Applied Acoustics 105 (2016) 24–34

Contents lists available at ScienceDirect

Applied Acoustics

journal homepage: www.elsevier.com/locate/apacoust

Source characterization using recordings made in a reverberant underwater channel

Kay L. Gemba *, Eva-Marie Nosal

Department of Ocean and Resources Engineering, School of Ocean and Earth Science and Technology, University of Hawai'i at Mānoa, Honolulu, HI 96822, United States

ARTICLE INFO

Article history:
Received 15 December 2014
Received in revised form 3 November 2015
Accepted 11 November 2015

Keywords:
Dereverberation
Pool experiment
Impulse response
Source characterization
Hydrophone calibration

ABSTRACT

The ability to accurately characterize an underwater sound source is an important prerequisite for many applications including detection, classification, monitoring and mitigation. Unfortunately, anechoic underwater recording environments required to make ideal recordings are generally not available. This paper presents a practical approach to source characterization when working in an imperfect recording environment; the source spectrum is obtained by equalizing the recording with the inverse of the channel's impulse response (IR). An experiment was conducted in a diving well (depth of 5.18 m) using a logarithmic chirp to obtain the IR. IR length is estimated using methods borrowed from room acoustics and inversion of non-minimum phase IR is accomplished separately in the time and frequency domain to allow for a direct comparison. Results indicate that the energy of controlled sources can be recovered with root-mean-square error of −70 dB (10–70 kHz band). Two equations, one coherent and the other incoherent, are presented to calculate source spectral levels of an unknown source in a reverberant environment. This paper introduces a practical procedure outlining steps to obtain an anechoic estimate of an unknown source using equipment generally available in an acoustic laboratory.

© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

Underwater source characterization is important for numerous applications. For example, passive acoustic detection and classification can be improved by knowledge of the sound characteristics of the object of interest. With knowledge of the source, array configuration and specifications can be optimized for monitoring. As another example, environmental compliance laws regulate an environment by putting limits on emitted acoustic energy, so that a sound source needs to be well understood before being used in the environment. Unfortunately, anechoic underwater recording environments required to make ideal recordings are generally not available or are cost-prohibitive.

An anechoic recording contains the direct arrival of acoustic energy from a source to the hydrophone with minimal noise or wall reflections. Sound levels estimated from recordings made in a reverberant environment (such as a test tank or pool) generally overestimate source levels due to additional wall reflections and noise. It was found [4,9] that the acoustic power of a source can be separated from reverberant energies by measuring the spectral pressure at one or more random locations in a reverberant enclosure (yielding spatial mean spectral levels). Recordings must be conducted in the far field of the source, e.g., the hydrophone is placed within the homogeneous and isotropic reverberant field. An estimate of the source is obtained by adjusting recorded levels with calculated reverberant energies. The reported error for a 100 Hz broadband white noise source [4] is ~1.5 dB and expected vs. calculated spectral levels for pure sinusoids differ by 0.1–5.8 dB. This approach provides an economic way [9] to estimate source power but is inherently limited to an incoherent estimate. To our knowledge, no other approaches exist for characterization of sound sources in underwater reverberant environments. Here, we follow a different ansatz using methods borrowed from room acoustics to estimate and invert the recording channel.

The recorded signal is the convolution of the source signal with the impulse response (IR) of the channel, hence, in principle, convolving the recorded signal with the inverse of the IR equalizes the channel (see Section 2), yielding an anechoic estimate of the source signal (Ref. [27] serves as an excellent introduction to deconvolution). The acoustic IR can be estimated with an excitation signal and by convolving an inverse filter with the received signal [24,6]. Theoretically, using an impulsive excitation signal is the preferred way to estimate the IR since an impulse freezes the system under investigation in time. In practice, when the test device is not purely electrical but has an acoustic path in the

* Corresponding author.
E-mail address: gemba@hawaii.edu (K.L. Gemba).

http://dx.doi.org/10.1016/j.apacoust.2015.11.008
0003-682X/© 2015 Elsevier Ltd. All rights reserved.

measurement chain, this procedure has to be adjusted because the transmitting transducer cannot realize an impulse. The excitation signal is selected and pre-colored to maximize signal to noise ratio (SNR) and the recorded signal reflects the states of the system over the playback duration. Popular signals include periodic signals such as maximum length sequences (MLS) and non-periodic signals such as linear or logarithmic sweeps. Once the IR is deconvolved and its length is estimated (see Section 3), it can be inverted.

Three primary methods have been investigated in room acoustic literature to coherently invert an acoustic IR: homomorphic deconvolution [31,21,25], single channel least squares (SCLS, a time domain method) [35,12,18], and inversion in the frequency domain [10]. In principle, homomorphic deconvolution is attractive because deconvolution of minimum phase signals in the time domain is division in the frequency domain and subtraction in the cepstrum domain [23]. However, non-minimum phase signals have cepstral overlap and the direct arrival cannot be easily separated from early reflections. It was found [21] that an IR has minimum phase only if the wall reflectivity coefficient is small enough (below approximately 0.4), otherwise its inverse will be acasual or unstable. The problem in room and underwater acoustics is the same: the IR is of non-minimum phase if partial energies (in the time domain) are not strictly decreasing. This is clearly the case for late reflections from a high impedance boundary (such as water–air). In addition, spectral zeros of the IR result in narrow band noise amplification and direct inversion is not desirable.

SCLS can address this problem and has been found to be more practical than homomorphic deconvolution [18]. The inverse of a mixed-phase IR in the least-squares sense can be significantly improved using a processing delay [21,17,3] to render it causal and improve stability. Even though only approximate equalization can be achieved [16], SCLS is robust to measurement noise and only partially equalizes deep spectral nulls [20], hence reducing narrow band noise amplification after equalization. In addition, it can easily evolve into a multi-channel method [16]. Ill-conditioned inverse problems can also efficiently be solved and regularized in the frequency domain [33,10]. The inverse filter is loaded with a small frequency dependent constant to improve the inversion. In this paper, we use and compare SCLS and frequency domain inversion.

In a preliminary experiment [8], a linear sweep was used to estimate the IR of an underwater reverberant recording channel. The inverted IR was used to remove reverberation effects to approximate source spectral levels (SSL) of a recorded SCUBA diver over an appropriate band. A follow up experiment was conducted to investigate and quantify dereverberation performance using control sources; these results are presented in the following sections.

This paper presents a practical procedure for underwater acoustic experimentation to recover an anechoic estimate of a source recorded in a reverberant environment. It is structured as follows: First, the problem is formulated in Section 2 and Section 3 describes the proposed experimental procedure which is validated by an experiment (Section 4). Methodology of data analysis is presented in Section 5 followed by results in Section 6. The paper concludes with a discussion in Section 7.

2. Mathematical formulation

Fig. 1(a) shows a diagram of the inverse problem in an underwater recording environment. The recording process can be modeled as the convolution of individual IRs. Here, the input signal of the source $d_i(t)$ is recorded in a reverberant channel $g(t)$ with a hydrophone $r_2(t)$. The hydrophone is connected to an analog to digital converter (ADC, denoted by $r_1(t)$) and the recorded output signal $d_o(t)$ is stored on a hard drive:

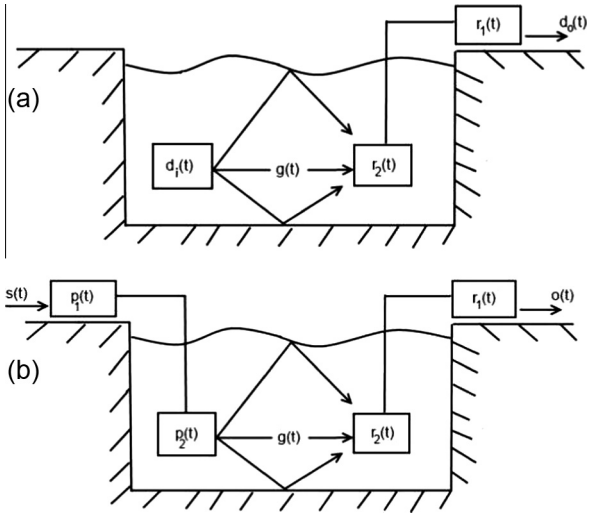


Fig. 1. Pool diagram showing schematics of (a) the inverse problem with an unknown source and (b) the forward problem with a known source.

$$d_o(t) = r_1(t) * r_2(t) * g(t) * d_i(t). \quad (1)$$

The problem of interest here is to estimate the input signal which is not immediately possible since both the source and the IR of the channel are unknown. To estimate the IR, the source signal is replaced by a known signal, shown in Fig. 1(b). For the forward problem, the source signal $s(t)$ is fed through a playback and pre-amp device $p_1(t)$ which is connected to a transmitting transducer $p_2(t)$. It is assumed that both the unknown source and the transmitting transducer have similar directionality and are of similar shape. The channel and the recording equipment is the same as in the inverse problem and the recorded signal is denoted by $o(t)$. For convenience in the rest of this paper, the total IR combining the playback and recording devices with the channel is abbreviated by the filter $h(t)$ (Eq. (2)).

$$h(t) = r_1(t) * r_2(t) * g(t) * p_2(t) * p_1(t) \quad (2)$$

$$o(t) = h(t) * s(t) \quad (3)$$

Our first task is to identify the IR of the system $h(t)$ which is convolved with the input $s(t)$ to the system to produce output $o(t)$ (Eq. (3)). Since the pool remains unchanged except for random fluctuations due to pool pumps and outside disturbances (such as wind), we assume that the resulting channel is an ergodic stochastic system. If we further assume that the in-phase and quadrature components of both amplitude and phase each have Gaussian distributions, the sinusoidal pressure in the channel follows a Rayleigh distribution [13] which is a function of absorption coefficient α_i, combined surface area (A_i) of the walls and water surface, and distance (r) from the source to the hydrophone. The 68% range of the sinusoidal sound pressure level (SPL) distribution (corresponding to approximately one standard deviation (SD), denoted by σ) was derived in [5] and is a linear approximation between σ and r for a poorly reverberant enclosure (this is where the constant in Eq. (4) comes from). Here, the original equation is slightly modified to average over six non-uniform absorption coefficients, corresponding to the boundaries of a rectangular enclosure:

$$\pm\sigma \approx 40r\left(1 - \sum_{i=1}^{6}\frac{\alpha_i}{6}\right)^{\frac{1}{2}}\left(\sum_{i=1}^{6}\alpha_i A_i\right)^{-\frac{1}{2}} \text{dB}. \quad (4)$$

The idea behind Eq. (4) is that square sound pressures are proportional to the energy density ratio of the reverberant and the direct sound field. The units of dB are relative to the mean reverberant SPL of a recorded signal. Therefore, the recorded system's IR $h(t)$ must be approximated by a sufficient number of realizations and its expectation will be denoted by $\mathbb{E}[h(t)]$. Note that the actual pressure distribution for broadband signals with different amplitudes is more complicated [34,15] and Eq. (4) will be used to approximate the SD (averaged over all frequencies) of the stochastic system.

Each realization of $h(t)$ can be estimated using an appropriate excitation signal [7] such as MLS, linear or logarithmic sweep, Golay sequence, and random noise. First, an inverse filter $f(t)$ is computed which equalizes the excitation signal (Eq. (5)). For a logarithmic sweep, $f(t)$ is the derivative of the time reversed excitation signal, which compensates for the non-white spectrum. In this case, an inverse filter's 3 dB per octave slope will equalize the pink spectrum of $s(t)$ (slope of -3 dB per octave) but will introduce a squaring of the magnitude spectrum and a pure delay [30]. The designed inverse filter must compensate for these effects to deconvolve the IR $h(t)$ (Eq. (6)). It is assumed that channel noise and self-noise of the electrical systems are not correlated with the excitation signal. Uncorrelated noise will at most contribute a constant [24] to the deconvolved IR: this constant is zero if the excitation signal has no trend, which can be ensured by proper signal design.

$$s(t) * f(t) = \delta(t) \tag{5}$$

$$o(t) * f(t) = h(t) * s(t) * f(t) = h(t) \tag{6}$$

After a sufficient number of realizations of $h(t)$ are obtained and $\mathbb{E}[h(t)]$ is estimated, deconvolution performance can be quantified objectively for a known source. A recorded test signal $o(t)$ is convolved with the IR's inverse $\mathbb{E}[h(t)]^{-1}$ (Section 3 discusses coherent inversion) to compute an estimate of the input $\hat{s}(t)$ (Eq. (7)). Note that the performance measure does not require explicit knowledge of the playback or recording equipment's IRs.

$$o(t) * \mathbb{E}[h(t)]^{-1} = h(t) * s(t) * \mathbb{E}[h(t)]^{-1} = \hat{s}(t) \tag{7}$$

The unknown source signal in the inverse problem can now be found by convolving the recorded signal in Eq. (1) with the inverse IR $\mathbb{E}[h(t)]^{-1}$. However, the inverse filter includes $p_1(t)$ and $p_2(t)$ and the recorded source signal must be adjusted by the IRs of the playback system and the transmitting transducer:

$$\begin{aligned} d_o(t) * \mathbb{E}[h(t)]^{-1} &= r_1(t) * r_2(t) * g(t) * d_i(t) * \mathbb{E}[h(t)]^{-1} \\ &= \hat{d}_i(t) * (p_2(t) * p_1(t))^{-1} \\ \hat{d}_i(t) &= d_o(t) * \mathbb{E}[h(t)]^{-1} * p_2(t) * p_1(t) \end{aligned} \tag{8}$$

If the phase response of the playback system is unknown, Eq. (8) can be computed incoherently to get sound pressure levels (Eq. (9)). The power spectral density (PSD) of the output signal equals the PSD of the input signal adjusted by the squared transfer function: $S_{out}(f) = S_{in}(f)|H^2(f)|$. As in the passive sonar equation [32] and using units of decibel, source spectral levels (SSL) can be computed by estimating the PSD of the recorded signal S_{d_o}, adjusted with the squared amplitude responses of the channel $|H|$, the transmitting transducer $|P_2|$ and the playback system $|P_1|$.

$$\begin{aligned} SSL = {} & 10\log_{10}(S_{d_o}) - 20\log_{10}(\mathbb{E}[|H|]) + 20\log_{10}(|P_2|) \\ & + 20\log_{10}(|P_1|) \end{aligned} \tag{9}$$

Units for Eq. (9) should be stated as $\pm\sigma$ [dB re 1 μPa^2/Hz at 1 m] for a channel length of 1 m.

3. Proposed experimental procedure

The following steps (Fig. 2) are proposed to estimate the IR in the forward problem in order to extrapolate the unknown source signal in the inverse problem. The first task is to estimate the length of the excitation signal which will be used to deconvolve the IR. In general, it is desirable to have a long duration excitation signal to increase the energy applied to the system and improve SNR for a single IR realization. A long duration signal also ensures that transducers have sufficient excitation time at lower frequencies (delayed low-frequency components). The overlap-and-add deconvolution approach [11] is superior to linear (the entire time signal) and circular (frequency-domain) deconvolution: it is faster and does not impose a minimum length restriction on the excitation signal in order to avoid circular aliasing. While linear deconvolution avoids time aliasing, the tail of the IR might be lost if the period of successively emitted excitation signals is shorter than the IR [30]. Therefore, we need a method to roughly approximate the length of the IR so that the excitation signal can be appropriately designed.

In room acoustics, reverberation refers to sound that reflects one or more times from the boundaries of an enclosure after excitation by a sound source [13]. If the room is large relative to the wavelengths of interest, it is sufficient to consider propagation of sound energy (i.e., phase information is not required) [22]. To determine the minimum length of the excitation signal, we need to know the time required for the signal and reverberant energy to decay to the noise floor. In room acoustics, reverberation time (denoted by T_{60}) is defined as the time it takes for the sound pressure level to fall by 60 dB after the cessation of sound (the sound is absorbed by interacting with the boundaries). We are also interested in a somewhat different duration here, which is the time for the sound pressure level to fall to the noise floor, since this is the portion of the IR that can be measured or deconvolved. We will call this the "signal to noise decay time" and denote it by T_{sn}.

The first step (Fig. 2(a)) in the dereverberation procedure is to estimate T_{60}, which will be used to approximate the actual reverberation time. This is accomplished using a formula borrowed from room acoustics [13]:

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

where c denotes soundspeed (1500 m/s for the freshwater pool here), V denotes the volume of the pool (in m^3), α_i is the absorption coefficient of each surface area A_i (in m^2), and S (in m^2) represents the combined area of all underwater walls and water/air boundary of the rectangular enclosure. The absorption coefficient α (we are interested in energy which is transmitted away from the system) for each boundary can be estimated using

$$\alpha_i = 1 - \left|\frac{z_w - z_i}{z_w + z_i}\right|, \tag{11}$$

where z_w represents the acoustic impedance of water and z_i the acoustic impedance of the boundary. Standard values for z are 415 N s/m^{-3} (air), 8×10^6 N s/m^{-3} (concrete) and 1.5×10^6 N s/m^{-3} (water) [14,19].

The estimated T_{60} can subsequently be used to design an excitation signal (e.g. a logarithmic sweep) $s(t)$ that is approximately 5–10 times longer (Fig. 2(b)) than T_{60}. The frequency range of the excitation signal should exceed the frequency range of interest to minimize transducer transients and its maximum sampling rate is given by the minimum sampling rate of either the playback or recording system.

The next step is to conduct the experiment (Fig. 2(c)). Noise sources such as pool pumps and water overflow mechanics should

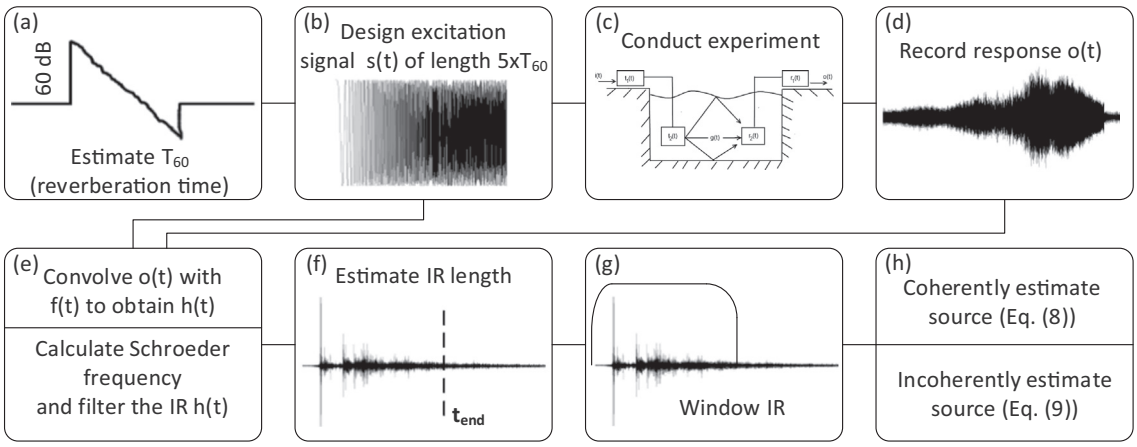


Fig. 2. Flowchart showing the proposed procedure to obtain (a)–(g) the impulse response in the forward problem and (h) the estimate of the unknown source in the inverse problem. A detailed description of each step is given in Section 3.

be eliminated. Results correspond to SSL if source-receiver separation is 1 m. However, recordings should be conducted in the far field region in reverberant enclosures [28] because SPL can fluctuate significantly in the near field. We recommend recording 100 successive realizations of the excitation signal each separated by a time sufficiently longer than T_{60} to allow energy to decay between recordings. In addition, we recommend recording a control signal to ensure that the IR's inverse is correctly scaled (see Eq. (7)). Once all excitation and control signals are recorded, the unknown source can be recorded in the same channel.

The IR $h(t)$ is deconvolved from the source signal (Fig. 2(e)) by convolving each recorded response (Fig. 2(d)) with the inverse filter $f(t)$ (Eq. (6)). Afterwards, the IR is filtered to remove resonance frequencies. As noted above, the room must be large relative to the wavelengths (lowest frequency) of interest. The task is to identify the upper bound of the transition region which separates lower frequency distinct modes from the high frequency region with statistical properties. The frequency for which this applies is called the Schroeder frequency f_s and is given by [4]

$$f_s = 0.6\sqrt{\frac{c^3 T_{60}}{V}}. \tag{12}$$

In practice, the IR is band-pass filtered at this step: while the lower bound is given by Eq. (12), actual cutoff frequencies should correspond to the bandwidth of interest or might be dictated by the frequency response of the equipment.

Once the IR is filtered, its length is estimated (Fig. 2(f)) with more accuracy than the estimate obtained using Eq. (10). We propose using the measure of echo density [1] to identify the transition region from high-energy early reflections to low-energy late reverberation. Removing the IR's late reverberant part reduces complexity for coherent inversion and does not significantly effect dereverberation performance. Echo density is computed by sliding a window over the IR and calculating the SD in each window. Early reflections correspond to a large SD with few outliers while the late reverberant part of the IR takes on a Gaussian distribution. The normalized echo density captures this difference by counting the percentage of values outside one SD: A value closer to zero indicates dominant energy due to early reflections while a value near one corresponds to the reverberant tail. Echo density is a function of time and can be plotted concurrently with the IR to identify the transition time (t_{end}) between early reflections and late reverberation before the IR is truncated using a window (Fig. 2(g)). Windowing tappers early and late samples smoothly to zero and eliminates a step-function response.

Once estimated for each realization, IRs are averaged to estimate the unknown source (Fig. 2(h)) either coherently (Eq. (8)) or incoherently (Eq. (9)). The incoherent formulation can subsequently be immediately applied. The coherent formulation requires inversion for the dynamics of the IR first. Inversion of mixed-phased IR is achieved using SCLS technique [27] and a detailed overview can be found in [8]. The optimum inverse $\hat{f}(t)$ in the least-squares sense is given by

$$\hat{f}(t) = [A^T A]^{-1} A^T z. \tag{13}$$

A is the circulant matrix of the IR and $z = [0, 0 \ldots, 1]^T$, where the spike (of value 1) occurs at the position of the delay. The processing delay improves inversion performance by shifting energies from the acausal part into the causal part of the IR.

In the frequency domain, the inverse of the IR is calculated using

$$\widehat{F}(\omega) = \frac{H^*(\omega)}{H^*(\omega)H(\omega) + \varepsilon(\omega)}, \tag{14}$$

where ω is the angular frequency. The complex conjugate of $H(\omega)$ is denoted by $H^*(\omega)$ and ε is a frequency dependent regularization parameter. We allow a gain of 0.1 dB for frequencies within the band of interest while dampening frequencies outside the band of interest by 6 dB.

4. Pool experiment

An experiment was conducted in the University of Hawai'i at Mānoa's diving well in June 2013 to quantify the performance of the proposed procedure for source characterization. The dimensions of the pool were 22.9 m by 22.9 m with a depth of 5.18 m, corresponding to primary resonance frequencies of 65 Hz and 290 Hz, respectively. To estimate the IR of the recording channel, a Fostex recorder (Tokyo, Japan, Model FR2-8347) was used for signal playback and signals were pre-amplified with a Roland OCTA-Capture device (Los Angeles, CA, Model UA1010). The channel gain of the pre-amplifier was adjusted depending on the signal from 0 to 6 dB. All signals were checked with an oscilloscope during playback, which was connected to a second output on the Roland OCTA-Capture device. The amplitude responses of both the Fostex and the Roland device are nearly uniform. A single CR1 Sensor Technology Limited transducer (Seattle, WA, SN: 09178-01) was connected to the pre-amp. The response of the transducer is band limited from 10 kHz to 100 kHz and not uniform. The minimum

Table 1
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 (+2 cm)	5	5	1	85	1	3
Mixed sinusoids (+4 cm)	5	5	1	85	1	3

and maximum sensitivity 111.5 dB re 1 μPa/V at 10 kHz and 136.5 dB re 1 μPa/V at 35 kHz, respectively.

A total of 9 TC4032-1 Teledyne-Reson hydrophones (Slangerup, Denmark) were used to record data. Four were placed in a spherical configuration around the CR1 at a distance of 1 m. Multiple receivers were used because either inversion method is capable of using multiple channels. However, only data from one of the four channels were used to compute results for this paper. The 5 remaining hydrophones where placed at random positions in the diving well, at least one wavelength away from reflective surfaces. The response of the Reson hydrophones is nearly flat at −170 dB re 1 V/μPa throughout the whole band of interest (10 kHz to 70 kHz). The minimum and maximum sensitivity are −172.6 dB re 1 V/μPa at 12.4 kHz and −168.6 dB re 1 V/μPa at 54.5 kHz. Data were recorded on nine channels of a custom Technologik ADC (Seattle, Washington) with a sampling rate of 264.60018 kHz. An analog high pass filter at 0.5 kHz and a 100 kHz analog low pass filter were used to pre-filter the signal to reject low frequency noise and additional energy at higher frequencies. The gain setting of the ADC was set to 20 such that the maximum amplitude of the recorded signal remained at about 0.6 V for most signals to avoid clipping.

Three different types of signals were played for IR calculations: linear and logarithmic sweeps and MLS. Sinusoids at different frequencies and white noise (10 kHz bandwidth) were also tested. There was a pause of 4 s after each signal to ensure that all input energy decayed below the noise floor. The distance of the spherical configured hydrophone was increased to 1.02 m and 1.04 m after all tests were completed and playback of sinusoids was repeated. This test was performed to investigate if incoherent dereverberation requires a strict channel geometry. Table 1 shows an overview of all signals played, their respective length, frequencies and total number of repetitions.

5. Analysis of data and performance measures

The following sub-sections discuss how T_{60} can be estimated from recorded data to validate Eq. (10) (Fig. 2(a)). The same method can be used to estimate T_{sn} of the IR which gives the upper bound for the IR's length (Fig. 2(f)). The following sub-sections discuss how the deconvolved IR is filtered (Fig. 2(e)), how its window is computed (Fig. 2(g)) and performance criteria for coherent inversion (Fig. 2(h)). The final sub-section discusses dereverberation performance measures for control sources.

5.1. Estimating T_{60} and T_{sn} from data

The standard way to obtain T_{60} experimentally is known as the method of backward integration [29] and was previously used [8] to estimate T_{sn} of an underwater channel. Schroeder showed that the ensemble average of the squared signal decay is equivalent to an integral over the squared IR. Here, a slightly modified method is used to plot remaining energy versus time of the deconvolved IR $h(t)$ to (a) estimate T_{sn} and dynamic range (measured from direct arrival until the IR decays into the noise floor) and (b) obtain a second estimate of T_{60} by extrapolating the linear region of the decay curve to a 60 dB drop. Knowledge of T_{sn} is required to estimate an upper bound of the IR length for analysis and the extrapolated value for a 60 dB drop will validate Eq. (10).

In particular, [2] modified Schroeder's method by subtracting an average noise term $\bar{\eta}^2$ from the squared IR g with additive noise η:

$$\langle h^2(t)\rangle = \int_t^\infty ([g(\tau)+\eta(\tau)]^2 - \bar{\eta^2})d\tau \qquad (15)$$
$$= \int_t^\infty (g^2(\tau) + 2g(\tau)\eta(\tau) + \eta^2(\tau) - \bar{\eta^2})d\tau$$

This method can be used to extend the decay slope and clearly separate the IR from the noise. As the noise $\eta(\tau)$ can be either positive or negative, the second term in the expanded binomial integrates to zero. When Eq. (15) is integrated over a time much longer than T_{sn}, the trend is dominated by the noise term because the integration time t is close to the upper integration limit and the IR has decayed into the noise floor. However, when the time t is close to the beginning of the signal, the trend is dominated by $g^2(\tau)$ and corresponds to a decay curve. The transition region between the decay curve and the noise marks the end of the measured IR. Signal to noise decay time and dynamic range of the IR can be estimated from its decay curve.

5.2. Filtering of data

After the IR is deconvolved from the recorded signal (Fig. 2(e)), it is filtered over the bandwidth of interest. The lower frequency bound is given by Eq. (12). Here, the Kaiser bandpass filter has a pass-band from 5 to 75 kHz. The lower limit is motivated by the frequency response of transducers (we used a high-frequency transducer well above the Schroeder frequency), the upper limit to reject noisy bands from the ADC. The filter has a ripple ratio of 0.1 dB, a stop-band attenuation of −60 dB and the transition bands are chosen to be 1 kHz.

The IR is windowed (Fig. 2(g)) using a combination of left half-kaiser window (right edge at direct arrival), rectangular window and right half-kaiser window (left edge at t_{end}). After windowing, all signals are downsampled to 150 kHz.

5.3. Coherent IR inversion

If a coherent estimate of the unknown source is desired (Fig. 2(h)), the IR needs to be inverted using a time or frequency domain method. Inversion performance is quantified after coherently convolving the computed inverse with the IR (essentially the inverse is an equalizer). Equalization performance is determined both in the frequency and time domain: for ideal equalization, the resulting signal $D(t) = \hat{f}(t) * h(t)$ is a delta function centered at the position of the delay m (same delay as the spike in vector z in Eq. (13)). Equalization performance in the time domain is a measure of both amplitude and phase and given by

$$\varepsilon_t = D(m). \qquad (16)$$

Equalization performance in the frequency domain is evaluated using the magnitude deviation [18] of the equalized IR and is used to measure incoherent equalization performance:

$$\varepsilon_f = \left[\frac{1}{I}\sum_{k=0}^{I-1}(10\ \log_{10}|\widehat{D}(k)| - \overline{D})^2\right]^{-1/2}, \quad (17)$$

where

$$\overline{D} = \frac{1}{I}\sum_{k=0}^{I-1}10\ \log_{10}|\widehat{D}(k)|. \quad (18)$$

In Eqs. (17) and (18), I corresponds to the length of the FFT with frequency bins k and Fourier coefficients $\widehat{D}$ of the equalized signal D. Magnitude deviation is invariant to the length of the FFT and, for ideal equalization, equates to zero.

5.4. Dereverberation performance

Dereverberation performance is quantified by adjusting the recorded control signal with the expectation of the inverse IR. Incoherently, an approximation to the PSD of the control signal $\widehat{S}_s$ is computed by adjusting the PSD of the output signal S_o with the ensemble average of the channel's amplitude response $|H|$:

$$10\ \log_{10}(\widehat{S}_s) = M[10\ \log_{10}(S_o) - 20\ \log_{10}(\mathbb{E}[|H|])]. \quad (19)$$

To reduce the variance due to the SPL distribution, adjusted PSDs are further smoothed on the decibel scale using a zero-phase moving average filter M. All results are computed with a 1 Hz resolution (this means, for example, that filter length of 200 points corresponds to a 200 Hz bandwidth).

The ensemble average of coherently inverted IRs (Eqs. (13) and (14)) is computed incoherently

$$10\ \log_{10}(\widehat{S}_s) = M[10\ \log_{10}(S_o) + 20\ \log_{10}(\mathbb{E}[|\widehat{F}|]) - 2\overline{D}], \quad (20)$$

where $|\widehat{F}|$ is the amplitude response of either $\hat{f}(t)$ or $\widehat{F}(\omega)$. The equation is further adjusted by a constant, equal to the spectral mean of the equalized IR ($\overline{D}$, given by Eq. (18)). The contribution from the regularization parameter in Eq. (14) must also be included in $\overline{D}$. Comparing results from Eq. (20) to results from Eq. (19) will allow an estimate of the additional error due to the coherent inversion procedure.

Dereverberation performance is measured using incoherent root-mean-square error (RMSE):

$$\mathrm{RMSE} = 10\ \log_{10}\left[\frac{1}{N}\sum_{k=1}^{N}||\widehat{S}_k - S_k||^2\right]^{-1/2}, \quad (21)$$

where $\widehat{S}$ and S are the PSD coefficients of the recovered and the original signal, respectively. The RMSE is computed over a spectral bandwidth of N coefficients whereas k denotes the frequency bin.

6. Results

6.1. IR estimation

Acoustic IRs were calculated using Eq. (6) for linear and logarithmic excitation methods. Fig. 3(a) shows an IR obtained using a logarithmic sweep. Fig. 3(b) shows a spectral comparison between two randomly selected realizations of 200 ms long IRs using a linear and a logarithmic sweep. Both were computed using a bin width of 3.3 Hz and smoothed using a moving average filter of 201 points. The range in Fig. 3(b) is less than one because the combined transfer function attenuates signals from the input (Fostex playback system) to the output (ADC). Both excitation methods produce a similar spectral shape except at about 42 kHz.

To estimate σ for the stochastic IR from data, the SD for each frequency (10–70 kHz band, 1 Hz resolution, 50 averaged realizations) was computed on the log scale (20 $\log_{10}$). Results were averaged over the entire band to yield the average SD. For the log. excitation, $\sigma = \pm 1.35$ dB and for the linear excitation, $\sigma = \pm 1.92$ dB. Using values for the diving well (given in Sections 4 and 5.1), the theoretical sinusoidal pressure distribution (Eq. (4)) yields $\sigma = \pm 1.94$ dB.

Magnitudes of the mixed sinusoids (Table 1) were averaged over a duration of 4 s and divided by the input amplitude. The result is plotted against the excitation methods in Fig. 3(b). Sinusoids follow the overall trend of the transfer functions, the biggest exception being the dip at 42 kHz. This indicates that the IR using logarithmic and linear excitation is correctly scaled. Similar plots were computed for recorded sinusoids offset by 2 and 4 cm (Table 1). Computed amplitude ratios are within 2σ of the exponential transfer function.

Fig. 4(a)–(c) demonstrates the effect of coherently averaging the IRs both in the time and frequency domain. The coherently averaged IRs were aligned by their maximum value before averaging in the time domain. The figure demonstrates a significant amount of destructive interference at higher frequencies, which is clearly visible above 60 kHz. In addition, destructive interference is also apparent in the 15–40 kHz band. Note that the AIRs could also be aligned by their maximum correlation coefficient to reduce destructive interference. However, coherent averaging of IR realizations is problematic due to high fluctuations of the waveform in the time domain. To avoid this problematic, we average incoherently in Eqs. (19) and (20).

6.2. T_{60} and T_{sn}

The theoretical T_{60} (Eq. (10)) for the pool is 282 ms using acoustic impedance values given in Section 5 and dimensions of the pool given in Section 4. The corresponding Schroeder frequency (Eq. (12)) is approximately 356 Hz, which is close to the resonant frequencies of the pool (<300 Hz). Fig. 5(a) shows calculated IR decay curves using Eq. (15). The linear trend of the IR is clearly visible after the approximately 8 dB drop due to the direct arrival. Best-fit linear regression lines computed for all limits in between 150 and 250 ms, using a 1 ms step size, indicate a T_{60} of 247 ms (corresponding to an upper integration limit of 225 ms and a residual of 0.9994).

As discussed in Section 5.1, Schroeder's method can be used to estimate T_{sn} and dynamic range of the deconvolved IR. If the selected upper integration limit in Eq. (15) is too short (e.g. 125 ms in Fig. 5(a)), not all energy of the IR is included and the linear range is not maximized. If the integration limit is much longer than T_{sn} (e.g. 400 ms), a secondary, linear trend above 225 ms is visible due to noise. Between 175 and 225 ms, the order of the IR terms are similar to the order of the noise, corresponding to a dynamic range of more than 45 dB. For inversion performance analysis, 175 ms was selected as the upper limit of the IR length. Fig. 5(b) shows an IR plotted with its echo density. The echo density indicates a diffuse room before the direct arrival marked at time 0. Afterwards, early reflections dominate the statistics of the echo density until the room is diffuse at about 80 ms.

To validate the modified method by Schroeder and the selected upper length of the IR, 50 clock aligned IRs recorded by the five far field hydrophones were averaged and shown in Fig. 5(c). The exponential decay of the late reverberation is evident in the figure. The direct arrivals of the far field hydrophones are aligned with respect to the direct arrival in Fig. 5(b). The plotted noise reference line is on the same order as the noise between 175 and 225 ms.

6.3. Coherent inversion of IR

Fig. 6 shows channel equalization performance versus processing delay of the least-squares filter using Eqs. (16) and (17) for a

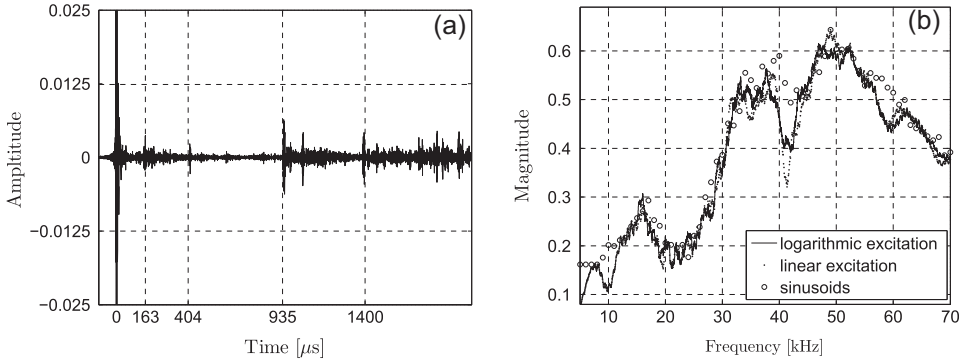


Fig. 3. (a) Acoustic IR of the diving well $h(t)$ with theoretical boundary reflection times. The scale is chosen to show details of the reflections, but cuts off the direct arrival which has a maximum amplitude between ±0.06. Ticks on the x-axis correspond to theoretical arrival time due to the boundaries (in order of arrival: direct arrival, floor, water surface, closest side wall, side wall). (b) Spectral comparison of $|H|$ using logarithmic, linear and pure sinusoidal excitations.

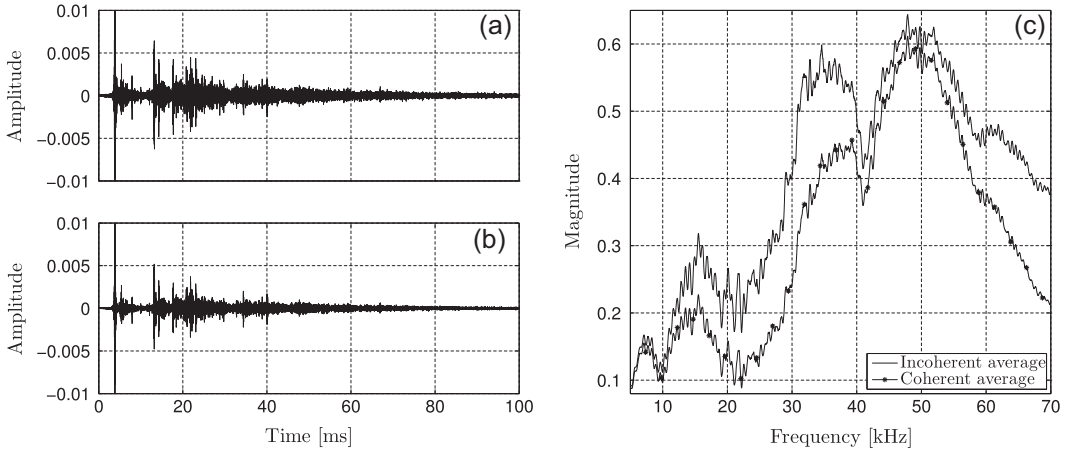


Fig. 4. (a) Single IR realization and (b) coherently averaged IR using 50 realizations. (c) Spectral comparison of incoherent vs coherent averaging using 50 realizations each. Comparing the IRs in the time domain, it can be observed that averaging reduces the variance but also attenuates reflection peaks. Effects of destructive interference are more apparent in (c) in the 10–40 kHz and above the 60 kHz band.

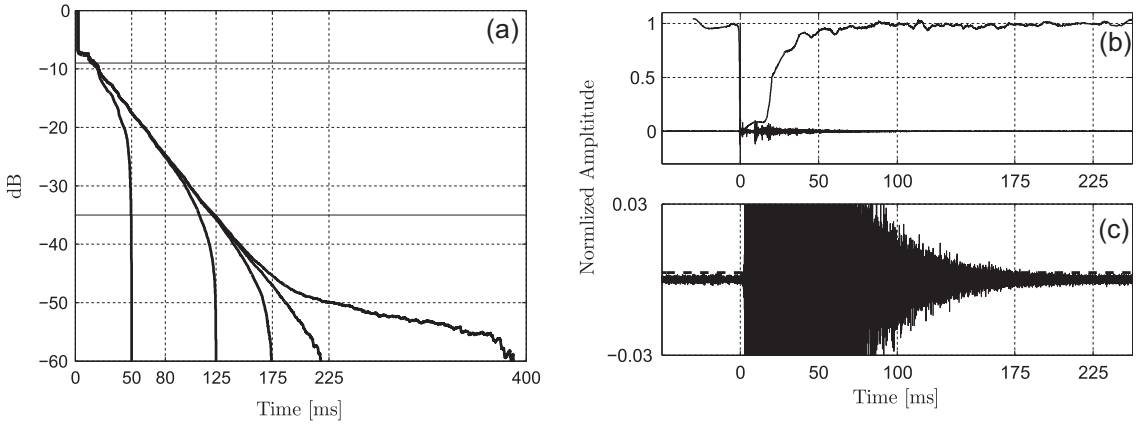


Fig. 5. (a) Decay curves of the IR with subtracted noise average. The ticks on the x-axis show selected upper integration limits, the two horizontal lines (−9 dB and −35 dB) correspond to the range over which T_{60} is calculated. (b) IR with echo density (top trace) showing transition time from early 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 ms.

randomly selected IR length of 152 ms (19 k samples, 10–70 kHz band). Equalization is significantly improved by increasing the processing delay: the greatest improvement is above 4 ms (500 samples points) and a noticeable improvement for ε_t is observed at 152 ms (same length as the inverted IR). Without a delay, $\varepsilon_f = 3.28$ and $\varepsilon_t \approx 0.001$. Results are similar for IR lengths ranging from 50 to 175 ms. In conclusion, the performance of the inverse filter improves significantly using a processing delay.

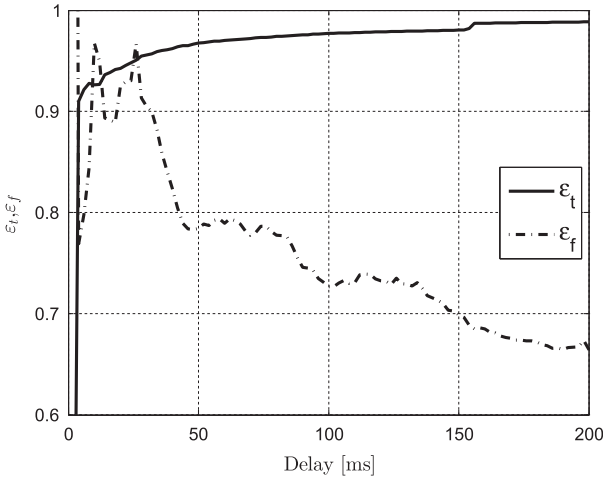


Fig. 6. Coherent SCLS inversion performance vs. processing delay for IR of length 152 ms.

6.4. Dereverberation results

We applied the dereverberation methods proposed here (Eq. (7)) to known sources. Doing so allowed us to establish the minimum length of the IR required to achieve reasonable dereverberation results, investigate the expectation operators in Section 5.4, and explore the length of the moving average filter. This section presents the results for the incoherent and coherent formulations.

For both formulations, the linear sweep was selected as the source (or control) signal because of its smooth spectra. The IR was calculated using the logarithmic excitation. First, the effect of varying the number of realizations for the ensemble average of the channel was explored. For these calculations, an IR length of 100 ms was used with no smoothing. The most significant RMSE reduction was achieved when increasing the number of realizations from 1 to 10: corresponding error decreased exponentially by about 7 dB. Using all 50 realizations, the error was reduced by an additional 1 dB. To minimize computational load for coherent inversion in the time domain, 10 realizations were selected for the ensemble average when investigating effects of smoothing and IR length.

6.4.1. Incoherent inversion

Incoherent dereverberation results were computed using Eq. (19) and 10 realizations were used to approximate the channel's amplitude response $|H|$. Results are shown in Fig. 7(a). Performance is function of both smoothing and IR length: error contours indicate that the error decreases as both parameters increase. Most improvement occurs as the length of the IR increases to 50 ms with minor additional improvement for further increases in IR length. The running average filter reduces the variance of both the transfer function and the recorded signal and significantly improves the RSME.

Fig. 8(a) shows the PSD of the recorded linear sweep which served as the source signal for Fig. 7. The recorded signal is adjusted with 1, 10 and 50 realization of the transfer function (Fig. 8(b), (c) and (d), respectively) using an IR length of 100 ms and smoothing filter length of 800 points. Increasing the number of realizations smooths out the dereverberated signal primarily in the 18–25 kHz band and approximately at 40 and 70 kHz. The improvement from 10 to 50 realizations is minor. Similar analysis for other test signals (i.e. white noise using various bandwidths) yields similar results.

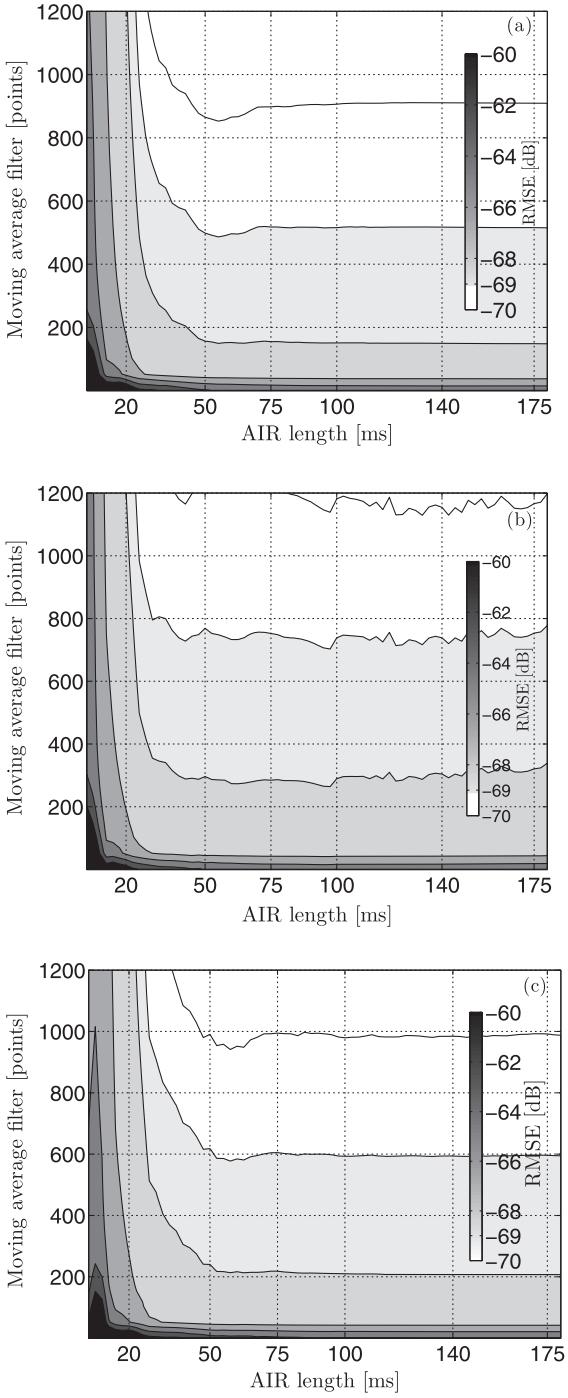


Fig. 7. RMSE in dB of dereverberated linear sweep. (a) Incoherently inverted IR, (b) coherently inverted IR using time domain (Eq. (13)) and (c) frequency domain (Eq. (14)) method. The expectation in Eqs. (19) and (20) is computed using 10 realizations of the inverted exponential sweep and RMSE ticks correspond to contour surfaces.

6.4.2. Coherent inversion

Dereverberation results using time-domain coherent inversion were computed using Eq. (20). First, the inverse of the IR is computed in the least-squares sense with a delay of 150 ms before averaging over 10 realization. RMSE results are shown in Fig. 7 (b). An IR length of 50 ms is required to achieve similar results to those using incoherent dereverberation. Comparing the length of the moving average filter at the −69 dB error contour in Fig. 7(b)

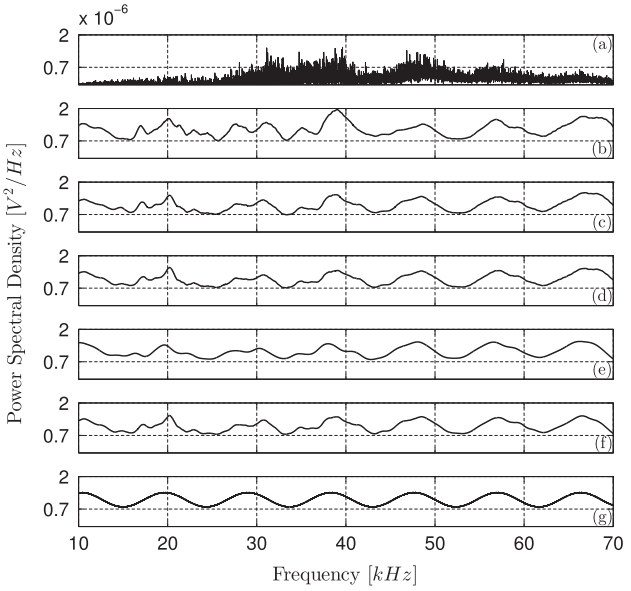


Fig. 8. (a) Recorded linear sweep. Incoherently adjusted linear sweep with (b) 1, (c) 10 and (d) 50 realizations in the ensemble average. (e) Coherently (Eq. (13)) adjusted linear chirp with 10 realizations in the ensemble average using a processing delay of 150 ms. (f) Coherently (Eq. (14)) adjusted linear chirp with 50 realizations in the ensemble average. Adjusted plots were computed using an IR length of 100 ms and moving average filter length of 800 points. The original swept-frequency cosine chirp is shown in (g).

(about 800) to Fig. 7(a) (about 500) indicates that dereverberation performance is reduced for the same length of the moving average filter. This result shows that trying to invert for the dynamics of the system increases the overall error. The PSD adjusted with the coherently inverted IR is shown in Fig. 8(e). This plot is computed using 10 realizations, an IR length of 100 ms, a delay of 100 ms and smoothed with an filter of 800 points. The shape is similar to the incoherent results but performance is poor below 35 kHz.

Fig. 7(c) shows the performance of the frequency domain inversion and allows a direct performance comparison to RMSE of the time domain method in Fig. 7(b). The error contours increase slightly when compared to the incoherent results in Fig. 7(a) but less than the time domain results in Fig. 7(b). This comparison indicates that the IR inverse should be calculated using Eq. (14) (at least from a magnitude persecutive). Fig. 8(f) shows results using frequency-domain coherent inversion. This plot is computed using 50 realizations (execution time is negligible) and the length of the inverse filter is the same as the length of the FFT. The performance increases lightly in comparison to Fig. 8(e) and results look very similar to the benchmark results in Fig. 8(d).

7. Discussion

The first task in applying the methods presented here for source characterization in a reverberant environment is to estimate T_{60} of the environment (Eq. (10)). This can be achieved without any prior knowledge or additional experiments, using only the dimensions and the approximate acoustic impedance values of the environment. The theoretical result for T_{60} of 282 ms compares well with the calculated result of 247 ms from data (Eq. (15)). The overestimate is probably caused by a decreased reflection coefficient with increasing frequency.

In the work presented here, the IR was estimated using both a linear and a logarithmic signal; it was not possible to obtain the IR using MLS. Both sweeps have similar standard deviations, $\sigma = \pm 1.35$ dB for the log. sweep and $\sigma = \pm 1.92$ dB for the linear sweep. These results are close to the theoretical sinusoidal distribution of approximately 2 dB. Eq. (4) can be used to approximate the SD of the broadband, *averaged* IR. Incoherent results may also be presented by plotting the adjusted PSD and its 68% confidence interval. The logarithmic sweep was selected for IR estimation for no particular reason except that it might be better suited for noisier environments due to its higher SNR at lower frequencies (which might be reflected by its reduced SD). Note that convolution of non-period signals to obtain the IR is trivial. In comparison, MLS require a strict time assumption of the system and precisely matching sampling rates of the recording and playback signal (we had fractional sampling rates for the ADC).

Eq. (15) can be used to identify T_{sn} and the dynamic range of the IR. Dereverberation performance is clearly a function of the IR length; the chosen filter has to include all of the early reflections and a good approximation seems to be quantifiable using echo density. While its range is function of the sliding window length (here 2500 taps), dereverberation performance is not too sensitive to the IR length. In the experiment here, an IR length above 50 ms and a smoothing filter length above 600 points yielded acceptable results. In the absence of any information, an IR length corresponding to an echo density close to one should be selected. This means that the dynamic range of the IR can be approximately 25 dB (see Fig. 5(a) at 80 ms) for a source with 45 dB of dynamic range (the linear sweep has a similar T_{sn} and dynamic range as the log. sweep in Fig. 5(c)), which is very reasonable in practice.

This paper presented two equations in Section 5.4 which can be used to approximate a signal in the forward problem. The incoherent equation does not require inversion for the dynamics of the channel using either the time domain or frequency domain formalism; subtracting the expectation of all incoherent realizations is sufficient. Results indicate that the RMSE for a broadband signal can approach −70 dB using moderate IR and running average filter lengths. −70 dB corresponds to a deviation of 1/10th of the average PSD's power. Overall, the range and trend of the amplitude of both the adjusted and original signal correlate well. Transducers with uniform frequency response will help improve SNR and dereverberation performance. For the experiment conducted here, Fig. 8 illustrates that performance is improved above 35 kHz, corresponding to the optimal frequency response of the transmitting transducer. Below 35 kHz, its amplitude response declines at about 17.5 dB per octave. RMSE computed for the 10–35 kHz band and 35–70 kHz band differed by approximately 4 dB favoring the higher frequency band. The poor performance below 35 kHz may be due to the channel geometry (note the improvement close to 10 kHz). Results further indicate that incoherent dereverberation might be invariant to small channel offsets and the point-source assumption can be relaxed. Therefore, once the IR is estimated for a given pool, any future recordings can be incoherently adjusted even if the recording has not exactly been performed in the original channel (this however requires additional verification and depends on the channel). We expect that for most practical purposes, the power of any source can be well approximated using the incoherent formulation.

Coherent equalization requires inversion for the dynamics of the system using Eq. (13) or (14). A processing delay (SCLS method, Fig. 6) shifts the acausal energies in the causal part of the signal and equalization improvements correspond to shifts past significant partial energies in the IR. For example, the greatest improvement here corresponded to a delay of 4 ms, which falls after the direct arrival (see Fig. 3(a) at 4.04 ms). Other major equalization improvements correspond to the dominant energies before 50 ms. This behavior of the spiking filter has already been identified in literature [27]. A distinct improvement is visible after a delay of 152 ms, being equal to the length of the IR. For maximum phase signals, a delay equal to the signal length minus one sample

corresponds to the best delay [36]. Here, the signal is of mixed phase and is seems that further improvement is possible but practically limited by the order of the pseudo-inverse (Eq. (13)). It should be noted that coherent equalization of IRs yields poor performance at offsets of fractions of a wavelength [26], which limits the method in practice. Depending on the geometry difference between the transducer and the source, it might be possible to compute a coherent estimate using lower frequencies only.

Fig. 7(b) and (c) shows coherent RMSE which is slightly increased in comparison to the incoherent case in Fig. 7(a). As expected, performance is similar for the length of the IR in both cases but inverting the dynamics of the system requires additional smoothing. An error of −70 dB is achieved for the time domain approach case using an IR length of 100 ms and a moving average filter with 1200 points. Depending on the nature of the signal, smoothing can significantly reduce narrow band signal features and care must be taken in selecting an appropriate filter length. It should be noted that coherent correlation of the IRs are poor and taking an ensemble average might result in destructive interference if clock management of consecutive recordings is not rigorously enforced. The destructive interference of coherently averaged IRs (Fig. 4) in the lower band might be due to fluctuating frequencies in this channel (see i.e. [13]), which complicates obtaining a coherent ensemble average. Increasing the length of the excitation signal will help to obtain a single IR realization which better approximates the mean of the pressure distribution. The channel's expectation in Eq. (8) can be taken before or after inversion and smoothing of the IR can be accomplished by use of an exponential decaying window to address the pressure fluctuation at the hydrophone. The constant in Eq. (20) is required since inverting a signal in the convolution sense yields a flat spectrum which is not necessarily unity. The same is true when designing the inverse for the logarithmic excitation in Eq. (6).

The time variance of the system and sinusoidal SPL distribution at the hydrophone must be addressed for both coherent and incoherent formulation. Including the expectation improved non-smoothed results by more than 7 dB and clearly helped to recover the shape of the control signal in Fig. 8. While the source signal can be reasonably recovered using 10 realizations for this experiment, the time variance might be larger for other environments. Recording 100 logarithmic realizations is recommended for any experiment which can be achieved in approximately 10 min. Results in Fig. 7 will further improve by including the expectation in Eqs. (8) and (9) on the recorded source signal. Dereverberation results here can therefore be interpreted as a lower performance bound when only one realization of the unknown signal is available.

Both the coherent time (Fig. 8(e)) and frequency (Fig. 8(f)) domain inversion methods produced comparable and acceptable results. However, the frequency domain method is faster by many orders of magnitude. For applications using narrow bandwidths, the SCLS formalism can produce artifacts such as pre-ringing in the equalized IR. While our excitation signal was very broadband (and an additional 5 kHz outside the band of interest), the frequency domain inversion might offer more regularization control for other applications. Presented results might further be improved by optimizing the regularization parameter and to allow for variation within the band of interest.

We demonstrated that it is possible to recover a control signal in the forward problem. Presented results are very fundamental from a system characterization point of view and methods as well as results (such as Schroeder's method and the smoothing filter) might aid in estimating the IR in more uncontrolled open ocean environments. Also, the inversion procedure might be applicable to localization applications for impulse like signals (such as mammal clicks) to improve event detection. Furthermore, the method presented here can possibly be used to calibrate transducers: once all amplitude responses of $h(t)$ are known, a transducer can be interchanged and the difference in amplitude response can be observed. To translate the results of the forward problem to the inverse problem, the point-source assumption and directionality requirement must be considered. Performance will decline if the source to be estimated and the transmitting transducer have different directionality, which will usually be the case. Fig. 3 indicates that the transmitting transducer might be directional: the magnitude of the high-impedance surface reflection is of lower order than the later side reflections. In addition, the forward problem neglects the adjustment due to the IRs of the playback equipment $p_1(t)$ and $p_2(t)$. The impedance mismatch can be kept to a minimum by selecting a pre-amp with small output impedance and a transmitting transducer with high impute impedance (voltage bridging). Results only compare the energy for the recovered signal and not its phase, which will be left for future investigation.

Acknowledgments

The authors would like to particularly thank one of the anonymous reviewers, who helped to improve the quality of the manuscript significantly. The authors thank research diver Troy Heitmann and Professor Volker Roeber for providing experimental help. 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. The views and conclusions contained in this document are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the U.S. Department of Homeland Security.

References

[1] Abel, J. and Huang, P. 2006. A Simple, Robust Measure of Reverberation Echo Density, in Proceedings of the 121st AES Convention, preprint 6985 1–10.
[2] Chu W. Comparison of reverberation measurements using Schroeder's impulse method and decay curve averaging method. J Acoust Soc Am 1978;63:1444–50.
[3] Clarkson P, Mourjopoulos J, Hammond J. Spectral, phase, and transient equalization for audio systems. J Audio Eng Soc 1985;33:127–32.
[4] Cochard N, Lacoume J, Arzelies P, Gabillet Y. Underwater acoustic noise measurement in test tanks. IEEE J Oceanic Eng 2000;25:516–22.
[5] Diestel HG. Probability distribution pressure of sinusoidal sound in a room. J Acoust Soc Am 1963;35:2019–22.
[6] Farina A. Simultaneous measurement of impulse response and distortion with a swept-sine technique. In: Audio engineering society convention, vol. 108. 2000. p. 1–24.
[7] Farina A. Advancements in impulse response measurements by sine sweeps. In: Audio engineering society convention 122, vol. 8(21). Vienna, Austria; 2007.
[8] Gemba KL, Nosal E-M, Reed TR. Partial dereverberation used to characterize open circuit scuba diver signatures. J Acoust Soc Am 2014;136:623–33.
[9] Hazelwood RA, Robinson SP. Underwater acoustic power measurements in reverberant fields. In: OCEANS 2007 – Europe. IEEE; 2007. p. 1–6.
[10] Kirkeby O, Nelson PA, Orduna-Bustamante F, Hamada H. Local sound field reproduction using digital signal processing. J Acoust Soc Am 1996;100 (3):1584–93.
[11] Kulp BD. Digital equalization using fourier transform techniques. In: Proceeding of 85th AES convention, vol. 2719. Los Angeles; 1988.
[12] Kumar K, Raj B, Singh R, Stern RM. An iterative least-squares technique for dereverberation. 2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), vol. 2. IEEE; 2011. p. 5488–91.
[13] Kuttruff KH. Room acoustics. 4th ed. London: Taylor & Francis; 2000. p. 368.
[14] Leighton TG, editor. The acoustic bubble. San Diego, CA: Academic Press, Inc; 1994. p. 640.
[15] Lubman D. Precision of reverberant sound power measurements. J Acoust Soc Am 1974;56:523–33.
[16] Miyoshi M, Kaneda Y. Inverse filtering of room acoustics. IEEE Trans Acoust Speech Signal Process 1988;36:145–52.
[17] Mourjopoulos J. Digital equalization of room acoustics. J Audio Eng Soc 1994;42:884–900.
[18] Mourjopoulos J, Clarkson P, Hammond J. A comparative study of least-squares and homomorphic techniques for the inversion of mixed phase signals. In: ICASSP '82. IEEE international conference on acoustics, speech, and signal processing, vol. 7. Paris, France: IEEE; 2003. p. 1858–61.

[19] Nawy EG. Concrete construction engineering handbook. 1st ed. Boca Raton, Florida: CRC Press LLC; 1997. p. 1232.
[20] Naylor P, Nikolay DG, editors. Speech dereverberation. London: Springer; 2010. p. 406.
[21] Neely ST, Allen JB. Invertibility of a room impulse response. J Acoust Soc Am 1979;66:165–9.
[22] Neubauer R, Kostek B. Prediction of the reverberation time in rectangular rooms with non-uniformly distributed sound absorption. Arch Acoust 2001;26:183–201.
[23] Oppenheim A, Schafer R. Dsp history – from frequency to quefrency: a history of the cepstrum. IEEE Signal Process Mag 2004;21:95–106.
[24] Polge R, Mitchell E. Impulse response determination by cross correlation. IEEE Trans Aerospace Electron Syst 1970;AES-6:91–7.
[25] Radlovic B, Kennedy R. Nonminimum-phase equalization and its subjective importance in room acoustics. IEEE Trans Speech Audio Process 2000;8:728–37.
[26] Radlovic B, Williamson R, Kennedy R. Equalization in an acoustic reverberant environment: robustness results. IEEE Trans Speech Audio Process 2000;8:311–9.
[27] Robinson EA, Trietel S. Geophysical signal analysis. Englewood Cliffs: Prentice-Hall; 1980. p. 466.
[28] Rossing T, editor. Springer handbook of acoustics. New York: Springer; 2007. p. 1182.
[29] Schroeder MR. New method of measuring reverberation time. J Acoust Soc Am 1965;37:409–12.
[30] Stan G, Embrechts J, Archambeau D. Comparison of different impulse response measurement techniques. J Audio Eng Soc 2002;50:249–64.
[31] Subramaniam S, Petropulu A, Wendt C. Cepstrum-based deconvolution for speech dereverberation. IEEE Trans Speech Audio Process 1996;4:392–6.
[32] Urick RJ. Principles of underwater sound for engineers. McGraw-Hill; 1967. p. 342.
[33] Van Buren AL. Theoretical design of nearfield calibration arrays. J Acoust Soc Am 1973;50:192–9.
[34] Waterhouse RV. Statistical properties of reverberant sound fields. J Acoust Soc Am 1968;43:1436–44.
[35] Xu G, Liu H, Tong L, Kailath T. A least-squares approach to blind channel identification. IEEE Trans Signal Process 1995;43:2982–93.
[36] Yom DH, Ann S. Spiking filtering of all-pass filter impulse responses. Proc IEEE 1987;75:1694–5.