

# Stochastic generation of complex crystal structures combining group and graph theory with application to carbon

Xizhi Shi,<sup>1,2</sup> Chaoyu He,<sup>1,2,\*</sup> Chris J. Pickard,<sup>3,4,†</sup> Chao Tang,<sup>1,2,‡</sup> and Jianxin Zhong<sup>1,2</sup>

<sup>1</sup>*Human Key Laboratory for Micro-Nano Energy Materials and Devices, Xiangtan University, Hunan 411105, P. R. China;*

<sup>2</sup>*School of Physics and Optoelectronics, Xiangtan University, Xiangtan 411105, China.*

<sup>3</sup>*Department of Materials Science & Metallurgy, University of Cambridge,  
27 Charles Babbage Road, Cambridge CB3 0FS, United Kingdom*

<sup>4</sup>*Advanced Institute for Materials Research, Tohoku University 2-1-1 Katahira, Aoba, Sendai, 980-8577, Japan*

(Dated: January 12, 2018)

A method is introduced to stochastically generate crystal structures with defined structural characteristics. Reasonable quotient graphs for symmetric crystals are constructed using a random strategy combined with space group and graph theory. Our algorithm enables the search for large-size and complex crystal structures with a specified connectivity, such as three-fold  $sp^2$  carbons, four-fold  $sp^3$  carbons, as well as mixed  $sp^2$ - $sp^3$  carbons. To demonstrate the method we randomly construct initial structures adhering to space groups from No.75 to No.230 and a range of lattice constants, and we identify 281 new  $sp^3$  carbon crystals. First-principles optimization of these structures show that most of them are dynamically and mechanically stable and are energetically comparable to those previously proposed. Some of the new structures can be considered as candidates to explain the experimental cold compression of graphite.

PACS numbers:

## I. INTRODUCTION

The prediction of crystal structures for a given composition is a central topic to the study of condensed matter<sup>1</sup>. Many methods have been developed, such as *ab initio* random structure searching<sup>2,3</sup>, evolutionary approaches<sup>4-6</sup>, particle swarm optimization methods<sup>7,8</sup>, periodic graph theory<sup>10?</sup>, metadynamics<sup>11,12</sup>, simulated annealing<sup>13,14</sup>, and basin or minima hopping methods<sup>15-17</sup>. Elemental carbon can form a large number of allotropes, such as graphite, cubic diamond, hexagonal diamond and other carbon phases with currently undetermined structures<sup>18-23</sup>. This is due to the propensity carbon to  $sp$ ,  $sp^2$  and  $sp^3$  hybridization, and it makes carbon a challenging target for crystal structure prediction. Over recent decades, many interesting aspects of elemental carbon have been explored. For example, much theoretical effort has been expended on the enumeration of hypothetical metastable carbon structures<sup>24-35</sup>, predicting potential super-hard carbon materials<sup>36-45</sup>, searching for possible super-dense carbon crystals<sup>46-48</sup>, solving the crystalline structures of the previously synthesized carbon phases<sup>49-59</sup>, determining the ultimate fate of carbon under extreme compression<sup>60</sup>, as well as designing direct-band gap carbons for solar cell application<sup>61,62</sup>. As documented in the Samara Carbon Allotrope Database (SACADA: <http://sacada.sctms.ru/> (accessed December 24, 2016).), there are at least 522 three-dimensional crystal structures that have been previously proposed for carbon. The number of potential carbon crystal structures increases rapidly<sup>63-66</sup>. In a very recent article, a large number of three-dimensional  $C_8$  carbon crystals<sup>67</sup> were proposed and it was noted that there are at least 460 articles concerning carbon structures in the literature.

In this paper, we describe a method for generating crystal structures with defined structural features. It is based on a random sampling strategy combined with space group and graph theory (RG<sup>2</sup>). We find that it efficiently yields reasonable quo-

tient graphs for symmetric crystals with specified coordination (such as  $sp^2$  carbons,  $sp^3$  carbons and mixed  $sp^2$ - $sp^3$  carbons) and unit cell. Using RG<sup>2</sup>, we systematically search for crystal structures of  $sp^3$  carbon in space groups from No.75 to No.230 with various lattice parameters and find 281 new  $sp^3$  carbon crystal structures. Our first principles results show that most of these new structures are dynamically and mechanically stable with comparable energetic stabilities to previously proposed carbon crystal structures. Some of them with low energy have been further found to be super-hard semiconductors. These results suggest that RG<sup>2</sup> is a powerful tool for the generation of crystal structures with defined geometrical features.

## II. METHOD

### A. Relation between crystal structure and quotient graph

As discussed by Wells<sup>68</sup>, each periodic crystal structure can be completely described by an infinite net. The translationally repeated part of the crystal structure can correspondingly represented as a finite quotient graph. Generally, each given quotient graph can be mapped back to a crystal structure through proper optimization. That is to say, we can predict crystalline structures through systematic enumeration of reasonable quotient graphs. Using such an idea, Winkler *et al* systematically enumerated many small  $sp^2$  carbon crystals with up to six atoms per unit cell in 2001<sup>30</sup>. In 2004, Strong *et al* applied this method to predict small  $sp^3$  carbon crystals containing four carbon atoms per unit cell and proposed eight new  $sp^3$  carbon crystals<sup>10</sup>. However, the method faces challenges in predicting large-size crystal structures. As the number of atoms per unit cell increases, the generation of possible quotient graphs and the embedding into three dimensions becomes more and more difficult. To the best of our knowledge, most of the pre-

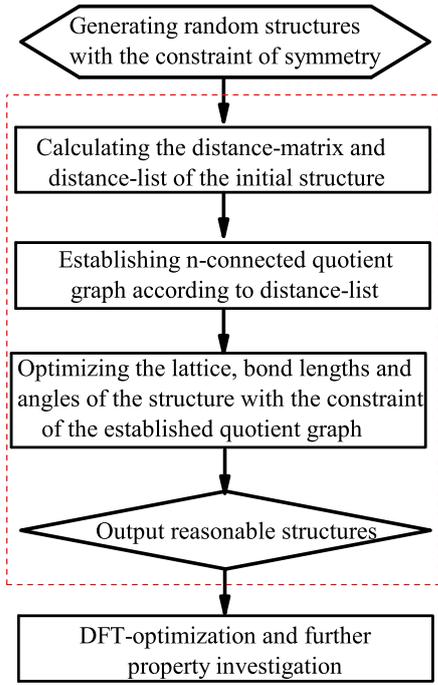


FIG. 1: The flow chart of our algorithm to generate n-connected crystalline network.

viously proposed large-size carbon crystals are generated by special techniques, such as segment recombination<sup>41,42,44</sup> and topological analysis<sup>69,70</sup>. Very recently, a systematic enumeration of carbon crystals with up to eight atoms per unit cell has been achieved through the artificial force induced reaction method (AFIR)<sup>67</sup>. The work suggests the construction of new quotient graphs from known equal-sized quotient graphs, thereby generating new crystals from known crystal structures. We propose a simple way to generate new feasible quotient graph for crystal structures with defined geometric features.

### B. Building the quotient graph and generating the structures

We have developed RG<sup>2</sup> to rapidly yield reasonable quotient graphs that will produce low energy carbon crystal structures. As shown in Fig. 1, we first generate a random initial structure in a given or randomly chosen space group. We do not initially worry about the quality of the structure and simply select the atomic positions so that the atoms are randomly and uniformly spread through the unit cell. We then give each atom in the central cell an index  $i$ , where ( $i = 1, N$ ), and calculate the distances between each  $i$ -atom and the other  $j$ -atoms in any of the  $hkl$ -cells, and so build a distance matrix for the

initial structure:

$$\begin{pmatrix} d_{11}^{hkl} & d_{12}^{hkl} & \dots & d_{1N}^{hkl} \\ \dots & \dots & \dots & \dots \\ \dots & \dots & d_{ij}^{hkl} & \dots \\ d_{N1}^{hkl} & d_{N2}^{hkl} & \dots & d_{NN}^{hkl} \end{pmatrix}^{DM} \quad (1)$$

where  $N$  is the number of atoms in the unit cell and  $d_{ij}^{hkl}$  is the distance between  $i$ -atom in the 000-cell and the  $j$ -atom in the  $hkl$ -cell. Based on the calculated distance matrix, we sort all the distances  $d_{ij}^{hkl}$ . We then build the bonds one by one by working through the unsaturated pairs, from small to large separations, until all the atoms are saturated or all the pairs have been considered. For each new bond, we record the corresponding cell-index  $[hkl]$  and keep track of the number of bonds for each atom. In such a way we construct the quotient graph for the structure as:

$$\begin{pmatrix} Q_{11} & Q_{12} & \dots & Q_{1N} \\ Q_{21} & Q_{22} & \dots & Q_{2N} \\ \dots & \dots & Q_{ij} & \dots \\ Q_{N1} & Q_{N2} & \dots & Q_{NN} \end{pmatrix}^{QG} \quad (2)$$

where  $Q_{ij}=[hkl]$  or zero. The zero value for  $Q_{ij}$  indicates that atom  $i$  and  $j$  are not connected and  $[hkl]$  indicates that  $i$ -atom is bonded to  $j$ -atom in the adjacent  $hkl$ -cell. We reject those structures with unsaturated atoms and generate a new random trial. After building a saturated quotient graph, we adjust the lattice parameters and atomic positions so as to improve the bond lengths and bond angles in the initially random structure. In the current work the ideal values are bond lengths of 1.55 Å and angles of 109.47°. The quotient graph is not updated during the optimization stage. Namely, we adjust the atomic positions according to only those distances  $d_{ij}^{hkl}$  with non-zero value of  $Q_{ij}$ . When the structure has bond lengths and angles sufficiently close to our ideal values, we recalculate the whole distance matrix of the structure and compare the values of all the non-bonding distances to the bonding distances. We reject structures which have non-bonding distances that are shorter than the bonding distances. The surviving structures are studied further through density functional theory (DFT) based first-principles optimization (see further details in the S1 part in our supplementary file<sup>71</sup>).

### III. RESULTS AND DISCUSSION

As summarized in SACADA, there are at least 522 carbon crystals, including pure  $sp^2$ , pure  $sp^3$  and mixed  $sp^2$ - $sp^3$  carbon crystals, that have been previously proposed. In our present work, the 95  $sp^3$  with tetragonal, hexagonal and cubic symmetry are considered as reference systems. The newly developed RG<sup>2</sup> is applied to systematically search for  $sp^3$  carbon crystals in space groups from No.75 to No.230 (tetragonal, hexagonal and cubic). The results show that most of the previously proposed  $sp^3$  carbon crystals can be generated excepting those large-size carbons structures constructed using specialised methods<sup>41,42,44,69,70</sup>. Otherwise, our approach

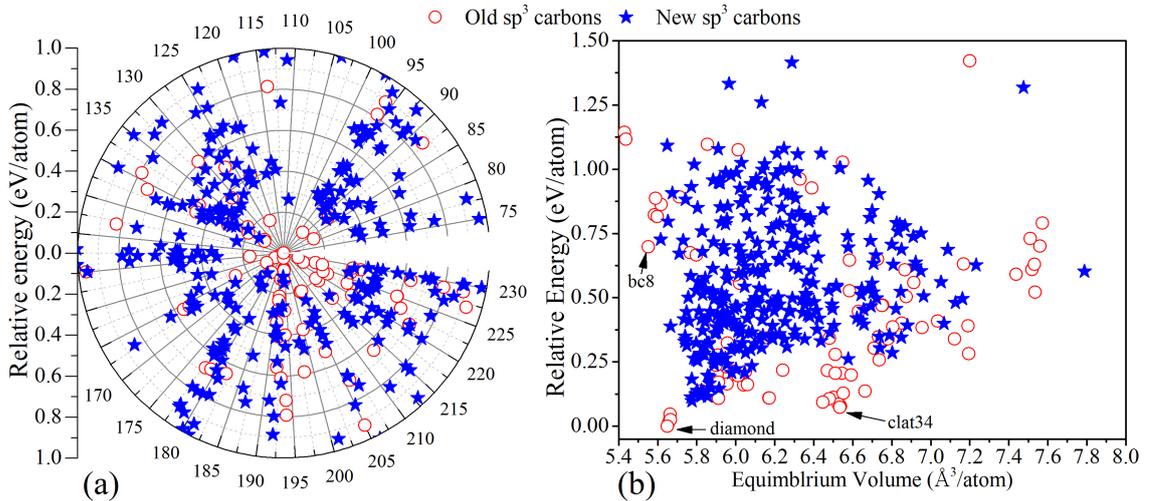


FIG. 2: The total-energies per atom of the  $sp^3$  carbon crystals belonging to space groups from No.75 to No. 230 relative to diamond (a). The total-energies are mapped as radius and the space group numbers are mapped as degree. The red circles and blue solid five-point stars mean the  $sp^3$  carbon crystals discovered in previous work and those newly discovered in our present work, respectively. The scatter plot of the relative average energy (eV/atom) against equilibrium volume ( $\text{\AA}^3/\text{atom}$ ) for  $sp^3$  carbons with tetragonal, hexagonal and cubic symmetry (b).

generates 281 new  $sp^3$  carbon crystals which are different to those in SACADA. Especially, about 129 of these newly discovered  $sp^3$  carbons contain atoms number per primitive cell larger than 40 (35 of them contain more than 100 atoms per primitive cell and 5 of them contain more than 200 atoms).

The results (before DFT-calculations) are obtained in just a few days, running on a single CPU. We then optimize all of these 281  $sp^3$  carbon crystals directly using the VASP code, which takes a few weeks on a cluster (100 CPUs). After that, some of these newly discovered  $sp^3$  carbons, with relatively lower energy or with special structural characteristics, are further investigated for their dynamical stabilities, electronic and mechanical properties. In Fig.2 (a) and (b) the average energies and the scatter plot of the average energies against equilibrium volume of these  $sp^3$  carbon crystals relative to diamond are plotted. The crystalline information files (CIF) of these 281 newly discovered  $sp^3$  carbons are provided as supplementary information and key data is listed in the supplementary file<sup>71</sup>, including their names, space groups, Hermann-Mauguin symbols, atom numbers, total energies, total volumes, average volumes and average energies relative to diamond. We have shown that RG<sup>2</sup> can quickly generate many new  $sp^3$  carbon crystals with energies around 100 to 1000 meV higher than diamond. Many of these structures have very large unit cells. Details of these results will be discussed below according to their crystal systems.

To conveniently manage and discuss such large numbers of structures generated by RG<sup>2</sup>, we require a naming strategy. The carbon structures are named according to key structural information including the space group number (No), the number of inequivalent atoms (Ne), the lattice constants (abc) of the corresponding crystalline cell, the hybridization type (h) and an additional index number (i), as "No-Ne-abc-h-i". The value of "No" can be 1 to 230 and the value of "Ne" can be any non-zero integers. The hybridization type "h" can be, for

example, 3, 4 and 34, which means pure  $sp^2$ , pure  $sp^3$  and mixed  $sp^2$ - $sp^3$  carbon structures, respectively. We project the lattice parameters a, b and c of the conventional cell onto the 26 letters (a-z and A-Z) in the following way:

$$\left( \begin{array}{cccccc} [0.5, 1] & [1, 1.5] & [1.5, 2] & [2, 2.5] & \dots & [25.5, 26] & [26, \infty) \\ \downarrow & \downarrow & \downarrow & \downarrow & \dots & \downarrow & \downarrow \\ a & A & b & B & \dots & z & Z \end{array} \right) \quad (3)$$

Further information is provided in the Supplementary file<sup>71</sup>.

### A. Cubic $sp^3$ carbons

The literature records that various cubic carbons<sup>21,72-77</sup> can be synthesized under high pressure conditions and some of them are still structurally unresolved. Much theoretical effort has been expended on predicting cubic carbon crystals over the past decades<sup>56-58,78-80</sup>. According to SACADA (New update from 2 May 2017), there about 38  $sp^3$  carbons belonging to the cubic system, including diamond, BC8<sup>47,81</sup>, T-carbon<sup>29</sup>, fcc-C34<sup>82</sup> and sc-C46<sup>82</sup>. In our present work, we can generate most of these previously proposed cubic carbons, excepting those containing very large numbers of atoms per cell (Such as KXI, KVI, Clathrate I-100, Clathrate I-280, Clathrate II-100, Clathrate II-280, Clathrate II+Iva and Clathrate I+II)<sup>69,70,83</sup>. As listed in the supplementary file<sup>71</sup>, we find 65 new  $sp^3$  cubic carbon crystals.

The relative energies of our cubic  $sp^3$  carbons are shown in Fig.2 and Fig.S1 (a). Nearly all of the newly discovered  $sp^3$  cubic carbons possess relative energies higher than that of diamond exceeding 300 meV/atom (excepting 205-6-iii-4-002 and 215-10-III-4-001). Our present work identifies some large-size carbon crystals. For example, our 230-7-000-4-001, 224-10-KKK-4-00, 224-10-kkk-4-001, 230-5-nnn-4-001, 230-5-MMM-4-001, 227-9-PPP-4-001, 221-7-

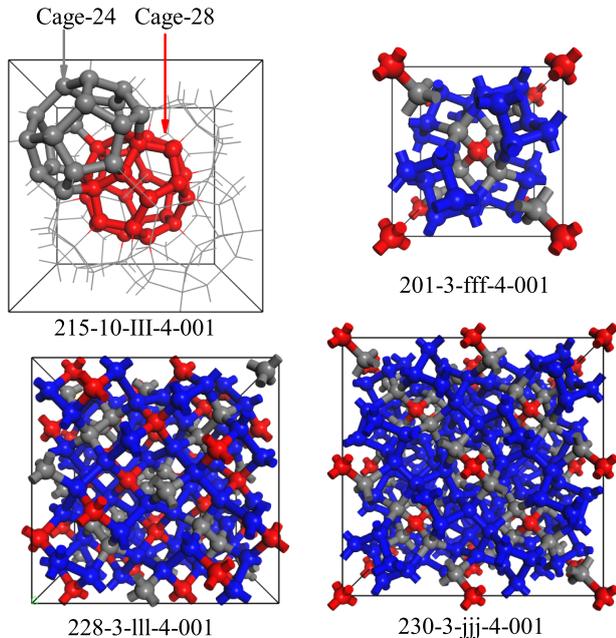


FIG. 3: The perspective crystalline views of the clathrate-like 215-10-III-4-001, the diamond-like 201-3-fff-4-001, 228-3-III-4-001 and 230-3-jjj-4-001 with only six-member rings.

JJJ-4-001, 230-4-LLL-4-002 and 230-4-LLL-4-001 contain 304, 226, 220, 208, 208, 170, 160, 156 and 156 atoms per cell, respectively. Some of the newly discovered  $sp^3$  cubic carbons (those name containing "nnn", "NNN", "ooo" and "OOO") possess lattice constants of nearly 14.7 Å. These results show that the method RG<sup>2</sup> possesses the ability to explore large-size carbon crystals, such as the one experimentally discovered in natural Posigai impact crater<sup>21</sup>. Unfortunately the simulated XRDs (Fig.S3) of these large-size cubic carbons do not explain the super-large cubic carbon discovered in the Posigai impact crater.

Some of our  $sp^3$  cubic carbons are structural intriguing and topologically interesting. For example, the clathrate-like 215-10-III-4-001 (It had been topologically predicted as odg<sup>84</sup> but not been investigated as a carbon phase.) belonging to space group No. 215. The calculated total energy of 215-10-III-4-001 is of about 296 meV/atom higher than that of diamond. It is more stable than the previously predicted T-carbon<sup>29</sup>, BC8<sup>47,81</sup> and BC12<sup>46</sup>. As shown in Fig.3 (a), 215-10-III-4-001 can be constructed from the cage-24 and cage-28, similar to fcc-C34, hex-C40 and sc-C46<sup>82</sup>. Both the dynamical and mechanical stabilities of 215-10-III-4-001 are confirmed through its vibrational spectra and elastic constants as discussed in the supplementary file<sup>71</sup>. As shown in Tab.S1 and Fig.S5, our results also indicate that it is a super-hard semiconductor. In Fig.3 (b)-(f), we show three new  $sp^3$  cubic carbons with only six-member carbon rings similar to diamond, BC8<sup>47,81</sup> and BC12<sup>46</sup>. They are 201-3-fff-4-001, 228-3-III-4-001 and 230-3-jjj-4-001 which possess total energies of about 1264 meV/atom, 931 meV/atom and 414 meV/atom higher than that of diamond, respectively.

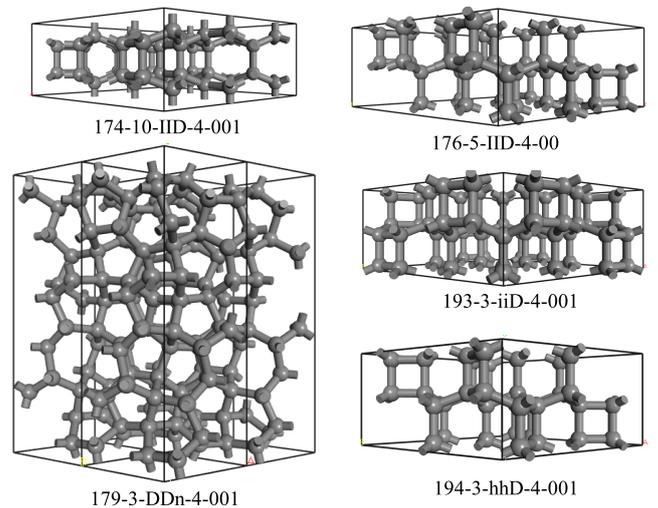


FIG. 4: The perspective crystalline views of 174-10-IIID-4-001, 179-3-DDn-4-001, 176-5-IIID-4-001, 193-3-iiD-4-001 and 194-3-hhD-4-001 can be structurally translated from graphite.

## B. Hexagonal $sp^3$ carbons

A well known hexagonal carbon crystal is hexagonal diamond. In the past decades, many hypothetical hexagonal carbons, such as 4H-diamond, 12R-diamond<sup>85</sup>, CFSs<sup>26</sup> and hex-C40<sup>82</sup>, have been proposed. As collected in SACADA, there are at least 35  $sp^3$  hexagonal carbon allotropes have been previously proposed. Our present work re-discovered most of these hexagonal  $sp^3$  carbons and generates 92 new hexagonal  $sp^3$  carbons. As shown in Fig.S1 (b), we can see that the most stable  $sp^3$  hexagonal carbons are hexagonal diamond and its modifications in space groups No.166, No.191 and No.194. Some of our  $sp^3$  hexagonal carbons possess a relatively low energy in comparison with most of the previously proposed ones. For example, the 174-10-IIID-4-001, 176-5-IIID-4-001, 179-3-DDn-4-001, 193-3-iiD-4-001 and 194-3-hhD-4-001 (see in Fig. 4) are structurally translated to graphite. Their energies are just 166 meV/atom, 125 meV/atom, 196 meV/atom, 114 meV/atom and 145 meV/atom higher than that of diamond, respectively. That is to say, they are energetically comparable to those previously proposed candidates (such as M-carbon<sup>4,55</sup>, W-carbon<sup>49</sup>, Z-carbon<sup>50-52</sup> and S-carbon<sup>53</sup>) for the super-hard carbons cold-compressed from graphite<sup>22</sup>.

As shown in the supplementary file<sup>71</sup> in Tab. S1 and Fig. S2, 174-10-IIID-4-001, 176-5-IIID-4-001, 179-3-DDn-4-001, 193-3-iiD-4-001 and 194-3-hhD-4-001 are dynamically stable. Their electronic and mechanical properties are shown in Tab. S1 and Fig. S5. We can see that they are carbon structures with semiconducting and super-hard properties. Our present search generates 14 hexagonal  $sp^3$  carbon crystals which can be structurally obtained from graphite. The simulated XRDs of these 14 hexagonal  $sp^3$  carbons are plotted in Fig.5. We can see that most of them can be considered as good candidates for explaining the super-hard carbons synthesized

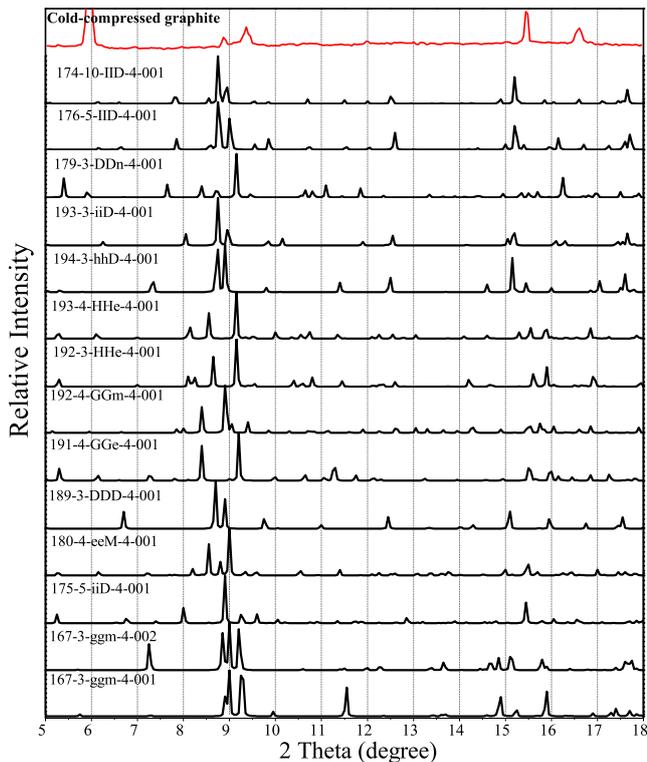


FIG. 5: The simulated XRDs of our newly discovered  $sp^3$  hexagonal carbons which can be considered as good candidates for explaining the super-hard carbon crystals synthesized by cold-compressing graphite.

in cold-compressing graphite<sup>22</sup>.

We find some intriguing hexagonal carbons with chiral framework structures (CFS) similar to the previously proposed CFSs<sup>26</sup>. These CFSs mainly distribute in chiral space groups 152 (154), 169 (170), 178 (179) and 180 (181). As shown in Fig.S4 are two typical pairs of CFSs, namely the previously proposed P6122 (P6522) and the newly discovered 169-4-FFe-4-001 (170-4-FFe-4-001). Our present work also discover some large-size  $sp^3$  hexagonal carbons with lattice constants very similar to those of the experimentally discovered Chaoite ( $a=8.948 \text{ \AA}$ ,  $c=14.078 \text{ \AA}$ )<sup>18–20</sup>. Examples are 193-6-IIIm-4-001 and 172-26-iiM-4-001 with lattice constants of ( $a=9.47 \text{ \AA}$ ,  $c=12.84 \text{ \AA}$ ) and ( $a=8.95 \text{ \AA}$ ,  $c=13.27 \text{ \AA}$ ), respectively. The relative energies of 193-6-IIIm-4-001 and 172-26-iiM-4-001 are 637 meV/atom and 362 meV/atom, respectively. We find that 172-26-iiM-4-001 can also be structurally constructed based on the precursor (graphite) used to synthesize Chaoite<sup>19,20</sup>. The simulated XRDs of these two new large-size  $sp^3$  hexagonal carbons cannot, however, explain the experimental XRD of Chaoite.

### C. Tetragonal $sp^3$ carbons

A tetragonal carbon structure had been reported experimentally. It has been named as P-diamond<sup>86</sup> and has lattice constants of  $a=3.68 \text{ \AA}$  and  $c=3.47 \text{ \AA}$ . The detailed structure of

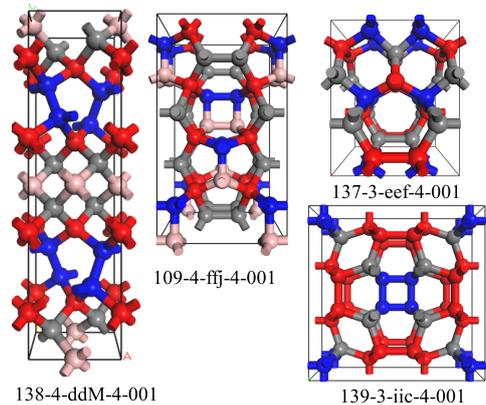


FIG. 6: The perspective crystalline views of the most stable four new tetragonal  $sp^3$  carbon allotropes discovered by RG<sup>2</sup>. They are 137-3-eef-4-001, 109-4-ffj-4-001, 139-3-iic-4-001 and 138-4-ddM-4-001.

such a simple tetragonal carbon phase is still an open question. On the other hand, there is experimental evidence that element silicon can form tetragonal crystals<sup>87,88</sup>. Most of these experimentally reported silicon phases are still structural unsolved. The prediction of a range of tetragonal carbon structures can provide candidates for explaining experimentally synthesized tetragonal silicon crystals and potential tetragonal carbon crystals which may be synthesized in future.

There are at least 28 theoretical  $sp^3$  tetragonal carbons which have been previously proposed, including the well-known bct-C4<sup>10,37–39,59</sup>, T12<sup>89</sup> and Clathrates III<sup>90</sup>, as well as the recently proposed P41212<sup>91</sup>. Among these  $sp^3$  tetragonal carbons, T12 (108 meV/atom), Clathrates III (113 meV/atom), P41212 (132 meV/atom) and I-4-carbon (161 meV/atom) are the most stable. In our present search, we find most of the previously proposed tetragonal  $sp^3$  carbons. Furthermore, we uncover 124 new  $sp^3$  tetragonal carbon structures. Some of them possess favourable energetic stability. For example, 138-4-ddM-4-001 (139 meV/atom), 137-3-eef-4-001 (207 meV/atom), 109-4-ffj-4-001 (261 meV/atom) and 139-3-iic-4-001 (264 meV/atom) as shown in Fig. 6. We investigate their vibrational, mechanical and electronic properties using VASP. As shown in Fig. S2 and Tab. S1, we can see that they are dynamically and mechanically stable carbon structures. Tab. S1 and Fig. S4, show that they are super-hard carbon semiconductors.

We also find some intriguing structures in the chiral tetragonal space groups 92 and 96. Shown in Fig.S6 are the newly discovered 92-6-HHC-4-001 (96-6-HHC-4-001) and 92-4-ggD-4-001 (96-4-ggD-4-001). Their relative energies are 318 meV/atom and 391 meV/atom, which are slightly less stable than the recently proposed P41212 (P43212) structures which have three inequivalent atoms. Other low energy tetragonal chiral  $sp^3$  carbons, such as 92-4-eei-4-001 (303 meV/atom), 96-3-eeE-4-001 (685 meV/atom) and 96-2-CCi-4-001 (338 meV/atom), can be found in the supplementary information provided in our supplementary file<sup>71</sup>.

#### IV. CONCLUSION

In this study, we present a simple approach to generate crystal structures with defined geometrical features. Based on this approach, we systematically search for crystal structures of  $sp^3$  carbon in space groups ranging from No.75 to No.230. We generate several hundred  $sp^3$  carbon structures, including most of the previously proposed ones and 281 new ones. The first principles results show that most of these 281  $sp^3$  carbons are dynamically and mechanically stable. These  $sp^3$  carbon crystals include some with very large cubic or hexagonal unit cells, some chiral framework structures and some low energy super-hard carbon structures. Some of them can be considered as candidates for explaining previous experimental results. These results suggest that  $RG^2$  is an effective approach for the generation crystal structures with defined geometrical features. This approach to the generation of structures can be straightforwardly extended to some other crystalline structures with defined geometrical features, such as the three-connected crystal structures for  $sp^2$  carbon and high pressure nitrogen ( $sp^3$ ),

four-connected crystal structures for  $sp^3$  silicon and germanium, as well as the crystal structures for mixed  $sp^2$ - $sp^3$  carbon,  $C_3N_4$  and  $Si_3N_4$ , by neglecting their chemical natures.

#### V. ACKNOWLEDGEMENT

This work is supported by the National Natural Science Foundation of China (Grant Nos. 11704319, 11647063, A040204 and 11204261), the National Basic Research Program of China (2012CB921303 and 2015CB921103), the Young Scientists Fund of the National Natural Science Foundation of China (Grant No. 11204260), the Natural Science Foundation of Hunan Province, China (Grant No. 2016JJ3118, the Scientific Research Found of HuNan Provincial Education department (No. 14C1095) and the Program for Changjiang Scholars and Innovative Research Team in University (IRT13093). CJP is supported by the Royal Society through a Royal Society Wolfson Research Merit award and the EPSRC through grants EP/P022596/1 and EP/J010863/2.

\* Electronic address: hechaoyu@xtu.edu.cn

† Electronic address: cjp20@cam.ac.uk

‡ Electronic address: tang\_chao@xtu.edu.cn

- <sup>1</sup> J. Maddox, *Nature* **335**, 201 (1988).
- <sup>2</sup> C. J. Pickard and R. Needs, *Physical Review Letters* **97**, 045504 (2006).
- <sup>3</sup> C. J. Pickard and R. Needs, *Journal of Physics: Condensed Matter* **23**, 053201 (2011).
- <sup>4</sup> A. R. Oganov and C. W. Glass, *The Journal of chemical physics* **124**, 244704 (2006).
- <sup>5</sup> C. W. Glass, A. R. Oganov, and N. Hansen, *Computer Physics Communications* **175**, 713 (2006).
- <sup>6</sup> A. O. Lyakhov, A. R. Oganov, H. T. Stokes, and Q. Zhu, *Computer Physics Communications* **184**, 1172 (2013).
- <sup>7</sup> Y. Wang, J. Lv, L. Zhu, and Y. Ma, *Computer Physics Communications* **183**, 2063 (2012).
- <sup>8</sup> Y. Wang, J. Lv, L. Zhu, and Y. Ma, *Physical Review B* **82**, 094116 (2010).  
J.-G. Eon, *Structural Chemistry* **23**, 987 (2012), ISSN 1572-9001, URL <https://doi.org/10.1007/s11224-012-0006-2>.
- <sup>9</sup> R. T. Strong, C. J. Pickard, V. Milman, G. Thimm, and B. Winkler, *Physical Review B* **70**, 045101 (2004).
- <sup>10</sup> R. Martoňák, A. Laio, and M. Parrinello, *Physical review letters* **90**, 075503 (2003).
- <sup>11</sup> R. Martoňák, A. Laio, M. Bernasconi, C. Ceriani, P. Raiteri, F. Zipoli, and M. Parrinello, *Zeitschrift für Kristallographie-Crystalline Materials* **220**, 489 (2005).
- <sup>12</sup> J. Pannetier, J. Bassas-Alsina, J. Rodriguez-Carvajal, and V. Caingnaert, *Nature* **346**, 343 (1990).
- <sup>13</sup> J. C. Schön and M. Jansen, *Angewandte Chemie International Edition* **35**, 1286 (1996).
- <sup>14</sup> D. J. Wales and J. P. Doye, *The Journal of Physical Chemistry A* **101**, 5111 (1997).
- <sup>15</sup> S. Goedecker, *The Journal of chemical physics* **120**, 9911 (2004).
- <sup>16</sup> M. Amsler and S. Goedecker, *The Journal of chemical physics* **133**, 224104 (2010).
- <sup>17</sup> A. El Goresy and G. Donnay, *Science* **161**, 363 (1968).
- <sup>18</sup> A. G. Whittaker and P. Kintner, *Science* **165**, 589 (1969).
- <sup>19</sup> A. G. Whittaker and G. M. Wolten, *Science* **178**, 54 (1972).
- <sup>20</sup> A. El Goresy, L. S. Dubrovinsky, P. Gillet, S. Mostefaoui, G. Graup, M. Drakopoulos, A. S. Simionovici, V. Swamy, and V. L. Masaitis, *Comptes Rendus Geoscience* **335**, 889 (2003).
- <sup>21</sup> W. L. Mao, H.-k. Mao, P. J. Eng, T. P. Trainor, M. Newville, C.-c. Kao, D. L. Heinz, J. Shu, Y. Meng, and R. J. Hemley, *Science* **302**, 425 (2003).
- <sup>22</sup> Z. Wang, Y. Zhao, K. Tait, X. Liao, D. Schiferl, C. Zha, R. T. Downs, J. Qian, Y. Zhu, and T. Shen, *Proceedings of the National Academy of Sciences of the United States of America* **101**, 13699 (2004).
- <sup>23</sup> P. A. Schultz, K. Leung, and E. Stechel, *Physical Review B* **59**, 733 (1999).
- <sup>24</sup> F. J. Ribeiro, P. Tangney, S. G. Louie, and M. L. Cohen, *Physical Review B* **74**, 172101 (2006).
- <sup>25</sup> C. J. Pickard and R. Needs, *Physical Review B* **81**, 014106 (2010).
- <sup>26</sup> S. Berber, E. Osawa, and D. Tománek, *Physical Review B* **70**, 085417 (2004).
- <sup>27</sup> R. Hoffmann, T. Hughbanks, M. Kertesz, and P. H. Bird, *Journal of the American Chemical Society* **105**, 4831 (1983).
- <sup>28</sup> X.-L. Sheng, Q.-B. Yan, F. Ye, Q.-R. Zheng, and G. Su, *Physical review letters* **106**, 155703 (2011).
- <sup>29</sup> B. Winkler, C. J. Pickard, V. Milman, and G. Thimm, *Chemical physics letters* **337**, 36 (2001).
- <sup>30</sup> M. A. Tamor and K. Hass, *Journal of Materials Research* **5**, 2273 (1990).
- <sup>31</sup> M. Itoh, M. Kotani, H. Naito, T. Sunada, Y. Kawazoe, and T. Adschiri, *Physical review letters* **102**, 055703 (2009).
- <sup>32</sup> Y. Yao, J. Tse, J. Sun, D. Klug, R. Martoňák, and T. Iitaka, *Physical review letters* **102**, 229601 (2009).
- <sup>33</sup> M. Côté, J. C. Grossman, M. L. Cohen, and S. G. Louie, *Physical Review B* **58**, 664 (1998).
- <sup>34</sup> M. óeffe, G. B. Adams, and O. F. Sankey, *Physical review letters* **68**, 2325 (1992).

- <sup>36</sup> Q. Zhu, Q. Zeng, and A. R. Oganov, *Physical Review B* **85**, 201407 (2012).
- <sup>37</sup> K. Umemoto, R. M. Wentzcovitch, S. Saito, and T. Miyake, *Physical review letters* **104**, 125504 (2010).
- <sup>38</sup> R. Baughman, A. Liu, C. Cui, and P. Schields, *Synthetic metals* **86**, 2371 (1997).
- <sup>39</sup> H. Domingos, *Journal of Physics: Condensed Matter* **16**, 9083 (2004).
- <sup>40</sup> F. Tian, X. Dong, Z. Zhao, J. He, and H.-T. Wang, *Journal of Physics: Condensed Matter* **24**, 165504 (2012).
- <sup>41</sup> R. Zhou and X. C. Zeng, *Journal of the American Chemical Society* **134**, 7530 (2012).
- <sup>42</sup> H. Niu, X.-Q. Chen, S. Wang, D. Li, W. L. Mao, and Y. Li, *Physical review letters* **108**, 135501 (2012).
- <sup>43</sup> S. Botti, M. Amsler, J. A. Flores-Livas, P. Ceria, S. Goedecker, and M. A. Marques, *Physical Review B* **88**, 014102 (2013).
- <sup>44</sup> C. He, L. Sun, C. Zhang, X. Peng, K. Zhang, and J. Zhong, *Physical Chemistry Chemical Physics* **14**, 8410 (2012).
- <sup>45</sup> X. Zhang, Y. Wang, J. Lv, C. Zhu, Q. Li, M. Zhang, Q. Li, and Y. Ma, *The Journal of chemical physics* **138**, 114101 (2013).
- <sup>46</sup> Z.-Z. Li, J.-T. Wang, L.-F. Xu, and C. Chen, *Physical Review B* **94**, 174102 (2016).
- <sup>47</sup> R. L. Johnston and R. Hoffmann, *Journal of the American Chemical Society* **111**, 810 (1989).
- <sup>48</sup> Q. Zhu, A. R. Oganov, M. A. Salvadó, P. Pertierra, and A. O. Lyakhov, *Physical Review B* **83**, 193410 (2011).
- <sup>49</sup> J.-T. Wang, C. Chen, and Y. Kawazoe, *Physical review letters* **106**, 075501 (2011).
- <sup>50</sup> D. Sellì, I. A. Baburin, R. Martoňák, and S. Leoni, *Physical Review B* **84**, 161411 (2011).
- <sup>51</sup> M. Amsler, J. A. Flores-Livas, L. Lehtovaara, F. Balima, S. A. Ghasemi, D. Machon, S. Pailhes, A. Willand, D. Caliste, S. Botti, et al., *Physical review letters* **108**, 065501 (2012).
- <sup>52</sup> Z. Zhao, B. Xu, X.-F. Zhou, L.-M. Wang, B. Wen, J. He, Z. Liu, H.-T. Wang, and Y. Tian, *Physical review letters* **107**, 215502 (2011).
- <sup>53</sup> C. He, L. Sun, C. Zhang, X. Peng, K. Zhang, and J. Zhong, *Solid State Communications* **152**, 1560 (2012).
- <sup>54</sup> D. Li, K. Bao, F. Tian, Z. Zeng, Z. He, B. Liu, and T. Cui, *Physical Chemistry Chemical Physics* **14**, 4347 (2012).
- <sup>55</sup> Q. Li, Y. Ma, A. R. Oganov, H. Wang, H. Wang, Y. Xu, T. Cui, H.-K. Mao, and G. Zou, *Physical review letters* **102**, 175506 (2009).
- <sup>56</sup> J.-T. Wang, C. Chen, and Y. Kawazoe, *Physical Review B* **85**, 214104 (2012).
- <sup>57</sup> Z.-Z. Li, C.-S. Lian, J. Xu, L.-F. Xu, J.-T. Wang, and C. Chen, *Physical Review B* **91**, 214106 (2015).
- <sup>58</sup> C. He, C. Zhang, H. Xiao, L. Meng, and J. Zhong, *Carbon* **112**, 91 (2017).
- <sup>59</sup> Y. Omata, Y. Yamagami, K. Tadano, T. Miyake, and S. Saito, *Physica E: Low-dimensional Systems and Nanostructures* **29**, 454 (2005).
- <sup>60</sup> M. Martinez-Canales, C. J. Pickard, and R. J. Needs, *Physical Review Letters* **108**, 045704 (2012).
- <sup>61</sup> Y.-Y. Zhang, S. Chen, H. Xiang, and X.-G. Gong, *Carbon* **109**, 246 (2016).
- <sup>62</sup> Y. J. Oh, S. Kim, I.-H. Lee, J. Lee, and K. Chang, *Physical Review B* **93**, 085201 (2016).
- <sup>63</sup> V. L. Deringer, G. Csányi, and D. M. Proserpio, *ChemPhysChem* **18**, 873 (2017).
- <sup>64</sup> I. A. Baburin, D. M. Proserpio, V. A. Saleev, and A. V. Shipilova, *Physical Chemistry Chemical Physics* **17**, 1332 (2015).
- <sup>65</sup> H.-J. Sung, S. Kim, I.-H. Lee, and K. J. Chang, *NPG Asia Materials* **9**, e361 (2017).
- <sup>66</sup> Y. Cheng, X. Feng, X. Cao, B. Wen, Q. Wang, Y. Kawazoe, and P. Jena, *Small* **13**, 1602894 (2017).
- <sup>67</sup> M. Takagi, T. Taketsugu, H. Kino, Y. Tateyama, K. Terakura, and S. Maeda, *Physical Review B* **95**, 184110 (2017).
- <sup>68</sup> A. F. Wells et al., *Three dimensional nets and polyhedra* (Wiley, 1977).
- <sup>69</sup> A. J. Karttunen, T. F. Fassler, M. Linnolahti, and T. A. Pakkanen, *Inorganic chemistry* **50**, 1733 (2011).
- <sup>70</sup> A. J. Karttunen and T. F. Fässler, *ChemPhysChem* **14**, 1807 (2013).
- <sup>71</sup> See supplemental material at <http://link.aps.org/supplemental/xxxx> for (s1) the dft calculation details, (s2) the explaining text for naming strategy, (s3) the supplementary figures, (s4) the supplementary tables, (s5) the information list of the 281 newly discovered  $sp^3$  carbon crystals, (s6) the optimized crystalline information files (cif) of these 281 newly discovered  $sp^3$  carbons.
- <sup>72</sup> R. Aust and H. Drickamer, *Science* **140**, 817 (1963).
- <sup>73</sup> P. Liu, Y. Cao, C. Wang, X. Chen, and G. Yang, *Nano letters* **8**, 2570 (2008).
- <sup>74</sup> V. Strel'itskii, V. Padalka, and S. Vakula, *Sov. Phys. Tech. Phys* **23**, 222 (1978).
- <sup>75</sup> H. Hirai and K. KONDO, *Science* **253**, 772 (1991).
- <sup>76</sup> H. Vora and T. Moravec, *Journal of Applied Physics* **52**, 6151 (1981).
- <sup>77</sup> K. Yamada, *Carbon* **41**, 1309 (2003).
- <sup>78</sup> V. Pokropivny and A. Pokropivny, *Physics of the Solid State* **46**, 392 (2004).
- <sup>79</sup> A. Pokropivny and S. Volz, *physica status solidi (b)* **249**, 1704 (2012).
- <sup>80</sup> M. Hu, F. Tian, Z. Zhao, Q. Huang, B. Xu, L.-M. Wang, H.-T. Wang, Y. Tian, and J. He, *The Journal of Physical Chemistry C* **116**, 24233 (2012).
- <sup>81</sup> R. Biswas, R. M. Martin, R. Needs, and O. Nielsen, *Physical Review B* **35**, 9559 (1987).
- <sup>82</sup> G. Benedek, E. Galvani, S. Sanguinetti, and S. Serra, *Chemical physics letters* **244**, 339 (1995).
- <sup>83</sup> H.-Y. Zhao, J. Wang, Q.-M. Ma, and Y. Liu, *The Journal of chemical physics* **138**, 164703 (2013).
- <sup>84</sup> O. Delgado-Friedrichs and M. O'Keeffe, *Acta Crystallographica Section A* **66**, 637 (2010).
- <sup>85</sup> J.-T. Wang, C. Chen, and Y. Kawazoe, *Physical Review B* **84**, 012102 (2011).
- <sup>86</sup> J. Kleiman, R. Heimann, D. Hawken, and N. Salansky, *Journal of applied physics* **56**, 1440 (1984).
- <sup>87</sup> L. Rapp, B. Haberl, C. J. Pickard, J. E. Bradby, E. G. Gamaly, J. S. Williams, and A. V. Rode, *Nature communications* **6**, 7555 (2015).
- <sup>88</sup> W. Chen, G. Hamon, R. Léal, J.-L. Maurice, L. Largeau, and P. Roca i Cabarrocas, *Crystal Growth & Design* **17**, 4265 (2017).
- <sup>89</sup> Z. Zhao, F. Tian, X. Dong, Q. Li, Q. Wang, H. Wang, X. Zhong, B. Xu, D. Yu, J. He, et al., *Journal of the American Chemical Society* **134**, 12362 (2012).
- <sup>90</sup> M. O'Keeffe, G. Adams, and O. Sankey, *Philosophical magazine letters* **78**, 21 (1998).
- <sup>91</sup> A. Mujica, C. J. Pickard, and R. J. Needs, *Physical Review B* **91**, 214104 (2015).