Journal of Materials Chemistry A

Reinforcement Learning-based Design of Shape-changing Metamaterials

Journal:	*Journal of Materials Chemistry A*
Manuscript ID	TA-ART-05-2023-003119.R2
Article Type:	Paper
Date Submitted by the Author:	09-Aug-2023
Complete List of Authors:	Bernaus Oliva, Sergi; Technical University of Denmark Bölle, Felix Tim; Technical University of Denmark, Department of Energy Conversion and Storage Las, A.T.; Technical University of Denmark Xia, Xiaoxing; Lawrence Livermore National Laboratory, Materials Engineering Division Castelli, Ivano; Danmarks Tekniske Universitet, Department of Energy Conversion and Storage

Reinforcement Learning-based Design of Shape-changing Metamaterials

Sergi Bernaus Oliva[1†], Felix T. Bölle[1†], A.T.Las[1], Xiaoxing Xia[2] and Ivano E. Castelli[1]

[1]Department of Energy Conversion and Storage, Technical University of Denmark, Street, Anker Engelundsvej 411, DK-2800, Kgs. Lyngby, Denmark.
[2]Materials Engineering Division, Lawrence Livermore National Laboratory, 7000 East Ave., Livermore, 94550, California, USA.

Contributing authors: xia7@llnl.gov; ivca@dtu.dk;
[†]These authors contributed equally to this work.

Abstract

During the last decade, artificially architected materials have been designed to obtain properties unreachable by naturally occurring materials, whose properties are determined by their atomic structure and chemical composition. In this work, we implement a new reinforcement learning (RL) method able to rationally design unique metamaterial structures at the nano-, micro-, and macroscale, which change shape during operational conditions. As an example, we apply this method to design nanostructured silicon anodes for Li-ion batteries (LIBs). The RL model is designed to apply different actions and predict change during operational conditions. The multi-component reward function comprises an increase in the total storage capacity of the resulting battery electrode and structural parameters, such as the minimum distance between the individual components of the nanostructure. Upon experimental validation using a polymer-based 3D printing technique, we expect that the newly discovered structures improve the current Si-based LIB anodes state-of-the-art by almost three times and almost ten times the current commercial LIB based on a graphitic anode. This RL-based optimization method opens up vast design space for other responsive metamaterials with tailored properties and pre-programmed structural transformation.

Keywords: metamaterials, batteries, reinforcement learning, silicon anode

1 Introduction

The development of new materials with better performance, novel functionality, and lower cost is crucial for meeting the ever-increasing societal need for a sustainable future. Traditional materials discovery is achieved by varying the chemical formulation and the crystallographic microstructure of a material until specific properties are optimized for the targeted application [1]. To expedite this largely "trial-and-error" approach [2], new materials can be designed and simulated first (often using atomistic simulations[3]), and then the best candidates are experimentally synthesised and tested afterwards. However, this is a lengthy and painstaking process as there is often a significant gap between what has been simulated and what could be synthesized[4]. To further expand the toolset for materials discovery with enhanced efficiency, researchers have been actively working on using materials' architecture as an additional turning knob to rationally design the so-called "metamaterials" to reach previously untapped areas in the materials property space. The properties and responsiveness of metamaterials are controlled not only by their chemical composition and atomic arrangement as in traditional materials but also through their artificial architectures at the nano-, micro-, and macroscale [5], which leads to intriguing applications in a variety of fields including energy storage[6, 7], structural mechanics [8], optics [9], acoustics [10], and bio-medical engineering [11]. This newly explored degree of freedom opens up a new vista of application opportunities for shape-changing materials other than the Si electrodes that will be discussed in this work. To induce a large degree of structural transformation, magnetically responsive materials such as ferromagnetic-particle-embedded composite polymers [12] and thermally responsive material such as shape memory polymer [13] and liquid crystal elastomers [14] can be used. The unique benefit of electrochemical modulation is that the shape change is non-volatile meaning that the structure will stay in the transformed geometry after the external stimuli are removed. A variety of electrochemically active materials exhibit large volumetric expansion and contraction associated with reversible redox reactions including but not limited to intermetallics [15], 2D materials [16], and conjugate polymers [17]. Liquid-update-induced swelling is another simple yet effective way to administer large volume changes in architected materials. A variety of polymers and composites swell significantly in aqueous and organic solvents and recover their original structures upon drying [18]. In particular, hydrogel is a versatile material that is sensitive to temperature, solvent type, and pH and ion concentration in the chemical environment [19]. For example, self-regulating microfluidic systems made out of shape changing metamaterials have been designed to maintain a stable temperature and pH [20]. Specific materials with various degrees of shape changes, mode of deformation, boundary conditions, stimulus types, and surrounding environment need to be carefully selected based on specific application need.

The design of the artificial structure of metamaterials is often inspired by conventional periodic lattice structures such as simple cubic, octet, or gyroid

structures and geometric motifs such as fractal or hierarchical structures [21], and then optimized by systematically varying geometric parameters such as beam diameter, wall thickness, and unit cell size. More recently, computational design methods such as topology optimization and machine learning (ML) have been implemented to accelerate the discovery process by generating virtual environments that can be used to make measurements without fabricating each structure. A few examples are the deep learning model used to discover relationships between metamaterial structures and their optical responses and to accelerate the inverse design of dielectric metasurfaces to miniaturise and develop optics [9, 22, 23]. These computational design methods are sufficient to optimize static materials properties but fall short in predicting metamaterials that change their shapes in response to external stimuli [24]. As illustrated in Fig. 1, such shape-changing metamaterials benefit a series of important applications, such as reconfigurable battery electrodes to store more energy while relieving internal stresses [25], tunable acoustic structures that can actively control sound propagation [26], stretchable kirigami polarization modulators for terahertz circular dichroism spectroscopy [27], and hydrogel metamaterials with responsive chirality to enable sensitive optical immunodetection of biomolecules [11]. To overcome the design challenge of such shape-changing metamaterials, we introduce a Reinforcement Learning (RL) method to optimize the their geometric structures based on specific application needs and the associated rewards.

In this work, we take shape-changing Li-ion battery electrodes [25] as an example, and implement a new RL model to design architected silicon (Si) electrodes for maximum storage capacity. Si is a next generation anode material for Li-ion batteries with a theoretical storage capacity 10 times greater than the currently dominating graphite anodes. However, Si undergoes up to 300% volume expansion during Li insertion leading to mechanical-failure-induced capacity fading and limited cycle life [28]. Recently, Xia and co-workers developed shaping-changing Si metamaterials that can accommodate large Si volume expansion and relieve the associated stresses by cooperative beam buckling [25]. Such Si architectures are composed of vertical posts connected with horizontal beams, as shown in the left part of Figure 1b. During lithiation, the beams cooperatively buckle, forming an in-plane sinusoidal pattern to accommodate the volume change as shown in the right part of Figure 1b. In response to the buckling torque, the vertical poles (or nodes) rotate in opposite directions, and the beam has the shape of a half-period sinusoidals. The change in shape happens in the 2D plane, and the final electrode material is obtained by stacking 2D layers with the same layout, which are connected by the vertical poles. These stacked square lattices show a high Si-mass-normalized capacity of 2010 mAh per g of Si after 50 cycles at C/6 but the total electrode-level capacity is low due to the low Si loading. Our RL model agent is designed to optimise the placement of these Si beams over an area in a virtual environment to find an optimised layout that maximises the theoretical storage capacity based on the input constraints, e.g., lengths of the beams and angles

formed by two beams, as well as, not overlapping beams both in the delithiated and lithiated shapes. Furthermore, we allow the RL model to select both non-frustrated and frustrated configurations, i.e., the beams connect two nodes rotating in the same direction, thus forming a full sinusoidal period, as shown later in Figure 2b. The most promising pattern found using this RL model has a theoretical capacity over four times larger than the original pattern proposed in Ref. [25]. To validate the RL-predicted design experimentally, we adopt a simple method in which 3D-printed polymer structures of the optimized metamaterial design are swollen by an organic solvent, which emulates the lithiation-induced volume expansion in Si-nanostructures and behaves as predicted in our model.

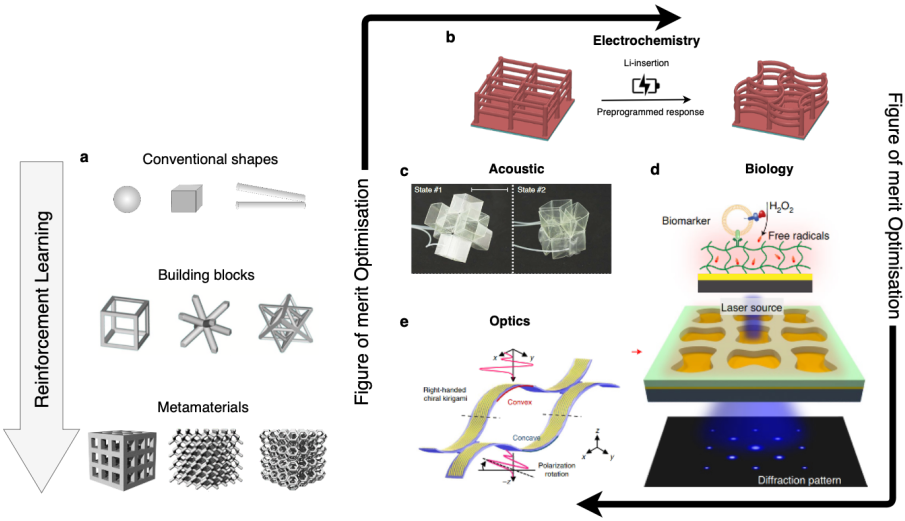

Fig. 1 Using reinforcement learning to build complex structures based on simple building blocks **(a)** we can achieve material properties not naturally present. Examples include increasing the storage capacity of a battery electrode **(b)**, controlling sound propagation direction using reconfigurable origami **(c)**, polarization modulation for terahertz wave **(d)**, as well as chirality-based optical immunodetection of biomolecules **(e)**. Images **(a)**[29], **(b)**[25], **(c)**[26], **(d)**[27], **(e)**[11] adapted with permission from their respective publishers and creators.

Even though we use these Si-nanostructure electrodes as an exemplar case, the RL-based design method is general. It can be applied to other shape-changing metamaterials as illustrated in Figure 1. Reconfigurable metamaterials can potentially benefit a series of important applications, such as shape-changing battery electrodes to store more energy, tunable phononic crystals that can actively control acoustic wave propagation, and adaptive membrane that can filter or release objects of different sizes on demand. RL-based topology optimization can predict improved metamaterial designs based on specific application needs and the associated rewards.

2 Methods

2.1 Reinforcement Learning

Reinforcement Learning (RL) has been successfully used to automate and accelerate the manual work of finding optimal structures given certain constraints [30]. RL algorithms are known for not requiring training beforehand, but instead of balancing exploration and exploitation based on the environment, configuration and current knowledge. In a standard RL model, an agent interacts with the environment to maximize a specified reward. The objective of our agent is to place Si-nodes and Si-beams connecting the nodes to create a pattern that changes shape upon lithiation (as shown in Ref. [25]). In RL literature, solving this problem has two approaches. One involves placing the nodes, known as the node placement problem, it has been successfully used in the design of the floor-planning of computer chips,[31] and the node linkage problem, which has been used in creating connections between users based on data from two different social networks.[32] The other consists of placing the nodes and edges together using smaller subgraph structures, e.g. single bonds[33] or larger molecular fragments.[34] In this work, we follow the latter approach. The packages used to build such a model are based on numpy, torch and gym, as well as matplotlib for the visualisations. Below here, we describe the design of the RL model in detail.

2.1.1 Environment and Subgraphs

The environment is modelled as an initially empty 2D canvas defined by its width and height, as shown in Figure 2a. Since we are interested in finding a repeatable pattern along all directions for maximum volumetric capacity, the canvas has periodic boundary conditions. While different unit cell shapes are possible, here, we constrain ourselves to an orthogonal unit cell. Nodes and actions are then added to the empty canvas. The nodes in the environment are characterized by their position and by a parameter that defines if they rotate to the left or right during lithiation (indicated with blue and red dots), forcing the connected beams to bend, forming a convex or concave shape, as shown in the figure 2a. During the training, the node rotation of the first placed node is set as left, and the rest of the nodes are based on the previously placed nodes. The beams are simply characterised by their lengths, which shape they form when lithiated (a sinusoidal with half or full period) and the nodes they connect. Experimentally the beams are designed to connect nodes rotating in a different direction, forming a sinusoidal of half-period [25]. Beams are placed as straight rods on the canvas, but during lithiation, they bend. The bending amount depends on the Li content of the beam, but to simplify the modelling, we approximate that they form a sinusoidal wave of a base amplitude of 0.49. With this amplitude, we obtain a length of the lithiated beams of 1.445, which closely resembles the bending observed in the fabricated materials [25].

The environment contains a list of subgraphs that can be added to the canvas. An example list can be seen in figure 2a. Each subgraph represents an

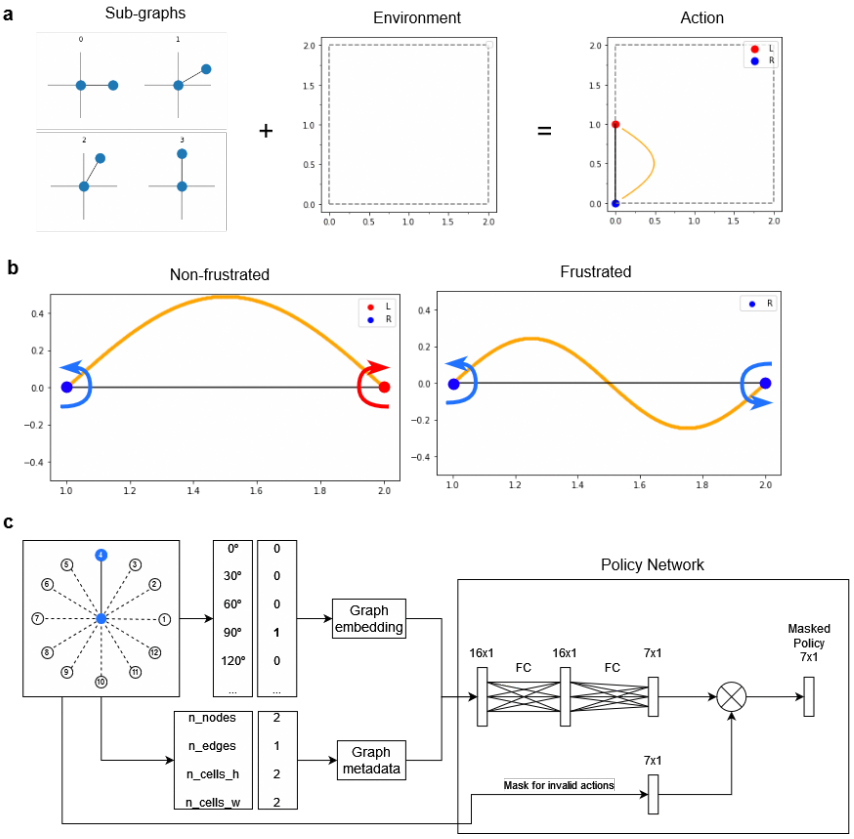

Fig. 2 **a**) example of the subgraphs available to the canvas for 30-degree increment angles and single length beam. Applying one of the subgraphs to the environment (canvas), we perform an action that applies a rotation to the nodes, creating the bent beam (orange). **b**) shape changes generated by the rotation of the nodes that the beam connects, a non-frustrated (frustrated) edge is represented by a half (full) sinusoidal curve with arrows indicating rotation direction of each beam (and Figure S1). **c**) schematics of the architecture of the Q-network model, the features include a graph embedding, that contains a representation of all the possible subgraph options that can be used starting from the given node, and the four metadata features: number of nodes, number of edges, width and height of the canvas. In this example there are twelve subgraphs, in addition to the the four metadata parameters, the sixteen features are then the input to the fully-connected (fc) neural network consisting of one additional ReLu layer as well as the output layer equal to the length of the possible actions. Subsequent to predicting the values for each possible action, a mask is applied to filter out illegal actions.

action that can be performed over a node and is defined by two nodes, connected by a beam. The parameters of a given subgraph are the angle and length that the beam forms between the two nodes. The list of subgraphs available to the environment is generated from the list of angles and beam lengths provided as input. Because of mirror symmetry, the angles are chosen between 0 and 180 degrees, included. Not all angles and beam lengths can generate patterns that are periodic within the canvas. To solve this, we allow the RL model to

adjust the input angles and lengths so that the actions can close the pattern within the canvas size, i.e., the beam lengths projected on the x- and y-axis of the canvas are dividends of the canvas dimensions. To calculate adjusting of angles and lengths, the initial angles and cell size are used as input. For each of the input angles, the sinus and cosine factors are calculated and rounded up to the closest even (for y and x axis respectively), after the values are found the length required is calculated using the Pythagorean theorem. This allows to have at least one valid closed shape based on the input parameters, as shown in Figure S2.

We note, that for symmetry reasons subgraphs with an angle in the (180,360) degree range are equivalent to (0,180). In most of the performed ML experiments, we define the angles in the [0,90] range and mirror them in the [90,180] range.

2.1.2 Actions

An action consists of applying a subgraph to a specific node, as illustrated in figure 2a. Starting from a selected node, the RL agent picks the next action to take, among a list of actions available for the selected node, masking out the actions that are illegal. An illegal action is defined as any action that results in two beams, either lithiated or not, colliding or a new node laying too close to already placed nodes. Since the illegal actions are filtered out, this results in the agent not learning from them, which has been found to improve the final solutions, as shown in Figure 3k.

In addition to these non-frustrated configurations, we allow exploring frustration, i.e., beams that connect nodes rotating in the same direction, which upon lithiation form beams shaped as full-period sinusoidals (Figure 2b). This can be done both within the RL model during training, and as post-processing after the structures have been optimised. If frustrations are allowed during training, when picking and filtering actions, edges that connect two nodes with the same direction of rotation are added as frustrated edges. When using post-processing frustration, we find all nodes of the same directionality that are close to each other and have no edge, and try to insert a frustrated edge in those positions, checking for collision with other beams beforehand, bypassing the requirement of needed a subgraph for each connection.

2.1.3 Agent and Rewards

The goal of RL agent is to find a set of state-action pairs, intended as the pair formed by a given state of the canvas and one action, that maximizes the total reward. The reward, η, is defined as the total beam length in the canvas, calculated as the sum of all the beam lengths before lithiation, divided by the canvas area. We employ Q-learning for training the agent,[35] more specifically the Delayed Q-learning, which aims at maximizing the reward by delaying any estimate until there is a statistically significant sample of observations. Q-learning is model-free, which means that the agent uses predictions

of the expected environment response to pick actions, it is based on a "trial-and-error" approach to learning instead of a reward system, which predicts the future reward and points towards increasing it. The actions taken by Q-learning are based on the Bellmann equation, which defines the quality of a state, intended as the current canvas layout. This suggests which actions to take based on how good the current state is. Since there is a huge amount of possible positions for the nodes, i.e. 2^{225} for a 2x2 canvas size, the values used by the Q-learning algorithm to determine actions (Q-values) are approximated using a neural network turning the agent into a Deep Q-network (DQN).[36]

Figure 2c shows a detailed overview of the neural network used to predict the Q-value. The input of the neural network consists of four metadata values (used for internal representation, number of nodes and edges, and width and height of the canvas) as well as a vector representing all the subgraphs for that environment, indicated by one or zero, depending on whether the subgraph is used or not in that state. For dealing with the exploration-exploitation problem, we employ the ϵ-greedy strategy[37], letting the agent take random actions with probability ϵ. We slowly decrease the ϵ value throughout an episode following an exponential decay. For determining the activation of neurons of the network, we use a ReLu curve, since it has the property of not activating all the neurons at the same time, this would provide a benefit due to the fact that our graph structures are symmetrical. Using a neural network to represent the Q-value can cause instabilities.[36] To counteract this fact we employ several measures to avoid divergent gradients including double Q-learning[38] and experience replay with a capacity of 1000, which was selected based on trial and error.[39]

3 Results

All the data produced for this paper[40] as well as the source code are available in the GitHub repository: https://gitlab.com/Feltbo/metabatt.

3.1 RL-designed Nanostructures

We now use the RL model described above to design Si-nanostructures aiming at increasing the volumetric capacity of the silicon anode. Because of limitations in the fabrication, as shown in Section 3.2, we assume that the minimum angle between beams is around 30 degrees. The angle was decided after trial and error from different angles, starting from 30 and 45 degrees. For now, we also assume that the length of each beam is equivalent in accordance with experiments performed in Ref. [25].

We define the score of the RL model as the sum of the lengths of beams that are in the canvas, such as that the more beams are placed, the higher the reward. In figure 3a, we can observe the evolution of the reward as a function of the number of episodes. For each edge added to the canvas, the reward is increased according to the length of the beam placed. This process can be clearly observed between steps **(c)** and **(d)**, where the agent finds a way to

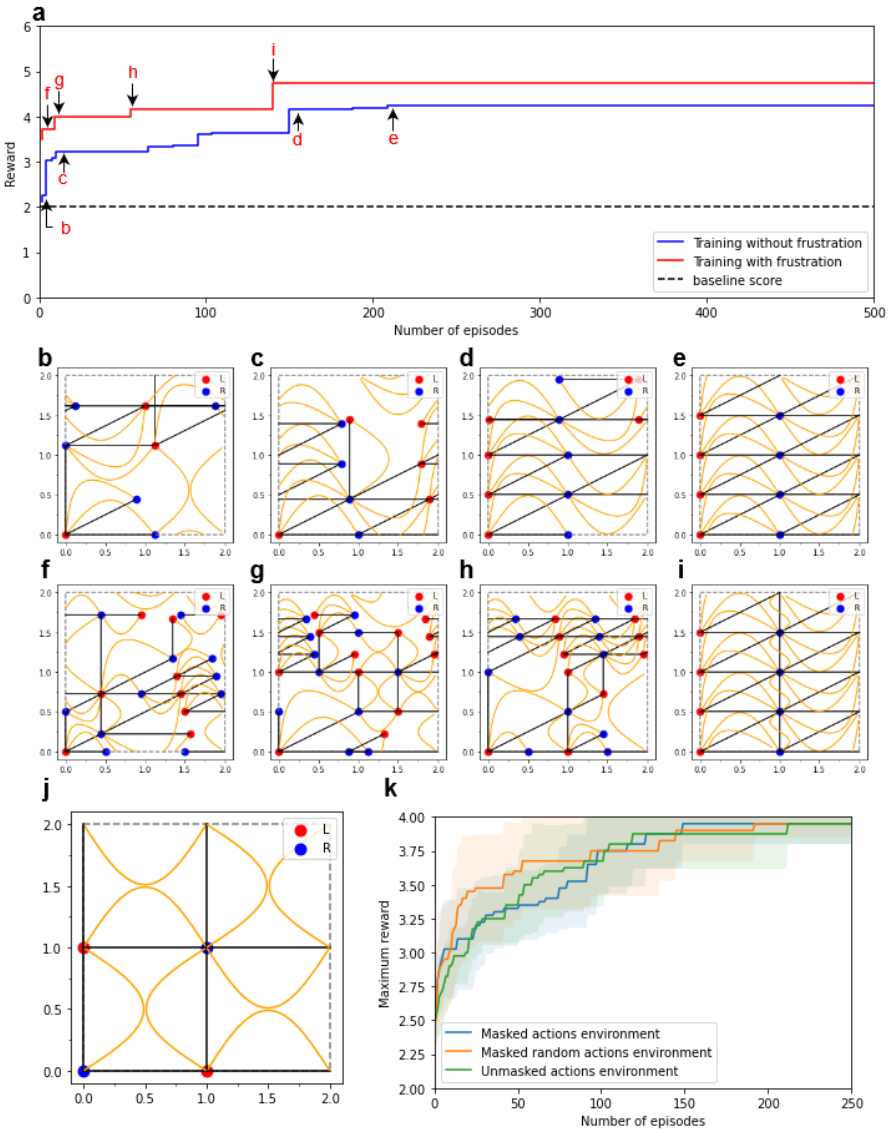

Fig. 3 **a**) Evolution of the reward for the training of subgraphs for 0, 30, 60, 90, 120, 150, 180 degrees during 500 episodes with and without frustration (in blue and red, respectively). **b-i**) nodes and beams configurations for selected improvements in the reward shown in **(a)**. **j**) the configuration corresponding to the results reported in Ref. [25] and taken as a baseline in this work. **(k)** comparison of the maximum reward achieved over the number of episodes for different algorithm configurations averaged out of 10 runs each, using [0, 30, 60, 90, 120, 150, 180] angles and fixed length of 1.

Angles	Not adjusted		Adjusted	
	Non frustrated	**Frustrated**	**Non frustrated**	**Frustrated**
0,15,75,90	7.50	7.50	8.12*	7.56*
0,30,60,90	4.00*	5.04*	4.62 (4.24*)	5.26 (4.74*)
0,45,90	2.50	3.18	2.50*	3.41*
0,90	2.00*	2.00*	2.00*	2.00*

Table 1 Maximum theoretical capacity calculated for various actions (angles), with both frustrated and not frustrated for adjusted and non-adjusted angles and lengths. The * indicates a closed shape.

optimise the positioning of the beams creating a shape that allows more beams to be added to the canvas. As a baseline, we take a tetragonal lattice shown in Figure 3j, which has a reward of 2.00. When considering 30-degree increments from 0 to 180, excluding frustrated configurations (blue line), we obtain the canvas of 3i, which has aligned nodes and a reward of 4.00. When frustrated connections are included (red line), we obtain a similar canvas, but with nodes rotating in the same direction that is connected, for a total reward of 4.24. We have implemented two approaches when handling the frustrated edges, the first one, which is the one that can be observed in 3a, tries to add frustrated edges to the canvas during the learning process. We note that because of the larger amount of possible connections, the RL-model becomes slower to train when frustrated connections are allowed. As mentioned before, we have also implemented a second method of including frustrated edges in the canvas, a post-processing tool, which tries to add frustrated connections to an optimised non-frustrated canvas, which increases the training speed. For the case shown here, the two methods give identical results.

Table 1 shows the reward for different sets of angles. The highest reward is obtained for small angles, e.g. 15 degrees. However, the small angles formed by the beams would most likely impact the cyclability of the anode material. We thus consider 30 degrees as the smallest angle, which can give a good trade-off between higher storage capacity and the long life of the electrode. We note that, in most cases, adding frustrations has a minor impact, below 20%, on the final reward (except for the 45-degree case). The frustration, however, increases the mechanical strain on the beam, thus reducing the lifetime and cyclability of the battery.

Due to this adjusting step, all the adjusted shapes found in 1 all have the * indicating that are closed shapes, but in some instances (30 degrees increments), the algorithm managed to find a shape that has an even higher score by using the beams in unique ways that were not predicted. The drawback is that the generated figures when this happens tend to not be closed shapes, as shown in Figure 4a. We have included both results in the table to be able to compare the increase in reward, at the cost of losing the closing property for the shape.

The angles used in the training for B in figure 4 are increments of 30 degrees (0, 30, 60, 90, 120, 150, 180), and a fixed length of 1. Meanwhile, the ones used for C are those same angles but adjusted ($\approx$ 26 degrees), and using multiple

lengths. As it can be observed, adjusting the angles and lengths to fit with a specific periodicity within the canvas produces a result that achieves a better score compared to the base angles used, 4.24 vs 4 respectively. Both results provide a high increase in the Si density compared to the results achieved in [25].

3.1.1 Evaluating the RL-model

Since the focus of this paper is the newly created simulation environment, we provide proof of the importance of only allowing legal actions. This incorporates knowledge about the environment, making it easier for the agent to learn the best possible actions. For comparison, we create an environment that does not filter out illegal actions. Every time an illegal action is played, the episode stops, and the reward at this state is returned. We call this environment "Unmasked actions Environment". We mark the environment only allowing legal actions, but acting randomly as "Masked Random Environment".

Figure 3k shows that the environment which does not filter out illegal actions achieves the final reward in 50 more episodes compared to the environment of the masked action. And since we are considering fewer actions, the masked environments process each episode faster than the random agent, which leads to a faster time to process each episode. Considering the relative simplicity of the action space used for this experiment (i.e., only one single edge length), it shows that an RL agent becomes even more valuable, if not inevitable, when trying to tackle problems with even larger phase spaces (i.e., more possible actions).

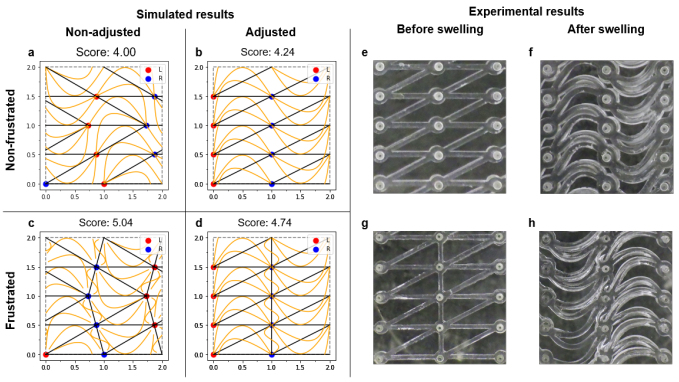

Fig. 4 Results of training with different input parameters. **(a,e)** use 30-degree increments input [0, 30, 60, 90, 120, 150, 180] and a single length, **(b,f)** use the same increments with adjusted angles and multiple lengths. Our baseline score is 2.00.**(c,g)** are pictures of the as-fabricated PDMS lattices based on the non-frustrated and frustrated designs.**(d,h)** are their transformed geometries after swollen by hexane, which agrees with our model's predictions.

3.2 Experimental validation

To validate the RL-optimized metamaterial design, we use additive manufacturing techniques to fabricate polydimethylsiloxane (PDMS) structures and emulate the swelling behaviour by immersing the structures in hexane. 2D PDMS lattices are fabricated based on the non-adjusted designs in Figure 4a and manufactured in Figure 4e. Acrylic-based templates with the desired geometries are 3D-printed using a Formlab Form 3 stereolithographic (SLA) printer and Formlab Clear V4 resin. Precursors of the SYLGARD 184 two-part PDMS kit are mixed and then poured into the template. After curing at 80C for two hours, the PDMS lattices were removed from the template. For the non-frustrated design, the node spacing is 20mm in the x direction and 10mm in the y direction. For the frustrated design, the node spacing is 26mm in the x direction and 13mm in the y direction. For both designs, the beam width is 2mm, and the beam thickness in the z direction is 4mm, and the nodes are approximated by doughnut-shaped rings with an inner diameter of 2mm and an outer diameter of 6mm as shown in Figure 4e and Figure 4g. Before the swelling experiments, the PDMS lattices are mounted onto a 3D-printed base plate with vertical posts that are inserted into all doughnut-shaped nodes. In doing so, the lateral positions of all the nodes in the PDMS lattices are constrained but the nodes are free to rotate in the right-handed or left-handed directions. During the swelling experiments, the PDMS lattices mounted on the base plates are immersed in hexane for 10min. The swollen PDMS lattices are then taken out of hexane so that their pictures can be taken with better contrast, as shown in Figure 4d and Figure 4h for the non-frustrated case and the frustrated case respectively (more detail are reported in the Supplementary Information and Figure S3). In this way, we have demonstrated the RL-optimized metamaterial designs would behave largely as predicted in physical experiments. The only discrepancy from the RL model is the buckling directions in the frustrated case are opposite, which will be discussed below.

4 Discussion

The discrepancy between the RL model and the swelling experiment can be attributed to the simplified physical parameters and constitutive equations used in the model, which we intentionally keep simple to demonstrate the general usage of the methodology. The reward in the current RL model is based only on the structural deformation of the beams, whose parameters have been obtained from Ref. [25], with the condition that two beams cannot touch each other and that the shape of the lithiated beam is simplified to be sinusoidal. Figures 4f and 4h indicate that in the solvent-swollen PDMS lattices, the turning point of buckled beams, where the radius of curvature is the smallest, is not at the mid-point of the beams, deviating from the sinusoidal shape used in the RL model. Another difference between the model and the polymer lattice experiment is that the nodes of the lattices are approximated by doughnut-shaped rings that are free to rotate when stationary vertical posts on the

substrate are inserted into them. Our model doesn't consider the deformation of the doughnut-shaped nodes, which are clearly stretched and deformed after swelling in Figures 4f and 4h.

Even in our simplified model, the predicted behaviour is accurate for the non-frustrated case, where all beams deform by Mode-I buckling in a cooperative way to minimize the elastic energy of the system. In the frustrated case, the frustrated beams deform by Mode-II buckling with a much higher elastic strain, which brings the energy penalty of each node to rotate in one direction to be very close to that in the other direction. Only in this situation, finer details of the deformation geometry and mechanical states start to play a noticeable role resulting in all nodes rotating in opposite directions as the RL model predicts. Even in this case, the model is largely correct because the two longer beams in the lattice deform by Mode-I (i.e., non-frustrated) and the shortest beam deforms by Mode-II buckling (i.e., frustrated) in both the experiment and the model.

The details of the RL model can be modified to improve its accuracy in the current nanostructured Si anode cases, but the framework used to optimize shape-changing metamaterials remains the same and can be adapted for various other metamaterials applications. For example, the deformed shape of the beams can be formulated more accurately using established solid mechanics beam theories. The exact structural and energetic response of the beam to lithiation or solvent-induced swelling could also be included by modelling the beams at the continuum level, through Finite Element Models (FEM). This would contribute to a more accurate description of the shape-changing canvas as well as include energetic information in the modelling, which can become critical when frustrated beams are present.

5 Conclusion

In this work, we have implemented a reinforcement learning method to rationally design shape-changing metamaterials using nano-structured Si anodes as a demonstration case. The RL model places pre-defined actions (nodes and beams) in an empty canvas optimising a multi-component reward function, which considers not only maximum theoretical capacity but also the structural response to electrochemical cycling—changes in beam shape during lithiation and delithiation processes. The efficiency of different RL environments (in which actions were selected randomly or filtered) has been evaluated and compared. The model discovered a new structure which can improve the capacity to 2.5 times that of the nanostructured Si anode reported in Ref. [25]. To validate if the RL-predicted metamaterial designs are indeed physical, we have used an additive manufacturing approach to create polymer lattices based on such designs and demonstrated largely identical shape-changing behaviours in response to solvent-induced swelling with the exception of the node rotating direction in the RL-designed frustrated lattices. The RL-based topology optimization framework used in this work can be adapted to predict improved

metamaterial designs based on specific demands and the associated rewards for a wide variety of future applications. For example, adaptive membranes can be designed to filter or release objects of specific sizes on demand, and their structures can be optimized by applying the RL model to increase efficiency and size selectivity. Tunable phononic or photonic crystals can also be designed to control the propagation of acoustic or optical wave of specific frequencies using the RL model. Complex structural transformation can be simplify into reduced-order steps such as node rotation and beam bending used in this work and then incorporated into the corresponding, computationally efficient RL models with application-specific rewards.

Acknowledgements

FTB and IEC acknowledge support from the Independent Research Fund Denmark (Green Transition Project 1, project "Reconfigurable Metamaterials for Next Generation High-capacity Batteries" under grant number 0217-00111B). X.X. acknowledges the financial support from Lawrence Livermore National Laboratory's Lab Directed Research and Development Program (22-ERD-004). Work at LLNL was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344.

References

[1] Curtarolo, S., Hart, G.L.W., Nardelli, M.B., Mingo, N., Sanvito, S., Levy, O.: The high-throughput highway to computational materials design. Nature Materials **12**(3), 191–201 (2013). https://doi.org/10.1038/nmat3568

[2] Alberi, K., Nardelli, M.B., Zakutayev, A., Mitas, L., Curtarolo, S., Jain, A., Fornari, M., Marzari, N., Takeuchi, I., Green, M.L., Kanatzidis, M., Toney, M.F., Butenko, S., Meredig, B., Lany, S., Kattner, U., Davydov, A., Toberer, E.S., Stevanovic, V., Walsh, A., Park, N.-G., Aspuru-Guzik, A., Tabor, D.P., Nelson, J., Murphy, J., Setlur, A., Gregoire, J., Li, H., Xiao, R., Ludwig, A., Martin, L.W., Rappe, A.M., Wei, S.-H., Perkins, J.: The 2019 materials by design roadmap. Journal of Physics D: Applied Physics **52**(1), 013001 (2018). https://doi.org/10.1088/1361-6463/aad926

[3] Schaarschmidt, J., Yuan, J., Strunk, T., Kondov, I., Huber, S.P., Pizzi, G., Kahle, L., Bölle, F.T., Castelli, I.E., Vegge, T., Hanke, F., Hickel, T., Neugebauer, J., Rêgo, C.R.C., Wenzel, W.: Workflow engineering in materials design within the BATTERY 2030+ project. Advanced Energy Materials **12**(17), 2102638 (2021). https://doi.org/10.1002/aenm.202102638

[4] Atkins, D., Ayerbe, E., Benayad, A., Capone, F.G., Capria, E., Castelli, I.E., Cekic-Laskovic, I., Ciria, R., Dudy, L., Edström, K., Johnson, M.R., Li, H., Lastra, J.M.G., Souza, M.L.D., Meunier, V., Morcrette, M., Reichert, H., Simon, P., Rueff, J.-P., Sottmann, J., Wenzel, W., Grimaud, A.: Understanding battery interfaces by combined characterization and simulation approaches: Challenges and perspectives. Advanced Energy Materials **12**(17), 2102687 (2021). https://doi.org/10.1002/aenm.202102687

[5] Cui, T.J., Smith, D., Liu, R. (eds.): Metamaterials. Springer, ??? (2010). 10.1007/978-1-4419-0573-4. https://doi.org/10.1007/978-1-4419-0573-4

[6] Xia, X., Di Leo, C.V., Gu, X.W., Greer, J.R.: In situ lithiation–delithiation of mechanically robust cu–si core–shell nanolattices in a scanning electron microscope. ACS Energy Letters **1**(3), 492–499 (2016)

[7] Narita, K., Citrin, M.A., Yang, H., Xia, X., Greer, J.R.: 3d architected carbon electrodes for energy storage. Advanced Energy Materials **11**(5), 2002637 (2021)

[8] Meza, L.R., Das, S., Greer, J.R.: Strong, lightweight, and recoverable three-dimensional ceramic nanolattices. Science **345**(6202), 1322–1326 (2014)

[9] Jiang, J., Fan, J.A.: Global optimization of dielectric metasurfaces using a physics-driven neural network. Nano letters **19**(8), 5366–5372 (2019)

[10] Zhu, Y., Gerard, N.J., Xia, X., Stevenson, G.C., Cao, L., Fan, S., Spadaccini, C.M., Jing, Y., Assouar, B.: Systematic design and experimental demonstration of transmission-type multiplexed acoustic metaholograms. Advanced Functional Materials **31**(27), 2101947 (2021)

[11] Zhao, H., Pan, S., Natalia, A., Wu, X., Ong, C.-A.J., Teo, M.C., So, J.B., Shao, H.: A hydrogel-based mechanical metamaterial for the interferometric profiling of extracellular vesicles in patient samples. Nature Biomedical Engineering **7**(2), 135–148 (2023)

[12] Kim, Y., Yuk, H., Zhao, R., Chester, S.A., Zhao, X.: Printing ferromagnetic domains for untethered fast-transforming soft materials. Nature **558**(7709), 274–279 (2018)

[13] Yang, C., Boorugu, M., Dopp, A., Ren, J., Martin, R., Han, D., Choi, W., Lee, H.: 4d printing reconfigurable, deployable and mechanically tunable metamaterials. Materials Horizons **6**(6), 1244–1250 (2019)

[14] Kotikian, A., Truby, R.L., Boley, J.W., White, T.J., Lewis, J.A.: 3d printing of liquid crystal elastomeric actuators with spatially programed

nematic order. Advanced materials **30**(10), 1706164 (2018)

[15] McDowell, M.T., Lee, S.W., Nix, W.D., Cui, Y.: 25th anniversary article: understanding the lithiation of silicon and other alloying anodes for lithium-ion batteries. Advanced materials **25**(36), 4966–4985 (2013)

[16] Acerce, M., Akdoğan, E.K., Chhowalla, M.: Metallic molybdenum disulfide nanosheet-based electrochemical actuators. Nature **549**(7672), 370–373 (2017)

[17] Jager, E.W., Inganas, O., Lundstrom, I.: Microrobots for micrometer-size objects in aqueous media: potential tools for single-cell manipulation. Science **288**(5475), 2335–2338 (2000)

[18] Sydney Gladman, A., Matsumoto, E.A., Nuzzo, R.G., Mahadevan, L., Lewis, J.A.: Biomimetic 4d printing. Nature materials **15**(4), 413–418 (2016)

[19] Jin, D., Chen, Q., Huang, T.-Y., Huang, J., Zhang, L., Duan, H.: Four-dimensional direct laser writing of reconfigurable compound micromachines. Materials Today **32**, 19–25 (2020)

[20] He, X., Aizenberg, M., Kuksenok, O., Zarzar, L.D., Shastri, A., Balazs, A.C., Aizenberg, J.: Synthetic homeostatic materials with chemo-mechano-chemical self-regulation. Nature **487**(7406), 214–218 (2012)

[21] Meza, L.R., Zelhofer, A.J., Clarke, N., Mateos, A.J., Kochmann, D.M., Greer, J.R.: Resilient 3d hierarchical architected metamaterials. Proceedings of the National Academy of Sciences **112**(37), 11502–11507 (2015)

[22] Ma, W., Cheng, F., Liu, Y.: Deep-learning-enabled on-demand design of chiral metamaterials. ACS Nano **12**(6), 6326–6334 (2018) https://doi.org/10.1021/acsnano.8b03569. https://doi.org/10.1021/acsnano.8b03569. PMID: 29856595

[23] Zhelyeznyakov, M.V., Brunton, S., Majumdar, A.: Deep learning to accelerate scatterer-to-field mapping for inverse design of dielectric metasurfaces. ACS Photonics **8**(2), 481–488 (2021) https://doi.org/10.1021/acsphotonics.0c01468. https://doi.org/10.1021/acsphotonics.0c01468

[24] Xia, X., Spadaccini, C.M., Greer, J.R.: Responsive materials architected in space and time. Nature Reviews Materials **7**(9), 683–701 (2022)

[25] Xia, X., Afshar, A., Yang, H., Portela, C.M., Kochmann, D.M., Di Leo, C.V., Greer, J.R.: Electrochemically reconfigurable architected materials.

Nature **573**(7773), 205–213 (2019)

[26] Overvelde, J.T.B., de Jong, T.A., Shevchenko, Y., Becerra, S.A., Whitesides, G.M., Weaver, J.C., Hoberman, C., Bertoldi, K.: A three-dimensional actuated origami-inspired transformable metamaterial with multiple degrees of freedom. Nature Communications **7** (2016). https://doi.org/10.1038/ncomms10929

[27] Choi, W.J., Cheng, G., Huang, Z., Zhang, S., Norris, T.B., Kotov, N.A.: Terahertz circular dichroism spectroscopy of biomaterials enabled by kirigami polarization modulators. Nature materials **18**(8), 820–826 (2019)

[28] McDowell, M.T., Lee, S.W., Nix, W.D., Cui, Y.: 25th anniversary article: Understanding the lithiation of silicon and other alloying anodes for lithium-ion batteries. Advanced Materials **25**(36), 4966–4985 (2013). https://doi.org/10.1002/adma.201301795

[29] Benedetti, M., du Plessis, A., Ritchie, R.O., Dallago, M., Razavi, S.M.J., Berto, F.: Architected cellular materials: A review on their mechanical properties towards fatigue-tolerant design and fabrication. Materials Science and Engineering: R: Reports **144**, 100606 (2021). https://doi.org/10.1016/j.mser.2021.100606

[30] Zhelyeznyakov, M.V., Brunton, S., Majumdar, A.: Deep Learning to Accelerate Scatterer-to-Field Mapping for Inverse Design of Dielectric Metasurfaces. ACS Photonics **8**(2), 481–488 (2021). https://doi.org/10.1021/acsphotonics.0c01468. Accessed 2022-06-06

[31] Mirhoseini, A., Goldie, A., Yazgan, M., Jiang, J.W., Songhori, E., Wang, S., Lee, Y.-J., Johnson, E., Pathak, O., Nazi, A., *et al.*: A graph placement methodology for fast chip design. Nature **594**(7862), 207–212 (2021)

[32] Kong, X., Zhang, J., Yu, P.S.: Inferring anchor links across multiple heterogeneous social networks. In: Proceedings of the 22nd ACM International Conference on Information & Knowledge Management, pp. 179–188 (2013)

[33] Simm, G., Pinsler, R., Hernández-Lobato, J.M.: Reinforcement learning for molecular design guided by quantum mechanics. In: International Conference on Machine Learning, pp. 8959–8969 (2020). PMLR

[34] Ståhl, N., Falkman, G., Karlsson, A., Mathiason, G., Bostrom, J.: Deep reinforcement learning for multiparameter optimization in de novo drug design. Journal of chemical information and modeling **59**(7), 3166–3176 (2019)

[35] Watkins, C.J.C.H.: Learning from delayed rewards (1989)

[36] Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A.A., Veness, J., Bellemare, M.G., Graves, A., Riedmiller, M., Fidjeland, A.K., Ostrovski, G., *et al.*: Human-level control through deep reinforcement learning. nature **518**(7540), 529–533 (2015)

[37] Tokic, M., Palm, G.: Value-difference based exploration: Adaptive control between epsilon-greedy and softmax. In: Bach, J., Edelkamp, S. (eds.) KI 2011: Advances in Artificial Intelligence, pp. 335–346. Springer, Berlin, Heidelberg (2011). https://doi.org/10.1007/978-3-642-24455-1_33

[38] Van Hasselt, H., Guez, A., Silver, D.: Deep reinforcement learning with double q-learning. In: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 30 (2016)

[39] Lin, L.-J.: Reinforcement Learning for Robots Using Neural Networks. Carnegie Mellon University, ??? (1992)

[40] Castelli, I.E., Arismendi-Arrieta, D.J., Bhowmik, A., Cekic-Laskovic, I., Clark, S., Dominko, R., Flores, E., Flowers, J., Frederiksen, K.U., Friis, J., Grimaud, A., Hansen, K.V., Hardwick, L.J., Hermansson, K., Königer, L., Lauritzen, H., Cras, F.L., Li, H., Lyonnard, S., Lorrmann, H., Marzari, N., Niedzicki, L., Pizzi, G., Rahmanian, F., Stein, H., Uhrin, M., Wenzel, W., Winter, M., Wölke, C., Vegge, T.: Data management plans: the importance of data management in the BIG-MAP project. Batteries &amp Supercaps **4**(12), 1803–1812 (2021). https://doi.org/10.1002/batt.202100117