Extreme Mechanics Letters 36 (2020) 100651

Contents lists available at ScienceDirect

Extreme Mechanics Letters

journal homepage: www.elsevier.com/locate/eml

Interactive inverse design of layered phononic crystals based on reinforcement learning

Chengcheng Luo, Shaowu Ning, Zhanli Liu *, Zhuo Zhuang

Applied Mechanics Lab., Department of Engineering Mechanics, School of Aerospace, Tsinghua University, Beijing 100084, China

ARTICLE INFO

Article history:
Received 15 January 2020
Received in revised form 14 February 2020
Accepted 16 February 2020
Available online 24 February 2020

Keywords:
Reinforcement learning
Phononic crystal
Bandgap
Elastic wave
Inverse design

ABSTRACT

As supervised learning has been successfully applied in mechanics, reinforcement learning is being attempted to be used to solve mechanical problems more intelligently. In this study, by imagining the mechanical design as a "game" to make clear what is the "score" to maximize, reinforcement learning is successfully applied to the design of layered phononic crystals with anticipated band structures, which can regulate elastic waves by blocking the waves in the range of bandgap. In order to get the desired bandgaps, it is necessary to design unique topological structure of phononic crystals. In this work, the topological structure of layered phononic crystals can evolve itself through interactive reinforcement learning algorithm, and finally reaches the topological structure which meets the given requirements. The reinforcement learning method performs very well both under the goal of maximizing the first-order bandgap width and designing the bandgap of the specified range, respectively. It is worth mentioning that the method is efficient and stable, that is independent of the initial state and target, and can finally learn an evolution route that will keep the objective function increasing. Inspired by the results of exploration, the theoretical analysis is also carried out to explain the design results and gives the feasible bandgap range in layered phononic crystals with given material properties. This reinforcement learning based interactive design scheme can be easily extended to other inverse design problems.

1. Introduction

Phononic crystal is a kind of artificial acoustic material that can control elastic waves. It has a vast application prospect in acoustic focusing, acoustic guidance, vibration isolation, noise reduction, acoustic stealth, and stress wave protection [1–8]. Inspired by photonic crystal [9], Kushwaha et al. proposed this kind of composite with periodic structure [10] to regulate the acoustic wave. The researches show that the periodic structure can form a discrete band structure. The bandgap is the range of frequencies in the band structure which there are no wave vectors corresponding to. If the material has infinite periods of unit cells, the harmonic elastic wave with the frequency in the bandgap cannot pass through the material at all. Although only phononic crystals with finite size can be produced in reality, the elastic waves in bandgap will have a very low transmission coefficient [11], thus still having a blocking effect on such waves. As a complement to traditional phononic crystals, Liu et al. put forward the concept of "resonant phononic crystals" [12], which absorb wave energy through local resonance and are usually composed of three phases of materials [3,5,6,13]. There are many topological structures of phononic crystals, as shown in Fig. 1. Different topological structures of the unit cell of phononic crystal correspond to different band structures. To control the waves of different ranges, researchers need to design different topological structures. However, there is no explicit relationship between band structures and topological structures. Most of the time, researchers get the band structure of various topological structures one by one through numerical calculations using FEM, PWE, FTDT, MST, etc. [14–20].

To adjust and control the acoustic waves in different frequency ranges, a variety of topological structures are designed, and the bandgap range can be adjusted by changing the geometry parameters in the structures [13,21–25]. In order to make tunable bandgaps, many researchers have studied the phononic crystals whose band structure can be controlled by mechanical and electromagnetic means [11,26–30]. Besides, some researchers have designed the phononic crystals using topology optimization to achieve the maximum bandgap width, the maximum relative bandgap width, or desired frequency spectra [31–34]. Although a lot of work has been done in the research of phononic crystals, the reverse design of phononic crystals is still an unsolved problem.

Recently, the introduction of machine learning methods into mechanics provides new ways to solve mechanical problems [35–43]. Machine learning methods can be divided into supervised

* Corresponding author.
E-mail address: liuzhanli@tsinghua.edu.cn (Z. Liu).

https://doi.org/10.1016/j.eml.2020.100651
2352-4316/

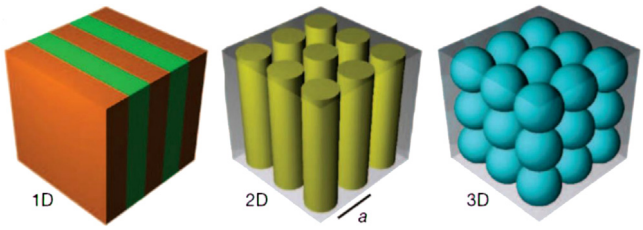

Fig. 1. Phononic crystals of different topological structures [44].

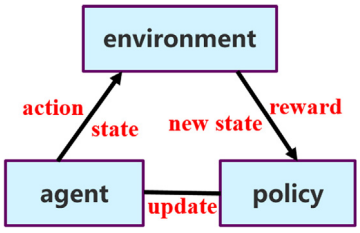

Fig. 2. The framework of reinforcement learning.

learning, unsupervised learning, and reinforcement learning. At present, the researches on combining machine learning and mechanics are mostly based on supervised learning, aiming to establish a set of mapping relationship between mechanical problems and parameters by neural network or other machine learning methods like supporting vector machine (SVM). However, these approaches always need a large number of data samples. After getting enough data samples, they are submitted to the neural network to train and use the trained neural network to predict new data samples. These methods separate the process of getting data and the process of machine learning, which are called "offline" machine learning methods. In fact, the machine learning process can be naturally combined with the process of getting data, which will reduce the number of needed data samples and make the process of getting data more intelligent. In recent years, the applications of reinforcement learning, such as alpha go, alpha zero, AI gaming, and robots, have attracted a lot of attention. As supervised learning has been successfully applied in mechanics [35,37–39,44–47], reinforcement learning can also be expected to solve mechanical problems. As a method that can find the optimal solution gradually by interacting with the environment, reinforcement learning has the potential to provide new solutions to mechanical optimization problems. In this paper, we made a preliminary exploration in this area, and take the inverse design of the band structure of layered phononic crystals as an example to show the application prospect of reinforcement learning method in mechanical problems.

The model in reinforcement learning includes three parts: environment, agent, and policy. The agent takes action to get feedback from the environment and use the feedback to update the policy, and the policy guides the agent to take the next action. The description of the agent's current situation is called the "state". The reinforcement learning aims to make the agent learn the policy from the exploration process, to maximize the rewards from the environment [36]. The process of interaction between agent and environment is shown in Fig. 2.

There are several reinforcement learning algorithms, such as Q-learning, A3C, PPO, TRPO, etc. [48–53], which are all based on the frame shown in Fig. 2, but different in the ways to update policy. As a representative algorithm in reinforcement learning, the Q-learning algorithm has made great achievements in the field of games, which made better achievements than humans in about 2600 Atari games [36]. Q-learning algorithm lets the agent play the game and take the score of each game as the feedback of the environment. The agent does not understand the game and may perform randomly and get low scores at first, but will gradually learn to choose a better action under each state and get a better final score. Making the score higher in the game can be analogized to the material design in mechanics, where we will want to get a material more in line with our wishes. In order to apply the Q-learning algorithm to mechanical problems, we need to imagine the mechanical design as a game, to make clear what is the "score" to maximize, and then to define the "state", "action" and "reward" in the game. The procedure of Q-learning algorithm is as follows:

Initialize $Q(s, a)$ *(we make all the values zero at first):*
Repeat (for each episode):
Initialize state s
Repeat (for each step):
Choose action a from state s using policy derived from Q-function(ϵ*-greedy*):
with the probability of ϵ:
Choose action a with the largest Q(s,a) (if several actions make the largest value, randomly choose one of them)
with the probability of 1-ϵ:
Choose action a randomly
Take action a, go to state s', get reward r
update Q-function:
$Q(s, a) = Q(s, a) + \alpha[r + \gamma \max_{a'} Q(s', a') - Q(s, a)]$
$s = s'$
Until the episode ends (e.g. reach the maximum steps).

Here $Q(s, a)$ is a function using state and action as variables. When the amount of data is small, Q function is recorded in the form of tables. When the amount of data is too large to be recorded in tables, usually millions of pieces of data [49], then it is derived into the DQN algorithm, where the neural network is used to store the values of the Q-function. As we do not face such a large number of data, we do not employ a neural network in this research. Table 1 shows the form of Q-function, and updating the Q-function is just updating the values in the table. The values of the Q-function are zero at the beginning, and in the following exploration process, the function is updated through reward, and the next action will be determined according to the current Q-function. Actually, the Q-function is just like the experience of the agent. On the one hand, the agent updates the experience through the interaction with the environment, on the other hand, it also determines what to do through the existing experience, although it is not brilliant at first. The algorithm to decide which action to choose is called $\varepsilon - greedy$ strategy. The agent will check the Q-function to see which action gets the largest value at the current state. The agent will select that action with the probability of ε, and choose the action randomly with the probability of $1 - \varepsilon.\varepsilon$ is between 0 and 1, and reflects the balance between exploration and optimization in reinforcement learning. The bigger the ε is, the faster you can find a satisfactory solution. However, it is not guaranteed to be the best solution globally. A smaller ε can better avoid the agent trapping into the local optimal solution, but will also take more time.is called the "discount coefficient". It is also between 0 and 1, and it is inspired by the human's psychology that one will think higher of the short-term reward than the long-term reward when the reward is the same. After an enough long exploration, the value of Q-function will converge to the expected final score of taking the action from the state.

By properly defining the state, action, and reward, the Q-learning algorithm can be used on a wide range of mechanical problems. In this paper, we employ layered phononic crystals design as an example, taking maximizing the first-order bandgap width and realizing any specified bandgap range as two tasks, to show the feasibility of reinforcement learning on mechanical inverse problems. This paper is organized in the following scheme.

Table 1
The representation of Q-function.

States	Value for action a_1	Value for action a_2
s_1	$Q(s_1, a_1)$	$Q(s_1, a_2)$
s_2	$Q(s_2, a_1)$	$Q(s_2, a_2)$
.....	.....	.....

Table 2
Physical parameters of materials.

Material	Shear modulus (GPa)	Density (kg/m^3)
Aluminum	28.7	2730
Epoxy	1.59	1180

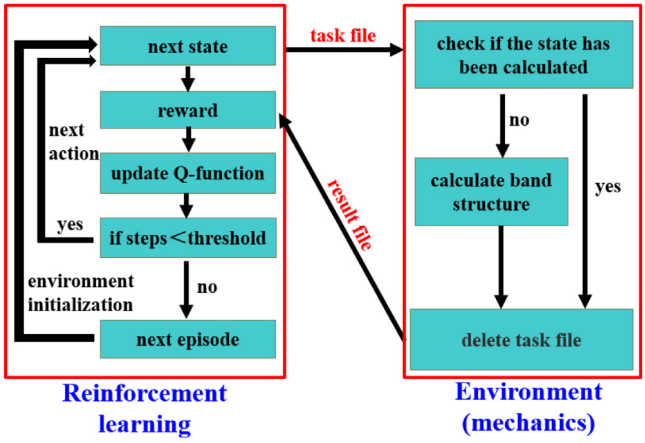

Fig. 3. The framework of reinforcement learning for layered phononic crystals.

Section 2 introduces the framework of using reinforcement learning in mechanics design problems and the way in our specific problems. Then the results of using reinforcement learning to maximize the first-order bandgap width and to realize specified bandgap range are given in Section 3. Section 4 discusses the characteristics of the evolution route, and the achievable bandgap range is analyzed when the materials are fixed. At last, the summary is concluded in Section 5.

2. Method and the framework of reverse design

2.1. Framework of using reinforcement learning in mechanical design

In this paper, we choose the bandgap of layered phononic crystal as the design object, whose band structure can be calculated by the transfer matrix method. We take epoxy and aluminum as materials, which are commonly used for phononic crystals [54]. Their physical parameters are shown in Table 2. By changing the thickness of the unit cell of two layers, different band structure can be obtained. Although we take the shear wave as an example, the method is valid for the longitudinal wave by replacing shear modulus with young's modulus or bulk modulus.

The framework of this design problem is shown in Fig. 3, where the reinforcement learning part is separated from the mechanical calculation part, which can be easily used for other problems. The environment part can be a commercial software like Comsol, Abaqus, but in this work, the bandgap is calculated using our own program based on the theory described in Section 2.2.

In a periodic unit cell, different combinations of thicknesses will bring different band structures and different bandgap width. The thicknesses of two layers in a unit cell are denoted by a_1 and a_2, respectively. We limit (a_1, a_2) in the range of:

$a_1 \in [0.02, 0.12]$ m
$a_2 \in [0.02, 0.12]$ m

The initial thickness is $a_1 = a_2 = 0.07$ m, and the increment for a step is:

$\Delta a_1 = 0.002$ m, $\Delta a_2 = 0.002$ m

In the reinforcement learning part, the agent can choose either a_1 or a_2 to increase or decrease with the fixed increment. This increment size is small enough so that only causes a slight change to the bandgap. In the given design area, there are $51 \times 51 = 2601$ kinds of combinations of geometry. In this paper, we successfully use the Q-learning method to achieve the desired bandgap quickly in the given design area.

The implementation process of the Q-learning algorithm is as follows. We define the "state" as (a_1, a_2), like a coordinate in the Cartesian coordinate system. The actions can be "a_1+", "a_1-", "a_2+" or "a_2-". They represent which thickness to change, and to increase or decrease. If the state exceeds the range of the given design area after an action, the new state is considered to be the same as the previous state. The "reward" is set as the increment of the objective function $\pi(a_1, a_2)$ from the old state to the new state, where $\pi(a_1, a_2)$ is determined by design objectives. The objective function must satisfy such a property: the larger the objective function is, the better the goal is achieved. The objective function is the score in this "game", and the reinforcement learning algorithm will always learn to make the value of objective function larger. We let the agent explore 70 steps in each episode, and take the result of the final state as the result of this episode. After an episode finished, the agent goes back to the initial state (0.07, 0.07) to begin the new episode. The experience of the previous episode has been recorded and the policy has been updated. The agent will perform better and better in the process of exploration.

In our work, we set maximizing the first-order bandgap and designing the bandgap of the specified range as the goals, respectively. To achieve these different goals, the only thing that needs to do is to change the objective function $\pi(a_1, a_2)$.

2.2. Calculation of band structure of layered phononic crystals

The sketch map of a layered phononic crystal is shown in Fig. 4. Each layer is infinitely long in the z-direction and y-direction, and each layer is a homogeneous material. We use μ_n, ρ_n, a_n to represents the shear modulus, density, and thickness of the material of the nth layer, respectively. Assuming a harmonic stress wave with an angular frequency of incidents at an angle of , we consider the steady-state solution of this problem by using the transfer matrix method [54].

From Snell's theorem in sound wave, the k_z, which is the wave vector along the z-direction, is constant in all layers. k_n represents the wave vector of x-direction in the nth layer, which can be calculated as $k_n = \sqrt{(\frac{\omega}{c_n})^2 - k_z^2}$. Here c_n represents the shear wave velocity in the nth layer, which can be calculated as $c_n = \sqrt{\frac{\mu_n}{\rho_n}}$. In each layer, the stress and displacement field in the k_n-space can be expressed as:

$$\sigma_n(x) = \sigma_n^+ e^{i(\omega t - k_n x)} + \sigma_n^- e^{i(\omega t + k_n x)} \tag{1}$$

$$u_n(x) = \frac{\sigma_n^+}{\rho_n c_n^2 k_n} e^{i(\omega t - k_n x)} - \frac{\sigma_n^-}{\rho_n c_n^2 k_n} e^{i(\omega t + k_n x)} \tag{2}$$

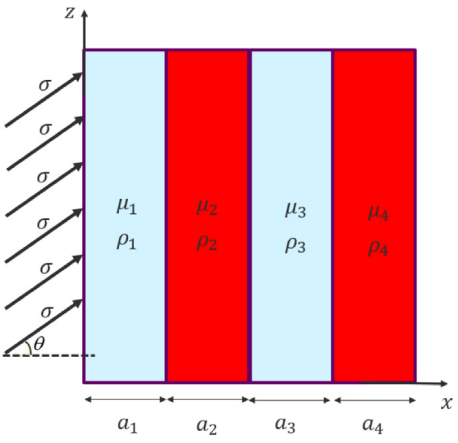

Fig. 4. Sketch map of layered phononic crystal.

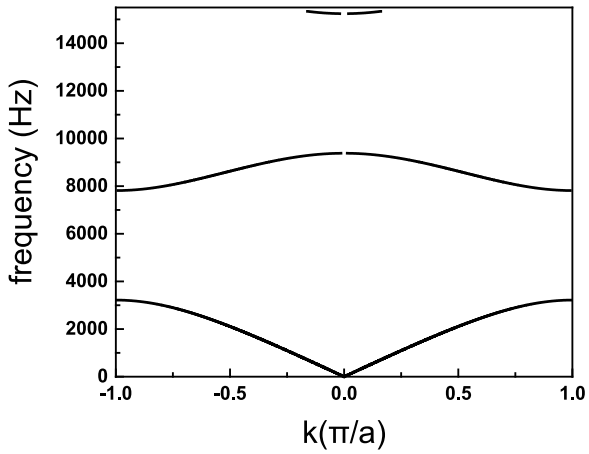

Fig. 5. The band structure when the thickness of two layers are both 0.07 m.

Here $\sigma_n^+, \sigma_n^-, u_n^+, u_n^-$ represent the amplitude of stress and displacement of waves forward and backward along y-direction, respectively.

Using the continuous condition of the displacement and stress on the boundary between the nth layer and the (n+1)-th layer, we have the equations:

$$\sigma_n(a_n) = \sigma_{n+1}(0) \tag{3}$$

$$u_n(a_n) = u_{n+1}(0) \tag{4}$$

Take Eq. (1)(2) into (3)(4), we have

$$\begin{bmatrix} \sigma_{n+1} \\ \sigma'_{n+1} \end{bmatrix} = T_n \begin{bmatrix} \sigma_n \\ \sigma'_n \end{bmatrix} \tag{5}$$

Here we call T_n as the transfer matrix from the nth layer to the (n+1)th layer, which can be expressed as:

$$T_n = \begin{bmatrix} 1 & 1 \\ \dfrac{\rho_n c_n^2 k_n}{\rho_{n+1} c_{n+1}^2 k_{n+1}} & \dfrac{\rho_n c_n^2 k_n}{\rho_{n+1} c_{n+1}^2 k_{n+1}} \end{bmatrix}^{-1} \begin{bmatrix} e^{i(-k_n a_n)} & e^{i(k_n a_n)} \\ e^{i(-k_n a_n)} & -e^{i(k_n a_n)} \end{bmatrix} \tag{6}$$

According to Eq. (5), we have the relationship between the (n+1)th layer and the first layer:

$$\begin{bmatrix} \sigma_{n+1} \\ \sigma'_{n+1} \end{bmatrix} = T \begin{bmatrix} \sigma_1 \\ \sigma'_1 \end{bmatrix} \tag{7}$$

$$T = T_n T_{n-1} \dots T_1$$

For a layered phononic crystal with n layers as a unit cell, the Bloch–Floquet periodic condition [55] can be expressed as:

$$\begin{bmatrix} \sigma_{n+1} \\ \sigma'_{n+1} \end{bmatrix} = e^{-ika} \begin{bmatrix} \sigma_1 \\ \sigma'_1 \end{bmatrix} \tag{8}$$

For a series of, calculate the wave vector k to make

$$\left| T - e^{-ika} I \right| = 0 \tag{9}$$

and the band structure is get. Those which have no real solution for k is in the bandgap. In the problem we consider, there are two layers in a unit cell, and the materials are aluminum and epoxy. When the thicknesses of the two layers are both 0.07 m, $\theta = 0°$, we calculate the band structure using the above method, and get the band structure as Fig. 5.

As can be seen from Fig. 5, the first-order bandgap range is from 3.2 KHz to 7.8 KHz. To verify this result, we create the model as shown in Fig. 6(a), which consists of three unit cells, where $a_1 = a_2 = 0.07$ m. By calculating the transfer coefficient of the composite, which is defined as the ratio of the amplitudes of stresses in the output section and in the input section, the blocking effect of the phononic crystal on waves can be quantified. The graph of transfer coefficient versus frequency is shown in Fig. 6(b). Compare Fig. 6(b) and Fig. 5, the first region in Fig. 6(b) with very low transmission coefficient exactly corresponds to the first-order bandgap in Fig. 5, and the second-order bandgap corresponds to the second region with low transmission coefficient in Fig. 6(b), which proves the correctness of our calculation method and show the validity of bandgap. It shows that with only three unit cells, the phononic crystal can significantly block the wave in the bandgap.

3. Implementation and inverse design results

In this section the interactive reinforcement learning algorithm is implemented to maximize the first-order bandgap width and to realize the specified bandgap range.

3.1. The maximization for first-order bandgap width

In many cases of applications, it is desired that the phononic crystal could block the waves in a wide range of frequency. Therefore, the first goal for reinforcement learning is to maximize the first-order bandgap width within a given range of thickness. We take the objective function $\pi(a_1, a_2)$ as the first-order bandgap.

At the initial state (0.07, 0.07), the first-order bandgap width is 4597 Hz. Through the Q-learning algorithm, as shown in Fig. 7(a), after several episodes, the final first-order bandgap width of each episode has converged to 16988 Hz, where $a_1 = 0.032$ m and $a_2 = 0.02$ m. As will be proved in Section 4.1, it is the maximum bandgap.

In Fig. 7(a) we can see, after about 40 episodes of exploration, the result converges to the optimal result. When the episodes have been convergent, the first-order bandgap width of steps in the last episode is shown in Fig. 7(b), from which we can see the bandgap width increases continually with steps.

Next we make more explorations to see if the result is stable. We make six tests, in which the agent should clear the experience and start from the same initial state. In all of the tests, the agent finally finds the same final state, as shown in Fig. 8. In addition, it can be found that the routes of evolution for the six tests are very similar. They are not the straight line from the start state to the final state, but all goes to about (0.08,0.05) first, then turn a corner. The principle of the route will be discussed in Section 4.1.

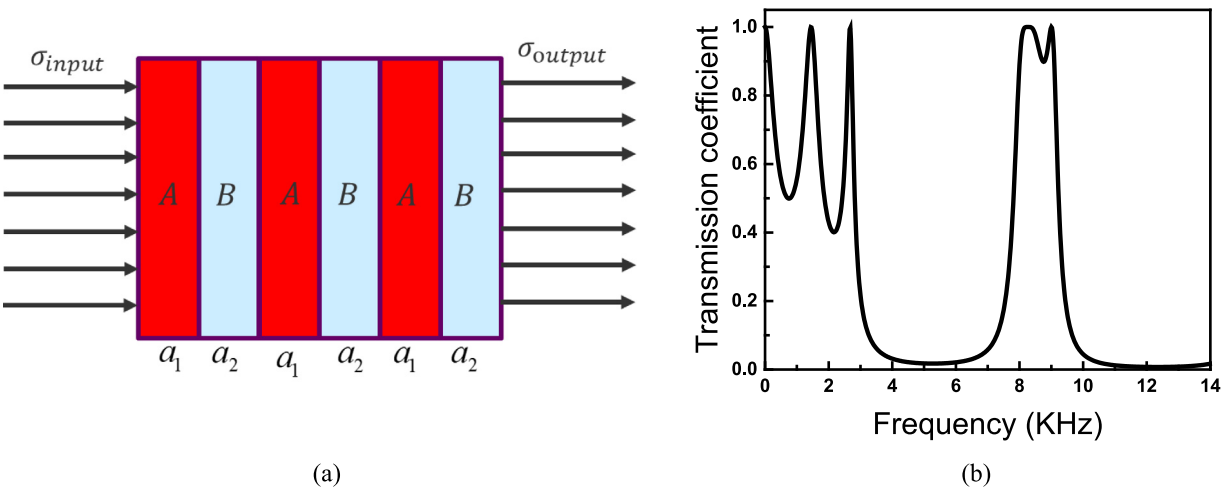

Fig. 6. (a) The phononic crystal with three unit cells; (b)The transfer coefficient versus frequency.

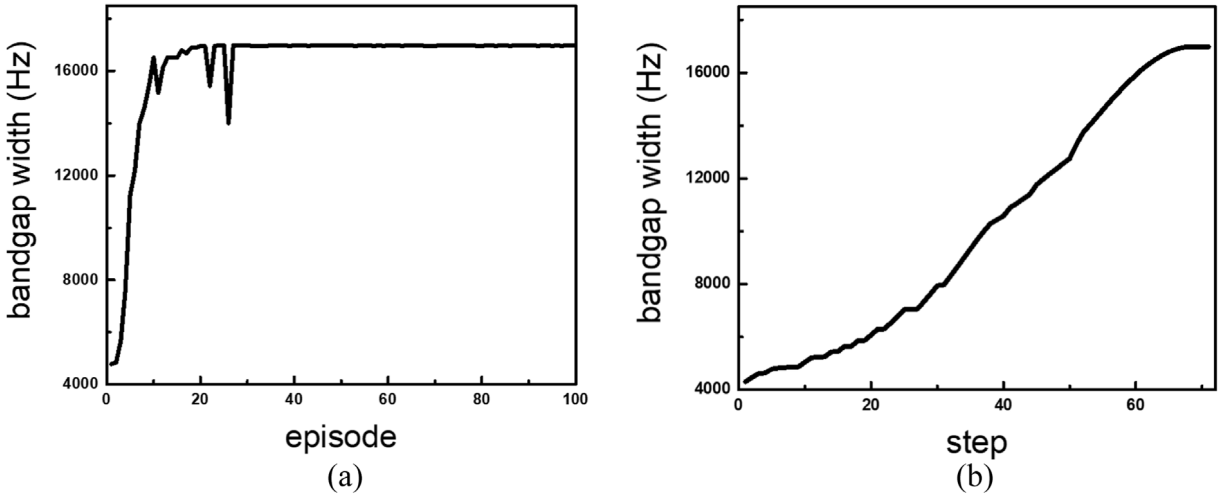

Fig. 7. (a) The final first-order bandgap width in each episode; (b)The bandgap width in every step at the final episode.

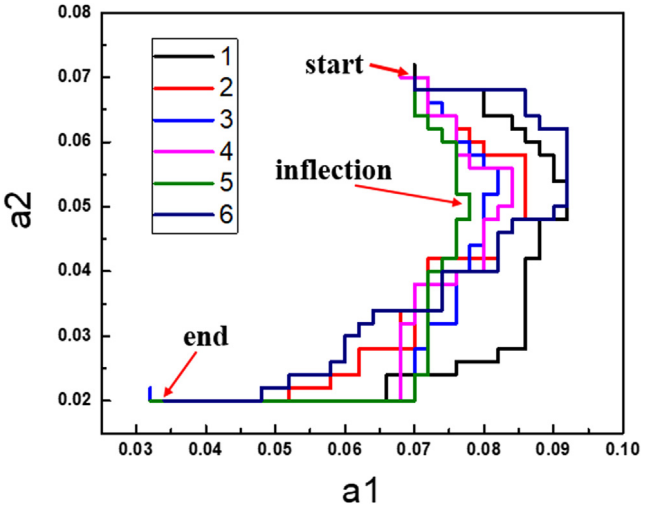

Fig. 8. Route of evolution at final episode for 6 tests.

In order to study whether the result is sensitive to the initial state, several tests from different initial states are made. In this study, we try four different initial points to see if the result is stable: (0.07,0.07), (0.12,0.12), (0.032,0.12) and (0.02,0.07). Though these four initial states vary substantially, all of these cases reached the same final state (0.032,0.02) finally. The evolution routes of different initial states are shown in Fig. 9(a). They all reach the same final state.

In order to quantify the efficiency of the algorithm on the problem, we define the parameter "states ratio" as:

$$states\ ratio = \frac{explored\ states}{all\ states}$$

The "explored states" includes all the arrived states from the beginning to the end. The lower the states ratio is, the less time it takes to converge to the final state. The states ratios of different initial states are shown in Fig. 9(b). It can be seen from the figure that although the initial state and the final state are far away, the final state can still be found in a very short relative time, which shows that reinforcement learning is feasible and efficient in this mechanical design problem. In addition, unlike many optimization algorithms [56–60], the reinforcement learning not only concerns about the final result but also concerns about the route of evolution. It learns the mapping from states to actions, which leads the route of revolution having principles behind. This will be discussed in Section 4.1.

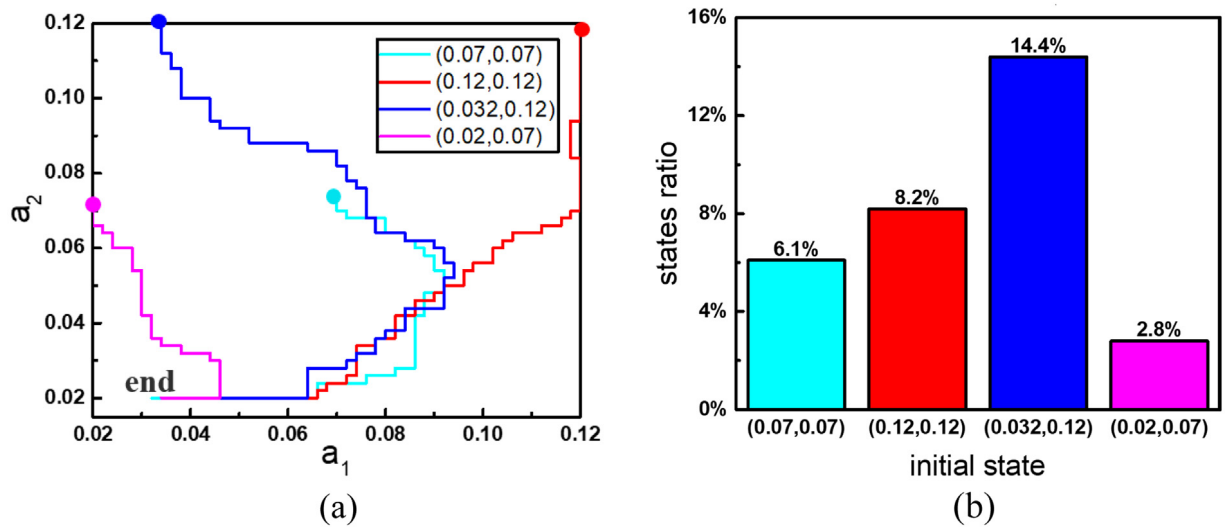

Fig. 9. (a) Routes of evolution of different initial states; (b) States ratio of different initial states.

3.2. Customized design of bandgap range

In many scenarios, the requirement of the phononic crystal is to block the waves in the specified frequency range. It requires designing a phononic crystal with a specified bandgap. We just modify the objective function as follows and achieve the goal. We record the bandgap range from x Hz to y Hz as [x, y]. If the specified bandgap range is [a, b], and the current first-order bandgap is $[x, y]$, we take the value of objective function as $-[(x-a)^2+(y-b)^2]$. The agent will make $[(x-a)^2+(y-b)^2]$ as small as possible, and the bandgap range will be close to the target bandgap.

We use [10000,20000], [6000,8000], [3000,9000], [6000, 20000], [4000,12000] as our specified bandgap ranges, respectively. These bandgap ranges are set so differently to see if reinforcement learning can achieve all kinds of goals.

As a result, they all reach the target bandgap range at a satisfactory accuracy, although some targets are limited by the design area. The evolution route and the bandgap of the final states are shown in Fig. 10(a). The goal of [10000,20000] achieves at [9291,19900], [6000,8000] achieves at [5328,8398], [3000,9000] achieves at [2817,8747], [6000,20000] achieves at [6522,19542]. Their final bandgap range is very close to the target bandgap. The first three goals finally stop at the boundary of the design area and may get better performance if the design area is larger. The state ratios of different optimization objectives are calculated and shown in Fig. 10(b). It can be seen that on this task, the Q-learning method still has very high exploration efficiency, even faster than the task in Section 3.1.

Through above research, it shows that the Q-learning method performs well in realizing the specified range of bandgap. In addition, the research shows that layered phononic crystals can achieve rich bandgap ranges. However, the goal of [6000,20000] is not achieved very accurately. In Section 4.2 we will prove that the goal is not achievable and discuss the achievable bandgap when the materials are given.

4. Discussion

From the above two tasks, it shows that Q-learning algorithm is applicable in finding the inverse design of mechanical problems, and has advantages in efficiency. For different goals, the only thing to do is to change the objective function, which greatly increases the adaptability of this method. On the basis of the above study, we make a further analysis about the characteristics of revolution route and the achievable bandgap range.

4.1. The characteristics of revolution route

The reinforcement learning learns the mapping from states to actions, so its choice in each state is of special meaning. The consistency of the routes in Fig. 8 proves that the evolution routes are not random. In order to study the characteristics of revolution routes in Fig. 8 and in Fig. 9(a), we need to know the bandgap width of every state. Through the derivation of band structure of the layered phononic crystal, it can be proved that the bandgap width $g(a_1, a_2)$ satisfies:

$$g(ka_1, ka_2) = \frac{1}{k} g(a_1, a_2) \quad \text{for all} \quad k > 0 \tag{10}$$

So by fixing a_2, study how the value of $g(a_1, a_2)$ changes with a_1, the value of $g(a_1, a_2)$ at every state can be deduced. We fix $a_2 = 0.02$, the curve of bandgap width with a_1 is shown in Fig. 11(a). Thus, the first-order bandgap width in the whole area can be calculated by applying Eq. (10) and is shown in Fig. 11(b). It can be seen that $a_1 = 0.032$, $a_2 = 0.02$ is the state with the largest first-order bandgap width, which suggests that the final result of reinforcement learning in this problem is the global optimal solution. In Fig. 7(b) we can see, the bandgap width is continuously increasing along the evolution route. And from Fig. 11(b) it can be inferred that when there are several actions that can increase the objective function, the agent will choose one of them with equal probability. In the path from (0.07,0.07) to about (0.08,0.05), increase a_1 or decrease a_2 can both increase the bandgap width, and from (0.08,0.05) to the final state, decrease a_1 or decrease a_2 can both increase the bandgap width. When the agent reaches the boundary, only one action could increase the bandgap. The choices with probability make the routes in Fig. 8 different but similar. Therefore, the evolution route learned from the Q-learning method, can be used in the scenarios where the continuous increase of the objective function is required.

4.2. The achievable bandgap range

From the research in Section 3.2, we get abundant bandgaps through reinforcement learning in the design area. And the result of explorations inspires us to study which bandgap ranges are achievable when the materials are given and the design area is not limited. Apply Eq. (9) to the unit cell of two layers, it can be

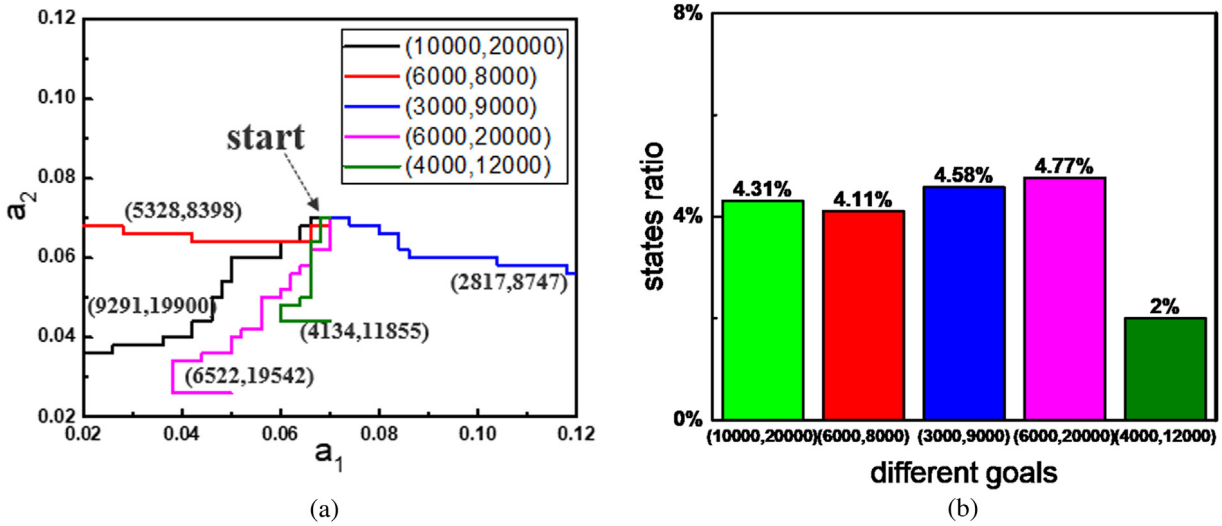

Fig. 10. (a) Evolution routes and final results of different goals; (b) States ratios of different goals.

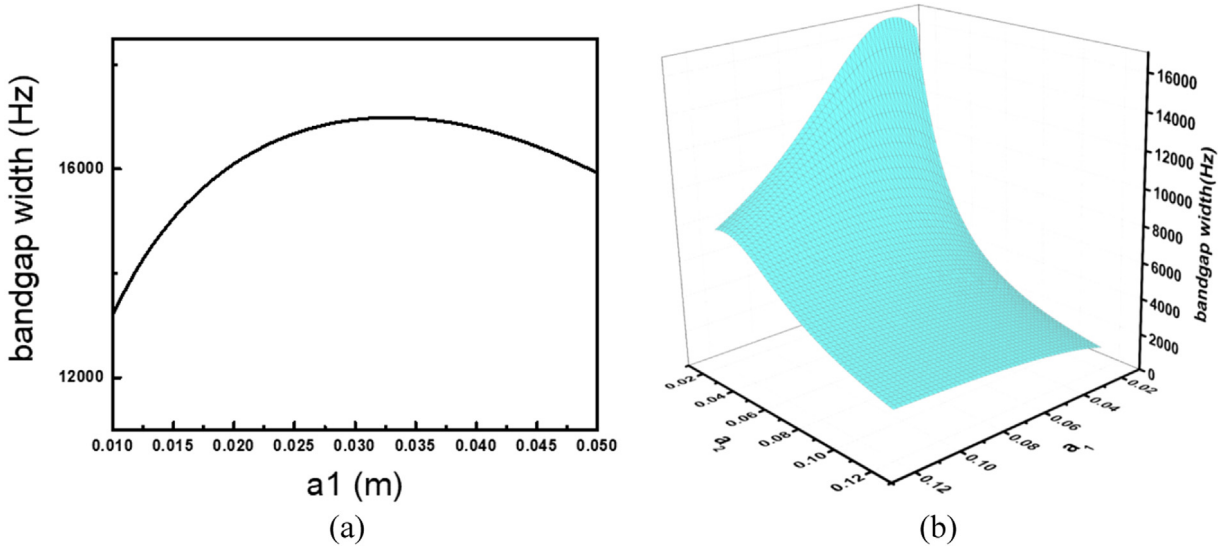

Fig. 11. (a) Bandgap width versus a_1 when $a_2 = 0.02$; (b) Bandgap width in the selected area.

deduced that solving the bandgap structure is just to solve this equation:

$$\cos ka = \cos(\frac{\omega a_1}{c_1})\cos(\frac{\omega a_2}{c_2}) - \frac{1}{2}(F + \frac{1}{F})\sin(\frac{\omega a_1}{c_1})\sin(\frac{\omega a_2}{c_2}) \quad (11)$$

where $F = \frac{\rho_1 c_1}{\rho_2 c_2}$ is the impedance ratio of the two materials. If an makes the right side of Eq. (11) between -1 and 1, then the is in pass band, otherwise it is in the bandgap. By defining the first-order relative bandgap as $2 \cdot \frac{\omega_{max} - \omega_{min}}{\omega_{max} + \omega_{min}}$, where , $\omega_{\min}$ are the upper boundary and the lower boundary of the first-order bandgap, respectively, it can be deduced from Eq. (10) that the relative bandgap is only related to $\frac{a_1 c_2}{c_1 a_2}$ and F, that is to say:

$$2 \cdot \frac{\omega_{max} - \omega_{min}}{\omega_{max} + \omega_{min}} = f(\frac{a_1 c_2}{c_1 a_2}, F) \quad (12)$$

The first variable in function f depends on the material and the geometry, and the second variable only depends on the material. Next we will give a theorem and prove it.

Theorem. *Whether the bandgap range [x,y] is achievable only depends on if its relative bandgap width,* $2 \cdot \frac{x-y}{x+y}$ *is achievable.*

Proof. If its relative bandgap width is achievable, it means that there exists an achievable bandgap range [x', y'], whose relative bandgap is the same as [x,y]. We assume the state of [x', y'] is $[a_1, a_2]$. Eq. (10) shows by dividing a_1, a_2 by k, we can obtain the bandgap range [kx',ky']. Find k to make ky'=y, then the following equation holds:

$$\frac{ky' - kx'}{ky' + kx'} = \frac{y - kx'}{y + kx'} = \frac{y - x}{y + x} \quad (13)$$

From Eq. (13) we can get

$$kx' = x \quad (14)$$

So the bandgap range [x,y] is achieved at the state of $[ka_1, ka_2]$. The proof finishes.

From the theorem, we only need to find the range of function in Eq. (12). In addition, it is obvious that $f(x, y) = f(\frac{1}{x}, y) = f(x, \frac{1}{y})$. So we only need to calculate the range of $f(x, y)|0 \le x \le 1, y \ge 1$.

By calculation we have the graph of relative bandgap width as shown in Fig. 12. It shows that the function f increases as x increases, and the larger the F(F≥1) is, the larger the range of

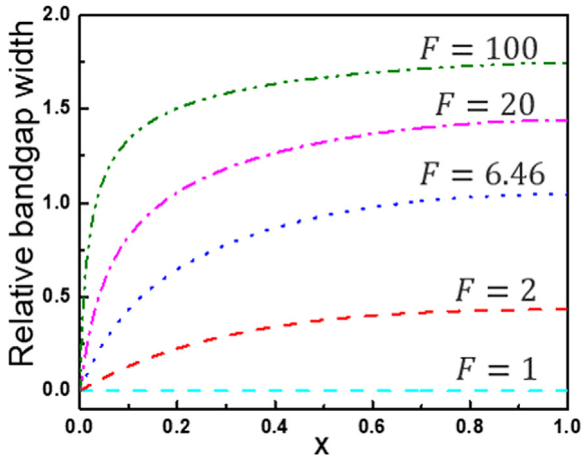

Fig. 12. Graph of relative first-order bandgap width at several impedance ratios.

bandgap is. When F=1, there is no bandgap at all. In the materials we choose, the F is 6.46. So it is proved that the maximum relative bandgap width is 1.05.

Among the bandgaps we specified in Section 3.2, the relative bandgap width of [6000,20000] is 1.08, which is out of the range of 1.05, while the other four goals are all in the range. It means the bandgap range [6000,20000] cannot be exactly realized. That is why the final bandgap range of this goal has still a little distance from the specified bandgap range, though it is still acceptable.

5. Conclusion

In this paper, the Q-learning algorithm, which is one of the reinforcement learning methods, is successfully applied in the inverse design of the band structure of layered phononic crystals. Taking maximizing the first-order bandgap width and achieving the specified bandgap range as examples, the application prospect of reinforcement learning in mechanical design problems is demonstrated. Conclusion remarks are drawn as follows:

(1) As can be seen from the flow of the Q-learning algorithm, the interactive inverse design method can be applied to various mechanical design problems by defining different states, actions, and rewards. For the problem of the same model, it is easy to achieve different goals by changing the objective function.
(2) The method has a great advantage in efficiency. Different from general optimization methods, the evolution route will keep the objective function increasing, which can be useful in some special scenarios. When there are several actions to increase objective function, the agent will randomly choose one of them.
(3) The achievable bandgap range of the layered phononic crystals with two materials has been discussed. It can be a reference for choosing the size and the materials to design phononic crystals.

Declaration of competing interest

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

Acknowledgments

This work is supported by the Science Challenge Project, China, No. TZ2018002, TZ2018001, National Natural Science Foundation of China, under Grant No. 11972205, 11722218 and 11921002, the National Key Research and Development Program of China (No. 2017YFB0702003), Opening Project of Applied Mechanics and Structure Safety Key Laboratory of Sichuan Province, China.

References

[1] S. Benchabane, A. Khelif, J.Y. Rauch, et al., Evidence for complete surface wave band gap in a piezoelectric phononic crystal, Phys. Rev. E 73 (6) (2006) 065601.
[2] X. Hu, Y. Shen, X. Liu, et al., Superlensing effect in liquid surface waves, Phys. Rev. E 69 (3) (2004) 030201.
[3] Acoustic Metamaterials and Phononic Crystals, Springer Science & Business Media, 2013.
[4] F. Cervera, L. Sanchis, J.V. Sánchez-Pérez, et al., Refractive acoustic devices for airborne sound, Phys. Rev. Lett. 88 (2) (2001) 023902.
[5] S. Nemat-Nasser, Refraction characteristics of phononic crystals, Acta Mech. Sinica 31 (4) (2015) 481–493.
[6] S.J. Mitchell, A. Pandolfi, M. Ortiz, Metaconcrete: designed aggregates to enhance dynamic performance, J. Mech. Phys. Solids 65 (2014) 69–81.
[7] J.N. Munday, C.B. Bennett, W.M. Robertson, Band gaps and defect modes in periodically structured waveguides, J. Acoust. Soc. Am. 112 (4) (2002) 1353–1358.
[8] S.M. Ivansson, Sound absorption by viscoelastic coatings with periodically distributed cavities, J. Acoust. Soc. Am. 119 (6) (2006) 3558–3567.
[9] E. Yablonovitch, Inhibited spontaneous emission in solid-state physics and electronics, Phys. Rev. Lett. 58 (20) (1987) 2059.
[10] M.S. Kushwaha, P. Halevi, L. Dobrzynski, et al., Acoustic band structure of periodic elastic composites, Phys. Rev. Lett. 71 (13) (1993) 2022.
[11] S. Babaee, N. Viard, P. Wang, et al., Harnessing deformation to switch on and off the propagation of sound, Adv. Mater. 28 (8) (2016) 1631–1635.
[12] Z. Liu, X. Zhang, Y. Mao, et al., Locally resonant sonic materials, Science 289 (5485) (2000) 1734–1736.
[13] X.N. Liu, G.K. Hu, G.L. Huang, et al., An elastic metamaterial with simultaneously negative mass density and bulk modulus, Appl. Phys. Lett. 98 (25) (2011) 251907.
[14] J. Zhao, Y. Li, W.K. Liu, Predicting band structure of 3D mechanical metamaterials with complex geometry via XFEM, Comput. Mech. 55 (4) (2015) 659–672.
[15] M. Åberg, P. Gudmundson, The usage of standard finite element codes for computation of dispersion relations in materials with periodic microstructure, J. Acoust. Soc. Am. 102 (4) (1997) 2007–2013.
[16] D.C. Dobson, An efficient method for band structure calculations in 2D photonic crystals, J. Comput. Phys. 149 (2) (1999) 363–376.
[17] F. Kobayashi, S. Biwa, N. Ohno, Wave transmission characteristics in periodic media of finite length: multilayers and fiber arrays, Int. J. Solids Struct. 41 (26) (2004) 7361–7375.
[18] M.M. Sigalas, N. Garcıa, Theoretical study of three dimensional elastic band gaps with the finite-difference time-domain method, J. Appl. Phys. 87 (6) (2000) 3122–3125.
[19] Y. Wang, F. Li, Y. Wang, et al., Tuning of band gaps for a two-dimensional piezoelectric phononic crystal with a rectangular lattice, Acta Mech. Sinica 25 (1) (2009) 65–71.
[20] F. Wu, Z. Liu, Y. Liu, Acoustic band gaps in 2D liquid phononic crystals of rectangular structure, J. Phys. D: Appl. Phys. 35 (2) (2002) 162.
[21] Y.F. Wang, Y.S. Wang, L. Wang, Two-dimensional ternary locally resonant phononic crystals with a comblike coating, J. Phys. D: Appl. Phys. 47 (1) (2013) 015502.
[22] Y.F. Wang, Y.S. Wang, X.X. Su, Large bandgaps of two-dimensional phononic crystals with cross-like holes, J. Appl. Phys. 110 (11) (2011) 113520.
[23] Q. Chen, A. Elbanna, Modulating elastic band gap structure in layered soft composites using sacrificial interfaces, J. Appl. Mech. 83 (11) (2016).
[24] S. Nemat-Nasser, H. Sadeghi, A.V. Amirkhizi, et al., Phononic layered composites for stress-wave attenuation, Mech. Res. Commun. 68 (2015) 65–69.
[25] Y. Huang, J. Li, W. Chen, et al., Tunable bandgaps in soft phononic plates with spring-mass-like resonators, Int. J. Mech. Sci. 151 (2019) 300–313.
[26] P. Wang, J. Shim, K. Bertoldi, Effects of geometric and material nonlinearities on tunable band gaps and low-frequency directionality of phononic crystals, Phys. Rev. B 88 (1) (2013) 014304.
[27] X. Chen, X. Xu, S. Ai, et al., Active acoustic metamaterials with tunable effective mass density by gradient magnetic fields, Appl. Phys. Lett. 105 (7) (2014) 071913.

[28] Y. Yuan, W. Zhou, J. Li, et al., Tuning bandgaps in metastructured beams: numerical and experimental study, J. Zhejiang Univ.-Sci. A 20 (11) (2019) 811–822.
[29] W. Zhou, W. Chen, Z. Chen, et al., Actively controllable flexural wave band gaps in beam-type acoustic metamaterials with shunted piezoelectric patches, Eur. J. Mech. A Solids 77 (2019) 103807.
[30] Z. Wang, Q. Zhang, K. Zhang, et al., Tunable digital metamaterial for broadband vibration isolation at low frequency, Adv. Mater. 28 (44) (2016) 9857–9861.
[31] H.W. Dong, X.X. Su, Y.S. Wang, et al., Topological optimization of two-dimensional phononic crystals based on the finite element method and genetic algorithm, Struct. Multidiscip. Optim. 50 (4) (2014) 593–604.
[32] O. Sigmund, J. Søndergaard Jensen, Systematic design of phononic band–gap materials and structures by topology optimization, Phil. Trans. R. Soc. A 361 (1806) (2003) 1001–1019.
[33] M.I. Hussein, G.M. Hulbert, R.A. Scott, Dispersive elastodynamics of 1d banded materials and structures: design, J. Sound Vib. 307 (3–5) (2007) 865–893.
[34] M.I. Hussein, K. Hamza, G.M. Hulbert, et al., Multiobjective evolutionary optimization of periodic layered materials for desired wave dispersion characteristics, Struct. Multidiscip. Optim. 31 (1) (2006) 60–75.
[35] C.T. Chen, G.X. Gu, Machine learning for composite materials, MRS Commun. 9 (2) (2019) 556–566.
[36] R.S. Sutton, A.G. Barto, Introduction To Reinforcement Learning, MIT press, Cambridge, 1998.
[37] Z. Yang, Y.C. Yabansu, R. Al-Bahrani, et al., Deep learning approaches for mining structure–property linkages in high contrast composites from simulation datasets, Comput. Mater. Sci. 151 (2018) 278–287.
[38] S. Tiryaki, A. Aydın, An artificial neural network model for predicting compression strength of heat treated woods and comparison with a multiple linear regression model, Constr. Build. Mater. 62 (2014) 102–108.
[39] F. Khademi, M. Akbari, S.M. Jamal, et al., Multiple linear regression, artificial neural network, and fuzzy logic prediction of 28 days compressive strength of concrete, Front. Struct. Civ. Eng. 11 (1) (2017) 90–99.
[40] B.A. Young, A. Hall, L. Pilon, et al., CaN the compressive strength of concrete be estimated from knowledge of the mixture proportions? New insights from statistical analysis and machine learning methods, Cem. Concr. Res. 115 (2019) 379–388.
[41] G.X. Gu, C.T. Chen, M.J. Buehler, De novo composite design based on machine learning algorithm, Extreme Mech. Lett. 18 (2018) 19–28.
[42] J.C. Michel, H. Moulinec, P. Suquet, Effective properties of composite materials with periodic microstructure: a computational approach, Comput. Methods Appl. Mech. Engrg. 172 (1–4) (1999) 109–143.
[43] C. Qi, A. Fourie, Q. Chen, Neural network and particle swarm optimization for predicting the unconfined compressive strength of cemented paste backfill, Constr. Build. Mater. 159 (2018) 473–478.
[44] X. Li, S. Ning, Z. Liu, et al., Designing phononic crystal with anticipated band gap through a deep learning based data-driven method, Comput. Methods Appl. Mech. Engrg. (2019) 112737.
[45] E. Samaniego, C. Anitescu, S. Goswami, et al., An energy approach to the solution of partial differential equations in computational mechanics via machine learning: Concepts, implementation and applications, 2019, arXiv preprint arXiv:1908.10407.
[46] A. Mosavi, M. Salimi, S. Faizollahzadeh Ardabili, et al., State of the art of machine learning models in energy systems, a systematic review, Energies 12 (7) (2019) 1301.
[47] S. Goswami, C. Anitescu, S. Chakraborty, et al., Transfer learning enhanced physics informed neural network for phase-field modeling of fracture, Theor. Appl. Fract. Mech. (2019) 102447.
[48] J. Schulman, F. Wolski, P. Dhariwal, et al., Proximal policy optimization algorithms, 2017, arXiv preprint arXiv:1707.06347.
[49] V. Mnih, K. Kavukcuoglu, D. Silver, et al., Playing atari with deep reinforcement learning, 2013, arXiv preprint arXiv:1312.5602.
[50] C.J.C.H. Watkins, P. Dayan, Q-learning, Mach. Learn. 8 (3–4) (1992) 279–292.
[51] V. Mnih, A.P. Badia, M. Mirza, et al., Asynchronous methods for deep reinforcement learning, in: International Conference on Machine Learning, 2016, pp. 1928–1937.
[52] J. Schulman, S. Levine, P. Abbeel, et al., Trust region policy optimization, in: International Conference on Machine Learning, 2015, pp. 1889–1897.
[53] M.G. Bellemare, W. Dabney, R. Munos, A distributional perspective on reinforcement learning, in: Proceedings of the 34th International Conference on Machine Learning-Volume 70, JMLR. org, 2017, pp. 449–458.
[54] R.E. Camley, B. Djafari-Rouhani, L. Dobrzynski, et al., Transverse elastic waves in periodically layered infinite and semi-infinite media, Phys. Rev. B 27 (12) (1983) 7318.
[55] C. Kittel, Introduction To Solid State Physics, Wiley, New York, 1976.
[56] A.A. Novotny, J. Sokołowski, Topological Derivatives in Shape Optimization, Springer Science & Business Media, 2012.
[57] B. Bourdin, A. Chambolle, Design-dependent loads in topology optimization, ESAIM Control Optim. Calc. Var. 9 (2003) 19–48.
[58] G. Allaire, F. Jouve, A.M. Toader, A level-set method for shape optimization, C. R. Math. 334 (12) (2002) 1125–1130.
[59] Y.M. Xie, G.P. Steven, A simple evolutionary procedure for structural optimization, Comput. Struct. 49 (5) (1993) 885–896.
[60] M.P. Bendsœ, Optimal shape design as a material distribution problem, Struct. Optim. 1 (4) (1989) 193–202.