

# Thermodynamics of diamond formation from hydrocarbon mixtures in planets

Received: 10 August 2022

Accepted: 15 February 2023

Published online: 27 February 2023

 Check for updatesBingqing Cheng<sup>1</sup>✉, Sebastien Hamel<sup>2</sup> & Mandy Bethkenhagen<sup>1,3</sup>

Hydrocarbon mixtures are extremely abundant in the Universe, and diamond formation from them can play a crucial role in shaping the interior structure and evolution of planets. With first-principles accuracy, we first estimate the melting line of diamond, and then reveal the nature of chemical bonding in hydrocarbons at extreme conditions. We finally establish the pressure-temperature phase boundary where it is thermodynamically possible for diamond to form from hydrocarbon mixtures with different atomic fractions of carbon. Notably, here we show a depletion zone at pressures above 200 GPa and temperatures below 3000 K–3500 K where diamond formation is thermodynamically favorable regardless of the carbon atomic fraction, due to a phase separation mechanism. The cooler condition of the interior of Neptune compared to Uranus means that the former is much more likely to contain the depletion zone. Our findings can help explain the dichotomy of the two ice giants manifested by the low luminosity of Uranus, and lead to a better understanding of (exo-)planetary formation and evolution.

Carbon and hydrogen are the fourth and the most abundant elements in the Universe<sup>1</sup>, and their mixture is the simplest basis to form organic compounds. In our Solar System, the Cassini mission revealed lakes and seas of liquid hydrocarbons on the surface of Titan<sup>2</sup>, and the New Horizon spacecraft detected methane frost on the mountains of Pluto<sup>3</sup>. In Neptune and Uranus, methane is a major constituent with a measured carbon concentration from around 2% in the atmosphere<sup>4</sup> and a concentration up to (assumed) 8% in the interior<sup>5</sup>. The methane in the atmosphere absorbs red light and reflects blue light, giving the ice giants their blue hues<sup>6</sup>. Moreover, numerous recently discovered extrasolar planets, some orbiting carbon-rich stars, have spurred a renewed interest in the high-pressure and high-temperature behaviors of hydrocarbons<sup>7</sup>.

Diamond formation from C/H mixtures is particularly relevant; the “diamonds in the sky” hypothesis<sup>8</sup> suggests that diamonds can form in the mantles of Uranus and Neptune. The diamond formation and the accompanying heat release may explain the long-standing puzzle that Neptune (but not Uranus) radiates much more energy than that it receives from the Sun<sup>9</sup>. Diamond is dense and will gravitate into the core of the ice giants. For white dwarfs, Tremblay et al.<sup>10</sup>

interpreted the crystallization of the carbon-rich cores to influence the cooling rate.

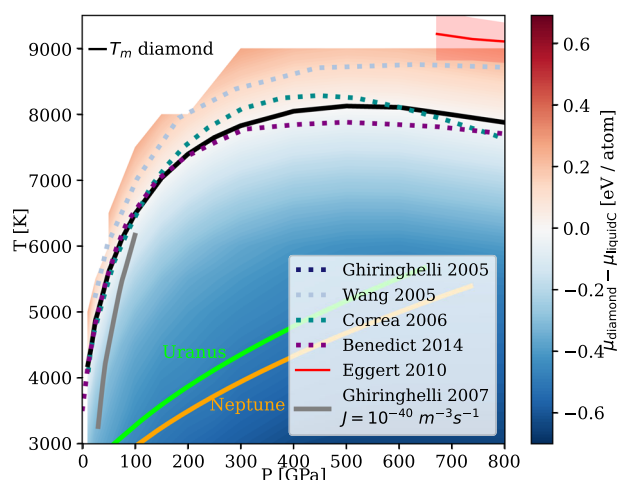
Many experimental studies have probed the diamond formation from C/H mixtures, but the experiments are extraordinarily challenging to perform and interpret because of the extreme thermodynamic conditions, kinetics, chemical inhomogeneities, possible surface effects from the sample containers, and the need to prove diamond formation inside a diamond anvil cell (DAC). Three DAC studies on methane disagree on the temperature range: Benedetti et al. reported diamond formation between 10 to 50 GPa and temperatures of about 2000 K to 3000 K<sup>11</sup>; between 10 to 80 GPa, Hirai et al. reported diamond formation above 3000 K<sup>12</sup>; while Lobanov et al. reported the observation of elementary carbon at about 1200 K, and a mixture of solid carbon, hydrogen, and heavier hydrocarbons at above 1500 K<sup>13</sup>. In methane hydrates, Kadobayashi et al. reported diamond formation in a DAC between 13 and 45 GPa above 1600 K but not at lower temperatures<sup>14</sup>. Laser shock-compression experiments found diamond formation in epoxy (C,H,Cl,N,O)<sup>15</sup> and polystyrene (C<sub>8</sub>H<sub>8</sub>)<sup>16</sup>, but none in polyethylene (C<sub>2</sub>H<sub>4</sub>)<sup>17</sup>.

<sup>1</sup>The Institute of Science and Technology Austria, Am Campus 1, 3400 Klosterneuburg, Austria. <sup>2</sup>Lawrence Livermore National Laboratory, Livermore, CA 94550, USA. <sup>3</sup>École Normale Supérieure de Lyon, Université Lyon 1, Laboratoire de Géologie de Lyon, CNRS UMR 5276, 69364 Lyon, Cedex 07, France.

✉ e-mail: [bingqing.cheng@ist.ac.at](mailto:bingqing.cheng@ist.ac.at)

Moreover, there is a mismatch between the experimental results and theoretical predictions particularly regarding the pressure range of diamond formation. Density functional theory (DFT) combined with crystal structure searches at the static lattice level predicted that diamond and hydrogen are stable at pressures above about 300 GPa<sup>18,19</sup>, while hydrocarbon crystals are stable at lower pressures<sup>18–22</sup>. Based on DFT molecular dynamics (MD) simulations of methane, Ancilotto et al. concluded that methane dissociates into a mixture of hydrocarbons below 100 GPa and is more prone to form diamond at above 300 GPa<sup>23</sup>, Sherman et al. classified the system into stable methane molecules (<3000 K), a polymeric state consisting of long hydrocarbon chains (4000–5000 K, 40–200 GPa), and a plasma state (>6000 K)<sup>24</sup>. However, these simulations are constrained to small system sizes and short time scales, so that it is impossible to distinguish between the formation of long hydrocarbon chains and the early stage of diamond nucleation. Using a semiempirical carbon model, Ghiringhelli et al.<sup>25</sup> determined that the diamond nucleation rate in pure liquid carbon is rapid at 85 GPa, 5000 K but negligibly small at 30 GPa, 3750 K, and then extrapolated the nucleation rate to mixtures employing an ideal solution model.

In this work, we go beyond the standard first-principles methods, and study the thermodynamics of diamond formation in C/H mixtures, by constructing and utilizing machine learning potentials (MLPs) trained on DFT data. To the best of our knowledge, this is the first MLP fitted for high-pressure mixtures, and the only one available for C/H mixtures with arbitrary compositions and applicable from low  $P$ - $T$  conditions to about 8000 K and 800 GPa. We first quantitatively estimate the coexistence line between diamond and pure liquid carbon at planetary conditions. We then reveal the nature of the chemical bonds in C/H mixtures at high-pressure high-temperature conditions. Finally, we determine the thermodynamic driving force of diamond formation in C/H mixtures, taking into account both the ideal and the non-ideal effects of mixing. We thereby establish the phase boundary where diamond can possibly form from C/H mixtures at different atomic fractions and  $P$ - $T$  conditions.



**Fig. 1 | Thermodynamics for diamond formation in pure liquid carbon.** The color scale shows the chemical potential of diamond,  $\Delta\mu_D$ , referenced to pure liquid carbon. The stability region of graphite at  $P \lesssim 10$  GPa is not shown. The melting curve  $T_m$  is compared to previous calculations using thermodynamic integration (TI) employing a semi-empirical potential by Ghiringhelli et al.<sup>31</sup>, TI using DFT by Wang et al.<sup>28</sup>, coexistence DFT simulations by Correa et al.<sup>29</sup>, an analytic free-energy model fitted to DFT data by Benedict et al.<sup>30</sup>, and a shock-compression experiment by Eggert et al.<sup>27</sup> with uncertainties indicated using the shaded area. The gray line shows the inferred threshold conditions with a diamond nucleation rate  $J$  of  $10^{-40} \text{ m}^{-3} \text{ s}^{-1}$  from pure liquid carbon by Ghiringhelli et al.<sup>25</sup>. The  $P$ - $T$  curves of planetary interior conditions for Uranus (green line) and Neptune (orange line) are from Ref. <sup>40</sup>.

## Results

### Diamond formation in pure liquid carbon

Although planets or stars typically contain a low percentage of carbon<sup>4</sup>, it is useful to start with a hypothetical environment of pure carbon. This is to establish the melting line of diamond and to facilitate the subsequent analysis based on C/H mixtures. Moreover, the high-pressure carbon system has experimental relevance in diamond synthesis and Inertial Confinement Fusion applications<sup>26</sup>.

Figure 1 shows the chemical potential difference  $\Delta\mu_D \equiv \mu_{\text{diamond}} - \mu_{\text{liquidC}}$  between the diamond and the pure liquid carbon phases calculated using our MLP at a wide range of pressures and temperatures. Our calculated melting line  $T_m$  of diamond in pure liquid carbon (solid black curve) is compared to other theoretical work and experimental shock-compression data (Fig. 1a). Our  $T_m$  is re-entrant at above 500 GPa, because liquid carbon is denser than diamond at higher pressures. This shape has been observed for the experimental melting line<sup>27</sup>. It was previously predicted using DFT simulations on smaller systems<sup>28–30</sup>, but not captured in the free energy calculations performed using a semi-empirical LCBO carbon model<sup>31</sup>.

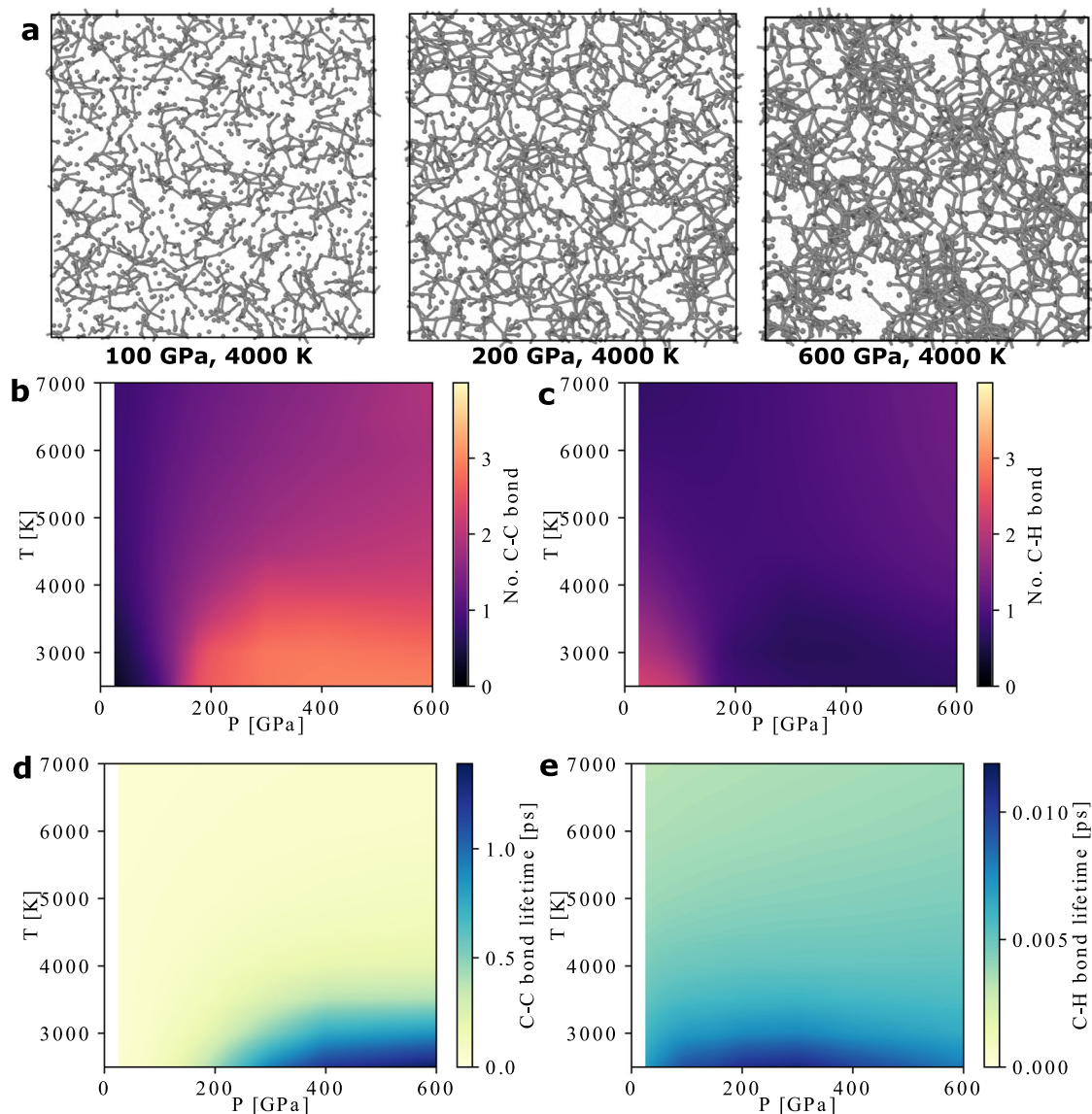
Although diamond solidification is thermodynamically favorable below the melting line, undercooled liquids can remain metastable for a long time as solidification is initiated by a kinetically activated nucleation process<sup>32</sup>. The only previous study that has quantified the diamond nucleation rate is by Ghiringhelli et al.<sup>25</sup> using the LCBO carbon model: the threshold  $J = 10^{-40} \text{ m}^{-3} \text{ s}^{-1}$  is indicated by the gray line in Fig. 1, and above this line the diamond formation rate is negligible even at the celestial scale. Overall, we find that the pure carbon system is deeply undercooled at the  $P$ - $T$  conditions in the two icy planets (green and orange lines in Fig. 1).

### The nature of C-H bonds

Going beyond the pure carbon case, we investigate the nature of the chemical bonds in C/H mixtures at conditions relevant for planetary interiors. The high-pressure behavior of hydrocarbons is also crucial in many shock-compression experiments for the development of fusion energy platforms and Inertial Confinement Fusion capsules<sup>33</sup>. The properties of the covalent C-C and C-H bonds are well-known at ambient conditions, but it is unclear how extreme conditions affect these bonds. DFT studies coupled with harmonic approximations have predicted a variety of hydrocarbon crystals to be stable at  $P \leq 300$  GPa<sup>18–22</sup>, but these studies are restricted to low temperatures as the melting lines of hydrogen and methane are below 1000 K and 2000 K<sup>12,34</sup>, respectively, while harmonic approximations break down completely for these liquids.

We performed MD simulations using our dissociable MLP for C/H mixtures over a wide range of thermodynamic conditions. We focus on the  $\text{CH}_4$  composition to directly compare to previous studies. Other compositions can be analyzed in the same way and yield qualitatively similar behaviors. At  $T < 2500$  K, the MD is not ergodic within the simulation time of 100 ps, and therefore analysis is performed only at temperatures above this threshold. Figure 2a shows the snapshots of carbon bonds from the MD simulations of the  $\text{CH}_4$  system. At 4000 K and  $P = 100$  GPa, 200 GPa, and 600 GPa, the system is primarily composed of various types of hydrocarbon chains. The formation of longer chains at higher pressures is consistent with the observations in previous DFT MD studies<sup>23,24</sup>, although the DFT simulations have severe finite size effects because polymer chains consisting of just a few carbon atoms can connect with their periodic images and become infinitely long. At high pressures, the chains assemble carbon networks, and the system shows more obvious signs of spatial inhomogeneity of carbon atoms.

In our chemical bond analysis, a C-C bond is identified whenever the distance between a pair of carbon atoms is within 1.6 Å, and a C-H bond is defined using a cutoff of 1.14 Å. The cutoffs are larger than the typical bond lengths to eliminate the misidentification of broken



**Fig. 2 | Bonding behavior in the CH<sub>4</sub> mixture at high-pressure high-temperature conditions.** **a** Snapshots from MD simulations using the MLP. Carbon atoms are shown as small gray spheres, while hydrogen atoms are not drawn for clarity.

Bonds are drawn for C-C pairs with distances within 1.6 Å. **b,c** Average number of C-C bonds (**b**) and C-H bonds (**c**) per carbon atom from MD simulations of CH<sub>4</sub> composition. **d,e** Average lifetimes of the C-C bonds (**d**) and C-H bonds (**e**).

bonds due to thermal fluctuations. The average number of C-C and C-H bonds at different conditions are shown in Fig. 2b, c. The number of bonds varies smoothly as a function of  $P$  and  $T$ .

The average number of the C-C bonds decreases with temperature. Moreover, as illustrated in the Supplementary Information, at a certain condition the carbon atoms in the system have varying number of C-C bonds, rather than all having the same number of bonds. These suggest that at  $T \geq 2500$  K the system is not made of hydrocarbon crystals that were predicted to be stable at low temperatures in previous DFT studies<sup>18–22</sup>. The average number of C-H bonds for each carbon atom is close to one at all conditions considered here even though the overall composition is CH<sub>4</sub>, indicating that most hydrogen atoms are not bonded to any carbons.

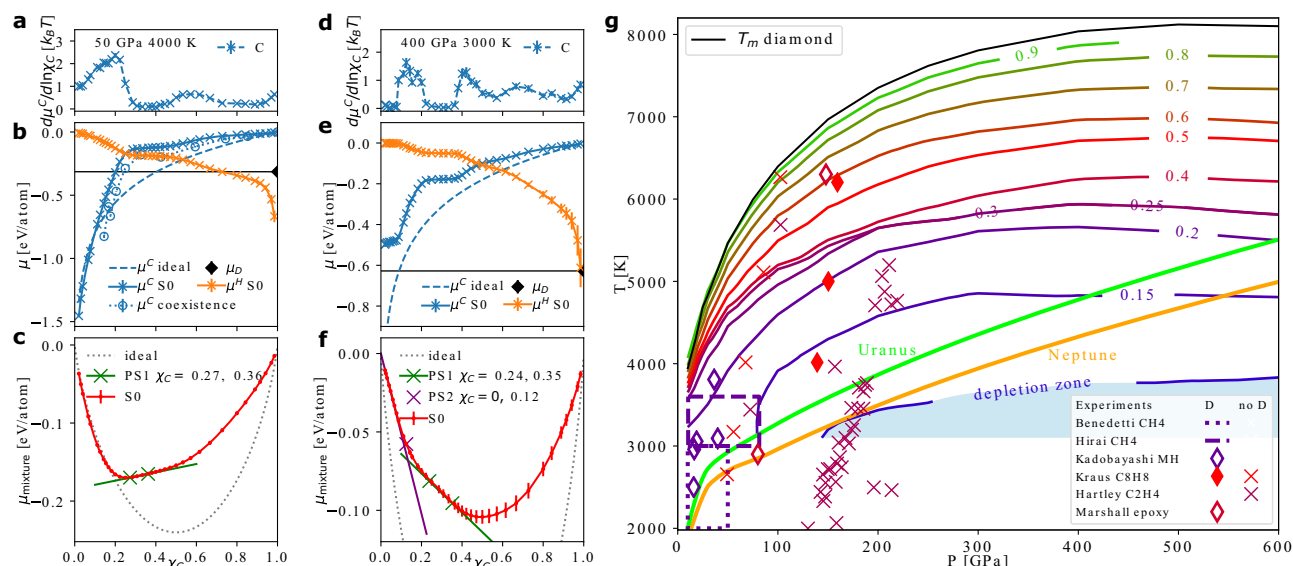
To determine the lifetimes of the C-C and the C-H bonds, we recorded the time it takes for a newly formed bond to dissociate during the MD simulations. Figure 2d, e show the average bond lifetimes. The C-H bond lifetimes are extremely short, less than about 0.01 ps. The C-C bonds are more long-lived, yet only have a mean lifetime of less than about 1 ps at all the conditions considered here. Such short lifetimes are consistent with previous DFT MD simulations of CH<sub>4</sub><sup>24</sup>.

The short bond lifetimes indicate that the hydrocarbon chains in the systems decompose and form quickly. In other words, the C/H mixture behaves like a liquid with transient C-C and C-H bonds.

### Thermodynamics of C/H mixtures

We then determine the chemical potentials of carbon in C/H mixtures,  $\Delta\mu^C(\chi_C)$ , as a function of the atomic fraction of carbon,  $\chi_C = N_C/(N_C + N_H)$ . This, combined with the chemical potential difference  $\Delta\mu_D$  between diamond and pure carbon liquid, establishes the thermodynamic phase boundary for diamond formation from C/H mixtures with varying atomic ratios.

Dilution will usually lower the chemical potential of carbon in a mixture, which can be understood from the ideal solution assumption:  $\mu_{id}^C = k_B T \ln(\chi_C)$ . However, the ideal solution model neglects the atomic interactions. To consider non-ideal mixing effects, we compute the chemical potentials of mixtures using the MLP. This is not an easy task, because traditional particle insertion methods<sup>35</sup> fail for this dense liquid system, and thermodynamic integration from an ideal gas state to the real mixture<sup>36</sup> is not compatible with the MLP. We employ the newly developed SO method which accounts for both the ideal and



**Fig. 3 | Thermodynamic behaviors of C/H mixtures.** **a,d**,  $d\mu^C/d\ln\chi_C$  computed from static structure factors using Eqn. (1), at 50 GPa, 4000 K (**a**) and at 400 GPa, 3000 K (**d**). **b,e** The chemical potential per carbon atom ( $\mu^C$ ) in C/H mixtures, estimated using the S0 method (blue crosses connected by solid curve), the coexistence method (blue hollow symbols connected by dotted curve), and the ideal solution approximation (blue dashed curve). The chemical potential per hydrogen atom ( $\mu^H$ ) is shown as orange line with crosses. The chemical potential of diamond ( $\mu_D$ ) is indicated by the black diamond symbol and the black horizontal line. **c,f** The free energy per atom of C/H mixtures with different carbon fractions  $\chi_C$ . The free energy of the ideal solution is shown as dashed gray curve, and the result from the S0 method is shown in red. The green line illustrates the constructed double tangent, and the two green crosses show the composition of the liquids forming after phase separation (PS1). The purple line is a tangent and the purple

cross shows the composition of the more carbon-rich phase after PS2. The standard mean errors are indicated using the error bars in **a-f,g**. The conditions below which diamond formation is thermodynamically possible at different carbon ratios (colored solid lines). The shaded area below the bright blue line indicates the depletion zone, where diamond formation is thermodynamically favorable regardless of the carbon ratio. Experimental observation of diamonds (D) is indicated by diamond symbols and dashed/dotted rectangular regions, while no diamond observation (no D) is marked by cross symbols: laser-heated DAC data for methane<sup>11,12</sup>, for methane hydrate (MH)<sup>14</sup>, and shock-compression experiments for polystyrene ( $-C_8H_8$ )<sup>16</sup>, polyethylene ( $-C_2H_4$ )<sup>17</sup>, and epoxy<sup>15</sup>. The symbols are colored according to the  $N_C/(N_C + N_H)$  ratio of the starting materials, using the same color scheme for the phase boundaries. The  $P$ - $T$  curves of planetary interior conditions for Uranus (green line) and Neptune (orange line) are from Ref. <sup>40</sup>.

the non-ideal contributions to the chemical potentials<sup>37</sup>:

$$\left(\frac{d\mu^C}{d\ln\chi_C}\right)_{T,P} = \frac{k_B T}{(1-\chi_C)S_{CC}^0 + \chi_C S_{HH}^0 - 2\sqrt{\chi_C(1-\chi_C)}S_{CH}^0}, \quad (1)$$

where  $S_{CC}^0$ ,  $S_{CH}^0$ , and  $S_{HH}^0$  are the values of the static structure factor between the said types of atoms at the limit of infinite wavelength<sup>37</sup>, which can be determined from equilibrium MD simulations of a C/H mixture with a given carbon fraction  $\chi_C$ .  $\mu^H$  is then fixed using the Gibbs-Duhem equation. Note that only the relative chemical potential is physically meaningful, and we conveniently select the reference states to be the pure carbon and hydrogen liquids, i.e.  $\mu^C(\chi_C=1)=0$  and  $\mu^H(\chi_C=0)=0$ . We obtained  $\mu^C$  and  $\mu^H$  at different  $\chi_C$  on a grid of  $P$ - $T$  conditions between 10 GPa–600 GPa and 3000 K–8000 K, by numerically integrating  $d\mu^C/d\ln\chi_C$ .

$d\mu^C/d\ln\chi_C$  at  $P=50$  GPa,  $T=4000$  K and  $P=400$  GPa,  $T=3000$  K are shown in Fig. 3a, d, respectively. For both sets, these values deviate from the ideal behavior (i.e. constant at 1), and have maxima and minima around certain compositions. The corresponding chemical potentials are plotted in Fig. 3b, e, while the results at other conditions are shown in Fig. 4 of the Methods. As an independent validation, we also computed  $\mu^C$  using the coexistence method described in the Methods, although this approach is in general less efficient and can become prohibitive if carbon concentration or diffusivity is low. The values from the coexistence method are shown as the hollow symbols in Fig. 3b, in agreement with the S0 method. As the statistical accuracy of the S0 method is much better compared to the coexistence approach, all the subsequent analysis is based on the former.

In both Fig. 3b and e,  $\mu^C$  has a plateau at  $\chi_C$  between about 0.25 and 0.35, and the same phenomenon is found at  $T \leq 5000$  K at 50 GPa, and

at even broader temperature range under increasing pressures, up to 8000 K at 600 GPa (see Fig. M4 of the Methods). At 50 GPa, 4000 K (Fig. 3b),  $\mu^C$  then decreases rapidly and approaches the ideal behavior at lower  $\chi_C$ . In contrast, at 400 GPa, 3000 K (Fig. 3e),  $\mu^C$  plateaus and reaches a constant value for  $\chi_C < 0.12$ . The plateaus at low  $\chi_C$  were observed at pressures between 200 GPa and 600 GPa and temperatures lower than 3500 K (see Fig. M4 of the Methods). In Fig. 3b,e, the chemical potentials of diamond,  $\mu_D$ , are indicated by black diamond symbols and horizontal lines. If  $\mu^C$  is larger than  $\mu_D$  at a given  $\chi_C$ , diamond formation is thermodynamically favorable.

To rationalize the plateaus, we express the per-atom chemical potential of the C/H mixture as

$$\mu_{mixture}(\chi_C) = \chi_C \mu^C(\chi_C) + (1-\chi_C) \mu^H(\chi_C), \quad (2)$$

and compare it to the ideal solution curve  $\mu_{mixture,id} = k_B T(\chi_C \log(\chi_C) + (1-\chi_C) \log(1-\chi_C))$ . Figure 3c shows  $\mu_{mixture}$  at 50 GPa, 4000 K. Compared with the ideal solution chemical potential (dashed gray curve) which is fully convex,  $\mu_{mixture}$  has two edges. One can thus perform a common tangent construction to the  $\mu_{mixture}$  curve to find out the coexisting liquid phases. The green line in Fig. 3d indicates the common tangent, and the two green crosses show the location of the edges. For C/H mixtures with  $\chi_C^1 = 0.27$  and  $\chi_C^2 = 0.36$  at the condition shown, a liquid-liquid phase separation (PS1) will occur and form two phases with the proportions determined by the lever rule. Here the region between the two edges is not concave but linear, which is because the phase separation has little activation barrier and already occurs during the MD simulations. In other words, a C/H mixture with a carbon fraction that is between the values of  $\chi_C^1$  and  $\chi_C^2$  will first undergo spontaneous

liquid-liquid phase separation, which explains the corresponding plateaus in  $\mu^C$  of Fig. 3b,e.

Furthermore, Fig. 3f shows that, at 400 GPa, 3000 K,  $\mu_{mixture}$  at low  $\chi_C$  significantly deviates from the ideal solution approximation (dashed gray curve), and one can construct a tangent as plotted in purple. This means that, besides the aforementioned PS1, C/H mixtures at a low C fraction can also phase separate (PS2) into a fluid of mostly hydrogen and another fluid with  $\chi_C \approx 0.12$  (purple cross). We show example snapshots of such phase separated configurations collected from the MD simulations in the Supplementary Information. This PS2 explains the plateau of  $\mu^C$  at low  $\chi_C$  in Fig. 3d, as the carbon concentrations in both phase-separated liquids stay the same, while only the proportions of the two liquids change. Supplementary Movie 1 shows the occurrence of PS2 in MD simulations. This phase separation has immense consequences: at pressures above 200 GPa and temperatures below 3000 K–3500 K, C in C/H mixtures will always have  $\mu^C > \mu_D$  even at very low C fraction due to PS2, and the carbon atoms will thus always be under a thermodynamic driving force to form diamond. We refer to these conditions as the “depletion zone”.

Figure 3g presents the thermodynamic phase boundaries, below which diamond formation is possible in C/H mixtures for each indicated carbon atomic fraction. This is obtained by combining the values of  $\mu^C(\chi_C)$  in C/H mixtures and  $\mu_D$  at a wide range of  $P$ - $T$  conditions. For lower and lower  $\chi_C$ , the boundaries deviate more and more from the  $T_m$  of diamond. At  $P < 100$  GPa, the locations of the boundaries are very sensitive to both temperature and pressure, whereas at higher  $P$  it is mostly independent of pressure. Figure 3g can also be read in another way: for a certain  $P$ - $T$  condition, it gives the minimal carbon ratio required to make diamond formation possible. Notice that the  $\chi_C = 0.25$  and  $\chi_C = 0.3$  lines almost overlap, which is due to the plateau of  $\mu^C$  induced by PS1. The light blue shaded area indicates the depletion zone, where diamond formation is always possible due to PS2. In this zone, carbon atoms will first form a carbon-rich liquid phase, and diamond can nucleate from this phase. Such process is similar with a two-step nucleation mechanism previously revealed in protein systems<sup>38</sup>.

Previous experimental measurements are included in Fig. 3g, with the conditions where diamonds were either found (diamond symbols or rectangular regions) or absent (cross symbols) indicated. At lower pressures, our calculations largely agree with the observation of diamond formation for methane in DAC experiments between 2000–3000 K<sup>11</sup> and above 3000 K<sup>12</sup>. We find less agreement with the shock-compression experiments at higher pressures<sup>16,17</sup>. We speculate that the disagreement may be because diamond formation needs to go through the activated nucleation process which may take longer than the short timescale of these rapid compression experiments, or may come from the difficulty in the temperature estimation of these experiments. The hollow diamond symbols in Fig. 3g show the diamond formation conditions from starting materials of more complex compositions: Marshall et al.<sup>15</sup> used epoxy (C:H:Cl:N:O  $\approx$  27:38:1:1:5) and Kadobayashi et al.<sup>14</sup> used methane hydrate. We find little agreement between Kadobayashi et al.<sup>14</sup> and our phase boundaries, if we compare solely in terms of  $\chi_C$  including all atomic species, although the agreement is improved if based on the  $\chi_C$  of methane alone, and indeed CH<sub>4</sub> may be an intermediate product in the experiment<sup>14</sup>. The liquid-liquid phase separations of C/H mixtures have not been previously observed, but they may be detected from speed of sound, mixed optical spectra, inhomogeneity in the diamond formation reaction, and hydrodynamic instability during compression experiments.

## Discussion

We first computed the melting line of diamond in pure liquid carbon. We then moved on to the C/H mixtures, and showed that they behave like liquids at  $T \geq 2500$  K. We finally precisely computed the thermodynamic boundary of diamond formation for different atomic ratios. Notably, we revealed the occurrence of phase

separations in C/H mixtures, which can greatly enhance diamond formation. For PS1, the C/H mixture will phase separate into two liquids with  $\chi_C$  of about 0.25 and 0.35. Both liquids have the same  $\mu^C$ , but their interfacial free energies with the diamond phase are different. Diamond will thus prefer to nucleate from the liquid with the lower interfacial free energy. At  $200 < P < 600$  GPa and  $T$  below 3000 K–3500 K, there is a depletion zone where C/H mixtures at a low C fraction can phase separate (PS2) into a fluid of mostly hydrogen and a more carbon-rich fluid with  $\chi_C \approx 0.12$ . In this zone, there is always thermodynamic driving force to form diamond from the carbon-rich phase.

Our phase boundaries in Fig. 3g put largely scattered experimental measurements<sup>11,12,15–17</sup> into context, and provide a mechanistic understanding of the thermodynamics involved. They also help gauge the accuracy of the experimental determination of diamond formation conditions, extrapolate between different experiments, and guide future efforts to validate these boundaries.

Note that our boundaries are solely based on the thermodynamic criterion, but the kinetic nucleation rate may play a role particularly in shock-compression experiments. An amount of undercooling may be needed for diamond to nucleate from C/H mixtures within finite time, depending on the magnitude of the nucleation activation barrier. In homogeneous nucleation, the magnitude of the interfacial free energy contribution is crucial<sup>25,32</sup>. In experiments, DAC are in close contact with the fluid samples, so heterogeneous nucleation may happen, which requires less undercooling compared with the homogeneous case. In addition, other elements (e.g. He, N, O) are also prevalent in icy planets, and we suggest future experiments to probe how they affect the phase boundaries of diamond formation.

The “depletion zone” can help explain the difference in the luminosity between Uranus and Neptune. Being similar in size and composition, Neptune has a strong internal heat source but Uranus does not<sup>39</sup>. The “diamonds in the sky” hypothesis<sup>8,11</sup> relates the heat source with diamond formation, but does not explain the dichotomy between the two planets. By comparing the  $P$ - $T$  conditions at different depths of the two ice giants from Ref.<sup>40</sup> with our calculated phase boundaries (in Fig. 3g), one can see that a relatively small difference in the planetary profile can drastically change the possibility of diamond formation: At the  $P$ - $T$  conditions in Uranus, diamond formation requires about 15% of carbon, which seems unlikely as less than 10% of carbon is believed to be present in its mantle<sup>4</sup>. As such, diamond formation in Uranus may be absent. In contrast, Neptune is a bit cooler so it is much more likely that its planetary profile may have an overlap with the depletion zone; at these conditions C/H mixtures will phase separate (PS2), and diamond formation is thermodynamically favorable regardless of the actual carbon fraction. If there is indeed an overlap, diamonds can in principle form in the depletion zone in the mantle of Neptune, and then sink towards the core while releasing heat. Although the mantle will become increasingly carbon-depleted, the diamond formation in the depletion zone can proceed until all carbon is exhausted. Moreover, the “diamond rain” will naturally induce a compositional gradient inside the planet, which is an important aspect in explaining the evolution of giant planets<sup>41,42</sup>.

Our carbon-ratio-dependent diamond formation phase boundaries can help estimate the prevalence and the existence criteria of extraterrestrial diamonds. Neptune-like exoplanets are extremely common according to the database of planets discovered<sup>43</sup>, and methane-rich exoplanets are modeled to have a carbon core, a methane envelope, and a hydrogen atmosphere<sup>7</sup>. Our boundaries can put a tight constraint on the structure and composition of these planets. Furthermore, diamond formation and liquid-liquid phase separation play a key role in the cooling process in white dwarfs<sup>10</sup>, and thus the precise determination for the onset of phase separation and crystallization is also crucial there.

## Methods

### DFT calculations

DFT is the workhorse of high-pressure equation-of-state calculations and has shown good agreement with several experiments on hydrocarbons and other systems<sup>44–46</sup> for measured thermodynamic properties in particular for Hugoniot curves. Single-point DFT calculations with VASP<sup>47–50</sup> were carried out for configurations with various C/H ratios to generate the training set of the MLP. The simulations were performed with the Perdew-Burke-Ernzerhof (PBE) exchange-correlation functional<sup>51</sup> employing hard pseudopotentials for hydrogen and carbon, a cutoff energy of 1000 eV, and a consistent k-point spacing of 0.2 Å<sup>-1</sup>. In addition, extensive PBE MD simulations for CH<sub>16</sub>, CH<sub>8</sub>, CH<sub>4</sub>, CH<sub>2</sub>, CH, C<sub>2</sub>H and C<sub>4</sub>H mixtures were performed, and together with previous PBE MD data for methane<sup>52</sup>, carbon<sup>30</sup> and hydrogen<sup>53</sup>, were used to benchmark the MLP. To approximate the impact of the thermal excitation of the electronic subsystem, we set the electronic temperature equal to the average ionic temperature during the DFT MD calculations as well as in the reference calculations used to train and test the MLP. The convergence tests of DFT and the influence of the electronic temperature are provided in the Supplementary Information.

### Machine learning potential

We generated flexible and dissociable MLPs for the high-pressure C/H system, employing the Behler-Parrinello artificial neural network<sup>54</sup>, and using the N2P2 code<sup>55</sup>. The total training set contains 92,185 configurations with a sum of 8,906,582 atoms, and was constructed using a combination of strategies, including DFT MD, random structure searches, adapting previous training sets for pure C<sup>56</sup> and H<sup>53</sup>, and active learning. The training set includes a large variety of structures: cubic/hexagonal diamond, graphite, graphane, carbon nanotubes, fullerenes, amorphous carbon, carbon structures with defects, liquid carbon, liquid hydrogen, many hydrogen crystalline polymorphs, hydrocarbon crystals, hydrocarbon liquids with varying carbon concentrations at a wide range of *P-T* conditions. Details on the construction and the benchmarks of the MLP are provided in the Supplementary Information. Note that the MLP has been extensively benchmarked for high-pressure liquid hydrogen, diamond/liquid carbon, and C/H mixtures based on energetic, thermodynamic and dynamic properties. However, we would like to caution the limitations of the current MLP: The MLP is not applicable to gas-phase hydrocarbons. For low-pressure carbon phases and diamond-graphite transitions, the current MLP has not been extensively tested, and we recommend users to employ the MLP from Ref. <sup>56</sup>. Long-range Van der Waals interactions in liquid methane at low density may be important<sup>57</sup> but are lacking in the current MLP. The comparison between PBE and MLP for structure and dynamic properties such as equation of states, radial distribution functions, diffusivity, vibrational density of states, and bond lifetimes are provided in the Supplementary Information. We recommend checking these comparisons before applying the MLP for a given C/H composition at certain conditions.

### MLP MD simulation details

All MD simulations were performed in LAMMPS<sup>58</sup> with a MLP implementation<sup>59</sup>. The time step was 0.25 fs for C/H mixtures, and 0.4 fs for pure carbon systems.

### Computing the chemical potentials of diamond and pure liquid carbon

We computed  $\Delta\mu_D$  using interface pinning simulations<sup>60</sup>, which were performed using the PLUMED code<sup>61</sup>. We used solid-liquid systems containing 1,024 C atoms at pressures between 10–800 GPa and at temperatures close to the melting line, employing the MLP. The Nosé-Hoover barostat was used only along the *z* direction that is perpendicular to the interface in these coexistence

simulations, while the dimensions of the supercell along the *x* and *y* directions were commensurate with the equilibrium lattice parameters of the diamond phase at the given conditions. We used the locally-averaged<sup>62</sup>  $Q_3$  order parameter<sup>63</sup> for detecting diamond structures, and introduced an umbrella potential to counter-balance the chemical potential difference and constrain the size of diamond in the system. We then used thermodynamic integration along isotherms and isobars<sup>64,65</sup> to extend the  $\Delta\mu_D$  to a wide range of pressures and temperatures.

### MLP MD simulation of CH<sub>4</sub>

The simulation cell contained 7,290 atoms (1,458 CH<sub>4</sub> formula units). Each simulation was run for more than 100 ps. The simulations were performed in the NPT ensemble, using the Nosé-Hoover thermostat and isotropic barostat. At each condition, two independent MD simulations were initialized using a starting configuration of either bonded CH<sub>4</sub> molecules on a lattice or a liquid. For  $T \geq 2500$  K, the two simulations provided consistent statistical properties. These simulations were the basis for the further analysis we performed. For  $T < 2500$  K the two runs gave different averages, meaning that under these conditions the system is not ergodic within the simulation time.

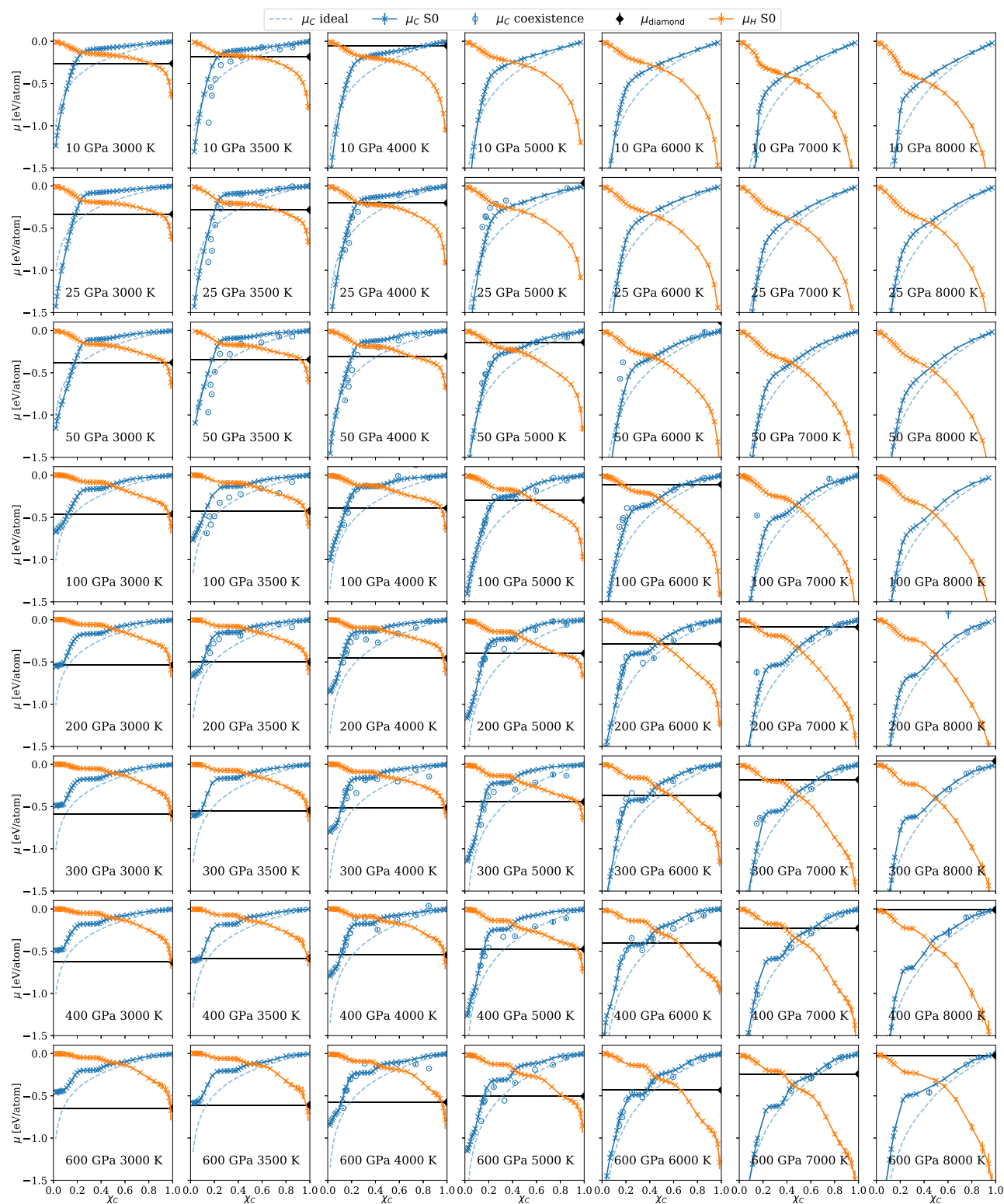
### Computing the chemical potentials of C in C/H mixtures

We used two independent methods for computing the chemical potentials of carbon in C/H mixtures at various conditions. The first is the S0 method<sup>37</sup> that uses the static structure factors computed from equilibrium NPT simulations. The S0 method uses the thermodynamic relationship between composition fluctuations and the derivative of chemical potential with respect to concentration, and accounts for both mixing entropy and enthalpy<sup>37</sup>. The simulations were performed on a grid of *P-T* conditions,  $P = 10$  GPa, 25 GPa, 50 GPa, 100 GPa, 200 GPa, 300 GPa, 400 GPa, 600 GPa, and  $T = 3500$  K, 4000 K, 5000 K, 6000 K, 7000 K, and 8000 K. At each *P-T* condition, MD simulations were run for systems at varying atomic ratios, on a dense grid of  $\chi_C$  from 0.015 to 0.98. The system size varied between 9,728 and 82,944 total number of atoms. We obtained the static structure factors at different wavevectors  $\mathbf{k}$  using the Fourier expansion on the scaled atomic coordinates, i.e.

$$S_{AB}(\mathbf{k}) = \frac{1}{\sqrt{N_A N_B}} \left\langle \sum_{i=1}^{N_A} \exp(i\mathbf{k} \cdot \hat{\mathbf{r}}_i(t)) \sum_{j=1}^{N_B} \exp(-i\mathbf{k} \cdot \hat{\mathbf{r}}_j(t)) \right\rangle \quad (3)$$

where  $AB$  can be CC, CH and HH, and  $\hat{\mathbf{r}}(t) = \mathbf{r}(t)/l(t)_{\text{NPT}}/l(t)$  and  $l(t)$  is the instantaneous dimension of the supercell. We then determined  $S_{CC}^0$ ,  $S_{CH}^0$  and  $S_{HH}^0$  by extrapolating  $S_{AB}(\mathbf{k})$  to the  $\mathbf{k} \rightarrow 0$  case using the Ornstein-Zernike form as described in Ref. <sup>37</sup>. Finally, we used numerical integration using Eqn. (3) of the main text to obtain the chemical potential of carbon for different atomic fractions, and get the chemical potential of H using the Gibbs-Duhem equation. All the chemical potential data are presented in Fig. 4.

The second approach is based on the coexistence method, similar to the setup used for computing the chemical potentials of the pure carbon systems. In this case, interface pinning simulations<sup>60,66</sup> were performed on a diamond-C/H liquid coexistence system containing 1024 C atoms and a varying number of H atoms at pressures 0–600 GPa. A snapshot of the coexistence system is provided in the Supplementary Information. The chemical potentials estimated using coexistence are shown in Fig. 4, and the errors shown are the standard errors of the mean estimated from the values of the CV. However, there are other sources of errors that are hard to estimate: finite size effects and ergodicity issues related to the explicit interface; the carbon concentration can vary in the liquid region of the simulation box.



**Fig. 4 | The chemical potentials of carbon and hydrogen in C/H mixtures at different temperature and pressure conditions.** The chemical potentials per carbon atom in C/H mixtures calculated using the S0 method are shown as blue crosses connected by solid curves, the values obtained using the coexistence method are shown as blue hollow symbols connected by dotted curves, and the

ideal solution approximation is indicated by blue dashed curves. The chemical potentials per hydrogen atom calculated using the S0 method are shown as orange crosses. The error bars show the standard mean errors. The chemical potential of diamond is indicated by the black diamond symbol and the black horizontal line.

## Data availability

All original data generated for the study, including the MLP, the training set, simulation input files, intermediate data, PYTHON notebook, are in the SI repository <https://github.com/BingqingCheng/high-ch><sup>67</sup>.

## Code availability

The MD simulations were performed using the LAMMPS code<sup>58</sup> with a MLP implementation<sup>59</sup>.

## References

1. Suess, H. E. & Urey, H. C. Abundances of the Elements. *Rev. Mod. Phys.* **28**, 53 (1956).
2. Mastrogiuseppe, M. et al. Deep and methane-rich lakes on Titan. *Nat. Astron.* **3**, 535 (2019).
3. Bertrand, T., Forget, F., Schmitt, B., White, O. L. & Grundy, W. M. Equatorial mountains on Pluto are covered by methane frosts resulting from a unique atmospheric process. *Nat. Commun.* **11**, 1 (2020).
4. Guillot, T. and Gautier, D., 10.16 - giant planets, in <https://doi.org/10.1016/B978-0-444-53802-4.00176-7> *Treatise on Geophysics (Second Edition)*, edited by Schubert, G. (2015) second edition ed., pp. 529–557.
5. Nettelmann, N. et al. Uranus evolution models with simple thermal boundary layers. *Icarus* **275**, 107 (2016).
6. Irwin, P. G. J. et al. Hazy blue worlds: A holistic aerosol model for Uranus and Neptune, including Dark Spots. *J. Geophys. Res. Planets* **127**, e2022JE007189 (2022).
7. Helled, R., Podolak, M. & Vos, E. Methane planets and their mass-radius relation. *Astrophys. J. Lett.* **805**, L11 (2015).
8. Ross, M. The ice layer in Uranus and Neptune—diamonds in the sky? *Nature* **292**, 435 (1981).
9. Pearl, J. C. & Conrath, B. J. The albedo, effective temperature, and energy balance of Neptune, as determined from Voyager data. *J. Geophys. Res. Space Phys.* **96**, 18921 (1991).
10. Tremblay, P.-E. et al. Core crystallization and pile-up in the cooling sequence of evolving white dwarfs. *Nature* **565**, 202 (2019).
11. Benedetti, L. R. et al. Dissociation of CH<sub>4</sub> at high pressures and temperatures: diamond formation in giant planet interiors? *Science* **286**, 100 (1999).
12. Hirai, H., Konagai, K., Kawamura, T., Yamamoto, Y. & Yagi, T. Polymerization and diamond formation from melting methane and their implications in ice layer of giant planets. *Phys. Earth Planet. Inter.* **174**, 242 (2009).
13. Lobanov, S. S. et al. Carbon precipitation from heavy hydrocarbon fluid in deep planetary interiors. *Nat. Commun.* **4**, 2446 (2013).
14. Kadobayashi, H. et al. Diamond formation from methane hydrate under the internal conditions of giant icy planets. *Sci. Rep.* **11**, 8165 (2021).
15. Marshall, M. C. et al. Diamond formation in double-shocked epoxy to 150 GPa. *J. Appl. Phys.* **131**, 085904 (2022).
16. Kraus, D. et al. Formation of diamonds in laser-compressed hydrocarbons at planetary interior conditions. *Nat. Astron.* **1**, 606 (2017).
17. Hartley, N. J. et al. Evidence for crystalline structure in dynamically-compressed polyethylene up to 200 GPa. *Sci. Rep.* **9**, 4196 (2019).
18. Gao, G. et al. Dissociation of methane under high pressure. *J. Chem. Phys.* **133**, 144508 (2010).
19. Naumova, A. S., Lepeshkin, S. V. & Oganov, A. R. Hydrocarbons under pressure: Phase diagrams and surprising new compounds in the c-h system. *J. Phys. Chem. C* **123**, 20497 (2019).
20. Liu, H., Naumov, I. I. & Hemley, R. J. Dense hydrocarbon structures at megabar pressures. *J. Phys. Chem. Lett.* **7**, 4218 (2016).
21. Conway, L. J. & Hermann, A. High pressure hydrocarbons revisited: from van der Waals compounds to diamond. *Geosciences* **9**, 227 (2019).
22. Ishikawa, T. & Miyake, T. Evolutionary construction of a formation-energy convex hull: Practical scheme and application to a carbon-hydrogen binary system. *Phys. Rev. B* **101**, 214106 (2020).
23. Ancilotto, F., Chiarotti, G. L., Scandolo, S. & Tosatti, E. Dissociation of methane into hydrocarbons at extreme (planetary) pressure and temperature. *Science* **275**, 1288 (1997).
24. Sherman, B. L., Wilson, H. F., Weeraratne, D. & Militzer, B. Ab initio simulations of hot, dense methane during shock experiments. *Phys. Rev. B* **86**, 224113 (2012).
25. Ghiringhelli, L. M., Valeriani, C., Meijer, E. J. & Frenkel, D. Local structure of liquid carbon controls diamond nucleation. *Phys. Rev. Lett.* **99**, 055702 (2007).
26. Zylstra, A. B. et al. Burning plasma achieved in inertial fusion. *Nature* **601**, 542 (2022).
27. Eggert, J. H. et al. Melting temperature of diamond at ultrahigh pressure. *Nat. Phys.* **6**, 40 (2010).
28. Wang, X., Scandolo, S. & Car, R. Carbon phase diagram from ab initio molecular dynamics. *Phys. Rev. Lett.* **95**, 185701 (2005).
29. Correa, A. A., Bonev, S. A. & Galli, G. Carbon under extreme conditions: Phase boundaries and electronic properties from first-principles theory. *Proc. Natl Acad. Sci. USA* **103**, 1204 (2006).
30. Benedict, L. X. et al. Multiphase equation of state for carbon addressing high pressures and temperatures. *Phys. Rev. B* **89**, 224109 (2014).
31. Ghiringhelli, L. M., Los, J. H., Meijer, E. J., Fasolino, A. & Frenkel, D. Modeling the phase diagram of carbon. *Phys. Rev. Lett.* **94**, 145701 (2005).
32. Kelton, K. F. Crystal nucleation in liquids and glasses, in [https://doi.org/10.1016/s0081-1947\(08\)60144-7](https://doi.org/10.1016/s0081-1947(08)60144-7) *Solid state physics*, Vol. 45 (Elsevier, 1991) pp. 75–177.
33. Betti, R. & Hurricane, O. A. Inertial-confinement fusion with lasers. *Nat. Phys.* **12**, 435 (2016).
34. Zha, C.-s., Liu, H., John, S. T. & Hemley, R. J. Melting and high p-t transitions of hydrogen up to 300 gpa. *Phys. Rev. Lett.* **119**, 075302 (2017).
35. Frenkel, D. Simulations: The dark side. *Eur. Phys. J.* **128**, 10 (2013).
36. Li, L., Totton, T. & Frenkel, D. Computational methodology for solubility prediction: Application to the sparingly soluble solutes. *J. Chem. Phys.* **146**, 214110 (2017).
37. Cheng, B. *J. Chem. Phys.* **157**, 121101 (2022).
38. ten Wolde, P. R. & Frenkel, D. Enhancement of protein crystal nucleation by critical density fluctuations. *Science* **277**, 1975 (1997).
39. Helled, R. & Fortney, J. J. The interiors of Uranus and Neptune: current understanding and open questions. *Philos. Trans. R. Soc. A* **378**, 20190474 (2020).
40. Scheibe, L., Nettelmann, N. & Redmer, R. Thermal evolution of Uranus and Neptune. *Astron. Astrophys.* **632**, A70 (2019).
41. Vazan, A. & Helled, R. Explaining the low luminosity of Uranus: a self-consistent thermal and structural evolution. *Astron. Astrophys.* **633**, A50 (2020).
42. Mankovich, C. R. & Fortney, J. J. Evidence for a dichotomy in the interior structures of Jupiter and Saturn from helium phase separation. *Astrophys. J.* **889**, 51 (2020).
43. Zeng, L. et al. Growth model interpretation of planet size distribution. *Proc. Natl Acad. Sci. USA* **116**, 9723 (2019).
44. Knudson, M. D. & Desjarlais, M. P. Interplay of high-precision shock wave experiments with first-principles theory to explore molecular systems at extreme conditions: A perspective. *J. Appl. Phys.* **129**, 210904 (2021).
45. Millot, M. et al. Experimental evidence for superionic water ice using shock compression. *Nat. Phys.* **14**, 297 (2018).
46. Kim, Y.-J. et al. Evidence for dissociation and ionization in shock compressed nitrogen to 800 gpa. *Phys. Rev. Lett.* **129**, 015701 (2022).



47. Kresse, G. & Hafner, J. Ab initio molecular dynamics for liquid metals. *Phys. Rev. B* **47**, 558 (1993).
48. Kresse, G. & Hafner, J. Ab initio molecular dynamics for open-shell transition metals. *Phys. Rev. B* **48**, 13115 (1993).
49. Kresse, G. & Hafner, J. Ab initio molecular-dynamics simulation of the liquid-metal-amorphous-semiconductor transition in germanium. *Phys. Rev. B* **49**, 14251 (1994).
50. Kresse, G. & Furthmüller, J. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Phys. Rev. B* **54**, 11169 (1996).
51. Perdew, J. P., Burke, K. & Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **77**, 3865 (1996).
52. Bethkenhagen, M. et al. Planetary ices and the linear mixing approximation. *Astrophys. J.* **848**, 67 (2017).
53. Cheng, B., Mazzola, G., Pickard, C. J. & Ceriotti, M. Evidence for supercritical behaviour of high-pressure liquid hydrogen. *Nature* **585**, 217 (2020).
54. Behler, J. & Parrinello, M. Generalized neural-network representation of high-dimensional potential-energy surfaces. *Phys. Rev. Lett.* **98**, 146401 (2007).
55. Singraber, A., Morawietz, T., Behler, J. & Dellago, C. Parallel multi-stream training of high-dimensional neural network potentials. *J. Chem. Theory Comput.* **15**, 3075 (2019).
56. Rowe, P., Deringer, V. L., Gasparotto, P., Csányi, G. & Michaelides, A. An accurate and transferable machine learning potential for carbon. *J. Chem. Phys.* **153**, 034702 (2020).
57. Veit, M. et al. Equation of state of fluid methane from first principles with machine learning potentials. *J. Chem. theory Comput.* **15**, 2574 (2019).
58. Plimpton, S. Fast Parallel Algorithms for Short-Range Molecular Dynamics. *J. Comput. Phys.* **117**, 1 (1995).
59. Singraber, A., Behler, J. & Dellago, C. Library-based LAMMPS implementation of high-dimensional neural network potentials. *J. Chem. Theory Comput.* **15**, 1827 (2019).
60. Pedersen, U. R., Hummel, F., Kresse, G., Kahl, G. & Dellago, C. Computing gibbs free energy differences by interface pinning. *Phys. Rev. B* **88**, 94101 (2013).
61. Tribello, G. A., Bonomi, M., Branduardi, D., Camilloni, C. & Bussi, G. Plumed 2: New feathers for an old bird. *Comput. Phys. Commun.* **185**, 604 (2014).
62. Lechner, W. & Dellago, C. Accurate determination of crystal structures based on averaged local bond order parameters. *J. Chem. Phys.* **129**, 114707 (2008).
63. Steinhardt, P. J., Nelson, D. R. & Ronchetti, M. Bond-orientational order in liquids and glasses. *Phys. Rev. B* **28**, 784 (1983).
64. Cheng, B. & Ceriotti, M. Computing the absolute Gibbs free energy in atomistic simulations: Applications to defects in solids. *Phys. Rev. B* **97**, 054102 (2018).
65. Reinhardt, A. & Cheng, B. Quantum-mechanical exploration of the phase diagram of water. *Nat. Commun.* **12**, 1 (2021).
66. Cheng, B., Dellago, C. & Ceriotti, M. Theoretical prediction of the homogeneous ice nucleation rate: Disentangling thermodynamics and kinetics. *Phys. Chem. Chem. Phys.* **20**, 28732 (2018).
67. Cheng, Bingqing <https://doi.org/10.5281/ZENODO.7578498> Data set for "thermodynamics of diamond formation from hydrocarbon mixtures in planets", (2023).

## Acknowledgements

BC thanks Daan Frenkel for stimulating discussions. We thank Aleks Reinhardt, Daan Frenkel, Marius Millot, Federica Coppari, Rhys Bunting, and Chris J. Pickard for critically reading the manuscript and providing useful suggestions. BC acknowledges resources provided by the Cambridge Tier-2 system operated by the University of Cambridge Research Computing Service funded by EPSRC Tier-2 capital grant EP/P020259/1. SH acknowledges support from LDRD 19-ERD-031 and computing support from the Lawrence Livermore National Laboratory (LLNL) Institutional Computing Grand Challenge program. Lawrence Livermore National Laboratory is operated by Lawrence Livermore National Security, LLC, for the U.S. Department of Energy, National Nuclear Security Administration under Contract DE-AC52-07NA27344. MB acknowledges support by the European Horizon 2020 program within the Marie Skłodowska-Curie actions (xICE grant number 894725), funding from the NOMIS foundation and computational resources at the North-German Supercomputing Alliance (HLRN) facilities.

## Author contributions

B.C., S.H., and M.B conceived the idea of studying high-pressure methane; B.C. designed the research; B.C. performed the simulations related to the MLP; B.C., S.H., and M.B. performed the DFT calculations; B.C., S.H., and M.B wrote the paper.

## Competing interests

All authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41467-023-36841-1>.

**Correspondence** and requests for materials should be addressed to Bingqing Cheng.

**Peer review information** : *Nature Communications* thanks the anonymous reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

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

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

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

© The Author(s) 2023