



# Stochastic dynamics of aircraft ground taxiing via improved physics-informed neural networks

Ying Zhang · Zhengrong Jin · Long Wang ·  
Kaixin Zheng · Wantao Jia 

Received: 12 October 2022 / Accepted: 23 October 2023 / Published online: 9 January 2024  
© The Author(s), under exclusive licence to Springer Nature B.V. 2024

**Abstract** In this paper, the stochastic propagation of the aircraft taxiing under the excitation of uneven runway is investigated based on physics-informed neural networks (PINNs). In particular, we successfully applied the PINNs with layer-wise locally adaptive activation functions (L-LAAF) and the learning rate decay strategy to address the challenging task of parameter identification for some aircraft systems. Specifically, the accuracy and effectiveness of the proposed method in solving the time-dependent Fokker–Planck equation for systems were first demonstrated. Subsequently, the proposed method is effectively utilized to identify the damping coefficient of landing gear and the aircraft body weight. Through numerical experiments and comparisons, we have demonstrated that incorporating L-LAAF and learning rate decay strategies can further enhance the performance of the network. The numerical simulation based on Monte Carlo fully validates the method. The development of physics-based deep learning techniques for aircraft system parameter identification research can help researchers better understand and control the behavior of systems, providing effective solutions for optimizing system design.

**Keywords** PINNs · Aircraft ground taxiing model · Fokker–Planck equations · Inverse problem

## 1 Introduction

As a widely used mode of transportation, ensuring the safety and comfort of aircraft has become a paramount concern for designers. Extensive attention is being given to addressing this issue effectively. It is worth noting that a significant portion of flight accidents occurs during take-off and landing on the ground. These critical phases expose the aircraft to various challenges, particularly when encountering uneven runway surfaces. The resultant random vibrations not only pose immediate risks but also contribute to long-term fatigue damage to the structural integrity of the aircraft. Consequently, the lifespan of the aircraft could be significantly compromised. Given the aforementioned challenges, conducting comprehensive research on the aircraft taxiing model becomes imperative. By analyzing and simulating the complex dynamics involved in the aircraft's movement on the ground, designers can develop more robust and resilient structures for aircraft, ultimately enhancing their safety and extending their operational lifespan. This not only ensures the safety and comfort of passengers but also establishes a solid foundation for the aviation industry's continuous growth and development.

In the field of aircraft engineering, researchers have made significant contributions to studying the dynamic

---

Y. Zhang · Z. Jin · K. Zheng · W. Jia (✉)  
School of Mathematics and Statistics, Northwestern Polytechnical University, Xi'an 710072, China  
e-mail: jiwantao@nwpu.edu.cn

L. Wang  
The State Key Laboratory of Nonlinear Mechanics, Institute of Mechanics, Chinese Academy of Sciences, Beijing 100190, China

behavior of the landing gear structure. Notably, Michael and Schlaefke [1–3] simplified the aircraft landing gear structure as a linear damping spring vibrator for theoretical study. This theoretical model has proven to be effective in analyzing the performance of the landing gear system. In 1953, Milwitzky and Cook [4] proposed a two-degree-of-freedom spring-mass model to study the dynamic responses of main landing gear during aircraft fall shock. Despite their simplicity, these models have stood the test of time and continue to be widely utilized in current research and practical applications. Their enduring popularity stems from their ability to accurately capture the essential characteristics and behavior of the landing gear under various loading and environmental conditions. By simplifying the complex dynamics of the landing gear into fundamental elements such as springs, masses and dampers, these models provide valuable insights into the performance and response of the landing gear structure during critical phases of flight.

Based on the two-mass block model, power spectral density method [5, 6], state space method [7] and deterministic numerical analysis [8] were proposed to obtain the responses of aircraft taxiing on uneven runway surface. The above methods mainly solve stochastic differential equations (SDEs) of the system under random excitation based on Newton's law of motion. Considering the aircraft taxiing model derived from Hamilton's principle [9], there are normally two ways to analyze the system responses. Specifically, one approach is to obtain the probability density of the system responses by Monte Carlo (MC) simulation from the perspective of SDEs of the system. However, this method requires a large number of samples. It will consume a large amount of computing time and resources if higher accuracy is required. Another approach is to analyze the dynamic behavior of the system by using the Fokker-Planck (FP) equations who govern the transient probability density of the system responses. The transient responses of an aircraft landing gear system play a pivotal role in understanding and analyzing the dynamic characteristics of the aircraft body. These responses provide valuable insights into how the landing gear interacts with the ground during taxiing, take-off and landing phases. As a result, studying and analyzing these transient responses prove to be of great significance for subsequent research on taxiing load. In general, it is difficult to get the exact solution of the FP equations. Solving high-dimensional FP equation is

still a question worthy of further investigation. Moreover, traditional numerical methods, such as the finite difference method and the finite element method, can be computationally expensive and time-consuming, especially when considering long-time simulations or solving FP equations repeatedly in parameter studies or optimization routines. The computational cost increases significantly as the grid or mesh resolution is increased to achieve more accurate results. This limits the efficiency and scalability of traditional methods, hindering their application to real-world problems with practical time constraints.

The landing gear buffer system is crucial for the safe landing process and passenger comfort in aircraft. Most modern aircraft use an oil-gas buffer, where the oil damping force plays a significant role. As the buffer moves, the oil flows through the damping hole at high speed, resulting in a damping effect. Therefore, the oil damping coefficient is a key factor in the design of the buffer system. In aircraft design, the weight of the aircraft body is an important factor that affects flight performance and economic feasibility. It ultimately determines the success or failure of the aircraft design. Additionally, the aircraft body weight has implications for subsequent load research. Thus, obtaining the characteristics of the aircraft structure from a limited set of measured data on aircraft skidding responses can help predict the oil damping coefficient of the landing gear buffer and the weight of the aircraft body. These predictions provide a theoretical basis for further analysis of data-driven model inversion and landing gear fatigue load. However, traditional numerical techniques cannot solve the problem of inversion of the physical driven differential model of aircraft sliding under the excitation of uneven ground. In addition, the inevitable noise makes reconstructing objects from available data a computationally intractable problem.

Neural networks (NNs) have been widely used in engineering fields, such as structural damage detection [10–13], fast tracking [14], structural parameter identification [15, 16] and so on. Physical information neural networks (PINNs) method is a general-purpose framework developed by Karniadakis for solving forward and inverse problems of partial differential equations (PDEs) [17]. This method is based on the PDEs formed by physical modeling and the solution is obtained by using neural networks as function approximation. Compared with traditional numerical methods, using deep learning technology to solve PDEs does not need

generate grids, and fewer points are required to train the network to approximate the solution, which can save a lot of time and avoid the dimension disaster problem when solving the high-dimensional PDEs. Xu et al. [18] introduced normalization conditions as supervision conditions and put forward a numerical method for solving FP equations using deep learning. Zhai et al. [19] introduced the differential operators of FP equations into the loss function, which significantly reduces the need for a large amount of data in the learning process. Wang et al. [20] proposed a learning rate annealing algorithm to balance different terms of the loss function so as to achieve higher accuracy of training results. Nowadays, the PINNs algorithm is widely applied to solve forward and inverse problems of PDEs in many fields, such as fluid mechanics [21–23], materials [24–26], power systems [27,28], biomedicine [29–31] and so on. The wide application of the PINNs algorithm across diverse domains underscores its potential for solving complex high-dimensional physical problems that are challenging for traditional numerical methods.

In this study, it is expected to obtain higher precision training results with fewer iterations. We try to make special processing of the activation function of the network to obtain more efficient training strategies. Typically, the activation function of hidden units in networks operates on affine transformations, and the introduction of activation functions introduces nonlinearity to neurons since most activation functions are nonlinear. This ensures that neural networks can handle more complex nonlinear models. Some researchers introduced adaptive activation function in neural networks, which greatly improved the training accuracy and accelerated the convergence speed [32–34]. Jagtap et al. [35,36] applied globally adaptive activation functions and locally adaptive activation functions based on the basic framework of PINNs, which proved the locally adaptive activation is superior to fixed and globally adaptive activation in training speed and accuracy. The learning rate plays a crucial role in searching for the global minimum [37]. A large learning rate may exceed the global minimum, while a small learning rate increases the computational cost. Usually, the learning rate decay strategy is adopted to achieve a better performance of the neural network. In this work, we employ layer-wise locally adaptive activation functions (L-LAAF) [36] to optimize the network structure. Given the varying learning abilities of each hidden layer, using locally defined activation slopes for

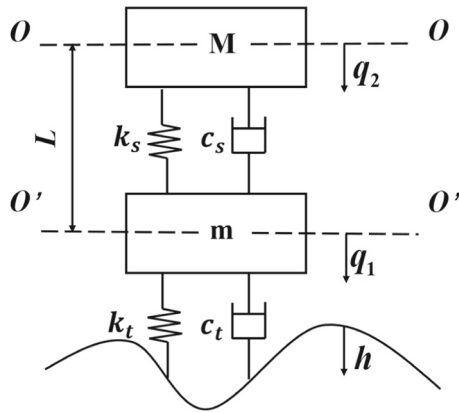
L-LAAF further improves the network's performance. Additionally, we adopt a piecewise constant learning rate decay strategy to accelerate the convergence of the model.

The rest of this paper is arranged as follows. In Sect. 2, the dissipative quasi-Hamiltonian system for aircraft ground sliding under Gaussian white noise excitation is derived. The main idea of solving FP equations using neural networks is briefly discussed in Sect. 3. The flow and steps of the proposed method are also introduced in detail. Section 4 provides results and detailed discussions for forward problems using the improved PINNs and learning rate decay strategy. In Sect. 5, the inverse problem is solved, that is, the experimental results of system parameter identification are given. Finally, Sect. 6 summarizes the work of this paper.

## 2 Model of aircraft landing gear

Most landing and skidding dynamic models are based on the two-mass system, i.e., the system is divided into elastic support mass and inelastic support mass, which can not only reflect the actual movement of the landing gear, but also greatly simplify the mathematical model. This part introduces the two-mass aircraft taxiing model under the excitation of uneven surface, which is derived from Hamilton's principle. Then, the FP equation governing the transient probability density is used to analyze the system responses and solve the inversion problem of the physical driven differential aircraft model under the excitation of uneven surface. Solving these problems is of great significance for understanding the dynamic characteristics of aircraft and predicting the design of aircraft structure.

The linearized model of the aircraft during taxing is shown in Fig. 1, which consists of two concentrated masses  $M$  and  $m$  [5].  $M$  is the sprung mass involving the mass of airframe, wing and buffer outer cylinder which is supported by the gear air.  $m$  is the unsprung mass including the mass of buffer piston rod, brake device and wheel which reacts the air spring force to the ground. The basic assumptions in the model is: the air spring force and oil damping force are linear; the wheel below the non-spring supported mass is simplified as a tire with linear damping and linear stiffness.



**Fig. 1** Linearized model of the landing gear

Let the O–O plane be a zero potential surface, selecting the displacement of sprung mass  $q_2$ , the displacement of unsprung mass  $q_1$  as generalized coordinates.

For the system shown in Fig. 1, the potential energy of the system is given by

$$V = V_a(q_1, q_2) + V_t(q_1, q_2) - mg(L + q_1) - Mgq_2, \tag{1}$$

where  $V_a(q_1, q_2)$  is the elastic potential energy of the air spring, and  $V_t(q_1, q_2)$  is the elastic potential energy of the tire.  $L$  represents the distance between two mass blocks. The potential energy in the linearized model can be expressed as follows:

$$\begin{cases} V_t = \frac{1}{2}k_t q_1^2 \\ V_a = \frac{1}{2}k_s (q_2 - q_1)^2 \end{cases}, \tag{2}$$

where  $k_s$  and  $k_t$  represent linear stiffness of the buffer and linear stiffness of the tire, respectively. The kinetic energy of system can be defined as

$$T = \frac{1}{2}m\dot{q}_1^2 + \frac{1}{2}M\dot{q}_2^2. \tag{3}$$

The Hamiltonian function of the system can be obtained from the Lagrange function as shown in the following [38]

$$\begin{aligned} H &= \left( \frac{p_1^2}{2m} + \frac{p_2^2}{2M} \right) \\ &+ [V_a(q_1, q_2) + V_t(q_1) - mg(L + q_1) - Mgq_2] \\ &= \left( \frac{p_1^2}{2m} + \frac{p_2^2}{2M} \right) + \frac{1}{2}k_s(q_2 - q_1)^2 \end{aligned}$$

$$+ \frac{1}{2}k_t q_1^2 - mg(l + q_1) - Mgq_2, \tag{4}$$

where  $p_1, p_2$  is the generalized momentum.

The stochastically excited and dissipated Hamiltonian system is given by Zhu et al. [39]

$$\begin{cases} \dot{q}_1 = \frac{\partial H}{\partial p_1} = \frac{p_1}{m} \\ \dot{q}_2 = \frac{\partial H}{\partial p_2} = \frac{p_2}{M} \\ \dot{p}_1 = -(k_t q_1 - k_s(q_2 - q_1) - mg) - c_{11}(\vec{q}, \vec{p}) \frac{\partial H}{\partial p_1} \\ \quad - c_{12}(\vec{q}, \vec{p}) \frac{\partial H}{\partial p_2} + f_{1k}(\vec{q}, \vec{p}) \xi_k(t) \\ \dot{p}_2 = -(k_s(q_2 - q_1) - Mg) - c_{21}(\vec{q}, \vec{p}) \frac{\partial H}{\partial p_1} \\ \quad - c_{22}(\vec{q}, \vec{p}) \frac{\partial H}{\partial p_2} \end{cases}. \tag{5}$$

where  $\xi_i(t) (i = 1, 2, \dots, k)$  is the description function of the runway pavement, which can be Gaussian white noise, broad band process or narrow band process.

However, the magnitude difference between the original system variables is huge, and variables related to the system momentum fluctuate in a large range, which bring the difficulty on solving the PDE numerically. To deal with this, the technique of transformation of variable is applied. We transform to rescale the variables as  $x_1 = q_1 + \frac{1}{k_t}(M - m)g, x_2 = q_2 + \frac{1}{k_t}(M - m)g - \frac{M}{k_s}g, x_3 = s_1 p_1, x_4 = s_2 p_2$ , here  $s_1$  and  $s_2$  are constants. Considering the linear damping  $c_{11} = c_t + c_s, c_{12} = -c_s, c_{21} = -c_s, c_{22} = c_s$ , the system is excited by the Gaussian white noise. So the Eq. (5) can be written as

$$\begin{cases} \dot{x}_1 = \frac{x_3}{s_1 m} \\ \dot{x}_2 = \frac{x_4}{s_2 M} \\ \dot{x}_3 = -[k_t s_1 x_1 - k_s s_1 (x_2 - x_1)] - (c_t + c_s) \frac{x_3}{m} \\ \quad + c_s \frac{x_4}{M} \frac{s_1}{s_2} + f_{1k} \xi_k(t) s_1 \\ \dot{x}_4 = -k_s s_2 (x_2 - x_1) + c_s \frac{x_3}{m} \frac{s_1}{s_2} - c_s x_4 \end{cases}, \tag{6}$$

### 3 Methodology

Consider a vector random process  $x(t) = [x_1(t), x_2(t), \dots, x_n(t)]^T$  governed by the following stochastic differential equation

$$\frac{d}{dt} x_j(t) = m_j(x, t) + \sum_{l=1}^m \sigma_{jl} W_l(t), \quad j = 1, 2, \dots, n, \tag{7}$$

where  $m_j$  is linear or nonlinear functions,  $\sigma_{jl}$  stands for constant, and  $W_l(t)$  is Gaussian white noise whose correlation function is

$$E[W_l(t)W_s(t + \tau)] = 2\pi K_{l_s}\delta(\tau), \tag{8}$$

here  $K_{l_s}$  is the spectral density.

The FP equation associated with Eq. (7) is shown as follows:

$$\frac{\partial p}{\partial t} + \sum_{j=1}^n \frac{\partial}{\partial x_j} (m_j p) - \frac{1}{2} \sum_{j,k=1}^n \frac{\partial^2}{\partial x_j \partial x_k} (\sigma_{jk} p) = 0, \tag{9}$$

where  $p = p(x, t)$  denotes the probability density function (PDF) of the  $x(t)$  at time  $t$ .

For convenience of analysis, rewrite Eq. (9) as

$$p_t + N[p, \lambda] = 0, x \in \Omega, t \in [0, T], \tag{10}$$

$$p(x, 0) = I(x), x \in \Omega, \tag{11}$$

$$p(x, t) = 0, x \in \partial\Omega, t \in [0, T], \tag{12}$$

where  $p(x, t)$  is defined on a domain  $\Omega \times [0, T]$ ;  $N[p, \lambda]$  is a differential operator with a parameter  $\lambda$  of the aircraft; and  $I(x)$  denotes the initial condition of the equation.  $\partial\Omega$  denotes the boundary of  $\Omega$ .

PINNs method is efficient for solving the forward and inverse problems involving nonlinear differential and integral equations with sparse, noisy and multi-fidelity data [5]. It can be used as a new class of numerical solvers for PDEs and to solve new data-driven model inversion and system identification problems. Consider a fully connected neural network with an input layer,  $K - 1$  hidden layers and an output layer. And the  $d$ th hidden layer contains  $N_d$  neurons. Each hidden layer accepts the output  $Z^{d-1} \in \mathbb{R}^{N_{d-1}}$  from the previous layer, affine transformation can be obtained from the following form

$$\mathcal{L}_d(Z^{d-1}) \triangleq w^d Z^{d-1} + b^d, \tag{13}$$

where the network weights term  $w^d \in \mathbb{R}^{N_d} \times \mathbb{R}^{N_{d-1}}$  and bias term  $b^d \in \mathbb{R}^{N_d}$  are related to the  $d$ th layer. We denote  $Z^0 = (x, t)$  as an input and  $Z^K = p(x, t)$  as the output value of final layer. Each of the transformed components is subjected to the nonlinear activation function  $\sigma(\cdot)$ , which serves as an input to the next layer. The utilization of L-LAAF makes each hid-

den layer have its own activation slope, the  $d$ th adaptive activation function is defined as follows:

$$\sigma(a_d(\mathcal{L}_d(Z^{d-1}))) (d = 1, 2, \dots, K - 1), \tag{14}$$

where  $\sigma$  is the activation function and  $a_d (d = 1, 2, \dots, K - 1)$  are additional  $K - 1$  parameters to be optimized. The improved neural network based on the L-LAAF can be expressed as [35,36]

$$p_{\bar{\Theta}}(Z) = (\mathcal{L}_K \circ \sigma \circ a_{K-1} \mathcal{L}_{K-1} \circ \sigma \circ a_{K-2} \mathcal{L}_{K-2} \circ \dots \circ \sigma \circ a_1 \mathcal{L}_1)(Z). \tag{15}$$

here, the trainable parameter  $\bar{\Theta}$  consists of  $\{w_d, b_d\}_{d=1}^K$  and  $\{a_d\}_{d=1}^{K-1}$ .

Compared with the original network, improved PINNs, i.e., PINNs with L-LAAF, are trained with a hyper-parameter  $a_d$  to each hidden layer for training. In the training process, we employed a segmented decay strategy for the learning rate to facilitate faster convergence of the network. The loss function is given by

$$Loss = Loss_f + Loss_B + Loss_I, \tag{16}$$

where  $Loss_f, Loss_B$  and  $Loss_I$  are defined as follows:

$$Loss_f = \frac{1}{N_f} \sum_{i=1}^{N_f} \left| \gamma(\hat{p}(x_i^f, t | \bar{\Theta})) \right|^2, \tag{17}$$

$$Loss_B = \frac{1}{N_B} \sum_{i=1}^{N_B} \left| \hat{p}(x_i^B, t | \bar{\Theta}) \right|^2, \tag{18}$$

$$Loss_I = \frac{1}{N_I} \sum_{i=1}^{N_I} \left| \hat{p}(x_i^I, t | \bar{\Theta}) - I(x_i) \right|^2. \tag{19}$$

Here,  $\gamma(\hat{p}) = \hat{p}_t + N(\hat{p}, x)$ ,  $Loss_f, Loss_B$  and  $Loss_I$  guarantee the approximate solution to satisfy Eqs. (10), (11) and (12), respectively.  $N_f, N_I$  and  $N_B$  denote the number of collation points, initial points and boundary points in training set, respectively. Algorithm 1 summarizes PINNs with L-LAAF.

#### 4 PDF Solving based on deep neural networks

In the follows, we use the above-mentioned PINNs with L-LAAF algorithm to investigate the solitons of the FP equation corresponding to Eq. (6) as shown in the following

**Algorithm 1** PINNs with L-LAAF.

**Input:** The collocations training set  $\{x_i^f, t\}$ , the initial training set  $\{x_i^I, 0\}$ , the boundary training set  $\{x_i^B, t\}$ .

**Output:**  $p(x, t | \bar{\Theta}^*)$

- 1: Initialize the weights and biases of the neural network.
- 2: Construct neural network and get the output  $\hat{p}(x, t | \bar{\Theta})$ .
- 3: Get the output's first and second derivatives  $\hat{p}_x(x, t | \bar{\Theta})$ ,  $\hat{p}_t(x, t | \bar{\Theta})$ ,  $\hat{p}_{xx}(x, t | \bar{\Theta})$ .
- 4: Update the neural network by minimizing the loss function

$$Loss = \frac{1}{N_f} \sum_{i=1}^{N_f} |\gamma(\hat{p}(x_i^f, t | \bar{\Theta}))|^2 + \frac{1}{N_B} \sum_{i=1}^{N_B} |\hat{p}(x_i^B, t | \bar{\Theta})|^2 + \frac{1}{N_I} \sum_{i=1}^{N_I} |\hat{p}(x_i^I, t | \bar{\Theta}) - I(x_i^I)|^2.$$

- 5: The optimal parameters  $\bar{\Theta}^*$  are obtained by appropriate optimization method to minimize the loss function.
- 6: **Return** results.

$$\begin{aligned} \frac{\partial p}{\partial t} + N(p, \lambda) = & \frac{\partial p}{\partial t} + \left[ \frac{\partial}{\partial x_1} \left( \frac{x_3}{s_1 m} \cdot p \right) \right. \\ & + \frac{\partial}{\partial x_2} \left( \frac{x_4}{s_2 M} \cdot p \right) \\ & - \frac{\partial}{\partial x_3} [k_t s_1 x_1 \cdot p - k_s s_1 (x_2 - x_1) \cdot p] \\ & - \frac{\partial}{\partial x_3} [(c_t + c_s) \frac{x_3}{m} \cdot p] \\ & + \frac{\partial}{\partial x_3} \left( c_s \frac{x_4 s_1}{M s_2} \cdot p \right) - \frac{\partial}{\partial x_4} [k_s (x_2 \\ & - x_1) s_2 \cdot p] + \frac{\partial}{\partial x_4} \left( c_s \frac{x_3 s_2}{m s_1} \cdot p \right) \\ & \left. - \frac{\partial}{\partial x_4} (c_s x_4 \cdot p) \right. \\ & \left. - \frac{1}{2} (\sqrt{2D})^2 (f_{1k})^2 \frac{\partial^2 p}{\partial x_3^2} \right] = 0, \end{aligned} \tag{20}$$

For the convenience of expression, the parameters of the aircraft system are denoted by  $\lambda$ , which are known in solving the forward problem and the parameter values of the aircraft system used in this section are shown in Table 1.

The space interval is  $[-2, 2] \times [-2, 2] \times [-2, 2] \times [-1.5, 1.5]$ , and the time interval is  $[0, 0.5]$ . The corresponding initial value conditions are as follows:

$$I(x, t) = \frac{1}{(2\pi)^2 |\Sigma|^{\frac{1}{2}}} \exp \left\{ -\frac{1}{2} (x^T - \mu)^T |\Sigma|^{-1} (x^T - \mu) \right\}, \tag{21}$$

where

$$x^T = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}, \mu = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \Sigma = \begin{pmatrix} \sigma_1 & 0 & 0 & 0 \\ 0 & \sigma_2 & 0 & 0 \\ 0 & 0 & \sigma_3 & 0 \\ 0 & 0 & 0 & \sigma_4 \end{pmatrix}, \tag{22}$$

$$x = (x_1, x_2, x_3, x_4) \in [-2.0, 2.0] \times [-2.0, 2.0] \times [-2.0, 2.0] \times [-1.5, 1.5], \tag{23}$$

and the left and right boundary conditions of Eq. (20) are shown as

$$B^l(x, t) = B^l(x_1, x_2, x_3, x_4, t) = 0, t \in [0, 0.5], \tag{24}$$

if  $x_1 = -2.0$  or  $x_2 = -2.0$  or  $x_3 = -2.0$  or  $x_4 = -1.5$ .

$$B^r(x, t) = B^r(x_1, x_2, x_3, x_4, t) = 0, t \in [0, 0.5], \tag{25}$$

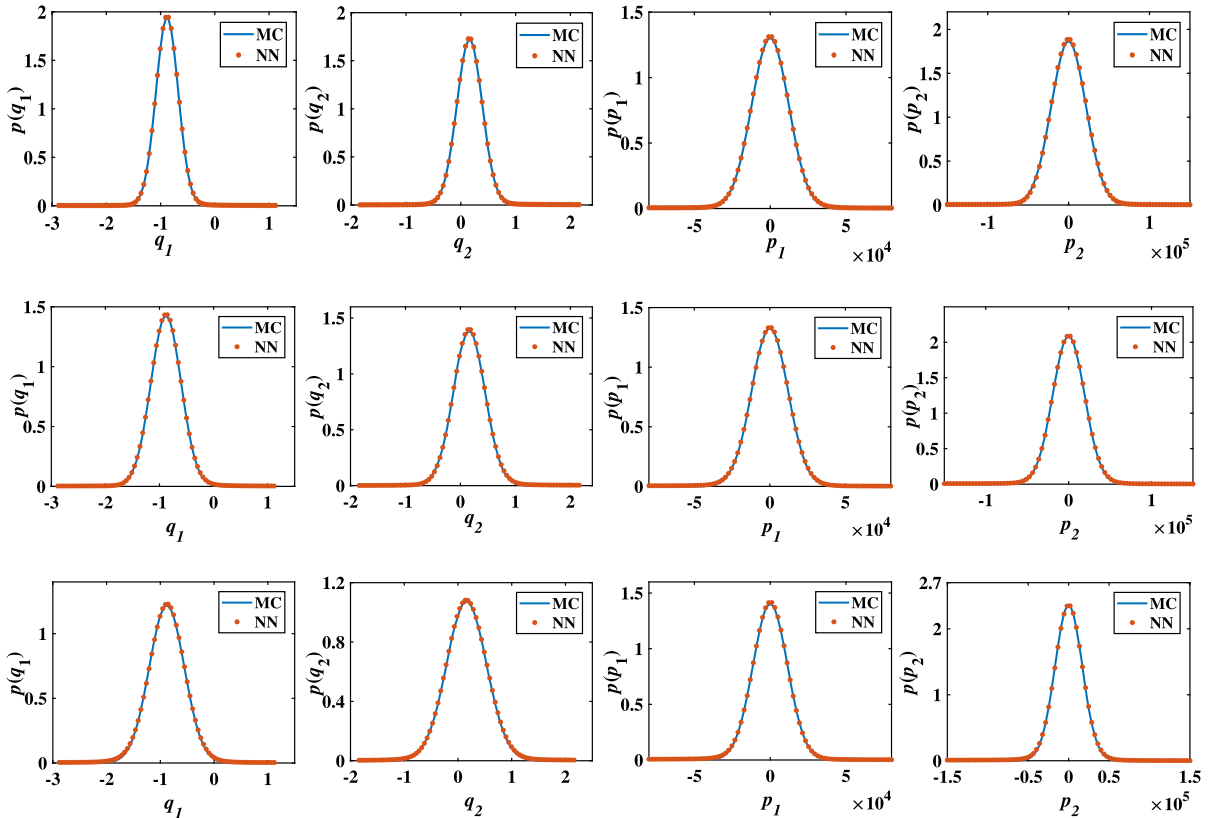
if  $x_1 = 2.0$  or  $x_2 = 2.0$  or  $x_3 = 2.0$  or  $x_4 = 1.5$ .

The structure of PINNs with L-LAAF is composed of four hidden layers with 20 neurons and the Adam optimizer is used to train the model. The activation function of the hidden layer adopts L-LAAF, and set the exponential function as the activation function for the output layer. The initial points training set  $N_I = 20000$  consists of two parts: 11000 points and 9000 points are selected in  $\Omega = [-2.0, 2.0] \times [-2.0, 2.0] \times [-2.0, 2.0] \times [-1.5, 1.5]$  and  $\Omega_1 = [-0.45, 0.45] \times [-0.45, 0.45] \times [-0.45, 0.45] \times [-0.3, 0.3]$ , respectively. Given a set of boundary points with  $N_B = 10000$  and a set of collation points with  $N_f = 90000$ , both of them are obtained by the Latin hypercube sampling (LHS) strategy. Setting the learning rate of the first 30,000 iterations as  $6 \times 10^{-3}$ , and the learning rate of the rest 170,000 iterations as  $6 \times 10^{-4}$ .

As a consequence, we can discover that the predict solutions approximate the MC solutions pretty well from Fig. 2, which exhibits the comparisons between MC solutions and predict solutions for  $t = 0.1$  (top

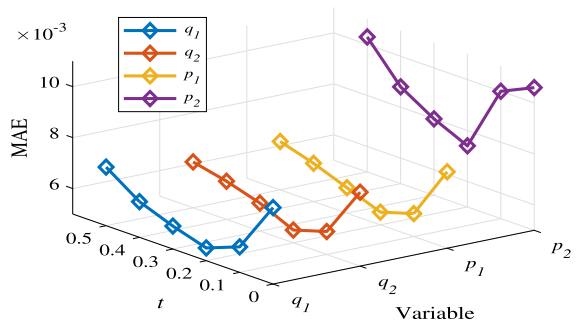
**Table 1** Aircraft system parameters

Parameter	$M$ (kg)	$m$ (kg)	$k_s$ (N m <sup>-1</sup> )	$k_t$ (N m <sup>-1</sup> )
Value	25345	7388	225000	$2 \times 10^5$
Parameter	$c_s$ (N m <sup>-1</sup> s)	$c_t$ (N m <sup>-1</sup> s)	$s_1$	$s_2$
Value	13940	400	$2.5 \times 10^{-5}$	$10^{-5}$



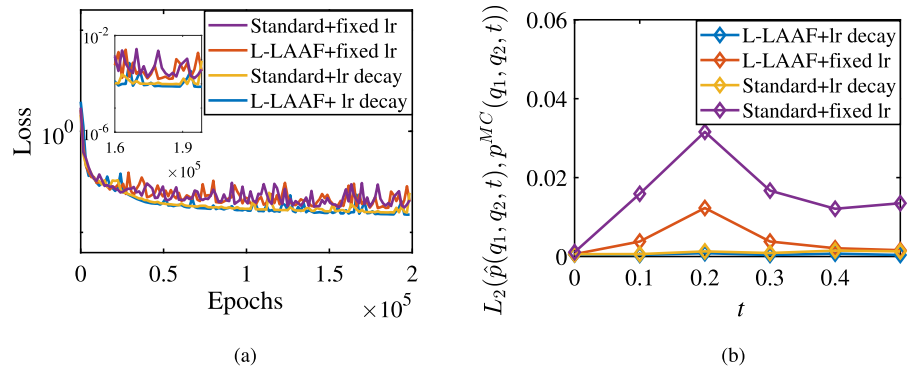
**Fig. 2** The comparison results between the solutions obtained from MC method and PINNs with L-LAAF algorithm at  $t = 0.1$  (top row),  $t = 0.3$  (middle row) and  $t = 0.5$  (bottom row)

row),  $t = 0.3$  (middle row) and  $t = 0.5$  (bottom row). And the mean absolute error between MC solutions and predict solutions defined by  $MAE = \frac{\sum_{i=1}^N |\hat{p}_i - p_i|}{N}$  is drawn in Fig. 3 when  $t = 0.0, 0.1, 0.2, 0.3, 0.4, 0.5$ . It shows that the error of each variable is extremely close to zero, which verifies the accuracy of our training results. To our pleasant surprise, the utilization of MC (Monte Carlo) method using approximately 10 million trajectories, which takes approximately 8h, whereas the proposed approach presented in this paper required only about 2h.



**Fig. 3** The mean absolute error of each variable between the solutions obtained from MC method and PINNs with L-LAAF algorithm

**Fig. 4** **a** Loss function; **b**  $L_2$  error with respect to  $p(q_1, q_2)$  for different training settings (standard tanh activation function and fix learning rate; L-LAAF and fixed learning rate; standard tanh activation function; L-LAAF and learning rate decay)



In what follows, we adopt the same training data set and basic network structure as the previous part to discuss the influence of the L-LAAF and the learning rate decay on network learning ability. In order to verify the effectiveness of PINNs enhanced by the L-LAAF and the learning rate decay more directly, we show the training results of four different settings (standard tanh activation and fixed learning rate ( $lr = 6 \times 10^{-3}$ ), L-LAAF using the tanh activation and fixed learning rate ( $lr = 6 \times 10^{-3}$ ), standard tanh activation and learning rate decay ( $lr = 6 \times 10^{-3}$  in the first 30000 iterations,  $lr = 6 \times 10^{-4}$  in the rest iterations), L-LAAF using the tanh activation and learning rate decay (the learning rate is consistent with the third case)) in Fig. 4.

Figure 4a illustrates the variation of the loss function with 20,000 Adam iterations. It shows that the loss function converges faster with the L-LAAF compared to the fixed tanh activation function under the same learning rate. Furthermore, the learning rate decay can make the value of the loss function converge less than the fixed learning rate when the activation function is consistent. The  $L_2$  error of the above four settings as indicated in Fig. 4b. We observe that the L-LAAF produces a lower prediction error than the fixed activation function under the same learning rate. It can also be seen from Fig. 4b that the learning rate decay can make  $L_2$  error reach a smaller value than fixed learning rate.

Figure 5 displays the comparison between MC results and PINNs results of four settings at  $t = 0.3$  with 200,000 Adam iterations. It is obvious that the results are consistent with those mentioned above. In a word, the PINNs with L-LAAF and the learning rate decay can effectively accelerate the convergence of the network, and achieve higher accuracy of training results.

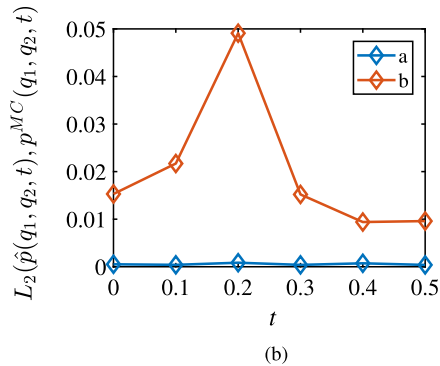
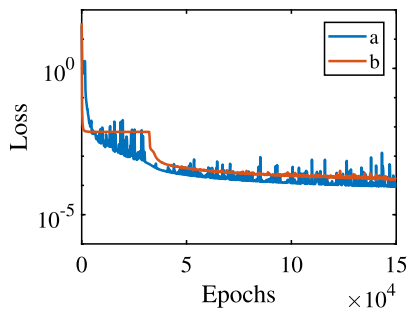
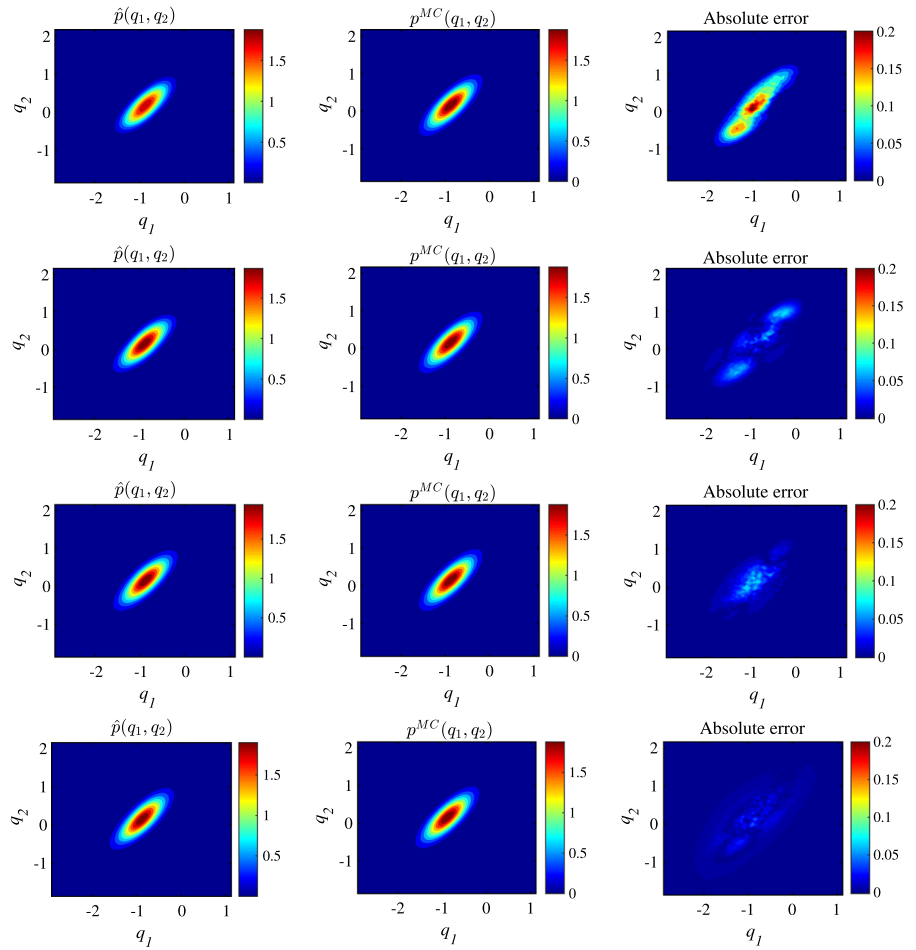
We further consider the influence of the selection of initial points on network. The boundary points set and collation points set used for computation are consistent with those discussed above, and we use the PINNs with L-LAAF algorithm and the learning rate decay to train the network. Figures 6 and 7 display the training results with different initial points selection settings (Case a. select 11,000 in  $\Omega$ , 9000 in  $\Omega_1$  by using LHS strategy Case b. select 20,000 in  $\Omega$  by using LHS strategy), and the total number of iterations is 150000. It can be clearly seen from the corresponding loss function converges faster and the fluctuation range is larger in Case a. From the  $L_2$  errors of joint PDF  $p(q_1, q_2)$  of each transient corresponding to different initial points selection settings in Fig. 6b. It manifests that  $L_2$  error of Case b is much larger than Case a. Figure 7a and b shows the comparison of PDF and corresponding absolute error of prediction with two different initial point selection methods at  $t = 0.4$ . Obviously, Fig. 7 reveals that the prediction accuracy is lower near the boundary, but getting higher accuracy where the gradient of PDF is larger under the Case a. And appropriately selecting more initial points near the middle part of the variable interval can make better prediction accuracy than uniformly selecting points.

## 5 Data-based parameter identification

In practical engineering, there are some parameters that cannot be directly measured in the aircraft structure. Thus, it is of great significance to identify them. Identifying unknown parameters in an aircraft model based on the observed data is the inverse problem to be studied here. The inverse problem is considered based on the trajectories of system variables at several discrete



**Fig. 5**  $\hat{p}(q_1, q_2)$  obtained by PINNs,  $p^{MC}(q_1, q_2)$  obtained by MC method and absolute error between them at  $t = 0.5$  for different training settings (first row): standard tanh activation function and fix learning rate; (second row): L-LAAF and fixed learning rate; (third row): standard tanh activation function and learning rate decay; (fourth row): L-LAAF and learning rate decay

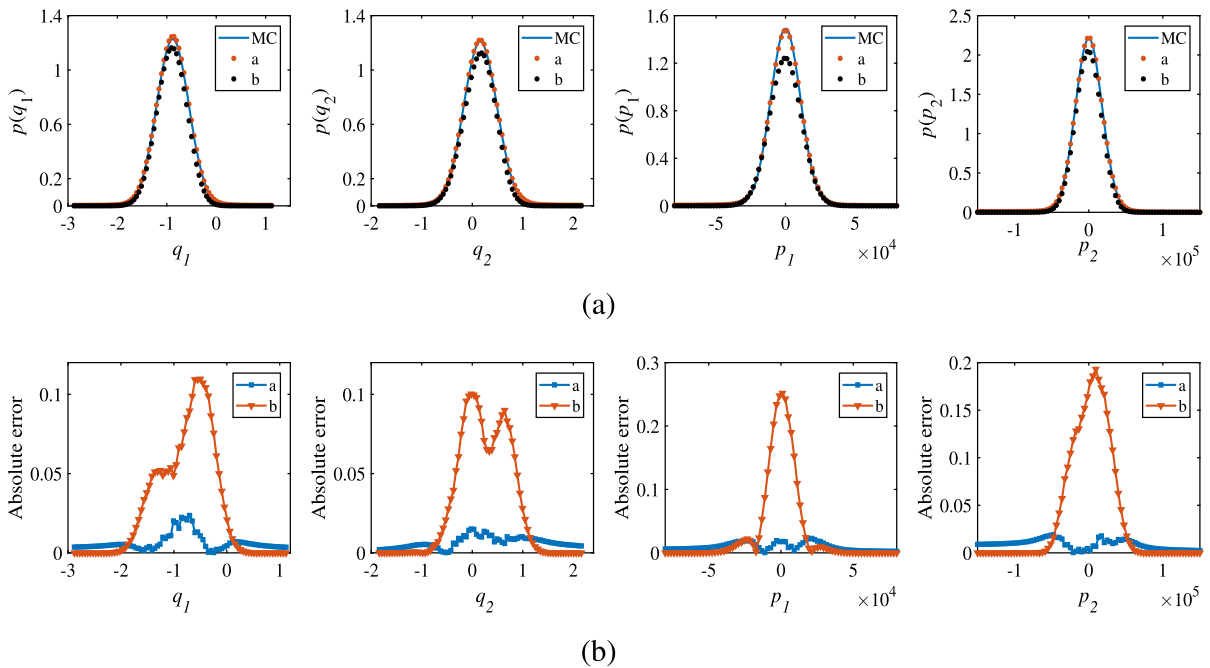


**Fig. 6** **a** Loss function; **b**  $L_2$  error with respect to  $p(q_1, q_2)$  for PINNs with L-LAAF algorithm and learning rate decay ( $lr = 6 \times 10^{-3}$  in the first 30,000 Adam iterations,  $lr = 6 \times 10^{-4}$  in the rest 120,000 Adam iterations) for different selections of

initial points (Case a. select 11,000 points in  $\Omega$ , 9000 points in  $\Omega_1$  by using LHS strategy Case b. select 20,000 points in  $\Omega$  by using LHS strategy)

moments. For the FP equation with unknown coefficients, the exact form of the equation can be derived by modifying the above PINNs with LAAF algorithm and

combining with the observed data, and then the PDF at any instant of time can be obtained. The specific implementation steps to solve the inverse problem are as fol-



**Fig. 7** **a** The comparison results between the solutions obtained from MC method and PINNs with L-LAAF algorithm with learning rate decay at  $t = 0.4$  for different selections of initial points; **b** The absolute error of each variable between the solutions obtained from MC method and PINNs with L-LAAF algorithm

with learning rate decay ( $lr = 6 \times 10^{-3}$  in the first 30,000 Adam iterations,  $lr = 6 \times 10^{-4}$  in the rest 120,000 Adam iterations) at  $t = 0.4$  for different selections of initial points (Case a. select 11,000 points in  $\Omega$ , 9000 points in  $\Omega_1$  by using LHS strategy Case b. select 20,000 points in  $\Omega$  by using LHS strategy)

lows: solutions of the differential equation of the system at these moments are obtained through MC simulation under the condition that the trajectories of several discrete time points of the system variables are known. Then the unknown parameters  $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_n)$  are setting as trainable variables of the network to be identified. Namely, the optimization parameter  $\Theta$  consists of  $\{w_d, b_d\}_{d=1}^K$ ,  $\{a_d\}_{d=1}^{K-1}$  and  $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_n)$  in the inverse problem. PDF of  $N_t$  points estimated by MC is selected as label data, and the loss function is defined as

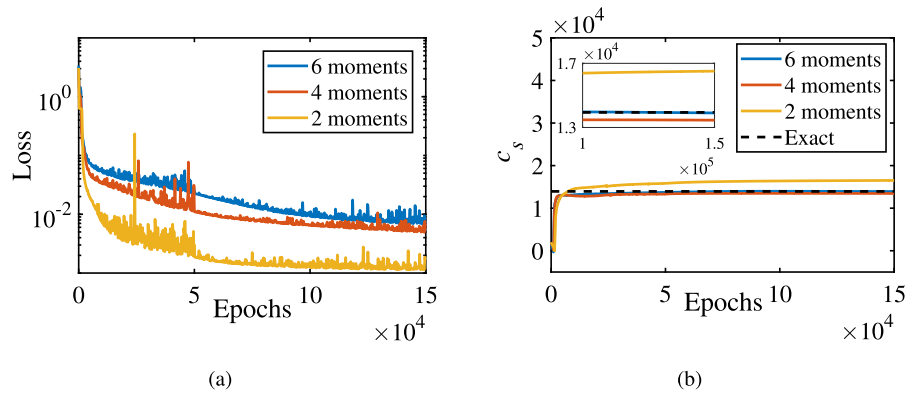
$$Loss = Loss_B + Loss_I + \frac{1}{N_t} \sum_{i=1}^{N_t} \left| p^{MC}(X, t) - \hat{p}(X, t) \right|^2. \quad (26)$$

Finally, the constructed network is optimized to obtain  $\Theta$  that minimizes the loss function.

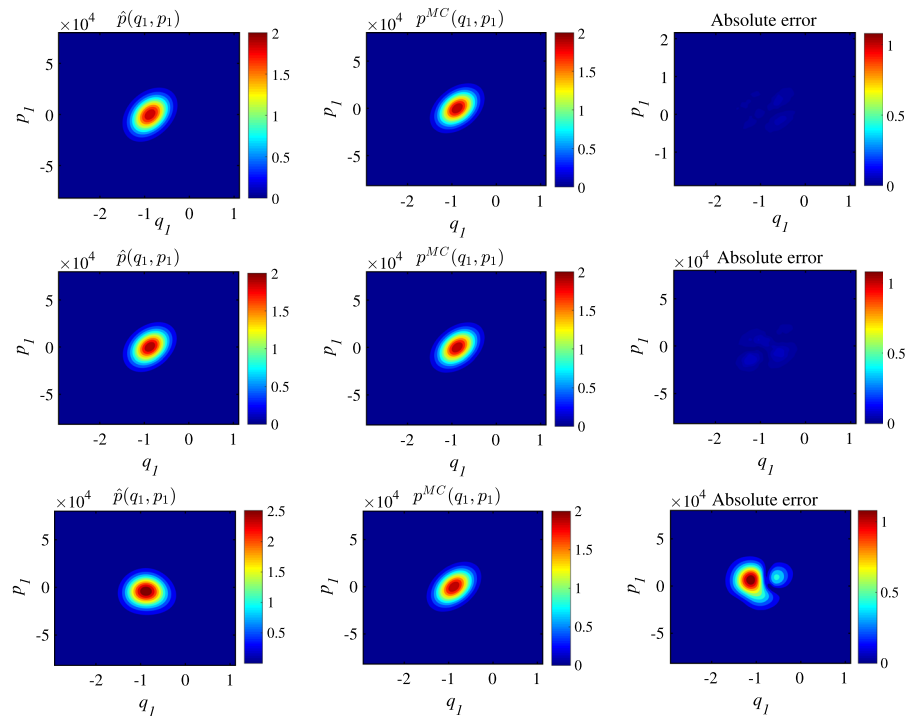
In this section, we turn our attention to the inverse problem of Eq. (20) to predict the design of landing gear structure parameters, namely, damping coefficient of landing gear  $c_s$  and aircraft body weight

$M$ . Here, we use the same network structure as the forward problem, and use the L-LAAF and learning rate decay strategy to train the network. On the premise of obtaining trajectories of system variables at individual time points, the MC method is used to simulate the PDF of the system, and then a small part of the PDF is selected as the collation points set. After the training data set is given, the PINNs with L-LAAF algorithm successfully predicted the potential data-driven unknown parameters by adjusting the loss function with 150,000 Adam iterations. The next part discusses the influence of the number of moments in the training set on network training under the total number of fixed 90000 training points remains unchanged. Figures 8, 9 and 10 specifically showcase the results of unknown parameter  $c_s$  by inputting training sets at six moments ( $t = 0.1, 0.2, 0.3, 0.4, 0.5$ ), 4 moments ( $t = 0.1, t = 0.2, t = 0.4, t = 0.5$ ), and 2 moments ( $t = 0.1, t = 0.5$ ), respectively. As can be seen from Fig. 8a, the loss function declines more slowly and its convergence value increases with the

**Fig. 8** **a** Loss function; **b** Variation of unknown coefficients  $c_s$  for PINNs with L-LAAF algorithm and learning rate decay ( $lr = 6 \times 10^{-3}$  in the first 30,000 Adam iterations,  $lr = 6 \times 10^{-4}$  in the rest 120,000 Adam iterations) for different number of moments (two moments; four moments; six moments)



**Fig. 9**  $\hat{p}(q_1, p_1)$  obtained by PINNs with L-LAAF and learning rate decay ( $lr = 6 \times 10^{-3}$  in the first 30,000 Adam iterations,  $lr = 6 \times 10^{-4}$  in the rest 120,000 Adam iterations),  $p^{MC}(q_1, p_1)$  obtained by MC method and absolute error between them at  $t = 0.3$  for (first row): the number of input moments is 6; (second row): the number of input moments is 4; (third row): the number of input moments is 2

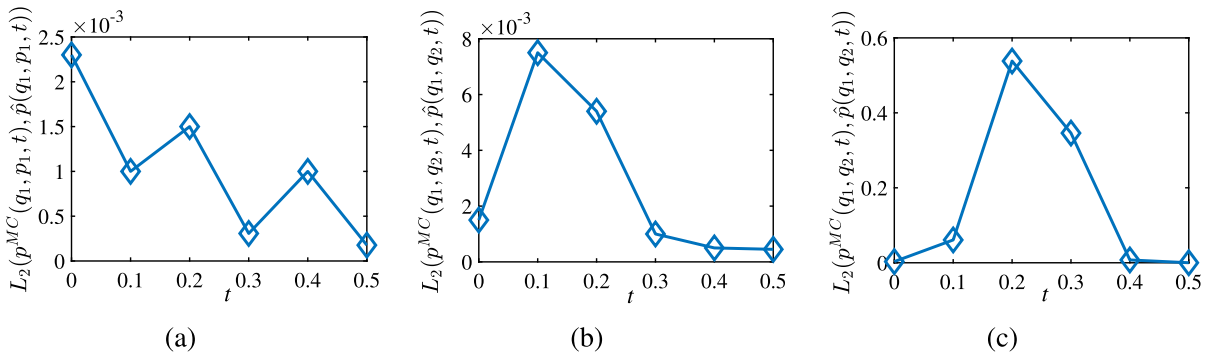


number of moments increases. Figure 8b illustrates the convergence of the estimated parameter to the real parameter. The relative error of the prediction reaches 0.2483% when the number of input moments is 6 from Table 2, indicating that the unknown parameter  $\lambda = c_s$  can be identified more accurately with more input moments.

The comparison between the predicted solution and MC solution of  $t = 0.3$  is shown in Fig. 9. And Fig. 10 exhibits  $L_2$  error related to  $p(q_1, q_2)$  at different moments. We observe that when the total number of input data is consistent, the more moments the input data are distributed, the higher the network train-

ing accuracy will be. To be specific, it is very important to integrate the data of multiple temporal moments in physical information learning. The relatively large amount of temporal data information can improve the accuracy of network prediction and parameter identification.

Due to the influence of uneven runway, the inevitable noise may have a certain influence on the test data, which makes parameter identification a thorny problem in calculation. Parameter identification of noisy label data based on PINNs with L-LAAF algorithm can further analyze the data-driven problem of model inversion. Then, parameters of FP equation are identi-



**Fig. 10** The comparison results of  $L_2$  error with respect to  $p(q_1, q_2)$  for different selections of initial points **a** The number of input moments is 6; **b** The number of input moments is 4; **c** The number of input moments is 2

**Table 2** Identification results of  $c_s$  for PINNs with L-LAAF algorithm and learning rate decay ( $lr = 6 \times 10^{-3}$  in the first 30,000 Adam iterations,  $lr = 6 \times 10^{-4}$  in the rest 120,000 Adam iterations) for different number of moments (two moments; four moments; six moments)

The number of moments	Identified (Correct $c_s = 13940$ )	Relative error (%)
6	13905.3	0.2483
4	13455.0	3.4788
2	16514.7	18.4699

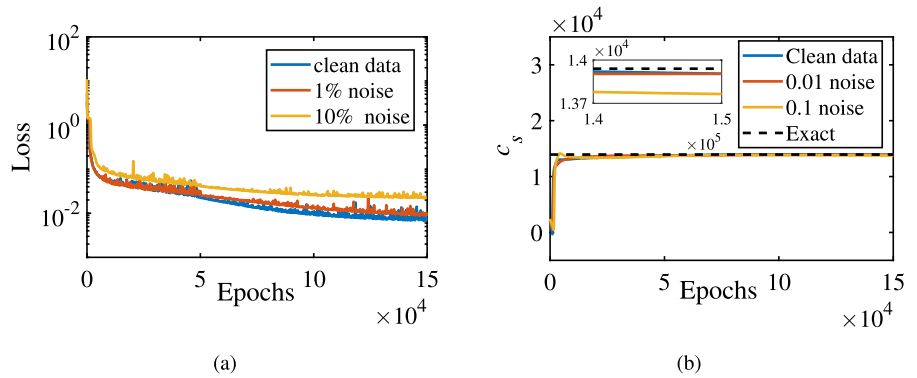
fied by clean data, 1% and 10% noise, respectively. As can be seen from Fig. 11a, the greater the noise intensity, the slower the loss function decreases and the larger the overall value of the loss function. Figure 11b and Table 3 indicate the variation curves of unknown parameter  $c_s$  in the iterative process and the identification results of  $c_s$  under different noise intensity. In the case of clean training data, the relative error of  $c_s$  is 0.2483%. In addition, the prediction is robust even if the training data is corrupted by 10% irrelevant Gaussian noise, and the relative error of  $c_s$  is 1.2620%. Specifically, we observe that the PINNs with L-LAAF algorithm can accurately identify  $c_s$ , even when the training data is corrupted by noise.

In the previous subsections, we have seen the advantages of using PINNs with L-LAAF algorithm and learning rate decay to solve forward problems. A natural question is whether this training strategy can play the same advantage in solving inverse problems. To verify our conjecture, we still choose a four hidden layers neural network, each hidden layer has 20 units. The unknown parameters  $\lambda = (c_s, M)$  in Eq. (20) are iden-

tified by using four PINNs training settings (standard tanh activation and fixed learning rate ( $lr = 6 \times 10^{-3}$ ), L-LAAF using the tanh activation and fixed learning rate ( $lr = 6 \times 10^{-3}$ ), standard tanh activation and learning rate decay ( $lr = 6 \times 10^{-3}$  in the first 30,000 iterations,  $lr = 6 \times 10^{-4}$  in the rest iterations), L-LAAF using the tanh activation and learning rate decay (the learning rate is consistent with the third case)), the total number of iterations is 150,000, and the corresponding inverse problem training results of the system are shown in Fig. 12.

Figure 12a provides the loss function with L-LAAF using the tanh activation decreases faster and the convergence value is smaller when the learning rate remains consistent. In the case of the same activation function, compared with the fixed learning rate, the decay of learning rate makes the loss function fluctuate less. Figure 12b, c and Table 4 present the variation curves of unknown parameters  $\lambda = (c_s, M)$  in the iterative process and the recognition results of two unknown parameters under the four training settings. Figure 12b, c clearly describes that training strategy of PINNs with L-LAAF and learning rate decay has the best performance, and the relative errors of  $c_s$  and  $M$  are 0.0916% and 1.7934%, respectively. And the PINNs with L-LAAF have higher accuracy in identifying two parameters when the learning rate is consistent. Furthermore, the learning rate decay can better identify  $c_s$ , but the change of learning rate does not have much effect on the identification of  $M$ , indicating that the relative error of  $c_s$  is more sensitive to the change of learning rate than  $M$ . In conclusion, it means that the L-LAAF and the learning rate decay training set-

**Fig. 11** **a** Loss function; **b** Variation of unknown coefficients  $c_s$  for PINNs with L-LAAF algorithm and learning rate decay ( $lr = 6 \times 10^{-3}$  in the first 30,000 Adam iterations,  $lr = 6 \times 10^{-4}$  in the rest 120,000 Adam iterations) for training data with different noise levels



**Table 3** Identification results of  $c_s$  for PINNs with L-LAAF algorithm and learning rate decay ( $lr = 6 \times 10^{-3}$  in the first 30,000 Adam iterations,  $lr = 6 \times 10^{-4}$  in the rest 120,000 Adam iterations) for training data with different noise levels

Noise intensity	Identified (Correct $c_s = 13940$ )	Relative error (%)
Clean data	13905.3	0.2483
1% noise	13904.2	0.2561
10% noise	13764.0	1.2620

ting contribute to the rapid learning process of PINNs related to inverse problems.

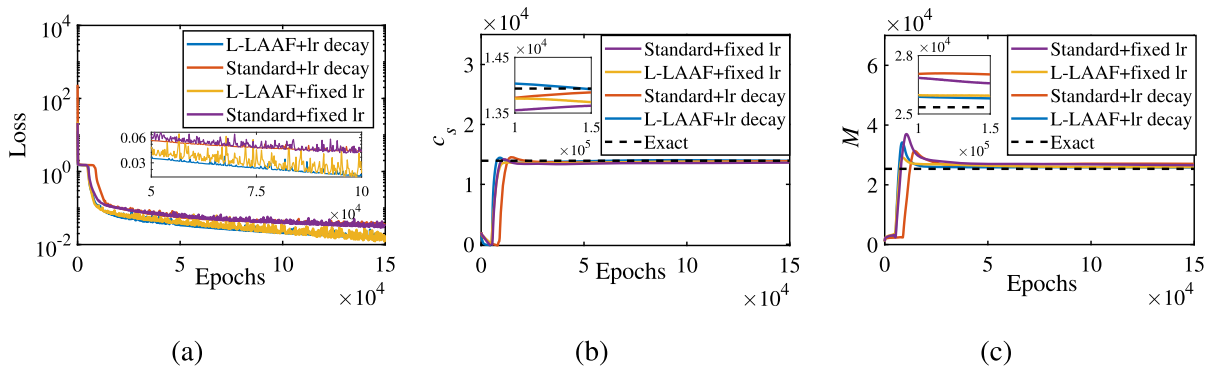
Finally, we discuss the two-parameter identification of the system with clean data, 1% and 10% noise, respectively. The network structure is composed of four hidden layers with 20 neurons in each layer, and the training strategy of L-LAAF and learning rate decay is adopted. The results after 150,000 iterations of training are shown in Fig. 13. As schematically illustrated in Fig. 13a, the corresponding loss function decreases more slowly with the increase in noise intensity. Figure 13b, c and Table 5 summarize the changes of unknown parameters  $c_s$  and  $M$  along with the iterative process and the identification results of  $c_s$  and  $M$ . We can clearly understand that although the training data is destroyed by 1% irrelevant Gaussian noise, the relative error of  $c_s$  is 0.4405%, and the relative error of  $M$  is 3.4452%, indicating that the PINNs with L-LAAF algorithm and the learning rate decay appear to be very robust with respect to noise levels in the data. For 10% noise of the data, the identification results of parameter  $c_s$  and  $M$  have relative errors of 2.1178% and 2.4859%, respectively, which points out that the parameter identification method has strong robustness

to data with noise. And the identification of  $M$  is less sensitive to the influence of noise intensity than the identification of  $c_s$ . Specifically, we observe that training strategy of PINNs with L-LAAF and learning rate decay can also correctly identify unknown parameters  $c_s$  and  $M$  with high accuracy, even if the training data is corrupted by noise.

Figure 13a illustrates that the corresponding loss function decreases more slowly with the increase in noise intensity. Figure 13b, c and Table 5 show the changes of unknown parameters  $c_s$  and  $M$  along with the iterative process and the identification results of parameters  $c_s$  and  $M$ . We can clearly understand that although the training data is destroyed by 1% irrelevant Gaussian noise, the relative errors of  $c_s$  and  $M$  are 0.4405% and 3.4452%, respectively, indicating that the prediction is still robust. For data with 10%, the relative error of  $c_s$  is 2.1178% and the relative error of  $M$  is 2.4859%, which imply that the PINNs with L-LAAF algorithm have strong robustness to data with noise, and the identification of  $M$  is less sensitive to the influence of noise intensity than that of  $c_s$ . Specifically, we observe that the PINNs with L-LAAF algorithm and the learning rate decay can also correctly identify unknown parameters  $c_s$  and  $M$  with very high accuracy, even if the training data is corrupted by noise.

### 6 Conclusion

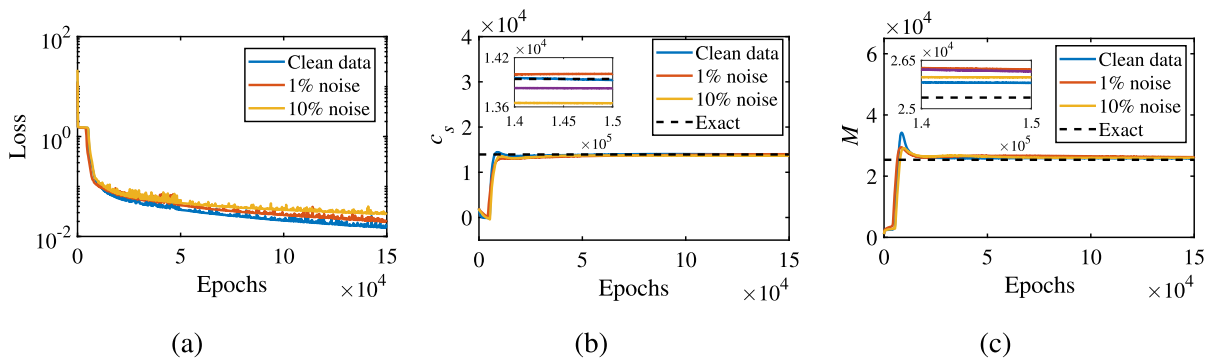
In this paper, we demonstrated a successful approach to solving the parameter identification problem of critical interest in aircraft systems. Specifically, we introduced and validated a PINNs framework with L-LAAF and learning rate decay strategy. Compared to the original PINNs, this improved framework effectively



**Fig. 12** **a** Loss function; **b** Variation of unknown coefficients  $c_s$ ; **c** Variation of unknown coefficients  $M$  for four different training strategies (standard tanh activation function and fix learning rate; L-LAAF and fixed learning rate; standard tanh activation function and learning rate decay; L-LAAF and learning rate decay)

**Table 4** Identification results of  $c_s$  and  $M$  for four different training strategies (standard tanh activation function and fix learning rate; L-LAAF and fixed learning rate; standard tanh activation function and learning rate decay; L-LAAF and learning rate decay)

Training strategy	Identified (Correct $c_s = 13940$ )	Relative error of $c_s$ (%)	Identified (Correct $M = 25345$ )	Relative error of $M$ (%)
Standard+fixed lr	13633.2	2.2011	26570.2	4.8339
L-LAAF+fixed lr	13696.7	1.7452	25939.0	2.3438
Standard+lr decay	13871.2	0.4934	27031.7	6.6519
L-LAAF+lr decay	13927.2	0.0916	25799.5	1.7934



**Fig. 13** **a** Loss function; **b** Variation of unknown coefficients  $c_s$ ; **c** Variation of unknown coefficients  $M$  for PINNs with L-LAAF algorithm and learning rate decay ( $lr = 6 \times 10^{-3}$  in the first 30000 Adam iterations,  $lr = 6 \times 10^{-4}$  in the rest 120,000 Adam iterations) for training data with different noise levels

enhanced network convergence and achieved higher accuracy. Our findings have been fully validated using MC numerical simulations. Additionally, we developed an improved PINNs framework for studying the probability density of the transient response of the aircraft system under uneven runway. We systematically investigated the impact of distribution of initial train-

ing points on prediction accuracy. Finally, we utilized the improved PINNs to solve the identification problem of damping coefficients and body mass parameters in the system. Notably, our results showed that, under consistent total input points, the relative error of the identification results of unknown parameter  $c_s$  is smaller when the training points are distributed at more

**Table 5** Identification results of  $c_s$  and  $M$  for PINNs with L-LAAF algorithm and learning rate decay ( $lr = 6 \times 10^{-3}$  in the first 30,000 Adam iterations,  $lr = 6 \times 10^{-4}$  in the rest 120,000 Adam iterations) for training data with different noise levels

Noise intensity	Identified (Correct $c_s = 13940$ )	Relative error of $c_s$ (%)	Identified (Correct $M = 25345$ ) (%)	Relative error of $M$ (%)
Clean data	13927.2	0.0916	25799.5	1.7934
1 noise	14001.4	0.4405	26218.2	3.4452
10 noise	13644.7	2.1178	25975.1	2.4859

moments. Furthermore, even when the training data was corrupted by 10% irrelevant Gaussian noise, our method accurately identified the parameters, demonstrating strong noise robustness of the improved PINNs. We believe that further development of the improved PINNs for parameter identification in aircraft systems can enhance modeling accuracy, accelerate the parameter identification process, widen design possibilities and provide valuable design guidance.

**Funding** This work was supported by the National Natural Science Foundation of China [grant numbers 12172286, 11872306]; and the Aeronautical Science Foundation of China [Grant Number 201941053004].

**Data availability** The data that supports the findings of this study is available from the corresponding author upon reasonable request.

#### Declarations

**Conflict of interest** The authors have no relevant financial or non-financial interests to disclose.

#### References

- Michael, F.: Theoretical and experimental principles of landing gear research and development. *Luftfahrtforschung* **14**, 387–419 (1937)
- Schlaefke, K.: Buffered and unbuffered impact on landing gear. *TB* **10**, 129–133 (1943)
- Schlaefke, K.: On force-deflection diagrams of airplane shock absorber struts. *NACA Tech. Memo.* **1373**, 109–113 (2015)
- Milwitzky, B., Cook, F.E.: *Analysis of Landing-Gear Behavior*. NASA, Washington, D.C. (1953)
- Nie, H., Kortum, W.: Analysis for aircraft taxiing at variable velocity on unevenness runway by the power spectral density method. *Nanjing Univ. Aeronaut. Astronaut.* **17**, 64–70 (2000)
- Liu, L.: Optimization of oleo-pneumatic shock absorber of aircraft. *Acta Aeronaut. Astronaut. Sin.* **13**, 206511 (1992)
- Hammond, J.K., Harrison, R.F.: Nonstationary response of vehicles on rough ground—a state space approach. *J. Dyn. Syst. Meas. Control* **103**, 245–250 (1981)
- Tung, C.C.: The effects of runway roughness on the dynamic response of airplanes. *J. Sound Vib.* **5**, 164–172 (1967)
- Zhang, Z.H., Zhu, S.J., Lou, J.J.: *Analysis for Aircraft Taxiing on Unevenness Runway Based on the Hamiltonian Systems*. In: The 9th National Conference on Vibration Theory and Application, Hangzhou (2007)
- Khatir, S., Tiachacht, S., Le, T.C., et al.: An improved artificial neural network using arithmetic optimization algorithm for damage assessment in FGM composite plates. *Compos. Struct.* **273**, 114287 (2021)
- Alazzawi, O., Wang, D.: Deep convolution neural network for damage identifications based on time-domain PZT impedance technique. *J. Mech. Sci. Technol.* **35**, 1809–1819 (2021)
- Ho, L.V., Nguyen, D.H., Mousavi, M., et al.: A hybrid computational intelligence approach for structural damage detection using marine predator algorithm and feedforward neural networks. *Compos. Struct.* **252**, 106568 (2021)
- Ho, L.V., Trinh, T.T., De, R.G., et al.: An efficient stochastic-based coupled model for damage identification in plate structures. *Eng. Fail. Anal.* **131**, 105866 (2022)
- Wang, S., Wang, H., Zhou, Y., et al.: Automatic laser profile recognition and fast tracking for structured light measurement using deep learning and template matching. *Measurement* **169**, 108362 (2021)
- Jiang, J., Chen, Z., Wang, Y., et al.: Parameter estimation for PMSM based on a back propagation neural network optimized by chaotic artificial fish swarm algorithm. *Int. J. Control.* **14**, 615–632 (2019)
- Chen, Y., Lu, L., Karniadakis, G.E., et al.: Physics-informed neural networks for inverse problems in nano-optics and metamaterials. *Opt. Express.* **28**, 11618–11633 (2020)
- Raissi, M., Perdikaris, P., Karniadakis, G.E.: Physics-informed neural networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *J. Comput. Phys.* **378**, 686–707 (2019)
- Xu, Y., Zhang, H., Li, Y.G., Liu, Q., Kurths, J.: Solving Fokker–Planck equation using deep learning. *Chaos* **30**, 013133 (2020)
- Zhai, J.Y., Dobson, M., Li, Y.: A deep learning method for solving Fokker–Planck equations. In: The 2nd Mathematical and Scientific Machine Learning Conference, Lausanne (2021)

20. Wang, S.F., Teng, Y.G., Perdikaris, P.: Understanding and mitigating gradient flow pathologies in physics-informed neural networks. *SIAM. J. Sci. Comput.* **43**, 3055–3081 (2020)
21. Raissi, M., Yazdani, A., Karniadakis, G.E.: Hidden fluid mechanics: learning velocity and pressure fields from flow visualizations. *Science* **367**, 1026–1030 (2020)
22. Raissi, M., Wang, Z., Triantafyllou, M.S., Karniadakis, G.E.: Deep learning of vortex-induced vibrations. *J. Fluid Mech.* **861**, 119–137 (2019)
23. Sun, L., Gao, H., Pan, S., Wang, J.X.: Surrogate modeling for fluid flows based on physics-constrained deep learning without simulation data. *Comput. Methods Appl. Mech. Eng.* **361**, 112732 (2020)
24. Zhu, Q., Liu, Z., Yan, J.: Machine learning for metal additive manufacturing: predicting temperature and melt pool fluid dynamics using physics-informed neural networks. *Comput. Mech.* **67**, 619–635 (2021)
25. Shukla, K., Di Leoni, P.C., Blackshire, J., Sparkman, D., Karniadakis, G.E.: Physics-informed neural network for ultrasound nondestructive quantification of surface breaking cracks. *J. Nondestr. Eval.* **39**, 1–20 (2020)
26. Yucesan, Y.A., Viana, F.A.: A physics-informed neural network for wind turbine main bearing fatigue. *Int. J. Progn. Health Manag.* **11**, 17–34 (2020)
27. Stiasny, J., Misyris, G.S., Chatzivasileiadis, S.: Physics-Informed Neural Networks for Non-linear System Identification applied to Power System Dynamics, Madrid (2021)
28. Misyris, G.S., Venzke, A., Chatzivasileiadis, S.: Physics-informed neural networks for power systems. In: 2020 IEEE Power and Energy Society General Meeting, Montreal (2020)
29. Kissas, G., Yang, Y., Hwuang, E., Witschey, W.R., Detre, J.A., Perdikaris, P.: Machine learning in cardiovascular flows modeling: predicting arterial blood pressure from non-invasive 4D flow MRI data using physics-informed neural networks. *Comput. Methods Appl. Mech. Eng.* **358**, 112623 (2020)
30. Sahli Costabal, F., Yang, Y., Perdikaris, P., Hurtado, D.E., Kuhl, E.: Physics-informed neural networks for cardiac activation mapping. *Front. Phys.* **8**, 42 (2020)
31. Jo, H., Son, H., Hwang, H.J., Kim, E.: Deep neural network approach to forward-inverse problems. *Math. NA* **15**, 247 (2020)
32. Yu, C.C., Tang, Y.C., Liu, B.D.: An adaptive activation function for multilayer feedforward neural networks. In: 2002 IEEE Region 10 Conference on Computers, Communications, Control and Power Engineering, Beijing (2002)
33. Shen, Y., Wang, B., Chen, F., Cheng, L.: A new multi-output neural model with tunable activation function and its applications. *Neural Process. Lett.* **20**, 85–104 (2004)
34. Dushkoff, M., Ptucha, R.: Adaptive activation functions for deep networks. *Electron. Imaging* **2016**, 1–5 (2016)
35. Jagtap, A.D., Kawaguchi, K., Karniadakis, G.E.: Adaptive activation functions accelerate convergence in deep and physics-informed neural networks. *J. Comput. Phys.* **404**, 109–136 (2020)
36. Jagtap, A.D., Kawaguchi, K., Karniadakis, G.E.: Locally adaptive activation functions with slope recovery for deep and physics-informed neural networks. *Proc. R. Soc. A.* **476**, 22–39 (2020)
37. Zhang, S., Song, Z.: An ethnic costumes classification model with optimized learning rate. In: Eleventh International Conference on Digital Image Processing, Guangzhou (2019)
38. Soize, C.: Exact stationary response of multi-dimensional non-linear Hamiltonian dynamical systems under parametric and external stochastic excitations. *J. Sound Vib.* **149**, 1–24 (1991)
39. Zhu, W.Q., Cai, G.Q., Lin, Y.K.: On exact stationary solutions of stochastically perturbed Hamiltonian systems. *Probab. Eng. Mech.* **5**, 84–87 (1990)

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.