# Letter

## Communication-Aware Mobile Relaying via an AUV for Minimal Wait Time: A Broad Learning-Based Solution

Wenqiang Cao, Jing Yan, *Senior Member, IEEE*, Xian Yang, Cailian Chen, and Xinping Guan, *Fellow, IEEE*

Dear Editor,

This letter studies the communication-aware mobile relaying via an autonomous underwater vehicle (AUV) for minimal wait time. Compared with the analysis-based channel prediction solution, the proposed discrete Kirchhoff approximation solution has a higher estimation accuracy. Different with the deep learning (DL), a semi-supervised broad learning system (BLS) based relaying controller can reduce training time. Major contributions of this letter lie in two aspects: 1) Construct a BLS-based channel estimator with obstacle scattering effect, where the accurate estimated channel can be obtained with low computational cost; 2) Design a semi-supervised BLS based relay controller, such that average wait time can be minimized.

**Related works:** Recently, the mobile relaying of an AUV has been widely used in marine applications [1], [2]. Since the distribution of underwater channels is uneven, some scholars are committed to use the spatial distribution of channel quality to guide an AUV to relay data in different positions, e.g., [3], [4]. However, the above algorithms ignore the impact of obstacles on channel distribution. In [5], [6], the Kirchhoff approximation-based numerical methods were developed to capture the impact of obstacles on the channel, however they are not suitable for the large objects due to the huge computational workload. To handle this issue, an analytical method was proposed in [7] to study the scattering of texture details, but it is not suitable for studying the scattering effect of distant obstacles.

Apart that, another task is to design an appropriate controller for AUV to reach the relay position. We have noticed that the supervised or unsupervised BLS based controllers, such as [8], [9], can reduce an amount of training time and achieve better training results compared to the mainstream deep learning based controllers [10]. However, the above controllers depend on labels or precise model parameters. Considering the scattering effect of obstacles, how to design a semi-supervised BLS based relay controller to achieve the relaying task is an open issue.

**Problem statement:** The underwater relay system comprises $n$ infinite capacity queues, where each queue is a source and destination sensor pair. Data arrived stochastically the source node must be transferred to the corresponding destination node, which is too far away for direct communication. To improve the transmission success rates of queues, the relay operation is performed by an AUV.

The characteristics of the multi-queue system are summarized as the following two aspects: 1) Data accumulation: Data accumulates at $\mathbf{p}_{i,s}$ according to a Poisson process with average rate $\lambda_i$ bits per second (bps) in queue $i$. 2) Relaying service: When serviced by an AUV, queue $i$ transfers data from source node at a rate of $B\times\xi$ bits per second (bps), where $B$ Hz denotes the fixed bandwidth of queue and $\xi$ bps/Hz denotes the data upload/offload speed of the AUV.

Specifically, the position and velocity vectors of AUV is defined as $\boldsymbol{\eta}=[x,y,z,\psi]^T$ and $\mathbf{v}=[u,v,w,r]^T$, respectively. $\boldsymbol{\tau}=[\tau_u,\tau_v,\tau_w,\tau_r]^T$ is the control input vector including surge force $\tau_u$, sway force $\tau_v$, heave force $\tau_w$, and yaw force $\tau_r$. The discrete form of its dynamic model is given as

$$\boldsymbol{\zeta}(k+1)=\mathbf{f}(\boldsymbol{\zeta}(k))+\mathbf{h}(\boldsymbol{\zeta}(k))\boldsymbol{\tau}(k) \tag{1}$$

where $\mathbf{f}(\boldsymbol{\zeta})=[\boldsymbol{\eta}+\delta\mathbf{J}(\psi)\mathbf{v};\mathbf{v}-\delta\mathbf{M}^{-1}(\mathbf{C}(\mathbf{v})\mathbf{v}+\mathbf{D}(\mathbf{v})\mathbf{v}+\mathbf{g}(\boldsymbol{\eta}))]$, $\mathbf{h}(\boldsymbol{\zeta})=[\mathbf{0}_{4\times4};\,\delta\mathbf{M}^{-1}]$, $\boldsymbol{\zeta}=[\boldsymbol{\eta};\mathbf{v}]$, $\delta$ is sampling interval, $\mathbf{J}(\psi)$ is rotation matrix, $\mathbf{M}$ is inertia matrix, $\mathbf{C}(\mathbf{v})$ is Coriolis-centripetal matrix, $\mathbf{D}(\mathbf{v})$ is damping matrix, and $\mathbf{g}(\boldsymbol{\eta})$ is hydrostatic force.

The position set of source-destination sensor pairs is defined by $\mathcal{E}=\{\mathbf{p}_{1,s},\mathbf{p}_{1,d},\ldots\mathbf{p}_{i,s},\mathbf{p}_{i,d},\ldots,\mathbf{p}_{n,s},\mathbf{p}_{n,d}\}$. In addition, ordinary sensor nodes are randomly deployed in underwater to gather channel measurements, whose position set is defined by $\mathbf{p}_c=\{\mathbf{p}_{c,1},\ldots,\mathbf{p}_{c,j},\ldots,\mathbf{p}_{c,N}\}$. Accordingly, $i\in\mathcal{V}_q=\{1,\ldots,n\}$, $j\in\mathcal{V}_o=\{1,\ldots,N\}$ and $N$ is the number of ordinary nodes. As a result, the measured SNR between AUV and source/destination nodes in queue $i$ can be shown as follows, i.e.,

$$\begin{aligned}\mathrm{SNR}_{\mathrm{dB}}(\mathbf{p},\mathbf{p}_{i,b})=&\,P_0(\mathbf{p},\mathbf{p}_{i,b})-10l_{i,b}\log_{10}(\alpha(f))-N^0_{\mathrm{dB}}\\&+\sum_{\bar{\mathbf{p}}\in S}P_0(\mathbf{p},\bar{\mathbf{p}})\nabla_{\bar{\mathbf{p}}}G(\mathbf{p},\bar{\mathbf{p}})\times\vec{\mathbf{n}}_{\mathrm{out}}\end{aligned} \tag{2}$$

where $G(\mathbf{p},\bar{\mathbf{p}})=\frac{\exp(-\iota\kappa\|\mathbf{p}-\bar{\mathbf{p}}\|)}{4\pi\|\mathbf{p}-\bar{\mathbf{p}}\|}$, $P_0(\mathbf{p},\mathbf{p}_{i,b})=K_{\mathrm{dB}}-10n_{\mathrm{PL}}\log_{10}(l_{i,b})+\mu_{\mathrm{MP}}(\mathbf{p},\mathbf{p}_{i,b})+\sigma_{\mathrm{SH}}(\mathbf{p},\mathbf{p}_{i,b})$, $l_{i,b}=\|\mathbf{p}-\mathbf{p}_{i,b}\|$, $\iota=\sqrt{-1}$, $b\in\{s,d\}$, $\nabla_{\bar{\mathbf{p}}}G=\frac{\partial G}{\partial\bar{\mathbf{p}}}$ and $10\log_{10}\alpha(f)=\frac{0.11f^2}{1+f^2}+\frac{44f^2}{4100+f^2}+2.75\times10^{-4}f^2+0.003$. Meanwhile, $\bar{\mathbf{p}}$ denotes the discretized point on the surface $S$ and $\vec{\mathbf{n}}_{\mathrm{out}}$ is the unit normal vector at the correspondent position $\bar{\mathbf{p}}$. Noting that the parameters $N^0_{\mathrm{dB}}$, acoustic frequency $f$, $K_{\mathrm{dB}}$, $n_{\mathrm{PL}}$, the statistical characteristics of multipath $\mu_{\mathrm{MP}}$ and shadow fading $\sigma_{\mathrm{SH}}$ are obtained in advance. Conversely, $\kappa$ is an scattering parameter to be estimated.

Some necessary definitions and assumption are given as follows.

Definition 1: Transmission success rate $F_{i,\mathrm{sd}}(\mathbf{p})$ is the successful probability of data transmission on the queue $i$ when the position of AUV is $\mathbf{p}$. Once AUV moves to an optimum position $\mathbf{p}^*$, we get the maximum data transmission success rate $F^{\max}_{i,\mathrm{sd}}(\mathbf{p}^*)$.

Definition 2: Average wait time $\Gamma$ is the mean duration between the arrival of data at queues and the end of its transmission, including the total service $\Gamma_1$ and switch-over time $\Gamma_2$ among relay positions.

Definition 3: The exhaustive service strategy means AUV serves the queue $i$ until the queue $i$ is empty before serving queue $i+1$.

Assumption 1: The obstacles are regarded as rigid objects, i.e., sphere, cylinder and cone. The detected set at the $k$-th time step is defined as $\Omega_k\subset\{B_1(O_1,\rho_1,h_1),\ldots,B_m(O_m,\rho_m,h_m),\ldots\}$, where $B_m(O_m,\rho_m,h_m)$ is the $m$-th ($m=1,2,\ldots$) obstacle; $O_m$, $\rho_m$ and $h_m$ are its center, radius and height, respectively.

Define $\Gamma_{2,\min}$ as the minimum value of $\Gamma_2$. The purposes of mobile relaying task are

1) Maximize transmission success rate: $F_{i,\mathrm{sd}}(\mathbf{p})\to F^{\max}_{i,\mathrm{sd}}(\mathbf{p}^*)$.

2) Minimize average waiting time: $\Gamma_2\to\Gamma_{2,\min}$.

**Design and analysis:** With the known incident wave $\mathbf{P}_0$, the discretization measurement process of SNR is as follows: 1) Discrete sampling of obstacle surfaces: We get the discretized point $\bar{\mathbf{p}}\in S$ and corresponding unit vector $\vec{\mathbf{n}}_{\mathrm{out}}$ using the mesh discretization method. 2) Calculate scattering field: Add up the inner product of the gradient $\nabla_{\bar{\mathbf{p}}}G(\mathbf{p},\bar{\mathbf{p}})$ and $\vec{\mathbf{n}}_{\mathrm{out}}$. 3) Obtain the measurement vector of SNR. Accordingly, the measured $\mathrm{SNR}_{\mathrm{dB}}(\mathbf{p}_{c,j},\mathbf{p})$ for $j\in\mathcal{N}_k$ is stacked into a vector form, i.e.,

$$\mathbf{Y}_{\mathrm{dB}}=\mathbf{P}_0+\mathbf{P}_{\mathrm{sca}}-\varepsilon \tag{3}$$

where $\mathbf{Y}_{\mathrm{dB}}=[\mathrm{SNR}_{\mathrm{dB}}(\mathbf{p}_{c,1},\mathbf{p}),\ldots,\mathrm{SNR}_{\mathrm{dB}}(\mathbf{p}_{c,|\mathcal{N}_k|},\mathbf{p})]^T$, $\mathbf{P}_0=\mathbf{H}\theta+\mu_{\mathrm{MP}}+\sigma_{\mathrm{SH}}$, $\mathbf{P}_{\mathrm{sca}}=\sum_{\bar{\mathbf{p}}\in S}\mathbf{P}_0\odot\nabla_{\bar{\mathbf{p}}}G(\mathbf{p}_c,\bar{\mathbf{p}})\times\vec{\mathbf{n}}_{\mathrm{out}}$, $\varepsilon=[1,\|\mathbf{p}_{c,1}-\mathbf{p}\|;\ldots;1,\|\mathbf{p}_{c,|\mathcal{N}_k|}-\mathbf{p}\|][N^0_{\mathrm{dB}};10\log_{10}(\alpha(f))]$. Meanwhile, $\theta=[K_{\mathrm{dB}},\ n_{\mathrm{PL}}]^T$, $\mathbf{H}=[1,-10\log_{10}(\|\mathbf{p}_{c,1}-\mathbf{p}\|);\ldots;1,-10\log_{10}(\|\mathbf{p}_{c,|\mathcal{N}_k|}-\mathbf{p}\|)]$, $\sigma_{\mathrm{SH}}=[\sigma_{\mathrm{SH}}(\mathbf{p}_{c,1},\mathbf{p}),\ldots,\ \sigma_{\mathrm{SH}}(\mathbf{p}_{c,|\mathcal{N}_k|},\mathbf{p})]^T$ and $\mu_{\mathrm{MP}}=[\mu_{\mathrm{MP}}(\mathbf{p}_{c,1},\mathbf{p}),\ldots,\ \mu_{\mathrm{MP}}(\mathbf{p}_{c,|\mathcal{N}_k|},\mathbf{p})]^T$.

We employ a first-order linear differential equation to depict the estimation procedure of $\kappa$, i.e., $\dot{\kappa}=\tau_\kappa$, where $\tau_\kappa$ is the increment input. Specifically, the $k$-th step cost function is given as

$$g_1(\kappa(k),\tau_\kappa(k))=\Delta(k)^T\mathbf{Q}_1\Delta(k)+R_1\tau_\kappa^2(k) \tag{4}$$

Corresponding author: Jing Yan.

Citation: W. Cao, J. Yan, X. Yang, C. Chen, and X. Guan, "Communication-aware mobile relaying via an AUV for minimal wait time: A broad learning-based solution," *IEEE/CAA J. Autom. Sinica*, vol. 11, no. 3, pp. 797–799, Mar. 2024.

W. Cao and J. Yan are with the Institute of Electrical Engineering, Yanshan University, Qinhuangdao 066004, China (e-mail: cwq@stumail.ysu.edu.cn; jyan@ysu.edu.cn).

X. Yang is with the Institute of Information Science and Engineering, Yanshan University, Qinhuangdao 066004, China (e-mail: xyang@ysu.edu.cn).

C. Chen and X. Guan are with the Department of Automation, Shanghai Jiao Tong University, Shanghai 200240, China (e-mail: cailianchen@sjtu.edu.cn; xpguan@sjtu.edu.cn).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JAS.2023.124095

where $\Delta(k)=\mathbf{Y}_{\rm dB}-\mathbf{H}\theta+\varepsilon-\mathbf{P}'_{\rm sca}(\kappa(k))$, $\mathbf{P}'_{\rm sca}=\sum_{\bar{\mathbf{p}}\in S}\mathbf{H}\theta\odot\nabla_{\bar{\mathbf{p}}}G(\mathbf{p}_c,\bar{\mathbf{p}})\times\vec{\mathbf{n}}_{\rm out}$, $\mathbf{Q}_1$ is a positive definite matrix and $R_1$ is a coefficient controlling the cost of increments.

Based on (4), the Bellman equation for the value function is $V_1(\kappa(k))=g_1(\kappa(k),\tau_\kappa(k))+\gamma_1 V_1(\kappa(k+1))$, where $\gamma_1\in(0,1]$ denotes the discount factor for the future reward. Accordingly, the optimization problem of $\kappa$ can be conducted as

$$\tau_\kappa^*(k)=\mathrm{argmin}_{\tau_\kappa}\{g_1(\kappa(k),\tau_\kappa(k))+\gamma_1 V_1(\kappa(k+1))\}. \qquad (5)$$

Referring to [11], the BLS-based network structure includes the "feature nodes" $\mathbf{Z}_\hbar(k)=X(k)\mathbf{W}_\ell+\boldsymbol{\rho}_\ell$, $\ell\in\{1,\ldots,n_f\}$ and "enhancement nodes" $\mathbf{E}_\hbar(k)=\phi(\mathbf{Z}^f(k)\mathbf{W}_\hbar+\boldsymbol{\rho}_\hbar)$, $\hbar\in\{1,\ldots,n_e\}$, where $\mathbf{Z}^f=[\mathbf{Z}_1(k),\ldots,\mathbf{Z}_{n_f}(k)]$ is the set of feature nodes. $\mathbf{W}_\ell$, $\mathbf{W}_\hbar$, $\boldsymbol{\rho}_\ell$and $\boldsymbol{\rho}_\hbar$ are random weight and bias matrices for $\mathbf{Z}_\ell(k)$ and $\mathbf{E}_\hbar(k)$, respectively. $n_f$ and $n_e$ are the group numbers of $\mathbf{Z}_\ell(k)$ and $\mathbf{E}_\hbar(k)$, respectively. Accordingly, the network structure is defined as $\mathbf{S}(k)=[\mathbf{Z}^f(k)|\,\mathbf{E}^e(k)]$, where $\mathbf{E}^e(k)=[\mathbf{E}_1(k),\ldots,\mathbf{E}_{n_e}(k)]$.

Accordingly, $V_1(\kappa(k))$ at $s$-th iteration is replaced as $V_1^s(\kappa(k))=(\mathbf{W}_1^s)^T\mathbf{S}^T(k)$, where $\mathbf{W}_1$ denotes the weight for the basis function $\mathbf{S}(k)$. Thus, we obtain the optimal increment policy, i.e.,

$$\tau_\kappa^s(k+1)=-\frac{1}{2}\frac{\partial V_1^s(\kappa(k))}{\partial\kappa(k)}=-\frac{1}{2}\frac{\partial\mathbf{S}(k)}{\partial\kappa(k)}\mathbf{W}_1^s. \qquad (6)$$

Next, we employ the following update law for $\mathbf{W}_1^{s+1}$, i.e., $\mathbf{W}_1^{s+1}=(\boldsymbol{\Omega}\boldsymbol{\Omega}^T)^{-1}\boldsymbol{\Omega}\chi$, where $\boldsymbol{\Omega}=[\Omega_1,\ldots\Omega_{\bar{\nu}},\ldots]^T$, $\chi=[\chi_1,\ldots,\chi_{\bar{\nu}},\ldots]^T$, $\chi_{\bar{\nu}}=g_1(\kappa(\bar{\nu}),\tau_\kappa(\bar{\nu}))$ and $\Omega_{\bar{\nu}}=\mathbf{S}^T(k_{\bar{\nu}})-\gamma_1\mathbf{S}^T(k_{\bar{\nu}}+1)$ for $\bar{\nu}\in\{1,\ldots,\nu\}$. Meanwhile, $\nu$ denotes the total number of date set and $k_{\bar{\nu}}$ is the time instant for the $\bar{\nu}$-th data element. With this process, $\tau_\kappa^*(k)$ can be obtained.

After obtaining the optimal parameter $\kappa^*$, we have

$$\begin{aligned}&\mathrm{Avg}(\mathrm{SNR}_{\rm dB}(\mathbf{p},\mathbf{p}_{i,b})\\&\quad=\bar{\mathbf{h}}\boldsymbol{\theta}(1+\breve{\mathfrak{R}})-\boldsymbol{\Xi}^T\boldsymbol{\Phi}^{-1}(\mathbf{Y}_{\rm dB}+\varepsilon-\mathbf{H}\theta\odot(\mathbf{1}_{|\mathcal{N}_k|}+\mathfrak{R}^T))-\varepsilon_{i,b}\\&\mathrm{Var}(\mathrm{SNR}_{\rm dB}(\mathbf{p},\mathbf{p}_{i,b})=(\xi^2+\rho^2)(1+\breve{\mathfrak{R}})-\boldsymbol{\Xi}^T\boldsymbol{\Phi}^{-1}\boldsymbol{\Xi}\end{aligned} \qquad (7)$$

where $\bar{\mathbf{h}}=[1,-10\log_{10}(\|\mathbf{p}-\mathbf{p}_{i,b}\|)]$, $\mathfrak{R}=[\sum_{\bar{\mathbf{p}}\in S}\nabla_{\bar{\mathbf{p}}}\mathbf{G}(\mathbf{p_{c,1}},\bar{\mathbf{p}})\times\vec{\mathbf{n}}_{\rm out},\ldots,\sum_{\bar{\mathbf{p}}\in S}\nabla_{\bar{\mathbf{p}}}\mathbf{G}(\mathbf{p}_{c,|\mathcal{N}_k|},\bar{\mathbf{p}})\times\vec{\mathbf{n}}_{\rm out}]$, $\breve{\mathfrak{R}}=\sum_{\bar{\mathbf{p}}\in S}\nabla_{\bar{\mathbf{p}}}G(\mathbf{p},\bar{\mathbf{p}})\times\vec{\mathbf{n}}_{\rm out}$, $\boldsymbol{\Xi}=\xi^2\times[\exp(-\|\mathbf{p}-\mathbf{p}_{c,1}\|/\beta),\ldots,\exp(-\|\mathbf{p}-\mathbf{p}_{c,|\mathcal{N}_k|}\|/\beta)]^T$ and $\varepsilon_{i,b}=10l_{i,b}\log_{10}\times(\alpha(f))+N^0_{\rm dB}$. Proof is shown in Theorem 1.

With the predicted channel, a semi-supervised BLS-based relay controller is proposed to achieve the relay task, as shown in Fig. 1. Under the exhaustive strategy, there is a specific value $\Gamma_{2,\min}=\sqrt{(\bar{\rho}(1-\bar{\rho})\varpi)/(\bar{\rho}^2-\sum_{i=1}^n\bar{\rho}_i^2)}$ that minimizes $\Gamma$, where $\bar{\rho}_i=\lambda_i B\xi$, $\bar{\rho}=\sum_{i=1}^n\bar{\rho}_i$, $\varpi=\sum_{i=1}^n\varpi_i^2$. Meanwhile, $\bar{\rho}_i$ is the traffic of the queue $i$ and $\varpi_i^2$ is the variance of the switch-over time.

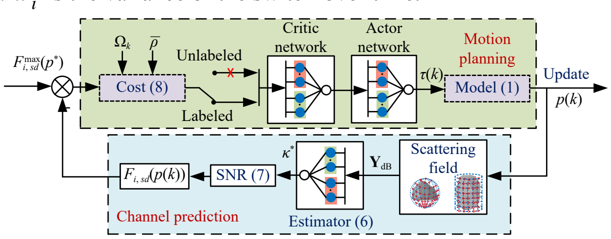


Fig. 1. Relationship between channel prediction and mobile relaying.

The cost function of motion planning is defined as

$$g_2(\mathbf{p}(k),\boldsymbol{\tau}(k))=\mathfrak{b}_1C_1(\mathbf{p}(k))+\mathfrak{b}_2C_2(k)+\mathfrak{b}_3C_3(\mathbf{p}(k))+\boldsymbol{\tau}^T(k)\mathbf{R}\boldsymbol{\tau}(k) \qquad (8)$$

where $C_1(\mathbf{p}(k))=\sum_{i=1}^n\frac{(1+\tanh(a_i(F_{i,\rm sd}(\mathbf{p}(k))-\epsilon_{i,1}))}{F_{i,\rm sd}^2(\mathbf{p}(k))}\times(1-\tanh(a_i(F_{i,\rm sd}(\mathbf{p}(k))-\epsilon_{i,2}))$, $C_2(\mathbf{p}(k))=\sum_{m\in\Omega_k}\frac{1}{(d_m-\rho_m)^2}$, $C_3(k)=(k-\Gamma_{2,\min})^2$ and $F_{i,\rm sd}(\mathbf{p}(k))=P_{i,s}(\mathbf{p}(k))\times P_{i,d}(\mathbf{p}(k))$. $\mathfrak{b}_1$, $\mathfrak{b}_2$, $\mathfrak{b}_3$, $a_i$, $\epsilon_{i,1}$ and $\epsilon_{i,2}$ are positive constants. $d_m=\|\mathbf{p}(k)-O_m\|$. Meanwhile, $\mathbf{R}=\mathrm{diag}\{[r_u,r_v,r_w,r_r]\}$ is a positive definite matrix. Of note, $P_{i,b}(\mathbf{p})=\mathbb{Q}\Big(\frac{\varrho-\mathrm{Avg}(\mathrm{SNR}_{\rm dB}(\mathbf{p},\mathbf{p}_{i,b}))}{\mathrm{Var}(\mathrm{SNR}_{\rm dB}(\mathbf{p},\mathbf{p}_{i,b}))}\Big)$, where $\mathbb{Q}(\cdot)$ is the complementary cumulative distribution function and $\varrho$ is a success rate threshold.

The Bellman equation for the value function of motion planning is $V_2(\boldsymbol{\zeta}(k))=g_2(\mathbf{p}(k),\boldsymbol{\tau}(k))+\gamma_2V_2(\boldsymbol{\zeta}(k+1))$, where $\gamma_2\in(0,1]$ denotes the discount factor for the future rewards. Accordingly, the optimization problem can be organized as

$$\boldsymbol{\tau}^*(k)=\mathrm{argmin}_{\boldsymbol{\tau}}(g_2(\mathbf{p}(k),\boldsymbol{\tau}(k))+\gamma_2V_2(\boldsymbol{\zeta}(k+1))). \qquad (9)$$

To solve (9), an admissible policy is used to collect labeled data $\{\mathbf{p}(k_\#),\ \Psi(k_\#)\}_{k_\#=1}^{k_\#^{\max}}$ and unlabeled data $\{\mathbf{p}(k_\#)\}_{k=k_\#^{\max}+1}^{k_{\max}}$, where $\Psi(k_\#)=g_2(\mathbf{p}(k_\#),\boldsymbol{\tau}(k_\#))$, $k_\#\in\{1_\#,\ldots,k_\#^{\max}\}$ is the $k_\#$-th sampling and $k_\#^{\max}$ is the total number of sampling.

Similarly, $V_2^{(s)}(\boldsymbol{\zeta}(k_\#))=\mathbf{W}_{cri}^{(s)T}\mathbf{S}_{cri}^T(k_\#)$ and $\boldsymbol{\tau}^{(s+1)}(k_\#)=[\tau_u^{(s+1)},\tau_v^{(s+1)},\tau_w^{(s+1)},\tau_r^{(s+1)}]^T$, where $\tau_\hbar^{(s+1)}=\mathbf{W}_\hbar^{(s+1)T}\mathbf{S}_\hbar^T(k_\#)$ denotes the actor network $\hbar\in\{u,v,w,r\}$ of AUV, $\mathbf{W}_{cri}^{(s)}$ and $\mathbf{W}_\hbar^{(s+1)}$ denote the $s$-th and $(s+1)$-th iterations of weights for the basis functions $\mathbf{S}_{cri}(k_\#)$ and $\mathbf{S}_\hbar(k_\#)$, respectively. The Bellman equation at the $s$-th iteration is $\Psi^{(s)}(k_\#)=-\mathbf{W}_{cri}^{(s)T}(\mathbf{S}_{cri}^T(\boldsymbol{\zeta}(k_\#+1))-\mathbf{S}_{cri}^T(\boldsymbol{\zeta}(k_\#)))-2\sum_\hbar\mathbf{W}_\hbar^{(s+1)T}\mathbf{S}_\hbar^T(\boldsymbol{\zeta}(k_\#))\times r_\hbar\bar{\tau}_\hbar(k_\#)=\mathbf{W}^{(s+1)T}\times\bar{\boldsymbol{\Omega}}(\boldsymbol{\zeta}(k_\#))$, where $\bar{\tau}_\hbar(k_\#)=\tau_\hbar(k_\#)-\tau_\hbar^{(s)}(k_\#)$, $\bar{\boldsymbol{\Omega}}(\boldsymbol{\zeta}(k_\#))=[\Delta\mathbf{S}_{cri}(\boldsymbol{\zeta}(k_\#));\ -2\mathbf{S}_u^T(\boldsymbol{\zeta}(k_\#))r_u\bar{\tau}_u(k_\#);\ -2\mathbf{S}_v^T(\boldsymbol{\zeta}(k_\#))r_v\bar{\tau}_v(k_\#);\ -2\mathbf{S}_w^T(\boldsymbol{\zeta}(k_\#))r_w\bar{\tau}_w(k_\#);-2\mathbf{S}_r^T(\boldsymbol{\zeta}(k_\#))r_r\bar{\tau}_r(k_\#)]$, $\mathbf{W}^{(s+1)}=[\mathbf{W}_{cri}^{(s)};\mathbf{W}_u^{(s+1)};\mathbf{W}_v^{(s+1)};\mathbf{W}_w^{(s+1)};\mathbf{W}_r^{(s+1)}]$ and $\Delta\mathbf{S}_{cri}(\boldsymbol{\zeta}(k_\#))=\mathbf{S}_{cri}^T(\boldsymbol{\zeta}(k_\#))-\mathbf{S}_{cri}^T(\boldsymbol{\zeta}(k_\#+1))$ is the difference in critic network.

Combined with $\bar{\boldsymbol{\Omega}}$ and (9), the weight update is solved by

$$\begin{aligned}\mathbf{W}^{(s+1)}=\mathrm{argmin}_{\mathbf{W}^{(s)}}\Big\{&\frac{\mathfrak{b}_4}{2}\big\|\vec{\boldsymbol{\Omega}}\mathbf{W}^{(s)}-\mathbf{Y}\big\|^2+\frac{1}{2}\big\|\mathbf{W}^{(s)}\big\|^2\\&+\frac{\mathfrak{b}_5}{2}tr(\mathbf{W}^{(s)\rm T}\vec{\boldsymbol{\Omega}}^T\mathbf{L}\vec{\boldsymbol{\Omega}}\mathbf{W}^{(s)})\Big\}\end{aligned} \qquad (10)$$

where $\vec{\boldsymbol{\Omega}}=[\bar{\boldsymbol{\Omega}}(\boldsymbol{\zeta}(1)),\ldots,\bar{\boldsymbol{\Omega}}(\boldsymbol{\zeta}(k_\#^{\max})),\ldots\bar{\boldsymbol{\Omega}}(\boldsymbol{\zeta}(k_{\max}))]^T$, $\mathbf{L}=\mathrm{diag}([\sum_{\vec{j}}^{k_{\max}}w_{1\vec{j}},\ldots,\sum_{\vec{j}}^{k_{\max}}w_{k_{\max}\vec{j}}])-[w_{\vec{i}\vec{j}}]_{k_{\max}\times k_{\max}}$, $w_{\vec{i}\vec{j}}=\exp(-0.5\|\mathbf{p}(\vec{i})-\mathbf{p}(\vec{j})\|/\bar{\sigma}^2)$, $\mathbf{Y}=[\Psi^{(s)}(0),\ldots,\Psi^{(s)}(k_\#^{\max}),0,\ldots,0]^T$, $\mathfrak{b}_4$, $\mathfrak{b}_5$ are positive constants, $\bar{\sigma}^2$ is the variance of Gaussian function and $\vec{i}$, $\vec{j}\in\{1,2,\ldots,k_{\max}\}$.

To solve (10), we set the gradient of (10) to 0, i.e.,

$$\Big(\mathfrak{b}_4\vec{\boldsymbol{\Omega}}^T\vec{\boldsymbol{\Omega}}+\mathbf{I}_{n_0}+\mathfrak{b}_5\vec{\boldsymbol{\Omega}}^T\mathbf{L}\vec{\boldsymbol{\Omega}}\Big)\mathbf{W}^{(s)}=\mathfrak{b}_4\vec{\boldsymbol{\Omega}}^T\mathbf{Y} \qquad (11)$$

where $\mathbf{I}_{n_0}$ is the $n_0$-dimensional diagonal identity matrix and $n_0$ is the number of neurons in the activation function $\bar{\boldsymbol{\Omega}}(\boldsymbol{\zeta}(\cdot))$.

If $k_{\max}>n_0$, $\mathbf{W}^{(s+1)}$ is calculated by considering (11), i.e., $\mathbf{W}^{(s+1)}=\mathfrak{b}_5(\mathfrak{b}_5\vec{\boldsymbol{\Omega}}^T\vec{\boldsymbol{\Omega}}+\mathbf{I}_{n_0}+\mathfrak{b}_6\vec{\boldsymbol{\Omega}}^T\mathbf{L}\vec{\boldsymbol{\Omega}})^{-1}\vec{\boldsymbol{\Omega}}^T\mathbf{Y}$, and hence, $\mathbf{W}^{(s+1)}$ can be obtained by the alternative solution, i.e., $\mathbf{W}^{(s+1)}=\vec{\boldsymbol{\Omega}}^T(\mathfrak{b}_6\mathbf{L}\vec{\boldsymbol{\Omega}}\vec{\boldsymbol{\Omega}}^T+\mathbf{I}_{k_{\max}}+\vec{\boldsymbol{\Omega}}\vec{\boldsymbol{\Omega}}^T)^{-1}\mathbf{Y}$.

Once the condition $\|\mathbf{W}^{(s+1)}-\mathbf{W}^{(s)}\|<\bar{\epsilon}_2$ is met, the optimal weight vector $\mathbf{W}^*$ can be obtained, where $\bar{\epsilon}_2>0$ is a threshold. Based on this, the optimal policy can be obtained, i.e.,

$$\boldsymbol{\tau}^*=\big[\mathbf{W}^{*T}\mathbf{S}_u^T;\mathbf{W}^{*T}\mathbf{S}_v^T;\mathbf{W}^{*T}\mathbf{S}_w^T;\mathbf{W}^{*T}\mathbf{S}_r^T\big]. \qquad (12)$$

Theorem 1: Given the channel measurement vector $\mathbf{Y}_{\rm dB}$ with obstacle scattering effect, which satisfies $\mathbf{Y}_{\rm dB}\sim\mathcal{N}(\mathbf{H}\theta-\varepsilon+\mathbf{P}'_{\rm sca},\ \boldsymbol{\Phi})$ and the predicted mean and variance of SNR are given as (7).

Proof: We employ the measurement vector $\mathbf{Y}_{\rm dB}$ to deduce the conditional probability density function of $\mathrm{SNR}_{\rm dB}(\mathbf{p},\mathbf{p}_{i,b})$.

Specifically, the probability density function (pdf) of $\boldsymbol{\vartheta}=[\mathbf{Y}_{\rm dB}+\varepsilon;\mathrm{SNR}_{\rm dB}(\mathbf{p},\mathbf{p}_{i,b})+\varepsilon_{i,b}]=[\boldsymbol{\vartheta}_{\rm blk1};\vartheta_{\rm blk2}]$ is reorganized as $f_0(\boldsymbol{\vartheta})=\frac{1}{\sqrt{(2\pi)^{|\mathcal{N}_k|+1}|\bar{\boldsymbol{\Phi}}|}}\exp\{-\frac{\boldsymbol{\Lambda}(\boldsymbol{\vartheta})}{2}\}$, where $\boldsymbol{\Lambda}(\boldsymbol{\vartheta})=(\boldsymbol{\vartheta}-\bar{\mathbf{H}}\theta\odot(\mathbf{1}_{|\mathcal{N}_k|+1}+\bar{\mathfrak{R}}^T))^T\times\bar{\boldsymbol{\Phi}}^{-1}(\boldsymbol{\vartheta}-\bar{\mathbf{H}}\boldsymbol{\theta}\odot(\mathbf{1}_{|\mathcal{N}_k|+1}+\bar{\mathfrak{R}}^T))$, $\bar{\mathbf{H}}=[\mathbf{H};\bar{\mathbf{h}}]$, $\bar{\mathfrak{R}}=[\mathfrak{R},\breve{\mathfrak{R}}]$, $\bar{\boldsymbol{\Phi}}=[\boldsymbol{\Phi},\boldsymbol{\Xi};\boldsymbol{\Xi}^T,\alpha]$ and $\alpha=(\xi^2+\rho^2)(1+\breve{\mathfrak{R}})$. Accordingly, $\boldsymbol{\Lambda}(\boldsymbol{\vartheta})$ is rearranged as

$$\boldsymbol{\Lambda}(\boldsymbol{\vartheta})=\boldsymbol{\Lambda}_1(\boldsymbol{\vartheta}_{\rm blk1})+\boldsymbol{\Lambda}_2(\boldsymbol{\vartheta}_{\rm blk2})$$

where $\boldsymbol{\Lambda}_1(\boldsymbol{\vartheta}_{\rm blk1})=(\boldsymbol{\vartheta}_{\rm blk1}-\mathbf{H}\theta\odot(\mathbf{1}_{|\mathcal{N}_k|}+\mathfrak{R}^T))^T\boldsymbol{\Phi}^{-1}(\boldsymbol{\vartheta}_{\rm blk1}-\mathbf{H}\theta\odot(\mathbf{1}_{|\mathcal{N}_k|}+\mathfrak{R}^T))$, $\bar{\boldsymbol{\Lambda}}_2(\boldsymbol{\vartheta}_{\rm blk1},\vartheta_{\rm blk2})=(\vartheta_{\rm blk2}-\mathrm{Avg})^T\mathrm{Var}^{-1}(\vartheta_{\rm blk2}-\mathrm{Avg})$, $\mathrm{Avg}=\bar{\mathbf{h}}\boldsymbol{\theta}(1+\breve{\mathfrak{R}})+\boldsymbol{\Xi}^T\boldsymbol{\Phi}^{-1}(\bar{\mathbf{w}}_{\rm blk1}-\mathbf{H}\theta\odot(\mathbf{1}_{|\mathcal{N}_k|}+\mathfrak{R}^T))$, and $\mathrm{Var}=(\xi^2+\rho^2)(1+\breve{\mathfrak{R}})-\boldsymbol{\Xi}^T\boldsymbol{\Phi}^{-1}\boldsymbol{\Xi}$.

Further, the pdf of the vector $\boldsymbol{\vartheta}$ can be changed as $f_0(\boldsymbol{\vartheta})=\frac{\exp\Big\{-\frac{\Lambda_1(\boldsymbol{\vartheta}_{\rm blk1})}{2}\Big\}\exp\Big\{-\frac{\Lambda_2(\boldsymbol{\vartheta}_{\rm blk1},\vartheta_{\rm blk2})}{2}\Big\}}{\sqrt{(2\pi)^{|\mathcal{N}_k|}|\boldsymbol{\Phi}|}\sqrt{2\pi\mathrm{Var}}}$. Accordingly, the marginal pdf of $\boldsymbol{\vartheta}_{\rm blk1}$

can be expressed as $f_{\boldsymbol{\vartheta}_{\text{blk1}}}(\boldsymbol{\vartheta}_{\text{blk1}}) = \frac{\exp\left\{-\frac{\Lambda_1(\boldsymbol{\vartheta}_{\text{blk1}})}{2}\right\}}{\sqrt{(2\pi)^{|N_k|}|\boldsymbol{\Phi}|}}$.

Finally, the conditional pdf of $\vartheta_{\text{blk2}}$ is $f_{2|1}(\vartheta_{\text{blk2}}|\boldsymbol{\vartheta}_{\text{blk1}}) = \frac{f(\boldsymbol{\vartheta})}{f_{\boldsymbol{\vartheta}_{\text{blk1}}}(\boldsymbol{\vartheta}_{\text{blk1}})} = \frac{\exp\left\{-\frac{\bar{\Lambda}_2(\boldsymbol{\vartheta}_{\text{blk1}},\vartheta_{\text{blk2}})}{2}\right\}}{\sqrt{2\pi|\text{Var}|}}$. ∎

**Simulation and experiment results:** In this section, we conduct simulation studies to verify the proposed relaying solution. It is assumed that three queues are deployed in the relaying area. The actual channel parameters is set as $\kappa = 1$. Besides that, some other channel parameters can be shown as $K_{\text{dB}} = -52$ dB, $n_{\text{PL}} = 0.5$, $\xi^2 = 4$, $\rho^2 = 2.25$ and $\beta = 1$. $\mathbf{Q}_1 = \mathbf{I}_{2500}$ and $R_1 = 1$. Accordingly, the iterative process for the unknown scattering parameter is shown in Fig. 2(a), where the estimated parameters converge to 1. Correspondingly, the measured SNR and predicted SNR are shown in Fig. 2(b). Clearly, the difference between the estimated SNR and the sampled SNR is very small. On the basis of this, the data transmission success rate (DTSR) for the queues is shown in Fig. 3(a). Compared to the analytical solution proposed in [7], our channel estimator has a higher estimation accuracy. Specifically, the analytical solution combines with the first and second kind spherical Bessel functions, i.e., $j_1(\cdot)$ and $j_2(\cdot)$, and Legendre order polynomial of order 2, i.e., $P_2(\cdot)$, to calculate the amplitude of scattering field. Accordingly, we calculate it at a distance of $r_0 = 5$ m from the center of the circle is expressed as $P_0 + j_1(\kappa r_0) + j_2(\kappa r_0) \times P_2(\cos(\theta))$, where $\theta \in (0, 2\pi]$ and $P_0 = -1$ dB. In addition, $j_1(\kappa r_0)$, $j_2(\kappa r_0)$ and $P_2(\cos(\theta))$ can be calculated by the besselj and legendre functions. Finally, the comparison result is shown in Fig. 2(c). The above results verify the effectiveness of the proposed channel estimation method.

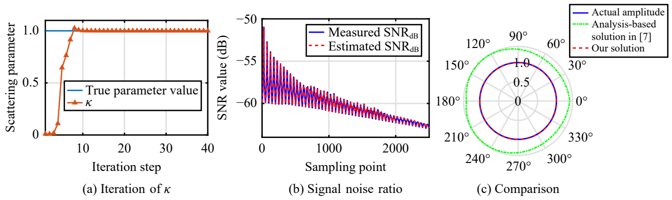


Fig. 2. Simulation results for the BLS-based online channel prediction.

With the predicted channel information, we verify the effectiveness of the semi-supervised relay controller. The initial state of AUV is $\boldsymbol{\zeta}(0) = [45, 20, -33, 5, 0, 0, 0, 0]^T$. Accordingly, the trajectory of AUV is presented in Fig. 3(a), whose position and orientation are shown in Fig. 3(b). Correspondingly, the optimal policy of AUV is shown in Fig. 3(c). Based on this, the DTSR of AUV is shown in Fig. 3(d). Clearly, $F_{1,\text{sd}}(\mathbf{p}(k))$, $F_{2,\text{sd}}(\mathbf{p}(k))$ and $F_{3,\text{sd}}(\mathbf{p}(k))$ converge to 0.25, which means AUV has been derived to the optimal relay positions for the three queues. The collision avoidance distances are shown in Fig. 3(e), where the collision avoidances are both greater than 6 m, 5 m and 3 m. Finally, the time taken for this control process is shown in Fig. 3(f), which approaches $\Gamma_{2,\min}$. These results verify the effectiveness of the proposed controller.

The field experiment is conducted in the pool of our lab. Due to condition limitation, we only verify the channel estimator. As shown by Fig. 4(a), the obstacle is captured by an cylinder according to Assumption 1. As mentioned above, we employ a digital hydrophone to obtain channel measurement data. Correspondingly, the iterative process of scattering parameter is shown in Fig. 4(c). In addition, the measured and estimated channel distribution is shown in Fig. 4(b). Clearly, the scattering parameter can be converged to 2.68. The estimated results are very close to the measured results, which proves the effectiveness of our proposed channel estimator.

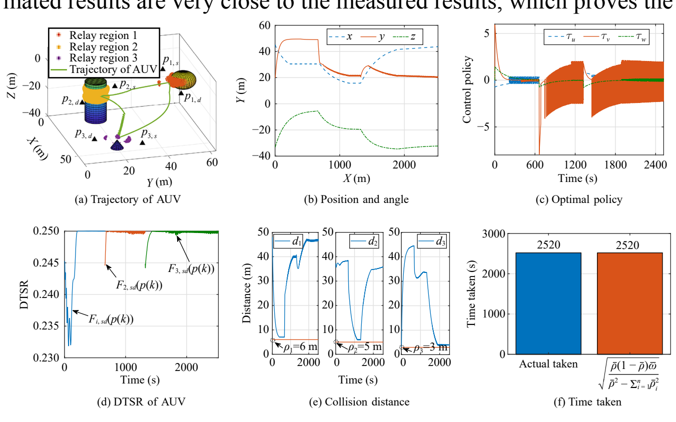


Fig. 3. Simulation results for the BLS-based online channel prediction.

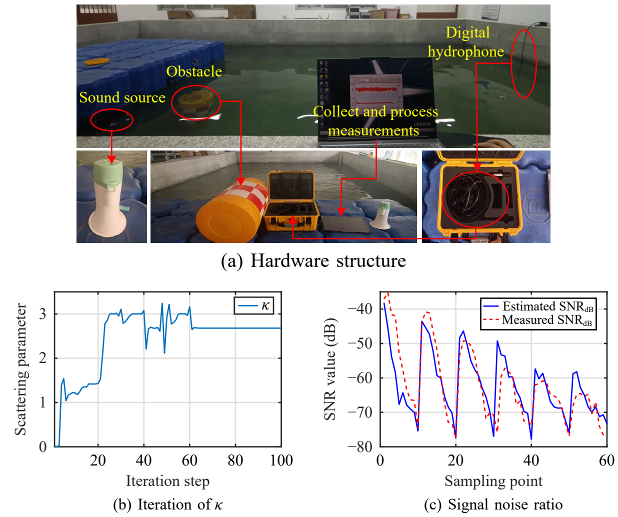


Fig. 4. Experimental results for the channel prediction.

**Conclusion and future works:** This letter studies the broad learning based mobile relaying solution of AUV for minimal wait time. It should be emphasized that our solution assumes that the obstacles are regarded as rigid objects, which is a prerequisite for using the Kirchhoff approximation method. However, it is valid only in a few special cases, where a underwater scatterer is rigid and immovable. Therefore, this is a limitation of our method. To solve this limitation, we will combine the controller with multiple sensing techniques (e.g., the sonar and binocular vision), such that the sensing information can be effectively fed back to the control systems. In future, we will implement the solution to the practical marine environment.

**Acknowledgments:** This work was supported in part by the Natural Science Foundation of China (62222314, 61973263, 62033011), the Youth Talent Program of Hebei (BJ2020031), the Distinguished Young Foundation of Hebei Province (F2022203001), and the Central Guidance Local Foundation of Hebei Province (226Z3201G).

## References


[1] W. Cao, J. Yan, X. Yang, X. Luo, and X. Guan, "Communication-aware formation control of AUVs with model uncertainty and fading channel via integral reinforcement learning," *IEEE/CAA J. Autom. Sinica*, vol. 10, no. 1, pp. 159–176, Jan. 2023.

[2] Z. Zhou, J. Liu, and J. Yu, "A survey of underwater multi-robot systems," *IEEE/CAA J. Autom. Sinica*, vol. 9, no. 1, pp. 1–18, Jan. 2022.

[3] M. Cheng, Q. Guan, F. Ji, J. Cheng, and W. Chen, "Mobile relaying-based reliable data collection in underwater acoustic sensor networks," *IEEE Wirel. Commun. Lett.*, vol. 11, no. 9, pp. 1795–1799, Sept. 2022.

[4] J. Yan, W. Cao, X. Yang, C. Chen, and X. Guan, "Communication-efficient and collision-free motion planning of underwater vehicles via integral reinforcement learning," *IEEE Trans. Neural Netw. Learn. Syst.*, 2022. DOI: 10.1109/TNNLS.2022.3226776.

[5] A. Bonomo, M. Isakson, and N. Chotiros, "A comparison of finite element and analytic models of acoustic scattering from rough poroelastic interfaces," *J. Acoust. Soc. Am.*, vol. 33, pp. 235–240, Mar. 2015.

[6] P. Nicolas, O. Ciaran, and B. Susann, "A Kirchhoff approximation-based numerical method to compute multiple acoustic scattering of a moving source," *Appl. Acoust.*, vol. 96, no. 1, pp. 108–117, Apr. 2015.

[7] N. Nadimi, J. Reza, and K. Layeghi., "An efficient acoustic scattering model based on target surface statistical descriptors for synthetic aperture sonar systems," *J. Mar. Sci. Appl.*, vol. 19, pp. 494–507, 2020.

[8] X. Gao, W. Bai, T. Li, L. Yuan, and Y. Long, "Broad learning systembased adaptive optimal control design for dynamic positioning of marine vessels," *Nonlinear Dyn.*, vol. 105, no. 1, pp. 1593–1609, Jul. 2021.

[9] C. Chen, D. Yu, and L. Liu, "Automatic leader-follower persistent formation control for autonomous surface vehicles," *IEEE Access*, vol. 7, no. 1, pp. 12146–12155, Dec. 2019.

[10] Y. Zhao, X. Qi, Y. Ma, Z. Li, R. Malekian, and M. Sotelo, "Path following optimization for an underactuated USV using smoothly convergent deep reinforcement learning," *IEEE Trans. Intel. Transp. Syst.*, vol. 22, no. 10, pp. 6208–6220, Oct. 2021.

[11] Y. Wang, K. Li, and Z. Chen, "Battery full life cycle management and health prognosis based on cloud service and broad learning," *IEEE/CAA J. Autom. Sinica*, vol. 9, no. 8, pp. 1540–1542, Aug. 2022.