ARC

Standard articles in Volume 1, 2023

Inaugural Issue of Aerospace Research Communications

EDITORS-IN-CHIEF

Prof. Yao Zheng

School of Aeronautics
and Astronautics,
Zhejiang University,
Hangzhou, China

Prof. Feng Liu

Department of
Mechanical & Aerospace
Engineering, University
of California, Irvine, USA

frontiers | Publishing Partnerships

ZHEJIANG
UNIVERSITY PRESS

ZHEJIANG UNIVERSITY 1897

ZITAPS

Scope & Mission

Aerospace Research Communications (ARC) is an open access, peer-reviewed international journal covering all aspects of aeronautics and astronautics. The journal publishes original papers and review articles related to all fields of aerospace research, in both theory and practice. The scope is considerably wide, covering research achievements in flight vehicles, propulsion systems, and experimental equipment, including fluid mechanics, flight mechanics, solid mechanics, vehicle conceptual design, avionics, control, material engineering and mechanical manufacturing.

Advisory and Editorial Boards

Aerospace Research Communications' Advisory Board and Associate Editors provide strategic and professional advice to the Editors-in-Chief. The journal recognises the following Advisory Board members and Associate Editors for their continuous guidance and support of the journal.

Advisory Board Members

George Em Karniadakis
Division of Applied Mathematics, Brown University, Providence, RI, USA

Roland Lewis
School of Engineering, Swansea University, Swansea, United Kingdom

Tom Shih
School of Aeronautics and Astronautics, Purdue University, West Lafayette, IN, USA

Wei Shyy
Department of Mechanical and Aerospace Engineering, Hong Kong University of Science and Technology, Hong Kong, China

Zhigong Tang
China Aerodynamics Research and Development Center, Mianyang, Sichuan, China

Zhijian Wang
Department of Aerospace Engineering, University of Kansas, Lawrence, KS, USA

Yi-Min Xie
School of Civil, Environmental and Chemical Engineering, RMIT University, Melbourne, Victoria, Australia

Kun Xu
Department of Mathematics, Hong Kong University of Science and Technology, Hong Kong, China

Yimin Xuan
College of Energy and Power Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing, China

Mingwu Yuan
Department of Mechanics and Engineering Science, Peking University, Beijing, China

Di Zhu
College of Mechanical and Electrical Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing, China

Editors-in-Chief

Yao Zheng
School of Aeronautics and Astronautics, Zhejiang University, Hangzhou, China

Feng Liu
Department of Mechanical and Aerospace Engineering, University of California, Irvine, USA

Assistant Editor-in-Chief

Jiafeng Zou
School of Aeronautics and Astronautics, Zhejiang University, Hangzhou, China

Associate Editors

Weiqiu Chen
School of Aeronautics and Astronautics, Zhejiang University, Hangzhou, China

Fangsen Cui
Department of Engineering Mechanics, Institute of High Performance Computing, Connexis, Singapore

Zhenping Feng
School of Energy and Power Engineering, Xi'an Jiaotong University, Xi'an, China

Song Fu
School of Aerospace Engineering, Tsinghua University, Beijing, China

Xiao-Wei Gao
School of Astronautics and Aeronautics, Dalian University of Technology, Dalian, China

Zhenghong Gao
National Key Laboratory of Aerodynamic Design and Research, Northwestern Polytechnical University, Xi'an, China

Hui Hu
Department of Aerospace Engineering, Iowa State University, Ames, IA, USA

Xiaodong Li
School of Energy and Power Engineering, Beihang University, Beijing, China

Yueming Li
State Key Laboratory for Strength and Vibration of Mechanical Structures, Xi'an Jiaotong University, Xi'an, China

Hong Liu
School of Aeronautics and Astronautics, Shanghai Jiaotong University, Shanghai, China

Hua Meng
School of Aeronautics and Astronautics, Zhejiang University, Hangzhou, China

Ning Qin
Department of Mechanical Engineering, University of Sheffield, Sheffield, United Kingdom

Maurizio Quadrio
Dipartimento di Scienze e Tecnologie Aerospaziali, Politecnico di Milano, Milano, Italy

Eusebio Valero
ETSI Aeronáutica y del Espacio, Universidad Politécnica de Madrid, Madrid, Spain

Yue Yang
Department of Mechanics and Engineering Science, Peking University, Beijing, China

Yufeng Yao
Department of Engineering Design and Mathematics, University of the West of England, Bristol, United Kingdom

Inaugural Issue of Aerospace Research Communications

ARC eBook
Copyright Statement

The copyright in the text of individual articles in this eBook is the property of their respective authors or their respective institutions or funders. The copyright in graphics and images within each article may be subject to copyright of other parties. In both cases this is subject to a license granted to Frontiers.
The compilation of articles constituting this eBook is the property of Frontiers.
Each article within this eBook, and the eBook itself, are published under the most recent version of the Creative Commons CC-BY licence. The version current at the date of publication of this eBook is CC-BY 4.0. If the CC-BY licence is updated, the licence granted by Frontiers is automatically updated to the new version.
When exercising any right under the CC-BY licence, Frontiers must be attributed as the original publisher of the article or eBook, as applicable.
Authors have the responsibility of ensuring that any graphics or other materials which are the property of others may be included in the CC-BY licence, but this should be checked before relying on the CC-BY licence to reproduce those materials. Any copyright notices relating to those materials must be complied with.
Copyright and source acknowledgement notices may not be removed and must be displayed in any copy, derivative work or partial copy which includes the elements in question.
All copyright, and all rights therein, are protected by national and international copyright laws. The above represents a summary only. For further information please read Frontiers' Conditions for Website Use and Copyright Statement, and the applicable CC-BY licence.

ISSN 1664-8714
ISBN 978-2-8325-4836-3
DOI 10.3389/978-2-8325-4836-3

Aerospace Research Communications (ARC) is an open-access, peer-reviewed international journal covering all aspects of aeronautics and astronautics. The journal publishes original Research and Review articles related to all fields of aerospace research, in both theory and practice.

This eBook collates the articles published in the journal in its first year, the year 2023. To date the journal has published 11 articles of those 11 articles, 4 are for the Flow Control Special Issue and 7 are standard articles in the journal.

This is the first issued eBook from the journal and collates all the standard articles accepted in 2023, the first year of publication that makeup Volume 1. We are looking forward to completing the Flow Control Special Issue later in the year and the publication of the eBook of the Special Issue.

Table of contents

AEROSPACE RESEARCH COMMUNICATIONS
REVIEW
published: 16 February 2023
doi: 10.3389/arc.2023.11130

Deep Reinforcement Learning: A New Beacon for Intelligent Active Flow Control

Fangfang Xie, Changdong Zheng, Tingwei Ji, Xinshuai Zhang, Ran Bi, Hongjie Zhou and Yao Zheng*

School of Aeronautics and Astronautics, Zhejiang University, Hangzhou, Zhejiang, China

The ability to manipulate fluids has always been one of the focuses of scientific research and engineering application. The rapid development of machine learning technology provides a new perspective and method for active flow control. This review presents recent progress in combining reinforcement learning with high-dimensional, non-linear, and time-delay physical information. Compared with model-based closed-loop control methods, deep reinforcement learning (DRL) avoids modeling the complex flow system and effectively provides an intelligent end-to-end policy exploration paradigm. At the same time, there is no denying that obstacles still exist on the way to practical application. We have listed some challenges and corresponding advanced solutions. This review is expected to offer a deeper insight into the current state of DRL-based active flow control within fluid mechanics and inspires more non-traditional thinking for engineering.

Keywords: deep learning, reinforcement learning, active flow control, model-free, fluid dynamics

INTRODUCTION

Despite many successful research efforts in the past decades, modifying the dynamics of flows to induce and enforce desired behavior remains an open scientific problem. In many industrial fields, researchers have placed great expectations on flow control techniques for engineering goals [1–3], such as drag reduction, noise suppression, mixing enhancement, energy harvesting. Due to the aggravation of carbon emissions and the greenhouse effect, controlling transportation drag or aerodynamic lift has become increasingly imperative.

Driven by the urgent demand from industry, active flow control (AFC) is being developed rapidly to harvest benefits for aviation or marine. As shown in **Figure 1**, Boeing and NASA tested a pneumatic sweeping-jet-based active flow control system on the vertical tail of the modified Boeing 757 ecoDemonstrator in April 2015. Active flow control was used to enhance the control authority of the rudder by mitigating flow separation on it at high rudder deflection, and side slip angles, which provided the required level of rudder control authority from a physically smaller vertical tail [4]. Whether using fluidic [5], micro blowing [6] or plasma actuators [7], the critical problem of active flow control is to design a reasonable control policy. The predetermined open-loop manner is the most straightforward choice. Still, the external actuation might be invalid if the evolution deviates from expectations and there are no corrective feedback mechanisms to modify the policy to compensate [8, 9]. A practical alternative is to adopt the closed-loop control manner [10–12], where the response is continuous compared with the desired result. Specifically, the control output to the process is informed by the sensors recording the flow information, then modified and adjusted to reduce the deviation, thus forcing the response to follow the reference.

***Correspondence:**
Tingwei Ji
zjjtw@zju.edu.cn

Received: *14 December 2022*
Accepted: *19 January 2023*
Published: *16 February 2023*

Citation:
Xie F, Zheng C, Ji T, Zhang X, Bi R, Zhou H and Zheng Y (2023) Deep Reinforcement Learning: A New Beacon for Intelligent Active Flow Control. Aerosp. Res. Commun. 1:11130. doi: 10.3389/arc.2023.11130

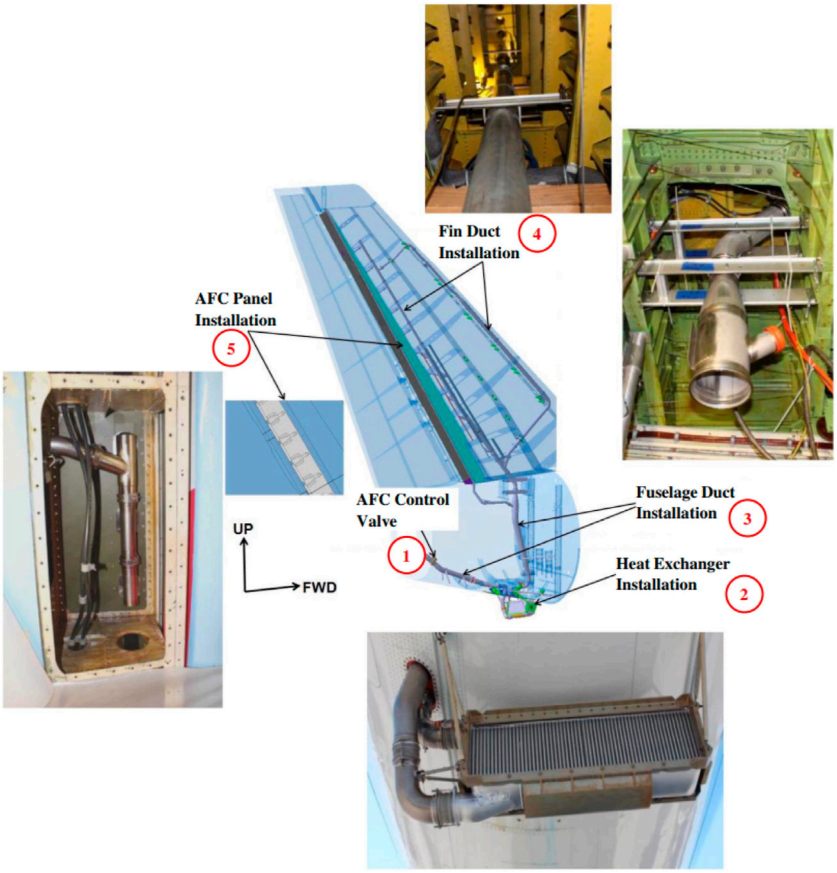

FIGURE 1 | A kind of flight-test AFC (Active Flow Control) system layout with photographs of hardware as installed [4].

In both ways, extensive work has been carried out by numerical simulations and experiments on exploring the nonlinear dynamics and underlying physical mechanisms of the controlled system with effective control law. For example, Xu et al. [13] investigated the separation control mechanism of a Co-flow Wall Jet, which utilized an upstream tangential injection and downstream streamwise suction simultaneously to achieve zero net-mass-flux flow control. It was found that the Co-flow wall Jet had a mechanism to grow its control capability with the increasing adverse pressure gradient. Sato et al. [14] conducted large-eddy simulations to study the separated flow control mechanism by a dielectric barrier discharge plasma actuator. From flow analysis, it was seen that an earlier and smoother transition case showed more significant improvements in the lift and drag coefficients. Moreover, the lift coefficient was improved since the actuation induced a large-scale vortex-shedding phenomenon.

While in many engineering applications, traditional large-scale physics-based models are intractable since it is required to evaluate the model to provide analysis rapidly and prediction [15–17]. The model reduction offers a mathematical foundation for accelerating physics-based computational models [18–20]. Alternatively, the model-free approach does not rely on any underlying model description of inputs to outputs. A significant advantage of a model-free manner in flow control is that it can avoid detailed identification of high-dimensional and nonlinear flow attractors, which would even shift during the regime. Moreover, with the development of machine learning techniques, it is possible to gain massive data. The control policy must grasp the embedded evolution rules and form data-driven logic. Namely, these model-free algorithms can simulate, extend and expand human intelligence to some degree.

As a critical branch of artificial intelligence, deep reinforcement learning (DRL) simplifies a stochastic dynamical system by using the framework of the Markov decision process (MDP) [22, 23]. DRL algorithms can explore and adjust control policies by interacting with the environment like a child, which gets a penalty when making mistakes. In a continuous process of trial and error, the control law in DRL learns how to get sweet lollipops (high reward) and avoid penalties. Besides, DRL utilizes the artificial neural network(ANN) as a function approximator [24]. Based on the such setting, the DRL is embedded as a state representation technology, which makes it possible to deal with high-dimensional complex problems, like Go, StarCraft, Robotics [21, 25–27]. As shown in **Figure 2**, Vinyals et al. [21] adopts a multi-agent reinforcement learning algorithm to train an agent named AlphaStar, in the full game of StarCraft II, through a series of online games against a human player. AlphaStar was rated at

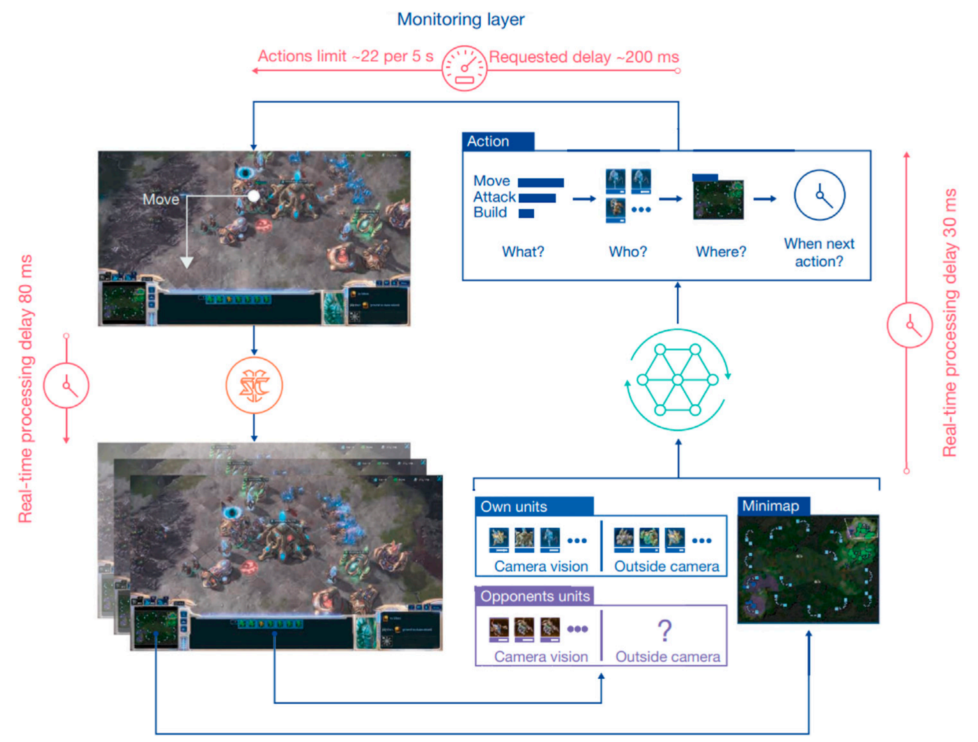

FIGURE 2 | Grandmaster level in StarCraft II using multi-agent reinforcement learning. The agent observes the game through an overview map and a list of units. To act, the agent outputs what action type to issue (for example, build), who it is applied to, where its objectives are, and when the next action will be issued [21].

Grandmaster level for all three StarCraft races and above 99.8% of officially ranked human players. Similarly, DRL has highlighted its strong potential in fluid dynamics applications, including drag reduction, collective swimming, and flight control. Many excellent pioneers have reviewed the work related to the application of reinforcement learning in active flow control [28–32]. The present paper will further review the latest developments and demonstrate some challenges to constructing robust and effective DRL-based active flow policies.

What needs to be mentioned is that there are many other algorithms still active on the stage that achieve great performance and have more potentials to exploit as well, such as a gradient-enriched machine learning control [33], Bayesian optimization control [34], RBF-NN adaptive control [35], ROM-based control [36]. In some work, reinforcement learning has also been compared with some algorithms, such as Bayesian optimization [37], genetic programming [38, 39], Lipschitz global optimization [39], etc. In particular, genetic programming algorithms, closely related to reinforcement learning, can achieve optimal decisions under the condition of an unknown model as well. Although the randomness of exploration brings low efficiency, evolutionary algorithms such as genetic programming are very popular in some problems like multi-objective optimization and global optimization [40, 41]. Not only in the field of flow control but algorithms combining evolutionary algorithms and reinforcement learning have also always been expected [42, 43]. For brevity, detailed comparison and discussion about the above model-free algorithms are not considered in this review, readers can refer to these papers [29, 39].

The rest of this review is organized as follows: Section *Deep Reinforcement Learning* presents some basic concepts and algorithms of DRL. Section *Applications of DRL-based Active Flow Control* offers the application of DRL on fluids problems, and Section *Challenges on DRL-Based Active Flow Control* shows some innovations and solutions to fluids problems to make DRL-based active flow control more effective. Finally, a summary and potential directions of DRL-based active flow control are drawn in Section *Conclusion*.

DEEP REINFORCEMENT LEARNING

This section introduces some basics concepts of typical reinforcement learning framework, and popular deep reinforcement learning algorithms, such as proximal policy optimization (PPO) [44] and soft actor-critic (SAC) [45]. First, the general terms and concepts are presented in Section *Markov Decision Process*. The optimization methods of reinforcement learning for policies are generally divided into Section *Value-Based Methods* and Section *Policy-Based Methods*. Either of the two methods has the ability to find the optimal control strategy. Still, their respective shortcomings must be addressed, like a relatively large gradient variance in policy method, etc. [46].

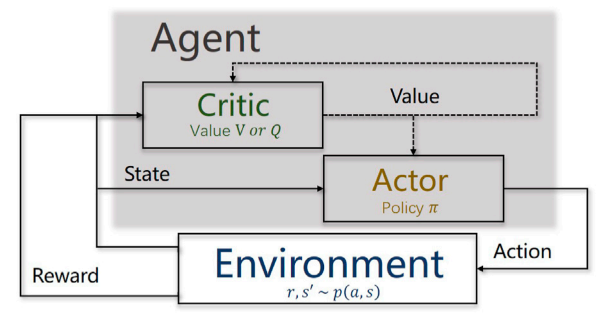

FIGURE 3 | The Markov decision process of reinforcement learning. According to the reward, the agent learns to evaluate the value of each action and gain an optimal policy that maximizes the expected return.

The actor-critic method discussed in Section *Actor-Critic Methods*, aims to combine the advantages of both ways and search for optimal policies using low-variance gradient estimates, which has been one of the most popular frameworks in reinforcement learning. Furthermore, two advanced deep reinforcement learning algorithms on the actor-critic framework are detailed in Section *Advanced Deep Reinforcement Learning Algorithms*.

Markov Decision Process

Reinforcement learning solves problems modeled as Markov decision processes (MDPs) [47]. The system state s, action a, reward r, time t, and reward discount factor γ are the basic concepts of MDPs. Under the intervention of action a, the system state s is transferred with a reward r. Reward r defines the goodness of action, and this transition is only related to action a and current state s, which refers to the memoryless property of a stochastic process. Mathematically, it means

$$p(s_{t+1}|s_t, a_t, s_{t-1}, a_{t-1}, \ldots, s_0, a_0) = p(s_{t+1}|s_t, a_t), \quad (1)$$

$$p(r_t|s_t, a_t, s_{t-1}, a_{t-1}, \ldots, s_0, a_0) = p(r_t|s_t, a_t). \quad (2)$$

Markov property helps simplify complex stochastic dynamics that are difficult to model in practice. The role of reinforcement learning is to search for an optimal policy telling which action to take in such an MDP. Specifically, the policy maps from state s and action a to the action probability distribution π, as $a_t \sim \pi(\cdot|s_t)$. In the discounted reward setting, the cost function J is equal to the expected value of the discounted sum of rewards for a given policy π; this sum is also called the expected cumulative reward

$$J(\pi) = E_{\tau\sim\pi}\left[\sum_{t=0}^{\infty}\gamma^t r_t\right]. \quad (3)$$

where the trajectory $\tau = (s_0, a_0, r_0, s_1, a_1, r_1, s_2, \cdots)$ is highly correlated to the policy π. And $\gamma \sim [0, 1)$ denotes the reward discount factor.

Over time, several RL algorithms have been introduced to search for an optimal policy with the greatest expected cumulative reward. They are divided into three groups [47]: actor-only, critic-only, and actor-critic methods, where the words actor and critic are synonyms for the policy and value function (policy-based and value-based), respectively. These algorithms are detailed in the following sections.

The Markov decision process can also be seen as a continuous interaction between the agent and the environment. The agent is a decision-maker that can sense the system state, maintain policies, and execute actions. Everything outside of the agent is regarded as the environment, including system state transition and action scoring [48], as shown in **Figure 3**. During the interaction, the agent dynamically adjusts the policy to learn behaviors with the most rewards.

Value-Based Methods

The value-based methods, such as Q-learning [49], SARSA [50], focus on the estimation of state value V^{π} or state-action value Q^{π} under the specified policy π, defined as:

$$V^{\pi}(s) = E_{a_t\sim\pi(\cdot|s_t)}\left[\sum_{t=0}^{\infty}\gamma^t r_t(s_t, a_t)|s_0 = s\right], \quad (4)$$

or

$$Q^{\pi}(s, a) = E_{a_t\sim\pi(\cdot|s_t)}\left[\sum_{t=0}^{\infty}\gamma^t r_t(s_t, a_t)|s_0 = s, a_0 = a\right]. \quad (5)$$

As its name suggests, it represents the ”value” of a state or state-action, which is mathematically the expected value of the discounted sum of rewards with initial state s or initial state-action $s - a$ for a given policy π. The state value $V^{\pi}(s_t)$ depends on the state s_t and assumes that the policy π is followed starting from this state. And the state-action value $Q^{\pi}(s, a)$ has specified additional action a_t, and the future selection of actions is under policy π.

According to the Markov property of the decision-making process, the Bellman equation, a set of linear equations, is proposed to describe the relationship among values of all states:

$$V^{\pi}(s) = E_{a\sim\pi(\cdot|s), s'\sim p(\cdot|s,a)}\left[r(s, a) + \gamma V^{\pi}(s')\right]. \quad (6)$$

where p represents the system dynamic. The values of states rely on the values of some other states or themselves, which is related to an important concept called bootstrapping.

Since state values can be used to evaluate policies, they can also define optimal policies. If $V(\pi_1) > V(\pi_2)$, π_1 is said better than π_2. Furthermore, if a policy is better than all the other possible policies in all states, then this policy is optimal. Optimality for state value function is governed by the Bellman optimality equation (BOE)

$$V^{*}(s) = \max_{\pi}\ E_{a\sim\pi, s'\sim p(\cdot|s,a)}\left[r(s, a) + \gamma V^{*}(s')\right]. \quad (7)$$

It is a nonlinear equation with a nice contraction property, and the contraction mapping theorem is applied to prove its convergence. The solution to the BOE always exists as the unique optimal state value, which is the greatest state value that can be achieved by any initial policy [47].

Similarly, the Bellman equation and Bellman optimality equation have expressions in terms of state-action values as

$$Q^{\pi}(s,a) = E_{s'\sim p(\cdot|s,a),a'\sim\pi(\cdot|s')}\left[r(s,a) + \gamma Q^{\pi}(s',a')\right], \tag{8}$$

and

$$Q^{\star}(s,a) = E_{s'\sim p(\cdot|s,a),a'\sim\pi(\cdot|s')}\left[r(s,a) + \underset{\pi}{max}\ \gamma Q^{\star}(s',a')\right]. \tag{9}$$

In practice, the state-action value plays a more direct role than the state value when attempting to find optimal policies. The Bellman optimality equation is a particular form of the Bellman equation. The corresponding state value is the optimal state value, and the related implicit optimal policy can be drawn from the greatest values. For example, the optimal policy $\pi^{\star}$ is calculated by using an optimization procedure over the value function:

$$\pi^{\star} = \underset{\pi}{argmax}\ Q^{\star}_{a\sim\pi}(s,a) \tag{10}$$

Policy-Based Methods

The value-based methods use value functions and no explicit functions for the policy. And the policy-based methods, such as REINFORCE [51], and SRV [52], do not utilize any form of a stored value function but work with a parameterized family of policies and optimize the objective function J directly over the parameter space. Assuming that the policy is represented by a parameterized function denoted as $\pi(a|s,\ \theta)$, which is differentiable concerning parameter vector θ, the gradient of the objective function J is described as

$$\nabla_{\theta} J = \frac{\partial J}{\partial \pi_{\theta}}\frac{\partial \pi_{\theta}}{\partial \theta}. \tag{11}$$

The objective function has different metrics leading to different optimal policies. There are many metrics candidates in the policy-based methods, such as average state value, average one-step reward. If the matric is the expected cumulative reward (6), it can apply gradient descent algorithm on policy parameter θ to gradually improve the performance of the policy π_{θ}, and the gradient is calculated as

$$\nabla_{\theta} J(\pi_{\theta}) = E_{\tau\sim\pi_{\theta}}\left[\sum_{t=0}^{\infty}\nabla_{\theta}\left(\log\pi_{\theta}(a_t|s_t)\right)Q^{\pi_{\theta}}(s_t,a_t)\right]. \tag{12}$$

Though in this form, the state-action value Q is called, which can be approximated by Monte Carlo estimation $Q^{\pi_{\theta}}(s_{t'},a_{t'}) = \sum_{t'=t}^{\infty}\gamma^{t'-t}r(s_{t'},a_{t'})$ in REINFORCE algorithm. Based on the gradient, the parameter θ is then adjusted in the direction of this gradient:

$$\theta_{t+1} = \theta_t + \alpha_t \nabla J_t. \tag{13}$$

where α is the optimization rate. Every update on parameter θ seeks for an increasement on the objective function $J(\pi_{\theta_{t+1}}) \geq J(\pi_{\theta_t})$. The main advantage of policy-based methods is their strong convergence property, which is naturally inherited from gradient descent methods. Convergence is obtained if the estimated gradients are unbiased and the learning rates α_k satisfy [47]

$$\sum_{t=0}^{\infty}\alpha_k = \infty, \qquad \sum_{t=0}^{\infty}{\alpha_k}^2 < \infty\ . \tag{14}$$

Different from the value-based methods, the policy π_{θ} is explicit, and actions are directly sampled from the optimal parameterized policy:

$$a^{\star} \sim \pi(\cdot|s,\theta^{\star}) \tag{15}$$

Actor-Critic Methods

Value-based methods rely exclusively on value function approximation and have a low variance in the estimates of expected returns. However, when nonlinear function approximators represent value functions, the approximation bias would lead to non-convergence during numerical iterations [53, 54]. The purpose of replay buffer and target value network techniques in Deep Q-learning Network [26, 55] algorithm ameliorate the above situation well, which achieves significant progress in Atari games. Besides, value-based methods must resort to an optimization procedure in every state encountered to find the action leading to an optimal value, which can be computationally expensive for continuous state and action spaces.

Policy-based methods work with a parameterized family of policies and optimize the objective function directly over the parameter space of the policy. One of this type's advantages is handling continuous state and action spaces with higher efficiency in terms of storage and policy searching [56]. However, a possible drawback is that the gradient estimation may have a significant variance due to the randomness of reward over time [56, 57]. Furthermore, as the policy changes, a new gradient is estimated independently of past estimates. Hence, there is no ”learning” in accumulating and consolidating older information.

Actor-critic methods aim at combining the value-based and policy-based methods [46, 58]. A parameterized function is proposed based on the value-based methods to learn state value V or state-action value Q as a critic. And the policy is not inferred from the value function. It uses a parameterized function as actor π_{θ}, which has good convergence properties in contrast with value-based methods and brings the advantage of computing continuous actions without the need for optimization procedures on a value function. At the same time, the critic supplies the actor with low-variance value knowledge $\hat{V}_{\phi}^{\pi_{\theta}}$ or $\hat{Q}_{\phi}^{\pi_{\theta}}$ and reduces the oscillation in the learning process.

Figure 4 shows the schematic structure of actor-critic methods. The agent consists of the critic and actor parts, which interact with the environment as presented in Section *Markov Decision Process*. During the collection of rewards, the critic is responsible for estimating value functions with parameterized function approximators like deep neural networks. The actor-critic methods often follow the idea of the bootstrap method to evaluate value function, whose objective function on state-action value is

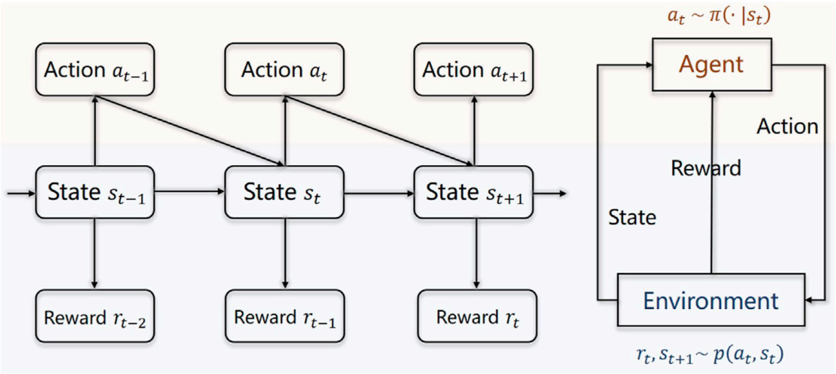

FIGURE 4 | The actor-critic methods framework.

$$J(\phi) = E_{\tau\sim\pi_\theta}\left[\frac{1}{2}\left(\sum_{t=0}^{\infty} r(s_t,a_t) + \gamma\hat{V}^{\pi_\theta}(s_{t+1};\phi) - \hat{V}^{\pi_\theta}(s_t;\phi)\right)^2\right], \tag{16}$$

or on state value

$$J(\phi) = E_{\tau\sim\pi_\theta}\left[\frac{1}{2}\left(\sum_{t=0}^{\infty} r(s_t,a_t) + \gamma\hat{Q}^{\pi_\theta}(s_{t+1},a_{t+1};\phi) - \hat{Q}^{\pi_\theta}(s_t,a_t;\phi)\right)^2\right]. \tag{17}$$

benefited from the bootstrap method, the estimation of value function $\hat{V}^{\pi_\theta}_{\phi}$ or $\hat{Q}^{\pi_\theta}_{\phi}$ is low-variance, which is a good choice for the gradient of actor's objective function

$$\nabla_\theta J(\pi_\theta) = E_{\tau\sim\pi_\theta}\left[\sum_{t=0}^{\infty}\nabla_\theta\left(\log\pi_\theta(a_t|s_t)\right)\hat{Q}^{\pi_\theta}(s_t,a_t)\right]. \tag{18}$$

It is worth noting that the value-based or policy-based methods are core reinforcement learning algorithms and have played a vital role. Many techniques, like delay policy updates [59], replay buffer [26], and target value network [55], is proposed to improve the efficiency of the algorithm. The actor-critic methods are the improvement of the policy-based methods in reducing the sample variance or the expansion of the value-based method in the continuous state-action space problem. Compared to value-based or policy-based methods, the actor-critic method shows many friendly properties, a popular template for researchers developing more advanced algorithms.

Advanced Deep Reinforcement Learning Algorithms

With the deepening of research, many advanced deep reinforcement learning algorithms on the actor-critic framework have been proposed, such as PPO [44], SAC [45], TD3 [59], DDPG [60] and so on. This section presents Proximal Policy Optimization (PPO) algorithm and Soft Actor-Critic (SAC) algorithm. Considering the length of the article, a brief introduction is given. For more details and principles, interested readers are suggested to refer to the original papers [44, 45].

Proximal Policy Optimization (PPO)

Proximal policy optimization (PPO) is a robust on-policy policy gradient method for reinforcement learning proposed by OpenAI [44]. Standard policy gradient methods perform one gradient update per data sampling. Still, PPO utilizes a novel objective function that enables multiple epochs of minibatch updates by importance sampling trick, which improves sample efficiency.

Typical trust-region methods constrain policy updates to a trust region, ensuring that the entire policy update process is monotonous. PPO suggests using a KL penalty instead of a constraint to solve the unconstrained optimization problem. The algorithm is based on an actor-critic framework, and its actor objective is modified as

$$J^{\theta^k}_{PPO}(\pi_\theta) = \sum_{s_t,a_t}\frac{p_\theta(a_t|s_t)}{p_{\theta^k}(a_t|s_t)}A^{\theta_k}(s_t,a_t) - \beta KL(\theta,\theta_k) \tag{19}$$

where k is reuse times on single batch of data; $A^\theta(s_t, a_t) = Q^\theta(s_t, a_t) - V^\theta(s_t, a_t)$ is the advantage function to reduce variance; β is the penalty factor of KL divergence.

PPO algorithm has the stability and reliability of trust-region methods [61]. But it is much simpler to implement, requiring only a few lines of code change to a vanilla policy gradient (VPG) implementation [47], which is applicable in general settings and has better overall performance.

Soft Actor-Critic (SAC)

Soft Actor-critic is an off-policy actor-critic deep RL algorithm based on the maximum entropy reinforcement learning framework [45]. In this framework, the actor aims to maximize the standard ultimate reward while also maximizing entropy. Maximum entropy reinforcement learning alters the RL objective [62], though the original aim can be recovered using a temperature parameter. More importantly, the maximum entropy formulation substantially improves exploration and robustness: maximum entropy policies are robust in the face of model and estimation errors, and they enhance exploration by acquiring diverse behaviors [45].

The maximum entropy objective (see, e.g., (Ziebart, 2010) generalizes the standard objective by augmenting it with an

TABLE 1 | Applications of DRL-based active flow control.

Category	Time	References	Algorithm	Objective
Flow stability	2018	[63]	DDPG	Control the Karman vortex shedding
Flow stability	2021	[37]	SAC,AL	Suppress the vortex-induced vibration
Flow stability	2021	[64]	PPO	Mitigate the hydrodynamic signature
Flow stability	2021	[65]	PPO	Enhance the vortex-induced vibration
Hydrodynamic Drag	2017	[66]	ML Actor-Critic	Build an implicit model and reduce drag
Hydrodynamic Drag	2019	[67]	PPO	Stabilize vortex alley and reduce drag
Hydrodynamic Drag	2021	[68]	PPO	Stabilization and drag reduction on DMD
Hydrodynamic Drag	2020	[69]	PPO	Control with small rotating cylinders
Hydrodynamic Drag	2020	[70]	PPO	Control over a range of Re numbers
Hydrodynamic Drag	2022	[71]	PPO	Control in weakly turbulent conditions
Hydrodynamic Drag	2022	[72]	PPO	Control the flow with Re = 1000
Hydrodynamic Drag	2020	[73]	TD3	Maximize the power gain efficiency
Hydrodynamic Drag	2022	[74]	single-step PPO	Control the wake of a 3D bluff body
Aerodynamic Performance	2020	[75]	PPO	Control on NACA0012 in pulsating inflow
Aerodynamic Performance	2022	[76]	PPO	Control lift on distributed sensors
Aerodynamic Performance	2020	[77]	DQN	Control flow separation
Aerodynamic Performance	2020	[78]	Ape-X DQN	Control flow separation
Aerodynamic Performance	2022	[79]	ApeX-DQN, ABN	Supress separation and visualize data area
Behavior Patterns	2021	[80]	DRQN	Study the behaviors of self-propelled fish
Behavior Patterns	2021	[81]	V-RACER	Learn escape under energy constraints
Behavior Patterns	2022	[82]	Q-learning	Explore collective locomotions
Behavior Patterns	2018	[83]	Q-learning	Learn glider soaring
Behavior Patterns	2019	[84]	RACER	Identify gliding and landing strategies

entropy term, such that the optimal policy additionally aims to maximize its entropy at each visited state:

$$J_{SAC}(\pi_\theta) = \sum_{t=0}^{T-1} E_{(s_t,a_t)\sim\rho_\pi}\left(r_t(s_t,a_t) + \alpha H(\pi(\cdot|s_t))\right) \tag{20}$$

where α is the temperature parameter determining the relative importance of the entropy term against the reward. H is the entropy of policy π.

In Ref. [45], it empirically showed that it matched or exceeded the performance of state-of-the-art model-free deep RL methods, including the off-policy TD3 algorithm and the on-policy PPO algorithm without any environment-specific hyperparameter tuning. And the real-world experiments indicated that soft actor-critic was robust and sample efficient enough for robotic tasks learned directly in the real world, such as locomotion and dexterous manipulation.

APPLICATIONS OF DRL-BASED ACTIVE FLOW CONTROL

For DRL-based active flow control, it is essential to construct a Markov Decision Process (MDP) from the flow phenomenon. If the state of flow and the reward of actions are well-selected, the reinforcement learning technique can solve the Bellman equation with high proficiency. Moreover, the artificial neural network applied to the above deep reinforcement learning algorithms has good approximation ability in high-dimensional space with less complexity than typical polynomial fitting. It has proven its advantages in many flow applications like prediction.

In the past 6 years, we have also seen many efforts to introduce deep reinforcement learning into the flow control field. From the initial tabular, e.g., Q-learning, to advanced deep learning, like Soft Actor-Critic (SAC) and Proximal Policy Optimization (PPO), DRL algorithms are equipped more smartly, and novel control phenomena have been explored. This section reviews recent flow control applications based on deep reinforcement learning, including Section *Flow Stability*, Section *Hydrodynamic Drag*, Section *Aerodynamic Performance*, and Section *Behavior Patterns*. For conciseness, a summary table is constructed as **Table 1**.

Flow Stability

Flow instability and transition to turbulence are widespread phenomena in engineering, and the natural environment [85–87]. The flow around a circular cylinder can be considered a prototype of the bluff body wakes, which is involved with various instability. In the cylinder wake, the transition from steady to periodic flow is marked by a Hopf bifurcation with critical Reynolds Re = 47, which is known as the first instability [88]. Three-dimensional fluctuations for higher Reynolds numbers further superimpose this vortex shedding. The onset of three-dimensionality occurs at the critical Reynolds number of Re = 175. These periodic behaviors can induce fluctuating hydrodynamic force on the bluff body, leading to vortex-induced vibrations, which can bring the challenge to structural fatigue performance or provide an opportunity for energy utilization [89, 90].

As early as 2018, Koizumi et al. [63] applied a deep deterministic policy gradient (DDPG) algorithm to control the Karman vortex shedding from a fixed cylinder. Compared with conventional model-based feedback control, the result of the DDPG also shows better control performance with reduced

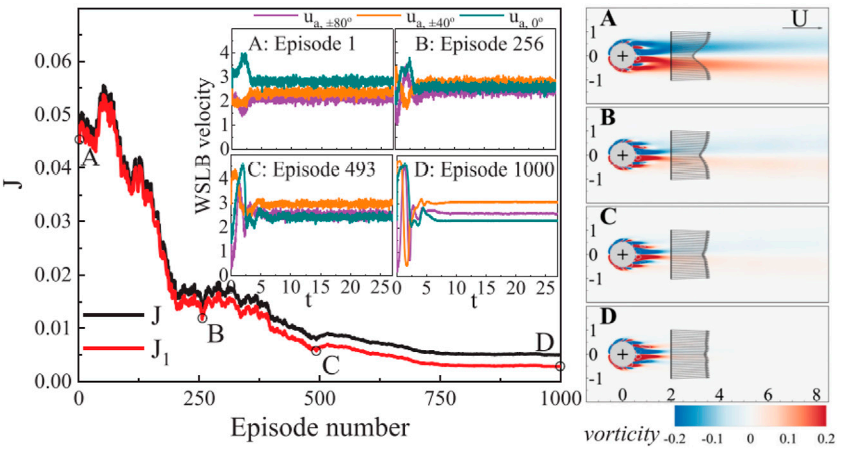

FIGURE 5 | Learning process represented by the variation of objective function values against episode number. Ren and Wang adopt a group of windward-suction-leeward-blowing (WSLB) actuators to stabilize both the wake of a fixed and flexible cylinder. Left: Learning process represented by the variation of cost function values against episode number. The four insets show WSLB actuations generated by the DRL agent at different stages of learning. Right: Instantaneous wake patterns and measured velocity profiles at the four selected stages [64].

lift. Later on, for the vortex-induced vibrations, some scholars have also tried deep reinforcement learning to eliminate them. By constructing a spring-mounted cylinder model, Zheng et al. [37] proposed a deep reinforcement learning active flow control framework to suppress the vortex-induced vibration of a cylinder immersed in uniform flow by a pair of jets placed on the poles of the cylinder as actuators. In training, the SAC agent is fed with a lift-oriented reward, which successfully reduces the maximum vibration amplitude by 81%. Ren et al. [64] further adopted windward-suction-leeward-blowing (WSLB) actuators to control the wake of an elastically mounted cylinder. They encoded velocity information in the VIV wake into the reward function of reinforcement learning, aiming at keeping pace with the stable flow. Only a 0.29% deficit in streamwise velocity is detected, which is a 99.5% reduction from the uncontrolled value, and the learning process is shown in **Figure 5**. Unlike the previous two cases, instead of reducing the intensity of the vortex shedding caused by the first instability, the essence of reinforcement learning flow control in Ren's work is to eliminate the vortex shedding caused by the first instability, which is the origin of the vortex-induced vibrations.

For the energy utilization of vortex-induced vibration, Mei et al. [65] proved that the performance of the active jet control strategy established by DRL for enhancing VIV is outstanding and promising. It is shown that the ANN can successfully increase the drag by 30.78% and the magnitude of fluctuation of drag and lift coefficient by 785.71% and 139.62%, respectively. Furthermore, the net energy output by VIV with jet control increased by 357.63% (case of water) compared with the uncontrolled situation.

Hydrodynamic Drag

In terms of hydrodynamic drag, it is the primary concern for modern hydrodynamic design. Namely, the potential benefits of an effective closed-loop active flow control for drag are highlighted for energy and transportation.

Like the flow stability topic, the early active flow control applications of reinforcement learning are within the deep neural network. Pivot and Mathelin [66] proposed a reinforcement learning active flow control framework whose value function and policy function are approximated with local linear models. Taking embedding and delayed effect of the action into consideration, the system's state is constructed carefully, and 17% of cylinder drag reduction is obtained by RL-controlled self-rotating. Then the artificial neural network technique is introduced into the field of active flow control on reducing hydrodynamics drag, which replaces the original way by using elaborately-designed state representation for the flow system. Rabault [67] was the first scholar to apply an artificial neural network trained through a deep reinforcement learning agent to perform active flow control for cylinder drag reduction. At Reynolds number of Re = 100, the drag can be reduced by approximately 8% shown in **Figure 6**. It was seen that the circulation area is dramatically increased, and the fluctuation of vortex shedding is reduced. Their forward-looking work provided a template for DRL-based active flow control in the fluid mechanics. Qin [68] modified the reward function with dynamic mode decomposition (DMD). With the data-driven reward, the DRL model can learn the AFC policy through the more global information of the field and the learning was improved. Xu [69] used DRL to control small rotating cylinders on the back of the controlled cylinder and achieved drag reduction, which successfully illustrated the adaptability of DRL to actuators in AFC problems.

To investigate the generalization performance of DRL, Tang [70] trained a PPO agent in a learning environment supporting four flow configurations with Reynolds numbers of 100, 200, 300, and 400, which effectively reduced the drag for any previously unrecognized value of the Reynolds number between 60 and 400.

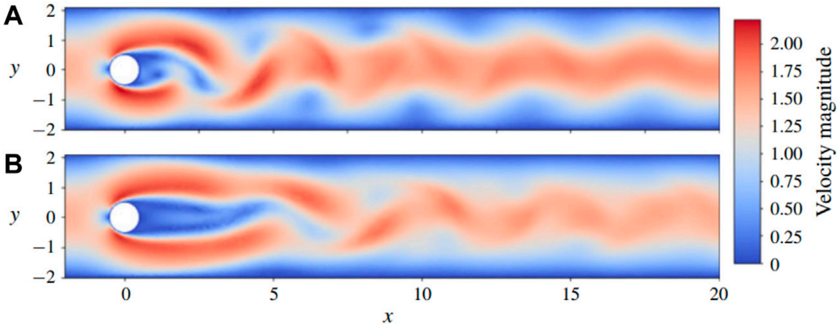

FIGURE 6 | Comparison of representative snapshots of the velocity magnitude in the case without actuation **(A)** and with active flow control **(B)**. The lower panel corresponds to the established pseudo-periodic modified regime, attained after the initial transient control [67].

Ren [71] extended the flow condition to a weakly turbulent state with Re = 1,000 and proved that the PPO agent can still find effective control policies but requires much more episodes in the learning. For larger Reynolds numbers like 2000, Varela [72] finds significantly different forms of nature in the control strategies from those obtained at lower Re. More importantly, the cross applications of agents both from Re = 1000 and Re = 1000 were conducted in a flow with Re = 2000. Two similar results with different natures of control strategies may indicate that the Reynolds number regime (Re = 2000) belongs to a transition towards a nature-different flow which would only admit a high-frequency actuation strategy to obtain drag reduction. The deep insight is waiting for future simulations on higher Reynolds numbers.

Later on, Fan et al. [73] demonstrated the feasibility and effectiveness of reinforcement learning (RL) in bluff body flow control problems in simulations and experiments by automatically discovering active control strategies for drag reduction in turbulent flow with two small rotating cylinders. It is a crucial step to identify the limitations of the available hardware when applying reinforcement learning in a real-world experiment. After an automatic sequence of tens of towing experiments, the RL agent is shown to discover a control strategy comparable to the optimal policy found through lengthy, systematically planned control experiments. Meanwhile, the flow mechanism for the drag reduction was also explored. Through verification by three-dimensional simulations, as seen in **Figure 7**, due to the gap between the large and small cylinders, a jet is informed within the hole, causing the change of flow topology in the cylinder wake. Therefore, compared to the non-rotating case, the pressure on the rear cylinder surface recovered to a negative value with a smaller magnitude, leading to a significant pressure drag reduction. Moreover, with the platform of a wind tunnel, Amico et al. [74] trained an agent capable of learning control laws for pulsed jets to manipulate the wake of a bluff body at Reynolds number Re = 10^5. It is the first application of a single-step DRL in an experimental framework at large values of the Reynolds number to control the wake of a three-dimensional bluff body.

Aerodynamic Performance

To make aviation greener, many efforts have been made to improve aircraft's aerodynamic performance to design a more effective, environmentally friendly air transport system [91]. Active flow control technology can potentially deliver breakthrough improvements in the aerodynamic performance of the aircraft, like enhanced lift; reduced drag; controlled instability; and reduced noise or delayed transition. This subsection will present recent studies on DRL-based active flow control for aerodynamic performance improvement.

Several scholars have applied reinforcement algorithms to achieve effective active flow strategies through numerical simulations or wind tunnel experiments to enhance lift and reduce drag. Wang [75] used the PPO algorithm on the synthetic jet control of flows over a NACA0012 airfoil at Re = 3,000 and embedded lift information into the reward function. The DRL agent can find a valid control policy with energy conservation by 83% under a combination of two different frequencies of inlet velocity. Guerra-Langan et al. [76] trained a series of reinforcement learning (RL) agents in simulation for lift coefficient control, then validated them in wind tunnel experiments. Specifically, an ANN aerodynamic coefficients estimator is trained to estimate lift and drag coefficients using pressure and strain sensor readings together with pitch rate. Results demonstrated that hybrid RL agents that use both distributed sensing data and conventional sensors performed best across the different tests.

To suppress or delay flow separation [92], Shimomura and Sekimoto [77] proposed a practical DRL-based flow separation control framework and investigated the plasma control effectiveness on a NACA0015 airfoil in a low-speed wind tunnel at a Reynolds number of 63000. As seen in **Figure 8**, based on deep Q-network(DQN), the closed-loop control keeps the flow attached and preserves it for a longer time by periodically switching the actuator on and off. With distributed executors and priority experience playback, they proved that the Ape-X DQN algorithm is more stable during training than the DQN algorithm in such plasma control problem [78]. Moreover, Takada et al. [79] investigated the performance of plasma control on the NACA0012 airfoil in compressible fluid numerical simulation,

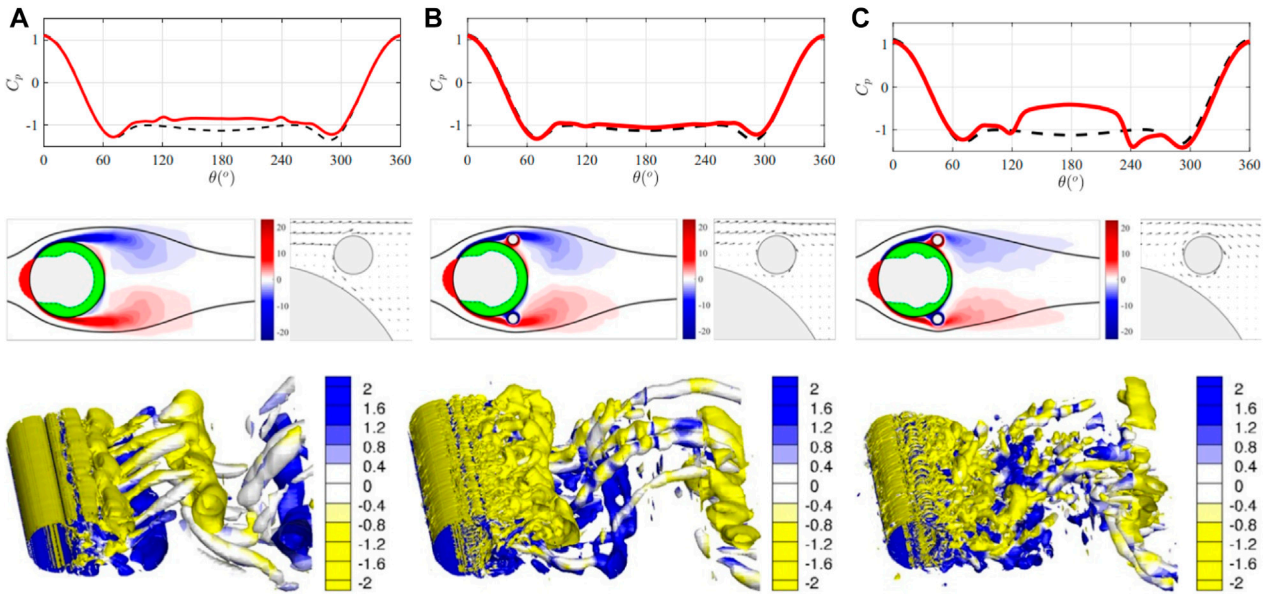

FIGURE 7 | Visualization of the vortical flow at three different stages of training. **(A)** Before training (both small cylinders are held still). **(B)** After 100 episodes (both small cylinders rotate at a medium speed). **(C)** After 500 episodes (both small cylinders rotate at about the maximum speed). (**A–C**, Top) Local pressure coefficient on the cylinder surface as a function of angle (*θ*), with the front stagnation point as zero degrees. The coefficient is shown by the red lines, with black dashed lines representing the reference coefficient of a single cylinder. (**A–C**, Middle Left) the z component of vorticity averaged spanwise and in time with the green/red area indicating the magnitude of negative/positive pressure on the main cylinder. (**A–C**, Middle Right) Velocity field near the small upper cylinder. (**A–C**, Bottom) Three-dimensional vortices. Note that to plot B, we restart the simulation from the flow snapshot saved at episode 100, keep the control cylinders rotating at the same speeds as those of episode 100, and continue to simulate over two vortex-shedding periods; similar procedures are performed to obtain C [73].

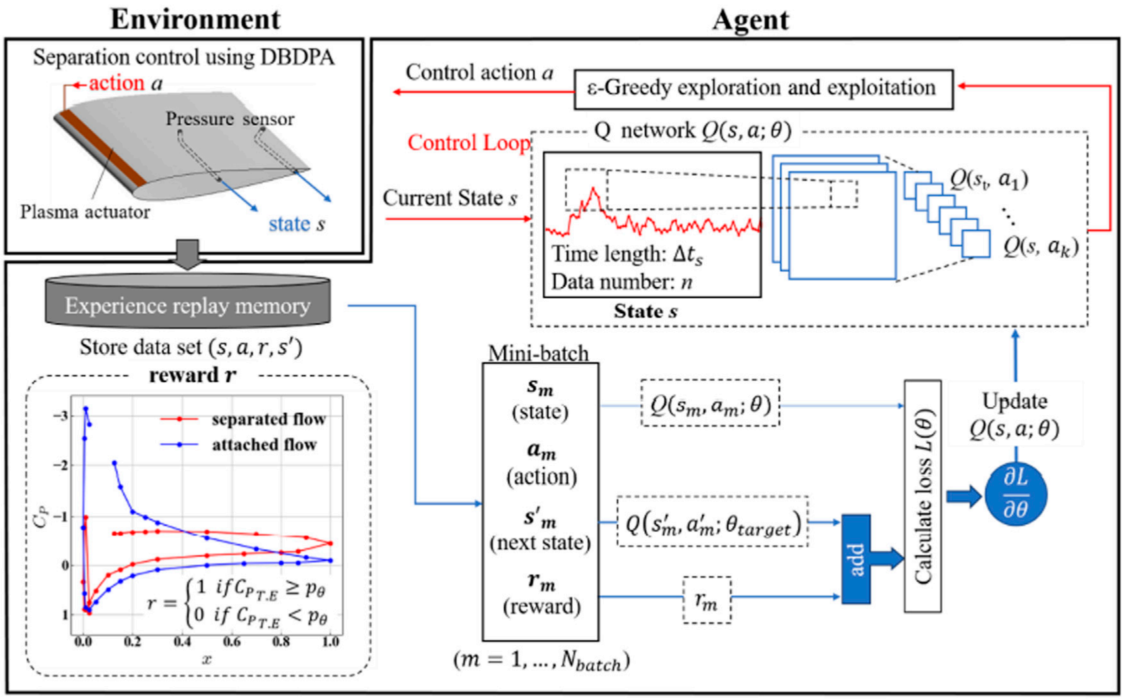

FIGURE 8 | An effective DRL-based flow separation control framework [77].

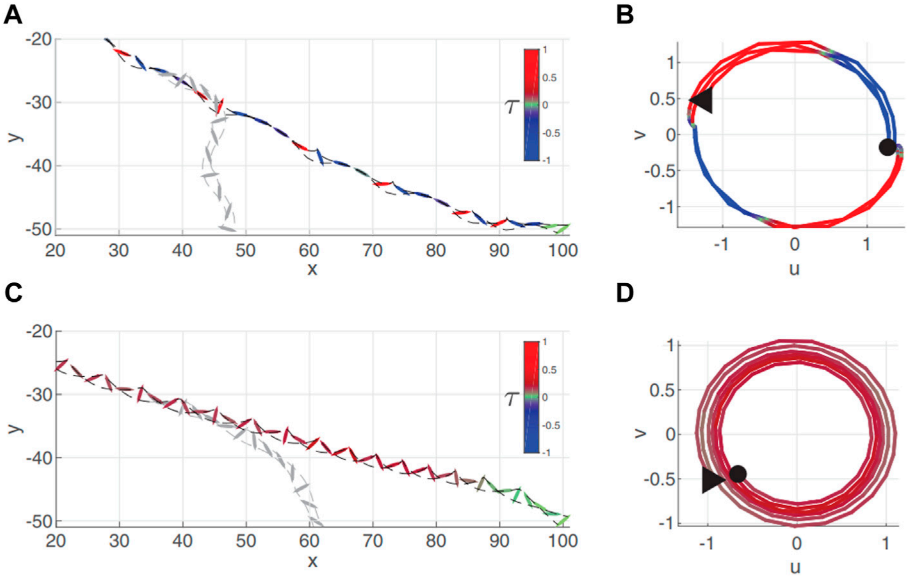

FIGURE 9 | Visualization of the two prevailing locomotion patterns adopted by RL agents for the active gliding model. Trajectories on the x-y plane for **(A)** bounding and **(C)** tumbling flight. The glider's snapshots are colored to signal the value of the control torque, and the dashed black lines track the ellipse's vertices. The grayed-out trajectories illustrate the glider's passive descent when abruptly switching off active control. **(B, D)** Corresponding trajectories on the u-v plane. The trajectories are colored based on the control torque, and a triangle and circle mark their beginning and end, respectively [84].

which obtained the qualitative characteristics of the control policy [77].

Behavior Patterns

Nature's creatures are the best teachers for researchers to discover the rule of behavior patterns, like gliders from birds that soar with thermal winds [93] or plant seeds that spread by gliding [94]. It is usually challenging to identify the internal mechanism of this adaptive pattern and generate corresponding behavior flow control strategies in another complex condition. Deep reinforcement learning has provided a new aspect to approach the goal.

In the identification and reproduction of fish adaption behaviors in complex environments, Zhu et al. [80] utilized deep recurrent Q-network (DRQN) algorithm with immersed boundary–lattice Boltzmann method to train the fish model and adapt its motion to optimally achieve a specific task, such as prey capture, rheotaxis and Kármán gaiting. Compared to existing learning models for fish, this work incorporated the fish position, velocity, and acceleration into the state space in the DRQN; it considered the amplitude and frequency action spaces and the historical effects. On the other hand, Mandralis et al. [81] deployed reinforcement learning to discover swimmer escape patterns constrained by the energy and prescribed functional form of the body motion, which can be transferred to the control of aquatic robotic devices operating under energy constraints. In addition, Yu et al. [82] numerically studied the collective locomotions of multiple undulatory self-propelled foils swimming by Q-learning algorithm. Especially swimming efficiency is the reward function, and visual information is included. It is found that the DRL algorithm can effectively discover various collective patterns with different characteristics, i.e., the staggered-following, tandem-following phalanx, and compact modes under two DRL strategies. The strategies are as follows: one is that only the following fish gets hydrodynamic advantages, and the other is that all group members take advantage of the interaction.

As for the gliding, there is also some work related to reinforcement learning, aiming at performing minimal mechanical work to control attitude. Reddy et al. [83] used Q learning to train a glider in the field to navigate atmospheric thermals autonomously, equipped with a flight controller that precisely controlled the bank angle and pitch, modulating these at intervals to gain as much lift as possible. The learned flight policy was validated through field experiments, numerical simulations, and estimates of the noise in measurements caused by atmospheric turbulence. Different from improving lift, Novati et al. [84] combined a two-dimensional model of a controlled elliptical body with DRL to achieve gliding with either minimum energy expenditure, or the fastest time of arrival, at a predetermined location. As seen in **Figure 9**, the model-free reinforcement learning led to more robust gliding than model-based optimal control policies with a modest additional computational cost. This study also demonstrated that the gliders with DRL can generalize their strategies to reach the objective location from previously unseen starting positions.

CHALLENGES ON DRL-BASED ACTIVE FLOW CONTROL

Modern control theory provides an essential basis for developing flow control methods from open-loop control to closed-loop

TABLE 2 | Challenges on DRL-based active flow control.

Category	Time	References	Algorithm	Key Words
Training Acceleration	2019	[95]	PPO	Parallelization of data collection
Training Acceleration	2021	[96]	PPO, TD3	Expert demonstrations, behavior cloning
Training Acceleration	2022	[97]	DQN	Transfer learning, *Pe* numbers
Training Acceleration	2022	[98]	PPO	Transfer learning, Re numbers
Training Acceleration	2022	[99]	SAC	Expert demonstrations, off-policy buffer
Control Delays	2022	[100]	ARP–DMDP–PPO	MDP, physics-informed delay, regressive
Sensor Configuration	2021	[71]	PPO	Sensitivity analysis
Sensor Configuration	2022	[101]	PPO, DPG	Global linear stability, sensitivity analyses
Sensor Configuration	2021	[102]	S-PPO-CMA	Sparse training, stochastic gate model
Sensor Configuration	2022	[38]	PPO	Linear genetic programming control
Sensor Configuration	2022	[79]	ApeX-DQN	Attention Branch Network
Partial Observables	2022	[103]	AC	Dissipative system, low-dimensional nature
Action Dimensionality	2019	[104]	PPO	Locality and invariance, densify reward

control. However, there may be better uses of time and resources than the detailed identification of a high-dimensional nonlinear fluid dynamical system for control. Alternatively, reinforcement learning with deep learning enables automatic feature engineering and end-to-end learning through gradient descent, so reliance on the flow mechanism is significantly reduced, shown in Section *Applications of DRL-based Active Flow Control*.

Though highlighted as a novel and promising direction, there are still some obstacles in the initial stage of DRL-based flow control. Some of these obstacles originate from the demand for practical reinforcement learning algorithms since direct numerical simulation, or experimental data are expensive to obtain in flow control problems. And others might be constrained by the flow control system's characteristics, such as control delay, sensor configuration, partial observation, etc. These obstacles have come to light during the application, and researchers have specified corresponding solutions with the knowledge of the physical system. More importantly, they have revealed potential problems and provided valuable references for similar issues, which are summarized in **Table 2**. The following section will focus on four aspects of challenges in using DRL-based active flow control: Section *Training Acceleration*, Section *Control Delays*, *Sensor Configuration*, Section *Partial Observables*, and Section *Action Dimensionality*.

Training Acceleration

Essentially, deep reinforcement learning is an optimization process based on parameterized policy (usually called "agent") through trial and error, which involves many interactions between the agent and the emulator. Therefore, compared to supervised/unsupervised learning, deep reinforcement learning is more time-consuming. Especially for the active flow control problem, the expensive data acquisition cost is required either in numerical simulation or wind tunnel experiment to represent the high dimensional flow state. On the other hand, the weakly-inductive-bias characteristic of reinforcement learning brings more possibilities and time consumption. To handle these issues, some works have been carried out on accelerating simulations or extracting prior knowledge from expert information for reinforcement learning, such as expert demonstrations, behavior cloning, or transfer learning.

From the perspective of accelerating simulation, Rabault et al. [95] demonstrated a perfect speedup by adapting the PPO algorithm for parallelization, which used several independent simulations running in parallel to collect experiences faster. As for extracting prior knowledge from expert information for reinforcement learning, Xie [96] firstly derived a simplified parametric control policy informed from direct DRL in sloshing suppression and then accelerated the DRL algorithm with a behavior cloning such simplified policy. Wang [98] transferred the DRL neural network trained with Re = 100, 200, 300 to the flow control tasks with Re = 200, 300, 1,000. As shown in **Figure 10**, it is due to the strong correlation between policy and the flow patterns under different conditions. Therefore a dramatic enhancement of learning efficiency can be achieved.

Furthermore, Konishi [97] introduced a physically reasonable transfer learning method for the trained mixer under different Péclet numbers. The balance transferability and fast learning on the Péclet number of the source domain were discussed. By filling the experience buffer with expert demonstrations, Zheng [99] proposed a novel off-policy reinforcement learning framework with a surrogate model optimization method, which enables data-efficient learning of active flow control strategies.

Control Delays

As the Reynolds number increases, temporal drag fluctuations under the DRL-controlled cylinder case tend to become increasingly more random and severe. Due to the appearance of turbulence in the state space, insufficient regression of the ANN with the time series during the decision process may result in deteriorating control robustness and temporal coherence. Due to the time elapse between actuation and response of flow, Mao [100] introduced the Markov decision process (MDP) with time delays to quantify the action delays in the DRL process by using a first-order autoregressive policy (ARP). This hybrid DRL method yielded a stable and coherent control, which resulted in a steadier and more elongated vortex formation zone behind the two-dimensional circular cylinder, hence, a much weaker vortex-shedding process and less fluctuating lift and drag forces. This

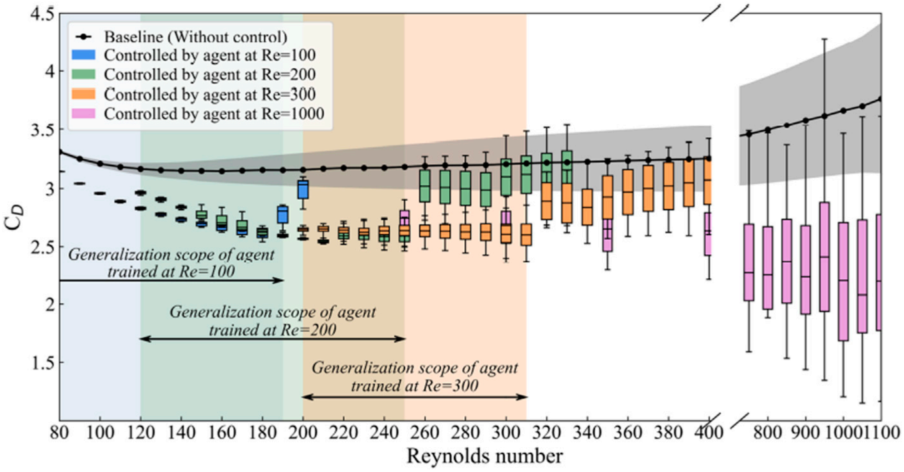

FIGURE 10 | Summary of the generalization reinforcement learning test. The black line and the mean line in boxes indicate the averaged drag coefficient in the flow without and with control, respectively, and the gray shaded area and box bodies show the range of oscillation of the drag coefficient at each corresponding Reynolds number [98].

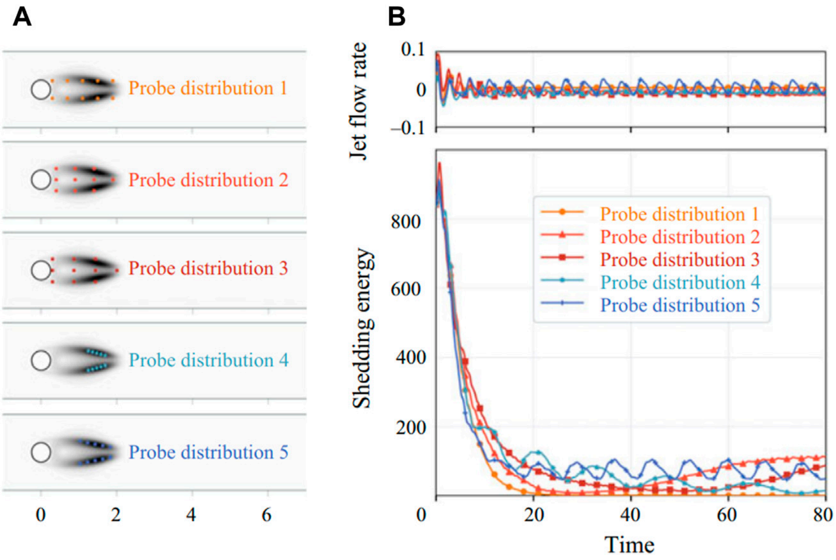

FIGURE 11 | RL control of the confined cylinder wake using ten probes. Different distributions of probes lead to a significant divergence in the control performance [101]. Panel **(A)** shows the five types of probe distribution, and panel **(B)** is the corresponding control performance, including the jet flow rate and the shedding energy.

method utilized the historical samples without additional training sampling than the standard DRL method. It can reduce drag and lift fluctuations by approximately 90% while achieving a similar level of drag reduction in the deterministic control at the same actuation frequency.

Sensor Configuration

In the closed-loop control framework, such as deep reinforcement learning, the sensor must be able to measure and provide a correct representation of the state of the flow system. The choice of the sensors, such as the type, number, and location, has a decisive effect on the maximum performance of the control policy. Extravagant sensor configuration is a huge and unnecessary burden in practical applications. The sensors measuring velocity, pressure, skin friction, and temperature in various resolutions, are mostly configured based on engineering experience. There is much room for improvement in adaptive algorithms, such as performing stability analysis or adopting novel optimization methods to obtain optimal and sensitive sensor locations.

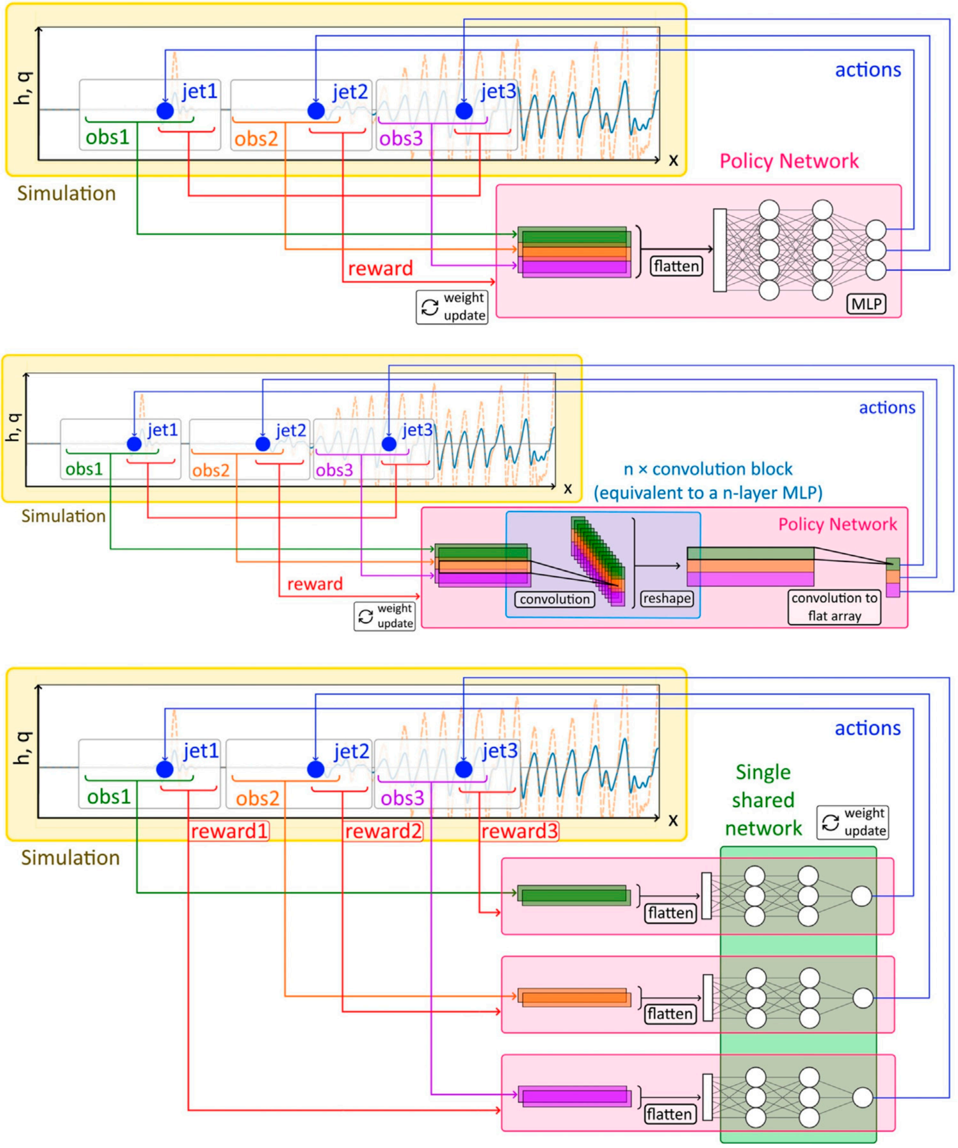

FIGURE 12 | Illustration of the 3 different methods for control of a system with translational invariance and locality. From top to bottom: M1, M2, and M3. M1 is the naive implementation of the DRL framework. M2 takes advantage of translation invariance of the system to reuse the network coefficients for the control of an arbitrary number of jets. M3 exploits both the translation invariance and the locality of the system by using a dense reward signal [104].

In terms of stability analysis, Ren et al. [71] performed a sensitivity analysis on the learned control policies to explore the layout of the sensor network by using the Python library SALib. It is concluded that the control had different sensitivity to locations and velocity components. Li et al. [101] conducted global linear stability and sensitivity analyses based on the adjoint method. It is found that the control is most efficient when the probes are placed in the most sensitive region, and it can be successful even when a few probes are properly placed in this manner. This work is a successful example of using and embedding physical flow information in the RL-based control. The Comparison between different probe distributions is shown in **Figure 11**.

As for the optimization methods, Paris et al. [102] introduced a novel algorithm (S-PPO-CMA) to optimize the sensor layout, focusing on the efficiency and robustness of the identified control policy. Along with a systematic study on sensor number and location, the proposed sparsity-seeking algorithm achieved a successful optimization with a reduced five-sensor layout while keeping state-of-the-art performance. Castellanos et al. [38] optimized the control policy by combining deep reinforcement learning and linear genetic programming control (LGPC) algorithm, which showed the capability of LGPC in identifying a subset of probes as the most relevant locations. In addition, Takada et al. [79] have adopted a new network structure named Attention Branch Network to visualize the activation area of the

DRL network, which provided references for sensor distribution. Especially, Attention Branch Network (ABN) [105] is a method to clarify the basis of the decision of neural networks, which enables the generation of an attention map to visualize the areas the neural network focuses on. It is clarified that the leading-edge pressure sensor is more important for determining the control action, and the trained neural network focused on the time variation of the pressure coefficient measured at the leading edge.

Partial Observables

Most current DRL algorithms assume that the environment evolves as a Markov decision process (MDP), and a learning agent can observe the environment state fully. However, in the real world, there are many cases where only partial observation of the state is possible. That is why existing reinforcement learning (RL) algorithms for fluid control may be inefficient under a small number of observables, even if the flow is laminar. By incorporating the dissipative system's low-dimensional space [106] of the learning algorithm, Kubo [103] resolved this problem and presented a framework for RL that can stably optimize the policy with a partially observable condition. In the practical application of a learning process in a fluid system like a learning agent without any information about flow state except rigid-body motion, the algorithm in this study can efficiently find the optimum control method.

Action Dimensionality

Sometimes it is difficult to handle high action space dimensionality on complex tasks. Applying reinforcement learning to those tasks requires tackling the combinatorial increase of the number of possible elements with the number of space dimensions. For example, for an environment with an N-dimensional action space and n discrete sub-actions for each dimension d, using the existing discrete-action algorithms, a total of $\prod_{d=1}^{N} n_d$ possible actions need to be considered. The number of actions that need to be explicitly represented grows exponentially with increasing action dimensionality [107].

Belus et al. [104] proposed a DRL framework to handle an arbitrary number of control actions (jets). This method relies on satisfactorily exploiting invariance and locality properties of the 1D falling liquid film system, which can be extended to other physics systems with similar properties. Inspired by the Convolutional Neural Networks (CNNs), three different methods for the DRL agent are designed as shown in **Figure 12**. This work set small regions in the neighborhood of each jet, where states and rewards were obtained. Methods 3 (”M3”) took into account this locality and extract N reward signals (the number of jets) to evaluate local behaviors with less dimension. Results showed both a good learning acceleration and easy control on an arbitrarily large number of jets and overcomed the curse of dimensionality on the control output size that would take place using a naive approach.

CONCLUSIONS

Exploring flow mechanisms and controlling flow has always been one of the most important and fruitful topics for researchers. The fluid system's high dimensionality, nonlinearity, and stochasticity limit the flow control policy exploration. It has yet to be widely applied in aviation or the marine industry. As a critical branch of artificial intelligence, reinforcement learning with deep learning enables automatic feature engineering and end-to-end learning through gradient descent so that reliance on domain knowledge is significantly reduced or even removed. Moreover, the deep and distributed representations in deep understanding can exploit the hierarchical composition of factors in data to combat the exponential challenges of the curse of dimensionality [108], which is a severe issue for the complex flow system.

Considerable research reviewed in Sections *Applications of DRL-based Active Flow Control* and *Challenges on DRL-Based Active Flow Control* has proved that deep reinforcement learning can achieve state-of-art performance in active flow control. Besides, there are other important topics which are not presented in the current review, such as optimization design [109–112], model discovery [113, 114], equation solving [115], microbiota behavior [116–119], plasmas magnetic control [120], convective heat exchange [121], chaotic system [122]. While there are some obstacles inevitably, like the demand to accelerate the training process (Section *Training Acceleration*) or the constraints related to the control system's characteristics, such as Section *Control Delays*, Section *Sensor Configuration*, partial observation (Section *Partial Observables*), Section *Action Dimensionality*, etc. This review has introduced five topics with their solutions, and more challenges are invisible below sea level, just like icebergs. We advocate that the physical information of the flow should be embedded into the DRL-based active flow control framework. More advanced data-driven methods should be fully utilized to discover the inherent association under big data. Efficient frameworks embedded with physical knowledge under practical background can promote the wide industrial application of intelligent, active flow control technology to the greatest extent. Based on the above research and our experience, it is inferred that the study of active flow control based on deep reinforcement learning in the future can be focused on the following five aspects:

(1) *Accelerate training speed and improve sample efficiency.* Compared with Atari, Go, and other traditional research fields of intensive learning, the cost of data acquisition is usually higher compared to numerical simulation or wind tunnel tests. Moreover, the high-dimensional feature extraction and random system characteristics are significant challenges to the convergence of these algorithms. It is of great significance to make more rational use of data, including offline paradigm [123], model building [124], data augmentation [125], etc.
(2) *Embed physical information into the reinforcement learning framework.* The pure AI algorithm neglects the dynamics and believes in the data-driven concept, which is also doomed to

be inefficient. It is brighter to combine the physical information into the DRL framework and develop artificial intelligence technology based on the full use of classical fluid mechanics research methods.

(3) *Attain interpretability from artificial intelligence decision.* Learning to control agents from high-dimensional inputs relies upon engineering problem-specific state representations, reducing the agent's flexibility. Embedding feature extraction in deep neural networks exposes deficiencies such as a lack of explanation, limiting the application of intelligent methods. Explainable AI methods [126] are advocated to improve the interpretability of intelligent control. With the help of such practices, further exploration of more fundamental physical connotations and scientific cognition of fluid mechanics is expected.

(4) *Transfer to the real world and eliminate the sim2real gap.* In practical applications like aircraft flight, it is unsafe to train agents directly by trial and error. However, the reality gap between the simulation and the physical world often leads to failure, which is triggered by an inconsistency between physical parameters (i.e., Reynolds number) and, more fatally, incorrect physical modeling (i.e., observation noise, action delay). Reducing or even eliminating the sim2real gap [127] is a crucial step in applying reinforcement learning to industrial applications.

(5) *Build up an open-source DRL-AFC community.* The rapid development of deep reinforcement learning in the field of active flow control owes to the fact that many predecessors published the code while publishing articles. At present, we can find the work of Rabault [67, 95], Jichao Li [101], Qiulei Wang [128] and others on the Github, including containers for full reproducibility. Such sharing and openness can not only let fluid mechanics researchers understand the latest release and update of DRL tools, but also let machine learning researchers understand the development direction of algorithms applied to complex physical systems. This review calls on researchers to further share code and open source benchmarks, build a multidisciplinary open source community, further strengthen cooperation, and promote the application of reinforcement learning in the field of fluid mechanics.

To summarize, deep reinforcement learning has established the beacon for active flow control, and its talent potential in complex flow system remain to be explored. Especially in the aviation industry, it is expected that this control mode can reach unprecedented heights and realize the impossible missions in many science fiction films, for example, rudderless aircraft controlled by jets, long-endurance vehicles with weak or even no drag, etc. It is no doubt there is still a long way before DRL-based flow control realizes real-world application, but it has promised us a bright future.

AUTHOR CONTRIBUTIONS

FX: Conceptualization; Investigation; Methodology; Project administration; Resources; Writing—review and editing. CZ: Conceptualization; Methodology; Writing—original draft. TJ: Conceptualization; Funding acquisition; Investigation; Project administration; Supervision. XZ: Methodology; Conceptualization. RB: Methodology; Writing—review and editing. HZ: Investigation; Writing—review and editing. YZ: Conceptualization; Funding acquisition; Investigation; Project administration; Supervision.

CONFLICT OF INTEREST

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

REFERENCES

1. Bower W, Kibens V. An overview of active flow control applications at the boeing company. In: 2nd AIAA Flow Control Conference (2004). p. 2624. doi:10.2514/6.2004-2624
2. Sudin MN, Abdullah MA, Shamsuddin SA, Ramli FR, Tahir MM. Review of research on vehicles aerodynamic drag reduction methods. *Int J Mech Mechatronics Eng* (2014) 14(02):37–47.
3. Zhang Y, Ye Z, Li B, Xie L, Zou J, Zheng Y. Numerical analysis of turbulence characteristics in a flat-plate flow with riblets control. *Adv Aerodynamics* (2022) 4(1):29–8. doi:10.1186/s42774-022-00115-z
4. Whalen EA, Shmilovich A, Spoor M, Tran J, Paul V, Lin JC, et al. Flight test of an active flow control enhanced vertical tail. *AIAA J* (2018) 56(9):3393–8. doi:10.2514/1.j056959
5. Glezer A, Amitay M. Synthetic jets. *Annu Rev Fluid Mech* (2002) 34(1): 503–29. doi:10.1146/annurev.fluid.34.090501.094913
6. Xie L, Zheng Y, Zhang Y, Ye ZX, Zou JF. Effects of localized micro-blowing on a spatially developing flat turbulent boundary layer. *Flow, Turbulence and Combustion* (2021) 107(1):51–79. doi:10.1007/s10494-020-00221-2
7. Cattafesta LN, III, Sheplak M. Actuators for active flow control. *Annu Rev Fluid Mech* (2011) 43:247–72. doi:10.1146/annurev-fluid-122109-160634
8. George B, Hussain F. Nonlinear dynamics of forced transitional jets: Periodic and chaotic attractors. *J Fluid Mech* (1994) 263:93–132. doi:10.1017/s0022112094004040
9. Koch CR, Mungal MG, Reynolds WC, Powell JD. Helical modes in an acoustically excited round air jet. *Phys Fluids A: Fluid Dyn* (1989) 1(9):1443. doi:10.1063/1.4738832
10. Kim J, Thomas RB. A linear systems approach to flow control. *Annu Rev Fluid Mech* (2007) 39(1):383–417. doi:10.1146/annurev.fluid.39.050905.110153
11. Bagheri S, Henningson DS, Hoepffner J, Schmid PJ. Input-output analysis and control design applied to a linear model of spatially developing flows. *Appl Mech Rev* (2009) 62(2). doi:10.1115/1.3077635
12. Brunton SL, Noack BR. Closed-loop turbulence control: Progress and challenges. *Appl Mech Rev* (2015) 67(5). doi:10.1115/1.4031175
13. Xu K, Ren Y, Zha G. Separation control by co-flow wall jet. In: AIAA AVIATION 2021 FORUM (2021). p. 2946.
14. Sato M, Aono H, Yakeno A, Nonomura T, Fujii K, Okada K, et al. Multifactorial effects of operating conditions of dielectric-barrier-discharge plasma actuator on laminar-separated-flow control. *AIAA J* (2015) 53(9): 2544–59. doi:10.2514/1.j053700
15. Farazmand MM, Kevlahan NKR, Protas B. Controlling the dual cascade of two-dimensional turbulence. *J Fluid Mech* (2011) 668:202–22. doi:10.1017/s0022112010004635

16. Semeraro O, Pralits JO, Rowley CW, Henningson DS. Riccati-less approach for optimal control and estimation: An application to two-dimensional boundary layers. *J Fluid Mech* (2013) 731:394–417. doi:10.1017/jfm.2013.352
17. Carini M, Pralits JO, Luchini P. Feedback control of vortex shedding using a full-order optimal compensator. *J Fluids Structures* (2015) 53:15–25. doi:10.1016/j.jfluidstructs.2014.11.011
18. Brunton SL, Kutz JN. *Data-driven science and engineering: Machine learning, dynamical systems, and control*. Cambridge: Cambridge University Press (2022).
19. Zhang X, Ji T, Xie F, Zheng C, Zheng Y. Data-driven nonlinear reduced-order modeling of unsteady fluid–structure interactions. *Phys Fluids* (2022) 34(5):053608. doi:10.1063/5.0090394
20. Zhang X, Ji T, Xie F, Zheng H, Zheng Y. Unsteady flow prediction from sparse measurements by compressed sensing reduced order modeling. *Comput Methods Appl Mech Eng* (2022) 393:114800. doi:10.1016/j.cma.2022.114800
21. Vinyals O, Babuschkin I, Czarnecki WM, Mathieu M, Dudzik A, Chung J, et al. Grandmaster level in starcraft ii using multi-agent reinforcement learning. *Nature* (2019) 575(7782):350–4. doi:10.1038/s41586-019-1724-z
22. Arulkumaran K, Deisenroth MP, Brundage M, Bharath AA. Deep reinforcement learning: A brief survey. *IEEE Signal Process. Mag* (2017) 34(6):26–38. doi:10.1109/msp.2017.2743240
23. François-Lavet V, Henderson P, Islam R, Bellemare MG, Pineau J. An introduction to deep reinforcement learning. *Foundations Trends® Machine Learn* (2018) 11(3-4):219–354. doi:10.1561/2200000071
24. Zou J, Han Y, So SS, Livingstone DJ. Overview of artificial neural networks. In: *Artificial neural networks*. In: *Methods in molecular Biology*™. Totowa, NJ, USA: Humana Press (2008). p. 14–22. doi:10.1007/978-1-60327-101-1_2
25. Silver D, Hubert T, Schrittwieser J, Antonoglou I, Lai M, Guez A, et al. A general reinforcement learning algorithm that masters chess, shogi, and go through self-play. *Science* (2018) 362(6419):1140–4. doi:10.1126/science.aar6404
26. Mnih V, Kavukcuoglu K, Silver D, Graves A, Antonoglou I, Wierstra D, et al. *Playing atari with deep reinforcement learning* (2013). *arXiv preprint arXiv:1312.5602*.
27. Kober J, Bagnell JA, Peters J. Reinforcement learning in robotics: A survey. *Int J Robotics Res* (2013) 32(11):1238–74. doi:10.1177/0278364913495721
28. Rabault J, Ren F, Zhang W, Tang H, Xu H. Deep reinforcement learning in fluid mechanics: A promising method for both active flow control and shape optimization. *J Hydrodynamics* (2020) 32(2):234–46. doi:10.1007/s42241-020-0028-y
29. Ren F, Hu HB, Tang H. Active flow control using machine learning: A brief review. *J Hydrodynamics* (2020) 32(2):247–53. doi:10.1007/s42241-020-0026-0
30. Vinuesa R, Lehmkuhl O, Lozano-Durán A, Rabault J. Flow control in wings and discovery of novel approaches via deep reinforcement learning. *Fluids* (2022) 7(2):62. doi:10.3390/fluids7020062
31. Garnier P, Viquerat J, Rabault J, Larcher A, Alexander K, Hachem E. A review on deep reinforcement learning for fluid mechanics. *Comput Fluids* (2021) 225:104973. doi:10.1016/j.compfluid.2021.104973
32. Viquerat J, Meliga P, Hachem E. *A review on deep reinforcement learning for fluid mechanics: An update* (2021). *arXiv preprint arXiv:2107.12206*.
33. Maceda GYC, Li Y, Lusseyran F, Morzyński M, Noack BR. Stabilization of the fluidic pinball with gradient-enriched machine learning control. *J Fluid Mech* (2021) 917.
34. Antoine BB, Maceda GYC, Fan D, Li Y, Zhou Y, Noack BR, et al. Bayesian optimization for active flow control. *Acta Mechanica Sinica* (2021) 37(12):1786–98. doi:10.1007/s10409-021-01149-0
35. Ren K, Chen Y, Gao C, Zhang W. Adaptive control of transonic buffet flows over an airfoil. *Phys Fluids* (2020) 32(9):096106. doi:10.1063/5.0020496
36. Gao C, Zhang W, Kou J, Liu Y, Ye Z. Active control of transonic buffet flow. *J Fluid Mech* (2017) 824:312–51. doi:10.1017/jfm.2017.344
37. Zheng C, Ji T, Xie F, Zhang X, Zheng H, Zheng Y. From active learning to deep reinforcement learning: Intelligent active flow control in suppressing vortex-induced vibration. *Phys Fluids* (2021) 33(6):063607. doi:10.1063/5.0052524
38. Castellanos R, Maceda GYC, de la Fuente I, Noack BR, Ianiro A, Discetti S. Machine-learning flow control with few sensor feedback and measurement noise. *Phys Fluids* (2022) 34(4):047118. doi:10.1063/5.0087208
39. Pino F, Schena L, Rabault J, Mendez MA. *Comparative analysis of machine learning methods for active flow control* (2022). *arXiv preprint arXiv:2202.11664*.
40. Banzhaf W, Nordin P, Keller RE, Francone FD. *Genetic programming: An introduction: On the automatic evolution of computer programs and its applications*. Burlington, MA, USA: Morgan Kaufmann Publishers Inc. (1998).
41. Langdon WB, Poli R. *Foundations of genetic programming*. Berlin, Germany: Springer Science & Business Media (2013).
42. Moriarty DE, Mikkulainen R. Efficient reinforcement learning through symbiotic evolution. *Machine Learn* (1996) 22(1):11–32. doi:10.1007/bf00114722
43. Salimans T, Ho J, Chen X, Sidor S, Sutskever I. *Evolution strategies as a scalable alternative to reinforcement learning* (2017). *arXiv preprint arXiv:1703.03864*.
44. Schulman J, Filip W, Dhariwal P, Radford A, Klimov O. *Proximal policy optimization algorithms* (2017). *arXiv preprint arXiv:1707.06347*.
45. Haarnoja T, Zhou A, Abbeel P, Levine S. Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor. In: International conference on machine learning. Stockholm, Sweden: PMLR (2018). p. 1861–70.
46. Konda V, Tsitsiklis J. Actor-critic algorithms. *Adv Neural Inf Process Syst* (1999) 12.
47. Sutton RS, Barto AG. *Reinforcement learning: An introduction*. Cambridge, MA, USA: MIT Press (2018).
48. Kaelbling LP, Littman ML, Moore AW. Reinforcement learning: A survey. *J Artif intelligence Res* (1996) 4:237–85. doi:10.1613/jair.301
49. Watkins CJCH, Dayan P. Q-learning. *Machine Learn* (1992) 8(3):279–92. doi:10.1007/BF00992698
50. Rummery GA, Niranjan M. *On-line Q-learning using connectionist systems, volume 37*. Cambridge, UK: University of Cambridge, Department of Engineering Cambridge, UK (1994).
51. Williams RJ. Simple statistical gradient-following algorithms for connectionist reinforcement learning. *Machine Learn* (1992) 8(3):229–56. doi:10.1007/bf00992696
52. Gullapalli V. A stochastic reinforcement learning algorithm for learning real-valued functions. *Neural networks* (1990) 3(6):671–92. doi:10.1016/0893-6080(90)90056-q
53. Tsitsiklis J, Van Roy B. Analysis of temporal-diffference learning with function approximation. *Adv Neural Inf Process Syst* (1996) 9.
54. Melo FS, Meyn SP, Ribeiro MI. An analysis of reinforcement learning with function approximation. In: Proceedings of the 25th international conference on Machine learning; July 5 - 9, 2008; Helsinki, Finland (2008). p. 664–71. doi:10.1145/1390156.1390240
55. Mnih V, Kavukcuoglu K, Silver D, Rusu AA, Veness J, Bellemare MG, et al. Human-level control through deep reinforcement learning. *Nature* (2015) 518(7540):529–33. doi:10.1038/nature14236
56. Sutton RS, McAllester D, Singh S, Mansour Y. Policy gradient methods for reinforcement learning with function approximation. *Adv Neural Inf Process Syst* (1999) 12.
57. Riedmiller M, Peters J, Schaal S. Evaluation of policy gradient methods and variants on the cart-pole benchmark. In: 2007 IEEE International Symposium on Approximate Dynamic Programming and Reinforcement Learning. Honolulu, HI, USA: IEEE (2007). p. 254–61. doi:10.1109/ADPRL.2007.368196
58. Grondman I, Busoniu L, Lopes GAD, Babuska R. A survey of actor-critic reinforcement learning: Standard and natural policy gradients. *IEEE Trans Syst Man, Cybernetics, C (Applications Reviews)* (2012) 42(6):1291–307. doi:10.1109/tsmcc.2012.2218595
59. Scott F, Hoof H, Meger D. Addressing function approximation error in actor-critic methods. In: International conference on machine learning. Stockholm, Sweden: PMLR (2018). p. 1587–96.
60. Lillicrap TP, Hunt JJ, Alexander P, Heess N, Tom E, Tassa Y, et al. *Continuous control with deep reinforcement learning* (2015). *arXiv preprint arXiv:1509.02971*.

61. Schulman J, Levine S, Abbeel P, Jordan M, Moritz P. Trust region policy optimization. In: International conference on machine learning. Stockholm, Sweden: PMLR (2015). p. 1889–97.
62. Haarnoja T, Tang H, Abbeel P, Levine S. Reinforcement learning with deep energy-based policies. In: International conference on machine learning. Stockholm, Sweden: PMLR (2017). p. 1352–61.
63. Koizumi H, Tsutsumi S, Shima E. Feedback control of karman vortex shedding from a cylinder using deep reinforcement learning. In: 2018 Flow Control Conference (2018). p. 3691.
64. Ren F, Wang C, Tang H. Bluff body uses deep-reinforcement-learning trained active flow control to achieve hydrodynamic stealth. *Phys Fluids* (2021) 33(9): 093602. doi:10.1063/5.0060690
65. Mei YF, Zheng C, Aubry N, Li MG, Wu WT, Liu X. Active control for enhancing vortex induced vibration of a circular cylinder based on deep reinforcement learning. *Phys Fluids* (2021) 33(10):103604. doi:10.1063/5.0063988
66. Pivot C, Cordier L, Mathelin L. A continuous reinforcement learning strategy for closed-loop control in fluid dynamics. In: 35th AIAA Applied Aerodynamics Conference; 5-9 June 2017; Denver, Colorado (2017). p. 3566. doi:10.2514/6.2017-3566
67. Rabault J, Kuchta M, Jensen A, Réglade U, Cerardi N. Artificial neural networks trained through deep reinforcement learning discover control strategies for active flow control. *J Fluid Mech* (2019) 865:281–302. doi:10.1017/jfm.2019.62
68. Qin S, Wang S, Rabault J, Sun G. *An application of data driven reward of deep reinforcement learning by dynamic mode decomposition in active flow control* (2021). *arXiv preprint arXiv:2106.06176*.
69. Xu H, Zhang W, Deng J, Rabault J. Active flow control with rotating cylinders by an artificial neural network trained by deep reinforcement learning. *J Hydrodynamics* (2020) 32(2):254–8. doi:10.1007/s42241-020-0027-z
70. Tang H, Rabault J, Alexander K, Wang Y, Wang T. Robust active flow control over a range of Reynolds numbers using an artificial neural network trained through deep reinforcement learning. *Phys Fluids* (2020) 32(5):053605. doi:10.1063/5.0006492
71. Ren F, Rabault J, Tang H. Applying deep reinforcement learning to active flow control in weakly turbulent conditions. *Phys Fluids* (2021) 33(3):037121. doi:10.1063/5.0037371
72. Varela P, Suárez P, Alcántara-Ávila F, Miró A, Rabault J, Font B, et al. Deep reinforcement learning for flow control exploits different physics for increasing Reynolds number regimes. *Actuators* (2022) 11:359. doi:10.3390/act11120359
73. Fan D, Liu Y, Wang Z, Triantafyllou MS, Karniadakis GEM. Reinforcement learning for bluff body active flow control in experiments and simulations. *Proc Natl Acad Sci* (2020) 117(42):26091–8. doi:10.1073/pnas.2004939117
74. Amico E, Cafiero G, Iuso G. Deep reinforcement learning for active control of a three-dimensional bluff body wake. *Phys Fluids* (2022) 34:105126. doi:10.1063/5.0108387
75. Wang YZ, Mei YF, Aubry N, Chen Z, Wu P, Wu WT. Deep reinforcement learning based synthetic jet control on disturbed flow over airfoil. *Phys Fluids* (2022) 34(3):033606. doi:10.1063/5.0080922
76. Guerra-Langan A, Araujo Estrada S, Windsor S. Reinforcement learning to control lift coefficient using distributed sensors on a wind tunnel model. In: AIAA SCITECH 2022 Forum; January 3-7, 2022; San Diego, CA & Virtual (2022). p. 0966. doi:10.2514/6.2022-0966
77. Shimomura S, Sekimoto S, Oyama A, Fujii K, Nishida H. Closed-loop flow separation control using the deep q network over airfoil. *AIAA J* (2020) 58(10):4260–70. doi:10.2514/1.j059447
78. Shimomura S, Sekimoto S, Oyama A, Fujii K, Nishida H. Experimental study on application of distributed deep reinforcement learning to closed-loop flow separation control over an airfoil. In: AIAA Scitech 2020 Forum; 6-10 January 2020; Orlando, FL (2020). p. 0579. doi:10.2514/6.2020-0579
79. Takada N, Ishikawa T, Furukawa T, Nishida H. Feedback control of flow separation over airfoil with deep reinforcement learning in numerical simulation. In: AIAA SCITECH 2022 Forum; January 3-7, 2022; San Diego, CA & Virtual (2022). p. 1365. doi:10.2514/6.2022-1365
80. Zhu Y, Tian FB, Young J, Liao JC, Lai J. A numerical study of fish adaption behaviors in complex environments with a deep reinforcement learning and immersed boundary–lattice Boltzmann method. *Scientific Rep* (2021) 11(1): 1691–20. doi:10.1038/s41598-021-81124-8
81. Mandralis I, Weber P, Guido N, Koumoutsakos P. Learning swimming escape patterns for larval fish under energy constraints. *Phys Rev Fluids* (2021) 6(9):093101. doi:10.1103/physrevfluids.6.093101
82. Yu H, Liu B, Wang C, Liu X, Lu XY, Huang H. Deep-reinforcement-learning-based self-organization of freely undulatory swimmers. *Phys Rev E* (2022) 105(4):045105. doi:10.1103/physreve.105.045105
83. Reddy G, Wong-Ng J, Celani A, Sejnowski TJ, Vergassola M. Glider soaring via reinforcement learning in the field. *Nature* (2018) 562(7726):236–9. doi:10.1038/s41586-018-0533-0
84. Guido N, Mahadevan L, Koumoutsakos P. Controlled gliding and perching through deep-reinforcement-learning. *Phys Rev Fluids* (2019) 4(9):093902. doi:10.1103/physrevfluids.4.093902
85. Drazin PG. *Introduction to hydrodynamic stability, volume 32*. Cambridge: Cambridge University Press (2002).
86. Schmid PJ, Henningson DS, Jankowski DF. Stability and transition in shear flows. applied mathematical sciences, vol. 142. *Appl Mech Rev* (2002) 55(3): B57–B59. doi:10.1115/1.1470687
87. Chandrasekhar S. *Hydrodynamic and hydromagnetic stability*. Chelmsford, MA, USA: Courier Corporation (2013).
88. Jan D, Le Gal P, Fraunié P. A numerical and theoretical study of the first hopf bifurcation in a cylinder wake. *J Fluid Mech* (1994) 264:59–80. doi:10.1017/s0022112094000583
89. Yue Y, Xie F, Yan H, Constantinides Y, Owen O, Karniadakis GEM. Suppression of vortex-induced vibrations by fairings: A numerical study. *J Fluids Structures* (2015) 54:679–700. doi:10.1016/j.jfluidstructs.2015.01.007
90. Xie F, Yue Y, Constantinides Y, Triantafyllou MS, Karniadakis GEM. U-shaped fairings suppress vortex-induced vibrations for cylinders in cross-flow. *J Fluid Mech* (2015) 782:300–32. doi:10.1017/jfm.2015.529
91. Abbas A, De Vicente J, Valero E. Aerodynamic technologies to improve aircraft performance. *Aerospace Sci Technol* (2013) 28(1):100–32. doi:10.1016/j.ast.2012.10.008
92. Munson BR, Okiishi TH, Huebsch WW, Rothmayer AP. *Fluid mechanics*. Singapore: Wiley (2013).
93. Newton I. *The migration ecology of birds*. Amsterdam, Netherlands: Elsevier (2010).
94. Shamoun-Baranes J, Leshem Y, Yom-Tov Y, Liechti O. Differential use of thermal convection by soaring birds over central Israel. *The Condor* (2003) 105(2):208–18. doi:10.1093/condor/105.2.208
95. Rabault J, Alexander K. Accelerating deep reinforcement learning strategies of flow control through a multi-environment approach. *Phys Fluids* (2019) 31(9):094105. doi:10.1063/1.5116415
96. Xie Y, Zhao X. Sloshing suppression with active controlled baffles through deep reinforcement learning–expert demonstrations–behavior cloning process. *Phys Fluids* (2021) 33(1):017115. doi:10.1063/5.0037334
97. Konishi M, Inubushi M, Goto S. Fluid mixing optimization with reinforcement learning. *Scientific Rep* (2022) 12(1):14268–8. doi:10.1038/s41598-022-18037-7
98. Wang YZ, Yue H, Aubry N, Chen ZH, Wu WT, Cui J. Accelerating and improving deep reinforcement learning-based active flow control: Transfer training of policy network. *Phys Fluids* (2022) 34(7):073609. doi:10.1063/5.0099699
99. Zheng C, Xie F, Ji T, Zhang X, Lu Y, Zhou H, et al. Data-efficient deep reinforcement learning with expert demonstration for active flow control. *Phys Fluids* (2022) 34:113603. doi:10.1063/5.0120285
100. Mao Y, Zhong S, Yin H. Active flow control using deep reinforcement learning with time delays in markov decision process and autoregressive policy. *Phys Fluids* (2022) 34(5):053602. doi:10.1063/5.0086871
101. Li J, Zhang M. Reinforcement-learning-based control of confined cylinder wakes with stability analyses. *J Fluid Mech* (2022) 932:A44. doi:10.1017/jfm.2021.1045
102. Paris R, Beneddine S, Dandois J. Robust flow control and optimal sensor placement using deep reinforcement learning. *J Fluid Mech* (2021) 913:A25. doi:10.1017/jfm.2020.1170
103. Kubo A, Shimizu M. Efficient reinforcement learning with partial observables for fluid flow control. *Phys Rev E* (2022) 105(6):065101. doi:10.1103/physreve.105.065101

104. Vincent B, Rabault J, Viquerat J, Che Z, Hachem E, Reglade U. Exploiting locality and translational invariance to design effective deep reinforcement learning control of the 1-dimensional unstable falling liquid film. *AIP Adv* (2019) 9(12):125014. doi:10.1063/1.5132378
105. Fukui H, Hirakawa T, Yamashita T, Fujiyoshi H. Attention branch network: Learning of attention mechanism for visual explanation. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. Long Beach, CA, USA: IEEE (2019). p. 10705–14. doi:10.1109/CVPR.2019.01096
106. Temam R. *Infinite-dimensional dynamical systems in mechanics and physics, volume 68*. Berlin, Germany: Springer Science & Business Media (2012).
107. Tavakoli A, Pardo F, Kormushev P. Action branching architectures for deep reinforcement learning. *Proc AAAI Conf Artif Intelligence* (2018) 32. doi:10.1609/aaai.v32i1.11798
108. Li Y. *Deep reinforcement learning: An overview* (2017). *arXiv preprint arXiv: 1701.07274*.
109. Hui X, Wang H, Li W, Bai J, Qin F, He G. Multi-object aerodynamic design optimization using deep reinforcement learning. *AIP Adv* (2021) 11(8): 085311. doi:10.1063/5.0058088
110. Lai P, Wang R, Zhang W, Xu H. Parameter optimization of open-loop control of a circular cylinder by simplified reinforcement learning. *Phys Fluids* (2021) 33(10):107110. doi:10.1063/5.0068454
111. Hassan G, Viquerat J, Larcher A, Meliga P, Hachem E. Single-step deep reinforcement learning for open-loop control of laminar and turbulent flows. *Phys Rev Fluids* (2021) 6(5):053902. doi:10.1103/physrevfluids.6.053902
112. Viquerat J, Rabault J, Alexander K, Hassan G, Larcher A, Hachem E. Direct shape optimization through deep reinforcement learning. *J Comput Phys* (2021) 428:110080. doi:10.1016/j.jcp.2020.110080
113. Bae HJ, Koumoutsakos P. Scientific multi-agent reinforcement learning for wall-models of turbulent flows. *Nat Commun* (2022) 13(1):1443–9. doi:10.1038/s41467-022-28957-7
114. Kim J, Kim H, Kim J, Lee C. *Deep reinforcement learning for large-eddy simulation modeling in wall-bounded turbulence* (2022). *arXiv preprint arXiv: 2201.09505*.
115. Wei S, Jin X, Li H. General solutions for nonlinear differential equations: A rule-based self-learning approach using deep reinforcement learning. *Comput Mech* (2019) 64(5):1361–74. doi:10.1007/s00466-019-01715-1
116. Qiu J, Mousavi N, Zhao L, Gustavsson K. Active gyrotactic stability of microswimmers using hydromechanical signals. *Phys Rev Fluids* (2022) 7(1):014311. doi:10.1103/physrevfluids.7.014311
117. Zhu G, Wen F, Zhu L. *Optimising low-Reynolds-number predation via optimal control and reinforcement learning* (2022). *arXiv preprint arXiv: 2203.07196*.
118. Borra F, Biferale L, Cencini M, Celani A. Reinforcement learning for pursuit and evasion of microswimmers at low Reynolds number. *Phys Rev Fluids* (2022) 7(2):023103. doi:10.1103/physrevfluids.7.023103
119. Tsang ACH, Tong PW, Nallan S, Pak OS. Self-learning how to swim at low Reynolds number. *Phys Rev Fluids* (2020) 5(7):074101. doi:10.1103/physrevfluids.5.074101
120. Jonas D, Felici F, Jonas B, Neunert M, Tracey B, Carpanese F, et al. Magnetic control of tokamak plasmas through deep reinforcement learning. *Nature* (2022) 602(7897):414–9. doi:10.1038/s41586-021-04301-9
121. Beintema G, Corbetta A, Biferale L, Toschi F. Controlling Rayleigh–bénard convection via reinforcement learning. *J Turbulence* (2020) 21(9-10): 585–605. doi:10.1080/14685248.2020.1797059
122. Bucci MA, Semeraro O, Alexandre A, Wisniewski G, Cordier L, Mathelin L. Control of chaotic systems by deep reinforcement learning. *Proc R Soc A* (2019) 475(2231):20190351. doi:10.1098/rspa.2019.0351
123. Levine S, Kumar A, Tucker G, Fu J. *Offline reinforcement learning: Tutorial, review, and perspectives on open problems* (2020). *arXiv preprint arXiv: 2005.01643*.
124. Moerland TM, Broekens J, Jonker CM. *Model-based reinforcement learning: A survey* (2020). *arXiv preprint arXiv:2006.16712*.
125. Connor S, Khoshgoftaar TM. A survey on image data augmentation for deep learning. *J big Data* (2019) 6(1):60–48. doi:10.1186/s40537-019-0197-0
126. Xu F, Uszkoreit H, Du Y, Fan W, Zhao D, Zhu J. Explainable ai: A brief survey on history, research areas, approaches and challenges. In: CCF international conference on natural language processing and Chinese computing; October 9–14, 2019; Dunhuang, China. Berlin, Germany: Springer (2019). p. 563–74.
127. Höfer S, Bekris K, Handa A, Gamboa JC, Mozifian M, Golemo F, et al. Sim2real in robotics and automation: Applications and challenges. *IEEE Trans automation Sci Eng* (2021) 18(2):398–400. doi:10.1109/tase.2021.3064065
128. Wang Q, Yan L, Hu G, Li C, Xiao Y, Xiong H, et al. *Drlinfluids–an open-source python platform of coupling deep reinforcement learning and openfoam* (2022). *arXiv preprint arXiv:2205.12699*.

AEROSPACE RESEARCH COMMUNICATIONS
ORIGINAL RESEARCH
published: 30 March 2023
doi: 10.3389/arc.2023.11194

Deep-Learning-Based Uncertainty Analysis of Flat Plate Film Cooling With Application to Gas Turbine

Yaning Wang[1,2], Xubin Qiu[2], Shuyang Qian[2], Yangqing Sun[2], Wen Wang[1,2] and Jiahuan Cui[1,2]*

[1]School of Aeronautics and Astronautics, Zhejiang University, Hangzhou, China, [2]ZJU-UIUC Institute, Zhejiang University, Haining, China

Nowadays, gas turbines intake jet air at high temperatures to improve the power output as much as possible. However, the excessive temperature typically puts the blade in the face of unpredictable damage. Film cooling is one of the prevailing methods applied in engineering scenarios, with the advantages of a simple structure and high cooling efficiency. This study aims to assess the uncertain effect that the three major film cooling parameters exert on the global and fixed-cord-averaged film cooling effectiveness under low, medium, and high blowing ratios *br*. The three input parameters include coolant hole diameter *d*, coolant tube inclination angle θ, and density ratio *dr*. The training dataset is obtained by Computational Fluid Dynamics (CFD). Moreover, a seven-layer artificial neural network (ANN) algorithm is applied to explore the complex non-linear mapping between the input flat film cooling parameters and the output fixed-cord-averaged film cooling effectiveness on the external turbine blade surface. The sensitivity experiment conducted using Monte Carlo (MC) simulation shows that the *d* and θ are the two most sensitive parameters in the low-blowing-ratio cases. The θ comes to be the only leading factor of sensitivity in larger blowing ratio cases. As the blowing ratio rises, the uncertainty of the three parameters *d, θ,* and *dr* all decrease. The combined effect of the three parameters is also dissected and shows that it has a more significant influence on the general cooling effectiveness than any single effect. The *d* has the widest variation of uncertainty interval at three blowing ratios, while the θ has the largest uncertain influence on the general cooling effectiveness. With the aforementioned results, the cooling effectiveness of the gas turbine can be furthermore enhanced.

Keywords: film cooling, gas turbine, deep learning, Sobol method, uncertainty quantification

***Correspondence:**
Jiahuan Cui
jiahuancui@intl.zju.edu.cn

Received: *13 January 2023*
Accepted: *23 February 2023*
Published: *30 March 2023*

Citation:
Wang Y, Qiu X, Qian S, Sun Y, Wang W and Cui J (2023) Deep-Learning-Based Uncertainty Analysis of Flat Plate Film Cooling With Application to Gas Turbine.
Aerosp. Res. Commun. 1:11194.
doi: 10.3389/arc.2023.11194

INTRODUCTION

In gas turbine applications, since the intake gas temperature positively correlates with the power output, the temperature of the intake gas is expected to be as high as possible in pursuit of a better power output of a gas turbine. However, such temperature typically exceeds the melting temperature of turbine blades, which would cause the blade to melt and even lead to potential dangers in a gas turbine (1). Therefore, it is crucial to develop effective cooling methods to prevent potential overheating problems and avoid operating in overheated environments. Among a diverse selection of cooling methods, film cooling is the preferred and widely accepted choice in practical applications. A jet of coolant is extracted from the compressor and sprayed out from the coolant hole drilled on the flat blade in a designed geometric orientation. This jet of cooling air soon conflates the mainstream and then quickly covers the top surface of the blade, serving as an interlayer between the superheated mainstream air and the blades to prevent the blades from making direct contact with the hot mainstream. This process helps to prolong the service life of the blade.

Nonetheless, the process of predicting the cooling effects from a given flat plate hole parameter set is complicated by the complexity and unpredictability of the vortex structure and gas mixing motion. Previous studies have shown that two categories of input parameters matter to the resulting film cooling efficiency. They include the property of the coolant jet and the coolant injection method. The thermal property of the coolant includes coolant temperature (2) or coolant-to-mainstream temperature ratio (3). Garg et al. (2) delved into the impact of coolant temperature exerted on adiabatic effectiveness of the gas turbine blades by applying Navier-Stokes codes. Han et al. (3) concluded that the mainstream-to-coolant jet temperature ratio impacts the film cooling effectiveness greatly, and better cooling performance is achieved at higher temperature. The dynamic coolant properties of coolant contain coolant density ratio (4) and blowing ratio (5). Sinha (6) found that increasing the density ratio would create a negative lifting effect that promotes the spreading of the coolant jet heavily, and thus, impaired the film cooling effectiveness significantly. Cao et al. (7) outlined that with the continuous increasing of blowing ratio, the film cooling effect first rises and then falls. The coolant injection method contains the cooling hole shape (8), compound angle (9), inclination angle (10), and exit lateral diffusion angles (11). Gritsch et al. (8) concluded that hole shapes significantly impact the film cooling effectiveness. Most studies have shown that coolant hole diameter, coolant tube inclination angle, density ratio, and blowing ratio play a good role in film cooling effectiveness. Thus, a performance analysis on these elements is warranted.

Uncertainty analysis is of great importance in gas turbines (12). Different parameters influence the behavior of the gas turbine differently. Even slight variations of some specific parameters could bring considerable differences in the performance of the gas turbine. However, most previous uncertainty quantification studies are conducted based on conventional and inefficient Polynomial Chaos Expansion (PCE) models. Akbar et al. (13) surveyed seven uncertainty parameters in total and performed the uncertainty quantification analysis by the PCE method for film cooling. Shi et al. (14) used a PCE method to evaluate the uncertain effect of the conical angles etc. on discharge coefficient and adiabatic cooling effectiveness. Mathiodakis et al. (15) implemented research on the effect of ambient humidity on gas turbine performance and found that under high working temperatures, the impact that humidity has on the gas turbine is much more severe compared with low working temperatures. Huang et al. (16) also applied uncertainty quantification in their study of the heat transfer performance on rotor blade squealer tips.

However, the computation time and computation load increase exponentially in cases of higher dimensions, and the traditional PCE methods are commonly used to solve the “single output” problem, i.e., it is more widely used to obtain the overall cooling temperature of the research region only, instead of the fix-cord-averaged temperature analysis. Even though theoretically, the PCE methods can also be utilized to produce laterally averaged results, the computational cost is relatively higher. Many beneficial attempts are conducted to solve the difficulties in high-dimensional cases, such as the surrogate-based optimization method and artificial neural network et al., to conclude the complex non-linear correlation between the input coolant parameter configurations and the resulting cooling effectiveness using semi-empirical correlations. Mellor et al. (17) accomplished that by finding and validating a semi-empirical correlation. However, the computation is still very complicated.

In recent years, deep learning has emerged and is making a favorable contribution in pushing the process of various application fields forward (18). In fluid mechanics, a surrogate model based on deep learning is a beneficial tool for setting up the complicated non-linear, and obscure relation between two data sets. Ma et al. (19) investigated the behavior of the combustion chamber in a rocket. They utilized a convolutional neural network to forecast relations between coolant jet film and mainstream hot jet. Dolati et al. (20) studied the film cooling effectiveness by building a GMDH-type neural network to model the plasma actuator effects over a flat plate. Yang et al. (21) employed convolution modeling to predict the plugging problems and cooling efficiency of transpiration film cooling in the study. Wang et al. (22) utilized a GRU neural network model to research a variety of cooling parameters in one dimension to forecast the trench film cooling effectiveness. It is deduced that it is accessible to use deep learning methods in film cooling research. Furthermore, Wang et al. also applied a supervised ANN on a SVG cooling configuration to explore the non-linear mapping between parameters and performance and conclude that when the blowing ratio is low, the radius of SVG dominates the cooling effectiveness (23).

This paper constructed and validated a deep-learning-based ANN model to obtain the dataset to identify a non-linear mapping to link the four cooling parameters to cooling effectiveness. The application of the ANN model greatly enhanced the reliability of the correlation between parameters and the performance of the flat film cooling. The four cooling parameters include coolant hole diameter d, density ratio dr, coolant tube inclination angle θ, and blowing ratio br, and then we conducted uncertain effects of these film-cooling parameters at different blowing ratios. *Test Case Definition and Turbulence Model Selection* section defines all the geometries and parameters related to the test case in detail. The training dataset is generated using CFD simulations. Then, in *Deep Learning Modeling and Validation* section, a seven-layer ANN algorithm is built to obtain a non-linear mapping. Finally, uncertainty quantification is conducted in *Uncertainty Analysis* section to compare the

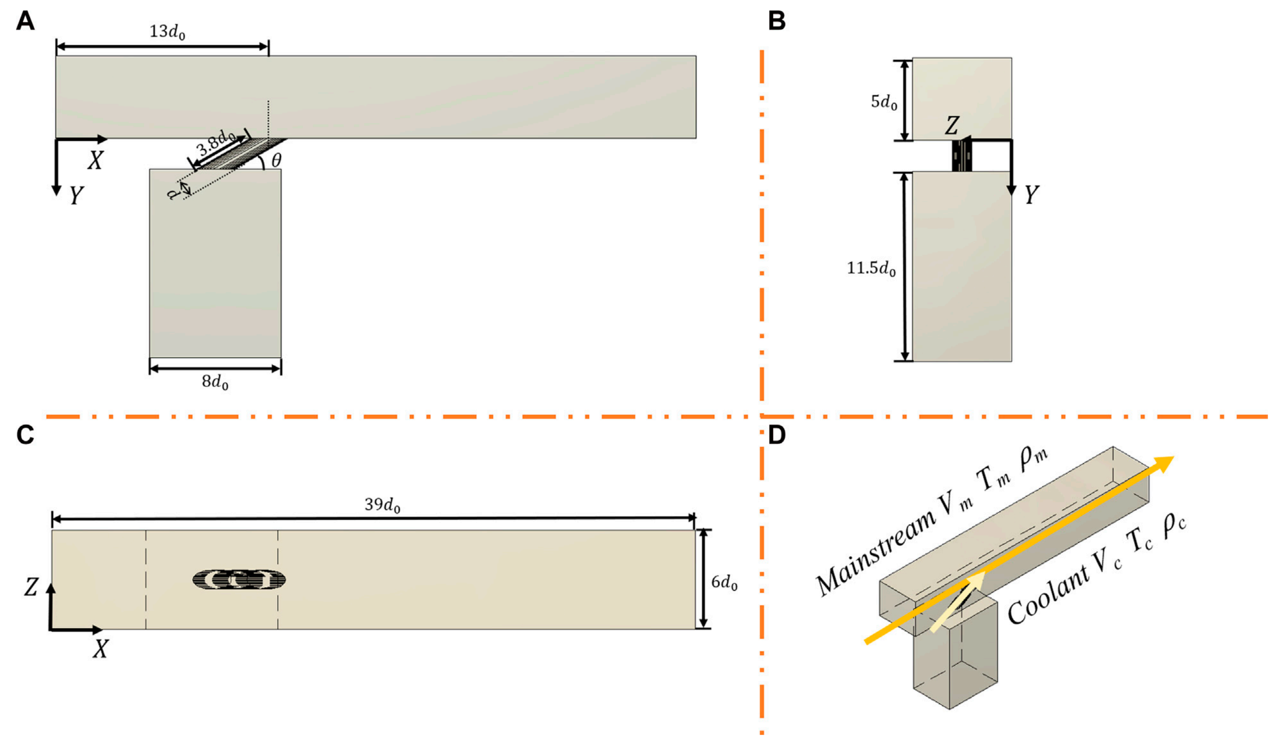

FIGURE 1 | Schematic of the flat plate film-cooling configuration: **(A)** Front view; **(B)** Left view; **(C)** Top view; **(D)** Stereogram.

effect of uncertain deviation of the three major film cooling diameters, including single hole diameter, inclination angle, and density ratio at different blowing ratios. Conclusions are drawn in *Conclusion* section.

TEST CASE DEFINITION AND TURBULENCE MODEL SELECTION

Test Case Geometry Setups

In previous research, the density ratio, blowing ratio, inclination angle, and diameter of the coolant tube hole are proven to have the most significant impact on the general temperature distribution near the external surface of the blade (24, 25). So in this paper, those four parameters were chosen to be researched. Given that successive combinations of several repeating units form the actual configuration of the blades, **Figure 1** shows the three-view drawing of the minimal periodic reference model. The mainstream chamber and the coolant chamber are modeled as two cuboids of dimensions $39d_0 \times 6d_0 \times 5d_0$ and $8d_0 \times 6d_0 \times 11.5d_0$, respectively. Here, d_0 stands for the standard diameter of the circular cross-section of the coolant inlet tube, and $d_0 = 12.5$ mm. $13d_0$ is measured from the mainstream inlet's front side to the coolant tube's central point. The coolant inlet tube is inclined at an angle θ concerning the x-z plane. The length of the coolant tube is $3.8d_0$. The mainstream inflow has a density of ρ_m and velocity of V_m, and the coolant inflow has a density of ρ_c and velocity of V_c. Due to the deviation of the cross-section area, this paper guarantees that the velocity of the jet inflow right at the coolant exit would accelerate to V_c. The flow direction of the mainstream flow is defined as going right, and the coolant jet flows upwards. All values mentioned above remain stationary as referenced except for the diameter, inclination angle, density ratio, and blowing ratio. **Table 1** summarizes the deviation interval of these four parameter inputs. These settings are the same as other scholars' research (19, 21, 24, 25). When one parameter is adjusted, to keep mono-variate, other parameters remain unchanged. This is achieved by slightly adjusting V_c and T_c. For example, when dr is changed from 1.1 to 1.2, V_c and T_c are adjusted to ensure that br, d, *and* θ are all unchanged.

TABLE 1 | Values of the film cooling parameters used in CFD.

Cooling parameters	Values
Blowing Ratio, *br*	[0.5,1.0,1.5]
Film Cooling Diameter, *d*/(mm)	[10.5, 11.5, 12.5, 13.5, 14.5]
Coolant Inclination Angle, *θ*/(degree)	[15, 25, 35, 45, 55]
Density Ratio, *dr*	[1.1, 1.2, 1.3]
Mainstream Temperature, T_m/(K)	313
Mainstream Velocity, V_m/(m/s)	20

Computational Assumptions and Variables

Ansys Fluent software is proven to have excellent performance in solving cooling problems (28-30). The problem is solved by applying steady-state solvers, which function by performing several iterations until the result converges. The standard to determine convergence is by comparing the continuity residual with 10^{-4} (27). If using a machine with 56 CPU to obtain the 225 datasets, it takes approximately 675 (225 × 3) hours. The mainstream and coolant inlet are assumed as velocity boundary conditions, and the mainstream outlet pressure is set to be 1 atm. Adiabatic and no-slip wall boundaries are applied for the mainstream and coolant chamber walls. Both flows are treated as superheated ideal gas, which follows the ideal gas law and is incompressible for low speed. The buoyancy effect is not assumed following Wang et al. (22, 27) and Yang et al. (26, 29). The Prandtl number of the turbulent is set to 0.667.

The *dr* and *br* are defined as shown in Eqs **1**, **2**. *dr* is defined as the temperature ratio at the mainstream versus the coolant jet. ρ_m and ρ_c mean the mainstream jet gas density and coolant jet gas density, respectively. The subscripts "m" and "c" denotes mainstream and coolant, which remain the same for temperature T and velocity V which are defined later. Furthermore, the definition of *br* can be interpreted as the product of *dr* and the velocity ratio at the coolant jet and mainstream.

$$dr = \frac{\rho_c}{\rho_m} = \frac{T_m}{T_c} \tag{1}$$

$$br = \frac{\rho_c \times V_c}{\rho_m \times V_m} \tag{2}$$

To avoid using a complicated high dimensional temperature matrix to represent the cooling efficiency, T^* and $\overline{T^*}$ are defined in Eqs **3**, **4**. T is the gauged temperature of the adiabatic and no-slip wall, with * denoting "dimensionless", and short bar overhead "–" labeling "fixed-cord-averaged." To represent the wall temperature, the output is interpolated as a 64 × 256-dimensional matrix from $0 < z < 6d_0$, $13d_0 < x < 39d_0$.

$$T^*(x,z) = \frac{T(x,z) - T_c}{T_m - T_c} \tag{3}$$

$$\overline{T^*}(x,z) = \frac{1}{64}\sum_{1}^{64} T^*(x,z)dz \tag{4}$$

Besides, the film cooling effectiveness at a single point, together with the fixed-cord-averaged and general film cooling effectiveness are derived according to dimensionless temperature T^*, as shown below in Eqs **5–7**.

$$\eta(x,z) = \frac{T_m - T}{T_m - T_c} = 1 - T^*(x,z) \tag{5}$$

$$\bar{\eta}(x) = \frac{1}{64}\sum_{z=1}^{z=64}\eta(x,z) \tag{6}$$

$$\eta_{av} = \left(\frac{1}{64}\right) \times \left(\frac{1}{256}\right)\sum_{z=1}^{z=64}\sum_{x=1}^{x=256}\eta(x,z) \tag{7}$$

Equation **8** defines the Mean Absolute Error (MAE), which is the index chosen to quantify the performance of a method. In the equations, m denotes the total sample size, a_i is the training data gained by Reynolds-Averaged Navier-Stokes (RANS) model, and y_i is the data acquired *via* the Large Eddy Simulation (LES) carried out by Wang et al. (31).

$$MAE = \frac{1}{m}\sum_{i=1}^{m}\left|a_i - y_i\right| \tag{8}$$

Turbulence Model Validation

Following other studies, the Fluent® 18.0 software is applied to all cases (21, 31). A validation experiment is conducted on the downstream central line's temperature and jet velocity distributions. This validation experiment aims to find the most suitable turbulence model for the following ANN and uncertainty quantification (UQ) analysis. The reference conditions are $d = d_0 = 12.5$ mm, $\theta = 35°$, $dr = 1.2$, and $br = 1.0$, which follows Wang et al.'s settings (27). The three candidate numerical methods are the SA model, the RNG k-ε model, the SST k-ω model, the Realizable k-ε model, and the experimental data from Ito (32, 33) and Sinha et al. (6) are used for comparison. The plot is shown in **Figure 2**, where x denotes the x-coordinate of the coolant jet outlet. The ratio x/d_0 is set as an x-axis parameter in two plots to realize parallel comparison under different d_0 selections. The coordinate interval researched is concentrated from $13d_0 < x < 39d_0$, $0 < z < 6d_0$. The data located in the region where $x < 13d_0$ is truncated because the constant mainstream temperature is assumed. In **Figure 2A**, the central-line film cooling effectiveness of the outer surface of the blade is computed *via* four numerical methods and two sets of experimental data. **Figure 2B** compares the four numerical models' mean average values (MAE). The results indicate that the MAEs for the SA and Realizable k-ε models are larger than the SST k-ω and RNG k-ε models. Furthermore, the trend of the RNG k-ε model results matched better with the data obtained from the experiments. The MAEs for the SST k-ω and the RNG k-ε models are 0.0106 and 0.0117, respectively. In comparison, the MAE for the Realizable k-ε model and SA model are 0.0241 and 0.0228, respectively, which are almost twice as large as the RNG k-ε and the SST k-ω model. Therefore, the RNG k-ε model is utilized in this study.

Grid Independence Study

In this study, the unstructured hybrid mesh is utilized. The y+ value for the near-wall cell is 1. Moreover, the grid stretch ratio is measured as 1.2 away from the solid wall. The grid cell number must be determined carefully since a massive number of grid cells raises the computational time meaninglessly, while too limited cell number conveys limited temperature distribution information and causes inaccuracy (34, 35). Therefore, the 2, 4.5, 6, and 7.5 million grid sizes are studied on the centerline of the flat plate model. The result is shown in **Figure 3**.

When the location is right downstream of the coolant hole, the 2-million case obviously differs from 4.5, 6, and 7.5-million cases, whereas the difference narrows as the distance increases. The 7.5 million grid cell case has the most significant fluctuation among all, which implies that the 7.5-million case is the most sensitive to react. This paper sets the grid cell number at a 6-million grid size for analyzing training and validation CFD data. **Figure 4** is the mesh schematic from three views: Axonometric,

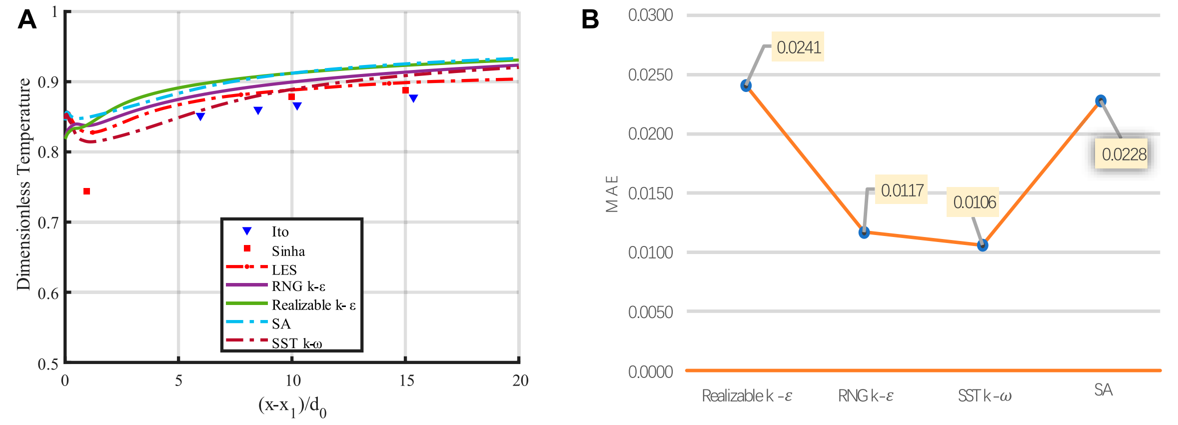

FIGURE 2 | Comparison between model validation: **(A)** Dimensionless temperature comparison between models on center line results; **(B)** MAE comparison between models.

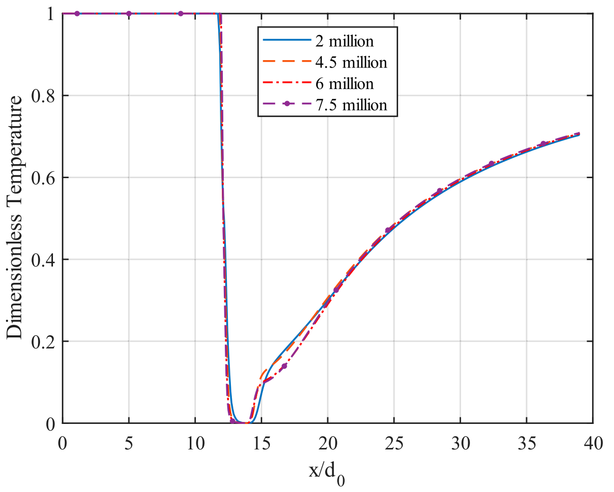

FIGURE 3 | Centerline dimensionless temperature with 2, 4.5, 6, and 7.5 million grid cells.

Front, and Top. It can be observed that a more accurate mesh resolution is located near the coolant hole.

DEEP LEARNING MODELING AND VALIDATION

Data Preparation

Given the enormous computational amount of the CFD method, this paper adopts ANN to reduce the computational burden. The input is a matrix containing coolant tube diameter, coolant tube inclination angle, density ratio, and blowing ratio. The output is a 64 × 256-dimensional matrix, with each entry representing the grid's temperature at the flat plate model's external surface. Cooling effectiveness can thus be obtained from this output matrix by applying Eqs **3–7**. The input data are normalized within (0, 1) before plugging into the input matrix. In Eq. **9**, each parameter with the subscript "norm" stands for the corresponding parameter after normalization. The learning domain is set in the rectangular region, with x and z coordinates fulfilling $13d_0<x < 39d_0$, $0<z < 6d_0$.

$$\begin{cases} \theta_{norm} = \dfrac{\theta - \theta_{min}}{\theta_{max} - \theta_{min}} = \dfrac{\theta - 15}{55 - 15} = \dfrac{\theta}{40} \\ br_{norm} = \dfrac{br - br_{min}}{br_{max} - br_{min}} = \dfrac{br - 0.5}{1.5 - 0.5} = br - 0.5 \\ d_{norm} = \dfrac{d - d_{min}}{d_{max} - d_{min}} = \dfrac{d - 10.5}{14.5 - 10.5} = \dfrac{d - 10.5}{4} \\ dr_{norm} = \dfrac{dr - dr_{min}}{dr_{max} - dr_{min}} = \dfrac{dr - 1.1}{1.3 - 1.1} = \dfrac{dr - 1.1}{1.2} \end{cases} \tag{9}$$

Structure of ANN

The ANN model is utilized to build the non-linear relationship between four flat plate configuration inputs, and the output temperature distribution matrix near the flat plate surface. The ANN model has seven layers in total, including one input layer, five hidden layers and one output layer. The first layer contains the four input parameters clarified above, and the information is propagated forward to the next layer through weighting, biasing, and activation (29). The activation of the first six layers is accomplished by applying the "Rectified Linear Unit" (ReLU) to intensify the non-linear regression before the output layer. A sigmoid function accomplishes the activation of the output layer.

Equation **10** shows the forward propagation process in the first six layers, where the weight matrix and bias matrix are denoted using W and b, respectively. The y_h denotes the output of the 2nd, 3rd, 4th, 5th, and 6th layers. Equation **11** shows the forward propagation in the output layer. The y_{out} stands for the output layer. In the output layer, final weighting and biasing are implemented. Then the result will be plugged into a sigmoid

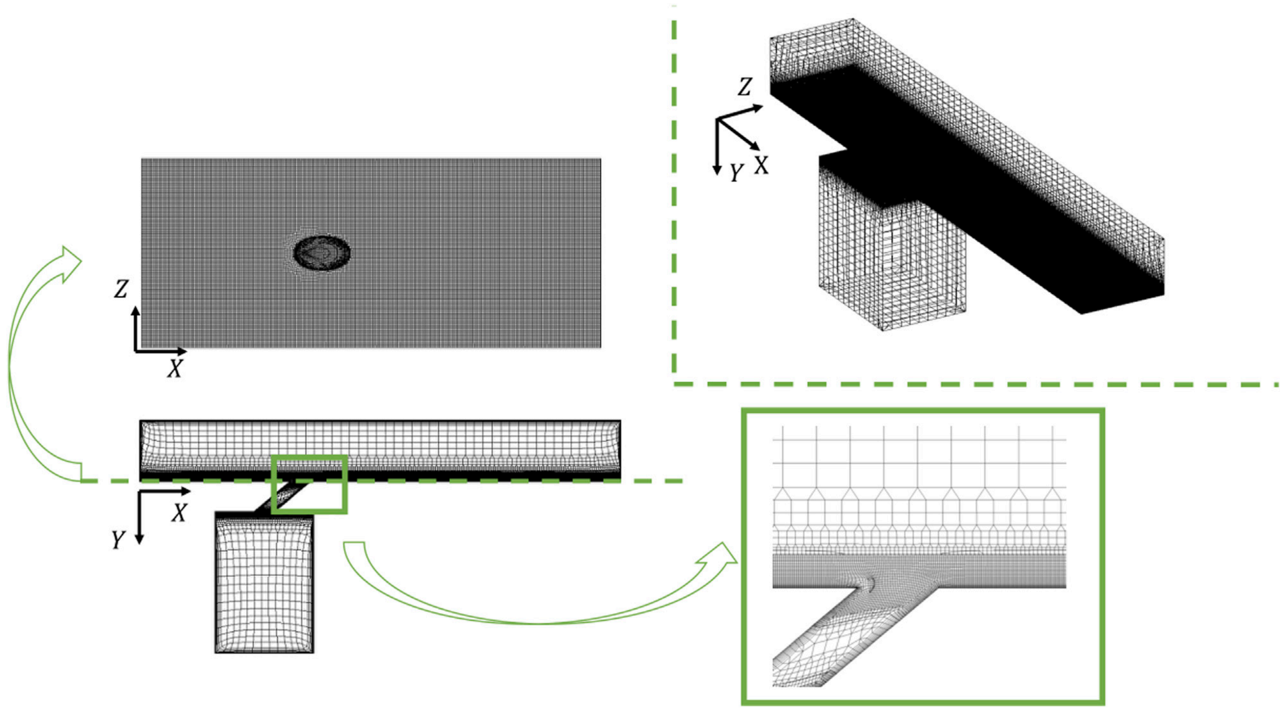

FIGURE 4 | Mesh schematic with enlarged part.

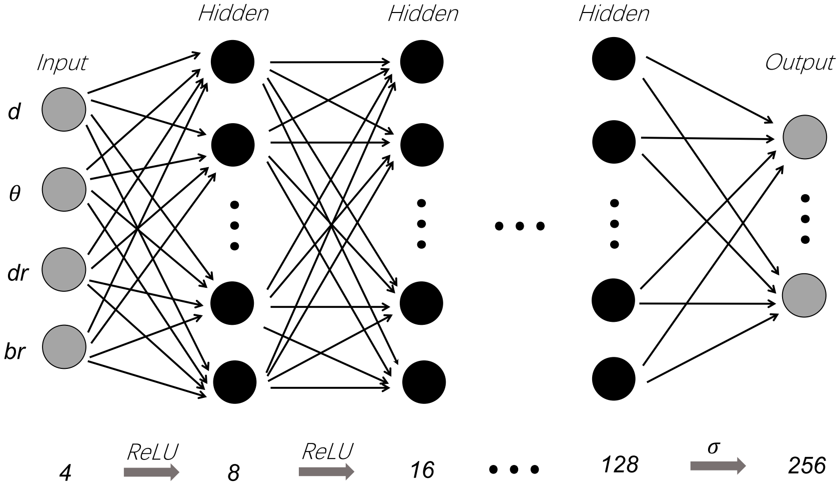

FIGURE 5 | Structure of the 7-layer ANN model.

function to generate the predicted temperature distribution as the output value.

Batch normalization and dropout are implemented in the five hidden layers to enhance learning and avoid the neural network collapsing by big data. Mean square error (MSE) is deployed as a loss function to assess the difference between predicted temperature T and the CFD simulated value P, as shown in Eq. **12**. The dropout is set to 0.1 in case of overfitting. The regression starts from 4 neurons and then multiplies until 256 neurons in the output layer. **Figure 5** graphically illustrates the regression process. 45,176 parameters in total are yet to be defined.

$$\begin{cases} \boldsymbol{ReLU}(\boldsymbol{x}) = \begin{cases} \mathbf{0}, \boldsymbol{x} < \mathbf{0} \\ \boldsymbol{x}, \boldsymbol{x} \geq \mathbf{0} \end{cases} \\ \boldsymbol{y_h} = \boldsymbol{ReLU}(\boldsymbol{Wx} + \boldsymbol{b}) \end{cases} \tag{10}$$

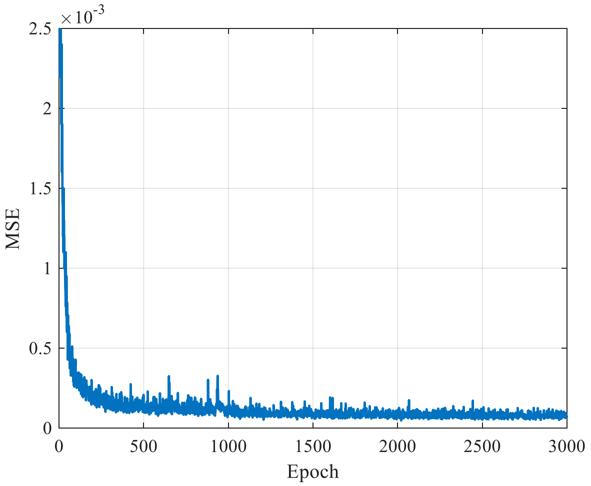

FIGURE 6 | Loss of the ANN model.

$$\begin{cases} \sigma(x) = \dfrac{1}{e^{-x}+1} \\ y_{out} = \sigma(Wx+b) \end{cases} \tag{11}$$

$$MSE = |T - P|^2 \tag{12}$$

Training and Validation

In the training process, the learning rate is one of the most significant hyperparameters that needs to be determined when applying ANN (36, 37). In this study, the learning rate is set to 0.01, and it is reset to 10% of its former value after every 1,000 epochs to keep convergence. The MSE is used to be the error index to evaluate whether the convergence of the model is accomplished. **Figure 6** shows that as the number of epochs increases, the MSE value converges. There is a noticeable fluctuation before the epoch reaches around 500. However, with the continuous growth of epochs, the value of MSE and the fluctuation keep falling.

According to the result, the 0.01 learning rate and reduction to 10% of its last value every 1,000 epochs is a good choice. Quoted error (QE) is proposed in Eq. **13** to measure the ANN model's absolute error. QE also serves as an index to help find the optimal hyperparameter settings. In the expression of QE, T is the temperature derived from CFD simulation, and P is the predicted temperature result at the final layer of the ANN. η^*_{min} is a fixed number 0, and the upper bound of η^*_{max} is 1. The result is scaled to a hundred percent of the corresponding QE value.

$$QE = \frac{|T - P|}{(\eta^*_{max} - \eta^*_{min})} \times 100\% \tag{13}$$

The ANN model uses QE to find the optimized hyperparameter settings containing splitting ratio, dropout, and batch size. The splitting ratio is acquired by dividing the amount of validation dataset by the whole dataset. A higher splitting ratio means a relatively lower portion of data is utilized in the training process. Thus, a balance in training and validation must be determined to achieve better prediction efficiency and accuracy. An experiment aiming at finding the optimal splitting ratio is conducted, and the result is shown in **Figure 7A**. The QE value of the case with splitting ratio = 0.2 is the highest compared with the situations whose splitting ratios are 0.1 and 0.3, which indicates that there is no simple proportional relationship between the magnitude of the error and the magnitude of splitting ratio. Besides, it is found that the QE value of the validation dataset is always higher than the training dataset by the ANN algorithm, regardless of which splitting ratio is chosen. Results show that a splitting ratio of 0.3 is believed to be the best among the three splitting ratios researched.

Similar experiments are designed for dropout and batch size. **Figure 7B** is the result of finding the ideal dropout. There is an apparent disparity between the case where the dropout is 0.1 compared to cases with 0.2 and 0.3 dropouts. The QE for the training dataset with 0.1 dropout is 0.31%, and the QE for the validation dataset with 0.1 dropouts is 0.33%. The QEs for the other two cases vary from 0.52% to 0.55%, which are much larger than those with a dropout of 0.1. This proves that the case with a dropout of 0.1 has the least error between the predicted temperature and the simulated value.

For Batch size, three different batch sizes are studied, and **Figure 7C** plots the result. The QE values of three different batch sizes are all located at [0.31%,0.38%], which implies that the batch size does not serve as a significant parameter with a considerable influence on the output in the ANN model. In the case with 32 batch size, the MSEs for the validation and training datasets are 0.33% and 0.31%, respectively. Therefore, the optimal batch size of 32 is chosen. For learning rate, as shown in **Figure 7D**, it is easily found that the varied learning rate could provide the best performance over the fixed learning rate such as 0.01, 0.001, and 0.0001. In all, the optimal hyperparameters are: dropout = 0.1, batch size = 32, splitting ratio = 0.3, and a varied learning rate.

Moreover, the structure of the ANN has been investigated to obtain the best performance. With these optimal parameters, the number of layers is determined to be 7 and the modes are determined according to the symmetry aiming for best training results. The information of different layers and nodes selection is shown in **Table 2** below. We chose 2, 3, 4, 7, and 9 layers because under these cases, the nodes arrangements are symmetric from 4 nodes to 256 nodes. In the experiment, we find that when the layer number comes to 7, the quoted error value (QE) reaches its minimum. However, we find that the quoted value of the 9-layer case increases, which is because too much layer increases the number of parameters and the iteration error. In the process of training, it is also reasonable in practice that some parameters are lost. After comparison, we find that the QE value of the 7-layer case is quite acceptable, so we consider 7 layers as the best layer number choice.

To further visualize the prediction accuracy using the designed ANN model, the fixed-cord-averaged film cooling effectiveness achieved by the ANN and CFD method simulation are compared for training and validation datasets. Six cases with diverse input parameter sets containing θ, d, dr, and br are randomly selected. The fixed-cord-averaged film cooling effectiveness on the upper

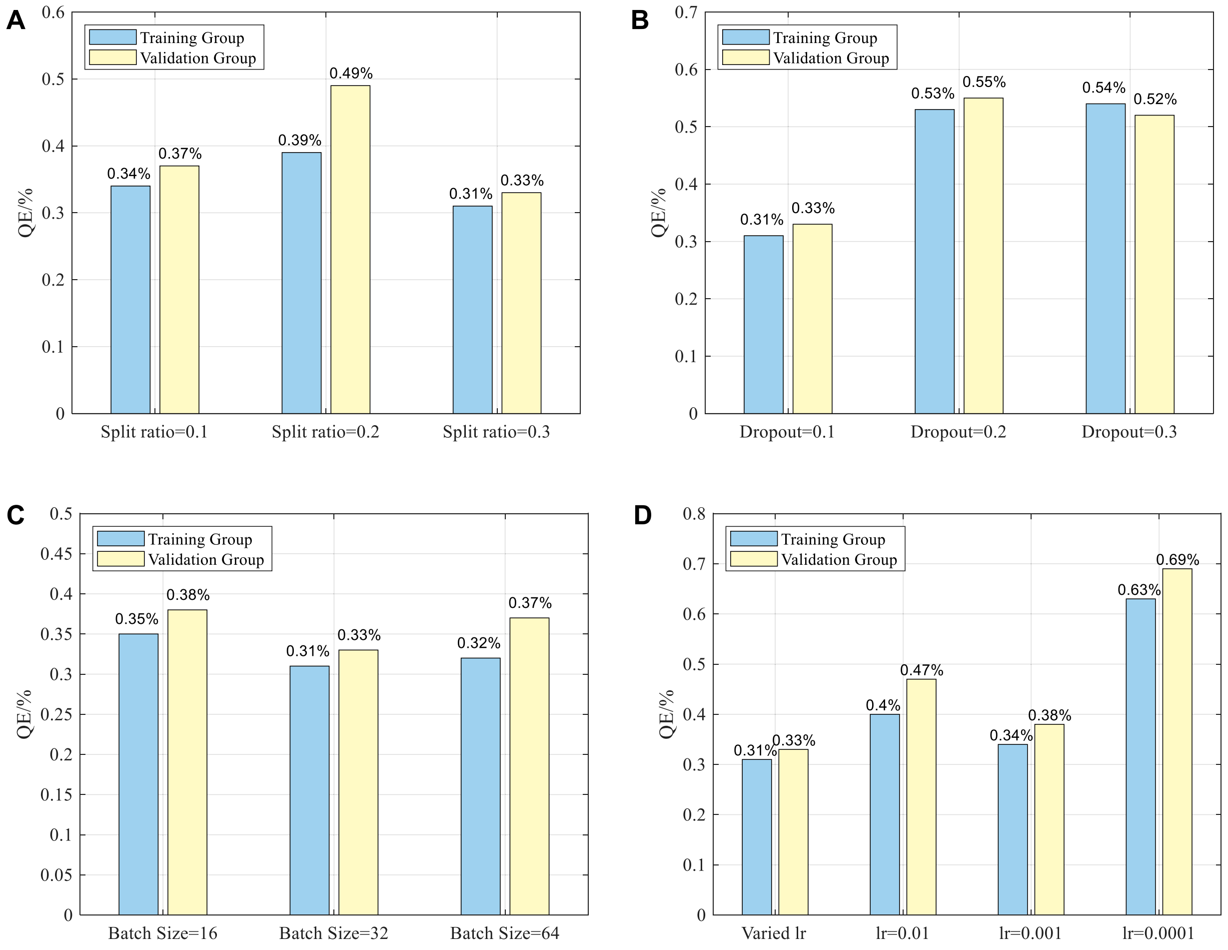

FIGURE 7 | QE values under different hyperparameters setups: **(A)** Splitting ratio; **(B)** Dropout; **(C)** Batch size; **(D)** Learning rate.

TABLE 2 | Layers and nodes information and corresponding QE values.

Number of layers	Nodes details	QE for training group (%)	QE for validation group (%)
2	4-256	0.91	0.96
3	4-128-256	0.35	0.40
4	4-16-64-256	0.32	0.36
7	4-8-16-32-64-128-256	0.31	0.33
9	4-8-16-16-32-64-64-128-256	0.34	0.45

surface of the turbine blade under six different cases are shown in **Figure 8**. All the six plots in **Figure 8** are about the fixed-cord-averaged cooling effectiveness results. The blue dotted lines in the plots represent the CFD cooling results, while the red line indicates the ANN results.

Figure 8A is for training datasets, and **Figure 8B** is for validation datasets. The result shows that the temperatures predicted using the ANN model for the training and validation datasets almost overlap with the temperatures obtained from CDF simulation. For example, in the case where $d = 10.5\ mm$, $\theta = 15$ °, $dr = 0.3$, $br = 0.5$, the fixed-cord-averaged cooling effectiveness derived by CFD is 0.108 when x/d_0 is positive zero, compared with the result derived by the ANN model of 0.130. While x/d_0 increases up to 15, the two curves almost overlap, and the error is negligible. After x/d_0 goes beyond 15, the error expands to its maximum value of 4.28%, where x/d_0 equals 17.4. Then, as x/d_0 keeps increasing, the error minimizes continuously. At the back end of the research region, the error is close to zero again. For fixed-cord-averaged film cooling effectiveness, the QE for training datasets is 0.29%, while for the validation dataset it is 0.32%. Conclusion can be drawn that even though fluctuation still occurs inevitably, the regression

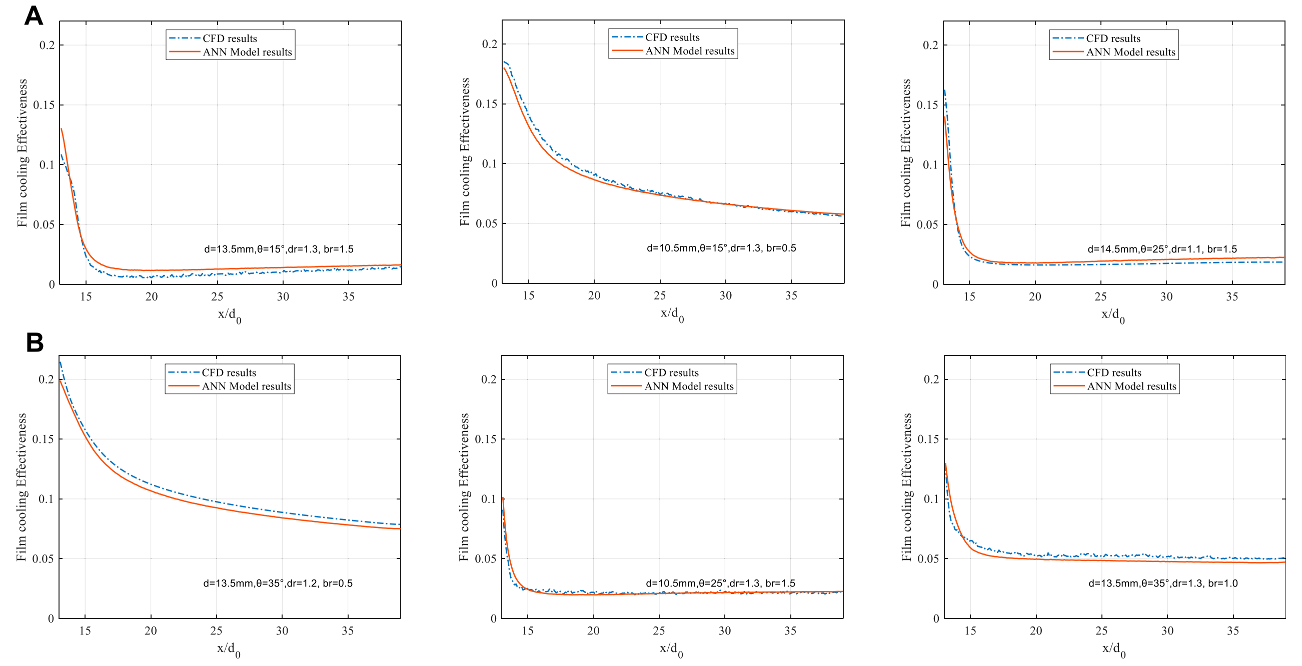

FIGURE 8 | Fixed-cord-averaged cooling effectiveness results comparison by the ANN and CFD model; **(A)** Training datasets; **(B)** Validation datasets.

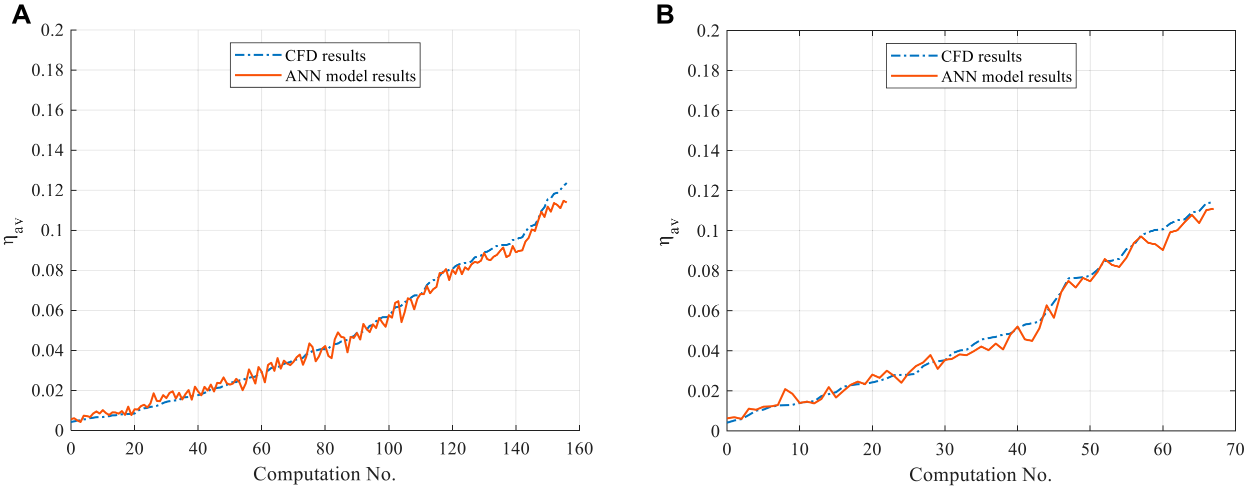

FIGURE 9 | General results comparison by the ANN and CFD model; **(A)** Training datasets; **(B)** Validation datasets.

performance of the ANN model in fixed-cord-averaged film cooling effectiveness is delighted under all randomly selected cases.

Besides the fixed-cord-averaged cooling effectiveness, the general cooling effectiveness is also studied to enhance the above conclusion. **Figure 9** compares the overall film cooling effectiveness from CFD and ANN methods downstream the coolant hole. The relationship between general film cooling effectiveness and computation numbers is plotted in **Figure 9**. 225 (3 × 5 × 5 × 3) cases of 0.3 splitting ratio are included, which means that **Figure 9A** includes the 156 training data, **Figure 9B** includes the 69 validation data. And they are ranked according to the magnitude of their general film cooling effectiveness to exhibit the results in a visual-friendly way. We labeled each data point using Computation No. from 1 to 156 and from 1 to 69 as shown on the horizontal axis. So, two lines are both discrete lines. Even though it is inevitable for the ANN and CFD results to have differences because the ANN's training data are derived from CFD. The results show that the general film cooling effectiveness derived from the ANN model is accurate enough for both training

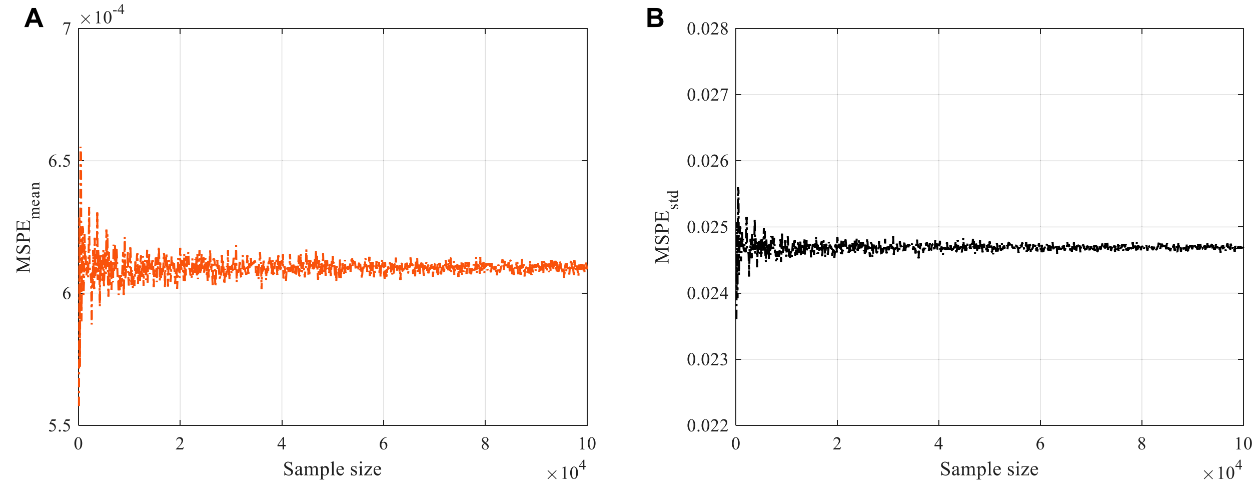

FIGURE 10 | MSPE values for MC for diverse sample size: **(A)** Mean of MSPE **(B)** Standard deviation of MSPE.

and validation data, regardless of the computation number. For general film cooling effectiveness, the QE for training datasets is 0.35%, while for the validation dataset, it is 0.30%.

UNCERTAINTY ANALYSIS

Monte Carlo Simulation and Sample Size

As the designing and manufacturing processes are deterministic, the variety of structures is usually not considered. For products with simple designs, the function could be achieved. However, the performance of sophisticated appliances such as gas turbines could vary significantly due to the uncertain deviation of their parameters. The best way to analyze and reduce unexpected uncertainty is to conduct an uncertainty quantification analysis of all related parameters (38, 39). In this paper, the parameters studied are d, θ, dr, and br.

Monte Carlo (MC) simulation is an extensively utilized technique to quantify the engineering field's uncertainty among various kinds of uncertainty quantification methods (40). MC simulation conducts statistical analysis to sample datasets obtained by repeated random sampling (41). MC simulation shows excellent advantage in mathematics and engineering due to its concise methodology, broad application domain, and various software choices (38). The mean and standard deviation of the mean squared pure error (MSPE), namely $MSPE_\mu$ and $MSPE_\sigma$ are studied instead to impair its harmful effect. The random samples are generated in the following way: Firstly, a sample vector $\vec{x}$ is formed from the normalized training dataset, which is $\vec{x} = (d, \theta, dr, br)$, and each entity is normalized with a value ranging from 0 to 1 according to Eq. **9**. Secondly, enough sample vectors $\vec{x}$ are generated in the same way and are arranged to form a distribution $\boldsymbol{X}$ for the following MC simulation. The sample size is defined as the number of sample vectors in each distribution. Therefore, parallel experiments are conducted, and the MSPE is proposed to quantify how well the MC simulation behaves. The mean and the standard of MSPE are defined in Eq. **14**. The experiment results are shown in **Figure 10**. It can be observed that even though the $MSPE_\sigma$ and $MSPE_\mu$ almost remain at a steady level, when the sample size is below 10,000, the fluctuation of both $MSPE_\sigma$ and $MSPE_\mu$ are very large, which represents that a sample size smaller than 10,000 is not suitable for MC simulation. As the sample size increases, however, the undulations of both the mean and the standard deviation narrow and remain stable continuously. Therefore, 10,000 is chosen as the optimal sample size for this research.

$$\begin{cases} y_{avg} = \dfrac{\sum_{i=1}^{n} y_i}{n} \\ MSPE_\sigma = \sqrt{\dfrac{1}{n-1} \times \sum_{i=1}^{n} \left(y_i - y_{avg}\right)^2} \\ MSPE_\mu = \dfrac{1}{n-1} \sum_{i=1}^{n} \left(y_i - y_{avg}\right)^2 \end{cases} \tag{14}$$

Sensitivity Analysis by Sobol Method

After the parameter distribution is settled, the Sobol method is utilized to investigate how the sensitivity of br, d, θ influences the film cooling effectiveness under three different values of dr. The Sobol method shows excellent performance in the sensitivity analysis (42, 43). The sample vector x is firstly transferred into the uncertainty input vector $\{X_1, X_2, X_3\}$, where X_1, X_2, *and* X_3 denote coolant hole diameter, inclination angle, and density ratio, respectively. A functional mapping $Y = f(X)$ is constructed to represent the relation between X and Y, where Y is the general film cooling effectiveness, η_{av}, defined in Eq. **7**.

Sobol indices are defined in Eq. **15**, where S_i and S_{Ti} are the first and total-effect variance-based Sobol indices, respectively. The Sobol index is used to quantify the sensitivity of a parameter. The larger the Sobol index is, the more significant impact its corresponding parameter has. X_i is the input parameter among the uncertainty

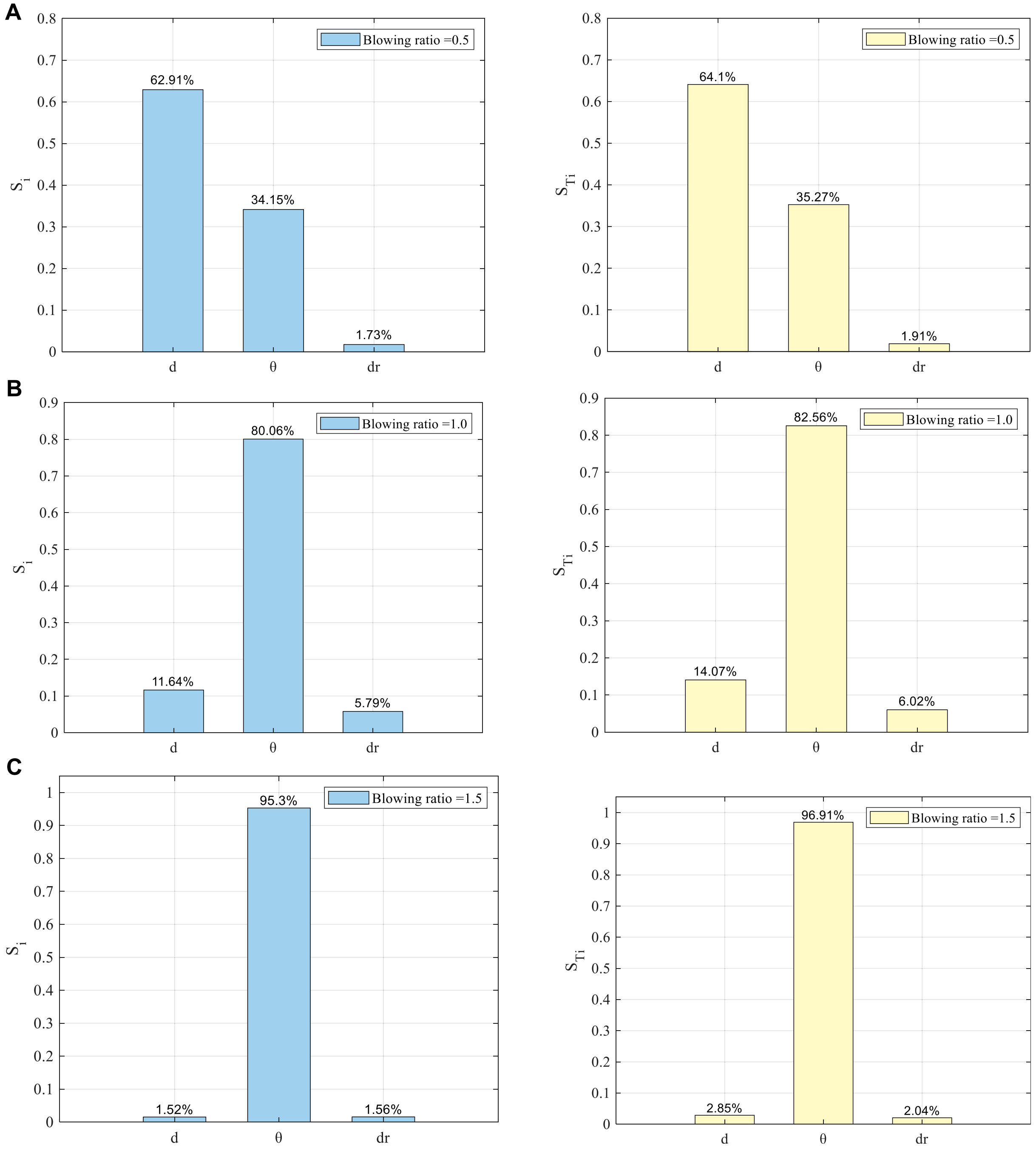

FIGURE 11 | Sodol Indices for *d*, θ, *dr* at **(A)** *br* = 0.5; **(B)** *br* = 1.0; **(C)** *br* = 1.5.

input set $\{X_1, X_2, X_3\}$, $X_{\sim i}$ denotes the set of all the input variables but X_i. By applying uncertainty deviation to each of the four-parameter sets, Y is affected to vary accordingly. Hence, the variances of Y under diverse scenarios are defined in Eq. **16**.

$$\begin{cases} S_i = \frac{1}{Cov(Y,Y)} \times Cov_{X_i}\left(E_{X_{\sim 1}}(Y|X_i), E_{X_{\sim 1}}(Y|X_i)\right) \\ S_{Ti} = \dfrac{Cov(Y,Y) - Cov_{X_i}\left(E_{X_{\sim i}}(Y|X_i), E_{X_{\sim i}}(Y|X_i)\right)}{Cov(Y,Y)} \end{cases} \tag{15}$$

$$\begin{cases} V_i = Cov_{X_i}\left(E_{X_{\sim i}}(Y|X_i), E_{X_{\sim i}}(Y|X_i)\right) \\ V_{ij} = Cov\left(E_{X_{\sim ij}}\left(Y|X_i,X_j\right), E_{X_{\sim ij}}\left(Y|X_i,X_j\right)\right) - V_i - V_j \\ Cov(Y,Y) = \sum_{i=1}^{n} V_i + \sum_{i<j}^{n} V_{ij} + V_{1,2,\ldots,n} \end{cases} \tag{16}$$

In addition to S_i and S_{Ti}, the temperature distribution near the blade also needs attention. If the blade surface is well covered by the cooling jet, its life span and reliability could be extended and increased heavily. To set up computation, the Probability Distribution Function (PDF) is introduced to represent the distribution of cooling effectiveness of every grid point on the flat surface of the turbine blade. The PDF are defined in Eq. **17**. Where y denotes the general film cooling effectiveness predicted by the ANN model, μ and σ is the mean value and standard deviation of the y.

$$\begin{cases} PDF(y) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(y-\mu)^2}{2\sigma^2}} \\ \mu = \frac{1}{n}\sum_{i=1}^{n} y_i \\ \sigma = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \mu)^2} \end{cases} \tag{17}$$

Test Case Definition and Turbulence Model Selection

By applying the uncertainty Sobol method to represent the three geometric parameters at blowing ratios of [0.5, 1.0, 1.5], the results of experiments are plotted in **Figures 11A–C**.

Figure 11A shows that at the low blowing ratio, the coolant hole diameter d has the largest value for both the S_i and S_{Ti}. It means that among the three geometric parameters, the deviation of coolant hole diameter has the largest uncertainty influence on the global film cooling effectiveness at a low blowing ratio. Followed by the hole diameter is the coolant jet inclination angle. Regardless of S_i or S_{Ti}, the uncertainty Sobol indices are both around 35%. The deviation of the density ratio is the weakest of all three parameters, the Sobol indices of which are less than 2%. In all, when the blowing ratio is set to 0.5, the trends of S_i and S_{Ti} for d and dr show high similarity. d is the dominant index which composites nearly 64%, θ yields around 35%, and dr has the most minimal effect on the uncertainty result.

For the medium-blowing ratio of 1.0, the order of the impact of the three geometric parameters changes utterly. **Figure 11B** shows that the difference between the S_i and the S_{Ti} is tiny. Still, the θ affects the result far more than the other two parameters. The S_i and S_{Ti} of θ both exceed 80%, which are approximately eight and fifteen times more than those of d and dr. It can be concluded that when the br is 1.0, the uncertain deviation of θ has the most significant impact on the general film cooling effectiveness.

Figure 11C indicates that for a large blowing ratio of 1.5, the corresponding result can be interpreted as a magnified one of a blowing ratio of 1.0. The coolant hole inclination angle is still the primary parameter, but its Sobol index is furthermore prominent than the medium-blowing ratio case. This time, the first-order and overall Sobol indices for θ outstrip 95%. Compared with the Sobol indices under the medium blowing ratio br, the hole diameter d and density ratio dr, which are all less than 3%, show negligible influence on the general cooling effectiveness. Moreover, it is found that when the blowing ratio increases, the S_i and S_{Ti} for all parameters increase accordingly.

In the gas turbine application, in a low blowing ratio case, both the coolant hole diameter and inclination angle significantly impact cooling effectiveness. However, the uncertain deviation of coolant hole diameter has a more significant effect. As for a high blowing ratio, usually more than 1.5, the coolant hole inclination angle needs special attention. In medium and high blowing ratio cases, the coolant inclination angle dominates the results, and its dominant effect increases as the blowing ratio increases, in all three cases. The trends of the first-order and the total-effect index are very similar. Thus, the same conclusion can be drawn.

Figure 12 shows the flow field of the flat plane. Since there are up to 225 cases in total, a reference case for the flow field example is defined with the following parameters: $d = 14.0$ mm, $\theta = 40^\circ$, $dr = 1.2$, $br = 0.5$. **Figure 12A** is the top view. It can be found that the flow field is highly symmetric on the top surface, and the green region is the widest temperature region, which occupies more than half the length of the region downstream of the confluence on the centerline. Moreover, the lateral diffusion is not evident since the width of the coolant stream keeps nearly equal to the hole diameter. **Figure 12B** is the front view of the flow field. It is detectable that right downstream the coolant hole, a reverse vortex pair is generated that drives the coolant jet away from the blade. A portion of the coolant on the top blends into the mainstream. The flow field shows that as the two jets move forward, the convergence of the two streams becomes more pronounced. Thus it is true that the coolant jet performs well only at a limited distance downstream of the confluence point. The reference case has a blowing ratio of less than 1, which means that the mainstream moves faster than the coolant jet, so **Figure 12C** indicates in the cross-section view that the general effect of gas motion is moving from the coolant jet to the hot mainstream. This leads to the coolant jet's diffusion and thus significantly impairs the cooling effectiveness. While in the cross-section view, we can also find that the coolant jet, which stays near the surface of the turbine blade, is rotating in a closed loop, that ensures that coolant is not lifted away from the blade, and achieves better cooling effectiveness at the surface layer of the blades.

During the experiment and analysis, it is found that under the small blowing ratio case of $br = 0.5$, the coolant hole diameter d has more impact sensitivity on the general film cooling effectiveness, compared with the coolant inclination angle. However, under a larger blowing ratio of 1.0, the coolant inclination angle has more effect than hole diameter. To further explain the flow physics that account for this result, the flow field of the cases under low and high blowing ratios, including $d = 10.5$ mm and 11.5 mm, coolant inclination angles of 20° and 30°, and in both x-z and y-z direction are compared and

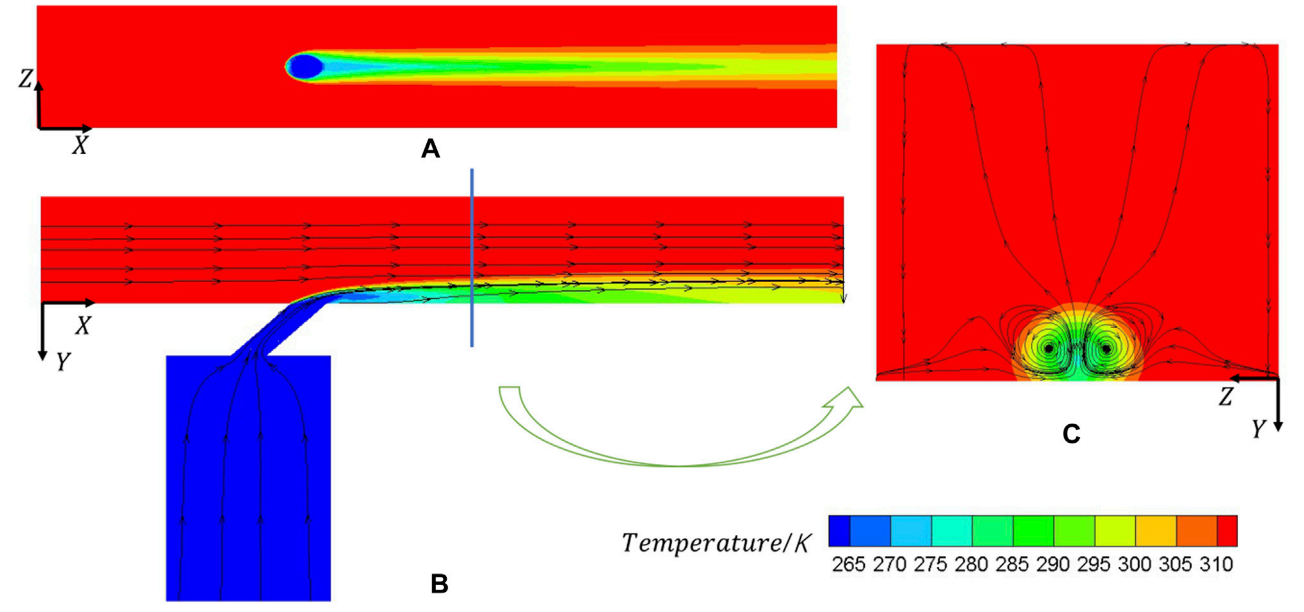

FIGURE 12 | Details of flow fields under different views: **(A)** X-Z; **(B)** X-Y; **(C)** Y-Z.

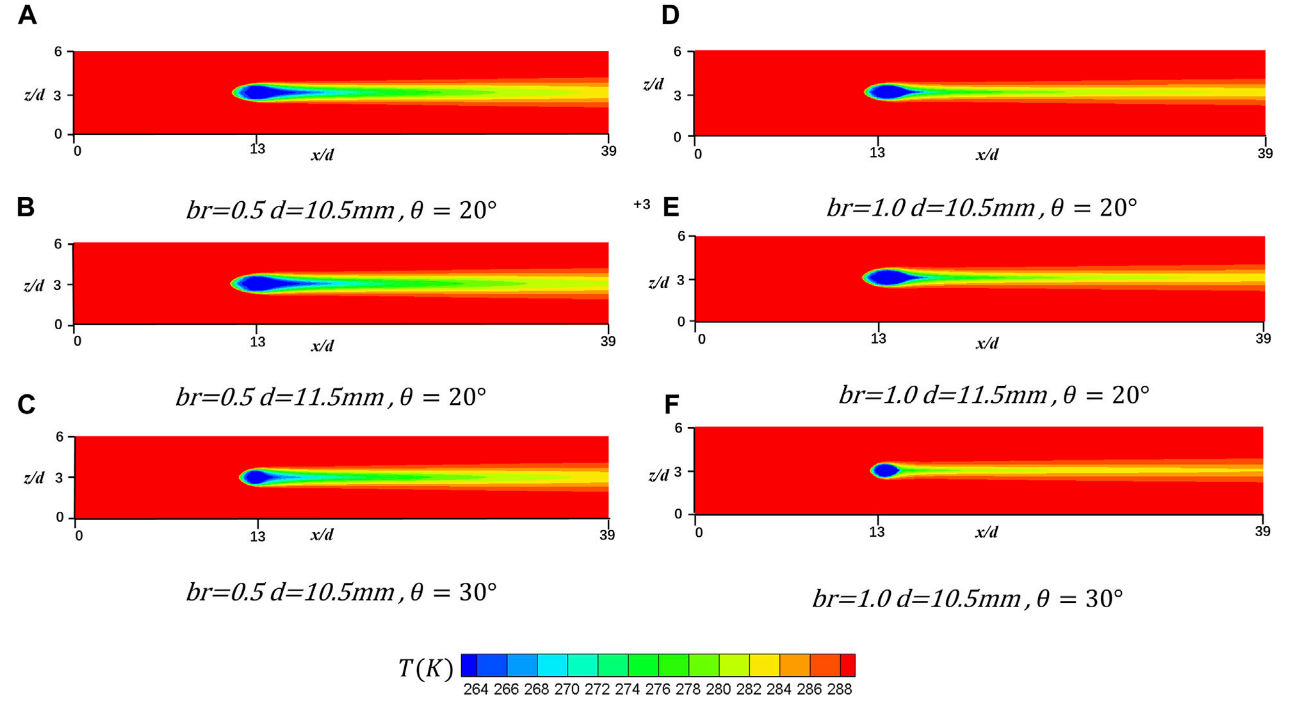

FIGURE 13 | Comparison of the top-view flow behavior at $y/d = 0$ in x-z plane. **(A)** $br = 0.5$, $d = 10.5$ mm, $\theta = 20^\circ$; **(B)** $br = 0.5$, $d = 11.5$ mm, $\theta = 20^\circ$; **(C)** $br = 0.5$, $d = 10.5$ mm, $\theta = 30^\circ$; **(D)** $br = 1.0$, $d = 10.5$ mm, $\theta = 20^\circ$; **(E)** $br = 1.0$, $d = 11.5$ mm, $\theta = 20^\circ$; **(F)** $br = 1.0$, $d = 10.5$ mm, $\theta = 30^\circ$.

analyzed. From a top view, by comparing **Figures 13A,B** and **Figures 13A,C**, we find that the low-temperature region at the center of the coolant hole expands to a larger degree under the influence of diameter change. This discovery implies that under a small blowing ratio of $br = 0.5$, the change in cooling effectiveness due to diameter change is larger than the change due to inclination angle. As for the case where the blowing ratio is as large as $br = 1.0$, **Figures 13D,E** and **Figures 13D,F** shows that

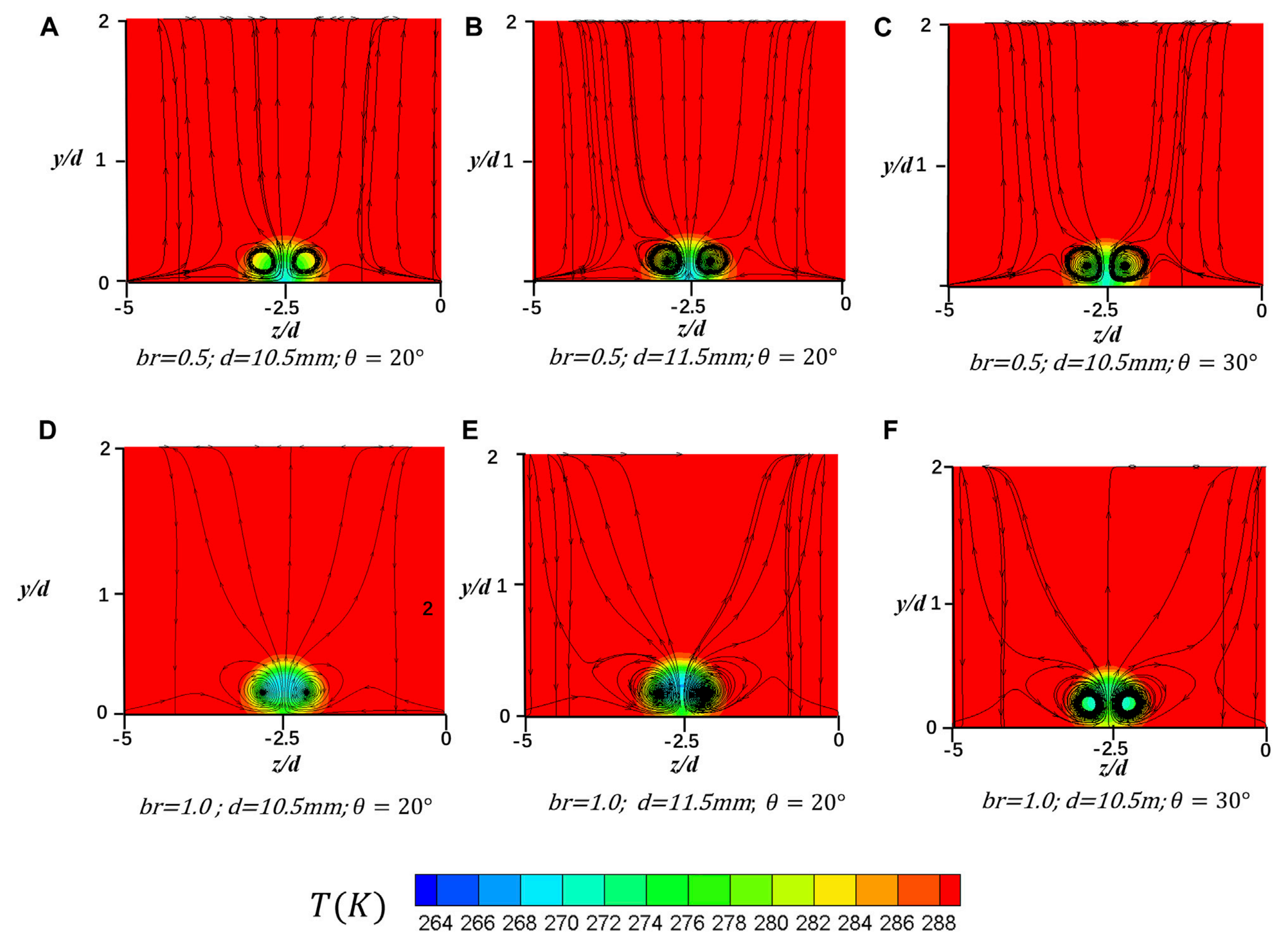

FIGURE 14 | Comparison of the side-view flow behavior at *x/d* = 16 in y-z plane. **(A)** *br* = 0.5, *d* = 10.5 mm, *θ* = 20°; **(B)** *br* = 0.5, *d* = 11.5 mm, *θ* = 20°; **(C)** *br* = 0.5, *d* = 10.5 mm, *θ* = 30°; **(D)** *br* = 1.0, *d* = 10.5 mm, *θ* = 20°; **(E)** *br* = 1.0, *d* = 11.5 mm, *θ* = 20°; **(F)** *br* = 1.0, *d* = 10.5 mm, *θ* = 30°.

the influence of the inclination angle is larger than the influence of the diameter, because the coolant flow in (F) shrinks dramatically, while the ones in (D) and (E) are similar to each other. **Figure 14** shows the cross-section view of the same 6 cases in **Figure 13**. When the blowing ratio is low, the difference of width of the CVP between **Figures 14A,B** is larger than the difference between **Figure 14A,C**; whereas, under a large blowing ratio, the coolant CVP is lifted to a larger degree in the case of increasing the angle instead of increasing thecoolant hole diameter. As shown in **Figure 14F**, the height of the CVP center climbs higher under the influence of the inclination angle compared with **Figure 14E**. Furthermore, the CVP is rotating upward, which results in the additional lifting of the coolant flow at the location of *x/d* = 19. This impairs the cooling effectiveness as well.

In addition to utilizing Sobol indices on behalf of the sensitivity, the control variate method is deployed to further research the individual effect of the three independent geometric parameters on the general effectiveness of cooling under three different blowing ratios in terms of the probability distribution(44). The reference values for *d*, *θ*, and *dr* are 12.5 mm, 35°, and 1.2 respectively. While varying one of the three, the others are set as the value in the reference case to achieve uniformity. **Figure 15** shows the probability distribution of general cooling effectiveness at different blowing ratios. The four subplots study the three single-parameter effects and the combined effect.

As shown in **Figure 15A**, the PDF distribution of the coolant hole diameter is studied. When the blowing ratio is 0.5, the 95% confidence interval due to hole diameter is [0.070, 0.113]. However, those for blowing ratios of 1.0 and 1.5 are [0.030, 0.049] and [0.014, 0.017], respectively. The corresponding interval length for the three blowing ratios are 0.043, 0.019, and 0.003, respectively. The result shows that as *br* is increasing, the 95% confidence interval due to hole diameter narrows rapidly, almost halving each time, and the mean of general cooling effectiveness decreases continuously, indicating that the uncertainty caused by hole diameter decreases.

For the hole inclination angle, as shown in **Figure 15B**, when the *br* increases, the uncertainty effect maintains a high level due to the hole inclination angle. The confidence intervals remain

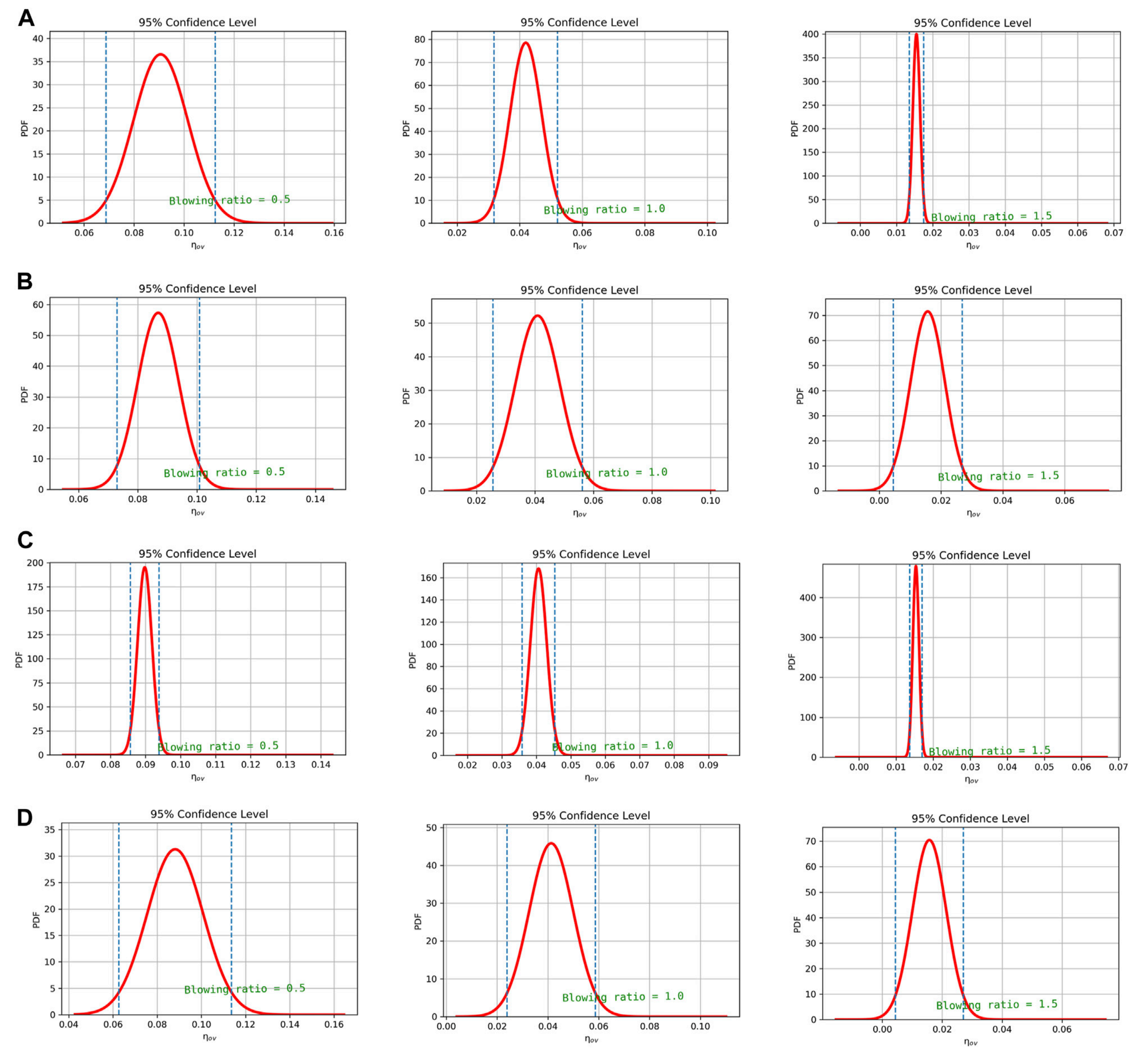

FIGURE 15 | PDF distribution of general film cooling effectiveness at three blowing ratios: **(A)** *d*; **(B)** θ; **(C)** *dr*; **(D)** Combined effect.

wide for all low, medium, and high blowing ratios. This is different from the case for hole diameter. With the increase of *br*, the range of variation of general film cooling effectiveness of inclination angle decreases from 0.030, 0.029 to 0.023 after simple calculation. It can be concluded that with the increasing of the blowing ratio, the uncertainty caused by the inclination angle θ reduces, but with a lower decreasing speed than the hole diameter.

Figure 15C shows that the magnitude of the length of the confidence interval of the density ratio is the smallest among all parameters. However, the *PDF* value of the density ratio is the highest, the minimum value of which still exceeds 160 in the case of the $br = 1.0$.

The uncertainty of the combined effect decreases as *br* rises, and the combined effect is more prominent than all three single effects. In **Figure 15D**, for example, the 95% confidence intervals are [0.062, 0.112], [0.023, 0.058], [0.004, 0.027], and the interval lengths are 0.060, 0.035, and 0.023, respectively. Compared with the single effect caused by the inclination angle θ, the interval length of the combination becomes wider.

CONCLUSION

This study aims to improve the gas turbine performance by strengthening the film cooling effectiveness, especially by focusing on the uncertainty of the three significant parameters, including single hole diameter, density ratio, and inclination angle, on the film cooling effectiveness under low, medium, and high blowing ratios. The uncertainty analysis was conducted using a deep-learning-based ANN model and uncertainty quantification method. Firstly, all related indices and research regions are defined at the beginning. Due to its best performance, the six-million grid size and the RNG k-ε model are chosen for the turbulence model. Secondly, a high-performance ANN model is delicately constructed for training and to seek the non-linear correlation between the parameter input and the cooling effectiveness output. CFD provides training and validation datasets. Finally, the sensitivity of three parameters is quantified, and uncertainty quantification is conducted to quantify the single and combined effect of the uncertainty of these three parameters on the general cooling effectiveness. The following conclusions are drawn.

1. After careful hyperparameter selection and training, the ANN model built in this study shows excellent performance in predicting the general and fixed-cord-averaged film cooling effectiveness according to input parameters compared with the data simulated by the CFD method. The QE value for fixed-cord-averaged film cooling effectiveness in training and validation datasets are 0.29% and 0.32%. The QE value for general film cooling effectiveness in training and validation datasets are 0.35% and 0.30%.
2. The Sobol method based on MC simulation shows that at a small blowing ratio, the coolant tube's diameter and inclination angle are two main factors to the cooling effectiveness, and the former has a more dominant effect. At medium and large blowing ratios, the inclination angle is the only leading factor to the film cooling effectiveness. Furthermore, the maximum effect of the inclination angle increases as the blowing ratio grows.
3. Uncertainty quantification reveals that the uncertainty of hole diameter, inclination angle, and density ratio all decrease as the blowing ratio rises. Moreover, the combined effect shows a higher impact on the general cooling effectiveness than any single effect. Within three parameters, the variation of the uncertainty interval of the hole diameter at three blowing ratios is the most obvious. Furthermore, the inclination angle θ has the most extensive uncertain influence on the general film cooling effectiveness among the three single parameters

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

AUTHOR CONTRIBUTIONS

YW: Conceptualization, data curation, formal analysis, investigation, methodology, resources, software, validation, and writing–original draft. XQ: Data curation, formal analysis, investigation, resources, visualization, and writing—original draft. SQ: Investigation, resources, visualization, and writing—original draft. YS: Data curation, formal analysis, methodology, visualization, and writing—original draft. WW: Resources, software, and validation. JC: Conceptualization, formal analysis, funding acquisition, investigation, project administration, software, supervision, visualization, writing—review and editing.

FUNDING

This study was supported in part by State Key Laboratory for Aerodynamics, the Zhejiang University/University of Illinois at Urbana-Champaign Institute and National Natural Science Foundation of China (Grant No. 52106060 and 92152202). It was led by Supervisor JC.

CONFLICT OF INTEREST

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

PUBLISHER'S NOTE

REFERENCES

1. Thamir KI, Ahmed NA. Improvement of gas turbine performance based on inlet air cooling systems: A technical review. *Int J Phys Sci* (2011) 6(4):620–7. doi:10.5897/IJPS10.563
2. Garg VK, Gaugler RE. Effect of coolant temperature and mass flow on film cooling of turbine blades. *Int J Heat mass transfer* (1997) 40(2):435–45. doi:10.1016/0017-9310(96)00040-3
3. Han JC, Ekkad S. Recent development in turbine blade film cooling. *Int J Rotating Machinery* (2001) 7(1):21–40. doi:10.1155/s1023621x01000033
4. Ekkad S, Han J-C. A Review of Hole Geometry and Coolant Density Effect on Film Cooling. In: *Proceedings of the ASME 2013 Heat Transfer Summer Conference*; 2013 Jul 14–19: Minneapolis, MN (2013).
5. Yang X, Zhang K, Wu J, Lei J, Su P, Fang Y. Numerical analysis of vane endwall film cooling and heat transfer with different mainstream turbulence intensities and blowing ratios. *Int J Therm Sci* (2022) 175:107482. doi:10.1016/j.ijthermalsci.2022.107482

6. Sinha AK, Bogard DG, Crawford ME. Film-cooling effectiveness downstream of a single row of holes with variable density ratio. *J Turbomach* (1991) 113. doi:10.1115/1.2927894
7. Cao N, Li X, Wu Z, Luo X. Effect of film hole geometry and blowing ratio on film cooling performance. *Appl Therm Eng* (2020) 165:114578. doi:10.1016/j.applthermaleng.2019.114578
8. Gritsch M, Schulz A, Wittig S. Adiabatic wall effectiveness measurements of film-cooling holes with expanded exits. *J Turbomach* (1998) 120. doi:10.1115/1.2841752
9. Guo X, Schröder W, Meinke M. Large-eddy simulations of film cooling flows. *Comput Fluids* (2006) 35(6):587–606. doi:10.1016/j.compfluid.2005.02.007
10. Cruse MW, Yuki UM, Bogard DG. Investigation of Various Parametric Influences on Leading Edge Film Cooling. In: *Proceedings of the ASME 1997 International Gas Turbine and Aeroengine Congress and Exhibition*; Jun 2–5; Orlando, FL. ASME (1997):V003T09A058.
11. Fu Z, Zhu H, Liu C, Wei J, Zhang B. Investigation of the influence of inclination angle and diffusion angle on the film cooling performance of chevron shaped hole. *J Therm Sci* (2018) 27(6):580–91. doi:10.1007/s11630-018-1070-8
12. Wang Y, Wang Z, Qian S, Wang W, Zheng Y, Cui J. Uncertainty quantification of the superposition film cooling with trench using supervised machine learning. *Int J Heat Mass Transfer* (2022) 198:123353. doi:10.1016/j.ijheatmasstransfer.2022.123353
13. Akbar M. The effects of coolant pipe geometry and flow conditions on turbine blade film cooling. *J Therm Eng* (2017) 3(3):1196–210. doi:10.18186/journal-of-thermal-engineering.314165
14. Shi W, Chen P, Li X, Ren J, Jiang H. Uncertainty quantification of the effects of small manufacturing deviations on film cooling: A fan-shaped hole. *Aerospace* (2019) 6(4):46. doi:10.3390/aerospace6040046
15. Mathioudakis K, Tsalavoutas T. Uncertainty reduction in gas turbine performance diagnostics by accounting for humidity effects. *J Eng Gas Turbines Power* (2002) 124(4):801–8. doi:10.1115/1.1470485
16. Huang M, Li Z, Li J, Song L. Uncertainty quantification and sensitivity analysis of aerothermal performance for the turbine blade squealer tip. *Int J Therm Sci* (2022) 175:107460. doi:10.1016/j.ijthermalsci.2022.107460
17. Mellor AM. Semi-empirical correlations for gas turbine emissions, ignition, and flame stabilization. *Prog Energ Combustion Sci* (1980) 6(4):347–58. doi:10.1016/0360-1285(80)90010-6
18. Awodele O, Jegede O. Neural Networks and Its Application in Engineering. In: *Proceeding of Informing Science + IT Education Conference 2009*; 2009 Jun 12–15, Macon, United States. InSITE (2009):83–95.
19. Wang C, Zhang J, Zhou J, Alting SA. Prediction of film-cooling effectiveness based on support vector machine. *Appl Therm Eng* (2015) 84:82–93. doi:10.1016/j.applthermaleng.2015.03.024
20. Dolati S, Amanifard N, Deylami HM. Numerical study and GMDH-type neural networks modeling of plasma actuator effects on the film cooling over a flat plate. *Appl Therm Eng* (2017) 123:734–45. doi:10.1016/j.applthermaleng.2017.05.149
21. Yang L, Chen W, Chyu MK. A convolution modeling method for pore plugging impact on transpiration cooling configurations perforated by straight holes. *Int J Heat Mass Transfer* (2018) 126:1057–66. doi:10.1016/j.ijheatmasstransfer.2018.06.068
22. Wang Y, Wang Z, Wang W, Li H, Shen W, Cui J. Predicting and optimizing multirow film cooling with trenches using gated recurrent unit neural network. *Phys Fluids* (2022) 34(4):045122. doi:10.1063/5.0088868
23. Wang Y, Qian S, Sun Y, Wang W, Cui J. Fast prediction and uncertainty analysis of film cooling with a semi-sphere vortex generator using artificial neural network. *AIP Adv* (2023) 13(1):015303. doi:10.1063/5.0132989
24. Wang Y, Wang Z, Wang W, Tao G, Shen W, Cui J. Two-dimensional prediction of the superposition film cooling with trench based on conditional generative adversarial network. *Int J Therm Sci* (2023) 184:107976. doi:10.1016/j.ijthermalsci.2022.107976
25. Wang Y, Wang W, Tao G, Zhang X, Luo S, Cui J. Two-dimensional film-cooling effectiveness prediction based on deconvolution neural network. *Int Commun Heat Mass Transfer* (2021) 129:105621. doi:10.1016/j.icheatmasstransfer.2021.105621
26. Yang L, Rao Y. Predicting the Adiabatic Effectiveness of Effusion Cooling by the Convolution Modeling Method. In: *Proceedings of the ASME Turbo Expo 2019: Turbomachinery Technical Conference and Exposition*; 2019 Jun 17–21. Phoenix, AZ. ASME (2019):V05AT12A004.
27. Wang Y, Luo S, Wang W, Tao G, Zhang X, Cui J. Film Cooling Prediction and Optimization Based on Deconvolution Neural Network. In: *High Performance Computing: ISC High Performance Digital 2021 International Workshops*; 2021 Jun 24–Jul 2; Frankfurt am Main, Germany. Springer International Publishing (2021):73–91.
28. Wang Y, Wang W, Tao G, Li H, Zheng Y, Cui J. Optimization of the semi-sphere vortex generator for film cooling using generative adversarial network. *Int J Heat Mass Transfer* (2022) 183:122026. doi:10.1016/j.ijheatmasstransfer.2021.122026
29. Yang L, Dai W, Rao Y, Chyu MK. A machine learning approach to quantify the film cooling superposition effect for effusion cooling structures. *Int J Therm Sci* (2021) 162:106774. doi:10.1016/j.ijthermalsci.2020.106774
30. Yepuri GB, Talanki Puttarangasetty AB, Kolke DK, Jesuraj F. Effect of RANS-type turbulence models on adiabatic film cooling effectiveness over a scaled up gas turbine blade leading edge surface. *J Inst Eng (India) Ser C* (2018) 99(4):393–400. doi:10.1007/s40032-016-0302-5
31. Wang W, Cui J, Qu S. Large-Eddy Simulation of Film Cooling Performance Enhancement Using Vortex Generator and Semi-Sphere. In: *Proceedings of the ASME Turbo Expo 2020: Turbomachinery Technical Conference and Exposition*; 2020 Sept 21–25. ASME (2020):V07BT12A028.
32. Guo X, Li W, Iorio F. Convolutional neural Networks for Steady Flow Approximation. In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*; 2016 Aug 13–17; San Francisco CA. ACM (2016):481–90.
33. Ito S, Goldstein RJ, Eckert ER. Film cooling of a gas turbine blade. *Int J Rotating Machinery.*(1978) 11. doi:10.1155/S1023621X01000033
34. Obara SY. Analysis of a fuel cell micro-grid with a small-scale wind turbine generator. *Int J Hydrogen Energ* (2007) 32(3):323–36. doi:10.1016/j.ijhydene.2006.07.032
35. Eggink H, Mertens P, Storm E, Giocomo LM. Hyperpolarization-activated cyclic nucleotide-gated 1 independent grid cell-phase precession in mice. *Hippocampus* (2014) 24(3):249–56. doi:10.1002/hipo.22231
36. McDonald A, Schrattenholzer L. Learning rates for energy technologies. *Energy policy* (2001) 29(4):255–61. doi:10.1016/s0301-4215(00)00122-1
37. Kumar M, Raghuwanshi NS, Singh R, Wallender WW, Pruitt WO. Estimating evapotranspiration using artificial neural network. *J irrigation drainage Eng* (2002) 128(4):224–33. doi:10.1061/(asce)0733-9437(2002)128:4(224)
38. Zou Z, Wang S, Liu H, Zhang W. *Axial turbine aerodynamics for aero-engines*. Singapore: Springer (2018).
39. Zang TA, Hemsch MJ, Hilburger MW, Kenny SP, Luckring JM, Maghami P, et al. *Needs and opportunities for uncertainty-based multidisciplinary design methods for aerospace vehicles*. Hampton, VA: Langley Res. Cent. Tech. Rep. NASA/TM-2002-211462 (2002).
40. Ferrenberg AM, Swendsen RH. Optimized Monte Carlo data analysis. *Comput Phys* (1989) 3(5):101–4. doi:10.1063/1.4822862
41. Raychaudhuri S. Introduction to Monte Carlo simulation. In: *2008 Winter Simulation Conference*; 2008 Dec 7-10; Miami, FL. IEEE (2008):91–100.
42. Sobol IM. Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Mathematics Comput simulation* (2001) 55(1-3):271–80. doi:10.1016/s0378-4754(00)00270-6
43. Gamannossi A, Amerini A, Mazzei L, Bacci T, Poggiali M, Andreini A. Uncertainty quantification of film cooling performance of an industrial gas turbine vane. *Entropy* (2019) 22(1):16. doi:10.3390/e22010016
44. Kucherenko S, Delpuech B, Iooss B, Tarantola S. Application of the control variate technique to estimation of total sensitivity indices. *Reliability Eng Syst Saf* (2015) 134:251–9. doi:10.1016/j.ress.2014.07.008

NOMENCLATURE

$\boldsymbol{d}$	film cooling diameter
$\boldsymbol{d_0}$	standard Value of d
$\boldsymbol{\theta}$	coolant tube inclination angle
$\boldsymbol{br}$	coolant-to-mainstream blowing ratio
$\boldsymbol{dr}$	coolant-to-mainstream density ratio
$\boldsymbol{\rho_c}$	density of coolant jet
$\boldsymbol{\rho_m}$	density of mainstream jet
$\boldsymbol{V_c}$	velocity of coolant jet
$\boldsymbol{V_m}$	velocity of mainstream jet
$\boldsymbol{T}$	gauged temperature
$\boldsymbol{T_c}$	temperature of coolant jet
$\boldsymbol{T_m}$	temperature of mainstream jet
$\boldsymbol{T^*}$	dimensionless temperature
$\overline{T^*}$	fixed-cord-averaged T^*
$\boldsymbol{\eta}$	film cooling effectiveness
$\bar{\eta}$	fixed-cord-averaged η
$\boldsymbol{\eta_{av}}$	general film cooling effectiveness
$\boldsymbol{MSE}$	mean square error
$\boldsymbol{QE}$	quoted error
$\boldsymbol{S_i}$	first-order sensitive index
$\boldsymbol{S_{Ti}}$	total-effect sensitive index
$\boldsymbol{MSPE}$	mean squared pure error
$MSPE_\sigma$	standard deviation of MSPE
$MSPE_\mu$	mean of the MSPE
$\boldsymbol{\mu}$	mean
$\boldsymbol{\sigma}$	standard deviation
PDF	probability distribution function

AEROSPACE RESEARCH COMMUNICATIONS
REVIEW
published: 17 July 2023
doi: 10.3389/arc.2023.11522

Overview of Advanced Numerical Methods Classified by Operation Dimensions

Xiao-Wei Gao*, Wei-Wu Jiang, Xiang-Bo Xu, Hua-Yu Liu, Kai Yang, Jun Lv and Miao Cui

School of Aeronautics and Astronautics, Dalian University of Technology, Dalian, China

In this article, the progress of frequently used advanced numerical methods is presented. According to the discretisation manner and manipulation dimensionality, these methods can be classified into four categories: volume-, surface-, line-, and point-operations–based methods. The volume-operation–based methods described in this article include the finite element method and element differential method; the surface-operation–based methods consist of the boundary element method and finite volume method; the line-operation–based methods cover the finite difference method and finite line method; and the point-operation–based methods mainly include the mesh free method and free element method. These methods have their own distinctive advantages in some specific disciplines. For example, the finite element method is the dominant method in solid mechanics, the finite volume method is extensively used in fluid mechanics, the boundary element method is more accurate and easier to use than other methods in fracture mechanics and infinite media, the mesh free method is more flexible for simulating varying and distorted geometries, and the newly developed free element and finite line methods are suitable for solving multi-physics coupling problems. This article provides a detailed conceptual description and typical applications of these promising methods, focusing on developments in recent years.

Keywords: finite element method, finite volume method, boundary element method, mesh free method, free element method, finite line method

***Correspondence:**
Xiao-Wei Gao
xwgao@dlut.edu.cn

***Received:** 28 April 2023
***Accepted:** 21 June 2023
***Published:** 17 July 2023

***Citation:**
Gao X-W, Jiang W-W, Xu X-B, Liu H-Y, Yang K, Lv J and Cui M (2023) Overview of Advanced Numerical Methods Classified by Operation Dimensions. Aerosp. Res. Commun. 1:11522. doi: 10.3389/arc.2023.11522*

INTRODUCTION

Most engineering problems can be represented by a set of second-order partial differential equations (PDEs) with relevant boundary conditions (B.C.), named the boundary value problem (BVP) of PDEs [1, 2]. For example, in thermal engineering, the diffusion-convection problem usually has the following BVP [3]:

$$\text{PDE: } \frac{\partial}{\partial x_i}\left(\lambda_{ij}(\boldsymbol{x})\frac{\partial T(\boldsymbol{x})}{\partial x_j}\right) + \frac{\partial \rho c v_i(\boldsymbol{x}) T(\boldsymbol{x})}{\partial x_i} + Q(\boldsymbol{x}) = 0,\ \boldsymbol{x} \in \Omega \tag{1}$$

$$\text{B.C.: } \begin{cases} T(\boldsymbol{x}) = \bar{T}(\boldsymbol{x}), & \boldsymbol{x} \in \Gamma_1 \\ -\lambda_{ij}(T,\boldsymbol{x})\dfrac{\partial T(\boldsymbol{x})}{\partial x_j} n_i(\boldsymbol{x}) = \bar{q}(\boldsymbol{x}), & \boldsymbol{x} \in \Gamma_2 \\ -\lambda_{ij}(T,\boldsymbol{x})\dfrac{\partial T(\boldsymbol{x})}{\partial x_j} n_i(\boldsymbol{x}) = h(\boldsymbol{x})(T(\boldsymbol{x}) - T_\infty), & \boldsymbol{x} \in \Gamma_3 \end{cases} \tag{2}$$

where T is the temperature, λ_{ij} the conductivity tensor, Q the heat source, v_i the velocity, $\bar{q}$ the specified heat flux, and h the heat transfer coefficient.

For solid mechanics problems, the BVP can be expressed as follows [4].

$$\text{PDE: } \frac{\partial}{\partial x_l}\left(D_{ijkl}(\boldsymbol{x})\frac{\partial u_k(\boldsymbol{x})}{\partial x_j}\right) + b_i(\boldsymbol{x}) = 0,\ \boldsymbol{x} \in \Omega \tag{3}$$

$$\text{B.C.: } \begin{cases} u_i(\boldsymbol{x}) = \bar{u}_i, & \boldsymbol{x} \in \Gamma_u \\ D_{ijkl}(\boldsymbol{x})n_j(\boldsymbol{x})\dfrac{\partial u_k(\boldsymbol{x})}{\partial x_l} = \bar{t}_i(\boldsymbol{x}), & \boldsymbol{x} \in \Gamma_t \end{cases} \tag{4}$$

where u_k is the displacement component, D_{ijkl} the constitutive tensor, b_i the body force, and $\bar{u}_i$ and $\bar{t}_i$ are the specified values of displacement and traction, respectively.

To solve the BVPs presented above, numerous numerical methods have been developed [5], which can be globally divided into four categories according to geometry discretisation and operation dimensions: volume-operation–based methods (including the finite element method [6, 7], finite block method [8], element differential method [4], etc.), surface-operation–based methods (including the boundary element method [9, 10], finite volume method [11], etc.), line-operation–based methods (including the finite difference method [12], finite line method [13], etc.), and point-operation–based methods (including the mesh free method [14], free element method [15], fundamental solution method [16], etc.). The classification of numerical methods into four categories as described above can help to deeply understand the innate characteristics of the various numerical methods. In these four types of numerical methods, most have two kinds of algorithms, the weak-form and strong-form algorithms [5]. As described in the article, the weak-form algorithms can be established by the weighted residual formulation, which requires integration over elements or divided sub-domains. Strong-form algorithms are based on the point collocation technique, which usually does not require integration computation. These four types of numerical methods will be described in the following sections.

VOLUME-OPERATION–BASED METHODS (VOBM)

Volume-operation–based methods refer to the methods performing the operations of PDEs based on a discretisation model that has the same size as the problem itself, i.e., 2 for two-dimensional (2D) and 3 for three-dimensional (3D) problems. The most commonly used VOBM is the finite element method (FEM), which is based on volume discretisation for 3D problems and plane region discretisation for 2D problems, respectively. In FEMs, the Galerkin FEM is the dominant method [6, 17], which establishes the solution scheme by using a variational principle in most publications. Nevertheless, as described in the article, all the weak-form algorithms, including the Galerkin FEM, can be derived by the weighted residual technique in a unified way, which will be described in the following.

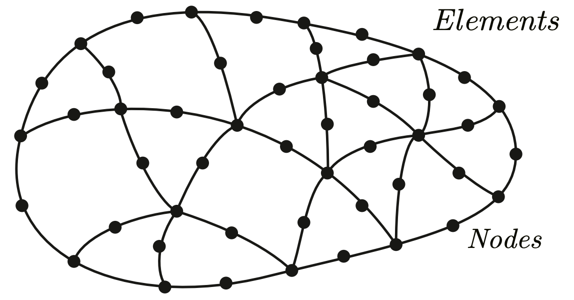

FIGURE 1 | Elements of a 2D FEM model.

Weighted Residual Formulation for Solving BVPs of PDEs

In the following, we take solid mechanics as a demonstration example for setting up the weighted residual formulation. To do this, multiplying the PDE **(3)** on both sides by a weight function w and integrating it through the computational domain Ω, it follows that

$$\int_\Omega w(\boldsymbol{x})\frac{\partial}{\partial x_l}\left(D_{ijkl}(\boldsymbol{x})\frac{\partial u_k(\boldsymbol{x})}{\partial x_j}\right)d\Omega + \int_\Omega w(\boldsymbol{x})b_i(\boldsymbol{x})d\Omega = 0 \tag{5}$$

Taking integration by parts and applying Gauss' divergence theorem to the first domain integral of Eq. **5**, the above equation becomes:

$$\int_\Omega \frac{\partial w(\boldsymbol{x})}{\partial x_l}D_{ijkl}(\boldsymbol{x})\frac{\partial u_k(\boldsymbol{x})}{\partial x_j}d\Omega = \int_\Gamma w(\boldsymbol{x})t_i(\boldsymbol{x})d\Gamma + \int_\Omega w(\boldsymbol{x})b_i(\boldsymbol{x})d\Omega \tag{6}$$

where t_i is the traction component on the boundary Γ of the domain Ω, which has the relationship with the displacement gradient shown in Eq. **4**.

In Eq. **6**, the basic physical variable u_k is mainly included in the volume integral of the left-hand side; therefore, it is called the volume-based weighted residual formulation. Taking integration by parts to the first domain integral of Eq. **6** and applying Gauss' divergence theorem again, the following equation can be obtained:

$$\int_\Gamma \frac{\partial w}{\partial x_l}D_{ijkl}n_j u_k d\Gamma - \int_\Omega \frac{\partial}{\partial x_j}\left(D_{ijkl}\frac{\partial w}{\partial x_l}\right)u_k d\Omega = \int_\Gamma w t_i d\Gamma + \int_\Omega w b_i d\Omega \tag{7}$$

In Eq. **7**, the basic physical variable u_k is included in both the surface and volume integrals of the left-hand side; therefore, it is called the surface-volume–based weighted residual formulation.

It is noted that Eqs **6**, **7** are valid for any sized closed domain Ω, and from this feature various weak-form solution algorithms can be generated, such as FEMs and BEMs, by taking different kinds of the weigh function w in an element or in the whole domain.

Weak-Form Finite Element Method (WFEM)

In FEM, the computational domain is discretized into a series of elements [6, 7] with a certain number of nodes. Usually, the nodes on the element interfaces should be linked point-to-point, as shown in **Figure 1**, for a 2D computation FEM model. Over each element, the displacement u_k is approximated using its nodal values u_k^α of the element by the shape function N_α as follows:

$$u_k = N_\alpha u_k^\alpha \tag{8}$$

where the repeated index α represents summation through all element nodes.

Galerkin Finite Element Method (GFEM)

In the Galerkin FEM, the weight function w in Eq. **6** is taken as the shape function N_c, i.e., $w = N_c$, with the subscript c representing the element nodal number corresponding to the collocation point c. Thus, for an element e with a domain Ω_e bounded by the boundary Γ_e, Eq. **6** results in the following element equation:

$$\underbrace{\int_{\Omega_e}\frac{\partial N_c}{\partial x_l}D_{ijkl}\frac{\partial N_\alpha}{\partial x_j}d\Omega\, u_k^\alpha}_{K_{ec}} = \underbrace{\int_{\Gamma_e} N_c t_i(\boldsymbol{x})d\Gamma}_{F_{ec}^t} + \underbrace{\int_{\Omega_e} N_c b_i d\Omega}_{F_{ec}^b} \tag{9}$$

where the left-hand side is related to the so-called element stiffness term and the right-hand side is the total equivalent load of element e.

We assume that the problem is discretized as N computational points, and each point is shared by a number of elements. Thus, for a point n, assembling all related elements' contributions from Eq. **9**, it follows that

$$\sum_{e_n c_n=1}^{E_n} K_{e_n c_n} = \sum_{e_n c_n=1}^{E_n} F_{e_n c_n}^t + \sum_{e_n c_n=1}^{E_n} F_{e_n c_n}^b \tag{10}$$

where E_n is the number of elements connected to point n and $e_n c_n$ represents the element e connected to point n at the element node c, which corresponds to point n. It is noted that for one element among all element nodes, only one node corresponds to point n.

In Eq. **10**, the equivalent traction load, the second term in Eq. **10**, has different values for interface and out boundary points, i.e.,

$$\begin{aligned} &\sum_{e_n c_n=1}^{E_n} F_{e_n c_n}^t = 0, \text{when } n \text{ is at interface nodes} \\ &\sum_{e_n c_n=1}^{E_n} F_{e_n c_n}^t = \sum_{e_n c_n=1}^{E_n} \int_{\Gamma_{e_n}} N_{c_n} t_i d\Gamma, \text{when } n \text{ is at out boundary nodes} \end{aligned} \tag{11}$$

where Γ_{en} is the out boundary of element e including point n and N_{c_n} is the shape function of the element node c corresponding to point n. The first equation in **(11)** comes from the fact that the equivalent traction loads from opposite surfaces of the related elements including the interface point are counteracted by one another, and the second equation in **(11)** relies on the characteristics of the shape function N_{c_n}, where its value is zero on the surfaces excluding node c_n.

When n in Eq. **10** goes through all the N points, the following system of equations can be produced in the matrix form:

$$\boldsymbol{K}\boldsymbol{u} = \boldsymbol{F} \tag{12}$$

where $\boldsymbol{K}$ is the global stiffness matrix, $\boldsymbol{u}$ the displacement vector, and $\boldsymbol{F}$ the total equivalent load vector.

The Galerkin FEM results in a symmetric and a banded sparse coefficient matrix $\boldsymbol{K}$ in the system of equations; this makes the method very efficient and stable. In particular, when some modern techniques are integrated into FEM, such as the control volume finite-element method [18, 19], isogemetric technique [20, 21], and gradient smoothing technique [22, 23], quite complicated engineering problems can be efficiently solved. Moreover, in recent years, a number of newly proposed FEMs have been developed, as described below.

Surface-Volume–Based Finite Element Method (SVFEM)

The Galerkin FEM presented above is derived based on the volume-based weighted residual formulation **(6)**. In the article, another type of FEM can be generated based on the surface-volume–based weighted residual formulation **(7)**, which has the same element discretisation as **Figure 1**. To do this, as done above, by applying Eq. **7** to an element domain, say element e, and by using $w = N_c$ and substituting Eq. **8** into Eq. **7**, it follows that

$$\begin{aligned} &\underbrace{\int_{\Gamma_e}\frac{\partial N_c}{\partial x_l}D_{ijkl}n_j N_\alpha d\Gamma u_k^\alpha}_{K_{ec}^\Gamma} - \underbrace{\int_{\Omega_e}\frac{\partial}{\partial x_j}\left(D_{ijkl}\frac{\partial N_c}{\partial x_l}\right)N_\alpha d\Omega\, u_k^\alpha}_{K_{ec}^\Omega} \\ &= \underbrace{\int_{\Gamma_e} N_c t_i(\boldsymbol{x})d\Gamma}_{F_{ec}^t} + \underbrace{\int_{\Omega_e} N_c b_i d\Omega}_{F_{ec}^b} \end{aligned} \tag{13}$$

Similar to Eq. **10**, for the computational point n, assembling all related elements' contributions from Eq. **13**, the following equation can be formed:

$$\sum_{e_n c_n=1}^{E_n} K_{e_n c_n}^\Gamma - \sum_{e_n c_n=1}^{E_n} K_{e_n c_n}^\Omega = \sum_{e_n c_n=1}^{E_n} F_{e_n c_n}^t + \sum_{e_n c_n=1}^{E_n} F_{e_n c_n}^b \tag{14}$$

The right-hand side of Eq. **14** is exactly the same as that of the standard Galerkin FEM, as shown in Eq. **11**.

When n in Eq. **14** goes through all the N discretized points, the following system of equations can be produced in the matrix form:

$$\tilde{\boldsymbol{K}}\boldsymbol{u} = \boldsymbol{F} \tag{15}$$

In the above equation, $\tilde{\boldsymbol{K}}$ is also called as the global stiffness matrix and is band-sparse, but it is not symmetric anymore; $\boldsymbol{u}$ and $\boldsymbol{F}$ are the same as in Eq. **12**.

Although $\tilde{\boldsymbol{K}}$ in SVFEM is not symmetric and this could reduce the solution efficiency to a certain level, Eq. **15** would give more accurate results than Eq. **12** since more mathematical treatments have been performed in Eq. **13** than in Eq. **9**.

The SVFEM presented above is included in the new research work by the article's authors. Some results have not yet been published in the literature; however, and they will be put in the public domain in the near future.

Strong-Form Volume-Operation–Based Methods

The above-described Galerkin FEM and SVFEM are weak-form algorithms, which require integration over elements to form the system of equations. In recent years, new types of strong-form FEM-like volume-operation–based methods have been proposed, which belong to a type of element collocation method and do not need integration computations. However, the stability of the strong-form algorithms is usually not as good as the weak-form algorithms, although for general problems these algorithms can still give satisfactory results.

Wen and Li et al. [8, 24] proposed the finite block method (FBM) in 2014, in which isoparametric element-like blocks are used to compute the first-order partial derivative of physical variables with respect to the global coordinates. FBM has the advantage of simple coding, and since element-like blocks are used, the stability of the solution is usually good. On the other hand, since all nodal values of physical variables over each block are independently inserted into the system of equations by introducing a consistent condition of physical variables and an equilibrium condition of the physical variable gradient in the system, there are more unknowns in the formed final system of equations than other frequently used numerical methods in the case of the same number of total nodes. In view of this issue, as few blocks as possible should be used when solving a problem using FBM to ensure that the final system of equations is not so large.

In the same period, Fantuzzi et al. [25, 26] proposed another type of strong-form FEM (SFEM), in which a set of formulations computing the first- and second-order spatial partial derivatives are derived for 2D problems and are used to collocate the governing PDEs in solid mechanics. In SFEM, the continuity condition among elements is determined by the compatibility, and a mapping technique is used to transform both the governing differential equations and the compatibility conditions between two adjacent sub-domains into the regular master element in the computational space. As in FBM, the treatment of the compatibility and equilibrium conditions between elements is still complicated; this makes SFEM not as flexible as GFEM when solving complicated engineering problems.

In 2017, Gao et al. [27] proposed a new type of strong-form FEM, called the Element Differential Method (EDM), for solving heat conduction problems, and later it was successfully used to solve solid mechanics [4, 28], electromagnetic [29], and thermo-mechanical-seepage coupled [30] problems. As in FBM, Lagrange polynomials are used to construct high-order elements in EDM. The essential difference between EDM and FBM is that both the first- and second-order partial derivatives were derived for 2D and 3D problems in EDM. The following is a brief review of EDM.

Looking back at Eq. **8** for physical variable interpolation over an element, the global coordinates can also be expressed by their nodal values and shape functions, as follows:

$$x_i = N_\alpha x_i^\alpha \tag{16}$$

Based on Eqs **8**, **16**, the following expressions can be derived for the first- and second-order partial derivatives [4, 27]:

$$\frac{\partial u_k}{\partial x_i} = d_i^{c\alpha'} u_k^{\alpha'} \tag{17}$$

$$\frac{\partial^2 u_k}{\partial x_i \partial x_j} = d_{ij}^{c\alpha''} u_k^{\alpha''} \tag{18}$$

where

$$d_i^{c\alpha'} = \frac{\partial N_\alpha}{\partial x_i} = [J]_{ik}^{-1} \frac{\partial N_\alpha}{\partial \xi_k} \tag{19a}$$

$$d_{ij}^{c\alpha''} = \frac{\partial^2 N_\alpha}{\partial x_i \partial x_j} = \left[[J]_{ik}^{-1} \frac{\partial^2 N_\alpha}{\partial \xi_k \partial \xi_l} + \frac{\partial [J]_{ik}^{-1}}{\partial \xi_l} \frac{\partial N_\alpha}{\partial \xi_k} \right] \frac{\partial \xi_l}{\partial x_j} \tag{19b}$$

where $[J] = [\partial \boldsymbol{x} / \partial \boldsymbol{\xi}]$ is the Jacobian matrix between the global coordinate $\boldsymbol{x}$ and the local coordinate $\boldsymbol{\xi}$ of the element. The detailed expressions for each term in Eqs **19a**, **19b** can be found in [27, 28].

The main advantage of the strong-form FEMs over weak-form FEMs is that the derived spatial partial derivatives can be directly substituted into the problem's PEDs and B.C. to set up the system of equations. For example, by using Eqs **17**, **18**, the PDE and B.C. for the solid mechanics shown in Eqs **3**, **4** can be directly used to generate the following discretized equations:

$$d_l^{c\beta'} D_{ijkl}\left(\boldsymbol{x}^{\beta'}\right) d_j^{\beta'\alpha'} u_k^{\alpha'} + b_i(\boldsymbol{x}^c) = 0, \boldsymbol{x}^c \in \Omega_e \tag{20}$$

$$\begin{cases} u_i(\boldsymbol{x}^c) = \bar{u}_i, \boldsymbol{x}^c \in \Gamma_u \\ D_{ijkl}(\boldsymbol{x}^c) n_j(\boldsymbol{x}^c) d_l^{c\alpha'} u_k^{\alpha'} = \bar{t}_i(\boldsymbol{x}^c), \boldsymbol{x}^c \in \Gamma_t \end{cases} \tag{21}$$

In the above equations, Γ_u and Γ_t are the out boundaries of the problem, over which the displacement and traction boundary conditions are specified. The big issue in the strong-form FEM is how to set up the discretized equation when the collocation point $\boldsymbol{x}^c$ is located on the interface Γ_I between elements. To solve this issue, Gao et al. [4, 27] proposed the summation-equilibrium technique for all related tractions, that is, $\sum\limits_{e=1,s=1} t_i^{es} = 0$, to form a single set of equations at an element interface node, which can be expressed as

$$\sum_{e=1,s=1}^{E_c} D_{ijkl}^{e}(\boldsymbol{x}^c) n_j^{es}(\boldsymbol{x}^c) d_l^{ec\alpha'} u_k^{e\alpha'} = 0, \boldsymbol{x}^c \in \Gamma_I \tag{22}$$

where E_c is the number of all elements connected with the interface collocation point c. Various numerical examples [27–31] have proved that the above equation can give correct results. The important point is that Eq. **22** allows the final system of equations to have the same size as the conventional FEM, which is much smaller than those in FBM [8] and SFEM [25].

SURFACE-OPERATION–BASED METHODS

Surface-operation–based numerical methods include the finite volume method (FVM), boundary element method (BEM), etc., which are operated mainly on the surfaces of a control volume or on the boundary of the considered problem.

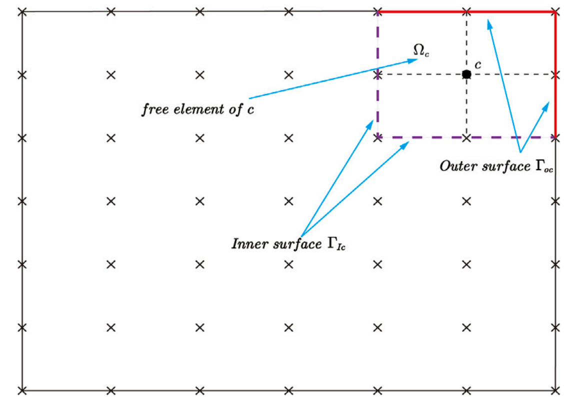

FIGURE 2 | Operation surfaces Γ_{Ic} and Γ_{oc} formed by the boundary of the free element built for collocation point *c*.

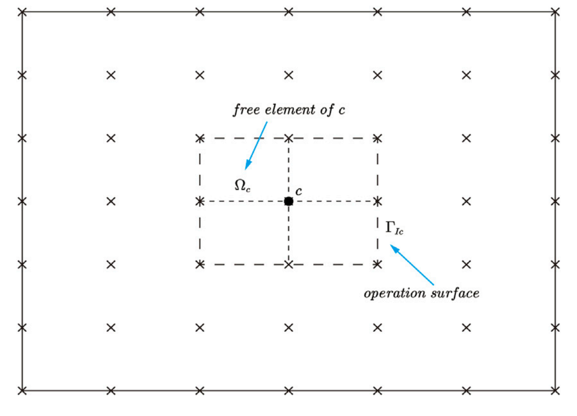

FIGURE 3 | Operation surface Γ_{Ic} of the free element completely included in the computational domain.

Finite Volume Method (FVM)

The FVM looks like a volume-based method [32–35]. However, in this article, it is classified into the category of surface-operation–based methods. This is because its main operation is over the surfaces of the control volume, rather than on the volume itself. To see this, let us take the weight function w to be 1 in Eq. **6**. This results in the following:

$$\int_{\Gamma} t_i d\Gamma + \int_{\Omega} b_i d\Omega = 0 \tag{23}$$

In FVM, the computational domain is discretized into a series of control volumes [32]. Applying Eq. **23** to each control volume, say volume Ω_c, and dividing its boundary $\partial\Omega_c$ into two parts, the inner boundary Γ_{Ic} and outer boundary Γ_{oc}, Eq. **23** can be written as:

$$\int_{\Gamma_{Ic}} D_{ijkl} n_j \frac{\partial u_k}{\partial x_l} d\Gamma + \int_{\Gamma_{oc}} \bar{t}_i d\Gamma + \int_{\Omega_c} b_i d\Omega = 0 \tag{24}$$

where $\Gamma_{Ic} \cup \Gamma_{oc} = \partial\Omega_c$.

Equation **24** is a typical formulation of FVM, from which we can see that the main computation is over the control surfaces of a control volume. The key work in FVM is the evaluation of the physical variable gradient $\partial u_k/\partial x_l$ included in the first control surface integral of Eq. **24** [36, 37]. In the conventional FVM, the interface Γ_{Ic} is taken as the mid-surface connected by the collocation point c and around-neighbour points; thus, $\partial u_k/\partial x_l$ at the mid-surface can be easily computed using the values of u_k between c and the neighbour points [38–40]. However, only the linear variation of u_k over the operation surface can be easily achieved. It is difficult to construct a high-order scheme to compute the value of $\partial u_k/\partial x_l$ on the operation surface. To overcome this problem, the free element [15, 31] can be used in FVM analysis.

Free Element-Based FVM (FEFVM)

In [15], the free element method (FrEM) was proposed for thermal-mechanical analysis. In FrEM, the isoparametric elements used in FEM are defined at each collocation point, as shown in **Figure 2**. The weak-form formulation of FrEM has the form shown in Eq. **24** [31]; however, the control volume is taken as the free element, as shown in **Figure 2**. Generally, for a free element, some of its operation surfaces are located inside the domain and some on the outer boundary of the problem, as shown in **Figure 2**. For this case, both the inner surface integral over Γ_{Ic} and outer surface integral over Γ_{oc} will appear in Eq. **24**.

However, when all the surfaces of the free element formed for the collocation c are located within the problem, as shown in **Figure 3**, only the inner surface integral exists.

In this case, Eq. **24** takes the following form:

$$\int_{\Gamma_{Ic}} D_{ijkl} n_j \frac{\partial u_k}{\partial x_l} d\Gamma + \int_{\Omega_c} b_i d\Omega = 0 \tag{25}$$

Since high-order free elements can be easily formed in FrEM [15], a high accuracy of $\partial u_k/\partial x_l$ in FEFVM can be easily achieved. A set of analytical expressions for computing $\partial u_k/\partial x_l$ over a free element have been derived in [4]. Although it is easy to set up a high-order free element, the accuracy of $\partial u_k/\partial x_l$ is not high. This is because its value is taken on the boundary of Ω_c, which is not as accurate as inside a free element. To overcome this drawback, the element-shell–strengthened FVM is proposed in the following section.

Element-Shell Enhanced FVM (ESFVM)

To improve the accuracy of $\partial u_k/\partial x_l$ included in the first boundary integral of Eq. **24**, additional free elements are formed for each side/surface of the collocation element Ω_c, which form an element ring/shell for 2D/3D control sides/surfaces, as shown in **Figure 4** for a 2D case.

In this strategy, the inner surface integral included in Eq. **24** can be written as follows:

$$\int_{\Gamma_{Ic}} D_{ijkl} n_j \frac{\partial u_k}{\partial x_l} d\Gamma = \sum_{s=1}^{N_{Is}} \int_{\Gamma_s} D_{ijkl} n_j \frac{\partial u_k}{\partial x_l} (x^s) d\Gamma(x^s) \tag{26}$$

where N_{Is} is the number of inner operation surfaces of the control volume Ω_c, which is 4 in the case shown in **Figure 4** and x^s denotes the coordinate of the integration point over a surface s. In

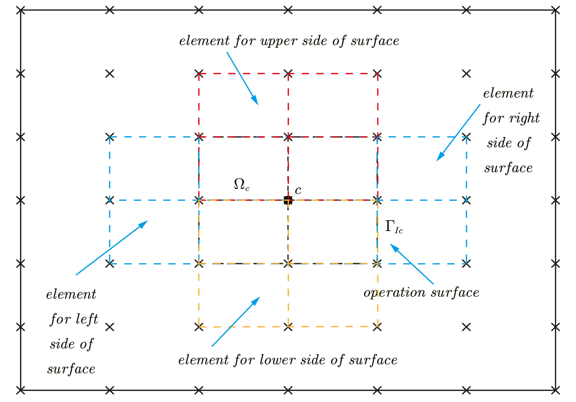

FIGURE 4 | Free elements around the operation surfaces.

ESFVM, the first spatial derivative term included in the right-hand side of Eq. **26** is evaluated by the free element of the element-shell including the operation surface under integration, and thus, using Eq. **17** it follows that

$$\int_{\Gamma_s} D_{ijkl} n_j \frac{\partial u_k}{\partial x_l}(x^s) d\Gamma(x^s) = \int_{\Gamma_s} D_{ijkl} n_j d_i^{s\alpha'}(x^s) d\Gamma(x^s) u_k^{\alpha'} \tag{27}$$

where $d_i^{s\alpha'}$ is computed using the free element formed for the surface s.

In ESFVM, since the operation surfaces of Ω_c are included in the additional formed elements, the accuracy of $\partial u_k / \partial x_l$ is higher than using the same element Ω_c built for point c.

Evaluation of the Domain Integral Involved in FVM

When the body force is considered in the computational problem, the FVM equations inevitably involve the domain integrals, as shown in Eqs **23–25**. Evaluation of the involved domain integrals is troublesome work. In conventional FVM, for achieving high efficiency, the domain integral is evaluated by assuming that the body force is constant throughout the control volume [41]. Thus, the domain integral can be simply written as

$$\int_{\Omega_c} b_i d\Omega = \bar{b}_i \, \Omega_c \tag{28}$$

where $\bar{b}_i$ is the average value of b_i in the control volume Ω_c.

Obviously, if b_i is sharply changeable in Ω_c, the above evaluation gives rise to a large error. To overcome this issue and for a universal scheme to accurately evaluate the domain integral, the Radial Integration Method (RIM) [42] can be employed to evaluate the domain integral in Eq. **28**, which can be expressed as

$$\int_{\Omega_c} b_i d\Omega = \int_{\partial\Omega_c} \frac{F_i}{r^n(c,\Gamma)} \frac{\partial r}{\partial n} d\Gamma \tag{29}$$

where $\partial\Omega_c$ is the boundary of the control volume Ω_c, $r(c,\Gamma)$ is the distance from the collocation point c to the boundary Γ, n is 1 or 2 for 2D or 3D problem, and

$$F_i = \int_0^{r(c,\Gamma)} b_i r^n dr \tag{30}$$

For most cases, b_i is a known function and Eq. **30** can be analytically integrated. For very complicated b_i, Eq. **30** can be evaluated using Gauss quadrature [42].

The above equations are suitable for any shaped control volume, regular or irregular, since the integration is over the boundary of Ω_c. In FEFVM and ESFVM, Ω_c is the free element domain formed for the collocation point c and $\partial\Omega_c$ is the boundary of the free element. Apart from the high accuracy, the main advantage of using RIM to evaluate the domain integral is that only the operation surfaces of the control volume are needed to evaluate the domain integral, with no need to perform volume integration over Ω_c.

Boundary Element Method (BEM)

In Eq. **7**, if the weight function w is taken as the displacement fundamental solution u_{mi}^*[9, 10], i.e., $w = u_{mi}^*$ the following BEM integral equation can be established:

$$\int_\Gamma \frac{\partial u_{mi}^*}{\partial x_l} D_{ijkl} n_j u_k d\Gamma - \int_\Omega \frac{\partial}{\partial x_j}\left(D_{ijkl}\frac{\partial u_{mi}^*}{\partial x_l}\right) u_k d\Omega = \int_\Gamma u_{mi}^* t_i d\Gamma + \int_\Omega u_{mi}^* b_i d\Omega \tag{31}$$

Recalling that the displacement fundamental solution u_{mi}^* satisfies the following equation:

$$\frac{\partial}{\partial x_j}\left(D_{ijkl}\frac{\partial u_{mi}^*}{\partial x_l}\right) + \delta_{mk}\delta(p,q) = 0 \tag{32}$$

Equation **31** becomes

$$u_m + \int_\Gamma t_{mk}^* u_k d\Gamma = \int_\Gamma u_{mi}^* t_i d\Gamma + \int_\Omega u_{mi}^* b_i d\Omega \tag{33}$$

where

$$t_{mk}^* = D_{ijkl} n_j \frac{\partial u_{mi}^*}{\partial x_l} \tag{34}$$

To evaluate the domain integral appearing in Eq. **33**, the conventional technique is to discretize the domain into internal cells [9]; however, this eliminates the advantage of BEM where only the boundary of the problem needs to be discretized into elements. To overcome this drawback, a transformation technique is usually employed to transform the domain integral into an equivalent integral. The most extensively used transformation technique is the Dual Reciprocity Method [DRM] [43]. Another technique used is the Radial Integration Method (RIM) [42], which can give more accurate results than DRM.

Using RIM, the domain integral in Eq. **33** can be expressed as

$$\int_\Omega u_{mi}^* b_i d\Omega = \int_\Gamma \frac{F_{mi}}{r^n(p,q)} \frac{\partial r}{\partial n} d\Gamma(q) \tag{35}$$

where the radial integral is

$$F_{mi} = \int_0^{r(p,q)} u_{mi}^* b_i r^n dr \tag{36}$$

where $n = 1$ for 2D problems and $n = 2$ for 3D problems. When b_i is a known function, Eqs **35**, **36** can give rise to a very accurate result. On the other hand, for the case where b_i is not very complicated, Eq. **36** can be analytically integrated [42]. However, when b_i is a complicated function, numerical integration should be performed [44].

LINE-OPERATION–BASED METHODS

Line-operation–based methods include the conventional finite difference method (FDM) [12] and the recently proposed finite line method (FLM) [3, 13]. In these methods, the computational domain is discretized into a series of points and lines formed by around points are then used to compute the spatial partial derivatives included in the PDEs, as shown in **Figure 5** for a 2D case. FDM constructs the first- and second-order partial derivatives using a line of points along the derivative directions. The main drawback of FDM is that if the lines that define the derivative directions are not orthogonal to one another in 2D or 3D problems, the accuracy of the cross-partial derivatives of different directions is usually very poor [45, 46]. This is why FDM cannot simulate irregular geometry problems well. In contrast, FLM has a much better performance in overcoming this drawback.

FLM uses a number of lines, named a line-set, to set up the solution scheme. Usually, at a collocation point, two lines (for 2D problems) or three lines (for 3D problems) are used to form the line-set, as shown in **Figure 5**. **Figure 6** shows the high-order line-sets of an internal collocation point for 2D and 3D problems.

Along a line of a line-set, the coordinates and physical variables can be expressed as

$$x_i = \sum_{\alpha=1}^{m} L^{\alpha}(l) x_i^{\alpha} \equiv L^{\alpha}(l) x_i^{\alpha} \tag{37}$$

$$u_k = \sum_{\alpha=1}^{m} L^{\alpha}(l) u_k^{\alpha} \equiv L^{\alpha}(l) u_k^{\alpha} \tag{38}$$

where m is the number of nodes defined along a line of the line-set, l is the arclength measured from node 1, and L^{α} is the Lagrange polynomial:

$$L^{\alpha}(l) = \prod_{\beta=1, \beta \neq \alpha}^{m} \frac{l - l^{\beta}}{l^{\alpha} - l^{\beta}}, \ (\alpha = 1 \sim m) \tag{39}$$

By differentiating Eqs **37**, **38** with respect to l, we can obtain expressions for computing the first- and second-order partial derivatives at the collocation point $\boldsymbol{x}^c$, as follows [3, 13]:

$$\frac{\partial u_k(\boldsymbol{x}^c)}{\partial x_i} = d_i^{c\alpha'} u_k^{\alpha'} \tag{40}$$

$$\frac{\partial^2 u_k(\boldsymbol{x}^c)}{\partial x_i \partial x_j} = d_{ij}^{c\alpha''} u_k^{\alpha''} \tag{41}$$

where

$$d_i^{c\alpha'} = \sum_{I=1}^{d} [J]_{iI}^{-1} \frac{\partial L_I^{\alpha}(l)}{\partial l}\bigg|_{l=l(\boldsymbol{x}^c)} \tag{42}$$

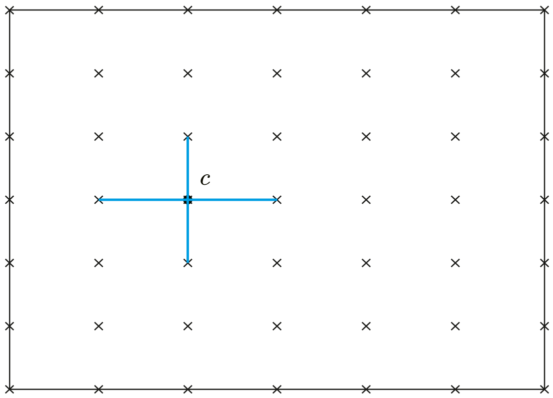

FIGURE 5 | Line-set consisting of two crossed lines for a 2D problem.

$$d_{ij}^{c\alpha''} = d_j^{c\beta'} d_i^{\beta'\alpha'} \tag{43}$$

where the repeated indexes represent summation, and $d = 2$ for 2D problems, $d = 3$ for 3D problems, and I represents the line number.

Using Eqs **40**, **41**, we can easily discretize a PDE and the related boundary conditions. For example, the PDE for the solid mechanics shown in Eqs **3**, **4** can directly generate a set of discretized equations which have the same forms as those shown in Eqs **20**, **21**.

POINT-OPERATION–BASED METHODS

Point-operation–based methods cover a number of numerical methods, such as the mesh free method (MFM) [5, 47–50], fundamental solution method [16], radial basis function method [51–53], and the newly developed free element method (FrEM) [15, 31]. In these types of methods, the computational domain is discretized into a series of points and solution schemes are established by collocating the governing PDEs or their integral forms at each collocation point. In MFM, the partial derivatives at the collocation point c are derived based on a group of scatter points within a specified support region around c, as shown in **Figure 7A**, while in FrEM, partial derivatives are derived based on an isoparametric element freely formed for point c, as shown in **Figure 7B**. In MFM and FrEM, both weak-form and strong-form solution schemes are available. In the following sections, the two schemes of FrEM will be described in detail.

Weak-Form Free Element Method (WFrEM)

In FrEM, a free element is independently formed for each collocation point c [31], with the domain denoted by Ω_c. The shape function shown in Eq. **8** is still employed for the formed free element. Let us apply the weighted residual formulation **(6)** to Ω_c and take the weight function as the shape function of the collocation point c, i.e., $w = N_c$. Thus, the following equation can be obtained:

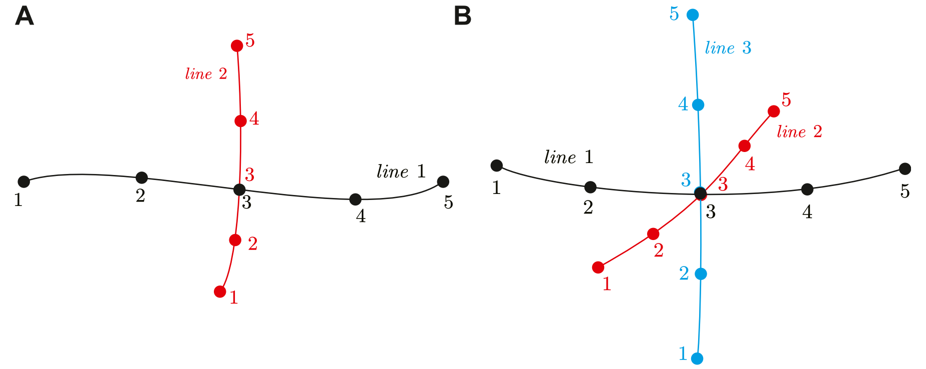

FIGURE 6 | Node distributions over line-sets defined at 2D or 3D internal points. **(A)** 2D problem **(B)** 3D problem.

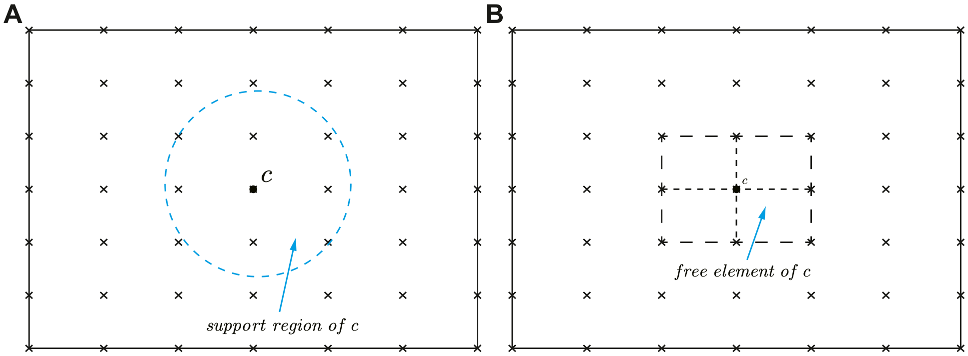

FIGURE 7 | 2D patterns for MFM **(A)** and FrEM **(B)** at a collocation point.

$$\int_{\Omega_c} \frac{\partial N_c}{\partial x_l} D_{ijkl} \frac{\partial u_k}{\partial x_j} d\Omega = \int_{\partial\Omega_c} N_c t_i(\boldsymbol{x}) d\Gamma + \int_{\Omega_c} N_c b_i d\Omega \tag{44}$$

where the derivatives of the shape functions are computed using Eq. **19a**.

Dividing the boundary $\partial\Omega_c$ of Ω_c into two parts; the inner boundary Γ_{Ic}, which is located within the problem and the outer boundary Γ_{oc}, which is located on the outer surface of the problem. Remembering that N_c is zero on the surfaces excluding point c, making the integral over Γ_{Ic} zero, as a result, Eq. **44** can be written as:

$$\int_{\Omega_c} \frac{\partial N_c}{\partial x_l} D_{ijkl} \frac{\partial N_\alpha}{\partial x_j} d\Omega \, u_k^\alpha = \int_{\Gamma_{oc}} N_c \bar{t}_i d\Gamma + \int_{\Omega_c} N_c b_i d\Omega \tag{45}$$

where $\Gamma_{oc} = \partial\Omega_c \cap \partial\Omega$, which is the outer boundary containing c.

From Eq. **45**, it can be seen that the form of the basic equation in WFrEM is similar to that in the conventional FEM. The essential difference between them is that the element in WFrEM is freely formed at each collocation point, the nodes of which are not restricted to any particular nodes of adjacent elements. It is also noteworthy that the free elements formed by around-collocation points are overlapped in FrEM since they are formed locally and independently at each point.

Strong-Form Free Element Method (SFrEM)

The SFrEM is a type of collocation method [15]. To achieve a highly accurate result, the collocation point c should be placed inside the formed free element. For this reason, the free elements used should have at least one internal node. In principle, any type of isoparametric elements with internal nodes can be utilized in SFrEM analysis [54–56]. For example, **Figure 8** shows new types of quadratic triangular and tetrahedral elements [55] and **Figure 9** shows a 21-noded block element [15].

The shape functions for the above triangular and tetrahedral elements can be found in [55] and that for the 21-noded quadratic block element in [15].

For higher-order elements, the best method is to use Lagrange elements. For example, **Figure 10** shows a 16-noded 2D third-order Lagrange element.

The shape functions of Lagrange elements for 2D and 3D problems can be constructed as follows [4]:

$$\begin{aligned} N_\alpha(\xi,\eta) &= L^I(\xi)L^J(\eta), \quad for\ 2D \\ N_\alpha(\xi,\eta,\zeta) &= L^I(\xi)L^J(\eta)L^K(\zeta), \quad for\ 3D \end{aligned} \tag{46}$$

where L^I, L^J, and L^K are determined by Eq. **39** and the superscript α is determined by the permutation of subscripts I, J, and K sequentially.

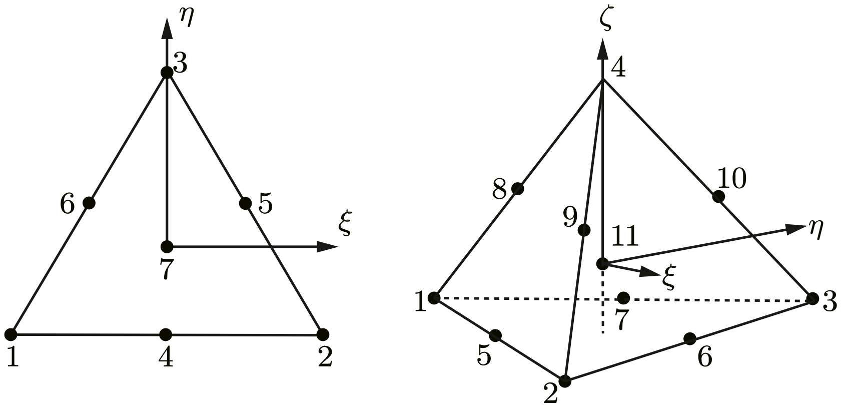

FIGURE 8 | Quadratic 7-noded triangular and 11-noded tetrahedral elements.

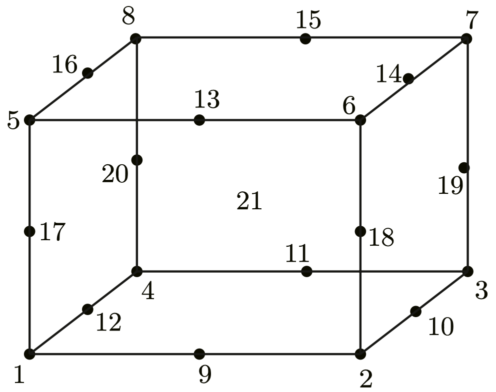

FIGURE 9 | 21-noded 3D quadratic block element.

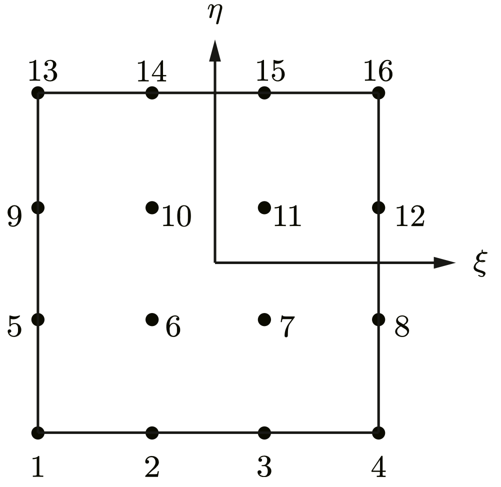

FIGURE 10 | 16-noded 2D third-order Lagrange element.

From the shape functions shown in Eq. **46**, the analytical expressions for computing the first- and second-order partial derivatives can be derived, which are the same as those shown in Eqs **17–19a**, **19b**. The collocation scheme to form the system of equations for the governing PDEs is the same as that in EDM, shown in Eqs **20**, **21**, the difference being that in SFrEM, only the interior and out boundary nodes are used.

NUMERICAL EXAMPLES

To demonstrate the performances of some of the numerical methods described in the article, a dual-struts supersonic combustor [57] is simulated in the following. The physical problem and relevant dimensions are shown in **Figure 11**. The Yong's modulus and Poison ratio of all materials in the problem are taken as $E = 206GPa$ and $\nu = 0.3$. To constrain the four frames supporting the combustor, the lower surface of the left frame is fixed and the lower surfaces of the other three frames are free only in the longitudinal direction (x-direction). All remaining outer surfaces of the combustor are imposed with a traction-free boundary condition. Pressure loads are applied on the inner surfaces of the combustor, which are distributed along the x-direction, as shown in **Figure 12**.

To simulate the problem using some of the described methods above, the whole structure is discretized into different numbers of nodes. **Figure 13** shows the computational mesh connected by all finite lines in the FLM model with 657,582 nodes. **Figure 14** shows a contour plot of the computed displacement amplitude over the deformed FLM mesh, with displacements enhanced ×20. For comparison, the problem is also computed using the FEM software ABAQUS, employing the same level nodes as those used in FLM. **Figure 15** shows the comparison of the computed displacement amplitude along the line MN using four different methods; ABAQUS (FEM), finite line method (FLM), weak-form free element method (WFrEM), and the strong-form method (SFrEM). For the last two methods, two meshes with different numbers of nodes are used, in which WFrEM-880k indicates the result of WFrEM using 880,000 nodes. To clearly examine the differences between the different methods, **Figure 16** shows the enhanced curves along two local parts of the line MN.

From **Figure 14**, it can be seen that the deformations of two areas after the struts are large. This is because the combustion

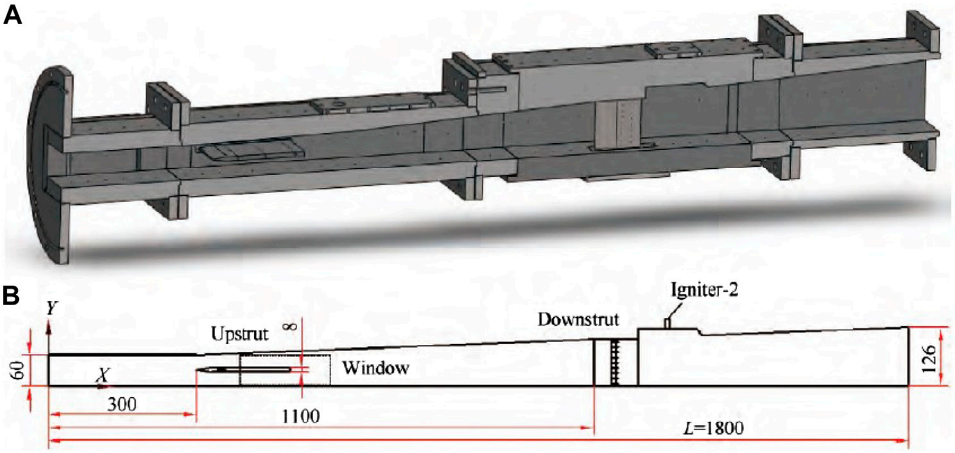

FIGURE 11 | Configurations of a combustion chamber with tow struts.

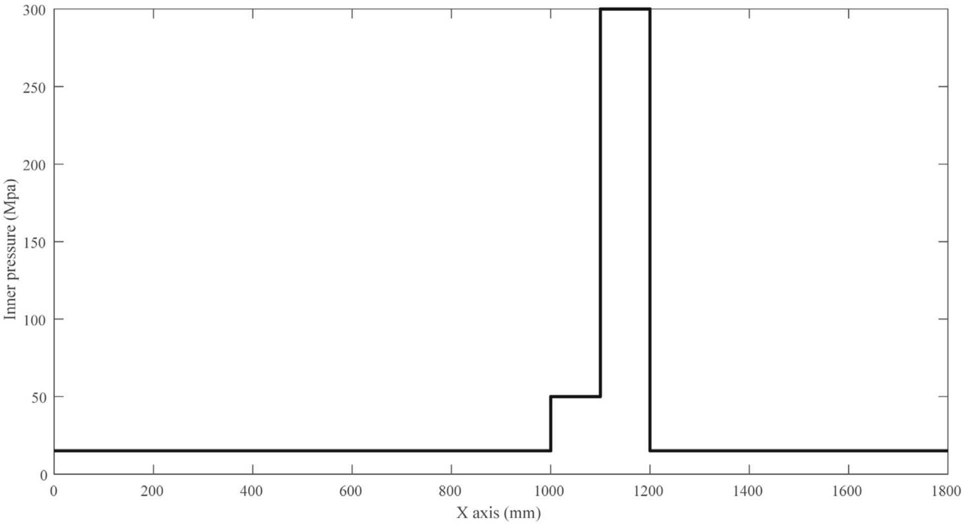

FIGURE 12 | Pressure applied on the inner surfaces of the combustor.

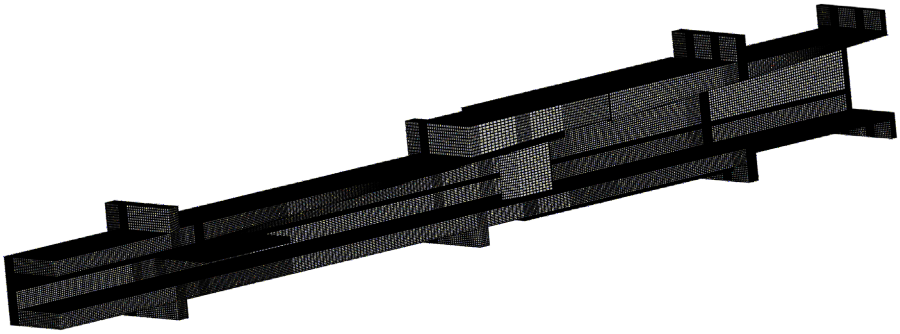

FIGURE 13 | Computational mesh used in FLM analysis.

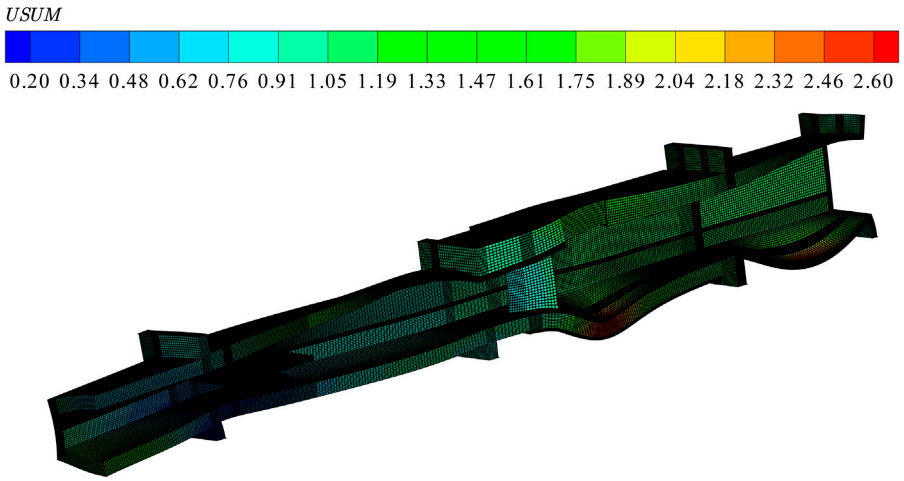

FIGURE 14 | Contours of the deformed mesh, with displacements multiplied ×20.

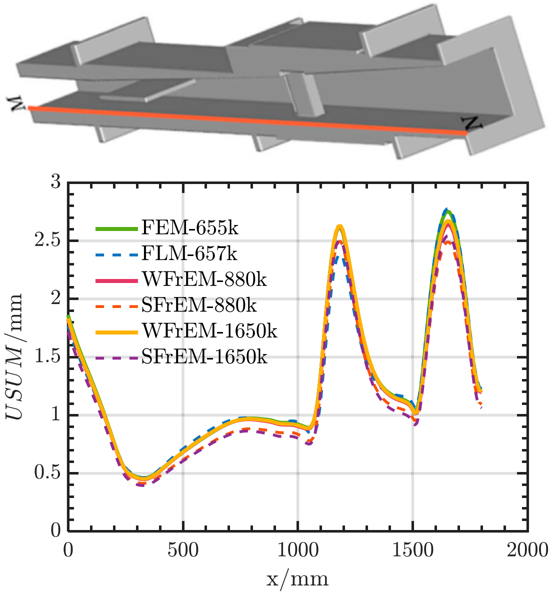

FIGURE 15 | Computed displacements along the line MN using the four different methods.

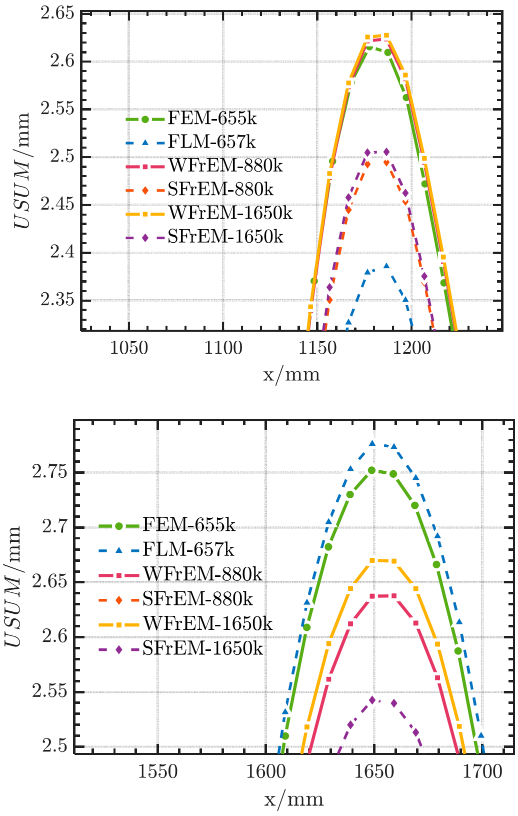

FIGURE 16 | Enhanced curves along two parts of line MN.

occurs immediately after the two struts; thus, the temperature and pressure are higher in these areas than other areas. Moreover, from **Figure 15**, it can be seen that all the computed results are in good agreement globally; this indicates that the methods presented in the article can handle real, complicated engineering problems. On the other hand, from **Figure 16**, it can be observed that the weak-form free element method (WFrEM) and the finite line method (FLM) give results closer to the finite element method (FEM). The essential reasons for this are that in WFrEM, numerical integration is performed over each free element, which is similar to FEM, giving very stable and accurate results. In FLM, the recursive technique is employed to evaluate the high-order derivatives, as shown in Eq. **43**, making more points contribute to each collocation point [3]; therefore, more stable and accurate results can be obtained using FLM over other strong-form solution schemes [50].

SUMMARY

In this article, four types of numerical methods are overviewed, most of which are newly proposed methods in recent years. Classification of all numerical methods into volume, surface, line, and point operation methods is performed for the first time in this article. This classification is conceptually clear and helpful for readers to understand the discretisation of problems and to realize the advantages and disadvantages of the different methods. Computational experience shows that the finite element, weak-form free element, and finite line methods have excellent performances.

AUTHOR CONTRIBUTIONS

X-WG: conceptualization; investigation; methodology; software; project administration; resources; writing- original draft; supervision. W-WJ: software; validation; writing—review and editing; X-BX, H-YL, KY, JL, and MC: writing—review and editing. All authors contributed to the article and approved the submitted version.

FUNDING

We gratefully acknowledge support of this investigation by the National Natural Science Foundation of China under Grant No. 12072064 and the Fundamental Research Funds for Central Universities under Grant No. DUT22YG204.

CONFLICT OF INTEREST

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

REFERENCES

1. Coleman C. J., On the use of radial basis functions in the solution of elliptic boundary value problems. *Comput Mech* (1996) 17:418–22. doi:10.1007/BF00363985
2. Wang Z. G., Liu L. S., Wu Y. H., The unique solution of boundary value problems for nonlinear second-order integral-differential equations of mixed type in Banach spaces. *Comput Math Appl* (2007) 54:1293–301. doi:10.1016/j.camwa.2007.04.018
3. Gao X. W., Ding J. X., Liu H. Y., Finite line method and its application in coupled heat transfer between fluid-solid domains. *Acta Phys Sin* (2022) 71(19):190201. doi:10.7498/aps.71.20220833
4. Gao XW, Li ZY, Yang K, Lv J, Peng HF, Cui M, et al. Element differential method and its application in thermal-mechanical problems. *Int J Numer Methods Eng* (2018) 113(1):82–108. doi:10.1002/nme.5604
5. Liu GR. An overview on meshfree methods: For computational solid mechanics. *Int J Comp Meth-sing* (2016) 13(5):1630001. doi:10.1142/S0219876216300014
6. Zienkiewicz OC, Taylor RL, Fox DD. *The finite element method for solid and structural mechanics*. Amsterdam: Elsevier (2014). p. 624.
7. Hughes TJR. *The finite element method: Linear static and dynamic finite element analysis*. Englewood Cliffs, NJ, USA: Prentice-Hall (1987). 704.
8. Wen PH, Cao P, Korakianitis T. Finite block method in elasticity. *Eng Anal Bound Elem* (2014) 46:116–25. doi:10.1016/j.enganabound.2014.05.006
9. Brebbia CA, Dominguez J. *Boundary elements: An introductory course*. London, UK: McGraw-Hill Book Co (1992). 860.
10. Gao XW, Davies TG. *Boundary element programming in mechanics*. Cambridge, UK: Cambridge University Press (2002). 253.
11. Onate E, Cervera M, Zienkiewicz OC. A finite volume format for structural mechanics. *Int J Numer Meth Eng* (1994) 37(2):181–201. doi:10.1002/nme.1620370202
12. Wang H, Dai W, Nassar R, Melnik R. A finite difference method for studying thermal deformation in a thin film exposed to ultrashort-pulsed lasers. *Int J Heat Mass Tran* (2006) 51:2712–23. doi:10.1016/j.ijheatmasstransfer.2006.01.013
13. Gao XW, Zhu YM, Pan T. Finite line method for solving high-order partial differential equations in science and engineering. *Part Diff Eq Appl Math* (2023) 7:100477. doi:10.1016/j.padiff.2022.100477
14. Atluri SN, Shen SP. *The meshless local Petrov-Galerkin (MLPG) method*. Henderson, NV, USA: Tech Sci (2002). 51.
15. Gao XW, Gao LF, Zhang Y, Cui M, Lv J. Free element collocation method: A new method combining advantages of finite element and mesh free methods. *Comput Struct* (2019) 215:10–26. doi:10.1016/j.compstruc.2019.02.002
16. Fan CM, Huang YK, Chen CS, Kuo SR. Localized method of fundamental solutions for solving two-dimensional Laplace and biharmonic equations. *Eng Anal Bound Elem* (2019) 101:188–97. doi:10.1016/j.enganabound.2018.11.008
17. Belytschko T, Liu WK, Moran B, Elkhodary K. *Nonlinear finite elements for continua and structures*. New York, USA: John Wiley and Sons (2000). 650.
18. Baliga BR. *A control-volume based finite element method for convective heat and mass transfer*. [PhD's thesis]. Minnesota: University of Minnesota (1978).
19. Schneider GE, Raw MJ. Control volume finite-element method for heat transfer and fluid-flow using colocated Variables.1.Computational procedure. *Numer Heat Tr* (1987) 11(4):363–90. doi:10.1080/10407798708552552
20. Hughes TJR, Cottrell JA, Bazilevs Y. Isogeometric analysis: CAD, finite elements, NURBS,exact geometry and mesh refinement. *Comput Methods Appl Mech Eng* (2005) 194:4135–95. doi:10.1016/j.cma.2004.10.008
21. Hou WB, Jiang K, Zhu XF, Shen Y., Li Y, Zhang X, et al. Extended Isogeometric Analysis with strong imposing essential boundary conditions for weak discontinuous problems using B++ splines. *Comput Method Appl Mech Eng* (2020) 370. doi:10.1016/j.cma.2004.10.008
22. Liu GR, Nguyen TT. *Smoothed finite element methods*. Boca Raton, FL, USA: Chemical Rubber Co (2010). 692.
23. Chen L, Liu GR, Zeng KY. A combined extended and edge-based smoothed finite element method (es-xfem) for fracture analysis of 2d elasticity. *Int J Comput Methods* (2011) 8(4):773–86. doi:10.1142/S0219876211002812
24. Li M, Wen PH. Finite block method for transient heat conduction analysis in functionally graded media. *Int J Numer Meth Eng* (2014) 99(5):372–90. doi:10.1002/nme.4693
25. Fantuzzi N, Tornabene F, Viola E, Ferreira AJM. A strong formulation finite element method (SFEM) based on RBF and GDQ techniques for the static and dynamic analyses of laminated plates of arbitrary shape. *Meccanica* (2014) 49:2503–42. doi:10.1007/s11012-014-0014-y
26. Fantuzzi N, Dimitri R, Tornabene F. A SFEM-based evaluation of mode-I Stress Intensity Factor in composite structures. *Compos Struct* (2016) 145:162–85. doi:10.1016/j.compstruct.2016.02.076
27. Gao XW, Huang SZ, Cui M, Ruan B, Zhu QH, Yang K, et al. Element differential method for solving general heat conduction problems. *Int J Heat Mass Tran* (2017) 115:882–94. doi:10.1016/j.ijheatmasstransfer.2017.08.039
28. Lv J, Song C, Gao XW. Element differential method for free and forced vibration analysis for solids. *Int J Mech Sci* (2019) 151:828–41. doi:10.1016/j.ijmecsci.2018.12.032
29. Gao LF, Gao XW, Feng WZ, Xu BB. A time domain element differential method for solving electromagnetic wave scattering and radiation problems.

Eng Anal Bound Elem (2022) 140:338–47. doi:10.1016/j.enganabound.2022.04.025

30. Zheng YT, Gao XW, Liu YJ. Numerical modelling of braided ceramic fiber seals by using element differential method. *Compos Struct* (2023) 304: 116461–1. doi:10.1016/j.compstruct.2022.116461
31. Jiang WW, Gao XW. Analysis of thermo-electro-mechanical dynamic behavior of piezoelectric structures based on zonal Galerkin free element method. *Eur J Mech A-solid* (2023) 99:104939. doi:10.1016/j.euromechsol.2023.104939
32. Moukalled F, Mangani L, Darwish M. *The finite volume method in computational fluid dynamics: An advanced introduction with OpenFOAM® and matlab*. Berlin,GER: Springer (2016). 791.
33. Ivankovic A, Demirdzic I, Williams JG, Leevers PS. Application of the finite volume method to the analysis of dynamic fracture problems. *Int J Fracture* (1994) 66(4):357–71. doi:10.1007/BF00018439
34. Cardiff P, Kara A, Ivankovic A. A large strain finite volume method for orthotropic bodies with general material orientations. *Comput Methods Appl Mech Eng* (2014) 268:318–35. doi:10.1016/j.cma.2013.09.008
35. Fallah N, Parayandeh-Shahrestany A. A novel finite volume based formulation for the elasto-plastic analysis of plates. *Thin Wall Struct* (2014) 77:153–64. doi:10.1016/j.tws.2013.09.025
36. Gong J, Xuan L, Ming P, Zhang W. An unstructured finite-volume method for transient heat conduction analysis of multilayer functionally graded materials with mixed grids. *Numer Heat Tr B* (2013) 63(3):222–47. doi:10.1080/10407790.2013.751251
37. Cavalcante MA, Marques SP, Pindera MJ. Computational aspects of the parametric finite-volume theory for functionally graded materials. *Comp Mater Sci* (2008) 44(2):422–38. doi:10.1016/j.commatsci.2008.04.006
38. Jasak H, Weller H. Application of the finite volume method and unstructured meshes to linear elasticity. *Int J Numer Methods Eng* (2000) 48(2):267–87. doi:10.1002/(SICI)1097-0207(20000520)48:2<267::AID-NME884>3.0.CO;2-Q
39. Bailey C, Cross M. A finite volume procedure to solve elastic solid mechanics problems in three dimensions on an unstructured mesh. *Int J Numer Methods Eng* (1995) 38:1757–76. doi:10.1002/nme.1620381010
40. Charoensuk J, Vessakosol P. A high order control volume finite element procedure for transient heat conduction analysis of functionally graded materials. *Heat Mass Tr* (2010) 46:1261–76. doi:10.1007/s00231-010-0649-8
41. Tao WQ. *Numerical heat transfer*. 2. Xian: Xian Jiaotong University Press (2001).
42. Gao XW. The radial integration method for evaluation of domain integrals with boundary-only discretization. *Eng Anal Bound Elem* (2002) 26:905–16. doi:10.1016/S0955-7997(02)00039-5
43. Nardini D, Brebbia CA. A new approach for free vibration analysis using boundary elements. In: Brebbia CA, editor. *Boundary element methods in engineering*. Berlin,GER: Springer (1982). 312–26.
44. Gao XW. A boundary element method without internal cells for two-dimensional and three-dimensional elastoplastic problems. *J Appl Mech* (2002) 69:154–60. doi:10.1115/1.1433478
45. Liszka T, Orkisz J. The finite difference method at arbitrary irregular grids and its application in applied mechanics. *Comput Struct* (1980) 11(1-2):83–95. doi:10.1016/0045-7949(80)90149-2
46. Thomas JW. *Numerical partial differential equations: Finite difference methods*. Berlin, GER: Springer Science and Business Media (2013). 437.
47. Liu WK, Belytschko T, Chang H. An arbitrary Lagrangian-Eulerian finite element method for path-dependent materials. *Comput Methods Appl Mech Eng* (1986) 58:227–45. doi:10.1016/0045-7825(86)90097-6
48. Li S, Liu WK. Meshfree and particle methods and their applications. *Appl Mech Rev* (2002) 55:1–34. doi:10.1115/1.1431547
49. Li W, Nguyen-Thanh N, Huang J, Zhou K. Adaptive analysis of crack propagation in thin-shell structures via an isogeometric-meshfree moving least-squares approach. *Comput Methods Appl Mech Eng* (2020) 358: 112613–3. doi:10.1016/j.cma.2019.112613
50. Wang DD, Wang JR, Wu JC. Superconvergent gradient smoothing meshfree collocation method. *Comput Methods Appl Mech Eng* (2018) 340:728–66. doi:10.1016/j.cma.2018.06.021
51. Wang JG, Liu GR. A point interpolation meshless method based on radial basis functions. *Int J Numer Methods Eng* (2002) 54:1623–48. doi:10.1002/nme.489
52. Hart EE, Cox SJ, Djidjeli K, Compact RBF. Compact RBF meshless methods for photonic crystal modelling. *J Comput Phys* (2011) 230(12):4910–21. doi:10.1016/j.jcp.2011.03.010
53. Zheng H, Yang Z, Zhang CH, Tyrer M. A local radial basis function collocation method for band structure computation of phononic crystals with scatterers of arbitrary geometry. *Appl Math Model* (2018) 60:447–59. doi:10.1016/j.apm.2018.03.023
54. Gao XW, Liu XY, Xu BB, Cui M& Lv J. Element differential method with the simplest quadrilateral and hexahedron quadratic elements for solving heat conduction problems. *Numer Heat Transfer, B* (2018) 73(4):206–24. doi:10.1080/10407790.2018.1461491
55. Gao XW, Liu HY, Lv J, Cui M. A novel element differential method for solid mechanical problems using isoparametric triangular and tetrahedral elements. *Comput Maths Appl* (2019) 78(11):3563–85. doi:10.1016/j.camwa.2019.05.026
56. Lv J, Zheng MH, Xu BB, Zheng YT, Gao XW. Fracture mechanics analysis of functionally graded materials using a mixed collocation element differential method. *Eng Fracture Mech* (2021) 244:107510. doi:10.1016/j.engfracmech.2020.107510
57. Ji PF, Zhu SH, Chen B, Xu X. Experimental investigation of dual-struts supersonic combustor with varying equivalence ratio injections. *J Propuls Tech* (2017) 38(9):2011–9. doi:10.13675/j.cnki.tjjs.2017.09.012

AEROSPACE RESEARCH COMMUNICATIONS
ORIGINAL RESEARCH
published: 20 November 2023
doi: 10.3389/arc.2023.12099

The Thermal Conductivity Coefficient of a Square Thermal Invisibility Cloak Cell and Its Application in Periodic Plate

*Yanyan Sun, Yijun Chai, Xiongwei Yang and Yueming Li**

State Key Laboratory for Strength and Vibration of Mechanical Structures, Shaanxi Key Laboratory of Environment and Control for Flight Vehicle, School of Aerospace, Xi'an Jiaotong University, Xi'an, China

In this paper, the thermal conductivities of the square thermal invisibility cloak are constructed in two ways. One is the direct method, another is the rotation matrix method. The thermal conductivity coefficients obtained by the two methods are the same. The cloud map of thermal conductivity coefficient of the thermal cloak is drawn, which can help us understand more intuitively how the thermal invisibility cloak works. Besides, to manipulate heat flow in a larger area, the cloaks are arranged periodically by introducing the position parameters into the calculation of the thermal conductivity coefficient of the thermal invisibility cloak. The heat insulation function of both the single thermal cloak and the thermal cloak periodic plate are tested under different heat boundary conditions using COMSOL Multiphysics. For different heat boundary conditions, heat flux direction of the simulation result is given. The results show that both the single thermal cloak and the thermal cloak periodic plate have the function of avoiding heat flow under different heat boundary conditions. The heat fluxes travel around the inner domain with good thermal stealth effect.

Keywords: thermal invisibility, periodic plate, thermal metamaterials, cloak, coordinate transformation

INTRODUCTION

Metamaterials can be used to reveal novel physics and develop new applications because of their novel artificial structures. The transformation principle can be used to control processes with governing equations that are form-invariant under coordinate transformations, enabling the control of diffusion processes by changing material parameters such as conductivity. Thermal metamaterials, together with the governing theories, make it possible to actively manipulate macroscopic heat phenomena of artificial systems, which enables people to change the heat phenomena. Such metamaterials or metamaterial-based devices refer to those artificial structures that yield novel functions in controlling heat transfer [1, 2].

The thermal invisibility cloak has a unique thermal insulation function, and its temperature gradient in the thermal stealth area is zero, which can avoid the heat flow very well. Therefore, thermal invisibility cloaks have a great application value in computer chips, spacecraft return modules, satellites, and other equipment.

In 2006, Pendry et al. [3], based on the form invariance of Maxwell's equation, designed an artificial metamaterial with anisotropic characteristic parameters to make electromagnetic waves propagate along a preset path, to achieve the purpose of electromagnetic invisibility. In 2008, Fan et al. [4] first introduced the concept of an electromagnetic invisibility cloak into the field of thermal

***Correspondence:**
Yueming Li
liyueming@mail.xjtu.edu.cn

Received: *23 September 2023*
Accepted: *31 October 2023*
Published: *20 November 2023*

Citation:
Sun Y, Chai Y, Yang X and Li Y (2023) The Thermal Conductivity Coefficient of a Square Thermal Invisibility Cloak Cell and Its Application in Periodic Plate.
Aerosp. Res. Commun. 1:12099.
doi: 10.3389/arc.2023.12099

science, and theoretically predicted the thermal invisibility cloak. In 2010, Li et al. [5] proposed the principle of using coordinate transformation to design multi-functional stealth cloak. The multi-functional invisibility cloak can not only realize the function of a thermal invisibility cloak, but also realize the function of an electric invisibility cloak. In 2012, Narayana et al. [6] used COMSOL software to simulate the heat flow distribution of two different materials coupled together. They built a prototype thermal cloak in the lab by combining two materials with vastly different thermal conductivities into a spiral-shaped multilayer structure. By measuring its temperature distribution in the temperature field, it is verified that its experimental performance is in complete agreement with the theoretical expectation, thus opening the door for experimental exploration of a thermal invisibility cloak. In 2013, Schittny et al. [7] constructed a new type of heat invisibility cloak using a single copper material. By machining thermal channels of different sizes on the copper sheet (each channel leads to different thermal conductivity due to different sizes), the heat flow in r<R2 is compressed to R1<r<R2. In 2014, Mao et al. [8] designed the expression of the thermal conductivity of the cylindrical heat cloak with non-conformal arbitrary cross-section based on the transformation of thermodynamics and verified that the heat cloak has the function of thermal protection for the stealth area through COMSOL Multiphysics simulation. In 2015, Li et al. [9] introduced the dependence of thermal conductivity on temperature when studying thermal superstructural materials, thus developing the transformation thermal theory (thermal conductivity is independent of temperature) and establishing the nonlinear transformation thermal theory, which then promoted the nonlinear thermal research based on thermal superstructural materials. The so-called nonlinear heat theory focuses on the thermal conductivity of materials, which is no longer a constant independent of temperature, but a physical quantity with specific response to temperature. Based on this response, they designed switching thermal cloaks that can respond intelligently to ambient temperatures and prepared macroscopic thermal diodes. In 2017, Xia et al. [10] deduced the general solution expression of the thermal conductivity of the three-dimensional heat cloak with arbitrary shape and carried out full wave simulation verification. In 2018, Xia et al. [11] deduced the expression of the thermal conductivity of the square heat cloak and designed the square heat cloak with any angle. In 2021, Sha et al. [12] address three long-standing challenges, i.e., transformation optics-induced anisotropic material parameters, the limited shape adaptability of experimental thermal meta devices, and *a priori* knowledge of back-ground temperatures and thermal functionalities. They took the local thermal conductivity tensors as input, resorted to topology optimization for the freeform designs of topological functional cell (TFCs), and then directly assembled and printed them. They designed and 3D-printed three freeform thermal meta devices (concentrator, rotator, and cloak).

The structural vibro-acoustic problem in a thermal environment [13, 14] is one of the most common problems in aerospace engineering. This paper studies the problem of heat avoidance in thermal environment, aiming at laying a foundation for solving the problem of thermal acoustic vibration. Previous studies have focused on a single cloak and did not propose how to arrange the cloaks periodically to study the manipulation of the heat flow by the periodic plate of the heat cloak. In this paper, two specific ways are given to obtain the transformation equation of the square heat cloak. In addition, the " position parameters " are used in the thermal conductivity of the cloak, to obtain the thermal cloak at any position in practical applications. Based on this, the thermal cloaks are arranged periodically. The regulating effect of the periodic plate of the thermal cloaks on the heat flow is simulated under different boundary conditions by using COMSOL software.

THEORETICAL BASIS

In nature, the heat flux always flows from hot areas to cold areas, which is decided by the heat conduction equation:

$$\rho \cdot c \frac{\partial T}{\partial t} = \nabla \cdot (\lambda \nabla T) + Q \tag{1}$$

where ρ, c, λ represent the density, thermal capacity, and the thermal conductivity of the medium respectively. T is the temperature. Q is the heat resource. If there is no heat resource and at steady state, the Eq. **1** is simplified as:

$$\nabla \cdot (\lambda \nabla T) = 0 \tag{2}$$

Since the Eq. **2** has formal invariance [15], it could be transformed in another space as follows:

$$\nabla' \cdot (\lambda' \nabla' T') = 0 \tag{3}$$

where λ', T' represent the thermal conductivity coefficients and temperatures in transformed spaces. According to the theory of transformation thermodynamics, the thermal conductivity relationship between the transformation space and the original space [16]:

$$\lambda' = \frac{A \cdot \lambda_0 \cdot A^T}{\det(A)} \tag{4}$$

where λ_0 represents the thermal conductivity coefficients in the original space. A is the Jacobian transformation matrix, reflecting the geometric changes from the original space to the transformation space, and its components are:

$$A_{ij} = \frac{\partial x_i'}{\partial x_j} \tag{5}$$

where x_i' denotes the three coordinate components *x', y', z'* in the new coordinate system. x_j represents the three coordinate components *x, y, z* in the original coordinate system.

The square thermal cloak is designed as the **Figure 1** shown. A square area with a side length of $2s_2$ in **Figure 1A** is compressed into a square torus in **Figure 1B** which is encircled by a square with side length of $2s_1$ and a square with side length of $2s_2$. Considering that the region has geometric symmetry, the region

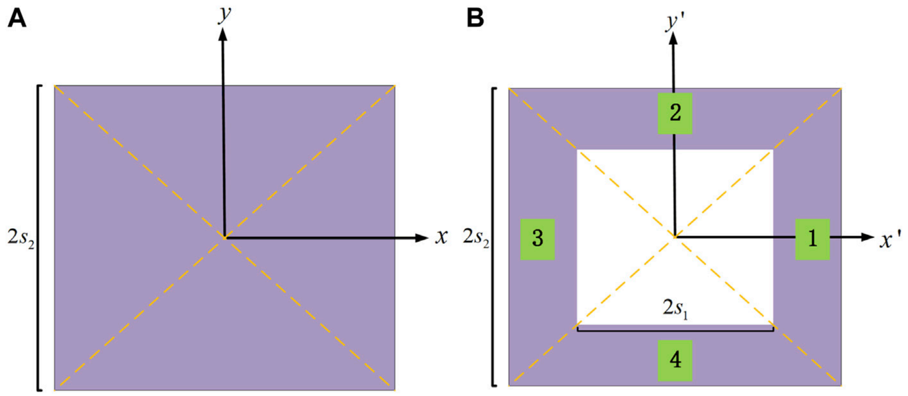

FIGURE 1 | The sketch map of the square thermal cloak (areas 1,2,3,4). **(A)** Schematic diagram of the area before transformation. **(B)** Schematic diagram of the transformed thermal invisibility cloak.

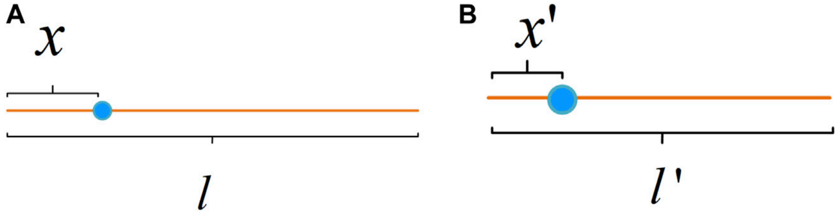

FIGURE 2 | Schematic diagram of compressed line segment. **(A)** Original line segment. **(B)** Compressed segment.

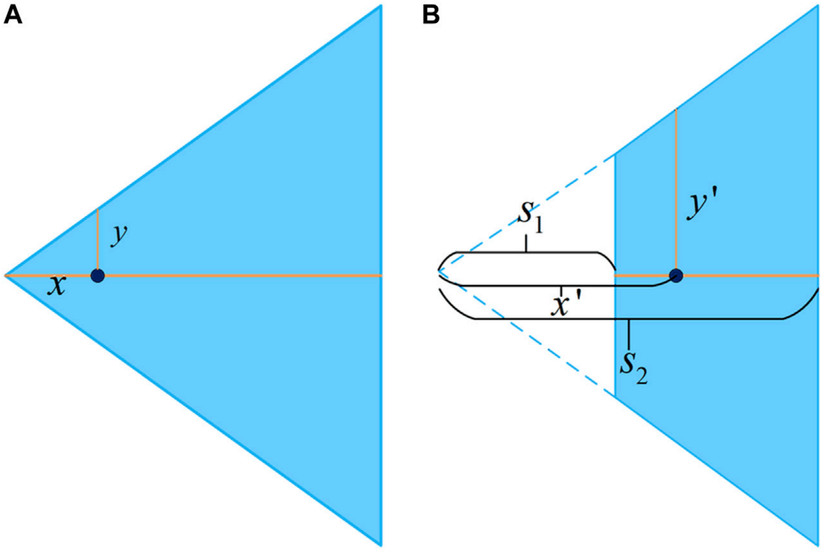

FIGURE 3 | The schematic diagram of the transformation of the area 1. **(A)** Area 1 before transformation. **(B)** Transformed area 1.

area could be divided into four triangular areas along its diagonal. The thermal conductivity is calculated according to the transformation equation of four triangular regions respectively [17].

Two approaches will be given to find the thermal conductivities of the square thermal cloak as follows. The first one is a direct method. To illustrate how the transformation equation of each region is obtained, consider the following case

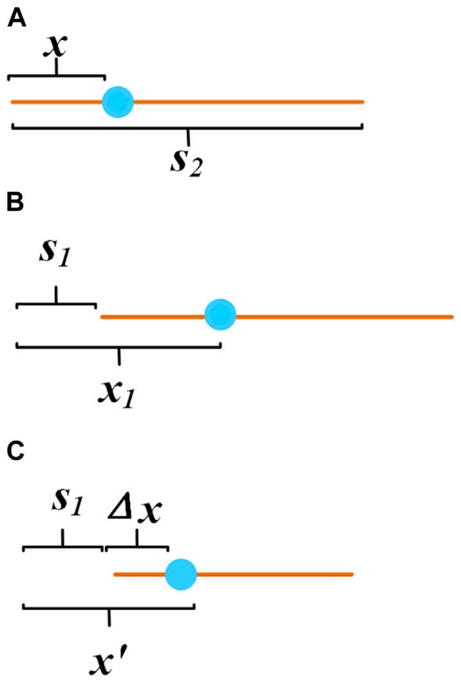

FIGURE 4 | Shift the whole segment to the right by s_1. **(A)** The segment before the move. **(B)** The line segment after the move. **(C)** Compress the line segment into $(s_2 - s_1)$.

Figure 2A: take a point x on a line segment, the length of which is l. When the segment is compressed to a length of l', as **Figure 2B** shows, the corresponding x becomes x'. It is not difficult to get the following formula:

$$\frac{x}{l} = \frac{x'}{l'} \tag{6}$$

Figures 3A, B are the schematic diagram of the area 1 before and after the transformation respectively. Take the left triangle area's median line, as **Figure 4A** shows, for example. For the first step shifting the whole segment to the right by s_1 (**Figure 4B**):

$$x_1 = x + s_1 \tag{7}$$

Second, compress the line segment into $(s_2 - s_1)$ (**Figure 4C**):

$$x' = s_1 + \Delta x \tag{8}$$

According to **formula 6**, it is obtained:

$$\frac{\Delta x}{s_2 - s_1} = \frac{x}{s_2} \tag{9}$$

Simultaneous Eqs **7–9**, it could be obtained:

$$x' = \frac{s_2 - s_1}{s_2} x + s_1 \tag{10}$$

It can be known from similar triangles:

$$\frac{x}{x'} = \frac{y}{y'}, y' = \left(\frac{s_2 - s_1}{s_2} + \frac{s_1}{x}\right) y \tag{11}$$

From the above example, it is easy to find that the transformation formula of area 1 is as follows:

$$x' = x\frac{s_2 - s_1}{s_2} + s_1, y' = \left(\frac{s_2 - s_1}{s_2} + \frac{s_1}{x}\right) y \tag{12}$$

In area 1, the Jacobian matrix A_1 and the thermal conductivity λ_1' of the transformation space are:

$$A_1 = \begin{bmatrix} \frac{s_2 - s_1}{s_2} & 0 \\ \frac{-s_1 y}{x^2} & \frac{s_2 - s_1}{s_2} + \frac{s_1}{x} \end{bmatrix}, \lambda_1{}' = \begin{bmatrix} 1 - \frac{s_1}{x'} & -\frac{s_1 y'}{x'^2} \\ -\frac{s_1 y'}{x'^2} & \frac{x'^4 + s_1{}^2 y'^2}{x'^4 - s_1 x'^3} \end{bmatrix} \lambda_0 \tag{13}$$

In a similar way, the transformation formula and the Jacobian matrix of areas 2, 3, and 4 could be found as the **formulae 14**, **15**, and **16**:

$$x' = x\left(\frac{s_2 - s_1}{s_2} + \frac{s_1}{y}\right), y' = y\frac{s_2 - s_1}{s_2} + s_1 \tag{14}$$

$$x' = x\frac{s_2 - s_1}{s_2} - s_1, y' = y\left(\frac{s_2 - s_1}{s_2} - \frac{s_1}{x}\right) \tag{15}$$

$$x' = x\left(\frac{s_2 - s_1}{s_2} - \frac{s_1}{y}\right), y' = y\frac{s_2 - s_1}{s_2} - s_1 \tag{16}$$

The corresponding Jacobian matrixes of areas 2, 3, and 4 are respectively as the **formulae 17**, **18**, and **19**:

$$A_2 = \begin{bmatrix} \frac{s_2 - s_1}{s_2} + \frac{s_1}{y} & -\frac{s_1 x}{y^2} \\ 0 & \frac{s_2 - s_1}{s_2} \end{bmatrix} \tag{17}$$

$$A_3 = \begin{bmatrix} \frac{s_2 - s_1}{s_2} & 0 \\ \frac{s_1 y}{x^2} & \frac{s_2 - s_1}{s_2} - \frac{s_1}{x} \end{bmatrix} \tag{18}$$

$$A_4 = \begin{bmatrix} \frac{s_2 - s_1}{s_2} - \frac{s_1}{y} & \frac{s_1 x}{y^2} \\ 0 & \frac{s_2 - s_1}{s_2} \end{bmatrix} \tag{19}$$

Finally, the thermal conductivities $\lambda_2', \lambda_3', \lambda_4'$ of each transformation area 2,3,4 are respectively as follows:

$$\lambda_2{}' = \begin{bmatrix} \frac{y'^4 + s_1{}^2 x'^2}{y'^4 - s_1 y'^3} & -\frac{s_1 x'}{y'^2} \\ -\frac{s_1 x'}{y'^2} & 1 - \frac{s_1}{y'} \end{bmatrix} \lambda_0 \tag{20}$$

$$\lambda_3{}' = \begin{bmatrix} 1 + \frac{s_1}{x'} & \frac{s_1 y'}{x'^2} \\ \frac{s_1 y'}{x'^2} & \frac{x'^4 + s_1{}^2 y'^2}{x'^4 + s_1 x'^3} \end{bmatrix} \lambda_0 \tag{21}$$

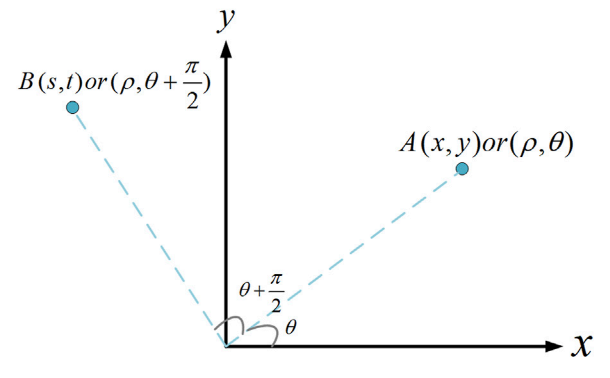

FIGURE 5 | Rotating schematic.

$$\lambda_4{}' = \begin{bmatrix} \dfrac{y'^4 + s_1{}^2 x'^2}{y'^4 + s_1 y'^3} & \dfrac{s_1 x'}{y'^2} \\ \dfrac{s_1 x'}{y'^2} & 1 + \dfrac{s_1}{y'} \end{bmatrix} \lambda_0 \tag{22}$$

The above is the direct way to get the thermal conductivity of the square thermal cloak.

The second way is using a rotation matrix to figure out the thermal conductivity of the four areas. The areas 2, 3, and 4 can be obtained by rotating the area1 counterclockwise by 90°, 180°, 270°. The rotation process can be implemented by the rotation matrix.

To begin with, let us introduce the rotation matrix P by the flowing example. Consider the situation shown in **Figure 5**: rotate point A 90° counterclockwise to get point B.

The coordinates of point A are:

$$x = \rho \cdot \cos\theta, y = \rho \cdot \sin\theta \tag{23}$$

where x and y indicate the horizontal and vertical coordinates of point A in the Cartesian coordinate system. ρ, θ are the polar diameter and polar angle of point A in the polar coordinate system.

The coordinates of point B are:

$$s = \rho \cdot \cos\left(\theta + \frac{\pi}{2}\right) = -\rho \cdot \sin\theta = -y, t = \rho \cdot \sin\left(\theta + \frac{\pi}{2}\right) = \rho \cdot \cos\theta = x \tag{24}$$

where s and t indicate the horizontal and vertical coordinates of point B in the Cartesian coordinate system.

The matrix form is used to clarify the relationship between A and B:

$$\begin{bmatrix} s \\ t \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \tag{25}$$

$$P = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \tag{26}$$

$$\begin{bmatrix} s \\ t \end{bmatrix} = P \begin{bmatrix} x \\ y \end{bmatrix} \tag{27}$$

P is the rotation matrix, which mean a transformation of rotating counterclockwise by 90°. For the area 1, the transformation formula is:

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \dfrac{s_2 - s_1}{s_2} & 0 \\ 0 & \dfrac{s_2 - s_1}{s_2} + \dfrac{s_1}{x} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} s_1 \\ 0 \end{bmatrix} \tag{28}$$

The transformation equation of area 2 can be obtained by rotating the transformation equation of area 1 90° counterclockwise. Multiply both sides of **formula 28** by the rotation matrix P:

$$P \begin{bmatrix} x' \\ y' \end{bmatrix} = P \begin{bmatrix} \dfrac{s_2 - s_1}{s_2} & 0 \\ 0 & \dfrac{s_2 - s_1}{s_2} + \dfrac{s_1}{x} \end{bmatrix} P^{-1} P \begin{bmatrix} x \\ y \end{bmatrix} + P \begin{bmatrix} s_1 \\ 0 \end{bmatrix} \tag{29}$$

B and B' are obtained by rotating A and A', so we can get:

$$\begin{bmatrix} s' \\ t' \end{bmatrix} = P \begin{bmatrix} x' \\ y' \end{bmatrix}, \begin{bmatrix} s \\ t \end{bmatrix} = P \begin{bmatrix} x \\ y \end{bmatrix} \tag{30}$$

Take the **formula 30** into the **formula 29**:

$$\begin{aligned} \begin{bmatrix} s' \\ t' \end{bmatrix} &= P \begin{bmatrix} \dfrac{s_2 - s_1}{s_2} & 0 \\ 0 & \dfrac{s_2 - s_1}{s_2} + \dfrac{s_1}{t} \end{bmatrix} P^{-1} \begin{bmatrix} s \\ t \end{bmatrix} + \begin{bmatrix} 0 \\ s_1 \end{bmatrix} \\ &= \begin{bmatrix} \left(\dfrac{s_2 - s_1}{s_2} + \dfrac{s_1}{t} \right) s \\ \dfrac{s_2 - s_1}{s_2} t + s_1 \end{bmatrix} \end{aligned} \tag{31}$$

Let s', t', s, t in the **formula 31** be represented by x', y', x, y respectively, the transformation equation of area 2 is obtained as follows:

$$x' = x\left(\frac{s_2 - s_1}{s_2} + \frac{s_1}{y} \right), y' = y \frac{s_2 - s_1}{s_2} + s_1 \tag{32}$$

In a similar way, by multiplying the rotation matrix P, the transformation equations of areas 3 and 4 can be found. Then, the thermal conductivity of each transformation space is calculated by **formula 4**. The results of the direct method and the rotation matrix method are the same. The two methods' principles are geometric transformation and matrix rotation respectively. The physical natures are the same, only the mathematical means are different. The cloud map of the thermal conductivity coefficient of the thermal cloak is drawn in **Figure 6**. **Figure 6A** is $\lambda_{xx}{}'$ and **Figure 6B** is $\lambda_{xy}{}'$.

The above thermal cloak is at the origin of the frame. Furthermore, the thermal conductivity of a thermal cloak at any position is calculated by introducing position parameters a, b. Move the coordinate system:

$$m = x + a, n = y + b, m' = x' + a, n' = y' + b \tag{33}$$

The transformation equations of area 1 are:

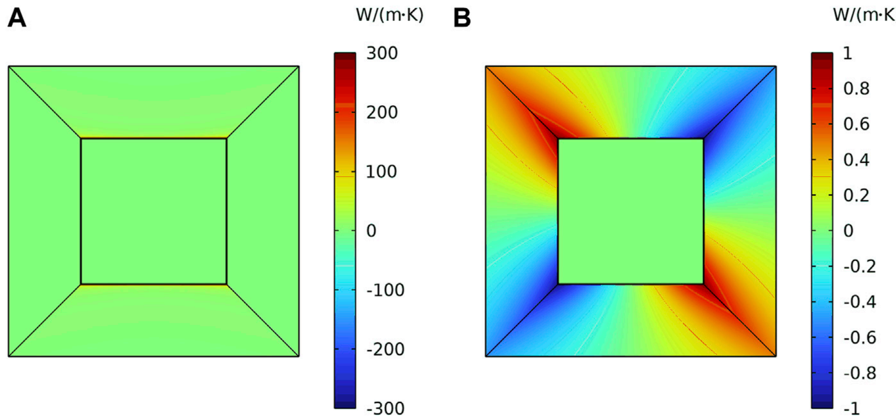

FIGURE 6 | The cloud map of thermal conductivity coefficient. **(A)** λ'_{xx}. **(B)** λ'_{xy}.

$$m' = (m-a)\frac{s_2-s_1}{s_2} + s_1 + a, n' = (n-b)\left(\frac{s_2-s_1}{s_2} + \frac{s_1}{x-a}\right) + b \tag{34}$$

Let m', n', m, n in the **formula 34** be represented by x', y', x, y respectively, the transformation equation of area 1 is obtained as follows:

$$x' = (x-a)\frac{s_2-s_1}{s_2} + s_1 + a, y' = (y-b)\left(\frac{s_2-s_1}{s_2} + \frac{s_1}{x-a}\right) + b \tag{35}$$

The corresponding Jacobian matrix and the thermal conductivity of area 1 at any position are:

$$A_1 = \begin{bmatrix} \frac{s_2-s_1}{s_2} & 0 \\ \frac{-s_1(y-b)}{(x-a)^2} & \frac{s_2-s_1}{s_2} + \frac{s_1}{x-a} \end{bmatrix},$$

$$\lambda_1' = \begin{bmatrix} 1 - \frac{s_1}{x'-a} & -\frac{s_1(y'-b)}{(x'-a)^2} \\ -\frac{s_1(y'-b)}{(x'-a)^2} & \frac{(x'-a)^4 + s_1{}^2(y'-b)^2}{(x'-a)^4 - s_1(x'-a)^3} \end{bmatrix}\lambda_0 \tag{36}$$

In a similar way, in the areas 2, 3, and 4, the thermal conductivity of a thermal cloak at any position can be obtained.

SIMULATION VERIFICATION

To verify the theoretical results of the above derivation, we use COMSOL to implement the simulation. The square-shaped thermal cloak is built in the heat transfer module. Set $s_1 = 3m$, $s_2 = 6m$. The thermal conductivity of the original space $\lambda_0 = 1w/(m \cdot k)$. Set the thermal conductivity calculated above in the material setup module. Two kinds of thermal boundaries are set for the simulation of one thermal cloak cell and the periodic plate.

Verification of One Thermal Cloak Cell

For one thermal cloak, the first thermal boundary conduction is as **Figure 7** shows, the left boundary is set to be 1000k, and the right boundary is set to be 0k. The upper and lower boundaries are set as thermal insulation. The **Figure 7A** is the heat flux direction, and the **Figure 7B** shows the diagram of isotherm. It is found that the heat flow stably flows around the heat cloak. It realizes the function of the thermal cloak.

Another is a diagonal thermal boundary, as **Figure 8** shows. The upper right corner is set as 1000k, and the lower left corner is set as 0k. Other boundaries are set as insulation. The **Figure 8A** is the heat flux direction, and the **Figure 8B** shows the diagram of isotherm. Besides, the position of the diagonal heat source is changed, so that the heat source is not on the diagonal of the heat cloak. The **Figure 9A** is the heat flux direction, and the **Figure 9B** shows the diagram of isotherm. It shows that the cloak acts as a shield against the heat flow.

Verification of a Periodic Plate With Thermal Cloak

As shown above, by introducing the position parameters a, b, the thermal conductivity tensor of the thermal invisibility cloak at any position can be obtained. According to this, the 8 × 8 periodic plate is built in the transfer module. The thermal conductivities of each cloak are set as deduced.

For the plate, similar to one thermal cloak, two kinds of the thermal boundaries are set. **Figure 10** and **Figure 11** are the simulation results of the two thermal boundaries respectively. The **Figure 10A** is the heat flux direction, and the **Figure 10B** shows the diagram of isotherm under the first thermal boundary. It indicates that the heat flows from the left to the right, and steadily bypasses each thermal cloak. The temperature gradient and heat flow inside each cloak are 0. The **Figure 11A** is the heat flux direction, and the **Figure 11B** shows the diagram of isotherm

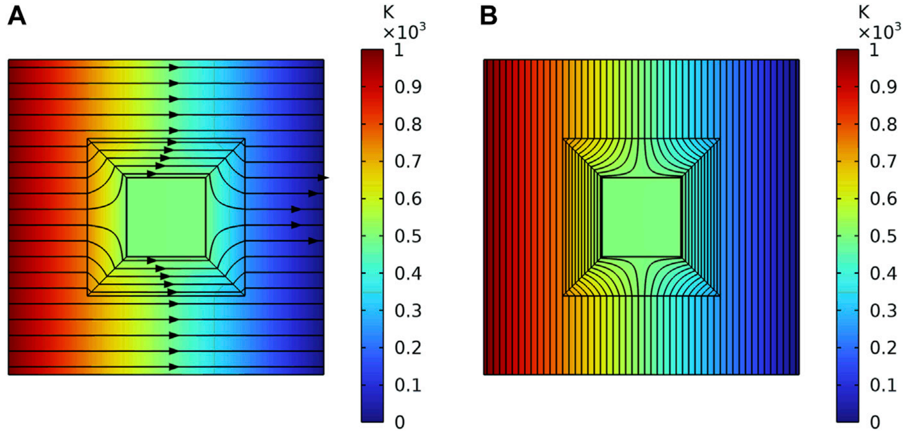

FIGURE 7 | The heat flux direction and the diagram of isotherm of one cloak under the first thermal boundaries. **(A)** The heat flux direction. **(B)** The diagram of isotherm.

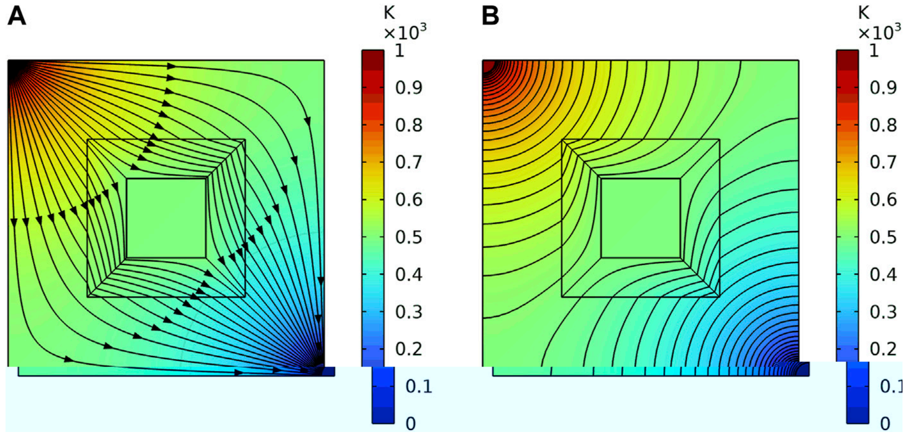

FIGURE 8 | The heat flux direction and the diagram of isotherm of one cloak under the second thermal boundaries. **(A)** The heat flux direction. **(B)** The diagram of isotherm.

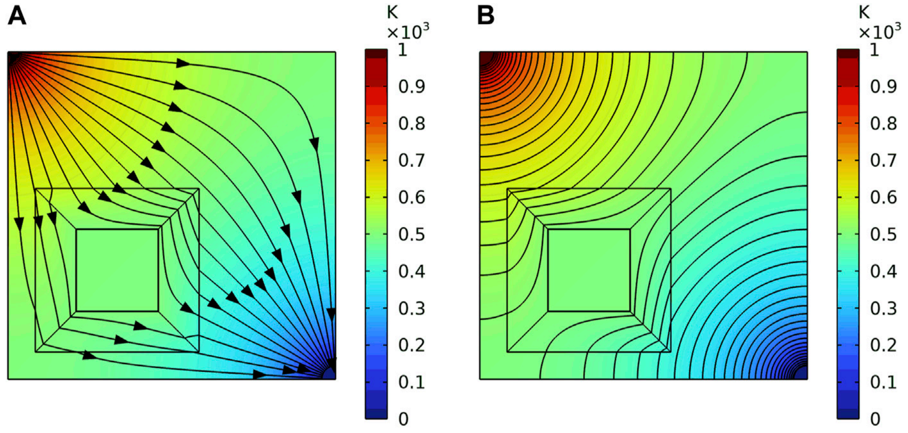

FIGURE 9 | The heat flux direction and the diagram of isotherm of one cloak under the second thermal boundaries when the heat source is not on the diagonal of the heat cloak. **(A)** The heat flux direction. **(B)** The diagram of isotherm.

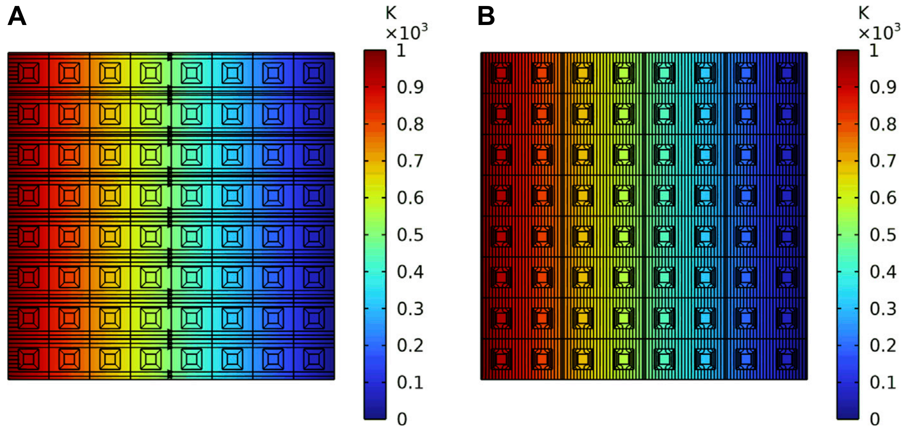

FIGURE 10 | The heat flux direction and the diagram of isotherm of the periodic plate under the first thermal boundary. **(A)** The heat flux direction. **(B)** The diagram of isotherm.

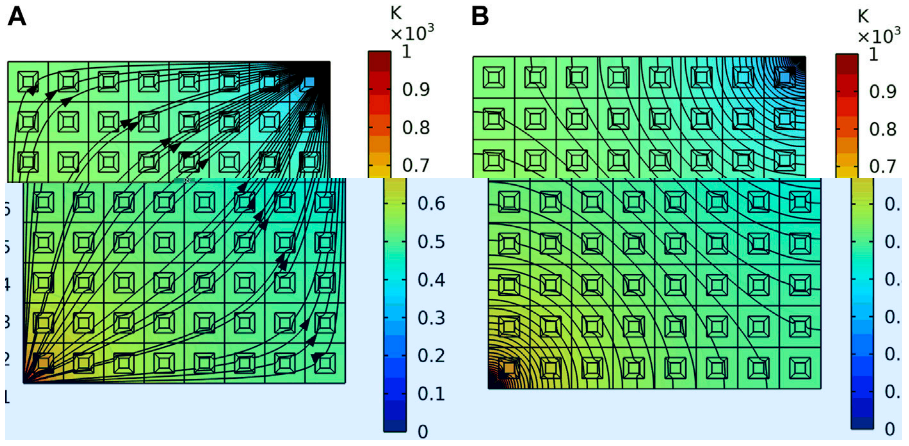

FIGURE 11 | The heat flux direction and the diagram of isotherm of the periodic plate under the second thermal boundary. **(A)** The heat flux direction. **(B)** The diagram of isotherm.

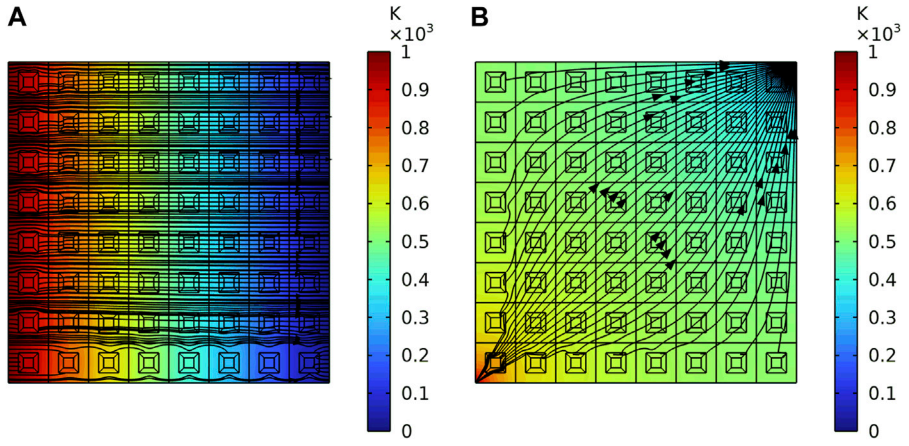

FIGURE 12 | The heat flux direction of the periodic plate under two thermal boundaries when the position parameters are abandoned. **(A)** Under the first thermal boundary. **(B)** Under the second thermal boundary.

under the diagonal thermal boundary. The heat flows steadily bypass each heat cloak. The heat flow in the center of the periodic plate is 0 and the closer to the diagonal heat source, the denser the isotherm. It proves that each cloak in the plate has the function of avoiding heat flow.

The position parameters a and b are not considered to prove their necessity. Two of the same kinds of thermal boundaries are set. **Figures 12A, B** are the heat flux direction under the two thermal boundaries when the position parameters a and b are abandoned. We know that when the position parameters a and b are not considered, the heat flow cannot stably bypass each thermal invisibility cloak. It can be found that the heat flux inside the cloaks that are located in the first column on the left and the last row on the bottom are closest to 0. That is because the thermal conductivity coefficient of the thermal invisibility cloak at those locations only changes one direction when transforming coordinates. As a result, the position parameters are not even a consideration, the heat cloak at these locations can still play a role in avoiding heat flux. Therefore, when arranging the thermal invisibility cloak periodically, the position parameters a and b must be considered.

DISCUSSION

Two methods have been given to construct the thermal conductivities of the square thermal invisibility cloak. The cloud map of the thermal conductivity coefficient of the thermal cloak has been drawn. Besides, by introducing the position parameters into the calculation of the thermal conductivity coefficient, the cloaks were arranged periodically so that they can manipulate heat flow in a larger area. The heat insulation function of both the single thermal cloak and the thermal cloak periodic plate were tested under different heat boundary conditions using COMSOL Multiphysics.

The results show that both the single thermal and the thermal cloak periodic plate have the function of avoiding heat flow under different heat boundary conditions. The heat fluxes travel around the inner domain with good thermal stealth effect. To prove the necessity of the position parameters, the comparative verification is done. The results show that the heat flow can stably bypass each thermal invisibility cloak only when the position parameters are considered.

In engineering application, vibration reduction and heat insulation have always been the focus of research. Phonon crystal plates can be used to reduce response. How to combine the phonon crystal plate with the thermal cloak periodic plate to play the role of shock absorption and heat insulation at the same time is worthy of further study.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

YYS conducted major research work and wrote this paper. YC and XY directed and participated in the discussion of the research work. YL directed the overall research work and guided the writing of the thesis. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the National Science Foundation of China (Grant No. 52192633).

CONFLICT OF INTEREST

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

REFERENCES

1. Zhang Z, Xu L, Qu T, Lei M, Lin ZK, Ouyang X, et al. Diffusion Metamaterials. *Nat Rev Phys* (2023) 5:218–35. doi:10.1038/s42254-023-00565-4
2. Yang S, Wang J, Dai G, Yang F, Huang J. Controlling Macroscopic Heat Transfer With thermal Metamaterials: Theory, Experiment and Application. *Phys Rep* (2021) 908:1–65. doi:10.1016/j.physrep.2020.12.006
3. Pendry JB, Schurig D, Smith DR. Controlling Electromagnetic Fields. *Science* (2006) 312:1780–2. doi:10.1126/science.1125907
4. Fan C, Gao Y, Huang J. Shaped Graded Materials With an Apparent Negative Thermal Conductivity. *Appl Phys Lett* (2008) 92:251907. doi:10.1063/1.2951600
5. Li J, Gao Y, Huang J. A Bifunctional Cloak Using Transformation Media. *J Appl Phys* (2010) 108:074504. doi:10.1063/1.3490226
6. Narayana S, Sato Y. Heat Flux Manipulation With Engineered Thermal Materials. *Phys Rev Lett* (2012) 108:214303. doi:10.1103/PhysRevLett.108.214303
7. Schittny R, Kadic M, Guenneau S, Wegener M. Experiments on Transformation Thermodynamics: Molding the Flow of Heat. *Phys Rev Lett* (2013) 110:195901. doi:10.1103/PhysRevLett.110.195901
8. Mao F, Li T, Huang M, et al. Research and Design of Heat Cloak With Non-Conformal Arbitrary Cross-Section. *Acta Phys Sin* (2014) 63:01. doi:10.7498/aps.63.014401
9. Li Y, Shen X, Wu Z, Huang J, Chen Y, Ni Y, et al. Temperature-Dependent Transformation Thermotics: From Switchable Thermal Cloaks to Macroscopic Thermal Diodes. *Phys Rev Lett* (2015) 115:195503. doi:10.1103/PhysRevLett.115.195503
10. Xia G, Yang L, Kou W, et al. Design and Research of Three-Dimensional Thermal Cloak With Arbitrary Shape Based on the Transformation Thermodynamics. *Acta Phys Sin* (2017) 66:10. doi:10.7498/aps.66.104401
11. Xia G, Yang L, Kou W. Design and Simulation of Square-Shaped Thermal Cloaks With Arbitrary Angles. *Journal Naval University Engineering* (2018) 30:02. doi:10.7495/j.issn.1009-3486.2018.02.017
12. Sha W, Xiao M, Zhang J, Ren X, Zhu Z, Zhang Y, et al. Robustly Printable Freeform Thermal Metamaterials. *Nat Commun* (2021) 12:7228. doi:10.1038/s41467-021-27543-7

13. Li Z, Wang X, Li Y. The Band gap Variation of a Two Dimensional Binary Locally Resonant Structure in Thermal Environment. *AIP Adv* (2017) 7:1. doi:10.1063/1.4973723
14. Zhu Y, Li Z, Li Y. The Lamb Wave Bandgap Variation of a Locally Resonant Phononic Crystal Subjected to Thermal Deformation. *AIP Adv* (2018) 8:5. doi:10.1063/1.5026523
15. Milton GW, Briane M, Willis JR. On Cloaking for Elasticity and Physical Equations With a Transformation Invariant Form. *New J Phys* (2006) 8:248. doi:10.1088/1367-2630/8/10/248
16. Shen X, Huang J. Transformation Thermotic: Thermal Metamaterials and Their Applications. *Acta Phys Sin* (2016) 65:17. doi:10.7498/aps.65.178103
17. Rahm M, Schurig D, Roberts DA, Cummer SA, Smith DR, Pendry JB. Design of Electromagnetic Cloaks and Concentrators Using Form-Invariant Coordinate Transformations of Maxwell's Equations. *J Photon Nanostructures: Fundamentals Appl* (2008) 6:87–95. doi:10.1016/j.photonics.2007.07.013

AEROSPACE RESEARCH COMMUNICATIONS
ORIGINAL RESEARCH
published: 12 December 2023
doi: 10.3389/arc.2023.12261

An Optical Attitude Measurement Method for Terminal Sensitive Projectile in Vertical Wind Tunnel Test

Xianshi Yuan[1,2], Chao Gao[1], Zhengke Zhang[1], Huan Wang[2], Yushuai Wang[1] and Yueqiang Li[1]*

[1]*School of Aeronautics, Northwestern Polytechnical University, Xi'an, China,* [2]*Xi'an Institute of Modern Control Technology, Xi'an, China*

An optical attitude (scanning angle and spin rate) measurement method for terminal sensitive projectile in vertical wind tunnel test is proposed. The principles of measuring scanning angle and spin rate simultaneously depend only on one fixed camera are presented in detail. The measurement processes, which mainly includes preparation, image acquisition, segmentation of projectile, contour extraction, and estimation of attitude, are described accordingly, and the key techniques involved in each step are expanded exactly. Moreover, the measurement errors are analyzed thoroughly. The advantages of the novel method is analyzed as follow: first, the mathematical model for calculating scanning angle of a steady-state terminal sensitive projectile is derived for the first time in literature, which makes it possible to measure scanning angle from images recorded by only a fixed camera; second, the algorithms of estimating scanning angle and spin rate depend only on the line-segments of the perspective cylinders in images are designed, which significantly improve the attitude measurement accuracy. Experiment was also conducted in laboratory to verify the novel method.

Keywords: terminal sensitive projectile, scanning angle, spin rate, optical measurement, vertical wind tunnel test

***Correspondence:**
Xianshi Yuan
yxsvilin@163.com

***Received:** 20 October 2023*
***Accepted:** 21 November 2023*
***Published:** 12 December 2023*

***Citation:**
Yuan X, Gao C, Zhang Z, Wang H, Wang Y and Li Y (2023) An Optical Attitude Measurement Method for Terminal Sensitive Projectile in Vertical Wind Tunnel Test. Aerosp. Res. Commun. 1:12261. doi: 10.3389/arc.2023.12261

INTRODUCTION

Terminal-sensitive projectile, also called sensor-detonated projectile, is a kind of cluster bomb which contains many smaller bombs and is very disruptive. It can be launched by cannons, long-range rockets or aircraft. When flying to the target over a certain height, the smaller bombs are thrown and they can automatically detect and recognize the targets and make accurate attacks when they reach steady-flying (or steady-scanning) state after a deceleration stage. Since the detection and attack operations are performed during the steady-scanning stage, the steady-scanning attitude such as falling velocity, scanning angle, and spin rate are very important parameters that affecting the performance and hit probability of the terminal sensitive projectile [1]. Therefore, it is necessary to measure the steady-scanning attitude of the terminal sensitive projectile through either vertical wind tunnel test or field test.

There are already many sensors [2] developed to measuring the steady-scanning attitude of the projectile. According to the installation position, existing sensors can be classified into two categories: projectile-borne sensors and non-projectile-borne sensors. Projectile-borne sensors [3–17] such as accelerometers, gyroscopes, geomagnetic sensors, and solar sensors are generally installed inside the body of the projectile. Schuler et al [3] proposed to use linear accelerometers to measure the spin rate of rotating objects. Without loss of generality, three or more linear accelerometers were placed at different axis of the projectile coordinate system, the linear acceleration along each axis were measured during high-speed flight, and the spin rate was thereby calculated based on the collected acceleration information. This type of sensors has the advantages of low cost, low-power consumption, and high reliability, etc. [4]. However, as the number of the linear accelerometers increases, the installation error increases, and more space is needed to place them which is not very practical because the body space of the projectile is limited. Moreover, the error of the calculated spin rate increases dramatically over time.

Gyroscopes, which often employed in inertial navigation systems, are another type of projectile-borne sensors. The traditional gyroscopes can be used to measure the scanning angle and spin rate of the projectile. They have the features of working all the time, and high accuracy and high stability in short term. Whereas, their volumes are usually large and their overload capacity are weak [5]. Recently, MEMS (Microelectromechanical system) is applied to the gyroscopes, and new MEMS gyroscopes, which have no moving parts, are developed to measure the attitude of the projectiles. The new MEMS gyroscopes usually have small volume and low cost. But their measuring ranges are often small and their long-term stability is low, the measuring error of the attitude increases over time as the linear accelerometers [6, 7].

Geomagnetic sensor, which obtain the attitude information through measuring the magnetic field vector information, is one of the most concerned research focuses among all attitude measuring sensors for spin missiles. Zhou et al. [8] presented an attitude calculation method based on geomagnetic sensor by establishing the geomagnetic model and considering the characteristics of steady scanning motion of terminal sensitive projectile. Deng et al. [9] studied a real-time roll angle estimation method based on phase-locked loop on signals from single-axis magnetometer for spinning projectile. The novel method is less dependent on the amplitude and able to reduce effect from geomagnetic blind area. Geomagnetic sensor owns the advantages of passive sensing, high sensitivity, as well as low power and cost [10]. Whereas, its accuracy depended on accurate geomagnetic model which usually difficult to obtain and vulnerable to environmental interference.

Solar sensor measures the spin rate, nutation angle and precession angle through recorded solar azimuth which is defined as the angle between missile rotational axis and the sunlight pass through the missile center. Gui and Yang [11] presented a scheme of practicable digital telemetering system for measuring the attitude of projectile in flight based on the principle of solar aspect angle. The system had been used in practice for the test of projectiles which survive as high as 18 thousand g. Although the principle of solar sensor is simple, it is severely limited by the sunlight and weather conditions, and its measuring accuracy is not high.

Owing to the limitations of each type of projectile-borne sensors, the multi-sensor systems which combine the accelerometers, gyroscopes, geomagnetic sensors and solar sensors are thereby designed [12–17]. Change et al. [12] estimated the attitude of projectiles using magnetometers and accelerometers. Harkins et al. [13] combined the solar sensors and magnetic sensor to measure the spin rate and scanning angle of spinning projectiles. Cao et al. [14] put forward a low-cost attitude detect system which consists of geomagnetic sensor and micro silicon gyro. Huang et al. [15] employed magnetometer, MEMS accelerometer and gyroscope to estimate attitude information. Since the multi-sensor systems can overcome the shortage of each type of sensors, either the robustness or the accuracy of the measuring system is improved significantly.

Non-projectile-borne sensors [1, 18–23], such as the optical measurement system (also known as high-speed video measuring technology or image-based method), Radar and global navigation satellite system (GNSS) etc., are usually placed outside of the body of the terminal sensitive projectile. The GNSS includes American GPS, Russian GLONASS, European Galileo, and Chinese BDS. Among them, the American GPS system is most widely used. Deng et al. [18] adopted a GPS (Global Positioning System) receiver with a single side-mounted antenna to detect the roll angle and rotational speed of a spinning vehicle A Frequency-Locked Loop (FLL) assisted Phase-Locked Loop (PLL) is designed to obtain the attitude information from GPS signals. Cao et al. [19] studied the roll angle calculation method based on Geomagnetic sensor and GNSS to meet the needs of high-frequency and high-accuracy roll angle measurement for the high-spin projectile. Because of limited body space of the projectile, either he layout of the baseline or the installation of the receiver is restricted, the measuring accuracy of the GNSS is not high.

Radar can also be used to measure the spin rate of high-spin projectile. Ferguson et al. [20] employed Doppler radar to measure the steady-state spin rates of the kinetic energy projectiles. Li et al. [21] presented a new method to measuring the rotation velocity of small arms projectiles by transforming Doppler radar data to the time domain on the premise of eliminating corresponding noise and clutter signals in the frequency domain. The disadvantage of Radar applied to measure the steady-state spin rates of terminal-sensitive projectile is that slot need to be carved at the bottom of the projectile.

Optical measurement instruments/systems are widely used in modern weapon test. Due to high-speed video technique, it is possible to smoothly view the steady-state rotation of a projectile spinning at 300 revolutions per second (HZ) with a camera recording at a rate of 10,000 frames per second. Gao et al. [22] used two cameras to measure the scanning angle and spin rate of two different types of terminal-sensitive projectiles in vertical wind tunnel test, and compared the measuring results with field test. Yakimenko et al. [1] presented a scheme of measuring spin rate of marked projectiles using computer vision analysis method. Black and white longitudinal stripes along the direction of the projectile axis should be painted on the surface of the projectile in advance. By counting the stripe pixel column of the recorded images, the spin rate of the projectile can be obtained exactly. Zhao et al. [23] proposed a model-based optimization algorithm to estimate the motion parameters of projectile from the recorded images from stereo linear array image. The optimal motion parameters are achieved by

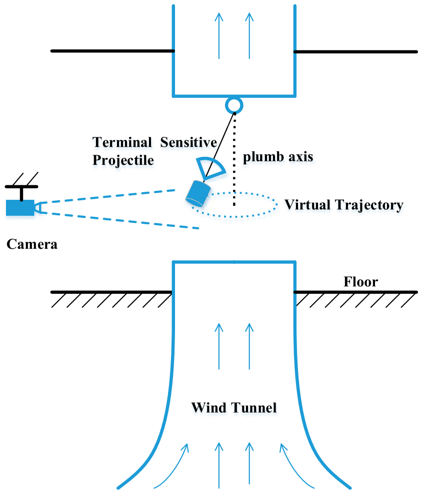

FIGURE 1 | Schematic diagram of vertical wind tunnel test.

matching the stereo projection of the projectile and that of the same size 3D model. The optical measurement instruments/systems usually have higher accuracy than other sensors in spin rate measurement, and are especially suitable for wind tunnel test. Whereas, they are often lack of precision in scanning angle measurement.

In this paper, an optical attitude measurement method for terminal sensitive projectile in vertical wind tunnel test is proposed. The optical measurement method is extremely simple since only one high-speed camera is employed. However, both scanning angle and spin rate can be measured accurately by using the novel method. The measurement principle, measurement process, and measurement errors are presented in detail. Compared with existing optical attitude measurement methods, the contribution of the novel method can be analyzed in two aspects: first, the mathematical model for calculating scanning angle of a steady-state terminal sensitive projectile is derived, which makes it possible to measure scanning angle from images recorded by only a fixed camera; second, the algorithms of estimating scanning angle and spin rate depend only on the line-segments of the perspective cylinders in images are designed, which significantly improve the measurement accuracy.

MEASUREMENT PRINCIPLE

Steady-Scanning Model

The experiment setup is shown in **Figure 1**. The length of the experimental section of the vertical wind tunnel is about 6.5 m, the diameter of the contraction section of the vertical wind tunnel is Ø4.5 m, the maximum wind speed for the experimental section is about 50 m/s (meter per second), and the minimum wind speed is about 5 m/s. During experiment, the terminal-sensitive projectile is hung in the air through a flexible rope over a pulley and rotate around a plumb axis, the steady spin rate ranges from 4 to 6 r/s (revolutions per second) to about 30 r/s (for different types of terminal-sensitive projectiles). It is widely recognized that the steady-state trajectory of a terminal-sensitive projectile follows an Archimedean spiral. However, at any given moment, the trajectory can also be visualized as a virtual cone. The generator line of this cone corresponds to the axis of the projectile, while the axis of the cone aligns with the plumb axis. The vertex of the cone represents the point of intersection between the plumb axis and the projectile's axis. Therefore, the scanning angle at each moment can be understood as the half-cone-angle of the corresponding virtual cone. Consequently, measuring the real-time scanning angle is equivalent to determining the cone-angle of the corresponding virtual cone.

Layout of Optical Measurement System

The configuration of the optical measurement system is particularly simple, as shown in **Figure 1**, only one high-speed camera is employed to record images of the terminal-sensitive projectile in flight. The camera can be installed on the wall, or on a tripod, but should be in a horizontal orientation and the optical axis of which should be orthogonal to the plumb axis. Moreover, the distance of the camera to the plumb axis should be far greater than the radius of the virtual cones, and some high-lumen LED lamps should be placed around the vertical wind tunnel to provide adequate lighting for the optical system.

Principle of Measuring Scanning Angle

As described in *Steady-Scanning Model* section, the steady-state trajectory of the terminal-sensitive projectile can be visualized as a virtual cone at each moment. Consequently, capturing the steady-scanning trajectory of the projectile on an image plane is essentially projecting the virtual cone onto that plane. In **Figures 2A, B**, the virtual cone at a specific moment is represented by a blue dashed line, with the generator line depicted as a blue solid line. The vertex of the cone is denoted as point S. To establish a coordinate system for the virtual cone, we introduce the coordinate system $O_w - X_w Y_w Z_w$, the origin O_w is located at the base center of the virtual cone, the axis Z_w is coincides with the axis of the virtual cone, and the axis X_w is parallel to the axis Z_c of the camera coordinate system. The intersection of the base border of the virtual cone and the axis Y_w is marked by points C and D. Furthermore, the camera coordinate system $O_c - X_c Y_c Z_c$ is defined, with the origin O_c representing the optical center of the camera, the axis Y_c is parallel to the axis Z_w, and the axis Z_c, which aligns with the optical axis of the camera, is orthogonal to the axis Z_w and intersects at point K.

According to the principles of Descriptive Geometry, the projection of a cone is an isosceles triangle, the two congruent legs of which are projection of the limit elements on the projection view. As shown in **Figure 2C**, The coordinate system o $-xy$ is the image coordinate system, the origin o is projection of the optical center O_c, the axis x is parallel to the axis X_c, and the axis y is parallel to the axis Y_c. The isosceles triangle Δsab on the image plane is projection of the virtual cone, and the two congruent legs sa and sb are projections of the limit elements SA and SB which are tangent to the rays O_cA and O_cB (it should be noted that the projection of the base of the virtual cone may be

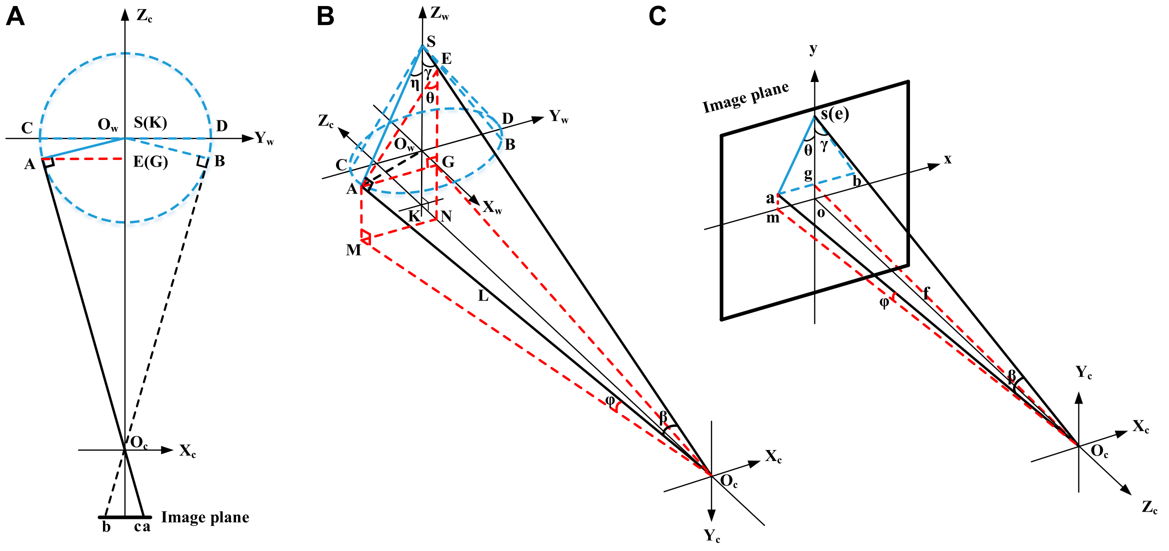

FIGURE 2 | Imaging principal of the steady-scanning trajectory of the terminal-sensitive projectile in vertical wind tunnel test. **(A)** Top-down view of the ray of the imaging system. **(B)** Isometric view of the ray from the optical center of the camera to the scene. **(C)** Isometric view of the ray from the optical center of the camera to the image plane.

not a line since the base would not be coincident with the plane of $O_c - X_cZ_c$, thereby the projection of the virtual cone may be not a triangle, to simply the analysis, the triangle is still used as instead here because the simplification does not cause any loss of the accuracy). Since the angle between any element on the conical surface and the axis of the cone is the half-cone-angle, if any one of the four angles $\angle ASO_w$, $\angle BSO_w$, $\angle CSO_w$, and $\angle DSO_w$ is calculated, the scanning angle is obtained accordingly. As a matter of fact, the projections of the two angles $\angle CSO_w$ and $\angle DSO_w$ on the image plane are equivalent to the themselves. Unfortunately, the projections of point C and point D cannot be identified on the image plane, so these two angles cannot be solved. The projections of the two angles $\angle ASO_w$ and $\angle BSO_w$ are the two angles $\angle aso$ and $\angle bso$ on the image plane, respectively.

Because there is perspective transformation between the angles and their projections, and the perspective transformation is difficult to obtain, the true values of the angles cannot be derived directly. In this section, both ray trace method and geometric analysis are employed to compute the true values (the values of the two angles $\angle ASO_w$ and $\angle BSO_w$ are equal, only the angle $\angle ASO_w$ is computed here). As shown in **Figure 2B**, create an auxiliary plane through point A that is perpendicular to the axis X_w and intersects with the ray O_cS and the axis X_w and Z_c at the points E, G and N, respectively. Then connect those intersection points and draw five auxiliary lines AE, AG, EG, GN and O_cG with red dotted line. Since the plane where the triangle ΔAGE is located is perpendicular to the axis Z_c, it is parallel to the image plane, hence the angle $\angle AEG$ is equal to its projection, i.e., $\angle AEG = \angle aeg$. Furthermore, according to the ray trace theory, the projections of point S and point E are coincident, and the projection of point G is located at the midpoint of the line ab on the image plane, if the angles $\angle aso$ is marked as θ, thus $\angle AEG = \angle aeg = \angle aso = \theta$. Obviously, the angle $\angle ASO_w$ is not equal to the angel $\angle AEG$, i.e., $\angle ASO_w \neq \angle AEG$, hence $\angle ASO_w \neq \theta$. That is the value of the angle $\angle ASO_w$ cannot be simply replaced with the value of its projection $\angle aso$.

Draw an auxiliary line AM through point A that is perpendicular to the plane $O_c - X_cZ_c$ and intersects at point M with red dotted line (**Figure 2B**). Meanwhile, draw an auxiliary line am through point a that is perpendicular to the axis x and intersects at point m with red dotted line (**Figure 2B**) According to the ray trace theory, the point m is projection of point M and there is a virtual ray between point m and point M (connected with a red dotted line). Moreover, connect the point A and point O_w with a black dotted line (**Figure 2B**). And mark the angle $\angle AO_cM$ as φ, the angle $\angle AO_cS$ as β, the angle $\angle O_cSK$ as γ , the angle $\angle ASO_w$ as η, and the length of the ray O_cA as L. As shown in **Figure 2B**, according to the Pythagorean theorem,

$$\begin{cases} \cos\eta = SO_w/SA \\ \tan\beta = SA/L \\ \cos\beta = L/SO_c \\ \cos\gamma = SK/SO_c \\ \sin\varphi = AM/L \end{cases}. \tag{1}$$

In addition, according to the geometric relationship,

$$\begin{cases} SK = SO_w + O_wK \\ O_wK = GN = AM \end{cases}. \tag{2}$$

By combing **Formulas 1**, **2**, the angle η can be computed

$$\begin{aligned} \eta &= \cos^{-1}\left(SO_w/SA\right) \\ &= \cos^{-1}\left(\left(SK - AM\right)/\left(L \times \tan\beta\right)\right) \\ &= \cos^{-1}\left(\left(L \times \cos\gamma/\cos\beta - L \times \sin\varphi\right)/\left(L \times \tan\beta\right)\right) \\ &= \cos^{-1}\left(\left(\cos\gamma/\cos\beta - \sin\varphi\right)/\tan\beta\right) \end{aligned}. \tag{3}$$

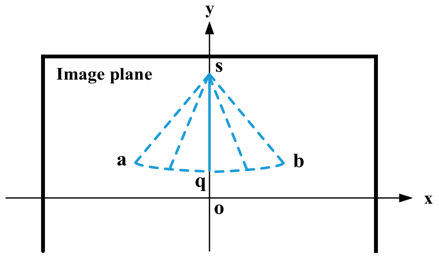

FIGURE 3 | Imaging of the steady-state motion of the terminal-sensitive projectile.

According to **Formula 3**, if the three angles β, γ and φ are solved, the angle η can be calculated exactly.

Assume the focal length of the camera is f, the pixel sizes are dx and dy along the x and y directions respectively, and the two-dimensional (2D) pixel coordinate of point a is (x_a, y_a). As shown in **Figure 2C**, according to the Pythagorean theorem,

$$\begin{cases} \sin\theta = |dx \times x_a|/sa \\ \cot\theta = sg/|dx \times x_a| \\ \cos\gamma = so/sO_c \\ \sin\varphi = dy \times y_a/aO_c \\ |sO_c|^2 = so^2 + f^2 \\ |aO_c|^2 = |dx \times x_a|^2 + |dy \times y_a|^2 + f^2 \end{cases}, \tag{4}$$

where the operator | | means taking absolute value. According to the Cosine theory,

$$\cos\beta = \left(|aO_c|^2 + |sO_c|^2 - sa^2\right)/\left(2 \times |aO_c| \times |sO_c|\right). \tag{5}$$

In addition, $so = sg + dy \times y_a$, by combining **Formulas 4, 5**, it can be inferred that

$$\begin{cases} \gamma = \cos^{-1}\left((|dx \times x_a| \times \cot\theta + dy \times y_a)/\sqrt{(|dx \times x_a| \times \cot\theta + dy \times y_a)^2 + f^2} \right) \\ \varphi = \sin^{-1}\left(dy \times y_a/\sqrt{|dx \times x_a|^2 + |dy \times y_a|^2 + f^2} \right) \\ \beta = \cos^{-1}\left(\dfrac{|dy \times y_a|^2 + |dx \times x_a| \times dy \times y_a \times \cot\theta + f^2}{\sqrt{|dx \times x_a|^2 + |dy \times y_a|^2 + f^2} \times \sqrt{(|dx \times x_a| \times \cot\theta + dy \times y_a)^2 + f^2}} \right) \end{cases}. \tag{6}$$

According to **Formula 6**, the three angles β, γ and φ are functions of the six variables f, dx, dy, x_a, y_a and θ. The 2D pixel coordinate (x_a, y_a) of point a and the angle θ can be achieved through image analysis exactly. The focal length f and the pixel sizes dx and dy can be obtained from the handbook of the high-speed camera. Thus, the three angles β, γ and φ can be computed exactly, and the value of half-cone-angle η can be obtained.

Note that the half-cone-angle η can be achieved only when the terminal-sensitive projectile rotates to position A and B as shown in **Figure 2A**, if rotates to other position, **Formulas 1, 3** are not valid any more, thus, the half-cone-angle η cannot be solved. That is in one complete revolution, the real time scanning angle of the terminal-sensitive projectile at two points can be measured in vertical wind tunnel test.

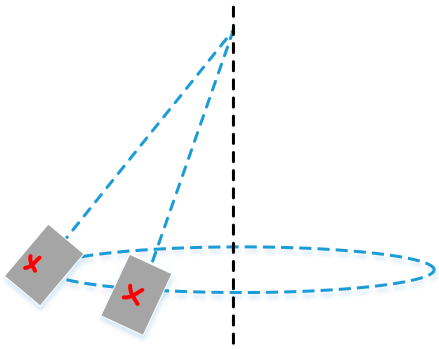

FIGURE 4 | View of the pattern "×" drawn on the surface of the terminal-sensitive projectile during rotation.

Principle of Measuring Spin Rate

Like the moon whose rotation rate is the same as the time it takes to make one revolution, the terminal-sensitive projectile rotates on its own axis in the same period of time that revolves around the plumb axis (**Figure 1**) during the steady-scanning stage in the vertical tunnel test, thus, the steady-state spin rate can be achieved by measuring either the rotation rate or the revolution speed of the terminal-sensitive projectile [1, 22].

The revolution speed can be solved by recording the time taken to make one revolution, or by counting the number of revolutions within a given time. The former can be seen as computing the real-time velocity while the former can be seen as computing the average velocity. Different from previous method that employs two cameras to track the revolution speed of the terminal-sensitive projectile, in this paper, only one camera is employed. As shown in **Figure 3**, the steady-state motion of the terminal-sensitive projectile on the image plane is similar to the pendulum motion which is a typical oscillating motion. Thus, the projection of the axis of the projectile must be coincident with the y-axis periodically. Mark the time of the ith coincidence as T_i where $i = 1, 2, \ldots, N$, thereupon the real-time velocity at time T_i can be expressed as

$$v_i = 1/\left(T_i - T_{i-2}\right), \tag{7}$$

and the average velocity can be expressed as

$$\bar{v} = 0.5 \times (N-1)/\left(T_n - T_1\right). \tag{8}$$

The rotation rate can be achieved by tracking features such as the natural texture on the surface of the terminal-sensitive projectile, pattern painted or drawn manually, corners, or specialized targets that widely used in optical measurement [1, 23]. Since the terminal-sensitive projectile rotates on its own axis, the shapes and coordinates of the features at different position are different when view from the optical axis as shown in **Figure 4**. Furthermore, the rotation rate is equal to the revolution speed, hence the shapes and coordinates of the features at the same position in different circles of rotation are almost identical. Therefore, the rotation rate can be obtained by tracking the

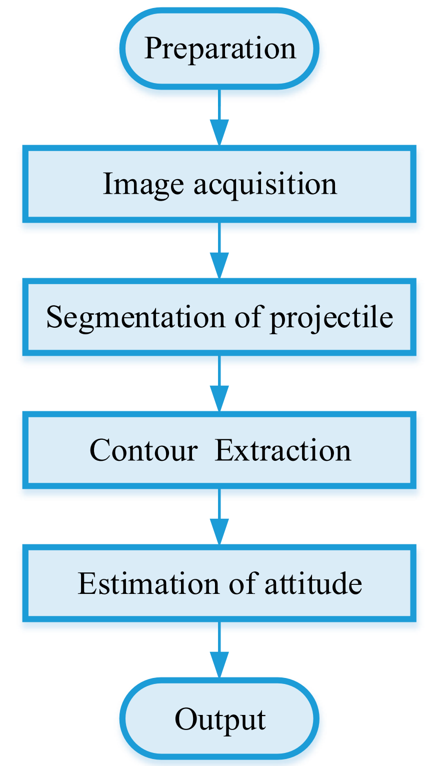

FIGURE 5 | Flowchart of measurement process.

shapes and coordinates of the features at some positions where they can be seen clearly using feature tracking algorithms or deep learning algorithms.

MEASUREMENT PROCESS

The measurement process mainly includes five steps: preparation, image acquisition, segmentation of projectile, contour extraction, and estimation of attitude, as shown in **Figure 5**. The key techniques involved in each step will be presented in detail.

Preparation and Image Acquisition

Before image acquisition, the camera should be adjusted to a ready state: adjust the camera to be horizontal by means of a spirit level, and align the vertical axis of the crosshair of the camera to be coincident with the stationary flexible rope which hanging the terminal-sensitive projectile in the air (the flexible rope is coincident with the plumb axis when it is in stationary state). Then record an image of the static scene as background image for segmentation step. Finally, begin to record images when the motion of the terminal-sensitive projectile is in a steady state.

Segmentation of Projectile

Segmentation of the projectile is essentially detecting moving object from complex scenes which is an important topic in the field of computer vision. Many methods, such as the frame difference [24], background subtraction [25], optical flow [26], or deep learning-based method [27], etc., have been proposed for moving object detection with either a stationary camera or a moving camera. Since the camera is fixed during the measurement, background subtraction is adopted here.

The background subtraction is the most effective method for detecting moving objects in video or a series of images [24]. The core of background subtraction is building the background model/image. In this paper, the background image is recorded in advance as mentioned in *Preparation and Image Acquisition* section. Suppose the gray value of the background image is given by $I(x, y, 0)$, and the gray value of the frame at time t is $I(x, y, t)$, where x and y are values of the 2D coordinate of pixel. The binary image operation results of background subtraction can be defined as

$$f(x, y, t) = \begin{cases} 1 & |I(x, y, t) - I(x, y, 0)| > T \\ 0 & |I(x, y, t) - I(x, y, 0)| \leq T \end{cases} \tag{9}$$

where the operator $||$ means taking absolute value, and T is the threshold. According to **Formula 9**, for each point in the image, if the difference value is greater than the threshold, then put the point as a foreground pixel, otherwise, regarding the point as a background pixel.

Because the flexible rope and parachute that linked with the terminal-sensitive projectile also rotate synchronously, they are also extracted during the segmentation. To separate the terminal-sensitive projectile from the flexible rope and parachute, after threshold, the morphological open operator is applied to the binary image several times until the projectile is separated completely. Finally, cluster the remain pixels in the binary image base on 4-neighbors and pick out the largest cluster which is the pure projectile.

Contour Extraction

The main body of the terminal-sensitive projectile is a cylinder and its projection on the image plane is a perspective 2D cylinder. To determine its orbit in the recorded images, the contour of the perspective 2D cylinder, which generally consists of two line-segments and two elliptic arcs, should be extracted to fit the axis of the cylinder. As accuracy is required for fitting, subpixel edge extraction operators [28] are used to extract subpixel contour points.

Because some of the contour points belong to the line-segments, some belong to the elliptic arcs, and a few belong to noise, it is necessary to separate them. Due to incompleteness, the previous works [29] have proven that the ellipse fitting may be inaccurate only rely on the contour points belong to the elliptic arcs. Therefore, the contour points belong to the elliptic arcs are not used here, only the contour points belong to the line-segments are adopted. To distinguish the contour points that belong to the line-segments, the random sample consensus (RANSAC) algorithm [30] is employed. The basic procedure is as follow:

Step 1. Randomly selecting two contour points from the extracted N contour points above;

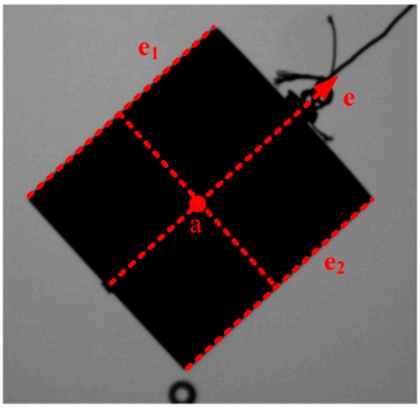

FIGURE 6 | Extraction of line-segments when the ray is orthogonal to the axis of the rotating cylinder.

Step 2. Calculating a line with the two points selected in Step 1;

Step 3. Calculating distance of the other $N-2$ contour points to the line obtained in Step 2, if the distance of a contour point is smaller than the threshold τ, it is labeled as an inlier, otherwise labeled as an outlier;

Step 4. Repeating Step 1~Step 3 until the number of repetitions exceeds $N/2$.

Step 5. Selecting the two lines with the most inliers.

To ensure accuracy, the threshold τ in the above procedure is set to one pixel, i.e., $\tau = 1$.

Estimation of Attitude

As analyzed in *Principle of Measuring Scanning Angle* section, the scanning angle can be solved only when the terminal-sensitive projectile rotates to position where the ray is orthogonal to the axis of the projectile. **Figure 6** shows an image of a cylinder rotated to that position. It can be seen that the two line-segments are parallel to each other, and the ellipse arcs almost become lines too. Mark the two line-segments as $\mathbf{e}_1$ and $\mathbf{e}_2$, the axis of the cylinder as **e**, then $\mathbf{e}_1 \parallel \mathbf{e}_2 \parallel \mathbf{e}$. Assume the angles between the line-segments $\mathbf{e}_1$, $\mathbf{e}_2$ and the y-axis of the image coordinate system is θ_1 and θ_2, respectively, thus the angle θ between the axis **e** and the y-axis can be calculated as

$$\theta = (\theta_1 + \theta_2)/2. \tag{10}$$

Once the two line-segments $\mathbf{e}_1$ and $\mathbf{e}_2$ are fitted, the angles θ_1 and θ_2 can be computed exactly, the angle θ thereby can be solved. That is the axis **e** does not need to be fitted any more.

The position of point a is determined as follow: project the center points of line-segments $\mathbf{e}_1$ and $\mathbf{e}_2$ to the axis **e**, then take the midpoint of the two projected points on the axis **e** as point a. Because the length of line-segments $\mathbf{e}_1$ and $\mathbf{e}_2$ are almost equal to each other, the projections of the center points on the axis **e** are almost coincident as shown in **Figure 6**. Once the position of point a is determined, its 2D coordinate (x_a, y_a) can be achieved exactly. Finally, by combing the parameters (the focal length f and the pixel sizes dx and dy) of the high-speed camera, the scanning angle η can be calculated exactly through **Formulas 3**, **6**.

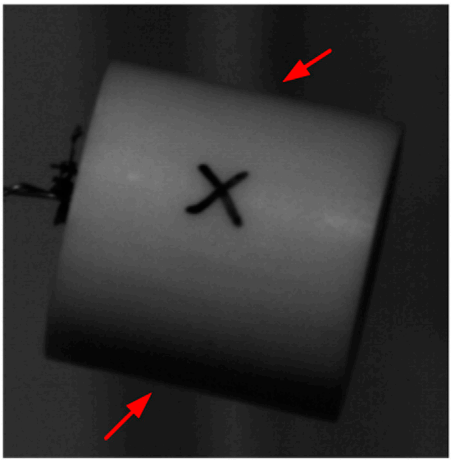

FIGURE 7 | A frame of a rotating cylinder whose surface is marked with an artificial pattern.

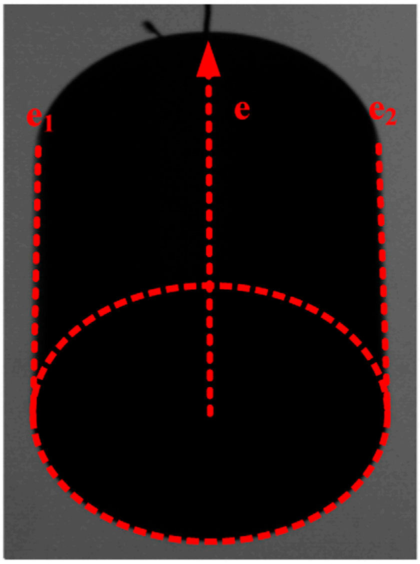

FIGURE 8 | A frame of a rotating cylinder whose axis is coincident with the y-axis of the image coordinate system.

As analyzed in *Principle of Measuring Spin Rate* section, the steady-state spin rate can be achieved by measuring either the rotation rate or the revolution speed of the terminal-sensitive projectile. To measure the rotation rate, the key point is tracking the features, so the surface of the terminal-sensitive projectile should be illuminated, as shown in **Figure 7**. Since the feature tracking algorithms have been widely studied and applied in computer vision, the details will

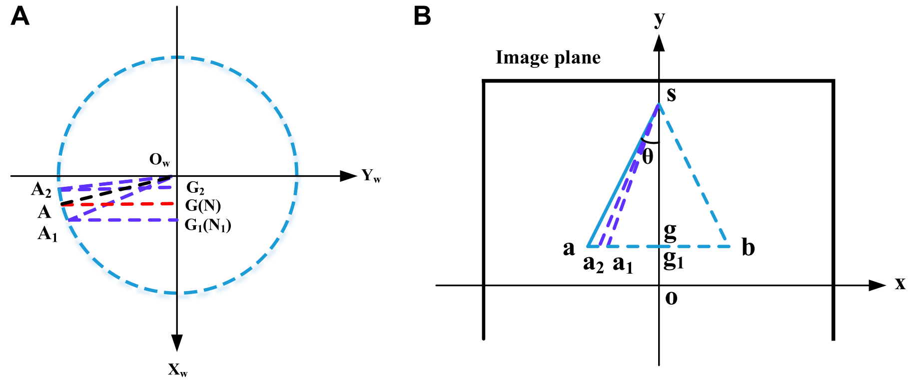

FIGURE 9 | Schematic of measurement error introduced by frame rate. **(A)** Partial view of the plane $O_w - X_wY_w$ in **Figure 2B**. **(B)** View of the trajectory on the image plane.

not be presented here. The disadvantage of measuring rotation rate by tracking features is that the contours of the cylinder in images may be quite vague (**Figure** 7), thus the line-segments would be difficult to distinguish or with a low separate accuracy, which would affect the scanning angle measurement accuracy.

To measure the revolution speed, the key is finding out the frames in which the axis of the projectile is coincident with the y-axis of the image coordinate system. As mentioned before, the axis of the projectile is difficult to fit accurately, hence only the line-segments are employed. As shown in **Figure 8**, due to perspective transformation, the line-segments $\mathbf{e}_1$ and $\mathbf{e}_2$ would be not parallel to each other while the axis $\mathbf{e}$ of the cylinder is coincident with the y-axis of the image coordinate system. To judge correctly, the follow formula is developed

$$\varepsilon = \left|\vec{\mathbf{e}_1}\cdot\vec{i} + \vec{\mathbf{e}_2}\cdot\vec{i}\right|, \tag{11}$$

where $\vec{\mathbf{e}_1}$ and $\vec{\mathbf{e}_2}$ are the normalized vector of line-segment $\mathbf{e}_1$ and $\mathbf{e}_2$, and $\vec{i} = (1, 0)$ is the basic unit vector. For a frame in which the axis $\mathbf{e}$ of the cylinder is coincident with the y-axis of the image coordinate system, the ε would be zero in theory. Whereas, it would be a very small value since there are errors. Thus, find out those frames with minimum ε, then use **Formulas 7**, **8** to compute the real-time and average velocity.

MEASUREMENT ERROR ANALYSIS

Measurement Error of Scanning Angle

There are many factors, such as the error of the focal length f and the pixel sizes dx and dy, frame rate, and contour extraction error, etc., may affect the scanning angle measuring accuracy. Since the subpixel edge extraction operators can reach an accuracy of 1/50 pixel, its influence is omitted. The error of the focal length f and the pixel sizes dx and dy may produce systematic error which can be corrected through some calibration procedure, so they are not discussed here. The frame rate has significant influence on the measurement accuracy, so it is discussed in detail. As shown in **Figure 9A**, a frame should be captured at position A in theory, whereas, it may be missed because there is a time interval between two frames. The maximum error is reached when position A is in a certain position between the two adjacent frames at position A_1 and A_2. Thus, either the 2D coordinate (x_a, y_a) of point a or the angle θ on the image would exist an error which would lead to an error of the scanning angle η according to **Formulas 3**, **6**.

Assuming the frame rate of the high-speed camera is ρ fps, and the spin rate of the terminal-sensitive projectile is ω r/s, then $\angle A_1O_wA + \angle A_2O_wA = \omega \times \pi/\rho$. That is $\angle A_1O_wA < \omega \times \pi/\rho$ and $\angle A_2O_wA < \omega \times \pi/\rho$. For ease of calculation, only the error at point A_1 is calculated and the angle $\angle A_1O_wA$ is set to the upper bound $\omega \times \pi/\rho$, i.e., $\angle A_1O_wA = \omega \times \pi/\rho$.

As shown in **Figures 2B, C**, according to ray trace theory

$$\begin{cases} GN/NO_c = og/f = y_a/f \\ GA/NO_c = ag/f = x_a/f \end{cases}. \tag{12}$$

Draw an auxiliary line G_1N_1 through point G_1 that is perpendicular to the axis Z_c (**Figure 9A**). Since the point G_1 is on the plane $O_w - X_wY_w$, thus $G_1N_1 = GN$. Assuming the projection of point G_1 on the image plane is g_1, the 2D coordinates of point a_1 is (x_{a1}, y_{a1}), and the angles $\angle a_1so$ is θ_1. Then

$$\begin{cases} G_1N_1/N_1O_c = GN/N_1O_c = og_1/f = y_{a1}/f \\ G_1A_1/N_1O_c = a_1g_1/f = x_{a1}/f \end{cases}. \tag{13}$$

According to **Figure 9A**, $NN_1 \ll NO_c$, hence $N_1O_c = NO_c - NN_1 \approx NO_c$, then

$$\begin{cases} y_{a1} = f \times GN/N_1O_c \approx f \times GN/NO_c = y_a \\ x_{a1} = G_1A_1 \times NO_c \times x_a/(GA \times N_1O_c) \approx (G_1A_1/GA) \times x_a \end{cases}. \tag{14}$$

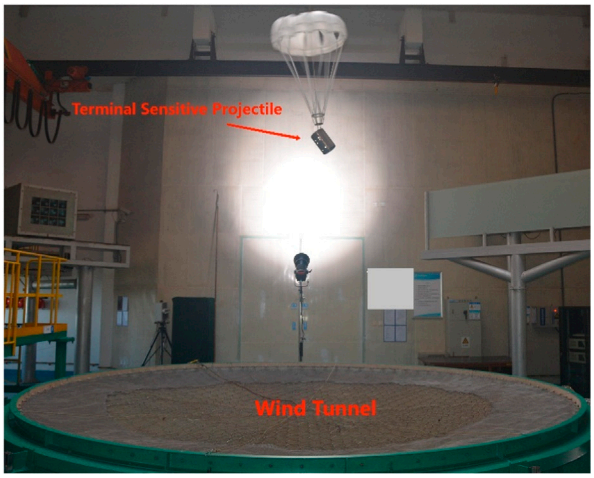

FIGURE 10 | Experiment setup used to test the attitude measurement algorithm in laboratory.

Moreover, according to the Pythagorean theorem,

$$\begin{cases} G_1A_1/GA = \sin\left(\angle AO_wG - \angle A_1O_wA\right)/\sin\angle AO_wGl \\ \angle AO_wG = \sin^{-1}\left(x_a/\left(\tan\beta\times\tan\eta\times\sqrt{x_a^2+y_a^2+f^2}\right)\right) \end{cases}. \tag{15}$$

Therefore, the 2D coordinate (x_{a1}, y_{a1}) of point a_1 can be computed exactly.

Once the 2D coordinate of point a_1 is solved, the angle θ_1 can be calculated according to the Pythagorean theorem as following (**Figure 9B**)

$$\theta_1 = \tan^{-1}\left(x_{a1}\times\tan\theta/x_a\right). \tag{16}$$

Finally, a new scanning angle η_1 can be calculated by inputting the parameters (x_{a1}, y_{a1}), θ_1, f, dx, and dy into **Formulas 3**, **6**, and the measuring error thereby can be obtained as

$$\Delta\eta = \left|\eta_1 - \eta\right|. \tag{17}$$

Measurement Error of Spin Rate

The frame rate and the time measurement error of the high-speed camera, and contour extraction error, etc., may affect the spin rate measuring accuracy. Among them, the frame rate is the major influence factor. As shown in **Figure 3**, a frame should be captured at position g where the axis of the projectile sg is coincident with the y-axis of the image coordinate system in theory, whereas, it may be missed because there is a time interval between two frames. The maximum error is just a frame. Therefore, the spin rate measurement error is

$$\Delta v = \left|w - \rho\times\omega/(\rho-\omega)\right| = \omega^2/(\rho-\omega), \tag{18}$$

FIGURE 11 | High-speed CMOS camera employed in the experiment.

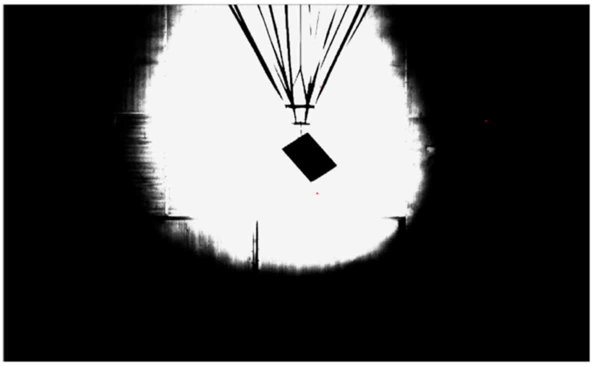

FIGURE 12 | A frame of the rotating terminal-sensitive projectile.

EXPERIMENT

Experiments were conducted in a laboratory to evaluate the proposed method for image segmentation. The experimental setup is illustrated in **Figure 10**. In the experiment, a terminal-sensitive projectile was suspended in the air using a flexible rope that passed over a pulley and rotated around a plumb axis. The projectile maintained a steady spin at a rate of approximately 3.5 revolutions per second. To ensure adequate illumination, a LED light source was employed. For image recording, a high-speed CMOS camera, specifically the MIKROTRON EoSens 3CXP as shown in **Figure 11**, was used. The camera has a maximum resolution of 1280 × 1024 pixels, with a pixel size of 8 × 8 μm. The focal length of the camera lens was set to 35 mm, and the camera operated at a frame rate of 1,000 frames per second.

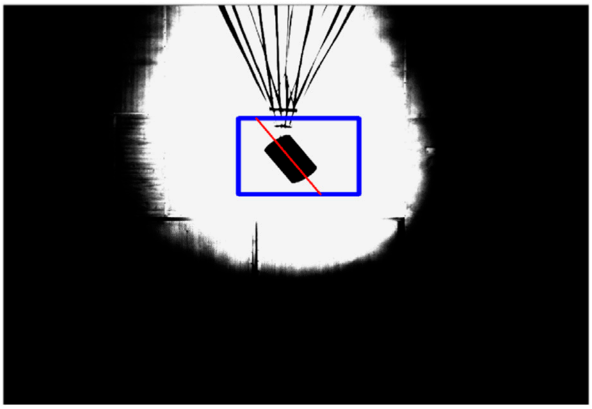

FIGURE 13 | Fitting the axis of the main body of the terminal-sensitive projectile.

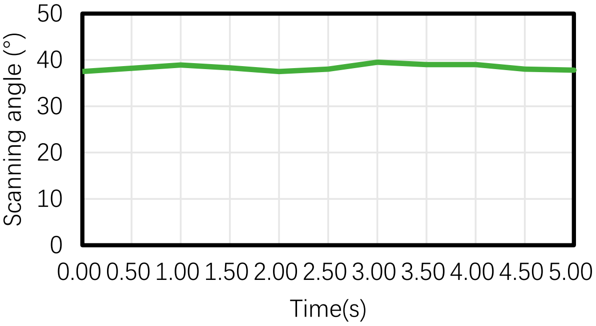

FIGURE 14 | Scanning angle of the rotating terminal-sensitive projectile.

During the experiment, two lighting schemes were tested to assess their impact on the quality of the captured images. The first scheme involved illuminating the main body of the terminal-sensitive projectile, while the second scheme involved illuminating the surrounding wall. The results showed that the latter scheme produced images with clearer edges compared to the former, which resulted in images with blurred edges. Therefore, the lighting scheme that illuminated the wall was ultimately chosen for further experimentation. The measurement process described in *Measurement Process* section was followed during the experiment. When the rotating terminal-sensitive projectile reached a steady state, the camera recorded a total of 5,000 images over a duration of 5 s. **Figure 12** displays a frame captured from the rotating terminal-sensitive projectile, providing a visual representation of the captured image. Additionally, **Figure 13** showcases the extracted axis of the main body of the terminal-sensitive projectile, which was obtained through the image processing and segmentation techniques discussed in the paper. These figures serve to illustrate the effectiveness of the proposed method in capturing and extracting relevant information from images of the rotating terminal-sensitive projectile. The clear edges in the images obtained using the selected lighting scheme contribute to accurate segmentation and subsequent analysis of the terminal-sensitive projectile's main body.

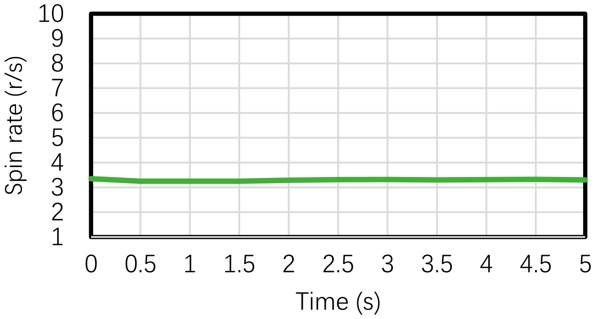

FIGURE 15 | Spin rate of the rotating terminal-sensitive projectile.

The measured scanning angles are depicted in **Figure 14**. The minimum scanning angle recorded was 37.5°, the maximum scanning angle was 39.2°, and the average scanning angle was 38.3°. The measurement error associated with the scanning angle was found to be less than 0.67°. Similarly, the measured spin rates are shown in **Figure 15**. The minimum spin rate observed during the experiment was 3.25 revolutions per second (r/s), the maximum spin rate was 3.35 r/s, and the average spin rate was 3.30 r/s. The measurement error for the spin rate was determined to be less than 0.05 r/s. These measurement results highlight the high accuracy achieved by the novel algorithm proposed in the paper. The algorithm effectively captures and analyzes the rotational motion of the terminal-sensitive projectile, providing precise measurements of both the scanning angles and spin rates. The low measurement errors indicate the reliability and robustness of the algorithm in accurately quantifying these parameters.

CONCLUSION

Optical measurement instruments/systems are widely used in modern weapon test. They usually have higher accuracy than other sensors in spin rate measurement, whereas lack of precision in scanning angle measurement. In this paper, an optical attitude measurement method for terminal sensitive projectile in vertical wind tunnel test is presented in detail. The optical measurement method is extremely simple since only one high-speed camera is employed. However, both scanning angle and spin rate can be measured accurately by using the novel method.

Compared with existing optical attitude measurement methods, the contribution of the novel method can be analyzed in two aspects: first, the mathematical model for calculating scanning angle of a steady-state terminal sensitive projectile is derived, which makes it possible to measure scanning angle from images recorded by only a fixed camera; second, the algorithms of estimating scanning angle and spin rate depend only on the line-segments of the perspective cylinders in images are designed, which significantly improve the measurement accuracy. The perspective cylinder, which is projection of the terminal sensitive projectile on image plane, is generally consist of two line-segments and two elliptic arcs. It is well known that the ellipse fitting may be inaccurate only rely on the contour points belong to the elliptic arcs. Therefore, only line-segments are adopted to compute both scanning angles and spin rate accurately.

The shortcoming of this paper is that the error of the focal length f and the pixel sizes dx and dy are not considered. In fact, the error of those parameters may produce considerable measurement error. Hence, they should be calibrated in advance in practical application.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

XY, CG, ZZ, and HW conducted the experiments, XY, YW, and YL wrote the manuscript. All authors contributed to the article and approved the submitted version.

CONFLICT OF INTEREST

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

REFERENCES

1. Yakimenko OA, Kolsch MN, Decker RJ. *Method and Apparatus for Computer Vision Analysis of Spin Rate of Marked Projectiles*. U.S. Patent 9911046 (2018). p. 3–6.
2. Shang J, Deng Z, Fu M, Wang ST. Advance and Perspective on Spin Rate Measurement Technology for Guided Projectile. *Acta Automatica Sinica* (2016) 42(11):1620–9.
3. Schuler AR, Grammatikos A, Fegley KA. Measuring Rotational Motion With Linear Accelerometers. *IEEE Trans aerospace Electron Syst* (1967)(3) 465–72. doi:10.1109/taes.1967.5408811
4. Lu YL, Pan YJ, Li L, Liu Y, Peng H. Measurement Method of Projectile's Heading and Pitching Angle Velocities Based on Biaxial Accelerometer. *J Chin Inertial Technol* (2015) 23:160–4.
5. Stroezel M, Saaro J, Stieler B. *Wind Tunnel Model Attitude Measurement With a Fiberoptic Gyro*. Stuttgart, Germany: Symposium Gyro Technology 1995 (1995). p. 14.
6. Liu N, Su Z, Li Q. Design and Experiment of a Novel Bellshaped Vibratory Gyro. *Sensors Actuators A: Phys* (2016) 238:37–50. doi:10.1016/j.sna.2015.10.036
7. Lin Z, Fu MY, Deng ZH, Liu N, Liu H. Frequency Split Elimination Method for a Solid-State Vibratory Angular Rate Gyro With an Imperfect Axisymmetric-Shell Resonator. *Sensors* (2015) 15(2):3204–23. doi:10.3390/s150203204
8. Zhou L, Liu R, Guo R. Attitude Calculation Method of Terminal Sensitive Projectile Based on Geomagnetic Sensor. *J Ordnance Equipment Eng* (2018) 3: 174–7.
9. Deng Z, Shen Q, Deng Z, Cheng J. Real-Time Estimation for Roll Angle of Spinning Projectile Based on Phase-Locked Loop on Signals From Single-Axis Magnetometer. *Sensors* (2019) 19(4):839. doi:10.3390/s19040839
10. Shang J, Deng Z, Fu M, Wang S. A High-Spin Rate Measurement Method for Projectiles Using a Magnetoresistive Sensor Based on Time-Frequency Domain Analysis. *Sensors* (2016) 16(6):894. doi:10.3390/s16060894
11. Yanning G, Yan Y. Telemetering for the Attitude of Projectile in Flight Based on the Principle of Solar Aspect Angle. *Acta Armamentarii* (2003) 24(2):250–2.
12. Changey S, Pecheur E, Brunner T. Attitude Estimation of a Projectile Using Magnetometers and Accelerometers: Experimental Validation. In: Proceedings of the IEEE/ION PLANS 2014. Monterey, CA. May 5 - 8, 2014. IEEE (2014). p. 1168–73.
13. Harkins T, Davis B, Hepner D. Novel Onboard Sensor Systems for Making Angular Measurements on Spinning Projectiles. Acquisition, Tracking, and Pointing XV. *Int Soc Opt Photon* (2001) 4365:176–87.
14. Hongsong C, Shunshan F. Researching Ammunition Attitude Detect Technique Combination of Geomagnetism and Gyro. *J Projectiles, Rockets, Missiles Guidance* (2006) 26(3):142–5.
15. Huang X, Wang CH, Yi GX, Wang YF. Extended Kalman Filter for IMU Attitude Estimation Using Magnetometer, MEMS Accelerometer and Gyroscope. *J Chin Inertial Tech (China)* (2005) 13(2):27–30.
16. Combettes C, Changey S, Adam R, Pecheur E. Attitude and Velocity Estimation of a Projectile Using Low Cost Magnetometers and Accelerometers. In: 2018 IEEE/ION Position, Location and Navigation Symposium (PLANS). 23-26 April 2018. Monterey, CA, USA. IEEE(2018). p. 650–7.
17. Yan A, Chen X, Zhou C, Zou HT. Attitude Measurement Based on Magnetometer and MEMS Gyroscope. *J Projectiles, Rockets, Missiles Guidance* (2016) 32(6):143–6.
18. DengShen ZQ, Deng Z. Roll Angle Measurement for a Spinning Vehicle Based on GPS Signals Received by a Single-Patch Antenna. *Sensors* (2018) 18:3479. doi:10.3390/s18103479
19. Cao P, Yu JY, Wang XM, Yao WJ, Wu YL. High-frequency Measurement and Calculation Study of Systematic Errors of High-Rolling Projectile Roll Angle Based on a Combination of MR/GNSS. *Acta Armamentarii* (2014) 35:795–800.
20. Ferguson EM, Bossoli RB, Jara EA. *Technique for Measuring the Spin Rate of Kinetic Energy Projectiles*. Adelphi: Technical Report of US Army Research Laboratory (1993). No. ARL-MR-60.
21. Li Y, Pang W, Wang Y, Wang F, Min BR. Measuring Method of Rotation Velocity for Small Arms Projectiles by Doppler Radar. *Chin J Radio Sci* (2007) 22(3):502–7.
22. Gao H, Zhao Y, Fang Y. Application of Vertical Wind Tunnel in Developing Sensor Fused Munitions. *J Projectiles, Rockets, Missiles Guidance* (2013) 33(5): 138–40.
23. Zhao Z, Wen G, Zhang X, Li D. Model-Based Estimation for Pose, Velocity of Projectile From Stereo Linear Array Image. *Meas Sci Rev* (2012) 12(3):104–10. doi:10.2478/v10048-012-0013-x
24. Yang K, Cai Z, Zhao L. Algorithm Research on Moving Object Detection of Surveillance Video Sequence. *Opt Photon J* (2013) 3(2):308–12. doi:10.4236/opj.2013.32b072
25. Elgammal A, Duraiswami R, Harwood D, Davis LS. Background and Foreground Modeling Using Nonparametric Kernel Density for Visual Surveillance. *Proc IEEE* (2002) 90: 1151–63.
26. Agarwal A, Gupta S, Singh DK. Review of Optical Flow Technique for Moving Object Detection. In: 2nd International Conference on Contemporary Computing and Informatics (IC3I). IEEE (2016). p. 409–13.
27. Redmon J, Divvala S, Girshick R, Farhadi A. You Only Look Once: Unified, Real-Time Object Detection. In: Computer Vision Pattern Recognition. IEEE (2016) 779–88.
28. Steger C. Subpixel-Precise Extraction of Lines and Edges. *Int Arch photogrammetry remote sensing* (2000) 33(3):141–56.
29. Huang JB, Chen Z, Chia TL. Pose Determination of a Cylinder Using Reprojection Transformation. *Pattern Recognition Lett* (1996) 17(10): 1089–99. doi:10.1016/0167-8655(96)00061-x
30. Martin AF, Robert CB. Random Sample Consensus: A Paradigm for Model Fitting With Applications to Image Analysis and Automated Cartography. *Comm ACM* (1981) 24:381–95.

AEROSPACE RESEARCH COMMUNICATIONS
REVIEW
published: 19 December 2023
doi: 10.3389/arc.2023.12394

State-Space Approaches to Complex Structures in Aerospace

Zinan Zhao[1,2], *Xudong Shen*[1], *Yipin Su*[1,2] *and Weiqiu Chen*[1,2,3]*

[1]Center for Soft Machines and Smart Devices, Huanjiang Laboratory, Zhuji, China, [2]Key Laboratory of Soft Machines and Smart Devices of Zhejiang Province, Department of Engineering Mechanics, Soft Matter Research Center, Zhejiang University, Hangzhou, China, [3]Faculty of Mechanical Engineering and Mechanics, Ningbo University, Ningbo, China

The state-space approach (SSA), traditionally utilized in modern control theory, has been successfully adopted over the last three decades to investigate the mechanical behaviors of complex structures composed of composite or smart materials. This is largely due to their increasing application across various fields, including aerospace, civil and marine engineering, and transportation vehicles. This paper provides a comprehensive review of the establishment of state-space formulations for structures of typical configurations, such as beams, plates, shells, and trusses, with a particular focus on their applications in the mechanical analyses of various complex aerospace or smart structures using the transfer matrix method. The paper first summarizes the three-dimensional SSAs applied to laminated structures without any assumptions on physical fields. By employing structural theories such as various beam, plate, and shell theories, simplified one-dimensional and two-dimensional SSAs for laminated structures are developed. The paper then outlines the advances in generating analytical solutions for the mechanical behaviors of laminated structures. For the sake of completeness, the paper also provides an account of SSAs applied to complex periodic structures, particularly in beam and truss forms. To overcome the limitations of conventional SSAs related to structures with specialized geometric configuration or under arbitrary boundary conditions, state-space based numerical methods have been proposed, for example, the state-space based differential quadrature method and state-space based finite-element method. The applications of these methods in the analyses of static and dynamic responses of complex structures are extensively reviewed. It is observed that there are still intriguing and potential research topics for the development of advanced SSAs with enhanced versatility and the studies on practical complex structures used in modern engineering, particularly in aerospace industry. Therefore, this review is expected to be beneficial for researchers in the fields of analytical and numerical methods, composite structures, aerospace, structural engineering, and more.

Keywords: state-space approach, complex structure, aerospace, structural theory, mechanical behavior, analytical and numerical methods

***Correspondence:**
Weiqiu Chen
chenwq@zju.edu.cn

Received: *08 November 2023*
Accepted: *30 November 2023*
Published: *19 December 2023*

Citation:
Zhao Z, Shen X, Su Y and Chen W (2023) State-Space Approaches to Complex Structures in Aerospace. Aerosp. Res. Commun. 1:12394. doi: 10.3389/arc.2023.12394

INTRODUCTION

Composite materials typically consist of two or more distinct components, each possessing unique mechanical, electric, magnetic, thermal, or chemical properties. As a result, structures composed of these composite materials, such as laminated structures, often exhibit superior performance characteristics compared to conventional structures made from homogeneous materials. These advantages include lightweight construction, high strength-to-weight and stiffness-to-weight ratios, excellent tailoring properties, and high resistant to corrosion and fatigue [1, 2]. Spacecraft and aircraft are typical weight- and strength-sensitive structures. According to a widely accepted rule of thumb, each pound of weight reduction in a primary structure results in an indirect saving of another pound elsewhere in the aircraft [3]. Today, structures composed of composite materials are increasingly becoming a cornerstone in modern aerospace and aviation industry [4, 5]. Moreover, these complex structures are also finding broader applications in fields such as mechanical and civil engineering, marine, transportation vehicles [6, 7]. Consequently, the burgeoning industrial uses of complex structures have spurred the development and innovation of necessary analytical and numerical techniques for analyzing the mechanical behaviors of these structures [8, 9].

The mechanics of three-dimensional (3D) continua are typically governed by partial differential equations (PDEs), wherein all physical fields, including the displacements, stresses, and strains, are functions of the three spatial coordinate variables and/or the time variable. Consequently, deriving exact analytical solutions for 3D elasticity structures can be challenging. To address this, various structural theories such as rod, beam, plate, and shell theories have been developed, introducing certain assumptions on physical fields regarding their space distributions. These theories have been widely used to study the mechanical behaviors of numerous engineering structures composed of homogeneously isotropic, transversely isotropic, or orthotropic materials. However, these theories sometimes fall short or face significant limitations when predicting the static and dynamic responses of structures composed of composite or smart materials, especially accurate distributions of displacements and stresses at interlayer interfaces. This is due to factors such as strong material anisotropy, multi-field coupling, heterogeneity, and structural complexity. As a result, conventional structural theories often lack the necessary generality and accuracy when studying complex structures.

An alternative strategy involves seeking solutions for complex structures directly based on analytical or numerical methods such as finite element methods (FEMs), displacement and stress functions, Taylor series expansions, and state-space approaches (SSAs), which are the focus of this article. SSAs, traditionally used for modeling control systems, describe the relationships between the output, input, and the state of the system. These three variables are only functions of the time. The state equation of the system is governed by first-order ordinary differential equations (ODEs) with respect to the state and input variables, and the output variables of the system can be fully determined when the state variables are known. By using state-space descriptions, the first-order ODEs of the system are transformed into a matrix expression, thereby enabling efficient utilization of high-performance computer facilities using engineering software such as MATLAB and Mathematica. Furthermore, due to the matrix expression of the system equation, the increase of state variables does not increase the complexity of system descriptions. Thus, SSAs offer better unity of expression and higher computational efficiency than other methods for modeling control systems.

Unlike the first-order ODEs of control systems, the governing equations of a 3D elastic body are PDEs with respect to the selected physical fields such as displacements, stresses, or strains. Therefore, the key to establishing the state-space formulations for elasticity problems lies in selecting the appropriate state variables and transforming the PDEs into a set of first-order ODEs. The practical application of the SSA to elastic bodies dates back to the study on vibrations of non-uniform beams by Thomson in 1950 [10]. By selecting the deflection, rotation, shear force, and bending moment as the state variables, they reformulated the governing equations, the fourth-order ODEs, of the beam deflection as a set of first-order ODEs based on the beam theory. For the non-uniform beam, the layerwise approximate technique (LAT) was employed by them to divide the beam into many segments such that the section of each segment could be treated as constant. Based on state-space descriptions, the continuity conditions at the interfaces and the boundary conditions at the two ends of the beam can be easily expressed by the state variables. Then a transfer relation between the state variables at the two ends of every segment can be established, and a global transfer matrix between the state variables at the two ends of the beam can be finally derived by eliminating all interfacial displacements and tractions based on interlayer continuity conditions. With the increase of the number of the divided segments, the approximate solutions obtained by the SSA approach the exact solutions for free vibration of the non-uniform beam. Evidently, in the SSA the dimensionality of the global transfer matrix of laminated structures is independent of the number and thickness of the layers, and thus the SSA is very powerful and efficient for seeking exact solutions of laminated structures no matter how many layers there are. Furthermore, the displacement and stress components at interlayer interfaces predicted by the SSA are accurately continuous, which cannot be realized by other numerical method like the FEM. Almost simultaneously, Thomson [11] also demonstrated that the SSA could be used for wave propagation analysis in an infinite stratified solid medium. By selecting the particle velocity and stresses as the state variables, the governing equations of the plane wave propagation were transformed into a matrix form, i.e., the state-space formulations for wave motions in a stratified medium.

Building on Thomson's pioneer work and the increasing use of laminated composite structures, there has been a significant development and application of SSAs in fields such as aerospace, geophysics [12], and structural and civil engineering [6] since 1990s. The 3D state-space formulations for materials with varying material symmetries and multi-field coupling effects, such as transversely isotropic, spherically isotropic, piezoelectric, and magneto-electro-thermo-elastic materials, have all been established for the studies of structures of different geometric configurations and with multi-field couplings. Apart from the general 3D problems, since no assumptions are adopted to simplify the physical fields for the states of plane strain or plane stress, the SSAs applied to solve plane-strain and plane-stress problems are also ascribed to the 3D SSAs as special cases in this article. Leveraging sophisticated structural theories, simplified one-dimensional (1D) and two-dimensional (2D) SSAs have also been developed for the analyses of laminated structures in beam, plate, and shell forms. The mixed SSAs, integrating the transfer matrix method for state vectors and theoretical or numerical methods, have been further

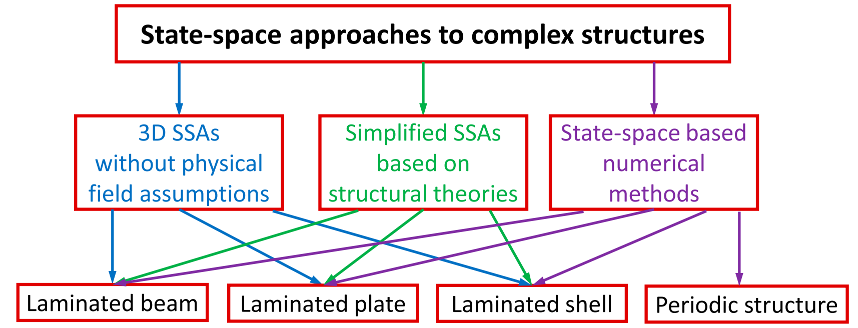

FIGURE 1 | Overall structural diagram of state-space approaches applied to complex structures.

developed to predict the mechanical behaviors of complex periodic structures composed of beam or truss units. In addition, various state-space based numerical methods, combining the advantages of SSAs and the associated numerical methods, have been proposed to solve more practical structural problems or address numerical difficulties associated with structures under arbitrary boundary conditions. For instance, in the state-space based differential quadrature method (SS-DQM), the DQM helps to overcome the limitations associated with conventional SSAs for plates and shells with arbitrary boundary conditions, while still retaining the feature of the SSA that the displacement and stress components are accurately continuous at interlayer interfaces. This feature is also inherited by the state-space based finite element method (SS-FEM), making the number of the unknow functions and unknow node variables independent of the layer number of laminated structures, which gives the SS-FEM an edge over conventional FEMs. Other specialized state-space based numerical methods, including the state-space based finite difference method (SS-FDM), state-space based boundary element method (SS-BEM), etc., have also been proposed to deal with structures with certain geometric configurations or to improve the computational efficiency. The overall structural diagram of SSAs applied to complex structures is illustrated in **Figure 1**.

This review aims to highlight the advances of SSAs to structures of typical configurations and their applications in the analyses of mechanical behaviors of complex structures composed of composite or smart materials. We first focus on the 3D SSAs for laminated structures of beam, plate, and shell configurations, in which no assumptions of physical fields are adopted. When structural theories, such as beam, plate, and shell theories, are employed, we present the establishment of 1D and 2D state-space formulations and their extensive applications in the mechanical analyses of laminated structures. We then provide an account of the applications of SSAs in studies on complex periodic structures composed of beam and truss units. The final focus of this article is the advances of state-space based numerical methods and their applications in static and dynamic analyses of complex structures. This review is intended to be beneficial for future studies on advanced SSAs and mechanical behaviors of practical complex structures in various fields, particularly for aeronautical, marine, civil, and transportation applications.

3D STATE-SPACE APPROACHES TO LAMINATED STRUCTURES

As previously mentioned, laminated structures have been found extensive applications in fields such as aerospace, aviation, structural engineering, and marine. Despite the flexibility offered by material and structural design and the ability to be composed of numerous layers for carrying large loads, laminated structures often grapple with issues such as the ply delamination and shear destruction due to the low ratio of transverse shear modulus to in-plane modulus. Since 1950s, 3D SSAs, which do not introduce any assumptions on physical fields, have been widely adopted to accurately predict mechanical behaviors of laminated structures. These approaches also provide useful benchmark results for comparing and clarifying various simplified theories or numerical methods. In this section, we will review the advances of 3D SSAs and their applications in analyzing static and dynamic responses of laminated structures of beam, plate, and shell configurations, as summarized in **Table 1**.

Laminated Beams

Beams, often utilized to represent slender structures with a length significantly greater than their cross-sectional dimensions, are a fundamental component in aeronautical applications. These applications include, but are not limited to, fuselage air wings, helicopter rotor blades, and turbine rotor blades. Laminated beams, in particular, are of paramount importance. This section reviews the research on 3D SSAs for beam configuration as well as their application in the accurate 3D analyses of both static and dynamic responses of laminated beams.

The SSA to elasticity, initially proposed by Bahar for the analysis of 2D plane stress problem in linear elasticity [13], has proven to be a potent and effective tool for deriving exact solutions of elasticity. Utilizing the state-space formulations established for plane-stress problem, Chen et al. [14] presented an exact solution for simply-supported cross-ply laminated

TABLE 1 | Application of 3D state-space approaches in the mechanical behavior analyses of complex structures of different configurations.

Configuration	Mechanical behavior	Reference
Beam	Bending and vibration	[14–16]
Beam	Bending	[6, 17–22]
Beam	Vibration	[23]
Plate	Bending and vibration	[24–29]
Plate	Bending	[30–48]
Plate	Vibration	[49–54]
Cylindrical shell	Bending, vibration, buckling	[55–57]
Cylindrical shell	Bending, vibration	[58–60]
Cylindrical shell	Bending	[61]
Cylindrical shell	Vibration	[62–68]
Cylindrical shell	Buckling	[69]
Cylindrical shell	Wave	[70–72]
Spherical shell	Bending	[73, 74]
Spherical shell	Vibration	[75–77]

beams with interlaminar bonding imperfections, described by a spring-layer model, as shown in **Figure 2A**. It can be seen from **Figure 2B** that the transverse shear stress τ_{xz} predicted by the 3D SSA for $R = 0$ was accurately continuous at the interfaces and the imperfect bonding significantly changed its distribution along the thickness direction. Moreover, numerical results indicated that an increase in the imperfection bonding parameter could lead to an increase in deflection and a decrease in the natural frequency, suggesting that interfacial imperfections result in a decrease in the effective stiffness of the laminated beam. The solutions obtained can serve as a valuable benchmark for simplified beam theories or numerical methods proposed for imperfect laminated beams. Chen et al. [6] also employed the SSA to analyze the static responses of laminated composite beam with partial shear interaction, a common structure in civil engineering, subjected to a combination of an arbitrary transverse load and a constant

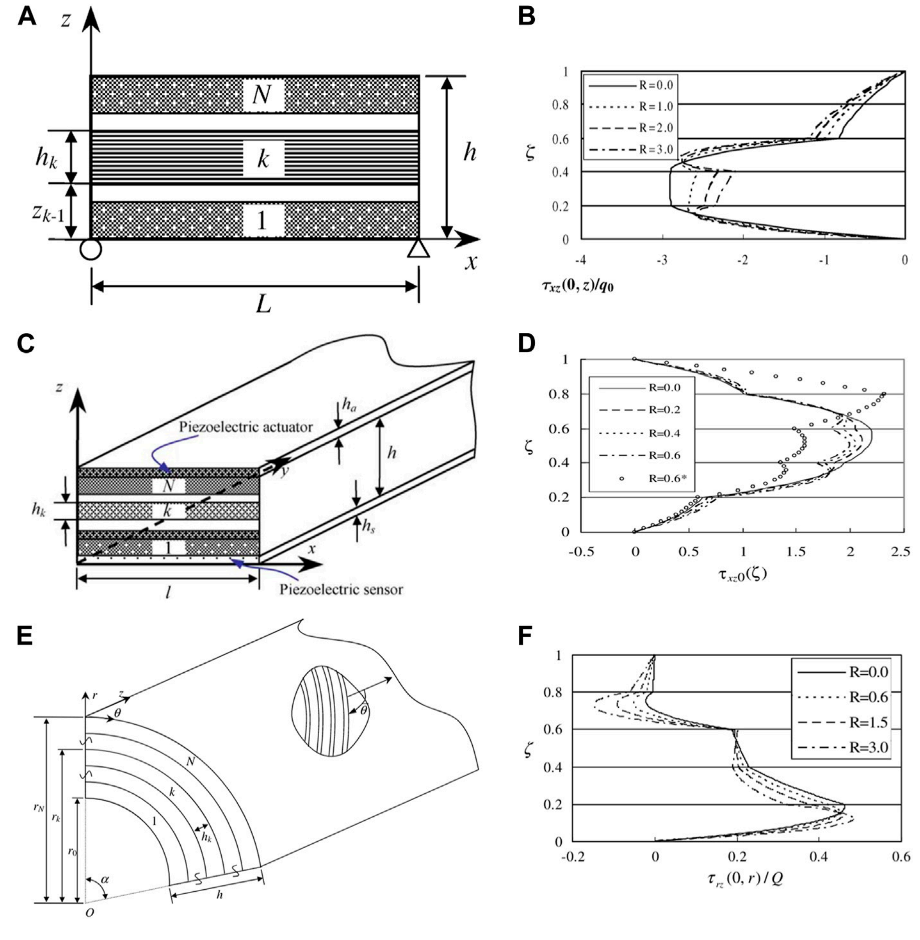

FIGURE 2 | Laminated structures of beam, plate, and shell configurations and their shear stress field predicted by the 3D state-space approach: **(A)** laminated beam with bonding imperfection [14]; **(B)** the distribution of transverse shear stress τ_{xz} of laminated beam in **(A)** along the thickness direction under different imperfect bonding parameter R; **(C)** imperfect laminated plate in cylindrical bending with surface piezoelectric layer [29]; **(D)** the distribution of transverse shear stress τ_{xz} of laminated plate in **(C)** along the thickness direction under different imperfect bonding parameter R; **(E)** laminated cylindrical panel with bonding imperfection [60]; **(F)** the distribution of shear stress τ_{rz} of laminated panel in **(E)** along the thickness direction under different imperfect bonding parameter R (Reproduced with permission from Chen et al. [14], copyright 2003 by AIAA; reproduced with permission from Chen et al. [29], copyright 2004 by Elsevier; reproduced with permission from Cai et al. [60], copyright 2004 by Elsevier).

axial force. To account for the interfacial slip, a non-continuous model of slip stiffness along the interface was integrated into the SSA. Numerical results suggested that the non-uniformity of the slip stiffness significantly influences the critical axial load and interactions of the composite beam. Furthermore, Xu and Wu [17] developed a new plane stress model of partial interaction composite beam with interlayer slips based on the SSA and presented analytical solutions for static responses of simply-supported beams. It was found that the 2D model predicts deflections more accurately than the 1D theory, which neglects the shear deformation of the beam. Additionally, the rigidity of the shear connectors for partial interaction has a substantial effect on the flexural stiffness of composite beams. Subsequently, they conducted an analytical study of composite beams strengthened by externally and adhesively bonded reinforcements, such as fiber-reinforced polymer (FRP) or steel with variable cross-sectional properties [18]. The interfacial shear stress and axial force in external reinforcement were particularly investigated, and it was found that the bonded FRP with tapered ends and the ways of tapering the FRP significantly alter the magnitude and distribution of the shear stress at the interfaces. The developed method provided an efficient tool for analyses of the debonding problem and design of external reinforcement of composite beams. Ying et al. [15] applied the SSA to study static bending and free vibration of functionally graded material (FGM) beams resting on a Winkler-Pasternak elastic foundation based on 2D plane-stress theory of elasticity. They found that the gradient index, aspect ratio, and foundation parameter significantly influence the deformation and natural frequencies of such an FGM beam, and these exact solutions could serve as a benchmark for the future studies on FGM beams resting on elastic foundations.

Unlike conventional laminated structures composed of elastic materials, smart structures, which integrate host laminated structures with smart components such as piezoelectric actuators and sensors, can react to environmental changes through actuating, sensing, and control [78–80]. By utilizing the SSA and the LAT, Bian et al. [16] conducted an exact analysis of the static and dynamic responses of simply-supported FGM beams integrated with surface piezoelectric actuator and sensor. The influence of the bonding adhesives between the host beam and piezoelectric actuator and sensor layers on the displacements, stresses, and natural frequency was systematically examined. The results demonstrated that imperfect bonding reduces the global stiffness of the entire beam structure, as evidenced by the increase in deflection and the decrease in natural frequency. Yan et al. [19] explored the time-dependent responses of this smart structure with viscoelastic interfaces, described by the Kelvin-Voigt viscoelastic model, by using the SSA and power series expansion. They discovered that the viscoelastic interfaces also contribute to the reduction of the effective stiffness of the entire structure, and the developed method can be applied for damage analyses in health monitoring of aeronautical and engineering structures. Furthermore, Yan et al. [23] investigated the dynamic responses of a simply-supported laminated beam bonded or embedded with piezoelectric layers based on the SSA. The electromechanical impedance (EMI) technique was employed to monitor interfacial defects, described by a linear spring-like model, in the laminated beam. Numerical results revealed that the output electric signal is sensitive to the bonding condition between the piezoelectric layer and the host beam, suggesting that monitoring the EMI can be used to conveniently detect the interfacial defects. In addition, Alibeigloo [20] provided the analytical solutions for simply-supported FGM beams integrated with piezoelectric actuator and sensor under an external electric load and thermo-mechanical load, based on the state-space formulations and Fourier series expansion. They found that the effect of thermal load on mechanical behaviors is greater than that of mechanical load, and the effect of applied voltage is strongly dependent on the length-to-thickness ratio. Using the SSA and Fourier series expansion, Qian et al. [21, 22] proposed an exact analytical method to predict thermal responses of simply-supported and clamped laminated beams subjected to non-uniform temperature boundary conditions. For the clamped support, an equivalent model was introduced to convert the clamped support to a simple support. Numerical simulations indicated that the surface temperature, length-to-thickness ratio, material properties, and the layer number of laminated beam significantly affect the distributions of the temperature, displacements, and stresses. Moreover, the comparison of the results with those from the FEM verified the accuracy and effectiveness of the proposed method.

Laminated Plates

Compared to beam configuration, the mechanical behaviors of laminated plate structures have garnered more attention from researchers. Since 1969, a variety of exact solutions for laminated composite plates have been derived using Pagano's approach [81–88]. However, this method necessitates the discussion of different eigenvalue cases to obtain the real solutions of the governing equations. In contrast to Pagano's approach, the SSA offers a uniform matrix expression and serves as a powerful and efficient tool for exact analyses of laminated plates, given that the number of the undetermined constants is independent of the number of structural layers. This subsection summarizes the applications of 3D SSAs in the studies on the mechanical behaviors of laminated plates.

Fan and Ye [30] utilized the 3D theory of elasticity to derive the state equations of an orthotropic elastic body and then presented a series solution of a simply-supported rectangular plate under various mechanical loads. Their numerical results indicated significant deviations between existing plate theories, such as Reissner's and Ambartsumyan's theories, and their exact results. Subsequently, they presented exact solutions for the static and dynamic responses of three-ply orthotropic thick plates with simply-supported edges [24]. Numerical simulations demonstrated that the SSA surpasses the conventional displacement method of elasticity for laminated plates, as the latter increases the number of equations and unknown constants with the increase of the plies. They also established the state-space formulations for the exact analyses of the axisymmetric free vibrations of transversely isotropic three-ply circular plates with simply-supported and clamped edges [49]. Comparisons with the Reissner and Mindlin theories revealed the insufficient accuracy of the existing plate theories. Zenkour [31] introduced an exact and unified tool for the symmetric and antisymmetric analyses of rectangular laminated plates subjected to sinusoidally/uniformly distributed loads based on the SSA. The results were compared with those in Refs [82, 84], demonstrating the computational efficiency and accuracy of the SSA. Qing et al. [50] presented a new solution governed by the state-space formulations for the dynamic analyses of 3D elastic laminated plates with damping effect, using the precise integration method and Muller method. They examined the effects of damping parameters on natural frequency, harmonic vibration, and complex frequency response, which showed the predictive capability of their approach.

In contrast to the perfect bonding in the aforementioned works, Chen et al. [25] investigated the influence of interlaminar bonding imperfections, described by a linear spring layer model, on the bending and free vibration of a simply-supported cross-ply laminated rectangular plate using the SSA. Their numerical results were compared with those predicted by the plate theory developed in Cheng et al. [89], indicating that the approximate theory performs well for moderately thick laminates with perfect bonding, but is inaccurate for imperfect bonding. Chen et al. [32, 33] then use the SSA to analyze the bending of a simply-supported angle-ply laminated plate and a thick orthotropic laminate with viscous interfaces under static load, respectively. They found that the static response of the laminate is highly sensitive to the presence of viscous interfaces. They also investigated the cylindrical bending and free vibration of a simply-supported angle-ply laminate with interfacial damage, described by the general spring-layer model, using the SSA [26]. Numerical results suggested that the deformation and frequency's sensitivity to interfacial damage could be used in nondestructive testing of structural damage. Qing et al. [34] used the SSA along with interfacial spring-layer model to study the energy release ratio of stiffened laminates with a planar delamination. They discussed the effect of the stiffeners on the distribution of the energy release modes, and the results revealed that their approach is a powerful and accurate alternative for solving the multiple delamination problems of laminated structures.

In addition to the previously mentioned studies on laminated plates composed of elastic materials, a series of theoretical research have been conducted on the mechanical behaviors of smart laminated plate structures. These structures combine host laminated plates with intelligent material layers to modify the structural responses through sensing, actuation, and control. For instance, Xu et al. [35] utilized the 3D SSA and Fourier series expansion to examine the coupled thermo-electro-elastic response of a smart laminated plate composed of fiber-reinforced cross-ply and piezo-thermo-elastic layers. They computed the sensitivity of the static response to variations in the mechanical, thermal, and piezoelectric material constants of the plate. Their results could serve as a benchmark for evaluating the accuracy and validity range of 2D models for such smart hybrid plates. Subsequently, they presented 3D analytical solutions for the free vibrations of laminated plates composed of the fiber-reinforced cross-ply and piezo-thermo-elastic layers, with initial stresses generated by either a temperature change or an electric load [51]. They obtained the frequency sensitivity to variations of various material constants and established a simple approximate formula for the sensitivity coefficients. This provided insight into the frequency dependence of initially stressed laminates on initial temperature, electric load, and geometric and material parameters. Lee and Jiang [36] first derived the state-space formulations for a 3D piezoelectric laminate of 6 mm material symmetry and performed an exact analysis of the coupled electroelastic behaviors of a simply-supported rectangular piezoelectric laminated plate subjected to a mechanical or electric load. The established state-space equations provided a useful means for the future studies on piezoelectric laminated plates with general interlayer (imperfect contact) or boundary conditions (bonded with foundations). Cheng and Batra [37] proposed an asymptotic expansion method to reduce the 3D equations to a hierarchy of 2D equations and used the SSA to investigate the deformations of multiple-electroded piezoelectric laminates subjected to surface tractions and surface and internal electric potentials. Numerical results showed excellent agreement with available exact solutions and demonstrated the errors of existing 2D piezoelectric plate models. Benjeddou and Deu [90] proposed a mixed SSA that retained the standard state displacement and transverse stress augmented with the electric potential and electric displacement to deal with the closed and open electric boundary condition. Its application to the exact analysis of transverse shear actuation and sensing of simply-supported three-layered plates with a piezoceramic core was then performed [38]. The influences of the plate aspect ratio, mechanical and electric excitation amplitudes, and the position and thickness of piezoceramic core on the static responses of the smart laminated plate were examined. The results were useful for comparison with approximate theories and numerical solutions. Vel and Batra [39] presented an exact 3D state-space solutions for the static cylindrical bending of simply-supported laminated plates embedded with shear mode piezoelectric actuators and subjected to mechanical and electric loads on the upper and lower surfaces. They then performed an exact analysis of the 3D deformations of such a smart laminated plate under mechanical and electric loads [40]. The exact displacement and stress solutions for different plate aspect ratios were compared with those obtained by the first-order shear deformation theory (FSDT). They found that the deviation between the FSDT and the exact solutions increases as the plate aspect ratio decreases. Additionally, Chen et al. [52] derived the state-space equations for free vibrations of transversely isotropic piezoelectric bodies, and numerical results illustrated that the results by 2D plate theory deviate from their exact solutions. Ding et al. [41] derived the axisymmetric state-space formulations of piezoelectric laminated circular plates based on 3D theory of piezoelectricity and the finite Hankle transform. They obtained the exact solutions for the rigidly-slipping support and elastically-simple support, which exhibited good agreements with those from the FEM. Following this, Ding et al. [53] established new state-space formulations with lower orders for transversely isotropic piezoelectricity by introducing two displacement functions and two stress functions. They then investigated the free vibration of a simply-supported rectangular piezoelectric laminated plate. The exact results obtained could serve as a benchmark for evaluating 2D approximate plate theories and numerical methods. Utilizing the Hankel transform, Wang [42] transformed the state-space formulation for the non-axisymmetric space problem of transversely isotropic piezoelectric media in cylindrical coordinate. They then presented the analytical solution of a semi-infinite piezoelectric medium based on the transfer matrix method. Subsequently, Wang et al. [43–45] similarly established the state-space formulations for axisymmetric problems of transversely isotropic piezoelectric media using the Hankel transform and transfer matrix method. Simulation results demonstrated that their novel approaches have higher computational efficiency than the classical displacement method and could be used to derive solutions for more complicated loads and boundary conditions. Qing et al. [54] established a state-space formulation for dynamic problems of simply-supported rectangular piezoelectric laminated plates, taking into consideration the transverse shear deformation and rotary inertia of laminate. This was based on the inversion of Laplace transform and modified mixed variational principle. Its application to the analysis of harmonic vibration and transient response of the laminate was presented to demonstrate its accuracy. The results can serve as benchmarks for testing thick plate theories and novel numerical methods. Using the SSA and recursive solution approach, Sheng et al. [46] presented an exact solution for thick orthotropic and piezoelectric laminated plates with clamped and electric open-circuited boundary conditions. The results, which agreed well with those of existing analytical solutions and finite element models, verified the effectiveness of the method. In contrast to the existing works on FGMs with a certain exponential law for the variations of physical fields along the thickness direction, Lezgy-Nazargah [47] performed an exact analysis of cylindrical bending of FGM piezoelectric laminates with arbitrary gradient compositions based on the SSA. The exact solutions, which exhibited excellent agreement with the published results, could be used to assess the accuracy of approximate plate theories and numerical methods.

In addition, Chen et al. [27–29] conducted a series of research on the effect of imperfect bonding on the mechanical behaviors of smart laminated plates, akin to previous research on composite laminates composed of elastic materials. Chen et al. [27, 28] developed 3D state-space formulations to investigate the bending and free vibration of simply-supported orthotropic and angle-ply piezoelectric rectangular laminates. These laminates had interlaminar bonding imperfections, which were modeled using a general spring layer. Concurrently, they explored the cylindrical bending and free vibration of simply-supported adaptive angle-ply laminates, which were bonded with surface piezoelectric actuator and sensor layers [29], as shown in **Figure 2C**. The interfaces between the host elastic laminate and piezoelectric layers were perfect, but the interlaminar bonding of the host laminate was imperfect. **Figure 2D** illustrated that the transverse shear stress τ_{xz} predicted by 3D SSA is exactly continuous at interlayer interfaces for $R = 0$ and the imperfect bonding parameter R has a substantial effect on the distribution of shear stress field along the thickness direction. The precise solutions, grounded in 3D exact theories of elasticity and piezoelectricity, are instrumental in the development and validation of 2D approximate theories and numerical methods. Wang et al. [48] formulated state-space equations for 3D, orthotropic, and linearly magneto-electro-elastic media and investigated the static deformations of a simply-supported laminated rectangular plate under mechanical and electric load. The numerical results were compared with those by Pan [91] to validate the SSA. This approach could be extended for analyses of such smart composite plates with general interlayer and boundary conditions.

Laminated Shells

Shells represent a common class of contemporary structures, with laminated shells offering the benefits of light weight, high strength, and excellent design flexibility in structural configurations. These structures have been extensively employed in critical components across modern aeronautical, aerospace, and transportation industries, including aircraft cabins, radome, rocket adapter, satellite central cylinder, missile fairings, and more. The application of state-space technique in exact analyses of cylindrical shell structure was realized by Soldatos and Hadjigeorgiou [62, 63] using a proposed successive approximation method. They segmented the cylindrical shell into numerous successive and coaxial sub-shells and transformed the Navier-type governing equation of 3D linear elasticity into a set of simpler differential equations for thin shells, based on power series expansion. Subsequently, they derived the state equation and obtained the exact solutions for the free vibration of simply-support homogeneous isotropic cylindrical shells. As the layer number of sub-shells increases, the solutions based on successive approximate method converge towards the exact solutions to the free vibration problems of the shell structure. This successive approximation method paves the way for exact analyses of laminated shells using the SSA. This subsection reviews the relevant research on the static and dynamic responses of 3D laminated shells and panels based on SSAs.

Utilizing the successive approximation method, Hawkes and Soldatos [64] conducted an exact analysis of longitudinal vibrations of homogeneously orthotropic and cross-ply laminated hollow cylinders. Numerical results validated the efficiency and rapid convergence of this approach for addressing vibration responses in laminated shell structures. Soldatos and Ye [55] further examined 3D static, dynamic, thermoelastic, and buckling behaviors of simply-supported homogeneous and cross-ply laminated hollow cylinders and cylindrical panels using successive approximation method. In-depth studies of 3D flexural vibrations [65], stress and displacement distributions [61], and buckling behaviors [69] of laminated hollow cylinders and open cylindrical panels of symmetric and antisymmetric cross-ply were carried out by them. They used the successive approximation method in conjunction with transfer matrix method, which consistently resulted in the solutions of a sixth-order system of algebraic equations, irrespective of the layer number of the structures under consideration.

Fan and Zhang [58] derived the state equations for orthotropic and doubly curved shells and presented the analytical solutions for static and dynamic responses under simply-supported boundary conditions. This was based on the Cayley-Hamilton method and the SSA. Numerical results indicated that the method can be used to study laminated shell structures with arbitrary thickness. It also has the advantage that the scale of the final algebraic equations is independent of the layer number of the structure. Ding et al. [56, 57] performed the exact analyses of axisymmetric deformation, vibration, and buckling behaviors of simply-supported thick laminated closed cylindrical shells using the SSA. Simulation results demonstrated the efficiency and rapid convergence of this method. By introducing the Hellinger-Reissner variational principle and auxiliary function at the lateral boundary, Ding and Tang [66] employed this method to investigate the 3D free vibration of thick laminated cylindrical shells with two clamped edges. The results were in good agreement with those by the FEM, indicating the accuracy of the developed method for analyses of laminated shells with complex boundary conditions. Chen et al. [67] established the state-space formulations for generally anisotropic materials with thermal effect and carried out a 3D vibration analysis of fluid-filled orthotropic FGM cylindrical shells under simply-supported boundary conditions using the SSA. The effects of geometric sizes, material gradient index, and fluid contact on natural frequencies were examined. Numerical results revealed that the natural frequencies predicted by the SSA agree well with the exact solution in [92], validating the effectiveness and accuracy of the present method. Moreover, the fluid contact and the decrease of material gradient index lower natural frequencies, and these relationships are significantly dependent on the thickness of cylindrical shells. Chen et al. [59, 60] carried out the bending and vibration analyses of simply-supported cross-ply and angle-ply laminated cylindrical panels with weak interfaces as shown in **Figure 2E**, described by the spring-layer model, using the SSA in conjunction with the LAT. It was found that the imperfect bonding significantly influenced the distribution of the physical fields, such as the shear stresses in **Figure 2F**, along the thickness direction of laminated shell. Numerical comparison showed that the SSA along with the LAT can be used for the exact prediction of interface damage in laminated panels through natural frequencies in health diagnosis of structures. Moreover, the existing shell theories derived for perfect laminated structures were no longer suitable for effective analyses of laminated structures with imperfect interfaces. In addition, Talebitooti et al. [70] presented 3D analytical solutions for wave propagation in simply-supported orthotropic cylindrical shells with arbitrary thickness subjected to subsonic external flow pressure using the SSA along with the LAT. The sound transmission loss (TL) of a thick-walled cylindrical shell at oblique incidence was calculated. It was found that for thick-walled shells and the plane wave in high frequencies, the significant difference of the obtained TL from the 3D analytical method and conventional shell theories can be observed due to the shear and rotation effects. Thus, the 3D analytical solutions predicted by the developed method could be a benchmark for validating the approximate shell theories when modeling waves propagating in thick laminated shells.

In addition to the aforementioned cylindrical shell structures, laminated spherical shells have also found extensive applications in aerospace, structural engineering, and geophysics [93–96]. By introducing several displacement and stress functions, Chen and Ding [73] established the state-space formulations of spherically isotropic elastic body. Then they presented the exact solutions for static deformations of a three-layered spherical shell subjected to distributed uniform pressure. They further employed this method to study the free vibration of multi-layered spherically isotropic hollow spheres [75]. The frequencies of vibration modes of different orders obtained through this method are in good agreement with the results predicted by conventional 3D method. These results underscored the accuracy and effectiveness of the SSA, particularly when dealing with spherical shell structures with multiple layers. These exact solutions also serve as a benchmark for validating approximate shell theories and numerical methods. Similarly, by introducing displacement and stress functions, two separated state-space formulations of a spherically isotropic piezoelectric body were derived by Chen et al. [74]. They then performed an exact static analysis of a laminated, spherically isotropic, and piezoelectric hollow sphere. The results obtained are in excellent agreement with the exact solutions provided by Heyliger and Wu [97], indicating that the method offers an efficient and powerful tool for 3D exact analyses of laminated piezoelectric spherical shells. Subsequently, Chen et al. [76] employed this SSA to study the free vibrations of FGM piezoceramic hollow spheres with radial polarization, based on the LAT. They found that the piezoelectric effect and material gradient index have significant influences on natural frequencies. Furthermore, the SSA in conjunction with the LAT could be applied to the 3D exact analysis of arbitrary nonhomogeneous spherical shell structures.

In recent years, dielectric elastomers (DEs), characterized by their electromechanical coupling and large deformation capability, have garnered increasing interest from researchers. These materials are being explored for the development of soft, lightweight, and small-scale actuators, sensors, transducers, and soft robotics [98, 99]. Efforts have been made to study vibration and wave behaviors in nonhomogeneous or laminated DE shells using the state-of-the-art SSAs. For instance, Zhu et al. [68] utilized the SSA in conjunction with the LAT to conduct an exact analysis of axisymmetric torsional and longitudinal vibration in an incompressible DE cylindrical shell. The shell was subjected to axial pre-stretch and inhomogeneous electric biasing fields induced by the radial voltage. It was found that the natural frequencies of the DE cylindrical shells predicted by the SSA are in good agreement with those from the exact solutions for different axial mode numbers. Wu et al. [71] investigated the guided circumferential shear-horizontal (SH) -type and Lamb-type waves in an incompressible DE cylindrical shell subjected to inhomogeneous biasing fields based on the SSA along with the LAT. Numerical results demonstrated the accuracy and efficiency of the SSA for analyses of wave propagations in inhomogeneous shell structures. The study also found that the propagation properties of two types of guided circumferential waves are significantly dependent on the biasing fields and geometric sizes. Subsequently, they further employed this method to analyze the axisymmetric torsional and longitudinal waves propagating in an FGM DE cylindrical shell under the complex loads of axial pre-stretch, radial pressure, and radial electric voltage [72]. Simulation results indicated that the pre-stretch, radial pressure difference, and electric load can be utilized to steer the propagation behaviors (e.g., frequency and wave velocity) of the axisymmetric guided waves in the FGM DE cylindrical shells. The material gradient index has a significant influence on the tunable capacity of these external loads. Additionally, Mao et al. [77] performed the 3D torsional and spheroidal vibration analyses of an incompressible DE spherical shells subjected to inhomogeneous biasing field induced by radial electric voltage and radial pressure difference. This work was also based on the SSA in conjunction with the LAT. Numerical results illustrated that the SSA is also suitable for vibration prediction of DE spherical shells with high accuracy. Furthermore, it was found that the natural frequencies of various vibration modes in DE spherical shell can be flexibly tuned by adjusting the external electric voltage and internal pressure.

SIMPLIFIED STATE-SPACE APPROACHES TO LAMINATED STRUCTURES

Drawing upon structural theories such as beam, plate, and shell theories, a variety of simplified state-space formulations have been developed. These formulations, tailored to specific structural configuration, have been extensively utilized to investigate the mechanical behaviors of laminated structures based on the transfer matrix method. This section provides a comprehensive review of the evolution of these simplified SSAs with reduced dimensions. Furthermore, it highlights the progress made in the corresponding applications to the analyses of static and dynamic responses of laminated structures in beam, plate, and shell configurations. A summary of these developments is provided in **Table 2**.

Laminated Beams

As early in 1950, Thomson [10] pioneered the establishment of state-space formulations for the depiction of beam bending, utilizing the simplified beam theory. This was followed by a free vibration analysis of a non-uniform beam, employing the LAT and transfer matrix method. After nearly four decades of gradual development, the SSAs have regained significant attention from the research community, spurred by the rapid emergence and widespread application of laminated composite structures.

Khdeir and Reddy [100, 101] presented analytical solutions for the buckling and free vibration of cross-ply laminated beam with arbitrary boundary conditions. These solutions were derived using the SSA in conjunction with the refined beam theories. Numerical results revealed that the discrepancies among different shear deformation theories are significantly less than that between any of them and the Euler-Bernoulli beam theory. Moreover, as the length-to-thickness ratio increased, the difference in critical buckling loads and natural frequencies between shear deformation theories and Euler-Bernoulli theory diminishes. Subsequently, Khdeir [103] presented an analytical solution for the transient response of antisymmetric cross-ply laminated beams with generalized boundary conditions and under arbitrary loads. This was achieved by using the classical, first-, and third-order shear deformation beam theories in tandem with the SSA. The study concluded that the results predicted by the high-order theory of Reddy (HOBT) and Timoshenko shear deformation beam theory are more accurate than those predicted by the Euler-Bernoulli theory for thick beams. Furthermore, Khdeir and Reddy [105] developed the analytical solutions for the bending of symmetric and antisymmetric cross-ply laminated beams with arbitrary boundary conditions and subjected to arbitrary loads. They found that the shear deformation, orthotropic ratio, symmetry of cross-ply, and boundary conditions significantly influence the deflection of laminated beams. They also employed this method to analyze the thermoelastic response of symmetric and antisymmetric cross-ply laminated beam under arbitrary boundary conditions and subjected to general temperature fields [106]. The study concluded that the effect of temperature fields on the deflection of composite beam is significantly dependent on layer number, shear deformation, cross-ply symmetry, and boundary conditions. The thermal buckling of cross-ply laminated beams subjected to uniform temperature distribution was also investigated by them [107] using the refined beam theories in conjunction with the SSA. The effects of length-to-thickness ratio, modulus ratio, thermal expansion coefficients, boundary conditions, and layer number on the critical buckling temperature were systematically discussed. In addition, Srinivasan et al. [115] developed a unified SSA based on different HOBTs taking both normal strains and shear deformation effects into consideration. They then presented analytical solutions for laminated composite beams subjected to transverse loads and under various boundary conditions. In the four HOBTs, shear deformation was described by four warping functions corresponding to exponential, hyperbolic, trigonometric, and parabolic functions. Comparative studies revealed that the models associated with the exponential and hyperbolic functions yield accurate and identical results in all studies. Ramaprasad et al. [116] established a unified state-space coupled field formulation for generating analytical solutions for thin-walled composite open section beams, which are extensively used in aerospace. All non-classical effects such as torsional warping, warping shear, material coupling, and shear deformation were taken into consideration. The results for arbitrarily laminated non-symmetric composite channel and I-section beams subjected to bending and torsional loads showed excellent agreement with the reported results in Ref. [140], validating the efficiency and accuracy of the proposed method.

Considerable research has been conducted to investigate the mechanical behaviors of smart beam systems, utilizing simplified

TABLE 2 | Summary of simplified state-space approaches based on structural theories and their applications in the mechanical analyses of complex structures in beam, plate, and shell configurations.

Configuration	Structural theory	Mechanical behavior	Reference
Beam	CBT, FOBT HOBT	Static and dynamic responses	[100–102]
Beam	CBT, FOBT, HOBT	Dynamic response	[103, 104]
Beam	CBT, FOBT, HOBT	Static response	[105–109]
Beam	CBT	Dynamic response	[10, 110–113]
Beam	Zig-zag beam theory	Dynamic response	[114]
Beam	HOBT	Static response	[115, 116]
Beam	FOBT	Dynamic response	[117]
Plate	CPT, SDPT	Static and dynamic responses	[118]
Plate	CPT, SDPT	Static response	[119, 120]
Plate	YNS plate theory	Static response	[121]
Plate	SDPT	Static response	[122, 123]
Plate	SDPT	Static and dynamic responses	[124–130]
Plate	SDPT	Dynamic response	[131]
Plate	TVPT	Static response	[132]
Plate	FVPT	Static and dynamic responses	[133]
Plate	FVPT	Dynamic response	[134]
C-shell	CST, FOST, TOST	Static and dynamic responses	[135]
C- and S-shells	CST, FOST, TOST	Static and dynamic responses	[136]
C- and S-shells	CST, FOST, TOST	Dynamic response	[137]
C-shell	FOST	Dynamic response	[138, 139]

Acronyms: CBT, classical beam theory; FOBT, first-order beam theory; TOBT, third-order beam theory; CPT, classical plate theory; SDPT, shear-deformation plate theory; YNS, Yang, Norris, and Stavsky; TVPT, two-variable plate theory; FVPT, four-variable plate theory; CST, classical shell theory; FOST, first-order shell theory; TOST, third-order shell theory; C-shell, Cylindrical shell; S-shell, spherical shell.

beam theories and the SSA. For instance, Pota and Alberts [110], building on the Euler-Bernoulli beam theory, developed an SSA for the free vibration analysis of a slewing beam system. This system was bonded with piezoelectric sensors and actuators, which were applied by an external voltage. For the special case where the hinged end of the structure was clamped, the results obtained agree with those presented in Ref. [141]. Based on the first-order beam theory (FOBT) and HOBT, Aldraihem and Khdeir [108] presented analytical solutions for the bending of beams bonded with thickness-shear and extension piezoelectric actuators under various boundary conditions using the SSA. They found that the operation modes of the piezoelectric actuator could result in significant difference between the deflections predicted from the FOBT and HOBT. Built upon the zig-zag beam theory in conjunction with the SSA, they also presented analytical solutions for the free vibration of soft-core sandwich beams with arbitrary boundary conditions [114]. The obtained results indicated that the support conditions have significant influences on the relationship between resonant frequency and length-to-thickness ratio. The comparison between the obtained results and the existing experimental, analytical, and numerical results in the literature showed that the zig-zag beam model provides accurate natural frequencies for sandwich beam with a soft core. These results also serve as useful benchmarks for approximate solutions from numerical methods like Rayleigh-Ritz method, FEM, etc. Furthermore, based on the Euler-Bernoulli beam theory, Palmeri and Adhikari [111] developed a Galerkin-type SSA for the transverse vibrations of slender double-beam system with a viscoelastic inner layer, described by standard linear solid model. Numerical results demonstrated the validity and accuracy of the proposed method in both frequency- and time-domain analyses. Subsequently, using this Galerkin-type analytical model, Palmeri and Ntotsios [112] investigated transverse vibrations of viscoelastic sandwich beams under different boundary conditions as shown in **Figure 3A**. It was found that the viscoelasticity significantly affect the frequencies and transverse mode shapes of the sandwich beams (see **Figures 3B, C**), and the numerical results verified the accuracy and versatility of the proposed method. Li et al. [113] also proposed a novel SSA for the transverse vibration analysis of a double-beam system. They introduced a mode-shape constant to the state space in the modeling framework to improve the computational accuracy and efficiency. Moreover, Sahmani and Ansari [109] investigated the buckling behaviors of nanobeams under various boundary conditions by utilizing nonlocal continuum beam models of different beam theories, including the CBT, FOBT, and Levinson beam theory (LBT), in conjunction with the SSA. The results obtained from the developed nonlocal beam model match those from the molecular dynamic simulation given by Ansari et al. [142], helping to predict the appropriate value of nonlocal parameter. It was also found that the nonlocal beam model and boundary conditions significantly influence the predicted value of nonlocal parameters. Additionally, Fazeli et al. [117] performed free and forced vibration analyses of smart orthotropic cross-ply laminated stepped beams integrated with a piezoelectric actuator, using the SSA based on the FOBT. The obtained natural frequency for piezoelectric sinusoidal excitation was compared with those from FEM and experimental study. It was found that the analytical method provides an efficient tool to predict the vibration behaviors of smart stepped laminated beam with high accuracy.

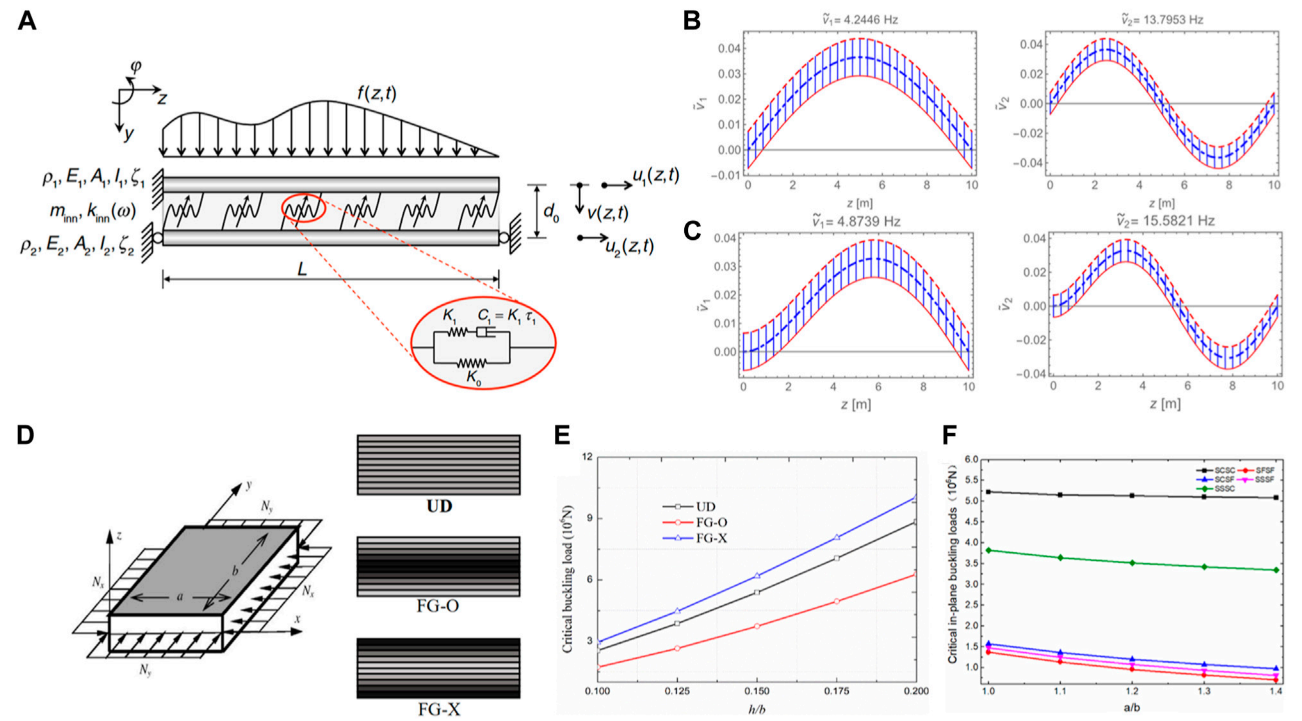

FIGURE 3 | Laminated structures of beam and plate configurations and the mechanical responses predicted by the simplified state-space approaches: **(A)** viscoelastic sandwich beams with its viscoelastic core modeled by an elastic spring in series with a viscous dashpot [112]; **(B)** first two transverse modes for undamped sandwich beam; **(C)** first two transverse modes for damped sandwich beam; **(D)** graphene reinforced composite functionally graded material plate with different gradient distributions along thickness direction [131]; **(E)** variations of critical buckling loads versus thickness-to-length ratio; **(F)** effect of the width-to-length ratio on critical in-plane buckling loads under different boundary conditions for FG-X pattern (Reproduced with permission from Palmeri et al. [112], copyright 2016 by ASCE, Open Access; Reproduced with permission from Lv et al. [131], copyright 2019 by IOP Publishing, Open Access).

This paragraph highlights the recent work on applications of the SSA to the mechanical analyses of FGM laminated beams. Based on the CBT, FOBT, and HOBT, Trinh et al. [104] developed analytical solutions for predicting the fundamental frequency of FGM sandwich beams under various boundary conditions using the SSA. Simulation results revealed that boundary conditions, material parameters, layer arrangement, and length-to-thickness ratio all significantly affect the fundamental frequency of such an FGM sandwich beam. Subsequently, they presented analytical solutions for the buckling and free vibration of FGM beams under mechanical and thermal loads using the SSA based on the HOBT [102]. The results demonstrated the efficiency and effectiveness of the analytical model. By utilizing the multi-term Kantorovich-Galerkin method for describing the displacement components, Wang et al. [143] proposed a 2D elasticity model in conjunction with the SSA to investigate the bending and free vibration of laminated graphene-reinforced composite (GPC) beams. Five different graphene distribution patterns were presented by altering the volume fraction of graphene in each layer. The study found that the laminated GPC beam with pattern X exhibits the smallest deflection and highest fundamental frequency for a high length-to-thickness ratio. Conversely, it displays the largest deflection and smallest fundamental frequency for a low aspect ratio due to the decrease in the transverse shear stiffness.

Laminated Plates

Chandrashekara and Santhosh [144] pioneered the SSA for the free vibration analysis of cross-ply laminated plates, utilizing the Maclaurin series expansion along the thickness direction. The efficacy of this method was substantiated by comparing it with the exact elasticity solution for plates under the plane strain assumption. Their findings revealed that the SSA, even with a limited number of series expansion terms, yields results in excellent concordance with the exact results for both thin and thick plates. In a similar vein, the research team of Khdeir and Reddy carried out extensive studies on the development of SSAs based on approximate plate theories, demonstrating their applicability to the analyses of mechanical behaviors of laminated plates. For instance, Khdeir [124, 125] utilized the SSA to formulate Lévy-type solutions for the buckling and free vibrations of antisymmetric angle-ply laminated plates, based on the first-order shear deformation plate theory (FSDPT). Moreover, Khdeir and Reddy [119] generated Lévy-type solutions for the bending of antisymmetric cross-ply rectangular laminates under sinusoidal transverse loads, employing the classical plate theory (CPT), FSDPT, and third-order shear deformation plate theory

(TSDPT) in conjunction with the SSA. Their numerical results highlighted significant discrepancies between the deflections predicted by the CPT and those from FSDPT and TSDPT under various boundary conditions. In another study, Khdeir [120] presented Lévy-type solutions for shear deformable antisymmetric angle-ply laminated plates, using the Yang, Norris, and Stavsky (YNS) theory in combination with the SSA. A comparative analysis between the Lévy-type solutions derived from the FSDPT and classical Kirchhoff plate theory in conjunction with the SSA [120] revealed significant difference between the results from the Kirchhoff plate theory and those from the FSDPT. In contrast to laminates with antisymmetric angle-plies, Librescu and Khdeir [122] proposed a higher-order plate theory and examined the static deformation of symmetric cross-ply laminated plates using the SSA. They further employed this method to investigate the buckling and free vibration of the same laminated plate under various boundary conditions [126]. Built upon the laminated anisotropic plate theory as considered in Refs [145, 146], Khdeir [127] generated Lévy-type solutions for the buckling and free vibration of symmetric cross-ply laminated plate utilizing the state-space technique. Khdeir [128] also used the SSA to develop the Lévy-type solutions for the buckling and free vibration of unsymmetric cross-ply laminated plates based on the refined shear deformation theory used in [147]. In [118], Khdeir commented on the application of the SSAs to bending, buckling, and free vibration analyses of composite laminates. It was concluded that for plate structures with the characteristic length-to-thickness ratio greater than 20, the results from [119–122, 124–126], [127, 128] obtained from different plate theories under various boundary conditions agree well with the exact solutions. Lastly, based on the FSDPT, Xing and Xiang [123] proposed an analytical method for the buckling of symmetric cross-ply composite laminates, using the separation-of-variable method in conjunction with the SSA to enhance the computational efficiency and circumvent numerically ill-conditioned problems. The accuracy and effectiveness of the analytical model were validated by comparing the numerical results with the those in the literature [126, 148].

SSAs have also been utilized to investigate the size-dependent behaviors of nanoplates. Leveraging the two-variable plate theory (TVPT) in conjunction with the Eringen's nonlocal elasticity theory, Sobhy [132] examined the effect of hygrothermal conditions on the bending of orthotropic nanoplate under various boundary conditions. The study concluded that the results predicted by the TVPT agree well with the published solutions, suggesting that the TVPT is more accurate than the CPT. Furthermore, the deflection of the nanoplate was found to be directly proportional to the temperature, moisture concentration, scale parameter, side-to-thickness ratio, and modulus ratio, while inversely proportional to the length-to-thickness ratio and mechanical load. Trinh et al. [129] developed the state-space Lévy solutions for size-dependent static, free vibration, and buckling behaviors of FGM sandwich plates under diverse boundary conditions. The analysis was based on a refined shear deformation theory in conjunction with the modified couple stress theory. It was revealed that the material distribution, geometric parameter, characteristic length parameter, and boundary conditions significantly alter the deflections, stresses, natural frequencies, and critical buckling loads of the sandwich plate. The results obtained provide a valuable benchmark for mechanical analyses of FGM microplates. In a similar vein, several research groups have focused on the mechanical behaviors of FGM plates. For instance, Zhang et al. [130] employed the state-space Lévy method to investigate the vibration and buckling behaviors of functionally graded carbon nanotube (FG-CNT) reinforced composite rectangular plate subjected to in-plane loads, based on the TSDPT. The simulation results indicated that the CNT distribution, length-to-thickness ratio, and boundary conditions have substantial influences on the natural frequencies, mode shapes, and critical buckling loads of the FG-CNT composite plates. Additionally, Lv et al. [131] employed this theory to explore the vibration characteristics of moderately thick FGM graphene reinforced composite plates with different gradient distributions in **Figure 3D** under different boundary conditions. The effects of graphene distribution, plate thickness-to-width ratio, length-to-width ratio, and external load on natural frequency and critical buckling loads were thoroughly examined and discussed. **Figure 3E** showed that the dependence of critical buckling load on thickness-to-length ratio is affected by the gradient distributions of the graphene. **Figure 3F** revealed that the relationship between the critical in-plane buckling loads and the width-to-length ratio is dependent on the support conditions of the FGM plate. Utilizing the four-variable plate theory (FVPT) in conjunction with the SSA, Demirhan and Taskin [133] investigated the bending and free vibration of porous FGM plate with two opposite simply-supported edges. The influence of porosity parameter, plate aspect ratio, material gradient index on the deflections and natural frequencies were scrutinized. The analytical method was validated by comparing the obtained results with those in the literature [149]. Similarly, Rouzegar et al. [134] presented the state-space Lévy solutions for the free vibration of a smart laminated plate, with its FGM core sandwiched by two piezoelectric layers, based on the FVPT. A comparison between the results and those available in Refs [150, 151] verified the accuracy of the analytical model. It was concluded that the piezoelectric layer thickness, material graded index, plate geometric parameter, and boundary conditions significantly influence the natural frequencies of the smart composite structure.

Laminated Shells

Recent years have seen a surge in research focused on the development of simplified state-space formulations based on various shell theories and their subsequent applications in the mechanical analyses of laminated shells. Built upon the classical, first-order, and third-order shell theories, Khdeir et al. [135] developed analytical solutions for the bending, free vibration, and buckling behaviors of cross-ply circular cylindrical shells under various boundary conditions, utilizing the stat-space technique. The study explored the effects of boundary conditions, layer number, and shear deformation on the deflection, natural frequencies, and critical buckling loads. Numerical results indicated a significant deviation between the results from the classical shell theory (CST) and those form the shear deformation theories. Furthermore, the inclusion of transverse shear strain results in the discrepancies between results from the first-order shell theory (FOST) and third-order shell theory (TOST). Subsequently, Khdeir and Reddy [136] presented Lévy-type solutions for the bending, free vibration, and buckling of doubly curved cylindrical and spherical shells under various boundary conditions. These solutions were based on three types of shell theories in conjunction with the SSA, and the conclusions drawn were similar to those in Ref. [135]. In a related study, Ye and Soldatos [69] established the state-space formulations for the buckling analysis of simply-supported laminated hollow cylinders and cylindrical panels subjected to the single or combined mechanical loads, by neglecting the initial shear stresses. Simulation results showed that the trends of the critical buckling load parameter as a function of the stiffness of the hollow cylinder or cylindrical shell agree with the 3D exact results. The SSA exhibits higher efficiency than the FEM for the shell structures with a large number of layers. Recently, Dozio [137] developed general state-space Lévy-type solutions for the free vibration of orthotropic multilayered cylindrical and spherical panels under different boundary conditions. Various shell theories can be integrated according to the shell thickness, the degree of anisotropy of the considered problems, and the desired accuracy. Compared with the existing Lévy-type vibration solutions [152, 153], the final matrix of the present solutions was invariant with respect to the 2D kinematic shell theories. A single-layer spherical shell was selected as a numerical example and the comparison of the displacement fields obtained from the 2D shell theory and 3D theory validated the accuracy of the method for the single-layered spherical shell. Moreover, the accuracy of the developed 2D analytical model for multilayered shells strongly depends on the refinement of the 2D shell theory, especially when the panel is thick and exhibits strong material anisotropy. Hosseini-Hashemi et al. [138] performed a free vibration analysis of FGM viscoelastic cylindrical panel made of polymeric foams using the SSA, based on the FOST. The effects of geometric and material parameters as well as boundary conditions on natural frequencies were examined. The agreement of the obtained results with those from the FEM verified the validity and accuracy of the developed method. Finally, Razgordanisharahi et al. [139] employed the SSA to investigate the free vibrations of a honeycomb sandwich panel consisting of a hexagonal honeycomb core layer and two face sheets, based on the FOST. The comparison of the obtained results with those in previous studies verified the accuracy of the model. The simulation results revealed that the natural frequencies of the sandwich panel are also significantly dependent on the geometric parameters of the honeycomb core.

STATE-SPACE APPROACHES TO COMPLEX PERIODIC STRUCTURES

In addition to the previously discussed laminated structures in beam, plate, and shell forms, complex periodic structures, composed of numerous identical units or periodic arrays in varying geometric configurations, have been extensively utilized in various fields such as aerospace, civil and structural engineering, and marine. These applications include, but are not limited to, space antenna, aircraft fuselages, and ship hulls. The multitude of potential unit configurations or unit array often complicates the mechanical analyses of these complex structures. To mitigate this complexity, units of different configurations are often modeled using approximate structural theories, which can lead to significant discrepancies in predicting the mechanical behaviors of entire structures [154]. The FEM is frequently employed as a robust and effective tool for analyzing these complex periodic structures due to its vast and comprehensive library of elements. However, the natural frequencies of complex periodic structures composed of many identical units are typically closely grouped within frequency bands. The number of natural frequencies within each band is equal to the number of the repeated units in the structure [155, 156]. Consequently, accurately predicting the natural frequencies and corresponding mode shapes of these structures using the FEM can be both time-consuming and potentially inaccurate. To address these challenges, mixed transfer matrix methods (also called mixed SSAs) have been proposed that combine numerical or theoretical methods with the transfer matrix method based on state-space formulations. These hybrid approaches are particularly useful for handling the mechanical problems associated with complex periodic structures that have intricate unit configurations. This section will review the development and application of these mixed SSAs in the dynamic responses of these complex periodic structures, as summarized in **Table 3**.

TABLE 3 | Summary of mixed state-space approaches and their applications in complex periodic structures composed of different unit configurations.

Unit configuration	Method	Dynamic response	Reference
Beam	TMM	Vibration	[157–160]
Beam	TMM and WPA	Vibration	[161]
Beam	FEM and TMM	Wave	[162]
Duct, rod, and beam	TMM and SM	Wave and vibration	[163]
Beam	TMM and WPA	Wave	[164–166]
Beam	TMM and WPA	Vibration	[167, 168]
Pipe, beam, and truss	TMM and WPA	Wave	[169]
Beam and truss	FEM, TMM, and WPA	Wave	[170]
Truss	FEM, TMM, and WPA	Vibration	[171–174]
Truss	FEM, TMM, and WPA	Wave	[175]

Acronyms: TMM, transfer matrix method; FEM, finite element method; WPA, wave propagation approach; SM, Spectral method.

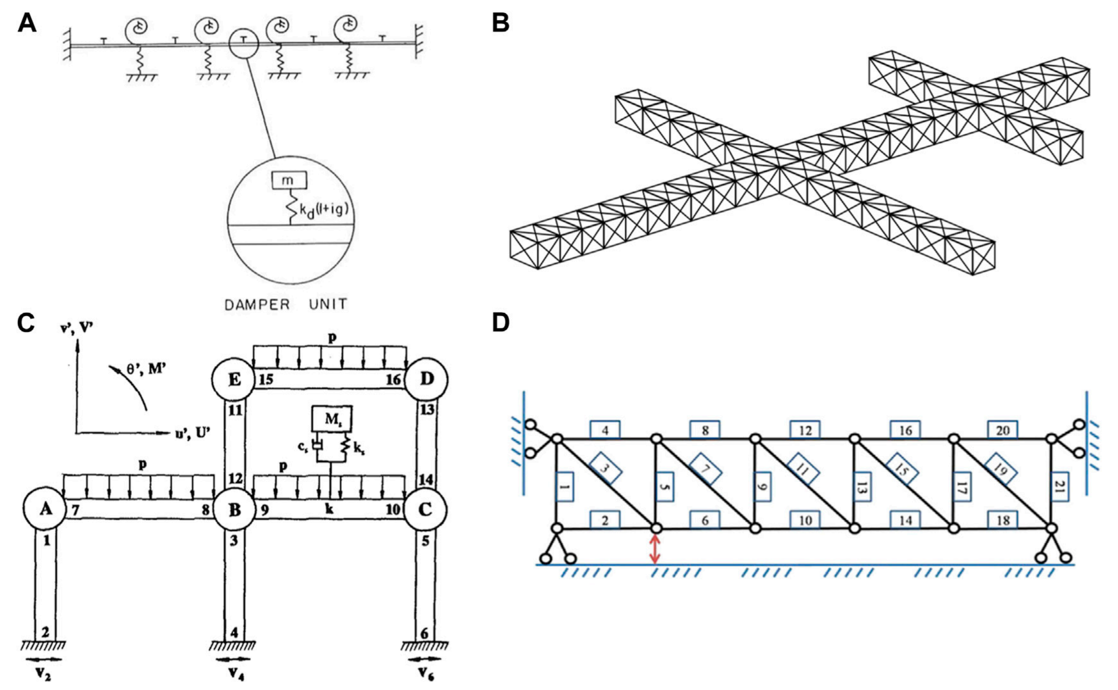

FIGURE 4 | Some cases of periodic complex structures. **(A)** Periodic beam structure with periodic elastic supports [157]; **(B)** schematic of truss-type space periodic structure; **(C)** frame structure based on beams [175]; **(D)** small-scale five-span planar truss-type structure [174] (Reproduced with permission from Lin and Mcdaniel [157], copyright 1969 by ASME; reproduced with permission from Cai and Lin [175], copyright 1991 by ASCE; reproduced with permission from Ren and Zhou [174], copyright 2014 by Elsevier, Open Access).

Beam-Type Periodic Structures

Professor Y. K. Lin and co-workers have conducted extensive research on the dynamic behaviors of complex periodic beam structures on various flexible supports, utilizing the transfer matrix method based on state-space formulations. For instance, Lin and Mcdaniel [157] developed an analytical method for predicting the frequency response of a finite periodic Euler-Bernoulli beam on multiple elastic supports as shown in **Figure 4A**, which was seen as a simplified version of a skin-stringer panel system used in flight vehicle design [176], based on the transfer matrix method. They presented the frequency-response functions for both undamped and damped systems and discussed the effect of the dampers, modeled by spring-suspended mass, on frequency responses. Subsequently, Vaicaitis et al. [158] investigated the spatial decay in the dynamic response of an infinite damped periodic Euler-Bernoulli beam structure, supported by identical elastic springs and subjected to sinusoidal excitations at a specific unit. They discovered that the distance of spatial decay in frequency response is dependent on the excitation frequency, the shape of spectral density, and the cross-spectral density of the forcing field. They also found that dampers can be used for vibration control of such a periodic beam structure. Furthermore, Vaicaitis and Lin [161] analyzed the frequency response of a finite N-span periodic Euler-Bernoulli beam on evenly spaced elastic supports, subjected to turbulent boundary-layer pressure excitation. They used the transfer matrix method in conjunction with the flexural wave propagation approach. Numerical results demonstrated the convenience and suitability of the transfer matrix method based on state-space formulations in studying the dynamic response of periodic structures subjected to random pressure fields. Lin and Yang [159] carried out a free vibration analysis of a simply-supported disordered periodic beam using the transfer matrix method for state vectors. They considered two types of disorder: the random deviation of unit length from an ideal identical value and the random fluctuation of the bending stiffness. They revealed that random unit lengths significantly influence natural frequencies of different orders, but the effect of random bending stiffness on frequencies is negligible for such a disordered multispan beam. Moreover, random unit lengths also led to variability in the normal modes. Yang and Lin [160] further examined the effect of random unit-length deviations of a disordered periodic beam on multiple hinge supports on its frequency responses. They applied two types of excitations to the beam structure, namely, a concentrated force or moment and a distributed frozen force. Their numerical results showed that considering the disorder in unit lengths decreases the amplitude of the statistical average of the frequency response function, particularly near the resonance frequencies, but increases its standard deviation when approaching these frequencies. Furthermore, some vibration modes that do not appear in the frequency spectrum of a perfect periodic beam can be actuated in this disordered structure. Later, Yeh and Chen [162] employed the finite element method and transfer matrix method to analyze the longitudinal wave propagating in a periodic sandwich beam structure, where the base beam was covered by the constrained layer and viscoelastic layer. The study revealed that the length ratio and base beam materials could be used to tune the location and width of the stop and pass bands of this periodic structure, thereby filtering certain bands of frequencies. Assis et al. [163] examined the complex band structures and forced response of a one-dimensional phononic system with arbitrary geometric and material profiles using the transfer matrix method. They derived the transfer matrix of the system from a new spectral approach based on a Riccati differential equation with the impedance as variable. They explored phononic systems in duct, rod, and beam configurations, and the obtained results are consistent with those from traditional plane wave expansion method and spectral element method, thus validating the high accuracy and efficient computational cost of the method.

In addition to the previously mentioned transfer matrix method, a wave propagation approach has also been adopted to investigate the dynamic behaviors of infinite or semi-infinite periodic structures. While this method falls outside the purview of this article, we provide a selection of representative literature [177–184] for the readers' convenience. By integrating the transfer matrix method and wave propagation approach, Lin and others proposed an innovative theoretical method to analyze the dynamic response of long periodic beams subjected to external excitations. For instance, Lin et al. [164] investigated the dynamic behaviors of a surface-mounted pipeline, modeled as an infinitely long Euler-Bernoulli beam, under seismic excitation. The seismic input was introduced to the pipeline structure through evenly spaced supports. Simulation results validated the efficacy of this method, which could be expended to examine the effect of random variation of the support spacings and the ground motion on the pipeline's response. Furthermore, Cai and Lin [165] explored the localization property of wave propagation in disordered periodic structures using the transfer matrix method in conjunction with the wave propagation approach. They proposed a new perturbation scheme to calculate the localization factor, defined as the average exponential decay ratio of wave propagation along the propagation direction, based on probability theory. The numerical example of a multi-supported Euler-Bernoulli beam with an additional torsional spring at each support highlighted two significant features of the developed method: its high accuracy and the inclusion of damping effect in practical structures. Moreover, they utilized the proposed method to study the high-level dynamic response near the point of dynamic excitation, induced by the structural disorder in a disordered periodic structure [167]. Their simulation results indicated that both the number of disordered units and the deviation level of disorder significantly influence the concentrated response near the excitation point. An increase in either one of these factors leads to larger standard deviations of responses. However, structural damping can help to reduce the response level and the likelihood for the responses to reach a much higher level. Additionally, Qiu and Lin [166] computed the localization factor of mono- and multi-channel waves in disordered periodic beam structures supported by linear torsional and translational springs. Their study found that the computed results agree well with those from Monte Carlo simulations, except near boundaries of each wave-passage band. Prof. Lin provided an insightful review of these works on wave propagation and frequency responses of disordered periodic beam structures [185]. Furthermore, Romeo and Paolone [169] employed the transfer matrix method to conduct the wave propagation analysis in three-coupled periodic structures among transversal, longitudinal, and torsional waves. They considered three types of periodic structures, namely, pipes, thin-walled beams, and truss beams. Their unique band structures, including the stop, pass, and complex domains, were computed and presented. Saeed and Vestroni [168] presented an exact closed-form method for frequency response analysis of periodic beam system under multiple excitations using the wave transfer matrix method. The wave solutions were formed by a free wave field incorporating the dynamic response of the entire structure and a forced wave field generated by constrained forces or external excitations. Two numerical examples, namely, a multi-constrained translating string and a beam on multiple supports, demonstrated the efficiency and accuracy of the proposed approach.

Truss-Type Periodic Structures

Complex truss-type structures composed of spatially periodic truss units are a prevalent form of large space structures as shown in **Figure 4B** used in aerospace and civil engineering due to their light weight, high strength, and design flexibility. The truss units are interconnected end-to-end, forming sections of a spatially periodic array, with different sections potentially intersecting or interconnecting. The dynamic analysis of these truss-type periodic structures presents a greater challenge compared to finite periodic beams for two primary reasons. First, the shared degrees of freedom at the interface between two adjacent truss units are larger than those between beam units. Second, the quantity of truss units surpasses that of beam units, and truss-type periodic structures are often considered infinitely long. To address these challenges, Lin and others proposed a novel method that integrates the FEM, the transfer matrix method, and the wave propagation approach. The FEM was employed to model a truss unit or each type of periodic unit array in the structure, accurately representing the dynamic characteristics of the unit, irrespective of its complexity. Traditional transfer matrices for state vectors were converted into the transfer matrix for wave vectors, enabling the identification of various types of waves, dependent on the number of degrees of freedom at the interface of two adjacent truss units, propagating along the structure in two opposite directions [186]. Moreover, the innovative concept of wave reflection and transmission matrices was introduced to replace single wave reflection matrix in the traditional wave propagation approach. This innovation circumvents computational errors associated with the apparent growing wave in the opposite propagation direction caused by the damping effect of practical structures.

Yong and Lin [170] applied this method to investigate the dynamic responses of long and complex periodic structures composed of beam and truss units under external excitation.

Their numerical examples demonstrated that this method is more computationally efficient than applying the FEM directly to the entire structure. They [171] further studied the dynamic behaviors of complex truss-type periodic structures, consisting of two intersecting arrays of truss-cell units, subjected to sinusoidal and impulse point excitations. The numerical results validated the efficiency, accuracy, and numerical robustness of this novel method for analyzing truss-type space structures with damping. Built upon this method, Cai and Lin [175] performed the dynamic response analysis of engineering frame structures composed of interconnected slender components as shown in **Figure 4C**, subjected to either concentrated or distributed loads. Every slender component was treated as a multi-channel waveguide, with its dynamic property characterized by wave-reflection and transmission matrices. The entire frame structure was then conceptualized by a network of these waveguides. Numerical examples, encompassing different boundary types and external loads, demonstrated numerical efficiency and theoretical accuracy of this approach for analyzing frame structures composed of interconnected members. In a further development, Yong and Lin [172] developed an enhanced computational scheme for dynamic response analysis of truss-type space periodic structures with complicated structural configurations and subjected to arbitrary excitations. This scheme featured three significant improvements: the characterization of individual truss-arrays via wave-scattering matrices, a new bookkeeping system to track transmitted, reflected, and re-directed waves in various direction, and a procedure to eliminate step-by-step superfluous unknowns for simplifying the calculation. Their research demonstrated that this innovative scheme offers a more efficient and general numerical tool for the response analyses of entire truss-type periodic structures and even individual structural members. In a parallel development, Luongo and Romeo [173] proposed a modified version of traditional complex wave vector approach for the dynamic analysis, including the free and forced vibrations, of long undamped periodic structures. This version introduced two significant features to address the numerical difficulties encountered by the complex wave vector method. First, the real transfer matrix for state vectors was transformed into a transfer matrix for wave vector, which remained real, thus avoiding an ill-posed problem. Second, the wave vectors were rearranged to let the computation to proceed in the propagation direction of waves, thereby avoiding the transfer matrix ill-conditioning. Numerical examples of an N-span Euler-Bernoulli beam structure validated the efficiency and superiority of this modified wave vector approach. Recently, Ren and Zhou [174] proposed a strain response estimation strategy for the fatigue monitoring of an offshore truss structure, using the state-space formulations and a Kalman filtering process. A small-scale planar truss periodic structure in **Figure 4D** subjected to deterministic and stochastic excitations was simulated to estimate the effectiveness and integration capacity of the proposed theoretical algorithms with the fatigue monitoring system.

STATE-SPACE BASED NUMERICAL METHODS

In general, the state equations of complex structures can be resolved either analytically or numerically. Traditional state-space analytical methods often encounter challenges when dealing with boundary conditions other than the simply-supported ones, such as the free and clamped end conditions. As a result, over the past two decades, several state-space based numerical methods have been proposed. These methods amalgamate the benefits of both the SSA and the corresponding numerical methods. Examples include the state-space based differential quadrature method (SS-DQM), the state-space based finite element method (SS-FEM), and the state-space based finite difference method (SS-FDM), among others. This section provides a succinct introduction to the state-space based numerical methods and reviews their applications in analyzing the mechanical behaviors of complex structures. These advances are then summarized in **Table 4** for easy reference.

State-Space Based Differential Quadrature Method

To circumvent the limitation that traditional state-space methods, which struggle with free and clamped boundary conditions, Chen et al. [187–189, 196, 197] proposed a semi-analytical approach that combines state-space formulations with the differential quadrature method (DQM). In this SS-DQM, state-space formulations are established with the transfer direction along the thickness of complex structures, and the state equations are discretized along the in-plane direction. This results in new state equations expressed by state variables at discrete points. The final system algorithm equation can then be obtained by applying various boundary conditions. Since its inception, the SS-DQM has been extensively applied to study the static and dynamic responses of laminated structures in beam, plate, and shell configurations, as summarized in this subsection.

Chen et al. [187] utilized the SS-DQM to present elasticity solutions for the free vibration of thin and thick laminated beams under arbitrary boundary conditions. The agreement between the obtained natural frequencies and those from other works [227, 228] validated the accuracy and effectiveness of the SS-DQM. Numerical results indicated that the SS-DQM predicts frequencies more accurately than the approximate beam theories. Subsequently, they investigated the bending and free vibration of arbitrarily thick beams resting on a Pasternak elastic foundation using the SS-DQM [188]. The SS-DQM was found to be superior to other numerical methods and the beam theories when compared with the exact solutions for simply-supported beams. Moreover, the foundation parameter significantly affects the natural frequencies of the beam system. They further employed this method to examine the free vibration of generally laminated beams, where the principal material axes of orthotropic plies do not coincide with the reference axes [189]. This promoted the

TABLE 4 | Summary of state-space based numerical methods to complex structures with different configurations.

State-space based numerical methods	Configuration	Reference
SS-DQM	Beam	[187–195]
SS-DQM	Plate	[196–201]
SS-DQM	Shell	[202]
SS-FEM	Beam	[203–208]
SS-FEM	Plate	[209–218]
SS-FEM	Shell	[219–221]
SS-BEM	Plate	[222]
SS-FDM	shell	[223]
SS-FSM	Plate	[224]
SS-RKP-FSM	Plate	[225]
SS-DSC	Beam	[226]

Acronyms: SS-DQM, state-space based differential quadrature method; SS-FEM, state-space based finite element method; SS-BEM, state-space based boundary element method; SS-FDM, state-space based finite difference method; SS-FSM, state-space based finite strip method; SS-RKP-FSM, state-space based reproducing kernel particle finite strip method; SS-DSC, state-space based discrete singular convolution.

applications of the state-space formulations for generally anisotropic materials in structural analysis. Additionally, the SS-DQM was employed by Xu and Wu [190] to investigate the buckling and free vibration of partially interactive composite beams with interlayer slip under various end conditions. This research served as an extension of the work presented in [17], further demonstrating the versatility of the SS-DQM in addressing diverse end conditions.

The SS-DQM has also been used to study the mechanical behaviors of FGM beams. For instance, Lü and Chen [191] studied the free vibration of FGM orthotropic beams with various boundary conditions using the LAT and the orthotropic elasticity theory of plane stress problem. They concluded that the SS-DQM can deal with arbitrary boundary conditions and meanwhile maintain good accuracy even for higher-order mode shapes of thick plates. Furthermore, the method in conjunction with the LAT allows for easy handling of arbitrary material inhomogeneity. Subsequently, they conducted a 2D thermoelasticity analysis of an FGM thick beam as shown in **Figure 5A**, whose material properties vary continuously and smoothly along the thickness direction, utilizing the SS-DQM along with the LAT [192]. In the numerical calculations, the joint coupling matrices (JCM) were adopted to express the continuity conditions at all mathematical interfaces, thereby avoiding numerical instability. It can be seen from **Figure 5B** that the excellent agreement between the stress field distributions obtained from the SS-DQM and the exact solutions validated the accuracy of the method. Also, these results could serve as useful benchmarks for analyses of FGM thick beams using other numerical methods. Next, based on the SS-DQM in conjunction with the LAT, they presented semi-analytical elasticity solutions for the bending and thermal deformations of a 2D FGM beam with bi-directional inhomogeneity under various end conditions and subjected to the temperature load [193]. The bi-directional inhomogeneity refers to the Young's modulus of the beam varying exponentially along both the thickness and longitudinal directions. Simulation results demonstrated that material inhomogeneity along the longitudinal direction significantly affects the deflection of the beam. Moreover, the introduction of bi-directional FGM materials can potentially reduce thermal stresses of the beam when subjected to temperature load. Li and Shi [194] extended the SS-DQM to investigate the free vibration of a functionally graded piezoelectric material (FGPM) beam under different boundary conditions. This was based on the elasticity and piezoelectricity theory in conjunction with the LAT. The strong agreement between the obtained results and those from the FEM validated the accuracy and reliability of the proposed method. Furthermore, the study revealed that the gradient index, boundary conditions, and beam aspect ratio have substantial influences on the fundamental frequency of FGPM beams. Additionally, the SS-DQM based on the elasticity theory was employed by Alibeigloo and Liew [195] to investigate the bending and free vibration of functionally graded carbon nanotube-reinforced composite (FG-CNTRC) beam sandwiched by two piezoelectric layers serving as the sensor and actuator. The effects of the CNT volume fraction, CNT distribution, beam aspect ratio, and boundary conditions on the deflection and natural frequencies were thoroughly examined and discussed.

Apart from the beam configuration, the SS-DQM has also been used to study the mechanical behaviors of complex structures of plate and shell configurations. For example, Chen and Lee [196] performed a free vibration analysis of a cross-ply laminated plate in cylindrical bending under different boundary conditions using the SS-DQM. Numerical results validated the efficiency of the method when dealing with the clamped and free-end conditions. Subsequently, Chen and Lü [197] developed the SS-DQM to investigate the 3D free vibration of a cross-ply laminated plate with one pair of opposite edges simply-supported and with arbitrary boundary conditions at the other pair of opposite edges. Numerous numerical examples of laminated plates, with both symmetric and unsymmetric cross-plies as well as different boundary conditions, were executed to demonstrate the efficiency of the SS-DQM. Lü et al. [198] also investigated the free vibration of generally supported rectangular Kirchhoff plates with internal rigid line supports under generally boundary conditions, utilizing the SS-DQM in conjunction with the JCM. The study underscored the superiority of the developed method over the traditional SSA. Leveraging the SS-DQM developed from the theory of piezoelasticity, Zhou et al. [199] conducted the static and dynamic analyses of orthotropic piezoelectric laminates, as shown in **Figure 5C**, in cylindrical bending under various boundary conditions. It can be noted from **Figure 5D** that the physical field distributions along the thickness direction obtained from the SS-DQM with the number of discrete points $M = 7$ agreed well with those from the exact solutions, validating the accuracy and effectiveness of the semi-analytical method. Subsequently, they extended the SS-DQM to study the effect of interfacial imperfection, described by a general spring layer model, on the mechanical behaviors of this laminate in cylindrical bending [200]. These studies concluded that the SS-DQM is highly accurate and efficient in dealing with non-simple support and imperfect interface problems. Yas and Moloudi [201]

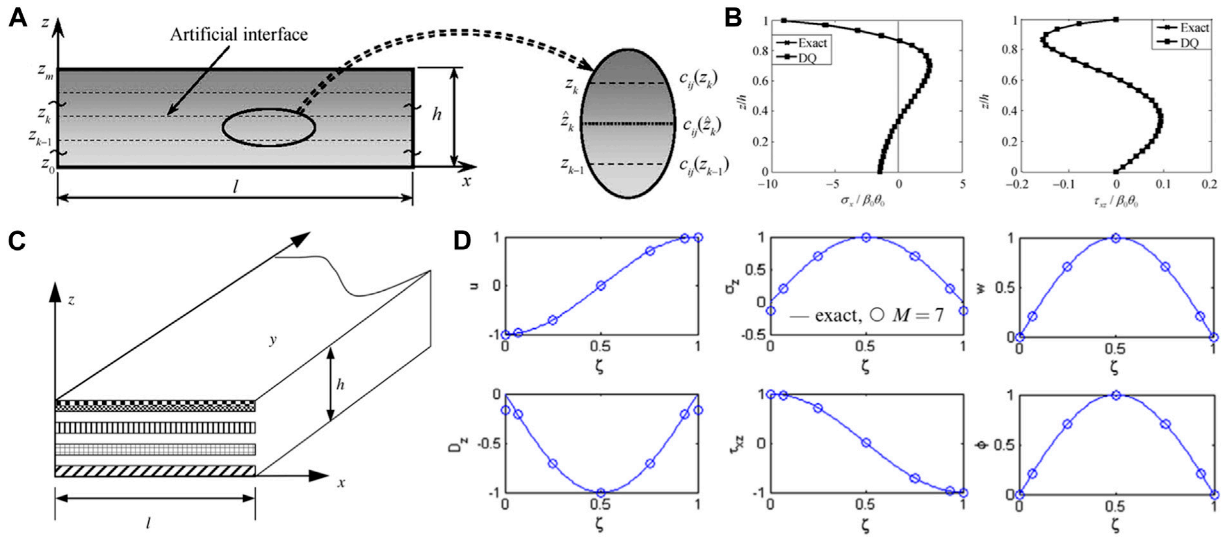

FIGURE 5 | Some cases of applications of the SS-DQM to mechanical analyses of laminated structures: **(A)** functionally graded material beam and its layerwise approximate model [192]; **(B)** comparison of stress fields of the beam in **(A)** along the thickness direction obtained from the SS-DQM and exact solutions; **(C)** piezoelectric laminated plate in cylindrical bending [199]; **(D)** comparison of physical fields of the plate in **(C)** along the thickness direction obtained from the SS-DQM and exact solutions (Reproduced with permission from Lü et al. [192], copyright 2006 by Springer Nature; reproduced with permission from Zhou et al. [199], copyright 2009 by Elsevier).

employed the SS-DQM to analyze the 3D free vibration of multi-directional FGPM annular plates resting on Pasternak elastic foundation under different boundary conditions. The influences of material gradient index along the thickness and radial directions, the Winkler and shear stiffness of the foundation, and boundary conditions on natural frequencies of the FGPM plate were examined. They found that a decrease in the multi-directional gradient index, an increase in plate thickness, and a change from clamped-clamped to clamped-simple all reduce the natural frequencies of the FGPM plate. The relationships between these factors and frequencies significantly depend on the Winkler and shear elastic coefficients.

Few effort has also been devoted to the mechanical analyses of complex shell structures based on the SS-DQM. Alibeigloo [202] performed static and free vibration analyses of an axisymmetric angle-ply laminated cylindrical shell with various boundary conditions. The accuracy of the method for laminated cylindrical shells was verified by comparing the obtained results with the analytical solutions for simply-supported shells. The effect of different boundary conditions on the static and vibration properties of laminated cylindrical shells was then studied.

State-Space Based Finite Element Method

The 3D finite element method serves as a robust and versatile tool for the mechanical analyses of complex structures under various boundary conditions and subjected to diverse external loads. However, it is widely recognized that the conventional 3D FEM only ensures the continuity of displacement components across all element boundaries. The stresses derived from the displacement components, based on the constitutive equations, are invariably discontinuous across the element boundaries. This discontinuity hampers the accurate prediction of the mechanical behaviors of laminated composite structures, particularly the stress distribution at interlayer interfaces, using the 3D FEM. To address this issue of stress discontinuity inherent in traditional 3D FEM, a semi-analytical method known as state-space based finite element method (SS-FEM) has been recently proposed. This method ingeniously integrates the 3D FEM with the state-space formulations [203–209]. In this approach, a laminated structure is partitioned into finite elements in the in-plane direction, while the displacement and stress distributions along the thickness direction are directly solved from the state equation of the structure. The SS-FEM is characterized by two significant features: 1) it ensures continuous displacement and stress distributions across all interlayer interfaces; and 2) the scale of the final system algorithm equation is independent of the number of layers in the laminated structures. This subsection provides a comprehensive overview of the development of the SS-FEMs for various laminated structures and their applications in the analyses of the mechanical behaviors of these structures.

Ali and Ahankar [203, 204] utilized the SS-FEM to investigate the time-domain dynamic responses of a laminated beam. Their findings were instrumental in developing a novel damage detection method for assessing the composite beam with delamination of varying sizes and positions, based on the Eigen-system Realization Algorithm (ERA). The method demonstrated remarkable accuracy in assessing the size and location of the delamination and predicting the reduction of the bending stiffness of the delamination element, with over 97% accuracy. Similarly, Li and Qing [205] presented a semi-analytical solution for the free vibration of laminated composite beams with delamination, using the SS-FEM based on the modified

Hellinger-Reissner (H-R) variational principle. They modeled the delamination between the upper and lower sub-laminates using a non-linear spring-layer model and examined the influences of the delamination size, depth, and boundary conditions on natural frequencies. Their model proved to be a potent and efficient tool for addressing multiple delamination issues in laminated structures. Zhou et al. [206] introduced the SS-FEM for the static response analysis of laminated beams fully sandwiched by piezoelectric actuating layers, based on a modified mixed variational principle. They presented the relationship between the deflection and the secondary converse piezoelectric effect (SCPE), finding that the SCPE significantly affects the deflection of the laminated beam fully covered with piezoelectric actuators. Subsequently, Zhou et al. [207] expanded the aforementioned SS-FEM to investigate the static deformation of laminated curved beams bonded with piezoelectric actuators on the top and bottom surfaces of the substrate. Their model was validated through several complex cases, such as multi-morph piezoelectric curve beams and clamped piezoelectric unimorph C-block beam, and comparison of the obtained results with those from the ANSYS demonstrated the efficiency and accuracy of the new mathematical model. dos Santos et al. [208] proposed a nonlinear SS-FEM based on the von Kármán strain-displacement relation to analyze the nonlinear responses of a laminated beam covered by a piezoelectric sensor, acting as an energy harvester, induced by the flow-structure interaction. The good agreement between the computational results and the experimental data underscored the importance of modeling structural and aerodynamic nonlinearity for accurately depicting the practical electro-aeroelastic behavior of the entire system.

In the context of plate configuration, Zou and Tang [210, 211] presented the semi-analytical solutions for the static deformation and thermal responses of laminated composite plates within the Hamilton system, using the SS-FEM. The quadrilateral element was used as the shape function. Numerical results validated the accuracy and efficiency of the SS-FEM for deformation analyses of laminated plates, particularly for extensive geometry and complex boundary conditions. Built upon the principle of virtual displacement, Sheng and Ye [212] developed the SS-FEM for 3D static deformation analysis of simply-supported three-layered composite plates, using an eight-node quadrilateral element. The comparison of the obtained results with the exact solution in [24] confirmed the accuracy and efficiency of the proposed method. They later applied the SS-FEM to perform stress analysis of orthotropic cross-ply multilayered composite plates, again using an eight-node quadrilateral element [209]. The agreement between the stress fields predicted by the SS-FEM and the existing 3D analytical solutions [229] as well as traditional FEM [230] validated the developed method. Numerical tests also revealed the superiority of the SS-FEM, which provides accurate continuous displacements and transverse stresses along the thickness direction, over the traditional FEM. Ye et al. [213] subsequently conducted the stress analysis of a laminated composite plate with free edges and subjected to transverse and in-plane loads based on the SS-FEM. He et al. [214] explored static displacements and stresses of laminated composite plates under three boundary conditions (free, clamped, and simply-supported) and subjected to thermal load, using the SS-FEM based on the mixed variational principle and an eight-node quadrilateral element. Numerical results again demonstrated the superiority of the SS-FEM over traditional FEMs. Qing et al. [215] developed a general SS-FEM to obtain semi-analytical solutions for the static deformation of magneto-electro-elastic laminated plates, based on the derived modified H-R mixed variational principle for magneto-electro-elastic bodies. They emphasized that the elastic, piezoelectric, and piezomagnetic laminated plates can be considered as special cases of the present model by neglecting certain material constants. They then used this mathematical model, in conjunction with the linear quadrilateral element, to investigate the static response of a simply-supported hybrid laminate and dynamic responses of a clamped aluminum plate with piezoelectric patches [216]. The comparison between the present results and those from the ANSYS validated the efficiency of the developed model. Furthermore, they studied the free vibration of laminated plates reinforced by ribs or beams as the stiffeners [217], and discussed in detail the effect of stiffener height and types on natural frequencies and mode shapes. The same research group also carried out a free vibration analysis of laminated piezoelectric plates with delamination, modeled by a nonlinear spring layer, using the SS-FEM based on the modified H-R variational principle [218]. These numerical simulations indicated that the SS-FEM is a powerful tool for dealing with complex stiffener reinforcement and delamination problems in complex structures.

Additionally, some effort has been invested in applying the SS-FEM to the mechanical analyses of laminated shell structures. For instance, Zhou and Yang [219] conducted a 3D analysis of simply-supported laminated cylindrical shell and panel with arbitrary shell thickness. Their study demonstrated that the proposed method could accurately provide 3D stress and displacement distributions of thin or thick composite shell structures. In a study parallel to the previous research on plate configuration, Sheng and Ye [220] utilized the SS-FEM to generate a semi-analytical solution for stress fields of cross-ply laminated composite shells by using a thin shell element. Simulation results showed that the distributions of stress fields along the thickness direction predicted from the SS-FEM agreed well with those from the exact solutions in [61], suggesting that the method is adept at handling stress problems of laminated shells. Also, the results could serve as a valuable benchmark for testing new shell theory and finite element codes. Built upon the SS-FEM for circular cylindrical shells, Qing et al. [221] proposed a mathematical model for the free vibration of thick double-shell systems, where two cylindrical shells are connected by the springs. Numerical results revealed that the proposed model is not only independent of the thickness and the number of layers in the shell system, but also capable of handling complex support boundaries and the arbitrarily distributed interconnecting springs.

Other State-Space Based Numerical Methods

Apart from the elegant SS-DQM and SS-FEM, other state-space based numerical methods have also been proposed to address

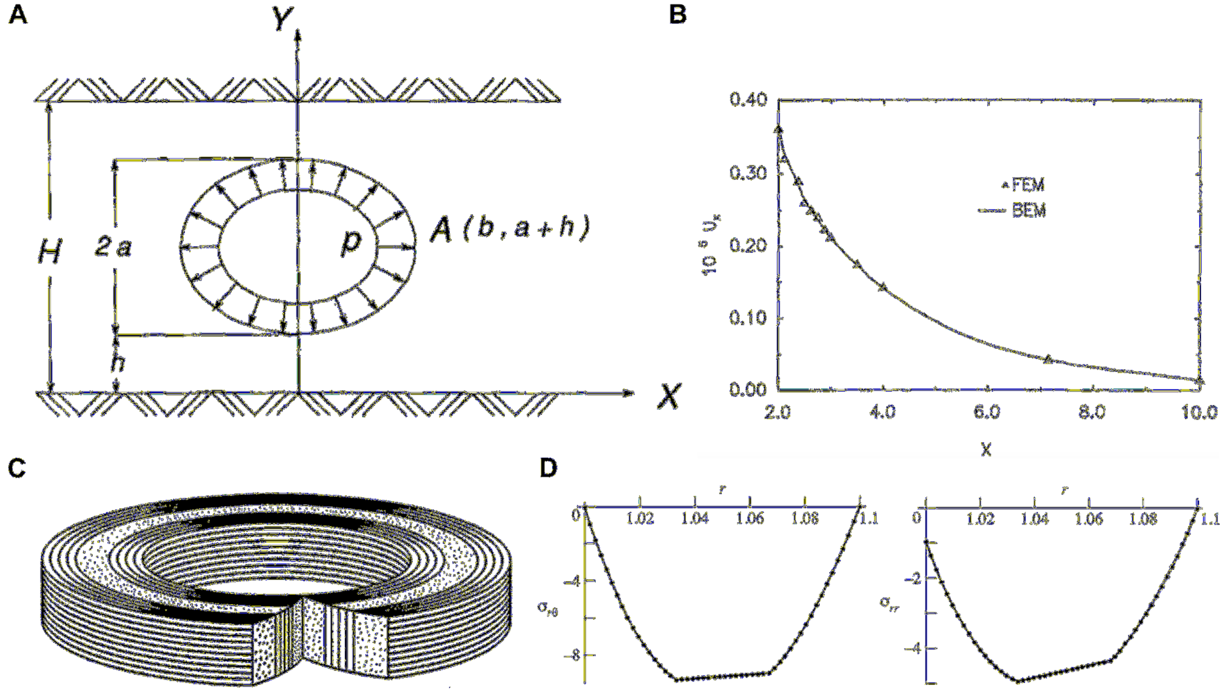

FIGURE 6 | Some cases of application of state-space based numerical method to mechanical analyses of laminated structures: **(A)** an orthotropic elastic layer with an elliptic cavity subjected to uniform internal pressure [222] **(B)** comparison of the horizontal displacement of the structure in **(A)** along the horizontal direction predicted by FEM and SS-BEM; **(C)** cross-section view of 3-ply laminated cylinder [223]; **(D)** variation of stress fields of the cylinder in **(C)** versus radius (solid line for analytical solutions and dot for state-space based finite difference method) (Reproduced with permission from Jiang and Lee; copyright 1994 by ASCE; reproduced with permission from Ruddock and Spencer [223], copyright 1997 by The Royal Society).

structures with unique geometric configurations, such as half-space, or large aspect ratio, or improve the convergence efficiency of numerical methods. In this subsection, we provide a succinct introduction to these state-space based numerical methods. Additionally, we summarize their applications in the mechanical analyses of complex structures.

For instance, Jiang et al. [222] proposed the state-space based boundary element method (SS-BEM) to study the stress distribution of multilayered anisotropic media or half-space under various loads. They employed a procrustean technique to enhance the speed of convergence. On one hand, the SSA can yield the fundamental solutions for orthotropic elastic layers under arbitrary boundary conditions, where the number of equations remains independent of the layer number of the structures. On the other hand, since these fundamental solutions satisfy the boundary conditions and continuity conditions, no elements are required on the boundaries and interfaces. Consequently, when these fundamental solutions are implemented in the spline BEM, the number of elements is significantly reduced without sacrificing numerical accuracy. Undoubtedly, the SS-BEM outperforms the traditional BEM in solving mechanical problems in multilayered anisotropic media, particularly in multilayered half-space. They selected a problem of a spring-supported anisotropic layer containing a uniformly pressurized elliptic cavity under two boundary conditions as the numerical example (schematic of the elastic layer with fixed boundaries was shown in **Figure 6A**). They found that even though only four elements are used for the entire boundary of the elliptic cavity, the results obtained from the SS-BEM still exhibit good agreement with those from the FEM as shown in **Figure 6B**. These results validated the accuracy and efficiency of the proposed SS-BEM.

To address the limitations of the SS-FEM and SS-BEM in analyzing structures with large aspect ratios, which necessitate numerous in-plane elements and consequently result in high computational costs, Ruddock and Spencer [223] introduced the state-space based finite difference method (SS-FDM). This method was specifically developed for the static deformation and stress analyses of laminated or inhomogeneous, anisotropic, elastic, and thermoelastic plates and shells subjected to various mechanical or thermal loads. The SS-FDM employs the finite-difference approximation to discretize the derivatives with respect to the in-plane directions, thereby formulating the elasticity equations as a system of linear first-order ODEs of the stress and displacement variables at the grid-points. By judiciously selecting proper state variables, the state-space formulations were established with the transfer direction along the thickness. Given the inherent efficiency and accuracy of the FDM, the SS-FDM, which is derived from the FDM, also exhibits high efficiency and accuracy, particularly for laminated or inhomogeneous structures. As a demonstration of its effectiveness, Ruddock and Spence conducted a stress analysis of a laminated circular thermoelastic and anisotropic cylinder in plane strain, as shown in **Figure 6C**, under four external loads. It can be seen from **Figure 6D** that the results obtained are in excellent agreement with those from the analytical solutions, thereby validating the viability and accuracy of the SS-FDM. However, they also emphasized the need for further research to enhance the robustness of the SS-FDM and to address the challenges associated with the incorporation of boundary conditions.

The finite strip method (FSM) is another well-established numerical method that is particularly suited for the 3D analysis of homogeneous structures with regular geometric configurations [231, 232]. By integrating the FSM and the SSA, Attallah et al. [224] presented a semi-analytical solution of simply-supported laminated composite plates. The FSM was employed to represent the displacement and stress components along the in-plane direction, utilizing polynomial shape functions, while the SSA was used to determine their distributions along the thickness direction of laminated plates. Recently, Khezri introduced a novel variant of FSM, known as the reproducing kernel particle finite strip method (RKP-FSM), to address issues related to abrupt thickness changes or material discontinuities [233]. Subsequently, the same research group [225] proposed an innovative numerical method for the static analysis of thick and orthotropic rectangular laminated composite plates under various boundary conditions. This method was based on the SSA in conjunction with the newly developed RKP-FSM. Similar to the classical FSM, the RKP-FSM was utilized to approximate the in-plane distributions of the displacements and stresses. The SSA was adopted to predict the displacements and stresses in the thickness direction of laminated plates. The SS-FSM and SS-RKP-FSM retained the advantages of the SSA, namely, the number of unknown variables in these methods is independent of the number of plate layers. A series of numerical examples were conducted to evaluate and validate the convergence, accuracy, and robustness of these methods.

Lastly, it is worth noting the state-space based discrete singular convolution (SS-DSC) algorithm proposed by Xin and Hu [226] for the free vibration analysis of laminated magneto-electro-elastic beams. In the SS-DSC approach, the DSC was employed to discretize the length direction of the beam, thereby transforming the original PDEs of a magneto-electro-elastic body into a set of first-order ODEs. By carefully selecting the state variables, the state equation of the beam was established with the transfer direction along the thickness direction. The thickness domain was then analytically solved by using the SSA. The combination of the DSC with the Taylor's series expansion enabled the implementation of various boundary conditions. Numerical examples demonstrated that the SS-DSC can accurately predict the natural frequencies of laminated composite beams with a low calculation cost. Furthermore, no special arrangement of grid points along the length direction is required, further enhancing the practicality of this method.

CONCLUSION AND OUTLOOKS

Conclusion

This paper presents a comprehensive review of the development and applications of SSAs (also known as transfer matrix methods for state vectors) in the analyses of the mechanical behaviors of complex composite structures of varying configurations, including the beams, plates, shells, and trusses. Utilizing the 3D SSA, simplified SSAs based on structural theories, and state-space based numerical methods, the static deformation and free vibration of laminated structures in beam, plate, and shell forms have been thoroughly investigated. For the sake of completeness, mixed SSAs for the dynamic response analyses of complex periodic structures have also been summarized. It can be generally concluded that SSAs offer a powerful and efficient tool for accurately predicting the static and dynamic responses of laminated and periodic structures. The results obtained could serve as valuable benchmarks for evaluating the developed approximate theoretical methods, numerical methods, and finite element codes. Compared to past research, significant advancements have been made in recent decades, which are summarized as follows:

1) State-space formulations for various material symmetries and geometrical configurations have been established. Leveraging these state-space formulations, transfer matrix methods have been integrated with a range of theoretical or numerical methods, including layerwise approximate technique, wave propagation method, structural theories, finite element method, and differential quadrature method, to facilitate the development of comprehensive SSAs. These approaches provide a more general, efficient, and accurate means of predicting the mechanical behaviors of complex composite and smart structures, particularly in aeronautical applications, under different boundary conditions and subjected to various external loads.
2) Applications of SSAs in the mechanical analyses of complex structures composed of novel smart and composite materials have been documented. These materials include dielectric elastomers, thermo-magneto-electro-elastic coupling materials, and fiber- or graphene-reinforced composite materials. It can be concluded that SSAs can accurately predict the influences of multi-field coupling, material gradient index, external electric, thermal, and mechanical loads, and reinforcement distributions on the mechanical behaviors of these structures.
3) SSAs have also been applied to novel complex structures, such as disordered periodic structures with random unit lengths or random elastic supports, phononic crystal systems with arbitrary geometric configuration, complex structures on various elastic supports or foundations. Novel frequency responses (e.g., high-level response) and wave properties (e.g., stop and pass bands) can be validly revealed and accurately predicted using SSAs.

It is worth mentioning that the state equations of the structures can be solved not only by the transfer matrix method, i.e., the SSA, in this article, but also by the method of reverberation-ray matrix (MRRM) [234–236]. The MRRM is particularly suitable for steady-state or transient response analyses of a truss-type complex structure or a multi-branched frame structure. Different from the system matrix in SSAs, the reverberation-ray matrix consists of two global matrices, namely, the global scattering matrix and the global phase matrix. The detailed differences between the SSA and MRRM can refer to Ref. [237]. In addition, in SSAs, the standard transfer matrix relating the displacements and stresses at the layer top and bottom would become problematic when the layer thickness or frequency increases, resulting in numerical instability of the transfer matrix method. Consequently, the stiffness matrix method (SMM) has been developed to solve this instability [238]. In SMM, the stresses at the top and bottom of the layered structure are expressed by the displacements at the top and bottom through the global stiffness matrix derived through a recursive algorithm based on layer stiffness matrix. In summary, the MRRM and SMM can be a viable alternative to SSAs for the mechanical analyses of certain complex structures in aeronautical and engineering applications.

Outlooks

Despite the substantial advancements that have been made, several significant issues listed in the following remain to be addressed in the future research.

1) SSAs have proven to be efficient and accurate in predicting the mechanical behaviors of low-frequency modes, including bending and extension, in complex composite and smart structures of various geometric configurations. However, high-frequency modes such as thickness-shear and thickness extensional modes exhibit more complex deformation patterns along both in-plane and thickness directions than their low-frequency counterparts [239, 240]. Consequently, the development of SSAs for analyzing the mechanical behaviors of high-frequency modes in these complex structures presents a valuable and intriguing avenue for the future research.
2) State-space techniques have been reported to accurately predict the wave properties of beam-type phononic crystals composed of elastic materials. However, the wave properties of these elastic material phononic crystals remain unalterable once their geometric parameters and configurations are established. Recently, the concept of dielectric elastomer phononic crystals has been introduced, offering the potential for tunable band structures through the application of mechanical and electrical excitations [99, 241, 242]. Consequently, there is a pressing need for the development of appropriate SSAs that can accurately predict the band structures of dielectric elastomer phononic crystals under mechanical and electrical loads. This area of research warrants further investigation.
3) While certain specialized state-space based numerical methods have been proposed to investigate the mechanical behaviors of complex structures of specialized geometric configurations or boundary conditions, their application to such structures remains limited. Consequently, the robustness and versatility of these methods remain uncertain, and potential numerical difficulties may arise in certain applications. This highlights a significant need for the further refinement, rigorous testing, and robustness validation of these state-space based numerical methods. This is a crucial step towards ensuring their reliability and applicability in a wide range of complex structural analyses.
4) The skin-stringer panel system, commonly used in aircraft fuselage, has been modeled as a 1D periodic beam structure on multiple elastic supports. Its dynamic responses have been effectively predicted using the transfer matrix method based on state-space formulations. However, the neglect of the curvature inherent of practical skin panels in the theoretical model may significantly influence the dynamic responses of the actual skin panel system [157]. As a result, investigating the dynamic behaviors of a curved periodic beam structure using the transfer matrix method for state vectors presents a valuable and challenging future research topic.
5) Fluid-solid interaction remains a crucial and challenging issue, particularly due to the extensive use of cylindrical shell and pipeline structures in maritime transportation. There is a pressing need to establish appropriate mechanical models that describe the interaction between the solid structures and various fluids. Built upon these models, the analyses of frequency responses and wave propagation in cylindrical shell structures or periodic pipe structures using SSAs is of vital importance. Such analysis is particularly relevant to the health monitoring and damage warning of these structures, and thus warrants further attention in the future research.
6) With the advancement of micro-electro-mechanical systems (MEMS) and nano-electro-mechanical systems (NEMS), novel laminated structures and devices at the nano or micro scale have been recently fabricated, exhibiting unique, size-dependent behaviors [243, 244]. However, current research on size-dependent behaviors of these miniaturized laminated structures is still in its infancy. Therefore, the development of SSAs in conjunction with various modified continuum theories [243, 245] to explore the size-dependent behaviors of such nanostructures are worth in-depth investigation in future research endeavors.

AUTHOR CONTRIBUTIONS

ZZ: Conceptualization, writing–original draft, writing–review and editing, funding acquisition. XS: Writing–review and editing. YS: Writing–review and editing. WC: Supervision, conceptualization, writing–review and editing, funding acquisition.

All authors contributed to the article and approved the submitted version.

FUNDING

The work was supported by the National Natural Science Foundation of China (Nos 12192211, 12072315, and 11872329), the Natural Science Foundation of Zhejiang Province (No. LD21A020001), the 111 Project (No. B21034), the National Postdoctoral Program for Innovative Talents (BX2021261), and the China Postdoctoral Science Foundation Funded Project (No. 2022M722745). This work was also partly supported by the specialized research projects of Huanjiang Laboratory, Zhuji, Zhejiang Province.

CONFLICT OF INTEREST

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

REFERENCES

1. Li D. *Analysis of Composite Laminates: Theories and Their Applications*. Amsterdam, Netherlands: Elsevier (2022).
2. Jawaid M, Thariq M, editors. *Sustainable Composites for Aerospace Applications*. Sawston, United Kingdom: Woodhead publishing (2018).
3. Rezende MC, Botelho EC. O Uso de Compósitos Estruturais na Indústria Aeroespacial. *Polímeros* (2000) 10:4–10. doi:10.1590/s0104-14282000000200003
4. Ezugwu EO. Key Improvements in the Machining of Difficult-To-Cut Aerospace Superalloys. *Int J Machine Tools Manufacture* (2005) 45(12-13):1353–67. doi:10.1016/j.ijmachtools.2005.02.003
5. Loewy RG. Recent Developments in Smart Structures With Aeronautical Applications. *Smart Mater Structures* (1997) 6(5):R11–42. doi:10.1088/0964-1726/6/5/001
6. Chen WQ, Wu YF, Xu RQ. State Space Formulation for Composite Beam-Columns With Partial Interaction. *Composites Sci Tech* (2007) 67(11-12): 2500–12. doi:10.1016/j.compscitech.2006.12.013
7. Zhu W, Yang H, Liu W, Shi B, Ling Z, Tao H. Experimental Investigation on Innovative Connections for Timber–Concrete Composite Systems. *Construction Building Mater* (2019) 207:345–56. doi:10.1016/j.conbuildmat.2019.02.079
8. Kapania RK, Raciti S. Recent Advances in Analysis of Laminated Beams and Plates, Part II: Vibrations and Wave Propagation. *AIAA J* (1989) 27(7): 935–46. doi:10.2514/3.59909
9. Noor AK, Burton WS. Assessment of Computational Models for Multilayered Composite Shells. *Appl Mech Rev* (1990) 43:67–97. doi:10.1115/1.3119162
10. Thomson WT. Matrix Solution for the Vibration of Nonuniform Beams. *J Appl Mech* (1950) 17:337–9. doi:10.1115/1.4010137
11. Thomson WT. Transmission of Elastic Waves Through a Stratified Solid Medium. *J Appl Phys* (1950) 21(2):89–93. doi:10.1063/1.1699629
12. Kennett B. *Seismic Wave Propagation in Stratified Media*. The Australian National University: ANU Press (2009).
13. Bahar LY. A State Space Approach to Elasticity. *J Franklin Inst* (1975) 299(1): 33–41. doi:10.1016/0016-0032(75)90082-4
14. Chen WQ, Ying J, Cai JB, Ye GR. Benchmark Solution of Laminated Beams With Bonding Imperfections. *AIAA J* (2004) 42(2):426–9. doi:10.2514/1.4776
15. Ying J, Lü CF, Chen WQ. Two-Dimensional Elasticity Solutions for Functionally Graded Beams Resting on Elastic Foundations. *Compos Structures* (2008) 84(3):209–19. doi:10.1016/j.compstruct.2007.07.004
16. Bian ZG, Lim CW, Chen WQ. On Functionally Graded Beams With Integrated Surface Piezoelectric Layers. *Compos Structures* (2006) 72(3): 339–51. doi:10.1016/j.compstruct.2005.01.005
17. Xu R, Wu YF. Two-Dimensional Analytical Solutions of Simply Supported Composite Beams With Interlayer Slips. *Int J Sol Structures* (2007) 44(1): 165–75. doi:10.1016/j.ijsolstr.2006.04.027
18. Xu R, Wu YF. Analytical Study of Beams Strengthened by Adhesively Bonded Reinforcement With Variable Properties Using State Space Method. *Composites Sci Tech* (2009) 69(11-12):1912–8. doi:10.1016/j.compscitech.2009.04.011
19. Yan W, Ying J, Chen WQ. Response of Laminated Adaptive Composite Beams With Viscoelastic Interfaces. *Compos Structures* (2006) 74(1):70–9. doi:10.1016/j.compstruct.2005.03.009
20. Alibeigloo A. Thermoelasticity Analysis of Functionally Graded Beam With Integrated Surface Piezoelectric Layers. *Compos Structures* (2010) 92(6): 1535–43. doi:10.1016/j.compstruct.2009.10.030
21. Qian H, Qiu Y, Lu C, Sha X. Analytical Solutions for Laminated Beams Subjected to Non-Uniform Temperature Boundary Conditions. *Compos Structures* (2022) 282:115044. doi:10.1016/j.compstruct.2021.115044
22. Qian H, Qiu Y, Lu C, Yang Y, Sha X. Thermal Analysis for Clamped Laminated Beams With Non-Uniform Temperature Boundary Conditions. *Thin-Walled Structures* (2022) 179:109693. doi:10.1016/j.tws.2022.109693
23. Yan W, Cai JB, Chen WQ. Monitoring Interfacial Defects in a Composite Beam Using Impedance Signatures. *J Sound Vibration* (2009) 326(1-2): 340–52. doi:10.1016/j.jsv.2009.04.030
24. Fan JR, Ye JQ. An Exact Solution for the Statics and Dynamics of Laminated Thick Plates With Orthotropic Layers. *Int J Sol Structures* (1990) 26(5-6): 655–62.
25. Chen WQ, Cai JB, Ye GR. Exact Solutions of Cross-Ply Laminates With Bonding Imperfections. *AIAA J* (2003) 41(11):2244–50. doi:10.2514/2.6817
26. Chen WQ, Lee KY. Three-Dimensional Exact Analysis of Angle-Ply Laminates in Cylindrical Bending With Interfacial Damage via State-Space Method. *Compos Structures* (2004) 64(3-4):275–83. doi:10.1016/j.compstruct.2003.08.010
27. Chen WQ, Cai JB, Ye GR, Wang YF. Exact Three-Dimensional Solutions of Laminated Orthotropic Piezoelectric Rectangular Plates Featuring Interlaminar Bonding Imperfections Modeled by a General spring Layer. *Int J Sol Structures* (2004) 41(18-19):5247–63. doi:10.1016/j.ijsolstr.2004.03.010
28. Chen WQ, Lee KY. Exact Solution of Angle-Ply Piezoelectric Laminates in Cylindrical Bending With Interfacial Imperfections. *Compos Structures* (2004) 65(3-4):329–37. doi:10.1016/j.compstruct.2003.11.008
29. Chen WQ, Ying J, Cai JB, Ye GR. Benchmark Solution of Imperfect Angle-Ply Laminated Rectangular Plates in Cylindrical Bending With Surface Piezoelectric Layers as Actuator and Sensor. *Comput Structures* (2004) 82(22):1773–84. doi:10.1016/j.compstruc.2004.05.011
30. Fan JR, Ye JQ. A Series Solution of the Exact Equation for Thick Orthotropic Plates. *Int J Sol Structures* (1990) 26(7):773–8.
31. Zenkour AM. Three-Dimensional Elasticity Solution for Uniformly Loaded Cross-Ply Laminates and Sandwich Plates. *J Sandwich Structures Mater* (2007) 9(3):213–38. doi:10.1177/1099636207065675
32. Chen WQ, Lee KY. Time-Dependent Behaviors of Angle-Ply Laminates With Viscous Interfaces in Cylindrical Bending. *Eur J Mechanics-A/Solids* (2004) 23(2):235–45. doi:10.1016/j.euromechsol.2003.12.004
33. Chen WQ, Kim GW, Lee KY. The Effect of Viscous Interfaces on the Bending of Orthotropic Rectangular Laminates. *Compos Structures* (2005) 67(3): 323–31. doi:10.1016/j.compstruct.2004.01.014
34. Qing G, Wang F, Liu Y. State Space Approach for Energy Release Rate Analysis of Delaminated Laminates With Stiffeners. *AIAA J* (2011) 49(10): 2123–9. doi:10.2514/1.j050610
35. Xu K, Noor AK, Tang YY. Three-Dimensional Solutions for Coupled Thermoelectroelastic Response of Multilayered Plates. *Comp Methods Appl Mech Eng* (1995) 126(3-4):355–71. doi:10.1016/0045-7825(95)00825-l
36. Lee JS, Jiang LZ. Exact Electroelastic Analysis of Piezoelectric Laminae via State Space Approach. *Int J Sol Structures* (1996) 33(7):977–90. doi:10.1016/0020-7683(95)00083-6

37. Cheng ZQ, Batra RC. Three-Dimensional Asymptotic Analysis of Multiple-Electroded Piezoelectric Laminates. *AIAA J* (2000) 38(2):317–24. doi:10.2514/2.959
38. Benjeddou A, Deu JF. Piezoelectric Transverse Shear Actuation and Sensing of Plates, Part 2: Application and Analysis. *J Intell Mater Syst Structures* (2001) 12(7):451–67. doi:10.1106/j8dx-r3at-ea6g-85k7
39. Vel SS, Batra RC. Exact Solution for the Cylindrical Bending of Laminated Plates With Embedded Piezoelectric Shear Actuators. *Smart Mater Structures* (2001) 10(2):240–51. doi:10.1088/0964-1726/10/2/309
40. Vel SS, Batra RC. Exact Solution for Rectangular Sandwich Plates With Embedded Piezoelectric Shear Actuators. *AIAA J* (2001) 39(7):1363–73. doi:10.2514/2.1455
41. Ding HJ, Xu RQ, Chi YW, Chen WQ. Free Axisymmetric Vibration of Transversely Isotropic Piezoelectric Circular Plates. *Int J Sol Structures* (1999) 36(30):4629–52.
42. Wang JG. State Vector Solutions for Nonaxisymmetric Problem of Multilayered Half Space Piezoelectric Medium. *Sci China Ser A: Math* (1999) 42(12):1323–31. doi:10.1007/bf02876034
43. Wang JG, Fang S. The State Vector Methods of Axisymmetric Problems for Multilayered Anisotropic Elastic System. *Mech Res Commun* (1999) 26(6): 673–8. doi:10.1016/s0093-6413(99)00077-4
44. Wang JG. The State Space Solution of Layered Space Axisymmetric Piezoelectric Media. *Acta Mech Sinica* (2001) 33(1):115–20.
45. Wang J, Fang S, Chen L. The State Vector Methods for Space Axisymmetric Problems in Multilayered Piezoelectric Media. *Int J Sol Structures* (2002) 39(15):3959–70. doi:10.1016/s0020-7683(02)00267-6
46. Sheng HY, Wang H, Ye JQ. State Space Solution for Thick Laminated Piezoelectric Plates With Clamped and Electric Open-Circuited Boundary Conditions. *Int J Mech Sci* (2007) 49(7):806–18. doi:10.1016/j.ijmecsci.2006.11.012
47. Lezgy-Nazargah M. A Three-Dimensional Exact State-Space Solution for Cylindrical Bending of Continuously Non-Homogenous Piezoelectric Laminated Plates With Arbitrary Gradient Composition. *Arch Mech* (2015) 67(1):25–51.
48. Wang J, Chen L, Fang S. State Vector Approach to Analysis of Multilayered Magneto-Electro-Elastic Plates. *Int J Sol Structures* (2003) 40(7):1669–80. doi:10.1016/s0020-7683(03)00027-1
49. Fan JR, Ye JQ. Exact Solutions for Axisymmetric Vibration of Laminated Circular Plates. *J Eng Mech* (1990) 116(4):920–7.
50. Qing G, Liu Y, Guo Q, Zhang D. Dynamic Analysis for Three-Dimensional Laminated Plates and Panels With Damping. *Int J Mech Sci* (2008) 50(1): 83–91. doi:10.1016/j.ijmecsci.2007.05.002
51. Xu K, Noor AK, Tang YY. Three-Dimensional Solutions for Free Vibrations of Initially-Stressed Thermoelectroelastic Multilayered Plates. *Comp Methods Appl Mech Eng* (1997) 141(1-2):125–39. doi:10.1016/s0045-7825(96)01065-1
52. Chen WQ, Xu RQ, Ding HJ. On Free Vibration of a Piezoelectric Composite Rectangular Plate. *J Sound Vibration* (1998) 4(218):741–8. doi:10.1006/jsvi.1998.1850
53. Ding HJ, Chen WQ, Xu RQ. New State Space Formulations for Transversely Isotropic Piezoelasticity With Application. *Mech Res Commun* (2000) 27(3): 319–26. doi:10.1016/s0093-6413(00)00098-7
54. Qing G, Xu J, Li P, Qiu J. A New Efficient Numerical Method and the Dynamic Analysis of Composite Laminates With Piezoelectric Layers. *Compos Structures* (2007) 78(4):457–67. doi:10.1016/j.compstruct.2005.11.002
55. Soldatos KP, Ye JQ. Three-Dimensional Static, Dynamic, Thermoelastic and Buckling Analysis of Homogeneous and Laminated Composite Cylinders. *Compos Structures* (1994) 29(2):131–43. doi:10.1016/0263-8223(94)90095-7
56. Ding KW, Fan J. Exact Solution for Axisymmetric Problem of Thick Laminated Continuous Closed Cylindrical Shells. *Eng Mech* (1994) 11(2): 8–19. (in Chinese).
57. Ding KW, Tang L, Fan J. Exact Analysis for Axisymmetric Vibration and Buckling of the Thick Laminated Closed Cylindrical Shells in a Hamilton System. *J Sound Vibration* (1997) 206(3):435–41. doi:10.1006/jsvi.1997.1085
58. Fan J, Zhang J. Analytical Solutions for Thick, Doubly Curved, Laminated Shells. *J Eng Mech* (1992) 118(7):1338–56. doi:10.1061/(asce)0733-9399(1992)118:7(1338)
59. Chen WQ, Wang YF, Cai JB, Ye G. Three-Dimensional Analysis of Cross-Ply Laminated Cylindrical Panels With Weak Interfaces. *Int J Sol Structures* (2004) 41(9-10):2429–46. doi:10.1016/j.ijsolstr.2003.12.018
60. Cai JB, Chen WQ, Ye GR. Effect of Interlaminar Bonding Imperfections on the Behavior of Angle-Ply Laminated Cylindrical Panels. *Composites Sci Tech* (2004) 64(12):1753–62. doi:10.1016/j.compscitech.2003.12.009
61. Ye JQ, Soldatos KP. Three-Dimensional Stress Analysis of Orthotropic and Cross-Ply Laminated Hollow Cylinders and Cylindrical Panels. *Comp Methods Appl Mech Eng* (1994) 117(3-4):331–51. doi:10.1016/0045-7825(94)90121-x
62. Soldatos KP, Hadjigeorgiou VP. Three-Dimensional Solution of the Free Vibration Problem of Homogeneous Isotropic Cylindrical Shells and Panels. *J Sound Vibration* (1990) 137(3):369–84. doi:10.1016/0022-460x(90)90805-a
63. Soldatos KP. An Iterative Solution of a Bessel Equation Based on Torsional Vibrations of Orthotropic Hollow Cylinders. *J Sound Vibration* (1991) 151(1):149–52. doi:10.1016/0022-460x(91)90657-6
64. Hawkes TD, Soldatos KP. Three-Dimensional Axisymmetric Vibrations of Orthotropic and Cross-Ply Laminated Hollow Cylinders. *AIAA J* (1992) 30(4):1089–98. doi:10.2514/3.11031
65. Ye JQ, Soldatos KP. Three-Dimensional Vibration of Laminated Cylinders and Cylindrical Panels With Symmetric or Antisymmetric Cross-Ply Lay-Up. *Composites Eng* (1994) 4(4):429–44. doi:10.1016/s0961-9526(09)80016-6
66. Ding K, Tang L. Three-Dimensional Free Vibration of Thick Laminated Cylindrical Shells With Clamped Edges. *J Sound Vibration* (1999) 1(220): 171–7. doi:10.1006/jsvi.1998.1897
67. Chen WQ, Bian ZG, Ding HJ. Three-Dimensional Vibration Analysis of Fluid-Filled Orthotropic FGM Cylindrical Shells. *Int J Mech Sci* (2004) 46(1): 159–71. doi:10.1016/j.ijmecsci.2003.12.005
68. Zhu F, Wu B, Destrade M, Chen W. Electrostatically Tunable Axisymmetric Vibrations of Soft Electro-Active Tubes. *J Sound Vibration* (2020) 483: 115467. doi:10.1016/j.jsv.2020.115467
69. Ye JQ, Soldatos KP. Three-Dimensional Buckling Analysis of Laminated Composite Hollow Cylinders and Cylindrical Panels. *Int J Sol Structures* (1995) 32(13):1949–62. doi:10.1016/0020-7683(94)00217-k
70. Talebitooti R, Ahmadi R, Shojaeefard MH. Three-Dimensional Wave Propagation on Orthotropic Cylindrical Shells With Arbitrary Thickness Considering State Space Method. *Compos Structures* (2015) 132:239–54. doi:10.1016/j.compstruct.2015.05.023
71. Wu B, Su Y, Chen W, Zhang C. On Guided Circumferential Waves in Soft Electroactive Tubes Under Radially Inhomogeneous Biasing fields. *J Mech Phys Sol* (2017) 99:116–45. doi:10.1016/j.jmps.2016.11.004
72. Wu B, Su Y, Liu D, Chen W, Zhang C. On Propagation of Axisymmetric Waves in Pressurized Functionally Graded Elastomeric Hollow Cylinders. *J Sound Vibration* (2018) 421:17–47. doi:10.1016/j.jsv.2018.01.055
73. Chen WQ, Ding HJ. A State-Space-Based Stress Analysis of a Multilayered Spherical Shell With Spherical Isotropy. *J Appl Mech* (2001) 68(1):109–14. doi:10.1115/1.1343913
74. Chen WQ, Ding HJ, Xu R. Three-Dimensional Static Analysis of Multi-Layered Piezoelectric Hollow Spheres via the State Space Method. *Int J Sol Structures* (2001) 38(28-29):4921–36. doi:10.1016/s0020-7683(00)00314-0
75. Chen WQ, Ding HJ. Free Vibration of Multi-Layered Spherically Isotropic Hollow Spheres. *Int J Mech Sci* (2001) 43(3):667–80. doi:10.1016/s0020-7403(00)00044-8
76. Chen WQ, Wang LZ, Lu Y. Free Vibrations of Functionally Graded Piezoceramic Hollow Spheres With Radial Polarization. *J Sound Vibration* (2002) 251(1):103–14. doi:10.1006/jsvi.2001.3973
77. Mao R, Wu B, Carrera E, Chen W. Electrostatically Tunable Small-Amplitude Free Vibrations of Pressurized Electro-Active Spherical Balloons. *Int J Non-Linear Mech* (2019) 117:103237. doi:10.1016/j.ijnonlinmec.2019.103237
78. Cheng ZQ, Lim CW, Kitipornchai S. Three-Dimensional Asymptotic Approach to Inhomogeneous and Laminated Piezoelectric Plates. *Int J Sol Structures* (2000) 37(23):3153–75. doi:10.1016/s0020-7683(99)00036-0
79. He XQ, Ng TY, Sivashanker S, Liew KM. Active Control of FGM Plates With Integrated Piezoelectric Sensors and Actuators. *Int J Sol Structures* (2001) 38(9):1641–55. doi:10.1016/s0020-7683(00)00050-0

80. Liew KM, He XQ, Ng TY, Kitipornchai S. Active Control of FGM Shells Subjected to a Temperature Gradient via Piezoelectric Sensor/Actuator Patches. *Int J Numer Methods Eng* (2002) 55(6):653–68. doi:10.1002/nme.519
81. Pagano NJ. Exact Solutions for Composite Laminates in Cylindrical Bending. *J Compos Mater* (1969) 3(3):398–411. doi:10.1177/002199836900300304
82. Pagano NJ. Exact Solutions for Rectangular Bidirectional Composites and Sandwich Plates. *J Compos Mater* (1970) 4(1):20–34. doi:10.1177/002199837000400102
83. Pagano NJ. Influence of Shear Coupling in Cylindrical. Bending of Anisotropic Laminates. *J Compos Mater* (1970) 4(3):330–43. doi:10.1177/002199837000400305
84. Pagano NJ, Hatfield SJ. Elastic Behavior of Multilayered Bidirectional Composites. *AIAA J* (1972) 10(7):931–3.
85. Heyliger P. Static Behavior of Laminated Elastic/Piezoelectric Plates. *AIAA J* (1994) 32(12):2481–4. doi:10.2514/3.12321
86. Heyliger P, Brooks S. Free Vibration of Piezoelectric Laminates in Cylindrical Bending. *Int J Sol Structures* (1995) 32(20):2945–60. doi:10.1016/0020-7683(94)00270-7
87. Heyliger P, Brooks S. Exact Solutions for Laminated Piezoelectric Plates in Cylindrical Bending. *J Appl Mech* (1996) 63(4):903–10. doi:10.1115/1.2787245
88. Heyliger P. Exact Solutions for Simply Supported Laminated Piezoelectric Plates. *J Appl Mech* (1997) 64(2):299–306. doi:10.1115/1.2787307
89. Cheng ZQ, Kennedy D, Williams WF. Effect of Interfacial Imperfection on Buckling and Bending Behavior of Composite Laminates. *AIAA J* (1996) 34(12):2590–5. doi:10.2514/3.13443
90. Benjeddou A, Deu JF. Piezoelectric Transverse Shear Actuation and Sensing of Plates, Part 1: A Three-Dimensional Mixed State Space Formulation. *J Intell Mater Syst Structures* (2001) 12(7):435–49. doi:10.1177/10453890122145258
91. Pan E. Exact Solution for Simply Supported and Multilayered Magneto-Electro-Elastic Plates. *J Appl Mech* (2001) 68(4):608–18. doi:10.1115/1.1380385
92. Chen WQ, Ding HJ. Natural Frequencies of Fluid-Filled Transversely Isotropic Cylindrical Shells. *Int J Mech Sci* (1999) 41(6):677–84. doi:10.1016/s0020-7403(98)00088-5
93. Raju PP. On Shallow Shells of Transversely Isotropic Materials. *J Press Vessel Tech* (1975) 97(3):185–91. doi:10.1115/1.3454293
94. Maiti M. Stresses in Anisotropic Nonhomogeneous Sphere. *J Eng Mech Division* (1975) 101(1):101–8. doi:10.1061/jmcea3.0001985
95. Montagner JP, Anderson DL. Constrained Reference Mantle Model. *Phys earth Planet Interiors* (1989) 58(2-3):205–27. doi:10.1016/0031-9201(89)90055-1
96. Ding ZY, Zou DQ, Ding HJ. A Study of Effects of the Earth's Radial Anisotropy on the Tidal Stress Field. *Tectonophysics* (1996) 258(1-4):103–14. doi:10.1016/0040-1951(95)00186-7
97. Heyliger P, Wu YC. Electroelastic Fields in Layered Piezoelectric Spheres. *Int J Eng Sci* (1999) 37(2):143–61. doi:10.1016/s0020-7225(98)00068-8
98. Pelrine R, Kornbluh R, Pei Q, Joseph J. High-Speed Electrically Actuated Elastomers With Strain Greater Than 100%. *Science* (2000) 287(5454):836–9. doi:10.1126/science.287.5454.836
99. Zhao Z, Chen Y, Hu X, Bao R, Wu B, Chen WQ. Vibrations and Waves in Soft Dielectric Elastomer Structures. *Int J Mech Sci* (2022) 239:107885. doi:10.1016/j.ijmecsci.2022.107885
100. Khdeir AA, Reddy JN. Free Vibration of Cross-Ply Laminated Beams With Arbitrary Boundary Conditions. *Int J Eng Sci* (1994) 32(12):1971–80. doi:10.1016/0020-7225(94)90093-0
101. Khdeir AA, Redd JN. Buckling of Cross-Ply Laminated Beams With Arbitrary Boundary Conditions. *Compos Structures* (1997) 37(1):1–3. doi:10.1016/s0263-8223(97)00048-2
102. Trinh LC, Vo TP, Thai HT, Nguyen TK. An Analytical Method for the Vibration and Buckling of Functionally Graded Beams Under Mechanical and Thermal Loads. *Composites B: Eng* (2016) 100:152–63. doi:10.1016/j.compositesb.2016.06.067
103. Khdeir AA. Dynamic Response of Antisymmetric Cross-Ply Laminated Composite Beams With Arbitrary Boundary Conditions. *Int J Eng Sci* (1996) 34(1):9–19. doi:10.1016/0020-7225(95)00080-1
104. Trinh LC, Vo TP, Osofero AI, Lee J. Fundamental Frequency Analysis of Functionally Graded Sandwich Beams Based on the State Space Approach. *Compos Structures* (2016) 156:263–75. doi:10.1016/j.compstruct.2015.11.010
105. Khdeir AA, Reddy JN. An Exact Solution for the Bending of Thin and Thick Cross-Ply Laminated Beams. *Compos Structures* (1997) 37(2):195–203. doi:10.1016/s0263-8223(97)80012-8
106. Khdeir AA, Reddy JN. Jordan Canonical Form Solution for Thermally Induced Deformations of Cross-Ply Laminated Composite Beams. *J Therm Stresses* (1999) 22(3):331–46.
107. Khdeir AA. Thermal Buckling of Cross-Ply Laminated Composite Beams. *Acta Mechanica* (2001) 149(1-4):201–13. doi:10.1007/bf01261672
108. Aldraihem OJ, Khdeir AA. Smart Beams With Extension and Thickness-Shear Piezoelectric Actuators. *Smart Mater Structures* (2000) 9(1):1–9. doi:10.1088/0964-1726/9/1/301
109. Sahmani S, Ansari R. Nonlocal Beam Models for Buckling of Nanobeams Using State-Space Method Regarding Different Boundary Conditions. *J Mech Sci Tech* (2011) 25:2365–75. doi:10.1007/s12206-011-0711-6
110. Pota HR, Alberts TE. Multivariable Transfer Functions for a Slewing Piezoelectric Laminate Beam. *J Dynamic Syst Meas Control* (1995) 117(3):352–9. doi:10.1115/1.2799126
111. Palmeri A, Adhikari S. A Galerkin-Type State-Space Approach for Transverse Vibrations of Slender Double-Beam Systems With Viscoelastic Inner Layer. *J Sound Vibration* (2011) 330(26):6372–86. doi:10.1016/j.jsv.2011.07.037
112. Palmeri A, Ntotsios E. Transverse Vibrations of Viscoelastic Sandwich Beams via Galerkin-Based State-Space Approach. *J Eng Mech* (2016) 142(7):04016036. doi:10.1061/(asce)em.1943-7889.0001069
113. Li Y, Xiong F, Xie L, Sun L. State-Space Approach for Transverse Vibration of Double-Beam Systems. *Int J Mech Sci* (2021) 189:105974. doi:10.1016/j.ijmecsci.2020.105974
114. Khdeir AA, Aldraihem OJ. Free Vibration of Sandwich Beams With Soft Core. *Compos Structures* (2016) 154:179–89. doi:10.1016/j.compstruct.2016.07.045
115. Srinivasan R, Dattaguru B, Singh G. Exact Solutions for Laminated Composite Beams Using a Unified State Space Formulation. *Int J Comput Methods Eng Sci Mech* (2019) 20(4):319–34. doi:10.1080/15502287.2019.1644394
116. Ramaprasad S, Dattaguru B, Singh G. Exact Solutions for Thin-Walled Composite Open Section Beams Using a Unified State Space Coupled Field Formulation. *Mech Based Des Structures Machines* (2022) 51(11):5983–6007. doi:10.1080/15397734.2021.2024846
117. Fazeli S, Stokes-Griffin C, Gilbert J, Compston P. An Analytical Solution for the Vibrational Response of Stepped Smart Cross-Ply Laminated Composite Beams With Experimental Validation. *Compos Structures* (2021) 266:113823. doi:10.1016/j.compstruct.2021.113823
118. Khdeir AA. A Remark on the State-Space Concept Applied to Bending, Buckling and Free Vibration of Composite Laminates. *Comput Structures* (1996) 59(5):813–7. doi:10.1016/0045-7949(95)00330-4
119. Khdeir AA, Reddy JN. Analytical Solutions of Refined Plate Theories of Cross-Ply Composite Laminates. *J Press Vessel Tech* (1991) 113(4):570–8. doi:10.1115/1.2928797
120. Khdeir AA. Comparison Between Shear Deformable and Kirchhoff Theories for Bending, Buckling and Vibration of Antisymmetric Angle-Ply Laminated Plates. *Compos Structures* (1989) 13(3):159–72. doi:10.1016/0263-8223(89)90001-9
121. Khdeir AA. An Exact Approach to the Elastic State of Stress of Shear Deformable Antisymmetric Angle-Ply Laminated Plates. *Compos Structures* (1989) 11(4):245–58. doi:10.1016/0263-8223(89)90090-1
122. Librescu L, Khdeir AA. Analysis of Symm≐tric Cross-Ply Laminated Elastic Plates Using a Higher-Order Theory: Part I—Stress and Displacement. *Compos Structures* (1988) 9(3):189–213. doi:10.1016/0263-8223(88)90014-1
123. Xing Y, Xiang W. Analytical Solution Methods for Eigenbuckling of Symmetric Cross-Ply Composite Laminates. *Chin J Aeronautics* (2017) 30(1):282–91. doi:10.1016/j.cja.2016.12.027
124. Khdeir AA. Free Vibration of Antisymmetric Angle-Ply Laminated Plates Including Various Boundary Conditions. *J Sound Vibration* (1988) 122(2):377–88. doi:10.1016/s0022-460x(88)80361-4
125. Khdeir AA. Stability of Antisymmetric Angle-Ply Laminated Plates. *J Eng Mech* (1989) 115(5):952–62. doi:10.1061/(asce)0733-9399(1989)115:5(952)
126. Khdeir AA, Librescu L. Analysis of Symmetric Cross-Ply Laminated Elastic Plates Using a Higher-Order Theory: Part II—Buckling and Free Vibration. *Compos Structures* (1988) 9(4):259–77. doi:10.1016/0263-8223(88)90048-7

127. Khdeir AA. Free Vibration and Buckling of Symmetric Cross-Ply Laminated Plates by an Exact Method. *J Sound Vibration* (1988) 126(3):447–61. doi:10.1016/0022-460x(88)90223-4
128. Khdeir AA. Free Vibration and Buckling of Unsymmetric Cross-Ply Laminated Plates Using a Refined Theory. *J Sound Vibration* (1989) 128(3):377–95. doi:10.1016/0022-460x(89)90781-5
129. Trinh LC, Vo TP, Thai HT, Nguyen TK, Keerthan P. State-Space Levy Solution for Size-Dependent Static, Free Vibration and Buckling Behaviours of Functionally Graded Sandwich Plates. *Composites Part B: Eng* (2018) 149: 144–64. doi:10.1016/j.compositesb.2018.05.017
130. Zhang LW, Song ZG, Liew KM. State-Space Levy Method for Vibration Analysis of FG-CNT Composite Plates Subjected to In-Plane Loads Based on Higher-Order Shear Deformation Theory. *Compos Structures* (2015) 134: 989–1003. doi:10.1016/j.compstruct.2015.08.138
131. Lv QH, Lei ZX, Zhang Y. Vibration Characteristic of Moderately Thick Graphene Reinforced Composite Functionally Graded Plate Based on State-Space Levy Method Combined With Higher Order Shear Deformation Theory. *IOP Conf Ser Mater Sci Eng* (2019) 531(1):012002. doi:10.1088/1757-899x/531/1/012002
132. Sobhy M. Hygrothermal Deformation of Orthotropic Nanoplates Based on the State-Space Concept. *Composites Part B: Eng* (2015) 79:224–35. doi:10.1016/j.compositesb.2015.04.042
133. Demirhan PA, Taskin V. Bending and Free Vibration Analysis of Levy-Type Porous Functionally Graded Plate Using State Space Approach. *Composites Part B: Eng* (2019) 160:661–76. doi:10.1016/j.compositesb.2018.12.020
134. Rouzegar J, Salmanpour N, Abad F, Li L. An Analytical State-Space Solution for Free Vibration of Sandwich Piezoelectric Plate With Functionally Graded Core. *Scientia Iranica* (2022) 29(2):502–33.
135. Khdeir AA, Reddy JN, Frederick D. A Study of Bending, Vibration and Buckling of Cross-Ply Circular Cylindrical Shells With Various Shell Theories. *Int J Eng Sci* (1989) 27(11):1337–51. doi:10.1016/0020-7225(89)90058-x
136. Khdeir AA, Reddy JN. Influence of Edge Conditions on the Modal Characteristics of Cross-Ply Laminated Shells. *Comput Structures* (1990) 34(6):817–26. doi:10.1016/0045-7949(90)90352-3
137. Dozio L. A Hierarchical Formulation of the State-Space Levy's Method for Vibration Analysis of Thin and Thick Multilayered Shells. *Composites Part B: Eng* (2016) 98:97–107. doi:10.1016/j.compositesb.2016.05.022
138. Hosseini-Hashemi S, Abaei AR, Ilkhani MR. Free Vibrations of Functionally Graded Viscoelastic Cylindrical Panel Under Various Boundary Conditions. *Compos Structures* (2015) 126:1–15. doi:10.1016/j.compstruct.2015.02.031
139. Razgordanisharahi A, Ghassabi AA, Hellmich C. Free Vibration Analysis of Cylindrical Honeycomb Sandwich Panels Using State-Space Levy Method. *Thin-Walled Structures* (2023) 182:110308. doi:10.1016/j.tws.2022.110308
140. Civalek Ö, Kiracioglu O. Free Vibration Analysis of Timoshenko Beams by DSC Method. *Int J Numer Methods Biomed Eng* (2010) 26(12):1890–8. doi:10.1002/cnm.1279
141. Alberts TE, Colvin JA. Observations on the Nature of Transfer Functions for Control of Piezoelectric Laminates. *J Intell Mater Syst Structures* (1991) 2(4): 528–41. doi:10.1177/1045389x9100200407
142. Ansari R, Sahmani S, Rouhi H. Rayleigh–Ritz Axial Buckling Analysis of Single-Walled Carbon Nanotubes With Different Boundary Conditions. *Phys Lett A* (2011) 375(9):1255–63. doi:10.1016/j.physleta.2011.01.046
143. Wang M, Xu YG, Qiao P, Li ZM. A Two-Dimensional Elasticity Model for Bending and Free Vibration Analysis of Laminated Graphene-Reinforced Composite Beams. *Compos Structures* (2019) 211:364–75. doi:10.1016/j.compstruct.2018.12.033
144. Chandrashekara S, Santhoshi U. Natural Frequencies of Cross-Ply Laminates by State Space Approach. *J Sound Vibration* (1990) 136(3):413–24. doi:10.1016/0022-460x(90)90453-7
145. Reddy JN. *Energy and Variational Methods in Applied Mechanics*. United States: John Wiley (1984).
146. Reddy JN. A Simple Higher-Order Theory for Laminated Composite Plates. *J Appl Mech* (1984) 51(4):745–52. doi:10.1115/1.3167719
147. Khdeir AA. *Analytical Solutions for the Statics and Dynamics of Rectangular Laminated Composite Plates Using Shearing Deformation Theories*. United States: Virginia Polytechnic Institute and State University (1986).
148. Xiang Y, Liew KM, Kitipornchai S. Exact Buckling Solutions for Composite Laminates: Proper Free Edge Conditions Under In-Plane Loadings. *Acta Mechanica* (1996) 117(1-4):115–28. doi:10.1007/bf01181041
149. Rezaei AS, Saidi AR, Abrishamdari M, Pour Mohammadi MH. Natural Frequencies of Functionally Graded Plates With Porosities via a Simple Four Variable Plate Theory: An Analytical Approach. *Thin-Walled Structures* (2017) 120:366–77. doi:10.1016/j.tws.2017.08.003
150. Jin G, Su Z, Shi S, Ye T, Gao S. Three-Dimensional Exact Solution for the Free Vibration of Arbitrarily Thick Functionally Graded Rectangular Plates With General Boundary Conditions. *Compos Structures* (2014) 108:565–77. doi:10.1016/j.compstruct.2013.09.051
151. Zaoui FZ, Ouinas D, Tounsi A. New 2D and Quasi-3D Shear Deformation Theories for Free Vibration of Functionally Graded Plates on Elastic Foundations. *Composites Part B: Eng* (2019) 159:231–47. doi:10.1016/j.compositesb.2018.09.051
152. Fazzolari FA. A Refined Dynamic Stiffness Element for Free Vibration Analysis of Cross-Ply Laminated Composite Cylindrical and Spherical Shallow Shells. *Composites Part B: Eng* (2014) 62:143–58. doi:10.1016/j.compositesb.2014.02.021
153. Hosseini-Hashemi S, Fadaee M. On the Free Vibration of Moderately Thick Spherical Shell Panel—A New Exact Closed-Form Procedure. *J Sound Vibration* (2011) 330(17):4352–67. doi:10.1016/j.jsv.2011.04.011
154. Noor AK, Anderson MS, Greene WH. Continuum Models for Beam-And Platelike Lattice Structures. *AIAA J* (1978) 16(12):1219–28. doi:10.2514/3.61036
155. Ayre RS, Jacobsen LS. Natural Frequencies of Continuous Beams of Uniform Span Length. *J Appl Mech* (1950) 17:391–5. doi:10.1115/1.4010165
156. Miles JW. Vibrations of Beams on Many Supports. *J Eng Mech Division* (1956) 82(1):1–9. doi:10.1061/jmcea3.0000001
157. Lin YK, McDaniel TJ. Dynamics of Beam-Type Periodic Structures. *J Eng Industry* (1969) 91:1133–41. doi:10.1115/1.3591761
158. Vaicaitis R, Doi K, Lin YK. Spatial Decay in the Response of Damped Periodic Beam. *J Aircraft* (1972) 9(1):91–3. doi:10.2514/3.44324
159. Lin YK, Yang JN. Free Vibration of a Disordered Periodic Beam. *J Appl Mech* (1974) 41(2):383–91. doi:10.1115/1.3423298
160. Yang JN, Lin YK. Frequency Response Functions of a Disordered Periodic Beam. *J Sound Vibration* (1975) 38(3):317–40. doi:10.1016/s0022-460x(75)80050-2
161. Vaicaitis R, Lin YK. Response of Finite Periodic Beam to Turbulent Boundary-Layer Pressure Excitation. *AIAA J* (1972) 10(8):1020–4. doi:10.2514/3.50288
162. Yeh JY, Chen LW. Wave Propagations of a Periodic Sandwich Beam by FEM and the Transfer Matrix Method. *Compos Structures* (2006) 73(1):53–60. doi:10.1016/j.compstruct.2005.01.026
163. Assis GF, Beli D, de Miranda EJ, Jr, Camino JF, Dos Santos JMC, Arruda JRF. Computing the Complex Wave and Dynamic Behavior of One-Dimensional Phononic Systems Using a State-Space Formulation. *Int J Mech Sci* (2019) 163:105088. doi:10.1016/j.ijmecsci.2019.105088
164. Lin YK, Zhang R, Yong Y. Multiply Supported Pipeline Under Seismic Wave Excitations. *J Eng Mech* (1990) 116(5):1094–108. doi:10.1061/(asce)0733-9399(1990)116:5(1094)
165. Cai GQ, Lin YK. Localization of Wave Propagation in Disordered Periodic Structures. *AIAA J* (1991) 29(3):450–6. doi:10.2514/3.10599
166. Qiu CX, Lin YK. Localization of Mono-And Multichannel Waves in Disordered Periodic Systems. *J Eng Mech* (1997) 123(8):830–6. doi:10.1061/(asce)0733-9399(1997)123:8(830)
167. Cai GQ, Lin YK. Statistical Distribution of Frequency Response in Disordered Periodic Structures. *AIAA J* (1992) 30(5):1400–7. doi:10.2514/3.11076
168. Saeed HM, Vestroni F. Simulation of Combined Systems by Periodic Structures: The Wave Transfer Matrix Approach. *J Sound Vibration* (1998) 213(1):55–74. doi:10.1006/jsvi.1997.1497
169. Romeo F, Paolone A. Wave Propagation in Three-Coupled Periodic Structures. *J Sound Vibration* (2007) 301(3-5):635–48. doi:10.1016/j.jsv.2006.10.017
170. Yong Y, Lin YK. Propagation of Decaying Waves in Periodic and Piecewise Periodic Structures of Finite Length. *J Sound Vibration* (1989) 129(1):99–118. doi:10.1016/0022-460x(89)90538-5

171. Yong Y, Lin YK. Dynamics of Complex Truss-Type Space Structures. *AIAA J* (1990) 28(7):1250–8. doi:10.2514/3.25202
172. Yong Y, Lin YK. Dynamic Response Analysis of Truss-Type Structural Networks: A Wave Propagation Approach. *J Sound Vibration* (1992) 156(1):27–45. doi:10.1016/0022-460x(92)90810-k
173. Luongo A, Romeo F. Real Wave Vectors for Dynamic Analysis of Periodic Structures. *J Sound Vibration* (2005) 279(1-2):309–25. doi:10.1016/j.jsv.2003.11.011
174. Ren P, Zhou Z. Strain Response Estimation for the Fatigue Monitoring of an Offshore Truss Structure. *Pac Sci Rev* (2014) 16(1):29–35. doi:10.1016/j.pscr.2014.08.005
175. Cai GQ, Lin YK. Wave Propagation and Scattering in Structural Networks. *J Eng Mech* (1991) 117(7):1555–74. doi:10.1061/(asce)0733-9399(1991)117:7(1555)
176. Lin YK, Donaldson BK. A Brief Survey of Transfer Matrix Techniques With Special Reference to the Analysis of Aircraft Panels. *J Sound Vibration* (1969) 10(1):103–43. doi:10.1016/0022-460x(69)90132-1
177. Heckl MA. Investigations on the Vibrations of Grillages and Other Simple Beam Structures. *The J Acoust Soc America* (1964) 36(7):1335–43. doi:10.1121/1.1919206
178. Mead DJ. Free Wave Propagation in Periodically Supported, Infinite Beams. *J Sound Vibration* (1970) 11(2):181–97. doi:10.1016/s0022-460x(70)80062-1
179. Gupta GS. Natural Flexural Waves and the Normal Modes of Periodically-Supported Beams and Plates. *J Sound Vibration* (1970) 13(1):89–101. doi:10.1016/s0022-460x(70)80082-7
180. Mead D. A General Theory of Harmonic Wave Propagation in Linear Periodic Systems With Multiple Coupling. *J Sound Vibration* (1973) 27(2):235–60. doi:10.1016/0022-460x(73)90064-3
181. Mead DJ. Wave Propagation and Natural Modes in Periodic Systems: I. Mono-Coupled Systems. *J Sound Vibration* (1975) 40(1):1–18. doi:10.1016/s0022-460x(75)80227-6
182. Mead D. Wave Propagation and Natural Modes in Periodic Systems: II. Multi-Coupled Systems, With and Without Damping. *J Sound Vibration* (1975) 40(1):19–39. doi:10.1016/s0022-460x(75)80228-8
183. Mead DJ, Markuš Š. Coupled Flexural-Longitudinal Wave Motion in a Periodic Beam. *J Sound Vibration* (1983) 90(1):1–24. doi:10.1016/0022-460x(83)90399-1
184. Zhu L, Elishakoff I, Lin YK. Free and Forced Vibrations of Periodic Multispan Beams. *Shock and Vibration* (1994) 1(3):217–32. doi:10.1155/1994/961272
185. Lin YK. Dynamics of Disordered Periodic Structures. *Appl Mech Rev* (1996) 49(2):57–64. doi:10.1115/1.3101888
186. Von Flotow AH. Disturbance Propagation in Structural Networks. *J Sound Vibration* (1986) 106(3):433–50. doi:10.1016/0022-460x(86)90190-2
187. Chen WQ, Lü CF, Bian EZ. Elasticity Solution for Free Vibration of Laminated Beams. *Compos Structures* (2003) 62(1):75–82. doi:10.1016/s0263-8223(03)00086-2
188. Chen WQ, Lü CF, Bian ZG. A Mixed Method for Bending and Free Vibration of Beams Resting on a Pasternak Elastic Foundation. *Appl Math Model* (2004) 28(10):877–90. doi:10.1016/j.apm.2004.04.001
189. Chen WQ, Lü CF, Bian ZG. Free Vibration Analysis of Generally Laminated Beams via State-Space-Based Differential Quadrature. *Compos Structures* (2004) 63(3-4):417–25.
190. Xu R, Wu YF. Free Vibration and Buckling of Composite Beams With Interlayer Slip by Two-Dimensional Theory. *J Sound Vibration* (2008) 313(3-5):875–90. doi:10.1016/j.jsv.2007.12.029
191. Lü CF, Chen WQ. Free Vibration of Orthotropic Functionally Graded Beams With Various End Conditions. *Struct Eng Mech* (2005) 20(4):465–76. doi:10.12989/sem.2005.20.4.465
192. Lü CF, Chen W, Zhong Z. Two-Dimensional Thermoelasticity Solution for Functionally Graded Thick Beams. *Sci China Ser G: Phys Mech Astron* (2006) 49:451–60. doi:10.1007/s11433-006-0451-2
193. Lü CF, Chen WQ, Xu RQ, Lim CW. Semi-Analytical Elasticity Solutions for Bi-Directional Functionally Graded Beams. *Int J Sol Structures* (2008) 45(1):258–75. doi:10.1016/j.ijsolstr.2007.07.018
194. Li Y, Shi Z. Free Vibration of a Functionally Graded Piezoelectric Beam via State-Space Based Differential Quadrature. *Compos Structures* (2009) 87(3):257–64. doi:10.1016/j.compstruct.2008.01.012
195. Alibeigloo A, Liew KM. Elasticity Solution of Free Vibration and Bending Behavior of Functionally Graded Carbon Nanotube-Reinforced Composite Beam With Thin Piezoelectric Layers Using Differential Quadrature Method. *Int J Appl Mech* (2015) 7(01):1550002. doi:10.1142/s1758825115400025
196. Chen WQ, Lee KY. On Free Vibration of Cross-Ply Laminates in Cylindrical Bending. *J Sound Vibration* (2004) 273(3):667–76. doi:10.1016/j.jsv.2003.08.003
197. Chen WQ, Lü CF. 3D Free Vibration Analysis of Cross-Ply Laminated Plates With One Pair of Opposite Edges Simply Supported. *Compos Structures* (2005) 69(1):77–87. doi:10.1016/j.compstruct.2004.05.015
198. Lü CF, Zhang ZC, Chen WQ. Free Vibration of Generally Supported Rectangular Kirchhoff Plates: State-Space-Based Differential Quadrature Method. *Int J Numer Methods Eng* (2007) 70(12):1430–50. doi:10.1002/nme.1929
199. Zhou YY, Chen WQ, Lü CF, Wang J. Free Vibration of Cross-Ply Piezoelectric Laminates in Cylindrical Bending With Arbitrary Edges. *Compos Structures* (2009) 87(1):93–100. doi:10.1016/j.compstruct.2008.01.002
200. Zhou YY, Chen WQ, Lü CF. Semi-Analytical Solution for Orthotropic Piezoelectric Laminates in Cylindrical Bending With Interfacial Imperfections. *Compos Structures* (2010) 92(4):1009–18. doi:10.1016/j.compstruct.2009.09.048
201. Yas MH, Moloudi N. Three-Dimensional Free Vibration Analysis of Multi-Directional Functionally Graded Piezoelectric Annular Plates on Elastic Foundations via State Space Based Differential Quadrature Method. *Appl Math Mech* (2015) 36:439–64. doi:10.1007/s10483-015-1923-9
202. Alibeigloo A. Static and Vibration Analysis of Axi-Symmetric Angle-Ply Laminated Cylindrical Shell Using State Space Differential Quadrature Method. *Int J Press Vessels Piping* (2009) 86(11):738–47. doi:10.1016/j.ijpvp.2009.07.002
203. Ali MY, Shankar K. Structural Health Monitoring of Composites Using System Identification. In: 15th Australian International Aerospace Congress (AIAC15); 25-28 February 2013; Melbourne, Australia (2013).
204. Ali MY, Shankar K. Agent Based Damaged Detection in Composite Laminates. *Nondestructive Characterization Compos Mater Aerospace Eng Civil Infrastructure, Homeland Security* (2013) 8694:528–33.
205. Li D, Qing GH. Free Vibration Analysis of Composite Laminates With Delamination Based on State Space Theory. *Mech Adv Mater Structures* (2014) 21(5):402–11. doi:10.1080/15376494.2012.697602
206. Zhou Y, Li S, Zhou H. State Space Finite Element Analysis for Piezoelectric Precision Positioning Considering Secondary Converse Piezoelectric Effect. *Finite Elem Anal Des* (2015) 102:85–94. doi:10.1016/j.finel.2015.04.010
207. Zhou Y, Nyberg T, Xiong G, Li S. State Space Finite Element Analysis for Piezoelectric Laminated Curved Beam With Variable Curvature. *Mech Adv Mater Structures* (2020) 27(4):265–73. doi:10.1080/15376494.2018.1472323
208. dos Santos CR, Pacheco DR, Taha HE, Zakaria MY. Nonlinear Modeling of Electro-Aeroelastic Dynamics of Composite Beams With Piezoelectric Coupling. *Compos Structures* (2021) 255:112968. doi:10.1016/j.compstruct.2020.112968
209. Sheng HY, Ye JQ. A State Space Finite Element for Laminated Composite Plates. *Comp Methods Appl Mech Eng* (2002) 191(37-38):4259–76. doi:10.1016/s0045-7825(02)00379-1
210. Zou GP, Tang LM. A Semi-Analytical Solution for Laminated Composite Plates in Hamiltonian System. *Comp Methods Appl Mech Eng* (1995) 128(3-4):395–404. doi:10.1016/0045-7825(95)00877-2
211. Zou GP, Tang LM. A Semi-Analytical Solution for Thermal Stress Analysis of Laminated Composite Plates in the Hamiltonian System. *Comput Structures* (1995) 55(1):113–8.
212. Sheng HY, Ye JQ. A Semi-Analytical Finite Element for Laminated Composite Plates. *Compos Structures* (2002) 57(1-4):117–23. doi:10.1016/s0263-8223(02)00075-2
213. Ye JQ, Sheng HY, Qin QH. A State Space Finite Element for Laminated Composites With Free Edges and Subjected to Transverse and In-Plane Loads. *Comput Structures* (2004) 82(15-16):1131–41. doi:10.1016/j.compstruc.2004.03.020
214. Han ZL, Cheng CZ, Sheng HY. A State-Space Finite Element Method for Laminated Composite Plates Under Thermal Loading. *J Therm Stresses* (2017) 40(10):1285–302. doi:10.1080/01495739.2017.1319256
215. Qing GH, Qiu JJ, Liu YH. Modified HR Mixed Variational Principle for Magnetoelectroelastic Bodies and State-Vector Equation. *Appl Math Mech* (2005) 26(6):722–8.

216. Qing GH, Qiu JJ, Liu YH. A Semi-Analytical Solution for Static and Dynamic Analysis of Plates With Piezoelectric Patches. *Int J Sol Structures* (2006) 43(6): 1388–403. doi:10.1016/j.ijsolstr.2005.03.048
217. Qing GH, Qiu JJ, Liu YH. Free Vibration Analysis of Stiffened Laminated Plates. *Int J Sol Structures* (2006) 43(6):1357–71. doi:10.1016/j.ijsolstr.2005.03.012
218. Liu YH, Wang BH, Qing GH. Free Vibration Analysis of Piezoelectric Laminated Plates With Delamination. *Eng Mech* (2012) 29(7):347–52.
219. Zhou J, Yang B. Three-Dimensional Analysis of Simply Supported Laminated Cylindrical Shells With Arbitrary Thickness. *AIAA J* (1996) 34(9):1960–4. doi:10.2514/3.13338
220. Sheng HY, Ye JQ. A Three-Dimensional State Space Finite Element Solution for Laminated Composite Cylindrical Shells. *Comp Methods Appl Mech Eng* (2003) 192(22-24):2441–59. doi:10.1016/s0045-7825(03)00265-2
221. Qing GH, Liu YH, Qiu JJ, Meng XJ. A Semi-Analytical Method for the Free Vibration Analysis of Thick Double-Shell Systems. *Finite Elem Anal Des* (2006) 42(10):837–45. doi:10.1016/j.finel.2006.01.002
222. Jiang LZ, Lee JS. Plane Orthotropic Layer by Transfer Matrix-Spline Boundary Element. *J Eng Mech* (1994) 120(5):1026–41. doi:10.1061/(asce)0733-9399(1994)120:5(1026)
223. Ruddock GJ, Spencer AJM. A New Approach to Stress Analysis of Anisotropic Laminated Elastic Cylinders. *Proc R Soc Lond Ser A: Math Phys Eng Sci* (1997) 453:1067–82. doi:10.1098/rspa.1997.0060
224. Attallah KMZ, Ye JQ, Sheng HY. Three-Aimensional Finite Strip Analysis of Laminated Panels. *Comput Structures* (2007) 85(23-24):1769–81. doi:10.1016/j.compstruc.2007.04.003
225. Khezri M, Gharib M, Bradford MA, Vrcelj Z. Analysis of Thick and Orthotropic Rectangular Laminated Composite Plates Using a State-Space-Based Generalised RKP-FSM. *Compos Structures* (2015) 133: 691–706. doi:10.1016/j.compstruct.2015.07.035
226. Xin L, Hu Z. Free Vibration of Layered Magneto-Electro-Elastic Beams by SS-DSC Approach. *Compos Structures* (2015) 125:96–103. doi:10.1016/j.compstruct.2015.01.048
227. Chandrashekhara K, Krishnamurthy K, Roy S. Free Vibration of Composite Beams Including Rotary Inertia and Shear Deformation. *Compos Structures* (1990) 14(4):269–79. doi:10.1016/0263-8223(90)90010-c
228. Rao MK, Desai YM, Chitnis MR. Free Vibrations of Laminated Beams Using Mixed Theory. *Compos Structures* (2001) 52(2):149–60.
229. Fan JR. *Exact Theory of Laminated Thick Plates and Shells*. Beijing: Beijing Science (1996).
230. Auricchio F, Sacco E. A Mixed-Enhanced Finite-Element for the Analysis of Laminated Composite Plates. *Int J Numer Methods Eng* (1999) 44(10): 1481–504. doi:10.1002/(sici)1097-0207(19990410)44:10<1481::aid-nme554>3.0.co;2-q
231. Cheung YK. The Finite Strip Method in the Analysis of Elastic Plates With Two Opposite Simply Supported Ends. *Proc Am Soc Civ Eng* (1968) 94:1365–78.
232. Fan SC, Cheung YK. Static Analysis of Right Box Girder Bridges by Spline Finite Strip Method. *Proc Inst Civil Eng* (1983) 75(2):311–23. doi:10.1680/iicep.1983.1507
233. Khezri M, Vrcelj Z, Bradford MA. Thin Plate Bending Analysis and Treatment of Material Discontinuities Using the Generalised RKP-FSM. *Comp Model Eng Sci (Cmes)* (2012) 87(4):271–305.
234. Pao YH, Keh DC, Howard SM. Dynamic Response and Wave Propagation in Plane Trusses and Frames. *AIAA J* (1999) 37(5):594–603. doi:10.2514/2.778
235. Pao YH, Su XY, Tian JY. Reverberation Matrix Method for Propagation of Sound in a Multilayered Liquid. *J Sound Vibration* (2000) 230(4):743–60. doi:10.1006/jsvi.1999.2675
236. Pao YH, Sun G. Dynamic Bending Strains in Planar Trusses With Pinned or Rigid Joints. *J Eng Mech* (2003) 129(3):324–32. doi:10.1061/(asce)0733-9399(2003)129:3(324)
237. Pao YH, Chen WQ, Su XY. The Reverberation-Ray Matrix and Transfer Matrix Analyses of Unidirectional Wave Motion. *Wave Motion* (2007) 44(6): 419–38. doi:10.1016/j.wavemoti.2007.02.004
238. Wang L, Rokhlin SI. Stable Reformulation of Transfer Matrix Method for Wave Propagation in Layered Anisotropic Media. *Ultrasonics* (2001) 39(6): 413–24. doi:10.1016/s0041-624x(01)00082-8
239. Zhao Z, Wang B, Qian Z, Yong YK. A Novel Approach to Quantitative Predictions of High-Frequency Coupled Vibrations in Layered Piezoelectric Plates. *Int J Eng Sci* (2020) 157:103407. doi:10.1016/j.ijengsci.2020.103407
240. Zhao Z, Chen W. Investigation on High-Frequency and Mode-Coupling Vibrations in Thickness-Extensional Piezoelectric Thin-Film Resonators With Initial Stress. *Appl Math Model* (2022) 112:78–90. doi:10.1016/j.apm.2022.07.030
241. Chen Y, Wu B, Su Y, Chen W. Effects of Strain Stiffening and Electrostriction on Tunable Elastic Waves in Compressible Dielectric Elastomer Laminates. *Int J Mech Sci* (2020) 176:105572. doi:10.1016/j.ijmecsci.2020.105572
242. Shmuel G, Debotton G. Band-Gaps in Electrostatically Controlled Dielectric Laminates Subjected to Incremental Shear Motions. *J Mech Phys Sol* (2012) 60(11):1970–81. doi:10.1016/j.jmps.2012.05.006
243. Zhao Z, Zhu J, Chen W. Size-Dependent Vibrations and Waves in Piezoelectric Nanostructures: A Literature Review. *Int J Smart Nano Mater* (2022) 13(3):391–431. doi:10.1080/19475411.2022.2091058
244. Stan G, Ciobanu CV, Parthangal PM, Cook RF. Diameter-Dependent Radial and Tangential Elastic Moduli of ZnO Nanowires. *Nano Lett* (2007) 7(12): 3691–7. doi:10.1021/nl071986e
245. Wang KF, Wang BL, Kitamura T. A Review on the Application of Modified Continuum Models in Modeling and Simulation of Nanostructures. *Acta Mechanica Sinica* (2016) 32:83–100. doi:10.1007/s10409-015-0508-4

AEROSPACE RESEARCH COMMUNICATIONS
ORIGINAL RESEARCH
published: 04 January 2024
doi: 10.3389/arc.2023.12420

Increasing Operational Resiliency of UAV Swarms: An Agent-Focused Search and Rescue Framework

Abhishek Phadke[1,2]* *and F. Antonio Medrano*[1,2]

[1]*Conrad Blucher Institute for Surveying and Science, Texas A&M University-Corpus Christi, Corpus Christi, TX, United States,*
[2]*Department of Computer Science, Texas A&M University-Corpus Christi, Corpus Christi, TX, United States*

Resilient UAV (Unmanned Aerial Vehicle) swarm operations are a complex research topic where the dynamic environments in which they work significantly increase the chance of systemic failure due to disruptions. Most existing SAR (Search and Rescue) frameworks for UAV swarms are application-specific, focusing on rescuing external non-swarm agents, but if an agent in the swarm is lost, there is inadequate research to account for the resiliency of the UAV swarm itself. This study describes the design and deployment of a Swarm Specific SAR (SS-SAR) framework focused on UAV swarm agents. This framework functions as a resilient mechanism by locating and attempting to reconnect communications with lost UAV swarm agents. The developed framework was assessed over a series of performance tests and environments, both real-world hardware and simulation experiments. Experimental results showed successful recovery rates in the range of 40%–60% of all total flights conducted, indicating that UAV swarms can be made more resilient by including methods to recover distressed agents. Decision-based modular frameworks such as the one proposed here lay the groundwork for future development in attempts to consider the swarm agents in the search and rescue process.

Keywords: resilience, UAV, swarm, search and rescue, UAS

INTRODUCTION

The use of UAV swarms is becoming more widespread due to the reduced costs of UAVs and their ability to accomplish tasks more quickly and effectively as a group rather than individually. Advancements in aircraft design and control, communication topologies, and battery systems have made coordinated UAV swarms possible. The use of UAV swarms in applications of military [1, 2], ecology [3], remote sensing [4, 5], disaster management [6], crowd control, emergency communication [7], agriculture [8, 9], and victim search [10, 11] are just some of the use cases. As individual and multi-robotic development and their interaction with real-world entities advances, these applications are only limited because of a lack of improvement, a discrepancy that exponentially decreases as time passes. With the increased diversity of swarm usage

***Correspondence**
Abhishek Phadke,
aphadke@islander.tamucc.edu

Received: *15 November 2023*
Accepted: *13 December 2023*
Published: *04 January 2024*

Citation:
Phadke A and Medrano FA (2024) Increasing Operational Resiliency of UAV Swarms: An Agent-Focused Search and Rescue Framework. Aerosp. Res. Commun. 1:12420. doi: 10.3389/arc.2023.12420

Abbreviations: AS, Application specific; FANET, Flying Ad Hoc Networks; GNSS, Global Navigation Satellite System; GCP, Ground control points; HBS, Heartbeat signal; IDS, Intrusion Detection System; MANET, Mobile Ad Hoc Networks; MPC, Model Predictive Control; MRS, Multi-Robot System; OLSR, Optimized Link State Routing; PSO, Particle Swarm Optimization; POV, Point of View; ROI, Region Of Interest; S.S., Swarm Specific; SSI, Signal Strength Indicator; SAR, Swarm Agent Specific; SAR, Search and Rescue; SMF, Swarm Methodology Focused; UAV, Unmanned Aerial Vehicle; UGV, Unmanned Ground Vehicle; UWSV, Unmanned Water Surface Vehicle; VPS, Visual positioning system; VRRZ, Variable Radius Risk Zones.

applications, research in UAV resiliency has also grown [12, 13]. Due to the close-knit topology of these swarms, the failure of agents above a certain threshold can often lead to cascaded systemic collapse and a pause on mission progress. The cause of this failure may be structural, such as in a leader-follower topology, wherein the followers may get disconnected if the leader fails. Additional uncertainty in this failure also exists, such as the possibility of a failing swarm agent crashing into other agents during its collapse. Resiliency is defined as the ability of a system to withstand disruptions. Broader definitions include the ability of a system to bounce back after a disruption [14]. D.D. Woods summarizes system resilience perfectly in its four core concepts [15]. They are resilience, such as rebound, robustness, extensibility, and adaptability. Previous work by the authors addresses systemic resilience in UAV swarms on a broader range by classifying UAV swarm operations into components and modules [16]. The resilience of UAV swarms is a complex topic that integrates multiple components of navigation [17], mapping [4], control [18], defensive and intrusion detection policies [19], agent welfare, and physical characteristics of the swarm agent [20] into an intricate system designed to create balance in a dynamic environment.

An MRS (Multi-Robot System) and a swarm are both concepts in robotics that involve the coordination and interaction of multiple robots. However, they have distinct differences in organization, control, and behavior. An MRS has a structured and explicit interaction scheme with a centralized control. Swarm agents, on the other hand, are more decentralized and self-organized. There might not be explicit communication between robots, and collective behavior emerges from simple interactions between individual robots following local rules.

Additionally, MRS may involve centralized controllers and planners to assign tasks to robots. Swarms rely on local interactions and distributed control. Each robot can typically make decisions based on its immediate surroundings or information gathered from nearby agents. This also involves a degree of autonomy in decision-making from completely autonomous to semi-autonomous. However, while swarms are expected to be inherently scalable, there is an ongoing debate on the minimum number of agents that must be present and acting collectively to label it as a swarm. Adding more robots to a swarm does not necessarily increase performance. Selecting the number of agents in a swarm has long been contested. However, approaches with agents as few as five and as many as 1,000 have been implemented and studied. Article [21] discusses how aspects such as system scalability, technical capabilities of individual agents, and financial or logical constraints influence the selection of the number of agents in the swarm.

These factors were crucial in selecting the number of agents for swarm response experiments performed in this study. The number of agents available for experimentation was limited. Additionally, some agents were designated as reserve and spare agents to ensure experiments continued in case of equipment failure. Space constraints allowed only a certain number of agents to fly in the designated airspace without the risk of agent collision and crashes due to induced airflow interactions. While all communication between agents was performed decentralized, primary communication and network protocols required for communication with GCS limited the number of agents connected to them.

Search and Rescue is a vast domain; focusing only on swarm agent welfare significantly narrows it. However, to concentrate results further, this SAR framework will be primarily described for exploratory swarm applications. Scenarios where a swarm of agents may be deployed over an area and, in the process, may lose contact with the swarm is the priority. This narrows down the framework focus as well as experiment design and validation. Two major types of SAR capabilities in UAV swarms are defined and categorized here. They are application-specific SAR [22] and swarm-specific SAR (SS-SAR). Although our study takes a different direction than a regular application-focused SAR use scenario, it remains an exploratory problem. The tracking, location, and Rescue of disabled swarm agents require other agents of the swarm to actively search the target space for the agent using techniques such as triangulated localization, computer vision, sensors on the ground, and the analysis of system-generated mission logs. A literature review reveals that most swarms lack the self-awareness needed to actively take care of their agents. More robust mechanisms for the welfare of UAV swarms [23] are needed as an additional means to increase systemic resilience.

To build robust applications and routine case scenarios that use UAV swarms, the swarm itself must be resilient to disruptions. Towards this goal concerning SAR swarm operations, the significant contributions of this paper are:

- A literature analysis that reveals a research gap in UAV swarm development related to the search and rescue of their agents.
- To address this gap, a novel UAV swarm framework, SS-SAR (Swarm Specific-SAR), is introduced to provide the ability to track, locate, and possibly rescue their agents. The framework uses a decentralized approach and local communication between neighboring agents and surrounding data to make semi-autonomous deployments of rescue craft that initiate direct communication with distressed agents.
- Experimental results show the SS-SAR framework's ability to reduce agent loss in swarm operations.
- Future framework upgrades and experiment designs are proposed to increase operational swarm resilience.

Using a decentralized approach for communication and agent decisions, this study aims to demonstrate scalable and robust responses of swarms to disruptive scenarios along with further scope for possible emergent behavior to avoid them altogether based on broad programmed constraints. The paper is arranged in the following way. Section *Introduction* gives a brief introduction of the area with research contributions of this study. Section *Summary of Recent Literature on Multi-Robot SAR* presents a categorization of current trends. Section *Swarm-Specific SAR Framework* presents the SS-SAR scenario

TABLE 1 | Summary of recent work on SAR using multi-agent UAV swarms categorized by approach.

Reference	Category	Description
[24]	AS	Using a modified fruit fly algorithm to improve the search efficiency of a multi-robot swarm
[25]	AS	Cooperative strategy for distributed UAV agents in a swarm performing unique functions for victim search and rescue operations
[26]	AS	Smart search for survivors using a genetic localization method to detect victim distress signals using autonomous maximum area searching UAV agents
[27]	AS	Collaboration between swarm agents for detecting victim presence
[28]	AS	Layered SAR based on disaster epicenter for improved victim detection using multiple agents
[29]	AS	Heterogeneous agent swarm based on ant colony optimization and agent decision process for victim searching at sea
[30]	AS	An open-source platform for managing drones for assistance in SAR operations
[31]	SMF	Using deep reinforcement learning to generate control commands for UAVs to search in an environment with an unknown number of targets
[32]	SMF	A dynamically varying number of swarm agents search for the target using MPC for generating cooperative search trajectories and maximizing performance
[33]	SMF	Creating target probability maps to guide swarm search actions based on flocking, velocity, and area coverage
[34]	SMF	Collaborative search function based on pigeon-inspired bio-inspired algorithm
[35]	SMF	Hexagonal grid decomposition of the search area for maximum efficiency during target search in a maritime rescue scenario
[36]	SMF	Planning using a Markov decision process and control using environmental exploration by deep learning for target detection
[37]	SMF	A bio-inspired algorithm based on fish schooling and foraging behaviors for improving target search functionality
[38]	SMF	A reinforcement learning-based concept to make a territory awareness map for generating cooperative search paths for multi-UAV swarms
[39]	SMF	A profit-driven adaptive search algorithm for moving targets using a UAV swarm capable of information exchange
[40]	SMF	PSO-MPC approach to solving and improving the efficiency of the SAR technique using multiple agents rather than a single agent
[41]	SS-SAR	A swarm-specific methodology for automatic replacement of any lost UAV during mission progress

description and framework workflow. Section *Performance Tests* describes both hardware and simulation performance tests and environmental parameters. Section *Results* presents the experimental results of the new SS-SAR framework. Section D*iscussion and Future Research Directions* provides future directions to approach the problem with suggested framework upgrades, and Section *Conclusion* provides concluding statements.

SUMMARY OF RECENT LITERATURE ON MULTI-ROBOT SAR

The multi-UAV SAR problem is a broad problem domain. This section categorizes current research into three distinct approaches. **Table 1** identifies research on the topic and categorizes the study as application or swarm-specific problems. Application-specific SAR (AS-SAR) and SS-SAR are the two categorizations previously discussed. Depending on how the SAR problem is approached, a third category is also included: Search Methodology Focused (SMF). Research in this category does not have a specific search target type; instead, the focus is on the general SAR methodology, where any internal or external search target can be assigned.

As one can see from **Table 1**, swarm-specific implementations are less explored in the literature. In addition to the above literature review, generalized methodologies exist that propose novel approaches that would improve facets of the SAR process. These include using bio-inspired algorithms for area coverage [42] formation tracking [43] and environment exploration [44], updated and merged observation maps or information exchange pathways [45], and efficient task planning [46, 47]. Frameworks such as [41] that propose automatic replacement of lost UAV agents are scarce. This example fits perfectly in this paper's proposed swarm-specific research category. To keep the literature analysis attainable, any approaches that do not directly describe the use of aerial vehicle swarms in the field for SAR have been eliminated. This includes broader research topics such as using machine learning methods to improve object recognition in aerial images taken by UAVs [48].

SWARM-SPECIFIC SAR FRAMEWORK

Workflow Description

Notations used in framework description and development are summarized in **Table 2**. This section briefly describes the broad workflow for the framework design process.

The SS-SAR workflow [49] is summarized in **Figure 1**. It is divided into four sections, with the first section defining the agent tracking phase, the second section containing the initial decision, section three having the primary decision process, and section four with the secondary decision process. The modular framework design assists in the testing and modification of one or more sections. This was especially useful in scenarios where the hardware and software test platforms could not simultaneously handle all the framework tests computationally or physically. For example, low-cost agents such as the DJI Tellos [50] used in the lab scenarios required testing individual sections piecewise rather than the entire framework simultaneously due to inefficient hardware and lack of sensors. The experiment section describes the modular experiments designed to test the workflow to the extent that the agents could handle it.

TABLE 2 | Notations.

n	Total number of agents in the swarm
i	Index of UAV agents (from 1 to n)
t	Equally spaced time interval between HBS signals
k	Index of HBS time sample
$HBS_{i,k}$	HBS signal from agent i at time k
Ind	Binary variable to denote the presence or absence of HBS signal
$Ind_{i,k}$	Binary value for HBS signal from agent i at time k
Ind_{all}	Binary variable based on an AND logical operation of all binary indicator variables
$HBS_{i,k} \rightarrow loc$	Location of agent i at time k (included in the corresponding HBS signal)
$HBS_{i,k} \rightarrow b_{level}$	Battery level of agent i at time k (included in the corresponding HBS signal)
$SSI_{(GC,\ i)}$	Signal strength indication of ground control to agent i
OG	On-ground indicator that is set to 1 if an agent is actively connected but is on the ground
p_0	Real-time pose check using the distressed agent camera
p_1	Real-time pose check using the rescue agent camera
p_2	Real-time pose check using IMU
tk_{id}	Task ID
tk_{cost}	Cost of completing a task
$R_{loc(i,k-1)}$	Denotes the rescue agent moving to the location of distressed agent i at k–1 time
rcm	Reconnection message

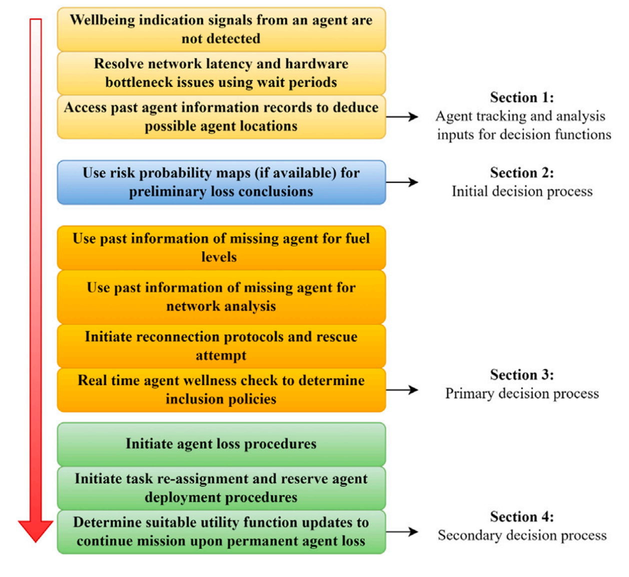

FIGURE 1 | Workflow of the SS-SAR framework.

The second advantage of the modular nature is that framework components can be upgraded, optimized, or changed. For example, while preliminary experiments for Section *Performance Tests* use an essential task re-assignment policy where only idle agents are given the tasks previously assigned to the lost agent, future iterations of the framework can use an optimized cost consideration, where characteristics of the task-receiving agent, such as its remaining fuel, are considered before re-assignment. An agent completing its task is only assigned the task of the fallen agent if its battery capacity allows it. The indicator tk_{cost} is used to determine the cost of completing the task that is estimated using the number of time intervals required t, and the expected change in battery level to complete the task, Δb_{level}.

Scenario Description

The generalized model in **Figure 2** was expanded into a specific scenario where a swarm of agents is performing a task, and one of the agents is in distress. This SS-SAR process is depicted in **Figure 2**.

The OLSR (Optimized Link State Routing) protocol has been extensively studied as an ideal routing protocol in SAR environments [51]. It routinely uses “Hello” and “Topology Control” messages to identify links and agent states. The heartbeat signal is often referred to as a modified hello message based on the base OLSR protocol. The heartbeat signal transmission is a small, quick transmission objective signal that each swarm agent can send at regular intervals. Various alternate implementations exist [26, 41, 52]; however, they follow a general structure that includes information denoting network I.D., transmitting agent ID, destination I.D., message type, security I.D., data segments, and error check. The HBS comprises location information, the battery level of the agent, signal strength indication, and the current task I.D.

$$HBS_{i,k} = \left[Loc_{i,k} + b_{level,i,k} + SSI_{(GC,i)} + tk_{id}\right]$$

A fixed number of agents, n, form the swarm. The HBS from every agent is expected after each time interval t, at a sample iteration denoted by k. An HBS is expected to be transmitted by every agent in the swarm after the time interval value of t. The signal is denoted by $HBS_{i,k.}$ This denotes the signal transmission from the i^{th} agent at the k^{th} time interval, in the range of $i = 1$ to n, and $k > 0$. The Ind binary variable indicates the presence or absence of the HBS signal for every $Ind_{i,k.}$ The value of 1 is recorded for every signal received and 0 if a signal is missing. The missing HBS agent id value determines which agent did not send the signal. The Ind_{all} is a logical operator set to 1 if all agents send a signal and 0 if HBS from an agent is missing.

Time intervals t are regular spaced and defined for transmitting the HBS. Careful consideration of this assigned value is required. A higher value of t can cause fewer HBS to be transmitted during mission time, i.e., a greater amount of time can elapse between a missing HBS and the system realization of an agent in distress. However, a lesser value of t can cause network bottlenecks if the system cannot receive and process HBS from all agents of the swarm. **Figure 3** shows HBS signal and sensor data transmission over a regular and disrupted time series. A longer period of HBS transmission intervals may result in delays between agent loss and system realization, $2t$. The disrupted time series shows the information delay for sensor data access of an agent by the operator. Since sensor data is sent at less frequent intervals than the HBS, the operator has access to information that may not give an exact interpretation of agent distress if the disruption occurs after a significant time interval after the last sensor data transmission.

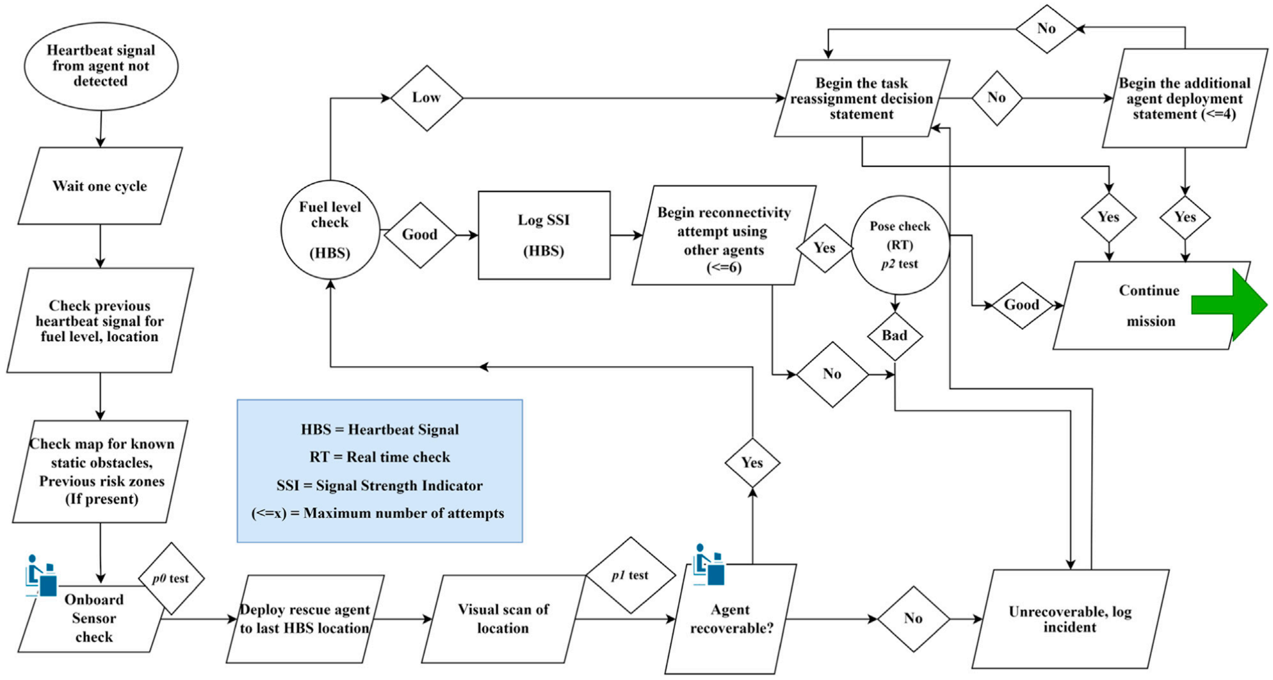

FIGURE 2 | SAR process for rescuing an agent in distress.

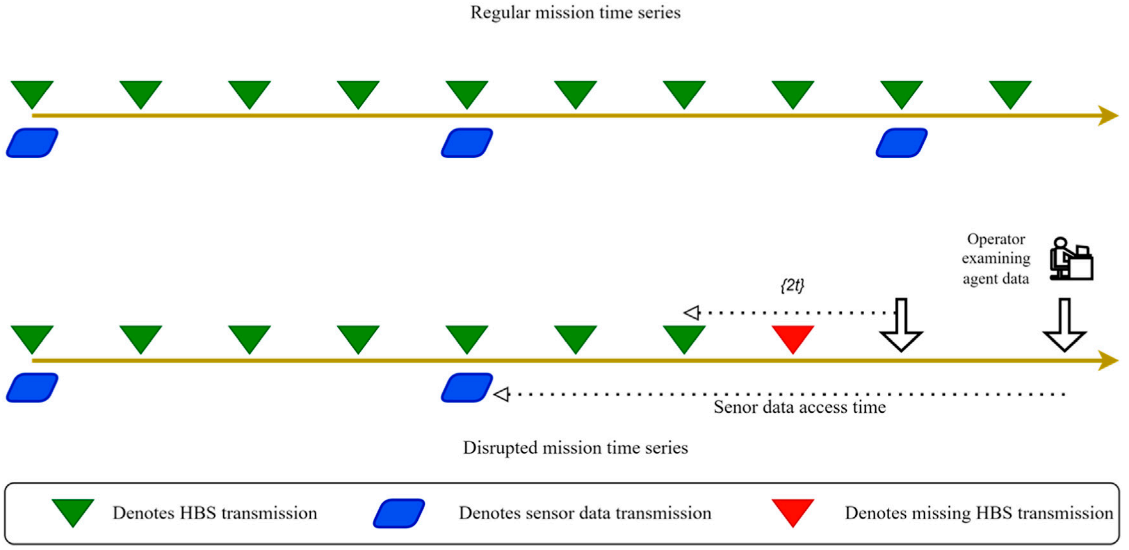

FIGURE 3 | HBS transmission and usage for regular and disrupted mission time series.

If an agent is missing, its past HBS record is retrieved and examined for its location during that transmission time interval. This location information may be outdated by a minimum value of the system realization time, that is, $2t$. A map overlay for known static obstacles is then used to determine if the agent was near obstacles during loss. A UAV agent can be distressed due to reasons such as collision with a static or dynamic obstacle, falling out of range from other agents in a mesh-based topology or with ground control in a directed topology, or issues with hardware components and fuel. Multiple pose checks are designed in the framework and conducted at each step to systematically eliminate the cause of disruptions. It is assumed that the agent, even when on the ground, has an open broadcast connection request to accept incoming connection requests from other agents or

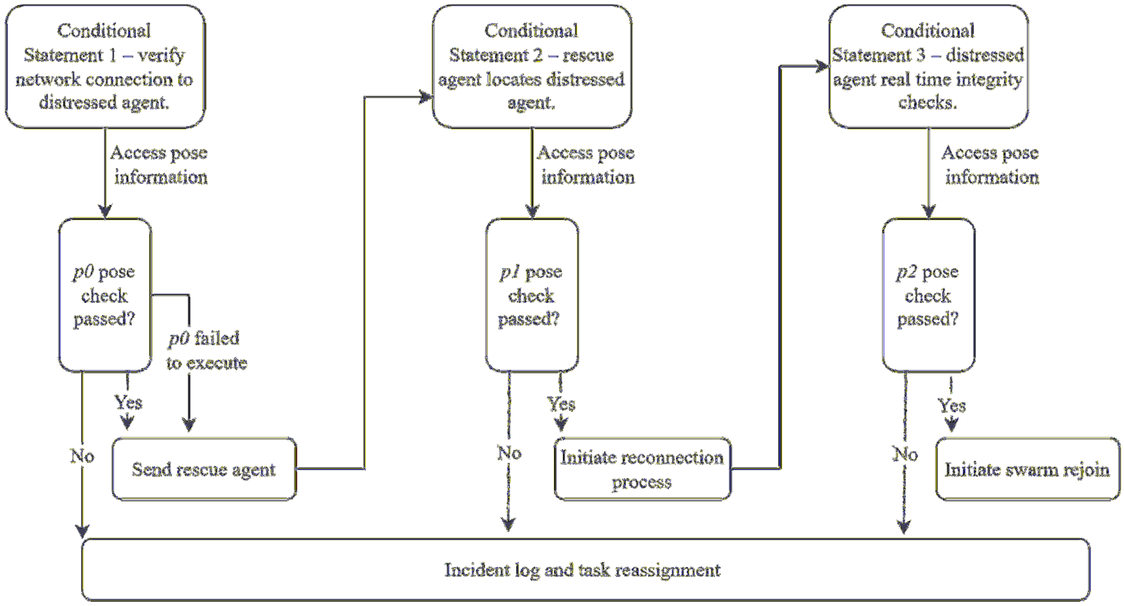

FIGURE 4 | Pose check workflow.

manual remote-control overrides. This feature is always present, even in basic agent builds. An agent that is connected to a network but has landed due to a collision, flock fragmentation, or getting stuck in an obstacle can still transmit an HBS with an on-ground indicator value of $O.G. = 1$. This indicates that an agent has landed but is still actively connected. However, this agent cannot be directly accepted back into the swarm. It is necessary to ensure that it can safely rejoin the swarm by gaining the necessary minimum altitude required by the agent to take a safe flight.

At specific decision points in the framework, agent status checks called pose checks are performed to gain additional information about the agent. The framework can perform three different checks: p_0, p_1, and p_2. Pose check flow conducted at different times during framework operation is shown in **Figure 4**.

Once an agent is realized to be in distress, an initial attempt is made to see if it is still possible to access its onboard vision sensor to conduct a preliminary pose check p_0. This checks if the agent has landed in such a position that it may be able to take off safely. Examples of passed and failed p_0 tests are shown in Section *Performance Tests*. The advantage of this method is that if the preliminary pose check fails, the framework can skip sending the rescue agent and directly move on to the unrecoverable agent process. However, this step is flexible: a rescue agent can still be deployed if the p0 check cannot be conducted.

If p_0 passes, rescue agent R moves to the location of the missing agent ($R_{loc(i,\ k-1)}$) and performs a visual scan of the location. The operator conducts real-time viewing of the rescue agent's camera data to conduct p_1. After an agent is located, the p_1 check using a rescue agent vision sensor is done to assess if the agent is in an environment from which it can take off safely. A fuel check using b_{level} and a network connection check using $SSI_{(G.C.,\ i)}$ are then performed. The SSI value contains agent connection data with ground control and neighboring agents. Depending on the network topology selected, an $SSI_{(G.C.,\ i)}$ value of 0 can be acceptable if the distressed agent connects to another agent rather than to ground control.

$$p_2 = (b_{level} > 40\%)\ \text{AND}\,(SSI_{(GC,i)} > 80\%)$$

This data is taken from the agent's previous HBS to create record logs of why the agent failed. This information is used to create risk zones as an information overlay in mission maps, a framework feature designed to reduce the failure of future agent movements on the same map.

If p_1 passes, the *rcm* messages are sent to reconnect with the agent. Once an agent is actively connected, a real-time pose check p_2 is conducted, which checks the current agent fuel level and network connection. In higher-level agents, this check can also take feedback from individual components onboard the vehicle, such as autopilot and motor sensors, to check for hardware integrity and orientation. If this pose check is passed, the distressed agent is deemed capable of rejoining the swarm. If the pose check fails, a log is created, and the agent's location is marked with an overlay that denotes the perceived reason for failure. A task re-assignment policy is then initiated to reassign the task of the lost agent to other swarm agents.

To date, probability maps have been a prevalent approach in SAR problems. Global or local maps are proposed that decompose ROI in grids [40], and a probability rate of the target being in each of the cells is calculated. Agents are encouraged to explore cells with a higher probability rating of the target being present in them. Similar approaches have been examined in [31] where agents not only create and maintain

TABLE 3 | Test observations and map used for the four performance tests.

Performance test (P.T.)	Test objectives	Map used	Number of flights
PT1	Observe time to distress, time to rescue, collision occurrence, and battery level check	M1	15
PT2	Observe p_1 and p_2, recovery, and operator log creation	M2	10
PT3	Observe p_0, p_1, p_2, recovery, and operator log creation	M3	10
PT4	Observe p_0, p_1, and p_2, time to distress, time to rescue, collision occurrence, and operator log creation	M4	10

observation map history, but maps from neighboring agents can also be combined. A similar logic is used in this case, where ground control creates and maintains a global risk map where each cell has an associated risk value. This is based on location data of previous agent loss, where an incident log is created every time an agent is lost in a particular area in the same map. This is especially useful in same-area routine flight scenarios where UAVs must visit the same area multiple times. Labeled hotspots can then be used as additional input constraints to path-generating algorithms by assigning proportional weights to high-risk zones, which the planning algorithms can then avoid or have issue mitigation resources ready if those areas are unavoidable.

Pose and orientation calculation can be upgraded with optimization loops coming from additional input sources. For example, the vision sensor data of the distressed UAV can be accessed, and an automatic pose orientation of the UAV can eliminate the need to dispatch a rescue UAV if the fallen UAV sends an unrecoverable camera pose. This was demonstrated during various experiments in which a human in the loop could access the sensor information of the distressed UAV to deduce its orientation. If determined to be unrecoverable, the agent's location is marked for post-mission recovery trials, and the swarm moves on directly to the task re-allocation phase of the fallen agent.

PERFORMANCE TESTS

Hardware and software tests were designed to test the proposed workflow under different conditions. Experiment range and series were selected considering the range and variability required to effectively demonstrate performance [53]. **Table 3** summarizes the primary objectives of each test, the map used, and the number of experimental flights performed. Overall, these tests represent a modular approach to developing and testing an SS-SAR framework for increasing the operational resiliency of a UAV swarm system.

Each performance test was associated with a map, as summarized in **Table 3**. The hardware tests were performed in maps M1 and M2, and the simulation tests were performed in maps M3 and M4. **Table 4** outlines characteristics of the map environments used in the performance tests.

The proposed framework is quite flexible regarding the agents that can be used. However, at minimum, lateral and downward vision sensors are required, along with either a GNSS module or capability for passive beacon georeferencing. Considering hardware and fly space limitations, an indoor location was used for hardware tests. The DJI EDU [50] UAV platform was chosen to perform hardware performance tests. These low-cost drones provide a basic environment for drone testing and flights. In the past, there have been multiple approaches to using Tello drones as platforms for singular and swarm development. The authors of [54] use Tello agents to demonstrate an automated swarm flight in a restricted flight space. A matrix formation control that uses Tello to display patterns was adopted in [55]. In [56], the DJI frame was used to build visionless sensing drones for obstacle avoidance and maze solving.

Related research such as this assisted in realizing the various limitations of the Tello platform during experiment design. The Tello agents are low-cost entry-level hardware and are intended for proof-of-concept experiments. Without a dedicated GNSS receiver, the agents rely on a VPS using the downward-facing camera module to localize using ground planes and additional GCPs. All recorded video and image data is streamed in real-time to ground control without storage and post-processing ability. While these constraints prevent executing a full-scale framework representation on these agents, our experiments modify the complete framework based on its modular structure. This modular and stepwise process permits testing smaller decision statements using simple Tello agents. **Table 5** contains manufacturer-provided specifications for the DJI drones. These specifications have been referenced from online user manuals [50].

Hardware Performance Tests

A modular and stepwise process was developed to test individual decision statements of the proposed SAR framework using simple Tello EDU agents. The objective of PT1 was to evaluate the time to distress, time to rescue, log any collision occurrence, and perform a battery level check. In PT1, two DJI agents were used as a part of the same swarm, with a rescue agent on active standby. A 2D visualization of M1 with initial agent positions and other mission information is shown in **Figure 5**.

One of the agents moving along the mission pads was forced to switch off its VPS to emulate a disruption condition. Meanwhile, the rescue agent was on active standby in the center of the mission area (**Figure 6**, left) and could take off once the distressed agent did not send an expected HBS (**Figure 6**, right). Using mission pad information transmitted by the distressed agent before it faced disruption, the rescue agent located the fallen agent (**Figure 7**, left), conducted pose checks, and sent *rcm* messages to the distressed agent. If the distressed agent received the messages, it switched on its VPS, allowing it to rejoin the swarm. The rescue agent then moved back to its deployment point to await the next distress event (**Figure 7**, right).

TABLE 4 | Map designations and properties.

Map designation	Map environment	Map design	Major map properties
M1	Indoor	Real-world	GCP referenced, residential, obstacle-free
M2	Indoor	Real-world	GCP referenced, office space, obstacles present
M3	Indoor	Simulation	GCP referenced office space and obstacles present. (M3 is the M2 space, recreated for simulation tests)
M4	Outdoor	Simulation	No GCP, outdoor terrain, obstacles present

TABLE 5 | DJI TELLO EDU spec sheet.

General parameters	Value
Weight (including propeller guards)	87 g
Maximum speed	17.8 mph (28.8 kph)
Maximum flight time	13 min
Camera	**Value**
Maximum image size	2,592 × 1,936
Video recording modes	H.D.: 1,280 × 720
Battery	**Value**
Capacity	1,100 mAh
Voltage	3.8 V
Energy	4.18 Wh

This performance test measured the time to distress, time to rescue and observed the number of collisions, rescue decisions, and the battery threshold value. The only pose check conducted to evaluate the b_{level} of the distressed agent.

PT2 again used two regular agents and one rescue agent on map M2. The objective of PT2 was to observe p_1 and p_2, and to attempt a recovery. A 2D top-down representation of M2 is in **Figure 8**. The number of GCPs could be increased or decreased with a maximum number of up to 20 GCPs placed in the fly space.

Figure 9 shows the M2 space where the three agents were released. The distress condition was simulated for one agent, where it landed behind the table. The rescue agent moved on location to conduct pose checks and begin recovery attempts. The distressed agent was not visible in the global view. However, various situations were observed using the rescue agent and p_1. **Figure 10** (top-right) shows the rescue agent's POV, where the distressed agent fell at an oblique angle.

Additionally, as the agent had strayed under the table, the height of the table prevented the agent from gaining the minimum altitude required to conduct a safe rejoin operation. This exemplifies how p_1 helps understand the distressed agent's situation. **Figure 10** (bottom-right) shows a different situation where the agent has landed in a pose that could allow it to take off. However, its minimum altitude rejoin value was still greater than the obstacle dimensions. In both situations, the operator recommended that further recovery operations be terminated.

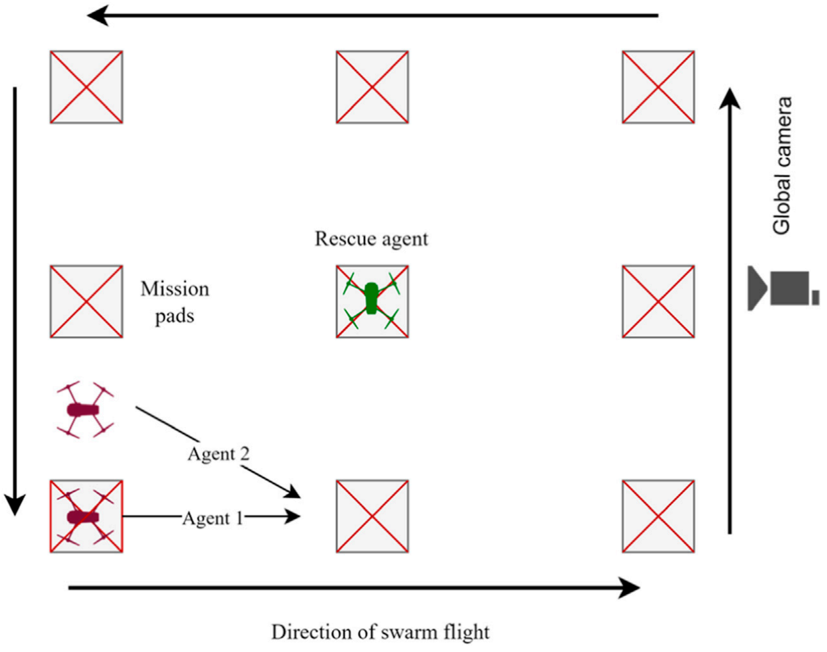

FIGURE 5 | 2D representation of M1 space.

FIGURE 6 | Preliminary framework test PT1 in the M1 space.

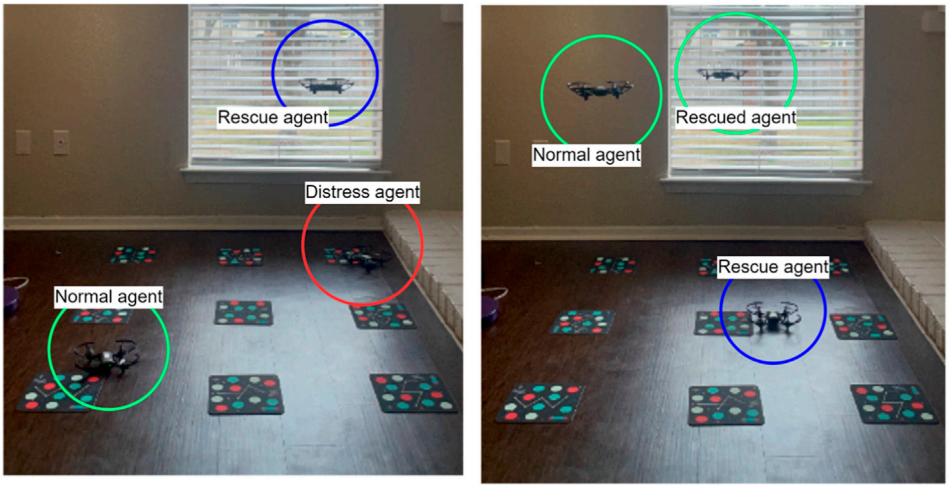

FIGURE 7 | PT1 in the M1 space shows a swarm agent's distress and recovery.

TABLE 6 | Simulation parameters for PT3.

Simulation parameters	Description
Target space	Close re-creation of physical space M2 for PT3
Simulation time	<500 s (Variable)
Size	20 m × 20 m

Figure 11 shows a third situation where the agent landed in a position from which it could take off, plus the right side figure shows the rescue agent's POV from which the operator determined that obstacle dimensions did not impede the distressed agent's safe rejoin procedure upon reconnection. In this situation, the operator recommended the framework to carry further rescue statements on the distressed agent.

Simulation Performance Tests

Indoor and outdoor scenarios to test the proposed framework were modeled in CoppeliaSim, formerly VREP [57]. **Table 6** outlines the basic simulation parameters for PT3 and PT4. PT3 was a simulation experiment carried out on M3 (**Figure 12A**), which is a close recreation of the M2 space used in the hardware experiments. The primary purpose of PT3 was to evaluate p_0, p_1, and p_2 and attempt a recovery.

A simple table and chair environment are used to show a failed p_1 and p_2 scenario (**Figure 12B**). Pose check p_1 used the agent vision sensor information to realize that the distressed agent failed in an inverted position. A normal decision cycle prevents the rescue agent from deploying on a failed p_1; however, a forced p_2 cycle using a rescue agent shows that the table dimensions would hinder a safe rejoin maneuver of

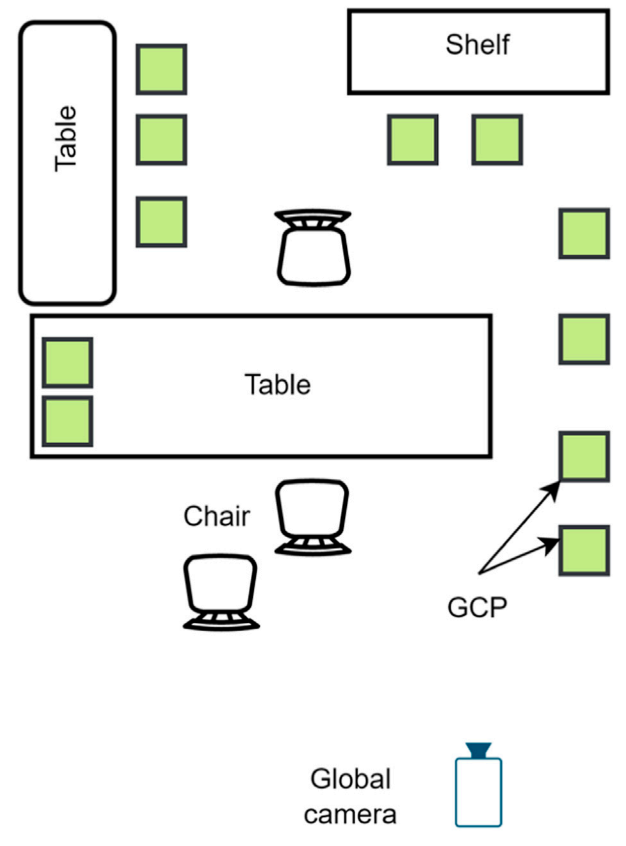

FIGURE 8 | A 2D representation of the M2 map (Not to scale).

the distressed agent even if the agent were not in an inverted position.

The primary purpose of PT4 was to evaluate p_0, p_1, and p_2, observe time to distress, time to rescue, and log any collision occurrence. **Table 7** outlines the basic simulation parameters for PT4. **Figure 13** shows PT4 on M4, where pose checks p_0, p_1, and p_2 were tested along with successful swarm rejoin scenarios. An abstract cube was placed in the field of view of the distressed agent to indicate its orientation for checking p_0. The rescue agent was then used to determine p_1 by observing the status of the distressed agent. Finally, a p_2 test evaluated if the distressed agent sensed it could rejoin the swarm.

Figure 14 shows an updated map where trees were present as obstacles that hindered agent progress and rejoin maneuvers. Here, multiple agents in the distress scenario were tested, where one agent landed in an inverted position and the other in a normal position. The rescue agent conducted a p_2 check on both agents to determine which agent could be safely recovered. Floating view windows in the figure show p_1 checks by both agents and a p_2 check by the rescue agent on one of the distressed agents.

RESULTS

The following section highlights observations recorded during each performance test and their analysis. **Figure 15** shows PT1 time to distress logs, the time when the agent first experienced an issue, and the time to rescue, which is the amount of time the rescue agent took to move to the position

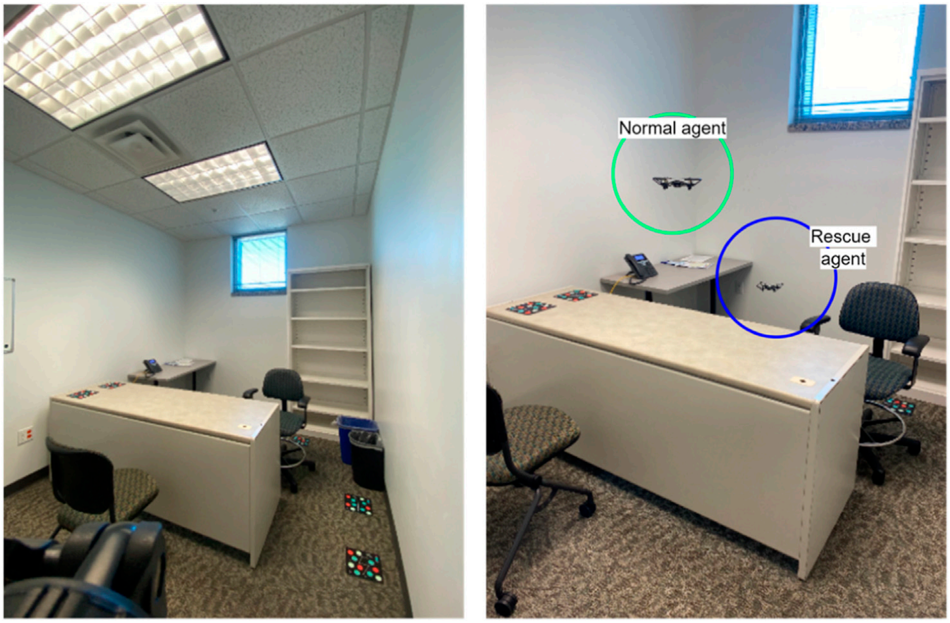

FIGURE 9 | Real-world M2 space with 2 regular and one rescue agent for PT2.

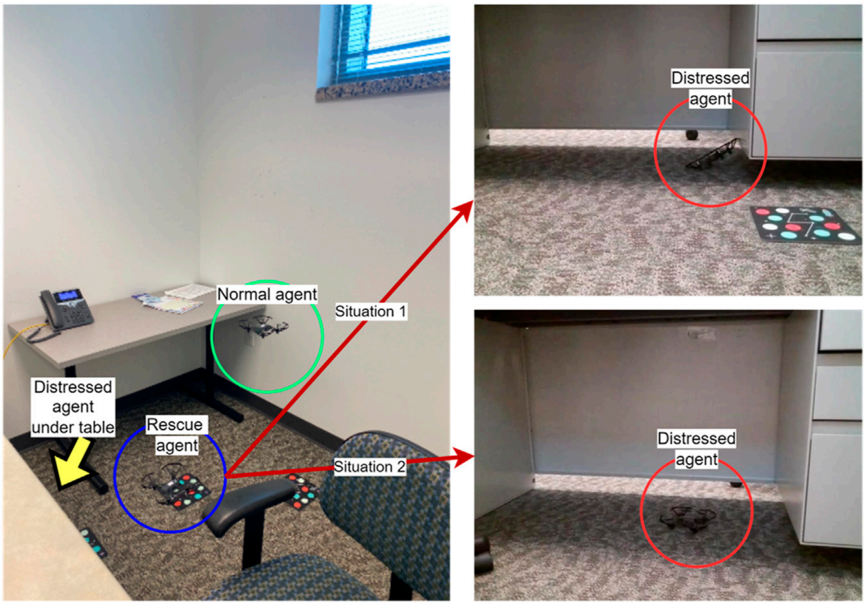

FIGURE 10 | A Different global view of the M2 space and two floating views from the rescue agent's POV.

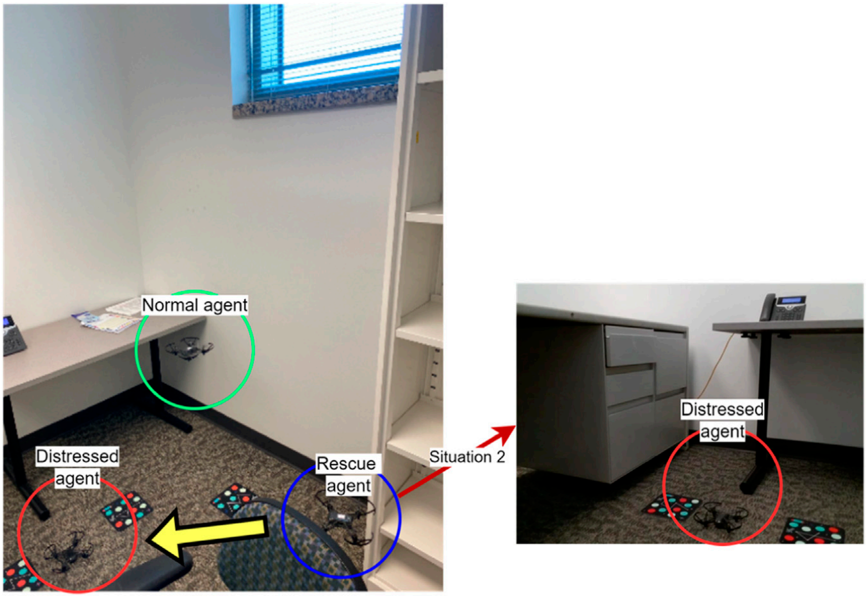

FIGURE 11 | A scenario where p_1 and p_2 are successful on M2 during PT2.

of the fallen agent and rescue it. Out of 15 flights, the rescue agent successfully rescues the distressed Tello agent nine times, as denoted by a green dot in **Figure 15**. A preliminary p_2 check was performed using the available battery percentage when *rcm* was successful. The battery values are in **Figure 16**. If the battery value after successful reconnection was established between the Rescue and the distressed agent was below the given threshold (50%), the distressed agent was deemed incapable of rejoining the swarm. This was observed during flights 5, 7, 12, and 14. Each flight was independent, and the battery was charged to maximum capacity before each flight. Collision occurrence was counted when the rescue agent experienced collision at any time during the rescue process. As such, those flights were recorded as an unsuccessful recovery. During flights 4 and 10, the rescue agent experienced a collision and could not recover the fallen agent successfully; these flights were logged as failures. It was observed

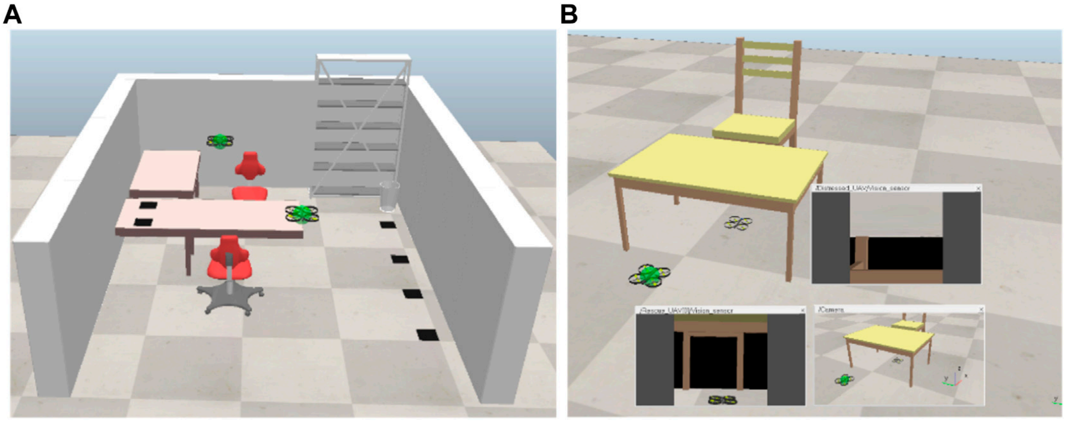

FIGURE 12 | M3 map **(A)** and basic furniture **(B)** designs for PT3.

TABLE 7 | Simulation parameters for PT4.

Simulation parameters	Description
Target space	Outdoor environment
Simulation time	<500 s (Variable)
Size	100 m × 100 m

that close interaction with swarm agents in the constrained airspace caused unpredictable drifts in agent movement due to induced airflow, resulting in collisions and crashes. Overall, a recovery rate of 60% was thus calculated for PT1.

Figure 17 shows PT2 performed on M2. For ten flights, p1 and p_2 rescue decisions were recorded. This test aimed to observe these pose checks and how they affect agent recovery. Flights 4, 7, 8, and 9 showed a successful recovery. Flights 1 to 3 and 5 failed p_1, where it was determined by the operator using the rescue agent that the fallen agent was not in a position from which it could safely take off. In M2, this scenario was due to indoor obstacles, such as furniture, that might prevent the agents' safe take-off ability. In Flight 6, the agent passed p_1, which denoted it was in an orientation and position that could enable safe take-off; however, it failed real-time pose check p_2. For flight 10, both p_1 and p_2 were successful. However, the agent could not take off due to an internal malfunction. A success rate of 40% for PT2 was observed.

In addition to previous pose checks, PT3 on M3 also performed preliminary pose check p_0 on the agents using the distressed agent's onboard vision sensor, as shown in **Figure 12**. **Figure 18** presents the ten flights performed. Flights 3, 5, 6, and 8 successfully recovered the distressed agent. For flights 1, 7, and 10, the agent failed p_0, indicating that the agent was not in a position to take off safely. As a result of the preliminary pose check failing, the rescue agent was not deployed to conduct

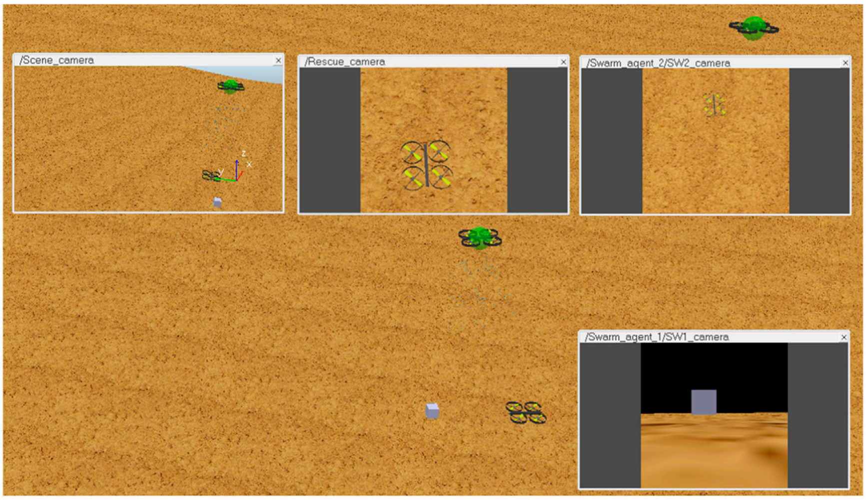

FIGURE 13 | Preliminary M4 map with no obstacles and abstract cube for p_0 checks.

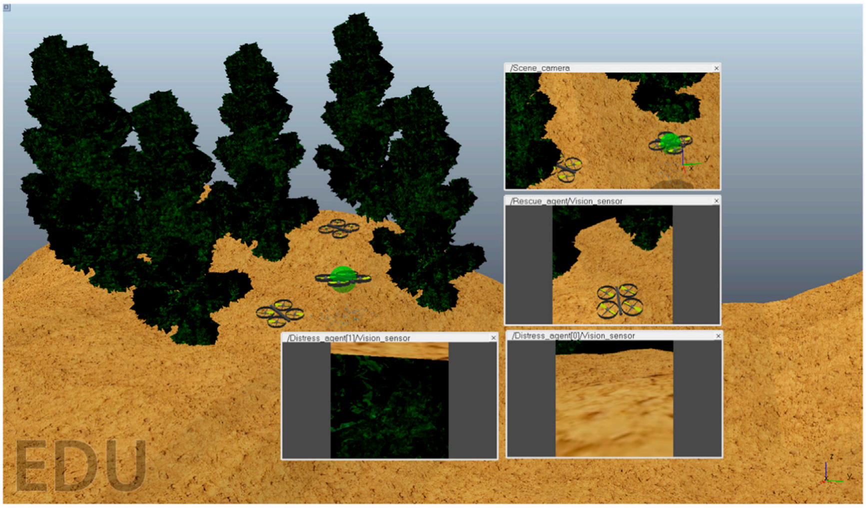

FIGURE 14 | M4 map with outdoor terrain and tree obstacles.

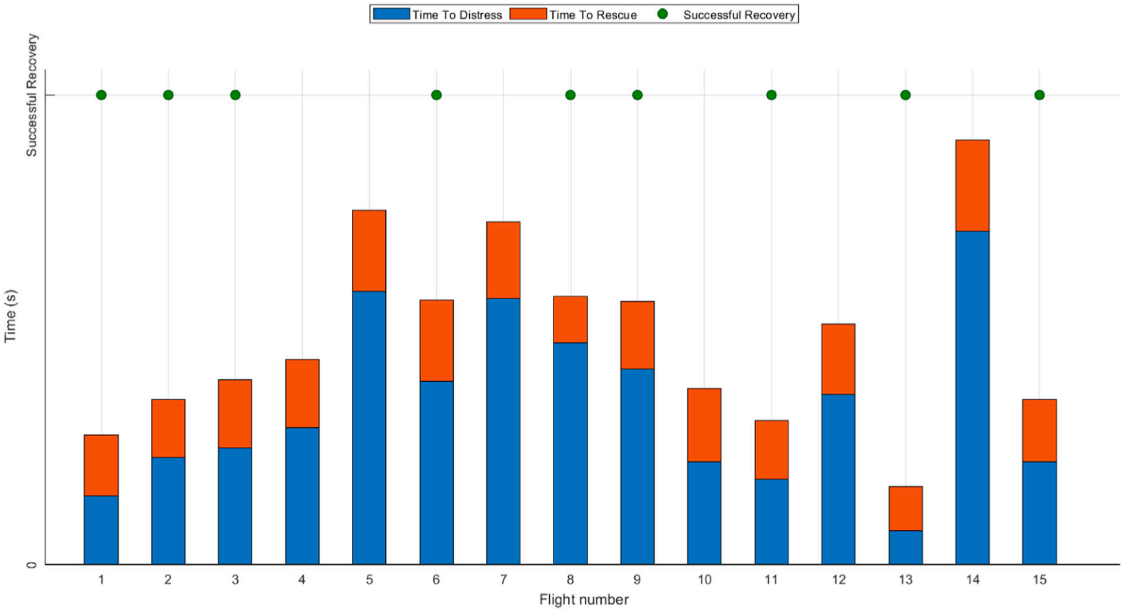

FIGURE 15 | PT1 time to distress and time to rescue with successful recovery decisions.

further observations. In Flight 2, p_0 passed. However, p_1 failed, leading to a failed rescue attempt. In Flights 4 and 9, a successful p_0 and p_1 were observed. However, the agent failed real-time check p_2 and was thus labeled unrecoverable. A success rate of 40% was observed for this test.

PT4 on M4 further examined an additional ten flights, and the results are shown in **Figure 19**, where time to distress logs the time a swarm agent experiences an issue, and time to rescue logs the time the rescue agent takes during rescue attempts. Flights 4, 5, 6, 7, and 9 showed the rescue agent's successful recovery of the distressed agent.

Further examination of operational parameters, as shown in **Figure 20**, gives additional failure information. In flights 1, 8, and 10, the distressed agent passed p_0, which denoted its orientation passed requirements for safe rejoin. However, p_1 failed. Since PT4 was performed on an outdoor terrain map that included tree obstacles, the primary reason for p_1 to fail was the tress obstructing safe rejoin maneuvers. In-flight 2, the distressed agent failed to p0 itself, as denoted by the onboard sensor that gave information regarding its orientation and crash position. The floating window views in **Figure 14** for the distressed agent one vision sensor FOV show an example of an agent that has

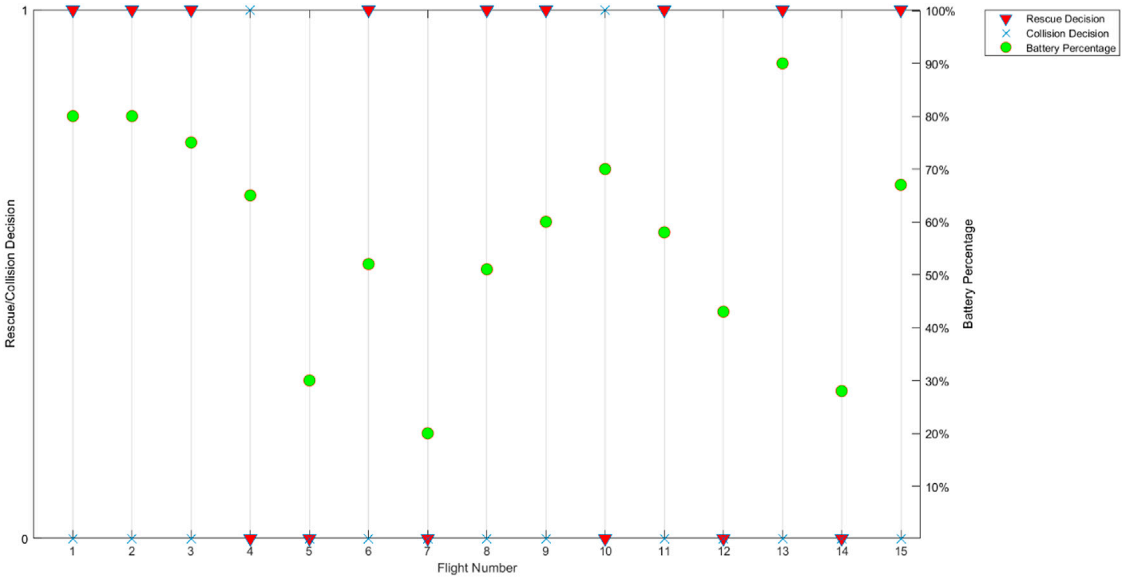

FIGURE 16 | PT1 rescue decision and collision occurrence plotted with battery percentage values.

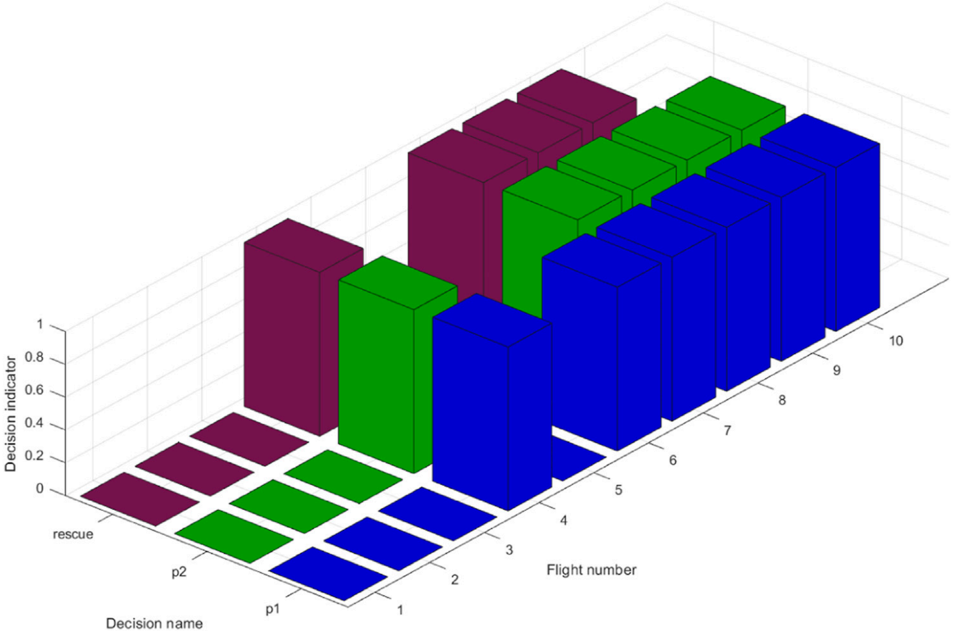

FIGURE 17 | PT2 pose checks and rescue decisions.

landed upside down. The inverted image shows the ground above and tree foliage below. In flight 3, the distressed agent passed all required pose checks; however, a collision with the rescue agent resulted in a failed attempt. Overall, a success rate of 50% was observed. A summary of success rates for all performance tests is shown in **Figure 21**.

A test on a sparse block map depicted in **Figure 22** was used to log agent loss and recovery incidents. Every distress signal was mapped as a triangle or square in that process. The triangles indicate agent loss due to network issues, and the squares indicate agent loss due to collision. Red shapes indicate an unsuccessful attempt at recovery, and blue shapes indicate successful agent recovery. Each shape is the result of a separate flight, and consecutive flights did not have prior risk zone information. However, creating such risk zones can then be used as future input parameters to create safe flight paths.

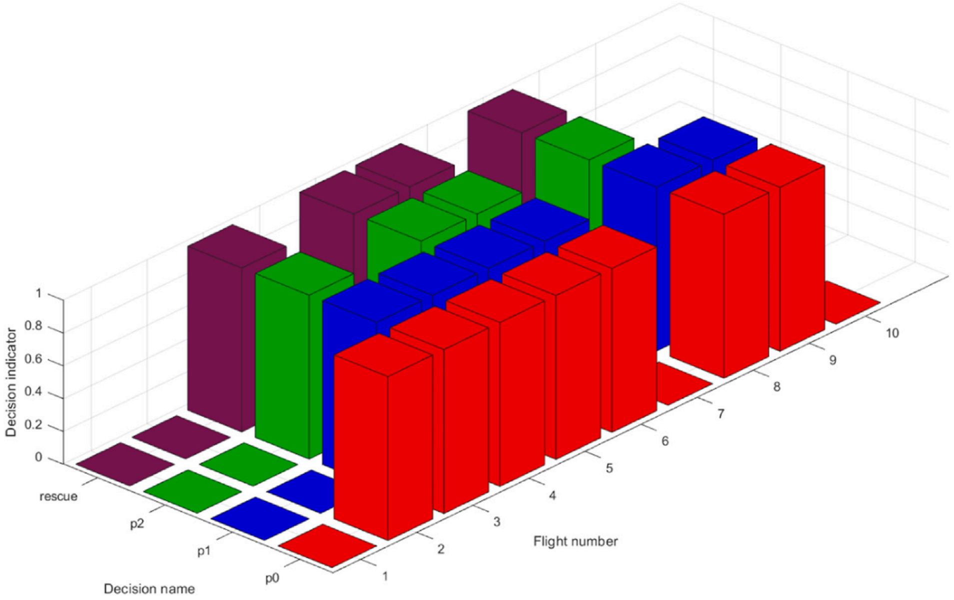

FIGURE 18 | PT3 pose checks and Rescue decisions.

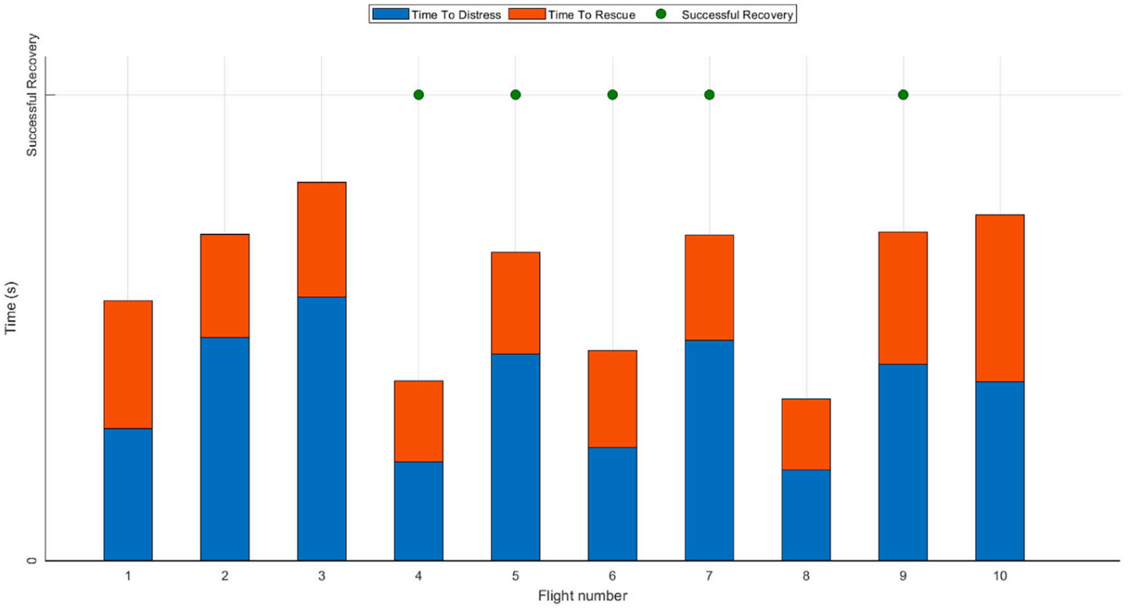

FIGURE 19 | PT4 time to distress and time to rescue with successful recovery decisions.

For example, if an area sees increased collisions due to dense obstacle geometry, a threat area can be modeled where agents entering that area do not venture below a preset altitude to avoid collisions. If agents moving to a particular area lose connection with ground control, the next iteration framework run will adjust the upper bound distance between the agents, which defines the maximum distance between two agents based on SSI_n. Adjusting the upper bounds will result in agents flying in close formation and using data hop pathways to connect to ground control and prevent agent loss due to network range limitations. Future work using this approach can demonstrate adaptability, robustness, and emergent behavior in the swarm based on simple governance rules.

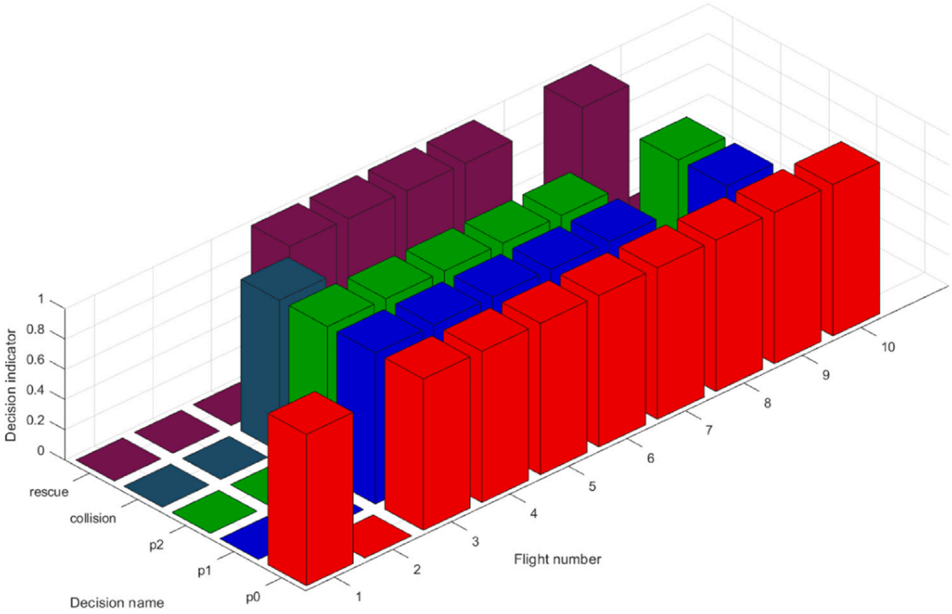

FIGURE 20 | PT4 pose checks and rescue decisions.

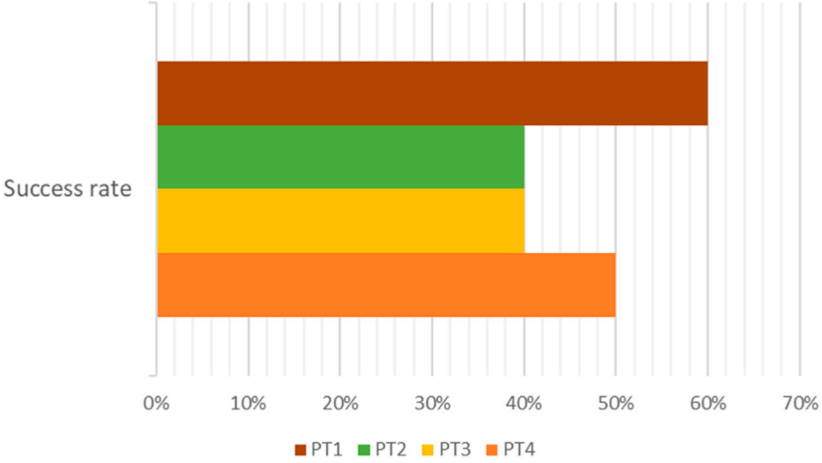

FIGURE 21 | Summary of PT results.

DISCUSSION AND FUTURE RESEARCH DIRECTIONS

The proposed framework is a preliminary step in developing robust methodologies for evaluating swarm awareness toward the wellbeing of its constituent agents. This includes testing capabilities such as keeping track of each agent's progress toward its task, realizing the occurrence of agents in distress, locating the distressed agents, and initiating rescue operations to enable them to rejoin the swarm. Several modifications could be implemented via the modular nature of the designed framework, as initiated by research directions summarized below.

It is crucial to consider the impact of emerging regulations on UAV operations, particularly the recent implementation of the FAA's Remote Identification (RID) rule [58]. This regulation mandates the use of Remote Identification modules on certain UAVs, allowing for the open broadcast and identification of these agents during flight. This rule ensures safer airspace and promotes regulated use of UAVs, UAV swarms, and their applications [59]. When integrated into our rescue framework, the potential for such information can significantly enhance tracking and rescue performance. By leveraging the real-time identification capabilities provided by RID, it is foreseen that such frameworks can precisely locate and rescue other agents within

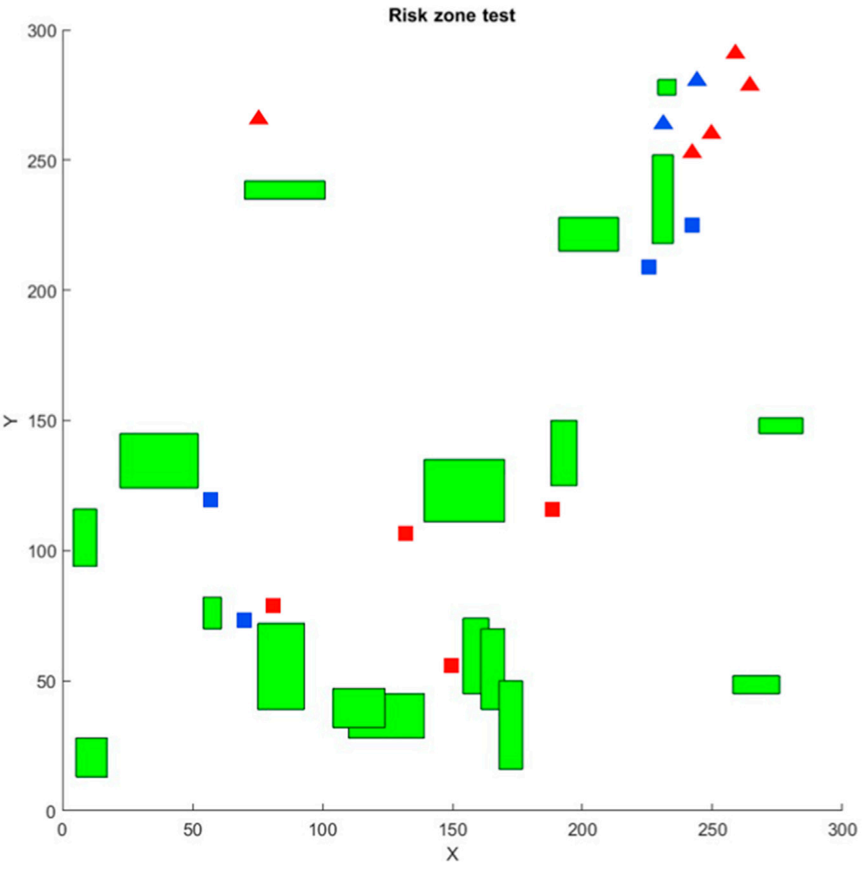

FIGURE 22 | Incident log overlay on a generic map as inputs for future iterations.

TABLE 8 | Summary of recent work on optimal abort policies, task rescheduling, and dynamic risk assessment.

Reference	Implementation	Description
[72]	General systemic deployment	Optimally aborting subtasks in heterogeneous swarms to increase overall unit survivability rate
[73]	General systemic deployment	Design the best abort strategy for multi-unit swarms based on the probability of external shocks damaging units
[74]	Single UAV focused	Design of replacement policies and maintenance cost for UAV reconnaissance system
[75]	Single UAV focused	Dynamic allocation of a fixed number of components to increase the mission completion rate by UAV in a reconnaissance scenario
[76]	UAV swarm focused	Considering the cost of damaged agents and unfinished tasks to compute abort policies
[77]	UAV swarm focused	Evaluate system mission reliability and suggest swarm maintenance strategies
[78]	UAV swarm focused	Incorporating abort policies in multi-UAV routing as a response to external shocks to ensure agent wellbeing
[79]	UAV swarm focused	Consider degradation level, mission time, and equipment health to create dynamic mission abort policies

the swarm more effectively, thus bolstering the overall efficiency and reliability of the proposed UAV rescue mechanism. Furthermore, exploring the compatibility and interoperability of our rescue system with other upcoming regulatory frameworks will be essential in ensuring the seamless integration and widespread adoption of our research in real-world UAV swarm applications.

Currently, some sections of the framework involve human decision-making. Most notable is the analysis of the distressed agent pose data transmitted by the rescue agent. The human operator observes the images to create a preliminary decision on the fallen agent's possibility of rejoining the swarm. The human in-loop component can be reduced by adding autonomous UAV detection capability that uses vision sensor data, deep learning, and image processing techniques. This is possible using approaches such as [60] that use agent vision sensors for target analysis. An additional upgrade involves multiple agents to capture disruption and distressed agent information from different angles to gather a richer dataset.

A modified task re-allocation algorithm would enable additional agents to join the swarm and take up the interrupted task of the fallen agent or a re-allocation scheme for existing swarm agents to assume responsibility for the incomplete task. Resource allocation implementations such as in [61–63] exist that could be implemented. Further experiments could be explored for loss rates in the same airspace with LiDAR (for obstacle detection) and preset waypoints in a map (using GNSS). This would expand the feature of the existing framework to create probability maps. While current risk zones were labeled using agent failure location data obtained from transmitted HBS, future experiments may include the presence of VRRZ. This is the system's ability to create variable radius risk zones. Each risk zone can have a variable radius, thus allowing the mapping of larger disruptive structures to be represented more accurately.

The above study uses agents with similar capabilities acting in the same operational space. Including a diverse range of agents identified by their differences in nature, hardware, or operational space introduces heterogeneity in the swarm. The impacts of such inclusion on the performance of SAR agents can also be explored. Some existing research investigates the possibility of using a swarm composed of heterogeneous agents for victim detection after a disaster [64]. Although their main goal was exploring how swarm heterogeneity can affect performance, they modeled a target search and rescue problem to study it. Their proposed technique differentiated between different agents and labeled them as heterogeneous using behavior trees. A positive correlation was produced between the swarm's heterogenous capability and the time to search and rescue the target. A similar approach can be explored in the future, where differently abled robots are introduced in the swarm and are tasked with looking for swarm agents whose operations have been disrupted during mission progress. In the above experiments, all distressed agents were located on the ground. Thus, adding a UGV to track and locate the fallen agents to create an in-depth pose check analysis would be a logical step for further exploration. Several implementations of heterogeneous swarms exist, such as UAV-UGV collaboration [65], UAV-UWSV [29], and UAV-UGV-UWSV [66], demonstrating promise for more effective results than a single operational space swarm.

Intrusion detection systems can be implemented on the UAV network as a backend process. While IDSs are most prevalent on traditional networks to deter unwanted network access and activity, current lightweight versions have been shown to run reasonably well on MANET and FANET deployments with acceptable performance [19, 67–70]. Various types of IDS are available depending on their makeup and method of detecting malicious entities [71]. IDS could detect external agents attempting to maliciously disturb swarm operations. Similar approaches could also address ground-based attempts to take over swarm networks. The possible advantage would be the existing periodically transmitted HBS signal that can be used as input to any IDS. Adding network transmission data from each agent in the HBS could be used to design either a rule-based or anomaly-based lightweight IDS, at the very least. In this way, the SAR framework could provide additional security features to the swarm using inherently built structures.

A different approach taken to designing robust behavior was observed as a way of defining reliability in systems. The methods used in this category implement preemptive strategies for maintenance, abort policies, or recovery actions. This alternative form of resilience integration calls for an independent study. However, the results of the brief survey conducted on it are summarized below. These methods can be viewed as possible implementations and upgrades to this proposed SS-SAR framework. **Table 8** summarizes the examined work based on their development focus being broader systemic implementations or UAV swarm-focused.

CONCLUSION

This research addresses gaps in current swarm resiliency research by addressing swarm-specific SAR rather than application-specific SAR. The approach was not to replace current SAR methodologies but to create an add-on that enables them to keep track of swarm agents while performing other functions. Modular experiments conducted on real-world hardware and simulations validated the need for, the possibility of, and the success rate of swarm-specific SAR approaches. While low-cost Tello drones were limited in their ability to handle a complete SS-SAR framework, they were crucial in testing the constituent process of the framework, such as reconnection protocols and pose check handlers. Simulation results provided a greater insight into how such frameworks can handle swarm agent loss. Experimental results prove that focusing on this approach to resiliency integration in multi-agent systems can produce the anticipated benefits. Recovery rates of distressed agents during and after the mission process increased drastically, especially in systems with no contingency rules. UAV swarms are complex and highly dynamic, making integrating resilience factors much more arduous. A system must exhibit awareness and diagnosis capability regarding its health before and after a disruption to efficiently produce solutions to mitigate said disruptions. This swarm-specific SAR framework is a crucial design step in that direction.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusion of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

Conceptualization, Methodology, Validation, Writing- Original draft- AP. Methodology, Writing- review and editing, Visualization, Supervision- FM. All authors contributed to the article and approved the submitted version.

CONFLICT OF INTEREST

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

REFERENCES

1. Siemiatkowska B, Stecz W. A Framework for Planning and Execution of Drone Swarm Missions in a Hostile Environment. *Sensors (Basel)* (2021) 21:4150. doi:10.3390/s21124150
2. Ko Y, Kim J, Duguma DG, Astillo PV, You I, Pau G. Drone Secure Communication Protocol for Future Sensitive Applications in Military Zone. *Sensors (Basel)* (2021) 21:2057. doi:10.3390/s21062057
3. Allan BM, Nimmo DG, Ierodiaconou D, VanDerWal J, Koh LP, Ritchie EG. Futurecasting Ecological Research: The Rise of Technoecology. *Ecosphere* (2018) 9. doi:10.1002/ecs2.2163
4. Nex F, Duarte D, Steenbeek A, Kerle N. Towards Real-Time Building Damage Mapping With Low-Cost UAV Solutions. *Remote Sensing* (2019) 11:287. doi:10.3390/rs11030287
5. Noor NM, Abdullah A, Hashim M. Remote Sensing UAV/Drones and Its Applications for Urban Areas: A Review. In: *IGRSM International Conference and Exhibition on Geospatial & Remote Sensing*. Malaysia: Kuala Lumpur (2018).
6. Hildmann H, Kovacs E. Review: Using Unmanned Aerial Vehicles (UAVs) as Mobile Sensing Platforms (MSPs) for Disaster Response, Civil Security and Public Safety. *Drones* (2019) 3:59. doi:10.3390/drones3030059
7. Klaine PV, Nadas JPB, Souza RD, Imran MA. Distributed Drone Base Station Positioning for Emergency Cellular Networks Using Reinforcement Learning. *Cognit Comput* (2018) 10:790–804. doi:10.1007/s12559-018-9559-8
8. Tsouros DC, Bibi S, Sarigiannidis PG. A Review on UAV-Based Applications for Precision Agriculture. *Information* (2019) 10:349. doi:10.3390/info10110349
9. Panagiotis Radoglou-Grammatikis PS, Lagkas T, Moscholios I. A Compilation of UAV Applications for Precision Agriculture. *Computer Networks* (2020) 172:107148. doi:10.1016/j.comnet.2020.107148
10. Lygouras E, Santavas N, Taitzoglou A, Tarchanidis K, Mitropoulos A, Gasteratos A. Unsupervised Human Detection With an Embedded Vision System on a Fully Autonomous UAV for Search and Rescue Operations. *Sensors (Basel)* (2019) 19:3542. doi:10.3390/s19163542
11. Dong J, Ota K, Dong M. UAV-Based Real-Time Survivor Detection System in Post-Disaster Search and Rescue Operations. *IEEE J Miniaturization Air Space Syst* (2021) 2:209–19. doi:10.1109/jmass.2021.3083659
12. Phadke A, Antonio Medrano F, Chu T. Engineering Resiliency in UAV Swarms–A Bibliographic Analysis. *J Phys Conf Ser* (2022) 2330:012007. doi:10.1088/1742-6596/2330/1/012007
13. Ordoukhanian EM, Azad M. Engineering Resilience Into Multi-UAV Systems. *Proced Comp Sci* (2019) 153:9–16. doi:10.1016/j.procs.2019.05.050
14. Hosseini S, Barker K, Ramirez-Marquez JE. A Review of Definitions and Measures of System Resilience. *Reliability Eng Syst Saf* (2016) 145:47–61. doi:10.1016/j.ress.2015.08.006
15. Woods DD. Four Concepts for Resilience and the Implications for the Future of Resilience Engineering. *Reliability Eng Syst Saf* (2015) 141:5–9. doi:10.1016/j.ress.2015.03.018
16. Phadke A, Medrano FA. Towards Resilient UAV Swarms—A Breakdown of Resiliency Requirements in UAV Swarms. *Drones* (2022) 6:340. doi:10.3390/drones6110340
17. Zhang J, Zhou W, Wang X. UAV Swarm Navigation Using Dynamic Adaptive Kalman Filter and Network Navigation. *Sensors (Basel)* (2021) 21:5374. doi:10.3390/s21165374
18. Bassolillo SR, D'Amato E, Notaro I, Blasi L, Mattei M. Decentralized Mesh-Based Model Predictive Control for Swarms of UAVs. *Sensors (Basel)* (2020) 20:4324. doi:10.3390/s20154324
19. Tan X, Su S, Zuo Z, Guo X, Sun X. Intrusion Detection of UAVs Based on the Deep Belief Network Optimized by PSO. *Sensors (Basel)* (2019) 19:5529. doi:10.3390/s19245529
20. Luo R, Zheng H, Guo J. Solving the Multi-Functional Heterogeneous UAV Cooperative Mission Planning Problem Using Multi-Swarm Fruit Fly Optimization Algorithm. *Sensors (Basel)* (2020) 20:5026. doi:10.3390/s20185026
21. Kwa HL, Philippot J, Bouffanais R. Effect of Swarm Density on Collective Tracking Performance. *Swarm Intelligence* (2023) 17:253–81. doi:10.1007/s11721-023-00225-4
22. Abhishek P, Medrano FA. Examining Application-Specific Resiliency Implementations in UAV Swarm Scenarios. *Intelligence & Robotics* (2023) 3:453–78. doi:10.20517/ir.2023.27
23. Phadke A, Medrano A. *A Resilient Multi-UAV System of Systems (SoS)*. Report no. 2771-9359 (2021).
24. Garg V, Tiwari R, Shukla A. Comparative Analysis of Fruit Fly-Inspired Multi-Robot Cooperative Algorithm for Target Search and Rescue. In: 2022 IEEE World Conference on Applied Intelligence and Computing (AIC); 17-19 June 2022; Sonbhadra, India (2022). p. 444–50.
25. Cao Y, Qi F, Jing Y, Zhu M, Lei T, Li Z, et al. Mission Chain Driven Unmanned Aerial Vehicle Swarms Cooperation for the Search and Rescue of Outdoor Injured Human Targets. *Drones* (2022) 6:138. doi:10.3390/drones6060138
26. Oh D, Han J. Smart Search System of Autonomous Flight UAVs for Disaster Rescue. *Sensors (Basel)* (2021) 21:6810. doi:10.3390/s21206810
27. Cardona GA, Calderon JM. Robot Swarm Navigation and Victim Detection Using Rendezvous Consensus in Search and Rescue Operations. *Appl Sci* (2019) 9:1702. doi:10.3390/app9081702
28. Alotaibi ET, Alqefari SS, Koubaa A. LSAR: Multi-UAV Collaboration for Search and Rescue Missions. *IEEE Access* (2019) 7:55817–32. doi:10.1109/access.2019.2912306
29. Yang T, Jiang Z, Dong J, Feng H, Yang C. Multi Agents to Search and Rescue Based on Group Intelligent Algorithm and Edge Computing. In: 2018 IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData); 30 July 2018 - 03 August 2018; Halifax, NS, Canada (2018). p. 389–94.
30. Ho Y-H, Tsai Y-J. Open Collaborative Platform for Multi-Drones to Support Search and Rescue Operations. *Drones* (2022) 6:132. doi:10.3390/drones6050132
31. Yan P, Jia T, Bai C. Searching and Tracking an Unknown Number of Targets: A Learning-Based Method Enhanced With Maps Merging. *Sensors (Basel)* (2021) 21:1076. doi:10.3390/s21041076
32. Papaioannou S, Kolios P, Theocharides T, Panayiotou CG, Polycarpou MM. Distributed Search Planning in 3-D Environments With a Dynamically Varying Number of Agents. *IEEE Trans Syst Man, Cybernetics: Syst* (2023) 53:4117–30. doi:10.1109/tsmc.2023.3240023
33. Liu S, Yao W, Zhu X, Zuo Y, Zhou B. Emergent Search of UAV Swarm Guided by the Target Probability Map. *Appl Sci* (2022) 12:5086. doi:10.3390/app12105086
34. Li L, Xu S, Nie H, Mao Y, Yu S. Collaborative Target Search Algorithm for UAV Based on Chaotic Disturbance Pigeon-Inspired Optimization. *Appl Sci* (2021) 11:7358. doi:10.3390/app11167358
35. Cho S-W, Park J-H, Park H-J, Kim S. Multi-UAV Coverage Path Planning Based on Hexagonal Grid Decomposition in Maritime Search and Rescue. *Mathematics* (2021) 10:83. doi:10.3390/math10010083
36. Walker O, Vanegas F, Gonzalez F. A Framework for Multi-Agent UAV Exploration and Target-Finding in GPS-Denied and Partially Observable Environments. *Sensors (Basel)* (2020) 20:4739. doi:10.3390/s20174739
37. Alhaqbani A, Kurdi H, Youcef-Toumi K. Fish-Inspired Task Allocation Algorithm for Multiple Unmanned Aerial Vehicles in Search and Rescue Missions. *Remote Sensing* (2020) 13:27. doi:10.3390/rs13010027
38. Yue W, Guan X, Wang L. A Novel Searching Method Using Reinforcement Learning Scheme for Multi-UAVs in Unknown Environments. *Appl Sci* (2019) 9:4964. doi:10.3390/app9224964
39. Li X, Chen J, Deng F, Li H. Profit-Driven Adaptive Moving Targets Search With UAV Swarms. *Sensors (Basel)* (2019) 19:1545. doi:10.3390/s19071545
40. de Alcantara Andrade FA, Reinier Hovenburg A, Netto de Lima L, Rodin CD, Johansen TA, Storvold R, et al. Autonomous Unmanned Aerial Vehicles in Search and Rescue Missions Using Real-Time Cooperative Model Predictive Control. *Sensors (Basel)* (2019) 19:4067. doi:10.3390/s19194067
41. Huang S, Teo RSH, Kwan JLP, Liu W, Dymkou SM. Distributed UAV Loss Detection and Auto-Replacement Protocol With Guaranteed Properties. *J Intell Robotic Syst* (2019) 93:303–16. doi:10.1007/s10846-018-0818-4
42. Godio S, Primatesta S, Guglieri G, Dovis F. A Bioinspired Neural Network-Based Approach for Cooperative Coverage Planning of UAVs. *Information* (2021) 12:51. doi:10.3390/info12020051

43. Xie Y, Han L, Dong X, Li Q, Ren Z. Bio-Inspired Adaptive Formation Tracking Control for Swarm Systems With Application to UAV Swarm Systems. *Neurocomputing* (2021) 453:272–85. doi:10.1016/j.neucom.2021.05.015
44. Wang C, Wang D, Gu M, Huang H, Wang Z, Yuan Y, et al. Bioinspired Environment Exploration Algorithm in Swarm Based on Lévy Flight and Improved Artificial Potential Field. *Drones* (2022) 6:122. doi:10.3390/drones6050122
45. Xu B, Liu T, Bai G, Tao J, Zhang Y, Fang Y. A Multistate Network Approach for Reliability Evaluation of Unmanned Swarms by Considering Information Exchange Capacity. *Reliability Eng Syst Saf* (2022) 219:108221. doi:10.1016/j.ress.2021.108221
46. Wang K, Zhang X, Qiao X, Li X, Cheng W, Cong Y, et al. Adjustable Fully Adaptive Cross-Entropy Algorithms for Task Assignment of Multi-UAVs. *Drones* (2023) 7:204. doi:10.3390/drones7030204
47. Zeng Y, Ren K, Sun Q, Zhang Y. *A Resilient Task Allocation Method for UAV Swarm Based on PSO-ILP Bilevel Nested Optimization* (2023). Available from: https://doi.org/10.21203/rs.3.rs-2765091/v1.
48. Bejiga M, Zeggada A, Nouffidj A, Melgani F. A Convolutional Neural Network Approach for Assisting Avalanche Search and Rescue Operations With UAV Imagery. *Remote Sensing* (2017) 9:100. doi:10.3390/rs9020100
49. Phadke A, Medrano A. Drone2Drone: A Search and rescue Framework for Finding Lost UAV Swarm Agents. Poster. In: *Symposium for Student Innovation, Research, and Creative Activities 2023*. United States: Texas A&M University-Corpus Christi (2023).
50. Tello DJI. *EDU User Manual* (2018).
51. Wheeb AH, Nordin R, Samah AA, Kanellopoulos D. Performance Evaluation of Standard and Modified OLSR Protocols for Uncoordinated UAV Ad-Hoc Networks in Search and Rescue Environments. *Electronics* (2023) 12:1334. doi:10.3390/electronics12061334
52. Li B, Jiang Y, Sun J, Cai L, Wen CY. Development and Testing of a Two-UAV Communication Relay System. *Sensors (Basel)* (2016) 16:1696. doi:10.3390/s16101696
53. Phadke A, Medrano FA, Sekharan CN, Chu T. Designing UAV Swarm Experiments: A Simulator Selection and Experiment Design Process. *Sensors* (2023) 23:7359. doi:10.3390/s23177359
54. Pohudina O, Kovalevskyi M, Pyvovar M. Group Flight Automation Using Tello EDU Unmanned Aerial Vehicle. In: 2021 IEEE 16th International Conference on Computer Sciences and Information Technologies (CSIT); 22-25 September 2021; LVIV, Ukraine (2021). p. 151–4.
55. Virbora N, Sokoeun U, Saran M, Sovicheyratana S, Channareth S, Saravuth S. Implementation of Matrix Drone Show Using Automatic Path Generator With DJI Tello Drones. In: 2022 International Conference on Engineering and Emerging Technologies (ICEET); 27-28 October 2022; Kuala Lumpur, Malaysia (2022). p. 1–5.
56. Pikalov S, Azaria E, Sonnenberg S, Ben-Moshe B, Azaria A. Vision-Less Sensing for Autonomous Micro-Drones. *Sensors (Basel)* (2021) 21:5293. doi:10.3390/s21165293
57. Rohmer E, Singh SPN, Freese M. CoppeliaSim (Formerly V-REP): A Versatile and Scalable Robot Simulation Framework. In: IEEE/RSJ International Conference on Intelligent Robots and Systems; 03-07 November 2013; Tokyo, Japan (2013). p. 1321–6.
58. FAA. Remote Identification of Unmanned Aircraft-Final Rule. In: FAA D, editor (2023). Available at: https://www.faa.gov/sites/faa.gov/files/2021-08/RemoteID_Final_Rule.pdf.
59. Phadke A, Boyd J, Medrano FA, Starek M. Navigating the Skies: Examining the FAA's Remote Identification Rule for Unmanned Aircraft Systems. *Drone Syst Appl* (2023) 11:1–4. doi:10.1139/dsa-2023-0029
60. Opromolla R, Inchingolo G, Fasano G. Airborne Visual Detection and Tracking of Cooperative UAVs Exploiting Deep Learning. *Sensors (Basel)* (2019) 19:4332. doi:10.3390/s19194332
61. Razzaq S, Xydeas C, Mahmood A, Ahmed S, Ratyal NI, Iqbal J. Efficient Optimization Techniques for Resource Allocation in UAVs Mission Framework. *PLoS One* (2023) 18:e0283923. doi:10.1371/journal.pone.0283923
62. Nguyen M-N, Nguyen LD, Duong TQ, Tuan HD. Real-Time Optimal Resource Allocation for Embedded UAV Communication Systems. *IEEE Wireless Commun Lett* (2019) 8:225–8. doi:10.1109/lwc.2018.2867775
63. Yu Z, Gong Y, Gong S, Guo Y. Joint Task Offloading and Resource Allocation in UAV-Enabled Mobile Edge Computing. *IEEE Internet Things J* (2020) 7:3147–59. doi:10.1109/jiot.2020.2965898
64. Sanjay Sarma VO, Parasuraman R, Pidaparti R. Impact of Heterogeneity in Multi-Robot Systems on Collective Behaviors Studied Using a Search and Rescue Problem. In: IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR). Abu Dhabi, United Arab Emirates; 04-06 November 2020; Abu Dhabi, United Arab Emirates (2020).
65. Chatziparaschis D, Lagoudakis MG, Partsinevelos P. Aerial and Ground Robot Collaboration for Autonomous Mapping in Search and Rescue Missions. *Drones* (2020) 4:79. doi:10.3390/drones4040079
66. López DS, Moreno G, Cordero J, Sanchez J, Govindaraj S, Marques MM, et al. Interoperability in a Heterogeneous Team of Search and Rescue Robots. In: *From Theory to Practice: Search and Rescue Robotics* (2017).
67. Condomines J-P, Zhang R, Larrieu N. Network Intrusion Detection System for UAV Ad-Hoc Communication: From Methodology Design to Real Test Validation. *Ad Hoc Networks* (2019) 90:101759. doi:10.1016/j.adhoc.2018.09.004
68. Choudhary G, Sharma V, You I, Yim K, Chen I-R, Cho J-H, et al. Intrusion Detection Systems for Networked Unmanned Aerial Vehicles: A Survey; 25-29 June 2018; Limassol, Cyprus (2018).
69. Basan E, Lapina M, Mudruk N, Abramov E. Intelligent Intrusion Detection System for a Group of UAVs. In: *Advances in Swarm Intelligence*. Cham: Springer (2021). p. 230–40.
70. Subbarayalu V, Vensuslaus MA. An Intrusion Detection System for Drone Swarming Utilizing Timed Probabilistic Automata. *Drones* (2023) 7:248. doi:10.3390/drones7040248
71. Phadke A, Ustymenko S. Updating the Taxonomy of Intrusion Detection Systems. In: 2021 IEEE 45th Annual Computers, Software, and Applications Conference (COMPSAC); 12-16 July 2021; Madrid, Spain (2021). p. 1085–91.
72. Levitin G, Finkelstein M, Xiang Y. Optimal Abort Rules and Subtask Distribution in Missions Performed by Multiple Independent Heterogeneous Units. *Reliability Eng Syst Saf* (2020) 199:106920. doi:10.1016/j.ress.2020.106920
73. Levitin G, Finkelstein M, Xiang Y. Optimal Abort Rules for Additive Multi-Attempt Missions. *Reliability Eng Syst Saf* (2021) 205:107245. doi:10.1016/j.ress.2020.107245
74. Zhang Q, Fang Z, Cai J. Preventive Replacement Policies With Multiple Missions and Maintenance Triggering Approaches. *Reliability Eng Syst Saf* (2021) 213:107691. doi:10.1016/j.ress.2021.107691
75. Zhao X, Dai Y, Qiu Q, Wu Y. Joint Optimization of Mission Aborts and Allocation of Standby Components Considering Mission Loss. *Reliability Eng Syst Saf* (2022) 225:108612. doi:10.1016/j.ress.2022.108612
76. Zhu X, Zhu X, Yan R, Peng R. Optimal Routing, Aborting and Hitting Strategies of UAVs Executing Hitting the Targets Considering the Defense Range of Targets. *Reliability Eng Syst Saf* (2021) 215:107811. doi:10.1016/j.ress.2021.107811
77. Guo J, Wang L, Wang X. A Group Maintenance Method of Drone Swarm Considering System Mission Reliability. *Drones* (2022) 6:269. doi:10.3390/drones6100269
78. Peng R. Joint Routing and Aborting Optimization of Cooperative Unmanned Aerial Vehicles. *Reliability Eng Syst Saf* (2018) 177:131–7. doi:10.1016/j.ress.2018.05.004
79. Liu L, Yang J. A Dynamic Mission Abort Policy for the Swarm Executing Missions and Its Solution Method by Tailored Deep Reinforcement Learning. *Reliability Eng Syst Saf* (2023) 234:109149. doi:10.1016/j.ress.2023.109149