Dynamic modeling and Learning based Path Tracking Control for ROV-based Deep-sea Mining Vehicle

Yuheng Chen[1], Haicheng Zhang[1*], Weisheng Zou[1], Haihua Zhang[2] Bin Zhou[2] and Daolin Xu[1]

(1. College of Mechanical and Vehicle Engineering, Hunan University, Changsha, Hunan 410082 China; 2 Taihu Laboratory of Deepsea Technological Science Lian Yun Gang Center, Jiangsu 222005 China

[*]Corresponding author e-mail address: chenyuheng@hnu.edu.cn, zhanghc@hnu.edu.cn, zouweisheng1@hnu.edu.cn, zhanghaihua@702sh.com, zhoubin@702cmes.com, dlxu@hnu.edu.cn.

Abstract: Track slippage and body sinking of the tracked mining vehicle in the traditional deep-sea mining system are the critical issues for operating stability. To solve this bottleneck problem, a novel ROV-based deep-sea mining system is proposed in this study, in which a remotely-operated vehicle (ROV) towering a sledge-shaped mining robot (MRT) named ROV-based Deep-sea Mining Vehicle (ROVDMV) is instead of the traditional tracked Deep-sea mining vehicle. The design of the ROVDMV can fundamentally overcome the bottleneck problem. However, the complex marine environment and multi-rigid-body design of the ROVDMV pose new challenges for its path-tracking control. Firstly, the dynamic model of the ROVDMV considering the ROV at a fixed depth is established based on the bicycle model, which is mainly used as the control object in the numerical simulation. Secondly, a learning-based path-tracking control strategy is proposed for the path-tracking control of the ROVDMV. In the control strategy, a novel nonparametric learning (NPL) method is introduced to learn the uncertain nonlinear dynamics considering the external disturbances and parametric uncertainty. The NPL method is proven to provide bounded estimated error. Besides, the enhanced NPL method can save approximately 33% of the computation time, and the average computation time for its optimization control problem is only 12.47ms. Finally, the numerical results show that the NPL method can learn nonlinear dynamics accurately, and the proposed strategy has proven to be effective.

Keywords: Deep-sea mining, dynamic modelling, learning-based model predictive control, ROV, path tracking

---

Nomenclature

b l, L the MRT's width and length, the horizontal projection length of the steel frame

u_{MRT} u_{ROV} v_{MRT} v_{ROV}, surge speed and sway speed

ψ ω_{MRT} ω_{ROV}, MRT's yaw angle and yaw rate

δ, the horizontal rotation angle of the ROV around the steel frame

L_g q K_{MRT} K_{ROV} E, Lagrangian function and its state kinetic energy, potential energy

---

F_{ROV}, F_{MRT} hydrodynamic force of the ROV, interaction force of the MRT

M_A, the add mass matrix of the ROV

$C\ D$, Centripetal and coriolis force matrix hydrodynamic damping matrix

$m_{MRT}\ m_{ROV}\ I_{MRT}\ I_{ROV}$, mass and rotational inertia

$X_{\dot{u}}\ Y_{\dot{v}}\ N_{\dot{r}}\ X_u\ X_{uu}\ Y_v\ Y_{vv}\ N_r\ N_{rr}$, hydrodynamic coefficient

$X_u\ X_{uu}\ Y_v\ Y_{vv}\ N_r\ N_{rr}$, hydrodynamic damping coefficient

$F_X\ F_Y\ T_N$, control input

$F_R\ F_{Rc}\ F_{Rb}$, longitudinal resistance, compaction resistance bulldozing resistance

M_{o_2}, the ROV's turning resistance moment

$f\ \mu_y$, lateral friction force and lateral friction coefficient

$W\ C_I$, underwater weight of the MRT terrain cone index

$E_1\ E_2\ E_3$, empirical coefficients

$F_R\ F_{Rc}\ F_{Rb}$, longitudinal resistance, compaction resistance and bulldozing resistance

$\Delta z\ c$, sinkage and apparent cohesion

r_s, ϕ , the density of the sediment and angle of internal shearing resistance

$k_{pr}\ k_{pc}$, coefficients of passive earth pressure

$N_r\ N_c$, Terzaghi bearing capacity coefficient

1. Introduction

The deep sea is rich in mineral resources, including polymetallic nodules, cobalt-rich crusts, and polymetallic sulfides. The efficient exploitation of these mineral resources can help alleviate the shortage of terrestrial mineral (Sha et al., 2023). To extract these resources, deep-sea mining systems serve as the essential equipment for harvesting minerals from the ocean floor, which lies thousands of meters beneath the surface. Currently, deep-sea mining remains in the research and exploration phase, primarily because deep-sea mining systems have not yet achieved the level of commercial viability. Since the concept of deep-sea mining was introduced, several collection methods have been proposed, ranging from ship-towed mining vehicles to Archimedes spiral self-propelled mining vehicles. At present, the hydraulic lift deep-sea mining system is regarded as a commercially viable option, as illustrated in Fig. 1(a). In this system, a deep-sea tracked mining vehicle must follow a designated mining path to collect mineral resources. The mineral ore, mixed with water, is then transported to the mining vessel through a buffer and lifting pipe by a lifting pump (H. Wu et al., 2023). The deep-sea tracked mining vehicle is one of the most central subsystems of the deep-sea mining system and must adhere to a predetermined mining path to gather the mineral resources. However, the traditional

heavy tracked mining vehicles face the challenge of significant grip and substantial subsidence when traversing on the seafloor sediments (Wang, Chen, Wang, Li, & Yang, 2023), and track slippage poses a key issue for accurate path tracking (Qin et al., 2021). Therefore, it is essential to develop an efficient and robust path tracking controller.

Path tracking controllers based on geometric methods, such as Fuzzy PID (Dai, Su, & Zhang, 2020), line of sight (Yeu et al., 2012) and pure pursuit 2013), are efficient and straightforward to design, as they directly calculate the path tracking control law. However, since no anti-slip strategy is considered, the performance of path tracking control may degrade due to track slippage. To address this issue, model predictive control (MPC) (W. Li et al., 2023; Li, He, Ma, Liu, & Liu, 2023; P. C. Wu, Wen, Chen, & Jin, 2017) and learning-based methods (An, Zhou, & Wang, 2024; Q. Chen et al., 2023) have been employed to account for track slippage in the design of path tracking controllers. By incorporating anti-slip constraints, an MPC-based optimal control problem is formulated to derive the path tracking control law (Y. Li et al., 2023). Chen et al. propose a learning-based path tracking controller based on deep reinforcement learning (Chen et al., 2023), where an improved deep deterministic policy gradient is explored. By fully considering the slip rate of the tracks in the reward function, this path tracking controller effectively mitigates track slip and ensures robust path tracking performance. It is evident that the anti-slip control strategy can enhance the path tracking performance to some extent. However, the track slippage, which is a critical issue for operational stability, cannot be completely eliminated. This limitation represents an inherent disadvantage of traditional deep-sea mining systems.

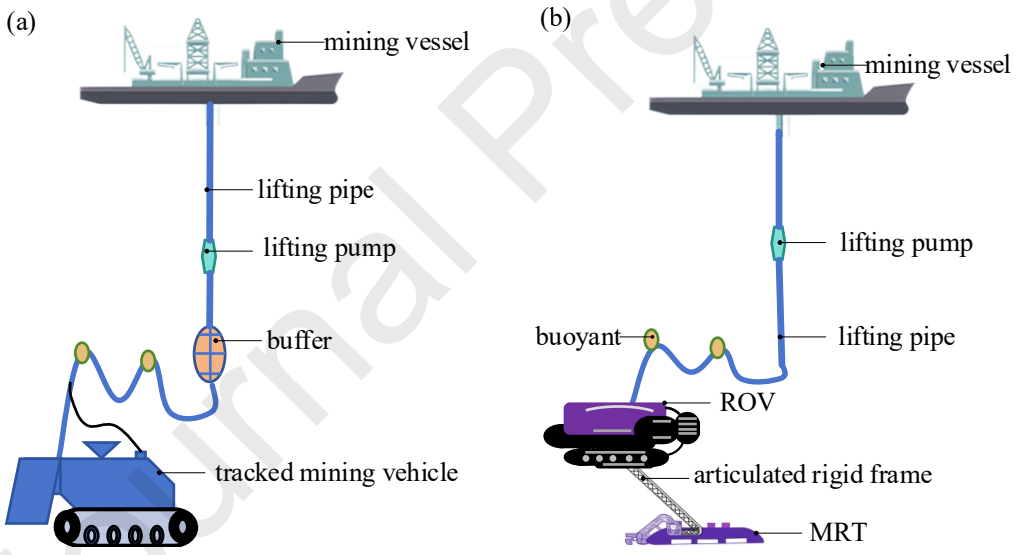

Fig.1 Comparison of two deep-sea mining systems (a, traditional deep-sea mining system; b, ROV-based deep-sea mining system)

To address the inherent disadvantages of traditional deep-sea mining systems, a remotely operated vehicle (ROV)-based deep-sea mining system is proposed, as illustrated in Fig. 1(b). In this collection method, the ROV-based deep-sea mining vehicle (ROVDMV), which consists of an ROV and a mining robot (MRT), replaces the tracked mining vehicle. The ROV tows the MRT, allowing it to glide smoothly along the seabed via an articulated steel frame, while the ROV maintains a specific height above the seabed. Compared to the tracked mining vehicle, this approach increases the contact area with the substrate and reduces ground pressure. In conclusion, the novel collection method effectively mitigates issues related to slippage and body subsidence. The ROVDMV must track a predetermined mining path to

collect the mineral resources in the ROV-based deep-sea mining system, and the path tracking strategy for the ROVDMV requires thorough investigation. Firstly, the dynamic model of the ROVDMV needs to be established as the controlled object in the numerical simulation of path tracking, which is more complex compared to traditional mining vehicles. In the force analysis, the hydrodynamic forces acting on the ROV and the interaction forces between the MRT and the seabed sediment must be taken into account. Secondly, the uncertain nonlinear dynamics present a challenge for the path tracking strategy, primarily due to model mismatch between the established dynamic model and the actual system. Accurately obtaining the hydrodynamic parameters of the ROV and the soil mechanical parameters of the seabed is difficult. Additionally, external current disturbances and process noise resulting from irregular seafloor topography can also contribute to model mismatch.

An efficient method for addressing model mismatch is based on a nonparametric learning (NPL) approach, wherein a mapping function is developed to estimate uncertain nonlinear dynamics (Calliess, Roberts, Rasmussen, & Maciejowski, 2020). The Kinky Inference (KI) prediction function, which utilizes the Lipschitz constant, is employed to accurately estimate the nonlinear unmodeled dynamics under random process noise. The Lazily Adapted Constant Kinky Inference (LACKI) rule is used to estimate the Lipschitz constant. However, the accuracy of this estimation is contingent upon a dense set of samples, resulting in high computational complexity. To address this issue, Kaikai Zheng proposes an event-triggering mechanism to enhance the method (Zheng, Shi, Shi, & Wang, 2023). For a known training set that encompasses all possible sampling points, the size of the training set can be reduced to ensure good real-time performance, with estimation accuracy remaining nearly uncompromised. It is important to note that the dynamics of the ROVDMV are coupled with its time-varying state, making it challenging to obtain a known training set in advance. Consequently, this method becomes less applicable. To tackle these challenges, the research focuses on studying the dynamics model of the ROVDMV and its learning-based path tracking control strategy. The main contributions of this paper are as follows.

1. A new method and equipment for deep-sea mining have been proposed, and the path tracking control of the ROVDMV is examined in the horizontal plane. Drawing inspiration from the bicycle model (Sun, Li, Li, & Li, 2022), a dynamic model of the ROVDMV has been developed as the controlled object in the numerical simulation.

2. A novel nonparametric learning method has been developed to address uncertain nonlinear dynamics, which encompass unknown external current disturbances and certain aspects of nonlinear dynamics affected by parametric uncertainty in the presence of random process noise. Based on the LACKI rule (Calliess et al., 2020), a new learning rule is introduced to enhance the learning performance. The training set, which consists of a limited number of sampling points, is dynamically updated by this learning rule to reduce online computational complexity. This approach effectively resolves the issue that the nonlinear programming (NLP) requires a previously known training set (Zheng et al., 2023). Furthermore, under the proposed learning rule, the estimated error of the nonlinear unmodeled dynamics is proven to be bounded.

3. A learning-based path tracking control strategy is proposed to achieve optimal path tracking performance. A virtual speed model predictive control (VSMPC) system is designed to compute the virtual speed control law, which aims to minimize the real-time path tracking deviation. Additionally, a learning-based model predictive control (LBMPC) framework is developed to determine the control input necessary for tracking the virtual speed control. In this framework, the learned uncertain nonlinear dynamics serve as the predictive model for the optimal control problem. With the estimated error proven to be bounded, the closed-loop stability can also be ensured (Zheng et al., 2023).

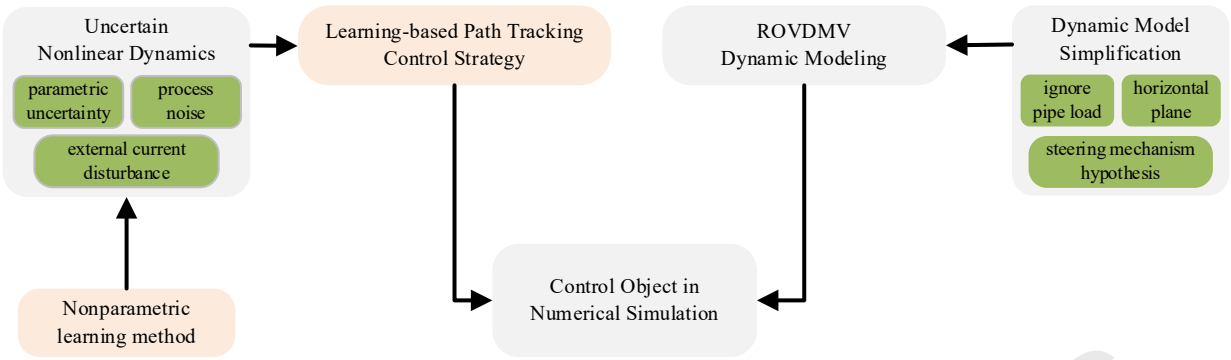

Fig.2 Research framework

The remainder of this paper is organized as follows: Section 2 introduces the dynamic model of the ROVDMV. Section 3 introduces a novel nonparametric learning method. Section 4 proposes a learning-based path tracking control strategy. Section 5 provides the results and discussion. The research framework of this paper is illustrated in Fig. 2.

Notations: denote $m \times n$ real matrices by $\mathcal{R}^{m\times n}$. Denote n-dimensional unit matrix and $m \times n$ zero matrix by I_n and $0_{m\times n}$. Given a positive definite matrix $P \in \mathcal{R}^{m\times m}$, define matrix norm $\|x\|_P^2 = x^{\mathrm{T}}Px, x \in \mathcal{R}^{m\times 1}$. Define set $\mathbb{K}_{N_1:N_2} = \{N_1, N_1+1, \cdots, N_2-1, N_2\}$. For a column vector $b \in \mathcal{R}^{m\times 1}$, b_j represents the element in j-th row. Denote the probability of an event A by $\mathrm{P}(A)$.

2. Dynamic Model of the ROVDMV

In this section, we derive the dynamic model of the ROVDMV for simulating its path tracking control. Compared to the towing force, the connecting force exerted by the flexible pipe on the ROV is relatively weak due to the saddle shape of the compensation soft pipe. Additionally, the mass of the steel frame is small in comparison to the mass of the ROV and MRT.

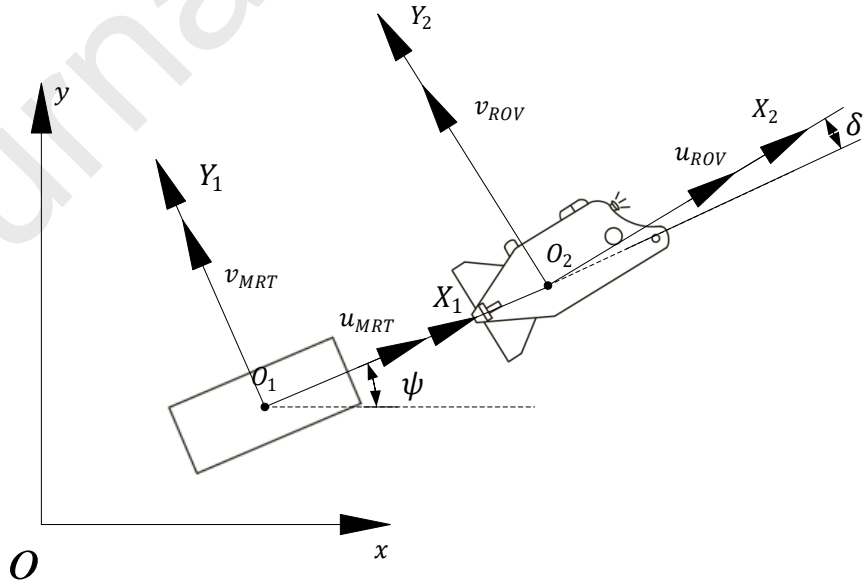

Fig.3 Coordinate frame of the ROVDMV

To simplify the model, the forces exerted by the flexible pipe and the mass of the steel frame are disregarded during the modeling process. Additional simplifications are also made in the preliminary research of the ROVDMV mining vehicle. It is assumed that the MRT and the steel frame are rigid, while the ROV and the steel frame are articulated. Drawing inspiration from the bicycle model in an Ackermann steering vehicle (Sun et al., 2022), the ROV and MRT are analogized to the front and rear wheels of a bicycle, respectively, with the ROV functioning as the steering mechanism. To ensure stable steering, both the steering angle of the ROV and its angular velocity around the steel frame are designed to be minimal. Consequently, the small angle assumption for the steering angle is introduced in the subsequent dynamic modeling.

Moreover, suppose the remotely operated vehicle (ROV) tows the mobile robotic transporter (MRT) at a constant depth, with the MRT sliding along the seafloor surface. The dynamic model can be considered in the horizontal plane; however, both the ROV and MRT possess three degrees of freedom (DOFs) each. To simplify the dynamic modeling, the kinematic relationship between the DOFs of the ROV and MRT is analyzed first. Consequently, it is only necessary to examine the dynamics of three DOFs of the ROV.

2.1 Kinematic Model

Firstly, the global coordinate and the local coordinate frame of the ROV and MRT are defined, as shown in Fig.3. $O-xy$ denotes the global coordinate system. $O_1-X_1Y_1$ denotes MRT's local coordinate system, and $O_2-X_2Y_2$ denotes ROV's local coordinate system. Then, surge speed u and sway speed v of the ROV and the MRT can be expressed as:

$$u_{MRT}=u_{ROV}\cos\delta-v_{ROV}\sin\delta\,,v_{MRT}=v_{ROV}\cos\delta+u_{ROV}\sin\delta \quad (1.)$$

where δ is the horizontal rotation angle of the ROV around the steel frame.

According to the bicycle model (Sun, Li, Li, & Li, 2022), MRT's yaw rate ω_{MRT} is related to the surge speed of the MRT u_{MRT} and the horizontal rotation angle of the ROV around the steel frame δ:

$$\dot{\psi}=\omega_{MRT}=\frac{u_{MRT}\delta}{L} \quad (2.)$$

where ψ is MRT's yaw angle. L is the horizontal projection length of the steel frame.

It can be seen both the ROV and the MRT have 3 degrees of freedom (DOFs), and the kinematic model of ROVDMV is given as follows:

$$\begin{cases}\dot{x}=u_{MRT}\cos\psi-v_{MRT}\sin\psi\\ \dot{y}=u_{MRT}\sin\psi+v_{MRT}\cos\psi\\ \dot{\psi}=\omega_{MRT}=\dfrac{u_{MRT}\delta}{L}\\ u_{MRT}=u_{ROV}\cos\delta-v_{ROV}\sin\delta\\ v_{MRT}=v_{ROV}\cos\delta+u_{ROV}\sin\delta\\ \dot{\delta}=\omega_{ROV}\end{cases} \quad (3.)$$

where x and y are MRT's global coordinates. ω_{ROV} is ROV's yaw rate.

2.2 Kinetic Model

The kinetic model of the ROVDMV consists of a kinetic equation of the ROV with 3 DOFs, which will be deduced using the Lagrangian method. Firstly, the kinetic energy of the ROV $K_{MRT}$and that of the MRT K_{ROV} are given as:

$$\begin{cases} K_{MRT} = \frac{1}{2}m_{MRT}\left(u_{MRT}^2 + v_{MRT}^2\right) + \frac{1}{2}I_{MRT}\omega_{ROV}^2 = \frac{1}{2}m_{MRT}\left(u_{ROV}^2 + v_{ROV}^2\right) + \frac{1}{2}I_{MRT}\left(\frac{u_{MRT}\delta}{L}\right)^2 \\ K_{ROV} = \frac{1}{2}m_{ROV}\left(u_{ROV}^2 + v_{ROV}^2\right) + \frac{1}{2}I_{ROV}\omega_{ROV}^2 \end{cases} \quad (4.)$$

where I_{ROV} and I_{MRT} denote the rotational inertia of the ROV and the MRT. m_{MRT} and m_{ROV} are the mass of the ROV and the MRT, respectively.

Note that the dynamics of the ROVDMV is considered in the horizontal plane, and the potential energy of the ROVDMV E can be treated as a constant. Then, the Lagrange function of the ROVDMV $L(q,\dot{q})$ is given as:

$$L_g(q,\dot{q}) = K_{MRT} + K_{ROV} - E = \frac{1}{2}(m_{MRT} + m_{ROV})\left(u_{ROV}^2 + v_{ROV}^2\right) + \frac{1}{2}I_{MRT}\left(\frac{u_{ROV}\delta}{L}\right)^2 + \frac{1}{2}I_{ROV}\omega_{ROV}^2 - E \quad (5.)$$

where the state q and its first order derivative $\dot{q}$ are denoted by $q = (x_{ROV}, y_{ROV}, \delta)^{\mathrm{T}}$ and $\dot{q} = (u_{ROV}, v_{ROV}, \omega_{ROV})^{\mathrm{T}}$.

Finally, the Lagrangian equation is given by:

$$\frac{d}{dt}\frac{L_g}{\partial\dot{q}} - \frac{L_g}{\partial q} = F_{ROV} + F_{MRT} + \tau + \tau_e + \tau_w \quad (6.)$$

where F_{ROV} is the generalized hydrodynamic force of the ROV in the local coordinate system $O_2 - X_2Y_2$. Denote the generalized interaction force between the MRT and seafloor sediments in the local coordinate system $O_2 - X_2Y_2$ by F_{MRT}. The control input is denoted by $\tau = (F_X, F_Y, T_N)^{\mathrm{T}}$, which represents the forces and torques corresponding to the 3 DOFs in the ROV's local coordinate system. $\tau_e \in \mathcal{R}^{3\times1}$ represents the unmodeled bounded external ocean current disturbance and model mismatch from parametric uncertainty. Denote the random bounded process noise force of seafloor sediments and the MRT from the rugged seafloor by $\tau_w \in \mathcal{R}^{3\times1}$ (Dai, Xue, Su, & Huang, 2021).

Substitute Eq. 5 into Eq. 6, and the left side of Eq. 6 becomes:

$$\frac{d}{dt}\frac{L_g}{\partial u_{ROV}} - \frac{L_g}{\partial x_{ROV}} = (m_{MRT} + m_{ROV})\dot{u}_{ROV} + \frac{I_{MRT}\delta^2}{L^2}\dot{u}_{ROV} = (m_{MRT} + m_{ROV} + m_\delta)\dot{u}_{ROV} \quad (7.)$$

$$\frac{d}{dt}\frac{L_g}{\partial v_{ROV}} - \frac{L_g}{\partial y_{ROV}} = (m_{MRT} + m_{ROV})\dot{v}_{ROV} \quad (8.)$$

$$\frac{d}{dt}\frac{L_g}{\partial \omega_{ROV}} - \frac{L_g}{\partial \delta_{ROV}} = I_{ROV}\dot{\omega}_{ROV} - \frac{I_{MRT}u_{ROV}^2}{L^2}\delta \quad (9.)$$

where $m_\delta = I_{MRT}\delta^2/L^2$ denotes the added mass of the MRT in the horizontal rotation angle DOF.

Based on previous research (Long, Hu, Qin, & Bian, 2022), an empirical formula can be utilized to express the generalized hydrodynamic force of the ROV F_{ROV}, which consists of the added mass force, centripetal and Coriolis force and hydrodynamic damping force:

$$F_{ROV} = M_A\dot{q} + (C + D)q(10.)$$

where the add mass matrix is denoted by $M_A = \mathrm{diag}\left(X_{\dot{u}}, Y_{\dot{v}}, N_{\dot{r}}\right)$, where $X_{\dot{u}}$, $Y_{\dot{v}}$ and $N_{\dot{r}}$ represent the hydrodynamic coefficients. Denote the hydrodynamic damping matrix by $D = \mathrm{diag}(X_u + X_{uu}|u_{ROV}|, Y_v + Y_{vv}|v_{ROV}|, N_r + N_{rr}|\omega_{ROV}|)$ where the symbols X_u, X_{uu}, Y_v, Y_{vv}, N_r and N_{rr} are hydrodynamic damping coefficients. Centripetal and Coriolis force matrix C can be expressed as:

$$C = \begin{bmatrix} 0 & 0 & (X_{vr} + m_{ROV})v_{ROV} + X_{rr}\omega_{ROV} \\ 0 & Y_{uv}u_{ROV} & (Y_{ur} - m_{ROV})u_{ROV} \\ 0 & N_{uv}u_{ROV} & N_{ur}u_{ROV} \end{bmatrix}(11.)$$

where X_{vr}, X_{rr}, Y_{uv}, Y_{ur}, N_{uv} and N_{ur} are hydrodynamic coefficients.

To further establish the Lagrangian equation (6), the interaction force F_{MRT} needs to be deduced. Since the soft seafloor sediment can be considered plastic soil (Chen et al., 2023), the MRT is subjected to longitudinal resistance F_R and lateral friction force f from a triangular load. Then, the mechanical sketch of the ROVDMV is established, as shown in Fig.4. Considering plastic soil, a lateral friction coefficient concerning the turning radius is used to express the lateral friction force (Al-Milli, Seneviratne, & Althoefer, 2010):

$$f = \mu_y \frac{W}{2}(12.)$$

where W represents the underwater weight of the MRT, and μ_y is the lateral friction coefficient:

$$\mu_y = E_1\left(1 - e^{\frac{C_I blE_2}{m_{MRT}}}\right)\left(1 - e^{\frac{C_I blE_3 r}{m_{MRT}}}\right)(13.)$$

in which l and b are the length and width of the MRT. C_I represents the terrain cone index. Denote the turning radius of the MRT by r. E_1, E_2 and E_3 are empirical coefficients.

As shown in Fig.4, the lateral friction force f can be equivalent to the ROV's turning resistance moment M_{o_2} with respect to O_2:

$$M_{o_2} = \left(\frac{3}{4}l + L\right)f - \left(\frac{1}{4}l + L\right)f = \frac{1}{2}lf = \frac{\mu_y W l}{4}(14.)$$

Then, the generalized interaction force F_{MRT} can be obtained:

$$F_{MRT} = \left(-F_R\cos\delta_{ROV}, F_R\sin\delta_{ROV}, -\mathrm{sg}(\omega_{\mathrm{MRT}})M_{o_2}\right)(15.)$$

where $\mathrm{sg}(\cdot)$ is a function denoted as $\mathrm{sg}(x) = 1, x > 0; \mathrm{sg}(x) = -1, x < 0; \mathrm{sg}(x) = 0, x = 0$.

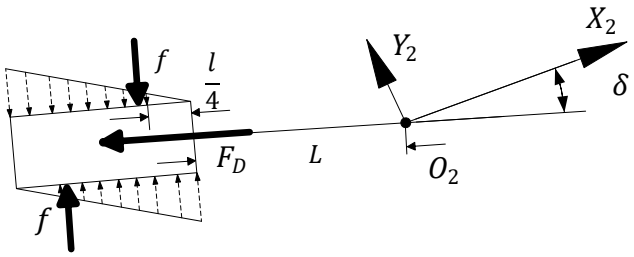

Fig.4 Mechanical sketch of the ROVDMV

Due to the sinkage on the soft seafloor sediment (Ruslan et al., 2023; Xu et al., 2022), the longitudinal resistance F_R consists of a compaction resistance F_{Rc} and a bulldozing resistance F_{Rb}:

$$F_R = F_{Rc} + F_{Rb} (16.)$$

where the underwater weight of the MRT and the soil properties will influence the sinkage characteristics (Yamada, Yamauchi, & Hashimoto, 2021).

Based on Bekker's pressure–sinkage relationship(Kar, 1987), the compaction resistance F_{Rc} can be expressed as follows:

$$F_{Rc} = \frac{b}{2f}\Delta z^2 - \frac{be}{\tilde{f}}\Delta z (17.)$$

where $\Delta z = e + \tilde{f}p$ is the sinkage. $p = W/bl$ is the normal pressure from the MRT acting on the soil. e and $\tilde{f}$ satisfy the empirical formula(Liu & Wong, 1996):

$$\tilde{f} = 1.99 - (18.)$$

$$\begin{matrix} e = 6.725 - 2.568\tau + 0.245\tau^2, \tau \geq 5\text{kPa} \\ e = 0, \tau < 5\text{kPa} \end{matrix} (19.)$$

in which τ is the shear strength of soft seafloor sediment:

$$\tau = c + p\tan\phi (20.)$$

where c is the apparent cohesion, and ϕ is the angle of internal shearing resistance.

Meanwhile, the bulldozing force can be expressed as a function between the mechanical properties of the soft seafloor sediment and the sinkage(Kar, 1987; Zeng et al., 2021):

$$F_{Rb} = \left(\frac{1}{2}r_s\Delta z^2 k_{pr} + c\Delta z k_{pc}\right)b (21.)$$

where r_s is the density of the sediment. Coefficients of passive earth pressure are denoted by k_{pr} and k_{pc}:

$$k_{pr} = \left(\frac{2N_r}{\tan\phi} + 1\right)\cos^2\phi, k_{pc} = (N_c - \tan\phi)\cos^2\phi (22.)$$

where N_r and N_c represent the Terzaghi bearing capacity coefficients.

3. Novel Nonparametric Learning Method

The kinetic model of the ROVDMV (6) employed in a path tracking strategy relies on prior knowledge, including model parameters and external ocean current disturbances. However, accurately obtaining model parameters, particularly the mechanical properties of the seabed soil, poses significant challenges. Additionally, the random process noise generated by seafloor sediments, along with unknown external ocean current disturbances, can further contribute to model mismatches.

To mitigate the effects of model mismatch, a novel nonparametric learning method has been developed based on the LACKI framework (Calliess et al., 2020; Zheng et al., 2023). This section introduces a learning rule designed to enhance the performance of the LACKI framework, ensuring both prediction accuracy and robust real-time performance. The innovative nonparametric learning (NPL) method effectively learns the values of uncertain nonlinear dynamics while accounting for unknown bounded external ocean current disturbances and uncharacterized process noise.

3.1 Methodology

The control model used in a path tracking strategy is considered as a discrete nonlinear model of the form:

$$\chi(k+1) = A\chi(k) + B\mathcal{U}(k) + f(\chi(k)) + \omega(k) \quad (23.)$$

where $\chi \in \mathbb{X} \subset \mathcal{R}^{n_x \times 1}$ is the state vector, and $\mathcal{U} \in \mathbb{U} \subset \mathcal{R}^{n_u \times 1}$ represents the control input vector. State matrix A and control matrix B are time-invariant. Nonlinear function $f(\chi(k)) \in \mathbb{Y} \subset \mathcal{R}^{n_x \times 1}$ represents the uncertain nonlinear dynamics with unknown bounded external disturbance, where $\mathbb{Y}$ is the output space. $\omega \in \mathbb{W} \subset \mathcal{R}^{n_x \times 1}$ is the bounded unknown process noise: $\|\omega\|_\infty \leq \bar{e}$, where the positive constant $\bar{e}$ is the upper bound, and $\mathbb{W}$ is the process noise space. Sets $\mathbb{X}$, $\mathbb{U}$ and $\mathbb{W}$ are assumed to be compact sets.

The estimated value of the uncertain nonlinear dynamic function f is denoted as $\hat{f}$. Define the sampled data set used for calculating the estimated value $\hat{f}$ by

$$D_n := \{(s(r), \tilde{f}(s(r))) | r \in \mathbb{K}_{1:N_n}\} \quad (24.)$$

where $s(k) := (\chi(k), \mathcal{U}(k)) \in \mathbb{S} \subset \mathcal{R}^{(n_x+n_u)\times 1}$ is the measured input value, and $\mathbb{S}$ is the input space. N_n is the number of recorded samples. For convenience, define the sampled input data set $\mathcal{G}_n := \{s(r) | r \in \mathbb{K}_{1:N_n}\}$. The measured output value of the uncertain nonlinear dynamic function f is denoted as $\tilde{f}(s(r)) := \chi(r) - A\chi(r-1) - B\mathcal{U}(r-1)$. The bounded unknown process noise holds that $\|\tilde{f} - f\|_\infty \leq \bar{e}$.

Assumption 1: Nonlinear function f is always Lipschitz continuous as:

$$\forall x, y \in \mathbb{K}_{1:N_n}, \|f(s(x)) - f(s(y))\|_\infty \leq L^* \|s(x) - s(y)\|_\infty$$

where L^* is the bounded Lipschitz constant, satisfying $L^* \leq \bar{L}^*$, where $\bar{L}^*$ is the upper bound.

Remark 1: Note that Lipschitz constant L^* is bounded, and sets $\mathbb{X}$, $\mathbb{U}$ and $\mathbb{W}$ are assumed to be compact sets. Function f is also bounded.

Definition 1 (Kinky Inference (KI) prediction function): With a sampled data set D_n obtained, $\hat{f}(s(k), L_n, D_n)$: $\mathbb{S}$

$\to\mathbb{Y}$ is calculated by a KI prediction function:

$$\hat{f}_j(s(k),L_n,D_n) := \frac{1}{2}\left(u_j\left(\tilde{f}_j(s(r)),s(k),L_n,N_n\right) + \ell_j\left(\tilde{f}_j(s(r)),s(k),L_n,N_n\right)\right), j \in \mathbb{K}_{1:n_x} \quad (25.)$$

$$u\left(\tilde{f}_j(s(r)),s(k),L_n,N_n\right) := \min_{r \in \mathbb{K}_{1:N_n}} \tilde{f}_j(s(r)) + L_n\|s(k) - s(r)\|_\infty + \bar{e} \quad (26.)$$

$$\ell\left(\tilde{f}_j(s(r)),s(k),L_n,N_n\right) := \max_{r \in \mathbb{K}_{1:N_n}} \tilde{f}_j(s(r)) - L_n\|s(k) - s(r)\|_\infty - \bar{e} \quad (27.)$$

where L_n is the estimated Lipschitz constant, and the time complexity for the KI prediction function is $\mathcal{O}(2N_n\mathcal{M})$. $\mathcal{M}$ is the time complexity for calculating the norm $\|\cdot\|_\infty$.

Definition 2 (LACKI rule): Estimated Lipschitz constant L_n is updated based on the LACKI rule:

$$L_n = \max\left\{L_{n-1}, \max_{\forall x,y \in \mathbb{K}_{1:N_n}, x \neq y} \frac{\|\tilde{f}(s(x)) - \tilde{f}(s(y))\|_\infty - \lambda}{\|s(x) - s(y)\|_\infty}\right\} \quad (28.)$$

where $\lambda > 0$ is a hyperparameter, and L_0 can be set as a small positive constant.

Remark 2: The positive hyperparameter λ is mainly used to deal with the unknown bounded process noise $w(k)$:

$$L_n \leq \frac{\|f(s(x)) - f(s(y))\|_\infty - (\lambda - 2\bar{e})}{\|s(x) - s(y)\|_\infty} \leq L^* + \frac{2\bar{e} - \lambda}{\|s(x) - s(y)\|_\infty}, \lambda \geq 2\bar{e}$$

if $L_n > L^*$, there will be an overfitting to the unknown process noise. That's why the minimum value of λ is $2\bar{e}$, and the estimated Lipschitz constant L_n is bounded:

$$L_n \leq \sup_{\forall x,y \in \mathbb{K}_{1:N_n}, x \neq y} \frac{\|f(s(x)) - f(s(y))\|_\infty - (\lambda - 2\bar{e})}{\|s(x) - s(y)\|_\infty} \leq L^* \quad (29.)$$

Definition 3 (minimal input space): Let the space $\mathcal{L} \subset \mathbb{S}$ be the minimum input space, containing all possible sample points, and denote the minimum number of sample points by $N_{\mathcal{L}}$.

Remark 3: Note that the real-time performance of the KI prediction function or the LACKI rule is related to the number of recorded samples N_n. If the number of recorded samples N_n is reduced, the computational burden of Eq. (25) and Eq. (28) can be reduced. To achieve good real-time performance, the nominal minimum number of sample points $N_{\bar{\mathcal{L}}}$ is used, which is assumed to be obtained from offline data. Denote the space $\bar{\mathcal{L}} \subset \mathbb{S}$ by the nominal minimal input space.

Here are two lemmas for the subsequent analysis.

Lemma 1 (Lipschitz Continuity, (Calliess et al., 2020)**):** The KI prediction function $\hat{f}(s(k),L_n,D_n)$ is also Lipschitz continuous with Lipschitz constant L_n:

$$\forall x,y \in \mathbb{K}_{1:N_n}, \left\|\hat{f}(s(x),L_n,D_n) - \hat{f}(s(y),L_n,D_n)\right\|_\infty \leq L_n\|s(x) - s(y)\|_\infty \quad (30.)$$

Lemma 2 (Sample-consistency, (Calliess et al., 2020)**):** If the minimum input space $\mathcal{L}$ is obtained**,** the LACKI rule is sample-consistent (up to $\frac{\lambda}{2}$), and the NPL estimated error $\|f-\hat{f}\|_\infty$ is bounded:

$$\forall r \in \mathbb{K}_{1:N_n}, \hat{f}(s(r),L_n,D_n) \in \mathcal{B}_{\frac{\lambda}{2}}\left(\tilde{f}(s(r))\right) \quad (31.)$$

$$\left\|f(s(r))-\hat{f}(s(r),L_n,D_n)\right\|_\infty \leq \frac{\lambda}{2}+\bar{e} \quad (32.)$$

where $\mathcal{B}_{\frac{\lambda}{2}}\left(\tilde{f}(s(r))\right)=\left\{y \in \mathbb{Y} \middle| \|y-\tilde{f}(s(r))\|_\infty \leq \frac{\lambda}{2}\right\}$ denotes the $\frac{\lambda}{2}$-ball around the measured output.

Learning rule: The nominal minimum number $N_{\bar{\mathcal{L}}}$ and the estimated Lipschitz constant L_n are two hyperparameters obtained offline, introduced in section 4.3. Different from the fixed sampled data set in (Zheng et al., 2023), the sampled data set $D_{N_{\bar{\mathcal{L}}}}$ is updated over time:

$$D_{N_{\bar{\mathcal{L}}}}=\left\{\left(s(r),\tilde{f}(s(r))\right) | r \in \mathbb{K}_{k-N_{\bar{\mathcal{L}}}+1:k}\right\} \quad (33.)$$

Remark 4: The sampled data set $D_{N_{\bar{\mathcal{L}}}}$, consists of sampled data from a continuous system arranged in time with the same sampling period. Under the learning rule, the distance between every two samples is bounded. At each sampling moment k, the oldest sample point $s\left(k-N_{\bar{\mathcal{L}}}\right)$ is discarded and the latest sample point $s(k)$ is retained. This construction method maintains the number of recorded samples $N_{\bar{\mathcal{L}}}$ by updating the sampled dataset. Under the **Learning rule**, the time complexity for the KI prediction function is $\mathcal{O}\left(2N_{\bar{\mathcal{L}}}\mathcal{M}\right)$.

Definition 6 (KI estimated error (Zheng et al., 2023)**):** If the minimum input space is obtained, KI estimated error $\left\|\mathfrak{b}\left(s(k),L_n,D_{N_{\mathcal{L}}}\right)\right\|_\infty$ is bounded:

$$\mathfrak{b}_j\left(\tilde{f}_j(s(r)),s(k),L_n,D_{N_{\mathcal{L}}}\right)=\frac{1}{2}\left(\mathcal{u}_j\left(\tilde{f}_j(s(r)),s(k),L_n,N_{\mathcal{L}}\right)-\ell_j\left(\tilde{f}_j(s(r)),s(k),L_n,N_{\mathcal{L}}\right)\right), j \in \mathbb{K}_{1:n_x}$$
$$\sup\left\|\mathfrak{b}\left(s(k),L_n,D_{N_{\mathcal{L}}}\right)\right\|_\infty=\bar{e} \quad (34.)$$

3.2 Theoretical Analysis

Lemma 2 and **Definition 6** have shown that if the minimum input space $\mathcal{L}$ is obtained, both the NPL estimated error and the KI estimated error are all bounded. It is important to note that the **Learning rule** is designed based on the nominal minimal input space. The conservative bounds of the NPL estimated error, as determined by the **Learning rule,** will be analyzed in the following sections. Here, we present two definitions and a lemma for further analysis:

Definition 4 (ε-convergent, (Zheng et al., 2023)**):** For a positive constant ε**,** denote a point set $\mathcal{s}_{\mathbb{N}}=\{s(r)|r \in \mathbb{K}_{1:\mathbb{N}}\}$ is ε-convergent to a point $s(r^*) \in \mathbb{S}$ by $\mathcal{s}_{\mathbb{N}} \xrightarrow{\varepsilon} s(r^*)$: iff $\exists m \in \mathbb{K}_{1:\mathbb{N}}, \forall r>m, \|s(r)-s(r^*)\|_\infty \leq \varepsilon$.

Definition 5 (ε-denseness, (Zheng et al., 2023)**):** For a positive constant ε, denote a set sequence $\mathcal{S}_{\mathbb{N}} = \{\mathcal{s}_r | r \in \mathbb{K}_{1:\mathbb{N}}\}$ is ε-dense to a point $s(r^*) \in \mathcal{s}_{\mathbb{N}}$ by $\mathcal{S}_{\mathbb{N}} \xrightarrow{\varepsilon} s(r^*)$: if $\exists n \in \mathbb{K}_{1:\mathbb{N}}$, $\mathcal{s}_n \xrightarrow{\varepsilon} s(r^*) \wedge \mathcal{s}_n \in \mathcal{S}_{\mathbb{N}}$. Denote a set sequence $\mathcal{S}_{\mathbb{N}}$ is ε-dense to a point set $\mathcal{s}_{\mathbb{N}}$ by $\mathcal{S}_{\mathbb{N}} \xrightarrow{\varepsilon} \mathcal{s}_{\mathbb{N}}$: $\forall s(r^*) \in \mathcal{s}_{\mathbb{N}}$:$\mathcal{S}_{\mathbb{N}} \xrightarrow{\varepsilon} s(r^*)$.

Lemma 3 (Proof given in Appendix.A)**:** Using the learning rule, for any sample point $s(k)$,r,$k \in \mathbb{K}_{k-N_{\overline{\mathcal{L}}}+1:k}$ and positive constants p, there always exists a ε^*-ball around $s(k)$, which is not all contained in the sampled data set $D_{N_{\mathcal{L}}}$, with $\varepsilon^* = p/L^*$:

$$\mathfrak{b}_j\left(\tilde{f}_j(r), s(k), L_n, N_{\overline{\mathcal{L}}}\right) \le L_n \varepsilon^* + \overline{e} = \frac{L_n}{L^*} p + \overline{e} \le p + \overline{e}, r \in \mathbb{K}_{k-N_{\overline{\mathcal{L}}}+1:k} \quad (35.)$$

which denotes the conservative bound of the KI estimated error by $\varepsilon = p + \overline{e}$.

Theorem 1: Using the learning rule, the nominal minimal input space $\overline{\mathcal{L}}$ is $\varepsilon_{\overline{\mathcal{L}}}$-dense to the minimal input space $\mathcal{L}$:

$$\mathrm{P}\left(\overline{\mathcal{L}} \xrightarrow{\varepsilon_{\overline{\mathcal{L}}}} \mathcal{L}\right) = (36.)$$

where $\varepsilon_{\overline{\mathcal{L}}} = p/L_n > \varepsilon^*$.

Proof: From **Lemma 3,** the ε^*-ball around a sample point $s(k) \in \mathbb{S}$,$k \in \mathbb{K}_{k-N_{\overline{\mathcal{L}}}+1:k}$ is equivalent to a neighborhood. Note that the unknown Lipschitz constant L^* is estimated by L_n using the hyperparameter λ in **Definition 2**. Then, for the nominal minimal input space, a larger neighborhood $\overline{\mathcal{L}}$-ball can be obtained:

$$\mathcal{B}_{\varepsilon^*}(s(k)) \subset \mathcal{B}_{\varepsilon_{\overline{\mathcal{L}}}}(s(k)) = \left\{s(r) \in \mathbb{S} \middle| \|s(k) - s(r)\|_\infty \le \varepsilon_{\overline{\mathcal{L}}}\right\}, k \in \mathbb{K}_{k-N_{\overline{\mathcal{L}}}+1:k}, r \in \mathbb{K}_{k-N_{\mathcal{L}}+1:k} \quad (37.)$$

In **Definition 5,** $\overline{\mathcal{L}} \xrightarrow{\varepsilon_{\overline{\mathcal{L}}}} \mathcal{L}$ also satisfies:

$$\forall s \in \mathcal{L}, \exists m \in \mathbb{K}_{k-N_{\mathcal{L}}+1:k}, \forall r > m: \mathrm{dist}\left(\overline{\mathcal{L}}, r\right) \le \varepsilon_{\overline{\mathcal{L}}} \quad (38.)$$

where the function $\mathrm{dist}\left(\overline{\mathcal{L}}, s\right)$ satisfies: $\mathrm{dist}\left(\overline{\mathcal{L}}, r\right) = \min_{s^* \in \overline{\mathcal{L}}} \|s^* - s(r)\|_\infty, r \in \mathbb{K}_{k-N_{\mathcal{L}}+1:k}$, and inequality (38) will be proved by contradiction. Give an assumption:

$$\exists s^* \in \mathcal{L}, \forall m \in \mathbb{K}_{k-N_{\mathcal{L}}+1:k}, \exists r > m: \mathrm{dist}\left(\overline{\mathcal{L}}, r\right) > \varepsilon_{\overline{\mathcal{L}}} \quad (39.)$$

which means there is no sample point recorded in the $\varepsilon_{\overline{\mathcal{L}}}$-ball around s^*.That is $\mathcal{B}_{\varepsilon_{\overline{\mathcal{L}}}}(s(k)) \cap \overline{\mathcal{L}} = \phi$. Define a probability by:

$$p_r(s^*) = \mathrm{P}\left(s(r) \in \mathcal{B}_{\varepsilon_{\overline{\mathcal{L}}}}(s(k))\right), r \in \mathbb{K}_{k-N_{\overline{\mathcal{L}}}+1:k} \quad (40.)$$

Then, inequality (39) is equivalent to

$$\mathrm{P}\left(\mathcal{B}_{\varepsilon_{\overline{\mathcal{L}}}}(s(k)) \cap \overline{\mathcal{L}} = \phi\right) = \prod_{r=k-N_{\overline{\mathcal{L}}}+1}^{k} (1 - p_r(s^*)) \quad (41.)$$

Note that the probability $p_r(s^*)$ is always nonnegative, and for $p_r(s^*) > 0$, it satisfies:

$$\lim_{\overline{\mathcal{L}} \to \infty} \mathrm{P}\left(\mathcal{B}_{\varepsilon_{\overline{\mathcal{L}}}}(s(k)) \cap \overline{\mathcal{L}} = \phi\right) = \quad (42.)$$

which means the assumption fails. Then, the proof is completed. □

Here gives a lemma for further analysis of the conservative bound of the NPL estimated error:

Lemma 4: (Proof given in Appendix.B) For any sample point $\in \overline{\mathcal{L}}$, denote $\xi = \underset{\xi \in \overline{\mathcal{L}}}{\operatorname{argmin}} \|s(\xi) - s(r)\|_\infty, r \in \mathbb{K}_{k-N_{\mathcal{L}}+1:k}$ by the nearest sample point in the nominal minimal input space $\overline{\mathcal{L}}$. With $\lambda = 2\overline{e}$, it holds that:

$$\ell_j\left(\tilde{f}_j(r), s(k), L_n, N_{\mathcal{L}}\right) \leq \tilde{f}_j(s(\xi)) + L_n \|s(\xi) - s(r)\|_\infty + \overline{e}, r \in \mathbb{K}_{k-N_{\mathcal{L}}+1:k} \quad (43.)$$

and for $\lambda > 2\overline{e}$, **Lemma 4** still holds.

Theorem 2: For the minimal input space $\mathcal{L}$, the NPL estimated error satisfies:

$$\left\|f(s(r)) - \hat{f}(s(k), L_n, D_{\mathcal{L}})\right\|_\infty \leq \frac{L_n + L^*}{L_n} p + 2\overline{e}, r \in \mathbb{K}_{k-N_{\mathcal{L}}+1:k} \quad (44.)$$

and using the **Learning Rule**, the NPL estimated that the error in inequality (44) still holds.

Proof: for the nearest sample point $s(\xi) \in \mathcal{L}$, it holds:

$$u\left(\tilde{f}_j(r), s(k), L_n, N_{\mathcal{L}}\right) \leq \tilde{f}_j(s(\xi)) + L_n \|s(\xi) - s(r)\|_\infty + \overline{e}, r \in \mathbb{K}_{k-N_{\mathcal{L}}+1:k} \quad (45.)$$

From **Lemma 4,** it holds:

$$\begin{aligned} \hat{f}_j(s(k), L_n, D_{\mathcal{L}}) &= \frac{1}{2}\left(u_j\left(\tilde{f}_j(r), s(k), L_n, N_{\mathcal{L}}\right) + \ell_j\left(\tilde{f}_j(r), s(k), L_n, N_{\mathcal{L}}\right)\right) \\ &\leq \tilde{f}_j(s(\xi)) + L_n \|s(\xi) - s(r)\|_\infty + \overline{e}, r \in \mathbb{K}_{k-N_{\mathcal{L}}+1:k} \quad (46.) \end{aligned}$$

which can be rewritten as:

$$\left\|\hat{f}_j\left(s(k), L_n, D_{\overline{\mathcal{L}}}\right) - \tilde{f}_j(s(\xi))\right\|_\infty \leq L_n \|s(\xi) - s(r)\|_\infty + \overline{e}, r \in \mathbb{K}_{k-N_{\mathcal{L}}+1:k} \quad (47.)$$

Combining the bounded process noise $\|\tilde{f} - f\|_\infty \leq \overline{e}$ and **Assumption 1**, it holds that:

$$\left\|\tilde{f}_j(s(\xi)) - f(s(k))\right\|_\infty \leq L^* \|s(\xi) - s(r)\|_\infty + \overline{e}, r \in \mathbb{K}_{k-N_{\mathcal{L}}+1:k} \quad (48.)$$

Based on triangle inequality, sum inequalities (47) and (48):

$$\begin{aligned}\left\|f(s(k))-\hat{f}(s(k),L_n,D_{\mathcal{L}})\right\|_\infty &\le (L^*+L_n)\|s(\xi)-s(r)\|_\infty+2\bar{e}\\ &\le \frac{L_n+L^*}{L_n}p+2\bar{e},r\in\mathbb{K}_{k-N_{\mathcal{L}}+1:k}(49.)\end{aligned}$$

where the second inequality holds in **Theorem 1.**

The validity of inequality (44) can be obtained by providing $\bar{\mathcal{L}}\overset{\varepsilon_{\bar{\mathcal{L}}}}{\to}\mathcal{L}$ using the **Learning Rule**. According to **Theorem 1,** the nominal minimal input space $\bar{\mathcal{L}}$ is $\varepsilon_{\bar{\mathcal{L}}}$-dense to the input space $\mathcal{L}$. Then, the proof is completed. □

4. Learning-based Path Tracking Control Strategy

The control objective is that the ROV tows the MRT accurately to track the reference path point. A hierarchical control strategy consisting of a VSMPC controller and an LBMPC controller is developed, whose scheme is shown in Fig.5.

Given a vector $\eta_r=(x_r,y_r,\psi_r,u_r)^{\mathrm{T}}$ that stands for the reference path point, where x_r and y_r are reference position coordinates. ψ_r is the reference yaw angle. u_r is the reference surge speed of the MRT. Given a vector $\eta=(x,y,\psi,u_{MRT})^{\mathrm{T}}$, the path tracking deviation is denoted by $e_\eta:=(\eta-\eta_r)=(e_x,e_y,e_\psi,e_u)^{\mathrm{T}}$. The VSMPC controller is developed based on kinematic Model (3) to obtain the ROV's virtual speed control law $\bar{v}_{ROV}$ to converge path tracking deviation.

With the uncertain nonlinear dynamics in Kinetic Model (6) learned through the proposed NPL method, the LBMPC controller is developed to compute the control input τ by formulating an optimal control problem aimed at tracking the ROV's virtual speed control law $\bar{v}_{ROV}$.

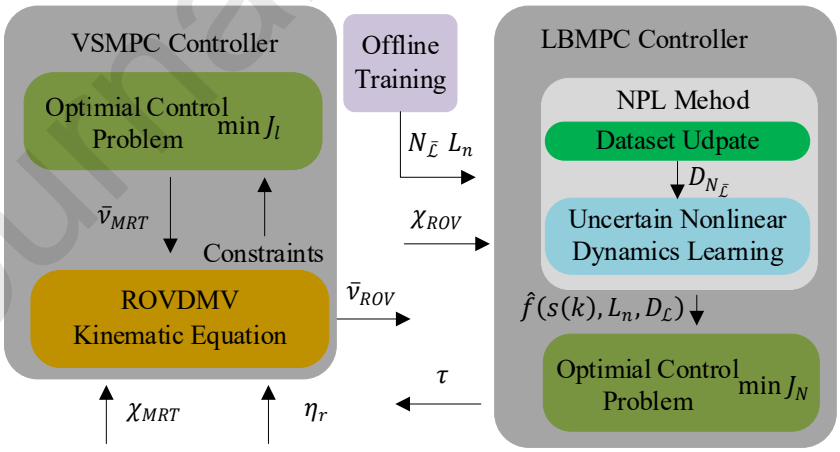

Fig.5 Scheme of the learning-based path tracking control strategy

4.1 VSMPC Controller

Based on the kinematic model of ROVDMV in Eq. (3), the discretized state space model of the MRT can be obtained:

$$\chi_{MRT}(k+1) = A_M\chi_{MRT}(k) + B_M\mathcal{U}_{MRT}(k) \quad (50.)$$

where increment of the MRT's speed $\mathcal{U}_{MRT} := (\Delta u_{MRT}, \Delta v_{MRT}, \Delta\omega_{MRT})^{\mathrm{T}}$ is the control input. Denote the state of the MRT by $\mathcal{X}_{MRT} := (\eta; \upsilon_{MRT})$. State matrix A_M and control matrix B_M satisfy:

$$A_M = \begin{bmatrix} I_3 & T_l\mathcal{J} \\ 0_{3\times 3} & I_3 \end{bmatrix}, B_M = \begin{bmatrix} T_l\mathcal{J} \\ I_3 \end{bmatrix} \quad (51.)$$

where T_l is the sampling time of the VSMPC controller. $\mathcal{J}$ represents a transformation matrix:

$$\mathcal{J} = \begin{bmatrix} \cos\psi & -\sin\psi & 0 \\ \sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (52.)$$

To minimize the path tracking deviation with smooth speed change, a cost function J_l is designed as follows:

$$J_l = \sum_{i=0}^{N_l-1} \|e_\eta(i|k)\|_{Q_l}^2 + \|\mathcal{U}_{MRT}(i|k)\|_{R_l}^2 + \|e_\eta(N_l|k)\|_{P_l}^2 \quad (53.)$$

where N_l is predictive horizon in the VSMPC controller. Q_l and R_l are weight matrices. Denote terminal weight matrix in the VSMPC controller by P_l.

Note that the MRT's speed and the increment of the MRT's speed are bounded, and constraints can be given as:

$$\upsilon_{MRT} \in \mathbb{V}_l := \{\upsilon_{MRT} | \upsilon_{MRT_{\min}} \le \upsilon_{MRT} \le \upsilon_{MRT_{\max}}\}, \mathcal{U}_{MRT} \in \mathbb{U}_l := \{\mathcal{U}_{MRT} | \mathcal{U}_{MRT_{\min}} \le \mathcal{U}_{MRT} \le \mathcal{U}_{MRT_{\max}}\} \quad (54.)$$

where $\upsilon_{MRT_{\max}}$ and $\upsilon_{MRT_{\min}}$ are known maximum and minimum MRT's speed. $\mathcal{U}_{MRT_{\max}}$ and $\mathcal{U}_{MRT_{\min}}$ represent known maximum and minimum increments of the MRT's speed.

Then, an optimal control problem is formulated to calculate the MRT's virtual speed control law ν_{MRT_d}:

$$\begin{aligned} &\min_{\mathcal{U}_{MRT}(i|k), i\in\mathbb{K}_{0:N_l-1}} \quad J_l \\ &\text{s.t.} \begin{array}{l} \chi_{MRT}(0|k) = \chi_{MRT}(k), \mathcal{U}_{MRT}(0|k) = \mathcal{U}_{MRT}(k) \\ \chi_{MRT}(i+1|k) = A_M\chi_{MRT}(i|k) + B_M\mathcal{U}_{MRT}(i|k) \\ \upsilon_{MRT}(i|k), \upsilon_{MRT}(N_l|k) \in \mathbb{V}_l \\ \mathcal{U}_{MRT}(i|k) \in \mathbb{U}_l \end{array} \end{aligned} \quad (55.)$$

where denote the solution to the optimal control problem by $\mathcal{U}^*_{MRT}(i|k), i \in \mathbb{K}_{0:N_l-1}$, and the MRT's virtual speed control law ν_{MRT_d} can be obtained:

$$\nu_{MRT_d} = \mathcal{U}^*_{MRT}(0|k) + \upsilon_{MRT}(k) \quad (56.)$$

According to the kinematic model of ROVDMV (3), the ROV's virtual speed control law ν_d can be calculated:

$$\nu_d = (\bar{u}_{ROV}, \bar{v}_{ROV}, \bar{\delta})^{\mathrm{T}} = \left(\bar{u}_{\mathrm{MRT}}, \bar{v}_{MRT}, \frac{\bar{\omega}_{MRT}}{\bar{u}_{\mathrm{MRT}}} L\right)^{\mathrm{T}} \quad (57.)$$

Remark 5: At any time k, with the state matrix A_M and the control matrix B_M in optimal control problem (55) assumed to be constant, the optimal control problem (55) is equivalent to a quadratic programming (QP) problem. It is important to note that solving a QP problem involves minimal computational burden. Consequently, calculating the

virtual speed control of the ROV will not adversely affect the real-time performance of the learning-based path tracking control strategy.

4.2 LBMPC Controller

Based on thedeductions from the kinetic model of the ROVDMV in Section 2.2, a discretized Lipschitz nonlinear model (Yu, Maier, Chen, & Allgöwer, 2013) can be obtained:

$$\chi_{ROV}(k+1) = A_R\chi_{ROV}(k) + B_R\mathcal{U}_{ROV}(k) + f(\chi_{ROV}(k)) + \mathcal{w}_{MRT}(k) \quad (58.)$$

where $\mathcal{w}_{MRT} \in \mathcal{R}^{4\times 1}$ is the unknown bounded process noise from the rugged seafloor. Denote the ROV state by $\chi_{ROV} := (u_{ROV}, v_{ROV}, \omega_{ROV}, \delta)^{\mathrm{T}}$. The state matrix A_R and the control matrix B_R satisfy:

$$A_R = \mathcal{M}^{-1}\begin{bmatrix} 1+T_R X_u & 0 & 0 & 0 \\ 0 & 1+T_R Y_v & 0 & 0 \\ 0 & 0 & 1+T_R N_r & 0 \\ 0 & 0 & T_R & 1 \end{bmatrix}, B_R = \mathcal{M}\begin{bmatrix} T_R & 0 & 0 \\ 0 & T_R & 0 \\ 0 & 0 & T_R \\ 0 & 0 & 0 \end{bmatrix},$$
$$\mathcal{M} = \begin{bmatrix} m_{MRT}+m_{ROV}+m_\delta - X_{\dot{u}} & 0 & 0 & 0 \\ 0 & m_{MRT}+m_{ROV}-Y_{\dot{v}} & 0 & 0 \\ 0 & 0 & I_{ROV}-N_{\dot{r}} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (59.)$$

where T_R is the sampling time of the LBMPC controller. $\mathcal{M}$ is the mass matrix. Note that the added mass m_δ is related to the horizontal rotation angle δ, and the state matrix A_R and the control matrix B_R are time varying. To match the Lipschitz nonlinear model (Yu et al., 2013), the added mass m_δ is simplified as a constant.

The function $f(\chi_{ROV})$, consisting of the nonlinear dynamics, $g(\chi_{ROV})$, the model mismatch from parametric uncertainty and the simplification of the added mass m_δ, $\tilde{g}$ and unknown external ocean current disturbance, h_e, is given:

$$f(\chi_{ROV}) = \mathcal{M}^{-1}(g(\chi_{ROV}) + \tilde{g}) + h_e \quad (60.)$$

where

$$g(\chi_{ROV}) = \begin{bmatrix} -F_R\cos\delta_{ROV} \\ F_R\sin\delta_{ROV} \\ -\mathrm{sg}(\omega_{\mathrm{MRT}})M_{o_2} \\ 0 \end{bmatrix} + \begin{bmatrix} (X_{vr}+m_{ROV})v_{ROV}\omega_{ROV} + X_{rr}\omega_{ROV}^2 + X_{uu}|u_{ROV}|u_{ROV} \\ Y_{uv}u_{ROV}v_{ROV} + (Y_{ur}-m_{ROV})u_{ROV}\omega_{ROV} + Y_{vv}|v_{ROV}|v_{ROV} \\ N_{uv}u_{ROV}v_{ROV} + N_{ur}u_{ROV}\omega_{ROV} + I_{MRT}u_{ROV}^2\delta/L^2 + N_{rr}|\omega_{ROV}|\omega_{ROV} \\ 0 \end{bmatrix}, h_e$$
$$= \mathcal{M}^{-1}\begin{bmatrix} \tau_e \\ 0 \end{bmatrix}$$

With the function $f(\chi_{ROV})$ learned by the novel NPL method, the state transition model used in the LBMPC controller is given as:

$$\chi_{ROV}(k+1) = A_R\chi_{ROV}(k) + B_R\mathcal{U}_{ROV}(k) + \hat{f}\left(s_{ROV}(k), L_n, D_{N_{\bar{I}}}\right) \quad (61.)$$

in which an extended state is given: $s_{ROV}(k) = (\chi_{ROV}(k), \mathcal{U}_{ROV}(k))$.

Denote the change of the control input by $\Delta\mathcal{U}_{ROV}(k) = \mathcal{U}_{ROV}(k+1) - \mathcal{U}_{ROV}(k)$. To minimize the virtual speed

deviation with smooth control input, the cost function J_n is designed as follows:

$$J_n = \sum_{i=0}^{N_n-1} \|e_v(i|k)\|_{Q_n}^2 + \sum_{i=0}^{N_n-2} \|\Delta\mathcal{U}_{ROV}(i|k)\|_{R_n}^2 + \|e_v(N_n|k)\|_{P_n}^2 \quad (62.)$$

where N_n is the predictive horizon in the VSMPC controller. Q_n and R_n are weight matrices. Denote terminal weight matrix in the LBMPC controller by P_n. Consider constraints on the control input and the ROV's speed:

$$\begin{aligned} \upsilon_{ROV} \in \mathbb{V}_n &:= \{\upsilon_{ROV} | \upsilon_{ROV_{\min}} \le \upsilon_{ROV} \le \upsilon_{ROV_{\max}}\}, \\ (\mathcal{U}_{ROV}, \Delta\mathcal{U}_{ROV}) \in \mathbb{U}_l &:= \{(\mathcal{U}_{ROV}, \Delta\mathcal{U}_{ROV}) | \mathcal{U}_{ROV_{\min}} \le \mathcal{U}_{ROV} \le \mathcal{U}_{ROV_{\max}}, \Delta\mathcal{U}_{ROV_{\min}} \le \Delta\mathcal{U}_{ROV} \le \Delta\mathcal{U}_{ROV_{\max}}\} \end{aligned} \quad (63.)$$

Then, an optimal control problem is formulated to calculate the control input:

$$\begin{aligned} &\min_{\mathcal{U}_{ROV}(i|k), i \in \mathbb{K}_{0:N_n-1}} J_n \\ &\chi_{ROV}(0|k) = \chi_{ROV}(k), \mathcal{U}_{ROV}(0|k) = \mathcal{U}_{ROV}(k) \\ \text{s.t.} \quad &\chi_{ROV}(i+1|k) = A_R\chi_{ROV}(i|k) + B_R\mathcal{U}_{ROV}(i|k) + \hat{f}\left(s_{ROV}(i|k), L_n, D_{N_{\bar{\mathcal{L}}}}\right) \\ &\upsilon_{ROV}(i|k), \upsilon_{ROV}(N_l|k) \in \mathbb{V}_n \\ &\{\mathcal{U}_{ROV}(i|k), \Delta\mathcal{U}_{ROV}(i|k)\} \in \mathbb{U}_n \end{aligned} \quad (64.)$$

where the solution to the optimal control problem is denoted by $\mathcal{U}^*_{ROV}(i|k), i \in \mathbb{K}_{0:N_n-1}$. Then, the control input can be obtained: $\tau(k) = \mathcal{U}^*_{ROV}(0|k)$.

Remark 6: There is no terminal constraint introduced in optimal control problems (55) and (64), and the stability analysis of the corresponding closed-loop system can be found in (Zheng et al., 2023).

4.3 Offline Training

In this section, the offline training in Fig.4 is introduced as two processes: coarse tuning and fine tuning. A nominal MPC controller is designed to conduct one path tracking numerical simulation in the coarse tuning process. Based on the numerical simulation results, the LACKI rule is utilized to preliminarily obtain an estimated Lipschitz constant L_n, and an offline training algorithm is developed to preliminarily select a nominal minimum number of sample points $N_{\bar{\mathcal{L}}}$. In the c process, the estimated Lipschitz constant L_n and the nominal minimum number $N_{\bar{\mathcal{L}}}$ are utilized in the KI prediction function to formulate LBMPC controller in section 4.2. Then, similar to the coarse tuning process, another algorithm to optimize the two hyper-parameters is developed.

The nominal MPC controller is used to calculate the ROV's control input, and optimal control problem (64) becomes:

$$\begin{aligned} &\min_{\mathcal{U}_{ROV}(i|k), i \in \mathbb{K}_{0:N_n-1}} J_n \\ &\chi_{ROV}(0|k) = \chi_{ROV}(k), \mathcal{U}_{ROV}(0|k) = \mathcal{U}_{ROV}(k) \\ s.t. \quad &\chi_{ROV}(i+1|k) = A_R\chi_{ROV}(i|k) + B_R\mathcal{U}_{ROV}(i|k) + g(\chi_{ROV}(i|k)) \\ &\upsilon_{ROV}(i|k), \upsilon_{ROV}(N_l|k) \in \mathbb{V}_n \\ &\{\mathcal{U}_{ROV}(i|k), \Delta\mathcal{U}_{ROV}(i|k)\} \in \mathbb{U}_n \end{aligned} \quad (65.)$$

when one path tracking numerical simulation is done, a sampled data set $D_{\bar{n}}$ can be obtained. Then, define the mean square error of the KI estimated error by:

$$\text{MSE} = \frac{1}{\bar{n}} \sum_{i=1}^{\bar{n}} \left\| \mathfrak{b}\left(s(i), L_n, D_{\bar{\mathcal{L}}}\right) \right\|_\infty \quad (66.)$$

, and then the two offline training algorithms are given:

Algorithm 1 offline training algorithm for coarse tuning

---

1. Based on the sampled data set $D_{\overline{n}}$, using **LACKI rule** to calculate the estimated Lipschitz constant L_n

2. set $\mathrm{MSE}_{\mathrm{last}}$ to a large positive constant, and give a small positive constant σ

3. for $i = 2{:}\overline{n}$

4. calculate mean square error of the KI estimated error

5. if $|\mathrm{MSE} - \mathrm{MSE}_{\mathrm{last}}| < \sigma \& \mathrm{MSE} \leq \mathrm{MSE}_{\mathrm{last}}$

6. break

7. end

8. $\mathrm{MSE}_{\mathrm{last}} = \mathrm{MSE}$

9. end

10. set $N_{\overline{\mathcal{L}}} = i - 1$

Remark 7: From **Lemma 3**, the estimated error of the KI is bounded based on a nominal minimal input space. Theoretically, if the nominal minimal input space is appropriate, a small mean square error for the KI's estimated error can be achieved. Steps 4-9 in **Algorithm 1** are employed to determine the suitable nominal minimum number of sample points, which subsequently allows for the identification of the appropriate nominal minimal input space.

Note that the state transition model in the nominal MPC controller does not account for external ocean current disturbances and parametric uncertainties, which limits the path tracking performance. In this case, sampled data set used in the **LACKI rule** may not be suitable enough, and the nominal minimum number $N_{\overline{\mathcal{L}}}$ can be optimized. Then, the **Algorithm 2** is introduced to search the optimal hyperparameters:

Algorithm 2 hyperparameter optimization algorithm for coarse tuning

---

1. set $N_{\overline{\mathcal{L}}}$ and the estimated Lipschitz constant L_n obtained from previous training.

2. perform one path tracking numerical simulation to obtain the data set $D_{\overline{n}}$

3. based on the sampled data set $D_{\overline{n}}$,using the **LACKI rule** to update estimated Lipschitz constant L_n

4. set $\mathrm{MSE}_{\mathrm{last}}$ to a large positive constant

5. for $i = N_{\overline{L}}{:}\overline{n}$

6. calculate mean square error of the KI estimated error

7. if $|\mathrm{MSE} - \mathrm{MSE}_{\mathrm{last}}| < \sigma \& \mathrm{MSE} \leq \mathrm{MSE}_{\mathrm{last}}$

8. break

9. end

10. $\mathrm{MSE}_{\mathrm{last}} = \mathrm{MSE}$

11. end

12. set $N_{\overline{L}} = i - 1$

Remark 8: Similar to **Algorithm 1,** steps 4-10 in **Algorithm 1** are also used to obtain the suitable nominal minimal input space. **Algorithm 2** need be iterated to obtain the optimal parameters $N_{\overline{L}}$ and L_n.

5. Results and discussion

Numerical simulations are conducted to evaluate the learning performance of the novel NPL method and to verify the path tracking control performance of the learning-based path tracking control strategy. The dynamic model of the ROVDMV, as deduced in section 2, is employed as the controlled object. This paper focuses on the learning-based path tracking control strategy, utilizing scale models for the ROVDMV, with parameter values provided in Section 5.1. The learning-based path tracking control strategy is applied to the ROVDMV's path tracking control, where the uncertain nonlinear dynamics learned by the NPL method are treated as unknown. Numerical simulations are performed using MATLAB, with an AMD Ryzen Threadripper PRO 3995WX 64-Core 2.70 GHz CPU and 256 GB of RAM running Windows 10. Optimal control problems (64) and (65) are solved using the MATLAB function "fmincon(·)," while optimal control problem (55) is addressed using the MATLAB function "quadprog(·)."

To verify the superiority of the proposed learning-based path tracking control strategy, a comparative numerical simulation was conducted. The only difference with the proposed learning-based path tracking control strategy is that optimal the in the comparative numerical simulation is the absence of the optimal strategy. Specifically, there is no approach to address model mismatch arising from parametric uncertainties and external disturbances under process noise. "MPC" denotes simulation results from the comparative numerical simulation, while "LBMPC" denotes simulation results from the proposed learning-based path tracking control strategy. The reference path of the MRT is generated by a sine curve:

$$\begin{cases} x_r = t \\ y_r = 5\sin 0.05t \\ \psi_r = \arctan\left(0.25\cos\left(0.05t\right)\right) \\ u_r = 1 \end{cases} \quad (67.)$$

5.1 Preliminary validation based on a degenerated individual ROV model

Indeed, validating the ROVDMV model and the learning-based path tracking control strategy is essential for the research's feasibility. However, the deep-sea mining system discussed in this paper is a complex, kilometer-scale giant system, making it challenging to establish an experimental model for the underwater mining system. Due to constraints in experimental conditions and funding, we utilize results from a degraded ROV model found in the existing literature (Tijjani, Chemori, & Creuze, 2021) to validate the proposed LBMPC controller. The yaw angle control with the proposed LMPC controller is performed in the presence of complex, unknown external ocean current disturbances.

Table.1 Yaw angle control performance

Method	Average yaw angle deviation (°)	Max yaw angle Deviation (°)
Proposed control strategy	0.05	0.01
Comparative control strategy	0.07	0.08

As the deviation results shown in Table 1, the achieved reductions yaw angle deviations demonstrate feasibility and advantages of the proposed LMPC controller. In the proposed learning-based path tracking control strategy, the VSMPC controller converts the MRT's control objective into the individual motion control for the ROV. Given that the LBMPC controller has been shown to be effective for the individual control of the ROV, it is theoretically applicable to the ROVDMV as well.

5.2 Parameters Set

The parameters of the ROVDMV scaled model are presented in Table 2. The hydrodynamic parameters of the ROV are referenced from Chen and Bian (2023), while the parameters for the soft seafloor sediment are derived from Al-Milli et al. (2010), Dai et al. (2021), and Zeng et al. (2021).

Table 2. Parameters of the dynamic model of ROVDMV

Parameter	Value	Parameter	Value	Parameter	Value
L	0.5 m	l	2 m	b	1.4 m
m_{MRT}	40.5 kg	m_{ROV}	48.85 kg	I_{ROV}	11.6 kgm^2
I_{MRT}	45.85 kgm^2	m_{δ}	4.12 kg	$X_{\dot{u}}$	-3.9 kg
$Y_{\dot{v}}$	-149.9 kg	$N_{\dot{r}}$	-53.87 kgm^2	X_u	-4.1 kg/m

N_r	-547 kgm²/s/rad	Y_{vv}	-553.4 kg/m	X_{uu}	-8.2 kg
N_{rr}	-1037 kgm²/s/rad	Y_v	-285.7 kg/m	X_{vr}	-149.9 kg
Y_{ur}	-120.8 kg	X_{rr}	-13.18kgm/rad	Y_{uv}	-120.8 kg
N_{ur}	-13.6 kg	N_{uv}	-163.9 kg	E_1	0.95
E_2	-0.1	E_3	-0.1	CI	420
W	244.5 N	ϕ	6.2 °	N_r	0.1
N_c	6.36	r_s	12.2 kN/m³	c	5.4 kPa

Based on discretized state space models (50) and (58), a dynamical model of the ROVDMV is deduced as the controlled object in the numerical simulation. To verify the robustness against parametric uncertainty, external disturbance and process, the parametric uncertainty, which leads to a perturbation of the nonlinear dynamics $\tilde{g}$, is modeled as $\tilde{g} = 0.1g$.

Moreover, a complex unknown external ocean current disturbance, which combines the constant disturbance and the time varying disturbance, is introduced:

$$h_e = \begin{cases} 0.1 * +0.2 * \sin(0.5t) \\ 0.1 + 0.15 * \sin(0.5t) \\ 0.05 + 0.1 * \sin(0.5t) \\ 0 \end{cases} \quad (68.)$$

The unknown uniformly distributed bounded process noise is defined as: $w_{MRT} = \left(w^1_{MRT}, w^2_{MRT}, w^3_{MRT}, 0\right)^{\mathrm{T}}$, $w^1_{MRT} \in [-0.005, 0.005]$, $w^2_{MRT} \in [-0.005, 0.005]$, $w^3_{MRT} \in [-0.005, 0.005]$。

Parameters in the VSMPC controller and the LBMPC controller greatly influence on the path tracking performance. In the optimal control problem (55), the weight matrices Q_l and P_l are used to minimize the path tracking deviation, while the weight matrix R_l is used to ensure smooth changes in the virtual ROV speed control law. Similarly, in optimal control problems (64) and (65), the weight matrices Q_n and P_n are used to minimize virtual speed deviation, and the weight matrix R_n is used to achieve smooth control input. To intuitively demonstrate the superiority of the proposed learning-based path tracking control strategy, the weight matrices and constraints in optimal control problems (64) and (65) are all set to be identical. Table 3 presents the parameters used in the three optimal control problems (55), (64) and (65).

Table.3 Parameters used in the optimal control problems

Parameter	Value	Parameter	Value	Parameter	Value
Q_l	$(44,52,1595,305)^{\mathrm{T}}$	P_l	$(84,72,1895,355)^{\mathrm{T}}$	R_l	$(84,72,1895,355)^{\mathrm{T}}$
Q_n	$(44,10,133)^{\mathrm{T}}$	P_n	$(54,15,143)^{\mathrm{T}}$	R_n	$(30,10,75)^{\mathrm{T}}$
T_l	0.05	T_n	0.05	$v_{MRT_{\min}}$	$-(0,0.3,0.02$
$v_{MRT_{\max}}$	$(1.5,0.3,0.02)^{\mathrm{T}}$	$\mathcal{U}_{MRT_{\min}}$	$-(0.05,0.01,0.002)$	$\mathcal{U}_{MRT_{\max}}$	$-\mathcal{U}_{MRT_{\min}}$
$v_{ROV_{\min}}$	$-(0,0.3,0.01)^{\mathrm{T}}$	$v_{ROV_{\max}}$	$(1.5,0.3,0.01)^{\mathrm{T}}$	$\mathcal{U}_{ROV_{\min}}$	$-\mathcal{U}_{ROV_{\max}}$
$\mathcal{U}_{ROV_{\max}}$	$(80,80,95)^{\mathrm{T}}$	$\Delta\mathcal{U}_{ROV_{\max}}$	$(10,1,10)^{\mathrm{T}}$	$\Delta\mathcal{U}_{ROV_{\min}}$	$-\Delta\mathcal{U}_{ROV_{\max}}$

5.3 Learning Performance Evaluation

Note that the proposed NPL method is mainly developed from (Zheng et al., 2023). To intuitively demonstrate its superiority, numerical simulations are conducted on the stabilization control of the system from the literature (Zheng et al., 2023). As the results recorded in Table.4, the mean square error of the proposed NPL method is smaller than that of the comparative NPL method. After offline training, the number of the sample points in the proposed NPL method is nearly half of the comparative NPL method. Since the time complexity of the KI prediction equation is positively correlated with the number of samples, the simulation time for the whole stabilization control of the proposed NPL Method can be reduced by approximately 33%.

Table.4 Performance of the two NPL Methods

	Proposed NPL Method	Comparative NPL Method
Number of the sample points	6	13
Mean square error	0.0856	0.1251
Simulation time (s)	82.27	123.18

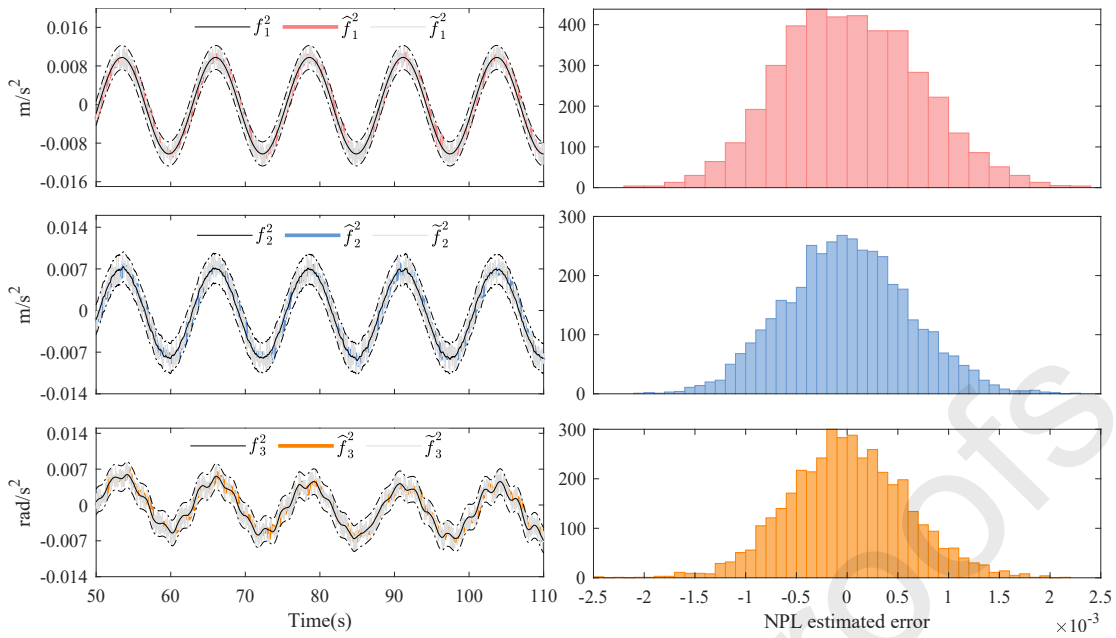

Fig.6 Results of the learned uncertain nonlinear dynamics (comparing the reference values, estimated values, actual values with noise (left) and results of histogram (right) of the NPL estimated error)

It can be seen the proposed NPL learning method exhibits better learning performance. Therefore, the following analysis focuses on the learning performance of the proposed NPL learning for the ROVDMV' uncertain nonlinear function.

According to **Remark 8,** with hyperparameter λ set to 0.0111**,** offline training is conducted, and the results are recorded in Table 4. The iteration step equals to 0 means the data set used in offline training is from simulation results of the comparative numerical simulation. The MSE converges to a constant value when the iteration step exceeds 3. Besides, the estimated Lipschitz constant L_n and the nominal minimum number of sample points $N_{\bar{\mathcal{L}}}$ also tend to be invariable. Then, $N_{\bar{\mathcal{L}}}$ is then set as 5, and Ln as 0.5578.

To visually evaluate the learning performance of the NPL method, with $N_{\bar{\mathcal{L}}}$ and L_n set according to the offline training results, a numerical simulation for path tracking is conducted using the proposed learning-based path tracking control strategy, with parameters set according to the offline training results. The numerical simulation results are presented in Figures 6 and 7, where the distribution of the NPL estimated error is illustrated in histograms (Figure 6, right). Since the values of the nonlinear function vary periodically, the results for the first 32 seconds are displayed in Figure 6 (left). The solid black line represents the actual values, while the solid gray line indicates the measured output values of the nonlinear function, which include process noise. To determine whether the NPL estimated error is bounded, the black dotted line—derived by adding and subtracting the upper bound of the process noise from the actual values—is utilized. The KI estimated error is depicted in Figure 7, where the scatterplot is shown on the left and the histogram is presented on the right.

Table 5 Offline training results

Iteration step	MSE	L_n	$N_{\bar{\mathcal{L}}}$
0	0.0043	0.2712	4
1	0.0032	0.5344	5
2	0.0031	0.5546	5
3	0.0029	0.5578	5
4	0.0029	0.5578	5
5	0.0029	0.5578	5

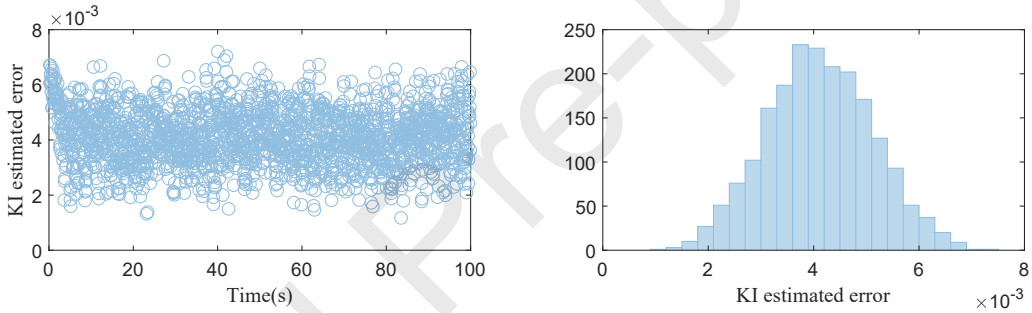

Fig.7 Results of scatterplot (left) and histogram (right) of the KI estimated error.

As shown in Fig.7, the estimated error for the KI is less than 0.8%, which satisfies the conservative upper bound analyzed in **Lemma 3**,set at 0.003. Additionally, as illustrated in Fig. 6, the estimated values align closely with the actual values, and the NPL estimated error remains bounded. The absolute value of the NPL estimated error is less than the conservative upper bound established in **Theorem 2**. When comparing the estimated values with the measured output values, their trends are consistent, indicating that the estimated values are closer to the actual values. This suggests that process noise is filtered to some extent by the NPL method. The histograms in Fig. 6 reveal a trend toward normal distribution in the estimated values, with the estimated errors primarily concentrated around zero. However, there is a time delay of one sampling period in the estimated values. According to the learning rule, the nominal minimal input space contains only sample points from previous moments, rather than encompassing all possible sample points. This explains the observed delay. Overall, the NPL method demonstrates good learning performance.

5.4 Path Tracking Performance Verification

To visually demonstrate the superiority of the proposed learning-based path tracking control strategy, Figure 8 illustrates the reference trajectory alongside the MRT's trajectory under both the and the path tracking deviation, where position deviation is denoted by $e_p = \sqrt{e_x^2 + e_y^2}$. Table 6 shows the average and maximum values of the path tracking

deviations, and the computing time. Compared with the MRT's trajectory under "LBMPC", the MRT's trajectory under "LBMPC" fits the reference trajectory better, especially at the amplitude of the sine curve. As shown in Fig.9 and Table 4, yaw angle deviations of the "MPC" and "LBMPC" are small, less than 1°. However, path tracking deviations of "LBMPC" are smaller than that of "MPC". Average position deviation of "LBMPC" is reduced from 0.32m to 0.01m, a reduction of about 97%. Maximum position deviation of "LBMPC" is reduced from 0.81m to 0.01m, a reduction of about 98%. It can be seen that the proposed learning-based path-tracking control strategy achieves ideal control performance and shows better robustness against strong disturbances and random noise. Moreover, the computing time of the "MPC" and "LBMPC" are almost the same. The difference is that the optimization control problem of LBMPC incorporates the KI prediction function in the state transition model. That means the NPL method will not bring much extra computational burden when solving the optimal control problem, and good real-time performance can be ensured.

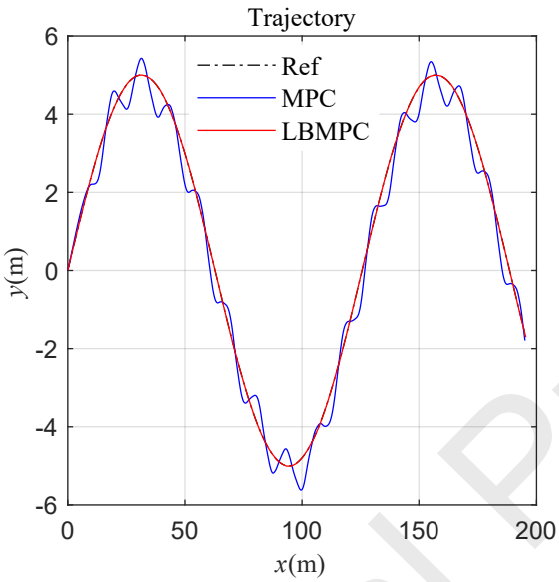

Fig. 8 MRT trajectory

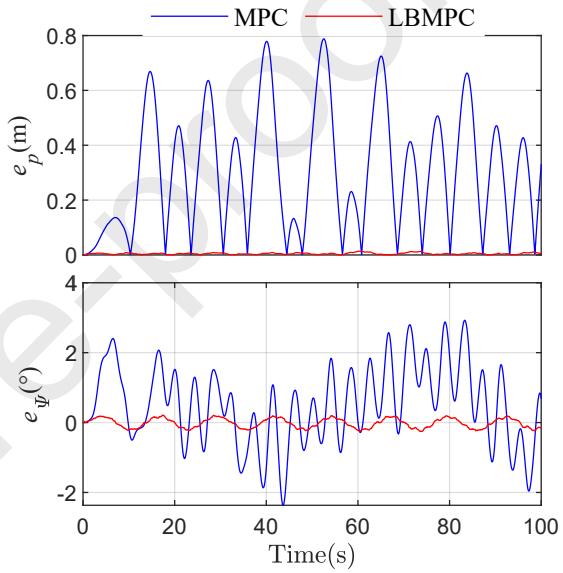

Fig. 9 Path tracking deviation

Table.6 Path tracking performance

	Average position deviation (m)	Max position deviation(m)	Average computing time(ms)
MPC	0.32	0.81	12.31
LBMPC	0.01	0.01	12.47

For a more detailed analysis, range of the ROV's speed and the control input are shown in Fig.10 and Fig.11. The mean squared error (MSE) values are used to evaluate the smoothness of the control input is recorded in Table.7. It can be seen that the F_Y's MSE of LBMPC is reduced about 60% compared with that of MPC. The T_N's MSE of LBMPC is reduced about 74% compared with that of MPC. As shown in Fig.11, the peak values of both F_Y and T_N have also Significantly decreased. Correspondingly, the sway velocity and the yaw rate of the LBMPC ranges more smoothly. Therefore, the proposed learning-based path tracking control strategy can enhance the smoothness of the control input.

However, the F_X's MSE is decreased significantly. Nevertheless, the ROV's surge speed of the LBMPC is also closer to the expected value that of the MPC. Note that the ROV need maintain a constant surge speed to drag the MRT to track the path. As the surge speed of the ROV is close to the desired surge speed, the control input F_x mainly addresses the external disturbance under the random unknown process noise to track the constant surge speed. This explains why the F_X's MSE has not decreased significantly. To illustrate more intuitively, another three types of control input trajectory are given: one without external disturbances and noise (NoOcdRan), one with only external disturbances (Ocd), and one with both external disturbances and noise present (OcdRan).

From Fig.12, it can be observed that the control input in the "NoOcdRan" scenario is relatively smooth. In contrast to the control input in the "NoOcdRan" case, the input under the "Ocd" scenario exhibits a sinusoidal trend. This is noteworthy because the ocean current disturbance mainly consists of a sinusoidal signal, and the control input reacts accordingly with a similar pattern. Comparing the control inputs from the "OcdRan" and "Ocd" cases, a random chattering trend is evident. This behavior indicates that the control input is also addressing external disturbances and process noise.

Table.7 MSE of the control input

	MPC	LBMPC
F_X	13.78	13.33
F_Y	25.49	62.82
T_N	8.21	31.72

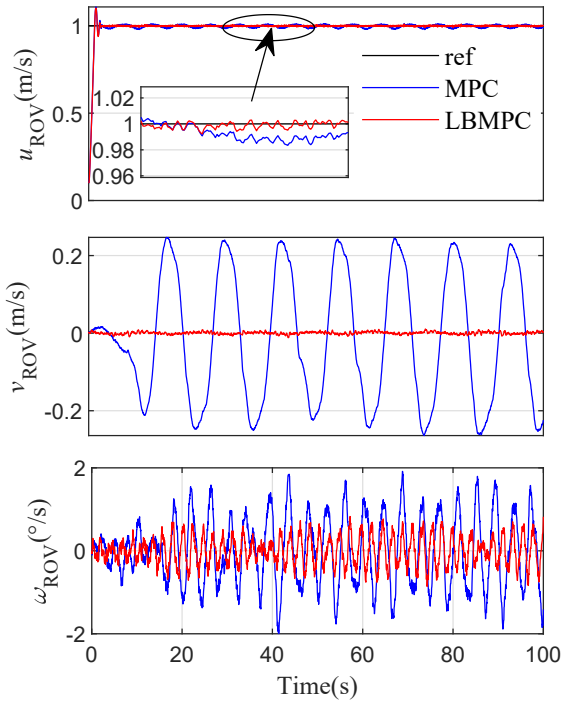

Fig. 10 ROV speed

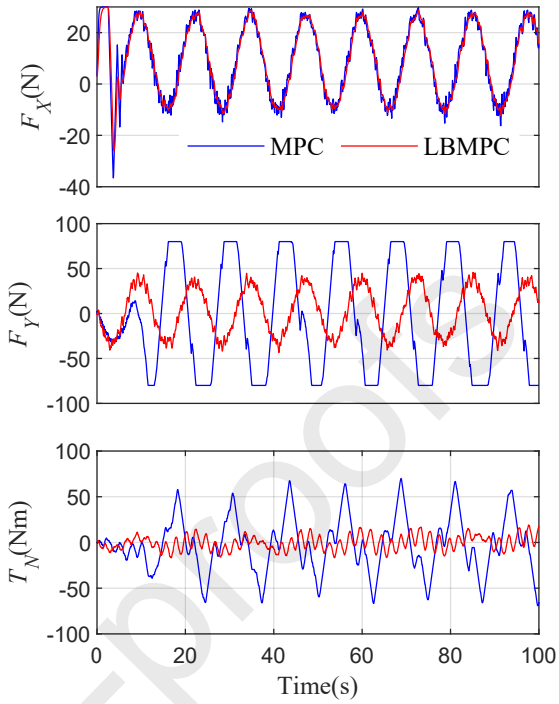

Fig. 11 Control input

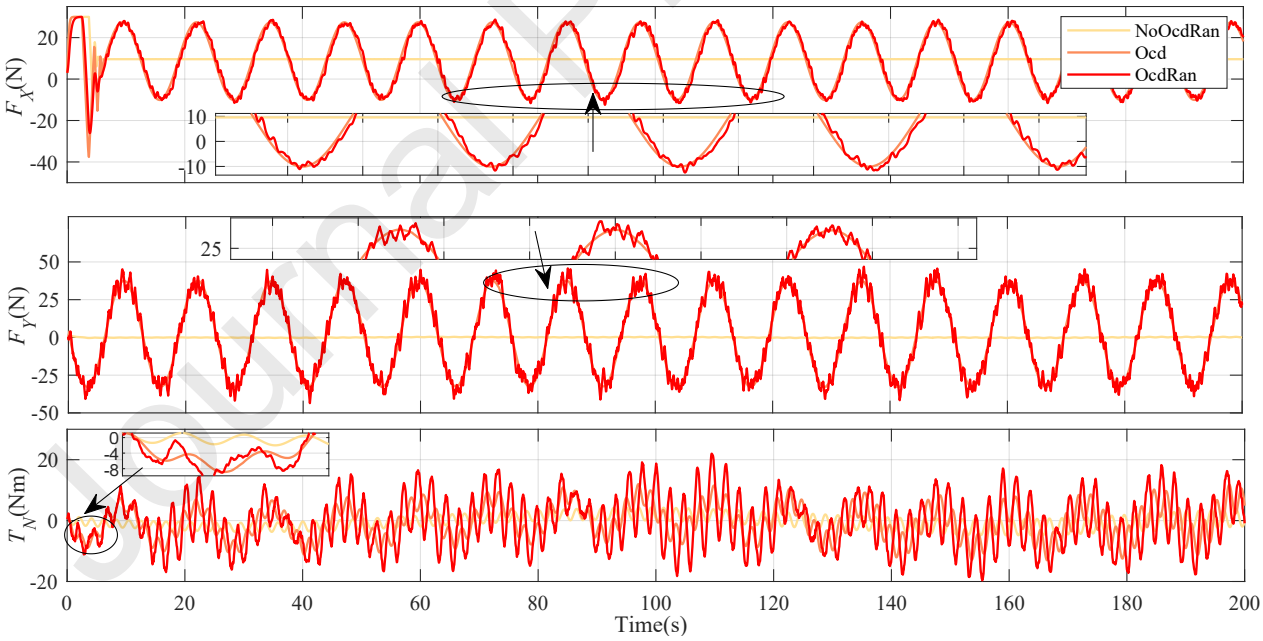

Fig. 12 Control inputs under different simulation cases

6. Conclusion

In this paper, we propose a novel method and equipment for deep-sea mining, focusing on the path tracking control of the ROVDMV in the horizontal plane. Inspired by a bicycle model, we establish a dynamic model of the ROVDMV

for further simulation analysis. Utilizing the KI prediction function and the LACKI rule, we employ a new nonparametric learning method to capture the nonlinear dynamics affected by external current disturbances. This nonparametric learning (NPL) method can filter out the process noise to some extent, yielding accurate estimated values. However, since the sample points are derived from past moments, the estimated values exhibit a time delay of one sampling period. Despite this delay, the NPL estimated error remains small and bounded. A learning-based path tracking control strategy is proposed for the ROVDMV, which consists of a Variable Speed Model Predictive Control (VSMPC) controller and a Learning-Based Model Predictive Control (LBMPC) controller. The VSMPC controller is responsible for calculating the virtual speed control law, while the LBMPC controller tracks this virtual speed control law, utilizing learned nonlinear dynamics as the state transition model in its optimal control problem. Numerical simulations demonstrate that the strategy achieves minimal path tracking deviation and smooth control inputs. Furthermore, the inherent robustness of Model Predictive Control (MPC) effectively addresses the estimated errors in Nonlinear Programming (NPL), confirming the efficacy of the proposed learning-based path tracking control strategy.

However, as demonstrated in Section 5.2, the estimated error of the NPL is significantly lower than the upper bound established in **Theorem 2.** This indicates that the upper bound is overly conservative. Consequently, the theoretical analysis of the upper bound of the NPL estimated error can be refined, and we intend to pursue this research in the future.

Author Declarations

The authors declare that they have no conflicts of interest to disclose.

Acknowledgements

This research work was supported by the National Natural Science Foundation of China (Grant No. 52071138, 12272128). I would like to express my gratitude to Associate Professor Yougang Bian of Hunan University for his valuable discussions on the conceptualization and methodology of the proposed learning-based path tracking control strategy. Additionally, I extend my thanks to PhD student Kaikai Zheng of the Beijing Institute of Technology for his insights on the convergence analysis of the proposed NLP method.

Appendix

A Proof of Lemma 3

Proof: Write the ball around $s(k)$ as: $\mathcal{B}_{\varepsilon^*}(s(k)) = \{s(r) \in \mathbb{S} | \|s(k) - s(r)\|_\infty \leq \varepsilon^*\}, k,r \in \mathbb{K}_{k-N_{\bar{\mathcal{I}}}+1:k}$.

For any sampled data $\tilde{s} \in \mathcal{B}_{\varepsilon^*}(s(k))$, it satisfies:

$$\begin{aligned} u_j\left(\tilde{f}_j(r),\tilde{s},L_n,N_{\mathcal{L}}\right) &= \min_{\tilde{s} \in \mathcal{B}_{\varepsilon^*}(s(k))} \tilde{f}_j(\tilde{s}) + L_n\|\tilde{s} - s(r)\|_\infty + \bar{e} \\ &\leq \tilde{f}_j(\tilde{s}) + L_n\|\tilde{s} - s(r)\|_\infty + \bar{e}, r \in \mathbb{K}_{k-N_{\bar{\mathcal{I}}}+1:k} (69.) \end{aligned}$$

$$\begin{aligned} \ell_j\left(\tilde{f}_j(r),\tilde{s},L_n,N_{\mathcal{L}}\right) &= \max_{\tilde{s} \in \mathcal{B}_{\varepsilon^*}(s(k))} \tilde{f}_j(\tilde{s}) - L_n\|\tilde{s} - s(r)\|_\infty - \bar{e} \\ &\geq \tilde{f}_j(\tilde{s}) - L_n\|\tilde{s} - s(r)\|_\infty - \bar{e}, r \in \mathbb{K}_{k-N_{\bar{\mathcal{I}}}+1:k} (70.) \end{aligned}$$

Then, the KI estimated error satisfies:

$$\mathfrak{b}_j\left(\tilde{f}_j(r),s(k),L_n,D_{N_{\overline{\mathcal{L}}}}\right)=\frac{1}{2}\left(\mathcal{u}\left(\tilde{f}_j(s(r)),s(k),L_n,N_{\overline{\mathcal{L}}}\right)-\ell\left(\tilde{f}_j(s(r)),s(k),L_n,N_{\overline{\mathcal{L}}}\right)\right)$$
$$\leq L_n\|\tilde{s}-s(r)\|_\infty+\overline{e}\leq L_n\varepsilon^*+\overline{e}=\frac{L_n}{L^*}p+\overline{e}\leq p+\overline{e}(71.)$$

Then, the proof is completed. □

B Proof of Lemma 4

Proof: firstly, it can be easily found that:

$$\mathcal{u}_j\left(\tilde{f}_j(r),s(k),L_n,N_{\overline{\mathcal{L}}}\right)\leq\tilde{f}_j(s(\xi))+L_n\|s(\xi)-s(r)\|_\infty+\overline{e},r\in\mathbb{K}_{k-N_{\mathcal{L}}+1:k}(72.)$$

if an inequation $\ell_j\left(\tilde{f}_j(r),s(k),L_n,N_{\mathcal{L}}\right)\leq\mathcal{u}_j\left(\tilde{f}_j(r),s(k),L_n,N_{\mathcal{L}}\right)$ can be proved, the proof is completed. The inequation will be proved by contradiction. Give an assumption:

$$\ell_j\left(\tilde{f}_j(r),s(k),L_n,N_{\mathcal{L}}\right)>\mathcal{u}_j\left(\tilde{f}_j(r),s(k),L_n,N_{\mathcal{L}}\right)(73.)$$

For convenience, two sample points are given:

$$r_0=\underset{r\in\mathbb{K}_{k-N_{\overline{\mathcal{L}}}+\mathbf{1}:k}}{\operatorname{argmin}}\tilde{f}_j(s(r))+L_n\|s(k)-s(r)\|_\infty+\overline{e}(74.)$$

$$r_1=\underset{r\in\mathbb{K}_{k-N_{\overline{\mathcal{L}}}+\mathbf{1}:k}}{\operatorname{argmax}}\tilde{f}_j(s(r))-L_n\|s(k)-s(r)\|_\infty-\overline{e}(75.)$$

Then, inequality (73) is equivalent to:

$$\tilde{f}_j(s(r_1))-L_n\|s(r_1)-s(r)\|_\infty-\overline{e}>\tilde{f}_j(s(r_0))+L_n\|s(k)-s(r_0)\|_\infty+\overline{e}(76.)$$

which can be rewritten as:

$$\tilde{f}_j(s(r_1))-\tilde{f}_j(s(r_0))-2\overline{e}>L_n(\|s(r_1)-s(r)\|_\infty+\|s(k)-s(r_0)\|_\infty)>L_n\|s(r_1)-s(r_0)\|_\infty(77.)$$

It can be obtained that:

$$\frac{\tilde{f}_j(s(r_1))-\tilde{f}_j(s(r_0))-2\overline{e}}{\|s(r_1)-s(r_0)\|_\infty}>L_n(78.)$$

Based on Remark 2, it holds:

$$\frac{\tilde{f}_j(s(r_1))-\tilde{f}_j(s(r_0))-2\overline{e}}{\|s(r_1)-s(r_0)\|_\infty}\leq L_n(79.)$$

, and it can be seen that inequality (78) and (79) are contradictory, which means the assumption fails. Then, the proof is completed.

For $\lambda > 2\bar{e}$, inequality (78) becomes:

$$\frac{\tilde{f}_j(s(r_1)) - \tilde{f}_j(s(r_0)) - \lambda}{\|s(r_1) - s(r_0)\|_\infty} > L_n - \frac{\lambda - 2\bar{e}}{\|s(r_1) - s(r_0)\|_\infty} \quad (80.)$$

and inequality (79) becomes:

$$\frac{\tilde{f}_j(s(r_1)) - \tilde{f}_j(s(r_0)) - \lambda}{\|s(r_1) - s(r_0)\|_\infty} \leq L_n - \frac{\lambda - 2\bar{e}}{\|s(r_1) - s(r_0)\|_\infty} \quad (81.)$$

Thus, for $\lambda > 2\bar{e}$, Lemma 4 still holds. □

References

Al-Milli, S., Seneviratne, L. D., & Althoefer, K. (2010). Track–terrain modelling and traversability prediction for tracked vehicles on soft terrain. *Journal of Terramechanics, 47*(3), 151-160. doi:https://doi.org/10.1016/j.jterra.2010.02.001

Calliess, J.-P., Roberts, S. J., Rasmussen, C. E., & Maciejowski, J. (2020). Lazily Adapted Constant Kinky Inference for nonparametric regression and model-reference adaptive control. *Automatica, 122*, 109216. doi:https://doi.org/10.1016/j.automatica.2020.109216

Chen, Q., Yang, J., Mao, J., Liang, Z., Lu, C., & Sun, P. (2023). A path following controller for deep-sea mining vehicles considering slip control and random resistance based on improved deep deterministic policy gradient. *Ocean Engineering, 278*, 114069. doi:https://doi.org/10.1016/j.oceaneng.2023.114069

Dai, Y., Su, Q., & Zhang, Y. (2020). A new dynamic model and trajectory tracking control strategy for deep ocean mining vehicle. *Ocean Engineering, 216*, 108162. doi:https://doi.org/10.1016/j.oceaneng.2020.108162

Dai, Y., Xue, C., Su, Q., & Huang, X. (2021). Numerical analysis on hydrodynamic characteristics of a deep-sea mining vehicle under three typical motions. *Ocean Engineering, 235*, 109446. doi:https://doi.org/10.1016/j.oceaneng.2021.109446

Kar, M. K. (1987). Prediction of track forces in skid-steering of military tracked vehicles. *Journal of Terramechanics, 24*(1), 75-84. doi:https://doi.org/10.1016/0022-4898(87)90060-7

Li, W., Cai, J., Duan, C., Chen, S., Ding, P., Lin, J., & Cui, D. (2023). Learning and ensemble based MPC with differential dynamic programming for nuclear power autonomous control. *Expert Systems with Applications, 215*, 119416. doi:https://doi.org/10.1016/j.eswa.2022.119416

Li, Y., He, D., Ma, F., Liu, P., & Liu, Y. (2023). MPC-based trajectory tracking control of unmanned underwater tracked bulldozer considering track slipping and motion smoothing. *Ocean Engineering, 279*, 114449. doi:https://doi.org/10.1016/j.oceaneng.2023.114449

Liu, C. H., & Wong, J. Y. (1996). Numerical simulations of tire-soil interaction based on critical state soil mechanics. *Journal of Terramechanics, 33*(5), 209-221. doi:https://doi.org/10.1016/S0022-4898(97)00005-0

Long, C., Hu, M., Qin, X., & Bian, Y. (2022). Hierarchical trajectory tracking control for ROVs subject to disturbances

and parametric uncertainties. *Ocean Engineering, 266*, 112733. doi:https://doi.org/10.1016/j.oceaneng.2022.112733

Qin, Z., Chen, L., Fan, J., Xu, B., Hu, M., & Chen, X. (2021). An Improved Real-Time Slip Model Identification Method for Autonomous Tracked Vehicles Using Forward Trajectory Prediction Compensation. *IEEE Transactions on Instrumentation and Measurement, 70*, 1-12. doi:10.1109/TIM.2020.3048801

Ruslan, N. A. I., Amer, N. H., Hudha, K., Kadir, Z. A., Ishak, S. A. F. M., & Dardin, S. M. F. S. (2023). Modelling and control strategies in path tracking control for autonomous tracked vehicles: A review of state of the art and challenges. *Journal of Terramechanics, 105*, 67-79. doi:https://doi.org/10.1016/j.jterra.2022.10.003

Sha, F., Xi, M., Chen, X., Liu, X., Niu, H., & Zuo, Y. (2023). A recent review on multi-physics coupling between deep-sea mining equipment and marine sediment. *Ocean Engineering, 276*, 114229. doi:https://doi.org/10.1016/j.oceaneng.2023.114229

Sun, C., Li, Q., Li, B., & Li, L. (2022). A Successive Linearization in Feasible Set Algorithm for Vehicle Motion Planning in Unstructured and Low-Speed Scenarios. *IEEE Transactions on Intelligent Transportation Systems, 23*(4), 3724-3736. doi:10.1109/TITS.2020.3041075

Tijjani, A. S., Chemori, A., & Creuze, V. (2021). Robust Adaptive Tracking Control of Underwater Vehicles: Design, Stability Analysis, and Experiments. *IEEE/ASME Transactions on Mechatronics, 26*(2), 897-907. doi:10.1109/TMECH.2020.3012502

Wang, L., Chen, X., Wang, L., Li, Z., & Yang, W. (2023). Mechanical properties and soil failure process of interface between grouser of tracked mining vehicle and deep-sea sediment. *Ocean Engineering, 285*, 115336. doi:https://doi.org/10.1016/j.oceaneng.2023.115336

Wu, H., Liu, W., Jiang, M., Li, C., Li, J., Chen, B., . . . Chen, Y. (2023). Numerical Simulation Investigation on Parameter Optimization of Deep-Sea Mining Vehicles. *Chinese Journal of Mechanical Engineering, 36*(1), 54. doi:10.1186/s10033-023-00871-z

Wu, P. C., Wen, H., Chen, T., & Jin, D. P. (2017). Model predictive control of rigid spacecraft with two variable speed control moment gyroscopes. *Applied Mathematics and Mechanics-English Edition, 38*(11), 1551-1564. doi:10.1007/s10483-017-2278-9

Xu, Z., Liu, Y., Yang, G., Xia, J., Dou, Z., Meng, Q., & Xu, X. (2022). Research on contact model of track-soft sediment and traction performance of four-tracked seabed mining vehicle. *Ocean Engineering, 259*, 111902. doi:https://doi.org/10.1016/j.oceaneng.2022.111902

Yamada, M., Yamauchi, G., & Hashimoto, T. (2021). Fundamental study on underwater trafficability for tracked vehicle. *Journal of Terramechanics, 98*, 42-49. doi:https://doi.org/10.1016/j.jterra.2021.07.001

Yeu, T.-K., Yoon, S.-M., Park, S.-J., Sup, H., Kim, H.-W., Lee, C.-H., . . . Ieee. (2012). *Study on Path Tracking Approach for Underwater Mining Robot.* Paper presented at the OCEANS MTS/IEEE Conference, Yeosu, SOUTH KOREA.

Yu, S., Maier, C., Chen, H., & Allgöwer, F. (2013). Tube MPC scheme based on robust control invariant set with

application to Lipschitz nonlinear systems. *Systems & Control Letters, 62*(2), 194-200. doi:https://doi.org/10.1016/j.sysconle.2012.11.004

Zeng, R., Kang, Y., Yang, J., Qin, B., Chen, S., & Cao, D. (2021). An Integrated Terrain Identification Framework for Mobile Robots: System Development, Analysis, and Verification. *IEEE/ASME Transactions on Mechatronics, 26*(3), 1581-1590. doi:10.1109/TMECH.2020.3024832

Zheng, K., Shi, D., Shi, Y., & Wang, J. (2023). Nonparameteric Event-Triggered Learning With Applications to Adaptive Model Predictive Control. *IEEE Transactions on Automatic Control, 68*(6), 3469-3484. doi:10.1109/TAC.2022.3191760

Yuheng Chen: Writing – original draft, Software, Resources, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Haicheng Zhang:** Writing – review & editing, Supervision, Software, Project administration, Methodology, Investigation, Funding acquisition, Data curation, Conceptualization. **Weisheng Zou:** Validation, Supervision, Methodology, Methodology, Conceptualization. **Haihua Zhang:** Supervision, Project administration, Investigation. **Bin Zhou:** Supervision, Project administration, Investigation. **Daoling Xu:** Supervision, Project administration, Investigation, Conceptualization.

Conflict of Interest:

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

ORCID: https://orcid.org/0000-0003-2452-3167

PhD Student Yuheng Chen

College of Mechanical and Vehicle Engineering Hunan University Changsha, Hunan, 410082, China

E-mail: chenyuheng@hnu.edu.cn

ORCID：https://orcid.org/0000-0002-5357-6332

Associate professor Haicheng Zhang

College of Mechanical and Vehicle Engineering Hunan University Changsha, Hunan, 410082, China

E-mail: zhanghc@hnu.edu.cn