

# VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

BRNO UNIVERSITY OF TECHNOLOGY

FAKULTA ELEKTROTECHNIKY A KOMUNIKAČNÍCH TECHNOLOGIÍ  
ÚSTAV TELEKOMUNIKACÍ

FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION  
DEPARTMENT OF TELECOMMUNICATIONS

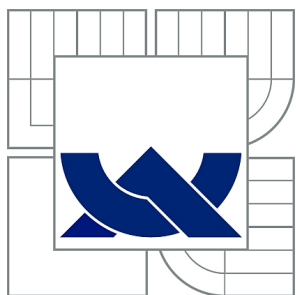
METODY DETEKCE AKUSTICKÝCH ZDROJŮ MIKROFONOVÝMI POLI

BAKALÁŘSKÁ PRÁCE  
BACHELOR'S THESIS

AUTOR PRÁCE  
AUTHOR

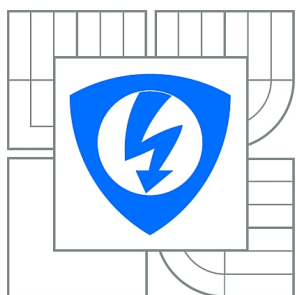
MICHAL DURKOŠ

BRNO 2011



**VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ**

BRNO UNIVERSITY OF TECHNOLOGY



**FAKULTA ELEKTROTECHNIKY A KOMUNIKAČNÍCH  
TECHNOLOGIÍ**

**ÚSTAV TELEKOMUNIKACÍ**

FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION  
DEPARTMENT OF TELECOMMUNICATIONS

## **METODY DETEKCE AKUSTICKÝCH ZDROJŮ MIKROFONOVÝMI POLI**

METHODS FOR DETECTING ACOUSTIC SOURCES USING MICROPHONE ARRAYS

**BAKALÁŘSKÁ PRÁCE**

BACHELOR'S THESIS

**AUTOR PRÁCE**

AUTHOR

**MICHAL DURKOŠ**

**VEDOUCÍ PRÁCE**

SUPERVISOR

**Ing. IVAN MÍČA**

BRNO 2011



VYSOKÉ UČENÍ  
TECHNICKÉ V BRNĚ

Fakulta elektrotechniky  
a komunikačních technologií

Ústav telekomunikací

# Bakalářská práce

bakalářský studijní obor  
Teleinformatika

**Student:** Michal Durkoš

**ID:** 115164

**Ročník:** 3

**Akademický rok:** 2010/2011

## NÁZEV TÉMATU:

**Metody detekce akustických zdrojů mikrofonomými poli**

## POKYNY PRO VYPRACOVÁNÍ:

Přesnost a spolehlivost lokalizace akustických zdrojů je do značné míry určena volbou lokalizační metody a uspořádání mikrofonomého pole. Na základě teoretického rozboru navrhnete pro konkrétní podmínky vhodné uspořádání mikrofonomů v poli a odpovídající lokalizační algoritmus. Přitom uvažujte praktické úlohy lokalizace jednoho či více úzkopásmových nebo širokopásmových zdrojů. Experimentálně ověřte dosahovanou přesnost a spolehlivost lokalizace.

## DOPORUČENÁ LITERATURA:

[1] BENESTY, J., CHEN, J., YITENG, H. Microphone Array Signal Processing. Berlin: Springer-Verlag, 2008. 250 s. ISBN 978-3-540-78612-2.

[2] EVEREST, F. A., The Master Handbook of Acoustics. McGraw/Hill, 2009. 528 s. ISBN 978-0071603324.

**Termín zadání:** 7.2.2011

**Termín odevzdání:** 2.6.2011

**Vedoucí práce:** Ing. Ivan Míča

**prof. Ing. Kamil Vrba, CSc.**

*Předseda oborové rady*

## UPOZORNĚNÍ:

Autor bakalářské práce nesmí při vytváření bakalářské práce porušit autorská práva třetích osob, zejména nesmí zasahovat nedovoleným způsobem do cizích autorských práv osobnostních a musí si být plně vědom následků porušení ustanovení § 11 a následujících autorského zákona č. 121/2000 Sb., včetně možných trestněprávních důsledků vyplývajících z ustanovení části druhé, hlavy VI. díl 4 Trestního zákoníku č.40/2009 Sb.

## **ABSTRAKT**

Táto práca porovnáva rozličné metódy na lokalizáciu zdroja zvuku pomocou lineárneho a guľového mikrofónového poľa. Zameriava sa na presnosť určenia smeru, z ktorého zvuk prichádza a na výpočetnú náročnosť algoritmov. Porovnávané sú algoritmy ESPRIT a MUSIC, ako aj metódy beamforming a Caponova metóda. Prvá časť teoreticky popisuje jednotlivé metódy a ich rovnice, v druhej časti sú uvedené výsledky experimentálneho overovania presnosti algoritmov. Výsledky sú čiastočne uvedené aj v prílohe, pretože sa jedná o väčšie množstvo grafov.

## **KLÍČOVÁ SLOVA**

DOA, lokalizácia zdroja, beamforming, Capon, MUSIC, ESPRIT, lineárne pole, guľové pole

## **ABSTRACT**

This paper compares different sound source localization methods exploiting linear and spherical microphone array. The focus is on direction of arrival accuracy and on computational cost of algorithms. Algorithms ESPRIT and MUSIC, as well as beamforming and Capon method are compared in this paper. The first part involves theoretical descriptions of particular methods, the other part involves results of experimental algorithm accuracy checking. Part of the results is listed in appendix, because of large amount of diagrams.

## **KEYWORDS**

DOA, source localization, beamforming, Capon, MUSIC, ESPRIT, linear array, spherical array

DURKOŠ, Michal *Metody detekce akustických zdrojů mikrofónovými poli: bakalářská práce*. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, Ústav telekomunikací, 2011. 50 s. Vedoucí práce byl Ing. Ivan Míča

## PROHLÁŠENÍ

Prohlašuji, že svou bakalářskou práci na téma „Metody detekce akustických zdrojů mikrofonovými poli“ jsem vypracoval samostatně pod vedením vedoucího bakalářské práce a s použitím odborné literatury a dalších informačních zdrojů, které jsou všechny citovány v práci a uvedeny v seznamu literatury na konci práce.

Jako autor uvedené bakalářské práce dále prohlašuji, že v souvislosti s vytvořením této bakalářské práce jsem neporušil autorská práva třetích osob, zejména jsem nezasáhl nedovoleným způsobem do cizích autorských práv osobnostních a jsem si plně vědom následků porušení ustanovení § 11 a následujících autorského zákona č. 121/2000 Sb., včetně možných trestněprávních důsledků vyplývajících z ustanovení § 152 trestního zákona č. 140/1961 Sb.

Brno .....

.....

(podpis autora)

Ďakujem vedúcemu práce Ing. Ivanovi Míčovi za veľmi užitočnú metodickú pomoc a cenné rady pri spracovaní bakalárskej práce.

# OBSAH

Úvod	10
<b>1 Metódy určenia polohy zdroja</b>	<b>11</b>
1.1 Beamforming	11
1.1.1 Delay-and-Sum beamforming	12
1.2 ESPRIT	16
1.3 Caponova metóda	20
1.4 MUSIC	20
<b>2 Experimentálne overenie presnosti algoritmov</b>	<b>22</b>
2.1 Pole zložené z 5 mikrofónov	22
2.2 Pole zložené z 8 mikrofónov	23
2.2.1 Lineárne pole	25
2.2.2 Guľové pole	28
<b>3 Záver</b>	<b>33</b>
Literatura	34
Seznam příloh	35
A Grafy pre metódy beamforming a Capon	36
B Fotografie usporiadania mikrofónových polí	49

# SEZNAM OBRÁZKŮ

1.1	Equidistantné lineárne pole so zdrojom $s(k)$ vo vzdialenom poli [1] . . .	13
1.2	Smerová odozva ekvidistantného poľa zloženého z piatich senzorov. . .	14
1.3	Ukážka vplyvu veľkosti medzier medzi senzormi na skreslenie signálu. . .	15
1.4	Simulácia výstupu mikrofónového poľa zloženého z piatich lineárne usporiadaných mikrofónov. . . . .	16
2.1	Poloha poľa a reproduktoru. . . . .	22
2.2	Usporiadanie reproduktorov. . . . .	24
2.3	Beamforming pre sínus s kmitočtom 2kHz - lineárne pole . . . . .	26
2.4	Metóda Capon pre sínus s kmitočtom 2kHz - lineárne pole . . . . .	27
2.5	Beamforming pre sínus s kmitočtom 2kHz - guľové pole . . . . .	30
2.6	Metóda Capon pre sínus s kmitočtom 2kHz - guľové pole . . . . .	31
A.1	Beamforming pre sínus s kmitočtom 4kHz - lineárne pole . . . . .	37
A.2	Metóda Capon pre sínus s kmitočtom 4kHz - lineárne pole . . . . .	38
A.3	Beamforming pre hudobný signál - lineárne pole . . . . .	39
A.4	Metóda Capon pre hudobný signál - lineárne pole . . . . .	40
A.5	Beamforming pre rečový signál - lineárne pole . . . . .	41
A.6	Metóda Capon pre rečový signál - lineárne pole . . . . .	42
A.7	Beamforming pre sínus s kmitočtom 4kHz - guľové pole . . . . .	43
A.8	Metóda Capon pre sínus s kmitočtom 4kHz - guľové pole . . . . .	44
A.9	Beamforming pre hudobný signál - guľové pole . . . . .	45
A.10	Metóda Capon pre hudobný signál - guľové pole . . . . .	46
A.11	Beamforming pre rečový signál - guľové pole . . . . .	47
A.12	Metóda Capon pre rečový signál - guľové pole . . . . .	48
B.1	Pole zložené z 5 mikrofónov . . . . .	49
B.2	Lineárne pole zložené z 8 mikrofónov . . . . .	49
B.3	Guľové pole zložené z 8 mikrofónov . . . . .	50

## SEZNAM TABULEK

2.1	Časové posuny a uhly dopadu pre jednotlivé posunutia vzoriek medzi dvomi susednými mikrofónmi. . . . .	23
2.2	Vzorkové oneskorenie a uhol dopadu pre jednotlivé dvojice mikrofónov.	23
2.3	Uhly dopadu pre jednotlivé reproduktory. . . . .	24
2.4	Určenie uhla dopadu pre signál sínus s kmitočtom 2kHz . . . . .	28
2.5	Určenie uhla dopadu pre signál sínus s kmitočtom 4kHz . . . . .	28
2.6	Určenie uhla dopadu pre hudobný signál . . . . .	29
2.7	Určenie uhla dopadu pre ženský hlas . . . . .	29
2.8	Určenie uhla dopadu pre signál sínus s kmitočtom 2kHz pri guľovom poli . . . . .	29
2.9	Určenie uhla dopadu pre signál sínus s kmitočtom 4kHz pri guľovom poli . . . . .	30
2.10	Určenie uhla dopadu pre hudobný signál pri guľovom poli . . . . .	31
2.11	Určenie uhla dopadu pre rečový signál pri guľovom poli . . . . .	32
A.1	Tabulka určujúca skutočné uhly dopadu pre jednotlivé časti obrázkov.	36

# ÚVOD

Lokalizácia akustických zdrojov sa využíva v rozličných situáciach, napríklad pri videokonferenciách na nasmerovanie kamery na práve rozprávajúceho rečníka, na nájdenie netesností karosérie auta vo vzduchovom tuneli a podobne. Presnosť lokalizácie je ovplyvnená polohami mikrofónov v mikrofónovom poli, ich vzájomnými vzdialenosťami a v neposlednej rade aj použitým lokalizačným algoritmom. Iné algoritmy sa používajú pri kruhovom poli, iné pri lineárnom poli.

V prvej kapitole sú opísané metódy založené na hľadaní maximálnej odozvy z určitého rozsahu smerov (hlavne  $-\pi/2$  až  $\pi/2$  - beamforming [8] a Caponova metóda [2], [4]), ako aj algoritmy pracujúce na základe rozkladu matíc signálov na charakteristické hodnoty a charakteristické vektory (algoritmy ESPRIT [7] a MUSIC [7]). V druhej časti je porovnávaná presnosť uvedených metód pomocou experimentálne nameraných dát v akustickom laboratóriu. Účinnosť algoritmov som porovnával so zreteľom na ich výpočetnú náročnosť. Výsledky sú zhrnuté v závere a v prílohe A, ktorá obsahuje grafické výsledky pre beamforming a Caponovu metódu.

# 1 METÓDY URČENIA POLOHY ZDROJA

## 1.1 Beamforming

Beamforming má dlhú históriu. Jeho použitie sa študovalo v množstve oblastí, z ktorých možno spomenúť radary, sonary, komunikácia, seizmológia, biomedicína a podobne. Používa sa na viacero úloh, najmä však na zistenie prítomnosti signálu, určenie smeru šírenia signálu (a teda aj na lokalizáciu zdroja) a na oddelenie žiadaneho signálu od šumu, iných signálov a odrazov, ktoré spravidla prichádzajú z iných smerov, ako požadovaný signál.

Schopnosť sensorového poľa priestorovej selektivity závisí na veľkosti aparatúry - ak sa aparatúra zväčšuje, zväčšuje sa aj selektivita. Avšak dôležitejším faktorom je veľkosť aparatúry vo vlnových dĺžkách. Vo všeobecnosti volíme vzdialenosti medzi jednotlivými senzormi tak, aby spĺňali podmienku  $d \leq \frac{\lambda}{2}$ , inak sa prejaví aliasing. Podľa požadovaného použitia môže veľkú rolu hrať aj geometria sensorového poľa. Napríklad pre správnu lokalizáciu zdroja musíme poznať geometriu poľa, dokonca jednoduchá geometria niekedy zjednoduší proces lokalizácie zdroja. Preto sa najčastejšie používajú kruhové a lineárne pole. Čoraz viac sa však objavujú aj trojdimenzionálne sférické polia, pretože lepšie zaznamenajú udalosti v zvukovom poli.

Beamformery fungujúce na princípe priestorových filtrov boli vyvinuté pre úzkopásmové signály ľahko definovateľné jednou frekvenciou. Avšak rečový signál, ktorým sa zaoberáme, je širokopásmový. Ak by sme použili úzkopásmové beamformery na takýto signál, smerový diagram pre jednotlivé kmitočty by bol rôzny a so zvyšujúcim sa kmitočtom by sa zmenšovala šírka hlavného laloku. Pri odlišnom nasmerovaní beamformera a uhla dopadu by bol požadovaný signál filtrovaný dolnou priepustou a okolitý šum by nebol správne utlmený v celom jeho spektre. Preto bolo treba vyvinúť iné spôsoby beamformingu na takéto signály. Dnes sa tento problém rieši rozložením požadovaného kmitočtového pásma do viacerých subpásiem a navrhnutím úzkopásmových beamformerov pre jednotlivé kmitočtové pásma.

Proces filtrovania môžeme rozdeliť do dvoch častí - synchronizácia a "weight-and-sum" (ďalej sumarizácia signálov). Synchronizácia spočíva v oneskorení signálu na výstupoch jednotlivých sensorov tak, aby jednotlivé komponenty signálu prichádzajúce z požadovaného smeru boli synchronizované - vo fáze. V druhom kroku sa jednotlivé komponenty signálu sčítajú do jedného spoločného výstupu. Oba kroky sú dôležité na správne ovládanie smerového diagramu (prvý krok ovplyvňuje nasmerovanie poľa, druhý krok šírku hlavného laloku a charakteristiky bočných lalokov), no v beamformingu sa kladie dôraz na určenie váhových koeficientov v druhom kroku. Tieto môžu byť definované tak, aby smerový diagram mal požadované vlastnosti,

no bežne je výhodnejšie určiť koeficienty adaptívnym algoritmom podľa vlastností signálu a šumu.

### 1.1.1 Delay-and-Sum beamforming

V sensorovom poli predpokladáme, že každý smer šírenia má svoje oneskorenie a svoj útlm. S týmto predpokladom zostavíme rovnicu vyjadrujúcu výstupy  $N$ -sensorového poľa v čase  $k$

$$\begin{aligned} y_n(k) &= \alpha_n s[k - t - F_n(\tau)] + v_n(k) \\ &= x_n(k) + v_n(k) \end{aligned} \quad (1.1)$$

kde ( $n=1, 2, \dots, N$ ), s rozsahom od 0 do 1, sú útlmové faktory závislé na šírení zvuku prostredím,  $s(k)$  je neznámy zdroj signálu,  $t$  je čas šírenia od zdroja k referenčnému senzoru 1,  $v_n(k)$  je šumový signál  $n$ -tého senzoru,  $\tau$  je relatívne oneskorenie medzi senzormi 1 a 2, a  $F_n(\tau)$  je relatívne oneskorenie medzi senzormi 1 a  $n$  (pre  $F_1(\tau) = 0$  a pre  $F_2(\tau) = \tau$ ) [1].

Výkonnosť DS beamformeru môžeme vyjadriť dvomi spôsobmi. Prvý z nich je určenie pomeru užitočného signálu a šumu (SNR) na vstupe a výstupe. Od beamformeru sa očakáva, že výstupné SNR bude vyššie než vstupné, teda že odstup užitočného signálu od šumu bude väčší. Druhý spôsob ukázania výkonnosti DS beamformeru je cez zodpovedajúci smerový diagram, ktorý nám úplne charakterizuje vstupno-výstupné správanie sa systému poľa. DS beamformer je vlastne  $N$ -bodový priestorový filter a jeho smerový diagram je definovaný ako amplitúda smerovej odozvy priestorového filtra. Podľa [1] je  $n$ -tý koeficient priestorového filtra  $\frac{1}{N} e^{j2\pi f F_n(\tau)}$ , kde  $f$  predstavuje kmitočet. Smerovú odozvu takéhoto filtra získame vykonaním Fourierovej transformácie. Keďže  $F_n(\tau)$  závisí rovnako na geometrii poľa aj na polohe zdroja, aj smerový diagram bude funkciou geometrie poľa a polohy zdroja. Takisto je závislý na počte sensorov a na kmitočte signálu.

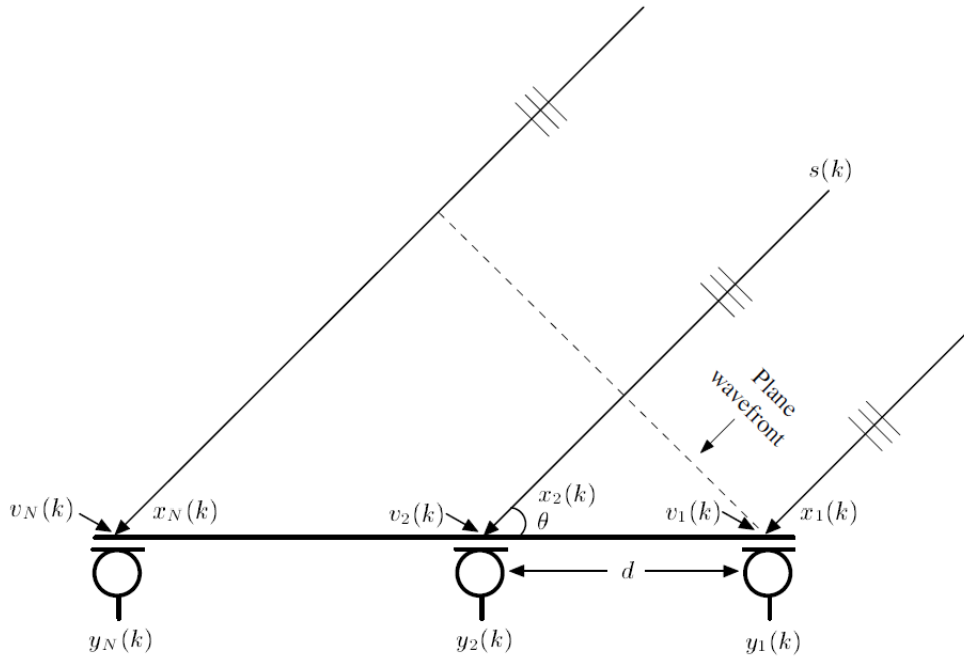
Predpokladajme lineárne pole pozostávajúce z  $N$  všesmerových sensorov (ako na obrázku 1.1). Označme vzdialenosti medzi susediacimi senzormi  $d$  a predpokladajme, že sensorové pole sa nachádza vo vzdialenom poli zdroja. Na pole potom dopadajú rovinné vlny pod uhlom  $\theta$ .

Relatívne oneskorenie medzi  $n$ -tým a referenčným sensorom (TDOA) potom môžeme vyjadriť ako

$$F_n(\tau) = (n - 1)\tau = (n - 1) \frac{d \cos \theta}{c} \quad (1.2)$$

kde  $c$  značí rýchlosť zvuku vo vzduchu. V tomto prípade môžeme smerovú odozvu DS filtra, ktorá je priestorovou Fourierovou transformáciou filtra, vyjadriť ako

$$S_{DS}(\Psi, \theta) = \frac{1}{N} \sum_{n=1}^N \left[ e^{j2\pi(n-1)f \frac{d \cos \theta}{c}} \right] e^{-j2\pi(n-1)f \frac{d \cos \theta}{c}} \quad (1.3)$$



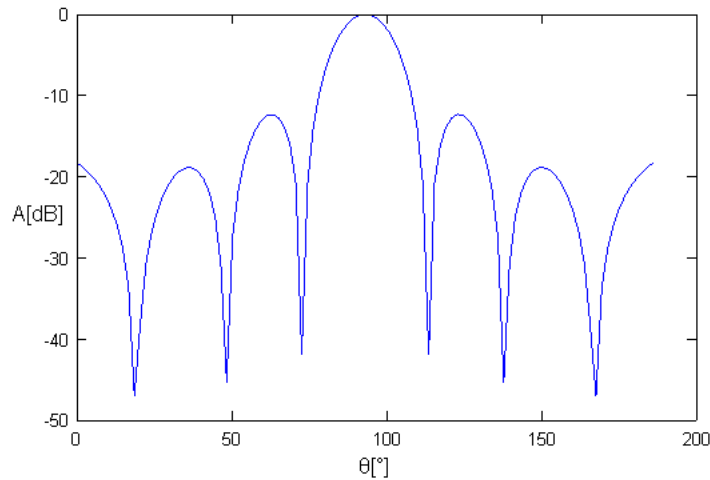
Obr. 1.1: Equidistantné lineárne pole so zdrojom  $s(k)$  vo vzdialenom poli [1]

kde  $\Psi (0 \leq \Psi \leq \pi)$ . Smerový diagram potom vyjadríme ako [1]

$$A_{DS}(\Psi, \theta) = |S_{DS}(\Psi, \theta)| = \left| \frac{\sin \left[ N \pi f \frac{d(\cos \Psi - \cos \theta)}{c} \right]}{N \sin \left[ \pi f \frac{d(\cos \Psi - \cos \theta)}{c} \right]} \right|. \quad (1.4)$$

Na obrázku 1.2 vidíme smerovú odozvu equidistantného lineárneho poľa zloženého z piatich sensorov, medzery medzi senzormi  $d = 5$  cm, signál dopadá pod uhlom  $\theta = 90^\circ$  a  $f = 2$  kHz. Lalok s najvyššou amplitúdou sa nazýva hlavný lalok, ostatné sa nazývajú bočné laloky. Dôležitým parametrom hlavného laloka je jeho šírka, definovaná ako úsek medzi prvými prechodmi nulou na každej strane hlavného laloka. U spomínaného lineárneho poľa sa dá tento parameter vypočítať pomocou vzorca  $2 \cos^{-1} \left( \frac{c}{Ndf} \right)$  [1]. Ako vidno zo vzorca, šírka laloka sa znižuje s rastúcim počtom sensorov, s rastúcimi medzerami medzi senzormi a s rastúcim kmitočtom. Výška bočných lalokov vyjadruje zisk šumu a ostatných nežiaducich zdrojov umiestnených mimo smer zamerania beamformeru. Pri návrhoch sensorových polí a beamformerov sa snažíme upraviť bočné laloky tak, aby boli čo najnižšie a teda nežiaduce signály prichádzajúce z iných smerov boli utlmené čo najviacej.

Avšak musíme si dať pozor na problém súvisjúci s medzerami medzi senzormi. Ako bolo vyššie spomenuté, šírka hlavného laloku sa znižuje s rastúcimi medzerami  $d$  medzi senzormi. Takže ak chceme ostrejší lalok, jednoducho zväčšíme medzery, čo vyústi vo väčšie rozmery pola. To by malo viesť k zníženiu šumu. Preto by sme teda

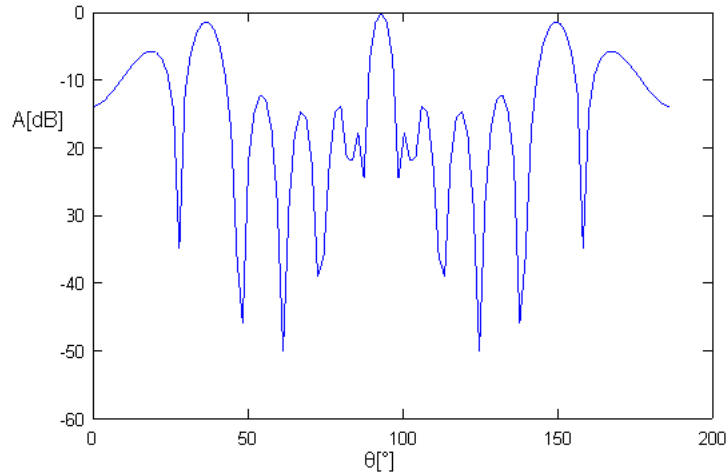


Obr. 1.2: Smerová odozva ekvidistantného poľa zloženého z piatich senzorov.

mali chcieť čo najväčšie medzery medzi senzormi. Avšak, ak  $d$  je väčšie ako  $\frac{\lambda}{2} = \frac{c}{2f}$ , kde  $\lambda$  je vlnová dĺžka signálu, vzrastie nám priestorový aliasing. Toto je vidno na obrázku 1.3, kde sme použili rovnaké pole ako v prípade obrázku 1.2, no medzeru medzi senzormi sme zvýšili na 20 cm. Tentokrát máme možnosť vidieť tri vysoké laloky - hlavný v strede, a dva krajné. Oba krajné laloky zobrazujú úroveň šumov, nežiaducich signálov a odrazov prichádzajúcich z iných smerov ako je požadované. Tieto signály by boli nerozoznateľné od signálu šíriaceho sa zo smeru hlavného laloku. Aby sme predišli tejto nejednoznačnosti, medzery medzi senzormi musia spĺňať podmienku  $d \leq \frac{\lambda}{2} = \frac{c}{2f}$ .

### Simulácia mikrofónového poľa

Túto časť projektu som spracoval vo voľne dostupnom programe GNU Octave, verzia 3.0.1. Cieľom bolo vytvoriť funkciu, ktorá by simulovala stav na výstupe mikrofónového poľa v prípade, že na pole dopadá zvuková vlna sínusového priebehu s určitým kmitočtom. Podľa teoretických úvah mala byť na výstupe najvyššia hodnota vtedy, ak dopadá zvuková vlna na mikrofónové pole kolmo. To je spôsobené tým, že na všetky mikrofóny dopadá zvuková vlna v rovnaký čas a nie sú žiadne časové posuny medzi jednotlivými mikrofónmi. Ako je viditeľné z obrázku 1.1, v prípade, že uhol dopadu  $\theta$  bude iný, bude sa líšiť čas dopadu rovinatej zvukovej vlny na jednotlivé mikrofóny, pretože k nasledujúcemu mikrofónu musí prejsť zvuková vlna dlhšiu trasu ako k referenčnému mikrofónu. Z uvedeného vyplýva, že signály nedopadnú na jednotlivé mikrofóny vo fáze, takže nebude získaný maximálny možný signál, ako pri



Obr. 1.3: Ukážka vplyvu veľkosti medzier medzi senzormi na skreslenie signálu.

kolmom dopade. Najskôr som teda zistil rozdiely v dráhe, ktoré musí uraziť rovinná vlna, než sa dostane k jednotlivým mikrofónom. Toto som dosiahol podľa rovnice

$$\Delta\lambda = d \cdot \cos\theta \quad (1.5)$$

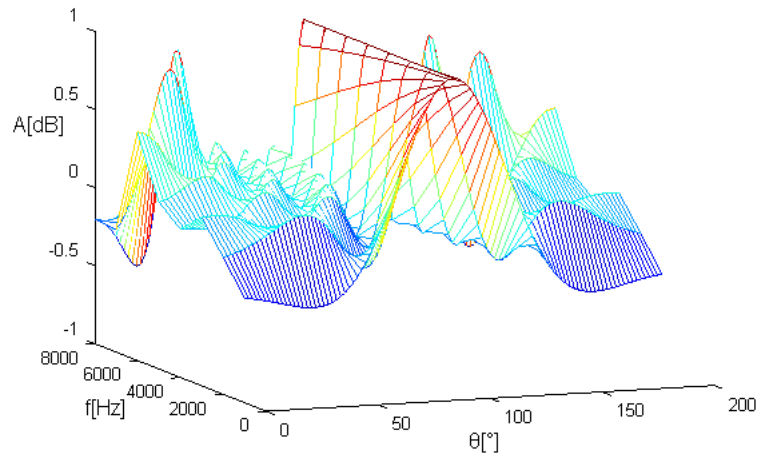
kde vzdialenosť mikrofónov  $d = 0,05$  m a uhol dopadu  $-\pi/2 \leq \theta \leq \pi/2$  sa mení podľa predpokladanej polohy zdroja. Jednotlivé mikrofóny zachytia signál s rozdielnou fázou. Tento rozdiel je vyjadrený rovnicou

$$\Delta\varphi = \frac{2 \cdot \pi \cdot f \cdot \Delta\pi}{c} \quad (1.6)$$

Pomocou tejto hodnoty som vypočítal okamžitú hodnotu signálu dopadajúceho na jednotlivé mikrofóny vzhľadom k rozdielnym dĺžkam dráh

$$y = \sin(\theta - \Delta\varphi). \quad (1.7)$$

Tento postup som opakoval pre všetky uhly  $-\pi/2 \leq \theta \leq \pi/2$  a pre kmitočty od 1000 Hz po 8000 Hz s krokom 1000 Hz. Výsledok je na obrázku 1.4 ( vzdialenosti medzi mikrofónmi  $d = 5$  cm, kmitočet  $1000 \text{ Hz} \leq f \leq 8000 \text{ Hz}$ , uhol dopadu  $-\pi/2 \leq \theta \leq \pi/2$ . Na ose z je vyjadrená maximálna amplitúda hodnotou 1 ). Na obrázku je vidno rastúci prejav aliasingu s rastúcim kmitočtom. Pre zníženie tohto prejavu na vyšších kmitočtoch je treba znížiť vzdialenosti medzi mikrofónmi tak, aby bola dodržaná podmienka  $d \leq \lambda/2 = \frac{c}{2f}$ . Pomocou tejto simulácie som vytvoril aj obrázky 1.2 a 1.3 použité v teoretickej časti.



Obr. 1.4: Simulácia výstupu mikrofónového poľa zloženého z piatich lineárne usporiadaných mikrofónov.

## 1.2 ESPRIT

ESPRIT - Estimation of Signal Parameters Via Rotational Invariance Techniques (určenie parametrov signálu pomocou techniky striedavej nemennosti) je algoritmus, ktorý bol prvýkrát predstavený v dizertačnej práci R. H. Roy na univerzite Stanford v roku 1987. Oproti dovtedy známym algoritmom (MUSIC, MLM, MEM) sa vyznačoval výrazným znížením výpočetných a úložných nákladov. V oblasti určenia smeru dopadu signálu (ďalej DOA - z anglického Direction of Arrival) sa tieto úspory dosiahli tým, že sensorové pole malo tzv. posunutú nemennosť - teda senzory sa vyskytovali vo vyrovnaných pároch s identickými vektormi posunu. ESPRIT je taktiež odolnejší voči nepresnostiam poľa, než bol MUSIC.

ESPRIT sa dá použiť na široké spektrum problémov, avšak v tejto práci sa venujeme jeho využitiu na určenie DOA. Predpokladáme izotropické nedisperzívne prostredie, v ktorom sa zvukové vlny šíria priamo, a zdroje signálu sú umiestnené vo vzdialenom poli sensorového poľa, teda zvukové vlny dopadajú na pole vo forme rovinatej vlny. Keďže pri úzkopásmových signáloch predpokladáme pevný známy stredný kmitočet  $\omega_0$ ,  $i$ -ty signál môžeme vyjadriť ako  $\tilde{s}_i(t) = u_i(t) \cos(\omega_0 t + v_i(t))$ , kde  $u_i(t)$  a  $v_i(t)$  sú pomaly premenlivé funkcie ktoré určujú amplitúdu a fázu  $\tilde{s}_i(t)$ . Pre takéto signály je pohodlnejšie použiť vyjadrenie vo forme komplexnej obálky, kde  $\tilde{s}_i(t) = \text{Re} \{s(t)\}$ , kde  $s(t) = u(t)e^{j(\omega_0 t + v(t))}$ . Keďže predpoklad úzkopásmového signálu naznačuje že  $u(t) \approx u(t - \tau)$  a  $v(t) \approx v(t - \tau)$  pre všetky možné zdržania signálu vplyvom šírenia  $\tau$ , vplyv časového posunu na dopadnuté vlny je obyčajný

posun fázy, teda  $s(t - \tau) \approx s(t)e^{-j\omega_0\tau}$ . Vo výsledku teda môžeme komplexný signál na výstupe  $k$ -teho senzoru v čase  $t$  napísať ako

$$x_k(t) = \sum_{i=1}^d a_k(\theta_i) s_i(t - \tau_k(\theta_i)), \quad (1.8)$$

kde  $\tau_k(\theta_i)$  je oneskorenie medzi referenčným bodom a  $k$ -tym sensorom pre  $i$ -tu vlnu dopadajúcu na pole zo smeru  $\theta_i$ ,  $a_k(\theta_i)$  je komplexná odozva odpovedajúceho senzoru (zisk a fáza) na kmitočte  $\omega_0$ , a predpokladáme  $d$  zdrojov. Keď použijeme vektorové označenie pre výstupy  $m$  sensorov, dostaneme tvar

$$\vec{x}(t) = \sum_{i=1}^d \vec{a}(\theta_i) s_i(t), \quad (1.9)$$

často ozačovaný aj ako *odozva poľa* alebo *vektor nasmerovania poľa* pre smer  $\theta_i$ . Ďalej definujeme  $\mathbf{A}(\theta) = [\vec{a}(\theta_1), \dots, \vec{a}(\theta_d)]$ ,  $\vec{s}(t) = [s_1(t), \dots, s_d(t)]^T$ , a pridaním šumu  $\vec{n}(t)$  dostaneme všeobecný dátový model merania podľa [7]

$$\vec{x}(t) = \mathbf{A}(\theta)\vec{s}(t) + \vec{n}(t) \quad (1.10)$$

Ďalej ešte predpokladáme, že počet sensorov  $m$  je vyšší ako počet zdrojov  $d$  a že senzory v jednotlivých pároch sú od seba vzdialené o vektor posunu  $\vec{\Delta}$ . Sensorové pole si rozdelíme na dve úplne identické čiastočné polia  $Z_x$  a  $Z_y$  vzájomne posunuté (nie otočené) o vektor posunu  $\vec{\Delta}$  o veľkosti  $\Delta$ . Signály prijaté  $i$ -tou dvojicou môžeme potom vyjadriť ako

$$\begin{aligned} x_i(t) &= \sum_{k=1}^d s_k(t) a_i(\theta_k) + n_x i(t), \\ y_i(t) &= \sum_{k=1}^d s_k(t) e^{j\omega_0 \Delta \sin \theta_k / c} a_i(\theta_k) + n_y i(t), \end{aligned} \quad (1.11)$$

kde  $\theta_k$  je DOA  $k$ -teho zdroja vzhľadom k smeru vektoru posunu  $\vec{\Delta}$ . Kombináciou výstupu každého senzoru v dvoch čiastočných poliach môžeme získané dátové vektory napísať nasledujúcim spôsobom

$$\vec{x}(t) = \mathbf{A} \vec{s}(t) + \vec{n}_x(t), \quad (1.12)$$

$$\vec{y}(t) = \mathbf{A} \Phi \vec{s}(t) + \vec{n}_y(t), \quad (1.13)$$

kde vektor  $\vec{s}(t)$  je  $d \times 1$  vektor dopadajúcich signálov pozorovaných na referenčnom senzore v čiastočnom poli  $Z_x$ . Matica  $\Phi$  je diagonálna matica fázových posunov medzi dvojicami sensorov pre  $d$  vlnoplôch s rozmermi  $d \times d$  a má tvar

$$\Phi = \text{diag} \{ e^{j\gamma^1}, \dots, e^{j\gamma^d} \}, \quad (1.14)$$

kde  $\gamma_k = \omega_0 \Delta \sin \theta_k / c$ .  $\Phi$  je unitárna matica, ktorá určuje vzťah medzi meraniami z čiastočného poľa  $Z_x$  a  $Z_y$ . Výstupy jednotlivých čiastočných polí môžeme skombinovať a dostaneme tak celkový výstupný vektor poľa

$$\vec{z}(t) = [\vec{x}(t), \vec{y}(t)]^T = \bar{\mathbf{A}}\vec{s}(t) + \vec{n}_z(t), \quad (1.15)$$

$$\bar{\mathbf{A}} = [\mathbf{A}, \mathbf{A}\Phi]^T, \vec{n}_z(t) = [\vec{n}_x(t), \vec{n}_y(t)]. \quad (1.16)$$

Základnou ideou ESPRIT-u je využitie rotačnej nemennosti východzieho signálového podpriestoru odvodeného posuvnou nemennosťou sensorového poľa.[7] Dôležitý je ten podpriestor, ktorý obsahuje výstupy z obidvoch spomínaných čiastočných polí. Simultánnym vzorkovaním výstupov získame dve sady vektorov  $\vec{E}_x$  a  $\vec{E}_y$  rozprestrených v tom istom signálovom podpriestore. Tento podpriestor sa dá získať zo znalosti kovariančnej matice  $\mathbf{R}_{ZZ} = \bar{\mathbf{A}}\mathbf{R}_{SS}\bar{\mathbf{A}}^* + \sigma^2 \Sigma_n$ . Platí

$$\mathbf{R}_{SS} = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N s(t)s^*(t).$$

$2m - d$  najmenších charakteristických hodnôt (anglicky *eigenvalues*) z  $(\mathbf{R}_{ZZ}, \Sigma_n)$  je rovných  $\sigma^2$ ,  $d$  vektorov charakteristických hodnôt (anglicky *eigenvectors*) odpovedajúcich  $d$  najväčším charakteristickým hodnotám sa použije, aby sme získali signálový podpriestor  $\mathbf{E}_S = \Sigma_n [e_1 | \dots | e_d]$ , kde  $\Re\{\mathbf{E}_S\} = \Re\{\bar{\mathbf{A}}\}$ .

Keďže platí  $\Re\{\mathbf{E}_S\} = \Re\{\bar{\mathbf{A}}\}$ , musí existovať jedinečná, nesingulárna matica  $\mathbf{T}$  [7], aby platilo

$$\mathbf{E}_S = \bar{\mathbf{A}}\mathbf{T}. \quad (1.17)$$

Nemenná štruktúra poľa umožňuje rozloženie  $\mathbf{E}_S$  na  $\mathbf{E}_X$  a  $\mathbf{E}_Y$  (porovnateľné s čiastočnými polami  $Z_X$  a  $Z_Y$ ) tak, že platí

$$\mathbf{E}_S = [\mathbf{E}_X, \mathbf{E}_Y]^T = [\mathbf{A}\mathbf{T}, \mathbf{A}\Phi\mathbf{T}]^T, \quad (1.18)$$

kde je vidno, že platí

$$\Re\{\mathbf{E}_X\} = \Re\{\mathbf{E}_Y\} = \Re\{\mathbf{A}\}. \quad (1.19)$$

Keďže  $\mathbf{E}_X$  a  $\mathbf{E}_Y$  zdieľajú spoločné stĺpcové priestory, rozmer  $\mathbf{E}_{XY} = [\mathbf{E}_X | \mathbf{E}_Y]$  je  $d$ , čo znamená, že existuje jedinečná matica  $\mathbf{F}$ , pre ktorú platí

$$\mathbf{0} = [\mathbf{E}_X | \mathbf{E}_Y] \mathbf{F} = \mathbf{E}_X \mathbf{F}_X + \mathbf{E}_Y \mathbf{F}_Y, \quad (1.20)$$

$$= \mathbf{A} \mathbf{T} \mathbf{F}_X + \mathbf{A} \Phi \mathbf{T} \mathbf{F}_Y. \quad (1.21)$$

Definujeme

$$\Psi = -\mathbf{F}_X [\mathbf{F}_Y]^{-1}, \quad (1.22)$$

potom sa rovnica 1.21 dá znázorniť ako

$$\mathbf{A} \mathbf{T} \Psi = \mathbf{A} \Phi \mathbf{T} \Rightarrow \mathbf{A} \mathbf{T} \Psi \mathbf{T}^{-1} = \mathbf{A} \Phi \quad (1.23)$$

Podľa [7], predpoklad, že matica  $\mathbf{A}$  má plnú hodnotu (najväčšia možná hodnota) naznačuje

$$\mathbf{T} \Psi \mathbf{T}^{-1} = \Phi. \quad (1.24)$$

Preto charakteristické hodnoty (*eigenvalues*)  $\Psi$  musia byť rovnaké ako diagonálne elementy  $\Phi$ , a stĺpce  $\mathbf{T}$  sú charakteristické vektory  $\Psi$ . Parametre signálu sú potom získané ako nelineárne funkcie charakteristických hodnôt operátora  $\Psi$ , ktorý rotuje jednu sadu vektorov ( $\mathbf{E}_X$ ) rozprestrených v  $m$ -rozmernom podpriestore signálu do inej sady ( $\mathbf{E}_Y$ ).

## Zhrnutie algoritmu ESPRIT

1. Získanie  $\mathbf{R}_{ZZ}$  z meraní  $\mathbf{Z}$
2. Výpočet všeobecného rozkladu  $\{\mathbf{R}_{ZZ}, \Sigma_n\}$  na charakteristické hodnoty (*general eigendecomposition*)  $\mathbf{R}_{ZZ} \bar{\mathbf{E}} = \bar{\mathbf{E}} \Sigma_n \Lambda$  kde  $\Lambda = \text{diag}\{\lambda_1, \dots, \lambda_{2m}\}$ ,  $\lambda_1 \geq \dots \geq \lambda_{2m}$ , a  $\bar{\mathbf{E}} = [e_1 | \dots | e_{2m}]$ .
3. Určenie počtu zdrojov  $d$
4. Určenie podpriestoru signálu  $S_Z = \Re\{\mathbf{E}_S\}$ , a jeho rozloženie na  $E_X$  a  $E_Y$ , kde platí

$$\mathbf{E}_S = \Sigma_n [e_1 | \dots | e_d] \Rightarrow [\mathbf{E}_X, \mathbf{E}_Y]^T.$$

5. Výpočet rozkladu na charakteristické hodnoty ( $\lambda_1 \geq \dots \geq \lambda_{2d}$ )  $\Sigma$ ,

$$\mathbf{E}_{XY}^* \mathbf{E}_{XY} = [\mathbf{E}_X^*, \mathbf{E}_Y^*]^T [\mathbf{E}_X | \mathbf{E}_Y] = \mathbf{E} \Lambda \mathbf{E}^*,$$

a rozdelenie  $\mathbf{E}$  na podmatice  $d \times d$ ,

$$\mathbf{E} = \begin{bmatrix} \mathbf{E}_{11} & \mathbf{E}_{12} \\ \mathbf{E}_{21} & \mathbf{E}_{22} \end{bmatrix}$$

6. Výpočet charakteristických hodnôt (*eigenvalues*) z  $\Psi = -\mathbf{E}_{12} \mathbf{E}_{22}^{-1}$ ,

$$\phi_k = \lambda_k(-\mathbf{E}_{12} \mathbf{E}_{22}^{-1}), \forall k = 1, \dots, d.$$

7. Určenie  $\theta_k = f^{-1}(\phi_k)$ , teda pre určenie DOA  $\theta_k = \sin^{-1}\{c \arg(\phi_k) / (\omega_0 \Delta)\}$ .

### 1.3 Caponova metóda

Táto metóda je často nazývaná aj metóda maximálnej pravdepodobnosti (maximum likelihood method - MLM)[2]. Jedná sa o hľadanie najvyššieho spektrálneho výkonu vlnoplôch prichádzajúcich z určitých smerov  $\theta$ . Táto metóda zahŕňa hľadanie cez všetky možné DOA (podľa veľkosti kroku) od  $-\pi/2$  do  $\pi/2$ . Predpokladajme  $n$  úzkopásmových vlnoplôch so spoločným stredným kmitočtom  $\omega$ . Každá je definovaná svojím azimutom a eleváciou. V tejto práci sa venujeme len určeniu azimutu, preto učenie elevácie ďalej vynechám. Znova využijeme rovnicu 1.9 ako výstup mikrofónového poľa, konkrétne maticu  $\mathbf{A}(\theta) = [\vec{a}(\theta_1), \dots, \vec{a}(\theta_d)]$ . Výkon z určitého smeru získame podľa

$$f_C(\theta) = [\vec{a}^H(\theta)R\vec{a}(\theta)]^{-1}, \quad (1.25)$$

kde  $R$  je kovariančná matica nameraných hodnôt a  $\vec{a}(\theta)$  je smerovací vektor [4]. Tento výkon určíme pre všetky  $\theta \in [-\pi/2, \pi/2]$  a smer, v ktorom je najvyšší určíme ako DOA.

### 1.4 MUSIC

MUSIC - MUltiple SIgnal Classification bol najpoužívanejší algoritmus na určenie parametrov signálu pred algoritmom ESPRIT. Bol vyvinutý v Lincolnových laboratóriách na americkej M.I.T. (Massachusetts Institute of Technology) v roku 1977. Hlavný prielom nastal v tom, že algoritmus MUSIC dokázal pracovať s ľubovoľným senzorovým poľom.

Pripomeňme si dátový model bez šumu  $\vec{x}(t) = \mathbf{A}(\theta)s(t)$ . Pre azimutový odhad DOA (len azimut bez elevácie) je rôznorodosť poľa jednoparametrová rôznorodosť, ktorú môžeme zobrazit ako *lano* prepletené subpriestorom  $C^m$  [7]. Kľúčovým pozorovaním je, že ak  $\mathbf{x}(t) = \mathbf{a}(\theta)s_\theta(t)$  je vhodný dátový model (bez šumu) pre jeden signál, dáta sú obmedzené jednorozmerným subpriestorom  $C^m$  charakterizovaným vektorom  $\mathbf{a}(\theta)$ . Pre  $d$  signálov sú dátové vektory obmedzené  $d$ -rozmerným subpriestorom  $C^m$  označovaným ako signálový subpriestor ( $S_x$ ). Keď pozorujeme  $d$  nezávislých vektorov,  $S_x$  je známe a prieniky pozorovaného subpriestoru a rôznorodosťou poľa, ktorá prekenuje pozorovaný subpriestor, parametry signálu sa určia hneď. Problém však nastane, ak máme k dispozícii len zašumené údaje  $\vec{x}(t) = \mathbf{A}(\theta)s(t) + \mathbf{n}(t)$ , pretože sa musí odhadnúť  $S_x$ .

Na odhadnutie subpriestoru signálu sa najčastejšie používa kritérium najmenších štvorcov. Cieľom je nájsť sadu  $d$  vektorov prekenujúcich subpriestor  $C^m$ , ktorá najlepšie vyhovuje pozorovaným dátam. Ak sú signály modelované ako nehybné stochastické procesy, predpokladáme že sú nekorelované so šumom a majú kladnú

kovariančnú maticu  $\mathbf{R}_{SS} \geq 0$ . Ak sa jedná o deterministický signál, predpokladá sa, že existuje predpísaná trvalá excitačná podmienka, t.j.

$$\mathbf{R}_{SS} = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N \mathbf{s}(t) \mathbf{s}^*(t)$$

a tá je kladne konečná. Podľa hore uvedených skutočností je potom kovariančná matica zmeraných údajov daná ako

$$\mathbf{R}_{XX} = E\{\mathbf{x}\mathbf{x}^*\} = \mathbf{A}\mathbf{R}_{SS}\mathbf{A}^* + \sigma^2 \sum_n \mathbf{e}_n. \quad (1.26)$$

Úlohou je nájsť sadu  $d$  lineárne nezávislých vektorov obsiahnutú v  $S_x = \Re\{\mathbf{A}\}$ . Jednou takou sadou je  $\mathbf{E}_S = \sum_n [\mathbf{e}_1 | \dots | \mathbf{e}_d]$ , kde  $\{\mathbf{e}_i, i = 1, \dots, d\}$  je sada charakteristických vektorov  $\mathbf{R}_{XX}$ . Vo veľkom množstve prípadov nie sú potrebné kovariančné matice známe a preto sa musia určiť z meraných hodnôt. Šumová kovariančná matica sa dá určiť z meraní, kde neboli prítomné užitočné signály.

Aj s perfektnou znalosťou signálového subpriestoru je hľadanie  $d$  prienikov rôznorodosti a  $S_x$  dosť náročné, hlavne ak chceme okrem azimutu určiť aj eleváciu a vzdialenosť. Nasledujúca funkcia ukazuje jedno možné meranie blízkosti elementu rôznorodosti k  $\hat{S}_x$  [7] :

$$\mathbf{P}_M(\theta) = \frac{\mathbf{a}^*(\theta)\mathbf{a}(\theta)}{\mathbf{a}^*(\theta)\mathbf{E}_N\mathbf{E}_N^*\mathbf{a}(\theta)}. \quad (1.27)$$

V neprítomnosti šumu je toto tzv. spektrum MUSIC-u nekonečné pre hľadané prieniky, v prítomnosti šumu má vrcholy v miestach, ktoré sú si najbližšie. Toto sa využíva na určenie parametrov signálu majúcich  $d$  najvyšších vrcholov v spektre. Avšak, zjavne najjednoduchšie jednorozmerné určovanie má aj nevýhody. Hlavný problém vyvstáva z faktu, že ak máme  $d$  prítomných známych signálov,  $d$  odhadov parametrov,  $\{\theta_1 | \dots | \theta_d\}$ , by sa malo hľadať skorej simultánne než zisťovaním po jednom ako v prípade hľadania pomocou  $\mathbf{P}_M(\theta)$ .

## Zhrnutie algoritmu MUSIC

1. Nazbieranie dát a odhad  $\mathbf{R}_{XX} = E\{\mathbf{x}\mathbf{x}^*\} = \mathbf{A}\mathbf{R}_{SS}\mathbf{A}^* + \sigma^2 \sum_n \mathbf{e}_n$
2. Riešenie pre charakteristický systém;  $\mathbf{R}_{XX}\mathbf{E} = \sum_n \mathbf{E}\Lambda$ , kde  $\Lambda = \text{diag}\{\lambda_1, \dots, \lambda_m\}$ ,  $\lambda_1 \geq \dots \geq \lambda_m$ , a  $\mathbf{E} = [\mathbf{e}_1 | \dots | \mathbf{e}_m]$ .
3. Odhad počtu zdrojov  $d$ .
4. Výpočet

$$\mathbf{P}_M(\theta) = \frac{\mathbf{a}^*(\theta)\mathbf{a}(\theta)}{\mathbf{a}^*(\theta)\mathbf{E}_N\mathbf{E}_N^*\mathbf{a}(\theta)},$$

kde  $\mathbf{E}_N = \sum_n [\mathbf{e}_{d+1} | \dots | \mathbf{e}_m]$ .

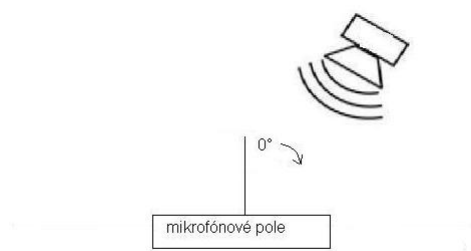
5. Nájdanie  $d$  (najvyšších) vrcholov  $\mathbf{P}(\theta)$  na určenie parametrov.

## 2 EXPERIMENTÁLNE OVERENIE PRESNOSTI ALGORITMOV

Experimentálne overenie presnosti algoritmov som vykonával v akustickom laboratóriu a v mŕtvej komore VUT v Brne.

### 2.1 Pole zložené z 5 mikrofónov

Tento experiment prebiehal v mŕtvej komore. Vytvoril som lineárne pole zložené z 5 mikrofónov s rovnakými vzdialenosťami medzi mikrofónmi  $d = 5\text{cm}$ . Na pole som púšťal z reproduktora umiestneného vo vzdialenosti 140cm od mikrofónového poľa. signál sínus s kmitočtom 2kHz s DOA  $10^\circ$  (obr. 2.1). Pre pole bola vytvorená aj teoretická simulácia (viď obr. 1.4). Mikrofóny som si očísloval od 1 do 5, keď 5 leží na strane kde by DOA bola  $90^\circ$ .



Obr. 2.1: Poloha poľa a reproduktora.

Na spracované dáta som ručne použil metódu delay-and-sum beamformingu. Základným problémom bolo určenie oneskorenia, s ktorým dopadne zvuková vlna na jednotlivé mikrofóny. Pri riešení tohto problému som vychádzal zo známej geometrie mikrofónového poľa a zo známych parametrov signálu. Pre použitý pokusný signál sínusového priebehu s kmitočtom 2000 Hz som podľa vzorca  $\lambda = \frac{c}{f}$  vypočítal vlnovú dĺžku signálu ( $c = 344 \text{ m*s}^{-1}$ ,  $f = 2000 \text{ Hz}$ , potom  $\lambda = 0,172 \text{ m}$ ). Počet vzoriek na jednu periódu signálu som vypočítal zo vzťahu  $n = \frac{f_{vz}}{f}$ , kde  $n$  je počet vzoriek na jednu periódu,  $f_{vz}$  je vzorkovací kmitočet a  $f$  je kmitočet vzorkovaného signálu ( $f_{vz} = 48000 \text{ Hz}$ ,  $f = 2000 \text{ Hz}$ , potom  $n = 24$  vzoriek na jednu periódu). Keďže vzdialenosti medzi mikrofónmi sú 5 cm, teda najväčší posun signálu môže byť 5 cm (pri uhle dopadu  $\theta = 0^\circ$ ), najväčšie oneskorenie medzi dvomi susednými mikrofónmi môže byť o 7 vzoriek (jedna vzorka znamená posun o 7,1 mm). Časový posun zodpovedajúci jednotlivým oneskoreniam podľa vzoriek som zistil z časového rozdielu

medzi vzorkami, teda zo vzorkovacej frekvencie. Uhol dopadu pre jednotlivé oneskorenia som získal zo vzťahu  $\theta = \arccos(\frac{c\delta t}{d})$ , kde  $c$  je rýchlosť zvuku,  $\delta t$  je časový posun medzi mikrofónmi a  $d$  je vzdialenosť medzi mikrofónmi). Zistené hodnoty sú uvedené v tabuľke 2.1. Mikrofónové pole ležalo vo vzdialenom poli zdroja. Pre jednotlivé

Tab. 2.1: Časové posuny a uhly dopadu pre jednotlivé posunutia vzoriek medzi dvomi susednými mikrofónmi.

$n_p[-]$	$t[\mu s]$	$\theta[^\circ]$
0	0	0
1	20,83	9,2
2	41,1	16,7
3	62,5	25,5
4	83,33	35
5	104,17	45,8
6	125	59,3
7	145,83	86,1

dvojice susedných mikrofónov som vypočítal uhol dopadu rovinatej vlny podľa vzorkového oneskorenia a prevodom cez tabuľku 2.1. Výsledky sú uvedené v tabuľke 2.2. Medzi 3. a 4. mikrofónom bol iný posun ako u ostatných dvojíc. Táto skutočnosť je

Tab. 2.2: Vzorkové oneskorenie a uhol dopadu pre jednotlivé dvojice mikrofónov.

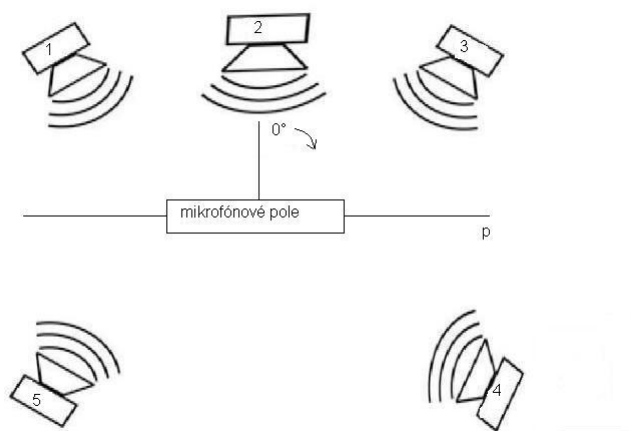
dvojica mikrofónov	vzorkový posun	uhol dopadu $\theta[^\circ]$
1-2	1	9,2
2-3	1	9,2
3-4	2	16,7
4-5	1	9,2

pravdepodobne spôsobená nepresným sformovaním mikrofónového poľa alebo zlým odhadom maxima korelačnej funkcie signálov zaznamenaných týmito mikrofónmi a následným vzorkovaním.

## 2.2 Pole zložené z 8 mikrofónov

Moja práca spočívala v umiestnení mikrofónov do mikrofónového poľa a následnom prehrávaní rozličných zvukových signálov z reproduktorov rozmiestnených na

kružnici s polomerom  $r = 2,3$  m. Mikrofónové pole bolo umiestnené v strede tejto kružnice. Jednotlivé reproduktory boli umiestnené podľa obrázku 2.2.



Obr. 2.2: Usporiadanie reproduktorov.

Uhly dopadu pre jednotlivé reproduktory sú uvedené v tabuľke 2.3.

Tab. 2.3: Uhly dopadu pre jednotlivé reproduktory.

číslo reproduktoru	uhol dopadu[°]
1	315 (-45)
2	0
3	45
4	135
5	225

Do reproduktorov som pustil 4 rôzne signály: úzkopásmový sínus ( $f = 2$  kHz,  $f = 4$  kHz), ľudskú reč (ženský hlas vyslovujúci slová "Zajtra musím pracovať") a hudobný signál. Signály som prehrával postupne z každého reproduktoru a zachytával mikrofónovým poľom. Výstup mikrofónového poľa som nahral pomocou programu Cubase 4 a uložil vo formáte ".wav" pre následné spracovanie programom Matlab. Na určovanie DOA som použil algoritmy vytvorené R. Mosesom v roku 1996, ktoré som mierne upravil pre vlastné potreby. V Matlabe som nahraté signály spracoval do vhodnej podoby a použil som na ne algoritmy na určenie DOA uvedené v teoretickej časti. Výsledky sú zaokrúhlené na desatiny stupňa, pretože na presnejšie výsledky by bolo treba veľmi presne zamerať polohy reproduktorov a vzdialenosti mikrofónov od seba.

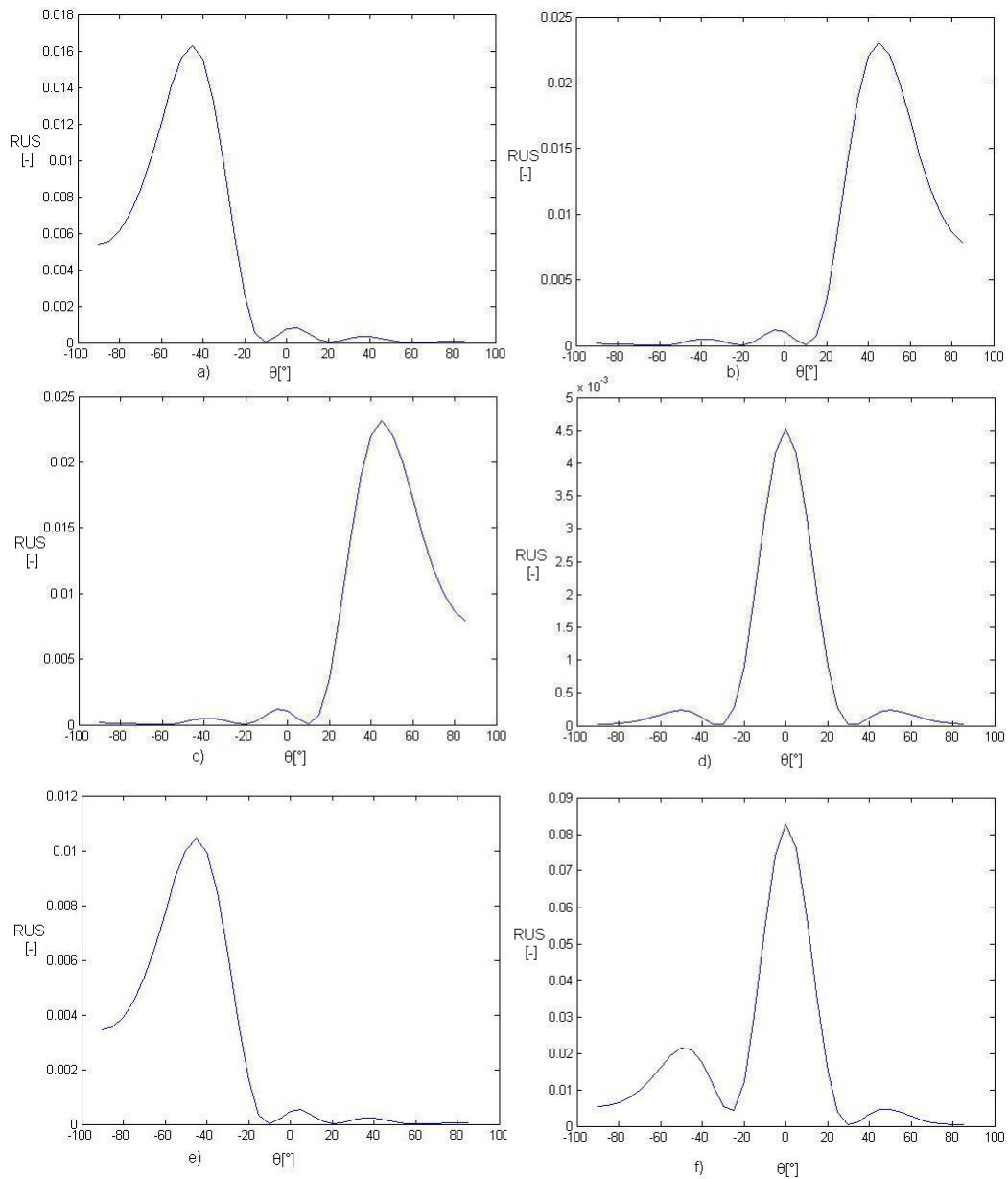
### 2.2.1 Lineárne pole

Použil som lineárne equidistantné mikrofónové pole zložené z 8 mikrofónov Behringer ECM 8000 so vzdialenosťou susedných mikrofónov  $d = 4$  cm (obr. 1.1). Ani jedna z metód beamforming a Capon nerozlišovala, či signál prichádza zpredu alebo zozadu. V prípade reproduktorov 4 a 5 algoritmus zobrazil výsledky tak, ako keby sme zrkadlovo prevrátili usporiadanie reproduktorov cez os p. Teda pre uhol  $135^\circ$  máme zobrazený uhol  $45^\circ$  a pre  $225^\circ$  ukazuje  $-45^\circ$ . Z uvedeného vyplýva, že pre použitie týchto algoritmov je vhodnejšie umiesniť mikrofónové pole tak, aby sa všetky potenciálne zdroje zvuku nachádzali v rozmedzí  $-\pi/2 \leq \theta \leq \pi/2$ . Obe metódy majú ďalšiu nevýhodu v tom, že hľadajú zdroj len na niekoľkých pozíciách, napríklad s krokom  $5^\circ$ . Ak by sme zvýšili počet krokov a tým znížili rozdiely medzi jednotlivými krokmi, zvýšili by sa výpočetné nároky. Ak zvolíme príliš veľký krok, neurčia nám algoritmy DOA dostatočne presne. Na grafoch vidno, že úzkopásmové signály oba algoritmy určili presne, no so širokopásmovými signálmi mali problém. Navyše pri hlasovom signále z dvoch reproduktorov súčasne ani jeden z algoritmov neurčil dva rozličné zdroje. Pri úzkopásmových signáloch však lepšie pôsobil algoritmus capon, ktorý mal oproti beamformingu strmšie priebehy v okolí skutočného DOA. Na obrázkoch 2.3 a 2.4 sú výsledky pre jednotlivé mikrofóny (DOA je pre a)  $-45^\circ$ , b)  $45^\circ$ , c)  $135^\circ$ , d)  $0^\circ$ , e)  $225^\circ$  a f)  $0^\circ$  a  $-45^\circ$ ). Z obrázkov je viditeľný strmší nábeh pre metódy Capon a žiadne zákmity mimo DOA. Z tohto dôvodu pokladám metódu Capon za výhodnejšiu. Výsledky pre ostatné signály sú uvedené v prílohe na obr. A.1 až A.6.

Algoritmy ESPRIT a MUSIC určili DOA úzkopásmových signálov veľmi presne, s maximálnou nepresnosťou  $0,6^\circ$  (sínus 2kHz) a  $0,3^\circ$  (sínus 4kHz). Rozptyl nepresností pri signále sínus s kmitočtom 2kHz bol pre ESPRIT 0,06 a pre MUSIC 0,008, smerodatné odchýlky boli pre ESPRIT 0,2449 a pre MUSIC 0,0894 a medián nepresností bol 0,3 pre ESPRIT a 0,16 pre MUSIC. Pri signále sínus s kmitočtom 4kHz bol rozptyl nepresností 0,1127 pre ESPRIT a 0,0024 pre MUSIC, smerodatné odchýlky 0,3357 pre ESPRIT a 0,0488 pre MUSIC a medián nepresností bol 0,2125 pre ESPRIT a 0,0714 pre MUSIC. Z uvedeného vyplýva, že MUSIC má nižšie nepresnosti ako aj ich rozptyl. Zistené hodnoty sú zhrnuté v tabulkách 2.4 a 2.5.

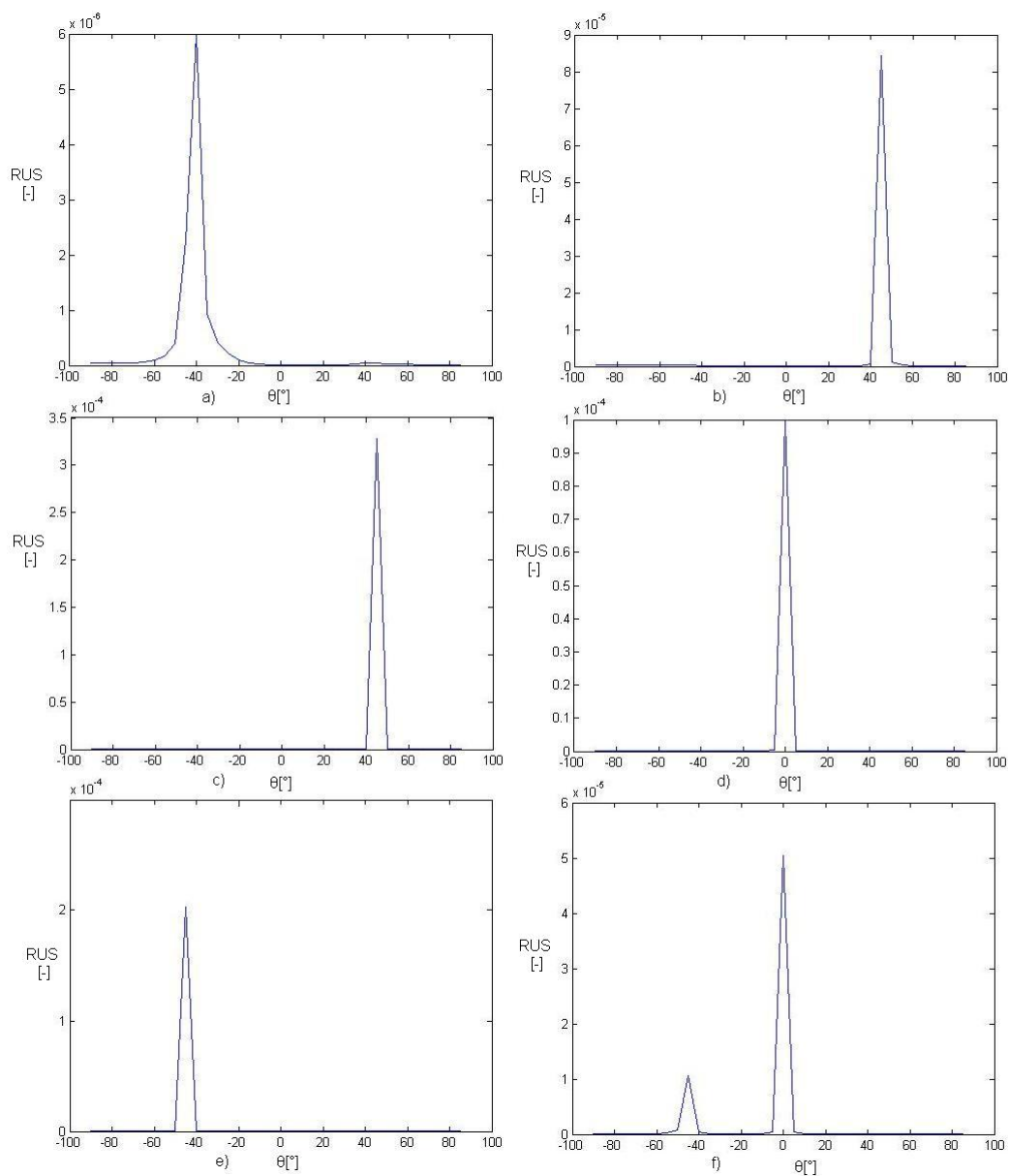
V určení DOA úzkopásmového signálu nebol problém. Oba algoritmy boli veľmi presné, no MUSIC má vyššiu výpočetnú náročnosť. Pri výbere z algoritmov sa treba rozhodnúť, či nám stačí o trochu nižšia presnosť ESPRITU, alebo chceme vyššiu presnosť MUSIC-u, avšak za vyššiu výpočetnú cenu.

Pri zisťovaní DOA širokopásmových signálov boli oba algoritmy menej presné ako v prípade úzkopásmových. V prípade hudobného signálu mal ESPRIT maximálnu odchýlku  $1,1^\circ$ , MUSIC len  $0,6^\circ$ . Oba algoritmy však dokázali určiť dva zdroje



Obr. 2.3: Beamforming pre sínus s kmitočtom 2kHz - lineárne pole

v rozličných polohách (tab. 2.6). V prípade hlasového signálu boli nepresnosti vyššie. ESPRIT mal najvyššiu odchýlku  $2,7^\circ$  pre jeden zdroj a až  $4,7^\circ$  v prípade dvoch zdrojov. MUSIC mal pre jeden zdroj najvyššiu odchýlku  $3,3^\circ$  a v prípade dvoch zdrojov až  $8,1^\circ$  (tab 2.7). Rozptyl nepresností bol  $0,1424$ (hudba) a  $1,9395$ (reč) pre MUSIC a  $0,0433$ (reč) a  $6,8462$ (reč) pre ESPRIT, smerodatná odchýlka  $0,3773$ (hudba) a  $1,3927$ (reč) pre MUSIC a  $0,2082$ (hudba) a  $2,6165$ (reč) pre ESPRIT a stredná hodnota nepresností  $0,4714$ (hudba) a  $1,9571$ (reč) pre MUSIC a  $0,3$ (hudba) a  $2,5571$ (reč)



Obr. 2.4: Metóda Capon pre sínus s kmitočtom 2kHz - lineárne pole

pre ESPRIT.

Z uvedeného vyplýva, že algoritmus MUSIC je vhodnejší pri hudobných signáloch, no na rečové sa viac hodí algoritmus ESPRIT. Presnejšie výsledky by sa dali získať rozložením širokopásmových signálov na niekoľko úzkopásmových a následné určenie DOA, no to by zvýšilo výpočetnú náročnosť.

Tab. 2.4: Určenie uhla dopadu pre signál sínus s kmitočtom 2kHz

uhol dopadu[°]	ESPRIT [°]	MUSIC [°]
-45	-44,6	-44,8
0	-0,4	-0,1
45	45,6	45,3
135	135,1	135,1
225	225,0	224,9
0 a -45	0,2 a -44,7	0,1 a -45,2

Tab. 2.5: Určenie uhla dopadu pre signál sínus s kmitočtom 4kHz

uhol dopadu[°]	ESPRIT [°]	MUSIC [°]
-45	-44,9	-44,9
0	-0,2	0,1
45	45,0	45,1
135	135,3	135,1
225	225,0	225,0
0 a -45	-0,1 a -44,9	0,1 a -45,0

### 2.2.2 Guľové pole

Na tvorbu guľového poľa som použil mikrofóny Behringer ECM 8000 umiestnené v rohoch kocky s hranou  $a = 8\text{cm}$ . Obe metódy vyhodnotili výsledné DOA pre reproduktory 4 a 5 zrkadlovo, ako aj v prípade lineárneho poľa. Bolo to spôsobené tým, že algoritmy boli nastavené na hľadanie zdrojov v rozmedzí  $-\pi/2, \pi/2$ . Po zvýšení hraníc na  $-\pi, \pi$  obe metódy zistili, že signál prichádza zozadu, no nepresnosti dosahovali až  $20^\circ$ . Aj v tomto prípade metóda Capon bola výhodnejšia. Hlavný rozdiel bol pri použití rečového signálu, u ktorého beamforming nedokázal určiť dva zdroje. Výsledky pre signál sínus 2kHz sú na obrázkoch 2.5 a 2.6. Aj v prípade guľového poľa má metóda Capon strmší priebeh a zákmity mimo DOA sú nulové. Výsledky pre ostatné signály sú uvedené v prílohe na obr. A.7 až A.12.

Algoritmy ESPRIT a MUSIC znovu veľmi presne určili DOA úzkopásmových signálov. Rozptyl nepresností bol 0(sínus 2kHz) a 0,0129(sínus 4kHz) pre ESPRIT a 0,1429(sínus 2kHz) a 0,0057(sínus 4kHz) pre MUSIC, smerodatná odchýlka bola 0(sínus 2kHz) a 0,1134(sínus 4kHz) pre ESPRIT a 0,3780(sínus 2kHz) a 0,0756(sínus 4kHz) pre MUSIC a stredná hodnota nepresnosti bola 0(sínus 2kHz) a 0,0429(sínus

Tab. 2.6: Určenie uhla dopadu pre hudobný signál

uhol dopadu[°]	ESPRIT [°]	MUSIC [°]
-45	-45,2	-44,9
0	-0,3	0,2
45	44,1	44,6
135	133,9	135,6
225	224,7	224,6
0 a -45	0,4 a -44,9	0,4 a -45,0

Tab. 2.7: Určenie uhla dopadu pre ženský hlas

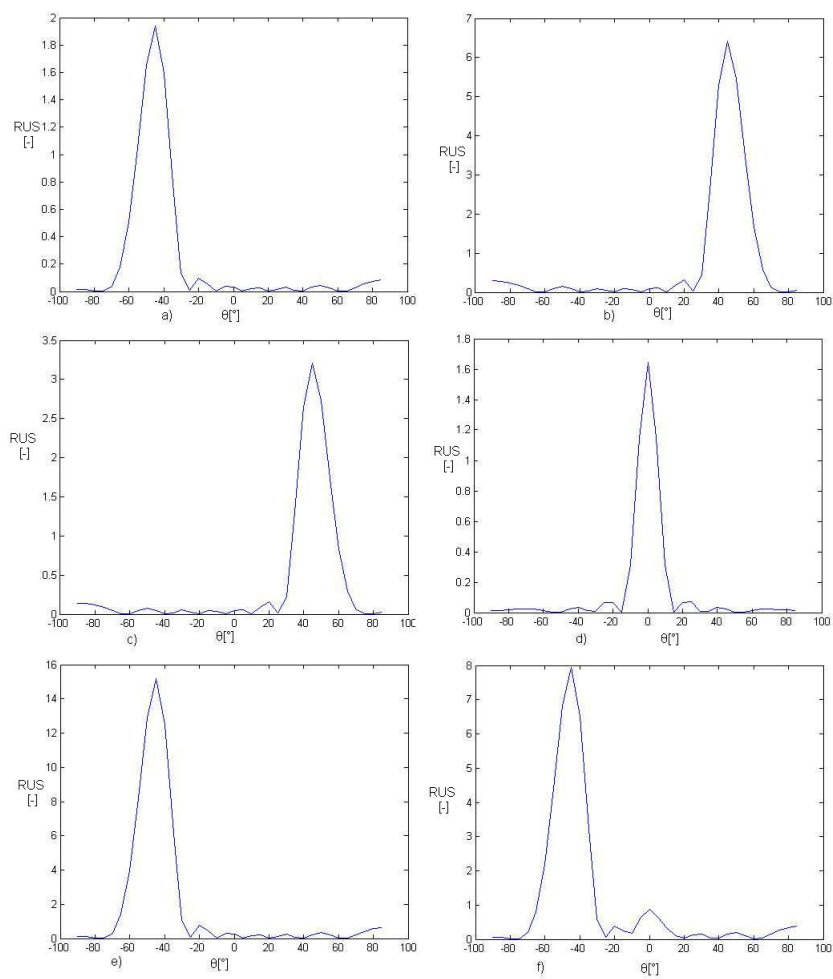
uhol dopadu[°]	ESPRIT [°]	MUSIC [°]
-45	-43,7	-43,5
0	-0,7	-0,2
45	44,2	43,6
135	136,5	136,3
225	222,3	228,3
0 a -45	-4,7 a -43,0	-8,1 a -42,9

4kHz) pre ESPRIT a 0,1429(sínus 2kHz) a 0.0289(sínus 4kHz) pre MUSIC. Zistené hodnoty sú uvedené v tabuľkách 2.8 a 2.9.

Tab. 2.8: Určenie uhla dopadu pre signál sínus s kmitočtom 2kHz pri guľovom poli

uhol dopadu[°]	ESPRIT [°]	MUSIC [°]
-45	-45,0	-45,0
0	0,0	0,0
45	45,0	45,0
135	135,0	134,0
225	225,0	225,0
0 a -45	0,0 a -45,0	-0,0 a -45,0

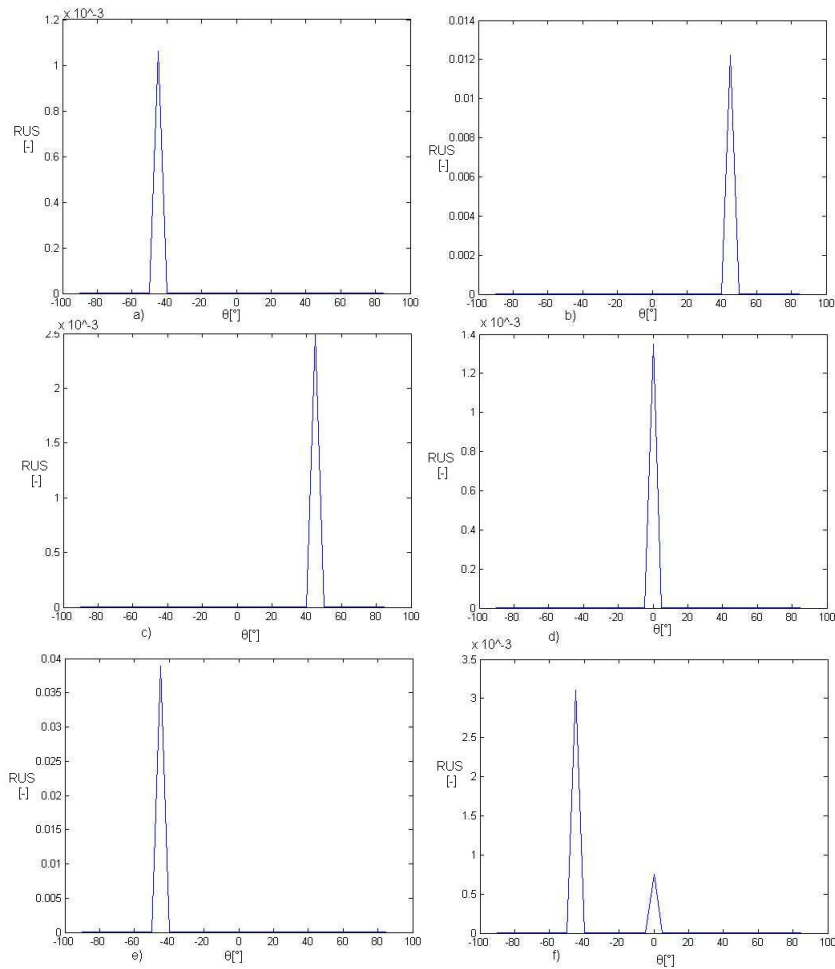
V prípade guľového poľa však boli aj širokopásmové signály určené veľmi presne. Pri hudobnom signále ukazovali oba algoritmy rovnaké výsledky. Rozptyl bol 0,0014, smerodatná odchýlka bola 0,0378 a stredná hodnota nepresností bola 0,0143. Najvyššie nepresnosti sa objavovali v prípade dvoch zdrojov zvuku (tab. 2.10).



Obr. 2.5: Beamforming pre sínus s kmitočtom 2kHz - guľové pole

Tab. 2.9: Určenie uhla dopadu pre signál sínus s kmitočtom 4kHz pri guľovom poli

uhol dopadu[°]	ESPRIT [°]	MUSIC [°]
-45	-45,0	-45,0
0	0,0	0,0
45	45,0	45,0
135	135,0	135,0
225	225,0	225,0
0 a -45	0,3 a -45,0	0,2 a -45,0



Obr. 2.6: Metóda Capon pre sínus s kmitočtom 2kHz - guľové pole

Tab. 2.10: Určenie uhla dopadu pre hudobný signál pri guľovom poli

uhol dopadu[°]	ESPRIT [°]	MUSIC [°]
-45	-45,0	-45,0
0	0,0	0,0
45	45,0	45,0
135	135,0	135,0
225	225,0	225,0
0 a -45	0,0 a -45,1	0,0 a -45,1

V prípade rečového signálu boli nepresnosti vyššie. ESPRIT mal maximálnu odchýlku 1,3°, MUSIC 0,7° (2.11). Rozptyl nepresností bol 0,1762 pre ESPRIT a 0,0595 pre MUSIC, smerodatná odchýlka bola 0,4198 pre ESPRIT a 0,244 pre MU-

SIC a stredná hodnota nepresností bola 0,7429 pre ESPRIT a 0,3429 pre MUSIC. V prípade guľového poľa algoritmy ESPRIT aj MUSIC veľmi presne určili polohu zdroja úzkopásmového i širokopásmového signálu. Pre sínus 2kHz bol presnejší ESPRIT, pre sínus 4kHz a pre rečový signál bol presnejší MUSIC. Pri hudobnom signále boli oba algoritmy rovnako presné. Znovu je treba sa rozhodnúť, aký signál budeme snímať, a či nám stačí presnosť ESPRIT-u s nižšími výpočetnými nárokmi, alebo vyžadujeme vyššiu presnosť MUSIC-u aj za cenu vyšších výpočetných nárokov.

Tab. 2.11: Určenie uhla dopadu pre rečový signál pri guľovom poli

uhol dopadu[°]	ESPRIT [°]	MUSIC [°]
-45	-44,7	-45,0
0	0,8	0,2
45	44,0	44,3
135	134,5	134,7
225	226,1	225,6
0 a -45	-0,2 a -46,3	0,2 a -45,4

### 3 ZÁVER

Porovnal som rôzne metódy lokalizácie akustických zdrojov podľa ich presnosti a výpočetnej náročnosti. Všetky štyri algoritmy (beamforming, Caponova metóda, ESPRIT a MUSIC) veľmi presne určili DOA zdrojov úzkopásmových signálov. Avšak beamforming a Caponova metóda sú len tak presné, ako nastavíme ich krok. Čím menší krok, tým horšia presnosť, i keď výpočetná náročnosť klesne. A naopak, čím väčší bude krok, tým presnejšie nám tieto metódy určia DOA signálu, avšak výpočet zaberie viacej času. Algoritmy ESPRIT a MUSIC boli približne rovnako presné, v prípade guľového pola dokonca úplne rovnaké. Preto je treba sa rozhodnúť, či sa spokojíme s o trochu vyššími nepresnosťami algoritmu ESPRIT pri nižšej výpočetnej náročnosti, alebo požadujeme vyššiu presnosť MUSIC-u aj za cenu vyššej výpočetnej náročnosti.

V prípade širokopásmových signálov sa presnosť určenia DOA znížila, najmä preto, že spektrum širokopásmových signálov bolo širšie a teda niektoré spektrálne zložky mohli spôsobovať aliasing a znižovať presnosť. Preto je výhodnejšie rozložiť spektrum signálu na niekoľko subpásiem a algoritmy potom aplikovať na jednotlivé subpásma. Najväčší problém mali algoritmy pri rečovom signále, ktorý nielen že je širokopásmový, no obsahuje aj medzislovné medzery, ktoré výpočet komplikovali. Beamforming a metóda Capon navyše pri rečovom signále nedokázali rozpoznať dva zdroje (okrem prípadu Capon-guľové pole, no aj tu je sa dá druhý zdroj popliesť s náhodnou nepresnosťou). ESPRIT a MUSIC mali znovu porovnateľné výsledky, no znovu sa treba rozhodnúť, či nám stačí o trochu menšia presnosť ESPRIT-u, alebo potrebujeme vyššiu presnosť MUSIC-u. Ak však predpokladáme prítomnosť rečového signálu, je vhodnejšie použiť MUSIC. Guľové pole malo vo všeobecnosti presnejšie výsledky ako lineárne.

## LITERATURA

- [1] BENESTY, J., CHEN, J., YITENG, H. Microphone Array Signal Processing. Berlin: Springer-Verlag, 2008. 250 s. ISBN 978-3-540-78612-2.
- [2] CAPON, J. High-resolution frequency-wavenumber spectrum analysis. Proceedings of the IEEE. August 1969, vol. 57, no. 8, s. 1408-1418.
- [3] EVEREST, F. A., The Master Handbook of Acoustics. McGraw/Hill, 2009. 528 s. ISBN 978-0071603324.
- [4] HAWKES, Malcolm; NEHORAI, Arye. Acoustic Vector-Sensor Beamforming and Capon Direction Estimation. IEE Transactions on Signal Processing. September 1998, vol. 46, no. 9, s. 2291-2304.
- [5] MARZETTA, Thomas L. A New Interpretation for Capon ´s Maximum Likelihood Method of Frequency-Wavenumber Spectral Estimation. IEEE Transactions on Acoustic, Speech and Signal Processing. April 1983, vol. ASSP-31, no. 2, s. 445-449.
- [6] OPPENHEIM, Alan V.; SCHAFER, Ronald W.; BUCK, John R. Discrete-Time Signal Processing. 2nd ed. New Jersey : Prentice-Hall, 1999. 870 s. ISBN 0-13-754920-2.
- [7] ROY, Richard; KAILATH, Thomas. ESPRIT - Estimation of Signal Parameters Via Rotational Invariance Techniques. IEEE Transactions on Acoustic, Speech and Signal Processing. July 1989, 37, 7, s. 984-995.
- [8] VAN VEEN, Barry D.; BUCKLEY, Kevin M. Beamforming: A Versatile Approach to Spatial Filtering. IEEE ASSP Magazine. 1988, April, s. 4-24.

## SEZNAM PŘÍLOH

A Grafy pre metódy beamforming a Capon	36
B Fotografie usporiadania mikrofónových polí	49

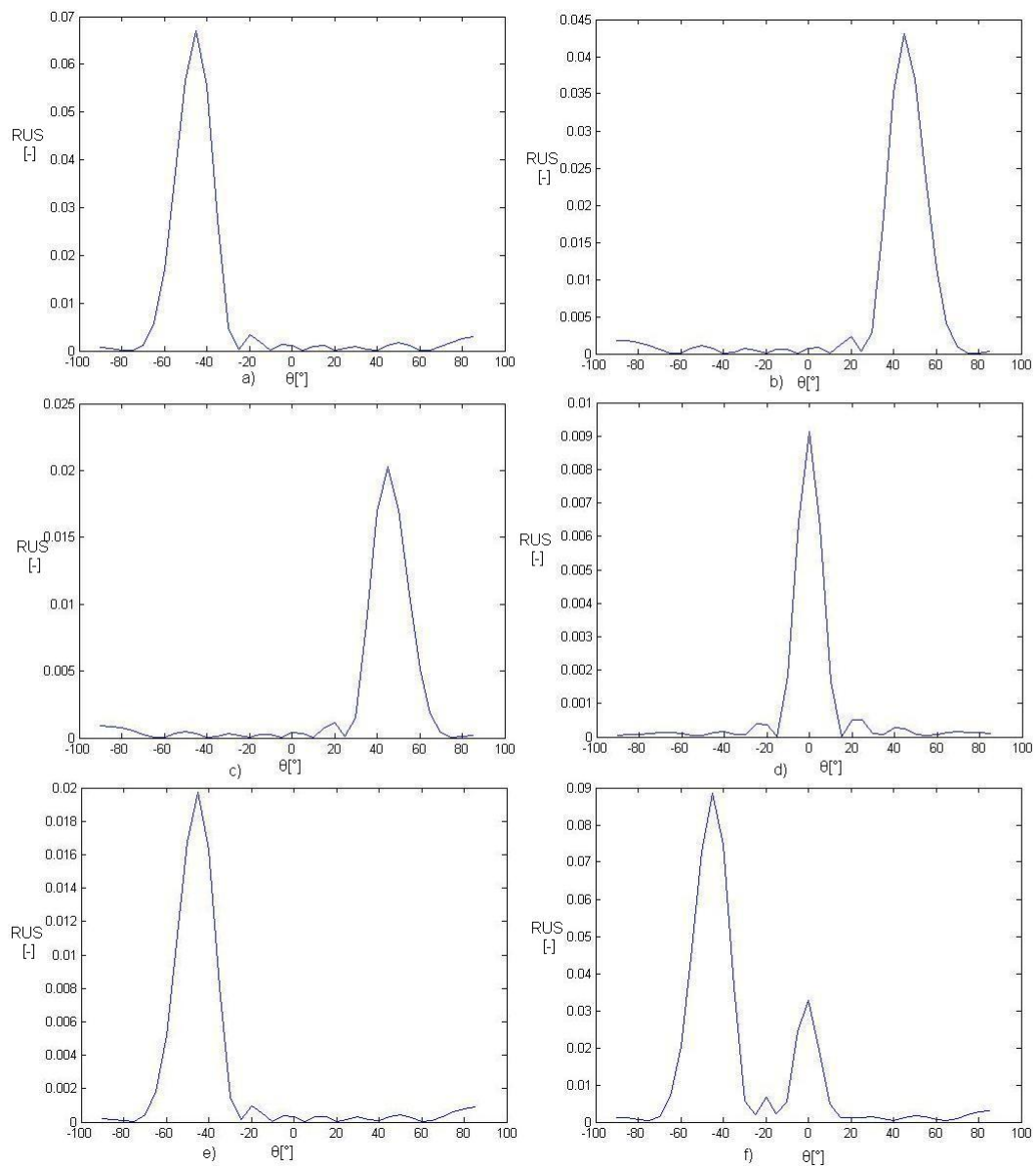
# A GRAFY PRE METÓDY BEAMFORMING A CAPON

Grafy zobrazujúce určenie DOA metódami beamforming a Capon. Vo všetkých obrázkoch platí nasledovné:

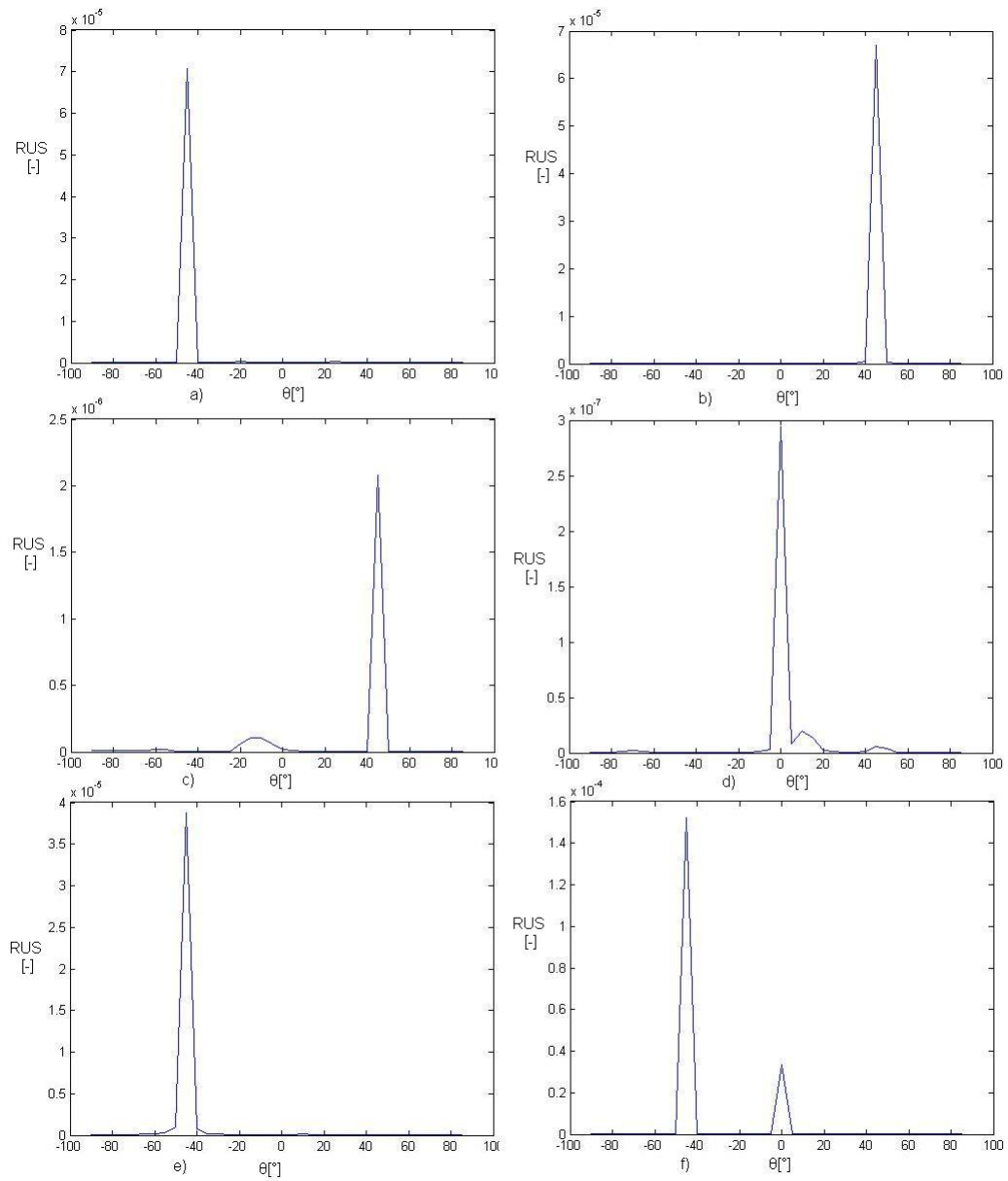
Tab. A.1: Tabuľka určujúca skutočné uhly dopadu pre jednotlivé časti obrázkov.

	uhol dopadu [°]
a)	315 (-45)
b)	45
c)	135
d)	0
e)	225
f)	0 + 315

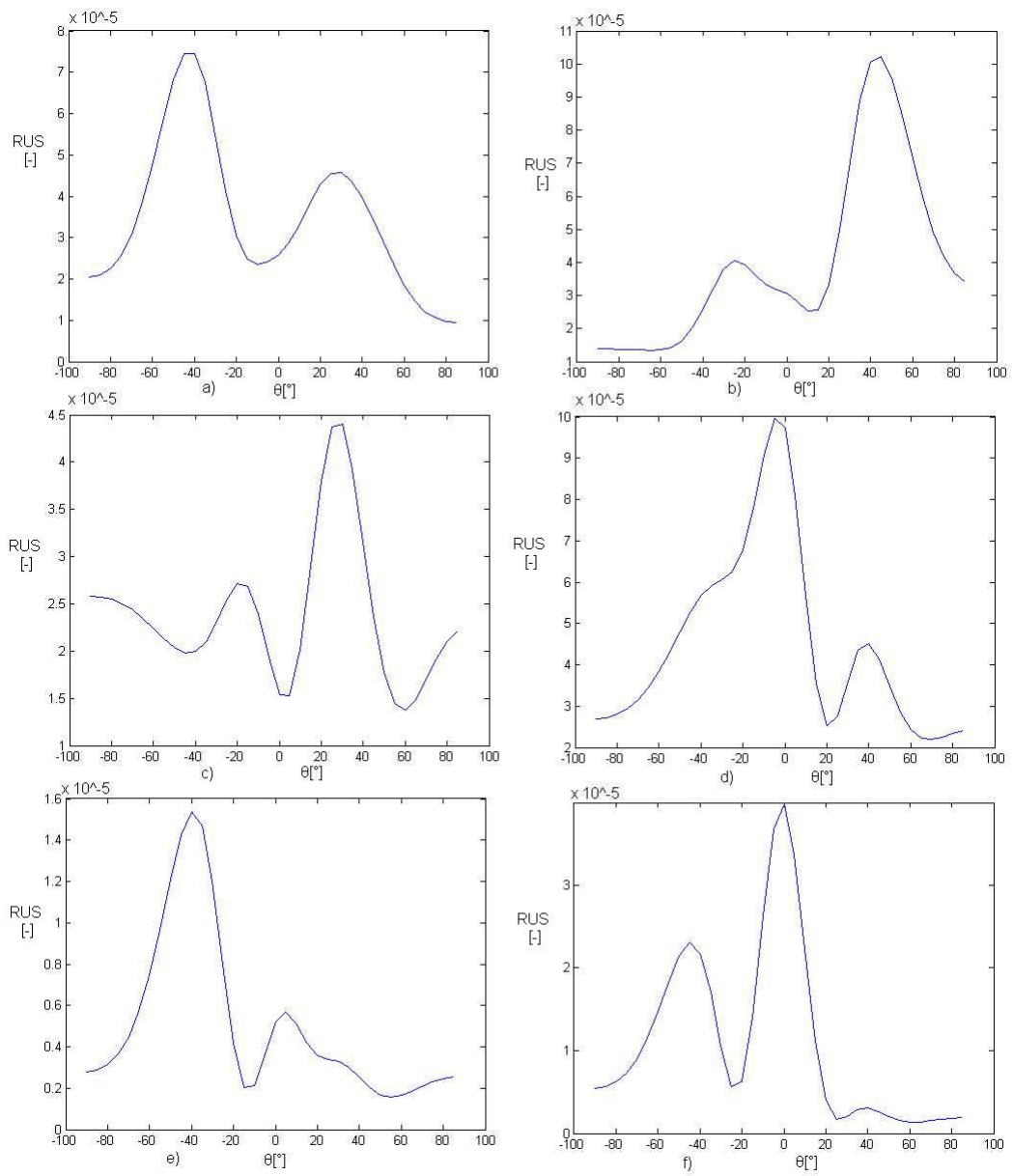
RUS je relatívna úroveň signálu (bezrozmerná).



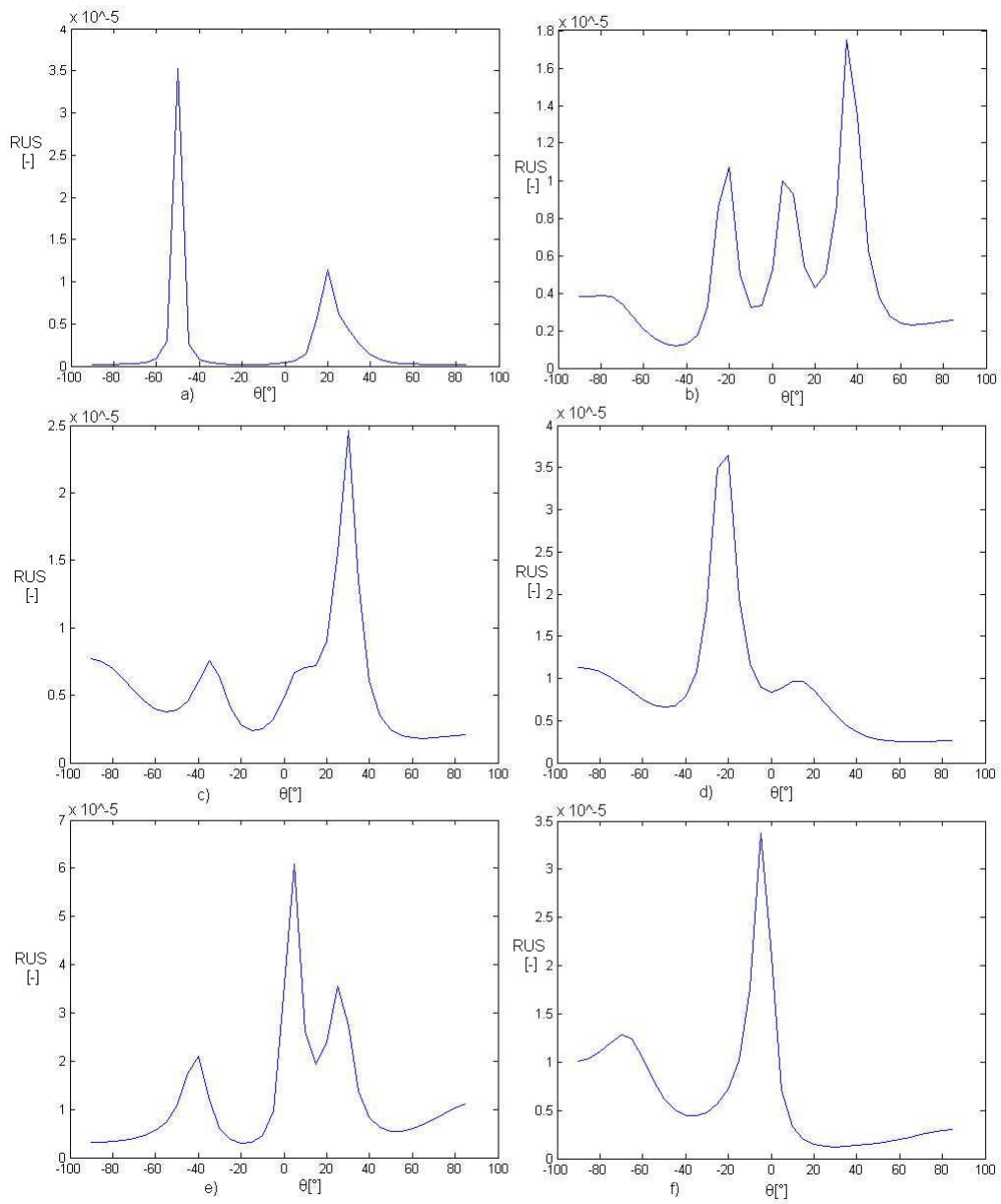
Obr. A.1: Beamforming pre sínus s kmitočtom 4kHz - lineárne pole



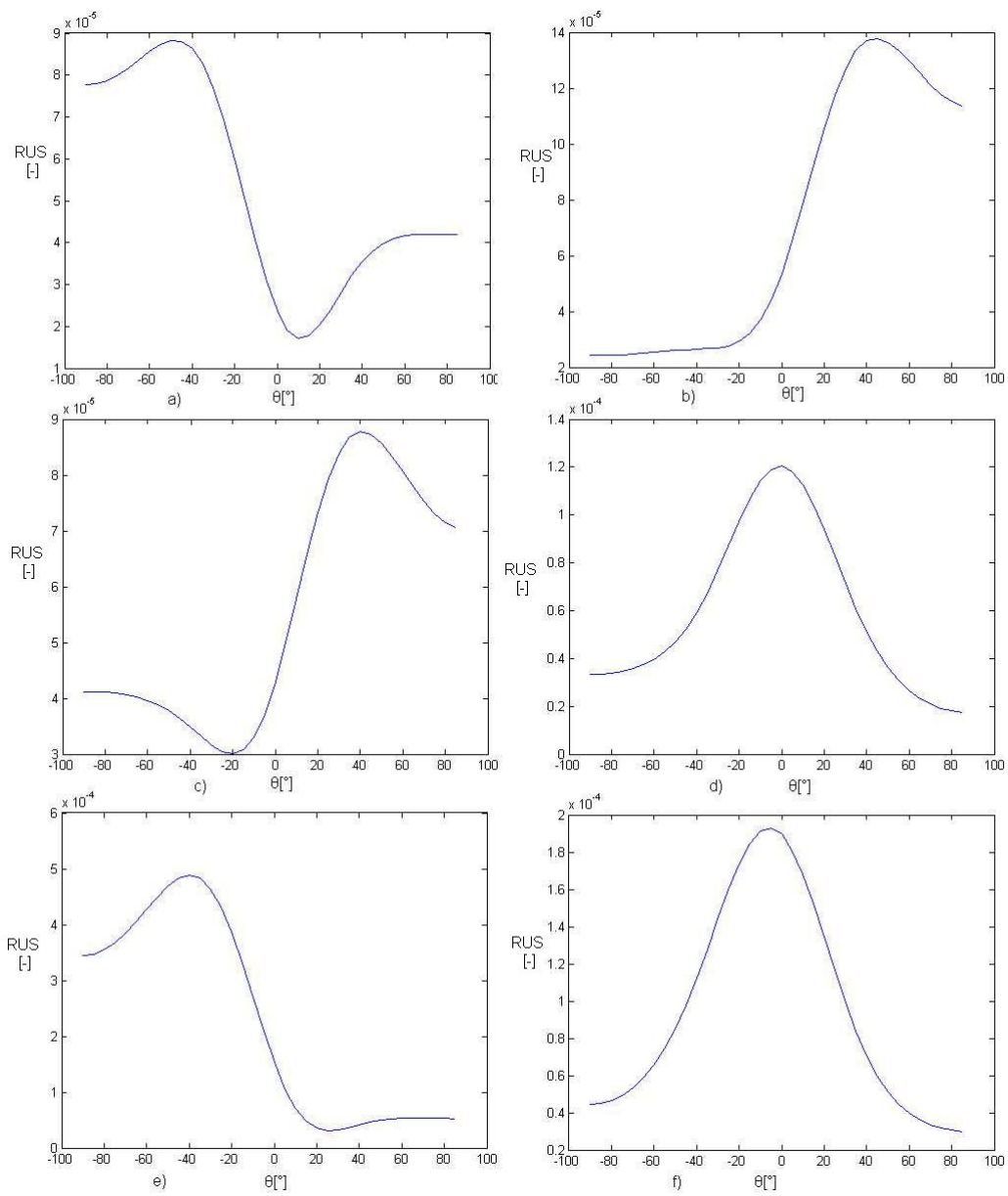
Obr. A.2: Metóda Capon pre sínus s kmitočtom 4kHz - lineárne pole



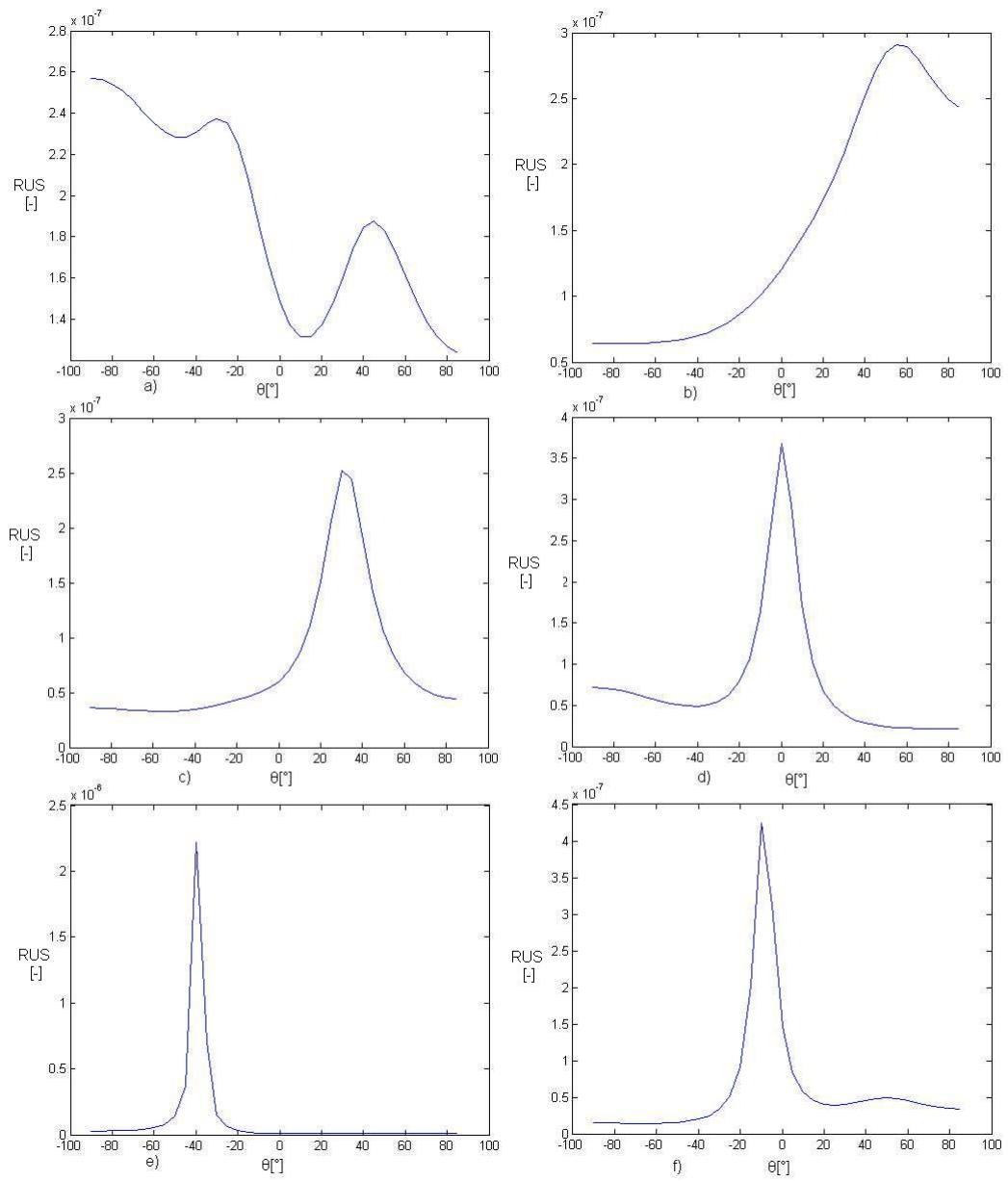
Obr. A.3: Beamforming pre hudobný signál - lineárne pole



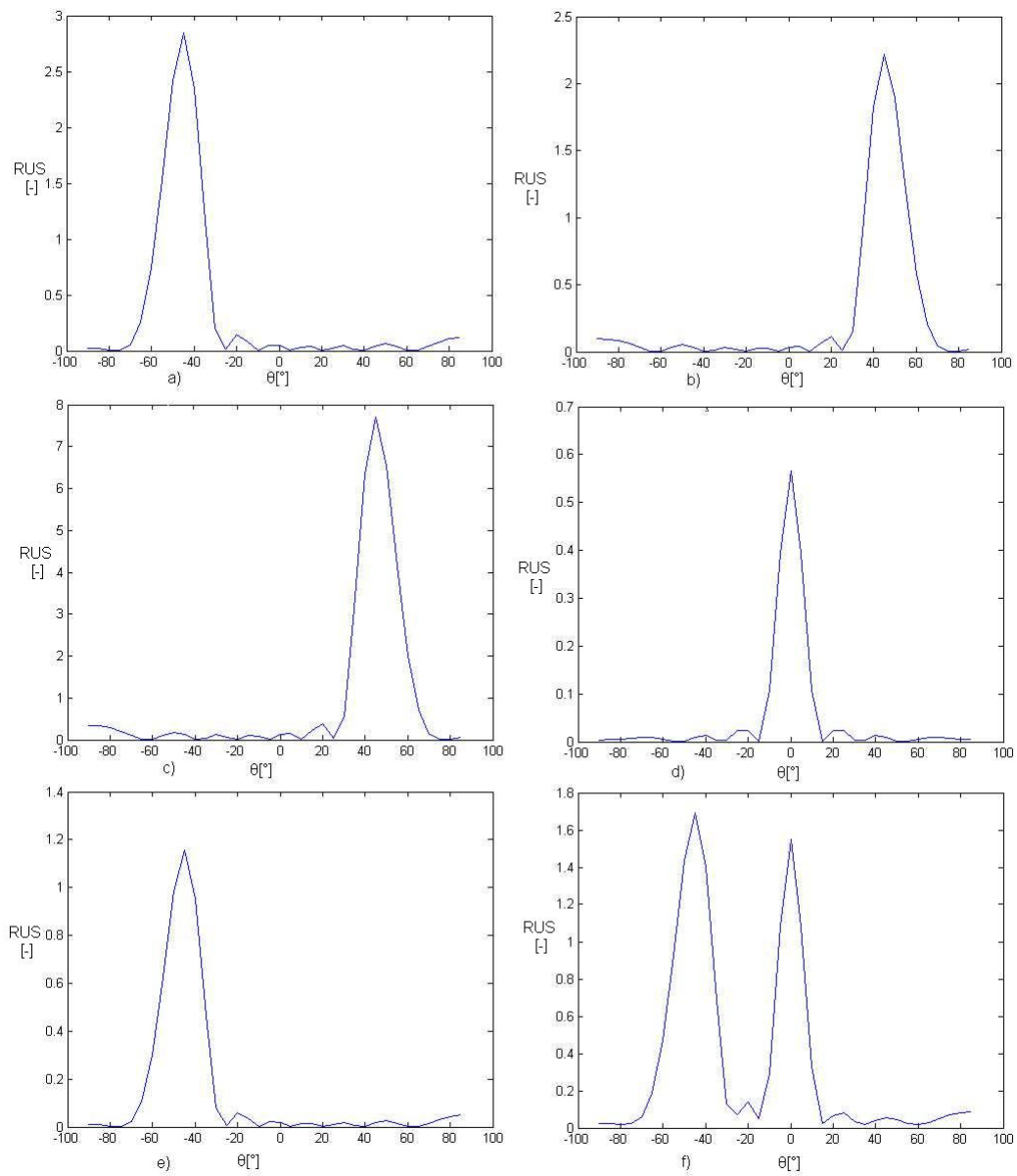
Obr. A.4: Metóda Capon pre hudobný signál - lineárne pole



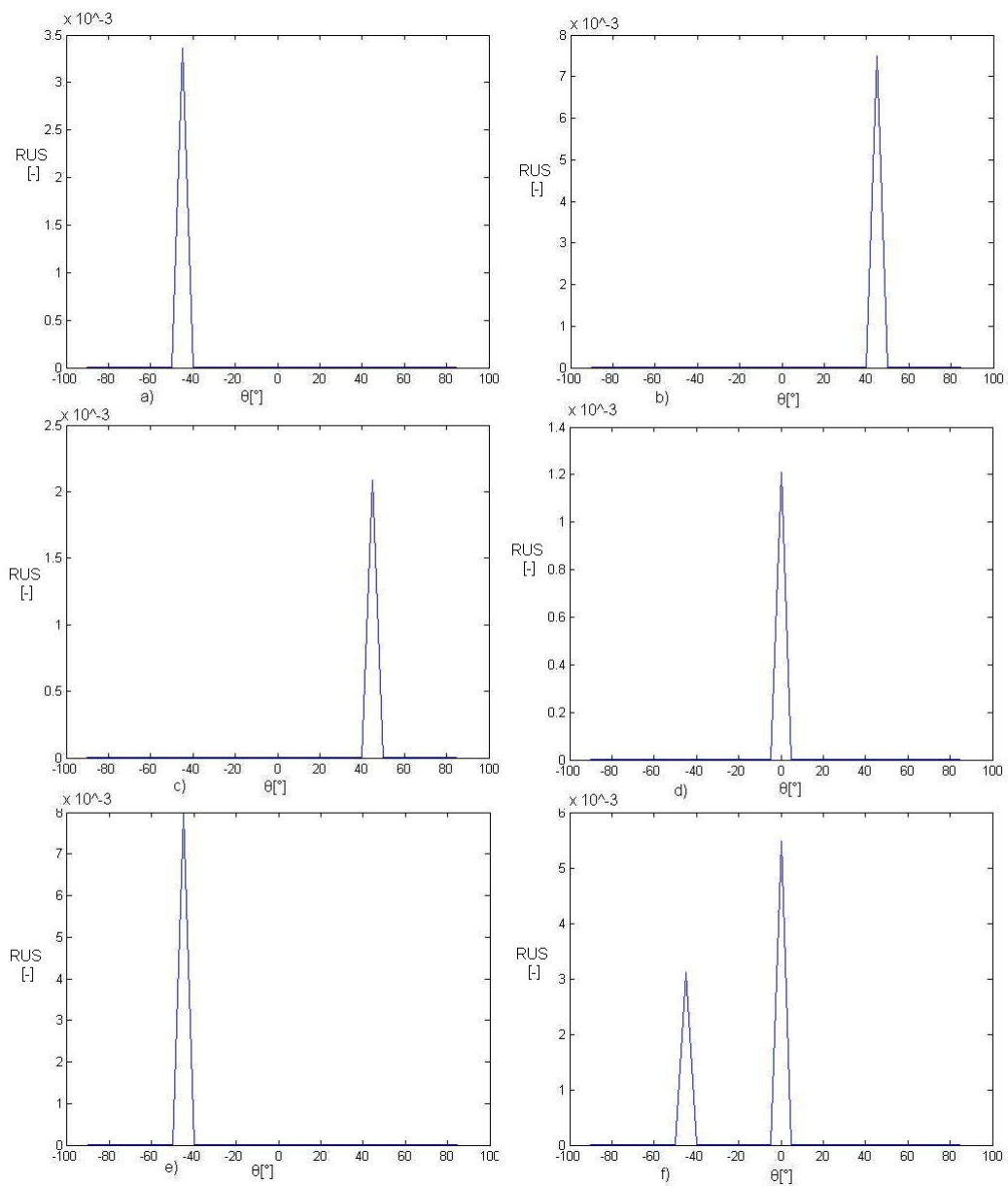
Obr. A.5: Beamforming pre rečový signál - lineárne pole



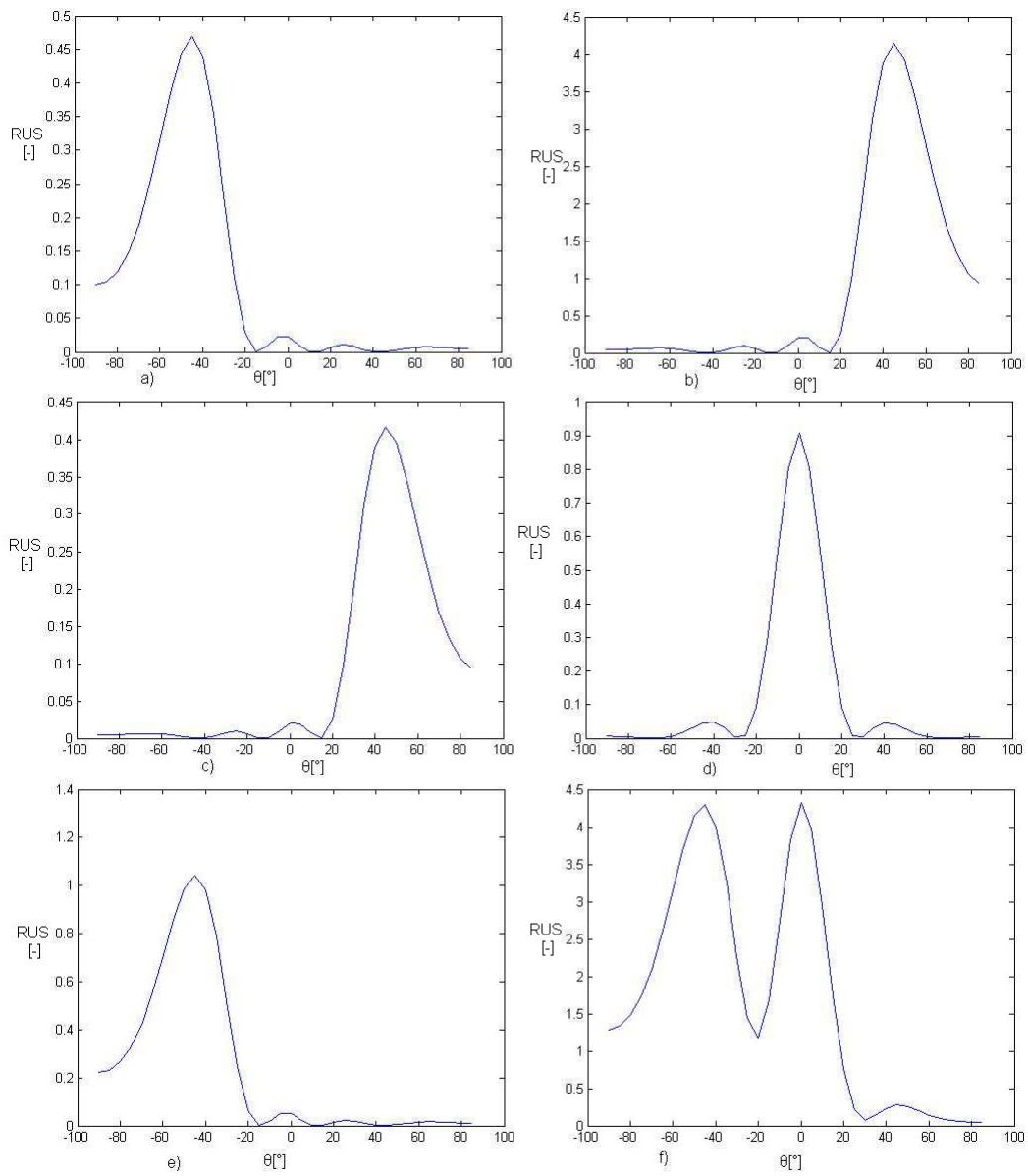
Obr. A.6: Metóda Capon pre rečový signál - lineárne pole



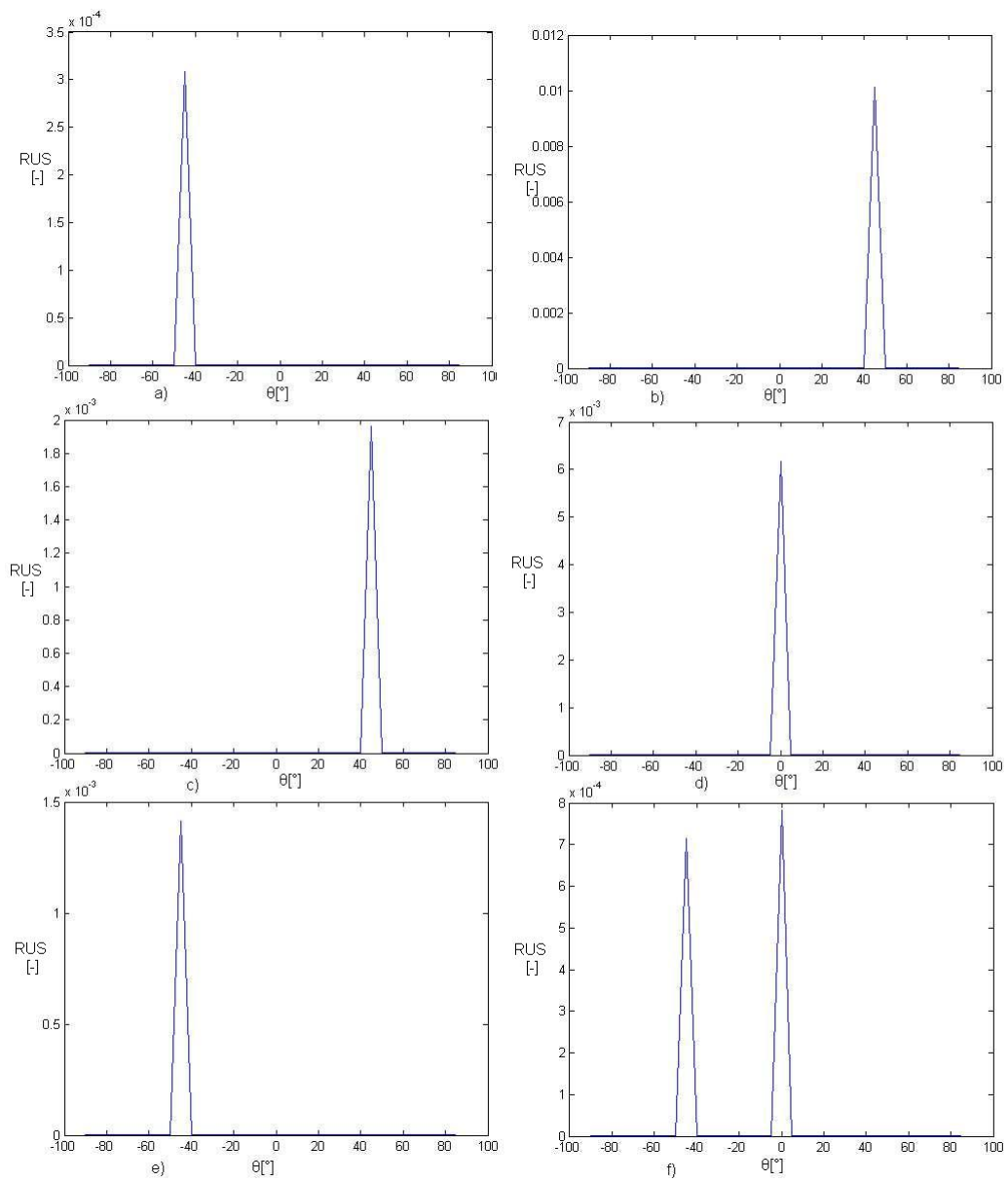
Obr. A.7: Beamforming pre sínus s kmitočtom 4kHz - guľové pole



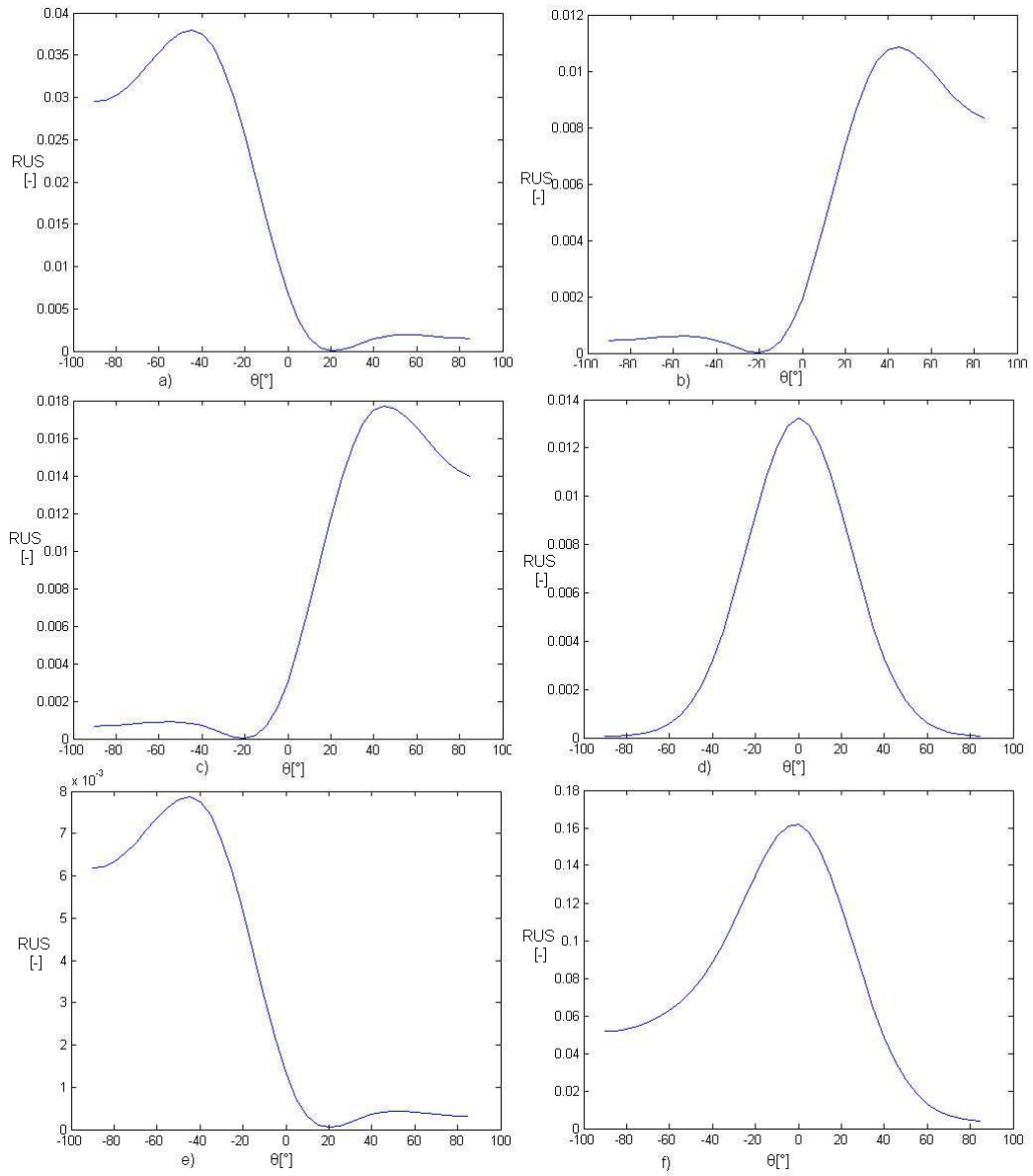
Obr. A.8: Metóda Capon pre sínus s kmitočtom 4kHz - guľové pole



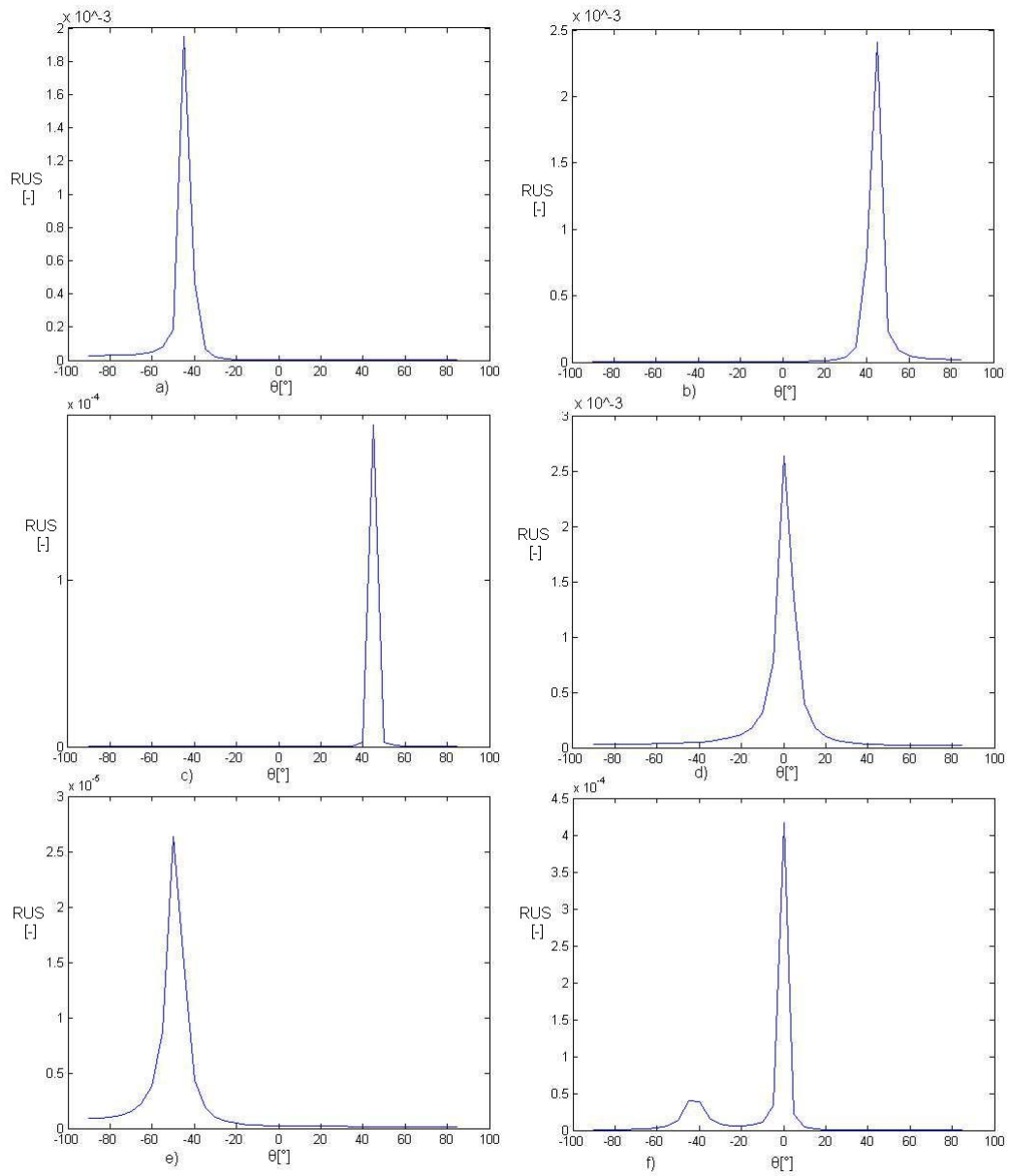
Obr. A.9: Beamforming pre hudobný signál - guľové pole



Obr. A.10: Metóda Capon pre hudobný signál - guľové pole



Obr. A.11: Beamforming pre rečový signál - guľové pole



Obr. A.12: Metóda Capon pre rečový signál - guľové pole

## B FOTOGRAFIE USPORIADANIA MIKROFÓNOVÝCH POLÍ



Obr. B.1: Pole zložené z 5 mikrofónov



Obr. B.2: Lineárne pole zložené z 8 mikrofónov



Obr. B.3: Guľové pole zložené z 8 mikrofónov