A New Compensation Method of Dynamic Lever-Arm Effect Error for Hypersonic Vehicles

BO YANG[1], HE YU[1], CHAOFAN LIU[1], XIANG WEI[1], ZICHEN FAN[2], AND JUN MIAO[3]
[1]School of Astronautics, Beihang University, Beijing 100191, China
[2]Beijing Institute of Control and Electronic Technology, Beijing 100038, China
[3]Qian Xuesen Laboratory of Space Technology, Beijing 100094, China

Corresponding author: He Yu (13261059095@163.com)

ABSTRACT For the Inertial Navigation System (INS) in hypersonic flight, when the centroid of the Inertial Measurement Unit (IMU) does not coincide with that of hypersonic vehicle, a lever-arm effect error will be generated, which will further expand in the dynamic complex environment. In order to improve the navigation accuracy, the lever-arm effect error needs to be compensated. Although much research has been done on the compensation of lever-arm effect, the existing methods are almost based on the fixed lever-arm length, which can not adapt to the changes of lever-arm in a dynamic and complex environment. In order to obtain the higher accuracy of navigation in the complex dynamic environment for a hypersonic vehicle, a more perfect dynamic lever-arm model is established, and a new compensation method for dynamic lever-arm effect based on reinforcement learning (RL-DLAC) is proposed. The method adopts the framework of reinforcement learning and obtains the optimal parameter estimation of the dynamic lever-arm model from the continuous action space according to the deep deterministic policy gradient (DDPG). The simulation results show that the proposed method can estimate the parameters of the dynamic lever-arm model with an error of only 0.31%, compensate the lever-arm effect error in complex dynamic environment and improve the navigation accuracy.

INDEX TERMS Hypersonic vehicle, inertial navigation system, dynamic lever-arm effect, reinforcement learning, complex dynamic environment.

I. INTRODUCTION

A hypersonic vehicle, as a cutting-edge weapon in modern combat environment, has attracted the attention of many countries because of its hypersonic and highly mobile characteristics [1], [2]. As one of the key technologies of hypersonic vehicles, precise navigation is important to ensure the feasibility and efficiency of them [3]. The Inertial Navigation System (INS) is generally selected as the main navigation system of hypersonic vehicle due to the advantage of autonomous navigation [4]. The error of its Inertial Measurement Unit (IMU) plays an important role in INS accuracy. After the error of the IMU is calibrated and compensated, the accuracy of navigation can be effectively improved [5].

The IMU is generally difficult to meet the requirements of coincidence with the centroid of a hypersonic vehicle during installation. Therefore, the lever-arm length between the IMU and the vehicle will produce the lever-arm effect error [6]. In the high speed and high maneuverability environment, due to the influence of elastic deformation and centroid change [7], the influence of lever-arm effect error generated by the IMU will be further highlighted. Therefore, it is of great significance to study the compensation of dynamic lever-arm effect error for hypersonic vehicles.

At present, many scholars at home and abroad have studied the error compensation of lever-arm effect, mainly including mechanical compensation method and digital filtering method. The mechanical compensation method generally relies on the corresponding angular motion scheme and the measurement system to estimate the lever-arm [8]–[12]. As a more practical on-line compensation method, digital filter methodology generally adopts the path of designing the relationship between lever-arm parameters and the navigation system model to calibrate the lever-arm through

The associate editor coordinating the review of this manuscript and approving it for publication was Erwu Liu.

filtering [13]–[16]. However, most of the current compensation methods for lever-arm effect focus on the premise of a constant lever-arm [17]–[20], which is difficult to adapt to the complex and dynamic flight environment. The current research on dynamic lever-arm compensation is generally based on the existing fixed lever-arm compensation methods, considering the dynamic change of the lever-arm caused by random flexible deformation, and using an Adaptive Kalman Filter (AKF) to compensate for the error from lever-arm effect [21]–[23]. However, the research does not consider the influence of other factors affecting the change of the lever-arm, and there is a lack of more detailed analysis of the model parameters, so its applicability needs to be improved.

Reinforcement learning (RL) has been applied in navigation systems because of its unique learning ability to interact with the environment [24]–[27]. For the actual complex dynamic environment, the change of lever-arm can also be transformed into a problem of interaction with the flight environment. Therefore, aiming at the lever-arm effect compensation for hypersonic vehicles in a complex dynamic environment, a new method of dynamic lever-arm effect based on reinforcement learning (RL-DLAC) is proposed. The method first establishes a more perfect elastic deformation dynamic lever-arm model under high maneuverability conditions, and then uses the environmental interaction mechanism of RL to carry out the adaptive estimation of the dynamic lever-arm, so as to effectively improve the navigation accuracy and adapt to the complex changes of the lever-arm.

The contribution of this study is that the change of dynamic lever-arm model is first analyzed on the basis of breaking through the traditional fixed lever-arm compensation, and a new method is developed for the compensation of dynamic lever-arm effect. The proposed method transforms the influence of different factors on the change of lever-arm into the change of dynamic lever-arm model parameters, which is an unknown term related to the environment. Thus, it is more suitable for compensating dynamic lever arm effect in complex dynamic environment of hypersonic vehicles.

The remainder of this paper is organized as follows. In Section 2, dynamic lever-arm error model of hypersonic vehicle is presented. The design for RL-DLAC is described in Section 3. The performance of RL-DLAC in calibration and compensation of the dynamic lever-arm for hypersonic vehicles is verified in Section 4. Finally, the conclusions are drawn in Section 5.

II. DYNAMIC LEVER-ARM ERROR MODEL
This section shows the limitations of the traditional method through the traditional lever-arm effect principle, and establishes a more perfect dynamic lever-arm error model based on the elastic deformation mechanism of hypersonic vehicle.

A. TRADITIONAL LEVER-ARM ERROR MODEL
Ideally, the IMU should measure the specific force information of the vehicle centroid. However, in the actual system, due to the limitation of the vehicle structure, the centroid of
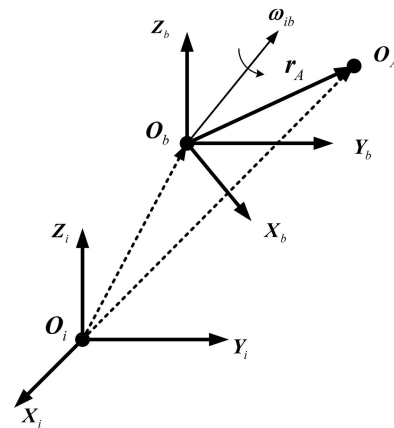

FIGURE 1. Schematic diagram of the lever-arm effect.

the accelerometer on the IMU does not necessarily coincide with the centroid of the vehicle. Therefore, the output of the accelerometer contains the lever-arm effect error when the vehicle is in angular motion (since the three accelerometers on the IMU are close, the inner lever-arm error between different accelerometers will not be considered here). The schematic diagram of that is shown in Fig. 1.

O_i and O_b are the centers of inertial coordinate system and body coordinate system respectively, as shown in Fig. 1. O_A is the center of mass of IMU. r_A is the vector of lever-arm length. ω_{ib} is the rotational angular velocity of the body coordinate system relative to the inertial coordinate system [22]. The accelerometer error caused by the lever-arm effect was written as ∇_{lever}, which is as follows:

$$\nabla_{lever} = \ddot{r}_A + 2\omega_{ib} \times \dot{r}_A + \dot{\omega}_{ib} \times r_A + \omega_{ib} \times (\omega_{ib} \times r_A)$$

$$(1)$$

In the traditional compensation method for the lever-arm effect, it is generally assumed that the relative position of the vehicle and IMU does not change, which is equivalent to $\ddot{r}_A = \dot{r}_A = 0$. And then the lever-arm length is extended to the state quantity in the inertial navigation error equation and estimated by Kalman filter. This method cannot guarantee the navigation accuracy in case of a dynamic lever-arm. Fig. 2 shows that a change in the lever-arm will lead to an increase in navigation error.

For the research on the change of lever-arm, there is only the method that the change in the lever-arm is regarded as random flexible deformation. Due to the change in the lever-arm, that means $\ddot{r}_A \neq 0, \dot{r}_A \neq 0$. However, the current model of flexible deformation is generally regarded as a second-order Markov process. The random deformation obviously does not have a clear physical meaning and is inaccurate. Therefore, it is necessary to improve the change model of the lever-arm. Besides, due to the change of environment, the parameters of an actual dynamic lever-arm model are also time-varying. At present, there is no compensation method for the change in the lever-arm model. Actually, the lever-arm model will
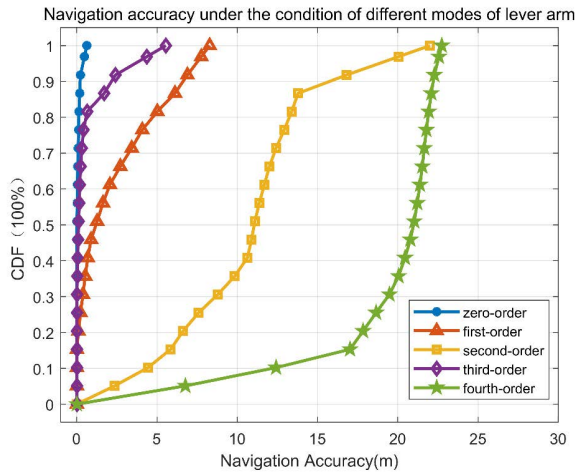
FIGURE 2. The influence of a dynamic lever-arm on navigation accuracy.

FIGURE 3. Schematic diagram of elastic deformation of the vehicle.

change due to the change of natural frequency and vibration mode in the actual flight environment of hypersonic vehicles. Therefore, it is necessary to find a compensation method for the change in the dynamic lever-arm model.

B. IMPROVED DYNAMIC LEVER-ARM ERROR MODEL

The structure of hypersonic vehicles generally has a certain stiffness, but structural deformation and bending vibration will occur under the high dynamic maneuver environment and the action of external forces, such as aerodynamic force. Therefore, the lever-arm length between the IMU and vehicle centroid can no longer be filtered and estimated as a constant. Similarly, it cannot be approximated by a simple second-order Markov process, because the actual lever-arm deformation in the vehicle structure is not a simple random process, and there is actual external force that needs to be considered. The derivation of an improved dynamic lever-arm model will be described in more detail.

The hypersonic vehicle is approximated as a slender beam structure, and the cross-sectional area along the longitudinal direction changes little. Therefore, the simplified beam can be used to replace the vehicle structure. Its transverse vibration in space can be regarded as the movement of a non-uniform elastic beam with free ends. The elastic displacement, flexural stiffness and mass inertia of the beam along the length are determined by the structural parameters of the vehicle. The elastic deformation diagram of the vehicle is shown in Fig. 3. The elastic deflection displacement equation can be written as

$$y(x, t) = \sum_{i=1}^{n} q_i(t) W_i(x) \quad i = 1, 2, \cdots, n \tag{2}$$

where $W_i(x)$ represents the relative lateral displacement relationship of each point on the longitudinal axis of the vehicle, which is called the *i*-th natural mode function. $q_i(t)$ is the generalized coordinate of elastic vibration, which can determine
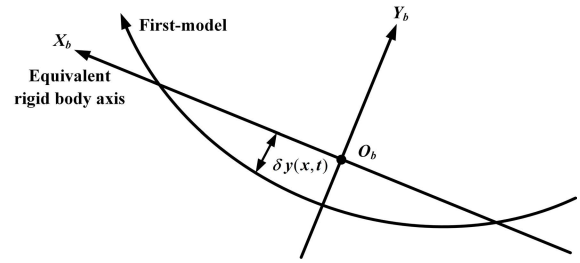
the amplitude of elastic vibration after the elastic vibration mode function is determined.

The generalized coordinate $q_i(t)$ is varying with time, which satisfies

$$\ddot{q}_i + 2\zeta_i \omega_i \dot{q}_i + \omega_i^2 q_i = \frac{Q_i}{M_i} \tag{3}$$

where ω_i is the natural frequency of the *i*-th mode. ζ_i is the damping coefficient of the *i*-th mode. Q_i is the generalized force of the *i*-th mode. M_i is the generalized mass of the *i*-th mode. Q_i and M_i satisfies

$$\begin{cases} Q_i = \int_0^l \sum f_{yb} W_i(x) \mathrm{d}x \\ M_i = \int_0^l m(x) W_i^2(x) \mathrm{d}x \end{cases} \tag{4}$$

where $f_{yb}(x)$ is the normal projection of the external force acting on the vehicle. $m(x)$ is the mass distribution along the longitudinal axis of the vehicle. l is the equivalent length of the beam. The external force of the vehicle is simplified to the action of gravity, thrust and aerodynamic force as given by

$$\sum f_{yb} = F_{gyb} + F_{kyb} + F_{ayb} \tag{5}$$

where F_{gyb}, F_{kyb} and F_{ayb} are the projection of gravity, thrust and aerodynamic force on the $O_b Y_b$-axis respectively.

According to the elastic deformation and random flexible deformation of the lever-arm of the vehicle, the generalized coordinates of the elastic vibration satisfy the relationship as follows:

$$\begin{aligned} \ddot{q}_i &= -2\zeta_i \omega_i \dot{q}_i - \omega_i^2 q_i + \frac{Q_i}{M_i} + w_{qi} \\ &= -2\zeta_i \omega_i \dot{q}_i - \omega_i^2 q_i + u_{qi} + w_{qi} \end{aligned} \tag{6}$$

where u_{qi} is the control quantity of elastic deformation. w_{qi} is the noise of elastic deformation.

Taking the third-order vibration mode as an example, the relationship of elastic deformation length satisfies

$$y(x, t) = q_1(t) W_1(x) + q_2(t) W_2(x) + q_3(t) W_3(x) \tag{7}$$

According to the geometric relationship, the relationship of lever-arm length is as follows:

$$\boldsymbol{r}_A = \boldsymbol{r}_{A0} + \boldsymbol{Y} \begin{bmatrix} q_1 \\ q_2 \\ q_3 \end{bmatrix}$$

$$= r_{A0} + Yq, \dot{r}_A = Y \begin{bmatrix} \dot{q}_1 \\ \dot{q}_2 \\ \dot{q}_3 \end{bmatrix} = Y\dot{q} \qquad (8)$$

$$\ddot{r}_A = Y \begin{bmatrix} \ddot{q}_1 \\ \ddot{q}_2 \\ \ddot{q}_3 \end{bmatrix}$$

$$= Y \left(K_1 \begin{bmatrix} \dot{q}_1 \\ \dot{q}_2 \\ \dot{q}_3 \end{bmatrix} + K_2 \begin{bmatrix} q_1 \\ q_2 \\ q_3 \end{bmatrix} + \begin{bmatrix} u_{q1} \\ u_{q2} \\ u_{q3} \end{bmatrix} + \begin{bmatrix} w_{q1} \\ w_{q2} \\ w_{q3} \end{bmatrix} \right)$$

$$= Y \left(K_1 \dot{q} + K_2 q + u_q + w_q \right) \qquad (9)$$

$$K_1 = diag \left(\begin{bmatrix} -2\zeta_1\omega_1 & -2\zeta_2\omega_2 & -2\zeta_3\omega_3 \end{bmatrix} \right) \qquad (10)$$

$$K_2 = diag \left(\begin{bmatrix} -\omega_1^2 & -\omega_2^2 & -\omega_3^2 \end{bmatrix} \right) \qquad (11)$$

$$Y = \begin{bmatrix} 0 & 0 & 0 \\ W_1(x) & W_2(x) & W_3(x) \\ 0 & 0 & 0 \end{bmatrix} \qquad (12)$$

According to the above equations, the error of lever-arm effect can be given by

$$\begin{aligned} \nabla_{lever} &= \ddot{r}_A + 2\omega_{ib} \times \dot{r}_A + \dot{\omega}_{ib} \times r_A + \omega_{ib} \times (\omega_{ib} \times r_A) \\ &= Y \left(K_1 \dot{q} + K_2 q + u_q + w_q \right) \\ &\quad + 2\omega_{ib} \times Y\dot{q} + \dot{\omega}_{ib} \times (r_{A0} + Yq) \\ &\quad + \omega_{ib} \times (\omega_{ib} \times (r_{A0} + Yq)) \\ &= \left(YK_1 + 2\omega_{ib}^\times Y \right) \dot{q} + \left(YK_2 + \dot{\omega}_{ib}^\times Y + \omega_{ib}^\times \omega_{ib}^\times Y \right) q \\ &\quad + Yu_q \\ &\quad + \left(\dot{\omega}_{ib}^\times + \omega_{ib}^\times \omega_{ib}^\times \right) r_{A0} + Yw_q \\ &= \left(YK_1 + 2\omega_{ib}^\times Y \right) \dot{q} + \left(YK_2 + \dot{\omega}_{ib}^\times Y + \omega_{ib}^\times \omega_{ib}^\times Y \right) q \\ &\quad + u_l + w_l \end{aligned} \qquad (13)$$

where r_{A0} is the initial lever-arm length (Initial calibration value of the default value which does not affect the compensation accuracy). q is the vector of generalized coordinate. u_l and w_l satisfy $u_l = Yu_q + \left(\dot{\omega}_{ib}^{b\times} + \omega_{ib}^{b\times} \omega_{ib}^{b\times} \right) r_{A0}$ and $w_l = Yw_q$, respectively.

In view of the extensive research on air-breathing hypersonic vehicles, one of those is taken as the simulation object [28]. The structure and system distribution diagram of the hypersonic vehicle is shown in Fig. 4.

The front and rear bodies of the hypersonic vehicle are equipped with subsystems, and liquid oxygen and hydrogen storage tanks are installed in the middle and both sides.

The natural frequency and mode shape are related to the bending stiffness, mass and beam structure mode shape function of hypersonic vehicle material. In high-speed flight, the mass of hypersonic vehicle decreases with the consumption of fuel. Besides, the increase of surface temperature will reduce the bending stiffness of material under the influence of aerodynamic heating on the surface of hypersonic vehicle. Therefore, the natural frequency of the vibration mode of hypersonic vehicle will change. The compensation of dynamic lever-arm model caused by the change of natural frequency will be described in detail in Section 3.

III. DYNAMIC LEVER-ARM EFFECT COMPENSATION METHOD BASED ON REINFORCEMENT LEARNING

This section introduces the design of a dynamic lever-arm effect compensation method. Reinforcement learning is used to learn the parameters of the dynamic lever-arm model, and the dynamic lever-arm is estimated by filtering to compensate for the lever-arm effect error.

A. PRINCIPLE OF DYNAMIC LEVER-ARM EFFECT COMPENSATION METHOD

According to the improved dynamic lever-arm error model in Section 2, the lever-arm model for hypersonic vehicles is variable and unknown in the actual flight process. Therefore, RL-DLAC is proposed to compensate for the error caused by the dynamic lever-arm effect, so as to obtain higher navigation accuracy. The principle of RL-DLAC is shown in Fig. 5.

$\tilde{\omega}_{ib}$ and $\tilde{f}_b$ are the outputs of the gyroscope and acceleration obtained through the IMU, respectively. After drift correction and dynamic lever-arm effect error compensation, ω_{ib} and f_b are obtained for inertial navigation solution. Given the initial state X_0 and the initial state covariance matrix P_0, if there is GNSS measurement input at the current time, the module will update the measurement to obtain the final error state estimation result $\hat{X}_k$, and then the current position estimation result $\hat{r}$ is obtained after inertial navigation correction. At the beginning, the natural frequency parameter ω_i in the lever-arm model is provided by the default value ω_i^0 (initial calibration result or previous estimation result). The later simulation takes the third-order vibration model as an example, and the ω_i^0 is written as $\begin{bmatrix} \omega_1^0 & \omega_2^0 & \omega_3^0 \end{bmatrix}^T$. Considering that the GNSS carrier phase differential positioning results have the characteristics of high accuracy but high environmental requirements, the method will consistently monitor the GNSS positioning status. Besides, the method will calculate the positioning error between the current integrated navigation system result $\hat{r}$ and r_G. If the calculation error is greater than the threshold, it indicates that there is a certain deviation in the natural frequency of the dynamic lever-arm model, resulting in the inaccurate results of the integrated navigation. Then, the method collects the IMU and GNSS data closest to the current time, which is added to the RL module for adaptive learning. Finally, the learned natural frequency is fed back to the previous dynamic lever-arm effect compensation system to continue filtering and correcting.

B. CONSTRUCTION OF REINFORCEMENT LEARNING

The goal of the reinforcement learning module is to learn the natural frequency of the dynamic lever-arm model, which can minimize the position error of the current condition obtained by RL-DLAC. The constituent elements of the reinforcement learning module are shown in Table 1.

The implemented tasks conform to the standard interface of an infinite-horizon discounted Markov decision process (MDP), defined by the tuple $(s, \alpha, P, r, \rho_0)$, where s is an infinite set of states, α is a set of actions, $r : s \times \alpha \to R$
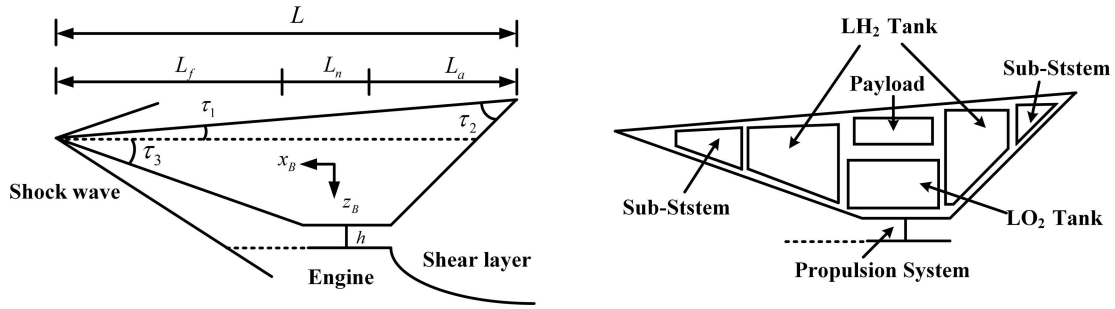
FIGURE 4. Structure and system distribution diagram of the hypersonic vehicle.
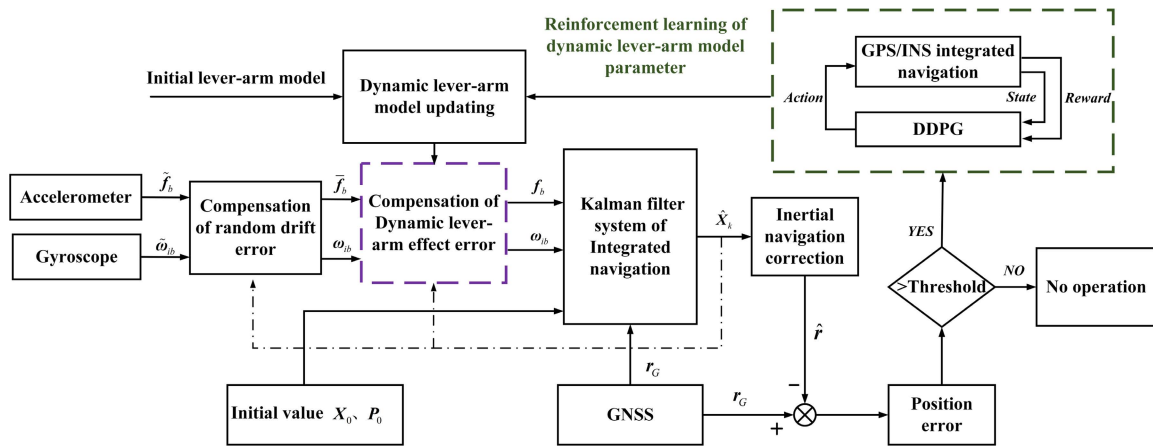

FIGURE 5. Schematic diagram of dynamic lever-arm effect compensation method based on RL.

TABLE 1. Constituent elements of RL.

Elements	Description
State: s	The natural frequency $\begin{bmatrix} \omega_1 & \omega_2 & \omega_3 \end{bmatrix}^T$ of dynamic lever-arm model.
Action: α	The change α_i of natural frequency ω_i in continuous real number space.
Reward: r	The opposite of the location error. The bigger the reward, the better the current navigation result.
Environment	The GNSS/INS integration navigation model. Once the environment receives a new action, it updates the current state and calculates the location error to get a reward. Then it passes the reward to the agent.
Agent	The decision model of RL. We use the Deep Deterministic Policy Gradient (DDPG) algorithm, which adopts a network to fit the strategy function in the aspect of action output.

is the transition probability distribution represented by the agent, $r : s \times \alpha \to R$ is the reward value obtained from the environment, and $\rho_0 : s \to R \geq 0$ is the initial state. These

characteristics allow the RL-DLAC to be formalized as an MDP where operator actions cause state transitions, which affect the ramp-up value of reward. Capturing the dependencies between actions and reactions allows the machine learning methods to be used to find good policies for the ramp-up of observed states, from which we can also obtain the learned optimal state.

1) ENVIRONMENT DEFINITION

The most commonly used and representative navigation system in hypersonic vehicles is the INS/GNSS integrated navigation [29]. Therefore, the INS/GNSS integrated navigation system is taken as the environmental element of RL, which will be used to evaluate the current dynamic lever-arm model.

In order to compensate for the lever-arm effect error, it is necessary to establish the IMU error equation under the influence of the lever-arm effect error (installation error and scale factor error are not considered here), which is given by

$$\begin{cases} \delta \boldsymbol{f}_b = \nabla + \nabla_{lever} \\ \delta \boldsymbol{\omega}_{ib} = \boldsymbol{\varepsilon}_b + \boldsymbol{\varepsilon}_r \end{cases} \quad (14)$$

where ∇ is the accelerometer drift error, $\boldsymbol{\varepsilon}_b$ and $\boldsymbol{\varepsilon}_r$ are the constant drift and first-order Markov process of gyroscope respectively.

The generalized coordinates of the dynamic lever-arm are extended to the state quantity in the integrated navigation system, which is

$$
\begin{aligned}
\boldsymbol{X} = \big[& \phi_E \ \phi_N \ \phi_U \ \delta V_E \ \delta V_N \ \delta V_U \ \delta L \ \delta\lambda \ \delta h \\
& \times \ \varepsilon_{bx} \ \varepsilon_{by} \ \varepsilon_{bz} \ \varepsilon_{rx} \ \varepsilon_{ry} \ \varepsilon_{rz} \ \nabla_x \ \nabla_y \ \nabla_z \\
& \times \ q_1 \ q_2 \ q_3 \ \dot{q}_1 \ \dot{q}_2 \ \dot{q}_3 \big]^T
\end{aligned} \tag{15}
$$

The state equation and measurement equation of the dynamic lever-arm effect error compensation system are

$$
\begin{cases}
\dot{\boldsymbol{X}}(t) = \boldsymbol{F}(t)\boldsymbol{X}(t) + \boldsymbol{B}(t)\boldsymbol{U}(t) + \boldsymbol{G}(t)\boldsymbol{W}(t) \\
\boldsymbol{Z}(t) = \boldsymbol{H}(t)\boldsymbol{X}(t) + \boldsymbol{V}(t)
\end{cases} \tag{16}
$$

where $\boldsymbol{F}(t)$ is the system matrix, $\boldsymbol{G}(t)$ is the system state noise coefficient matrix, $\boldsymbol{W}(t)$ is the system noise matrix, $\boldsymbol{H}(t)$ is the measurement coefficient matrix, $\boldsymbol{V}(t)$ is the measurement noise matrix, and $\boldsymbol{U}(t)$ is the control matrix. The system matrix $\boldsymbol{F}(t)$ is given by

$$
\boldsymbol{F}(t) = \begin{bmatrix}
\boldsymbol{F}_1(t)_{9\times9} & \boldsymbol{F}_2(t)_{9\times9} & \boldsymbol{F}_4(t)_{9\times6} \\
\boldsymbol{0}_{9\times9} & \boldsymbol{F}_3(t)_{9\times9} & \boldsymbol{0}_{9\times6} \\
\boldsymbol{0}_{6\times9} & \boldsymbol{0}_{6\times9} & \boldsymbol{F}_5(t)_{6\times6}
\end{bmatrix} \tag{17}
$$

$\boldsymbol{F}_1(t)_{9\times9}$ is the inertial navigation error matrix, and the other matrices are in (18)–(23), as shown at the bottom of the next page, where T_g and T_a are the correlation time of the gyroscope and accelerometer, respectively.

The error of the dynamic lever-arm effect is mainly reflected in the error equation of speed and position, so the system measurement equation is

$$
\boldsymbol{Z}(t) = \boldsymbol{H}(t)\boldsymbol{X}(t) + \boldsymbol{V}(t) = \begin{bmatrix}
R_M \delta L \\
R_N \cos L \delta\lambda \\
\delta h \\
\delta V_E \\
\delta V_N \\
\delta V_v
\end{bmatrix} + \begin{bmatrix}
v_1 \\
v_2 \\
v_3 \\
v_4 \\
v_5 \\
v_6
\end{bmatrix} \tag{24}
$$

$$
\boldsymbol{H}(t) = \begin{bmatrix}
\boldsymbol{0}_{3\times6} & diag\begin{bmatrix} R_M & R_N \cos L & 1 \end{bmatrix} & \boldsymbol{0}_{3\times3} & \boldsymbol{0}_{3\times12} \\
\boldsymbol{0}_{3\times3} & diag\begin{bmatrix} 1 & 1 & 1 \end{bmatrix} & \boldsymbol{0}_{3\times6} & \boldsymbol{0}_{3\times12}
\end{bmatrix} \tag{25}
$$

RL-DLAC uses a Kalman filter with deterministic control, which can be expressed as

$$
\boldsymbol{X}_k = \boldsymbol{\Phi}_{k,k-1}\boldsymbol{X}_{k-1} + \boldsymbol{B}_{k-1}\boldsymbol{u}_{k-1} + \boldsymbol{\Gamma}_{k-1}\boldsymbol{W}_{k-1} \tag{26}
$$

$$
\hat{\boldsymbol{X}}_{k,k-1} = \boldsymbol{\Phi}_{k,k-1}\hat{\boldsymbol{X}}_{k-1} + \boldsymbol{B}_{k-1}\boldsymbol{u}_{k-1} \tag{27}
$$

$$
\begin{aligned}
\hat{\boldsymbol{X}}_k = {} & \boldsymbol{\Phi}_{k,k-1}\hat{\boldsymbol{X}}_{k-1} + \boldsymbol{B}_{k-1}\boldsymbol{u}_{k-1} \\
& + \boldsymbol{K}_k \left(\boldsymbol{Z}_k - \boldsymbol{H}_k\boldsymbol{\Phi}_{k,k-1}\hat{\boldsymbol{X}}_{k-1} - \boldsymbol{H}_k\boldsymbol{B}_{k-1}\boldsymbol{u}_{k-1} \right)
\end{aligned} \tag{28}
$$

$$
\boldsymbol{K}_k = \boldsymbol{P}_{k/k-1}\boldsymbol{H}_k^{\mathrm{T}} \left(\boldsymbol{H}_k\boldsymbol{P}_{k/k-1}\boldsymbol{H}_k^{\mathrm{T}} + \boldsymbol{R}_k \right)^{-1} \tag{29}
$$

$$
\boldsymbol{P}_{k/k-1} = \boldsymbol{\Phi}_{k,k-1}\boldsymbol{P}_{k-1}\boldsymbol{\Phi}_{k,k-1}^{\mathrm{T}} + \boldsymbol{\Gamma}_{k-1}\boldsymbol{Q}_{k-1}\boldsymbol{\Gamma}_{k-1}^{\mathrm{T}} \tag{30}
$$

$$
\boldsymbol{P}_k = (\boldsymbol{I} - \boldsymbol{K}_k\boldsymbol{H}_k)\boldsymbol{P}_{k/k-1} \tag{31}
$$

where $\boldsymbol{X}_k$ is the state of the estimated time k, $\boldsymbol{Z}_k$ is the measurement of the estimated time k, $\boldsymbol{\Phi}_{k,k-1}$ is the state

TABLE 2. The initial value of natural frequency.

Natural frequency	Initial value (rad/s)
ω_1^0	20.8453
ω_2^0	50.0123
ω_3^0	98.7743

transition matrix, $\boldsymbol{H}_k$ is the measurement matrix, $\boldsymbol{R}_k$ is the measurement noise covariance matrix, $\boldsymbol{P}_k$ is the covariance matrix of estimated state, and $\boldsymbol{K}_k$ is the gain matrix of filter.

2) STATE QUANTITY OF RL
Generally, a set of default calibration values is provided after production of the vehicle for corresponding types of structures. The initial value of the natural frequency of elastic deformation is used as an example, as shown in Table 2. In the actual flight process, due to the harsh flight environment of hypersonic vehicles, its natural frequency will change. Therefore, the natural frequency is taken as the state quantity of reinforcement learning.

3) REWARD FUNCTION OF RL
When the environment receives an action that causes a state change, we need to define a sufficient reward mechanism to evaluate the current action. Since the state quantity of reinforcement learning determines the integrated navigation accuracy, we modeled the reward as the opposite of the location error over the N seconds of IMU and GNSS data. The reward value is configured to the opposite of the final localization error, expressed as $\boldsymbol{r} = -\left|\hat{\boldsymbol{r}} - \boldsymbol{r}_G\right|$. By maximizing the reward, we will obtain the natural frequency of elastic deformation, which can minimize the positioning error.

4) AGENT OF RL
DDPG is an actor–critic, model-free algorithm based on the deterministic policy gradient that can operate over continuous action spaces and it includes two neural networks (NNs): a critic NN introduced to evaluate the long-time performance of the designed control in the current time step and an action NN used to output continuous action in the corresponding state [30]. Moreover, in order to improve learning efficiency and prevent local optimality, an experience replay memory is created to store historical samples, each time a certain number of samples are randomly selected for training. In each episode, the DDPG executes a certain time step. At each time step, the action α_i is selected by the action evaluation network of the ACTOR module, which is based on the cost function that maximizes the reward from taking accurate action. Although there are some designs of actor-critic network that can reduce the computational cost of reinforcement learning [31]–[33], DDPG is still selected as the agent of

reinforcement learning because the actual parameters will not change dramatically in a certain period of time.

DDPG builds an actor network for an agent to select actions via the actor–critic system, instead of using the traditional greedy algorithm. This method has been proven to enable learning of good policies for many tasks using low-dimensional observations. The network structure used in DDPG is Multi-Layer Perception (MLP) through multiple fully connected (FC) layers. The structure of the actor part is not the same as that of the critic part since the meanings of the input and output are different.

The actor evaluation network and the actor target network are three-layer fully connected networks. The dimension of the input layer is 3 and that of the hidden layer is 30, followed by a ReLU activation function. The output layer size is equal to the action. After applying a tanh activation function to the neurons of the output layer, we can finally obtain the predicted action. The key parameters in DDPG are listed in Table 3.

TABLE 3. The key parameters in DDPG.

Parameters	Value
Actor. learning rate	0.001
Critic. learning rate	0.002
DDPG. max_episodes	50
DDPG. max_episodes_steps	5
DDPG. memory_capacity	250
DDPG. soft_update	0.01

IV. SIMULATION ANALYSIS AND DISCUSSION

In this section, we present examples to verify the performance of the RL-DLAC for compensation of dynamic lever-arm effect error. Comparing the RL-DLAC method with the traditional fixed lever-arm compensation method based on Kalman filter (KF-FLAC), it can be found that RL-DLAC

$$F_2(t)_{9\times9} = \begin{bmatrix} C_b^n & C_b^n & 0_{3\times3} \\ 0_{3\times3} & 0_{3\times3} & C_b^n \\ & 0_{3\times9} & \end{bmatrix},$$

$$F_5(t)_{6\times6} = \begin{bmatrix} 0_{3\times3} & I_{3\times3} \\ K_2 & K_1 \end{bmatrix} \tag{18}$$

$$F_3(t)_{9\times9} = diag\left(0 \quad 0 \quad 0 \quad -\frac{1}{T_{gx}} \quad -\frac{1}{T_{gy}} \quad -\frac{1}{T_{gz}} \quad -\frac{1}{T_{ax}} \quad -\frac{1}{T_{ay}} \quad -\frac{1}{T_{az}}\right) \tag{19}$$

$$F_4(t)_{9\times6} = \begin{bmatrix} 0_{3\times3} & 0_{3\times3} \\ C_b^n\left(YK_2 + \dot{\omega}_{ib}^\times Y + \omega_{ib}^\times\omega_{ib}^\times Y\right) & C_b^n\left(YK_1 + 2\omega_{ib}^\times Y\right) \\ 0_{3\times3} & 0_{3\times3} \end{bmatrix} \tag{20}$$

$$\omega_{ib}^\times = \begin{bmatrix} 0 & -\omega_{ibz} & \omega_{iby} \\ \omega_{ibz} & 0 & -\omega_{ibx} \\ -\omega_{iby} & \omega_{ibx} & 0 \end{bmatrix}, \dot{\omega}_{ib}^\times$$

$$= \begin{bmatrix} 0 & -\dot{\omega}_{ibz} & \dot{\omega}_{iby} \\ \dot{\omega}_{ibz} & 0 & -\dot{\omega}_{ibx} \\ -\dot{\omega}_{iby} & \dot{\omega}_{ibx} & 0 \end{bmatrix} \tag{21}$$

$$B(t)_{24\times6} = \begin{bmatrix} 0_{3\times3} & 0_{3\times3} \\ 0_{3\times3} & C_b^n \\ 0_{15\times3} & 0_{15\times3} \\ I_{3\times3} & 0_{3\times3} \end{bmatrix}, \quad U(t) = \begin{bmatrix} u_q \\ u_r \end{bmatrix} \tag{22}$$

$$G(t)_{24\times15} = \begin{bmatrix} C_b^n & 0_{3\times3} & 0_{3\times3} & 0_{3\times3} & 0_{3\times3} \\ 0_{3\times3} & 0_{3\times3} & 0_{3\times3} & 0_{3\times3} & C_b^n \\ 0_{6\times3} & 0_{6\times3} & 0_{6\times3} & 0_{6\times3} & 0_{6\times3} \\ 0_{3\times3} & I_{3\times3} & 0_{3\times3} & 0_{3\times3} & 0_{3\times3} \\ 0_{3\times3} & 0_{3\times3} & I_{3\times3} & 0_{3\times3} & 0_{3\times3} \\ 0_{3\times3} & 0_{3\times3} & 0_{3\times3} & 0_{3\times3} & 0_{3\times3} \\ 0_{3\times3} & 0_{3\times3} & 0_{3\times3} & I_{3\times3} & 0_{3\times3} \end{bmatrix},$$

$$W(t) = \begin{bmatrix} w_g \\ w_r \\ w_a \\ w_q \\ w_l \end{bmatrix} \tag{23}$$

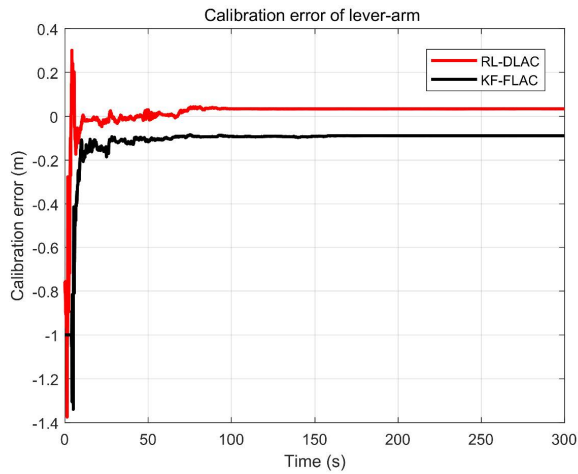**FIGURE 6.** Simulation diagram of fixed lever-arm calibration.

FIGURE 7. Flight trajectory of the hypersonic vehicle.

has more obvious advantages in the compensation of dynamic lever-arm effect.

For the performance evaluation of the proposed RL-DLAC method, two different applications are studied in the simulation analysis, which are as follows.

A. APPLICATION OF FIXED LEVER-ARM EFFECT COMPENSATION

The change in the dynamic lever-arm is small, when Q_i is small during flight. Besides, the influence of fuel consumption on the change of mass center is small in a period of time, so it can be approximately considered that the change in the lever-arm of the vehicle is 0. The calibration results of the RL-DLAC and KF-FLAC for the fixed lever-arm error are compared, as shown in the Fig. 6.

Fig. 6 shows that both RL-DLAC and KF-FLAC can calibrate the fixed lever-arm, but RL-DLAC has higher accuracy.

B. APPLICATION OF DYNAMIC LEVER-ARM EFFECT COMPENSATION

In the actual flight process, the hypersonic vehicle is facing a complex and dynamic environment. As shown in Fig. 7, a dynamic flight trajectory is simulated according to actual hypersonic vehicle flight, which involves various maneuvers such as climbing, pitching, rolling and turning.

The initial position of the hypersonic vehicle was at north latitude 30.011°, east longitude 120.094° and 30km altitude. The initial velocities along the three axes of the E-N-U navigation frame were 0m/s, 1700m/s and 0m/s, respectively. The initial position error was 5m, 5m, and 10m, initial velocity error was 0.2m/s, 0.2m/s, and 0.2m/s and initial attitude error was 1′, 1′, 1′. The gyro's constant drift and white noise were 0.1°/h and 0.01°/h. The accelerometer's zero bias and white noise were 10^{-3}g and 10^{-4}g. The root mean square errors (RMSEs) of GNSS position and velocity were 5m, and 0.1m/s. The sampling periods of INS and GNSS were 0.02s and 0.1s. The simulation time was 300s and the filtering period of the INS/GNSS integration was 0.1s.
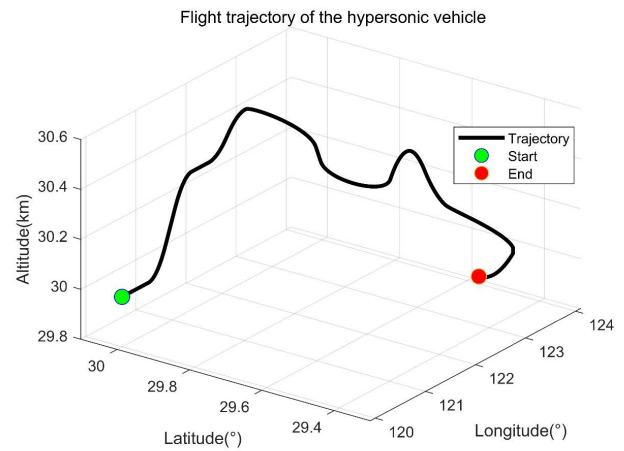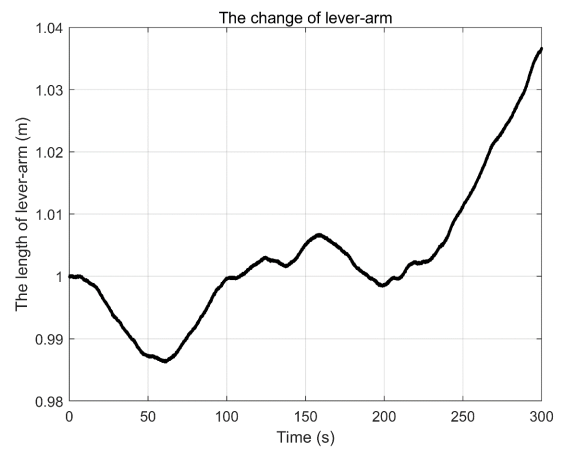

FIGURE 8. The change of lever-arm in the high dynamic environment.

In the high dynamic environment, the change of lever-arm is as shown in the Fig. 8.

It can be seen from Section II that the increase of surface temperature will reduce the bending stiffness of material under the influence of aerodynamic heating on the surface of hypersonic vehicle. Therefore, the natural frequency of the vibration mode of hypersonic vehicle will change. The initial value of the natural frequency of the hypersonic vehicle structure has been given in Table 2. It will change with the extension of flight time and the change of flight state. In order to make the simulation conform to reality, the change of natural frequency is simulated by using a set of test data from hypersonic vehicles, as shown in Table 4.

The RL-DLAC is used to identify the dynamic lever-arm, learn the change of its natural frequency, and improve the dynamic lever-arm model in the actual flight process. The natural frequency learning results are shown in Table 5.

It can be found from Table 5 that RL-DLAC has a good natural frequency learning effect, which determines the dynamic lever-arm model. The average error is only 0.31%.

The RL-DLAC and KF-FLAC are used to calibrate the dynamic lever-arm respectively, as shown in Fig. 9.

TABLE 4. The real value of the natural frequency.

Natural frequency	Initial value (rad/s)	Real value(rad/s)
ω_1^0	20.8453	20.1658
ω_2^0	50.0123	55.3610
ω_3^0	98.7743	113.0664

TABLE 5. The learning results of natural frequency.

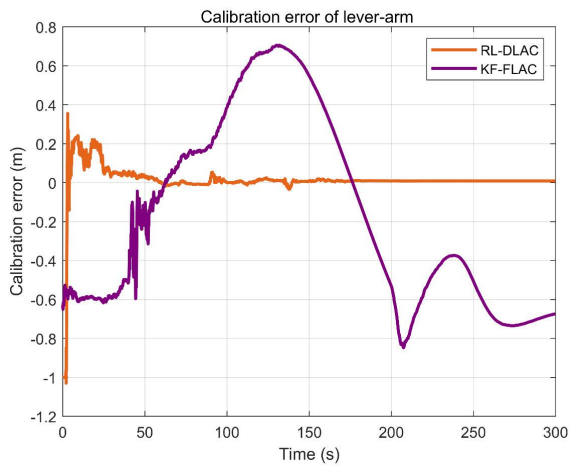| Natural frequency	Real value (rad/s)	Learned value (rad/s)	Error
ω_1^0	20.1658	20.1389	0.13%
ω_2^0	55.3610	55.1664	0.35%
ω_3^0	113.0664	113.5816	0.46%

FIGURE 9. Simulation diagram of dynamic lever-arm calibration.

Fig. 9 shows that RL-DLAC has better calibration results for dynamic lever-arm, and the traditional method is no longer applicable. The reason is that the lever-arm model designed by the traditional method is static, which ignores the influence of the lever-arm change in the actual flight environment. The RL-DLAC not only considers the change in the lever-arm, but also considers the change in parameters in the lever-arm model. Through reinforcement learning, the parameters of the lever-arm model are determined, and a more perfect dynamic lever arm model is established. Therefore, the compensation effect for the lever-arm is better.

The navigation errors of the two methods for hypersonic vehicles are shown in Fig. 10 and Fig. 11.

Obviously, Fig. 10 and Fig. 11 show that the navigation accuracy of using the RL-DLAC method to compensate for the dynamic lever-arm effect error is improved when
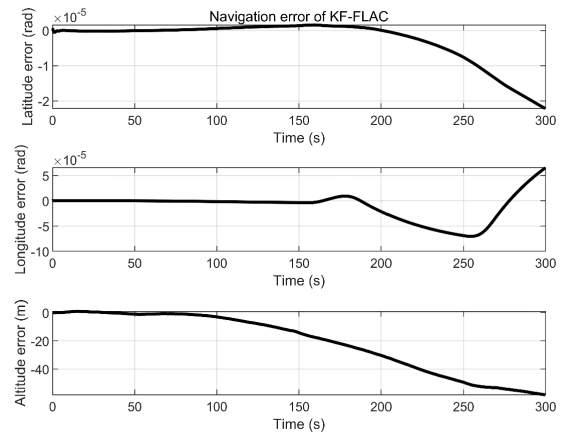

FIGURE 10. Navigation error of hypersonic vehicle with KF-FLAC.
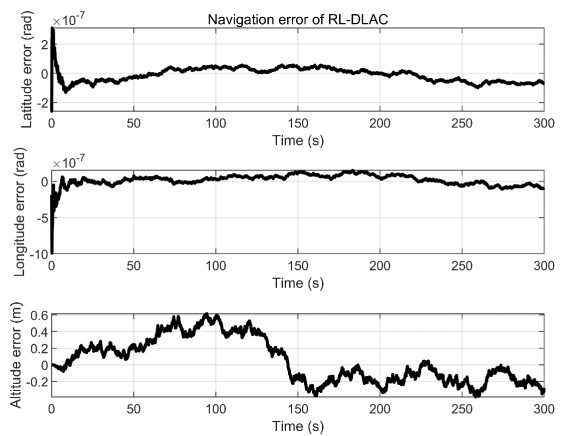

FIGURE 11. Navigation error of hypersonic vehicle with RL-DLAC.

compared with the traditional KF-FLAC method. This is because the lever-arm effect error caused by the dynamic lever-arm is not eliminated in the traditional method. Besides, due to the cumulative effect of errors, the navigation errors of the traditional method will further diverge over time.

Compared with the divergence effect of the traditional method in the face of the dynamic lever-arm effect, the RL-DLAC method proposed in this paper has better navigation accuracy. The main reason is that we have established a more perfect dynamic lever-arm model and the model parameters are learned through reinforcement learning to ensure the accuracy of the model during flight, which suppresses the divergence of navigation error.

V. CONCLUSION

In order to improve the navigation accuracy of hypersonic vehicles, a new compensation method for the dynamic lever-arm effect based on reinforcement learning (RL-DLAC), which can adapt to the complex flight environment, is proposed in this paper.

RL-DLAC has two significant features. (A) Focusing on the change of the lever-arm in a high dynamic environment, a more perfect error model of the lever-arm effect is

established based on elastic deformation theory, so that the influence of the lever-arm effect error caused by the change in the lever-arm can be reduced by compensation. (B) According to the changes of parameters in the lever-arm effect error model, a reinforcement learning method is designed based on the natural frequency, which realizes the adaptive estimation of the parameters of the lever-arm effect error dynamic model and enhances the robustness of the method. The RL-DLAC method can use the deep deterministic policy gradient (DDPG) to obtain the best state estimation and natural frequency from continuous actions. The simulation results show that the natural frequency of vehicle structure estimated by the method can obtain high-precision navigation results through the dynamic lever-arm error model. Compared with the traditional methods, it improves the compensation process of dynamic lever-arm effect and suppresses the divergence of navigation error caused by the change of lever-arm model.

The application of the intelligent method will also produce some limitations, such as high-performance computing power, which is a key challenge to the hardware system. If we want to realize the real-time control, a small network structure and design of the basis function need be developed, but it will be difficult to select the initial value for convergence. The DDPG is adopted in this paper, which is suitable for practical problems with strong nonlinearity. In order to reduce the computational cost, this paper has reduced the learning parameters to three, but even so, due to the existence of training in reinforcement learning, it will inevitably lead to the difficulty of real-time performance. Fortunately, due to the slow influence of model parameters caused by the environment, there will be no drastic change within a certain time range, which can maintain the high positioning accuracy of the navigation system and make up for the problem of long training time. If the flight time is short in practical application, we can also collect data first, adopt offline training, and then use the learned parameters when flying in the same environment. Future research work will focus on the improvement of the RL-DLAC. We expect to propose a dynamic lever-arm effect compensation method based on the model-free learning method.

REFERENCES

[1] H. An, Q. Wu, G. Wang, Y. Kao, and C. Wang, "Adaptive compound control of air-breathing hypersonic vehicles," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 56, no. 6, pp. 4519–4532, Dec. 2020.

[2] H. An, Q. Wu, C. Wang, and X. Cao, "Scramjet operation guaranteed longitudinal control of air-breathing hypersonic vehicles," *IEEE/ASME Trans. Mechatronics*, vol. 25, no. 6, pp. 2587–2598, Dec. 2020.

[3] G. Hu, L. Ni, B. Gao, X. Zhu, W. Wang, and Y. Zhong, "Model predictive based unscented Kalman filter for hypersonic vehicle navigation with INS/GNSS integration," *IEEE Access*, vol. 8, pp. 4814–4823, 2020.

[4] H. He, B. Zhu, and F. Zha, "Particle swarm optimization-based gyro drift estimation method for inertial navigation system," *IEEE Access*, vol. 7, pp. 55788–55796, 2019.

[5] S. Wang, G. Yang, and L. Wang, "An improve hybrid calibration scheme for strapdown inertial navigation system," *IEEE Access*, vol. 7, pp. 151669–151681, 2019.

[6] Z. Xiong, H. Peng, J. Wang, R. Wang, and J.-Y. Liu, "Dynamic calibration method for SINS lever-arm effect for HCVs," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 51, no. 4, pp. 2760–2771, Oct. 2015.

[7] J. Lu, L. Xie, and B. Li, "Applied quaternion optimization method in transfer alignment for airborne AHRS under large misalignment angle," *IEEE Trans. Instrum. Meas.*, vol. 65, no. 2, pp. 346–354, Feb. 2016.

[8] S. Hong, M. H. Lee, H.-H. Chun, S.-H. Kwon, and J. L. Speyer, "Experimental study on the estimation of lever arm in GPS/INS," *IEEE Trans. Veh. Technol.*, vol. 55, no. 2, pp. 431–448, Mar. 2006.

[9] Z. Liu, L. Wang, K. Li, and J. Gao, "A compensation method of lever arm effect for tri-axis hybrid inertial navigation system based on fiber optic gyro," *Meas. Sci. Technol.*, vol. 28, no. 5, May 2017, Art. no. 055103.

[10] Q. Wu, K. Li, and J. Liu, "The asynchronous gimbal-rotation-based calibration method for lever-arm errors of two rotational inertial navigation systems," *IEEE Access*, vol. 7, pp. 4653–4663, 2019.

[11] Q. Wu, K. Li, and W. Liang, "An improved calibration and compensation method for lever-arm errors between two rotational inertial navigation systems," in *Proc. IEEE Int. Conf. Mechatronics Autom. (ICMA)*, Tianjin, China, Aug. 2019, p. 2457.

[12] X. Wang, Y. Han, and L. Xie, "A single-axis rotation modulation initial alignment algorithm based on inner lever arm effect compensation in accelerometers," in *Proc. 3rd Int. Conf. Unmanned Syst. (ICUS)*, Harbin, China, Nov. 2020, p. 162.

[13] S. Yang, X. Hu, X. Wang, and D. Li, "A structure scheme to reduce the influence of the accelerometer inner lever-arm effect on tri-axis inertial navigation systems," *Meas. Sci. Technol.*, vol. 32, no. 5, Mar. 2021, Art. no. 055103.

[14] Y. H. Cho, W. J. Park, and C. G. Park, "Novel methods of mitigating lever arm effect in redundant IMU," *IEEE Sensors J.*, vol. 21, no. 7, pp. 9465–9474, Apr. 2021.

[15] B. Du, Z. Shi, M. Ding, L. Han, and J. Song, "The calibration method for accelerometers in the redundant MEMS inertial navigation system," *Meas. Sci. Technol.*, vol. 32, no. 9, Sep. 2021, Art. no. 095004.

[16] R. Wang, X. Hou, F. Liu, and Y. Yu, "GPS/INS integrated navigation for quadrotor UAV considering lever arm," in *Proc. 35th Youth Academic Annu. Conf. Chin. Assoc. Autom. (YAC)*, Zhanjiang, China, Oct. 2020, p. 132.

[17] T. Song, K. Li, Q. Wu, Q. Li, and Q. Xue, "An improved self-calibration method with consideration of inner lever-arm effects for a dual-axis rotational inertial navigation system," *Meas. Sci. Technol.*, vol. 31, no. 7, Apr. 2020, Art. no. 074001.

[18] G.-B. Chang and M. Liu, "Hybrid Kalman and unscented Kalman filters for INS/GPS integrated system considering constant lever arm effect," *J. Central South Univ.*, vol. 22, no. 2, pp. 575–583, Feb. 2015.

[19] L. Zhang and T. Zhang, "A fast calibration method for dynamic lever-arm parameters for IMUs based on the backtracking scheme," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 9, pp. 6946–6956, Sep. 2020.

[20] Q. Wu, K. Li, and T. Song, "The calibration for inner and outer lever-arm errors based on velocity differences of two RINSs," *Mech. Syst. Signal Process.*, vol. 160, Nov. 2021, Art. no. 107868.

[21] X. Liu, X. Xu, L. Wang, Y. Li, and Y. Liu, "H∞ filter for flexure deformation and lever arm effect compensation in M/S INS integration," *Int. J. Nav. Archit. Ocean Eng.*, vol. 6, no. 3, pp. 626–637, Sep. 2020.

[22] T. Wang, J. Cheng, D. Guan, Y. Kang, and W. Zhang, "Modified compensation algorithm of lever-arm effect and flexural deformation for polar shipborne transfer alignment based on improved adaptive Kalman filter," *Meas. Sci. Technol.*, vol. 28, no. 9, Aug. 2017, Art. no. 095101.

[23] Q. Cao, M. Zhong, and J. Guo, "Non-linear estimation of the flexural lever arm for transfer alignment of airborne distributed position and orientation system," *IET Radar, Sonar Navigat.*, vol. 11, no. 1, pp. 41–51, May 2017.

[24] A. Anwar and A. Raychowdhury, "Autonomous navigation via deep reinforcement learning for resource constraint edge nodes using transfer learning," *IEEE Access*, vol. 8, pp. 26549–26560, 2020.

[25] O. Bouhamed, H. Ghazzai, H. Besbes, and Y. Massoud, "Autonomous UAV navigation: A DDPG-based deep reinforcement learning approach," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, Sevilla, Spain, Oct. 2020, p. 1.

[26] Y. Liu, H. Wang, T. Wu, Y. Lun, J. Fan, and J. Wu, "Attitude control for hypersonic reentry vehicles: An efficient deep reinforcement learning method," *Appl. Soft Comput.*, vol. 123, Jul. 2022, Art. no. 108865.

[27] N. Wang, Y. Gao, H. Zhao, and C. K. Ahn, "Reinforcement learning-based optimal tracking control of an unknown unmanned surface vehicle," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 7, pp. 3034–3045, Jul. 2021.

[28] A. Culler, T. Williams, and M. Bolender, "Aerothermal modeling and dynamic analysis of a hypersonic vehicle," in *Proc. AIAA Atmos. Flight Mech. Conf. Exhib.*, Hilton Head, South Carolina, Aug. 2007, p. 6395.

[29] G. Hu, B. Gao, Y. Zhong, L. Ni, and C. Gu, "Robust unscented Kalman filtering with measurement error detection for tightly coupled INS/GNSS integration in hypersonic vehicle navigation," *IEEE Access*, vol. 7, pp. 151409–151421, 2019.

[30] X. Gao, H. Luo, B. Ning, F. Zhao, L. Bao, Y. Gong, and J. Jiang, "RL-AKF: An adaptive Kalman filter navigation algorithm based on reinforcement learning for ground vehicles," *Remote Sens.*, vol. 12, no. 11, pp. 1704–1729, May 2020.

[31] X. Bu and Q. Qi, "Fuzzy optimal tracking control of hypersonic flight vehicles via single-network adaptive critic design," *IEEE Trans. Fuzzy Syst.*, vol. 30, no. 1, pp. 270–278, Jan. 2022.

[32] X. Bu, Y. Xiao, and H. Lei, "An adaptive critic design-based fuzzy neural controller for hypersonic vehicles: Predefined behavioral nonaffine control," *IEEE/ASME Trans. Mechatron.*, vol. 24, no. 4, pp. 1871–1881, May 2019.

[33] Q. Qi and X. Bu, "Adaptive dynamic programing design for the neural control of hypersonic flight vehicles," *J. Franklin Inst.*, vol. 358, no. 16, pp. 8169–8192, Oct. 2021.

BO YANG received the Ph.D. degree in aerospace engineering from Beihang University (BUAA), in 2009. She is currently an Associate Professor with the School of Astronautics, Beihang University. Her current research interests include astronomical autonomous navigation technology of spacecraft, autonomous navigation and guidance technology for deep space exploration, and GNC key technology of repeatable launch vehicle.


HE YU received the bachelor's degree in detection, guidance, and control from Beihang University, Beijing, China, in 2018, where he is currently pursuing the Ph.D. degree in aerospace engineering with the School of Astronautics. His current research interests include integrated navigation technology and INS error compensation technology.

CHAOFAN LIU received the bachelor's degree in detection, guidance, and control from Beihang University, Beijing, China, in 2020, where he is currently pursuing the master's degree in electronic information with the School of Astronautics. His current research interest includes aero-optical effects.

XIANG WEI received the bachelor's degree in detection, guidance, and control from Beihang University, Beijing, China, in 2020. He is currently pursuing the master's degree in electronic information with the School of Astronautics. His current research interest includes dynamic control allocation.
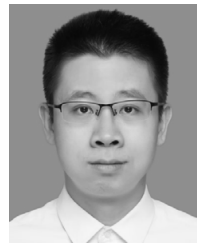
ZICHEN FAN received the Ph.D. degree in flight vehicle design and engineering from Beihang University (BUAA), in 2021. He is currently working with the Beijing Institute of Control and Electronic Technology. His current research interest includes astronomical autonomous navigation technology of spacecraft.

JUN MIAO received the Ph.D. degree in flight vehicle design and engineering from Beihang University (BUAA), in 2018. He is currently working with the Qian Xuesen Laboratory of Space Technology. His current research interests include spacecraft attitude and orbit control.

• • •