



M Ű E G Y E T E M 1 7 8 2

**Budapesti Műszaki és Gazdaságtudományi Egyetem**  
Villamosmérnöki és Informatikai Kar  
Hálózati Rendszerek és Szolgáltatások Tanszék

Csóka Bence

**HANGFORRÁSOK  
LOKALIZÁCIÓJA  
MIKROFONRENDSZEREKSEL**

KONZULENS

**Dr. Fiala Péter**

BUDAPEST, 2019

# Tartalomjegyzék

<b>Kivonat</b> .....	<b>4</b>
<b>Abstract</b> .....	<b>5</b>
<b>1 Bevezetés</b> .....	<b>6</b>
1.1 Motiváció, cél .....	6
1.2 Korábbi kutatások a témában.....	6
<b>2 Nyalábformáló algoritmusok</b> .....	<b>8</b>
2.1 A nyalábformálás alapjai .....	8
2.1.1 Fókuszálás és forráslokalizáció .....	8
2.1.2 A mikrofonok és a vászon elhelyezkedése .....	9
2.1.3 A vizsgálati frekvencia .....	11
2.2 Conventional Beamforming.....	12
2.3 Multiple Signal Classification .....	13
2.4 Compressive Beamforming .....	14
2.4.1 Compressive Beamforming zajmentes esetben .....	14
2.4.2 Compressive Beamforming zajos esetben .....	15
2.4.3 Compressive Beamforming pontszórásfüggvényekkel .....	16
<b>3 Szimulációk</b> .....	<b>17</b>
3.1 A szimulációs keretrendszer .....	17
3.2 Szimulációk két dimenzióban.....	18
3.2.1 Conventional Beamforming.....	19
3.2.2 Multiple Signal Classification .....	20
3.2.3 Compressive Beamforming .....	21
3.3 Szimulációk három dimenzióban.....	24
3.3.1 Conventional Beamforming.....	25
3.3.2 Multiple Signal Classification .....	25
3.3.3 Compressive Beamforming .....	26
<b>4 Mérések, eredmények</b> .....	<b>29</b>
4.1 Süketszobás mérések .....	29
4.1.1 Fehérzajt sugárzó hangforrások .....	29
4.1.2 Két különböző zeneszámot sugárzó hangforrások.....	31
4.2 Szabadtéri mérések .....	33

4.2.1 Flying Ranger .....	34
4.3 Továbbfejlesztési lehetőségek .....	37
<b>5 Összefoglalás.....</b>	<b>38</b>
<b>Köszönetnyilvánítás .....</b>	<b>39</b>
<b>Irodalomjegyzék.....</b>	<b>40</b>

# Kivonat

Jelen dolgozat témája hangforrások pozíciójának vagy pályájának automatizált meghatározása mikrofonrendszerekkel történő akusztikai mérésekkel. A mikrofonrendszer által a felvett hangjelek megfelelő erősítésével és késleltetésével egy adott irányba fókuszálhatunk különböző nyalábformáló algoritmusok segítségével.

A dolgozat három nyalábformáló algoritmus részletezésével foglalkozik, ezek közül kettő a szakirodalom által részletesen tárgyalt. A CBF (Conventional Beamforming) módszer a három közül a legegyszerűbb, ugyanakkor a legkevésbé pontos is. A MUSIC (Multiple Signal Classification) algoritmus a vett jelek keresztspektrum-mátrixának a jel- és zajterekre való felbontásán alapul.

Újdonságként tárgyalásra kerül a CS (Compressive Sensing) elven alapuló Compressive Beamforming algoritmus. A Compressive Beamforming a forráslokalizálást konvex optimalizálási problémára vezeti vissza, és egy ritkasági kényszerfeltételnek köszönhetően pontosabb eredményt ad, mint a lineáris algebrai elven működő algoritmusok.

A különböző algoritmusokat MATLAB környezetben elvégzett szimulációk segítségével zajmentes és zajos esetekben is összehasonlítjuk. Méréseket mutatunk be, melyekkel a gyakorlatban is értékelhetjük a nyalábformáló algoritmusok működését, ezzel is igazolva használhatóságukat különböző hangforrások (például drónok) lokalizációjához.

## **Abstract**

The topic of this thesis is the automatized determination of the location and course of sound sources by acoustic measurements with microphone systems. It is possible to focus in a specific direction with the microphone system by amplifying and delaying the received signals in a suitable manner, with the help of beamforming algorithms.

The thesis details three types of beamforming algorithms, two of which have already been discussed extensively in the scientific literature. CBF (Conventional Beamforming) is the simplest of the three, but it is also the least accurate. MUSIC (Multiple Signal Classification) is based on the separation of the received signal's cross-spectral matrix to signal and noise subspaces.

As a novelty, the CS (compressive sensing) based Compressive Beamforming is also studied. This method localizes sources by solving a convex optimization problem, and it offers a more accurate solution compared to linear algebraic algorithms thanks to a sparsity condition.

These algorithms are comparable through simulations in the MATLAB environment, both in noiseless or noisy conditions. We can also evaluate the performance of these methods in practice by performing measurements, and prove their usability in localizing sound sources.

# 1 Bevezetés

## 1.1 Motiváció, cél

Napjainkban sokszor felmerül az igény arra, hogy akusztikai kameraként használt mikrofonrendszerek segítségével, akusztikai módszerekkel tudjuk meghatározni egy hangforrás pozícióját. Ez a módszer jól használható alternatívát jelent olyan esetekben, amikor különböző okokból nem lehet vagy nem érdemes szenzorokat elhelyezni a hangot lesugárzó objektumon. A mikrofonok által vett jelek nyalábformáló algoritmusokkal feldolgozva felhasználhatók arra, hogy megbecsüljük a forráseloszlást, és ez alapján amplitúdótérképeket készítsünk. Mindezt álló és mozgó hangforrás esetén is megtehetjük, utóbbinál az idő függvényében is vizsgálhatjuk a forráseloszlás és a rá végzett becslések változását.

Munkám során a céloom két nyalábformáló algoritmus megismerése és tesztelése, ezek a Conventional Beamforming (CBF) és a Multiple Signal Classification (MUSIC). Ezeket összehasonlítom egy újszerű módszerrel, a Compressive Sensing (CS) alapú Compressive Beamforming-gal, ami jelenleg is nemzetközileg kutatás alatt áll. Az algoritmusokat MATLAB környezetben implementálva felhasználom szimulációkhoz és mérési eredmények feldolgozásához, ezáltal kipróbálom a pontosságukat és a megbízhatóságukat különböző hangforrások és körülmények esetén.

## 1.2 Korábbi kutatások a témában

Ebben a szakaszban röviden megemlítek néhány újabb kutatási eredményt a Compressive Beamforming területén a teljesség igénye nélkül, és ezeket összevetem a saját munkámmal, a legfőbb hasonlóságokat és különbségeket keresve.

Xenaki, Gerstoft és Mosegaard 2014-es cikke többek között a CBF-et, a MUSIC-ot és a CS-t tárgyalja, különös figyelmet fordítva ez utóbbira [1]. Az algoritmusok bemutatása mellett kétdimenziós szimulációkkal hasonlítják össze őket. Konklúzióként levonják, hogy a CS nagyfelbontású iránybecslést tud adni más módszerekhez képest.

Xenaki és Gerstoft egy másik, 2015-ös cikke [2], illetve Park, Choo és Seong egy 2018-ban megjelent publikációja [3] olyan módszert tárgyal, amely a

hagyományosan alkalmazott diszkrét szögtartomány helyett folytonos szögtartományt használ az iránybecslésben, ezáltal kiküszöbölik a módszer bizonyos hátrányait. Mindkét cikkben kétdimenziós szimulációkkal igazolják az eljárásuk működését, az utóbbiban ehhez több pillanatképet együttesen is felhasználnak.

A CS számos területen alkalmazható a gyakorlatban, például az orvostudományban [4], ultrahangos képalkotásban [5], csatornakódolás során hibajavításban [6], szeizmikus képalkotásban [7] és képrekonstrukcióban [8].

Jelen dolgozatban összehasonlítom a Compressive Beamforming-ot a hagyományos nyalábformálással és a MUSIC algoritmussal, és egy új változatát mutatom be. A becslésekhez diszkrét szögtartományt használok. A módszerekkel két- és háromdimenziós szimulációkat is végzek, és mikrofonrendszerekkel végzett mérések eredményét is felhasználok a tesztelésükre.

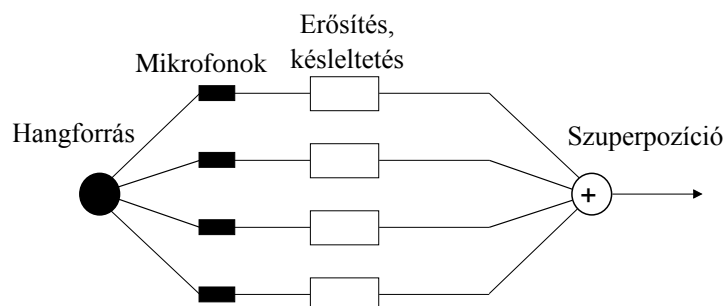
## 2 Nyalábformáló algoritmusok

Ebben a fejezetben a nyalábformálás elvét és a három vizsgált algoritmust részletezem. A 2.1 szakaszban bemutatom a használt Delay and Sum módszert, ennek véges és végtelen fókusz távolságú eseteit, a mikrofonok és a mérési pontok elhelyezkedésének hatását, és a választott mérési frekvencia alsó és felső határait. A 2.2, 2.3 és 2.4 szakaszok tárgyalják a Conventional Beamforming, a Multiple Signal Classification és a Compressive Beamforming elvét, illetve ez utóbbinak néhány kiegészítését.

### 2.1 A nyalábformálás alapjai

#### 2.1.1 Fókuszálás és forráslokalizáció

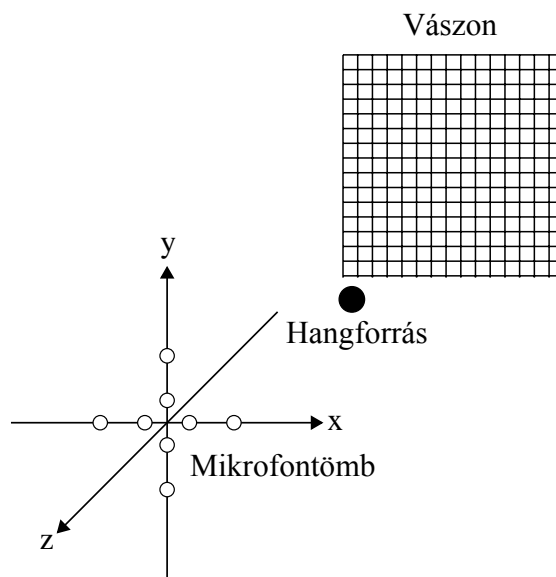
A nyalábformáló algoritmusok használatával és a mikrofonrendszerrel két fő feladatot kell elvégeznünk, ezek az akusztikai fókuszálás és a forráslokalizáció. Az akusztikai fókuszálás a Delay and Sum nevű módszeren alapszik, aminek lényege, hogy gömbkarakterisztikájú mikrofonokat többféle különböző elrendezésben elhelyezhetünk, majd ezek vett jelét külön-külön erősítjük és késleltetjük, majd ezeket a módosított jeleket szuperponáljuk. A jelek egymáshoz képesti fázisviszonyai megfelelő beavatkozás esetén azt eredményezik, hogy a tér egy adott irányából érkező hangot kiemeljük velük, más irányokból érkezőket pedig elnyomjuk. A mikrofonok térbeli elrendezésével és a vett jelek erősítésével és késleltetésével lehet kialakítani egy számunkra megfelelő iránykarakterisztikát, a késleltetések további befolyásolásával pedig hasonló hatást érhetünk el, mintha elforgatnánk a mikrofontömböt, ezáltal többféle irányra lehet fókuszálni.



1. ábra: A Delay and Sum módszer. A mikrofonok vett jeleit külön-külön erősítjük és késleltetjük, majd szuperponáljuk, ezzel egy adott irányra fókuszálunk.



A másik fő feladat, a forráslokalizáció (vagy nyálábformálás) során a térben lévő hangforrások irányát és teljesítményét becsüljük meg a nyálábformáló algoritmusok segítségével. A térben kijelölünk egy ponthalmazt, amik közösen alkotják az úgynevezett akusztikai vásznat, és az ezeken a pontokon elhelyezett virtuális források közül azokat választjuk ki, amelyek által létrehozott hangtér a leginkább hasonlít a mikrofontömb által vett valós hangjelekhez. A két feladat (fókuszálás és nyálábformálás) egymástól függetlenül is végrehajtható, de jelen munka során mindkettőt együttesen használva hajthatjuk végre az iránybecslést és a forrásjel kiemelését.



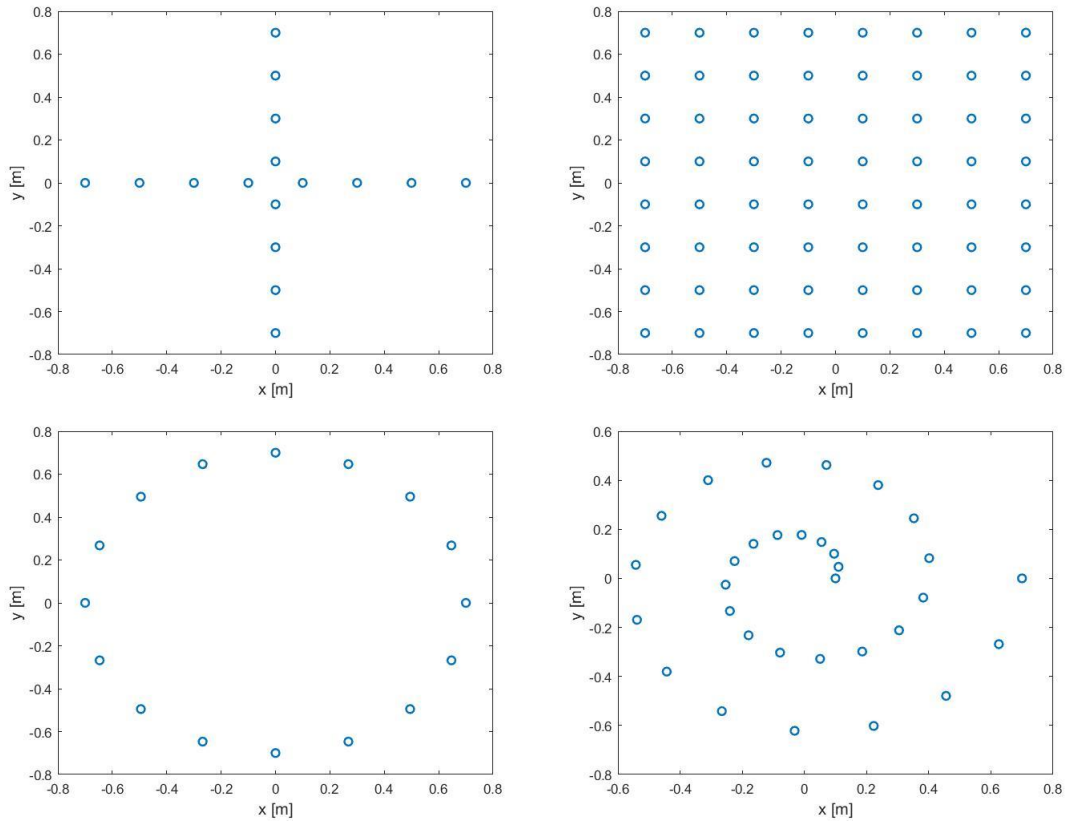
2. ábra: Általános mérési elrendezés.

### 2.1.2 A mikrofonok és a vászon elhelyezkedése

A mikrofonokat többféle elrendezésben helyezhetjük el, például vonal, kereszt, téglalap, négyzetrács, kör vagy spirál formában. Az elrendezés formája és az egyes mikrofonok pozíciói is meghatározzák a vett jeleket módosító beavatkozó függvényt, illetve a mikrofontömb iránykarakterisztikáját is. Két dimenzióban, ahol csak egy síkon belül, egy szögparaméterrel leírható irányokra fókuszálunk, a vonal elrendezést használjuk. Három dimenzióban, ahol az egész térben, két szögparaméterrel leírható irányokra fókuszálhatunk, a drágán megvalósítható négyzetrács elrendezés helyett érdemes téglalapot vagy keresztet (esetleg kört vagy spirált) alkalmazni.

Az akusztikai vászon pontjai/irányai általában egy egyenesen/körön, vagy egy síkban/gömbfelületen helyezkednek el egyenletesen. A hangnak a vászon és a

mikrofonrendszer közötti terjedésében figyelembe kell venni a levegő mint közvetítő közeg bizonyos tulajdonságait, mint például a légnyomást és a hőmérsékletet, amik alapján kiszámolható a hangsebesség.

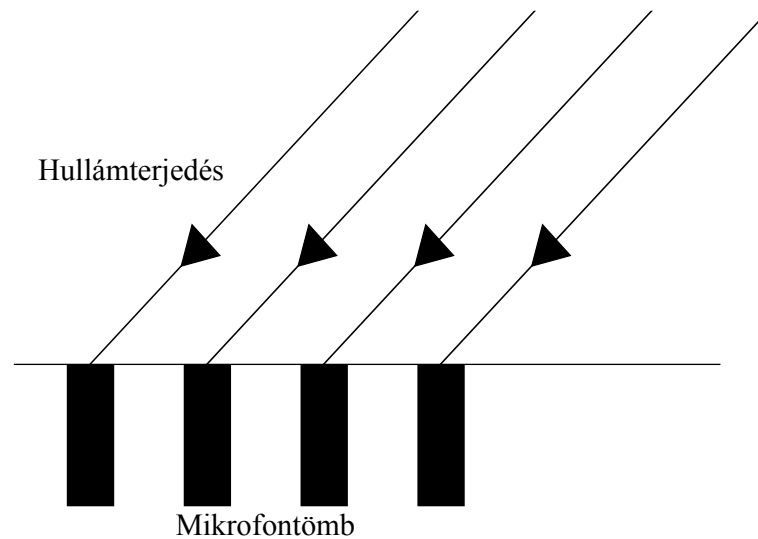


**3. ábra: Kereszt, négyzetrács, kör és spirál elrendezés.**

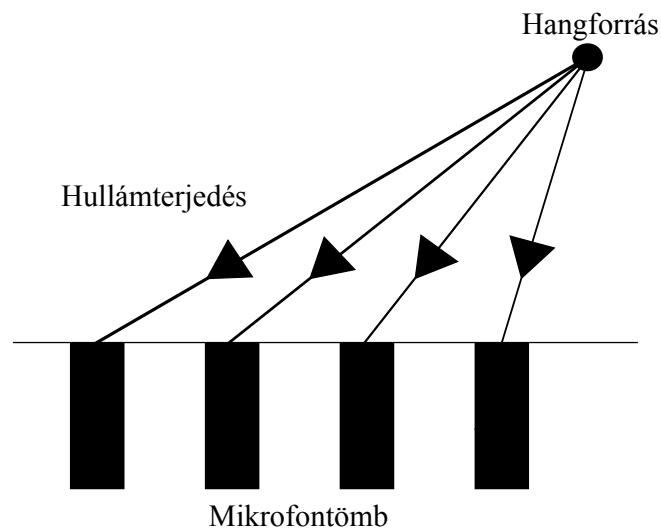
A nyalábformálást kétféleképpen is elvégezhetjük attól függően, hogy a forrás milyen messze helyezkedik el a mikrofontömbtől. Ha ehhez a távolsághoz képest elhanyagolható a tömb mérete, akkor alkalmazhatjuk a végtelen fókusz távolságú nyalábformálást. Ebben az esetben a forrás által kibocsátott hang hullámfrontja síkfelületnek tekinthető, és a mikrofonokhoz közel azonos amplitúdójú és beesési szögű hullámok érkeznek, de különböző fázisban. Itt a nyalábformálás során csak a beesési szöget kell figyelembe venni, a forrás és a mikrofon távolságát nem, és a vászon definiálásakor is elég csak irányokat (szögeket) megadni.

Ha a mikrofonok és a forrás távolsága összemérhető a mikrofonrendszer méretével, akkor véges fókusz távolságú nyalábformálást kell használnunk. Ilyenkor a hullámfront egy gömbfelület, azaz a mikrofonok által vett jelek különböző amplitúdójúak és fázisúak, és a beesési szög is különböző minden mikrofonnál. A nyalábformálás során a beesési szöget és a távolságot is figyelembe kell venni, a vásznat

pontok koordinátáinak összességéként kell definiálni. Érdeemes megvizsgálni azt is, milyen mérési és szimulációs eredmények adódnak akkor, ha a forrás távolsága és a vászon távolsága a mikrofontömbtől különbözik.



4. ábra: Végtelen fókusz távolságú nyalábformálás.



5. ábra: Véges fókusz távolságú nyalábformálás.

### 2.1.3 A vizsgálati frekvencia

A nyalábformáló algoritmusokat frekvenciatartományban alkalmazzuk, ami ez esetben azt jelenti, hogy egy adott pillanatban (rövid időintervallumban) minden mikrofon vett jelének egy adott frekvenciájú komponensét használjuk fel a becsléshez. A megfelelő frekvencia választása kulcsfontosságú, ugyanis bizonyos határfrekvenciák által behatárolt tartományon kívül a mérési elvbe hiba csúszik.

Amennyiben a frekvencia megfelelően lett kiválasztva, a forrás irányára fókuszálva a mikrofonok vett jelét késleltetés után összeadva azok erősítik egymást, mert fázisban vannak, más irányokra pedig gyengítik, vagy akár ki is oltják egymást a fáziskülönbségek miatt. Ha a frekvencia túl nagy (a hullámhossz túl kicsi), akkor térbeli átlapolódás történik, a mikrofonok jele olyan irányra fókuszálva is fázisban lehet, ahol valójában nincs forrásunk. Ha a frekvencia túl kicsi (a hullámhossz túl nagy), akkor ugyan nem látunk szellemforrásokat, de rossz irányra fókuszálva sem lesz nagy a fáziskülönbség a jelek között, és csak nagyon durva becslést tudunk adni. A felső határfrekvencia az (1) összefüggéssel számítható ki:

$$f < \frac{c}{2d}, \quad (1)$$

ahol  $c$  a hangsebesség,  $d$  pedig a mikrofonok egymástól való távolsága (egyenletes elrendezést feltételezünk). Minél közelebb vannak egymáshoz a mikrofonok, annál nagyobb a felső határfrekvencia. Az alsó határfrekvenciára nincs konkrét összefüggés, minél kisebb a frekvencia, úgy fokozatosan annál elmosódottabb lesz az amplitúdótérkép.

## 2.2 Conventional Beamforming

A Delay and Sum módszerben a mikrofonok vett jelének az erősítését és késleltetését minden mikrofonhoz egy külön beavatkozó függvénnyel tehetjük meg. A beavatkozó függvényt úgy is fel lehet fogni, hogy valódi forráseloszlás-vektor ( $\mathbf{x}$ ) és a mikrofonok által vett jelek vektora ( $\mathbf{y}$ ) között a kapcsolatot egy  $\mathbf{A}$  érzékelési mátrix valósítja meg. Ennek a mátrixnak a segítségével tudunk becslést adni a vett jelekből a forráseloszlásra a nyálábformáló algoritmusokkal. Az érzékelési mátrix a mikrofonok és a vászon elhelyezkedése, illetve a választott mérési frekvencia alapján határozható meg. Az  $\mathbf{x}$  és  $\mathbf{y}$  vektorok frekvenciatartománybeli információt hordoznak, azaz azt mondják meg, hogy a kisugárzott/vett jeleknek a kiválasztott vizsgálati frekvencián mekkora amplitúdójú komponensük van egy rövid időszakaszon.

Az érzékelési mátrix elemei a következő formulák alapján határozhatók meg, a (2)-es egyenlet kétdimenziós, a (3)-as pedig háromdimenziós esetre vonatkozik [1]:

$$\mathbf{A}(i, j) = \frac{1}{\sqrt{M}} e^{jk\alpha_i \sin \eta_j}, \quad (2)$$

$$\mathbf{A}(i, j) = \frac{1}{\sqrt{M}} e^{jk(\alpha_i \sin \theta_j \sin \varphi_j + \beta_i \cos \theta_j \sin \varphi_j)}, \quad (3)$$

ahol  $M$  a mikrofonok száma ( $1/\sqrt{M}$ -mel normalizálunk),  $k$  pedig a hullámszám (a vizsgálati körfrekvencia és a hangsebesség hányadosa).  $\alpha_i$  és  $\beta_i$  az  $i$ -dik mikrofon koordinátái derékszögű koordináta-rendszerben az  $\alpha$  és a  $\beta$  tengelyek mentén (a mikrofonok az  $\alpha$ - $\beta$  síkban helyezkednek el),  $\eta_j$ ,  $\theta_j$  és  $\varphi_j$  pedig a vászon  $j$ -dik pontjának az irányát azonosítják, végtelen fókusz-távolság esetén ( $\eta$  az  $\alpha$  tengellyel bezárt szög mínusz  $90^\circ$ ,  $\varphi$  az irány  $\alpha$ - $\beta$  síkra levetítésének az  $\alpha$  tengellyel bezárt szöge,  $\theta$  pedig a  $\gamma$  tengellyel bezárt szög, a koordináta-rendszer középpontja egyben a mikrofonrendszer geometriai középpontja is). Az  $A$  mátrixnak összesen annyi oszlopa van, ahány pontból áll a vászon, és annyi sora, ahány mikrofonból áll a mikrofontömb. Az érzékelési mátrix az alábbi módon hozza létre a kapcsolatot a forrás ( $\mathbf{x}$ ) és a vett jelek ( $\mathbf{y}$ ) között:

$$\mathbf{y} = \mathbf{A}\mathbf{x}. \quad (4)$$

Az egyik legegyszerűbb nyálábformálási módszer a CBF, ami során az  $\mathbf{y}$ -t balról szorozzuk az  $\mathbf{A}$  mátrix konjugált transzponáltjával:

$$\hat{\mathbf{x}} = \mathbf{A}^H \mathbf{y}. \quad (5)$$

Ennek a módszernek az előnye az egyszerűsége és kis számításigénye, hátránya azonban a pontatlansága, ugyanis az antenna iránykarakterisztikájának a főnyalábja véges, de nullánál nagyobb szélességű, és a valódi forráspozíció közvetlen közelében is zérustól jelentősen eltérő értékeket becslünk, az amplitúdótérkép elmosódott lesz.

## 2.3 Multiple Signal Classification

A MUSIC algoritmus egy lineáris algebrai módszer [9], amelynek során a vett jelekből képzett spektrális keresztkorrelációs mátrixot szétválasztjuk jel és zaj alterekre, majd a zaj sajátvektorainak és az érzékelési mátrixnak a segítségével becslünk.

A keresztkorrelációs mátrixot ( $\mathbf{R}$ ) az alábbiak szerint határozhatjuk meg:

$$\mathbf{R} = \mathbf{y}\mathbf{y}^H. \quad (6)$$

A mátrix sajátvektorait sorba rendezzük a hozzájuk tartozó sajátértékek szerint növekvő sorrendbe. Ezek közül a  $K$  legnagyobb sajátértékhez tartozó sajátvektor alkotja a jel alterét ( $\mathbf{U}_s$ ), a többi pedig a zajét ( $\mathbf{U}_n$ ), ahol  $K$  a források becsült száma (ezt például úgy is megbecsülhetjük, hogy megnézzük, hogy a sorba rendezett sajátértékek közül melyik két szomszédos hányadosa a legnagyobb, és ott lesz a jel és a zaj alterek határa). Az érzékelési mátrixszal és a zaj alterrel a következő formula szerint végezzük a becslést:

$$\mathbf{P}_{MUSIC} = \frac{1}{\mathbf{A}^H \mathbf{U}_n \mathbf{U}_n^H \mathbf{A}}, \quad (7)$$

vagyis az  $\mathbf{A}^H \mathbf{U}_n \mathbf{U}_n^H \mathbf{A}$  mátrix minden elemének vesszük a reciprokát.

A módszer hátránya, hogy előre meg kell becsülni a források számát, illetve az, hogy csak inkoherens forrásokra működik. Előnye azonban a kis számításigény mellett a hagyományos nyálábformáláshoz képest nagyobb pontosság.

## 2.4 Compressive Beamforming

A Compressive Beamforming a Compressive Sensing (CS) eljárásán alapul. Ez a módszer a forráslokalizációt egy konvex optimalizálási problémára visszavezetve oldja meg, azonban a forrásnak meg kell felelnie egy ritkasági kényszerfeltételnek annak érdekében, hogy a módszer pontos becslést adhasson [1], [10], [11]. A Compressive Beamforming előnye a pontossága a kényszerfeltétel teljesülése esetén, hátránya viszont a nagy számításigény, ezért valós időben ritkán használható.

### 2.4.1 Compressive Beamforming zajmentes esetben

A ritkasági kényszerfeltétel azt jelenti, hogy kevés forrásunk van a vászon pontjainak számához képest (amik elég sűrűn helyezkednek el ahhoz, hogy ne adódjon probléma a források rossz illeszthetősége miatt). Ezt matematikailag úgy is megfogalmazhatjuk, hogy a forrást leíró vektornak (vagy mátrixnak) kevés nem-nulla eleme van. Ha ez a feltétel teljesül, akkor a célunk az, hogy a forrásra végzett becslésünknek minél kevesebb zérustól különböző értéke legyen, miközben  $\mathbf{Ax} = \mathbf{y}$  teljesül:

$$\min_{\mathbf{x} \in \mathbb{C}^N} \|\mathbf{x}\|_0 \text{ úgy, hogy } \mathbf{Ax} - \mathbf{y} = 0, \quad (8)$$

ahol  $\mathbf{x}$  a becslésünk a forrásra,  $\mathbf{y}$  a vett jelek mátrixa,  $\mathbf{A}$  az előre meghatározott és fix ( $\mathbf{x}$ -től és  $\mathbf{y}$ -től nem függő) érzékelési mátrix,  $\|\mathbf{x}\|_0$  pedig az  $\mathbf{x}$  vektor  $l_0$ -normája. Egy adott vektor  $l_p$ -normája általánosan:

$$\|\mathbf{x}\|_p = \left( \sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}}, \quad (9)$$

ami  $p=0$ -ban éppen a nem-zérus elemek számát adja meg:

$$\|\mathbf{x}\|_0 = \sum_{i=1}^n 1_{x_i \neq 0}. \quad (10)$$

A (8) probléma megoldására hatékony algoritmus nem létezik, azonban ha az  $l_0$ -norma szerinti feltételt relaxáljuk  $l_1$ -normára, akkor egy jóval könnyebben megoldható problémához jutunk:

$$\min_{\mathbf{x} \in \mathcal{C}^N} \|\mathbf{x}\|_1 \text{ úgy, hogy } \mathbf{Ax} - \mathbf{y} = 0. \quad (11)$$

A Compressive Sensing módszerek területén nagy áttörés volt annak a bebizonyítása, hogy megfelelően ritka forrás esetén (jóval kevesebb forrás van, mint vizsgálati pont a vásznon) a (11) ekvivalens a (8) problémával, és felhasználható nyálábformáló algoritmusokban [12], [13], [14].

### 2.4.2 Compressive Beamforming zajos esetben

A gyakorlatban a mikrofonok által vett jelre zaj is rakódik, amit a nyálábformáló algoritmusban figyelembe kell venni. Additív zajt feltételezve:

$$(\mathbf{y} = \mathbf{Ax} + \mathbf{n}) \quad (12)$$

a konvex optimalizálási problémánk a következőképpen módosul [15]:

$$\min_{\mathbf{x} \in \mathcal{C}^N} \|\mathbf{x}\|_1 \text{ úgy, hogy } \|\mathbf{Ax} - \mathbf{y}\|_2 \leq \varepsilon. \quad (13)$$

Az  $\varepsilon$  a zaj miatt bevezetett hibahatár a becslésre, amire teljesülnie kell az  $\|\mathbf{n}\|_2 \leq \varepsilon$  feltételnek.

Ha a zajszint ismeretlen, nekünk kell becslést adnunk rá, azonban felmerül a kérdés, hogy az alul- vagy a felülbecslés közül melyik a kedvezőbb. A zajszint alulbecslése esetén előfordulhat, hogy a megoldás kevésbé lesz ritka a valóságnál, azaz olyan helyeken is forrást látunk, ahol valójában nincsenek. Felülbecsléskor ennek az ellenkezője történik, azaz lehet, hogy a megoldás túl ritka lesz, és a kisebb erősségű forrásokat nem találjuk meg. A kettő közül az alulbecslés a kedvezőbb, ugyanis a ritkaság csökkenése kiküszöbölhető újrasúlyozott  $l_1$  minimalizálással [16], és szintkorrekcióval.

Az újrasúlyozott  $l_1$  minimalizálás során először egy kezdeti becslést kell végrehajtani a (13) formula segítségével. Ezután iteratíván megoldjuk az alábbi súlyozott problémát:

$$\min \|\mathbf{W}\mathbf{x}\|_1 \text{ úgy, hogy } \|\mathbf{Ax} - \mathbf{y}\|_2 \leq \varepsilon, \quad (14)$$

ahol  $\mathbf{W}$  egy diagonális mátrix, az elemeinek értéke a kezdeti becslés előtt 1, utána pedig a legutóbbi becslés alapján adhatóak meg:

$$w_i = \frac{1}{\hat{x}_i + \xi}. \quad (15)$$

Itt  $\xi$  egy nullánál nagyobb szám a legkisebb várható amplitúdó nagyságrendjében, ami biztosítja, hogy a súlyozás egyik eleme se legyen végtelen,  $\hat{x}_i$  a forrásra végzett

legutóbbi becslés  $i$ -dik eleme. Az iterációt akkor hagyjuk abba, amikor az utolsó két becslés közel megegyezik, ez általában két iterációval megvalósul:

$$|w_i x_i|^{k+1} = \begin{cases} \frac{\hat{x}_i}{\hat{x}_i + \xi} \approx 1, \text{ ha } |x_i| \geq 0, \\ 0, \text{ egyébként.} \end{cases} \quad (16)$$

Az újrásúlyozott megoldás tovább javítható szintkorrekció alkalmazásával. Ha az érzékelési mátrixnak csak azokat az oszlopait hagyjuk meg, amelyek a megoldásban csúcértéknek felelnek meg, a többinek az elemeit pedig nullával helyettesítjük, akkor a kapott  $\mathbf{A}_a$  mátrixot felhasználva oldjuk meg a következő egyenletrendszert [17]:

$$\hat{\mathbf{x}}_a = \mathbf{A}_a^+ \mathbf{y}, \quad (17)$$

ahol  $\mathbf{A}_a^+$  a mátrix pszeudoinverze.

### 2.4.3 Compressive Beamforming pontszórásfüggvényekkel

A Compressive Beamforming módszernek bevezethetjük egy olyan változatát, amelynek során egy másfajta konvex optimalizálási problémát oldunk meg a pontszórás-függvényeket felhasználásával. A pontszórás-függvényeket minden virtuális forráspozícióra definiálhatjuk úgy, hogy feltételezünk egy egységnyi teljesítményű forrást az adott pontban, majd pedig megnézzük, hogy a vászon összes pontjára mekkora forrás-teljesítményt becsülhetünk hagyományos nyalábformálással (CBF). A pontszórás-függvényekből létrehozhatunk egy mátrixot, aminek annyi oszlopa (és sora) van, ahány virtuális forráspozíció, és a  $k$ -edik oszlopa a  $k$ -edik ponthoz tartozó függvény vektora. Ez a mátrix az érzékelési mátrixból így számítható ki:

$$\mathbf{A}_{PSF} = \mathbf{A}^H \mathbf{Y} = \mathbf{A}^H \mathbf{A} \mathbf{E} = \mathbf{A}^H \mathbf{A}. \quad (18)$$

A módosított konvex optimalizálási probléma során a hagyományos nyalábformálás eredményét próbáljuk előállítani a pontszórás-függvények lineáris kombinációjaként:

$$\min_{\mathbf{x} \in \mathbb{C}^N} \|\mathbf{x}\|_1 \text{ úgy, hogy } \|\mathbf{A}_{PSF} \mathbf{x} - \mathbf{y}\|_2 \leq \varepsilon \text{ és } \mathbf{x} \geq 0, \quad (19)$$

ahol  $\mathbf{y}$  a becslés hagyományos nyalábformálással, és az  $\mathbf{x} \geq 0$  feltétel fizikailag azt jelenti, hogy egy hangforrásnak a teljesítménye nemnegatív.



## 3 Szimulációk

A fent részletezett nyálábformálási módszerek összehasonlíthatóak szimulációk elvégzésével. MATLAB környezetben elvégezhetőek olyan számítások, amik akár kétdimenziós, akár háromdimenziós térbeli mérési elrendezéseket szimulálnak.

### 3.1 A szimulációs keretrendszer

A munkám során felhasznált MATLAB környezetben megírt szimulációs keretrendszer objektum-orientált, különböző osztályokból létrehozott elemek segítségével lehet a szükséges számításokat elvégezni. A keretrendszer néhány fontosabb, általam is felhasznált osztálya:

- Environment: az akusztikus környezetet valósítja meg (amely általában levegő), annak olyan tulajdonságait tárolja, mint például a légnyomást, a hőmérsékletet és a relatív páratartalmat, illetve ki tudja számítani a hangsebességet az adott közegben;
- Source: a hangforrás osztálya, amelyet egy trajektória és egy jel ír le;
- SoundField: egy Environment és egy Source elemből áll, képes kiszámolni egy adott pozícióban az időfüggő hangnyomást;
- MicrophoneArray: a mikroföntömböt megvalósító osztály, a mikrofonok koordinátáit tárolja;
- Canvas: a vászon pontjainak koordinátáit tároló osztály;
- Propagation: az akusztikus környezet, a vászon és a mikroföntömb alapján képes kiszámítani az érzékelési mátrixot;
- Mixdown: a vett jel alapsávba való lekeverését végzi el;
- CsmEstimator: a spektrális keresztkorrelációs mátrixot számítja ki;
- ConventionalBeamformer: a hagyományos nyálábformálást elvégző osztály, egy Propagation és egy CsmEstimator elemet kell megadni neki;

- Music: a MUSIC iránybecslést elvégző osztály, egy-egy Propagation és CsmEstimator elemeken kívül még a források előre megbecsült számát kell megadni;
- CompressedBeamformer: a Compressed Beamforming algoritmust elvégző osztály, egy Propagation és egy MixDown elemet kell megadni neki.

A nyálábformáló algoritmusokban a mikrofonok által vett jelnek egy pillanatban egy adott frekvenciájú komponensét használjuk fel (2.1.3 alfejezet). A spektrumot egy adott pillanatban úgy közelíthetjük, hogy a jelnek egy rövid szakaszán végzünk Fourier-transzformációt, ennek a módszernek azonban korlátozott a pontossága, ugyanis az időablak hossza és a kapott spektrum felbontása fordítottan arányos:

$$\Delta f = \frac{1}{T}, \quad (20)$$

itt a  $T$  az időablak hossza (ez a mostani szimulációk során 0,2 másodperc),  $\Delta f$  pedig a maximális frekvenciafelbontás. Ez azt jelenti, hogy minél rövidebb időtartamra akarunk becslést végezni, minél inkább pillanatszerűen akarjuk megkapni a spektrumot, annál alacsonyabb lesz a felbontása. Lehetőségünk van arra, hogy meghatározzuk pontosabban a jel egy adott frekvenciájú komponensét, anélkül, hogy megnöveljük az időablak méretét. Ezt a feladatot a MixDown osztály hajtja végre a következő lépéseken keresztül [18]:

1. A kiválasztott  $f_0$  frekvenciát felhasználva az eredeti jelünket egyoldali frekvenciamodulációval eltoljuk balra:

$$x_{mod}(t) = x(t)e^{-j2\pi f_0 t}. \quad (21)$$

2. A modulált jelet egy aluláteresztő szűrővel szűrjük, aminek a vágási frekvenciája jóval kisebb, mint a választott frekvencia.
3. A modulált és szűrt jelet időtartományban dolgozzuk fel. A jel amplitúdója a szűrő beállása után közel állandó, és ezt az értéket vesszük az adott frekvenciájú komponens amplitúdójának.

## 3.2 Szimulációk két dimenzióban

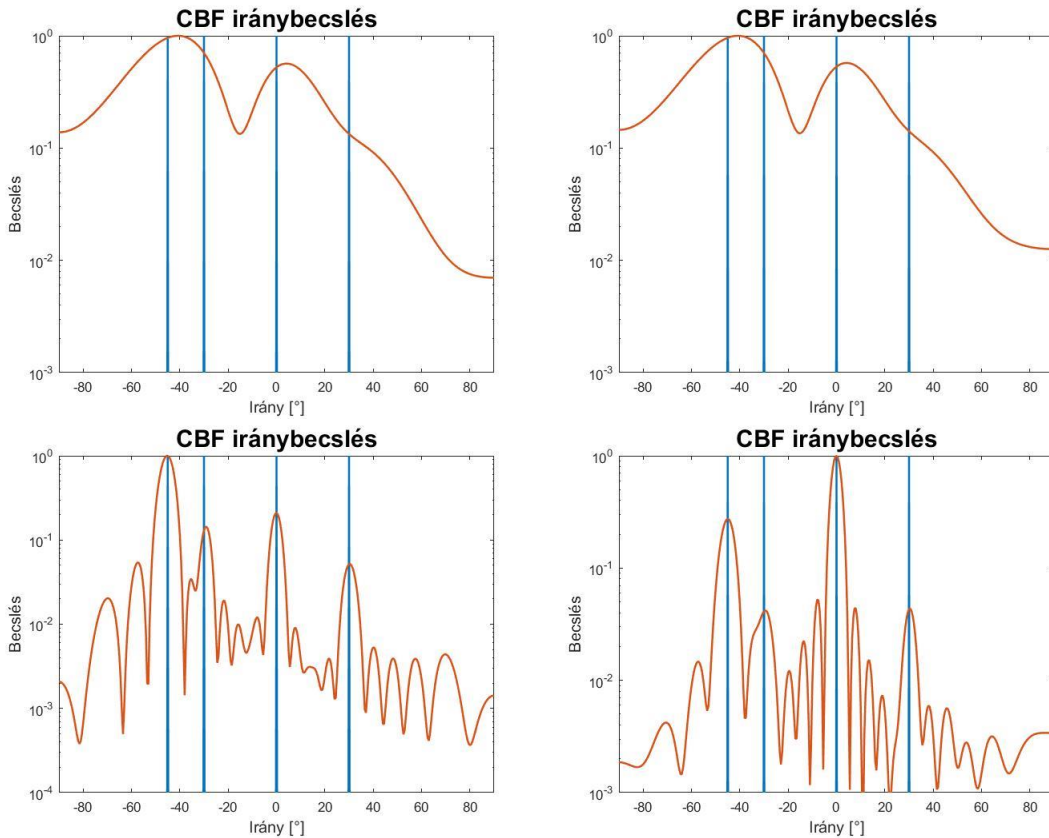
Kétdimenziós szimulációk esetén 24 darab mikrofont használunk vonal elrendezésben, két szomszédos mikrofon távolsága 6,087 centiméter. A hangsebességet

343,4486 m/s-nak vesszük. Ennek megfelelően a felső határfrekvencia az (1) formula alapján 2821,16 Hz. Érdekes megvizsgálni nyalábformálás eredményét ennél lényegesen kisebb, lényegesen nagyobb, illetve nagyjából ekkora frekvencián is megvizsgálni és ez alapján is értékelni a különböző algoritmusok használhatóságát különböző frekvenciákon. Vizsgálati frekvenciáknak 500, 2500 és 5000 Hz-et választottam. A vászon 361 pontból áll, ezek egyenletesen helyezkednek el egy félkörön, aminek a középpontjában van a mikrofontömb és a sugara 100 méter. A vászon által behatárolt szögtartomány  $[-90^\circ, +90^\circ]$ . A térben négy pontszerű forrás helyezkedik el 150 méter távolságban és  $-45^\circ$ ,  $-30^\circ$ ,  $0^\circ$  és  $+30^\circ$  irányokban (a negyedik forrás 5000 Hz-es vizsgálati frekvencia esetében szükség szerint át lett helyezve  $+60^\circ$ -ra, hogy bizonyos jelenségek jobban vizsgálhatók legyenek). A források fehérzajt bocsátanak ki (a mintavételi frekvencia 48 kHz), szórásuk rendre 3, 1, 2 és 1 egységnyi (azaz teljesítményük 9, 1, 4 és 1 egységnyi). Mindhárom nyalábformáló algoritmus működését vizsgálom zajmentes és zajos esetben is, utóbbinál 20 dB-es jel-zaj viszonytal. Az eredményüket egy véletlenszerűen kiválasztott pillanatban (0,2 másodperces időtartamra vonatkoztatva) ábrázolom. Az ábrákon a valódi forrásirányokat kézzel, az algoritmusok becslését pirossal jelölöm, ezek a maximumukhoz viszonyított relatív értékeként értelmezendők.

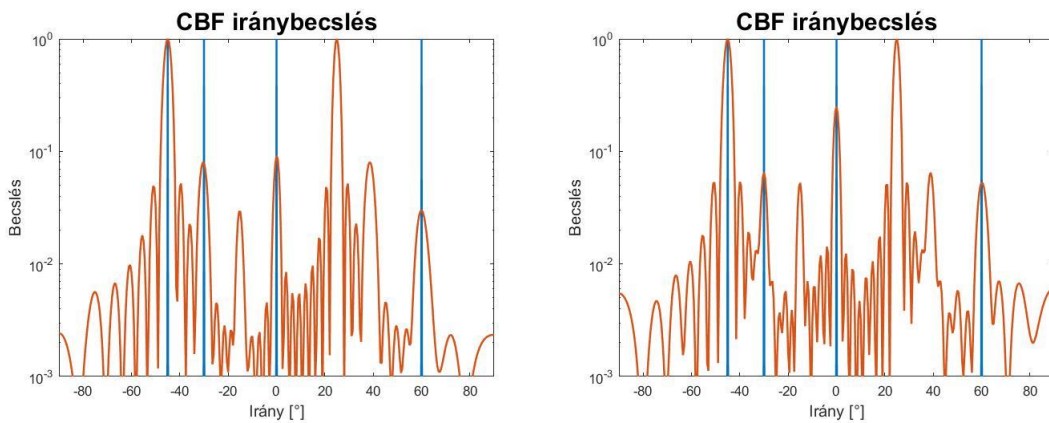
### 3.2.1 Conventional Beamforming

A 6. ábra a hagyományos nyalábformálás eredménye látható, kézzel a valóságos forrásirányok, pirossal pedig a becslések eredménye van ábrázolva. 500 Hz-es vizsgálati frekvencián a módszer nem alkalmas a források irányának becslésére. 2500 Hz-en zajtalan esetben az első három forráshoz tartozó főnyalábok még kiemelkednek a melléknyalábok közül, de a negyedik már nagyjából a legnagyobb melléknyalábok szintjén van, ráadásul a második és a negyedik forrás irányának a becslése kicsit pontatlan. Zajos esetben már csak két főnyaláb emelkedik ki a melléknyalábok közül.

5000 Hz-en már egyáltalán nem használható ez a módszer jó jel-zaj viszonyban sem, ugyanis a térbeli átlapolódás miatt megjelenik 20-30 fok környékén egy szellemforrás, ami a legnagyobb főnyalábbal van egy szinten, emiatt nem lehet helyesen meghatározni a források irányát (7. ábra). Ez a jelenség feltűnik a többi módszernél két és három dimenzióban is, ezért azoknak az eredményét nem ábrázolom külön 5000 Hz-es vizsgálati frekvencia mellett.



**6. ábra: CBF iránybecslés zajmentes (bal) és 20 dB jel-zaj viszonyú (jobb) esetben, 500 Hz (fent) és 2500 Hz (lent) vizsgálati frekvencián.**

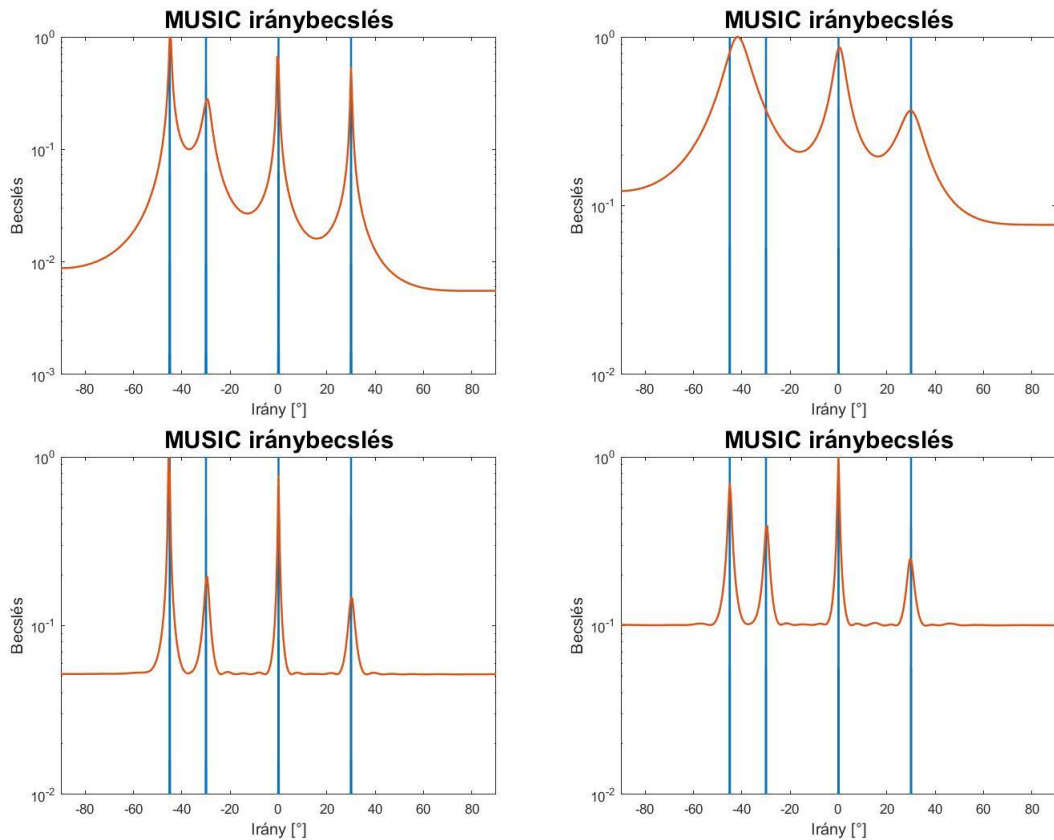


**7. ábra: CBF iránybecslés zajmentes (bal) és 20 dB jel-zaj viszonyú (jobb) esetben, 5000 Hz vizsgálati frekvencián.**

### 3.2.2 Multiple Signal Classification

A 8. ábra mutatja a MUSIC algoritmus alapú iránymeghatározás eredményét. A CBF-fel ellentétben 500 Hz-en kicsit pontatlan, de használható becslést kapunk, csak a zaj bevezetésekor nem sikerül mind a négy forrást elkülöníteni egymástól. 2500 Hz-es

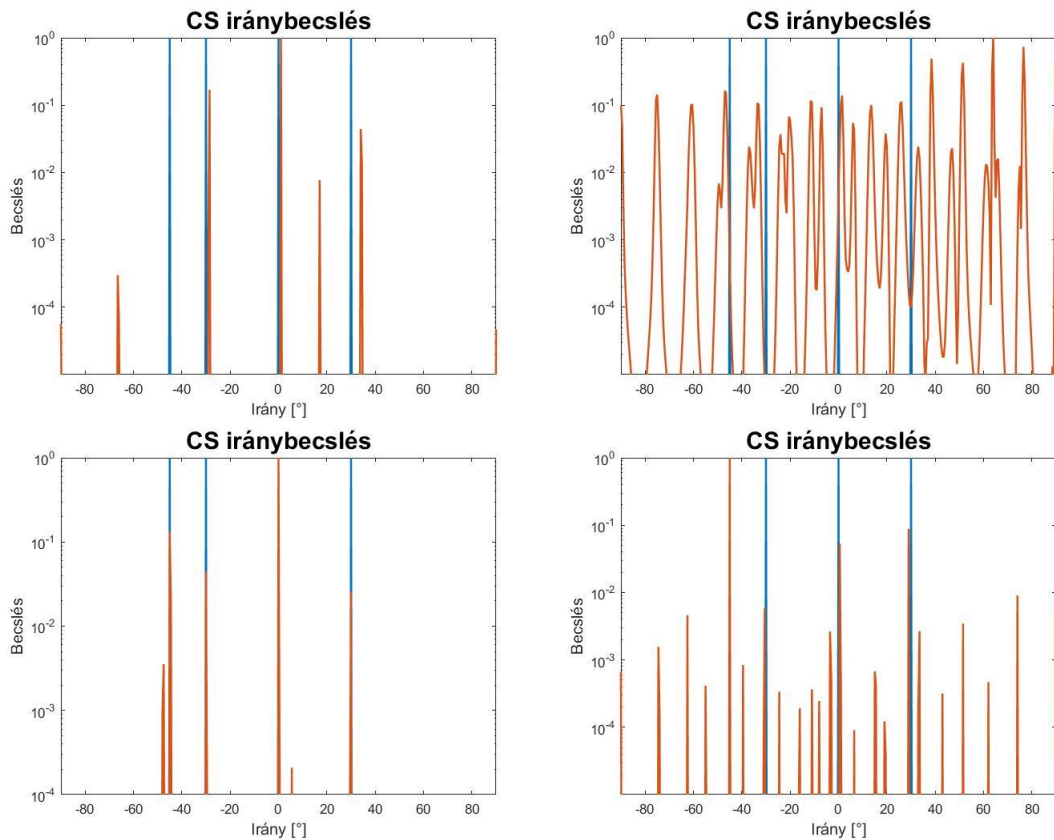
vizsgálati frekvencián még zajjal együtt is viszonylag nagy pontosságú megbecsülhetjük eredményt kapunk.



**8. ábra: MUSIC iránybecslés zajmentes (bal) és 20 dB jel-zaj viszonyú (jobb) esetben, 500 Hz (fent) és 2500 Hz (lent) vizsgálati frekvencián.**

### 3.2.3 Compressive Beamforming

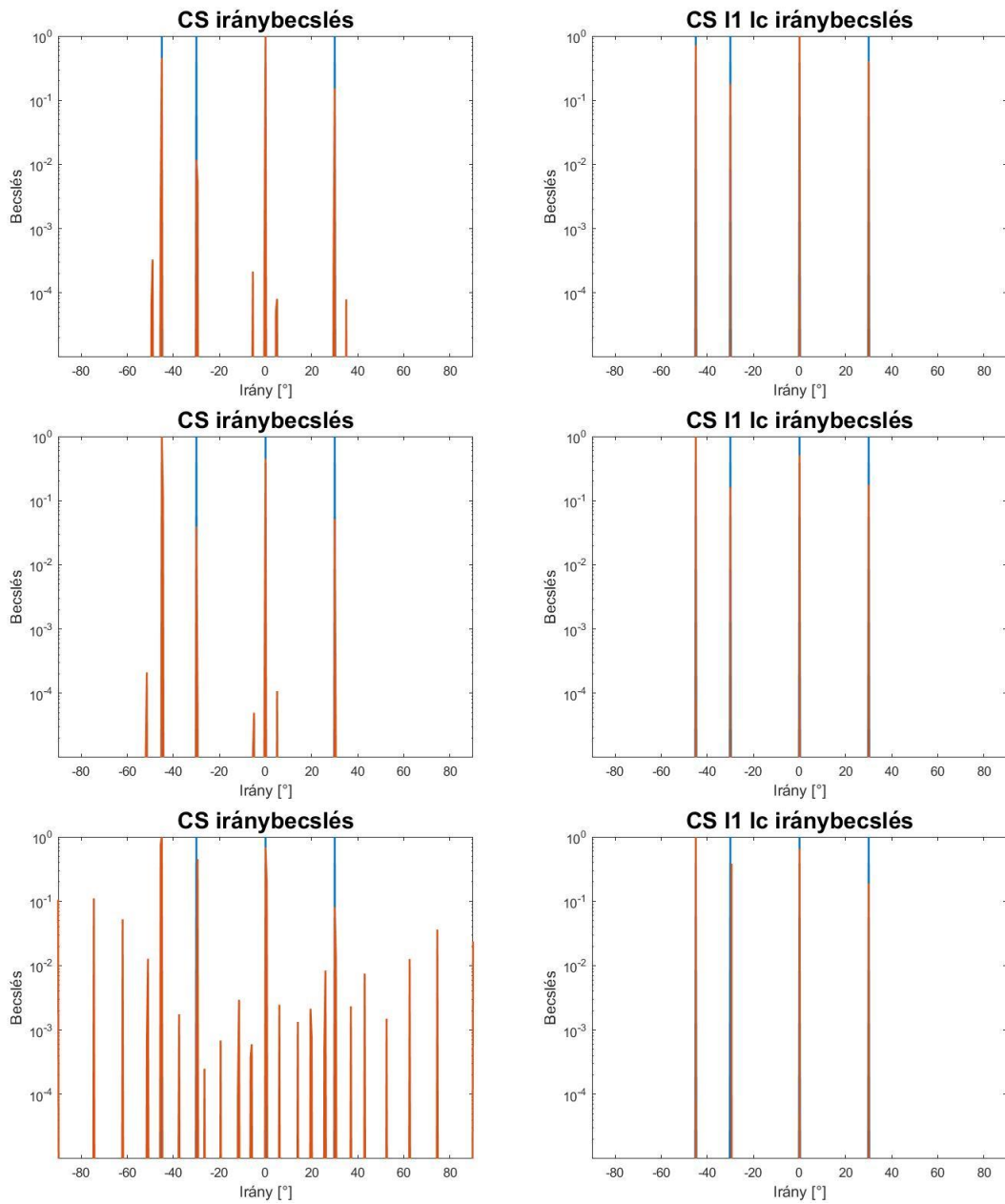
A Compressive Beamforming módszer eredménye a 9. ábra látható. 500 Hz-en az algoritmus nem tud helyes becslést adni, csak 2500 Hz-en. Megfelelő vizsgálati frekvencia esetén is a zajtűrése jóval kisebb, mint a MUSIC-nak. Ha viszont elég jó a jel-zaj viszony, a Compressive Beamforming nagy pontosságú és ritkaságú megoldást képes adni.



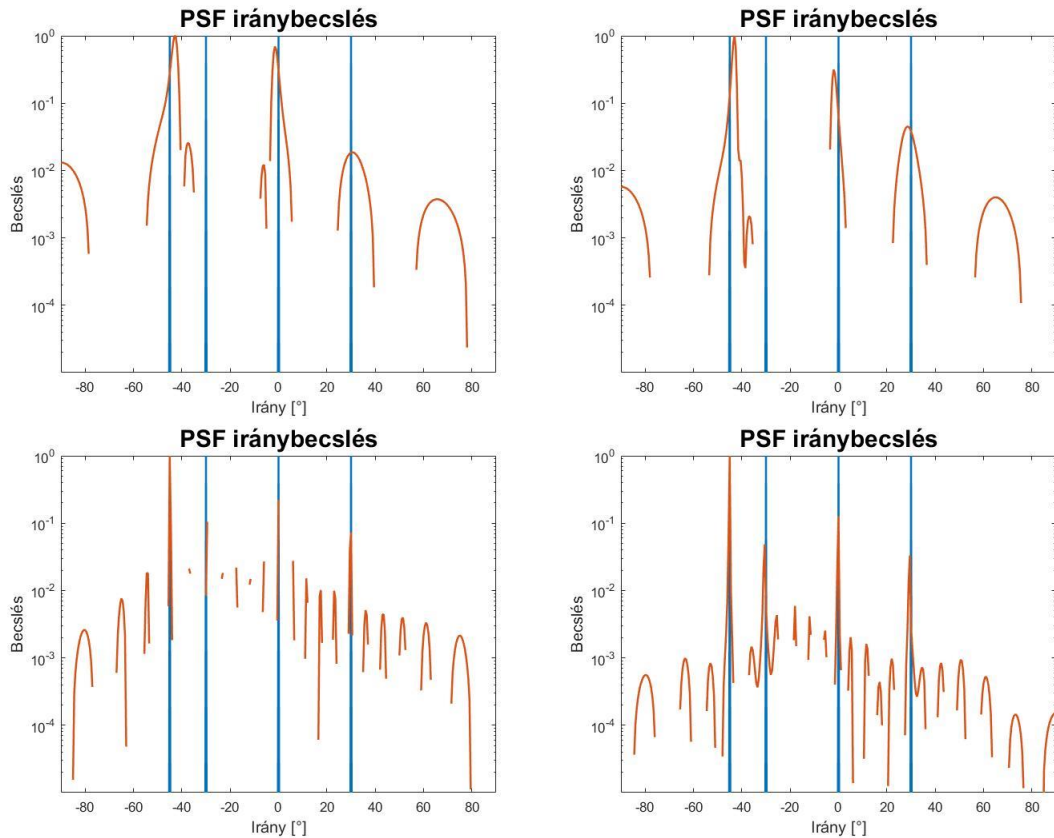
**9. ábra: CS iránybecslés zajmentes (bal) és 20 dB jel-zaj viszonyú (jobb) esetben, 500 Hz (fent) és 2500 Hz (lent) vizsgálati frekvencián.**

A Compressive Beamforming-ot kiegészítve  $l_1$  minimalizálással és szintkorrekcióval sokkal jobb zajtűrést és tisztább képet kapunk (10. ábra). Az algoritmus még 20 dB jel-zaj viszony mellett is viszonylag pontosan megtalálja a forrásokat az esetek többségében.

A pontszórás-függvény alapú Compressed Beamforming módszert használva az eredetihez hasonló eredményhez jutunk (11. ábra). Lényeges különbség azonban, hogy 2500 Hz-en ennek a módszernek jóval nagyobb a zajtűrése, mint a hagyományos Compressed Beamforming-nak.



10. ábra: CS iránybecslés (bal) és kiegészítésekkel (jobb), 100 dB (fent), 60 dB (középen) és 20 dB (lent) jel-zaj viszony mellett, 2500 Hz vizsgálati frekvencián.



11. ábra: CS iránybecslés zajmentes (bal) és 20 dB jel-zaj viszonyú (jobb) esetben, 500 Hz (fent) és 2500 Hz (lent) vizsgálati frekvencián.

### 3.3 Szimulációk három dimenzióban

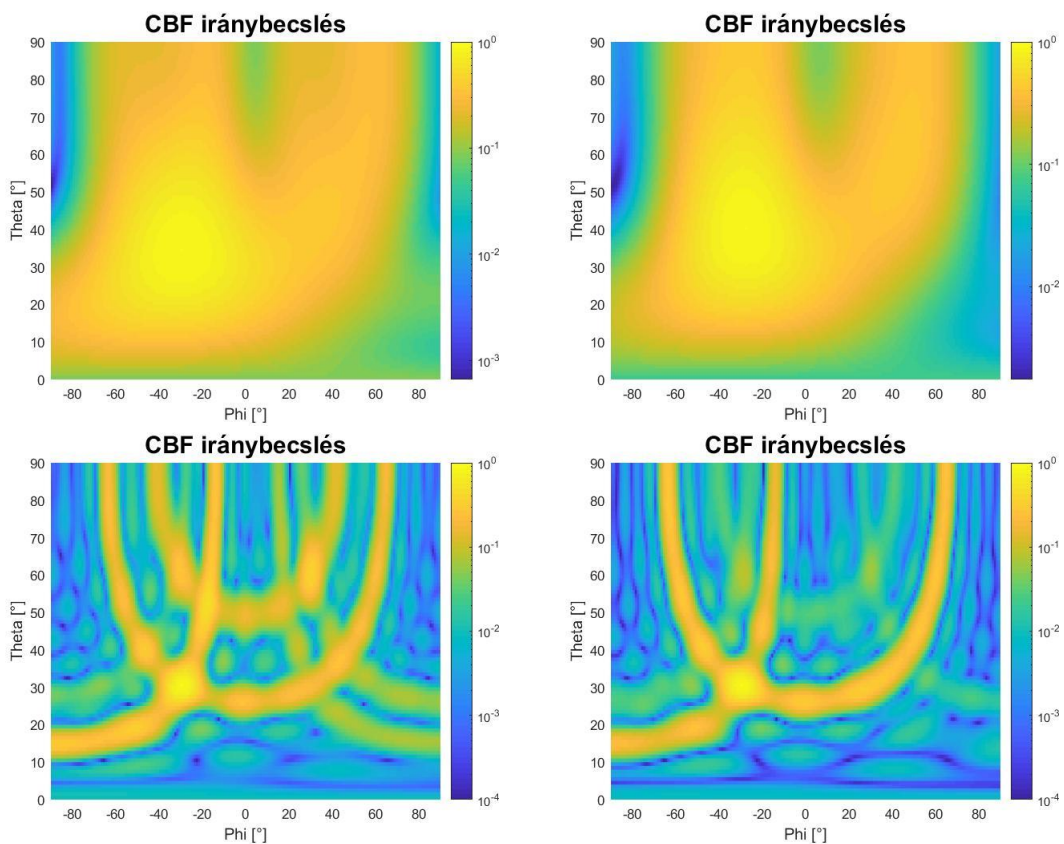
Háromdimenziós szimulációk esetén összesen 48 darab mikrofont használunk kereszt elrendezésben (azaz két egymásra merőleges vonalban, mik egyenként 24 mikrofonból állnak), két szomszédos mikrofon távolsága ebben az elrendezésben is 6 centiméter. A felső határfrekvencia ebben az esetben 2862,07 Hz. Vizsgálati frekvenciáknak itt is 500, 2500 és 5000 Hz-et választottam. A vászon 32851 pontból áll, ezek egy gömb felületén helyezkednek el, aminek a középpontjában van a mikrofontömb és a sugara 100 méter. A vászon pontjainak irányai két szöggel írhatók le, gömbi koordinátarendszerben, ahol  $\theta$  a z tengellyel bezárt szög,  $\varphi$  pedig az irány x-y síkra vett vetületének az x tengellyel bezárt szöge. A vászon által behatárolt szögtartomány:  $\varphi=[-90^\circ, +90^\circ]$  és  $\theta=[0^\circ, +90^\circ]$ . A térben négy pontszerű forrás helyezkedik el 150 méter távolságban és  $(\varphi, \theta)=(-30^\circ, 30^\circ), (30^\circ, 30^\circ), (30^\circ, 60^\circ), (-30^\circ, 60^\circ)$  irányokban. A források fehérzajt bocsátanak ki (a mintavételi frekvencia 48 kHz), szórásuk rendre 3, 1, 2 és 1 egységnyi. Mindhárom nyálábformáló algoritmus működését vizsgálom zajmentes és zajos esetben is, utóbbinál 20 dB-es jel-zaj



viszonnyal. Az eredményüket egy véletlenszerűen kiválasztott pillanatban (0,2 másodperces időtartamra vonatkoztatva) ábrázolom.

### 3.3.1 Conventional Beamforming

A 12. ábra mutatja a hagyományos nyalábformálás eredményét. A kétdimenziós szimulációhoz hasonlóan, 500 és 5000 Hz-en nem használható jól a módszer, ugyanis előbbinél az elmosódott kép, utóbbinál a térbeli átlapolódás révén létrejött szellemforrások miatt nem tudunk pontos iránybecslést végezni. 2500 Hz-en, zaj nélkül és zajjal is csak pontatlanul tudjuk megbecsülni az irányokat.

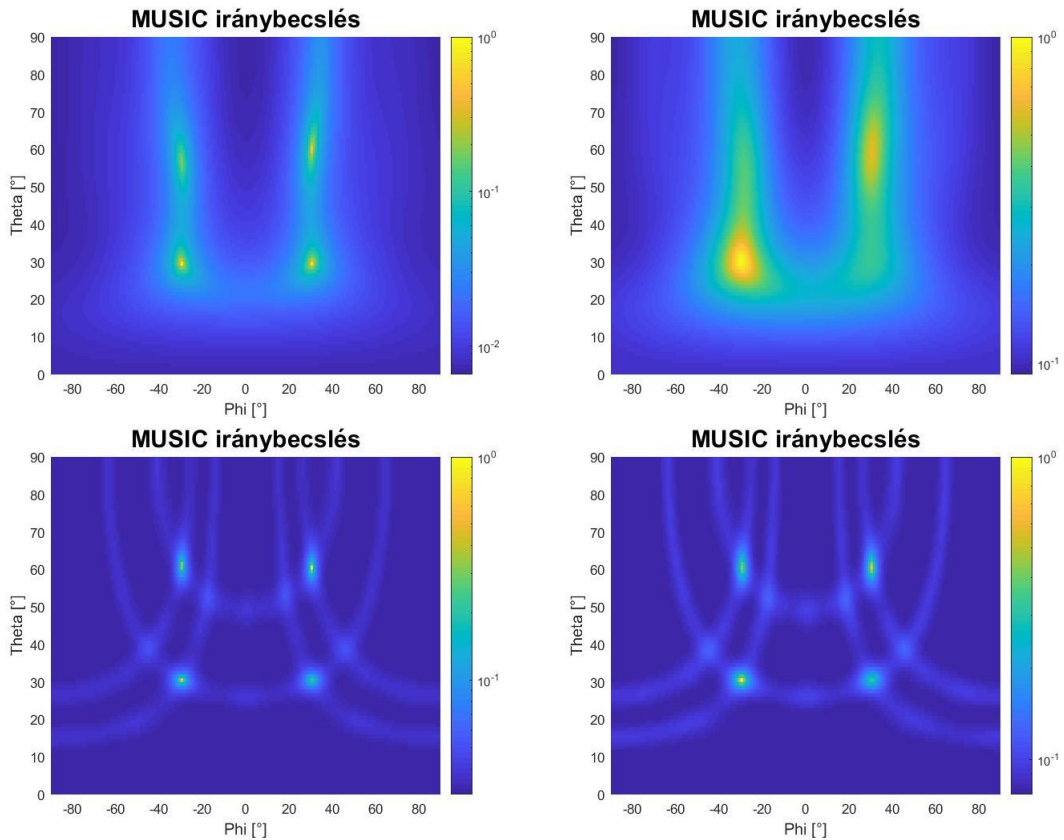


12. ábra: CBF iránybecslés zajmentes (bal) és 20 dB jel-zaj viszonyú (jobb) esetben, 500 Hz (fent) és 2500 Hz (lent) vizsgálati frekvencián.

### 3.3.2 Multiple Signal Classification

A MUSIC algoritmus is hasonló eredményeket ad két és három dimenzióban, ahogy a 13. ábra is látható. 500 Hz-es vizsgálati frekvencián, kellően jó jel-zaj viszony mellett jól használható a módszer. 2500 Hz-en, a határfrekvencia közelében viszonylag jó a zajtűrése, 20 dB-es jel-zaj viszonytal együtt is képes meglehetősen pontos

iránymeghatározásra. Amikor a MUSIC és a CBF is használható, a MUSIC pontosabb és nagyobb bizonyosságú becslést tud adni.



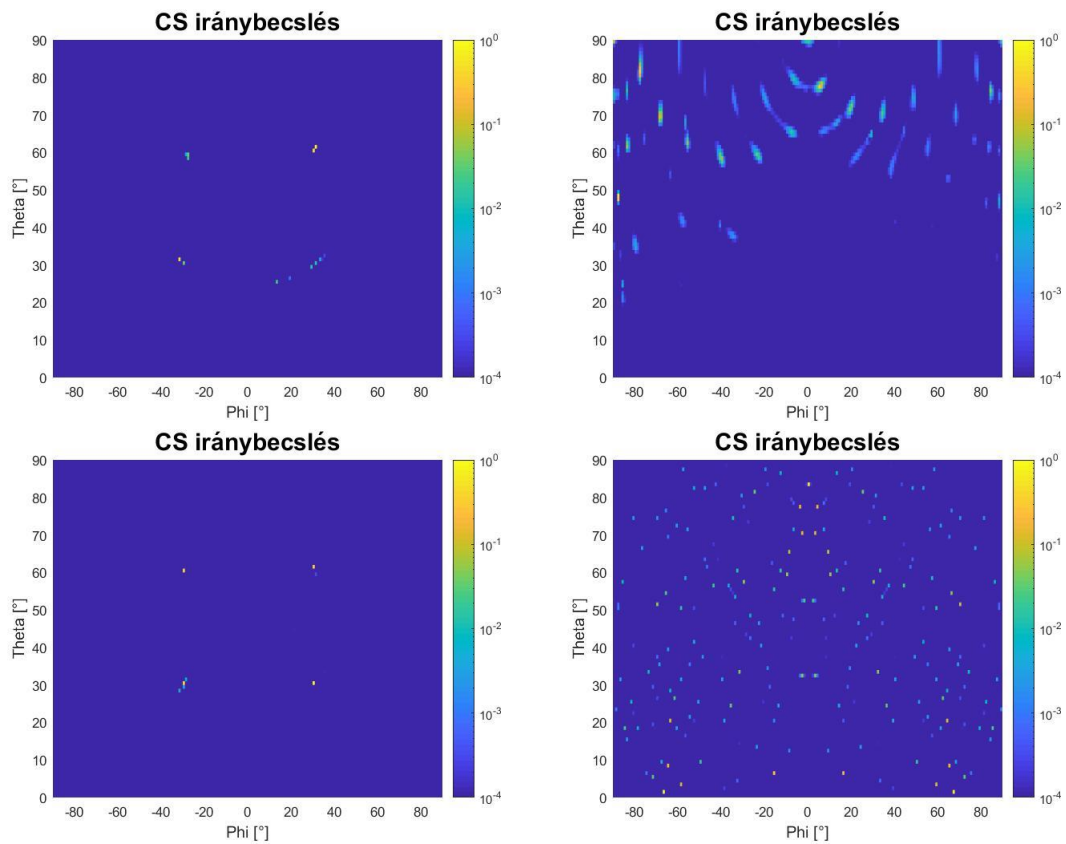
**13. ábra: MUSIC iránybecslés zajmentes (bal) és 20 dB jel-zaj viszonyú (jobb) esetben, 500 Hz (fent) és 2500 Hz (lent) vizsgálati frekvencián.**

### 3.3.3 Compressive Beamforming

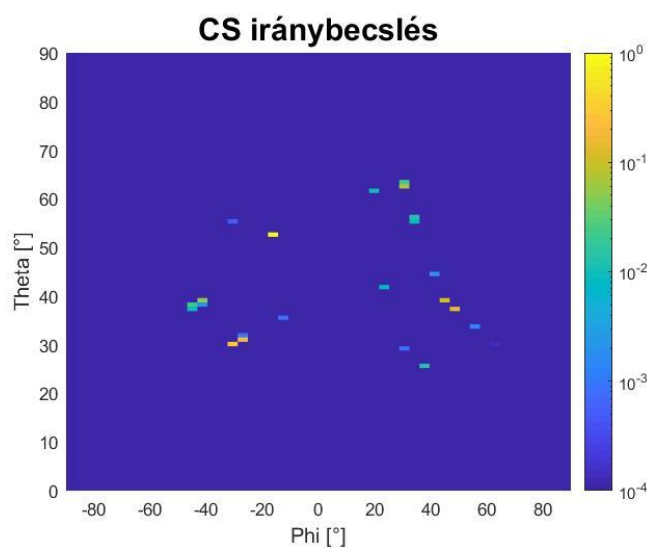
Bár a szakirodalomban még nincsenek háromdimenziós szimulációs eredmények a Compressive Beamforming-ről, sok a hasonlóság a kétdimenziós esettel. Három dimenzióban is csak 2500 Hz-en, kedvező jel-zaj viszony mellett vagy zaj nélkül tud jól használható kameraképet eredményezni (14. ábra). Figyelni kell továbbá arra is (kettő és három dimenzióban egyaránt), hogy a vászon pontjai/irányai kellően sűrűn helyezkedjenek el. Ha a források irányai nem illeszkednek jól a vászonra (ezt a problémát angolul „basis mismatch”-nek nevezik [19]), akkor előfordulhat, hogy a kapott eredmény helytelen (15. ábra). Ez a jelenség csak a Compressive Beamforming-ot érinti, a CBF-et és a MUSIC-ot nem.

$L_1$  minimalizálás és szintkorrekció bevezetésével (16. ábra) ismét egy jobb zajtűréssel rendelkező eljárást kapunk. A kétdimenziós szimulációkkal szemben azonban 20 dB jel-zaj viszony esetén a MATLAB cvx toolbox-a nem tudta megoldani a

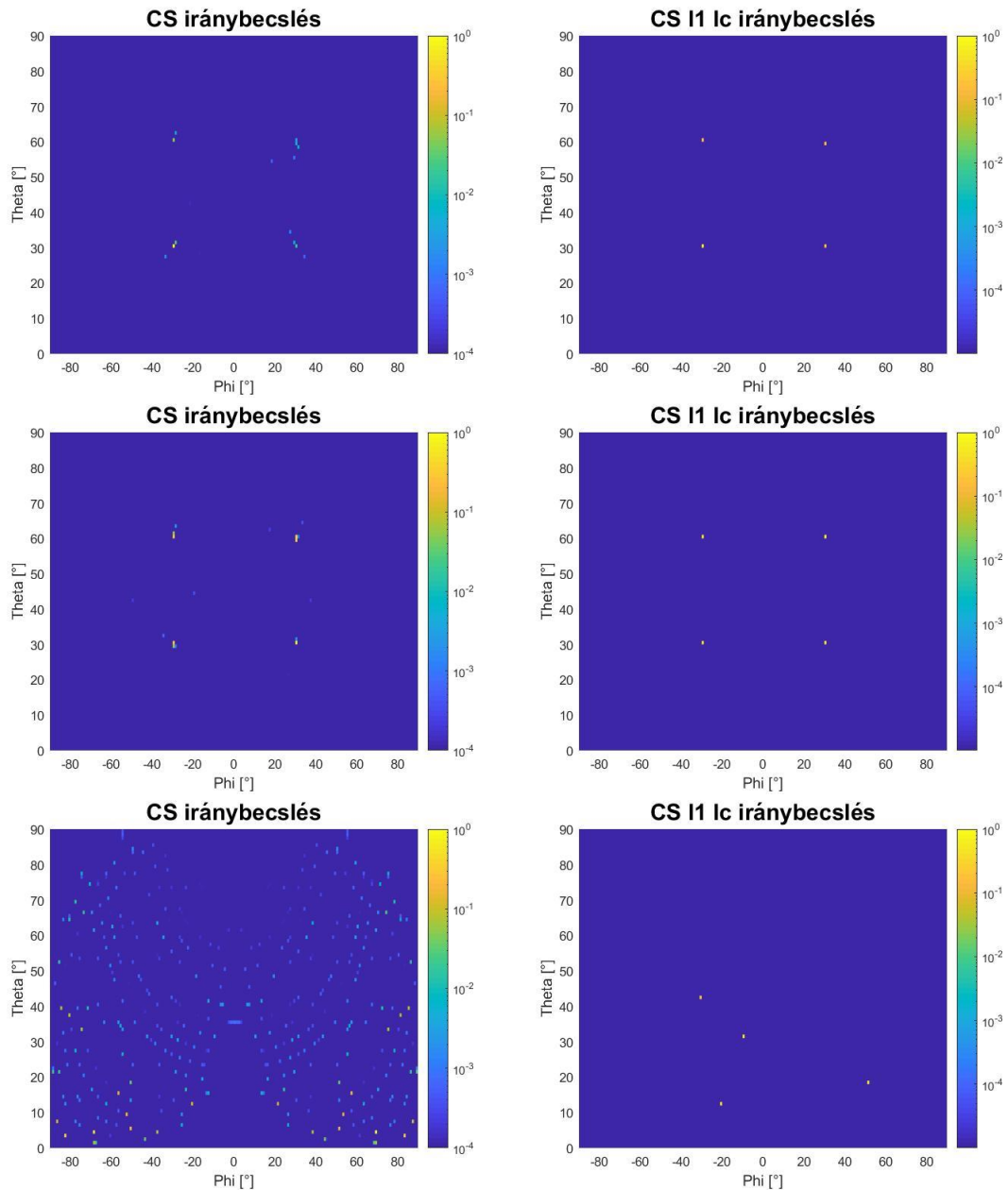
megadott konvex optimalizálási problémát. A ritkasági kényszerfeltétel ugyan teljesül, de nagyon sok zérustól különböző eleme van a megoldásnak, ezekből pedig nem tudjuk eldönteni, hogy hol van valójában a forrás.



**14. ábra: CS iránybecslés zajmentes (bal) és 20 dB jel-zaj viszonyú (jobb) esetben, 500 Hz (fent) és 2500 Hz (lent) vizsgálati frekvencián.**



**15. ábra: A „basis mismatch” miatti hibás kamerakép, 2500 Hz vizsgálati frekvencián, zajmentes esetben, 101-szer 51-es méretű vászonnal.**



**16. ábra: CS iránybecslés (bal) és kiegészítésekkel (jobb), 100 dB (fent), 60 dB (középen) és 20 dB (lent) jel-zaj viszony mellett, 2500 Hz vizsgálati frekvencián.**

A pontszórás-függvény alapú Compressive Beamforming három dimenzióban a PSF-mátrix nagy mérete miatt igencsak számításigényes (egy 0,2 másodperces időablakra végezve a becslést akár több óráig is eltarthat az iteráció), a többi módszernél sokkal lassabb, valós időben nem lehet alkalmazni.

## 4 Mérések, eredmények

Ebben a fejezetben a hagyományos nyálábformálás, a MUSIC és a Compressive Beamforming által feldolgozott mérési eredményeket mutatok be. Jelenleg még nem érhető el a szakirodalomban olyan eredmény, amelynek része a Compressive Beamforming gyakorlati mérésekben való alkalmazása.

### 4.1 Süketszobás mérések

A következő mérések a BME Akusztikai és Stúdiótechnikai Laboratórium süketszobájában lettek elvégezve. A körülmények az ideálshoz viszonylag közeliek, a jel-zaj viszony kedvező. A mérési eredményeket a szimulációkhoz is felhasznált keretrendszer segítségével dolgoztam fel.

A 4.1.1 szakaszban egy, illetve kettő fehérzaj-forrás által keltett hangok feldolgozásának eredményeit részletezem. Mindkét mérés egy 48 csatornás mikrofontömb segítségével történt, amelynek a szomszédos mikrofonjai közötti távolság 6 centiméter. Ez azt jelenti, hogy a térbeli átlapolódás határfrekvenciája 2862,07 Hz, ezért a vizsgálati frekvenciát 2500 Hz-nek választottam.

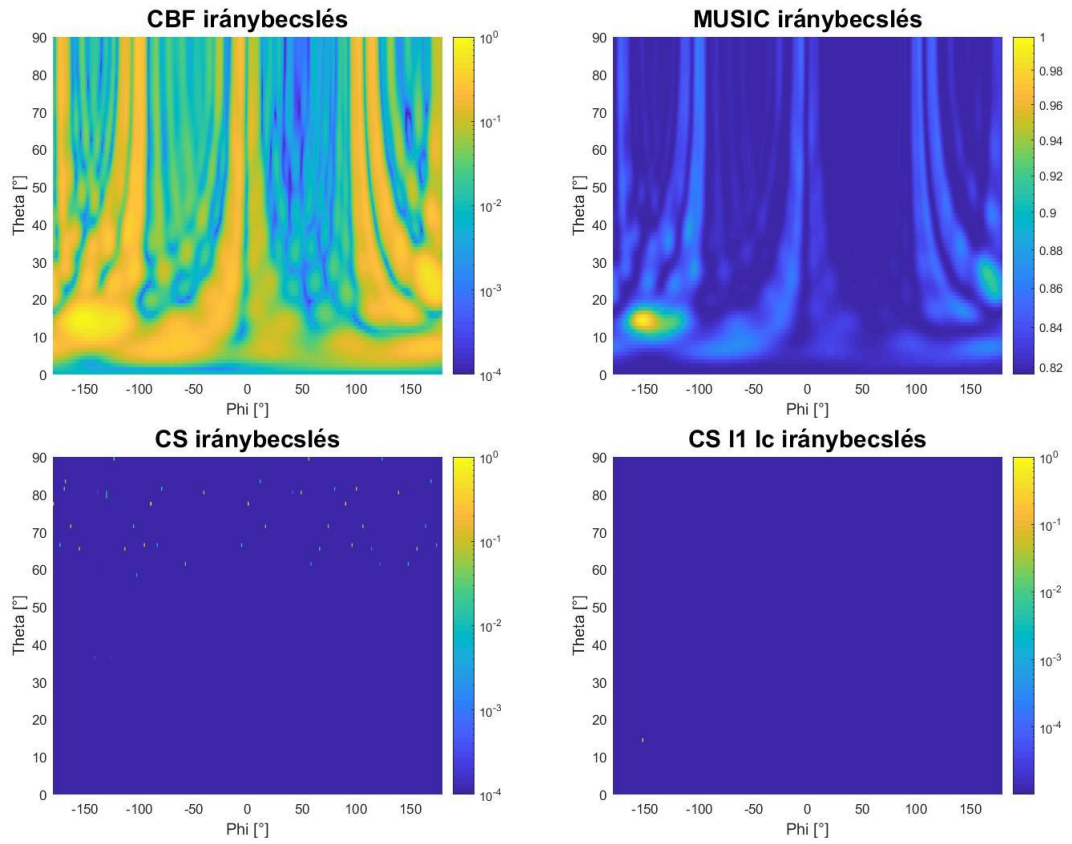
A 4.1.2 alfejezetben két különböző hangot kiadó forrás lokalizációját végzem el. A két hangszóró két különböző zeneszámot játszik le, és ezeket a hangokat egy 24 és 48 csatornás mikrofontömb veszi fel. A 48 csatornás rendszer ugyanaz, mint a fehérzajos mérés során használt, a 24 csatornásban viszont a szomszédos mikrofonok távolsága 8 centiméter. A határfrekvencia itt 2146,55 Hz, ennek megfelelően a vizsgálati frekvenciát 2000 Hz-re csökkentettem.

A mintavételi frekvencia mindkét esetben 48 kHz, a vászon szögtartománya pedig  $\varphi=[-180^\circ, +180^\circ]$  és  $\theta=[0^\circ, +90^\circ]$  gömbi koordináta-rendszerben, azaz a vászon pontjai egy félgömb felületen helyezkednek el.

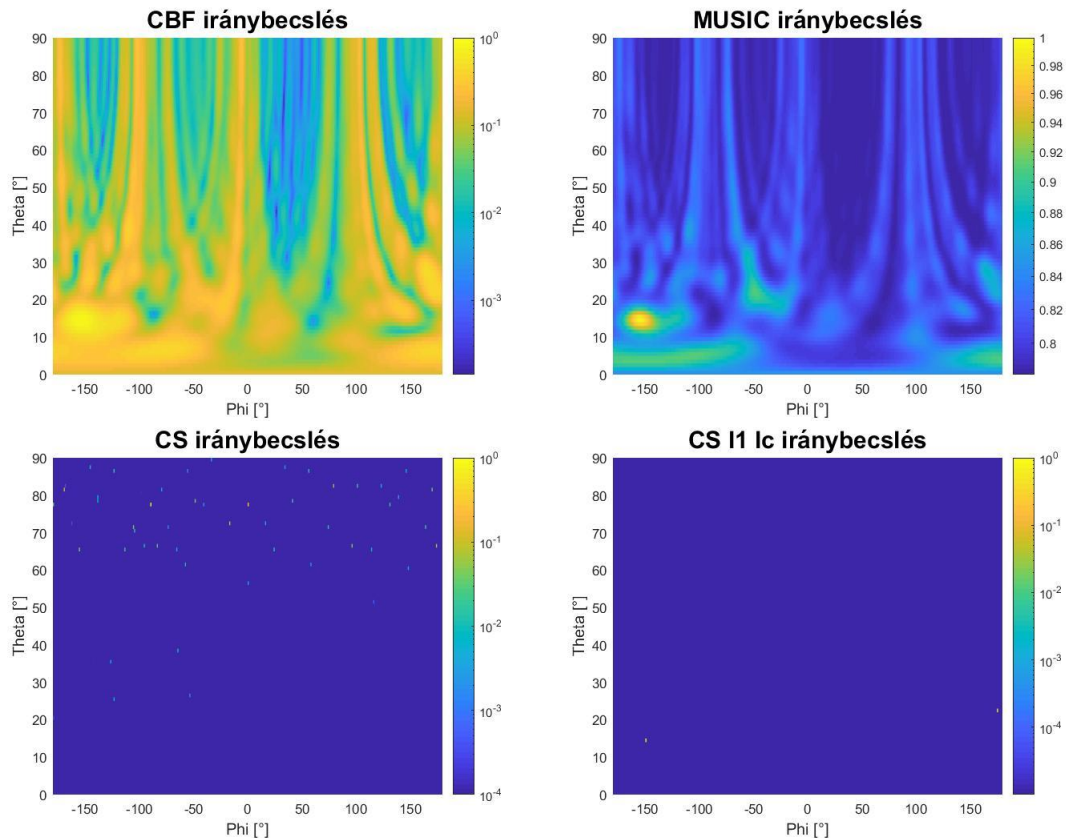
#### 4.1.1 Fehérzajt sugárzó hangforrások

A fehérzaj-források lokalizációja során hasonló eredményre juthatunk, mint a szimulációk esetén (17. ábra, 18. ábra). A Compressive Beamforming csak akkor tud helyesen becsülni, ha ki van egészítve újrasúlyozott  $l_1$ -minimalizálással és

szintkorrekcióval. A kiegészített CS egy forrásra pontos becslést tud adni, kettőre viszont csak akkor, ha az egyik forrás nem erősebb lényegesen a másikonál.



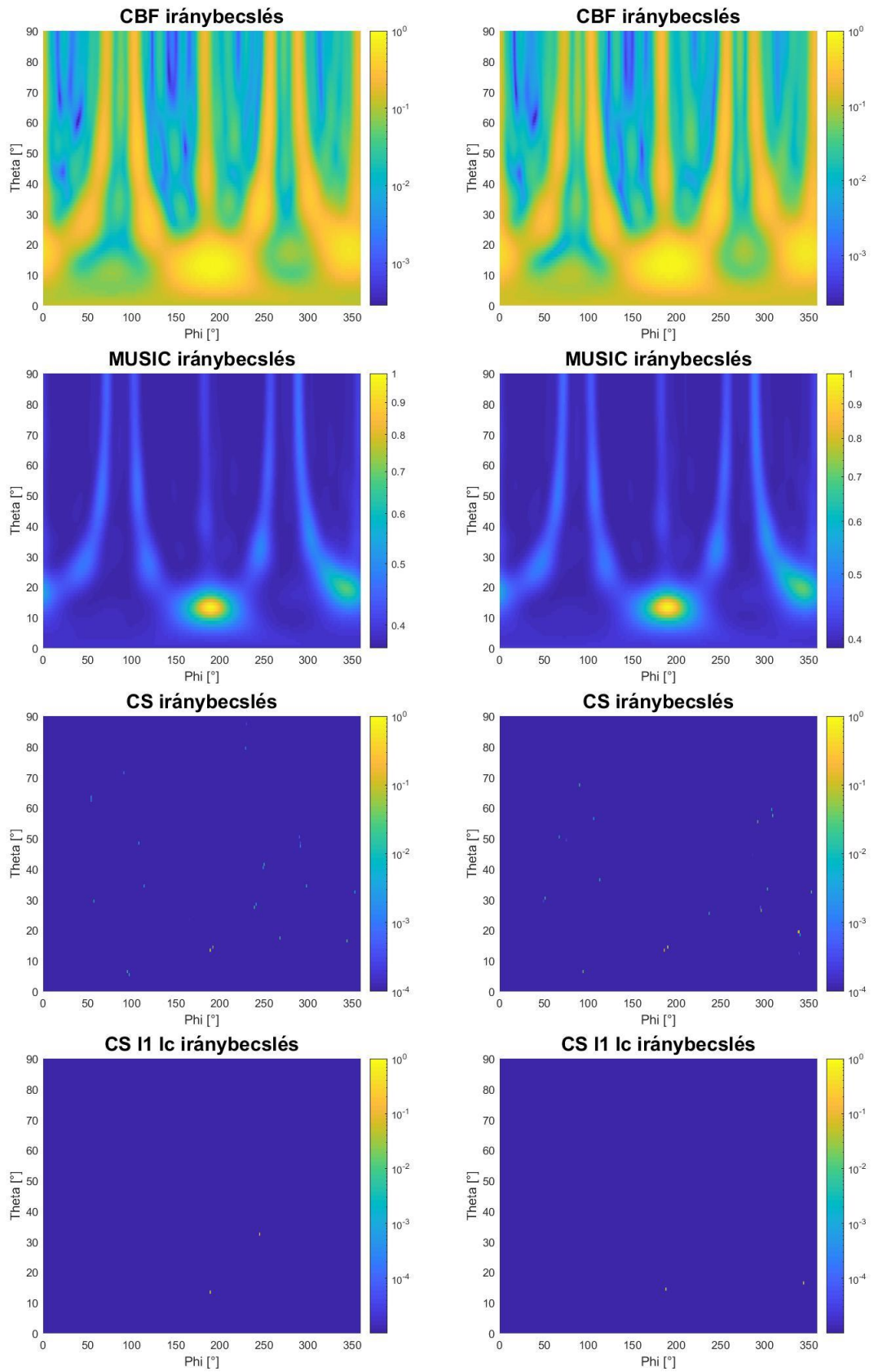
17. ábra: Egy fehérzaj-forrás lokalizációja.



18. ábra: Két fehérzaj-forrás lokalizációja.

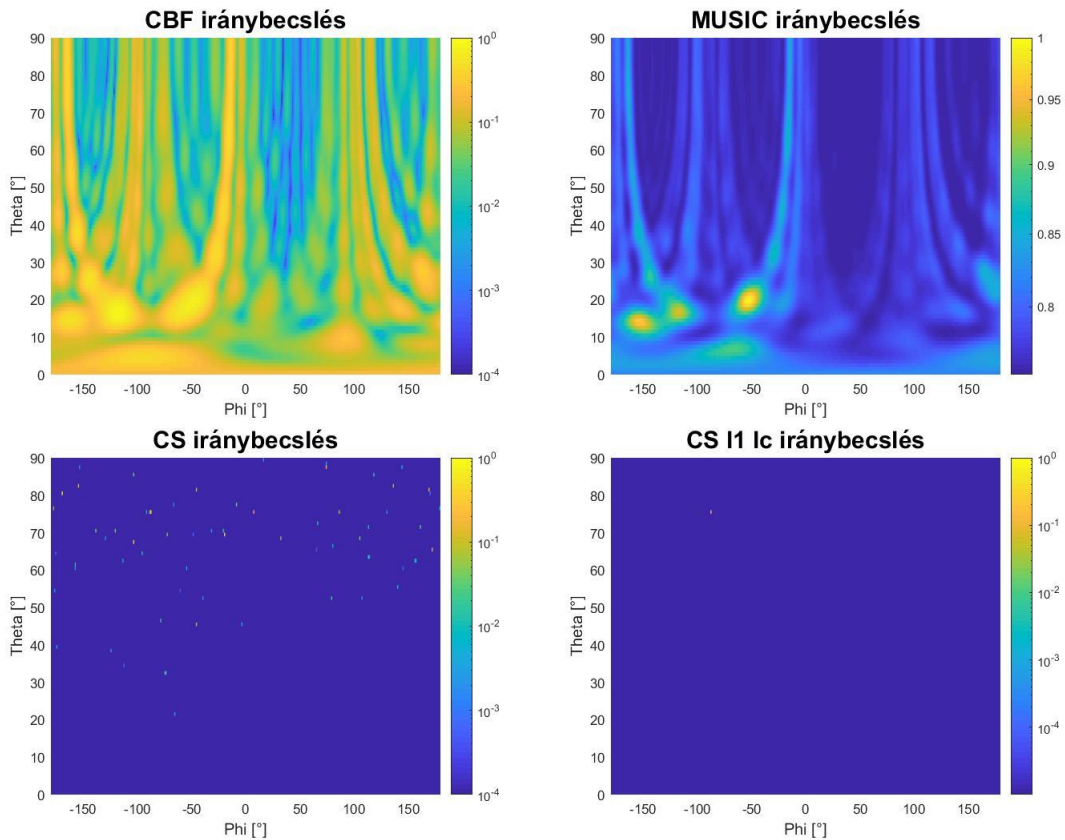
#### 4.1.2 Két különböző zeneszámot sugárzó hangforrások

Ezen mérés során a két hangszóró két különböző zeneszámot sugárzott (az egyikből egy Lady Gaga szám, a másikkól pedig egy Guns N' Roses szám). A 24 csatornás mérés eredményét két különböző időpontban ábrázoltam a 19. ábra. A kiegészített Compressive Beamforming ezek közül csak a másodikban vezetett helyes eredményre. Ez feltehetőleg azért történt így, mert az elsőben a hangforrások erőssége között jelentős eltérés volt, és az erősebben sugárzó hangszóró által kiadott hang valamilyen felületen visszaverődött, és ezt a reflektált hangot a mikrofonrendszer erősebbnek érzekelte, mint a másik hangforrás hangját. (Megjegyzés: itt a szögtartományt  $\varphi=[0^\circ, +360^\circ]$  és  $\theta=[0^\circ, +90^\circ]$ -ra módosítottam, hogy az eredmény jobban látható legyen.) A 48 csatornával történő iránybecslés (20. ábra) során ebben a véletlenszerűen kiválasztott időpontban nem sikerült ezzel a módszerrel helyes becslést adni, valószínűleg hasonló okokból.



19. ábra: Két különböző hangforrás lokalizációja 24 mikrofonnal, két különböző időpontban.





20. ábra: Két különböző hangforrás lokalizációja 48 mikrofonnal.

## 4.2 Szabadtéri mérések

Munkám során lehetőségem adódott arra, hogy részt vegyek egy olyan mérésben, ahol pilóta nélküli légi járművek (röviden drónok) röptetése során azok lokalizációja volt a cél, többféle módszerrel (pl. radar, optikai, akusztikai stb.). A mérést a korábban már részletezett 48 csatornás mikrontömbbel végeztük el. A süketszobához képest a körülmények itt kevésbé ideálisak, a jel-zaj viszony jóval rosszabb.



**21. ábra: A mérések ideje alatt reptetett drónok, balról jobbra: DJI Phantom4, Quad480, Eurotor4, Tarot680, Secopx8 és Flying Ranger.**



**22. ábra: A mérésekhez felhasznált 48 csatornás mikrofontömb.**

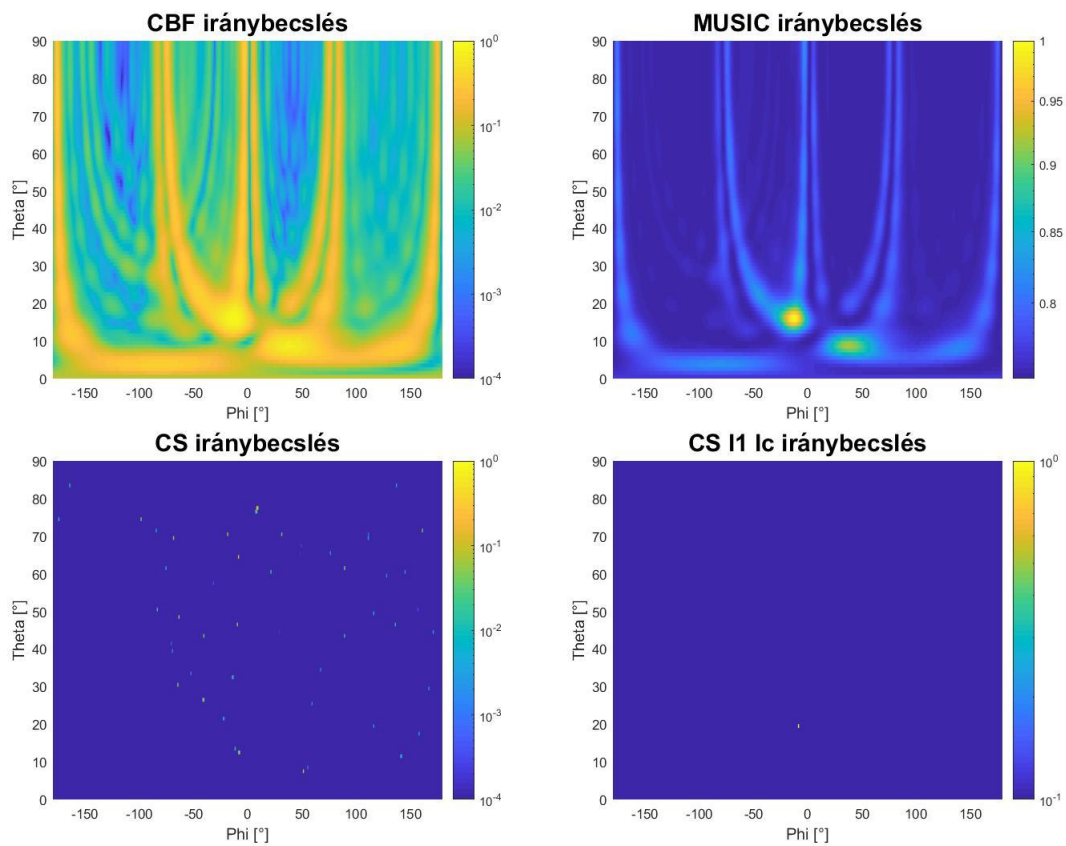
### **4.2.1 Flying Ranger**

A Flying Ranger nevű drónhoz tartozó kameraképek közül hármát választottam ki, amiket alább ábrázolok:

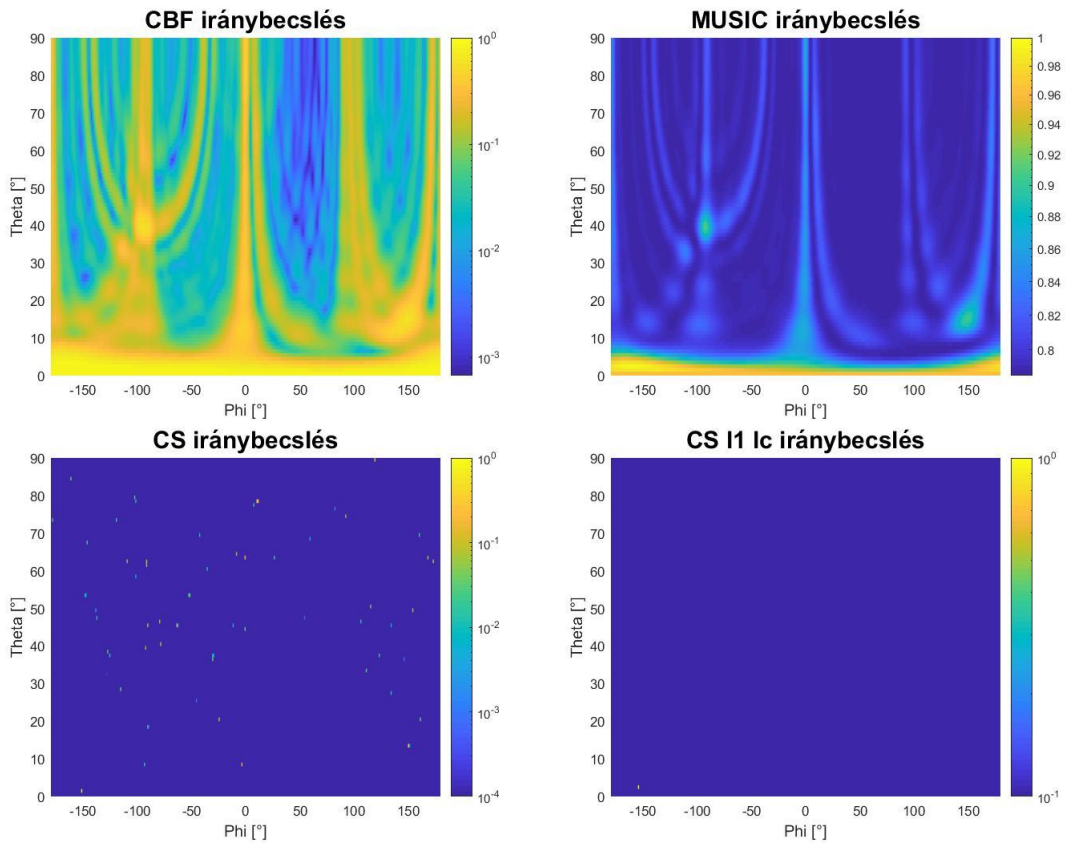
- Az első kameraképhez tartozó időpont a mérés kezdete utáni 1 perc 15 másodperc, amikor a drón a leszállási terület felett volt egy helyben (23. ábra).

- A második kameraképhez tartozó időpont a mérés kezdete utáni 1 perc 35 másodperc, amikor a drón a leszállási területtől néhány méterre, a mikrofontömb előtt volt (24. ábra).
- Az harmadik kameraképhez tartozó időpont a mérés kezdete utáni 2 perc 20 másodperc, amikor a drón a leszállási terület felett volt egy helyben, az első kamerakép helyénél magasabban forgott egy helyben (25. ábra).

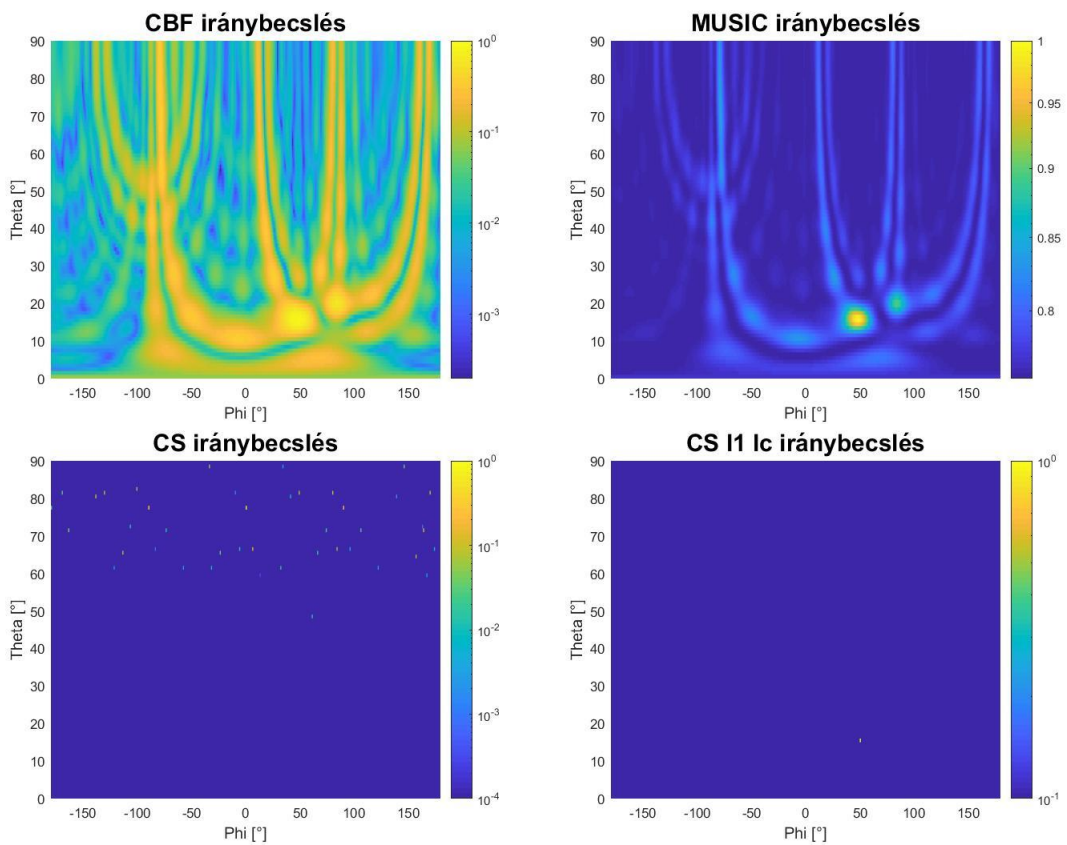
A MUSIC továbbra is jó kameraképet alkot, a CS viszont csak kiegészítésekkel együtt tud jó minőségű becslést adni.



23. ábra: Kamerakép a mérés kezdete után 1 perc 15 másodperccel.



24. ábra: Kamerakép a mérés kezdete után 1 perc 35 másodperccel.



25. ábra: Kamerakép a mérés kezdete után 2 perc 20 másodperccel.

### 4.3 Továbbfejlesztési lehetőségek

Jelen munka többféleképpen is kibővíthető, továbbfejleszhető. Érdeemes megfigyelni, hogy az algoritmusok miként viselkednek különböző típusú hangforrásokra a fehérzajon kívül (pl. azonos frekvenciájú szinuszjeleket sugárzó források lokalizációjára nem alkalmas a MUSIC, mert azok nem korrelálatlanok). A kör és a spirál mikrofon elrendezések is a gyakorlatban megvalósítható elrendezések, ezért érdemes ezekkel is szimulációkat és méréseket elvégezni. Az  $l_1$ -minimalizálással és szintkorrekcióval kiegészített Compressive Beamforming megfelelő vizsgálati frekvencián pontos iránybecslésre képes, de hasznos lehet megvizsgálni, hogy alacsony frekvenciákon mennyire működésképes a módszer. Az rosszul definiált akusztikai vászon esetén fellépő Basis Mismatch probléma (ezen belül is a virtuális forráspozíciók rossz iránya és távolsága), és ennek kiküszöbölése a CS folytonossá tételével szintén érdekes továbbfejlesztési irány lehet. Bevezethetjük a CS-nek egy olyan változatát, ahol a vásznat leszűkítjük azoknak a pontoknak/irányoknak a halmazára, ahol a CBF vagy a MUSIC által forrást találtunk (például olyan virtuális forráspozíciókra, amikre a legnagyobb becsült forrásamplitúdóhoz képest az elem egy mínusz 10 dB-es küszöbön felül van).

## 5 Összefoglalás

Dolgozatomban céloom nyalábformáló algoritmusok bemutatása, illetve azok tesztelése volt szimulációkkal. Az algoritmusokat akusztikai kameraként használt mikrofonok vett jelen alkalmazva a mikrofontömb alkalmassá válik arra, hogy hangforrásokot lokalizáljon.

Három főbb nyalábformálási módszert mutattam be, ezek a hagyományos nyalábformálás (CBF), a MUSIC, és a Compressive Sensing alapú Compressive Beamforming. Ezek közül a CBF a legegyszerűbb, ugyanakkor ez az algoritmus eredményezi a legkisebb felbontású kameraképet, és kisfrekvenciás hangforrások helybecslésére nem alkalmas. A MUSIC egy lineáris algebrai módszer, aminek előnye a CBF-fel szemben, hogy jobb a zajtűrése, kisfrekvencián is képes forráslokalizációra, és nagyobb a felbontása, hátránya viszont, hogy csak inkohereus források esetén működik. A Compressive Beamforming a nyalábformálást egy konvex optimalizálási problémára vezeti vissza, számításigénye jóval nagyobb, mint a CBF-é és a MUSIC-é, és a zajtűrése a MUSIC-nál kisebb. Az itt bevezetett ritkasági kényszerfeltételnek köszönhetően azonban nagyon nagy felbontású kameraképet kaphatunk vele. Az módszert kiegészíthetjük újrasúlyozott  $l_1$ -minimalizálással és szintkorrekcióval, illetve bevezethetjük egy olyan változatát, amely pontszórásfüggvényeken alapul. Mindkettő nagyobb zajtűrésű eljárást eredményez.

Az algoritmusokkal két- és háromdimenziós szimulációkat végeztem, ahol a forrásokot fehérzajnak feltételeztem. Két dimenzióban vonal, három dimenzióban pedig kereszt alakú mikrofonelrendezést használtam. Végül süketszobai és szabadtéri mérési eredményekkel is teszteltem a nyalábformáló módszereket.

A Compressive Beamforming nyalábformálási módszer szimulációk során pontszerű forrásokot feltételezve, reflexiók nélkül, megfelelő jel-zaj viszonytal és kiegészítésekkel szép kameraképet produkált. Gyakorlatban, mérések során azonban több zavartényező is ronthatja az eredményt, mint például a nem megfelelően definiált vászon, a reflexiók vagy rossz jel-zaj viszony, amik miatt ez az új módszer nem mindig tud teljesen megbízhatóan működni, és a megoldás ritkasága ellenére nem mindig előnyös a hagyományos algoritmusokkal szemben.

## **Köszönetnyilvánítás**

Szeretném megköszönni konzulensemnek, Dr. Fiala Péternek a segítségét, tanácsait és iránymutatását. Köszönöm Dr. Rucz Péternek is a munkám haladását segítő ötleteit, észrevételeit. Végül megköszönném családomnak támogatásukat, építő kritikáikat és érdeklődésüket.

## Irodalomjegyzék

- [1] A. Xenaki, P. Gerstoft, K. Mosegaard: „Compressive beamforming”. The Journal of the Acoustical Society of America, Vol. 136 (1), 2014, 260-271.
- [2] A. Xenaki, P. Gerstoft: „Grid-free compressive beamforming”. The Journal of the Acoustical Society of America, Vol. 137 (4), 2015, 1923-1935.
- [3] Y. Park, Y. Choo, W. Seong: „Multiple snapshot grid free compressive beamforming”. The Journal of the Acoustical Society of America, Vol. 143, 2018, 3849-3859.
- [4] M. Lustig, D. Donoho, J. M. Pauly: „Sparse MRI: The application of compressed sensing for rapid MR imaging.” Magn. Reson. Med. 58, 2007, 1182–1195.
- [5] N. Wagner, Y. C. Eldar, Z. Friedman: „Compressed beamforming in ultrasound imaging”. IEEE Transactions on Signal Processing Vol. 60, 2012, 4643–4657.
- [6] E. J. Candés, T. Tao: „Decoding by linear programming”. IEEE Trans. Inf. Theory 51, 4203–4215 (2005).
- [7] H. Yao, P. M. Shearer, P. Gerstoft: „Compressive sensing of frequency-dependent seismic radiation from subduction zone megathrust ruptures”. Proc. Natl. Acad. Sci. U.S.A. Vol. 110, 2013, 4512–4517.
- [8] J. Romberg: „Imaging via compressive sampling”. IEEE Signal Processing Magazine Vol. 25, 2008, 14–20.
- [9] R. Schmidt: „Multiple emitter location and signal parameter estimation”. IEEE Transactions on Antennas and Propagation Vol. 34, 1986, 276–280.
- [10] E. J. Candes, M. B. Wakin: „An introduction to compressive sampling”. IEEE Signal Processing Magazine Vol. 25, 2008, 21–30.
- [11] R. G. Baraniuk: „Compressive sensing”. IEEE Signal Processing Magazine Vol. 24, 2007, 118–121.
- [12] E. J. Candés, J. Romberg, T. Tao: „Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information”. IEEE Transactions on Information Theory Vol. 52, 2006, 489–509.
- [13] M. Elad: „Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing Pages”. Springer, New York, 2010, pp. 1–359.
- [14] S. Foucart, H. Rauhut: „A Mathematical Introduction to Compressive Sensing”. Springer, New York, 2013, pp. 1–589.
- [15] J. A. Tropp: „Just relax: Convex programming methods for identifying sparse signals in noise”. IEEE Transactions on Information Theory Vol. 52, 2006, 1030–1051.



- [16] E. J. Candés, M. B. Wakin, S. Boyd: „Enhancing sparsity by reweighted  $l_1$  minimization,”. J. Fourier Anal. Appl. Vol. 14, 2008, 877–905.
- [17] C. F. Mecklenbräuker, P. Gerstoft, A. Panahi, and M. Viberg: „Sequential Bayesian sparse signal reconstruction using array data”. IEEE Transactions on Signal Processing Vol. 61, 2013, 6344–6354.
- [18] Fiala Péter, Rucz Péter: „Hangjelek digitális feldolgozása az akusztikai gyakorlatban”. Mérési leírás.
- [19] Y. Chi, L. L. Scharf, A. Pezeshki, A. R. Calderbank: “Sensitivity to basis mismatch in compressed sensing”. IEEE Transactions on Signal Processing Vol. 59, 2011, 2182–2195.