UA 2014
2nd International Conference and Exhibition
on Underwater Acoustics
22-27 June 2014, Rhodes, Greece

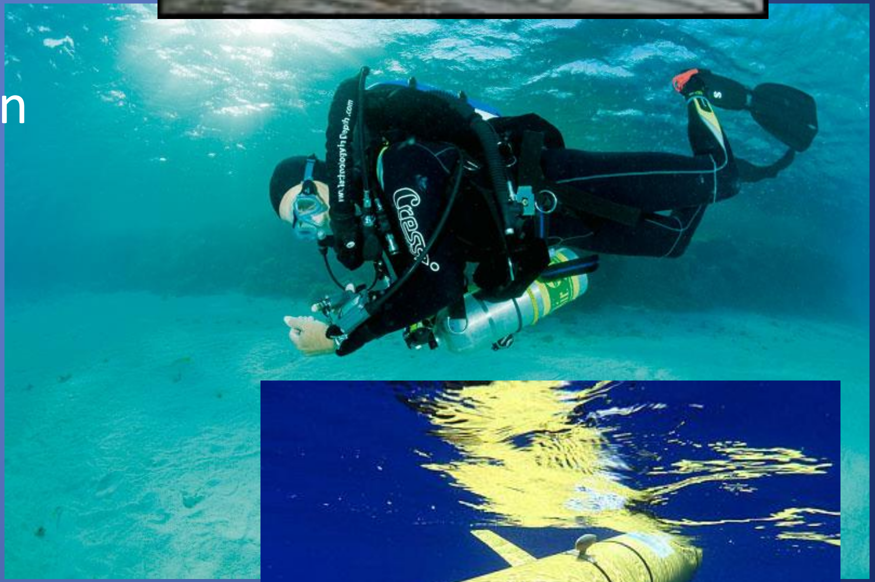
Estimating source spectra from recordings made in a reverberant underwater channel

Kay Gemba and Eva-Marie Nosal
Dept. of Ocean and Resources Engineering
Honolulu, HI 96822
gemba@hawaii.edu

This material is based upon work supported by the U.S. Department of Homeland Security, Science and Technology Directorate, Office of University Programs, under Grant Award Number 2008-ST-061-ML0002.

Motivation

- The performance of a passive acoustic detector can be improved by knowledge of the anechoic signature of the target and the noise environment
- No procedure for source characterization in reverberant environments
  - Anechoic facilities are not available
- Goal: Develop a robust and practical dereverberation method for underwater pool experiments
  - Method should be applicable to a variety of sources such as AUVs, Surface robots, Gliders, UBA and others without any special equipment or configurations



Problem Formulation

- Source levels recorded in reverberant environments are overestimate due to early reflections and late reverberation

- A solution is to estimate the impulse response (IR) of the recording channel and remove additional reverberant energy by inverting the IR.

- Assumptions:
  - Linear but not necessary time invariant system
  - Noise is stationary and uncorrelated
  - Sources have same directionality
  - Ergodic

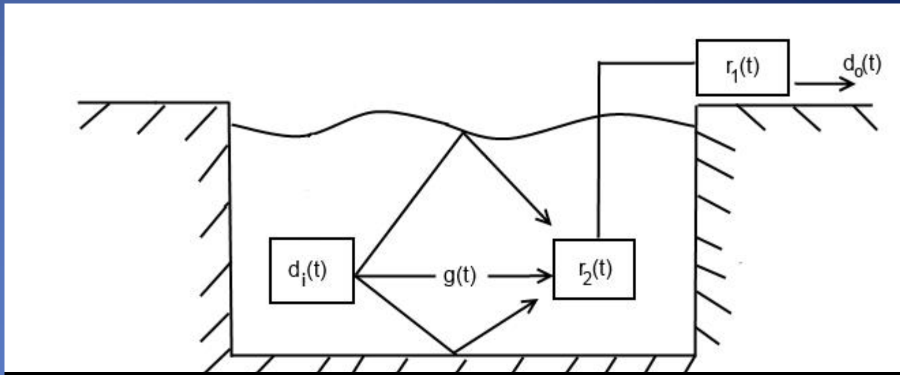


Fig.1: Pool diagram (backward problem)

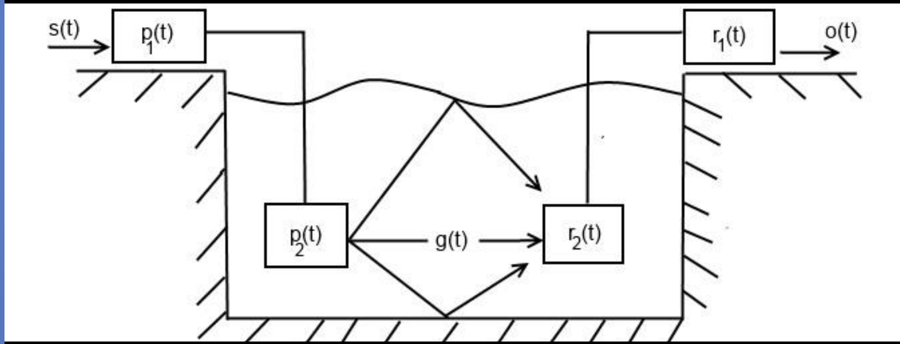


Fig.2: Pool diagram (forward problem)

Component	Time domain
Unknown SCUBA diver	$d_i(t)$ / $d_o(t)$
Input / Output Signal	$s(t)$ / $o(t)$
Playback elements	$p_1(t)$, $p_2(t)$
Receiver elements	$r_1(t)$, $r_2(t)$

Incoherent vs. coherent magnitude performance

1. h(t) is the combined impulse response from the playback system to the ADC
2. Backward Problem: The diver can be found by convolving the recorded signal with the expectation of the inverse and with the 'playback' impulse responses

$$\mathbf{h(t) = r_1(t) * \mathit{r}_2(t) * g(t) * p_2(t) * p_1(t)} \quad (1)$$

$$\boxed{d_i(t) = d_o(t) * E[h^{-1}(t)] * p_2(t) * p_1(t) \quad (2)}$$

3. Forward Problem: The PSD of a recorded control $\mathbf{S_o}$ signal is adjusted incoherently by the ensemble average of the transfer function and smoothed with a zero-phase moving average filter M
4. For a "coherent" comparison, the PSD of control signal is adjusted by the optimum-inverse in the least-squares sense $|\widehat{F}|$ and by a constant (mean of equalized signal).

$$\mathbf{o(t) = h(t) * s(t)}$$

$$\mathbf{10\log_{10}\hat{S}_s = M[10\log_{10}S_o - E[20\log_{10}|H|]\,]} \quad (3)$$

$$\mathbf{10\log_{10}\hat{S}_s = M[10\log_{10}S_o + E\left[20\log_{10}|\widehat{F}|\right] - 2\hat{D}]} \quad (4)$$

Coherent Inversion

- Coherent inversion is achieved in the least-squares sense using a processing delay

$$\hat{f} = \arg\min_f ||Af - z||_2^2$$

$$\hat{f} = [A^T A]^{-1} A^T z$$

- In practice, two parameters are varied to minimize the error:
  - Length of the IR
  - Delay of the spike

$$\sigma_t = D(l)$$

$$\sigma_f = \left[\frac{1}{I} \sum_{k=0}^{I-1} (10 \log_{10} |\hat{D}(k)| - \bar{D})^2 \right]^{-1/2}$$

$$\bar{D} = \frac{1}{I} \sum_{k=0}^{I-1} 10 \log_{10} |\hat{D}(k)|$$

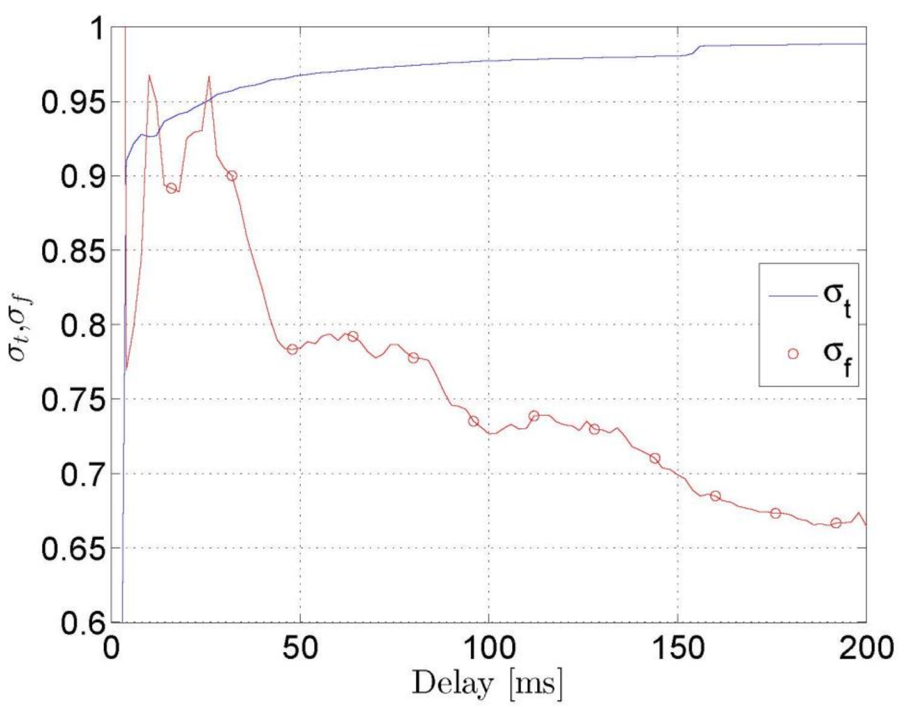


Fig.3: Inversion performance vs. delay l

“Forward” Experiment

UNIVERSITY of HAWAI‘I MĀNOA

Fig.4: Reverberation experiment: Playback and recording equipment

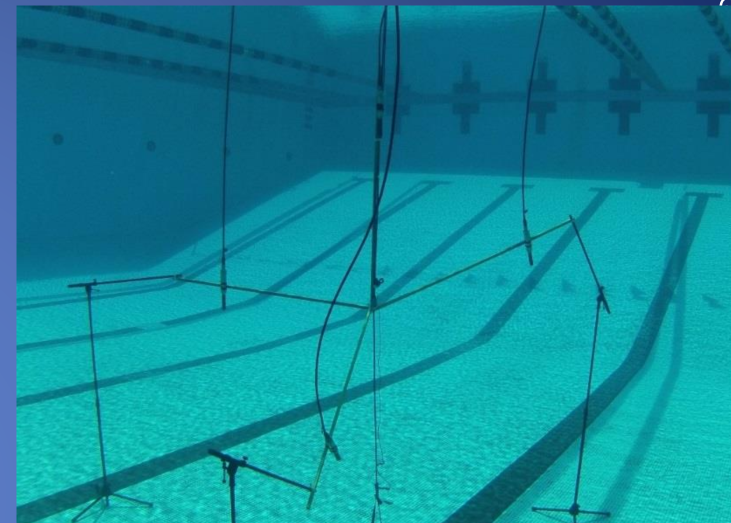


Fig.5: Spherical array

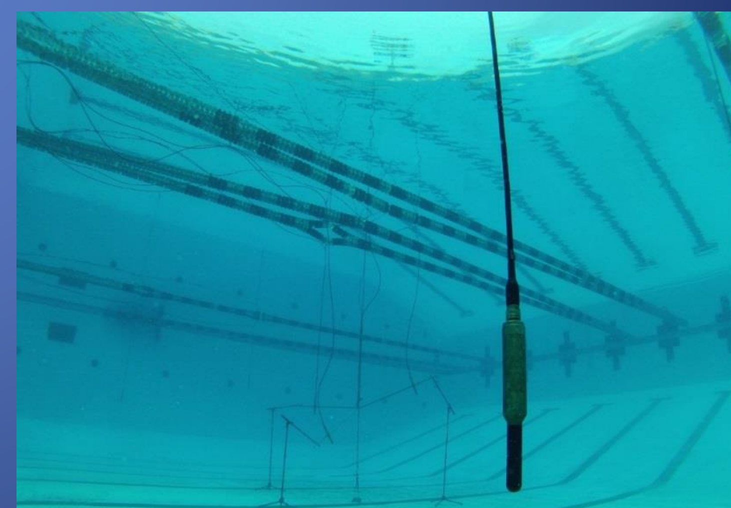


Fig.6: Random hydrophone

- Pool dimensions: 22.9 x 22.9 x 5.2 m
- 4 spherical array hydrophones (at 1m)
  - Only use one channel but can be extended
- 5 random hydrophones

Recorded signals and IR estimation

TABLE I. Overview of Recorded Signals

Signal Type	Duration [s]	Start [kHz]	Step [kHz]	Stop [kHz]	Repetition	Pre-amp Gain [dB]
Linear sweep	3	1	-	85	50	3
Logarithmic sweep	3	1	-	85	50	3
M-Sequence	5	1	-	85	50	0
Sinusoids	5	5	5	85	10	3
Mixed sinusoids	5	5	1	85	1	3
White noise	4	10	10	80	10	6
Mixed Sinusoids (+2cm)	5	5	1	85	1	3
Mixed Sinusoids (+4cm)	5	5	1	85	1	3

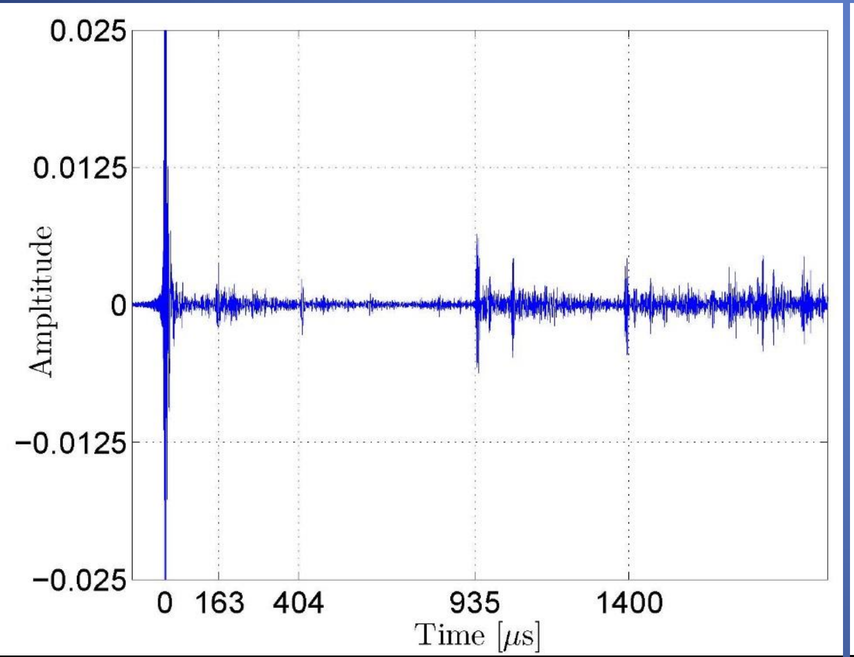


Fig.7: Impulse response with theoretical boundary reflection times

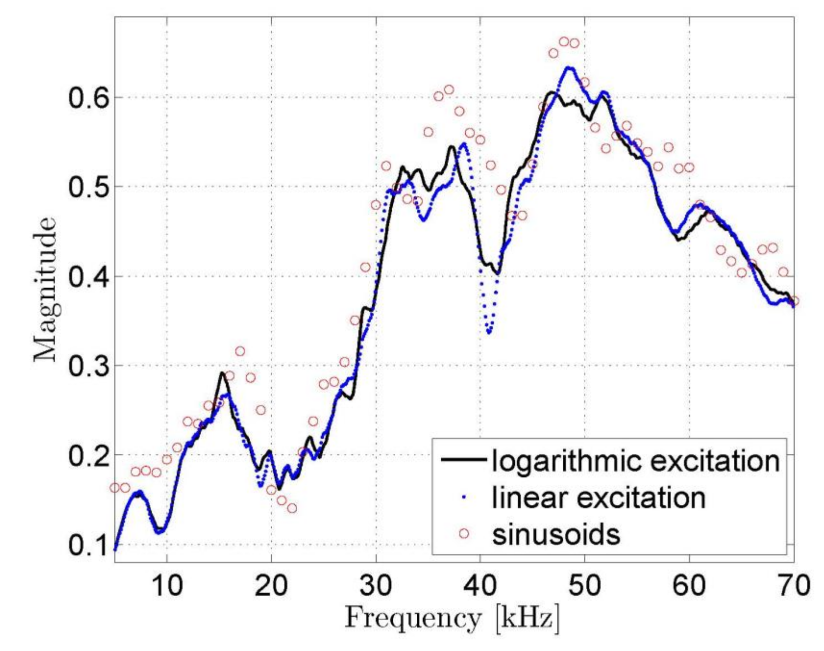


Fig.8: Spectral comparison of excitation methods and sinusoids

Estimation of IR length

- Modified Schroeder's method of backward integration

$$< g^2(t) > = \int_t^{\infty} \left[(k(\tau) + \eta(\tau))^2 - \overline{\eta^2} \right] d\tau$$

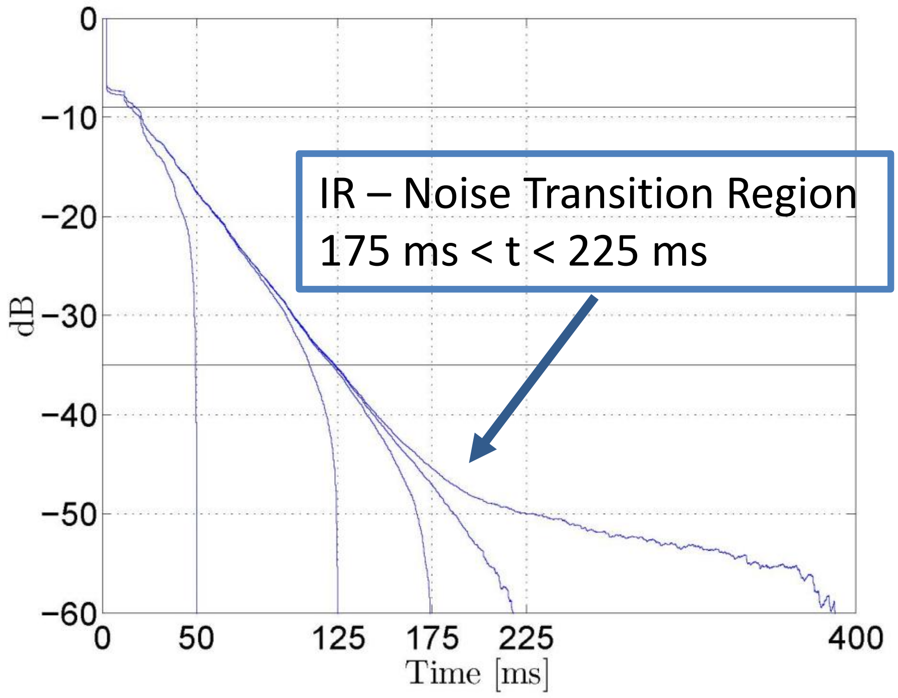


Fig.9: Decay slopes of IR with subtracted noise average

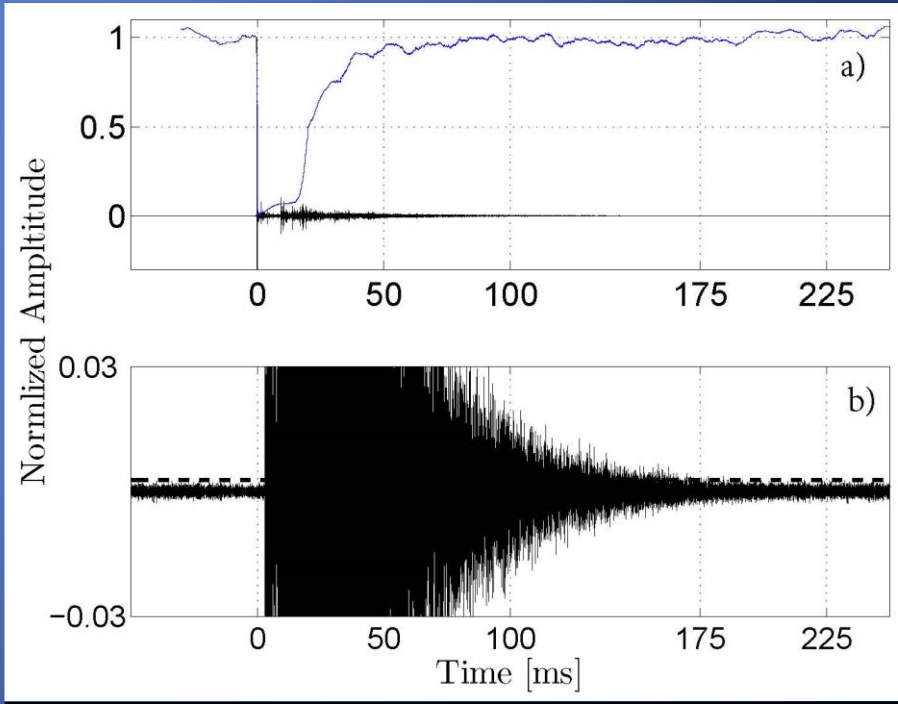


Fig.10: a) IR with echo density (top trace)
b) clock aligned avg. of IR w/ noise line

Coherent vs. Incoherent Dereverberation Comparison

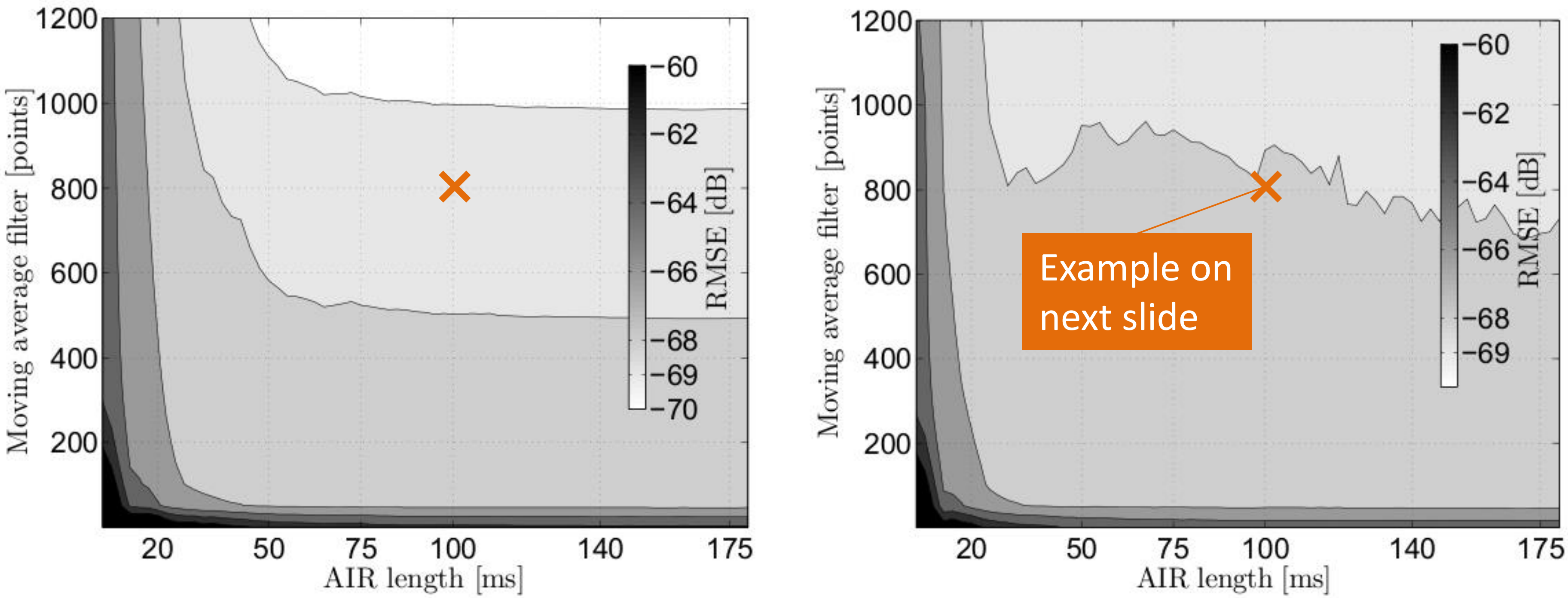


Fig.11: RMSE of dereverberated linear sweep: Incoherent (left) and coherent (right) inverse using 10 realizations. RMSE ticks correspond to contour surfaces

$$\boldsymbol{RMSE} = \mathbf{10}\log_{10}\sqrt{\frac{1}{N}\sum_{k=1}^{N}||\widehat{S_k} - S_k||^2}$$

RSME is computed for PSD coefficients over 10-70 kHz band (1 Hz resolution)

Dereverberation Example

Fig.12: Dereverberation example using an IR length of 100 ms and a moving average filter length of 800 points

a) Recorded linear sweep (1 realization)

b) Incoherent adjustment with 1 TF

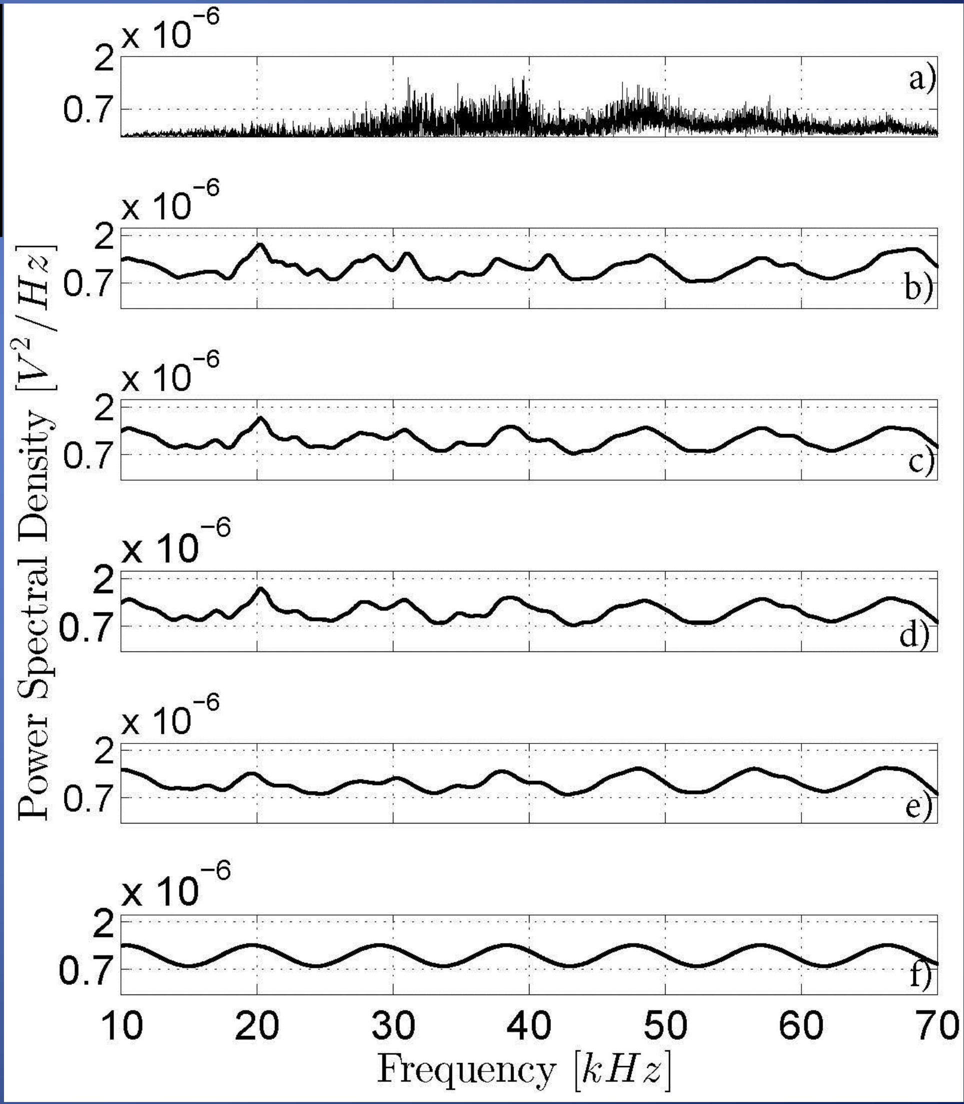
c) Incoherent adjustment with 10 TF

d) Incoherent adjustment with 48 TF

e) Coherent adjustment with 10 TF

f) Original linear sweep

- Optimal frequency range of transmitting transducer > 35 kHz



Dereverberation Procedure

1. Calculate theoretical reverberation time over 60 dB (T_{60}) of pool (no a priori recordings required)

$$T_{60} = 0.0368 \frac{V}{-S \log(1 - \sum_{i=1}^{6} \frac{\alpha_i A_i}{S})}$$

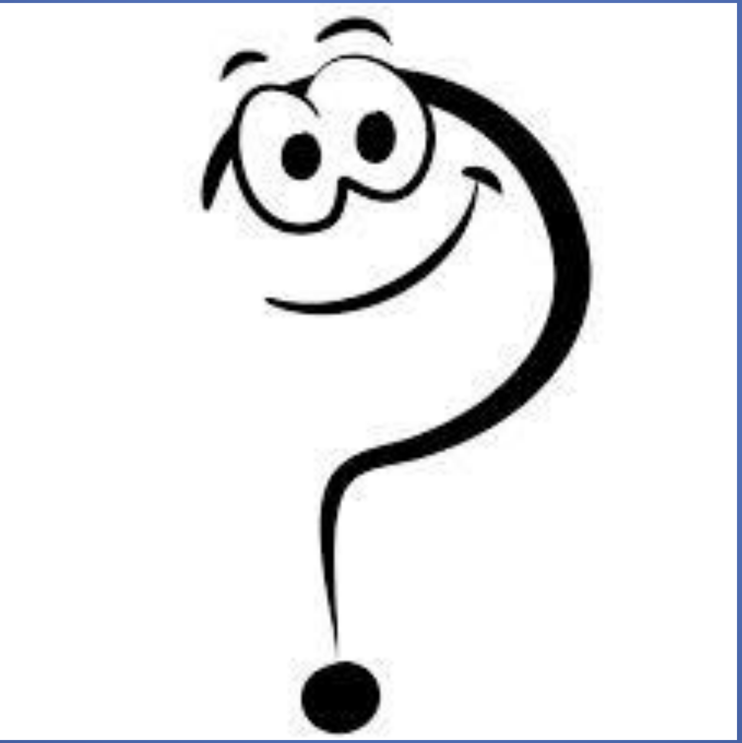
2. Design exponential sweep (approx. 5-10 times longer) and properly scaled inverse
3. Record 100 realizations (10 min) in the same channel (length of 1m) as the unknown source

Dereverberation Procedure

4. Deconvolve IR and estimate its length using Schroeder's method and/or echo density. Window IR accordingly.
5. Compute incoherent average of the transfer function and adjust PSD of unknown source to obtain SSL: $\pm\,\sigma\ [dB\ re\ 1\mu Pa^{2}/_{Hz}\ at\ 1m]$

$$\pm\sigma \approx 39.8r\left(1-\sum_{i=1}^{6}\frac{\alpha_i}{6}\right)^{\frac{1}{2}}\left(\sum_{i=1}^{6}\alpha_i A_i\right)^{-\frac{1}{2}}\ \mathrm{dB}$$

Questions?



Example: Scaling of Excitation Sweep

- Log Sweep Properties:
  - Frequency [1 to 85 kHz]
  - Length: 3 seconds
  - FS: 264600.18
  - Amplitude: 0.4
- The procedure is simplified for the linear sweep (scaling only)

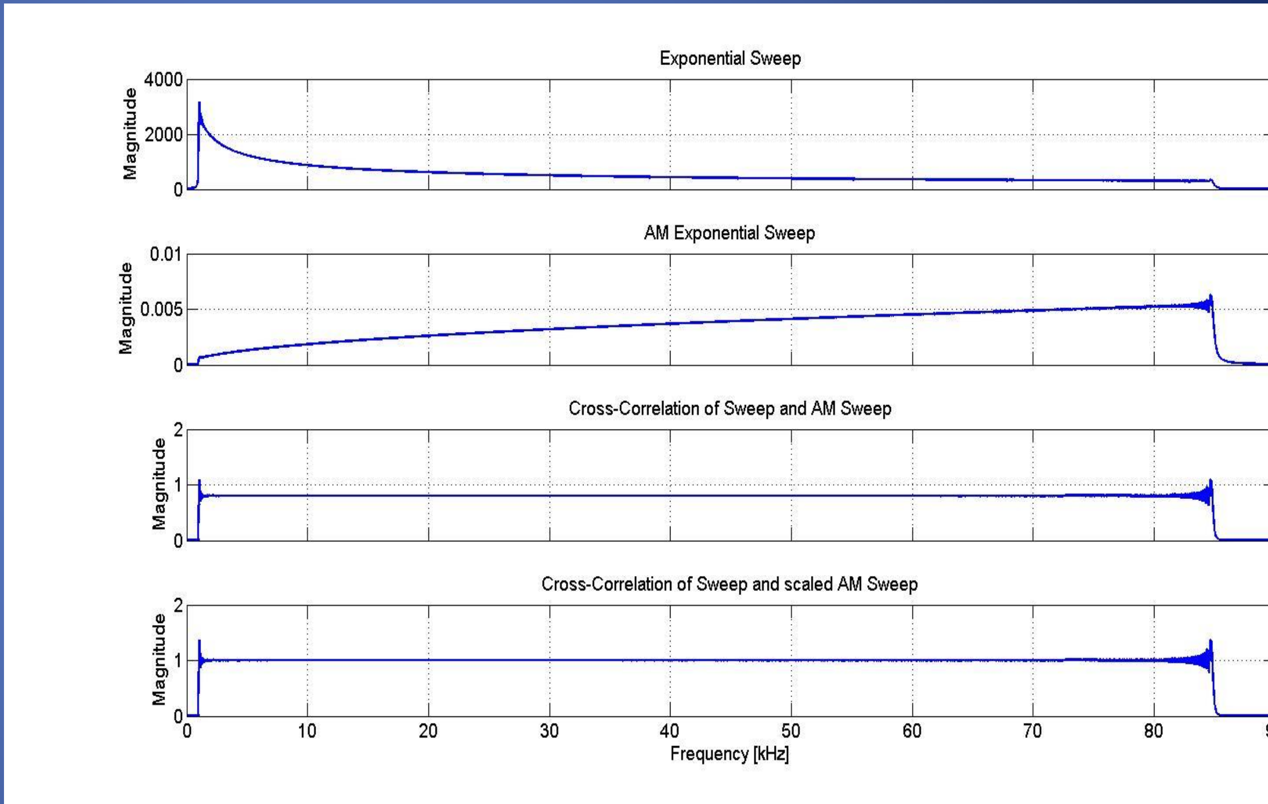


Fig.: scaled autocorrelation

Estimate g(t)

- Goal is to separate g(t) from the transducer transfer functions (assumed unknown in phase and amplitude)
- 1. convolve a synthetic AIR (1) w/ an unknown TF, resembling the IR of the electrical equipment (2)
- 2. Apply a cascaded scaling process to equalize frequency dependent gain and recover an estimate of the AIR (3)
- 3. Calculate error (4)

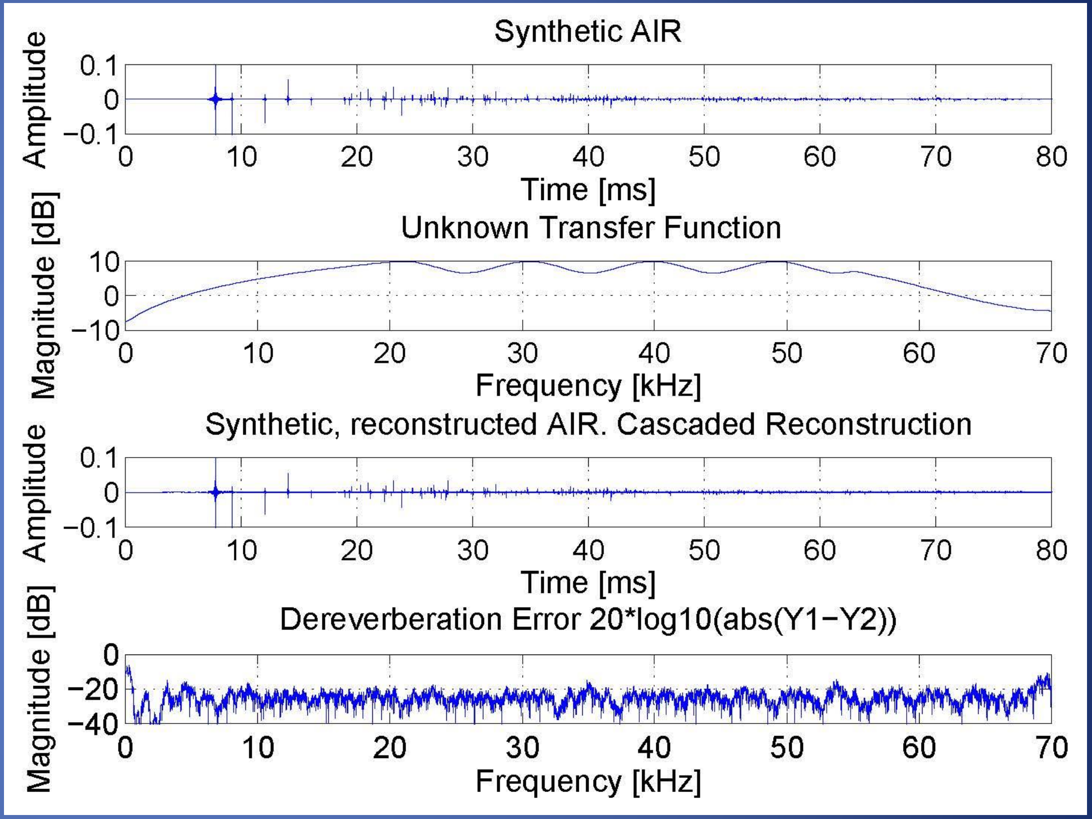
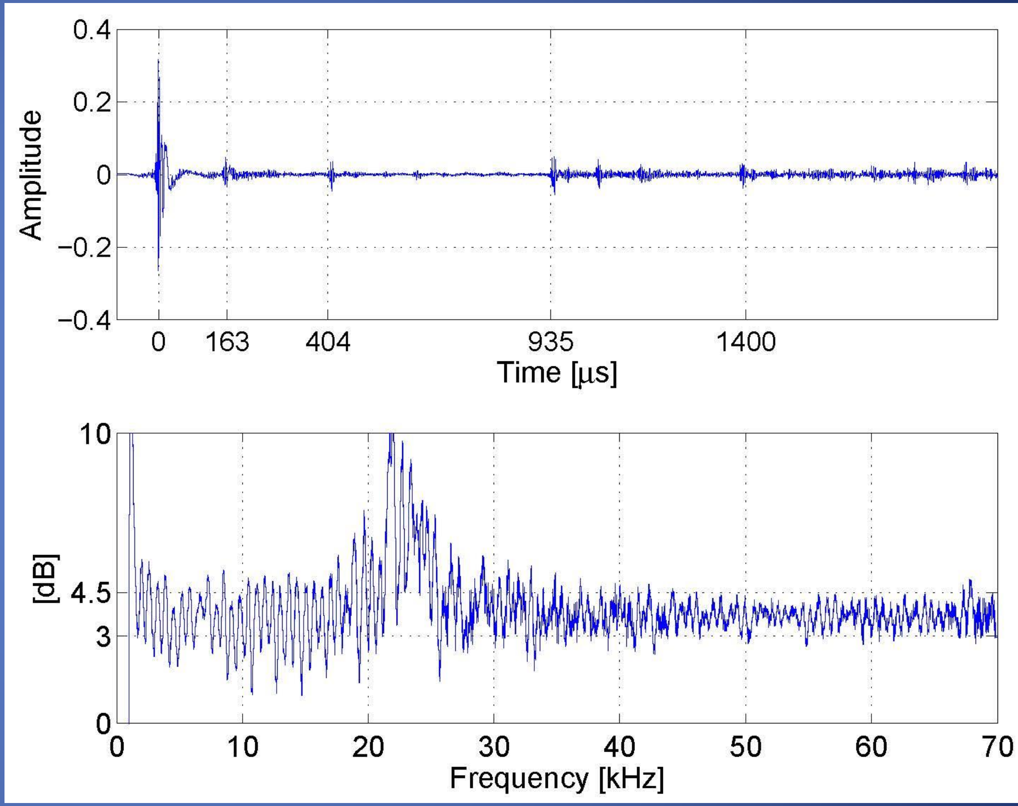


Fig.: Scaling process and error

g(t) estimate

- Time domain reflection correlate well
- |G| is smoothed on the decibel scale (251 points)
- What happens at 23 kHz?



Inversion using Least-Squares (LS)

- Provide only approximate equalization by minimizing the squared error
  - Only partially equalize spectral nulls which reduces narrowband noise amplification
  - Less sensitive to noise and inexact IR estimates
- LS inverse filters are very long and non-causal
  - Equalization results improve significantly when using a delay (processing delay)
  - Length of filter depends on reverberation time, sampling rate, delay (> 20k coefficients)
  - IR is non-minimum phase for reflection coefficients > 0.4 (reflection coefficient for water/air boundary > 0.4)
- The single channel least-squares formalism is extendable to a multichannel equalization method. The non-minimum phase problem is eliminated and if there are no common zeros, exact equalization can be achieved. (MINT = Mutiple-input/output INverse Theorem).

The Spike Filter

Equalize a channel with the inverse filter f of delay m	$h(t) * f_m(t) = \delta(t-m)$
Impulse response	$h = [h_0\ h_1\ h_2 \cdots h_{n-1}]^T$
Inverse filter	$f = [f_0\ f_1\ f_2 \cdots f_{n-1}]^T$
Spike filter [n+m-1]	$z = [0\ 0 \cdots 1]^T$
Best approximation	$\hat{f} = [\hat{f}_0\ \hat{f}_1\ \hat{f}_2 \cdots \hat{f}_{n-1}]^T$

$$\hat{f} = arg \min_f ||Hf - z||_2^2$$

$$\hat{f} = [H^T H]^{-1} H^T z$$

$$H = \begin{bmatrix} h_1 & 0 & \cdots & 0 \\ h_2 & h_1 & \cdots & 0 \\ \vdots & h_2 & \ddots & \vdots \\ h_{n-1} & \vdots & \ddots & 0 \\ 0 & h_{n-1} & \vdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \ddots & h_{n-2} \\ 0 & 0 & 0 & h_{n-1} \end{bmatrix}$$

- In practice, two parameters are varied to minimize the error:
  - Length of the impulse response [n]
  - Delay of the spike [m]
- Dimensions of H are (n+m-1) x (n)
  - Circulant Toepliz structure (inverse is positive definite and symmetric)

Inversion Performance

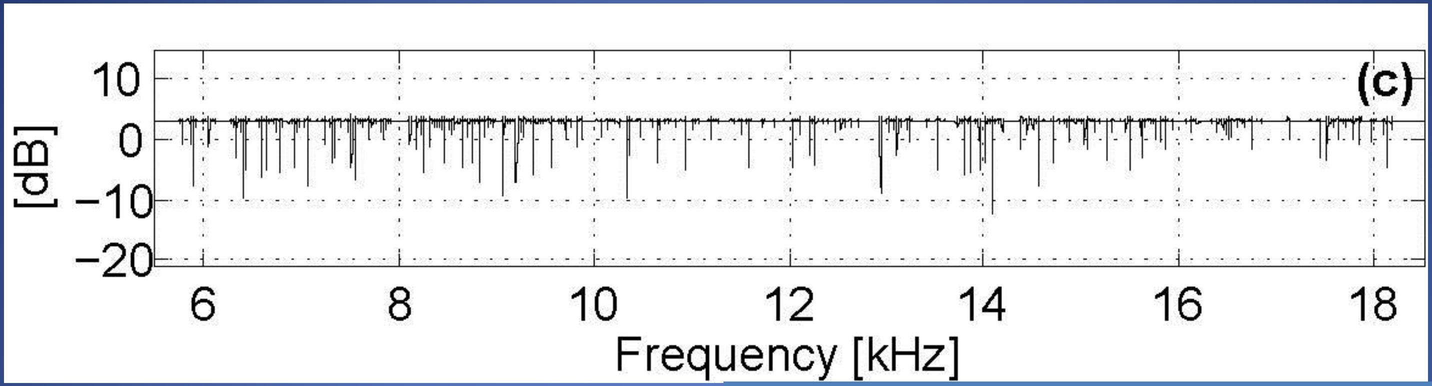
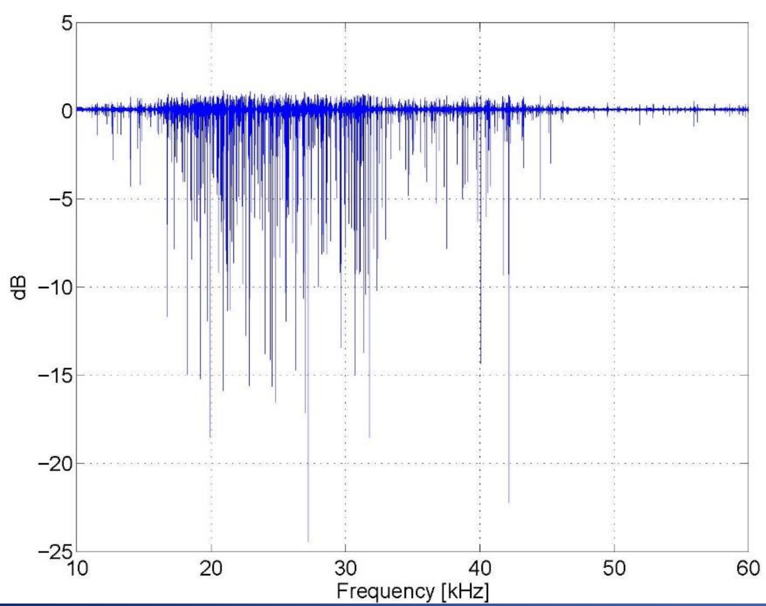


Fig.: Channel equalization using Lubell Speaker



- Equalization of spectral zeros seems to be a problem
- Performance might improve by minimizing the dynamic range of the IR
- Transmitting transducer has largest range
- CR1 is ‘optimum’ for 35-60kHz band
- Perform inversion over sub-bands?

Fig.: Channel equalization using CR1 Transducer