The Journal of Chemical Physics ARTICLE pubs.aip.org/aip/jcp

A graph-based statistical model for carbon nanostructures

Cite as: J. Chem. Phys. 162, 154104 (2025); doi: 10.1063/5.0244219
Submitted: 17 October 2024 • Accepted: 24 March 2025 •
Published Online: 15 April 2025

Chang-Chun He,[1] Shao-Gang Xu,[2,3] Jiarui Zeng,[1] Weijie Huang,[1] Yao Yao,[1,4] Yu-Jun Zhao,[1] Hu Xu,[2,3,a)] and Xiao-Bao Yang[1,b)]

AFFILIATIONS

[1] Schol of Physics and Optoelectronics, South China University of Technology, Guangzhou 510640, China
[2] Department of Physics, Southern University of Science and Technology, Shenzhen 518055, China
[3] Quantum Science Center of Guangdong-Hong Kong-Macao Greater Bay Area (Guangdong), Shenzhen 518045, People's Republic of China
[4] State Key Laboratory of Luminescent Materials and Devices, South China University of Technology, Guangzhou 510640, China

[a)] **Electronic mail:** xuh@sustech.edu.cn
[b)] **Author to whom correspondence should be addressed:** scxbyang@scut.edu.cn

ABSTRACT

Energy degeneracy in physical systems may be induced by symmetries of the Hamiltonian, and the resonance of degeneracy states in carbon nanostructures can effectively enhance the stability of the system. Combining the octet rule, we introduce a statistical model to determine the physical properties by lifting the energy degeneracy in carbon nanostructures. This model offers a direct path to accurately ascertain electron density distributions in quantum systems, akin to how charge density is used in density functional theory to deduce system properties. Our methodology diverges from traditional quantum mechanics, focusing instead on this unique statistical model by maximizing bonding entropy to determine the fundamental properties of materials. Applied to carbon nanoclusters and graphynes, our model not only precisely predicts bonding energies and electron density without relying on external parameters but also enhances the prediction of electronic structures through bond occupancy numbers, which act as effective hopping integrals. This innovation offers insights into the structural properties and quantum behavior of electrons across various dimensions.

I. INTRODUCTION

Energy degeneracy resulting from latent symmetry is a crucial aspect of quantum systems, contributing to various quantum phenomena.[1–3] The latent symmetry is attributed to the symmetry of the isospectral effective Hamiltonian obtained through subsystem partitioning, and the rotational symmetries can be broken in a controlled manner while maintaining the more fundamental symmetry.[4] Lifting the degeneracy by external perturbations can lead to a lower energy state with energy level splitting. For instance, the Jahn–Teller effect describes how local geometric distortions can lower the overall energy of a system by removing degeneracy.[5] In carbon nanostructures, quantum resonance provides another mechanism for lifting energy degeneracy by the combination of possible degenerate Kekulé structures, emerging delocalized electron states, and aromaticity to further stabilize the system.[6–8] For the benzene molecule, there are two Kekulé structures with C_3 symmetry that both meet the octet rule, exhibiting the twofold energy degeneracy. Resonance theory shows that the uniform superposition of Kekulé structures with C_6 symmetry corresponds to the true electron density of benzene,[9] lifting the energy degeneracy with additional resonance energy.

Generally, a polycyclic aromatic hydrocarbon (PAH) molecule has a series of Kekulé structures that satisfy the octet rule in a degenerate energy state. To construct the most valuable one to represent all Kekulé structures, Fries rule is a simple and intuitive tool used to predict the most dominant Kekulé structures of PAHs,[10–13] which is valuable for understanding their stability, reactivity, and aromaticity. Clar's rule identifies the key structure by maximizing the number of nonadjacent π-sextets.[14] The nonadjacent π-sextets have two forms with alternating single and double bonds, independently contributing a statistical weight of 2, and the statistical weight is 2^N for N π-sextets. Therefore, Clar's rule provides an avenue to lift energy degeneracy by favoring the structure

with the largest statistical weight, which best approximates the relative stability of various carbon nanostructures.[15–18] The sextets predicted by Clar's rule generally correspond to aromatic rings, aligning with various theoretical methods,[19,20] although exceptions exist in irregularly shaped PAHs.[21] In addition, PAH isomers with more sextets tend to exhibit higher kinetic stability, lower reactivity, and larger energy gaps.[22] Despite the success in interpreting chemical bonding, the empirical rules limit quantitative electron distribution analysis for lack of a rigorous mathematical foundation in quantum theory.[23–25] A single Clar structure often cannot accurately reproduce the actual electron density distribution of a PAH molecule, as the symmetry of the Clar structure often disagrees with the realistic molecular symmetry.[26] Semi-empirical methods provide a cost-effective approach for calculating explicit, anharmonic vibrational frequencies in large molecules such as PAHs.[27] However, their default parameters often lack sufficient accuracy for experimental comparison, particularly in the hydride stretching region, which is crucial for observations. The resonance theory describes chemical bonding by combining several key contributing structures with specific weights into a resonance hybrid,[28] where the weights are derived using *ab initio* methods derived from valence bond theory[29] or from the natural bond orbital approaches.[30] To best approximate the actual wave function, a superposition of all possible Clar resonators is constructed, where the weights of the resonators are determined through a projection-weighted symmetric orthogonalization scheme from the density functional theory (DFT) wave function.[23]

Clar's rule, as a model, falls short in accurately predicting the electron density along with structural stabilities and electronic properties of carbon nanostructures. Current quantitative models require training data from DFT to acquire model parameters and precisely predict the properties of unknown structures, which belong to a first-principles-based (second-principles) scheme.[31] To address electronic properties, the tight-binding (TB) model offers a simpler method to construct an effective Hamiltonian,[32] but it relies on DFT or experimental data for integral parameters.[33] Several deep learning methods have been proposed to predict structural properties after fitting millions of parameters based on massive data from DFT.[34–36] The inherent chemical origin and deeper insights into specific systems are obscured by the intricate neural network model.[37] Therefore, there is a need for a more direct method that can both conceptualize and quantify the electron density distribution, providing a deeper understanding of chemical bonding in carbon nanostructures.

The octet rule is a chemical rule of thumb that reflects the theory that main-group elements tend to bond in such a way that each atom has eight electrons in its valence shell, giving it the same electronic configuration as a noble gas.[38] The carbon atoms will bond to the nearest atoms rather than the next-nearest neighboring atoms, and the interactions between the nearest neighboring atoms will be the strongest. Generally, for stable PAH molecules, all carbon atoms should obey the octet rule. Although there are 175 linearly independent resonance Lewis structures in benzene,[39] only two Kekulé structures satisfy the octet rule. Therefore, we only consider the most important resonance structures, using this approximation to simplify the otherwise tedious task of determining all the resonance structures. Many ionic resonance structures represent bonding configurations where not all carbon atoms conform to the octet rule. In this work, we do not consider these structures and only aim to understand the PAHs that meet the octet rule. In addition, a modified octet rule considering unpaired electrons is required to understand those molecules that do not obey the octet rule, which will be focused on in our future work.

In this work, we introduce a streamlined bonding entropy model (BEM) for determining electron density distribution by lifting the energy level in carbon nanostructures. With the combination of the octet rule and maximal entropy principle, the predicted structural properties of this model remarkably align with both experimental findings and first-principles calculations. By integrating the octet rule with the BEM, we can accurately evaluate bonding energies, uncovering significant contributions of bonding entropy to structural stability in carbon nanostructures. Furthermore, the occupancy numbers (ONs) of each C–C bond derived from our model predict the electron density and bond length. These ONs can serve as hopping integrals in the TB models, thereby simplifying accurate predictions of electronic structures, including energy gaps, energy levels, and the corresponding spatial distribution of wave functions. Local aromaticity is defined by the mean value of ONs in each six-membered ring, offering a simple yet effective method to measure electron delocalization in the π-conjugated system. Our method provides valuable insights into the structural properties and quantum behavior of electrons across different dimensional carbon nanostructures.

II. TERMINOLOGY DEFINITION

Below, we list the definitions of frequently referred to terms in the work.

Terminology	Definition
Bonding entropy S_b	$-\sum_i p_i \log(p_i)$, where p_i is the probability of electrons occupying bond i
Occupancy numbers (ONs)	n_i is the electron numbers of C–C bonds, ON is $n_i/2$

III. RESULTS AND DISCUSSION

A. The derivation of bonding entropy model

To understand the interatomic interactions in carbon nanostructures, we start with the simple benzene molecule, which has two different forms of alternating single and double bonds that obey the octet rule, as shown in the top panel of Fig. 1(a). The degenerated Kekulé structures hold three-fold symmetry, where the fictional C–C double bond is stronger than the C–C single bond, violating the identical C–C bonds in the benzene molecule. To lift the two-fold degenerate energy level, the equally weighted superposition of two Kekulé structures can satisfy the intrinsic symmetry as depicted in the middle panel of Fig. 1(a), which has the lowest energy due to the additional resonance energy. From the view of molecular theory, the optimal state ψ_1 is the resonance of two equal Kekulé structures; thus, the resonance weight is $1/\sqrt{2}$ to ensure that the ONs of all C–C bonds are identical to 1.5, in agreement

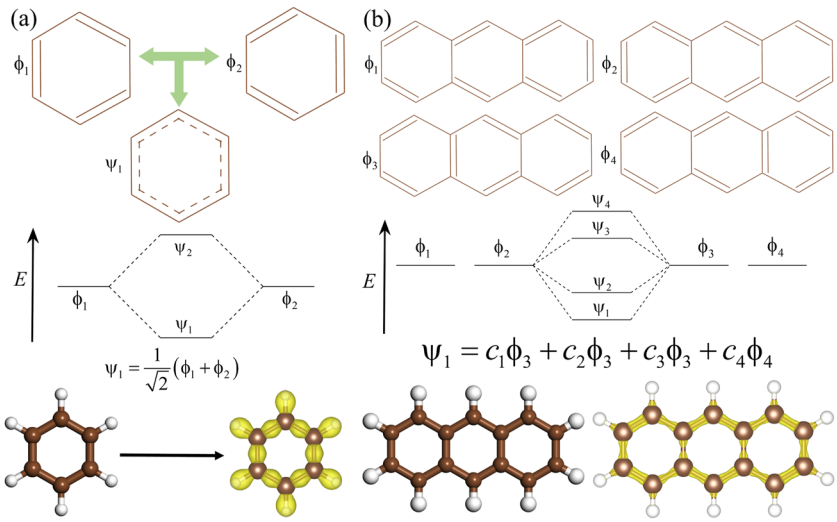

FIG. 1. (a) Top part is two Kekulé structures (denoted by ϕ_1, ϕ_2) and the average structure of benzene (denoted by ψ_1), middle part is the energy level diagram of two degenerate states, ϕ_1, ϕ_2, and bottom part is the structure and charge distribution of benzene calculated by DFT. (b) Top part is four Kekulé structures of anthracene (denoted by ϕ_1–ϕ_4), middle part is the energy level diagram of four degenerate states ϕ_1–ϕ_4 with the undetermined parameters, and bottom part is the structure and charge density of anthracene calculated by DFT.

with the charge density distribution derived from DFT, as shown in the bottom part of Fig. 1(a). For any carbon nanostructures, the ground state can be represented by the linear combinations of a series of degenerate Kekulé structures,[40] but the resonance weights cannot be directly derived unless from first-principles calculations. For instance, anthracene has four Kekulé structures [see Fig. 1(b)]. Particularly, the charge density from DFT shows that the edge C–C bonds are stronger than the inner C–C bonds in the bottom part of Fig. 1(b), but we cannot determine the accurate ONs of C–C bonds, while the proposed BEM can easily calculate the resonance weights with external parameters.

For a given carbon nanostructure, a series of possible degenerate Kekulé structures satisfy the octet rule. We can assume that all electrons are delocalized across all chemical bonds, and the number of electrons in each bond is variable. In this framework, the total N_{total} electrons in the system are allocated across each bond as demonstrated in Fig. 2(a), where each bond contains n_i electrons. We propose a bonding entropy model (BEM), where all electrons naturally adopt the most uniform distribution, corresponding to the maximization of Shannon entropy across the bonds. This enhances electron delocalization and maximizes stability, in turn reflecting the system's tendency to favor the most stable configuration through optimal electron distribution. Therefore, the optimal electron distribution across N_{bond} C–C bonds corresponds to the maximum bonding entropy,

$$S_b = -\sum_{i=1}^{N_{\text{bond}}} p_i \log(p_i), \quad p_i = \frac{n_i}{N_{\text{total}}}, \tag{1}$$

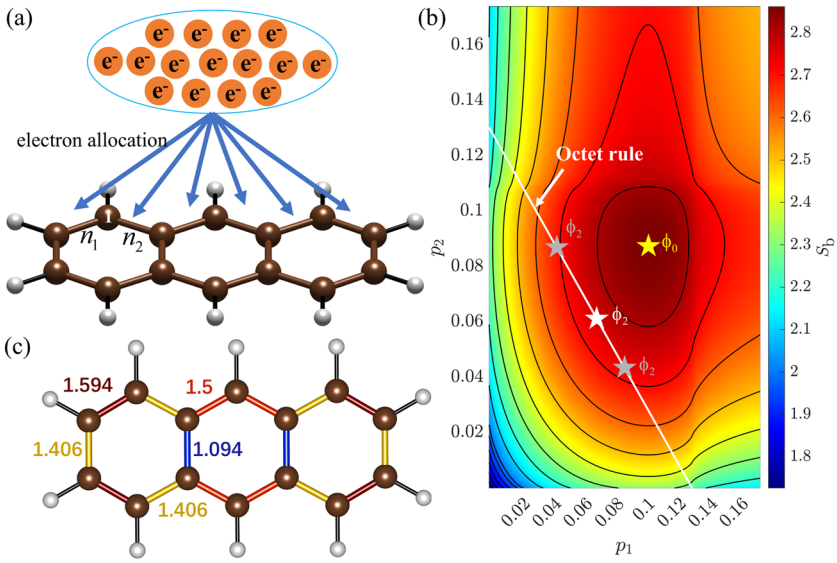

FIG. 2. (a) Total electron allocation model for anthracene molecule. (b) The bonding entropy S_b surface varied with p_1 and p_2 for the anthracene molecule. (c) The optimal occupancy number of each bond for the anthracene molecule and the ON number is marked near the bond.

which is detailed in the section "The derivation of bonding entropy model" of the supplementary material.

We can construct a series of bonding configurations in which each carbon atom obeys the octet rule. In the statistical view, the most probable bonding configuration is the one with the maximal multiplicity, corresponding to the maximal entropy. The maximum bonding entropy of S_b can decide the ground state electron density for carbon nanostructures. Here, we take anthracene as an example to show how to determine the p_i of each bond. The variation of S_b with ϕ_1 and ϕ_2 is given in Fig. 2(b), where the global maximum ϕ_0 without any constraints is marked by the yellow pentagon. ϕ_0 represents the equal electron probability for each bond, which contradicts the actual electron distribution of the molecule due to varying bonding environments for each carbon atom.

The identification of the maximal S_b is essential, particularly under local bonding constraints, ensuring each carbon atom adheres to the octet rule.[41,42] For instance, we apply the constraint $n_1 + n_2 = 6$ to carbon atom 1, taking into account the two-electron requirement of the C–H bond to satisfy the duplet rule for hydrogen. Figure 1(b) shows four Kekulé structures that satisfy the octet rule. The white line represents the area compliant with the octet rule in Fig. 1(b), and the red pentagon indicates the optimal electron density of ϕ_*, as depicted in Fig. 2(c). The S_b for ϕ_* distribution is notably larger than that of the Kekulé structures, indicating greater stability in the resonance structure compared to a single Lewis structure. Similar to bond order, the ONs of each bond, defined by $n_i/2$ for C–C bonds in two-center two-electron bonds, are indicated near each inequivalent bond. Note that the predicted ONs are in good agreement with the molecular symmetry compared with the Clar structure.[43] In addition, the ONs also agree with the bond order determined by the natural resonance theory, as detailed in the supplementary material.

For a PAH molecule, there are a series of bonding configurations (states) that obey the octet rule. Among these states, the state with the greatest number of microstates is the most probable from the statistical view. The greater number of indistinguishable microstates corresponds to the larger entropy. Therefore, the state with the maximum entropy is the optimal bonding configuration under the constraint of the full valence shell. The BEM can be interpreted through the maximization of Shannon information entropy for electrons distributed across covalent bonds. The combination of the maximal entropy principle and the octet rule can well determine the optimal electron density for PAHs. In π-conjugated systems, electrons inherently prefer a uniform distribution across all bonds while adhering to the octet rule, aligning with the principle of maximized Shannon entropy. This uniform distribution facilitates optimal electron density, thereby enhancing structural stability through effective electron delocalization.

B. Energy prediction ability of BEM

According to the BEM, the bond energies of carbon nanostructures can be expressed by S_b. Thus, the relative stability of isomers can be distinguished by the global maximum S_b. As reflected in Fig. 3(a), the first-principles calculations based on the highly precise CCSD(T)/CBS method[44] prove that triphenylene with threefold symmetry is the most stable among $C_{18}H_{12}$ molecules. Interestingly, the S_b of triphenylene is larger than that of chrysene, aligning well with the Coupled Cluster with Single, Double, and perturbative Triple excitations/Complete Basis Set (CCSD(T)/CBS) method. Thus, S_b effectively describes PAH stability without external parameters, offering a distinct perspective on isomerization energies. As shown in the inset of Fig. 3(a), the ONs in these structures reflect the symmetry of the molecules and align with the charge density distribution obtained from DFT, implying that the ON can reflect the local electron distribution. Moreover, Fig. 3(b) reveals a strong linear relationship between the ON of bonds and bond length, where shorter bond lengths corresponding to higher ON values with stronger interactions. Triphenylene, which is identified as the most stable molecule, has the lowest ON deviation from 1.5, whereas other molecules show significant deviations, contributing to their relative instability.

Beyond the stability evaluation, the resonance weights W_R (satisfying $W_R > 0$, $\sum_R W_R = 1$) of all resonance structures can be defined by the optimal resonance-type representation of ϕ_* determined by the minimal BEM,[45]

$$W_R = \arg\min \|\phi_* - \sum_R W_R \phi_R\|, \tag{2}$$

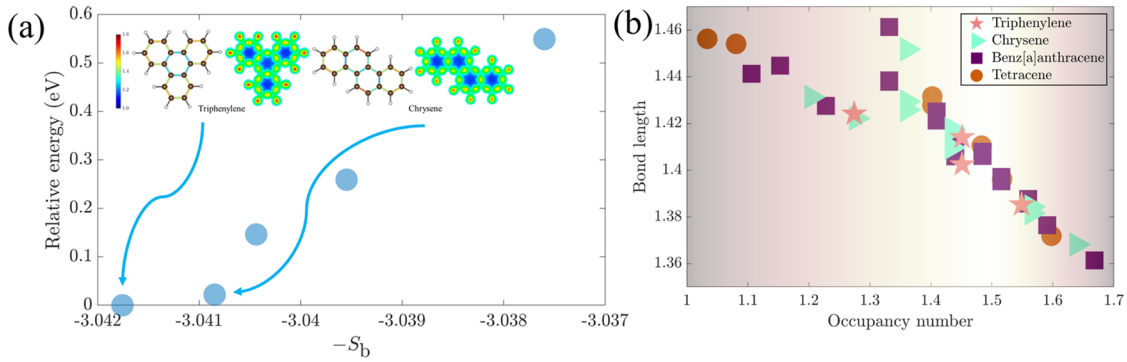

FIG. 3. (a) The linear relationship between the relative energy predicted by the DFT and $-S_b$ calculated by BEM, where the charge density and ON of two stable $C_{18}H_{12}$ molecules are plotted in the inset. (b) The relationship between the ON predicted by BEM and bond length determined by the DFT.

where ϕ_R represents the ON of resonance structures (the C–C single bond is 1, the C–C double bond is 2, and the C–C partial double bond is 1.5 like benzene), and the ϕ_* represents the ON of the structures derived by BEM. The weights for anthracene in Fig. 1(b) are $c_1 = c_4 = 40.6\%, c_2 = c_3 = 9.4\%$.

To validate the generality of the BEM, we also considered other carbon-rich clusters as shown in Fig. 4. These molecules were evaluated using DFT calculations,[46,47] in which the stability of structures with low energy is confirmed by the high-level Heyd–Scuseria–Ernzerhof (HSE06) calculation,[48,49] as detailed in the supplementary material. We take $C_{96}H_{48}$ with 24 hexagonal rings as an example to demonstrate the predictive power in cycloarenes, which mainly consist of two types of structures: kekulene and clarene.[50,51] The kekulenes and clarenes are generated by the GenInfi[52] program. Figure 4(a) clearly shows that the S_b has good linear consistency with the relative energy from DFT calculations. The most stable structure is marked in the left-top part of Fig. 4(a), in which the hexagonal rings stack with staggered alignment rather than being arranged in a single row since the ON of each bond is more uniform to gain more bonding entropy. In contrast, the structure of the highest energy in the right-bottom of Fig. 4(a) is composed of three elongated PAH clusters, losing more bonding entropy. Figure 4(b) shows that the carbon nanobelts, which are rigid and thick segments of carbon nanotubes,[53] can also be accurately predicted. The carbon nanobelts are generated by the GenCNB program.[54] Similar to cycloarenes, the most stable carbon nanobelt has the hexagonal rings in a staggered arrangement, while the most unstable nanobelt also has elongated PAH clusters. Therefore, the elongated PAH clusters will be unstable.

Next, we focus on the carbon nanostructures with pentagon rings, as shown in Fig. 4(c), in which the structure with isolated pentagons has the lowest energy as predicted by BEM and DFT calculations. The isolated pentagons can lead to a more even distribution of ONs in the carbon nanostructure. The more adjacent pentagons there are, the less stable the structure becomes, as depicted in Fig. 4(c). In addition to C–H systems, C–H–O systems can also be described by BEM, where the most stable structure occurs when the two oxygen atoms are in the *ortho*-position, as shown in Fig. 4(d). Because the C–O is a double bond, the bond distribution will be completely different with the introduction of oxygen atoms. Note that the carbon atom connected to the oxygen atom will be bonded to two other carbon atoms by single bonds; the *ortho*-oxygen atoms will result in the entire system having only three C–C single bonds, thereby reducing one single bond, which effectively increases the bonding entropy and enhances the structural stability. Therefore, the BEM can excellently predict the stability among widespread carbon nanostructures.

With a given carbon nanostructure, there are numerous combinations that meet the octet rule, and the charge density is attributed to all these combinations according to BEM. The most likely electron distribution is deduced through ergodic theory, which entails averaging over all possible distributions.[55] The maximal S_b thus determines the optimal electron distribution for a system, suggesting a uniform electron density to maximize resonance energy. Viewed as preliminary information, the number of the total electron constraint and octet rule aptly reflects the most informed current understanding, in line with the principle of maximum entropy.[56] The previous study has pointed out the connection between Shannon information

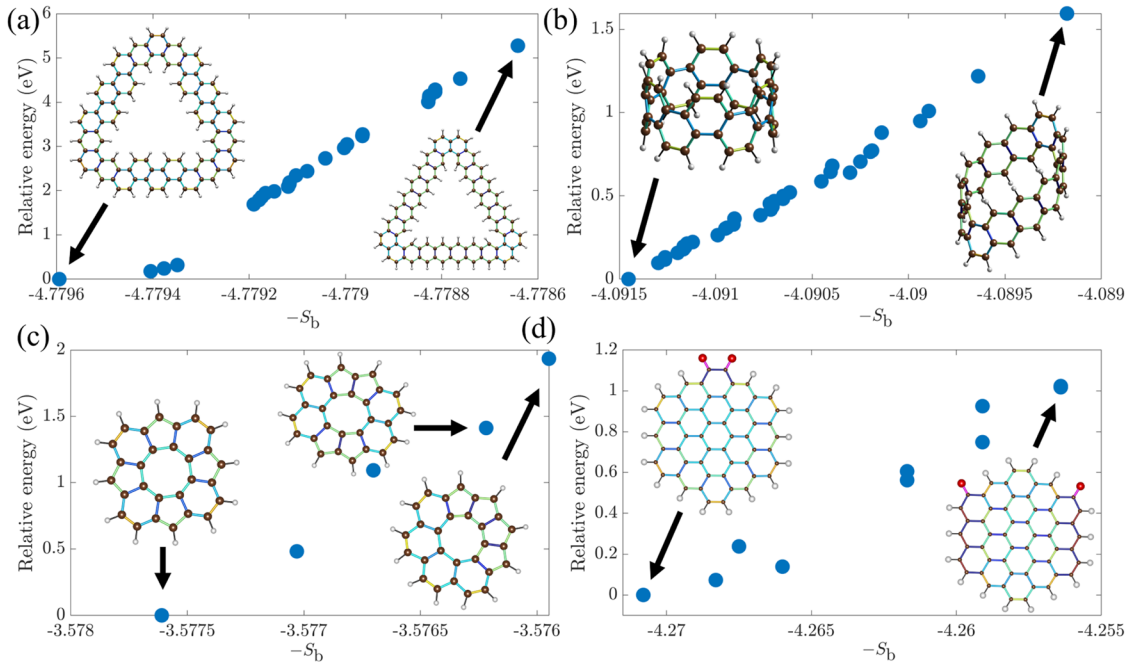

FIG. 4. Relationship between relative energy calculated by DFT and bonding entropy for (a) cycloarenes, (b) carbon nanobelts, (c) carbon nanostructures with pentagon rings, and (d) C–H–O systems.

entropy and molecular aromaticity and stability,[57–59] where the electron density is calculated by DFT. Our BEM directly determines the electron density of PAHs and emphasizes the important role of bonding entropy in the carbon nanostructures, where the stabilities are dominated by delocalized electrons.

C. Electronic structure prediction capability of BEM

Due to the outstanding optoelectronic properties and applications in organic field-effect transistors,[60–62] the electronic structures of carbon nanostructures have attracted significant attention in both experimental and theoretical domains.[63,64] The TB model is typically employed to describe the electronic structures, which often requires DFT calculations to fit integral parameters. Previous research indicates that accurately predicting energy gaps in PAHs requires up to 13 hopping integral parameters,[33] essential for differentiating various C–C bonds. In the self-consistent Hückel method, the hopping integrals are adjusted self-consistently based on bond orders and atomic charges.[65] Similarly, the Su–Schrieffer–Heeger (SSH) model demonstrates that alternating single and double bonds exhibit different hopping integral parameters.[66] Herein, we assume the hopping parameter is proportional to the ON of the bond between carbon atoms i and j, constructing the TB Hamiltonian for PAHs with nearest-neighbor couplings as follows:

$$H = -\sum_{i,j} t_{i,j}(c_i^\dagger c_j + h.c.), \tag{3}$$

where the on-site energy is set to zero, $t_{i,j}$ is the hopping integral, the subscript i,j represents the indices of the nearest neighboring carbon atoms, and $c_i^\dagger$ and c_j are the creation and annihilation operators of the π electron at sites i and j, respectively. The predicted energy gap by the TB model shows a linear correlation with the DFT-calculated energy gap in the HSE06 level, as depicted in Fig. 5(a). This relationship is expressed as $E_g^{\mathrm{DFT}} = \gamma \times E_g^{\mathrm{TB}} + E_g^0$, with parameters $\gamma = 2.355$ and $E_g^0 = 0.01$ eV. Notably, over 70% of the energy gaps predicted by the TB model deviate less than 0.1 eV from those of the DFT calculations, as shown in Fig. 5(b), using only two external parameters, namely, γ and E_g^0. This underscores the efficacy of the ON in reflecting the hopping integral and bond strength. In addition, the energy levels of the magic PAH clusters are also accurately predicted, as plotted in Fig. 6. Therefore, the ON of bonds serves as a reliable descriptor for predicting the electronic structures of PAHs, effectively capturing both structural stability and charge density.

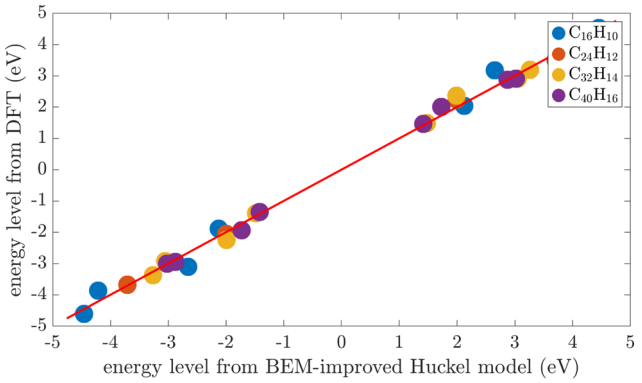

FIG. 6. Excellent agreement between energy level predicted by the BEM-improved Hückel model and DFT.

To investigate the wavefunction of molecular orbitals, we choose some typical molecular orbitals of the $C_{26}H_{16}$ molecule to demonstrate the accuracy of our TB model. The coefficients of each atomic orbital are the eigenvectors corresponding to the energy level as detailed in the "Molecular orbital predictions for the $C_{26}H_{16}$ molecule" section of the supplementary material. It is worth noticing that the BEM-improved Hückel model cannot apparently improve the Hückel model for describing π molecular orbitals.

As shown in Fig. 7, tetracene with a linear arrangement of acenes is less stable than triphenylene with Y-type acenes, and the energy gap of tetracene is smaller than that of triphenylene.[22] Here, we define the mean value κ of ONs on six-membered rings to describe the local aromaticity. We can assert that the ring is akin to the "fully aromatic" benzene if κ approaches 1.5; otherwise, it exhibits less aromaticity. It is apparent that κ values in triphenylene are larger than those in tetracene, which suggests additional aromaticity and stability.[17] The defined κ offers an avenue to describe the aromaticity, providing insight into the electronic properties of PAHs. Furthermore, triphenylene can self-assemble into a hexabenzokekulene with sixfold symmetry,[67] which has the largest energy gap among all isomers. Note that the hexabenzokekulene holds

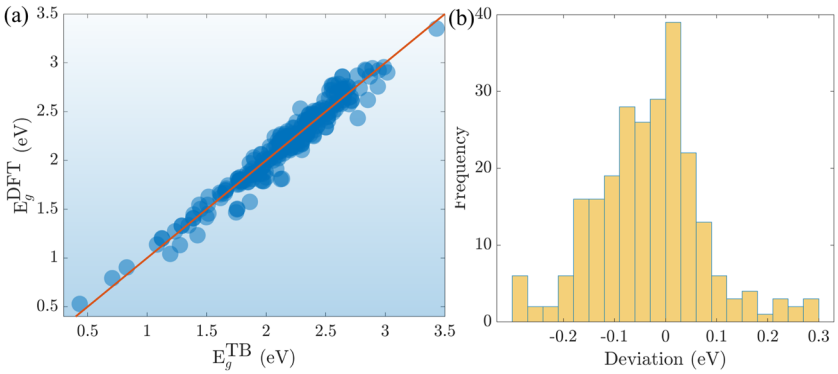

FIG. 5. (a) The linear relationship between the energy gaps predicted by the TB method and those calculated by DFT. (b) The frequency distribution of deviation between the TB method and the DFT method.

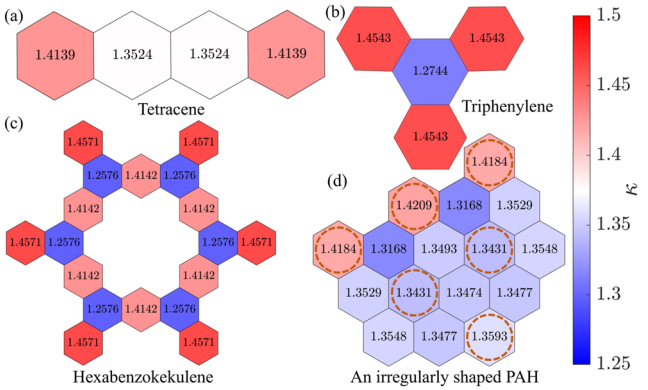

FIG. 7. Defined local aromaticity for four representative molecules: (a) tetracene, (b) triphenylene, (c) hexabenzokekulene, and (d) an irregularly shaped PAH, where the color corresponds to the right colorbar.

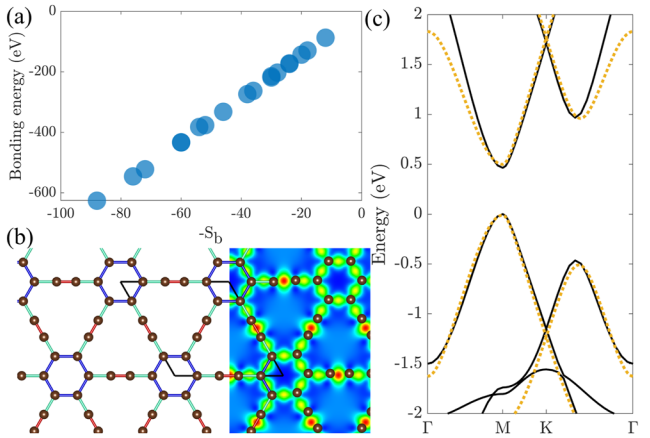

FIG. 8. (a) The linear correlation between S_b and bonding energy derived from DFT. (b) Distribution of ON and charge density for the γ-graphyne structure. (c) The band structure determined by the DFT method (solid line) and the TB method (dashed line).

12 highly aromatic rings with the red coloring of the 18 hexagonal rings, in which a particularly high proportion of aromatic rings results in a larger bandgap and high stability. Figure 7 shows the large, irregularly shaped PAHs as Clar's rule-breaking examples,[21] where Clar's rule shows that the A and C rings are more aromatic than rings B and D. However, according to several aromaticity descriptors on the rings [$MCI^{1/n}$, HOMA, and $NICS(1)_{zz}$], the aromaticity of rings A–D is comparable,[23] which is also in good agreement with the local aromaticity description in our model. For a carbon nanostructure of regularly fused pentagon–heptagon pairs, our model can also predict the accurate aromaticity[68] as detailed in the supplementary material. It must be mentioned that our model's description of anti-aromaticity is not sufficiently accurate, which requires further research in future work. We have provided some perspectives on this issue in the supplementary material.

For sp^2 bonded carbon nanoclusters with 2c–2e bonds, the combined use of the S_b and the octet rule has been effective in predicting various properties. We have extended this approach to sp–sp^2 hybridized graphyne structures,[69,70] showing the capability of S_b in forecasting structural properties in periodic systems. Unlike isolated PAH structures, periodic boundary conditions need to be considered in the unit cell for graphyne. We have systematically generated a range of nonequivalent candidates for sp–sp^2 hybridized carbon allotropes, covering numbers of carbon atoms from 12 to 60.[71] The definition of S_b for graphyne is where the globally maximal S_b corresponds to the bonding energy. Consequently, the S_b can elucidate structural stability in both molecular and periodic systems, highlighting its wide-ranging applicability.

Moving beyond structural stability, we focus on the electronic structures of graphynes. For the synthesized γ-graphyne,[72] we present the predicted ON of each bond alongside the charge density in Fig. 8(b). It is observed that bonds between two-coordinate carbon atoms are stronger than those between three-coordinate carbon atoms, as indicated by the higher ON and the red area in the charge density distribution. By optimizing the ON of each bond to maximize S_b, we construct a TB model for graphyne, contributed by p_z electrons, to analyze electronic structures. In this model, the hopping integral is proportional to the exponential of the ON. The TB-based band structures (dashed line), as given in Fig. 8(c), show considerable agreement with DFT results (solid line) near the Fermi level. A single proportionality coefficient is fine-tuned to match the bandgap determined by DFT. This matching process underscores the effectiveness of ON in differentiating hopping integrals between the nearest carbon atoms. The band structure predictions for other graphyne structures are further detailed in Fig. S15 of the supplementary material. Here, the optimal ON for each bond serves as a reliable measure of the bond strength and effectively describes the electronic structures of various dimensional systems.

IV. CONCLUSIONS

In conclusion, we have developed a graph-based statistical model that determines the electron density and related properties in carbon nanostructures across various dimensions. Without relying on pre-training data from DFT, the bonding entropy defined in our model shows a linear correlation with the bonding energy derived from DFT. This highlights the crucial role of bonding entropy in stabilizing carbon nanostructures with delocalized electrons. The optimal occupancy number of each bond accurately reflects bond strength and closely correlates with the bond length and charge density distribution. Furthermore, the mean value of the occupancy number offers a measure of local aromaticity for six-membered rings, aligning well with current calculations for aromaticity. By incorporating a bonding entropy model, we provide a statistical perspective to understand the quantum behaviors of electrons, offering a concise yet precise approach for determining structural stability and electronic properties in carbon nanoclusters and graphynes, which may also hold potential for broad applicability in other systems as well. In the end, we summarize the key limitations of the current model and possible directions for future work.

- The anti-aromatic carbon nanostructures are not described accurately because the electron localization contrasts with the assumption by the maximum entropy principle. Thus,

proper correction terms should be introduced to balance the electron delocalization and localization.
- In the B/N doped carbon nanostructures, the electron delocalization is maintained. The weights of C–N bonds are different from those of C–B bonds, and extra parameters should be fit. For instance, the weights of C–N bonds are different from those of C–B bonds. The heteroatom cannot destroy the electron delocalization of systems.
- The BEM may be extended to describe diradical characteristics in the carbon-rich and boron-rich systems, which are dominated by electron delocalization.

SUPPLEMENTARY MATERIAL

The supplementary material contains information on the computational methods, the detailed derivation of the bonding entropy model, the energy level and aromaticity prediction of PAH molecules, the comparison between existing models, the extension for BEM, the perspective for understanding anti-aromatic systems, and the band structure prediction of graphynes.

ACKNOWLEDGMENTS

This work is supported by the Guangdong Basic and Applied Basic Research Foundation (Grant Nos. 2023A1515110894 and 2023A1515010734), the Guangdong Provincial Key Laboratory of Functional and Intelligent Hybrid Materials and Devices (Grant No. 2023-GDKLFIHMD-04), the National Natural Science Foundation of China (Grant Nos. 12474228, 12374182 and 12204224), and the Shenzhen Science and Technology Program (Grant No. RCYX20200714114523069).

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

C.-C.H. and S.-G.X. contributed equally to this work.

Chang-Chun He: Conceptualization (equal); Data curation (equal); Formal analysis (equal); Investigation (equal); Methodology (equal); Visualization (equal); Writing – original draft (equal). **Shao-Gang Xu**: Conceptualization (equal); Data curation (equal). **Jiarui Zeng**: Methodology (equal); Writing – original draft (equal). **Weijie Huang**: Data curation (equal); Formal analysis (equal). **Yao Yao**: Supervision (equal); Writing – review & editing (equal). **Yu-Jun Zhao**: Writing – original draft (equal); Writing – review & editing (equal). **Hu Xu**: Project administration (equal); Writing – review & editing (equal). **Xiao-Bao Yang**: Project administration (equal); Supervision (equal); Writing – review & editing (equal).

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

[1]B. T. Matthias, T. H. Geballe, and V. B. Compton, "Superconductivity," Rev. Mod. Phys. **35**, 1–22 (1963).
[2]K. von Klitzing, "Quantum hall effect: Discovery and application," Annu. Rev. Condens. Matter Phys. **8**, 13–30 (2017).
[3]R. Tsu and L. Esaki, "Tunneling in a finite superlattice," Appl. Phys. Lett. **22**, 562–564 (1973).
[4]M. Röntgen, M. Pyzh, C. V. Morfonios, N. E. Palaiodimopoulos, F. K. Diakonos, and P. Schmelcher, "Latent symmetry induced degeneracies," Phys. Rev. Lett. **126**, 180601 (2021).
[5]H. A. Jahn and E. Teller, "Stability of polyatomic molecules in degenerate electronic states - I—orbital degeneracy," Proc. R. Soc. London, Ser. A **161**, 220–235 (1937).
[6]P. Muller, "Glossary of terms used in physical organic chemistry (IUPAC recommendations 1994)," Pure Appl. Chem. **66**, 1077–1184 (1994).
[7]A. Narita, X.-Y. Wang, X. Feng, and K. Müllen, "New advances in nanographene chemistry," Chem. Soc. Rev. **44**, 6616–6643 (2015).
[8]X.-Y. Wang, A. Narita, and K. Müllen, "Precision synthesis versus bulk-scale fabrication of graphenes," Nat. Rev. Chem **2**, 0100 (2017).
[9]Y. Liu, P. Kilby, T. J. Frankcombe, and T. W. Schmidt, "The electronic structure of benzene from a tiling of the correlated 126-dimensional wavefunction," Nat. Commun. **11**, 1210 (2020).
[10]K. Fries, "Über bicyclische Verbindungen und ihren Vergleich mit dem Naphtalin. III. Mitteilung," Adv. Cycloaddit. **454**, 121–324 (1927).
[11]K. Fries, R. Walter, and K. Schilling, "Über tricyclische verbindungen, in denen naphtalin mit einem heterocyclus anelliert ist," Adv. Cycloaddit. **516**, 248–285 (1935).
[12]A. Ciesielski, T. M. Krygowski, and M. K. Cyrański, "How to find the fries structures for benzenoid hydrocarbons," Symmetry **2**, 1390–1400 (2010).
[13]P. W. Fowler, W. Myrvold, and W. H. Bird, "Counterexamples to a proposed algorithm for fries structures of benzenoids," J. Math. Chem. **50**, 2408–2426 (2012).
[14]E. Clar, "The aromatic sextet," in *Mobile Source Emissions Including Policyclic Organic Species*, edited by D. Rondia, M. Cooke and R. K. Haroz (Springer, Netherlands, Dordrecht, 1983), pp. 49–58.
[15]Y.-W. Son, M. L. Cohen, and S. G. Louie, "Energy gaps in graphene nanoribbons," Phys. Rev. Lett. **97**, 216803 (2006).
[16]T. Cao, F. Zhao, and S. G. Louie, "Topological phases in graphene nanoribbons: Junction states, spin centers, and quantum spin chains," Phys. Rev. Lett. **119**, 076401 (2017).
[17]M. Randić, "Aromaticity of polycyclic conjugated hydrocarbons," Chem. Rev. **103**, 3449–3606 (2003).
[18]T. Wassmann, A. P. Seitsonen, A. M. Saitta, M. Lazzeri, and F. Mauri, "Clar's theory, π-electron distribution, and geometry of graphene nanoribbons," J. Am. Chem. Soc. **132**, 3440–3451 (2010).
[19]M. Randić and A. T. Balaban, "Partitioning of π-electrons in rings for clar structures of benzenoid hydrocarbons," J. Chem. Inf. Model. **46**, 57–64 (2006).
[20]Y. Ruiz-Morales, "The agreement between Clar structures and nucleus-independent chemical shift values in pericondensed benzenoid polycyclic aromatic hydrocarbons: An application of the Y-rule," J. Phys. Chem. A **108**, 10873–10896 (2004).
[21]N. Nishina, M. Makino, and J.-i. Aihara, "Aromatic character of irregular-shaped nanographenes," J. Phys. Chem. A **120**, 2431–2442 (2016).
[22]K. Strutyński, A. Mateo-Alonso, and M. Melle-Franco, "Clar rules the electronic properties of 2d π-conjugated frameworks: Mind the gap," Chem. - Eur. J. **26**, 6569–6575 (2020).
[23]Y. Wang, "Quantitative resonance theory based on the Clar sextet model," J. Phys. Chem. A **126**, 164–176 (2022).
[24]Y. Wang, "Extension and quantification of the fries rule and its connection to aromaticity: Large-scale validation by wave-function-based resonance analysis," J. Chem. Inf. Model. **62**, 5136–5148 (2022).
[25]I. Gutman, S. Radenkovic, M. Antic, and J. DJurdjevic, "A test of Clar aromatic sextet theory," J. Serb. Chem. Soc. **78**, 1539–1546 (2013).
[26]S. j. Cyvin and I. Gutman, "Kekulé structures and their symmetry properties," Comput. Math. Appl. **12**, 859–876 (1986).

[27]B. R. Westbrook and R. C. Fortenberry, "Taming semi-empirical methods for PAHS and vibrational spectra," J. Mol. Spectrosc. **398**, 111846 (2023).
[28]W. C. Herndon and M. L. Ellzey, Jr., "Resonance theory. V. Resonance energies of benzenoid and nonbenzenoid .pi. Systems," J. Am. Chem. Soc. **96**, 6631–6642 (1974).
[29]R. M. Parrondo, P. Karafiloglou, R. R. Pappalardo, and E. Sanchez Marcos, "Calculation of the weights of resonance structures of molecules in solution," J. Phys. Chem. **99**, 6461–6467 (1995).
[30]E. D. Glendening and F. Weinhold, "Natural resonance theory: I. General formalism," J. Comput. Chem. **19**, 593–609 (1998).
[31]P. García-Fernández, J. C. Wojdeł, J. Íñiguez, and J. Junquera, "Second-principles method for materials simulations including electron and lattice degrees of freedom," Phys. Rev. B **93**, 195137 (2016).
[32]W. M. C. Foulkes and R. Haydock, "Tight-binding models and density-functional theory," Phys. Rev. B **39**, 12520–12536 (1989).
[33]Z.-P. Cao, Y.-J. Zhao, J.-H. Liao, and X.-B. Yang, "Gap maximum of graphene nanoflakes: A first-principles study combined with the Monte Carlo tree search method," RSC Adv. **7**, 37881–37886 (2017).
[34]A. Chandrasekaran, D. Kamal, R. Batra, C. Kim, L. Chen, and R. Ramprasad, "Solving the electronic structure problem with machine learning," Npj Comput. Mater. **5**, 22 (2019).
[35]P. B. Jørgensen and A. Bhowmik, "Equivariant graph neural networks for fast electron density estimation of molecules, liquids, and solids," Npj Comput. Mater. **8**, 183 (2022).
[36]B. G. del Rio, B. Phan, and R. Ramprasad, "A deep learning framework to emulate density functional theory," Npj Comput. Mater. **9**, 158 (2023).
[37]Y. Zhao, L. Wang, J. Luo, T. Huang, S. Tao, J. Liu, Y. Yu, Y. Huang, X. Liu, and J. Ma, "Deep learning prediction of polycyclic aromatic hydrocarbons in the high arctic," Environ. Sci. Technol. **53**, 13238–13245 (2019).
[38]W. Kossel, "Über molekülbildung als frage des atombaus," Ann. Phys. **354**, 229–362 (1916).
[39]R. D. Harcourt, "The electronic structure of the benzene molecule," Nature **329**, 491–492 (1987).
[40]Y. Zhang, F. K. Sheong, and Z. Lin, "Natural fragment bond orbital method for interfragment bonding interaction analysis," J. Am. Chem. Soc. **146**(50), 34591–34599 (2024).
[41]I. Langmuir, "The octet theory of valence and its applications with special reference to organic nitrogen compounds," J. Am. Chem. Soc. **42**, 274–292 (1920).
[42]S. Xu, C. He, Y. Zhao, X. Yang, and H. Xu, "Generalized octet rule with fractional occupancies for boron," J. Am. Chem. Soc. **145**, 25003–25009 (2023).
[43]M. Solà, "Forty years of Clar's aromatic π-sextet rule," Front. Chem. **1**, 22 (2013).
[44]A. Karton, "How reliable is DFT in predicting relative energies of polycyclic aromatic hydrocarbon isomers? Comparison of functionals from different rungs of Jacob's ladder," J. Comput. Chem. **38**, 370–382 (2017).
[45]E. D. Glendening, C. R. Landis, and F. Weinhold, "Resonance theory reboot," J. Am. Chem. Soc. **141**, 4156–4166 (2019).
[46]G. Kresse and J. Furthmüller, "Efficient iterative schemes for *ab initio* total-energy calculations using a plane-wave basis set," Phys. Rev. B **54**, 11169–11186 (1996).
[47]G. Kresse and D. Joubert, "From ultrasoft pseudopotentials to the projector augmented-wave method," Phys. Rev. B **59**, 1758–1775 (1999).
[48]J. Paier, M. Marsman, K. Hummer, G. Kresse, I. C. Gerber, and J. G. Ángyán, "Screened hybrid density functionals applied to solids," J. Chem. Phys. **124**, 154709 (2006).
[49]J. Heyd, G. E. Scuseria, and M. Ernzerhof, "Hybrid functionals based on a screened Coulomb potential," J. Chem. Phys. **118**, 8207–8215 (2003).
[50]K. Du and Y. Wang, "Infinitenes as the most stable form of cycloarenes: The interplay among π delocalization, strain, and π–π stacking," J. Am. Chem. Soc. **145**, 10763–10778 (2023).
[51]K. Du and Y. Wang, "Generalized kekulenes and clarenes as novel families of cycloarenes: Structures, stability, and spectroscopic properties," Phys. Chem. Chem. Phys. **26**, 7877–7889 (2024).
[52]Y. Wang, *The geninfi program*, https://github.com/yangwangmadrid/GenInfi (2023).
[53]Y. Wang, Y. Zhou, and K. Du, "Enumeration, nomenclature, and stability rules of carbon nanobelts," J. Chem. Inf. Model. **64**, 1261–1276 (2024).
[54]Y. Wang, *The gencnb program*, https://github.com/yangwangmadrid/GenCNB (2023).
[55]C. C. Moore, "Ergodic theorem, ergodic theory, and statistical mechanics," Proc. Natl. Acad. Sci. U. S. A. **112**, 1907–1911 (2015).
[56]J. S. Rowlinson, "Probability, information and entropy," Nature **225**, 1196–1198 (1970).
[57]S. Noorizadeh and E. Shakerzadeh, "Shannon entropy as a new measure of aromaticity, shannon aromaticity," Phys. Chem. Chem. Phys. **12**, 4742–4749 (2010).
[58]C. Rong, B. Wang, D. Zhao, and S. Liu, "Information-theoretic approach in density functional theory and its recent applications to chemical problems," WIREs Comput. Mol. Sci. **10**, e1461 (2020).
[59]F. Fratev, D. Bonchev, and V. Enchev, "A theoretical information approach to ring and total aromaticity in ground and excited states," Croat. Chem. Acta **53**, 545–554 (1980).
[60]S. Chen, P. Slattum, C. Wang, and L. Zang, "Self-assembly of perylene imide molecules into 1D nanostructures: Methods, morphologies, and applications," Chem. Rev. **115**, 11967–11998 (2015).
[61]W. Jiang and Z. Wang, "Molecular carbon imides," J. Am. Chem. Soc. **144**, 14976–14991 (2022).
[62]L. Ma, Y. Han, Q. Shi, and H. Huang, "The design, synthesis and application of rubicene based polycyclic aromatic hydrocarbons (PAHS)," J. Mater. Chem. C **11**, 16429–16438 (2023).
[63]Y. Ruiz-Morales, "HOMO−LUMO gap as an index of molecular size and structure for polycyclic aromatic hydrocarbons (PAHs) and asphaltenes: A theoretical study. I," J. Phys. Chem. A **106**, 11283–11308 (2002).
[64]Y. Ruiz-Morales, "Application of the Y-rule and theoretical study to understand the topological and electronic structures of polycyclic aromatic hydrocarbons from atomic force microscopy images of soot, coal asphaltenes, and petroleum asphaltenes," Energy Fuels **36**, 8725–8748 (2022).
[65]F. E. Harris, "Self-consistent methods in Hückel theory," J. Chem. Phys. **48**, 4027–4028 (1968).
[66]W. P. Su, J. R. Schrieffer, and A. J. Heeger, "Solitons in polyacetylene," Phys. Rev. Lett. **42**, 1698–1701 (1979).
[67]S. J. Cyvin, J. Brunvoll, and B. N. Cyvin, "Enumeration and classification of coronoid hydrocarbons. 10. Double coronoids," J. Chem. Inf. Comput. Sci. **30**, 210–222 (1990).
[68]I. C.-Y. Hou, Q. Sun, K. Eimre, M. Di Giovannantonio, J. I. Urgel, P. Ruffieux, A. Narita, R. Fasel, and K. Müllen, "On-surface synthesis of unsaturated carbon nanostructures with regularly fused pentagon–heptagon pairs," J. Am. Chem. Soc. **142**, 10291–10296 (2020).
[69]D. Malko, C. Neiss, F. Viñes, and A. Görling, "Competition for graphene: Graphynes with direction-dependent Dirac cones," Phys. Rev. Lett. **108**, 086804 (2012).
[70]B. G. Kim and H. J. Choi, "Graphyne: Hexagonal network of carbon with versatile Dirac cones," Phys. Rev. B **86**, 115435 (2012).
[71]S.-G. Xu, X.-T. Li, Z.-J. Chen, C.-C. He, C. He, X.-B. Yang, and H. Xu, "Toward hidden materials with directional bonds," Phys. Rev. Mater. **7**, 084202 (2023).
[72]Y. Hu, C. Wu, Q. Pan, Y. Jin, R. Lyu, V. Martinez, S. Huang, J. Wu, L. J. Wayment, N. A. Clark, M. B. Raschke, Y. Zhao, and W. Zhang, "Synthesis of γ-graphyne using dynamic covalent chemistry," Nat. Synth. **1**, 449–454 (2022).