CHARACTERIZATION OF UNDERWATER ACOUSTIC SOURCES RECORDED IN REVERBERANT ENVIRONMENTS WITH APPLICATION TO SCUBA SIGNATURES

A DISSERTATION SUBMITTED TO THE GRADUATE DIVISION OF THE UNIVERSITY OF HAWAI‘I AT MĀNOA IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

IN

OCEAN AND RESOURCES ENGINEERING

DECEMBER 2014

By
Kay Leonard Gemba

Dissertation Committee:

Eva-Marie Nosal, Chairperson
Dilmurat M. Asimov
Brian Bingham
Todd R. Reed
Bruce Howe

The scriptures speak of three Holy rivers Within.

These are Existence, Consciousness, and Bliss.

Being beyond thought or effort,

they cannot be objectified or subjectified.

They are so dear, so near,

behind the retina and before the breath.

You need not see; This, you are it.

-Papaji

Acknowledgments

I would like to thank my Mother Hanni, my Father Uwe, my sister Kathy, and my friends Jeff, Melissa, Patrick, Jacob, Jacobcito, Felix, Troy, Volker, Alex, Kim, Christina, and Marcel for all the wonderful support (and distractions) they have given me during this time. Professor Eva-Marie Nosal, Professor Todd R. Reed, and Professor Neil Frazer contributed significantly to my education with many insightful and passionate discussions and I am grateful for their support and trust.

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. Current methods adjust source recordings with spatially averaged estimates of reverberant levels. However, adjustments can introduce significant errors due to a high degree of energy variability in reverberant enclosures and solutions are inherently limited to incoherent approximations. This dissertation introduces an approach towards a practical, improved procedure to obtain an anechoic estimate of an unknown source recorded in a reverberant environment. Corresponding research is presented in three self-contained chapters.

An anechoic estimate of the source is obtained by equalizing the recording with the inverse of the channel's impulse response (IR). The IR is deconvolved using a broadband logarithmic excitation signal. The length of the IR is estimated using methods borrowed from room acoustics and inversion of non-minimum phase IR is accomplished in the least-squares sense. The proposed procedure is validated by several experiments conducted in a reverberant pool environment. Results indicate that the energy of control sources can be recovered coherently and incoherently with root-mean-square error (RMSE) of $\sim$ -70 dB (10 - 70 kHz band).

The proposed method is subsequently applied to four recorded SCUBA configurations. Results indicate that reverberation added as much as 6.8 dB of energy. Mean unadjusted sound pressure levels (0.3 - 80 kHz band) were 130 $\pm$ 5.9 dB re 1 μPa at 1 m. While the dereverberation method is applied here to SCUBA signals, it is generally applicable to other sources if the impulse response of the recording channel can be obtained separately.

This dissertation also presents an approach to separate all coloration from the deconvolved IR. This method can be used to estimate the channel's IR or the magnitude spectrum of the combined electrical equipment. The procedure is validated using synthetic results of an image-source model and the channel's IR is recovered over the full band with a RMSE of -31 dB.

Table of Contents

List of Tables

List of Figures

Chapter 1

Problem motivation and dissertation organization

1.1 Introduction

The work presented in this dissertation is motivated by the need of researchers using passive acoustics to characterize underwater sources. Source characterization is a pre-requisite for many applications such as detection, classification, monitoring, and mitigation. Sources of interest include small aquatic species, mechanical items such as autonomous underwater vehicles, and divers using an underwater breathing apparatus such as SCUBA or rebreathers. However, researchers usually do not have access to an underwater anechoic facility and blue water experiments are impractical and expensive. With this constraint in mind, a problem regarding characterization arises since recordings in a controlled pool environment include reverberant energies and noise.

The problem is addressed in the literature (1, 2) by (a) estimating reverberant energy (spatial mean spectral levels) in the enclosure and (b) adjusting spectral levels of the unknown source with the estimate. This approach has three disadvantages. First, spectral levels can vary by more than 10 dB in reverberant enclosures (3). Computed reverberant mean levels might therefore significantly underestimate or overestimate true levels. Second, the method adjusts mean levels by a frequency independent constant corresponding to the time required for energies to decay to the noise floor. This approximation does not include frequency dependent decay and therefore introduces additional errors. Third, results are inherently limited to an incoherent estimate. The procedure to estimate and remove reverberant energies can improve by addressing these three points.

This dissertation introduces a preliminary method which addresses the above shortcomings. An impulse response (IR) is a complete description of the recording channel's spatial coordinates and contains frequency information of additional reverberant energy. It can be deconvolved using appropriate excitation techniques (i.e., linear or exponential sweeps) and truncated before decay to the noise floor. Once estimated, the IR is inverted either coherently (in the least-squares sense) or incoherently over an appropriate bandwidth. The calculated inverse is used to adjust a source recorded in the same channel. When inverted coherently, the IR can be used (in theory) as an equalizer to deconvolve the original signal, yielding a time domain waveform. In practice, the IR's inverse might not be stable or casual; corresponding problematic and required assumptions will be addressed in the appropriate sections.

In summary, it is the goal of this dissertation to propose a practical procedure outlining steps to estimate the anechoic signal of a source recorded in an underwater reverberant environment. Once the best estimate of the source is calculated, it is possible to exploit its spectral characteristics for a specific application. The primary goal is to recover an estimate of the magnitude spectrum of an unknown source. Furthermore, a method is introduced to approximate the channel's IR without a priori knowledge of the electrical equipment's IR (which might not be available). In theory, this method can be used to obtain an estimate of the unknown source's time domain waveform (which includes phase information) if a calibrated recording system is available.

1.2 Reverberation and background noise

In an ideal situation for sound source characterization, energy propagating directly from the source to the receiver could be recorded for analysis without contribution from other sources of energy. Recordings made in a blue water environment or in an anechoic environment approximate this ideal. For recordings made in a non-anechoic enclosure, received energy consists of the direct energy of interest but also of early reflections, reverberant energy and noise.

Early reflections are wall reflections, which arrive milliseconds after the direct arrival. They are generally sparse and their arrival pattern depends on the channel geometry with respect to the enclosure. As the waves keep reflecting and diffusing from one wall to another, the sound field as a whole starts to be uncorrelated with respect to to the original signal. At that time, the phase and frequency content of the original signal is considered to be random. This part of the reverberation phenomenon is called late reverberation. In order to characterize the source in a reverberant environment, the additional energy due to early reflections and late reverberations must

be considered and removed from the recording. The degradation of the recording is considered to be of convolutional nature. There are, however, other sources besides reverberation which contribute to additional energy in the recording.

Unlike reverberation, noise can be considered to be additive and is handled by different techniques. Background noise can be a factor when recording sources of low signal to noise ratio (SNR). This noise can originate from pool pumps, flow noise, nearby machinery, rain, or other activity close to the pool. Some background noise is usually inevitable and must be included in the analysis. However, it is possible to reduce levels of background noise significantly by switching off pool pumps and other flow related noise sources and increasing SNR as much as possible. A noise discussion is included in both the dereverberation procedure (chapter 3) and channel IR estimation method (chapter 4).

1.3 Dissertation organization

The content of this dissertation is organized in three chapters. Each chapter corresponds to a paper, which is either already published or has been submitted to the *Journal of Acoustical Society of America* (JASA). Publications have been modified to fit the format of this dissertation.

Chapter 2 (4, published in JASA, Aug. 2014) discusses the need for SCUBA diver characterization to enhance security for near port environments. SCUBA signatures, sound pressure levels and source spectrum levels are presented and a method is introduced to coherently invert the recording channel in order to remove reverberant energy. This research was a first attempt to coherently characterize sound sources in a reverberant environment but does not investigate channel inversion performance.

Chapter 3 (5, submitted to JASA, Sep. 2014) introduces a dereverberation procedure for underwater reverberant environments and objectively quantifies performance using known sound sources. The procedure is explicitly derived and each step is discussed in detail using a flow chart for illustration purposes. Each step of the procedure is validated using experimental data. The impulse response of an underwater channel is estimated, inverted both coherently and incoherently, and used to adjust a control signal recorded in the same channel. The adjusted signal is then compared to the original signal, yielding objective performance results. However, the obtained impulse response includes other impulse responses (from transducers and electrical equipment) and can not be used directly to remove reverberant energies from an unknown source.

Chapter 4 (6, submitted to IEEE Signal Processing Letter, Nov. 2014) introduces a method to remove colorations of transducers and other electrical equipment from the impulse response of the underwater channel. This method is validated using an image-source model and is subsequently used to estimate the impulse response of the underwater reverberant channel from chapter 3.

In conclusion, the impulse response including transducer effects can be obtained using the procedure outlined in chapter 3. A coherent estimate of the channel's impulse response can then be obtained with the method introduced in chapter 4. These two methods offer an attractive solution to estimate the reverberant energy of the channel which can be used to obtain source spectral levels of an unknown source. Shortcomings and suggested future research is discussed in chapter 5.

Chapter 2

Partial dereverberation used to characterize open circuit scuba diver signatures

2.1 Abstract

The use of passive acoustics to detect self-contained underwater breathing SCUBA divers is useful for near-shore and port security applications. While the performance of a detector can be optimized by understanding the signal's spectral characteristics, anechoic recording environments are generally not available or are cost-prohibitive. A practical solution is to obtain the source spectra by equalizing the recording with the inverse of the channel's impulse response. This paper presents a dereverberation method for signal characterization that is subsequently applied to four recorded SCUBA configurations. The inverse impulse response is computed in the least-square sense, and partial dereverberation of SCUBA is performed over the 6 to 18 kHz band. Results indicate that early reflections and late reverberation added as much as 6.8 dB of energy. Mean unadjusted sound pressure levels computed over the 0.3 - 80 kHz band were 130 $\pm$ 5.9 dB re 1 μPa at 1m. Bubble noise carries a significant amount of the total energy and masks the regulator signatures from 1.3 to 6 kHz, depending on the regulator configuration. While the dereverberation method is applied here to SCUBA signals, it is generally applicable to other sources if the impulse response of the recording environment can be obtained separately.

2.2 Introduction

PROTECTION of critical infrastructure is a priority for security agencies. About 90% of trade is accomplished with cargo ships globally (7). Oceanic trade routes are connected to land-

based transportation systems by ports, with some large regions supplied by a few major ports, such as the Port of New York or Los Angeles. Disruption of port traffic by a device, such as dirty bomb, and the resulting contamination and loss of life and goods would have significant impact on these regions. In addition, military assets such as nuclear submarines or aircraft carriers use port facilities and require increased security.

Protection can be achieved using a layered approach of technological systems (8, 9). Generally speaking, technologies are layered from wide range, low resolution systems to narrowly focused, high resolution systems. These layers overlap and ideally allow technologies to communicate with one other. In particular, the layers to protect ports might be arranged as follows: satellites, Automated Identification System, radar, high frequency radar, optical systems, automated underwater vehicles and underwater acoustics. Most technologies used for harbor security do not penetrate the water surface which limits a layered systems approach for Self-Contained Underwater Breathing Apparatus (SCUBA) detection. Fortunately, acoustics can be used to monitor below the water surface.

Acoustic monitoring technologies can be broadly separated into active and passive systems. Several active acoustic systems are available for port security applications (10, 11, 12), which have the ability to detect low signal to noise (SNR) sources, such as divers. In favorable multipath environments SCUBA divers were detected at about 500 m (12) and intercepted at 350 m. However, the use of active systems can be limited in reverberant, nearshore environments (13). An active system may falsely classify sources with similar scattering characteristics, it requires an operator and has a higher up front cost than a passive system. The problem of a high false alarm rates is magnified in a multipath environment (14). Passive acoustic systems can complement some of these shortcomings and have no acoustic impact on their environment. This is an advantage in protected environments and when considering the bio-effects of noise on marine life (15). A passive system can also improve source classification performance if the underlying signal is known. Unfortunately port and near shore environments are typically very noisy areas: pleasure boats, commercial vessels (16), noise from shore, snapping shrimp (17), breaking waves (18) as well as wind and rain (19) all contribute to background noise levels, making acoustic detection of low SNR sources such as SCUBA difficult. Nevertheless, a well-designed passive acoustic system has potential to provide useful detection information, particularly when integrated with autonomous underwater and surface vehicles (20) and/or active acoustic systems.

A passive acoustic system relies on an accurate characterization of the source signal. A few past studies have published information about SCUBA source characteristics. Recordings of

several underwater breathing apparatus such as SCUBA and rebreathers were made by Radford (21) to quantify the effects of SCUBA noise during studies of mobile aquatic animals. Sound pressure levels (SPLs) and spectrum levels were reported for a window length of 10 s, which included at least one diver breath and bubble noise. Mean reported source levels (SLs) ranged from 164 to 158 dB re 1 μPa at 1 m for SCUBA using a bandwidth of 50 Hz-5 kHz. Depending on the environment, it was estimated that fish can detect SCUBA at over 200 m for noisy ocean conditions. The authors concluded that continuous broad band noise (up to 1.3 kHz) in their SCUBA recordings are due to exhaled bubbles. In another study, SPLs were reported for several unclassified SCUBA systems by 22, ranging from 131 to 147 dB re 1 μPa at 1 m. Configurations varied in terms of equipment, tank pressure, kicking intensity and breathing intensity.

Diver detection methods employing passive acoustics were analyzed for several environments in previous literature. A method to detect divers using a single hydrophone with a multi-band matched filter (23) was investigated in a tank and an estuarine environment (Hudson River, New York). Results indicated detection distances of about 50 m in the estuary. Extensive work in the same estuarine environment was conducted by 24, 25, 26, 14, 27 giving detection distance of more than 100 m for single hydrophone using a noncoherent envelope processor. 28, 29, 30 conducted several sea trials with single hydrophones, a six hydrophone array and electric underwater sensors. Detection distances for the port of Gothenburg (Sweden) were on the order of 30 meters for a single hydrophone. Nearshore reef environments (Kilo Nalu, Island of Oahu, Hawai‘i) were analyzed (13) using two 24-element L-shaped hydrophone arrays, with resulting detection ranges of about 20-30 m. Integrated passive detection systems to detect underwater sources (27) have been investigated, as have systems to deter swimmers (31).

To our knowledge, no spectrum levels or source signatures have been published for various SCUBA configurations in the open literature. This paper presents an analysis of different configurations of SCUBA diving equipment, including sound signatures, SPL, and spectrum levels. A method to remove reverberation due to the underwater recording environment (in this case, a pool) over a particular frequency band is presented and applied. SPLs are adjusted within a practical subband in the least-square sense using dereverberation techniques borrowed from speech processing.

This paper is structured as follows: First, an introduction to open circuit regulators is given. Sections 2.3 discuss relevant SCUBA equipment followed by a short introduction to different SCUBA designs. The experimental setup is introduced in Sec. 2.4 and methodology of data

analysis in Sec. 2.5. Signatures of different SCUBA systems are presented in Sec. 2.6, including SLs and dereverberation results, as well as several observations and recommendations.

2.3 SCUBA open circuit mechanics

A SCUBA setup enables a diver to breathe a gas mixture autonomously below the water surface. ”Open” refers to the state of the loop: in an open system, the diver exhales the gas while for a closed configuration (i.e., rebreathers), the gas is recycled. The gas mixture in an open SCUBA system is usually air, composed of 20.95 % oxygen, 78.09 % nitrogen and small amounts of trace gases by volume (32). However, divers may use different kinds of gas blends, depending on the application of the dive. Popular mixtures include enriched air commonly referred to as Nitrox 32 % and Nitrox 36 %. The amount of oxygen is increased (e.g. to 32 % oxygen) to decrease nitrogen absorption in the blood and increase dive time (32). The term ”gas” will be used instead of air when addressing the internal flow through the SCUBA gear.

Acoustically relevant parts of the gear are the first and second stages of the regulator and the pressure hose connecting the two. The first stage is the assembly which attaches to the tank valve and reduces high pressurized gas from the tank to an intermediate pressure. The intermediate pressure hose then delivers the gas to the second stage. The second stage, also called the primary regulator, is the part from which the diver breathes. Here, the intermediate pressurized gas coming from the first stage is reduced to ambient pressure which the diver can breathe (32). The first and second stage are likely to have different acoustic signatures, but it is beyond the scope of this paper to analyze them separately.

The signature of the SCUBA signal can be decomposed into two components: the demand (or gas intake) and the exhale. The demand fluctuates during a normal breath (33), which likely causes the signature to vary with time. The specific structure of the internal and moving mechanism of the regulator might also influence the demand signature. As the diver exhales, bubbles form and ascend to the surface; these create the exhale component of the SCUBA signature. Ideally both components would be recorded separately in order to quantify their individual signatures. However, bubbles take longer to reach the surface than the time needed between breaths even at shallow depths (i.e., several meters). Therefore, a separate recording of the inhale signature is impractical since any field experiment recording will have a mixed signature. In contrast, the exhale signature alone can be analyzed between two consecutive breaths.

As the diver demands gas, the first stage releases gas from the tank to its chamber. The pressure for an aluminum 80 cubic feet tank ranges up to around 3000 psia (pounds per square inch absolute, relative to a vacuum), while the intermediate pressure in the chamber is nominally 150 psig (pounds per square inch gauge, relative to atmospheric pressure), which may vary with equipment (32). The first part of the demand signature depends on the mechanism within the first stage which opens the high pressure valve. The second part of the demand signature is dominated by the flow noise of the gas through the chamber, pressure hose and primary regulator. As the gas demand ends, the intermediate pressure stabilizes and the high pressure valve closes. This is the third and final part of the demand signature. An complete analysis of SCUBA configuration must consider different mechanism and designs.

Generally speaking, two different configurations of SCUBA systems are commercially available: diaphragm and piston design. Each system can be either balanced or unbalanced. Both the first and second stage of a regulator might be balanced or unbalanced and might have either a piston or diaphragm design (32). A short introduction to each type is given in Secs. 2.3.1-2.3.4.

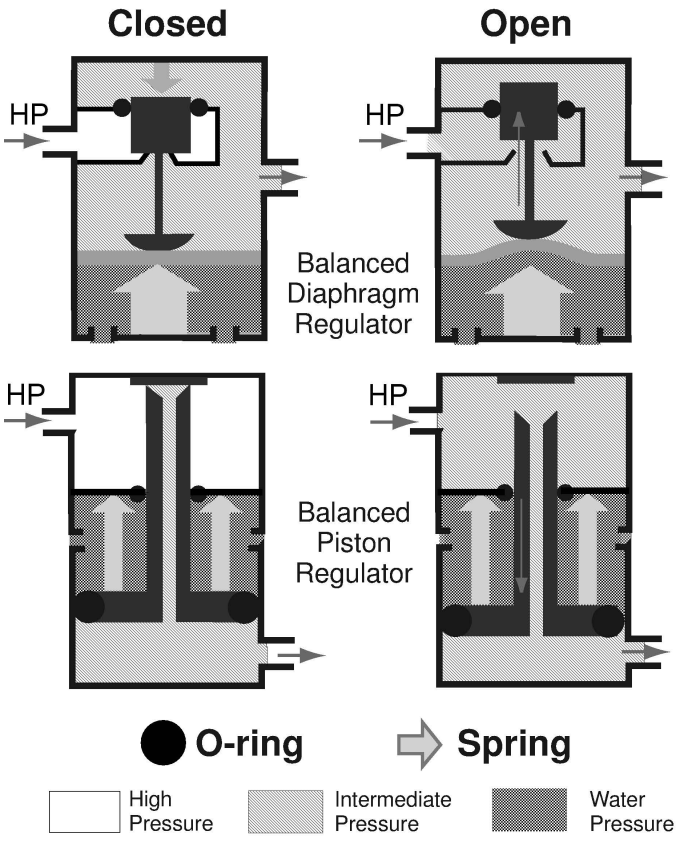

Figure 2.1. Comparison of balanced diaphragm and balanced piston regulator. Figure published with permission from NAUI Worldwide.

2.3.1 Balanced diaphragm

The diaphragm is a flexible rubber disk attached to a bias spring (see top diagram in Fig. 2.1). The disk separates water at ambient pressure and air pressure in the intermediate pressure chamber. Furthermore, the diaphragm connects to a lever piston, which opens or closes the high pressure (marked HP in the diagram) valve. The pressure valve is closed in its default position. As the diver demands gas, the pressure in the chamber decreases. The ambient water pressure pushes the diaphragm inward and opens the high pressure valve. The pressure difference required to open the high pressure valve is called the cracking pressure. The mechanism is reversed as the intermediate pressure increases when the diver stops demanding gas: the diaphragm moves to its resting position which closes the high pressure valve. This process is independent of tank pressure (hence the term ”balanced” regulator), which means that the diver needs to produce the same pressure differential to start the breathing process regardless of tank pressure.

2.3.2 Balanced piston

The mechanics of the balanced piston process are almost identical to those of the balanced diaphragm. The main difference is that the piston sits in a high pressure seat assembly in its resting position (see bottom diagram in Fig. 2.1). This assembly is attached to the chamber wall. When the diver inhales, the piston moves impulsively out of its seat. When the demand comes to a halt the process is reversed; the piston slams back into the high pressure seat assembly rather than bending back. This process is also independent of tank pressure.

2.3.3 Unbalanced diaphragm

The main difference between the unbalanced diaphragm and the balanced configurations above is the force balance within the first stage. For a balanced configuration, the intermediate chamber pressure opposes the bias spring and the ambient water pressure. For an unbalanced configuration, the intermediate chamber pressure and the tank pressure oppose the bias spring and the ambient water pressure. As the tank pressure is reduced, less force pushes the assembly into opening position. Therefore, the diver has to create a higher cracking pressure.

2.3.4 Unbalanced piston

Analogous to the difference between the balanced and unbalanced diaphragm, the unbalanced piston has a different force balance than the balanced piston does. The unbalanced piston mechanism uses the same configuration that the balanced piston does.

2.4 Experimental overview

Several experiments were conducted in a swimming pool (diving well) at the University of Hawai'i at Mānoa in December 2011, July 2012, August 2012 and June 2013. Pool pumps elevated the overall background noise level in the lower frequencies, caused flow noise as water flushed out of the pool, and produced transients caused by movement of the plastic door mechanisms attached to the outflow areas. Background noise was minimized by shutting off the pool pumps in the 2012 trials (facility management did not allow to shut off pool pumps during other experiments). The dimensions of the pool were 22.9 m x 22.9 m with a depth of 5.18 m, corresponding to resonance frequencies of 65 Hz and 290 Hz, respectively. SCUBA setups were selected to represent one of the four design combinations (balanced or unbalanced, diaphragm or piston). A total of four sets of regulators were tested (with approximate free flow rates in parenthesis):

1. Apeks XTX 200, balanced diaphragm (<300 Liter/min),
2. Oceanic SP-5 unbalanced piston first stage, unbalanced 2nd stage (250 Liter/min),
3. ScubaPro MK25 balanced, flow-though piston first stage and balanced 2nd stage G250 (<300 Liter/min),
4. Royal Mistral, unbalanced diaphragm, approx. manufacturing date 1962 (600-700 Liter/min).

SCUBA systems were recorded with a single 6050-C International Transducer Corporation (ITC, Santa Barbara, CA) hydrophone. Both hydrophone and SCUBA setup were suspended from the swimming pool surface to nominally 1.5 meter above the pool bottom, facing each other 2 meters apart. Fig. 2.2 shows a schematic of the underwater recording environment including dimensions. The letter S denotes the source (either diver or transmitting transducer) and H the hydrophone. A scientific diver descended with her own equipment and approached a suspended aluminum 80 cubic foot SCUBA tank attached to one of the test regulators. She recorded the tank working pressure, switched primary regulators to use the test regulator and started breathing for

two minutes while the hydrophone recorded her breaths. The diver was advised to breathe as regularly as possible. After 2 min passed, the diver switched regulators and dumped pressure out of the suspended tank until it reached the next pressure of interest and recording continued. After three iterations, the suspended SCUBA equipment was replaced with a fresh tank and regulator. Table 2.1 shows the number of recorded breaths at each pressure.

Table 2.1. Tank pressure and number of breaths recorded

Regulator	Rec. 1		Rec. 2		Rec. 3	
	[psi]	Breaths	[psi]	Breaths	[psi]	Breaths
Apeks	2600	24	2200	28	1200	13
Oceanic	3000	20	2000	13	1000	28
ScubaPro	1200	19	750	24	500	18
Mistral	1800	13	-	-	-	-

Data were recorded on a single channel of a custom analog to digital converter (ADC, 34) with a sampling rate of 192 kHz, a 10 Hz high pass filter and an antialiasing filter. The gain setting was chosen empirically such that the maximum amplitude of the recorded signal remained at $\sim$0.6 V to avoid clipping. The response of the 6050-C ITC hydrophone used in the recordings is nearly flat at -158 dB re 1 $V_{rms}/\mu Pa$ at 2m until about 30 kHz. Sensitivity increases in the 30-70 kHz band to $\sim$ -153 dB re $V_{rms}/\mu Pa$ at 2m. There is a resonance frequency at 50 kHz. The final band (70-96 kHz) is characterized by a steep roll-off.

To estimate the impulse response of the recording channel, the suspended SCUBA equipment was replaced with a Lubell speaker (Model LL916c, Lubell Labs, Inc., Columbus, OH). Just as the SCUBA equipment, the distance of the Lubell speaker to the hydrophone was 2 m with a height of 1.5 m above the floor. The Lubell speaker's channel was therefore the same as the recording channel of the SCUBA diver. The useful frequency range of the speaker ranged from 1 kHz to 18 kHz. Test signals included five 0.5 s linear frequency sweeps from 0.3 to 22 kHz followed by three 10 s of pink noise. This sequence of test signals was repeated three times.

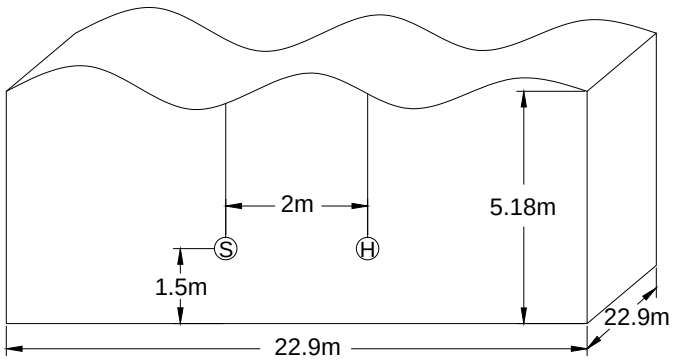

Figure 2.2. Schematic diagram of the diving well. The source S represents either the diver or transmitting transducer and H represents the hydrophone.

2.5 Analysis of data

2.5.1 Source levels

Source spectrum levels (SSL) represent the acoustic pressure as a distance of 1m away from the source. SSL are calculated using the passive sonar equations (2.5.1) as given by 35 and are a function of frequency in units of dB re 1 μPa^2/Hz at 1m. First, the power spectral density (PSD, denoted by S_{xx}) is computed after convolving the extracted diver signal with the inverse of the channel's impulse response. Afterwards, the PSD is adjusted for hydrophone sensitivity $|M_h|$ and ADC gain, G. Calibration curves and measured amplitude responses are calibrated using $V_{p2p/2}$ (volts peak-to-peak divided by a factor of 2).

$$SSL[f] = S_{xx} + |M_h| - 20\log(G) + 20\log(R) + \alpha(R), \quad (2.5.1)$$

$$SL = 10\log_{10}\sum_f 10^{SSL[f]/10}. \quad (2.5.2)$$

If recordings are made at a distance R $>$ 1 m, the PSD is further adjusted for spherical spreading and frequency dependent attenuation. The absorption coefficient, α, with units of dB/m is calculated for fresh water (36) with a temperature $> 20^\circ$ C, corresponding to a fresh water heated pool environment. To convert SSL to SL, computed SSLs (1 Hz bin width) are integrated using Eq. (2.5.2) and units are stated as dB re 1 μPa at 1m, including the bandwidth of integration.

When the PSD is computed without the inverse, the terminology *spectrum levels* and *SPLs* is used, respectively (note that the same units are applicable). Spectrum levels and SPLs therefore represent the energy 1m away of the acoustic source including reverberation effects. To emphasize a particular comparison when not including the inverse, the adjective *unadjusted* (i.e.,

unadjusted SPL) will be used. On the other hand, to emphasize a particular comparison when the inverse is included, the adjective *dereverberated* (i.e., dereverberated SPL) will be used.

To compute SCUBA signals, samples of 1.5 s duration were extracted for analysis. Each sample contained a single diver breath (about 1.4 s long) preceded and followed by a brief period of background noise. To calculate PSD the samples were windowed with a Kaiser window (37) using a beta value of 6.5 and normalized by a broadband normalization factor (38) to account for window effects. Samples were not filtered. Spectrum levels were adjusted for hydrophone sensitivity, ADC gain, spherical spreading and frequency dependent attenuation. Spectrum levels were integrated from 300 Hz to 80 kHz for SPL calculations. The lower frequency bound of 300 Hz was chosen to eliminate contributions from resonance frequencies of the pool. The upper bound of 80 kHz was selected to reduce system noise, allow for filter transitions near the Nyquist frequency, and reject samples with low SNR due to the roll-off of the hydrophone response.

2.5.2 Dereverberation of recorded signals

SCUBA signals were recorded in an environment subject to early reflections and late reverberations and, as a result, calculated SPLs are overestimated when calculated directly from the recorded signals. One way to remove additional energy in the recordings is to invert the acoustic impulse response (AIR) of the recording channel and to correlate the recorded signal with the resulting inverse. The inverse of the single hydrophone mixed-phase AIR can be significantly improved using a delay (39, 40, 41) to render it casual and improve stability. However, 42 found that equalization of AIR yields poor performance at offsets of fractions of a wavelength for a given channel, yielding incoherent dereverberation. Even though only approximate equalization can be achieved using single-channel least-squares (SCLS) methods (43), this technique can be efficiently employed in practical applications (44): SCLS filters are more robust to measurement noise and only partially equalize deep spectral nulls (45), reducing narrow band noise amplification after equalization.

The first step in dereverberating a recorded signal is to choose an appropriate length for the AIR (or correspondingly, the applicable reverberation time), which is based on several factors. First and most importantly, the AIR should include all of the early reflections and most of the late reverberant energy to account for the overall additional energy in the recorded signal of interest. However, the reverberant energy decays below the noise floor and poor SNR samples in the AIR should not be considered. Second, a delay will be added to the AIR, which increases the

dimensions of the matrix to be inverted. To minimize computational demand, the added delay and choice of decay time should be kept as short as possible.

Usefully, methods developed for room acoustics can be modified for underwater recordings in enclosed spaces since the same physical principles apply. In room acoustics, the term "reverberation" describes the reflected energy within an enclosure (3). After cutoff of a sound source, the energy that arrives at some point in the room decays due to attenuation. Plotting SPL against time gives decay curve. Reverberation time (T_{60}) is defined as the time it takes for the energy to reach one millionth of its initial value after the cessation of sound, and corresponds to a sound pressure level drop of 60 dB. In practice reverberation can be estimated by extrapolating the linear region of the decay curve to a 60 dB drop (3).

A simple and intuitive way to obtain a single decay curve is by playing colored noise (46) over a duration sufficiently long to allow the room to reach a steady state after which it is suddenly turned off; the decay curve is estimated by plotting SPL against time after the cessation. Decay curves from multiple realizations can be averaged to minimize random fluctuations; we averaged over three realizations in this study.

A more sophisticated way to obtain the decay curve is known as the the method of backward integration (47). Schroeder showed that the ensemble average of the squared signal decay $\langle h^2 \rangle$ is equivalent to an integral over the squared impulse response g,

$$\langle h^2(t) \rangle = \int_t^\infty [g(x)]^2 dx = \int_0^\infty [g(x)]^2 dx - \int_0^t [g(x)]^2 dx. \tag{2.5.3}$$

In practice, the upper bound of the integral is chosen to minimize the effect of the noise tail. For a single realization, Schroeder's method represents an improvement over the colored noise method because it gives the ensemble average of decay curves and is consequently insensitive to random fluctuations and more efficient (3).

In the work presented here, the AIR is obtained by correlating the recorded linear frequency sweep with the original sweep, both extended with zeros to twice their original length before correlation. A line of best fit applied to the linear portion of the resulting ensemble averaged decay curve is used to estimate the reverberation time (3). The tail of the AIR is subsequently truncated to reflect the reverberation time. The first arrival in the impulse response is set as the first sample point and the impulse response is scaled to unity.

Inversion of mixed-phased AIR is achieved using single channel least-squares (SCLS) techniques (see 48 for a good discussion on the related spiking filter and 45 for a general overview)

and the inverse solution is given by

$$\hat{f} = [H^T H]^{-1} H^T z. \tag{2.5.4}$$

In Eq. (2.5.4), $\hat{f}$ is the optimum inverse of the AIR in the least-squares sense. H is the circulant matrix of the AIR and $z = [0, 0..., 1, ...0, 0]^T$ where the spike (of value 1) occurs at the position of the delay. Originally, Robinson includes a noise prior in Eq. (2.5.4). The prior can be used to render the inversion more stable, e.g. to improve conditioning under poor SNR recordings. Since our recordings are made well above the noise floor, the noise prior is not used in the analysis. The performance of the inverse for a particular delay in the frequency domain is evaluated using the magnitude deviation (45) of the equalized impulse response:

$$\sigma = \left[\frac{1}{I} \sum_{k=0}^{I-1} (10 \log_{10} |\hat{D}(k)| - \bar{D})^2 \right]^{-1/2}, \tag{2.5.5}$$

where

$$\bar{D} = \frac{1}{I} \sum_{k=0}^{I-1} 10 \log_{10} |\hat{D}(k)|. \tag{2.5.6}$$

In equation 2.5.5 and 2.5.6, I corresponds to the length of the fast Fourier transform (FFT) with frequency bins k and Fourier coefficients, $\hat{D}$, of the inverted AIR. Magnitude deviation is invariant to the length of the FFT and, for ideal equalization, equates to zero. An appropriate delay is selected by plotting magnitude deviation versus delay and choosing a value corresponding to a minima. Afterwards, the inverse, delayed AIR is windowed using a combination of a Kaiser and rectangular window: a half-Kaiser window of 0.005 seconds is applied to the beginning and tail of the inverse AIR whereas all remaining coefficients are unaffected. Windowing the edges reduces undesired edge effects of the deconvolved signal. Deconvolution is achieved by convolving the SCUBA signal with the inverse impulse response in the time domain. The AIR is filtered before inversion over an appropriate sub-band, as discussed in Sec. 2.6.3. The filters have a ripple ratio of 0.1 dB, a stop-band attenuation of -60 dB, and transition bands of 500 Hz on both sides. After filtering, all signals are down-sampled to reduce computational load.

2.6 Results

In what follows, dB is used to abbreviate dB re 1 μPa at 1m when reporting sound pressure levels. For plotting, individual spectrum levels are smoothed (on the decibel scale) with a running average filter of 20 points unless otherwise noted.

2.6.1 Exhale signature

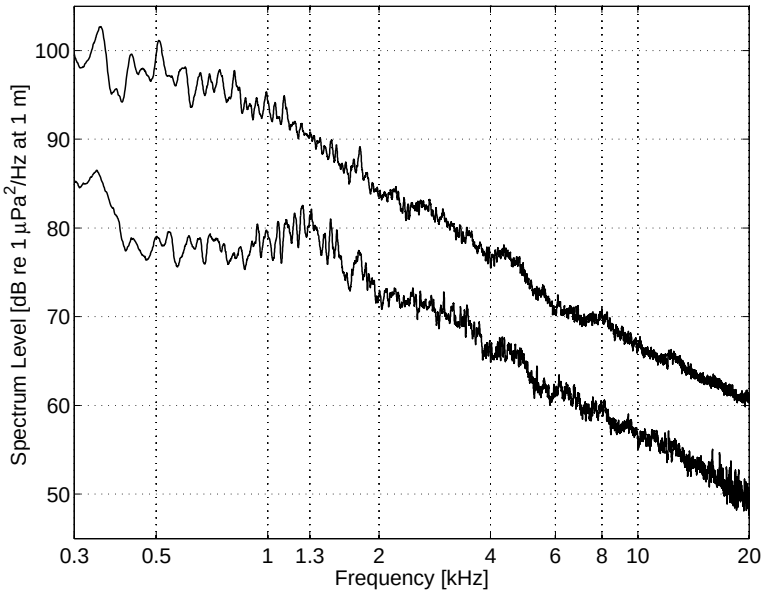

Figure 2.3. Ensemble averaged noise due to exhaled bubbles (top) and control (bottom).

Exhaled bubbles are present in all recordings at all times when a diver is present in the pool. Since bubbles from the previous breath are present when the diver takes a subsequent breath, the demand signature is inseparable from the exhale signature. However, the exhale signature can be measured in the absence of a diver breath by calculating spectrum levels between diver breaths.

Figure 2.3 shows bubble spectrum levels and control background spectrum levels (recorded before the diver entered the pool). Levels are ensemble averaged over 30 recordings from the Apeks regulator. Results indicate that the bubble signature is broadband and dominant below $\sim$1.3 kHz, with significant energy in higher frequencies as well. Spectrum levels of bubble signatures roll-off linearly past 1.3 kHz with a slope similar to that of the control. Bubble SPL is 127.8 dB in the 0.3-3.5 kHz band. 3.5 kHz is empirically determined (from Fig. 2.4) as the transition point between the exhale and the demand signature for the Apeks regulator.

Figure 2.4 shows spectrum levels of all four regulators against respective exhale bubble spectrum levels. 30 samples were used to compute regulator ensembles, except the Royal Mistral for which only 9 samples were used. As above, the exhale signatures are computed using samples taken between diver breaths. At lower frequencies all regulator ensembles closely follow the exhale signatures. The demand signature diverges from the exhale signature at $\sim$3.5 kHz for the Apeks

regulator, 4.5 kHz for the the Oceanic regulator, 6 kHz for the Scuba Pro regulator, and 1.3 kHz for the Royal Mistral. Above 6 kHz, no bubble noise is present for any of the regulators.

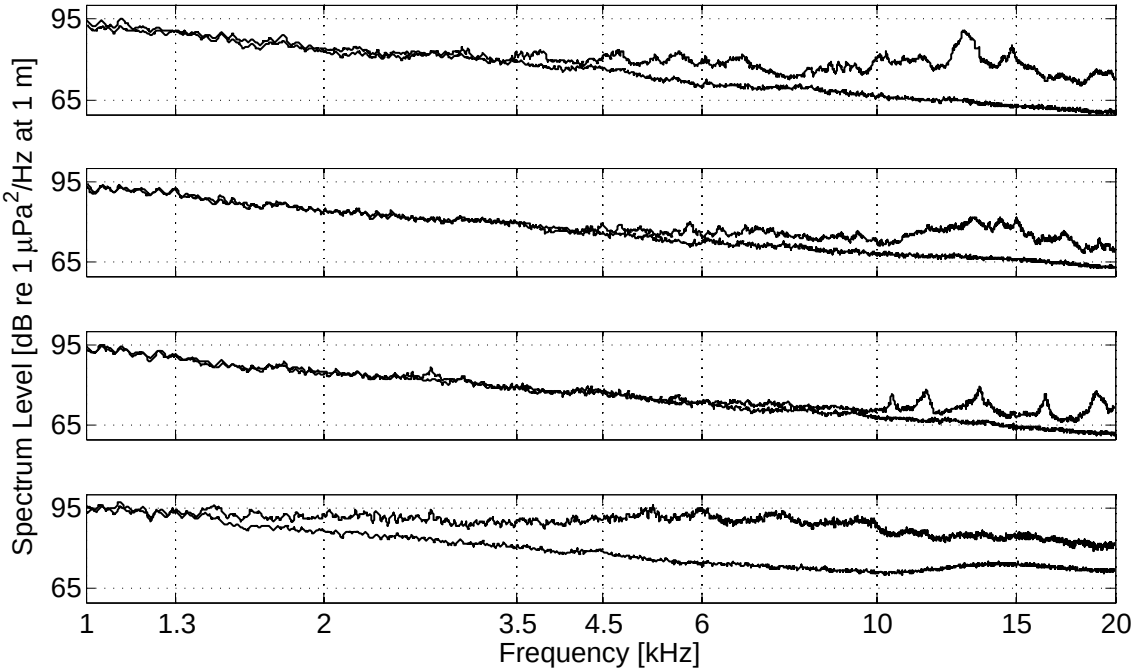

Figure 2.4. 4 Regulator spectrum levels (Apeks, Oceanic, ScubaPro and Mistral from top to bottom) showing transition from the exhale signature (bubble noise, bottom trace) at 3.5, 4.5, 6, and 1.3 kHz respectively from regulator signature

2.6.2 Sound pressure levels and signatures

Unadjusted signatures (waveforms, spectra and spectrograms) of all four regulators are presented in Fig. 2.5 and Fig. 2.6. All signals are broadband and likely extend beyond the 80 kHz limit of our recording system. A single breath can easily be identified: the sudden onset is abrupt over the whole band as the diver begins to inhale. Several bands display dominant energies throughout the breath, which vary from one regulator to another. All multi-stage regulators show narrowband energy peaks between 10 and 20 kHz and some energy peaks at higher frequencies (e.g. ScubaPro at 40 kHz and $>$ 60 kHz). These peaks can be exploited for detection and classification purposes. The Mistral contains a significant amount of energy between 4 and 10 kHz and a narrowband energy peak at 50 kHz (though the 50 kHz peak could be caused by the ITC resonance at this frequency, this is unlikely since the other regulators did not exhibit a peak at 50 kHz despite containing high-frequency energy).

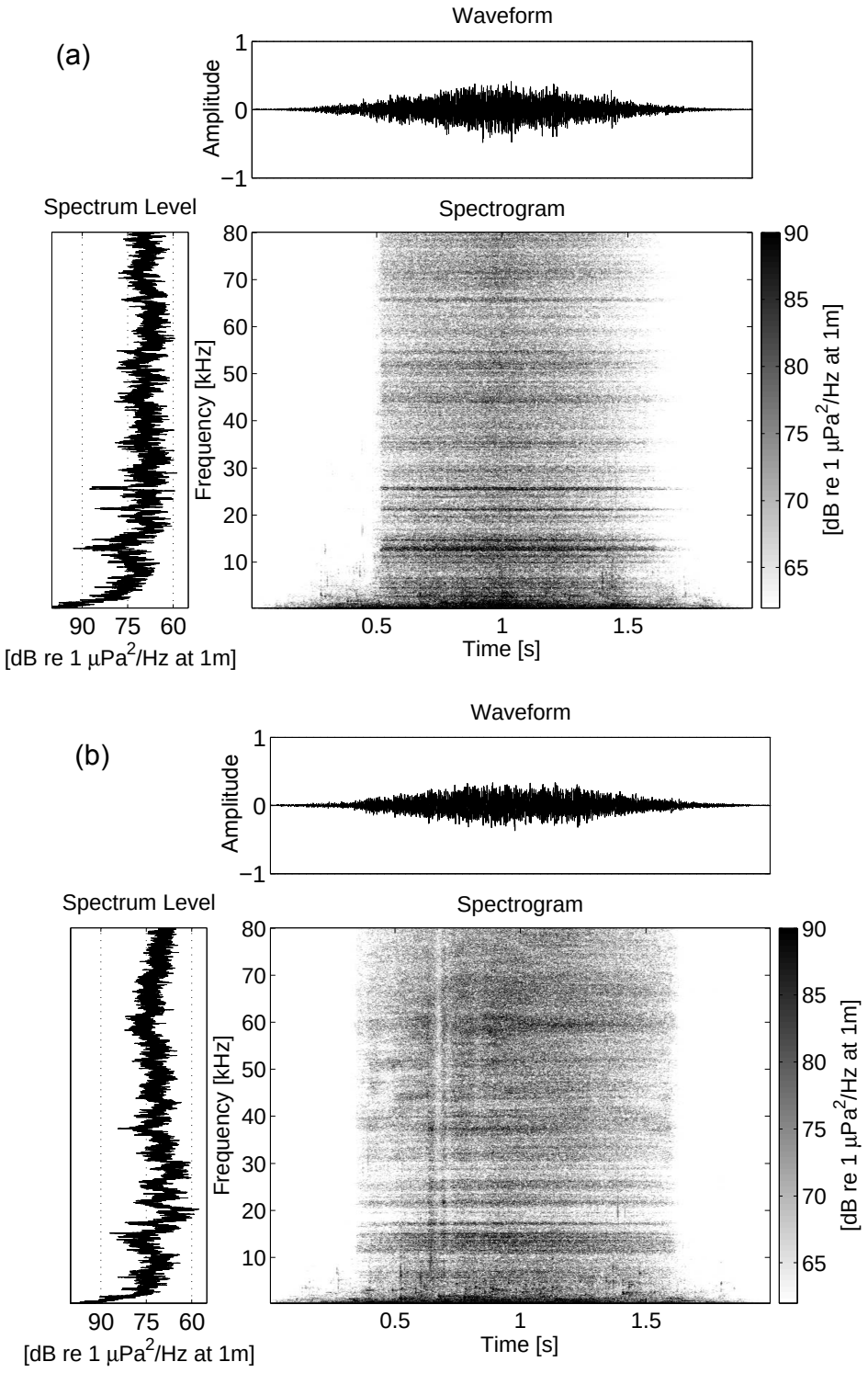

Figure 2.5. Spectrogram (2048 frequency bins at 50% overlap) of SCUBA signals: (a) Apeks XTX 200 and (b) Oceanic SP-5

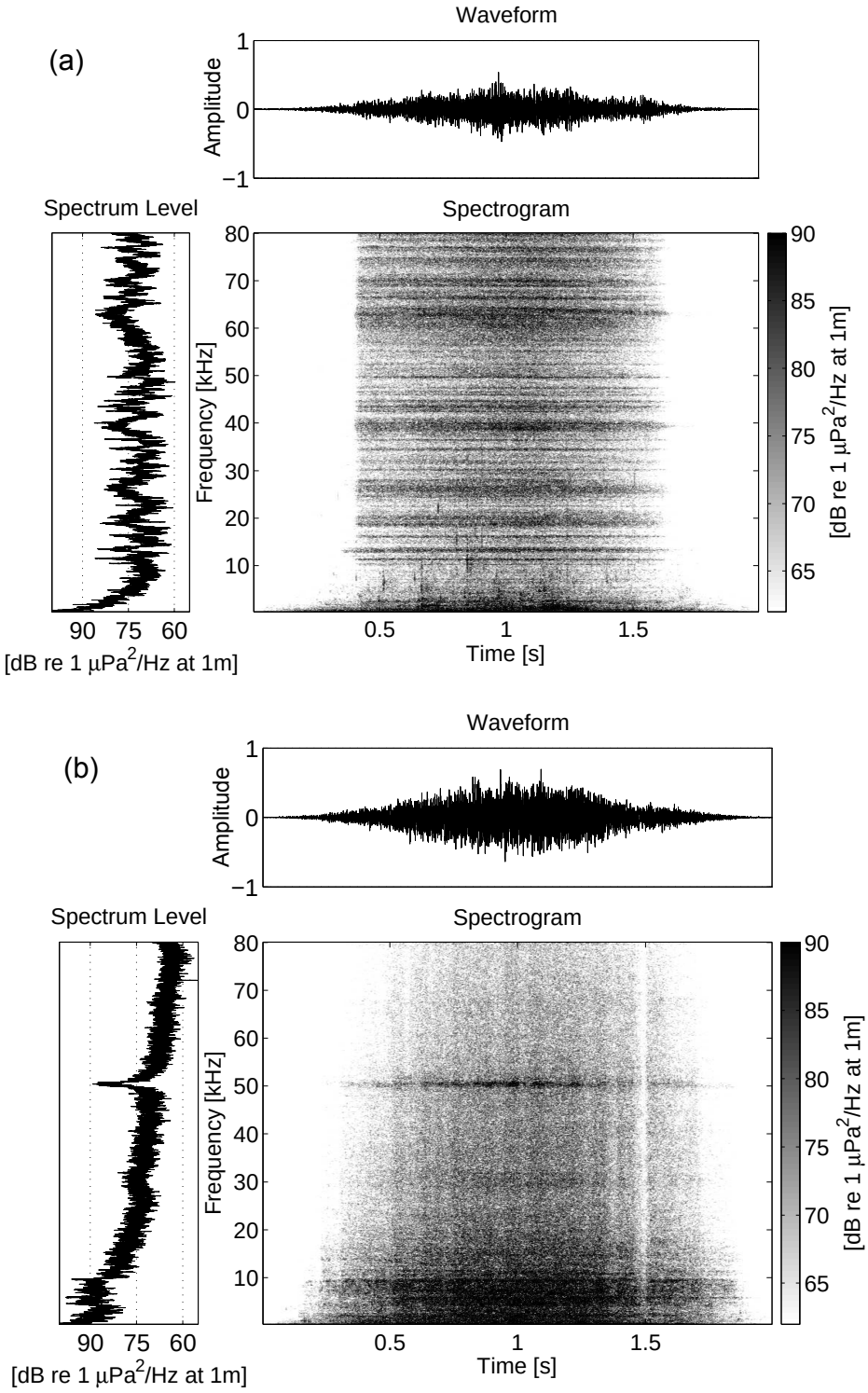

Figure 2.6. Spectrogram (2048 frequency bins at 50% overlap) of SCUBA signals: (a) Scuba Pro MK25 and (b) Royal Mistral

Table 2.2. SCUBA system sound pressure levels [dB re 1 μPa at 1m]. Parentheses indicate the number of recorded breaths. Top left: Total SPL (demand plus exhale signatures). Top right: Demand signature integrated over the 6-80 kHz band. Bottom left: Unadjusted recordings integrated over the 6-18 kHz band. Bottom right: Dereverberated recordings integrated over the 6-18 kHz band.

Regulator	Total SPL (0.3-80 kHz)			Regulator SPL (6-80 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	SPL before dereverb. (6-18 kHz)			SPL after dereverb. (6-18 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

SPLs are presented in Table 2.2. Mean unadjusted levels computed over the whole band are almost identical for all regulators, showing a 0.2 dB variation in SPL for the first three modern regulators (Apeks, Oceanic, and ScubaPro). For these three regulators, total mean levels are 130.3 $\pm$ 0.1 dB. The SPL of the Royal Mistral is $>$4 dB higher than that of the other regulators. The minimum computed SPL is 125.0 dB and the maximum is 135.9 dB. The 6-80 kHz band levels exclude the contribution from bubble noise and are computed to compare demand levels only. Mean levels for the first three regulators are 127.5 $\pm$ 0.5 dB. The SPL of the Mistral in this band is 3.9 dB higher than that of the other regulators. Several significant lower outliers in SPL were identified in the second recordings of the Oceanic (2 breaths) and ScubaPro (10 breaths) regulators, with a minimum of 110 dB. It is believed that these outliers were a result of unusually weak breaths made by the scientific diver. Overall, variations in SPL of the demand signature (i.e. 22 dB for the ScubaPro) are much larger than the difference computed over the whole band (i.e., 8 dB for ScubaPro).

The experiment in 2013 was conducted to investigate the relationship between breathing intensity and SCUBA SPL. It used the Apeks regulator only. Results were integrated over the full bandwidth of 0.3-80 kHz. Computed SPLs range from 116.3 to 131.7 dB. The difference between

peak demand and quiet breathing is 15.4 dB. Results integrated over the 6-80 kHz band (regulator signature only) range from 110.5 to 130.7 dB (a 20.2 dB difference).

2.6.3 Dereverberated signatures

The SPLs for the pink noise recordings are 25 dB above the noise floor and range from 104 to 129 dB. The decaying signal reached the noise floor in ~ 0.255 seconds, giving a reverberation time (T_{60}) estimate of 0.55 s. In comparison, backward integration method by Schroeder yields a reverberation time close to 0.4 s. A 0.25 s AIR length (corresponding to an energy decay of 37.5 dB) is used for the inversion procedure: the AIR included most of the reverberant energy, it was kept as short as possible for inversion, and reduced low SNR contributions from the tail (at ~ 0.25 s, mean regulator SPLs generally fell below the noise floor). 0.25 seconds corresponds to a filter length of 9250 taps. To select the appropriate delay for AIR inversion, magnitude deviation of the equalized AIR is plotted as a function of inversion delay, using Eq. (2.5.5) (Fig. 2.7). A delay of 226.5 ms (corresponding to a logarithmic deviation of 0.68 dB) is selected for the inverse; further performance improvement is minor and costly in terms of matrix inversion.

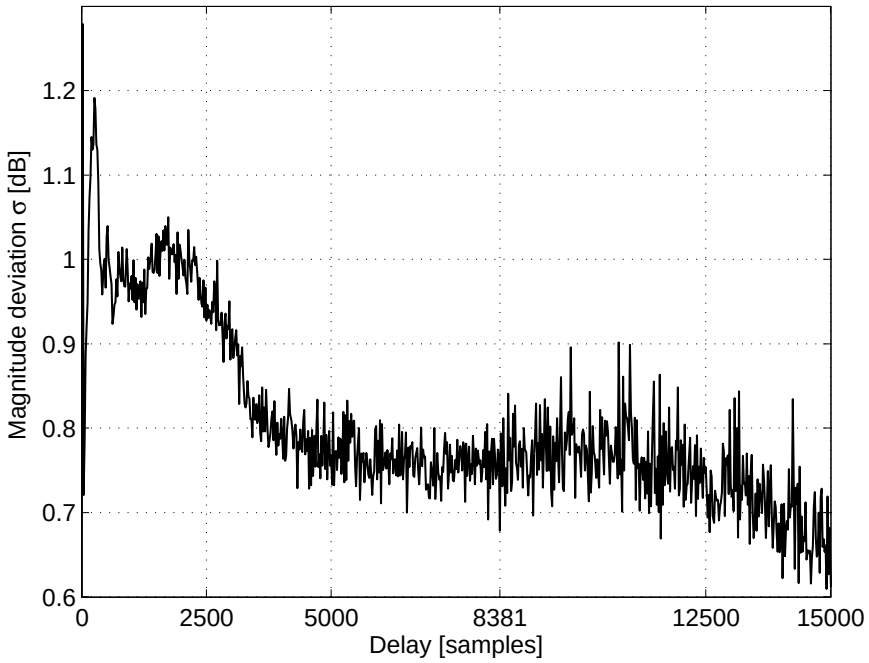

Figure 2.7. Magnitude deviation of equalized AIR versus delay.

Figure 2.8 shows the recording channel AIR, the AIR inverse, and the equalized AIR. Sparse early reflection are present in the first 10 ms of the AIR, while the later response becomes

more diffuse. The pool-floor reflection arrives at ~ 2 ms after the direct arrival, while the first wall and surface reflections arrive 6 ms after the first arrival. The inverse of the AIR is shown in Fig. 2.8b, centered and zoomed in at the point of delay. The equalized AIR in the frequency domain is shown in Fig. 2.8c. Several frequencies are not perfectly equalized and deviate from the mean by more than 10 dB. The dereverberation technique is only applicable for a fixed channel, whereas the

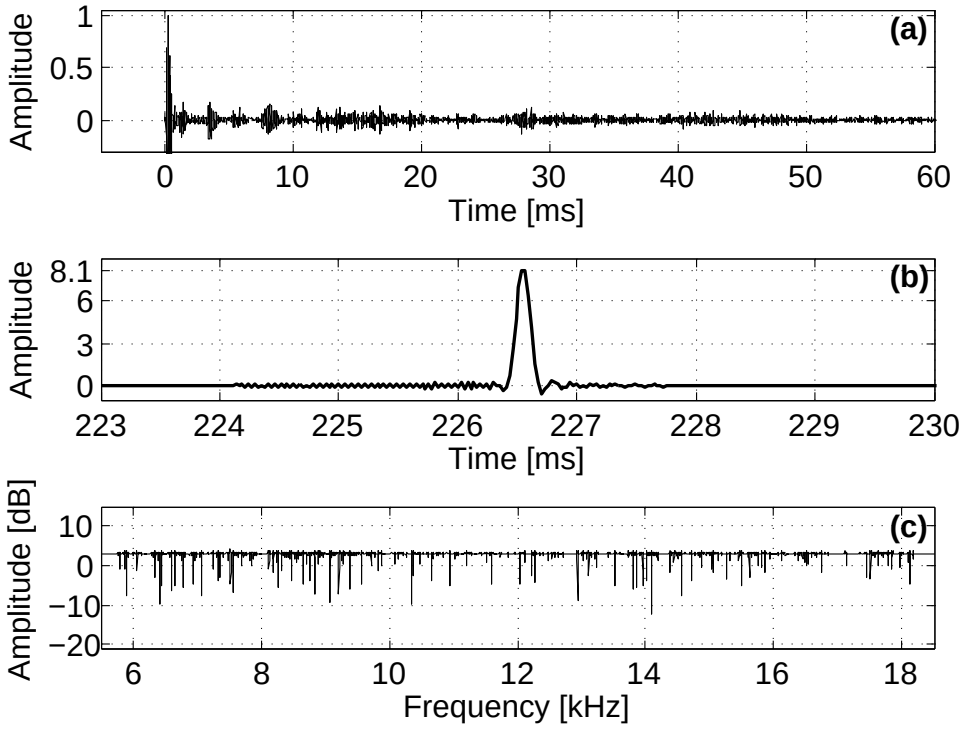

Figure 2.8. Acoustic impulse response (a), AIR inverse (b), and equalized AIR (c).

exhale signature is not fixed: the bubbles do not stay within the recording channel, but ascend to the surface. Consequently, the dereverberation method cannot be applied to the part of the acoustic signature in which the exhale signature is dominant. 6 kHz is the highest transition frequency between bubble and regulator signature for all regulators and is selected as the lower frequency bound for the dereverberation procedure (see Fig. 2.4). The usable frequency range of the speaker limits the upper bound to 18 kHz. The SCUBA signals and the AIR are bandpass filtered between 6 and 18 kHz and down-sampled to 37 kHz to reduce complexity of the system for dereverberation. SLs are given in Table 2.2. SLs are between 5.5 dB (Oceanic) and 6.8 dB (Mistral) lower than the SPLs before dereverberation. The new minima is 101.6 dB and the new maxima is at 124.3 dB.

Figure 2.9 shows a comparison between the unadjusted Apeks spectrum levels and the dereverberated SSLs. Mean source levels are 6.3 dB lower (6-18 kHz band) for the Apeks regulator.

The outstanding energy spike at 13 kHz seems to be invariant to the dereverberation procedure and, therefore, seems to be a property of the acoustic signature of the regulator rather than an artifact of reverberation.

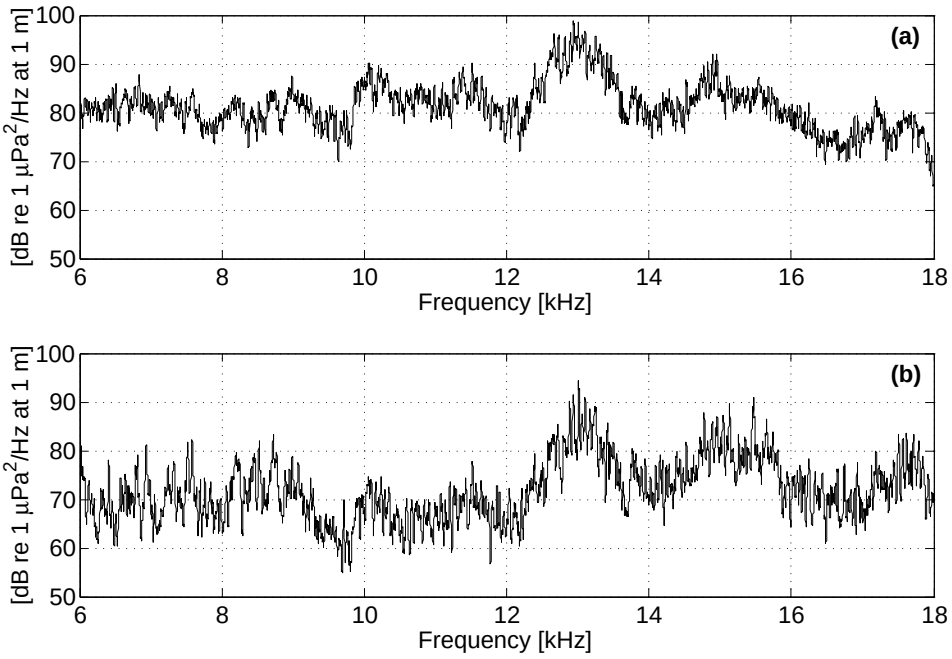

Figure 2.9. Apeks unadjusted spectrum levels (a) and source spectrum levels (b).

2.7 Discussion

SPLs are similar for all modern regulators with an overall mean close to 130 dB for the unadjusted level with a SCUBA diver breathing normally. The higher mean SPL of $\sim$ 4.5 dB for the Mistral regulator might be due to higher flow rate, a single stage design, or because a different scientific diver performed that particular test (increased breath length within a typical interval of integration or more bubbles in the water column). Variations due to the mechanical design could not be observed and might require an in depth investigation using higher time-frequency resolution techniques.

Comparing the combined and inhale only parts of the signature leads to several observations. The lower frequency band (0.3-6 kHz) carries about half the acoustic energy in comparison to the full band (0.3-80 kHz). Second, SCUBA SPL (without exhale contribution) changes as much as 22 dB (ScubaPro, 6-80 kHz). The diver can (at least partially) control this range in SPL by vary-

ing breathing intensity. The combined signature (0.3-80 kHz) changes by 8.1 dB (ScubaPro, 0.3 kHz-80 kHz). The diver does not have much control over this part of the signature and bubble noise is a significant part of the energy below 6 kHz (but note that the position of this bound depends on the regulator under consideration). A diver who wants to reduce his/her SPL would therefore choose a system which reduces the amount of acoustic energy due to the bubble signature to avoid detection by a sensor or aquatic species.

The mean unadjusted SPLs reported here are 1 dB lower than the lowest SPLs recorded by 22 (131-147 dB re 1 μPa at 1 m), while the upper end of SPL ranges reported here fall well within their range. 22 made measurements in a small tank (312 ft long, 12 ft wide and 6 ft deep), and it is possible that additional energy due to reverberation and resonance might have elevated their levels. Their bounds of integration for SPL calculations are unknown and it is possible that additional energies $<$ 300 kHz significantly changed their results. 22 cited a range of about 16 dB caused by intentional change in demand intensity by the diver. Results in Table 2.2 show a SPL range of 10.9 dB in the full band (0.3-80 kHz) and 22 dB in the sub-band (6 kHz-80 kHz), however the diver was breathing as regularly as possible for these recordings. Our 2013 recording for the Apeks regulator give a SPL difference of 15.4 dB which is very close to the 16 dB difference reported by Donskoy et al.

The SPLs reported by 21 are $\sim$ 30 dB higher than mean unadjusted levels measured here over the full band. The factors contributing to this difference cannot be confirmed, however their experiment setup differed significantly from ours as did the objective of their work, which was to estimate the distance at which SCUBA can be detected by aquatic animals. Radford et al. made recordings in a 20 m water column which could have resulted in higher levels of bubble noise due to a longer water column. The lower frequency bound for calculating SPL in our experiment is 300 Hz while 21 used 50 Hz; it is likely that considerable additional energy exists in the low frequencies [indeed, the spectrograms presented by Radford et al. show a significant amount of energy below 2 kHz]. Their observation that bubble noise is dominant until about 1.3 kHz is consistent with our results (see Fig. 2.3).

Our results indicate that dereverberation of recorded underwater, stationary point-sources can be achieved in a controlled environment if the AIR of the channel can be estimated. Reverberation time is estimated using two methods which produced significantly different results [0.4 s with the method by Schroeder and 0.55 s with the pink noise decay]. The method by Schroeder is an ensemble average of all individual decay curves and eliminates random fluctuations so that resulting reverberation time estimates are more reliable.

The inversion procedure given by Eq. (2.5.4) resulted in a square, recursive Toepliz matrix on the order of 15,000 by 15,000. Such large matrices can be computationally problematic for moderately long AIR using high sampling rates. In addition, the method used here is sensitive to differences in position between the control source (used to measure the AIR) and test source. The method is only applicable to fixed sources since it requires that the channel is fixed. In addition, particular care must be taken for mobile sources (such as divers) to ensure that the test source is in the same location as the control source. Constant frequency signals exceeding the length of the AIR are less sensitive to exact alignment. Additional errors arise since most sources are not point-sources and have different directionality characteristics than the transducer. Figure 2.8c shows that perfect equalization is not achieved since several frequencies are not equally attenuated. The drawback of the dereverberation method used here is the ill conditioned inverse of the SCLS solution, yielding inaccurate equalization performance.

2.8 Summary

The SCUBA SPLs and spectrum levels reported here represent a first step towards designing a passive acoustic SCUBA detector for near-shore and port security applications. Our results suggest that past published levels overestimate the energy of SCUBA regulators. For the three modern regulators tested, mean unadjusted SPLs were close to 130 dB re 1 μPa at 1m. Given the transition between exhale and demand signature (1.3-6 kHz), it may be possible to exploit the characteristics of the bubble signature for classification purposes (e.g. SCUBA vs non-SCUBA). Since bubbles lag the diver position, their signature might be exploited to produce an additional vector for short duration tracking.

A method to remove reverberation for the use of underwater passive acoustic experiment was presented and used to remove additional energies due to reflections of the pool, accounting for as much as 6.8 dB over the 6-18 kHz band. The method can be used (and extended) for any type of recording environment to characterize a source. Since SCUBA and especially rebreathers (21) have low SPL and the ambient noise field can be unpredictable, it seems plausible to focus on detection rather than tracking. To increase gain, a linear array analog to a tripwire (29) seems to be the optimum solution to detect divers in a noisy harbor or nearshore environment.

2.9 Future directions

The dereverberation procedure can be improved in terms of robustness while keeping simplicity in mind. Such a method will be applicable for a wide range of underwater pool experiments, which could be used for sources such as autonomous underwater vehicles or similar. It is desirable to quantify equalization mismatch using different signals for transfer function estimation and different decay times. The SCLS filter can then be applied to control signals, yielding objective performance results.

2.10 Acknowledgments

The authors would like to thank the reviewers for the very helpful comments and suggestions. The authors thank research divers Troy Heitmann, Lauren Tuthill and Keo Lopez. John Allen and Tyler Hee Wai contributed the Lubell speaker and recording support for experiments conducted in 2012. Volker Roeber helped with the experimental setup. Guidance from Margo Edwards, Director of the National Center for Island, Maritime and Extreme Environmental Security, helped to direct the study. Keoki Stender and David Pence from UH Diving Safety Office helped with diving related questions. This paper is dedicated to Kevin Flanagan†, our SCUBA instructor and friend. 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.

Chapter 3

Source characterization using recordings made in a reverberant underwater channel

3.1 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 various excitation methods to obtain the IR. IR length is estimated using methods borrowed from room acoustics and inversion of non-minimum phase IR is accomplished in the least-squares sense. 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.

3.2 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 (1, 2) 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 (1) is $\sim$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 (2) 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 convolving the recorded signal with the inverse of the IR equalizes the channel and can be used to obtain an anechoic estimate of the source signal. Estimating the acoustic IR can be accomplished via a matched filter by correlating the excitation signal with the received signal (49, 50). 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 measurement chain, this procedure has to be adjusted because the transmitting transducer can not realize an impulse. The excitation signal is selected or 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. The IR is deconvolved by exciting the system under investigation with the excitation signal and

correlating the output with the input (see Sec. 3.3). Once the IR is deconvolved and its length is estimated (see Sec. 3.4), it can be inverted.

Two main methods have been investigated in room acoustic literature to coherently invert an acoustic IR: homomorphic deconvolution (51, 39, 52) and single channel least squares (SCLS) (53, 54, 44). 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 (55). However, non-minimum phase signals have cepstral overlap and the direct arrival can not be easily separated from early reflections. It was found (39) 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 (44). The inverse of a mixed-phase IR in the least-squares sense can be significantly improved using a processing delay (39, 40, 41) to render it casual and improve stability. Even though only approximate equalization can be achieved (43), SCLS is robust to measurement noise and only partially equalizes deep spectral nulls (45), hence reducing narrow band noise amplification after equalization. In addition, it can easily evolve into a multi-channel method (43).

In a preliminary experiment (4), 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 Sec. 3.3 and Sec. 3.4 described the proposed experimental procedure which is validated by an experiment (Sec. 3.5). Methodology of data analysis is presented in Sec. 3.6 followed by results in Sec. 3.7. The paper concludes with a discussion in Sec. 3.8.

3.3 Mathematical Formulation

Figure 3.1(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:

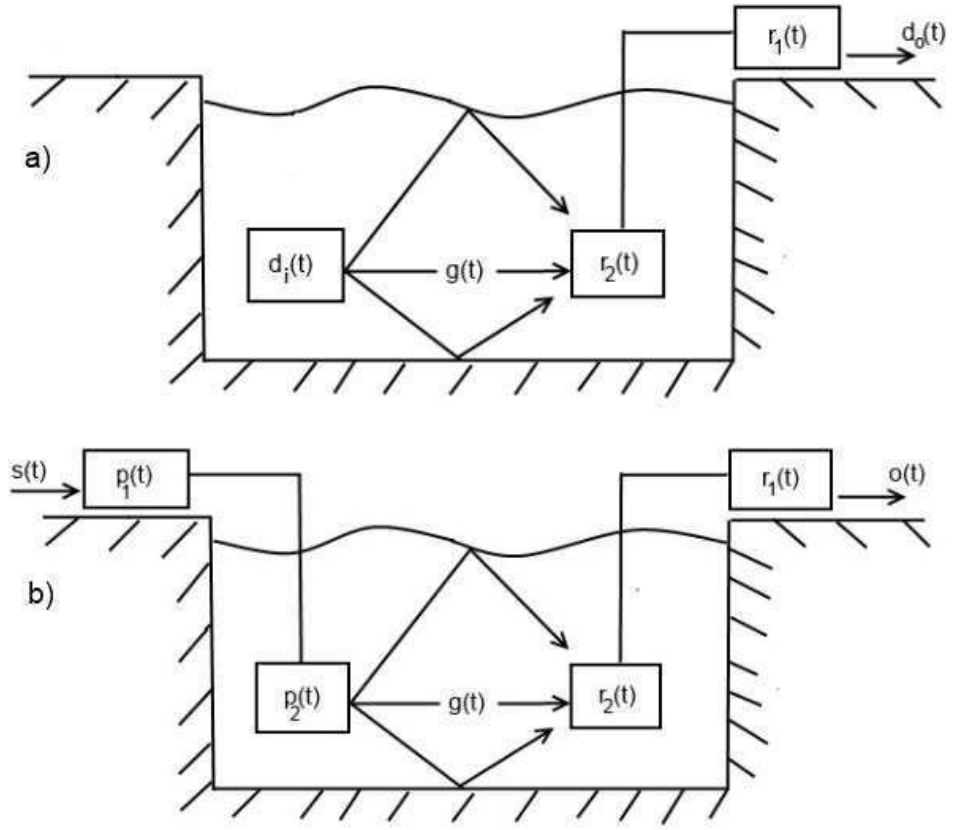

Figure 3.1. Pool diagram showing schematics of (a) the inverse problem with an unknown source and (b) the forward problem with a known source.

$$d_o(t) = r_1(t) * r_2(t) * g(t) * d_i(t). \tag{3.3.1}$$

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. 3.1(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. (3.3.2)).

$$h(t) = r_1(t) * r_2(t) * g(t) * p_2(t) * p_1(t) \tag{3.3.2}$$

$$o(t) = h(t) * s(t) \tag{3.3.3}$$

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. (3.3.3)). 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 distribution of both amplitude and phase is Gaussian, the sinusoidal pressure in the channel then follows a Rayleigh distribution (3) 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 the the hydrophone. The 68% range of the sinusoidal sound pressure level (SPL) distribution (corresponding to approximately one standard deviation (SD), denoted by σ) is given by 56. Here, this equation is slightly modified to average over non-uniform absorption coefficients:

$$\pm\,\sigma \approx 40r \left(1 - \sum_{i=1}^{6} \frac{\alpha_i}{6}\right)^{\frac{1}{2}} \left(\sum_{i=1}^{6} \alpha_i A_i\right)^{-\frac{1}{2}} \text{ dB.} \tag{3.3.4}$$

Note that the actual pressure distribution for broadband signals with different amplitudes is more complicated (57, 58) and Eq. (3.3.4) will be used to approximate the SD (averaged over all frequencies) of the stochastic system. The system's IR h(t) must be approximated by a sufficient number of realizations and its expectation will be denoted by $\mathbb{E}[h(t)]$.

Each realization of h(t) can be estimated using an excitation signal (59): the output signal is cross-correlated with the input signal (Eq. (3.3.5)), in a manner analogous to a matched filter. The complex conjugate of $s(-t)$ is denoted by $\overline{s(-t)}$. This procedure scales h(t) by the autocorrelation of the input signal. It is most practical to scale the excitation signal after the experiment (which might have been altered to avoid clipping) by its autocorrelation (the scaled signal is denoted by $s_u(-t)$) to deconvolve the IR h(t).

$$o(t) * \overline{s(-t)} = h(t) * s(t) * \overline{s(-t)} \tag{3.3.5}$$

$$o(t) * \overline{s_u(-t)} = h(t) \tag{3.3.6}$$

This method ensures that h(t) is invariant to the amplitude of the input signal and Eq. (3.3.5) reduces to Eq. (3.3.6). Temporal inversion causes a phase inversion and the cross-correlation results in a pure delay of h(t). Hence h(t) has no amplitude or phase contributions due to the excitation method. 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 (49) to the deconvolved IR: this constant is zero if the excitation signal has no trend, which can be ensured by proper signal design.

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 $\mathbb{E}[h(t)]^{-1}$ (Sec. 3.4 discusses coherent inversion) to compute an estimate of the input $\hat{s}(t)$ (Eq. (3.3.7)). Note that the performance measure does not require explicit knowledge of the playback or recording equipment's IRs.

$$o(t) * \mathbb{E}[h(t)]^{-1} = h(t) * s(t) * \mathbb{E}[h(t)]^{-1} = \hat{s}(t) \tag{3.3.7}$$

The unknown source signal in the inverse problem can now be found by convolving the recorded signal in Eq. (3.3.1) 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:

$$\begin{aligned} d_o(t) * \mathbb{E}[h(t)]^{-1} &= r_1(t) * r_2(t) * g(t) * d_i(t) * \mathbb{E}[h(t)]^{-1} \\ &= \hat{d}_i(t) * (p_2(t) * p_1(t))^{-1} \\ \hat{d}_i(t) &= d_o(t) * \mathbb{E}[h(t)]^{-1} * p_2(t) * p_1(t) \end{aligned} \tag{3.3.8}$$

If the phase response of the playback system is unknown, Eq. (3.3.8) can be computed incoherently to get sound pressure levels (Eq. (3.3.9)). As in the passive sonar equation (35), source spectral levels (SSL) can be computed by estimating the power spectral density (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 $|P_1|$.

$$SSL = 10\log_{10}(S_{d_o}) - 20\log_{10}(\mathbb{E}[|H|]) + 20\log_{10}(|P_2|) + 20\log_{10}(|P_1|) \tag{3.3.9}$$

Units for this discrete equation as well as its standard deviation σ should be stated as $\pm\sigma$ [dB re 1 μPa^2/Hz at 1m] for a channel length of 1 m.

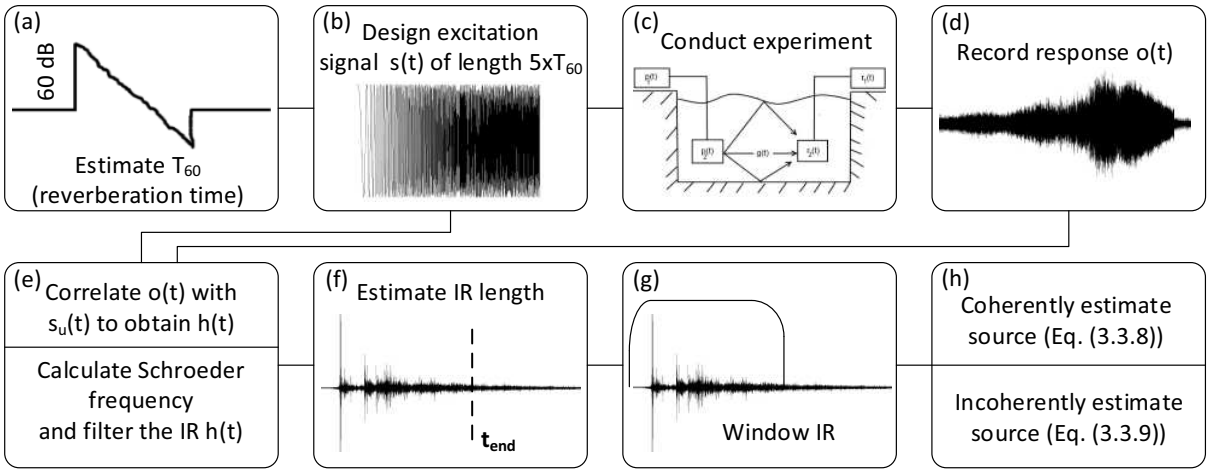

Figure 3.2. Flowchart showing the proposed procedure to obtain (a) - (g) the impulse response in the forward problem and (h) the estimate of the unknown source in the inverse problem. A detailed description of each step is given in Sec. 3.4

3.4 Proposed experimental procedure

The following steps (Fig. 3.2) 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 practise, the IR is obtained by circular deconvolution rather than by linear deconvolution (i.e. Eq. (3.3.6) is computed in the frequency domain). Circular deconvolution is possible for excitation signals that are longer than the IR because distortion products appear in the noise floor and can be discarded by windowing the signal. Shorter excitation signal will fold back into the IR (circular aliasing). In addition, it is desirable to have a long excitation signal to ensure transducers have sufficient excitation time at lower frequencies (delayed low-frequency components). For this reason, 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 (3). If the room is large relative to the lowest frequency of interest, it is sufficient to consider propagation of sound energy (i.e. phase information is not required) (60).

To determine the minimum length of the excitation signal, we need to know that 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. 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. 3.2(a)) in the dereverberation procedure is to estimate T_{60}. This is accomplished using a formula borrowed from room acoustics (3):

$$T_{60} = \frac{24ln(10)}{c} \frac{V}{-S\log(1 - \sum_{i=1}^{6} \alpha_i A_i / S)}, \tag{3.4.1}$$

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), and S (in m^2) represents the combined area of all underwater walls and water/air boundary of the rectangular enclosure. The absorption coefficient for each boundary can be estimated using

$$\alpha_i = 1 - \left| \frac{z_w - z_i}{z_w + z_i} \right|, \tag{3.4.2}$$

where z_w represents the acoustic impedance of water and z_i the acoustic impedance of the boundary. Standard values for z are 415 Ns/m^{-3} (air), 8×10^6 Ns/m^{-3} (concrete) and 1.5×10^6 Ns/m^{-3} (water) (61, 62).

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. 3.2(b)) than T_{60}. The longer excitation signal increases the likelihood that circular aliasing is avoided when using circular deconvolution for fast computation times. 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. 3.2(c)). Noise sources such as pool pumps and water overflow mechanics should 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 (63) 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. (3.3.7)). 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. 3.2(e)) by correlating the recorded response for each recorded realization (Fig. 3.2(d)) with the scaled excitation signal (Eq. (3.3.6)). Afterwards, the IR is filtered to remove resonance frequencies. As noted above, the room must be large relative to the lowest frequency of interest. The lowest frequency for which this applies is called the Schroeder frequency f_s and is given by (1)

$$f_s = 0.6\sqrt{\frac{c^3 T_{60}}{V}}. \tag{3.4.3}$$

In practice, the IR is band-pass filtered at this step: while the lower bound is given by Eq. (3.4.3), 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. 3.2(f)) with more accuracy than the estimate obtained using Eq. (3.4.1). We propose using the measure of echo density (64) 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 function of time and can be plotted concurringly 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. 3.2(g)).

Once estimated for each realization, IRs are averaged to estimate the unknown source (Fig. 3.2(h)) either coherently (Eq. (3.3.8)) or incoherently (Eq. (3.3.9)). 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 (48) and a detailed overview can be found in 4. The optimum inverse $\hat{f}$ (MoorePenrose pseudoinverse) in the least-squares sense is given by

$$\hat{f} = [A^T A]^{-1} A^T z. \tag{3.4.4}$$

A is the circulant matrix of the IR and $z = [0, 0..., 1]^T$, 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.

3.5 Pool experiment

Table 3.1. Overview of recorded signals with lower frequency f_0, step size f_s and stop frequency f_1. Frequency units are in kHz.

Signal Type	Length [s]	f_0	f_s	f_1	Repetition	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
Sinusoids (+2cm)	5	5	1	85	1	3
Sinusoids (+4cm)	5	5	1	85	1	3

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 and maximum sensitivity 111.5 dB re 1 $\mu Pa/V$ at 10 kHz and 136.5 dB re 1 $\mu Pa/V$ 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 since SCLS can evolve into a multi-channel method, 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 $V/\mu Pa$ throughout the whole band of interest (10 kHz to 70 kHz). The minimum and maximum

sensitivity are -172.6 dB re 1 $V/\mu Pa$ at 12.4 kHz and -168.6 dB re 1 $V/\mu Pa$ 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 volts 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 seconds 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 3.1 shows an overview of all signals played, their respective length, frequencies and total number of repetitions.

3.6 Analysis of data and performance measures

The following sub-section discusses how T_{60} can be estimated from recorded data to validate Eq. (3.4.1) (Fig. 3.2(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. 3.2(f)). The following sub-sections discuss how the deconvolved IR is filtered (Fig. 3.2(e)), how its window is computed (Fig. 3.2(g)) and performance criteria for coherent inversion (Fig. 3.2(h)). The final sub-section discusses dereverberation performance measures for control sources.

3.6.1 Estimating T_{60} and T_{sn} from data

A sophisticated way to obtain T_{60} experimentally is known as the the method of backward integration (47) and was previously used (4) 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. (3.4.1).

In particular, 65 modified Schroeder's method by subtracting an average noise term $\bar{\eta^2}$ from the squared IR g with additive noise η:

$$\begin{aligned} \langle h^2(t) \rangle &= \int_t^\infty ([g(\tau) + \eta(\tau)]^2 - \bar{\eta^2}) d\tau \\ &= \int_t^\infty (g^2(\tau) + 2g(\tau)\eta(\tau) + \eta^2(\tau) - \bar{\eta^2}) d\tau \end{aligned} \tag{3.6.1}$$

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. (3.6.1) is integrated over a time much longer than T_{sn}, the trend is dominated by the noise term when 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 (Eq. (3.6.1)).

3.6.2 Filtering of Data

After the IR is deconvolved from the recorded signal (Fig. 3.2(e)), it is filtered over the bandwidth of interest. The lower frequency bound is given by Eq. (3.4.3). Here, the Kaiser bandpass filter has a pass-band from 5-75 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. 3.2(g)) using a combination of left half-kaiser window (right edge at direct arrival), rectangular window and right half-kaiser window (right edge at t_{end}). After windowing, all signals are downsampled to 150 kHz.

3.6.3 Coherent IR inversion

If a coherent estimate of the unknown source is desired (Fig. 3.2(h)), the IR needs to be inverted using SCLS formalism. 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 $D = \hat{f} * h$ is a delta function centered at the position of the delay m (same delay as the spike in vector z in Eq. (3.4.4)). Equalization performance in the time domain is a measure of both amplitude and phase and given by

$$\varepsilon_t = D(m). \tag{3.6.2}$$

Equalization performance in the frequency domain is evaluated using the magnitude deviation (44) of the equalized IR and is used to measure incoherent equalization performance:

$$\varepsilon_f = \left[\frac{1}{I}\sum_{k=0}^{I-1}(10\log_{10}|\hat{D}(k)| - \bar{D})^2\right]^{-1/2}, \tag{3.6.3}$$

where

$$\bar{D} = \frac{1}{I}\sum_{k=0}^{I-1} 10\log_{10}|\hat{D}(k)|. \tag{3.6.4}$$

In Eq. (3.6.3) and Eq. (3.6.4), I corresponds to the length of the FFT with frequency bins k and Fourier coefficients $\hat{D}$ of the equalized signal D. Magnitude deviation is invariant to the length of the FFT and, for ideal equalization, equates to zero.

3.6.4 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 $\hat{S}_s$ is computed by adjusting the PSD of the output signal S_o with the ensemble average of the channel's amplitude response $|H|$:

$$10\log_{10}(\hat{S}_s) = M[10\log_{10}(S_o) - 20\log_{10}(\mathbb{E}[|H|])]. \tag{3.6.5}$$

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 (Eq. (3.4.4)) is computed incoherently

$$10\log_{10}(\hat{S}_s) = M[10\log_{10}(S_o) + 20\log_{10}(\mathbb{E}[|\hat{F}|]) - 2\bar{D}], \tag{3.6.6}$$

where $|\hat{F}|$ is the amplitude response of $\hat{f}$. The equation is further adjusted by a constant, equal to the spectral mean of the equalized IR ($\bar{D}$, given by Eq. (3.6.4)). Comparing results from Eq.

(3.6.6) to results from Eq. (3.6.5) 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):

$$\text{RMSE } = 10\log_{10}\left[\frac{1}{N}\sum_{k=1}^{N}||\hat{S}_k - S_k||^2\right]^{-1/2}, \tag{3.6.7}$$

where $\hat{S}$ 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.

3.7 Results

3.7.1 IR estimation

Acoustic IRs were calculated using Eq. (3.3.6) for linear and logarithmic excitation methods. Figure 3.3(a) shows an IR obtained using a logarithmic sweep. Figure 3.3(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 an moving average filter of 201 points. The range in Fig. 3.3(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 realizations) was computed on the log scale ($20log_{10}$). 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 Sec. 3.5 and Sec. 3.6.1), the theoretical sinusoidal pressure distribution (Eq. (3.3.4)) yields $\sigma = \pm 1.94$ dB. Magnitudes of the mixed sinusoids (Table 3.1) were averaged over a duration of 4 s and divided by the input amplitude. The result is plotted against the excitation methods in Fig. 3.3b. 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 3.1). Computed amplitude ratios are within 2σ of the exponential transfer function.

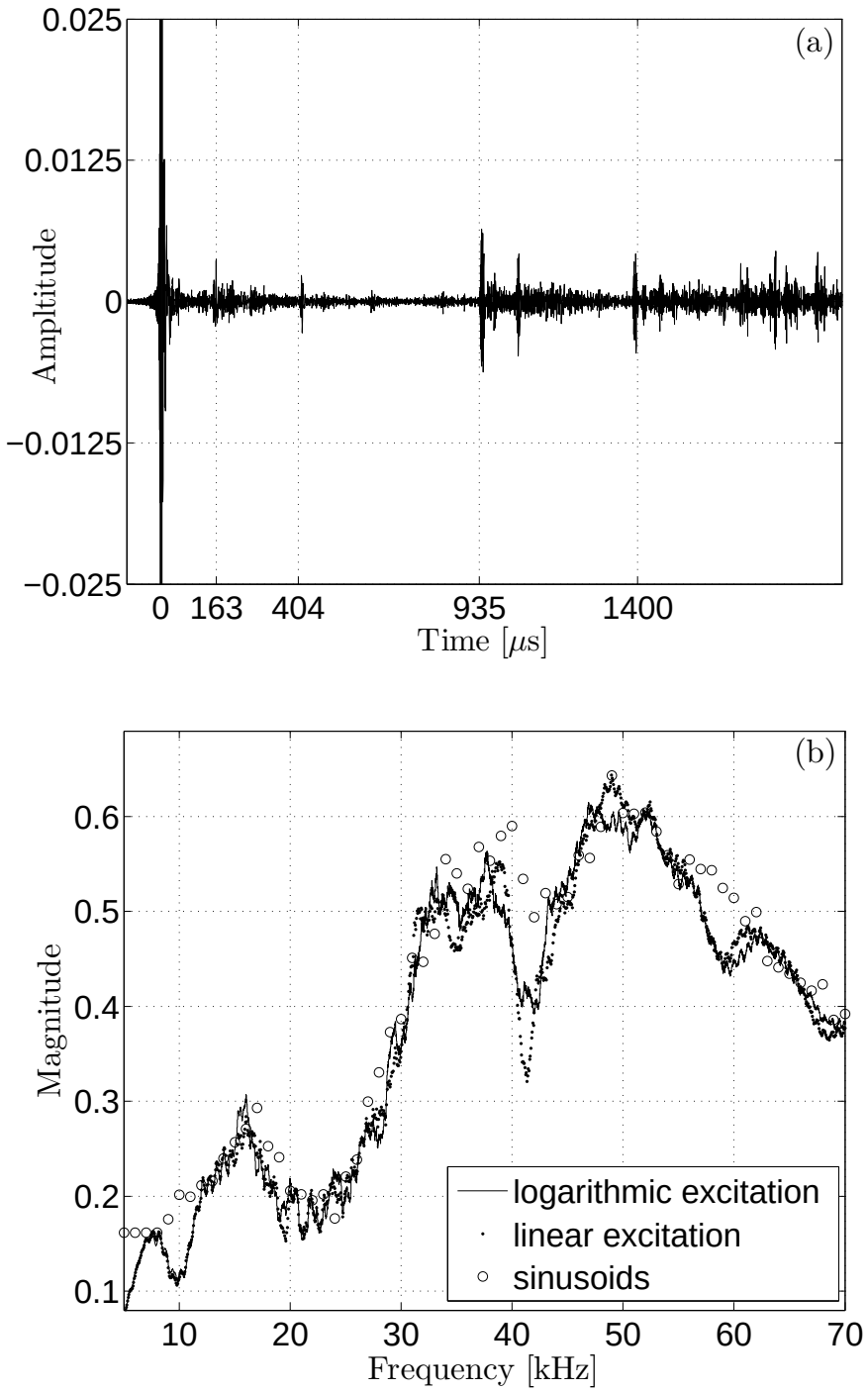

Figure 3.3. (a) Acoustic IR of the diving well h(t) with theoretical boundary reflection times. The scale is chosen to show details of the reflections, but cuts off the direct arrival which has a maximum amplitude between ± 0.06. Ticks on the x-axis correspond to theoretical arrival time due to the boundaries (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

3.7.2 T_{60} and T_{sn}

The theoretical T_{60} (Eq. (3.4.1)) for the pool is 282 ms using acoustic impedance values given in Sec. 3.6 and dimensions of the pool given in Sec. 3.5. The corresponding Schroeder frequency (Eq. (3.4.3)) is approximately 356 Hz, which is close to the resonant frequencies of the pool (< 300 Hz). Figure 3.4(a) shows calculated IR decay curves using Eq. (3.6.1). 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 150-250 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 Sec. 3.6.1, 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. (3.6.1) is too short (e.g. 125 ms in Fig. 3.4(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 175-225 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. Figure 3.4(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. 3.4(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. 3.4(b). The plotted noise reference line is on the same order as the noise between 175 and 225 ms.

3.7.3 Coherent inversion of IR

Figure 3.5 shows channel equalization performance versus processing delay of the least-squares filter using Eq. (3.6.2) and Eq. (3.6.3) for a randomly selected IR length of 152 ms (19k 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 $\varepsilon_t \approx$

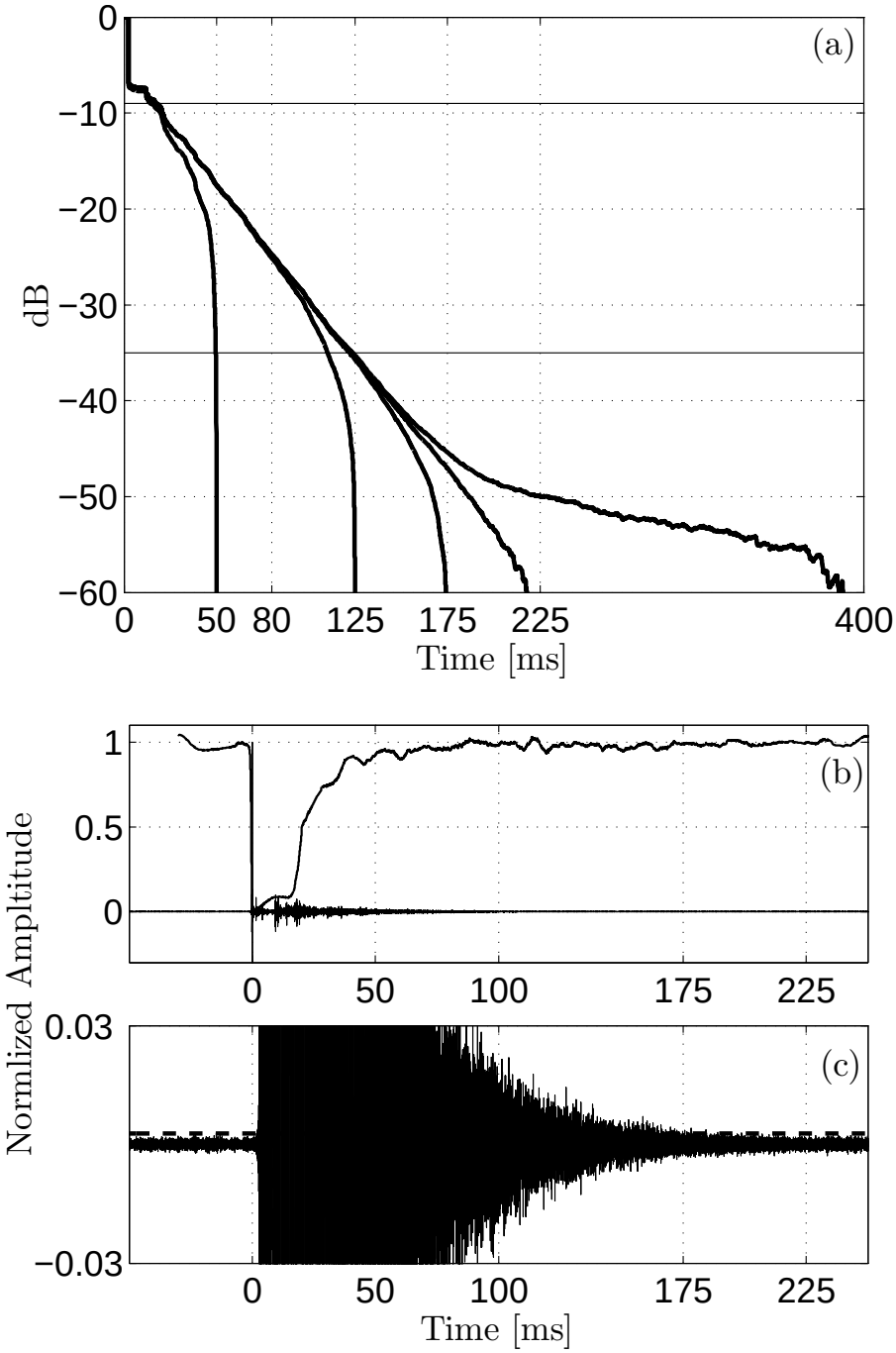

Figure 3.4. (a) Decay curves of the IR with subtracted noise average. The ticks on the x-axis show selected upper integration limits, the two horizontal lines (-9 dB and -35 dB) correspond to the range over which T_{60} is calculated. (b) IR with echo density (top trace) showing transition time from 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.

0.001. Results are similar for IR lengths ranging from 50-175 ms. In conclusion, the performance of the inverse filter improves significantly using a processing delay.

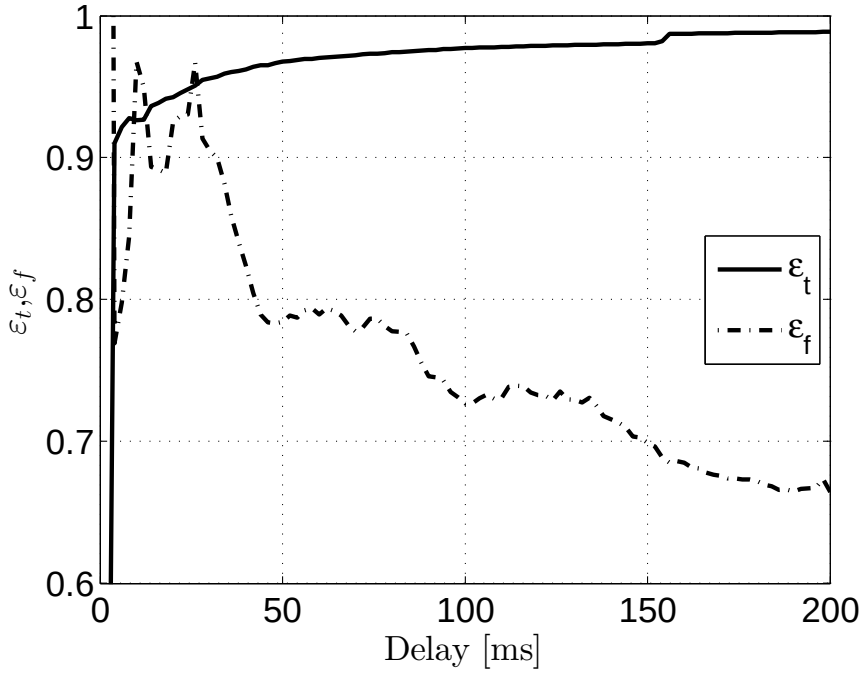

Figure 3.5. Coherent inversion performance vs. processing delay for IR of length 152 ms

3.7.4 Dereverberation results

We applied the dereverberation methods proposed here (Eq. (3.3.7)) 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 Sec. 3.6.4, 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 signal because of its smooth spectra. The logarithmic sweep was selected to compute the IR. 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, 10 realizations were selected for the ensemble average when investigating effects of smoothing and IR length.

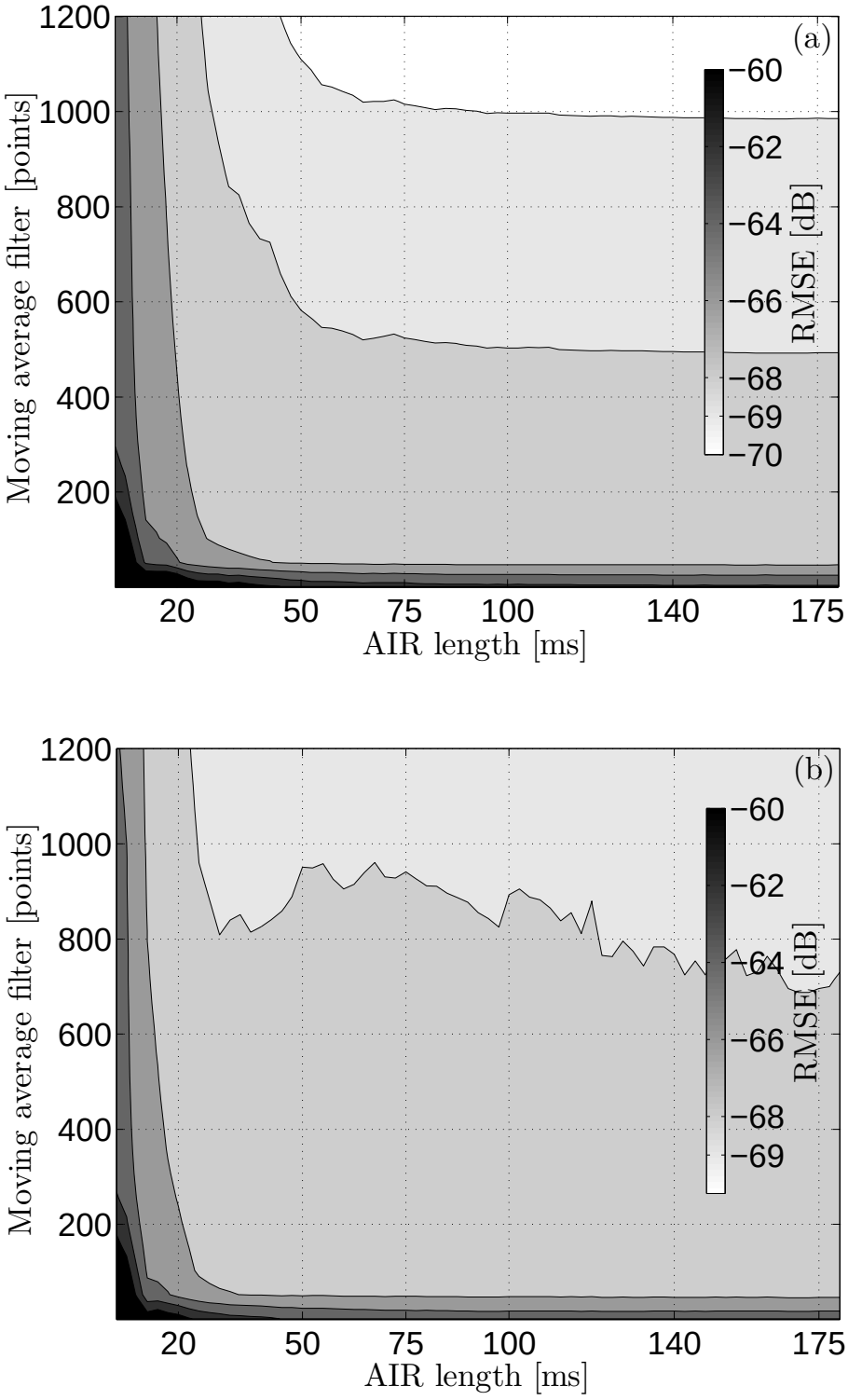

Figure 3.6. RMSE in dB of (a) incoherently and (b) coherently inverted IR of dereverberated linear sweep using Eq. (3.6.5) and (3.6.6), respectively. The expectation is computed using 10 realizations and RMSE ticks correspond to contour surfaces.

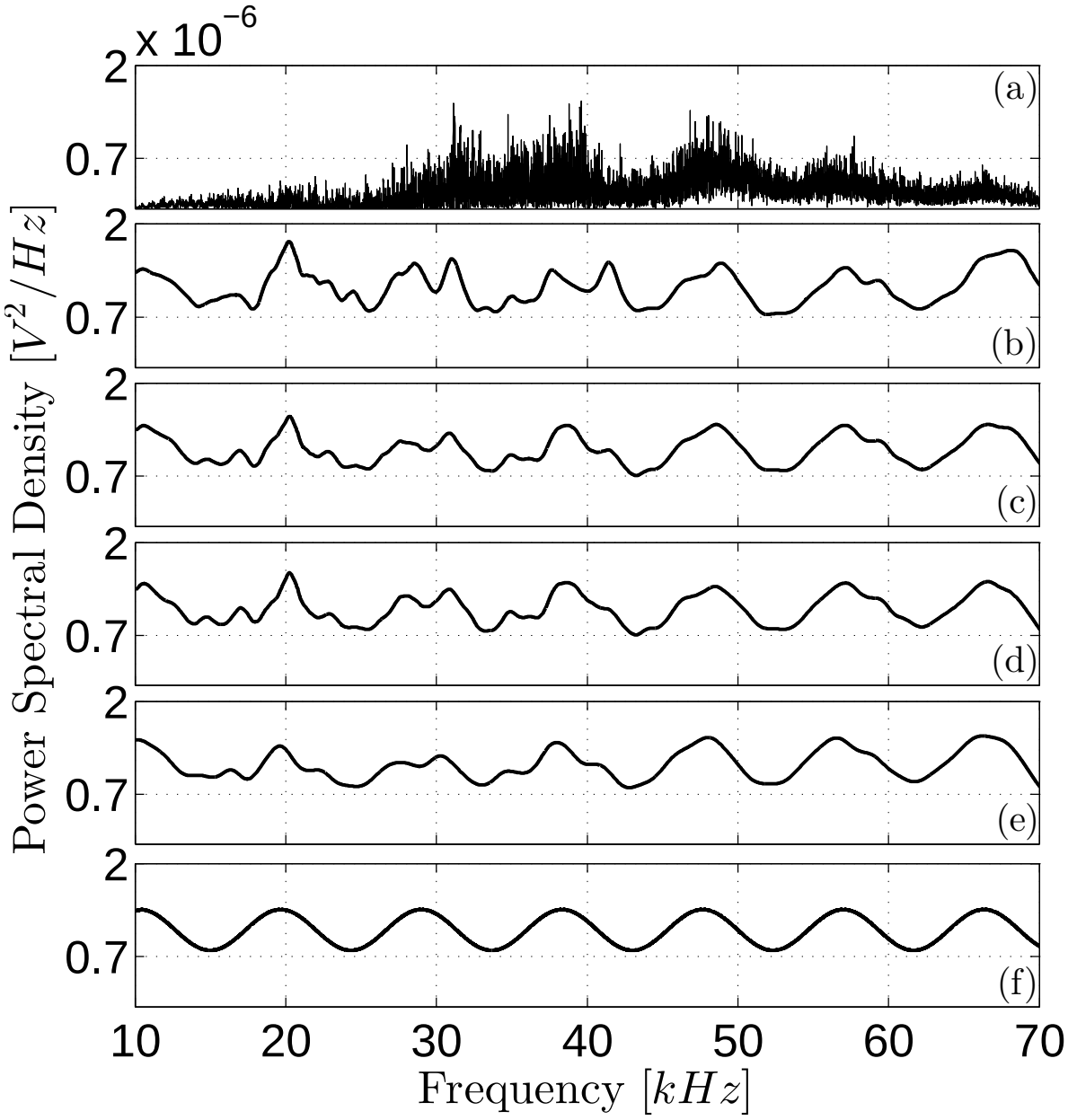

Figure 3.7. (a) Recorded linear sweep. Incoherently adjusted linear sweep with (b) 1, (c) 10 and (d) 50 realizations in the ensemble average. (e) Coherently adjusted linear chirp with 10 realizations in the ensemble average using a processing delay of 150 ms. All plots were computed using an IR length of 100 ms and moving average of 800 points. The original swept-frequency cosine chirp is shown in (f).

Incoherent inversion

Incoherent dereverberation results were computed using Eq. (3.6.5) and 10 realizations were used to approximate the channel's amplitude response $|H|$. Results are shown in Fig. 3.6(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.

Figure 3.7(a) shows the PSD of the recorded linear sweep which served as the source signal for Fig. 3.6. The recorded signal is adjusted with 1, 10 and 50 realization of the transfer function (Fig. 3.7(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 10-50 retaliations is minor. Similar analysis for other test signals (i.e. white noise using various bandwidths) yields similar results.

Coherent inversion

Dereverberation results using coherent inversion were computed using Eq. (3.6.6). First, the inverse of the IR is computed in the least-squares sense with a delay of 150 ms before averaging over 10 realization. Fig. 3.6(b) shows RMSE for coherent inversion. 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. 3.6(b) (about 800) to Fig. 3.6(a) (about 600) 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. 3.7(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.

3.8 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. (3.4.1)). 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. (3.6.1)). 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. Equation (3.3.4) can therefore 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 correlation 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). MLS further require a strict linearity assumption: transducer non-linearities cause distortion peaks (66) and care must be taken to ensure that the transmitting transducer is not excited into its non-linear region. Sweeps are invariant to transducer non-linearities because non-linearities deconvolve before the direct arrival and can be removed with a window (46). Generally speaking, logarithmic sweeps seem to be the more practical excitation alternative, especially from a system point of view. Experiments in room acoustics (46) confirmed that for environments with non-correlated noise, sweep performance was superior due to higher initial SNR. The MLS caused clipping in this experiment at a pre-amp setting of +3 dB and SNR needed to be reduced. In addition, when using sweeps, samples suffering from correlated noise (the main advantage of MLS) can easily be identified by plotting the spectrogram of the IR: energy is observable before and after the direct arrival over a particular frequency band (here, the ADC had some correlated self noise above 75 kHz). Bad frequency bands or samples can be excluded from the analysis or the noise source can be eliminated (e.g. door mechanisms for water overflow devices that cause impulsive noise can be fixed in place).

Equation (3.6.1) 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 the echo density is function of the sliding window length (here 2500 points) and its range can change slightly, 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. 3.4(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. 3.4(a)), which is very reasonable in practice.

This paper presented two equations in Sec. 3.6.4 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 the least-squares 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. 3.7 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. Results further indicate that incoherent dereverberation is 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. 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. (3.4.4), shown for a particular case in Fig. 3.5. The processing delay 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. 3.3(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 (48). 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 (67). 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. (3.4.4)). Most of the magnitude deviation ε_f corresponds to non-equalized frequencies primarily below 35 kHz. These frequencies likely correspond to partially equalized spectral zeros, resulting from sinusoidal oscillatory behavior of the transfer function over the linear decreasing amplitude response (5-35 kHz) of the transmitting transducer. This suggests that inversion performance can be improved using the optimum bandwidth of a transmitting transducer: reducing the dynamic range of the combined IRs (Eq. (3.3.2)) will improve equalization performance (45). It should be noted that coherent equalization of IRs yields poor performance at offsets of fractions of a wavelength (42), 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. 3.6(b) shows coherent RMSE which is slightly increased in comparison to the incoherent case in Fig. 3.6(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 coherent case using an IR length of 150 ms and a moving average filter with 1400 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 channel's expectation in Eq. (3.3.8) 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. (3.6.6) 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. (3.3.6).

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. 3.7. 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 minutes. Results in Fig. 3.6 will further improve by including the expectation in Eq. (3.3.8) and

Eq. (3.3.9) 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.

We demonstrated that it is possible to recover a control signal in the forward problem. 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. 3.3 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.

3.9 Acknowledgments

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.

Chapter 4

Estimating and removing colorations from the deconvolved impulse response of an underwater acoustic channel

4.1 Abstract

The impulse response (IR) of an acoustic channel can be obtained using a known excitation signal. However, the deconvolved IR is colored by transducers and other electrical equipment. This letter presents a method to separate all coloration from the channel's IR using a pseudo quadrature mirror filter bank in the time domain. The method is validated using synthetic results of an image-source model and the channel's IR is recovered over the full band with a root-mean-square error of -31 dB (spanning -14 dB to -107 dB). The estimated IR of a reverberant channel, recorded in a pool environment, is presented.

4.2 Motivation

RECORDINGS of acoustic sources conducted with electrical equipment such as transducers and analog to digital converters do not reflect the true levels of the source. Actual levels are amplified or reduced due to the equipment's (non-uniform) frequency response: this effect is referred to as coloration. Colorations are generally undesirable and often impossible to remove from recordings when using non-calibrated equipment. Manufacturers using pulse-gating techniques to calibrate transducers usually supply amplitude responses only, rather than a time waveform of a

transducer's impulse response (IR). In addition, impedance mismatches between connected equipment can significantly alter the combined response of individually calibrated components. For many applications, however, it is desirable to obtain coherent or at least incoherent source level estimates, and both require knowledge of the equipment's combined impulse response $u(t)$.

This paper presents a method to estimate the combined frequency response of the electrical equipment if the system's transient response can be obtained. In Eq. (4.2.1), the recording $r(t)$ of an acoustic source $s(t)$ in a reverberant underwater channel $g(t)$ is colored by $u(t)$. The transient response $h(t)$ can be recovered via deconvolution if $s(t)$ is known (49). In this letter, we hypothesise a method to recover $g(t)$, which is convolved with the unknown combined impulse responses of all electrical systems $u(t)$ to form $h(t)$, Eq. (4.2.2).

$$r(t) = s(t) * g(t) * u(t) \tag{4.2.1}$$

$$h(t) = g(t) * u(t) \tag{4.2.2}$$

4.3 Approach

In order to recover an estimate of $g(t)$, the following assumptions and observations are made. We require that the system is linear and that the first (direct) arrival and the second (reflected) arrival in $g(t)$ are sufficiently separated in time (discussed further below). Furthermore, it is assumed that the direct arrival of $g(t)$ corresponds to a scaled delta function: $k\delta(t-t_d)$, where k is a constant and t_d is the direct arrival time. To illustrate this requirement, suppose an experiment is conducted in a fully anechoic environment without frequency dependent attenuation or spreading losses. In this ideal example, $g(t) = k\delta(t-t_d)$ and $h(t) = k\delta(t-t_d) * u(t) = ku(t-t_d)$. Replacing the anechoic with a reverberant environment, the direct arrival still correspond to a scaled delta function but the IR also includes an additional reverberant part. A spreading loss model can be used to estimate the (frequency independent) constant. In other words, we think of the direct arrival as an all-pass filter with an unknown, pure delay t_d (corresponding to the channel length), which does not alter the recorded signal's phase or magnitude. Without loss of generality, we set $k = 1$ and $t_d = 0$ in what follows.

The delta function is key to recovering $g(t)$ because it has unit area, which corresponds to a flat spectrum with unit magnitude response. When the channel's transfer function (TF) $G(\omega)$ is multiplied by the unknown equipment's TF $U(\omega)$, the resulting spectrum $H(\omega)$ corresponds to the unknown TF when the fast Fourier transform (FFT) is computed over the duration of the direct

arrival. However, spectral division ($H(\omega)/U(\omega)$) might result in narrow-band noise amplification if $U(\omega)$ is poorly conditioned. Furthermore, $U^{-1}(\omega)$ is acausal if $U(\omega)$ displays non-minimum phase behavior. To circumvent problems associated with stability and acausality, the following discussion focuses on the area of the direct arrival in the time domain.

To recover $g(t)$, we filter $h(t)$ into $i = \{1, 2, ..., n\}$ band-limited versions using a pseudo quadrature mirror filter bank (PQMFB, 68) with near-perfect reconstruction properties (the reconstruction error depends on the filter order and can reach -100 dB). All filters in the bank (denoted by $f(t)$) are real, linear, of equal bandwidth and, ideally, their reconstruction error vanishes except for a pure delay (which can be ignored for zero-phase forward and reverse filtering): $f(t) = \sum_i^n f_i(t) \approx \delta(t)$. The selected bandwidth and number of filters correspond to the maximum slope of $|U(\omega)|$ such that its magnitude spectrum is approximately constant in each band. Forward and reverse filtering requires that the direct arrival is separated from the first reflection by twice the filter's IR length.

Convolving $f(t)$ with Eq. (4.2.2) yields Eq. (4.3.1):

$$\begin{aligned} f(t) * h(t) &= f(t) * g(t) * u(t) \\ &= g(t) * \sum_{i=1}^{n} f_i(t) * u(t). \end{aligned} \tag{4.3.1}$$

Next, we compute the absolute area in band i over the duration of the direct arrival, denoted by limits t_0 and t_e:

$$\int_{t_0}^{t_e} |f_i(t) * h(t)|dt = \int_{t_0}^{t_e} |g(t) * f_i(t) * u(t)|dt = \int_{t_0}^{t_e} |f_i(t) * u(t)|dt = b_i. \tag{4.3.2}$$

Within these limits, $g(t)$ corresponds to a delta function. When a delta function is convolved with a sinusoid, the resulting area corresponds to the amplitude of the sinusoid. Here, the absolute area (pressure and filter coefficients can be negative) approximates $|U(\omega)|$ in band i (similar to a step function), and coefficients are denoted by b_i. The overlap between adjacent bands contributes to errors in b_i. It should be noted that the bandwidth for each filter is not narrow ($>$ 1 kHz) since equipment manufacturers design their systems (among other constrains) to minimize variations in $|U(\omega)|$.

Once all coefficients are estimated, all n bands are divided by their respective scale factors and summed up to recover the signal over the full band (an FFT computed over the direct arrival yields a flat spectrum with unit magnitude response). The resulting signal has the desired magnitude spectrum but phase contributions from $u(t)$ remain. To estimate and remove phase due

to $u(t)$, computing an FFT over the direct arrival yields coefficients used to design an all-pass filter with inverse phase response. Convolving the filter with the scaled signal recovers an estimate of $g(t)$ (denoted by $\hat{g}(t)$).

4.4 Validation

We validate our method using an image-source model (69) to generate a channel's IR $g(t)$ (Fig. 4.1(a)) which is subsequently convolved with an unknown IR $u(t)$. The magnitude response of $u(t)$ is shown in Fig. 4.1(b) and approximates a piezoelectric transducer (optimum frequency response $\sim$20-50 kHz) amplified by other equipment for a total dynamic range close to 20 dB. In practice, $\hat{g}(t)$ is recovered (Fig. 4.1(d)) using a cascaded process of several sets of filter banks with different transition frequencies. Here, we used four sets of filter banks from most broad-band to narrow-band: 10 filters (order 60), 25 filters (order 120), 30 filters (order 180) and 60 filters (order 240). This method is used because the error (Fig. 4.1(e)) is largest between two adjacent bands (at the transition region of each step-function). Designing an additional set of filters with center frequencies at the transition frequencies of the previous set reduces the error further and improves the estimate. For extrema (i.e., 35 and 54 kHz in Fig. 4.1(b)) the error is increased but can be controlled with a well-designed cascaded filtering process. The error will probably always be significant at band-edges (0 kHz and 70 kHz). Unwrapped phase responses of individual IRs are shown in Fig. 4.1(c) and phase error in Fig. 4.1(f). Figure 4.2 shows details of the recovery process: Fig. 4.2(a) shows a typical plot when $f(t)$ is convolved with $h(t)$ including integration limits, and Fig. 4.2(b-c) show the magnitude spectra of the direct arrival in $h(t)$ and $\hat{g}(t)$, respectively.

4.5 Application and Example: University of Hawai‘i at Mānoa diving well

An experiment (similar to 4) was conducted in a reverberant pool environment (22.9 m by 22.9 m with a depth of 5.2 m) using an exponential sweep (5-85 kHz) as the source signal with transducer separation of 1 m to estimate $\hat{g}(t)$. Instruments where rigidly mounted nominally in the center of the diving well. Channel geometry is the same as in the image-source model (so we can use the same filter length) and the unknown TF is estimated to be smoother and to have less

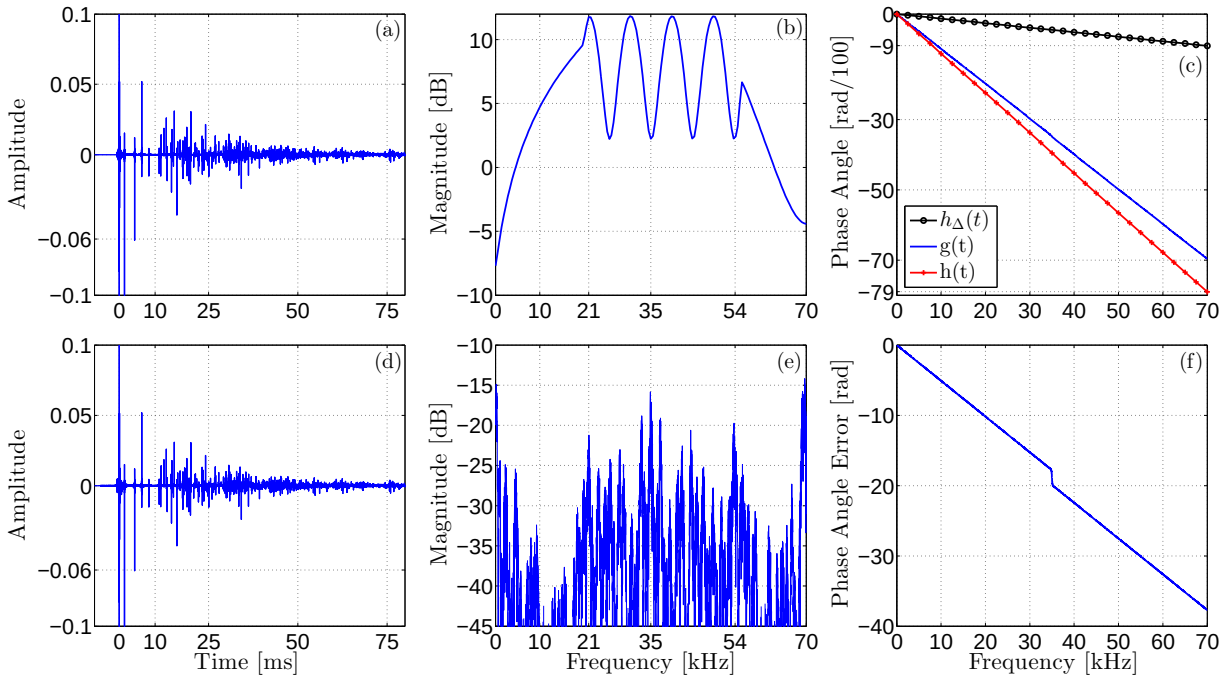

Figure 4.1. (a) Synthetic IR g(t) obtained form the image source model sampled at 140 kHz and (b) $|U(\omega)|$ of unknown transfer function. (c) Phase responses of $g(t)$, $h(t)$, and the all-pass filter using the direct arrival in $h(t)$ (denoted by $h_\Delta(t)$). (d) Recovered IR $\hat{g}(t)$ and (e) error of recovered IR using cascaded PQMFBs. The error between (a) and (d) is computed on the spectrum using $20log_{10}(||G(\omega)| - |\hat{G}(\omega)||)$ (root-mean-square error -31 dB, max. -14 dB, min. -107 dB). (f) Phase angle error ($\angle G(\omega) - \angle\hat{G}(\omega)$) of recovered signal.

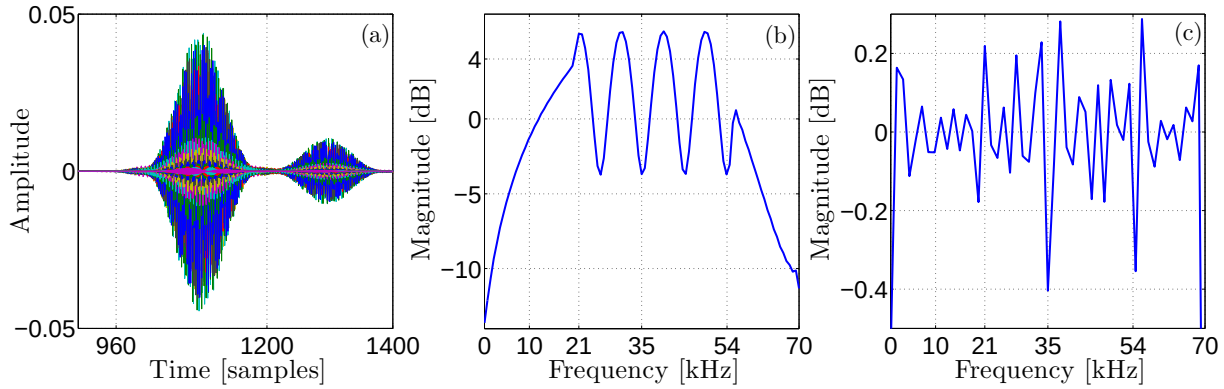

Figure 4.2. (a) $f(t) * h(t)$ (n=30, each color corresponds to a different band) with integration limits $t_0 = 960$ samples and $t_e = 1200$ samples. The direct arrival of h(t) is located at $\sim$1100 samples, the first reflection at $\sim$1300 samples. Note that the IR of each filter decays to zero well before its length of 180 samples (so "sufficient separation" between direct arrival and first reflection can be less than twice the highest filter order). (b) $|H(\omega)|$ measured over integration limits. Note the similarity to Fig. 4.1(b) but with a different range. (c) $|\hat{G}(\omega)|$ measured over integration limits. Ideally, the response should be 0 dB. Note that the deviations at i.e. 35 and 54 kHz and band edges correspond to errors in Fig. 4.1(e).

dynamic range (due to an unknown impedance mismatch) than Fig. 4.1(b). Signal to noise ratio (SNR) was ∼45 dB. The recovered IR $\hat{g}(t)$ is presented in Fig. 4.3(a) and an ensemble average of its magnitude spectrum in Fig. 4.3(b). The ensemble average is required to approximate the pressure distribution (57) in a reverberant environment. The spectrum has been plotted within the optimum frequency response (35-68 kHz) of the CR1 Sensor Technology Limited transducer (Seattle, WA, SN: 09178-01) using a sampling frequency of 140 kHz. Figure 4.3(b) indicates that reverberation accounts for ≤ 0.55 dB in the 35-68 kHz band and is larger at lower frequencies.

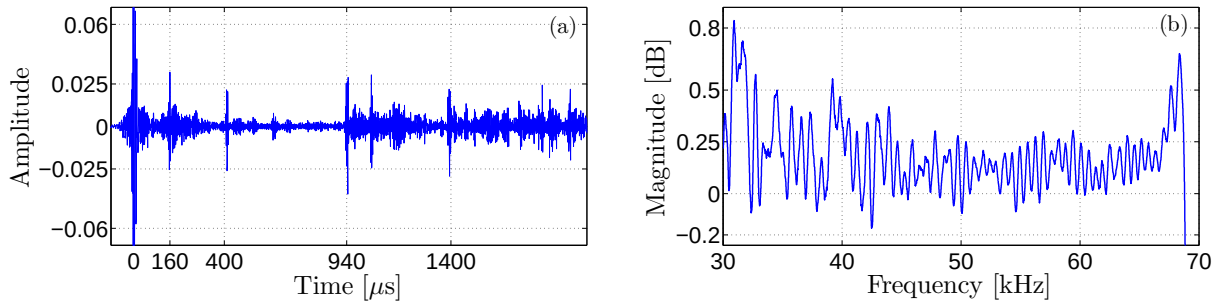

Figure 4.3. (a) IR estimate of recording channel with x-ticks corresponding to theoretical boundary reflection times. The scale is chosen to show details of the reflections, but cuts of the direct arrival, which has a maximum amplitude between ± 0.2. Note that this method modifies the noise-profile around the direct arrival, which is also evident in Fig. 4.1(d). The spectrum of the primary arrival has ∼ unit magnitude response, similar to Fig. 4.2(c). The signal is filtered over the 28-68 kHz band using a Kaiser bandpass filter. (b) Magnitude spectrum $M[20\log_{10}(\mathbb{E}[|\hat{G}|])]$ computed with 10 realizations (1 Hz resolution, 400 point zero-phase moving average filter $M[\cdot]$, ensemble average is denoted by $\mathbb{E}[\cdot]$). The positive slope and steep roll-off in the 65-68 kHz band is caused by the bandpass filter's IR.

4.6 Conclusions

We have shown that it is possible to (a) separate $g(t)$ and $u(t)$, (b) estimate the shape of the unknown TF, and (c) recover an estimate of a channel's IR under favorable SNR conditions. Future work will be conducted to compare the performance of the presented method to a deconvolution approach (e.g., compute $U^{-1}(\omega)$ explicitly). Since the performance of this method depends on the slope of the unknown TF (in particular transducers), it should be tested using a variety of different equipment. This method can further improve with regular filtering instead of the zero-phase filter implementation: the reconstruction error decreases and filter length can be increased by a factor of two (the trade-off is that integration limits will change by a constant during the cas-

caded approach). The error can be further reduced by maximizing filter lengths and optimizing the number of filters in the cascaded approach as well as by using non-uniform bandwidths.

Technically, no information about the unknown TF is required. However, it is recommended to pre-color the excitation signal used to deconvolve the combined IR (Eq. (4.2.2)) by the inverse amplitude response of equipment with significant non-uniform amplitude responses (such as transducers). This will (a) increase SNR of h(t), (b) decrease the dynamic range and smooth $u(t)$, and subsequently improve the estimate of $\hat{g}(t)$. We neglected to discuss noise: uncorrelated noise will not contribute to the IR if the excitation signal has zero mean. Correlated, stationary noise (assuming sufficient SNR at the hydrophone and not excessively long excitation signals) will be part of $u(t)$ and removed. The method presented in this letter can be applied to recover $\hat{g}(t)$ and estimate $u(t)$ recorded in a pool or in an open ocean environment. Matlab code will be made available upon request.

Chapter 5

Shortcomings and suggested future research

The proposed source characterization method requires that dimensions and directionality are similar for both the unknown source and the transmitting transducer. This limits the method in practice since directionality of most sources is unknown and dimensions are likely to be different. The directionality for different SCUBA configurations is likely similar: regulator dimensions are almost identical and the internal function at each stage is the same (i.e., to achieve the pressure drop). This is not necessarily true for closed systems (rebreathers): commercial system are only a few years old and not as rigorously tested as SCUBA equipment. The complex design can vary significantly from one manufacture to another and might contribute to a different directionality. Furthermore, if the source spectrum is broadband, the IR needs to be measured with different transducers, which likely have different directionalities. For example, a low frequency Lubell speaker is highly directional and, generally speaking, high-frequency transducers are almost omnidirectional. Current research can be extended by introducing a method to equalize directionality differences between transducers such that all equipment is omnidirectional. Shadow regions created by the unknown source might introduce an additional bias and should be addressed.

Further research can focus on the analytic pressure distribution for broadband sources in reverberant environments. Small changes in the recording channel may correspond to significant changes in the frequency dependent pressure distribution. For a given environment, channel equalization might be accomplished by measuring the IR at different locations corresponding to the dimensions of the unknown source. A channel can subsequently be selected within a field of minimum variation and data can be used to validate an analytic model. The model has to include the unknown source since its presence might also alter the pressure field.

Furthermore, the analytic equation will yield a frequency dependent estimate of the variance of a recording channel. The distributions will change from the near- to the far-field, which

should be included in the analysis. The outcomes would validate position of expectation operators and logarithms in incoherent formulations, similar to:

$$SSL = M_1[10\log_{10}(\mathbb{E}[S_o])] - M_2[20\log_{10}(\mathbb{E}[|\hat{G}|])] - |R|,$$

where S_o is the PSD of the recorded signal, $|\hat{G}|$ and $|R|$ are the amplitude responses of the channel's TF and recording equipment, and M_1 and M_2 are moving averages of different lengths. It is not clear at this point where to compute the expectation when estimating the PSD. Perhaps it is reasonable to compute the assembly average incoherently before the signal is squared.

The proposed method is for a single hydrophone and can evolve into a multi-hydrophone least-squares method. The corresponding impulse response has the potential to be inverted exactly, i.e., if there are no common spectral zeros between individual channels. Perhaps it is possible to average such a combined impulse response not just for a particular source point but over a volume corresponding to the unknown source.

The method presented in chapter 4 requires experimental verification (i.e., comparison of a signal recorded in an anechoic and reverberant environment). Reverberation likely adds significant energies at lower frequencies, which needs to be investigated. Furthermore, the method can still be optimized. For known sources, the method can be used to estimate spreading energy and derive a relationship between spherical and cylindrical spreading as a function of distance in reverberant environments. In other words, for a given environment, the scale constant for the delta function can be measured (see Sec. 4.3). This research would yield the evolution from purely spherical to purely cylindrical spreading as a function of source-receiver separation and boundary distance. In addition, the method can be used to extract coloration information. This can easily be validated with a calibrated system.

Bibliography

(1) N. Cochard, J. Lacoume, P. Arzelies, and Y. Gabillet, "Underwater acoustic noise measurement in test tanks," *IEEE Journal of Oceanic Engineering*, vol. 25, pp. 516–522, Oct. 2000.

(2) R. A. Hazelwood and S. P. Robinson, "Underwater acoustic power measurements in reverberant fields," in *OCEANS 2007 - Europe*, pp. 1–6, IEEE, June 2007.

(3) K. H. Kuttruff, *Room Acoustics*. London: Taylor & Francis, 4th ed., 2000.

(4) K. L. Gemba, E.-M. Nosal, and T. R. Reed, "Partial dereverberation used to characterize open circuit scuba diver signatures," *J. Acoust. Soc. Am.*, vol. 136, pp. 623–633, Aug. 2014.

(5) K. L. Gemba and E.-M. Nosal, "Source characterization using recordings made in a reverberant underwater channel," *Manuscript submitted for publication in J. Acoust. Soc. Am.(copy on file with author)*.

(6) K. L. Gemba, E.-M. Nosal, and T. R. Reed, "Estimating and removing colorations from the deconvolved impulse response of an underwater acoustic channel," *Manuscript submitted for publication in IEEE Signal Processing Letter (copy on file with author)*.

(7) P. Kaluza, A. Kölzsch, M. T. Gastner, and B. Blasius, "The complex network of global cargo ship movements.," *J. R. Soc.*, vol. 7, pp. 1093–103, July 2010.

(8) M. Smookler, B. Clark, and J. Ostrander, "Underwater Detection and Surveillance Technology for Commercial Port and Vessel Security. Who is Going to Pay for It," in *Proceedings of OCEANS 2005 MTS/IEEE*, (Washington, D.C), pp. 935 – 940, IEEE, 2005.

(9) M. Bruno, K. W. Chung, H. Salloum, A. Sedunov, N. Sedunov, A. Sutin, H. Graber, and P. Mallas, "Concurrent use of satellite imaging and passive acoustics for maritime domain awareness," in *2010 International WaterSide Security Conference*, (Marina di Carrara, Italy), pp. 1–8, IEEE, Nov. 2010.

(10) K. Shaw, R. Scott, and G. Holdanowicz, "Sonar sentinels on guard for submerged swimmers," *Janes Navy International*, vol. 110, no. 8, pp. 10–18, 2005.

(11) T. Folegot, G. Martinelli, P. Guerrini, and J. M. Stevenson, "An active acoustic tripwire for simultaneous detection and localization of multiple underwater intruders.," *J. Acoust. Soc. Am.*, vol. 124, pp. 2852–2860, Nov. 2008.

(12) D. Suchman and T. Meurling, "Integrated underwater intruder detection system," in *2010 International WaterSide Security Conference*, (Marina di Carrara, Italy), pp. 1–8, IEEE, Nov. 2010.

(13) J. Gebbie, M. Siderius, and J. S. Allen, "Passive Acoustic Array Harbor Security Applications," *Marine Technology Society Journal*, vol. 45, pp. 103–110, May 2011.

(14) B. Borowski, A. Sutin, H.-S. Roh, and B. Bunin, "Passive acoustic threat detection in estuarine environments," in *Proc. SPIE 6945, Optics and Photonics in Global Homeland Security IV* (C. S. Halvorson, D. Lehrfeld, and T. T. Saito, eds.), vol. 6945, (San Diego), pp. 694513–694513–11, Apr. 2008.

(15) D. P. Nowacek, L. H. Thorne, D. W. Johnston, and P. L. Tyack, "Responses of cetaceans to anthropogenic noise," *Mammal Review*, vol. 37, pp. 81–115, Apr. 2007.

(16) J. A. Hildebrand, "Anthropogenic and natural sources of ambient noise in the ocean," *Marine Ecology Progress Series*, vol. 395, pp. 5–20, Dec. 2009.

(17) M. W. Legg, A. J. Duncan, A. Zaknich, and M. V. Greening, "Analysis of impulsive biological noise due to snapping shrimp as a point process in time," in *OCEANS 2007 - Europe*, (Aberdeen, Scotland), pp. 1–6, IEEE, June 2007.

(18) W. Carey, "Oceanic low frequency ambient noise," in *OCEANS 2000 MTS/IEEE Conference and Exhibition*, vol. 1, (Providence, RI), pp. 453–458, IEEE, 2000.

(19) G. M. Wenz, "Acoustic Ambient Noise in the Ocean: Spectra and Sources," *J. Acoust. Soc. Am.*, vol. 34, pp. 1936–1956, Dec. 1962.

(20) B. S. Bingham, E. F. Prechtl, and R. A. Wilson, "Design Requirements for Autonomous Multivehicle Surface-Underwater Operations," *Marine Technology Society Journal*, vol. 43, pp. 61–72, May 2009.

(21) C. A. Radford, A. G. Jeffs, C. T. Tindle, R. G. Cole, and J. C. Montgomery, "Bubbled waters: The noise generated by underwater breathing apparatus," *Marine and Freshwater Behaviour and Physiology*, vol. 38, pp. 259–267, Dec. 2005.

(22) D. M. Donskoy, N. A. Sedunov, A. N. Sedunov, and M. A. Tsionskiy, "Variability of SCUBA diver's acoustic emission," in *Proceedings of SPIE* (C. S. Halvorson, D. Lehrfeld, and T. T. Saito, eds.), vol. 6945, (Orlando, Florida), pp. 694515–694515–11, Apr. 2008.

(23) X. Chen and U. Tureli, "Passive Acoustic Detection of Divers Using Single Hydrophone," in *2006 Fortieth Asilomar Conference on Signals, Systems and Computers*, (Pacific Grove, CA), pp. 554–558, IEEE, 2006.

(24) R. Stolkin, A. Sutin, S. Radhakrishnan, M. S. Bruno, B. Fullerton, A. Ekimov, and M. Raftery, "Feature based passive acoustic detection of underwater threats," in *Proc. SPIE 6204, Photonics for Port and Harbor Security II*, vol. 620408, (Orlando, Florida), pp. 620408–620408–10, 2006.

(25) R. Stolkin and I. Florescu, "Probabilistic analysis of a passive acoustic diver detection system for optimal sensor placement and extensions to localization and tracking," in *OCEANS 2007*, (Vancouver, BC), pp. 1–6, IEEE, 2007.

(26) B. Bunin, A. Sutin, and M. S. Bruno, "Maritime security laboratory for maritime security research," in *Proceedings of SPIE*, vol. 6540, (San Diego), pp. 65400S–65400S–8, 2007.

(27) A. Sutin, B. Bunin, A. Sedunov, N. Sedunov, L. Fillinger, M. Tsionskiy, and M. S. Bruno, "Stevens Passive Acoustic System for underwater surveillance," in *2010 International Water-Side Security Conference*, (Marina di Carrara, Italy), pp. 1–6, IEEE, Nov. 2010.

(28) R. K. Lennartsson, E. Dalberg, L. Persson, and S. Petrovic, "Passive Acoustic Detection and Classification of Divers in Harbor Environments," in *OCEANS 2009, MTS/IEEE Biloxi - Marine Technology for Our Future: Global and Local Challenges*, (Biloxi, MS), pp. 1–7, IEEE, 2009.

(29) A. T. Johansson, R. K. Lennartsson, E. Nolander, and S. Petrovic, "Improved passive acoustic detection of divers in harbor environments using pre-whitening," in *OCEANS 2010 MTS/IEEE SEATTLE*, (Seattle, WA), pp. 1–6, IEEE, Sept. 2010.

(30) R. K. Lennartsson, E. Dalberg, A. T. Johansson, L. Persson, S. Petrovic, and E. Rabe, "Fused passive acoustic and electric detection of divers," in *2010 International WaterSide Security Conference*, (Marina di Carrara, Italy), pp. 1–8, IEEE, Nov. 2010.

(31) A. Sutin and Y. Sinelnikov, "Time Reversal Acoustic approach for non-lethal swimmer deterrent," in *2010 International WaterSide Security Conference*, (Marina di Carrara, Italy), pp. 1–5, IEEE, Nov. 2010.

(32) J. Heine, J. Bookspan, and P. Oliver, *NAUI Master Scuba Diver*. Riverview: Naui Association of Underwater Divers, 2nd ed., 2004.

(33) P. Riegel, "Breathing Resistance in Scuba Regulators," in *OCEANS 1976*, (Washington, DC), pp. 556–561, IEEE, 1976.

(34) T. Fedenczuk and E.-M. Nosal, "Hawaii Experimental Acoustics Range (HEAR) for shallow water applications," *Marine Technology*, vol. 45, no. 3, pp. 69–76, 2011.

(35) R. J. Urick, *Principles of Underwater Sound for Engineers*. McGraw-Hill, 1967.

(36) R. E. Francois, "Sound absorption based on ocean measurements: Part I: Pure water and magnesium sulfate contributions," *J. Acoust. Soc. Am.*, vol. 72, no. 3, pp. 896–907, 1982.

(37) J. F. Kaiser, "Nonrecursive Digital Filter Design Using the Io-sinh Window Function," in *IEEE Int. Symp. Circuits and Systems (ISCAS74)*, (San Francisco), pp. 20–23, 1974.

(38) D. Havelock, S. Kuwano, and M. Vorländer, eds., *Handbook of Signal Processing in Acoustics*. New York: Springer, 2009.

(39) S. T. Neely and J. B. Allen, "Invertibility of a room impulse response," *J. Acoust. Soc. Am.*, vol. 66, no. 1, pp. 165–169, 1979.

(40) J. Mourjopoulos, "Digital equalization of room acoustics," *J. Audio Eng. Soc.*, vol. 42, no. 11, pp. 884–900, 1994.

(41) P. Clarkson, J. Mourjopoulos, and J. Hammond, "Spectral, phase, and transient equalization for audio systems," *J. Audio Eng. Soc.*, vol. 33, no. 3, pp. 127–132, 1985.

(42) B. Radlovic, R. Williamson, and R. Kennedy, "Equalization in an acoustic reverberant environment: robustness results," *IEEE Transactions on Speech and Audio Processing*, vol. 8, pp. 311–319, May 2000.

(43) M. Miyoshi and Y. Kaneda, "Inverse filtering of room acoustics," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 36, no. 2, pp. 145–152, 1988.

(44) J. Mourjopoulos, P. Clarkson, and J. Hammond, "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), pp. 1858–1861, IEEE, 2003.

(45) P. Naylor and D. G. Nikolay, eds., *Speech Dereverberation*. London: Springer, 2010.

(46) S. Müller and P. Massarani, "Transfer-Function Measurement with Sweeps," *J. Audio Eng. Soc.*, vol. 49, pp. 443–471, June 2001.

(47) M. R. Schroeder, "New Method of Measuring Reverberation Time," *J. Acoust. Soc. Am.*, vol. 37, no. 6, pp. 409–412, 1965.

(48) E. A. Robinson and S. Trietel, *Geophysical Signal Analysis*. Englewood Cliffs: Prentice-Hall, 1980.

(49) R. Polge and E. Mitchell, "Impulse Response Determination by Cross Correlation," *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-6, pp. 91–97, Jan. 1970.

(50) A. Farina, "Simultaneous measurement of impulse response and distortion with a swept-sine technique," in *Audio Engineering Society Convention*, vol. 108, pp. 1–24, 2000.

(51) S. Subramaniam, A. Petropulu, and C. Wendt, "Cepstrum-based deconvolution for speech dereverberation," *IEEE Transactions on Speech and Audio Processing*, vol. 4, no. 5, pp. 392–396, 1996.

(52) B. Radlovic and R. Kennedy, "Nonminimum-phase equalization and its subjective importance in room acoustics," *IEEE Transactions on Speech and Audio Processing*, vol. 8, no. 6, pp. 728–737, 2000.

(53) G. Xu, H. Liu, L. Tong, and T. Kailath, "A least-squares approach to blind channel identification," *IEEE Transactions on Signal Processing*, vol. 43, no. 12, pp. 2982–2993, 1995.

(54) K. Kumar, B. Raj, R. Singh, and R. M. Stern, "An iterative least-squares technique for dereverberation," in *2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, vol. 2, pp. 5488–5491, IEEE, May 2011.

(55) A. Oppenheim and R. Schafer, "Dsp history - From frequency to quefrency: a history of the cepstrum," *IEEE Signal Processing Magazine*, vol. 21, pp. 95–106, Sept. 2004.

(56) H. G. Diestel, "Probability Distribution Pressure of Sinusoidal Sound in a Room," *J. Acoust. Soc. Am.*, vol. 35, no. December, pp. 2019–2022, 1963.

(57) R. V. Waterhouse, "Statistical Properties of Reverberant Sound Fields," *J. Acoust. Soc. Am.*, vol. 43, no. 6, pp. 1436–1444, 1968.

(58) D. Lubman, "Precision of reverberant sound power measurements," *J. Acoust. Soc. Am.*, vol. 56, no. 2, pp. 523–533, 1974.

(59) A. Farina, "Advancements in impulse response measurements by sine sweeps," in *Audio Engineering Society Convention 122*, vol. 8, (Vienna, Austria), p. 21, Oct. 2007.

(60) R. Neubauer and B. Kostek, "Prediction of the reverberation time in rectangular rooms with non-uniformly distributed sound absorption," *Archives of Acoustics*, vol. 26, pp. 183–201, 2001.

(61) T. G. Leighton, ed., *The Acoustic Bubble*. San Diego, CA: Academic Press, Inc, 1st ed., 1994.

(62) E. G. Nawy, *Concrete Construction Engineering Handbook*. Boca Raton, Florida: CRC Press LLC, 1st ed., 1997.

(63) T. Rossing, ed., *Springer Handbook of Acoustics*. New York: Springer, 1st ed., 2007.

(64) J. Abel and P. Huang, "A Simple, Robust Measure of Reverberation Echo Density," *in Proceedings of the 121st AES Convention, preprint 6985*, pp. 1–10, Nov. 2006.

(65) W. Chu, "Comparison of reverberation measurements using Schroeder's impulse method and decay curve averaging method," *J. Acoust. Soc. Am.*, vol. 63, no. May 1978, pp. 1444–1450, 1978.

(66) G. Stan, J. Embrechts, and D. Archambeau, "Comparison of different impulse response measurement techniques," *J. Audio Eng. Soc.*, vol. 50, no. 4, pp. 249–264, 2002.

(67) D. H. Yom and S. Ann, "Spiking filtering of all-pass filter impulse responses," *Proceedings of the IEEE*, vol. 75, no. 12, pp. 1694–1695, 1987.

(68) T. Nguyen, “Near-perfect-reconstruction pseudo-QMF banks,” *IEEE Transactions on Signal Processing*, vol. 42, no. 1, pp. 65–76, 1994.

(69) E. A. Lehmann and A. M. Johansson, “Prediction of energy decay in room impulse responses simulated with an image-source model.,” *J. Acoust. Soc. Am.*, vol. 124, pp. 269–77, July 2008.