The Watermark Benchmark for Underwater Acoustic Modulation Schemes

— Source link ⧉

Paul van Walree, Francois-Xavier Socheleau, Roald Otnes, Trond Jenserud

Institutions: Norwegian Defence Research Establishment

Topics: Watermark and Benchmark (computing)

Related papers:

- Statistical Characterization and Computationally Efficient Modeling of a Class of Underwater Acoustic Communication Channels

- Underwater acoustic communication channels: Propagation models and statistical characterization

- Multicarrier Communication Over Underwater Acoustic Channels With Nonuniform Doppler Shifts

- Underwater Acoustic Network Simulation With Lookup Tables From Physical-Layer Replay

- A new image watermarking technique based on noise modulation

The Watermark Benchmark for Underwater Acoustic Modulation Schemes

Paul van Walree, François-Xavier Socheleau, Roald Otnes, Tron Jenserud

The Watermark benchmark for underwater acoustic modulation schemes

Paul A. van Walree, François-Xavier Socheleau, Roald Otnes, and Trond Jenserud

P. A. van Walree, R. Otnes, and T. Jenserud are with the Maritime Systems Division at Norwegian Defence Research Establishment (FFI), P.O. box 115, NO-3191 Horten, Norway.

F.-X. Socheleau is with IMT Atlantique, Lab-STICC, Univ. Bretagne Loire, 29238 Brest, France.

Abstract

Watermark is a freely available benchmark for physical-layer schemes for underwater acoustic communications. It allows researchers to test and compare algorithms for the physical layer under realistic and reproducible conditions. The benchmark is a shell around the validated FFI channel simulator Mime, which is driven by at-sea measurements of the time-varying impulse response. The first release of Watermark is issued with a library of channels measured in Norway (two sites), France, and Hawaii, offering three frequency bands (4–8, 10–18, and 32.5–37.5 kHz), single-hydrophone and array receivers, and play times varying from 33 seconds to 33 minutes.

Index Terms

Benchmark testing, channel simulator, wideband systems, signal fluctuations, acoustic communication.

I. INTRODUCTION

Numerous modulation schemes have been proposed for underwater acoustic communications during the past decades. In the majority of cases it is impossible to compare the performances of these schemes in a meaningful way, as field tests are performed in different environments. Moreover, the variation in channel characteristics is large even in a given environment [1], [2]. Unlike terrestrial radio-frequency communications, underwater acoustic communications is a technological field deprived of standard test channels. A variety of channel models and channel simulators have been proposed for acoustic communications, e.g. [3]–[10]. These models have varying degrees of realism, completeness, and availability, and in most cases they are tuned to a single environment and platform configuration.

The underWater AcousTic channEl Replay benchMARK (Watermark) is a realistic simulation tool now made available to the underwater communications community. It is a shell around the validated FFI channel simulator "Mime" [7], [11], which convolves user waveforms with measured channels. The initial release is provided with a library of channels from four environments. The primary objective of the present initiative is to provide an *easy-to-use* benchmark that enables development, testing, and comparison of physical-layer algorithms under *realistic* and *reproducible* conditions.

Watermark was presented at UComms 2016 with two test channels for a single receiver [12]. Following the presentation at UComms, IMT Atlantique (formerly known as Telecom Bretagne)

generously offered array data [13] for release with the benchmark, and, soon after, Scripps Institution of Oceanography approved a request to include array data from the Kauai Acomms MURI 2011 (KAM11) experiment [14], [15].

This Technical Communication summarizes the principle of channel replay, presents the available test channels, discusses the operation of Watermark, and gives examples of its use. Format descriptions and detailed user instructions are provided in a separate report [16]. The software, test channels and user manual are available for download from [17]. Watermark is programmed in MATLAB and has been successfully tested on Windows, Linux, and Mac OS.

II. THE MIME CHANNEL SIMULATOR

Mime distorts input waveforms by convolving them with measured channels, or statistical variations thereof. Its operation principle is known as channel replay

$$y(t) = \int_{-\infty}^{\infty} \hat{h}(t,\tau)\, x(t-\tau)\, \mathrm{d}\tau + n(t)\;, \tag{1}$$

where $x(t)$ is the input signal, $\hat{h}(t,\tau)$ a time-varying impulse response (TVIR), $n(t)$ a noise term, and $y(t)$ the distorted output signal. Channel replay can be performed in several ways [7], [8], [11], [18], [19], and recently the term has also been used for network simulations with measured packet error statistics [20].

Mime has a direct-replay mode where $\hat{h}(t,\tau)$ is the measured channel, which is an estimate of the true channel $h(t,\tau)$. The maximum simulation (play) time with direct replay equals the measurement time. Mime also has a stochastic replay mode, which generates TVIR realizations that are statistically consistent with the measured channel. This mode can run forever on a single input measurement. The realism of direct-replay simulations depends primarily on the quality of the channel measurement. The measurement errors have been well documented in the case of correlative channel sounders [21], and the combination of Mime and correlative sounders has been well validated [7], [11]. Simulated bit and packet error ratios, and receiver output signal-to-noise ratio (SNR), are close to the corresponding values measured at sea. The stochastic replay mode has also yielded good validation results, but has limitations regarding the reproduction of non-stationary channels, time-varying delays, and correlated scattering. To guarantee realistic simulations under all conditions, Watermark is released only with the direct-replay mode of Mime. This mode faithfully reproduces all effects mentioned in [2], including

channel non-stationarity, correlated scattering, time-varying delays, and ultra-wideband channels. The one type of channel that cannot be reproduced is the overspread channel, because it cannot be measured properly.

Another advantage of direct replay over stochastic replay is that the code is simple, fast, and free of user parameters. Its disadvantage is that a significant play time requires a significant measurement time, and a correspondingly bulky channel library. Acoustic channels are usually not stationary over longer periods, for reasons such as platform motion, weather variability, movement of the tides, etc. Non-stationarity of test channels is a price paid for the realism, just like testing at sea has to deal with channel non-stationarity. For a fair and reproducible comparison between modulation schemes, it is recommended that performance results reported for Watermark channels are averaged over the full play time of the replay channel.

III. Channel library

Watermark is released with sounding (TVIR measurement) data for five channels. Table I summarizes measurement conditions and test channel parameters, while Fig. 1 provides schematic drawings of the platform deployment. The transmitter is bottom-mounted in NOF1 (Norway — Oslofjord) and NCS1 (Norway — Contintental Shelf), suspended in the water column in BCH1 (Brest Commercial Harbor), and towed by a surface ship in KAU1 and KAU2 (Kauai). The receiver is a bottom-mounted hydrophone (NOF1, NCS1) or a vertically suspended hydrophone array (BCH1, KAU1, KAU2). NOF1 and NCS1 are single-input single-output (SISO), whereas BCH1, KAU1, and KAU2 are single-input multiple-output (SIMO). Both types of channels are represented by a set of channel files. A channel file contains a single-hydrophone TVIR measurement, typically with a duration of 30–60 seconds. A set of channel files represents consecutive soundings in the SISO case, and a simultaneous sounding on consecutive hydrophones in the SIMO case.

The frequency band and sounding duration in Table I are properties of the employed channel probe signal, but they also give the nominal band and the maximum signal duration for Watermark simulations, as explained in Sec. IV-B. Probing waveforms are cyclic transmissions of pseudonoise (m-sequence) or linear frequency modulated (LFM) pulses. Both signal types have a (root) raised-cosine spectrum [22, pp. 607–608] with a small roll-off factor, which yields a flat spectrum over a large part of the given band. The importance of a flat probe signal spectrum is

TABLE I

TEST CHANNEL PARAMETERS AND SOUNDING CONDITIONS.

Name	NOF1	NCS1	BCH1	KAU1	KAU2
Environment	Fjord	Shelf	Harbour	Shelf	Shelf
Time of year	June	June	May	July	July
Range	750 m	540 m	800 m	1080 m	3160 m
Water depth	10 m	80 m	20 m	100 m	100 m
Transmitter depl.	Bottom	Bottom	Suspended	Towed	Towed
Receiver depl.	Bottom	Bottom	Suspended	Suspended	Suspended
Probe signal type	LFM train	Pseudonoise	Pseudonoise	LFM train	LFM train
-3 dB freq. band	10–18 kHz	10–18 kHz	32.5–37.5 kHz	4–8 kHz	4–8 kHz
Roll-off fact.	1/8	1/8	1/10	1/8	1/8
Sounding duration	32.9 s	32.6 s	59.4 s	32.9 s	32.9 s
Delay coverage	128 ms	32 ms	102 ms	128 ms	128 ms
Doppler coverage	7.8 Hz	31.4 Hz	9.8 Hz	7.8 Hz	7.8 Hz
Type	SISO	SISO	SIMO	SIMO	SIMO
# hydrophones	1	1	4	16	16
Element spacing	—	—	1 m	3.75 m	3.75 m
# cycles	60	60	1	1	1
Cycle time	400 s	600 s	—	—	—
Total play time	33 min	33 min	1 min	33 s	33 s

explained in Sec. IV-B.

The delay coverage is the pulse repetition period of the channel sounder, and the Doppler coverage is its reciprocal. These numbers give the maximum possible delay spread and Doppler spread in the replay channel. Energy in the delay-Doppler spread function of the true channel that is outside these limits, is aliased in the channel estimate [21], [23]. Aliasing is often unavoidable in acoustic channels probed with cyclic sounders, but the measurement error is small when the aliased energy is a small fraction of the total energy. The channels in Table I are well behaved and offer realistic replay.

The next subsections characterize the test channels with delay-Doppler spread functions, i.e., two-dimensional power densities which give the distribution of received signal power vs. time delay τ and frequency shift υ. The figures zoom in on the main multipath arrivals and the mean Doppler shift has been removed. More comprehensive channel analyses are shown in the user
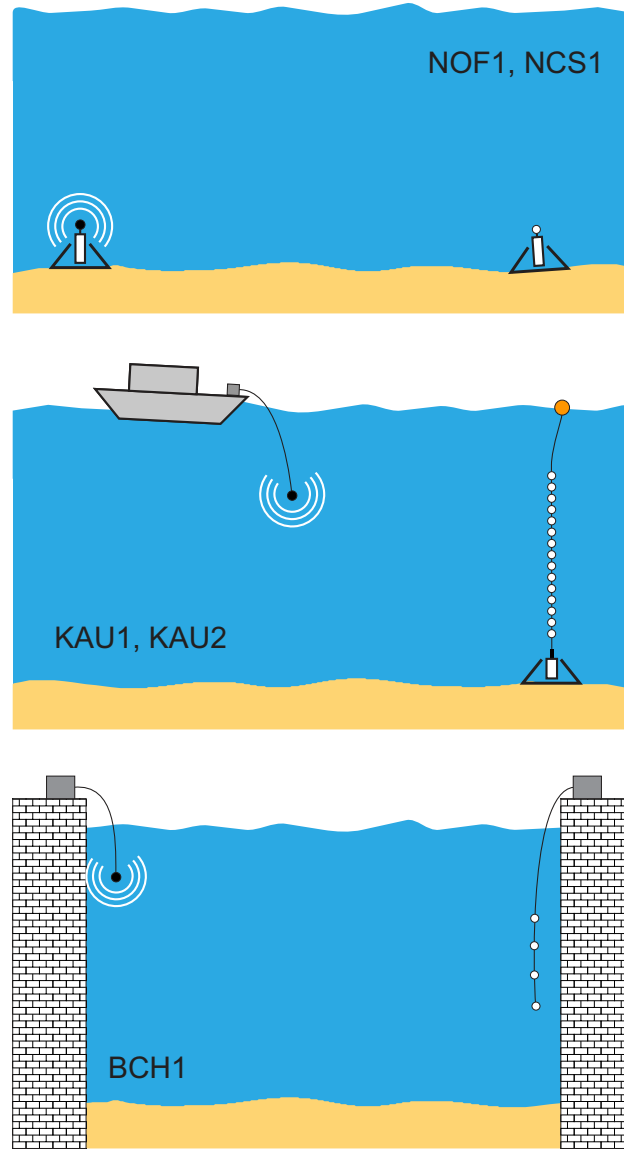
Fig. 1. Sketch of the deployment setups.

manual [16].

A. Norway – Oslofjord (NOF1)

NOF1 is a channel measured in a shallow stretch of Oslofjorden between a stationary source and a stationary single-hydrophone receiver. The data consist of a TVIR measurement over 32.9 s, repeated 60 times at 400-s intervals, yielding a total play time of $\approx$ 33 minutes. Note
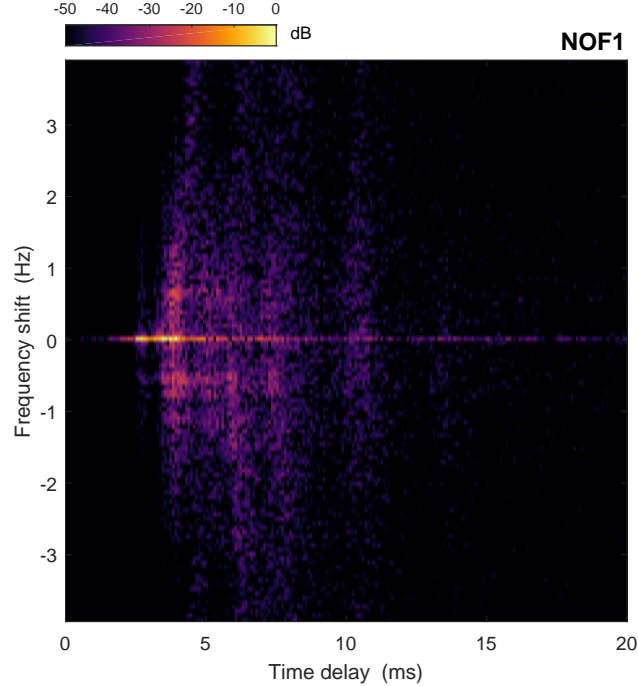
Fig. 2. Delay-Doppler spread function of the first NOF1 sounding.

that the play time is discontinuous. For instance, a 5-s communication packet fits 6 times in a single TVIR (see Sec. V-C) and yields $60 \times 6 = 360$ independent packets in simulation. The first 6 packets are transmitted between t_0 and $t_0 + 32.9$ s, packets 7–12 between $t_0 + 400$ s and $t_0 + 432.9$ s, etc.

Figure 2 characterizes the channel, where zero delay is arbitrarily placed just before the first arrival. This arrival has no frequency spread, whereas later arrivals are Doppler spread due to sea surface interactions. There is also a streak of delayed clutter at $v = 0$, due to reflection off stationary scatterers. Despite the scattering, most energy is concentrated in a narrow delay-Doppler window (between $\tau = 3$ ms and $\tau = 5$ ms, at $v = 0$ Hz). NOF1 is a relatively benign communication channel.

B. Norway — Continental Shelf (NCS1)

NCS1 was measured on Norway's continental shelf between a stationary source and a stationary single-hydrophone receiver. The data consist of a TVIR measurement over 32.6-s, repeated 60 times at 600-s intervals, yielding a total play time of ≈ 33 minutes. This play time is
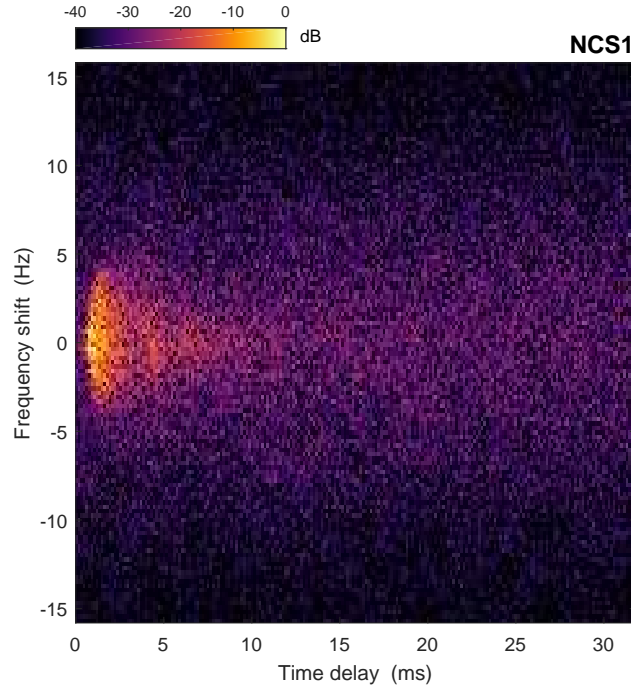
Fig. 3. Delay-Doppler spread function of the first NCS1 sounding.

discontinuous. For instance, a 5-s communication packet fits 6 times in a single TVIR and yields $60 \times 6 = 360$ independent packets in simulation. The first 6 packets are transmitted between t_0 and $t_0 + 32.6$ s, packets 7–12 between $t_0 + 600$ s and $t_0 + 632.6$ s, etc.

Figure 3 characterizes the channel, which has in common with NOF1 that most energy is concentrated at the start of the impulse response. Otherwise it is a completely different channel. There are no stable paths, and the coherence time is of order 100 ms [16]. NCS1 is more challenging than NOF1, in particular for coherent communication schemes which need to estimate the phase of a signal.

Challenging channels are difficult to measure, and NCS1 has the largest measurement errors among the Watermark test channels. The true channel has long diffuse tails in delay and in Doppler, which do not suddenly stop at the boundaries of the measurement box in Fig. 3. These tails are aliased in the measurement. However, the fraction of the total energy that is aliased is not necessarily large. A validation has been performed to assess the impact of aliasing and other measurement errors for NCS1, concluding that Watermark simulation with NCS1 yields realistic communication performance [12]. At any rate, a channel library with only perfect measurements
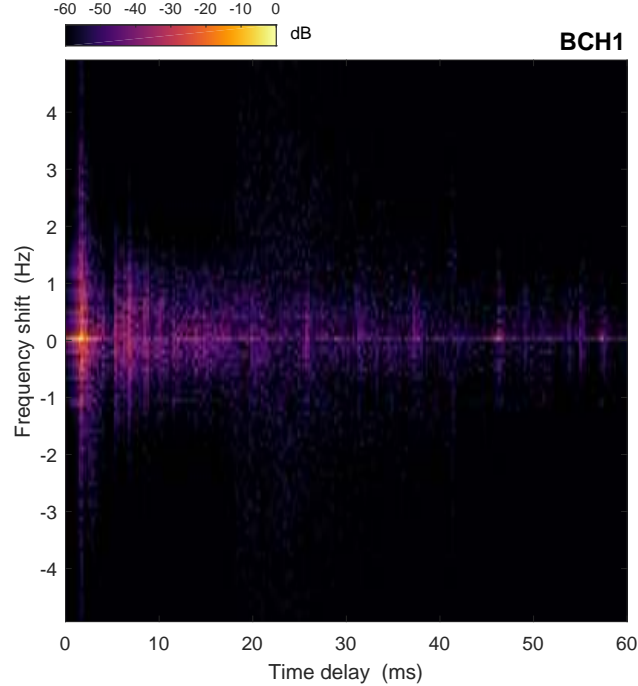
Fig. 4. Delay-Doppler spread function of BCH1, hydrophone channel 2.

would imply a library with only benign channels. Some measurement errors have to be tolerated to give the community a taste of less forgiving channels.

C. Brest Commercial Harbor (BCH1)

BCH1 is a SIMO channel measured in the commercial harbour of Brest, France. A source and a 4-element array were lowered into the water column from two docks [13]. A single probe transmission over a range of 800 m resulted in a TVIR measurement over 59.4 s, simultaneously recorded on the four hydrophones. Figure 4 characterizes the channel. Similarly to NOF1, the channel is a mixture of stable and fluctuating arrivals, but with a larger number of distinct trailing paths.

The spatial correlation is relatively high. For instance, for the two main arrivals ($\tau = 1.8$ and $\tau = 7$ ms) the Pearson correlation coefficient across the array ranges from 0.5 to 0.9. Although correlated, the SIMO channel shows some spatial selectivity, e.g., a ≈ 9 dB difference in propagation loss between hydrophone channels 1 and 3.

D. Kauai 1 (KAU1)

KAU1 is a SIMO channel from the Kauai Acomms MURI 2011 (KAM11) experiment. This experiment was conducted in shallow water off the western side of Kauai, HI, USA [14], [15]. The channel was probed between 4 and 8 kHz, between a towed source and a vertically suspended array with 16 receivers (one of which was broken). The play time is 33 s. Figure 5 characterizes the channel for a hydrophone near the center of the array. Arrivals clusters are observed, whose Doppler spread increases with an increasing delay. This is caused by an increasing number of surface bounces occurring at increasing grazing angles.

In addition to Doppler spreading due to variability of the medium, there is a time-varying Doppler shift (TVD) due to tow ship motion. This also affects the direct path and explains its Doppler spread. This can be verified by resampling the received probe signal with a time-varying resampling factor derived from the phase of the direct path [7], [23]. The resulting channel is shown in Fig. 6, revealing elimination of the Doppler spread from the first arrival. Although difficult to observe, the resampling also reduces the Doppler spread on the trailing clusters. In KAU1, communication schemes may benefit from adaptive resampling or a phase-locked loop (PLL) [24], assisting the receiver in dealing with TVD. Note that TVD removal is not included in Watermark; it is applied separately here to illustrate a property of the channel.

E. Kauai 2 (KAU2)

KAU2 is another SIMO channel from the KAM11 experiment. It concerns the same probe signal transmission as KAU1, but recorded on a different array. The main difference with KAU1 is a longer range, which attenuates delayed arrival clusters more rapidly: Figure 7. TVD contributes significantly to the total Doppler spread in this channel, similarly to KAU1.

IV. Technical remarks

A. Propagation loss and time synchronization

Watermark is not calibrated for propagation loss, but propagation loss differences between consecutive soundings (NOF1, NCS1) or consecutive hydrophones (BCH1, KAU1, KAU2) are preserved. This means that the signal level of simulated waveforms may fluctuate in time and depth. Watermark also preserves time synchronization of signals received on the vertical arrays.
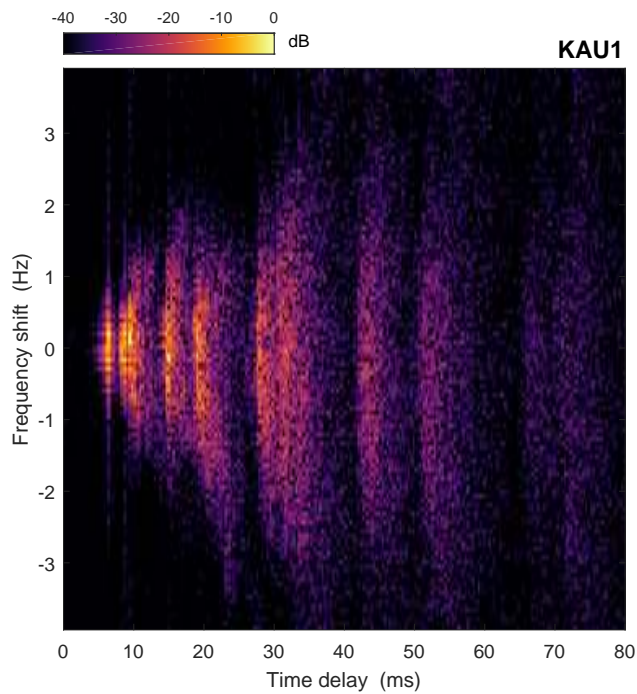
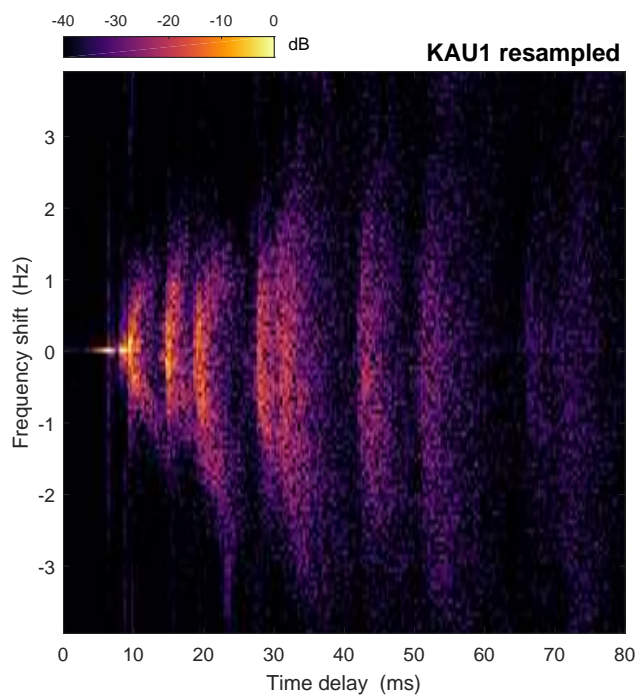Fig. 5. Delay-Doppler spread function of KAU1, hydrophone channel 8.

Fig. 6. Delay-Doppler spread function of KAU1, hydrophone channel 8, after TVD removal.
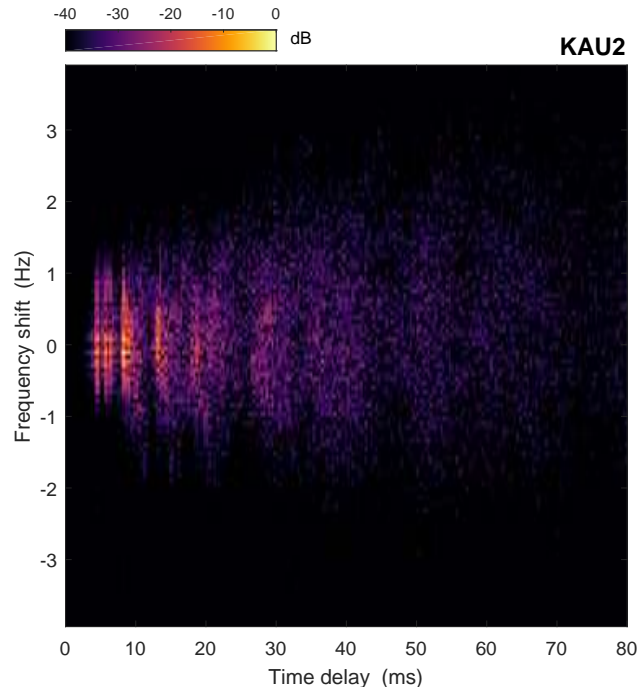
Fig. 7. Delay-Doppler spread function of KAU2, hydrophone channel 8.

This means, for example, that beamforming Watermark data gives correct results. Whether the correct results are desirable results is a different question, as the element spacing is much larger than the acoustic wavelength in the three SIMO channels.

B. Transfer function

Input signals have to comply with the frequency bands listed in Table I. Watermark does not use explicit filters, but the channel files contain an implicit bandpass filter. Correlative sounders yield ultra-wideband channel estimates $\hat{h}(t, \tau)$ [25], where the transfer functions of the channel, the sounder hardware, and the probe signal are superimposed on one another. The composite transfer function is reproduced in direct replay. For this reason, the probe signals of Table I have a near rectangular spectrum. In the flat region of the probe signal spectrum (e.g., 10.5–17.5 kHz for NOF1 and NCS1), the spectral distortion of received signals is not affected by probing waveform characteristics, but entirely due to frequency-dependent propagation loss and hardware effects.

The probe signal spectrum, and thus the composite transfer function, falls off rapidly outside

the nominal band (10–18 kHz in the case of NOF1 and NCS1). Excess bandwidth of user waveforms will be strongly attenuated. Communication schemes using a different band than the nominal band of a test channel should be adapted, because Watermark will not adapt to these schemes.

The frequency response is revealed by passing white noise over a test channel. Figure 8 illustrates this for an NCS1 sounding. The top panel shows a spectrogram of the received probe signal, and the middle panel that of the input signal $x(t)$ (see Eq. 1).[1] A spectrogram of the output signal $y(t)$ is shown in the bottom panel. It strongly resembles that of $x(t)$, as it should, and shows that simulation is only possible where there is signal energy in the probe signal spectrogram. This makes sense, as it is impossible to reproduce something which has not been measured. The main difference between the spectrograms is the weak reverberation tail, seen at the end of the probe signal, which is not reproduced in simulation. This energy is aliased in the channel estimate.[2]

C. Doppler effects

The mean Doppler shift V_0 is removed from the sounding data before the channel estimation. This is done by resampling the raw acoustic data by a resampling factor $1 - V_0/c$, where $c = 1500$ m/s is the nominal sound speed. The used sign convention is such that a positive velocity corresponds to a positive range rate, i.e., time dilation of the signal. This Doppler shift is missing from the estimates $\hat{h}(t, \tau)$ in the channel library, and from Figs. 2–7 for that matter.

The reason for removing the mean Doppler shift is that this minimizes the measurement errors of the channel sounder. Instantaneous Doppler spreading and time-varying Doppler shifts around the mean value are reproduced in direct replay, whereas the mean shift V_0 is reinstalled when Watermark retrieves the simulated packets. These packets are resampled by a factor $1/(1-V_0/c)$. This ensures that all Doppler effects experienced by the channel probe signal in the real channel are transferred onto the input signal $x(t)$ in the simulated channel. The Doppler spread in a replay channel may be slightly reduced compared with the original channel [11].

---

[1] Note that the noise is the input signal $x(t)$ in this case, not the additive term $n(t)$.

[2] Actually Watermark does not allow input signals whose length exceed that of the probe signal; the parts of $y(t)$ before $t = 4$ s and after $t = 37$ s have not been simulated.

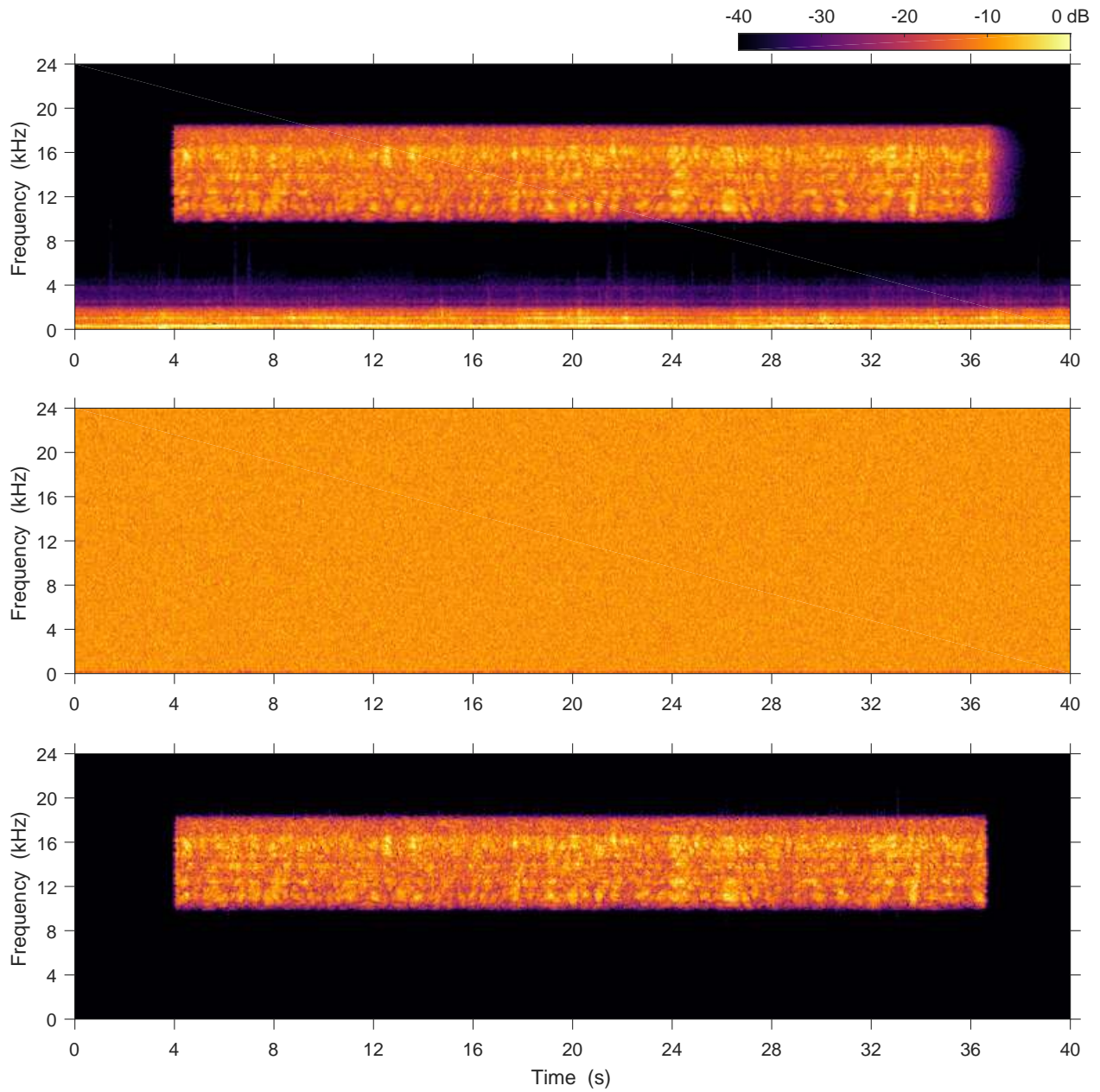Fig. 8. Spectrograms of a recorded probe signal (top), white Gaussian noise as the input $x(t)$ to Watermark (middle), and the output signal $y(t)$ (bottom).

Watermark does not discriminate between true Doppler shifts due to platform motion, and apparent Doppler shifts due to clock frequency offsets (CFOs). Both effects affect the duration of received acoustic signals. A CFO is a mismatch between the nominal (assumed) and the

true sample rate of a device. There can be an offset in both sender and receiver, resulting in an apparent Doppler shift V_0 whose magnitude depends on the quality of the clocks. Offsets corresponding to an apparent velocity of several cm/s are not uncommon.

Watermark reproduces the at-sea conditions, including hardware effects, and hence simulated packets can be subject to a small Doppler shift even between bottom-mounted senders and receivers. Communication systems need to deal with such shifts in simulation, just like they need to deal with such shifts in subsea applications.

V. PRINCIPLES OF OPERATION

This section describes the operation of Watermark, omitting detailed instructions concerning file formats, function calls, etc. These are documented in the user manual [16].

A. System requirements

Watermark runs in MATLAB and works under Windows, Linux, and Mac OS. It uses the Signal Processing toolbox and requires release R2012b or more recent. The reinsertion of small Doppler shifts (Sec. IV-C) requires systems with sufficient random-access memory. Contemporary work stations should be fine for the channels in Table I.

B. Input signal

The first step is preparation of a passband waveform, typically a communication packet, for transmission over a given test channel. The packet should fit into a channel estimate in both time and frequency (Table I). Packets longer than the measured TVIR are refused and frequency contents outside the nominal band is strongly attenuated in simulation. Users also need to supply the number of bits in the packet, defined as the number of physical-layer information bits available to higher layers in the protocol stack. It does not include overhead due to synchronization, training, physical-layer header, and error-correcting codes. But if a communication protocol stack is considered, it does include overhead due to layers above the physical layer.

C. Channel replay

Watermark stacks multiple copies of the input signal packet, as many as fit into the TVIR. It uses a spacing between successive packets to avoid reverberation from one packet spilling into
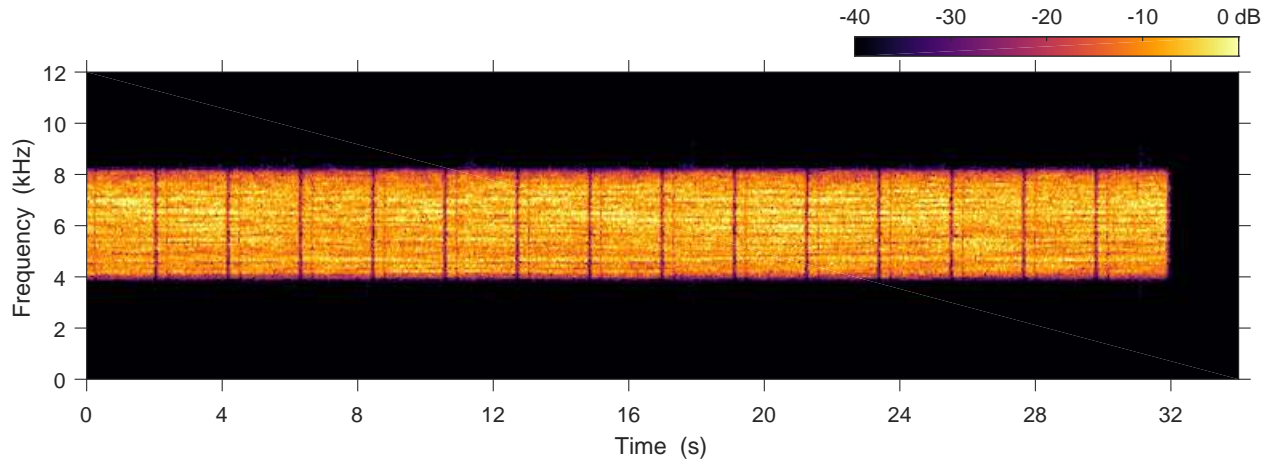
Fig. 9. Spectrogram of a filtered packet train in KAU2 (hydrophone 8) for a signal with a length of 2 s.

the next one. Since the replay channel has a measurement-imposed upper limit on the delay spread, the required spacing is precisely known. The packet train is convolved with the TVIR using direct replay.

The number of packets returned by Watermark for a single TVIR depends on the input signal duration T_{s}, the play time T_{play}, the pulse repetition time of the channel sounder T_{rep}, and a safety margin $\Delta = 4$ ms, which is used accommodate possible sidelobes occurring in the filtering process:

$$N = \lfloor T_{\mathrm{play}}/(T_{\mathrm{s}} + T_{\mathrm{rep}} + \Delta) \rfloor . \tag{2}$$

For instance, a communication packet with length 2 s would fit 16 times in a KAU2 sounding with play time 32.9 s, but with the required spacing of $T_{\mathrm{rep}} + \Delta = 132$ ms it fits only 15 times: Fig. 9. For the total number of packets returned by Watermark, N has to be multiplied by the number of sounding cycles (SISO channels) or the number of hydrophones (SIMO channels). However, the number of SIMO packets available for array processing is just N.

D. Retrieving packets

Filtered signal trains (such as in Fig. 9) are stored on disk by Watermark. Functions are available to retrieve simulated packets for SISO and SIMO applications. SISO packets can be

retrieved with additive white Gaussian noise (AWGN) at a specified E_b/N_0 value, where E_b is the signal energy per information bit and N_0 the noise power spectral density. The noisy waveform is ≈ 10 seconds longer than the actual packet, which arrives with a random delay after 4–6 seconds. Detection and synchronization are essential tasks of modems and modulation schemes, and Watermark aims to enable a realistic comparison of autonomous receivers.

The characteristics of ambient noise in the oceans vary greatly with environment and conditions [5], [26]–[28]. Gaussian statistics may occur in regions where the noise is dominated by sea surface agitation, but it will generally be coloured. The reason for using AWGN in Watermark is that this permits the use of the unambiguous SNR per bit (E_b/N_0) metric. SISO packets can also be retrieved without noise, giving users freedom to apply different noise models, or measured noise, if available. SIMO packets can only be retrieved without noise, because Watermark version 1 lacks a noise model with realistic spatial correlation properties.

Note that retrieving multiple SISO packets at a fixed SNR is not equivalent to at-sea signaling with a constant source level, because both signal level and noise level may vary in the field. Signal level fluctuations are reproduced by Watermark, which simply scales the noise accordingly to realize the target SNR for each packet.

E. Reporting results

When Watermark is used to demonstrate the performance of a modulation scheme, it is good practice to produce an unambiguous report that allows other people to compare their systems in a meaningful way. There is a lack of standard terminology in underwater acoustic communications, causing people to use different definitions for seemingly trivial quantities such as SNR, bandwidth and bit rate. It is recommended to report, in addition to whatever one would like to share about the proposed algorithms:

- The name of the test channel and the Watermark version.
- The number of information bits in the packet.
- The effective bit rate reported by Watermark.
- The signal bandwidth.
- The total number of packets in the test channel.
- The type of noise in simulation (e.g. the default AWGN or something else) .

- Whether communication results are averaged over all packets or a specific subset. Averaging types other than the arithmetic mean should be stated.

- Used definitions of bandwidth, input SNR, output SNR, etc., as appropriate.

- Whether pre-existing knowledge is used about signal start or Doppler shift, or whether the receiver has to figure out everything itself.

- To what extent signal and receiver parameters are tuned to the test channel under consideration.

- Anything else which helps to reproduce or appreciate the results.

VI. EXAMPLES

This section provides three example Watermark applications. The objective is to illustrate the use of Watermark, not to present novel modulation schemes or state-of-the art performances. Descriptions of the schemes are therefore kept short. The modulation format is essentially given by the direct-sequence spread spectrum (DSSS) scheme described in [15], including the non-spread case of a spreading code with length 1. Modifications to this format include fitting into the frequency bands of the test channels, and varying preamble lengths and message sizes.

A. Comparing schemes in NOF1 and NCS1

Table II specifies signals of two modulation schemes which are tested and compared in NOF1 and NCS1. The first signal is a single-carrier waveform with a quadrature phase-shift keyed (QPSK) symbol constellation. The second signal is DSSS, which is also single-carrier QPSK, but with a 15-chip spreading code for enhanced robustness. Both schemes employ a rate-1/2 convolutional code. Root raised-cosine pulses are used with a roll-off factor of 1/3 at a rate of 5600 s^{-1}.

Both signals carry a message of 256 information bits. They are equipped with a detection preamble and training symbols, which are a considerable overhead at this short message length. The effective bit rates are thus relatively low. The QPSK packet lasts only 0.21 s, which results in thousands of independent packets in NOF1 and NCS1.

The receivers use a filter matched to a bank of Doppler-shifted preamble replicas (with velocities spanning the range from -5 to $+5$ m/s) for detection, synchronization and Doppler estimation. The knowledge that NOF1 and NCS1 are channels between fixed stations is not used.

TABLE II

SIGNAL PARAMETERS.

	QPSK	DSSS
Carrier frequency	14 kHz	14 kHz
-3 dB bandwidth B	5.6 kHz	5.6 kHz
# information bits	256	256
Effective bit rate R_{eff}	1222 b/s	180 b/s
# packets in NOF1	5760	1260
# packets in NCS1	7920	1320

The chosen detection threshold results in about 1 false alarm per day in AWGN. Demodulation is performed with adaptive equalization. The SISO packet retrieval function is called from a receiver batch job cycling through a range of E_b/N_0 values. At each SNR, all available packets are retrieved with AWGN and a random delay, and fed to the receiver.

The packet error ratio (PER) is considered as the performance metric for this investigation. A packet error occurs when one or more bits are in error at the decoder output, or when the packet is not detected. The outcome of the simulations is shown in Fig. 10, leading to several observations:

- NCS1 is a more challenging channel than NOF.
- DSSS is more robust than QPSK.
- DSSS works well in both channels, but requires a 3–6 dB higher SNR in NCS1 to achieve the same performance as in NOF1.
- The performance of QPSK in NCS1 is severely limited by delay-Doppler spread.

Results can also be examined versus the SNR metric that is more common in underwater acoustics (ratio of signal power to noise power, measured in the frequency band of the signal) with the conversion $\text{SNR} = (R_{\text{eff}}/B) \times (E_b/N_0)$. This increases the gap between DSSS and QPSK, and confirms the notion that data communication at low SNR requires low-rate signaling.

The observation that a spreading code increases the robustness to channel distortions and noise is not surprising, but without realistic simulations it would be hard to quantify the improvements. A high-rate coherent scheme that delivers in NCS1 emerges as a challenge.
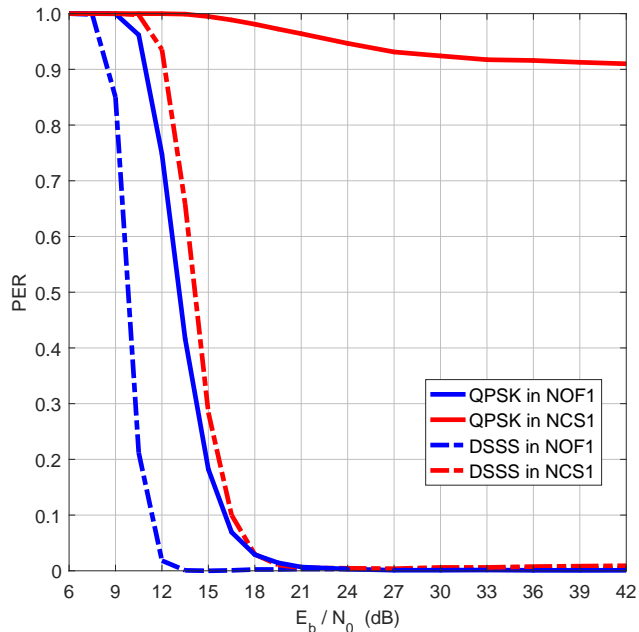
Fig. 10.　PER in NOF1 and NCS1 versus E_b/N_0.

B. Understanding system performance

The QPSK scheme of the previous section is adapted to KAU1 by lowering the carrier frequency to 6 kHz, and by lowering the symbol rate to 2800 s^{-1}, which reduces the -3 dB bandwidth to 2.8 kHz. The number of information bits is increased to 2048. The signal length becomes 1.4 s, and the effective data rate 1460 b/s.

Watermark returns 21 SIMO packets, which are fed to an adaptive multichannel equalizer in AWGN. The noise is spatially uncorrelated, and is added at a high (30 dB acoustic) SNR. The detector runs on the top hydrophone, whereas the receiver is applied to the top hydrophone, the top two hydrophones, the top four hydrophones, etc. One hydrophone was broken during the data collection, so the maximum number is 15.

Detection is performed with a Doppler-bank matched filter, cycling through velocities from -5 to $+5$ m/s in steps of 0.1 m/s. The filter is matched to the signal preamble, which is an m-sequence with a length of 255 chips.

The receiver output SNR is computed as one over the mean square symbol error, after removal of equalizer bias. The blue curve in Fig. 11 shows the output SNR and bit error ratio (BER),
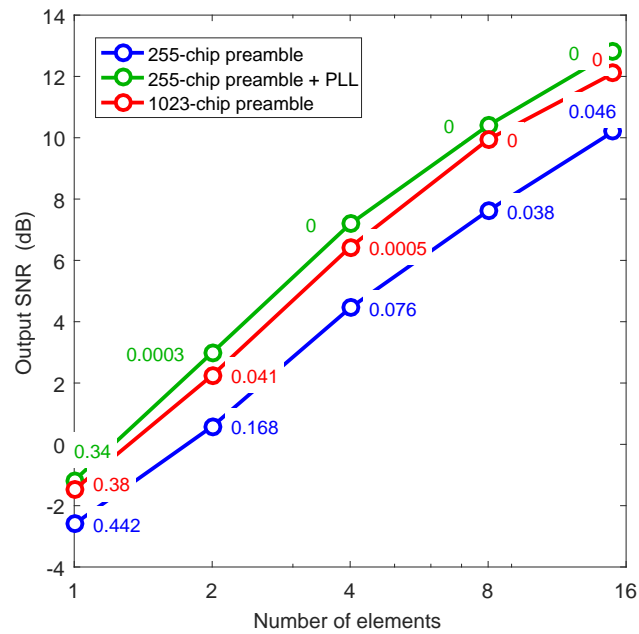
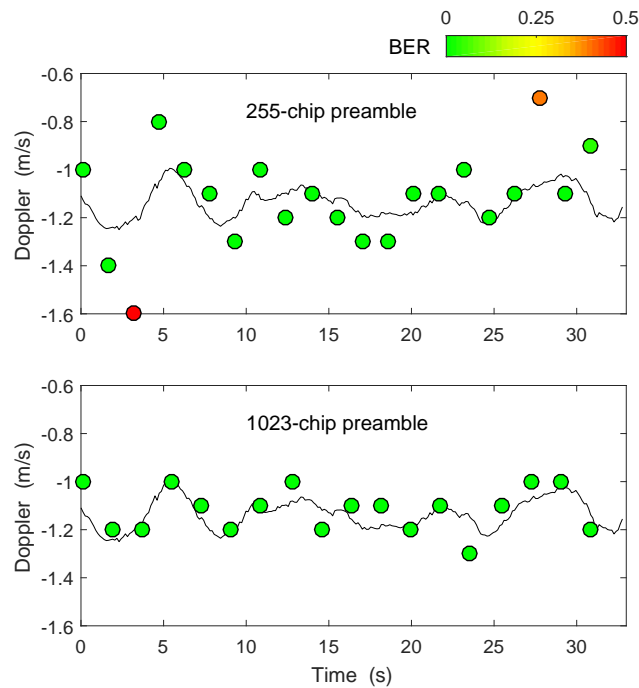Fig. 11. Receiver output SNR and BER, averaged over 21 (18) packets in KAU1.

Fig. 12. Preamble Doppler estimates, and BER for 15 elements. The black curve is the ground truth extracted from $\hat{h}(t, \tau)$.

averaged over the 21 Watermark packets. The mean BER is 0.44 for the top hydrophone. Adding elements improves the BER, but it does not drop to zero. Inspection of individual packets reveals that there are two packets with high BER, which do not improve at all upon adding elements. The cause is revealed in Fig. 12, which shows the detector Doppler estimate. A TVD estimate extracted from the known channel is included as the ground truth. The two BER outliers concern the two packets with the largest Doppler mismatch.

The receiver removes the detector Doppler estimate by resampling, and leaves the residual Doppler to the equalizer. It appears that a mismatch of ≈ 0.4 m/s is too large a burden for the equalizer, whose taps cannot keep track with the phase rotation of symbols.

Two strategies are explored to improve the performance. The first alternative is to improve the Doppler resolution by increasing the preamble length to 1023 chips. This requires re-running Watermark, and yields zero BER for all packets (Figs. 11, 12). A disadvantage is that the effective bit rate drops from 1460 to 1220 b/s, and that the complexity of the detector increases. The number of Watermark packets also drops, from 21 to 18. The second alternative is to equip the receiver with a PLL [24]. This low-complexity device deals explicitly with the phase rotation, and allows the equalizer to concentrate on path amplitude fluctuations and differential Doppler. Figure 11 shows that alternative 2 outperforms alternative 1. However, note that a long preamble would be required anyway for detection in the low-SNR regime. The two alternatives can also be combined, yielding a minor further improvement in the present high-SNR scenario (not shown).

C. Preamble detection in BCH1 and KAU2

In the previous section, Watermark was used to examine the effect of preamble Doppler resolution. In the present section, the benchmark is used to examine preamble detection statistics. Communication schemes deployed in different environments may require different decision statistics for optimal operation.

In underwater acoustic scenarios, the noise variance as well as the Doppler shift and the arrival time of preambles are unknown at reception. In this context, packet detection classically relies on a bank of constant false alarm rate (CFAR) matched filters [29, ch. 4] made of Doppler-shifted preamble replicas. The decision statistic is computed by combining the filter outputs. Several combination alternatives can be chosen, and, depending upon the channel properties, they may behave differently, so compromises must be found.

For single-hydrophone detection, a standard approach is to compare, at each instant, the maximum value of the matched filters output to a threshold (the maximum being taken over all Doppler branches). This approach, which considers each sample individually, implicitly assumes that the channel contains a single dominant path, which is not always the case as shown in Figs. 5 and 7. Another approach, more robust to multipath, is to consider the contribution of the different paths by averaging each matched filter output over a given time delay window [30, ch. 5.7]. The decision is then made by comparing the maximum value of these averages to a threshold. The optimal averaging duration requires a priori knowledge of the channel power delay profile, which is not available in practice. Therefore, only empirical choices can be made.

When several hydrophones are available at reception, the single-hydrophone statistics can be added up across the array to make a decision. However, if the space diversity offered by the array is known to be deficient because of correlation between impulse responses, other strategies may provide better performance. For instance, the space-time matrix obtained by stacking matched filter outputs can have a rank lower than the dimension of the array. If available, this prior information can be used to apply a multichannel detector of signals having a known rank [31], [32].

For the tests, the DSSS scheme of Sec. VI-A is adapted to BCH1 and KAU2 by changing the carrier frequency to 35 kHz and 6 kHz, respectively, and by lowering the chip rate to 2800 s^{-1}. A 511-chip m-sequence is used as a preamble and each packet carries 256 information bits. The packet duration is 2.925 s so that the effective bit rate becomes 87.5 b/s. Watermark returns 19 SIMO packets for BCH1 and 10 for KAU2, which are then fed to a preamble detector. Spatially uncorrelated WGN is added to the SIMO packets, where the SNR is defined as the ratio of average received signal power (across time and space) to noise power, measured in the -3 dB frequency band of the signal (2.8 kHz). The noise variance as well as the Doppler shift and the arrival time are assumed unknown at reception. The bank of Doppler-shifted replicas spans the velocity range from -5 to $+5$ m/s in 0.1-m/s steps. The detection thresholds are set to get 1 false alarm per day on average in AWGN. The results are averaged over all Watermark packets and 500 noise realizations are generated for each packet.

Fig. 13 illustrates the impact of time delay averaging on the correct detection rate for single-hydrophone detection in BCH1 (channel #2) and KAU2 (channel #8). While a 10-ms averaging window does not provide any performance improvement in BCH1, it is highly beneficial for
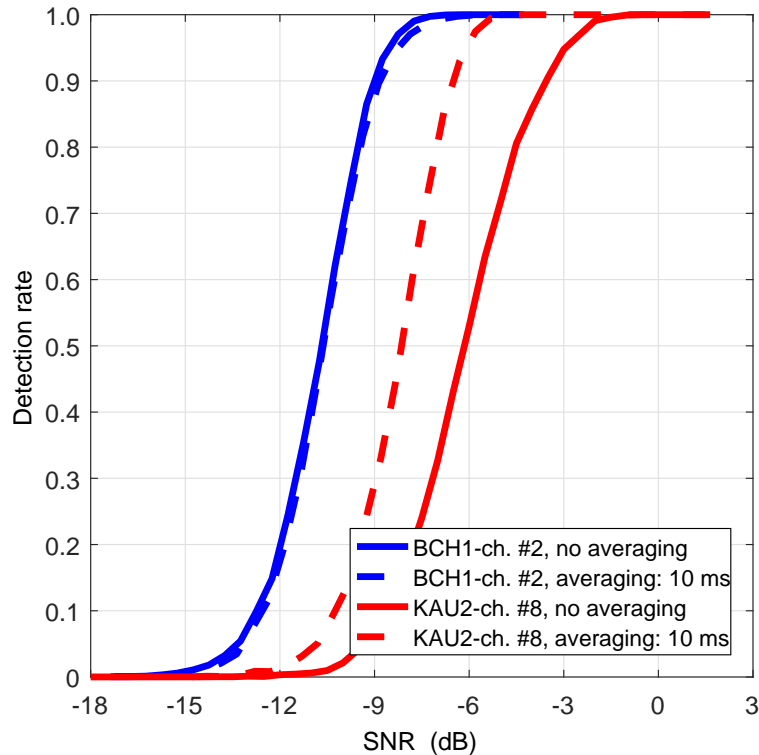
Fig. 13. Correct detection rate versus SNR for single-hydrophone detection in BCH1 (channel #2) and KAU2 (channel #8): impact of time delay averaging.

KAU2. The reason is that BCH1 contains one dominant path, whereas KAU has several significant arrivals with a spread in arrival times. The optimal averaging duration for both channels can be deduced from Fig. 14. This figure shows, for a correct detection rate set to 90%, the SNR gain obtained for several averaging durations compared with the case without time-delay averaging. For BCH1, this optimal value is less than 1 ms and averaging over more than 10 ms can be worse than not averaging at all. KAU2 is less sensitive to the averaging duration once it is greater than 5 ms. This is explained by the slower decay of the power delay profile of KAU2 compared with BCH1.

Fig. 15 shows the performance of multichannel detection for BCH1 with 4 hydrophones and KAU2 with 16 hydrophones, using 100 noise realizations at each SNR value. KAU2 outperforms BCH1, in contrast to Fig. 13, owing to a higher array gain. Several assumptions are made with respect to the rank of the space-time matrix obtained by aggregating the matched filters outputs
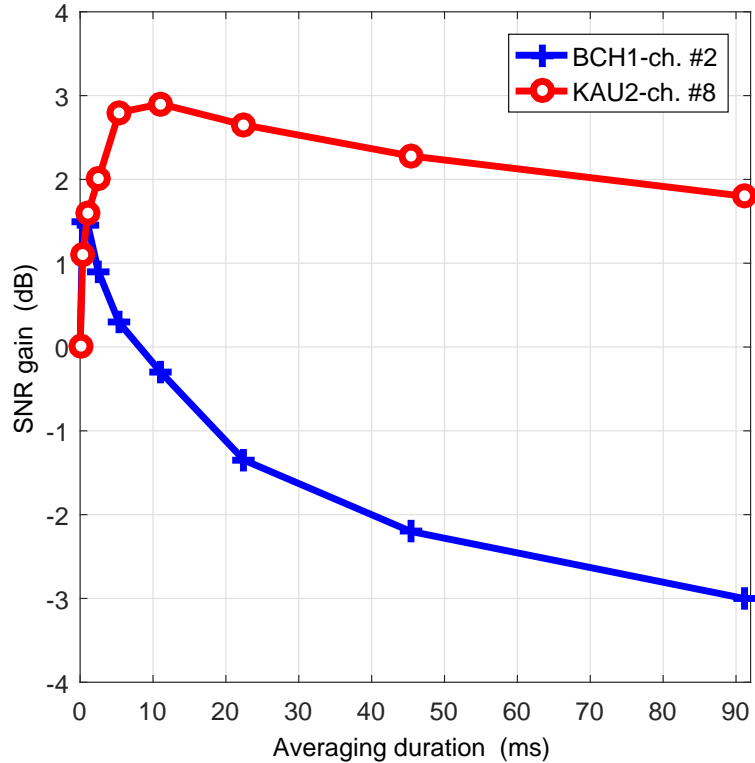
Fig. 14. SNR gain provided by averaging the matched-filters output of BCH1 (channel #2) and KAU2 (channel #8) over time delay. This gain is measured for a correct detection rate of 90%.

of the array. For BCH1, assuming a lower rank than the dimension of the array is beneficial for detection. For instance, a 1 dB improvement is achieved by limiting the rank to 2 instead of 4. As discussed in [33], although the hydrophone spacing is greater than 20 wavelengths, the correlation between hydrophones is quite high, which explains this performance improvement. KAU2 has much more spatial diversity than BCH1. Therefore, reducing the rank does not provide any gain.

D. Summary

The above examples illustrated applications of Watermark, and the insight it can provide. A given SISO scheme performed satisfactorily in one channel but not in another. A spreading code was required to ensure a low PER in the tougher channel, even at high SNR. The cause of a few persistent high-BER packets in a SIMO set-up was pinpointed to insufficient preamble Doppler resolution, and the scheme was improved correspondingly. Preamble detection statistics
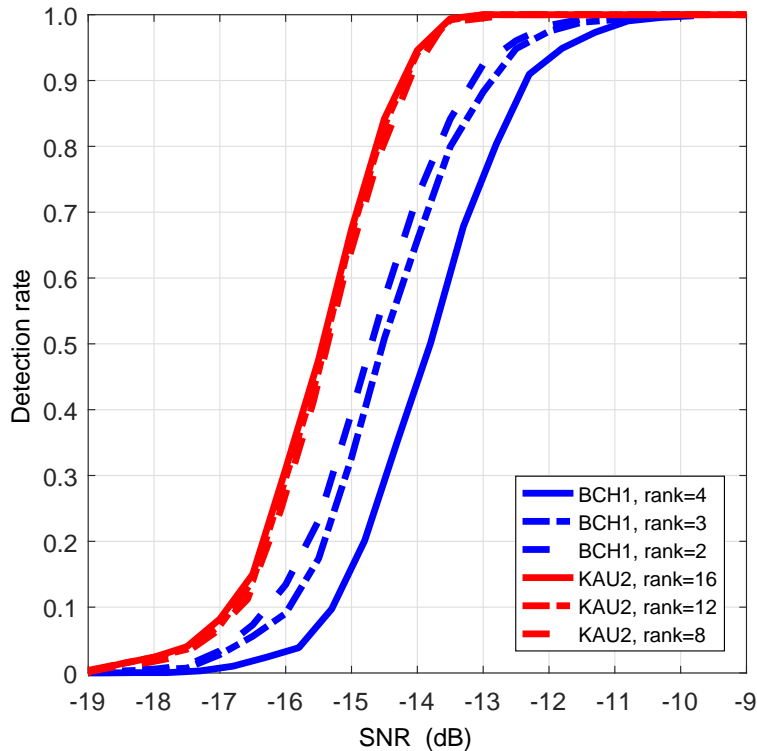
Fig. 15. Correct detection rate versus SNR for multiple hydrophone detection in BCH1 (4 channels) and KAU2 (16 channels): impact of rank limitation. For the simulation, matched filters outputs are aggregated over a 10-ms delay window.

were shown to vary between channels, and different methods to improve detection had different effects in different channels.

Watermark makes it easier to understand communication and detection performance, since the channels are known. Other differences with at-sea signaling are the possibility to improve schemes and re-transmit signals over the original channel, and the possibility to send multiple signals in parallel in the same frequency band.

VII. Future additions to Watermark

Watermark version 1 includes a set of channels with different parameters and characteristics, fulfilling several needs. Needs not covered by the first release require extensions. Possible extensions and additions are as follows.

The library of test channels can be extended by performing suitable soundings, using good acoustic measurements practice [34]. Extensions may concern different frequency bands, envi-

ronments, channel properties, deployment geometries, etc. It is recommended to use a flat probe signal spectrum, whose frequency band falls in a well-behaved region of the hardware transfer function. Correlative sounders are straightforward to use and have documented measurement errors [21], but other channel estimation algorithms may be used as well. More important than the type of sounder, is the requirement that measurement errors are small. Perhaps the most important requirement to a replay communication channel, is that communication performance in simulation is comparable to the communication performance at sea. It is therefore recommended to transmit different types of communication signals between sounding signals to enable validations like in [11]. Even validated acoustic propagation models may be used with Watermark, provided that their output is in the form of a TVIR at a suitably high sampling rate, and in the desired format [16].

More realistic noise models will add to the simulation fidelity in the regime of low SNR. Since the characteristics of ambient noise in the oceans vary widely [5], [26]–[28], different types of noise are needed to design robust communication systems. Acoustic noise models should be validated, but it is also possible to use measured noise, with a long measurement time to be able to draw a large number of independent realizations.

End-to-end calibration of sounding equipment and noise recorders will enable calibrated channel simulation resulting in simulated signals and noise at sound pressure levels corresponding to their ocean values. This could be used, for instance, to find the minimum source level that delivers adequate system performance in a given set-up.

VIII. Conclusions

A realistic benchmark is now available for underwater acoustic communications. It is based on a replay channel simulator driven by measurements of the time-varying impulse response. Its initial library of five test channels comprises four geographical areas and three frequency bands. Three of these test channels offer reception on a vertical line array. The primary use of Watermark is benchmarking of the physical layer of acoustic communication systems, but in principle it can also be used for different sonar applications involving the one-way transmission of sound.

Watermark brings the realism of at-sea signaling into the office while adding reproducibility. Applications include:

- Test and improve detection and communication algorithms.

- Compare different modulation schemes in the same channel.

- Examine the performance of a given scheme in different channels.

- Document the performance of modulation schemes and the effect of parameter settings, both for in-house evaluations and scientific publications.

- The quest for a robust high-rate system.

- Extract error statistics for network simulations.

- Study channel characteristics and signal fluctuations.

The benchmark can be extended with channels from different environments and frequency bands, either for own use or for general distribution, depending on the willingness of third parties to perform suitable measurements and share the data.

IX. ACKNOWLEDGMENTS

REFERENCES

[1] J. Preisig, "Acoustic propagation considerations for underwater acoustic communications network development," in *WUWNet'06*, Los Angeles, California, USA, September 2006.

[2] P. A. van Walree, "Propagation and scattering effects in underwater acoustic communication channels," *IEEE J. Ocean. Eng.*, vol. 38, no. 4, pp. 614–631, October 2013.

[3] A. Zielinksi, Y.-H. Yoon, and Lixue Wu, "Performance analysis of digital acoustic communication in a shallow water channel," *IEEE J. Ocean. Eng.*, vol. 20, no. 4, pp. 293–299, October 1995.

[4] C. Bjerrum-Niese and R. Lützen, "Stochastic simulation of acoustic communication in turbulent shallow water," *IEEE J. Ocean. Eng.*, vol. 25, no. 4, pp. 523–532, October 2000.

[5] M. Chitre, "A high-frequency warm shallow water acoustic communications channel model and measurements," *J. Acoust. Soc. Am.*, vol. 122, no. 5, pp. 2580–2586, November 2007.

[6] M. Siderius, M. B. Porter, P. Hursky, V. McDonald, and the KauaiEx Group, "Effects of ocean thermocline variability on noncoherent underwater acoustic communications," *J. Acoust. Soc. Am.*, vol. 121, no. 4, pp. 1895–1908, April 2007.

[7] P. A. van Walree, T. Jenserud, and M. Smedsrud, "A discrete-time channel simulator driven by measured scattering functions," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 9, pp. 1628–1637, December 2008.

[8] F.-X. Socheleau, C. Laot, and J.-M. Passerieux, "Stochastic replay of non-WSSUS underwater acoustic communication channels recorded at sea," *IEEE Trans. Sig. Proc.*, vol. 59, no. 10, pp. 4838–4849, October 2011.

[9] H. S. Dol, M. E. G. D. Colin, M. A. Ainslie, P. A. van Walree, and J. Janmaat, "Simulation of an underwater acoustic communication channel characterized by wind-generated surface waves and bubbles," *IEEE J. Ocean. Eng.*, vol. 38, no. 4, pp. 642–654, October 2013.

[10] P. Qarabaqi and M. Stojanovic, "Statistical characterization and computationally efficient modeling of a class of underwater acoustic communication channels," *IEEE J. Ocean. Eng.*, vol. 38, no. 4, pp. 701–717, October 2013.

[11] R. Otnes, P. A. van Walree, and T. Jenserud, "Validation of replay-based underwater acoustic communication channel simulation," *IEEE J. Ocean. Eng.*, vol. 38, no. 4, pp. 689–700, October 2013.

[12] P. van Walree, R. Otnes, and T. Jenserud, "Watermark: A realistic benchmark for underwater acoustic modems," in *2016 IEEE Third Underwater Communications and Networking Conference (UComms)*, Aug 2016, pp. 1–4.

[13] F.-X. Socheleau, A. Pottier, and C. Laot, "WATERMARK: BCH1 dataset description," Research report 17331, hal-01404491, Institut Mines-Telecom; TELECOM Bretagne, UMR CNRS 6285 Lab-STICC, 2016.

[14] W. S. Hodgkiss and J. C. Preisig, "Kauai Acomms MURI 2011 (KAM11) experiment," in *Proc. ECUA'12*, Edinburgh, UK, July 2012, pp. 993–1000.

[15] R. Otnes, P. A. van Walree, H. Buen, and H. Song, "Underwater acoustic network simulation with lookup tables from physical-layer replay," *IEEE J. Ocean. Eng.*, vol. 40, no. 4, pp. 822–840, October 2015.

[16] P. van Walree, R. Otnes, and T. Jenserud, "The Watermark manual and user's guide," FFI-rapport 2016/01378, Forsvarets Forskningsinstitutt, 2016.

[17] "The Watermark acoustic modem benchmark," www.ffi.no/watermark, Online.

[18] Y. Isukapalli, H. C. Song, and W. S. Hodgkiss, "Stochastic channel simulator based on local scattering function," *JASA Express Lett.*, vol. 130, no. 4, pp. EL200–EL205, October 2011.

[19] F.-X. Socheleau, C. Laot, and J.-M. Passerieux, "Parametric replay-based simulation of underwater acoustic communication channels," *IEEE J. Ocean. Eng.*, vol. 40, no. 4, pp. 796–806, October 2015.

[20] C. Petrioli, R. Petroccia, J. R. Potter, and D. Spaccini, "The SUNSET framework for simulation, emulation and at-sea testing of underwater wireless sensor networks," *Ad Hoc Networks*, vol. 34, pp. 224–238, 2015.

[21] G. Matz, A. F. Molisch, F. Hlawatsch, M. Steinbauer, and I. Gaspard, "On the systematic measurement errors of correlative mobile radio channel sounders," *IEEE Trans. Commun.*, vol. 50, no. 5, pp. 808–821, May 2002.

[22] J. G. Proakis and M. Salehi, *Digital Communications*, McGraw-Hill, fifth edition, 2008.

[23] P. van Walree, "Channel sounding for acoustic communications: Techniques and shallow-water examples," FFI-rapport 2011/00007, Forsvarets Forskningsinstitutt, 2011.

[24] M. Stojanovic, J. A. Catipovic, and J. G. Proakis, "Phase-coherent digital communications for underwater acoustic channels," *IEEE J. Ocean. Eng.*, vol. 19, no. 1, pp. 100–111, January 1994.

[25] P. A. van Walree and R. Otnes, "Ultrawideband underwater acoustic communication channels," *IEEE J. Ocean. Eng.*, 2013, DOI: 10.1109/JOE.2013.2253391.

[26] G. M. Wenz, "Acoustic ambient noise in the ocean: Spectra and sources," *J. Acoust. Soc. Am.*, vol. 34, no. 12, pp. 1936–1956, December 1962.

[27] P. H. Dahl, J. H. Miller, D. H. Cato, and R. K. Andrew, "Underwater ambient noise," *Acoust. Today*, vol. 3, no. 1, pp. 23–33, January 2007.

[28] G. B. Deane and J. C. Preisig, "Very high frequency noise sources in the littoral zone," in *2016 IEEE Third Underwater Communications and Networking Conference (UComms)*, August 2016, pp. 1–4.

[29] L. L. Scharf, *Statistical Signal Processing: Detection, Estimation, and Time Series Analysis*, Addison-Wesley, pp. 1 - 524, Reading, Massachusetts, 1991.

[30] S. M. Kay, *Fundamentals of Statistical Signal Processing, Volume II, Detection Theory, 14th printing*, Prenctice Hall, pp. 1-560, 2009.

[31] O. Besson, S. Kraut, and L. L. Scharf, "Detection of an unknown rank-one component in white noise," *IEEE Transactions on Signal Processing*, vol. 54, no. 7, pp. 2835–2839, July 2006.

[32] S. Sirianunpiboon, S. D. Howard, and D. Cochran, "Multiple-channel detection of signals having known rank," in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, May 2013, pp. 6536–6540.

[33] F. X. Socheleau, A. Pottier, and C. Laot, "Stochastic replay of simo underwater acoustic communication channels," in *OCEANS 2015 - MTS/IEEE Washington*, Oct 2015, pp. 1–6.

[34] J. R. Potter, M. B. Porter, and J. C. Preisig, "UComms: A conference and workshop on underwater communications, channel modeling, and validation," *IEEE J. Ocean. Eng.*, vol. 38, no. 4, pp. 603–613, October 2013.