Structure prediction of boron-doped graphene by machine learning

Thaer M. Dieb, Zhufeng Hou (iD), and Koji Tsuda (iD)

COLLECTIONS

Paper published as part of the special topic on Data-Enabled Theoretical Chemistry

ARTICLES YOU MAY BE INTERESTED IN

Structure prediction of boron-doped graphene by machine learning

Thaer M. Dieb,[1,2,3,a)] Zhufeng Hou,[2,a)] and Koji Tsuda[1,2,3,b)]

[1]*Graduate School of Frontier Sciences, The University of Tokyo, Kashiwa, Japan*
[2]*National Institute for Materials Science (NIMS), Tsukuba, Japan*
[3]*RIKEN, Center for Advanced Intelligence Project, Tokyo, Japan*

Heteroatom doping has endowed graphene with manifold aspects of material properties and boosted its applications. The atomic structure determination of doped graphene is vital to understand its material properties. Motivated by the recently synthesized boron-doped graphene with relatively high concentration, here we employ machine learning methods to search the most stable structures of doped boron atoms in graphene, in conjunction with the atomistic simulations. From the determined stable structures, we find that in the free-standing pristine graphene, the doped boron atoms energetically prefer to substitute for the carbon atoms at different sublattice sites and that the para configuration of boron-boron pair is dominant in the cases of high boron concentrations. The boron doping can increase the work function of graphene by 0.7 eV for a boron content higher than 3.1%. © *2018 Author(s).*
*All article content, except where otherwise noted, is licensed under a Creative Commons Attribution (CC BY) license (*http://creativecommons.org/licenses/by/4.0/*).* https://doi.org/10.1063/1.5018065

I. INTRODUCTION

Chemical composition modification of materials by incorporating additional elements is one of the commonly used approaches to tune the material properties. To achieve the desired material properties, not only the optimal composition but also the spatial distribution of incorporated atoms is demanded to be discovered. Because of the complexity introduced by the interplay between structural and chemical degrees of freedom, the atomic structure search of dopant atoms in the host material is very challenging, particularly for the non-diluted doping level.

Graphene, a single layer of sp^2-bonded carbon atoms arranged in a honeycomb lattice, is the most explored representative of two-dimensional (2D) materials.[1,2] However, the electrical conduction of pristine monolayer graphene cannot be switched off because of its zero band gap in the electronic density of states, limiting the range of potential applications.[2,3] The perfectly flat graphene is chemically inert and not of practical interest in the applications involving chemical reactions.[4] To expand the applications of graphene, many efforts have been made to modify graphene by chemical functionalization, chemical doping, structure engineering, and so on.[2,4–6] Among these approaches, incorporation of heteroatoms into graphene is shown to be a versatile method for controllable tuning of its physical and chemical properties.[5–10] Boron (B) and nitrogen (N) have attracted much great attention for chemical doping of graphene since they are neighboring to carbon (C) and their atomic radii are similar to that of carbon.[11,12] Compared with nitrogen, substitutional doping of boron in graphene was predicted to be more energetically favorable[13,14] and found to exhibit a p-type doping effect.[14,15] Recent experiments have

shown that boron can be embedded in graphene at higher concentrations (up to 20%).[10,14,16–18] It was also found that the hexagonal structure exists in boron carbides (B_xC_{1-x}), thin films with B contents less than 0.5.[19] For B_xC_{1-x}, thin films with B content at $x = 0.25$, an ordered layer structure was proposed.[20,21] These make boron-doped graphene (B-graphene) a unique and very attractive material from both fundamental and practical viewpoints. However, the local structural form of B-graphene with relatively high B concentration is not yet well understood. The spatial distribution of doped B in graphene is strongly affected by the synthesis methods and conditions. For the epitaxial B-graphene grown by chemical vapor deposition, the B dopants in graphene can be a preferential substitution of carbon in only one of the graphene sublattices[10,18] and be completely random,[22] depending on the metal substrate used. To understand the energetic stability and to find out possible ordered structures of B-graphene, we carry out a global search for the local structure configurations of the B dopants in graphene with a series of concentrations by means of the supercell model and the atomistic simulations.

In the case of substitutional doping of B for C in defect-free graphene, the number of possible structure configurations can be written as C_{n+m}^m if the symmetry is broken, where n and m are the numbers of C and B atoms, respectively. It is evident that the exhaustive search of all possible structure configurations for multiple B dopants in graphene would be prohibitive from the efficiency. To search the stable atomic structures of B-graphene more efficiently, in this work we propose to use a machine learning approach that utilizes Monte Carlo tree search,[23] which was recently utilized for materials and chemical design[24,25] with a rollout schema depending on Bayesian optimization (BO).[26] Based on the structure search of B-graphene, we discuss the performance of this proposed scheme of global optimization.

[a)]T. M. Dieb and Z. Hou contributed equally to this work.
[b)]tsuda@k.u-tokyo.ac.jp

The remainder of this paper is organized as follows. In Sec. II, we briefly introduce the machine learning methods used for the search of atomic structures of B-graphene. The details of computation setup are given in Sec. III. The results for the optimization performance of the machine learning methods, the stabilities of B-graphene, and the electronic structures of B-graphene are presented in Sec. IV. The conclusions are given in Sec. V.

II. MACHINE LEARNING BASED OPTIMIZATION METHOD

Determination of optimal material structure with certain quality metrics traditionally depended on the experience of domain knowledge experts and trial-and-error experiments. Several machine learning-based methods have been proposed to accelerate this process with as few experiments as possible.[24,27–31] In such methods, the structure prediction is often formulated as a selection of optimal solution from a candidates space that maximizes or minimizes a black-box function (usually the target property).[32,33] Experiments are commonly replaced by simulators such as first-principles calculations.

Currently available efficient methods such as Bayesian optimization (BO)[26,34] are not scalable enough. Due to its exceptional performance in computer Go game,[23,35] Monte Carlo tree search (MCTS)[23] has recently gained attention in material design[24] offering a superior scalability with efficiency trade-off.

MCTS employs a shallow search tree where a node is a possible assignment of an atom into a position in the structure. Iteratively, the tree expands towards the optimal solution in 4 steps: selection, expansion, simulation, and backpropagation. A full solution is obtained from this shallow tree using the random rollout (completion) technique in the simulation step. To increase the MCTS efficiency, we propose to use Bayesian learning to engineer the rollout[24] instead of

the random selection. We briefly discuss this rollout operation here. See the supplementary material for details on the proposed method. The source code is freely available at https://github.com/tsudalab/MDTS.

Bayesian optimization methods maintain a surrogate model of the black-box function, most commonly Gaussian process (GP). A pool of candidate S' is generated at a node of level p_ℓ where each data point represents a full structure of N positions, with positions $p_1, \ldots, p_\ell$ being determined along the path from the root to the selected node and positions $p_{\ell+1}, \ldots, p_N$ being randomly generated. Data points in S' are then vectors of binary values (0, 1). If no data point from S' is previously observed, GP starts with an initial set of randomly selected data points from S'. GP is updated as more data points are observed. An acquisition function is then used to determine the next optimal solution. Within computational budget, the selected solutions are returned (Fig. 1).

III. COMPUTATIONAL DETAILS

To study B-graphene, a supercell constructed by the 4×4 extension of the hexagonal unit cell of graphene was employed and the substitution of carbon by boron was considered. To avoid the spurious interaction between graphene layers, a vacuum thickness in the supercell was set to 20.0 Å. The in-plane lattice constant of graphene supercell was fixed at $4a_0$, where $a_0 = 2.464$ Å for the calculated lattice constant of graphene. Figure 2 shows the atomic structure of the graphene supercell used in the present study. The number of B dopants was considered from 1 to 10, which correspond to the boron concentration variation from 3.125% to 31.25%. The number of possible structure configurations for different numbers of doped B atoms in a 4×4 graphene supercell is summarized in Table I. Because of the large number of structure configurations for multiple B dopants ($n_B \geqslant 4$) in graphene, we employed a combinatorial
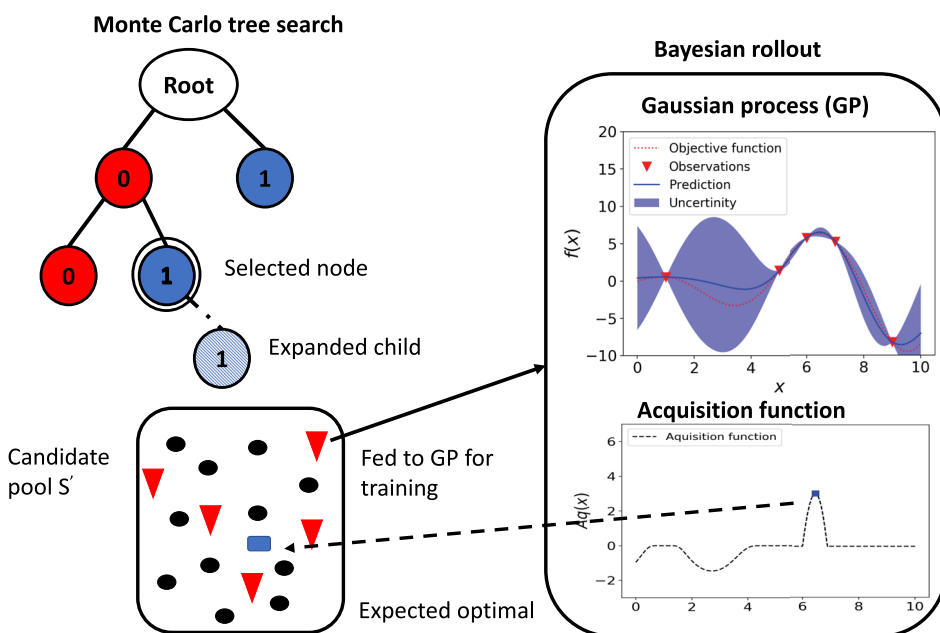

FIG. 1. Monte Carlo tree search (MCTS) with Bayesian rollout for binary atom assignment. MCTS uses a shallow tree search. At a certain iteration, only partial solution is determined in the tree. To obtain full solutions, a pool of full solution candidates is enumerated. GP is fed with previous observations if available or initial random selection (red triangles). An acquisition function is applied to determine the next optimal observation from the pool.
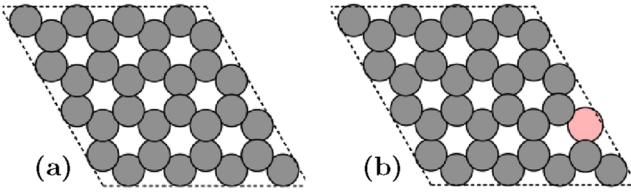
FIG. 2. The atomic structures of a 4 × 4 supercell of monolayer graphene (a) without and (b) with one doped B atom.

structure-generation approach for the local-level modeling of atomic substitutions and partial occupancies in crystals to reduce the number of potential configurations.[36] Our proposed algorithm is then applied to search the stable structure configurations of B-doped graphene. During the structure search, the atomic positions in each structure configuration were optimized by the limited-memory Broyden-Fletcher-Goldfarb-Shanno (L-BFGS) method.[37] The energy and atomic force were calculated by the classical force field method,[38] as implemented in the ATK-classical simulator of Atomistix ToolKit (ATK).[39,40] The interactions between the atoms in the system under study were described by the Tersoff bond order potentials developed by Matsunaga et al.[41] for C–B and C–C interactions, which have been widely used to study the B-doped carbon nanostructures. For the top ten structures with lower energies found by the global search methods for each case of multiple B dopants, we carried out density functional theory (DFT) calculations to optimize the atomic position and calculate the energy.

Density functional theory calculations have been performed with the Vienna *Ab initio* Simulation Package (VASP).[42,43] Perdew-Burke-Ernzerh of (PBE) exchange-correlation functional[44] within the generalized gradient approximations (GGA) was used in conjunction with the projector-augmented wave (PAW) method.[45,46] A plane wave basis set with an energy cutoff of 500 eV was used and a k-point mesh was sampled with a 11 × 11 × 1 Monkhorst-Pack scheme.[47] The criterions for total energy and force convergence were set to 10^{-5} eV and 10^{-2} eV/Å, respectively. Because of the fairly delocalized nature for the defect states induced by doped boron atoms,[48,49] the spin polarization of B-graphene is energetically unfavorable, which has been confirmed in the test calculations of spin-polarized DFT. Therefore, we report here the results from the non-spin-polarized calculations. Since the well-known underestimation of band gap in the GGA-PBE calculations due to self-interaction errors, we have also performed the HSE06 hybrid functional calculations[50] to study the electronic structures of B-graphene in some cases.

IV. RESULTS AND DISCUSSION

A. Obtaining optimal structures

As listed in Table I, the number of symmetry non-equivalent atomic structures for multiple doped B atoms considered in a 4 × 4 graphene supercell can reach several hundred thousands. For such large space of candidate structures of B-graphene, the atomic structure search was carried out by the MCTS with Bayesian rollout. We present in Fig. 3 the evolution of lattice energy during the structure search. In the cases considered here, we can see that the change of lattice energies for found optimal structures is less than 0.02 eV per supercell as the search reaches the predefined maximum steps, which indicates a good convergence for the structure search. For the cases of search space with thousand candidate structures, the finding of optimal structure takes 40% of the search space. But for the case of search space with candidate structures more than one hundred thousand, the optimal structure of B-graphene could be found within 2% of the search space. Therefore, our proposed optimization scheme might be very suitable for the larger search space that is hard to be treated, for example, the use of larger supercell size with a kept number of dopant atoms.

B. Structure stability of boron-doped graphene

To evaluate the stabilities of doped B atoms in graphene, we calculate their formation energies according to the following definition:

$$\Delta E_{\text{form}} = \frac{1}{m_{\text{B}}} \left[E_t(\text{BG}) - E_t(\text{PG}) - m_{\text{B}} \mu_{\text{B}} + m_{\text{B}} \mu_{\text{C}} \right], \quad (1)$$

where $E_t(\text{BG})$ and $E_t(\text{PG})$ are the total energies of graphene supercell with and without B dopants, m_{B} is the number of doped B atoms, and μ_{B} and μ_{C} are the chemical potentials of boron and carbon, which are taken as the total energies per atom of α-boron crystal and pristine monolayer graphene, respectively. The formation of single B substitution in the 4 × 4 graphene supercell is about 1.186 eV, which is in good agreement with the result (1.12 eV) reported in the previous study.[51]

It is well known that the honeycomb lattice of graphene consists of two sublattices, namely, A and B sublattices. To understand the occupancy preference of multiple-doped B atoms in graphene, it is necessary to first examine the interaction of two B substitutions in graphene. For this purpose, we calculate the interaction energy of two B substitutions as defined in the following equation:

$$\Delta E_{\text{int}} = E_t(2\text{B}) + E_t(\text{PG}) - 2E_t(1\text{B}), \quad (2)$$

TABLE I. The number of possible structure configurations for m_{B} (m_{B} = 1, ..., 10) doped B atoms in a 4 × 4 graphene supercell after merging symmetry.[36] The corresponding concentration of doped B in graphene is $c_{\text{B}} = m_{\text{B}}/32$.

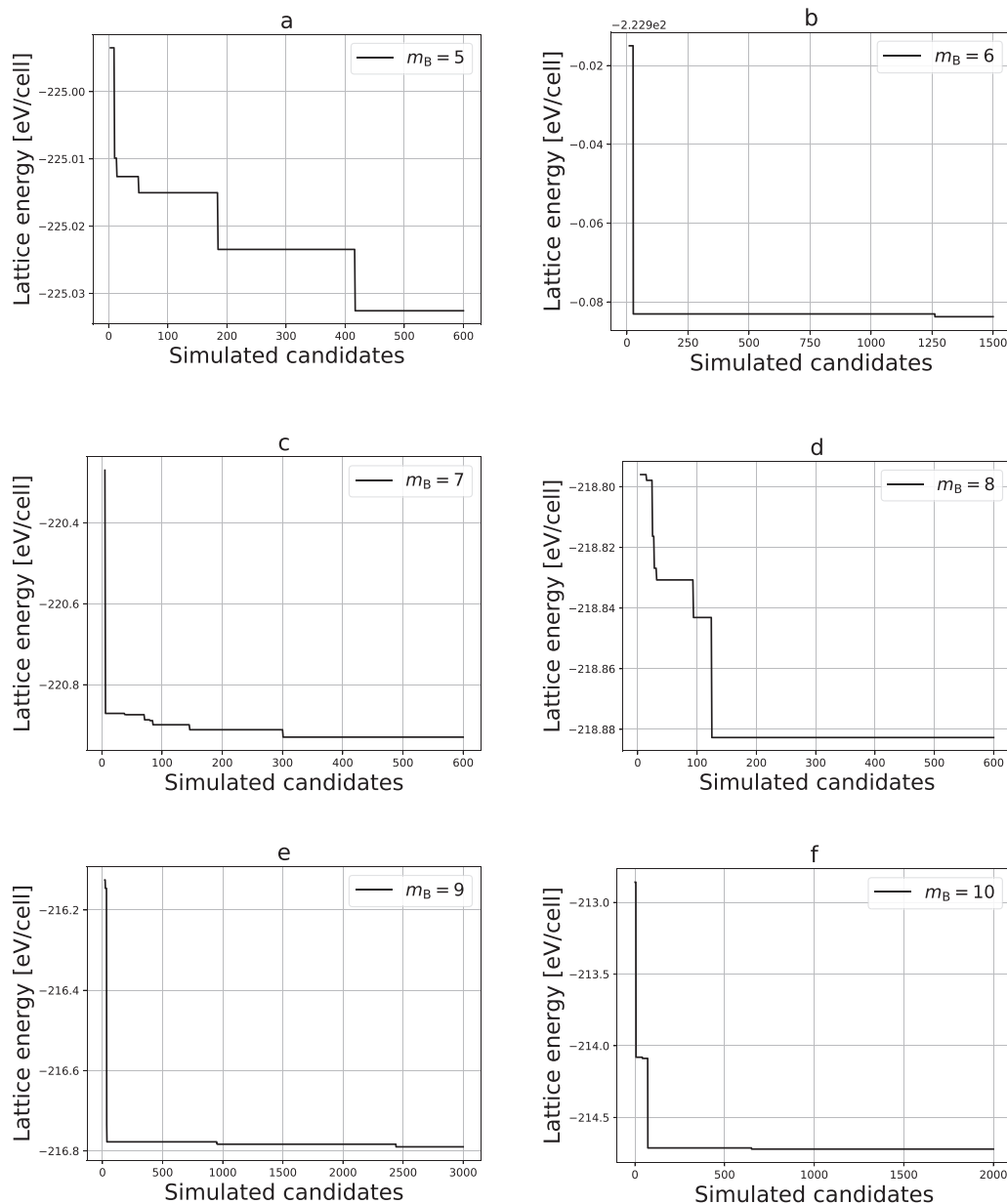| m_{B}	1	2	3	4	5	6	7	8	9	10
c_{B} (%)	3.125	6.25	9.375	12.5	15.625	18.75	21.875	25	28.125	31.25
$N_{\text{conf.}}$	1	8	37	241	1129	5002	17 929	55 817	147 362	338 741

FIG. 3. The evolution of lattice energy during the structure search of B-graphene by the optimization scheme based on MCTS with Bayesian rollout for the different numbers of doped B atoms in a 4 × 4 supercell: (a) $m_B = 5$, (b) $m_B = 6$, (c) $m_B = 7$, (d) $m_B = 8$, (e) $m_B = 9$, and (f) $m_B = 10$.

where E_t(PG), E_t(1B), and E_t(2B) are the total energies of graphene supercells with zero, one, and two doped B atoms, respectively. The positive (negative) sign of the interaction energy defined in Eq. (2) indicates two substitutional B dopants in graphene that repel (attract) each other. It also implies the increase (decreasing) in the formation energy of two B substitutions with respect to that of single B substitution. In the 4 × 4 graphene supercell considered here, the most stable configuration of two B substitutions is depicted in Fig. 4(a), where two doped B atoms occupy the different sublattices, and the calculated interaction energy is about 0.022 eV. This indicates that the interaction of two substitutional B atoms in graphene is repulsive. For two B substitutions at the first nearest-neighboring (NN) lattice sites (called ortho configuration) in graphene, they show the strongest repulsive interaction and the corresponding interaction energy is 1.262 eV. For two

B substitutions at the second and third NN lattice sites (called meta and para configurations, respectively) in graphene, their interaction energies are 0.507 eV and 0.075 eV, respectively. As the para configuration of the B-B pair is more stable than the meta configuration, it has been found in the chemically synthesized B-doped nanographene by using the boron-containing polycyclic aromatic hydrocarbon (PAH).[52,53] The B-graphene with such less stable meta configuration of B-B pair was synthesized by chemical vapor deposition (CVD) on a nickel or cobalt substrate that has strong interaction with graphene.[10,18]

The atomic structures for the most stable configurations of B-graphene at different B concentrations are depicted in Fig. 4. They have been confirmed in the calculations that take into account the relaxation of in-plane lattice constants. We have verified their stabilities in the calculations of
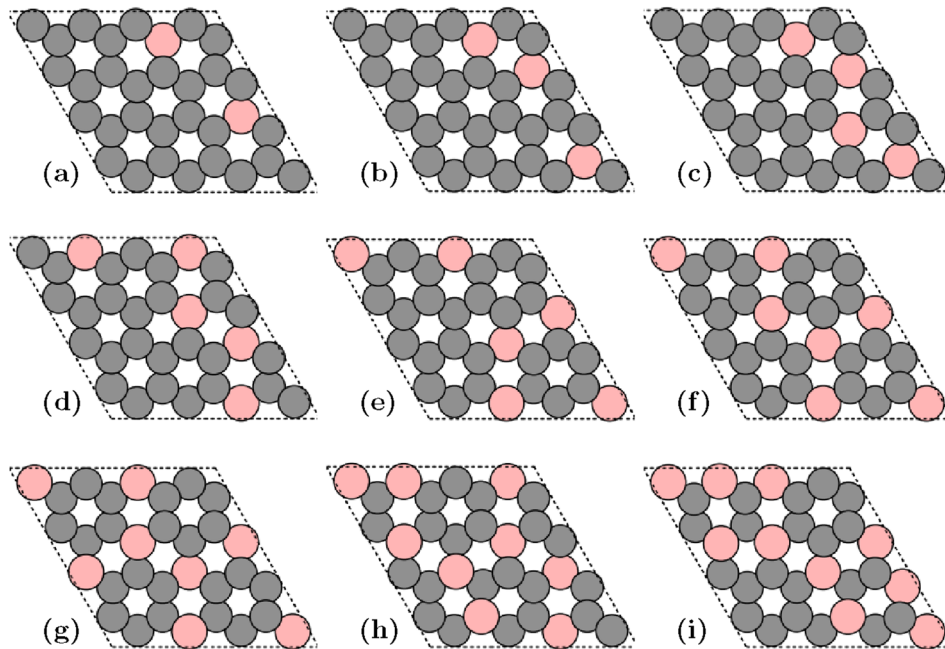
FIG. 4. The most stable structure configurations for doped B atoms in the 4×4 supercell of graphene at different concentrations. From (a) to (i), the number of doped B atoms changes from two to ten. The gray and red balls represent C and B atoms, respectively.

local-density approximations (LDAs) and meta-GGA. The detailed comparison of the results is given in the supplementary material. The corresponding formation energies of B substitutions are presented in Fig. 5. In these structure configurations, the para configuration of the B-B pair is dominant, indicating the substitutional B atoms prefer to occupy different sublattices in the free standing pristine graphene. The unbalanced sublattice doping observed in experiments may be attributed to the strong interaction of metal substrate and graphene.[10,18] In particular, it is noticed that the formation energy of four B substitutions is lower than those of three and five ones. In the case of eight B substitutions, which corresponds to BC_3, its formation energy is lower than those of seven and nine ones. For the most stable structure configuration of eight B substitutions, as shown in Fig. 4(g), all of the benzene-like C6 rings are separated by B-B pairs and thus its aromaticity is very high. This structure configuration has been proposed in the literature[20,21] for the synthesized BC_3 compound.
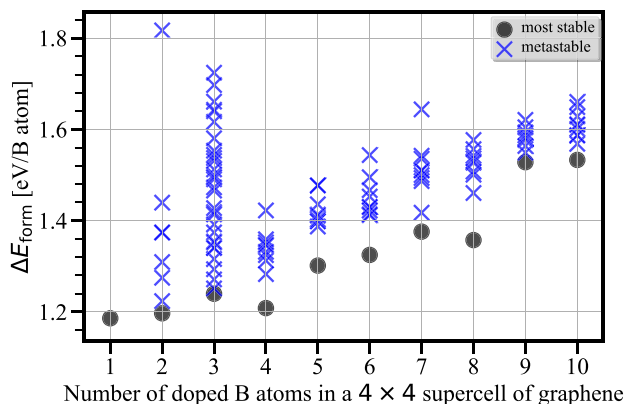

FIG. 5. The formation energy of doped B atom in a 4×4 supercell of graphene as a function of the number of doped B atoms.

C. Electronic structures of boron-doped graphene

To explore the electronic structure of graphene modified by B doping, we present in Fig. 6 the total density of states (DOS) of B-graphenes at a series of B concentrations. It is well known that the Fermi level (E_F) of perfect graphene coincides with the Dirac point. Because boron has one less electron than carbon, the B substitution leads to the shifting of E_F down into the valence bands. We also notice that the B doping opens up a band gap at the Dirac point of graphene, which is supported by the electrical conductivity measurement of B-graphene.[14] In the case of eight B atoms doped in the 4×4 graphene supercell, a band gap is open around the E_F and thus this structure of B-graphene is a semiconductor, which is consistent with the previous results.[54,55] The band gaps of BC_3 predicted by the GGA-PBE and HSE06 hybrid functional calculations in the present study are about 0.17 eV and 1.38 eV, respectively. To quantitatively characterize the doping level of carriers, we present the work function of B-graphene in Fig. 7. The work function of perfect monolayer graphene predicted by the GGA-PBE calculations in the present study is about 4.25 eV, which is in good agreement with the experiment results (4.272 eV for CVD-grown graphene measured by UV photoelectron spectroscopy[56] and 4.57 ± 0.05 eV for mechanically exfoliated graphene measured by a scanning Kelvin probe microscope,[57] respectively.). It also agrees well with the previous result (4.20 eV) of GGA-PBE calculations in the study of Lazar et al.[58] We can see that the boron doping increases the work function of graphene significantly. The similar trend was also found in the study of Lazar et al.[58] Considering the underestimation of band gap in the GGA-PBE calculations, we have also performed the HSE06 hybrid functional calculations to check the work functions of perfect monolayer graphene and B-graphene, which includes the cases of singe B substitution in 4×4 supercell and BC_3. The obtained work functions for these three cases are 4.23, 5.23, and 5.80 eV, respectively.
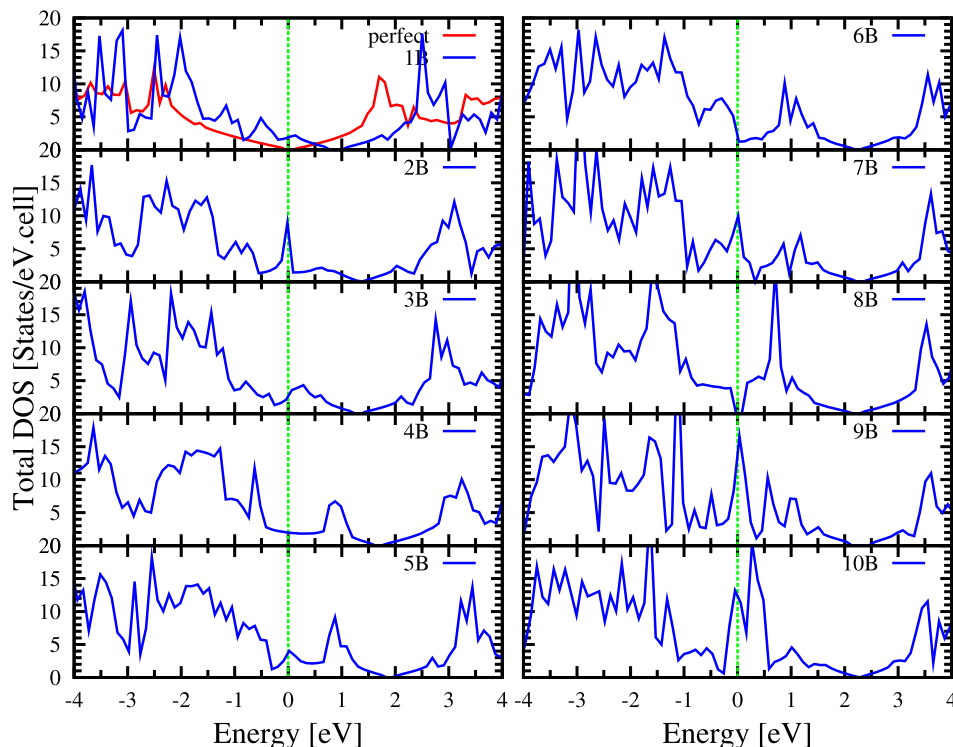
FIG. 6. The total density of states (DOS) of B-graphene with different numbers of doped B atoms in a 4 × 4 supercell for their atomic structures shown in Figs. 2 and 4. The Fermi level is set as zero and indicated by the dashed line.
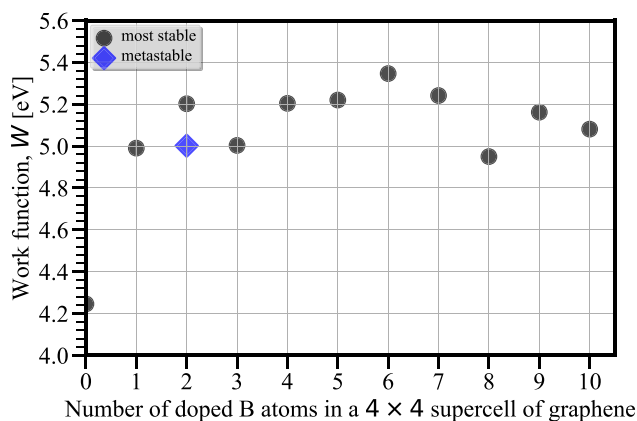

FIG. 7. The work function of B-graphene as a function of the number of doped B atoms in a 4 × 4 supercell. In the case of two B dopants, the work function for the metastable structure of B-B pair in the para configuration is also given.

Depending on the doped B concentration, the GGA-PBE calculations show that the work function of graphene can be tuned by 0.7 eV–1.0 eV.

V. CONCLUSION

In summary, we have employed the Monte Carlo tree search (MCTS) with Bayesian rollout to search the stable structures of B-graphene for the boron concentration up to 31.25%. Compared with the sole use of Bayesian optimization, this integrated optimization method shows a superiority of better scalability. Our results show that in the free-standing graphene, the doped boron atoms energetically prefer to substitute for the carbon atoms at different sublattices. For B-graphene with high boron concentration, the para B-B pairs

are dominant. Because of the repulsive interaction between boron substitutions in graphene, the doped boron would tend to be segregated. The doped boron can open a band gap at the Dirac point of graphene. Particularly in the concentration of doped B at 25%, namely, BC_3, the found stable structure exhibits semiconducting behavior. We also find that boron doping can lead to an increase of the work function of graphene in the range between 0.7 eV and 1.0 eV. Our results would be very helpful to further explore the application of B-graphene. In addition, the proposed method can be applied to multiple atom types assignment problems such as boron and nitrogen codoped graphene by allowing more than two branches in the search tree and using a descriptor based on one-hot encoding.

SUPPLEMENTARY MATERIAL

See supplementary material for the details on the proposed machine learning method and the detailed comparison of the results obtained by the different exchange-correlation functionals in DFT calculations.

ACKNOWLEDGMENTS

[1] A. K. Geim and K. S. Novoselov, Nat. Mater. **6**, 183 (2007).

[2] M. F. Craciun, I. Khrapach, M. D. Barnes, and S. Russo, J. Phys.: Condens. Matter **25**, 423201 (2013).

[3] P. Avouris, Nano Lett. **10**, 4285 (2010).

[4] A. Eftekhari and H. Garcia, Mater. Today Chem. **4**, 1 (2017).

[5] L. K. Putri, W.-J. Ong, W. S. Chang, and S.-P. Chai, Appl. Surf. Sci. **358**, 2 (2015).

[6] H. Liu, Y. Liu, and D. Zhu, J. Mater. Chem. **21**, 3335 (2011).

[7] X. Wang, G. Sun, P. Routh, D.-H. Kim, W. Huang, and P. Chen, Chem. Soc. Rev. **43**, 7067 (2014).

[8] J. Duan, S. Chen, M. Jaroniec, and S. Z. Qiao, ACS Catal. **5**, 5207 (2015).

[9] H. Cui, Z. Zhou, and D. Jia, Mater. Horiz. **4**, 7 (2017).

[10] D. Y. Usachov, A. V. Fedorov, A. E. Petukhov, O. Y. Vilkov, A. G. Rybkin, M. M. Otrokov, A. Arnau, E. V. Chulkov, L. V. Yashina, M. Farjam, V. K. Adamchuk, B. V. Senkovskiy, C. Laubschat, and D. V. Vyalikh, ACS Nano **9**, 7314 (2015).

[11] H. Wang, T. Maiyalagan, and X. Wang, ACS Catal. **2**, 781 (2012).

[12] S. Agnoli and M. Favaro, J. Mater. Chem. A **4**, 5002 (2016).

[13] L. S. Panchakarla, K. S. Subrahmanyam, S. K. Saha, A. Govindaraj, H. R. Krishnamurthy, U. V. Waghmare, and C. N. R. Rao, Adv. Mater. **21**, 4726 (2009).

[14] Y.-B. Tang, L.-C. Yin, Y. Yang, X.-H. Bo, Y.-L. Cao, H.-E. Wang, W.-J. Zhang, I. Bello, S.-T. Lee, H.-M. Cheng, and C.-S. Lee, ACS Nano **6**, 1970 (2012).

[15] W.-C. Yen, H. Medina, J.-S. Huang, C.-C. Lai, Y.-C. Shih, S.-M. Lin, J.-G. Li, Z. M. Wang, and Y.-L. Chueh, J. Phys. Chem. C **118**, 25089 (2014).

[16] W. Zhao, J. Gebhardt, K. Gotterbarm, O. Höfert, C. Gleichweit, C. Papp, A. Görling, and H.-P. Steinrück, J. Phys.: Condens. Matter **25**, 445002 (2013).

[17] J. Gebhardt, R. J. Koch, W. Zhao, O. Höfert, K. Gotterbarm, S. Mammadov, C. Papp, A. Görling, H.-P. Steinrück, and T. Seyller, Phys. Rev. B **87**, 155437 (2013).

[18] D. Y. Usachov, A. V. Fedorov, O. Y. Vilkov, A. E. Petukhov, A. G. Rybkin, A. Ernst, M. M. Otrokov, E. V. Chulkov, I. I. Ogorodnikov, M. V. Kuznetsov, L. V. Yashina, E. Y. Kataev, A. V. Erofeevskaya, V. Y. Voroshnin, V. K. Adamchuk, C. Laubschat, and D. V. Vyalikh, Nano Lett. **16**, 4535 (2016).

[19] I. Caretti, R. Gago, J. M. Albella, and I. Jiménez, Phys. Rev. B **77**, 174109 (2008).

[20] J. Kouvetakis, R. B. Kaner, M. L. Sattler, and N. Bartlett, J. Chem. Soc., Chem. Commun. **1986**, 1758.

[21] H. Yanagisawa, T. Tanaka, Y. Ishida, M. Matsue, E. Rokuta, S. Otani, and C. Oshima, Phys. Rev. Lett. **93**, 177003 (2004).

[22] L. Zhao, M. Levendorf, S. Goncher, T. Schiros, L. Pálová, A. Zabet-Khosousi, K. T. Rim, C. Gutiérrez, D. Nordlund, C. Jaye, M. Hybertsen, D. Reichman, G. W. Flynn, J. Park, and A. N. Pasupathy, Nano Lett. **13**, 4659 (2013).

[23] C. B. Browne, E. Powley, D. Whitehouse, S. M. Lucas, P. I. Cowling, P. Rohlfshagen, S. Tavener, D. Perez, S. Samothrakis, and S. Colton, IEEE Trans. Comput. Intell. AI Games **4**, 1 (2012).

[24] T. M. Dieb, S. Ju, K. Yoshizoe, Z. Hou, J. Shiomi, and K. Tsuda, Sci. Technol. Adv. Mater. **18**, 498 (2017).

[25] X. Yang, J. Zhang, K. Yoshizoe, K. Terayama, and K. Tsuda, Sci. Technol. Adv. Mater. **18**, 972 (2017).

[26] D. R. Jones, M. Schonlau, and W. J. Welch, J. Global Optim. **13**, 455 (1998).

[27] D. Reker and G. Schneider, Drug Discov. Today **20**, 458 (2015).

[28] A. R. Oganov and C. W. Glass, J. Chem. Phys. **124**, 244704 (2006).

[29] M. Ahmadi, M. Vogt, P. Iyer, J. Bajorath, and H. Fröhlich, J. Chem. Inf. Model. **53**, 553 (2013).

[30] S. Ju, T. Shiga, L. Feng, Z. Hou, K. Tsuda, and J. Shiomi, Phys. Rev. X **7**, 021024 (2017).

[31] R. Gómez-Bombarelli, D. K. Duvenaud, J. M. Hernández-Lobato, J. Aguilera-Iparraguirre, T. D. Hirzel, R. P. Adams, and A. Aspuru-Guzik, ACS Cent. Sci. **4**(2), 268 (2018).

[32] A. Seko, A. Togo, H. Hayashi, K. Tsuda, L. Chaput, and I. Tanaka, Phys. Rev. Lett. **115**, 205901 (2015).

[33] P. V. Balachandran, D. Xue, J. Theiler, J. Hogden, and T. Lookman, Sci. Rep. **6**, 19660 (2016).

[34] J. Snoek, H. Larochelle, and R. P. Adams, in *Advances in Neural Information Processing Systems 25*, edited by F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger (Curran Associates, Inc., 2012), pp. 2951–2959.

[35] D. Silver, A. Huang, C. J. Maddison, A. Guez, L. Sifre, G. van den Driessche, J. Schrittwieser, I. Antonoglou, V. Panneershelvam, M. Lanctot, S. Dieleman, D. Grewe, J. Nham, N. Kalchbrenner, I. Sutskever, T. Lillicrap, M. Leach, K. Kavukcuoglu, T. Graepel, and D. Hassabis, Nature **529**, 484 (2016).

[36] K. Okhotnikov, T. Charpentier, and S. Cadars, J. Cheminf. **8**, 17 (2016).

[37] J. Nocedal and S. J. Wright, *Numerical Optimization*, 2nd ed. (Springer, New York, 2006).

[38] J. Schneider, J. Hamaekers, S. T. Chill, S. Smidstrup, J. Bulin, R. Thesen, A. Blom, and K. Stokbro, Modell. Simul. Mater. Sci. Eng. **25**, 085007 (2017).

[39] See www.quantumwise.com for QuantumWise A/S, Atomistix ToolKit version 2016.01.

[40] M. Griebel and J. Hamaekers, Comput. Methods Appl. Mech. Eng. **193**, 1773 (2004).

[41] K. Matsunaga, C. Fisher, and H. Matsubara, Jpn. J. Appl. Phys., Part 2 **39**, L48 (2000).

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

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

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

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

[46] G. Kresse and D. Joubert, Phys. Rev. B **59**, 1758 (1999).

[47] H. J. Monkhorst and J. D. Pack, Phys. Rev. B **13**, 5188 (1976).

[48] K. E. Kweon and G. S. Hwang, Phys. Rev. B **82**, 195439 (2010).

[49] B. Zheng, P. Hermet, and L. Henrard, ACS Nano **4**, 4165 (2010).

[50] J. Heyd, G. E. Scuseria, and M. Ernzerhof, J. Chem. Phys. **124**, 219906 (2006).

[51] I. Choudhuri, N. Patra, A. Mahata, R. Ahuja, and B. Pathak, J. Phys. Chem. C **119**, 24827 (2015).

[52] C. Dou, S. Saito, K. Matsuo, I. Hisaki, and S. Yamaguchi, Angew. Chem., Int. Ed. **51**, 12206 (2012).

[53] S. Osumi, S. Saito, C. Dou, K. Matsuo, K. Kume, H. Yoshikawa, K. Awaga, and S. Yamaguchi, Chem. Sci. **7**, 219 (2016).

[54] D. Tomanek, R. M. Wentzcovitch, S. G. Louie, and M. L. Cohen, Phys. Rev. B **37**, 3134 (1988).

[55] X. Chen, Y. Yao, H. Yao, F. Yang, and J. Ni, Phys. Rev. B **92**, 174503 (2015).

[56] S. Bae, H. Kim, Y. Lee, X. Xu, J.-S. Park, Y. Zheng, J. Balakrishnan, T. Lei, H. Ri Kim, Y. I. Song, Y.-J. Kim, K. S. Kim, B. Özyilmaz, J.-H. Ahn, B. H. Hong, and S. Iijima, Nat. Nanotechnol. **5**, 574 (2010).

[57] Y.-J. Yu, Y. Zhao, S. Ryu, L. E. Brus, K. S. Kim, and P. Kim, Nano Lett. **9**, 3430 (2009).

[58] P. Lazar, R. Zboril, M. Pumera, and M. Otyepka, Phys. Chem. Chem. Phys. **16**, 14231 (2014).