

Graphene with vacancies: Supernumerary zero modes

Norman Weik,^{1,2,*} Johannes Schindler,^{1,2,†} Soumya Bera,³ Gemma C. Solomon,^{4,5} and Ferdinand Evers⁶

¹*Institute of Nanotechnology, Karlsruhe Institute of Technology, Campus North, D-76344 Karlsruhe, Germany*

²*Institut für Theorie der Kondensierten Materie, Karlsruhe Institute of Technology, Campus South, D-76128 Karlsruhe, Germany*

³*Max-Planck-Institut für Physik komplexer Systeme, D-01187 Dresden, Germany*

⁴*Nano-Science Center, University of Copenhagen, DK-2100 Copenhagen, Denmark*

⁵*Department of Chemistry, University of Copenhagen, DK-2100 Copenhagen, Denmark*

⁶*Institute of Theoretical Physics, University of Regensburg, D-93050 Regensburg, Germany*

(Received 15 April 2016; revised manuscript received 22 July 2016; published 26 August 2016)

The density of states $\rho(E)$ of graphene is investigated within the tight-binding (Hückel) approximation in the presence of vacancies. They introduce a nonvanishing density of zero modes n_{zm} that act as midgap states, $\rho(E) = n_{zm}\delta(E) + \text{smooth}$. As is well known, the actual number of zero modes per sample can, in principle, exceed the sublattice imbalance, $N_{zm} \geq |N_A - N_B|$, where N_A, N_B denote the number of carbon atoms in each sublattice. In this paper, we establish a stronger relation that is valid in the thermodynamic limit and that involves the *concentration* of zero modes, $n_{zm} > |c_A - c_B|$, where c_A and c_B denote the concentration of vacancies per sublattice; in particular, n_{zm} is nonvanishing even in the case of balanced disorder, $N_A/N_B = 1$. Adopting terminology from benzoid graph theory, the excess modes associated with the current carrying backbone (percolation cluster) are called *supernumerary*. In the simplest cases, such modes can be associated with structural elements such as carbon atoms connected with a single bond, only. Our result suggests that the continuum limit of bipartite hopping models supports nontrivial “supernumerary” terms that escape the present continuum descriptions.

DOI: [10.1103/PhysRevB.94.064204](https://doi.org/10.1103/PhysRevB.94.064204)

I. INTRODUCTION

As is well known, Dirac particles are realized, e.g., in clean graphene, which exhibits two Dirac cones at the \mathbf{K} and \mathbf{K}' points of the Brillouin zone [1]. Another important incarnation is met in energy spectra of quasiparticles over [two-dimensional (2D)] condensates of fermions with p -type pairing [2,3]. In three dimensions, the Dirac dispersion can be realized in Weyl- and Dirac-type semimetals [4,5].

One out of many interesting aspects of Dirac fermions relates to their topological properties [1,6]. When punching a hole (“defect”) into Dirac gases, a zero-energy state can form at its boundary. Such zero modes indicate a demarcation line that separates two topologically distinct phases (e.g., condensate and trivial vacuum) from each other. The number of zero modes relates, via the Atiyah-Singer index theorem [7], to topological charges of certain gauge field configurations that represent the holes as vortices in continuum descriptions.

A natural lattice representation of topological defects are vacancies in tight-binding models of graphene. Indeed, an isolated vacancy carries a zero-energy mode that exhibits the characteristic $1/(x + iy)$ decay away from the defect [8]. We can associate with every vortex a unit of topological charge, with a sign that is positive for one and negative for the other sublattice. Then, a first expectation based on the index theorem familiar from the continuum theory could be that a mismatch of vacancy concentrations in the two sublattices c_A, c_B would induce a density $n_{zm} = |c_A - c_B|$ of zero modes. Moreover, in

the limit of balanced sublattices, $c_A = c_B$, one would expect a cancellation of topological charges and therefore $n_{zm} = 0$.

In this paper, we are mostly interested in the spectra of tight-binding models of large (but finite) graphene flakes. In this case, a kind of lattice analog of the continuum version of the Atiyah-Singer index theorem applies,

$$\zeta \geq |N_A - N_B|, \quad (1)$$

where N_A, N_B denote the number of sites in the respective sublattices and ζ is the number of zero modes. Equation (1) has been derived for general bipartite lattices by Inue, Trugman, and Abrahams and assumes a connected lattice graph [9]. Since (1) is an inequality, the index theorem can deliver only a lower bound and in the perfectly balanced situation, $N_A/N_B = 1$, it does not give any information at all, strictly speaking. As we will demonstrate in the following, the tight-binding (TB) model for graphene flakes exhibits a finite concentration of zero modes even in the situation of balanced sublattices. A discussion of possible consequences for graphene’s material properties will be offered.

A few preparatory remarks are in order. There are two reasons why the estimate based on “=,” as adopted in the continuum index theorem, tends to be a very crude one, when considering a graphene (TB) lattice decorated with vacancies.

(i) At every nonzero concentration of vacancies, $c_A, c_B > 0$, there is a nonvanishing probability that a sequence of vacancies punched into pristine graphene forms a closed line, i.e., a loop. This loop cuts the sample into two separate fragments (“cluster”). The estimate Eq. (1) applies to each fragment (indices 1,2) separately, $N_{zm} \geq |N_{A1} - N_{B1}| + |N_{A2} - N_{B2}|$, $N_{zm} = \zeta_1 + \zeta_2$, which, in general, gives a nonvanishing result for the total number of zero modes N_{zm} per sample, even if $|N_{A1} + N_{A2} - N_{B1} - N_{B2}| = 0$. Generic continuum

*Present affiliation: Institute of Transport Science, RWTH Aachen University, D-52056 Aachen, Germany.

†Present affiliation: Institut für Technische Optik, Pfaffenwaldring 9, D-70569 Stuttgart, Germany.

treatments do not incorporate the effects of correlated disorder and therefore fragmentation escapes their scope.

(ii) Inue *et al.* were interested in the localization properties of fermions on *connected* bipartite lattices (single cluster, no fragments) as representatives of the chiral universality class [9,10]. Their result is, in fact, a special case of a stronger inequality that has been proven earlier. Longuet-Higgins [11] showed in 1949 that the number of zero modes is bounded from below by $\zeta \geq N - 2N_{\text{db}}$, where N_{db} denotes the maximum number of double bonds that can be placed on a given graphene cluster (“flake”) with $N = N_A + N_B$ sites [12]. This latter result, though derived for graphene flakes, is valid for the much more general situation of hopping on bipartite graphs (see Appendix A for further details). Recent research on benzoidal graph theory sharpens the statement further: On the honeycomb lattice the equality sign holds true, so that we have the exact statement [13]

$$\zeta = N - 2N_{\text{db}}. \quad (2)$$

As a consequence, and as we further exemplify below, even a single balanced cluster $N_{A_i} = N_{B_i}$ (cluster index i) exhibits, in general, a number of “extra” modes due to the mismatch $N_{\text{db}} < N_A, N_B$. We call these modes “supernumerary,” while the conventional ones (due to sublattice imbalance) are called “predictable” [14]. Only in the simplest cases the supernumerary zero modes (SZMs) can be traced back to particular structural elements, such as singly connected lattice sites (coordination number one, “dangling sites”). In the present paper we establish, using numerical simulations, that SZMs come with a macroscopic abundance on large isolated clusters; in particular, their concentration on the percolation cluster is finite.

A. Lattice animals, edge motifs, and topological modes

To illustrate the efficiency of the equality Eq. (2) we have investigated all clusters (“lattice animals”) up to seven sites, i.e., mass $s = 7$, and their associated spectra. In Fig. 1 we show as an example the smallest animal that supports extra zero modes, $\zeta = 2$, in this case. This result is verified from the associated connectivity matrix,

$$h = \begin{pmatrix} 1 & & & \\ 1 & & & \\ 1 & 1 & 1 & \end{pmatrix}. \quad (3)$$

It displays two identical pairs of columns, hence $\zeta = 2$.

The example of small animals illustrates that the proper placing of double bonds requires knowledge of the topology

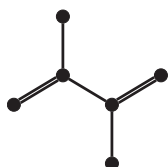


FIG. 1. The simplest example of a balanced lattice animal $N_A = N_B = 3$ that carries SZMs. There are no predictable zero modes because $N_A = N_B$. However, two double bonds can be placed and therefore we have two SZMs: $\zeta = 6 - 4 = 2$.

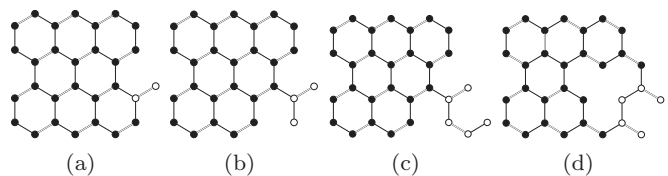


FIG. 2. Edge motifs: (a) Dangling site. (b) Double-dangling site that is associated with a zero-energy state, as can be seen from the fact that one site without double bonds is identified. (c) Double-dangling chain. (d) “U structure.”

of the full cluster. To find N_{db} for very large clusters therefore is not a trivial task; it is, in fact, exponentially hard. However, the structure of matrix (3) suggests a recipe for generating structural elements of much larger clusters that carry types of zero modes that do not require scrutinizing the full object in order to be predicted. Such elements are shown in Fig. 2: dangling sites (left) and double-dangling sites (right). Double-dangling sites constitute two singly connected sites that share a common internal port site. It is easy to see directly from placing double bonds that double-dangling sites are associated with an extra zero mode. By inspecting the connectivity matrix, the conclusion is confirmed immediately.

Hence, one has a structure

$$H = \left(\begin{array}{cc|ccc} 1 & 1 & \star & \dots & \star \\ 0 & 0 & \star & \dots & \star \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \star & \dots & \star \end{array} \right). \quad (4)$$

The two linearly dependent column vectors reflect the occurrence of a zero-energy state.

Analogous arguments can be given for more complex edge structures involving motifs similar to double-dangling sites (Fig. 2, right panels). For instance, it can be shown that all double-dangling sites where both *arms* consist of an odd number of sites lead to linearly dependent column vectors and consequently cause zero modes. Another edge motif associated with zero modes, the “U structure,” is depicted in Fig. 2 (right). It consists of a site of degree 2 in the center which is connected to two sites adjacent to dangling sites. In this case the dangling sites can also be expanded to arms leading to a whole class of structures. As for the double-dangling sites, structures having odd-numbered arms also produce zero modes.

The examples discussed thus far suggest the following picture: Equation (2) predicts zero modes of two different kinds. The number of predictable modes equals the sublattice mismatch $|N_A - N_B|$. They are expected to be topological. There are additional numbers of supernumerary zero modes that come in pairs and that exist even in the balanced case, $N_A = N_B$.

B. Percolation cluster

We discuss the implications of our findings for the spectrum of a spanning cluster. In the limit of vanishing vacancy concentration, $n_{\text{vac}} \rightarrow 0$, the percolation cluster is dense, exhibiting mostly isolated vacancies; correlated formations of vacancies that would induce pairs of extra zero modes, e.g.,

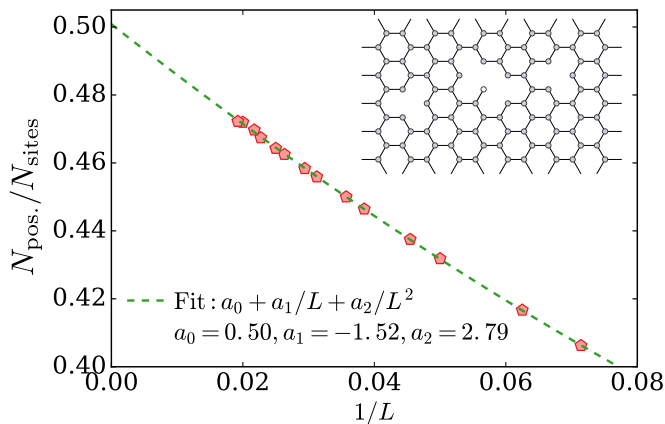


FIG. 3. Supernumerary zero modes (SZMs) in balanced graphene sheets with a single dangling site and two isolated vacancies. The plot shows that the number of possibilities $N_{\text{pos.}}$, for all placements of the pair of isolated vacancies that yield two SZMs, grows proportional to the sample area (number of lattice sites N_{sites}). Inset: Example with size 10×20 that exhibits two SZMs with double-periodic boundary conditions, which is seen by either exact diagonalization or by bond placing.

due to edge motifs, are rare:

$$\zeta(s, n_{\text{vac}}) = \zeta_{\text{p}}^{\text{pre}} + \zeta_{\text{p}}^{\text{sup}}, \quad (5)$$

where $\zeta_{\text{p}}^{\text{pre}} = |N_{\text{Ap}} - N_{\text{Bp}}|$ and $\zeta_{\text{p}}^{\text{sup}} = \mathcal{F}[n_{\text{vac}}]s$, and $\mathcal{F}[n_{\text{vac}}] \propto n_{\text{vac}}^{\alpha}$ in the limit of vanishing concentration $n_{\text{vac}} \rightarrow 0$. Here the exponent α is a positive number and $s = N_{\text{Ap}} + N_{\text{Bp}}$ is the cluster mass. To estimate α , we recall that the leading corrections are due to correlated disorder configurations. If the structures dominating the extra modes are given, e.g., by double-dangling sites, we have a scaling with n_{vac}^4 . Hence, one expects an inequality $\alpha \leq 4$.

Now, consider the limit of very large clusters for which we expect a scaling $|N_{\text{Ap}} - N_{\text{Bp}}| \propto s^{1/2}$. Therefore, the second term in Eq. (5) dominates for large enough systems, $s \gg \xi^2$. The crossover length ξ exhibits a scaling with vacancy concentration, $\xi \propto n_{\text{vac}}^{-\alpha}$.

Motivated by the investigation of the example displayed in the inset of Fig. 3, we propose that α is smaller than four. The plot shows a balanced sheet with two vacancies per sublattice. One convinces oneself, e.g., by placing double bonds or by exact diagonalization, that this example exhibits two supernumerary modes, $\zeta_{\text{p}}^{\text{sup}} = 2$. They are brought about by the combined effect of a dangling site and two single vacancies. Importantly, there is a large number of possibilities for placing the isolated vacancies with respect to the dangling site all yielding two supernumerary zero modes. For a given lattice size L we have determined this number via exact enumeration assuming a toroidal geometry (double-periodic boundary conditions). As seen from the main plot of Fig. 3, it is proportional to the sheet size as quantified, e.g., by the number of constituting lattice sites $L^2 - 4$. Since this implies that getting two supernumerary modes out of four balanced vacancies simply is proportional to the probability of forming a dangling site, the example suggests $\alpha = 2$.

We consider the analysis (Fig. 3) suggestive, but clearly it is far from conclusive, because any argument employing a

virial-type expansion is underlying an assumption of separability. It is not obvious at this point to what extent such an assumption will hold for SZMs.

To illustrate the difficulty, we change the boundary conditions cutting the torus open, so it forms a zigzag carbon nanotube. Then, two zero modes are found even in the absence of vacancies. Under the assumption of perfect separability, one would expect that zigzag nanotubes carrying a dangling site and two balancing vacancies should exhibit vacancy configurations with four zero modes, two boundary modes plus two extra modes related to the vacancies. This, however, is not the case; in all cases only two zero modes are observed. Apparently, boundary modes tend to mix with the vacancy-induced modes and a naive notion of additivity does not hold. To avoid boundary-related complications, our focus has been on the torus geometry.

The experience with boundary modes we interpret as a hint that separability is not guaranteed, at least not within the small sample sizes L available to us. This provides an additional motivation for us to proceed with the numerical analysis and investigate graphene sheets with a finite concentration of balanced vacancies.

II. ZERO MODES ON PERCOLATION CLUSTERS: NUMERICAL SIMULATION

We confirm our analytical considerations and demonstrate numerically that for compensated disorder the largest fraction of zero modes is of the supernumerary kind, i.e., not predictable. Our focus is on percolating clusters that we obtain in the following way: A pristine flake of size $L \times L$ is generated with periodic boundary conditions. Vacancies are added at a given concentration, keeping $N_{\text{A}} = N_{\text{B}}$. The percolation cluster is the biggest fragment that remains from the original flake. We identify it by employing the Hoshen-Kopelman algorithm (see Appendix B for details).

A. Results

Figure 4 displays an ensemble averaged number of zero modes on the percolation cluster as a function of the cluster size at a fixed vacancy concentration $n_{\text{vac}} = 20\%$. We make three basic observations: (i) The expected scaling $\sim \sqrt{s}$ for the predictable modes is confirmed. (ii) The total number of zero modes scales with $\sim s$, as expected. (iii) A considerable fraction of these modes is due to the simple (nontopological) edge motifs (dangling sites and “U”-type structures, see Fig. 3; for additional details, see Ref. [15]) that we discussed before. The probability of finding such structures scales with a power in n_{vac} of order 4 or higher. Due to our expectation $\alpha \lesssim 4$, we do not expect edge motifs, such as Figs. 2(a)–2(c), to give the dominant contribution in the low concentration regime.

In Fig. 5 we display how the number of supernumerary zero modes associated with the percolation cluster disappears with decreasing vacancy concentration. Percolation clusters have been drawn from systems of sizes $L = 100, \dots, 140$. At vacancy densities exceeding $\sim 7\%$ the data can safely be extrapolated into the macroscopic limit. However, at smaller concentrations, finite size effects are substantial. As a consequence, the asymptotic limit of low n_{vac} , which would allow

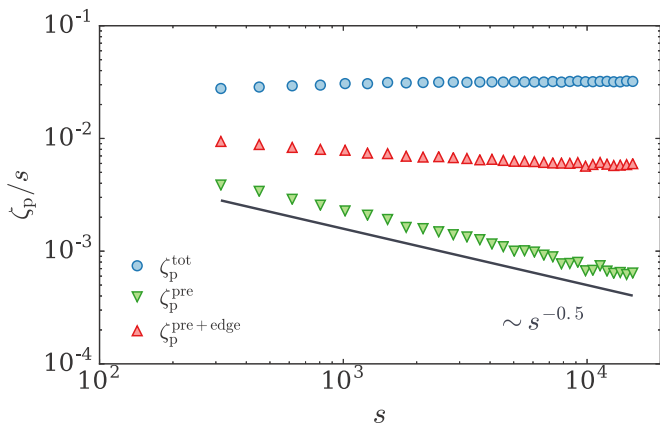


FIG. 4. Average number of zero modes associated with the percolation cluster obtained in systems of size $L = 20, 22, \dots, 128$ over its (average) mass s . The number of zero modes ζ_p^{pre} based on sublattice imbalance and based on the combined effects of sublattice imbalance and edge structures are compared ($n_{\text{vac}} = 20\%$; average of 1000 realizations for systems up to 10 000 sites and over 100 realizations for systems exceeding that size). In order to highlight the effect of edge motifs (Fig. 3) at large concentrations, we also show a curve (Δ) where they have been added to the predictable modes.

for a numerical determination of α , is not readily observable. Notice, however, that ζ_p^{sup} increases by about a factor of 3 with L changing by a factor of ~ 1.4 near vacancy concentration $\sim 3\%$. We would like to interpret this as an indication that ζ_p^{sup} is overshooting the dotted-dotted-dashed line in Fig. 5 in the limit of large L , signaling that indeed $\alpha < 4$. Hence, we believe that a relatively small value for the exponent that would be consistent with the data at concentrations exceeding $\sim 10\%$, roughly $\alpha \approx 2$, could conform with the existing numerical data also at lower vacancy concentrations.

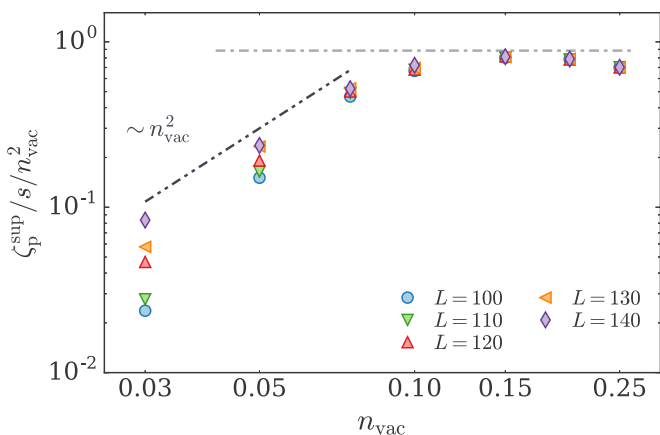


FIG. 5. Average number of supernumerary zero modes associated with the percolation cluster obtained in systems of size $L = 100, 110, \dots, 140$ over the vacancy concentration n_{vac} in a double logarithmic scale. The dashed-dotted-dotted line is a guide indicating a conservative upper bound ($\alpha = 4$) for the power after finite size extrapolation. A value $\alpha = 2$ corresponds to the horizontal line as shown in the figure (dashed-dotted-dashed).

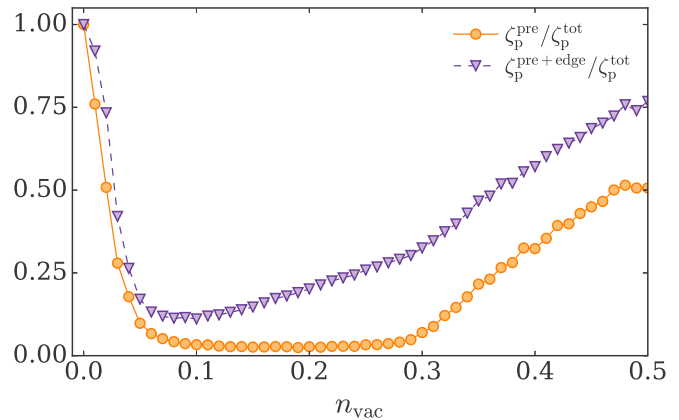


FIG. 6. Zero modes on a percolation cluster. The total number of modes (∇) and the fraction assigned to only sublattice imbalance (\circ) is shown. Data are normalized to the total number of zero modes in the graphene sample (size: $L = 100$; average over 500 disorder realizations).

Figure 6 shows once again an overview over a large range of vacancy concentrations. The fraction of zero modes that one finds on average on the percolation cluster is compared with the number of modes, $\zeta_p^{\text{pre}} = |N_{A_p} - N_{B_p}|$, predictable from the sublattice imbalance (at a given system size, $L = 100$). In the clean limit, $n_{\text{vac}} \rightarrow 0$, the crossover length exceeds our system size, $\xi \gg L$, and all zero modes are seen to be predictable. Towards larger concentrations the crossover length ξ decreases and, eventually in the range $n_{\text{vac}} \approx 1\% - 3\%$ ξ , exceeds the system size $L = 100$. In that regime, the supernumerary zero modes dominate and at concentrations above 7% they contribute more than 80%.

B. Discussion: Zero modes and graphene's material properties

Before the discussion of the impact of zero modes, we recall that vacancies in tight-binding lattices (TB vacancies) are not meant to model missing carbon atoms (real vacancies) in the graphene material. For example, real vacancies give rise to strong lattice distortions and also, possibly, to edge reconstruction which is not included in our TB modeling. Instead, TB vacancies are better thought of as a model of carbon atoms that have been promoted from sp^2 to sp^3 hybridization via chemical functionalization [16,17].

Unbalanced fragments. Despite being conceptually trivial, the “excess modes” resulting from unbalanced fragments ($N_{A_i} \neq N_{B_i}$) can make an important contribution to the thermodynamic density of states (DOS). Therefore, they enter generic thermodynamic properties of the sample, and thus may be relevant, e.g., for magnetism. Indeed, the impact of zero modes for the electronic states and ground-state magnetism has been investigated intensively for small, isolated graphene flakes with edge hydrogenation [18]. For instance, a pair of supernumerary modes may give rise to an antiferromagnetic ground state by virtue of Hund's rule coupling in a bowtie shaped fragment called “Clar's goblet” [19]. To what extent this conclusion carries over to larger systems is, however, not clear. Part of the difficulty is that supernumerary zero modes

may not benefit from topological protection and therefore can, in principle, be strongly susceptible to interaction effects (beyond mean field) [20]. We mention that electronic structure effects related to supernumerary modes have been reported to also manifest in preferred binding geometries [21].

Percolation cluster. The transport current is carried by the percolating (“spanning”) cluster; smaller fragments are irrelevant. Analogous to the thermodynamic DOS, also the transport DOS, which is associated with the percolation cluster, exhibits a singular behavior.

As any other fragment, the spanning cluster is unbalanced, in general. Therefore, it supports a number of predictable zero modes, $\zeta_p^{\text{pre}} = |N_{A_p} - N_{B_p}|$, even with (global) compensation $N_A = N_B$. (“p” denotes the spanning/percolation cluster.) Interestingly, our results indicate that the fraction of SZMs contained in ζ_p always dominates over the predictable modes at large enough cluster sizes, $\zeta_p^{\text{pre}} \ll \zeta_p^{\text{sup}}$. Therefore, unless SZMs are typically localized they could leave a signature in transport calculations, which may not have been identified as of yet. In particular, it might be very difficult to resolve it in numerical studies relying, e.g., on the Kubo formula due to the smearing induced by the imaginary frequency shift [22].

III. CONCLUSIONS

The implications of the existence of extra zero modes for our understanding of graphene lattice models with vacancy disorder have not been previously analyzed for macroscopic systems. The main finding of the present work is that supernumerary zero modes exist in macroscopic abundance even in compensated graphene lattice models. These modes are not incorporated in the present continuum theories of this material. Therefore, the question can be raised regarding what aspects of the hydrodynamic theory of bipartite lattice models generally coincide with the present continuum theories, as they have been studied, e.g., by Gade and Wegener [23,24] and many others [10,25–29].

Our recent research has identified a low energy regime in which Gade-Wegner scaling of the DOS is violated [16]. The similarity of the numerical data with concurrent field theoretical results was interpreted as evidence that local imbalances of the vacancy distribution produce very strong disorder effects that drive the Gade-Wegner fixed point unstable [16,30]. Supernumerary zero modes, however, have not been considered in this work.

One would expect that the spectral properties away from zero energy should feel a certain impact of the (supernumerary) zero modes. We give two obvious reasons. First, the concentration of spectral weight at zero energy must lead to a depletion of spectral weight in other spectral regions. Second, eigenvectors at nonvanishing energies must be orthogonal to the (macroscopic) degenerate subspace at $E = 0$. What this implies for physical observables, and whether a new energy regime at very low energies is introduced, remains to be seen.

Note added. Recently, we became aware of the study performed by Sanyal, Damle, and Motrunich [31] which also reports a finite density of zero modes in graphene with vacancies.

ACKNOWLEDGMENTS

Discussions with P. Ostrovsky, I. Gornyi, Ch. Stafford, and R. Moessner are gratefully acknowledged. We are also grateful to V. Häfner for his contribution at the early stages of this work. S.B. thanks S. Roy for helping with the visualization. We thank I. Kondov and the Jülich Supercomputer Center (JUROPA, project HKA12) for computational assistance and resources.

APPENDIX A: ZERO MODES OF ISOLATED CLUSTERS—ANALYTICAL RESULTS

We consider the following representation of all zero-energy modes in a graphene sample,

$$N_{\text{zm}} = \sum_{\mathcal{C}} \zeta(\mathcal{G}_{\mathcal{C}}) N_{\mathcal{C}}(c_A, c_B), \quad (\text{A1})$$

where the sum is over all types of clusters \mathcal{C} (“lattice animals”) within the sample (including the percolating cluster) and $N_{\mathcal{C}}$ denotes the number of clusters of a certain type. We have introduced here the number of zero modes associated with each cluster ζ that is determined by the cluster’s graph \mathcal{G} . The graph of a cluster is defined via the connectivity matrix, which may be thought of as the piece of the tight-binding Hamiltonian $H_{\mathcal{C}}$ associated with the cluster subspace. All clusters sharing the same $H_{\mathcal{C}}$ (up to rotations) are representatives of the same graph \mathcal{G} .

The cluster numbers $N_{\mathcal{C}}$ and their dependency on the vacancy concentrations are typical objects of percolation theory [32]. They incorporate all statistical aspects of vacancy disorder in (A1) and, in particular, a classification of clusters according to their mass (total number of sites) and perimeter (number of sites forming the external boundary) [32,33]. Quantum mechanics enters (A1) via the weight factor ζ , which is an intrinsic property of a given graph.

1. Inequalities and double bonds

In order to reveal the connection between zero modes and placing of double bonds, we follow the original argument by Longuet-Higgins and reproduce his proof. To be specific, we consider a cluster of carbon atoms that form a graphene flake. Call the sublattice with the majority number of sites the A sublattice, so $N_A > N_B$. The goal is to show that

$$\zeta \geq N_A + N_B - 2N_{\text{db}} \quad (\text{A2})$$

for the number of zero modes associated with this cluster.

Proof. The tight-binding Hamiltonian H associated with the cluster can be represented as a matrix of the dimension $\dim H = N_A + N_B$ that takes the special block off-diagonal form

$$H = \begin{pmatrix} 0 & h \\ h^\dagger & 0 \end{pmatrix}, \quad (\text{A3})$$

where h is a $N_A \times N_B$ matrix that connects the A/B sublattices with each other. We start by recalling that

$$\zeta = \dim H - \text{rank } H. \quad (\text{A4})$$

The rank of H can be defined in several equivalent ways. One definition is to say that $\text{rank } H$ is given by the largest order of

any nonzero minor of H . The use of this definition will in the end connect the proof to the concept of double bonds.

Namely, consider any term that arises when calculating a minor of H ; it has the structure

$$\prod_{\bar{a}}^P \bar{h}_{\bar{a}\bar{b}} \prod_{\bar{b}}^Q h_{\bar{a}\bar{b}}. \quad (\text{A5})$$

(For notational simplicity, we define here the matrix elements $[h^\dagger]_{ab} = \bar{h}_{ab}$, $a = 1, \dots, N_A$, $b = 1, \dots, N_B$.) To visualize this term, we can imagine the set of lattice points $\{a\}, \{b\}$ that represent the sublattices A and B, respectively. Let $\bar{h}_{\bar{a}\bar{b}}$ denote a double bond that connects the pair of sites (\bar{a}, \bar{b}) . Similarly, $h_{\bar{a}\bar{b}}$ connects the pair of lattice points (a, b) via a single bond. Importantly, since the minor is a determinant (of a submatrix of H) it is guaranteed that there cannot be a site which attaches to more than one double bond or to more than one single bond.

Here, we have attributed a covering of the lattice sites with a bond pattern that follows a simple rule. We can also work the other way around and convert a lattice covering with single and double bonds, that follows these rules, into a term that appears when evaluating minors of H . This is what we will do now. The goal is to find a covering such that the number of factors appearing in Eq. (A5), $P + Q$, becomes as large as possible. This largest possible value coincides with rank H .

The covering we are after is identified as follows: We cover the set of points a, b with the maximum number of double bonds that we can place, so $P = N_{\text{db}}$. The single bonds we place along the same pairs, $Q = N_{\text{db}}$. Therefore, rank $H \leq 2N_{\text{db}}$ and the proof is complete. Note that the proof holds for any bipartite lattice, not only for the honeycomb one.

APPENDIX B: NUMERICAL METHOD

The quantitative analysis combines two numerical approaches. On the one hand, a brute force diagonalization of the TB Hamiltonian with vacancies based on Lapack's singular value decomposition (SVD) routine is performed. On the other hand, the microscopic structure of the sample is investigated by performing a cluster analysis based on a modified Hoshen-Kopelman cluster labeling algorithm [34]. By combining the two approaches the spectral properties of single clusters and microscopic structures on these clusters can be analyzed. Our present analysis is limited to $O(10^4)$ sites due to the vanishing level spacing making the identification of exact zero modes challenging in larger systems.

1. The Hoshen-Kopelman algorithm

For the microscopic analysis of the generation of zero modes, a modified Hoshen-Kopelman cluster labeling algorithm for honeycomb structures [34] is employed. It allows one to identify and label all clusters and to determine their sublattice imbalance.

The lattice algorithm scans linewise for nonvacant sites from the upper left to the lower right corner (see Fig. 7, left). If a given nonvacant site is connected to an already labeled site, it is given the same label. It is given a new label in case the adjacent sites known so far are vacancies. However, a difficulty arises if the site is adjacent to two labeled sites having different labels (see Fig. 7, right). In this case two

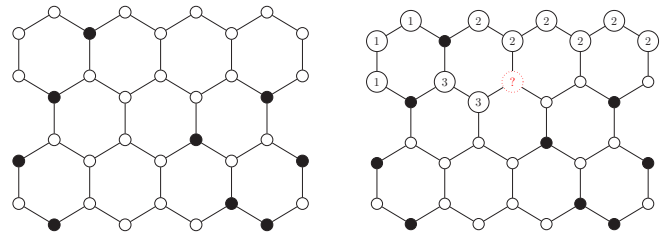


FIG. 7. Cluster labeling. Left: An exemplary graphene sample with vacancies (black). Right: Schematics of the Hoshen-Kopelman algorithm. The sites are labeled linewise: A given site is attributed the label of the sites to the left or the line above if one of them is nonvacant or they both have the same label. The marked site points out a conflict where the label to the left and the top do not match and the clusters with labels 2 and 3 get merged.

formerly independent clusters are merged, requiring backward relabeling of sites, which is numerically very expensive. This problem was a major concern in the early days of percolation theory in the 1950's and it was common belief that "the direct simulation of percolation is out of the question" [35] at that time.

The problem was solved in the middle of the 1970's with the development of the very efficient Hoshen-Kopelman algorithm [34], which allows one to label all clusters with a single sweep through the lattice. It avoids backward labeling by introducing an additional array of cluster *markers* which (1) counts the number of sites belonging to a given cluster and (2) keeps track of the relabeling of clusters which have been merged with other (larger) clusters by storing a pointer to the larger cluster it has gotten merged with. In particular, the labels of clusters which have become merged with larger ones are no longer attributed to new sites (see Fig. 8). The sample may be cleaned by adding a second sweep through the lattice and replacing "antiquated labels" (such as 3 and 4 in Fig. 8). However, it is not required since all the information is stored in the Marker array.

For more details on the Hoshen-Kopelman algorithm, the reader is referred to the original paper by Hoshen and Kopelman [34].

2. Modifications to the Hoshen-Kopelman algorithm

In this work some minor modifications were made to the Hoshen-Kopelman algorithm.

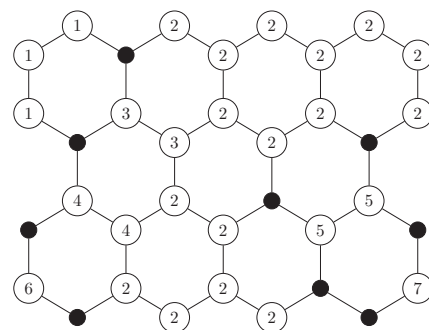


FIG. 8. The final labeling after the Hoshen-Kopelman algorithm has been performed.

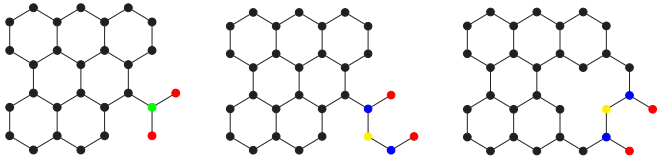


FIG. 9. Illustration of the algorithm for the identification of edge motifs using a coloring scheme. First step (red): Identification of corner sites. Second step (blue): Coloring of sites adjacent to corner sites. In case an already blue site is to be colored, this site is colored green. Third step (yellow): Coloring of sites which are adjacent to one vacancy and two blue (or green) sites.

(a) Periodic boundary conditions (PBCs) are used instead of fixed ones in order to avoid zero modes due to zigzag edges already discussed in the context of carbon nanoribbons [36]. They are integrated by adding an additional sweep along the sample's left and bottom edge once all sites have been labeled.

(b) In addition to the storage of each cluster's number of sites in the marker array, its number of carbon atoms in the two sublattices are also kept such that the sublattice imbalance of each cluster can be determined.

(c) For investigating the properties of specific clusters such as the largest (spanning) one, a second scan of the sample is added, removing all but the desired cluster(s) from the sample. This cluster spectrum can then be determined by treating

all other sites as vacant and diagonalizing the resulting TB Hamiltonian. Hence we can directly determine the spectral properties of single clusters in the sample.

3. Identification of edge motifs

For the identification of edge motifs generating zero modes such as double-dangling sites, an additional analysis is required once all clusters have been identified using the Hoshen-Kopelman algorithm. Our analysis of edge structures includes all structures up to double-dangling bonds with one arm of length 1 and one of length 3 and U structures with arms of length 1 (see Fig. 2). These structures all involve (coordination one) corner sites that are connected to only one other lattice site.

Starting with these sites, the algorithm consists of three steps: (1) All corner sites are colored red. (2) The sites connected to red sites are colored blue. In case a site has already been colored blue, it is colored green. (3) Sites adjacent to two blue-colored (or green-colored) sites and a vacancy are colored yellow.

Obviously, the green sites mark the position of double-dangling sites with arms of length 1 since they are adjacent to two (red) corner sites. The yellow sites mark the positions of double-dangling sites with arms of length 1 and 3 or U structures (see Fig. 9).

-
- [1] A. H. Castro Neto, F. Guinea, N. M. R. Peres, K. S. Novoselov, and A. K. Geim, The Electronic Properties of Graphene, *Rev. Mod. Phys.* **81**, 109 (2009).
 - [2] N. Read and D. Green, Paired states of fermions in two dimensions with breaking of parity and time-reversal symmetries and the fractional quantum Hall effect, *Phys. Rev. B* **61**, 10267 (2000).
 - [3] G. E. Volovik, *The Universe in a Helium Droplet*, International Series of Monographs on Physics (Oxford University Press, Oxford, U.K., 2009).
 - [4] X. Wan, A. M. Turner, A. Vishwanath, and S. Y. Savrasov, Topological semimetal and Fermi-arc surface states in the electronic structure of pyrochlore iridates, *Phys. Rev. B* **83**, 205101 (2011).
 - [5] P. Hosur and X. Qi, Recent developments in transport phenomena in Weyl semimetals, *C. R. Phys.* **14**, 857 (2013).
 - [6] S. Das Sarma, S. Adam, E. H. Hwang, and E. Rossi, Electronic Transport in Two-Dimensional Graphene, *Rev. Mod. Phys.* **83**, 407 (2011).
 - [7] Y. Aharonov and A. Casher, Ground state of a spin- $\frac{1}{2}$ charged particle in a two-dimensional magnetic field, *Phys. Rev. A* **19**, 2461 (1979).
 - [8] V. M. Pereira, J. M. B. Lopes dos Santos, and A. H. Castro Neto, Modeling disorder in graphene, *Phys. Rev. B* **77**, 115109 (2008).
 - [9] M. Inui, S. A. Trugman, and E. Abrahams, Unusual properties of midband states in systems with off-diagonal disorder, *Phys. Rev. B* **49**, 3190 (1994).
 - [10] F. Evers and A. D. Mirlin, Anderson Transitions, *Rev. Mod. Phys.* **80**, 1355 (2008).
 - [11] H. C. Longuet Higgins, Some studies in molecular orbital theory I. Resonance structures and molecular orbitals in unsaturated hydrocarbons, *J. Chem. Phys.* **18**, 265 (1950).
 - [12] We here follow the nomenclature of the original literature. The maximum number of placeable double bonds N_{db} may also be thought of as the maximum number of nonadjacent edges β . Alternatively, N_{db} relates to the total number of sites N and the maximum number of nonadjacent sites α via $N_{db} = N - \alpha$.
 - [13] S. Fajtlowicz, P. E. John, and H. Sachs, On maximum matchings and eigenvalues of benzenoid graphs, *Croat. Chem. Acta* **78**, 195 (2005).
 - [14] We here follow the mathematical literature [37] deviating from the one used by Longuet-Higgins [11] and Bonfanti *et al.* [21] who refer only to those modes as supernumerary that exist in excess of $N - 2N_{db}$.
 - [15] N. Weik, Diploma thesis, Karlsruhe Institute of Technology, 2013.
 - [16] V. Häfner, J. Schindler, N. Weik, T. Mayer, S. Balakrishnan, R. Narayanan, S. Bera, and F. Evers, Density of States in Graphene with Vacancies: Midgap Power Law and Frozen Multifractality, *Phys. Rev. Lett.* **113**, 186802 (2014).
 - [17] T. O. Wehling, S. Yuan, A. I. Lichtenstein, A. K. Geim, and M. I. Katsnelson, Resonant Scattering by Realistic Impurities in Graphene, *Phys. Rev. Lett.* **105**, 056802 (2010).
 - [18] O. V. Yazyev, Emergence of magnetism in graphene materials and nanostructures, *Rep. Prog. Phys.* **73**, 056501 (2010).
 - [19] W. L. Wang, O. V. Yazyev, S. Meng, and E. Kaxiras, Topological Frustration in Graphene Nanoflakes: Magnetic Order and Spin Logic Devices, *Phys. Rev. Lett.* **102**, 157201 (2009).

- [20] M. Kinza, J. Ortloff, and C. Honerkamp, Effective low-energy Hamiltonians for interacting nanostructures, *Phys. Rev. B* **82**, 155430 (2010).
- [21] M. Bonfanti, S. Casolo, G. F. Tantardini, A. Ponti, and R. Martinazzo, A few simple rules governing hydrogenation of graphene dots, *J. Chem. Phys.* **135**, 164701 (2011).
- [22] A. Ferreira and E. R. Mucciolo, Critical Delocalization of Chiral Zero Energy Modes in Graphene, *Phys. Rev. Lett.* **115**, 106601 (2015).
- [23] R. Gade and F. Wegner, The $n = 0$ replica limit of $U(n)$ and $U(n)SO(n)$ models, *Nucl. Phys. B* **360**, 213 (1991).
- [24] R. Gade, Anderson localization for sublattice models, *Nucl. Phys. B* **398**, 499 (1993).
- [25] S. Ryu, C. Mudry, C. Y. Hou, and C. Chamon, Masses in graphenelike two-dimensional electronic systems: Topological defects in order parameters and their fractional exchange statistics, *Phys. Rev. B* **80**, 205319 (2009).
- [26] A. D. Mirlin, F. Evers, I. V. Gornyi, and P. M. Ostrovsky, Anderson transitions: Criticality, symmetries and topologies, *Int. J. Mod. Phys. B* **24**, 1577 (2010).
- [27] P. M. Ostrovsky, M. Titov, S. Bera, I. V. Gornyi, and A. D. Mirlin, Diffusion and Criticality in Undoped Graphene with Resonant Scatterers, *Phys. Rev. Lett.* **105**, 266803 (2010).
- [28] E. J. König, P. M. Ostrovsky, I. V. Protopopov, and A. D. Mirlin, Metal-insulator transition in two-dimensional random fermion systems of chiral symmetry classes, *Phys. Rev. B* **85**, 195130 (2012).
- [29] S. Ryu, C. Mudry, A. W. W. Ludwig, and A. Furusaki, Global phase diagram of two-dimensional Dirac fermions in random potentials, *Phys. Rev. B* **85**, 235115 (2012).
- [30] P. M. Ostrovsky, I. V. Protopopov, E. J. König, I. V. Gornyi, A. D. Mirlin, and M. A. Skvortsov, Density of States in a Two-Dimensional Chiral Metal with Vacancies, *Phys. Rev. Lett.* **113**, 186803 (2014).
- [31] S. Sanyal, K. Damle, and O. I. Motrunich, Vacancy-induced low-energy states in undoped graphene, [arXiv:1602.09085](https://arxiv.org/abs/1602.09085).
- [32] M. B. Isichenko, Percolation, Statistical Topography, and Transport in Random Media, *Rev. Mod. Phys.* **64**, 961 (1992).
- [33] D. Stauffer and A. Aharony, *Introduction To Percolation Theory* (Taylor & Francis, London, 1994).
- [34] J. Hoshen and R. Kopelman, Percolation and cluster distribution. I. Cluster multiple labeling technique and critical concentration algorithm, *Phys. Rev. B* **14**, 3438 (1976).
- [35] J. M. Hammersley, D. C. Handscomb, and G. Weiss, Monte Carlo methods, *Phys. Today* **18**, 55 (1965).
- [36] K. Nakada, M. Fujita, G. Dresselhaus, and M. S. Dresselhaus, Edge state in graphene ribbons: Nanometer size effect and edge shape dependence, *Phys. Rev. B* **54**, 17954 (1996).
- [37] M. J. Rigby and R. B. Mallion, On the eigenvalues and eigenvectors of certain finite, vertex-weighted, bipartite graphs, *J. Comb. Theory, Ser. B* **27**, 122 (1979).