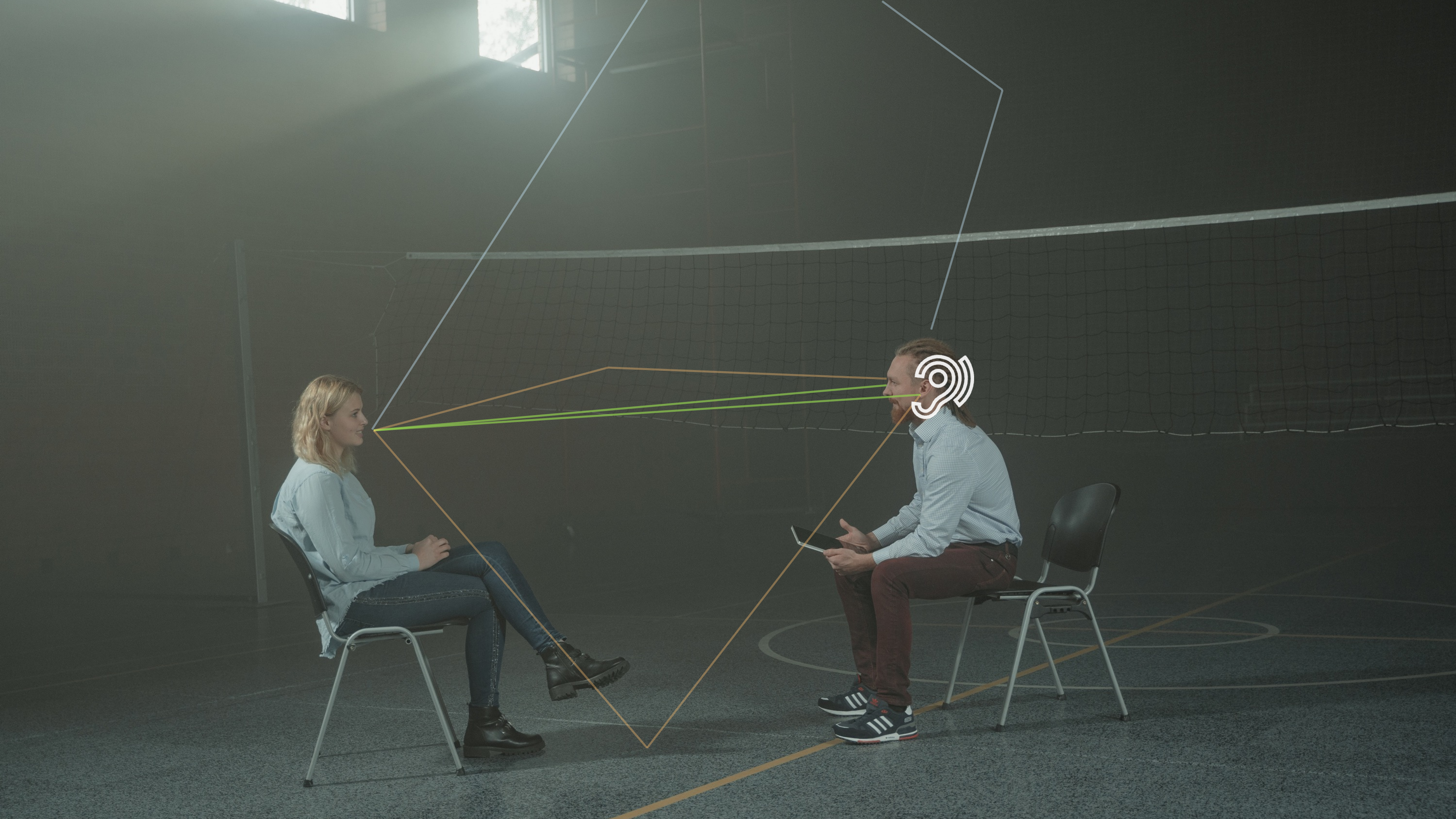
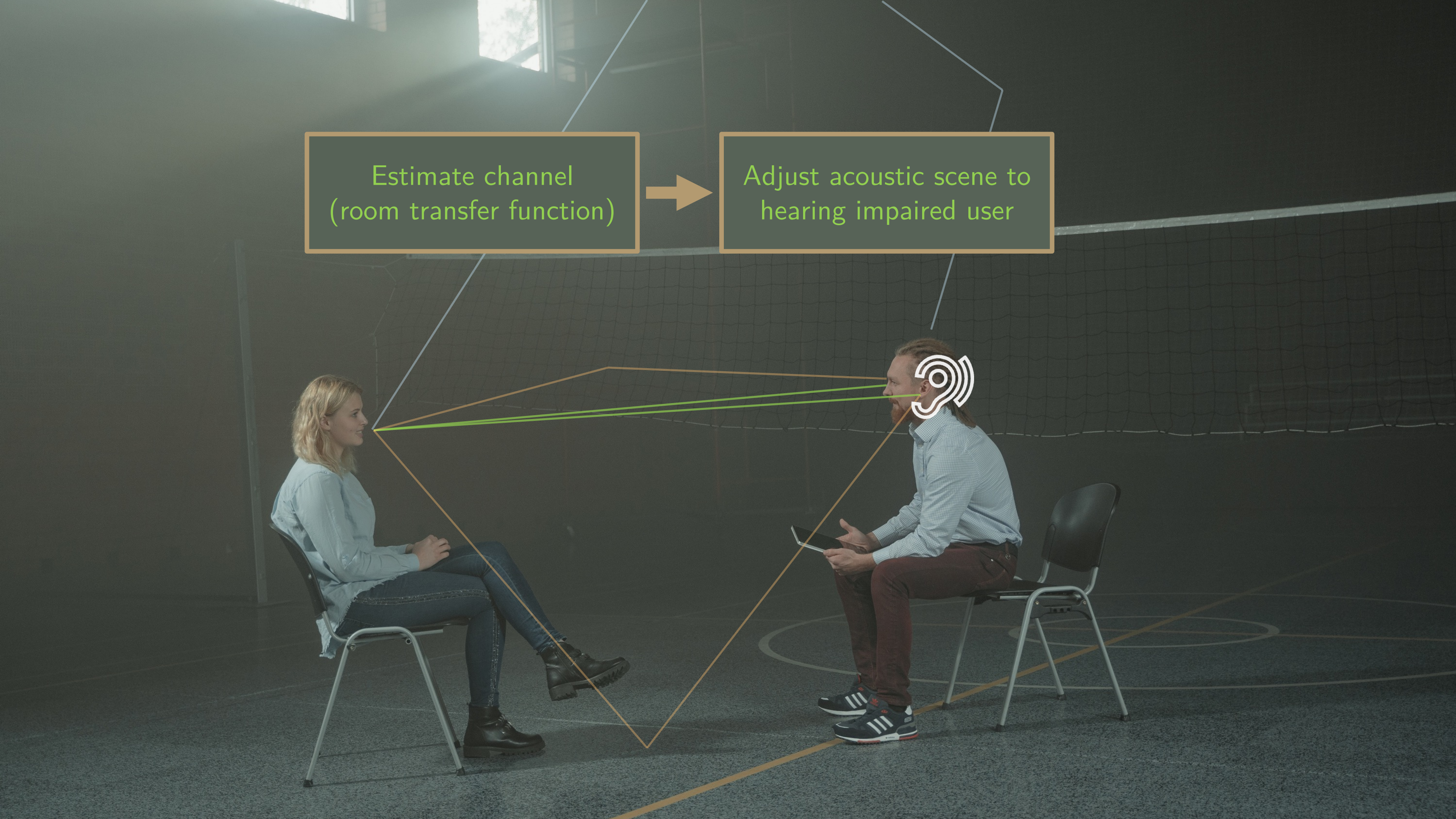
Who am I?

NORTH SEA
British Isles
ATLANTIC OCEAN
IRELAND
DUBLIN
Belfast
Cork
UNITED KINGDOM
SCOTLAND
Wick
Aberdeen
Perth
Edinburg
Glasgow
Newcastle
Liverpool
Manchester
WALES
ENGLAND
Birmingham
Cardiff
LONDON
Plymouth
Southampton
Stavanger
Kristiansand
Göteborg
Kalmar
Gotland
Oland
DENMARK
COPENHAGEN
Malmö
BALTIC SEA
TALLIN
ESTONIA
Saint Petersburg
Novgorod
Iaroslavl
RIGA
LATVIA
Tver
LITHUANIA
VILNIUS
MOSCOW
Riazan
Kaluga
Tula
MINSK
BELARUS
Oriol
Gomel
Kursk
Belgorod
Kharkiv
KIEV
UKRAINE
Dnieper
Dnipropetrovsk
Donetsk
Lviv
MOLDOVA
CHISINAU
Odessa
Sea of Azov
Sevastopol
BLACK SEA
Bosporus
NETHERLANDS
ASTERDAM
BRUSSEL
Antwerpen
BELGIUM
LUX.
Hamburg
BERLIN
Elba
Leipzig
Düsseldorf
GERMANY
Frankfurt
Stuttgart
Danube
Munich
POLAND
Oder
Vistula
WARSAW
Kraków
PRAGUE
CZECH REP.
SLOVAKIA
BRATISLAVA
Carpathians Mts.
VIENNA
AUSTRIA
BUDAPEST
HUNGARY
ROMANIA
BUCHAREST
Danube
Le Havre
PARIS
Brest
Rennes
Nantes
FRANCE
Loire
Lyon
Bordeaux
Rhone
Toulouse
Montpellier
Pyrenees
Marseille
BERNE
LIECHT.
SWITZ.
Alps
SLOVENIA
LJUBJANA
ZAGREB
Milano
Venice
Torino
Po
Florence
Monaco
CROATIA
BOSNIA & H
SARAJEVO
BELGRADE
SERBIA
ADRIATIC SEA
MONTENEGRO
PODGORICA
SKOPJE
FYROM
SOFIA
BULGARIA
ITALY
ROME
Corsica
Naples
Bari
TIRANA
ALBANIA
Thessaloniki
Sardinia
Palermo
Messina
Sicily
Peloponnesus
ATHENS
GREECE
AEGEAN SEA
Crete
VALETTA
MALTA
ANKARA
TURKEY
Izmir
Konya
Kayseri
Adana
Antalya
Aleppo
NICOSIA
CYPRUS
SYRIA
LEBANON
BEIRUT
Vigo
Bilbao
AND.
Porto
Caimbra
Zaragoza
Barcelona
MADRID
Tago
Valencia
Balearic Is.
PORTUGAL
SPAIN
LISBON
Cordoba
Seville
Gibraltar(UK)
Ceuta
Tangier
Melilla
RABAT
Fes
Oran
Chlef
Constantie
ALGIERS
ALGERIA
TUNIS
TUNISIA
MEDITERRANEAN SEA




Estimate channel
(room transfer function)
Adjust acoustic scene to
hearing impaired user

Estimate channel (room transfer function) → Adjust acoustic scene to hearing impaired user

Why are hearing-aids not so good?

1. Signals change fast (non-stationarity)
2. Variety of reverberant environments
3. Real-time constraints
4. Battery constraints

Applications

https://images.squarespace-cdn.com/content/v1/54d696e5e4b05ca7b54cff5c/1625158518155-U2HP9YODRMV1I3T83M4J/Higher+Order+Ambisonics+Microphones.jpg

Cramér-Rao bound for acoustic transfer function estimation

Giovanni Bologni (G.Bologni@tudelft.nl), Richard C. Hendriks, Richard Heusdens

31 May, 2024

On the menu today

Part 1 - **Cramér-Rao bound** for acoustic transfer function estimation

1. Parameter estimation & Cramér-Rao bound (CRB)
2. Case study – ATF estimation

Part 2 - Acoustic transfer function estimation with **inter-frequency correlation**

1. Channel estimation algorithm
2. Experiments

Parameter estimation

A quantity θ needs to be estimated

Given a (random) model, what is best achievable accuracy on estimating unknown quantity?

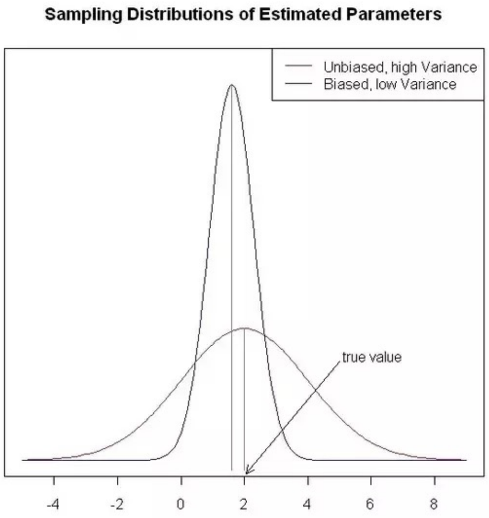
Parameter estimation

A quantity θ needs to be estimated

Given a (random) model, what is best achievable accuracy on estimating unknown quantity?

Best accuracy = minimum MSE = minimum variance (unbiased estimator)

$$\mathrm{MSE}(\theta, \hat{\theta}) = \mathrm{var}(\hat{\theta}) + \mathrm{bias}^2(\hat{\theta}, \theta).$$



Pic: quora.com

Cramér-Rao bound

Under regularity assumptions on probability distribution $p(x;\theta)$,

$$\mathrm{var}(\hat{\theta}) \geq I^{-1}(\theta) = \frac{1}{-\mathbf{E}\left[\frac{\partial^2 \ln p(x;\theta)}{\partial\theta^2}\right]}.$$

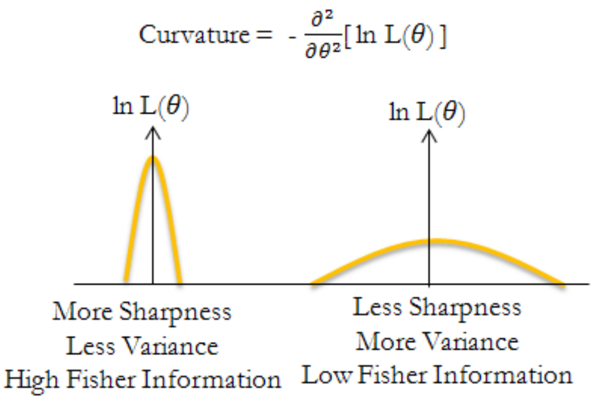
Pic: gaussianwaves.com

Cramér-Rao bound

Under regularity assumptions on probability distribution $p(x;\theta)$,

$$\text{var}(\hat{\theta}) \geq I^{-1}(\theta) = \frac{1}{-\mathbf{E}\left[\frac{\partial^2 \ln p(x;\theta)}{\partial \theta^2}\right]}.$$

If PDF $\mathrm{p}(x;\theta)$ is influenced by parameter more, estimation will be more accurate



Pic: gaussianwaves.com

Deterministic function of parameter

Suppose we want to estimate a function $g(\theta)$ of the parameter

Example *sensor measures a quantity θ, but the instantaneous power θ^2 is needed:*

$$g(\theta) = \theta^2$$

In this case,

$$\text{var}(\hat{\theta}) \geq \frac{\left(\frac{\partial g}{\partial \theta}\right)^2}{-\mathbf{E}\left[\frac{\partial^2 \ln p(x;\theta)}{\partial \theta^2}\right]}.$$

Cramér-Rao bound – multiple parameters

Cramér-Rao bound – multiple parameters

The Fisher information matrix (FIM) is the negative expected Hessian of the log-likelihood function:

$$\mathbf{I}_\theta = -\mathbf{E}\big[\nabla_\theta \, \nabla_\theta^H \ln p(\mathbf{x})\big] = -\mathbf{E}\big[\nabla_\theta^2 \ln p(\mathbf{x})\big], \tag{1}$$

where the expectation is taken with respect to $p(\mathbf{x})$ and

$$[\nabla_\theta f]_i = \partial f / \partial \theta_i, \qquad [\nabla_\theta^2 f]_{ij} = \partial f / \partial \theta_i \partial \theta_j^*.$$

Cramér-Rao bound – multiple parameters

The Fisher information matrix (FIM) is the negative expected Hessian of the log-likelihood function:

$$\mathbf{I}_\theta = -\,\mathbf{E}\big[\nabla_\theta\, \nabla_\theta^H \ln p(\mathbf{x})\big] = -\,\mathbf{E}\big[\nabla_\theta^2 \ln p(\mathbf{x})\big], \tag{1}$$

where the expectation is taken with respect to $p(\mathbf{x})$ and

$$[\nabla_\theta f]_i = \partial f / \partial \theta_i, \qquad [\nabla_\theta^2 f]_{ij} = \partial f / \partial \theta_i \partial \theta_j^*.$$

The covariance matrix $\mathbf{R}_{\hat{\theta}}$ of any unbiased estimator $\hat{\theta}$ of θ satisfies $\mathbf{R}_{\hat{\theta}} \succeq \mathbf{I}_\theta^{-1}$.

$\mathbf{A} \succeq \mathbf{B}$ means $\mathbf{A} - \mathbf{B}$ is positive semidefinite with $\mathbf{A}$ and $\mathbf{B}$ Hermitian: $\mathbf{A} = \mathbf{A}^H$, $\mathbf{B} = \mathbf{B}^H$.

CRB – complex parameters

CRB described until now holds for real parameters.
How to extend to complex parameters z?

Bos, A. van den. "A Cramer-Rao Lower Bound for Complex Parameters.", 1994
Brandwood, D. H. "A Complex Gradient Operator and Its Application in Adaptive Array Theory.", 1983

CRB – complex parameters

CRB described until now holds for real parameters.
How to extend to complex parameters z?

Two equivalent approaches:

- Consider real $Re(z)$ and imaginary part $Im(z)$ separately (cumbersome)
- Consider complex number z and its conjugate z^* (also cumbersome ☺ but generally easier)

Bos, A. van den. "A Cramer-Rao Lower Bound for Complex Parameters.", 1994
Brandwood, D. H. "A Complex Gradient Operator and Its Application in Adaptive Array Theory.", 1983

On the menu today

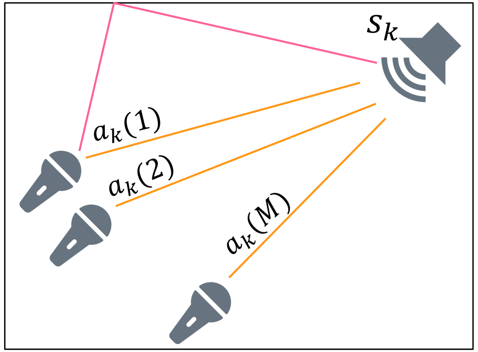
Cramér-Rao bound for acoustic transfer function estimation

1. Parameter estimation & Cramér-Rao bound (CRB)
2. Case study – ATF estimation

Acoustic transfer function estimation with **inter-frequency correlation**

1. Channel estimation algorithm
2. Experiments

Case study – channel estimation

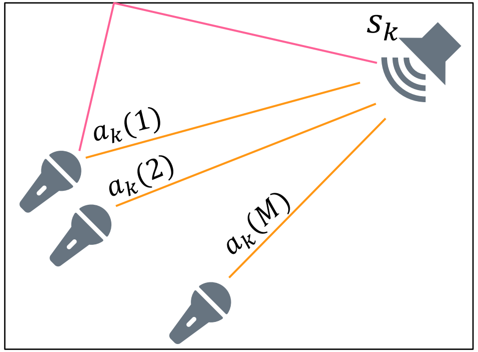


Signal model – frequency domain

Let a point source emit sound. The sound is measured by an array of M sensors. The received signal in the short-time Fourier transform (STFT) domain is

$$\mathbf{x}(l) = \mathbf{s}(l) + \mathbf{v}(l) = s(l)\,\mathbf{a} + \mathbf{v}(l) \in \mathbb{C}^M, \quad l = 1, \ldots, L.$$

Our goal: recover transfer function $\mathbf{a}$ from noisy recording $\mathbf{x}$



Deriving CRB for channel estimation

$$\mathrm{var}(\hat{\theta}) \geq I^{-1}(\theta) = \frac{1}{-\mathbf{E}\left[\frac{\partial^2 \ln p(x;\theta)}{\partial \theta^2}\right]}.$$

Likelihood function

Collect IID measurements in data matrix $\boldsymbol{X}$

Assume noise $\boldsymbol{v}$ is complex circular Gaussian process

Unknown parameters $\theta = [\mathbf{a}^T \mathbf{a}^H]^T \in \mathbb{C}^{2M}$

Conditional likelihood is

$$p(\mathbf{X}; \theta, s(l)) = \frac{1}{|\pi \mathbf{R}|^L} \exp\left(-\sum_{l=1}^{L} (\mathbf{x}(l) - s(l)\mathbf{a})^H \mathbf{R}^{-1} (\mathbf{x}(l) - s(l)\mathbf{a}) \right),$$

Stoica, P., and A. Nehorai. "Performance Study of Conditional and Unconditional Direction-of-Arrival Estimation." 1990

Log-likelihood and its derivatives

Define log-likelihood

$$L(\theta) = \ln p(\mathbf{X};\theta) = -L\ln|\pi\mathbf{R}| - \sum_{l=1}^{L}(\mathbf{x}(l) - s(l)\mathbf{a})^H\mathbf{R}^{-1}(\mathbf{x}(l) - s(l)\mathbf{a})$$

We then have

$$\nabla_{\mathbf{a}^*}L(\theta) = \mathbf{R}^{-1}\sum_{l=1}^{L}(s(l)^*\mathbf{x}(l) - |s(l)|^2\mathbf{a})$$

$$\nabla_{\mathbf{a}}L(\theta) = (\nabla_{\mathbf{a}^*}L(\theta))^*$$

$$-\mathbf{E}\big[\nabla_{\mathbf{a}^*}\nabla_{\mathbf{a}^*}^H L(\theta)\big] = -\mathbf{E}\big[\nabla_{\mathbf{a}^*}\nabla_{\mathbf{a}}^T L(\theta)\big] = \mathbf{E}\left[\mathbf{R}^{-1}\sum_{l=1}^{L}|s(l)|^2\right] = E_s L\mathbf{R}^{-1}$$

$$-\mathbf{E}\big[\nabla_{\mathbf{a}^*}\nabla_{\mathbf{a}}^H L(\theta)\big] = \mathbf{0}$$

Fisher information matrix

With this, the Fisher information matrix is

$$\mathbf{I}_\theta = \begin{bmatrix} E_s L \mathbf{R}^{-1} & \mathbf{0} \\ \mathbf{0} & E_s L \mathbf{R}^{-*} \end{bmatrix}. \tag{1}$$

The block-diagonal matrix can be easily be inverted, leading to

$$\mathbf{I}_\theta^{-1} = \begin{bmatrix} \frac{1}{E_s L}\mathbf{R} & \mathbf{0} \\ \mathbf{0} & \frac{1}{E_s L}\mathbf{R}^* \end{bmatrix} \tag{2}$$

Variance is finally bounded as:

$$\text{var}(\hat{a}_i) \geq \frac{[R]_{ii}}{E_s L}, \quad i = 1, \ldots, M. \tag{3}$$

Deterministic function of parameter

ATFs are often estimated in relation to a reference microphone r, as in $g(\theta) = g(\mathbf{a}, \mathbf{a}^*) = \mathbf{a}/a_r$. In this case,

$$\mathbf{R}_{g(\theta)} - (\nabla_\theta \mathbf{g}) \mathbf{I}_\theta^{-1} (\nabla_\theta^H \mathbf{g}) \geq 0, \tag{1}$$

where $\mathbf{R}_{g(\theta)}$ is the covariance matrix of $g(\theta)$. Choosing $r = 1$, Jacobian is

Deterministic function of parameter

ATFs are often estimated in relation to a reference microphone r, as in $g(\theta) = g(\mathbf{a}, \mathbf{a}^*) = \mathbf{a}/a_r$. In this case,

$$\mathbf{R}_{g(\theta)} - (\nabla_\theta \mathbf{g})\mathbf{I}_\theta^{-1}(\nabla_\theta^H \mathbf{g}) \geq 0, \tag{1}$$

where $\mathbf{R}_{g(\theta)}$ is the covariance matrix of $g(\theta)$. Choosing $r = 1$, Jacobian is

$$\nabla_\theta \mathbf{g} = \begin{bmatrix} \nabla_{\mathbf{a}} \mathbf{g} & \nabla_{\mathbf{a}^*} \mathbf{g} \end{bmatrix} \tag{2}$$

$$= \left[\begin{array}{ccccc|c} 0 & 0 & 0 & & 0 & \\ -\theta_2/|\theta_1|^2 & 1/\theta_1 & 0 & \dots & 0 & \\ -\theta_3/|\theta_1|^2 & 0 & 1/\theta_1 & \dots & 0 & \mathbf{0}_{M\times M} \\ \vdots & & & \ddots & \vdots & \\ -\theta_M/|\theta_1|^2 & 0 & \dots & 0 & 1/\theta_1 & \end{array}\right], \tag{3}$$

where $[\nabla_\theta \mathbf{f}]_{ij} = \partial f_i / \partial \theta_j$.

Experiments

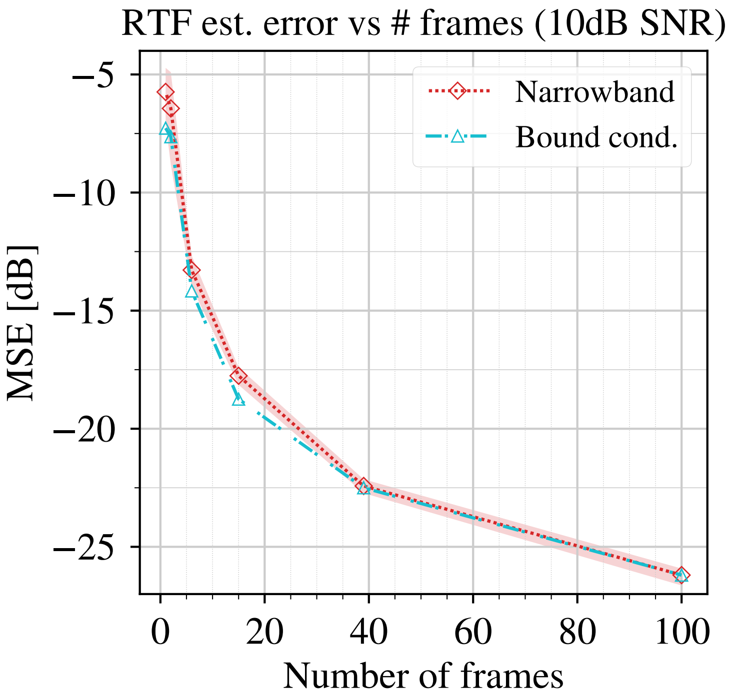
Settings

- 3 microphones,
  random ATF,
  covariance matrices estimated N snapshots,
  results averaged over 1000 experiments

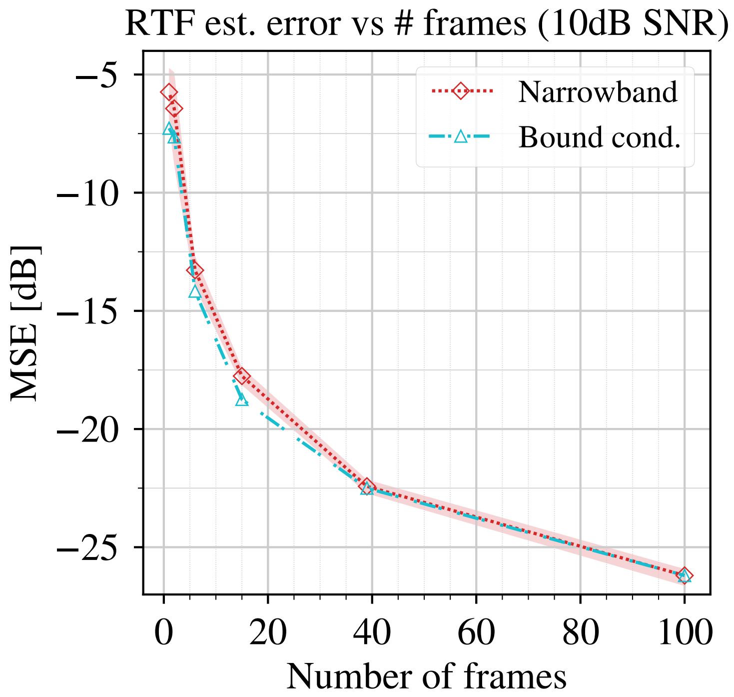
Performance metric

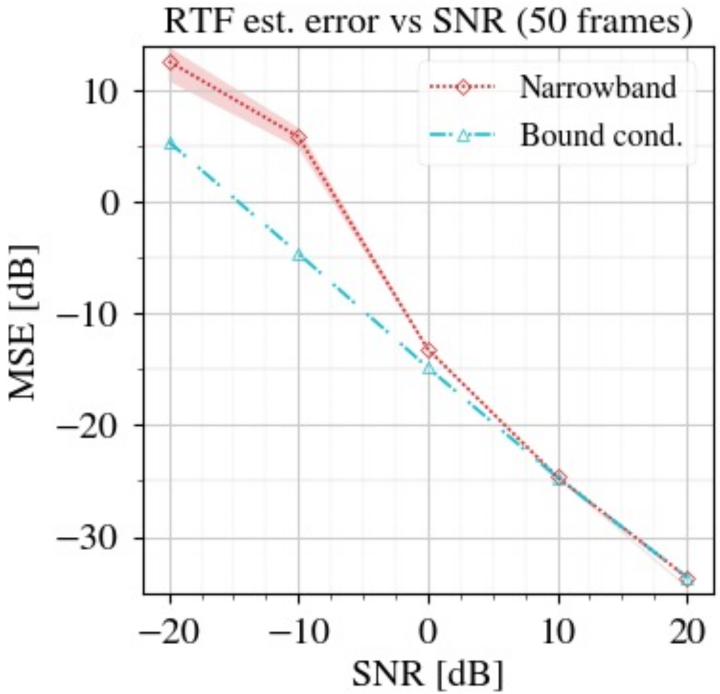
- Mean-squared error (in dB) between actual and estimated RTFs

Experiments



Experiments





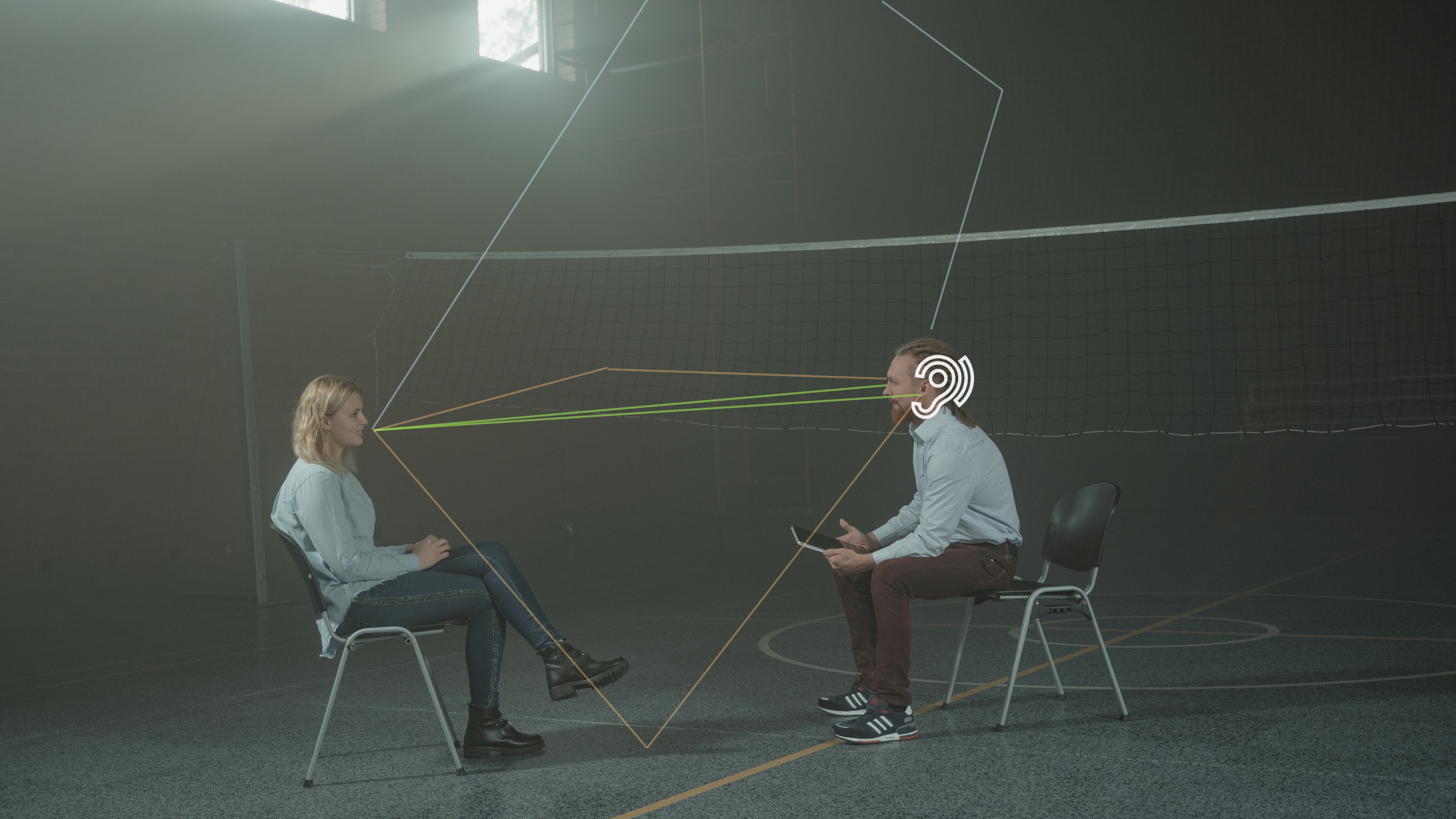
On the menu today

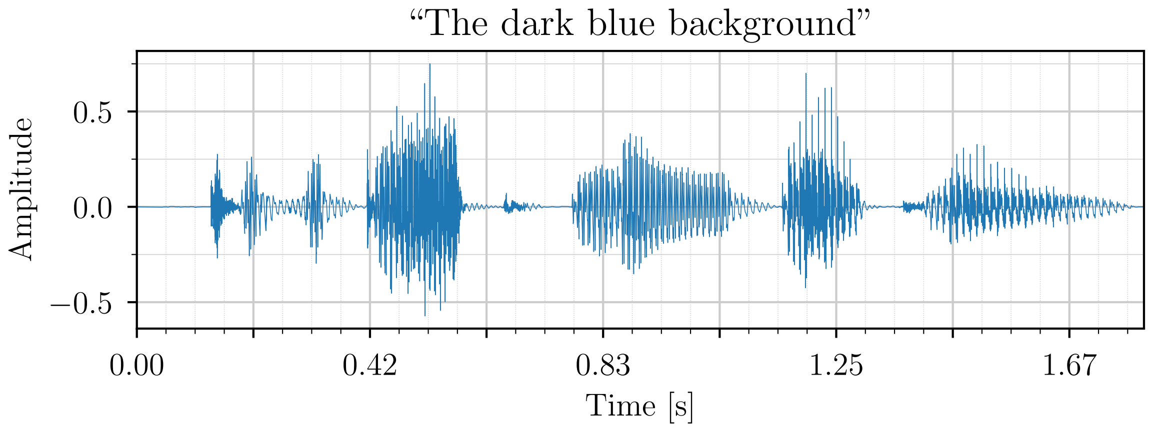
Cramér-Rao bound for acoustic transfer function estimation

1. Parameter estimation & Cramér-Rao bound (CRB)
2. Case study – ATF estimation

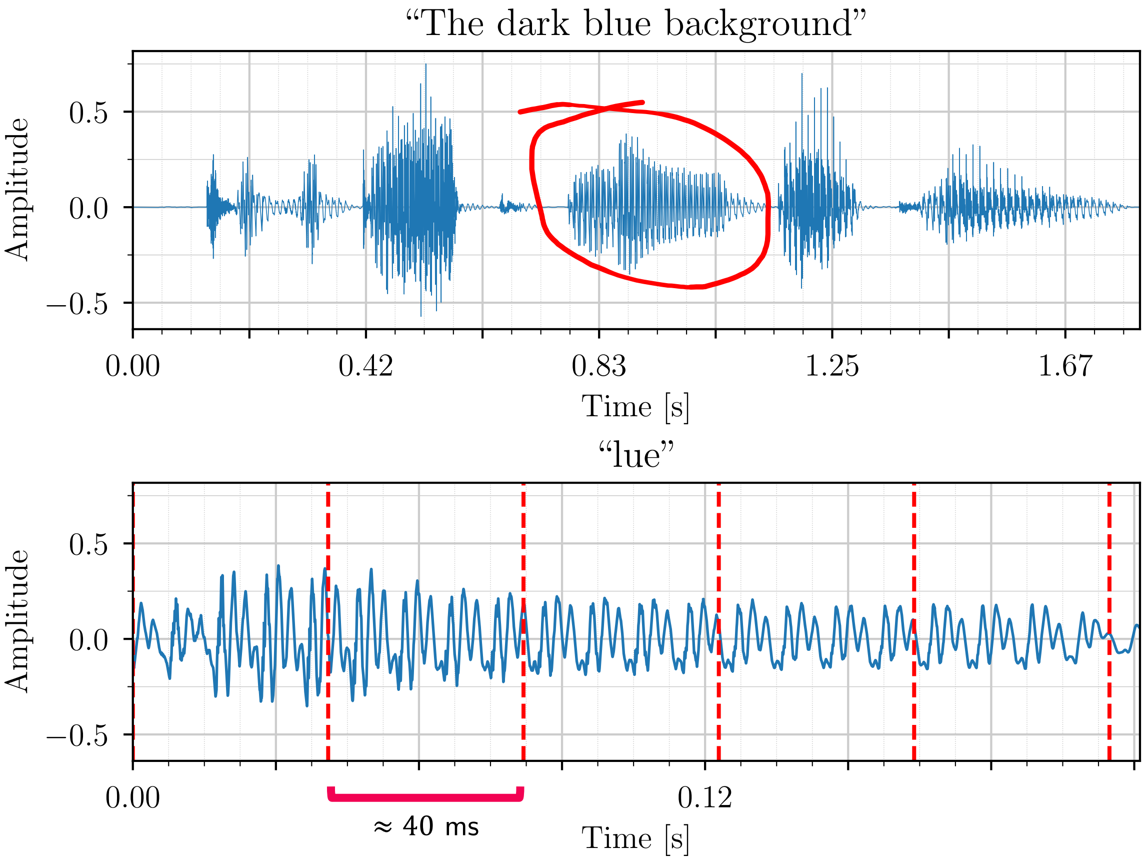
Acoustic transfer function estimation with **inter-frequency correlation**

1. Channel estimation algorithm
2. Experiments





“The dark blue background”
Amplitude
0.5
0.0
−0.5
0.00
0.42
0.83
1.25
1.67
Time [s]



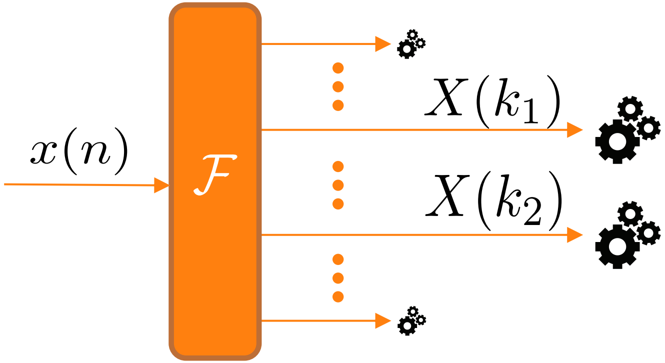
G. Bologni, R. Heusdens, and R. C. Hendriks "Harmonics to the rescue: Why voiced speech is not a WSS process", *submitted for consideration at the 18th International Workshop on Acoustic Signal Enhancement (IWAENC 2024)*

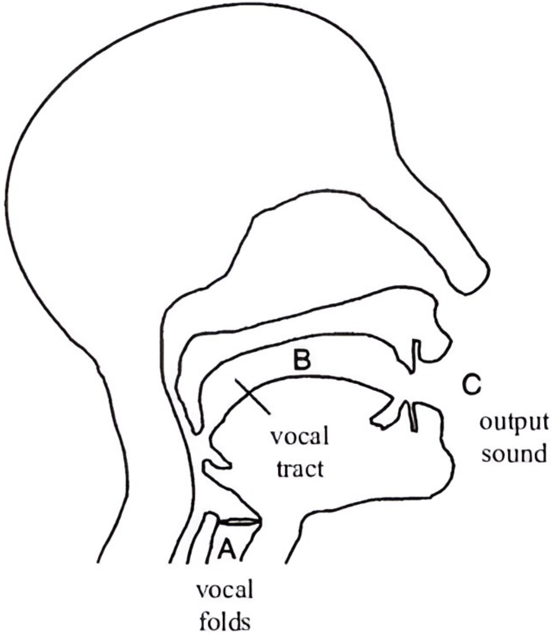
Inter-frequency correlations

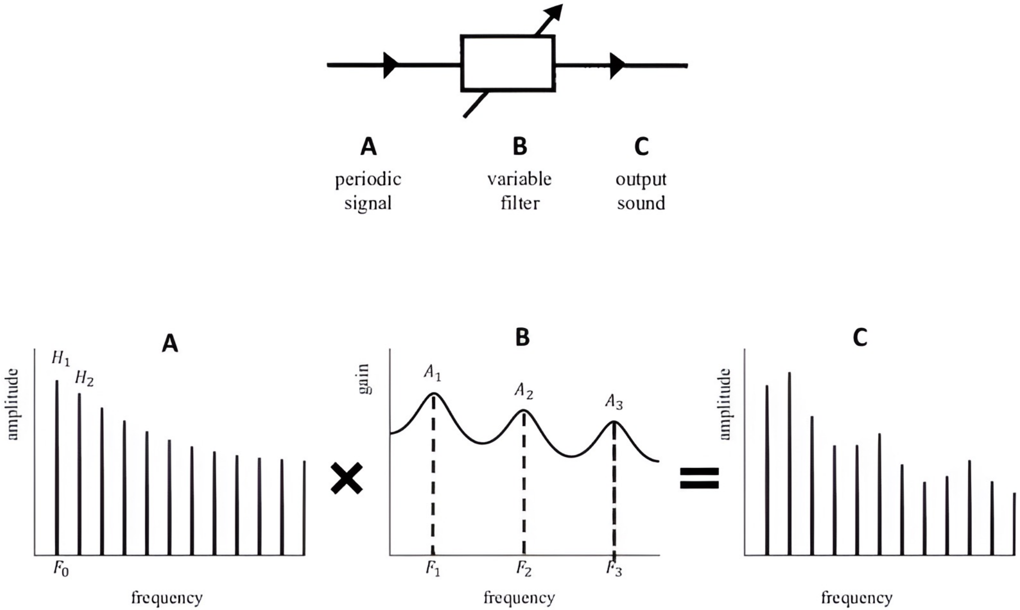
Signal processing algorithms (Wiener filter, MVDR) operate on **frequency-by-frequency** basis.

Frequency bins often assumed **mutually uncorrelated.**

Is assumption verified?





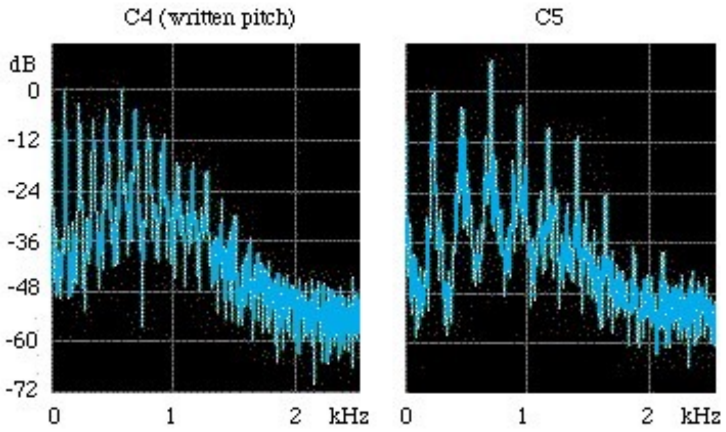


[1] S. J. Park, "Towards Understanding Voice Discrimination Abilities of Humans and Machines," PhD thesis, 2019.

Inter-frequency correlations

Inter-frequency correlations found in

1. Speech, wind instruments
2. Windowed signals ("frequency leakage" effect)
3. Non-stationary signals



Picture adapted from https://www.phys.unsw.edu.au/jw/brassacoustics.html

Benesty, Jacob, Jingdong Chen, and Emanuël A. P. Habets. "The Bifrequency Spectrum in Speech Enhancement." 2012.

Can acoustic parameter estimation be improved by exploiting “hidden” correlations across frequencies?

Inter-frequency correlations

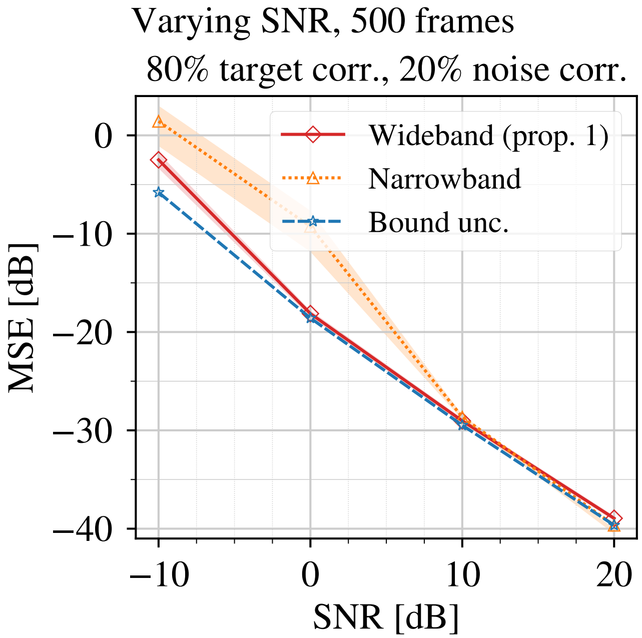
Inter-frequency correlations found in

1. Speech, wind instruments
2. Windowed signals (“frequency leakage” effect)
3. Non-stationary signals

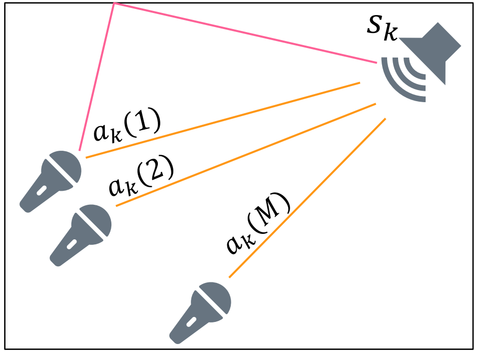
Picture adapted from https://www.phys.unsw.edu.au/jw/brassacoustics.html

Benesty, Jacob, Jingdong Chen, and Emanuël A. P. Habets. “The Bifrequency Spectrum in Speech Enhancement.” 2012.

Can acoustic parameter estimation be improved by exploiting “hidden” correlations across frequencies?



Channel estimation



Signal model – all frequencies

Noisy coefficients corresponding to single time frame l, at K different frequencies, can be stacked in a column as

$$\mathbf{x} = \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \vdots \\ \mathbf{x}_K \end{bmatrix} \in \mathbb{C}^{KM}.$$

Clean speech coefficients $\mathbf{s}$ and noise coefficients $\mathbf{v}$ can be obtained in the same way.

Signal model – all frequencies

The spatio-frequency correlation matrix $\mathbf{R}_\mathbf{x}$ can then be expressed as

$$\mathbf{R}_\mathbf{x} = \mathbf{E}\left[\mathbf{x}\mathbf{x}^H\right] = \begin{bmatrix} \mathbf{r}_x(1,1) & \mathbf{r}_x(1,2) & \cdots & \mathbf{r}_x(1,K) \\ \mathbf{r}_x(2,1) & \mathbf{r}_x(2,2) & \cdots & \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{r}_x(K,1) & \mathbf{r}_x(K,2) & \cdots & \mathbf{r}_x(K,K) \end{bmatrix} \in \mathbb{C}^{KM \times KM}, \qquad (1)$$

where the bifrequency spatial correlation $\mathbf{r}_x(i,j)$ between noisy vectors at two arbitrary frequencies, i and j, is the spatial correlation matrix

$$[\mathbf{R}_\mathbf{x}]_{ij} = \mathbf{r}_x(i,j) = \mathbf{E}[\mathbf{x}_i \mathbf{x}_j] \in \mathbb{C}^{M \times M}. \qquad (2)$$

Signal model – second order

$$\mathbf{R}_x = \mathbf{E}\left[\mathbf{x}\mathbf{x}^H\right] \tag{1}$$

$$= \mathbf{E}\left[(\mathbf{A}\mathbf{s} + \mathbf{v})(\mathbf{A}\mathbf{s} + \mathbf{v})^H\right] \tag{2}$$

$$= \mathbf{A}\,\mathbf{E}\left[\mathbf{s}\mathbf{s}^H\right]\mathbf{A}^H + \mathbf{E}\left[\mathbf{v}\mathbf{v}^H\right] \tag{3}$$

$$= \mathbf{A}\mathbf{R}_s\mathbf{A}^H + \mathbf{R}_v \tag{4}$$

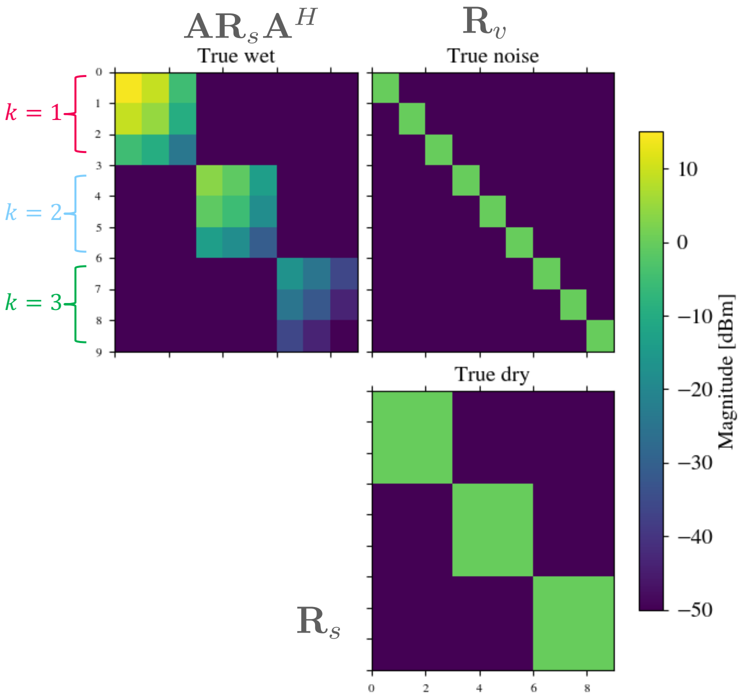
$$= \mathbf{R}_d + \mathbf{R}_v \tag{5}$$

$$\mathbf{A} = \mathrm{diag}(\mathbf{a}) = \mathrm{diag}(a_{11}, \ldots, a_{1M}, a_{21}, \ldots a_{KM}) \tag{6}$$

Signal model – second order

$$\mathbf{R}_x = \mathbf{E}\left[\mathbf{x}\mathbf{x}^H\right] \quad (1)$$
$$= \mathbf{E}\left[(\mathbf{A}\mathbf{s} + \mathbf{v})(\mathbf{A}\mathbf{s} + \mathbf{v})^H\right] \quad (2)$$
$$= \mathbf{A}\,\mathbf{E}\left[\mathbf{s}\mathbf{s}^H\right]\mathbf{A}^H + \mathbf{E}\left[\mathbf{v}\mathbf{v}^H\right] \quad (3)$$
$$= \mathbf{A}\mathbf{R}_s\mathbf{A}^H + \mathbf{R}_v \quad (4)$$

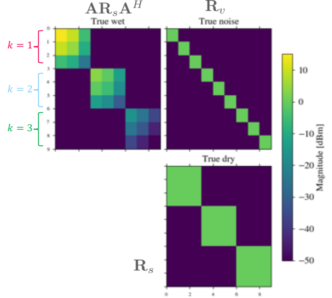
$$\mathbf{A} = \operatorname{diag}(\mathbf{a}) = \operatorname{diag}(a_{11}, \ldots, a_{1M}, a_{21}, \ldots a_{KM}) \quad (5)$$



RTF estimation – SVD direct

- We propose new RTF estimation algorithm that exploits inter-frequency correlations

- The algorithm is not ‘optimal’, but it sometimes achieve the CRB
  - Q: what would it take for ‘optimality’?



SVD-direct – key observation

When the number of frequencies is $K = 2$, the clean covariance matrix can be decomposed as:

$$\mathbf{R_d} = \begin{bmatrix} \sigma_{s_1}^2 \mathbf{a_1}\mathbf{a_1}^H & \mathbf{E}[s_1 s_2^*]\mathbf{a_1}\mathbf{a_2}^H \\ \mathbf{E}[s_2 s_1^*]\mathbf{a_2}\mathbf{a_1}^H & \sigma_{s_2}^2 \mathbf{a_2}\mathbf{a_2}^H \end{bmatrix} = \begin{bmatrix} \mathbf{R_d^1} \\ \mathbf{R_d^2} \end{bmatrix}, \tag{1}$$

Notice that $\mathbf{R_d^1}$ is rank-1 matrix.
Left principal singular vector of $\mathbf{R_d^1}$ is $\mathbf{a_1}$. Applying SVD,

$$\mathbf{R_d^1} = \mathbf{SDV}^H \approx d\,\mathbf{a}_1\mathbf{v}^H \tag{2}$$

SVD-direct – algorithm

1. Estimation of spatio-frequency covariance matrices: noisy $\hat{\mathbf{R}}_{\mathbf{x}}$, noise $\hat{\mathbf{R}}_{\mathbf{v}}$.

$$\hat{\mathbf{R}}_{\mathbf{x}} = \frac{1}{L}\sum_{l=1}^{L}\mathbf{x}(l)\mathbf{x}(l)^H \tag{1}$$

$$\hat{\mathbf{R}}_{\mathbf{v}} = \frac{1}{L}\sum_{l=1}^{L}\mathbf{v}(l)\mathbf{v}(l)^H \tag{2}$$

2. Estimation of $\hat{\mathbf{R}}_{\mathbf{d}}$ from generalized eigenvector:

$$\hat{\mathbf{R}}_{\mathbf{x}}\mathbf{U} = \hat{\mathbf{R}}_{\mathbf{v}}\mathbf{U}\mathbf{\Lambda} \tag{3}$$

$$(\hat{\mathbf{R}}_{\mathbf{d}} + \hat{\mathbf{R}}_{\mathbf{v}})\mathbf{U} = \hat{\mathbf{R}}_{\mathbf{v}}\mathbf{U}\mathbf{\Lambda} \tag{4}$$

$$\hat{\mathbf{R}}_{\mathbf{d}} \approx \hat{\mathbf{R}}_{\mathbf{v}}\mathbf{U}_1(\mathbf{\Lambda}_1 - \mathbf{I})\mathbf{U}_1^{-1} \tag{5}$$

SVD-direct – algorithm

3. Partition in K "fat" $M \times KM$ blocks:

$$\hat{\mathbf{R}}_{\mathbf{d}} = \begin{bmatrix} \hat{\mathbf{R}}_{\mathbf{d}}^{(1)} \\ \hat{\mathbf{R}}_{\mathbf{d}}^{(2)} \\ \vdots \\ \hat{\mathbf{R}}_{\mathbf{d}}^{(K)} \end{bmatrix}. \tag{6}$$

4. SVD on individual subblocks

$$\hat{\mathbf{R}}_{\mathbf{d}}^{(k)} = \mathbf{P}^{(k)} \mathbf{D}^{(k)} \mathbf{Q}^{(k)}. \tag{7}$$

5. Rescale left principal singular vectors

$$\hat{\mathbf{a}}^{(k)} = \text{Normalize}(\mathbf{p}_1^{(k)}).$$

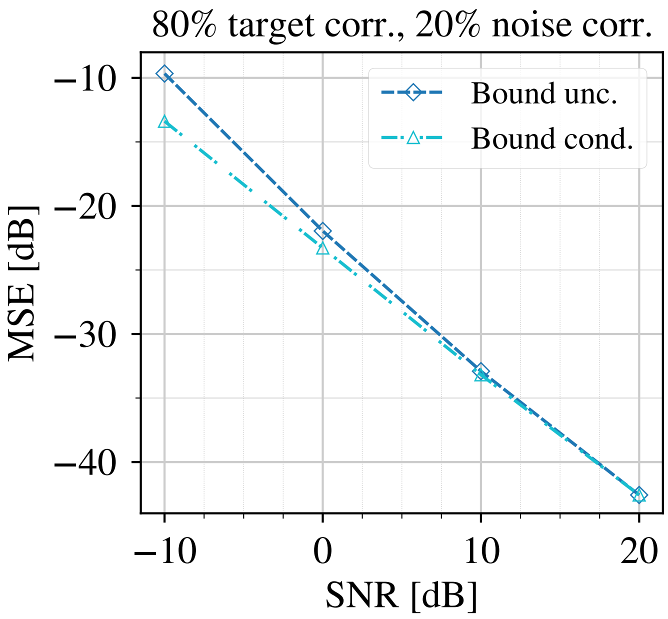
Two different CRBs

Conditional bound (*see prev. slides*)

True target received signal $\boldsymbol{s}$ is known

Unconditional bound

received signal $\boldsymbol{s}$ is unknown, but its first- and second-order statistics $\boldsymbol{R}_s$ are known



On the menu today

Cramér-Rao bound for acoustic transfer function estimation

1. Parameter estimation & Cramér-Rao bound (CRB)
2. Case study – ATF estimation

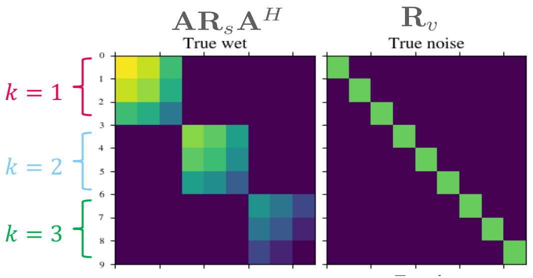
Acoustic transfer function estimation with **inter-frequency correlation**

1. Channel estimation algorithm
2. Experiments

How to simulate data

- To verify correctness of CRBs, we need access to the **true signal statistics**
- We generate random covariance matrices

Remember the signal model



$$\mathbf{R}_x = \mathbf{A}\mathbf{R}_s\mathbf{A}^H + \mathbf{R}_v$$

$$\mathbf{A} = \mathrm{diag}(\mathbf{a}) = \mathrm{diag}(a_{11}, \ldots, a_{1M}, a_{21}, \ldots a_{KM}) \quad (2)$$

Controlling correlations

Example for K=2 frequencies, M=2 sensors. Variable power.

$$\mathbf{R}_v^{\text{both}} = \begin{bmatrix} \sigma_{v_{11}}^2 & \rho\sqrt{\sigma_{v_{11}}^2 \sigma_{v_{12}}^2} & \cdots & \rho\sqrt{\sigma_{v_{11}}^2 \sigma_{v_{22}}^2} \\ & \sigma_{v_{12}}^2 & \cdots & \vdots \\ & & \sigma_{v_{21}}^2 & \rho\sqrt{\sigma_{v_{21}}^2 \sigma_{v_{22}}^2} \\ & & & \sigma_{v_{22}}^2 \end{bmatrix}, \quad (1)$$

where $\sigma_{v_{ii}}^2 \sim \mathcal{U}(5e-4, 5e-1), \forall i = 1, ..., MK$ and $\rho \in [0, 1]$.

Experiments

$M = 2$ microphones
$K = 5$ frequency components,
Random RTF $\boldsymbol{a}/a_{ref}$, real and imaginary parts drawn from $Uniform(-1, 1)$
Covariance matrix $\boldsymbol{R}_x$ estimated over multiple snapshots

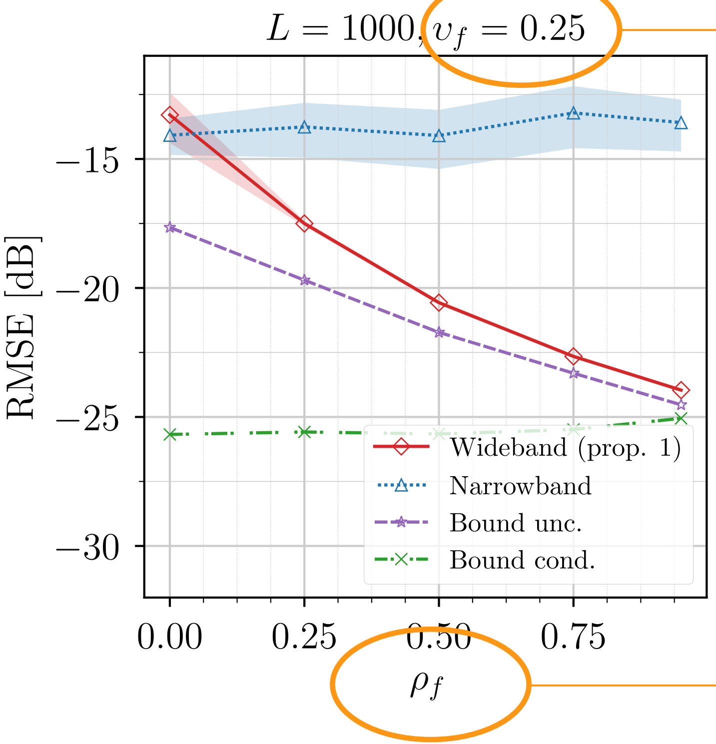
Oracle knowledge of $\boldsymbol{R}_v$

Target $\boldsymbol{s}(l)$ and noise $\boldsymbol{v}(l)$ signals drawn from complex Gaussian distributions with given covariance matrices

Performance metric

- Mean-squared error (in dB) between actual and estimated RTFs

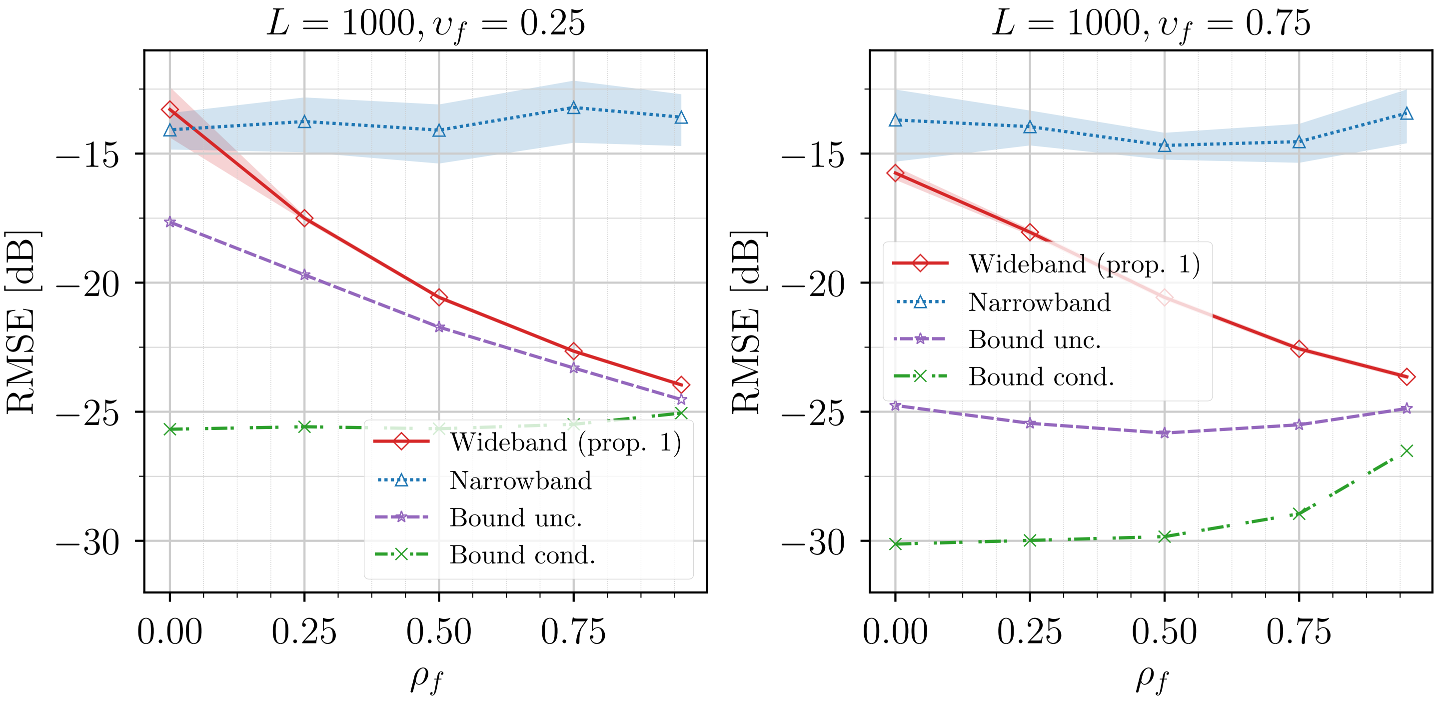
Synthetic data – varying target correlation



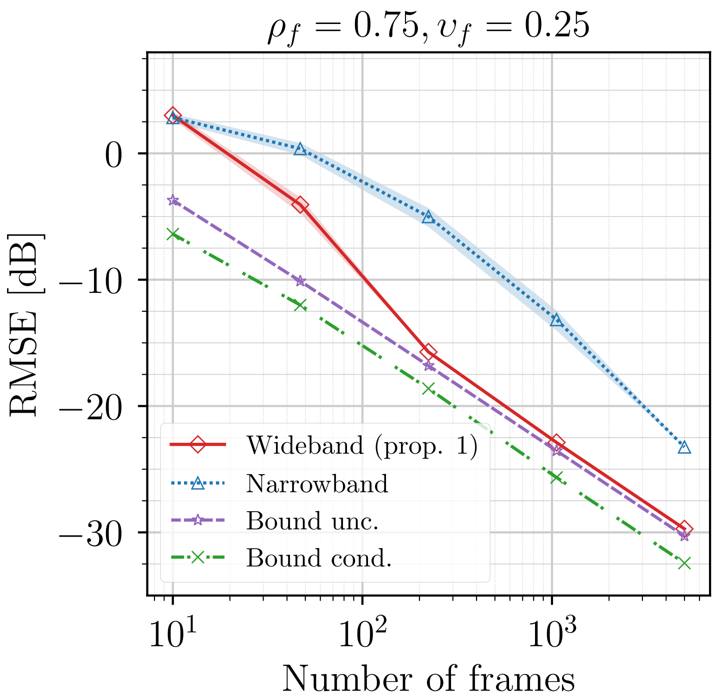
Noise frequency correlation

Target frequency correlation

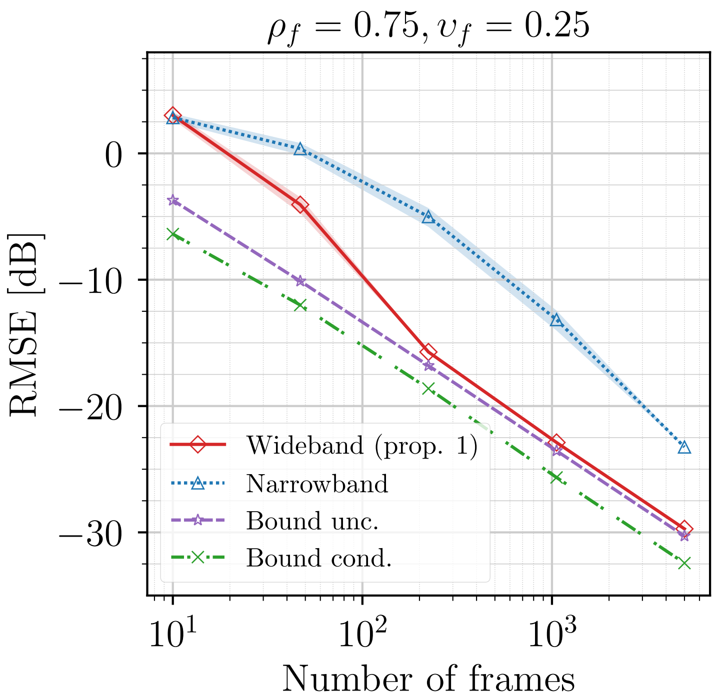
Synthetic data – varying target correlation

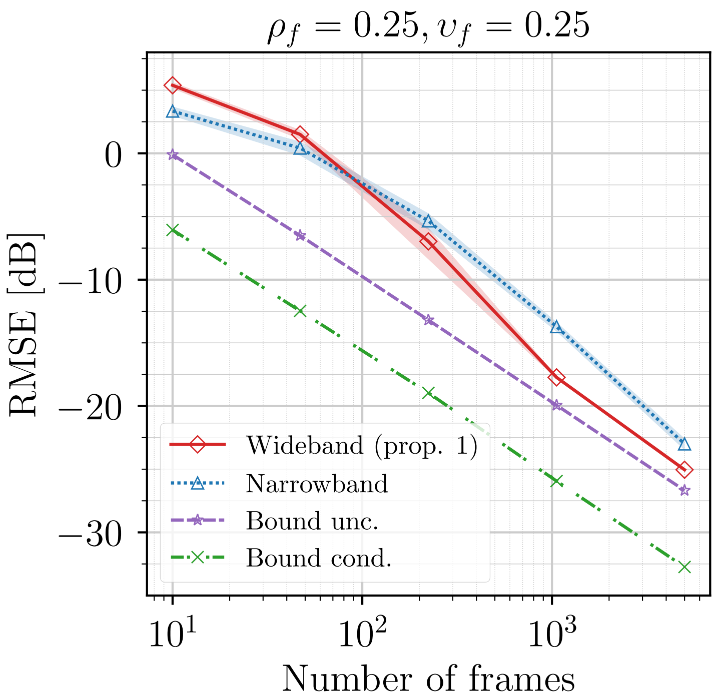


Synthetic data – varying number of frames

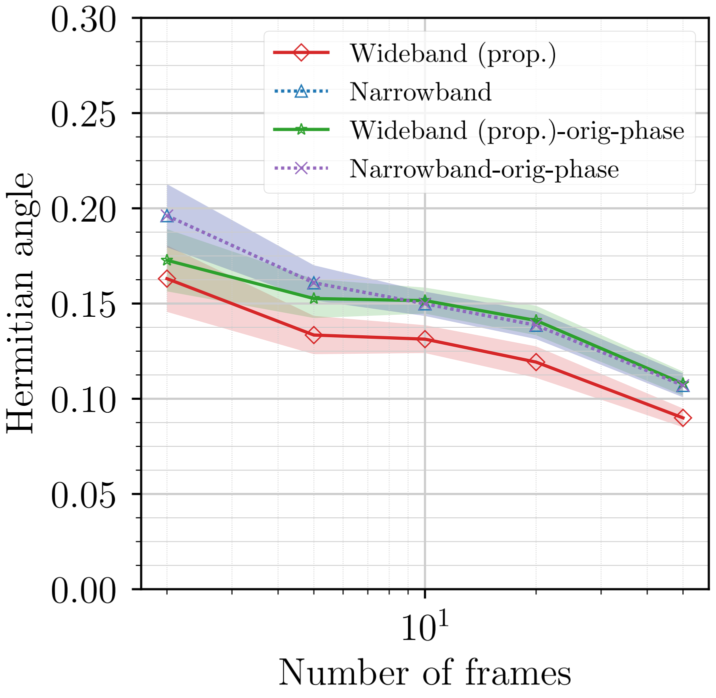


Synthetic data – varying number of frames

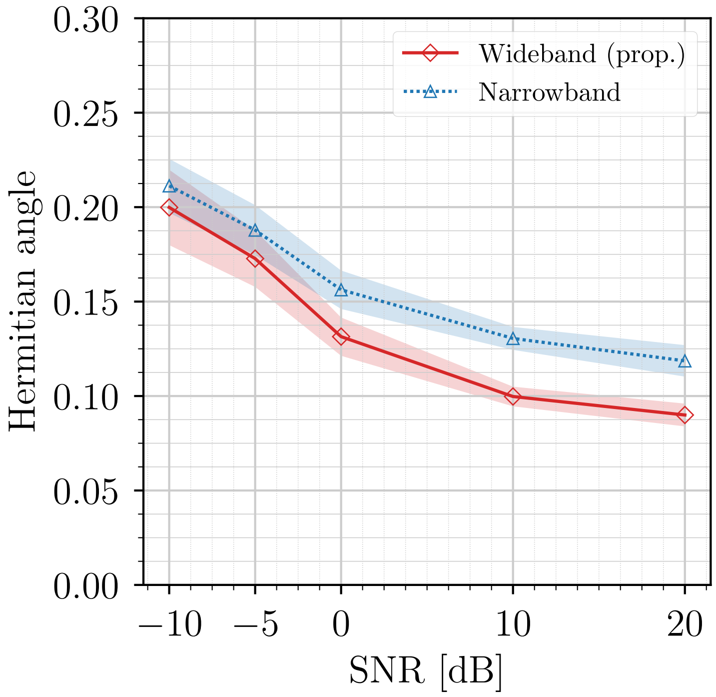




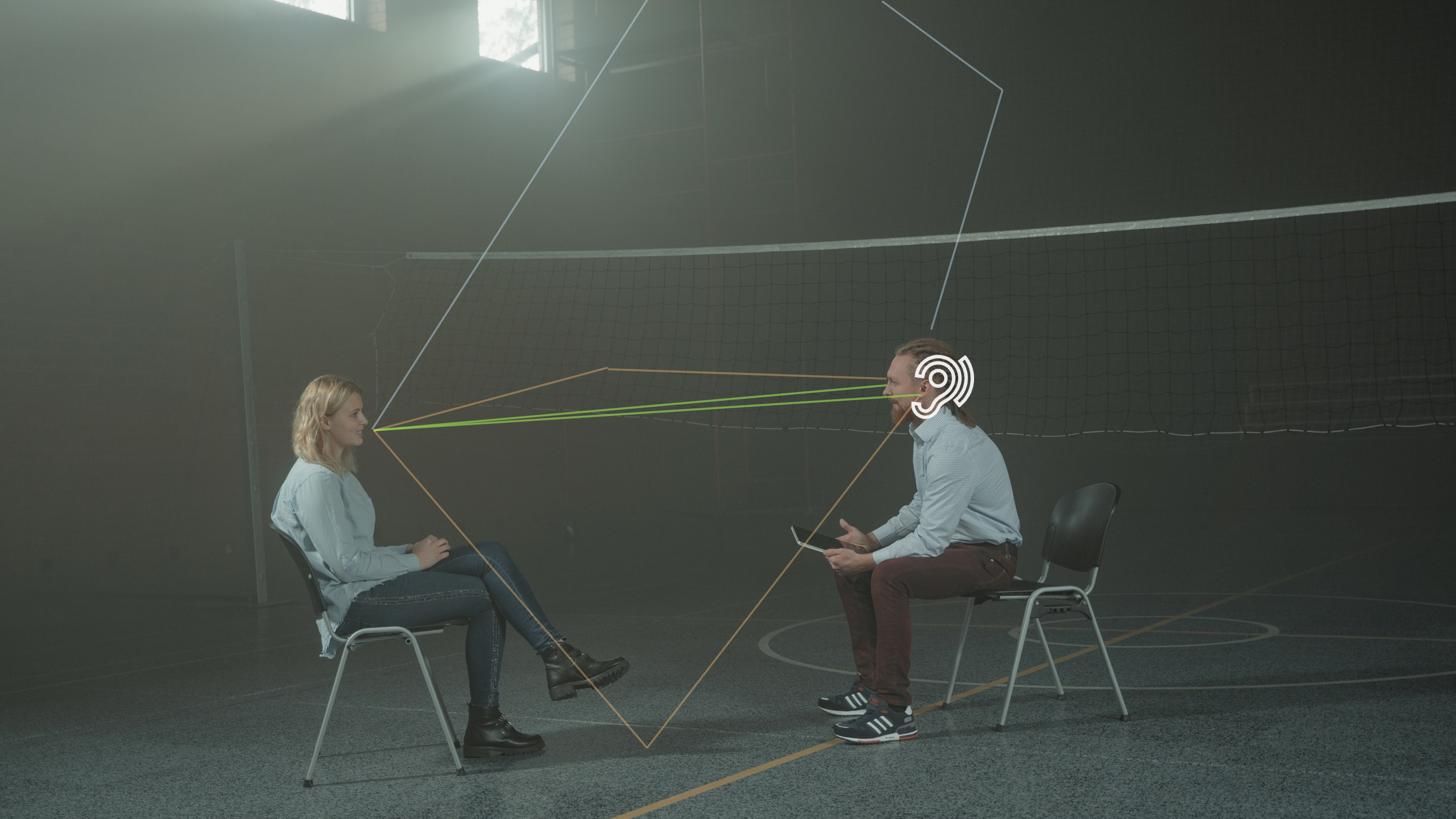
Real data



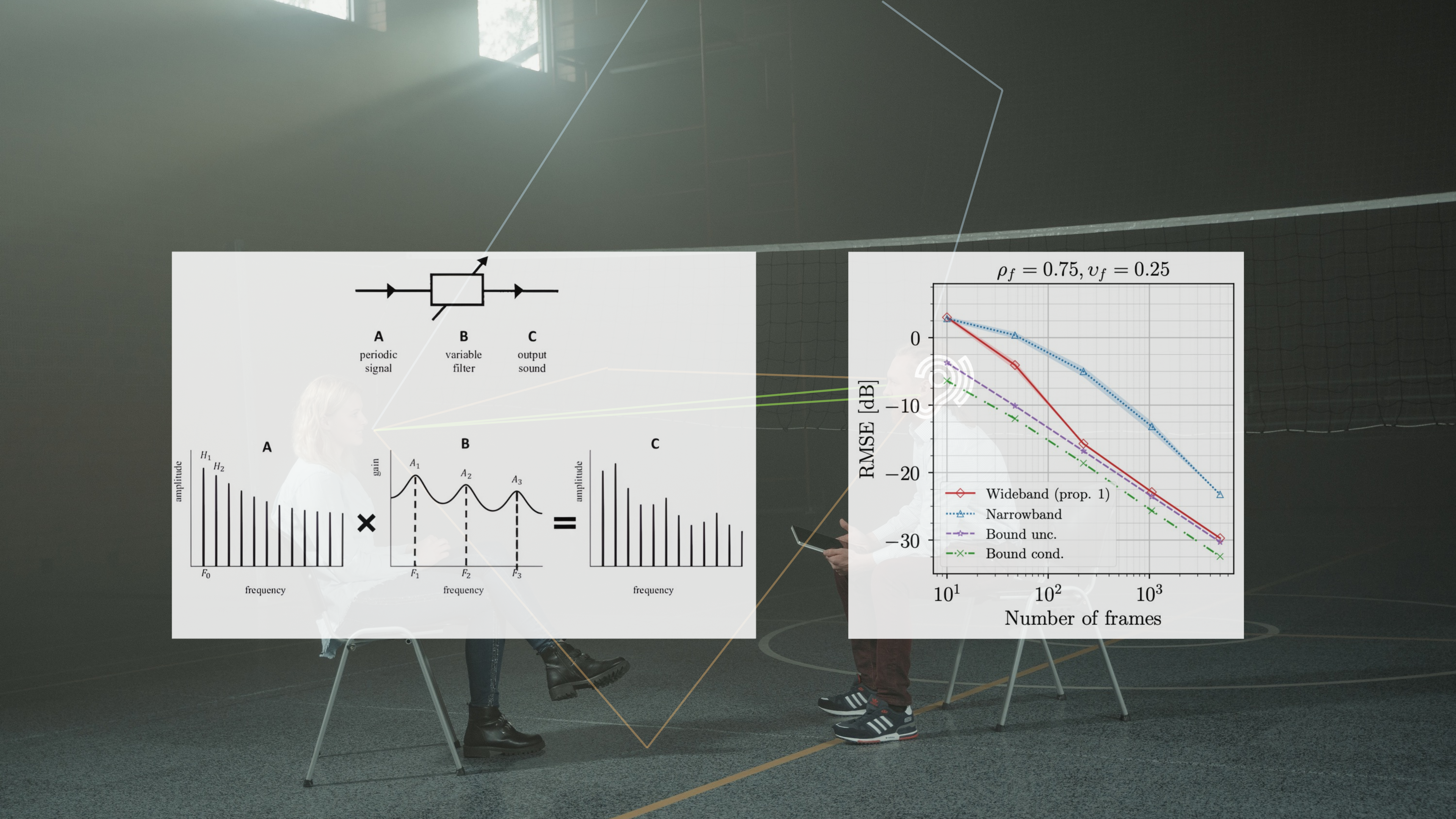
0.30
0.25
0.20
0.15
0.10
0.05
0.00
Hermitian angle
10^1
Number of frames
Wideband (prop.)
Narrowband
Wideband (prop.)-orig-phase
Narrowband-orig-phase



0.30
0.25
0.20
0.15
0.10
0.05
0.00
Hermitian angle
−10
−5
0
10
20
SNR [dB]
Wideband (prop.)
Narrowband




A
periodic signal
B
variable filter
C
output sound
A
amplitude
H1
H2
F0
frequency
×
B
gain
A1
A2
A3
F1
F2
F3
frequency
=
C
amplitude
frequency


A
periodic signal
B
variable filter
C
output sound
A
amplitude
H_1
H_2
F_0
frequency
B
gain
A_1
A_2
A_3
F_1
F_2
F_3
frequency
C
amplitude
frequency
$\rho_f = 0.75, \upsilon_f = 0.25$
RMSE [dB]
0
−10
−20
−30
10^1
10^2
10^3
Number of frames
Wideband (prop. 1)
Narrowband
Bound unc.
Bound cond.


MONDIAL MOVERS
ROADWORX
ROADWORX

Open questions - theses projects

- Collection of binaural impulse response database in our brand-new audio lab

- Wideband beamforming
  - *Wideband* MVDR with cue preservation [1]
  - ATF informed DNN [2]

- Improve DNN-based speech enhancement by including frequency-shifted waveforms as input

Do you have an idea to win the next Fields medal? Come talk to us!

[1] R. Chopra, D. Ghosh, and D. K. Mehra, "Spectrum Sensing for Cognitive Radios Based on Space-Time FRESH Filtering," IEEE Transactions on Wireless Communications, vol. 13, no. 7, pp. 3903–3913, Jul. 2014, doi: 10.1109/TWC.2014.2314125.

[2] A. Briegleb, M. M. Halimeh, and W. Kellermann, "Exploiting Spatial Information with the Informed Complex-Valued Spatial Autoencoder for Target Speaker Extraction," ICASSP 2023

Recap

Cramér-Rao bound for acoustic transfer function estimation

1. Parameter estimation & Cramér-Rao bound (CRB)
2. Case study – ATF estimation

Acoustic transfer function estimation with **inter-frequency correlation**

1. Channel estimation algorithm
2. Experiments

Cramér-Rao bound for acoustic transfer function estimation

Giovanni Bologni (G.Bologni@tudelft.nl), Richard Hendriks, Richard Heusdens

31 May, 2024

TU Delft