Characterization of underwater acoustic sources recorded in reverberant environments (with application to SCUBA signatures)

K.L. Gemba

Marine Physical Laboratory
Scripps Institution of Oceanography
gemba@ucsd.edu

Presentation Outline

- Problem motivation and formulation
- Proposed source characterization method
  - Estimating the channel's impulse response (IR)
  - Inverting the dynamics of the IR
  - Validation experiment
  - Results (inversion & recovery of control signals)
- (Application: SCUBA characterization)
- (Shortcomings and future research)
- (Estimating and removing colorations from the deconvolved IR)

Motivation

- Characterization of underwater sound sources is an important pre-requisite for many applications, including detection, classification, and monitoring
- It helps to know the characteristics of the signal which you are looking for or are interested in
- Anechoic facilities (minimize noise and boundary reflections) are generally not available
- Goal: Develop a robust and practical method for characterizing sources in reverberant environments

Prior Characterization Approach

- Measure spectral pressure at random locations adjusted by the theoretical reverberant energy
  - Yields spatial mean spectral levels
  - Reported errors for selected sinusoids are < 6 dB
- Shortcomings:
  - Limited to an incoherent estimate
  - Levels can vary significantly in space and in frequency

1

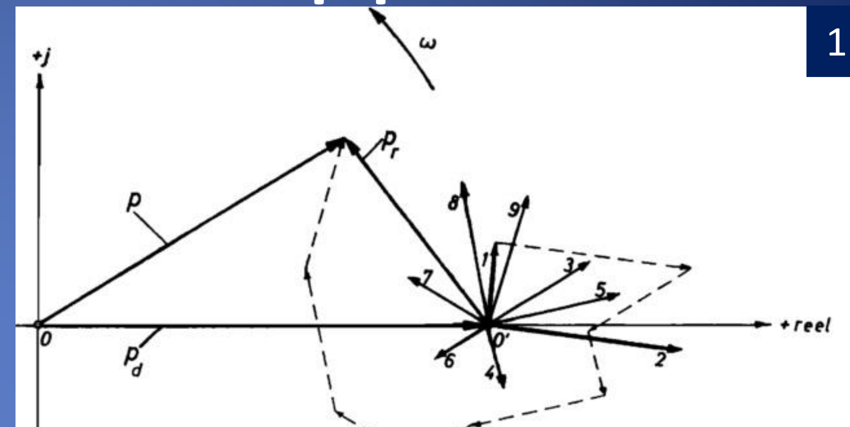


Fig.1: Random combination of reverberant sound pressure components

2

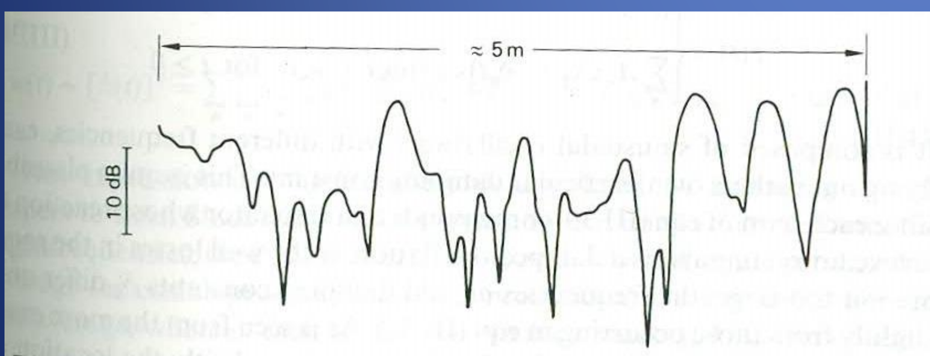


Fig. III.9. Logarithmic record of the steady state sound pressure level in a small lecture room at 1000 Hz; abscissa is the location of the measuring microphone on a certain line.

2

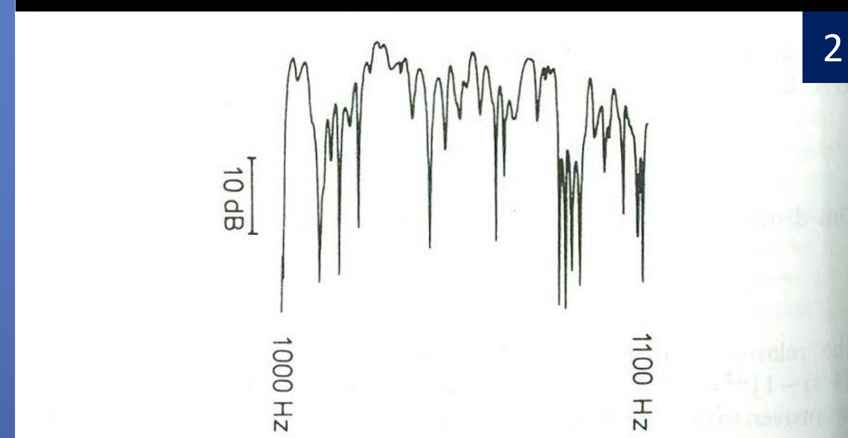


Fig. III.8. Logarithmic record of the sound pressure level at steady state conditions from 1000 to 1100 Hz (frequency curve), measured in a lecture room.

1. Diestel, H. G. (1963). Prob. Distribution Pressure of Sinusoidal Sound in a Room. *J. Acoust. Soc. Am.* 35
2. Kuttruff, K. H. (2000). Room Acoustics, 4th edition (Taylor & Francis, London), 368.

The Reverberant Environment

- No procedure exists for source characterization in a reverberant environments to estimate the signals:
  - Time domain waveform
  - Source Spectral Levels (SSL)

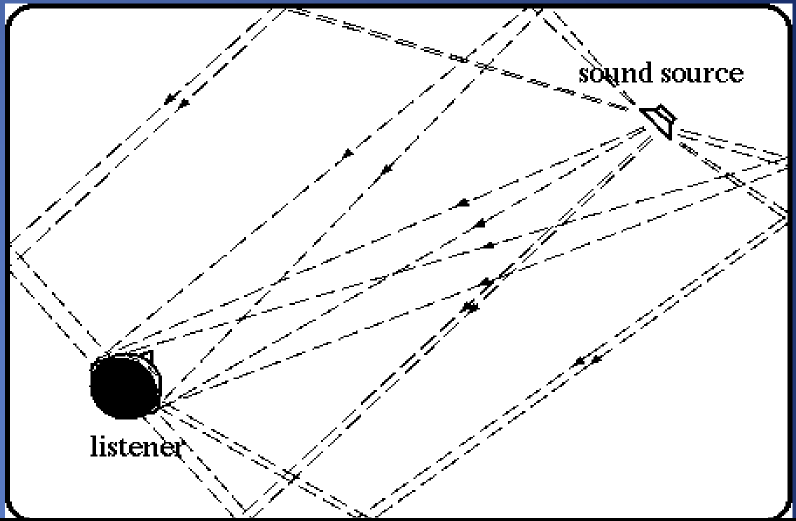


Fig.2: Reverberant environment showing multipath from a source to a listener

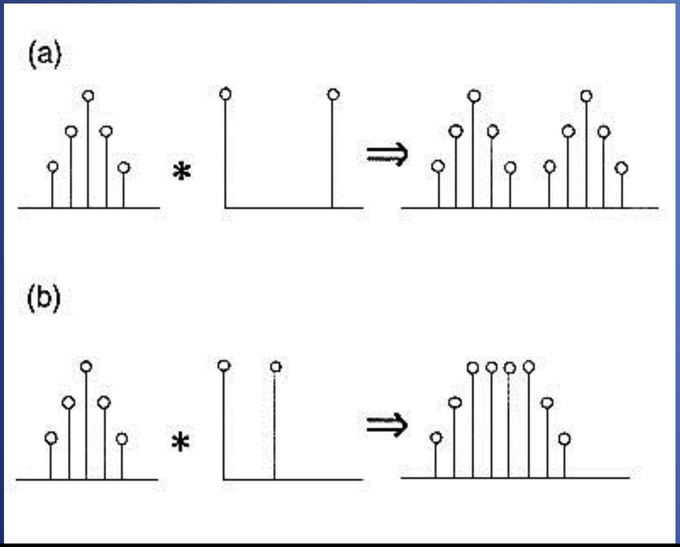


Fig.3: Convolution of a signal with 2 multi-path arrivals (direct arrival and echo)

- Reverberation degrades the recorded source signal by overlapping events
- This degradation is modeled by convolving the source signal with a map of reflections
- If we can estimate this map, we might be able to invert it and remove the overlapping events

Problem Formulation

- Source levels recorded in reverberant environments do not reflect true levels of the recorded source
- A solution is to estimate the impulse response (IR) of the recording channel and remove reverberant energy by inverting the IR (deconvolution)
- Assumptions:
  - Linear but not necessary time invariant system (need expectation)
  - Noise is stationary and uncorrelated
  - Sources have similar directionality
  - Sources have similar dimensions

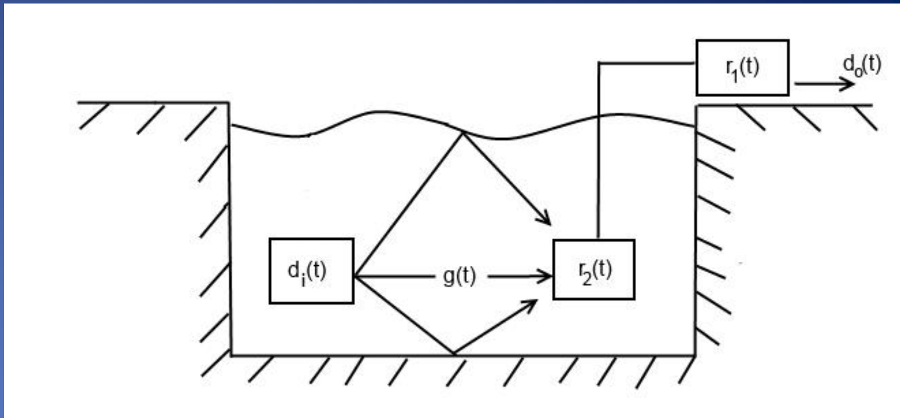


Fig.3: Pool diagram (inverse problem)

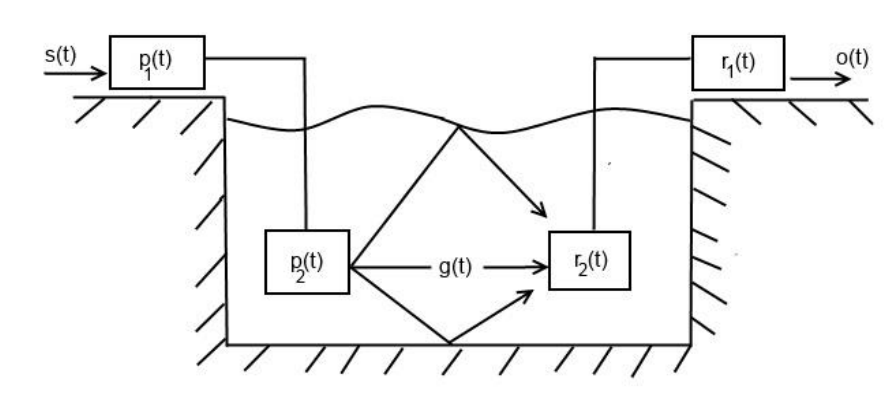


Fig.4: Pool diagram (forward problem)

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	$r_1(t)$, $r_2(t)$

Deconvolution Procedure: Overview

- 1. Estimate the acoustic impulse response h(t) of the linear stochastic system (we need to compute its expectation)
- 2. Invert the impulse response
- 3. Equalize the recording channel and recover an estimate (Magnitude and Phase) of a known signal
- 4. Quantify inversion / deconvolution performance
- 5. Recover source levels of an unknown signal

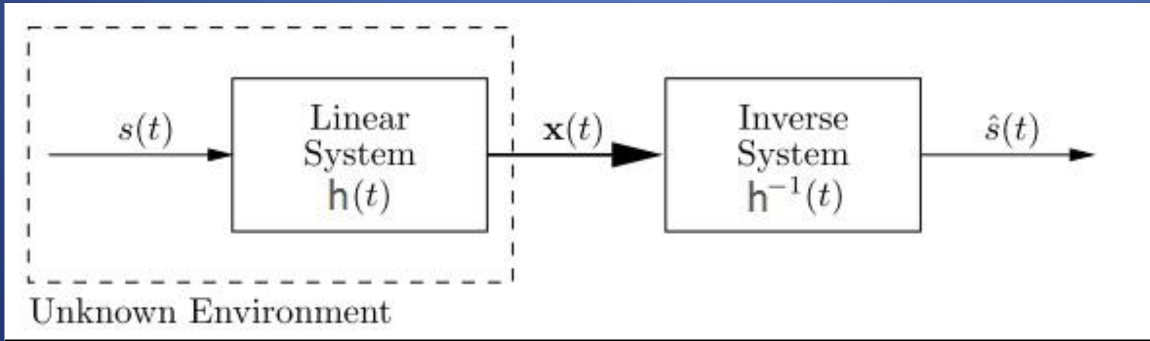


Fig. 5: Channel equalization and recovery of input signal (estimate)

Mathematical Formulation

Knowns: $o(t), s(t), |R_1(f)|, |R_2(f)|, |P_2(f)|, |P_1(f)|$
Unknows: $r_1(t), r_2(t), g(t), p_2(t), p_1(t)$

$$o(t) = r_1(t) * r_2(t) * g(t) * p_2(t) * p_1(t) * s(t)$$

$$h(t) = r_1(t) * r_2(t) * g(t) * p_2(t) * p_1(t)$$

$$o(t) = h(t) * s(t)$$

$$o(t) * f(t) = h(t)$$

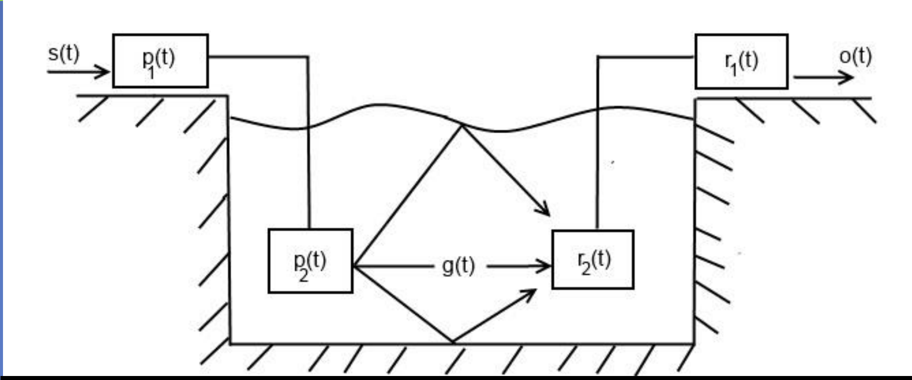


Fig. 6: Pool diagram (forward problem)

- The IR is deconvolved by cross-correlation (the matched filter detects changes in amplitude and phase of the system under investigation)
- The excitation signal should not affect magnitude or phase of the combined system (denoted by h(t))
  - The autocorrelation of the input signal must be a delta function
    → This is all in the inverse filter f(t)
  - Phase contribution is a pure delay

Mathematical Formulation

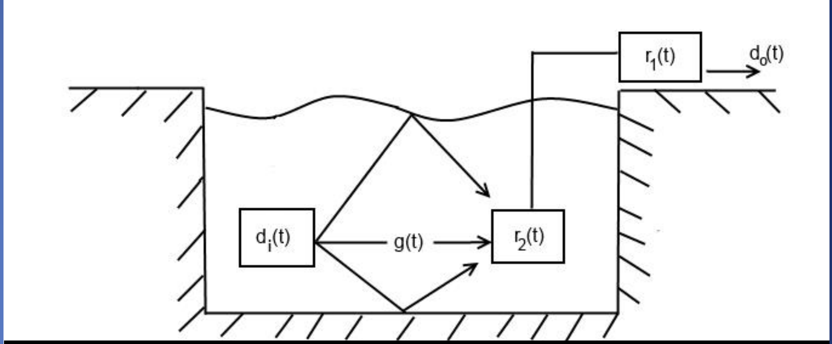


Fig. 7: Pool diagram (inverse problem)

Forward problem: performance

$$h(t) = r_1(t) * r_2(t) * g(t) * p_2(t) * p_1(t)$$

$$o(t) = h(t) * s(t)$$

$$o(t) * E[h^{-1}(t)] = h(t) * s(t) * E[h^{-1}(t)] = \hat{s}(t)$$

Inverse problem solution:

$$d_o(t) = r_1(t) * r_2(t) * g(t) * d_i(t)$$

$$d_o(t) * E[h^{-1}(t)] = \hat{d}_i(t) * [p_2(t) * p_1(t)]^{-1}$$

$$\hat{d}_i(t) = d_o(t) * E[h^{-1}(t)] * p_2(t) * p_1(t)$$

For an incoherent estimate, the PSD of the recorded diver $\boldsymbol{D_o(f)}$ is adjusted by the channel's IR and by the amplitude response of the playback equipment (in dB)

- Expectation denoted by E[ ]
- Recorded at 1m for source spectral levels

$$\left|\hat{\boldsymbol{D}}_{\boldsymbol{i}}(\boldsymbol{f})\right| = 10\log_{\mathbf{10}} \boldsymbol{D_o}(\boldsymbol{f}) - \mathbf{20}\log_{\mathbf{10}} \boldsymbol{E}[|\boldsymbol{H}(\boldsymbol{f})|] + \mathbf{20}\log_{\mathbf{10}}|\boldsymbol{P_2}(\boldsymbol{f})| + \mathbf{20}\log_{\mathbf{10}}|\boldsymbol{P_1}(\boldsymbol{f})|$$

Estimate h(t)

- System requires excitation
  - Excitation signal needs to be longer than the IR of the channel to avoid aliasing by circular deconvolution → approximated using the reverberation time (denoted by T_{60})
  - Excitation signal must be realizable (i.e. logarithmic and linear sweep)

$$T_{60} = \frac{24 \ \ln 10}{c} \frac{V}{-S \log(1 - \sum_{i=1}^{6} \frac{\alpha_i A_i}{S})}$$

- Estimate length of IR (decay into noise floor)
  - Schroeder's method of backward integration: ensemble average of the squared signal decay is equivalent to an integral over the squared impulse response:

$$< \boldsymbol{\psi^2(t)} > = \int_{t}^{\infty} [\boldsymbol{h(\tau)}]^2 \boldsymbol{d\tau}$$

Inverting the IR: Least-Square (LS) Method

- LS provides only approximate equalization
  - 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
  - IR is non-minimum phase for reflection coefficients > 0.4 (late reflection from water/air boundary)
  - Equalization results improve significantly when using a delay (processing delay)
  - Length of filter depends on reverberation time, sampling rate, delay (> 20k coefficients)
- 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 (Wave-Shaping 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)

Forward Experiment

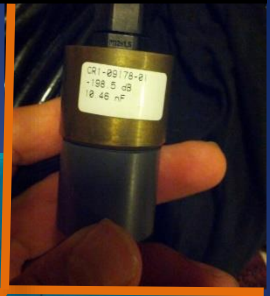


Fig.5: CR1 Transducer

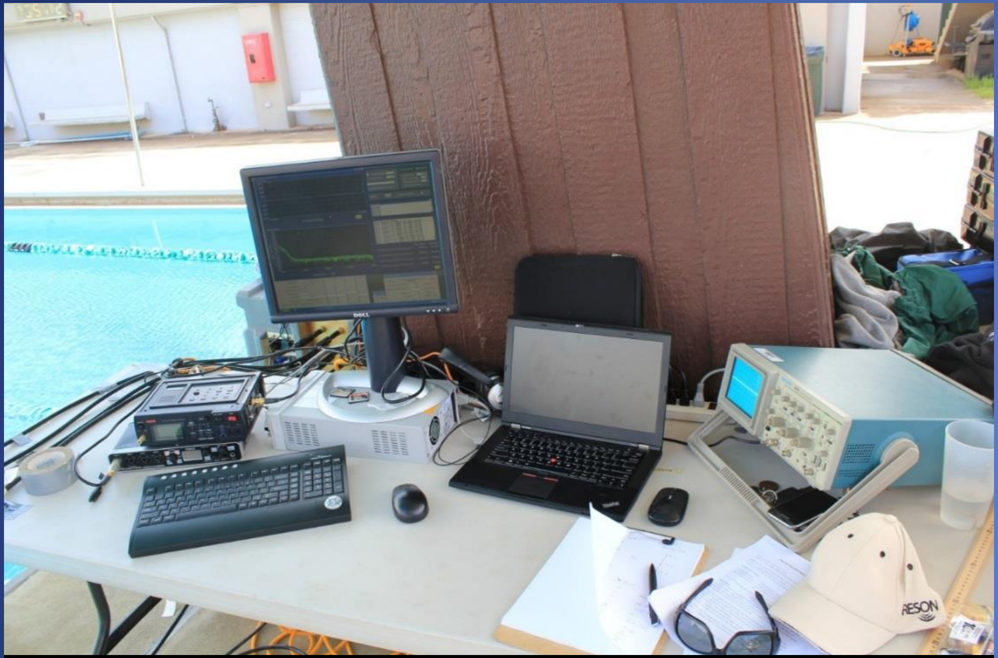


Fig.8: Playback and recording equipment

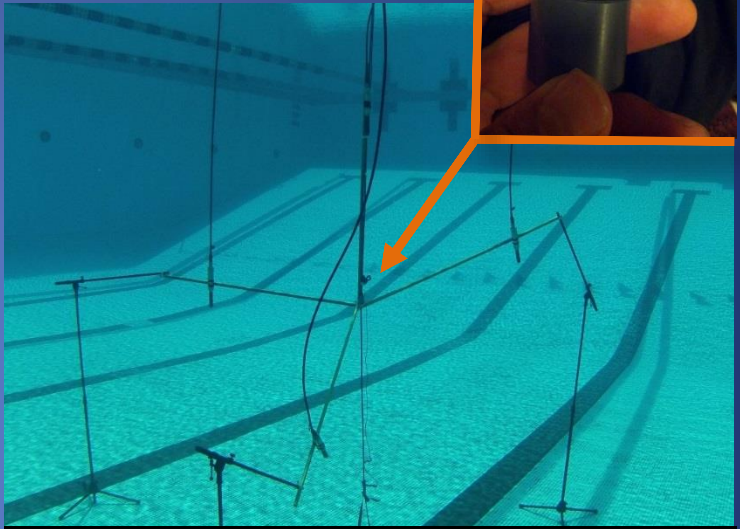


Fig.9: Spherical array

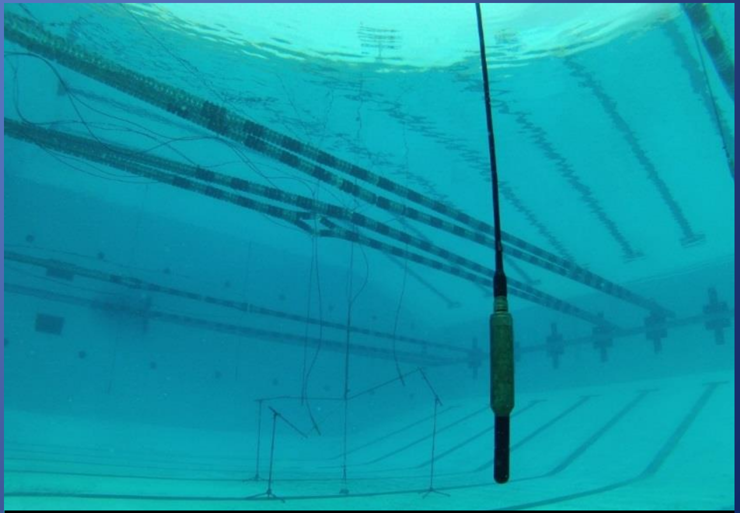


Fig.10: Random hydrophone

- 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 to multi-channel method
- 5 random far-field 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

Log Sweep

Sinusoid

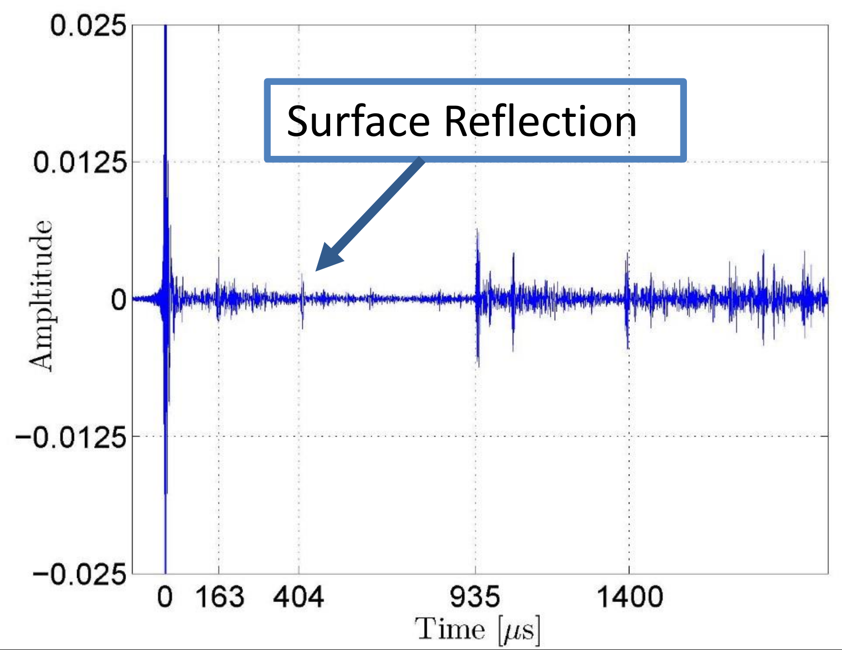


Fig.11: Impulse response h(t) with theoretical boundary reflection times

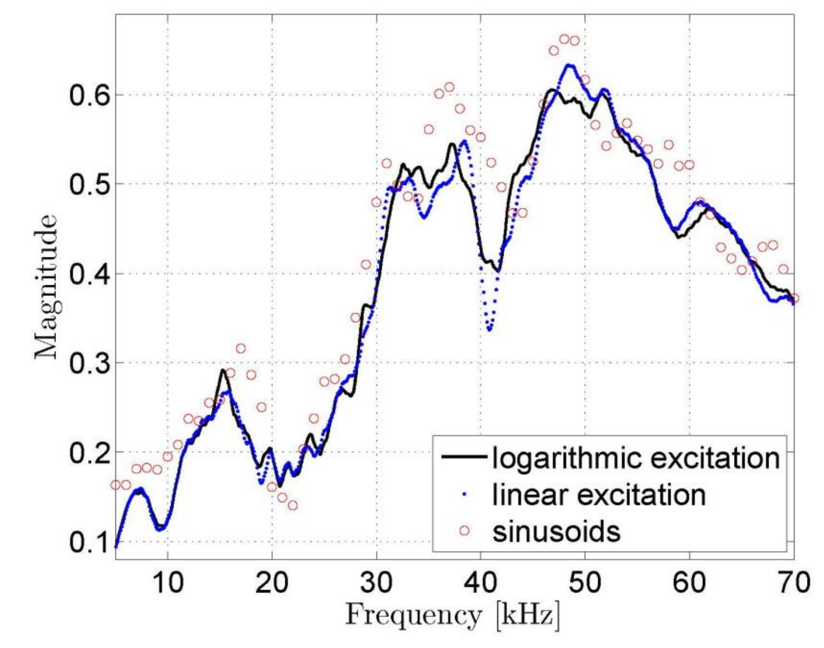


Fig. 12: Spectral comparison of excitation methods and sinusoids

Estimation of IR Length

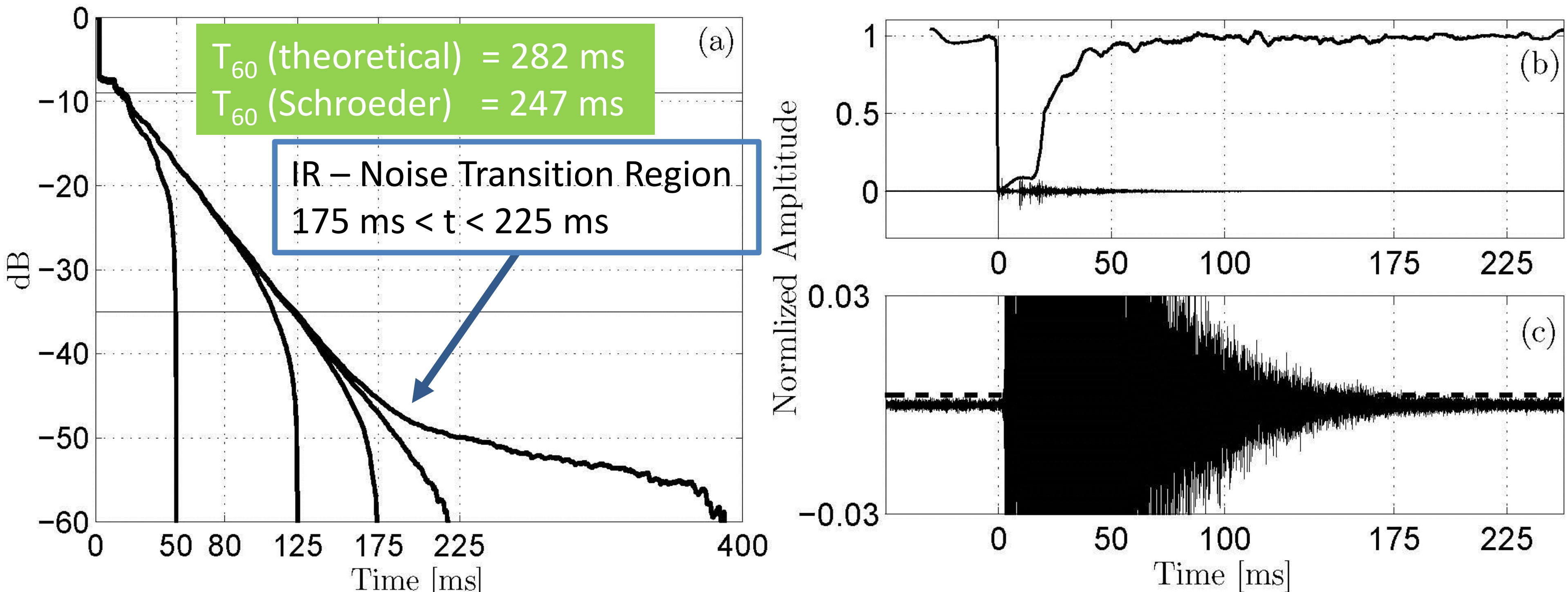


Fig. 13: a) Decay curves of the IR with subtracted noise average. The ticks on the x-axis show selected upper integration limits, the two horizontal lines (-9 dB and -35 dB) correspond to the range over which T_{60} is calculated. (b) IR with echo density (top trace) showing transition time from early reflections to late reverberations at approx. 80 ms. (c) Zoomed in ensemble averaged IRs of far field hydrophones aligned with respect to the direct arrival in (b) with dashed noise reference line showing decay into the noise floor at nominally 175

Inverting for the dynamics of g(t)

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

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

$$D = \hat{f} * h$$

$$\varepsilon_t = D(m)$$

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

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

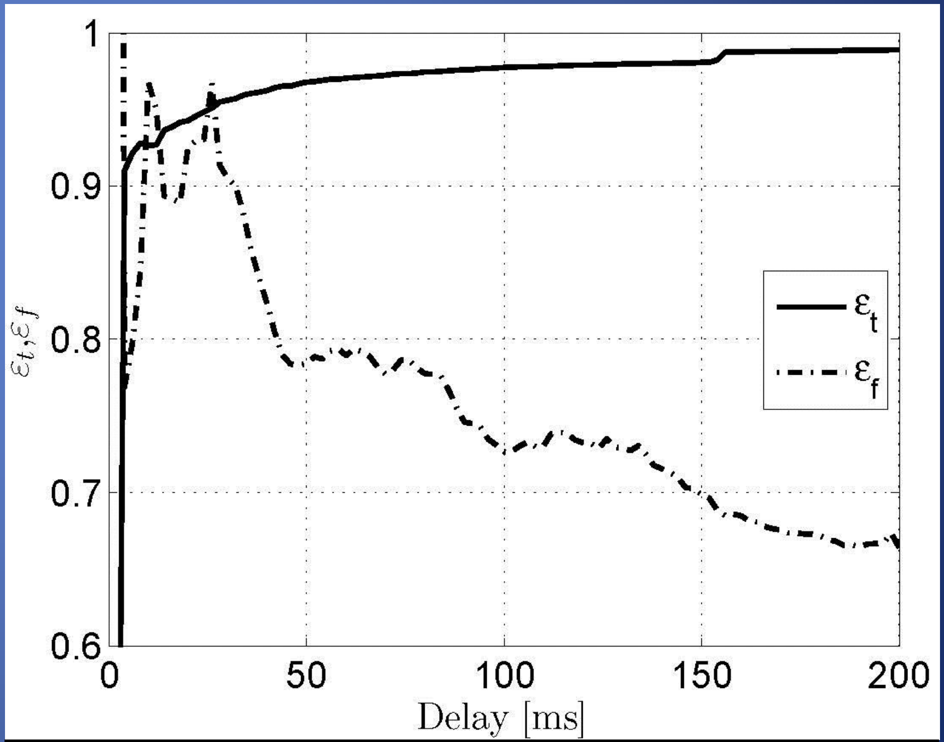


Fig.14: Inversion performance vs. delay for an IR length of 152 ms

Invert g(t) Frequency Performance: ε_f

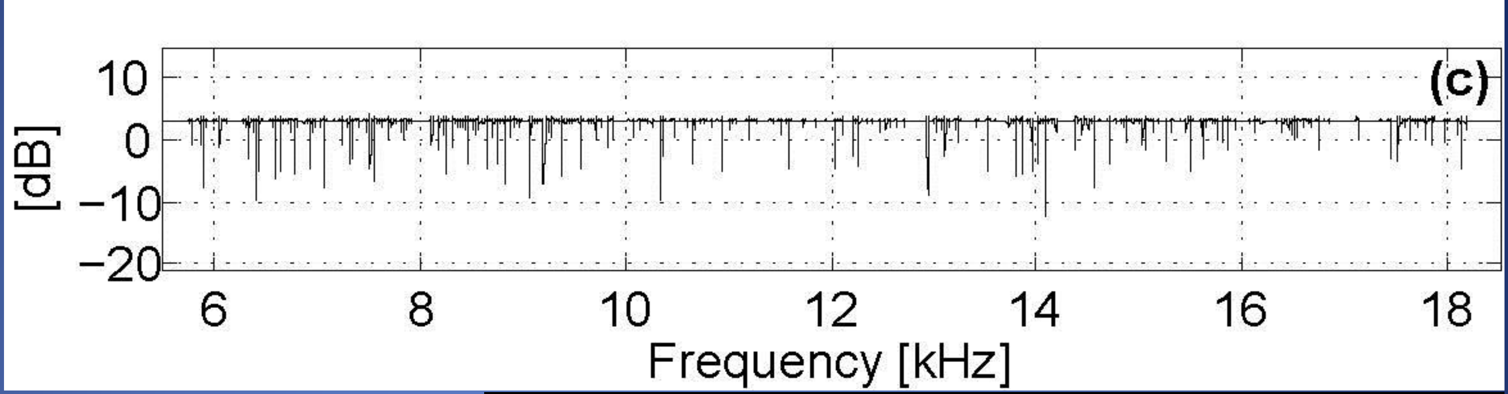


Fig. 15: Channel equalization using Lubell speaker

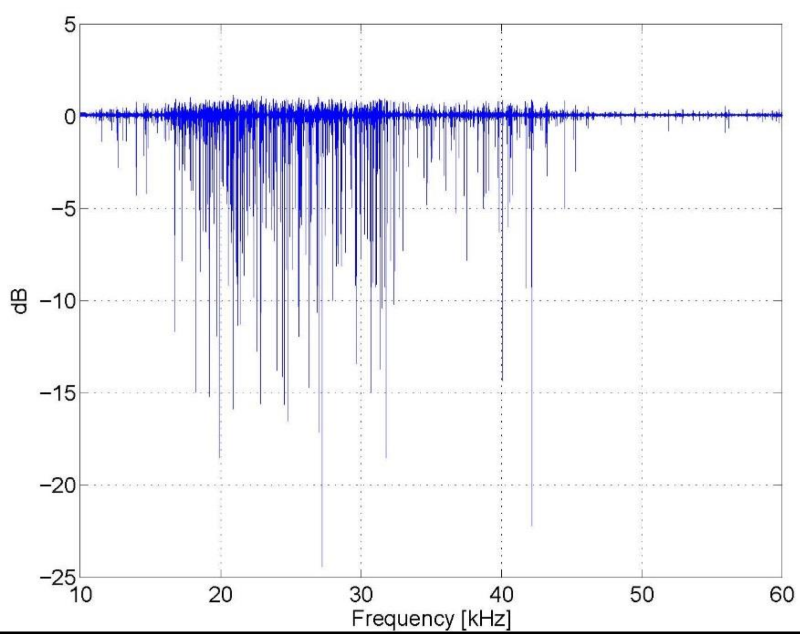


Fig. 16: Channel equalization using CR1 transducer

- Equalization of spectral zeros seems to be a problem
- Performance might improve by minimizing the dynamic range of the IR
- Transmitting transducer has largest dynamic range
- CR1 is 'optimum' for 35-60kHz band
- Note that the equalized signal for the Lubell speaker is not at 0 dB

Incoherent vs. "Coherent" Magnitude Performance

1. Forward Problem
2. The PSD of a recorded control signal $\mathbf{S_o}$ is adjusted incoherently by the ensemble average of the transfer function and smoothed with a zero-phase moving average filter M
3. For a "coherent" comparison, the PSD of a control signal is adjusted by the best estimate of the amplitude response of the inverse in the least-squares sense $|\widehat{\boldsymbol{F}}|$ and by a constant (mean of equalized signal).
4. The recovered signal $\hat{\mathbf{S}}_{\mathbf{s}}$ is then compared to the original signal, yielding objective performance results.

$$\mathbf{o(t) = h(t) * s(t)} \quad (1)$$

$$\mathbf{10\log_{10}\hat{S}_s = M[10\log_{10} S_o - 20\log_{10} \boldsymbol{E}[|H|]} \quad (2)$$

$$\mathbf{10\log_{10}\hat{S}_s = M[10\log_{10} S_o + 20\log_{10} \boldsymbol{E}[|\widehat{F}|] - 2\bar{\boldsymbol{D}}]} \quad (3)$$

Incoherent vs. Coherent Dereverberation Comparison

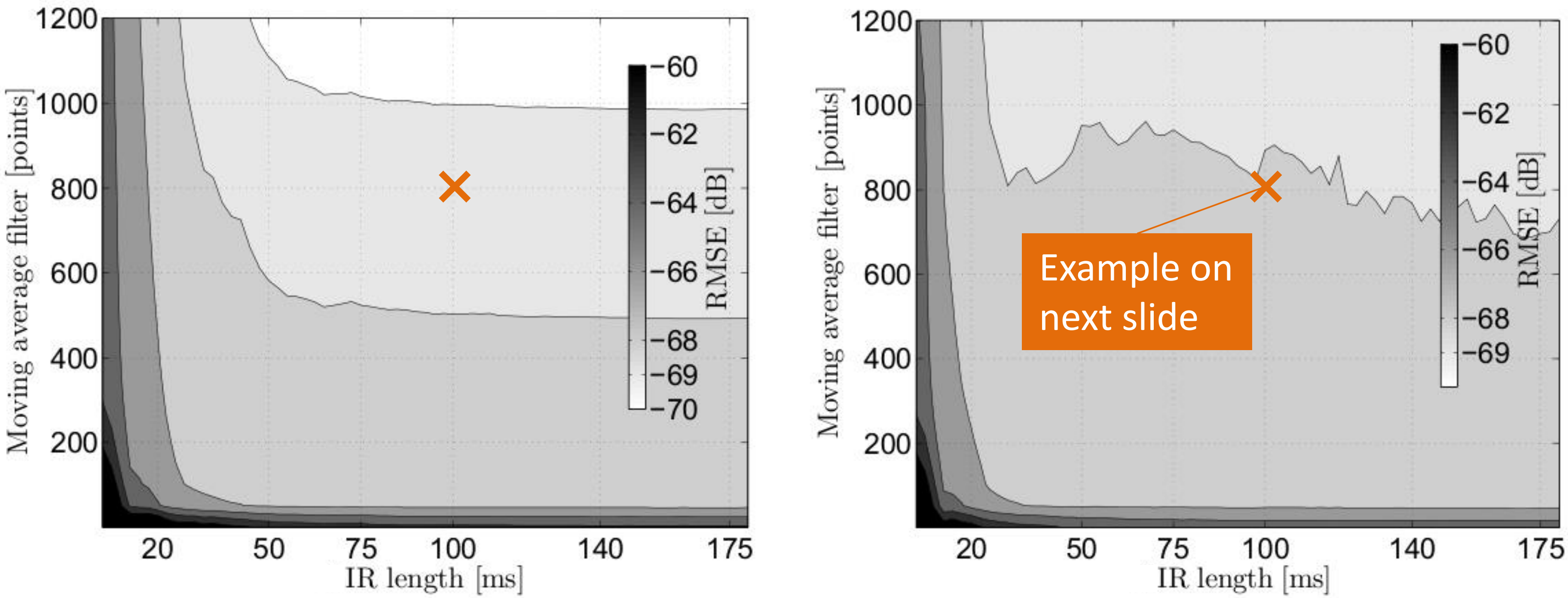


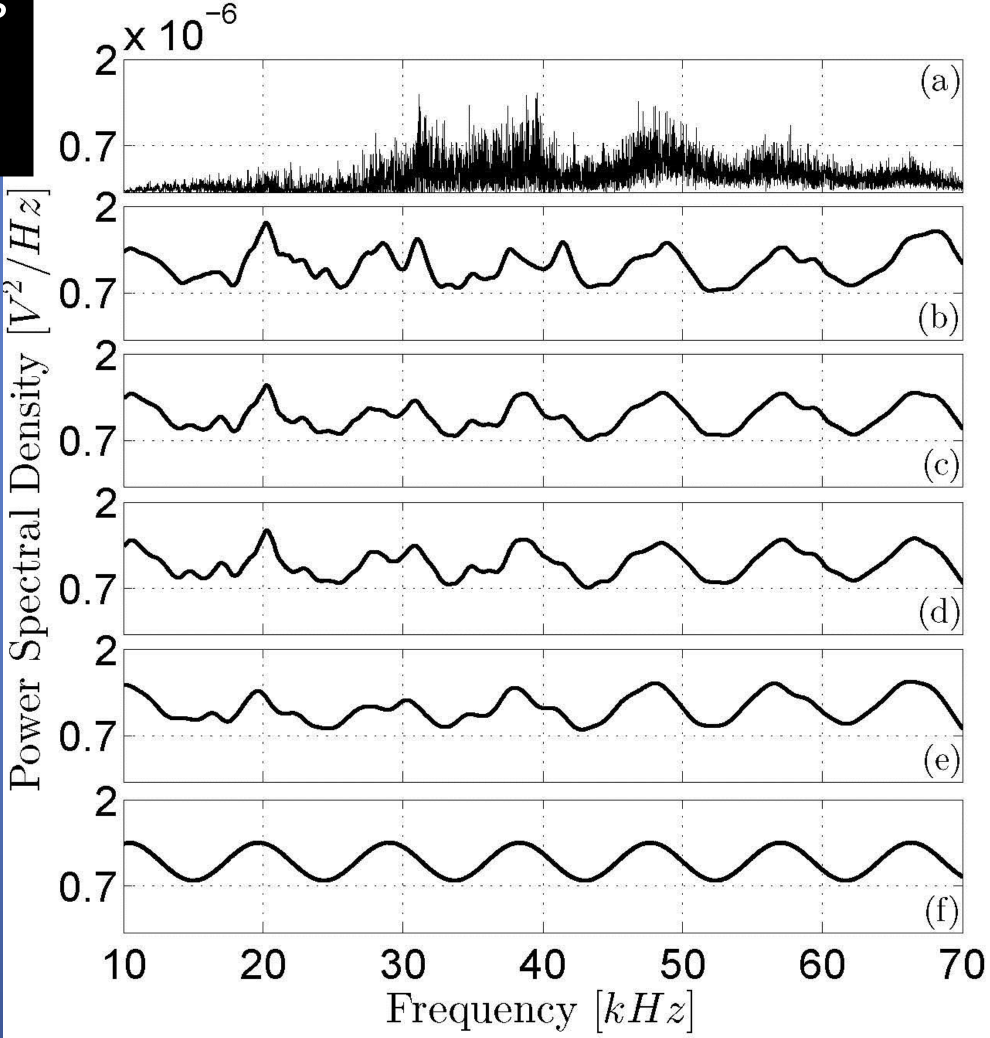
Fig.17: RMSE of dereverberated linear sweep: Incoherent (left) and coherent (right) inverse using 10 realizations. RMSE ticks correspond to contour surfaces

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

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

Dereverberation Example

Fig.18: 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 50 TF
e) Coherent adjustment with 10 TF
f) Original linear sweep

- Optimum frequency range of transmitting transducer: > 35 kHz

Generalized procedure*

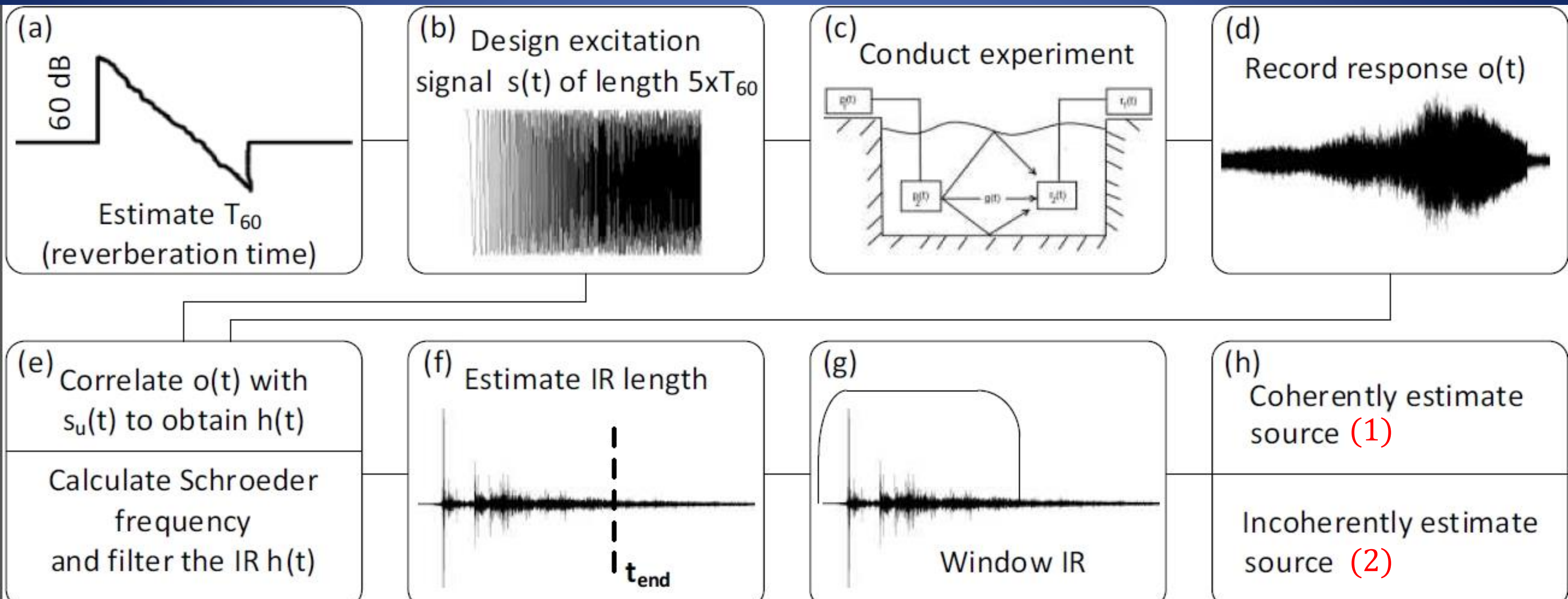


Fig. 19: Flowchart showing the proposed procedure to obtain (a) - (g) the impulse response in the forward problem and (h) the estimate of the unknown source in the inverse problem.

(1) $\boldsymbol{d_i(t) = d_o(t) * E\left[\hat{f}(t)\right] * p_2(t) * p_1(t)}$

(2) $\boldsymbol{|D_i(f)| = M[D_o(f) - E[|H(f)|\,] + |P_2(f)| + |P_1(f)|]}$

Incoherent results correspond to SSL if recorded at 1m and should be stated as:

$\boldsymbol{|D_i(f)|} \pm \sigma\ [dB\ re\ 1\mu Pa^{\,2}/_{Hz}\ at\ 1m]$

*Gemba K L and Nosal E-M "Source characterization using recordings made in a reverberant underwater channel", Manuscript submitted for consideration for publication in Applied Acoustics.

Fig. 26: Mahalo!

Application: SCUBA Characterization

- Pool resonance freq.: < 300 Hz
- Hydrophone: ITC 6050-C
  - Test Distance: 2 m for both Diver and Lubell Speaker
- ADC: 192 kHz at 24 bits
  - 10 Hz digital high-pass filter

Fig. 20: Lubell Speaker LL916

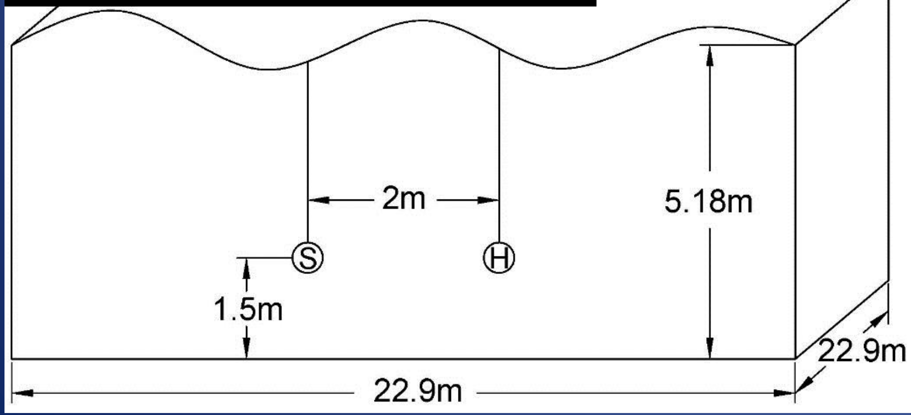


Fig. 21: Pool schematic

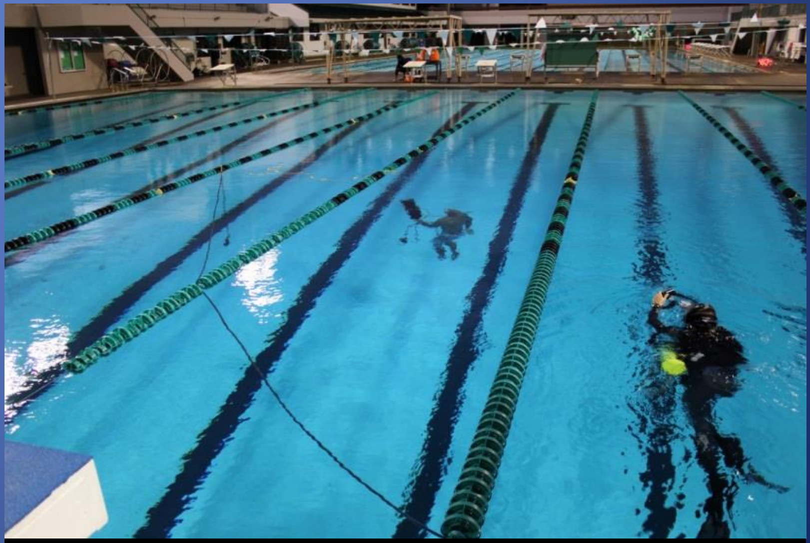
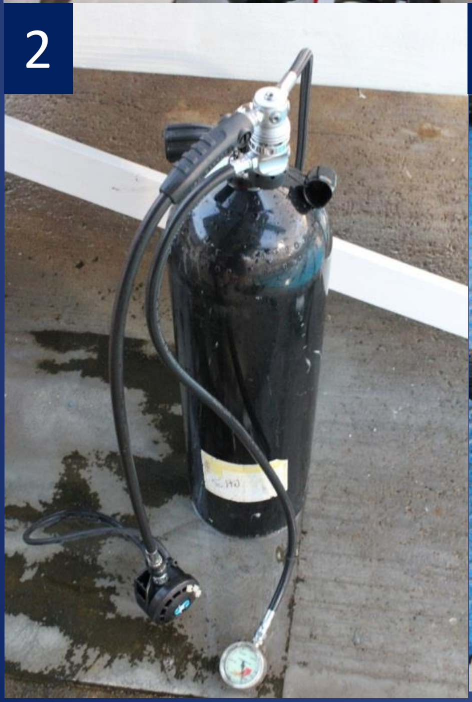


Fig. 22: Diving well, equipment setup

Tested Regulators

1. Oceanic SP-5, unbalanced Piston
2. ScubaPro MK25, balanced Piston
3. Apeks XTX 200, balanced Diaphragm
4. Royal Mistral, unbalanced Diaphragm, and only Single Stage Design

1


2


3


4


Apeks Regulator*

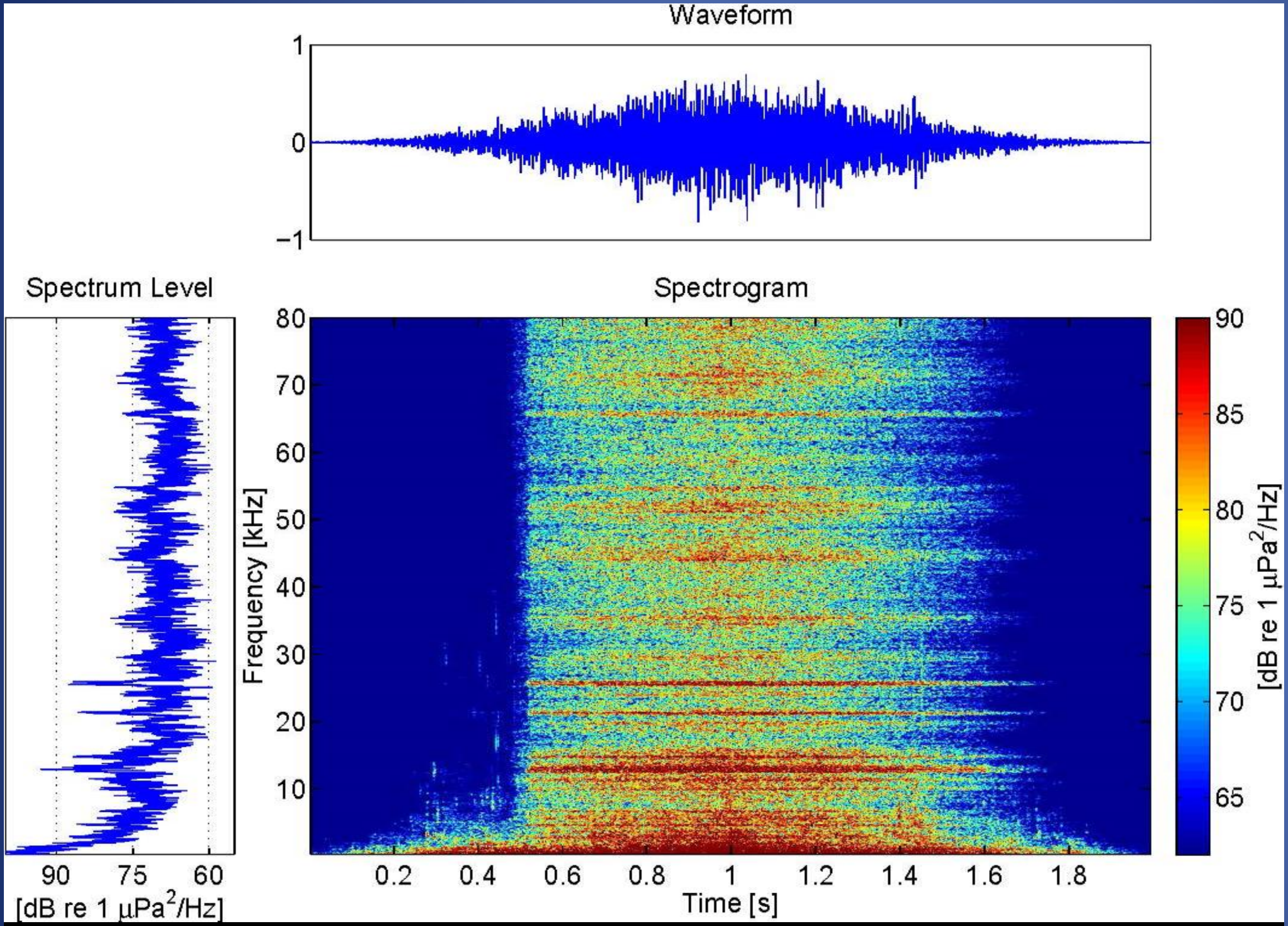


Fig. 23: Spectrogram (2048 frequency bins), Spectrum Levels and Waveform of Apeks Regulator

10 sec Recording, 0.3 kHz – 22 kHz

10 sec Recording, 3.5 kHz – 22 kHz

*Gemba K L, Nosal E-M and Reed T R (2014). Partial dereverberation used to characterize open circuit SCUBA signatures. J. Acoust. Soc. Am., 136(2), 623-633

Combined vs. Demand Signature

- Band: 0.3-80kHz (combined signature)
  - Mean SPL: 130 dB re 1μPa at 1m
  - Breathing range: 16 dB
- Band: 6-80kHz (demand signature)
  - Mean SPL: 127 dB re 1μPa at 1m
  - Breathing range: 22 dB
- Bubbles carry a lot of energy: a diver would choose a closed system to avoid detection

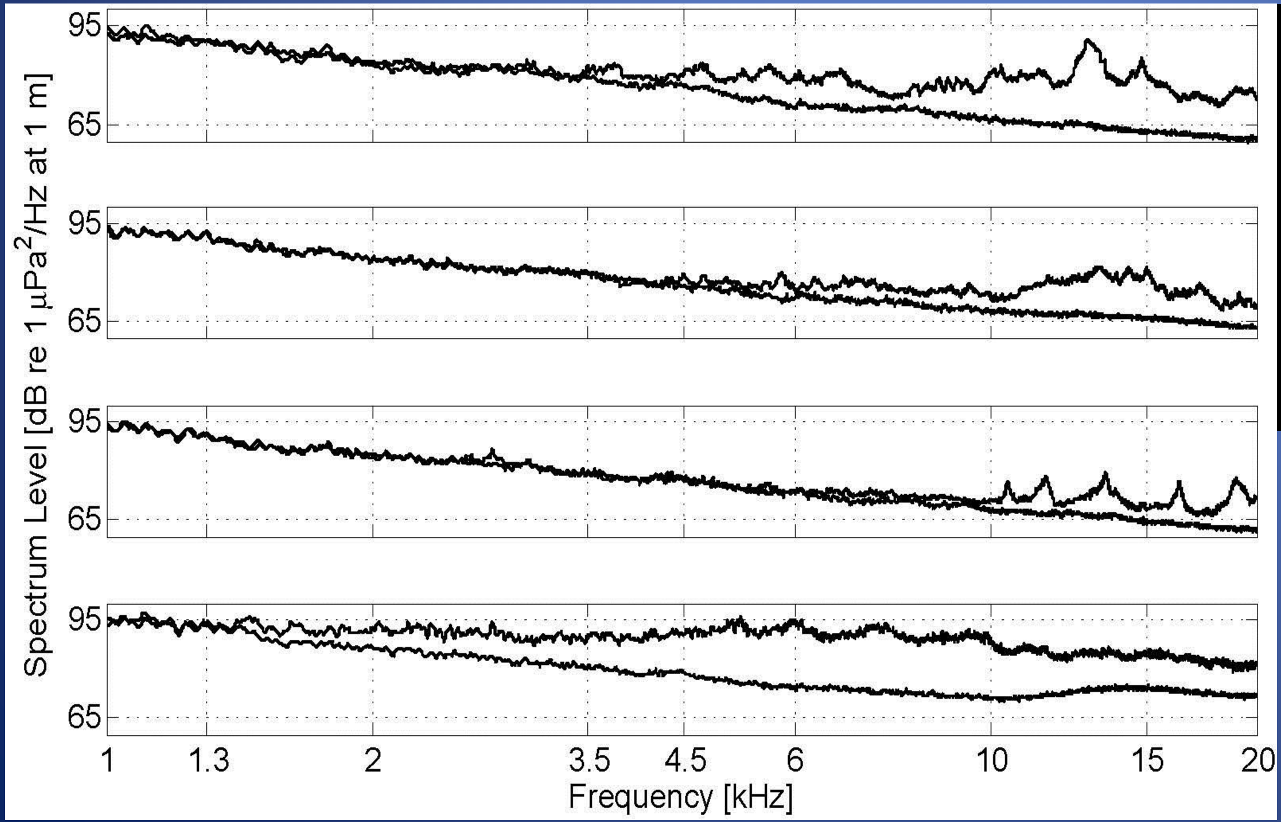


Fig.24: 4 Regulator spectra (Apeks, Oceanic, ScubaPro and Mistral from top to bottom) showing their signature transition from the exhale signature

Coherently Adjusted SCUBA Signal

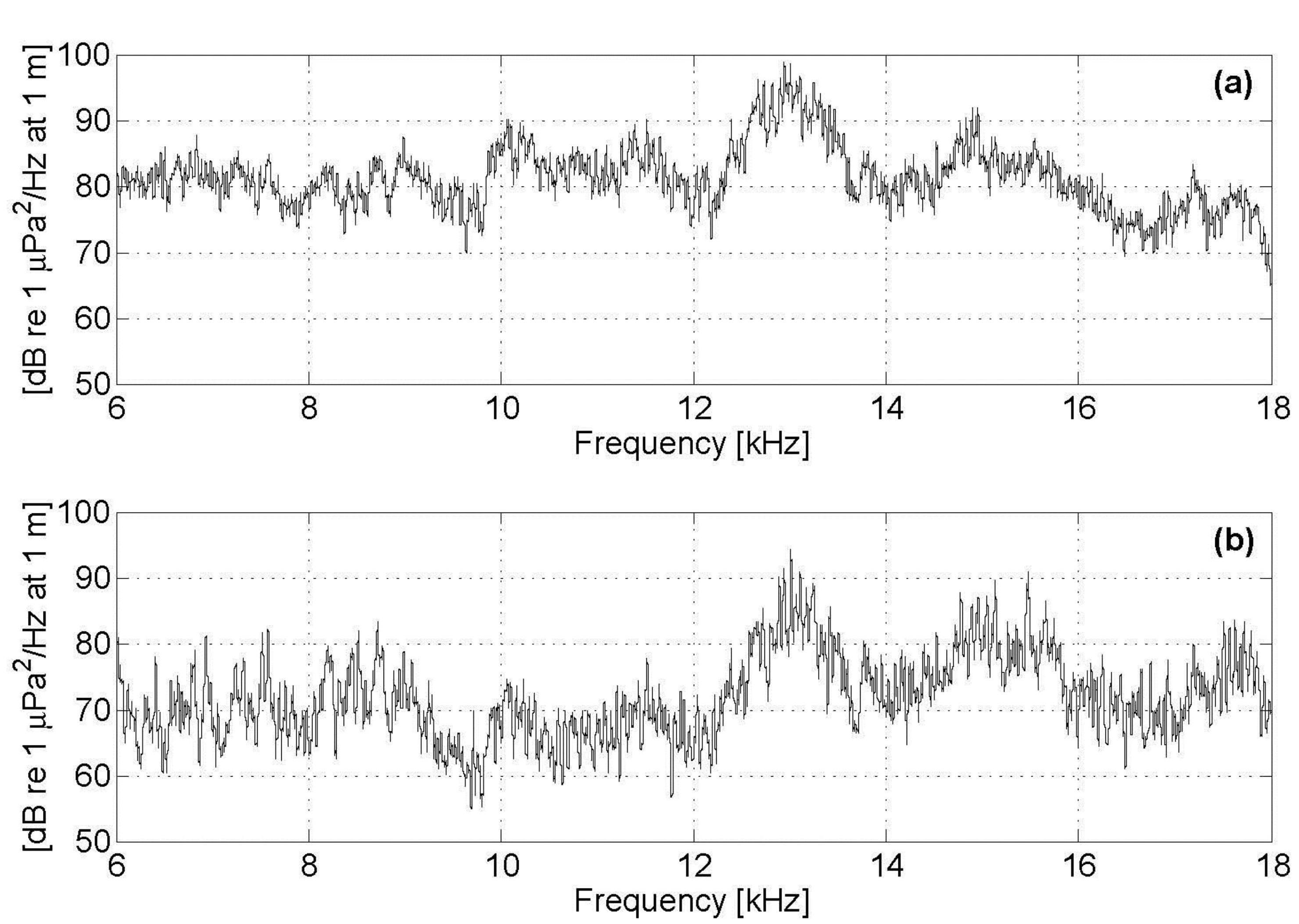


Fig. 25: (a) Original signal and (b) coherently adjusted signal with impulse response of the recording channel. The spike at nominally 13 kHz seems to be part of the equipment.

Regulator SPL

TABLE II. Sound Pressure Levels [dB re 1μPa] for all tested SCUBA regulators

Regulator	Combined SPL (0.3 kHz to 80 kHz)			Regulator SPL (6 kHz to 80 kHz)		
	Mean	Minimum	Maximum	Mean	Minimum	Maximum
Apeks (47)	130.4	127.8	133.3	127.2	124.5	131.1
Oceanic (51)	130.2	125.0	133.5	128.0	112.4	132.2
Scubapro (44)	130.4	125.5	133.6	127.3	110.4	132.5
Mistral (9)	134.9	133.6	135.9	131.4	129.8	132.3

Regulator	Sub-band original SPL (6 kHz to 18 kHz)			Adjusted SPL with AIR (6 kHz to 18 kHz)		
	Mean	Minimum	Maximum	Mean	Minimum	Maximum
Apeks (47)	122.4	119.8	124.7	116.1	113.2	119.0
Oceanic (51)	119.4	109.6	128.5	113.9	101.9	123.0
Scubapro (44)	116.0	108.2	120.0	109.6	101.6	114.4
Mistral (9)	129.7	128.1	131.1	122.9	121.0	124.3

- Signal Length: 1.5 s
- About half the energy is carried by the bubble signature
- Reverberation adds about 6 dB [6kHz to 18 kHz]
- Variation due to breathing: 16 dB

Shortcomings and future research

- Only a first step towards a practical method to characterize an unknown acoustic source
  - Dimensions of transmitting transducer and unknown source are likely not the same
  - Directionality is likely not similar either
- How well does the forward problem translate to the inverse problem? (need anechoic comparison)
- Coherent formulation needs further improvement
- The expectation applies also to the test signal, but was not yet investigated (this will improve results)
- This method can be extended from a single to a multi-hydrophone method

Pressure Distribution

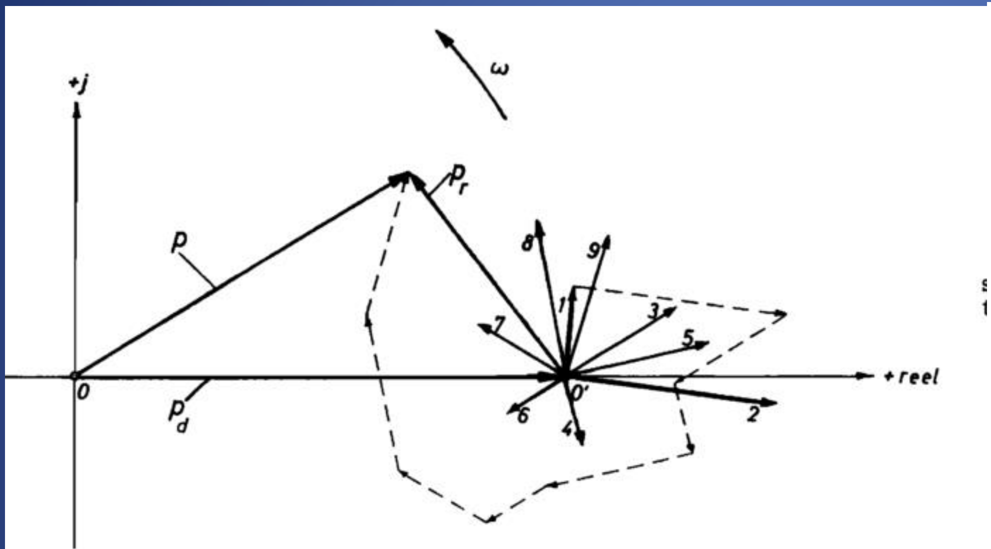
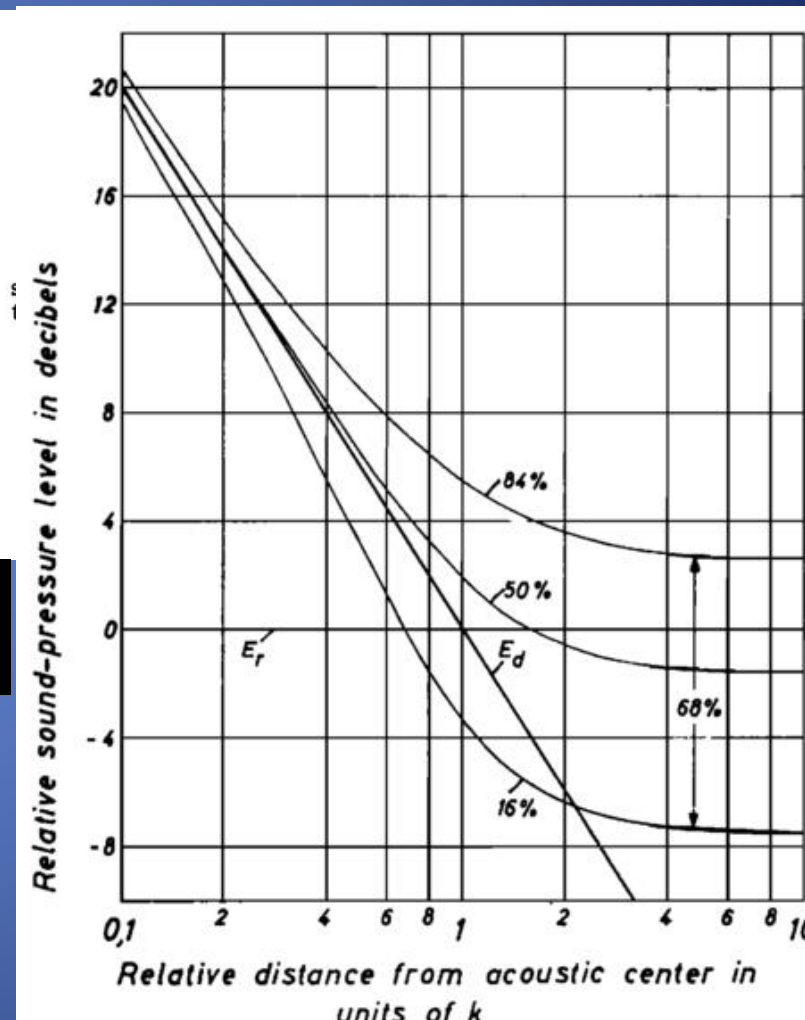


Fig. : Combination of the sound pressure in the complex time-vector plane



- Steady state excitation?
- Time variance of the system

Estimating and Removing Colorations

$$h(t) = r_1(t) * r_2(t) * g(t) * p_2(t) * p_1(t) = g(t) * u(t)$$

- The deconvolved IR of the channel g(t) is colored by the electrical equipment r(t) and p(t).
- Even if the magnitude response of individual equipment is known, connected equipment might behave differently (impedance) → treat coloration as unknown u(t).
- G(t) could be integrated into the passive sonar equation (incoherent) to remove reverberant energy of the channel and estimate SSL:

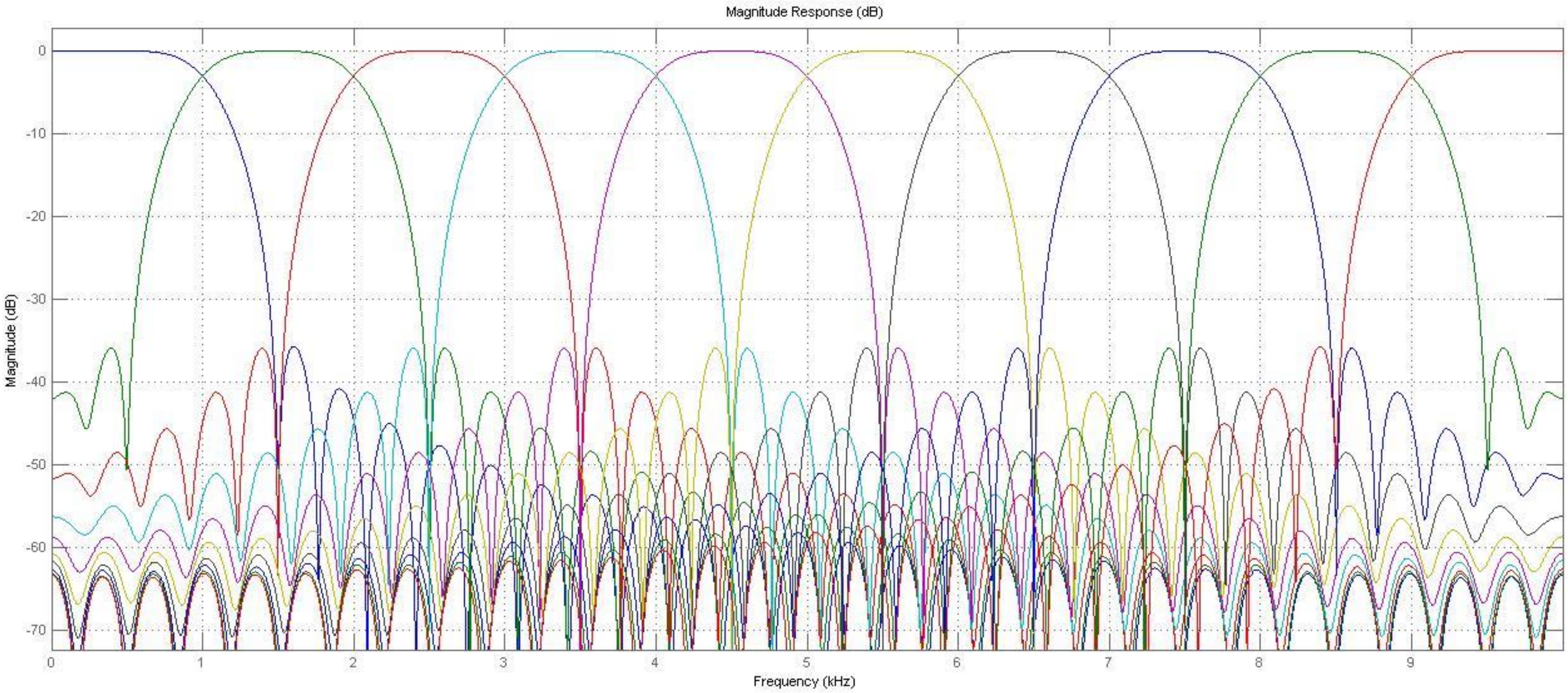
$$\boldsymbol{SSL[f]} = \mathbf{10}\log \boldsymbol{S_{xx}} + \mathbf{20}\log|\boldsymbol{M_h}| - \mathbf{20}\log|\boldsymbol{ADC}| + 20\log(R) + \boldsymbol{\alpha(R)}$$

$$\boldsymbol{SSL[f]} = \mathbf{10}\log \boldsymbol{S_{xx}} - \mathbf{20}\log|\boldsymbol{G}| + \mathbf{20}\log|\boldsymbol{M_h}| - \mathbf{20}\log|\boldsymbol{ADC}| + \cancel{20\log(R)} + \boldsymbol{\alpha(R)}$$

Approach to recover g(t)

- Observation: direct arrival in g(t) is a scaled delta function.
- All information about u(t) are contained in the direct arrival and can be removed incoherently.
- To estimate g(t), u(t) must be inverted.
  - Narrowband noise amplification? Causal?
- Proposed solution: Estimate coefficients of u(t) with a Pseudo-QMF bank and recover an estimate of g(t).

Pseudo-QMF Bank, 10 filters



Validation using Image-Source Model

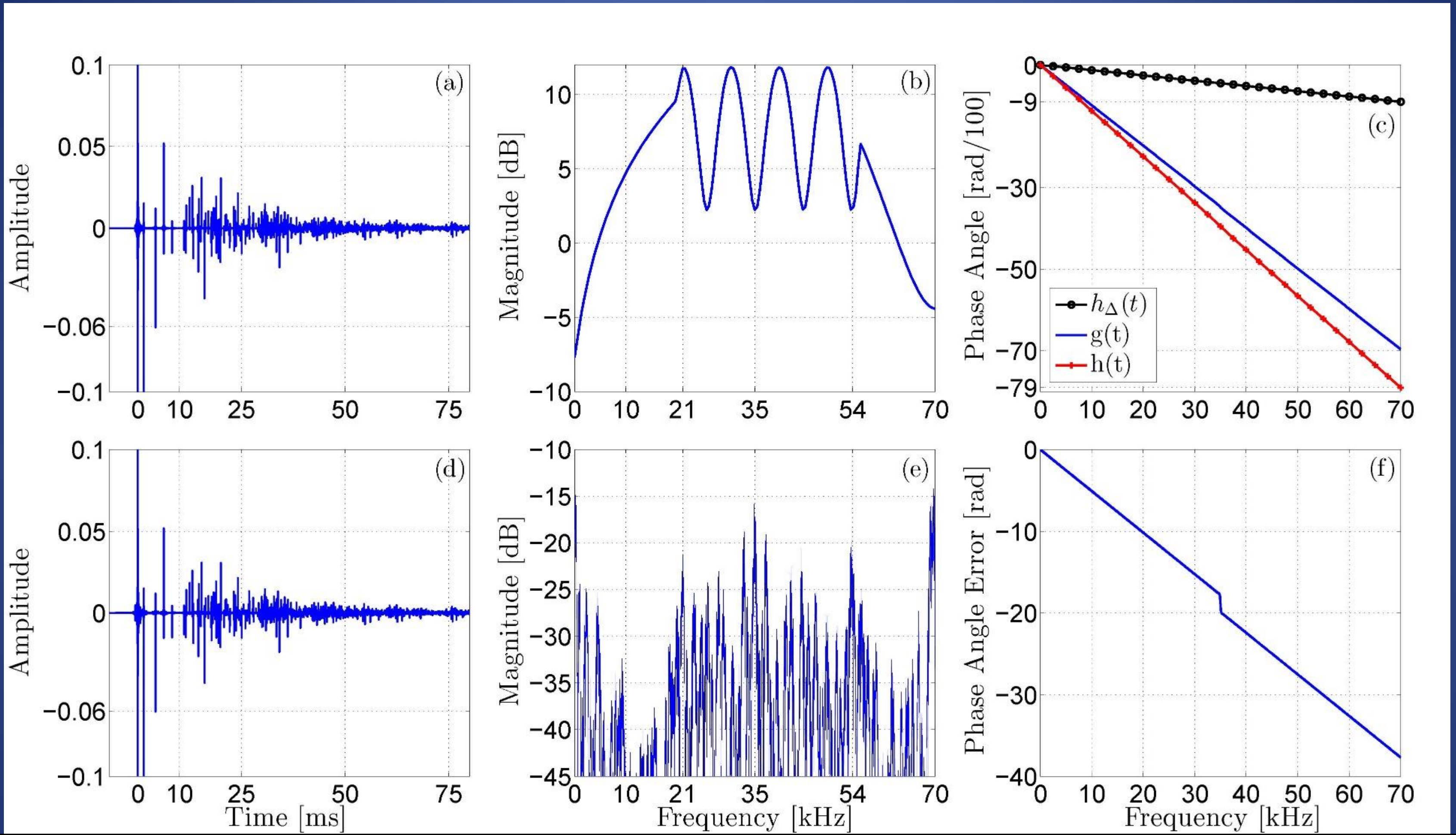


Fig.: (a) Synthetic IR g(t) obtained form the image source model sampled at 140 kHz and (b) |U(f)| of unknown transfer function. (c) Phase responses of g(t), h(t), and the all-pass filter using the direct arrival in h(t) (denoted by $h_\Delta(t)$). (d) Recovered IR $\hat{g}(t)$ and (e) error of recovered IR using cascaded process. The error between (a) and (d) is computed on the spectrum using 20log10($|\hat{G}(f)|-|G(f)|$) (RMSE -31 dB, ranging. -14 to -107 dB). (f) Phase angle error of recovered signal.

Recovery Process

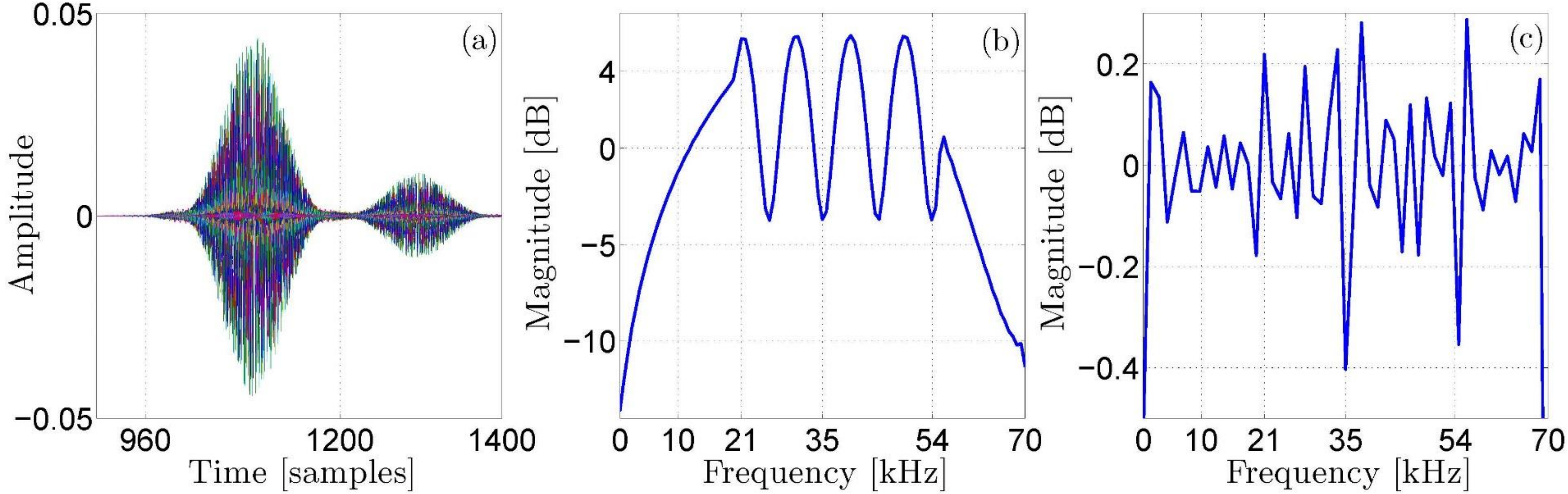


Fig.: (a) $f(t) * h(t)$ (n=30, each color corresponds to a different band) with integration limits t_0 = 960 samples and t_e = 1200 samples. The direct arrival of h(t) is located at 1100 samples, the first reflection at 1300 samples. (b) |H(f)| measured over integration limits. (c) $|\hat{G}(f)|$ measured over integration limits. Ideally, the response should be 0 dB.

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: $\pm\,\sigma\ [dB\ re\ 1\mu Pa^{2}/_{Hz}\ at\ 1m]$

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

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)

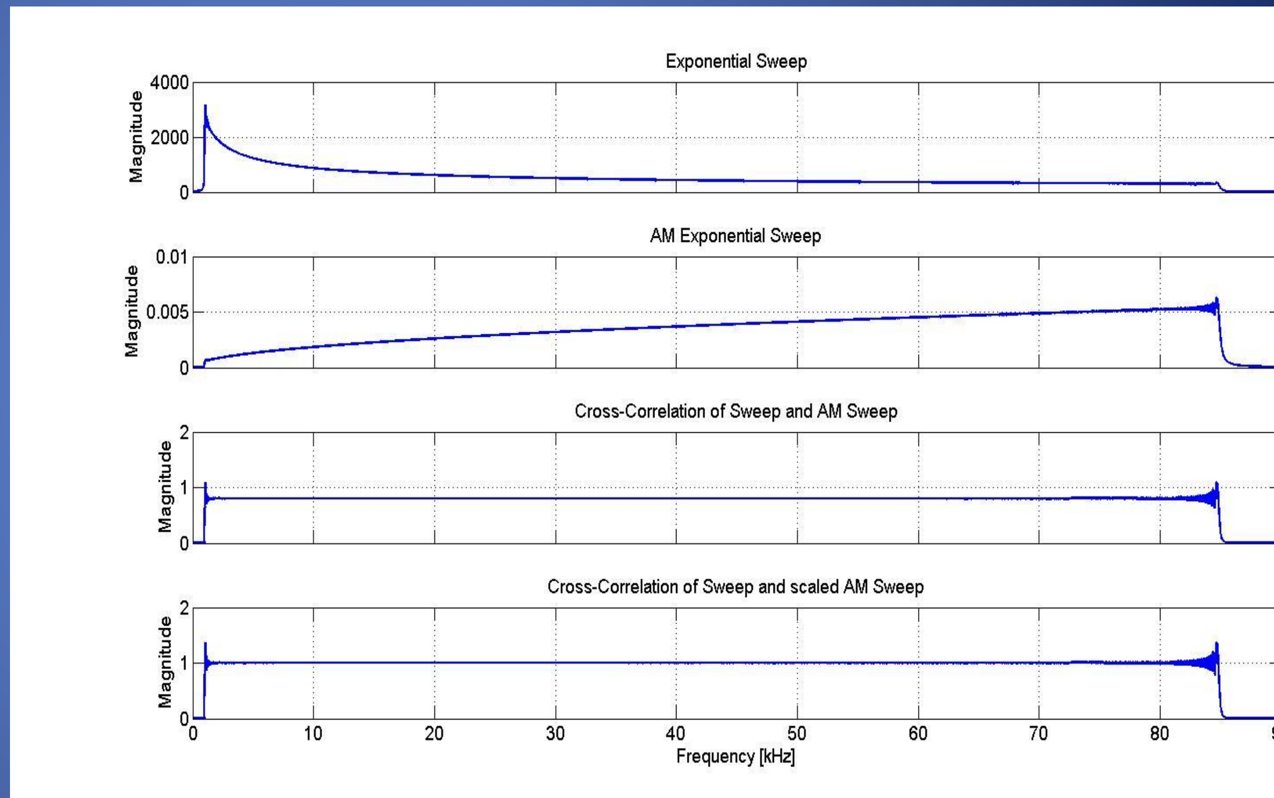


Fig.: scaled autocorrelation