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

Kay L. Gemba[a)]
Marine Physical Laboratory of the Scripps Institution of Oceanography, University of California at San Diego, La Jolla, California 92093-0238, USA
gemba@ucsd.edu

Eva-Marie Nosal
Department of Ocean and Resources Engineering, School of Ocean and Earth Science and Technology, University of Hawai'i at Mānoa, Honolulu, Hawaii 96822, USA
nosal@hawaii.edu

Todd R. Reed
Department of Electrical Engineering, University of Hawai'i at Mānoa, Honolulu, Hawaii 96822, USA
trreed@hawaii.edu

Abstract: The impulse response (IR) of an acoustic channel can be obtained by cross-correlating the received signal with the broadband excitation signal in unfavorable noise conditions. However, the deconvolved IR is colored by the IRs of the combined electrical equipment. This letter presents a time domain approach using pre-computed filters to whiten the unknown coloration in order to obtain the channel's time domain waveform. The method is validated with an image-source model and the IR of the channel is recovered with spectral root mean square error of −27 dB. Data results obtained from a pool experiment with non-calibrated equipment yield a whitened IR with standard deviation of 0.9 dB (30–68 kHz band).

[CFG]
Date Received: July 11, 2016 **Date Accepted:** December 8, 2016

1. 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. Source 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. Manufactures using pulse-gating techniques to calibrate transducers supply its amplitude response rather than its impulse response (IR), resulting in source level estimates lacking phase information. In addition, impedance mismatches between connected equipment may significantly alter the system's response originally estimated from individual response curves. However, for applications including source characterization and detection, it is desirable to obtain the time domain waveform of the source.

This letter presents a method to estimate and remove all coloration if the system's transient response can be obtained. In Eq. (1), the recording $r(t)$ of a source or excitation signal $s(t)$ (here: 30–68 kHz) in a reverberant underwater channel $g(t)$ is colored by $u(t)$ with additive noises from the source $n_s(t)$, the channel $n_c(t)$, and the receiver $n_r(t)$,

$$r(t) = [[[[s(t) + n_s(t)] * g(t)] + n_c(t)] * u(t)] + n_r(t) \tag{1}$$

$$\approx s(t) * g(t) * u(t) + n_t(t), \tag{2}$$

$$h(t) = g(t) * u(t) + n_h(t). \tag{3}$$

Assuming zero mean uncorrelated noise, the model simplifies to Eq. (2), where $n_t(t)$ contains all noise contributions. The colored IR $h(t)$ in Eq. (3) can be obtained via an inverse filter (Polge and Mitchell, 1970), which changes the power of the noise [denoted by $n_h(t)$].

[a)]Author to whom correspondence should be addressed.

2. Approach

In this work, the channel $g(t)$ refers to the acoustic propagation between a fixed source-receiver pair in a pool or tank. The proposed method might also be useful for moored transducers. 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 scalar and t_d is the direct arrival time. For example, the assumption that the entire $g(t)$ is reasonably independent of frequency is accurate to the first order when ray theory is applicable (Jensen *et al.*, 2011; Schroeder, 1996).

To illustrate the requirement further, suppose an experiment is conducted in an anechoic environment. In this ideal example, $g(t)=k\delta(t-t_d)$, so $h(t)=k\delta(t-t_d) * u(t)=ku(t-t_d)$. Replacing the anechoic with a reverberant environment, the IR includes a multipath tail (a sequences of impulses corresponding to boundary reflections) in additional to the scaled direct arrival. An appropriate spreading loss model can be used to estimate k (or vice versa). In other words, we think of the direct arrival as a scaled all-pass filter with an unknown, pure delay t_d (corresponding to the channel length), which does not alter the recorded signal's phase. 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. This flat spectrum is modified when the channel's transfer function (TF) $G(\omega)$ is multiplied by the unknown equipment's TF $U(\omega)$. The character of this modified spectrum corresponds to the unknown TF $U(\omega)$ when computing the fast Fourier transform (FFT) over the duration of the direct arrival in $h(t)$. However, spectral division $[H(\omega)/U(\omega)]$ to remove coloration is sometimes undesirable and might result in narrowband noise amplification if $U(\omega)$ is poorly conditioned. Furthermore, $U^{-1}(\omega)$ is acausal if $u(t)$ displays non-minimum phase behavior (Robinson and Treitel, 1980; Gemba and Nosal, 2016). To circumvent problems associated with stability and causality, the following discussion focuses on the area of the direct arrival in the time domain. In other words, we want to recover the real waveform $g(t)$ without *spectral* division.

To recover $g(t)$, we filter $h(t)$ into $i=\{1,2,\ldots,n\}$ band-limited versions using a pseudo quadrature mirror filter bank (PQMFB, Nguyen, 1994) with near-perfect reconstruction properties. The reconstruction error depends on the filter order and can fall below -100 dB. All filters in the bank $f(t)$ 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 choice of a PQMFB is evident: after filtering the signal of interest, each bandlimited signal can be appropriately *whitened*. All bandlimited signals are subsequently summed up to yield a manipulated version of the original signal over the full bandwidth. We select the number of filters by computing the FFT over the direct arrival found in $h(t)$ such that the resulting magnitude spectrum is approximately constant in each band (similar to a step function). Forward and reverse filtering requires that the direct arrival is separated from the first reflection by twice the filter's IR length. This requirement either limits the size of the filter bank since a narrow band filter requires more coefficients or imposes limitations on the use of this method in practice.

To uncolor the IR of the channel, we convolve the filter bank $f(t)$ with Eq. (3):

$$\begin{aligned} f(t)*h(t) &= f(t)*g(t)*u(t)+f(t)*n_h(t) \\ &= \sum_{i=1}^{n} f_i(t)*g(t)*u(t)+\sum_{i=1}^{n} f_i(t)*n_h(t). \end{aligned} \tag{4}$$

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

$$\begin{aligned} &\int_{t_0}^{t_e} |g(t)*f_i(t)*u(t)+f_i(t)*n_h(t)|dt \\ &= \int_{t_0}^{t_e} |\delta(t)*f_i(t)*u(t)|dt+\int_{t_0}^{t_e} |f_i(t)*n_h(t)|dt = b_i+\sigma_i. \end{aligned} \tag{5}$$

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

$|U_i(\omega)|$ similar to a step function, and coefficients are denoted by b_i with additive noise σ_i. The overlap between adjacent bands contributes to errors in b_i. It should be noted that the bandwidth for each selected filter is not narrow (>1 kHz). A common design requirement for electrical equipment is to keep the frequency response as flat as possible.

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 $\hat{g}(t)$.

3. Simulation

We validate our method using a room response, image-source model (ISM, Lehmann and Johansson, 2008). The pool dimensions are 22.9 m by 22.9 m with a depth of 5.2 m, the soundspeed is a constant 1500 m/s. The channel is 1.5 m above the bottom, approximately at the center of the pool and has a length of 1 m. This configuration yields a direct arrival to first reflection separation of ∼1.5 ms. The ISM is used to generate an IR $g(t)$ which is subsequently convolved with $u(t)$. We add white Gaussian noise at signal-to-noise ratio (SNR) = 10 dB to the linear chirp with convolved IRs [Eq. (2)] before cross correlation and scaling by the squared magnitude spectrum of the excitation signal to obtain $h(t)$ [Fig. 1(a)]. The sampling rate is 140 kHz, the signal and the noise are each 1 s in duration and cover the same bandwidth. The magnitude response of $u(t)$ is shown in Fig. 1(b) and approximates a piezoelectric transducer amplified by other equipment for a total dynamic range close to 20 dB. Several resonance peaks are included to achieve a broadband response. A response with increased coloration character is assumed for this simulation when compared to the factory supplied response of our experimental transducer.

$\hat{g}(t)$ is recovered [Fig. 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 broadband to narrowband: 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. 1(e)] is largest between two adjacent filters (at the transition region of each band or 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. SNRs of 20, 10, and 5 dB yield spectral root-mean-square errors of −35, −27, and −22 dB. The gain of the matched filter as well as the SNR determines when the IR decays into the noise floor. The noise variance in Fig. 1(d) determines how many IR peaks can be resolved. To exclude noise, we only use values $\geq |0.01|$ of the entire IR to compute the error $[20\log_{10}(|G(\omega) - \hat{G}(\omega)|)]$ in Fig. 1(e). The "S" in SNR [Eq. (2)] favors stronger IR peaks and a different SNR measure (e.g., considering the weakest multipath of

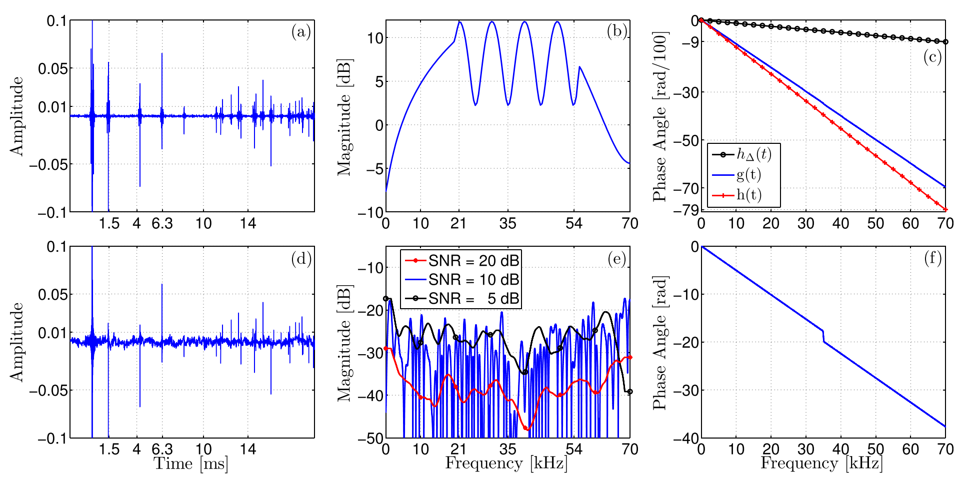

Fig. 1. (Color online) SNR = 10 dB results: (a) $h(t)$ and (b) $|U(\omega)|$ of the unknown transfer function. (c) Phase responses of $g(t)$, $h(t)$, and the all-pass filter using the direct arrival in $h(t)$ [denoted by $h_\Delta(t)$]. (d) Recovered IR $\hat{g}(t)$ and (e) recovery error. Error curves for SNR 20 and 5 dB scenarios are smoothed to allow for a comparison. (f) Phase angle error $[\angle G(\omega) - \angle \hat{G}(\omega)]$.

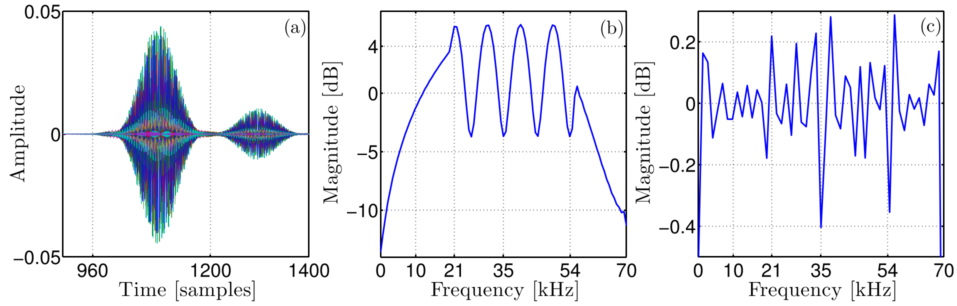

Fig. 2. (Color online) (a) $f(t) * h(t)$ (n = 30, each color corresponds to a different band) with integration limits $t_0 = 960$ samples and $t_e = 1200$ samples. The direct arrival of $h(t)$ is located at ∼1100 samples, the first reflection at ∼1300 samples. Note that the IR of each filter decays to 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. 1(b) but with a different range. (c) $|\hat{G}(\omega)|$ measured over integration limits. Ideally, the response should be 0 dB.

interest) could serve as a practical way to define SNR. Unwrapped phase responses of individual IRs are shown in Fig. 1(c) and phase error in Fig. 1(f). Figure 2 shows details of the recovery process: Fig. 2(a) displays a typical plot when $f(t)$ is convolved with $h(t)$, and Fig. 2(b) and 2(c) show the magnitude spectrum of the direct arrival in $h(t)$ and $\hat{g}(t)$, respectively. The magnitude spectrum in Fig. 2(c) is used as an empirical performance measure; it indicates if more filtering is required or if a certain accuracy is achieved.

4. Channel estimation in an underwater pool environment

An experiment (for more details, see Gemba *et al.*, 2014; Gemba and Nosal, 2016) was conducted in a reverberant pool environment (concrete wall overlaid with square tiles) using a linear chirp (5–85 kHz) as the excitation signal with transducer separation of 1 m to estimate $\hat{g}(t)$. The chirp is not windowed and was designed to well exceed the bandwidth of interest to reduce edge transients. The deconvoled IR is windowed using a combination of a rectangular and a Kaiser window to tapper early and late samples smoothly to zero and eliminate a step-function response. The channel geometry is the same as in the model with first and second arrivals similar to Fig. 2(a), so we can use the same filter length. The unknown TF is estimated to be smoother and to have less dynamic range than Fig. 1(b). In particular, all electrical components were individually or factory calibrated but an unknown impedance mismatch changed the overall response. We did not correct for using individual calibrated components but treat the experienced response as the unknown TF.

Recordings were conducted in the presence of light winds and turned off pool pumps, yielding a direct arrival in $h(t)$ of ∼45 dB over the noise floor. The recovered IR $\hat{g}(t)$ is presented in Fig. 3(a). Only some of the early wall reflections match the reflections in the simulations because one of the pool walls is slanted. The spectrum computed over the direct arrival of the colored $h(t)$ and whitened $\hat{g}(t)$ is shown in Fig. 3(b). The optimum frequency response of the CR1 Sensor Technology Limited transducer (Seattle, WA, SN: 09178-01) dictates SNR and is 30–68 kHz. Results show that the dynamic range of the whitened IR is significantly decreased within the equipment's optimum

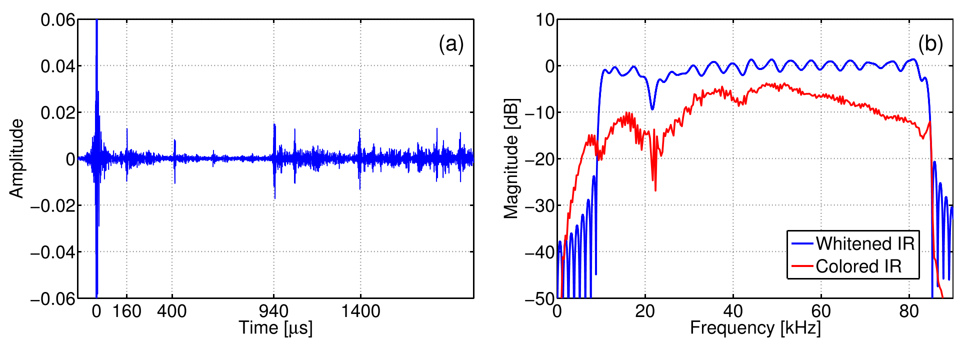

Fig. 3. (Color online) (a) IR estimate of recording channel with x-ticks corresponding to theoretical boundary reflection times. The scale is chosen to show details of the reflections, but cuts off the direct arrival, which has a maximum amplitude between ±0.2. The whitened IR in (b) has a standard deviation of 0.9 dB (30–68 kHz). The spectrum of the colored IR has an increased dynamic range.

frequency response. The oscillatory character is likely an artifact of the filter bank. This result only displays one realization and an ensemble average is required to capture the pressure distribution in a reverberant environment (Waterhouse, 1968; Gemba and Nosal, 2016).

5. Conclusions

We have shown that it is possible to separate the IR $g(t)$ from the colorations $u(t)$ and recover an estimate of a channel's IR. This method can be used to calibrate an unknown transducer in a controlled environment using frequency modulated signals. It might further be valuable in order to observe calibration changes on long-term deployed systems subject to biofouling. The presented time domain solution to the method might be replaced with a frequency domain approach [e.g., see Eq. (14) in Gemba and Nosal, 2016].

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. (3)] by the inverse amplitude response of equipment with significant character (such as transducers) or with a power spectral density estimate of the noise. Pre-coloring will serve as a pre-equalizer and increase the SNR of $h(t)$. Furthermore, pre-coloring will add a shading term to the frequency equivalent of Eq. (3), decreasing the dynamic range of $|U(\omega)|$ and possibly improving the estimate of the channel $\hat{g}(t)$.

References and links

Gemba, K. L., and Nosal, E.-M. (**2016**). "Source characterization using recordings made in a reverberant underwater channel," Appl. Acoust. **105**, 24–34.

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

Jensen, F., Kuperman, W. A., Porter, M., and Schmidt, H. (**2011**). *Computational Ocean Acoustics*, 2nd ed. (Springer Science, New York), 764 pp.

Lehmann, E. A., and Johansson, A. M. (**2008**). "Prediction of energy decay in room impulse responses simulated with an image-source model," J. Acoust. Soc. Am. **124**, 269–277.

Nguyen, T. (**1994**). "Near-perfect-reconstruction pseudo-QMF banks," IEEE Trans. Signal Process. **42**, 65–76.

Polge, R., and Mitchell, E. (**1970**). "Impulse response determination by cross correlation," IEEE Trans. Aerospace Electron. Syst. **AES-6**, 91–97.

Robinson, E. A., and Treitel, S. (**1980**). *Geophysical Signal Analysis* (Prentice-Hall, Englewood Cliffs, NJ), 466 pp.

Schroeder, M. R. (**1996**). "The 'Schroeder frequency' revisited," J. Acoust. Soc. Am. **99**, 3240–3241.

Waterhouse, R. V. (**1968**). "Statistical properties of reverberant sound fields," J. Acoust. Soc. Am. **43**, 1436–1444.