Structural phase transitions in a $MoWSe_2$ monolayer: Molecular dynamics simulations and variational autoencoder analysis

Pankaj Rajak, Aravind Krishnamoorthy, Aiichiro Nakano, Priya Vashishta, and Rajiv Kalia*

Collaboratory for Advanced Computing and Simulations, Department of Physics & Astronomy, Department of Computer Science, Department of Chemical Engineering & Materials Science, Department of Biological Sciences, University of Southern California, Los Angeles, California 90089-0242, USA

(Received 20 February 2019; published 30 July 2019)

Electrical and optoelectronic properties of two-dimensional (2D) transition metal dichalcogenides (TMDCs) can be tuned by exploiting their structural phase transitions. Here semiconducting (2H) to metallic (1T) phase transition is investigated in a strained $MoWSe_2$ monolayer using molecular dynamics (MD) simulations. Novel intermediate structures called α and β are found between the 2H and 1T phases. These intermediate structures are similar to those observed in a 2D MoS_2 by scanning transmission electron microscopy. A deep generative model, namely the variational autoencoder (VAE) trained by MD data, is used to generate novel heterostructures with α and β interfaces. Quantum simulations based on density functional theory show that these heterostructures are stable and suitable for novel nanoelectronics applications.

DOI: 10.1103/PhysRevB.100.014108

I. INTRODUCTION

A two-dimensional $MoWSe_2$ system was recently synthesized with chemical vapor deposition (CVD) and structural changes owing to mechanical deformation were examined with an *in situ* setup of a Raman spectrophotometer combined with a nanoindenter and scanning probe microscope [1]. The experiment reveals that the heterostructure contains randomly distributed nanoscale patches of WSe_2 in a matrix of $MoSe_2$. Under ambient conditions, Raman spectra of the system indicate 2H structure in the entire sample of the $MoSe_2$ matrix containing WSe_2 patches. The 2H monolayer has trigonal prismatic (D_{3h}) symmetry with an ABA atomic stacking sequence, and it is a direct band-gap semiconductor [2]. Under strain, cracks develop in the alloy and the Raman spectra show not only features corresponding to the 2H structure but also new features indicative of the 1T structure. The metallic 1T phase possesses octahedral (O_h) symmetry and an ABC atomic stacking sequence.

In this paper we examine atomistic mechanisms of polymorphism in a single, strained monolayer of $MoWSe_2$ using molecular dynamics (MD) simulations [3]. The system consists of randomly distributed WSe_2 patches in a $MoSe_2$ matrix. The system size is $0.5\,\mu\text{m} \times 0.5\,\mu\text{m}$. The $MoWSe_2$ sample is precracked and subjected to mode-1 strain. With the onset of crack propagation under an applied strain, we observe a 2H $\rightarrow$ 1T structural transformation around the crack front. The MD simulations reveal novel structures between 2H and 1T phases. These structures include a stable precursor structure called α which consists of three-to-four zigzag chains of Mo-Se or W-Se atoms. The nearest-neighbor (nn) Mo-Mo or W-W separation in these chains is constricted by about 14% relative to the nn separation in the 2H structure. Between 2H and 1T phases, the simulation also reveals another novel structure called β which has four coordinated Se atoms instead of three in the α phase [4]. These results are consistent with scanning tunneling electron microscopy (STEM) data, which reveal α and β as boundary-layer structures in the 2H $\rightarrow$ 1T transition in MoS_2.

In recent years, supervised and unsupervised machine learning models have shown tremendous success in material science for efficient discovery of new materials with desired properties [5–9]. Data-driven force field models based on several different approaches [10,11], including neural network and kernel ridge regression, have shown quantum-mechanical density functional theory (DFT)-level accuracy in the prediction of atomic forces and energies [12–14]. Similarly, unsupervised learning algorithms, namely deep generative models such as autoencoders and adversarial networks, have been applied to *in silico* synthesis of molecular design with desired properties [15,16]. Here we use a deep generative model based on a variational autoencoder (VAE) to learn the relationship among the complex phases that are generated during the fracture of $MoWSe_2$ heterostructure. With enough training examples, VAE can learn probability distribution functions and sample structures from those distributions. The VAE trained by the MD data is used to synthesize two types of devices involving 2H/α and 2H/β interfaces. Quantum simulations based on DFT indicate that these semiconductor-metal heterostructures are stable and suitable for bottom-up fabrication of Schottky diodes, nanoscale transistors [17], and quantum leads for one-dimensional charge and spin transport [18].

II. RESULTS AND DISCUSSION

Crack propagation and the resulting changes in the structure of $MoWSe_2$ are studied for two orientations of the initial 2H lattice relative to the direction of the applied strain. Here the primary focus is on the (11) orientation in which one

*Corresponding author: rkalia@usc.edu

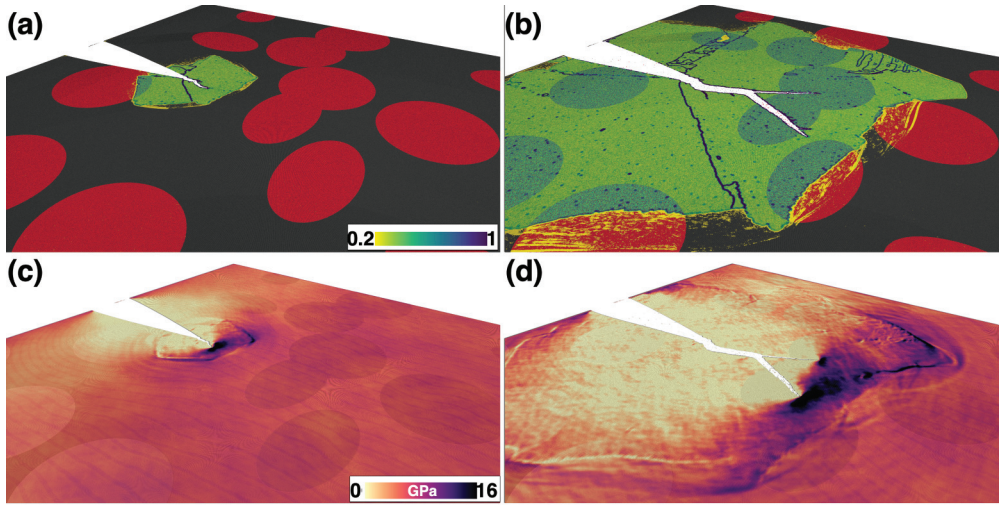

FIG. 1. Structural phase transformation induced by crack propagation in a 0.5 μm × 0.5 μm $MoWSe_2$ heterostructure. Snapshot showing local strain analysis in 2H → 1T transformation (a) at the onset of crack propagation and (b) after fracture. The post-fracture heterostructure consists of multiple crack branches, grain boundaries, and strain-induced transformed regions consisting of 1T structure and defects. Black and red regions in (a) and (b) represent $MoSe_2$ and WSe_2 in the 2H crystal structure. Light and dark green regions in (a) and (b) represent the 1T crystal structure in $MoSe_2$ and WSe_2, respectively. Dark blue regions indicate defects in grain boundaries. The color bar in (a) and (b) indicates local strain in the system [7]. (c) and (d) Snapshots of the stress distribution (σ_{yy}) in the material at the onset of crack propagation and after fracture, respectively.

of the Mo-Se bonds in the 2H structure is parallel to the direction of the applied strain, namely the y axis. Results for the (10) orientation, where one of the Mo-Se bonds in the 2H lattice is perpendicular to the y axis, are presented in the Supplemental Material [19]. The precrack begins to propagate when the applied strain in the (11) orientation is increased to 2.5%, which is consistent with the experimental observation of the response of $MoWSe_2$ heterostructure under strain [1]. Figure 1(a) presents a snapshot of the primary crack at the critical strain corresponding to the onset of crack propagation. The tip of the primary crack is in the 2H structure of the $MoSe_2$ matrix before the onset of crack propagation. The crack tip velocity is 2.1 km/s, which is 65% of the Rayleigh wave speed in $MoSe_2$ (see Sec. 4 in the Supplemental Material [19]).

Crack propagation in $MoSe_2$ is accompanied by a rapid 2H → 1T phase transformation in which the 2H/1T phase boundary propagates at an average speed of 2.6 ± 0.5 km/s. The crack and the 2H/1T phase boundary slow down slightly when they enter a nanoscale patch of WSe_2. The slowing down is caused by the nucleation of new intermediate structures and defects at the $MoSe_2/WSe_2$ interface. Experimental studies of strain-induced deformation and fracture reveal that 2H → 1T phase transformation and the presence of WSe_2 patches combine to enhance toughening of the $MoSe_2$ matrix [1].

The primary crack bifurcates into two secondary crack branches at an applied strain of 2.5%, see Fig. 1(b). This bifurcation occurs with the formation of new structures and point defects inside the 2H/1T phase boundary. The speed of the primary crack at the onset of bifurcation is 65% of the Rayleigh wave speed. This has also been observed in a MD simulation study of fracture along the (10) orientation in graphene [20,21]. We observe that 2H/1T phase transformation along each secondary crack branch also propagates at a higher speed than the secondary cracks. Secondary cracks also slow down when they enter WSe_2 patches.

Figure 1(c) shows stress distribution in $MoWSe_2$ at the onset of primary crack propagation (see movie S1 in the Supplemental Material [19]). The local fracture toughness calculated from the stress at the crack front and the corresponding crack length is close to the calculated K_{1C} (1.54 MPa $\sqrt{\text{m}}$) of $MoSe_2$ in the 2H phase [22]. Figure 1(d) shows stress distribution in $MoWSe_2$ when secondary cracks nucleate from the primary crack in the $MoSe_2$ matrix. The secondary cracks are inclined at an angle of $50 \pm 5^\circ$ (see Fig. 1(b) and the movie S2 in the Supplemental Material [19]), which is due to the fact that the maximum stress at the primary crack front causing Mo-Se bond breaking is along these two directions. We observe closure of some of the secondary cracks, which reflects toughening of the $MoWSe_2$ heterostructure.

Figure 2 shows intermediate structures between 2H and 1T phases in the $MoSe_2$ matrix and WSe_2 nanoscale patches. The α phase is always a precursor structure in the 2H → 1T transition in both $MoSe_2$ and WSe_2. The key signature of the α structure is zigzag chains consisting of three-to-four Mo-Se or W-Se bonds in which Mo atoms are five coordinated and Mo-Mo and W-W nn distances are constricted relative to their respective nn separation in unstrained $MoSe_2$ and WSe_2. The α structure between 2H and 1T phases has been observed in STEM studies of MoS_2 [4].

We also observe another intermediate structure called β shown in Fig. 2(b) and the movie S3 in the Supplemental Material [19]. Se atoms in the β structure are fourfold coordinated. This is consistent with STEM observations of the 2H → 1T phase transformation in MoS_2 [4]. The 2H/1T transformation and the intermediate α and β structures in the STEM experiment arise from elevated temperatures (400−700 °C). In our simulation, the local temperature near the 2H → 1T transformation front reaches 900 °C, which is

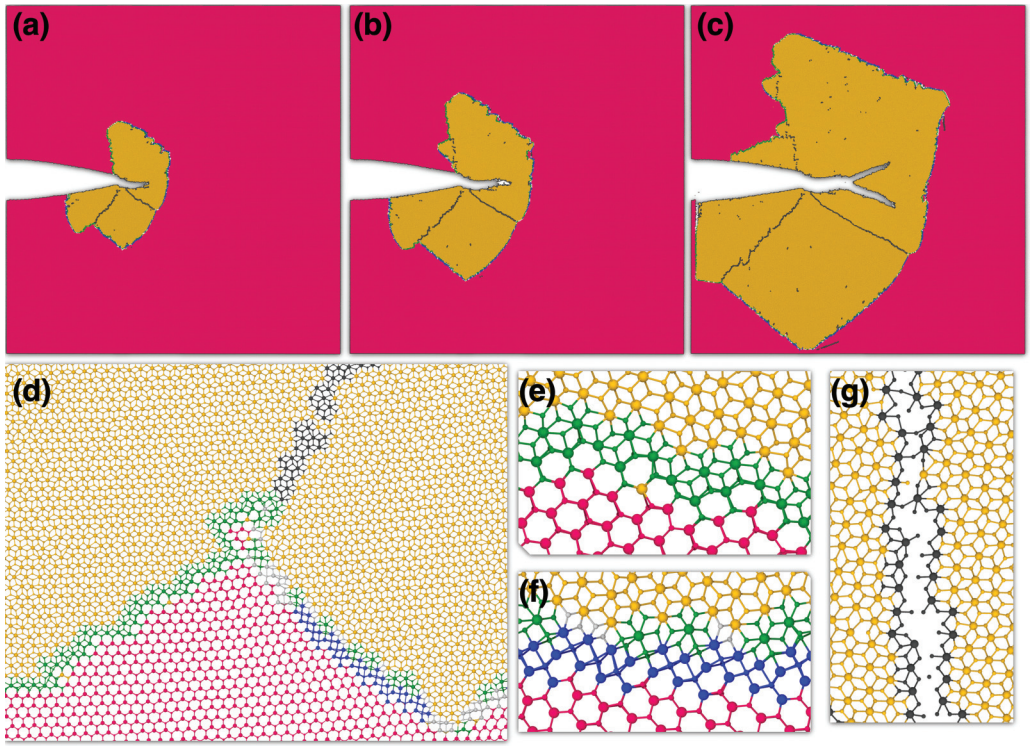

FIG. 2. Dynamics of interfacial structures at the 2H/1T boundary. (a)–(c) The evolution of interfacial regions between 2H → 1T as the crack propagates though the $MoWSe_2$ heterostructure. Here the magenta area is the 2H structure and yellow regions are fracture-induced 1T structure, which also contains grain boundaries and defects (shown in black). (d) An enlarged view of the interfacial regions consisting of α (green) and β (blue) structures. Interfacial regions without α and β structures are shown in gray. (e)–(g) Atomic views of α, β, and defects in the crack, respectively. In the α structure, each Mo/W is either 5 or 7 coordinated and the neighboring Se is 3 coordinated. In the β structure, each Se is 4 coordinated and the neighboring Mo is 6 coordinated.

higher than the temperature in strained $MoWSe_2$ regions far removed from the process zone.

We have further analyzed the relationship among the complex phases that are generated during the fracture of $MoWSe_2$ heterostructure using a deep generative model based on VAE. The input data for VAE are atomic trajectories generated by MD simulations. We examine transition pathways between different phases of $MoWSe_2$ by training the VAE using 40 000 examples consisting of uniform distributions of 2H, 1T, α, β, defects, and crack branches from the MD simulation. The model is trained for 200 epochs. The reconstruction loss and KL divergence loss converge as shown in Fig. S4 in the Supplemental Material [19]. Details of the VAE architecture and the training process are given in Sec. III and the Supplemental Material [19].

The VAE model, tensor representation of the MD data, and visualization of the latent space after training are shown schematically in Fig. 3(a). The latent space dimension (10) is significantly smaller than the dimension of the original input data $64 \times 64 \times 3 = 12288$. Hence, to reconstruct the original input data, VAE must learn the most important statistical features of the data representation in the latent dimension. The t-distributed stochastic neighbor embedding (t-SNE) visualization after training shows a clear separation of 2H, 1T, α, β, and defects in the latent space, see Fig. 3(a) [23]. Figure 3(b) shows the principal component analysis (PCA) visualization of the latent space, which unlike t-SNE captures the global relationship among various clusters present in the data. The PCA visualization shows that the encoding of α, β, and defect structures lies between the encoding of 2H and 1T structures. This leads to the inference that VAE can correctly identify α and β phases between 2H → 1T transformation.

We further examine the phase transformation pathways using latent space algebra [24]. We first calculate the mean of 2H (1T) structures in the latent space, i.e., μ_{2H} (μ_{1T}), by randomly selecting 20000 2H (1T) patches and averaging their z value. Subsequently, we sample structures from the decoder where the input z value is computed using linear mixing:

$$z = a\mu_{2H} + (1 - a)\mu_{1T}. \tag{1}$$

The value of a controls the relative sizes of 2H and 1T structures in the sample. Figures 3(b) and 3(c) show the transformation pathway 2H → 1T using the mixing rules applied to the structures in the latent space. Similar transformation pathway for 2H → 1T is observed for randomly selected encoding of 2H and 1T and applying linear mixing to z instead of μ in Eq. (1). However, the range of z values corresponding to α or β structure is still unknown. Hence, it is difficult to generate a specific structure using VAE without knowing the values of μ and the width of the Gaussian distribution σ.

In order to build a generative model that can create any specific type of structure, we train conditional VAE (CVAE) [25] with a loss function

$$\text{Loss} = E_{X\sim D}[E_{z\sim Q}[\log P(X|z,c) - D_{KL}[Q(z|X,c)\|P(z|c)]]]. \tag{2}$$

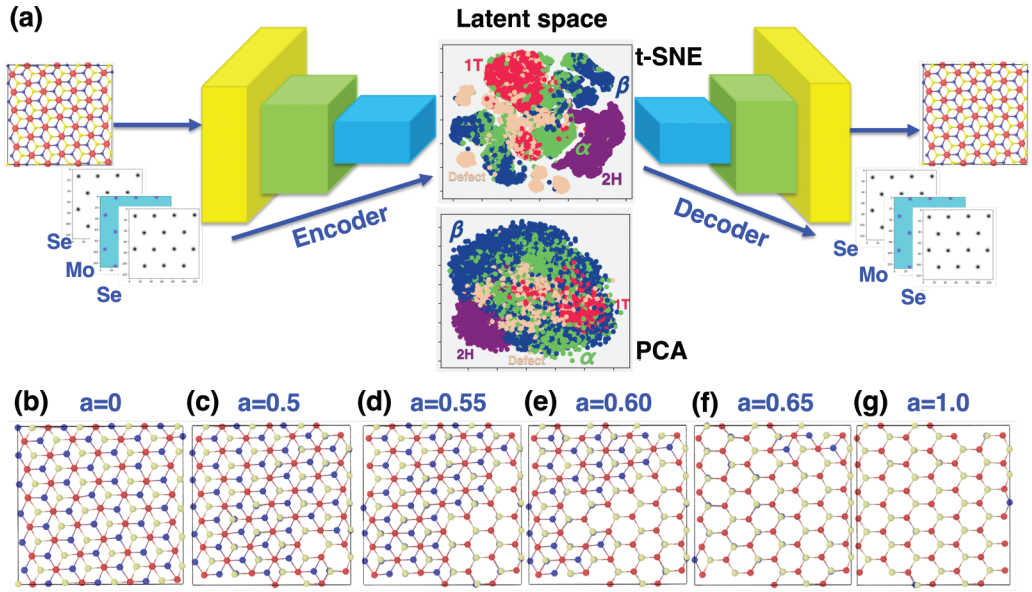

FIG. 3. Deep generative model for strain-induced structural transformation in $MoWSe_2$. (a) The figure shows how VAE unfolds the phase-transformation pathway in $MoWSe_2$ heterostructure. The latent space visualization of the VAE after training reveals structural representations of 2H, 1T, α, β, and various defects including the complex relationships among them. (b)–(g) 2H $\rightarrow$ 1T transformation pathways. Here samples are synthesized by mixing 2H and 1T in the latent space using Eq. (1).

This loss function is similar to the one used in VAE except that it contains a conditional variable c which controls the nature of the generated output. During training, the conditional variable (c) is appended to the input training data and a z value is generated by the encoder. Details of the training process are given in the Supplemental Material [19]. After training, the decoder takes as input a random number generated from a normal distribution (input z value) and a conditional variable c to generate a specific structure. Figures 4(a) and 4(e) show α and β structures generated by CVAE. In our model, values of $c = 0,\ 1,\ 2,\ 3,\ 4$ control the distribution of the generated 2H, 1T, α, β, and defects, respectively.

To validate the stability of these interfaces, we perform *ab initio* simulations based on DFT to relax the α and β interfacial structures generated by VAE and CVAE (see Sec. III for further details). Figures 4(b) and 4(f) depict atomic displacements during the relaxation of as-generated interfacial structures to their local minima. The VAE and CVAE generate high-quality interfaces that differ from the DFT-relaxed atomic positions by less than 0.18 Å and lie within 70 meV/atom of the local energy minima. The relatively small atomic displacements demonstrate that the CVAE can accurately reproduce high-quality interfacial structures close to local energy minima. Figures 4(b) and 4(f) also show that atomic relaxation arises mostly from the accommodation of the 2H/1T interfacial strain by the mechanically softer 1T WSe_2 structure. This behavior is consistent with the observed mechanism of α-interface formation in TMDCs such as MoS_2, where interfacial strain is accommodated by contraction in interplanar distances in the 1T structure. Additionally, the dynamic stability of the VAE-generated interfaces was confirmed by *ab initio* MD simulations at room and high temperatures, which showed negligible structural change from the as-generated interfaces from VAE (see Fig. S5).

Figures 4(c) and 4(g) depict the DFT-calculated density of electronic states of a metal-semiconductor-metal configuration containing α and β structures, respectively. The density of states of this material configuration, common in electronic devices like transistors, is characterized by the presence of valence and conduction bands arising from energy levels belonging to the $MoSe_2$ 2H crystal and gap-spanning states belonging to the energy levels from interfacial and bulklike (i.e., far from the interface) 1T WSe_2 regions. These states span the forbidden energy range of the 2H structure, closing the 1.3 eV band gap [Figs. 4c) and 4(g)] characteristic of the pristine $MoSe_2$ monolayer. Figures 4(d) and 4(h) show that Fermi-level states in the metal-semiconductor-metal structure are highly localized on the semimetallic interfacial and bulk 1T WSe_2 regions with negligible contribution from the semiconducting 2H $MoSe_2$ crystal.

In summary, MD simulations reveal that crack propagation in a $MoWSe_2$ monolayer induces not only 2H $\rightarrow$ 1T transformation but also novel intermediate structures α and β. The α structure is characterized by zigzag chains in which Mo-Mo nearest-neighbor separation is constricted by 14%. This is consistent with STEM observations showing 15% Mo-Mo constriction in MoS_2. In the β structure, each Se is 4 coordinated and the neighboring Mo are 6 coordinated which is again in accord with STEM data on MoS_2. The STEM experiment on MoS_2 also reveals another intermediate structure called γ between 2H and 1T phases. We also observe a few γ structures in our MD simulation but the data are not sufficient to train the VAE. We have used conditional VAE to computationally synthesize heterogeneous α and β phases in $MoWSe_2$. DFT simulations show minimal relaxation in as-generated atomic configurations, indicating that deep generative models such as VAE can accurately synthesize high-quality metal-semiconductor interfaces for electronic devices.

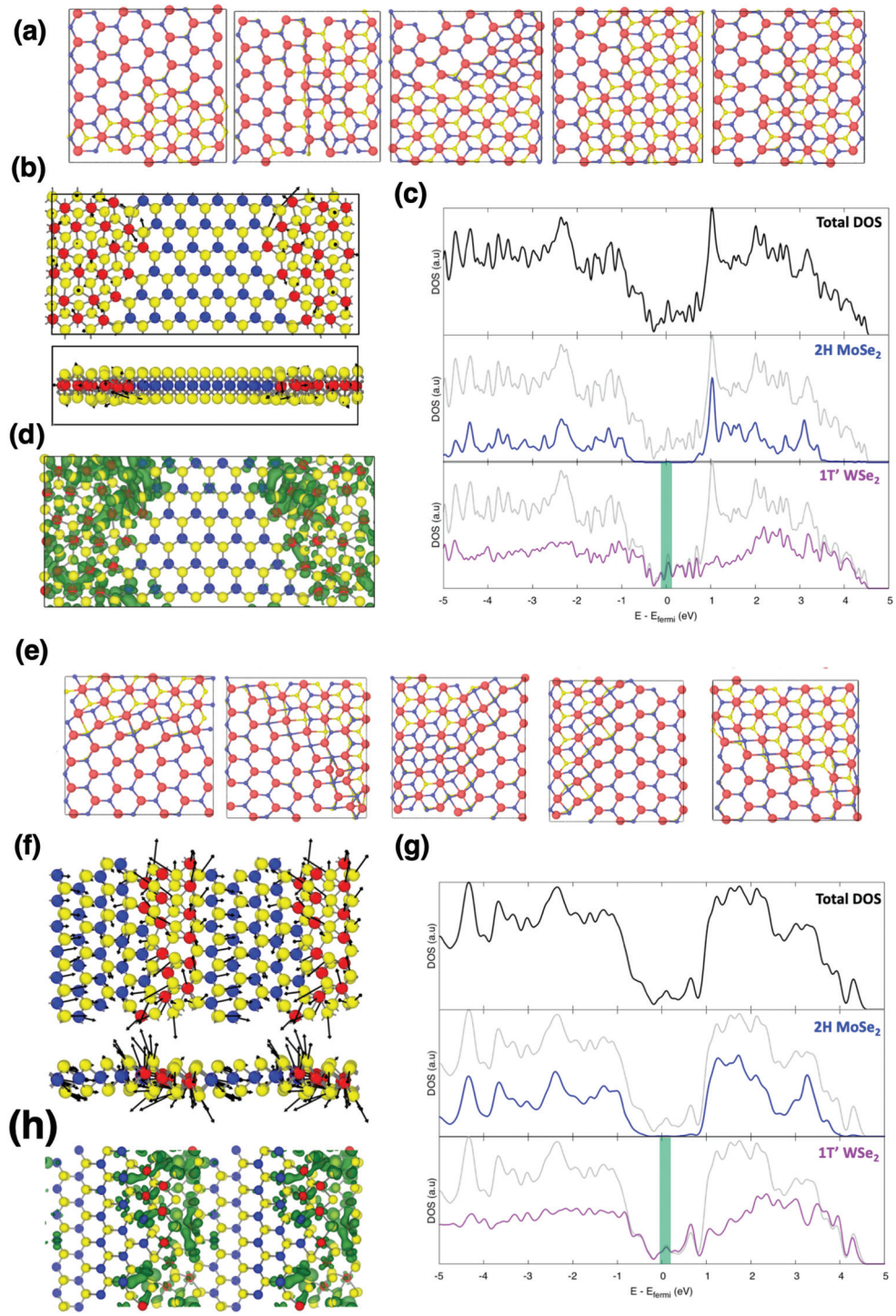

FIG. 4. Stability and electronic structure of VAE-synthesized interfaces. (a) CVAE based synthesis of heterogeneous α phase in $MoWSe_2$, which is validated by DFT simulations (b)–(d). The relatively small atomic displacement vectors [(b), magnified 5× for clarity] show minimal relaxation from as-generated atomic configurations, indicating that CVAE can accurately reproduce high-quality interfacial structures close to local energy minima. (c) Total (top) and spatially decomposed (middle and bottom) density of electronic states indicate that the interfacial and bulklike distorted 1T metallic states close the band gap in the 2H crystal. Spatial distribution of electronic states around the Fermi level (energy range highlighted in green) are described in (d), which shows that these states are localized at the 2H-1T interface and the distorted 1T region with negligible intensity in the semiconducting 2H region. (e) CVAE-based synthesis of β interfacial structure in $MoWSe_2$ validated by DFT simulations (f)–(h). Atomic displacement vectors [(f), magnified 5× for clarity] show only small relaxations from CVAE-generated atomic configurations. (g) Total (top) and spatially decomposed (middle and bottom) density of electronic states indicate that the interfacial and bulklike distorted 1T metallic states close the band gap in the 2H crystal. Spatial distribution of electronic states around the Fermi level (energy range highlighted in green) are described in (h), which shows that these states are localized at the 2H-1T interface and the distorted 1T region with negligible intensity in the semiconducting 2H region.

III. METHODS

A. Molecular dynamics methodology

We create a $MoWSe_2$ heterostructure similar to the one synthesized by the CVD process [1]. The system consists of a $MoSe_2$ matrix of dimensions $0.5\,\mu m \times 0.5\mu m$ in which W atoms are substituted randomly on Mo sites to create nanoscale patches of WSe_2. We insert a precrack of length 0.17 μm and clamp around 20 Å from the edges of the boundaries of the heterostructure. Molecular dynamics (MD) simulations are performed on the precracked $MoWSe_2$ with Stillinger-Weber type force field trained by DFT calculations and validated by experimental data for mechanical properties of $MoSe_2$ and WSe_2. The Stillinger-Weber force field models covalent interatomic interactions in the $MoSe_2$ and WSe_2 crystal structures using two-body and three-body terms that mimic energy profiles for bond-stretching and bond-bending (i.e., bond-angle distortion) interactions. Details of force-field parametrization and validation are given in the Supplemental Material [19].

Before studying mode-1 fracture of $MoWSe_2$ heterostructure, the system with the precrack is first relaxed using conjugate gradient and then heated to 100 K in the NVT ensemble for 100 ps. It is further thermalized for another 100 ps in the NVE ensemble. The time step is 1 fs during heating, thermalization, and mechanical straining. The thermalized system at 100 K is then subjected to mode-1 fracture, and no thermostat is applied at this time. During mode-1 fracture, the tensile strain is applied perpendicular to the precrack at a strain rate of $2.25 \times 10^{-7}\,s^{-1}$, and the system is relaxed for 100 ps after each incremental strain of 0.225%

We have further studied the effect of system size, distribution of WSe_2 patches in $MoSe_2$ matrix, and presence of interfacial point defects on the failure behavior of $MoWSe_2$ heterostructure. Details of these simulations are provided in Sec. 2 in the Supplemental Material [19].

B. Variational autoencoder (VAE) architecture and training

Atomic configurations from the MD simulation are used to create input data for VAE. The input consists of randomly selected patches of size 2 nm × 2 nm. Each patch is converted into a $64 \times 64 \times 3$ tensor, where the grid size is 0.031 nm × 0.031 nm. Channels 1 and 3 of the input tensors map the atomic coordinates of top and bottom Se layers, respectively, and channel 2 encodes the atomic information of the Mo layer. We use an exponential kernel to transform the atomic coordinates into a tensor representation:

$$T_{c=1,2,3}(x,y) = \sum_{i=1}^{N} \exp\left(-\frac{(x-x_i)^2 + (y-y_i)^2}{\eta^2}\right), \quad (3)$$

where c refers to the channel of the input tensor, (x, y) correspond to the center of a specific grid of the tensor T, (x_i, y_i) are the coordinates of Mo or Se atoms, N is the total number of atoms within 2 nm square from the origin ($x = 0,\ y = 0$), and η is the width of the exponential kernel (taken as 0.02 nm).

The VAE architecture is build using TensorFlow. The VAE encoder consists of three convolutional layers and a fully connected layer. The dimension of the first, second, and third convolutional layers of the encoder are $32 \times 32 \times 32$, $16 \times 16 \times 64$, and $8 \times 8 \times 128$, respectively. The output of the third convolutional layer is flattened and converted into a ten-dimensional vector using a fully connected layer. The decoder consists of two fully connected networks with 64 and 8192 hidden units, followed by two convolutional layers of dimension $16 \times 16 \times 64$ and $32 \times 32 \times 32$, and an output layer of dimensions $64 \times 64 \times 3$. In each convolutional layer, filter of size 4×4 is used along with a stride of 2, and leaky relu ($\alpha = 0.2$) as a nonlinear activation function. The output generated from the decoder is transformed from its tensor representation into Cartesian coordinate by doing depth first search (DFS) on each channel of output tensor. DFS search returns clusters of nonempty grids, where each grid is connected to its immediate eight neighbors. After that, the center of each cluster is used as an atomic coordinate. Details of VAE training is given in the Supplemental Material [19]. The VAE code used for training can be found at [26].

C. *Ab initio* DFT simulations

Interfacial structures and ground-state energies are optimized with density functional theory (DFT) using the projector augmented wave (PAW) [27] implementation in the Vienna *ab initio* simulation package (VASP) [28,29]. Exchange and correlation effects are calculated using the Perdew-Burke-Ernzerhof form of the generalized gradient approximation to the exchange-correlation functional [30]. Valence electron wave functions are constructed using a plane wave basis set containing components up to a kinetic energy of 450 eV and the reciprocal space is sampled at the Γ point with a 0.1 eV Gaussian smearing of orbital occupancies. DFT simulations are performed on systems containing 216 atoms, corresponding to 33 formula units of $MoSe_2$ and 39 formula units of WSe_2, in a simulation cell of dimensions 18.0×39.375 Å along the a and b directions, respectively. The simulation cell spans 20 Å along the c axis and contains a 15 Å thick vacuum to remove spurious image interactions. DFT calculations are run until the energy converges to within 1×10^{-6} eV/atom and forces on ions are under 5×10^{-2} eV/Å in each self-consistency cycle.

ACKNOWLEDGMENTS

This work was supported as part of the Computational Materials Sciences Program funded by the U.S. Department of Energy, Office of Science, Basic Energy Sciences, under Award No. DE-SC0014607. The simulations were performed at the Argonne Leadership Computing Facility under the DOE INCITE and Aurora Early Science programs and at the Center for High Performance Computing of the University of Southern California. We acknowledge helpful discussions with Amey Apte, ChandraSekhar Tiwary, Vidya Kochat, and Professor Ajayan at Rice University.

P.R. carried out the MD simulation and the development of VAE model using MD data. A.K. developed the force field and performed the DFT simulation. R.K.K., A.N., and P.V. advised this work. All authors contributed to the analysis, discussion, and preparation of the manuscript.

[1] A. Apte, V. Kochat, P. Rajak, A. Krishnamoorthy, P. Manimunda, J. A. Hachtel, J. C. Idrobo, S. A. S. Amanulla, P. Vashishta, A. Nakano, R. K. Kalia, C. S. Tiwary, and P. M. Ajayan, ACS Nano **12**, 3468 (2018).

[2] A. K. Geim and I. V. Grigorieva, Nature (London) **499**, 419 (2013).

[3] D. P. Kingma and M. Welling, arXiv:1312.6114.

[4] Y. C. Lin, D. O. Dumcenco, Y. S. Huang, and K. Suenaga, Nat. Nanotechnol. **9**, 391 (2014).

[5] R. Ramprasad, R. Batra, G. Pilania, A. M. Kanakkithodi, and C. Kim, Npj Comput. Mater. **3**, 54 (2017).

[6] A. M. Kanakkithodi, T. D. Huan, and R. Ramprasad, Chem. Mater. **29**, 9001 (2017).

[7] E. D. Cubuk, R. J. S. Ivancic, S. S. Schoenholz, D. J. Strickland, A. Basu, Z. S. Davidson, J. Fontaine, J. L. Hor, Y. R. Huang, Y. Jiang, N. C. Keim, K. D. Koshigan, J. A. Lefever, T. Liu, X. G. Ma, D. J. Magagnosc, E. Morrow, C. P. Ortiz, J. M. Rieser, A. Shavit, T. Still, Y. Xu, Y. Zhang, K. N. Nordstrom, P. E. Arratia, R. W. Carpick, D. J. Durian, Z. Fakhraai, D. J. Jerolmack, D. Lee, J. Li, R. Riggleman, K. T. Turner, A. G. Yodh, D. S. Gianola, and A. J. Liu, Science **358**, 1033 (2017).

[8] L. Bassman, P. Rajak, R. K. Kalia, A. Nakano, F. Sha, J. Sun, D. J. Singh, M. Aykol, P. Huck, K. Persson, and P. Vashishta, Npj Comput. Mater. **4**, 74 (2018).

[9] A. Kaczmarowski, S. J. Yang, I. Szlufarska, and D. Morgan, Comput. Mater. Sci. **98**, 234 (2015).

[10] A. C. T. van Duin, S. Dasgupta, F. Lorant, and W. A. Goddard, J. Phys. Chem. A **105**, 9396 (2001).

[11] T. Liang, Y. K. Shin, Y. T. Cheng, D. E. Yilmaz, K. G. Vishnu, O. Verners, C. Zou, S. R. Phillpot, S. B. Sinnott, and A. C. T. van Duin, Annu. Rev. Mater. Res. **43**, 109 (2013).

[12] T. D. Huan, R. Batra, J. Chapman, S. Krishnan, L. Chen, and R. Ramprasad, Npj Comput. Mater. **3**, 37 (2017).

[13] B. Onat, E. D. Cubuk, B. D. Malone, and E. Kaxiras, Phys. Rev. B **97**, 094106 (2018).

[14] E. D. Cubuk, B. D. Malone, B. Onat, A. Waterland, and E. Kaxiras, J. Chem. Phys. **147**, 024104 (2017).

[15] R. G. Bombarelli, J. N. Wei, D. Duvenaud, J. M. H. Lobato, B. S. Lengeling, D. Sheberla, J. A. Iparraguirre, T. D. Hirzel, R. P. Adams, and A. A. Guzik, ACS Cent. Sci. **4**, 268 (2018).

[16] B. S. Lengeling and A. A. Guzik, Science **361**, 360 (2018).

[17] S. G. Louie and M. L. Cohen, Phys. Rev. B **13**, 2461 (1976).

[18] D. Le and T. S. Rahman, J. Phys. Condens. Matter **25**, 312201 (2013).

[19] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevB.100.014108 for detailed information about force-field parametrization, mode-1 fracture of $MoWSe_2$ with interfacial defect for (11) and (10) orientation, VAE training and dynamic stability of VAE-generated 2H-1T interfacial structures, and movies showing strain incuded phase transformation in $MoWSe_2$ during mode-1 fracture.

[20] A. Omeltchenko, J. Yu, R. K. Kalia, and P. Vashishta, Phys. Rev. Lett. **78**, 2148 (1997).

[21] M. J. Buehler, H. Tang, A. C. T. van Duin, and W. A. Goddard, III, Phys. Rev. Lett. **99**, 165502 (2007).

[22] Y. C. Yang, X. Li, M. R. Wen, E. Hacopian, W. B. Chen, Y. J. Gong, J. Zhang, B. Li, W. Zhou, P. M. Ajayan, Q. Chen, T. Zhu, and J. Lou, Adv. Mater. **29**, 1604201 (2017).

[23] L. van der Maaten and G. Hinton, J. Mach. Learn. Res. **9**, 2579 (2008).

[24] A. Radford, L. Metz, and S. Chintala, arXiv:1511.06434.

[25] K. Sohn, H. Lee, and X. Yan, *Proceedings of the 28th International Conference on Neural Information Processing Systems, Montreal, Canada, 2015* (MIT Press, Cambridge, MA, 2015), p. 3483.

[26] https://github.com/rajak7/Generative_model_for_2H-1T_phase-transformation_in_MoWSe2_heterostructure

[27] P. E. Blöchl, Phys. Rev. B **50**, 17953 (1994).

[28] G. Kresse and J. Furthmüller, Phys. Rev. B **54**, 11169 (1996).

[29] G. Kresse and J. Furthmüller, Comput. Mater. Sci. **6**, 15 (1996).

[30] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).