INTERNATIONAL AUDIO LABORATORIES ERLANGEN
A joint institution of Fraunhofer IIS and Universität Erlangen-Nürnberg

AUDIO LABS

Fifty Years of Reverberation Reduction

From analog signal processing to machine learning

Emanuël Habets

AES 60th Conference on DREAMS

FAU FRIEDRICH-ALEXANDER UNIVERSITÄT ERLANGEN-NÜRNBERG

Fraunhofer IIS

Acknowledgment

Jacob Benesty

Israel Cohen

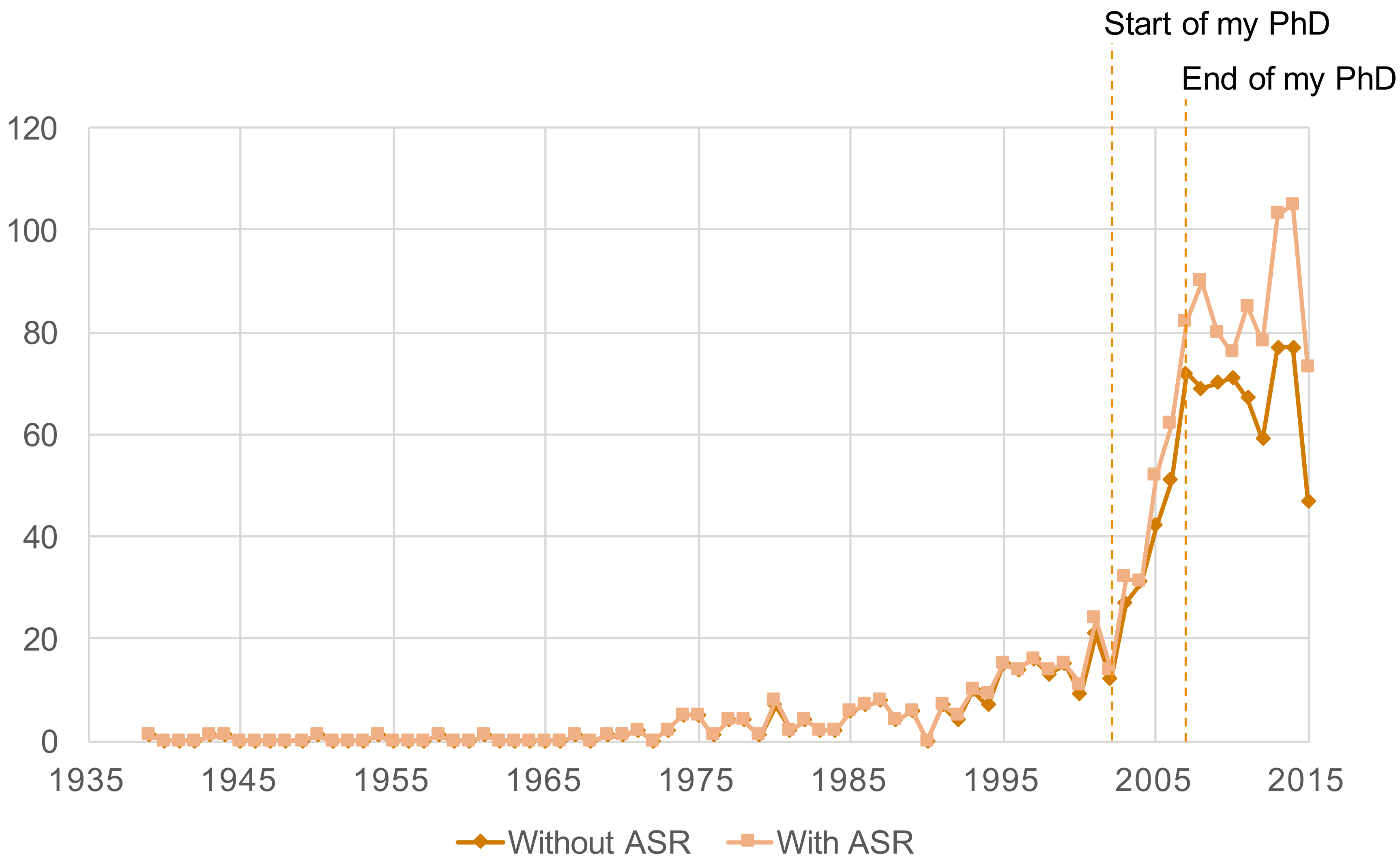
Sharon Gannot

Patrick Naylor

Former and Present PhD Students and Post-Docs at
Bar-Ilan University
Imperial College London
International Audio Laboratories Erlangen

Members of the REVERB 2014 Organization Committee

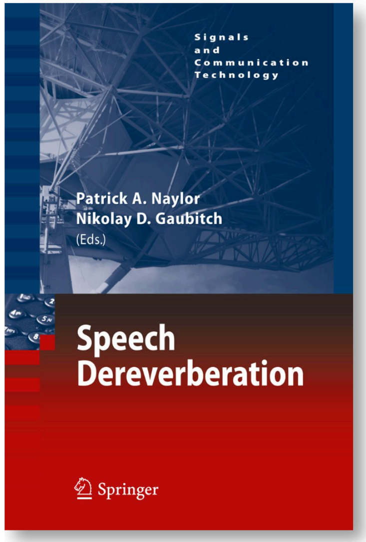
Dereverberation Related Publications[1]



[1]Source: Scopus

Interest

- Book
  - "Speech Dereverberation" by Naylor and Gaubitch (Eds.), 2010
- Challenges (selected)
  - **RE**verberant **V**oice **E**nhancement and **R**ecognition **B**enchmark 2014
  - **A**coustic **C**haracterization of **E**nvironments 2015
  - **A**utomatic **S**peech recognition **I**n **R**everberant **E**nvironments 2015
- Tutorials (selected)
  - EUSIPCO 2010 by Naylor, Habets, and Evers
  - ICASSP 2012 by Nakatani, Sehr, and Kellermann
  - 54th AES 2014 by Naylor, Moore, and Evers
- EU projects such as SCENIC, DREAMS and EARS



Motivation

1938 Patent "Improvements in Electric Signal Amplifiers Incorporating Voice-operated Devices" by Ryall

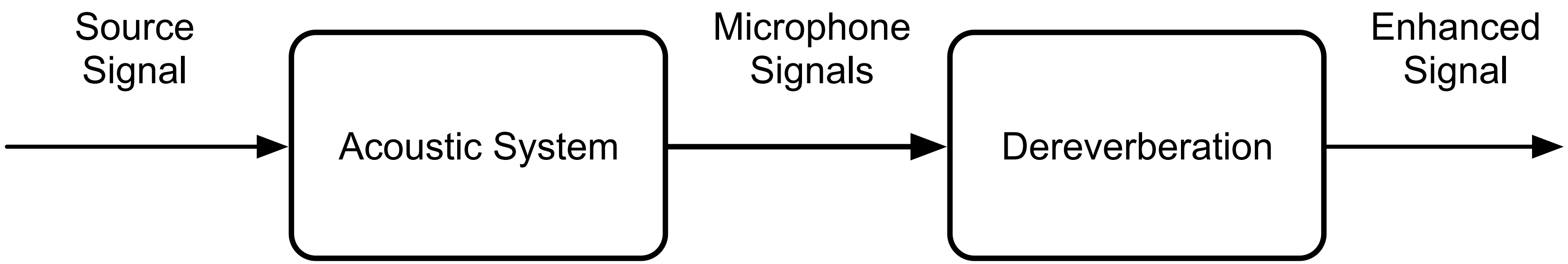
below 600 cycles per second. If the distant end transmitter is a distant speaking microphone, for instance of the moving coil type as is often employed in practice to give the necessary stability, it is preferable in any event to reduce the low frequency signal response in order to prevent the effects of room reverberation from causing discomfort. In order to be certain that the output from the loud-

Applications

- Human-Machine Interfaces
- Hands-free Communication
- Hearing-aids
- Music Post-Production

Dereverberation Approaches

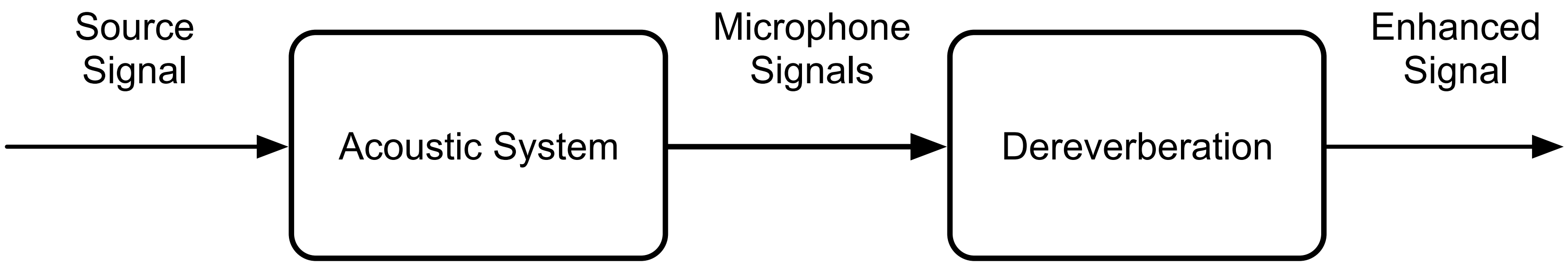
- Three fundamentally different approaches
  1. Model the acoustic system, estimate the model parameters by treading the source signal as a nuisance, and then estimate the source signal
  2. Model the reverberation as an additive process, and then estimate the source signal
  3. Directly estimate the source signal from the microphone signals by treading the acoustic system as unknown



Dereverberation Approaches

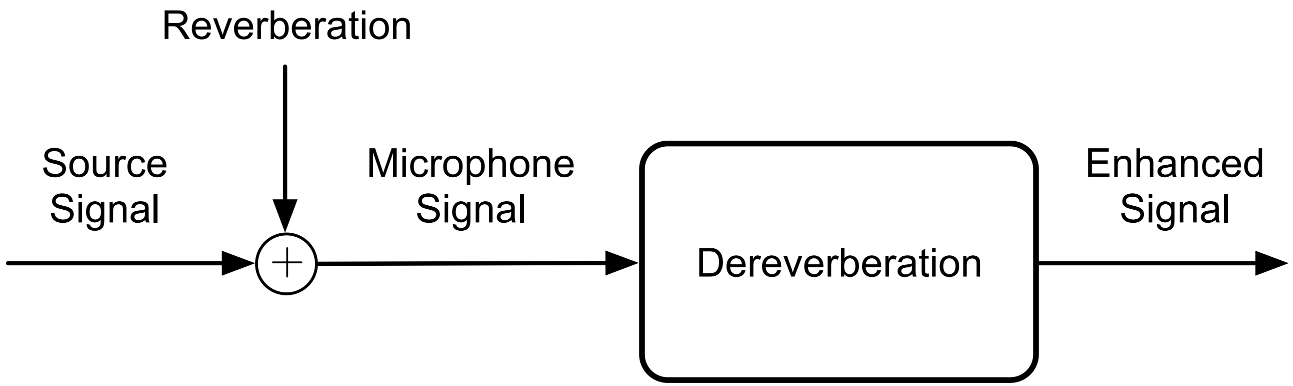
- Three fundamentally different approaches
  1. Model the acoustic system, estimate the model parameters by treading the source signal as a nuisance, and then estimate the source signal
  2. Model the reverberation as an additive process, and then estimate the source signal
  3. Directly estimate the source signal from the microphone signals by treading the acoustic system as unknown

Reverberation Cancellation



Dereverberation Approaches

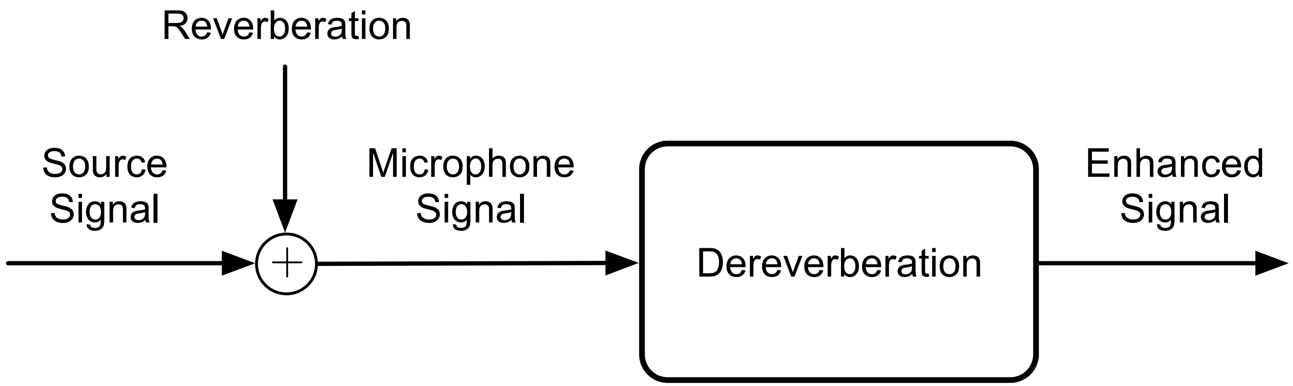
- Three fundamentally different approaches
  1. Model the acoustic system, estimate the model parameters by treading the source signal as a nuisance, and then estimate the source signal
  2. Model the reverberation as an additive process, and then estimate the source signal
  3. Directly estimate the source signal from the microphone signals by treading the acoustic system as unknown



Dereverberation Approaches

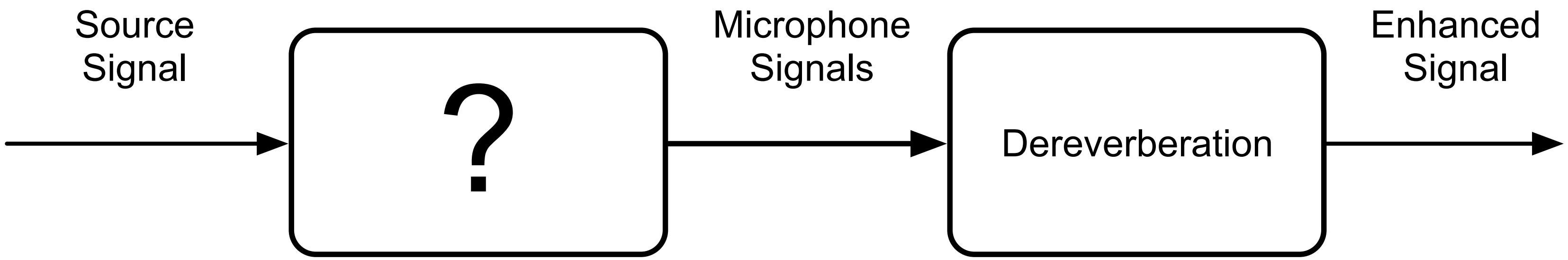
- Three fundamentally different approaches
  1. Model the acoustic system, estimate the model parameters by treading the source signal as a nuisance, and then estimate the source signal
  2. Model the reverberation as an additive process, and then estimate the source signal
  3. Directly estimate the source signal from the microphone signals by treading the acoustic system as unknown

Reverberation Suppression



Dereverberation Approaches

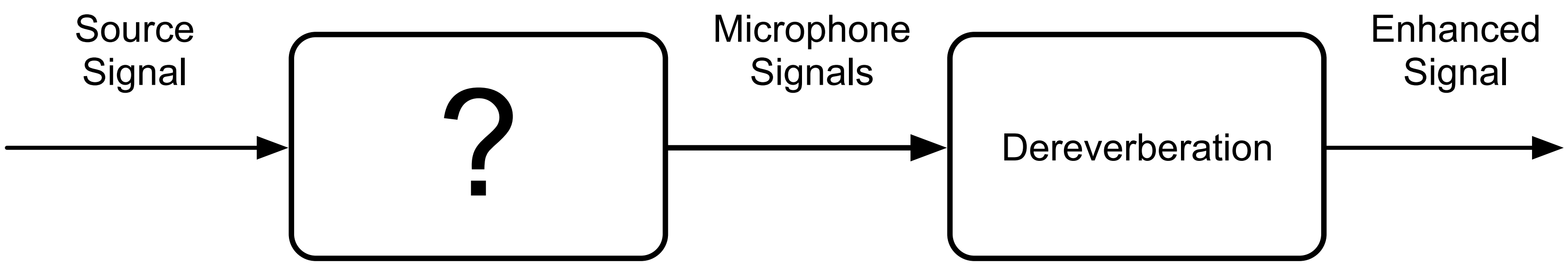
- Three fundamentally different approaches
  1. Model the acoustic system, estimate the model parameters by treading the source signal as a nuisance, and then estimate the source signal
  2. Model the reverberation as an additive process, and then estimate the source signal
  3. Directly estimate the source signal from the microphone signals by treading the acoustic system as unknown



Dereverberation Approaches

- Three fundamentally different approaches
  1. Model the acoustic system, estimate the model parameters by treading the source signal as a nuisance, and then estimate the source signal
  2. Model the reverberation as an additive process, and then estimate the source signal
  3. Directly estimate the source signal from the microphone signals by treading the acoustic system as unknown

Direct Estimation



Outline

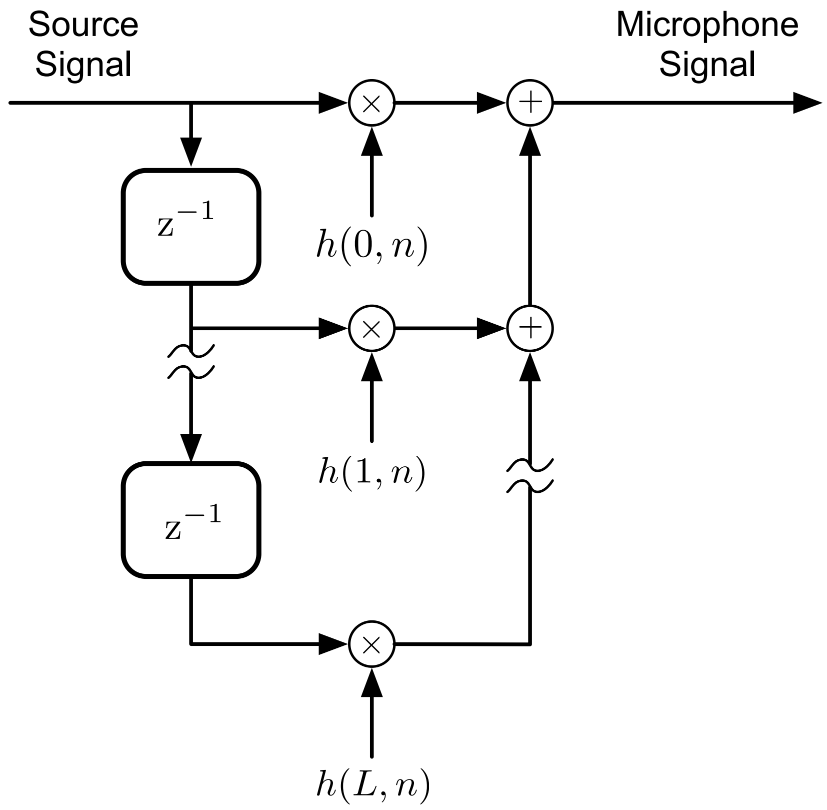
- Reverberation Cancellation
- Reverberation Suppression
- Direct Estimation
- Practical Challenges
- Conclusions and Future Challenges

Outline

- Reverberation Cancellation
- Reverberation Suppression
- Direct Estimation
- Practical Challenges
- Conclusions and Future Challenges

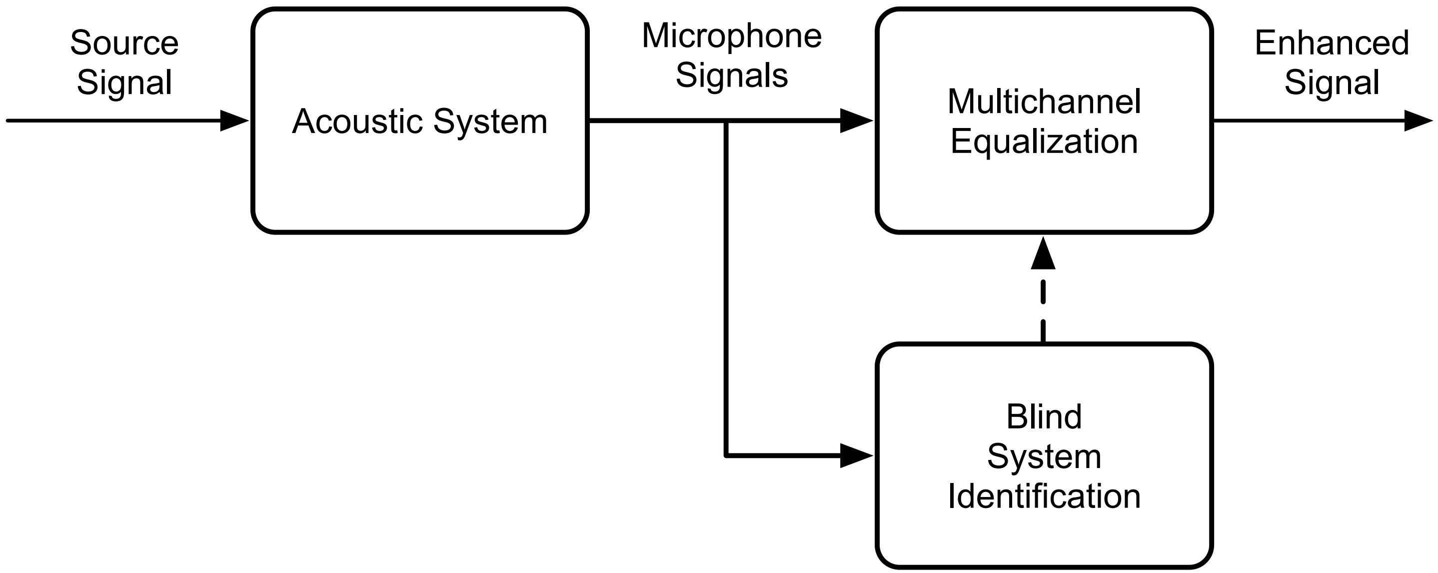
Reverberation Cancellation
Moving Average Process (Time-Domain)

- The acoustic channels are modeled as finite impulse responses
- The reverberation signal equals a moving average process



Reverberation Cancellation
Moving Average Process (Time-Domain)

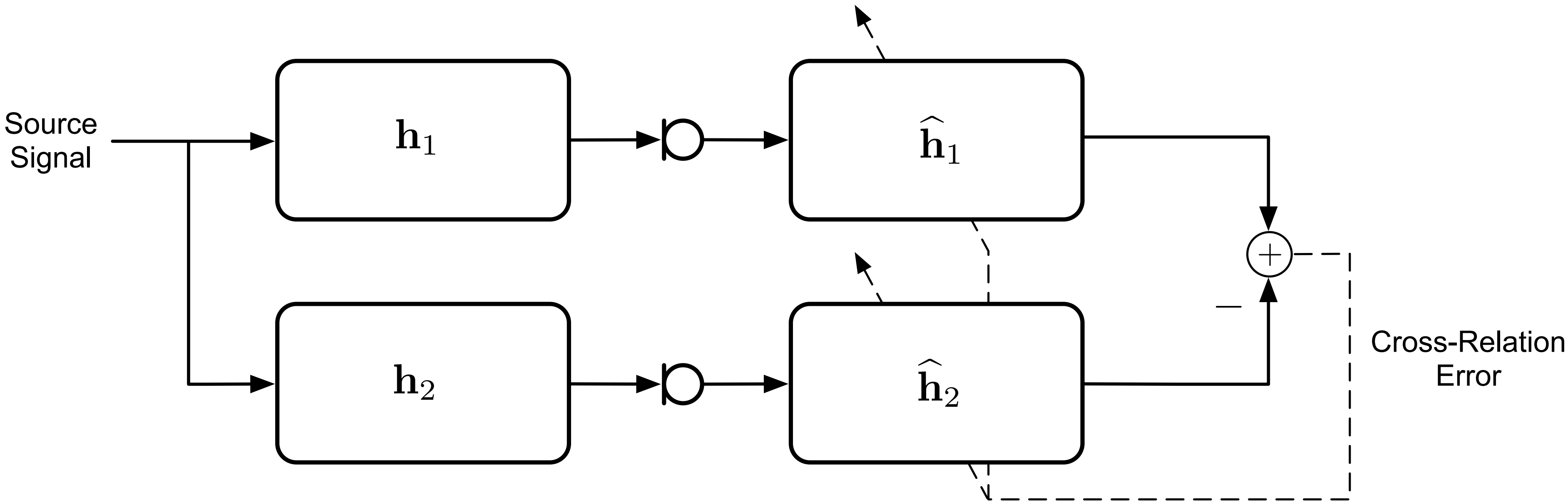
- The desired signal is a delayed or filtered version of the source signal
- To obtain an estimate of the desired signal:
  1. Blindly identify the model parameters of the acoustic system
  2. Estimate the desired signal by applying a multichannel equalizer



Reverberation Cancellation
Moving Average Process (Time-Domain)

- Identification approaches for two or more microphones
  1. From the null space of the microphone signals' correlation matrix
     (Gürelli and Nikias, 1995) and (Gannot and Moonen, 2003)

  2. By minimizing the cross-relation error (Xu et al. 1995)



Reverberation Cancellation
Moving Average Process (Time-Domain)

- Identification approaches for two or more microphones
  1. From the null space of the microphone signals' correlation matrix (Gürelli and Nikias, 1995) and (Gannot and Moonen, 2001)

  2. By minimizing the cross-relation error (Xu et al. 1995)
     - MC LMS (Huang and Benesty, 2002)
     - MC Newton (Huang and Benesty, 2002)
     - Normalized MC Frequency-Domain LMS (Huang and Benesty, 2003)
     - Noise Robust MC Frequency-Domain LMS (Haque and Hasan, 2008)
     - MC Quasi-Newton (Habets and Naylor, 2010)
     - State-space Frequency-Domain Adaptive Filter (Malik et al. 2012)
     - …

Reverberation Cancellation
Moving Average Process (Time-Domain)

- Multichannel equalization
  - MINT (Miyoshi and Kaneda, 1988)
  - Matched Filtering (Flanagan et al., 1993)
  - Channel Shortening (Kallinger and Mertins, 2006)
  - Weighted Least Squares (Zhang and Naylor, 2008)
  - Relaxed Multichannel Least Squares (Zhang et al., 2010)
  - Infinity- and p-Norm Optimization (Mertins et al., 2010)
  - RMCLS with Constrained Initial Taps (Lim and Naylor, 2012)
  - Partial MINT (Kodrasi and Doclo, 2012a)
  - Regularized Partial Multichannel Equalization (Kodrasi and Doclo, 2012b)
  - Perceptually Constrained Channel Shortening (Kodrasi et al., 2013)
  - …

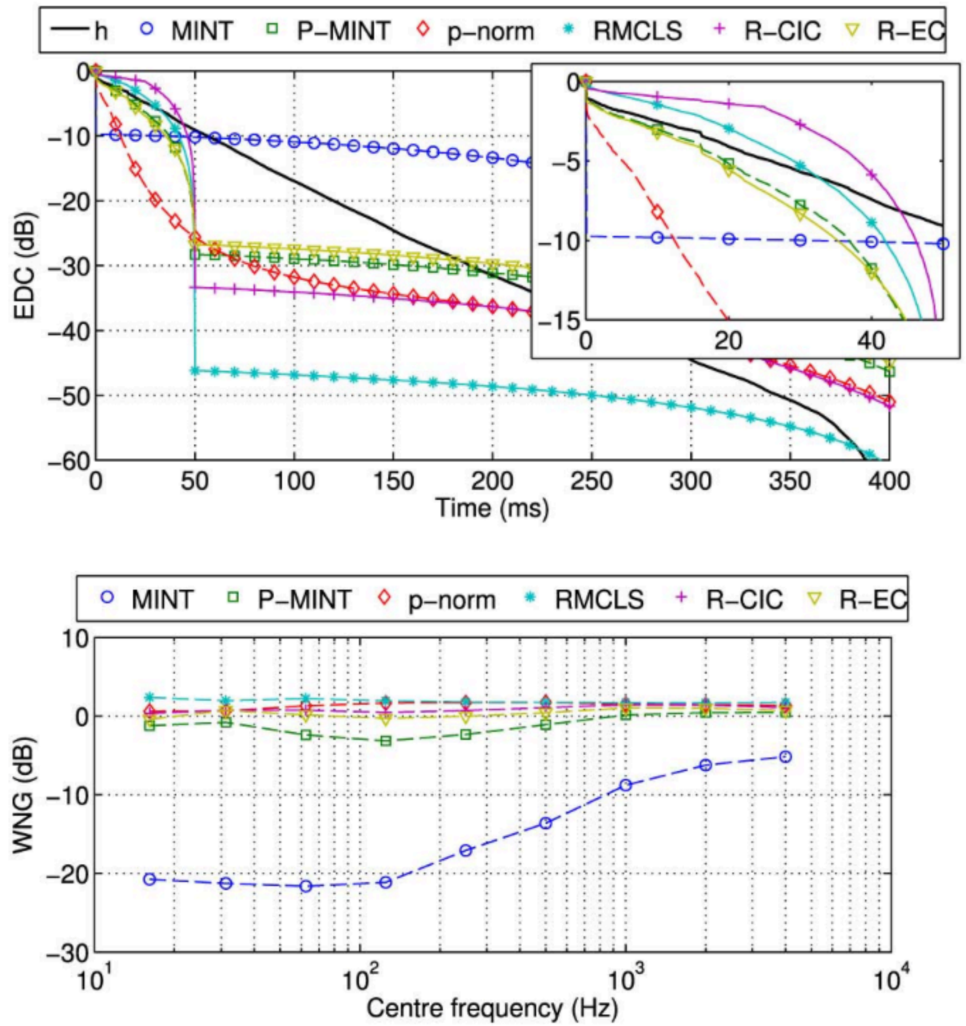
Reverberation Cancellation
Moving Average Process (Time-Domain)

- Example (Lim et al., 2014)
  - Normalized projection misalignment was -33 dB
  - Reverberation time was 0.4 seconds
  - Equalized impulse response

$$\mathrm{EIR}(n) = \sum_{m=1}^{M} h_m(n) * g_m(n)$$

  - Energy decay curve

$$\mathrm{EDC}(n) = \sum_{i=n}^{L} \mathrm{EIR}^2(i)$$



Reverberation Cancellation

Moving Average Process (Time-Domain)

- Opportunities
  - Ability to control the equalized impulse response
  - Localization in reverberant environments

- Challenges / Open Issues
  - Both identification and equalization approaches assume there are no common-zeros across the channels. These common-zeros can lead to undesired signal coloration.
  - Robustness against estimation errors
  - High and unknown channel order

Reverberation Cancellation
Moving Average Process (TF-Domain)

- In (B. Schwartz et al., 2015) the microphone signals were modeled in the STFT domain as a moving average process

$$X_m(n,k) = \sum_{\ell=0}^{L} H_m(\ell,k)\, S(n-\ell,k)$$

AUDIO LABS

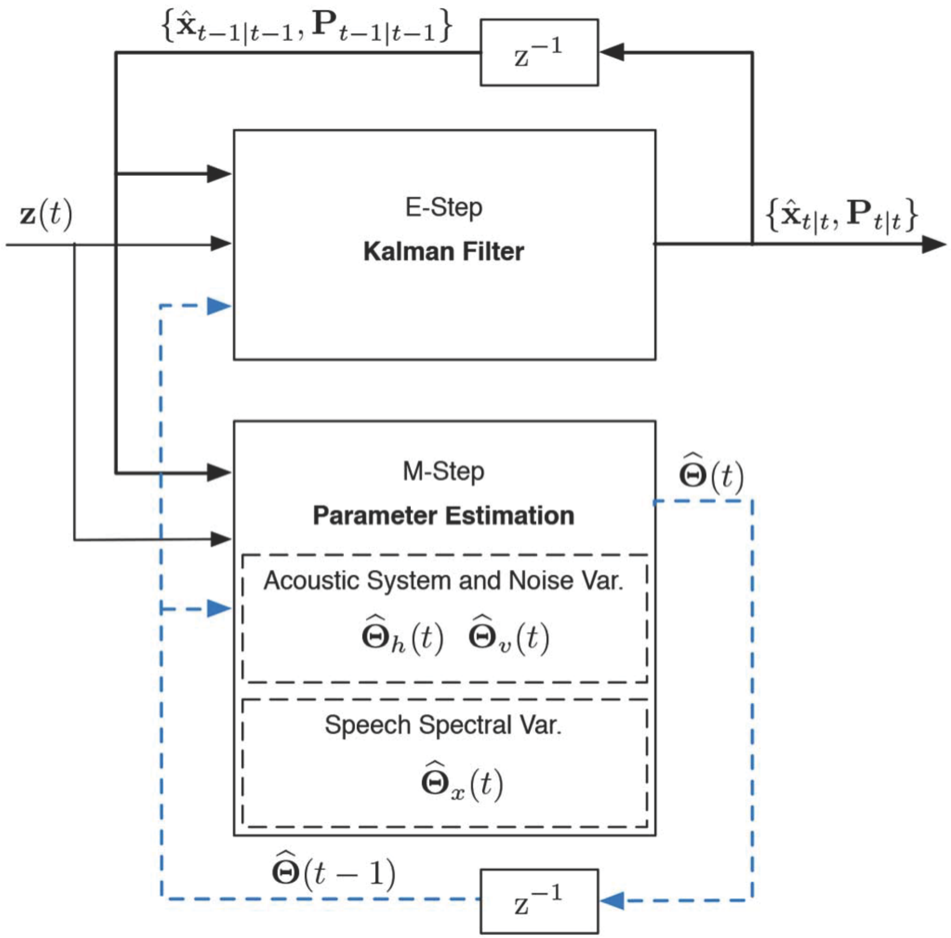
Reverberation Cancellation
Moving Average Process (TF-Domain)

- In (B. Schwartz et al., 2015) the microphone signals were modeled in the STFT domain as a moving average process

$$X_m(n,k) = H_m(0,k)\, S(n,k) + \sum_{\ell=1}^{L} H_m(\ell,k)\, S(n-\ell,k)$$

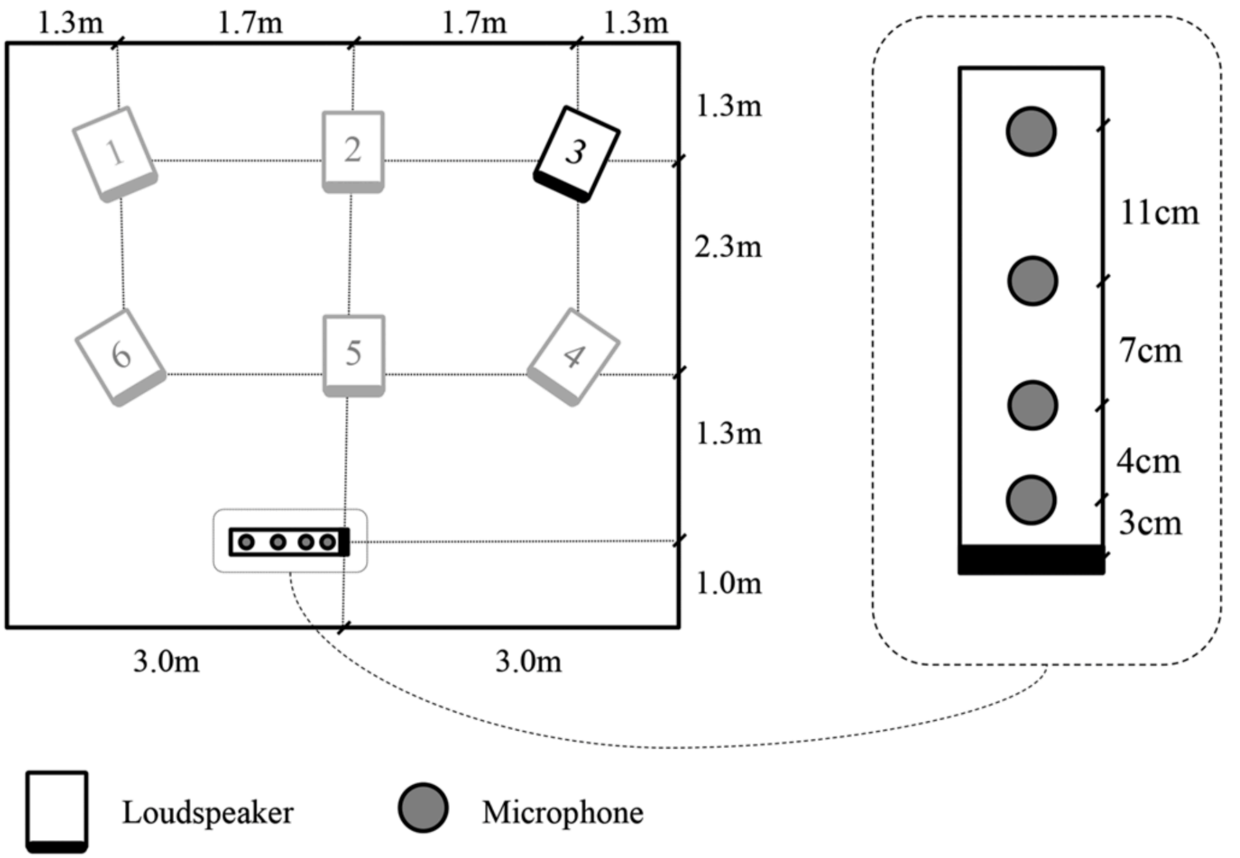
Reverberation Cancellation
Moving Average Process (TF-Domain)

- A recursive expectation-maximization scheme is used to estimate the model parameters of the system, speech, and noise online
- In the E-Step, a Kalman filter is used to estimate the desired speech signal (and the error covariance matrix)



Source: (B. Schwartz et al., 2015)

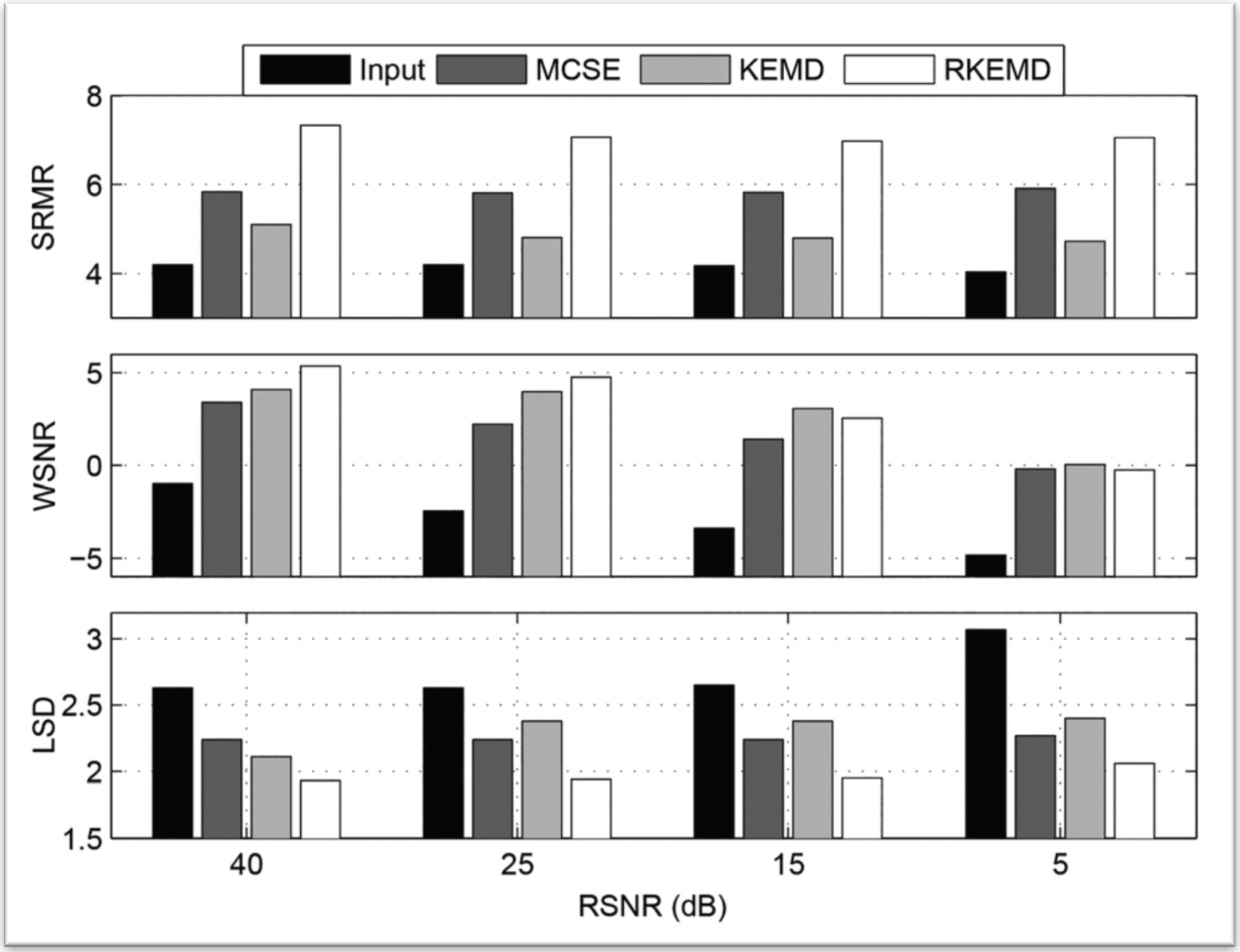
Reverberation Cancellation
Moving Average Process (TF-Domain)



Source: (B. Schwartz et al., 2015)

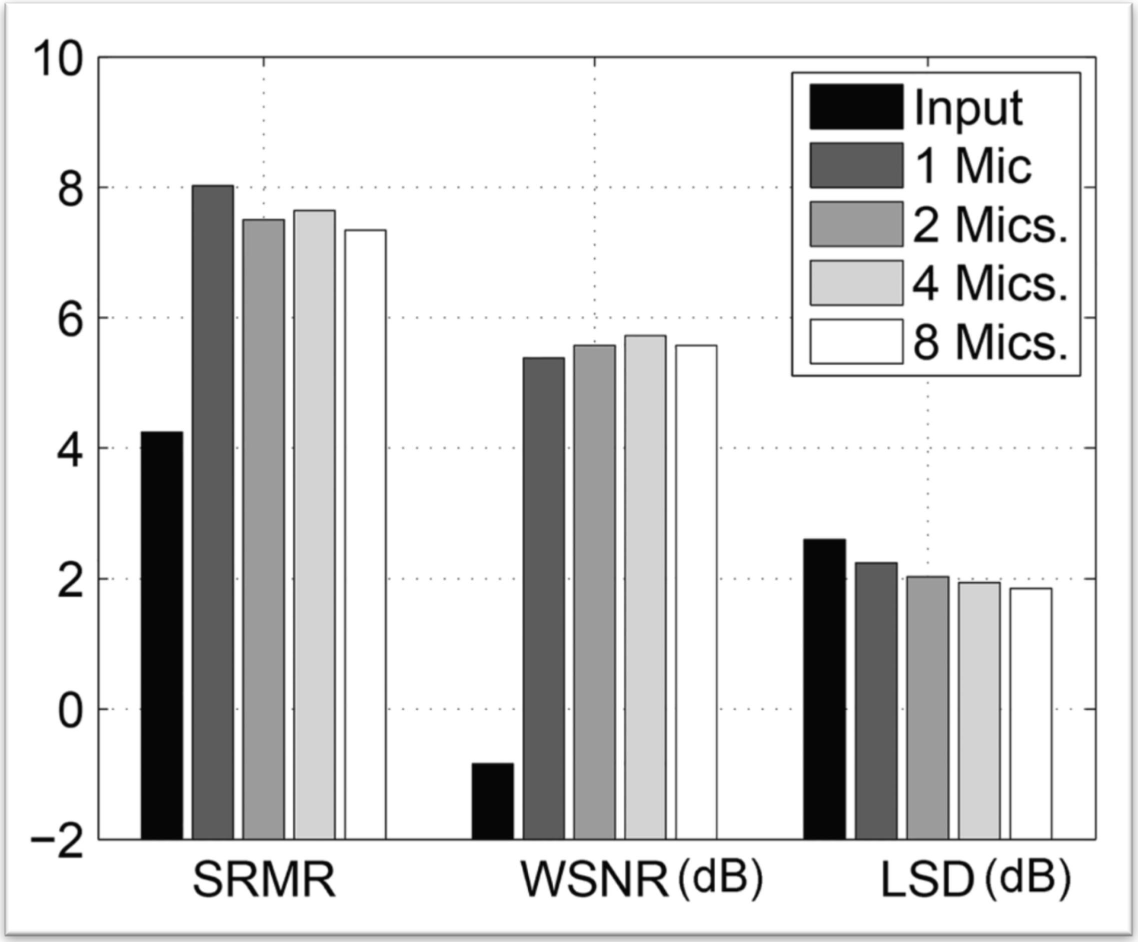
Reverberation Cancellation

Moving Average Process (TF-Domain)



Source: (B. Schwartz et al., 2015)

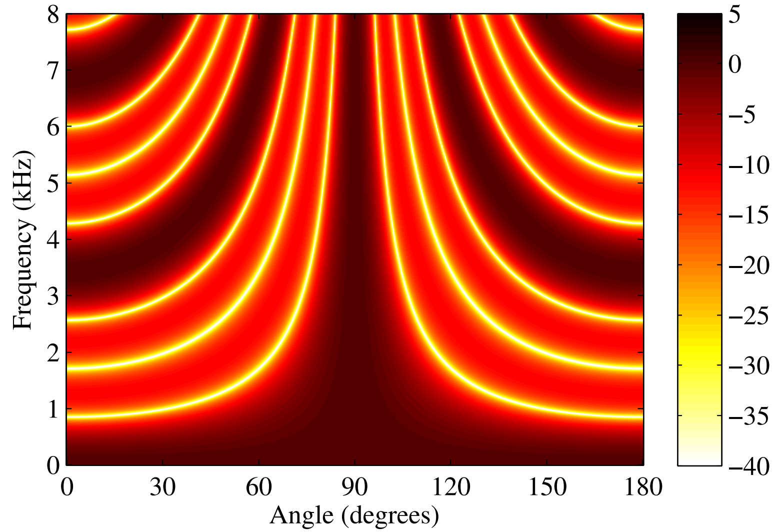
Reverberation Cancellation
Moving Average Process (TF-Domain)

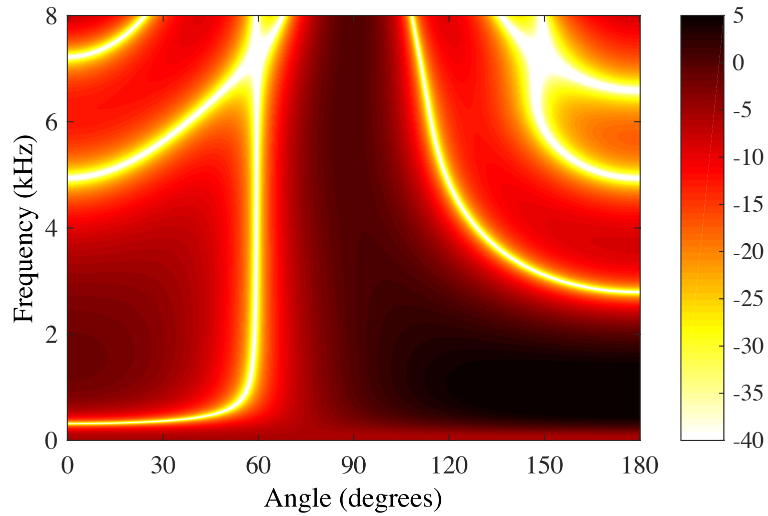


Source: (B. Schwartz et al., 2015)

Reverberation Cancellation
Spatial Filtering

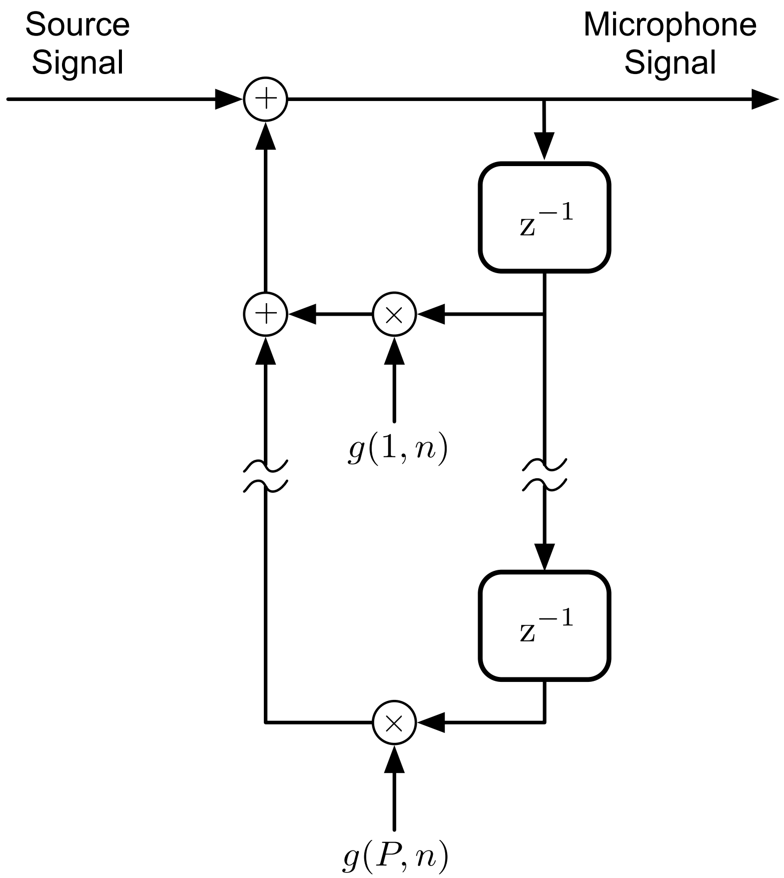
- The direction-of-arrival (DOA) of most early reflections differ from that of the direct path
- These early reflections can be reduced using a spatial filter
- When available, the DOAs of the early reflections can be added as additional constraints (Peled and Rafaely, 2013)



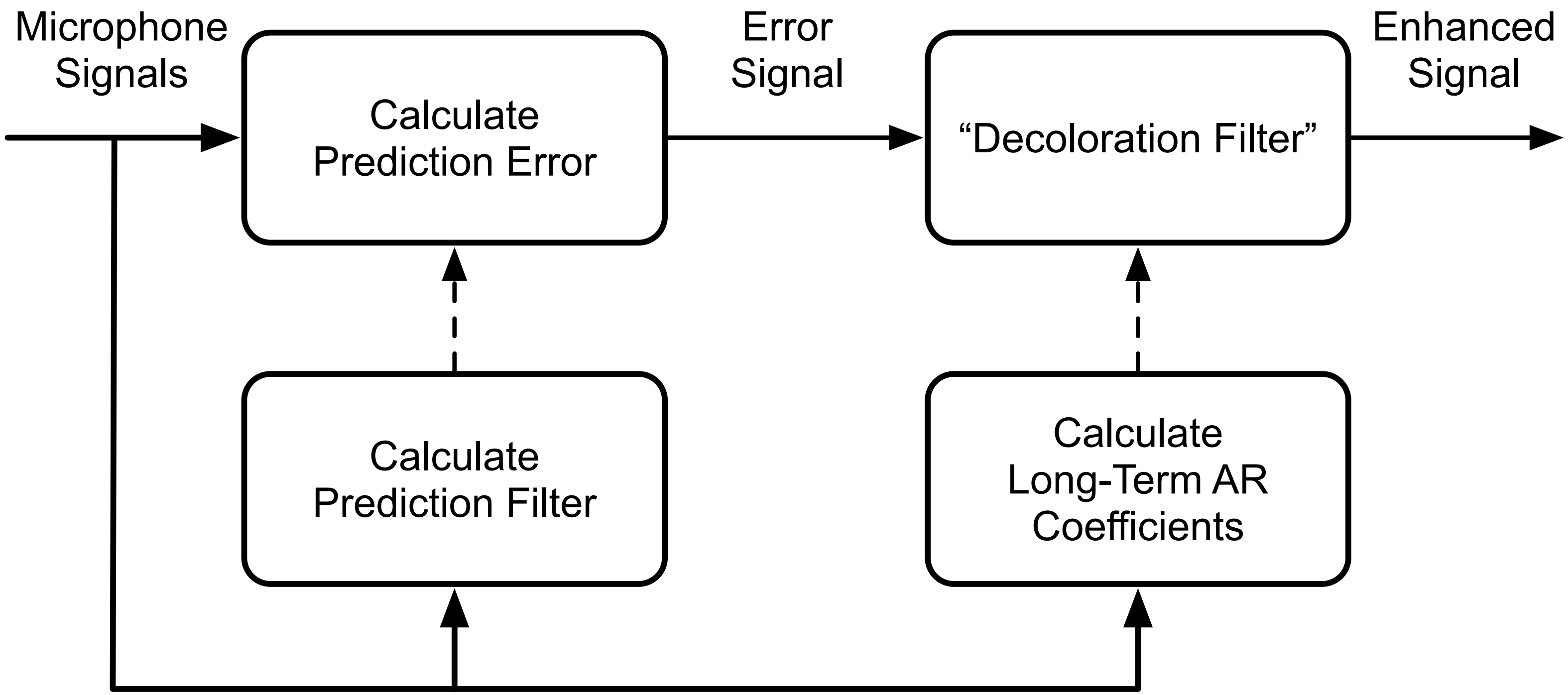


Reverberation Cancellation
Autoregressive Process

- The acoustic channels are modeled as infinite impulse responses
- The reverberation signal equals an autoregressive process

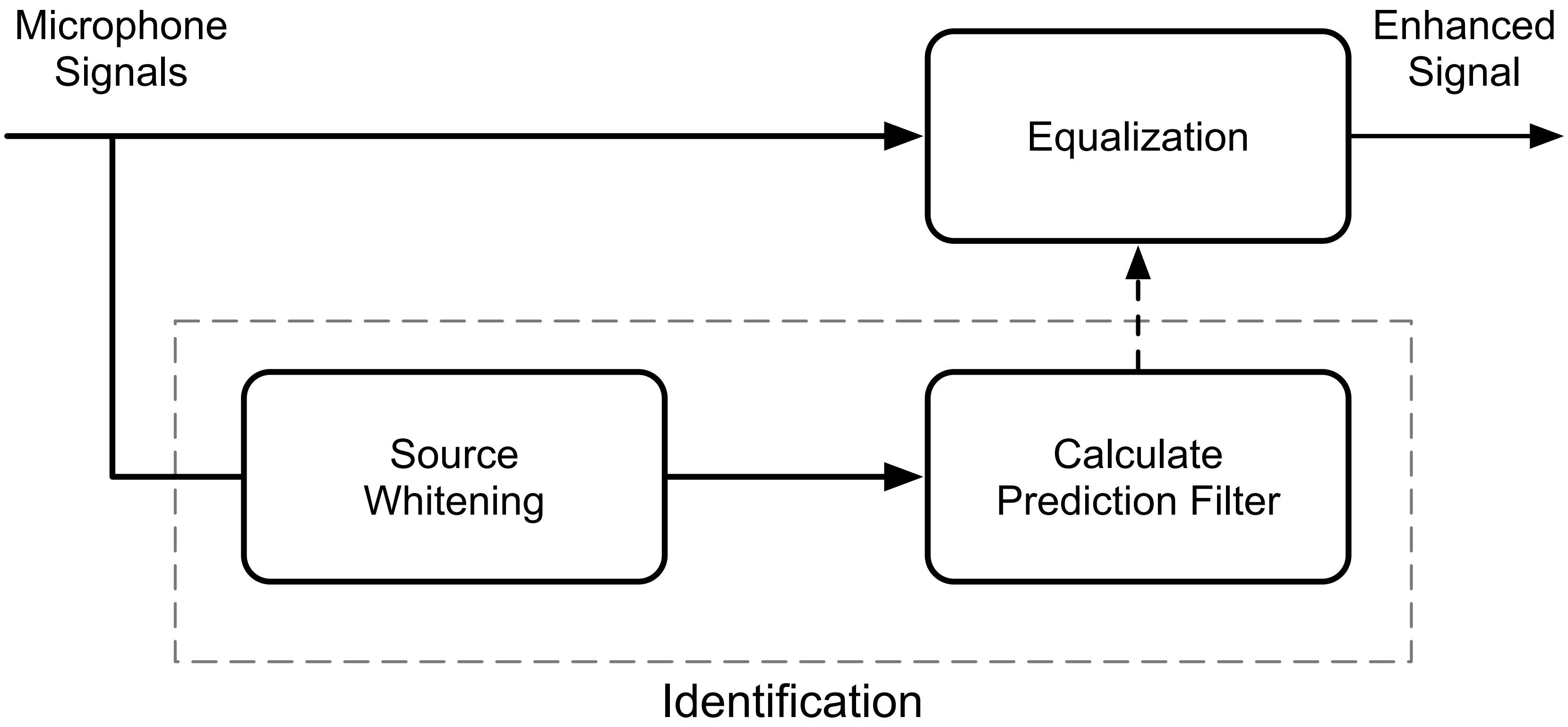


Reverberation Cancellation
Autoregressive Process (Time-Domain)



(Delcroix et al., 2004)

Reverberation Cancellation
Autoregressive Process (Time-Domain)



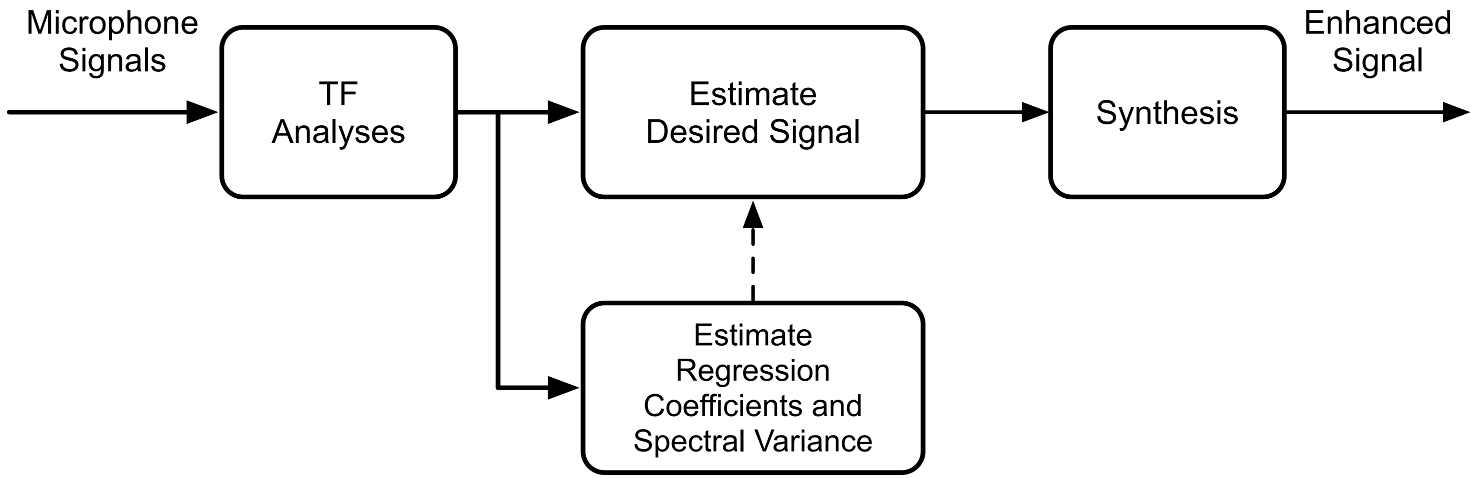
(Triki and Slock, 2005)

Reverberation Cancellation
Autoregressive Process (TF-Domain)

- The reverberation signal at the first microphone is modeled as

$$X_1(n,k) = \sum_{m=1}^{M} \sum_{p=D}^{P} c_m^*(p,k)\, X_m(n-p,k) + D_1(n,k)$$

where $c_m(p,k)$ denotes the *p*-th regression coefficient at the *m*-th microphone, *D* the prediction delay, and *P* the regression order



Reverberation Cancellation
Autoregressive Process (TF-Domain)

- Approaches to estimate the regression coefficients and spectral variance of the desired signal
  - The desired signal is modeled as a time-varying complex Gaussian process. The parameters are iteratively estimated in the least-squares sense (Yoshioka et al., 2009)
  - The parameters are estimated using an Expectation-Maximization scheme (Togami et al., 2013)
  - The desired signal is modeled as a time-varying complex generalized Gaussian process. For a given shape parameter, the parameters are iteratively estimated in the least-squares sense (Jukić et al., 2014)
  - The parameters are estimated by minimizing an auxiliary cost function that measures the inter-frame dependence (Yoshioka et al., 2012)
  - ….

Reverberation Cancellation
Autoregressive Process

- Opportunities
  - Has been successfully applied also in the context of ASR
  - The model holds also for multiple sources

- Challenges / Open Issues
  - To reduce the effect of early reflections
  - To control the desired signal

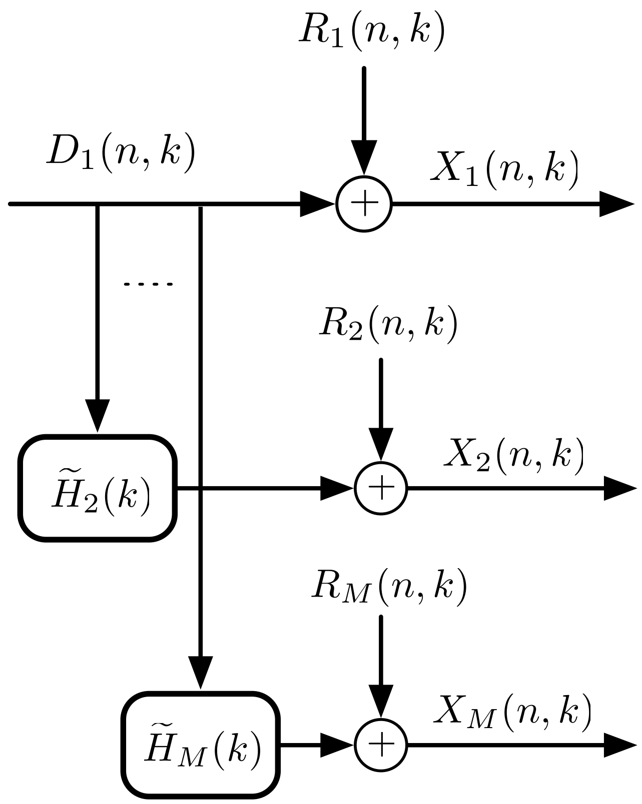
Outline

- Reverberation Cancellation
- Reverberation Suppression
- Direct Estimation
- Practical Challenges
- Conclusions and Future Challenges

Reverberation Suppression

- It is assumed that
  1. reverberation is an additive process
  2. the desired signal and the reverberation signal are uncorrelated
  3. the reverberation signal can be modeled as

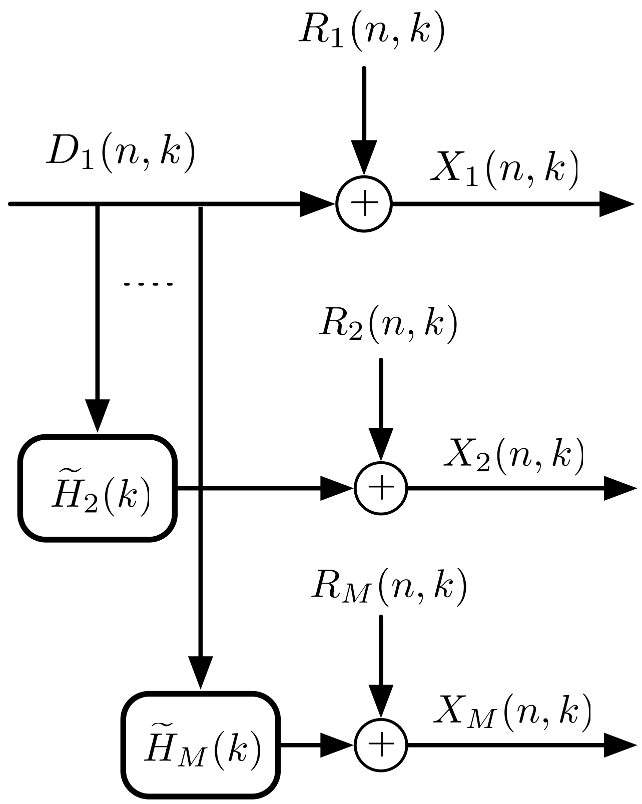
$$\mathbf{r}(n,k) \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \mathbf{\Phi}_{\mathbf{r}}(n,k))$$



Reverberation Suppression

- It is assumed that
  1. reverberation is an additive process
  2. the desired signal and the reverberation signal are uncorrelated
  3. the reverberation signal can be modeled as

$$\mathbf{r}(n,k) \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \phi_R(n,k)\,\boldsymbol{\Gamma}(k))$$



Reverberation Suppression

- From this perspective, we can develop
  - Data-Independent Spatial Filtering Techniques
  - Single-Channel Spectral Enhancement Techniques
  - Multi-Channel Spectral Enhancement Techniques
  - Data-Dependent Spatial Filtering Techniques

Reverberation Suppression
Data-Independent Spatial Filtering

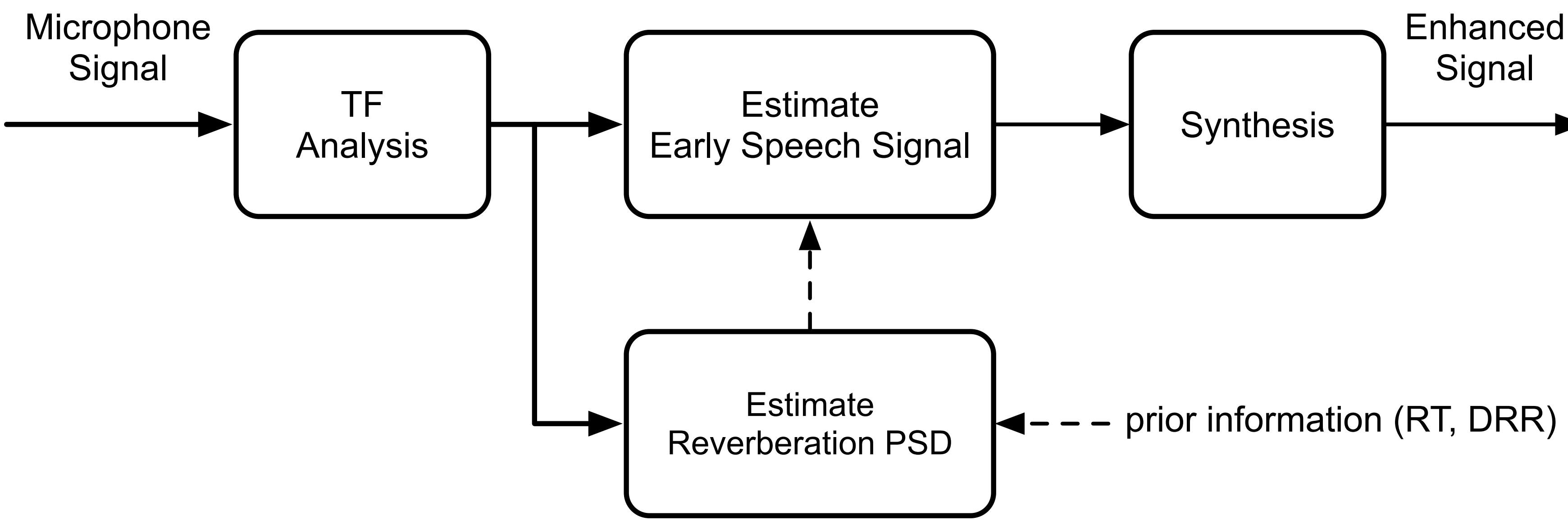
- Apply a spatial filter $\mathbf{w}(k)$ that maximizes the directivity index

$$\mathbf{w}(k) = \underset{\mathbf{w}}{\operatorname{argmax}} \frac{1}{\mathbf{w}^{\mathrm{H}}\boldsymbol{\Gamma}(k)\mathbf{w}} \quad \text{subject to} \quad \mathbf{w}^{\mathrm{H}}\mathbf{h}(k) = 1$$

- This leads to the well-known super-directive beamformer
- The reverberation reduction strongly depends on the number of microphones, array geometry, and the DOA of the source
- At high frequencies the SDR increases by approximately 3 dB when doubling the number of microphones

Reverberation Suppression
Single-Channel Spectral Enhancement

- Single-channel spectral enhancement techniques commonly require an estimate of the clean speech PSD and the interference PSD
- Statistical models for the acoustic channel can be used to derive estimators for the late reverberation PSD

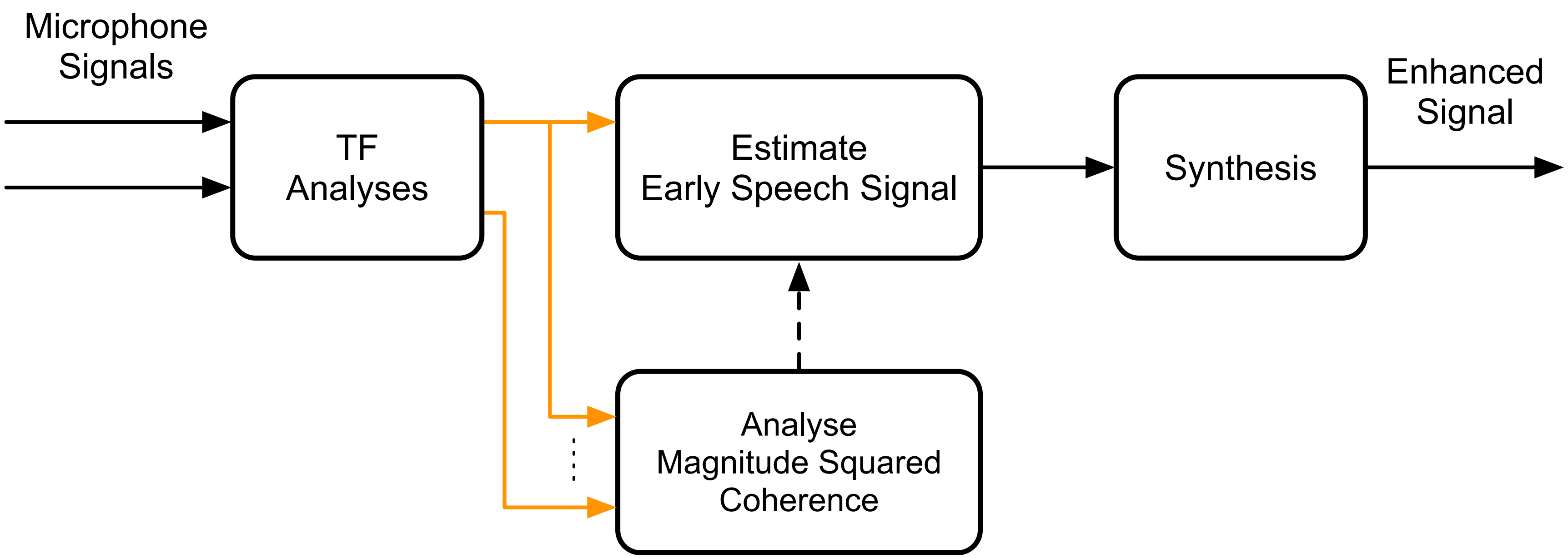


Reverberation Suppression
Single-Channel Spectral Enhancement

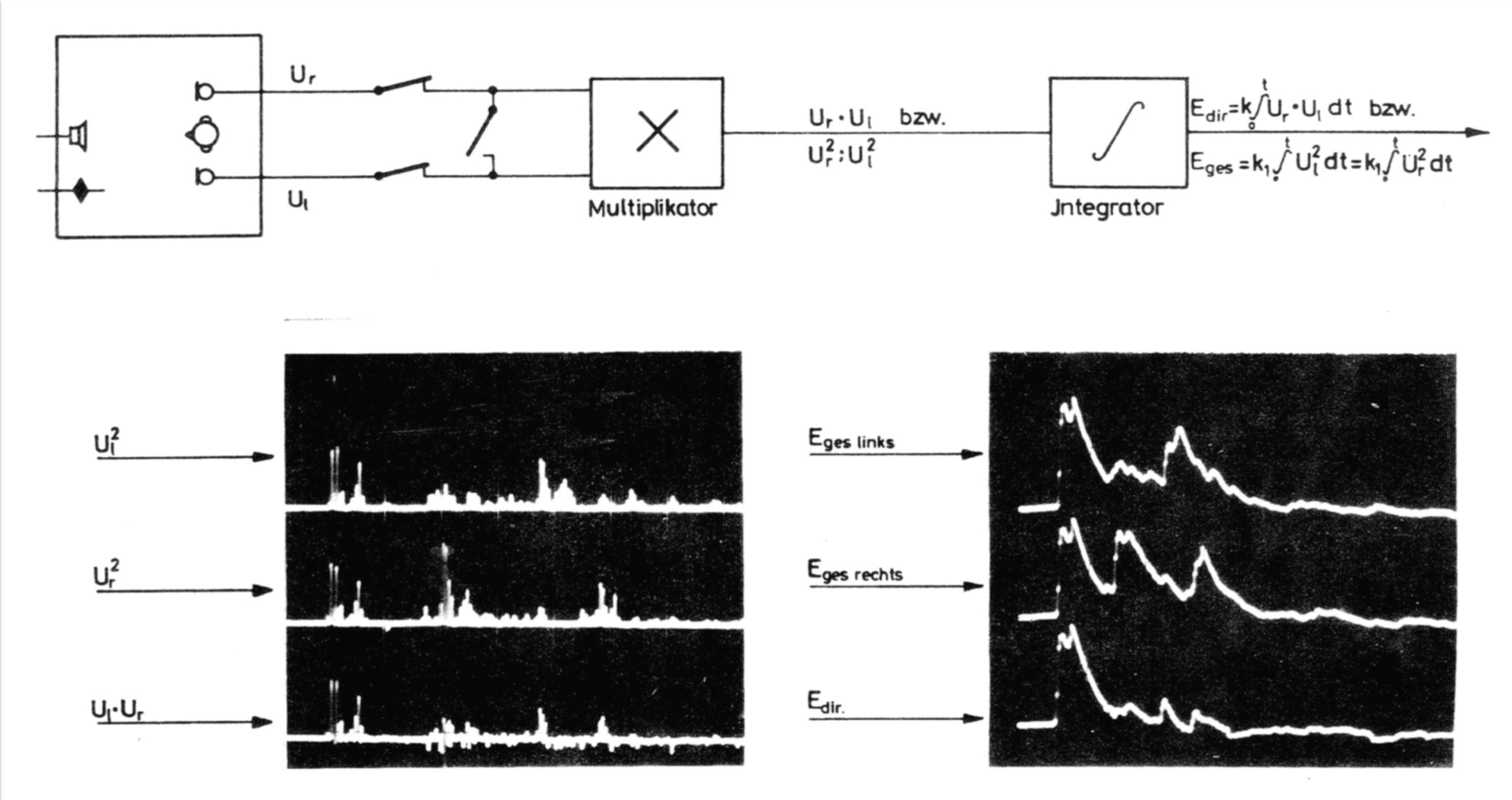
- Selected approaches to estimate the late reverberation PSD:
  - Lebart et al. (2001) used Moorer's model and a frequency independent reverberation time (RT)
  - Habets (2004) used Polack's model and a frequency dependent RT
  - Habets et al. (2007/2009) proposed a generalized statistical model that depends on the direct-to-reverberation ratio (DRR) and RT
  - Erkelens et al. (2010) proposed a correlation-based PSD estimator
- This let to new challenges such as blindly estimating the DRR and RT which were also part of the recent ACE 2015 Challenge

Reverberation Suppression
Multi-Channel Spectral Enhancement

- In early works, a sub-band gain was used that was computed based on the magnitude squared coherence between two microphones
- An analog system was proposed by Danilenko (1968), and a digital system was later proposed by Allen et al. (1977)



Reverberation Suppression
Multi-Channel Spectral Enhancement



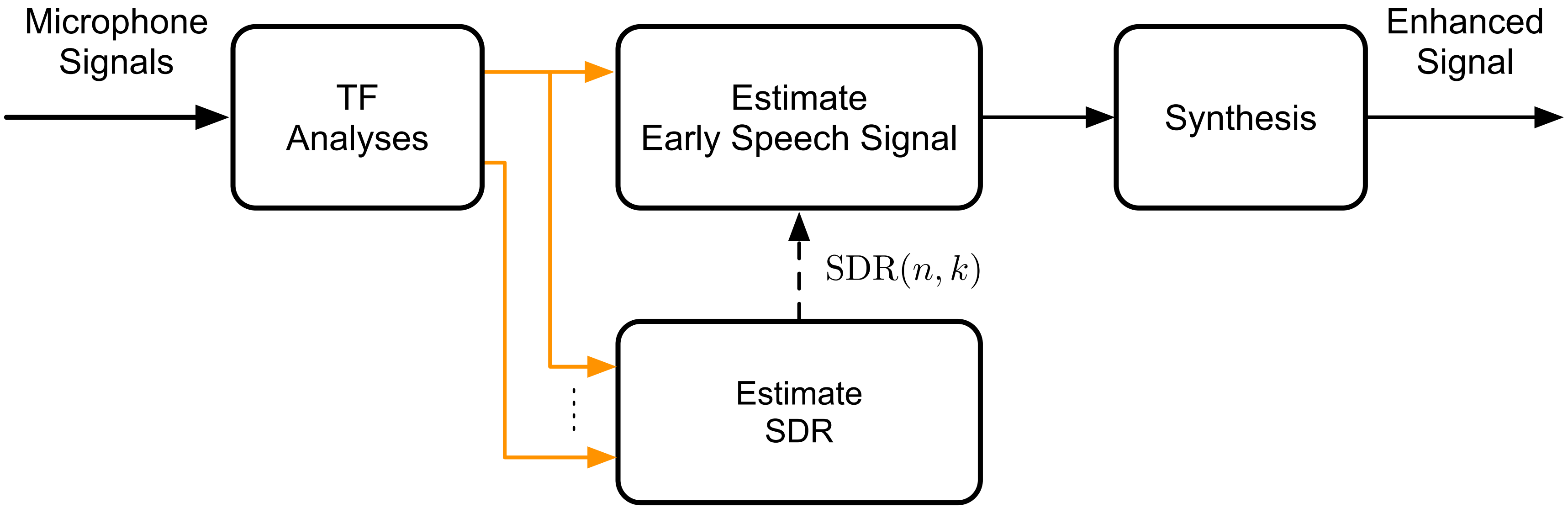
Source: L. Danilenko, “Binaural hearing in non-stationary diffuse sound field,” Dissertation, RWTH Aachen University, 1968.

Reverberation Suppression
Multi-Channel Spectral Enhancement

- Alternatively, multiple microphones can be used to estimate the signal-to-diffuse ratio (SDR) that is defined as

$$\mathrm{SDR}(n,k) = \frac{\mathrm{E}\left\{|D_1(n,k)|^2\right\}}{\mathrm{E}\left\{|R_1(n,k)|^2\right\}} = \frac{\phi_{D_1}(n,k)}{\phi_{R_1}(n,k)}$$

- The early speech can be estimated using a single-channel Wiener filter

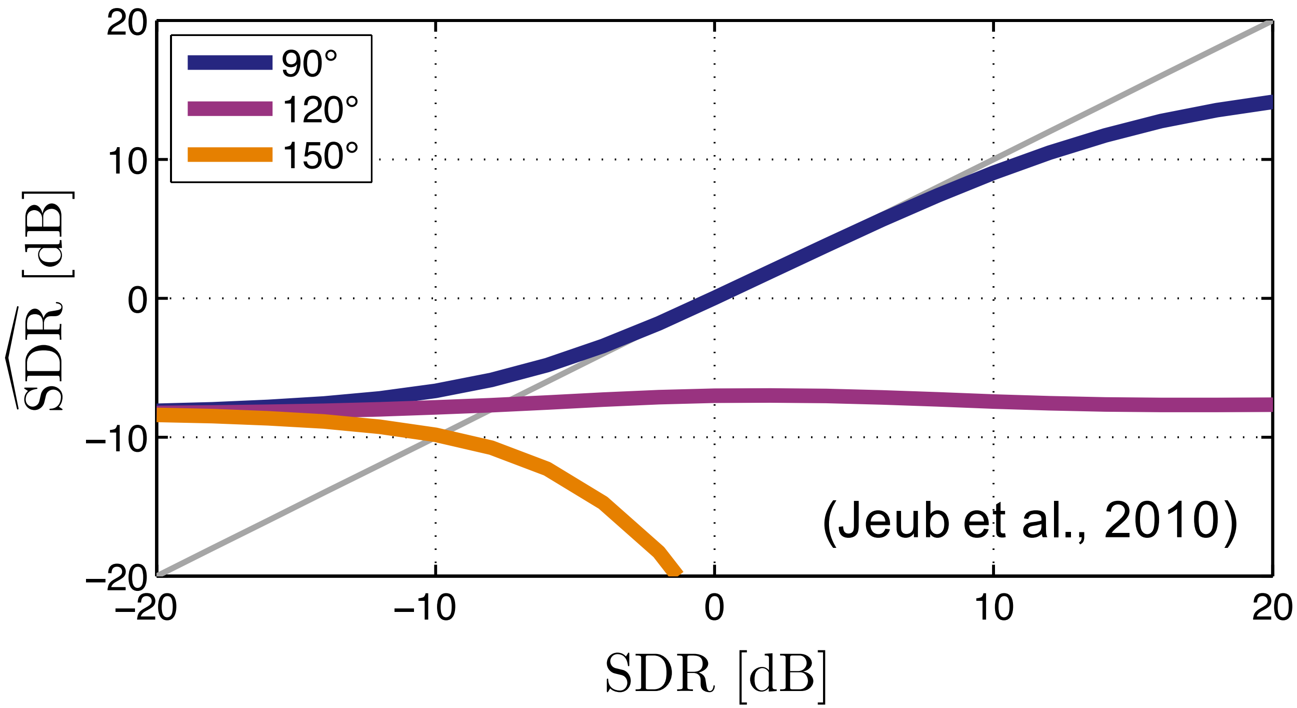


Reverberation Suppression
Multi-Channel Spectral Enhancement

- Selected SDR estimators:
  - B-format microphones (Ahonen and Pulkki, 2009)
  - Omni-directional microphones using the real part of the complex coherence and a source at 90 degrees (Jeub et al. 2010)
  - Omni-directional microphones (complex coherence) (Thiergart et al., 2012a)
  - First-order directional microphones (complex coherence) (Thiergart et al., 2012b)
  - Power-based SDR estimation for directional microphones (Thiergart et al., 2014)
  - Unbiased SDR estimation (Schwarz and Kellermann, 2015)

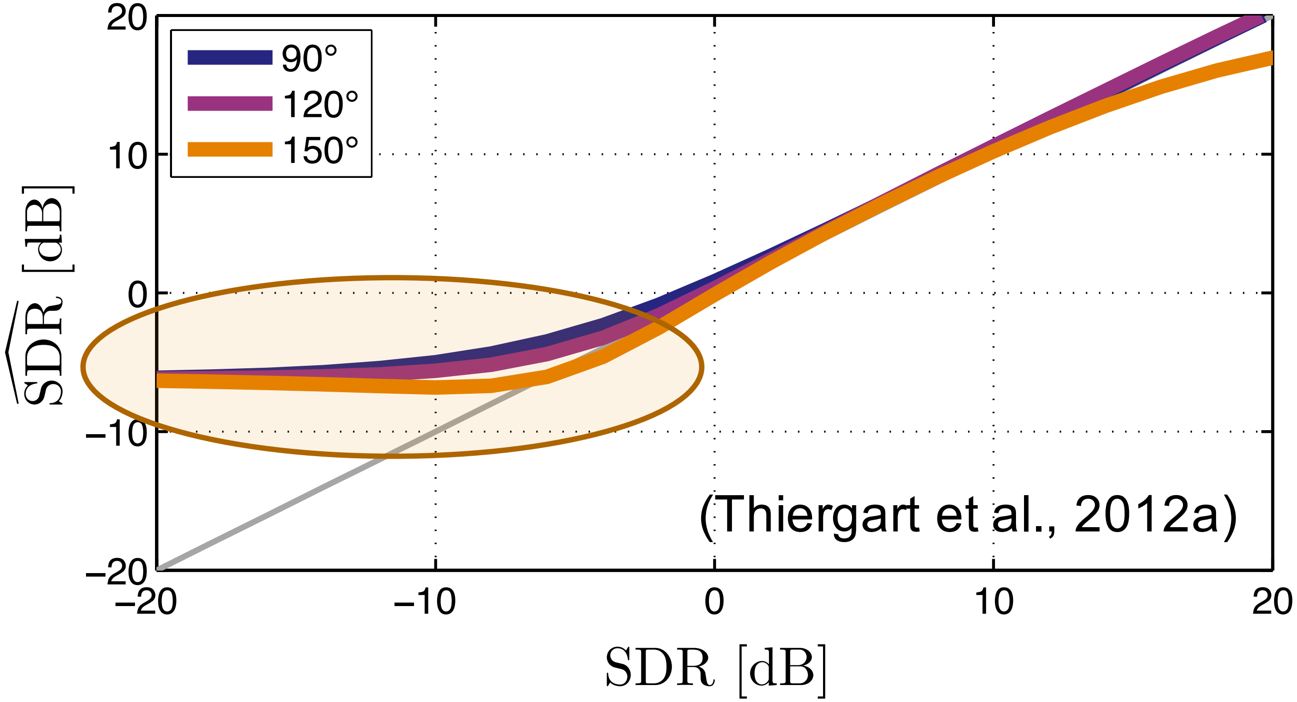
Reverberation Suppression
Multi-Channel Spectral Enhancement

- Example:
  - SDR estimation at 2.5 kHz using two omni-directional microphones



Reverberation Suppression
Multi-Channel Spectral Enhancement

- Example:
  - SDR estimation at 2.5 kHz using two omni-directional microphones



Reverberation Suppression
Multi-Channel Spectral Enhancement

- Rather than estimating the SDR, the reverberation PSD and/or speech PSD can be estimated separately

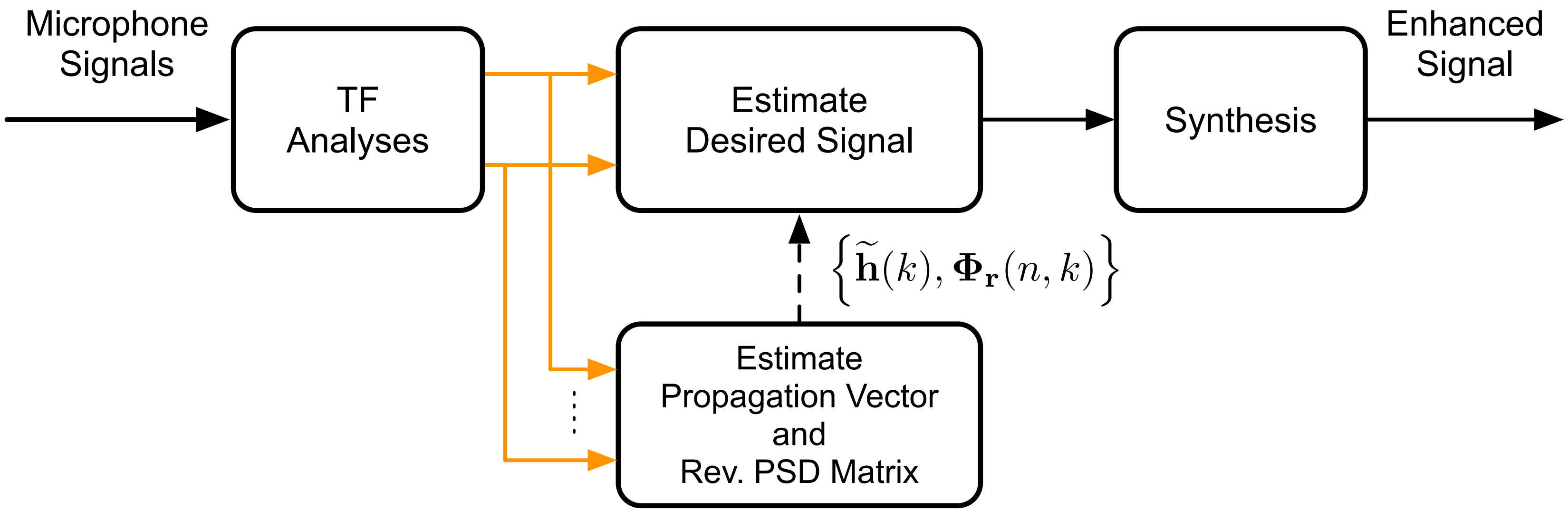
- Selected PSD estimators:
  - Two-channel estimation using a reference signal (Habets and Gannot, 2007)
  - Two-channel BSS algorithm with a null-constraint toward the desired source (Schwarz et al., 2012)
  - Maximum likelihood (ML) estimate of the reverberation PSD using reference signals (Braun et al., 2013/2015)
  - Joint ML estimate of the clean and reverberation PSD (Kuklasinski et al., 2014)
  - ML estimation of reverberation PSD in noisy environments using a Newton approach (O. Schwartz et al., 2015)
  - Joint ML estimation of speech and reverberation PSD in noisy environments using a Newton approach (O. Schwartz et al., 2016)

Reverberation Suppression
Multi-Channel Spectral Enhancement

- Previously discussed methods can also be used to estimate the late reverberation PSD
- Some examples:
  - Estimate the late reverberation PSD using a long-term multi-step linear prediction technique (Kinoshita et al., 2009)
  - Estimate the late reverberation signal and its PSD using a partial acoustic system equalizer (Cauchi et al. , 2015)
  - …

Reverberation Suppression
Data-Dependent Spatial Filtering

- Using the propagation vector and reverberation PSD matrix



- MVDR filter and post-filter (Lefkimmiatis et al., 2007)
- MC-MMSE filter in the spherical harmonic domain (Braun et al., 2013)
- MC-MMSE filter using relative direct-path transfer functions (Thiergart et al., 2014)
- MC-MMSE filter using relative early transfer functions (O. Schwartz et al., 2015)

Reverberation Suppression

- Opportunities
  - Robust technique that has been successfully applied to several practical scenarios
  - The estimated PSDs can be used to control the spatial filter
- Challenges / Open Issues
  - To achieve even more reverberation reduction
  - To handle anisotropic / unknown reverberant sound fields

Outline

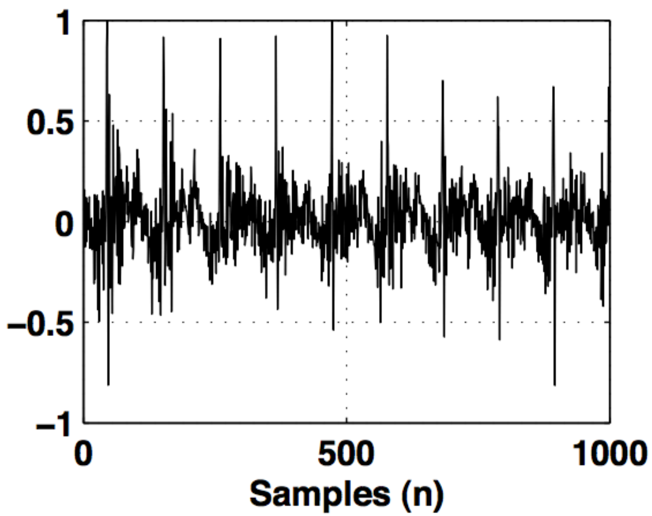
- Reverberation Cancellation
- Reverberation Suppression
- Direct Estimation
- Practical Challenges
- Conclusions and Future Challenges

Direct Estimation

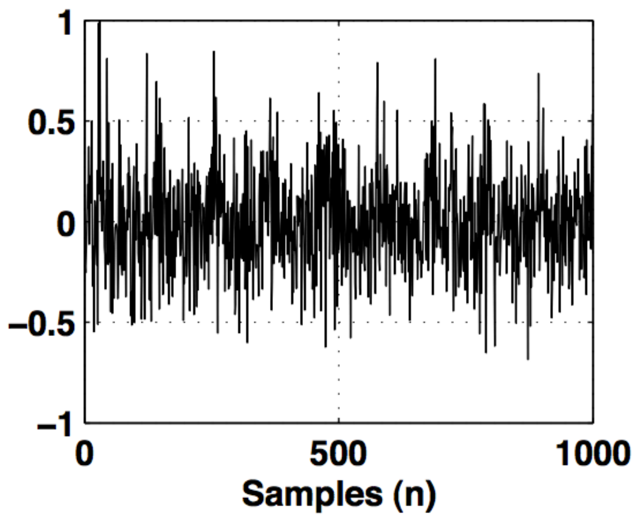
LP Residual Processing

- Based on the well-known source-filter model for speech production
- The all-pole filter coefficients that model the vocal tract are less affected by reverberation (Gaubitch et al., 2006)
- The prediction residual signal contains random peaks between the larynx-cycles due to reverberation

LP residual without reverberation



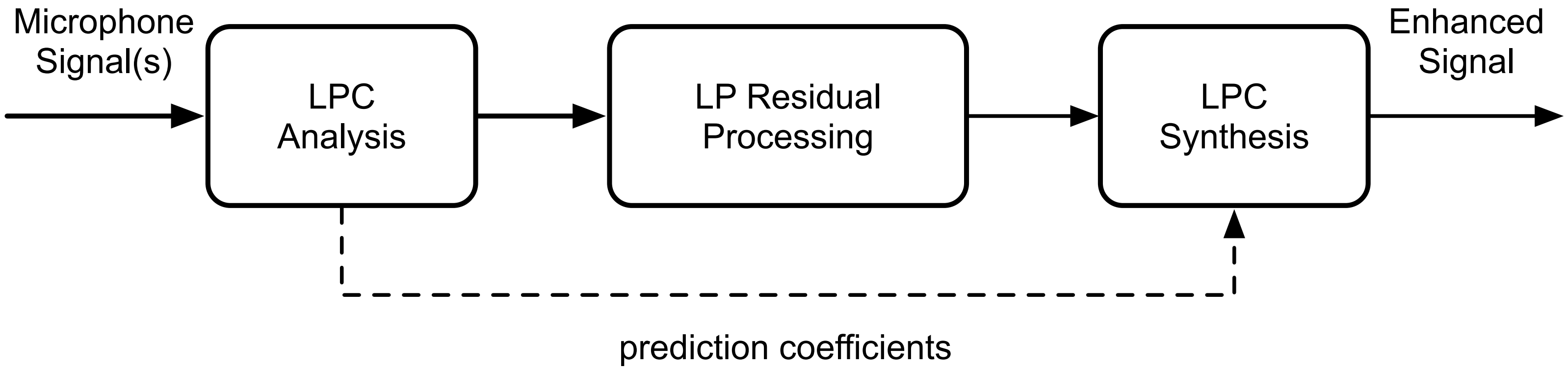
LP residual with reverberation



Direct Estimation
LP Residual Processing

- In general, the processing consists of the following steps
  1. Estimate the all-pole coefficients that model the vocal tract
  2. Compute the linear prediction residual signal
  3. Process the linear prediction signal
  4. Reconstruct the clean speech using the prediction coefficients and the enhanced linear prediction signal

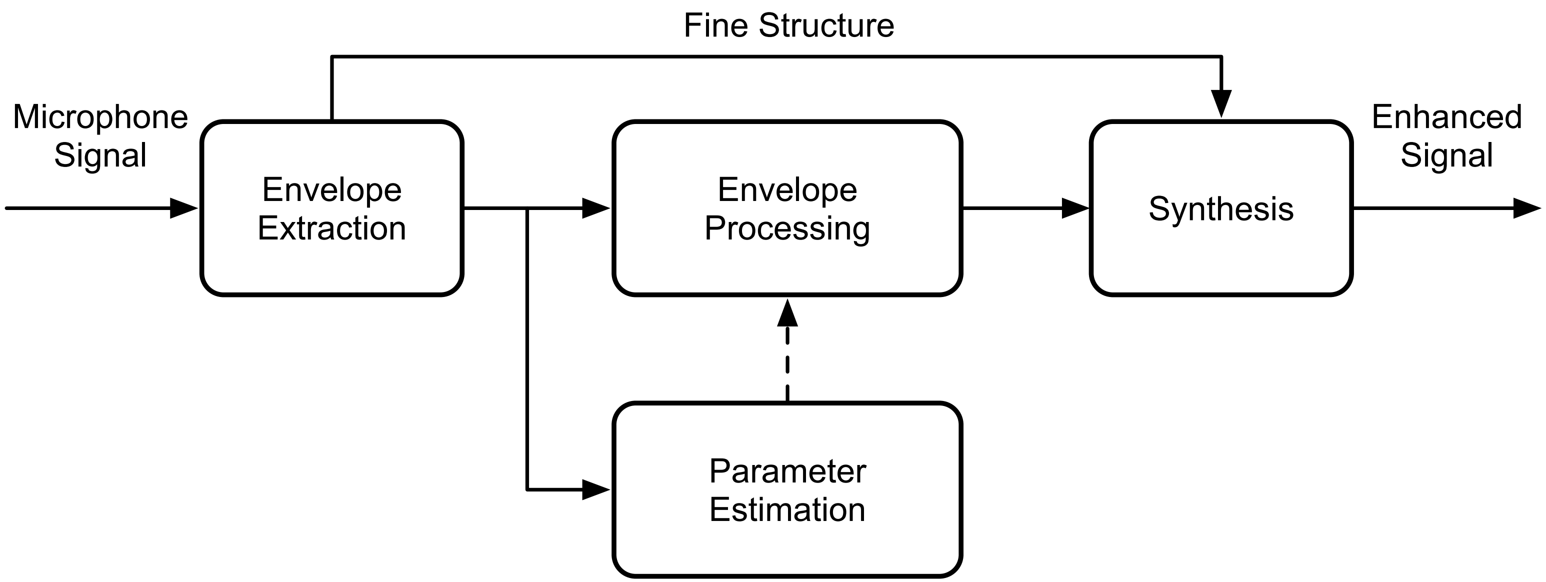
Direct Estimation
LP Residual Processing



- Several LP residual enhancement approaches:
  - Synthesis a Clean Residual Signal (Allen, 1974)
  - Regional Weighting Function (Yegnanarayana, 1998)
  - Wavelet Extrema Clustering (Griebel and Braindstein, 1999)
  - Kurtosis Maximization (Gillespie et al., 2001)
  - Larynx-Cycle Averaging (Gaubitch and Naylor, 2003)

Direct Estimation
Temporal Envelope Processing

- Reverberation has been found to reduce temporal modulations
- The reverberation signal can be decomposed in
  1. A temporal envelope
  2. A fine structure



Direct Estimation
Temporal Envelope Processing

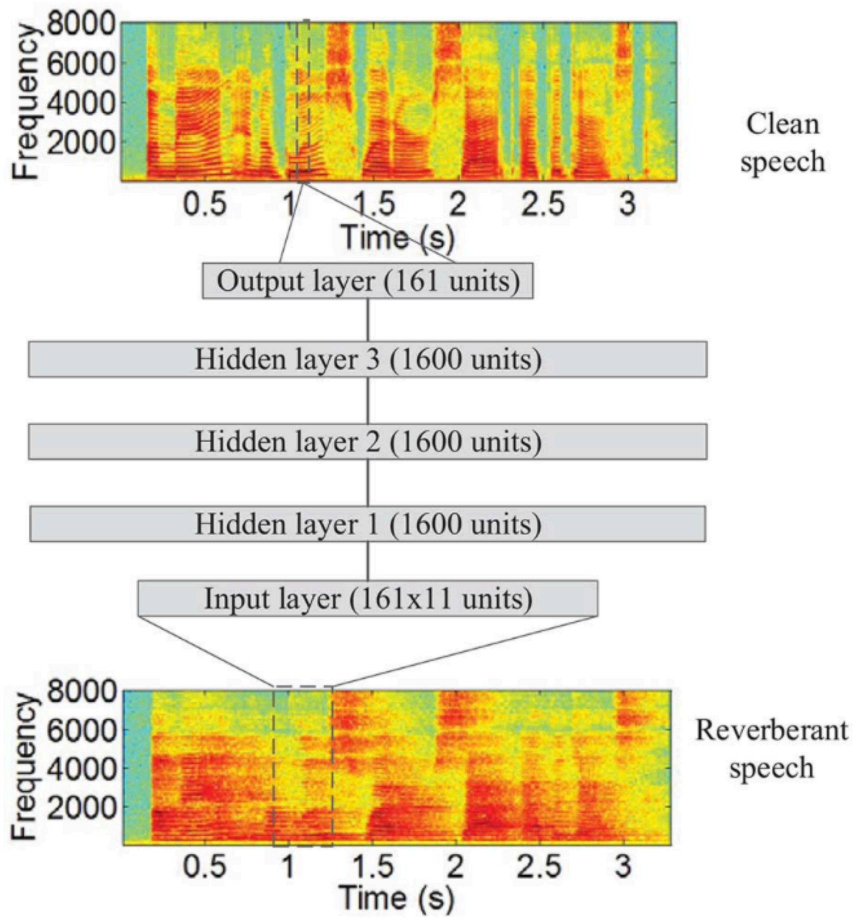
- Selected envelope processing approaches:
  - Analog center clipping of the envelope (Berkley and Mitchell, 1978)
  - Inverse filtering / deconvolution of the envelope
    - Langhans and Strube (1982)
    - Mourjopoulos and Hammond (1983)
  - …

Direct Estimation
Neural Networks

- DNN-based spectral mapping
  (Han et al, 2014) and (Han et al., 2015)

- Opportunities
  - Sufficient data
  - Processing (not training) can be performed in real-time

- Open Issues
  - How well does it generalize?



Any many more…

- Harmonicity based Dereverberation Method (HERB) (Nakatani et al., 2004)
- Inverse Modulation Transfer Function (Unoki et al., 2004)
- Bayesian inference of speech and channel model parameters (Hopgood, 2005)
- Convolutive Non-Negative Matrix Factorization (Kameoka, 2009)
- Variational Bayesian Inference for Multichannel Dereverberation and Noise Reduction (Schmid et al., 2014)
- ….

Outline

- Reverberation Cancellation
- Reverberation Suppression
- Direct Estimation
- Practical Challenges
- Conclusions and Future Challenges

Practical Challenges

- Dereverberation decreases the perceived loudness
  - Challenging particularly in the case of multiple sources
- Coloration due to early reflections
- Unbalance between early reflections and (lack of) late reverberation
- Additional interferences (e.g., echo, noise, interfering sources)
- Application specific requirements
  - Real-time processing
  - Low latency

Outline

- Reverberation Cancellation
- Reverberation Suppression
- Direct Estimation
- Practical Challenges
- Conclusions and Future Challenges

Conclusions and Future Challenges

- The field of dereverberation has significantly advanced during the last 15 years
- In the context of speech enhancement and ASR, dealing with reverberation remains one of the more challenging problems

- Future Challenges
  - Reduce early reflections / flutter echo
  - Larger source-receiver distances → Lower DRR and SNR
  - Application specific challenges (e.g., binaural dereverberation)
  - Performance measures (REVERB 2014)

References (1/7)

- M. I. Gürelli and C. L. Nikias, “EVAM: An eigenvector-based algorithm for multichannel blind deconvolution of input colored signals,” IEEE Trans. Signal Process., vol. 43, no. 1, pp. 134–149, Jan. 1995.
- S. Gannot and M. Moonen, “Subspace methods for multimicrophone speech dereverberation,” EURASIP Journal on Applied Signal Processing, vol. 2003, no. 11, pp. 1074–1090, 2003.
- G. Xu, H. Liu, L. Tong, and T. Kailath, “A least-squares approach to blind channel identification,” IEEE Trans. Signal Process., vol. 43, no. 12, pp. 2982–2993, Dec. 1995.
- Y. Huang and J. Benesty, “Adaptive multi-channel least mean square and Newton algorithms for blind channel identification,” Signal Processing, vol. 82, pp. 1127–1138, Aug. 2002.
- Y. Huang and J. Benesty, “A class of frequency-domain adaptive approaches to blind multichannel identification,” IEEE Trans. Signal Process., vol. 51, no. 1, pp. 11–24, Jan. 2003.
- M. Haque and M. Hasan, “Noise robust multichannel frequency-domain LMS algorithms for blind channel identification,” IEEE Signal Process. Lett., vol. 15, pp. 305–308, 2008.
- E.A.P. Habets and P.A. Naylor, “An online quasi-Newton algorithm for blind SIMO identification,” Proc. of the IEEE Intl. Conf. on Acoustics, Speech, and Signal Processing (ICASSP 2010), Mar. 14-19, 2010.
- S Malik and G Enzner, “State-space frequency-domain adaptive filtering for nonlinear acoustic echo cancellation,” IEEE Trans. Acoust., Speech, Signal Process., Vol. 20, Issue 7, pp. 2065-2079, 2012.
- M. Miyoshi and Y. Kaneda, “Inverse filtering of room acoustics,” IEEE Trans. Acoust., Speech, Signal Process., vol. 36, no. 2, pp. 145–152, Feb. 1988.
- J. L. Flanagan, A. C. Surendran, and E. E. Jan, “Spatially selective sound capture for speech and audio processing,” Speech Communication, vol. 13, no. 1-2, pp. 207–222, Oct. 1993.
- M. Kallinger and A. Mertins, “Multi-channel room impulse response shaping - a study,” Proc. IEEE Intl. Conf. Acoust., Speech, Signal Processing (ICASSP), 2006.
- W. Zhang and P. A. Naylor, “An algorithm to generate representations of system identification errors,” Research letters in signal processing, vol. 2008, 2008.

References (2/7)

- W. Zhang, E. A. P. Habets, and P. A. Naylor, “On the use of channel shortening in multichannel acoustic system equalization,” Proc. Intl. Workshop Acoust. Echo Noise Control (IWAENC), Tel Aviv, Israel, Aug. 2010.
- A. Mertins, T. Mei, and M. Kallinger, “Room impulse response shortening/reshaping with infinity- and p-norm optimization,” IEEE Trans. Audio, Speech, Lang. Process., vol. 18, no. 2, pp. 249–259, Feb. 2010.
- F. Lim and P. A. Naylor, “Relaxed multichannel least squares with constrained initial taps for multichannel dereverberation,” Proc. Intl. Workshop Acoust. Signal Enhancement (IWAENC), 2012.
- I. Kodrasi and S. Doclo, “Robust partial multichannel equalization techniques for speech dereverberation,” Proc. IEEE Intl. Conf. on Acoustics, Speech and Signal Processing (ICASSP), Kyoto, Japan, Apr. 2012.
- I. Kodrasi, S. Goetze, and S. Doclo, “Increasing the robustness of acoustic multichannel equalization by means of regularization,” Proc. Intl. Workshop Acoust. Signal Enhancement (IWAENC), 2012, pp. 161–164.
- I. Kodrasi, S. Goetze, and S. Doclo. "A perceptually constrained channel shortening technique for speech dereverberation," Proc. IEEE Intl. Conf. on Acoustics, Speech and Signal Processing (ICASSP), 2013.
- F. Lim, W. Zhang, E.A.P. Habets and P.A. Naylor, "Robust multichannel dereverberation using relaxed multichannel least squares," IEEE/ACM Trans. on Audio, Speech, and Language Processing, Vol. 22, Issue 9, pp. 1379-1390, Sep. 2014.
- B. Schwartz, S. Gannot and E.A.P. Habets, "Online speech dereverberation using Kalman filter and EM algorithm," IEEE/ACM Trans. on Audio, Speech, and Language Processing, Vol. 23, Issue 2, pp. 394-406, 2015.
- Y. Peled and B. Rafaely, "Linearly-constrained minimum-variance method for spherical microphone arrays based on plane-wave decomposition of the sound field,” IEEE Trans. on Audio, Speech, and Language Processing, Vol. 21, Issue 12, pp. 2532-2540, 2013.
- M. Delcroix, T. Hikichi, and M. Miyoshi, “Dereverberation of speech signals based on linear prediction,” Proc. of the 8th Intl. Conf. on Spoken Language Processing (ICSLP), Korea, 2004.
- M. Triki and D.T.M. Slock, “Blind dereverberation of quasi-periodic sources based on multichannel linear prediction,” Proc. of the International Workshop on Acoustic Echo and Noise Control (IWAENC), Netherlands, 2005.

References (3/7)

- T. Yoshioka, T. Nakatani, and M. Miyoshi, "Integrated speech enhancement method using noise suppression and dereverberation," IEEE Trans. on Audio, Speech, and Language Processing, Vol. 17, Issue 2, pp. 231-246, 2009.
- M. Togami, Y. Kawaguchi, R. Takeda, Y. Obuchi, N. Nukaga, "Optimized speech dereverberation from probabilistic perspective for time varying acoustic transfer function," IEEE Trans. on Audio, Speech, and Language Processing, Vol. 21, Issue 7, pp. 1369-1380, 2013.
- A. Jukić and S.Doclo, "Speech dereverberation using weighted prediction error with Laplacian model of the desired signal," Proc. of the Intl. Conf. on Acoustics, Speech and Signal Processing (ICASSP), 2014.
- T. Yoshioka and T. Nakatani, "Generalization of multi-channel linear prediction methods for blind MIMO impulse response shortening," IEEE Trans. On Audio, Speech, and Language Processing, Vol. 20, Issue 10, pp. 2707-2720, 2012.
- K. Lebart, J.M. Boucher, and P.N Denbigh, "A new method based on spectral subtraction for speech dereverberation," Acta Acoustica, vol. 87, no. 3, pp. 359–366, 2001.
- E.A.P. Habets, "Single-channel speech dereverberation based on spectral subtraction," Proc. of the 15th Annual Workshop on Circuits, Systems and Signal Processing (ProRISC), Netherlands, Nov. 25-26, 2004.
- E.A.P. Habets, S. Gannot and I. Cohen, "Late reverberant spectral variance estimation based on a statistical model," 2014 SPS Signal Processing Letters Best Paper Award, IEEE Signal Processing Letters, Vol. 16, No. 9, pp. 770-773, Sep. 2009.
- E.A.P. Habets, "Single- and Multi-Microphone Speech Dereverberation using Spectral Enhancement," Ph.D. Thesis, Technische Universiteit Eindhoven, The Netherlands, Jun. 25, 2007.
- J.S. Erkelens and R.Heusdens, "Correlation-based and model-based blind single-channel late-reverberation suppression in noisy time-varying acoustical environments." IEEE Trabs. on Audio, Speech, and Language Processing, Vol . 18, Issue 7, pp. 1746-1765, 2010.
- L. Danilenko, "Binaural hearing in non-stationary diffuse sound field," Dissertation, RWTH Aachen University, 1968.

References (4/7)

- J. Allen, D. Berkley, and J. Blauert, “Multimicrophone signal-processing technique to remove room reverberation from speech signals,” J. Acoust. Soc. Am., Vol. 62, No 4, pp. 912-915, 1977.
- J. Ahonen and V. Pulkki, “Diffuseness estimation using temporal variation of intensity vectors,” Proc. of the Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA), New Paltz, 2009.
- M. Jeub, C.M. Nelke, C. Beaugeant, P. Vary, ”Blind estimation of the coherent-to-diffuse energy ratio from noisy speech signals.” Proc. of the European Signal Processing Conf. (EUSIPCO), Spain, 2011).
- O. Thiergart, G. Del Galdo and E.A.P. Habets, "Signal-to-reverberant ratio estimation based on the complex spatial coherence between omnidirectional microphones," Proc. of the IEEE Intl. Conf. on Acoustics, Speech, and Signal Processing (ICASSP 2012), Kyoto, Japan, Mar. 25-30, 2012.
- O. Thiergart, G. Del Galdo and E.A.P. Habets, "Diffuseness estimation with high temporal resolution via spatial coherence between virtual first-order microphones," Proc. of the IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA 2011), New York, USA, Oct. 16-19, 2011.
- O. Thiergart, T. Ascherl and E.A.P. Habets, "Power-based signal-to-diffuse ratio estimation using noisy directional microphones," Proc. of the IEEE Intl. Conf. on Acoustics, Speech, and Signal Processing (ICASSP), Florence, Italy, May 4-9, 2014.
- A. Schwarz and W. Kellermann, "Coherent-to-diffuse power ratio estimation for dereverberation," IEEE/ACM Trans. on Audio, Speech, and Language Processing, Vol. 23, Issue 6, pp. 1006-1018, 2015
- A. Schwarz, K. Reindl, and W. Kellermann, "A two-channel reverberation suppression scheme based on blind signal separation and Wiener filtering," Proc. of the IEEE Intl. Conf. on Acoustics, Speech and Signal Processing (ICASSP), 2012.
- S. Braun and E.A.P. Habets, "Dereverberation in noisy environments using reference signals and a maximum likelihood estimator," Proc. of the 21th European Signal Processing Conference (EUSIPCO), Marrakech, Morocco, Sep. 2-9, 2013.
- S. Braun and E.A.P. Habets, "A multichannel diffuse power estimator for dereverberation in the presence of multiple sources," EURASIP Journal on Audio, Speech, and Music Processing, Vol. 2015:34, 2015.

References (5/7)

- A. Kuklasinski, S. Doclo, S.H. Jensen, J. Jensen, “Maximum likelihood based multi-channel isotropic reverberation reduction for hearing aids,” Proc. European Signal Processing Conf. (EUSIPCO), Lisbon, Portugal, pp. 61–65, 2014.
- O. Schwartz, S. Braun, S. Gannot and E.A.P. Habets, ”Maximum likelihood estimation of the late reverberant power spectral density in noisy environments,” Proc. of the IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA), 2015.
- O. Schwartz, S. Gannot and E.A.P. Habets, "Joint maximum likelihood estimation of late reverberant and speech power spectral density in noisy environments," Proc. of the IEEE Intl. Conf. on Acoustics, Speech, and Signal Processing (ICASSP), Shanghai, China, 2016.
- K. Kinoshita, M. Delcroix, T. Nakatani, and M. Miyoshi, "Suppression of late reverberation effect on speech signal using long-term multiple-step linear prediction, " IEEE Trans. on Audio, Speech, and Language Processing, Vol. 17, Issue 4, pp. 534-545, 2009.
- B. Cauchi, P.A. Naylor, T. Gerkmann, S. Doclo, and S. Goetze, "Late reverberant spectral variance estimation using acoustic channel equalization," Proc. European Signal Processing Conference (EUSIPCO), pp. 2481-2485, 2015.
- S. Lefkimmiatis and P. Maragos, "A generalized estimation approach for linear and nonlinear microphone array post-filters," Speech Communication, Vol. 49, Issue 7, pp. 657-666, 2007.
- S. Braun, D.P. Jarrett, J. Fischer and E.A.P. Habets, "An informed spatial filter for dereverberation in the spherical harmonic domain," Proc. of the IEEE Intl. Conf. on Acoustics, Speech, and Signal Processing (ICASSP), Vancouver, Canada, May 26-31, 2013.
- O. Thiergart, M. Taseska and E.A.P. Habets, "An informed parametric spatial filter based on instantaneous direction-of-arrival estimates," IEEE/ACM Trans. on Audio, Speech, and Language Processing, Vol. 22, Issue 12, pp. 2182-2196, Dec. 2014.
- O. Schwartz, S. Gannot and E.A.P. Habets, "Multi-microphone speech dereverberation and noise reduction using relative early transfer functions," IEEE/ACM Trans. on Audio, Speech, and Language Processing, Vol. 23, Issue 2, pp. 240-251, Jan. 2015.

AUDIO LABS

References (6/7)

- N.D. Gaubitch, D.B. Ward, and P.A. Naylor, “Statistical analysis of the autoregressive modeling of reverberant speech,” Journal of the Acoustical Society of America, vol. 120, pp. 4031–4039, Dec. 2006.
- J.B. Allen, “Synthesis of pure speech from a reverberant signal,” U.S. Patent No. 3786188, 1974.
- B. Yegnanarayana, “Enhancement of reverberant speech using LP residual,” Proc. of the IEEE Intl. Conf. on Acoustics, Speech, and Signal Processing (ICASSP), 1998.
- S. Griebel and M. Brandstein, “Wavfelet tranform extrama clustering for multi-channel speech deverberation,” Proc. of the IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA’99), 1999.
- B.W. Gillespie, H. Malvar, and D. Florencio, “Speech dereverberation via maximum- kurtosis subband adaptive filtering,” Proc. of the IEEE Intl. Conf. on Acoustics, Speech, and Signal Processing (ICASSP), 2001.
- N. D. Gaubitch, P. A. Naylor, and D. B. Ward, “On the use of linear prediction for dereverberation of speech,” Proc. Intl. Workshop Acoust. Echo Noise Control (IWAENC), pp. 99–102, 2003.
- D.A. Berkley and O.M.M. Mitchell, “Removing reverberative echo components in speech signals,” U.S. Patent No. 4166924, 1979.
- T. Langhans and H. Strube, “Speech enhancement by nonlinear multiband envelope filtering,” Proc. of the IEEE Intl. Conf. on Acoustics, Speech, and Signal Processing (ICASSP), 1982.
- J. Mourjopoulos and J.K. Hammond, “Modelling and enhancement of reverberant speech using an envelope convolution method,” Proc. of the IEEE Intl. Conf. on Acoustics, Speech and Signal Processing (ICASSP), 1983.
- K. Han, Y. Wang, and D. Wang, "Learning spectral mapping for speech dereverberation," Proc. IEEE Intl. Conf. on Acoustics, Speech and Signal Processing (ICASSP), 2014.
- K. Han, Y. Wang, D. Wang, W.S. Woods, I. Merks, and T. Zhang, "Learning spectral mapping for speech dereverberation and denoising," IEEE/ACM Trans. on Audio, Speech, and Language Processing, Vol. 23, Issue 6, pp. 982-992, 2015.

References (7/7)

- T. Nakatani, B. Juang, K. Kinoshita, and M. Miyoshi, “Harmonicity based derever- beration with maximum a posteriori estimation,” Proc. of the IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA’05), 2005, pp. 94–97.
- K. Sakata M. Unoki, M. Furukawa and M. Akagi, “A speech dereverberation method based on MTF concept in power envelope restoration,” Acoustical Science and Technology, vol. 25, no. 4, pp. 243–254, 2004.
- J. R. Hopgood, "Models for blind speech dereverberation: A subband all-pole filtered block stationary autoregressive process," Proc. European Signal Processing Conference (EUSIPCO), 2005.
- H. Kameoka, T. Nakatani, and T. Yoshioka, "Robust speech dereverberation based on non-negativity and sparse nature of speech spectrograms," Proc. IEEE Intl. Conf. on Acoustics, Speech and Signal Processing (ICASSP), 2009.
- D. Schmid, G. Enzner, S. Malik, D. Kolossa, R. Martin, “Variational Bayesian inference for multichannel dereverberation and noise reduction,” IEEE/ACM Trans. on Audio, Speech, and Language Processing, Vol. 22, Issue 8, pp. 1320-1335, 2014.