

ARTICLE OPEN



Autonomous reinforcement learning agent for stretchable kirigami design of 2D materials

Pankaj Rajak^{1,2}, Beibei Wang^{2,3}, Ken-ichi Nomura^{2,4}, Ye Luo¹, Aiichiro Nakano^{1,2,3,4,5}, Rajiv Kalia^{2,3,4,5} and Priya Vashishta^{1,2,3,4,5}

Mechanical behavior of 2D materials such as MoS₂ can be tuned by the ancient art of kirigami. Experiments and atomistic simulations show that 2D materials can be stretched more than 50% by strategic insertion of cuts. However, designing kirigami structures with desired mechanical properties is highly sensitive to the pattern and location of kirigami cuts. We use reinforcement learning (RL) to generate a wide range of highly stretchable MoS₂ kirigami structures. The RL agent is trained by a small fraction (1.45%) of molecular dynamics simulation data, randomly sampled from a search space of over 4 million candidates for MoS₂ kirigami structures with 6 cuts. After training, the RL agent not only proposes 6-cut kirigami structures that have stretchability above 45%, but also gains mechanistic insight to propose highly stretchable (above 40%) kirigami structures consisting of 8 and 10 cuts from a search space of billion candidates as zero-shot predictions.

npj Computational Materials (2021)7:102; <https://doi.org/10.1038/s41524-021-00572-y>

INTRODUCTION

Two-dimensional (2D) materials have been highlighted in recent studies for their promising applications in various fields such as catalysis, photonics, optoelectronic and spintronic devices, including sensors and high-performance electrodes^{1–6}. Their outstanding functionality and tunability have been the spotlight of research in nanoscience community. Among 2D materials, transition metal dichalcogenide (TMDC) is drawing a great deal of attention and are considered promising candidates for the next generation of thermoelectric and optoelectronic devices due to their excellent electrical properties, mechanical strength, and chemical stability^{7–9}. For example, molybdenum disulfide (MoS₂), a well-known solid lubricant, has become an excellent candidate for wearable electronic devices due to its high electron mobility, wide band gap, flexibility, and fracture toughness^{10–12}. Numerous studies of TMDC materials have reported surprising structure-property relationships due to defects, mechanical straining, and vertical stacking of heterogenous TMDC layers^{13–16}. For example, Ajayan and collaborators have shown experimentally that under strain an MoWSe₂ heterostructure undergoes a structural transformation near the crack tip¹⁷. They find that the transformation from the semi-conducting 2H to metallic 1T phases increases the fracture toughness.

A popular approach to engineering mechanical properties is to design a structural pattern inside materials. Mechanical meta-materials utilize hierarchical patterning in their structure. For example, kagomé lattice and octet truss exhibit negative Poisson's ratio, negligible shear modulus, and negative compressibility^{18–21}. Origami or kirigami design based on the ancient paper crafting technique are employed to change the mechanical behavior of 2D materials^{22–25}. For example, graphene is brittle in nature but its flexibility can be substantially enhanced by introducing cut patterns in the graphene sheet, thereby enabling stretchable electronics^{26,27}. Blees et al. have successfully created nanodevices based on kirigami-patterned graphene

and demonstrated its excellent flexibility²⁷. Unlike the traditional composite materials that are based on molecular assembly, designed patterns and structures can be introduced into materials using the kirigami technique. As a consequence, a variety of applications of kirigami are found in nanocomposites, soft electronics, and functional and optical metamaterials^{28–32}. However, the mechanical properties of kirigami-based 2D materials are sensitive to a number of parameters in a kirigami pattern, e.g., the total number of cuts, cut location, length, and width, as well as the interaction between cuts^{33,34}. A carefully designed kirigami pattern must incorporate both in-plane and out-of-plane deformation effects subjected to a strain in order to render high stretchability in the 2D structure. Zhu et al. have experimentally shown that graphene can be stretched up to 50%³⁵. Fracture mechanics in a 2D nanosheet with many cuts is non-trivial. Furthermore, the exponential size of the parameter space makes it impractical for experimental and theoretical studies to find the optimum kirigami pattern with the maximum stretchability.

Recently, machine learning (ML) methods have shown tremendous success in material science, especially in designing data-driven force field models for materials simulation, building ML models to predict materials property from their atomic structure, and even build models to create atomic structure for materials that exhibit a set of desired properties^{36–42}. Neural network and Gaussian processes have been used to develop data-driven interatomic potentials that enable molecular dynamics (MD) simulations with nearly the accuracy of first-principles quantum mechanical calculations^{43–46}. Also, ML models based on kernel ridge regression, support vector machines, and graph neural network have been used to predict a wide range of material properties such as elastic constants, dielectric constants, and thermoelectric properties, where the model input is a representation of the material under consideration, and the output from the model is a target property^{47–53}. In fact, recently ML models have also been used to predict toughness and

¹Argonne Leadership Computing Facility, Argonne National Laboratory, Argonne, IL, USA. ²Collaboratory for Advanced Computing and Simulations, University of Southern California, Los Angeles, CA, USA. ³Department of Physics and Astronomy, University of Southern California, Los Angeles, CA, USA. ⁴Department of Chemical Engineering and Materials Science, University of Southern California, Los Angeles, CA, USA. ⁵Department of Computer Science, University of Southern California, Los Angeles, CA, USA. ✉email: rajak@usc.edu; priyav@usc.edu

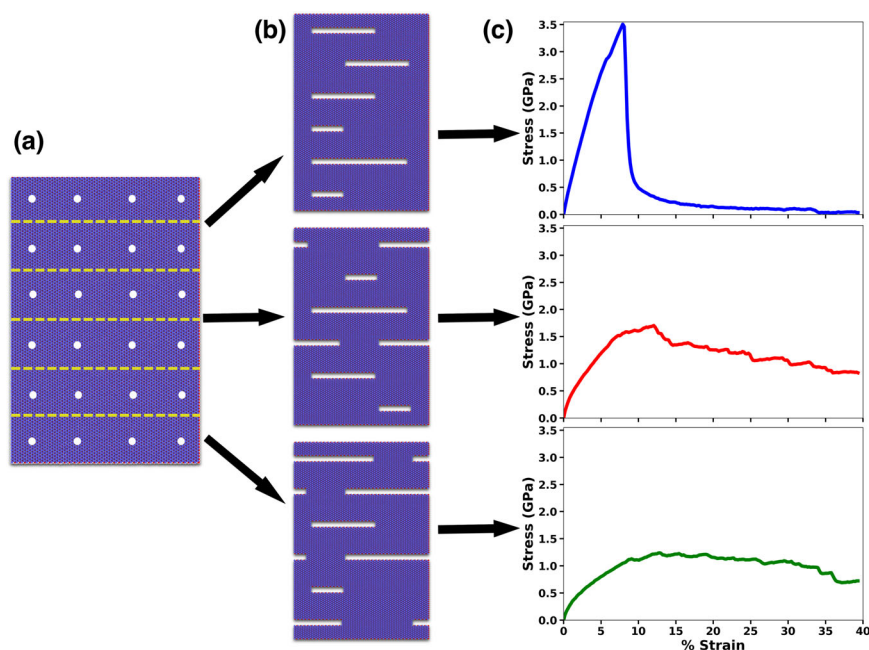


Fig. 1 MD simulation of MoS₂ kirigami structures. **a** Schematic of 20 nm × 30 nm MoS₂ structure showing the vertical partitioning of system into 6 groups (yellow dashed lines) for the synthesis of kirigami structures. In each row, a cut can be placed at one of the four locations (white circle). **b** Three kirigami structures created by different choices of cut length and location and **c** their corresponding stress–strain curves by uniaxial tensile deformation simulation using molecular dynamics.

strength of composite structures and predict crack propagation path in crystalline solids during dynamic fracture^{54,55}. Alternatively, in many applications where we are only interested in materials with optimal properties, ML methods based on Bayesian optimization (BO), active learning, and reinforcement learning (RL) are better suited^{56–58}. BO finds the optimal value from a black-box function involving a few parameters with minimum function evaluation, where each function evaluation is very expensive. Application of BO includes discovery of battery materials and 2D materials with maximum band gap and thermoelectric properties^{59,60}. In contrast, RL is more suitable for discrete problems that involve finding an optimal value and sequential decision-making under uncertainty in a huge parameter space. RL is heavily utilized in robotics, games, and machine translation. In material science, RL has been used to predict reaction pathways, optimal condition for chemical reactions, and design of molecules with desired chemical properties^{61–65}.

In this study, we have successfully applied RL to optimize the kirigami pattern in 2D MoS₂ and demonstrated that the maximum stretchability exceeds 45% for a system consisting of up to 6 cuts. We have used MD simulations to create the training data for the RL model. The location, length, and separation between kirigami cuts are variables in each simulation. Specifically, we have 13 different choices depending on the length and location of each cut in the material. In the RL model, placement of cuts inside the MoS₂ structures is posed as a sequential decision-making process. The goal of the RL agent is to place up to 6 cuts inside the MoS₂ and maximize the total reward at the end, which is proportional to the stretchability of the material. Further, the kirigami design strategies learned from the training data for 4 and 6 cuts, the RL agent successfully makes zero-shot predictions of kirigami structures with 8 and 10 cuts from a search space of a billion candidate structures without any training data. The MD calculations confirm that the stretchability of 8- and 10-cut structures exceed 40%.

RESULTS

MoS₂ kirigami state space

A schematic of the MoS₂ kirigami design and the corresponding stress–strain curves during uniaxial tensile simulation by MD are shown in Fig. 1a–c. Here, the stretchability of the material is defined as the maximum strain at which the material fails, i.e., the point of sudden drop in stress in the stress–strain curve of the material, Fig. 1c, which shows that the failure point of kirigami materials is very sensitive to the topology of the cut patterns. The stretchability of these systems is a function of the total number of cuts, cut length, and their location on the nanosheet. Figure 1a shows possible choices of kirigami cut patterns in a 20 nm × 30 nm MoS₂ nanosheet. The strategy for inserting a cut is that the system is first vertically partitioned into n rows of equal width, where each row may contain a horizontal cut. In these rows, a single cut can be placed at one of the 4 equally-spaced points chosen horizontally along the central axis of the row, and the length of these inserted cut is either 5, 10, or 15 nm. Furthermore, if the end point of the inserted cut extends beyond the MoS₂ nanosheet width (20 nm), it is wrapped back from the other side, see Fig. 1b (bottom). Thus, we have 13 different choices to place a cut in each row: 4 different location × 3 cut length + no cut, which makes the total number of possible kirigami structures to be 13^n , where n is the total number of vertical rows in an MoS₂ nanosheet. The state space of kirigami structures grows exponentially with the number of rows containing cuts. The number of possible structures is 28,561 for $n=4$, 4,826,809 for $n=6$ and 137,858,491,849 for $n=10$. It is impossible to search this vast space and find structures with high stretchability *via* experiments or MD simulation of random structures. Instead of the brute-force search to find optimal structures, we have created an RL model to find structures with high stretchability after training. Previously, Hanakata et al. used an alternative approach based on active learning and iterative search to design graphene kirigami with maximum stretchability⁵⁰. Here, we formulated the MoS₂ kirigami design as a reward-based continuous optimization problem in the framework of RL, where the goal of the RL agent is to learn an optimal policy that creates structures with high

stretchability. Training data for the RL model is generated by performing MD simulation of kirigami structures under uniaxial tensile stress for all possible 28,561 structures of $n = 4$, and for randomly sampled 70,000 structures from 13^5 structures in the case of $n = 6$. The details of the MD simulation are given in the “Methods” section.

RL framework for kirigami design

Since each MD simulation of a kirigami structure is expensive and takes ~ 6 h, it is not feasible to use MD simulations to provide feedback in the form of reward during the training of the RL agent. To address this problem, we designed our RL agent using model-based offline reinforcement learning, where the agent does have access to the environment during training and learns the optimal policy from randomly sampled suboptimal offline data from the environment⁶⁶. Figure 2a shows the schematic of our offline RL framework, which consists of an RL agent and a dynamic model of kirigami structure deformation as policy evaluator. Here, the dynamic model is constructed using deep convolutional neural network (CNN) and trained by MD simulation data to predict stretchability (S) of input kirigami structures, which is fed into the model as a $64 \times 64 \times 1$ binary tensor. Details of the dynamics model’s network architecture, conversion of kirigami structures into binary tensor, and its training and test accuracy are given in the “Methods” section and in Supplementary Discussion on Training and Validation Results of the Dynamic Model.

In the next stage of the RL framework consisting of a trained dynamic model using offline data, the RL agent formulates the construction of kirigami structures as a sequential decision-making process, where the decision about the placement of cut inside the nanosheet is made one by one in a raster scan fashion inspired by laser fabrication of stretchable kirigami device⁶⁷. Here, the synthesis of a single kirigami structure from a pure MoS_2 nanosheet (containing no cut) is called an episode of RL, where the length of each episode is kept fixed at $T = 4$ or 6 (equal to the total number of vertical rows), see Fig. 2b. During an episode at each time step t , the RL agent makes a decision about the location and length of the next cut in the t th row from the 13 available choices (action space). The input to the RL agent is the

intermediate structure from the $(t-1)$ th time step. Using the action proposed by the RL agent, a cut is created at the specified location of proposed length in the t th row, which serves as an input state to the RL agent at the next time step. At each time step, the dynamic model rewards the RL agent. The reward is 0 for $t < T$ and at the terminal state ($t = T$) it is proportional to the stretchability (S) of the material if it is above a certain threshold:

$$\text{Reward}(t) = \begin{cases} 0 & (t \leq T - 1 \text{ or } (t = T \text{ and } S < S_T)) \\ 0.2S & (t = T \text{ and } S \geq S_T) \end{cases} \quad (1)$$

The RL agent is a deep Q-network (DQN)⁵⁶, whose input is a $64 \times 64 \times 2$ tensor that summarizes the intermediate state (s_t) of the structure until time step t and outputs the state-action value function, $Q(s_t, a_t)$, for all the actions (a) we can take from s_t , which is 13 different cut choices to place in the t th row of the material. $Q(s_t, a_t)$ is an estimate of the total expected reward for each state-action pair from the t th time step to the end of the episode under a given policy. The goal of the DQN is to learn a policy (π) that maximizes the $Q(s_t, a_t)$ for each state-action pair (Eq. 2). Details of the network architecture, and the training of DQN are given in the “Methods” section and in the Supplementary Method and Supplementary Table 1.

$$Q(s_t, a_t) = \max_{\pi} E[r_t + \gamma r_{t+1} + \gamma^2 r_{t+2} + \dots + \gamma^{T-t-1} r_N] \\ \approx \max_{\pi} E_{s'} [r_t + \gamma \max_{a_{t+1}} Q(s_{t+1}, a_{t+1})] \quad (2)$$

Two different RL agents are constructed using the same DQN architecture for $n = 4$ and 6 such that they are trained to learn a policy that proposes structures with $S \geq S_T$ (threshold stretchability). The expected reward per episode during training for these models is given in Fig. 3a and Fig. S1. Here, the value of S_T in the reward function (Eq. 1) is a hyper-parameter and needs to be tuned during training to screen out structures with lower S values quickly but at the same time provide sufficient reward signals for learning. Since the objective of the RL agent is to maximize its total reward, the final optimal policy converges near the structure that has maximum S value as long as S_T is chosen carefully. We observe that smaller value of S_T leads to longer training time and lower expected reward, whereas large value of S_T causes no learning at all as most structures give zero reward that means no gradient signals from the reward function to the RL agent to update its parameters. The effect of different S_T values on training is discussed in the Supplementary Methods, and the best result in terms of expected reward is obtained with $S_T = 15\%$ and 20% for $n = 4$ and 6 , respectively, Supplementary Fig. 2. After training, the RL agents follow an ϵ -greedy policy to generate structures in each episode and choose an action at each time step that maximizes $Q(s_t, a_t)$ with $1 - \epsilon$ probability or a random action with $\epsilon = 0.1$ probability. Figure 3b shows the probability density function (PDF) of S for all the 28,561 structures with $n = 4$ cuts and PDF for S constructed by sampling 500 structures using a random search by the RL agent. In the case of a random search, each structure is chosen with equal probability. We observe in Fig. 3b that the true distribution of S is highly sparse, where 95% of the structures have $S < 15\%$. Thus, a random search cannot find structures with superior mechanical properties, which is also evident in Fig. 3b where most of the structures have $S < 15\%$. On the other hand, after training, the RL agent proposes structures with high S . In fact, most of the structures proposed by the RL agent have $S \geq 30\%$ for 80% of the time, which is closer to structures near S_{max} with maximum probability at S_{max} for $n = 4$. Figure 3c, d show several structures proposed by the RL agent and the random structures along with their S values.

In the case of $n = 6$, we have only used 70,000 structures ($\sim 1.45\%$ of all possible structures) to train the dynamic model first and then the RL agent for 12,000 episodes. After training, we use the RL agent to find structures with the highest value of S . Figure 4a shows the cumulative density function (CDF) of S for

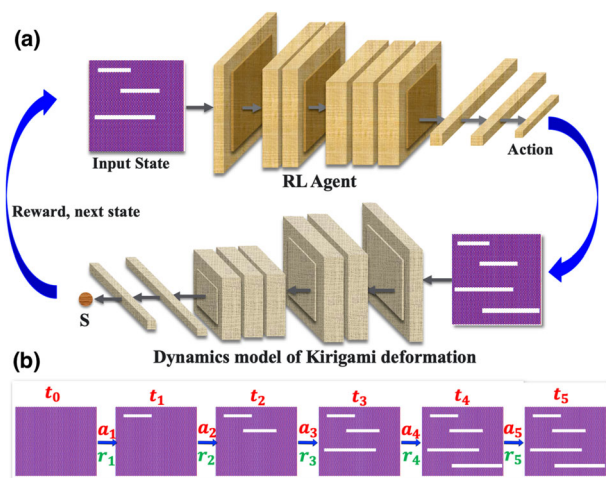


Fig. 2 RL agent for kirigami design. **a** A schematic of the RL agent and a dynamic model for the synthesis of MoS_2 kirigami structures. The dynamic model predicts the stretchability (S) of kirigami structure proposed by the RL agent and gives a reward to the RL agent which is used by agent to update its policy so as to maximize its total reward and creates structures with high S . **b** An episode of RL agent showing initial, final and intermediate structures and the associated actions taken and rewards received during the synthesis of MoS_2 kirigami structure.

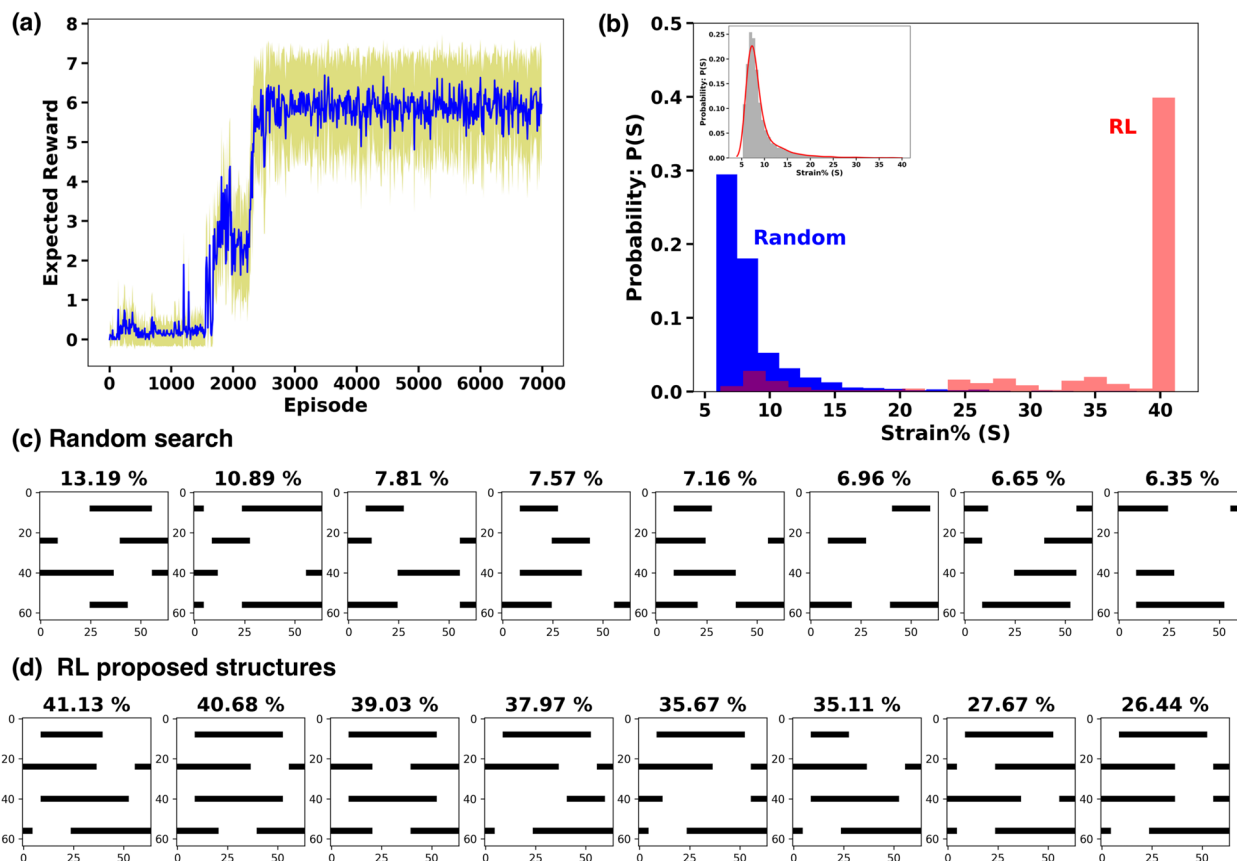


Fig. 3 Training and test results of RL agent with 4 cuts. **a** Expected reward (blue lines) and its variance (yellow) during the training of the RL agent for $n = 4$ with $S_T = 15\%$ during trainin. **b** Probability density function (PDF) constructed by sampling 500 structures using random search (blue) and RL agent (red). Here, the inset shows the true PDF of strain for all the possible structures for $n = 4$, i.e., a structure state space consisting of 13^4 structures. **c** Several samples generated during random search along with their stretchability and **d** shows few samples proposed by the RL agent along with their stretchability.

500 structures proposed by an RL agent, and Fig. 4b shows 18 structures with $S \geq 20\%$. These structures show ductile fracture and thus have high S (also shown in Fig. 4b), whereas structures with smaller $S \leq 10\%$ show brittle failure. Figure 4c, d shows the atomistic failure mechanism and the stress–strain curve for a system with low and high values of S , respectively. In the case of structures with low S (Fig. 4c), the local tensile stress near the cut edges increases with the applied strain. When any of these local stress concentrators exceeds a critical value (bond strength of Mo–S bond), it results in Mo–S bond breaking and rapid crack propagation originating from those high-stress regions. The crack propagation also causes a sudden drop in stress (Fig. 4c) and shows brittle fracture. In strained systems with high S values (Fig. 4f), we also observe compressive stresses in addition to local tensile stresses near the cut edges in intermediate regions of the cut both along the x and y directions (see Fig. 4g, h). Due to the absence of constraints along the z axis, these compressive stresses cause out-of-plane motion along the z axis in those regions, which releases the system’s stored energy and delays the tensile stress near the cut edges to reach Mo–S bond strength. These out-of-plane deformations of localized regions due to compressive stress varies with the cut design in the system and that directly affects the stretchability of the system (Fig. 4f). Supplementary information videos S1.mov and S2.mov show the failure mechanism of these two systems.

Zero-shot prediction of kirigami structures with 8–10 cuts

The sequence in which cuts are placed inside the material greatly affects the stretchability S . In fact, we can quantify the importance

of each cut in terms of its state-action value function, $Q(s_t, a_t)$. Figure 5a, b shows the principal component analysis (PCA) of the learned feature in the last layer of DQN for all the intermediate structures, and the associated action that has maximum $Q(s_t, a_t)$ at time $t = 2$ and 3 for $n = 6$, i.e., the second and third cut places inside the material. We observe that there are very few design choices available with respect to the placement of the 2nd and 3rd cuts inside the material, which will eventually lead to the creation of structures with high S values. Two such intermediate structures with high S value are also shown in Fig. 5a, b, whereas a proposed kirigami structure by the RL agent *via* these two intermediate structures is shown in Supplementary Fig. 3. Using $Q(s_t, a_t)$, we can also determine the best location and length of the very first cut inside the material. Figure 5c shows $Q(s_t, a_t)$ for all the 13 choices associated with the placement of the first cut, and a kirigami structure that is generated using them as first cut and then the follow-up RL policy for the remaining cut. We observe that S for generated structures is as low as 21%, reaches a maximum value of 47%, and is highly dependent upon the placement of the first cut.

The above analysis shows that the RL agent has learned strategies of kirigami design with high S value during training. At test time, this mechanistic understanding can be exploited by the RL agent to make zero-shot predictions of high S kirigami structures with $n = 8$ and 10 cuts, where the search space is above a billion candidates. Even though no training data is provided from $n \geq 8$, the RL agent can still propose structures containing more than 8 cuts by simply increasing the length of the RL episode at test time. We have used this strategy of longer



Fig. 4 MD validation of RL proposed kirigami structures with 6 cuts. **a** Cumulative density function (CDF) of the stretchability of systems proposed by RL for $n = 6$, i.e., a structure state space consisting of 13^6 structures. **b** Top 18 structures proposed by the RL agent for $n = 6$. Here, structures are sorted by their stretchability (S) and blue and red values above the structures are their S value computed by the dynamic model and MD simulation, respectively. Stress–strain curve, deformation mechanism, and local stress distribution, p_{xx} and p_{yy} , inside the materials before failure for system showing brittle failure and have low stretchability in (c–e) and for system showing ductile failure and have high stretchability in (f–h).

episode lengths to generate 8 structures each for $n = 8$ and 10 and performed MD simulations to compute their true S values, see Supplementary Discussion on zero-shot prediction for details. Here again, the RL agent successfully proposes structures with $S \geq 40\%$ for $n \geq 8$ as shown in Fig. 5d, e and Supplementary Figs. 8–9.

Comparison of RL with other baselines models. We have also compared the performance of our RL framework with two other machine learning methods of material design with optimal properties: active learning using iterative search similar to Hanakata et al. and conditional variational autoencoder (CVAE)^{50,68}. Details of these two models and their performance in finding structures with maximum S are given in the Supplementary Discussion under Baselines. We observe that CVAE fails to generate structures with $S \geq 35\%$, Supplementary Figs. 10–11. This happens because for both $n = 4$ and 6 most structures have $S \geq 20\%$. Thus, CVAE does not accurately learn the extremely sparse region of structures ($S \geq 30\%$) as very little training data is present there and after training only generates structures with $S \leq 30\%$. Active learning discovers the first structure with $S > 40.0\%$ and the best structure ($S = 47.3\%$) with 500 and 1300 structure evaluation respectively, whereas the RL agent also discovers the optimal structure with $S = 47.3\%$ around 1000 episode, which is equivalent to 1000 structure evaluations and uses the rest of the training episodes to construct its policy in high S regions, Supplementary

Figs. 12–13. We observe that the active learning approach is sample efficient in terms of finding the optimal structure but is also computationally more expensive in time due to its sequential nature and requires more memory for training as each step of active learning requires the evaluation of entire 4.8 million structures for $n = 6$. In comparison, RL requires less memory with lower training time since it does not evaluate all 4.8 million structures at each episode and beside learning the optimal policy also provides mechanistic insight into the placement of cuts in terms of their state-active value function as described in the previous section. These mechanistic insights of kirigami design are difficult to infer from the active learning approach. Further, the RL agent can exploit these learned strategies for completely different objectives such as completion of partially created kirigami structures (see the Supplementary Discussion on partial task completion) and zero-shot prediction of optimal structures from a billion candidate search space ($n \geq 8$) that has data distribution different from the training data ($n = 6$); see Supplementary Figs. 8–9 and Supplementary Discussion on zero-shot prediction.

DISCUSSION

In summary, we have shown that RL can efficiently generate 2D kirigami structures with high stretchability from an extremely large search space consisting of millions of structures. Further,

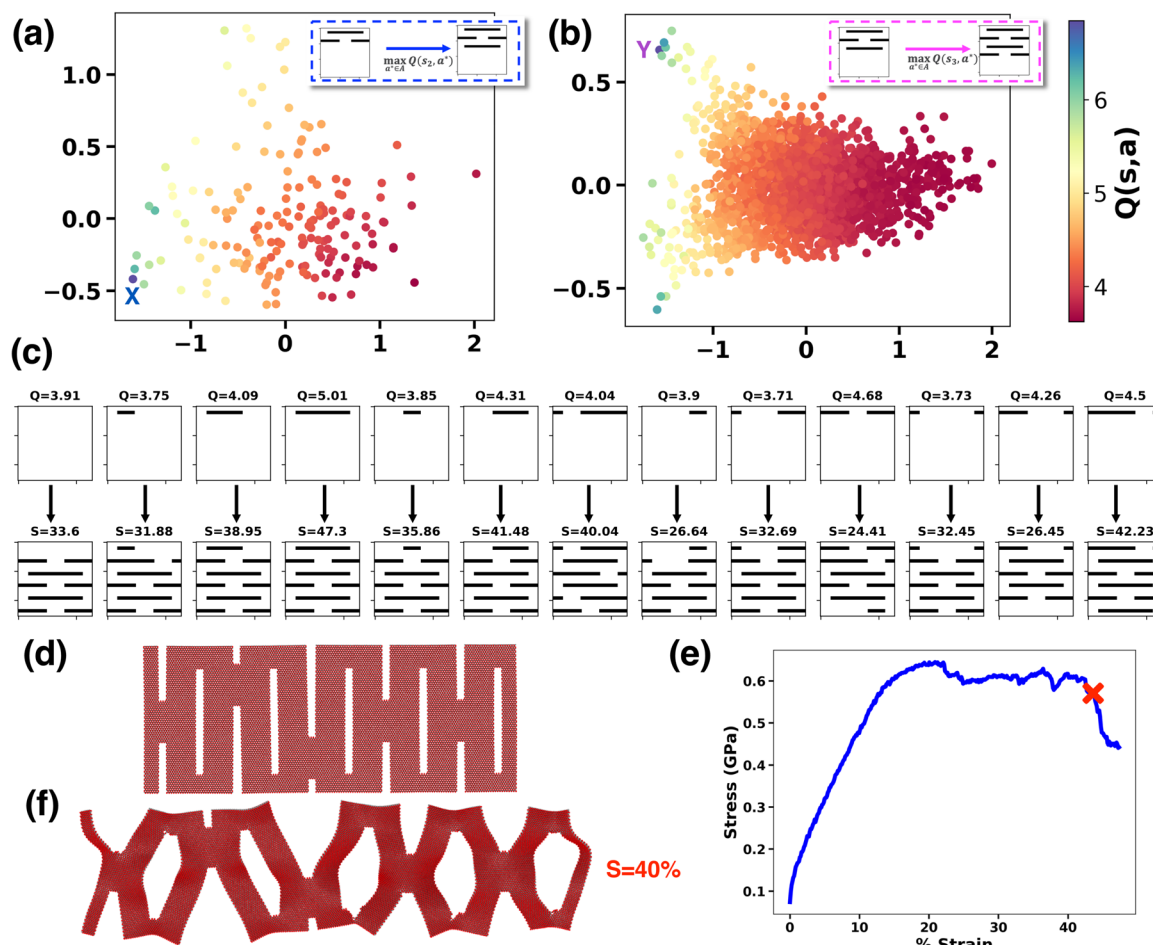


Fig. 5 Mechanistic insight of kirigami design learned by RL agent and zero-shot prediction of 10-cut structures. **a, b** Principal component analysis (PCA) of the learned features of the intermediate states at time $t = 2$ (**a**) and time $t = 3$ (**b**) by the RL agent that has maximum $Q(s, a)$, where each of these states are colored by that $Q(s_t, a)$ value. State-action pair that has maximum $Q(s_t, a)$ is also highlighted in (**a**) and (**b**). **c** All possible cut choices to place a cut inside the material at $t = 1$ and the associated $Q(s, a)$ for that state-action pair. It also shows a proposed structure along with its stretchability by the RL agent for each of the associated action taken at $t = 1$ and then afterward following the RL policy. **d** zero-shot prediction of kirigami structures with $n = 10$. **e** Stress–strain curve computed by MD simulation and **f** atomic structure of the system just above 40% strain (red cross).

an RL agent provides insight into the placement and length of cuts inside the material that will eventually create structure with superior mechanical properties in terms of its state-action value function, $Q(s_t, a_t)$. In general, the optimal structure design of materials is high dimensional, involves sequential decision-making, and sometime discrete in nature. For example, designing crystal structure with high hardness involves the selection of very specific set of elements and their placement inside the crystal at specific locations. Known ML methods such as regression and Bayesian optimization are difficult to apply because we cannot take the gradient of the model due to its discrete nature of search space, and manual evaluation of each structures becomes computationally expensive and memory-bound when the size of the search space increases to billions of structures. Previously, ML models for optimal structure search in a discrete space used active learning with iterative search by Hanakata et al., whereas Bombarelli et al. have mapped the discrete representations of molecules on a continuous latent space using variational autoencoder and then used that latent space representation for optimal molecule design^{50,69}. Here, our RL formulation provides another valuable technique to convert these challenging problems into a reward-based continuous optimization, which directly proposes not only the best structures after training but can also exploit its learned

mechanistic insight to generate structures that are outside the training data distribution. Thus, RL can be used to design a wide class of metamaterials where black-box function optimization is difficult to apply.

METHODS

Molecular dynamics simulation

The 2H crystal structure of MoS_2 is used here to create kirigami structures by removing Mo and S atoms from a 2D- MoS_2 nanosheet in the 2H phase such that the stoichiometry of the system is maintained. The width of each cut is 1 nm. Each system has lateral dimensions $20 \text{ nm} \times 30 \text{ nm}$, containing $\sim 22,000$ atoms. We relax each kirigami structure by minimizing the energy configuration using the conjugate gradient method. After that, we relax the system at a temperature of 100 K for 200 ps under NVE ensemble. The system is then subjected to uniaxial tension, following loading and relaxation phases. During the loading phase, each system is homogeneously expanded along the y -axis for 1 ns, followed by the relaxation phase, where the entire system is relaxed for another 5 ps in the NVE ensemble. The whole cycle is repeated to apply larger strains. This procedure amounts to an effective strain rate of $4 \times 10^8 \text{ s}^{-1}$. Periodic boundary condition is applied only along the y -axis, and the tensile stress of the material is computed during the relaxation phase. Details of the stress calculation and interaction potential are given in the Supplementary Notes.

Dynamics model for mechanical deformation

We use a CNN network consisting of three convolutional layers of dimensions $32 \times 32 \times 16$, $16 \times 16 \times 32$, and $32 \times 32 \times 64$ followed by three fully-connected layers of dimensions 4096×1024 , 1024×512 , and 512×1 . The input is a $64 \times 64 \times 1$ binary tensor representing a kirigami structure and the output is stretchability S . Here, the input binary tensor is created from the 2D atomic nanosheets of size $20 \text{ nm} \times 30 \text{ nm}$ by first partitioning the nanosheets into bins of size $\frac{20}{64} \text{ nm} \times \frac{30}{64} \text{ nm} = 0.3125 \text{ nm} \times 0.46875 \text{ nm}$ and then each bin that does not contain any atom is given a value of 1 otherwise 0. In each convolutional layer, we use padding and strides equal to 1×1 and the kernel size is 3×3 . ReLU is used as a non-linear activation function between hidden layers. We use training data consisting of 8000 and 70,000 structures for $n = 4$ and 6, respectively. During training, we use Adam optimizer with a learning rate of 0.005 and train the model for 100 epochs with a batch size of 64 and the dropout probability of 0.5 in between the second to last fully-connected layers. The training and test accuracy along with error analysis of the model are given in the Supplementary Figs. 4–7 of the Supplementary Discussion.

Reinforcement learning agent (DQN Network)

DQN is another CNN network with network architecture similar to that of the dynamic model. However, the input to the DQN network is a $64 \times 64 \times 2$ tensor representing intermediate structure state at time t and the output $Q(s_t, a_t)$ for all the 13 actions for state s_t . Here, the first dimension of the input tensor is a binary tensor of the structure generated at time step $t-1$, and the second dimension of the tensor is y th coordinate of the cut added in time step $t-1$. During training, we use experienced replay with replay memory of size 10,000. The training algorithm for the DQN is given in Supplementary Table 1. We trained two different RL agents for $n = 4$ and 6. The expected reward during training for these models is given in Fig. 3a and Supplementary Fig. 1.

DATA AVAILABILITY

The example training dataset of MoS₂ kirigami structures used for training is distributed along with the code.

CODE AVAILABILITY

The PyTorch code used in this study can be found at https://github.com/rajak7/RL_kirigami.git.

Received: 6 August 2020; Accepted: 7 May 2021;

Published online: 09 July 2021

REFERENCES

- Voiry, D., Yang, J. & Chhowalla, M. Recent strategies for improving the catalytic activity of 2D TMD nanosheets toward the hydrogen evolution reaction. *Adv. Mater.* **28**, 6197–6206 (2016).
- Lin, H. et al. Chalcogenide glass-on-graphene photonics. *Nat. Photonics* **11**, 798–805 (2017).
- Xie, L. & Cui, X. Manipulating spin-polarized photocurrents in 2D transition metal dichalcogenides. *Proc. Natl Acad. Sci. USA* **113**, 3746 (2016).
- Tang, H. et al. Growth of polypyrrole ultrathin films on MoS₂ monolayers as high-performance supercapacitor electrodes. *Adv. Mater.* **27**, 1117–1123 (2015).
- Le, D., Rawal, T. B. & Rahman, T. S. Single-layer MoS₂ with sulfur vacancies: structure and catalytic application. *J. Phys. Chem. C* **118**, 5346–5351 (2014).
- Apte, A. et al. 2D electrets of ultrathin MoO₂ with apparent piezoelectricity. *Adv. Mater.* **32**, 2000006 (2020).
- Radisavljevic, B., Radenovic, A., Brivio, J., Giacometti, V. & Kis, A. Single-layer MoS₂ transistors. *Nat. Nanotechnol.* **6**, 147–150 (2011).
- Bertolazzi, S., Brivio, J. & Kis, A. Stretching and breaking of ultrathin MoS₂. *ACS Nano* **5**, 9703–9709 (2011).
- Komsa, H.-P. & Krashennikov, A. V. Two-dimensional transition metal dichalcogenide alloys: stability and electronic properties. *J. Phys. Chem. Lett.* **3**, 3652–3656 (2012).
- Conley, H. J. et al. Bandgap engineering of strained monolayer and bilayer MoS₂. *Nano Lett.* **13**, 3626–3630 (2013).
- Lee, G.-H. et al. Flexible and transparent MoS₂ field-effect transistors on hexagonal boron nitride-graphene heterostructures. *ACS Nano* **7**, 7931–7936 (2013).

- Wang, S. et al. Atomically sharp crack tips in monolayer MoS₂ and their enhanced toughness by vacancy defects. *ACS Nano* **10**, 9831–9839 (2016).
- Cao, Z., Harb, M., Lardhi, S. & Cavallo, L. Impact of interfacial defects on the properties of monolayer transition metal dichalcogenide lateral heterojunctions. *J. Phys. Chem. Lett.* **8**, 1664–1669 (2017).
- Guzman, D. M. & Strachan, A. Role of strain on electronic and mechanical response of semiconducting transition-metal dichalcogenide monolayers: an ab-initio study. *J. Appl. Phys.* **115**, 243701 (2014).
- Gong, Y. et al. Vertical and in-plane heterostructures from WS₂/MoS₂ monolayers. *Nat. Mater.* **13**, 1135–1142 (2014).
- Mann, J. et al. 2-Dimensional transition metal dichalcogenides with tunable direct band gaps: MoS_{2(1-x)}Se_{2x} monolayers. *Adv. Mater.* **26**, 1399–1404 (2014).
- Apte, A. et al. Structural phase transformation in strained monolayer MoWSe₂ alloy. *ACS Nano* **12**, 3468–3476 (2018).
- Wei, K., Peng, Y., Qu, Z., Pei, Y. & Fang, D. A cellular metastructure incorporating coupled negative thermal expansion and negative Poisson's ratio. *Int. J. Solids Struct.* **150**, 255–267 (2018).
- Nicolaou, Z. G. & Motter, A. E. Mechanical metamaterials with negative compressibility transitions. *Nat. Mater.* **11**, 608–613 (2012).
- Guo, K. & Buehler, M. J. Nature's way: hierarchical strengthening through weakness. *Matter* **1**, 302–303 (2019).
- Xu, W. et al. Reversible MoS₂ origami with spatially resolved and reconfigurable photosensitivity. *Nano Lett.* **19**, 7941–7949 (2019).
- Rafsanjani, A. & Bertoldi, K. Buckling-induced kirigami. *Phys. Rev. Lett.* **118**, 084301 (2017).
- Hwang, D.-G. & Bartlett, M. D. Tunable mechanical metamaterials through hybrid kirigami structures. *Sci. Rep.* **8**, 3378 (2018).
- Shyu, T. C. et al. A kirigami approach to engineering elasticity in nanocomposites through patterned defects. *Nat. Mater.* **14**, 785–789 (2015).
- Isobe, M. & Okumura, K. Initial rigid response and softening transition of highly stretchable kirigami sheet materials. *Sci. Rep.* **6**, 24758 (2016).
- Zhang, P. et al. Fracture toughness of graphene. *Nat. Commun.* **5**, 3782 (2014).
- Blees, M. K. et al. Graphene kirigami. *Nature* **524**, 204–207 (2015).
- Xu, L., Shyu, T. C. & Kotov, N. A. Origami and kirigami nanocomposites. *ACS Nano* **11**, 7587–7599 (2017).
- Choi, S., Han, S. I., Kim, D., Hyeon, T. & Kim, D.-H. High-performance stretchable conductive nanocomposites: materials, processes, and device applications. *Chem. Soc. Rev.* **48**, 1566–1595 (2019).
- Wang, C., Wang, C., Huang, Z. & Xu, S. Materials and structures toward soft electronics. *Adv. Mater.* **30**, 1801368 (2018).
- Liu, Z. et al. Nano-kirigami with giant optical chirality. *Sci. Adv.* **4**, eaat4436 (2018).
- Jing, L. et al. Kirigami metamaterials for reconfigurable toroidal circular dichroism. *NPG Asia Mater.* **10**, 888–898 (2018).
- Hanakata, P. Z., Qi, Z., Campbell, D. K. & Park, H. S. Highly stretchable MoS₂ kirigami. *Nanoscale* **8**, 458–463 (2016).
- Wang, B., Nakano, A., Vashishta, P. D. & Kalia, R. K. Nanoindentation on monolayer MoS₂ kirigami. *ACS Omega* **4**, 9952–9956 (2019).
- Zhu, S., Huang, Y. & Li, T. Extremely compliant and highly stretchable patterned graphene. *Appl. Phys. Lett.* **104**, 173103 (2014).
- Ramprasad, R., Batra, R., Piliand, G., Mannodi-Kanakkithodi, A. & Kim, C. Machine learning in materials informatics: recent applications and prospects. *npj Comput. Mater.* **3**, 54 (2017).
- Butler, K. T., Davies, D. W., Cartwright, H., Isayev, O. & Walsh, A. Machine learning for molecular and materials science. *Nature* **559**, 547–555 (2018).
- Huan, T. D. et al. A universal strategy for the creation of machine learning-based atomistic force fields. *npj Comput. Mater.* **3**, 37 (2017).
- Rajak, P., Krishnamoorthy, A., Nakano, A., Vashishta, P. & Kalia, R. Structural phase transitions in a MoWSe₂ monolayer: molecular dynamics simulations and variational autoencoder analysis. *Phys. Rev. B* **100**, 014108 (2019).
- Cubuk, E. D. et al. Structure-property relationships from universal signatures of plasticity in disordered solids. *Science* **358**, 1033 (2017).
- Yu, C.-H., Qin, Z. & Buehler, M. J. Artificial intelligence design algorithm for nanocomposites optimized for shear crack resistance. *Nano Futures* **3**, 035001 (2019).
- Jain, A. et al. Commentary: The Materials Project: a materials genome approach to accelerating materials innovation. *APL Mater.* **1**, 011002 (2013).
- Behler, J. Constructing high-dimensional neural network potentials: a tutorial review. *Int. J. Quantum Chem.* **115**, 1032–1050 (2015).
- Mailoa, J. P. et al. A fast neural network approach for direct covariant forces prediction in complex multi-element extended systems. *Nat. Mach. Intell.* **1**, 471–479 (2019).
- Bartók, A. P., Payne, M. C., Kondor, R. & Csányi, G. Gaussian approximation potentials: the accuracy of quantum mechanics, without the electrons. *Phys. Rev. Lett.* **104**, 136403 (2010).

46. Mocanu, F. C. et al. Modeling the phase-change memory material, $\text{Ge}_2\text{Sb}_2\text{Te}_5$, with a machine-learned interatomic potential. *J. Phys. Chem. B* **122**, 8998–9006 (2018).
47. Xie, T. & Grossman, J. C. Crystal graph convolutional neural networks for an accurate and interpretable prediction of material properties. *Phys. Rev. Lett.* **120**, 145301 (2018).
48. Mannodi-Kanakkithodi, A., Pilania, G., Huan, T. D., Lookman, T. & Ramprasad, R. Machine learning strategy for accelerated design of polymer dielectrics. *Sci. Rep.* **6**, 20952 (2016).
49. Gu, T., Lu, W., Bao, X. & Chen, N. Using support vector regression for the prediction of the band gap and melting point of binary and ternary compound semiconductors. *Solid State Sci.* **8**, 129–136 (2006).
50. Hanakata, P. Z., Cubuk, E. D., Campbell, D. K. & Park, H. S. Accelerated search and design of stretchable graphene kirigami using machine learning. *Phys. Rev. Lett.* **121**, 255304 (2018).
51. Ren, F. et al. Accelerated discovery of metallic glasses through iteration of machine learning and high-throughput experiments. *Sci. Adv.* **4**, eaq1566 (2018).
52. Ubaru, S., Międlar, A., Saad, Y. & Chelikowsky, J. R. Formation enthalpies for transition metal alloys using machine learning. *Phys. Rev. B* **95**, 214102 (2017).
53. Bernstein, N. et al. Quantifying chemical structure and machine-learned atomic energies in amorphous and liquid. *Silicon Angew. Chem. Int. Ed.* **58**, 7057–7061 (2019).
54. Gu, G. X., Chen, C.-T. & Buehler, M. J. De novo composite design based on machine learning algorithm. *Extrem. Mech. Lett.* **18**, 19–28 (2018).
55. Hsu, Y.-C., Yu, C.-H. & Buehler, M. Using deep learning to predict fracture patterns in crystalline solids. *Matter* **3**, 197–211 (2020).
56. Mnih, V. et al. Human-level control through deep reinforcement learning. *Nature* **518**, 529–533 (2015).
57. Shahriari, B., Swersky, K., Wang, Z., Adams, R. P. & Freitas, N. D. Taking the human out of the loop: a review of Bayesian optimization. *Proc. IEEE* **104**, 148–175 (2016).
58. Sutton, R. S. & Barto, A. G. *Reinforcement Learning: An Introduction* 2nd edn (MIT Press, in the press) (2018).
59. Bassman, L. et al. Active learning for accelerated design of layered materials. *npj Comput. Mater.* **4**, 74 (2018).
60. Attia, P. M. et al. Closed-loop optimization of fast-charging protocols for batteries with machine learning. *Nature* **578**, 397–402 (2020).
61. Sanchez-Lengeling, B. & Aspuru-Guzik, A. Inverse molecular design using machine learning: Generative models for matter engineering. *Science* **361**, 360 (2018).
62. Popova, M., Isayev, O. & Tropsha, A. Deep reinforcement learning for de novo drug design. *Sci. Adv.* **4**, eaap7885 (2018).
63. Zhou, Z., Li, X. & Zare, R. N. Optimizing chemical reactions with deep reinforcement learning. *ACS Cent. Sci.* **3**, 1337–1344 (2017).
64. Segler, M. H. S., Preuss, M. & Waller, M. P. Planning chemical syntheses with deep neural networks and symbolic AI. *Nature* **555**, 604–610 (2018).
65. Kearnes, S., Li, L. & Riley, P. Decoding molecular graph embeddings with reinforcement learning. Preprint at <https://arxiv.org/abs/1904.08915> (2019).
66. Levine, S., Kumar, A., Tucker, G. & Fu, J. Offline reinforcement learning: tutorial, review, and perspectives on open problems. Preprint at <https://arxiv.org/abs/2005.01643> (2020).
67. Groeger, D. & Steimle, J. *LASEC: Instant Fabrication of Stretchable Circuits Using a Laser Cutter* (2019).
68. Doersch, C. Tutorial on variational autoencoders. Preprint at <https://arxiv.org/abs/1606.05908> (2016).
69. Gómez-Bombarelli, R. et al. Automatic chemical design using a data-driven continuous representation of molecules. *ACS Cent. Sci.* **4**, 268–276 (2018).

ACKNOWLEDGEMENTS

This work was supported by National Science Foundation, Future Manufacturing Program, Award 2036359. This research was partly supported by Aurora Early Science programs and used resources of the Argonne Leadership Computing Facility, which is a DOE Office of Science User Facility supported under Contract DE-AC02-06CH11357. Computations were performed at the Argonne Leadership Computing Facility under the DOE INCITE and Aurora Early Science programs and at the Center for Advanced Research Computing of the University of Southern California. We would like to thank one of the reviewers for asking us to examine if RL can be used to propose a high-stretchability kirigami structure with 10 cuts, which led to zero-shot predictions for 8- and 10-cut structures that have stretchability above 40%.

AUTHOR CONTRIBUTIONS

P.R. analyzed and developed the reinforcement learning model, and P.R. and B.W. performed the MD simulations and analysis of MD data. P.R., K.N., R.K.K., A.N., and P.V. designed the RL assisted Kirigami design. R.K.K., A.N., and P.V. advised on this research work. All authors contributed to analyzing the results and writing the manuscript.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41524-021-00572-y>.

Correspondence and requests for materials should be addressed to P.R. or P.V.

Reprints and permission information is available at <http://www.nature.com/reprints>

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



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021