Chemical Science

ROYAL SOCIETY OF CHEMISTRY

EDGE ARTICLE
Volker L. Deringer *et al.*
Exploring the configurational space of amorphous
graphene with machine-learned atomic energies

ROYAL SOCIETY
OF CHEMISTRY

EDGE ARTICLE

Exploring the configurational space of amorphous graphene with machine-learned atomic energies

Zakariya El-Machachi, [ID] [a] Mark Wilson [ID] [b] and Volker L. Deringer [ID] *[a]

Two-dimensionally extended amorphous carbon ("amorphous graphene") is a prototype system for disorder in 2D, showing a rich and complex configurational space that is yet to be fully understood. Here we explore the nature of amorphous graphene with an atomistic machine-learning (ML) model. We create structural models by introducing defects into ordered graphene through Monte-Carlo bond switching, defining acceptance criteria using the machine-learned local, atomic energies associated with a defect, as well as the nearest-neighbor (NN) environments. We find that physically meaningful structural models arise from ML atomic energies in this way, ranging from continuous random networks to paracrystalline structures. Our results show that ML atomic energies can be used to guide Monte-Carlo structural searches in principle, and that their predictions of local stability can be linked to short- and medium-range order in amorphous graphene. We expect that the former point will be relevant more generally to the study of amorphous materials, and that the latter has wider implications for the interpretation of ML potential models.

Introduction

The study of the amorphous state has long been of fundamental research interest.[1] It is also increasingly important to understand structure–property correlations in amorphous materials, owing to ubiquitous applications in solar cells,[2] transparent electronic devices,[3] or phase-change memories.[4] Whilst bulk amorphous phases are challenging to study structurally, two-dimensional (2D) amorphous materials can be directly visualized by atomistic imaging techniques such as high-resolution transmission electron microscopy (HRTEM).[5–11] And just like graphene is the prototypical ordered 2D material, there is ongoing research interest in its disordered analogue(s). Indeed, Toh *et al.* recently synthesized a centimeter-scale sample of free-standing monolayer amorphous carbon[12] and characterized the structure based on the interpretation of HRTEM images.

The amorphous forms of carbon have been widely studied using computer simulations. Most commonly, these studies have been carried out in the framework of molecular dynamics (MD), from early work on melt-quenching[13–19] to direct simulations of thin-film growth by ion deposition.[20–23] Very recently, the properties of "amorphous graphite", as a 3D extended arrangement of individual amorphous graphene sheets, were investigated with density-functional theory (DFT) based MD and electronic-structure computations.[24] Earlier simulation studies

had emphasized the connection between low-density amorphous carbon and the idealized case of 2D amorphous graphene ("aG" in the following).[25,26]

Graphene itself, as a 2D system, is rather well-confined and allows for Monte Carlo (MC) simulations to model topological defects. The Wooten–Winer–Weaire (WWW) algorithm, initially proposed for silicon,[27] remains a simple yet robust approach to generating continuous random network (CRN) models. In this, disorder is gradually introduced through local bond transpositions and structural relaxation. To simulate aG, bond transpositions are introduced as Stone–Wales (SW) defects,[28] with moves being accepted or rejected based on a suitable Metropolis criterion, typically the total-energy difference between the new and old configuration. SW defects are created by a formal in-plane 90° rotation of two atoms around the mid-point of the bond[28] and are the foundational example of (intrinsic) topological disorder in 2D carbon.[29–32] This approach has long been used to study aG.[33] Various implementations of the WWW algorithm exist: D'Ambrosio *et al.* showed that the majority of bond transpositions are rejected during annealing, and that an early decision scheme can enhance computation speed by rejecting unfavorable transpositions;[34] Ormrod Morley *et al.* constructed Metropolis criteria from topological metrics such as ring distributions (that is, energy changes were not considered in that case).[35]

Machine learning (ML) based interatomic potential models are increasingly being used to accelerate materials simulations.[36–38] ML potentials are typically "trained" with DFT data and can achieve similar accuracy for a small fraction of the cost. One key assumption in many of these methods is that the total energy can be separated into sums of machine-learned

[a]Department of Chemistry, Inorganic Chemistry Laboratory, University of Oxford, Oxford OX1 3QR, UK. E-mail: volker.deringer@chem.ox.ac.uk

[b]Department of Chemistry, Physical and Theoretical Chemistry Laboratory, University of Oxford, Oxford OX1 3QZ, UK

atomic energies.[39,40] Whilst being an approximation in the first place, it was argued that these atomic energies may in fact be amenable to interpretation: a connection between ML atomic energies and local chemical structure was made for partly occupied crystallographic sites in β-rhombohedral boron,[41] and for atomic environments with different coordination numbers in amorphous silicon.[42] ML models for other properties of local atomic environments have recently been investigated, ranging from the local electronic density of states[43,44] to local distortion factors in grain boundaries.[45] The nature of these local ML properties (including atomic energies), and their usefulness in predicting physical properties, remains an interesting research question. (See, e.g., ref. 46 for a discussion of ML atomic energies in a chemically complex system.)

In the present work, we explore the configurational space of amorphous graphene based on an ML potential model that gives access to total and local energies. We use an MC bond-switching algorithm where ML atomic energies are used in the acceptance criterion and show that doing so leads to physically sound structural models. Depending on the details of the algorithm, we obtain CRN-like or paracrystalline structures. Our work shows that ML atomic energies can be used in different ways to "drive" MC simulations based on local and nearest-neighbor (NN) energy contributions, with implications for research on amorphous graphene and likely on other disordered structures and materials.

Methods

Potential-energy models

A common ansatz in developing potential-energy models is that the total energy, E, can be separated into a sum of atomic contributions:

$$E = \sum_i^{\text{atoms}} \varepsilon_i, \tag{1}$$

where $\varepsilon_i = \varepsilon(\{\mathbf{r}_{ij}\})$, with i, j being atomic indices, and $r_{ij} \leq r_{\text{cut}}$. This ansatz can be applied to many systems as short- and medium-range interactions predominantly determine the total energy. However, the question of how to formulate $\varepsilon(\{\mathbf{r}_{ij}\})$ is not trivial: in carbon, there is a vast configurational space with a subtle interplay between structure and energetics, such as in dihedral and torsional forces, weak interlayer interactions, and so on. Moreover, locality can depend on the structure: numerical experiments showed a large difference in the locality of atomic forces for diamond versus graphite.[47] (We note in passing that extracting local energies directly from DFT is not trivial, although attempts have previously been made.[48])

Empirical potentials are typically parameterized for a specific composition and phase of a material. For example, the original Tersoff potential for carbon[13] was parameterized by fitting parameters of repulsive and attractive pair potentials to cohesive energies of carbon polymorphs along with the lattice parameter and bulk modulus of diamond. This potential[13] used a bond-order approach, where bond strengths are modified according to the number of neighbors. The original reactive empirical bond-

order (REBO) potential was an update of the Tersoff potential incorporating hydrogen,[49] with REBO-II adding further improvements.[50] A long-range Lennard-Jones term was added to REBO-II, creating the adaptive intermolecular reactive bond order (AIR-EBO) potential.[51] The long-range bond order potential (LCBOP)[52] is similar to AIREBO, having a long-range term in conjunction with a bond-order description—albeit here it is built in from the beginning. The environment-dependent interaction potential (EDIP) for carbon[53,54] was initially developed from an earlier silicon EDIP model.[55,56] EDIP has been shown to be successful in describing various properties of amorphous carbon, including the graphitization at low and high densities.[57–59]

Gaussian approximation potential (GAP)

The Gaussian approximation potential (GAP) framework[40] is used to "machine-learn" interatomic potential models from quantum-mechanical data, often based on DFT. Unlike empirical potentials which are constructed based on physical knowledge, GAP makes a non-parametric fit. This means that the model can adjust to complex input data—however, it also means that the selection and quantity of reference ("training") data is critically important.[60]

In brief, the local energy for a GAP model is

$$\varepsilon(\mathbf{q}) = \sum_{t=1}^{N_t} \alpha_t K(\mathbf{q}, \mathbf{q}_t), \tag{2}$$

where the sum runs over N_t training configurations represented by the local-environment descriptor $\mathbf{q}_t$, with a corresponding weighting coefficient α_t attributed during fitting. K is a covariance kernel which measures the similarity between the input and training configurations, represented by $\mathbf{q}$ and $\mathbf{q}_t$, respectively. A commonly used approach for the latter task is the Smooth Overlap of Atomic Positions (SOAP) descriptor and kernel.[61]

For the present work we use the amorphous carbon potential, GAP-17.[47] This model has been shown to predict energies within tens of meV per atom compared to DFT, as well as providing a good description of structural and elastic properties of amorphous carbon.[47]

Initialization

We generate a pristine, 200-atom layer of crystalline graphene using the experimental bond length of 1.42 Å (ref. 62 and 63). A spacing of 20 Å between layers ensures that there is no interaction between periodic replicas. Where noted, NN-energy-based runs start from structural models that have been "thermalized", i.e., disordered, using local energies at $\beta = 2.0$ eV^{-1} (see below). The cell parameters are allowed to relax prior to the first bond switch, and then kept fixed for the duration of the MC simulation. The structures are kept planar, simplifying the present study to the idealized 2D case, and noting that puckering can lower the energy further.[64]

Monte-Carlo protocol

A kinetic MC algorithm is used to generate aG structural models. The simulations begin either from cG or from

a thermalized structure. At each step, a random atom is chosen along with a random neighbor, determined using a cutoff of 1.85 Å. The atoms undergo an SW rotation (90° about the bond center) and the new structure is then relaxed using the conjugate-gradient algorithm. For initial testing of structural relaxations, energy evaluations, and MC runs, we used the Atomic Simulation Environment (ASE)[65] interfaced to quippy[66,67] (**https://github.com/libAtoms/QUIP**). The production MC runs shown in this work used LAMMPS.[68] The force tolerance for structural relaxations was 1 meV Å^{-1} with a maximum of 150 relaxation steps. A topological constraint is imposed where the newly relaxed structure must be 3-coordinate, otherwise the move is rejected. (This constraint is included because there will be moves which create stubborn coordination defects, and we found that these defects hindered the progress of MC annealing where the simulation would get "stuck" in local minima.) If the constraints are met, the new structure is accepted with the following probability:

$$w = \begin{cases} 1, & \text{if } \Delta E \leq 0, \\ \exp(-\beta \Delta E), & \text{if } \Delta E > 0, \end{cases} \quad (3)$$

where $\beta = (k_B T)^{-1}$ and $\Delta E = E_{new} - E_{old}$. If $\Delta E > 0$, we generate a random number z in [0,1), and if $w > z$, then the move is accepted, else it is rejected and the previous configuration is kept. In this work, β has no physical relation to temperature since ΔE is not measured for all atoms about their equilibrium position, as the atoms are being held fixed after the bond transposition and thus thermal fluctuations are ignored (unlike with MD). Fixing particle positions at every move results in ergodicity being broken and thus samples are not taken from a Boltzmann distribution.[69] Furthermore, the local-energy and NN-energy framework explicitly do not include all atomic energies which is another source of ergodicity being broken. This is not a concern for the present work as temperature-driven dynamics are not relevant (see ref. 69 for an approach to introducing ergodicity to a similar problem) and results in β becoming a tunable parameter. β was initially chosen to be 2.0 eV^{-1} to correspond to the study by Toh *et al.*,[12] and then tuned heuristically for the local- and total-energy-based MC runs.

We use the atomic energies for the defect pair in the Metropolis criterion, defining the local defect energy as

$$\varepsilon_{local} = \varepsilon_1 + \varepsilon_2 \quad (4)$$

and the NN defect energy as

$$\varepsilon_{NN} = \varepsilon_{local} + \sum_i^{NN} \varepsilon_i, \quad (5)$$

where ε_i is the local atomic energy of the i-th atom, with $i = 1$ and $i = 2$ denoting the two atoms in the SW defect pair, and $i = 3\ldots6$ the topological NNs of the SW defect pair (Fig. 1c and d). MC runs based on total energies were also carried out for reference. For the 200-atom systems, 25 independent and parallel simulations were conducted over 10 000 MC steps each for different β values and for the 612-atom system, 20 000 MC steps were taken.

Fig. 1 Stone–Wales defect in graphene and definition of nearest-neighbor (NN) atoms. (a) Schematic of an in-plane single-bond rotation in graphene, leading to an SW defect. (b) A bond rotation by 90° creates two 5-membered and two 7-membered rings. Definitions are given for the defect bond length, d_1, the defect–NN bond length, d_2, and the defect bond angle, θ (*cf.* Table 1). (c) Definition of local energy for the Metropolis criterion. The gray atoms labeled **1** and **2** are the SW defect pair, with shading indicating the overlap of the two cutoff spheres up to which atoms contribute to the local energy (this value is 3.7 Å for the GAP-17 model, larger than sketched here). (d) The green atoms (**3–6**) are the topological NNs of the defect pair. Green lines indicate the overlap of the cutoff spheres for each NN atom.

Structural analysis

Ring statistics were determined using a shortest-path algorithm[70] as implemented in Matscipy.[71] Topological metrics typically used in network analysis were applied to the 612-atom structures, namely, Lemaître's law[72] and the assortativity[73] (as discussed below). For SOAP structural analysis, we compared each individual atom in a given aG structural model to an atom in cG. The SOAP parameters were: radial cutoff, 5.5 Å; cutoff transition width, 0.5 Å; neighbor-density smoothness, $\sigma_{atom} = 0.5$ Å; basis-set convergence parameters, $n_{max} = l_{max} = 16$; dot-product kernel raised to a power of $\zeta = 4$. All structures were visualized using OVITO.[74]

Phonon calculations

A geometry optimization was performed on the local, NN and Toh *et al.* structures keeping the cell fixed ($F_{max} < 1$ meV Å^{-1}). We applied a finite-difference approach with a displacement of 0.01 Å. The forces in the displaced structures were computed using GAP-17 and second-order force constants were computed using phonopy.[75] A 50 × 50 × 1 k-point mesh was used to determine the vibrational density of states (VDOS). We observed minor imaginary modes in the aG models, presumably due to the constraint of planarity; these are omitted from VDOS plots and will be investigated in future work.

Results and discussion

Local energies for a single defect

Fig. 2 characterizes the structure and energetics of a single SW defect in graphene. GAP-17 predicts that the energies of the SW defect pair of atoms are considerably lower in energy than those of the NN atoms. This behavior is in marked contrast with that for the empirical potentials (Fig. 2b and Table 1): the latter suggest that the SW-pair atoms are relatively higher in energy (with the exception of LCBOP), with atoms in the 5-membered rings being higher in energy overall and those in 7-membered rings being lower. LCBOP and GAP-17 predictions are similar when averaged over the pentagon-pair environments, only differing slightly in structural parameters. REBO and AIREBO show slightly larger SW-pair energies and more strained structures (Table 1).

The total defect formation energy ΔE_{total} for GAP-17 is similar overall to that predicted by most empirical potentials—and to DFT values, which we obtained using CASTEP.[79] In an earlier study, hybrid-DFT and quantum Monte Carlo calculations predicted values of 5.69 eV and 5.92 ± 0.03 eV, respectively.[30] EDIP, LCBOP, and REBO were parameterized for sp^2 environments, and their ΔE_{total} is in close agreement with that for GAP-17. AIREBO was also parameterized for sp^2 environments, however the added long-range term may not be optimized for 2D planar carbon. In contrast, $\Delta \varepsilon_{\text{local}}$ and $\Delta \varepsilon_{\text{NN}}$ are not in close agreement, where $\Delta \varepsilon_{\text{local}}$ differs the most across all the potentials (except for REBO and AIREBO which are parameterized identically for short-range terms). Overall, GAP-17 predicts the lowest energy for the NN defect formation energy among the potential models investigated.

Table 1 Structure and formation energy of a single SW defect in graphene from ML and empirical potentials as well as DFT. The table shows the computed bond length in pristine graphene (d_0), the defect-pair (d_1) and defect-NN bond lengths (d_2), the defect bond angle (θ; cf. Fig. 1b), and three energies associated with defect formation: the local energy change in the rotated atoms only ($\Delta \varepsilon_{\text{local}}$), the NN energy change ($\Delta \varepsilon_{\text{NN}}$), and the total defect formation energy (ΔE_{total})

	d_0 (Å)	d_1 (Å)	d_2 (Å)	θ (deg)	$\Delta \varepsilon_{\text{local}}$ (eV)	$\Delta \varepsilon_{\text{NN}}$ (eV)	ΔE_{total} (eV)
GAP-17	1.41	1.32	1.44	121.89	0.14	2.25	5.58
EDIP	1.50	1.39	1.58	118.26	1.69	3.32	5.58
LCBOP	1.42	1.34	1.45	121.76	0.73	2.39	5.33
REBO	1.42	1.34	1.46	120.64	0.99	2.85	5.61
AIREBO	1.40	1.32	1.44	120.88	1.09	3.15	6.27
LDA[76]	1.41	1.31	1.45	122.34	—	—	5.14
PBE[77]	1.43	1.32	1.46	122.49	—	—	5.03
PBEsol[78]	1.42	1.32	1.46	122.40	—	—	5.02

Structural parameters as predicted by the respective potentials are also shown in Table 1. The defect-pair bond length (d_1), defect–NN bond length (d_2), and defect bond angle, θ (Fig. 1b) are in close agreement across all potentials—with the exception of EDIP, which notably increases the cell parameters upon relaxation. This is shown by the difference in cG bond length (d_0) between EDIP and the other potentials when the cell parameters are allowed to relax, with EDIP predicting an elongated d_0. GAP-17 and LCBOP are in close agreement with regard to θ, predicting a slightly wider bond angle for the 5|7 pair, whereas EDIP predicts it slightly lower, notwithstanding the fact that the EDIP defect formation energy is in close agreement with that for GAP-17. Furthermore, GAP-17 is in excellent agreement with the bond lengths and angles from our DFT calculations.

Fig. 2 Structure and local energy of a Stone–Wales defect in graphene. The figure compares local energies (top row) and nearest-neighbor averaged energies ("NN", bottom row) for a single Stone–Wales (SW) defect as relaxed using the respective interatomic potential. (a, b) Local atomic energies from GAP-17 and empirical potentials, respectively. (c, d) Same but for NN energies (eqn (5)), from GAP-17 and empirical potentials, respectively.

Using local energies to drive Monte-Carlo annealing

Fig. 3 shows the energy profiles of independent parallel MC simulations *via* successive SW transformations, which we use to create structural models of aG. Panel (a) shows the evolution of the average energy of the ensemble when using (only) the atomic energies of the SW defect pair in the Metropolis criterion. Clearly, using a β value of 2.0 eV^{-1} leads to a highly disordered 3-coordinate structure (as shown in the inset), with energy convergence occurring after about 1000 MC steps. Such structures contain high-energy 3- and 4-membered rings with significant strain. We call the resulting structures "thermalized" and use them as starting points for the runs characterized in blue in Fig. 3b, as indicated by an arrow.

Nearest-neighbor (NN) atoms are defined topologically in the present work, and so we include the sum over the SW pair and its bonded neighbors, including six atoms in total (eqn (5)). The energy difference before and after the SW transformation is used in the Metropolis criterion. Results from two protocols, starting either from cG or from the previously thermalized structures, converge to similar energy values (green and blue curves in Fig. 3b). It seems that runs from both starting points tend toward a mutual limiting distribution. This is reflected in Fig. 3b, as both curves have means and standard deviations in very close agreement after approximately 4000 MC steps. However, it is apparent that this protocol is not ergodic, as discussed in the Methods section. The principle of ergodicity states that all possible configurations of the system should be attainable[80] and thus it is clear that given the topological constraints imposed, it is impossible for certain proposed moves to be accepted even though they may be energetically favorable.

Fig. 3c illustrates the effect of varying β on the resulting energy profile when using the local-energy criterion. It is clear

Fig. 3 Evolution of disordered graphene structures during Monte-Carlo simulations. Lines indicate mean energies for an ensemble of 25 separate runs, with shaded regions indicating a standard deviation of 1σ. Representative structural snapshots are shown, with atoms color-coded by local energy. (a) The atomic energies of the two atoms involved in the SW transformation are used for the Metropolis criterion, creating a "thermalized" structure rapidly at $\beta = 2.0$ eV^{-1}. These structures are used as initial configurations in (b). (b) MC annealing *via* NN energies ($\beta = 2.0$ eV^{-1}) for the Metropolis criterion. The curves converge at slightly above the energy of an MC-annealed paracrystalline structure from Toh *et al.* (ref. [12]), which is included for reference (red line). (c) The effect of β on the MC energy profile. It is seen that a smaller β (higher simulation "temperature") leads to a higher-energy structure and faster convergence. We emphasize that despite the high standard deviations, the lowest occurring energy value is $\Delta E = 0$ eV at.$^{-1}$ corresponding to ideal crystalline graphene. (d) Heuristically tuning β to find a mutual limiting distribution between MC approaches (within statistical fluctuations).

that using a lower β value corresponds to a higher-energy final structure as well as rapid convergence: for $\beta = 2.0$ eV^{-1}, convergence occurs in fewer than 1000 MC steps (black line). Increasing β results in lower-energy structures and slower convergence, as shown by the $\beta = 50.0$ eV^{-1} and $\beta = 60.0$ eV^{-1} results. For $\beta = 60.0$ eV^{-1}, convergence has not occurred after 10 000 steps, however the mean appears to be tending towards that for $\beta = 50.0$ eV^{-1}, implying that there might be a minimum level of disorder that is "stable" for the local-energy criterion. Additionally, as β was increased beyond 60.0 eV^{-1}, proposed moves were always rejected, further suggesting a minimum convergence energy using this framework.

Fig. 3d shows the results from three different runs with heuristically tuned β values for local-, NN-, and total-energy criteria, respectively, converging towards a mutual distribution (within statistical fluctuations). This result allows the direct comparison of structures generated by the different frameworks side-by-side, as the effect of β has been removed. Such a comparison is particularly instructive for lower-energy structures, as we will show below: they show a richer configurational space, including medium-range order (as compared to the highly disordered high-energy structures), and this space may be traversed differently by different MC runs.

The configurational space of amorphous graphene

Having explored different protocols for generating aG structural models, we next created larger-scale structures: the system size was increased to 612 atoms and structures from local- and NN-energy-based searches were studied. Additionally, the 610-atom structural model published by Toh et al. (ref. 12) was included for comparison. Fig. 4 shows these three configurations. As we keep the cells fixed throughout the Monte-Carlo searches (Methods), we computed the tensile stress for the final structures, using GAP-17, and found them to correlate with the degree of disorder. The structure generated using local energies had an in-plane tensile stress of 3.53 GPa; the NN structure had 2.85 GPa, and for the Toh et al. structure we obtained 0.87 GPa.
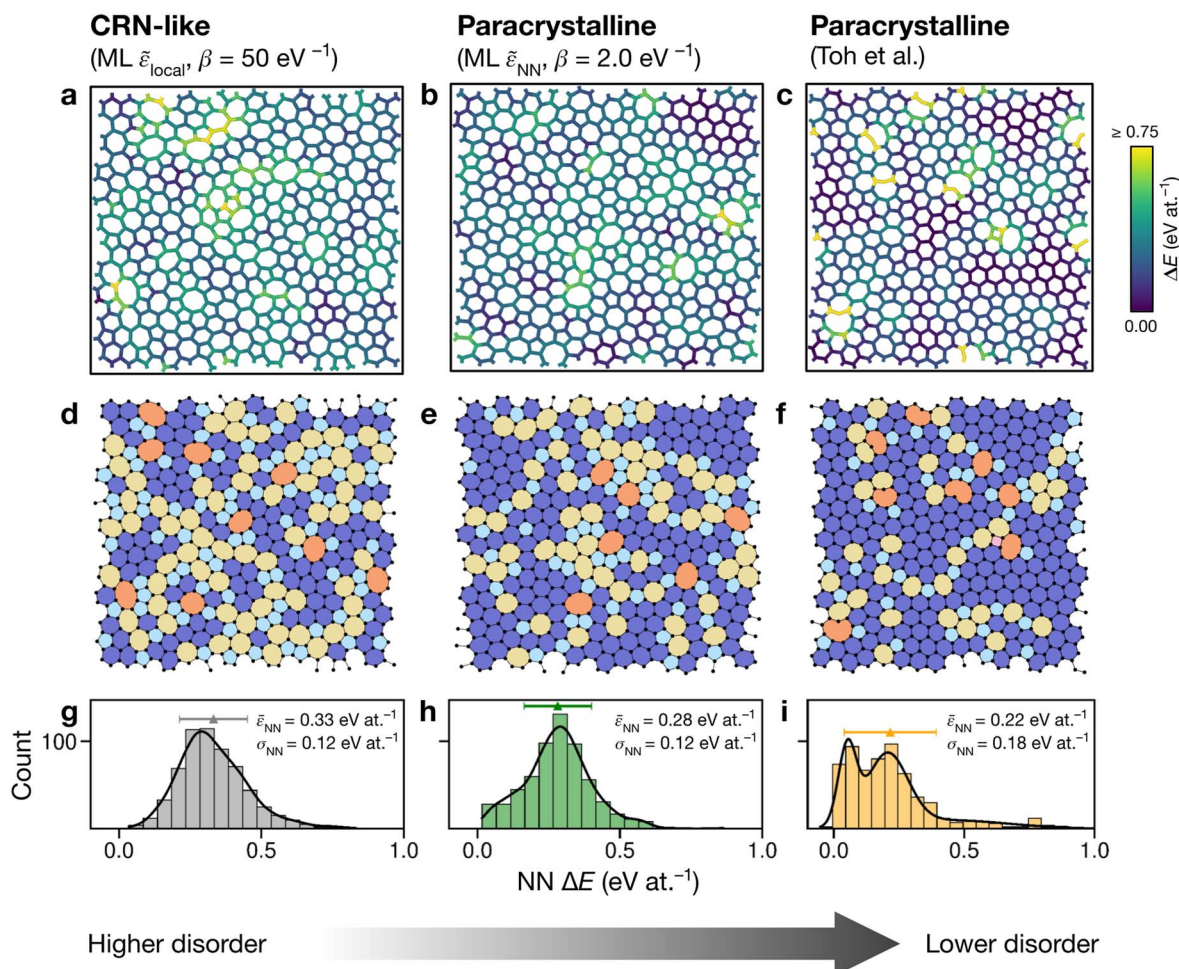

Fig. 4 Structural models of amorphous graphene. Structures in panels (a–c) are color-coded according to the average atomic energy over a chosen atom and the corresponding NNs. (a) Structure after the final MC step for the local-energy criterion and $\beta = 50$ eV^{-1}. (b) As for (a), but with NN energies used in the Metropolis criterion. (c) Structure taken from Toh et al.[12] and optimized with GAP-17. (d–f) Structures as in (a–c), now color-coded by ring size (pink, ≤ 4; light blue, 5; dark blue, 6; yellow, 7; orange, ≥ 8). (g–i) Distributions of NN energies in the respective structures, with mean values and standard deviations given. Curves were obtained as kernel density estimates.

For the local-energy framework, the absence of crystal-like pockets is clear as seen in panels (a) and (d). The structure resembles a CRN, with chains of 5- and 7-membered rings running across the cell, and there are more large rings compared with the other structures. The pronounced disorder is likely a result of the low ML energy for the SW defect itself (Fig. 2a), and thus of the low energy cost for these transpositions (given that medium-range order is not captured in the local energies alone).

With NN energies used in the Metropolis criterion, panels (b) and (e) suggest that using an NN criterion encourages small pockets of crystal-like regions forming, indicating that direct contributions from NNs maintain medium-range order, whilst retaining an amorphous structure at $\beta = 2.0$ eV^{-1}. Visually, the locally averaged energies (up to NNs) in panels (a–c) suggest that using just the local defect-pair energy gives more regions of higher energy (shown in yellow in Fig. 4a) compared with the NN framework (Fig. 4b). This appears consistent with the nature of CRNs *versus* paracrystalline structures, and color-coding by NN energies shows a clear difference between the two structures.

In the NN-energy-based structure (Fig. 4b), there are pockets of crystallinity, indicated by regions of ordered 6-membered rings. Interestingly, there is an aggregation of more disordered regions, with 5-membered rings surrounding larger 7- and 8-membered rings. For the structure shown in Fig. 4c, the authors started from a randomized, hard-sphere constrained structure and worked towards a paracrystalline sample using the AIREBO potential. In the resulting structure, we note the presence of coordination defects, since the randomized initial structure was not topologically constrained. There is also a 4-membered ring (pink in Fig. 4f). It is evident that this sample is paracrystalline with regions of locally crystal-like order separated by 5|7 grain boundaries, and that it is more ordered than the NN-based structure (Fig. 4e). Larger defects appear to congregate as seen, for example, in the top of the figure. These defective environments are high in energy, as shown by the color-coding.

Fig. 4g–i shows the distribution of NN energies for the respective structures. The average energy for the atoms in the local-energy-based structure is 0.33 eV at.$^{-1}$ with a standard deviation of $\sigma = 0.12$ eV at.$^{-1}$, and the energy of the NN-based structure is 0.28 ± 0.12 eV at.$^{-1}$ above pristine graphene. Both values are close to the mean energy resulting from the independent runs for the 200-atom system (Fig. 3b and d): hence, an individual 200-atom run will not likely be sufficient to describe aG, but an ensemble of multiple independent runs will be—just like many small-scale structural models of 3D amorphous carbon have been used in the fitting of GAP-17.[47] The distribution in the locally averaged energies is similar overall between Fig. 4g and h.

For Fig. 4i, we find a bimodal distribution in the energy, indicating that this structure shows higher paracrystalline order compared to that generated by the NN framework. There exist pockets of more ordered, more energetically favorable regions in between more disordered ones, as seen in the structural model shown in panels (c) and (f). Very recently, a paracrystalline sample of diamond has been synthesized.[82] This discovery, along with the synthesis and characterization of paracrystalline

graphene[12,83] suggests that the landscape of disorder in carbon may have a link between "fully" amorphous and crystalline. This is further reflected in the NN energy distributions in Fig. 4, where panel (g) clearly shows a single distribution, panel (i) a bimodal distribution, and panel (h) a small hint of one. These ML energy distributions may provide a quantitative distinction between CRN and paracrystalline graphene, where a prominent bimodal distribution indicates the latter and the lack thereof indicates the former.

Quantifying disorder

We show further, quantitative structural indicators in Fig. 5. The bond-angle distribution (panel a) characterizes medium-range order. The local-energy-based structures (gray) display a wide range of bond angles, reflecting the relatively high level of disorder. The absence of a clear peak at 120° also emphasizes the reduction of medium-range order as is characteristic of CRNs. The NN curves (green) are narrower and centered around 120°, consistent with locally crystal-like environments. A shoulder peak at $\approx 109°$ shows the strong presence of 5-membered rings. The Toh *et al.* structure has a pronounced peak at 120° with a smaller shoulder at $\approx 109°$, indicative of the larger degree of locally crystal-like order. The bond angles in the 200-atom models (solid lines) agree well with those in the 612-atom ones (dashed).

The count of shortest-path rings is another metric for medium-range order. Ring counts for the local-energy MC runs reflect the large disorder, showing more 5- and 7-membered rings than in the other structures. The structures from the NN runs have ring counts centered around 6, as for the Toh *et al.* structure, with a larger count suggesting increased crystallinity in the latter case. As with the bond-angle distributions (Fig. 5a), the ring statistics for the 200-atom *versus* 612-atom structures are in close agreement within each other, within the standard deviation of the values for the former (Fig. 5a).

SOAP is a structural similarity metric for atomic environments.[61] 2D plots can reveal correlations between SOAP similarity on the one hand, and locally averaged energies on the other hand.[42] Fig. 5d shows the distribution of locally averaged energies (up to NN) and SOAP similarity to cG. The wider the spread in both axes, the greater the structural disorder. Comparing local-energy- with NN-energy-based data, both KDE curves are in close agreement, being skewed slightly toward the left (SOAP) and to higher energies for the runs using local energies only. This is expected, since the ML model predicts relatively low energies for the SW pair (Fig. 2), allowing structures to become more disordered. When averaging over NN environments, the distribution shifts and narrows as we observe fewer highly disordered environments. Data for the structure from Toh *et al.*[12] are provided to show how ordered this structure is. Defective environments are clearly identified with a series of data points at higher energies. As with Fig. 4, there is a clear bimodal distribution in the KDE curve. As seen for the structural indicators characterized in Fig. 5a and b, the dashed lines representing the 612-atom structures agree well with the set of separate 200-atom structures.

Fig. 5 Local structure and stability in models of amorphous graphene. Results for 25 local-energy-based ($\beta = 50.0$ eV^{-1}, grey) and NN-energy-based ($\beta = 2.0$ eV^{-1}, green) runs, *i.e.*, for 25×200 atoms each, and for the 612-atom structures based on local ($\beta = 50.0$ eV^{-1}, gray dashed) and NN ($\beta = 2.0$ eV^{-1}, green dashed) energies are given. The structural model from Toh *et al.* is also analyzed (orange, 610 atoms). Data from the 612-atom and Toh *et al.* structures are scaled arbitrarily to fit. (a) Bond-angle distributions for the different structures. (b) Ring statistics for the different structures. (d) An NN energy *vs.* SOAP plot, following ref. 42. Kernel density estimates (KDEs) are used to show distributions of properties. The bandwidth is determined following Scott's rule,[81] with a grid size of 200. Additionally, panels (c) and (e) show plots of network-topology metrics, namely, Lemaître's law and assortativity, respectively.

In addition to the well-established ring statistics and the SOAP similarity, we analyzed the structural models with two topological metrics typically used in network theory, *viz.* Lemaître's law[72] and the assortativity metric.[73] Lemaître's law connects the second moment of the ring distribution ($\mu_2 = \langle k^2 \rangle - \langle k \rangle^2$, where k is the ring size) with the fraction of six-membered rings, p_6. Since we impose the constraint that all atoms must be 3-fold connected, it follows that the distribution can be explained well by a single maximum-entropy distribution leading to a characteristic curve.[72] As $\lim_{p_6 \to 0.6}$ from $p_6 = 1$, μ_2 increases linearly as $1 - p_6$ in the region of $p_6 \gtrsim 0.6$. This line is extended (red line) to show that most of the points from both local ($\beta = 50.0$ eV^{-1}) and NN frameworks are located on this curve. This curve corresponds to the maximum-entropy solution if only 5-, 6-, and 7-membered rings are present, which is the case in both these frameworks. μ_2 then increases exponentially beyond this where structures from the local framework at $\beta = 2.0$ eV^{-1} are located. These structures show greater disorder and a wider spread of μ_2 values, as expected due to the

presence of 3- and 4-membered rings, where the maximum-entropy solution follows an exponential profile. For the 612-atom structures, the local framework ($\beta = 50.0$ eV^{-1}) based structure is more disordered than the NN-based one and the Toh *et al.* structure (orange). The latter does not yield a data-point on the curve as it has a (small) number of 2-coordinate sites and hence would be located on a different Lemaître curve.[35]

The assortativity, r, measures how likely a large ring is to be next to a smaller ring (disassortative, $r < 0$) or to other large rings (assortative, $r > 0$). In many physical systems, one finds a preference for disassortative configurations. This is, here, reflected in the range of r values (Fig. 5e). For the 200-atom systems, we find no discernible correlation between r and p_6. For the 612-atom ones, both the local and NN frameworks yield structures with similar r values. The Toh *et al.* structure has a slightly less negative r, implying a slight preference for a more random arrangement. This may be due to coordination defects in the structure skewing r toward zero.

Fig. 6 Signatures of structural disorder. (a) Vibrational density of states (VDOS) for the different structures and pristine graphene (black line). The gray arrows indicate trends in peak shifts. (b) Radial distribution functions, $g(r)$, for the different structures.

We finally considered properties that can be compared to previous or future experiments. Vibrational properties of graphene are readily characterized by Raman spectroscopy[84,85] and computations,[26,86] and we show predicted VDOS in Fig. 6a. For pristine graphene (as reference), we obtain a characteristic peak at 1510 cm^{-1} and a smaller one at 1410 cm^{-1}, in close agreement with previous computations[26,86] and experimental spectra.[84,85] As disorder increases, the 1510 cm^{-1} peak shifts to higher frequencies, as shown in Fig. 6a: to about 1520, 1570, and finally 1610 cm^{-1}.

The peaks at 1510–1610 cm^{-1} seen in Fig. 6a for the disordered structures are close to Raman data by Toh *et al.*, with the corresponding peak at 1558 cm^{-1}.[12] The blue shift of the 1510 cm^{-1} peak with increasing disorder is a well-known effect and has been observed in VDOS and Raman spectra.[26,86] Furthermore, we observe an "amorphous" peak in the acoustic (low-frequency) region, at 814–836 cm^{-1}, red-shifting with increasing disorder.

Another way of validating against experiment is to evaluate the radial distribution function (RDF) and X-ray diffraction (XRD) patterns: in disordered structures, characteristic peaks will become wider and less intense.[12,87] However, when comparing amorphous structures, differences in RDF and XRD are subtle:[18,87] in Fig. 6b, the second RDF peak at 2.41 Å broadens with increasing disorder, but this effect is subdued for the remaining peaks at 2.82 and 3.62 Å for the local and NN

structures. The Toh *et al.* structure has longer-range order, reflected in more prominent RDF peaks at greater distances (4.23 and 4.93 Å). The local and NN structures have peak locations comparable to that of the pair correlation function recreated from TEM in ref. 12.

Conclusions

Atomic energies predicted by the GAP ML framework can be used to drive Monte-Carlo structural exploration in principle. We have shown this by creating structural models of amorphous graphene, one of the prototypical disordered systems in physics and chemistry. We found that using (only) ML atomic energies leads to structures resembling continuous random networks, whereas including nearest-neighbor energies tends to drive the simulations toward some degree of paracrystalline order. We suggest that histograms of local energies, as shown in Fig. 4, can give insight into the degree of "randomness" in different amorphous networks: deviations from random order are visible as peaks at either low (paracrystalline) or high (coordination defects) energy.

The demonstrated ability to use ML local energies in MC annealing indicates potential for future research on amorphous materials. Describing local environments using ML methods can provide insight into the relation between atomic structure and energetics, and therefore structural stability, in amorphous materials—including amorphous carbon, which has emerging applications in biosensing[88] or batteries.[89] The fact that NN-averaged energies yield reasonable, partly paracrystalline structural models may be attributed to the fact that they provide "smoothing" over the variance in local atomic energies. This finding is consistent with earlier findings for the electronic DOS[43,44] and might have wider consequences for ML predictions of local properties, which are yet to be fully explored.

Data availability

Author contributions

Z. E.-M. developed the Monte-Carlo protocols and carried out the computational work. V. L. D. initiated and coordinated the study. All authors made substantial contributions to data analysis and discussions. Z. E.-M. and V. L. D. wrote the paper with input from M. W., and all authors approved the final version.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

References

1 A. C. Wright, The great crystallite *versus* random network controversy: A personal perspective, *Int. J. Appl. Glass Sci.*, 2014, **5**, 31–56.

2 D. Shi, Z. Guo and N. Bedford, 10 - Nanoenergy Materials, in *Nanomaterials and Devices*, William Andrew Publishing, Oxford, 2015, pp. 255–291.

3 J. Zhao, G. Zhu, W. Huang, Z. He, X. Feng, Y. Ma, X. Dong, Q. Fan, L. Wang, Z. Hu, Y. Lü and W. Huang, Synthesis of large-scale undoped and nitrogen-doped amorphous graphene on mgo substrate by chemical vapor deposition, *J. Mater. Chem.*, 2012, **22**, 19679–19683.

4 M. L. Gallo and A. Sebastian, An overview of phase-change memory device physics, *J. Phys. D: Appl. Phys.*, 2020, **53**, 213002.

5 J. Kotakoski, A. V. Krasheninnikov, U. Kaiser and J. C. Meyer, From point defects in graphene to two-dimensional amorphous carbon, *Phys. Rev. Lett.*, 2011, **106**, 105505.

6 P. Y. Huang, S. Kurasch, A. Srivastava, V. Skakalova, J. Kotakoski, A. V. Krasheninnikov, R. Hovden, Q. Mao, J. C. Meyer, J. Smet, D. A. Muller and U. Kaiser, Direct imaging of a two-dimensional silica glass on graphene, *Nano Lett.*, 2012, **12**, 1081–1086.

7 P. Y. Huang, S. Kurasch, J. S. Alden, A. Shekhawat, A. A. Alemi, P. L. McEuen, J. P. Sethna, U. Kaiser and D. A. Muller, Imaging atomic rearrangements in two-dimensional silica glass: Watching silica's dance, *Science*, 2013, **342**, 224–227.

8 Z. Yang, J. Hao, S. Yuan, S. Lin, H. M. Yau, J. Dai and S. P. Lau, Field-effect transistors based on amorphous black phosphorus ultrathin films by pulsed laser deposition, *Adv. Mater.*, 2015, **27**, 3748–3754.

9 W.-J. Joo, J.-H. Lee, Y. Jang, S.-G. Kang, Y.-N. Kwon, J. Chung, S. Lee, C. Kim, T.-H. Kim, C.-W. Yang, U. J. Kim, B. L. Choi, D. Whang and S.-W. Hwang, Realization of continuous Zachariasen carbon monolayer, *Sci. Adv.*, 2017, **3**, e1601821.

10 Z. Yang, J. Hao and S. P. Lau, Synthesis, properties, and applications of 2D amorphous inorganic materials, *J. Appl. Phys.*, 2020, **127**, 220901.

11 S. Hong, C.-S. Lee, M.-H. Lee, Y. Lee, K. Y. Ma, G. Kim, S. I. Yoon, K. Ihm, K.-J. Kim, T. J. Shin, S. W. Kim, E.-c. Jeon, H. Jeon, J.-Y. Kim, H.-I. Lee, Z. Lee, A. Antidormi, S. Roche, M. Chhowalla, H.-J. Shin and H. S. Shin, Ultralow-dielectric-constant amorphous boron nitride, *Nature*, 2020, **582**, 511–514.

12 C.-T. Toh, H. Zhang, J. Lin, A. S. Mayorov, Y.-P. Wang, C. M. Orofeo, D. B. Ferry, H. Andersen, N. Kakenov, Z. Guo, I. H. Abidi, H. Sims, K. Suenaga, S. T. Pantelides and B. Özyilmaz, Synthesis and properties of free-standing monolayer amorphous carbon, *Nature*, 2020, **577**, 199–203.

13 J. Tersoff, Empirical interatomic potential for carbon, with applications to amorphous carbon, *Phys. Rev. Lett.*, 1988, **61**, 2879–2882.

14 G. Galli, R. M. Martin, R. Car and M. Parrinello, Structural and electronic properties of amorphous carbon, *Phys. Rev. Lett.*, 1989, **62**, 555–558.

15 D. A. Drabold, P. A. Fedders and P. Stumm, Theory of diamond like amorphous carbon, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1994, **49**, 16415–16422.

16 N. A. Marks, D. R. McKenzie, B. A. Pailthorpe, M. Bernasconi and M. Parrinello, Microscopic structure of tetrahedral amorphous carbon, *Phys. Rev. Lett.*, 1996, **76**, 768–771.

17 D. G. McCulloch, D. R. McKenzie and C. M. Goringe, Ab initio simulations of the structure of amorphous carbon, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2000, **61**, 2349–2355.

18 A. Kumar, M. Wilson and M. F. Thorpe, Amorphous graphene: a realization of Zachariasen's glass, *J. Phys.: Condens. Matter*, 2012, **24**, 485003.

19 D. R. Robinson and M. Wilson, The liquid↔amorphous transition and the high pressure phase diagram of carbon, *J. Phys.: Condens. Matter*, 2013, **25**, 155101.

20 H.-P. Kaukonen and R. M. Nieminen, Molecular-dynamics simulation of the growth of diamond like films by energetic carbon-atom beams, *Phys. Rev. Lett.*, 1992, **68**, 620–623.

21 N. A. Marks, Thin film deposition of tetrahedral amorphous carbon: a molecular dynamics study, *Diamond Relat. Mater.*, 2005, **14**, 1223–1231.

22 M. A. Caro, V. L. Deringer, J. Koskinen, T. Laurila and G. Csányi, Growth mechanism and origin of high sp^3 content in tetrahedral amorphous carbon, *Phys. Rev. Lett.*, 2018, **120**, 166101.

23 M. A. Caro, G. Csányi, T. Laurila and V. L. Deringer, Machine learning driven simulated deposition of carbon films: From low-density to diamond like amorphous carbon, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2020, **102**, 174201.

24 R. Thapa, C. Ugwumadu, K. Nepal, J. Trembly and D. A. Drabold, Ab initio simulation of amorphous graphite, *Phys. Rev. Lett.*, 2022, **128**, 236402.

25 B. Bhattarai, A. Pandey and D. A. Drabold, Evolution of amorphous carbon across densities: An inferential study, *Carbon*, 2018, **131**, 168–174.

26 B. Bhattarai, P. Biswas, R. Atta-Fynn and D. A. Drabold, Amorphous graphene: a constituent part of low density amorphous carbon, *Phys. Chem. Chem. Phys.*, 2018, **20**, 19546–19551.

27 F. Wooten, K. Winer and D. Weaire, Computer generation of structural models of amorphous Si and Ge, *Phys. Rev. Lett.*, 1985, **54**, 1392–1395.

28 A. J. Stone and D. J. Wales, Theoretical studies of icosahedral C_{60} and some related species, *Chem. Phys. Lett.*, 1986, **128**, 501–503.

29 J. C. Meyer, C. Kisielowski, R. Erni, M. D. Rossell, M. F. Crommie and A. Zettl, Direct imaging of lattice atoms and topological defects in graphene membranes, *Nano Lett.*, 2008, **8**, 3582–3586.

30 J. Ma, D. Alfè, A. Michaelides and E. Wang, Stone-Wales defects in graphene and other planar sp^2-bonded materials, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2009, **80**, 033407.

31 F. Banhart, J. Kotakoski and A. V. Krasheninnikov, Structural defects in graphene, *ACS Nano*, 2011, **5**, 26–41.

32 F. L. Thiemann, P. Rowe, A. Zen, E. A. Müller and A. Michaelides, Defect-dependent corrugation in graphene, *Nano Lett.*, 2021, **21**, 8143–8150.

33 V. Kapko, D. A. Drabold and M. F. Thorpe, Electronic structure of a realistic model of amorphous graphene, *Phys. Status Solidi B*, 2010, **247**, 1197–1200.

34 F. D'Ambrosio, J. Barkema and G. T. Barkema, Efficient structural relaxation of polycrystalline graphene models, *Nanomaterials*, 2021, **11**, 1242.

35 D. Ormrod Morley and M. Wilson, Controlling disorder in two-dimensional networks, *J. Phys.: Condens. Matter*, 2018, **30**, 50LT02.

36 J. Behler, First principles neural network potentials for reactive simulations of large molecular and condensed systems, *Angew. Chem., Int. Ed.*, 2017, **56**, 12828–12840.

37 V. L. Deringer, M. A. Caro and G. Csányi, Machine learning interatomic potentials as emerging tools for materials science, *Adv. Mater.*, 2019, **31**, 1902765.

38 P. Friederich, F. Häse, J. Proppe and A. Aspuru-Guzik, Machine-learned potentials for next-generation matter simulations, *Nat. Mater.*, 2021, **20**, 750–761.

39 J. Behler and M. Parrinello, Generalized neural-network representation of high-dimensional potential-energy surfaces, *Phys. Rev. Lett.*, 2007, **98**, 146401.

40 A. P. Bartók, M. C. Payne, R. Kondor and G. Csányi, Gaussian approximation potentials: The accuracy of quantum mechanics, without the electrons, *Phys. Rev. Lett.*, 2010, **104**, 136403.

41 V. L. Deringer, C. J. Pickard and G. Csányi, Data-driven learning of total and local energies in elemental boron, *Phys. Rev. Lett.*, 2018, **120**, 156001.

42 N. Bernstein, B. Bhattarai, G. Csányi, D. A. Drabold, S. R. Elliott and V. L. Deringer, Quantifying chemical structure and machine-learned atomic energies in amorphous and liquid silicon, *Angew. Chem., Int. Ed.*, 2019, **58**, 7057–7061.

43 C. Ben Mahmoud, A. Anelli, G. Csányi and M. Ceriotti, Learning the electronic density of states in condensed matter, *Phys. Rev. B*, 2020, **102**, 235130.

44 V. L. Deringer, N. Bernstein, G. Csányi, C. Ben Mahmoud, M. Ceriotti, M. Wilson, D. A. Drabold and S. R. Elliott, Origins of structural and electronic transitions in disordered silicon, *Nature*, 2021, **589**, 59–64.

45 X. Song and C. Deng, Atomic energy in grain boundaries studied by machine learning, *Phys. Rev. B*, 2022, **6**, 043601.

46 M. Eckhoff and J. Behler, From molecular fragments to the bulk: Development of a neural network potential for MOF-5, *J. Chem. Theory Comput.*, 2019, **15**, 3793–3809.

47 V. L. Deringer and G. Csányi, Machine learning based interatomic potential for amorphous carbon, *Phys. Rev. B*, 2017, **95**, 094203.

48 N. Chetty and R. M. Martin, First-principles energy density and its applications to selected polar surfaces, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1992, **45**, 6074–6088.

49 D. W. Brenner, Empirical potential for hydrocarbons for use in simulating the chemical vapor deposition of diamond films, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1990, **42**, 9458–9471.

50 D. W. Brenner, O. A. Shenderova, J. A. Harrison, S. J. Stuart, B. Ni and S. B. Sinnott, A second-generation reactive empirical bond order (REBO) potential energy expression for hydrocarbons, *J. Phys.: Condens. Matter*, 2002, **14**, 783–802.

51 S. J. Stuart, A. B. Tutein and J. A. Harrison, A reactive potential for hydrocarbons with intermolecular interactions, *J. Chem. Phys.*, 2000, **112**, 6472–6486.

52 J. H. Los and A. Fasolino, Intrinsic long-range bond-order potential for carbon: Performance in Monte Carlo simulations of graphitization, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2003, **68**, 024107.

53 N. A. Marks, Generalizing the environment-dependent interaction potential for carbon, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2000, **63**, 035401.

54 N. A. Marks, Modelling diamond-like carbon with the environment-dependent interaction potential, *J. Phys.: Condens. Matter*, 2002, **14**, 2901–2927.

55 J. F. Justo, M. Z. Bazant, E. Kaxiras, V. V. Bulatov and S. Yip, Interatomic potential for silicon defects and disordered phases, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1998, **58**, 2539–2550.

56 M. Z. Bazant, E. Kaxiras and J. F. Justo, Environment-dependent interatomic potential for bulk silicon, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1997, **56**, 8542–8552.

57 C. de Tomas, I. Suarez-Martinez and N. A. Marks, Graphitization of amorphous carbons: A comparative study of interatomic potentials, *Carbon*, 2016, **109**, 681–693.

58 C. de Tomas, I. Suarez-Martinez, F. Vallejos-Burgos, M. J. López, K. Kaneko and N. A. Marks, Structural prediction of graphitization and porosity in carbide-derived carbons, *Carbon*, 2017, **119**, 1–9.

59 T. B. Shiell, D. G. McCulloch, D. R. McKenzie, M. R. Field, B. Haberl, R. Boehler, B. A. Cook, C. de Tomas, I. Suarez-Martinez, N. A. Marks and J. E. Bradby, Graphitization of glassy carbon after compression at room temperature, *Phys. Rev. Lett.*, 2018, **120**, 215701.

60 V. L. Deringer, A. P. Bartók, N. Bernstein, D. M. Wilkins, M. Ceriotti and G. Csányi, Gaussian process regression for materials and molecules, *Chem. Rev.*, 2021, **121**, 10073–10141.

61 A. P. Bartók, R. Kondor and G. Csányi, On representing chemical environments, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2013, **87**, 184115.

62 G. Yang, L. Li, W. B. Lee and M. C. Ng, Structure of graphene and its disorders: a review, *Sci. Technol. Adv. Mater.*, 2018, **19**, 613–648.

63 D. R. Cooper, B. D'Anjou, N. Ghattamaneni, B. Harack, M. Hilke, A. Horth, N. Majlis, M. Massicotte, L. Vandsburger, E. Whiteway and V. Yu, Experimental review of graphene, *ISRN Condens. Matter Phys.*, 2012, 501686.

64 Y. Li, F. Inam, A. Kumar, M. F. Thorpe and D. A. Drabold, Pentagonal puckering in a sheet of amorphous graphene, *Phys. Status Solidi B*, 2011, **248**, 2082–2086.

65 A. Hjorth Larsen, J. J. Mortensen, J. Blomqvist, I. E. Castelli, R. Christensen, M. Dułak, J. Friis, M. N. Groves, B. Hammer, C. Hargus, E. D. Hermes, P. C. Jennings, P. B. Jensen, J. Kermode, J. R. Kitchin, E. L. Kolsbjerg, J. Kubal, K. Kaasbjerg, S. Lysgaard, J. B. Maronsson, T. Maxson, T. Olsen, L. Pastewka, A. Peterson, C. Rostgaard, J. Schiøtz, O. Schütt, M. Strange, K. S. Thygesen, T. Vegge, L. Vilhelmsen, M. Walter, Z. Zeng and K. W. Jacobsen, The atomic simulation environment—a Python library for working with atoms, *J. Phys.: Condens. Matter*, 2017, **29**, 273002.

66 G. Csányi, S. Winfield, J. R. Kermode, A. De Vita, A. Comisso, N. Bernstein and M. C. Payne, Expressive programming for computational physics in Fortran 95+, *IoP Comput. Phys. Newsletter*, Spring, 2007.

67 J. R. Kermode, f90wrap: an automated tool for constructing deep python interfaces to modern Fortran codes, *J. Phys.: Condens. Matter*, 2020, **32**, 305901.

68 A. P. Thompson, H. M. Aktulga, R. Berger, D. S. Bolintineanu, W. M. Brown, P. S. Crozier, P. J. in't Veld, A. Kohlmeyer, S. G. Moore, T. D. Nguyen, R. Shan, M. J. Stevens, J. Tranchida, C. Trott and S. J. Plimpton, Lammps - a flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales, *Comput. Phys. Commun.*, 2022, **271**, 108171.

69 R. L. C. Vink, A finite-temperature Monte Carlo algorithm for network forming materials, *J. Chem. Phys.*, 2014, **140**, 104509.

70 D. S. Franzblau, Computation of ring statistics for network models of solids, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1991, **44**, 4925–4930.

71 Matscipy, version 0.7.0 **https://github.com/libAtoms/matscipy**, 2022.

72 A. Gervois, J. P. Troadec and J. Lemaître, Universal properties of Voronoi tessellations of hard discs, *J. Phys. A: Math. Gen.*, 1992, **25**, 6169–6177.

73 M. E. J. Newman, Assortative mixing in networks, *Phys. Rev. Lett.*, 2002, **89**, 208701.

74 A. Stukowski, Visualization and analysis of atomistic simulation data with OVITO–the Open Visualization Tool, *Modell. Simul. Mater. Sci. Eng.*, 2010, **18**, 015012.

75 A. Togo and I. Tanaka, First principles phonon calculations in materials science, *Scr. Mater.*, 2015, **108**, 1–5.

76 J. P. Perdew and A. Zunger, Self-interaction correction to density-functional approximations for many-electron systems, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1981, **23**, 5048–5079.

77 J. P. Perdew, K. Burke and M. Ernzerhof, Generalized gradient approximation made simple, *Phys. Rev. Lett.*, 1996, **77**, 3865–3868.

78 J. P. Perdew, A. Ruzsinszky, G. I. Csonka, O. A. Vydrov, G. E. Scuseria, L. A. Constantin, X. Zhou and K. Burke, Restoring the density-gradient expansion for exchange in solids and surfaces, *Phys. Rev. Lett.*, 2008, **100**, 136406.

79 S. J. Clark, M. D. Segall, C. J. Pickard, P. J. Hasnip, M. J. Probert, K. Refson and M. Payne, First principles methods using CASTEP, *Z. Kristall.*, 2005, **220**, 567–570.

80 D. Landau and K. Binder, *A Guide to Monte Carlo Simulations in Statistical Physics*, Cambridge University Press, 2005.

81 D. W. Scott, *Multivariate density estimation: theory, practice, and visualization*, John Wiley & Sons, 2015.

82 H. Tang, X. Yuan, Y. Cheng, H. Fei, F. Liu, T. Liang, Z. Zeng, T. Ishii, M.-S. Wang, T. Katsura, H. Sheng and H. Gou, Synthesis of paracrystalline diamond, *Nature*, 2021, **599**, 605–610.

83 Z. Liu, D. Panja and G. T. Barkema, Structural dynamics of polycrystalline graphene, *Phys. Rev. E*, 2022, **105**, 044116.

84 A. C. Ferrari and J. Robertson, Interpretation of Raman spectra of disordered and amorphous carbon, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2000, **61**, 14095–14107.

85 A. C. Ferrari, J. C. Meyer, V. Scardaci, C. Casiraghi, M. Lazzeri, F. Mauri, S. Piscanec, D. Jiang, K. S. Novoselov, S. Roth and A. K. Geim, Raman spectrum of graphene and graphene layers, *Phys. Rev. Lett.*, 2006, **97**, 187401.

86 S. N. Shirodkar and U. V. Waghmare, Electronic and vibrational signatures of Stone-Wales defects in graphene: First-principles analysis, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2012, **86**, 165401.

87 Y. Wang, Z. Fan, P. Qian, T. Ala-Nissila and M. A. Caro, Structure and pore size distribution in nanoporous carbon, *Chem. Mater.*, 2022, **34**, 617–628.

88 T. Laurila, S. Sainio and M. A. Caro, Hybrid carbon based nanomaterials for electrochemical detection of biomolecules, *Prog. Mater. Sci.*, 2017, **88**, 499–594.

89 E. Olsson, J. Yu, H. Zhang, H.-M. Cheng and Q. Cai, Atomic-scale design of anode materials for alkali metal (Li/Na/K)-ion batteries: Progress and perspectives, *Adv. Energy Mater.*, 2022, **12**, 2200662.