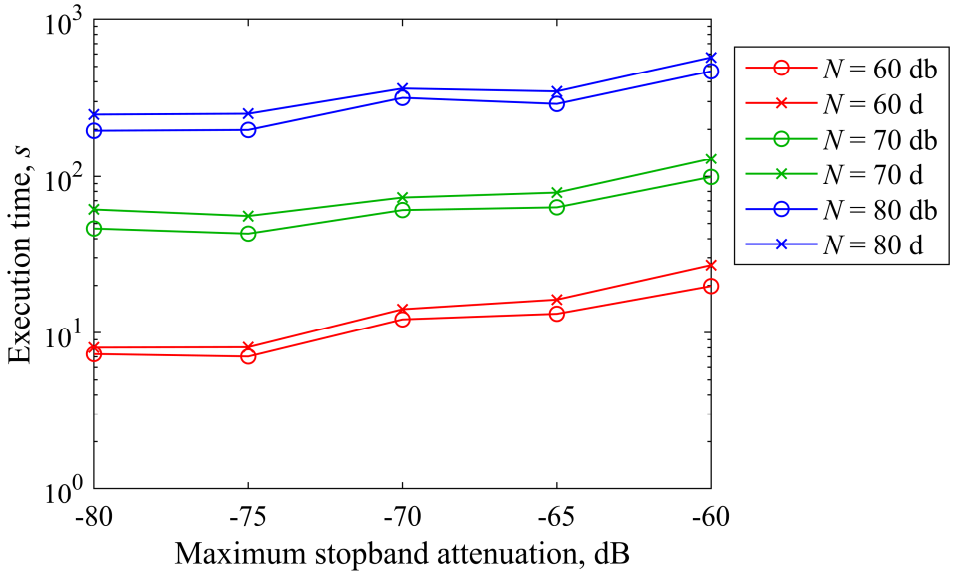

Figure 5.3 Execution time for design of various filters obtained by depth-first (d) and combination of depth- and breadth-first (db) search strategy.

5.5.2 Comparison With Exhaustive Search

Here, the comparison of the proposed method and the exhaustive search is presented in the scope of number of subproblems and execution time. The comparison is made for filters with orders $N = 60$, 70, and 80, passband and stopband edge frequencies $\omega_p = 0.3\pi$ and $\omega_s = 0.5\pi$, passband approximation error $\delta_p = 0.001$ dB, and stopband approximation errors $\delta_s = -60$ dB, -70 dB, and -80 dB. The initial sparsity used in the branch and bound method was obtained by applying Algorithm 5.1.

For the proposed method, the number of subproblems, N_{db}, and execution time, t_{db}, are obtained for the design performed on a personal computer with quad-core Intel i7 processor operating at the clock of 3.6 GHz. For the exhaustive search, the number of subproblems, N_{es}, and execution time, t_{es}, can only be estimated. Namely, the exhaustive search requires solving 2^N problems in (5.9), (5.23) or (5.25), which lasts too long for measurement. Therefore, the estimation is made assuming that solving one optimization problem requires 1 ms. The results are shown in Table 5.1. It is evident that branch and bound enables solving of problems for which solving with exhaustive search would not finish in a lifetime.

Table 5.1 Number of subproblems and execution time for design of various filters obtained by exhaustive search, N_{es} and t_{es}, and combination of depth- and breadth-first, N_{db} and t_{db}, search strategy.

N	δ_s, dB	N_{es}	t_{es}	N_{db}	t_{db}
60	–60	1.2e18	3.6e7 year	1182	21.1 s
	–70	1.2e18	3.6e7 year	743	13.3 s
	–80	1.2e18	3.6e7 year	412	7.5 s
70	–60	1.2e21	3.7e10 year	4706	1.7 min
	–70	1.2e21	3.7e10 year	3036	1.1 min
	–80	1.2e21	3.7e10 year	2170	49.2 s
80	–60	1.2e24	3.8e13 year	17295	7.7 min
	–70	1.2e24	3.8e13 year	12421	5.4 min
	–80	1.2e24	3.8e13 year	7633	3.6 min

5.6 Design Examples

All examples show the design of lowpass filters with desired amplitude response

$$d = \begin{cases} 1, & 0 \le \omega \le \omega_p \\ 0, & \omega_s \le \omega \le \pi \end{cases} \tag{5.33}$$

In peak-error constraints, total number of points $Q = 10N$ is used unless otherwise noted. The desired response, when defined in finite number of frequency points, is obtained as

$$d_q = \begin{cases} 1, & q = 1, 2, ..., Q_p \\ 0, & q = Q_p + 1, Q_p + 2, ..., Q \end{cases} \tag{5.34}$$

where Q_p is a number of frequency points in the passband.

In quadratic constraints, $w = 1$ is used in all care bands.

5.6.1 Lowpass Filters With Peak-Error Constraints

5.6.1.1 Example 1

In this section, filters are designed with the specifications given in Example 1 in Section 4 and in [39]. These specifications are here repeated for convenience. They are $N = 60$, 70 and 80, $\omega_p = 0.3\pi$, $\omega_s = 0.5\pi$, $\delta_p = 0.001$ dB, and $\delta_s = -60$ dB, −65 dB, −70 dB, −75 dB, −80 dB. For the initial number of nonzero coefficients, L_{NZ} from [39] is used.

Table 5.2 shows the comparison of the obtained filters with their counterparts designed by SGP method in Section 4 and with the filters in [39]. The branch and bound method is global, as opposed to other two methods. Therefore, it shows how close these nonglobal methods are to the global solutions. It is interesting to note that the branch and bound method in all filters provides the same number of nonzero coefficients as does the SGP method. Furthermore, the proposed method gives two more zeroes then does the method in [39] in six filters. The paper referred to does not provide the information on the impulse response length. Therefore the comparison of W_{IR} is given only for SGP method. To compare W_{IR} with the one obtained by the SGP method, the shortest impulse response among all obtained solutions is listed. It is interesting to observe that by increasing filter order resulting distributions of nonzero coefficients remain the same.

Table 5.2 Number of obtained sparse filters, N_S, length of impulse response, W_{IR} and number of nonzero coefficients, L_{NZ} of proposed filters compared with corresponding filters obtained by SGP method in Section 4 and with filters in [39]. The improvements are marked in bold.

N	δ_s, dB	N_S	W_{IR}	L_{NZ}	W_{IR} SGP	L_{NZ} SGP	L_{NZ} [39]
60	–60	5	**47**	37	53	37	37
70	–60	5	**47**	37	53	37	37
80	–60	5	**47**	37	53	37	37
60	–65	1	53	**37**	53	37	39
70	–65	1	53	**37**	53	37	39
80	–65	1	53	**37**	53	37	39
60	–70	4	**47**	39	53	39	39
70	–70	4	**47**	39	53	39	39
80	–70	4	**47**	39	53	39	39
60	–75	1	53	**39**	53	39	41
70	–75	1	53	**39**	53	39	41
80	–75	1	53	**39**	53	39	41
60	–80	1	49	41	49	41	41
70	–80	1	49	41	49	41	41
80	–80	1	49	41	49	41	41

To illustrate the properties of the filters from Table 5.2, the responses of a filter obtained for $N = 60$ and $\delta_s = -70$ dB are shown. Note that, after branch and bound optimization has been performed, all filters with the same sparsity are further refined as described in Section 5.4. Among them, the filter with the smallest approximation error is selected. Figure 5.4 and Table 5.3 show its impulse response samples in graphical and numerical form. The corresponding magnitude response is shown in Figure 5.5, with enlarged passband in Figure 5.6. Clearly, the maximum approximation errors in passband and stopband are smaller than the required δ_p and δ_s. This effect can be observed in Figures 5.5 and 5.6 where the required attenuations are shown in green.

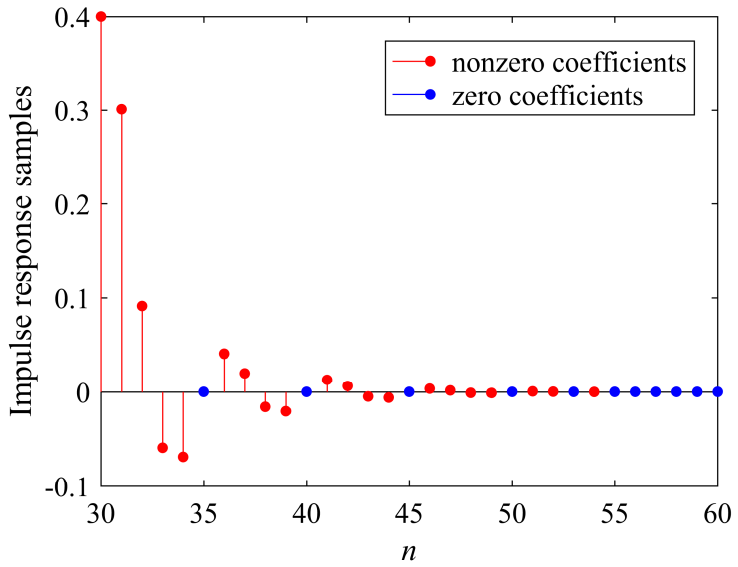

Figure 5.4 Impulse response samples of sparse lowpass filter from Table 5.2 with $N = 60$ and $\delta_s = -70$ dB. Only right-hand side is shown.

Table 5.3 Coefficients of right-hand side of sparse lowpass filter from Table 5.2 with $N = 60$ and $\delta_s = -70$ dB

n	$h(n) = h(N-n)$	n	$h(n) = h(N-n)$	n	$h(n) = h(N-n)$
30	0.39999	41	0.01294	52	0.00016
31	0.30094	42	0.00629	53	0
32	0.09136	43	–0.00489	54	–0.00010
33	–0.05912	44	–0.00607	55	0
34	–0.06880	45	0	56	0
35	0	46	0.00340	57	0
36	0.04064	47	0.00153	58	0
37	0.01988	48	–0.00106	59	0
38	–0.01585	49	–0.00120	60	0
39	–0.02048	50	0		
40	0	51	0.00049		

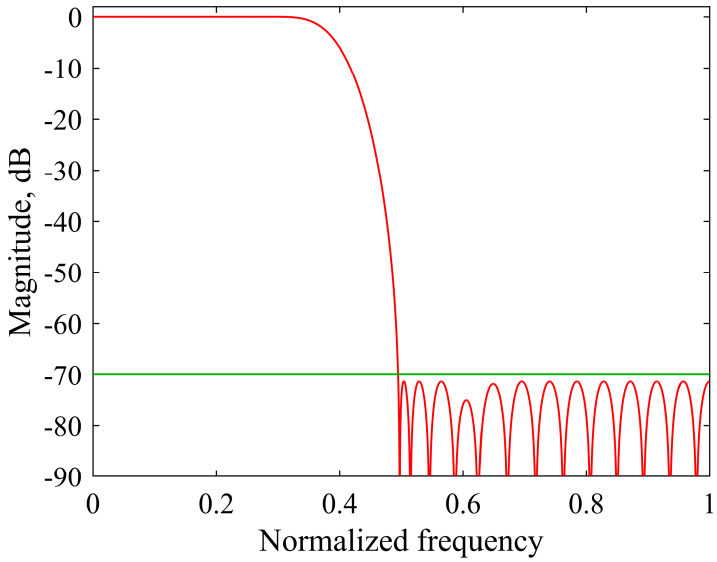

Figure 5.5 Magnitude response of sparse lowpass filter from Table 5.2 with order $N = 60$ and required maximum stopband approximation error $\delta_s = -70$ dB. Required δ_s is shown in green.

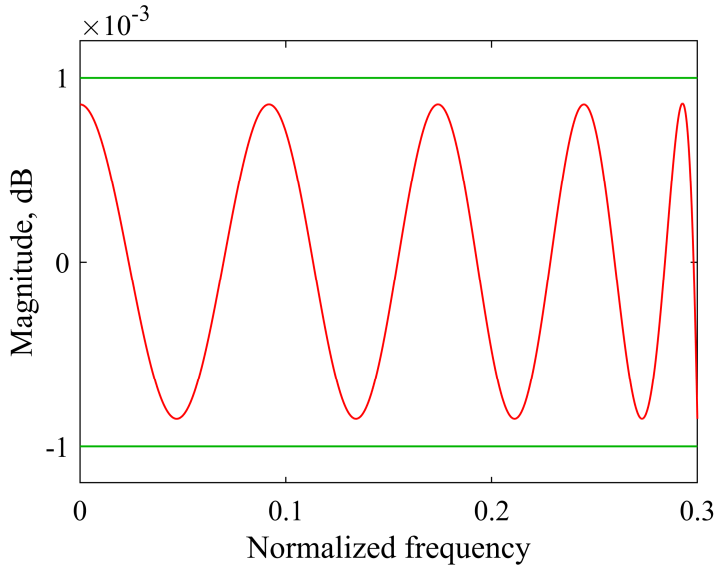

Figure 5.6 Enlarged passband magnitude response of sparse lowpass filter from Table 5.2 with order $N = 60$ and required maximum passband approximation error $\delta_p = 0.001$ dB. Required δ_p is shown in green.

5.6.1.2 Example 2

The filters presented in this example are designed with the specifications given in Example 2 in Section 4 and in [40]. These specifications are $N = 60$, 70 and 80, $\omega_p = 0.3\pi$, $\omega_s = 0.5\pi$, $\delta_p = 0.5$ dB, and $\delta_s = -60$ dB, −65 dB, −70 dB, −75 dB, −80 dB. For the initial number of nonzero coefficients, L_{NZ} from [40] is used. Note that required passband approximation error of 0.5 dB is a rather loose requirement. Consequently, higher sparsities than those obtained in the previous example are expected.

Table 5.4 shows the comparison of the obtained filters with the corresponding filters designed by SGP method in Section 4 and with the filters in [40]. The results show that the proposed branch and bound method provides much higher sparsities and shorter impulse responses then do the SGP method and the method in [40].

Table 5.4 Number of obtained sparse filters, N_S, length of impulse response, W_{IR}, and number of nonzero coefficients, L_{NZ}, of proposed filters compared with corresponding filters obtained by SGP method in Section 4 and with filters in [40]. The improvements are marked in bold.

N	δ_s, dB	N_S	W_{IR}	L_{NZ}	W_{IR} SGP	L_{NZ} SGP	L_{NZ} [40]
60	–60	1	**23**	**21**	33	25	29
70	–60	1	**23**	**21**	33	25	29
80	–60	1	**23**	**21**	33	25	25
60	–65	4	**25**	**23**	31	25	29
70	–65	4	**25**	**23**	31	25	29
80	–65	4	**25**	**23**	41	29	29
60	–70	14	**25**	**25**	31	25	33
70	–70	14	**25**	**25**	31	25	33
80	–70	14	**25**	**25**	31	25	31
60	–75	1	**29**	**25**	35	27	33
70	–75	1	**29**	**25**	35	27	33
80	–75	1	**29**	**25**	35	27	33
60	–80	6	**29**	**27**	37	29	37
70	–80	6	**29**	**27**	39	29	37
80	–80	6	**29**	**27**	39	29	35

5.6.1.3 Example 3

This example presents a set of lowpass symmetric FIR filters, which have low attenuations in the stopband and steep transition bands. Such filters are considered in many papers, including [38] and [44]. The filters are obtained with $N = 80, 90$ and 100, $\omega_p = 0.0436\pi$, $\omega_s = 0.0872\pi$, $\delta_p = 0.5$ dB, and $\delta_s = -20$ dB, –25 dB, –30 dB, –35 dB, and –40 dB. For the initial number of nonzero coefficients, L_{NZ} from [38] is used.

Table 5.5 shows the comparison of the obtained filters and their counterparts in [38] and [44]. The proposed method results in the same number of nonzero coefficients as does the method in [38]. Furthermore, two more zeros are obtained in three filters when compared with [44]. Authors in [44] optimized sparsity and filter's length. Therefore, the impulse response length, W_{IR}, of the filters obtained is also included in the comparison. The proposed method achieves shorter impulse responses in six filters. It is interesting to observe the number of sparse filters N_S. As expected, smaller attenuations in the stopband give more degrees of freedom. Furthermore, many filters with equal minimum number of nonzero coefficients are found.

Table 5.5 Number of obtained sparse filters, N_S, number of nonzero coefficients, L_{NZ}, and length of impulse response, W_{IR}, of proposed filters compared with corresponding filters presented in [38] and [44].

N	δ_s, dB	N_S	L_{NZ} / W_{IR}	L_{NZ} [38]	L_{NZ} / W_{IR} [44]
80	–20	32	29/45	29	29/45
90	–20	31	29/45	29	29/45
100	–20	31	29/45	29	29/45
80	–25	14	37/49	37	37/49
90	–25	14	37/49	37	37/49
100	–25	14	37/49	37	37/49
80	–30	25	47/55	47	47/55
90	–30	25	47/55	47	47/55
100	–30	25	47/55	47	47/55
80	–35	5	55/**75**	55	55/77
90	–35	6	55/**75**	55	55/77
100	–35	6	55/**75**	55	55/77
80	–40	8	**65/79**	65	67/81
90	–40	20	**65/79**	65	67/81
100	–40	20	**65/79**	65	67/81

5.6.2 Lowpass Filters With Quadratic Constraints

In this section, a set of lowpass FIR filters with quadratic energy constraint in care bands is designed. Filters with the same specifications were used in [45]. These specifications are $N = 88$, $\omega_p = 0.0436\pi$, $\omega_s = 0.0872\pi$, and $Q = 5N$. The initial number of nonzero coefficients is set to L_{NZ} from [45]. Note that the paper referred to optimizes weighted least-squares energy by solving the problem

$$\underset{\mathbf{z}}{\text{minimize}} \quad \int_0^{\pi} W(\omega)\left[H(\mathbf{z},\omega) - H_d(\omega)\right]^2 d\omega \tag{5.35}$$

where $W(\omega)$ is equal to 1 in care bands, and 0 otherwise. In this dissertation, energy is constrained to some specified value, as shown in (5.10). The constraint in (5.10) and the objective function in (5.35) differ in factor $1/\pi$. To keep the physical meaning of energy, weighted least square value obtained in [45] is here multiplied by factor $1/\pi$. This value is used as γ in the proposed design.

Table 5.6 shows the comparison of the proposed filters with the corresponding filters in [45]. In paper referred to, the authors list the obtained peak approximation error in the passband and stopband although they do not optimize it. Therefore, peak-errors δ_p and δ_s obtained with the proposed method are also provided, for comparison. As clear from the table, the quadratic approximation error and the number of nonzero coefficients obtained by the proposed method are better in all seven filters. The obtained peak errors are somewhat larger than those in [45]. However, lowering the peak-error is further analyzed in Section 5.6.3, which considers simultaneous application of peak and quadratic constraints.

The impulse response samples of filter number 6 from Table 5.6 are shown in Figure 5.7. Their numerical values are given in Table 5.7, for convenience. Magnitude response of this filter is presented in Figure 5.8.

Table 5.6 Number of obtained sparse filters, N_S, as well as number of nonzero coefficients, L_{NZ}, optimum quadratic approximation error, ζ_{opt}, and passband and stopband peak errors, δ_p and δ_s, of proposed filters compared to corresponding filters in [45]. The improvements are marked in bold.

Filter no.	N_S	L_{NZ}	L_{NZ} [45]	ζ_{opt}	γ [45]	proposed δ_p / δ_s	$\delta_p = \delta_s$ [45]
1	2	**85**	89	**5.2972e–6**	5.7550e–6	2.9595e–2/2.5846e–2	1.9680e–2
2	9	**77**	79	**1.0233e–5**	1.2939e–5	2.6978e–2/4.0383e–2	2.4121e–2
3	5	**75**	77	**1.3251e–5**	1.5053e–5	3.1818e–2/4.4267e–2	3.0025e–2
4	1	**73**	75	**1.6069e–5**	1.7271e–5	3.0851e–2/4.8029e–2	3.2719e–2
5	1	**71**	73	**1.9938e–5**	2.0923e–5	3.6358e–2/5.2708e–2	3.6368e–2
6	1	**67**	71	**3.0152e–5**	3.0641e–5	4.7126e–2/5.9601e–2	3.1554e–2
7	6	**67**	69	**3.0152e–5**	3.3238e–5	4.7126e–2/5.9601e–2	4.3849e–2

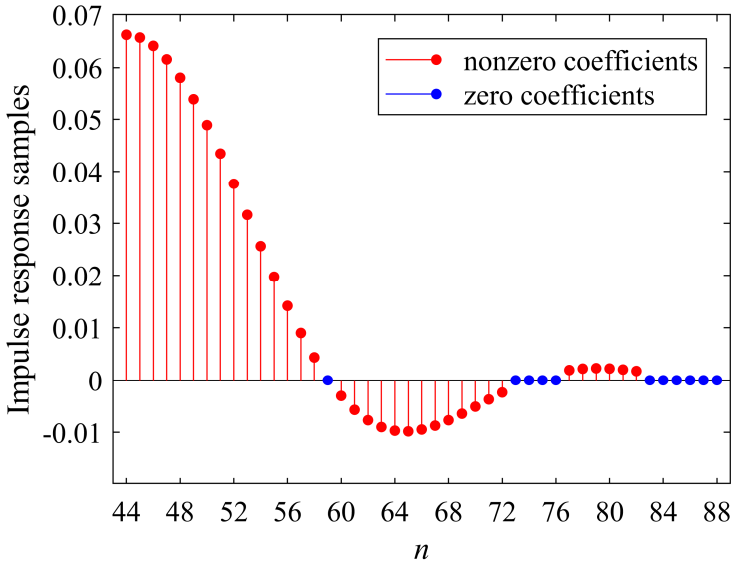

Figure 5.7 Impulse response samples of filter number 6 from Table 5.6. Only right-hand side is shown.

Table 5.7 Coefficients of the right-hand side of filter number 6 from Table 5.6.

n	$h(n) = h(N–n)$	n	$h(n) = h(N–n)$	n	$h(n) = h(N–n)$	n	$h(n) = h(N–n)$
44	0.06620	56	0.01417	68	–0.00768	80	0.00225
45	0.06567	57	0.00901	69	–0.00644	81	0.00206
46	0.06410	58	0.00437	70	–0.00508	82	0.00178
47	0.06155	59	0	71	–0.00370	83	0
48	0.05808	60	–0.00302	72	–0.00237	84	0
49	0.05381	61	–0.00570	73	0	85	0
50	0.04887	62	–0.00768	74	0	86	0
51	0.04342	63	–0.00899	75	0	87	0
52	0.03762	64	–0.00968	76	0	88	0
53	0.03163	65	–0.00981	77	0.00196		
54	0.02562	66	–0.00946	78	0.00223		
55	0.01975	67	–0.00872	79	0.00232		

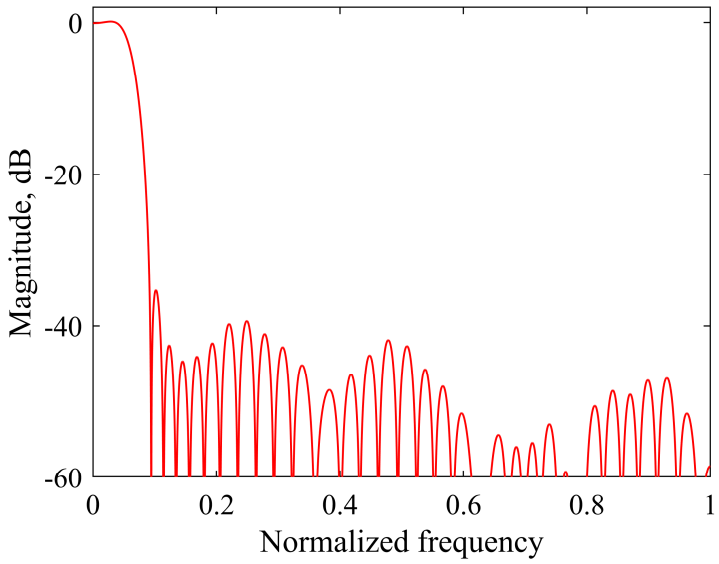

Figure 5.8 Magnitude response of filter number 6 from Table 5.6.

5.6.3 Lowpass Filters With Peak-Error and Quadratic Constraints

This example presents lowpass FIR filters obtained with the specifications from [45], which have already been used in the previous section. However, the peak-error constraints are here incorporated together with the quadratic-error constraints. Since the paper [45] shows the peak-errors of the filters obtained, these values are used here as the required δ_p and δ_s.

Table 5.8 shows the comparison of the proposed filters with the corresponding filters in [45]. Clearly, better filters in the terms of peak-error then those in the previous section are obtained. However, due to reduced design freedom, slightly lower sparsities are encountered. They are equal to those in [45] in four and improved in three filters. The quadratic error is additionally improved by the refinement shown in Section 5.4 and it is lower in all cases.

As in the previous example, the impulse response samples of filter number 6 from Table 5.8 are shown in Figure 5.9 and Table 5.9. The magnitude response of this filter in comparison with the filter obtained in the previous section is shown in Figure 5.10. In addition, in Figure 5.11 the transition band of magnitude response is enlarged to illustrate the benefits of adding the peak-error constraints.

Table 5.8 Number of obtained sparse filters, N_S, as well as number of nonzero coefficients, L_{NZ}, optimum quadratic approximation error, ζ_{opt}, and passband and stopband peak-errors δ_p and δ_s of proposed filters compared to corresponding filters in [45]. The improvements are marked in bold.

Filter no.	N_S	L_{NZ}	L_{NZ} [45]	ζ_{opt}	γ [45]	proposed δ_p / δ_s	$\delta_p = \delta_s$ [45]
1	1	**87**	89	**4.9959e–6**	5.7550e–6	1.9666e–2/1.9596e–2	1.9680e–2
2	1	**77**	79	**1.2526e–5**	1.2939e–5	2.2963e–2/2.4121e–2	2.4121e–2
3	14	77	77	**1.1162e–5**	1.5053e–5	2.4410e–2/3.0024e–2	3.0025e–2
4	6	75	75	**1.4456e–5**	1.7271e–5	2.9416e–2/3.2714e–2	3.2719e–2
5	6	73	73	**1.7360e–5**	2.0923e–5	2.8255e–2/3.6368e–2	3.6368e–2
6	13	71	71	**2.4446e–5**	3.0641e–5	3.1216e–2/3.1551e–2	3.1554e–2
7	1	**67**	69	**3.2868e–5**	3.3238e–5	4.3527e–2/4.3830e–2	4.3849e–2

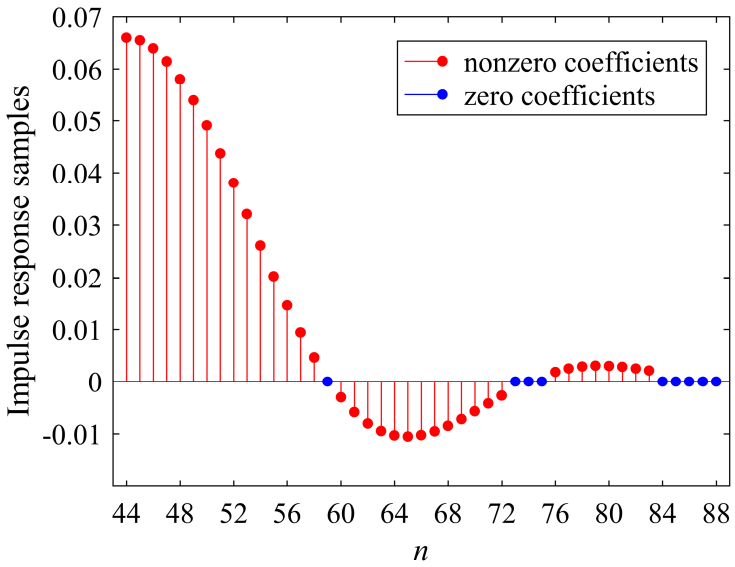

Figure 5.9 Impulse response samples of filter number 6 from Table 5.8. Only right-hand side is shown.

Table 5.9 Coefficients of right-hand side of filter number 6 from Table 5.8.

n	$h(n) = h(N-n)$	n	$h(n) = h(N-n)$	n	$h(n) = h(N-n)$	n	$h(n) = h(N-n)$
44	0.06598	56	0.01461	68	–0.00849	80	0.00306
45	0.06547	57	0.00938	69	–0.00718	81	0.00287
46	0.06394	58	0.00464	70	–0.00573	82	0.00255
47	0.06145	59	0	71	–0.00421	83	0.00214
48	0.05807	60	–0.00303	72	–0.00271	84	0
49	0.05391	61	–0.00587	73	0	85	0
50	0.04908	62	–0.00801	74	0	86	0
51	0.04372	63	–0.00948	75	0	87	0
52	0.03800	64	–0.01030	76	0.00189	88	0
53	0.03207	65	–0.01054	77	0.00252		
54	0.02610	66	–0.01026	78	0.00291		
55	0.02023	67	–0.00954	79	0.00308		

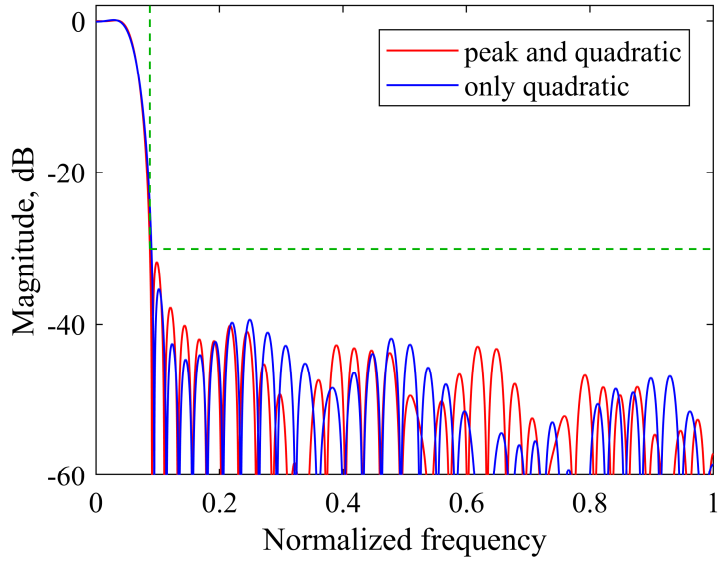

Figure 5.10 Magnitude responses of filter number 6 from Tables 5.6 (blue) and 5.8 (red). Vertical green line is drawn at stopband edge frequency, whereas horizontal green line represents the required maximum stopband peak-error.

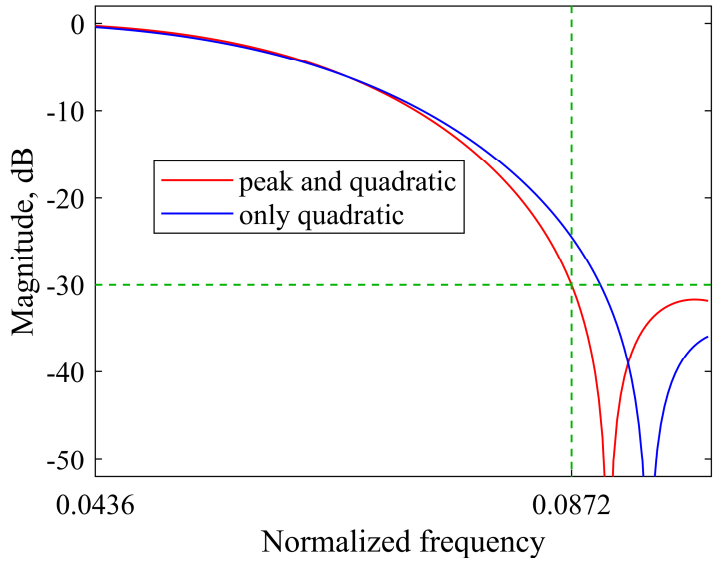

Figure 5.11 Enlarged transition bands of filter number 6 from Tables 5.6 (blue) and 5.8 (red). Vertical green line is drawn at stopband edge frequency, whereas horizontal green line represents the required maximum stopband peak-error.

6 Global Optimization of Spatial Filters With Constrained Coefficients' Dynamic Range Ratio

6.1 Optimization-Based Design of Pencil Beams with Minimax Sidelobes

The design of a pencil beam whose sidelobes are optimal in minimax sense is given in (2.34). If the second constraint the problem referred to is evaluated on a finite set of equidistant points θ_q, $q = 1, 2, ..., Q$, $\theta_q \in [\theta_s, \pi/2]$, the optimization takes the form [76]

$$\begin{aligned} &\underset{\mathbf{a},\delta}{\text{minimize}} \quad \delta \\ &\text{subject to} \quad \sum_{k=1}^{N} a_k = 1 \\ &\qquad\qquad \left\| \mathbf{A}_q \mathbf{a} \right\| \le \delta, \quad q = 1, 2, ..., Q \end{aligned} \tag{6.1}$$

where

$$\mathbf{A}_q = \begin{bmatrix} \cos\left(u_q x_1\right) & \cos\left(u_q x_2\right) & \dots & \cos\left(u_q x_N\right) \\ \sin\left(u_q x_1\right) & \sin\left(u_q x_2\right) & \dots & \sin\left(u_q x_N\right) \end{bmatrix} \tag{6.2}$$

$$u_q = \frac{2\pi}{\lambda} \sin\left(\theta_q\right), \quad q = 1, 2, \dots, Q \tag{6.3}$$

The problem in (6.1) is a convex second-order cone program with one equality and Q second-order-cone constraints, which can be easily solved. The solution results in a beampattern which can also be obtained analytically, by the Dolph-Chebyshev method. However, the optimization-based approach enables incorporating additional requirements, such as the requirement for a low dynamic range ratio of excitation coefficients.

6.2 DRR Constrained Design

Incorporating constraints for low DRR into a pencil beam design is illustrated in (2.48). If these constraints are applied in (6.1), the optimization problem becomes

$$\begin{aligned}
\underset{\mathbf{a},\delta,t}{\text{minimize}} \quad & \delta \\
\text{subject to} \quad & \sum_{k=1}^{N} a_k = 1 \\
& \left\| \mathbf{A}_q \mathbf{a} \right\| \le \delta, \quad q = 1,2,...,Q \\
& \left| a_k \right| \le Dt, \quad k = 1,2,...,N \\
& \left| a_k \right| \ge t, \quad k = 1,2,...,N \\
& t \ge 0
\end{aligned} \tag{6.4}$$

In this problem, constraints $|a_k| \ge t$, $k = 1, 2, \ldots, N$ are not convex. However, if all coefficient signs are known in advance, such constraints could be rewritten in a convex form. Assuming $S^+ \subseteq \{1, 2, ..., N\}$ and $S^- \subseteq \{1, 2, ..., N\}$ are the set of indices corresponding to positive and negative coefficients, the problem in (6.4) can be expressed as

$$\begin{aligned}
\underset{\mathbf{a},\delta,t}{\text{minimize}} \quad & \delta \\
\text{subject to} \quad & \sum_{k=1}^{N} a_k = 1 \\
& \left\| \mathbf{A}_q \mathbf{a} \right\| \le \delta, \quad q = 1,2,...,Q \\
& a_k \le Dt, \quad k \in S^+ \\
& a_k \ge t, \quad k \in S^+ \\
& -a_k \le Dt, \quad k \in S^- \\
& -a_k \ge t, \quad k \in S^- \\
& t \ge 0
\end{aligned} \tag{6.5}$$

The problem in (6.5) is convex. It contains one equality, Q second-order cone and $2N + 1$ inequality constraints.

To solve the problem in (6.4) globally, (6.5) should be applied for all combinations of coefficient signs. This leads to an exhaustive search with 2^N optimization runs. The exhaustive search can be performed in an acceptable time only for arrays with small number of elements. For more complex arrays, the search should be performed in an efficient manner. Here, one such approach based on branch and bound algorithm will be presented [60].

To perform a search of all combinations of coefficient signs, a tree is formed. Figure 6.1 shows such a tree built for an antenna array with four elements. At the root, denoted by Level 0, all coefficients are assumed positive. At Level 1, one negative coefficient is placed at each possible position, 1, 2, 3 and 4. At a new level, additional negative coefficient is added at the positions b+1, b+2, ..., 4, where b marks the position at which last negative coefficient has

been assigned. Such a tree has b uniquely specified left-hand signs in each node. In Figure 6.1, the values of b, as well as signs specified in each node are marked in red.

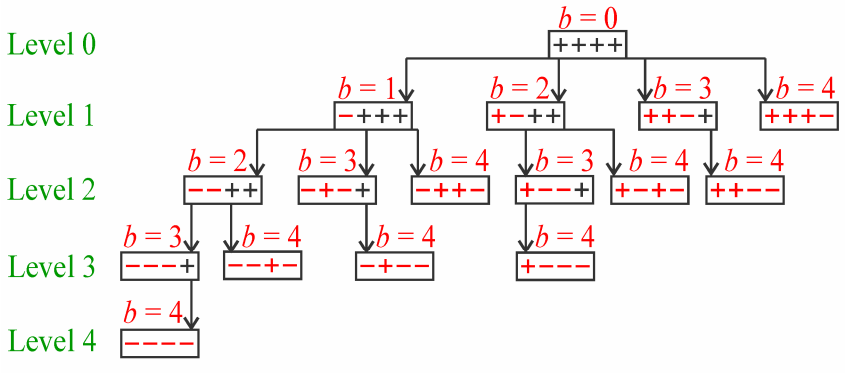

Figure 6.1 Example of tree with all combinations of coefficient' signs in antenna array with four elements. Signs which are uniquely specified in particular node are marked in red.

The tree is explored by using a depth-first search. Let variables δ_{opt} and $\mathbf{a}_{opt}$ denote the maximum sidelobe level and optimum antenna coefficients. At the beginning, the value of δ_{opt} is set to ∞. At each node, two tests are made. In the first test, the problem (6.5) is solved with the coefficient signs shown in Figure 6.1. If the problem is feasible, new δ and $\mathbf{a}$ are obtained as a solution. If $\delta < \delta_{opt}$, an update is made $\delta_{opt} = \delta$ and $\mathbf{a}_{opt} = \mathbf{a}$. In the second test, all branches leaving the node are simultaneously tested by solving the following problem

$$
\begin{aligned}
\underset{\mathbf{a},\delta,t}{\text{minimize}} \quad & \delta \\
\text{subject to} \quad & \sum_{k=1}^{N} a_k = 1 \\
& \left\| \mathbf{A}_q \mathbf{a} \right\| \le \delta, \quad q = 1,2,\ldots,Q \\
& a_k \le Dt, \quad k \in S^+ \;\&\; k \le b \\
& a_k \ge t, \quad k \in S^+ \;\&\; k \le b \\
& -a_k \le Dt, \quad k \in S^- \;\&\; k \le b \\
& -a_k \ge t, \quad k \in S^- \;\&\; k \le b \\
& a_k \le Dt \quad k = b+1, b+2, \ldots, N \\
& -a_k \le Dt \quad k = b+1, b+2, \ldots, N \\
& t \ge 0
\end{aligned}
\tag{6.6}
$$

The problem (6.6) is a relaxed version of problem in (6.5). At each node, the signs of first b coefficients are known. The remaining $N–b$ coefficients, which are in Figure 6.1 marked in black, can be assigned any possible sign, resulting in 2^{N-b} combinations. In (6.6) upper- and lower-bound constraints are tested only for known (red) signs. On the other hand, for the other (black) signs, nonconvex constraints $|a_k| \geq t$ are omitted and only convex $|a_k| \leq Dt$, $k = b+1, b+2, ..., N$ are solved. This relaxation makes the problem in (6.6) a convex one and therefore solvable globally. It is evident that solution of (6.6) might not satisfy the DRR constraints. However, if $\delta > \delta_{opt}$ is obtained, going deeper in the tree cannot give a solution better than the existing, therefore corresponding branch can be pruned. It should be noted that in such case δ_{opt} is not updated because solution of relaxed problem is not a solution to an original problem.

For two antenna arrays which have same values of coefficients, but whose coefficients are reversed in time, the resulting sidelobe level will be the same. This implies that both solutions can be obtained as a result. Furthermore, it implies that one node testing can be omitted. If sign combinations are expressed in a binary form, then the coefficients with negative sign are given the value of 0 and coefficients with positive sign are given the value of 1. B and B_r are the binary values of the original and reversed combination of coefficient signs, respectively. If $B > B_r$, the node testing is redundant.

6.3 Design Examples

In the following examples, pencil beams with isotropic elements and spacing of $\lambda / 2$ are considered. The number of frequency points is set to $Q = 10N$ and the beamwidth of 12° is used unless otherwise noted.

6.3.1 Example 1

In the first example, the proposed method is compared with the exhaustive search. The number of elements, N, and the required DRR, D, are given in Table 6.1 together with the number of subproblems, N_{es}, and design time, t_{es}, which are estimated for the exhaustive search, as well as the number of subproblems, N_{bnb}, and design time, t_{bnb}, obtained by the proposed method. Exhaustive search requires solving of problem (6.5) 2^N times. On a personal computer with quad-core Intel i7 processor operating at the clock of 3.6 GHz, solving each optimization problem takes 1 ms in average. This value is used in calculation of t_{es}.

It is evident that the proposed method significantly decreases number of subproblems that require solving. Consequently, it reduces the design time. Clearly, the obtained t_{bnb} indicates the possibility of interactive design when experimenting with various parameters is needed.

Table 6.1 Comparison of number of subproblems and design time for exhaustive search, N_{es} and t_{es}, and proposed method, N_{bnb} and t_{bnb}, for arrays with various numbers of elements, N, and required DRR, D.

N	D	N_{es}	t_{es}	N_{bnb}	t_{bnb}
10	2	1024	1.024 s	42	0.21 s
	1.5	1024	1.024 s	48	0.21 s
	1	1024	1.024 s	39	0.16 s
20	2	1.049e6	17.4 min	20	0.20 s
	1.5	1.049e6	17.4 min	20	0.21 s
	1	1.049e6	17.4 min	112	1 s
30	2	1.074e9	298 hours	142	2.63 s
	1.5	1.074e9	298 hours	291	5.31 s
	1	1.074e9	298 hours	1595	24.53 s
40	2	1.100e12	34.9 years	80	2.50 s
	1.5	1.100e12	34.9 years	155	4.89 s
	1	1.100e12	34.9 years	2104	49.52 s
50	2	1.126e15	35702 years	244	11.65 s
	1.5	1.126e15	35702 years	777	37.19 s
	1	1.126e15	35702 years	11451	393.8 s

6.3.2 Example 2

This example provides the comparison of optimum sidelobe level (SLL) in DRR-unconstrained and DRR-constrained design for various numbers of antenna elements. In DRR-unconstrained design, optimum coefficients and SLL are obtained by solving (6.1), resulting in Dolph-Chebyshev beam patterns with DRR values of D_{DC}. The comparison is given in Figure 6.2 for pencil beams with $N = 5, 10, 15, 20, 30$ and 40 elements. The last point in each graph corresponds to DRR-unconstrained design, with DRR equal to D_{DC}. By adding constraints which require $D < D_{DC}$, SLL deteriorates. It is interesting to observe that deterioration of SLL is larger for larger antenna arrays.

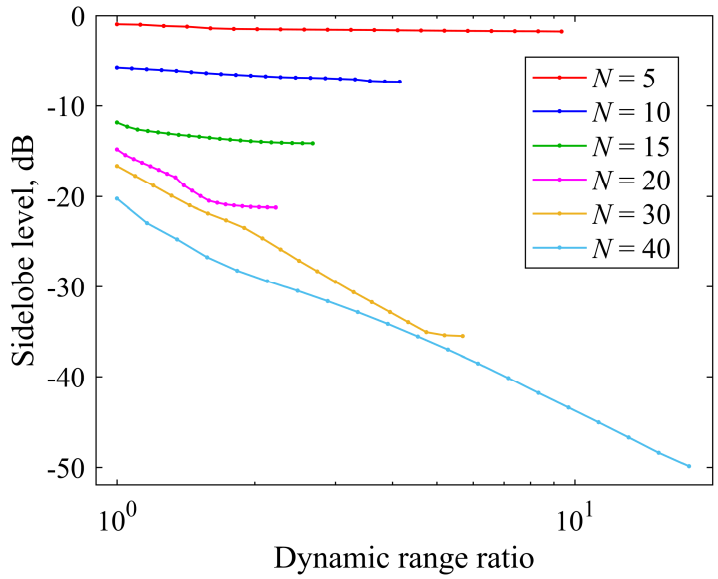

Figure 6.2 Optimum SLL obtained with proposed method for various numbers of elements, N, and required DRR, D.

6.3.3 Example 3

In this example, the influence of constraining DRR on the sidelobe level in antenna arrays with various numbers of elements is further investigated. The proposed method is applied to arrays with beamwidth of 10°, $N = 10, 15, 25$ and 40 and DRR constrained to values from 1 to D_{DC} logarithmically distributed in 20 points. Obtained results are compared with the method which minimizes the differences between coefficients assuming they are positive [61]. The comparison is shown in Figure 6.3. It is evident that the possibility of coefficients to take positive and negative values, which is ensured by the proposed method, results in a lower SLL.

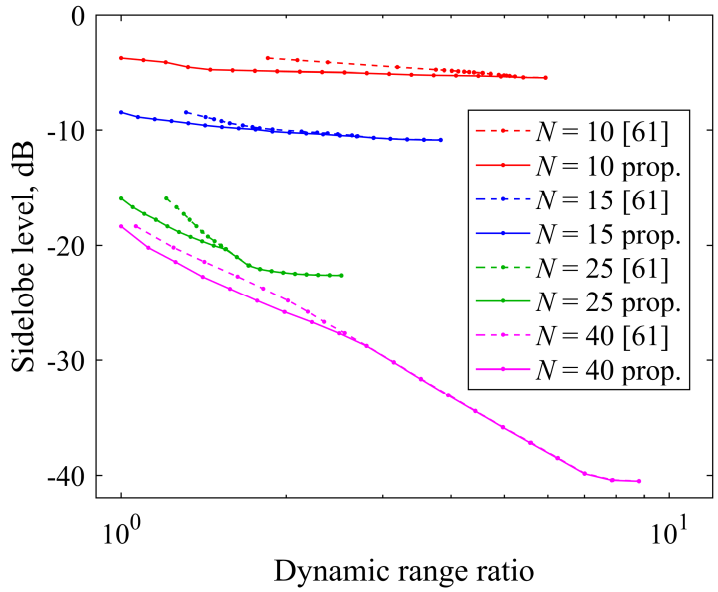

Figure 6.3 Comparison of SLL obtained with proposed method and method in [61] for antenna arrays with beamwidth of 10°, various N and DRR constrained in range from 1 to D_{DC}.

6.3.4 Example 4

Here, antenna array with 38 elements, beamwidth of 6° and DRR constrained to 1.7 is designed by using proposed method. The design results in an array with the sidelobe level of –20 dB. This sidelobe level is used as a desired value in a design of array with the method which allows only positive coefficients [61]. The coefficients obtained in both designs are shown in Figures 6.4 and 6.5. Coefficients of the proposed array are given in Table 6.2, for convenience. The beampatterns of both arrays are shown in Figure 6.6.

The method in [61] results in DRR of 2.079 and SLL of –19.974 dB whereas the proposed method provides DRR of 1.7 and the SLL of –20 dB. Clearly, the proposed method achieves a gain in DRR by allowing the coefficients to take negative values.

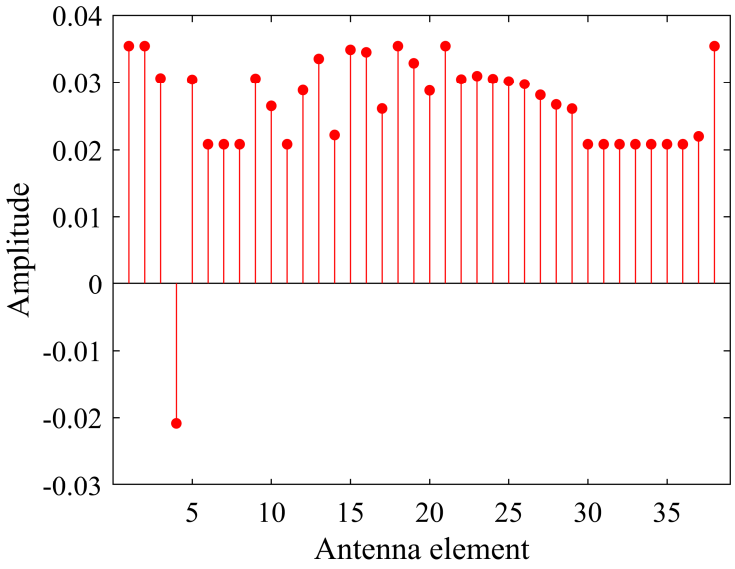

Figure 6.4 Coefficients of antenna array with 38 elements, beamwidth of 6°, DRR constrained to 1.7, and optimum SLL, obtained with proposed method.

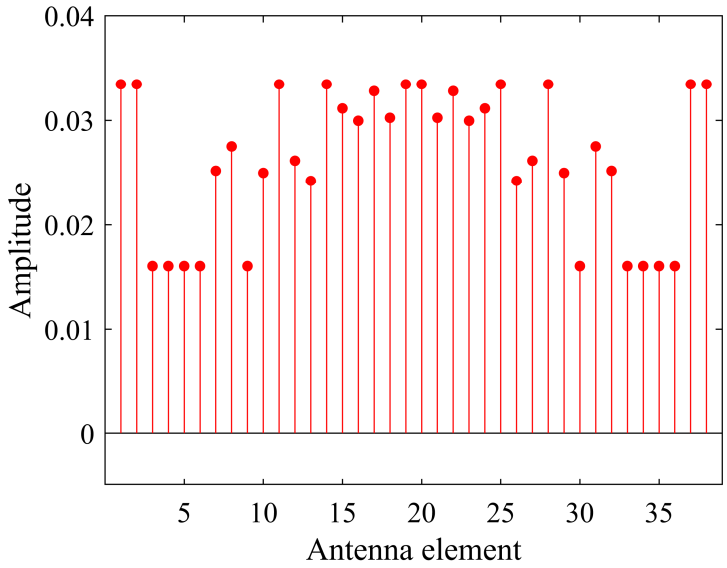

Figure 6.5 Coefficients of antenna array with 38 elements, beamwidth of 6°, specified SLL of −20 dB, and minimum DRR, obtained with method which allows only positive coefficients [61].

Table 6.2 Coefficients of pencil beam with $N = 38$, beamwidth of 6° and DRR constrained to 1.7, obtained with proposed method.

k	$a(k)$	k	$a(k)$	k	$a(k)$	k	$a(k)$
1	0.03543	11	0.02084	21	0.03543	31	0.02084
2	0.03543	12	0.02881	22	0.03046	32	0.02084
3	0.03062	13	0.03353	23	0.03096	33	0.02084
4	–0.02084	14	0.02220	24	0.03052	34	0.02084
5	0.03042	15	0.03487	25	0.03016	35	0.02084
6	0.02084	16	0.03451	26	0.02974	36	0.02084
7	0.02084	17	0.02609	27	0.02810	37	0.02199
8	0.02084	18	0.03543	28	0.02669	38	0.03543
9	0.03056	19	0.03288	29	0.02607		
10	0.02647	20	0.02876	30	0.02084		

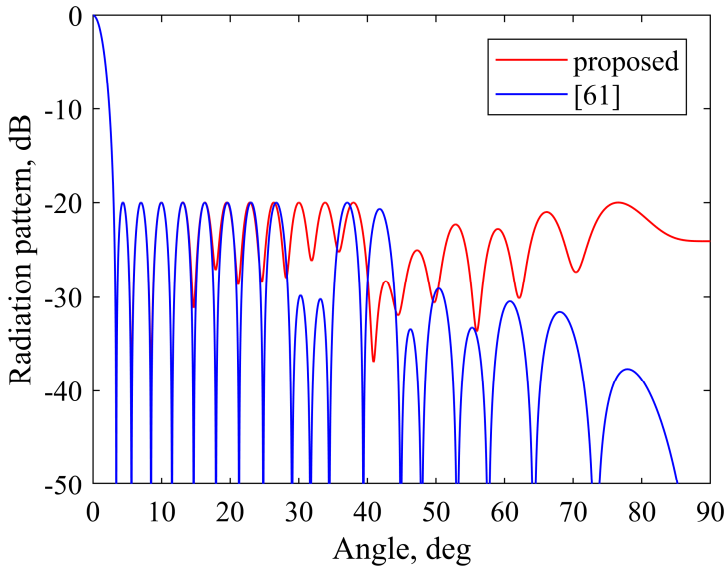

Figure 6.6 Radiation pattern of antenna array with 38 elements, beamwidth of 6° and SLL of –20 dB, obtained with proposed method and method in [61].

6.3.5 Example 5

The proposed method allows the design of arrays with equal absolute values of excitation coefficients. These coefficients are obtained when DRR is constrained to the value of $D = 1$. One such example is presented in Figure 6.7, which shows the optimum coefficients and Figure 6.8 which illustrates the corresponding radiation pattern of an array with 38

elements and beamwidth of 6°. The SLL of −17.7 dB is obtained. It should be noted that the set of available patterns for $D = 1$ is discrete. Consequently, it is not always possible to realize required beamwidth. Therefore, a required value of the beamwidth is in this case considered a targeted value, rather than exact specification.

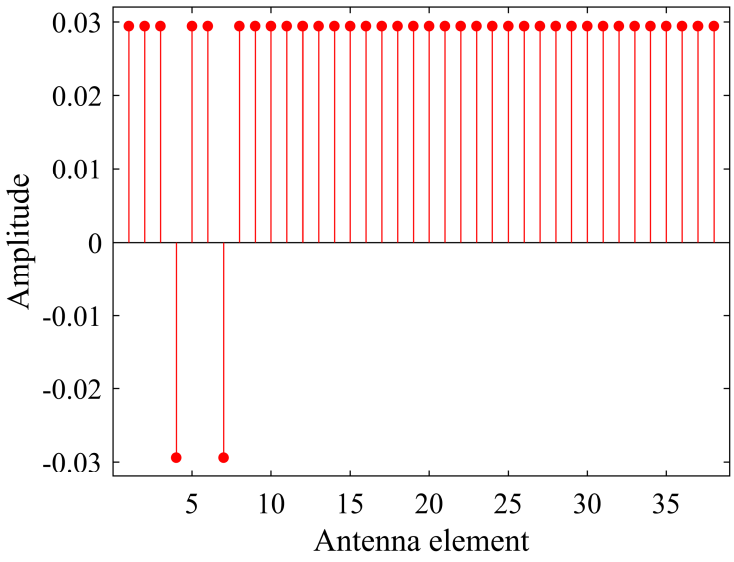

Figure 6.7 Coefficients of array obtained with proposed method for $N = 38$, beamwidth of 6° and $D = 1$.

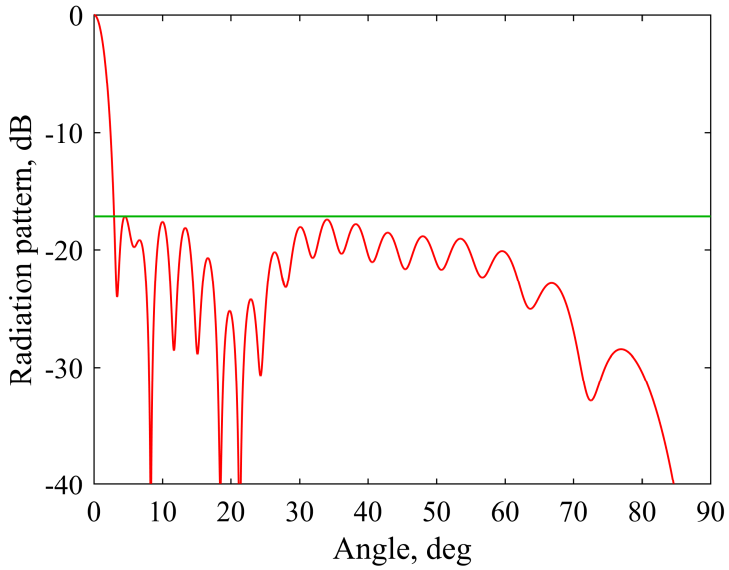

Figure 6.8 Radiation pattern of array obtained with proposed method for $N = 38$, beamwidth of 6° and $D = 1$. Green line represents the obtained SLL.

7 Global Optimization of Sparse Spatial Filters With Constrained Coefficients' Dynamic Range Ratio

A DRR-unconstrained design can result in a high value of DRR. Such is especially the case for arrays with steep slope of the main beam. Inevitably, some of the coefficients in such arrays take low magnitudes. On the other hand, forcing DRR to a small value causes small coefficients to rise, thus deteriorating the radiation pattern. The deterioration can be reduced if certain coefficients are allowed to take zero values. These coefficients do not participate in DRR and consequently increase design's freedom. Clearly, such an approach leads to sparse design.

7.1 DRR Constrained Design With Sparsity

In the nonsparse design (6.4) which is in Section 6.2 solved by using branch and bound algorithm, only positive and negative signs of coefficients are allowed. Here, the problem is extended by allowing coefficients to take zero values. However, this increases the complexity of design problem. If $S^+, S^-, S^0 \subseteq \{1, 2, ..., N\}$ are sets of indices denoting positive, negative and zero coefficients, the optimization problem in (6.5) becomes

$$
\begin{aligned}
\underset{\mathbf{s},\delta,t}{\text{minimize}} \quad & \delta \\
\text{subject to} \quad & \sum_k a_k = 1 \qquad k \in \left(S^+ \cup S^-\right) \\
& \left\|\mathbf{B}_q\mathbf{s}\right\| \le \delta, \quad q = 1,2,...,Q \\
& a_k \le Dt, \qquad k \in S^+ \\
& a_k \ge t, \qquad k \in S^+ \\
& -a_k \le Dt, \qquad k \in S^- \\
& -a_k \ge t, \qquad k \in S^- \\
& t \ge 0
\end{aligned}
\tag{7.1}
$$

where $\mathbf{s}$ is a vector containing a_k, $k \in (S^+ \cup S^-)$ and $\mathbf{B}_q$ is matrix composed of the columns $k \in (S^+ \cup S^-)$ taken from $\mathbf{A}_q$ defined in (6.2).

Figure 7.1 shows a part of the tree with coefficient signs and zero values corresponding to an array with $N = 4$. Here, the branching covers positive, negative and zero coefficients.

Like in the previous section, the branch and bound method with efficient pruning is utilized to find the array with minimum SLL [72].

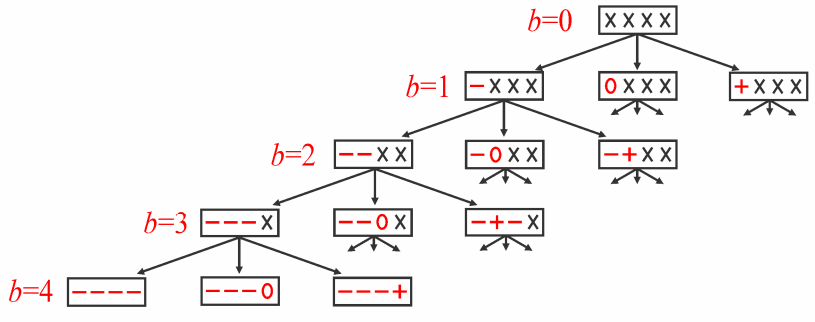

Figure 7.1 Part of tree with combinations of specified (+, –, 0) and unspecified (X) signs and zero values in antenna array with four elements. Specified signs and zero values in particular node are marked in red.

Compared to the tree described in section 6.2, allowing zero-valued coefficients rises the complexity of an algorithm from 2^N to 3^N. However, the proposed method allows efficient pruning. The method starts by setting the optimum SLL to $\delta_{opt} = \infty$. Then, similarly to a nonsparse design, nodes are processed by applying two tests. In the first test, the optimization problem in (7.1) is solved with the assumption that all unspecified (X) coefficients' elements, which are in Figure 7.1 marked in black, are positive (+). If $\delta < \delta_{opt}$ is obtained, an update $\delta_{opt} = \delta$ is made. In the second test, branches leaving the node are tested by solving the problem

$$
\begin{aligned}
\underset{\mathbf{s},\delta,t}{\text{minimize}} \quad & \delta \\
\text{subject to} \quad & \sum_k a_k = 1 \qquad k \in \left(S^+ \cup S^-\right) \\
& \left\|\mathbf{B}_q \mathbf{s}\right\| \le \delta, \quad q = 1,2,...,Q \\
& a_k \le Dt, \qquad k \in S^+ \;\&\; k \le b \\
& a_k \ge t, \qquad k \in S^+ \;\&\; k \le b \\
& -a_k \le Dt, \qquad k \in S^- \;\&\; k \le b \\
& -a_k \ge t, \qquad k \in S^- \;\&\; k \le b \\
& a_k \le Dt \qquad k = b+1, b+2, ..., N \\
& -a_k \le Dt \qquad k = b+1, b+2, ..., N \\
& t \ge 0
\end{aligned}
\tag{7.2}
$$

where b denotes the number of specified signs and zero values, as shown in Figure 7.1. The expression in (7.2) presents relaxed version of problem in (7.1). It should be noted that zero

coefficients are omitted from the constraints and objective function in both – (7.1) and (7.2). They are added into the response after the search has been finished, corresponding to the indices from the set S^0. If $\delta > \delta_{opt}$ is obtained by solving (7.2), it is certain that better SLL cannot be found below that node, so the corresponding node can be pruned. However, update must not be made, because solution of (7.2) solves relaxed and not original problem.

7.2 Design Examples

In the following examples, pencil beams with isotropic elements and spacing of $\lambda / 2$ are considered. The number of frequency points is set to $Q = 10N$ and the beamwidth of 12° is used unless otherwise noted.

7.2.1 Example 1

This example compares the proposed method and the exhaustive search. The comparison is given in Table 7.1 where the number of elements, N, and required DRR, D, are given together with the number of subproblems, N_{es}, and design time, t_{es}, which are estimated for the exhaustive search, as well as the number of subproblems, N_{bnb}, and design time, t_{bnb}, obtained by the proposed method. Here, the complexity of exhaustive search method is increased by introducing zero as a possible coefficient value. It requires solving of problem (7.1) 3^N times. Assuming the same personal computer is used as in a nonsparse design, it is estimated that solving each optimization problem requires 1 ms on average, so this value is used for calculating t_{es}.

Table 7.1 Comparison of number of subproblems and design time for exhaustive search, N_{es} and t_{es}, and proposed method, N_{bnb} and t_{bnb}, for sparse arrays with various numbers of elements, N, and required DRR, D.

N	D	N_{es}	t_{es}	N_{bnb}	t_{bnb}
10	2	59049	59.049 s	312	2.071 s
	1.5	59049	59.049 s	552	3.588 s
	1	59049	59.049 s	612	3.458 s
20	2	3.487e9	968.6 h	76	1.077 s
	1.5	3.487e9	968.6 h	228	3.212 s
	1	3.487e9	968.6 h	6368	1.239 min
30	2	2.059e14	6258.8 years	9196	3.627 min
	1.5	2.059e14	6258.8 years	40580	15.590 min
	1	2.059e14	6258.8 years	198832	1.092 h

By comparing tables 6.1 and 7.1, it is evident that sparse design is much more demanding. It should be kept in mind that both methods have combinatorial complexity, but for pencil beam arrays with small to medium number of elements, which are frequently used in communication systems, both methods give global results in an acceptable time.

7.2.2 Example 2

In this example, the optimum SLL obtained for both nonsparse and sparse arrays for various required DRR, D, and various numbers of antenna elements, N, are compared. The comparison is given in Figure 7.2 for pencil beams with $N = 10$, 15, 20 and 40. It is evident that by allowing some of the coefficients to be zero, SLL is improved in all cases. The improvement is bigger in arrays with higher number of elements and arrays with smaller DRR. Apparently, allowing coefficients to take zero values pays off.

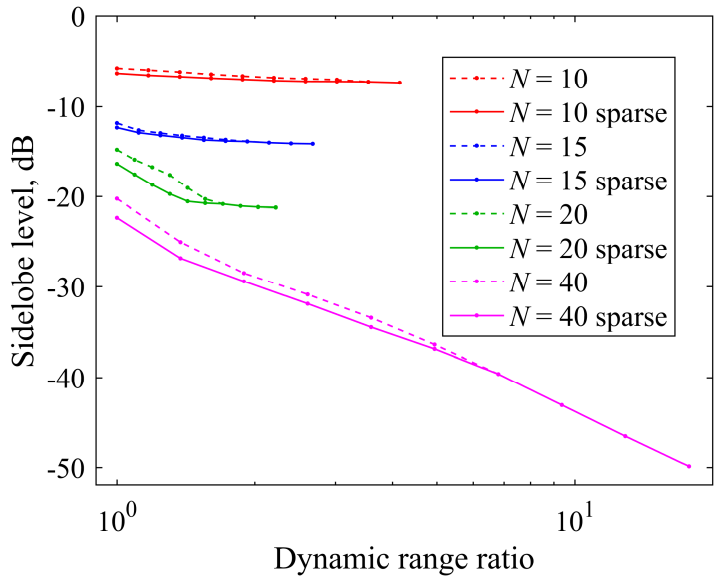

Figure 7.2 Optimum SLL obtained for nonsparse and sparse arrays for various numbers of elements, N, and required DRR, D.

7.2.3 Example 3

In this example, two arrays are designed, both with $N = 20$ and $D = 1.4$. One array is designed by using nonsparse method from chapter 6.2 whereas the other is obtained by using sparse design. Figures 7.3 and 7.4 show the coefficients' amplitudes, whereas Figure 7.5 shows the radiation patterns of both arrays. Both methods result in the same DRR, which is equal to D. However, the nonsparse method provides the SLL of −18.764 dB, whereas the

sparse method achieves –20.434 dB. It is evident that by allowing the coefficients to take zero value, new degrees of freedom are introduced. Therefore, better results both in terms of SLL and number of nonzero elements are obtained.

Coefficients of the proposed sparse array are given in Table 7.2, for convenience.

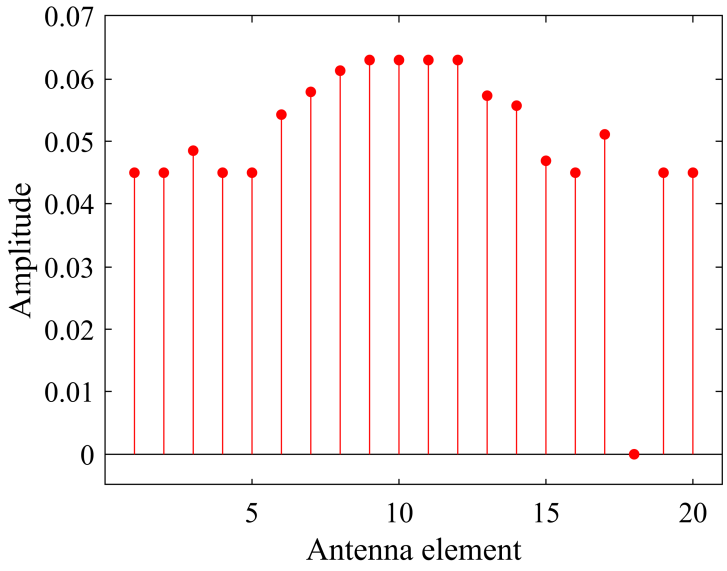

Figure 7.3 Coefficients of sparse pencil beam array with $N = 20$, beamwidth of 12° and $D = 1.4$.

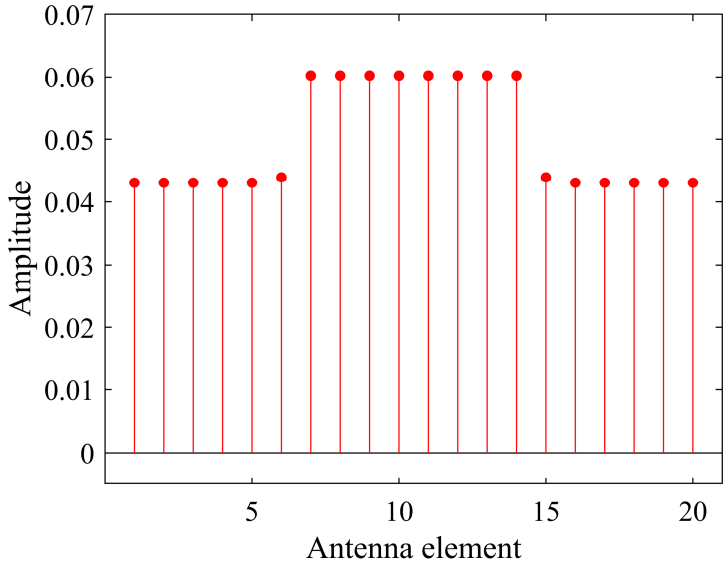

Figure 7.4 Coefficients of nonsparse pencil beam array with $N = 20$, beamwidth of 12° and $D = 1.4$.

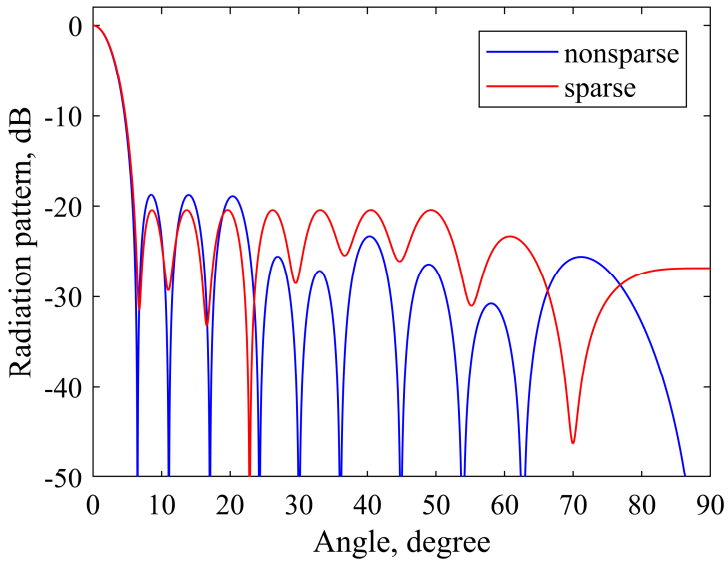

Figure 7.5 Radiation pattern of sparse and nonsparse pencil beam array with $N = 20$, beamwidth of 12° and $D = 1.4$.

Table 7.2 Coefficients of sparse pencil beam with $N = 20$, beamwidth of 12° and DRR constrained to 1.4, obtained with proposed method.

k	$a(k)$	k	$a(k)$
1	0.04501	11	0.06302
2	0.04501	12	0.06302
3	0.04851	13	0.05724
4	0.04501	14	0.05566
5	0.04501	15	0.04691
6	0.05423	16	0.04501
7	0.05784	17	0.05110
8	0.06135	18	0
9	0.06302	19	0.04501
10	0.06302	20	0.04501

7.2.4 Example 4

Here, the comparison of sparse and nonsparse methods is extended with the case when DRR is constrained to 1. Two arrays with $N = 20$, beamwidth of 10° and $D = 1$ are designed, one with the nonsparse method and one with the proposed sparse method. Pencil beam coefficients and radiation pattern for both nonsparse and sparse arrays are given in Figures 7.6, 7.7, and 7.8. The sparse array contains two zero elements. Instead, the nonsparse method

pushes one coefficient to negative value. Consequently, the former array exhibits the SLL of −15.065 dB, whereas the latter achieves the SLL of −13.443 dB.

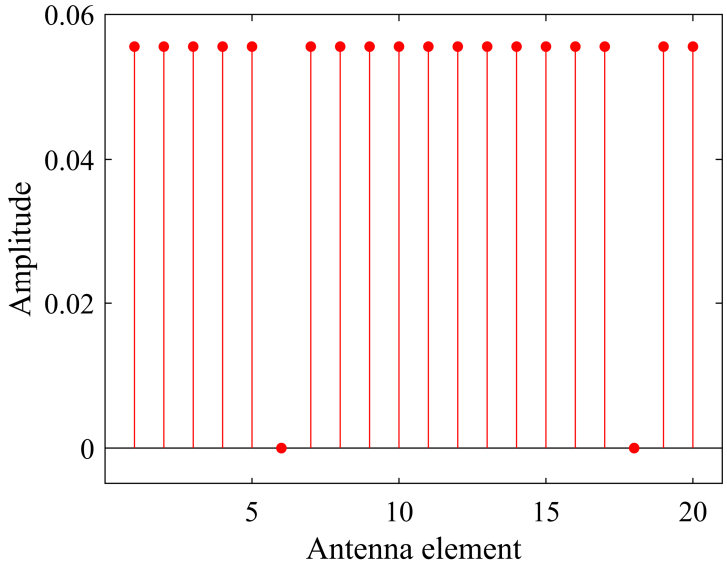

Figure 7.6 Coefficients of sparse pencil beam array with $N = 20$, beamwidth of 10° and $D = 1$.

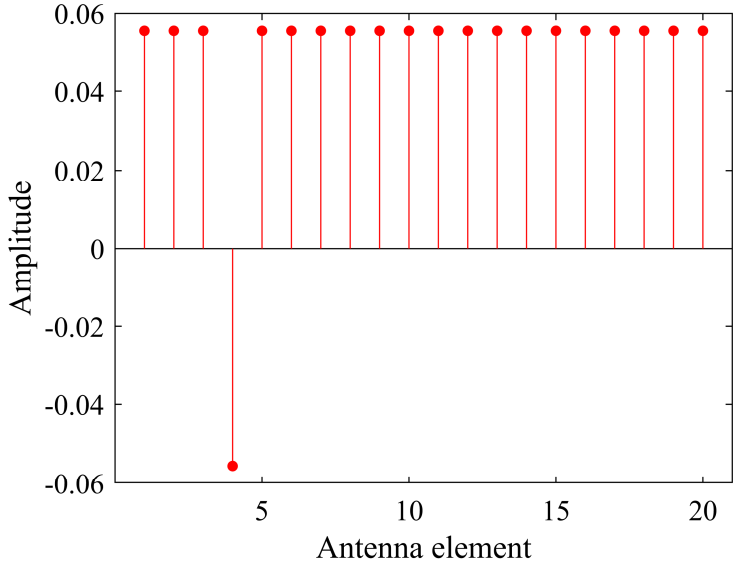

Figure 7.7 Coefficients of nonsparse pencil beam array with $N = 20$, beamwidth of 10° and $D = 1$.

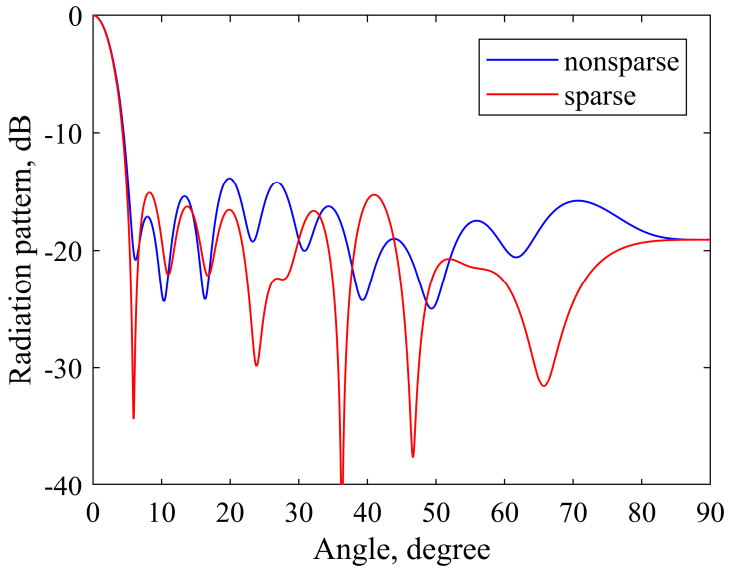

Figure 7.8 Radiation pattern of sparse and nonsparse pencil beam array with $N = 20$, beamwidth of 10° and $D = 1$.

8 Conclusion

In this dissertation, new methods for the design of efficient systems based on optimization techniques have been developed. The efficiency has been considered to lower complexity of digital and spatial filter design. Knowledge collected from the theory of compressed sensing has been applied to conventional filter and antenna array design. In this context, a contribution is achieved which consists of three parts.

In the first part, a new method for the design of sparse FIR filters based on signomial programming has been developed. The requirement for sparsity is usually expressed with l_0-norm. Here, the optimization problem has been approximated by relaxing the l_0-norm with an l_p-norm where $0 < p < 1$. The problem has been solved by forming the sequence of convex subproblems. This method, although iterative and local, provides good results. It is suitable for the design of low and medium order filters. The method has been successfully applied in the design of sparse FIR filters with linear phase and the filters with no phase specifications.

Although iterative methods are fast and robust, they cannot guarantee the optimum's globality. Therefore, an improvement has been achieved in global method for sparse filter design based on branch and bound algorithm. The method has been applied to sparse FIR filters with linear phase. Efficient pruning has enabled fast calculation of optimum filters in both peak- and quadratic-error sense, as well as in both senses. The method gives solutions for small and medium filter orders in an acceptable time. Furthermore, it provides the measure for testing local methods, as it can show how close their solutions are to global ones.

Finally, knowledge acquired in proposed FIR filter design has been applied in the design of spatial filters. A global optimization method has been developed which utilizes branch and bound algorithm in the design of pencil beams with minimax sidelobes and constrained dynamic range ratio. Furthermore, it has been shown that the obtained radiation pattern can be improved by introducing sparsity.

9 References

[1] S. K. Mitra, “Digital signal processing – a computer-based approach.”, New York: Mc-Graw Hill, 2011.

[2] A. V. Oppenheim, A. S. Willsky, and S. N. Hamid, “Signals & Systems 2nd edition”, Prentice-Hall Inc., 1996.

[3] T. W. Parks, and J. H. McClellan, “Chebyshev Approximation for Nonrecursive Digital Filters with Linear Phase,” IEEE Transactions on Circuit Theory, vol. 19, no. 2, pp. 189–194, March 1972.

[4] J. W. Adams, “FIR digital filters with least-squares stopbands subject to peak-gain constraints,” IEEE Transactions on Circuits and Systems, vol. 39, no. 4, pp. 376–388, April 1991.

[5] S. J. Orfanidis, “Electromagnetic waves and antennas,” Rutgers University, 2016.

[6] P. – S. Kildal, “Foundations of antenna engineering: A unified approach for line-of-sight and multipath,” Kildal Antenn AB, Gothenburg, Sweeden, 2015.

[7] Y. Lim, “Frequency-response masking approach for the synthesis of sharp linear phase digital filters,” IEEE Transactions on Circuits and Systems, vol. 33, no. 4, pp. 357–364, April 1986.

[8] E. J. Candes., J. Romberg, and T. Tao, “Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information,” IEEE Transactions on Information Theory, vol. 52, no. 2, pp. 489–509, January 2006.

[9] D. L. Donoho, M. Elad, and V. N. Temlyakov, “Stable recovery of sparse overcomplete representations in the presence of noise,” IEEE Transactions on Information Theory, vol. 52, no. 1, pp. 6-18, January 2006.

[10] D. L. Donoho, “Compressed sensing,” IEEE Transactions on Information Theory, vol. 52, no. 4, pp. 1289–1306, April 2006.

[11] R. G. Baraniuk, “Compressive sensing [Lecture notes],” IEEE Signal Processing Magazine, vol. 24, no. 4, pp. 118–121, July 2007.

[12] E. J. Candes, and M. B. Wakin, “An introduction to compressive sampling,” IEEE Signal Processing Magazine, vol. 25, no. 2, pp. 21–30, March 2008.

[13] M. F. Duarte, and Y. C. Eldar, “Structured Compressed Sensing: From Theory to Applications,” IEEE Transactions on Signal Processing, vol. 59, no. 9, pp. 4053–4085, September 2011.

[14] I. Rish, and G. Grabarnik, “Sparse modeling: theory, algorithms, and applications,” 1st edition, CRC press, 2014.

[15] S. Qaisar, R. M. Bilal, W. Iqbal, M. Naureen, and S. Lee, “Compressive sensing: From theory to applications, a survey,” Journal of Communications and Networks, vol. 15, no. 5, pp. 443–456, October 2013.

[16] M. F. Duarte, M. A. Davenport, D. Takhar, J. N. Laska, T. Sun, K. F. Kelly, and R. G. Baraniuk, “Single-pixel imaging via compressive sampling,” IEEE Signal Processing Magazine, vol. 25, no. 2, pp. 83–91, March 2008.

[17] M. J. Sun, M. P. Edgar, G. M. Gibson, B. Sun, N. Radwell, R. Lamb, and M. J. Padgett, “Single-pixel three-dimensional imaging with time-based depth resolution,” Nature Communications, vol. 7, pp. 1–6, May 2016.

[18] M. Lustig, D. Donoho, and J. M. Pauly, “Sparse MRI: The application of compressed sensing for rapid MR imaging,” Magnetic Resonance in Medicine, vol. 58, no. 6, pp. 1182–1195, October 2007.

[19] M. Lustig, D. L. Donoho, J. M. Santos, and J. M. Pauly, “Compressed Sensing MRI,” IEEE Signal Processing Magazine, vol. 25, no. 2, pp. 72–82, March 2008.

[20] M. Jurisic Bellotti, and M. Vucic, “Depth reconstruction of multiple light sources based on compressed sensing,” Proceedings of 2018 4th International Conference on Frontiers of Signal Processing, Poitiers, France, 24–27 September 2018, pp. 89–93.

[21] D. Wei, “Non-convex optimization for the design of sparse FIR filters,” in Proceedings of the 2009 IEEE/SP 15th Workshop on Statistical Signal Processing, Cardiff, UK, 31 August–3 September 2009, pp. 117–120.

[22] J. T. Kim, W. J. Oh, and Y. H. Lee, “Design of nonuniformly spaced linear-phase FIR filters using mixed integer linear programming,” IEEE Transactions on Signal Processing, vol. 44, no. 1, pp. 123–126, January 1996.

[23] J. L. H. Webb, and D. C. Munson, “Chebyshev optimization of sparse FIR filters using linear programming with an application to beamforming,” IEEE Transactions on Signal Processing, vol. 44, no. 8, pp. 1912–1922, August 1996.

[24] Y.-S. Song, and Y. H. Lee, “Design of sparse FIR filters based on branch-and-bound algorithm,” in Proceedings of the 1997 40th Midwest Symposium on Circuits and Systems, 3–6 August 1997, Sacramento, CA, USA, pp. 1445–1447.

[25] D. Mattera, F. Palmierl, and S. Haykin, “Efficient sparse FIR filter design,” in Proceedings of 2002 IEEE International Conference on Acoustics, Speech and Signal Processing, Orlando, FL, USA, 13–17 May 2002, pp. II–1537–II–1540.

[26] O. Gustafsson, L. S. DeBrunner, V. DeBrunner, and H. Johansson, “On the design of sparse half-band like FIR filters,” in Proceedings of the 2007 Conference Record of the Forty-First Asilomar Conference on Signals, Systems and Computers, Pacific Grove, CA, USA, 4–7 November 2007, pp. 1098–1102.

[27] T. Baran, D. Wei, and A. V. Oppenheim, “Linear programming algorithms for sparse filter design,” IEEE Transactions on Signal Processing, vol. 58, no. 3, pp. 1605–1617, March 2010.

[28] W. S. Lu, and T. Hinamoto, “Digital filters with sparse coefficients,” in Proceedings of the 2010 IEEE International Symposium on Circuits and Systems, Paris, France, 30 May–2 June 2010, pp. 169–172.

[29] C. Rusu, and B. Dumitrescu, “Iterative reweighted L1 design of sparse FIR filters,” Signal Processing, vol. 92, no. 4, pp. 905–911, April 2012.

[30] A. Jiang, H. K. Kwan, and Y. Zhu, “Peak-error-constrained sparse FIR filter design using iterative SOCP,” IEEE Transactions on Signal Processing, vol. 60, no. 8, pp. 4035–4044, August 2012.

[31] A. Jiang, and H. K. Kwan, “WLS design of sparse FIR digital filters,” IEEE Transactions on Circuits and Systems I: Regular Papers, vol. 60, no. 1, pp. 125–135, January 2013.

[32] D. Wei, C. K. Sestok, and A. V. Oppenheim, “Sparse filter design under a quadratic constraint: low-complexity algorithms,” IEEE Transactions on Signal Processing, vol. 61, no. 4, pp. 857–870, February 2013.

[33] D. Wei, and A. V. Oppenheim, “A branch-and-bound algorithm for quadratically-constrained sparse filter design,” IEEE Transactions on Signal Processing, vol. 61, no. 4, pp. 1006–1018, February 2013.

[34] H. Zhao, W. B. Ye, and Y. J. Yu, “Sparse FIR filter design based on genetic algorithm,” in Proceedings of IEEE International Symposium on Circuits and System, Beijing, China, 19–23 May 2013, pp. 97–100.

[35] R. Matsuoka, T. Baba, and M. Okuda, “Constrained design of FIR filters with sparse coefficients,” in Proceedings of Asia-Pacific Signal and Information Processing Association Annual Summit and Conference, Siem Reap, Cambodia, 9–12 December 2014, pp. 1–4.

[36] A. Jiang, H. K. Kwan, Y. Zhu, X. Liu, N. Xu, and Y. Tang, “Design of sparse FIR filters with joint optimization of sparsity and filter order,” IEEE Transactions on Circuits and Systems I: Regular Papers, vol. 62, no. 1, pp. 195–204, January 2015.

[37] Y. Yang, W. Zhu, and D. Wu, “Design of sparse FIR filters based on reweighted l_1-norm minimization,” in Proceedings of 2015 IEEE International Conference on Digital Signal Processing, Singapore, Singapore, 21–24 July 2015, pp. 858–862.

[38] W. Ye, and Y.J. Yu, “Greedy algorithm for the design of linear-phase FIR filters with sparse coefficients,” Circuits, Systems, and Signal Processing, vol. 35, no. 4, pp. 1427–1436, April 2016.

[39] A. Jiang, H. K. Kwan, Y. Zhu, X. Liu, N. Xu, and X. Yao, “Peak-error-constrained sparse FIR filter design using iterative l_1 optimization,” in Proceedings of 24th European Signal Processing Conference, Budapest, Hungary, 29 August–2 September 2016, pp. 180–184.

[40] L. Zheng, A. Jiang, and H. K. Kwan, “Sparse FIR filter design via partial l_1 optimization,” in Proceedings of IEEE International Symposium on Circuits and System, Baltimore, MD, USA, 28–31 May 2017, pp. 1–4.

[41] R. Matsuoka, S. Kyochi, S. Ono, and M. Okuda, “Joint sparsity and order optimization based on ADMM with non-uniform group hard thresholding,” IEEE Transactions on Circuits and Systems I: Regular Papers, vol. 65, no. 5, pp. 1602–1613, May 2018.

[42] H. K. Kwan, J. Liang, and A. Jiang, “Sparse FIR Filter Design using Iterative MOCSA,” in Proceedings of 2018 IEEE 61st International Midwest Symposium on Circuits and Systems (MWSCAS), Windsor, ON, Canada, 5–8 August 2018, pp. 952–955,

[43] Z. G. Feng, K. F. C. Yiu, and S.Y. Wu, “Design of sparse filters by a discrete filled function technique,” Circuits, Systems and Signal Processing, vol. 37, no. 10, pp. 4279–4294, October 2018.

[44] W. Chen, M. Huang, and X. Lou, “A branch-and-bound algorithm with reduced search space for sparse filter design,” in Proceedings of the 2018 IEEE Asia Pacific Conference on Circuits and Systems, Chengdu, China, 26–30 October 2018, pp. 329–332.

[45] H. K. Kwan, J. Liang, and A. Jiang, “Sparse Linear Phase FIR Filter Design using Iterative CSA,” in Proceedings of 2018 IEEE 23rd International Conference on Digital Signal Processing (DSP), Shanghai, China, 19–21 November 2018, pp. 1–4.

[46] T. Hirakawa, and M. Nakamoto, “A constraint optimization of low-delay and low-operation driven FIR digital filters for big data signal processing,” IEEJ Transactions on Electronics, Information and Systems, vol. 138, no. 4, pp. 299–305, 2018.

[47] W. Chen, M. Huang and X. Lou, “Design of sparse FIR filters with reduced effective length,” IEEE Transactions on Circuits and Systems I: Regular Papers, vol. 66, no. 4, pp. 1496–1506, April 2019.

[48] M. Jurisic Bellotti, and M. Vucic, “Design of nonlinear-phase FIR-filters based on signomial programming,” in Proceedings of the 2019 11th International Symposium on Image and Signal Processing and Analysis, Dubrovnik, Croatia, 23–25 September 2019, pp. 141–146.

[49] M. Jurisic Bellotti, and M. Vucic, “Sparse FIR filter design based on signomial programming,” Elektronika Ir Elektrotechnika, vol. 26, no. 1, pp. 40–45, February 2020.

[50] H. Wang, Z. Zhao, and L. Zhao, “Matrix Decomposition Based Low-Complexity FIR Filter: Further Results,” IEEE Transactions on Circuits and Systems I: Regular Papers, vol. 67, no. 2, pp. 672–685, February 2020.

[51] W. Chen, M. Huang, W. Ye, and X. Lou, “Cascaded Form Sparse FIR Filter Design,” IEEE Transactions on Circuits and Systems I: Regular Papers, vol. 67, no. 5, pp. 1692–1703, May 2020.

[52] A. Jiang, H. K. Kwan, Y. Tang, and Y. Zhu, “Sparse FIR Filter Design via Partial L_1-Norm Optimization,” IEEE Transactions on Circuits and Systems II: Express Briefs, vol. 67, no. 8, pp. 1482–1486, August 2020.

[53] X. Xi, and Y. Lou, “Sparse fir filter design with k-max sparsity and peak error constraints,” IEEE Transactions on Circuits and Systems II: Express Briefs, vol. 68, no. 4, pp. 1497–1501, April 2021.

[54] W. P. M. N. Keizer, “Low-sidelobe pattern synthesis using iterative fourier techniques coded in MATLAB [EM programmer's notebook],” IEEE Antennas and Propagation Magazine, vol. 51, no. 2, pp. 137–150, April 2009

[55] F. E. S. Santos, and J. A. R. Azevedo, “Adapted raised cosine window function for array factor control with dynamic range ratio limitation,” in Proceeding of 11th European Conference on Antennas and Propagation, Paris, France, 19–24 March 2017, pp. 2020–2024.

[56] G. Buttazzoni, and R. Vescovo, “Gaussian approach versus Dolph-Chebyshev synthesis of pencil beams for linear antenna arrays,” Electronic Letters, vol. 54, no. 1, pp. 8–10, January 2018.

[57] M. Matijascic, and G. Molnar, “Design of linear arrays forming pencil beams based on derivatives of chebyshev polynomials,” in Proceedings of 42nd International Convention on Information and Communication Technology, Electronics and Microelectronics, Opatija, Croatia, 20–24 May 2019, pp. 117–121.

[58] R. Vescovo, “Consistency of constraints on nulls and on dynamic range ratio in pattern synthesis for antenna arrays,” IEEE Transactions on Antennas and Propagation, vol. 55, no. 10, pp. 2662–2670, October 2007.

[59] G. K. Mahanti, A. Chakraborty, and S. Das, “Design of fully digital controlled reconfigurable array antennas with fixed dynamic range ratio,” Journal of Electromagnetic Waves and Applications, vol. 21, pp. 97–106, January 2007.

[60] M. Jurisic Bellotti, and M. Vucic, Global Optimization of Pencil Beams with Constrained Dynamic Range Ratio, in Proceedings of the 14th European Conference on Antennas and Propagation (EuCAP 2020), Copenhagen, Denmark, 15–20 March 2020, pp. 1–5.

[61] B. Fuchs, and S. Rondineau, “Array pattern synthesis with excitation control via norm minimization,” IEEE Transactions on Antennas and Propagation, vol. 64, no. 10, pp. 4228–4234, October 2016.

[62] X. Fan, J. Liang, and H. C. So, “Beampattern synthesis with minimal dynamic range ratio,” Signal Processing, vol. 152, pp. 411–416, November 2018.

[63] X. Fan, J. Liang, Y. Zhang, H. C. So, and X. Zhao, “Shaped power pattern synthesis with minimization of dynamic range ratio,” IEEE Transactions on Antennas and Propagation, vol. 67, no. 5, pp. 3067–3078, May 2019.

[64] R. M. Leahy, and B. D. Jeffs, “On the design of maximally sparse beamforming arrays,” IEEE Transactions on Antennas and Propagation, vol. 39, no. 8, pp. 1178–1187, August 1991.

[65] B. Fuchs, “Synthesis of Sparse Arrays With Focused or Shaped Beampattern via Sequential Convex Optimizations,” IEEE Transaction on Antennas and Propagation, vol. 60, no. 7, pp. 3499–3503, July 2012.

[66] G. Buttazzoni, and R. Vescovo, “Compressive sensing approach to the synthesis of sparse antenna arrays,” in Proceedings of the 12th European Conference on Antennas and Propagation, London, UK, 9–13 April 2018.

[67] D. Pinchera, M. D. Migliore, F. Schettino, M. Lucido, and G. Panariello, “An effective compressed-sensing inspired deterministic algorithm for sparse array synthesis,” IEEE Transactions on Antennas and Propagation, vol. 66, no. 1, pp. 149–159, January 2018.

[68] S. E. Nai, W. Ser, Z. L. Yu, and H. Chen, “Beampattern synthesis for linear and planar arrays with antenna selection by convex optimization,” IEEE Transactions on Antennas and Propagation, vol. 58, no. 12, pp. 3923–3930, December 2010.

[69] G. Buttazzoni, and R. Vescovo, “Pencil beam constrained synthesis of linear sparse arrays in presence of coupling effects,” in Proceedings of 2018 IEEE International Symposium on Antennas and Propagation, Boston, MA, 8–3 July 2018, pp. 2197–2198.

[70] X. Wang, M. Yao, D. Dai, and F. Zhang, “Synthesis of linear sparse arrays based on dynamic parameters differential evolution algorithm,” IET Microwaves, Antennas & Propagation, vol. 13, no. 9, pp. 1491–1497, July 2019.

[71] J. L. A. Quijano, M. Righero, and G. Vecchi, “Sparse 2-D array placement for arbitrary pattern mask and with excitation constraints: A simple deterministic approach,” IEEE Transactions on Antennas and Propagation, vol. 62, no. 4, pp. 1652–1662, April 2014.

[72] M. Jurisic Bellotti, and M. Vucic, “Global Optimization of Sparse Pencil Beams with Constrained Dynamic Range Ratio,” in Proceedings of the 2020 IEEE International Symposium on Antennas and Propagation and North American Radio Science Meeting (AP-S/URSI 2020), Montreal, Quebec, Canada, 5–10 July 2020, pp. 129–130.

[73] J. Nocedal, and S. J. Wright, “Numerical Optimization”, Springer, New York, NY, USA, 2006.

[74] A. Antoniou, and W. S. Lu, “Practical optimization: Algorithms and engineering applications,” Springer US, 2007.

[75] S. Boyd, and L. Vandenberghe, “Convex optimization,” Cambridge University Press, New York, USA, 2004.

[76] M. S. Lobo, L. Vandenberghe, S. Boyd, and H. Lebret, “Applications of second-order cone programming,” Linear Algebra and its Applications, vol. 284, no. 1, pp. 193–228, November 1998.

[77] S. Boyd, S.-J. Kim, L. Vandenberghe, and A. Hassibi, “A tutorial on geometric programming,” Optical Engineering, vol. 8, no. 1, pp. 67–127, March 2007.

[78] G. Xu, “Global optimization of signomial geometric programming problems,” European Journal of Operational Research, vol. 233, no. 3, pp. 500–510, March 2014.

[79] A. H. Land, and A. G. Doig, “An Automatic Method of Solving Discrete Programming Problems,” Econometrica, vol. 28, no. 3, pp. 497–520, July 1960.

[80] D. R. Morrison, S. H. Jacobson, J. J. Sauppe, and E. C. Sewell, “Branch-and-bound algorithms: A survey of recent advances in searching, branching, and pruning”, Discrete Optimization, vol. 19, pp. 79–102, 2016.

[81] MOSEK ApS, “The MOSEK optimization toolbox for MATLAB manual. Version 8.1,” 2017.

[82] J. F. Sturm, “Using SeDuMi1.02, a MATLAB toolbox for optimization over symmetric cones,” Optimization Methods and Software, vol. 11–12, pp. 625–653, 1999.

[83] M. Grant, and S. Boyd, “Graph implementations for nonsmooth convex programs,” Recent Advances in Learning and Control (a tribute to M. Vidyasagar), V. Blondel, S. Boyd, and H. Kimura, Lecture Notes in Control and Information Sciences, pp. 95-110, 2008.

[84] M. Grant, and S. Boyd, “{CVX}: Matlab Software for Disciplined Convex Programming, version 2.1,” March 2014.

[85] A. Mutapcic, K. Koh, S. J. Kim, and S. Boyd, “ggplab version 1.00: A Matlab toolbox for geometric programming,” 2006.

Biography

Maja Jurišić Bellotti was born in 1987 in Zenica, Bosnia and Herzegovina. She received the master's degree in Electronic and Computer Engineering Program from the University of Zagreb Faculty of Electrical Engineering and Computing, Zagreb, Croatia, in 2011.

From 2011 to 2016 she has been working as a Project Engineer in Altpro d.o.o., Zagreb. Since April 2016 she has been employed as a Research and Teaching Assistant at the University of Zagreb Faculty of Electrical Engineering and Computing. On the same faculty she enrolls in Doctoral study in winter semester 2016.

The scope of her doctoral research covers methods for the design of sparse conventional and spatial filters. She participated in the research project Beyond Nyquist Limit, IP-2014-09-2625, funded by Croatian Science Foundation led by Prof. Damir Seršić, PhD from 2016 to 2019. She is currently a researcher on the project Efficient Signal Processing Systems for Software Defined Radio, IP-2019-04-4189, led by Prof. Mladen Vučić, PhD.

Maja Jurišić Bellotti has been involved in teaching activities in the following courses: Signal Processing for Communications, Embedded System Design, Tools for Digital Design and Embedded Systems.

She is a member of the IEEE Association.

List of publications:

Journal papers

[1] M. Matijascic, M. Jurisic Bellotti, M. Vucic, and G. Molnar, *Optimum Synthesis of Pencil Beams with Constrained Dynamic Range Ratio,* International Journal of Antennas and Propagation, vol. 2022, pp. 1-14, November 2022.

[2] M. Jurisic Bellotti, and M. Vucic, *Sparse FIR Filter Design Based on Signomial Programming,* Elektronika Ir Elektrotechnika, vol. 26, no. 1, pp. 40–45, February 2020.

Papers in conference proceedings

[1] K. Vodvarka, M. Jurisic Bellotti, and M. Vucic, Synthesis of L1 Pencil Beams with Constrained Sidelobe Level and Dynamic Range Ratio, in Proceedings of the 16th European Conference on Antennas and Propagation (EuCAP 2022), Madrid, Spain, 27 March–1 April 2022, pp. 1–5.

[2] M. Jurisic Bellotti, and M. Vucic, *Global Optimization of Sparse Pencil Beams with Constrained Dynamic Range Ratio*, in Proceedings of the 2020 IEEE International Symposium on Antennas and Propagation and North American Radio Science Meeting (AP-S/URSI 2020), Montreal, Quebec, Canada, 5–10 July 2020, pp. 129–130.

[3] M. Jurisic Bellotti, and M. Vucic, *Global Optimization of Pencil Beams with Constrained Dynamic Range Ratio,* in Proceedings of the 14th European Conference on Antennas and Propagation (EuCAP 2020), Copenhagen, Denmark, 15–20 March 2020, pp. 1–5

[4] M. Jurisic Bellotti, and M. Vucic, *Design of Nonlinear-Phase FIR-Filters Based on Signomial Programming*, in Proceedings of the 11th International Symposium on Image and Signal Processing and Analysis, Dubrovnik, Croatia, 23–25 September 2019, pp. 141–146.

[5] M. Jurisic Bellotti, and M. Vucic, *Depth Reconstruction of Multiple Light Sources Based on Compressed Sensing*, in Proceedings of 2018 4th International Conference on Frontiers of Signal Processing, Poitiers, France, 24–27 September 2018, pp. 89–93.

Abstracts in conference proceedings

[1] K. Vodvarka, M. Jurisic Bellotti, and M. Vucic, *Global Design of L1 Pencil Beams With Multiple Constraints*, Abstract book of Seventh International Workshop on Data Science (IWDS 2022), Zagreb, Croatia, 26th October 2022, pp. 15–17

[2] K. Vodvarka, M. Jurisic Bellotti, and M. Vucic, *Global Design of Linear Antenna Arrays With Constrained Dynamic Range Ratio*, Abstract book of Sixth International Workshop on Data Science (IWDS 2021), Zagreb, Croatia, 24th November 2021, pp. 14–16

[3] M. Jurisic Bellotti, and M. Vucic, *Design of Sparse FIR Filters with no Phase Specifications Using Signomial Programming*, Abstract book of Fourth International Workshop on Data Science (IWDS 2019), Zagreb, Croatia, 15th October 2019, pp. 22–23

[4] M. Jurisic Bellotti, and M. Vucic, *Retrieving Scene Depth from Defocused Image Using Compressed Sensing*, Abstract book of Third International Workshop on Data Science (IWDS 2018), Zagreb, Croatia, 16th October 2018, pp. 12–13

[5] M. Jurisic Bellotti, and M. Vucic, *Compressed Sensing and its Application in Range Imaging*, Abstract book of Second International Workshop on Data Science (IWDS 2017), Zagreb, Croatia, 30th November 2017, pp. 15–16

Životopis

Maja Jurišić Bellotti rođena je 1987. u Zenici, Bosna i Hercegovina. Diplomirala je 2011. godine na Sveučilištu u Zagrebu, Fakultet elektrotehnike i računarstva, smjer Elektroničko i računalno inženjerstvo.

Od 2011. do 2016. zaposlena je kao projektni inženjer u firmi Altpro d.o.o. Od travnja 2016. zaposlena je kao znanstveni novak u suradničkom zvanju asistenta na Zavodu za elektroničke sustave i obradbu informacija Fakulteta elektrotehnike i računarstva. Na istom fakultetu upisuje postdiplomski doktorski studij elektrotehnike u zimskom semestru 2016. godine.

U okviru doktorskog studija bavi se metodama za dizajn rijetkih konvencionalnih i prostornih filtara. Sudjelovala je kao istraživač na znanstveno-istraživačkom projektu Iznad Nyquistove granice, IP-2014-09-2625, financiranom od Hrvatske zaklade za znanost pod vodstvom prof. dr. sc. Damira Seršića od 2016. do 2019. godine. Trenutno sudjeluje kao istraživač na projektu Učinkoviti sustavi za obradu signala namijenjeni programski definiranom radiju, IP-2019-04-4189, pod vodstvom prof. dr. sc. Mladena Vučića.

Pored znanstvenog rada, Maja Jurišić Bellotti uključena je u nastavne aktivnosti na predmetima Obrada signala u komunikacijama, Projektiranje ugradbenih računalnih sustava, Alati za razvoj digitalnih sustava te Ugradbeni računalni sustavi.

Član je udruženja IEEE.