niverzitet u Sarajevu
Elektrotehnički fakultet u Sarajevu
Odsjek za automatiku i elektroniku

Rekonstrukcija signala sa mogućnošću kompresije

Završni rad
I ciklusa studija

Mentor:

Red. prof. dr Melita Ahić-Đokić

Kandidat:

Nadir Kapetanović

Sarajevo, septembar 2012

Red. prof. dr Melita Ahić-Đokić, dipl.el.inž.
Viši asistent mr Emir Sokić, dipl.el.inž.
Odsjek za automatiku i elektroniku
Sarajevo, 30.11.2011.

Tema za završni rad

studenta I ciklusa studija koji studira na ETF-u u skladu sa principima Bolonjskog procesa na Odsjeku za automatiku i elektroniku (akad.g. 2011/12)

Tema: Rekonstrukcija signala sa mogućnošću kompresije
Student: Kapetanović Nadir

Postavka rada:

Fundamentalna Nyquist-Shannonova teorema o uzimanju uzoraka nam daje dovoljan uvjet koju mora zadovoljiti frekvencija uzorkovanja kako bi bilo moguće izvršiti dovoljno pouzdanu rekonstrukciju originalnog signala iz uzoraka. Međutim, za posebnu klasu signala tzv. signala sa mogućnošću kompresije, posljednja istraživanja su pokazala da je potreban uvjet za rekonstrukciju značajno povoljniji od dovoljnog uvjeta, odnosno da potrebna frekvencija uzorkovanja može biti značajno manja od Nyquistove. Ova oblast se naziva još i komprimirano uzorkovanje.

U okviru rada potrebno je napraviti pregled matematičkih osnova na kojima počiva komprimirano uzorkovanje, te modela koji se često koriste pri akviziciji i rekonstrukciji, sa naglaskom na algoritme koji ih omogućavaju.

Također, u okviru rada je poželjno ilustrirati algoritam akvizicije i rekonstrukcije poduzorkovanih signala (vremenski signal, 2D slika).

Koncept i metode rješavanja:

Za izradu softverskog dijela rada i računarsku analizu preporučuje se upotreba MATLAB softverskog paketa, mada je dozvoljeno korištenje i drugih okruženja (Mathcad, GNU Octave, Scilab, LabView, C/C++, Java itd.).

Polazna literatura:

1. Mark Davenport, Marco Duarte, Yonina Eldar, and Gitta Kutyniok, "*Compressed Sensing: Theory and Applications* ", Cambridge University Press , 2012.
2. Emmanuel Candès and Michael Wakin, *"* An introduction to compressive sampling", IEEE Signal Processing Magazine, 25(2), pp. 21 - 30, March 2008.
3. Marco F. Duarte, Yonina C. Eldar, " *Structured Compressed Sensing:From Theory to Applications*", IEEE Transactions on Signal Processing, 2011
4. David Donoho and Yaakov Tsaig, *„Extensions of compressed sensing.*", Signal Processing, 86(3), pp. 533-548, March 2006.
5. A. C. Gilbert, M. J. Strauss, J. A. Tropp, and R. Vershynin, *" One sketch for all: Fast algorithms for compressed sensing "*, University of Michigan, 2006.
6. Simon Foucart, *„Notes on Compressive Sensing"*, Vanderbilt University Department of Mathematics, 2009.
7. Melita Ahić-Đokić, *"Signali i sistemi",* Elektrotehnički fakultet u Sarajevu, 2010.

Potpis mentora:

Red. prof. dr Melita Ahić-Đokić

Sažetak

Teorija Compressive sensing-a/sampling-a predstavlja novu i izuzetno zanimljivu paradigmu uzorkovanja, koja se praktično u potpunosti suprotstavlja dosada korištenoj metodi uzorkovanja čije su temelje postavili Kotelnikov, Shannon, Nyquist i Whittaker, a koja daje samo dovoljan uslov za frekvenciju uzorkovanja. Naime, ova teorija zagovara uzorkovanje frekvencijom manjom od Nyquistove i uz to nudi načine rekonstrukcije originalnog signala na osnovu malog broja linearnih neadaptivnih mjerenja, koja su nerijetko slučajna. U cilju rekonstrukcije signala koji se najčešće susreću u praksi (pa su zato i od najvećeg značaja za ovakvu teoriju), ali i onih signala koji spadaju u domen teoretske analize, ovaj pristup procesu uzorkovanja predlaže strukturnu podjelu signala prema njihovoj prorijeđenosti u svom „prirodnom“ ili pak nekom transformacionom domenu. Ovo znači, da ukoliko je poznat domen u kojem signal ima mogućnost kompresije (jer su istinski prorijeđeni signali jako rijetki u praksi), to se on može predstaviti pomoću jako malo koeficijenata transformacije, a potom uzorkovati pomoću slučajne ili pak unaprijed struktuirane matrice mjerenja. Dakle, osnovni cilj CS-a jeste smanjenje broja mjerenja koja su potrebna da bi se izvršila tačna rekonstrukcija originalnog signala, te razvijanje novih sistema za uzorkovanje koji vrše slučajna mjerenja koja ipak odražavaju osobine originalnog signala i omogućavaju tačnu rekonstrukciju signala na osnovu takvih mjerenja.

Abstract

Compressive sensing/sampling theory poses a new and extremely interesting paradigm in the field of signal sampling and reconstruction, and is practically entirely opposite to the sampling method used today, the theory developed by Kotelnikov, Shannon, Nyquist and Whittaker, and which gives only a sufficient condition for sampling frequency. Namely, this theory promotes sampling process at a frequency lower than the Nyquist frequency, and offers different ways of signal reconstruction based on a small set of linear non-adaptive measurements, which are often random. In effort to reconstruct signals which are most commonly found in practice (and are most important for this theory), but also signals which have only a theoretical value, this approach to the sampling process suggests structural division of signals in respect to sparsity in their “natural”, or perhaps some transformation domain. This means that, if a domain in which a signal is compressive (truly sparse signals are really rare in practice) is known, then that signal can be represented by a very few transformation coefficients, and then sampled using a random or some more structured sensing matrix. Hence, the main goal of CS is to reduce the number of measurements needed to correctly reconstruct the original signal, and also to develop new sensing systems which implement random measurements that reflect the properties of the original signal, and futhermore enable precise signal reconstruction based on these measurements.

Sadržaj

1. Uvod 5

2. Proriježeni signali i signali sa mogućnošću kompresije 7
   - 2.1. Pojam signala i uzorkovanja 7
   - 2.2. Uvod u vektorske prostore 11
   - 2.3. Osnove i okviri vektorskih prostora 13
   - 2.4. Načini predstavljanja prorijeđenosti 15
   - 2.5. Signali sa mogućnošću kompresije 18

3. Matrice mjerenja 21
   - 3.1. Kreiranje matrica mjerenja 21
   - 3.2. Uslovi za nul-prostor 24
   - 3.3. Svojstvo ograničene izometrije 25
   - 3.4. Veza između svojstva ograničene izometrije i svojstva nul-prostora 26
   - 3.5. Koherentnost 27

4. Rekonstrukcija prorijeđenih signala l_1 minimizacijom 28
   - 4.1. Rekonstrukcija signala l_1 minimizacijom 28
   - 4.2. Rekonstrukcija signala bez šuma 31
   - 4.3. Rekonstrukcija signala sa šumom 32

5. Algoritmi za rekonstrukciju sparse signala (sparse recovery) 34
   - 5.1. Algoritmi za rekonstrukciju sparse signala (sparse recovery) 34
   - 5.2. Metode bazirane na konveksnoj optimizaciji 35
   - 5.3. Greedy algoritmi 37
   - 5.4. Kombinatorički algoritmi 38

6. Primjena signala sa mogućnošću kompresije u praksi 40
   - 6.1. Medicina: Rekonstrukcija slika magnetne rezonance, elektroencefalografija 40
   - 6.2. Jednopikselna kamera 43

7. Rekonstrukcija signala sa mogućnošću kompresije u Matlabu 46
   - 7.1. Rekonstrukcija signala koji je sparse u vremenskom domenu 46
   - 7.2. Rekonstrukcija signala koji je sparse u frekventnom domenu 51

8. Zaključak 57

9. Literatura 58

Lista korištenih skraćenica 60

1. Uvod

Revolucija digitalizacije čiji smo svakodnevni svjedoci promoviše razvoj i primjenu novih sistema mjerenja, tj. akvizicije podataka, sa sve višom pouzdanošću, brzinom i rezolucijom. U centru dosada primjenjivanih sistema za analogno-digitalnu konverziju, odnosno uzorkovanje (općenito nazvano mjerenjem), nalazi se upravo Teorema o uzorkovanju za koju su zaslužni Kotelnikov, Shannon, Nyquist i Whittaker koji su nezavisno jedan od drugog, i za različite oblasti primjene, dali dovoljan uslov koji treba postaviti na frekvenciju uzorkovanja da bi rekonstrukcija frekventno ograničenog signala, koji je kontinualan u vremenu, bila moguća. Bazirajući se na ovom principu, digitalizacija bilo koje vrste signala, bilo da je riječ o zvučnom signalu, slici, videu ili dr., je omogućila konstrukciju sistema uzorkovanja i obrade koji su robusniji, fleksibilniji, jeftiniji, a samim tim i puno češće korišteni u odnosu na njihove analogne ekvivalente.

Veliki nedostatak ovakve paradigme jeste to što u cilju povećanja kvalitete digitaliziranih podataka (npr. slike ili videa) dolazi do velikog porasta Nyqistove frekvencije, što za rezultat često ima prevelik broj uzoraka, od kojih u većina nije nužno potrebna, pa se daljom kompresijom takvi uzorci i odbacuju. I pored ogromnih napredaka na polju računarske moći, u praksi se često mogu susresti primjeri gdje nije moguće napraviti uređaje koji mogu uzorkovati zahtijevanom frekvencijom, jer je previsoka, ili je konstrukcija takvih uređaja jednostavno tehnološki neisplativa. Takođe, uzastopni procesi uzorkovanja, a potom kompresije su često jako zahtjevni sa stanovišta računarskih resursa, jer se vrši akvizicija ogromnog broja uzoraka od kojih većina naknadno biva otklonjena procesom kompresije. Ovo se protivi zdravom razumu i neefikasno je.

Nyquistov uslov je samo dovoljan, ali ne i potreban uslov za izbor frekvencije uzorkovanja pa se može postaviti pitanje: „Da li je moguće vršiti uzorkovanje frekvencijom manjom od Nyquistove, tj. da li je moguće vršiti mjerenja samo 'bitnih' vrijednosti signala tako da je rekonstrukcija na osnovu tih malobrojnih mjerenja zaista moguća?“. Odgovor na ovo pitanje je potvrdan, što otvara sasvim nove oblasti istraživanja i povezivanja primijenjene matematike i informacionih nauka sa medicinom, tehnologijom snimanja fotografija, video snimaka i muzike, te mnogim drugim naučim oblastima. Nova teorija o uzorkovanju, simbolično nazvana uzorkovanjem sa mogućnošću kompresije (eng. Compressive sensing/sampling, skraćeno CS) predstavlja sasvim novi okvir za akviziciju signala i dizajn senzora. CS omogućava potencijalno veliku redukciju broja potrebnih uzoraka za preciznu rekonstrukciju signala na osnovu istih, te posljedično i smanjenje računarskih resursa potrebnih za rekonstrukciju signala.

Osnovna ideja koju CS podcrtava jeste zaobilaženje „duplog posla“ uzorkovanja i kompresije, upravo direktnim uzorkovanjem u komprimiranoj formi, tj. pri nižoj frekvenciji uzorkovanja. Za razvoj teorije CS-a zaslužni su Emmanuel Candes, Justin Romberg, Terence Tao i David Donoho, koji su pokazali da se signali konačne dimenzionalnosti koji ispunjavaju određene uslove, mogu rekonstruisati na osnovu malog broja linearnih, neadaptivnih mjerenja. Pored toga, u radu će biti pokazano da je rekonstrukcija signala na osnovu slučajnih mjerenja moguća, što se kosi sa uobičajenim načinom razmišljanja, a što leži u samoj srži CS-a.

U Poglavlju 2. definiše se pojam signala i uzorkovanja, zajedno sa isječcima iz naučnih radova Kotelnikova [1], Shannona [2], Nyquista [3] i Whittakera [4] koji su postavili temelje moderne i dosad najčešće korištene metode uzorkovanja. Takođe, tu je i uvod u vektorske prostore, osnove i okvire vektorskih prostora kao matematička osnova za daljnje razmatranje osnovne ideje koju CS promoviše. Nadalje se daje definicija svojstva prorijeđenosti i posljedično prorijeđenih signala, te signala sa mogućnošću kompresije koji će u nastavku rada u većini slučajeva i biti razmatrani (pošto su signali istinski prorijeđeni u svom „prirodnom“ domenu izuzetnu rijetki, i skoro se nikada ne susreću u praksi).

Poglavlje 3. sadrži osnovne ideje o dizajnu matrica mjerenja, koje bi omogućile uzorkovanje upravo samo „bitnih“ vrijednosti signala, i teoretsku osnovu istih, tj. uslove i svojstva koja takva matrica mora da zadovoljava. Sva svojstva navedena u poglavlju su i matematički opisana, te su date osnovne teoreme vezane za ovaj segment rada, kao dokaz da je ono što mnogi smatraju nemogućim i matematički neopisivim u suštini moguće ma koliko se protivilo „zdravom razumu“ i dosadašnjem smjeru razmišljanja (o uslovima uspješne rekonstrukcije na osnovu uzorkovanih podataka), koje je jako zatvoreno i na neki način ograničeno sopstvenom teorijom koja ga opisuje.

Poglavlja 4. i 5. se bave rekonstrukcijom prorijeđenih signala i signala sa mogućnošću kompresije, s tim da je u Poglavlju 3. data matematička osnova ovakve rekonstrukcije, tj. njen matematički opis, dok su u Poglavlju 4. dati algoritmi koji rade na osnovu ovih teoretskih postavki, te tako povezuju primijenjenu matematiku sa informacionim naukama.

Praktična primjena CS-a je obrađena i Poglavlju 6., s posebnim osvrtom na njegovu primjenu u medicini za rekonstrukciju slika magnetne rezonance i elektroencefalografije. U ovim primjerima je podcrtana očita prednost CS-a u poređenju sa standardnim Nyquistovim uzorkovanjem u pogledu kako broja potrebnih uzoraka tako i računarskih resursa potrebnih za rekonstrukciju ovakvih medicinskih snimaka, pri čemu je vrijeme pregleda drastično skraćeno.

Poglavljem 7. ovaj rad je zaokružen, jer ovo poglavlje daje praktične rezultate teorije koja je predstavljena u prvih pet poglavlja. Naime, u ovom poglavlju su obrađeni primjeri rekonstrukcije signala sa mogućnošću rekonstrukcije, pomoću programskog paketa Matlab, u dva reprezentativna primjera signala sa karakterističnim osobinama.

Na kraju je dat popis literature i korištenih skraćenica.

U radu se nije ulazilo u detalje matematičke teorije koja podupire obrađenu temu nego su navedene samo osnovne teoreme (bez dokaza) i pojmovi nužni za shvatanje koncepata koji stoje iza teme rada, upravo iz razloga što su mnoge stvari izuzetno komplicirane, te u većini slučajeva čak i izlaze iz okvira gradiva Prvog studijskog ciklusa (BSC).

2. Proriježđeni signali i signali sa mogućnošću kompresije

2.1. Pojam signala i uzorkovanja

Postupak rekonstrukcije signala sa mogućnošću kompresije ne može biti obrađen bez definiranja osnovnih pojmova vezanih za analizu signala i sistema kao što su signal, kontinualni i diskretni signali, te teorema o uzorkovanju koja predstavlja dovoljan uslov za mogućnost rekonstrukcije kontinualnog signala na osnovu uzoraka tog istog signala. Cilj ovog rada jeste da pokaže da je za neke klase signala moguće, uz nepoštivanje teoreme o uzorkovanju, izvršiti rekonstrukciju.

Naime, signal u užem smislu (čija će se definicija koristiti u nastavku) predstavlja vremenski zavisnu fizičku veličinu. Ako se ovakva definicija poopšti, signal se može definirati i kao fizička veličina koja može biti zavisna ne samo od vremena, nego i od prostornih koordinata, ili pak nekih drugih nezavisno promjenljivih. Signal sam po sebi ne nosi informaciju, nego je informacija sadržana u promjeni signala, pa se tako vrši pretvorba informacije u signal kada ju je potrebno obraditi, pohraniti ili pak prenijeti na daljinu. Bitno je napomenuti da informacija ne može postojati bez promjene prethodnog stanja. Od dimenzionalnosti informacije zavisi i dimenzionalnost signala koji nosi tu informaciju. Tako se za zvučni signal kaže da je jednodimenzionalni (1-D) signal kod kojeg je vrijeme nezavisna promjenljiva, dok je npr. signal slike dvodimenzionalni (2-D) signal kod kojeg postoje dvije prostorne promjenljive.

S obzirom na informacioni aspekt, moguća je podjela svih signala na determinističke i slučajne signale. Deterministički signal je onaj signal za koji je njegov fizički opis u potpunosti poznat, bilo da je dat analitički (matematički) ili grafički. Deterministički signali se nadalje mogu podijeliti na periodičke i aperiodičke signale. Slučajni (eng. *random*) signali su određeni samo u domenu vjerovatnog opisa, tj. sve vrijednosti takvih signala se ne mogu u potpunosti opisati pomoću nekog analitičkog izraza, pa se za opisivanje ovakvih signala koristi funkcija vjerovatnoće, kao što su srednja vrijednost, srednje kvadratna vrijednost, varijansa, funkcija raspodjele i dr. S obzirom na to da buduća vrijednost slučajnog signala nije unaprijed poznata, lako se zaključuje da zapravo samo slučajni signali nose informaciju.

Za neki signal $f(t)$ se kaže da je kontinualan ukoliko je $f(t)$ realna ili kompleksna jednoznačno određena vremenska funkcija, definirana u svakom vremenskom trenutku t. Kontinualni signali se mogu podijeliti u više klasa, i to na periodičke i aperiodičke, zatim na parne i neparne, te konačno na signale snage i energije.

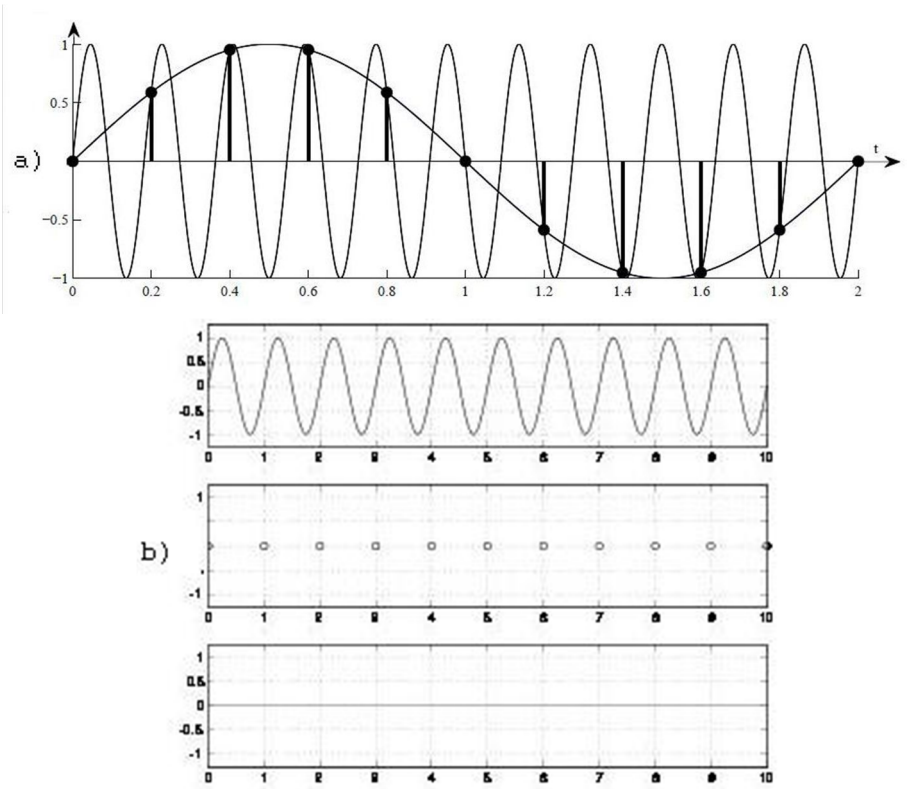
Diskretni signal, za razliku od kontinualnog, nije definiran u svakom vremenskom trenutku t, nego samo u diskretnim vremenskim trenucima $t = kT$ u kojima je izvršeno uzorkovanje kontinualnog signala, pri čemu je k cijeli broj, a T predstavlja interval uzorkovanja. Dakle, diskretni signal je predstavljen nizom (sekvencom) vrijednosti kontinualnog signala u trenucima odabiranja. Odabiranje, tj. postupak diskretizacije se može vršiti uniformno i neuniformno, s tim da se uniformna diskretizacija najčešće koristi iz

praktičnih razloga, ali i zbog jednostavnijih matematičkih alata koji su potrebni. Sekvenca f[n] u diskretnom vremenu n se definira kao: $f[n] = \lim_{\varepsilon \to 0} f[kT + \varepsilon]$, $\varepsilon > 0$, odnosno kao granična vrijednost kontinualnog signala sa desne strane [5].

U nastavku slijedi kratka teorijska osnova postupka uzorkovanja koja je sve do skora zauzimala centralnu ulogu u analizi signala i sistema, pogotovo s razvojem digitalne tehnologije koja može operisati samo sa diskretnim signalima upravo zbog ograničenosti memorije, ali i ograničene dužine memorijskih registara (digitalnih riječi), što opet povlači mogućnost postizanja samo konačne, u zadatom opsegu željene tačnosti.

Poznato je da se uzorkovanje može posmatrati kao proces modulacije koji kontinualni signal transformiše u povorku impulsa (odabiraka) čije su površine jednake amplitudama kontinualnog signala u odgovarajućim trenucima. Bitno je napomenuti da nije svejedno kako se bira period uzorkovanja, jer npr. u slučaju loše odabranog perioda uzorkovanja, dvije ili više kontinualnih sinusoida različitih frekvencija mogu prilikom uzorkovanja davati istu povorku impulsa (slika 2.1. a)). Također, ukoliko je period uzorkovanja jednak periodu sinusoide, dobija se sekvenca istih vrijednosti, te je nemoguće na osnovnu takve sekvence razlučiti da li je originalni kontinualni signal sinusoidalan ili pak konstantan (slika 2.1. b)).

Očito je da bilo koji diskretni signal ima beskonačno mnogo „originalnih“ kontinualnih signala, pa je upravo ova pojava poznata kao aliasing efekat [5].



Slika 2.1. (a) Aliasing usljed uzorkovanja signala periodom uzorkovanja većim od perioda signala (Slika preuzeta iz [6]) (b) Aliasing efekat usljed uzorkovanja signala periodom uzorkovanja jednakim periodu signala (Slika preuzeta iz [5])

Teoreme i leme korištene u ovom radu će biti navođene bez dokaza .

Lema 2.1. Da bi aliasing bio izbjegnut, signal koji se uzorkuje mora na neki način biti frekventno ograničen, drugim riječima njegova Fourierova transformacija (spektar) mora biti jednaka nuli izvan konačnog opsega učestanosti, tj. ukoliko je $\mathrm{f(t)} \overset{F.T.}{\longleftrightarrow} F(j\omega)$, to mora biti zadovoljeno $F(j\omega) = 0$ za učestanosti $|\omega| > \omega_m$, gdje je ω_m maksimalna učestanost u spektru [5].

U naučnoj zajednici je općeprihvaćeno mišljenje da je prvu preciznu formulaciju teoreme o uzorkovanju (Teorema 2.1.) za primjene u komunikacijskom inžinjeringu postavio Vladimir Aleksandrovich Kotelnikov [4]:

Teorema 2.1. Bilo koja funkcija $F(t)$ koja se sastoji od frekvencija od 0 do f_1 $[Hz]$ se može predstaviti pomoću sljedećeg reda:

$$F(t) = \sum_{-\infty}^{\infty} D_k \frac{\sin \omega_1 \left(t - \frac{k}{2f_1}\right)}{t - \frac{k}{2f_1}}, \tag{2.1.}$$

gdje je k cijeli broj, $\omega_1 = 2\pi f_1$, D_k – niz konstanti koje zavise od $F(t)$. Jasno, svaka funkcija $F(t)$ koje je predstavljena redom (2.1.) sadrži frekvencije od 0 do f_1 $[Hz]$ [1].

Claude Elwood Shannon je 1948. godine objavio svoju verziju teoreme o uzorkovanju, iako ju je napisao već 1940. godine, a dugo vrijeme potrebno za objavljivanje je bilo posljedica situacije nakon Drugog svjetskog rata. U notaciji i formulaciji Shannona, teorema o uzorkovanju predstavljena je Teoremom 2.2. i glasi:

Teorema 2.2. Neka f(t) ne sadrži frekvencije veće od W. Tada je

$$f(t) = \sum_{-\infty}^{\infty} X_n \frac{sin\pi(2wt - n)}{\pi(2wt - n)}, \tag{2.2.}$$

gdje je $X_n = f\left(\frac{n}{2W}\right)$.

Očito je da nakon linearnih transformacija i odgovarajućeg odabira parametara, formulacija teoreme o uzorkovanju postaje ista kao Kotelnikova formulacija, odnosno za $f(t) = C(t)$ dobija se notacija po E.T.Whittakeru, tj.:

$$C(t) = \sum_{-\infty}^{\infty} f(a + k\omega) \frac{sin\frac{\pi}{\omega}(t - a - k\omega)}{\frac{\pi}{\omega}(t - a - k\omega)}. \tag{2.3.}$$

Zanimljivo je napomenuti da se često za E.T.Whittakera smatra da je inicirao teoremu o uzorkovanju, no međutim njegov rad je u mnogome potpomognut radom japanskog matematičara K. Ogura-e iz 1915. godine u kojem je naveo teorem o uzorkovanju u sličnom obliku koji je i danas poznat, te sugerirao način izvođenja strogog dokaza koristeći rezultate iz više matematike – reziduuma [4].

Na osnovu svega rečenog može se zaključiti da, ukoliko je zadovoljena Lema 2.1., kontinualni signal f(t) (slika 2.2. (a)) se može rekonstruisati iz spektra

$$F^*(j\omega) = \frac{1}{T}\sum_{k=-\infty}^{k=+\infty} F(j\omega - jk\omega_0) \tag{2.4.}$$

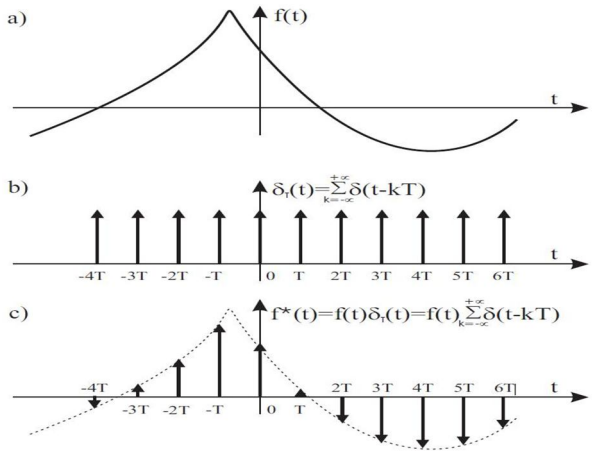
(slika 2.3. (b)). $F^*(j\omega)$ je F.T. zvjezdaste funkcije

$$f^*(t) = f(t)\delta_T(t) = f(t)\sum_{k=-\infty}^{k=+\infty}\delta(t-kT) = \sum_{k=-\infty}^{k=+\infty} f(kT)\delta(t-kT) \tag{2.5.}$$

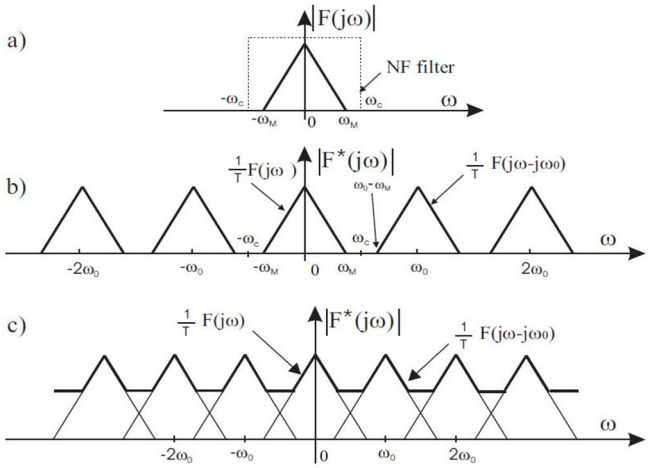
(slika 2.2. (c)). Rekonstrukcija signala $f(t)$ iz $F^*(j\omega)$ je moguća ukoliko je učestanost ω_0 izabrana tako da u spektru $F^*(j\omega)$ nema preklapanja između dva susjedna spektra, odnosno dovoljno je da vrijedi:

$$\omega_0 - \omega_M \geq \omega_M \Rightarrow \omega_0 \geq 2\omega_m \text{ ili } f_0 \geq 2f_M \text{ ili } T \leq \frac{\pi}{\omega_M} = \frac{1}{2f_M} \tag{2.6.}$$

gdje je T period uzorkovanja, f_0 frekvencija uzorkovanja, a f_M maksimalna (granična) frekvencija spektra F($j\omega$) kontinualnog signala f(t). Uslov (2.6.) se naziva **Nyquistov uslov**, te on određuje upravo maksimalan period uzorkovanja, odnosno minimalnu frekvenciju uzorkovanja [5].



Slika 2.2. Idealni diskretizirani signal - zvjezdasta funkcija (a) Kontinualni signal $f(t)$ (b) Povorka impulsa $\delta_T(t)$ (c) Zvjezdasta funkcija $f^*(t)$ (Slika preuzeta iz [5])



Slika 2.3. (a) Spektar originalnog kontinualnog signala (b) spektar zvjezdaste funkcije spektralno ograničenog kontinualnog signala bez preklapanja spektara (c) spektar zvjezdaste funkcije spektralno ograničenog kontinualnog signala sa preklapanjem spektara (Slika preuzeta iz [5])

2.2. Uvod u vektorske prostore

U svrhu lakšeg shvatanja matematičkih alata koji predstavljaju srž teorije o uzorkovanju signala koji imaju mogućnost kompresije (eng. ***compressive sensing,*** u daljem tekstu CS), u ovom poglavlju se definišu pojmovi norme, vektorskog i normiranog vektorskog prostora. Bez poznavanja ovih osnovnih matematičkih pojmova i njihovog dobrog razumijevanja nije moguće ulaziti u bilo kakvo dublje razmatranje teme ovog rada.

U realnim vektorskim prostorima $\mathbb{R}^2$ *i* $\mathbb{R}^3$ definira se intenzitet (modul) vetora kao dužina duži kojom je taj vektor predstavljen, odnosno kao rastojanje vrha vektora od ishodišta. Međutim, pojam intenziteta vektora $\mathbb{R}^2$ *ili u* $\mathbb{R}^3$ se proširuje i na vektore u proizvoljnom vektorskom prostoru.

Definicija 2.1. Neka je X vektorski prostor nad poljem skalara $\mathbb{R}$ ili $\mathbb{C}$. **Norma** na X je svako preslikavanje $\| \|: X \to \mathbb{R}$, koje zadovoljava sljedeće uslove (*aksiome norme*) :

a) $\|x\| \geq 0, \|x\| = 0 \Leftrightarrow x = 0_x$ (0_x – neutralni/nula element u skupu X),
b) $\|\lambda x\| = |\lambda| \cdot \|x\|$ (homogenost norme),
c) $\|x + y\| \leq \|x\| + \|y\|$ (nejednakost trougla),

za sve $x, y \in \mathbb{R}$ i za svaki skalar $\lambda \in \mathbb{R}$ (odnosno $\lambda \in \mathbb{C}$). Uređeni par $(X, \| \|)$ vektorskog (realnog, odnosno kompleksnog) prostora X i norme na X zove se **normirani** (realni, odnosno kompleksni) **(vektorski) prostor**. Vrijednost $\|x\|$ za $x \in \mathbb{R}$ zove se **norma vektora** x [7].

Za realan broj $p \geq 1$, ℓ_p norma se definira relacijom:

$$\|x\|_p = \begin{cases} \left(\sum_{i=1}^{N} |x_i|^p \right)^{\frac{1}{p}}, & p \epsilon [1, \infty); \\ \max\limits_{i=1,2,\ldots,N} |x_i|, & p = \infty. \end{cases} \qquad (2.7.)$$

Analiza signala se uglavnom fokusirala na signale koje proizvode fizički sistemi. Mnogi prirodni signali i sistemi koje je napravio čovjek se mogu modelirati kao linearni. Dakle, prirodno je razmatrati modele sistema koji komplementiraju ovakvoj linearnoj strukturi. Ovakva predodžba je sadržana u modernoj analizi signala upravo modeliranjem signala kao ***vektora*** koji postoje u odgovarajućem ***vektorskom prostoru***. Ovo zadržava linearnu strukturu koja se često priželjkuje, tj. da ako se saberu dva signala dobija se treći signal, koji ima fizičko značenje. Takođe, vektorski prostori dozvoljavaju primjenu intuicije i alata iz geometrije u $\mathbb{R}^3$ kao što su dužine, razdaljine, te uglovi, da bi se opisali i uporedili signali koji su od interesa. Ovo je korisno čak i kada signali egzistiraju u višedimenzionalnim ili čak beskonačnodimenzionalnim prostorima.

U ovom radu, signali će biti razmatrani kao realne funkcije čiji su domeni bilo kontinualni ili diskretni, te konačni ili beskonačni. Ove pretpostavke će biti naglašene u svakom poglavlju po potrebi. U nastavku slijedi kratak pregled nekih od ključnih koncepata vezanih za vektorske prostore koji će biti potrebni za razvijanje teorije CS-a. Po pravilu će biti

razmatrani **normirani vektorski prostori**, tj. vektorski prostori koji imaju normu. U slučaju diskretnog, konačnog domena, signali se mogu posmatrati kao vektori u N-dimenzionalnom Euklidskom prostoru (označenom sa $\mathbb{R}^N$).

Kada se bude bavilo vektorima u $\mathbb{R}^N$ Euklidskom prostoru, često će biti korištene ℓ_p norme, koje se za $p\epsilon[1,\infty]$ definiraju relacijom (2.7.).

U Euklidskom prostoru se također može razmatrati standardni unutrašnji proizvod u $\mathbb{R}^N$, koji se obilježava sa:

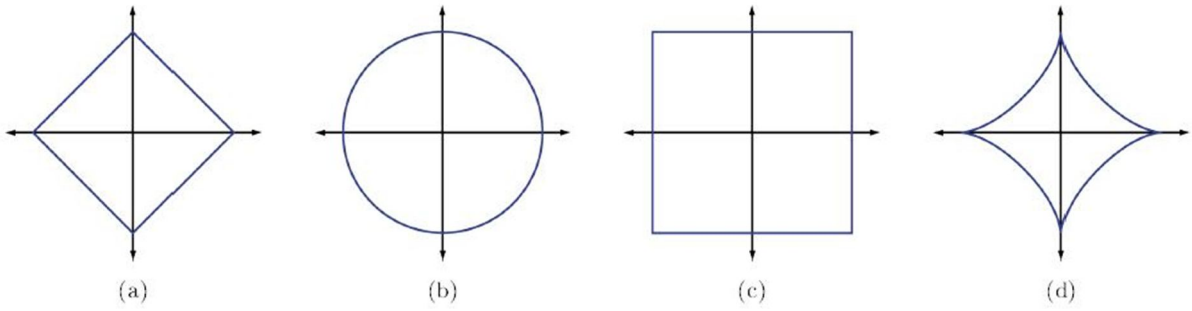
$$< x, z > = z^T x = \sum_{i=1}^{N} x_i z_i. \tag{2.8.}$$

Ovaj unutrašnji proizvod vodi do ℓ_2 norme: $\|x\|_2 = \sqrt{< x, x >}$.

U nekim kontekstima je korisno proširiti pojam ℓ_p normi za slučaj kada je $p < 1$. U tom slučaju, „norma" definirana relacijom (2.7.) ne zadovoljava nejednakost trougla, pa se naziva ***kvazinormom***. Često će biti korištena oznaka $\|x\|_0 := |supp(x)|$, gdje $supp(x) = \{i: x_i \neq 0\}$ označava skup nenultih vrijednosti x, a $|supp(x)|$ označava kardinalnost od $supp(x)$. Primijetimo da $\|\cdot\|_0$ nije čak ni kvazinorma, ali se lako dokazuje da je:

$$\lim_{p\to 0} \|x\|_0 \, x^p = \ |supp(x)|, \tag{2.9.}$$

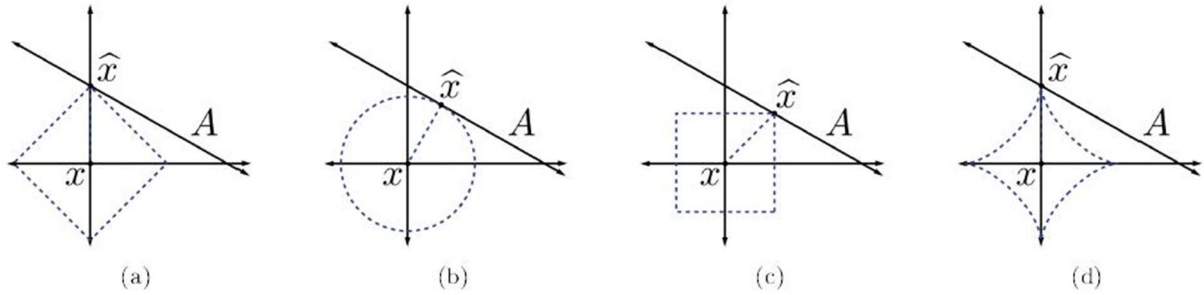
što opravdava ovakav izbor označavanja. ℓ_p (kvazi)norme imaju bitno drugačija svojstva za različite vrijednosti p. Ovo se može ilustrirati slikom 2.4., na kojoj je prikazana jedinična sfera, tj. $\{x: \|x\|_p = 1\}$, navedena za svaku od ovih normi u $\mathbb{R}^2$. Uočljivo je da je za $p < 1$ odgovarajuća jedinična sfera nekonveksna (što upravo odražava nezadovoljenje nejednakosti trougla kod kvazi-norme).



Slika 2.4. Jedinične kružnice u $\mathbb{R}^2$ Euklidskom prostoru za ℓ_p norme za $p = 1, 2, \infty$, i za ℓ_p normu za $p = 0.5$.
(a) Jedinična kružnica za ℓ_1 normu (b) Jedinična kružnica za ℓ_2 normu (c) Jedinična kružnica za ℓ_∞ normu
(d) Jedinična kružnica za ℓ_p kvazi-normu

Obično se norme koriste kao mjera jačine signala, ili veličine greške. Npr. pretpostavimo da je dat neki signal $x \in \mathbb{R}^2$ i da želimo da ga aproksimiramo koristeći tačku u jednodimenzionalnom afinom prostoru A. Ako mjerimo grešku aproksimacije koristeći ℓ_p normu, onda je naš zadatak da pronađemo $\hat{x} \in A$ koje minimizira $\|x - \hat{x}\|_p$. Izbor p će imati bitan utjecaj na svojstva rezultirajuće greške aproksimacije. Primjer je ilustriran na slici 2.5. Da bi se izračunala najbliža tačka u A u odnosu na x koristeći ℓ_p normu, možemo zamisliti rast ℓ_p sfere centrirane na x sve dok ne dođe do presjeka sa A. Ovo će biti upravo tačka

$\hat{x} \in A$ koja je najbliža u odnosu na x, u odgovarajućoj ℓ_p normi. Može se primijetiti da veće p ima tendenciju da ravnomjernije rasprostire grešku između dva koeficijenta, dok manje p dovodi do greške koja je neravnomjernije raspoređena i ima tendenciju da se raspršuje. Ova intuicija se generalizira i za više dimenzije, te igra jako bitnu ulogu u razvoju teorije CS-a [8].



Slika 2.5. Najbolja aproksimacija tačke u $\mathbb{R}^2$ Euklidskom prostoru jednodimenzionalnim podprostorom koristeći ℓ_p norme za $p = 1, 2, \infty$, i za ℓ_p normu za $p = 0.5$. (a) Aproksimacija u ℓ_1 normi (b) Aproksimacija u ℓ_2 normi (c) Aproksimacija u ℓ_∞ normi (d) Aproksimacija u ℓ_p kvazi-normi [8]

2.3. Osnove i okviri vektorskih prostora

Skup $\Psi = \{\psi_i\}_{i \in I}$ se naziva bazom/osnovom za vektorski prostor v konačne dimenzionalnosti, ako su vektori u skupu širine v i ako su linearno nezavisni. Ovo implicira da se svaki vektor u prostoru može predstaviti kao linearna kombinacija (manjeg, osim trivijalnog slučaja) skupa baznih vektora i to na jedinstven način. Nadalje, koeficijenti ove linearne kombinacije se mogu naći na osnovu unutrašnjeg proizvoda signala i vektora koji u vektorskom prostoru predstavljaju funkcije iz skupa funkcija kojima se vrši aproksimacija signala. U diskretnim uslovima, razmatrat će se samo realni Hilbertovi prostori konačnog broja dimenzija gdje je v=$\mathbb{R}^N$ i $I = \{1, \ldots, N\}$ [8].

Bitno je napomenuti da je kontinualni signal $x(t)$ u ovom i svim sljedećim teoretskim razmatranjima CS-a zapravo predstavljen kao realni, jednodimenzionalni diskretni signal $x[n]$ konačnog trajanja (pri čemu su prilikom uzorkovanja/diskretizacije zadovoljeni Lema 2.1. i uslov 2.6., tj. zadovoljena je Teorema o uzorkovanju) koji može biti predstavljen kao vektor kolona dimenzija $Nx1$ koji se nalazi u $\mathbb{R}^N$ vektorskom prostoru, a čiji su elementi vrijednosti x(n) sekvence $x[n]$, gdje je $n = 1,2, \ldots, N$. Signali više dimenzionalnosti, kao što je npr. slika, se pretvaraju u dugi jednodimenzionalni vektor [9].

Matematički, bilo koji signal $x \in \mathbb{R}^N$ se može izraziti kao:

$$x = \sum_{i \in I} a_i \widetilde{\psi_i}, \tag{2.10.}$$

gdje se koeficijenti računaju kao $a_i = < x, \psi_i >$, a skup $\{\psi_i\}_{i \in I}$ su vektori koji čine dualnu bazu. Drugi način označavanja baze i dualne baze jeste način na koji oni operišu nad x. Ovdje se dualna baza $\widetilde{\Psi}$ naziva **bazom sinteze** (koristi se za rekonstrukciju signala relacijom 2.8.), a Ψ **bazom analize**.

Ortonormirana baza (u daljem tekstu će biti korištena skraćenica ONB) je definirana kao skup vektora $\Psi = \{\psi_i\}_{i \in I}$ koji formiraju bazu i čiji su elementi ortogonalni i imaju jediničnu normu. Drugim riječima, $<\psi_i, \psi_j>=0$ ako je $i \neq j$ a $<\psi_i, \psi_j>=1$ ako je $i = j$. U slučaju ONB-a, baza sinteze je jednostavno Hermitska adjunkta (adjungovana matrica) baze analize, tj. vrijedi da je $\widetilde{\Psi} = \Psi^T$.

Često je korisna generalizacija koncepta baze koja omogućava da skupovi sadrže moguće linearno zavisne vektore, što za rezultat ima ono što se naziva **okvir** (eng. *frame*). Formalnije, okvir je skup vektora $\{\Psi_i\}_{i=1}^{n}$ u $\mathbb{R}^d$, d<n koji odgovara matrici $\Psi \in \mathbb{R}^{dxn}$, takvoj da za sve vektore $x \in \mathbb{R}^d$ vrijedi:

$$A\|x\|_2^2 \leq \|\Psi^T\|_2^2 \leq B\|x\|_2^2 \tag{2.11.}$$

gdje je $0 < A \leq B < \infty$. Primjećuje se dakako da uslov A>0 implicira da redovi matrice Ψ moraju biti linearno nezavisni. Kada se A izabere kao najveća moguća vrijednost, a B kao najmanja moguća vrijednost, da bi ove nejednakosti bile zadovoljene, onda ih nazivamo **(optimalnim) granicama okvira** (eng. o*ptimal frame bounds*). Ako se A i B mogu odabrati tako da vrijedi A=B, onda se okvir naziva A-uskim (engl. *A-tight*), a ako je A=B=1, onda je Ψ **Parsevalov okvir**. Okvir se naziva jednakim po normi, ako postoji neko λ>0 takvo da je $\|\Psi_i\|_2 = \lambda$ za svako i=1, ..., N, a jedinična je norma ukoliko je λ=1. Primjetno je da, iako je koncept okvira jako generalan i može biti definiran u beskonačnodimenzionalnim prostorima, u slučaju kada je Ψ matrica dimenzija dxN, A i B jednostavno odgovaraju najmanjoj i najvećoj sopstvenoj vrijednosti izraza $\Psi\Psi^T$, respektivno.

Okviri mogu pružiti bogatije predstavljanje podataka s obzirom na njihovu redundantnost: za dati signal x postoji beskonačno mnogo vektora koeficijenata α takvih da je $x = \Psi\alpha$. Svaki izbor dualnog okvira $\widetilde{\Psi}$ pruža drugačiji izbor vektora koeficijenata α. Formalnije, bilo koji okvir koji zadovoljava relaciju:

$$\Psi\widetilde{\Psi}^T = \widetilde{\Psi}\Psi^T = I \tag{2.12.}$$

se naziva **(alternativni) dualni okvir** (eng. *alternate dual frame*). Specijalni izbor $\widetilde{\Psi} = (\Psi\Psi^T)^{-1}\Psi$ se naziva **kanonski dvojni okvir** (eng. *canonical dual frame*). Također je poznat pod nazivom Moore-Penrose pseudoinverzna matrica (eng. *pseudoinverse*). Očito je da, pošto uslov A>0 zahtijeva da Ψ ima linearno nezavisne redove, to je osigurano da je moguće naći inverznu matricu od proizvoda $\Psi\Psi^T$, tj. da je $\widetilde{\Psi}$ nesingularna. Dakle, jedan način da je dobije skup podesnih koeficijenata jeste izraz:

$$\alpha_d = \Psi^T(\Psi\Psi^T)^{-1}x. \tag{2.13.}$$

Može se dokazati da je ova sekvenca zapravo sekvenca najmanjih koeficijenata u ℓ_2 normi, tj. $\|\alpha_d\|_2 \leq \|\alpha\|_2$ za svako α za koje vrijedi da je $x = \Psi\alpha$.

Konačno, uobičajeno je da se u literaturi koja se bavi problemom aproksimacije prorijeđenih signala (signala koji, kada se predstave u obliku vektora, pretežno imaju nultu

vrijednosti u domenu u kojem su predstavljeni; u daljem tekstu eng. *sparse*), baza ili okvir nazivaju **rječnikom** ili **redundantnim rječnikom** respektivno, s tim da se elementi rječnika nazivaju **atomima** [8].

2.4. Načini predstavljanja prorijeđenosti

Transformiranje signala po novoj bazi ili okviru može omogućiti predstavljanje signala u konciznijoj, sažetoj formi. Rezultirajuća kompresija je korisna za smanjenje količine pohranjenih i poslanih podataka, što u nekim primjenama može predstavljati priličnu uštedu resursa. Dakle, poželjno je jednostavno poslati koeficijente analize dobijene iz baznog razvoja ili razvoja okvira, umjesto slanja njegove višedimenzionalne međusobne veze. U slučajevima kada je broj nenultih koeficijenata mali, kaže se da postoji tzv. sparse predstavljanje signala. Modeli sparse signala omogućavaju postizanje kompresije većih razmjera, a u slučaju CS-a se može koristiti spoznaja da je signal sparse u poznatoj bazi ili okviru da bi rekonstruisali originalni signal na osnovu malog broja mjerenja. Za sparse podatke, samo se nenulti koeficijenti pohranjuju ili šalju u mnogim slučajevima, dok se ostali koeficijenti mogu smatrati nultim.

Matematički, za signal x se kaže da je K-sparse kada ima najviše K nenultnih koeficijenata, tj. $\|x\|_0 \leq K$. Neka:

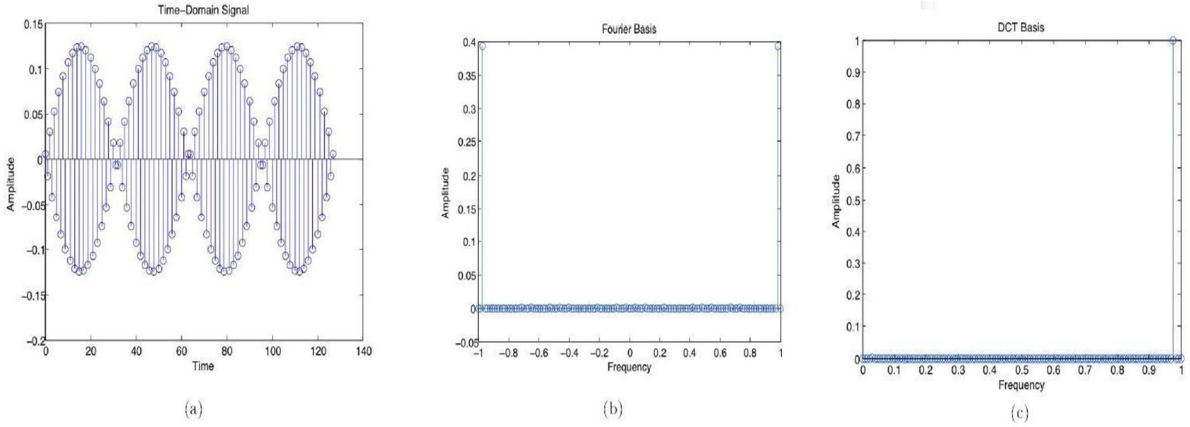
$$\sum\nolimits_K = \{x: \|x\|_0 \leq K\} \tag{2.14.}$$

označava skup svih K-sparse signala. U praksi se najčešće ne sreću signali koji su sparse po prirodi, ali je njihova reprezentacija u nekoj bazi Ψ često sparse (npr. kompresija slike). U ovom slučaju za signal x se opet kaže da je K-sparse, jasno imajući u vidu da se x može izraziti kao $x = \Psi\alpha$ gdje je $\|\alpha\|_0 \leq K$.

Prorijeđenost, tj. svojstvo signala da, kada se u nekom domenu (bilo vremenskom ili npr. *wavelet* domenu) predstave u obliku vektora pretežno imaju nulte vrijednosti (u nastavku teksta engl. *sparsity*), se dugo koristila u analizi signala i teoriji aproksimacije za potrebe procesa kompresije, uklanjanja šuma, kao i u statistici i teoriji učenja kao metod izbjegavanja prevelikog preklapanja. Sparsity također igra bitnu ulogu u oblastima kao što su teorija statističkih procjena i odabira modela, u studijama ljudskog sistema organa vida. Također je masovno iskorištena u procesima obrade slike, s obzirom da *multiscale wavelet transformacija* pruža skoro pa sparse predstavljanja za prirodne slike. U nastavku su opisani neki jednodimenzionalni i dvodimenzionalni primjeri sparse signala, odnosno sparse predstavljanja signala [8].

2.4.1. 1-D modeli signala

U ovom dijelu će biti prikazan primjer tri bazna razvoja koja daju različite nivoe sparsity-a za isti signal. Jednostavni periodički signal je uzorkovan i predstavljen kao povorka impulsa čije su pojedinačne površine proporcionalne vrijednosti signala u trenutku uzorkovanja (slika 2.6.(a)). Uzorkovanje se može predstaviti kao bazna ekspanzija pri čemu su elementi baze impulsi koji se nalaze u periodičkim tačkama duž vremenske ose. Dualna baza se u ovom slučaju sastoji od Sa funkcija korištenih za rekonstrukciju signala na osnovu uzoraka. Ovakvo predstavljanje signala sadrži mnogo nenultih koeficijenata, a s obzirom na periodičnost signala mnogo je redundantnih mjerenja. Predstavljanje ovog signala u Fourierovoj bazi, u drugu ruku, zahtijeva samo dva nenulta vektora pravilno skalirana na pozitivnim i negativnim frekvencijama (slika 2.6.(b)). U cilju smanjenja broja potrebnih koeficijenata može se primijeniti diskretna kosinusna transformacija (u daljem tekstu: DCT) ovog signala, što zahtijeva samo jedan nenultni koeficijent u razvoju (slika 2.6.(c)). Jednačina DCT-a glasi $X_k = \sum_{n=0}^{N-1} x_n \cos(\frac{\pi}{N}(n+\frac{1}{2})k)$ gdje su $k = 0,\ldots,N-1$, x_n ulazni signal, a N broj uzoraka diskretne transformacije [8].



Slika 2.6. Modulirani kosinusni signal predstavljen na tri različita načina: (a) Povorka impulsa (b) Fourierova baza (c) DCT baza (Slika preuzeta iz [8])

2.4.2. 2-D modeli signala

Koncept iz poglavlja 2.4.1. se može proširiti i na 2-D signale. Npr, binarna slika noćnog neba je sparse u standardnom prostornom domenu jer su većina piksela zapravo crni pikseli nulte vrijednosti [8]. Slično ovome primjeru, slike prirode takođe karakterišu velike glatke površine ili površine sa teksturom i relativno malo oštrih rubova. Za signale sa ovakvom strukturom se zna da su jako blizu sparse signala kada se predstave pomoću multiscale wavelet transformacije. Wavelet transformacija se sastoji od rekurzivne podjele slike na niskofrekventne i visokofrekventne komponente. Komponente na najnižim frekvencijama daju grubu skaliranu aproksimaciju slike, dok komponente na višim frekvencijama dodaju detalje i razlučuju ivice. Ono što se primjećuje kada se računaju

koeficijenti *wavelet transformacije* tipične slike prirode, kao što je prikazano na slici 2.7., jeste da je većina koeficijenata jako male vrijednosti. Dakle, postoji mogućnost da se dobije dobra aproksimacija signala postavljajući male koeficijente na nulu, odnosno vršeći odbacivanje koeficijenata koji su manji od predefinirane vrijednosti (u daljem tekstu eng. t*resholding*) koeficijenata da bi se dobila k-sparse predstava signala. Prilikom mjerenja greške aproksimacije koristeći ℓ_p normu, ova procedura daje najbolju aproksimaciju originalnog signala u k-smislu, tj. najbolju aproksimaciju signala korištenjem samo k elemenata baze [10].

(a) (b)

Slika 2.7. Sparse predstavljanje slike pomoću multiscale wavelet transformacije. (a) Originalna slika (b) Wavelet prikaz. Veliki koeficijenti su predstavljeni svijetlim pikselima, dok su mali koeficijenti predstavljeni tamnim pikselima. Primijećuje se da većina wavelet koeficijenata ima vrijednost blisku nuli. (Slika preuzeta iz [10])

Slika 2.8. pokazuje primjer jedne takve slike i njene najbolje aproksimacije u k-smislu. Ovo predstavlja suštinu nelinearne aproksimacije koja je nelinearna upravo zato što izbor koeficijenata koji se zadržavaju u aproksimaciji zavisi od samog signala. Slično, s obzirom da slike prirode imaju približno sparse wavelet transformacije, ova ista thresholding operacija služi kao efektivan metod za odbacivanje određenih uobičajenih vrsta šuma, koje obično nemaju sparse *wavelet transformaciju*. Thresholding daje najbolju k-term aproksimaciju signala u odnosu na ortonormiranu bazu. Kada se koriste redundantni okviri, mora se osloniti na algoritme za sparse aproksimaciju poput onih opisanih u poglavlju 5 [10].

(a) (b)

Slika 2.8. Sparse aproksimacija slike prirode. (a) Originalna slika. (b) Aproksimacija slike dobijena zadržavanjem samo najvećih 10% wavelet koeficijenata. [10]

2.5. Signali sa mogućnošću kompresije

2.5.1. Mogućnost kompresije i aproksimacija sa K članova

Bitna pretpostavka koja se koristi u kontekstu CS-a jeste da signali pokazuju određen stepen strukturalnosti. Do sada je sparsity jedina struktura koja je bila razmotrena, tj. broj nenultih koeficijenata koje signal posjeduje kada se predstavi u ortonormiranoj bazi Ψ. Za signal se smatra da je sparse ako ima jako malo nenultih vrijednosti u poređenju sa njegovom ukupnom dužinom.

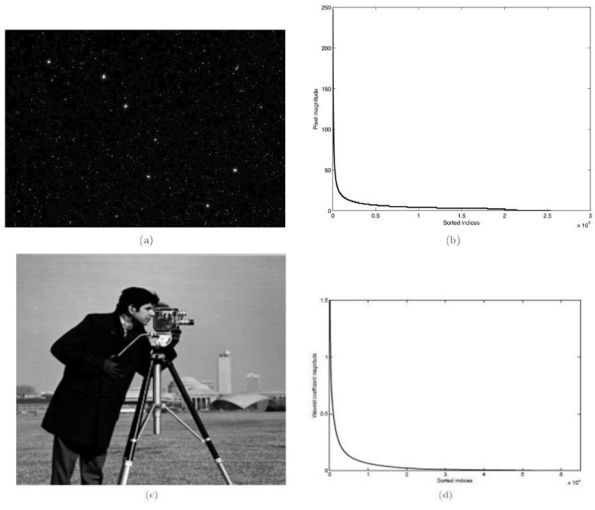
Malo struktuiranih signala je istinski sparse, prije su to signali sa mogućnošću kompresije. Za signal se kaže sa ima mogućnost kompresije ako amplitude njegovih sortiranih koeficijenata u Ψ jako brzo opadaju. Da bi se ovo razmotrilo matematički, neka je x signal koji ima mogućnost kompresije u bazi Ψ:

$$x = \Psi\alpha, \tag{2.15.}$$

gdje su α koeficijenti signala x u bazi Ψ. Ako x ima mogućnost kompresije, onda amplitude sortiranih koeficijenata α_s opadaju po stepenom zakonu:

$$|\alpha_s| \leq C_1 s^{-q}\,, s = 1, 2, \ldots \tag{2.16.}$$

Signal sa mogućnošću kompresije je onaj koji ispunjava zakon (2.16.) stepenog pada. Što je q veće, to amplitude brže opadaju, pa signal ima veću mogućnost kompresije. Slika 2.9. prikazuje slike koje imaju mogućnost kompresije u različitim bazama [8].



Slika 2.9. Slika u lijevom gornjem uglu je signal koji ima mogućnost kompresije u prostoru. Kada se vrijednosti piksela sortiraju od najveće do najmanje, postoji oštar pad. Slika u donjem lijevom uglu nema mogućnost kompresije u prostoru, ali ima mogućnost kompresije u wavelet domenu, s obzirom da njeni wavelet koeficijenti pokazuju osobinu stepenog pada. (Slika preuzeta iz [8])

S obzirom da koeficijenti signala sa mogućnošću kompresije opadaju tako brzo, to se ovakvi signali mogu predstaviti pomoću $K \ll N$ koeficijenata. Najbolja aproksimacija signala u smislu K je ona pri kojoj je K najvećih koeficijenata zadržano, pri čemu su svi ostali jednaki nuli. Razlika između originalnog signala i njegove aproksimacije u smislu K se označava greškom aproksimacije u K-smislu, kao $\sigma_{\mathrm{K}}(\mathrm{x})$ koja je definirana relacijom:

$$\sigma_{\mathrm{K}}(\mathrm{x}) = \arg \min_{\alpha \in \sum_K} \|x - \Psi\alpha\|_2. \tag{2.17.}$$

Za signale sa mogućnošću kompresije, možemo ustanoviti granicu kao zakon stepenog opadanja:

$$\sigma_{\mathrm{K}}(\mathrm{x}) \leq C_2 K^{\frac{1}{2}-s}. \tag{2.18.}$$

Ustvari, može se pokazati da će $\sigma_{\mathrm{K}}(\mathrm{x})_2$ opadati kao $\mathrm{K}^{-\mathrm{r}}$ ako i samo ako sortirani koeficijenti α_{i} opadaju kao $\mathrm{i}^{-\mathrm{r}+\frac{1}{2}}$. Slika 2.10. pokazuje primjer slike i njene aproksimacije u K-smislu [8].

(a) (b)

Slika 2.10. Sparse aproksimacija prirodne slike. (a) Originalna slika (b) Aproksimacija slike dobijena zadržavanjem samo najvećih 10% wavelet koeficijenata. Upravo zato što su slike prirode signali sa mogućnošću kompresije u wavelet domenu, to je aproksimacija ove slike u smislu njenih najvećih wavelet koeficijenata jako dobra aproksimacija. (Slika preuzeta iz [8])

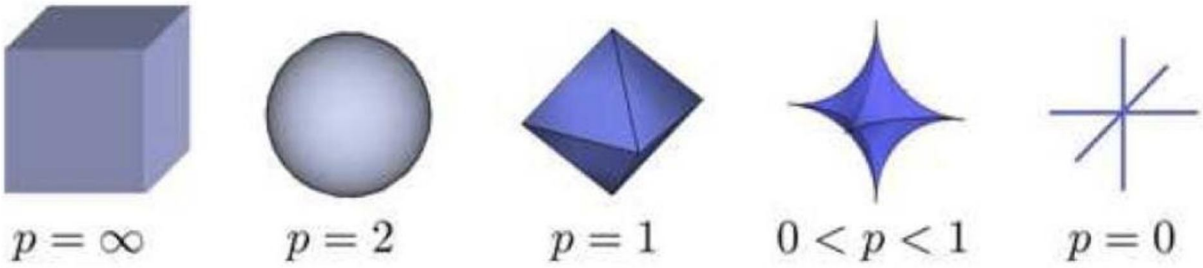
2.5.2. Mogućnost kompresije i ℓ_p prostori

Mogućnost kompresije nekog signala je povezana sa ℓ_p prostorom kojem signal pripada. Beskonačna sekvenca $x[n]$ je element ℓ_p prostora za određenu vrijednost p ako i samo ako je njegova ℓ_p norma konačna, tj. ako je zadovoljena nejednakost:

$$\|x\|_p = \left(\sum_i |x_i|^p\right)^{\frac{1}{p}} < \infty. \tag{2.19.}$$

Što je p manje, to vrijednosti sekvence moraju brže opadati da bi ona konvergirala, tako da je njena norma ograničena. U graničnom slučaju kada je $p = 0$, „norma“ je zapravo pseudonorma i ona ima određen broj nenultih vrijednosti. Kako p opada, to veličina

odgovarajućeg ℓ_p prostora takođe opada. Slika 2.11. prikazuje različite ℓ_p jedinične sfere (sve sekvence čija je ℓ_p norma jednaka 1) u tri dimenzije.



Slika 2.11. Kako se vrijednost p smanjuje, veličina odgovarajućeg ℓ_p prostora takođe opada. Ovo se može slikovito vidjeti kada se upoređuju veličine prostora signala, u tri dimenzije, za koje je ℓ_p norma manja ili jednaka jedan. Zapremina ℓ_p sfera se smanjuje sa p.

Ako se pretpostavi da je signal uzorkovan beskonačno dobro, i ako se označi sa $x[n]$, onda vrijedi: Da bi $x[n]$ sekvenca imala ograničenu ℓ_p normu, njeni koeficijenti moraju imati tendenciju pada po stepenom zakonu sa $q > \frac{1}{p}$. Odatle, signal koji se nalazi u ℓ_p prostoru za koji vrijedi da je $p \leq 1$ zadovoljava uslov pada po stepenom zakonu, pa prema tome ima sposobnost kompresije [8].

3. Matrice mjerenja

3.1. Kreiranje matrica mjerenja

Postoje dva glavna teoretska pitanja koja se nameću u CS-u. Prvo, kako kreirati matricu mjerenja Φ tako da ona očuva informaciju sadržanu u signalu x? Drugo, kako se originalni signal x može rekonstruisati na osnovu mjerenja y? U slučaju kada su podaci dobiveni mjerenjem sparse ili imaju mogućnost kompresije, to je moguće formirati matrice Φ takve da vrijedi $M \ll N$, koje pritom osiguravaju mogućnost precizne i efikasne rekonstrukcije originalnog signala koristeći mnoge praktične algoritme.

Ovo poglavlje se odnosi na dio CS-a vezan za kreiranje matrice mjerenja Φ. Postupci kreiranja ovakvih matrica neće biti odmah predstavljeni, nego će se prvo analizirati nekoliko poželjnih svojstava koje bi matrica Φ trebala imati, nakon čega će biti navedeno nekoliko primjera konstrukcije takve matrice.

Da bi se dobilo na konkretnosti, u ovom poglavlju će se razmatrati samo standardni modeli CS-a konačne dimenzionalnosti. Preciznije, ako za signal x vrijedi $x \in \mathbb{R}^N$, to će se analizirati sistemi mjerenja koji vrše M linearnih mjerenja. Ovaj proces se može opisati matematički relacijom:

$$y = \Phi x, \tag{3.1.}$$

gdje je Φ matrica dimenzija MxN, a $y \in \mathbb{R}^M$. Matrica Φ predstavlja redukciju dimenzionalnosti, tj. ona mapira prostor $\mathbb{R}^N$, gdje je N obično veliko, u $\mathbb{R}^M$, pri čemu je M obično mnogo manje od N [8].

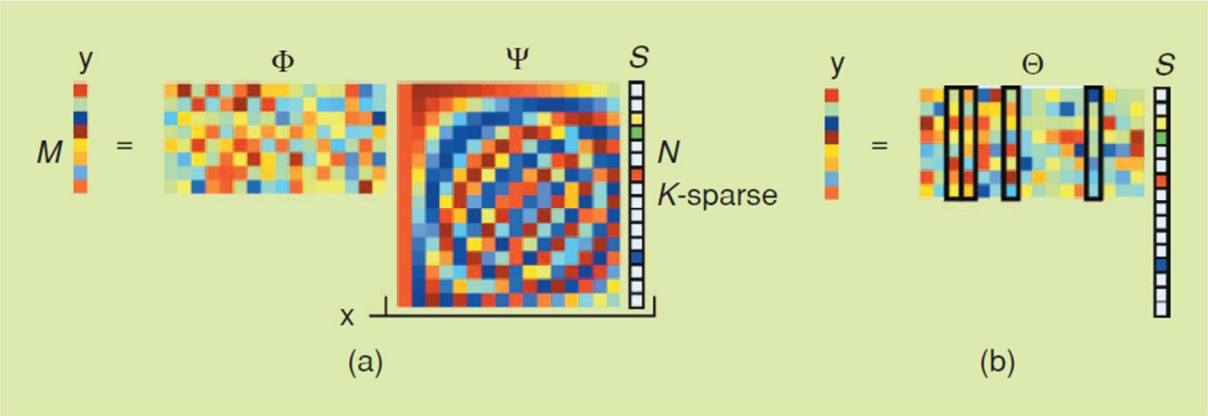
Može se primijetiti da se u okviru standardnog CS-a pretpostavlja da su mjerenja neadaptivna, što znači da je broj redova matrice Φ unaprijed fiksiran i ne zavisi od prijašnjih mjerenja. U određenim okruženjima, adaptivna mjerenja mogu dovesti do značajnih poboljšanja performansi.

Iako standardni CS pretpostavlja da je x vektor konačne dužine sa indeksima koji su diskretne vrijednosti (kao npr. vrijeme ili prostor), u praksi je često interesantno dizajnirati sisteme mjerenja namijenjene za kontinualno indeksirane signale kao što su kontinualni signali čija je nezavisna promjenljiva npr. vrijeme. Zasad će se x jednostavno analizirati kao vektor konačne dužine, koji sadrži uzorke dobivene uzorkovanjem kontinualnog signala x Nyquistovom učestanošću, te će se privremeno ignorisati problem načina dobijanja mjerenja koja imaju mogućnost kompresije, bez da se prvo vrši uzorkovanje Nyquistovom učestanošću [8].

Ukoliko se relacija (2.15.) uvrsti u relaciju (3.1.), dobija se:

$$y = \Phi x = \Phi\Psi\alpha = \theta\alpha, \tag{3.2.}$$

gdje je $\theta = \Phi\Psi$ matrica dimenzija MxN, Ψ baza u kojoj je signal x sparse ili ima mogućnost kompresije, a α koeficijenti signala x u bazi Ψ [9]. Dakle, vrši se korelacija signala koji se uzorkuje sa valnim oblicima $\varphi_k(t), \mathrm{k} = 1, 2, \ldots, \mathrm{M}$, odnosno članovima matrice Φ (tačnije njenim kolonama). Ovo predstavlja standardnu postavku predstavljanja vektora mjerenja preko matrice mjerenja Φ i signala x čije se vrijednosti i mjere, tj. uzorkuju [11]. Grafički prikaz relacije (3.2.) prikazan je na slici 3.1.



Slika 3.1. (a) Mjerenje u CS-u pomoću slučajne Gaussianske matrice mjerenja Φ i matrice Ψ koeficijenata DCT-a. Vektor koeficijenata s je sparse sa $K = 4$. (b) Proces mjerenja pomoću matrice $\theta = \Phi\Psi$. Četiri kolone ove matrice (uokvirene kolone) odgovaraju nenultim vrijednostima vektora s; vektor mjerenja y je linearna kombinacija ovih kolona (Slika preuzeta iz [9])

Ako su valni oblici mjerenja Diracove delta funkcije, onda je y vektor uzorkovanih vrijednosti signala x u vremenskom ili prostornom domenu. Kada su valni oblici mjerenja indikatorske funkcije piksela, onda je y skup podataka o slici koji se tipično dobijaju pomoću senzora u digitalnom fotoaparatu. Ukoliko su valni oblici mjerenja sinusoide, tada je y vektor Fourierovih koeficijenata. Ovo su samo neki od mnogih primjera valnih oblika mjerenja, čija priroda u suštini zavisi od prirode članova matrice mjerenja Φ i matrice baze Ψ [11].

Da bi rekonstrukcija originalnog signala x dimenzija $Nx1$ na osnovu M mjerenja (pri čemu je $M \ll N$) bila moguća, matrice Φ i Ψ moraju zadovoljavati određene uslove. Problem ovakve rekonstrukcije u opštem slučaju predstavlja jako komplikovan, praktično nemoguć problem, s obzirom da za $M \ll N$, relacija (3.2.) predstavlja sistem od M linearnih jednačina N nezavisno promjenljivih, pa postoji beskonačno mnogo rješenja $\hat{x}$ ovog sistema za koje vrijedi $y = \Phi\hat{x}$. No međutim, ovaj problem je moguće riješiti uz zadovoljenje sljedećih uslova za matrice Φ i Ψ.

Prvo, vektor x ne smije biti u nul-prostoru matrice Φ, tj. Φ mora biti takva matrica da ni za jedno x ne vrijedi da je $y = \Phi\mathrm{x} = 0$, što bi značilo da rekonstrukcija signala x na osnovu nul-vektora mjerenja y nema smisla, tj. nije moguća. Dakle, Φ mora zadovoljavati svojstvo nul-prostora. (Detaljnije o uslovima i svojstvu nul-prostora, te svojstvu spark pogledati Poglavlje 3.2.).

Svojstvo nul-prostora, je i potreban i dovoljan uslov za utvrđivanje uslova koji obezbjeđuju mogućnost rekonstrukcije K-sparse signala na osnovu M mjerenja, pri čemu je

$M \geq K$, ali ovi uslovi ne vrijede za šum. Kada mjerenja sadrže šum, ili su netačna usljed greške kvantizacije, korisno je uzeti u obzir nešto strožije uslove, pa se uvodi sljedeći uslov na izometriju matrica Φ koji predstavlja bitan dio teorije CS-a [8]. (Detalje o svojstvu ograničene izometrije, stabilnosti rekonstrukcije, te granicama mjerenja pogledati u Poglavlju 3.3.). Takođe, može se pokazati da postoji veza između svojstva nul-prostora i svojstva ograničene izometrije (Detalje o vezi između svojstva nul-prostora i svojstva ograničene izometrije pogledati u Poglavlju 3.4.).

Drugo, matrica mjerenja Φ i matrica valnih oblika baze Ψ ne smiju biti međusobno korelirane, tj. njihova korelacija mora biti što manja. Matematička osobina matrica koja kvantificira ovo svojstvo naziva se koherentnost matrica. Kao što postoji veza između svojstava ograničene izometrije i nul-prostora matrice, tako postoji i veza između koherentnosti matrica i svojstva spark, te svojstva ograničene izometrije matrice. (Detalje o koherentnosti i vezi ovog svojstva sa ostalim svojstvima matrice mjerenja pogledati u Poglavlju 3.5.).

Neka je Φ kanonska baza, odnosno baza Diracovih delta funkcija $\varphi_k(t) = \delta(\mathrm{t} - \mathrm{k})$, a Ψ je Fourierova baza redova $\Psi_{\mathrm{j}} = \frac{1}{\sqrt{N}} e^{i\frac{2\pi}{N}jt}$. Kako je Φ matrica mjerenja, to ovo odgovara klasičnoj shemi uzorkovanja u vremenskom ili prostornom domenu. Ovaj par vrijeme-frekvencija zadovoljava da je njihova koherentnost minimalna, što je i bio cilj. Takođe, Diracove delta funkcije i sinusoide su maksimalno nekoherentne (koherentnost ima minimalnu vrijednost) ne samo u jednoj dimenziji, nego i u proizvoljno-dimenzionalnom prostoru. Postoji još mnogo primjera međusobno nekoherentnih matrica Φ i Ψ, ali jedan od najvažnijih, ako ne i najvažniji primjer jesu slučajne matrice mjerenja Φ za koje vrijedi da su skoro nekoherentne sa bilo kojom fiksnom bazom Ψ. Nadalje, ukoliko su elementi matrice Φ nezavisno identično raspodijeljeni, tj. ako su to vrijednosti Gaussove raspodjele vjerovatnoće, ili binarne vrijednosti ± 1, onda matrica Φ pokazuje jako nisku koherentnost sa bilo kojom bazom Ψ.

Ovdje se nameće prilično čudna implikacija – ako je mjerenje nekoherentnim sistemima dobro, onda efikasni mehanizmi rekonstrukcije trebaju pronaći korelacije sa slučajnim valnim oblicima, tj. bijelim šumom [11].

3.2. Uslovi za nul-prostor

Polazna tačka utvrđivanja uslova koji moraju biti postavljeni na matricu Φ u kontekstu kreiranja matrice mjerenja, je upravo razmatranje nul-prostora matrice Φ, koji je definiran relacijom:

$$\aleph(\Phi) = \{z: \Phi \mathrm{z} = 0\}. \tag{3.3.}$$

Ukoliko se žele rekonstruisati svi sparse signali x na osnovu mjerenja Φx, onda je očito da za bilo koji par različitih vektora $x, x' \in \sum_K = \{x: \|x\|_0 \leq K\}$, mora biti zadovoljeno $\Phi x \neq \Phi x'$, jer u suprotnom ne bi bilo moguće razlikovati x od x' samo na osnovu mjerenja y. Formalnije, iz $\Phi x = \Phi x'$ slijedi $\Phi(x - x') = 0$ gdje $x - x' \in \sum_{2K}$, pa je očito da Φ jedinstveno predstavlja sve $x \in \sum_K$ ako i samo ako $\aleph(\Phi)$ ne sadrži vektore u $\sum_{2K}$. Postoji više ekvivalentnih načina opisivanja ovog svojstva, a jedan od najčešćih je poznat kao spark [12].

3.2.1. Spark

Definicija 3.1. Spark neke zadate matrice Φ je najmanji broj kolona matrice Φ koje su linearno zavisne.

Ovakva definicija omogućava definiranje sljedeće teoreme.

Teorema 3.1. Za bilo koji vektor $y \in \mathbb{R}^M$ postoji najviše jedan signal $x \in \sum_K$ takav da je $y = \Phi x$, ako i samo ako vrijedi da je $spark(\Phi) > 2K$.

Lako se uočava da $spark(\Phi) \in [2, M + 1]$. Prema tome, Teorema 3.1. daje donju granicu za M, tj. mora vrijediti $M \geq 2K$ [12].

3.2.2. Svojstvo nul-prostora

Da bi se ilustrirao dublji smisao NSP-a u kontekstu sparse rekonstrukcije, prvo se mora analizirati način određivanja efikasnosti i preciznosti algoritama za sparse rekonstrukciju kada se radi o generalnom signalu x koji nije sparse. Neka izraz Δ: $\mathbb{R}^M \to \mathbb{R}^N$ predstavlja specifičnu metodu rekonstrukcije. Primarno se treba fokusirati na uslove koji obezbjeđuju zadovoljavanje za formu:

$$\|\Delta(\Phi \mathrm{x}) - \mathrm{x}\|_2 \leq C \frac{\sigma_K(x)_1}{\sqrt{K}} \tag{3.4.}$$

za svako x, pri čemu je $\sigma_K(x)_1$ u opštem slučaju definirano relacijom:

$$\sigma_K(x)_P = \min_{\hat{x} \in \sum_K} \|x - \hat{x}\|_p. \tag{3.5.}$$

Ovo osigurava istovjetnu rekonstrukciju svih mogućih K-sparse signala, ali isto tako osigurava nivo robusnosti za signale koji nisu sparse, a koja direktno zavisi od toga koliko dobro su signali aproksimirani K-sparse vektorima. Takvi uslovi koji obezbjeđuju zadovoljavanje se zovu optimalnim za instancu (eng. *instance-optimal*) upravo zato što oni osiguravaju optimalne performanse za svaku instancu signala x [13]. Ovo ih odvaja od uslova koji obezbjeđuju zadovoljavanje koji vrijede samo na određenom podskupu mogućih signala, kao što su sparse signali ili signali sa mogućnošću kompresije – kvalitet uslova koji obezbjeđuju zadovoljavanje se prilagođava određenom izboru x-a. Takođe im se često daje naziv uniformnih uslova koji obezbjeđuju zadovoljavanje pošto vrijede uniformno za svaki signal x.

Adaptacija teoreme iz [13] pokazuje da, ako postoji ikakav algoritam rekonstrukcije koji zadovoljava (3.4.), onda Φ mora nužno zadovoljavati NSP reda $2K$.

Teorema 3.2. Neka izraz $\Phi: \mathbb{R}^N \to \mathbb{R}^M$ označava matricu mjerenja, a izraz $\Delta: \mathbb{R}^M \to \mathbb{R}^N$ proizvoljan algoritam rekonstrukcije. Ako uređeni par (Φ, Δ) zadovoljava relaciju (3.4.), onda Φ zadovoljava NSP reda $2K$ [13].

3.3. Svojstvo ograničene izometrije

Definicija 3.2. Matrica Φ zadovoljava **svojstvo ograničene izometrije** (eng. *restricted isometry property*, u daljem tekstu RIP) reda K ako postoji $\delta_K \in (0,1)$ takvo da je relacija

$$(1 - \delta_K)\|x\|_2^2 \leq \|\Phi x\|_2^2 \leq (1 + \delta_K)\|x\|_2^2 \tag{3.6.}$$

zadovoljena za svako $x \in \sum_K = \{x: \|x\|_0 \leq K\}$.

Ako matrica Φ zadovoljava RIP reda $2K$, onda se relacija (3.6.) može interpretirati ovako: Matrica Φ približno čuva udaljenost između bilo koja dva para K-sparse vektora, što znači da K-sparse vektori ne mogu biti u nul-prostoru matrice Φ. Ekvivalentna interpretacija RIP-a jeste da su svi podskupovi od K kolona matrice Φ približno ortogonalni. Kolone matrice Φ ne mogu biti striktno ortogonalne pošto matrica Φ ima više kolona nego redova. Ovo ima fundamentalne implikacije kada je riječ o robusnosti na šum [14].

3.3.1. Svojstvo ograničene izometrije i stabilnost

U Poglavlju 5 ovog rada će se pokazati da, ako matrica Φ zadovoljava RIP, onda je to dovoljno za mnoštvo algoritama da mogu uspješno rekonstruisati sparse signal iz mjerenja koja imaju superponiran šum. U ovom poglavlju će se prvo pobliže ispitati opravdanost korištenja RIP-a. Jasno je da je donja granica u RIP-u potreban uslov ukoliko se želi biti u mogućnosti rekonstruisati sve sparse signale na osnovu mjerenja Φx, iz istog razloga iz kojeg je NSP potreban. Može se čak i više reći o neophodnosti RIP-a ako se razmotri sljedeća definicija stabilnosti.

Definicija 3.3. Neka izraz $\Phi: \mathbb{R}^N \to \mathbb{R}^M$ označava matricu mjerenja, a izraz $\Delta: \mathbb{R}^M \to \mathbb{R}^N$ proizvoljan algoritam rekonstrukcije. Za uređeni par (Φ, Δ) se kaže da je C-stabilan ako za bilo koje $x \in \sum_K$ i bilo koje $e \in \mathbb{R}^M$ vrijedi relacija:

$$\|\Delta(\Phi \mathrm{x} + \mathrm{e}) - \mathrm{x}\|_2 \leq C\|\mathrm{e}\| \tag{3.7.}$$

Ova definicija kaže da, ako se doda mala količina šuma na mjerenja, onda njegov utjecaj na rekonstruisani signal ne bi trebao biti velik.

3.3.2. Granice mjerenja

Broj mjerenja potrebnih da bi se postigao RIP se također može uzeti u obzir. Ako se zanemari utjecaj konstante δ i ako se dimenzije problema (N, M i K) stave u centar pažnje, moguće je postići jednostavnu donju granicu.

Teorema 3.3. Neka je Φ matrica dimenzija MxN koja zadovoljava RIP reda $2K$ sa konstantom $\delta \in (0, \frac{1}{2}]$. Tada vrijedi:

$$M \geq CKln\left(\frac{N}{K}\right) \tag{3.8.}$$

gdje je $C = \frac{1}{2\ln(\sqrt{24}+1)} \approx 0.28$.

Restrikcija $\delta \leq \frac{1}{2}$ je proizvoljna i napravljena je zbog pogodnosti – manje modifikacije utvrđuju granicu kao $\delta \leq \delta_{max}$ za bilo koje $\delta_{max} < 1$. Također, iako nije uložen nikakav napor da bi se konstante optimizirale, bitno je primijetiti da su one i ovako prilično razumne. Teorema 3.3. ne daju preciznu zavisnost M za željenu RIP konstantu δ [10].

3.4. Veza između svojstva ograničene izometrije i svojstva nul-prostora

U ovom paragrafu će biti pokazano da ako matrica zadovoljava RIP, onda ona zadovoljava i NSP. Dakle, RIP je strikno jače svojstvo od NSP-a.

Teorema 3.4. [10] Neka matrica Φ zadovoljava RIP reda $2K$ sa $\delta_{2K} < \sqrt{2} - 1$. Tada Φ zadovoljava i NSP reda $2K$ sa konstantom:

$$C = \frac{\sqrt{2}\delta_{2K}}{1 - \left(1 + \sqrt{2}\right)\delta_{2K}}. \tag{3.9.}$$

3.5. Koherentnost

Dok spark, NSP i RIP pružaju garante rekonstrukcije sparse signala, provjera da li neka generalna matrica Φ zadovoljava ijedan od ovih uslova ima kombinatoričku računsku kompleksnost, pošto se u svakom slučaju mora uzeti u obzir $\binom{N}{K}$ submatrica. U mnogim slučajevima je bolje koristiti svojstva matrice Φ koja se mogu lako izračunati, da bi se pružili konkretniji garanti rekonstrukcije. Koherentnost matrice je jedno takvo svojstvo.

Definicija 3.4. Koherentnost matrice Φ, $\mu(\Phi)$ je najveći skalarni (unutrašnji) proizvod između bilo koje dvije kolone ϕ_i, ϕ_j matrice Φ:

$$\mu(\Phi) = \max_{1\le i<j\le N} \frac{|<\phi_i,\phi_j>|}{\|\phi_i\|_2\|\phi_j\|_2}. \tag{3.13.}$$

Moguće je pokazati da je koherentnost matrice uvijek u opsegu $\mu(\Phi) \in \left[\sqrt{\frac{N-M}{M(N-1)}}, 1\right]$; donja granica je poznata pod imenom Welchova granica. Očito je da donja granica, kada je $N \gg M$, približno iznosi $\mu(\Phi) \ge \frac{1}{\sqrt{M}}$ [10].

Definicija 3.5. Neka je dat uređeni par (Φ, Ψ) ortogonalnih baza u $\mathbb{R}^N$. Φ se koristi za za uzorkovanje signala x, a Ψ se koristi za predstavljanje signala x u određenoj bazi. Restrikcija na uređene parove nije od esencijalne važnosti, ali se uvodi zbog pojednostavljenja analize. **Koherentnost između matrica Φ i Ψ** je određena relacijom:

$$\mu(\Phi,\Psi) = \sqrt{N}\max_{1\le k,<j\le N}\left|\langle\varphi_k,\psi_j\rangle\right|. \tag{3.14.}$$

Dakle, koherentnost mjeri najveću korelaciju između matrica Φ i Ψ. Ako Φ i Ψ sadrže međusobno korelirane članove, onda je koherentnost velika. U suprotnom, koherentnost je mala. Iz linearne algebre slijedi da se mjera koherentnosti nalazi u intervalu $\mu(\Phi,\Psi) \in \left[1,\sqrt{N}\right]$. CS se uglavnom bavi parovima matrica Φ i Ψ sa niskom koherentnosti [11].

Koherentnost se ponekad može dovesti u vezu sa sparkom, NSP-om i RIP-om i to pomoću sljedećih lema.

Lema 3.1. Za bilo koju matricu Φ vrijedi:

$$spark(\Phi) \ge 1 + \frac{1}{\mu(\Phi)}. \tag{3.15.}$$

Lema 3.2. Ako Φ ima kolone čija je norma jedinična i koherentnost $\mu = \mu(\Phi)$, onda Φ zadovoljava RIP reda K sa $\delta = (K-1)\mu$ za svako $K < \frac{1}{\mu}$ [10].

4. Rekonstrukcija prorijeđenih signalna ℓ_1 minimizacijom

4.1. Rekonstrukcija signala ℓ_1 minimizacijom

U nastavku ovog rada pokazat će se da postoji čitava lepeza pristupa rekonstrukciji sparse signala x na osnovu malog broja linearnih mjerenja. Počet će se sa „najprirodnijim“ pristupom problemu rekonstrukcije sparse signala.

Neka su data mjerenja $y = \Phi x$ i neka je poznato da je originalni signal x sparse ili pak ima mogućnost kompresije, onda je sasvim logično pokušati rekonstruisati x rješavajući optimizacioni problem oblika:

$$\hat{x} = arg\min_z \|z\|_0 \ , pri\ čemu\ z \in \mathcal{B}(y), \tag{4.1.}$$

gdje $\mathcal{B}(y)$ osigurava da je $\hat{x}$ konzistentno sa mjerenjima y. Poznato je da je $\|z\|_0 = |supp(z)|$, te da ova relacija jednostavno daje broj nenultih članova u z, pa odatle slijedi da relacija (4.1.) traži signal koji je najviše sparse, a koji je u isto vrijeme konzistentan sa posmatranim mjerenjima. Npr., ako su mjerenja tačna i bez šuma, onda se može postaviti $\mathcal{B}(y) = \{z: \Phi \mathrm{z} = \mathrm{y}\}$. Ako su u pitanju mjerenja sa malom količinom ograničenog superponiranog šuma, onda se može postaviti $\mathcal{B}(y) = \{z: \|\Phi \mathrm{z} - \mathrm{y}\| \leq \varepsilon\}$. U oba slučaja relacija (4.1.) pronalazi signal x koji je najviše sparse, a koji je konzistentan sa mjerenjima y.

Može se primijetiti da se u relaciji (4.1.) inherentno pretpostavlja da je x samo po sebi sparse. U nešto češćoj postavci u kojoj je $x = \Psi\alpha$, ovakav pristup se lako modificira, pa se sada razmatra optimizacioni problem oblika:

$$\hat{\alpha} = arg\min_z \|z\|_0 \ , pri\ čemu\ z \in \mathcal{B}(y), \tag{4.2.}$$

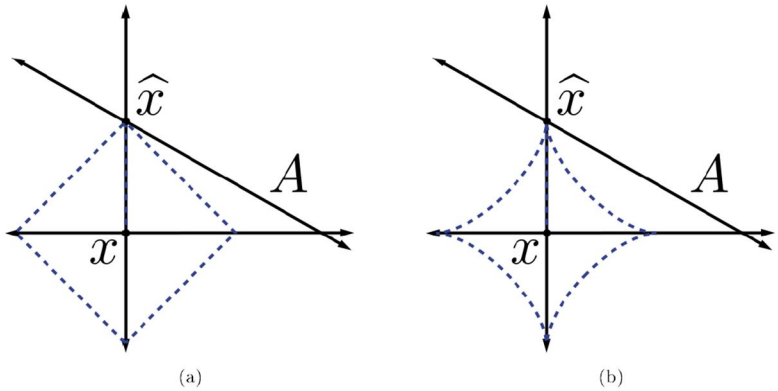
gdje je $\mathcal{B}(y) = \{z: \Phi\Psi \mathrm{z} = \mathrm{y}\}$ ili $\mathcal{B}(y) = \{z: \|\Phi\Psi \mathrm{z} - \mathrm{y}\| \leq \varepsilon\}$. Smjenom $\widetilde{\Phi} = \Phi\Psi$ vidi se da su relacije (4.1.) i (4.2.) u suštini identične. U mnogim slučajevima uvođenje Ψ ne komplikuje kreiranje matrica Φ takvih da $\widetilde{\Phi}$ posjeduje željena svojstva. Prema tome, u nastavku ovog rada pažnja će uglavnom biti posvećena samo za slučaj kada je $\Psi = \mathrm{I}$. Bitno je napomenuti međutim, da ovakva restrikcija postavlja određene granice u analizi kada je Ψ generalizovani rječnik, a ne ortonormirana baza. Npr., u slučaju kada je $\|\hat{x} - x\|_2 = \|\Psi\hat{\mathrm{c}} - \Psi \mathrm{c}\|_2 \neq \|\hat{\alpha} - \alpha\|_2$, pa se ograničenje na $\|\hat{c} - c\|_2$ ne može direktno prevesti u ograničenje na $\|\hat{x} - x\|_2$, što je često metrika koja je od interesa za analizu.

Iako je moguće analizirati ponašanje relacije (4.1.) za pogodne pretpostavke Φ, ova strategija nije od interesa s obzirom da ciljna funkcija $\|\cdot\|_0$ nije konveksna, pa je postupak (4.1.) potencijalno jako teško riješiti. Zapravo, može se pokazati da je za generaliziranu matricu Φ traženje rješenja koje aproksimira pravi minumum u suštini NP-težak. Jedan od

načina prevođenja ovog problema u nešto što se može lakše riješiti jeste zamjena $\|\cdot\|_0$ s njenom konveksnom aproksimacijom $\|\cdot\|_1$. Dakle, razmatra se optimizacioni problem:

$$\hat{x} = arg\min_z \|z\|_1 \ , pri\ čemu\ z \in \mathcal{B}(y), \tag{4.3.}$$

Ukoliko je $\mathcal{B}(y)$ konveksno, onda je problem (4.3.) računski izvodiv. Tačnije, kada je $\mathcal{B}(y) = \{z: \Phi\Psi z = y\}$, rezultirajući problem se može predstaviti kao problem linearnog programiranja [15].



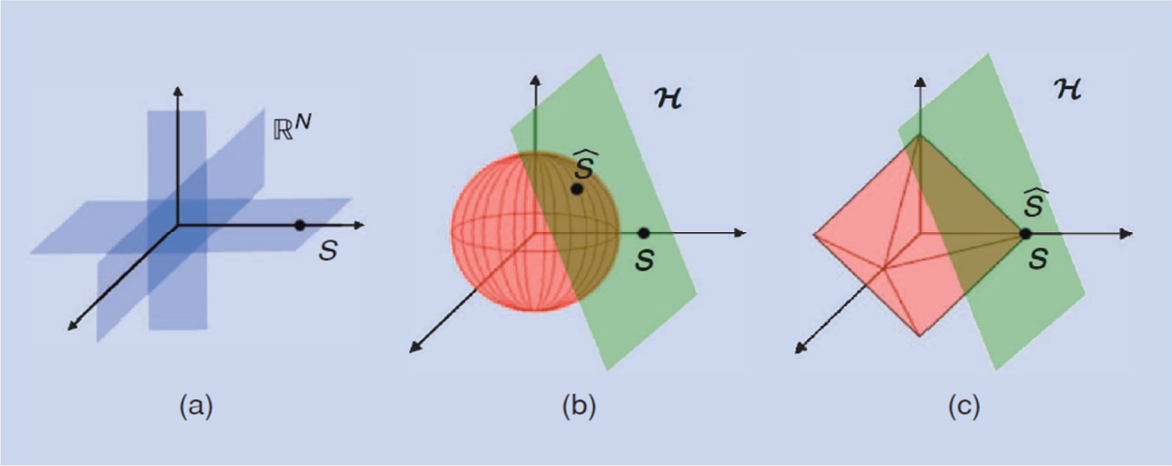
Slika 4.1. Najbolja aproksimacija tačke u $\mathbb{R}^2$ pomoću jednodimenzionalnog potprostora (pomoću prave) korištenjem ℓ_1 norme i ℓ_p norme za $p = \frac{1}{2}$. (a) Aproksimacija u ℓ_1 normi (b) Aproksimacija u ℓ_p kvazinormi (Slika preuzeta iz [8])

Očito je da zamjena (4.1.) sa (4.3.) pretvara problem koji je računski praktično nemoguće riješiti u rješiv problem, s tim da možda nije odmah očito da je rješenje optimizacionog problema (4.3.) imalo slično rješenju optimizacionog problema (4.1.). Međutim, postoje intuitivni razlozi za prihvatanje činjenice da će upotreba ℓ_1 minimizacije unaprijediti prorijeđenost. Kao primjer se može dati problem koji je predstavljen na slici 4.1. U ovom slučaju rješenje ℓ_1 minimizacijskog problema se u potpunosti podudara sa rješenjem ℓ_p minimizacijskog problema za bilo koje $p < 1$, te je uz to i očito sparse [8].

Prednost ovakve minimizacije jeste to što se kao njen rezultat dobije tačna rekonstrukcija signala x na osnovu skupa kondenzovanih podataka (vektora mjerenja y), i to postupkom minimizacije konveksnog funkcionala koji ne pretpostavlja nikakvu informaciju o broju nenultih vrijednosti x-a, njihovim indeksima, ili pak amplitudama. Za sve ove informacije se a priori pretpostavlja da su u potpunosti nepoznate. Jednostavno, alogoritam se pokrene i ako je dobijeni signal dovoljno sparse, dolazi do tačne rekonstrukcije.

Bitno je napomenuti da u ovom radu nije ni razmatrana ℓ_2 minimizacija, zato što ovakva minimizacija skoro nikada ne pronalazi K-sparse rješenje, nego upravo suprotno – daje rješenje koje ima mnogo nenultih vrijednosti. U suštini, ℓ_2 minimizacija kao rješenje daje vektor koji ima minimalnu energiju u transliranom nul-prostoru [11]. Grafički prikaz relacija koje opisuju rekonstrukciju signala ℓ_0, ℓ_1 i ℓ_2 minimizacijom u vektorskom prostoru više dimenzionalnosti, tj. u $\mathbb{R}^3$, i poređenje njihove tačnosti dat je na slici 4.2, dok slika 4.3.

prikazuje praktičan primjer sparse signala i njegove rekonstrukcije pomoću ℓ_1 i ℓ_2 minimizacije, uz poređenje istih.

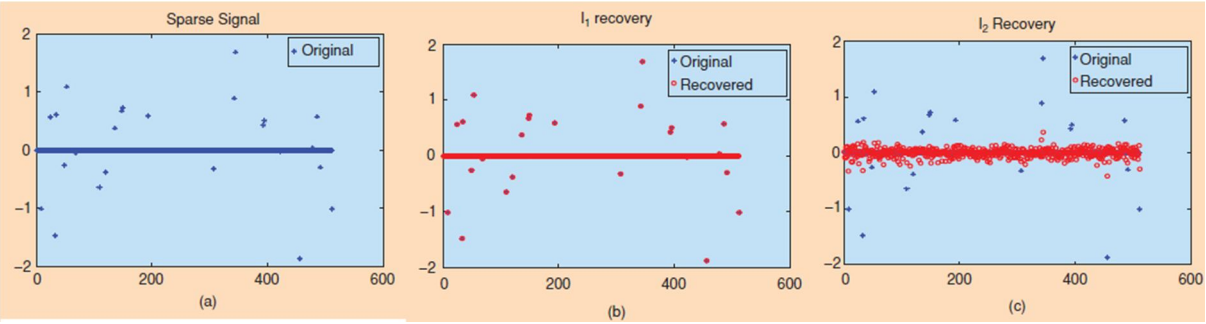


Slika 4.2. (a) Podprostori koji sadrže sparse vektore u $\mathbb{R}^3$ leže blizu koordinatnih osa. (b) Prikaz ℓ_2 minimizacije koja pronalazi presjek između tačke koja pronalazi tačku presjeka $\hat{s}$ (koja nije sparse) između ℓ_2 lopte (crvene hipersfere) i nul-prostora translirane matrice mjerenja (zelene ravni). (c) Prikaz rješenja ℓ_1 minimizacije koja pronalazi sparse tačku presjeka $\hat{s}$ sa visokom vjerovatnosti, zahvaljujući špicastoj geometriji ℓ_1 „lopte“. (Slika preuzeta iz [9])

Korištenje ℓ_1 minimizacije u svrhu unaprijeđenja ili iskorištenja sparsity-a ima dugu historiju, počev od rada Beurlinga na temu Ekstrapolacije Fourierove transformacije na osnovu djelomičnih opservacija.

1965. godine Logan je u nešto drugačijem kontekstu pokazao da se signal s ograničenim spektrom može savršeno rekonstruisati čak i sa prisustvom proizvoljnih grešaka na malom intervalu. Metode rekonstrukcije su se i ovdje sastojale od traženja signala sa ograničenim spektrom koji je najbliži posmatranom signalu i to u ℓ_1 normi. Ovo se može posmatrati kao dodatna validacija intuicije koja se stiče na osnovu Slike 4.1., tj. da je ℓ_1 norma pogodna za sparse greške [8].

Historijski, korištenje ℓ_1 minimizacije u svrhu rješavanja velikih problema je konačno postalo primijenjeno u praksi kada je došlo do naglog povećanja računarske moći tokom kasnih 70-ih i ranih 80-tih. U jednoj od prvih primjena, pokazano se da se geofizički signali, koji se sastoje od dosta „špiceva“ (koji ukazuju na značajne promjene između podzemnih slojeva zemljišta), mogu rekonstruisati samo na osnovu njihovih visokofrekventnih komponenti i to upravo korištenjem ℓ_1 minimizacije. Konačno, tokom 90-tih, u naučnoj zajednici je došlo do ponovnog interesa za ovakve pristupe, upravo u svrhu pronalaženja sparse aproksimacija signala i slika kada se oni predstave u redundantnim rječnicima ili unijama baza [15].



Slika 4.3. (a) Realan sparse signal. (b) Rekonstrukcija sparse signala na osnovu 60 kompleksnih Fourierovih koeficijenata pomoću ℓ_1 minimizacije. (c) Rekonstrukcija signala minimalne energije dobivena zamjenom ℓ_1 norme sa ℓ_2 normom; ℓ_1 i ℓ_2 daju jako različite rezultate. Rješenje ℓ_2 minimizacije očito ne daje dobru aproksimaciju originalnog signala. (Slika preuzeta iz [11])

Dakle, postoji niz razloga zbog kojih se može pretpostaviti da ℓ_1 minimizacija pruža precizan metod za rekonstrukciju sparse signala. Ono što je još i važnije, jeste da ovakav metod omogućava računski izvediv pristup problemu rekonstrukcije sparse signala. U nastavku ovog poglavlja analizirat će se teorijske osnove ℓ_1 minimizacije signala sa i bez superponiranog šuma [8].

4.2. Rekonstrukcija signala bez šuma

Analiza problema rekonstrukcije signala bez superponiranog šuma prilikom mjerenja počinje od relacije:

$$\hat{x} = arg\min_z \|z\|_1 \ , pri\ čemu\ z \in \mathcal{B}(y), \tag{4.4.}$$

za specifične odabire $\mathcal{B}(y)$. Da bi to bilo moguće, potrebna je sljedeća teorema.

Teorema 4.1. Neka Φ zadovoljava svojstvo ograničene izometrije (skraćeno RIP) reda $2K$ sa $\delta_{2K} < \sqrt{2} - 1$ i neka su data mjerenja u formi $y = \Phi x$. Ako je $\mathcal{B}(y) = \{z: \Phi\Psi z = y\}$, onda rješenje $\hat{x}$ za (4.4.) zadovoljava relaciju:

$$\|\hat{x} - x\|_2 \leq C_0 \frac{\sigma_K(x)_1}{\sqrt{K}} \tag{4.7.}$$

Teorema 4.1. je jako važna. Razmatrajući slučaj kada $x \in \sum_K = \{x: \|x\|_0 \leq K\}$ može se zaključiti da matrica Φ zadovoljava RIP, što dozvoljava da broj mjerenja bude reda $O(Klog(\frac{N}{K}))$, pa je moguća tačna rekonstrukcija K-sparse signala. Ova posljedica se čini nevjerovatnom, pa se može očekivati da je ovakva procedura izrazito osjetljiva na šum, no međutim, u nastavku će biti pokazano da je ovakav algoritam zaista stabilan [8]. Dakle, ukoliko je x K-sparse signal, onda Teorema 4.1. osigurava da je kvalitet rekonstruisanog signala isti kao kad bi se unaprijed znali indeksi K najvećih vrijednosti x-a i da su kao takvi direktno izmjereni. Drugim riječima, rekonstrukcija je približno dobra kao postupak

predviđanja u kojem se, uz potpuno poznavanje signala x, vrši ekstrakcija K najvažnijih podataka, tj. najvećih vrijednosti signala x [11].

Primjetno je da Teorema 4.1. pretpostavlja da Φ zadovoljava RIP. Ova pretpostavka se može vrlo lako zamijeniti pretpostavkom da Φ zadovoljava NSP. Tačnije, ako je od interesa samo rekonstrukcija signala bez šuma, u tom slučaju h leži u nul-prostoru matrice Φ. Dakle, u ovakvoj postavci problema rekonstrukcije je očito da ako Φ zadovoljava NSP, onda ona zadovoljava i istu granicu greške (kao i u slučaju kada zadovoljava RIP) [8].

4.3. Rekonstrukcija signala sa šumom

Mogućnost tačne rekonstrukcije sparse signala na osnovu mjerenja bez šuma predstavlja obećavajući rezultat u naučnoj oblasti analize signala i sistema. Međutim, u većini realnih sistema mjerenja sadržavaju neki oblik šuma. Npr., da bi se podaci u računaru obradili, prvo ih je potrebno predstaviti koristeći konačan broj bita, pa zbog toga mjerenja tipično sadržavaju grešku kvantizacije. Takođe, sistemi koji su implementirani na bazi fizičkog hardware-a podliježu pojavi niza različitih tipova šuma, zavisno od okoline.

Koliko god iznenađujuće zvučalo, moguće je modificirati optimizacijski problem:

$$\hat{x} = arg\min_{z}\|z\|_1 \ , pri\ čemu\ z \in \mathcal{B}(y)\ i\ \|\Phi\hat{x} - y\|_2 \leq \varepsilon \tag{4.8.}$$

tako da se omogući rekonstrukcija sparse signala uz prisustvo mnogih uobičajenih modela šuma. Kao što se i može očekivati, svojstvo ograničene izometrije (RIP) je jako korisno za uspostavljanje uslova koji osiguravaju zadovoljavanje performansi kod šuma [8].

4.3.1. Ograničen šum (Bounded noise)

U ovom poglavlju će se prvo definirati granica za performanse uniformno ograničenog šuma u najgorem slučaju.

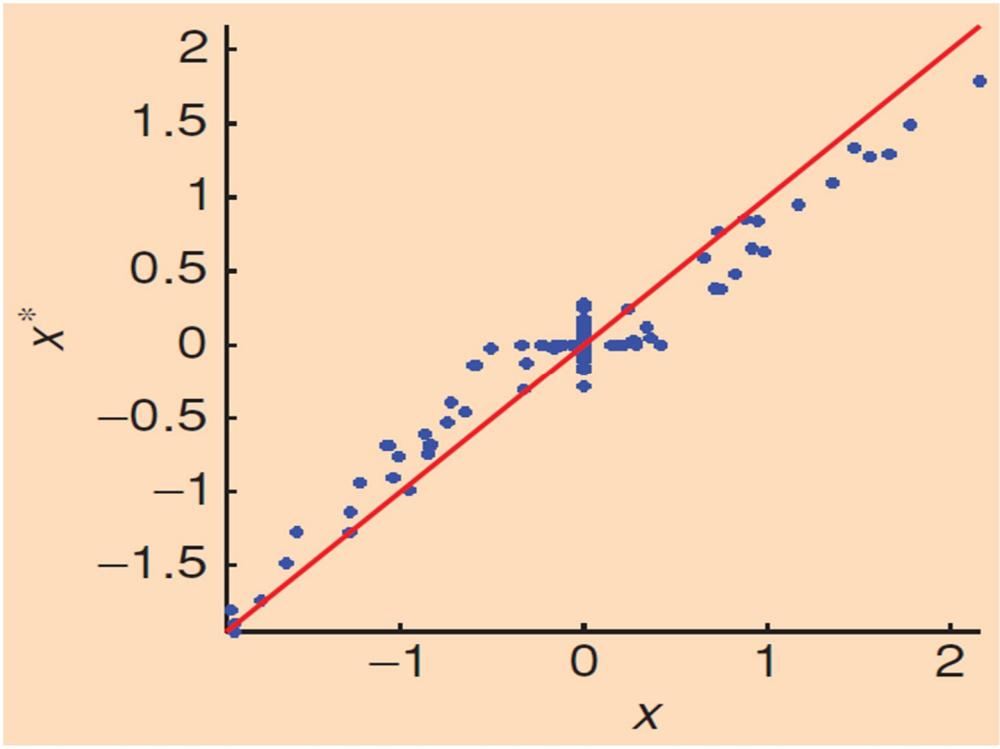
Teorema 4.2. Neka Φ zadovoljava svojstvo ograničene izometrije (skraćeno RIP) reda $2K$ sa $\delta_{2K} < \sqrt{2} - 1$ i neka su data mjerenja u obliku $y = \Phi\mathrm{x} + \mathrm{e}$, gdje je $\|e\|_2 \leq \varepsilon$ (ε ograničava količinu šuma). Kada je $\mathcal{B}(y) = \{z\colon \|\Phi\mathrm{z} - \mathrm{y}\| \leq \varepsilon\}$, onda rješenje $\hat{x}$ optimizacijskog problema (4.8.) zadovoljava uslov:

$$\|\hat{x} - x\|_2 \leq C_0 \frac{\sigma_K(x)_1}{\sqrt{K}} + C_2\varepsilon, \tag{4.9.}$$

gdje su

$$C_0 = 2\frac{1 - \left(1 - \sqrt{2}\right)\delta_{2K}}{1 - \left(1 + \sqrt{2}\right)\delta_{2K}}, C_2 = 4\ \frac{\sqrt{1 + \delta_{2K}}}{1 - \left(1 + \sqrt{2}\right)\delta_{2K}}. \tag{4.10.}$$

Iz relacije (4.9.) je očito da je greška pri rekonstrukciji ograničena sumom dva izraza. Prvi član ove sume predstavlja grešku do koje bi došlo da se radi o rekonstrukciji signala bez šuma, dok je drugi član proporcionalan sa razinom šuma. Nadalje, konstante C_0 i C_2 su obično male. Npr., za $\delta_{2K} = \frac{1}{4}$, vrijedi $C_0 \leq 5.5$ i $C_2 \leq 6$. Slika 4.4. prikazuje primjer rekonstrukcije signala sa šumom [11].



Slika 4.4. Signal x (horizontalna osa) i njegova rekonstrukcija x^* (vertikalna osa) dobijena pomoću teoreme 4.2.. U ovom primjeru $N = 512$ i $M = 256$, pri čemu je signal 64-sparse (tj. $K = 64$). Kao matrica mjerenja korištena je slučajna matrica, a smetnju predstavlja Gaussov bijeli šum. Za ovaj primjer vrijedi da greška prilikom reonstrukcije iznosi $\|\hat{x} - x\|_2 \approx 1.3\varepsilon$. (Slika preuzeta iz [11])

5. Algoritmi za rekonstrukciju sparse signala

5.1. Algoritmi za rekonstrukciju sparse signala

Glavni problema CS-a jeste rekonstrukcija sparse signala x na osnovu mjerenja $y = \Phi x + e$ koja sadrže šum. Značajni napori su uloženi u cilju razvoja algoritama koji se izvršavaju brzo, te daju tačnu i stabilnu rekonstrukciju signala x na osnovu mjerenja y. Kao što je već rečeno u prethodnim poglavljima, „dobra“ CS matrica Φ tipično zadovoljava određene geometrijske uslove, kao što je npr. svojstvo ograničene izometrije (RIP). Praktični algoritmi iskorištavaju ovu činjenicu na razne načine u cilju smanjenja broja mjerenja, omogućavanja brže rekonstrukcije, te osiguravanja robusnosti i na numeričke i na stohastičke greške. Dizajn algoritama za rekonstrukciju sparse signala je vođen određenim kriterijima. Neki od najbitnijih su pobrojani u nastavku.

- **Minimalan broj mjerenja.** Za algoritme rekonstrukcije sparse signala mora biti potrebno približno isti broj mjerenja koji je potreban za stabilno zapisivanje K-sparse signala.
- **Robusnost na šum prilikom mjerenja i nepodudaranje modela.** Algoritmi za rekonstrukciju sparse signala moraju biti stabilni s obzirom na perturbacije ulaznog signala, kao i dodati šum prilikom mjerenja. Oba navedena tipa greške prirodno proističu iz praktičnih sistema.
- **Brzina.** Algoritmi za rekonstrukciju sparse signala moraju stremiti ka korištenju minimalnih računarskih resursa, iako se u mnogim primjenama CS-a javljaju visokodimenzionalni signali.
- **Uslovi koji osiguravaju zadovoljenje performansi.** U prethodnim poglavljima su navedeni razni uslovi koji osiguravaju zadovoljenje performansi rekonstrukcije signala korišenjem ℓ_1 minimizacije. Prilikom evaluacije drugih algoritama, vršit će se ista razmatranja. Npr., može se odabrati dizajniranje algoritama koji posjeduju uslove koji osiguravaju zadovoljenje optimalnosti za instancu ili zadovoljenje neke vjerovatnoće. Takođe, može se odabrati fokus na performanse algoritama za rekonstrukciju tačno K-sparse signala x, ili se mogu razmatrati performanse algoritama rekonstrukcije generaliziranih signala. Mogu se razmatrati i algoritmi koji su popraćeni uslovima koji osiguravaju zadovoljenje performansi bilo u slučaju rekonstrukcije signala sa ili bez šuma.

U literaturi postoje mnogi algoritmi koji zadovoljavaju neke (ili čak sve) gore navedene kriterije. Kako je nemoguće opisati sve te algoritme rekonstrukcije, u ovom poglavlju će biti analizirani samo neki. Generalno govoreći, metode rekonstrukcije se dijele na tri kategorije:

a) metode koje su bazirane na konveksnoj optimizaciji,
b) *greedy* metode,

c) kombinatoričke metode.

Ostatak ovog poglavlja analizira nekoliko primjera algoritama rekonstrukcije sparse signala i njihovih svojstava [8].

5.2. Metode bazirane na konveksnoj optimizaciji

Bitna klasa algoritama za rekonstrukciju spada u područje primjene konveksne optimizacije. Ovi algoritmi optimiziraju konveksnu funkciju $f(\cdot)$ nepoznate promjenljive x pomoću (možda čak i neograničenog) podskupa od $\mathbb{R}^N$. Detaljniji opis i matematička podloga ovakvih algoritama data je u nastavku iz razloga što su ovi algoritmi korišteni za rekonstrukciju signala sa mogućnošću kompresije u praktičnom dijelu ovog rada (Detaljnije u Poglavlju 7).

5.2.1. Postavka algoritma

Neka je $J(x)$ konveksna funkcija cilja koja daje prednost *sparsity*-u, tj. $J(x)$ je malo kada je x sparse. Da bi se rekonstruisao signal $\hat{x}$ na osnovu mjerenja $y = \Phi x, \Phi \in \mathbb{R}^{\mathrm{MxN}}$, treba riješiti optimizacijski problem:

$$\min_x \{ J(x): y = \Phi x \} \tag{5.1.}$$

kada mjerenja ne sadrže šum, ili pak

$$\min_x \{ J(x): H(\Phi x, y) \leq \varepsilon \} \tag{5.2.}$$

kada mjerenja sadrže šum. Ovdje je H funkcija greške koja daje udaljenost između vektora Φx i y. Za pogodan parametar μ, relacija (5.2.) je ekvivalentna neograničavajućoj formulaciji optimizacijskog problema:

$$\min_x \{J(x) + \mu H(\Phi x, y)\} \tag{5.3.}$$

za neko $\mu > 0$. Parametar μ se može birati po principu pokušaja i grešaka, ili pomoću statističkih tehnika poput kros-validacije. U algoritmima konveksnog programiranja, J i H se najčešće biraju kao: $J(x) = \|x\|_1$, dakle $\ell_1$norma od x, a $H(\Phi x, y) = \frac{1}{2}\|\Phi x - y\|_2^2$, tj. kao $\ell_2$norma greške između posmatranih mjerenja i linearnih projekcija ciljnog vektora x. U statistici se problem minimizacije H pod uslovom da je $\|x\|_1 \leq \delta$ naziva *Lasso problem*. Generalno govoreći, $J(\cdot)$ se ponaša kao regulacioni uslov i može se zamijeniti drugim, kompleksnijim funkcijama.

Upotreba konvencionalnih paketa za konveksnu optimizaciju za formulacije (5.1.), (5.2.) i (5.3.) može biti vrlo nepraktična. Ipak, navedeni problemi optimizacije nameću dva glavna izazova koji su specifični za praktične probleme koji se susreću u CS-u: a) primjene u realnim uslovima su neizbježno jako komplikovane (slika rezolucije $1024x1024$ piksela vodi ka optimizaciji preko milion promjenljivih, što je izvan dosega bilo kojeg standardnog optimizacijskog software-skog paketa); b) ciljna funkcija nije glatka, a standardne tehnike zaglađivanja ne daju jako dobre rezultate. Dakle, za ovakve probleme konvencionalni algoritmi (koji tipično uključuju množenje matrica) nisu praktični, pa čak ni primjenljivi. Ovi jedinstveni izazovi koji se susreću u kontekstu CS-a su unutar naučne zajednice koja se bavi optimizacijom doveli do značajnog interesa za razvijanje poboljšanih algoritama za rekonstrukcija sparse signala [8].

5.2.2. Algoritmi koji se baziraju na ograničenom optimizacijskom problemu

U slučaju kada mjerenja ne sadrže šum, problem ℓ_1 minimizacije (koji se dobije za $J(x) = \|x\|_1$) se može predstaviti kao problem linearnog programiranja (skraćeno LP) sa jednadžbenim ograničenjima. Ovakvi problemi se mogu riješiti unutar polinomskog vremena $(O(N^3))$ korištenjem standardnih metoda unutrašnjih tačaka. Naime, ovo je bio prvi izvediv rekonstrukcijski algoritam koji se razvio za potrebe CS-a i on ima jake uslove koji osiguravaju ispunjenje, kao što je pokazano ranije u ovom radu.

Kada mjerenja ne sadrže šum, problem optimizacije se može predstaviti kao konusni problem drugog reda (eng. *second-order cone program,* skraćeno *SOCP*) sa kvadratnim jednadžbama kao ograničenjima. Rješavanje LP-ova i SOCP-ova predstavlja početni podsticaj na polju optimizacijskog istraživanja, jer je njihova primjena u praktičnim problemima CS-a ograničena zbog činjenice da su i N (broj članova vektori predstavljenog signala x) i M (broj mjerenja, tj. ograničenja) veoma veliki brojevi u mnogim slučajevima. Bitno je istaći da i LP i SOCP odgovaraju formulacijama (5.1.) i (5.2.), te se rješavaju koristeći metode unutrašnjih tačaka prvog reda [8].

5.2.2. Algoritmi koji se baziraju na neograničenom optimizacijskom problemu

Takođe se koriste i algoritmi koji za razliku od LP-a i SOCP-a rješavaju problem optimizacije bez ograničenja dat relacijom (5.3.). Jedan od tih algoritama je i *fixed-point continuation* algoritam (skraćeno FPC) čija je prednost upravo to što je iterativni algoritam, pa ga je lakše i računati, ali i isprogramirati. Predložene su mnoge varijacije, poboljšanja i generalizacije ovakvog pristupa u cilju efikasnosti osnovne iteracije, ali i u cilju proširivanja njegove primjenljivosti za razne odabire kriterija J. Ovaj algoritam ne bi bio efikasan da se parametar μ iz relacije (5.3.) postepeno ne smanjuje svakom sljedećom iteracijom, što dovodi do finalnog optimalnog rješenja.

Još jedna od bitnijih vrsta algoritama iz ove klase jesu Bregmanove iteracijske metode, koje se kao i FPC zasnivaju na rješavaju neograničenog optimizacijskom problema

datog relacijom (5.3.). Rješenja koja se dobiju ovakvom iteracijom jako brzo konvergiraju ka stvarnom rješenju za zatvorene i konveksne funkcije $J(x)$. Broj iteracija potrebnih za dobijanje tačne rekonstrukcije originalnog signala na osnovu mjerenja je tipično manji od 5 ukoliko se parametar μ pogodno odabere [8].

5.3. Greedy algoritmi

Nasuprot rješavanju programa konveksne optimizacije (što može biti računarski jako zahtjevno), rekonstrukcija signala sa mogućnošću kompresije se može izvršiti i metodama sparse aproksimacije. Cilj sparse rekonstrukcije je da se na osnovu linearnih mjerenja y rekonstruiše vektor x koji je najviše sparse. Drugim riječima, ovo se može predstaviti kao (nekonveksni) problem:

$$\min_{I}\{\,|I|: y = \sum_{i \in I} \phi_i x_i\} \tag{5.4.}$$

gdje I predstavlja određeni podskup indeksa $i = 1, \ldots, N$, a ϕ_i označava i-tu kolonu matrice mjerenja Φ. Poznato je da je potraga za optimalnim podskupom I^* (u skupu kojeg čine kolone matrice Φ) sa najmanjom kardinalnošću problem koji je NP-težak. Umjesto ovoga, klasične metode sparse rekonstrukcije rješavaju ovaj problem „greedily“ odabirajući kolone matrice Φ i formirajući sukcesivno bolje aproksimacije signala x na osnovu mjerenja y.

Metoda ***Matching Pursuit*** (skraćeno MP), koju su Mallat i Zhang imenovali i predstavili naučnoj zajednici iz oblasti obrade signala, je iterativni *greedy* algoritam koji vrši dekompoziciju signala u linearnu kombinaciju elemenata rječnika, pri čemu je matrica mjerenja $\Phi \in \mathbb{R}^{MxN}$ zapravo taj rječnik. Konceptualno, MP je jako jednostavan. Glavna veličina u ovom algoritmu rekonstrukcije je ostatak $r \in \mathbb{R}^M$ koji predstavlja dio mjerenja koji „još uvijek nije objašnjen“. U svakoj iteraciji algoritma se bira vektor iz rječnika koji je maksimalno koreliran sa ostatkom r, te se nakon određenih računanja (koja u ovom radu nisu navedena zbog ograničenog prostora) dobija bolja aproksimacija originalnog signala x, a potom se vrijednosti ostataka i aproksimiranog signala ažuriraju. Iterativni proces prestaje onda kada norma ostatka r postane manja od neke unaprijed zadate vrijednosti. Ovaj algoritam se može pokazati kao računarski neizvediv za mnoge probleme, s obzirom da kompleksnost MP-a raste linearno sa brojem iteracija. Dakako, maksimalan broj MP iteracija se može i ograničiti s gornje strane, a tako unaprijeđen MP algoritam naziva se ***Orthogonal Matching Pursuit*** (skraćeno OMP) koji konvergira za najviše K iteracija, gdje je K sparsity originalnog signala. Tropp i Gilbert su dokazali da se OMP može koristiti za rekonstrukciju sparse signala i to sa visokom vjerovatnoćom. Ovaj algoritam zahtijeva veću računarsku moć, pa se može pokazati da je ukupna kompleksnost OMP algoritma reda $O(MNK)$. U slučaju signala koji nije izraženo sparse, OMP nije efikasan pošto kompleksnost tada raste po kvadratnom zakonu sa brojem (K) nenultih vrijednosti originalnog signala.

Još jedan jako bitan *greedy* algoritam koji se koristi za rekonstrukciju sparse signala jeste ***Compressive Sampling Matching Pursuit*** (skraćeno CoSaMP) koji koristi osobinu

ograničene izometrije (RIP) matrice mjerenja da bi kao rezultat dao originalni signal x. Dakle, ukoliko matrica mjerenja Φ zadovoljava RIP, to znači da ona sadrži K kolona koje su međusobno približno ortogonalne, pa se ova osobina koristi za dokazivanje jake konvergencije rezultata ovog algoritma. Za razliku od MP i OMP algoritama, ovaj algoritam ima svojstvo da se indeksi u aproksimiranom signalu mogu dodavati, ali je moguće i brisati indekse iz skupa vrijednosti dobijenih prethodnom iteracijom (što nije moguće ukoliko se koriste MP ili OMP). Uzevši u obzir neke generalne pretpostavke, kompleksnost ovog algoritma je $O(MN)$, što je očito nezavisno od sparsitya originalnog signala, pa ovaj algoritam predstavlja poboljšanje u odnosu na konveksne metode i greedy algoritme. Iako je ovaj CoSaMP remek-djelo sparse rekonstrukcije on ipak posjeduje i jednu manu: algoritam zahtijeva poznavanje vrijednosti sparsitya K signala koji se rekonstruiše. Netačan unos sparsitya može dovesti do većih grešaka aproksimacije u odnosu na slučaj kada se koristi slabiji algoritam, npr. OMP. Granice stabilnosti koje prate CoSaMP osiguravaju da je greška usljed pogrešno odabranih parametara ograničena, ali još uvijek nije poznato kako se ove granice primjenjuju u praksi [8].

5.4. Kombinatorički algoritmi

Pored konveksne optimizacije i greedy algoritama, postoji još jedna bitna klasa algoritama za rekonstrukciju sparse signala koji se nazivaju **kombinatoričkim algoritmima**. Ovi algoritmi, najčešće razvijeni od strane naučne zajednice teorijske informatike, se u mnogim slučajevima pojavljuju i prije začetka literature na temu CS-a, ali si jako relevantni za problem rekonstrukcije sparse signala.

Najstariji kombinatorički algoritmi su se razvili u kontekstu testiranja grupe. Pri rješavanju ovog problema, pretpostavi se da postoji ukupno N članova grupe, od čega je nepoznat podskup od K članova koji imaju anomaliju, a koje treba pronaći. Npr, može se postaviti problem pronalaženja proizvoda sa defektom u industrijskom postrojenju, ili pak identifikacije podskupa zaraženih uzoraka tkiva u medicinskom kontekstu. U oba ova slučaja vektor x indicira koji elementi imaju anomaliju, tj. $x_i \neq 0$ za K elemenata sa anomalijom, a $x_i = 0$ u suprotnom. Cilj je osmisliti nekoliko testova koji omogućavaju da se identificira support (a po mogućnosti i vrijednosti nenultih elemenata) vektora x, dok se u isto vrijeme minimizira broj izvršenih testiranja. U najjednostavnijoj praktičnoj postavci ovi testovi su predstavljeni binarnom matricom mjerenja Φ čiji su elementi ϕ_{ij} jednaki 1 ako i samo ako je j-ti član korišten u i-tom testu. Ako je rezultat testa linearan s obzirom na ulazne podatke, onda je problem rekonstrukcije vektora x u biti isti kao i standardni problem rekonstrukcije sparse signala.

Nekoliko kombinatoričkih algoritama za rekonstrukciju sparse signala je razvijeno u literaturi. Neki od tih algoritama su: Random Fourier Sampling, HHS Pursuit, te Sparse Sequential Matching Pursuit. Zbog ograničenosti prostora, u ovom radu neće biti analizirani pojedini kombinatorički algoritmi, nego samo njihova glavna ideja. Ono što je bitno

napomenuti za kombnatoričke algoritme jeste to da je za tačnu rekonstrukciju originalnog signala potrebno poznavati i imati punu kontrolu nad matricom mjerenja Φ, što je upravo suprotno od principa na kojim počivaju konveksna optimizacija i greedy algoritmi (koji rade sa proizvoljnom matricom koja zadovoljava generički uslov kao što je RIP). Ovaj dodatni stepen slobode može voditi do puno bržih algoritama [8].

6. Primjena signala sa mogućnošću kompresije u praksi

6.1. Medicina: Rekonstrukcija slika magnetne rezonance, elektroencefalografija

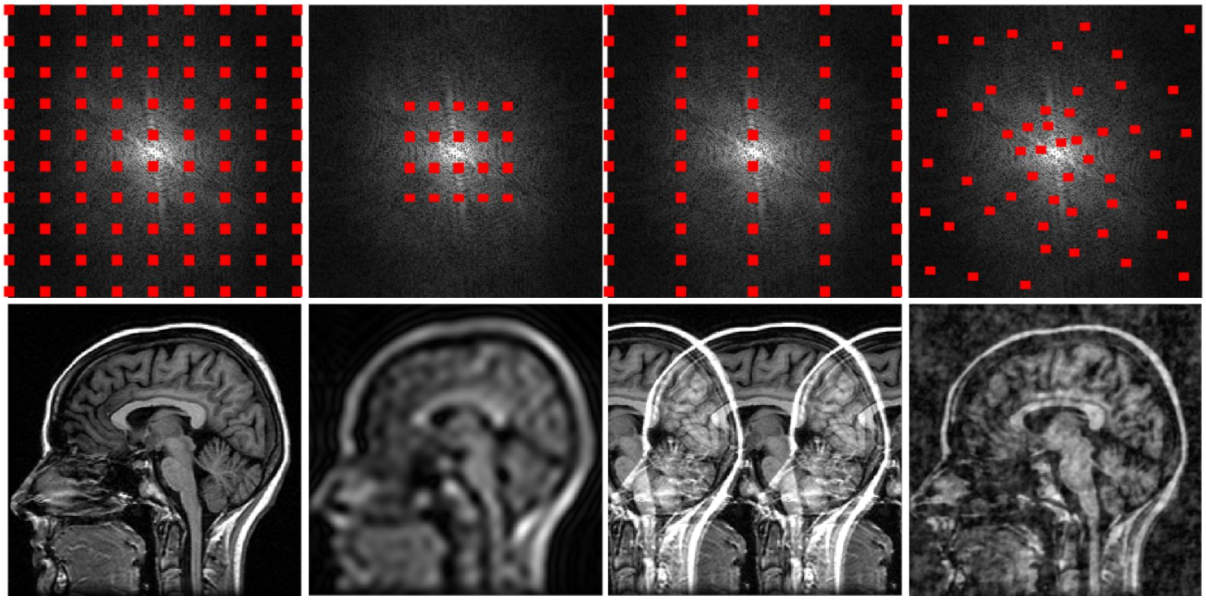
Magnetna rezonanca (eng. *Magnetic Resonance Imaging*, u daljem tekstu MRI) je medicinska tehnika „slikanja“ bazirana na principu da se protoni u molekulama vode u ljudskom tijelu usklađuju pod djelovanjem magnetnog polja. Uređaji za MRI neprestano pulsiraju magnetna polja da bi se molekule vode u ljudskom tijelu dezorijentisale a onda ponovo preorijentisale, što izaziva odašiljanje radiofrekvencija koje se mogu detektovati. Pretpostavlja se da je objekat koji se „slika“ zapravo skup trodimenzionalnih piksela (u daljem tekstu eng. *voxel*). Magnetni pulsevi se šalju inkrementalno duž gradijenta što dovodi do drugačijeg enkodiranja faze i frekvencije za svaku kolonu i red voxela, respektivno. Zanemarujući detalje samog fizikalnog procesa, magnetno polje mjereno prilikom MRI-a odgovara Fourierovim koeficijentima objekta koji se „slika“. Objekat se potom može rekonstruisati pomoću inverzne Fourierove transformacije, pa se zbog toga MRI može posmatrati i kao mjerenje Fourierovih koeficijenata.

Veliko ograničenje procesa MRI-a je linearna veza između broja izmjerenih podataka i vremena skeniranja. MRI skeniranja koja dugo traju psihološki negativno utječu na pacijenta, izazivaju nelagodu, a i skupa su. Prema tome, minimiziranje vremena skeniranja bez gubitka kvalitete slike je od velikog značaja za medicinu.

Teorija CS-a se može primijeniti na rekonstrukciju MRI slike, upravo koristeći činjenicu da su takve slike zapravo sparse 2-D signali u transformacijskom domenu. Prilikom standardne rekonstrukcije MRI slike, podsempliranje u Fourierovom domenu rezultira *aliasing*-om, što znači da tako rekonstruisana slika nije podudarna sa željenom originalnom slikom (Slika 6.1. b) i c)). Međutim, kada je slika objekta u poznatom domenu sparse ili ima mogućnost kompresije, onda se ona može rekonstruisati metodama i algoritmima rekonstrukcije čiji je pregled i kratak opis dat u Poglavlju 5. Za rekonstrukciju MRI slika se najčešće koriste DCT i Wavelet transformacija za računanje elemenata matrice transformacije Ψ, jer kao što je rečeno, u tim domenima su ovi 2-D signali sparse ili imaju mogućnost kompresije [16].

Kao primjer učinkovitosti primjene principa CS-a na uzorkovanje i rekonstrukciju MRI slika data je slika 6.1 koja predstavlja MRI snimak/sliku mozga. Međutim, zavisno od rasporeda i broja piksela koji se uzorkuju zavisi i kvaliteta rekonstrukcije originalne slike. Rezolucija slike je određena mjerom prekrivenosti K-prostora. Slike 6.1. u gornjem redu prikazuju raspored uzorkovanih piksela, dok slike 6.1. u donjem redu predstavljaju rekonstrukcije slike na osnovu tih mjerenja. Na slici 6.1. a) dat je ravnomjeran raspored mjerenja koji je u skladu sa Nyquistovim kriterijem za frekvenciju uzimanja uzoraka nekog signala. Rekonstruisana slika na osnovu ovih mjerenja je jako dobre kvalitete, ali je mana

ovakvog postupka upravo dugo trajanje skeniranja, veliki zahtjevi za memorijom da bi se pohranili svi izmjereni podaci, te čest osjećaj neugodnosti i klaustrofobije kod pacijenata. Na slici 6.1. b) dato je prostorno nejednako raspoređeno uzorkovanje koje pri tom ne zadovoljava Nyquistov uslov, pa je i rekonstruisana slika slabe kvalitete, tj. mutna je. Na slici 6.1. c) dato je prostorno jednako raspoređeno uzorkovanje koje pri tom ne zadovoljava Nyquistov uslov. U ovom slučaju je došlo do aliasinga (originalna slika je preklopljena svojim istovjetnim kopijama sa obje strane) jer je očito frekvencija uzorkovanja (širina između pojedinih kolona uzorkovanih piksela) manja od dvostruke frekvencije promjene ovakvog 2-D signala. Na slici 6.1. d) dat je primjer pseudo-slučajnog rasporeda uzorkovanih piksela, pri čemu je rekonstrukcija originalne slike na osnovu ovakvih mjerenja jako visoke kvalitete. Prednosti ovakve metode, koju u principu CS i zagovara, jesu: kraće trajanje pregleda, tj. skeniranja zbog manjeg broja potrebnih mjerenja, manji zahtjevi za memorijom koja je potrebna za pohranjivanje uzetih uzoraka, te smanjenje nelagode i klaustrofobije kod pacijenta.



a) b) c) d)

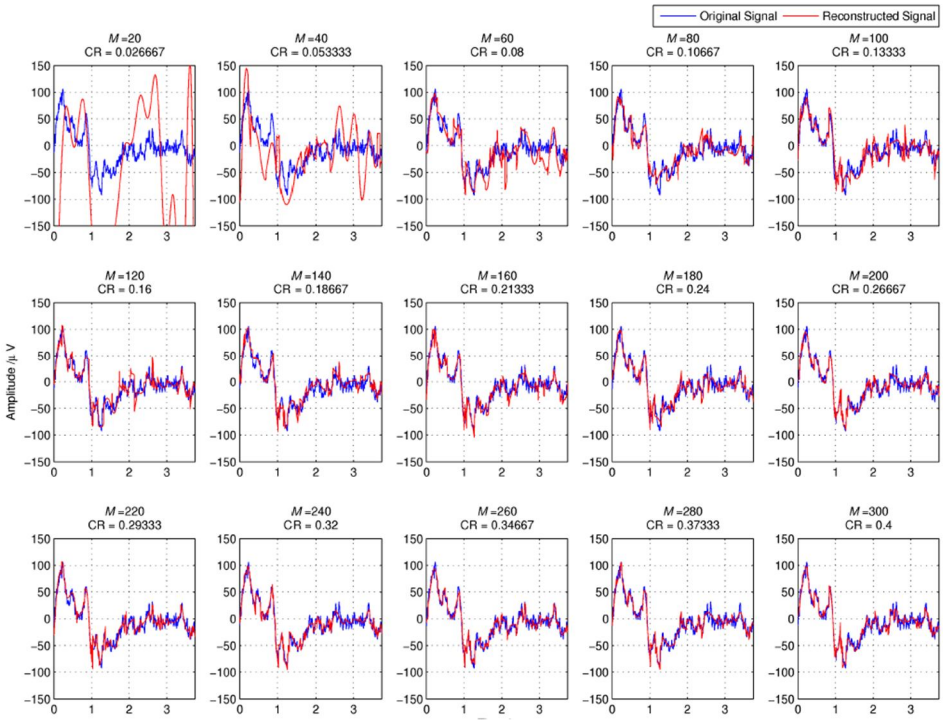
Slika 6.1. (a) Rekonstrukcija slike na osnovu uzorkovanja koje zadovoljava Nyquistov kriterij. (b) Prostorno nejednako raspoređeno uzorkovanje koje pri tom ne zadovoljava Nyquistov uslov. (c) Prostorno jednako raspoređeno uzorkovanje koje pri tom ne zadovoljava Nyquistov uslov. (d) Pseudo-slučajno raspoređena mjerenja (Slika preuzeta iz [17])

Elektroencefalografija (eng. *electroencephalography*, u daljem tekstu EEG) i **magnetoencefalografija** (eng. *magnetoencephalography*, u daljem tekstu MEG) su dvije jako popularne neinvazivne metode opisivanja funkcionisanja mozga koje se postiže mjerenjem raspodjela električnog potencijala i magnetnih polja skalpa uzrokovanih neuronskim okidanjima. EEG i MEG daju vremensku rezoluciju reda milisekunde na vremenskoj karakteristici neuronske aktivnosti, te mogu pomoći prilikom određivanja izvora struje u mozgu, i to rješavajući inverzni problem.

Modeli za neuromagnetne izvore smatraju da je neuronska aktivnost često prostorno ograničena. Na osnovu ove ideje, algoritmi poput FOCUSS-a (eng. *Focal Underdetermined System Solution*) se koriste za identifikaciju izrazito lokaliziranih izvora. Oni pretpostavljaju sparse model da bi riješili problem sistema koji ima više nepoznatih veličina nego jednadžbi koje ga opisuju.

FOCUSS je rekurzivna procedura linearne procjene bazirana na težinskom pseudo-inverznom rješenju. Algoritam pridružuje struju (sa nelinearnim parametrima lokacije struje) svakom elementu unutar jednog područja, tako da se vrijednosti nepoznate struje mogu linearno povezati sa mjerenjima. Težinski koeficijenti se pri svakom koraku računaju na osnovu rješenja prethodnog iteracijskog koraka. Algoritam konvergira prema raspodjeli izvora struje u kojoj broj parametara potreban za opisivanje struja tih izvora ne prekoračuje broj mjerenja. Inicijalizacija određuje prema kojem od lokaliziranih rješenja algoritam zapravo konvergira [18].

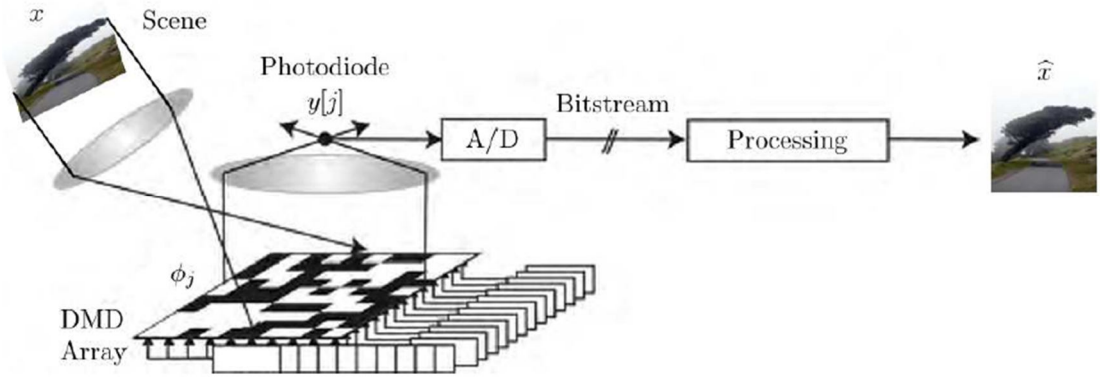
Pokazano je da je za pogodno odabran broj mjerenja M, pri čemu je $M \ll N$ (N je ukupan broj vrijednosti originalnog signala pretvorenog u vektor) moguće izvršiti tačnu rekonstrukciju originalnog signala samo na osnovu tih mjerenja. Slika 6.2. ilustruje jedan takav primjer gdje je ukupan broj vrijednosti vektora koji predstavlja razliku potencijala mjerenu na ljudskom skalpu $N = 750$, a prihvatljivo dobra rekonstrukcija se dobije već na osnovu samo $M = 120$ mjerenja, što upravo pokazuje veliku prednost CS-a nad do sada korištenim modelima uzorkovanja i rekonstrukcije signala.



Slika 6.2. Kvalitativna ilustracija rekonstrukcijskih performansi CS-a primijenjenih na EEG signale skalpa za različite faktore kompresije (eng. *Compression Ratio*, skraćeno CR) i različit broj mjerenja. Plavom bojom je označen originalni signal, dok je rekonstruisani signal crvene boje. (Slika preuzeta iz [19])

6.2. Jednopikselna kamera

U posljednjih nekoliko godina predloženo je nekoliko hardware-skih arhitektura za primjenu CS-a u tehnologiji fotoaparata. U ovom radu će se obraditi samo tzv. jednopikselna kamera. Ovakav fotoaparat je optički računar koji sekvencijalno mjeri unutrašnje proizvode $y(j) = \langle x, \phi_j \rangle$ između N-pikselne uzorkovane verzije slučajnog svjetlosnog polja iz scene koja se posmatra (obilježene sa x) i skupa od N-pikselnih testnih funkcija $\{\phi_j\}_{j=1}^{M}$. Arhitektura ovakvog fotoaparata je ilustrovana na slici 6.3., a prikaz fizičke realizacije je dat na slici 6.4. Kako je pokazano na ovim slikama, svjetlosno polje je fokusirano pomoću leće (označene sa „Lens 1“ na slici 6.4.) i to ne na CCD ili CMOS matricu uzorkovanja nego na prostorni modulator svjetlosti (eng. *spatial light modulator*, u daljem tekstu SLM). SLM modulira intenzitet svjetlosne zrake zavisno od upravljačkog signala. Jednostavan primjer transmisivnog SLM-a koji ili propušta ili blokira dijelove svjetlosne zrake je dodatna transparentnost, a drugi primjer jeste LCD projektor.

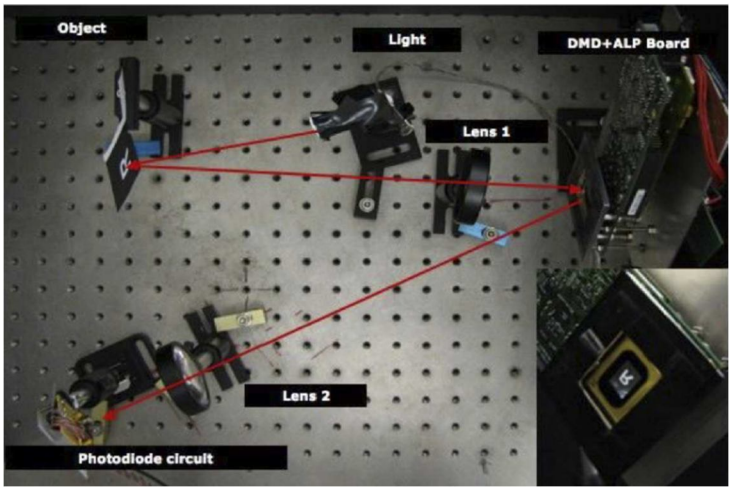


Slika 6.3. Blok dijagram jednopikselnog fotoaparata. (Slika preuzeza iz [20])

Digitalni mikroogledalni uređaj (eng. *micromirror device*, u daljem tekstu DMD) koji proizvodi kompanija Texas Instruments je reflektivni SLM koji selektivno preusmjerava dijelove svjetlosne zrake. DMD se sastoji od matrice elektrostatički aktuiranih mikroogledala veličine bakterije, gdje je svako ogledalo u matrici suspendirano iznad individualne ćelije memorije slučajnog pristupa (eng. *random access memory*, u daljem tekstu SRAM). Svako ogledalo se rotira oko zgloba i može se pozicionirati u jedan od dva položaja (± 10 stepeni u odnosu na horizontalu) u zavisnosti od toga koji bit je upisan u SRAM ćeliju. Tako se svjetlost koja pada na DMD može reflektovati u dva pravca u zavisnosti od orijentacije ogledala.

Svaki element SLM-a odgovara određenom elementu ϕ_j (i njegovom odgovarajućem pikselu u x). Za dato ϕ_j, odgovarajući element SLM-a se može orijentisati bilo „prema“ (što odgovara vrijednosti 1 kod tog elementa ϕ_j), ili pak „od“ (što odgovara vrijednosti 0 kod tog elementa ϕ_j) druge leće (označene sa „Lens 2“ na slici 6.4.). Ovo drugo sočivo skuplja reflektovanu svjetlost i fokusira je na jedan fotonski detektor (tj. na jedan piksel) koji integrira proizvod x i ϕ_j da bi izračunao mjerenje $y(j) = \langle x, \phi_j \rangle$ kao izlazni napon. Ovaj

napon se onda digitalizira pomoću A/D konvertora. Vrijednosti ϕ_j između 0 i 1 se mogu dobiti podrhtavanjem ogledala na jednu pa na drugu stranu u toku vremena integracije fotodiode. Predstavljanjem x-a kao vektor kolonu i ϕ_j kao vektor red(ove), moguće je modelirati ovakav sistem kao računanje proizvoda $y = \Phi x$, gdje ϕ_j odgovara svakom redu matrice Φ. Da bi se dobila slučajna mjerenja, orijentacije ogledala ϕ_j se zadaju potpuno slučajno koristeći generator pseudoslučajnih brojeva, potom se izračuna $y(j)$, te se ovaj proces ponovi M puta da bi se dobio vektor mjerenja y.



Slika 6.4. Prototip jednopikselnog fotoaparata (Slika preuzeta iz [20])

Dizajn jednopikselne kamere smanjuje potrebnu veličinu, kompleksnost i cijenu matrice fotonskih detektora na nivo samo jedne ćelije, što omogućava korištenje egzotičnih detektora koji se ne bi mogli koristiti u konvencionalnom digitalnom fotoaparatu. Neki od primjera ovih detektora uključuju fotomnožačku cijev (eng. *fotomultiplier tube*), ili lavinsku fotodiodu za slikanje u slučaju lošeg osvjetljenja, te spajanje nekoliko fotodioda osjetljivih na različite talasne dužine svjetlosti za multimodalno mjerenje, spektrometre za hiperspektralno slikanje itd.

Pored fleksibilnosti mjerenja, praktične prednosti jednopikselnog fotoaparata uključuje činjenice da je kvantna efikasnost fotodiode veća nego kod pikselnih senzora tipične CCD ili CMOS matrice, te za faktor ispune DMD-a može doseći i 90%, dok kod CCD/CMOS matrica on dostiže vrijednost od samo 50%. Bitna prednost koju treba naglasiti jeste da svako CS mjerenje primi oko $N/2$ puta više fotona nego obični pikselni senzor, što značajno smanjuje distorziju slike usljed tamnog šuma i šuma očitanja.

Jednopikselni fotoaparat spada u klasu multiplekserskih kamera. Osnovni standard za multipleksiranje je klasično raster skeniranje, gdje su testne funkcije ϕ_j sekvenca delta funkcija $\delta[n-j]$ koje okreću svako ogledalo. Postoje bitne prednosti rada u CS u odnosu na raster način skeniranja, uključujući manji broj mjerenja (M mjerenja kod CS-a u odnosu na N mjerenja kod raster skeniranja) i značajno smanjen tamni šum.

Slika 6.5. (a) i (b) ilustrira originalnu sliku (slovo R bijele boje na crnoj podlozi) x i rekonstruisanu sliku $\hat{x}$ uslikanu prototipom jednopikselnog fotoaparata sa slike 6.4., pri čemu je $N = 256x256$, a $M = N/50$.

Slika 6.5. (c) prikazuje sliku Mandrillovog testa, uslikanu u uvjetima lošeg osvjetljenja koristeći RGB filtere i fotomnožačku cijev, dimenzija $N = 256x256$ dobijenu pomoću jednopikselnog fotoaparata i to na osnovu $M = N/10$ mjerenja.

U oba slučaja je korištena minimizacija totalne varijacije (koja je u bliskoj vezi sa ℓ_1 minimizacijom *wavelet* koeficijenata) za rekonstrukciju slika [20].

Slika 6.5. Primjeri rekonstrukcija slika sa jednopikselnog fotoaparata. (a) $256x256$ konvencionalna slika crno-bijelog slova „R“. (b) Rekonstruisana slika na osnovu $M = 1300$ mjerenja jednopikselnog fotoaparata ($50x$ manje od Nyquistovog kriterija za uzorkovanje). (c) $256x256$pikselna rekonstrukcija slike Mandrillovog testa u boji, uslikanu u uvjetima lošeg osvjetljenja koristeći RGB filtere i fotomnožačku cijev, na osnovu samo $M = 6500$ slučajnih mjerenja. (Slika preuzeza iz [20])

7. Rekonstrukcija signala sa mogućnošću kompresije u Matlabu

U ovom poglavlju obrađen je praktični dio ovog rada kao dokaz da sav teoretski dio rekonstrukcije signala sa mogućnošću kompresije uistinu funkcioniše i u praksi. Korišten je programski paket Matlab, te besplatni *toolbox " l1 magic*" koji je razvio Emmanuel Candes, jedan od vodećih naučnika na polju Compressive sensing-a. L1-magic je zapravo skup programskih rutina napravljenih za Matlab, a služi upravo za rekonstrukciju signala sa mogućnošću kompresije i to koristeći metode konveksne optimizacije čiji je matematički opis dat u poglavlju 5.1. Algoritmi su bazirani na standardnoj metodi unutrašnjih tačaka, te su pogodni i za optimizacione probleme velikih razmjera. U nastavku su obrađena dva primjera rekonstrukcije signala i to signala koji su sparse u vremenskom, a potom i u frekventnom domenu.

7.1. Rekonstrukcija signala koji je sparse u vremenskom domenu

Prvi primjer rekonstrukcije signala sa mogućnošću kompresije je signal koji je sparse u vremenskom domenu. Originalni signal x zapravo predstavlja puls konačne dužine, odnosno x ima nenulte vrijednosti samo za prvih $K = 50$ uzoraka, dok je ostatak vrijednosti signala x jednak nuli. Bitno je napomenuti da je ukupni broj vrijednosti signala x jednak $N = 1024$. Ovo znači da x u vremenskom domenu ima samo 4.88% nenultih vrijednosti, dok su preostale vrijednosti kao što je ranije rečeno – jednake nuli. Dakle, zadovoljen je uslov da je $K \ll N$, pri čemu je K mjera sparsitya signala, a N ukupan broj vrijednosti signala x. Dakle zadati signal se može opisati sljedećim relacijama:

Broj uzoraka:
$N = 1024$

Sparsity u vremenskom domenu:
$K = 50$

Vremenski oblik signala:
$x = [x_1, x_2, x_3, \dots, x_i, \dots, x_N]^T, i \in [1, N]$

$$x = \begin{cases} 1, i \epsilon [2, K-1] \\ 0.5, i \in \{1, K\} \\ 0, i \in [K+1, N] \end{cases}$$

Problem određivanja minimalnog broja pseudoslučajnih mjerenja M na osnovu kojih je teoretski moguće izvršiti rekonstrukciju originalnog signala x je riješen analizom teorema

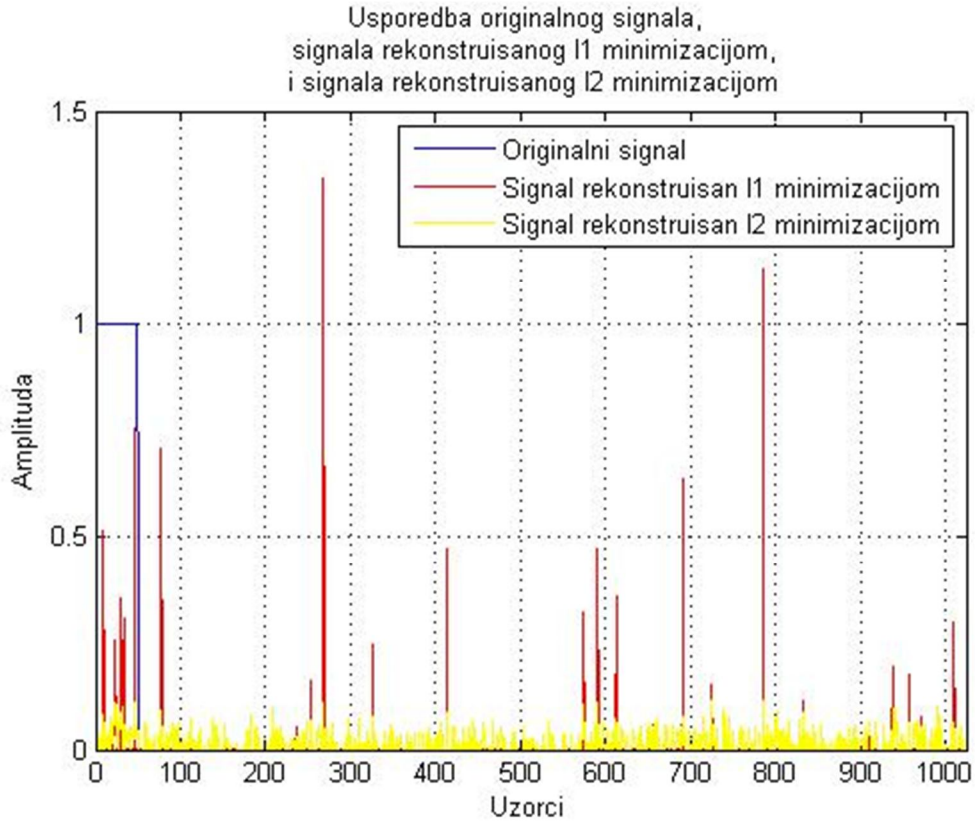
datih u Poglavlju 3. Teorema 3.1. implicira da mora biti zadovoljeno $M \geq 2K$, to znači da mora biti zadovoljena nejednakost $M \geq 100$. Nadalje, na osnovu Teoreme 3.3. koja kaže da mora biti zadovoljen uslov $M \geq CKln\left(\frac{N}{K}\right)$, što pri usvojenoj konstanti $C = 0{,}28$ kako je to pomenuto uz teoremu znači da je drugi uslov donje granice za broj mjerenja $M \geq 43$. Sada će biti razmotrene kvalitete rekonstrukcije originalnog signala za različite vrijednosti mjerenja M. Pseudokod optimizacijskog algoritma glasi:

```
N = 1024; % postavljanje ukupnog broja vrijednosti originalnog signala
M = 20; % broj izvršenih mjerenja

x = zeros(N,1);  % kreiranje pulsa konačne širine u vremenskom domenu
x(1)=0.5;
for i=2:49
    x(i)=1;
end
x(50)=0.5;

A = randn(M,N); % kreiranje matrice mjerenja A kao slučajne matrice
                % dimenzija MxN

A = orth(A')';  % sada su sve kolone matrice mjerenja A ortogonalne

y = A*x;    % kreiranje vektora slučajnih mjerenja

x0 = A'*y; % početna pretpostavka za početak rekonstrukcije originalnog
           % signala je data u obliku rješenje inverznog problema kojeg
           % Matlab rješava %kao L2 minimizacijski problem, pa se kao
           % rezultat dobije signal minimalne energije koji nije sparse

% rješavanje linearnog programa
tic
% Koristi se l1eq_pd funkcija za rješavanje problema linearnog
% programiranja koja kao argumente uzima početnu pretpostavku aproksimacije
% x0, matricu mjerenja A, transponovanu matricu A koja je nepotrebna pa se
% funkciji prosljeđuje kao prazna matrica, potom vektor mjerenja y i
% konačno zahtijevana granica greške aproksimacije. Funkcija tic pokreće
% tajmer koji mjeri vrijeme potrebno za rekonstrukciju originalnog signala,
% pa se taj tajmer zaustavlja pozivom funkcije toc i to onda kada je
% rekonstrukcija originalnog signala izvršena.
xp = l1eq_pd(x0, A, [], y, 1e-3);
toc

norm(xp-x)  % L2 norma razlike aproksimiranog i originalnog signala kao
            % mjera greške aproksimacije l1 minimizaciaje

norm(x0-x)  % L2 norma razlike aproksimiranog i originalnog signala kao
            % mjera greške aproksimacije l2  minimizacije
```

Pomoću koda iz Matlaba urađeno je nekoliko testnih rekonstrukcija za slučajeve $M = 50, 150, 200$ mjerenja. Svaki od primjera je ukratko analiziran po pitanju greške aproksimacije čiju mjeru matematički predstavlja ℓ_2 norma razlike aproksimiranog i originalnog signala, potom vizuelnog poklapanja pomenutih signala, te usporedba performansi rekonstrukcije bazirane na ℓ_1 minimizacijskom problemu sa rekonstrukcijom

baziranom na ℓ_2 minimizacijskom problemu koji predstavlja aproksimaciju signala minimalne energije na osnovu datih mjerenja, te se intuitivno unaprijed pretpostavlja da će performanse ovakvog optimizacionog algoritma biti jako loše. Signal x_1 neka označava rekonstruisani signal pomoću ℓ_1 minimizacije, a x_2 neka označava rekonstruisani signal pomoću ℓ_2 minimizacije zbog preglednosti i lakše uočljivosti performansi ove dvije metode.

Test 1: Rekonstrukcija signala x dužine $N = 1024$ na osnovu $M = 50$ mjerenja



Slika 7.1. Rekonstrukcija originalnog signala na osnovu 50 mjerenja

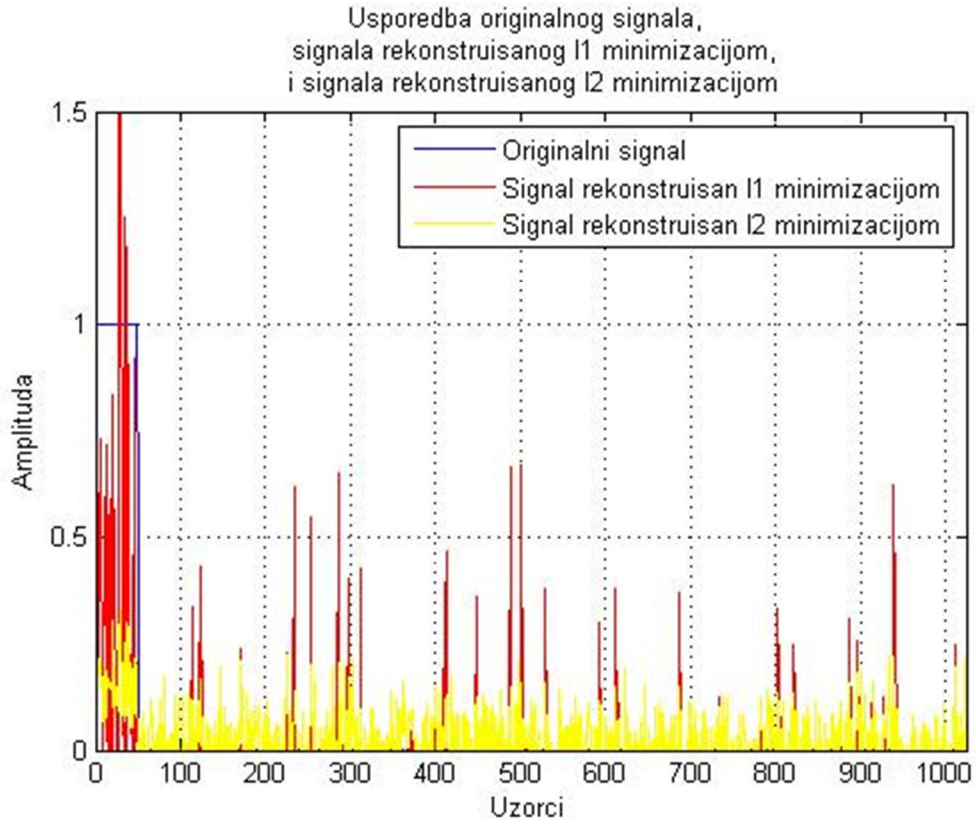
Sa slike 7.1. se vidi da nijedna od dvije metode rješavanja optimizacijskom problema nije postigla željene rezultate, čak šta više aproksimirani signal dobiven bilo jednom ili drugom metodom uopšte ne odgovara originalnom signalu. Greške pri aproksimaciji ℓ_1 minimizacijom, odnosno ℓ_2 minimizacijom iznose :

$$\|\hat{x}_1 - x\|_2 = \ 7.2609$$
$$\|\hat{x}_2 - x\|_2 = \ 6.8031$$

U ovom testnom primjeru se dolazi do zaključka da pri ovako malom broju mjerenja ℓ_2 minimizacija daje manju grešku u odnosu na ℓ_1 minimizaciju, ali bitno je uzeti u obzir da niti jedna od ovih minimizacijskih metoda ne daje ni približno zadovoljavajuće rezultate.

Zaključak ovog testa je da je potrebno povećati broj mjerenja da bi rekonstrukcija bila približnija.

Test 2: Rekonstrukcija signala x dužine $N = 1024$ na osnovu $M = 150$ mjerenja



Slika 7.2. Rekonstrukcija originalnog signala na osnovu 150 mjerenja

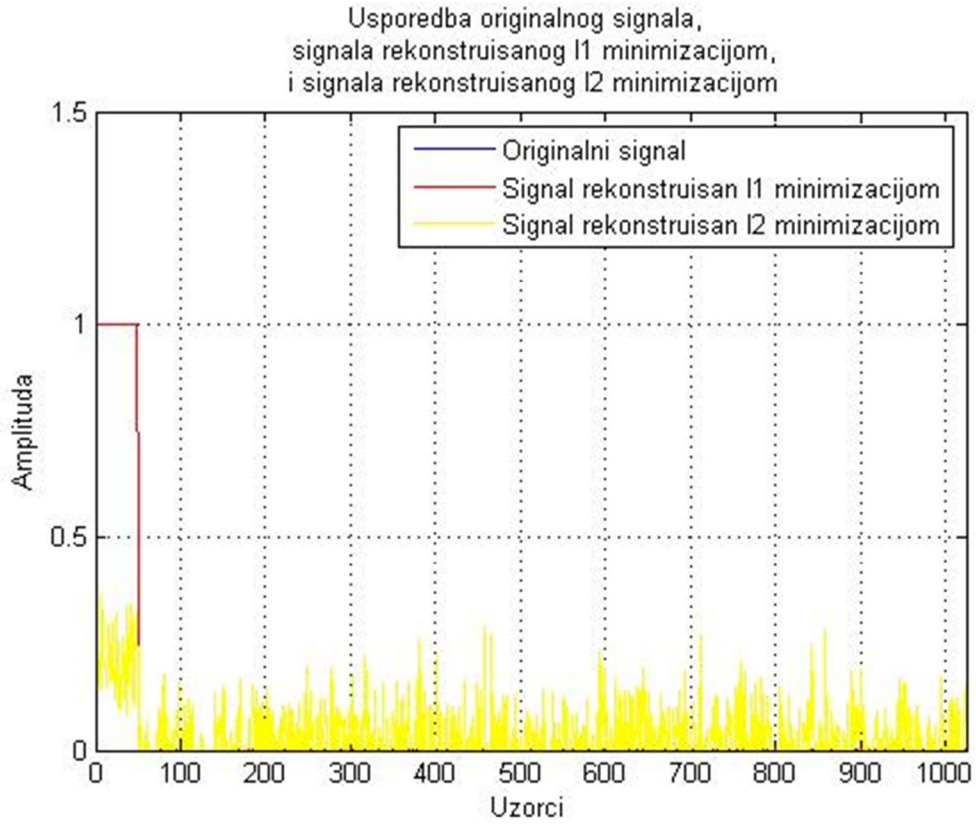
Sa slike 7.2. se vidi da nijedna od dvije metode rješavanja optimizacijskom problema nije postigla željene rezultate, čak šta više aproksimirani signal dobiven bilo jednom ili drugom metodom uopšte ne odgovara originalnom signalu kao i u prethodnom primjeru, iako je broj mjerenja povećan $7{,}5x$ u odnosu na inicijalni. Greške pri aproksimaciji ℓ_1 minimizacijom, odnosno ℓ_2 minimizacijom iznose :

$$\|\hat{x}_1 - x\|_2 = 5.9357$$
$$\|\hat{x}_2 - x\|_2 = 6.5057$$

U ovom primjeru dolazi do obrata po pitanju odnosa grešaka aproksimacije ℓ_1 minimizacije i ℓ_2 minimizacije, tj. već u ovom primjeru je greška pri ℓ_1 minimizaciji manja u odnosu na grešku pri ℓ_2 minimizaciji, ali još uvijek niti jedna od ovih minimizacijskim metoda ne daje ni približno zadovoljavajuće rezultate, odnosno greške aproksimacije su mnogo veće od dozvoljene.

Zaključak ovog testa je da je ponovo potrebno povećati broj mjerenja da bi rekonstrukcija bila tačna.

Test 3: Rekonstrukcija signala x dužine $N = 1024$ na osnovu $M = 200$ mjerenja



Slika 7.3. Rekonstrukcija originalnog signala na osnovu 200 mjerenja

Sa slike 7.3. se vidi da pri povećanju broja mjerenja $4{,}65x$ u odnosu na vrijednost koja predstavlja apsolutni teoretski minimum dolazi do tačne rekonstrukcije originalnog signala, pa se vizuelno originalni i rekonstruisani signal praktično ni ne razlikuju osim možda u tačkama gdje originalni signal ima nulte vrijednosti. Međutim, ovakav subjektivan sud naravno nije dovoljno dobar i egzaktan pokazatelj kvalitete rekonstrukcije odnosno greške pri aproksimaciji, pa tako greška aproksimacije za ℓ_1 minimizaciju, odnosno ℓ_2 minimizaciju u ovom slučaju iznose :

$$\|\hat{x}_1 - x\|_2 = 7.3110e-005$$
$$\|\hat{x}_2 - x\|_2 = 6.2468$$

U ovom primjeru ℓ_1 minimizacija pokazuje svu svoju moć i daje grešku za 5 redova veličine manju u odnosu na ℓ_2 minimizaciju što dokazuje princip da ℓ_2 minimizacija za odgovarajući broj mjerenja nužnih za tačnu rekonstrukciju signala daje signal minimalne energije, a ne signal koji je najviše prorijeđen, dok ℓ_1 minimizacija upravo promoviše prorijeđenost kao svojstvo signala čiju rekonstrukciju vrši.

Ono što je još zanimljivo jeste da sa porastom broja mjerenja četiri puta u odnosu na prvi testni primjer, greška pri aproksimaciji ℓ_1 minimizacijom padne za pet redova veličine, dok greška pri aproksimaciji ℓ_2 minimizacijom padne za manje od 1, što ponovo dokazuje osobinu ℓ_2 minimizacije da kao rješenje daje signal minimalne energije, dok ℓ_1 minimizacija

(naravno za odgovarajući broj mjerenja) daje najprorjeđeniji signal, što je upravo i glavni razlog korištenja ove metode kako je rečeno i u Poglavlju 4. Primijećeno je da za isti broj mjerenja algoritam ne daje iste rezultate što dokazuje da je rekonstrukcija zagarantovana samo u domenu vjerovatnoće, koja je jako visoka, ali nije $100\% - tna$, što znači predstavlja problem ukoliko se traži ponovljivost performansi algoritma rekonstrukcije prilikom svakog pokretanja algoritma.

7.2. Rekonstrukcija signala koji je sparse u frekventnom domenu

U drugom primjeru rekonstrukcije sparse signala dati su testovi rekonstrukcije signala koji je sparse u frekventnom domenu. Očit primjer ovakvog signala je sinusoidalni signal i to diskretna sinusna sekvenca $x[i]$ pošto Matlab jasno sve varijable pohranjuje u obliku matrica, odnosno u ovom slučaju vektora (kao specijalnog slučaja matrice dimenzija $Nx1$) dužine $N = 1024$. Diskretna sinusoida je zadata kao zbir tri sinusne sekvence frekvencija $29, 100\, i\, 60$. Ukoliko se ove frekvencije pomnože faktorom $\frac{2\pi}{N}$, to se dobiju vrijednosti odgovarajućih radijan frekvencija sinusnih sekvenci. U domenu Fourierove transformacije, ovi signali će predstavljati 6 impulsa, što znači da sparsity ovakvog signala iznosi $K = 6$.

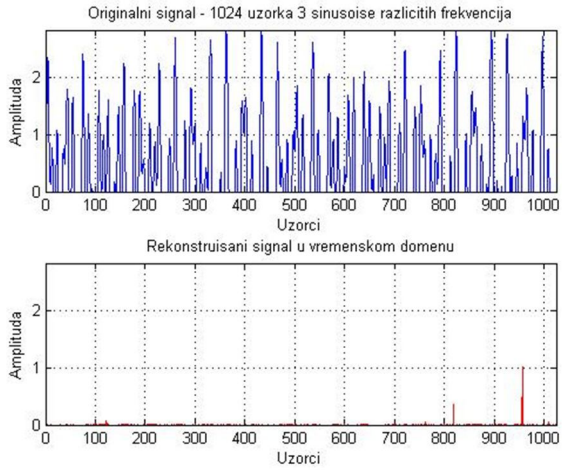
Na osnovu Teoreme 3.1. mora biti zadovoljen uslov $M \geq 2K$, što u ovom slučaju daje nejednakost $M \geq 12$. Još jedan uslov minimalnog broja mjerenja potrebnog za tačnu rekonstrukciju originalnog signala dat je relacijom (3.8.) u Teoremi 3.3., prema kojoj broj mjerenja mora zadovoljavati uslov $M \geq CKlog\left(\frac{N}{K}\right)$, tj. uz pretpostavljeno $C = 0.28$, dobija se $M \geq 8.63, tj. M \geq 9$. Dio koda specijaliziranog za rekonstrukciju signala koji su sparse u frekventnom domenu dat je u nastavku:

```
N=1024; %broj vrijednosti originalnog signala

M=256; %broj mjerenja (promjenljivo)

k1=29; %frekvencije sinusoida
k2=100;
k3=60;

n=0:N-1; %vektor indeksa diskretne sinusoide koja predstavlja zbir 3 sinusa

%originalni signal
x=sin(2*pi*(k1/N)*n)+sin(2*pi*(k2/N)*n)+sin(2*pi*(k3/N)*n);

% Kreiranje matrice Diskretne Fourierove transformacije dimenzija NxN, koja
% predstavlja matricu kompleksnih brojeva duž jedinične kružnice. Kada se
% ova matrica pomnoži vektorom koji predstavlja originalni signal kao
% rezultat se dobije vektor koeficijenata DCT-a originalnog signala
B=dftmtx(N);
Binv=inv(B);  % računanje matrice za inverznu Diskretnu Fourierovu
              % transformaciju
```

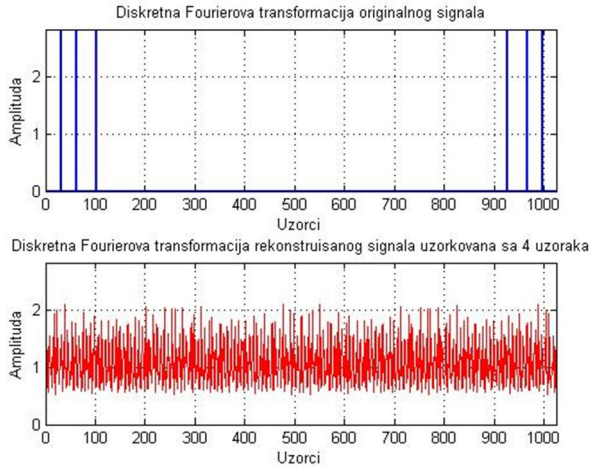
```
% računanje koeficijenata Diskretne Fourierove transformacije originalnog
% signala
xf=B*x';

q=randperm(N); % vektor q je vektor slučajno permutiranih cijelih brojeva
od 1 do N uključivo

A=Binv(q(1:M),:); %kreiranje matrice mjerenja A na način da se iz matrice
Binv uzme M slučajno odabranih redova čiji su indeksi određeni sa prvih M
vrijednosti vektora q (vektora slučajno permutiranih cijelih brojeva od 1
do N uključivo)

y=(A*xf);  %uzimanje slučajnih mjerenja koeficijenata DCT-a originalnog
           % signala

x0=A'*y;   % početna pretpostavka za početak rekonstrukcije koeficijenata
           % DCT-a originalnog signala je data u obliku rješenje inverznog
           % problema kojeg Matlab rješava kao L2 minimizacijski problem,
           % pa se kao rezultat dobije signal minimalne energije koji nije
           % sparse

%pokretanje algoritma rekonstrukcije
tic
% Koristi se l1eq_pd funkcija za rješavanje problema linearnog
% programiranja koja kao argumente uzima početnu pretpostavku aproksimacije
% x0, matricu mjerenja A, transponovanu matricu A koja je nepotrebna pa se
% funkciji prosljeđuje kao prazna matrica, potom vektor mjerenja y i
% konačno zahtijevana granica greške aproksimacije. Funkcija tic pokreće
% tajmer koji mjeri vrijeme potrebno za rekonstrukciju koeficijenata DCT-a
% originalnog signala, pa se taj tajmer zaustavlja pozivom funkcije toc i
% to onda kada je rekonstrukcija originalnog signala izvršena.
xp=l1eq_pd(x0,A,[],y,1e-5);
toc

xprec=real(Binv*xp);  %rekonstruisani signal u vremenskom domenu dobiven
% kao realni dio proizvoda matrice inverznog DCT-a i koeficijenata
% dobivenih algoritmom rekonstrukcije

 norm(xprec-x') %greška aproksimacije kao l2 norma razlike rekonstruisanog
                % i originalnog signala
```

Pomoću koda iz Matlaba urađeno je nekoliko testnih rekonstrukcija za slučajeve $M = 9, 12, 256$ mjerenja. Svaki od primjera je ukratko analiziran po pitanju greške aproksimacije čiju mjeru matematički predstavlja ℓ_2 norma razlike aproksimiranog i originalnog signala, potom vizuelnog poklapanja pomenutih signala i njihovih spektara.

Test 1: Rekonstrukcija signala x dužine $N = 1024$ na osnovu $M = 9$ mjerenja



(a)



(b)

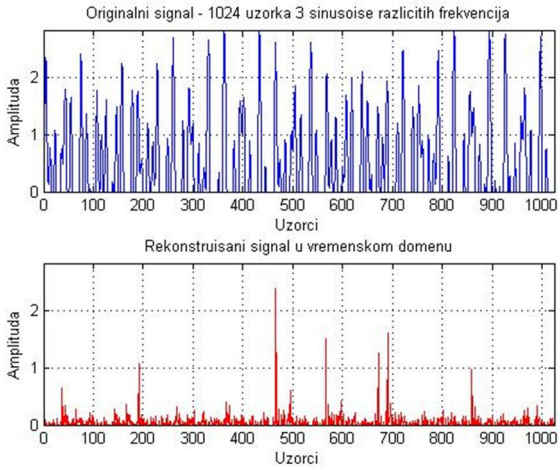
Slika 7.4. (a) Vremenski oblici originalnog i rekonstruisanog isgnala. (b) Diskretna Fourierova transformacija originalnog i rekonstruisanog signala. (c)

Poredeći originalni signal sa rekonstruisanim signalnom (Slika 4.5. (a)) na osnovu teoretskog minimuma na osnovu Teoreme 3.3. je očito da je aproksimacija jako loša, pa se logično ni spektri ovakvih signala uopšte ne podudaraju što se vidi na Slici 7.4. (b). Naime greška aproksimacije predstavljena kao ℓ_2 norma razlike aproksimiranog i originalnog signala iznosi:

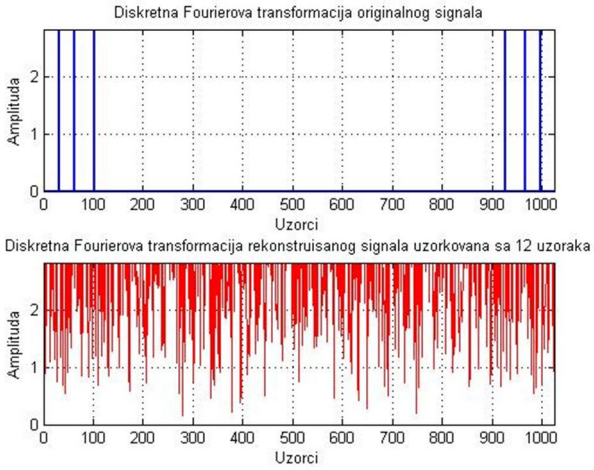
$$\|\hat{x} - x\|_2 = 39.1931$$

što je jako veliko. Ovo znači da je potrebno povećati broj mjerenja da bi aproksimacija originalnog signala ℓ_1 minimizacijom bila tačnija.

Test 2: Rekonstrukcija signala x dužine $N = 1024$ na osnovu $M = 12$ mjerenja



(a)



(b)

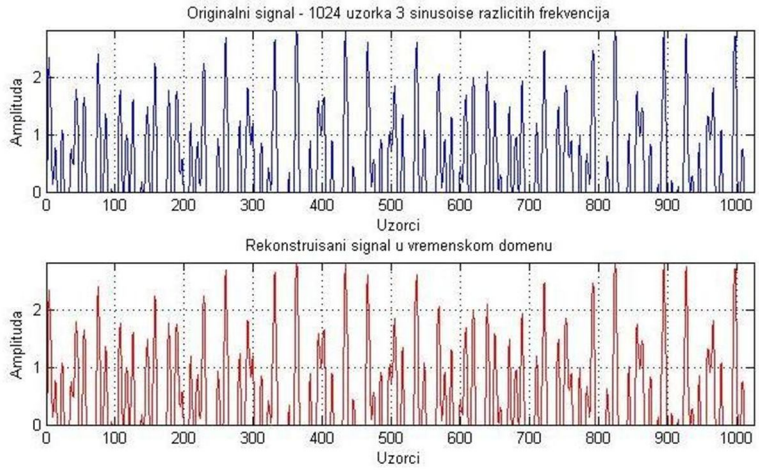
Slika 7.5. (a) Vremenski oblici originalnog i rekonstruisanog isgnala. (b) Diskretna Fourierova transformacija originalnog i rekonstruisanog signala. (c)

Poredeći originalni signal sa rekonstruisanim signalnom (Slika 7.5. (a)) na osnovu teoretskog minimuma na osnovu Teoreme 3.1. je očito da je aproksimacija loša, pa se logično ni spektri ovakvih signala uopšte ne podudaraju što se vidi na Slici 7.5. (b). Naime greška aproksimacije predstavljena kao ℓ_2 norma razlike aproksimiranog i originalnog signala iznosi:

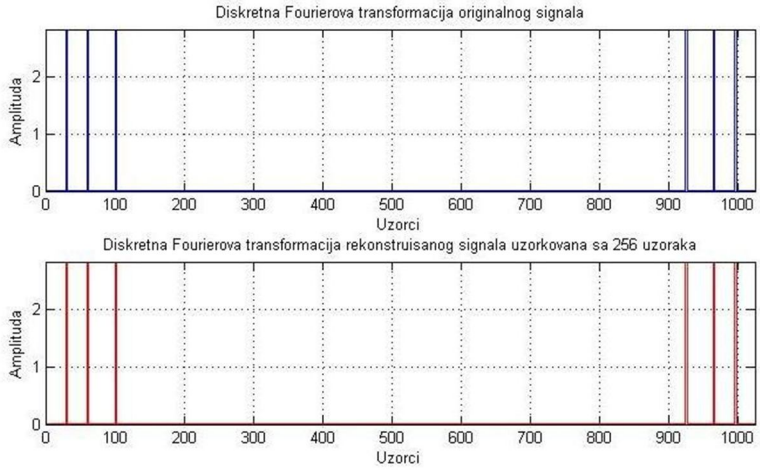
$$\|\hat{x} - x\|_2 = 39.1660$$

što je jako veliko, ali opet manje od greške aproksimacije u prošlom testu. Ovo znači da je potrebno dodatno povećati broj mjerenja da bi aproksimacija originalnog signala ℓ_1 minimizacijom bila tačnija.

Test 3: Rekonstrukcija signala x dužine $N = 1024$ na osnovu $M = 256$ mjerenja



(a)



(b)

Slika 7.6. (a) Vremenski oblici originalnog i rekonstruisanog isgnala. (b) Diskretna Fourierova transformacija originalnog i rekonstruisanog signala. (c)

Poredeći originalni signal sa rekonstruisanim signalom (Slika 7.6. (a)) na osnovu $M = 256$ mjerenja jasno je da je aproksimacija dobra, pa se signali u vremenskom domenu skoro u potpunosti podudaraju, a logično je da se i spektri ovakvih signala podudaraju što se vidi na Slici 7.6. (b). Naime greška aproksimacije predstavljena kao ℓ_2 norma razlike aproksimiranog i originalnog signala iznosi:

$$\|\hat{x} - x\|_2 = 2.7501\mathrm{e} - 004$$

što je zadovoljavajuće, čak šta više jako dobro. Ovo znači da nije potrebno dalje povećavati broj mjerenja, te da se na osnovu samo $1/4$ ukupnog broja vrijednosti originalnog signala može izvršiti rekonstrukcija sa zadovoljavajućom tačnosti.

Kao i u slučaju rekonstrukcije signala koji je sparse u vremenskom domenu, primijećeno je da za isti broj mjerenja performanse algoritma rekonstrukcije mogu varirati od jako dobrih do loših, jer je tačnost rekonstrukcije teoretski zagarantovana određenom

vjerovatnoćom, koja nije 100%-tna, ali je jako blizu te vrijednosti. I u ovom razmatranju, ukoliko se od algoritma rekonstrukcije zahtijeva ponovljivost, upitno je da li će ovaj ili bilo koji algoritam dati iste rezultate za ponovljene parametre rekonstrukcije, što može predstavljati problem u nekim praktičnim situacijama gdje se iziskuje velik stepen ponovljivosti.

8. Zaključak

U ovom radu obuhvaćena je teorija rekonstrukcije signala sa mogućnošću kompresije, od same definicije signala i rekonstrukcije, preko definiranja prorijeđenih signala i signala sa mogućnošću kompresije. Bitno je napomenuti da su u ovaj rad uvršteni i uslovi koje matrica mjerenja mora zadovoljavati da bi rekonstrukcija signala sa mogućnošću kompresije bila moguća na osnovu relativno malo mjerenja, te da slučajna mjerenja pokazuju izuzetno dobre performanse u smislu postotka vjerovatnoće da će rekonstrukcija signala sa mogućnošću kompresije biti tačna na osnovu takvih mjerenja. Takođe, nakon postavljanja uslova na matricu mjerenja, obrađeni su i načini rekonstrukcije signala sa i bez šuma, gdje je pokazano kako ovakav način rekonstrukcije signala sa superponiranim šumom ima zavidne performanse u odnosu na rekonstrukciju istog signala standardnom Nyquistovom metodom uzorkovanja. U Poglavlju 7. je i praktično dokazano da teorija rekonstrukcije signala sa mogućnošću kompresije daje jako dobre rezultate, te da je samim tim vjerovatnoća tačne rekonstrukcije velika, a greška aproksimacije pri istoj jako mala.

Koncepti koji se pojavljuju u ovom radu suštinski nisu novi, ali su primijenjeni na sasvim suprotan način u odnosu na dosadašnju teoriju vezanu za problem uzorkovanja. Ono što je revolucionarno jeste novo viđenje mogućnosti primjene nekih matematičkih aparata koji naizgled nemaju nikakve veze sa temom. Ono što je možda i najviše zapanjujući aspekt ovakvog pristupa procesima uzorkovanja i rekonstrukcije signala jeste pojava slučajnih mjerenja, te tačne rekonstrukcije na osnovu istih, što naizgled možda i izgleda nemoguće, ali je teorijski opisano i dokazano.

Ovaj rad predstavlja samo mali dio svega onoga što se može napisati o odabranoj temi, ali zbog ograničenog prostora i ponekad zbunjujuće i komplicirane matematičke teorije koja izlazi iz domena dosadašnjih spoznaja na Prvom ciklusu studija, odabrane su samo najbitnije i fundamentalne stvari prije svega.

9. Literatura

[1] V. A. Kotelnikov, "*On the transmission capacity of the 'ether' and of cables in electrical communications*", All-Union Conf. Technological Reconstruction of the Communications Sector and the Development of Low-current Engineering, 1933.

[2] C. E. Shannon, "*Communications in the presence of noise*", Proc. IRE vol. 37, 1949.

[3] H. Nyquist, "*Certain topics in telegraph transmission theory*", Trans. AIEE vol. 47, 1928.

[4] Raromir S. Stankovic, Jaakko T. Astola, Mark G. Karpovsk, „*Some Historic Remarks on Sampling Theorem*", Academy of Finland, Finnish Center of Excellence Programme, Grant No. 44876, 2008.

[5] *Melita Ahić-Đokić, "Signali i sistemi"*, Elektrotehnički fakultet Sarajevo, 2010.

[6] www.tonmeister.ca/main/textbook/intro_to_sound_recordingch9.html

[7] Huse Fatkić, *Predavanja iz predmeta "Inženjerska matematika 2",* Elektrotehnički fakultet Sarajevo, 2010.

[8] Richard Baraniuk, Mark A. Davenport, Marco F. Duarte, Chinmay Hegde, *„An Introduction to Compressive Sensing"*, Course col11133, 2011.

[9] Richard Baraniuk, *„Compressive sensing"*, IEEE Signal Processing Magazine, , 2007.

[10] Mark A. Davenport, Marco F. Duarte, Yonina C. Eldar, Gitta Kutyniok, *„Introduction to compressed sensing",* Stanford University, 2011.

[11] Emmanuel J. Candès, Michael B. Wakin, *„An introduction to Compressive Sampling"*, IEEE SIGNAL PROCESSING MAGAZINE ,2008.

[12] D. Donoho and M. Elad, *„Optimally sparse representation in general (nonorthogonal) dictionaries via l1 minimization"*, Proc. Natl. Acad. Sci., 100(5):21978211;2202, 2003.

[13] A. Cohen, W. Dahmen, R. DeVore, *„Compressed sensing and best k-term approximation",* J. Amer.Math. Soc., 22(1):2118211;231, 2009.

[14] Massimo Fornasier, Holger Rauhut, *„Compressive Sensing"*, Austrian Academy of Sciences Johann Radon Institute for Computational and Applied Mathematics,2011.

[15] S. Chen, D. Donoho, M. Saunders, *„Atomic decomposition by basis pursuit",* SIAM J. Sci. Comp.,20(1):338211;61, 1998.

[16] M. Lustig, D. Donoho, J. Pauly, *„Rapid mr imaging with compressed sensing and randomly undersampled 3dft trajectories“,* Proc. Annual Meeting of ISMRM, Seattle, WA, 2006.

[17] Michael Lustig, David L. Donoho, Juan M. Santos, John M. Pauly, *„Compressed Sensing MRI*“, Magnetic Resonance Systems Research Laboratory, Department of Electrical Engineering, Stanford University, 2007.

[18] I. Gorodnitsky and B. Rao, "*Convergence analysis of a class of adaptive weighted norm extrapolation algorithm“, In Proc. Asilomar Conf. Signals, Systems, and Computers*“, Paci_c Grove, A, 1993.

[19] Amir M. Abdulghani, Alexander J. Casson and Esther Rodriguez-Villegas, *“Quantifying the performance of compressive sensing on scalp EEG signals“*, Electrical and Electronic Engineering Department, Imperial College London, 2010.

[20] M. Duarte, M. Davenport, D. Takhar, J. Laska, T. Sun, K. Kelly, R. Baraniuk, *„Single-pixel imaging via compressive sampling“*, IEEE Signal Processing Mag., 25(2):838211;91, 2008.

Lista korištenih skraćenica

CCD - Charge-coupled device
CMOS - Complementary metal–oxide–semiconductor
CoSaMP - Compressive Sampling Matching Pursuit
CS – uzorkovanje signala sa mogućnošću kompresije (eng. *compressive sensing)*
DCT - diskretna kosinusna transformacija
DMD - Digitalni mikroogledalni uređaj (eng. *micromirror device)*
EEG - Elektroencefalografija (eng. *electroencephalography)*
FOCUSS - *Focal Underdetermined System Solution* algoritam
FPC - *fixed-point continuation* algoritam
LCD - Liquid crystal display
LP - linearni program
MEG - Magnetoencefalografija (eng. *magnetoencephalography*,
MP - Matching Pursuit metoda rekonstrukcije
MRI - Magnetna rezonanca (eng. *Magnetic Resonance Imaging)*
NSP - svojstvo nul-prostora (eng. *null space propety)*
OMP - Orthogonal Matching Pursuit
ONB – ortonormirana baza vektorskog prostora
RIP - svojstvo ograničene izometrije (eng. *restricted isometry property)*
SLM - prostorni modulator svjetlosti (eng. *spatial light modulator*)
SNR - odnos signal-šum (eng. *signal-to-noise ratio*)
SOCP - konusni problem drugog reda (eng. *second-order cone program,*