The impact of topology on the open-shell characters of benzenoid polycyclic hydrocarbons: a data-driven study

Yuyi Yan†, Fengru Zheng†, Jiayi Lu†, Hao Jiang†, Zhiwen Zhu†, Qiang Sun*†

†Materials Genome Institute, Shanghai University, 200444 Shanghai, China

ABSTRACT: Open-shell benzenoid polycyclic hydrocarbons (BPHs) with unpaired electrons can serve as functional materials for future quantum applications. However, there is a grand challenge in the search and realization of molecules with desired open-shell characters from a gigantic chemical space of BPHs, which calls for strategies to direct both the fundamental understandings and experimental progress. Machine learning (ML) has recently become one of the most exciting tools entering materials science for accelerating data analysis and data mining. In this work, we have built up a structure database of BPHs through graphical enumeration, and performed data-driven analysis through the ML model. Combined with a simple model Hamiltonian calculation, we discover that the number of the internal vertices of the BPH graphs is strongly correlated with their open-shell characters. We further demonstrate the role of the phenalenyl subgraph in predicting the magnetic ground state. In addition, we conceive a class of interesting open-shell BPHs whose spin densities are globally distributed over the molecules, which may benefit their chemical stability. These data-driven results not only provide a database of open-shell BPHs, but also extend the well-known Lieb’s theorem and Ovchinnikov’s rule and provide a simple way for designing open-shell carbon nanostructures, which may facilitate the exploration of emerging quantum phases and the realization of technologically relevant magnetic materials.

INTRODUCTION

Carbon-based nanomaterials are promising components for next-generation mechanical and electronic devices due to their diverse chemical structures and flexible electronic properties. Exploring the spin degrees of freedom of unpaired electrons in open-shell molecules or radicals affords controls of quantum information and thus can be harnessed for future information storage and processing. Although most of the current magnetic materials involve d- or f-block elements, magnetisms originated from p or π electrons in carbon-based materials have several key fundamental advantages. For instance, in contrast to the localized nature of magnetic moments in metal atoms, unpaired π electrons could extend over several or tens of nanometers benefiting transferring and manipulation of spin information. Their magnetic configurations including the spin quantum number S and magnetic exchange interactions can be precisely predicted by simple approximations, and are highly tunable through variations in molecular size, shape and edge structure.

Benzenoid polycyclic hydrocarbons (BPHs), also known as polycyclic π-conjugated hydrocarbons or nanographenes, are fragments of graphene with the carbons at the edge passivated by mono-hydrogens. Each carbon atom in BPHs hosts one conjugated π electron, which may contribute to the magnetic properties. The studies on open-shell benzenoid polycyclic hydrocarbons date back more than a century ago and have remained a very active area of research hitherto. However, experimental progress in this area has been facing the challenges of synthesis and long-term stability owing to their reactive natures. It is not until recently, the advances in on-surface synthesis methods in ultrahigh-vacuum conditions have made it possible to synthesize highly reactive open-shell PBHs including triangulene, zethrene, heptathrene, Clar’s goblet, acenes, periacenes, perihenes, as well as other open-shell molecules, some of which were never realized before.

Despite the long-standing research activities on BPHs, most of the molecules including both open-shell and closed-shell ones remain synthetically unavailable or even theoretically unexplored because of their enormous structural diversity. The structures of BPHs can be mathematically abstracted as subgraphs of the hexagonal lattice with edges (carbon-carbon bonds) and vertices (conjugated carbon atoms) of degree 2 or 3, which underlies most of the mathematical algorithms for enumerating BPHs. The enumeration of the number N of BPHs with the number h of hexagons has been an important topic at the interface of mathematics and chemistry for a long history.
The first enumeration of the benzenoids in a chemical context may be traced back to 1968. With access to modern computers, this field immediately booms. Until 1989, the number of benzenoids with up to $h = 12$ was known. Ten years later, this number of $h$ had been improved to 21. Another breakthrough occurred in 2002 with the advent of a new algorithm, the number of $h$ has been raised to 35, and the number of BPH structures is revealed to be formulated as $N = k^h$ ($4.789 \leq k \leq 5.905$). Although the topology of PBHs is known to impact its open-shell characters, these mathematical algorithms are typically not meant to predict their magnetic properties. Through quantum chemistry methods, magnetic and electronic properties of BPHs can be calculated. However, the gigantic number makes it impractical to screen open-shell BPHs and to extract the key features determining their magnetic properties through quantum chemistry.

In recent years, machine learning (ML) has been one of the most exciting tools entering materials science, which could accelerate data analysis and data mining. As a more efficient material design technique than conventional computational methods, it builds statistic models from available datasets to make predictions and explore underlying rules. With a good balance between efficiency and cost, ML has been successfully applied for materials discovery, structure analysis, property prediction and reverse designs. Although the open-shell characters of BPHs are strongly structure-dependent, it has not been explored by a statistic or machine learning model to the best of our knowledge. Therefore, it is of utmost interest to study the open-shell characters of BPHs via the ML method and eventually develop designing rules for BPHs of interesting open-shell characters.

In this work, we established a database of structures of open-shell BPHs being fused with less than 10 hexagons. An approach of constructive nature was applied to enumerate all the possible structures. Open-shell BPHs were screened from the tight-binding (TB) quantum mechanical calculation, and all of their structures were output. We extracted seven structural features from the database as inputs for the following machine learning analysis. The importance and influence of features on magnetism were analyzed through an interpretable strategy combining machine learning with the Shapley additive explanations (SHAP) technique. Consequently, we established a simple rule which could predict the magnetic ground states of BPHs. Several kinds of open-shell BPHs were discussed and analyzed by the Hubbard model. Our results not only extend the application of the well-established Lieb’s theorem and Ovchinikov’s rule, but also provides a simple way based on structural topology for the design of open-shell BPHs.

Figure 1. Enumeration of benzenoid polycyclic hydrocarbons (BPHs). (a) Numbers of BPHs fused with $h$ hexagons ($h$ up to 9). Helicenoids that are not subgraphs of the hexagonal lattice are not included. The colors indicate the numbers of the total spin quantum number $S$ ($S = 0$, blue; $S = 1/2$, green; $S = 1$, yellow; $S = 3/2$, red). The BPHs with the Fjord edges (see red examples in Supporting section 1) are displayed in semitransparent colors in the lower row of each class. (b) A few examples of the BPH graphs with different numbers of $S$ and $h$. 

![Figure 1](image-url)
RESULTS AND DISCUSSION

To build up the molecular database in an efficient and constructive manner, we limit the number of benzene rings of BPHs \( h \) with less than 10. We enumerate all the different combinations of the hexagonal rings from \( h = 3 \) to \( h = 9 \) via the permutation-combination approach computationally, such that all the possible structures are obtained. Then, we remove the duplicated structures by screening the combinations with the same structural features (see Supporting section 3 for more details). On the basis of the computer-assisted structural enumeration, we could acquire not only all the graphs of BPHs but also the feature vectors used in the subsequent machine-learning model. To avoid the complexity of discussions, structures of helicenoids that are not subgraphs of the hexagonal lattice are not considered. The numbers (\( N \)) of BPHs composed of \( h \) hexagons we have obtained are consistent with the results determined previously (\( N = 81 \) when \( h = 6 \); \( N = 331 \) when \( h = 7 \); \( N = 1435 \) when \( h = 8 \); \( N = 6505 \) when \( h = 9 \))\(^{39} \), thus validating the reliability of our method (Fig. 1a and Table S1). In the summary plot of Fig. 1a, we pick out the structures with the Fjord edges which has two vertices in the closest proximity but being disconnected, and their numbers are displayed in the lower row of each class. Since BPHs of such structures suffer particularly strong distortions due to the steric hindrance resulting in highly non-planar configurations, we are not considering them in the following data-driven analysis.

A simple but particularly useful model for studying the magnetic properties in \( \pi \)-conjugated carbon materials is the tight-binding (TB) model.\(^{37} \) The tight-binding Hamiltonian of a bipartite honeycomb lattice can be analyzed using a mathematically rigorous approach of the benzenoid graph theory,\(^{38} \) which predicts the number of zero-energy states (also referred to as non-bonding or mid-gap states) equal to the graph’s nullity \( \eta = 2\alpha - N \), where \( N \) represents the total number of sites and \( \alpha \) is the maximum possible number of non-adjacent sites. The emergence of magnetism in the system is a result of the instability of the zero-energy states (the so-called Stoner criterion), which is determined by the competition of the exchange energy gain and the kinetic-energy penalty due to the spin polarization.\(^{39} \) Therefore, when we consider electron-electron interactions with on-site coulomb repulsion \( U \), the original zero states will split by \( \Delta_s = \frac{U}{2} \sum_i n_i^2 \), where \( \sum_i n_i^2 \) determines the degree of localization of the corresponding electronic states. The zero-energy states undergo spin polarization at any \( U > 0 \) irrespective of their degree of localization, in chemistry words, an open-shell state is generated.

Although the benzenoid graph theory and TB calculations are able to predict the occurrence of zero-energy states, it is not clear how the electron spins align in these states. The complementary knowledge can be supplied by Lieb’s theorem\(^{40} \) which determines the total \( S \) of a bipartite lattice described by the Hubbard model. This theorem states that in the case of repulsive electron-electron interactions, a bipartite system has the ground state characterized by the total spin \( S = \frac{1}{2} |N_A - N_B| \), where \( N_A \) and \( N_B \) are the numbers of sites in sublattices A and B, respectively. The two counting rules are linked by the relation, \( \eta \geq |N_A - N_B| \).\(^{41} \)

Due to the direct relation between the open-shell characters of BPHs and the spin quantum number \( S \), we group the structures of BPHs by \( S \), and the result is shown in Fig. 1a and table S1. From the summarized plot, it is evident that the number of BPHs with open-shell characters of \( S \geq \frac{1}{2} \) increases rapidly with the increasing number of the hexagonal rings \( h \). Among them, BPHs with high-spin ground states are of particular interests.\(^{7, 8, 10} \) Our statistics suggest two smallest BPHs with \( S = 1 \) when \( h = 6 \) and three smallest BPHs with \( S = 3/2 \) when \( h = 9 \) (see structures in Fig. 1b). Interestingly, with the increase of \( h \), the proportion of the BPHs with \( S \geq \frac{1}{2} \) rises from 31.8% for \( h = 5 \) to 50.5% for \( h = 9 \). Given that a plethora of open-shell singlet BPHs are not counted (vide infra), open-shell BPHs account for a major diversity despite that they are less stable in reality. This finding marks BPH as a rich platform for exploring \( \pi \)-magnetism. In Fig. 1b, we show the structures of a few BPHs which are classified by their \( S \). A full list of the structures of open-shell BPHs with \( h \leq 9 \) is attached in the Supporting section 1.

Machine learning has been widely applied in various fields of materials science and technology in recent years.\(^{32, 33} \) The combination of big data and artificial intelligence has led to remarkable progress in the development of new materials and obtained deeper understanding of their structure-property relationships.\(^{32, 42} \) In this work, the abundant structures of the BPHs, over 6000, constitute a valuable molecular database which could benefit from the machine learning technique. Here, we apply an interpretable strategy combining a machine learning algorithm with the SHAP approach to explore the potential relationship between structural features and open-shell properties of BPHs.

Figure 2. The relationship between structural features and magnetic properties studied by machine learning model. (a) The SHAP summary plot showing the impacts of structural features on the open-shell characters of BPHs. The positive values represent the open-shell character and the negative values represent the closed-shell character. Red/blue points indicate high/low feature values. (b) The impact of feature 1, \( N_4 - N_6 \), on the output SHAP value. (c) The impact of feature 4, the number of the internal vertex \( N_i \), on the output SHAP value. Red and green backgrounds mark \( N_i \) being odd and even. (d) The structure of the triangulene molecule (i) having four internal vertices.
Prior to the statistical analysis by machine learning models, it is necessary to define the feature vectors of samples as inputs to the machine learning model. We have therefore determined seven structural features, whose combination is unique to each BPH while capturing their topological characteristics. The descriptions of each feature are given in Fig. S1. The dataset includes a total of 6392 structures with 3217 multiplet characters (S ≥ ½) and 3175 singlet characters (S = 0). The samples are divided into 80% as the training set and 20% as the testing set. Four different classification algorithms of machine learning (Random Forest, MLP Classifier, Logistic Regression, Support Vector Machines) as implemented in the SCIKIT-LEARN python library are tested and all of them achieve excellent performances (results in Fig. S3).

The SHAP analysis is then applied to identify the key factors leading to the open-shell characters. As shown in the SHAP dot diagram in Fig. 2a, features 1 and 4 have the highest influences on the open-shell properties. The feature 1, which is the number of the imbalanced sublattice atoms, reflects Lieb’s theorem. Open-shell characters emerge as long as feature 1 is not zero (Fig. 2b). Surprisingly, the feature 4, which corresponds to the number of the internal vertex of BPH graphs (N_{iv}, example in Fig. 2d), not only correlates with the open-shell characters but also implies the multiplicity of BPHs. The SHAP analysis reveals that a zero number of N_{iv} leads to a singlet, an odd number of N_{iv} predicts a multiplet and an even number predicts either a singlet or multiplet (Fig. 2c).

Figure 3. Illustration of the triangle counting rule (TCR). (a) Left: the numbers of A/B sublattices, internal vertex (N_{iv}) and phenalenyl triangle in the phenalenyl molecule (2). Every phenalenyl subgraph in the BPH graph can be represented by a triangle and corresponds to one internal vertex. Right: the TB and MFH computed energy levels of 2. An on-site Coulomb energy (U) of 3.0 eV is added in the MFH model and results in a spin-split gap Δ_s. (b) Two BPHs with multiple internal vertices and phenalenyl triangles. Two kinds of triangles with different orientations can be identified, which determines the magnetic ground state of the BPHs. (c) Two BPHs in which the phenalenyl triangles are overlapped with each other. The magnetic ground state can also be determined. (d) Counting the phenalenyl triangles in high-order molecular triangles. In (b-d), the MFH calculated spin density distributions of the corresponding BPH structures are shown on the right. The size of the spheres indicates the magnitude of the spin density.

The finding of the relationship of the internal vertex to the open-shell character is remarkably inspiring. It implies that feature 4 is intrinsically linked to the magnetic ground state of BPHs. A close inspection of the topological graph indicates that every internal vertex corresponds to one phenalenyl subgraph as shown in Fig. 3. Thus, the number of feature 4 (N_{iv}) represents the number of the phenalenyl moiety within the PBH structures. Phenalenyl or phenalenyl radical, as the smallest open-shell BPH, has a doublet ground state with S = ½. The TB calculated energy spectrum exhibits one zero-energy state which leads to the spin-split gap after considering the Coulomb energy U in the MFH model (Fig. 3a). For larger BPHs, two kinds of phenalenyl triangles with different orientations can be discernible as indicated by orange and cyan triangles in Fig. 3b and 3c. The total spin number S can be determined by half of the difference between the numbers of two kinds of triangles. For instance, two orange triangles can be identified from uthrene (3) leading to S = 1, whereas molecule 4 with two orange and one cyan triangles has S = 1/2 (Fig. 3b). The same rule applies to molecules in which overlapped phenalenyl triangle can be derived (two examples in Fig. 3c). An interesting case is the triangulene molecule, which can be represented by a large triangle. As displayed in Fig. 3d, three orange and one cyan phenalenyl triangle are found in [3]-triangulene structure (1) leading to S = 1, while
six orange and three cyan phenalenyl triangle are found in [4]-triangulene structure (9) leading to $S = \frac{3}{2}$.

Figure 4. Applying the triangle counting rule to BPHs with high spins. (a) The effect of adding triangles on the magnetic properties—adding hexagons one by one from [3]-triangulene (i). (i – iv) The newly added triangles are coupled with the original triangles, and the spin densities are distributed over the remaining triangles. (b) from (i) to (iv), the effect of adding hexagons without the inclusion of more triangles on the magnetic properties. We start with the structure of uthrene (3). Top row: the structures of BPHs. Middle row: MFH computed spin density distributions. Bottom row: MFH computed energy spectra of the corresponding structures.

This counting rule is particularly simple and straightforward as compared to the widely used Lieb’s theorem which counts all the vertices from the two interpenetrating sublattices, since it only needs to count the number of the internal vertex. In addition, the associated phenalenyl triangle can be used to predict the distributions of spin densities in open-shell BPHs. We start with the popular [3]-triangulene molecule (1). By sequential adding hexagons in the structure graph of 1, we show how its magnetic properties evolve. As indicated in Fig. 4a, adding a hexagon shown in red on the edge of 1 produces an additional internal vertex and therefore a phenalenyl triangle in cyan (8). Since this triangle is different from the major phenalenyl triangle in 1, the $S$ is determined to be $\frac{1}{2}$. This additional phenalenyl triangle quenches the spin density distributed around the region where two triangles overlap (compare the spin density distributions of 1 and 8 in Fig. 4a). Further attaching a hexagon in 8 yields two internal vertices and phenalenyl triangles. Since the two triangles have opposite orientations, the $S$ of 9 remains to be $\frac{1}{2}$. While, the overlapping between the triangles reduces the weights of spin densities around the overlapped regions (Fig. 4a-iii). Fig. 4b shows another example of molecules in which the introduction of additional hexagons does not produce extra internal vertices and phenalenyl triangles. In such case, the open-shell ground states of PBHs will not be affected as clearly reflected from the MFH computed spin density distributions and energy spectra.

Besides the magnetic properties originating from the zero-energy state, some of the singlet BPHs with $S = 0$ exhibit an open-shell ground state due to the spin-polarization of low-energy states or competitions between the aromatic stabilization and penalty of forming radicals.43 However, the predictions of the open-shell singlets have been more challenging than the multiplet species. One of the most representative classes of BPHs with ground states of open-shell singlet is zethrene and its derivatives. They have gained intensive research interest for a long time as diradicaloids. Previous experimental44-46 and computational works47, 48 reveal that zethrenes develop an increasing open-shell character through lateral and longitudinal $\pi$-extensions.

Figure 5. Applying the triangle counting rule to BPHs of open-shell singlet. (a) The structure of zethrene (14) consisting of two different phenalenyl triangles. Its absolute magnetization and the diradical character $y_0$ are zero. (b)-(d) Structures of zethrene derived BPHs. Their spin density maps, absolute magnetizations and the diradical characters are also calculated. (e) Summary of the absolute spin magnetization of
the zethrene-derived BPHs. \( n \) denotes the number of connecting hexagons. (f) Summary of the diradical character \( y_0 \) with the increasing number of connecting hexagons.

As shown in Fig. 5, the smallest zethrene molecule 14 consists of two phenalenyl triangles of opposite orientation, which yields \( S = 0 \). The magnetic ground state of zethrene molecules can be evaluated by the diradical character \( y_0 \), which is defined as \( y_0 = 1 - \frac{2T_i}{1 + T_i} \), where \( T_i \) is the orbital overlap between the corresponding orbital pairs. The diradical character takes a value between 0 and 1, which correspond to absolute closed-shell and diradical states, respectively. Compared to the results obtained by the DFT method (14: \( y_0 = 0.035; 15: y_0 = 0.035; 17: y_0 = 0.37; 19: y_0 = 0.5 \)), the present values calculated by constructing a density matrix from the TB and MFH models are relatively smaller (14: \( y_0 = 0.035; 15: y_0 = 0.37; 17: y_0 = 0.37; 19: y_0 = 0.135 \). This difference comes from the choice of the Coulomb energy \( U \) in the MFH model.

Here, 14 is calculated to be a purely closed-shell molecule with \( y_0 = 0 \). By elongating zethrene with central bridging hexagons as molecules 15 to 20, the number of the phenalenyl triangles and their corresponding orientations remain unchanged, preserving the total spin quantum number \( S = 0 \). Meanwhile, zethrenes develop increasing open-shell characters with the elongation of the central bridge (Fig. 5b-5d) as manifested by the increased values of \( y_0 \) and the absolute magnetization from the MFH model. We have computed different zethrene derivatives with up to 3 central hexagons (structures shown in Fig. 5g), and the results are summarized in Figs. 5e and 5f. We find that the threshold number of the bridging hexagons to give rise to open-shell singlet is 2. In addition, the shape of the bridging hexagons influences the diradical character \( y_0 \), i.e., the bridging hexagons with zigzag-type edges (17, 19) yield lower values of \( y_0 \) than with the armchair-type edges (18, 20). This can be interpreted by the tendency of spin delocalization along the zigzag edges.22–33 These findings are useful for the design of open-shell BPHs with open-shell singlet ground state.

Figure 6. BPHs with topological frustrations. (a-i) Left: the structure of Clar’s goblet (21). Right: the spin density distribution of 21. (a-ii) Left: the TB and MFH computed energy spectra of 21. Right: MFH computed energy spectrum of 21 in its excited triplet state. (a-iii) MFH-calculated wave functions of two SOMOs of its open-shell singlet. (b), (c) Two additional BPHs with topological frustration which are also constructed by 11 hexagons. All the three molecules have relatively small excitation energies and spin density distributions. Red/blue isosurfaces denote opposite signs of the spin direction or wave function.

There is another kind of BPH with the magnetic ground state of open-shell singlet. Their magnetic properties derive from the topological frustration, meaning that kekulé structures cannot be drawn despite the absence of sublattice imbalance.50–54 BPHs with topological frustrations are subject to any value of \( U \) and exhibit the open-shell ground state. Fig. 6a shows the structure of Clar’s goblet (21) along with its spin density map. Although the total spin of the molecule \( S = 0 \), the system’s nullity amounts to \( \eta = 2 \). This is corroborated with TB calculations, which predicts the presence of two zero-energy states (Fig. 6a-ii). The zero-energy states become spin-polarized as Coulomb interactions are included with the Hubbard model (Fig. 6a-iii). By drawing the phenalenyl triangles, we find that such topologically
frustrated BPHs could be considered as two BPHs of opposite spin moments connected by a phenyl bridge with no spin density distributions. The diradical characters \( y_0 \) of this kind of BPH are all 1, and there is no overlap between the wave functions of its SOMOs orbitals (Fig. 6a–iii), which further suggests the pure diradical property. For the topological frustrated BPHs, it requires at least eleven hexagons (six for the two phenalenyl triangles and five for the perylene bridge) to build the open-shell singlet BPHs. This is consistent with the fact that in our BPH database with \( h \) less than 10 we could not find any singlet molecules with zero-energy states. Based on this design rule, we conceive seven molecules of topological frustration comprising 11 hexagons (Fig. 6a–iii). The TB and MFH calculations confirm that all of them have two zero energy states resulting in two singly occupied molecular orbitals with antiparallel aligned electrons. It is worth mentioning that the spin excitation energies of these molecules from the ground state of open-shell singlet to the excited state of open-shell triplet are relatively small (around a few meV) by MFH calculations with a \( U = 3 \) eV. This observation implies that their magnetic properties are likely to be experimentally identified by the state-of-the-art scanning probe microscopy techniques.11

The discussions about the two kinds of nanographene edges, the armchair and the zigzag edges, have been scrutinized in numerous works.55–57 PBHs of zigzag-edge tend to host low-energy states, which shift towards the Fermi level as their size increases.56–57 Coulomb repulsion triggers spin polarizations and induces magnetism in zigzag-PBHs. One of the most representative cases is acenes which are linearly fused benzenes (Fig. S2). In this work, we also notice the importance of the two different edges in the BPHs magnetism. We find that the zigzag edge tends to act as a magnetically interactive bridge, in contrast to the armchair edge which tends to act as a magnetically isolating bridge. For the zethrenes 17 and 19 (structures in Fig. 5) which are connected by the bridging hexagons of zigzag edge, the spin densities with opposite orientations will extend along the zigzag edge. This behavior leads to a weakening of the spin polarization, the relatively low diradical character \( y_0 \) and low absolute spin magnetization. In contrast, for their counterparts 18 and 20 connected by the bridge hexagons of armchair edge, the spin densities with opposite orientations will be isolated in their respective phenalenyl triangles. This leads to a strengthening of spin polarization, and in turn leads to a relatively high diradical character \( y_0 \) and high absolute spin magnetization (Fig. 5c,5d). For BPHs with topological frustrations, the key to the formation of topological frustrations also lies in the construction of ‘bridges’ with armchair edges (highlighted in Fig. 6) to isolate the opposite spin densities in different parts. These observations further demonstrate the roles of two different edges in the construction of magnetic PBHs, and provide implications for the designing rule.

Figure 7. Open-shell BPHs with globally delocalized spin densities. (a-i) Left: The structure of BPH (24) in which two different triangles are partially overlapped with each other. Right: the spin densities are uniformly distributed over the molecule. (a-ii) The MFH computed energy spectrum of the molecule. The spin-split gap is relatively small due to the delocalization of the spin densities. (a-iii) MFH-calculated wave function of SOMO. Red/blue isosurfaces denote opposite signs of the wave function. (b),(c) Two additional BPHs with globally delocalized spin densities having \( S = 1/2 \) and \( S = 1 \), respectively. Their energy spectra and wave functions of SOMOs are shown.
Throughout our PBH database, we discover a class of BPHs with intriguing magnetic properties, whose spin densities are not localized on specific edges or segments, but evenly distributed over the entire molecule (molecules 24-26 in Fig. 7 and others in Fig. 9). The nature of the delocalized spin densities tends to offer relatively small spin-split gaps compared to molecules of similar size. Moreover, the molecules are likely to be chemically more stable since unpaired electrons in these molecules tend to delocalized over the molecule.

We first inspect the smallest molecule 24 within this class of BPHs whose topology is displayed in Fig. 7a. Three internal vertices and accordingly three phenalenyl triangles overlapped on each other are found in 24, leading to a $S = \frac{1}{2}$ magnetic ground state. The spin densities are evenly distributed on one set of the sublattice of 24 as shown in Fig. 7a, this is consistent with the MFH-calculated wave function of the SOMO orbital (Fig. 7a-iii). We find a spin-split Coulomb gap of only 0.401 eV (Fig. 7a-ii) determined by the MFH calculation. For comparison, the Coulomb gaps for 8 and olympicene which have the same magnetic ground states, are 0.555 eV and 0.565 eV, respectively. Applying the same design route, i.e., the constituting phenalenyl triangles being partially overlapping, we conceive two other molecules with $S = 1/2$ and 1 (Fig. 7b and 7c), one with a C$_{3h}$ symmetry and the other with a D$_{3h}$ symmetry. Both of them exhibit a uniform spin density distribution over the entire molecule (Fig. 7b-iii and 7c-iii). For larger structures having more hexagons, the same strategy also works (Fig. 5g). Given that open-shell BPHs are usually highly unstable due to the unpaired electrons, it is reasonable to believe that such a class of molecules with a uniform spin density distribution will survive for a longer time, which is advantageous for the applications of open-shell BPHs.

CONCLUSION
To conclude, we have built up a structure database of BPHs with up to 9 fusing hexagons by graph enumeration. We screened the open-shell BPHs by computing the TB eigenvalues of the BPH graphs. We have disclosed a significant number of open-shell BPHs whose structures are listed in Supporting section 1, and are surprising to find that open-shell BPHs outnumber the closed-shell ones which are known to be more stable. Given that only a small number of open-shell BPHs have been realized in reality, this collection of open-shell BPHs grouped by the number of constituting hexagons and the total spin number S may serve to guide the experimental explorations of BPHs with desired magnetic properties.

Moreover, the establishment of the database allows for a data-driven analysis through the ML model. Different structural features were extracted from the database as inputs into the ML model, and we scored their importance on the magnetic properties through an interpretable SHAP method. We identified that in addition to the well-established magnetic fingerprint of sublattice imbalance, which underlies the Lieb's theorem, the number of internal vertices of the BPH graph is strongly correlated with their open-shell properties. We further demonstrate the role of the phenalenyl subgraph in predicting the magnetic ground state, which we called the triangle counting rule. This rule is particularly simple but more useful as compared to the widely used Lieb's theorem: (1) Only the number of the internal vertex needs to be counted. (2) The spin number S of the magnetic ground state can be determined by the different numbers of two differently oriented triangle subgraphs within the BPHs. (3) It provides hints to the open-shell properties of those BPHs with $S = 0$ while having the phenalenyl subgraphs like zethrenes, which are fundamentally more challenging for theoretical predictions. Finally, we have conceived a type of new open-shell BPHs whose spin densities are globally distributed over the molecules, which may benefit the chemical stability of these reactive species. Therefore, we conclude that our results extend the well-known Lieb's theorem and Ovchinnikov's rule and provide a simple way for designing open-shell carbon nanostructures. With the recent advances on emerging synthetic routes,88-90 we foresee the experimental realizations of the predicted structures and their ultimate applications in further quantum technologies.

ASSOCIATED CONTENT
The Supporting Information is available free of charge on the ACS Publications website at DOI:

AUTHOR INFORMATION
Corresponding Authors
*qiangsun@shu.edu.cn

Author Contributions
Y.Y. and Q.S. wrote the paper with contributions from all co-authors. All authors discussed the results and commented on the manuscript at all stages.

Notes
The authors declare no competing financial interests.

ACKNOWLEDGMENT
This work was supported by National Natural Science Foundation of China (No. 22072086).

REFERENCES
(1) Zhao, G.; Li, X.; Huang, M.; Zhen, Z.; Zhong, Y.; Chen, Q.; Zhao, X.; He, Y.; Hu, R.; Yang, T.; Zhang, R.; Li, C.; Kong, J.; Xu, J.-B.; Ruoff, R. S.; Zhu, H., The physics and chemistry of graphene-on-surfaces. Chem. Soc. Rev. 2017, 46 (15), 4417-4449.
(2) Wu, J.; Pisula, W.; Müllen, K., Graphenes as Potential Material for Electronics. Chem. Rev. 2007, 107 (3), 718-747.
(3) Nakahara, M.; Ota, Y.; Rahimi, R.; Kondo, Y.; Tada-Umezaki, M., Molecular Realizations of Quantum Computing 2007. WORLD SCIENTIFIC: 2009; Vol. Volume 2, p 284.
(4) Rahimi, R.; Sato, K.; Furukawa, K. O. U.; Toyota, K.; Shimoo, D.; Nakamura, T.; Kitagawa, M.; Takui, T., PULSED ENDOR-BASED QUANTUM INFORMATION PROCESSING. International Journal of Quantum Information 2005, 03 (suppo), 197-204.
(5) Tombros, N.; Jozsa, C.; Popinciuc, M.; Jonkman, H. T.; van Wees, B. J., Electronic spin transport and spin precession in single graphene layers at room temperature. Nature 2007, 448 (7153), 571-574.
(6) Pavlice, N.; Mistry, A.; Majzik, Z.; Moll, N.; Meyer, G.; Fox, D. J.; Gross, L., Synthesis and characterization of trianguene. Nat Nanotechnol 2017, 12 (4), 308-312.
(7) Su, J.; Telychko, M.; Hu, P.; Macam, G.; Mutombo, P.; Zhang, H.; Bao, Y.; Cheng, F.; Huang, Z.-Q.; Qiu, Z.; Tan, S. J. R.; Lin, H.; Jelinek, P.; Chuan, F.-C.; Wu, J.; Lu, J., Atomically precise bottom-up synthesis of &shREW ;extended [5]trianguene. 2019, 5 (7), eaaV7717.
(48) Marković, S.; Radenković, S.; Marković, Z.; Gutman, I., DFT study on singlet diradical character of zethrenes. *Russian Journal of Physical Chemistry A* **2011**, *85* (13), 2368-2372.

(49) Nakano, M.; Kishi, R.; Nakagawa, N.; Ohta, S.; Takahashi, H.; Furukawa, S.-i.; Kamada, K.; Ohta, K.; Champagne, B.; Botek, E.; Yamada, S.; Yamaguchi, K., Second Hyperpolarizabilities ($\gamma$) of Bismidazole and Bistriazole Benzenes: Diradical Character, Charged State, and Spin State Dependences. *The Journal of Physical Chemistry A* **2006**, *110* (12), 4238-4243.

(50) Zeng, W.; Sun, Z.; Heng, T. S.; Gonçalves, T. P.; Gopalakrishna, T. Y.; Huang, K.-W.; Ding, J.; Wu, J., Superheptazethrene. *Angew. Chem. Int. Ed.* **2016**, *55* (30), 8615-8619.

(51) Eimre, K., eimrek/tb-mean-field-hubbard: v2.1.0. Zenodo **2022**.

(52) Yang, L.; Park, C.-H.; Son, Y.-W.; Cohen, M. L.; Louie, S. G., Quasiparticle Energies and Band Gaps in Graphene Nanoribbons. *Phys. Rev. Lett.* **2007**, *99* (18), 186801.

(53) Magda, G. Z.; Jin, X.; Hagymási, I.; Vancsó, P.; Osváth, Z.; Nemes-Incze, P.; Hwang, C.; Biró, L. P.; Tapasztó, L., Room-temperature magnetic order on zigzag edges of narrow graphene nanoribbons. *Nature* **2014**, *514* (7524), 608-611.

(54) Wang, W. L.; Yazyev, O. V.; Meng, S.; Kaxiras, E., Topological Frustration in Graphene Nanoflakes: Magnetic Order and Spin Logic Devices. *Phys. Rev. Lett.* **2009**, *102* (15), 157201.

(55) Enoki, T.; Kobayashi, Y.; Fukui, K.-I., Electronic structures of graphene edges and nanographene. *Int. Rev. Phys. Chem.* **2007**, *26* (4), 609-645.

(56) Nakada, K.; Fujita, M.; Dresselhaus, G.; Dresselhaus, M. S., Edge state in graphene ribbons: Nanometer size effect and edge shape dependence. *Physical Review B* **1996**, *54* (24), 17954-17961.

(57) Fernández-Rossier, J.; Palacios, J. J., Magnetism in Graphene Nanoislands. *Phys. Rev. Lett.* **2007**, *99* (17), 177204.

(58) Grill, L.; Hecht, S., Covalent on-surface polymerization. *Nature Chemistry* **2020**, *12* (2), 115-130.

(59) Lu, J.; Jiang, H.; Yan, Y.; Zhu, Z.; Zheng, F.; Sun, Q., High-Throughput Preparation of Supramolecular Nanostructures on Metal Surfaces. *ACS Nano* **2022**.

(60) Sun, Q.; Jiang, H.; Yan, Y.; Fasel, R.; Ruffieux, P., Molecular heterostructure by fusing graphene nanoribbons of different lengths through a pentagon ring junction. *Nano Research* **2022**.