entropy

MDPI

Article

Passive Continuous Variable Measurement-Device-Independent Quantum Key Distribution Predictable with Machine Learning in Oceanic Turbulence

Jianmin Yi [1], Hao Wu [1] and Ying Guo [1,2,*]

[1] School of Automation, Central South University, Changsha 410083, China; 214612196@csu.edu.cn (J.Y.); 214612208@csu.edu.cn (H.W.)
[2] School of Computer Science, Beijing University of Posts and Telecommunications, Beijing 100876, China
* Correspondence: guoying@bupt.edu.cn

Abstract: Building an underwater quantum network is necessary for various applications such as ocean exploration, environmental monitoring, and national defense. Motivated by characteristics of the oceanic turbulence channel, we suggest a machine learning approach to predicting the channel characteristics of continuous variable (CV) quantum key distribution (QKD) in challenging seawater environments. We consider the passive continuous variable (CV) measurement-device-independent (MDI) QKD in oceanic scenarios, since the passive-state preparation scheme offers simpler linear elements for preparation, resulting in reduced interaction with the practical environment. To provide a practical reference for underwater quantum communications, we suggest a prediction of transmittance for the ocean quantum links with a given neural network as an example of machine learning algorithms. The results have a good consistency with the real data within the allowable error range; this makes the passive CVQKD more promising for commercialization and implementation.

Keywords: continuous variable quantum key distribution; measurement-device-independent; oceanic turbulence model; neural network

Citation: Yi, J.; Wu, H.; Guo, Y. Passive Continuous Variable Measurement-Device-Independent Quantum Key Distribution Predictable with Machine Learning in Oceanic Turbulence. *Entropy* **2024**, *26*, 207. https://doi.org/10.3390/e26030207

Academic Editor: Osamu Hirota

Received: 15 January 2024
Revised: 13 February 2024
Accepted: 26 February 2024
Published: 27 February 2024

1. Introduction

Quantum key distribution is a kind of encrypted means of communication [1,2], which uses the principle of quantum mechanics to enable legitimate parties to exchange secret keys securely. Continuous variable quantum key distribution (CVQKD) has been developed over decades due to its efficient source preparations and compatibility with current devices. Recently, a kind of meliorative protocol called the continuous variable measurement-device-independent (CV-MDI) protocol [3,4] has been proposed, in which a third party Charlie performs Bell state measurement on the quantum states prepared by Alice and Bob, and then broadcasts the result to Alice and Bob to generate the secret key. This detection strategy could counter an attack on practical devices because the measurement is performed by an untrusted third party rather than on Alice or Bob's side. However, for the classical CVQKD protocol, the quantum states are prepared actively, which requires high precision modulators to reduce modulation error and achieve a complex modulation format, making it expensive for practical implementations.

Currently, a kind of quantum key distribution has been suggested with passive state preparations [5]. Compared with active state preparations, which require high extinction ratio modulators, passive states can be derived from a thermal source for the passive CVQKD. If the initial thermal state generated by the source is strong enough, this scheme can tolerate high detector noise on Alice's side. Additionally, the output of the source does not need to be single-mode, as an optical homodyne detector can selectively measure a single mode determined by the local oscillator. Since then, passive state preparation has attracted much attention [6–9]. In 2018, passive states were applied to one-way classical

quantum communication [10], and this has been experimentally demonstrated [11,12]. There have been many results of passive state preparations in recent years, such as security analysis [13] and applications [14]. In 2019, passive states were used for the CV-MDI QKD protocol [15].

Over time, the CV-MDI system has expanded from the free space channel to the ocean quantum links [16–18]. However, in the implementation of the ocean quantum links, many factors, such as seawater salinity, oceanic turbulence, and chlorophyll concentration, have an affect on the propagation of light beams [19]. To solve these difficulties, we propose a machine learning-based prediction of ocean transmittance to provide data reference for engineering applications in practice. In recent years, in the field of QKD, machine learning has been paid more and more attention. In 2020, Z. A. Ren et al. employed machine learning methods to select an optimal QKD protocol [20]; in the same year, a random forests algorithm was used to directly predict the optimal parameters of the QKD system [21]. Two years later, Zhou et al. used neural networks to construct a secure key rate prediction model for discrete modulation continuous variable systems [22]. In 2023, Ahmadian. M et al. used machine learning to improve the polarization tracking compensation scheme of a QKD system [23]. The organization of this paper is as follows. In Section 2, CV-MDI QKD with passive state preparation is suggested. In Section 3, we analyze the characteristics of the oceanic channel and propose a machine learning-assisted model based on an oceanic turbulence model for transmittance prediction. In Section 4, the secret key rate in the oceanic scenario is derived. Section 5 shows the simulation results, and then Section 6 draws the conclusions.

2. CV-MDI QKD with Passive State Preparation

The Gaussian-modulated coherent states (GMCS) QKD protocol is implemented based on the prepare-and-measure scheme. And from Eve and Bob's points of view, the state from Alice is a single-mode thermal state with an average photon number of a half of modulation variance. In fact, the security of the GMCS QKD is commonly proved based on an equivalent entanglement-based protocol [24], where Alice performs conjugate homodyne detection on one mode of a two-mode squeezed vacuum state and sends the other mode to Bob. In this picture, the state from Alice is indeed thermal.

Here, we prepare the passive state by using a thermal source. There is a relationship between the value of the number of photons output at the Alice terminal and the modulation variance V in the GMCS QKD protocol, and the protocol with passive states requires a Gaussian modulator with a modulation variance of V. The preparation of passive states is implemented by taking advantage of a thermal source, beam splitters, optical attenuators, and homodyne detectors rather than the amplitude and phase modulators. The CV-MDI QKD protocol with passive state preparation is depicted in Figure 1, and its implementation can be described as follows.

Step 1: Alice and Bob each prepare a thermal sources. They use a 50:50 beam splitter to split the optical signal into two correlated spatial modes (the average number of photons output by each source is n_0), denoted by (Mod_{A1}, Mod_{A2}) (for Alice's side) and (Mod_{B1}, Mod_{B2}) (for Bob's side), respectively. Next, Alice (Bob) attenuates the average photon number of Mod_{A1} (Mod_{B1}) down to a half of the variance of V_A (V_B) by using an optical attenuator. The modulated signals are then transmitted to a third party, Charlie.

Step 2: Alice (Bob) performs heterodyne detection on both the X and P quadratures of mode Mod_{A2} (Mod_{B2}). They broadcast the measurement results to Charlie. The quadratures of Mod_{A1} (Mod_{B1}) at Charlie's side have the relation with Mod_{A2} (Mod_{B2}) as follows ($X_{A1} = \sqrt{\frac{2\eta_A}{\eta_D}} X_{A2}$, $P_{A1} = \sqrt{\frac{2\eta_A}{\eta_D}} P_{A2}$) and ($X_{B1} = \sqrt{\frac{2\eta_B}{\eta_D}} X_{B2}$, $P_{B1} = \sqrt{\frac{2\eta_B}{\eta_D}} P_{B2}$). Here, η_A and η_B represent the transmittance of the attenuator, while η_D represents the efficiency of the practical homodyne detector.

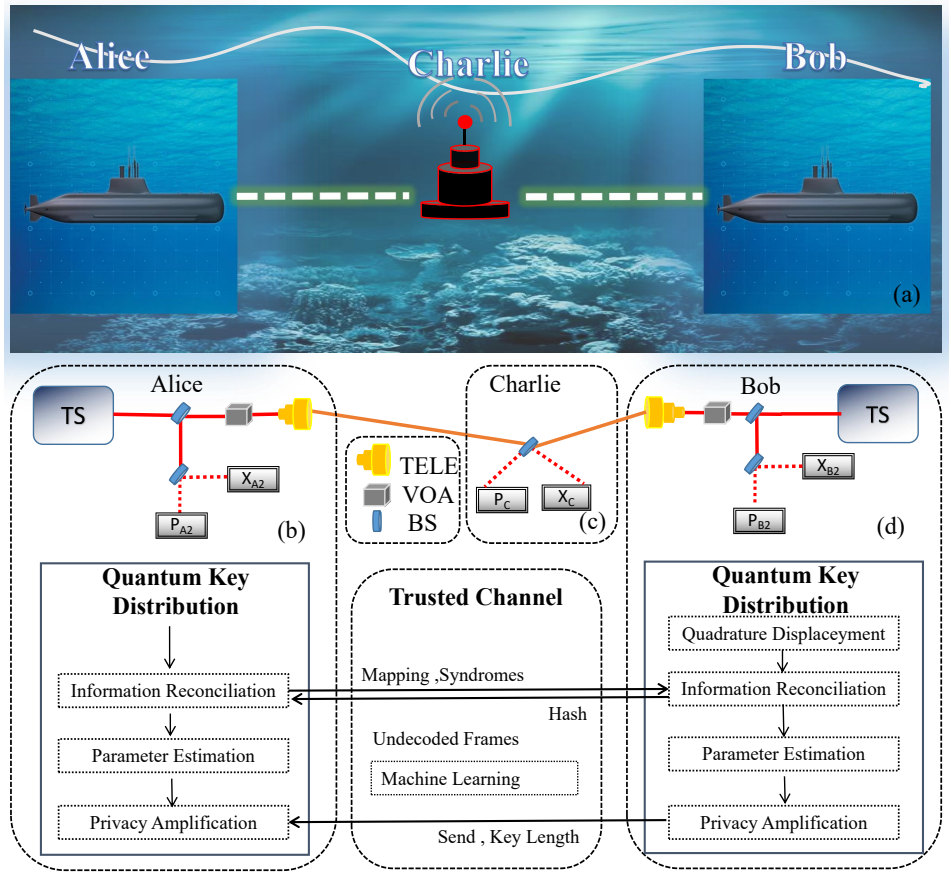

Figure 1. (**a**) Schematic diagram of the application of the underwater CV-MDI protocol. (**b**) Alice's side. (**c**) Charlie's side. (**d**) Bob's side. In the data processing stage, the machine learning module is used to predict transmittance. The specific explanation of this algorithm can be found in part 3. VOA, variable optical attenuator; TS, thermal source; the red dashed line denotes conjugate homodyne detection. TELE, telescope.

Step 3: Charlie mixes the received Mod_{A1} and Mod_{B1} on a balanced beam splitter and conducts Bell state measurement on them. The results are detected by conjugate homodyne detection at the output ports. After that, Charlie broadcasts the quadratures (X_C,P_C) over a classical public channel to Alice and Bob.

Step 4: After repeating these steps several times, Alice and Bob obtain a string of raw keys. Next, they apply post-processing operations such as privacy amplification and error correction to filter the data of (X_{A1}, P_{A1}), (X_{B1}, P_{B1}), and (X_C, P_C). Then, Alice and Bob get the final secret keys. The process is similar to the traditional CV-MDI QKD protocol with active state preparation, where communication parties can obtain final secret keys if the detected total noise falls below a certain threshold value. Compared with the Gaussian state, the passive state does not require the participation of a high-precision Gaussian modulator, which reduces the complexity and cost of the system.

3. Transmittance Prediction with Machine Learning

In this section, we first analyze the effect of the oceanic turbulence channel on light propagation, then we suggest a machine learning-based prediction model that can be used as a reference for practical underwater quantum communication systems.

3.1. Optical Propagation Characteristics of the Oceanic Turbulence Channel

Based on the seawater chlorophyll model and the elliptical model, which have been described in [25]—with the exception of the seawater extinction coefficient T—the Monte Carlo method used in the elliptic beam model for the oceanic turbulence channel has

general applicability to any other ocean. The seawater extinction coefficient is the sum of the ocean absorption factor t_{abs} and scattering factor t_{sca}, which have an effect on the absorption and scattering of light in the ocean. t_{abs} and t_{sca} are functions of the ocean depth d and wavelength λ, the specific function varies depending on the type of ocean, and we have analyzed the optical propagation characteristics in ocean type S1. Mathematically, t_{abs} and t_{sca} has the form:

$$\begin{aligned} t_{\text{abs}} &= l_c^0[u_c(d)]^{0.602} + l_w + l_f^0 u_f(d)\mathrm{e}^{-k_f\lambda} + l_h^0 u_h(d)\mathrm{e}^{-k_h\lambda}, \\ t_{\text{sca}} &= m_s^0 u_s(d) + m_l^0 u_l(d) + m_w, \end{aligned} \tag{1}$$

where l_w represents the absorption due to pure water in relation to wavelength λ, and l_w, corresponding to different wavelengths, is given by [26]. l_h^0 denotes the absorption coefficient of chlorophyll α in relation to λ, and l_h^0 corresponding to different wavelengths is given by [27]. The details of Equation (1) are given in Appendix A. By fitting the function, we get the functional relationship between l_w, l_h^0, and λ, respectively. It should be noted that this function does not contain quantum noise; parameters like the absorption coefficient of chlorophyll α, the loss of light propagation in pure water, etc. are all related physical factors that affect light propagation, and they quantified the effect of seawater on the propagation of light.

The relationship between wavelengths, depth, absorption factor, and scattering factor are given in Figure 2a,b.

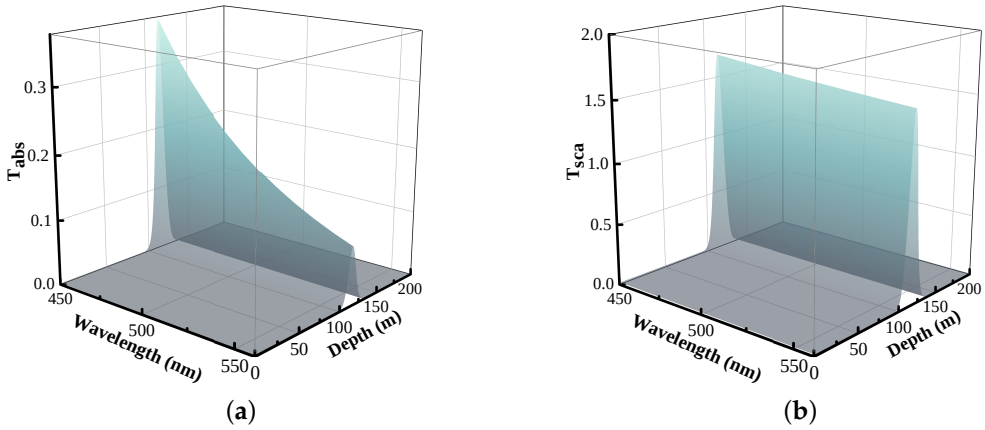

Figure 2. Variation of absorption and scattering factors with wavelength and depth. (**a**) Absorption factor. (**b**) Scattering factor.

The characteristics of S1 can be clearly seen in Figure 2. The absorption and scattering factors of the S1 ocean increase significantly at 100–150 m due to high chlorophyll concentration and plankton enrichment at this depth, which leads to a sharp decrease in the secret key rate near this depth, as detailed in Section 5.

3.2. Transmittance Prediction of Seawater Channel

In practice, the estimation of transmittance and excess noise requires the two legitimate parties to sacrifice part of the raw keys for the parameter estimation procedure; the more the raw keys are consumed, the more accurate the estimation of the transmittance and the excess noise is. However, sacrificing too many raw keys will affect the efficiency of communication. Meanwhile, the estimation for the transmittance in the parameter estimation is intended to estimate its lower bound as much as possible to ensure the absolute security of communication. In this scheme, we do not discard the parameter estimation step. Instead of the transmittance obtained from the parameter estimation step, we use the transmittance predicted by machine learning to participate in the estimation of the secret key rate; the former is more accurately close to the true value than the lower

bound of the transmittance (T_{low}) and thus we can obtain a higher secret key rate without sacrificing more raw keys.

This approach allows the CVQKD system to maintain stable performance in various environmental conditions, thereby improving the system's reliability. The structure of the Elman neural network is illustrated in Figure 3, and the prediction procedure is outlined in Figure 4.

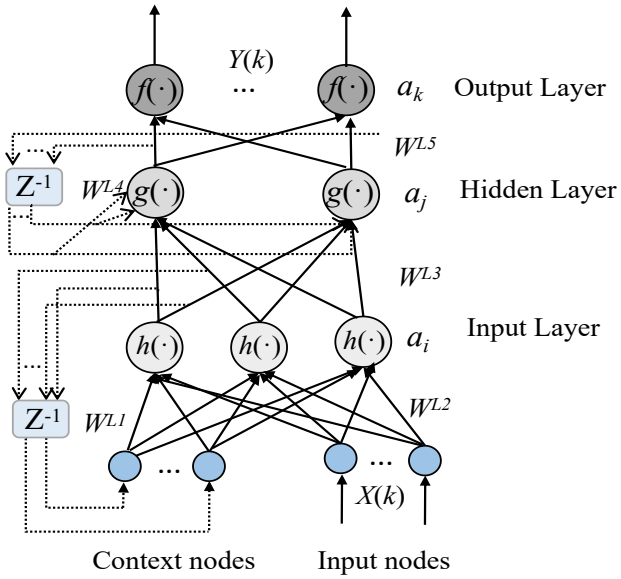

Figure 3. Structure of the Elman neural network. $X(\mathrm{k})$, $Y(\mathrm{k})$, input and output vectors; W^{Li} (where i = 1, 2,..., 5), connection weights; $h(\cdot)$, $g(\cdot)$, $f(\cdot)$, the nodal activation functions; a_i, a_j, a_k, the thresholds; Z^{-1}, the unity delay.

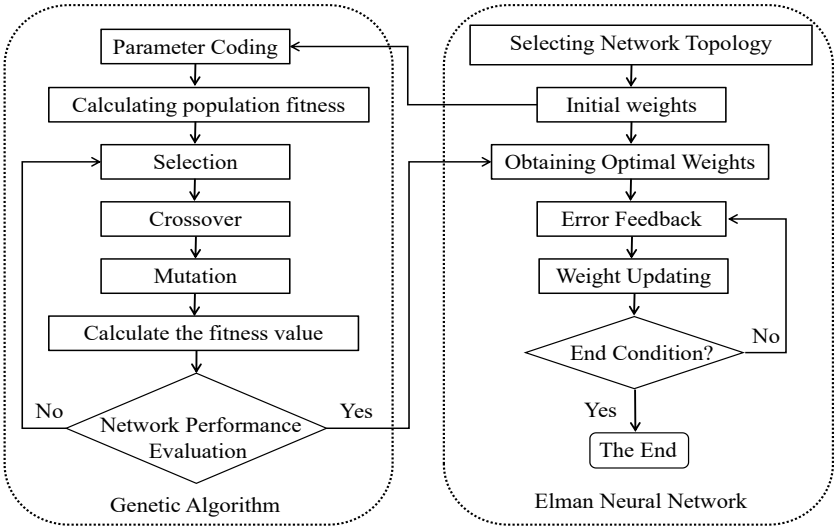

Figure 4. Flowchart of GA-Elman algorithm.

The Elman neural network is a type of recurrent neural network (RNN) that is used for time series prediction and sequence modeling [28]. It has a feedback loop that allows information from previous time steps to be fed back into the network, enabling it to capture temporal dependencies in the input sequence.

We present below an overview of its structure and training process. An Elman network typically consists of three main layers: an input layer, which receives external inputs and feeds them into the next layer; a hidden layer, which is a set of neurons that perform computations based on both current input and a context vector received from the previous time step. Each neuron has self-recurrent connections along with feedforward connections

from the input layer; an output layer, which processes the outputs of the hidden layer neurons to generate the final output; context units—a distinctive characteristic of the Elman network is the context layer or 'memory' units, which are a copy of the hidden layer activations at one time step, which are then fed back into the hidden layer during the next time step. This feedback loop enables the network to maintain some form of short-term memory that can influence its future predictions.

There are basic steps for training an Elman network. Initialization: Assign random initial weights and biases to all connections between layers; Forward Propagation: For each sequence step, input the current time step data into the input layer. The hidden layer computes its activations based on the current input and the previously stored context. The output layer generates its prediction using the hidden layer activations; Backpropagation through time (BPTT): After making predictions for an entire sequence, calculate the loss function comparing predicted outputs to target values across the sequence. BPTT extends standard backpropagation by unrolling the network over time and computing gradients through the unfolded network; Calculate the error gradient for each time step and update the weight matrices and bias vectors accordingly; Parameter Update: Using an optimization algorithm like stochastic gradient descent (SGD) or variants such as Adam, update the network parameters according to the computed gradients, aiming to minimize the total sequence loss; Iterative Training: Repeat this process over many iterations (epochs) until the performance on a validation set stabilizes or starts to degrade, indicating convergence or potential overfitting; Regularization: If necessary, apply regularization techniques to control model complexity and prevent overfitting. During the training process, it is crucial to monitor the learning curves, adjusting hyperparameters such as learning rate, batch size, and the number of hidden units if needed, to ensure efficient and accurate learning of temporal patterns in the data.

The GA-Elman algorithm [29] is a hybrid approach that combines the Elman recurrent neural network with an adaptive genetic algorithm to optimize the network's parameters for time series prediction. The main idea is to use the genetic algorithm to search for the optimal combination of weights and biases in the Elman network to minimize the prediction error. The algorithm starts by initializing the Elman network with random weights and biases. The training data are then fed into the network, and the output is computed. The genetic algorithm is used to optimize the weights and biases based on the prediction error. The genetic algorithm creates a population of candidate solutions, which are evaluated based on their fitness, i.e., how well they minimize the prediction error. The fittest solutions are selected for reproduction, and their offspring inherit their genetic traits through crossover and mutation. The Elman network is trained using the optimized weights and biases, and the process is repeated until the prediction error converges or the maximum number of iterations is reached. The trained network is then used to predict future values of the time series. The GA-Elman algorithm has several advantages over other time series prediction methods. It can handle nonlinear and non-stationary time series, and it can adapt to changing environments. The genetic algorithm allows for a global search of the parameter space, which can lead to better solutions than gradient-based methods.

The relationship between t_{abs}, t_{sca} and transmittance T is $T = e^{-(t_{abs}+t_{sca})z}$ [30], z is the transmission distance. Combined with Equation (1), we can find that the transmittance is a binary function whose independent variables are depth and transmission distance, so the inputs to the machine learning model are the depth and transmission distance.

The transmittance prediction of the Elman and GA-Elman algorithms on transmission can be seen in Figure 5. To provide a quantitative analysis of the performance improvements made to the Elman algorithm, we present the prediction errors of GA-Elman and Elman for 1200 sets of test samples, as shown in Figure 6. Prediction error refers to the difference between the value of the transmittance output of the predicted model and the real value when we input the depth and distance.

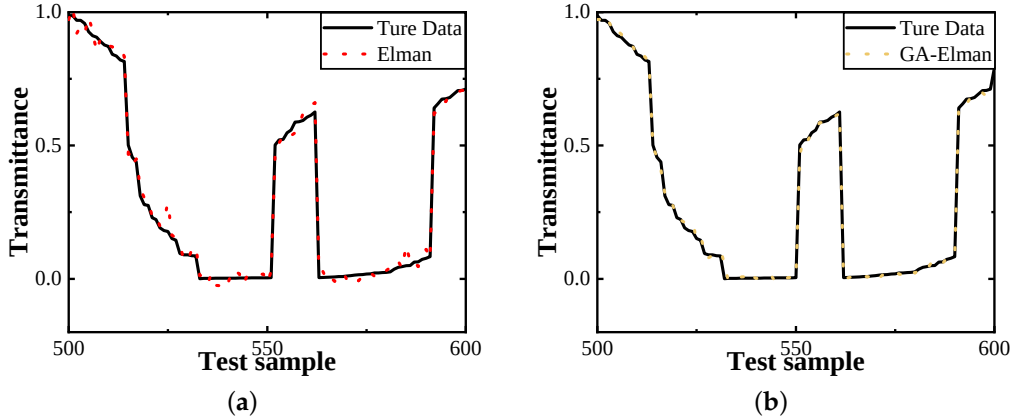

Figure 5. Predictions for transmittance. (**a**) Performance of Elman algorithm. (**b**) Performance of GA-Elman algorithm.

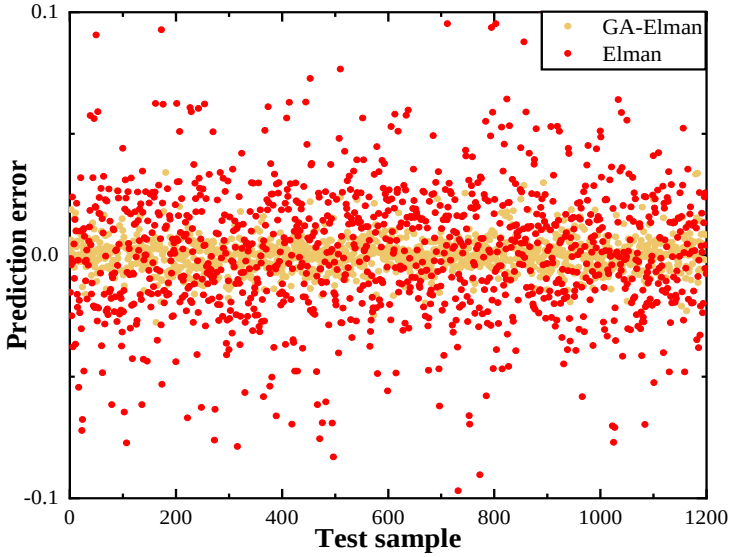

Figure 6. Prediction error of Elman and GA-Elman algorithm. The average absolute percentage error between Elman and GA-Elman are 2.814% and 0.506%, respectively.

The numerical analysis demonstrates that this model is capable of predicting the fluctuation of transmittance within an acceptable error range, which has a strong correlation with the actual transmission values. The prediction results can be used to assist actual derivation and calculation under certain circumstances.

The elliptic model provides the probability density function (PDF) of the transmittance, and an estimate of the transmittance is obtained by solving the inverse function of the cumulative distribution function, but the value of the actual measured transmittance can be any arbitrary value within the range of the PDF, and is not exactly equivalent to the former. According to [25], the variance of the transmittance is estimated to be on the order of 10^{-5}, with a transmittance of at least 0.4 or higher at effective underwater communication distances. Therefore, when the model predicts the transmittance based on actual measurements, the error in the model's prediction is within acceptable limits, even if the actual value of the transmittance at the predicted location is the value of the transmittance corresponding to a very small probability in the PDF.

The protocol performance under the transmittance prediction model based on machine learning is shown in Figure 7.

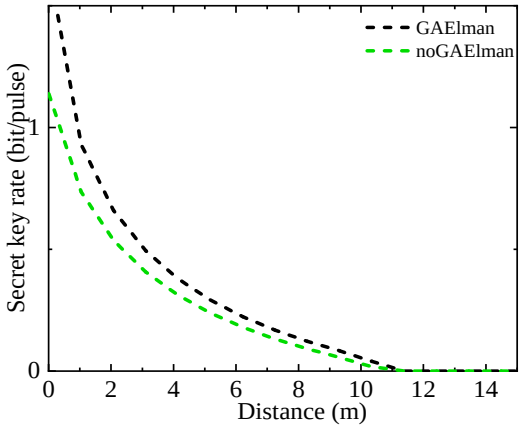

Figure 7. Performance improvement diagram of QKD system assisted by machine learning model.

The black dashed line represents the secret key rate curve after the transmittance predicted by machine learning is applied to the parameter estimation, and the green dashed line represents the secret key rate curve when the parameter estimation takes the lower bound of the transmittance without the application of the machine learning model.

4. Security Analysis

According to the above-mentioned processing, we obtain the transmittance in seawater channels, and hence we can establish the correlation among average transmittance, ocean depth, and transmission distance. Moreover, the passive state for the CV-MDI QKD protocol usually leads to excess noise, which provides an opportunity for eavesdropping through joint attacks. From [3], we assume that Eve adopts the most general joint attack against the protocol, which involves using the joint two-mode attack strategy that targets both links simultaneously. This approach is considered more effective than a single-mode attack strategy, which involves an additional layer of complexity to the security analysis.

4.1. Secret Key Rate in Asymptotic Scenarios

To begin with, we assume that the preparation sides have the same variance, thus the covariance matrix (CM) can be written as $V_{A_1B_1|C_1} = V_{A_1 \oplus B_1} - ZC^{-1}Z^T$, where $V_{A_1 \oplus B_1}$ represents the reduced covariance matrix (CM) of Alice and Bob's modes, while C denotes the outcome CM of Charlie, and Z represents the complex correlations between these CMs. Regarding the eavesdropping strategy, assuming a Gaussian distribution, Eve has the potential to intercept the traveling modes A_1 and B_1, which are mixed with two quantum-correlated ancillary modes. The reduced state $V_{E_1E_2}$ can be written in the normal form:

$$V_{E_1E_2} = \begin{pmatrix} \varpi_1 \mathrm{I}_2 & G \\ G & \varpi_2 \mathrm{I}_2 \end{pmatrix}; \tag{2}$$

ϖ_1 and ϖ_2 are the variance of the thermal excess noise disturbing the corresponding link, with $I_2 = \mathrm{diag}(1,1)$, with $G = \mathrm{diag}(\mathrm{g}, -\mathrm{g})$, where

$$g = \min\left[\sqrt{(\varpi_1 - 1)(\varpi_2 + 1)}, \sqrt{(\varpi_1 + 1)(\varpi_2 - 1)}\right] \tag{3}$$

is set to minimize the secret key rate. From [3,31], the simplified covariance matrix between Alice and Bob can be calculated as:

$$V_{A_1B_1|C} = \begin{bmatrix} \left(V - \frac{T_A(V^2-1)}{\vartheta}\right) I_2 & \frac{\sqrt{T_A T_B}(V^2-1)}{\vartheta}\sigma_Z \\ \frac{\sqrt{T_A T_B}(V^2-1)}{\vartheta}\sigma_Z & \left(V - \frac{T_B(V^2-1)}{\vartheta}\right) I_2 \end{bmatrix}, \tag{4}$$

with $\sigma_Z = \mathrm{diag}(1, -1)$, $V = V_A + 1 = V_B + 1$, and

$$\vartheta = V(T_A + T_B) + \varpi_1(1 - T_A) + \varpi_2(1 - T_B) - 2g\sqrt{(1 - T_A)(1 - T_B)}. \tag{5}$$

Let X_E denote the information that Eve can get by using the two-mode attack, and it is given by

$$X_E = S\left(\rho_{A_1B_1|C}\right) - S\left(\rho_{B_1|C\alpha}\right), \tag{6}$$

where $S\left(\rho_{A_1B_1|C}\right)$ and $S\left(\rho_{B_1|C\alpha}\right)$ can be calculated as:

$$\begin{aligned} S\left(\rho_{A_1B_1|C}\right) &= H(\lambda_1) + H(\lambda_2), \\ S\left(\rho_{B_1|C\alpha}\right) &= H\left[\sqrt{\det\left(V_{A_1B_1|C}\right)}\right], \end{aligned} \tag{7}$$

with $H(x) = \frac{1+x}{2}\log_2\left(\frac{1+x}{2}\right) - \frac{x-1}{2}\log_2\left(\frac{x-1}{2}\right)$. Here, λ_1 and λ_2 are the symplectic eigenvalues of $V_{A_1B_1|C}$. The mutual information between Alice and Bob is given by:

$$I_{AB} = \log_2 \frac{V}{X_{total}}, \tag{8}$$

where X_{total} can be divided into $X_{total} = X_{loss} + \varepsilon_E$. The pure loss in channel from senders to Charlie is defined as X_{loss}, which has the form $X_{loss} = 2\frac{T_A+T_B}{T_AT_B}$, and the total excess noise $\varepsilon_E = \varepsilon_P + \varepsilon_0$, where ε_0 is the background noise, and ε_P is the total excess noise in the process of passive state preparation. Therefore, we have:

$$\begin{aligned} \varepsilon_P &= \varepsilon_A + \varepsilon_B, \\ \varepsilon_A &= \frac{2\,\mathrm{V}_A}{\eta_D n_0}(1 + V_{el}) - \frac{V_A}{n_0}, \\ \varepsilon_B &= \frac{2\,\mathrm{V}_B}{\eta_D n_0}(1 + V_{el}) - \frac{V_B}{n_0}, \end{aligned} \tag{9}$$

where V_A and V_B are the modulation variance, V_{el} is the electronic noise of the homodyne detector, η_D is the efficiency of the homodyne detector, and n_0 is the average number of photons output by the thermal source.

4.2. Secret Key Rate in the Finite-Size Case

In the finite-size condition, the secret key rate is given by [32,33]:

$$K = \frac{n}{N}\left[K_\infty\left(T_A^{low}, T_B^{low}, \varepsilon_{X_C}^{high}, \varepsilon_{P_C}^{high}\right) - \Delta(n)\right], \tag{10}$$

where the signals exchanged by Alice and Bob are N. Due to the effects of finite size, Alice and Bob should conduct the parameter estimation by using a number of m keys in the practical condition. The remaining number n, which has the correlation with m, is given by $n = N - m$, which is used to generate the secret key. The correction term $\Delta(n)$ is simplified as:

$$\Delta(n) = 7\sqrt{\frac{\log_2 2/\epsilon_{P_A}}{n}}. \tag{11}$$

The estimation of error in privacy amplification ϵ_{P_A} is set to 10^{-10}. The noise terms of X_C and P_C has the form:

$$\varepsilon X_C = \varepsilon P_C = \varepsilon_P + \frac{1}{2}[\varpi_1(1 - T_A) + \varpi_2(1 - T_B)] - \mathrm{g}\sqrt{(1 - T_A)(1 - T_B)}, \tag{12}$$

$$\delta_{\varepsilon X_C} = \delta_{\varepsilon P_C} = \sqrt{\frac{2\varepsilon_{X_C}}{m}}. \tag{13}$$

The maximum noise of X_C and P_C generated in Charlie's detection is given by:

$$\begin{aligned} \varepsilon_{X_C}^{high} &= \varepsilon_{X_C} + 6.5\delta_{\varepsilon X_C}, \\ \varepsilon_{P_C}^{high} &= \varepsilon_{P_C} + 6.5\delta_{\varepsilon P_C}. \end{aligned} \tag{14}$$

Considering the security of the protocol, the channel transmittance is considered in the worst case because Alice and Bob are the same as Charlie and, consequently, only the case between Alice and Charlie is introduced. The expression can be conducted from:

$$T_A^{low} = \frac{1}{2}\left(X_2^{low} - X_1^{up}\right), \tag{15}$$

with

$$\begin{aligned} &X_2^{low} = X_2 - 6.5\sqrt{\mathrm{Var}(X_2)}, X_1^{up} = X_1 + 6.5\sqrt{\mathrm{Var}(X_1)}, \\ &X_1 = \langle T_A \rangle - \left\langle \sqrt{T_A} \right\rangle^2, X_2 = \langle T_A \rangle + \left\langle \sqrt{T_A} \right\rangle^2, \\ &\mathrm{Var}(X_1) = \mathrm{Var}(X_2) = \sigma^2_{\langle T_A \rangle} + 2\sigma^4_{\langle \sqrt{T_A} \rangle}\left[1 + 2\frac{\mu^2\sqrt{T_A}}{\sigma^2_{\langle \sqrt{T_A} \rangle}}\right]. \end{aligned} \tag{16}$$

Here, we have notations $\sigma^2_{\langle T_A \rangle} = \int (P(T_A))^2 \sigma^2_{\hat{T}_A}$, $\sigma^2_{\langle \sqrt{T_A} \rangle} = \int (P(T_A))^2 \mathrm{Var}\left(\sqrt{\hat{T}_A}\right)$, and $\mu_{\sqrt{T_A}} = \int P(T_A)\mathbb{E}\left(\sqrt{\hat{T}_A}\right)$. Taking $\sqrt{\hat{T}_A} = \sqrt{\frac{2(\hat{C}_{AC})^2}{\eta V_A^2}}$; we obtain

$$\begin{aligned} \delta^2_{\widehat{T_A}} &= \frac{\mathrm{Var}(\hat{C}_{AC})8T_A}{\eta V_A^2}, \\ \mathbb{E}\left(\sqrt{\hat{T}_A}\right) &= \sqrt{T_A}, \\ \mathrm{Var}\left(\sqrt{\hat{T}_A}\right) &= \frac{2\,\mathrm{Var}(\hat{C}_{AC})}{\eta V_A^2}. \end{aligned} \tag{17}$$

The variance of $\hat{C}_{AC}$ is $\frac{(\eta V_A^2 T_A + V_A V_N)}{m}$ and V_N is the variance of X_N.

5. Simulation Results

In what follows, we demonstrate the performance of the CV-MDI QKD system in terms of the secret key rate and transmission distance. In numerical simulations, we set the average number of output photons to 800 per pulse and the modulation variance to 60. For simplicity, we assume that the homodyne detector is noiseless and has an efficiency of 0.95 [15]. The communication block size is 10^8, and thermal noise $\varpi_1 \sim \varpi_2 \sim 1.01$. Depending on the distance between Alice and Bob, and Charlie, it can be classified as asymmetric and symmetric. It is worth mentioning that the system performs better in the asymmetric case when Alice and Charlie are closer to each other than in the symmetric case. Therefore, we will only present the performance in the former case.

The variation relationship between the secret key rate as the dependent variable and the transmission distance and depth as the independent variables is shown in Figure 8.

The title 'distance' in Figure 8 means the distance between Alice and Bob, namely the effective distance between two underwater communication parties. The green dashed line represents the change of the secret key rate with depth when the communication distance between the two communication parties is 1.65 m, so that the change of the dent of the three-dimensional surface in a certain depth segment can be more clearly seen, which just validates the previous analysis of the extinction coefficient, that is, the extinction coefficient increases sharply in this depth segment, and naturally the corresponding secret key rate should decline sharply in this depth segment.

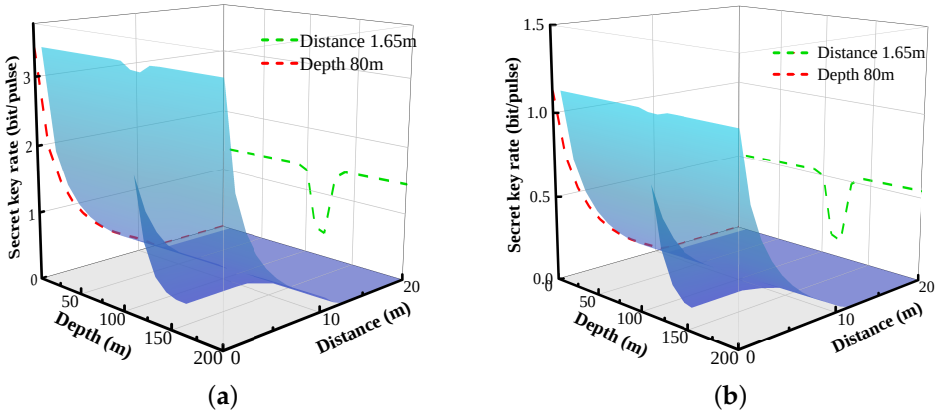

Figure 8. Secret key rate in asymmetric case.(**a**) Asymptotic case. (**b**) Finite-size case.

Additionally, we find how the secret key rate correlates with depth and transmission distance. The green curves demonstrate that the impact of ocean depth on the secret key rate is significant for a given transmission distance. This phenomenon is attributed to the presence of a strong optical fading effect at a specific depth in the ocean, which is denoted in Section 3. Therefore, it is advisable to avoid deploying communication devices in areas where the extinction factor is concentrated. The red curve illustrates the accepted phenomenon that the secret key rate decreases with increasing transmission distance for a given depth. The trend of the secret key rates in the finite-size case is similar to that in ideal conditions. However, the rates are lower due to the effects of the finite size.

6. Conclusions

We have proposed passive state CV-MDI to ocean scenarios. Then, we analyzed the optical propagation characteristics of the oceanic turbulence channel; moreover, we have presented a transmittance prediction model using the GA-Elman neural network. This model exhibits a high level of predictive accuracy for quantum communication in oceanic turbulence. The machine learning-assisted CVMDI protocol with passive states has improved its performance, although, limited by the complexity of the underwater environment and the attenuation of light propagation, the transmission distance has not been significantly improved, which is the limitation of this paper. However, the method for improving the lower bound of transmittance in parameter estimation by predicting transmittance is also suitable for free-space channels. In the future, with the proposal of quantum communication protocols with higher performance, it is expected to provide a new idea for auxiliary QKD systems.

The secret key rates in the asymptotic and finite-size cases are derived and the performance of the scheme is calculated. These above findings contribute to the advancement of passive CV-MDI QKD in challenging underwater environments. The ability to accurately predict transmittance in oceanic turbulence can enhance the security and reliability of quantum communication systems operating in such conditions.

Author Contributions: Conceptualization, J.Y. and Y.G.; methodology, J.Y. and H.W.; formal analysis, J.Y.; data curation, H.W.; writing—original draft preparation, J.Y.; writing—review and editing, H.W.; validation, Y.G. All authors have read and agree to the published version of the manuscript.

Funding: This research is backed by the National Natural Science Foundation of China (Grant Nos. 62103388), as well as the Natural Science Foundation of Hunan Province (Grant No. 2023JJ50268, 2023JJ50269) and the Key Research and Development Program of Hunan Province (Grant Nos. 2022GK2016) and the Scientific Research Fund of Hunan Provincial Education Department (Grant Nos. 22C0446, 21A0470, 22A0669).

Data Availability Statement: Data are contained within the article.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A. The Seawater Chlorophyll Model

Many factors such as seawater density, turbulence, bubble surface have important effects on light propagation in the ocean quantum links. From [25], the deterministic losses caused by the ocean extinction has an effect on the transmittance

$$T_{ext} = e^{-zt}, \tag{A1}$$

T_{ext} is the extinction-induced transmittance, z denotes the transmission distance and t is the seawater extinction coefficient which is related to the wavelength λ, and it is defined by $t = t_{abs} + t_{sca}$. t_{abs} is the ocean absorption factor which has the form

$$t_{abs} = l_c^0[u_c(d)]^{0.602} + l_w + l_f^0 u_f(d)\mathrm{e}^{-k_f\lambda} + l_h^0 u_h(d)\mathrm{e}^{-k_h\lambda}, \tag{A2}$$

u_c is the chlorophyll a content and it is defined as

$$u_c(d) = u_b + ds + \frac{h\sqrt{2\pi}}{\varsigma}\exp\left(-\frac{(d-d_{\max})^2}{2\varsigma^2}\right). \tag{A3}$$

The standard deviation of the concentration of chlorophyll ς is given by

$$\varsigma = \frac{h}{\sqrt{2\pi(u_{\mathrm{chl}} - u_b - d_{\max}s)}}, \tag{A4}$$

the content of fulvic acid is defined as $u_f(d) = 1.74098u_c(d)\mathrm{e}^{0.12327u_c(d)}$, the concentration of humic acid has the form $u_h(d) = 0.19334u_c(d)\mathrm{e}^{0.12343u_c(d)}$, where $t_{sca} = m_s^0 u_s(d) + m_l^0 u_l(d) + m_w$ is the scattering factor, and the small particles' concentration is defined as $u_s(d) = 0.01739u_c(d)\mathrm{e}^{0.11631u_c(d)}$, and $u_l(d) = 0.76284u_c(d)\mathrm{e}^{0.03092u_c(d)}$ is the large particles' concentration. the meaning and parameter of these variables are summarized in Table A1.

Table A1. Variables of ocean model.

	Meaning of the Variates	Parameter
l_c^0	The absorption coefficient of chlorophyll a at wavelength λ	0.009 m^2/mg
l_w	The loss of light propagation in pure water	0.0507 m^{-1}
l_f^0	The fulvic acid's absorption coefficient	35.959 m^2/mg
k_f	The fulvic acid's exponential coefficient	0.0189 nm^{-1}
λ	The wavelength	532 nm
l_h^0	The humic acid's absorption of coefficient	18.828 m^2/mg
k_h	The humic acid's exponential coefficient	0.01105 nm^{-1}
u_b	The surface's background chlorophyll content	0.0429 mg/m^3
s	The vertical gradient of concentration	-0.000103 mg/m^2
h	The total chlorophyll a above the background levels	11.87 mg
$d_{\max}$	The depth of the deep chlorophyll maximum	115.4 m
u_{chl}	The maximum chlorophyll concentration at the chlorophyll maximum layer	0.708 mg/m^3
m_s^0	The scattering coefficient of small particulate matter	$1.1513(400/\lambda)^{1.7}$
m_l^0	The scattering coefficient of large particulate matter	$0.3411(400/\lambda)^{0.3}$
m_w	The scattering coefficient of the pure water	$0.005826(400/\lambda)^{4.322}$
d	The depth of ocean	

References

1. Weedbrook, C.; Pirandola, S.; García-Patrón, R.; Cerf, N.J.; Ralph, T.C.; Shapiro, J.H.; Lloyd, S. Gaussian quantum information. *Rev. Mod. Phys.* **2012**, *84*, 621–699. [CrossRef]
2. Pirandola, S.; Andersen, U.L.; Banchi, L.; Berta, M.; Bunandar, D.; Colbeck, R.; Englund, D.; Gehring, T.; Lupo, C.; Ottaviani, C.; et al. Advances in quantum cryptography. *Adv. Opt. Photonics* **2020**, *12*, 1012–1236. [CrossRef]
3. Pirandola, S.; Ottaviani, C.; Spedalieri, G.; Weedbrook, C.; Braunstein, S.L.; Lloyd, S.; Gehring, T.; Jacobsen, C.S.; Andersen, U.L. High-rate measurement-device-independent quantum cryptography. *Nat. Photonics* **2015**, *9*, 397–402. [CrossRef]
4. Li, Z.; Zhang, Y.-C.; Xu, F.; Peng, X.; Guo, H. Continuous-variable measurement-device-independent quantum key distribution. *Phys. Rev. A* **2014**, *89*, 052301. [CrossRef]

5. Mauerer, W.; Silberhorn, C. Quantum key distribution with passive decoy state selection. *Phys. Rev. A* **2007**, *75*, 050305. [CrossRef]
6. Adachi, Y.; Yamamoto, T.; Koashi, M.; Imoto, N. Simple and efficient quantum key distribution with parametric down-conversion. *Phys. Rev. Lett.* **2007**, *99*, 180503. [CrossRef] [PubMed]
7. Zhang, Y.; Chen, W.; Wang, S.; Yin, Z.-Q.; Xu, F.-X.; Wu, X.-W.; Dong, C.-H.; Li, H.-W.; Guo, G.-C.; Han, Z.-F. Practical non-Poissonian light source for passive decoy state quantum key distribution. *Opt. Lett.* **2010**, *35*, 3393–3395. [CrossRef] [PubMed]
8. Curty, M.; Ma, X.; Lo, H.-K.; Lütkenhaus, N. Passive sources for the Bennett-Brassard. quantum-key-distribution protocol with practical signals. *Phys. Rev. A* **2010**, *82*, 052325. [CrossRef]
9. Sun, S.-H.; Tang, G.-Z.; Li, C.-Y.; Liang, L.-M. Experimental demonstration of passive-decoy-state quantum key distribution with two independent lasers. *Phys. Rev. A* **2016**, *94*, 032324. [CrossRef]
10. Qi, B.; Evans, P.G.; Grice, W.P. Passive state preparation in the Gaussian-modulated coherent-states quantum key distribution. *Phys. Rev. A* **2018**, *97*, 012317. [CrossRef]
11. Huang, P.; Wang, T.; Chen, R.; Wang, P.; Zhou, Y.; Zeng, G. Experimental continuous-variable quantum key distribution using a thermal source. *New J. Phys.* **2021**, *23*, 113028. [CrossRef]
12. Qi, B.; Gunther, H.; Evans, P.G.; Williams, B.P.; Camacho, R.M.; Peters, N.A. Experimental passive-state preparation for continuous-variable quantum communications. *Phys. Rev. Appl.* **2020**, *13*, 054065. [CrossRef]
13. Xu, S.; Li, Y.; Wang, Y.; Mao, Y.; Wu, X.; Guo, Y. Security Analysis of a Passive Continuous-Variable Quantum Key Distribution by Considering Finite-Size Effect. *Entropy* **2022**, *23*, 1698. [CrossRef] [PubMed]
14. He, Y.-Q.; Mao, Y.; Zhong, H.; Huang, D.; Guo, Y. Hybrid linear amplifier-involved detection for continuous variable quantum key distribution with thermal states. *Chin. Phys. B* **2020**, *29*, 050309. [CrossRef]
15. Bai, D.; Huang, P.; Ma, H.; Wang, T.; Zeng, G. Passive-state preparation in continuous-variable measurement-device-independent quantum key distribution. *J. Phys. B At. Mol. Opt. Phys.* **2019**, *52*, 135502. [CrossRef]
16. Feng, Z.; Li, S.; Xu, Z. Experimental underwater quantum key distribution. *Opt. Express* **2019**, *29*, 8725–8736. [CrossRef] [PubMed]
17. Hiskett, P.A.; Lamb, R.A. Underwater optical communications with a single photon-counting system. *Adv. Photon Count. Tech. VIII* **2014**, *9114*, 113–127.
18. Hu, C.-Q.; Yan, Z.-Q.; Gao, J.; Li, Z.-M.; Zhou, H.; Dou, J.-P.; Jin, X.-M. Decoy-state quantum key distribution over a long-distance high-loss air-water channel. *Phys. Rev. Appl.* **2021**, *15*, 024060. [CrossRef]
19. Guo, Y.; Xie, C.; Huang, P.; Li, J.; Zhang, L.; Huang, D.; Zeng, G. Channel-parameter estimation for satellite-to-submarine continuous-variable quantum key distribution. *Phys. Rev. A* **2018**, *97*, 052326. [CrossRef]
20. Ren, Z.; Chen, Y.; Liu, J.; Ding, H.; Wang, Q. Implementation of Machine Learning in Quantum Key Distributions. *IEEE Commun. Lett.* **2020**, *25*, 940–944. [CrossRef]
21. Ding, C.; Wang, Q. Predicting optimal parameters with random forest for quantum key distribution. *Quantum Inf. Process.* **2020**, *19*, 2. [CrossRef]
22. Zhou, M.; Liu, Z.; Liu, W.; Li, C.; Bai, J.; Xue, Y.; Fu, Y.; Yin, H.; Chen, Z. Neural network-based prediction of the secret-key rate of quantum key distribution. *Sci. Rep.* **2022**, *12*, 8879. [CrossRef] [PubMed]
23. Ahmadian, M.; Ruiz, M.; Comellas, J.; Velasco, L. Cost-Effective ML-Powered Polarization-Encoded Quantum Key Distribution. *J. Light. Technol.* **2023**, *13*, 4119–4128. [CrossRef]
24. Frédéric, G.; Gilles, V.A.; Jérôme, W.; Rosa, B.; Nicolas, C.; Philippe, J.G. Quantum key distribution using gaussian-modulated coherent states. *Nature* **2003**, *6920*, 238–241.
25. Zuo, Z.; Wang, Y.; Mao, Y.; Ruan, X.; Guo, Y. Security of quantum communications in oceanic turbulence. *Phys. Rev. A* **2021**, *104*, 052613. [CrossRef]
26. Haltrin, V.I.; Kattawar, G.W. *Effects of Raman Scattering and Fluorescence on Apparent Optical Properties of Seawater*; Texas A&M University: College Station, TX, USA, 1991.
27. Yentsch, C.S. The influence of phytoplankton pigments on the colour of sea water. *Deep. Sea Res. 1953* **1960**, *7*, 1–9. [CrossRef]
28. Elman, J.L. Finding Structure in Time. *Cogn. Sci.* **1990**, *2*, 179–211. [CrossRef]
29. Jia, W.; Zhao, D.; Zheng, Y.; Hou, S. A novel optimized GA–Elman neural network algorithm. *Neural Comput. Appl.* **2019**, *31*, 449–459. [CrossRef]
30. Xiang, Y.; Wang, Y.; Ruan, X.; Zuo, Z.; Guo, Y. Improving the discretely modulated underwater continuous-variable quantum key distribution with heralded hybrid linear amplifier. *Phys. Scr.* **2021**, *6*, 065103. [CrossRef]
31. Ottaviani, C.; Spedalieri, G.; Braunstein, S.L.; Pirandola, S. A Continuous-variable quantum cryptography with an untrusted relay: Detailed security analysis of the symmetric configuration. *Phys. Rev. A* **2015**, *91*, 022320. [CrossRef]
32. Ruppert, L.; Usenko, C.V.; Filip, R. Long-distance continuous-variable quantum key distribution with efficient channel estimation. *Phys. Rev. A* **2014**, *6*, 062310. [CrossRef]
33. Pirandola, O.S. Finite-size analysis of measurement-device-independent quantum cryptography with continuous variables. *Phys. Rev. A* **2017**, *96*, 4aPta1.