

ETLYAK Eksamensprojekt

Sum & delay beamforming.

Adam Høj AU617583 201803767
Emil Seilov AU634145 201907490

15. juni 2022



Indhold

1	Introduktion	2
2	Projektets formål	2
3	Teori	3
3.1	Beam pattern	3
3.2	Grating lobes	6
3.3	Steering	8
3.4	Forsinkelse af signaler	9
3.4.1	FIR Filter	10
3.5	Korrigteret DAS-output.	11
4	Test af projektet	13
5	Resultater	17
5.1	Test i lyddødt rum	17
5.1.1	5 mics, 20 cm	18
5.2	Test i lytterummet	22
6	Diskussion	23
7	Konklusion	24
8	Bilag	26

1 Introduktion

Delay and Sum (DAS)-beamforming er en teknik hvor et mikrofon array med m mikrofoner opfanger lyd simultant. Efterfølgende bruges en signalbehandlings-algoritme til at forsinke de forskellige mikrofoninputs. Forsinkelsenstiden er direkte koblet til placering af den givne lydkilde. Algoritmen giver mulighed for at vælge et punkt i et plan og lytte til dette.

Projektet vil efterprøve teorien omkring DAS-beamforming ved at opstille teorien, en række test-cases der vil blive behandlet og diskuteret, for til sidst at komme med en konklusion.

Lydfiler til projektet findes i den vedlagte zipfil og på Github.

2 Projektets formål

Der vil blive bygget et array med fem mikrofoner, fordelt med 20 cm afstand. Teorien baserer sig på artikler af Greensted, Grythe og Anguera[1, 2, 3]. Vi vil tage udgangspunkt i teorien og opstille en test, hvor vi undersøge om teorien virker i praksis. Forsøgene udføres i AudioLab på Aarhus Universitet hvor det lyddøde rum og lyttestuen vil blive brugt.

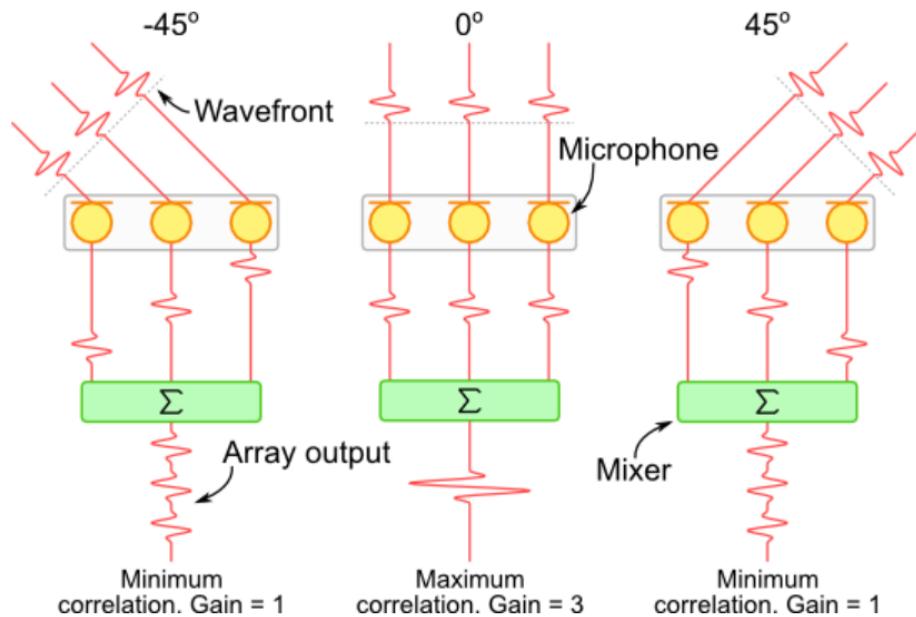


Fig. 1: Med tre mikrofoner og en mixer, vil udgangssignalet være afhængig af bølgefrontens vinkel

Beamforming handler om at opnå korrelation mellem et array af mikrofoner, ved at forsinke de målte signaler. Figur 1 viser, at der er max korrelationer, når bølgefronten står vinkelret på mikrofon-arrayet og man kan på den måde opnå en forstærkning af det indgående signal.

Det bør vise sig muligt at optage signaler fra de 5 mikrofoner for derefter at bruge digital signalbehandling til at udføre *beamforming* som vist i figur 8. Ved beamforming kan man retningsbestemme sin optagelse. Det bygger på den omvendte teknologi af line array systemer, hvor man retningsbestemmer lyd med højtalere.

3 Teori

I dette afsnit gennemgås den forskellige teori man skal have kendskab til, for at beamform.

3.1 Beam pattern

En beam pattern er mic-arrayets karakteristik. Mikrofonerne i projektet er ideelt set omnidirektionelle. Når man sætter flere mikrofoner sammen vil der opstå interferens mellem mikrofonerne[3]. Den teoretiske karakteristik for mikrofon-arrayets beam pattern kan beregnes med eq. 1.

$$z_m = 20 \cdot \log_{10} \left(\frac{1}{M} \sum_{m=0}^{M-1} \exp \left(\frac{j2\pi \cdot f \cdot m \cdot l \cdot \sin(\theta)}{c} \right) \right) \quad (1)$$

hvor

- M : antal af mikrofoner
- l : afstand mellem to mikrofoner [m]
- f : Mic arrayets gain ved frekvensen f [Hz]
- θ : Planbølges indfaldsvinkel på array (målt vinkelret på arrayet)
- c : lydens hastighed i luft [$343.3 \frac{\text{m}}{\text{s}}$]

Resultatet af eq. 1 er en beam pattern med et main lobe, side lobes og grating lobes. Det vises på figur 2, som er et 3D rum, at main lobe peger med et punkt på et plan, hvorpå lydgivere kan befinde sig. Hvad der er indkapslet indenfor det blå areal, er hvad der bliver modtaget højest. Det betyder ikke at der ikke modtages lyd fra andre vinkler. Disse lobes er frekvens- og retningsbestemte samt en kombination af antallet af mikrofoner og deres indbyrdes afstand.

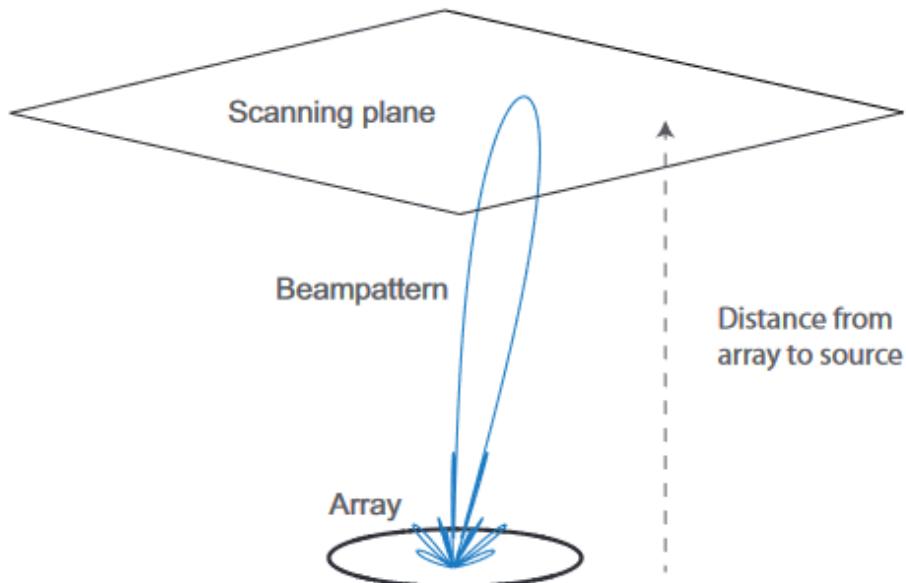


Fig. 2: Figur fra Beamforming algorithms - beamformers[2]

Eq. 1 bruges nu til at kigge på forskellige situationer for vores eget array. Når man kigger på en graf for beam patterns, så er forstærkningen højest tættest på kanten og lavest mod midten. Vinkel 0° er main lobe og de andre vinkler peger ud mod lydgiveres retning, vinkelret på arrayet.

På figur 3 kigges på 5 mikrofoner med 20 cm afstand mellem hver mikrofon. Det ses at lavfrekvent materiale ser omnidirekktionelt ud. Det er mellem 250 Hz og 500 Hz at der begynder at opstå interferens mellem mikrofonerne. Det betyder også at det først er fra dette frekvensområde, man kan begynde at beamform. Kigger man på 4000 Hz ses mange side lobes ift. figur 4 og figur 5. Dæmpningen på figur 3's side lobes er større og der er længere mellem grating lobes (som forklares i næste afsnit).

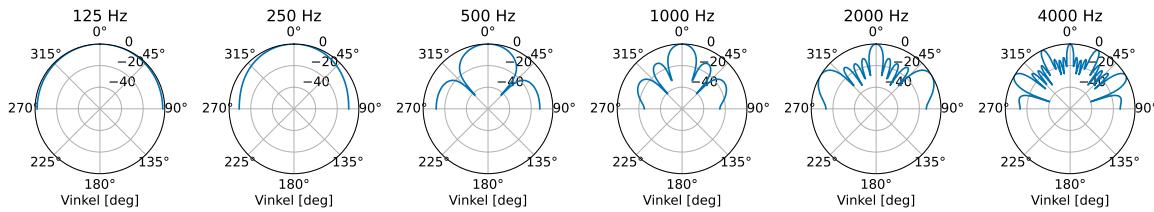


Fig. 3: Polarpatterns for 5 mikrofoner med 20 cm afstand. Vist i oktavbånd fra 125 Hz til 4000 Hz

Figur 4 viser 3 mikrofoner placeret med 20 cm mellem hinanden. Teoretisk set bliver mikrofonarrayet dårligere end de to andre figurer. Frekvensen man kan beamform fra ligger mellem 500 Hz og 1000 Hz

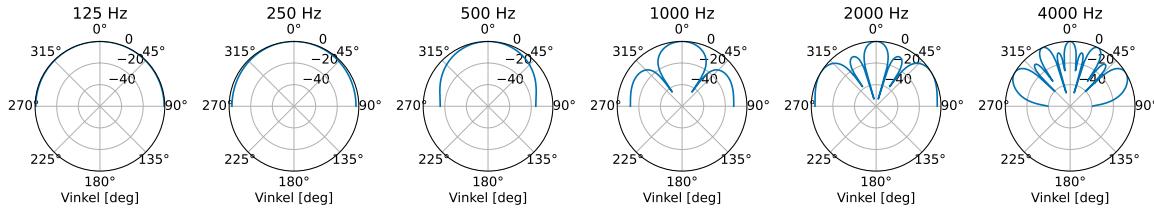


Fig. 4: Polarpatterns for 3 mikrofoner med 20 cm afstand. Vist i oktavbånd fra 125 Hz til 4000 Hz

Hvis man øger afstanden mellem de 3 mikrofoner til 40 cm, så får man figur 5. Når afstanden mellem mikrofoner øges, kan man beam form længere ned i frekvens. Ved 250 Hz begynder dens beam pattern at krumme indad mod midten, hvor figuren med 20 cm afstand stadig ligger ude på kanten. Den større afstand mellem mikrofonerne giver flere grating lobes når man går op i frekvens.

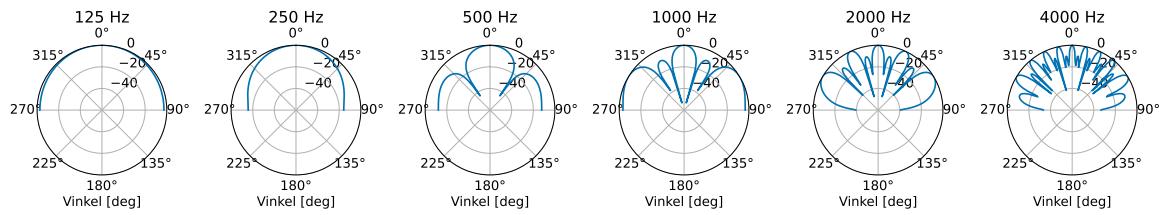


Fig. 5: Polarpatterns for 3 mikrofoner med 40 cm afstand. Vist i oktavbånd fra 125 Hz til 4000 Hz

Fælles for alle beam patterns er, at der opstår kamfiltrering (comb filtering) lige så snart man har mere end én mikrofon. Kamfiltrering er lyden af, at noget er ude af fase. Det lyder *hult* og man kan høre at der mangler noget.

3.2 Grating lobes

For et lineært, 1 dimensionelt array opstår der ekstra *lobes* til forskellige vinkler, med samme forstærkning som main lobe. Grating lobes er uønskede i et system, da de opfanger signaler fra andre retninger end main lobe, uden dæmpning.

Grating lobes opstår når den ekstra distance en bølgefront flytter sig, er et multiplum af signalets bølgelængde. I denne situation er det modtagende signal for hvert element (mikrofon) korrelerede og derfor opstår ingen dæmpning af signalet.

Vinklerne, hvortil grating lobes opstår, kan beregnes [1, Grating lobes]:

$$\theta = \sin^{-1} \left(\frac{n \cdot c}{f \cdot l} \right) \quad (2)$$

hvor

- l : afstand mellem mikrofoner
- c : lydens hastighed i luft
- f : frekvens for indgående signal
- n : et heltal på grating lobe nummer

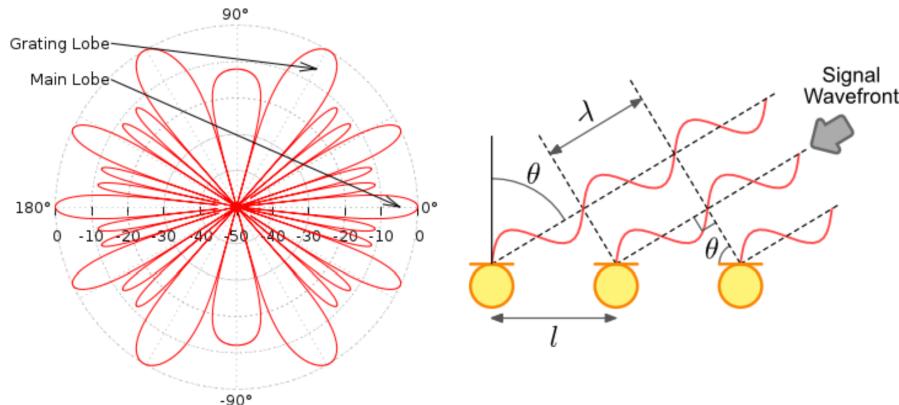


Fig. 6: Grating Lobes. Figur fra Labbook Pages[1]

Ud fra eq. 2 beregnes vinklerne til alle grating lobes. Resultaterne ses i tabel 1 og 2. Det giver ikke mening at kigge på frekvenser, hvor grating lobes ikke opstår. Derfor beregnes der for 1 kHz, 2 kHz og 4 kHz. Når man forsøger at regne grating lobes større end 90° får man et komplekst tal. Beam patterns gentager sig i hver kvadrant og derfor kan man ikke beregne for 360°.

Tabel 1: Grating lobes ved 20 cm mikrofon afstand

1 kHz		2 kHz		4 kHz	
n	angle (deg)	n	angle (deg)	n	angle (deg)
0	0	0	0	0	0
1	~	1	59.04	1	25.39
2	~	2	~	2	59.04

Tabel 2: Grating lobes ved 40 cm mikrofon afstand

1 kHz		2 kHz		4 kHz	
n	angle (deg)	n	angle (deg)	n	angle (deg)
0	0	0	0	0	0
1	59.04	1	25.39	1	12.38
2	~	2	59.04	2	25.39
3	~	3	~	3	40.03
4	~	4	~	4	59.04

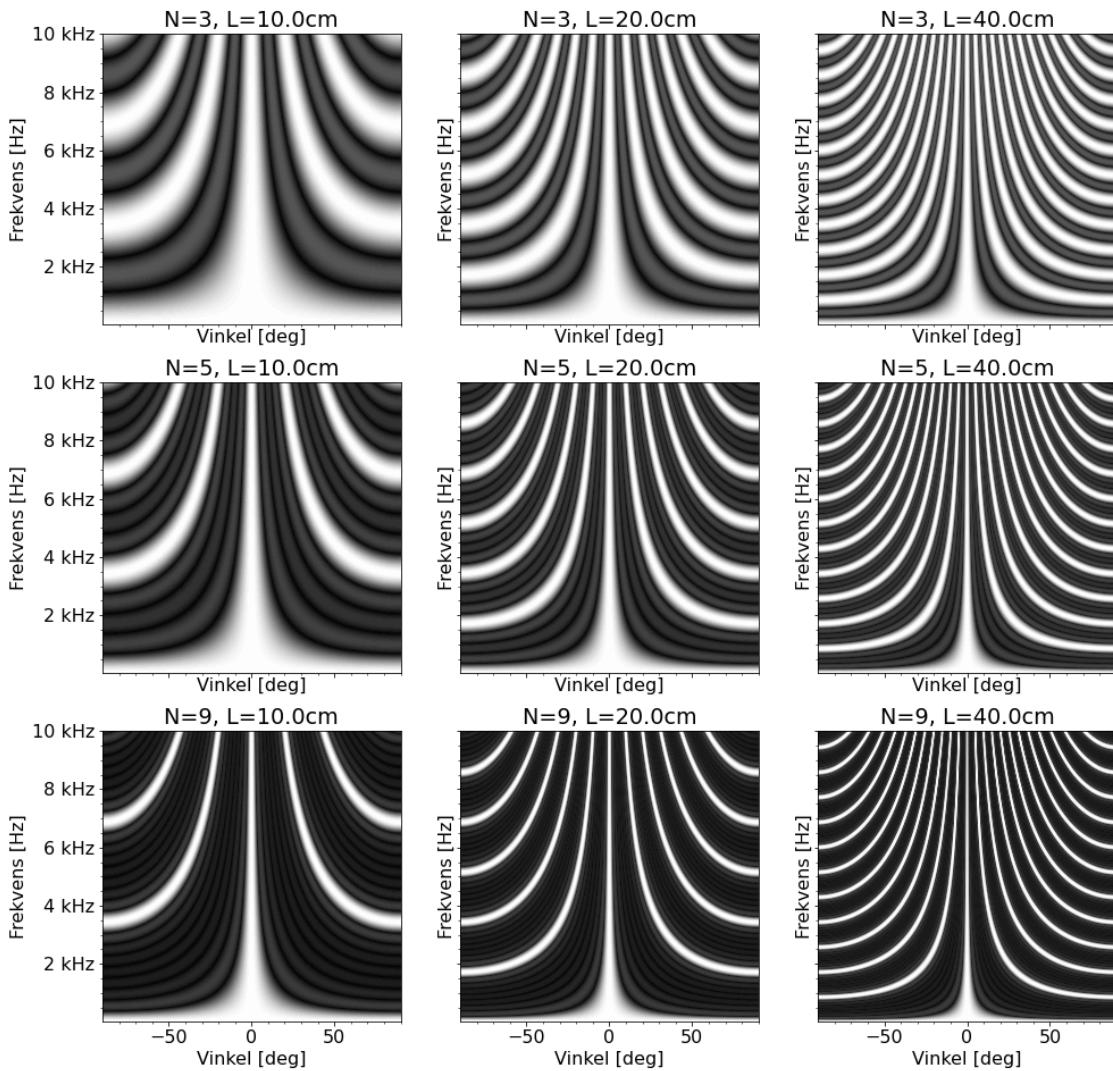


Fig. 7: Ved at variere antallet af mikrofoner og afstanden mellem disse, er det muligt at vise sammenhængen mellem placering og *grating lobes*. I plottet ses 3, 5 og 9 mikrofoner placeret med afstande 10, 20 og 40 cm

Det viser sig, at med større afstand mellem mikrofoner, opstår der flere grating lobes til samme frekvens.

3.3 Steering

For nu er main lobe sat til at være vinkelret på mikrofon arrayet. Dette betyder at, lydkilder kan stå indimellem side og græde lobes, hvilket vil skabe interferens og forvrængning af signalet man ønsker at lytte på. Elektronisk kan man dog korrigere for denne fejl ved at *flytte* main lobe mod lydgiveren, uden at flytte sit array.

Den simpleste metode til at flytte main loben er ved at indsætte forsinkelser til hver mikrofon, så den indgående bølgefront "rammer" hver mikrofon på samme tid¹, inden de bliver summeret. At flytte main lobe en bestemt retning kaldes "steering". Man kan meget nøjagtigt beregne, hvor hvor meget man skal styre mod en lydgiver ved brug af Pythagoras' læresætninger[1, Steering][2]. Selve signalbehandlingen vil blive behandlet i et senere afsnit.

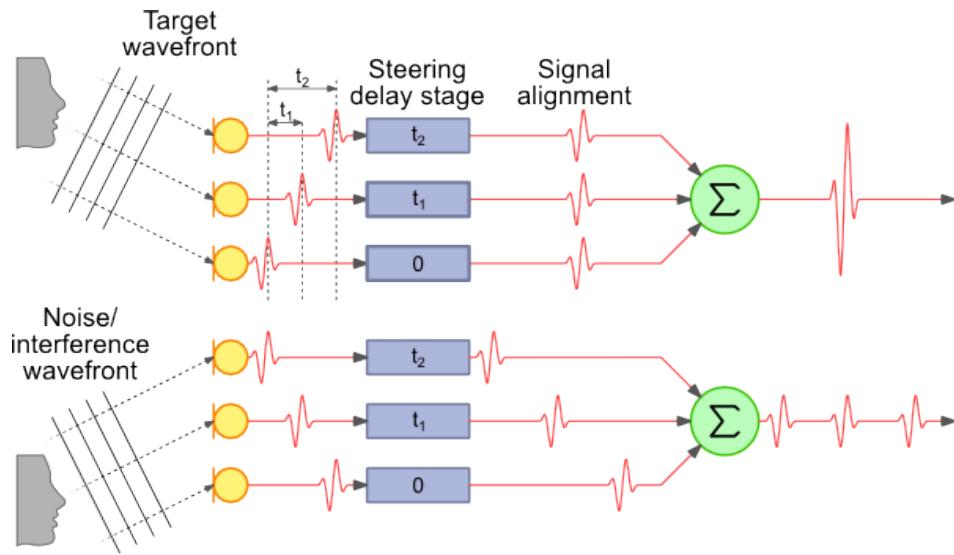


Fig. 8: Ved at lave digitale delaylines ses det muligt at styre retningen der optages fra. Figur fra Labook-Pages[1]

¹Det antages at bølgefronten er plan.

3.4 Forsinkelse af signaler

Forsinkelserne skal implementeres som samples, der forsinkes. De fleste formler baserer sig på trigonometriske funktioner ud fra figur 9 [1, Fractional Delays].

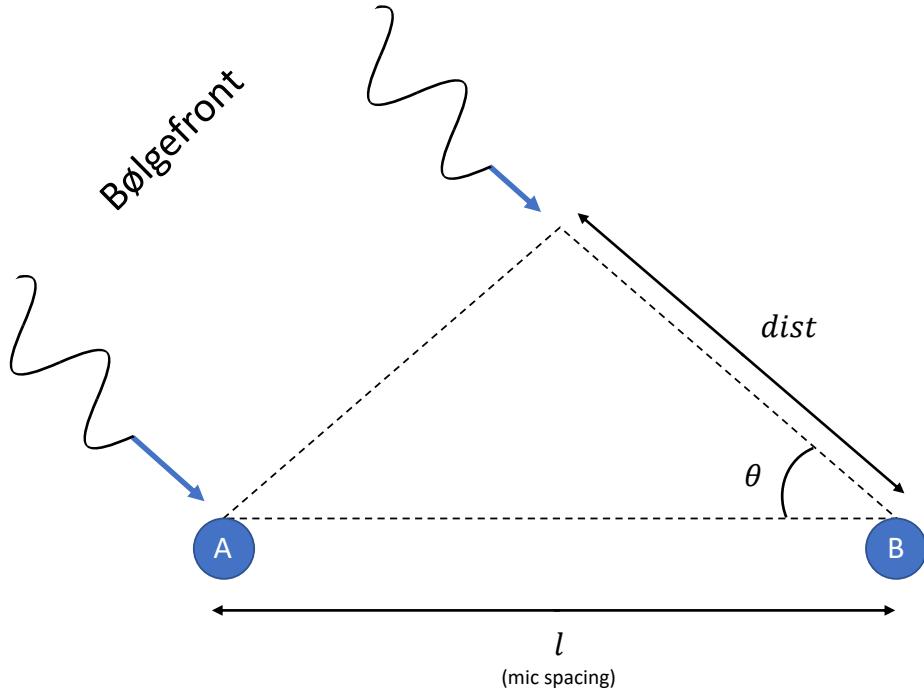


Fig. 9: Forsinkelserne implementeres som samples og baserer sig på disse trigonometriske forhold

Distance mellem bølgefronten og en mikrofon kaldes *dist*. Distancen måles i meter:

$$dist = l \cdot \cos(\theta^\circ) \quad [\text{m}]$$

Bølgefronten skal altså rejse yderligere *dist* længere for at nå mikrofon B, efter den har ramt mikrofon A. Hvis lydens hastighed i luft er $343.3 \frac{\text{m}}{\text{s}}$ må forsinkelsen i sekunder være givet ved:

$$\tau_t = \frac{l \cdot \cos(\theta^\circ)}{c} = \frac{dist}{c} \quad [\text{s}]$$

Det ønskes at have forsinkelsen i samples frem for tid. Der samples med 48 kHz

$$f_s = 48 \text{ kHz}$$

Dette giver en samplingstid på

$$t_s = \frac{1}{f_s} = \frac{1}{48 \text{ kHz}} = 20.8333 \mu\text{s}$$

Forsinkelsen i tid deles med samplingstiden, for at finde antallet af samples

$$\tau_s = \frac{\tau_t}{t_s} = \frac{f_s \cdot l \cdot \cos(\theta^\circ)}{c} \quad [\text{samples}]$$

Heltalsdelen af τ_s er nem at implementere ved et cirkulært array, hvor et antal bagerste samples i signalet, bliver flytte om foran. På den måde *rulles* signalet. Den resterende decimal-del i samples

flyttes med et FIR filter. Heltals delayet ville også kunne justeres ved hjælp af et FIR filter — dette vil dog ikke blive berørt i dette projekt.
 Heltalsværdien og decimalværdien adskilles ved

$$\tau_{int} = \lfloor \tau_s \rfloor$$

$$\tau_{frac} = \tau_s \% 1$$

3.4.1 FIR Filter

Decimalforskydningen udføres ved hjælp af et FIR filter, hvor der bruges en forskudt sinc funktion.

$$x_m = t - \tau_{frac_m} \quad (t = 0, 1, 2 \dots N-1)$$

$$center = \left\lfloor \frac{N}{2} \right\rfloor$$

$$\text{sinc}_m = \frac{\sin(\pi \cdot (x_m - center))}{\pi \cdot (x_m - center)}$$

$$h_m(t) = \text{sinc}_m \cdot w(t)$$
(3)

hvor $w(t)$ er en vinduesfunktion. I den efterfølgende implementering er et blackman vindue brugt, forsøg med Hamming og Hanning vinduer viste sig også effektive.
 N for antallet af koefficenter/FIR tappe. Subscriptet m indikere hvilken mikrofon der behandles, et højere antal tappe giver større præcision og indsnævre det notch-filter der ses i figur 11 og figur 10.

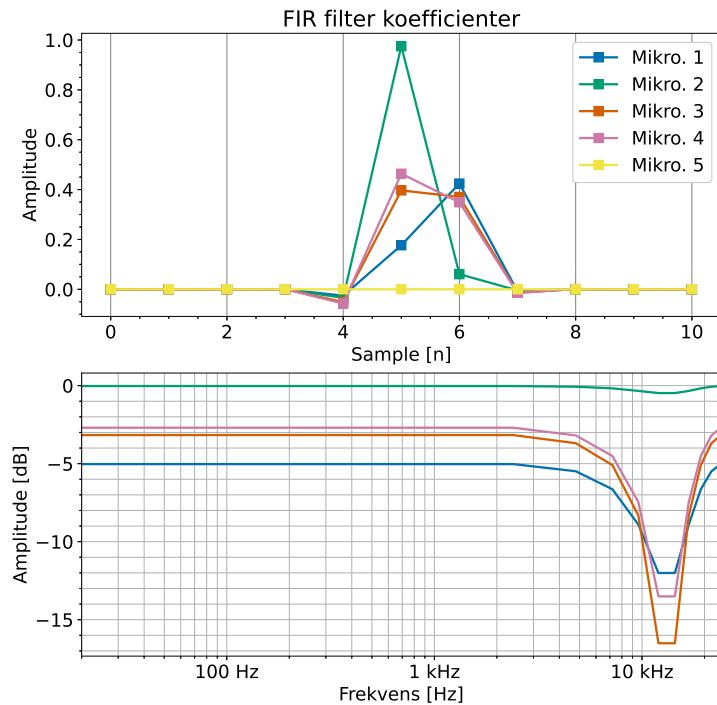


Fig. 10: 5 forskellige 11 tap FIR-filttere brugt til forsinkelse af signaler. Det bemærkes at der opstår et 11kHz notchfilter, med en dæmpning fra -18 dB til -1 dB

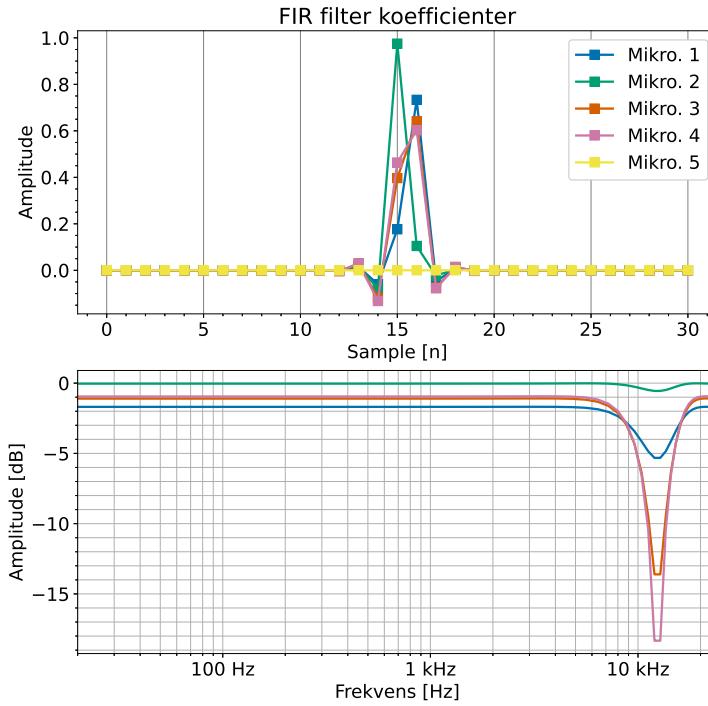


Fig. 11: 5 forskellige 31 tap FIR-filtere brugt til forsinkelse af signaler. 11kHz notchfilter bemærkes igen, dog med en langt mere snæver karakteristik.

Ved denne foldning opstår et ekstra delay der beskrives $\frac{N-1}{2}$.
Dette delay trækkes fra τ_{int} for at undgå et yderligere delay i den proces.

3.5 Korrigert DAS-output.

Et korrigert mikrofonoutput kan derfor beskrives som

$$y_m(t) = x_m \left[t - \left(\tau_{int} - \frac{N-1}{2} \right) \right] * h_m(t), \quad (4)$$

$$\begin{aligned} z(t) &= \sum_{m=1}^M y_m(t) \\ z(t) &= \frac{1}{M} \sum_{m=1}^M x_m \left[t - \left(\tau_{int} - \frac{N-1}{2} \right) \right] * h_m(t) \end{aligned} \quad (5)$$

hvor M er antallet af mikrofoner og y_m er et korrigert mikrofonoutput. Der medregnes *ikke* energifald over den rejste distance, og der laves en antagelse om at bølgefronterne er plane idet at de rammer mikrofon-arrayet[4].

Udfra figur 7 samt, figur 3, vælges der at udføre en højpasfiltrering, da mulighederne for retningsbestemmelse under 400 Hz anses som værende umuligt, udfra de afstandsmuligheder der er tiltænkt.

Der vælges et 4 ordens butterworth-filter, da denne type ikke påvirker signalet i en voldssom grad, men dog dæmper mængden af lavfrekvent materiale — andre typer filtre som Chebyshev ville muligvis kunne bruges, men er ikke afprøvet i forbindelse med vores forsøg.

4 Test af projektet

Der ønskes at lave en række testcases, hvor der er opsat forskellige lyttepositioner. Ved at afspille 2 lydkilder simultant i et kendt rum vil der kunne testes om det er muligt at styre retningen for mikrofonarrayet, hvorved en enkelt kilde kan isoleres - uden at ændre på den mekaniske funktion/position på mikrofoner. Testen implementeres i Python og forventes ikke at kunne implementeres i et realtime format.

Forsøget er udført i AudioLab på Aarhus Universitet med følgende udstyr

- Mikrofontype : Behringer ECM8000
- Lydkilder : Genelec 8020
- Lydkort : Behringer FCA1616
- Preamps : MOTU 8Pre
- DAW : Ableton Live 10

5 mikrofoner blev fastmonteret på en træplade med 20 cm afstand. Dette gav mulighed for følgende testvarianter vist i 3. Opstillingen ses i figur 12



Fig. 12: 5 · Behringer ECM8000 mikrofoner monteret med 20 cm afstand.

Mikrofonerne blev tilkoblet MOTU 8Pre og blev via ADAT overført til FCA1616 lydkort. Afspilning blev startet i Ableton Live og der blev optaget en række sekvenser som vist i figur 13. Det optagne materiale blev eksporteret fra Ableton Live til et format (WAV) der kunne bearbejdes i Python.

Tabel 3: De 3 varianter der er mulige at teste med projektets opstilling. 2a og 2b er hhv. med hhv. højre og venstre ydremikrofon som udgangspunkt

Variant	Mikrofoner	Afstand [cm]
1	5	20
2a	3	20
2b	3	20
3	3	40

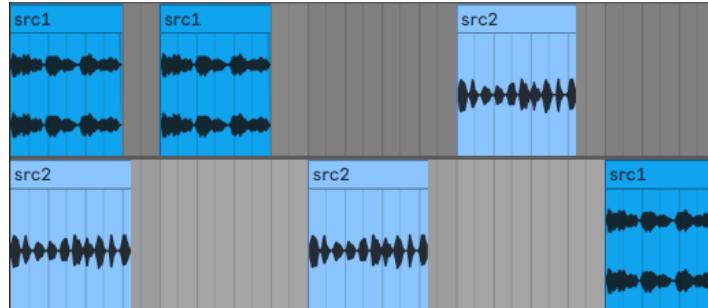


Fig. 13: Ableton afspilning. Begge kilder afspilles simultant, hvorefter der laves en enkelt afspilning af hver kilde. Til sidst ombyttes de to kilder. Det øverste spot af spille af kilde #1 og det nederste spor afspilles af kilde #2

Mikrofoner og højtalere blev placeret, og afstanden målt og noteret. Denne teknik gik igen ved alle forsøgsopstillinger. I de følgende afsnit vises resultater for forsøget foretaget i det lyddøde rum og lyttestue.

De to talesignaler, der bruges i forsøget ses på figur 14 og figur 15. Opsætningen af forsøget i lyttestuen er vist på figur 16.

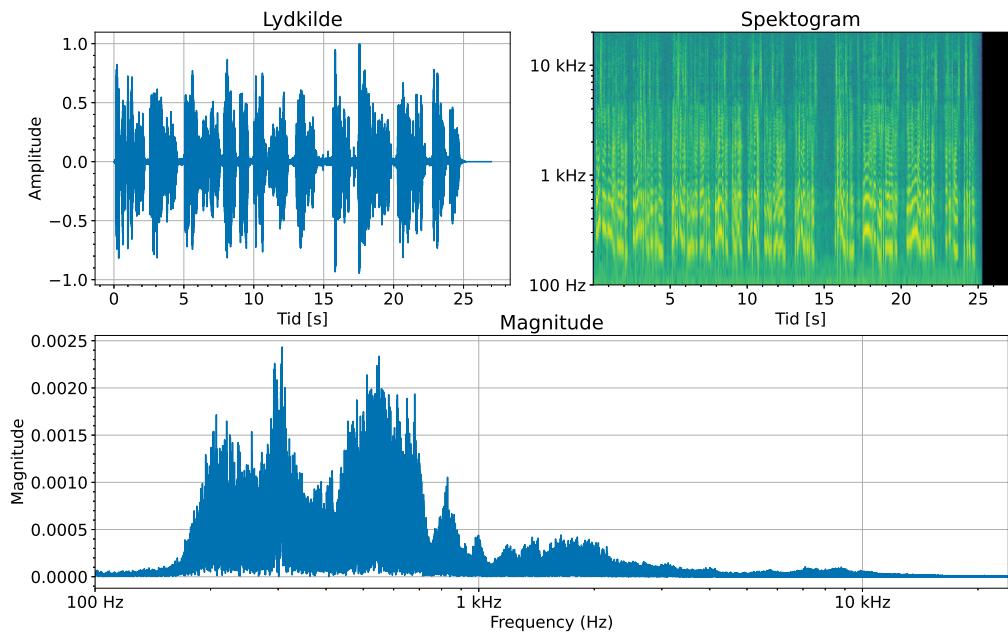


Fig. 14: Kilde #1: En kvindestemme oplæser en historie.[5]

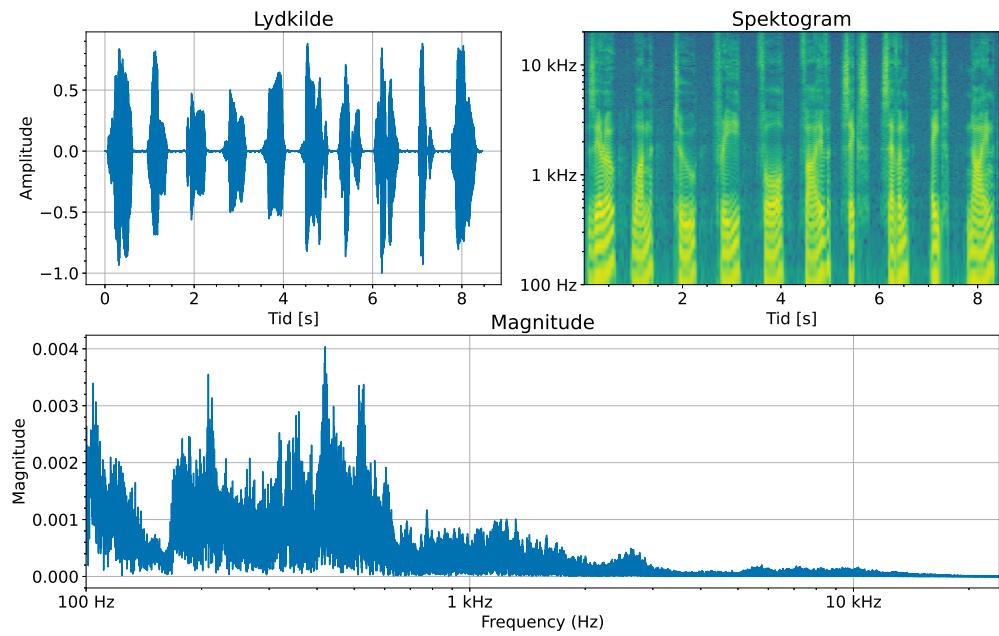


Fig. 15: Kilde #2: En mandestemme tæller fra 0 til 9.[5]



Fig. 16: Opsætning af test i lyttestue

5 Resultater

Lydfilerne der er genereret i denne test, kan findes i den vedlagte zipfil, eller via GitHub repositoriet ETLYAK-DAS, hvor der desuden findes en Jupyter Notebook der udfører signalbehandlingen, samt et Python script med selve signalbehandlingsmetoderne.

Dette afsnit viser resultaterne for det lyddøderum hvor der blev afspillet 2 stemme stimuli, som beskrevet tidligere i rapporten. Der findes resultater for lytterummet i bilaget, 8.

5.1 Test i lyddødt rum

Tabel 4: Forsinkelsesresultater for 2 kilder i et lyddødt rum(3). Mikrofonopsætningen bestod af 5 mikrofoner med 20 cm afstand. Kildeplacering ses i figur 17.

	Kilde 1 (-1.1,2.95) [m]			Kilde 2 (0.95,2.5) [m]		
	θ_m	τ_{int_m}	τ_{fract_m}	θ_m	τ_{int_m}	τ_{fract_m}
m_1	76.65	34	0.9055	61.63	0	0
m_2	76.65	26	0.1781	61.63	10	0.0485
m_3	76.65	17	0.4528	61.63	20	0.0971
m_4	76.65	8	0.7264	61.63	30	0.1457
m_5	76.65	0	0	61.63	40	0.1942

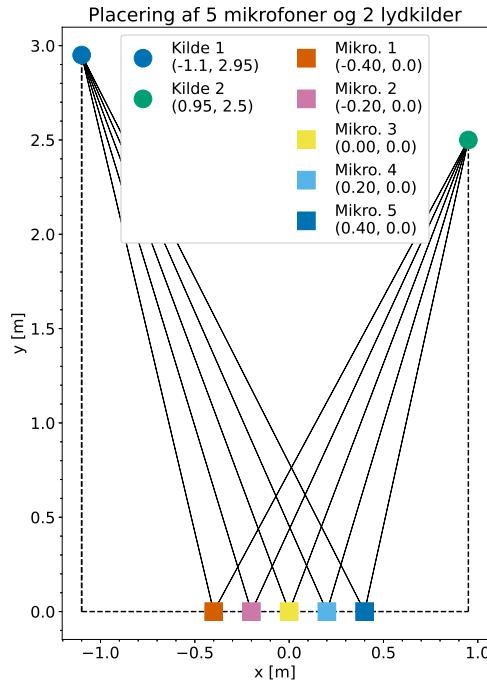


Fig. 17: Illustration der viser opstillingen, udført i AudioLabs lyddøde-rum.

5.1.1 5 mics, 20 cm

I denne test bruges 5 mikrofoner med 20 cm afstand mellem mikrofonerne. På figur 18 vises det ikke-korrigerede signal af en kvindestemme i højre højtalere og en mandestemme i venstre højtalere. Begge signaler er summeret sammen da man kan lave superposition af kilder i et lyddødt rum.

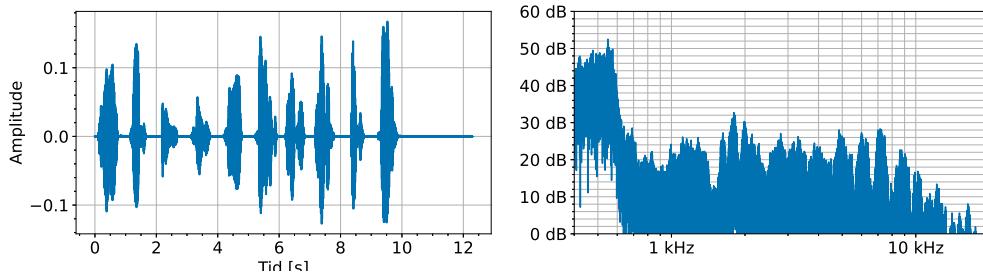


Fig. 18: Ikke-korrigeret talesignaler

Det er tydeligt at se, at størstedelen af energien i signalet ligger mellem 400 Hz og 700 Hz. Ved sådan et signal forventes kamfiltrering. Ift. inputsignalene (figur 14 og figur 15), så er der mistet en del højfrekvent materiale. Dette kan skyldes kvalitet af optagelse eller kamfiltrering. Auditivt er optagelsen hul, som om der tales ned i en dåse. Der opstår interferens mellem mikrofonerne. Det lyder ret "dårligt" og de to signaler forstyrre hinanden en del. De to signaler adskiller sig dog i forvejen fra hinanden, ved at ligge i forskellige frekvensområder. Han er mere lavfrekvent end hende.

Begge signaler er hørbare med ca. samme lydstyrke. Den lavfrekvente del fra 0-400 Hz er dæmpet meget pga. butterworth-filteret.

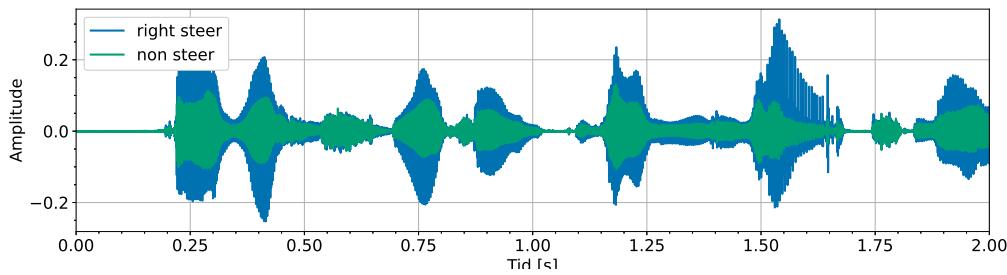


Fig. 19: Højrestyrede talesignal - fokus på mandestemmen

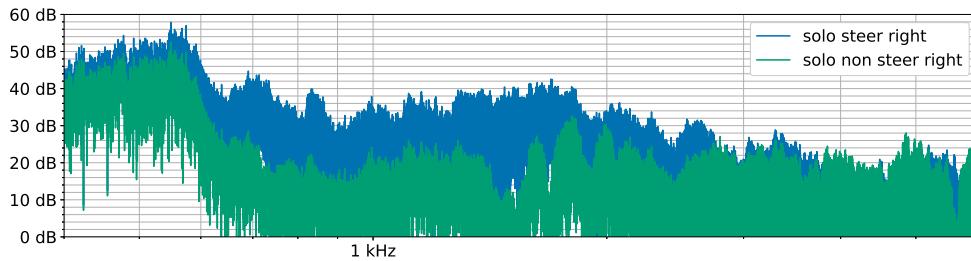


Fig. 20: Styring mod højre kilde

Styring mod højre Der styres mod højre højtalere. Resultatet vises i figur 19. figur 15 viser det rå signal af mandestemmen. Tager man det tidslige signal i betragtning, kan man se, at mandestemmen er mindre betydningsfuld, efter der er højrestyret. Peak-værdien på FFT'en omkring 600 Hz går fra ca. 50 dB (ikke-korrigeret) til ca. 56 dB (korrigeret), hvilket svarer til en fordobling. I tidssignalet svinger peak-værdierne $\pm 0,1$ ved det ikke-korrigerede signal. Efter højrestyringen er det ca. $\pm 0,2$. Det lavfrekvente materiale er dæmpet mere på det højrestyrede signal i relation til det ikke-korrigerede signal.

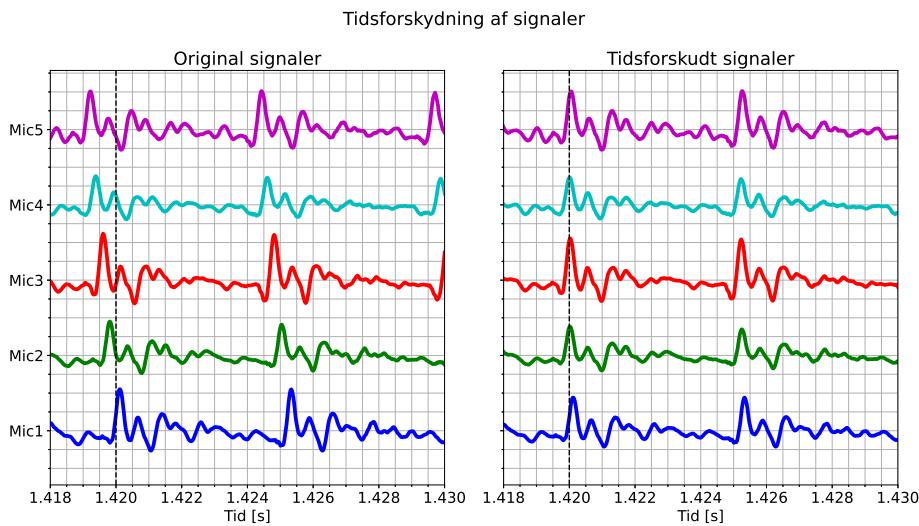


Fig. 21: Resultatet af de shiftede signaler. Her ses Mikrofon 1 som fixpunkt for forsinkelsen.

Der opnåes en forstærkning på 5.6 dB ved styring mod kvindestemen iforhold til det ikke styrede signal, hvilket ca. svarer til en fordobling.

Auditivt er der mistet en lille smule af den højfrekvente del af hende, hvilket får hende til at lyde nasal, grundet en overrepræsenteret mellemtone. Lyden af hendes stemme minder om en telefonefekt hvor man har fjernet noget af diskanten og bunden og forstærket mellemtonen. Til gengæld er hun klarere og tydeligere, da han er blevet dæmpet en del. Specielt lyder hans signal mere "tyndt". Den lavfrekvente del af ham er dæmpet. Den højrestyrede højtaler, hvor hun er i fokus, lyder næsten dobbelt så høj som ham - hvilket formegentligt relaterer sig til den dobbelte lydstyrke fra det ikke-korrigerede signal og det korrigerede signal, og de 5.6 dB. Det er en usædfuld korregering, da forskellen på dette og det ikke-korrigerede signal er så stor.

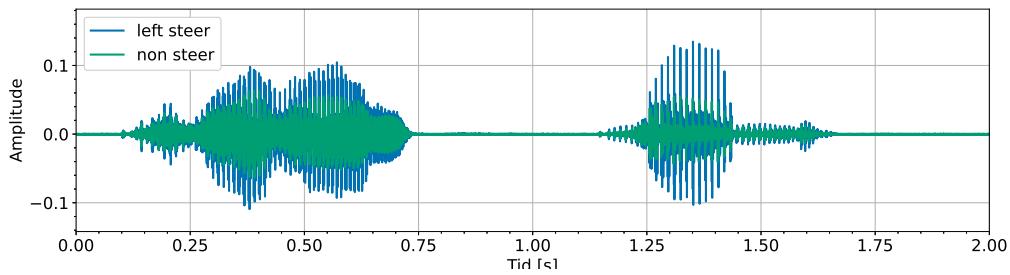


Fig. 22: Vestrestyret talesignal - fokus på herrestemmen

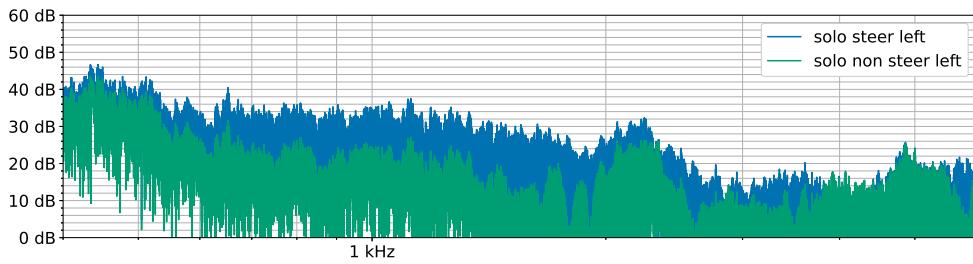


Fig. 23: Styring mod venstre kilde

Styring mod venstre På figur 22 styres der mod mandestemmen i venstre højttaler. En visuel analyse af tidssignalet siger os, at herrestemmen er mere fremtræden end det ikke-korrigerede signal. Mandestemmen tæller fra 0 til 9 og man kan tyde disse stakkerede signaler med ca. 1 sekund mellemrum. På frekvensresponset ses mere ”lavfrekvent”aktivitet (400 Hz - 500 Hz) end på det højrestyrede og ikke-korrigerede signal, hvilket stemmer overens med indholdet af kildematerialet.

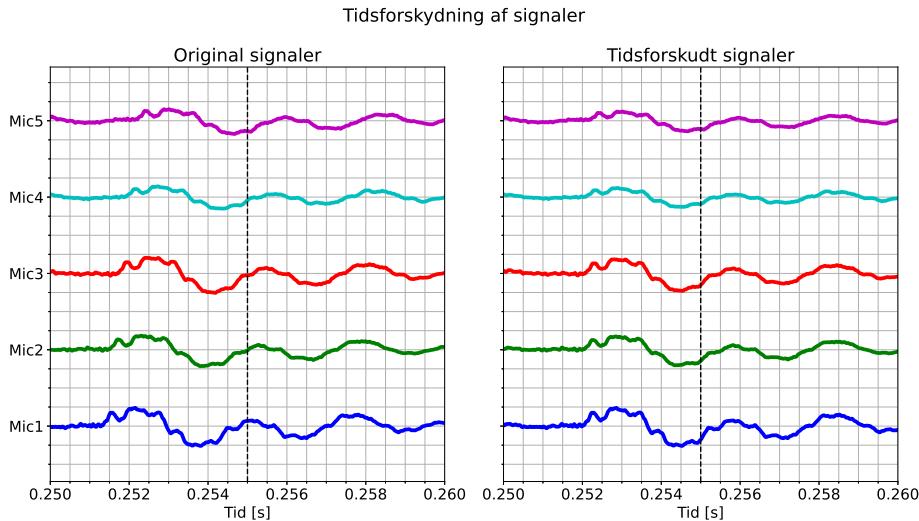


Fig. 24: Resultatet af de shiftedede signaler, hvor Mikrofon 5 er brugt som fixpunkt

Der opnåes en forstærkning på 4.4 dB ved styring mod mandestemmen.

Auditivt er der tydeligt en forskel på den indbyrdes dynamik mellem signalerne. Hendes stemme er blevet meget "tynd" hvor hun ikke forstyrre den venstre højtaler meget. Man kan tydeligt høre hvordan kamfiltreringen ødelægger hendes signal. Hun lyder meget ukorreleret med sig selv - typisk lyd af faseproblemer. Han lyder ikke lige så nasal som hende, på det højrestyrede signal, grundet hans signal i forvejen ikke er lige så mellemtone-agtigt. Til gengæld kan man stadig fornemme telefonefekteten hvor bund og top er fjernet. Efter signalet er korrigteret er han mere tydelig og ligger højere i lydbilledet end hende. Dette passer også med, hvad der ses på tidssignalet.

5.2 Test i lytterummet

Ved gennemlytning af optagelserne fra lytterummet, var begge stemmer tydelige i optagelserne, dog var der en hørbar forskel på de forskellige styringer og de virkede efter hensigten.

Plots for denne test ses i bilag 8.

6 Diskussion

Der var udført en række tests med sinustoner, hvor vi afspillede to forskellige toner (1 kHz og 2 kHz), hvor vi ønskede at eftervise, at vi kunne styre lydstyrken på et signal ved at "zoom" ind på én tone ad gangen. Testen slog desværre fejl, da der ikke var taget højde for eventuelle grating lobes og andre usikkerheder (harmoni mellem toner, vinkel/beam pattern efter korrigering). I stedet for at bruge rentoner, kunne man efterprøve teorien ved at afspille hvid støj i en højtalere og et sweep/chirp signal i den anden højtalere. Hvidstøjssignalet kunne forhåbenligt også vise noget om kamfiltreringen ved forskellige konfigurationer af mic arrays og evt. beam pattern. Ved at bruge et sweep vil der signalet ramme dødzoner og grating lobes. Man vil altså kunne vide sig sikker på, at nogle toner kan høres, og nogle kan ikke. Dette skulle også gerne passe med grafen vist på figur 7.

Talesignalerne ligger i to forskellige frekvensområder og der vil naturligvis være en auditiv adskilelse af signalerne på forhånd. Det ville være spændende at bruge den samme person til at indspille forskellige lydfiler, sådan så man fjerner en måleusikkerhed, der hedder frekvensadskillelse og kan fokusere på tydeligheden/dynamikken af to signaler efter korrigering.

Indtil nu er der ikke taget hensyn til placering af højtalere ift. grating lobes og/eller afstande til mic array. Der er også stadig en vis usikkerhed på, hvorvidt mic arrayet's main lobe (0°) "peger" ud fra center-mikrofonen eller andet steds. Der var et ønske om at eftervise en figur i artiklen fra Grythe ([2, III. MINIMUM VARIANCE BEAMFORMING - steered response]) hvor man ser en beam pattern efter korrigering. Dette kunne sige noget om de indfaldende kilder, når man har styret sin main lobe til at pege over mod en lydkilde. Dette hører sig til fremtidigt arbejde, som bestemt er værd at undersøge nærmere.

Efter at have kigget på beam patterns over et frekvensspektrum (figur 7) synes vi det ser spændende ud, når man går ned i afstand mellem mikrofoner, og op i antallet af mikrofoner. Bredden på main lobe er større, når man har placeret mikrofonerne med kort indbyrdes afstand. Til gengæld er frekvensen man kan beamform fra højere. Dog er der færre grating lobes og generel mindre dæmpning på signalerne udenfor main lobe. I og med, at main lobe er bredere, er måleusikkerheder på højtaleralægning måske mindre betydnende og forhåbentlig kan det give mere tolerance til det højfrekvente område.

7 Konklusion

Det har været muligt at eftervise teorien fra på Labbook Pages. Vi har kunnet isolere en lydkilde med et mikrofon array. Det er ikke uden bekostning af hovedsignalet man ønsker at høre på. Grundet grating lobes og interferrens til forskellige frekvenser, måleusikkerheder osv. bliver hovedsignalet i sit main lobe påvirket. I vores tilfælde oplever vi en "telefon-effekt" på hovedsignalet. Til gengæld stiger lydstyrken i sit main lobe, og dæmper omringliggende lydkilder. Det er lykkedes at forstærker kvindestemmen 5.6 dB og mandestemmen 4.4 dB, hvilket for kvindestemmen er svarende til en fordobling. De korrelerede signaler er meget mere klare end de ikke-korrigerede signaler. Med 5 mikrofoner og en træplade har det altså været muligt at styre et signal hen til en lydkilde i efterprocessing.

Litteratur

- [1] D. A. Greensted, "Labbook Pages - Microphone Array Beamforming," 2010. [Online]. Available: <http://www.labbookpages.co.uk/audio/beamforming.html>
- [2] J. Grythe, "Beamforming algorithms - beamformers," *Norsonic Tech. NOTE*, pp. 1–5, 2015. [Online]. Available: <https://web2.norsonic.com/wp-content/uploads/2016/10/TN-beamformers.pdf>
- [3] X. Anguera, "Multichannel Acoustic Beamforming for Meetings," 2008. [Online]. Available: <http://www.xavieranguera.com/phdthesis/node85.html>
- [4] T. Roelandts, "How to Create a Fractional-Delay Filter," 2019. [Online]. Available: <https://tomroelandts.com/articles/how-to-create-a-fractional-delay-filter>
- [5] D. Murphy, J. Rees-Jones, and Audiolab - Department of Electronics - University of York, "Openair." [Online]. Available: https://www.openair.hosted.york.ac.uk/?page_id=310

8 Bilag

Resultater fra Lytterum

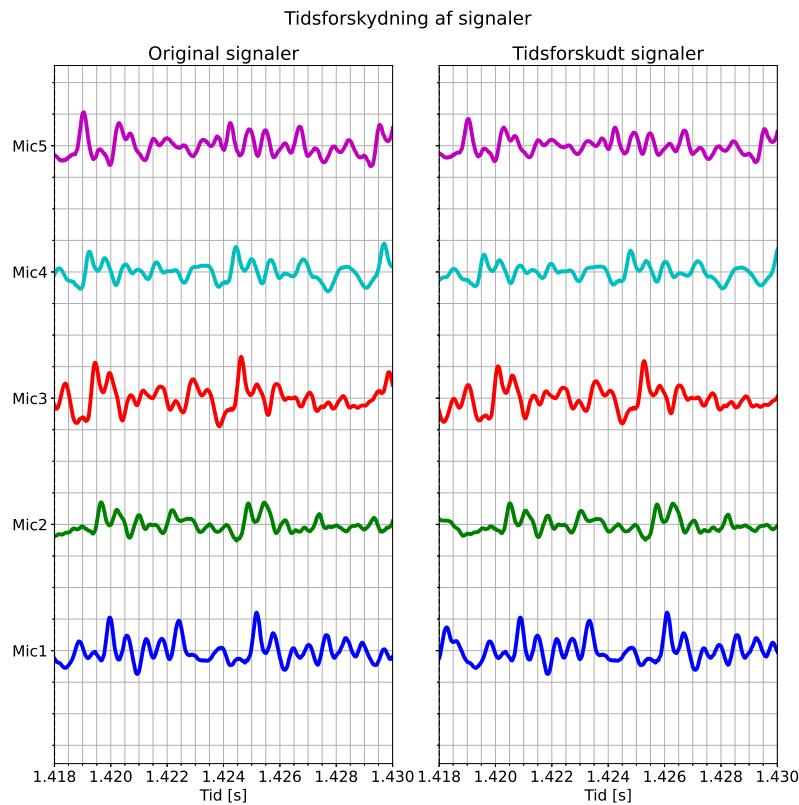


Fig. 25: Venstrekorrigert lyttestuesignal

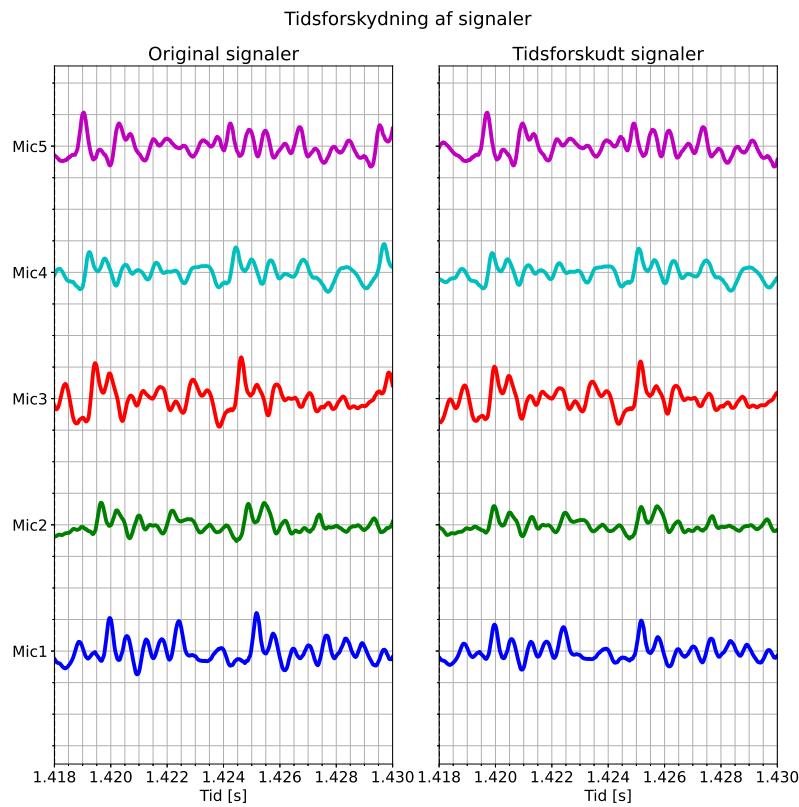


Fig. 26: Højrekorrigeret lyttestuesignal

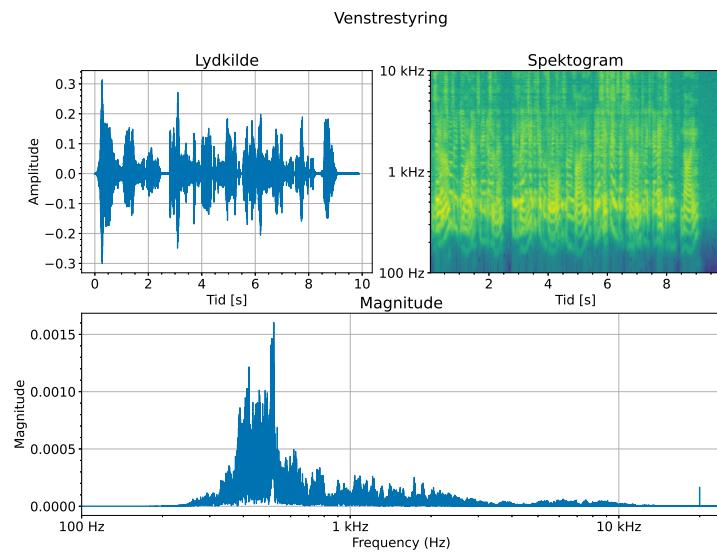


Fig. 27: Venstrekorrigert lyttestuesignal

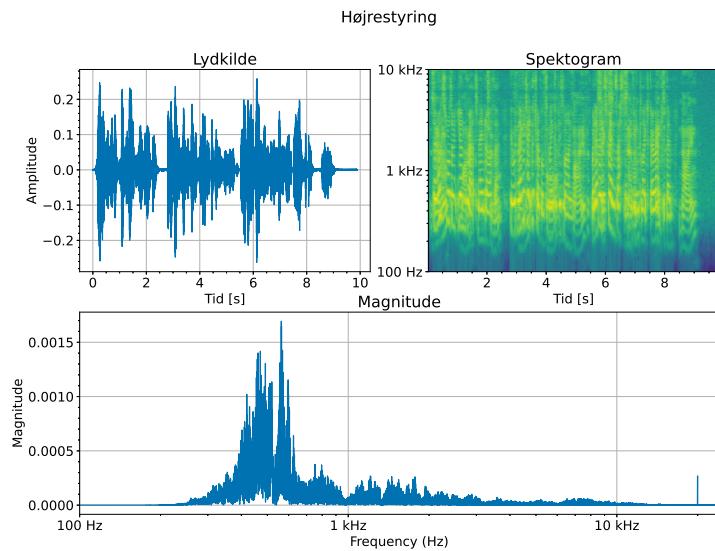


Fig. 28: Højrekorrigeret lyttestuesignal

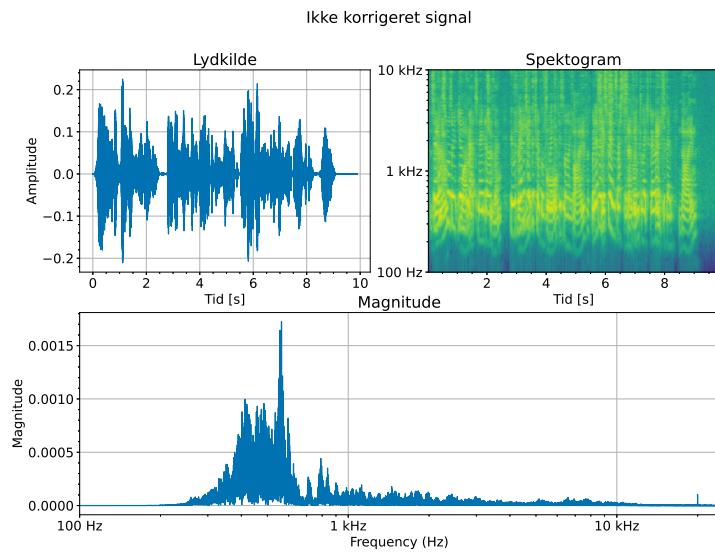


Fig. 29: Ukorrigert lyttestuesignal

Resultater for 3 mikrofoner med 40 cm afstand i det lyddøderum

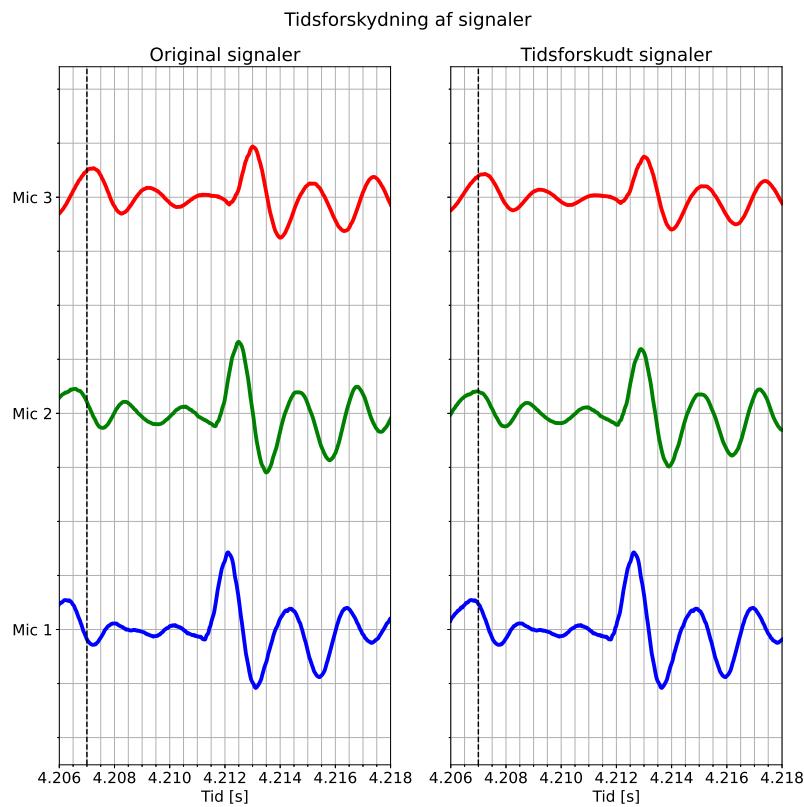


Fig. 30: Signalkorrigert mod venstre side

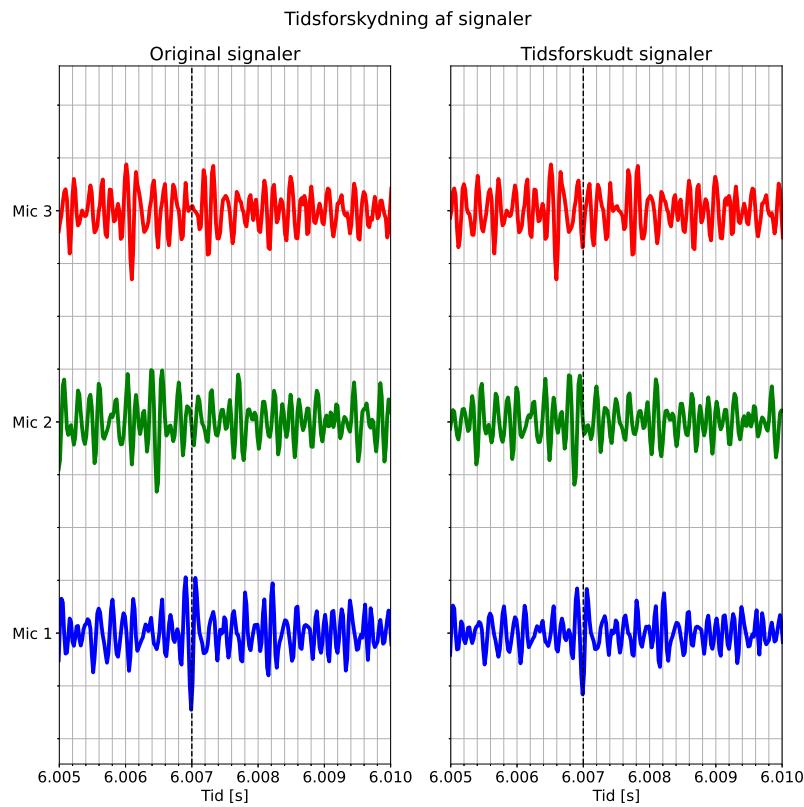


Fig. 31: Signalkorrigert mod højresiden

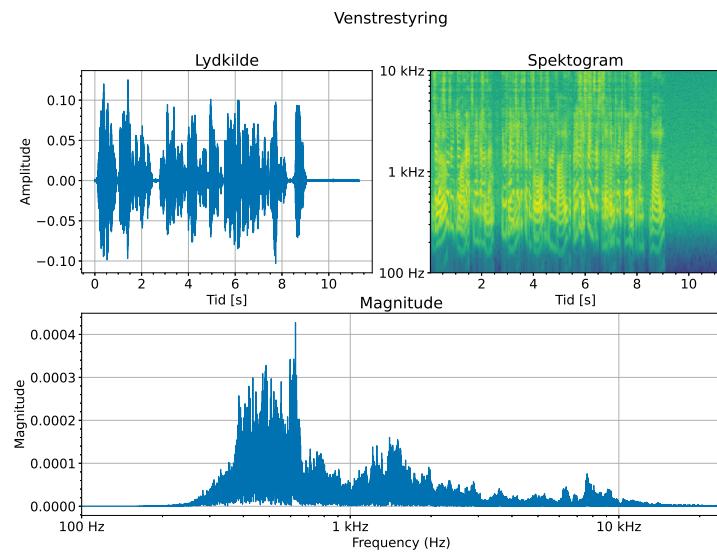


Fig. 32: Signalkorrigert mod venstre side

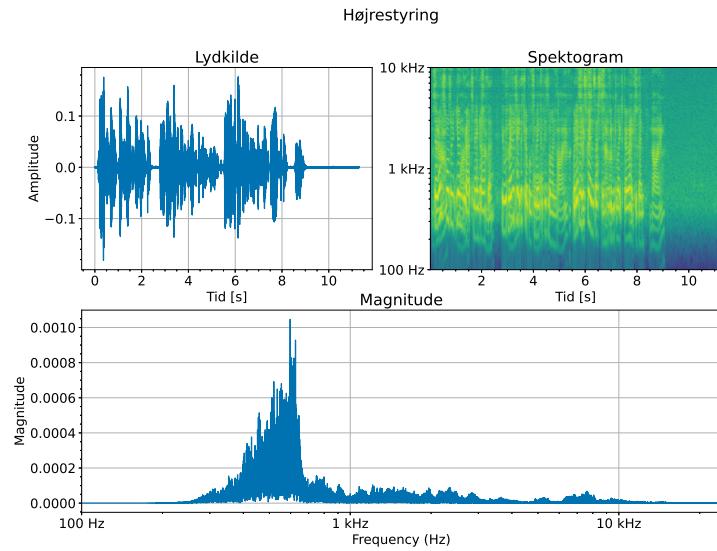


Fig. 33: Signalkorrigert mod højre side

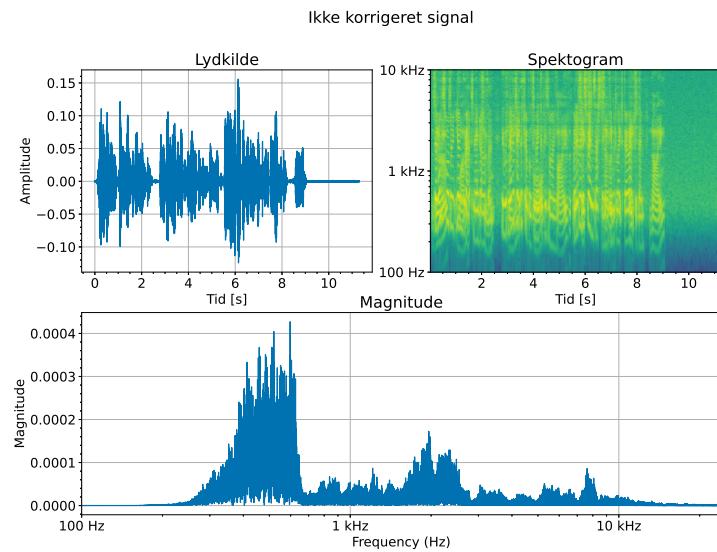


Fig. 34: Ukorrigert signal