Introduction to distributed speech enhancement algorithms for ad hoc microphone arrays and wireless acoustic sensor networks

Part I: Array Processing in Acoustic Environments

Sharon Gannot[1] and Alexander Bertrand[2]

[1]Faculty of Engineering, Bar-Ilan University, Israel

[2]KU Leuven, E.E. Department ESAT-STADIUS, Belgium

EUSIPCO 2013, Marrakesh, Morocco

Acoustic Spatial Processing

Multi-Microphone Solutions

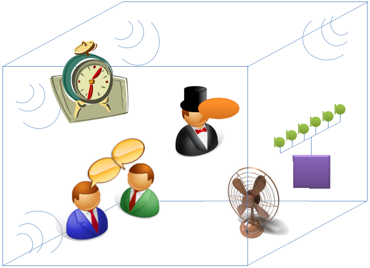
- Add the spatial domain to the time/frequency domain.
- Allow spatially selective algorithms for signal separation and noise suppression, which outperform single-microphone algorithms.
- Adopt array processing techniques to the acoustic world.

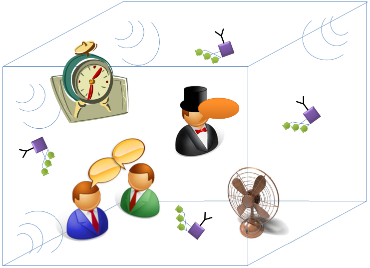
Distributed Microphone Arrays

- Microphones can be placed randomly, avoiding tedious calibration.
- Utilization of very large microphone number is possible, hence increased spatial resolution may be expected.
- High probability to find microphones close to a relevant sound source.
- Improved sound field sampling.

Challenges of Distributed Beamforming

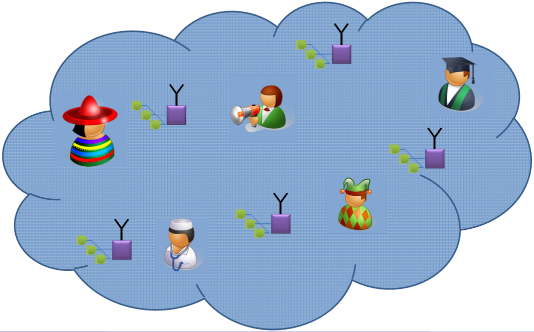
- Distributed microphone array beamforming:
    - Ad hoc sensor networks.
    - Large volume (and many nodes).
- Robustness:
    - High fault percentage.
    - Arbitrary deployment of nodes.
    - Sampling rate mismatch.





Tutorial Outline

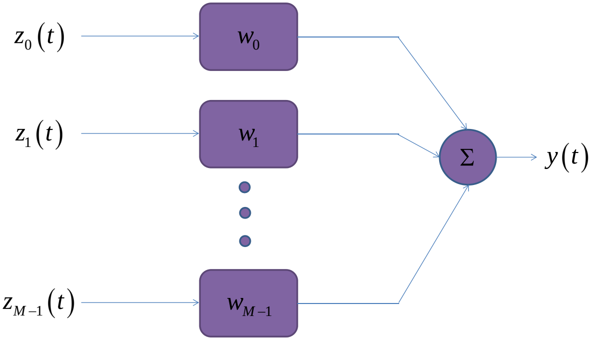
- **Part I** Array Processing in acoustic environment.
- **Part II** DANSE-based distributed speech enhancement in WASNs.
- **Part III** GSC-based distributed speech enhancement in WASNs.
- **Part IV** Random microphone deployment: Performance & Sampling rate mismatch.



Spatial Filters

Beamforming (Narrowband Signals):

$$y(t) = \mathbf{w}^H(t)\mathbf{z}(t).$$

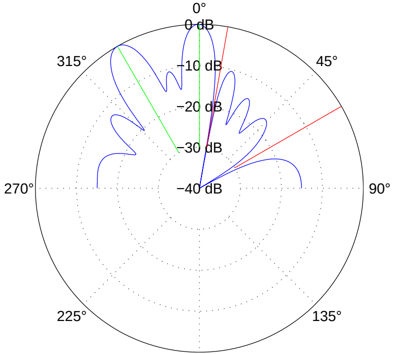


$\mathbf{w}$: $M \times 1$ beamforming vector of complex gains.

Beampattern Control

Beamformers

- Discriminate between angles.
- Can be steered by setting $\mathbf{w}$.
- Depends on the ratio $\frac{d}{\lambda_0}$.



Room Acoustics Essentials

Sound Fields

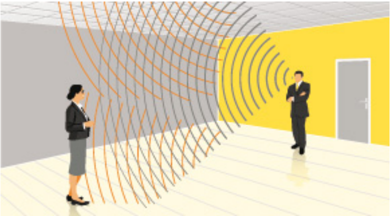
- Directional Room impulse response relates source and microphones.
- Uncorrelated Signals on microphone are uncorrelated.
- Diffused Sound is coming from all directions

[Dal-Degan and Prati, 1988];

[Habets and Gannot, 2007].

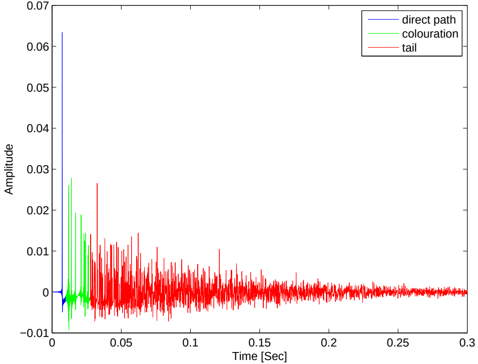
Reverberation

- Late reflections tend to be diffused.
- Deteriorates intelligibility.
- Degrades ASR performance.
- Beamforming becomes a cumbersome task.



The Room Impulse Response (RIR)

[Allen and Berkley, 1979]; simulator: [Habets, 2006]; [Polack, 1993]; [Jot et al., 1997]



3 Parts:

- Direct path.
- Colouration (early arrivals).
- Reverberation tail (late arrivals).

Reverberation should be taken into consideration while designing the algorithms even if it does not deteriorate speech quality and intelligibility.

From Geometry to Linear Algebra

Array Design for Speech Propagating in Acoustic Environments

- Beampattern: Array response as a function of the angle of arrival (AoA).
- In reverberant environments (especially for low DRR), sound propagation is more involved than merely the AoA.
- The steering vector generalizes to acoustic transfer function (ATF). Beampattern becomes meaningless.
- The ATF summarizes all arrivals of the speech signals.
- The vector of received signals is treated as a vector in an abstract linear space.
- Linear Algebra methods are utilized to construct beamformers.
- It is a cumbersome task to blindly estimate the ATFs.

Array Processing in Speech Applications I

1. Fixed beamforming Combine the microphone signals using a time-invariant filter-and-sum operation (data-independent)

   [Jan and Flanagan, 1996]; [Doclo and Moonen, 2003].

2. Blind Source Separation (BSS) Considers the received signals at the microphones as a mixture of all sound sources filtered by the RIRs. Utilizes Independent Component Analysis (ICA) techniques

   [Makino et al., 2007]; TRINICON, [Buchner et al., 2004].

3. Adaptive Beamforming Combine the spatial focusing of fixed beamformers with adaptive suppression of (spectrally and spatially time-varying) background noise

   General reading: [Cox et al., 1987]; [Van Veen and Buckley, 1988]; [Van Trees, 2002].

4. Computational Auditory Scene Analysis (CASA) Aims at performing sound segregation by modelling the human auditory perceptual processing [Wang and Brown, 2006].

Array Processing in Speech Applications II

Beamforming Criteria

1. Adaptive optimization [Sondhi and Elko, 1986]; [Kaneda and Ohga, 1986]; [Brandstein and Ward, 2001].
2. Minimum variance distortionless response (MVDR) and GSC [Van Compernolle, 1990]; [Affes and Grenier, 1997]; [Nordholm et al., 1993]; [Hoshuyama et al., 1999]; [Gannot et al., 2001]; [Herbordt, 2005]; [Gannot and Cohen, 2008].
3. Minimum mean square error (MMSE) - GSVD based spatial Wiener filter [Doclo and Moonen, 2002].
4. Speech distortion weighted multichannel Wiener filter (SDW-MWF) [Doclo and Moonen, 2002]; [Spriet et al., 2004]; [Doclo et al., 2005].
5. Maximum signal to noise ratio (SNR) [Warsitz and Haeb-Umbach, 2007].
6. Linearly constrained minimum variance (LCMV) [Markovich et al., 2009].

Array Processing in Speech Applications III

Some Books

1. Acoustic signal processing for telecommunication [Gay and Benesty, 2000].
2. Microphone Arrays: Signal Processing Techniques and Applications [Brandstein and Ward, 2001].
3. Speech Enhancement [Benesty et al., 2005].
4. Blind speech separation [Makino et al., 2007].
5. Microphone Array Signal Processing [Benesty et al., 2008a].
6. Springer handbook of speech processing [Benesty et al., 2008b].
7. Handbook on array processing and sensor networks [Haykin and Liu, 2010].
8. Speech processing in modern communication: Challenges and perspectives [Cohen et al., 2010].

Multiple Wideband Signals (e.g. Speech)

Multiplicative Transfer Function (MTF) Approximation

$t \stackrel{\text{STFT}}{\Longrightarrow} \{\ell, k\}$; Convolution $\stackrel{\text{STFT}}{\Longrightarrow}$ Multiplication (for long enough frames).

Microphone Signals ($m = 0, \ldots, M-1$):

$$z_m(\ell, k) = \sum_{j=1}^{P_d} s_j^d h_{jm}^d + \sum_{j=1}^{P_i} s_j^i h_{jm}^i + \sum_{j=1}^{P_n} s_j^n h_{jm}^n + n_m$$

Vector Formulation

$$\mathbf{z}(\ell, k) = \mathbf{H}^d \mathbf{s}^d + \mathbf{H}^i \mathbf{s}^i + \mathbf{H}^n \mathbf{s}^n + \mathbf{n} \triangleq \mathbf{H}\mathbf{s} + \mathbf{n}.$$

$$P = P_d + P_i + P_n \ \leq M$$

Beamforming in the STFT Domain

Apply filter & sum beamforming independently for each frequency bin.

Linearly Constrained Minimum Variance Beamformer

[Er and Cantoni, 1983]; [Van Veen and Buckley, 1988]

LCMV Criterion

- $y(\ell, k) = \mathbf{w}^H(\ell, k)\mathbf{z}(\ell, k)$.
- Let $\boldsymbol{\Phi}_{nn} = E\{\mathbf{n}\mathbf{n}^H\}$ be the $M \times M$ correlation matrix of the unconstraint sources.
- Minimize noise power $\mathbf{w}^H\boldsymbol{\Phi}_{nn}\mathbf{w}$
  Such that a linear constraint set is satisfied: $\mathbf{C}^H\mathbf{w} = \mathbf{g}$.
- $\mathbf{C}$: $M \times P$ constraints matrix.
- $\mathbf{g}$: $P \times 1$ response vector.

Closed-form Solution

$$\mathbf{w}(\ell, k) = \boldsymbol{\Phi}_{nn}^{-1}\mathbf{C}\left(\mathbf{C}^H\boldsymbol{\Phi}_{nn}^{-1}\mathbf{C}\right)^{-1}\mathbf{g}$$

Linearly Constrained Minimum Power (LCMP) Beamformer

[Van Trees, 2002]

LCMV vs. LCMP

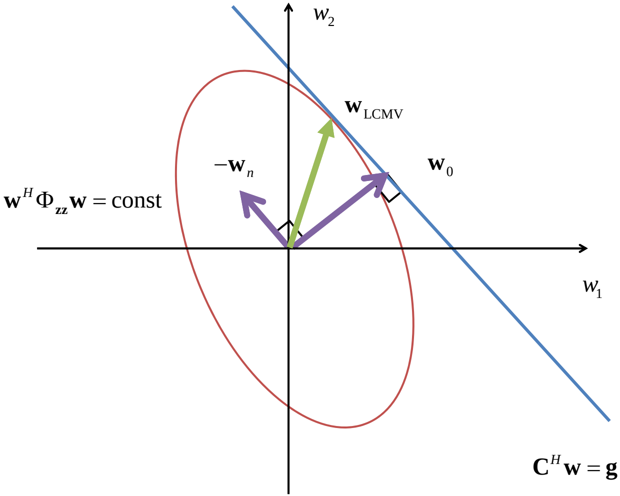
- Assume $\mathbf{C} = \mathbf{H}$ (all directional signals constrained).

$$\begin{aligned}
\mathbf{w}_{\mathrm{LCMP}} &= \underset{\mathbf{w}}{\operatorname{argmin}}\{\mathbf{w}^H \mathbf{\Phi}_{zz}\mathbf{w} \text{ s.t. } \mathbf{H}^H\mathbf{w} = \mathbf{g}\} \\
&= \underset{\mathbf{w}}{\operatorname{argmin}}\{\mathbf{w}^H (\mathbf{H}\mathbf{\Phi}_{ss}\mathbf{H}^H + \mathbf{\Phi}_{nn})\mathbf{w} \text{ s.t. } \mathbf{H}^H\mathbf{w} = \mathbf{g}\} \\
&= \underset{\mathbf{w}}{\operatorname{argmin}}\{\mathbf{g}^H \mathbf{\Phi}_{ss}\mathbf{g} + \mathbf{w}^H\mathbf{\Phi}_{nn}\mathbf{w} \text{ s.t. } \mathbf{H}^H\mathbf{w} = \mathbf{g}\} \\
&= \underset{\mathbf{w}}{\operatorname{argmin}}\{\mathbf{w}^H \mathbf{\Phi}_{nn}\mathbf{w} \text{ s.t. } \mathbf{H}^H\mathbf{w} = \mathbf{g}\} = \mathbf{w}_{\mathrm{LCMV}}
\end{aligned}$$

- If $\mathbf{H}$ is not accurately estimated, the LCMP beamformer exhibits self-cancellation and hence severe speech distortion.
- It is quite common in the literature to use only the term LCMV for both beamformers.

LCMV Minimization

Graphical Interpretation [Frost III, 1972]



The Minimum Variance Distortionless Beamformer

[Affes and Grenier, 1997]; [Hoshuyama et al., 1999]; [Gannot et al., 2001]

Beamformer Design:

- One desired signal $\Rightarrow$ Single constraint ($P = 1$).
- "Steer a beam" to desired source and minimize other directions.
- $\mathbf{C} = \mathbf{h}^d; \quad \mathbf{g} = 1.$

Closed-form Solution (MPDR eq. MVDR):

$$\mathbf{w}(\ell, k) = \frac{\mathbf{\Phi}_{zz}^{-1}\mathbf{h}^d}{(\mathbf{h}^d)^H\mathbf{\Phi}_{zz}^{-1}\mathbf{h}^d} = \frac{\mathbf{\Phi}_{nn}^{-1}\mathbf{h}^d}{(\mathbf{h}^d)^H\mathbf{\Phi}_{n}^{-1}\mathbf{h}^d}$$

Output signal:

$$y = s^d + \text{residual noise and interference signals}$$

The Relative Transfer Function [Gannot et al., 2001]

Modified Constraint Set:

$$\mathbf{C}(\ell, k) = \mathbf{h}^d(\ell, k); \quad \tilde{\mathbf{g}}(\ell, k) = (h_1^d(\ell, k))^*$$
$$\Rightarrow (\mathbf{h}^d(\ell, k))^H \mathbf{w} = (h_1^d(\ell, k))^*$$

Equivalent to:

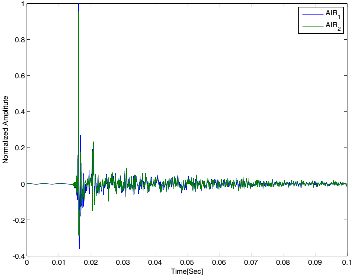
$$\tilde{\mathbf{C}}(\ell, k) = \tilde{\mathbf{h}}^d(\ell, k) \triangleq \frac{\mathbf{h}^d}{h_1^d} = \left[1 \ \frac{h_2^d}{h_1^d} \ \ldots \ \frac{h_M^d}{h_1^d} \right]^T; \ \mathbf{g}(\ell, k) = 1.$$

with $\tilde{\mathbf{h}}^d(\ell, k)$ the relative transfer function - the ratio of all ATFs to the reference ATF (#1 in this case)

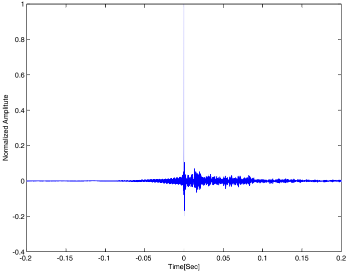
Output signal:

$$y = h_1^d s^d + \text{residual noise and interference signals}$$

The Importance of the RTF



(a) Room Impulse Responses



(b) Relative Impulse Response

Features

- Can be blindly estimated from data.
- No need to know microphone position (crucial in ad hoc applications).
- Multitude estimation procedures exists.
- Usually exhibits “better behaviour” than the ATF.
- Drawback: Non-causal (in severe cases can cause “pre-echo”).

RTF Estimation Procedures

- Utilizing speech non-stationarity and noise stationarity [Shalvi and Weinstein, 1996]; [Gannot et al., 2001].
  - An extension to two nonstationary sources in stationary noise exists [Reuven et al., 2008].
- Utilizing speech presence probability and spectral subtraction [Cohen, 2004].
- Based on eigenvalue decomposition (EVD) of the spatial correlation matrix for the multiple sources case [Markovich et al., 2009]. Nonconcurrent desired and interference sources.
  - An extension to concurrent desired and interference source, based on ICA (TRINICON), exists [Reindl et al., 2013].
- Recursive extensions exist:
  - Single source: use PASTd [Yang, 1995] to recursively track the rank-1 eigenvector [Affes and Grenier, 1997].
  - Multiple sources: use generalization of PASTd to recursively track the rank-P eigenvectors with arbitrary activity pattern [Markovich-Golan et al., 2010].

Multiple Speech Distortion Weighted Multichannel Wiener Filter (MSDW-MWF)[Markovich-Golan et al., 2012]

Notation (Reminder)

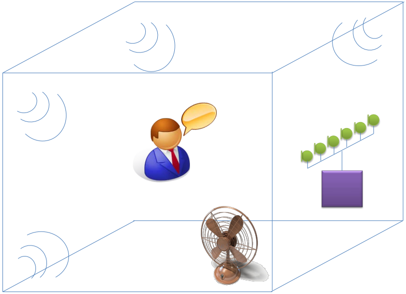
- Received signals: $\mathbf{z}(\ell, k) = \mathbf{H}\mathbf{s} + \mathbf{n}$.
- $P < M$ constrained sources: $\mathbf{s}(\ell, k) \triangleq \begin{bmatrix} s_1 & \cdots & s_P \end{bmatrix}^T$ and respective ATFs: $\mathbf{H}(\ell, k) \triangleq \begin{bmatrix} \mathbf{h}_1 & \cdots & \mathbf{h}_P \end{bmatrix}$.
- Sources covariance matrix: $\mathbf{\Phi}_{ss} = \operatorname{diag}\left\{\phi_{s_1 s_1}, \ldots, \phi_{s_P s_P}\right\}$.
- Microphones covariance matrix: $\mathbf{\Phi}_{zz} \triangleq \mathbf{H}\mathbf{\Phi}_{ss}\mathbf{H}^H + \mathbf{\Phi}_{nn}$.

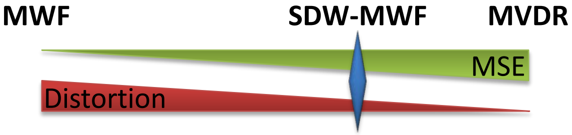
MSDW-MWF

- Control the distortion of each individual source.
- Minimize the weighted mean square error (MSE).
- Desired response for all constrained signals: $d(\ell, k) \triangleq \mathbf{g}^H \mathbf{s}(\ell, k)$.
- The beamformer output: $y(\ell, k) = \mathbf{w}^H \mathbf{z}(\ell, k)$.
- MSE: $\mathrm{E}\left\{|d(\ell) - y(\ell)|^2\right\}$.

Speech enhancement with a Single Source I

Speech Distortion Weighted Multichannel Wiener Filter (SDW-MWF) [Doclo and Moonen, 2002]; [Spriet et al., 2004]; [Doclo et al., 2005]





Speech enhancement with a Single Source II

Speech Distortion Weighted Multichannel Wiener Filter (SDW-MWF) [Doclo and Moonen, 2002]; [Spriet et al., 2004]; [Doclo et al., 2005]

The Multichannel Wiener Filter (MWF) Criterion

$$J_w \triangleq \mathrm{E}\left\{ \left| d(\ell) - y(\ell) \right|^2 \right\} = \left| g - (\mathbf{h}^d)^H \mathbf{w} \right|^2 \phi_{s^d s^d} + \mathbf{w}^H \mathbf{\Phi}_{nn} \mathbf{w}$$

The Speech Distortion Weighted (SDW)-MWF Criterion

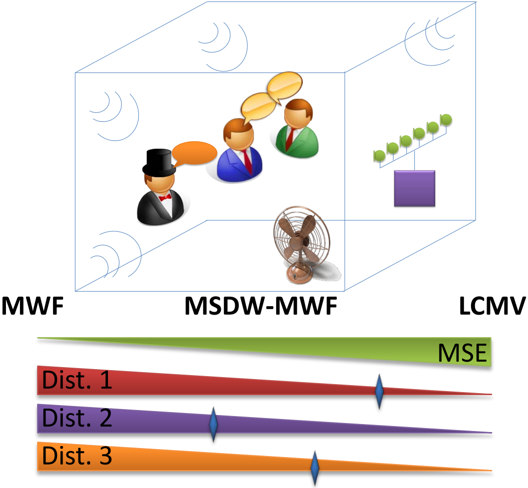
$$J_{\mathrm{SDW\text{-}MWF}} = \left| g - (\mathbf{h}^d)^H \mathbf{w} \right|^2 \phi_{s^d s^d} + \mu \mathbf{w}^H \mathbf{\Phi}_{nn} \mathbf{w}$$

SDW-MWF Solution(Requires VAD)

$$\mathbf{w} = \frac{\phi_{s^d s^d} \mathbf{\Phi}_{nn}^{-1} \mathbf{h}^d}{\mu + \phi_{s^d s^d} (\mathbf{h}^d)^H \mathbf{\Phi}_{nn}^{-1} \mathbf{h}^d} g$$

Speech Enhancement with Multiple Sources I

[Markovich-Golan et al., 2012]



Speech Enhancement with Multiple Sources II

[Markovich-Golan et al., 2012]

The MSDW-MWF Criterion

$$J_{\text{MSDW-MWF}} \triangleq \left(\mathbf{g} - \mathbf{H}^H\mathbf{w}\right)^H \mathbf{\Lambda}\mathbf{\Phi}_{ss}\left(\mathbf{g} - \mathbf{H}^H\mathbf{w}\right) + \mathbf{w}^H\mathbf{\Phi}_{nn}\mathbf{w}$$

- Diagonal weights matrix: $\mathbf{\Lambda} \triangleq \text{diag}\left\{\lambda_1, .., \lambda_P\right\}$.

MSDW-MWF Beamformer (Requires VAD)

$$\mathbf{w} \triangleq \left(\mathbf{H}\mathbf{\Lambda}\mathbf{\Phi}_{ss}\mathbf{H}^H + \mathbf{\Phi}_{nn}\right)^{-1}\mathbf{H}\mathbf{\Lambda}\mathbf{\Phi}_{ss}\mathbf{g}$$

Special Cases of $\boldsymbol{\Lambda}$

MWF

- $\boldsymbol{\Lambda} = \mathbf{I}$.
- $\mathbf{w} = \boldsymbol{\Phi}_{zz}^{-1}\mathbf{H}\boldsymbol{\Phi}_{ss}\mathbf{g}$.

SDW-MWF (Reminder: Single Source of Interest)

- $\boldsymbol{\Lambda} = \mu^{-1}$.
- $\mathbf{w} = \left(\mathbf{h}^d \phi_{s^d s^d} (\mathbf{h}^d)^H + \mu \boldsymbol{\Phi}_{nn}\right)^{-1} \mathbf{h}^d \phi_{s^d s^d} g$.
- $\lim_{\mu \to 0} \mathbf{w} = \frac{\boldsymbol{\Phi}_{nn}^{-1}\mathbf{h}^d}{(\mathbf{h}^d)^H \boldsymbol{\Phi}_{nn}^{-1}\mathbf{h}^d} g$ (MVDR eq. MPDR).

LCMV

- $\boldsymbol{\Lambda} = \mu^{-1}\boldsymbol{\Phi}_{ss}^{-1}$.
- $\lim_{\mu \to 0} \mathbf{w} = \boldsymbol{\Phi}_{nn}^{-1}\mathbf{H}\left(\mathbf{H}^H \boldsymbol{\Phi}_{nn}^{-1}\mathbf{H}\right)^{-1}\mathbf{g}$ (LCMV eq. LCMP).

References and Further Reading I

Affes, S. and Grenier, Y. (1997).
A signal subspace tracking algorithm for microphone array processing of speech.
IEEE transactions on Speech and Audio Processing, 5(5):425–437.

Allen, J. and Berkley, D. (1979).
Image method for efficiently simulating small-room acoustics.
J. Acoustical Society of America, 65(4):943–950.

Benesty, J., Chen, J., and Huang, Y. (2008a).
Microphone array signal processing.
Springer.

Benesty, J., Huang, Y., and Sondhi, M., editors (2008b).
Springer handbook of speech processing.
Springer Verlag.

Benesty, J., Makino, S., and Chen, J., editors (2005).
Speech Enhancement.
Signals and Communication Technology. Springer, Berlin.

Brandstein, M. S. and Ward, D. B., editors (2001).
Microphone Arrays: Signal Processing Techniques and Applications.
Springer-Verlag, Berlin.

Buchner, H., Aichner, R., and Kellermann, W. (2004).
TRINICON: A versatile framework for multichannel blind signal processing.
In *IEEE International Conference on Acoustics, Speech, and Signal Processing*, volume 3, pages iii–889, Montreal, Canda.

References and Further Reading II

Cohen, I. (2004).
Relative transfer function identification using speech signals.
IEEE Transactions on Speech and Audio Processing, 12(5):451–459.

Cohen, I., Benesty, J., and Gannot, S., editors (2010).
Speech processing in modern communication: Challenges and perspectives.
Topics in signal processing. Springer.

Cox, H., Zeskind, R., and Owen, M. (1987).
Robust adaptive beamforming.
IEEE trans. on Acoustics, Speech and Signal Proc., 35(10):1365–1376.

Dal-Degan, N. and Prati, C. (1988).
Acoustic noise analysis and speech enhancement techniques for mobile radio application.
Signal Processing, 15(4):43–56.

Doclo, S. and Moonen, M. (2002).
GSVD-based optimal filtering for single and multimicrophone speech enhancement.
IEEE Trans. on Signal Processing, 50(9):2230–2244.

Doclo, S. and Moonen, M. (2003).
Design of far-field and near-field broadband beamformers using eigenfilters.
Signal Processing, 83(12):2641 – 2673.

Doclo, S., Spriet, A., Wouters, J., and Moonen, M. (2005).
Speech Enhancement, chapter Speech Distortion Weighted Multichannel Wiener Filtering Techniques for Noise Reduction, pages 199–228.
In [Benesty et al., 2005].

References and Further Reading III

Er, M. and Cantoni, A. (1983).
Derivative constraints for broad-band element space antenna array processors.
IEEE Transactions on Acoustics, Speech and Signal Processing, 31(6):1378–1393.

Frost III, O. L. (1972).
An algorithm for linearly constrained adaptive array processing.
Proceedings of the IEEE, 60(8):926–935.

Gannot, S., Burshtein, D., and Weinstein, E. (2001).
Signal enhancement using beamforming and nonstationarity with applications to speech.
IEEE Transactions on Signal Processing, 49(8):1614–1626.

Gannot, S. and Cohen, I. (2008).
Springer Handbook of Speech Processing and Speech Communication, chapter Adaptive beamforming and postfiltering.
In [Benesty et al., 2008b].

Gay, S. L. and Benesty, J., editors (2000).
Acoustic signal processing for telecommunication.
Kluwer Academic.

Habets, E. and Gannot, S. (2007).
Generating sensor signals in isotropic noise fields.
The Journal of the Acoustical Society of America, 122:3464–3470.

Habets, E. A. P. (2006).
Room impulse response (RIR) generator.
http://home.tiscali.nl/ehabets/rir_generator.html.

References and Further Reading IV

Haykin, S. and Liu, K. R., editors (2010).
Handbook on array processing and sensor networks, volume 63.
Wiley-IEEE Press.

Herbordt, W. (2005).
Sound capture for human/machine interfaces - Practical aspects of microphone array signal processing, volume 315 of *Lecture Notes in Control and Information Sciences*.
Springer, Heidelberg, Germany.

Hoshuyama, O., Sugiyama, A., and Hirano, A. (1999).
A robust adaptive beamformer for microphone arrays with a blocking matrix using constrained adaptive filters.
IEEE trans. on Signal Proc., 47(10):2677–2684.

Jan, E. and Flanagan, J. (1996).
Sound capture from spatial volumes: Matched-filter processing of microphone arrays having randomly-distributed sensors.
In *IEEE Int. Conf. Acoust. Speech and Sig. Proc. (ICASSP)*, pages 917–920, Atlanta, Georgia, USA.

Jot, J.-M., Cerveau, L., and Warusfel, O. (1997).
Analysis and synthesis of room reverberation based on a statistical time-frequency model.
In *Audio Engineering Society Convention 103*. Audio Engineering Society.

Kaneda, Y. and Ohga, J. (1986).
Adaptive microphone-array system for noise reduction.
IEEE Trans. Acoustics, Speech, and Signal Processing, 34(6):1391–1400.

References and Further Reading V

Makino, S., Lee, T.-W., and Sawada, H. (2007).
Blind speech separation.
Springer Heidelberg.

Markovich, S., Gannot, S., and Cohen, I. (2009).
Multichannel eigenspace beamforming in a reverberant noisy environment with multiple interfering speech signals.
IEEE Transactions on Audio, Speech, and Language Processing, 17(6):1071–1086.

Markovich-Golan, S., Gannot, S., and Cohen, I. (2010).
Subspace tracking of multiple sources and its application to speakers extraction.
In *The IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP)*, pages 201–204, Dallas, Texas, USA.

Markovich-Golan, S., Gannot, S., and Cohen, I. (2012).
A weighted multichannel Wiener filter for multiple sources scenarios.
In *The IEEE 27th Convention of IEEE Israel (IEEEI)*, Eilat, Israel.
best student paper award.

Nordholm, S., Claesson, I., and Bengtsson, B. (1993).
Adaptive Array Noise Suppression of Handsfree Speaker Input in Cars.
IEEE trans. on Vehicular tech., 42(4):514–518.

Polack, J.-D. (1993).
Playing billiards in the concert hall: The mathematical foundations of geometrical room acoustics.
Applied Acoustics, 38(2):235–244.

References and Further Reading VI

Reindl, K., Markovich-Golan, S., Barfuss, H., Gannot, S., and Kellermann, W. (2013).
Geometrically constrained TRINICON-based relative transfer function estimation in underdetermined scenarios.
In *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, New Paltz, USA.

Reuven, G., Gannot, S., and Cohen, I. (2008).
Dual-source transfer-function generalized sidelobe canceller.
IEEE Transactions on Audio, Speech, and Language Processing, 16(4):711–727.

Shalvi, O. and Weinstein, E. (1996).
System identification using nonstationary signals.
IEEE Trans. Signal Processing, 44(8):2055–2063.

Sondhi, M. and Elko, G. (1986).
Adaptive optimization of microphone arrays under a nonlinear constraint.
In *IEEE Int. Conf. Acoust. Speech and Sig. Proc. (ICASSP)*, volume 11, pages 981–984, Tokyo, Japan.

Spriet, A., Moonen, M., and Wouters, J. (2004).
Spatially pre-processed speech distortion weighted multi-channel wiener filtering for noise reduction.
Signal Processing, 84(12):2367–2387.

Van Compernolle, D. (1990).
Switching adaptive filters for enhancing noisy and reverberant speech from microphone array recordings.
In *Proc. Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, pages 833–836, Albuquerque, New Mexico, USA. IEEE.

Van Trees, H. L. (2002).
Detection, Estimation, and Modulation Theory, volume IV, Optimum Array Processing.
Wiley, New York.

References and Further Reading VII

Van Veen, B. D. and Buckley, K. M. (1988).
Beamforming: A versatile approach to spatial filtering.
IEEE Acoustics, Speech and Signal Proc. magazine, pages 4–24.

Wang, D. and Brown, G. J. (2006).
Computational auditory scene analysis: Principles, algorithms, and applications.
Wiley interscience.

Warsitz, E. and Haeb-Umbach, M. (2007).
Blind acoustic beamforming based on generalized eigenvalue decomposition.
IEEE Transactions on Audio, Speech, and Language Processing, 15(5):1529–1539.

Yang, B. (1995).
Projection Approximation Subspace Tracking.
IEEE transactions on Signal Processing, 43(1):95–107.