SCIENCE CHINA
Technological Sciences

• Article •

February 2024 Vol. 67 No. 2: 435–448
https://doi.org/10.1007/s11431-023-2463-0

Extremum seeking control for UAV close formation flight via improved pigeon-inspired optimization

YUAN GuangSong & DUAN HaiBin*

State Key Laboratory of Virtual Reality Technology and Systems, School of Automation Science and Electrical Engineering, Beihang University (BUAA), Beijing 100083, China

Received April 13, 2023; accepted July 4, 2023; published online December 19, 2023

This paper proposes a comprehensive design scheme for the extremum seeking control (ESC) of the unmanned aerial vehicle (UAV) close formation flight. The proposed design scheme combines a Newton-Raphson method with an extended Kalman filter (EKF) to dynamically estimate the optimal position of the following UAV relative to the leading UAV. To reflect the wake vortex effects reliably, the drag coefficient induced by the wake vortex is considered as a performance function. Then, the performance function is parameterized by the first-order and second-order terms of its Taylor series expansion. Given the excellent performance of nonlinear estimation, the EKF is used to estimate the gradient and the Hessian matrix of the parameterized performance function. The output feedback of the proposed scheme is determined by iterative calculation of the Newton-Raphson method. Compared with the traditional ESC and the classic ESC, the proposed design scheme avoids the slow continuous time integration of the gradient. This allows a faster convergence of relative position extremum. Furthermore, the proposed method can provide a smoother command during the seeking process as the second-order term of the performance function is taken into account. The convergence analysis of the proposed design scheme is accomplished by showing that the output feedback is a supermartingale sequence. To improve estimation performance of the EKF, a improved pigeon-inspired optimization (IPIO) is proposed to automatically tune the noise covariance matrix. Monte Carlo simulations for a three-UAV close formation show that the proposed design scheme is robust to the initial position of the following UAV.

unmanned aerial vehicle, close formation, extremum seeking control, Newton-Raphson method, improved pigeon-inspired optimization

Citation: Yuan G S, Duan H B. Extremum seeking control for UAV close formation flight via improved pigeon-inspired optimization. Sci China Tech Sci, 2024, 67: 435–448, https://doi.org/10.1007/s11431-023-2463-0

1 Introduction

Close formation of unmanned aerial vehicle (UAV) has been increasingly studied and researched in the field of UAV formations [1–6]. Its potential benefits include enhanced coordination, reduced radar cross section, improved obstacle avoidance and especially increased energy efficiency. In close formation flight, the following UAV can take advantage of the upwash of the wake vortex induced by the leading UAV to enhance lift and decrease drag [7]. Therefore, saving energy can be realized by the following UAV, which has been validated through theoretical investigation [8,9], observed in wind tunnel experiments [10], and substantiated by flight tests [11]. However, the drag reduction achieved by the following UAV is highly sensitive to the position relative to the leading UAV (an error of just 10% of the wingspan can reduce the benefits by half) [12]. Furthermore, the meandering of the leading UAV and random wind turbulence can affect the characteristics of the wake vortex, which leads to the uncertainty of the optimal relative position [13, 14]. Therefore, in order to maintain a consistently high reduction in drag, it is crucial for the following UAV to continuously locate the optimal relative

*Corresponding author (email: hbduan@buaa.edu.cn)

position.

To achieve this control objective, significant efforts have been made with various control schemes [15–19]. The model-based open loop control scheme has been proposed to achieve drag reduction [20–24]. In this strategy, the optimal relative position is predetermined by conducting an investigation of the wake vortex using the established model. Additionally, this predetermined optimal relative position is assumed to be unchanged in close formation flight. Inevitably, such a scheme can not provide the maximal drag reduction. This is because its effectiveness is limited by the uncertainty of the optimal relative position and high sensitivity of the drag reduction to position. Therefore, it is desirable to develop a model-free adaptive feedback control that can generate a dynamic estimation of the optimal relative position during close formation flight. To this end, extremum search control (ESC) is introduced into the design of the close formation controller. In ref. [25], the traditional ESC was utilized to estimate the optimal relative position dynamically. Its fundamental concept is to find the extremum points of the performance function by the process of modulation and demodulation. However, the traditional ESC based on modulation and demodulation exhibits slow convergence speed when estimating the optimal relative position. To improve the convergence speed, ref. [26] takes the classic ESC to perform the estimation of optimal relative position. In the classic ESC, the extended Kalman filter (EKF) is used to estimate the gradient information. While the ESC based on the EKF has been a significant improvement in the convergence speed of the optimal relative position, there are occurrences of highly nonsmooth behaviour during the convergence. The reason behind this issue may be that only estimating the gradient can not adequately capture the complex nonlinear dynamics of the wake vortex model.

Motivated by the unresolved problems aforementioned, this paper proposes a Newton-Raphson method based ESC to estimate the optimal relative position Dynamically. According to the Newton-Raphson method, the performance function is assumed to be approximated by a quadratic function. This is accomplished by parameterizing the performance function using first-order and second-order terms of a Taylor series expansion. Given the excellent performance of nonlinear estimation, the EKF is still utilized to estimate the gradient and Hessian matrix of the parameterized performance function [27, 28]. Compared with the traditional and classic ESC, the proposed scheme avoids the slow continuous time integration by iterative calculation. This allows a faster convergence of optimal relative position. Furthermore, the proposed scheme can provide a smoother convergence as the second-order term of the performance function is taken into account. To form a feedback, a continuous horseshoe vortex method with high accuracy is used to establish the wake vortex model [29]. Additionally, the established wake vortex model is also used to perform the static analysis of optimal relative position and yield powerful insight into the ESC of the relative position. The drag coefficient induced by the wake vortex is considered as the performance function.

It has been widely recognized that tuning the free parameters of the noise covariance matrix plays a crucial role in improving the estimation performance of the EKF [30]. By optimizing the free parameters, the EKF can adapt more effectively to the dynamic characteristics of the wake vortex model and the uncertainties in the wake vortex measurements. Due to the limitation of knowledge and experience, the manual tuning method for free parameters is generally not only excessively time-intensive, but also considerably difficult to acquire the optimal parameter combination. Different from the manual tuning method, the evolutionary algorithm has been proven to be feasible and effective for automatic tuning of the free parameters [31, 32]. However, basic evolutionary algorithms, such as particle swarm optimization (PSO) [33, 34], genetic algorithms (GA) [35], and pigeon-inspired optimization (PIO) [36] are generally suitable for single-modal optimization problems with a limited number of optimization parameters [37, 38]. Due to the requirement of estimating the gradient and Hessian matrix in our proposed Newton-Raphson method based ESC, the number of optimization parameters in the EKF is significantly increased. Furthermore, the highly nonlinear characteristic of the wake vortex model gives rise to a multi-modal optimization problem. For multi-variable and multi-modal characteristics of the proposed Newton-Raphson method based ESC, therefore, an improved PIO (IPIO) is co-proposed to escape local optima and enable a more global search. The IPIO can effectively handle the complex nonlinear relationship between the free parameters of the EKF and the performance of the proposed Newton-Raphson method based ESC. As a result, it has the potential to enhance the performance.

The main contributions of this paper are summarized as follows.

(1) A Newton-Raphson method based ESC is proposed to estimate the optimal relative position of UAV close formation dynamically.

(2) A continuous horseshoe vortex method with high accuracy is used to establish the wake vortex model.

(3) The EKF is used to estimate the gradient and the Hessian matrix.

(4) The IPIO is co-proposed to automatically tune the free parameters of the EKF.

2 Aerodynamic modeling

The aerodynamics of multiple UAVs in a close formation flight is much more complex compared with the aerodynamics in free formation flight because of aerodynamic interactions. As illustrated in Figure 1, the wake vortex generated by a leading UAV induces the regions of downwash and upwash behind its wings. A following UAV can utilize the upwash of the induced velocity to reduce its induced drag at the fixed lift, which is the mechanism of the drag reduction in close formation flight. To quantitatively describe the reduction of the induced drag, the wake vortex of the leading UAV is modeled firstly by only considering the wing. Then, the primary effects of the wake vortex on the following UAV are evaluated by transforming the induced velocity into the induced drag coefficient.

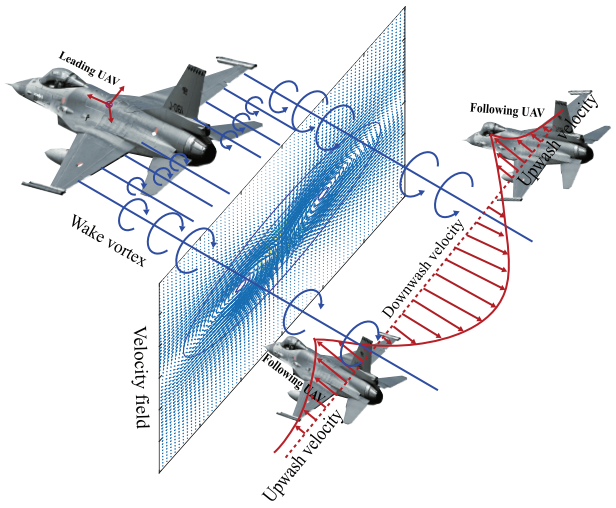

Figure 1 (Color online) UAVs close formation flight.

2.1 Wake vortex model

Considering both accuracy of estimation and the effectiveness of calculation for the real-time application, a continuous horseshoe vortex method is utilized to develop the wake vortex model of the leading UAV [29]. As shown in Figure 2, the wake vortex can be modeled as an infinite number of semi-infinite horseshoe vortices that are continuously distributed. Based on the structural characteristics of the horseshoe vortex, the wake vortex can be divided into a bound vortex and a free-trailing vortex. The bound vortex is attached to the wing surface and its filaments align with the quarter-chord line. On the other hand, the free-trailing vortex separates from the wing surface and its filaments extend downstream to infinity in parallel with the velocity vector of the leading UAV.

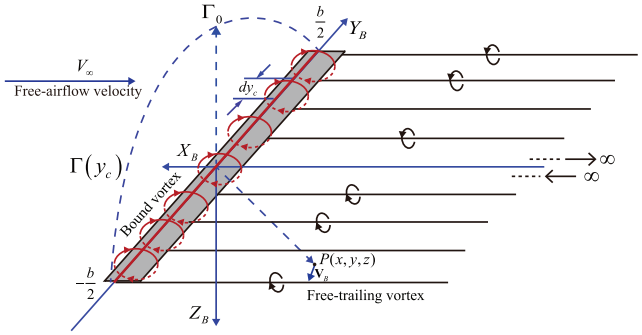

Figure 2 (Color online) Composition of the wake vortex.

For an approximately rectangular wing, the lift of a leading UAV can be considered to exhibit an elliptical distribution along the quarter-chord line. Accordingly, the circulation distribution surrounding the wing is assumed to be

$$\Gamma(y_c)=\Gamma_0\sqrt{1-\left(\frac{2y_c}{b}\right)^2},\quad -b/2\leqslant y_c\leqslant b/2, \tag{1}$$

where y_c is the lateral coordinate of the point on the quarter-chord line; $\Gamma_0=2V_\infty SC_L/(b\pi)$ denotes the circulation at the wing root with V_∞ being the magnitude of the free-airflow velocity, S and b being the area and span of the wing, and C_L representing the lift coefficient.

In accordance with the continuous horseshoe vortex method, the magnitude of the wake velocity induced by the wake vortex is determined by evaluating the induced velocity of the single straight vortex filament along the quarter-chord line. Therefore, the magnitude of the wake velocity at an arbitrary point $P(x,y,z)$ is formulated as follows:

$$\begin{cases} V_{bX_B}=\displaystyle\int_{-b/2}^{0}\kappa\cdot\mu\cdot\frac{\Gamma(y_c)z}{4\pi(x^2+z^2)} \\ \qquad\cdot\left(\frac{y-y_c}{\sqrt{x^2+(y-y_c)^2+z^2}}-\frac{y+y_c}{\sqrt{x^2+(y+y_c)^2+z^2}}\right)\cdot \mathrm{d}y_c, \\ V_{bZ_B}=\displaystyle\int_{-b/2}^{0}\kappa\cdot\mu\cdot\frac{\Gamma(y_c)x}{4\pi(x^2+z^2)} \\ \qquad\cdot\left(\frac{y-y_c}{\sqrt{x^2+(y-y_c)^2+z^2}}-\frac{y+y_c}{\sqrt{x^2+(y+y_c)^2+z^2}}\right)\cdot \mathrm{d}y_c, \end{cases} \tag{2}$$

$$\begin{cases} V_{fY_B}=\displaystyle\int_{-b/2}^{b/2}\kappa\cdot\mu\cdot\frac{\Gamma(y_c)z}{4\pi[(y-y_c)^2+z^2+r_c^2]} \\ \qquad\cdot\left(1-\frac{x}{\sqrt{x^2+(y-y_c)^2+z^2}}\right)\cdot \mathrm{d}y_c, \\ V_{fZ_B}=\displaystyle\int_{-b/2}^{b/2}\kappa\cdot\mu\cdot\frac{\Gamma(y_c)(y-y_c)}{4\pi[(y-y_c)^2+z^2+r_c^2]} \\ \qquad\cdot\left(1-\frac{x}{\sqrt{x^2+(y-y_c)^2+z^2}}\right)\cdot \mathrm{d}y_c, \end{cases} \tag{3}$$

where V_{fY_B} and V_{fZ_B} are induced by the free-trailing vortex on the Y_B-axis and Z_B-axis, respectively; V_{bX_B} and V_{bZ_B} are

induced by the bound vortex on the X_B-axis and Z_B-axis, respectively; r_c is the core radius of the free-trailing vortex aimed to resolve any potential issues of singularity; κ represents the rate of decay on the strength of the wake vortex; μ denotes the interaction coefficient of filaments. The wake velocity of the point P is defined as $\boldsymbol{V}_B = [V_{X_B}, V_{Y_B}, V_{Z_B}]$, where V_{X_B} ,V_{Y_B}, and V_{Z_B} are the velocity components on the X_B, Y_B, and Z_B axes, respectively. Therefore, $V_{X_B} = V_{b_{X_B}}$,$V_{Y_B} = V_{f_{Y_B}}$, and $V_{Z_B} = V_{b_{Z_B}} + V_{f_{Z_B}}$. Note that V_{X_B}, V_{Y_B}, and V_{Z_B} are commonly called the backwash, sidewash, and upwash velocity, respectively.

2.2 Induced drag coefficient

If flying in the upwash region, a following UAV can experience forward tilt in its lift vector owing to the change of the local angle of attack caused by the upwash as demonstrated in Figure 3. Let $\Delta\alpha$ represent the change of the local angle of attack at arbitrary point on the quarter-chord line of the following UAV, which can be calculated by $\Delta\alpha = -(V_{Z_B}/V_\infty)$. Let $\boldsymbol{L}$ denotes the fixed lift and $\boldsymbol{L}'$ represents the effective lift under the change of the local angle of attack. Using the statistical average strategy, the induced lift coefficient is determined by

$$\Delta C_L = \frac{1}{N}\left(\sum_{i=1}^{N} C_{L_\alpha}\Delta\alpha_i\right), \tag{4}$$

where N represents the number of statistical points, C_{L_α} is the lift-curve slope, $\Delta\alpha_i$ is the change of the local angle of attack at the i-th statistical point. Eventually, the induced drag coefficient is estimated according to the induced lift coefficient. This can be formulated as

$$\Delta C_D = \frac{(C_L + \Delta C_L)^2 - C_L^2}{\pi e_0 AR}, \tag{5}$$

where C_L is the lift coefficient, e_0 is the Oswald efficiency number, AR denotes the aspect ratio that can be calculated by $AR = b^2/S$.

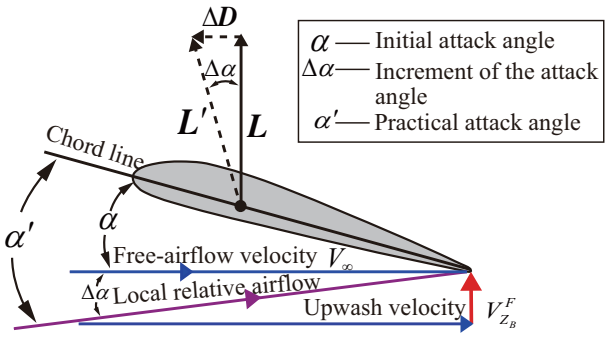

Figure 3 (Color online) Principle of the induced drag.

3 Newton-Raphson method based ESC

In this section, the Newton-Raphson method based ESC is proposed to to dynamically estimate the optimal position relative to the leading UAV. The following UAV is then commanded toward this relative position. The entire design scheme is demonstrated in Figure 4. In the proposed design scheme, the Newton-Raphson method is utilized to estimate the coordinates of the extremum of the performance function. Given the iterative nature of the Newton-Raphson method, the gradient and Hessian of the performance function is estimated using the EKF. Finally, the convergence of the proposed Newton-Raphson method based ESC is analyzed.

3.1 Extremum location estimation

According to published work, the induced drag coefficient is mainly determined by the lateral and vertical relative separations between the leading and following UAV, respectively. Furthermore, the induced drag coefficient with respect to these two variables exhibits a unimodal behavior. Therefore, the induced drag coefficient is considered as the performance function $f(\boldsymbol{x})$ for the extremum seeking. The Newton-Raphson method is used to estimate the extremum position of the performance function. In accordance with the Newton-Raphson method, the $f(\boldsymbol{x})$ is assumed to be approximated by a quadratic function at any point. Then, it is parameterized using first-order and second-order terms of its Taylor series expansion.

Considering any arbitrary point around a particular point $\boldsymbol{x}_k$, the performance function is formulated as

$$f(\boldsymbol{x}) \approx f(\boldsymbol{x}_k) + \boldsymbol{g}_k^{\mathrm{T}}(\boldsymbol{x} - \boldsymbol{x}_k) + \frac{1}{2}(\boldsymbol{x} - \boldsymbol{x}_k)^{\mathrm{T}}\boldsymbol{H}_k(\boldsymbol{x} - \boldsymbol{x}_k), \tag{6}$$

where $\boldsymbol{x} = [y, z]^{\mathrm{T}}$ is the relative position inputs that y and z are the lateral and vertical respectively, $\boldsymbol{g}_k$ and $\boldsymbol{H}_k$ are the gradient and Hessian at $\boldsymbol{x}_k$ respectively. For notational simplicity, the origin of the coordinate system can be shifted to the point $\boldsymbol{x}_k$ by redefining the above equation with respect to distance between $\boldsymbol{x}_k$ and any other point. This can be written as

$$\Delta f_k = \boldsymbol{g}_k^{\mathrm{T}}\Delta\boldsymbol{x}_k + \frac{1}{2}\Delta\boldsymbol{x}_k^{\mathrm{T}}\boldsymbol{H}_k\Delta\boldsymbol{x}_k, \tag{7}$$

where $\Delta f_k = f(\boldsymbol{x}) - f(\boldsymbol{x}_k)$ and $\Delta\boldsymbol{x}_k = \boldsymbol{x} - \boldsymbol{x}_k$. At extremum coordinate system, the extremum point of the equation can be found by making the derivative equal to zero with respect to $\Delta\boldsymbol{x}_k$. Perform

$$\frac{\partial\Delta f_k}{\partial\Delta\boldsymbol{x}_k} = \boldsymbol{g}_k + \boldsymbol{H}_k\Delta\boldsymbol{x}_k = \boldsymbol{0}. \tag{8}$$

By solving this equation, one can derive

$$\Delta\boldsymbol{x}_k = -\boldsymbol{H}_k^{-1}\boldsymbol{g}_k. \tag{9}$$

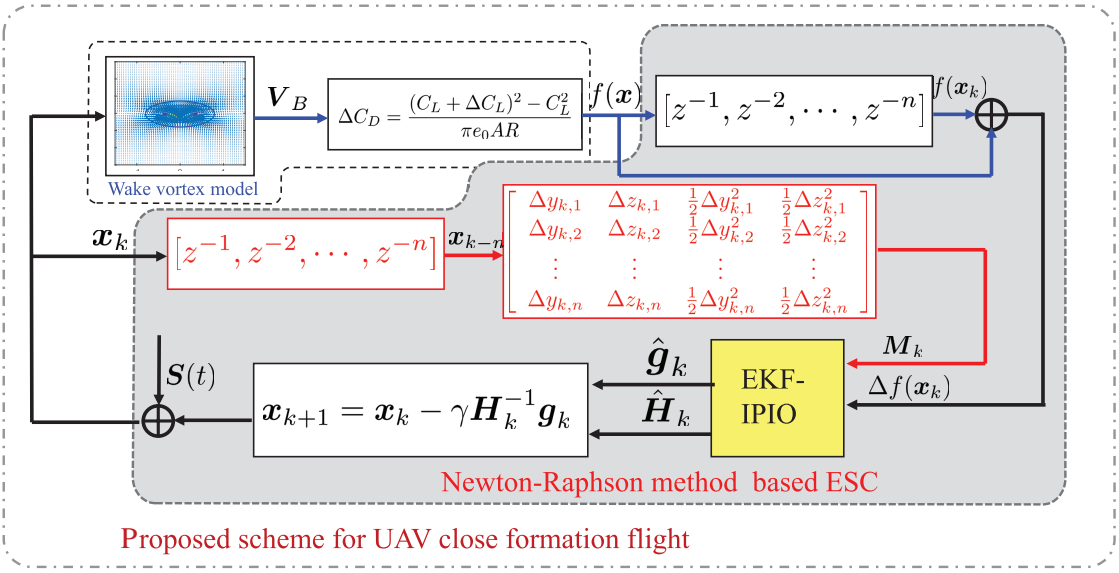

Figure 4 (Color online) Framework of Newton-Raphson method based ESC.

Then, the extremum coordinate of the performance function is found by using $\Delta \boldsymbol{x}_k = \boldsymbol{x} - \boldsymbol{x}_k$. Defining the estimated extremum position as $\boldsymbol{x}_{k+1}$, the above process is repeated as

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k - \gamma \boldsymbol{H}_k^{-1} \boldsymbol{g}_k \tag{10}$$

until a sufficiently precise value is reached. In the iterative equation, γ is the step size. This is done to ensure that the Wolfe conditions are satisfied at each step.

3.2 Gradient and Hessian estimation

In proposed design scheme, the estimation of the gradient and Hessian of the performance function is recognized as a crucial component that directly influences convergence speed and accuracy. To minimize the estimation variance and accelerate convergence, the EKF is considered to estimate the gradient and Hessian.

The gradient and Hessian at the relative position input $\boldsymbol{x}_k = (y_k, z_k)$ are denoted as $\boldsymbol{g}_k = [g_{1_k}, g_{2_k}]^{\mathrm{T}}$ and $\boldsymbol{H}_k = \mathrm{diag}([H_{11_k}, H_{22_k}])$, respectively. Since the gradient of the performance function has significant change in the neighborhood of the extremum position, the components associated with the coupling terms of $\boldsymbol{H}_k$ have been excluded to prevent potential oscillatory convergence. Then, eq. (7) can be reformulated in a more understandable form as

$$\Delta f_k = \begin{bmatrix} \Delta y_k & \Delta z_k & \frac{1}{2}\Delta y_k^2 & \frac{1}{2}\Delta z_k^2 \end{bmatrix} \begin{bmatrix} g_{1_k} \\ g_{2_k} \\ H_{11_k} \\ H_{22_k} \end{bmatrix}, \tag{11}$$

where $\Delta y_k = y_k - y_{k-1}$ and $\Delta z_k = z_k - z_{k-1}$. Clearly, the gradient and Hessian of the performance function are defined as the states of the EKF, namely $\boldsymbol{s}_k = [g_{1_k}, g_{2_k}, H_{11_k}, H_{22_k}]^{\mathrm{T}}$. Furthermore, these states are assumed to remain constant in a deterministic sense. To ensure observability, a minimum of four measurements at different time instants are required to satisfy that the rank of the observability Gramian is equal to the rank of the output matrix. This can be written as

$$\mathrm{rank}\,(\boldsymbol{F}) = \mathrm{rank} \begin{pmatrix} \boldsymbol{M}_k \\ \boldsymbol{M}_k \boldsymbol{I}_k \end{pmatrix} = \mathrm{rank}\,(\boldsymbol{M}_k), \tag{12}$$

where $\boldsymbol{F}$ is the observability Gramian matrix, $\boldsymbol{I}_k$ and $\boldsymbol{M}_k$ are the state-transition and measurement matrix, respectively. Note that $\boldsymbol{I}_k$ is identity matrix. The difference of a measure $\boldsymbol{x}$ between at iteration k and $k-n$ is defined as $\Delta \boldsymbol{x}_{k,n} = \boldsymbol{x}_k - \boldsymbol{x}_{k-n}$ with $n = 1, 2, \ldots, N$. Thus, the recursive state and measurement equations of the EKF are constructed as

$$\boldsymbol{s}_{k+1} = \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}}_{\boldsymbol{I}_k} \boldsymbol{s}_k + \boldsymbol{\omega}_k, \tag{13}$$

$$\Delta \boldsymbol{f}_{k+1} = \underbrace{\begin{bmatrix} \Delta y_{k,1} & \Delta z_{k,1} & \frac{1}{2}\Delta y_{k,1}^2 & \frac{1}{2}\Delta z_{k,1}^2 \\ \Delta y_{k,2} & \Delta z_{k,2} & \frac{1}{2}\Delta y_{k,2}^2 & \frac{1}{2}\Delta z_{k,2}^2 \\ \vdots & \vdots & \vdots & \vdots \\ \Delta y_{k,n} & \Delta z_{k,n} & \frac{1}{2}\Delta y_{k,n}^2 & \frac{1}{2}\Delta z_{k,n}^2 \end{bmatrix}}_{\boldsymbol{M}_k} \boldsymbol{s}_{k+1} + \boldsymbol{\nu}_{k+1}, \tag{14}$$

where $\boldsymbol{\omega}_k$ and $\boldsymbol{\nu}_{k+1}$ represent the process noise and the measurement noise, respectively. Generally, they are assumed to be independent normally distributed Gaussian white noise with covariance matrices $\boldsymbol{Q} = \mathrm{diag}([Q_{11}, Q_{22}, Q_{33}, Q_{44}])$ and $\boldsymbol{R} = \mathrm{diag}([R_{11}, R_{22}, \cdots, R_{nn}])$, respectively. It is worth noting that the process noise and measurement noise here essentially reflect the confidence level of the model and measurement, respectively.

Based on the established state-transition and measurement equations, the estimation $\hat{\boldsymbol{g}}_k$ and $\hat{\boldsymbol{H}}_k$ of $\boldsymbol{g}_k$ and $\boldsymbol{H}_k$ are calculated by the EKF algorithm. Substituting these two estimations into eq. (15), the actual command input to the following UAV is obtained by

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k - \gamma \hat{\boldsymbol{H}}_k^{-1} \hat{\boldsymbol{g}}_k. \tag{15}$$

Remark 1 The number of measurement differences, namely the size of N, largely determines the EKF's tolerance to measurement noise. However, a too-large N may slow convergence. Therefore, it is value should fully consider the noise tolerance and convergence speed.

3.3 Convergence analysis

The convergence of the proposed Newton-Raphson method based ESC is analyzed in this subsection. Before proceeding to the formal analysis, the lemma on the EKF of the proposed design scheme is introduced.

Lemma 1 Consider the designed dynamical system eqs. (13) and (14)

$$\boldsymbol{s}_{k+1} = \boldsymbol{I}_k \boldsymbol{s}_k + \boldsymbol{\omega}_k, \tag{16a}$$

$$\Delta \boldsymbol{f}_{k+1} = \boldsymbol{M}_k \boldsymbol{s}_{k+1} + \boldsymbol{\nu}_{k+1}. \tag{16b}$$

If the system is uniformly completely observable and uniformly completely controllable, and if the initial value of error covariance matrix $\mathcal{P}_0 \geqslant 0$, then $\mathcal{P}_k$ is uniformly bounded for all $k \geqslant \mathcal{N}$.

Proof. The proof of the lemma was previously published in ref. [39] and is omitted here for brevity.

The convergence analysis will be stated below. To show convergence, the sequence $f(\boldsymbol{x}_k)$ will be proven to be a supermartingale. Furthermore, convergence analysis will also show that search behavior will not longer occur once the system reaches an area around the extremum of the performance function, which is partly determined by the error covariance bound of the EKF. The main results are formally presented in the following theorem.

Theorem 1 Consider the proposed Newton-Raphson method based ESC. The control system will converge to a neighbourhood of the extremum position, if $\boldsymbol{K}^{\mathrm{T}}\boldsymbol{H}_k\boldsymbol{K} + 2\boldsymbol{K}^{\mathrm{T}} < 0$ for $\boldsymbol{e}_k = 0$ or $\hat{\boldsymbol{g}}_k^{\mathrm{T}}(\boldsymbol{K}^{\mathrm{T}}\boldsymbol{H}_k\boldsymbol{K} + 2\boldsymbol{K}^{\mathrm{T}})\hat{\boldsymbol{g}} + \mathrm{tr}(\boldsymbol{H}_k\boldsymbol{K}\boldsymbol{P}_k\boldsymbol{K}^{\mathrm{T}}\boldsymbol{H}_k^{\mathrm{T}}) < 0$ for $\boldsymbol{e}_k \neq 0$, where $P_k = \mathrm{diag}([\mathcal{P}_{k_{11}}, \mathcal{P}_{k_{22}}])$ is the error covariance of the gradient estimation, K is the gain for gradient estimation, $\boldsymbol{e}_k$ is the estimation error for the gradient, $\mathrm{tr}(\cdot)$ represents the trace of matrix.

Proof. The proof is completed by validating that the sequence generated by $f(\boldsymbol{x}_k)$ is a supermartingale. Thus, the mathematical expectation of the extremum searching process needs to be proven to satisfy

$$\begin{aligned} &\bullet\ E[f(\boldsymbol{x}_k)] < \infty, \quad k \geqslant 0, \\ &\bullet\ E[f(\boldsymbol{x}_{k+1}) | \mathscr{F}_k] < f(\boldsymbol{x}_k), \quad k \geqslant 0, \end{aligned} \tag{17}$$

where $\mathscr{F}_k$ is the completion of the σ-field in probability space generated by the measurement history for $f(\boldsymbol{x})$ and $\boldsymbol{x}$ up to step k. The first item is guaranteed by definition of the performance function. The second item will be confirmed in the following.

First, only the gradient of the performance function estimated by the EKF is considered to generate the iterative scheme, namely

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \boldsymbol{K}\hat{\boldsymbol{g}}_k, \tag{18}$$

where $\hat{\boldsymbol{g}}_k$ is an estimation of the true value $\boldsymbol{g}_k$. The error between $\hat{\boldsymbol{g}}_k$ and $\boldsymbol{g}_k$ is defined as $\boldsymbol{e}_k = \boldsymbol{g}_k - \hat{\boldsymbol{g}}_k$. Then, the iterative scheme based on the true value of the gradient is given by

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \boldsymbol{K}(\hat{\boldsymbol{g}}_k + \boldsymbol{e}_k). \tag{19}$$

Furthermore, the iterative scheme assumes that $f(\boldsymbol{x}_k)$ is approximated by a quadratic function as defined in eq. (6)

$$f(\boldsymbol{x}) - f(\boldsymbol{x}_k) \approx \boldsymbol{g}_k^{\mathrm{T}}(\boldsymbol{x} - \boldsymbol{x}_k) + \frac{1}{2}(\boldsymbol{x} - \boldsymbol{x}_k)^{\mathrm{T}}\boldsymbol{H}_k(\boldsymbol{x} - \boldsymbol{x}_k). \tag{20}$$

By substituting $\boldsymbol{x} = \boldsymbol{x}_{k+1}$, one can derive

$$\begin{aligned} &f(\boldsymbol{x}_{k+1}) - f(\boldsymbol{x}_k) \\ =&(\hat{\boldsymbol{g}}_k + \boldsymbol{e}_k)^{\mathrm{T}}\boldsymbol{K}^{\mathrm{T}}\hat{\boldsymbol{g}}_k + \frac{1}{2}(\hat{\boldsymbol{g}}_k + \boldsymbol{e}_k)^{\mathrm{T}}\boldsymbol{K}^{\mathrm{T}}\boldsymbol{H}_k\boldsymbol{K}(\hat{\boldsymbol{g}}_k + \boldsymbol{e}_k) \\ =&\hat{\boldsymbol{g}}_k^{\mathrm{T}}\boldsymbol{K}^{\mathrm{T}}\hat{\boldsymbol{g}}_k + \boldsymbol{e}_k^{\mathrm{T}}\boldsymbol{K}^{\mathrm{T}}\hat{\boldsymbol{g}}_k + \frac{1}{2}(\hat{\boldsymbol{g}}_k + \boldsymbol{e}_k)^{\mathrm{T}}\boldsymbol{K}^{\mathrm{T}}\boldsymbol{H}_k\boldsymbol{K}(\hat{\boldsymbol{g}}_k + \boldsymbol{e}_k) \\ =&\hat{\boldsymbol{g}}_k^{\mathrm{T}}\boldsymbol{K}^{\mathrm{T}}\hat{\boldsymbol{g}}_k + \boldsymbol{e}_k^{\mathrm{T}}\boldsymbol{K}^{\mathrm{T}}\hat{\boldsymbol{g}}_k + \frac{1}{2}\hat{\boldsymbol{g}}_k^{\mathrm{T}}\boldsymbol{K}^{\mathrm{T}}\boldsymbol{H}_k\boldsymbol{K}\hat{\boldsymbol{g}}_k + \frac{1}{2}\boldsymbol{e}_k^{\mathrm{T}}\boldsymbol{K}^{\mathrm{T}}\boldsymbol{H}_k\boldsymbol{K}\boldsymbol{e}_k \\ =&\frac{1}{2}\hat{\boldsymbol{g}}_k^{\mathrm{T}}(\boldsymbol{K}^{\mathrm{T}}\boldsymbol{H}_k\boldsymbol{K} + 2\boldsymbol{K}^{\mathrm{T}})\hat{\boldsymbol{g}}_k + \frac{1}{2}\boldsymbol{e}_k^{\mathrm{T}}\boldsymbol{K}^{\mathrm{T}}\boldsymbol{H}_k\boldsymbol{K}\boldsymbol{e}_k + \boldsymbol{e}_k^{\mathrm{T}}\boldsymbol{K}^{\mathrm{T}}\hat{\boldsymbol{g}}_k. \end{aligned} \tag{21}$$

For the case of $\boldsymbol{e}_k = \boldsymbol{0}$, the control system will converge if

$$\boldsymbol{K}^{\mathrm{T}}\boldsymbol{H}_k\boldsymbol{K} + 2\boldsymbol{K}^{\mathrm{T}} < 0. \tag{22}$$

In this case, the condition for convergence does not involve $\hat{\boldsymbol{g}}_k$. When $\boldsymbol{e}_k \neq \boldsymbol{0}$, $\hat{\boldsymbol{g}}_k$ gets entangled in the condition of convergence. Calculating the conditional expectation yields

$$\begin{aligned} E[f(\boldsymbol{x}_{k+1}) - f(\boldsymbol{x}_k) | \mathscr{F}_k] =& \frac{1}{2}\hat{\boldsymbol{g}}_k^{\mathrm{T}}(\boldsymbol{K}^{\mathrm{T}}\boldsymbol{H}_k\boldsymbol{K} + 2\boldsymbol{K}^{\mathrm{T}})\hat{\boldsymbol{g}}_k \\ &+ \frac{1}{2}\mathrm{tr}(\boldsymbol{H}\boldsymbol{K}\boldsymbol{P}_k\boldsymbol{K}^{\mathrm{T}}\boldsymbol{H}^{\mathrm{T}}). \end{aligned} \tag{23}$$

According to the previous assumption of normal distribution for the error, one can gain $E[\boldsymbol{e}_k^{\mathrm{T}}\hat{\boldsymbol{g}}_k|\mathscr{F}_k]=0$ and $E[\boldsymbol{e}_k^{\mathrm{T}}\boldsymbol{e}_k|\mathscr{F}_k]=\boldsymbol{P}_k$. In the second case, therefore, the condition for convergence is

$$\hat{\boldsymbol{g}}_k^{\mathrm{T}}(\boldsymbol{K}^{\mathrm{T}}\boldsymbol{H}_k\boldsymbol{K}+2\boldsymbol{K}^{\mathrm{T}})\hat{\boldsymbol{g}}_k+\mathrm{tr}(\boldsymbol{H}_k\boldsymbol{K}\boldsymbol{P}_k\boldsymbol{K}^{\mathrm{T}}\boldsymbol{H}_k^{\mathrm{T}})<0. \tag{24}$$

This completes the proof.

Remark 2 The significance of the convergence condition for $\boldsymbol{e}_k\neq\boldsymbol{0}$ is not readily found in the above form. By assuming $\boldsymbol{H}_k$ is positive definite and setting the gain $\boldsymbol{K}=-\boldsymbol{H}_k^{-1}$, one can find

$$E[f(\boldsymbol{x}_{k+1})-f(\boldsymbol{x}_k)|\mathscr{F}_k]=\frac{1}{2}\mathrm{tr}[\boldsymbol{H}_k(-\hat{\boldsymbol{g}}_k\hat{\boldsymbol{g}}_k^{\mathrm{T}}+\boldsymbol{P}_k)\boldsymbol{H}_k^{\mathrm{T}}]. \tag{25}$$

This condition implies that the covariance of the gradient estimation must be greater than the covariance of the error for continuous convergence, which has a very intuitive sense. If this requirement is not met, the control system will not continue to converge in a direction. Meanwhile, this also indicates that the convergence boundary of the control system around the extremum is defined by $\boldsymbol{P}_k$.

4 Noise covariance optimization via the IPIO

The noise covariance matrices $\boldsymbol{Q}$ and $\boldsymbol{R}$ are the parameters of the EKF to be tuned based on the noise characteristics of the system. They are used to control the confidence level of the model and measurement. Their values directly influence dynamic performance and estimation accuracy of the EKF. However, obtaining optimal parameters is a challenging task due to the strong correlation among them, especially for large systems. To take into account all parameters together, their tuning is defined as a optimization problem. To address the potential multimodality, the IPIO algorithm is proposed to optimize the parameters.

4.1 Basic PIO algorithm

The basic PIO algorithm consists of two operators, namely the map and compass operator and the landmark operator. The combination of the two operators can enhance search capabilities for finding the global optimum. Assuming an optimization problem of D^{p}-dimension, there are N^{p} pigeons available to conduct the search. The position and velocity of the i-th pigeon are denoted by $\boldsymbol{X}_i^{\mathrm{p}}=[x_{i_1}^{\mathrm{p}},x_{i_2}^{\mathrm{p}},...,x_{i_{D^{\mathrm{p}}}}^{\mathrm{p}}]$ and $\boldsymbol{V}_i^{\mathrm{p}}=[v_{i_1}^{\mathrm{p}},v_{i_2}^{\mathrm{p}},...,v_{i_{D^{\mathrm{p}}}}^{\mathrm{p}}]$, respectively. Their values will be updated in each iteration.

Map and compass operator In the map and compass operator, the new position and velocity of the pigeon i at the t-th iteration are determined using the following formula

$$\boldsymbol{V}_i^{\mathrm{p}}(t)=\boldsymbol{V}_i^{\mathrm{p}}(t-1)\mathrm{e}^{-\lambda t}+\mathrm{rand}\cdot(\boldsymbol{X}_{\mathrm{gbest}}^{\mathrm{p}}-\boldsymbol{X}_i^{\mathrm{p}}(t-1)), \tag{26a}$$

$$\boldsymbol{X}_i^{\mathrm{p}}(t)=\boldsymbol{X}_i^{\mathrm{p}}(t-1)+\boldsymbol{V}_i^{\mathrm{p}}(t), \tag{26b}$$

where superscript p denotes the abbreviations of the pigeon, λ is the map and compass factor; rand represents a random number with rand$\in[0,1]$; $\boldsymbol{X}_{\mathrm{gbest}}^{\mathrm{p}}$ denotes the current global optimal position, which can be acquired by evaluating the fitness values of all pigeon positions.

Landmark operator In the landmark operator, the pigeon with the poor fitness value will be discarded. This makes the number of pigeons decrease by a half during each iteration and accelerate the convergence. The position center of the remaining pigeons is considered as the global optimal position. Therefore, the remaining pigeons will update their positions by moving towards this position center, which can be formulated as

$$N^{\mathrm{p}}(t)=\frac{N^{\mathrm{p}}(t-1)}{2},$$

$$\boldsymbol{X}_c^{\mathrm{p}}(t)=\frac{\sum_{i=1}^{N^{\mathrm{p}}(t)}\boldsymbol{X}_i^{\mathrm{p}}(t-1)\mathrm{fit}(\boldsymbol{X}_i^{\mathrm{p}}(t-1))}{\sum_{i=1}^{N^{\mathrm{p}}(t)}\mathrm{fit}(\boldsymbol{X}_i^{\mathrm{p}}(t-1))}, \tag{27a}$$

$$\boldsymbol{X}_i^{\mathrm{p}}(t)=\boldsymbol{X}_i^{\mathrm{p}}(t-1)+\mathrm{rand}\cdot(\boldsymbol{X}_c^{\mathrm{p}}(t)-\boldsymbol{X}_i^{\mathrm{p}}(t-1)), \tag{27b}$$

where $\boldsymbol{X}_c^{\mathrm{p}}$ is the position center of remaining pigeons. It should be noted that $\mathrm{fit}(\boldsymbol{X}_i^{\mathrm{p}}(t-1))=1/\mathrm{fit}(\boldsymbol{X}_i^{\mathrm{p}}(t-1))$, where $\mathrm{fit}(\cdot)$ denotes the fitness function.

4.2 Improvement strategy

In the map and compass operator of the basic PIO, the map and compass factor λ is responsible for maintaining a balance between global exploration and local exploitation. If it is relatively large, it will strengthen the global exploration, whereas a relatively small value can enhance the local exploitation. However, a constant λ is always employed in the map and compass operator of the basic PIO. As a result, the contribution of λ to balancing global exploration and local exploitation has not been fully realized. Therefore, the sine-powered chaos is utilized to dynamically adjust it with iteration. This adjustment strategy can be written as

$$\begin{cases}\lambda(t)=\tau(t)\cdot\sin(\pi\lambda(t-1)),\\ \tau(t)=\tau_{\max}-t\cdot(\tau_{\max}-\tau_{\min})/t_{\max},\end{cases} \tag{28}$$

where τ is the control parameter defined within the range of $(0,1)$; $\tau_{\max}$ and $\tau_{\min}$ are predetermined as the upper and lower bounds, respectively; $t_{\max}$ denotes the maximum number of iterations. The chaotic characteristic of λ is gradually diminished as the iterations proceed. This implies that the dynamic adjustment strategy maintains the benefits of both traversal and randomness of the chaotic weight during the initial iterations, which encourages the pigeons to perform a extensive global exploration. In the subsequent iterations, the pigeons

converge towards the global optimal position and the reduced flight velocity allows for a more fine search.

Moreover, in the map and compass operator of the basic PIO, the new positions of pigeons are calculated only by adding the previous position $\boldsymbol{X}_i^{\mathrm{p}}(t-1)$ and the current velocity $\boldsymbol{V}_i^{\mathrm{p}}(t)$. This may lead to a conservative balance between global exploration and local exploitation only although dynamically adjusting λ. Taking the relationship between position and velocity into consideration [40], the improved position update formula is given as follows:

$$\begin{aligned}\boldsymbol{X}_i^{\mathrm{p}}(t) =&\omega_i^{\mathrm{p}}(t-1)\cdot\boldsymbol{X}_i^{\mathrm{p}}(t-1)+(1-\omega_i^{\mathrm{p}}(t-1))\cdot\boldsymbol{V}_i^{\mathrm{p}}(t)\\&+\mathrm{rand}\cdot\sigma_i^{\mathrm{p}}(t-1)\cdot\boldsymbol{X}_{\mathrm{gbest}}^{\mathrm{p}},\end{aligned}\tag{29}$$

where ω_i^{p} represents the dynamic weights associated with both position and velocity, σ_i^{p} is the acceleration coefficient. They are calculated by $\omega_i^{\mathrm{p}}(t-1)=\sigma_i^{\mathrm{p}}(t-1)=\frac{\mathrm{e}^{\mathrm{fit}(\boldsymbol{X}_i^{\mathrm{p}}(t-1))/\overline{\mathrm{fit}}(t-1)}}{(1+\mathrm{e}^{-\mathrm{fit}(\boldsymbol{X}_i^{\mathrm{p}}(t-1))/\overline{\mathrm{fit}}(t-1)})^{t-1}}$, where $\overline{\mathrm{fit}}(t-1)$ denotes the mean fitness value of all pigeons at the iteration of $(t-1)$. It should be noted that the basic PIO improved by the proposed two improvement strategies is referred to as IPIO for brevity.

Remark 3 In the landmark operator of the basic PIO, all pigeons are directed towards the position center, resulting in a clear movement direction for each individual pigeon. Therefore, there is little potential room for improvement according to the functioning mechanism of the landmark operator. Given this fact, the IPIO will continue to utilize the landmark operator of the basic PIO.

4.3 Fitness function

During the parameter tuning process, the optimization objective is to minimize the state estimation error covariance. Furthermore, different weights are assigned to the estimation errors based on the importance of the estimated states. Hence, the fitness function is formulated as

$$\begin{aligned}J(N_k)=&\sum_{k=1}^{N_k}(\omega_{g_1}P_{11}(k)+\omega_{g_2}P_{22}(k)\\&+\omega_{H_{11}}P_{33}(k)+\omega_{H_{22}}P_{44}(k))\Delta h,\end{aligned}\tag{30}$$

where P_{11}, P_{22}, P_{33} and P_{44} are the state estimation error covariance; ω_{g_1}, ω_{g_2}, $\omega_{H_{11}}$ and $\omega_{H_{22}}$ are the weight coefficients; Δh is the sample time; N_k is the number of samples that can be calculated by $N_k=t/\Delta h$ with t being running time of simulation system. Note that the formulated fitness function contains two essential factors: "time" and "error", which enable the evaluation of both the dynamic and steady-state performance of the EKF. This ensures the the response speed and the steady-state accuracy of the proposed Newton-Raphson method based ESC simultaneously. Furthermore, the formulated fitness function emphasizes accumulation of error covariance throughout the entire estimation process, which can effectively mitigate the impact of the large initial error covariance.

4.4 Optimization procedure

The optimization procedure of the IPIO for the EKF mainly includes two parts: Initialization of position and velocity and execution of iterative operators. The purpose of initialization is to increase the diversity of the pigeons and avoid premature convergence. The iterative operators are used to perform the global exploration and local exploitation. The detailed optimization procedure is summarized as follows.

Step 1 Input the parameters of the IPIO, define the search space and set the boundaries.

Step 2 Initialize the state of each pigeon randomly in the search space, including the position and its associated velocity.

Step 3 Drive ESC simulation system using the parameters from Step 2 and calculate the fitness value. Evaluate the results of all pigeons and find the current global optimal value.

Step 4 Execute the iteration by the map and compass operator and the landmark operator of the IPIO independently.

Step 5 Check the stop criterion. If the maximum iterations is reached, output the global optimal value and stop the algorithm. Otherwise, go to Step 4.

Furthermore, the pseudocode of the above steps is given in Algorithm 1. It provides an instructional frame on how these steps can be implemented.

5 Simulation results and analysis

In this section, the simulation verification is carried out for a leading-following UAV close formation based on the F-16 aircraft model developed by ref. [41]. This will serve two main purposes. The first is to analyze the location of the optimal relative position for the following UAV in close formation flight, which is accomplished by examining the strength variations of the wake vortex with the relative position to the leading UAV. The second purpose is to validate the convergence performance of the proposed Newton-Raphson method based ESC optimized via IPIO. To demonstrate the advantages, the proposed design is compared with the traditional ESC based on modulation and demodulation [25] and the class ESC based on the EKF [26]. For the sake of fairness, the parameters of these two comparison controllers are also obtained by using the IPIO.

Algorithm 1 IPIO employed by the ESC

▷ % Search space of parameters is defined by $\boldsymbol{Q}_{\min}$, $\boldsymbol{Q}_{\max}$ and $\boldsymbol{R}_{\min}$, $\boldsymbol{R}_{\max}$ %
Input: D^{p}, N^{p}, $\tau_{\max}$, $\tau_{\min}$, $t_{1\max}$, $t_{2\max}$, $\boldsymbol{Q}_{\min}$, $\boldsymbol{Q}_{\max}$, $\boldsymbol{R}_{\min}$, $\boldsymbol{R}_{\max}$
Output: $\boldsymbol{X}^{\mathrm{p}}_{\mathrm{gbest}}$ ▷ % Store optimal parameters %

Initialize $\boldsymbol{X}^{\mathrm{p}}$ and $\boldsymbol{V}^{\mathrm{p}}$; ▷ % Random distribution %
$t \leftarrow 0$;
for $i = 1 \rightarrow N^{\mathrm{p}}$ **do**
  ▷ % $p(i)$=Simulink.SimulationInput(ESC simulation system)%
  Create ESC simulation system object;
  ▷ % $p(i)$=$p(i)$.setVariable(X_i^{p})%
  Substitute $\boldsymbol{X}_i^{\mathrm{p}}$ into the established object;
end
▷ % parpool('local') and outP=parsim(p) %
Run ESC simulation system objects and calculate J;
Evaluate $\boldsymbol{X}^{\mathrm{p}}$ and record $\boldsymbol{X}^{\mathrm{p}}_{\mathrm{gbest}}$;
▷ % Map and compass operator %
while $t \le t_{1\max}$ **do**
  ▷ % First improvement measure %
  Adjust λ dynamically by eq. (28);
  for $i = 1 \rightarrow N^{\mathrm{p}}$ **do**
    Calculate ω_i^{p} and σ_i^{p};
    ▷ % Second improvement measure %
    Update $\boldsymbol{V}_i^{\mathrm{p}}$ and $\boldsymbol{X}_i^{\mathrm{p}}$ by eqs. (26) and (29);
    Create ESC simulation system object;
    Substitute $\boldsymbol{X}_i^{\mathrm{p}}$ into the established object;
  end
  Run ESC simulation system objects and calculate J;
  Evaluate $\boldsymbol{X}^{\mathrm{p}}$ and Update $\boldsymbol{X}^{\mathrm{p}}_{\mathrm{gbest}}$;
  $t \leftarrow t + 1$; ▷ % Iterations plus one %
end
▷ % Landmark operator %
while $t \le (t_{1\max} + t_{2\max})$ **do**
  Sort $\boldsymbol{X}^{\mathrm{p}}$ according to J;
  Eliminate pigeons with poor fitness values;
  Calculate $\boldsymbol{X}_c^{\mathrm{p}}$ by eq. (27);
  for $i = 1 \rightarrow N^{\mathrm{p}}$ **do**
    ▷ % All pigeons fly towards $X_c{}^p$ %
    Update $\boldsymbol{X}_i^{\mathrm{p}}$ by eq. (27);
    Create ESC simulation system object;
    Substitute $\boldsymbol{X}_i^{\mathrm{p}}$ into the model object;
  end
  Run ESC simulation system objects and calculate J;
  Evaluate $\boldsymbol{X}^{\mathrm{p}}$ and Update $\boldsymbol{X}^{\mathrm{p}}_{\mathrm{gbset}}$;
  $t \leftarrow t + 1$;
end

In the simulation, the leading UAV is assumed to be always under a level and straight stable flight with the velocity of 152 m/s at the attack angles of 4.5°. At the beginning of close formation, the flight velocity and attitude of the following UAV are considered to be the same as those of the the leading UAV. Under the position command generated by the proposed ESC scheme, the following UAV flies towards the optimal relative position autonomously using the autopilot. Note that the autopilot is assumed to be able to accurately track the given the position command. Therefore, the influence of the position tracking error is not considered when analyzing the convergence performance of the extremum seeking controller.

5.1 Location analysis of the optimal relative position

The magnitude of the free-airflow velocity is equal to that of the leading UAV velocity without considering the wind speed, i.e., $V_\infty = 152$ m/s. Other parameters required for modeling the wake vortex are set to $C_L = 0.2134$, $C_{L_\alpha} = 0.0489$, $\mu = 1.6$, $e_0 = 0.95$, $b = 9.144$ m, $S = 27.8709$ m^2.

According to the continuous horseshoe vortex method, the strength of the wake vortex can be characterized as a positional function relative to the leading UAV. The variation of the dimensionless induced wake velocity field with the longitudinal position is illustrated in Figure 5. It is evident that the magnitude of the induced wake velocity field experiences a notable reduction as the longitudinal position undergoes negative increments. Furthermore, the intensity diminishes almost to zero when $x < -10b$, which indicates the complete disappearance of the wake vortex effect. Therefore, the downstream longitudinal relative position should not exceed $10b$ if the following UAV intends to use the wake vortex effect of the leading UAV to decrease the drag. The location of the optimal relative position relative is denoted as (x_S, y_S, z_S). Considering the minimum safe separation between the UAVs and maximizing the wake vortex effect, the longitudinal relative location is set to $x_S = -3b$. For this reason, the variation of the induced lift and drag coefficient of the following UAV with the lateral and vertical separation relative to the leading UAV is investigated at $x_S = -3b$. The results are demonstrated in Figure 6. The points of maximum induced lift and minimum induced drag coefficient can be clearly observed

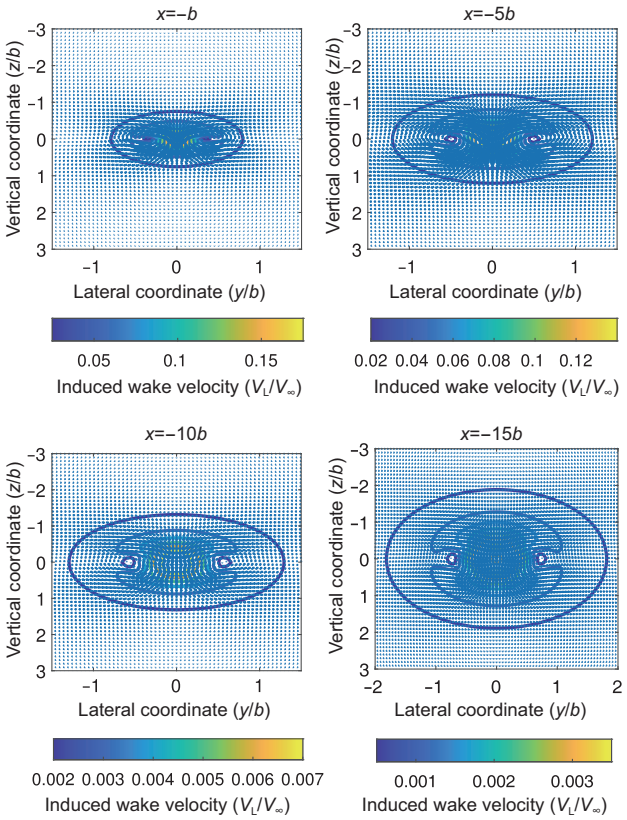

Figure 5 Sectional views of dimensionless velocity field induced by the wake vortex at different longitudinal positions.

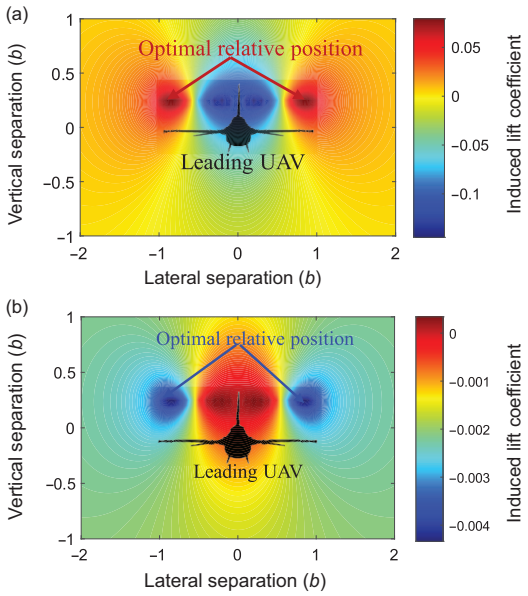

Figure 6 Variation of the induced coefficient with lateral and vertical position ($x = -3b$).

and appear at the same location. This just validates the relevant conclusion that the increase of induced lift promotes the decrease of the induced drag [29]. Therefore, the lateral and vertical relative location of the optimal relative position are set to $y_S = \pm 0.85b$ and $z_S = 0.25b$, respectively. The extremum seeking controller is required to find the optimal relative position location and converge to the optimal relative position eventually. Note that it can also be seen from the figure that the location of the optimal relative position is very close to the area where the induced drag coefficient increases, which indicates that the controller may transfer to this area if it can not converge stably near the optimal relative position. This will make UAV close formation lose the original significance of reducing drag.

5.2 Parameters optimization of the EKF using IPIO

To balance noise tolerance and convergence speed, the difference number of the measurement matrix is set to $N = 10$, i.e., the measurement matrix is composed of 10 time delay row vectors. The step size is set to $\gamma = 0.0004$. The extremum seeking controller start at $(y, z) = (2.4b, 0.6b)$ and $(y, z) = (2.6b, 0.7b)$ to verify the robustness of optimization algorithms to the initial values. Through difference calculation of the performance function, the initial states of the EKF is approximately $\boldsymbol{s}_0 = [0.2624, 0.1368, -0.0342, 0.0327]^{\mathrm{T}} \times 10^{-4}$ with the initial covariance matrix of $10^{-6} \times \boldsymbol{I}_4$. The perturbation signals are set as $0.003b\sin(30t)$ and $0.003b\cos(30t)$. The weight coefficients of the fitness function are set to $\omega_{g_1} = \omega_{g_2} = 100$ and $\omega_{H_{11}} = \omega_{H_{22}} = 1$. The search space is determined as the set of pairs $(\boldsymbol{Q}, \boldsymbol{R})$, subject to $\boldsymbol{Q} \in [1, 10] \times [10^{-4}, 10^{-0}]$ and $\boldsymbol{R} \in [1, 10] \times [10^{-9}, 10^{-7}]$. Note that this is also the domain that can make the extremum search controller get stability. To verify the competence of the IPIO, the genetic algorithm (GA) [35], the basic PIO [36] and particle swarm optimization (PSO) [33] are introduced to serve as the comparison algorithms. The maximum number of iteration for all algorithms is set to 15 with the population size of 32. Other parameters are listed in Table 1. Each optimization algorithm was run repeatedly for 10 times and the optimization result with the smallest fitness value is selected to represent the optimization performance.

From the output curve of the performance function, the convergence process of the the proposed ESC by using the IPIO is reflected in Figure 7. It can be clearly reflected from $t = 0$ that the pigeons are randomly distributed in the search space and the positions of most pigeons are worse. With the iteration, most pigeons converge towards the global optimal position. When $t = 8$, the IPIO basically stops global exploration and starts local exploitation. However, an unexpected result has occurred at $t = 12$. Compared with other pigeons' position, the convergence speed of the performance function at the global optimal position is slowest. This may be caused by the high measurement confidence level of the EKF. On the other hand, if the pigeons make the state-transition model get high confidence level, this will accelerate the convergence but also lead to large estimation error. From $t = 14$, while the SCPIO performs local exploitation, a small number of pigeons jump out of the global optimal position to avoid falling into local optimum. Through the continued search of the landmark operator, the global optimal position further evolves to a lower fitness value as demonstrated in $t = 20$. As previously analyzed, the optimization algorithm is required to find the optimal balance of the confidence level between the state-transition model and measurement of the EKF so as to accelerate the convergence and reduce the fitness value.

The comparison results of four optimization algorithms are shown in Figure 8. From the optimization results, the GA has

Table 1 Parameters of each optimization algorithm

Algorithm	Parameter	Description	Value
GA [35]	p_c	Crossover probability	0.9
	p_m	Mutation probability	0.1
PSO [33]	ω	Inertia weight	0.4
	c_1	Self learning factor	2
	c_2	Group learning factor	2
PIO [36] SCPIO	$t_{1_{\max}}$	Maximum iteration of operator 1	12
	$t_{2_{\max}}$	Maximum iteration of operator 2	3
	λ	Map and compass factor	0.4
	$[\tau_{\min}, \tau_{\max}]$	Control parameter	[0.4, 0.6]

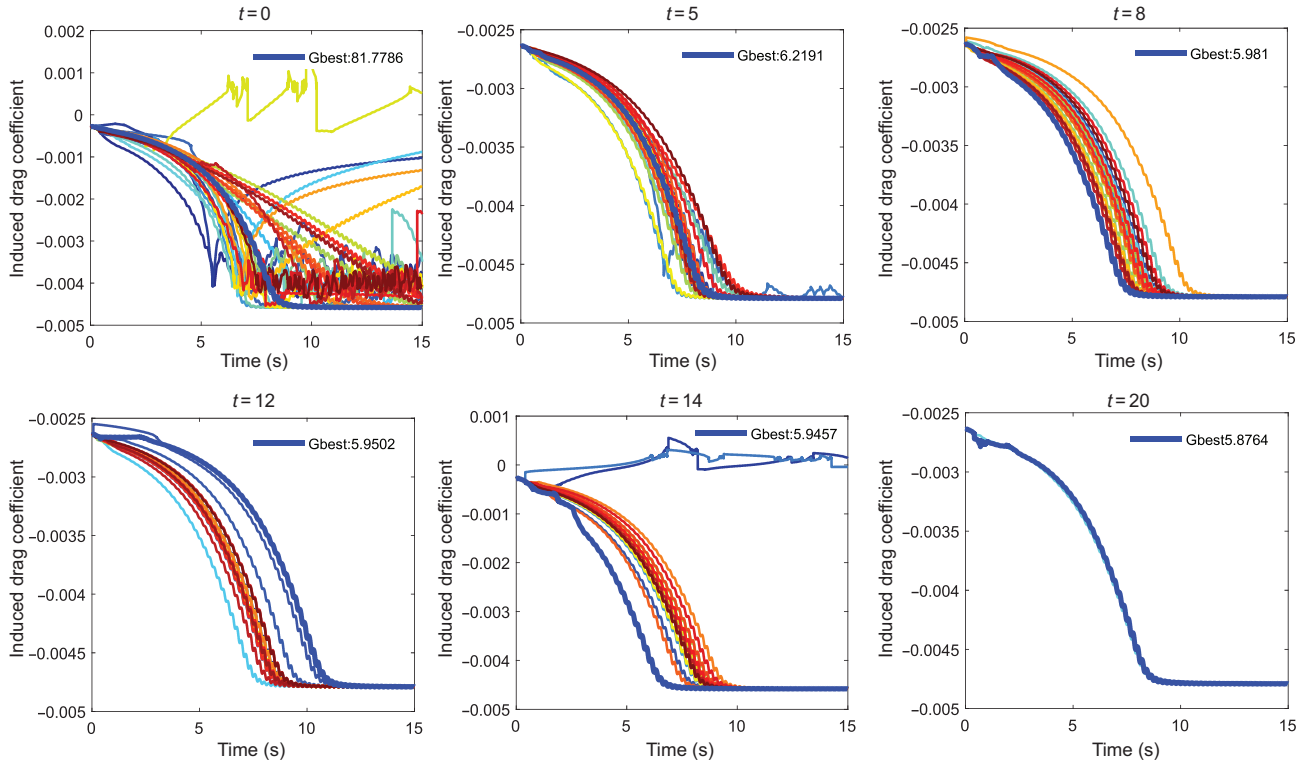

Figure 7 Evolutionary process of the global optimal position of the IPIO reflected by the output curve of the performance function.

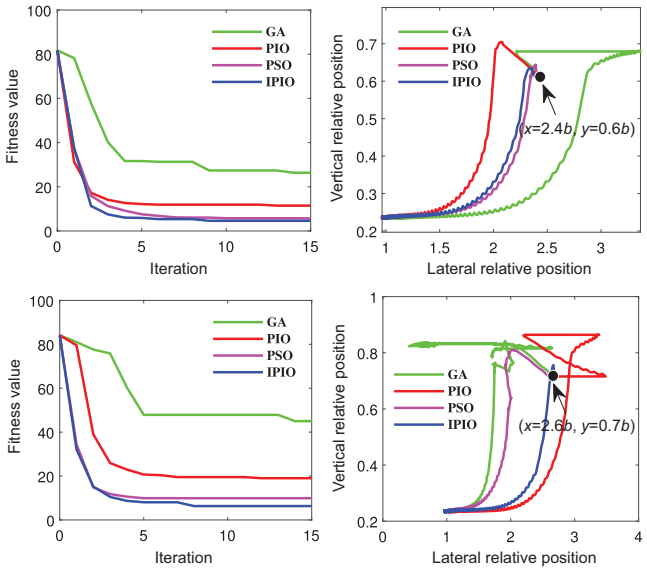

Figure 8 Comparison of four optimization algorithms for the different initial position ($x = 2.4b, y = 0.6b$ and $x = 2.6b, y = 0.7b$).

the weakest optimization ability to tuning the parameters of the EKF in the proposed ESC. This implies that the GA may have weak capabilities in handling multi-modal optimization problems. Compared to the PSO, the basic PIO exhibits a relatively lower convergence accuracy but faster convergence speed. This is also the reason why the PIO algorithm was chosen for improvement and subsequent utilization in this paper. From the trajectory curves of the following UAV, furthermore, it can be clearly seen that the proposed IPIO algorithm has good robustness to the initial values. This is mainly attributed to the introduced improvements, which effectively balance global exploration and local exploitation.

5.3 Validation of convergence performance

The convergence speed and accuracy are important aspects in evaluating the control performance of a extremum seeking controller. A total of three extremum seeking controllers are used for comparison, i.e., the traditional ESC based on modulation and demodulation, the class ESC based on the EKF and the proposed Newton-Raphson method based ESC. The high-pass filter cutoff frequency and integral gain of the traditional ESC are 5 rad/s and 10^6, respectively. The integral gain of the class ESC is 3×10^4 and other parameters are identical with those of the proposed ESC. Note that the reason why these two controllers have high gain is that the amplitude change of the performance function is smaller after the system is perturbed. It should be noted that the class ESC uses EKF to only estimate the gradient of the performance function and the command is obtained by integrating the gradient. The proposed ESC not only performs gradient estimation but also estimates the Hessian matrix and the command

is iteratively acquired by Newton-Raphson method.

The convergence performance of three extremum seeking controllers is demonstrated in Figure 9. It can be seen that three extremum seeking controllers have the identical convergence accuracy. However, there is a significant difference in terms of convergence speed. The traditional ESC has the slowest convergence speed. This may be caused by the continuous time integration of the small demodulation signal. Increasing the integral gain will accelerate convergence, but this can lead to a large steady-state oscillation. Compared with the class ESC based on the EKF, the proposed ESC produces a slightly slower convergence. This indicates that the introduction of Hessian matrix will increase the confidence level of measurement, which reduces the convergence speed. Further observation, the forms of output curves under the extremum seeking controllers by using the traditional and class EKF are almost identical. If the output lateral and vertical separation in the form of approximate step are used as the reference command of the following UAV, this may cause severe oscillation to the underdamped UAV system. On the contrary, the output form of the proposed extremum seeking controllers is close to the quadratic function curve owing to the introduction of Hessian matrix. This form of reference command is conducive to non-overshoot tracking of the following UAV.

The estimation output and error variance of the proposed ESC and the classic ESC are illustrated in Figures 10 and 11, respectively. The classic ESC generates a transient pulse at around $t = 5.5$ s for the gradient estimation, which is also the reason why the output curves of the

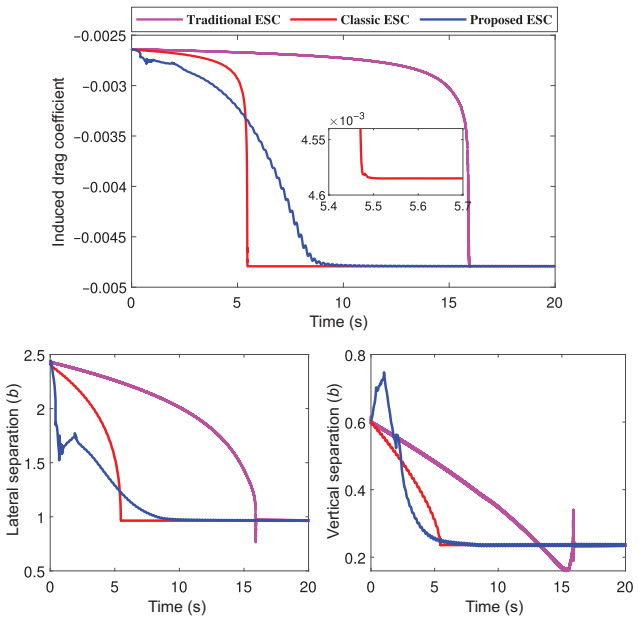

Figure 9 Convergence performance of three extremum seeking controllers.

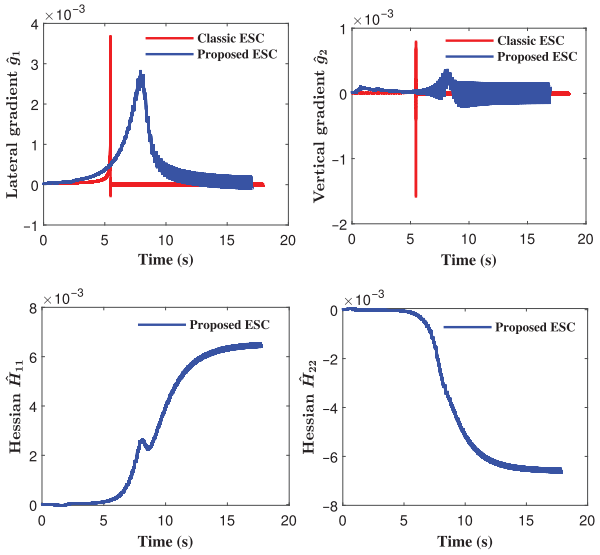

Figure 10 Estimation output of the proposed ESC and the classic ESC.

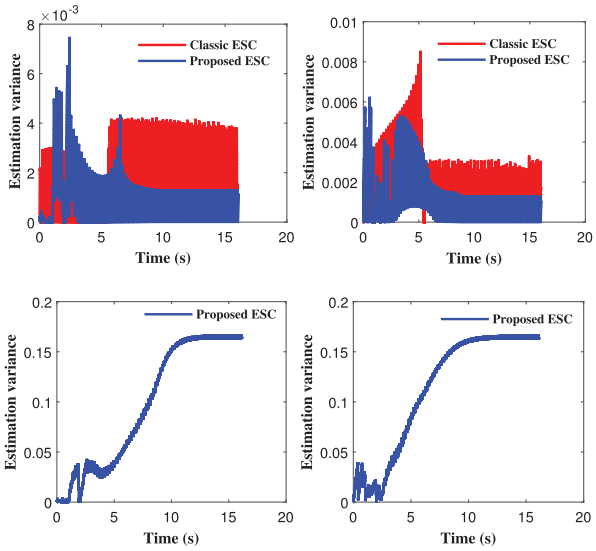

Figure 11 Estimation variance of the proposed ESC and the classic ESC.

extremum seeking controller are approximate to the step form. Because of the the addition of Hessian matrix estimation, the gradient estimation of the proposed ESC does not change dramatically near the convergence position. According to the error variance of estimation states, one can conclude that the proposed ESC has higher estimation accuracy than the classical ESC. However, the proposed ESC shows a large estimation error for the Hessian matrix, which leads to a lower confidence level for the state transition. Conversely, this further confirms that the introduction of Hessian matrix will increase the confidence level of measurement.

To further explore the convergence behavior of the proposed extremum seeking controller, a Monte Carlo analysis

was performed where the initial position of the following UAV was randomly selected for a three-UAV close formation. The following UAV 1 is initialized to the left of the leading UAV and the following UAV 2 is placed on the right. The UAV 1 and UAV 2 both fly in the wake vortex of the leading UAV that is flying straight and level. Their flight paths are depicted in Figure 12. Obviously, the convergence behavior of the proposed extremum seeking controller is independent of the initial position of the following UAV.

6 Conclusion

In this paper, a comprehensive design scheme for the ESC of UAV close formation has been developed to achieve maximum drag reduction. The method mainly includes four aspects of design. To form a feedback, a continuous horseshoe vortex method with high accuracy is employed to model the wake vortex. The Newton-Raphson iterative method is designed as the output feedback of ESC. The EKF is used to estimate the gradient and the Hessian matrix of the parameterized performance function. The IPIO is applied to automatically tune the noise covariance matrix of the EKF. Simulation results indicate that the proposed scheme allows a faster convergence of relative position extremum. Furthermore, the proposed scheme can provide a smoother command during the seeking process. Finally, Monte Carlo analysis for a three-UAV close formation shows that the proposed design scheme is robust to the initial position of the following UAV. It is necessary to clarify that the proposed Newton-Raphson method based ESC was validated based on the assumption of an ideal autopilot. Therefore, the current simulation validation may not fully consider the physical characteristics and constraints of real fixed-wing UAVs, which could lead to the potential limitations in the convergence performance of the controller. Due to the nonlinearity and complexity deriving from the physical characteristics and constraints, the extreme seeking process may require a larger number of iterative steps to achieve convergence. Furthermore, the lack of consideration for the quality of sensor data may result in error accumulation during the search process. Consequently, the controller may fail to converge precisely to the global optimal position. In future work, we aim to incorporate the physical characteristics and constraints of real fixed-wing UAVs into the design of close formation ESC. The aerodynamic characteristics, dynamic properties, and inertia coupling, etc., will be carefully taken into consideration. Moreover, the influence of sensor errors on the design of the ESC will be further explored.

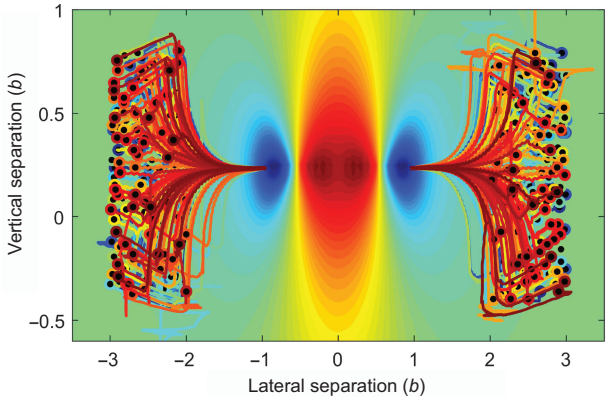

Figure 12 Convergence behavior resulting from a set of random initial positions.

This work was supported by the National Natural Science Foundation of China (Grant Nos. 91948204, U20B2071, T2121003 and U1913602), and Open Fund/Postdoctoral Fund of the Laboratory of Cognition and Decision Intelligence for Complex Systems, Institute of Automation, Chinese Academy of Sciences (Grant No. CASIA-KFKT-08).

1 Zhang Q, Liu H H T. Robust nonlinear close formation control of multiple fixed-wing aircraft. J Guidance Control Dyn, 2021, 44: 572–586
2 DeVries L, Paley D A. Wake sensing and estimation for control of autonomous aircraft in formation flight. J Guidance Control Dyn, 2016, 39: 32–41
3 Wang Y, Wang D. Tight formation control of multiple unmanned aerial vehicles through an adaptive control method. Sci China Inf Sci, 2017, 60: 070207
4 Fan Y M, Qiu H X, Duan H B. Unmanned aerial vehicle close formation cooperative control based on predatory escaping pigeon-inspired optimization (in Chinese). Sci Sin Tech, 2015, 45: 559–572
5 Luo Y, Bai A, Zhang H. Distributed formation control of UAVs for circumnavigating a moving target in three-dimensional space. Guid Navigat Control, 2021, 01: 2150014
6 Bai T T, Wang D B, Masood R J. Formation control of quad-rotor UAV via PIO. Sci China Tech Sci, 2022, 65: 432–439
7 Mirzaeinia A, Hassanalian M, Lee K, et al. Energy conservation of V-shaped swarming fixed-wing drones through position reconfiguration. Aerospace Sci Tech, 2019, 94: 105398
8 Halaas D J, Bieniawski S R, Whitehead B, et al. Formation flight for aerodynamic benefit simulation development and validation. In: Proceedings of the 52nd Aerospace Sciences Meeting. National Harbor: AIAA, 2014
9 Kent T E, Richards A G. Analytic approach to optimal routing for commercial formation flight. J Guidance Control Dyn, 2015, 38: 1872–1884
10 Bangash Z A, Sanchez R P, Ahmed A, et al. Aerodynamics of formation flight. J Aircraft, 2006, 43: 907–912
11 Hanson C E, Pahle J, Reynolds J R, et al. Experimental measurements of fuel savings during aircraft wake surfing. In: Proceedings of the 2018 Atmospheric Flight Mechanics Conference. Atlanta: AIAA, 2018
12 Zhang Q, Liu H H T. Aerodynamics modeling and analysis of close formation flight. J Aircraft, 2017, 54: 2192–2204
13 Ransquin I, Caprace D G, Duponcheel M, et al. Wake vortex detection and tracking for aircraft formation flight. J Guidance Control Dyn, 2021, 44: 2225–2243
14 Vechtel D, Fischenberg D, Schwithal J. Flight dynamics simulation of formation flight for energy saving using LES-generated wake flow fields. CEAS Aeronaut J, 2018, 9: 735–746
15 Shan J, Liu H T. Close-formation flight control with motion synchronization. J Guidance Control Dyn, 2005, 28: 1316–1320

16 Zhai S, Li C, Wang C, et al. Vertically optimal close formation flight control based on wingtip vortex structure. J Aircraft, 2020, 57: 964–973
17 Chao H, Gu Y, Tian P, et al. Wake vortex detection with UAV close formation flight. In: Proceedings of the AIAA Atmospheric Flight Mechanics Conference. Dallas: AIAA, 2015
18 Irigireddy A S C R, Moncayo H. Vision based relative navigation for close-formation flight missions. In: Proceedings of the AIAA Scitech 2020 Forum. Orlando: AIAA, 2020
19 Wang Y S, Chen Z Q, Sun M W, et al. Design and stability analysis of a generalized reduced-order active disturbance rejection controller. Sci China Tech Sci, 2022, 65: 361–374
20 Liu C, Jiang B, Zhang K. Adaptive fault-tolerant H-infinity output feedback control for lead-wing close formation flight. IEEE Trans Syst Man Cybern Syst, 2018, 50: 2804–2814
21 Zhang Q, Liu H H. Integrator-augmented robust adaptive control design for close formation flight. In: Proceedings of the AIAA Guidance, Navigation, and Control Conference. Grapevine: AIAA, 2017
22 Zhang Q, Liu H H T. UDE-based robust command filtered backstepping control for close formation flight. IEEE Trans Ind Electron, 2018, 65: 8818–8827
23 Lei Y F, Li J M, Zhao A L. Adaptive neural networks control for uncertain parabolic distributed parameter systems with nonlinear periodic time-varying parameter. Sci China Tech Sci, 2022, 65: 1482–1492
24 Han H G, Wang T, Sun H Y, et al. Fuzzy super-twisting sliding mode control for municipal wastewater nitrification process. Sci China Tech Sci, 2022, 65: 2420–2428
25 Binetti P, Ariyur K B, Krstic M, et al. Formation flight optimization using extremum seeking feedback. J Guidance Control Dyn, 2003, 26: 132–142
26 Brodecki M, Subbarao K. Autonomous formation flight control system using in-flight sweet-spot estimation. J Guidance Control Dyn, 2015, 38: 1083–1096
27 Huang Y, Zhang Y, Xu B, et al. A new adaptive extended kalman filter for cooperative localization. IEEE Trans Aerosp Electron Syst, 2018, 54: 353–368
28 Calise A J. Enforcing an algebraic constraint in extended kalman filter design. J Guidance Control Dyn, 2017, 40: 2229–2236
29 Yuan G, Xia J, Duan H. A continuous modeling method via improved pigeon-inspired optimization for wake vortices in UAVs close formation flight. Aerospace Sci Tech, 2022, 120: 107259
30 He L, Hu M K, Wei Y J, et al. State of charge estimation by finite difference extended Kalman filter with HPPC parameters identification. Sci China Tech Sci, 2020, 63: 410–421
31 Zhu K, Han B, Zhang T. Multi-UAV distributed collaborative coverage for target search using heuristic strategy. Guid Navigat Control, 2021, 01: 2150002
32 Tong B, Wei C, Shi Y. Fractional order darwinian pigeon-inspired optimization for multi-UAV swarm controller. Guid Navigat Control, 2022, 02: 2250010
33 Kennedy J, Eberhart R. Particle swarm optimization. In: Proceedings of the ICNN'95-international conference on neural networks. Perth: IEEE, 1995
34 Li Z X, Shi X L, Cao J D, et al. CPSO-XGBoost segmented regression model for asphalt pavement deflection basin area prediction. Sci China Tech Sci, 2022, 65: 1470–1481
35 Lambora A, Gupta K, Chopra K. Genetic algorithm-A literature review. In: Proceedings of the 2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon). Faridabad: IEEE, 2019
36 Duan H, Qiao P. Pigeon-inspired optimization: a new swarm intelligence optimizer for air robot path planning. Int J Intelligent Computing Cybernetics, 2014, 7: 24–37
37 Liu C B, Ma Y H, Yin H, et al. Human resource allocation for multiple scientific research projects via improved pigeon-inspired optimization algorithm. Sci China Tech Sci, 2021, 64: 139–147
38 Yu Y P, Liu J C, Wei C. Hawk and pigeon's intelligence for UAV swarm dynamic combat game via competitive learning pigeon-inspired optimization. Sci China Tech Sci, 2022, 65: 1072–1086
39 Krener A J. The convergence of the extended Kalman filter. In: Directions in Mathematical Systems Theory and Optimization. Berlin: Springer, 2003. 173–182
40 Chen K, Zhou F, Liu A. Chaotic dynamic weight particle swarm optimization for numerical function optimization. Knowledge-Based Syst, 2018, 139: 23–40
41 Sonneveldt L. Nonlinear F-16 model description. Technical Report. Control & Simulation Division, Faculty of Aerospace Engineering, Delft University of Technology, 2006