1

Introduction

1.1 Background and Context

1.1.1 Early Exploration of Underwater Acoustics

The Earth is a water planet, with two-thirds of the surface covered by water. Exploration of the mysterious underwater world has never ceased in human history. As early as 400 BC, Aristotle had noted that sound could be heard in water as well as in air. In AD 1490, Leonardo da Vinci wrote: "If you cause your ship to stop and place the head of a long tube in the water and place the other extremity to your ear, you will hear ships at great distances" [268]. In 1826, Charles Sturm and Daniel Colladon made the first accurate measurement of sound speed in water at Lake Geneva, Switzerland. The first practical application of underwater sound appeared in the 1900s: the underwater bells equipped on lightships were simultaneously sounded with a fog horn to measure the offshore distance based on the difference of the airborne and waterborne arrivals, and meanwhile the stereo headphones were also used for directions [397]. With the sinking of Titanic in 1912, L. F. Richardson successively filed a patent of echo ranging with sound in air and a patent application of echo ranging in water.

Along with the application of submarine and underwater mines in World War I (1914–1918), considerable progress has been made in underwater acoustics, especially on the underwater echo ranging for submarine and mine detection. In 1914, Constantin and Chilowski conceived the idea of submarine detection by underwater echo ranging. Based on the discovery of the piezoelectric effect by Jacques Curie and Pierre Curie in 1880, Paul Langevin in 1918 used quartz (piezoelectric) transducers as source and receiver to extend one-way sound transmission to 8 km, and for the first time observed clear echoes from a submarine at distances as large as 1500 m. Between World War I and World War II, scientists started to understand some fundamental concepts of sound in water, such as sound refraction due to changes of water temperature, salinity and pressure. Development of underwater sound applications during this period can be found in echo ranging for commercial use, underwater tomography and fisheries acoustics. The research effort on underwater acoustics during World War II (1941–1945) was mainly focused on improving echo ranging systems which were later coined as "sonar" (for SOund Navigation And Ranging). During this period, topics relative to sonar system performance were extensively investigated, including the high-frequency acoustics, low-frequency sound propagation, ambient noise, etc. By the end of World War II, the underwater sound had

OFDM for Underwater Acoustic Communications, First Edition. Shengli Zhou and Zhaohui Wang.

been primarily used for navigation and threat-finding. In 1945, an underwater telephone, which was developed by the Navy Underwater Sound Laboratory in the United States for the purpose of communication with submerged submarines, was the first application of underwater sound for communications [321]. Since then, development on underwater acoustic communications has been made in various underwater acoustic applications.

1.1.2 Underwater Communication Media

To establish communications among underwater assets and systems floating on the surface, four different communication media have been used.

- Cables. There have been many cabled observatories established over the years. Cables provide robust communication performance; however, the deployment and maintenance cost is very high. This motivates the use of wireless data transmission.
- Acoustic waves. For underwater wireless communication systems, acoustic waves are used as the primary carrier due to their relatively low absorption in underwater environments. However, acoustic waves have low propagation speed and a very limited frequency band.
- Electromagnetic (EM) waves. The use of EM waves in the radio frequency band has several advantages over acoustic waves, mainly faster velocity and high operating frequency (resulting in higher bandwidth). The key limitation of using EM waves for underwater communication is the high attenuation due to the conductive nature of seawater [255].
- Optical waves. Using optical waves for communication obviously has a big advantage in data rate. However, there are a couple of disadvantages for optical communication in water. Firstly, optical signals are rapidly absorbed in water. Secondly, optical scattering caused by suspended particles and plankton is significant. Thirdly, the high level of ambient light in the upper part of the water column is another adverse effect for using optical communication.

Apparently, each of the three physical waves as wireless information carrier has its own advantages and disadvantages. For a more intuitive comprehension, we summarize the major characteristics of acoustic, electromagnetic and optical carriers in Table 1.1. Acoustic waves propagate well in seawater and can reach a far distance. This justifies using acoustic waves for most underwater wireless communications.

Table 1.1 Comparison of acoustic, EM and optical waves in seawater environments

	Acoustic	Electromagnetic	Optical
Nominal speed (m/s)	~ 1500	$\sim 33\,333\,333$	$\sim 33\,333\,333$
Power loss	relatively small	large	$\propto$ turbidity
Bandwidth	$\sim$ kHz	$\sim$ MHz	~ 10–150 MHz
Frequency band	$\sim$ kHz	$\sim$ MHz	$\sim 10^{14}$–10^{15} Hz
Antenna size	~ 0.1 m	~ 0.5 m	~ 0.1 m
Effective range	$\sim$ km	~ 10 m	~ 10–100 m

Source: Liu 2008 [255], Table 2, p. 984. Reproduced with permission of Wiley.

1.1.3 Underwater Systems and Networks

Along with the tremendous scientific and technology advances in last several decades, a wide range of underwater exploration and applications have emerged. The scientific exploration spans across multiple disciplines, such as physical oceanography, marine biology, and deep sea archaeology (e.g., discovery of the wreck of the *Titanic*). Environmental applications involve studies in pollution monitoring, climate change, and global warming. Commercial applications of underwater technologies can be found in, e.g., offshore oil/gas field monitoring, fishery industries, and treasure discovery. Military applications of underwater technologies include tactical surveillance in coastal areas, harbors and ports etc.

In recent years, development of underwater vehicles of various sizes and capabilities, such as sea gliders and autonomous underwater vehicles (AUVs), has enabled underwater applications without human interaction. For example, sea gliders can be deployed in lakes or oceans to collect data samples of water over a large time period, and then send the data back to a control center for scientific studies. A fleet of underwater vehicles can form an underwater network, in which vehicles can collaborate to accomplish predetermined tasks. As more intelligent systems are deployed in underwater applications, the need of communications and networking keeps growing.

1.2 UWA Channel Characteristics

Given the complexity of underwater acoustic medium and the low propagation speed of sound in water, the underwater acoustic channel is commonly regarded as one of the most challenging channels for communication. Next we will look into several distinguishing characteristics of underwater acoustic channels. Comparisons between the underwater acoustic channel and the terrestrial radio channel are made along with the descriptions of underwater acoustic channel characteristics.

1.2.1 Sound Velocity

The extremely slow propagation speed of sound through seawater is an important factor that differentiates it from electromagnetic propagation. The speed of sound in water depends on the water properties of temperature, salinity and pressure; illustrative plots of the three parameters as functions of water depth are shown in Figure 1.1 [305, Chap. 9]. A typical speed of sound in water near the ocean surface is about 1520 m/s, which is more than 4 times faster than the speed of sound in air, but five orders of magnitude smaller than the speed of light. The speed of sound in water grows with increasing water temperature, increasing salinity and increasing depth. Approximately, the sound speed increases 4.0 m/s for water temperature rising 1°C. When salinity increases one practical salinity unit (PSU), the sound speed in water increases to 1.4 m/s. As the depth of water (therefore also the pressure) increases to 1 km, the sound speed increases roughly to 17 m/s. It is noteworthy to point out the above assessments are only for rough quantitative or qualitative discussions, and the variations in sound speed for a given property are not linear in general.

A typical sound speed profile as a function of depth in deep water, is shown in Figure 1.2 [305, Chap. 9]. Depending on the depth, the profile can be divided into four layers.

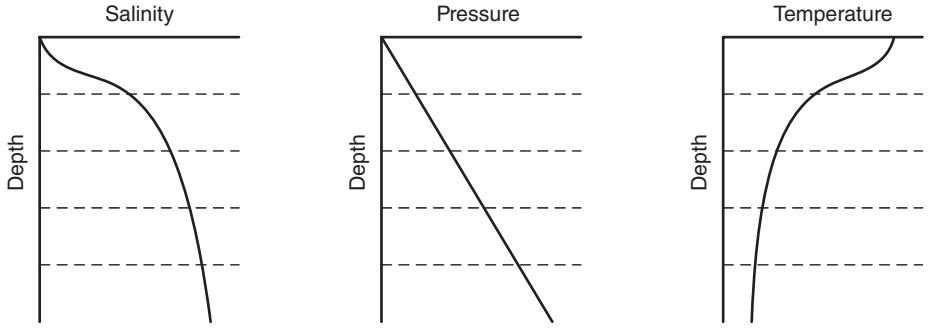

Figure 1.1 Variations of environmental parameters as functions of depth.

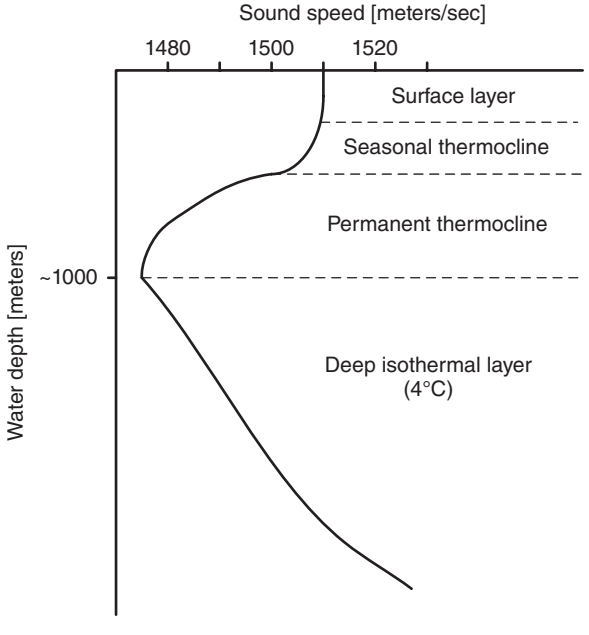

Figure 1.2 Typical sound speed profile in deep water.

- *Surface layer.* The surface layer usually has a water depth of a few tens of meters. Due to the mixing effect of wind, both temperature and salinity in this layer tend to be homogeneous, which leads to a constant sound velocity. This layer is also called a *mixed layer*.
- *Seasonal and permanent thermocline layers.* In the *thermocline* layers, the water temperature decreases as the water depth grows; as illustrated in Figure 1.1. In these two layers, the effect of increases in pressure and salinity cannot compensate the effect of temperature decrease. Therefore, there is a negative gradient of the sound speed profile in depth. In the seasonal thermocline layer, the negative gradient varies with seasons, while it is less seasonal in the permanent thermocline layer.

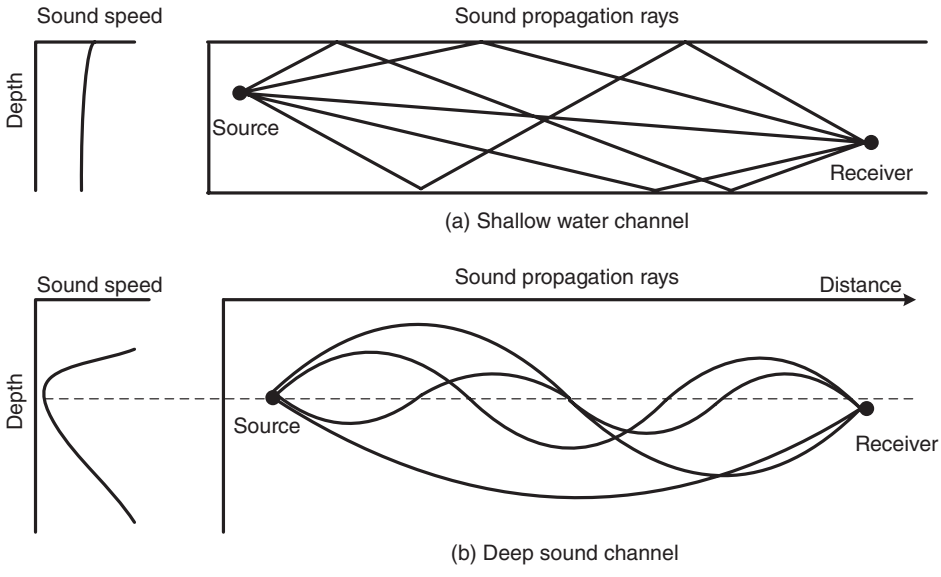

Figure 1.3 Multiple propagation paths.

- *Deep isothermal layer.* The water temperature is nearly constant around 4°C. The sound speed is therefore mainly determined by the water pressure, which leads to a positive gradient of sound speed in depth.

According to Snell's law, a ray of sound bends toward the direction of low propagation speed. In shallow water, the sound speed is usually constant throughout the water column. The acoustic signal usually propagates along straight lines, as illustrated in Figure 1.3(a). The sound speed profile of deep water channels diversifies the sound propagation paths. In particular notice that there is a minimal sound speed at a particular water depth (named the *channel axis*) between the permanent thermocline layer and the deep isothermal layer. For an acoustic signal transmitted at the channel axis, a ray of sound will be bent downward when propagating to the permanent thermocline layer and bent upward when propagating to the isothermal layer, thus being trapped within the two layers without interacting with the sea surface and bottom, as illustrated in Figure 1.3(b). This type of channel is called the *deep sound channel*, and the corresponding propagation is called SOFAR (for SOund Fixing And Ranging). An interesting phenomenon of SOFAR propagation is that a path traveling a longer distance could have a shorter travel time. Due to the refraction caused by inhomogeneous sound speed, there exist both shadow zones and convergence zones in the acoustic field, where a *shadow zone* denotes an area which cannot be penetrated by direct sound paths, and a *convergence zone* denotes an area which is insonified intensively by a bundle of sound paths.

1.2.2 Propagation Loss

There are three primary mechanisms of energy loss during the propagation of acoustic waves in water: (i) absorptive loss, (ii) geometric spreading, and (iii) scattering loss. Note that the

discussions in this section are based on empirical approaches, while discussions based on propagation models will be provided in Section 1.2.4.

1.2.2.1 Frequency-Dependent Absorption

During propagation, wave energy may be converted to other forms and absorbed by the medium. The absorptive energy loss is directly controlled by the material imperfection for the type of physical wave propagating through it. For EM waves, the imperfection is the electric conductivity of seawater. For acoustic waves, this material imperfection is the inelasticity, which converts the wave energy into heat.

The absorptive loss for acoustic wave propagation is frequency-dependent, and can be expressed as $e^{\alpha(f)d}$, where d is the propagation distance and $\alpha(f)$ is the absorption coefficient at frequency f. For seawater, the absorption coefficient at frequency f in kHz can be written as the sum of chemical relaxation processes and absorption from pure water [4, 266]:

$$\alpha(f) = \frac{A_1 P_1 f_1 f^2}{f_1^2 + f^2} + \frac{A_2 P_2 f_2 f^2}{f_2^2 + f^2} + A_3 P_3 f^2, \tag{1.1}$$

where the first term on the right side is the contribution from boric acid, the second term is from the contribution of magnesium sulphate, and the third term is from the contribution of pure water; A_1, A_2, and A_3 are constants; the pressure dependencies are given by parameters P_1, P_2 and P_3; and the relaxation frequencies f_1 and f_2 are for the relaxation process in boric acid and magnesium sulphate, respectively. Please refer to [266, Chap. 2] for formulations of the coefficients A_1, A_2, A_3, P_1, P_2, P_3, f_1, and f_2 as functions of temperature, salinity and water depth.

In underwater acoustic communications, Thorp's formula can be used as a simplified absorption model for frequencies less than 50 kHz,

$$\alpha(f) = \frac{0.11 f^2}{1 + f^2} + \frac{44 f^2}{4100 + f^2} + 2.75 \times 10^{-4} f^2 + 0.003 \tag{1.2}$$

where f denotes frequency in kHz [51, 195].

1.2.2.2 Geometric Spreading Loss

Geometric spreading is the local power loss of a propagating acoustic wave due to energy conservation. When an acoustic impulse propagates away from its source with longer and longer distance, the wave front occupies larger and larger surface area. Hence, the wave energy in each unit surface (also called *energy flow*) becomes less and less. For the spherical wave generated by a point source, the power loss caused by geometric spreading is proportional to the square of the distance. On the other hand, the cylindrical waves generated by a very long line source, the power loss caused by geometric spreading is proportional to the distance. For a practical underwater setting, the geometric spreading is a hybrid of spherical and cylindrical spreading, with the power loss to be proportional to d^β, where β is between 1, for cylindrical spreading, and 2, for spherical spreading [397]. Provided that the sound propagation in real channels can hardly be classified into either of the two spreading models, a practical value of the spreading exponent can be taken as $\beta = 1.5$. Note that geometric spreading is frequency-independent.

1.2.2.3 Scattering Loss

Scattering is a general physical process in which the incident wave is reflected by irregular surfaces in many different directions. The sound scattering in underwater environments can be attributed to the nonuniformities in the water column and interactions of acoustic waves with nonideal sea surfaces and bottoms. Obstacles in the water column include point targets such as fish and plankton, and scattering volumes such as fish schools and bubble clouds. The corresponding scattering loss depends on the acoustic wavelength and target size. In particular, the scattering loss increases as the acoustic wavelength decreases. The scattering property of sea surface and bottom is mainly determined by the interface roughness. High interface roughness induces large spatial energy dispersion. The roughness of sea surface is due to the capillary waves caused by wind, the amplitude of which ranges from centimeters to meters (e.g., swells). The roughness of sea bottom depends on the geology, including e.g., the roughness of rocks, sand ripples, and organisms in sediments. The amplitude of the roughness also varies from centimeters to meters. Similar to the target scattering in the water column, scattering loss at sea surface and bottom is also frequency-dependent.

In real environments, the two types of scattering processes coexist. For example, in the presence of a high wind speed, the wind-generated waves increases the roughness of sea surface, and breaking waves can create bubble clouds of a large size. Both types of scattering losses happen when the acoustic wave interacts with both sea surface and bubble clouds. Moreover, the wind-generated waves become moving reflectors of acoustic waves, thus introducing energy dispersion not only in the spatial domain but also in the frequency domain.

1.2.2.4 Propagation Loss Parametrization

Denote ξ as the scattering loss. For an acoustic signal at frequency f, the attenuation after propagating a distance of d can be formulated as

$$P_{\text{att}}(f, d) = \xi d^{\beta} e^{\alpha(f)d}. \tag{1.3}$$

Different from the propagation loss in the terrestrial radio channel which only has spreading loss with an exponent $2 \sim 6$ [145], the propagation loss in underwater acoustic channels is spreading-loss dominant in the near-distance transmission and absorption-loss dominant in the long-distance transmission.

1.2.3 Time-Varying Multipath

An acoustic wave can reach a certain point through multiple paths. In a shallow water environment, where the transmission distance is larger than the water depth, wave reflections from the surface and the bottom generate multiple arrivals of the same signal. In deep water applications, surface and bottom reflections may be neglected. However, the wave refractions due to the spatially varying sound speed cause significant multipath phenomena.

Assume that there are N_{pa} paths, and let ξ_p denote the scattering loss, d_p the propagation distance and τ_p the propagation delay of the pth path. Then the pass loss along the pth path can be written as

$$P_{\text{att}}(f, d_p) = \xi_p d_p^{\beta} e^{\alpha(f)d_p} \tag{1.4}$$

which combines the effects of spreading loss, absorptive loss, and scattering loss. For a channel which is *time-invariant* within a certain time interval, the channel transfer function at frequency f can be described as

$$H(f) = \sum_{p=1}^{N_{\text{pa}}} \frac{1}{\sqrt{P_{\text{att}}(f, d_p)}} e^{-j2\pi f \tau_p}. \tag{1.5}$$

The channel transfer function in (1.5) reveals that the overall channel attenuation is dependent not only on the distance, but also on the frequency. Since $\alpha(f)$ increases as f increases, high frequency waves will be considerably attenuated within a short distance, while low frequency acoustic waves can travel far. As a result, the bandwidth is extremely limited for long-range applications, while for short-range applications, several tens of kHz bandwidth could be available (a thorough study on the relationship between bandwidth (capacity) and distance is reported in [362]).

1.2.3.1 Large Delay Spread

The channel delay spread is defined as the maximal difference in the times-of-arrival of channel paths,

$$D_\tau := \max\{|\tau_p - \tau_q|\}, \qquad \forall p, q. \tag{1.6}$$

The slow speed of acoustic waves and significant multipath phenomena cause very large channel delay spread. For example, two physical arrivals that differ 15 meters in path length lead to an arrival time difference of 10 ms (here we assume the propagation speed of sound is 1500 m/s). In shallow water, the typical delay spread is around several tens of milliseconds; an example is shown in Figure 1.4(a), but occasionally delay spread can be as large as 100 ms [208]. In deep water, the delay spread can be of the order of seconds. For underwater acoustic communications, the large delay spread leads to severe intersymbol interference due to the waveform time-dispersion (also called *time-spreading*).

1.2.3.2 Large Doppler Spread

Time variability is one of the most challenging features of underwater acoustic channels. Due to medium instability, such as the current-induced platform motion and wind-generated waves as time-varying reflectors, different propagation paths could have different time-variabilities. For example, the direct path without reflections could be very stable, while the sea surface reflected paths could have time variations incurred by the motion of surface waves. The different time variabilities lead to different Doppler scaling effects or Doppler shifts of the transmitted signal.

Denote v_p as the Doppler rate of the pth path, namely the change rate of the propagation length of the pth path. The channel Doppler rate spread is defined as the maximal difference of the Doppler rates of channel paths,

$$D_{\text{d}} = \max\left\{\frac{|v_p - v_q|}{c}\right\}, \qquad \forall p, q \tag{1.7}$$

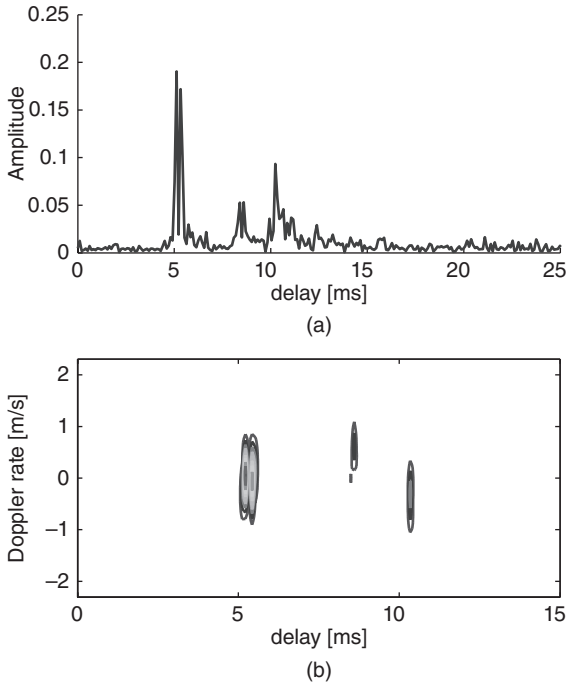

Figure 1.4 Channel profiles from a shallow water stationary experiment of SPACE'08. (a) an example channel impulse response as a function of delay; (b) an example channel scattering function on the delay-Doppler plane.

where c is the sound speed in water. The slow propagation speed of sound introduces large Doppler spread or shifts. For example, consider $v = 1.5$ m/s and $f_c = 30$ kHz, where v is the rate of change of the propagation path length (e.g, the platform velocity), and f_c is the system center frequency. The Doppler frequency shift at f_c is given by $f_d = v/cf_c = 30$ Hz.

On the outset, large Doppler spread results in a reduction in the *channel coherence time* (the time period when the channel can be viewed as static) or an apparent increase in the rate of channel fluctuation [316]. The large Doppler spread causes severe interference among different frequency components of the signal (also referred to as *frequency-spreading*). An example of the channel scattering function is shown in Figure 1.4(b).

1.2.3.3 Sparsity of Channel Paths

Despite the large delay spread and Doppler rate spread, multiple paths in the underwater acoustic channel tend to be sparse, with channel energy concentrated in a few paths; see Figure 1.4. Although the joint presence of large delay and Doppler rate spread entails a complex communication channel, the multipath sparsity is one key feature to be exploited in communication system design.

1.2.4 Acoustic Propagation Models

Given environmental parameters, the acoustic propagation in the three-dimensional underwater environment can be characterized by the wave equation

$$\nabla^2 p = \frac{1}{c^2(x,y,z)} \frac{\partial^2 p}{\partial t^2} \tag{1.8}$$

where (x, y, z) represent the 3D coordinate of one point in water, p, t, and $c(x, y, z)$ denote the sound pressure, time, and sound speed in water, respectively, and ∇^2 is the Laplacian operator. For a sinusoidal wave of frequency f_0, the wave equation can be written as the Helmholtz equation,

$$\nabla^2 p + k^2(x,y,z)p = 0 \tag{1.9}$$

where $k(x, y, z) := 2\pi f_0/c(x, y, z)$ is the wave number.

Despite of simplicity of the wave equation, finding its solution is a complicated task. Depending on applications, several typical solutions are available to characterize the acoustic field.

- Ray theory: By assuming that the phase varies much faster than the amplitude, this method takes the three-dimensional sound pressure as a product of an amplitude function and a phase function which are independent and thus can be solved individually. This above assumption limits the application of the ray theory to the high-frequency systems. Compared to other methods, the ray theory provides an appealing intuition of sound propagation. A commonly used code is the Bellhop ray tracing program [313].
- Normal mode solutions: This method provides an exact solution of the wave equation, but is restricted to the horizontally stratified channel which only has variation of sound velocity in the depth direction and assumes a flat and horizontal bottom. Hence, the solution is range independent. This method is often used in the time-reversal processing and the matched field processing. One general code is KRAKEN based on the KRAKEN normal mode model [312].
- Wave number integration: Similar to the normal mode solution, this method assumes a stratified channel, and computes the acoustic field using wave number integration. In particular, using the fast Fourier transform (FFT), the fast field program (FFP) can directly evaluate the integral solution to obtain a numerical solution of the wave equation. Although this method can provide accurate solution, it fails to provide physical interpretation of acoustic fields relative to the other three methods. An example of the FFP program is OASES (for Ocean Acoustic and Seismic Exploration Synthesis) [335].
- Parabolic approximation: Considering only the forward propagation direction, this method approximates the Helmholtz equation in (1.9) by a parabolic equation (PE) which can be evaluated numerically. A large number of PE approximations have been developed since the 1970s [195]. The PE method is suitable for calculating acoustic field in a range-dependent environment. In various PE codes, the bottom topography and surface roughness can be accounted for. An example of the PE codes is MMPE based on the Monterey-Miami PE model [344]. A recent application to high frequency acoustic transmission can be found in [336].

Notice that the acoustic field is to be described with a resolution on the level of wavelength. The latter three methods are mainly suitable to the low-frequency domain in which the acoustic

field is stable for observation. For low-frequency communications, the Parabolic approximation can be used to simulate the underwater acoustic channel, while the ray-tracing theory is commonly adopted for high-frequency systems. Please refer the textbook [195] for detailed presentations of the theory and to the website [161] for variants and program codes of the above four solutions.

1.2.5 Ambient Noise and External Interference

Noise is used to denote a signal that distorts the desired ones. Depending on applications, underwater acoustic noise consists of different components. Specific to the underwater acoustic communication system, the acoustic noise can be grouped into two categories: ambient noise and external interference.

Ambient noise is one kind of background noise which comes from a myriad of sources. The common sources of ambient noise in water include volcanic and seismic activities, turbulence, surface shipping and industrial activities, weather processes such as wind-generated waves and rain, and thermal noise [266]. Due to the multiple sources, ambient noise can be approximated as Gaussian, but it is not white. The level of underwater ambient noise may have large fluctuations upon a change with time, location or depth. For short-range acoustic communication, the level of ambient noise may be well below the desired signal. For long-range or covert acoustic communication, the noise level would be a limiting factor for communication performance.

External interference is an interfering signal which is recognizable in the received signal. Corresponding sources include marine animals, ice cracking, and acoustic systems working in the same environment. For example, snapping shrimp in warm water and ice cracking in polar regions generate impulsive interferences [78]. Sonar operations could occasionally happen at the same time with communications, creating an external interference which is highly structured [422]. Relative to ambient noise, external interferences are neither Gaussian nor white. The presence of this kind of noises may cause highly dynamic link error rate or even link outage.

It should be noted that the noise level is highly frequency-dependent. The noise power spectrum density almost monotonically decreases as frequency increases, until up to about 100 kHz when terminal noise becomes dominant. Thus, when selecting a suitable frequency band for communication, besides the frequency-dependent path loss as shown in (1.3), noise should also be taken into account [316, 362].

1.3 Passband Channel Input–Output Relationship

A diagram for the transmitter and receiver in the presence of underwater acoustic channels is shown in Figure 1.5. Since acoustic signals have low frequency, the passband samples $\tilde{x}[n]$ are often directly generated. After digital-to-analog (D/A) conversion, the passband signal $\tilde{x}(t)$ is amplified, and passed to matching circuits, matched to the transducer. At the receiver side, the weak signal is increased in level by a pre-amplifier, filtered by a simple bandpass filter, and sampled at the passband. From the signal processing point of view, the channel includes the imperfections of the transmitter and receiving circuits. All the modules that are lumped together are called channel between $\tilde{x}(t)$ and $\tilde{y}(t)$.

This book establishes the channel input–output relationship directly in the passband. This allows us to capture the wideband channel effect: (a) the propagation effect is frequency

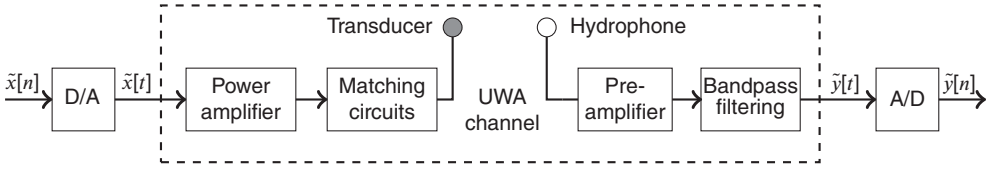

Figure 1.5 The UWA system in the passband.

dependent, (b) each transducer has its own transmit voltage response (TVR), and the matching is not uniform in the signal band, and (c) the Doppler distortion is frequency dependent. In all these considerations, the absolute values of the frequency band do matter.

Assume that the channel is linear time-varying channel. Then we can represent it by $h(t;\tau)$. From the signal processing perspective, the overall channel between the transmitter and the receiver is

$$\begin{aligned}\tilde{y}(t) &= \tilde{x}(t) \star h(t;\tau) + \tilde{n}(t) \\ &= \int h(t;\tau)x(t-\tau)d\tau + \tilde{n}(t)\end{aligned} \tag{1.10}$$

where $\star$ denotes the convolution operation.

1.3.1 Linear Time-Varying Channel with Path-Specific Doppler Scales

The channel $h(t;\tau)$ is general with no specified structure. From the system identification point of view, we need to parameterize the channel. We start with the assumption that the channel consists of N_{pa} discrete paths,

$$h(t;\tau) = \sum_{p=1}^{N_{\text{pa}}} \mathcal{A}_p(t)\delta(\tau - \tau_p(t)) \tag{1.11}$$

where $\mathcal{A}_p(t)$ and $\tau_p(t)$ are the time-varying amplitude and delay for the pth path, respectively.

For a short block of length T_{bl}, one can assume that $\mathcal{A}_p(t)$ and $\tau_p(t)$ as slowly varying. For this, one can adopt the following assumptions

- AS1): The amplitude is constant within a short block

$$\mathcal{A}_p(t) = A_p. \tag{1.12}$$

- AS2): The delay variation within one block can be approximated by a first-order polynomial

$$\tau_p(t) \approx \tau_p - a_p t, \qquad t \in [0, T_{\text{bl}}], \tag{1.13}$$

  where τ_p is the initial delay and a_p is first order derivative of $\tau_p(t)$.

The parameter a_p is often termed the Doppler scaling factor. Based on assumptions AS1) and AS2), we have a time varying channel with different Doppler scales on different paths as

$$h(t;\tau) = \sum_{p=1}^{N_{\text{pa}}} A_p \delta\left(\tau - (\tau_p - a_p t)\right) \tag{1.14}$$

The received passband signal is related to the transmitted passband signal as

$$\tilde{y}(t) = \sum_{p=1}^{N_{\text{pa}}} A_p \tilde{x}\left(\left(1 + a_p\right) t - \tau_p\right) + \tilde{w}(t) \tag{1.15}$$

where the equivalent noise $\tilde{w}(t)$ contains both ambient and model-mismatch noises as

$$\tilde{w}(t) = \tilde{n}(t) + \underbrace{\tilde{x}(t) \star h(t;\tau) - \sum_{p=1}^{N_{\text{pa}}} A_p \tilde{x}\left(\left(1 + a_p\right) t - \tau_p\right)}_{\text{signal-dependent model mismatch noise}} \tag{1.16}$$

Note that a physical channel might not be able to be represented exactly, but it can be approximated by (1.14) from a signal processing point of view.

Equations (1.14) and (1.15) are the foundations used by the receiver designs in this book, where the input output relationship is parameterized by N_{pa} triplets $\{A_p, a_p, \tau_p\}$; see Figure 1.6 for illustrations on the general and special cases.

The following special cases are often used.

1.3.2 Linear Time-Varying Channels with One Common Doppler Scale

Assume that all the paths have the same Doppler scale factor. The channel is simplified to

$$h(t;\tau) = \sum_{p=1}^{N_{\text{pa}}} A_p \delta\left(\tau - \left(\tau_p - at\right)\right). \tag{1.17}$$

This channel is parameterized by N_{pa} pairs $\{A_p, \tau_p\}$ plus the common Doppler scale. A common Doppler scale can be readily removed through a resampling operation.

1.3.3 Linear Time-Invariant Channel

Assume that all the paths are stable with no delay variations. The channel is simplified to

$$h(\tau) = \sum_{p=1}^{N_{\text{pa}}} A_p \delta\left(\tau - \tau_p\right). \tag{1.18}$$

The channel is parameterized by N_{pa} pairs $\{A_p, \tau_p\}$.

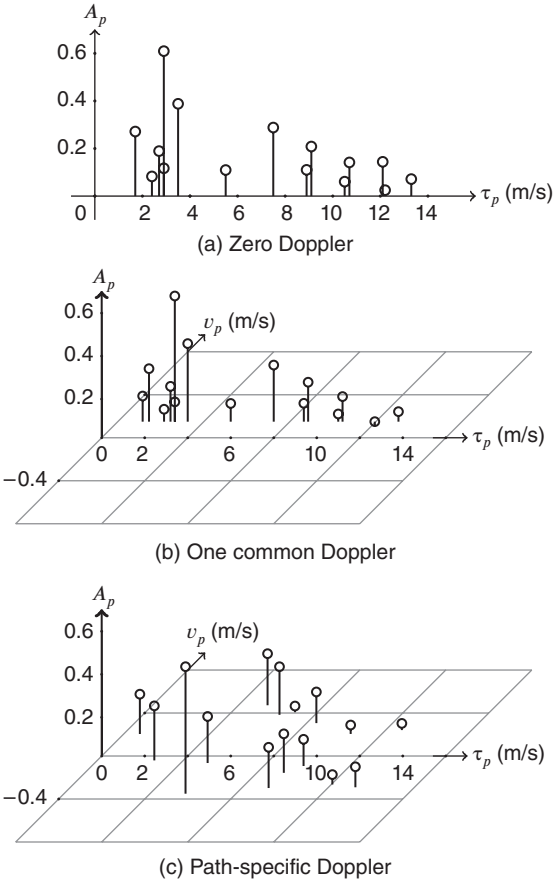

Figure 1.6 Illustration of three channel models. v_p denotes the Doppler speed associated with the pth path.

1.3.4 Linear Time-Varying Channel with Both Amplitude and Delay Variations

One can certainly extend the work to be more general, by using a Taylor expansion on the amplitude and delay respectively. For example, the approximation on the amplitude can be the N_{amp}-th order. The approximation on the delay can be on the N_{delay}-th order.

- AS1) The amplitude variation within one block can be approximated by a polynomial up to order N_{amp}:

$$\begin{aligned}\mathcal{A}_p(t) &\approx A_p^{(0)} - A_p^{(1)}t + \frac{1}{2}A_p^{(2)}t^2 + \cdots + \frac{(-1)^{N_{\text{amp}}}}{N_{\text{amp}}!}A_p^{(N_{\text{amp}})}t^{N_{\text{amp}}} \\ &= \sum_{n=0}^{N_{\text{amp}}} \frac{(-1)^n}{n!}A_p^{(n)}t^n \end{aligned} \tag{1.19}$$

- AS2) The delay variation within one block can be approximated by a polynomial up to order N_{delay}:

$$\tau_p(t) \approx a_p^{(0)} - a_p^{(1)} t + \frac{1}{2} a_p^{(2)} t^2 + \cdots + \frac{(-1)^{N_{\text{delay}}}}{N_{\text{delay}}!} a_p^{(N_{delay})} t^{N_{\text{delay}}}$$

$$= \sum_{n=0}^{N_{\text{delay}}} \frac{(-1)^n}{n!} a_p^{(n)} t^n \tag{1.20}$$

Channel parameterization based on these two polynomials is quite general, with special cases available in the literature.

Setting $N_{\text{delay}} = 1$ in AS2), the multipath channel is approximated as:

$$h(t;\tau) \approx \sum_{p=1}^{N_{\text{pa}}} \left(\sum_{n=0}^{N_{\text{amp}}} \frac{(-1)^n}{n!} A_p^{(n)} t^n \right) \delta\left(\tau - \left(\tau_p - a_p t\right)\right). \tag{1.21}$$

Receiver designs based on up to the second-order polynomial amplitude fitting and up to the first-order polynomial delay fitting have been reported in [442].

1.3.5 Linear Time-Varying Channel with Frequency-Dependent Attenuation

In a practical system as shown in Figure 1.5, the transmitter voltage response (TVR) is usually not constant due to imperfect circuit matching to the transducer across all the frequency band. Meanwhile, the signal attenuation in underwater acoustic channels is frequency-dependent. One could extend the channel in (1.11) as

$$h(t;\tau) = \sum_{p=1}^{N_{\text{pa}}} \mathcal{A}_p(t) \gamma_p\left(\tau - \tau_p(t)\right) \tag{1.22}$$

where $\gamma_p(t)$ represents a combined effect of TVR and the frequency-dependent acoustic channel attenuation. In practical systems, the TVR can be measured [34], and the propagation pattern can be determined analytically or experimentally. Recent experimental results in [404, 405] have further confirmed the wideband nature of the underwater acoustic channels. Incorporating the frequency-dependent templates into practical receiver designs is yet to be demonstrated.

1.4 Modulation Techniques for UWA Communications

The main techniques used in underwater acoustic communications are: frequency hopped FSK, direct sequence spread spectrum, single carrier transmission, sweep-spread carrier modulation, and multicarrier modulation.

1.4.1 Frequency Hopped FSK

In FSK modulation, information bits are used to select the carrier frequencies of the transmitted signal. The receiver compares the measured power at different frequencies to infer what has

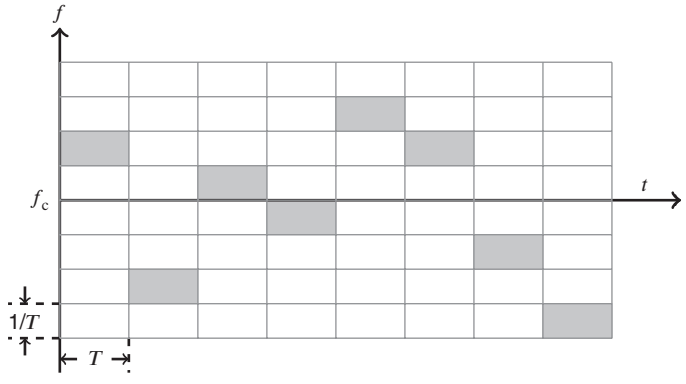

Figure 1.7 An illustration of frequency hopped FSK.

been sent. Using only an energy detector at the receiver, this scheme bypasses the need for channel estimation, and is thus robust to channel variations. However, guard bands are needed to avoid the interference caused by frequency-spreading, and guard intervals are needed to avoid the interference caused by time-spreading.

Frequency hopped (FH) FSK avoids the waiting of the channel clearing corresponding to the previous symbol, by hopping to a different frequency, as illustrated in Figure 1.7. The transmitted signal in passband is

$$\tilde{x}(t) = 2\Re\left\{\sum_{i=-\infty}^{\infty} e^{j2\pi f(i;s[i])(t-iT)} g(t-iT)\right\} \tag{1.23}$$

where T is the time duration of each tone with a frequency support of $1/T$, the tone of the ith symbol is determined by $f(i; s[i])$ which is a function of both the time slot index i and the data symbol $s[i]$, and $g(t)$ is the pulse shaper. Within a bandwidth B, the total number of tones that can be used is BT. Since BT is larger than one, there is a gain on noise suppression on the order of BT, a benefit from frequency-hopping spread spectrum. Due to the bandwidth expansion via frequency hopping, the overall bandwidth efficiency is low, typically less than 0.1 bits/sec/Hz.

1.4.2 Direct Sequence Spread Spectrum

In DSSS modulation, a narrowband waveform of bandwidth W is spread to a large bandwidth B before transmission. This is achieved by multiplying each symbol with a spreading code of length $N = \lfloor B/W \rfloor$, and transmitting the resulting sequence at a high rate as allowed by bandwidth B. For coherent DSSS, the baseband signal is

$$x(t) = \sum_{i=\infty}^{\infty} s[i] \sum_{n=0}^{N-1} c[i;n] p(t-(iN+n)T_c) \tag{1.24}$$

where $s[i]$ is the information-bearing symbol, $c[i; n]$ is the chip sequence for the ith symbol, T_c is the chip duration, and $p(t)$ is the pulse shaper on the chip level. The corresponding passband

signal is

$$\tilde{x}(t) = 2\Re\{x(t)e^{j2\pi f_c t}\}. \tag{1.25}$$

Multiple arrivals at the receiver side can be separated via the de-spreading operation which suppresses the time-spreading induced interference, thanks to the nice auto-correlation properties of the spreading sequence. Channel estimation and tracking are needed as phase-coherent modulation is used to map information bits to symbols before spreading [128, 443].

For noncoherent DSSS, information bits can be used to select different spreading codes to be used, and the receiver compares the amplitudes of the outputs from different matched filters, with each one matched to one choice of spreading code. This avoids the need for channel estimation and tracking.

1.4.3 Single Carrier Modulation

One major step towards high rate communication is single carrier transmission of information symbols from constellations such as phase-shift-keying (PSK) and quadrature-amplitude-modulation (QAM) [367]. With symbols $s[i]$ and pulse shaping filter $p(t)$, the transmitted signal is

$$x(t) = \sum_{i=-\infty}^{\infty} s[i]p(t - iT), \tag{1.26}$$

where T is the symbol period. The corresponding passband signal can be similarly obtained as in (1.25).

The channel introduces intersymbol interference (ISI) due to multipath propagation. When data symbols are transmitted at a high rate, the same physical channel leads to more channel taps in the discrete-time equivalent model. Advanced signal processing at the receiver side is used to suppress the interference; this process is termed *channel equalization*. Although widely used for slowly-varying multipath channels in radio applications, channel equalization for fast-varying underwater channel is a significant challenge.

There are various receiver designs developed for single carrier transmissions.

- The canonical receiver in [367] successfully combined a second-order phase-locked-loop to track channel phase variations with an adaptive decision feedback equalizer to suppress intersymbol interference. Multichannel DFE is adopted when there are multiple receiving elements [366].
- In the time-reversal approach, the signals from multiple elements are combined, followed by a single channel equalizer, where the equalizer can be linear or based on decision feedback [108, 148, 350, 355, 451].
- The complexity of time-domain equalization grows quickly as the number of channel taps increases, which will eventually limit the rate increase for single-carrier phase-coherent transmission. A frequency-domain equalization approach can effectively deal with channels with a large number of taps [468].

Iterative channel equalization and data decoding can be carried out for decoding performance improvement, which leads to the so-called Turbo equalization.

1.4.4 Sweep-Spread Carrier (S2C) Modulation

In all the previous modulation schemes, the carrier frequency stays constant for the whole data burst, or at least within each symbol duration as in FH-FSK. A new signaling method based on the implementation of a sweep-spread carrier, which entails rapid fluctuation of carrier frequency, has been proposed in [207]. Let f_L and f_H denote the lower and higher ends of the signal band, and T_{sw} be the sweep time. The frequency variation rate is then $m = (f_H - f_L)/T_{sw}$. The sweep-spread carrier consists of a succession of sweeps as [207]

$$c(t) = \exp\left[j2\pi\left(f_L\bar{t} + \frac{1}{2}m\bar{t}^{\,2}\right)\right], \quad \bar{t} = t - \left\lfloor\frac{t}{T_{sw}}\right\rfloor T_{sw}. \tag{1.27}$$

The baseband signal $s(t)$, which could be coherent or differentially modulated, is converted to the passband through

$$\tilde{x}(t) = 2\Re\{s(t)c(t)\}. \tag{1.28}$$

The receiver carried out the carrier demodulation through the multiplication of the received signal with an appropriately varying gradient-heterodyne signal having the same sweep cycle and the same slope of the frequency variation [207]. In the presence of a multipath fading channel, the signals with different arrival times lead to different residual frequencies after carrier demodulation. Bandpass filtering is used to separate different arrival paths. Signals along stable paths are typically selected for data demodulation. This way, the interference from unstable paths is suppressed. Note that a signal bandwidth several times larger than the symbol rate is often used to ensure a large frequency variation rate m for a good multipath resolution; hence, this method can be regarded as one special form of spread spectrum communication, which entails some noise reduction capabilities.

1.4.5 Multicarrier Modulation

The idea of multicarrier modulation is to divide the available bandwidth into a large number of subbands, where each subband has its own (sub)carrier. Within each band, the symbol rate is reduced with an increased symbol duration, so that intersymbol interference can be less severe, which helps to simplify the receiver complexity of channel equalization. There are many variants of multicarrier modulation. One way to characterize them is to check whether the subbands are overlapping or nonoverlapping, as illustrated in Figure 1.8.

Due to the existence of guard bands between neighboring subbands in the multicarrier approach with nonoverlapping subbands, bandpass filtering can be used to separate the signals from different subbands. Hence, this approach is essentially a frequency-division-multiplexing (FDM) approach, by splitting a large bandwidth into smaller pieces. Within each band, one can adopt signaling schemes such as M-ary frequency-shift keying, single carrier transmission, or another multicarrier modulation with overlapping subcarriers in a small bandwidth.

Orthogonal frequency-division multiplexing (OFDM) is one prevailing example of multicarrier modulation with overlapping subcarriers. The waveform is carefully designed to maintain orthogonality even after propagating over a long multipath channel to eliminate the need for an equalizer [44, 418]. Precisely due to this advantage, OFDM has prevailed in recent

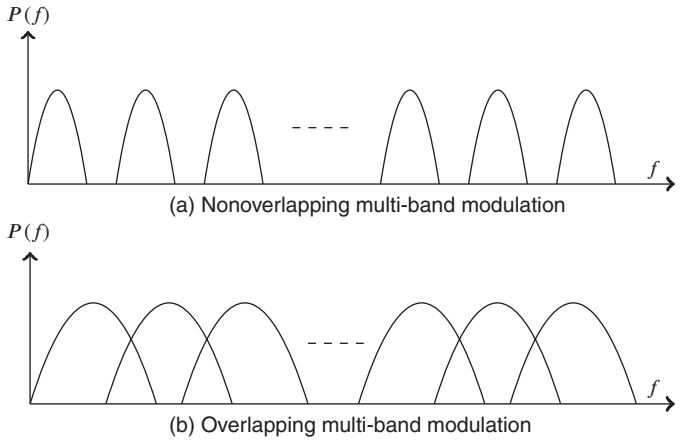

Figure 1.8 Illustration of multicarrier modulation techniques. $P(f)$ denotes the power spectrum density.

broadband wireless radio applications, including digital audio/video broadcasting, wireless local/metropolitan area networks, and fourth generation cellular networks. Filterbank based approaches also belong to the category of multicarrier modulation with overlapping subbands, making extensions to the Fourier bases used in OFDM.

The fast variations of underwater acoustic channels entail large Doppler spread which introduces significant interference among OFDM subcarriers. Signal processing specialized to underwater acoustic channels are needed to make OFDM work in underwater environments.

1.4.6 Multi-Input Multi-Output Techniques

A wireless system that employs multiple transmitters and multiple receivers is referred to as a multiple-input multiple-output (MIMO) system. It has been shown that the channel capacity in a scattering-rich environment increases linearly with $\min(N_t, N_r)$, where N_t and N_r are the numbers of transmitters and receivers, respectively [123, 379]. Such a drastic capacity increase does not incur a penalty on precious power and bandwidth resources; but instead it comes from the utilization of spatial dimension virtually creating parallel data pipes. Hence, MIMO modulation is a promising technology to offer yet another fundamental advance on high data rate underwater acoustic communication [209, 210]. MIMO has been applied in both single carrier transmission and multicarrier transmission.

MIMO introduces additional interference among parallel data streams from different transmitters. Also, each receiver has more channels to estimate, which requires more overhead spent on training symbols. For fast varying underwater channels, the number of transmitters might not be large for best rate-and-performance tradeoff. In addition to co-located antennas, distributed MIMO is also possible if clustered single-transmitter nodes could cooperate [398]. Certainly, implementation of distributed MIMO needs to address challenging practical issues such as node synchronization and cooperation.

1.4.7 Recent Developments on Underwater Acoustic Communications

The development of underwater acoustic communications prior to year 2000 has been summarized in a number of overview papers published over the years [14, 65, 208, 360]. It is often viewed that the first milestone in underwater telemetry is the introduction of digital techniques in early 1980s. The second milestone is the introduction of phase coherent processing for single carrier transmissions in early 1990, based on the seminal work in [367].

Since 2000, there have been extensive investigations on underwater acoustic communications [15, 79, 160, 343, 448]. Towards high data rate communications, the following three major directions have been pursued.

- Various receivers have been designed to improve the performance of single carrier transmissions. Example approaches include the time reversal processing in [107, 121, 330, 354, 355, 449, 450], the combination of time reversal and decision feedback equalization (DFE) in [108, 148, 346, 349, 350, 353, 355, 356, 451], the frequency domain equalization in [468], and the joint channel estimation and data detection approach in [249, 411].
- Multicarrier modulation has been successfully applied to underwater acoustic channels. For coherent demodulation, example approaches include the block-by-block based OFDM receivers in [149, 201, 232, 235, 257, 375, 457] and the adaptive OFDM receiver in [361, 363]. Explicit intercarrier interference (ICI) cancellation is one key element for enhanced receiver performance [38, 180, 392]. Noncoherent on-off keying and differential encoding have been explored in [140] and [12, 365] for OFDM systems, respectively. Adaptive modulation and coding have been recently studied for underwater OFDM [324, 413]. Variants of the multicarrier modulation other than OFDM have also been pursued [185, 287].
- MIMO techniques have attracted a lot of attention to increase the spectrum efficiency. For MIMO single carrier transmissions, example receivers include the multi-channel DFE based approach [209, 210, 331, 369], time-reversal combined with single channel DFE [347, 348, 352, 357], frequency domain equalization [460, 461], successive interference cancellation [250, 251], and iterative (Turbo) equalizations [325, 374, 377, 415, 416]. For MIMO OFDM, example approaches include the block-by-block receiver design in [172, 181, 233] and the adaptive receiver in [67, 364]. Recent receiver designs for MIMO systems with transmissions from multiple spatially distributed users can be found in [82, 83, 84, 183, 391, 423].

1.5 Organization of the Book

This book is solely focused on OFDM and MIMO OFDM for underwater acoustical channels. It is worthwhile to point out that different modulation schemes have their strengths in different situations. Generally speaking, FHSS and DSSS are good candidates for low data rates and robust operations. For low-medium rates and longer range, single carrier transmission is a suitable choice. S2C lies between DSSS and single carrier, depending on the spreading factor. For short-range and large data rates over long channels, OFDM has its competitive advantage. An intuitive illustration is shown in Figure 1.9, where the boundary and the overlapping regions should not be interpreted quantitatively.

There are many books available on OFDM and MIMO OFDM in wireless radio channels. However, as pointed out in this chapter, the challenges of radio and underwater acoustic channels are drastically different, and the receiver designs for underwater acoustic OFDM

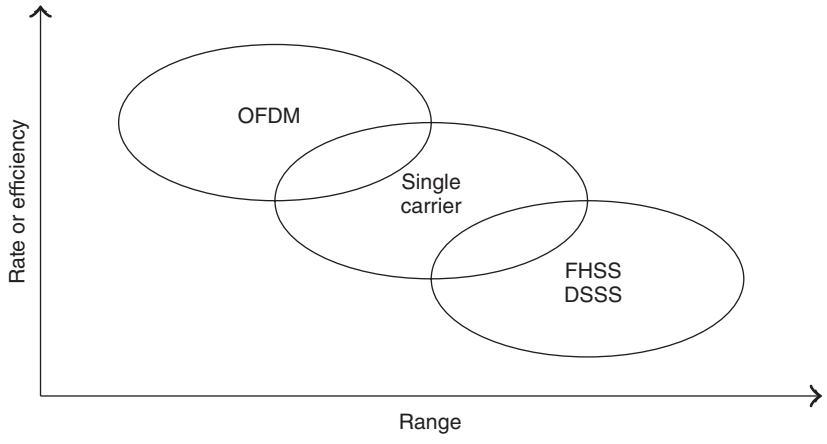

Figure 1.9 Illustration of relative merits of different techniques.

have some unique features. Specifically, the receiver design adopts a novel signal processing based model to parameterize UWA channels in (1.14), which is richer than that sufficient for radio channels. This book necessarily uses the passband formulation rather than the baseband formulation in existing books. This captures the wideband nature of underwater acoustic channels. Channel sparsity is a key element of the book, where the latest advances in compressive sensing have been incorporated into the channel estimation module. In addition, explicit intercarrier-interference consideration is adopted in this book, and iterative receiver processing is a central piece.

The chapters and appendices are arranged in the following order:

- *Part I: Basics.* Chapter 2 presents the basic principle of OFDM modulation and demodulation. Chapter 3 presents nonbinary low-density parity-check (LDPC) coded OFDM, where the LDPC codes are used within numerical and experimental results in this book. Chapter 4 discusses a property regarding to the peak-to-average-power-ratio of a transmitted OFDM signal.
- *Part II: Receiver components.* Chapter 5 presents an overview of OFDM transceiver design for point-to-point communications. Chapter 6 details the synchronization and Doppler estimation modules. Chapter 7 presents the channel and noise variance estimation algorithms. Chapter 8 contains the data detection algorithms under different channel input−output relationships.
- *Part III: Single-transmitter communications*. Chapter 9 develops a block-by-block progressive receiver along with its performance results. Chapter 10 describes a block-to-block adaptive receiver with clustered channel adaptation. Chapter 11 deals with channels having significantly separated multipath clusters, a concern in deep water horizontal communications. Chapter 12 presents an OFDM receiver in the presence of external interference.
- *Part IV: Multi-input multi-output communications*. Chapter 13 presents receiver design for co-located MIMO OFDM. Chapter 14 deals with distributed MIMO OFDM, with quasi-synchronous reception. Chapter 15 presents a receiver design for completely asynchronous multiple user communication.

- *Part V: Receiver design in relay networks*. Chapter 16 presents two cooperative relay protocols in OFDM modulated underwater acoustic networks. Chapter 17 considers the use of physical layer network coding.
- *Part VI: Modem development and underwater localization*. Chapter 18 describes the advances on the OFDM modem development. Chapter 19 covers underwater ranging and localization solutions.
- *Part VII: Appendixes*. Appendix A contains various sparse channel estimation algorithms. Appendix B describes the setup of two major experiments, from which data sets were collected to validate many algorithms described in this book.