Counting the number of excited states in organic semiconductor systems using topology

Michael J. Catanzaro,[1] Tian Shi,[2] Sergei Tretiak,[3,a] and Vladimir Y. Chernyak[1,2,b]

[1]*Department of Mathematics, Wayne State University, 656 W. Kirby, Detroit, Michigan 48202, USA*
[2]*Department of Chemistry, Wayne State University, 5101 Cass Ave., Detroit, Michigan 48202, USA*
[3]*Theoretical Division, Center for Nonlinear Studies, and Center for Integrated Nanotechnologies,*
Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA

Exciton scattering theory attributes excited electronic states to standing waves in quasi-one-dimensional molecular materials by assuming a quasi-particle picture of optical excitations. The quasi-particle properties at branching centers are described by the corresponding scattering matrices. Here, we identify the topological invariant of a scattering center, referred to as its winding number, and apply topological intersection theory to count the number of quantum states in a quasi-one-dimensional system. © *2015 AIP Publishing LLC.* [http://dx.doi.org/10.1063/1.4908560]

I. INTRODUCTION

A quantum system that resides on a network/graph forms an important generalization of a particle in a box problem, where the quantum eigenstates can be attributed to standing waves. In our earlier work, we have developed a multi-scale Exciton Scattering (ES) approach for photo-excitations in large, conjugated molecules. Excitations are described as quasi-particles (excitons) that move along the linear segments, getting scattered at molecular termini, joints, and branching centers, represented by graph edges and vertices, respectively.[1–5] Excited states can then be calculated by solving the ES equations [see Eq. (13)], once the ES parameters, namely, exciton dispersion $\omega(k)$ for graph edges and scattering matrices $\Gamma_a(k)$ for graph vertices a, are known. Both ingredients $\omega(k)$ and $\Gamma_a(k)$ can be extracted from the reference quantum chemistry calculations of relatively small molecular fragments with reasonably low computational cost.[3,5] Thus, extracting the ES parameters and solving the ES equations are two fundamental steps in the multi-scale modeling of electronic excitations in branched, organic semi-conductors. The ES approach is asymptotically exact for long enough linear segments[2,5] and shows an excellent agreement with reference quantum chemistry calculations[4,5] even in the case of short linear segments. Our previous studies have demonstrated accurate calculations of excited state transition energies[4,5] and optical spectra[6] in very large, low-dimensional molecular systems with negligible computational cost using this technique. Moreover, this framework enables an efficient characterization of excited-state electronic structure modifications due to different branching patterns and donor/acceptor substitutions.[7,8] We reiterate that all information on the system Hamiltonian and further approximations (e.g., the quantum chemistry method involved, and basis sets) is fully contained in the tabulated ES parameters [$\omega(k)$ and $\Gamma_a(k)$], so once they are extracted, there is no need to specify the aforementioned Hamiltonian, as well as supporting information.

It is worth noting that the ES approach expresses electronic excitations in branched, conjugated molecules in terms of excited states in infinite polymers with perfect geometry, and the scattering properties of molecular joints. Due to discrete translational symmetry, electronic excitations in an infinite polymer chain possess a good quantum number, referred to as quasimomentum $k \in [-\pi, \pi]$, that resides in a 1 dimensional Brillouin zone. This Brillouin zone is represented by a circle, obtained from the segment $[-\pi, \pi]$ by implementing periodic boundary conditions. The continuous spectrum of electronic excitations in an infinite polymer chain is, therefore, represented by exciton bands, each characterized by its dispersion $\omega(k)$, i.e., the dependence of the excitation energy ω on its quasimomentum k. The ES methodology implements a well-known particle in a box concept by viewing a branched, conjugated molecule as a quasi-one-dimensional box, where the linear segments are represented by sub-boxes, while the scattering matrices $\Gamma_a(k)$ play the role of the proper boundary conditions obtained from this microscopic approach. In a finite branched structure, standing waves are formed due to interference of waves from quasi-particle (exciton) scattering at molecular joints and termini. This results in a discrete spectrum of electronic excitations, in full analogy with the standard particle in a box problem. The exciton energies for a finite molecule are obtained by solving a generalized spectral problem in an $n = 2n_s$-dimensional vector space, with n_s being the number of linear segments (graph edges), referred to as the system of the ES equations.[2] In general terms, the ES approach splits the energy range into spectral regions, spanned by different exciton bands, and further identifies the exciton spectra within the above regions by solving the ES equations. Obviously, the exciton bands are properties of infinite polymers, or stated differently, properties of the repeat unit, whereas the excitation energies also depend on the scattering matrices of the molecular joints, and the molecular

a)Electronic address: serg@lanl.gov
b)Electronic address: chernyak@chem.wayne.edu

structure, i.e., the linear segment lengths L_α and the way the segments are connected via the joints.

The spectrum of electronic excitations is a key fingerprint that defines molecular photoinduced dynamics, including energy transfer and relaxation processes. Naturally, an important property of the excited state manifold determined by a specific exciton band is the total number N of excited states in the given spectral region. According to the ES theory, N is determined by the spectrum $\omega(k)$ of the relevant band, the scattering matrices $\Gamma_a(k)$ of the molecular joints and termini, the linear segment lengths L_α (that can be measured as the numbers of repeat units within a segment), and finally the molecular topology (i.e., the way the linear segments are connected). The molecular topology can be formally described by a graph, whose vertices and edges represent the joints (including termini) and linear segments, respectively. Topologically speaking, a joint a is determined by its degree r_a. Equivalently, all termini (degree one), all double joints (degree two), all triple joints (degree three), and so on, are topologically identical; the differences are encoded in their scattering matrices. A natural question arises: Is there a simple relation that expresses the number N of excitations within the spectral region of a given exciton band in terms of the inputs of the ES approach, i.e., the graph that describes the topological structure of a branched molecule, the dispersion $\omega(k)$, scattering matrices $\Gamma_a(k)$, and the segment lengths L_α, where a and α stand for the graph vertices and edges, respectively?

In our previous work, we made the first attempt to address the above question within the ES approach. We introduced a topological invariant (coining winding number) to count the total number N of excited states within the spectral region of a single exciton band in symmetric, conjugated molecules.[9] The counting was achieved by associating to a scattering matrix $\Gamma_a(k)$ of a highly symmetric molecular scattering center, an integer-valued topological invariant Q_a, coined topological charge. Further, we implemented topological intersection theory in its very elementary, albeit very intuitive form, resulting in a lower bound,

$$N \geq \sum_\alpha L_\alpha + \sum_a Q_a. \tag{1}$$

We have also demonstrated that for molecules with long enough linear segments, the bound in Eq. (1) becomes tight, i.e., the inequality turns into an equality, so that the count is solely dependent on the number of repeat units and the scattering matrices of joints. The obtained result allows for a natural interpretation: In a finite-size molecule, the number of electronic excitations within the spectral region of the exciton band is given by the total number of repeat units in the linear segments plus the sum of topological charges of the scattering centers. In other words, each repeat unit provides a state to the full count, whereas the number of states provided by a scattering center (a molecular terminal or a joint) is defined by its topological charge Q_a. The latter is fully determined by the topological properties of the relevant scattering matrix $\Gamma_a(k)$. We applied this method to investigate the scattering matrix of the "X" joint of phenylacetylene (PA) based molecules, and successfully predicted one additional resonance state with

energy inside the band and two bound excited states with energy outside the band contributed from the joint.[9]

This methodology has already found applications in building effective tight binding models for electronic excitations in low-dimensional conjugated systems.[10] Since the number of lattice sites for a scattering center/repeat unit in these models depends explicitly on the topological invariant, we have a systematic way to construct the lattice model, and therefore, the Frenkel type exciton Hamiltonian. For example, every repeat unit of PA based molecules can naturally be represented by a lattice site, since each of them contributes exactly one excited state to an exciton band. Another nontrivial example arises from scattering at *ortho* and *meta* joints contributing one and zero additional states, translating into one and zero lattice sites for them, respectively. The number of lattice sites for the "X" joint is one, since it provides one additional state. Such assignment of lattice sites guarantees the correctness and accuracy of the tight binding models.

It is important to note that the results on the counting, discussed above, were ultimately dependent on the high symmetry of the molecules under study. Not only should the joints possess high symmetry, but also should the latter be preserved by the actual molecule. Stated differently, the aforementioned results are valid only for linear oligomers and for molecules where a symmetrical double, triple, or quadruple joint connects linear segments of equal length. The latter condition is necessary for the approach we used in order to maintain the high symmetry of the system. This limitation is imposed by restrictive assumptions: The main result (Eq. (1)), has actually been derived for linear oligomers, followed by extending it to highly symmetric molecules, listed above, using the classification of electronic excitations by their symmetry, thus, mapping the counting problem in a symmetric molecule to its linear segment counterpart. Therefore, the approach presented in Ref. 9 is not extendable to an arbitrary molecular topology. Moreover, it provides no insights into how the notion of topological charge could be extended to the case of an arbitrary joint, or if such extension is possible at all.

In this manuscript, we address the question of counting electronic excitations within the spectral region of a single exciton band, explicitly formulated earlier, in full generality, i.e., for a branched molecule described by an arbitrary graph, as well as general type scattering centers, not necessarily possessing a high degree of symmetry. Specifically, we introduce a topological invariant by identifying the winding number for an *arbitrary* scattering center and relate the segment lengths L_α along with the winding numbers l_a of the vertices to the total number of excited states N in a single exciton band. We obtain an explicit lower bound for N in terms of L_α and l_a [Eqs. (16) and (20)]. This is achieved by formulating the ES equations as an intersection problem, followed by applying topological intersection theory, which provides simple expressions for the *intersection index*, the latter bounding the number of solutions of the ES equations. Furthermore, this bound becomes exact when the molecular arms are long enough. We also demonstrate the formulated concepts by analyzing excited state structure in branched conjugated polyfluorenes. We emphasize that the theory

presented here applies generally to any quasi-one-dimensional quantum system, as our results are based on the following generic properties: (i) unitarity of scattering processes, (ii) phase change from propagation along linear segments, and (iii) quasi-momenta residing in a one-dimensional Brillouin zone.

The manuscript is organized as follows. In Sec. II, we introduce the approach we have taken, including a simple and intuitive description of topological intersection theory, with an emphasis on its application to the problem under study (Sec. II B). We also describe in some detail the main results presented in the manuscript (Sec. II A), as well as intuitive arguments that stand behind the formal derivations. We present a simple picture of the winding number (Sec. II C), together with intuitive arguments in support of the index theorem. Moreover, we rationalize a relation between solutions of the ES equation and actual excitons (Sec. II D). The reader not interested in the details of these derivations can then bypass this section, jumping directly to Sec. IV, where the applications are presented. Sec. III contains a detailed derivation of our main results, consisting of several subsections, and focuses on specific steps of the derivation. In Subsection III A, we rewrite the ES equations as an eigenvalue problem [Eq. (14)], simplifying their description and allowing for a clear formulation in terms of intersection theory. Using this representation, we relate the number of solutions to the number of excitons in Eq. (16). In Subsection III B, we introduce a topological invariant of a scattering vertex, as well as the local intersection indices. This allows us to relate the number of solutions to local intersection indices via the index theorem [Eq. (18)]. We then formulate counting of excitons in terms of intersection theory in Subsection III C, along with sketching a proof of the index theorem. In Subsection III D, we apply perturbation theory to determine the situations in which our approximation becomes exact. Finally, in Sec. IV, we exemplify the presented theory to a class of polyflourene-based molecules. A brief summary of the results, as well as a list of problems to be addressed in the future is provided in Sec. V.

II. METHODOLOGY AND STATEMENTS OF THE RESULTS

In this section, we present our main results on counting the number of electronic excitations in a branched, conjugated molecule within the spectral region of a single exciton band, and introduce the key structures involved in the main statements. We further present clear, intuitive arguments that stand behind the formal derivations, doing so on a conceptual level. In this way, a reader who is interested in the results themselves, as well as the concepts that stand behind them, can skip the forthcoming Sec. III, where the details of the derivation are presented, and jump directly to Sec. IV, where an application of our approach is presented. We also introduce the essential concepts behind topological intersection theory, and how it can be used to solve systems of equations. This is done by providing a simple example of intersection theory in terms of single variable calculus, followed by showing how it can be adapted to solve the ES equations.

A. ES equations and the main results

Within the ES approach, a branched conjugated molecule is described by the following data. (i) The first ingredient is the graph that determines the molecule's topological structure. Generally, a graph is a collection of vertices connected to one another by edges. Our graph models the conjugated molecule under study, by assigning a vertex to each scattering center, and an edge to each molecular linear segment. The graph is equipped with the following additional data: (ii) the number of repeat units L_α in a linear segment α, with $\alpha = 1, \ldots, n_s$, often referred to as the segment lengths, (iii) the exciton spectrum (or, equivalently, dispersion) $\omega(k)$, i.e., the dependence of the exciton energy ω on its momentum k in an infinitely long polymer with the same repeat unit, and (iv) the quasimomentum-dependent exciton scattering matrix $\Gamma_a(k)$ of vertex a, where a varies over the set of molecule scattering centers (including the termini).

The main exciton counting result is formulated in terms of the winding numbers l_a associated with the molecular scattering centers (graph vertices) labeled by a and two additional integers d_k, for $k = 0, \pi$, referred to as bandedge indices. A vertex winding number l_a is an integer number completely determined by the scattering matrix $\Gamma_a(k)$ of a vertex a. The bandedge indices d_k are integer numbers that reflect the properties of the bandedge (i.e., $k = 0, \pi$) solutions of the ES equations, i.e., the relation between the number of solutions and the actual bandedge excitons. The quantities l_a and d_k will be introduced later in this section.

The main counting result is the lower bound,

$$N \geq \sum_\alpha L_\alpha + \frac{1}{2} \sum_a l_a + \frac{d_0 + d_\pi}{2}, \qquad (2)$$

for the number N of electronic excitations within the spectral region of a single exciton band, combined with the statement that the bound becomes tight (i.e., the inequality turns into an equality) for a molecule with long enough linear segments. There is a very important particular case, which turns out to represent the generic situation: for all molecular vertices whose scattering matrices we have extracted from numerical data in our previous studies, as well as for all scattering matrices calculated analytically within the framework of lattice models (except for some special values of there parameters), we find $\Gamma_a(0) = \Gamma_a(\pi) = -1$. In this case, as will be shown below, we have $d_0 = d_\pi = -n_s$; therefore, we introduce the *topological charge* of any vertex a as

$$Q_a = \frac{l_a - r_a}{2} \qquad (3)$$

(we reiterate that r_a is the vertex degree, i.e., the number of molecule linear segments attached to the scattering center). We can further make use of an obvious relation $\sum_a r_a = 2n_s$, which is true in an graph, and recast Eq. (2) in the form of Eq. (1). Hence, we can interpret the topological charge Q_a of a vertex as the number of additional states the corresponding scattering center brings to the spectral region of a given exciton band.

We are now in a position to explain what kind of observations made it possible to come up with such a simple and universal expression for counting the number of electronic

excitations within a properly defined spectral region for an arbitrary branched conjugated molecule with long enough linear segments. The first step is bringing in the ES approach; as of today, this step is not too surprising, since the ES approach was introduced almost a decade ago, and is well understood by now. Still, it provides a vast simplification of the problem, since (a) it allows one to summarize all the relevant properties of electron interactions and correlations in a branched conjugated molecule, as well as the details of the system Hamiltonian, by introducing relatively simple objects, namely, the excitons spectrum $\omega(k)$ and scattering matrices $\Gamma_a(k)$, and (b) the problem of identification of electronic excitations in a branched structure is reduced to solving the system of ES equations that can be described as a generalized spectral problem in a vector space of a relatively low dimension $n = 2n_s$.

The second and most non-trivial step, is reducing the problem of solving the ES equations to an *intersection problem*. This is achieved by representing the ES equations in a form,

$$\tilde{\Gamma}(k)\psi = \psi, \qquad (4)$$

where $\tilde{\Gamma}(k)$ is a quasimomentum dependent $2n_s \times 2n_s$ scattering matrix, whereas ψ is a $2n_s$-dimensional vector of amplitudes of the incoming exciton waves at the scattering centers. The matrix $\tilde{\Gamma}(k)$ can be represented in a natural way as a product of three matrices [for an explicit expression see Eq. (14)]. A reader, interested in a detailed derivation of Eq. (4) can find them in Sec. III A; here we present some simple, intuitive arguments. Note that here we use an abbreviated notation ψ instead of $\psi^{(+)}$, used in Sec. III A. The matrix $\Gamma(k)$ is block-diagonal, with the blocks represented by $\Gamma_a(k)$.[20] The matrix $\Gamma(k)$ describes scattering at molecular vertices by transforming the $2n_s$-dimensional vector of incoming exciton amplitudes to the vector $\Gamma(k)\psi$ of the outgoing counterparts. The diagonal matrix $\Lambda(k)$ with the eigenvalues $\exp(ikL_\alpha)$ describes the exciton propagation along the linear segments. The block-diagonal matrix P, with 2×2 blocks, accounts for the fact that an outgoing wave, after propagation along a linear segment, turns into an incoming wave at the opposite end of the latter. Obviously, all blocks are identical and k-independent. It is easy to see now that Eq. (4) is nothing else than a consistency condition for a standing wave: scattering at the vertices, followed by propagation along the segments, should result in the same wave. Stated a bit more formally, Eq. (4) is an alternative way of recasting the ES equations in one of its standard forms [see Eq. (13)].

The advantage of the representation, given by Eq. (4), is that the solutions of the ES equations can be considered as intersections. Indeed, Eq. (4) states that for a certain value k of quasimomentum, the unitary $n \times n$ (we reiterate that $n = 2n_s$) matrix $\tilde{\Gamma}(k)$ has an eigenvalue equal to 1. Therefore, the solutions of the ES equations correspond to the intersections of the closed curve, $\tilde{\Gamma}(k)$, parameterized by the quasimomentum k, with the subspace of $U(n)$ represented by those unitary matrices that have an eigenvalue equal to 1, with the intersections occurring in the space $U(n)$ of all unitary $n \times n$ matrices. This subspace of $U(n)$, whose matrices have an eigenvalue equal to 1 is, hereafter referred to as

$D_1 U(n)$. The remaining steps would not be too surprising to someone familiar with the basics of algebraic topology: as briefly stated in Sec. I, we apply topological intersection theory to compute a simpler quantity, namely, the intersection index, which provides a lower bound on the number of intersections (i.e., the number of the solutions to the ES equation). Next, we relate the number of ES equation solutions to the number N of electronic excitations (excitons), which involves dealing with over- and under-counting, and finally, show the tightness of the bound when the linear segments are long enough.

It should be noted that the standard topological intersection theory is not directly applicable to our case, and a proper extension/generalization should be developed. In Sec. II B, we present in a simple and intuitive way the basic concepts of topological intersection theory, rationalize why it is not directly applicable to our case, and develop a proper extension.

B. Topological intersection theory and its generalization

The main computational tool we have used in this manuscript to count the number N of electronic excitations within the spectral region of an exciton band is topological intersection theory. In this formulation, each equation of a system is interpreted as, generally speaking, some multi-dimensional surface, or speaking more mathematically, a manifold. In this setting, a solution to a system of equations manifests itself as an intersection of the aforementioned manifolds. Finding these intersections (and specifically, the number of such) can be a difficult task due to the generic nature of the equations. An easier approximation to compute is a topological intersection index $q = \sum_j q_j$, consisting of a sum over all intersections of local intersection indices $q_j = \pm 1$, i.e., the intersections are weighted with a ± 1 sign factor. The topological nature of this index implies that not only is it fixed under deformations of the curves but also that there should (and do) exist expressions for solutions (in our case, excitons) in purely topological terms. Denoting the total number of intersections by m (i.e., solutions of a system of equations), we obviously have $m \geq q$, so that topological intersection theory provides a lower bound for the number m of solutions in terms of the topological intersection index q, which is much easier to compute.

As an example, consider the family of cubics $z(t) = x(t) + 8y^2 + 8y^3$ together with $4z = -x^2 - y^2$, where x is parameterized by t, shown in Fig. 1. We see that for different values of t, this family of curves will have a different number of intersections with the surface, and therefore, the system will have a different number of solutions. However, when intersections are weighted with $+1$ or -1 (depending on orientations at crossings), we find that the sum of these weights is always 1 (the $t = 0$ case is not stable, since it bifurcates into the other cases under slight perturbations).

At the end of Sec. II A, we identified the solutions of the ES equations with the intersections of a closed curve, defined by $\tilde{\Gamma}(k)$ in the space of the unitary group $U(n)$, with the subspace $D_1(U(n))$ of unitary matrices with at least one unit eigenvalue. There is, however, an obstacle to straightforward application of standard topological intersection theory: the
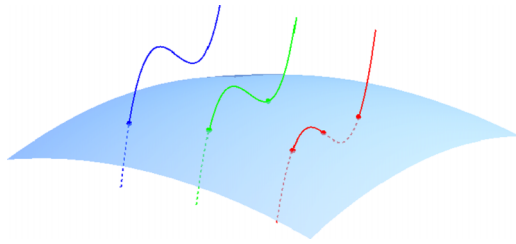
FIG. 1. Solving systems of equations by means of intersection theory. The parametric system of equations $z = x + 8y^2 + 8y^3$, $x = t$, $4z = -x^2 - y^2$ is solved. The first two equations describe a curve ($t = -1$: blue, $t = 0$: green, and $t = 1$: red), while the third equation describes a surface. We weight intersections with $+1$ if the curve comes from underneath or -1 if the curve comes from above, yielding an intersection index equal to $+1$. This appreciates the fact that we will have at least one solution in general.

latter is valid in the case of intersecting *manifolds*, namely, spaces that locally look like open neighborhoods of euclidean spaces $\mathbb{R}^n$. In our case, while the curve $\tilde{\Gamma}(k)$ is a manifold, the subspace $D_1(U(n))$ is not, the latter statement being due to the points, represented by unitary matrices with a degenerate unit eigenvalues. Stated differently, the complexity of the situation originates from degenerate solutions of the ES equations.

To treat our situation, we extend ordinary intersection theory by generalizing a notion of the local intersection index q_j, associated with an intersection point, so that it admits integer values, rather than just ± 1. The generalized version of q_j is defined as follows. Consider an intersection (equivalently a solution to the ES equation) at $k = k_j$, so that $\tilde{\Gamma}(k_j)$ has exactly m_j unit eigenvalues; hereafter, we refer to m_j as the multiplicity of an intersection. Recall that all eigenvalues belong to the circle $|z| = 1$ in the complex plane. In the case of an isolated intersection (which we always have since $\tilde{\Gamma}(k)$ is an analytic function of k), when $k \neq k_j$ is close to k_j, $\tilde{\Gamma}(k)$ does not have unit eigenvalues. So, when k goes through k_j, starting with $k < k_j$ and ending up with $k > k_j$, there is a certain number of eigenvalues, denoted q_j, that pass through the point $z = 1$, moving from below to above, the rest being reflected from $z = 1$. We refer to q_j as the local intersection index; note that q_j can be positive, negative, or zero, and obviously, $-m_j \leq q_j \leq m_j$. A more formal definition of the local intersection index is given in Sec. III B. We further define the number m of the ES equation solutions (with the degeneracy accounted for) and the global intersection index,

$$m = \sum_j m_j, \quad q = \sum_j q_j, \quad (5)$$

so that, obviously, $m \geq q$, and the intersection index provides a lower bound for m.

One might ask a question: Why was the local intersection index defined as it was? The answer is with the above definition, the global intersection index q becomes a topological invariant, i.e., it depends only on the topological class of $\tilde{\Gamma}(k)$. In other words, it is not changed upon deformations of the curve $\tilde{\Gamma}(k)$ and, therefore, is easily computable. Topological invariance of q is established via the index theorem [Eq. (11)].

C. Winding number associated with a scattering matrix and the index theorem

We start with introducing our main computational tool, namely, the winding number associated with a scattering matrix, that will allow us to derive the index theorem that relates the topological intersection index, introduced in Sec. II B, to the winding number of the scattering matrix $\tilde{\Gamma}(k)$, as well as explicitly compute the latter. Hereafter, we use the term scattering matrix for any quasimomentum k dependent $n \times n$ unitary matrix $f(k)$, where n can be arbitrary. Examples of scattering matrices are $r_a \times r_a$ scattering matrices $\Gamma_a(k)$ at the molecular vertices, as well as $2n_s \times 2n_s$ matrices $\tilde{\Gamma}(k)$, $\Gamma(k)$, $\Lambda(k)$, and P, introduced in Sec. II A, as well as, in a more formal way in Sec. III A. Note that P is actually k-independent. The integer-valued winding number $w(f)$ is defined by making use of the fact that the determinant of a unitary matrix is a unimodular complex number, i.e., we can represent $\det f(k) = e^{i\varphi(k)}$,

$$w(f) = \int_{-\pi}^{\pi} \frac{dk}{2\pi i} (\det f(k))^{-1} \frac{d}{dk} \det f(k)$$
$$= \int_{-\pi}^{\pi} \frac{dk}{2\pi i} \frac{d \ln (\det f(k))}{dk} = \int_{-\pi}^{\pi} \frac{dk}{2\pi} \frac{d\varphi(k)}{dk}. \quad (6)$$

Since φ is a phase (angular) variable, i.e., it is defined up to an integer multiple of 2π, the integral in the r.h.s. of Eq. (6) is an non-zero integer number; stated differently, it reflects the multi-valued nature of the logarithm, as a function of a complex argument.

Explicit computation of the generalized intersection index q, introduced in Sec. II B, is possible due to the following algebraic property of the winding number. Consider scattering matrices $f(k)$, $g(k)$, and $h(k)$ of the sizes $n \times n$, $n \times n$, and $m \times m$, respectively. Denote by $f \cdot g$ and $f \oplus h$ the $n \times n$ matrix product of f and g, and the $(n + m) \times (n + m)$ block-diagonal matrix with the blocks, given by f and h, respectively. The well-known multiplicative properties of the determinant, namely, $\det(f \cdot g) = (\det f)(\det g)$ and $\det(f \oplus h) = (\det f)(\det h)$, immediately imply the desired algebraic properties of the winding number,

$$w(f \cdot g) = w(f) + w(g), \quad w(f \oplus h) = w(f) + w(h). \quad (7)$$
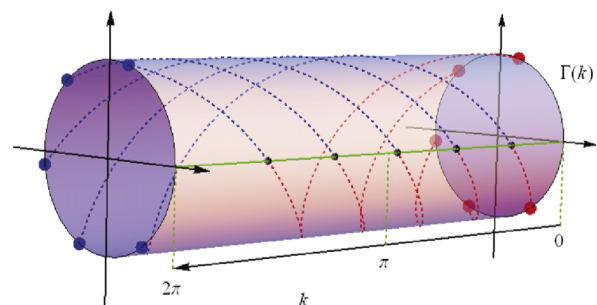

FIG. 2. Two equivalent interpretations of the flux that result in the index theorem. The eigenvalues form a periodic trajectory, evolving over time (k), where each eigenvalue has charge $+1$. To calculate the total flux, one can count the full number of rotations of all particles, which amounts to calculating $w(\tilde{\Gamma})$. Alternatively, one could calculate the total charge passing through any cross-section, e.g., $z = 1$ (green line), yielding $q = \sum q_j$. The comparison of the two interpretations gives the index theorem [Eq. (11)].

The algebraic properties [Eq. (7)] allow for a concise computation of the winding number $w(\tilde{\Gamma})$,

$$w(\tilde{\Gamma}) = w(\Lambda) + w(\Gamma) + w(P) = 2\sum_{\alpha} L_{\alpha} + \sum_{a} w(\Gamma_a). \quad (8)$$

In deriving Eq. (8), we have used the fact that P is k-independent, which due to Eq. (6), implies $w(P) = 0$. A direct calculation shows $w(\Lambda_{\alpha}) = L_{\alpha}$ for a 1×1 scattering matrix $\Lambda_{\alpha}(k) = e^{ikL_{\alpha}}$. Recalling the definition $l_a = w(\Gamma_a)$, we arrive at

$$w(\tilde{\Gamma}) = 2\sum_{\alpha} L_{\alpha} + \sum_{a} l_a, \quad l_a \equiv w(\Gamma_a). \quad (9)$$

The lower bound,

$$m = \sum_{j} m_j \geq \sum_{j} q_j = q = 2\sum_{\alpha} L_{\alpha} + \sum_{a} l_a, \quad (10)$$

for the number of solutions to the ES equations, which becomes tight for long enough segments, is obtained by combining the explicit expression given by Eq. (9) with the index theorem,

$$q = \sum_{j} q_j = w(\tilde{\Gamma}). \quad (11)$$

Note that the tightness of the bound [Eq. (10)] follows from the statement that $q_j = m_j$ for long enough segments. We comment on the above statement in Sec. II D and provide a more formal derivation in Sec. III D.

We conclude this subsection with simple intuitive arguments in support of the index theorem [Eq. (11)], a sketch of a formal proof, for a reader, interested in details, is presented in Sec. III C.

The intuition relies on two equivalent representations of charged particles current/flux (see Fig. 2). Consider the eigenvalues $\lambda_1, \ldots, \lambda_n$ (with possible degeneracy properly accounted for) of $\tilde{\Gamma}$ as a system of n undistinguishable particles with charge $+1$ that reside in a unit circle $|z| = 1$ in the complex plane $\mathbb{C}$, and let us think of k as time. Then, the scattering matrix $\tilde{\Gamma}(k)$ provides a periodic trajectory of our n-particle system. To calculate the total charge flux in the system over a period, we count the total number of full rotations of all particles in the system. Since $\det\tilde{\Gamma}(k) = \prod_{i=1}^{n} \lambda_i(k)$ with $\lambda_i(k)$ being the eigenvalues of $\tilde{\Gamma}(k)$, the total flux, due to Eq. (7) is given by $q = \sum_{j} q_j = \sum_{j} w(\lambda_j) = w(\tilde{\Gamma})$. On the other hand, the total charge flux q is given by the total charge that went through any cross section, e.g., the total charge traveled through the point $z = 1$ of the circle. For the second interpretation, the charge travels through $z = 1$ at times $k = k_j$, and, according to the definition of the local intersection index, given in Sec. II B, in the amount of q_j, so that $q = \sum_{j} q_j$. The index theorem is a reflection of the fact that these two interpretations of charge flux are the same (taking orientation into account).

D. From solutions of the ES equation to excitons and the long-arm limit

In this section, we finalize our computation by relating the number of solutions of the ES equations m to the actual number of electronic excitations N within the spectral region defined by the exciton band, which results in the lower bound for N [Eq. (2)]. We further comment on the tightness of the bound for the case of long enough segments.

We start by noting that any exciton corresponds to a solution of the ES equation. On the other hand, if we have a solution of the ES equation with $k = k_j$, we should expect to have a solution at $k = -k_j$, corresponding to the same exciton. This is evident since the latter is represented by a standing wave, so that we have two plane waves at each linear segment, propagating in the opposite directions. More precisely, these two plane waves have opposite quasimomenta, due to time-reversal symmetry. The above statement is formalized in Sec. III A, where we demonstrate that if ψ is a solution of the ES equation [Eq. (4)] for $k = k_j$, then $\Gamma(k_j)\psi$ is a solution with $k = -k_j$, which corresponds to the same exciton. This implies that the solutions of the ES equation with $k \neq 0$, π count each exciton twice. The situation when $k = 0$, π is a just a bit more complicated. In particular, due to $k = -k$ for $k = 0$, π, the two plane waves on each segment are actually one wave. As demonstrated in Sec. III A, for $k = 0$, π, the vector space of solutions to the ES equation is invariant with respect to the action of $\Gamma(k)$, and $(\Gamma(k))^2 = 1$. This implies that the bandedge, i.e., $k = 0$, π, solutions have a well-defined parity, i.e., even and odd solutions with $\Gamma(k)\psi = \psi$, and $\Gamma(k)\psi = -\psi$. Denote the number of even and odd solutions by d_k^+ and d_k^-, respectively, and further introduce the bandedge indices $d_k = d_k^+ - d_k^-$. It is easy to understand that the odd solutions are unphysical, since they correspond to a complete zero standing wave (the plane waves with $\pm k$ cancel each other), whereas even solutions correspond to actual excitons. Therefore, we arrive at a simple relation,

$$N = \frac{m + d_0 + d_{\pi}}{2} \quad (12)$$

between the number of electronic excitations N within the spectral region of an exciton band and the number of solutions to the ES equations m. The main counting result [Eq. (2)] is readily obtained by simply combining Eqs. (10) and (12).

The tightness of the bound for long enough linear segments is due to the fact that in this limit, $q_j = m_j$. According to the definition of the local intersection index, this means that if we have an m_j-fold degenerate solution of the ES equations for $k = k_j$, i.e., $\tilde{\Gamma}(k_j)$ has exactly m_j eigenvalues equal to 1, they all move above and below the point $z = 1$ on their circle of residence, when k becomes slightly larger and slightly smaller than k_j, respectively. This is demonstrated in Sec. III D by applying standard quantum mechanical perturbation theory.

III. DETAILS OF THE DERIVATION

In this section, we present the details of the derivation of our results on counting electronic excitations in branched conjugated molecules.

A. The ES equations

We start by considering a quasi-one-dimensional quantum system on a graph (a collection of vertices connected by edges), whose states (excitons) are quasiparticles.[1,11,12] The latter are represented by plane waves that reside on the linear

segments (graph edges α), getting scattered at the scattering centers (graph vertices a), which leads to the formation of standing waves that describe discrete quantum states of the system. We consider the case when the linear segments possess discrete translational symmetry, which is broken by the finite segment length only. Therefore, an exciton in an infinite chain has a well-defined quantum number, namely, quasimomentum $k \in [-\pi, \pi]$ that resides in the Brillouin zone, and the exciton properties are described by its spectrum $\omega(k)$.[2,3,5] The time-reversal symmetry implies $\omega(-k) = \omega(k)$. We also assume that for all ω_0 inside the exciton band (excluding the band edges $k = 0, \pi$) there are exactly two values of quasimomentum, i.e., $\pm k_0$ so that $\omega(\pm k_0) = \omega_0$.

Every scattering center a is described by a frequency-dependent scattering matrix Γ_a that relates the amplitudes of the outgoing plane waves to the incoming counterparts.[2,5] The scattering matrices $\Gamma_a(k)$ are unitary ($\Gamma_a(-k) = \Gamma_a^\dagger(k)$) and $\Gamma_a(k)\Gamma_a(-k) = 1$, since the quasimomenta of the incoming and outgoing waves are different just by the sign due to time-reversal symmetry. The scattering matrix can be viewed as a map $\Gamma_a : S^1 \to U(n_a)$ of the circle $S^1 \subset \mathbb{C}$, naturally embedded into the complex plane (by $z = e^{ik}$), which represents the Brillouin zone to the unitary group, with n_a being the vertex degree (the number of edges attached to it). The scattering matrices are analytical functions of $z \in S^1$, which implies they are complex analytical functions of z in some neighborhood of S^1 in $\mathbb{C}$.

Denoting by $\psi_{a\alpha}^{(\pm)}$ the amplitudes of the incoming/outgoing waves at vertex a from/to edge α, the ES equations are given by

$$\psi_{a\alpha}^{(-)} = \sum_\beta \Gamma_{a,\alpha\beta}(k)\psi_{a\beta}^{(+)}, \quad \psi_{a\alpha}^{(+)} = e^{ikL_\alpha}\psi_{b\alpha}^{(-)}, \qquad (13)$$

where $L_\alpha \in \mathbb{N}$ is the (integer) length of a segment α, and in the second set of equations, $\{a, b\}$ is the border of edge α. The first set of equations in (13) connects the amplitudes of the outgoing waves to the incoming counterparts at scattering centers, whereas the second set describes the change of the wave phase as a result of traveling along a linear segment.

We can further treat $\psi^{(\pm)} = \sum_{a\alpha} \psi_{a\alpha}^{(\pm)} e_{a\alpha}$ as vectors in the vector space V, spanned on the set of oriented vertices $e_{a\alpha}$ of our graph. We naturally represent $V = \oplus_a V_a = \oplus_\alpha V_\alpha$, where V_a and V_α are the vector spaces spanned on the vectors $\{e_{a\alpha}\}$ (α is attached to a) and $\{e_{b\alpha}, e_{c\alpha}\}$ (α connects b and c), respectively. Thus, $\Gamma_a(k)$ is a family of unitary matrices parameterized by k acting in V_a. This allows us to define a family $\Gamma(k) = \oplus_a \Gamma_a(k)$ of unitary matrices acting in the space V. We further introduce a unitary matrix $P = \oplus_\alpha P_\alpha$ acting in V with $P_\alpha e_{b\alpha} = e_{c\alpha}$ when α connects b and c, and a family of unitary matrices $\Lambda(k) = \oplus_\alpha e^{ikL_\alpha} \mathrm{id}_{V_\alpha}$. Using the above notations, the ES equations can be combined into one vector equation for $\psi^{(+)}$ only,

$$\tilde{\Gamma}(k)\psi^{(+)} = \psi^{(+)}, \quad \tilde{\Gamma}(k) \equiv \Lambda(k)P\Gamma(k), \qquad (14)$$

and we can interpret $\tilde{\Gamma} : S^1 \to U(n)$, with $n = \sum_a n_\alpha$, as a map of the Brillouin zone to the unitary group $U(n)$ (see the Appendix for a brief overview of topological properties of the unitary group). Representing the ES equations in the form (14) is crucial in obtaining the main result of this manuscript.

Since $\tilde{\Gamma}(k)$ is an analytic function of quasimomentum k, Eq. (14) can hold only for a finite number of quasimomenta $k_j \in S^1$ (hereafter, k is identified with $z = e^{ik}$). We further denote the number of linearly independent solutions of Eq. (14) for $k = k_j$ by m_j, referred to as the multiplicity. It is natural to refer to $m = \sum_j m_j$ as the total number of solutions of the ES equations (accounting for the degeneracy). Our strategy is (i) finding a relation between m and the number of quantum states/excitons, and (ii) establishing a lower bound for m, as well as showing that the bound is tight when the segment lengths L_α are large enough. Task (i) involves elementary linear algebra only. Achieving task (ii) necessitates topological intersection theory and constitutes the main technical result of this manuscript.

If $\psi^{(+)}$ is a solution of Eq. (14) with $k = k_j$, so that $k_j \neq 0, \pi$, then $\Gamma(k_j)\psi^{(+)}$ is also a solution with $k = -k_j$, which follows from $\tilde{\Gamma}(-k_j)\Gamma(k_j)\psi^{(+)} = \Lambda(-k_j)P\psi^{(+)} = P\Lambda(-k_j)\psi^{(+)} = \Gamma(k_j)\psi^{(+)}$. Here, we have used $\Gamma(-k)\Gamma(k) = 1$, $\Lambda(-k)\Lambda(k) = 1$, $P^2 = 1$, and $[P, \Lambda(k)] = 0$. Obviously, these two solutions represent the same standing wave, i.e., the same exciton. Therefore, the number of excitons with $k \neq 0, \pi$ is equal to the half of number of solutions of Eq. (14) with $k \neq 0, \pi$. The cases of $k = 0, \pi$ need a bit more care. First of all, a definition of $\Gamma_a(0)$ and $\Gamma_a(\pi)$ is not obvious, since in these cases, we have $k = -k$, and there is no distinction between the incoming and outgoing waves. We can still define $\Gamma_a(0)$ and $\Gamma_a(\pi)$ by using the analytical continuation from the Brillouin zone with $k \neq 0, \pi$. This leads to the matrices $\Gamma(k) \in U(n)$ with $k = 0, \pi$ that satisfy $(\Gamma(k))^2 = 1$. Introducing $\tilde{P}(k) = \Lambda(k)P$, we have $(\tilde{P}(k))^2 = 1$. Therefore, Eq. (14) in the cases $k = 0, \pi$ can be recast in an equivalent form $\Gamma(k)\psi^{(+)} = \tilde{P}(k)\psi^{(+)}$.

The excitons[1,11,12] (i.e., quantum states, rather than the solutions of the ES equations) are described by a vector $\psi \in V$ of amplitudes, rather than a pair $\psi^{(\pm)}$ of vectors (due to $k = -k$), and the equations adopt a form $\Gamma(k)\psi = \psi$, $\tilde{P}(k)\psi = \psi$, which can be recast as

$$\Gamma(k)\psi = \tilde{P}(k)\psi, \quad \Gamma(k)\psi = \psi. \qquad (15)$$

The first equation (15) is the $k = 0, \pi$ version of the ES equations. Therefore, the vector spaces $W_k^+ \subset W_k$ of the bandedge excitons are the subspaces of the spaces W_k of solutions to the ES equation that satisfy the second equation (15). It is straightforward to verify that W_k is an invariant subspace of $\Gamma(k)$. Therefore, we have $W_k = W_k^+ \oplus W_k^-$, where W_k^- is the unphysical space defined by the condition $\Gamma(k)\psi = -\psi$. These correspond to solutions that provide zero everywhere wavefunctions of the standing waves (excitons). Introducing $d_k^\pm = \dim W_k^\pm$ and recalling our consideration of the $k \neq 0, \pi$ case, we obtain for the number of excitons N,

$$N = \frac{1}{2}\left(m + (d_0^+ - d_0^-) + (d_\pi^+ - d_\pi^-)\right). \qquad (16)$$

Two comments are in place. First, identification of $d_k^\pm$ is much easier compared to finding m, since it is a linear problem, compared to a generalized spectral problem. Second, in a generic situation, we have $\Gamma_a(k) = -1$ for $k = 0, \pi$, which yields $d_k^+ = 0$, $d_k^- = n/2 = n_s$.

B. The index theorem

To establish a topological bound for the multiplicity m, we need to introduce a natural topological invariant, hereafter referred to as the winding number. Any continuous function $S^1 \to S^1$ has a natural topological invariant, known as its winding number. Intuitively, it counts how many times the domain S^1 wraps around the codomain S^1. Due to continuity, this must be an integer number (taking orientation into account). In our case, this integer $w(f)$, associated with a map $f : S^1 \to U(n)$ of the Brillouin zone to the unitary group, is defined by

$$w(f) = \int_{-\pi}^{\pi} \frac{dk}{2\pi i} (\det f(k))^{-1} \frac{d}{dk} \det f(k). \quad (17)$$

An interpretation of $w(f)$ that demonstrates its integer character is as follows. For a described map f, we have a map $\det f : S^1 \to U(1)$, since the determinant of a unitary matrix is a unimodular number. Then, $w(f)$, defined by Eq. (17), is the winding number of $\det f : S^1 \to U(1)$, i.e., the number of times $\det f(k)$ winds over the circle $U(1) \cong S^1$, while k goes once over the Brillouin zone S^1. The winding number is a topological (homotopy) invariant, meaning it does not change upon continuous deformations of the map f. In our further derivations, we will make substantial use the following key algebraic properties of the winding number, outlined in Sec. II C: For any maps $f, g : S^1 \to U(n)$ and $h : S^1 \to U(m)$, the relations, given by Eq. (7) hold. We reiterate that $f \cdot g : S^1 \to U(n)$ and $f \oplus h : S^1 \to U(n + m)$ are obtained by point-wise multiplication of matrices and forming a block-diagonal matrix, respectively, as well as that the relations of Eq. (7) are a direct consequence of the properties of determinants, namely, $\det(f \cdot g) = (\det f)(\det g)$ and $\det(f \oplus h) = (\det f)(\det h)$, combined with the definitions of Eq. (17).

To formulate the topological bound, we begin with a naive observation of any solution $k = k_j$ of the ES equations [Eq. (14)]: $\tilde{\Gamma}(k_j)$ should be unitary and have an eigenvalue equal to 1. If we let $D_1 U(n)$ denote the set of $n \times n$ unitary matrices that have at least one eigenvalue equal to 1, then solutions of the ES equations are associated with the intersections of the 1-dimensional cycle in $U(n)$ defined by the map $\tilde{\Gamma} : S^1 \to U(n)$ with the $(n^2 - 1)$-dimensional subspace $D_1 U(n) \subset U(n)$. Note that $\dim U(n) = n^2$. To compute m, the number of solutions of the ES equation, we need to count the intersections weighted with their multiplicities m_j. To do so, we further introduce the intersection index that counts the intersections weighted with local intersection indices q_j. Generally, q_j is different from m_j, whereas the total index $q = \sum_j q_j$ is a topological invariant, meaning it depends only on the topological classes of the intersecting cycles, in our case $D_1 U(n)$ and $\tilde{\Gamma}$. A statement that explicitly computes q as a topological invariant and relates it to the sum of the local intersection indices is usually referred to as an *index theorem*.[13] In our case, it adopts the form,

$$\sum_j q_j = w(\tilde{\Gamma}). \quad (18)$$

In what follows, we (a) define the local indices q_j and present plausible arguments on the validity of Eq. (18), (b) show that

$|q_j| \leq m_j$, (c) further show that when the segment lengths L_α are large enough we have $q_j = m_j$, and (d) explicitly compute

$$w(\tilde{\Gamma}) = 2 \sum_\alpha L_\alpha + \sum_a l_a, \quad l_a \equiv w(\Gamma_a). \quad (19)$$

Combining (b), (c), and Eq. (19) with Eq. (18), we obtain an explicit bound,

$$m \geq 2 \sum_\alpha L_\alpha + \sum_a l_a \quad (20)$$

that becomes tight for long enough segments.

Derivation of Eq. (19) is straightforward. We start with a representation for $\tilde{\Gamma}$ given by Eq. (14) with explicit expressions for the three factors in the r.h.s.,

$$\tilde{\Gamma} = (\oplus_\alpha e^{ikL_\alpha} \mathrm{id}_{V_\alpha}) \cdot (\oplus_\alpha P_\alpha) \cdot (\oplus_a \Gamma_a) \quad (21)$$

and apply the properties of the winding number [Eq. (7)], combined with explicit computations $w(e^{ikL_\alpha} \mathrm{id}_{V_\alpha}) = 2L_\alpha$, $w(P_\alpha) = 0$, and the definition of l_a, given in Eq. (19).

The local intersection indices are defined in the following way. Consider a solution of Eq. (14) [or equivalently an intersection of $\tilde{\Gamma}$ with $D_1 U(n)$] of multiplicity m_j at $k = k_j$. The analytic nature of $\tilde{\Gamma}(k)$ implies solutions is isolated. Therefore, there is a small neighborhood $(k_j - \varepsilon, k_j + \varepsilon)$ of k_j for which there are no other solutions. For any k, in this neighborhood, there will be m_j eigenvalues of $\tilde{\Gamma}(k)$ that are close to 1 that can be distinguished from the other eigenvalues. If $k \neq k_j$, then all of these m_j eigenvalues must be close, but not equal, to 1 (any eigenvalue precisely equal to 1 would signify a solution and contradict the isolated nature of the solutions). Let m_j^+ be the number of eigenvalues close to 1 with positive imaginary part for $k \in (k_j, k_j + \varepsilon)$ and m_j^- be the number of eigenvalues with positive imaginary part for $k \in (k_j - \varepsilon, k_j)$. Define the local intersection index $q_j := m_j^+ - m_j^-$. Naturally, q_j can be interpreted as the number of eigenvalues of $\tilde{\Gamma}(k)$ that move through the point $1 \in U(1)$ in the counterclockwise direction while $k \in (k_j - \varepsilon, k_j + \varepsilon)$ goes through k_j in the counterclockwise direction. Obviously, $|q_j| \leq m_j$.

C. Sketch of a proof of index theorem

Intuitive arguments in support of validity of the index theorem [Eq. (18)] have been presented at the end of Sec. II C. In this section, we present a sketch of a proof. Consider a map $f : S^1 \to U(n)$ that intersects $D_1 U(n)$ at a finite number of (isolated) points $k_j \in S^1$. The main idea is to replace the segments of the curve $f : S^1 \to U(n)$ for the values of the parameter $k_j - \varepsilon < k < k_j + \varepsilon$ close to the intersection points k_j with the segments that miss $D_1 U(n)$ and compute the change in the winding number due to the aforementioned replacement. The resulting curve $\tilde{f}$ misses $D_1 U(n)$, so that $w(\tilde{f}) = 0$, and the winding number $w(f)$ is described by the total change of the winding number due to the all replacements.

Note that $f(k_j)$ has exactly m_j unit eigenvalues. Denote $Y_j^\varepsilon = \{k \in S^1 : k_j - \varepsilon < k < k_j + \varepsilon\}$, and $Z_\delta = \{e^{i\varphi} : |\varphi| < \delta\}$. Obviously, we can choose $\varepsilon > 0$ and $\delta > 0$ to be small enough, so that $f(k)$ has exactly m_j eigenvalues that belong to Z_δ for all $k \in Y_j^\varepsilon$, and Y_j^ε do not intersect for different j. Choose some $k_j^-, k_j^+ \in Y_j^\varepsilon$ so that $k_j^- < k_j < k_j^+$ and replace

the curve segments $f|_{[k_j^-, k_j^+]} : [k_j^-, k_j^+] \to U(n)$ with the curves $\tilde{f}_j : [k_j^-, k_j^+] \to U(n)$ that miss $D_1U(n)$ in the following way. Fix basis sets $e_i^{(j)\pm}$ that diagonalizes the matrices $f(k_j^\pm)$ and let $\lambda_i^{(j)\pm}$ be the corresponding eigenvalues. Consider two sets of the eigenvalue trajectories $\lambda_i^{(j)} : [k_j^-, k_j^+] \to U(1)$ that start and end at $\lambda_i^{(j)-}$ and $\lambda_i^{(j)+}$, respectively. Note that any set of the eigenvalue trajectories can be extended to a path $[k_j^-, k_j^+] \to U(n)$ that starts and ends at $f(k_j^-)$ and $f(k_j^+)$, respectively, by choosing basis set trajectories $e^{(j)} : [k_j^-, k_j^+] \to U(n)$ (here, we identify the sets of orthonormal basis sets with the elements of the unitary group $U(n)$) that start and end at $e^{(j)-}$ and $e^{(j)+}$, respectively. Denote these two extensions by $f_j^{(r)} : [k_j^-, k_j^+] \to U(n)$ with $r = 0, 1$. The first set of the eigenvalue trajectories connect the initial eigenvalues to their final counterparts via the paths of the minimal length. The corresponding extensions $f_j^{(0)}$ to paths in $U(n)$ are topologically (homotopy) equivalent to the original path segments $f|_{[k_j^-, k_j^+]}$. The second set connects the initial to the final eigenvalues in a way that the trajectories miss the point $\lambda = 1$. This can be achieved in the following way. Among the first set of trajectories, there are exactly $|q_j|$ ones that go through the point $\lambda = 1$ to change the number of eigenvalues, which lie in Z_δ and have a positive imaginary part, from m_j^- to m_j^+. We replace these $|q_j|$ trajectories with the ones that go over the circle in the opposite direction, missing the point $\lambda = 1$.

The contributions of both types of trajectory segments to the integral representation for the winding number [Eq. (5)] are

$$w(f_j^{(r)}) = \int_{k_j^-}^{k_j^+} \frac{dk}{2\pi} (\det f_j^{(r)}(k))^{-1} \frac{d}{dk} \det f_j^{(r)}(k)$$
$$= \sum_i \int_{k_j^-}^{k_j^+} \frac{dk}{2\pi} (\lambda_i^{(j,r)}(k))^{-1} \frac{d}{dk} (\lambda_i^{(j,r)}(k)). \quad (22)$$

Comparing the integrals under the summation sign in Eq. (22) for $r = 1$ and $r = 0$, we can see that they are identical for all i, except for those that correspond to the eigenvalue trajectories that were replaced to avoid going through the $\lambda = 1$ point. In the latter case, the difference between two integrals between the $r = 1$ and $r = 0$ cases can be represented as an integral over the whole circle with the eigenvalue winding exactly once over $U(1)$ in the clockwise (counterclockwise) direction for the replaced segment that corresponds to going through the $\lambda = 1$ point in the counterclockwise (clockwise) direction, respectively. This yields

$$w(f_j^{(1)}) - w(f_j^{(0)}) = -q_j. \quad (23)$$

Denote by $\tilde{f} : S^1 \to U(n)$ the curve obtained by replacing the original segments $f|_{[k_j^-, k_j^+]}$ with $f_j^{(1)}$. As shown earlier, replacing the original segments with $f_j^{(0)}$ does not change the winding number. Therefore, it follows from Eq. (23) that

$$w(\tilde{f}) - w(f) = -\sum_j q_j. \quad (24)$$

Since $\tilde{f}$, by construction, misses $D_1U(n)$, we have $w(\tilde{f}) = 0$, so that Eq. (24) implies the statement of the index theorem [Eq. (18)].

A rigorous proof of the index theorem can be given using homological arguments, which is beyond the scope of this manuscript. This formal argument will be given in a later paper.

D. The long arm limit

To demonstrate the tightness of the bound in Eq. (20), we show that for long enough segments at any intersection, we have $m_j^+ = m_j$ and $m_j^- = 0$, so that $q_j = m_j$, which validates (c). This can be done by applying quantum mechanical perturbation theory in the degenerate case, i.e., in a small neighborhood of an intersection k_j,

$$\tilde{\Gamma}(k) \approx \tilde{\Gamma}(k_j) + (k - k_j) \frac{d\Lambda(k_j)}{dk_j} \Lambda^{-1}(k_j) \tilde{\Gamma}(k_j)$$
$$+ (k - k_j) \tilde{\Gamma}(k_j) \Gamma^{-1}(k_j) \frac{d\Gamma(k_j)}{dk_j}. \quad (25)$$

The first term in the perturbation grows linearly with the segment length, whereas the second one is segment length independent and can be neglected for long enough segments. So that Eq. (25) can be approximated as

$$d\tilde{\Gamma}(k) \approx i(k - k_j)L\tilde{\Gamma}(k_j), \quad (26)$$

where $L = \oplus_\alpha L_\alpha \mathrm{id}_{V_\alpha}$ is a Hermitian operator whose eigenvalues are the segment lengths L_α. According to quantum mechanical perturbation theory, the first-order correction to a (possibly degenerate) eigenvalue is given by the eigenvalues of the projection of the perturbation to the operator onto the subspace of eigenvectors that correspond to the zero-order eigenvalue, which in our case is the m_j-degenerate unit eigenvalue. Since $\tilde{\Gamma}(k_j)$ acts as the unit operator in the relevant eigenvector subspace, we need to inspect the eigenvalues of the projection of L. According to the quantum mechanical variational principle, all eigenvalues of the projection of a Hermitian operator exceed the lowest eigenvalue of the operator itself, which in the case of L is the minimal segment length. Therefore, in a small neighborhood of k_j, where the perturbation theory is applicable, the imaginary part of all close to one eigenvalues is positive/negative for $k > k_j$ and $k < k_j$, respectively, which completes the argument.

IV. APPLICATION TO POLYFLOURENE-BASED MOLECULE

We further illustrate the above theoretical models using quantum-chemical calculations of branched polyfluorenes, a class of technologically important conjugated polymers.[14,15] A Y-shaped family of molecules Y_{abc} is shown in Fig. 3(a), where a, b, and c are the segment lengths in repeat units. The ground state molecular geometries were optimized at the AM1 level[16] with Gaussian09 package,[17] followed by computing the vertical excitation energies and transition density matrices of up to 48 excited states using the Collective Electronic Oscillator (CEO) method.[18,19] We select the first exciton band by checking the transition density matrices. Following our previous work,[4–7,9] we further extracted the dispersion $\omega(k)$ (Fig. 4(a)) and scattering matrices $\Gamma(k)$ for $k \in [0, \pi]$. For the
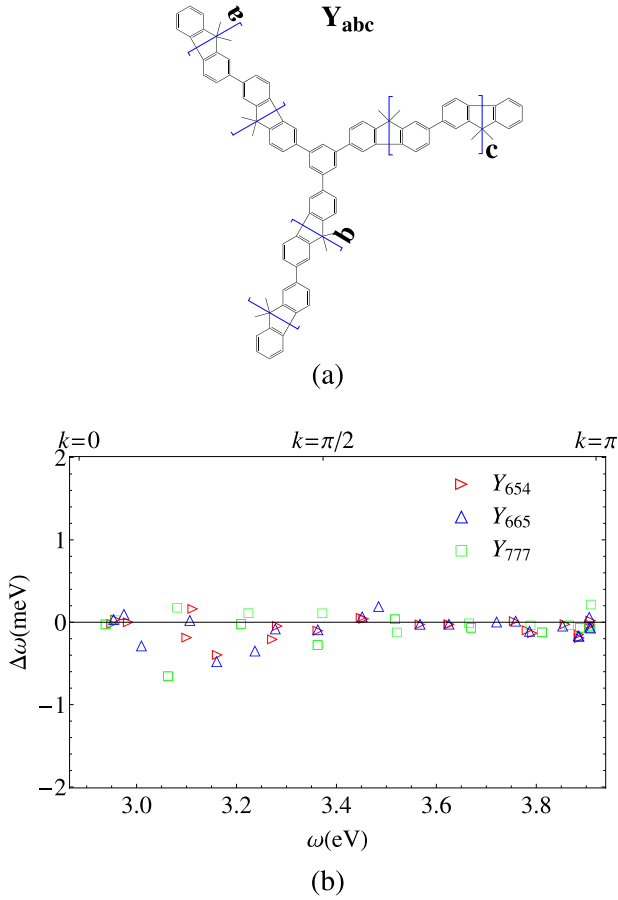
(a)

(b)

FIG. 3. (a) The Y-shape molecule Y_{jmn}. (b) Differences between excitation energies obtained by ES approach and the CEO method. $\Delta\omega = \omega_{ES} - \omega_{CEO}$.
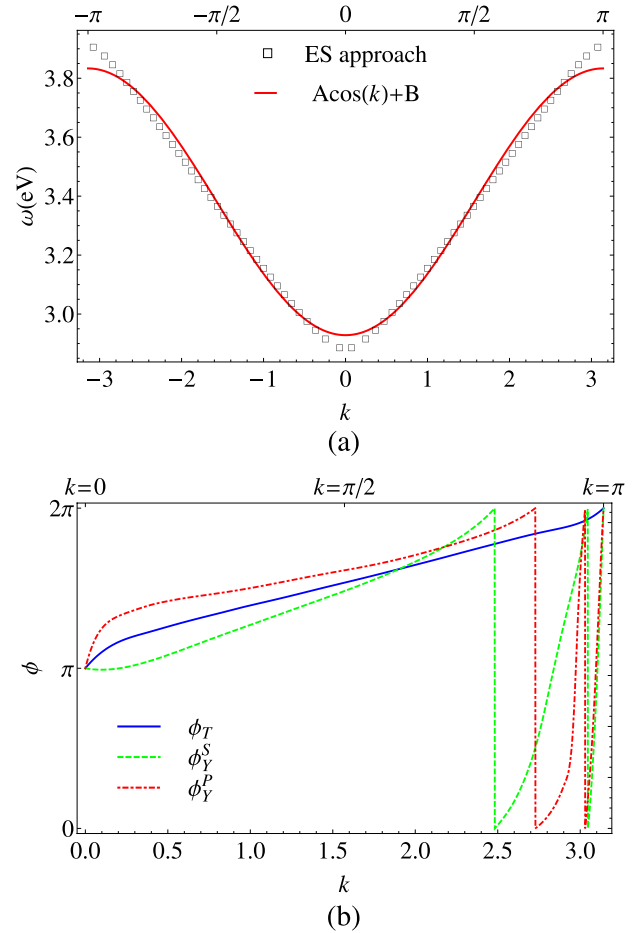

(a)

(b)

FIG. 4. (a) The dispersion of the first exciton band fit with a single cosine $A\cos(k) + B$, $A = -0.452$ and $B = 3.381$. (b) The scattering phases ϕ_T of the terminals, and ϕ_Y^S and ϕ_Y^P of the Y joint. Both ϕ_Y^S and ϕ_Y^P are shifted by -4π.

scattering matrix Γ_T of the terminal and Γ_Y of Y joint, we have $\det(\Gamma_T) = e^{i\phi_T}$ and $\det(\Gamma_Y) = e^{i\phi_Y^S}e^{2i\phi_Y^P}$, where ϕ_Y^S and ϕ_Y^P are scattering phases of the Y-joint corresponding to a singlet and 2-fold degeneracy, respectively (see Ref. 9 for details). The scattering parameters in Fig. 4 allow for an accurate extraction of the excitation energies for an arbitrary Y-shaped molecules using the ES equation (14), see, for example, Fig. 3(b).

From the scattering phases (Fig. 4(b)), we extract the winding numbers $w(\Gamma_T) = 1$ and $w(\phi_Y^S) = w(\phi_Y^P) = 5$, the latter following from each having 2 kinks (sharp 2π jumps) over half of the Brillouin zone, and a total shift of 2π over the full Brillouin zone. Therefore, $w(\Gamma_Y) = w(\phi_Y^S) + 2w(\phi_Y^P) = 15$ and according to Eq. (10), the number of solutions for the ES equation is

$$m = w(\tilde{\Gamma}) = 2(a + b + c) + 3w(\Gamma_T) + w(\Gamma_Y)$$
$$= 2(a + b + c) + 18, \qquad (27)$$

for the molecule Y_{abc} shown in Fig. 3(a). There are a total of 6 solutions for $k = 0$ and $k = \pi$, which do not correspond to physical states. Hence, the total number of states with excitation energies inside the exciton band is

$$N = (m - 6)/2 = a + b + c + 6, \qquad (28)$$

where the factor $1/2$ arises due to time reversal symmetry, in full accordance with Eq. (12).

The result presented in Eq. (28) can be obtained directly by using the general counting formula [Eq. (2)] with $l_T = 1$, $l_Y = 15$, $d_0 = d_\pi = n_s = 3$, or its generic counterpart [Eq. (1)] with $Q_T = 0$, $Q_Y = 6$, where the vertex topological charges were computed using Eq. (3).

V. CONCLUSION

In summary, we have identified an integer-valued topological invariant l_a, associated with an arbitrary scattering center a of a branched molecular structure (schematically represented by a graph), which is completely determined by the corresponding scattering matrix $\Gamma_a(k)$. The integer l_a was motivated by a well known concept from algebraic topology, and has a simple integral form,

$$l_a = \int_{-\pi}^{\pi} \frac{dk}{2\pi i} (\det \Gamma_a(k))^{-1} \frac{d}{dk} \det \Gamma_a(k). \qquad (29)$$

We have reduced the problem of counting the number of solutions m of the ES equations to an intersection problem, and further extended topological intersection theory to come up with a simple lower bound for m in terms of the segment lengths L_α and vertex winding numbers l_a [Eq. (10)]. This allowed us to evaluate the total number of electronic

excitations N within a spectral region of a single exciton band [Eq. (2)]. Additionally, this permitted a simple and easy to interpret formula [Eq. (1)] that works for the generic case. Furthermore, it does not contain the band-edge indices d_k, and counts the number of excitations in terms of the vertex topological charges $Q_a = (l_a - r_a)/2$. We also demonstrated that both lower bounds become exact for long enough linear segments. Quantum chemical calculations of Y-shaped branched oligofluorenes provide direct illustration of our theory. Being formulated in terms of dispersion relations and scattering matrices, the presented counting expressions are very general. Therefore, this approach can be applied to any quantum quasi-one-dimensional systems, or even more generally to any wave phenomena in networks, e.g., optical communications.

The introduced concepts carry a number of possible implications on studies of photo-physical phenomena in branched conjugated structures. The presented theory provides an explicit link between the graph topology (connectivity) of the branched system and the interlining molecular electronic and optical properties defined by the spectrum of electronic excitations: The excited-state electronic structure in a finite molecule is a result of complex interplay between excitations in linear segments, determined by the exciton band structure in infinite polymers and electronic excitations in scattering centers. Chemical coupling between the linear segments and scattering centers modifies the excitations in the latter, resulting in bound and resonant states.[8] The resonant states that fall inside the exciton band show up as resonances in the density of states, or equivalently in sharp 2π-rotations of the det $\Gamma_a(k)$ in the narrow regions of k around the resonance of $\omega(k)$ with a state at the scattering center, referred to as kinks[8,10] are much harder to analyze, compared to bound states. Our counting result allows for a following interpretation: the number N_{res} of resonant states in a finite structure is given by just the sum $N_{res} = \sum_a Q_a$ of the topological charges, associated with the scattering centers. In particular, when the connectivity of a conjugated structure is changed, with the scattering centers being kept the same, the resonant states do not appear or disappear, but rather their structure is modified. In the future, we will study the dependence of the resonant states structure on the graph topology by examining the position dependent exciton density of states.

Finally, we note that the possible increase in the number of states N inside the exciton band for shorter segments (i.e., in the case of inequality in Eq. (1)) can be explained by noting some of the bound states get "pushed" inside the exciton band due to the effects of quantum confinement of the exciton center of mass motion. We would like to emphasize that the inequality that forms a lower bound for N [Eq. (1)], which follows from topological considerations, reflects a physical/chemical trend: Effects of quantum confinement of the exciton center of mass motion are sometimes turning the bound states into the resonant counterparts and not the other way around.

ACKNOWLEDGMENTS

APPENDIX: TOPOLOGICAL PROPERTIES OF THE UNITARY GROUPS

In this appendix, we describe some relevant topological properties of the unitary groups $U(n)$ in a simple and self-consistent fashion.

First, taking the determinant of a unitary matrix defines a map $\det : U(n) \to U(1)$. We can define a map $s : U(1) \to U(n)$, referred to as a section, by defining $s(\lambda)$ to be a unitary matrix with the eigenvalues $(\lambda, 1, \ldots, 1)$. Note that $\det \circ s = \mathrm{id}$. A choice of a section defines an isomorphism (of topological spaces, as well as smooth manifolds, but not groups) $u : U(n) \to U(1) \times SU(n)$, given by $u(g) = (\det g, g(s(\det g))^{-1})$. We are interested in topological invariants of maps $S^1 \to U(n)$, or more precisely the homotopy equivalence classes of such maps. Since $U(n) \cong U(1) \times SU(n)$, a map $S^1 \to U(n)$ is described by a pair of maps $S^1 \to U(1)$ and $S^1 \to SU(n)$. If $[X, Y]$ denotes the set of homotopy classes of maps $X \to Y$, then we have $[S^1, U(n)] \cong [S^1, U(1)] \times [S^1, SU(n)]$. Using the fact that $[S^1, SU(n)]$ consists of a single point, each 1-dimensional cycle in $SU(n)$ is contractible (meaning topologically equivalent to the constant map). As a topological space $U(1) \cong S^1$, so that $[S^1, U(1)] \cong [S^1, S^1] \cong \mathbb{Z}$, with the homotopy class of a map given by its winding number. Therefore $[S^1, U(n)] \cong \mathbb{Z}$, i.e., maps from a circle to a unitary group have exactly one topological invariant that takes values in the integers $\mathbb{Z}$. This also implies that the determinant map $\det : U(n) \to U(1)$ generates an isomorphism $[S^1, U(n)] \to [S^1, U(1)] \cong \mathbb{Z}$. Formulated in a simpler way, to identify the topological invariant of a map $f : S^1 \to U(n)$ one needs just to compute the winding number $w(\det(f))$, where $\det(f) = \det \circ f$. Therefore, we refer to the topological invariant $w(f) = w(\det(f))$ of a map $f : S^1 \to U(n)$ as its winding number.

In proving the index theorem (see a sketch in Sec. III C), we will make use of the fact that the space $U(n) \setminus D_1 U(n)$ is contractible to a point. We reiterate that $D_1 U(n) \subset U(n)$ is a closed subspace of unitary matrices that have at least one unit eigenvalue. In particular, if a map $f : S^1 \to U(n)$ misses $D_1 U(n)$, i.e., can be represented as a composition $S^1 \to U(n) \setminus D_1 U(n) \to U(n)$ (where the second map is the inclusion), then $w(f) = 0$. The contractibility can be demonstrated by presenting a continuous deformation (homotopy) of the identity map $\mathrm{id} : U(n) \setminus D_1 U(n) \to U(n) \setminus D_1 U(n)$ to a constant map, the latter mapping the whole domain to a single point. The spectra of matrices $x \in U(n) \setminus D_1 U(n)$ are contained in an open interval $(0, 2\pi)$ in the sense that the eigenvalues are given by $\lambda_s = e^{i\varphi_s}$ with $\varphi_s \in (0, 2\pi)$. Consider an obvious contraction of the interval to its center, given by the point $\pi \in (0, 2\pi)$. The contraction/deformation of the interval defines a corresponding contraction/deformation of

the spectrum of any matrix $x \in U(n) \setminus D_1 U(n)$. The described above deformation of the spectrum provides a deformation of the corresponding matrix: we deform the spectrum keeping the eigenspaces, related to the eigenvalues unchanged. This describes a contraction of $U(n) \setminus D_1 U(n)$ to the point $-\mathrm{id} \in U(n) \setminus D_1 U(n)$.

[1]C. Wu, S. V. Malinin, S. Tretiak, and V. Y. Chernyak, Nat. Phys. **2**, 631 (2006).

[2]C. Wu, S. V. Malinin, S. Tretiak, and V. Y. Chernyak, J. Chem. Phys. **129**, 174111 (2008).

[3]C. Wu, S. V. Malinin, S. Tretiak, and V. Y. Chernyak, J. Chem. Phys. **129**, 174112 (2008).

[4]C. Wu, S. V. Malinin, S. Tretiak, and V. Y. Chernyak, J. Chem. Phys. **129**, 174113 (2008).

[5]C. Wu, S. V. Malinin, S. Tretiak, and V. Y. Chernyak, Phys. Rev. Lett. **100**, 057405 (2008).

[6]H. Li, S. V. Malinin, S. Tretiak, and V. Y. Chernyak, J. Chem. Phys. **132**, 124103 (2010).

[7]H. Li, C. Wu, S. V. Malinin, S. Tretiak, and V. Y. Chernyak, J. Phys. Chem. Lett. **1**, 3396 (2010).

[8]H. Li, M. J. Catanzaro, S. Tretiak, and V. Y. Chernyak, J. Phys. Chem. Lett. **5**, 641 (2014).

[9]H. Li, C. Wu, S. V. Malinin, S. Tretiak, and V. Y. Chernyak, J. Phys. Chem. B **115**, 5465 (2011).

[10]H. Li, S. V. Malinin, S. Tretiak, and V. Y. Chernyak, J. Chem. Phys. **139**, 064109 (2013).

[11]E. Collini and G. D. Scholes, Science **323**, 369 (2009).

[12]G. D. Scholes and G. Rumbles, Nat. Mater. **5**, 683 (2006).

[13]E. H. Spanier, *Algebraic Topology* (Springer-Verlag, New York, 1981).

[14]U. Scherf and E. J. W. List, Adv. Mater. **14**, 477 (2002).

[15]S. Becker, C. Ego, A. C. Grimsdale, E. J. W. List, D. Marsitzky, A. Pogantsch, S. Setayesh, G. Leising, and K. Mullen, Synth. Met. **125**, 73 (2001).

[16]M. J. S. Dewar, E. G. Zoebisch, E. F. Healy, and J. J. P. Stewart, J. Am. Chem. Soc. **107**, 3902 (1985).

[17]M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson *et al.*, GAUSSIAN *09, Revision A.01*, Gaussian Inc., Wallingford, CT, 2009.

[18]S. Mukamel, S. Tretiak, T. Wagersreiter, and V. Chernyak, Science **277**, 781 (1997).

[19]S. Tretiak and S. Mukamel, Chem. Rev. **102**, 3171 (2002).

[20]Using the standard linear algebra notation this can be represented as $\Gamma(k) = \oplus_a \Gamma_a(k)$.