

# Izvještaj za drugi ciklus iz predmeta DOG

---

*Domaća zadaća za drugi ciklus:*

*Domaća zadaća drugog ciklusa je izvještaj o obavljenom samostalnom radu na preostalih sedam vježbi iz udžbenika „Uvod u digitalnu obradbu govora“. To su vježbe::*

8.	Autokorelacijski postupak izračunavanja prediktora .....	2
9.	Levinson – Durbinov algoritam i analiza dobitka predikcijskog kodiranja .....	9
10.	Kovarijantni postupak linearne predikcije .....	11
11.	Utjecaj kvantizacije koeficijenata LPC filtra .....	19
12.	Homomorfna analiza govornog signala .....	23
13.	Određivanje zvučnosti i osnovne frekvencije titranja glasnica .....	38
14.	Primjer kompletnog koda govornog signala .....	40

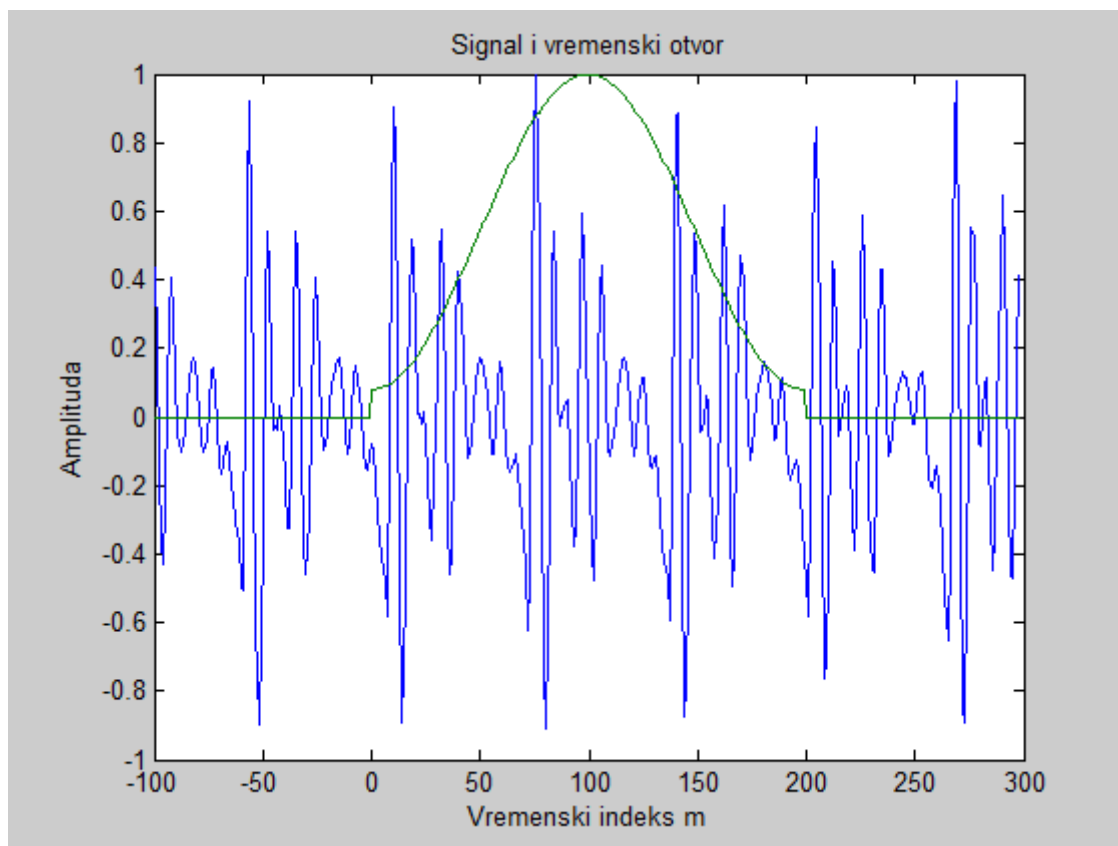
## 8. Autokorelacijski postupak izračunavanja prediktora

### Zadatak 1.

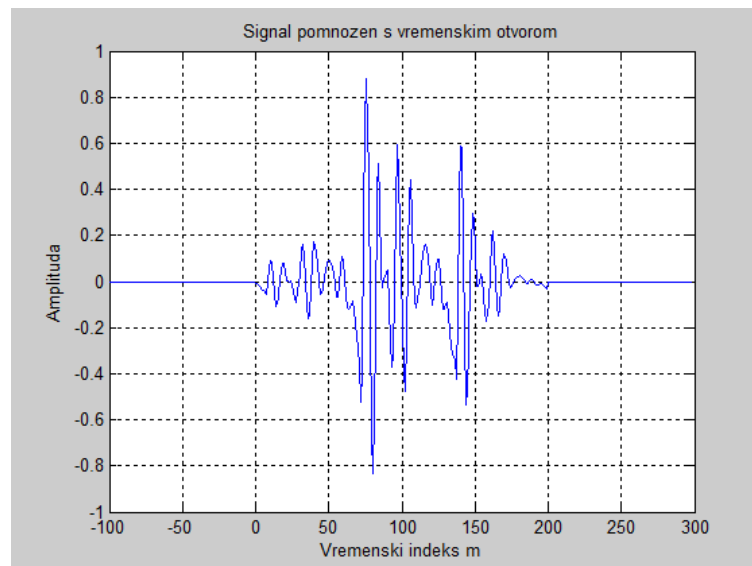
Koristeći prediktor desetoga reda ( $P=10$ ) kakav je korišten u praktičnom primjeru ovog poglavlja, ponoviti postupak za neki drugi samoglasnik. Rezultate koje Matlab ispisuje u komandni prozor možete snimiti tako da prije pokretanja programa `dog3.m` zadate naredbu `diary`.

Kolika je sumarna kvadratna pogreška predikcije u tom slučaju?

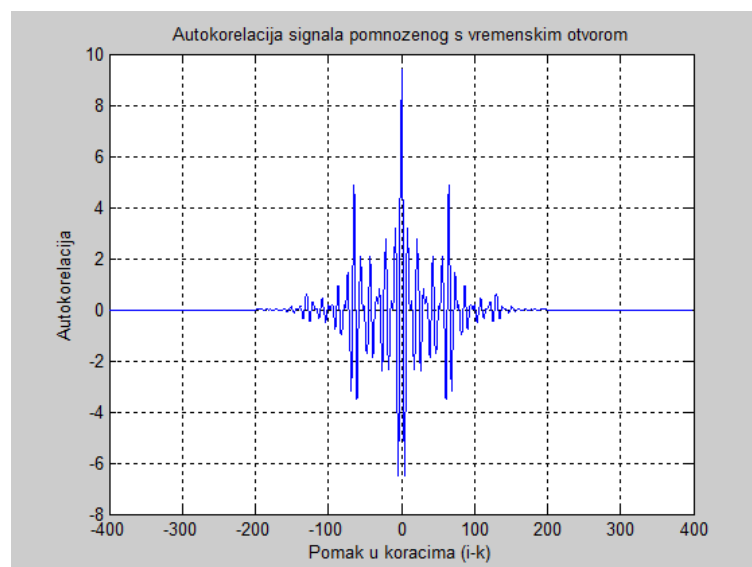
Sumarna kvadratna pogreska predikcije = 0.714273



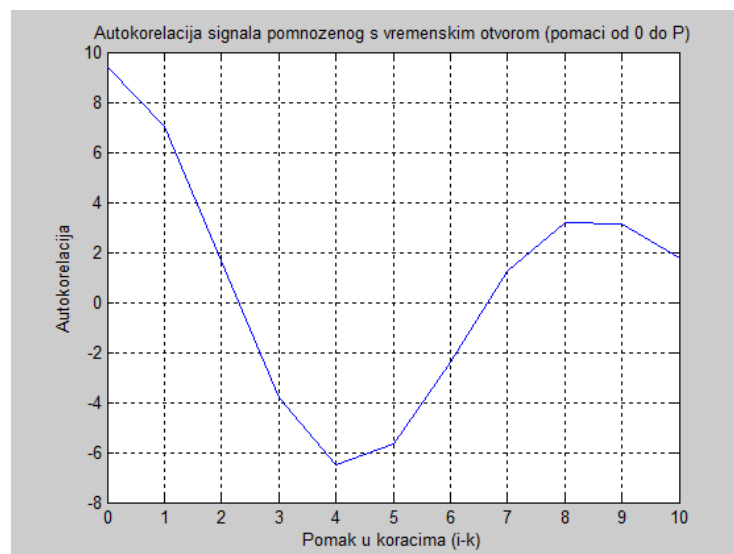
8.1 Prikaz signala glasa "a" i vremenskog otvora

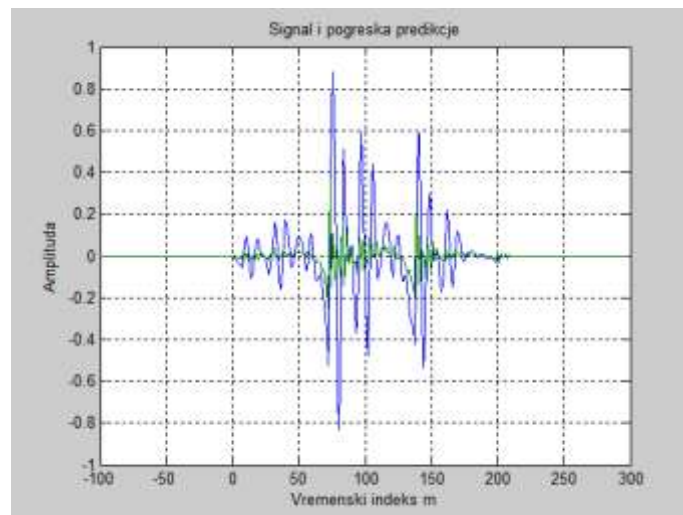


## 8.2 Signal glasa "a" pomnožen s vremenskim otvorom

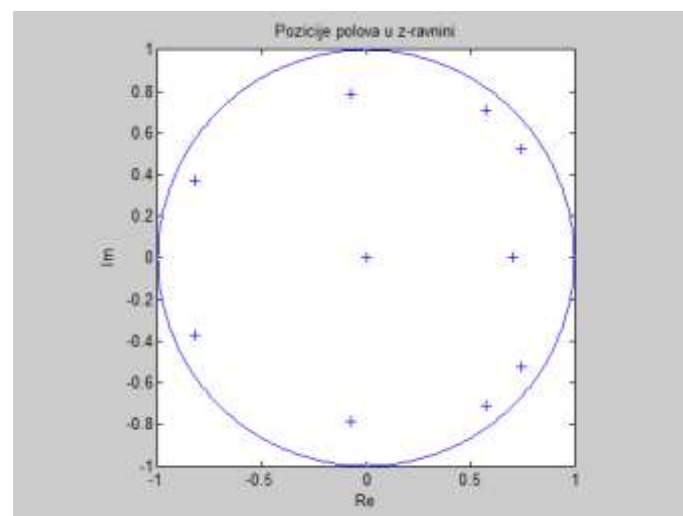


## 8.3 Izračunata autokorelacija signala pomnožena istim vremenskim otvorom

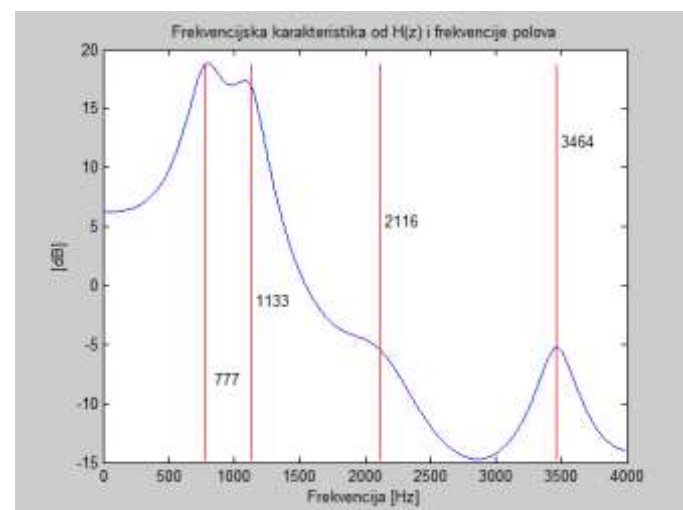




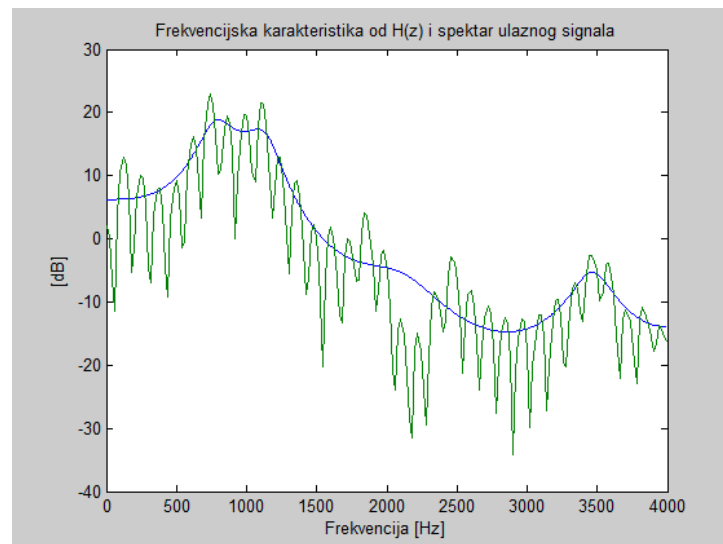
8.5 Prikaz signala(plava) i predikcijske pogreške(zelena)



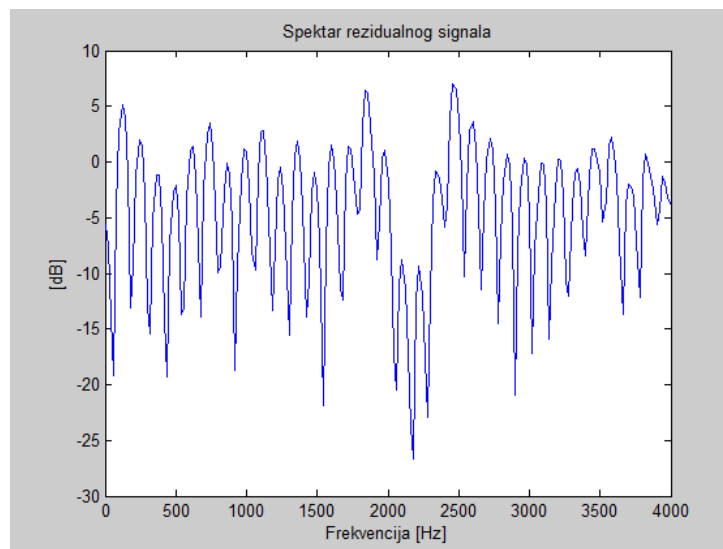
8.6 Položaj polova u Z ravnini (polova ima koliki je i red prediktora, 10)



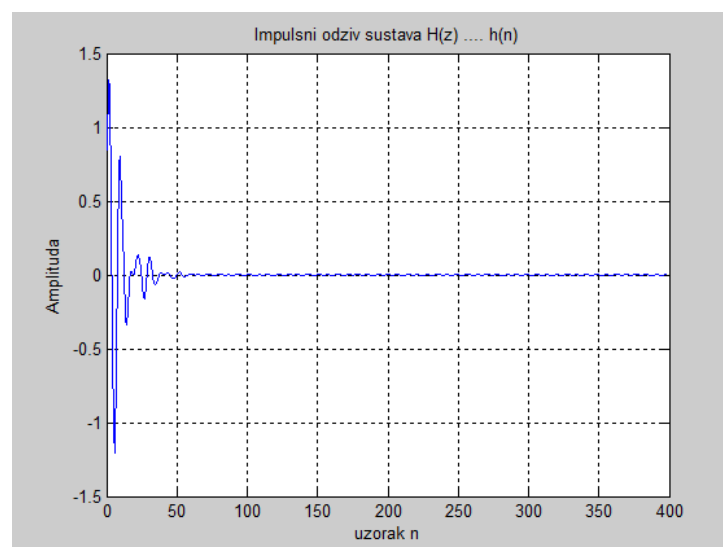
8.7 Prikaz frekvencijske karakteristike prijenosne funkcije i frekvencije polova



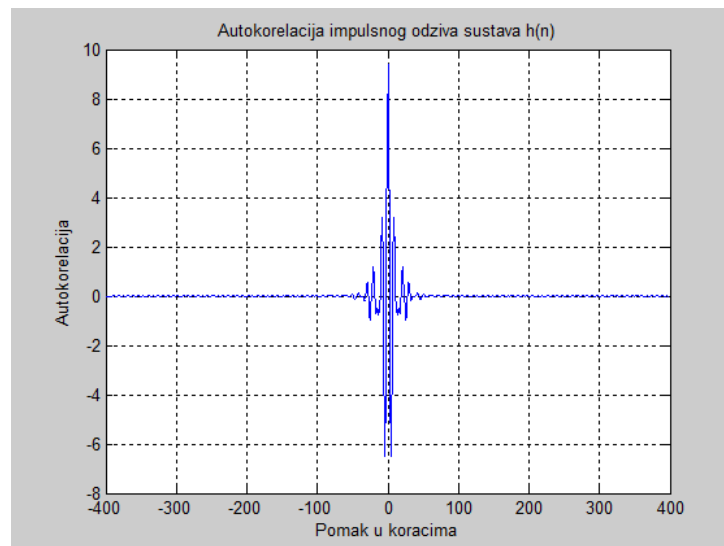
8.8 Uz frekvencijsku karakteristiku prijenosne funkcije(plavo) prikazan je i spektar ulaznog signala (zeleno)



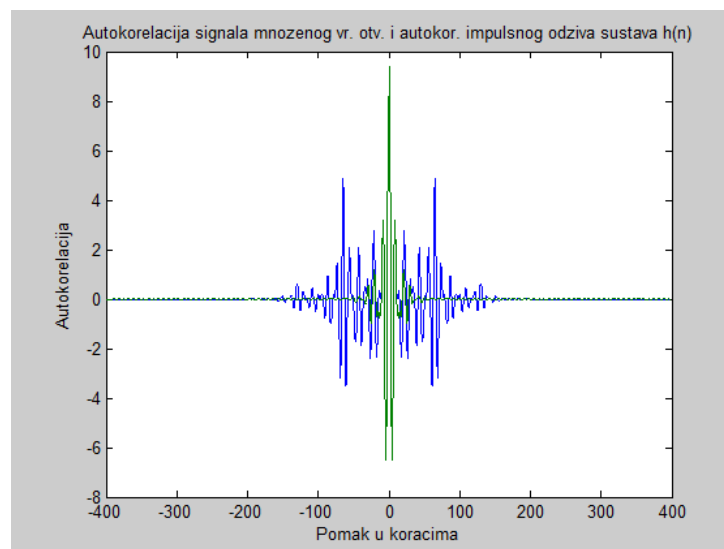
8.9 Prikaz spektra rezidualnog signala



8.10 Impulsni odziv našeg sustava



8.11 Prikaz autokorelacije impulsnog odziva našeg sustava



8.12 Prikaz autokorelacije signala pomnoženog vremenskim otvorom(plavo) i autokorelacije impuslog odziva sustava (zeleno)

### Zadatak 2.

Ponoviti postupak iz praktičnog primjera sa samoglasnikom *e*, ali za neki drugi red prediktora (npr.  $P=4$ ,  $P=6$ ,  $P=14$ ,  $P=20$ ) i ustanoviti kakav utjecaj red prediktora ima na sumarnu kvadratnu pogrešku predikcije.

Dobiveni rezultati za glas „e“:

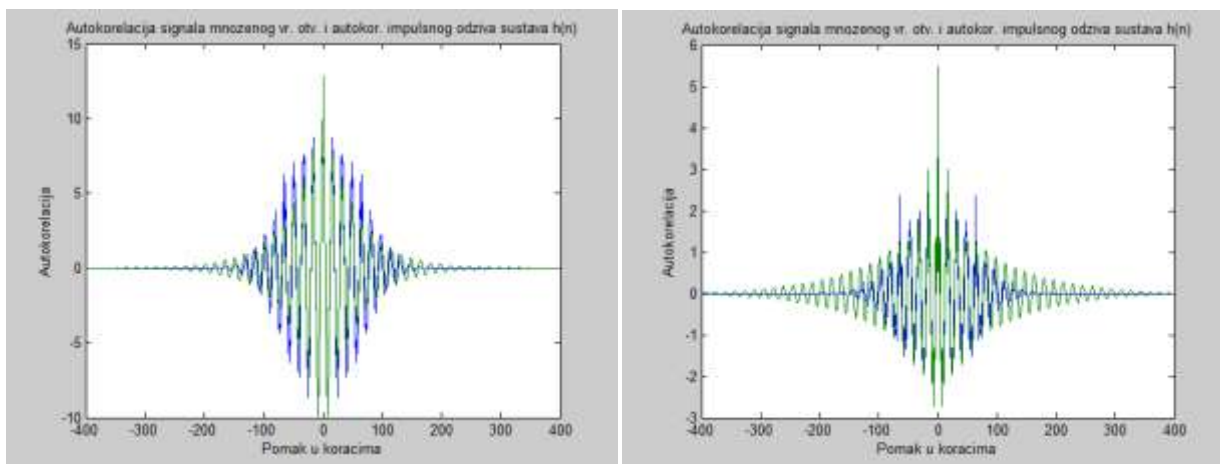
- $P=4$ ; Sumarna kvadratna pogreska predikcije = 3.679071
- $P=6$ ; Sumarna kvadratna pogreska predikcije = 1.844881
- $P=14$ ; Sumarna kvadratna pogreska predikcije = 1.553142
- $P=20$ ; Sumarna kvadratna pogreska predikcije = 1.507948

Očigledno povećanjem reda prediktora dobivamo manju sumarnu kvadratnu pogrešku predikcije. Također jer vidljivo kako povećanjem reda prediktora, sumarna kvadratna pogreška predikcije opada eksponencijalno.

### Zadatak 3.

Ponoviti postupak iz praktičnog primjera i za vašu snimku samoglasnika. Diskutirati razlike u frekvenciji titranja glasnica za vaše primjere u odnosu na govornu sekvencu korištenu u zajedničkom dijelu. Diskutirati učinkovitost prediktora za vaše glasove.

Najmanja pogreška za glas *e* i red prediktora  $P=8$  za koji sumarna kvadratna pogreska predikcije iznosi 1.354087, dok za primjerice za glas *o* i  $P=10$  sumarna kvadratna pogreska predikcije = 0.786576. Vrhunac autokorelacije se pojavljuje na  $\pm 64$  pomaka (za glas *e*).



8.133 Usporedba autokorelacijska signala za glas „o“ i glas „e“

Na slici 8.13 vidljivo je kako je za moje glasove prediktor relativno učinkovit, odnosno nije baš previše učinkovit, ali nije niti za odbaciti.

**Zadatak 4.**

*Za jedan odabrani glas i fiksirani red prediktora istražiti utjecaj tipa vremenskog otvora na spektar rezidualnog signala linearne predikcije. Također istražiti utjecaj visine glasa (frekvencije titranja glasnica) na oblik ovog spektra.*

Vremenski otvor utječe na način da se može dogoditi preklapanje latica spektara. Primjerice korištenjem pravokutnog otvora se povećava valovitost počnih latica što primjerice kod dubokih muških glasova može dovesti do preklapanja latica spektra otvora s laticama susjednih harmonika.

Povećanjem reda prediktora vidljivo je kako se na frekvencijska karakteristici od  $H(z)$  spektar odabranog okvira govornog signala se spektar sve više izravnavava. Sama visina glasa utječe na način da se povećanjem visine glasa povećava jasno i frekvencija te su time latice šire kod rezidualnog signala.



## 9. Levinson – Durbinov algoritam i analiza dobitka predikcijskog kodiranja

### **Zadatak 1.**

*U datoteci „samo.wav“ snimljeni su samo samoglasnici i neutralni samoglasnik glas 3. Zadatak je napraviti novu wav-datoteku u kojoj će biti snimljeni i neki drugi glasovi (nova datoteka također neka sadrži samoglasnike i glas 3 kako bi se moglo uspoređivati različite glasove iste osobe). Napraviti analizu za njih i komentirati uspješnost algoritma za pojedine glasove s obzirom na pogrešku predikcije.*

*Napomena: U dog4.m potrebno je kao argument kod naredbe wavread umjesto „samo.wav“ staviti ime nove datoteke. Osim toga, kad se ne koristi „samo.wav“, na izborniku se mora odabrati opcija 'Ostalo', dok ostale opcije ne vrijede bez korekcije indeksa okvira analize.*

*Diskutirati numeričku složenost izvedbe Levinson-Durbinovog algoritma u usporedbi s klasičnim metodama rješavanja sustava linearnih jednadžbi (npr. s metodom Gauss-Jordanove eliminacije).*

Za moju snimku samoglasnika (a,e,i,o,u, 3):

Glas „a“ : Sumarna kvadratna pogreska predikcije = 1.613042

Glas „e“ : Sumarna kvadratna pogreska predikcije = 1.247928

Glas „i“ : Sumarna kvadratna pogreska predikcije = 0.444640

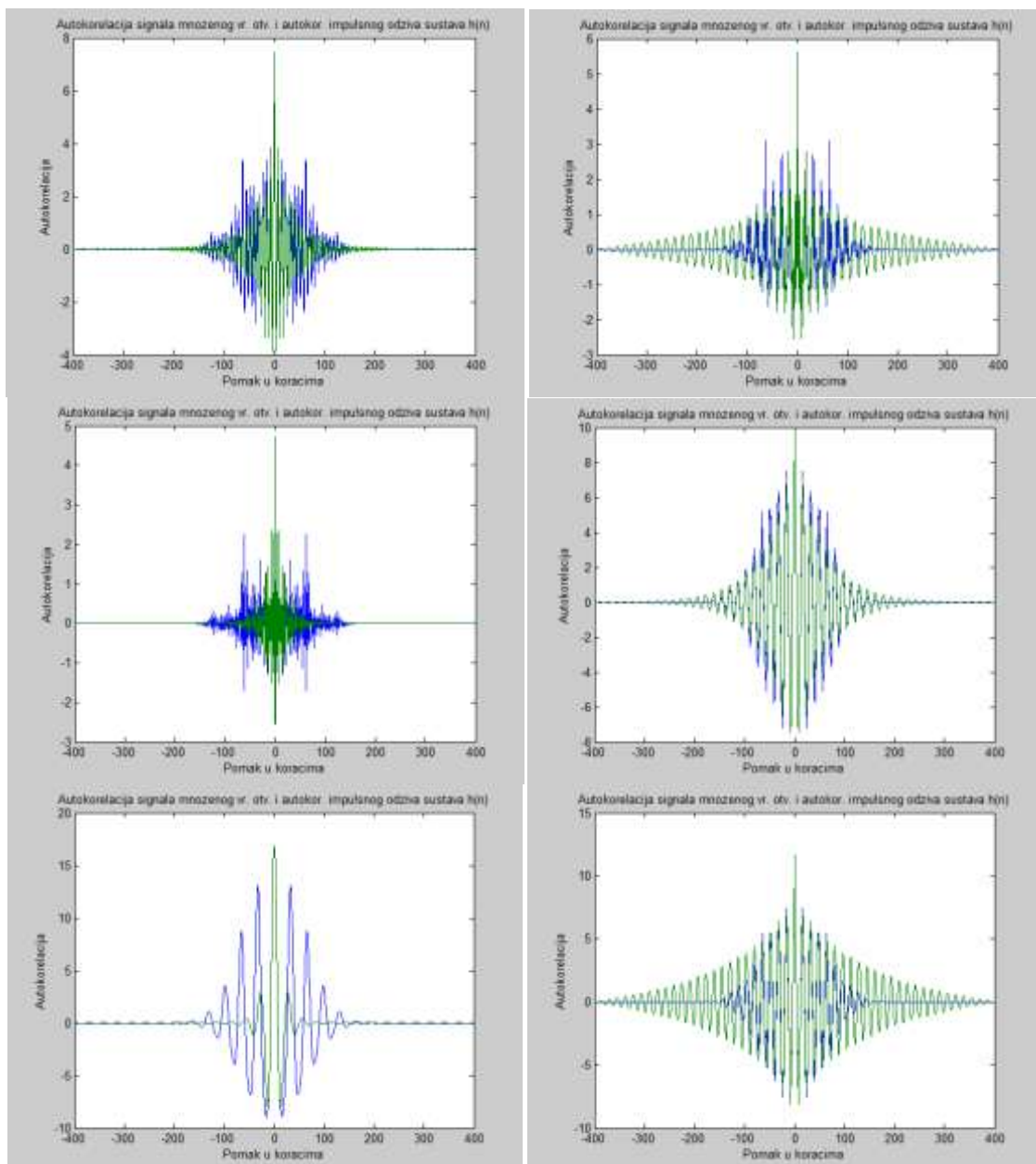
Glas „o“ : Sumarna kvadratna pogreska predikcije = 0.483119

Glas „u“ : Sumarna kvadratna pogreska predikcije = 0.383305

Glas „3“ : Sumarna kvadratna pogreska predikcije = 3.155856

Prema ovim izračunatim sumarnim kvadratnim pogreškama predikcije vidljivo je znatno poboljšanje s obzirom za prethodni zadatak. Primjerice najbolje za glas „e“ dobila se sumarna kvadratna pogreška predikcije u iznosu od 1.507948 za poprilično visok red prediktora ( $P=20$ ), dok je Levinson – Durbinovim algoritmom iznos sumarne kvadratne pogreške predikcije 1.247928, ali to za duplo manji red prediktora ( $P=10$ ). Dakle ovom metodom moguće je poboljšati predikciju, odnosno umanjiti predikcijsku pogrešku bez povećanja reda prediktora.

Što se tiče složenosti izvedbe Levinson-Durbinovog algoritma u usporedbi s klasičnim metodama rješavanja sustava linearnih jednadžbi, primjerice s metodom Gauss-Jordanove eliminacije, složenost Levinson-Durbinovog algoritma je složenosti  $O(n^2)$  dok složenost Gauss-Jordanove eliminacije je  $O(n^3)$  za matrice  $n \times n$ . Jedini nedostatak Levinson-Durbinovog algoritma je naravno taj što se on može primjenjivati samo na matricama Toeplitz strukture.



9.14 Slika izračuna autokorelacije signala za pojedina samoglasnike (a,e,i,o,u, ɜ)

## 10. Kovarijantni postupak linearne predikcije

### Zadatak 1.

- a) Nadopuni Matlab funkciju „cholesky\_dado.m“ za izračunavanje koeficijenata predikcije  $\alpha$  metodom Cholesky dekompozicije, te verificiraj pravilnost rada korištenjem klasične inverzije matrice  $f_i$ .

Cijela funkcija zajedno s nadodanim dijelovima (valjda):

```
function [al,y,d]=cholesky_dado(fi,psi)

p=length(psi);
if max(abs(size(fi)-[p p])),
    error('Wrong dimensions');
end;

% Elementi dijagonalne matrice D
d=zeros(p,1);
% Pomocna matrica V
v=eye(p);

% Nadji prvi element od D(1,1) i prvi stupac matrice V (ispod dijagonale)
d(1)=fi(1,1);
v(2:p,1)=fi(2:p,1)/d(1);

% Za ostale elemente D(k,k) i ostale stupce matrice V
for k=2:p,
    i=[k+1:p]; % indeksi pod-dijagonalnih redaka
    j=[1:k-1];
    d(k)=fi(k,k)-(v(k,j).^2)*d(j); % nadji D(k,k)
    v(i,k)=(fi(i,k)-v(i,j)*diag(d(j))*v(k,j)')/d(k); % i V(i,k)
end;

% Nadjimo pomocno rjesenje y supstitucijom unaprijed
y=zeros(p,1);
y(1)=psi(1); % prvi clan
for i=2:p,
    y(i)=psi(i) - v(i,1:i-1)*y(1:i-1); % nadodano
end;

% Podijelimo y sa dijagonalnim elementima od D
y2=y./d;

% Nadjimo konacno rjesenje al supstitucijom unazad
al=zeros(p,1);
al(p)=y2(p); % zadnji clan
for i=p-1:-1:1,
    al(i)=y2(i)-v(i+1:p,i)*al(i+1:p); % nadodano
end;
```

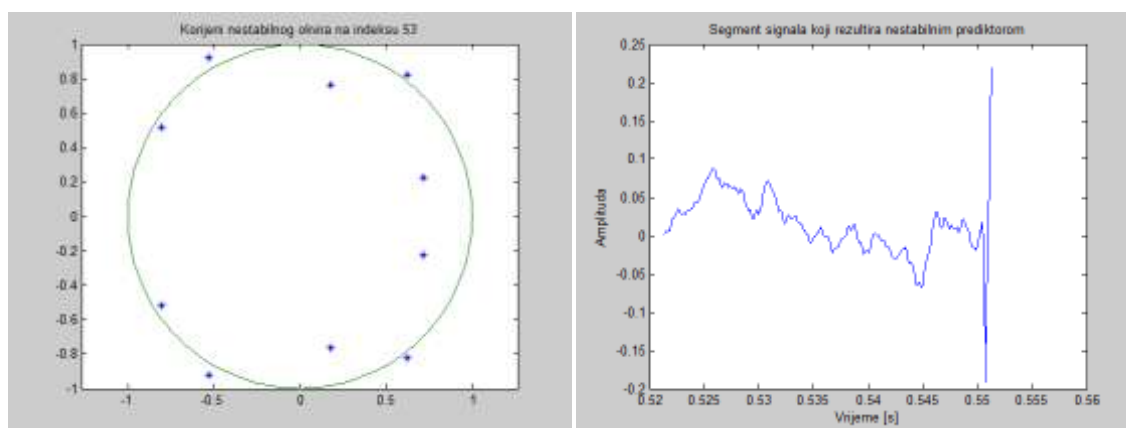
- b) Pomoću programa „nestab“ prebrojiti sve nestabilne okvire, tj. odrediti postotak nestabilnih LPC okvira za zvučne datoteke koje se nalaze u radnom folderu poglavlja.

Datoteka	Broj nestabilnih okvira i njihov postotak
<b>drago1.wav</b>	Ukupni broj nestabilnih okvira: <b>1</b> , što iznosi <b>0.28 %</b> nestabilnih okvira
<b>drago2.wav</b>	Ukupni broj nestabilnih okvira: <b>2</b> , što iznosi <b>0.54 %</b> nestabilnih okvira
<b>goran1.wav</b>	Ukupni broj nestabilnih okvira: <b>2</b> , što iznosi <b>0.47 %</b> nestabilnih okvira
<b>mama1.wav</b>	Ukupni broj nestabilnih okvira: <b>3</b> , što iznosi <b>1.53 %</b> nestabilnih okvira
<b>mama2.wav</b>	Ukupni broj nestabilnih okvira: <b>3</b> , što iznosi <b>0.73 %</b> nestabilnih okvira
<b>samo.wav</b>	Ukupni broj nestabilnih okvira: <b>4</b> , što iznosi <b>0.77 %</b> nestabilnih okvira

- c) Grafički prikazati nestabilne okvire analize s najvećim odstupanjem van jedinične kružnice za svaku zvučnu datoteku koja ima takve nestabilne okvire. Prikazati položaj polova u z-ravnini i valni oblik signala tih nestabilnih okvira. Kako spol govornika utječe na vjerojatnost pojave nestabilnih okvira?

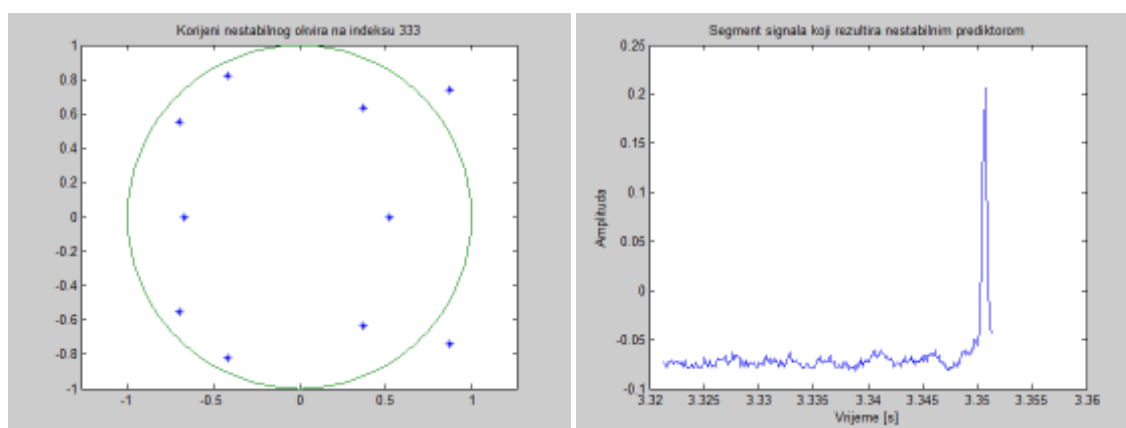
U nastavku slijedi grafički prikaz pojedinih korijena nestabilnih okvira te prikaz segmenata koji rezultiraju nestabilnim prediktorom.

drago1.wav:



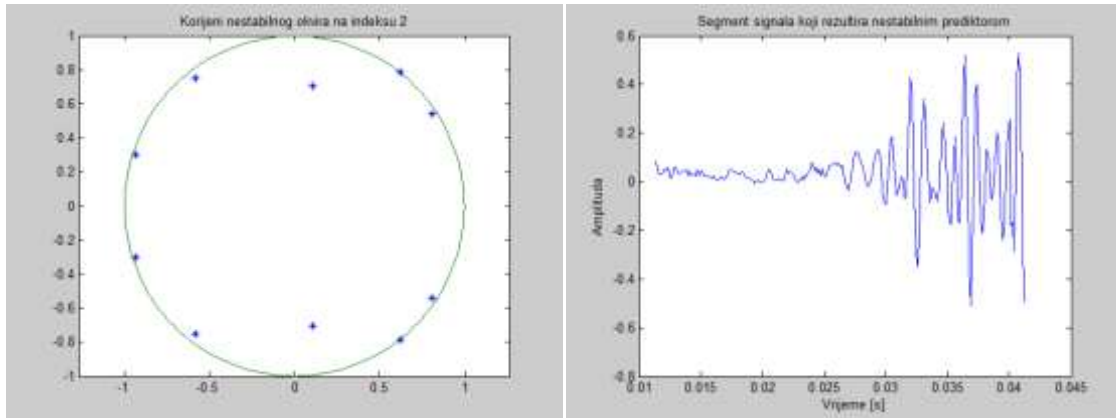
10.1 Korijeni nestabilnog okvira s najvećim odstupanjem(lijevo) i segment signala koji rezultira nestabilnim prediktorom(desno)

drago2.wav:



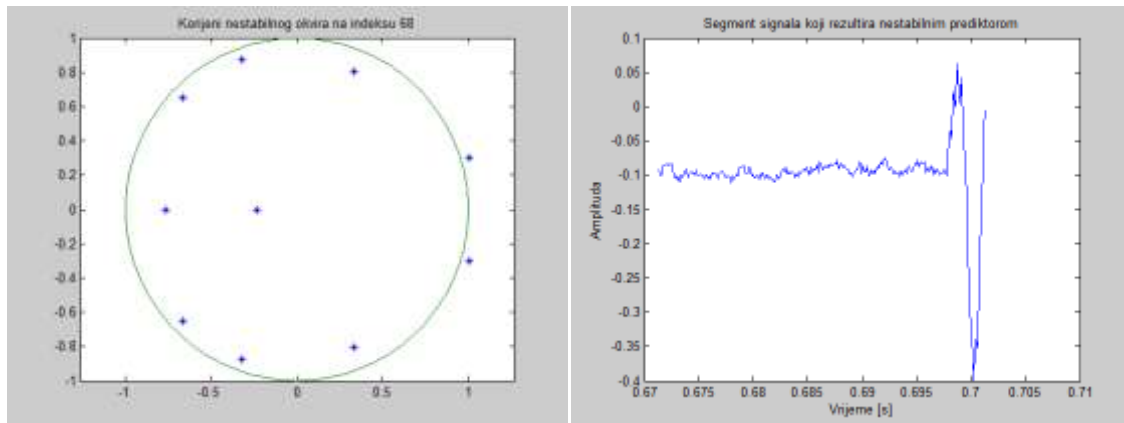
10.2 Korijeni nestabilnog okvira s najvećim odstupanjem(lijevo) i segment signala koji rezultira nestabilnim prediktorom(desno)

goran1.wav:



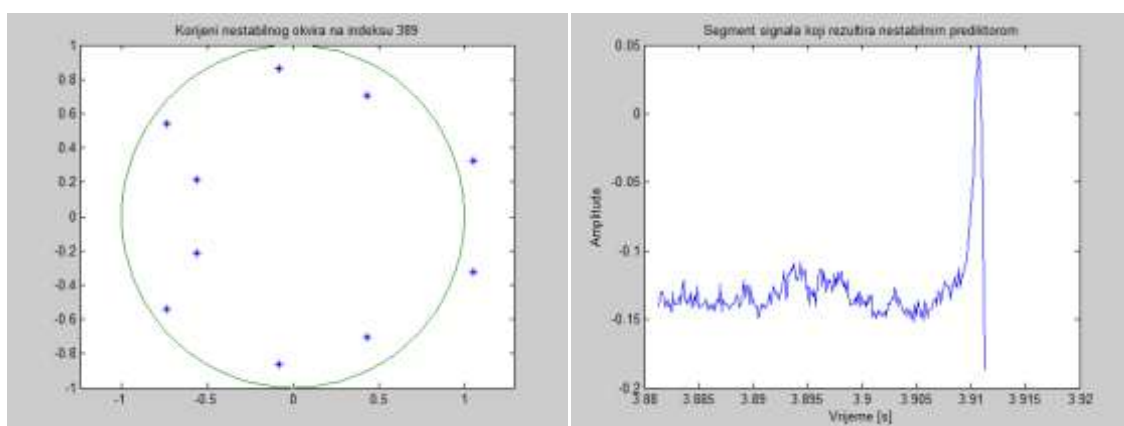
10.3 Korijeni nestabilnog okvira s najvećim odstupanjem(lijevo) i segment signala koji rezultira nestabilnim prediktorom(desno)

mama1.wav:



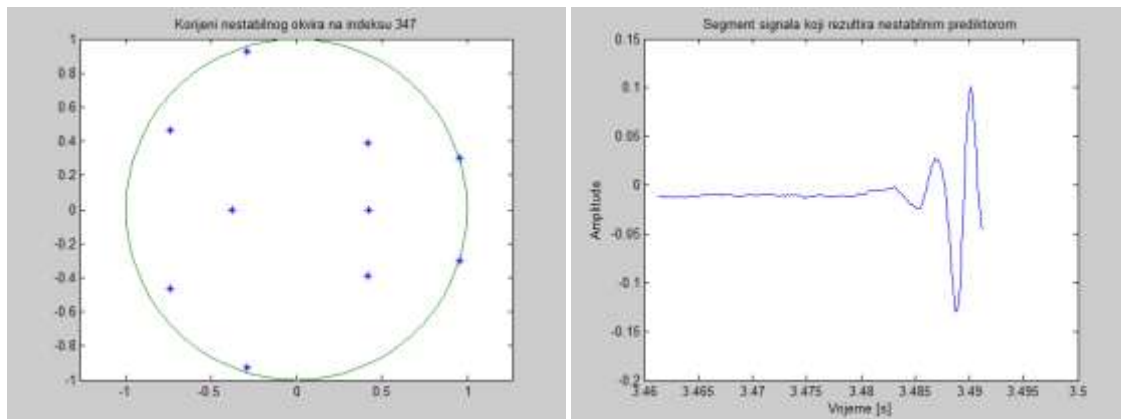
10.4 Korijeni nestabilnog okvira s najvećim odstupanjem(lijevo) i segment signala koji rezultira nestabilnim prediktorom(desno)

mama2.wav:



10.5 Korijeni nestabilnog okvira s najvećim odstupanjem(lijevo) i segment signala koji rezultira nestabilnim prediktorom(desno)

samo.wav:



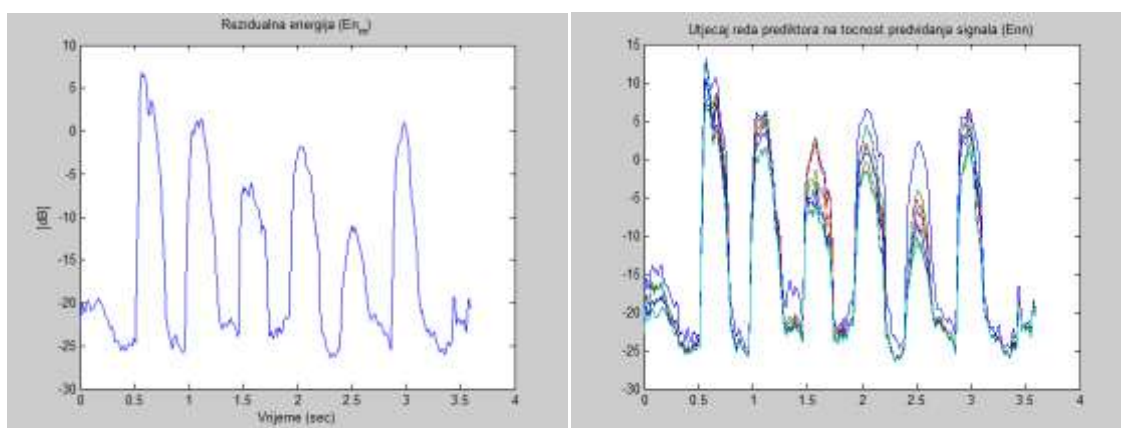
#### 10.6 Korijeni nestabilnog okvira s najvećim odstupanjem(lijevo) i segment signala koji rezultira nestabilnim prediktorom(desno)

Po rezultatima koji se nalaze u donjoj tablici rekao bih da spol pojedinog glasa ne utječe na pojavu nestabilnosti sustava.

Datoteka	Broj okvira s najvećim odstupanjem
drago1.wav	53
drago2.wav	333
goran1.wav	2
mama1.wav	68
mama2.wav	235
samo.wav	260

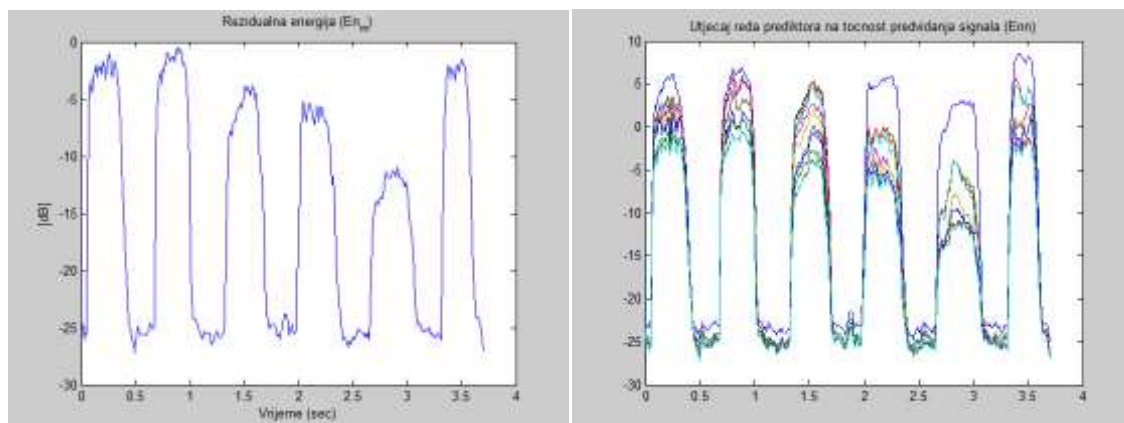
- d) Grafički prikazati ovisnost rezidualne energije o redu predikcije i rezidualnu energiju za najviši red predikcije. Prikaze napraviti u logaritamskom mjerilu, a ne u linearnom radi jasnijeg prikaza utjecaja reda prediktora.

drago1.wav:



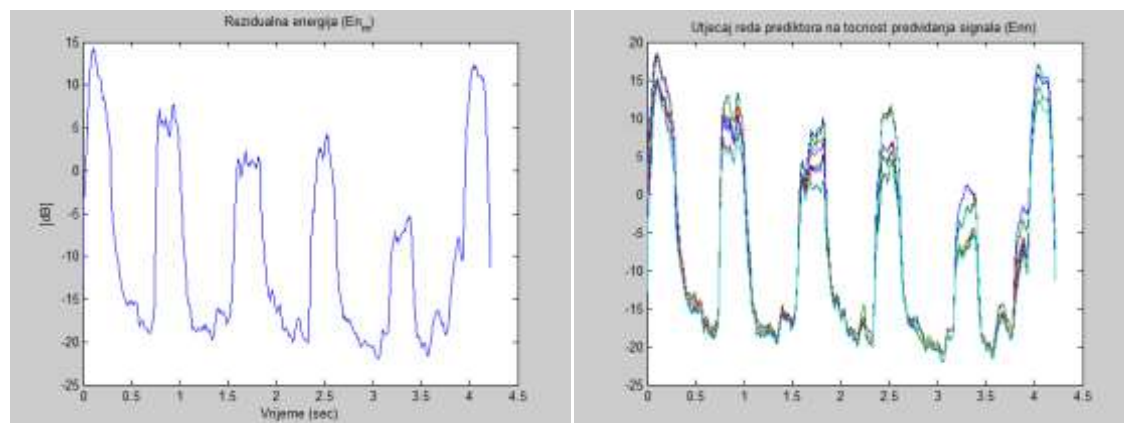
#### 10.7 Rezidualna energija ( $E_{m_n}$ ) (lijevo) i prikaz utjecaja reda prediktora na točnost predviđanja signala (desno)

drago2.wav:



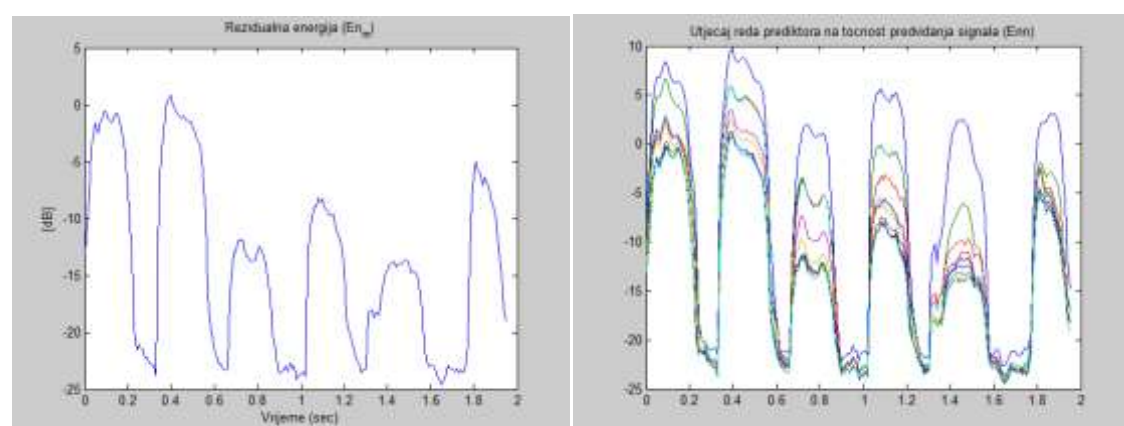
10.8 Rezidualna energija ( $E_{m_n}$ ) (lijevo) i prikaz utjecaja reda prediktora na točnost predviđanja signala (desno)

goran1.wav:



10.9 Rezidualna energija ( $E_{m_n}$ ) (lijevo) i prikaz utjecaja reda prediktora na točnost predviđanja signala (desno)

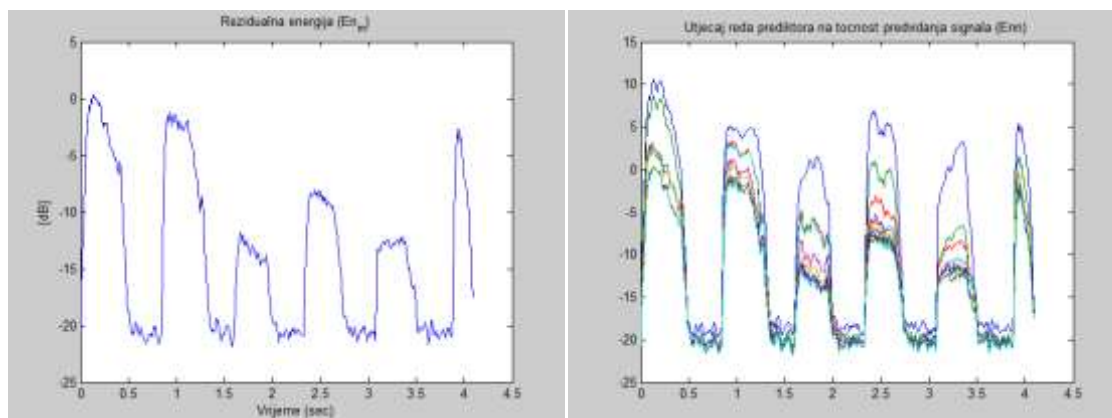
mama1.wav:



10.10 Rezidualna energija ( $E_{m_n}$ ) (lijevo) i prikaz utjecaja reda prediktora na točnost predviđanja signala (desno)

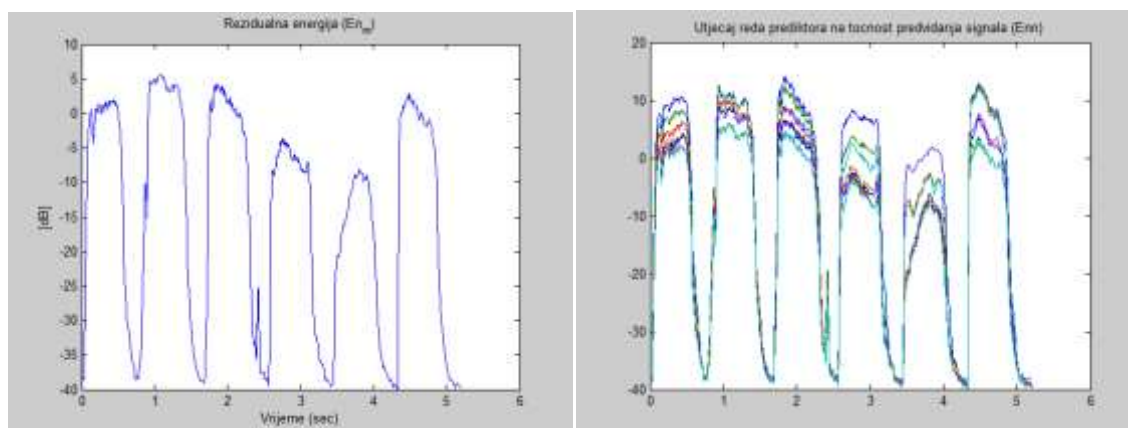


mama2.wav:



10.11 Rezidualna energija ( $E_{m_n}$ ) (lijevo) i prikaz utjecaja reda prediktora na točnost predviđanja signala (desno)

samo.wav:



10.12 Rezidualna energija ( $E_{m_n}$ ) (lijevo) i prikaz utjecaja reda prediktora na točnost predviđanja signala (desno)

- e) Izraziti numeričku složenost cijelog algoritma (ukupni broj operacija zbrajanja i množenja) za sve dijelove algoritma (izračun prvog stupca, formiranje cijelog sustava jednačbi, rješavanje sustava metodom Cholesky dekompozicije), kao funkciju parametara analize (faktor, korak,  $p$ ,  $N = \text{faktor} * \text{korak}$ ).

Proračun korelacijske matrice zahtijeva približno  $N * p$  množenja dok rješavanje sustava matrice jednačbe (Cholesky dekompozicija) zahtijeva broj množenja proporcionalan sa  $p^3$ .

Ako definiramo sljedeće:

$N$  – broj uzoraka analiziranog segmenta govora,  $p$  – red predikcije

- izračun prvog stupca operacija  $\text{korak} * (p + 1)$
- izračun  $d_i$  i  $v_{ik}$  elemenata matrice  $p * (p + 1)$  operacija
- izračun vrijednosti  $\alpha$  zahtijeva  $p * (p + 1) / 2$  operacija
- izračun  $Y$ -a  $p * (p + 1) / 2$  operacija

Tada je ukupan broj operacija :

$$\text{korak} * (p + 1) + 2 * p * (p + 1)$$



- f) Za zadanu frekvenciju otipkavanja  $f_s=8\text{kHz}$ , na osnovu gornje procjene broja operacija odrediti potrebu brzinu procesora izraženu kao MOPS (milijuna operacija u sekundi), za izvođenje algoritma u stvarnom vremenu gdje pretpostavljamo da se zbrajanje i množenje provodi svako kao jedna operacija procesora (uz ignoriranje svih drugih operacija potrebnih za stvarnu izvedbu algoritma, kao što su dohvat ili pohrane podataka). Pri tome pretpostaviti slijedeće vrijednosti parametara analize faktor=3, korak=80,  $N=240$ ,  $p=10$ .

$$\text{korak} * (p + 1) + 2 * p * (p + 1) = 80 * 11 + 2 * 10 * 11 = 1100$$

Budući da se u sekundi provodi analiza 100 puta, potrebno je 0.11 MOPS-a.

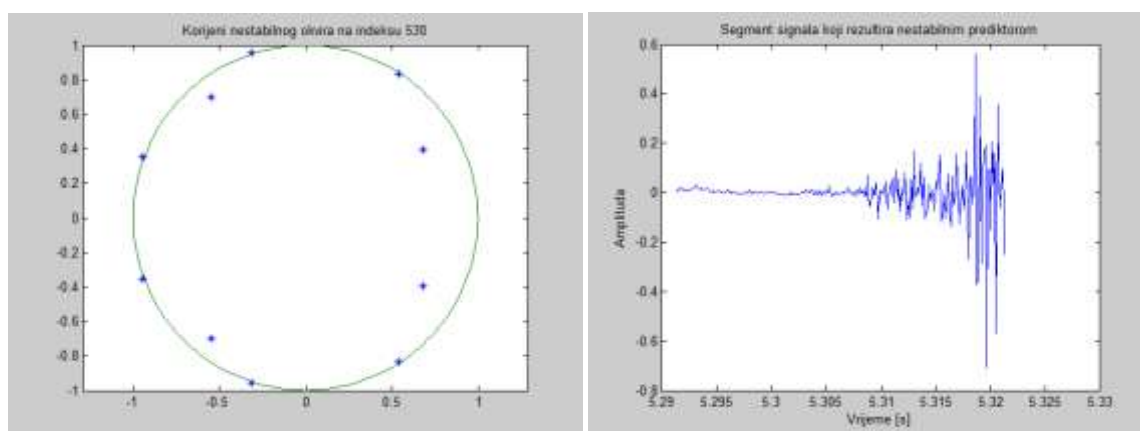
## Zadatak 2.

- g) Ponoviti analizu iz zadatka 1b) do 1d) za vlastite zvučne datoteke i usporediti rezultate.

1b) Ukupni broj nestabilnih okvira: 6, što iznosi 0.93 posto nestabilnih okvira

Za moju snimku je broj nestabilnih okvira veći od zadanih snimki. Razlogu tomu vjerojatno leži i u kvaliteti same snimke koja i nije baš reprezentativna.

1c) U nastavku slijedi grafički prikaz pojedinih korijena nestabilnih okvira te prikaz segmenata koji rezultiraju nestabilnim prediktorom.



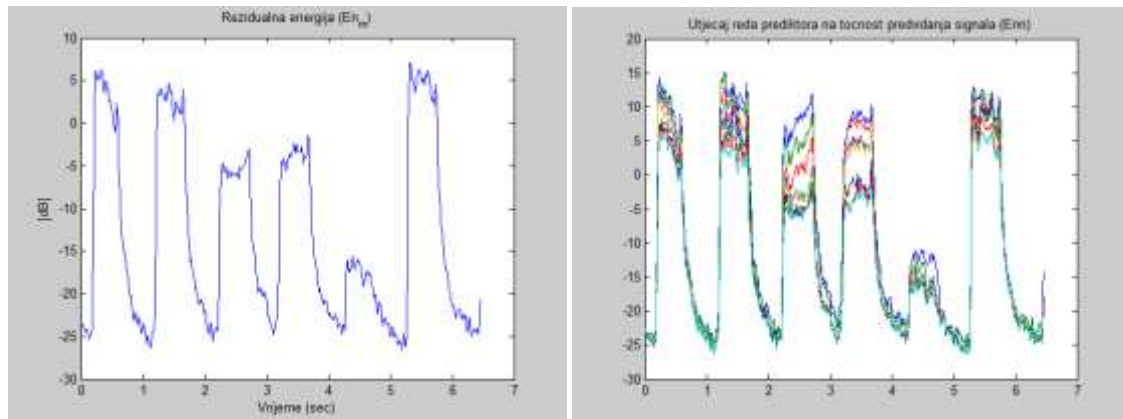
10.13 Korijeni nestabilnog okvira s najvećim odstupanjem(lijevo) i segment signala koji rezultira nestabilnim prediktorom(desno)

Kao i kod prethodnog zadatka vidljivo je kako se nestabilnost pojavljuje pred kraj.

1d)

Prikaz rezidualne energije ( $Em_n$ ) i prikaz utjecaja reda prediktora na točnost predviđanja signala za vlastitu snimku.

samoglasnici.wav:



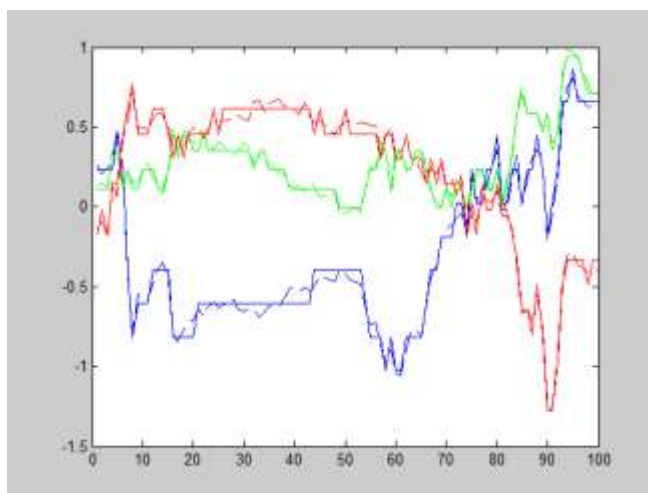
10.14 Rezidualna energija ( $Em_n$ ) (lijevo) i prikaz utjecaja reda prediktora na točnost predviđanja signala (desno)

## 11. Utjecaj kvantizacije koeficijenata LPC filtra

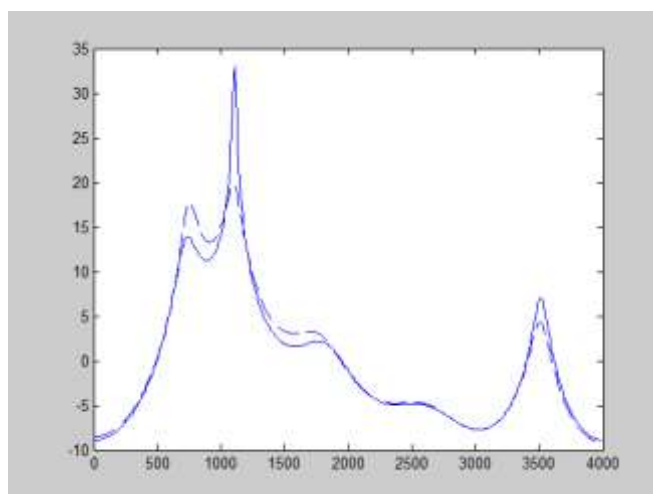
### Zadatak za samostalni rad 1.

a) Napiši Matlab program za određivanje  $SD^2(okv)$ ,  
✓ Naziv datoteke: „pog11\_zad1.m“

b) provjeri rezultate usporedbom sa slikom 15.2-4,  
Provjera:



11.1 Prva tri a-koeficijenta LPC filtra prilikom izgovora glasa „a“ i Amplitudno frekvencijske karakteristike originalnog (crtkano)



11.2 frekvencijske karakteristike originalnog (crtkano) i kvantiziranog (puna linija) filtra za 30-ti okvir analize (sredina izgovora glasa 'a')

c) Odredi točni iznos  $SD^2 = \underline{\quad 2.3894 \quad}$

d) Odredi ovisnost  $SD^2$  o broju bita  $b$ , te je prikaži tabelarno (grafički) za  $b$  od 3 do 7

$b = 3$	$SD^2 = 5.4713$
$b = 4$	$SD^2 = 2.3894$
$b = 5$	$SD^2 = 1.6935$
$b = 6$	$SD^2 = 0.4948$
$b = 7$	$SD^2 = 0.0871$

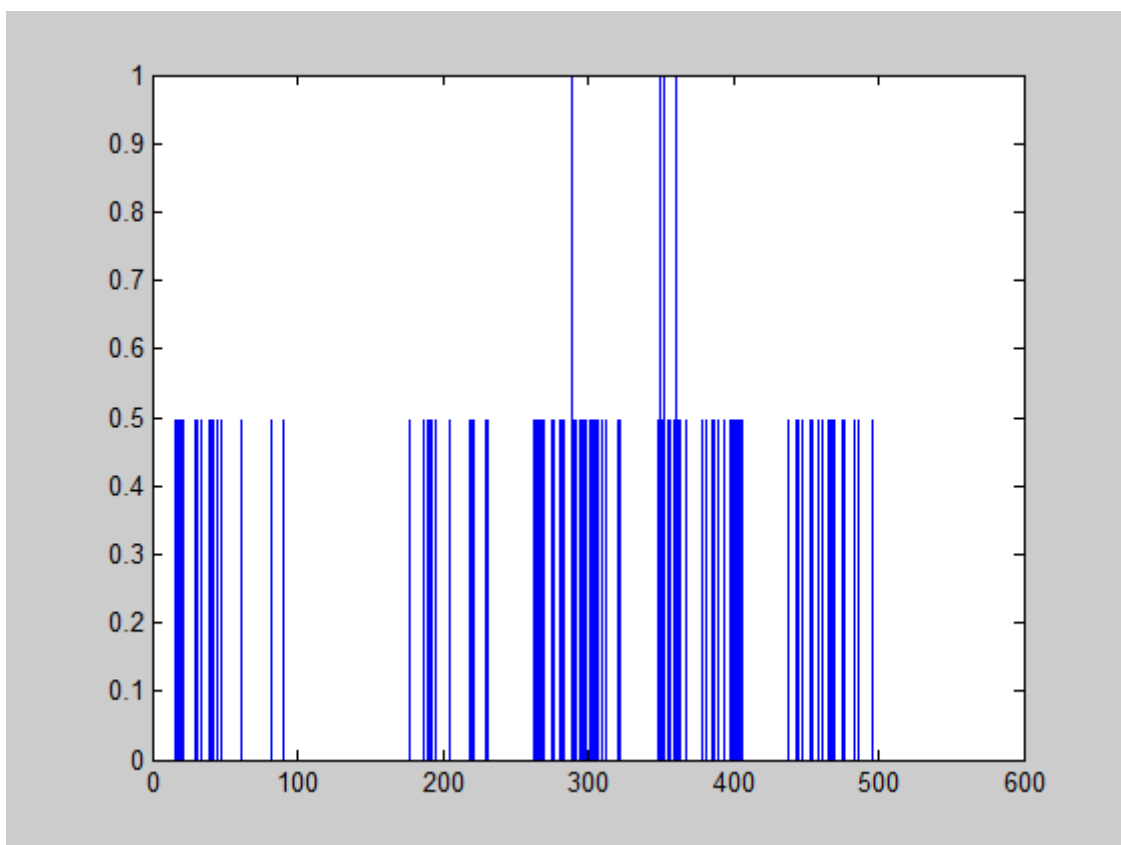
e) Pri kojem broju bita se ostvaruje uvjet na nečujnost izobličenja  $b_{min} = \underline{\underline{6}}$

### Zadatak za samostalni rad 2.

a) Opisani postupak određivanja broja nestabilnih konjugirano kompleksnih parova polova treba provesti za sve okvire analize,  $okv=1$  do  $b_f$ , tako da se u stupac  $unst$  sa  $b_f$  redaka upiše broj nestabilnih formata svakog pojedinog okvira. Napiši pripadni program u Matlabu koji to ostvaruje.

✓ Naziv datoteke: „pog11\_zad2.m“

b) Izvrši ga i grafički prikaži rezultate (uz  $b=4$ ).



11.3 Grafički prikaz rezultata uz  $b = 4$

c) Za koje glasove je broj nestabilnih okvira najvedi, a za koje najmanji ?

Kao što se može vidjeti na slici 11.3 vidljivo je kako je za glasove „o“ i „u“ najveći broj nestabilnih okvira, dok za glas „e“ je najmanji broj nestabilnih okvira.

d) Potrebno je ustanoviti postotak nestabilnih LPC okvira prema izrazu:

```
post_unst=mean(unst>0)*100;
```

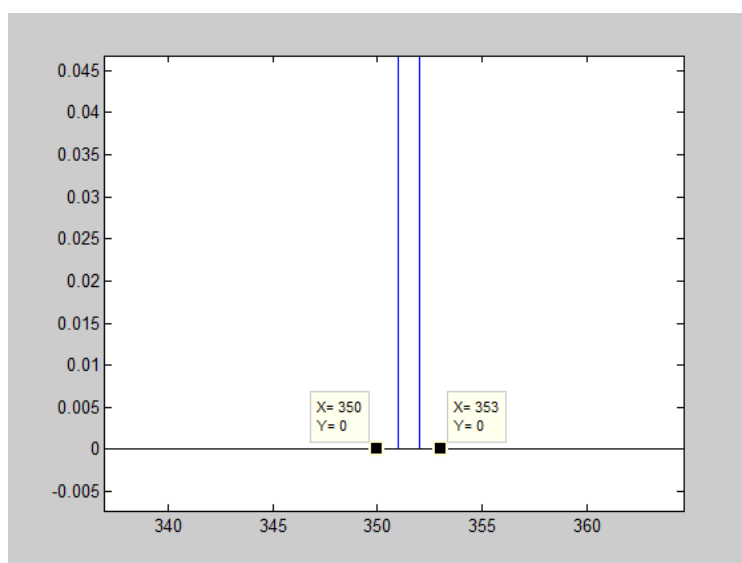
$post\_unst = 20.2677$

e) Odrediti ovisnost  $post\_unst$  o broju bita  $b$ , te je prikazati grafički i tabelarno za  $b$  od 3 do 7

Broj bitova	Postotak nestabilnih LPC okvira
3	49,7132
4	20,2677
5	6,5010
6	0,9560
7	0,3824

f) Koliko ima nestabilnih okvira uz  $b=7$ , koji su to okviri i kojem glasu pripadaju?

g)



11.4 15 Prikaz nestabilnih okvira

Uvidom na gornju sliku vidljivo je kako su smao okviri 351 i 352 nestabilni okviri. Oni pripadaju glasu slova „u“.

### Zadatak za samostalni rad 3.

Potrebno je ponoviti postupak kvantizacije koji je opisan u poglavlju 15.2, ali sada na skupu  $k$ -koeficijenata. Prvi korak je izračunavanje matrice  $k\_m$  na osnovu  $a\_m$  koja je učitana iz datoteke *samo.mat* (PAZI: u funkciju *a2k\_new* ne ulazi koeficijent  $a_0$  koji se originalno nalazi u matrici  $a\_m$ ). Ponovi zadatak 1. u cijelosti, te prikaži rezultate iz stavke 1.d) usporedno s onima za slučaj kvantizacije  $a$ -koeficijenata s istim brojem bita.

Tablica ukazuje na bolju kvalitetu kvantizacije za korištenje PARCOR koeficijenata:

Broj bitova	$SD^2$ LPC koef.	$SD^2$ PARCOR koef.
3	11,3143	6,7098
4	4,5021	1,9770
5	1,4617	0,5577
6	0,3905	0,1303
7	0,1013	0,0344

**Zadatak za samostalni rad 4.**

*Modificiraj program iz zadatka 3. za kvantizaciju na LAR koeficijentima i usporedi ove rezultate s prethodnima.*

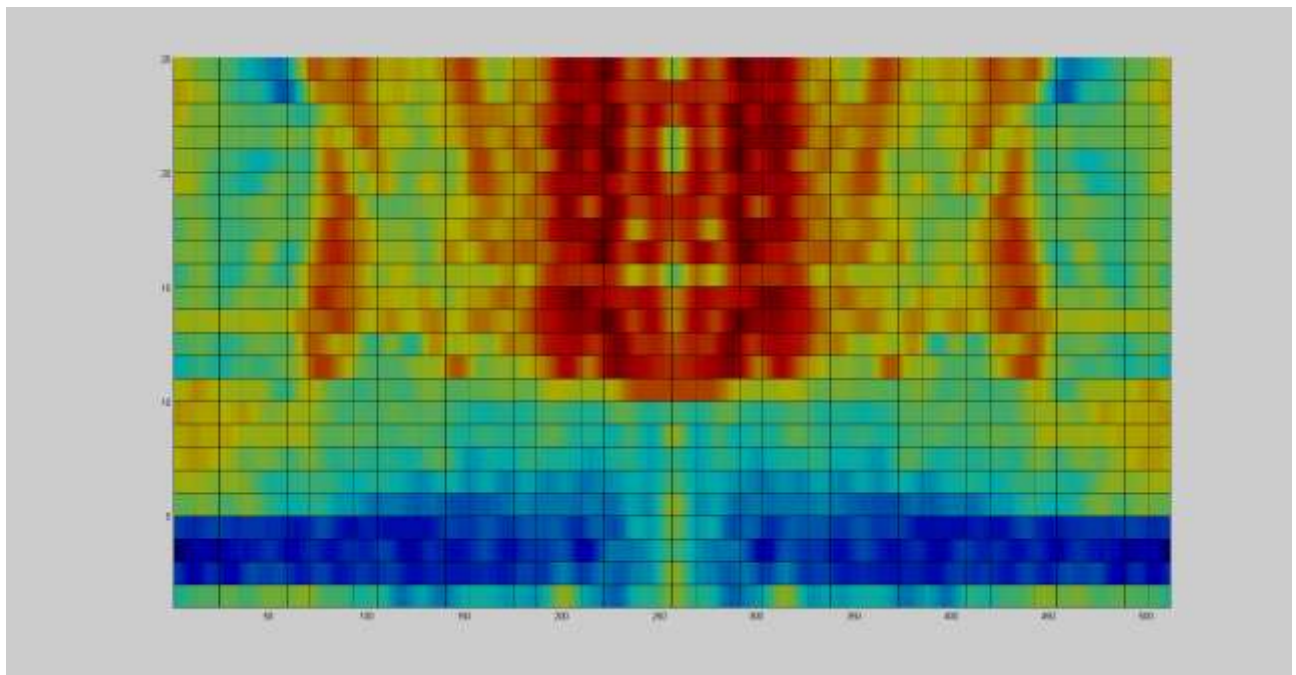
Još jedna tablica koja ukazuje na najbolju kvalitetu s LAR koeficijentima:

Broj bitova	$SD^2$ LPC koef.	$SD^2$ PARCOR koef.	$SD^2$ LAR koef.
3	11,3143	6,7098	1,1785
4	4,5021	1,9770	0,3205
5	1,4617	0,5577	0,0787
6	0,3905	0,1303	0,0210
7	0,1013	0,0344	0,0049

## 12. Homomorfna analiza govornog signala

### Zadaci za samostalno izvođenje

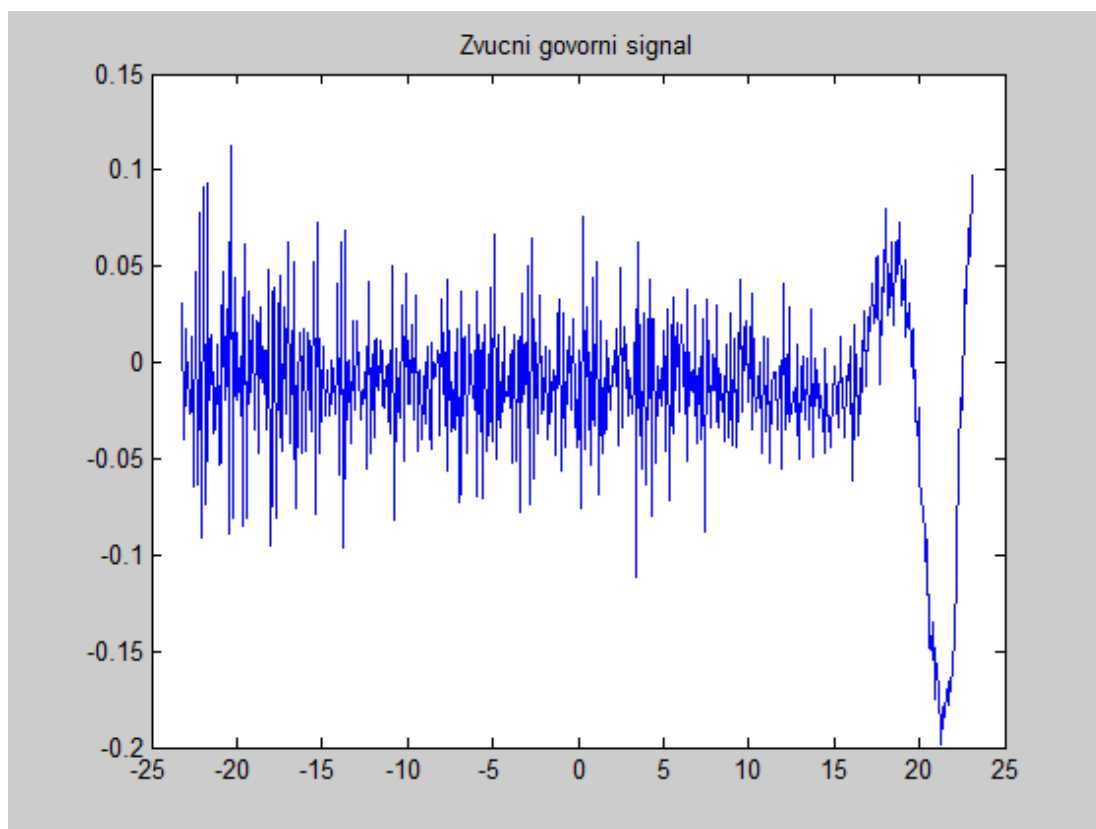
- a) *Kombiniranjem svih odsječaka Matlab programa napisati cjelokupni program za homomorfnu obradbu govora (učitavanje zvučne datoteke, odabir dijela zvučne sekvence za obradbu, prikaz svih međurezultata).*  
✓ Naziv datoteke: „pog12\_zad1.m“
- b) *Modificirati program radi provođenja analize preko svih okvira analize u učitanoj zvučnoj datoteci za odabrani korak analize, uz pohranu spektralne ovojnice svakog okvira (DFT od LTP-a) u veliku matricu.*  
✓ Naziv datoteke: „pog12\_zad2.m“
- c) *Ostvariti prikaz tako izračunate spektralne ovojnice u obliku spektrograma govora, tj. primjenom naredbe pcolor, te dobivene rezultate usporediti s rezultatima temeljenim na LPC analizi.*  
✓ Naziv datoteke: „pog12\_zad2.m“



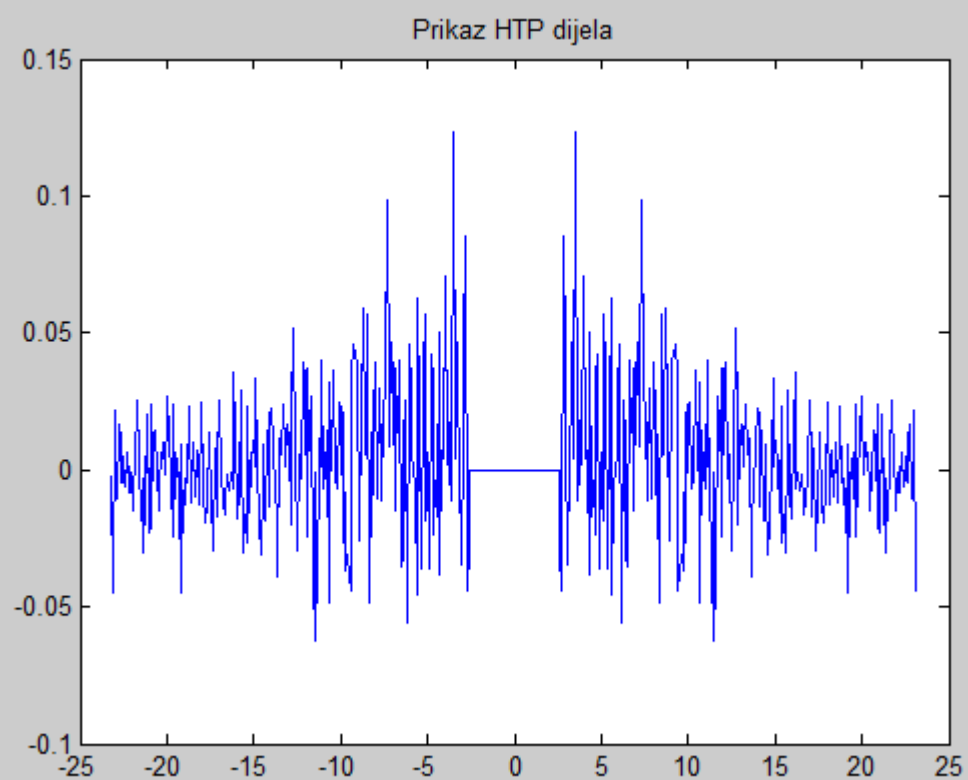
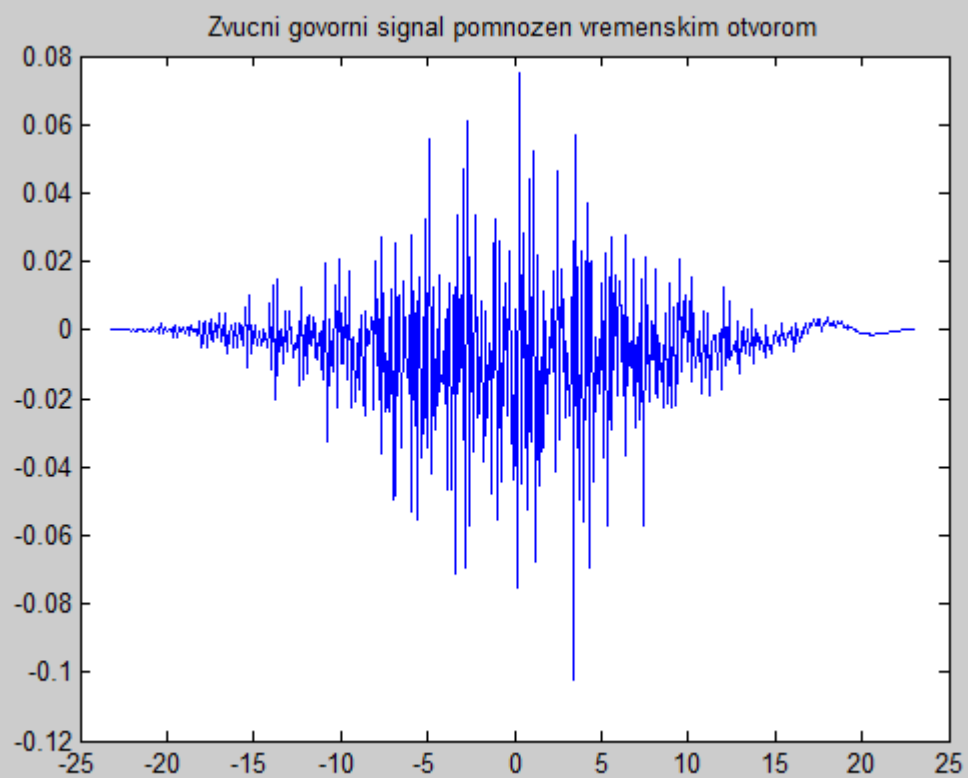
- d) Snimiti vlastitu zvučnu datoteku s istim parametrima kao „sa2.wav“, tj.  $f_s=11025\text{Hz}$ , 16bita, mono, te homomorfnu analizu provesti na njoj za primjer jednog zvučnog i jednog bezvučnog glasa.

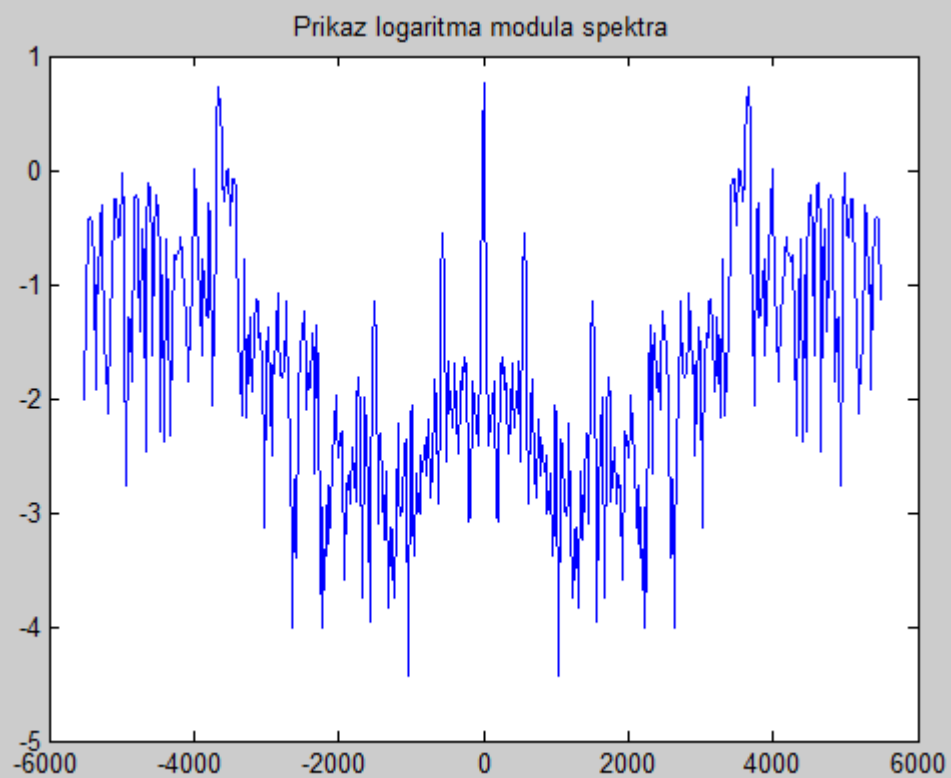
Snimka se sastoji od glasova „s“ i „a“ kao i u primjeru iz zadatka 1 a). Svaka slika ima na sebi napisan naziv što se nalazi na njoj tako da nisam pisao opise pojedinih slika.

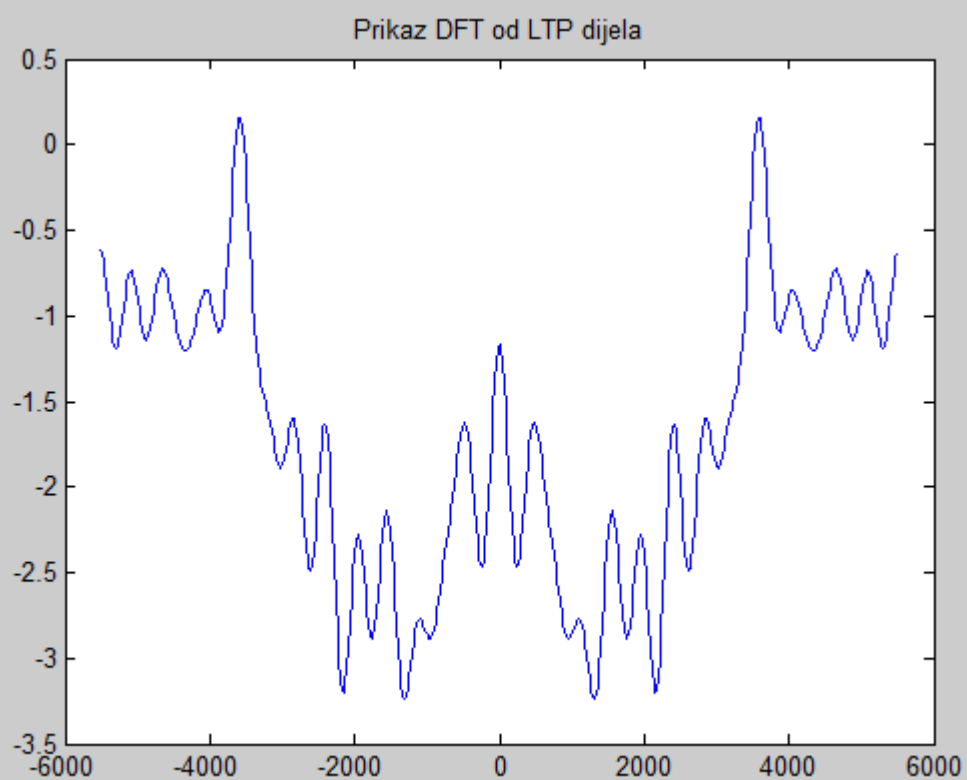
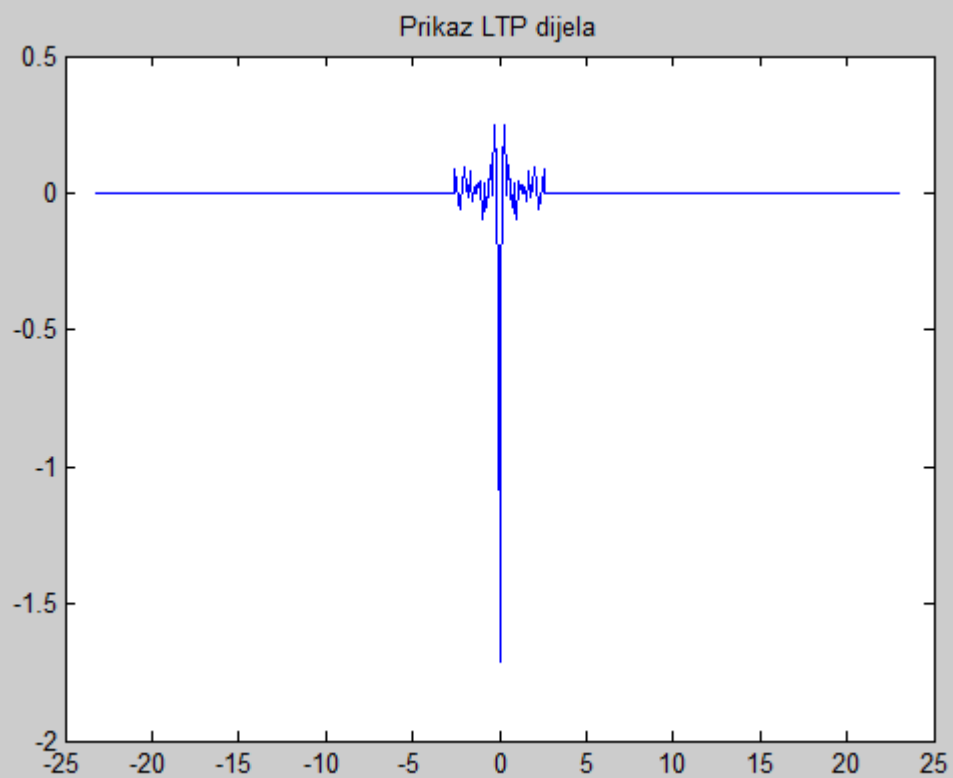
Glas „S“:

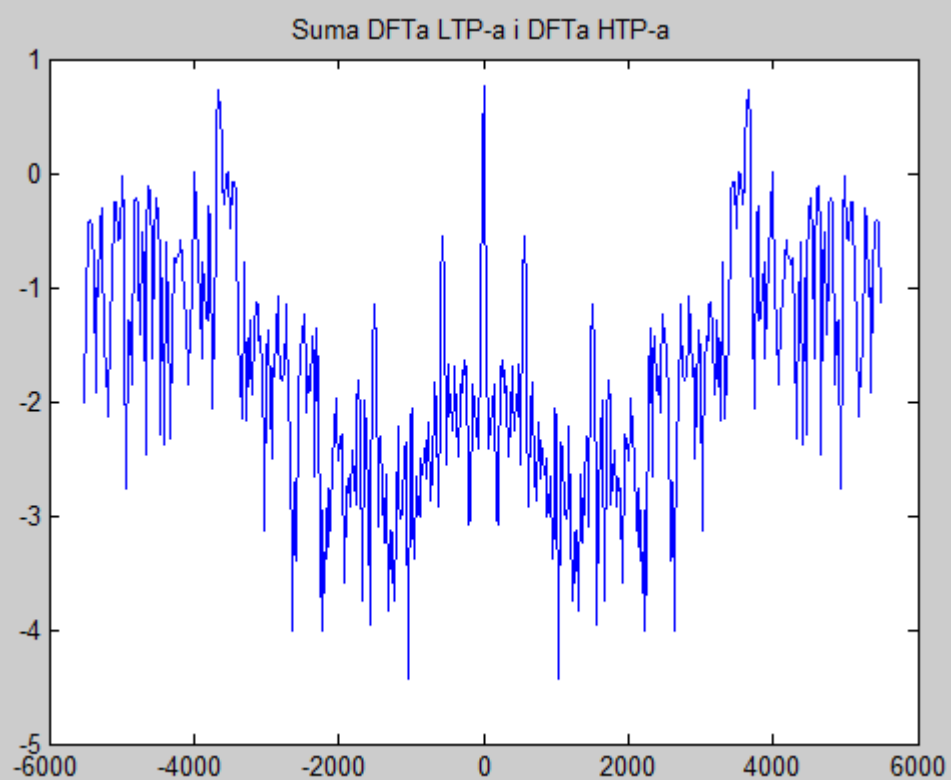
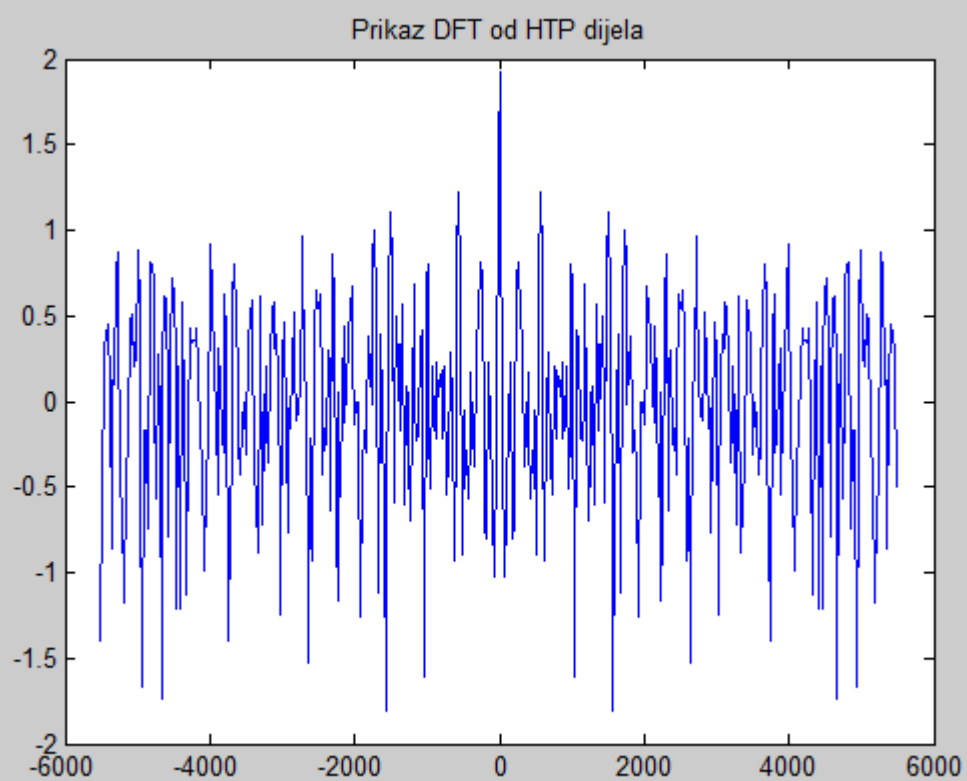


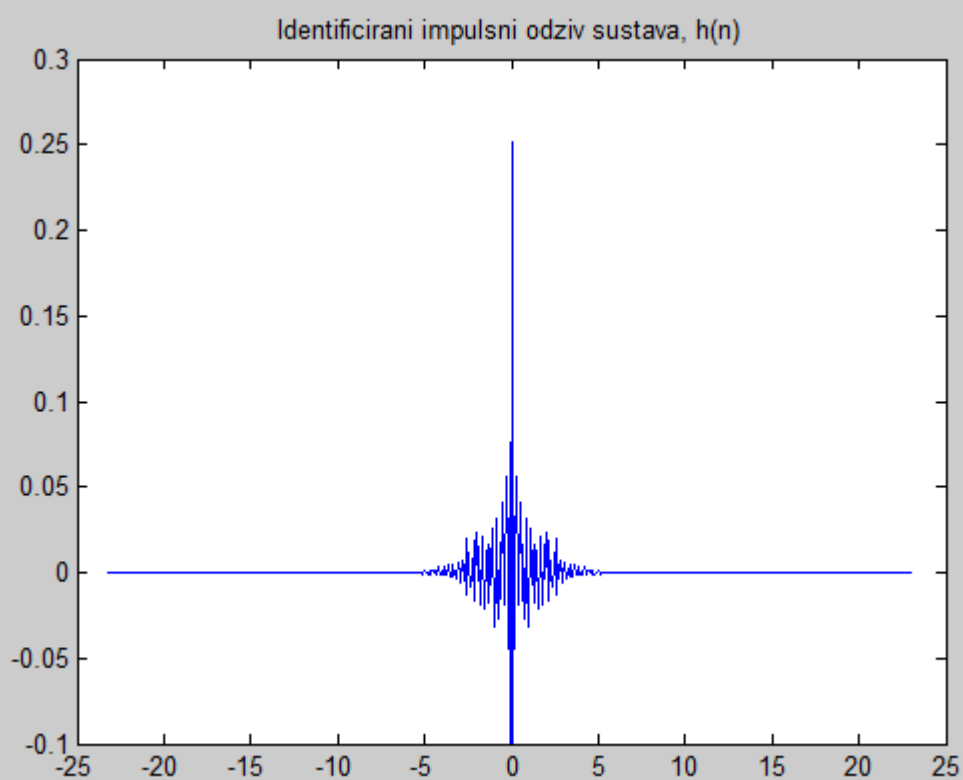
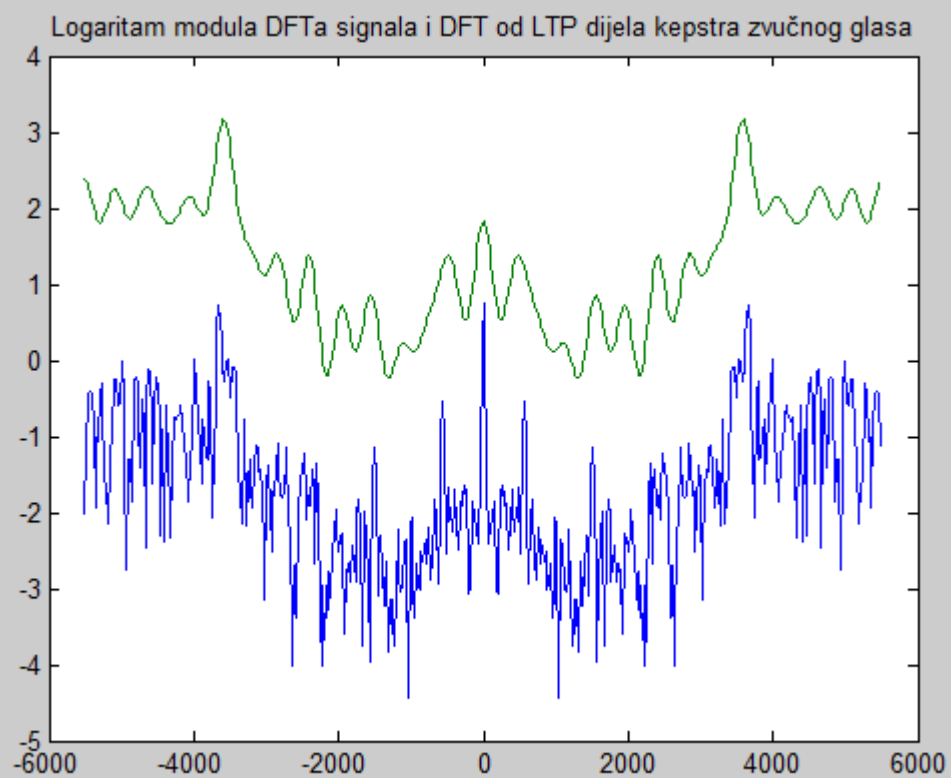


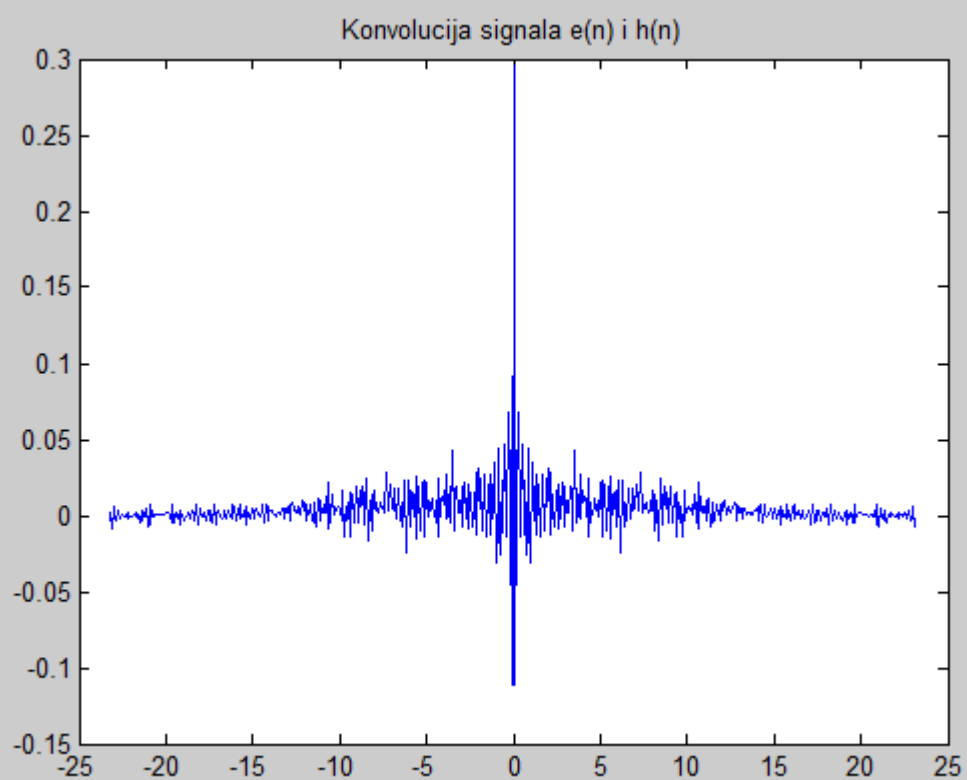
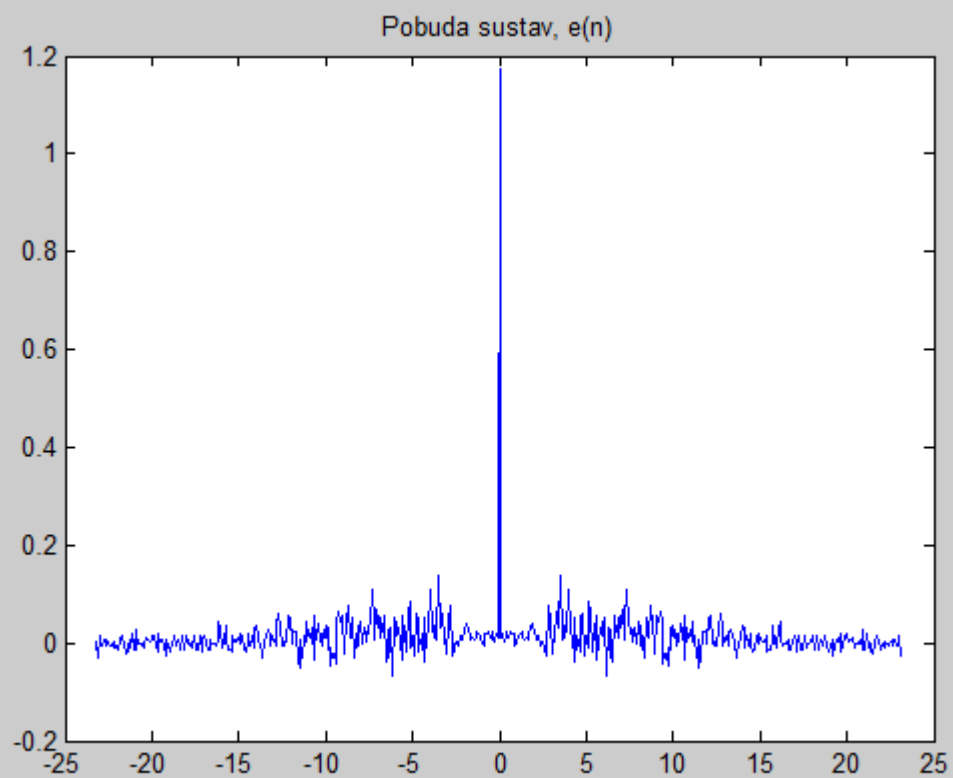




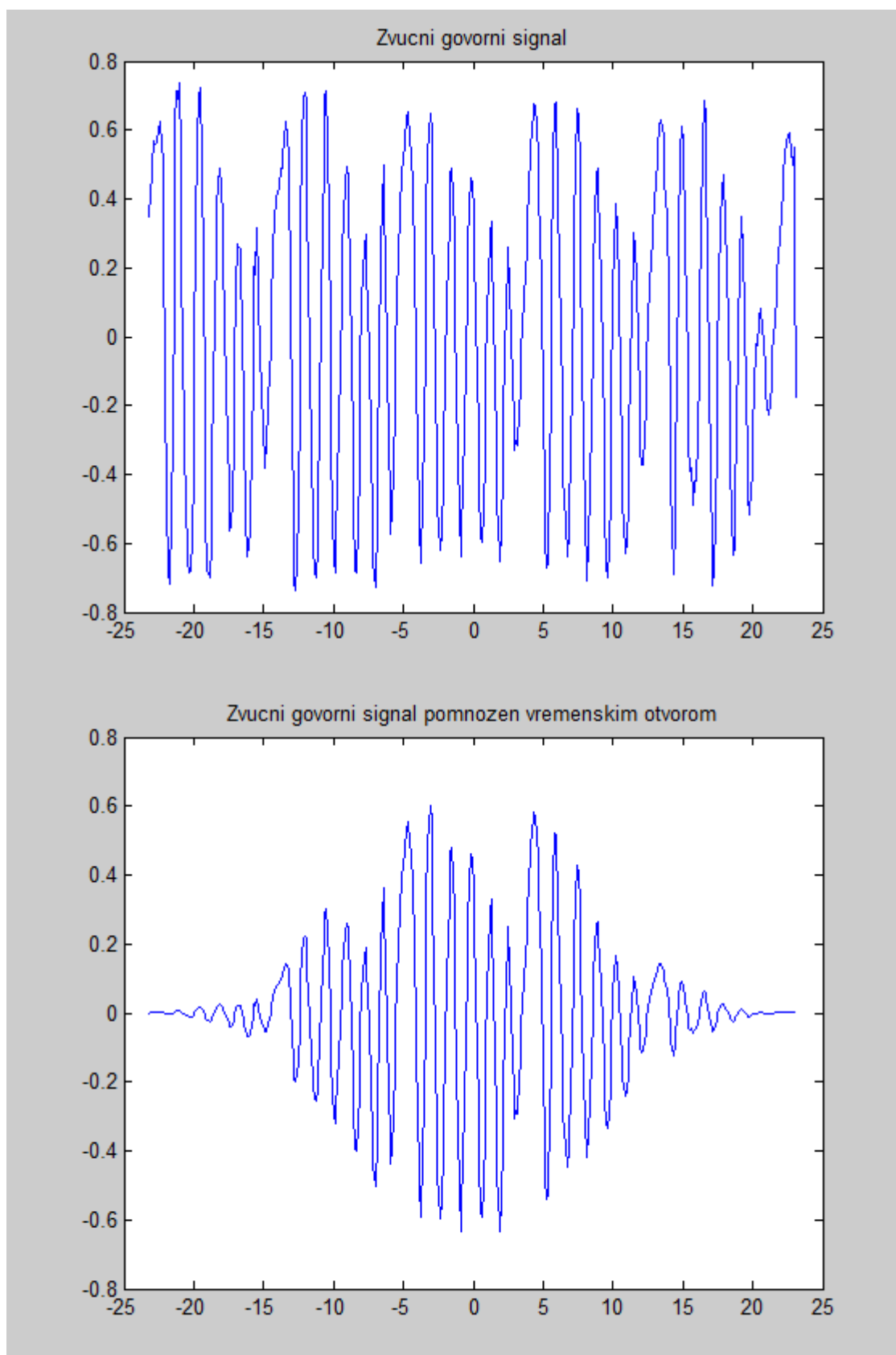


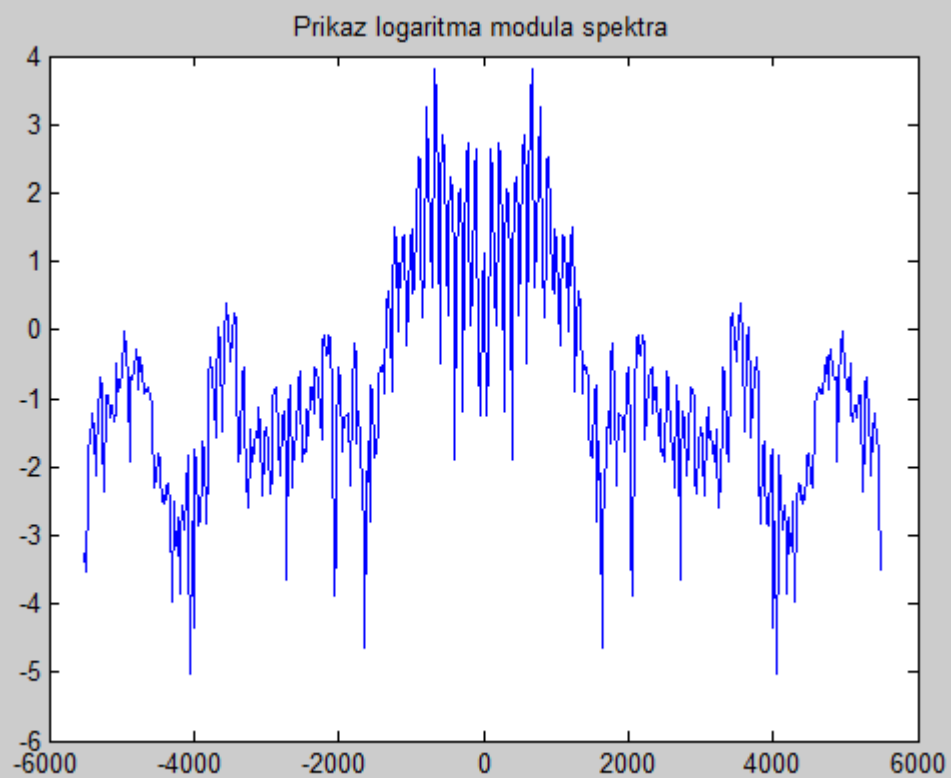




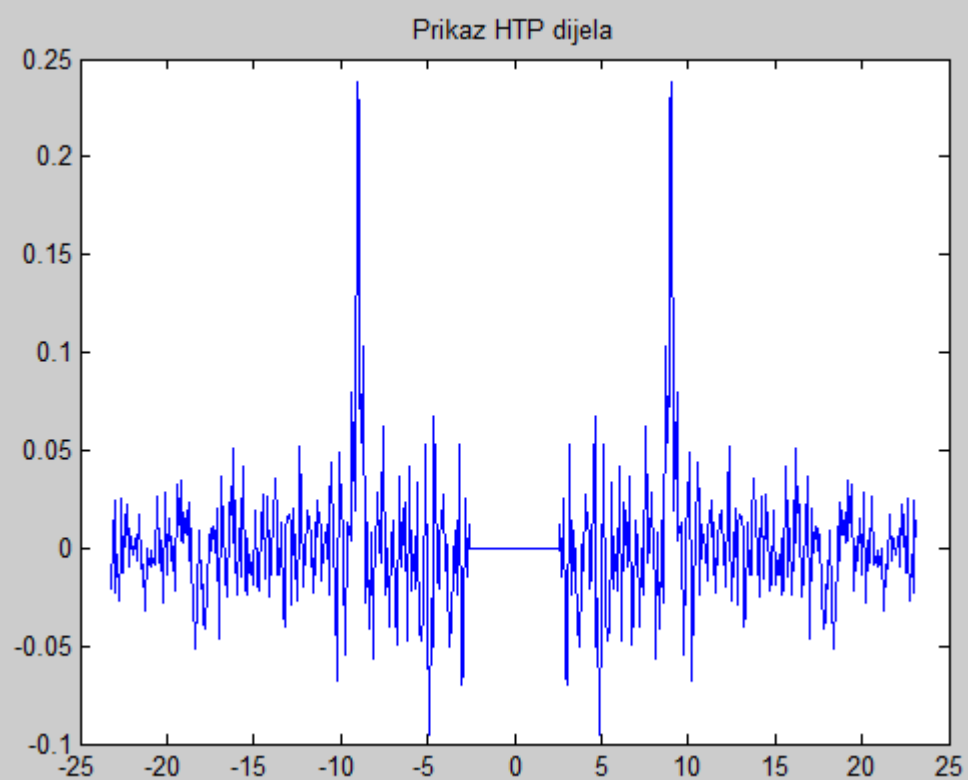
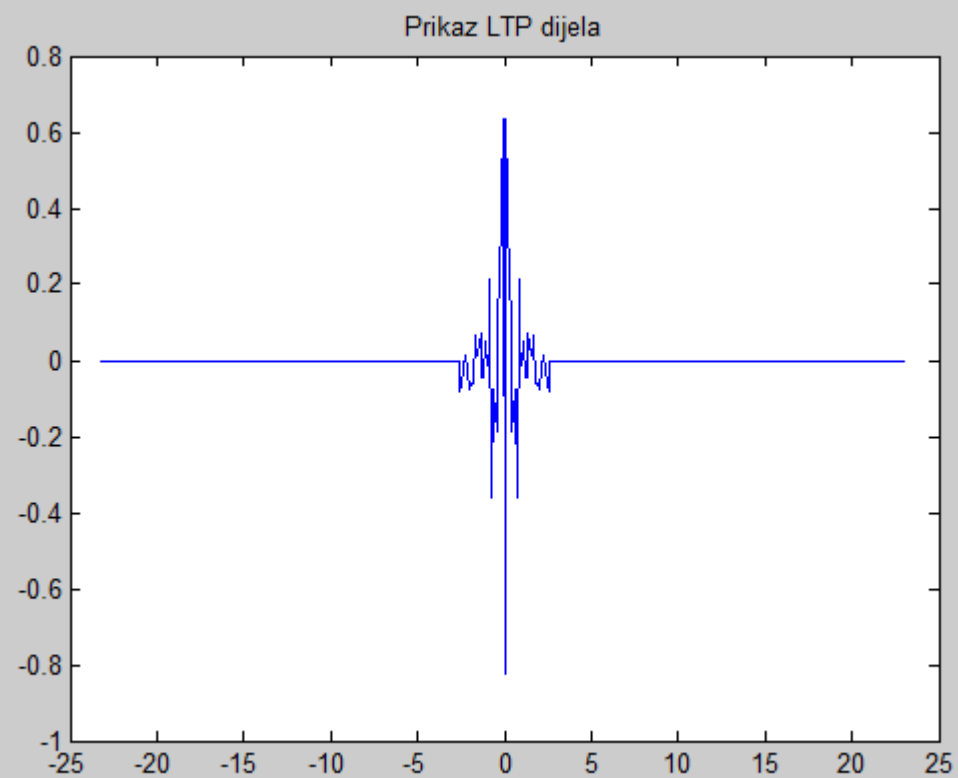


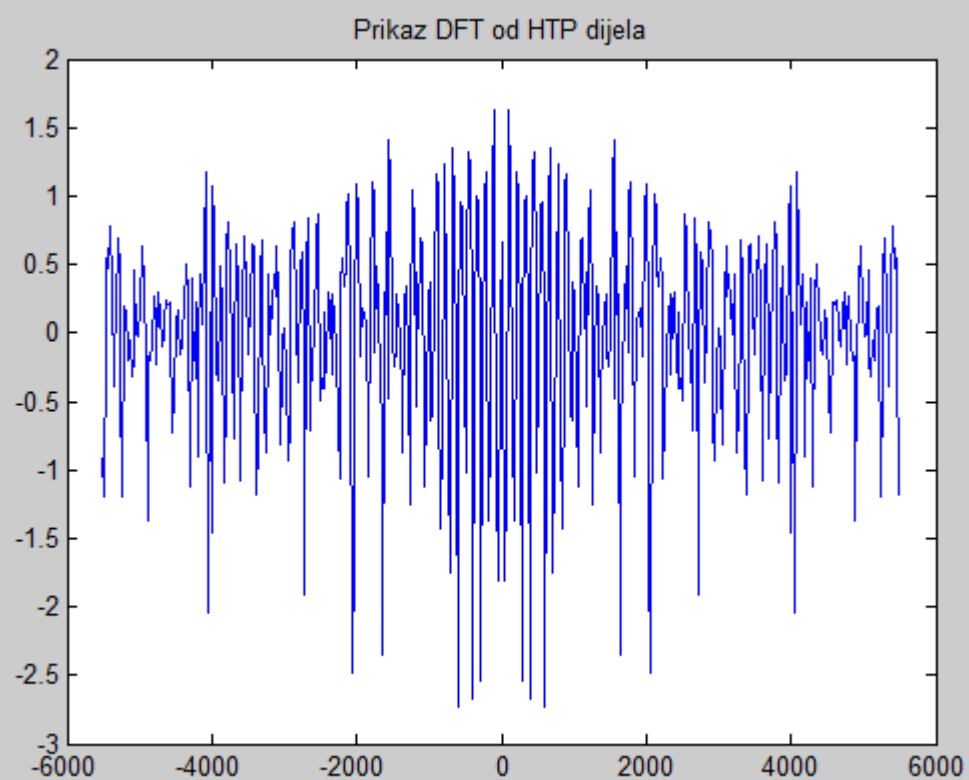
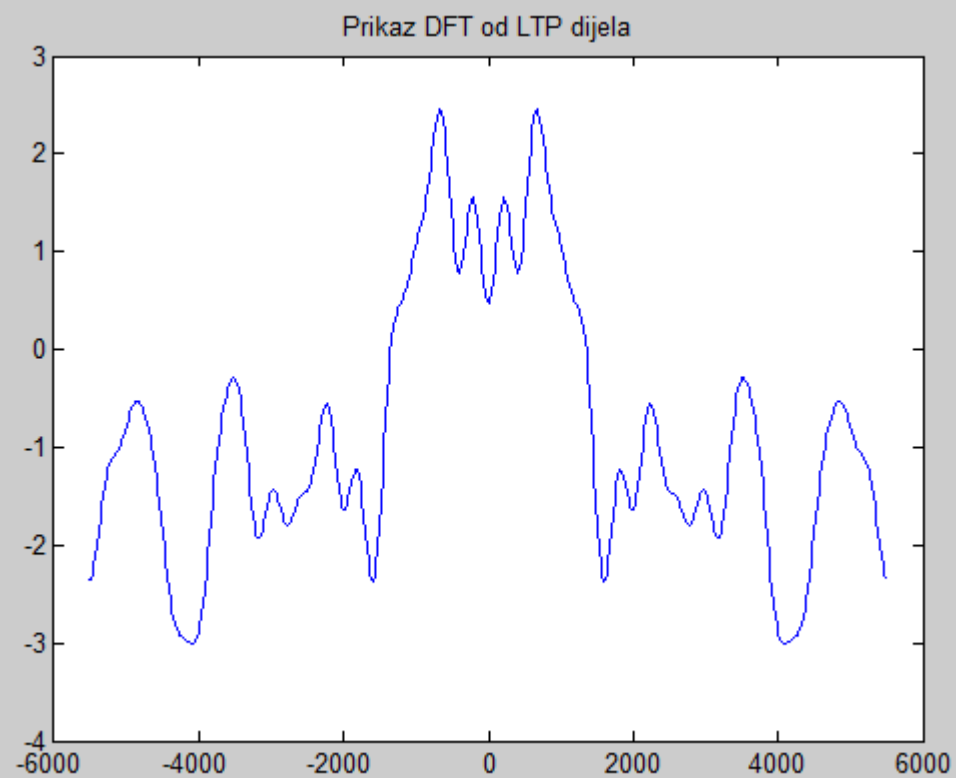
A sad za glas „a“:

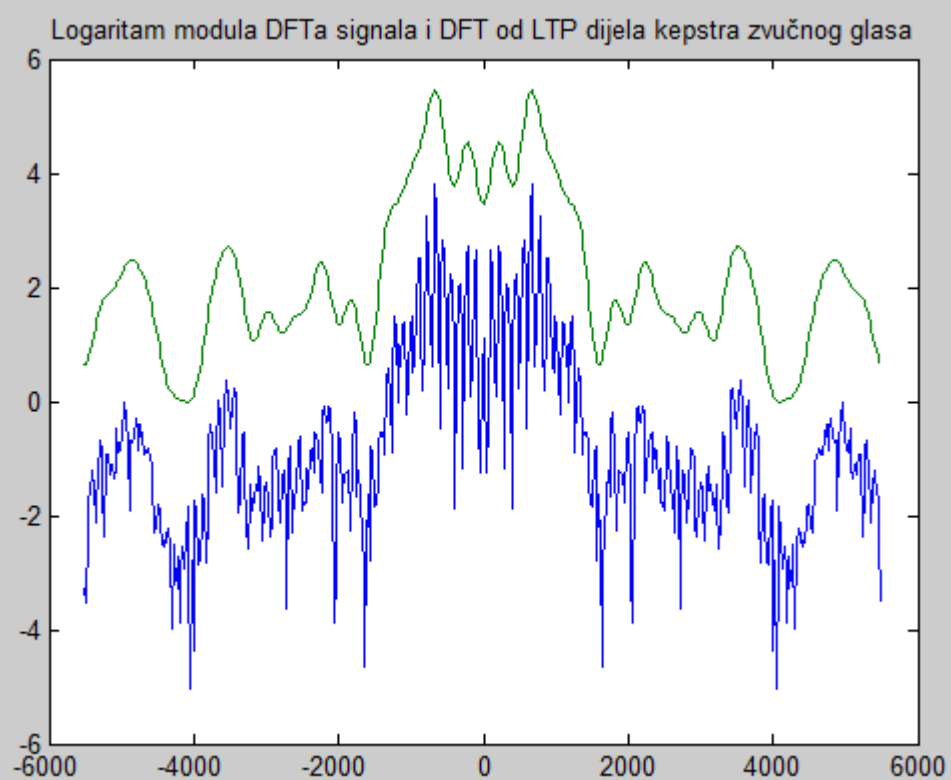
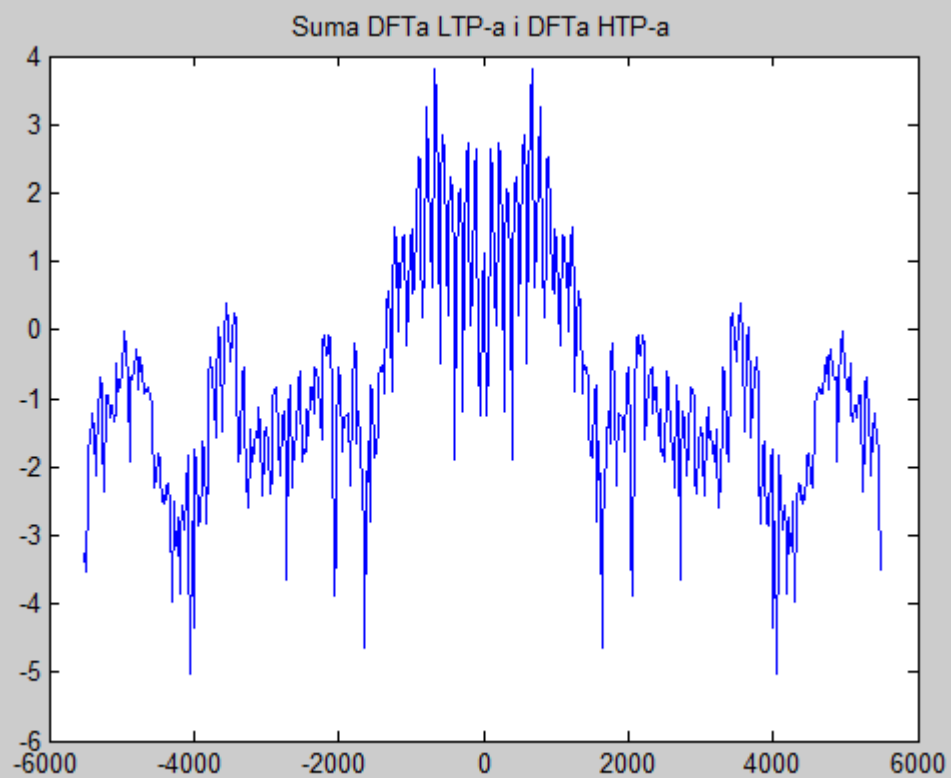


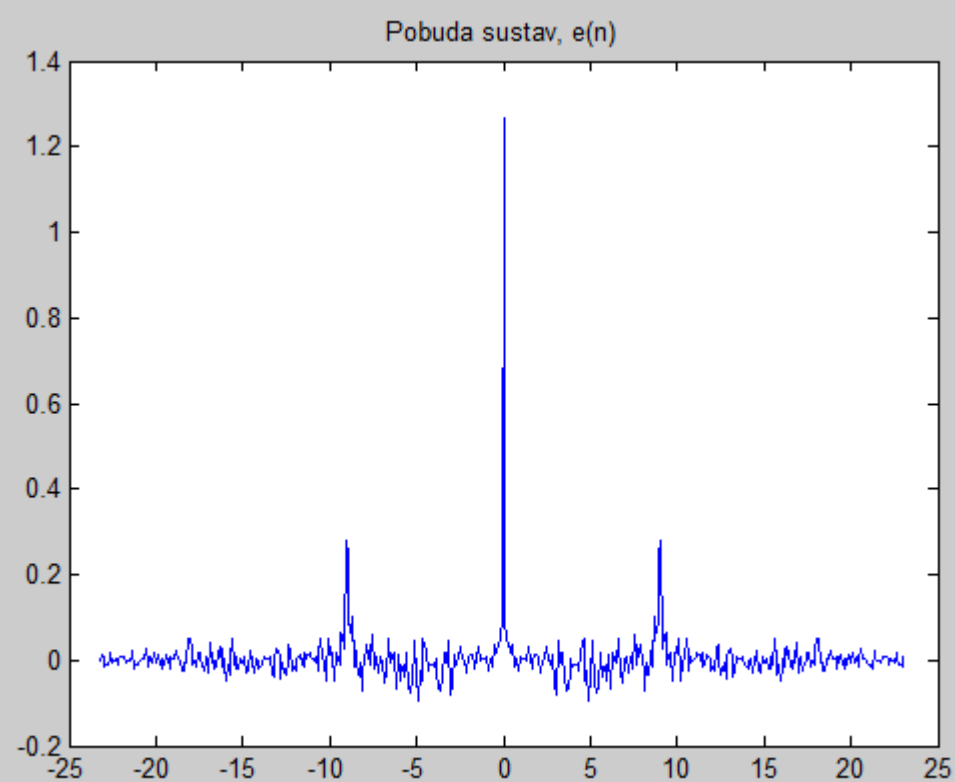
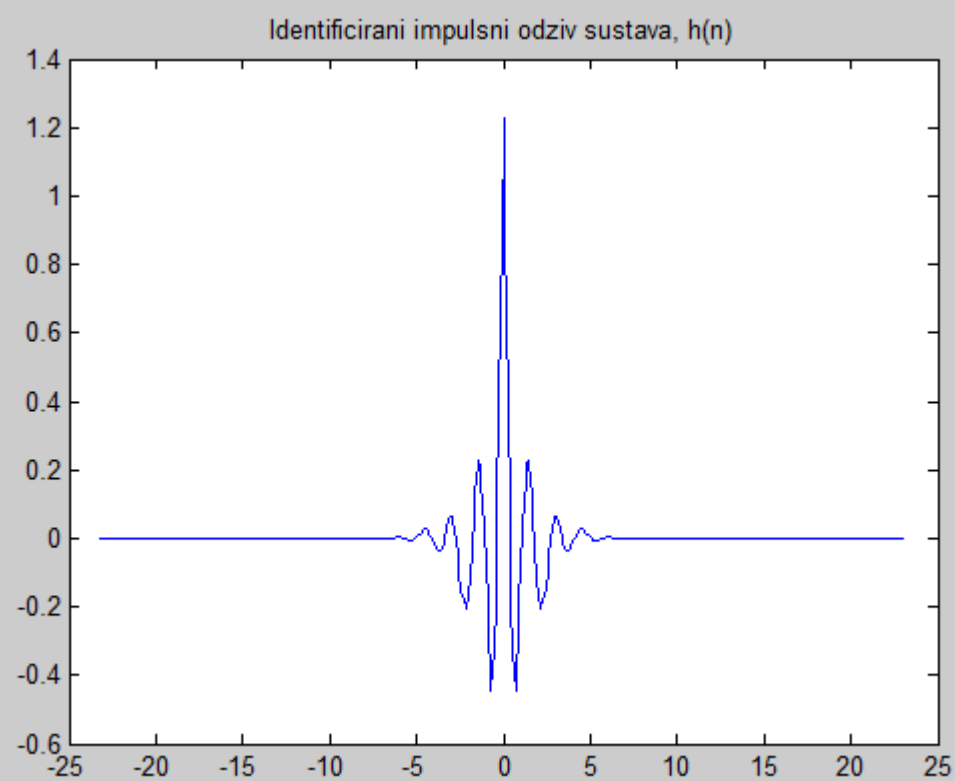


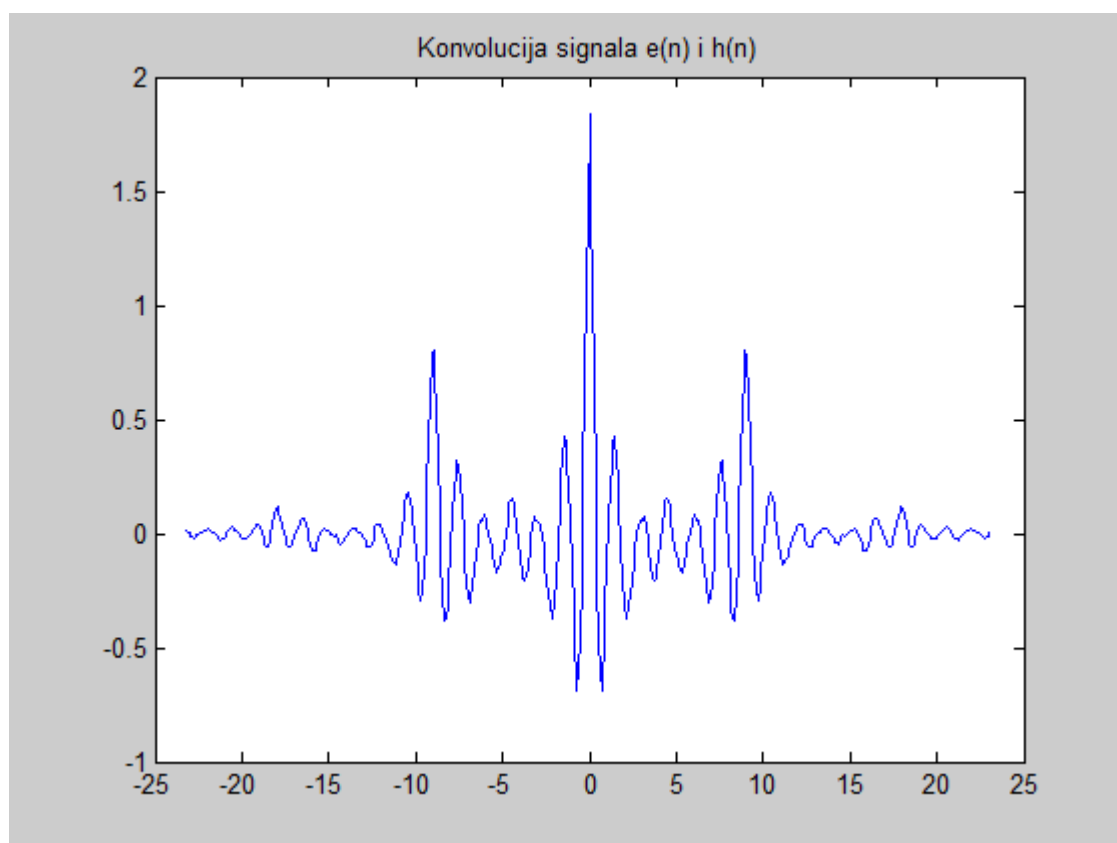










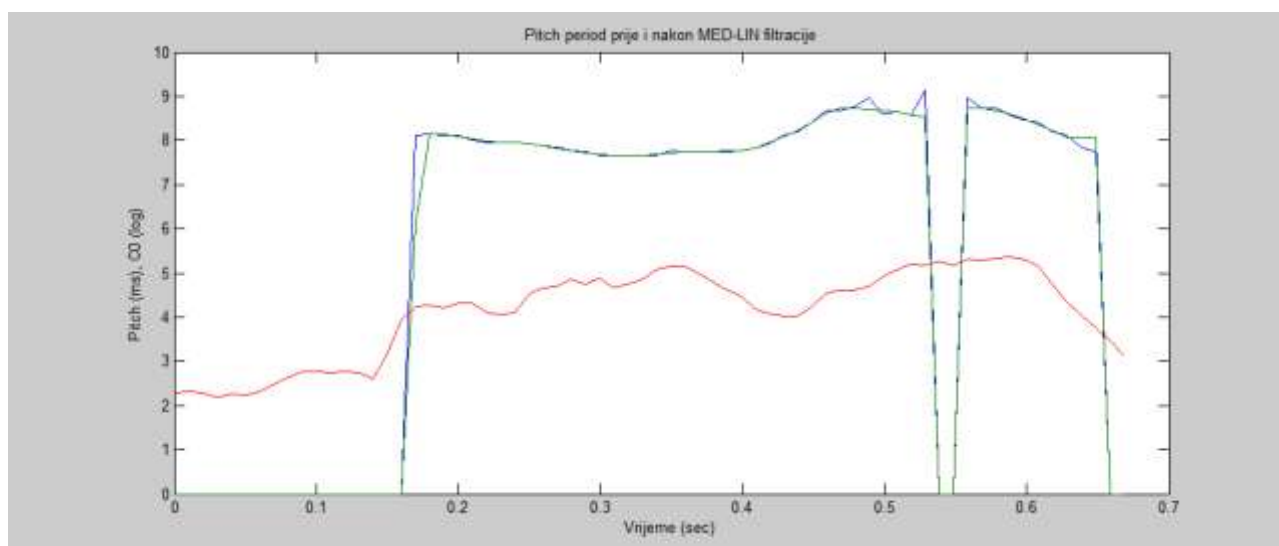


## 13. Određivanje zvučnosti i osnovne frekvencije titranja glasnica

### Zadatak 1.

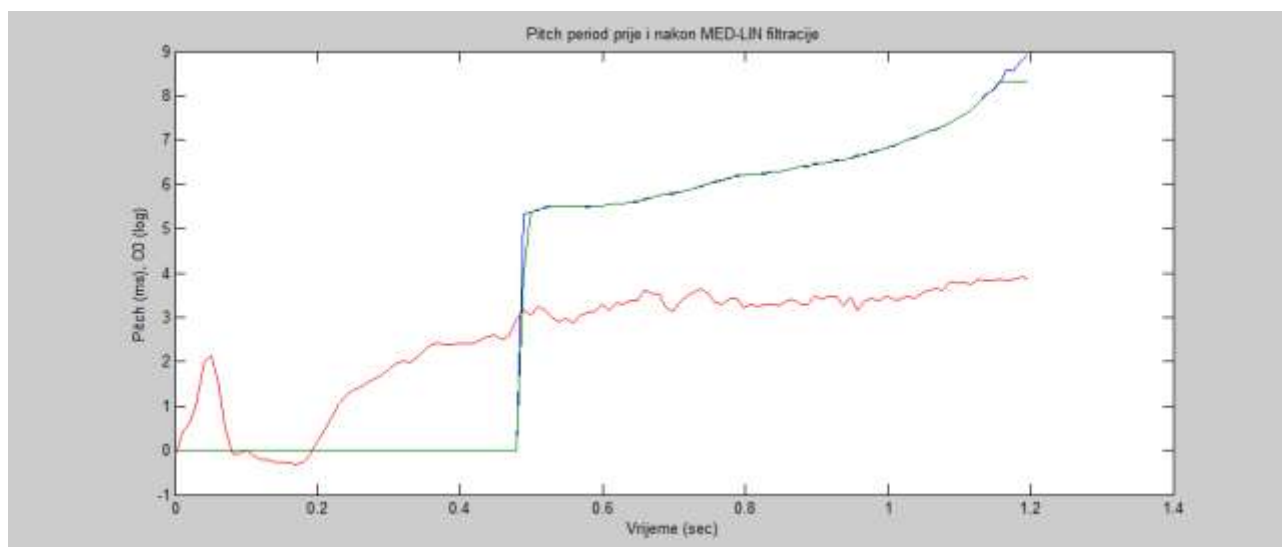
Koristeći prediktor Izvesti postupak određivanja zvučnosti i perioda titranja glasnica za govorni zapis „glava.wav“, koji se nalazi u radnom folderu ovog poglavlja. Odrediti periode titranja glasnica za prvi i drugi glas 'a' u ovoj riječi. Kakva je zvučnost preostalih glasova u ovoj riječi? Ponoviti analizu za zapis „sa2.wav“ koji je korišten u praktičnim primjerima prošlog poglavlja.

Za govornu datoteku „glava.wav“ prikazan je pitch period na slici 13.1. Na njemu se može uočiti kako je pitch period za prvi glas „a“ počeo na oko 7,5ms te na kraju završio na oko 8.5 ms. Drugi glas „a“ je počeo na oko 8.5ms te pao na 8ms. Razlog ovog pada leži u tome što ako poslušamo datoteku „glava.wav“ možemo čuti kako drugi glas „a“ zvuči uzlazno, odnosno zvuči nešto kao „aaAA“. Što se tiče ostalih glasova, glas „g“ je bezvučan jer kod njega ne titraju glasnice, dok su sonati „l“ i „v“ zvučni jer kod njih glasnice titraju.



13.1 Pitch periodi za datoteku "glava.wav"

Za datoteku „sa2.wav“ situacija je nešto drukčija što se može vidjeti i na slici 13.2. Naime, ako poslušamo datoteku „sa2.wav“ ona zvuči padajuće, odnosno zvuči nešto kao „SSAAAAaaaa..“ dakle potpuno suprotno od drugog glasa „a“ u prethodnoj datoteci. Vidljivo je na slici 13.2 kako to propadanje glasa „a“ se očituje na grafu kao porast pitch perioda. Što se tiče glasa „s“ on je očito bezvučan.

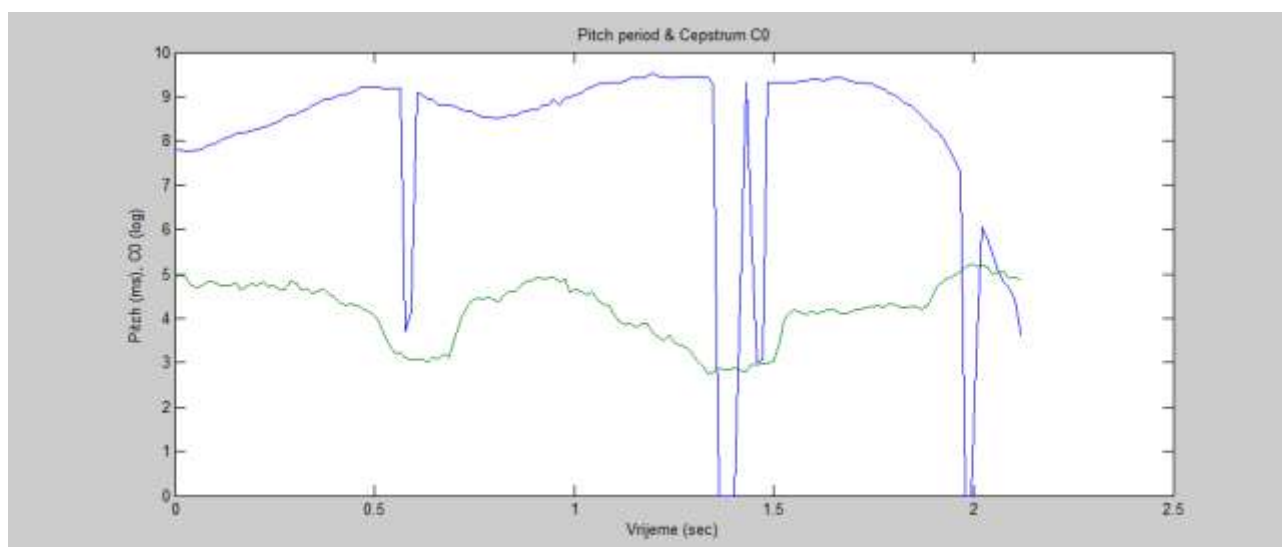


13.2 Pitch periodi za datoteku "sa2.wav"

### Zadatak 2.

*Snimiti vlastiti govor (jednu riječ) ali tako da se za neki zvučni glas (npr. samoglasnik) frekvencija titranja glasnica varira, tj. visina glasa se mijenja tokom izgovora tog glasa. Neka promjena frekvencije bude izražena. Za tako snimljenu sekvencu provedite keprstralnu analizu te provjerite da li je informacija o zvučnosti pojedinih okvira analize određena sukladno sa zvučnosti tih glasova.*

Za potrebe ovog zadatka odabrana je riječ banana koja je izgovorena i spremljena u datoteku „banana.wav“. Snimka je snimljen na način da je prvi samoglasnik „a“ padajući, drugi glas „a“ prvo poraste pa počne padati dok treći glas „a“ samo raste. Ti porasti i padovi mogu se vidjeti i na slici 13.3.



13.3 Pitch periodi za datoteku "banana.wav"

## 14. Primjer kompletnog koda govornog signala

### Zadatak 1.

- a) Proučite detaljnije kako radi funkcija za sintezu „synt.m“

✓

- b) Varirajte slijedeće parametre i utvrdite kako utječu na kvalitetu sintetiziranog govornog signala:

1. korak analize ..... param(2) pokušaj 220, 330, 440

Povećanjem pokušaja snimka postaje sve gora i gora. Najzanimljivija snimka je za korak analize namješten na 440, kada se pred kraj čuje nešto u stilu osobe koja šapuće.

2. v/uv prag ..... param(7) pokušaj 0.1, 0.25

Povećanjem ovog parametra, također dolazi do sličnih artefakata na riječi „glas“ gdje se čuje kao da netko šapuće, vjerojatno zvučno prelazi u bezvučno. Snižavanjem vrijednosti praga nije došlo do znatnijeg pogoršanja niti poboljšanja.

3. stupanj LPCa .... param(10) pokušaj 6, 16

Smanjenjem parametra 10 snimka postaje nerazumljivija na način da snimka zvuči kao da netko govori kroz hrpu snijega, dok povećanjem tog parametra snimka dobiva prirodniji zvuk, odnosno zvuči realnije i vjerodostojnije, kao da nije toliko sintetiziran glas, već samo loša snimka.

- c) Korištenjem postupaka skalarne kvantizacije koeficijenata koji su istraženi u poglavlju 15, prije sinteze provedite kvantizaciju parametara LPC modela (bilo a-reprezentacije, bilo k-reprezentacije), te istražite utjecaj pogrešaka kvantizacije LPC modela na sintetizirani govor. Pri tome koristite rezolucije koeficijenata kojima se postiže prosječno spektralno izobličenje od 1dB (što bi trebalo biti nečujno), odnosno 4 dB što bi trebalo uzrokovati čujne artefakte.

✓ „pog14\_zad1.m“

- d) Sličan postupak primijenite i na kvantizaciju pitch-perioda, odnosno logaritma rezidualne energije predikcije  $E_n$ .

✓ „pog14\_zad2.m“

- e) Kvantizirane parametre pohranite u novu mat-datoteku, te je proslijedite kao ulaz funkciji za sintezu. Kolika je ukupna potrebna brzina prijenosa za sve kvantizirane parametre modela uz koje je razlika između idealne sinteze i sinteze temeljene na kvantiziranim parametrima nezamjetna?

Skripta „pog14\_zad1.m“ opisuje kvantizaciju jednaku onoj u 15. poglavlju. Uz korištenje 3 bita za kvantizaciju čuje se lošija kvaliteta sintetiziranog govora jer je više pogrešaka, odnosno 3 bita je premalo, dok se za 4 bita ne čuje prevelika razlika, odnosno snimka je poprilično slična originalnoj sintetiziranoj snimci uz dodan šum.



**Zadatak 2.**

*Za svoju govornu sekvencu koja je korištena i u prvom primjeru provesti opisani postupak manipulacije vocoder parametarskog modela.*

Provedeno i spremljeno u arhivu, datoteke su istog naziva kao i u zadatku(glas\_du.wav ...)