Underwater Acoustic Communications: Design Considerations on the Physical Layer

Milica Stojanovic

Massachusetts Institute of Technology

Cambridge, MA 02139

millitsa@mit.edu

Abstract—Acoustic propagation is characterized by three major factors: attenuation that depends on the signal frequency, multipath propagation, and low speed of sound (1500 m/s). The channel has a sparse impulse response, where each physical path acts as a time-varying low-pass filter, and motion introduces additional Doppler spreading and shifting. Because propagation is best supported at low frequencies, acoustic communication systems are inherently wideband. The way in which these facts influence the design of signal processing methods is considered for single-carrier and multi-carrier systems. Moreover, the facts that the available bandwidth and transmission power depend heavily on the distance, and that channel latency is high, bear important implications on the design of network architectures and related protocols.

I. INTRODUCTION

Underwater acoustic channels are generally recognized as one of the most difficult communication media in use today. Acoustic propagation is best supported at low frequencies, and the bandwidth available for communication is extremely limited. For example, an acoustic system may operate in a frequency range between 10 and 15 kHz. Although the total communication bandwidth is very low (5 kHz), the system is in fact ultra-wideband, in the sense that bandwidth is not negligible with respect to the center frequency. Sound propagates underwater at a very low speed of 1500 m/s, and propagation occurs over multiple paths. Delay spreading over tens or even hundreds of milliseconds results in a frequency-selective signal distortion, while motion creates an extreme Doppler effect. The worst properties of radio channels–poor physical link quality of a mobile terrestrial radio channel and high latency of a satellite channel–are combined in an underwater acoustic channel.

As the history of underwater acoustic communications testifies, major advances in signal processing were made when the physical nature of propagation was respected through proper channel modeling. Examples that illustrate this fact include combined modeling of multipath and phase distortion for equalization in single-carrier wideband systems [1], a method used in a real-time acoustic modem [2]. More recently, detection of multi-carrier signals has been shown to benefit from explicit Doppler shift modeling, while sparse channel estimation, which recognizes the fact that underwater multipath is not contiguous but consists of isolated signal arrivals, is being used to improve the performance of both single-carrier and multi-carrier systems.

In this paper, we take a tutorial overview of the channel properties, aiming to reveal those aspects of acoustic propagation that are relevant for the design of communication systems. While a complete and accurate model of an underwater acoustic channel remains elusive, we focus on a simplified treatment that may provide a communications engineer with sufficient detail for effective system design. In particular, we take a fresh look at modeling a point-to-point channel response, beginning from first principles.

The paper is organized into three sections that address (1) attenuation and noise, (2) multipath propagation, and (3) the Doppler effect. Implications of acoustic propagation extend beyond the physical layer, and we conclude the paper by considering their impact on the design of future underwater networks.

II. ATTENUATION AND NOISE

Perhaps the most distinguishing property of acoustic channels is the fact that path loss depends on the signal frequency. This dependence is a consequence of absorption, i.e. transfer of acoustic energy into heat. In addition to the absorption loss, signal experiences a spreading loss which increases with distance. The overall path loss is given by

$$A(l, f) = (l/l_r)^k a(f)^{l-l_r} \tag{1}$$

where f is the signal frequency, and l is the transmission distance, taken in reference to some l_r. The path loss exponent k models the spreading loss, and its usual values are between 1 and 2 (for cylindrical and spherical spreading, respectively). The absorption coefficient $a(f)$ is an increasing function of frequency, which can be obtained using an empirical formula [3]. It may be interesting to note that $10 \log a(f) \approx \alpha_0 + \alpha_1 f + \alpha_2 f^2$ for frequencies up to about 50 kHz.

Noise in an acoustic channel contains the ambient noise and site-specific noise. The ambient noise, which is always present, may be modeled as Gaussian, but it is not white. Its power spectral density decays at approximately 18 dB/decade.

If we define a narrow band of frequencies of width Δf around some frequency f, the signal-to-noise ratio (SNR) in this band can be expressed as

$$SNR(l, f) = S_l(f)/A(l, f)N(f) \tag{2}$$

where $S_l(f)$ is the power spectral density of the transmitted signal, whose power may be adjusted according to the dis-

tance. For any given distance, the narrow-band SNR is thus a function of frequency, as shown in Fig.1. From this figure it becomes obvious that the acoustic bandwidth depends on the transmission distance. In particular, the bandwidth and the power needed to achieve a pre-specified SNR over some distance can be approximated as $B(l) = b \cdot l^{-\beta}$, $P(l) = p \cdot l^{\psi}$, where $\beta \in (0,1)$, and $\psi \geq 1$ [4]. The bandwidth is severely limited at longer distances: at 100 km, only about a kHz is available. At shorter distances, the bandwidth increases, but it will ultimately be limited by that of the transducer. The fact that bandwidth is limited implies the need for bandwidth-efficient modulation methods if more than a bps/Hz is to be achieved over these channels.

Another important observation to be made is that the acoustic bandwidth is centered at low frequencies. In fact, the acoustic bandwidth B is often on the order of its center frequency f_c, which makes an acoustic communication system inherently wideband. This fact in turn bears significant implications on the design of signal processing methods, as it prevents one from making the narrowband assumption ($B << f_c$), on which many radio communication principles are based. Respecting the wideband nature of the system is particularly important in multi-channel (array) processing [5] and, as we shall discuss, in synchronization for mobile acoustic systems.
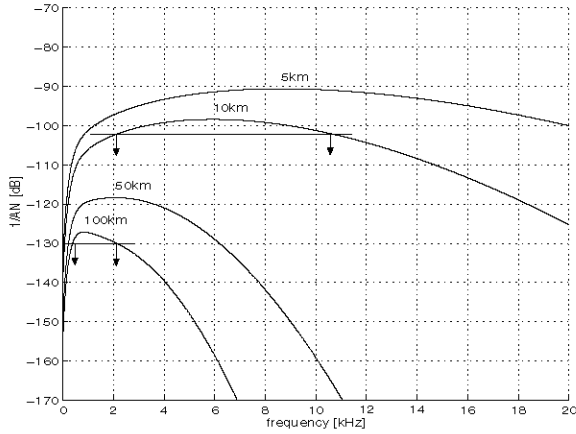

Fig. 1. Signal-to-noise ratio in an acoustic channel depends on the frequency and distance through the factor $1/A(l,f)N(f)$.

The fact that the acoustic bandwidth depends on the distance has important implications on the design of underwater networks. Specifically, it makes a strong case for multi-hopping, since dividing the total distance between a source and destination into multiple hops enables transmission at a higher bit rate over each (shorter) hop. The same fact helps to offset the delay penalty involved in relaying [6]. Since multi-hopping also ensures lower total power consumption, its benefits are doubled from the viewpoint of energy-per-bit consumption on an acoustic channel.

III. MULTIPATH

Multipath formation in the ocean is governed by two effects: sound reflection at the surface, bottom and any objects, and

sound refraction in the water. The latter is a consequence of sound speed variation with depth, which is mostly evident in deep water channels. Fig.2 illustrates the two mechanisms.
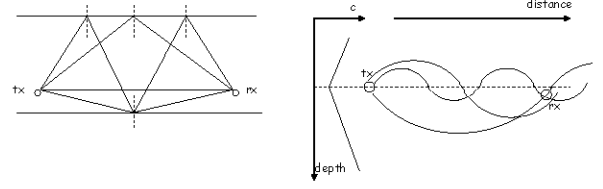

Fig. 2. Multipath formation in shallow and deep water.

The impulse response of an acoustic channel is influenced by the geometry of the channel and its reflection properties, which determine the number of significant propagation paths, their relative strengths and delays. Strictly speaking, there are infinitely many signal echoes, but those that have undergone multiple reflections and lost much of the energy can be discarded, leaving only a finite number of significant paths.

To put a channel model in perspective, let us denote by l_p the length of the p-th propagation path, with $p = 0$ corresponding to the first arrival (which is not necessarily the strongest). In shallow water, where sound speed can be taken as a constant c, path lengths can be calculated using plain geometry, and path delays can be obtained as l_p/c. In reference to the strongest path, whose arrival time is estimated to some value t_0 which determines the reference time at the receiver, we can also define relative path delays as $\tau_p = l_p/c - t_0$.

Surface reflection coefficient equals -1 under ideal conditions, while bottom reflection coefficients depend on the type of bottom (hard, soft) and the grazing angle [7]. If we denote by Γ_p the cumulative reflection coefficient along the pth propagation path, and by A_p the propagation loss associated with this path, then $h_p = \Gamma_p/\sqrt{A_p}$ represents the gain of this path. At this point, we may be tempted to express the channel response as

$$h(t) = \sum_p h_p \delta(t - \tau_p) \qquad (3)$$

as it is usually done for a radio channel [8]. However, as we have seen in Sec.II, the path gain in an acoustic channel is a function of frequency. Hence, it is constant only for a single frequency signal, i.e. a tone. For a broadband signal, each frequency will experience a different attenuation. Using the attenuation (1), we obtain the frequency response of the p-th path,

$$H_p(f) = \frac{\Gamma_p}{\sqrt{A(l_p, f)}} \qquad (4)$$

Hence, *each path* of an acoustic channel acts as a low-pass filter, introducing its own dispersion. Path dispersion is much less than the combined multipath spread of all paths, but a complete model must account for it nonetheless. Any approximations that may result from the general model will depend on the spectral occupancy of the signal.

The overall channel response in the frequency domain is

$$H(f) = \sum_p H_p(f)e^{-j2\pi f\tau_p} \tag{5}$$

and the corresponding impulse response can be expressed as

$$h(t) = \sum_p h_p(t - \tau_p) \tag{6}$$

where $h_p(t)$ is the inverse Fourier transform of $H_p(f)$.

To gain insight into the multipath effects, let us look at an example of a system with transmitter and receiver placed near the bottom at a depth of 75 m and separated by 3 km. The spreading factor is taken to be $k = 1.5$, and a 3 dB loss is associated with each bottom reflection. Fig.3 illustrates the results for the first six paths. Shown in the top row are the individual path transfer functions $H_p(f)$ and responses $h_p(t)$. The overall transfer function and response (magnitudes $|H(f)|$ and $|h(t)|$) are shown below.

The total multipath spread is governed by the longest path delay, which is on the order of tens of ms, a value that is typically observed in experiments. Marked by the asterisks are the values $h_p(0)$ at delays τ_p. On this scale, spreading on individual paths is not obvious, but it can be clearly seen from the path responses above. The shape of the path responses may at first sight resemble a Gaussian curve. This is not exactly so, but if one were to approximate the absorption coefficient as $e^{\alpha f^2}$, a Gaussian shape would follow. In any case, there obviously exists some cut-off frequency f_p which can be used to describe the bandwidth of each path. Corresponding to this frequency is the duration $1/f_p$ of each path's response. In the example shown, this frequency is on the order of 50 kHz. Hence, path dispersion cannot be neglected for signals of comparable bandwidth.

Except for scaling by the reflection coefficient, the responses of individual paths are very similar in our example, because the path lengths do not differ much. If one were to look at a system with shorter range but same depth, the path length difference would be more pronounced, causing the path responses to differ more noticeably.

Let us now consider transmission of a baseband signal $u(t)$ modulated onto a carrier of frequency f_c. The transmitted signal $s(t)$ and the received signal $r(t)$ are expressed as

$$s(t) = Re\{u(t)e^{j2\pi f_c t}\} \tag{7}$$
$$r(t) = Re\{v(t)e^{j2\pi f_c t}\} = s(t) * h(t) \tag{8}$$

The baseband received signal can be expressed as

$$v(t) = u(t) * c(t) \tag{9}$$

where

$$c(t) = h(t)e^{-j2\pi f_c t} \tag{10}$$

is the baseband response of the channel, taken with respect to the frequency f_c. This response can also be expressed as

$$c(t) = \sum_p c_p(t - \tau_p) \tag{11}$$
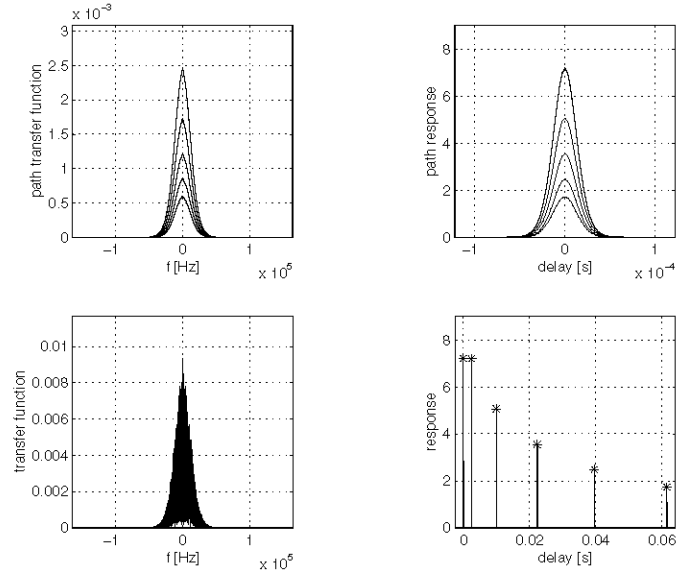

Fig. 3. Channel response functions.

where

$$c_p(t) = h_p(t)e^{-j2\pi f_c\tau_p}e^{-j2\pi f_c t} \tag{12}$$

is the baseband response of the pth path. The corresponding transfer functions are defined as

$$C(f) = H(f + f_c) = \sum_p C_p(f)e^{-j2\pi f\tau_p}$$
$$C_p(f) = H_p(f + f_c)e^{-j2\pi f_c\tau_p} \tag{13}$$

Note that $|c_p(t)| = h_p(t)$, and, hence, the duration of $c_p(t)$ equals that of the true response $h_p(t)$. Fig.4 illustrates the baseband response of the direct path, $c_0(t)$, for f_c=5, 10 and 15 kHz. Shown in the figure are the magnitude (solid), and the real and imaginary parts (dashed even and odd, respectively). Note that $c_p(t)$ will closely resemble $h_p(t)$ if $f_c << f_p$.
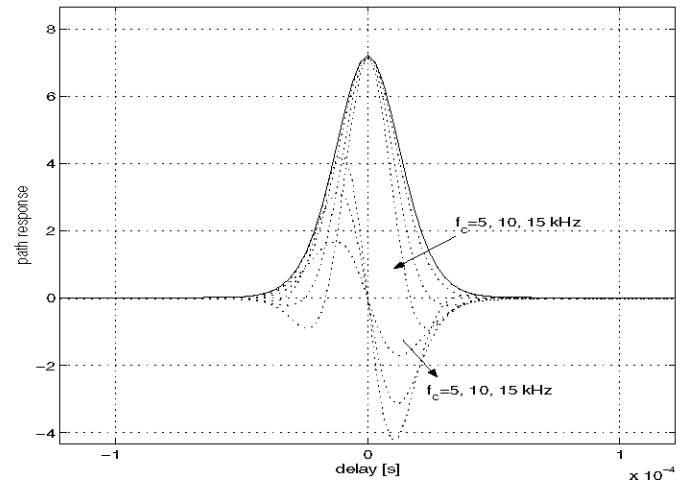

Fig. 4. Baseband path response for several values of the carrier frequency.

If the simplified channel model (3) were used, it would corresponds to a baseband discrete-path impulse response

$$c(t) = \sum_p c_p \delta(t - \tau_p) \qquad (14)$$

with $c_p = h_p e^{-j2\pi f_c \tau_p}$. An acoustic channel mandates a more complicated model (11), but a question naturally arises as to whether it is possible to simplify this model to match the discrete-path form (14). The answer to this question depends on the bandwidth occupancy of the transmitted signal. If the signal $u(t)$ changes slowly with respect to the response $c_p(t)$, i.e. if the signal bandwidth Δf is narrow enough that $C_p(f)$ remains approximately constant over $f \in [-\Delta f/2, \Delta f/2]$, then the received signal can be approximated as

$$v(t) = \sum_p c_p u(t - \tau_p) \qquad (15)$$

with

$$c_p = H_p(f_c) e^{-j2\pi f_c \tau_p} \qquad (16)$$

Hence, from the viewpoint of a narrowband signal, an *equivalent* baseband channel response is given in the form (14).

As we have seen, an acoustic communication signal is not likely to be narrowband. Nonetheless, so long as it is band-limited, say to some frequency region $f \in [-B/2, B/2]$, we can define an *equivalent* baseband response. This is the channel response that encompasses only those frequencies that are relevant for the signal, i.e.

$$\tilde{C}(f) = C(f) F_B(f) = \sum_p \tilde{C}_p(f) e^{-j2\pi f \tau_p} \qquad (17)$$

where $F_B(f) = 1$ for $f \in [-B/2, B/2]$ and 0 otherwise. The corresponding equivalent baseband response is

$$\tilde{c}(t) = c(t) * f_B(t) = \sum_p \tilde{c}_p(t - \tau_p) \qquad (18)$$

where $f_B(t) = B \text{sinc}(\pi B t)$ and

$$\tilde{c}_p(t) = \int_{-B/2}^{+B/2} H_p(f + f_c) e^{-j2\pi f_c \tau_p} e^{+j2\pi f t} df \qquad (19)$$

Fig.5 illustrates the response functions of the equivalent baseband channel model for the same system parameters as before, a carrier frequency of 15 kHz, and a bandwidth of 5 kHz. Shown in the top row are the path transfer functions and responses (magnitudes $|\tilde{C}_p(f)|$ and $|\tilde{c}_p(t)|$), while $|\tilde{C}(f)|$ and $|\tilde{c}(t)|$ are shown below. Marked by the asterisks are the values $|\tilde{c}_p(0)|$ at delays τ_p. Marked by the circles are the values $|c_p|$ that correspond to the narrowband model (16). Note that the two are *not* equal in general, although in this example they differ negligibly. The fact that they appear similar even though the signal is *not* narrowband is most easily explained by looking at the transfer functions $\tilde{C}_p(f)$. In our example, these functions exhibit an approximately odd symmetry around the zero frequency; hence, their integral (19) may be approximated by that of an equivalent flat function whose value equals $\tilde{C}_p(0)$ over $f \in [-B/2, B/2]$.

The equivalent baseband response is useful for obtaining a discrete-path channel model (one whose response can be represented in terms of delta functions). This is done by applying the sampling theorem to the band-limited channel response. One approach is to represent each path response as

$$\tilde{c}_p(t) = \sum_i \tilde{c}_p(i/B) \text{sinc} \pi B(t - i/B) \qquad (20)$$

This representation leads to the received signal in the form

$$v(t) = \sum_i \sum_p \frac{1}{B} \tilde{c}_p(i/B) u(t - i/B - \tau_p) \qquad (21)$$

The underlying discrete-path equivalent baseband response is

$$c'(t) = \sum_i \sum_p \frac{1}{B} \tilde{c}_p(i/B) \delta(t - i/B - \tau_p) \qquad (22)$$

Another approach is to represent the overall response as

$$\tilde{c}(t) = \sum_i \tilde{c}(i/B) \text{sinc} \pi B(t - i/B) \qquad (23)$$

The received signal can now be represented as

$$v(t) = \sum_i \frac{1}{B} \tilde{c}(i/B) u(t - i/B) \qquad (24)$$

and the underlying equivalent discrete-path response is

$$c''(t) = \sum_i \frac{1}{B} \tilde{c}(i/B) \delta(t - i/B) \qquad (25)$$

The two discrete-path channel models (22) and (25) are not the same, although they result in the same (band-limited) received signal. The first model has non-uniformly spaced coefficients (taps), which are the samples of $\tilde{c}_p(t)$, taken every $1/B$. Fig.6 illustrates these coefficients for our example. Note that only a few coefficients suffice to represent each path.

In the second model, the coefficients are uniformly spaced at $1/B$. However, because the path delays are not explicitly included in this model, more coefficients may be needed to capture the entire channel response. This fact is illustrated in Fig.7. Shown in this figure on top are the coefficients of the non-uniformly spaced model, which are centered around (continuous-valued) path delays τ_p, while below them are the coefficients of the uniformly-spaced model. The difference can be appreciated by looking at marked values $|\tilde{c}_p(0)|$ at delays τ_p.

As the signal bandwidth *increases*, so does the resolution at which the channel is represented, and the two models become more similar. The dominant coefficients of the uniformly-spaced model can then be associated with the physical propagation paths; however, we must keep in mind that their values are *not* to be identified with the coefficients (16) of the narrowband response; on the contrary, they must be computed over the wide range of frequencies using (19).
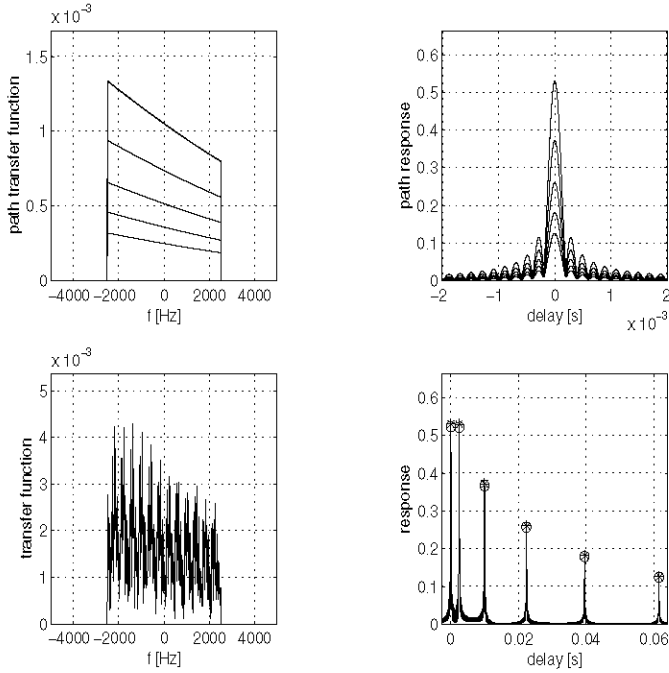
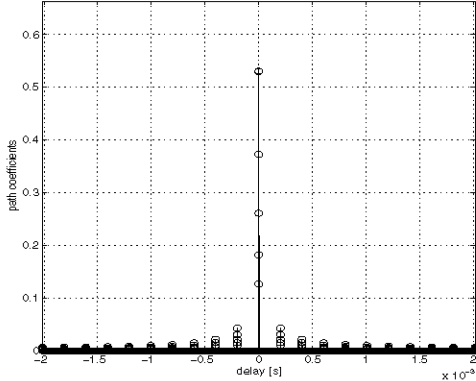Fig. 5. Equivalent baseband channel response functions.

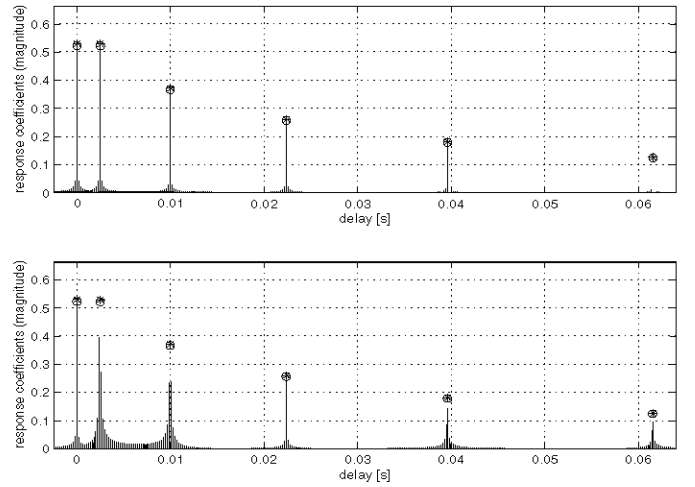Fig. 6. Path coefficients $c_p(i/B)$, $i = 0, \pm1, \ldots$.

Fig. 7. Equivalent discrete-path models with non-uniform (top) and uniform (bottom) tap spacing.

although the *total* delay span is large, only a few coefficients may suffice to represent the channel response. Channel modeling thus becomes an important aspect of equalization, and sparsing has been investigated for decision-feedback equalization [9], [10], and, more recently, for turbo equalization [11]. So far, sparse channel estimation has been considered in the context of the uniformly spaced model. It remains for the future to tell whether additional gains can be extracted (at a reasonable cost in complexity) by considering the non-uniform model. In pursuing these gains, fine details of signal design must be appreciated. In particular, the greater the bandwidth (and the symbol rate), the grater will be the ISI span, but the resolution will improve *both* in delay and in time. The former will enable more efficient sparsing (more similarity between the two discrete-path models), while the latter will enable faster adaptation (more frequent observations of the time-varying channel).

A. Single-carrier systems

Multipath dispersion of an underwater acoustic channel creates a frequency-selective signal distortion that must be equalized at the receiver. The fact that the multipath spread may be on the order of tens of ms or more implies that the inter-symbol interference (ISI) in a single-carrier broadband system may span tens or even hundreds of symbol intervals, a situation very different from that typically found in radio systems, where ISI may involve a few symbols only. Due to the length of ISI, maximum likelihood sequence detection is often abandoned in favor of computationally feasible equalization methods. Nonetheless, long adaptive equalizers may still be needed, which have considerable complexity, noise enhancement, and sensitivity to tracking constants. This problem can be alleviated by realizing that the equalizer coefficients are a function of the channel response, which is *sparse*. Namely,

B. Multi-carrier systems

A different approach to overcoming the frequency selective distortion is through multi-carrier modulation. Orthogonal frequency division multiplexing (OFDM) for acoustic channels has recently gained interest as the first experiments have shown successful high-rate transmission [13], [14]. Future efforts will undoubtedly rely on channel modeling, and this task should be approached keeping in mind the broadband nature of acoustic systems. Towards this goal, let us cast our current channel model into the framework of a broadband OFDM signal. Let us denote the signal transmitted on the k-th subcarrier of frequency $f_k = f_0 + k\Delta f$ as

$$s_k(t) = Re\{d_k g(t) e^{j2\pi f_k t}\} = Re\{u_k(t) e^{j2\pi f_0 t}\} \qquad (26)$$

where $g(t)$ is a unit-amplitude pulse of duration $T = 1/\Delta f$ and d_k is the signal (data symbol). The broadband signal,

consisting of K carriers occupying a bandwidth $B = K\Delta f$ is

$$s(t) = \sum_{k=0}^{K-1} s_k(t) = Re\{u(t)e^{j2\pi f_0 t}\} \qquad (27)$$

Let us assume that Δf is indeed small enough that $H_p(f) \approx H_p(f_k)$ for $f \in [f_k - \Delta f/2, f_k + \Delta f/2]$. Given the fact that the OFDM symbol duration T has to be greater than the multipath spread of the channel if the system is going to make efficient use of bandwidth, we have that Δf will be on the order of 100 Hz (or less) for a multipath spread of about 10 ms. For this Δf, the narrowband assumption is justified. Then, we have that the received signal is

$$r(t) = \sum_{k=0}^{K-1}\sum_{p} H_p(f_k)s_k(t - \tau_p) = Re\{v(t)e^{j2\pi f_0 t}\} \quad (28)$$

It may again be interesting to compare this model to the simpler case (3) which would give the received signal as

$$r(t) = \sum_{p} h_p s(t - \tau_p) \qquad (29)$$

For the two models to be equal, the coefficients $H_p(f_k)$ would need to be independent of k, i.e. the path transfer functions would need to be flat over the entire bandwidth B. However, as we have seen from Fig.3, this is not the case in a broadband acoustic system.

Equivalently in baseband, we have that

$$v(t) = \sum_{k=0}^{K-1} d_k \left[\sum_{p} H_p(f_k)e^{-j2\pi f_k \tau_p} g(t - \tau_p) \right] e^{j2\pi k\Delta f t} \tag{30}$$

Demodulation is performed so as to ensure that all the received signal components have been taken into account, yielding

$$y_k = \frac{1}{T} \int_{\hat{\tau}_0 - \epsilon}^{\hat{\tau}_0 + T' - \epsilon} v(t)e^{-j2\pi k\Delta f t}dt = H(f_k)d_k \qquad (31)$$

where $\hat{\tau}_0$ is the estimate of τ_0, ε is a safety margin, and $T' = T + T_g$, where $T_g \geq \hat{\tau}_{P-1} - \hat{\tau}_0 + 2\varepsilon$ is a guard interval that is sufficient to accommodate for the delay spreading. The signal at the output of the demodulator is thus represented in the usual form, $y_k = H_k d_k$, but the frequency-dependence remains in the terms

$$H_k = H(f_k) = \sum_{p} H_p(f_k)e^{-j2\pi f_k \tau_p} \qquad (32)$$

Channel estimation for an OFDM system can be performed in the frequency domain (direct estimation of the coefficients H_k) or in the time domain. The later targets estimation of the channel response $h(t)$, or, more precisely, the baseband equivalent taken with respect to f_0, $b(t) = h(t)e^{-j2\pi f_0 t}$. Namely, we can define the discrete Fourier pair

$$H_k = \sum_{l=0}^{K-1} b_l e^{-j2\pi kl/K} \qquad (33)$$

where the time-domain coefficients b_l are related to the samples of the baseband response $b(t)$ taken at intervals $1/B$.

Time domain channel estimation enables pilot-based detection of OFDM signals, but it can also be used to implement channel sparsing, which in turn will lead to an improved performance. Work is in progress on this issue for acoustic channels [15]. Similarly as with single-carrier systems, it remains to be seen whether explicit estimation of path gains *and* delays can provide further gains in performance.

C. Time variability

There are two sources of the channel's time variability: inherent changes in the propagation medium, and those that occur because of the transmitter/receiver motion. Inherent changes range from those that occur on very large time scales that do not affect the instantaneous level of a communication signal (e.g., monthly changes in the temperature) to those that occur on short time scales and affect the signal. Prominent among the latter are the changes induced by the surface waves, which effectively cause the displacement of the reflection point, resulting in both scattering of the signal, and Doppler spreading due to the changing path length.

It is out of the scope of the present treatment to summarize what is known about statistical characterization of these apparently random changes in the channel response. Suffice it to say that unlike in a radio channel, where a number of models for both the probability distribution (e.g., Rayleigh fading) and the power spectral density of the fading process (e.g., the Jakes' model) are well accepted and even standardized, there is no consensus on statistical characterization of acoustic communication channels. Experimental results suggest that some channels may just as well be characterized as deterministic, while others seem to exhibit Rice or Rayleigh fading [12]. Channel coherence times below 100 ms have been observed, but not often. For a general-purpose design, one may consider worst case coherence times on the order of seconds. In the absence of good statistical models for simulation, experimental demonstration of candidate communication schemes remains a *de facto* standard.

IV. THE DOPPLER EFFECT

Motion of the transmitter or receiver contributes additionally to the changes in the channel response. This occurs through the Doppler effect which causes frequency shifting as well as additional frequency spreading. The magnitude of the Doppler effect is proportional to the ratio $a = v/c$ of the relative transmitter/receiver velocity to the speed of sound. Because the speed of sound is very low as compared to the speed of electromagnetic waves, motion-induced Doppler distortion of an acoustic signal can be extreme. The only comparable situation in radio communications occurs in the Low Earth Orbiting satellite systems, where the relative velocity of satellites flying overhead is extremely high. (The channel there, however, is not nearly as dispersive.) Autonomous underwater vehicles (AUVs) move at speeds that are on the order of few m/s, but even without intentional motion, underwater instruments are subject to drifting with waves, currents and tides, which may occur at comparable velocities. In other words, there is always

some motion present in the system, and a communication system has to be designed taking this fact into account. The major implication of the motion-induced distortion is on the design of synchronization algorithms.

As the transmitter and receiver move relative to each other, the distance between them changes, and so does the signal delay. As a consequence, the leading edge of a transmitted signal may experience one delay, while the trailing edge will experience another. This situation is illustrated in Fig.8. Focusing on a single path, and neglecting the path dispersion, let us look at a single pulse $g(t)$ modulated onto a carrier of frequency f_c. For a constant velocity v, the received signal is

$$s'(t_0 + t) = s(t_0 + t - \frac{l(t_0) - vt}{c}) \qquad (34)$$

where $l(t_0)$ is the distance traveled by the signal arriving at t_0. Setting this time as the reference at the receiver, we have that $r(t) = s'(t_0 + t)$, i.e.,

$$r(t) = s(t + at - \tau) = Re\{g(t + at - \tau)e^{j2\pi f_c(t + at - \tau)}\} \qquad (35)$$

where $\tau = l(t_0)/c - t_0$. With respect to the center frequency f_c, the baseband received signal is

$$f(t) = e^{-j2\pi f_c \tau} g(t + at - \tau)e^{j2\pi a f_c t} \qquad (36)$$

Not counting the phase $2\pi f_c \tau$, this signal is distorted in two ways: first, it is scaled in time by $(1 + a)$, so that a transmitted pulse of duration T is observed at the receiver as having duration $T/(1 + a)$. Equivalently, its bandwidth B is observed as $(1 + a)B$. Second, a frequency offset $a f_c$ is introduced. The first type of distortion accounts for motion-induced Doppler spreading, while the second accounts for Doppler shifting.
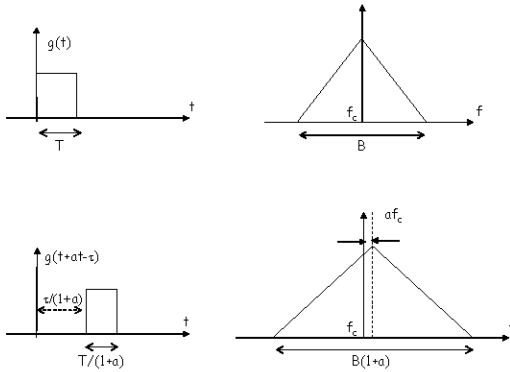

Fig. 8. Motion causes changes in the signal duration and frequency. The Doppler factor $a = v/c$ in an acoustic channel can be several orders of magnitude greater than in a radio channel.

The way in which these distortions affect signal detection depends on the actual value of the factor a. For comparison, let us look at a highly mobile radio system. At 160 km/h (100 mph), we have that $a = 1.5 \cdot 10^{-7}$. This value is low enough that Doppler spreading can be neglected. In other words, there is no need to account for it explicitly in symbol

synchronization. The error made in doing so is only 1/1000 of a bit per 10,000 bits. Hence, a simplified model can be adopted using an approximation $g(t + at) \approx g(t)$. In contrast to this situation, a stationary acoustic system may experience unintentional motion at 0.5 m/s (1 knot), which would account for $a = 3 \cdot 10^{-4}$. For an AUV moving at several m/s (submarines can move at much greater velocities), the factor a will be on the order of 10^{-3}, a value that cannot be ignored. In such a case, the approximation $g(t + at) \approx g(t)$ cannot be justified.

Non-negligible motion-induced Doppler spreading thus emerges as another major factor that distinguishes an acoustic channel from the mobile radio channel, and dictates the need for explicit delay synchronization in all but stationary systems. The approach that has demonstrated successful performance in single-carrier broadband acoustic systems is that of coupled equalization and synchronization [1]. In this approach, the phase offset, which account for the Doppler shift, is estimated jointly with the equalizer coefficients (i.e. the channel response). The so-obtained phase estimate can also be used to compute the Doppler factor, which is then used to resample the incoming signal, thus explicitly performing delay synchronization by decompressing the signal in time. The entire procedure is performed adaptively and in a closed loop. This procedure has been implemented in a real-time acoustic modem [2], and shown robust performance in a variety of conditions.

In multi-carrier systems, equalization is accomplished easily in the frequency domain, but the Doppler effect creates a particularly severe distortion. In radio systems, where time compression/dilation can be neglected, the only distortion remaining is the frequency offset. Since the system is narrowband, the Doppler shift appears as almost equal for all subcarriers. This fact greatly eases the task of synchronization, and many efficient synchronization algorithms have been developed for OFDM radio systems. In a wideband system, the situation is quite different. Here, each frequency f_k is shifted by an amount that cannot be approximated as equal for all subcarriers. The effect is that of an accordion, as illustrated in Fig.9. This figure offers an exaggerated view, but nonetheless one that serves to illustrate the fact that the Doppler effect in a wideband acoustic system causes *non-uniform* frequency shifting.
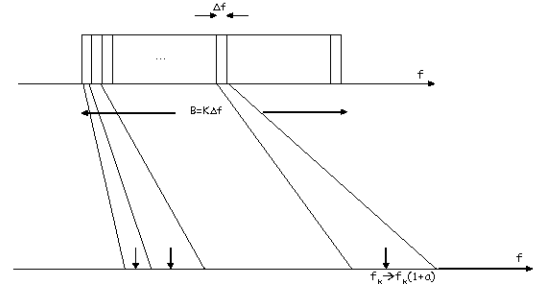

Fig. 9. Motion-induced Doppler shift is not uniform in a wideband system.

In a mobile acoustic system, the signal must first be re-sampled in order to reduce the Doppler factor to a reasonably low value. This can be done using an independently obtained estimate of the Doppler factor. Since large values are at hand to begin with, there is likely to be some error that cannot be neglected. If the original signal is characterized by a Doppler factor a', and resampling is performed using an estimate a'', then the resampled signal will be characterized by a residual factor $a = (a' - a'')/(1 + a'')$. Once the residual Doppler shift is much smaller than the subcarrier spacing, $af_k << \Delta f$, FFT demodulation can be performed.

In a multipath channel, there exists a possibility that each path will experience a different Doppler effect. To develop a channel model for this situation, one can follow the steps of Sec.III, including a residual Doppler factor a_p for the p-th path. Time variation can also be included, and we do so by assuming that the channel transfer functions $H_p(f)$ do not change much during one OFDM block, but may change from one block to another. We denote the values observed during the n-th OFDM block by $H_p(f_k, n)$. A similar approximation is made for the Doppler factors; namely, we assume that motion occurs at a constant velocity over one block, although it may change to a slightly different value in the next block. These assumptions are easily justified for block lengths on the order of tens of ms. Note, however, that their validity is compromised as the block length increases, and this ultimately limits the bandwidth efficiency that can be achieved using simple post-FFT processing on a time-varying acoustic channel.

To model the Doppler effect, let us express the signal transmitted on the k-th carrier as

$$s_k(t) = Re\{\sum_n d_k(n)g(t - nT')e^{j2\pi k\Delta f(t-nT')}e^{j2\pi f_0 t}\} \tag{37}$$

This signal travels over the channel in which the path length variation over an interval of time Δt within one block is modeled as

$$l_p(t_{n,p} + \Delta t) = l(t_{n-1,p}) - v_p(n)\Delta t \tag{38}$$

where $t_{n,p}$ is the time of arrival of the leading signal edge in block n, i.e.

$$t_{n,p} = t_{n-1,p} + \frac{T'}{1 + a'_p(n)} = t_{0,p} + T'\sum_{i=0}^{n-1}\frac{1}{1 + a'_p(i)} \tag{39}$$

The velocity $v_p(n)$, i.e. the Doppler factor $a'_p(n) = v_p(n)/c$, can be calculated from the channel geometry, i.e. the angle of signal arrival, and the actual motion of the transmitter, receiver and reflection points.

For the first incoming block, $n = 0$, we have that $t_{0,p} = l_p(t_{0,p})/c$. Corresponding to the strongest path is the time t_0, at which the receiver sets its local time to 0. In reference to this time, the relative initial path delays are $\tau_p(0) = t_{0,p} - t_0$.

The Doppler effect on the signal $s_k(t)$ is modeled as

$$s'_{k,p}(t_{n,p} + \Delta t) = s_k(t_{n,p} + \Delta t - \frac{l_p(t_{n,p}) - v_p(n)\Delta t}{c}) \tag{40}$$

for $\Delta t \in (0, T'/(1 + a'_p(n)))$. Setting $t_{n,p} + \Delta t = t_0 + t$, this signal can be expressed as

$$s'_{k,p}(t_0 + t) = s_k(t + a'_p(n)t - \tau_p(n) - a'_p(n) \cdot nT') \tag{41}$$

for $t \in (t_{p,n} - t_0, t_{p,n+1} - t_0)$, where the relative path delays are given by

$$\tau_p(n) = (1 + a'_p(n))[\tau_p(0) - \sum_{i=0}^{n-1}\frac{a'_p(i)}{1 + a'_p(i)}T'] \tag{42}$$

Substituting for the transmitted signal (37), the composite received signal is obtained as

$$\begin{aligned}
r'_k(t) &= \sum_p H_p(f_k, n)s'_{k,p}(t_0 + t) = Re\{\sum_n \sum_p d_k(n) \\
&\cdot g(t + a'_p(n)t - \tau_p(n) - a'_p(n) \cdot nT' - nT') \\
&\cdot e^{j2\pi k\Delta f(t + a'_p(n)t - \tau_p(n) - a'_p(n) \cdot nT' - nT')} \\
&\cdot e^{j2\pi f_0(t + a'_p(n)t - \tau_p(n) - a'_p(n) \cdot nT')}\}
\end{aligned} \tag{43}$$

and, summing all the carriers, we obtain $r'(t) = \sum_{k=0}^{K-1} r'_k(t)$.

Suppose now that we have an estimate a'' of the dominant Doppler factor. This estimate can be obtained from dedicated channel probes, or directly from the signal by estimating the duration (compression/dilation) of each block. When the system velocities remain approximately constant over several consecutive blocks (AUVs will not change their speed significantly over a few seconds), averaging can be performed to improve the estimate. Resampling[1] yields the signal

$$r(t) = r'(\frac{t}{1 + a''}) = Re\{\sum_{k=0}^{K-1} v_k(t)e^{j2\pi f_0 t}\} \tag{44}$$

that will be fed to the demodulator. Corresponding to k-th carrier of this signal is the component

$$\begin{aligned}
v_k(t) &= \sum_n \sum_p H_p(f_k, n)e^{-j2\pi f_k(\tau_p(n) + a'_p(n) \cdot nT')}d_k(n) \\
&\cdot g(t + a_p(n)t - \tau_p(n) - a'_p(n) \cdot nT' - nT') \\
&\cdot e^{j2\pi k\Delta f(t - nT')} \cdot e^{j2\pi f_k a_p(n)t}
\end{aligned} \tag{45}$$

where

$$a_p(n) = \frac{a'_p(n) - a''}{1 + a''} \tag{46}$$

is the residual Doppler factor of the p-th path.

Assuming that each received block is now confined to its own interval of duration T', FFT demodulation can be performed in the usual manner. Because of the residual Doppler effect, however, there will still be some shifting from one subcarrier to another, which causes inter-carrier interference (ICI). The demodulated signal in the k-th subband is thus a superposition

$$y_k(n) = \sum_l y_{kl}(n) \tag{47}$$

---

[1] We will assume that the the same value a'' is used for multiple blocks, although in general a different value $a''(n)$ can be used for each block.

where

$$y_{k,l}(n) = \frac{1}{T} \int_{\hat{\tau}_0(0)+nT'-\epsilon}^{\hat{\tau}_0(0)+nT'+T'-\epsilon} v_l(t)e^{-j2\pi k\Delta f(t-nT')}dt \quad (48)$$

In the presence of Doppler effect, the safety margin ϵ must account for additional delay spreading. Specifically, this will accumulate to

$$\epsilon_p(n) = a''\tau_p(0) - \sum_{i=0}^{n-1}\frac{a'_p(i)-a''}{1+a'_p(i)}T' \quad (49)$$

over n blocks, which should be a small fraction of T' under normal conditions (e.g., even for $a'_p(i)-a''$ on the order of 10^{-4} and, say, 100 blocks). Then, we have that

$$y_{k,l}(n) = d_l(n)\sum_p \rho^{(p)}_{k,l}(n)H_p(f_l,n)e^{-j2\pi f_k(1+a'')\tau_p(0)}e^{j\theta^{(p)}_k(n)} \quad (50)$$

where

$$\rho^{(p)}_{k,l}(n) = \frac{1}{1+a_p(n)}\text{sinc}\phi^{(p)}_{k,l}(n) \cdot e^{j\phi^{(p)}_{k,l}(n)}$$
$$\phi^{(p)}_{k,l}(n) = \frac{a_p(n)2\pi f_l + (l-k)2\pi\Delta f}{1+a_p(n)} \cdot \frac{T}{2} \quad (51)$$

The factor $\rho^{(p)}_{k,l}(n)$ models the ICI which results from the residual Doppler effect on the p-th path. Clearly, in order for the ICI not be destructive, we need $a_p(n)f_l << \Delta f, \forall l$. The demodulator in the k-th subband will then yield[2]

$$y_k(n) \approx d_k(n)\sum_p H_p(f_k,n)e^{-j2\pi f_k(1+a'')\tau_p(0)}e^{j\theta^{(p)}_k(n)}+z_k(n) \quad (52)$$

where any residual ICI is treated as independent additive noise $z_k(n)$, and the only distortion remaining is contained in the phase $\theta^{(p)}_k(n)$, which is given by

$$\theta^{(p)}_k(n) = 2\pi f_k\sum_{i=0}^{n-1}\frac{a'_p(i)-a''}{1+a'_p(i)}T' \quad (53)$$

For a constant a'', we also have that

$$\theta^{(p)}_k(n+1) = \theta^{(p)}_k(n) + 2\pi f_k\Delta a_p(n)T' \quad (54)$$

where

$$\Delta a_p(n) = \frac{a'_p(n)-a''}{1+a'_p(n)} \quad (55)$$

If the channel geometry is such that the range is much greater than depth (which often is the case) and if the Doppler effect is predominantly caused by the transmitter/receiver motion, i.e. not motion of a reflection point, then it is reasonable to assume that the Doppler factor is approximately the same for all the paths, $a_p(n) = a(n)$. The received signal then simplifies to the model [13],

$$y_k(n) \approx d_k(n)H_k(n)e^{j\theta_k(n)} + z_k(n) \quad (56)$$

[2]We are neglecting amplitude scaling by $1 + a_p(n)$.

This model clearly indicates the major distortion caused by the motion, which is the time-varying phase

$$\theta_k(n+1) = \theta_k(n) + \Delta a(n)2\pi f_kT' \quad (57)$$

i.e. the non-uniform Doppler shift $\Delta a(n)f_k$. Whereas the channel gain may vary in time as

$$H_k(n) = \sum_p H_p(f_k,n)e^{-j2\pi f_k(1+a'')\tau_p(0)} \quad (58)$$

it does so more slowly than the phase $\theta_k(n)$, which can change by $2\pi\Delta a(n)f_kT'$ from one block to another. For example, if $\Delta a(n) = 10^{-4}$, f_k=15 kHz, and T'=50 ms, the phase change will be $2\pi/10$.

In a conventional approach to OFDM signal detection, each block is detected independently. This can be done by allocating null subcarriers for frequency offset estimation, and pilot subcarriers for channel estimation. This approach takes no advantage of coherence between adjacent blocks, and it is preferred when the channel is varying rapidly [14]. Another approach is based on the model (56), (57). In this approach, $\Delta a(n)$ is estimated adaptively and used to compute the phases for all subcarriers, without reserving any null subcarriers. A *single* parameter is thus needed for synchronization of *all* K subcarriers, which may be many (e.g. 1024) in a bandwidth efficient acoustic system. This model-based approach has shown good performance with experimentally recorded data [13], [15].

At this time, no attempt has been made to develop receivers that accommodate the possibility of path-specific Doppler shifts. However, experimental recordings of acoustic channels demonstrate that such situations are possible, including those in which the Doppler shift has different *sign* on different paths [16]. In addressing receiver design for such channels, the model given by (52) and (54) may be a useful first step.

V. Design Considerations for Acoustic Networks

Implications of acoustic propagation extend beyond the physical layer, affecting all the layers of a network architecture. We have already seen that the bandwidth-distance dependence builds a strong case for underwater multi-hopping. In addition to the bandwidth, there are two major factors that influence the design of network protocols: the transmission power and the low speed of sound.

In an acoustic system, power required for transmitting is much greater than power required for receiving. Transmission power depends on the distance, and its typical values are on the order of tens of Watts.[3] In contrast, the power consumed by the receiver is much lower, with typical values ranging from about 100 mW for listening or low-complexity detection, to no more than a few Watts required to engage a sophisticated processor for high-rate signal detection. In sleep mode, from

[3]An acoustic signal propagates as a pressure wave, whose power is measured in Pascals (commonly, in dB relative to a micro Pascal). In seawater, 1 Watt of radiated acoustic power creates a sound field of intensity 172 dB re μPa 1 m meter away from the source.

which a node can be woken on command, no more than 1 mW may be needed.

Underwater instruments are battery-powered, and, hence, it is not simply the power, but the energy consumption that matters. This is less of an issue for mobile systems, where the power used for communication is a small fraction of the total power consumed for propulsion, but it is important for networks of fixed bottom-mounted nodes, where the overall network lifetime is the figure of merit.

One way to save the energy is by transmitting at a higher bit rate. (This is one more reason to investigate efficient processing methods for wideband signals.) For example, the WHOI modem [2] has two modes of operation: high rate at 5 kbps and low rate at 80 bps. This modem will require about 60 *times* less energy per bit (18 dB) in the high-rate mode. The receiver's energy consumption will also be lower, although it requires 3 W for detection of high-rate signals as opposed to 80 mW for detection of low-rate signals (the difference is about 2 dB).

Another way to save the energy is by minimizing the number of retransmissions. In random access networks, which are suitable for serving a varying number of users that transmit in a bursty manner, this task is made difficult by high channel latency. For example, the basic principle of carrier sensing multiple access – that a node should transmit only if it hears no on-going transmissions – is compromised in an acoustic channel where the packets propagate slowly, and the fact that none are overheard does not mean that some are not present in the channel. Multiple access with collision avoidance (MACA) has been used in the early acoustic network trials [17], and a number of variants have since been proposed [18], [19]. A different approach has been sought through the design of coordinated sleeping schedules for underwater nodes [20], [21]. Apart from the protocol design, it must be kept in mind that selection of power and bit rate will influence its performance: reducing the power reduces the level of interference; increasing the bit rate makes the packets shorter and reduces the chances of collision. Low speed of sound further challenges the throughput efficiency of any data link control scheme that requires automatic repeat request (ARQ), because current technology supports only half-duplex operation. Careful consideration of the physical layer parameters can help to design data packets so as to take maximal advantage of limited resources [22]. The implications on routing protocols are similarly important.

At this time, it is not certain in which direction the underwater networks will develop, as possible applications depend on the network capabilities, which are still developing, and the question of network capacity remains open. Both ad hoc networks, and infrastructure-based ones can be envisioned. In either case, acoustic propagation implies design principles that may be quite different from those used in radio networks (for a cellular underwater network, this issue is discussed in [23]), while the harshness of the environment dictates systems that are neither small nor easily deployable, and certainly not inexpensive or disposable [24].

REFERENCES

[1] M.Stojanovic, J.Catipovic and J.Proakis, "Adaptive multichannel combining and equalization for underwater acoustic communications," *Journal of the Acoustical Society of America*, vol.94 (3), Sept. 1993, pp.1621-1631.

[2] L. Freitag, M. Grund, S. Singh, J. Partan, P. Koski and K. Ball, "The WHOI micro-modem: An acoustic communications and navigation system for multiple platforms," in Proc. *IEEE Oceans Conf.*, 2005.

[3] L.Berkhovskikh and Y.Lysanov, *Fundamentals of Ocean Acoustics*, " Springer, 1982.

[4] M.Stojanovic, "On the relationship between capacity and distance in an underwater acoustic channel," in Proc. *First ACM International Workshop on Underwater Networks (WuwNet/MobiCom)*, 2006.

[5] M.Stojanovic, J.Catipovic and J.Proakis, "Reduced-complexity multichannel processing of underwater acoustic communication signals," *J. Acoust. Soc. Am.*, 98(2), Pt.1, pp.961-972, Aug. 1995.

[6] M.Stojanovic, "Capacity of a relay acoustic link," in Proc. *IEEE Oceans Conf.*, Oct. 2007.

[7] F.Jensen, W.Kuperman, M.Porter and H.Schmidt, *Computational Ocean Acoustics*, Springer Verlag, 1994.

[8] J.Proakis, *Digital Communications*, McGraw Hill, 1994.

[9] M.Stojanovic, L.Freitag and M.Johnson, "Channel-estimation-based adaptive equalization of underwater acoustic signals," in Proc. *IEEE Oceans Conf.*, Sept. 1999.

[10] W.Li and J.Preisig, "Estimation and equalization of rapidly varying sparse acoustic communications channels," in Proc. *IEEE Oceans Conf.*, Sept. 2006.

[11] S.Roy, T.Duman and V.McDonald, "Error rate improvement in underwater MIMO communications using sparse partial response equalization," in Proc. *IEEE Oceans Conf.*, 2006.

[12] M.Chitre, "A high-frequency warm shallow water acoustic communications channel model and measurements," *Journal of the Acoustical Society of America*, vol. 122 (5), Nov. 2007, pp. 2580-2586.

[13] M.Stojanovic, "Low complexity OFDM detector for underwater acoustic channels," in Proc. *IEEE Oceans Conf.*, Sept. 2006.

[14] B.Li, S.Zhou, M.Stojanovic, L.Freitag and P.Willet, "MIMO OFDM over an underwater acoustic channel," in Proc. *IEEE Oceans Conf.*, Oct. 2007.

[15] M.Stojanovic, "OFDM for underwater acoustic communications: adaptive synchronization and sparse channel estimation," submitted to *ICASSP*, 2008.

[16] J.Preisig, "Acoustic propagation considerations for underwater acoustic communications network development," in Proc. *First ACM International Workshop on Underwater Networks (WuwNeT/Mobicom)*, Sept. 2006.

[17] J.Rice, "SeaWeb acoustic communication and navigation networks," in Proc. *International Conference on Underwater Acoustic Measurements: Technologies and Results*, June 2005.

[18] M.Molins and M.Stojanovic, "Slotted FAMA: a MAC protocol for underwater acoustic networks," in Proc. *IEEE Oceans Conf.*, May 2006.

[19] B.Peleato and M.Stojanovic,"Distance aware collision avoidance protocol for ad-hoc underwater acoustic sensor networks," *IEEE Communication Letters*, to appear.

[20] V.Rodoplu and M.Park, "An energy-efficient MAC protocol for wireless acoustic networks," in Proc. *IEEE Oceans Conf.*, Sept. 2005.

[21] W.Ye, J.Heidemann and D.Estrin, " Medium access control with coordinated adaptive sleeping for wireless sensor networks," *IEEE/ACM Trans. Networking*, vol.12, No.3, June 2004, , pp.493-506.

[22] M.Stojanovic, "Optimization of a data link protocol for an underwater acoustic channel," in Proc. *IEEE Oceans Conf.*, 2005.

[23] M.Stojanovic, "Frequency reuse underwater: capacity of an acoustic cellular network," in Proc. *Second ACM International Workshop on Underwater Networks (WuwNet/MobiCom)*, 2007.

[24] J.Partan, J.Kurose and B.Levine, "A survey of practical issues in underwater networks," in Proc. *First ACM International Workshop on Underwater Networks (WuwNeT/Mobicom)*, Sept. 2006.