WILEY | Hindawi

Research Article

Reinforcement Learning-Based Adaptive Switching Scheme for Hybrid Optical-Acoustic AUV Mobile Network

Hanjiang Luo (iD),[1] **Ziyang Xu,**[1] **Jinglong Wang,**[1] **Yuting Yang,**[1] **Rukhsana Ruby,**[2] **and Kaishun Wu**[2]

[1]*College of Computer Science and Engineering, Shandong University of Science and Technology, Qingdao, China*
[2]*College of Computer Science and Software Engineering, Shenzhen University, Shenzhen, China*

Correspondence should be addressed to Hanjiang Luo; luo.hj@foxmail.com

In an autonomous underwater vehicles– (AUVs–) based optical-acoustic hybrid network, it is critical to achieve ultra high-speed reliable communications, in order to reap the benefits of the complementary systems and perform high-bandwidth and low-latency operations. However, as the mobile AUVs operate in harsh oceanic environments, it is essential to design an effective switching algorithm to execute flexible hybrid acoustic-optical communications and increase the network throughput. In this paper, we propose a Q-learning-based adaptive switching scheme to maximize the network throughput by capturing the dynamics of the varying channels as well as the mobility of AUVs. In order to address the challenge associated with partial observations of the optical channel and improve the switching efficiency in extreme conditions, a blind optical channel estimation method is designed and implemented with the Extended Kalman Filter (EKF), in which the relationship between the underwater acoustic and optical channels is utilized to improve the channel prediction accuracy. Based on this environmental status, a reinforcement learning approach is leveraged to build a near-optimal switching strategy for the hybrid network. We conduct numerical simulations to verify the performance of the scheme, and the simulation results demonstrate that the proposed switching scheme is effective and robust.

1. Introduction

The deployment of underwater sensor networks (USNs) has enabled extensive marine activities of ocean monitoring and exploring [1, 2], in which both underwater acoustic communication (UAC) and underwater wireless optical communication (UWOC) are utilized for underwater networking. Although UAC is the only reliable and dominated technology that currently enables medium and long-range underwater wireless communications, it also suffers from several shortcomings (e.g., limited bandwidth, low-speed, high propagation delay, and high energy consumption [3]), which may make it difficult to meet growing application demands, such as transferring underwater real-time ultra high-definition videos and conducting real-time remotely controlled operations [4, 5].

To alleviate the limitations of UAC, as an alternative and complementary technology, UWOC has developed rapidly recently that enables high-speed, low-delay, and low-energy consumption networks, which may compensate for the deficiencies of UAC in terms of latency and bandwidth [6]. Nevertheless, UWOC has its own drawbacks communicating only with relatively short range, and it is also affected by hazardous oceanic environments (e.g., underwater obstacles, turbulence, turbidity, and light noise) [7]. To address these problems, a paradigm of multimodal networking is proposed to integrate multiple communicating systems (e.g., optical and acoustic) in a hybrid network, in order to mitigate shortcomings and take advantage of complementary technologies [8–10]. Along this line, the integration of optical and acoustic systems has been explored in both simulations and experiments [11–14], and these optical-acoustic hybrid systems

are leveraged to retrieve monitored data [9, 15], control coordination [16], build multiple routing paths [17, 18], and transfer real-time video [9], which demonstrate the benefit and effectiveness of such hybrid optical-acoustic systems [19].

Since a fleet of AUVs are desirable mobile vehicles to perform flexible underwater tasks under operational scenarios, such as environment monitoring, security surveillance, and underwater asset investigation [1, 20], the study focused on leveraging a swarm of AUVs to form underwater mobile networks has been developing rapidly in the recent years [3, 21]. By networking a swarm of AUVs equipped with both optical and acoustic communication systems, such a hybrid network enables high-bandwidth and low-latency underwater video transmission and real-time underwater operation and control [9, 22]. However, implementing such a network is not trivial as the associated challenges should be addressed carefully due to the limitations of the system, the hostile environments, as well as the mobility of AUVs [23, 24]. Among them, one critical issue is to maximize the hybrid network throughput by exploiting the complementary communicating technologies in terms of both acoustic and optical, especially the optical waves, as UWOC suffers from limited communication range [23]. Therefore, the optical link needs to be switched ON or OFF complementarily to the acoustic link in an automatic manner [13]. However, the optical system is affected by the communication distance, beam width, and the mobility of AUVs. Moreover, as the AUVs are operated in the harsh environments, the temporal and spatial dynamics of the channel (e.g., the turbidity of the seawater, the obstacles, and the turbulence and currents) may degrade the performance of the communication significantly [7]. Consequently, it is desirable to monitor and even predict the changing conditions of the channels in order to carry out the switching scheme in an adaptive and effective manner [25, 26].

There are a few pioneering studies dealing with the switching issue and providing precious investigations of underwater multimodal wireless networks [13, 27–29]. In [13], the authors explored the multimodal switching issue to maximize instantaneous network throughput using the range-based triggering mechanism to proactively switch among different physical layers (PHYs). However, the switching strategy uses a preset threshold to execute the switching policy. In [27], the authors discussed the statistical relationship between acoustic and optical channels and proved feasibility of predicting the optical channel state based on the properties of acoustic channels. Although a switching scheme based on optical signal-to-noise ratio (SNR) threshold was presented, no further specific experiments and simulations were conducted to validate the performance. Nevertheless, an effective switching strategy should be adaptive to the environmental dynamics, and as reinforcement learning (RL) has the capability of interacting with the environment and could gradually learn an optimal or near-optimal action policy, it is a promising artificial intelligence (AI) tool to develop switching strategy for underwater sensor networks. In a recent research, a model-free RL method is adopted to deal with the dynamics of the channels in order to smoothly switch among different types of acoustic modems

in an adaptive manner [28, 29]. However, the switching policy in a hybrid optical-acoustic communication system, involving two different types of links, has not been studied much. Furthermore, as the UWOC has more restrictions compared to the acoustic communication, these limitations should be carefully addressed in the switching scheme.

To tackle the aforementioned challenges, in order to maximize the throughput of the hybrid AUV mobile network, as depicted in Figure 1, in this paper, we leverage the Q-learning and Extended Kalman Filter (EKF) tools to deal with the severe oceanic environments and propose an adaptive switching scheme. Compared to these learning-ignorant schemes, the proposed scheme does not require prior knowledge of the environments. Furthermore, we leverage the relationship between the acoustic and optical channels to enhance the effectiveness of switching strategy by capturing the dynamics of the varying channels [27]. To the best of our knowledge, this is the first study to provide an adaptive switching scheme for the AUV-based hybrid optical-acoustic network. To summarize, the major contributions of this paper are listed as follows:

(1) We propose an adaptive switching scheme for an AUV-based optical-acoustic hybrid network that leverages both the Q-learning and EKF tools to increase the network throughput

(2) A blind estimation method for underwater optical channel state is implemented with the EKF tool to improve the channel prediction accuracy for effective proactive switching

(3) The critical factors (e.g., acoustic and solar noise, water turbidity, AUV mobility, and optical beam width) which affect the effectiveness and robustness of the switching scheme are investigated via numerical simulations

2. Preliminaries

In this section, we introduce the preliminaries for the proposed adaptive switching scheme, which include the following: network model, acoustic channel, optical channel, and localization of AUVs.

2.1. Basics of Network Model. The hybrid underwater acoustic-optical network includes a swarm of AUVs and an underwater positioning system, as shown in Figure 1. Each AUV is equipped with both the acoustic and optical communication technologies. Moreover, the positioning system is deployed either on an assistant ship or on a floating buoy, which is near the AUV operating area. As depicted in Figure 2, the hybrid system consists of an underwater wireless acoustic communication (UWAC) link and an UWOC link. The UWAC link is dedicated for conveying the feedback information of channel states and the positioning information of the AUVs. Therefore, the acoustic channel is used for both the control and low-rate data transmission purposes.

As the swarm of AUVs are mobile, the designated switching scheme aims to choose the UWAC link or UWOC link
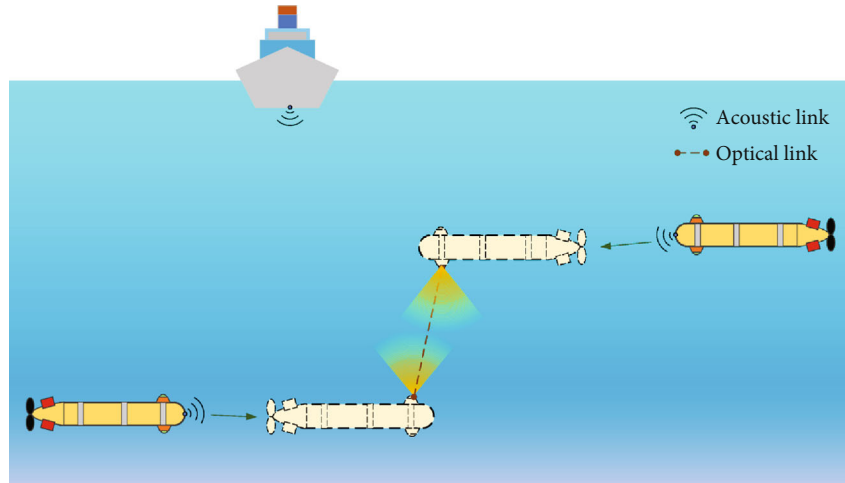
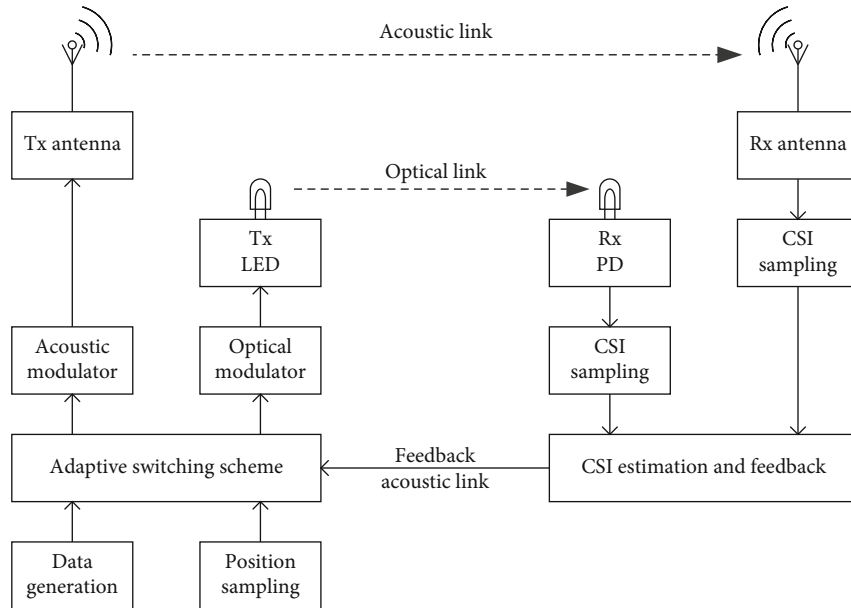FIGURE 1: A sample structure of the hybrid optical-acoustic network.

FIGURE 2: A sample block diagram of hybrid optical-acoustic systems.

for data transmission in an adaptive manner based on the channel conditions and the AUV positions. Once the conditions for the optical link are favorable, a well-designed scheme should switch it ON, which means the initiation of optical transmission for many tasks, such as pointing, acquisition, and tracking (PAT) system [30]. During the process, the receiver sends back the feedback information (e.g., channel state information (CSI)) into the acknowledge string (ACK) via acoustic links until the UWOC link has been established.

2.2. Basics of Acoustic Channel. To monitor the channel state and measure the noise level, the signal-to-noise ratio (SNR) is selected as CSI to characterize the channel. The path loss in the underwater acoustic channel is given by [31]: $A(r, f) = A_0 r^b a(f)^r$, where A_0 is the normalizing constant, r is the

distance between the transmitter and the receiver, b is the spreading factor, f is the acoustic signal frequency, and a (f) is the absorption coefficient estimated by Thorp's derivation which is constant for the specific frequency sound wave.

By ignoring the changes of noise and transmission power in a short time period, the observation of SNR of time-varying and frequency-selective channel can be formalized as Markov chains in discrete time domain, which implies that the current theoretical prediction of SNR depends on the past prediction. The SNR of transmission performed at frequency over distance between a transmitter and a receiver can be expressed as follows [31]:

$$SNR_A = \frac{P}{A(r, f)N_0(f)B},$$ (1)

where P is the transmit power, N_0 is the noise power, and B is the narrow band around the signal frequency.

Although the underwater acoustic noise can be described as the white Gaussian noise, the received signal always suffers from Non-Gaussian noise (e.g., alpha noise and Middleton Class A and Class B noise) consisting of short spike pulses generated by external interference in the underwater channel [32, 33], which result in the probability density function (PDF) of the noise with longer and extended tails. According to these studies, it is known that the Non-Gaussian noise follows the symmetrical alpha-stable (SaS) distribution class [34]. The characteristic function of the SaS distribution class is given by

$$\varphi(x) = \exp\left(i\mu x - \gamma|x|^{\alpha}\right) \tag{2}$$

where α is the characteristic exponent that determines the degree of pulse characteristic of the distribution and $0 < \alpha \leq 2$ holds. The terms μ and γ are the location and dispersion parameters, which are similar to the mean and variance of the Gaussian distribution. When $\alpha = 2$ holds, the α stable distribution is equivalent to the Gaussian distribution.

To suppress the spikes and make the noise conform to the Gaussian distribution, some filters have been proposed such as U-filter's Gaussianization process [35] and median filter [36]. By leveraging the aforementioned filters, the noise of the acoustic channel $Noise_A$ is defined as

$$Noise_A \sim N\left(SNR_A, \sigma_{AC}^2\right), \tag{3}$$

where σ_{AC}^2 represents the mean error variance.

2.3. Basics of Optical Channel. To build an UWOC link among the mobile AUVs, an LED is adopted as the transmitter in this work, which has a wide beam angle to reduce the strict alignment requirement and guarantee the communication between AUVs [27].

Correspondingly, we also choose SNR as the main feature of the optical CSI, since in a real-time control system (RTCS), the optical PAT is driven by affluent SNR. Note that when the AUV is equipped with an optical noise sensor, the short-term noise dynamics can be taken into consideration. The SNR can be modelled as a Markov chain [37] as follows:

$$SNR_O = \left[\frac{P_t e^{-c(\lambda)r} D^2 \cos\phi}{4r^2 NEP \tan^2\theta}\right]^2, \tag{4}$$

where P_t is the optical transmit power, r is the distance between the transmitter and the receiver, $c(\lambda)$ stands for the attenuation coefficient which consists of absorption coefficient $a(\lambda)$ and scattering coefficient $b(\lambda)$: $c(\lambda) = a(\lambda) + b(\lambda)$, λ is the beam wavelength, D is the receiver aperture diameter, ϕ is the incident angle between the optical axis of the receiver and the line-of-sight (LOS) direction, θ is the half angle transmitter beam width, and NEP is the noise equivalent power.

Under the consumption of unchanged transmit power and fixed optical system parameters, the noise of optical channel $Noise_O$ is defined by

$$Noise_O \sim N\left(SNR_O, NEP^2\right), \tag{5}$$

where the noise of optical NEP^2 can be modelled as the sum of a series of Gaussian noises. It includes the thermal noise in the signal amplification process, the quantum shot noise generated at the receiving end, the photodetector dark current noise generated by the photodetector electrical current leakage, and the background noise caused by environment optical clutter.

2.4. Localization of AUVs. Since both the acoustic and the optical channels are affected by the communication distance, a positioning system is required to determine the positions of AUVs [3]. As shown in Figure 3, we use the ultra short baseline positioning (USBL) technique to locate AUVs by measuring the phase difference of the target's acoustic signal to the hydrophone through the shipborne array probe as follows:

$$\begin{cases} x = d \cos\alpha \\ y = d \cos\beta \end{cases}, \tag{6}$$

where d represents the distance between the shipborne base station and the positioning target, α is the azimuth angle along the x-axis, and β is the azimuth angle along the y-axis. Then, the position information is broadcast to the transmitter AUV Tx and the receiver AUV Rx. After that, Tx can calculate the relative position $P_{Tx \longrightarrow Rx}^t$ by the following formula [38]:

$$P_{Tx \longrightarrow Rx}^t = \begin{bmatrix} x_{Rx} - x_{Tx} \\ y_{Rx} - y_{Tx} \\ z_{Rx} - z_{Tx} \end{bmatrix}, \tag{7}$$

where z_{Tx} is the depth of the transmitter and z_{Rx} is the depth of the receiver, which are updated after completing each communication round. We assume that the trajectory of the AUV is preset by the shipborne base station. Thus, the relative position in the next time slot can be predicted in advance.

3. Adaptive Switching Scheme

In this section, we present the adaptive switching scheme design based on the RL technique. The overall diagram of switching scheme is shown in Figure 2. We first provide the fundamental mechanism of the proposed switching policy, then propose an optical channel estimation based on the EKF in detail, and finally provide the switching algorithm at the end of the section.

3.1. The Design of the Switching Scheme. As shown in Figure 4, we briefly describe the primary mechanisms of
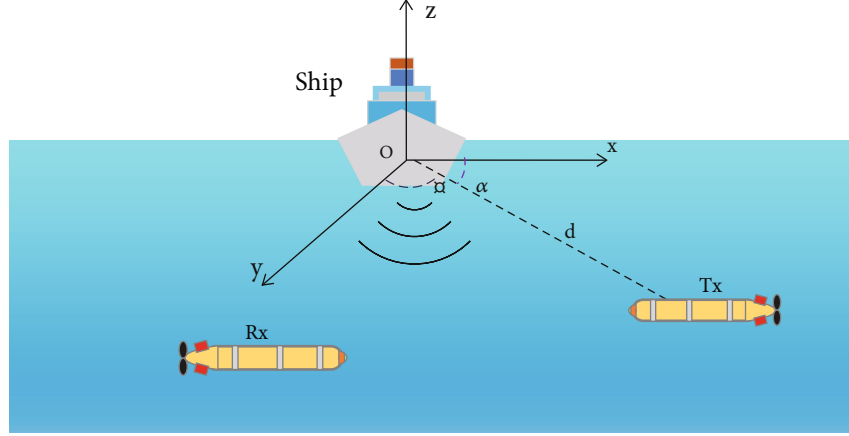
FIGURE 3: The positioning system by ultra short baseline positioning technology.
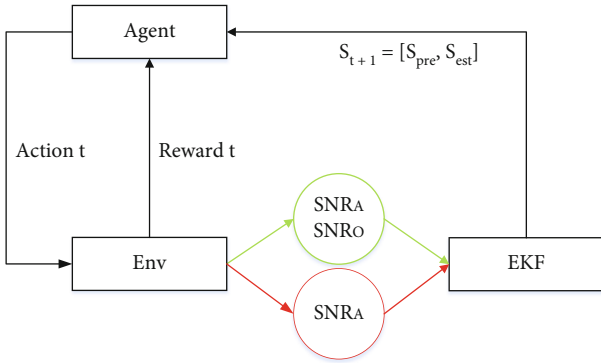

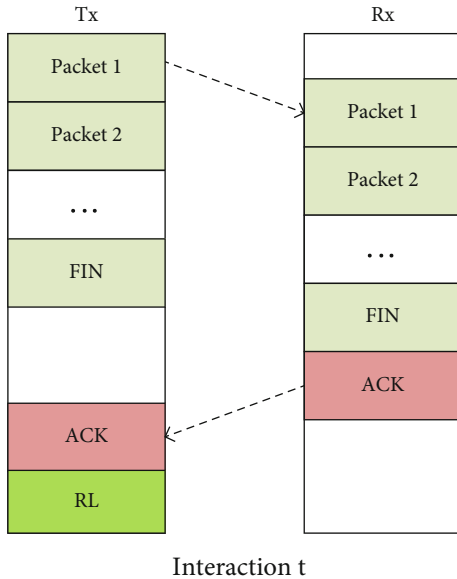FIGURE 4: The illustration of the adaptive switching scheme.

FIGURE 5: Illustration of CSI estimation and feedback process.

the switching scheme. To design an effective switching policy, the AUV should consider the positioning of the target transmitter, the states of acoustic and optical channels. The channel states are denoted as SNR_A and SNR_O, respectively.

These sensed information are then fed into the EKF tool for further processing before sending to the agent as part of the environment states. It should be noted that SNR_O is only measured and collected after the UWOC link is switched ON.

More precisely, the agent in an AUV generates and updates the policy function $\pi(s)$. The inputs to the policy function come from the communications with the target AUV via acoustic link, including state S and reward R. The state set S consists of acoustic channel state estimation outcome S_{pre} and the optical channel state estimation outcome S_{est} as shown in Figure 4. The output of the policy function is only a single action, which affects the measurement of the next state.

Optical state information is more difficult to acquire compared to acoustic channel states because of the extra cost of alignments. Therefore, we keep apart the acoustic and optical communications in the learning process, which in turn requires a reward function R specifically for the optical communication, but not for the acoustic communication. Subsequently, the AUV alternating policy function $\pi(s)$ is equivalent to the exploration of the successful connections for the optical communications. To realize the exploration, we design a blind estimation model for the optical channel. The basic idea of the estimation process is that the receiver estimates the optical CSI by combining the theoretical value of optical CSI and the acoustic measurement outcome when the optical information is not available via the relationship between the acoustic and optical channel states [27].

The proposed optical channel state estimation model is developed based on the EKF tool, and the optical CSI measurement is used to estimate the optical channel SNR_O. When the optical CSI measurement is not available, we obtain the optical estimation state S_{est} by updating the observation matrix with the acoustic CSI tracking information SNR_A and its theoretical value.

3.2. Design of the RL Technique. The interaction between an AUV and the optical channel can be formulated as a Markov Decision Process (MDP). The UWOC channel state transition matrix is estimated and updated by acoustic channel

TABLE 1: The settings of the SNR levels.

State	Setting	
	dB	Level
S_{pre}	$SNR_A \leq 5$	0
	$5 < SNR_A \leq 10$	1
	$10 < SNR_A \leq 15$	2
	$SNR_A > 15$	3
S_{est}	$SNR_O \leq 5$	0
	$5 < SNR_O \leq 10$	1
	$10 < SNR_O \leq 15$	2
	$15 < SNR_O \leq 20$	3
	$SNR_O > 20$	4

observations while switching. Therefore, there is a trade-off between exploration and exploitation in AUV. Exploration implies updating the model frequently so as to obtain a more accurate future prediction, which incurs additional costs. On the other hand, the higher degree of exploitation may result in a local optimal solution. For this reason, an effective quantification function is proposed to adjust ε, which urges the agent to have more explorations when the constructed model is ineffective. The following elements are included in the design of the proposed scheme:

(i) Environment: The environment contains the acoustic and optical channels and the receiver, and it generates reward when the optical communication is ON

(ii) State: The state S describes the current environment. We combine the SNR S_{pre} of the acoustic channel and the optical estimated SNR S_{est} obtained by the EKF tool as states S, which is denoted as S = $[S_{pre}, S_{est}]$. To simplify the design, the state quantities are discretized into N and M levels in a total of N × M states

(iii) State value function: The estimated value of current state $V_\pi(s)$ is calculated as the expectations of the future rewards

(iv) Action: The action set Action = [0, 1] is the UWOC link binary controller bit, where 0 means OFF and 1 means ON

(v) Reward: The reward function R_t can be set according to the different tasks of the AUV. An agent seeks the optimal policy to maximize the value function V_π while leveraging the feedback information from the environment. As there is only a unique task, the reward function R_t is set to a fixed value, which is only obtained during the UWOC link. If the UWOC is successfully established, the agent will get the reward χ to update Q-value of action ON. While UWOC is failed to set up, the agent will get the reward φ to modify Q-value of action OFF:

$$R_t = \begin{cases} \chi, & Success \\ \varphi, & Fail \\ 0, & Otherwise \end{cases}, \tag{8}$$

The policy function π is a greedy strategy, which balances the relationship between the exploration and exploitation as follows:

$$\pi(s) = \begin{cases} argmax_{\tau \in Action} Q(s, \tau), & 1 - \varepsilon \\ random, & \varepsilon \end{cases}, \tag{9}$$

where the action τ to maximize the Q-value is chosen with probability $1 - \varepsilon$ and the action is selected randomly with probability ε.

Bellman equation is used to describe the relationship between the current state value and the consequent state value that the current state value is equal to the expectation of the sum of discount next state value and the instantaneous reward [29]:

$$V_\pi(s) = E_\pi(R_{t+1} + \gamma V_\pi(S_{t+1}) | S_t = s), \tag{10}$$

where γ is the discount factor. The learned action value function Q directly approximates the optimal action and maximizes over all possible actions in the next state as follows:

$$Q(s, \tau) = (1 - \delta)Q(s, \tau) + \delta \left(R_{t+1} + \gamma \max_{\tau'} Q(s', \tau') \right): \tag{11}$$

where δ is the learning rate. We assume that the environmental information is observable. Then, we set the reward function R_t according to the network requirement. For example, as we want to increase the successful trial of the switching process, we continuously update Q-table in interaction with the environment. The Q-function $\pi(s)$ uses a tabular approximation method, and the CSI information is approximated into several levels.

3.3. Optical Channel Estimation. As discussed earlier, the state of the RL technique consists of two kinds of information: SNR_A and SNR_O. Since the acoustic link is used for the control channel, its SNR can be easily obtained. The actual optical SNR, especially when there is no optical communicating link, cannot be obtained directly. In [27], the possibility of using acoustic SNR to predict optical SNR is discussed. Inspired by this research, we use the EKF tool to estimate optical SNR from the acoustic SNR which require a small amount of prior knowledge and is adapted to the AUV movement.

Initialization: Learning rate δ, exploration and exploitation threshold ε state-action-value-function $Q(s, \tau) = 0$.
1: **for** $t = 1, 2, 3, \cdots$ **do:**
2:　　Obtain the relative position $P^t_{Tx \longrightarrow Rx}$ using the relation in (7)
3:　　Predict the hybrid channel state mean $\widehat{X}_{t|t-1}$
4:　　Predict the variance $P_{t|t-1}$ using the relation in (23).
Blind Estimation Stage
5:　　**if** $\tau = 0$ **then**
6:　　　　Estimate the observation Z^b_t using the relation in (17).
7:　　　　Obtain the channel estimation $\widehat{X}_{t|t}$ using the relation in (24).
8:　　**end if**
Feedback Stage
9:　　**if** $\tau = 1$ **then**
10:　　　　Observe the channel state Z^f_t using the relation in (20).
11:　　　　Estimate the channel state $\widehat{X}_{t|t}$ using the relation in (24).
12:　　　　Obtain the reward R_t.
13:　　**end if**
Online Learning Stage
14:　　　　Update $S_{t+1} = [S_{pre}, S_{est}]$
15:　　　　Update Q-Value $Q(s, \tau)$ using the relation in (11).
16:　　　　Choose action $\tau \in Action$ using the relation in (9).
17:　　　　Update the EKF parameters.
18: **end for**

ALGORITHM 1: The proposed adaptive switching algorithm.

TABLE 2: UWOC parameters.

Parameters		Value
Receiver sensitivity	$\Re$	$0.26 \, A/W$
Receiver aperture diameter	D	$1.1 \, mm^2$
Half angle transmitter beam width	θ	$0.5 \, rad$
Electron charge	q	$1.6 \times 10^{-19} \, C$
Wavelength	λ	$532 \, nm$
Downwelling irradiance	E	$1440 \, watts/m^2$
Radiance factor	L_{fac}	2.9
Diffuse attenuation coefficient	T_A	$\exp(-\tau_0)$
Boltzmann's constant	k	1.381×10^{-23}
Equivalent temperature	T_e	$290 \, K$
Dark noise	I_{DC}	$1.226 \times 10^{-9} \, Ampere$
Transmit power		$30 \, W$
Beam attenuation coefficient		1.0
Signal bandwidth		$100 \, kHz$
Planck's constant		$6.6261 \times 10^{-34} \, Js$
Underwater reflectance		1.25%
Speed of light		2.25257×10^8
Radiant absorption factor		0.5
Noise figure of system		4
Load resistance		$100 \, \Omega$

We first define the state space vector X_t at time slot t as follows [38]:

$$X_t = \begin{bmatrix} P^t_{A \longrightarrow B} \\ SNR^t_O \\ SNR^t_A \end{bmatrix}_{5 \times 1}, \quad (12)$$

where $P^t_{A \longrightarrow B}$ is the relative position vector of AUVs and SNR^t_O and SNR^t_A are the SNRs of optical and acoustic channels, respectively. Then, we derive the state transition function from (4) as follows:

$$X_t = f(X_{t-1}) = F_t X_{t-1} + \omega_t,$$
$$= \begin{bmatrix} I_{3 \times 3} & 0 & 0 \\ \cdots & G^t_O & 0 \\ 0 & \cdots & G^t_A \end{bmatrix}_{5 \times 5} X_{t-1} + \omega_t, \quad (13)$$

$$G^t_O = \frac{1}{\left(e^{c(|P_t| - |P_{t-1}|)} \times (|P_t|/|P_{t-1}|)^3 \times (x_{t-1}{}^2 + y_{t-1}{}^2/x_t{}^2 + y_t{}^2) \right)}, $$
$$G^t_A = \frac{1}{(1 + (\Delta r / r_{t-1}))^b \times a(f)^{\Delta r}}, \quad (14)$$

where F_t is the transfer matrix and is used to adjust the prior estimation covariance matrix $P_{t|t-1}$ and ω_t is the process noise which obeys the Gaussian distribution with mean 0 and covariance Q. Secondly, the observations can be divided into two stages based on the availability of optical channel
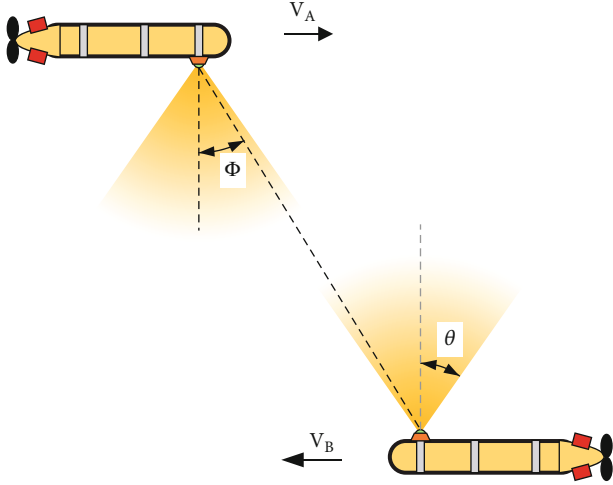
FIGURE 6: A sample optical communication scenario between two AUVs.

observations, such as blind estimation stage and feedback stage.

In the blind estimation stage, due to the lack of the optical channel observation, we resort the relationship between acoustic and optical channels to estimate the optical SNR via the acoustic measurements. At this stage, the observation (or measurement) vector Z_t^b at time t can be expressed as

$$Z_t^b = \begin{bmatrix} P_{A \longrightarrow B}^t \\ SNR_A^t \end{bmatrix}_{4 \times 1}, \tag{15}$$

where $P_{A \longrightarrow B}^t$ and SNR_A^t are the measurement of relative positions of AUVs and the acoustic SNR. Then, the observation is expressed as

$$Z_t^b = h_b(X_t) = H_t^b X_t + v_t^b b = \begin{bmatrix} I_{3 \times 3} & 0 & 0 \\ 0 & 0 & L^t \end{bmatrix}_{4 \times 5} X_t + v_t^b, \tag{16}$$

$$L^t = C_b SNR_O \frac{r^{4-b} \exp^{2cr} NEP^2}{a^r \cos^2 \varnothing}, \tag{17}$$

where $h_b(\bullet)$ represents the mapping at the blind estimation stage, which converts the 5×1 state vector X to the corresponding 4×1 measurement vector Z_b. H_t^b is observation matrix and v_t^b is the observation noise, which is assumed to be zero mean Gaussian white noise with covariance R. C_b is a constant related to signal amplifier circuits and the underwater environment.

In the feedback stage, we use the optical channel measurement as feedback to the EKF tool, in order to improve the accuracy of the estimation. The observation (or measurement) vector Z_t^f at time t of this stage can be expressed as

$$Z_t^f = \begin{bmatrix} P_{A \longrightarrow B}^t \\ SNR_O^t \\ SNR_A^t \end{bmatrix}_{5 \times 1}, \tag{18}$$

where SNR_A^t is the measured value of the optical SNR. Afterwards, the observation is expressed as

$$Z_t^f = h_f(X_t) = H_t^f X_t + v_t^f, \tag{19}$$

$$H_t^f = I_{5 \times 5}. \tag{20}$$

where $h_f(\bullet)$ represents the mapping at the feedback stage, which converts the 5×1 state vector X to the corresponding 5×1 measurement vector Z^f. H_t^f is an identity matrix and v_t^f is the observation noise, which is assumed to be zero mean Gaussian white noise with covariance R. According to the aforementioned descriptions, the observation function $h(\bullet)$ and the observation matrix H_t of the system can be abbreviated as the following relation:

$$Z_{t+1} = h(X_t) = \begin{cases} h_b(X_t) + v_t^b, & at\ Blind\ Stage \\ h_f(X_t) + v_t^f, & at\ Feedback\ Stage \end{cases}, \tag{21}$$

$$H_t = \begin{cases} H_t^b, & at\ Blind\ Stage \\ H_t^f, & at\ Feedback\ Stage \end{cases}, \tag{22}$$

Due to the nonlinearity nature of the system view state and observation function, EKF is employed for channel state estimation since it is a nonlinear version of Kalman filter. Standard EKF tool generally consists of two phases: prediction and updating. There are three covariance matrices: P, Q, and R [39]. The Q and R are both positive definite matrices which depended on the environment settings, and the $P_{0|0}$ is initialized as an identity matrix. The state vector and its covariance matrix can be iteratively updated by the following relations [39]:

Prediction:

$$\hat{X}_{t|t-1} = f\left(\hat{X}_{t-1|t-1}\right),$$
$$P_{t|t-1} = F_t P_{t-1|t-1} F_t^T + Q. \tag{23}$$

Updating:

$$\tilde{y}_t = z_t - h\left(\hat{X}_{t|t-1}\right),$$
$$S_t = H_t P_{t|t-1} H_t^T + R,$$
$$K_t = P_{t|t-1} H_t^T S_t^{-1}, \tag{24}$$
$$\hat{X}_{t|t} = \hat{X}_{t|t-1} + K_t \tilde{y}_t,$$
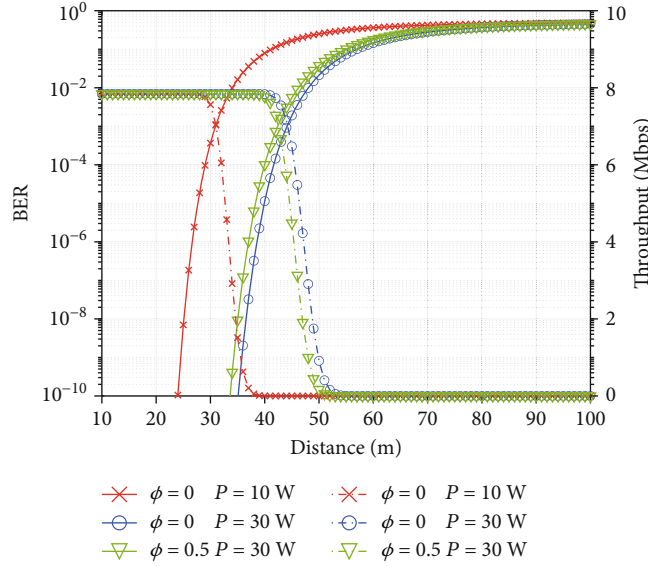$$P_{t|t} = (I - K_t H_t) P_{t|t-1}.$$

FIGURE 7: The characteristic of UWOC with different distances.
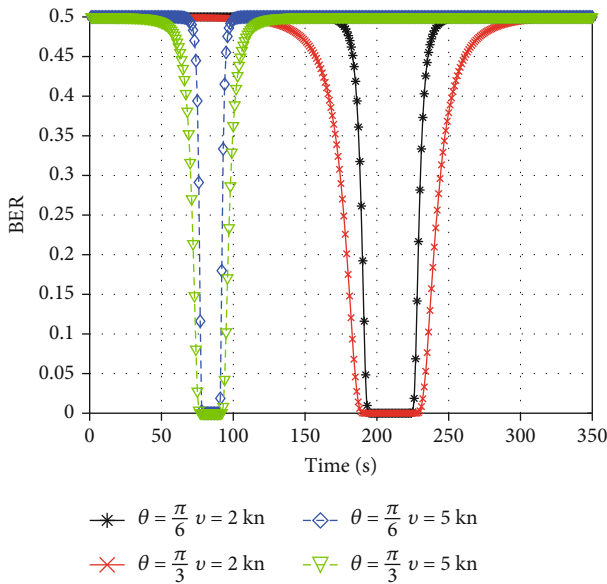

FIGURE 8: The characteristic of UWOC with different AUV speeds and half angles.

After the observations, such as relative position, acoustic SNR, or optical SNR, have been processed, the EKF tool provides an estimation of X as the input state for the RL technique to make decisions.

3.4. The Description of the Proposed Switching Algorithm.
The proposed algorithm is deployed in the controllers of the AUVs, and the switching policy is stored in a Q-table and updated by communicating with other AUVs. The discrete time slot is denoted as the period which begins at the packet transmission until Tx has received the ACK from Rx and updates the Q-table, and the sequences are related to the packet transmission process as shown in Figure 5.

We assume that the environment remains relatively stable over a short period of time. At the beginning of each interaction, the transmitter first receives the positioning information $P_{Tx \longrightarrow Rx}$, and the receiver tracks and estimates the channel state X_t using the EKF tool while receiving the packets. Once the transmission is initiated, the receiver estimates the optical channel state $\widehat{X}_t$ and places the estimation value into the ACK. Finally, the transmitter progresses to the online learning stage and updates the policy function $\pi(s)$. The initial parameters are the learning rate δ, the exploration constant ε, and the system error σ. Before the packet transmission, the position of the target is determined by the positioning system which resides in the shipborne base station or a buoy.

Then, the receiver calculates the prediction of the hybrid channel state $\widehat{X}_t$ using the relation in (23) and collects the observations. The value of the observation function is defined in (21). While the optical communication is OFF, the measurement function Z_t is obtained by the blind estimation which combines the theoretical value of optical SNR_O and the acoustic measurement SNR_A using the relation in (17). Then, the observation is updated with the optical measurement outcome using the relation in (20) under the process of optical communication. We choose SNR to represent CSI [23]; thus, we have a double state tuple S_t consisting of both the acoustic SNR estimation outcome SNR_A^t and the optical SNR estimation outcome SNR_O^t. The quantified levels are shown in Table 1, and we set the state $S_t = [S_{pre}^t, S_{est}^t]$ for the RL algorithm.

The condition for judging optical communication successful is based on an SNR threshold which the PD on the receiver conceives the optical signal is real [23]. It should be noted that the optical signal measurement of the receiver is sent back to the transmitter via the acoustic channel. Therefore, the online learning stage can be reached at both the transmitter and the receiver ends through the bi-directional
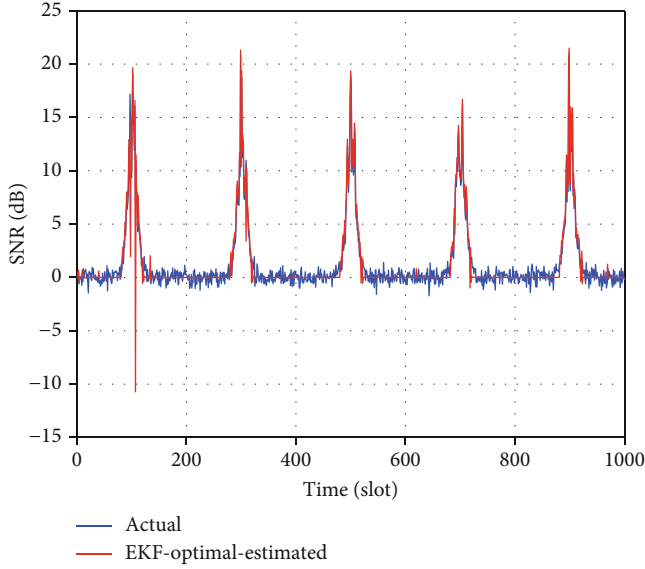
FIGURE 9: The convergence of the estimation and the actual state.

communication process, and it reduces the costs of learning. After the successful switching operation, the optical SNR is used to assist the PAT procedure [39].

The adaptive switching process is summarized in Algorithm 1. Once the system is initialized, the algorithm is mainly divided into three parts: blind estimated stage, feedback stage, and online learning stage. At the beginning of each interaction, the relative position P of each agent is obtained to predict the channel state $\widehat{X}_{t|t-1}$ and the covariance P. When the optical channel is not obtained by the agent, the blind estimation stage is used to estimate the optical channel observation Z by the acoustic observations using the relation in (17). Otherwise, when the UWOC is switched ON, the feedback stage is used to estimate the optical channel state Z by the optical observations. Subsequently, the corresponding agent obtains the reward R at the end of UWOC switching trials. After getting the channel estimation $\widehat{X}_{t|t}$ from the blind estimation stage or the feedback stage where the optical state is measured at the feedback stage but absent at the blind estimation stage, the agent updates the next state S_{t+1} using SNR_A and SNR_O in Table 1. In the online learning stage, the agent updates the Q-table and EKF parameters, and then the action of the next interaction is chosen using the relation in (9).

4. Performance Evaluation

In this section, we evaluate the performance of the proposed switching scheme through simulations. The operating area of AUVs is set to $2000\,m \times 2000\,m \times 500\,m$. The primary simulation parameters and settings are provided in Table 2. We first simulate the characteristics of optical communication between mobile AUVs. Then, we verify the effectiveness of the proposed switching algorithm. Finally, we testify the robustness of the proposed scheme.

4.1. The Characteristics of Optical Communication between Mobile AUVs.
The performance of the optical communication of mobile AUVs is affected by several factors. As shown in Figure 6, when two mobile AUVs encounter, the chance of performing reliable visible light communication (VLC) via LOS link depends on several factors (e.g., the distance between two AUVs, the transmit power P, the half-angle FOV θ, the incident angle of the receiver $\varnothing$, and the speed of AUV v).

As shown in Figure 7, we illustrate the performance variations under different communication distances between two AUVs, and also with different levels of transmit power P and the incident angle $\varnothing$. The half angle is fixed to $\theta = \pi/6$. The collision avoidance range of AUVs is set to 10 m [40].

As shown in the figure, the solid lines describe the throughput, and the dotted lines represent the bit error rate (BER). The distance between two AUVs is varying from 10 m to 100 m. We take different settings of transmit power and incident angle to compare the performance, such as the transmit power is set to 10 W and 30 W, and the incident angle is set to 0 rad and 0.5 rad. Figure 7 shows that the optical communication throughput has good performance which can reach the magnitude of Mbit. However, the throughput declines rapidly when the distance is larger than a number in between the range of 30 m and 50 m. Consequently, we can take some measures to improve the optical transmission capability, such as improving the transmit power. The figure also shows that the BER has a significant performance for applications associated with the communication distance of 40 m.

Figure 8 shows the UWOC characteristics of two mobile AUVs with respect to different speeds and half angles. The speed of AUVs is set to 2 and 5 knots, and the half angle θ is set to $\pi/6$ and $\pi/4$. As shown in the figure, it can be observed that under the same half angle θ, the time window is larger with a slow speed of AUVs in terms of BER communication performance. The delay that an AUV experiences while communicating with another one via the optical link is longer with a wide half angle compared to the narrow half angle under the same speed.

4.2. The Effectiveness of the Switching Scheme.
In order to verify the effectiveness of the proposed scheme, we simulate the performance of the EKF and Q-learning tools. In the simulations, two AUVs are deployed in the depth of 15 m and 25 m. The trajectories of these AUVs are in a linear roundtrip patrol along a straight line with a range span about 500 m. The initial positions of the AUVs are on the left and right sides of their patrol lines, and the AUVs are moving towards each other at a speed of 2 knot and 4 knot, respectively. Each interaction process is defined as a time slot as shown in Figure 5. When the two AUVs patrol along the preset lines, the SNR of UWOC gradually increases and decreases in accordance with the change of distance between two AUVs.

As shown in Figure 9, the EKF tool has a good performance when the optical SNR is large enough, and both states are generated under the simulated environments. Overall, the EKF is relatively stable when estimating the states, and
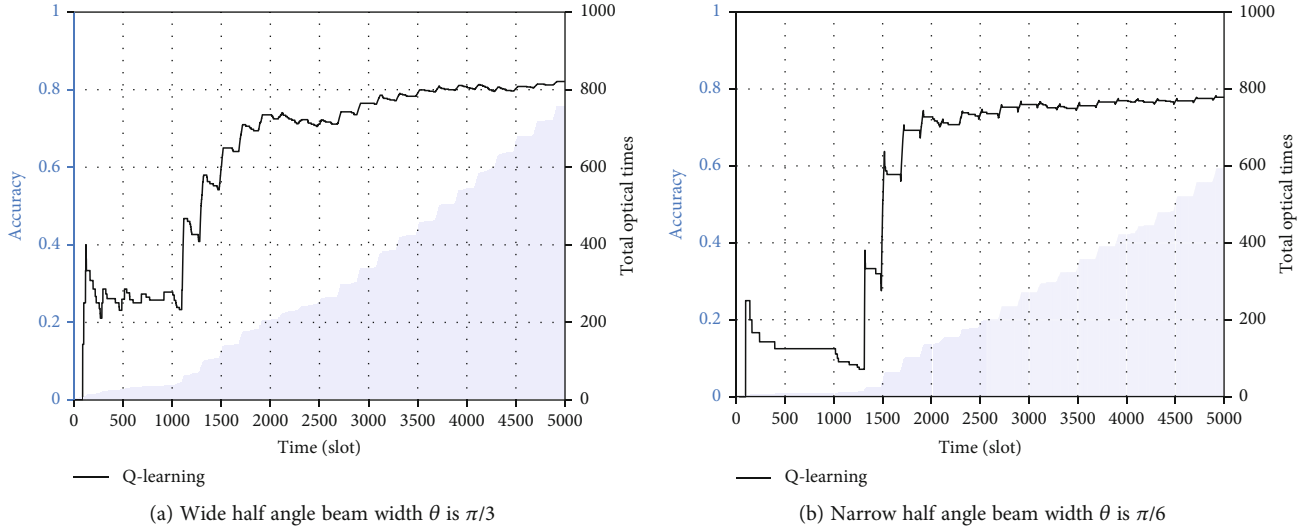
(a) Wide half angle beam width θ is $\pi/3$

(b) Narrow half angle beam width θ is $\pi/6$

FIGURE 10: The convergence of Q-learning process with the varying half angle beam width θ.
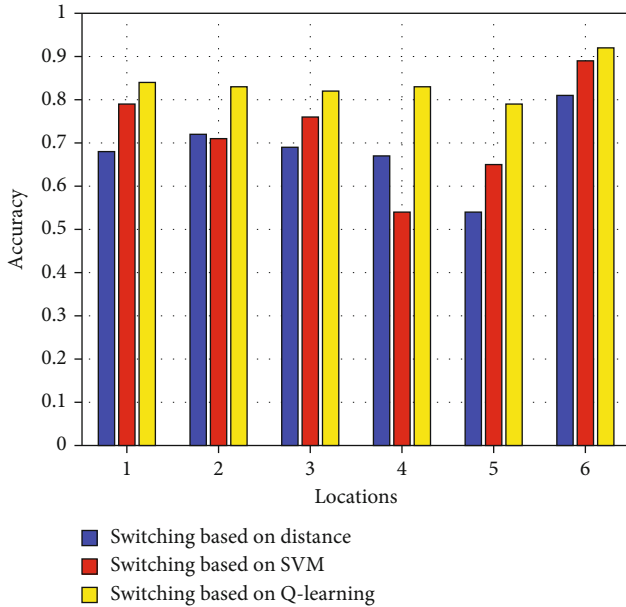

FIGURE 11: The comparison of the proposed method with other switching methods.

the actual states are generally match the estimated states, especially when there is no optical measured values.

The effectiveness of the Q-learning process is simulated and analyzed, and the corresponding results are shown in Figure 10. The accuracy is defined as the ratio of the successful switching time to the total switching trials. The initial learning rate δ is set to 0.01, and the reward discount γ is 0.7. When the optical communication is successful, the corresponding AUV gets the reward $\chi = 1$, and the punishment is $\varphi = 0.8$ for the other case. The ε of the greedy strategy is initially set to 0.06, and it decreases with the increasing level of accuracy.

To compare the performance of the learning process, we set different half angles in Figures 10(a) and 10(b) to

evaluate the convergence process of Q-function. Overall, Figure 10 shows that the accuracy of the switching scheme has significant improvement with the increasing training rounds. During the beginning stage of the training process, the optical communication delay becomes less due to the few number of training samples. However, after the convergence of the policy function, the optical communication delay is increasing, while the accuracy of switching policy is stabilized.

The figures also show that the half angle has an impact on the learning process. Compared to Figures 10(a) and 10(b), the narrow half angle means the shorter switching time window, and so the exploration trials at the beginning of the learning process are fewer. However, with the increasing of the training trials, the accuracy with narrow half angle converges to 80%, which is almost identical to the wide one.

Figure 11 shows the switching accuracy performance of the proposed method in yellow color compared with traditional methods which include distance-based method in blue color [41] and SVM-based switching method in red color [27]. There are six different underwater communication locations considered in our simulation which include harbor, rough sea, calm sea, calm sea with working boat, turbid waters, and clear waters corresponding to locations 1, 2, 3, 4, 7, and 8 in [27]. It is worth noting that there are different kinds of mechanical noise interference in the harbor, obvious spikes in the rough sea, fixed frequency noise in the calm sea with working boat, and large attenuation coefficient of beam in turbid water. As shown in Figure 11, the distance-based switching method is insensitive to acoustic noise but vulnerable to the turbidity of the water, and the SVM-based method depending on acoustic SNR is mainly affected by the acoustic noise. Compared to these traditional methods, our proposed method achieves more than 75% switching accuracy in all cases.

4.3. The Robustness for Switching Scheme. In this part, we evaluate the robustness of the switching scheme under
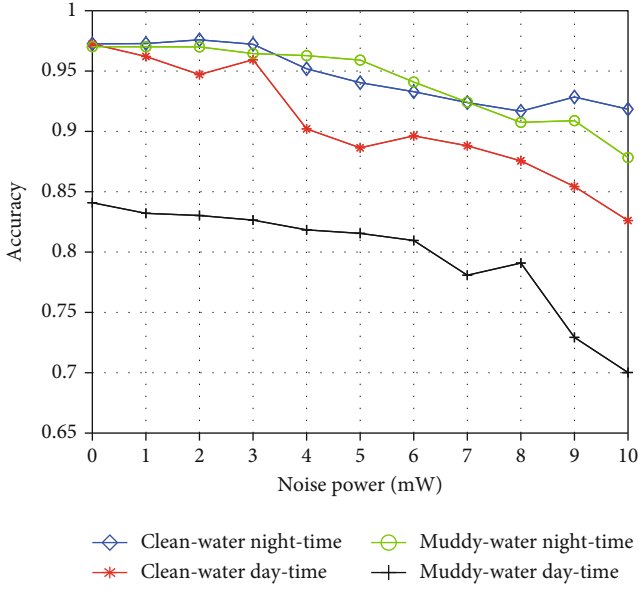
FIGURE 12: The switching accuracy under different environmental conditions.
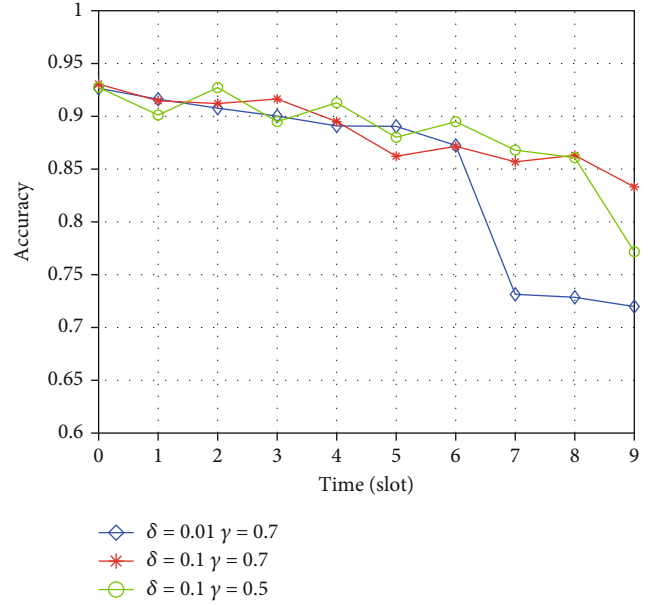

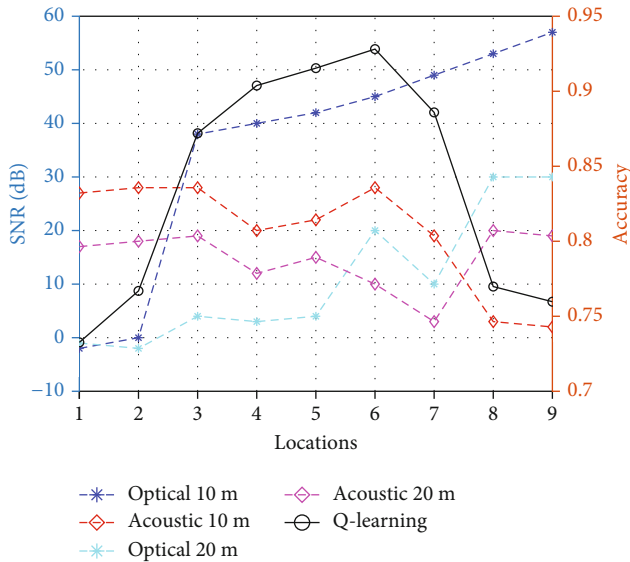FIGURE 14: The performance under different delays, learning rates, and discount factors.

FIGURE 13: The switching accuracy under different SNR levels.

influence and thus improves the adaptivity caused by the variation of optical attenuation coefficient.

Figure 13 shows the performance of the proposed scheme under different SNR levels. The left y-axis is the value of SNR sorting the optical SNR from the smallest to the largest according to the open dataset in [27] corresponding to locations 7, 8, 4, 5, 6, 3, 2, 9, and 1, respectively. The dotted line is defined as the mean value of SNR of the acoustic and optical channels in the communication range of 10 m and 20 m. The right y-axis is the switching accuracy, and the solid line shows the performance of the proposed method. It can be seen that the switching accuracy is lower when either the acoustic or optical SNR is weak. This is because the high attenuation coefficient of the optical channel can lead to transmission link instability, while the low acoustic SNR can lead to poor estimation of the optical channel. Overall, the proposed method has a good performance under different SNR levels.

The learning process generally involves the optical alignment procedure, and as the alignment process of the optical system needs a certain duration of time to perform the task, it requires the switching scheme to remain effective for that duration of time. Therefore, the time delay caused by the alignment procedure affects the successful ratio of the switching. As shown in Figure 14, the time tolerance of the Q-learning-based strategy is tested with a greedy coefficient of 0.01, and the average value of the accuracy is calculated with different time delays, learning rates, and discounter factors. As depicted in the figure, the larger time delay leads to less belief in the past experience, which implies that a larger learning rate has a better time tolerance. Since the discount factor indicates the weight to the future reward, as shown in the figure, a large discount factor γ has a better adaptation to time delay.

different types of water and sunlight-noise levels for realistic underwater wireless communication environments.

As the real underwater wireless communication environment is very complex, we focus on evaluating the robustness of the switching scheme under different noises and optical attenuation coefficients. The parameters related to the underwater environment are taken from the work in [42]. The simulated results are shown in Figure 12, and we can observe that the environmental noise affects the overall performance of the switching scheme. Moreover, as the radiation noise of sunlight increases, the performance of the switching scheme decreases. However, under the condition of strong noise, increasing the transmit power alleviates the

5. Conclusions

In this paper, we proposed an adaptive switching scheme for the underwater AUV-based acoustic-optical hybrid mobile network while combining the long-range but low-rate acoustic and the high-rate but short-range optical communications. In this scheme, we leveraged a RL-based method and the EKF tool to improve the adaptivity of the switching method. In response to the challenge associated with the intermittent feature of the optical channel, a blind estimation method based on EKF was proposed to estimate the optical channel state using the acoustic channel measurement. To deal with the harsh ocean environments, in the scheme, the relationship between the acoustic and optical channels, the channel variations and the mobility of the AUVs were considered and integrated into the learning process of the agent. We also conducted numerous simulations to verify the effectiveness and robustness of the proposed switching algorithm by considering the AUV speed, the environmental noise, the half angle beam width, and the optical alignment delay. In the future, we will apply multi-agent RL techniques in the switching scheme to improve the overall throughput further.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] H. Luo, K. Wu, R. Ruby, Y. Liang, Z. Guo, and L. M. Ni, "Software-defined architectures and technologies for underwater wireless sensor networks: a survey," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 4, pp. 2855–2888, 2018.

[2] H. Luo, J. Wang, B. Fanfeng, R. Ruby, K. Wu, and Z. Guo, "Recent progress of air/water cross-boundary communications for underwater sensor networks: a review," *IEEE Sensors Journal*, pp. 1–23, 2022.

[3] A. Peter, J. Sticklus, and A. Harlakin, "Underwater optical wireless communications in swarm robotics: a tutorial," *IEEE Communications Surveys & Tutorials*, vol. 23, no. 4, pp. 1–30, 2021.

[4] H. Esmaiel, Z. A. H. Qasem, H. Sun, J. Qi, J. Wang, and Y. Gu, "Wireless information and power transfer for underwater acoustic time-reversed Noma," *IET Communications*, vol. 14, no. 19, pp. 3394–3403, 2020.

[5] Z. A. H. Qasem, H. Esmaiel, H. Sun, J. Qi, and J. Wang, "Deep learning-based spread-spectrum fgsm for underwater communication," *Sensors*, vol. 20, no. 21, pp. 6123–6134, 2020.

[6] N. Saeed, A. Celik, T. Y. al-Naffouri, and M.-S. Alouini, "Underwater optical wireless communications, networking, and localization: a survey," *Ad Hoc Networks*, vol. 94, p. 101935, 2019.

[7] A. Celik, N. Saeed, B. Shihada, T. Y. Al-Naffouri, and M.-S. Alouini, "End-to-end performance analysis of underwater optical wireless relaying and routing techniques under location uncertainty," *IEEE Transactionson Wireless Communications*, vol. 19, no. 2, pp. 1167–1181, 2019.

[8] R. Ruby, S. Zhong, B. M. ElHalawany, H. Luo, and K. Wu, "SDN-Enabled energy-aware routing in underwater multimodal communication networks," *IEEE/ACM Transactions on Networking*, vol. 29, no. 3, pp. 965–978, 2021.

[9] S. Han, Y. Noh, U. Lee, and M. Gerla, "Optical-acoustic hybrid network toward real-time video streaming for mobile underwater sensors," *Ad Hoc Networks*, vol. 83, pp. 1–7, 2019.

[10] F. Campagnaro, R. Francescon, P. Casari, R. Diamant, and M. Zorzi, "Multimodal underwater networks: recent advances and a look ahead," in *In Proceedings of the International Conference on Underwater Networks & Systems*, pp. 1–8, New York, 2017.

[11] C. Moriconi, G. Cupertino, S. Betti, and M. Tabacchiera, "Hybrid acoustic/optic communications in underwater swarms," in *In OCEANS 2015-Genova*, pp. 1–9, Genova, Italy, 2015.

[12] N. Farr, A. Bowen, J. Ware, C. Pontbriand, and M. Tivey, "An integrated, underwater optical/acoustic communications system," in *In OCEANS 2010 IEEE-Sydney*, pp. 1–6, Sydney, NSW, Australia, 2010.

[13] F. Campagnaro, F. Guerra, F. Favaro et al., "Simulation of a multimodal wireless remote control system for underwater vehicles," in *InProceedings of the 10th International Conference on Underwater Networks & Systems*, pp. 1–8, New York, 2015.

[14] F. Campagnaro, F. Favaro, F. Guerra, V. S. Calzado, M. Zorzi, and P. Casari, "Simulation of multimodal optical and acoustic communications in underwater networks," in *In OCEANS 2015-Genova IEEE*, pp. 1–7, Genova, Italy, 2015.

[15] D. Zhang, I. N'Doye, T. Ballal, T. Y. Al-Naffouri, M. S. Alouini, and T. M. Laleg-Kirati, "Localization and tracking control using hybrid acoustic–optical communication for autonomous underwater vehicles," *IEEE Internet of Things Journal*, vol. 7, no. 10, pp. 10048–10060, 2020.

[16] T. Hu and Y. Fei, "Murao: a multi-level routing protocol for acoustic-optical hybrid underwater wireless sensor networks," in *In 9th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks (SECON)*, pp. 218–226, Seoul, Korea (South), 2012.

[17] E. P. M. C. Júnior, L. F. M. Vieira, and M. A. M. Vieira, "Captain: a data collection algorithm for underwater optical-acoustic sensor networks," *Computer Networks*, vol. 171, no. 2020, p. 107145, 2020.

[18] A. Celik, N. Saeed, B. Shihada, T. Y. Al-Naffouri, and M.-S. Alouini, "Opportunistic routing for opto-acoustic internet of underwater things," *IEEE Internet of Things Journal*, vol. 9, no. 3, pp. 1–15, 2021.

[19] J. Wang, W. Shi, L. Xu, L. Zhou, Q. Niu, and J. liu, "Design of optical-acoustic hybrid underwater wireless sensor network,"

Journal of Network and Computer Applications, vol. 92, pp. 59–67, 2017.

[20] Z. A. Qasem, H. A. Leftah, H. Sun, J. Qi, J. Wang, and H. Esmaiel, "Deep learning-based code indexed modulation for autonomous underwater vehicles systems," *Vehicular Communications*, vol. 28, no. 3, p. 100314, 2021.

[21] J. Connor, B. Champion, and M. A. Joordens, "Current algorithms, communication methods and designs for underwater swarm robotics: a review," *IEEE Sensors Journal*, vol. 21, no. 1, pp. 153–169, 2021.

[22] C. Lodovisi, P. Loreti, L. Bracciale, and S. Betti, "Performance analysis of hybrid optical–acoustic auv swarms for marine monitoring," *Future Internet*, vol. 10, no. 7, pp. 1–13, 2018.

[23] N. D. Hardy, H. G. Rao, S. D. Conrad et al., "Demonstration of vehicle-to-vehicle optical pointing, acquisition, and tracking for undersea laser communications," *In Free-Space Laser Communications SPIE*, vol. 10910, pp. 205–214, 2019.

[24] M. Kong, Y. Guo, M. Sait et al., "Toward automatic subsea operations using real time underwater optical wireless sensor networks," *IEEE Photonics Journal*, vol. 14, no. 1, pp. 1–9, 2021.

[25] S. K. Shrivastava, S. Sengar, and S. P. Singh, "A new switching scheme for hybrid fso/rf communication in the presence of strong atmospheric turbulence," *Photonic Network Communications*, vol. 37, no. 1, pp. 53–62, 2019.

[26] S. K. Shrivastava, S. Sengar, and S. P. Singh, "On the effect of incorrect channel condition information on modified switching scheme of hybrid fso/rf system," *IEEE Transactions on Cognitive Communications and Networking*, vol. 5, no. 4, pp. 1208–1217, 2019.

[27] R. Diamant, F. Campagnaro, M. De Filippo De Grazia et al., "On the relationship between the underwater acoustic and optical channels," *IEEE Transactions on Wireless Communications*, vol. 16, no. 12, pp. 8037–8051, 2017.

[28] J. Lin, S. Wei, X. Liang, and X. Jiang, "Adaptive modulation switching strategy based on q-learning for underwater acoustic communication channel," in *In Proceedings of the Thirteenth ACM International Conference on Underwater Networks & Systems*, pp. 1–5, New York, 2018.

[29] S. Wei, J. Lin, K. Chen, X. Liang, and E. Cheng, "Reinforcement learning-based adaptive modulation and coding for efficient underwater communications," *IEEE Access*, vol. 7, pp. 67539–67550, 2019.

[30] X. Sun, C. H. Kang, M. Kong et al., "A review on practical considerations and solutions in underwater wireless optical communication," *Journal of Lightwave Technology*, vol. 38, no. 2, pp. 421–431, 2020.

[31] M. Stojanovic, "On the relationship between capacity and distance in an underwater acoustic communication channel," *ACM SIGMOBILE Mobile Computing and Communications Review*, vol. 11, no. 4, pp. 34–43, 2007.

[32] X. Zhang, W. Ying, and B. Yang, "Parameter estimation for class a modeled ocean ambient noise," *Bandung Institute of Technology*, vol. 50, no. 3, pp. 330–345, 2018.

[33] X. Feng, M. Zhou, X. Zhang, K. Ye, J. Wang, and H. Sun, "Variational Bayesian inference based direction of arrival estimation in presence of shallow water non-Gaussian noise," *Journal of Electronics and Information Technology*, vol. 44, pp. 1–10, 2022.

[34] X. Zhang, W. Ying, P. Yang, and M. Sun, "Parameter estimation of underwater impulsive noise with the class B model,"

IET Radar, Sonar & Navigation, vol. 14, no. 7, pp. 1055–1060, 2020.

[35] P. Wang and L. Feng, "U-filter's Gaussianization function for interference background," in *In 2009 International Conference on Wireless Communications & Signal Processing*, pp. 1–4, Nanjing, China, 2009.

[36] J. Wang, J. Li, S. Yan et al., "A novel underwater acoustic signal denoising algorithm for Gaussian/non-Gaussian impulsive noise," *Transactions on Vehicular Technology*, vol. 70, no. 1, pp. 429–445, 2021.

[37] J. W. Giles and I. N. Bankman, "Underwater optical communications systems. part 2: basic design considerations," in *In MILCOM 2005-2005 IEEE Military Communications Conference*, pp. 1700–1705, Atlantic City, NJ, USA, 2005.

[38] G. Cario, A. Casavola, G. Gagliardi, M. Lupia, U. Severino, and F. Bruno, "Analysis of error sources in underwater localization systems," in *In OCEANS 2019-Marseille*, pp. 1–6, Marseille, France, 2019.

[39] A. Alalwan, T. Mohamed, M. Chakir, and T. M. Laleg, "Extended Kalman filter based linear quadratic regulator control for optical wireless communication alignment," *IEEE Photonics Journal*, vol. 12, no. 6, pp. 1–12, 2020.

[40] H. Bjørn-Olav, M. B. Eriksen, K. Y. Pettersen, and M. S. Wiig, "A modified dynamic window algorithm for horizontal collision avoidance for AUVs," in *In 2016 IEEE Conference on Control Applications (CCA),*, pp. 499–506, Buenos Aires, Argentina, 2016.

[41] F. Campagnaro, F. Guerra, P. Casari, R. Diamant, and M. Zorzi, "Implementation of a multi-modal acoustic-optical underwater network protocol stack," in *In OCEANS 2016-Shanghai*, pp. 1–6, Shanghai, China, 2016.

[42] L. J. Johnson, R. J. Green, and M. S. Leeson, "Underwater optical wireless communications: depth dependent variations in attenuation," *Applied Optics*, vol. 52, no. 33, pp. 7867–7873, 2013.