Prediction of molecular single-photon emitters: A materials-modeling approach

Prediction of molecular single-photon emitters: A materials-modeling approach

Erik Karlsson Öhman [1], Daqing Wang [2], R. Matthias Geilhufe [1], and Christian Schäfer [1,3,*]
[1]*Department of Physics, Chalmers University of Technology, 412 96 Göteborg, Sweden*
[2]*Institute of Applied Physics, University of Bonn, Wegelerstraße 8, 53115 Bonn, Germany*
[3]*Institute of Applied Physics, TU Wien, Vienna, Austria*

(Received 22 December 2025; accepted 4 March 2026; published 26 March 2026)

Interfacing light with quantum systems is an integral part of quantum technology, with the most essential building block being single-photon emitters. Although various platforms exist, each with its individual strengths, molecular emitters boast a unique advantage—namely, the flexibility to tailor their design to fit the requirements of a specific task. However, the characteristics of the vast space of possible molecular configurations are challenging to understand and explore. Here, we present a theoretical and computational framework to initiate exploration of this vast potential by integrating database analysis with microscopic predictions. Using a model system of dibenzoterrylene in an anthracene host as benchmark, our approach identifies promising new candidates, among them a chiral molecular emitter. Future extensions of our approach integrated with machine learning routines hold the promise of ultimately unlocking the full potential of molecular quantum light-matter interfaces.

DOI: 10.1103/11bh-ywcc

I. INTRODUCTION

Interfacing quantum light with quantum emitters promises avenues for secure communication [1], efficient simulation of quantum-mechanical systems [2], quantum computing [3], nondestructive imaging techniques [4], and alternative avenues for material design [5,6]. An ideal single-photon emitter (SPE) delivers single photons of a specific wavelength on demand via a two-level transition with Fourier-limited linewidth. Moreover, SPEs need to be interfaced to resonant optical structures to enable near-unity-efficiency coupling to a single optical mode [7], permitting gigahertz operation rates [8]. Multiple emitters can be cooperatively coupled, promising a flexible handle to modify their emission characteristics. This might range from enhanced emission speed [9], thus hastening operation times, over topologically reduced emission linewidth [10], the controlled generation of entangled photons [11], to the creation of super- and subradiant states of coupled molecules [12,13].

Various platforms have been studied over the last decade. Among the most promising platforms are semiconductor quantum dots [14], color centers in solids, such as the widely known nitrogen-vacancy centers in diamond [15], defects in two-dimensional materials [16], and aromatic molecules embedded in organic hosts [7]. An excellent example of the latter being dibenzoterrylene (DBT) molecules embedded in anthracene [17]. Single DBT molecules are photostable, feature near lifetime-limited linewidth, have negligible intersystem-crossing yield, and high branching ratio. Compared to other platforms, molecular SPEs cover a broad range of wavelengths over the visible to near-infrared spectrum [18] and feature a unique strength of the flexibility to design a molecule to our liking, i.e., providing an exceptional degree of control over the emitter's characteristics. However, utilizing this flexibility requires us to understand and identify promising molecules and suitable hosts for such applications [19]—a highly nontrivial task given the vast space of possible combinations.

In this work, we describe a first attempt for a fully predictive theoretical approach to scan the vast available configuration space and identify ideal combinations of emitter and host molecules for specific applications. A schematic overview of our approach is illustrated in Fig. 1. We combine database analysis with similarity investigations and microscopic calculations to identify suitable alternatives to established matches, here in particular to DBT in anthracene. The vast majority of the touched database entries has been experimentally synthesized. In this work, we focus on varying the guest molecule embedded in anthracene due to its prominent use in molecular SPE designs. The outlined methodology is suitable for changing both guests and hosts. On the methodology side, we incorporate (time-dependent) density-functional theory (TDDFT) and machine learning potentials. Various promising candidates are identified through our approach. Among them are new candidates, one being an inherently chiral emitter—a potential building block for chiral photonics [20]. Our approach is validated by the prediction of terrylene, a well-known match. We begin by providing in Sec. II a brief introduction into required theoretical tools before discussing our results and promising candidates in Sec. III.

*Contact author: christian.schaefer@tuwien.ac.at

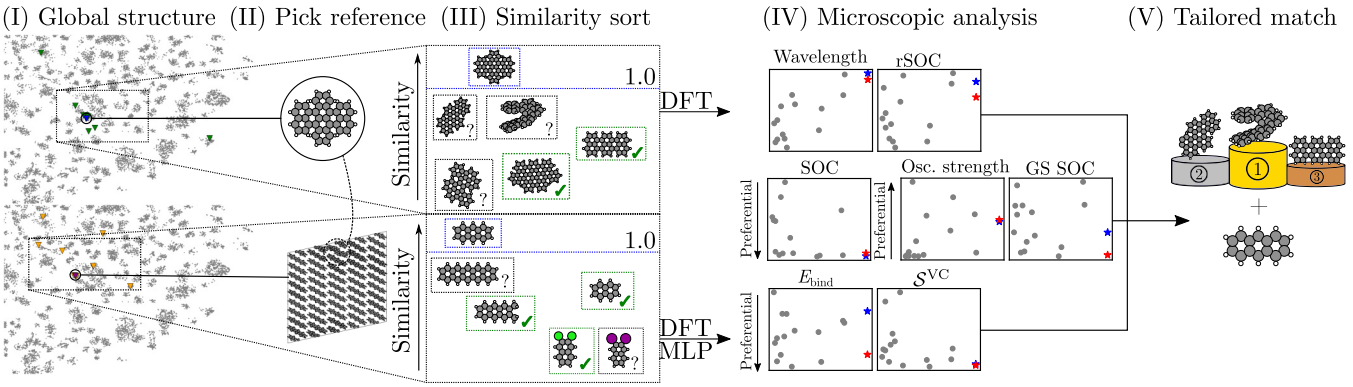

FIG. 1. Hunting emitter-host pairs: (I) Entries from the COD that are available for substructure search by SMILES (approximately 2×10^5 structures) are collected. The SMILES strings are converted to bitvectors of fixed size and dimensionality reduction and clustering is performed on the dataset by using the t-SNE and HDBSCAN algorithms. (II) A suitable, known emitter-host pair is chosen as a reference. We will limit ourselves here to DBT in anthracene (illustrated), which is among the most studied and promising candidates. (III) The Tanimoto index is used as a metric and potential replacement emitters and hosts are ranked by their similarity to the references. Suitable high-scoring candidates are examined and selected. (IV) Further microscopic analysis, such as DFT and molecular dynamics, is performed for a set of relevant observables. This includes, for example, emission wavelength, oscillator strength, various metrics for spin-orbit coupling, vibronic coupling, and formation energies. (V) Results are evaluated and emitter-host pairs are proposed.

II. THEORY

Identifying suitable combinations of emitters and hosts can utilize a variety of tools. Here, we develop an approach that combines routines from cheminformatics, electronic structure theory, molecular dynamics, and machine learning. Our brief introduction will not suffice to deliver a comprehensive overview of such a broad set of frameworks. Readers are therefore encouraged to investigate the cited references for further details.

A. Clustering and similarity

In a first step, we apply clustering to explore the global structure of organic molecules with respect to the distribution of known emitters and hosts. A prerequisite for this is to define a descriptor that uniquely identifies a molecule and a metric that defines a relative distance between two molecules. We use the simplified molecular input line entry system (SMILES) descriptor in combination with Morgan fingerprints and similarity measured by the Tanimoto index, as detailed below.

1. SMILES

Our dataset consists of all entries from the Crystallography Open Database (COD) that are available for substructure search by SMILES [21] plus selected known emitter molecules [18]—totaling 172 228 entries at the time of writing. SMILES characterizes the chemical composition and connectivity of a molecule [22]. Atoms and bonds of the molecule are encoded with a series of characters in a string, as illustrated in Fig. 2(a). This means that spatial information on, for example, bond distances and angles, is not preserved other than what can be inferred from the abstract chemical bond notation of the molecule. SMILES are not unique; i.e., one molecule can have two different SMILES, but can be canonicalized [23]. Essentially complementary to SMILES are local descriptors, such as the smooth overlap of atomic positions (SOAPs), which provide an explicit representation of atomic coordinates [24]. Defining simple similarities solely on SOAPs averaged per atom lacks information about the global structure (see Sec. I of the Supplemental Material [25]). SOAP (or other local descriptors) encodes local information, e.g., about the conjugation system [26], and a combination with SMILES will be the subject of future research.

2. Similarity and fingerprints

To be able to quantify the similarity of two molecules, the SMILES strings have to be encoded into a fingerprint, a binary vector of fixed size. We use the open-source PYTHON cheminformatics toolkit RDKit [27], and its implementation of the Morgan algorithm [28,29], for this purpose. The creation of the fingerprint depends on two parameters: the bitsize, i.e., the length of the fingerprint that is generated, and the radius, how many neighbors of each atom should be encoded into the fingerprint. For instance, if the radius is set to zero, all atoms are encoded into the fingerprint without any information about their connectivity. Setting the radius to 1 allows the fingerprint to encode information about each atom and its immediate neighbors, setting the radius to 2 encodes information about the atom and its two nearest neighbors, and so on. For the purpose of this study, we have set the bitsize to 1024 and the radius to 2. The similarity of two molecules to each other will be defined via the Tanimoto index $S_{A,B} = c/(a + b - c)$, a well-established metric in the field of cheminformatics [30]. Here, $a(b)$ is the number of bits set to 1 in fingerprint $A(B)$, and c is the intersection of all bits set to 1 in fingerprint A and B. At most $S_{A,B} = 1$ for two identical fingerprints and $S_{A,B} = 0$ if the fingerprints have no common features.

3. Dimensionality reduction and clustering

The general structure of our dataset is revealed using the t-distributed stochastic neighbor embedding (t-SNE) algorithm. As the name suggests, t-SNE is a stochastic approach to project data onto a low-dimensional representation that is ideal for clustering and visual assessment. To enhance in-

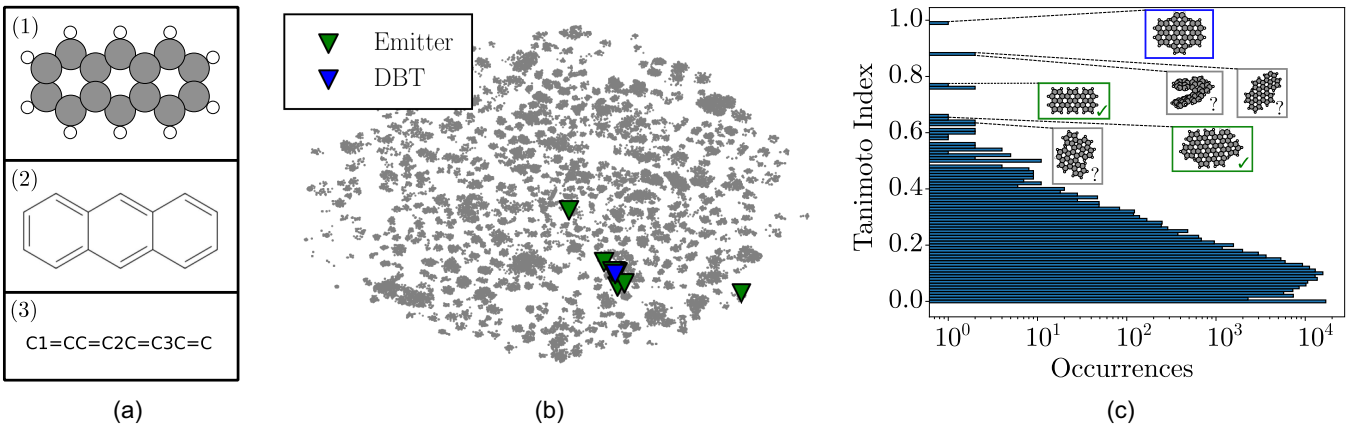

FIG. 2. (a) Generation of SMILES strings: Schematic showing the generation of SMILES strings from (1) a molecular structure, which is converted into (2) a graph representing the atomic connectivity, and finally into (3) a linear text-based SMILES representation of the molecule. (b) Global structure of dataset: t-SNE plot showing the distribution of molecular data. Known emitters are indicated by green triangles, and DBT is highlighted as a blue triangle. Similarity is quantified by converting SMILES into Morgan fingerprints, which are then set in relation using the Tanimoto metric (details in text). (c) Similarity ranking: Logarithmic distribution of the Tanimoto index of the materials in the Crystallography Open Database with respect to DBT. The insets show relevant samples from the distribution. They are, from top to down and left to right, as follows: DBT, terrylene, 4127216, 2000909, 1555531, and BDPB, where the numbers are COD IDs. The blue square denotes the reference material, green squares denote a known emitter, although not necessarily in anthracene, and the gray squares denote previously unknown candidates.

terpretability, we use the Hierarchical Density-Based Spatial Clustering of Applications with Noise (HDBSCAN) algorithm to identify clusters and eliminate outliers for ease of visualization. Both algorithms are implemented through the scikit-learn library and a perplexity value of 50 is used for t-SNE and a minimum cluster size of 10 is used for HDBSCAN [31].

Systematic structure identification is conceptually similar to identifying a preferred location. Our global map [panel (I) in Fig. 1] provides a coarse-grained overview of the terrain and allows to separate districts by specific characteristics. An intuitive approach to finding the "best SPE in town" is to pick a suitable candidate and begin exploring similar structures, as shown in steps (II) and (III). Candidates similar enough are then sampled [step (IV)]; i.e., we obtain microscopic information with the tools detailed in the following. Subsequent ranking brings particularly promising candidates to light [step (V)]. First steps in the direction of machine learning classification techniques (see Sec. III D) pave a way toward more explorative approaches in the future.

B. Molecular structure

Estimating key merits of an SPE, including oscillator strength, transition energy, and intersystem-crossing yield, requires a solution of the electronic Schrödinger equation. Density-functional theory (DFT) strikes a balance between accuracy and computational cost and is therefore the predominant approach to describe molecular and crystal structures. The linear many-body Schrödinger equation, scaling exponentially in system size, is mapped onto a set of coupled nonlinear single-particle Kohn-Sham equations [32] that can be solved with subcubic scaling in the case of semilocal, and quartic scaling in the case of the here used hybrid functionals. Throughout this work, we utilize ORCA [33] to solve the latter. For an emitter of interest, we begin with identifying the optimal ground-state configuration by performing geometry relaxation using the B3LYP functional [34,35] with D4 dispersion correction [36] using the def2-TZVPD basis. [37] Based on those structures, we perform Casida linear-response time-dependent DFT calculations [38] (including Tamm-Dancoff approximation) to obtain electronic transition strength, frequency, rotary strength, and spin-orbit coupling (SOC). Response TDDFT and geometry relaxation are then combined to estimate Stokes shifts and a figure of merit for vibronic coupling, as detailed in Sec. II B 2.

Changes in electronic configuration are coupled with nuclear motion of both the guest molecules and the host crystal. Vice versa, thermal occupation of host phonons will result in scattering events and induce homogeneous broadening. We estimate the molecular motion with the foundation machine learning potential MACE-OFF [39] including dispersive interactions. This affordable approach allows us (1) to deposit emitter molecules in the chosen host crystal and (2) to estimate the coupling strength between local vibrations (to be excited through Franck-Condon transitions) and the delocalized phonons.

1. Emitter embedding

The ideal position and orientation of the guest molecule in a host crystal is determined by Gibbs free energy. Given that most applications are implemented at temperatures near zero Kelvin, we discard the entropic components and deposit the guest molecules such that the formation energy is minimal [40]. To estimate the formation energy of the guest-host complex, an emitter is randomly inserted into a $5 \times 5 \times 5$ anthracene supercell with two anthracene molecules per unit cell. We remove then anthracene molecules that would be overlapping with the inserted emitter. A minimum of two and

a maximum of five molecules were removed when generating the structures. For each emitter and each number of molecules removed, up to 25 structures were generated, or alternatively until 10 000 trials had been made. A detailed algorithm can be found in Sec. II of the Supplemental Material [25]. The combined guest-host complex, as well as the emitter and host individually, are relaxed using MACE-OFF, and the formation (or binding) energy is calculated [41]:

$$E_{\text{bind}} = E_{\text{complex}} - E_{\text{emitter}} - \frac{N_{\text{atoms}}^{\text{complex}} - N_{\text{atoms}}^{\text{emitter}}}{N_{\text{atoms}}^{\text{supercell}}} E_{\text{supercell}}. \tag{1}$$

Here, E_{complex} is the energy of the combined complex (host with embedded molecules), E_{emitter} is the energy of the emitter only, and $E_{\text{supercell}}$ is the energy of the pure host (full supercell). The last contribution accounts for the molecules removed from the supercell to provide space for the embedded emitter molecule. A structure with the lowest formation energy is deemed the most stable configuration of the complex for the specific combination.

2. Vibronic coupling

A key merit of an SPE is the fraction of photons emitted into the 0-0-zero-phonon line, where no vibration or phonon is excited, and which is often also referred to as zero-phonon line, relative to the overall emission. This quantity, also referred to as the branching ratio, is more intricate in the case of single molecules embedded in a host material. First, the coupling of the electronic transition from $S_1 \rightarrow S_0$ with the vibrational modes of the guest molecule gives rise to a series of vibronic zero-phonon-line transitions, described by the canonical Franck-Condon physics. Second, the electronic transition of the guest molecule also couples to the vibrational modes (phonons) of the host material, described by Debye-Waller physics, which leads to phonon sidebands associated with each of the zero-phonon lines (also known as Debye-Waller factor). Thus, the branching ratio of the pure electronic component of the $S_1 \rightarrow S_0$ transition is diminished by both the coupling to the vibrational modes of the molecule and to the phonons of the host.

The strength of the 0-0-zero-phonon line, i.e., no vibrational and no phononic excitation under emission, can be theoretically estimated from the Huang-Rhys factors of the entire supercell. This requires (1) calculation of the vibrational modes of the combined system and (2) relaxation of the entire supercell using time-dependent DFT calculations for the excited S_1 state—an extremely costly undertaking for large supercells. In this work, we propose an approximate metric for the branching ratio that is inspired by a combination of molecular Huang-Rhys factor and the probability of scattering in and out of the corresponding mode. First, the eigenmodes of the emitter-host complex and the eigenmodes of the emitter molecule are calculated with MACE-OFF.

The vibrational mode (ν_i) of the isolated emitter can be projected into the basis of vibrational-phononic modes (j) of the entire emitter-host complex, defining the diagonal components of the reduced density matrix:

$$\rho_j^{\nu_i} = tr_j\left(\left|\nu_{i,\text{emitter}}^{\text{isolated}}\right\rangle \left\langle \nu_{i,\text{emitter}}^{\text{isolated}}\right| \right) = \left| \left\langle \nu_{i,\text{emitter}}^{\text{isolated}} \middle| \nu_{j,\text{emitter}}^{\text{embedded}} \right\rangle \right|^2, \tag{2}$$

where the integration appears only over nuclear coordinates associated with the emitter molecule. The elements $\rho_j^{\nu_i}$ range between 0 (no overlap) and 1 for a vibrational mode ν_i that is perfectly represented in the complex mode j. The von Neumann entropy of this projection,

$$S_i^P = -\sum_j^{N_{\text{complex}}} \rho_j^{\nu_i} \ln \rho_j^{\nu_i}, \tag{3}$$

provides a measure of how broad the overlap distribution is between local vibrations and the modes of the emitter-host complex. In other words, S_i^P quantifies the degree of mode delocalization [42] and the strength of scattering between the phononic background and local distortions. Equation (3) is loosely related to the canonical entropy of the scattering into and out of the vibrational mode ν_i [43]. The vibrational mode of an embedded molecule that is perfectly isolated from the phononic background will feature minimal scattering and $S_i^P = 0$. A related approach has been explored in Ref. [44].

Finally, we need to account for the relative contribution of each molecular eigenmode to the optical emission process, i.e., the vibronic coupling. We perform excited-state calculations using TDDFT of the emitter and relax their excited-state geometry (resulting in a Stokes shift between absorption and emission wavelength). The relaxed excited state configuration is then inserted into a ground-state DFT calculation to obtain the forces F, which are then projected onto the eigenmodes of the isolated emitter:

$$g_i = \left| \left\langle F \middle| \nu_{\text{emitter},\,i}^{\text{isolated}} \right\rangle \right|^2. \tag{4}$$

In other words, a ground-state energy calculation with the relaxed excited-state geometry gives rise to vibrational forces under emission. Within the double-harmonic approximation, the equivalent vibrational force would be obtained when inserting the S_0 geometry into the excited S_1 state calculation. We have decided here to perform the entire relaxation in order to provide emission wavelengths and Huang-Rhys factors for validation.

The weighted entropy, from here on called the vibronic coupling entropy, is then defined as

$$\mathcal{S}^{\text{VC}} = \sum_i^{N_{\text{modes}}^{\text{isolated}}} g_i S_i^P. \tag{5}$$

A low $\mathcal{S}^{\text{VC}}$ implies therefore that local vibrational modes are only weakly displaced after optical emission, and the modes that are displaced couple weakly to the phononic background.

Appendixes A and B demonstrate the strong correlation between common Huang-Rhys factors and force projection, the agreement between DFT and MACE-OFF, and the impact of entropic weighting. Appendix A provides additional indications why $\mathcal{S}^{\text{VC}}$ is preferable to direct estimates of the Franck-Condon related Huang-Rhys factors.

3. Spin transitions

Intersystem crossing, i.e., the transition between singlet and triplet space, plays a central role in the design of SPEs. Large intersystem crossing combined with a notable triplet lifetime will severely hamper the utilization of the emitter

since any transition into the triplet manifold blocks the desired $S_1 \leftrightarrow S_0$ emission over the phosphorescence lifetime. On the other hand, an appropriate yield of intersystem crossing is a desired feature for quantum memory applications, organic light-emitting diodes [45], and marker molecules used in biology [46]. In this work, we will use the magnitude of SOC as a primer for intersystem-crossing probabilities. All such SOC calculations are for isolated molecules, but it should be emphasized that the host material can mediate intermolecular intersystem-crossing channels, a subject for future studies [47,48].

The ideal feature in our approach contains information that correlates with relevant physical process, yet remains computationally simple enough to allow evaluation for many possible guest and host candidates. We define three different measures for intersystem-crossing events, all represented by spin-orbit couplings calculated using TDDFT. Kasha's rule suggests that intersystem transitions will always appear from the lowest excited state of a spin manifold and our focus will be set on the corresponding spin-orbit couplings. First, we estimate exothermic transition events from the excited $|S_1\rangle$ singlet state into the manifold of lower energetic triplet states (up to the tenth excited triplet state)

$$\mathrm{SOC} = \sqrt{\sum_i^{10} \begin{cases} |\langle T_i| H_{\mathrm{SO}} |S_1\rangle|^2, & E_1^S > E_i^T, \\ 0, & \text{else}, \end{cases}} \tag{6}$$

with spin-orbit Hamiltonian H_{SO}. This provides a simple figure of merit for energetically favorable, and thus intuitively dominant, transitions. We note that this is a simplified measure and that resonant singlet-to-triplet transitions are more likely than strict transition into the T_1 state [49]. Moreover, the reverse and ground-state SOC (GS SOC) may be defined as

$$\mathrm{rSOC} = \sqrt{\sum_i^{10} \begin{cases} |\langle T_i| H_{\mathrm{SO}} |S_1\rangle|^2, & E_1^S < E_i^T, \\ 0, & \text{else} \end{cases}} \tag{7}$$

and

$$\mathrm{GS\ SOC} = |\langle T_1| H_{\mathrm{SO}} |S_0\rangle|. \tag{8}$$

The reverse spin-orbit coupling (rSOC) estimates endothermic processes, i.e., transitions into higher-energetic triplet states. Such processes are deemed dominant in specific guest-host systems [47]. Finally, the phosphorescence rate correlates with the SOC from the lowest triplet state into the singlet ground state.

The ideal SPE for photon generation would feature low SOC and rSOC but notable GS SOC. While none of the three measures is redundant, the SOC is most relevant for SPE applications, and all SOCs tend to be closely linked via the symmetry of the electronic system.

III. RESULTS AND DISCUSSION

The merits defined in the previous section provide us with a set of indicators for the performance of a molecular emitter. Let us now begin our hunt for molecular quantum light-matter interfaces.

A. Similiarity

From Fig. 2(b), we see that most of the known emitters are localized to a single region in the overall parameter space of the dataset. In other words, most known emitters have similar molecular structures. This is a first indication that filtering potential emitter candidates by molecular structure could be a meaningful approach. In fact, the notion that similar molecules share similar properties is a cornerstone in the field of cheminformatics, wherein it is sometimes referred to as the *similar property principle* [50]. Furthermore, when filtering the dataset with respect to DBT, we can see in Fig. 2(c) that most structures have a Tanimoto index under 0.4. In fact, 99.9% of structures in our dataset has a Tanimoto index of 0.4 or lower. Excluding the reference molecule, only 45 molecules (0.026%) have a Tanimoto index of 0.5 or higher, and only two of those molecules (COD IDs: 2000909 and 4127216) have a Tanimoto index over 0.85. It is therefore sensible to begin the search for candidates in the vicinity of DBT, all hydrocarbons with the exception of diacenaphtheno[7,8-b:7^1, 8^1-d]thiophene. We would like to add that outliers in Fig. 2(b) suggest that future global optimization strategies should include an explorative component.

B. Characterizing emitter candidates

The 15 highest-scoring emitters of the similarity analysis are selected for further calculations and evaluations; all structures are visualized in Fig. S4 of the Supplemental Material [25]. A glance over Fig. 3 reconfirms that structures with higher similarity (>0.8) have comparable performance, while a low Tanimoto index (<0.6) does not allow any conclusion. The collected information with molecular identifiers and rotary strength can be found in Table I. We begin our detailed evaluation with the most fundamental optical features, i.e., brightness and wavelength [Figs. 3(a) and 3(b)]. The excitation wavelengths are positively correlated with the Tanimoto index since a comparable size of the conjugated π system implies a higher similarity in SMILE strings. Oscillator strength is largely uncorrelated with our chosen similarity measure but tends to be of comparable size within our domain of investigation. If we compare the highlighted points in Figs. 3(a) and 3(b), we notice a clear correlation: A longer wavelength results in a weaker oscillator strength.

DBT and terrylene defend their place as top emitters via exceptionally small vibronic couplings. The SOC, rSOC, and the GS SOC of the emitter candidates are plotted against their respective Tanimoto indices in Figs. 3(e)–3(g). Notice that DBT underperforms in SOC and rSOC, while terrylene is nearly ideal. This might allow two different interpretations: (1) A good emitter should overperform in a wide range of features; success in only one relevant attribute is insufficient. (2) Metrics that utilize the spin-orbit coupling elements are insufficient as reasonable primers for intersystem-crossing rates. Our goal is therefore to find a suitable trade-off for new candidates, where the relative weight between features depends on the specific task at hand and remains to be identified. The impact of SOC will be discussed separately for all predictions in order to draw conclusions with and without consideration of SOC. We have observed no clear correlation between emitter size and SOC value.

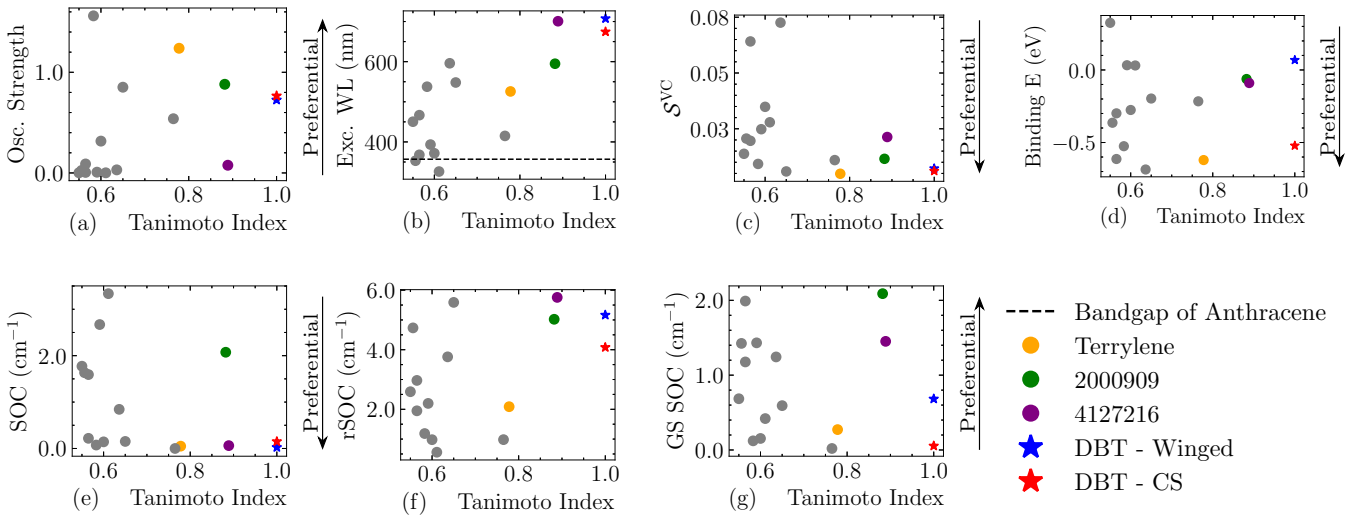

FIG. 3. Microscopic analysis: Figures showing (a) the oscillator strength of emission, (b) the excitation wavelength (absorption), (c) the vibronic coupling entropy (with anthracene as host), (d) the binding energy (with anthracene as host), (e) the intersystem crossing, (f) the reverse intersystem crossing, and the (g) ground-state intersystem crossing plotted for emitters against their Tanimoto index. The reference emitter (DBT) and three especially interesting candidates (terrylene, 2000909, 4127216) are highlighted.

TABLE I. Microscopic analysis: Tanimoto index, oscillator strength, excitation wavelength, absorption circular dichroism, intersystem crossing, reverse intersystem crossing, ground-state intersystem crossing, vibronic coupling entropy, and binding energy of investigated emitters in anthracene. Details on the two computationally identified conformers of DBT (DBT-centrosymmetric, short DBT-CS and DBT-winged) are discussed in Sec. III of the Supplemental Material [25]. The structures of all candidates are illustrated in Fig. S4 of the Supplemental Material [25].

Name/COD ID	Tanimoto index	f_{osc}^{abs}	f_{osc}^{em}	$\lambda_{emitter}^{abs}$ (nm)	$\lambda_{emitter}^{em}$ (nm)	R (10^{40}cgs)	SOC (cm^{-1})	rSOC (cm^{-1})	GS SOC (cm^{-1})	$\mathcal{S}^{VC}$	E_{bind} (eV)	Refs.
DBT- CS[a]	1.00	0.77	0.77	674.6	706.0	−0.17	0.15	4.08	0.05	0.0061	−0.52	[17]
DBT-winged[a]	1.00	0.73	0.73	707.5	738.6	0.17	0.02	5.17	0.68	0.0072	0.07	[51]
Hexa-peri-hexabenzo[7] helicene/ 4127216	0.89	0.13	0.08	700.8	1100.7	276.16	0.06	5.76	1.45	0.0213	−0.09	[52]
Tetrabenzo[de,hi,op,st] pentacene/2000909	0.88	0.84	0.88	594.9	630.3	73.96	2.07	5.02	2.09	0.0115	−0.06	[53]
Terrylene/–[b]	0.78	1.17	1.24	525.8	551.7	0.37	0.05	2.09	0.27	0.0048	−0.62	[47]
Perylene/–[c]	0.77	0.50	0.54	415.0	447.8	−0.05	0.0	0.98	0.02	0.0110	−0.22	[54]
Benzodiphenanthrobisanthene (BDPB)/-[c]	0.65	0.79	0.85	548.1	589.4	−3.98	0.15	5.59	0.59	0.0059	−0.20	[55]
Diindenozethrene/1555531	0.64	0.52	0.03	596.0	820.3	59.20	0.84	3.76	1.24	0.0726	−0.69	[56]
–/4062916	0.61	0.00	0.00	326.2	353.1	0.46	3.34	0.56	0.42	0.0279	0.03	[57]
Benzo[*ghi*]perylene/1554212	0.60	0.26	0.32	371.7	403.0	4.28	0.14	0.98	0.15	0.0348	−0.28	[58]
Naphtho[1,2-i]pentahelicene/7155241	0.59	0.00	0.01	393.4	426.1	5.24	2.67	2.20	1.43	0.0248	0.03	[59]
2.3.7.8-di-(peri-naphthylen)-pyren (DPNP)/–[c]	0.58	1.46	1.56	537.6	562.6	0.75	0.07	1.19	0.12	0.0093	−0.53	[60]
Diacenaphtheno[7,8-b:7^1, 8^1-d]thiophene/2203348	0.57	0.19	0.09	466.7	597.2	6.25	0.22	2.97	1.18	0.0641	−0.30	[61]
–/4118733	0.57	0.01	0.01	367.9	399.5	−0.34	1.59	1.95	1.99	0.0196	−0.61	[62]
Dibenzo[g,p]chrysene/4107152	0.56	0.01	0.02	353.1	386.4	1.67	1.63	4.73	1.42	0.0206	−0.36	[63]
Pentabenzo[a,d,g,j,m] coronene/4037514	0.55	0.00	0.00	450.4	482.0	−0.06	1.77	2.59	0.68	0.0138	0.32	[64]

[a]Reference emitter.
[b]Known emitter in anthracene.
[c]Known emitter in other host.

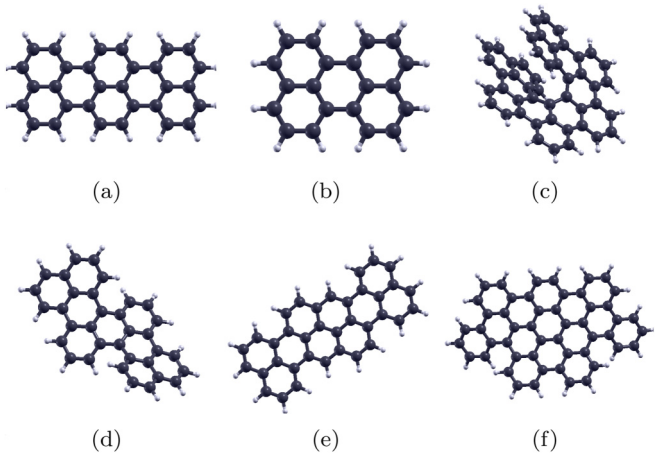

FIG. 4. Emitter candidates: (a) The recovery of terrylene serves as validation for our exploitation strategy. (b) Perylene illustrates the relevance of intersystem-crossing events. (c) Hexa-peri-hexabenzo[7]helicene (COD ID: 4127216) is a chiral emitter with decent performance at long wavelength, a promising building block for chiral photonics. (d) Tetrabenzo[de,hi,op,st]pentacene (COD ID: 2000909) follows a similar synthesis route as DBT and presents an ideal validation point for future experimental studies. Additional validation candidates: (e) DPNP and (f) BDPB.

C. Promising candidates

As a result, we have chosen to highlight six emitters: terrylene and perylene (known in anthracene [47,65]), 2000909 (unknown), 4127216 (unknown), benzodiphenanthrobisanthene (BDPB) (unknown in anthracene), and di-(peri-naphthylen)-pyren (DPNP) (unknown in anthracene), all illustrated in Fig. 4.

1. Terrylene: A validation

Experimental research demonstrated that also terrylene performs well in anthracene [47]. We may thus place extra emphasis on observables for which terrylene and DBT exhibit similar behavior. Most notably, as shown in Fig. 3(c), terrylene and DBT have the lowest and second-lowest vibronic coupling entropy among all emitters embedded in anthracene, underlining the relevance of this measure. Furthermore, terrylene features an exceptionally high oscillator strength. Overall, our analysis correctly ranks terrylene as one of the best emitters in anthracene.

A second known guest in anthracene is the smaller-sized perylene. Smaller size reduces its oscillator strength while increasing $\mathcal{S}^{\mathrm{VC}}$—perylene is therefore clearly a worse SPE than terrylene. As we shall see in the next section, this moves perylene almost horizontally along the first principal component (PC1) closer to the decision boundary between "good" and "bad" emitters. Perylene and terrylene illustrate therefore nicely the underlying physics; i.e., strong single-photon emission demands foremost a bright and sharp zero-phonon line. Any coupling to local vibrations or delocalized phonons will negatively impact the performance. However, more challenging to predict are intersystem crossings mediated by the host crystal. Direct SOC is vanishingly small in perylene (see Table I), yet experiments find sizable intersystem crossing that is likely mediated by the anthracene host [65]. By itself, this results in blinking, i.e., turning the SPE on and off, a feature present for terrylene in anthracene. In the case of perylene, however, the low GS SOC (see Table I) traps the system for such an extended time in the triplet manifold that it entirely prevents the utilization of perylene as SPE.

2. Exciting candidates

Out of the small number of investigated emitters, two are deemed especially interesting for further investigation: 2000909 and 4127216, pictured in Fig. 4. The vibronic coupling entropies of 2000909 and 4127216 are among the lowest of the emitters with transition energy smaller than the band gap of anthracene.

We can branch our discussion into two possible directions from here on. First, tetrabenzo[de,hi,op,st]pentacene (2000909) exhibits an oscillator strength higher than that of DBT at a wavelength domain just between terrylene and DBT—filling the gap between these two wavelength domains. The trade-off is larger SOC elements. It can be produced analogously to DBT [66] and serves therefore as an excellent first test candidate.

The second direction provides an alternative perspective than established approaches. Hexa-peri-hexabenzo[7]helicen (COD ID: 4127216) has a lower oscillator strength with longer wavelength and acceptable SOC couplings. Controlling the larger vibronic coupling entropy, also indicated by the notable Stark shift, will be more challenging. However, its emission is strongly chiral (see rotary strength R in Table I), which may offer unique opportunities to design inherently chiral SPEs. Such chiral SPEs would be ideal for near-field chiral sensing and serve as a promising component for chiral photonics [67–70].

3. Validation candidates

Other well-tested molecular SPEs are DPNP and BDPB, although no studies have been conducted in anthracene. Those two candidates might serve as excellent validation cases for our approach, as they can be easily synthesized, are well studied, and their relative performance in experiment would support the theoretical development.

D. Principal components and classification

No emitter is perfect in every metric. Which trade-off is ideal depends heavily on the desired task of the emitter. Central challenges are therefore (1) to identify the relative importance of the various microscopic features (vibronic coupling, spin-orbit coupling, oscillator strength) for a specific task and (2) to judge whether a candidate might be worth investigating in further detail based on readily accessible information on existing databases.

Our focus in this work is solely on optical coherence. Four features stand out in their relevance: (1) high oscillator strength, (2) an absorption wavelength longer than that of anthracene to allow optical excitation of the emitter, (3) low intersystem crossing, and (4) a low vibronic coupling entropy. An emitter that performs better than average in all those criteria will be considered "good," all others "bad."

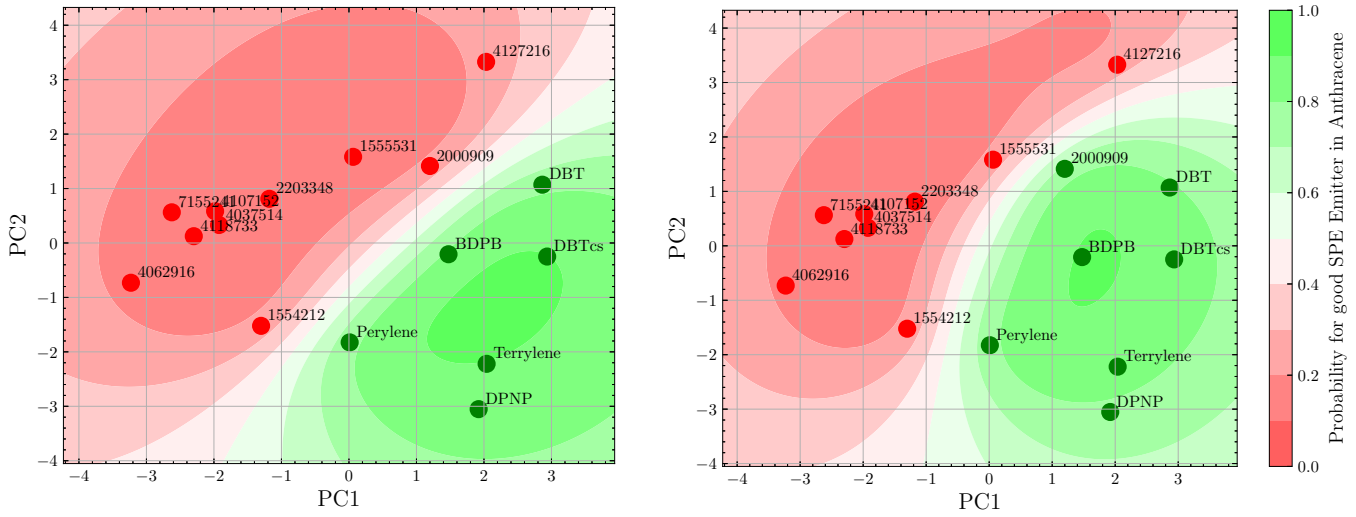

FIG. 5. Gaussian process classification: The ten-dimensional space {Tanimoto index, $f_{\text{osc}}^{\text{abs}}$, $f_{\text{osc}}^{\text{em}}$, $\lambda_{\text{emitter}}^{\text{abs}}$, $\lambda_{\text{emitter}}^{\text{em}}$, SOC, rSOC, GS SOC, $\mathcal{S}^{\text{VC}}$, E_{bind}} (Table I) is reduced to the two dominant principal components (PC1 and PC2). An emitter is labeled "good" if it performs better than average in the most relevant categories (left) $f_{\text{osc}}^{\text{em}} > \overline{f_{\text{osc}}^{\text{em}}}$, $\lambda_{\text{emitter}}^{\text{abs}} > \lambda_{\text{anthracene}}^{\text{abs}}$, $\mathcal{S}^{\text{VC}} < \overline{\mathcal{S}^{\text{VC}}}$, and SOC $< \overline{\text{SOC}}$. Right-hand side, the SOC is ignored in the labeling; i.e., we do not include any primer for intersystem crossing.

Such *labels* allow the use of machine learning approaches for classification and pattern recognition. To illustrate the results and understand the inherent relevance of each component, we can calculate the two dominant principal components (PC1 and PC2) from the ten-dimensional space in Table I (except the rotary strength). These data, including the above labels color-coded good (green) and bad (red), are visualized in Fig. 5. We then apply a Gaussian process classifier via the PYTHON package scikit-learn [31,71]. The kernel is composed of a constant kernel (1.0 and 10^{-3} to 10^{3}) and a radial basis function kernel with length scale (1.0 and 10^{-2} to 10^{2}). The Gaussian process classifier aims to find an optimal separation into good and bad candidates, drawing a decision boundary between both clusters. A more intense green (red) color corresponds to a higher probability of a potential candidate being good (bad). Left (right) side of Fig. 5 includes (excludes) the SOC in the labeling process. Please note that we train the Gaussian process classifier model here with all data points to gather a fundamental understanding. Future research will focus on expanding, validating, and testing this model more rigorously.

DBT and terrylene [Fig. 4(a)] are correctly identified as good emitters, consistent with existing experimental observations. DPNP is identified as another strong candidate, while perylene and BDPB are shifted along $-$PC1 close to the decision boundary. Consideration of the SOC degree of freedom has here a substantial impact, with a more restrictive prediction when considering SOC (left). Host-mediated intersystem-crossing events are missing in our current study, resulting in the false labeling of perylene as promising emitter. The candidates 2000909 and our chiral emitter (4127216) are shifted diagonally along $-$PC1 + PC2 and remain close to the decision boundary. Especially, 2000909 should represent an ideal candidate for experimental validation as its synthesis is similar to DBT. All other candidates have a low probability to perform on a comparable level. Our chiral emitter (4127216) is of special interest as it might represent an inherently chiral SPE—an extremely interesting direction for sensing and chiral photonics.

IV. CONCLUSION AND OUTLOOK

Quantum light-matter interfaces are a key building block to drive the development of next-generation quantum technology. Molecules are a promising platform due to their brightness, low inhomogeneous broadening, and low cost. Most strikingly, their enormous flexibility allows us to tailor the features of an SPE to a specific task at low cost. Thus far, this potential has not been leveraged, as a lack of predictive theory and resource-intensive experiments prohibited systematic exploration of the available chemical space.

Our work tackles this limitation with a combination of database research, microscopic analysis, and initial steps to apply machine learning. First, we extract SMILES from the Crystallography Open Database and isolate structures similar to DBT, one of the top performers in anthracene. Second, various critical features, such as oscillator strength, vibronic coupling, and spin-orbit coupling strength, are calculated using density-functional theory and machine learning potentials. Finally, those quantities serve as figures of merit for optical coherence and are then ultimately used to classify possible candidates.

The first instance of this methodology identifies up to six alternatives for DBT in anthracene. First, it correctly predicts terrylene, a known match, which confirms that our predictions are meaningful. Among all the emitters considered, tetrabenzo[de,hi,op,st]pentacene (COD ID: 2000909) performs well above average for all metrics and can be synthesized following a similar route as DBT—an excellent candidate to validate theory with moderate effort. Tetrabenzo[de,hi,op,st]pentacene has an excitation wavelength

that lies between those of terrylene and DBT, thereby filling the gap in that spectral regime. An especially exciting emitter of interest is hexa-peri-hexabenzo[7]helicene (COD ID: 4127216), which performs comparably to tetrabenzo[de,hi,op,st]pentacene across all categories, with the exception of oscillator strength. Despite its lower oscillator strength, hexa-peri-hexabenzo[7]helicene merits further investigation due to its distinctly chiral character, which could prove valuable for applications in chiral photonics and chiral near-field sensing [70].

In addition to the identification of promising molecular SPE candidates, our approach provides additional context to the established empirical guidelines for "good" SPEs. Our figures of merits, e.g., the vibronic coupling entropy, are computationally affordable, and the consistency in preliminary classification underlines their usefulness. Although additional work will be required for a reliable assessment, our current results support the established development guidelines.

Based on this article, various extensions can be incorporated in the future. One is to connect our current approach to the true optical coherence of the SPE, i.e., developing a predictive framework for the photon statistics and coherence of the emitter. This would involve a consistent treatment of the quantum system and structured bath in the presence of nanophotonic elements—a demanding but reachable goal due to recent progress in the efficient description of many-body system-bath dynamics [72,73]. Validating the suitability of our predicted chiral SPE could have considerable implications for chiral photonics [20,74] and chiral sensing. A second development direction is to establish a global exploration strategy. We focused on the chemical neighborhood around known matches, an idea loosely based on the *similar property principle* that exploits our existing knowledge. However, combinations with (multifidelity) Bayesian optimization suggest that our technique can be extended with a global, i.e., explorative, component [75,76].

This article lays the first stone for a predictive design strategy for molecular quantum light-matter interfaces, leveraging their chemical flexibility to identify ideal quantum systems that can be tailored to a wide range of tasks.

ACKNOWLEDGMENTS

The authors acknowledge funding from the Nano Area of Advance, an Excellence Cluster of Chalmers. C.S. acknowledges funding from the Horizon Europe research and innovation program of the European Union under the Marie Skłodowska-Curie Grant Agreement No. 101065117. R.M.G. acknowledges funding from the Swedish Research Council (VR starting Grant No. 2022-03350), the Olle Engkvist Foundation (Grant No. 229-0443), the Royal Physiographic Society in Lund (Horisont), and the Knut and Alice Wallenberg Foundation (Grant No. 2023.0087). D.W. acknowledges financial support from Germany's Excellence Strategy—Cluster of Excellence Matter and Light for Quantum Computing (ML4Q) EXC 2004/1-390534769 and the European Union (ERC, MSpin, 101077866). The authors acknowledge the TU Wien Bibliothek for financial support through its Open Access Funding Programme.

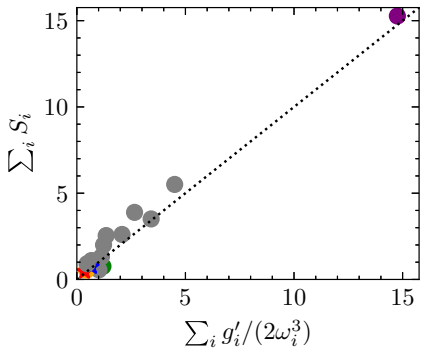

FIG. 6. Equivalence of metrics: Sums over Huang-Rhys factors defined in Eq. (B2) vs reweighted sums over g_i'. Each g_i' is here the $1/\sqrt{M}$ reweighted equivalent to g_i (see text). The small remaining deviations arrive from the inconsistencies between normal modes obtained with MACE-OFF and shifts and forces calculated with TDDFT.

DATA AVAILABILITY

The data that support the findings of this article are openly available [77]. All input files and scripts to produce, postprocess, and evaluate/plot the data are available in the Zenodo repository.

APPENDIX A: COMPARISON OF HUANG-RHYS FACTOR AND VIBRONIC COUPLING ENTROPY $\mathcal{S}^{\mathrm{VC}}$

The local vibrational contribution to the vibrational coupling entropy is closely related to the Huang-Rhys factor, supporting the evaluation in this article. In principle, the two can be brought into exact agreement [44] by reweighting the force used in the projection with $1/\sqrt{M}$ and summing the frequency- weighted projections as

$$g_i' = \left|\left\langle M^{-0.5} F \middle| \nu_{\mathrm{emitter},i}^{\mathrm{isolated}} \right\rangle\right|^2, \tag{A1}$$

$$\sum_i S_i = \sum_i \frac{1}{2\omega_i^3} g_i'. \tag{A2}$$

Figure 6 illustrates that this agreement is in our case only true up to small deviations due to the inconsistency between the equilibrium shift and forces, calculated with TDDFT, and the normal modes, which are obtained with MACE-OFF. We note that the vibronic coupling entropy is therefore closely related to the Huang-Rhys factors, yet allows to account for entropic effects due to the phonon bath. This is further emphasized in Figs. 7(a) and 7(b) , which show clear positive correlations. Coupling between local vibrational and global phononic modes plays a subordinate role in the entropy but provides valuable information on temperature-sensitive homogeneous broadening due to phonon scattering. Those effects result in slight shifts among the datapoints and ensure DBT and terrylene are the best performing guests in anthracene (see Fig. 8).

We present in Fig. 7(c) an alternative figure of merit. Focusing solely on the Franck-Condon contribution of the molecule and its coupling to the extended supercell, we can

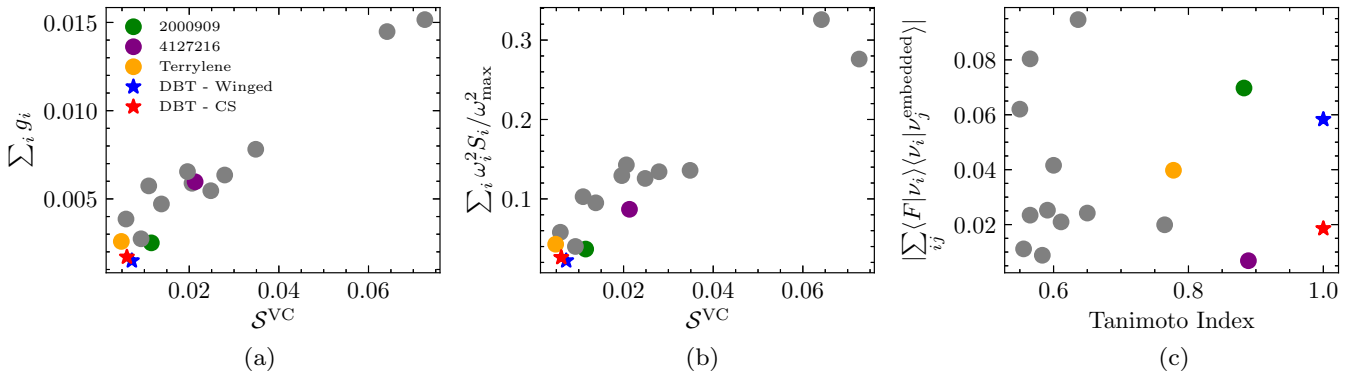

FIG. 7. (a) Sum over emission-force projected on vibrational modes vs vibronic coupling entropy: A clear positive correlation is found with only small variations due to the coupling between local vibrations and global phonons. (b) Weighted Huang-Rhys sum vs vibronic coupling entropy: A clear positive correlation is found. This comes as no surprise due to the conceptual equivalence illustrated in Figs. 6 and 7(a). The entropy includes, however, an additional estimate of the coupling between local vibrational and global phononic modes, thus accounting for scattering events that dictate temperature progression (see also Fig. 8). The maximal frequency is here chosen to be the maximal frequency of all investigated emitters. (c) Absolute of the sum over Franck-Condon vibronic coupling with combined system: The Franck-Condon vibronic coupling, i.e., the vibrational force due to electronic transition, is projected directly onto the entire supercell. This is closely related to the Huang-Rhys factor of the combined guest-host system if we ignore the displacement of the phonons due to the charge-displacement of the molecule. In other words, the forces include only molecular contributions that scatter with the phonons of the superlattice. The smaller the value, the stronger the 0-0 phonon-line and the more useful the molecular guest as SPE. See text for further details.

estimate the strength of the 0-0 phonon line as

$$\left|\sum_j \left\langle F \middle| \nu_j^{\mathrm{embedded}} \right\rangle\right|. \tag{A3}$$

Inserting unity $\sum_i^{\mathrm{isolated}} |\nu_i\rangle\langle\nu_i|$, we obtain the equivalent formulation

$$\left|\sum_{ij} \langle F|\nu_i\rangle \left\langle \nu_i \middle| \nu_j^{\mathrm{embedded}} \right\rangle\right|. \tag{A4}$$

A larger value implies larger scattering into vibrational modes, i.e., a weaker zero-phonon line. Overall, a decent agreement can be observed with $\mathcal{S}^{\mathrm{VC}}$ (compare to Fig. 3). However, especially the established emitters Terrylene and DBT perform considerably worse than in the vibronic coupling entropy. Given that those are established candidates, it seems reasonable to conclude that the entropic estimate of the scattering in $\mathcal{S}^{\mathrm{VC}}$ is partially correcting for the missing Debye-Waller term and thus provides an overall better description of the involved physics. Recall that relaxation of the entire supercell with TDDFT is computationally prohibitive. Future work will explore the agreement with embedding approaches such as ONIOM [78] and others [44].

APPENDIX B: HUANG-RHYS FACTORS—ORCA VERSUS ASE/MACE-OFF

A vibrational system settled in equilibrium will be pushed out of the latter after an electronic transition of the system. This is a natural consequence of the (nearly) instantaneous changes in electronic structure, i.e., the potential-energy surface, and the following forces acting on the nuclei. This effect is known as vibronic coupling and often described in the Franck-Condon approximation, which assumes indeed an

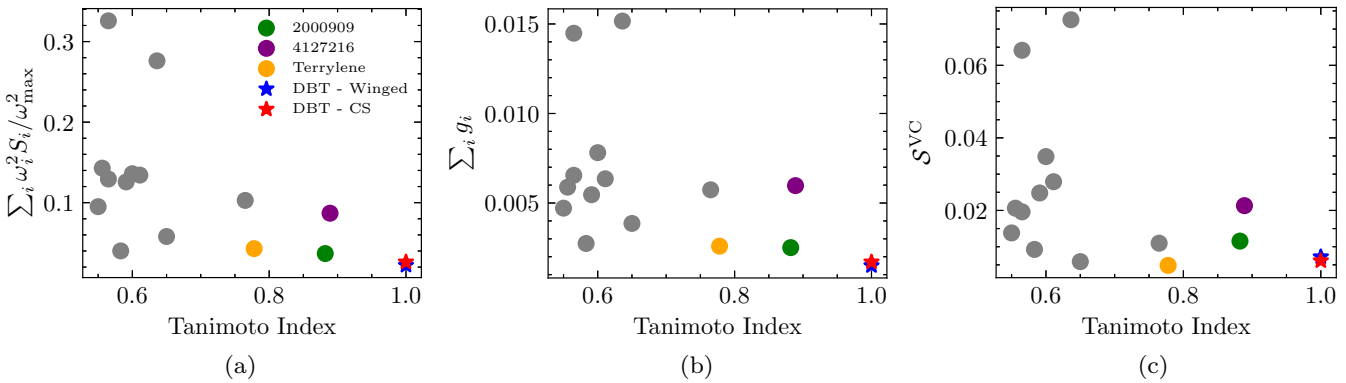

FIG. 8. Comparison for vibronic coupling estimates: (a) Sum over weighted Huang-Rhys factors, (b) sum over emission-force projected onto vibrational modes, and (c) vibronic coupling entropy. All three quantities are well correlated, with a slightly larger change between panels (b) and (c), where local-to-phonon-coupling is incorporated. Local vibronic features seem thus more relevant for the overall performance compared to phononic effects, while the latter will impact especially the temperature sensitivity and homogeneous linewidth broadening.

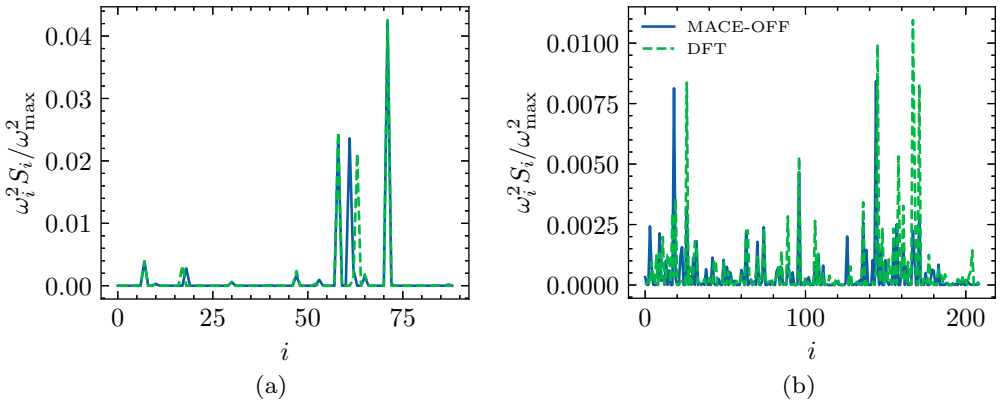

FIG. 9. Benchmark DFT vs MACE-OFF: Weighted Huang-Rhys factors are calculated using the same GS and excited-state configurations (obtained with DFT) but different vibrational modes. (a) Perylene: The planar structure and simplicity of perylene are accurately captured by MACE-OFF. Merely observable are slight switches in the modeordering and minor deviations in the strength of Huang-Rhys factors. (b) Chiral emitter: While the overall structure is well preserved, we can identify deviations in the low- and high-frequency domains. Notable normal modes in the low-frequency domain involve springlike oscillations that are mediated via the weak van der Waals interaction along the helical axis. Notable high-energy modes are dominated by hydrogen oscillations. It is not surprising that MACE-OFF performs worse in predicting weaker bonds, yet the overall structure is well preserved, which indicates that our conclusions remain valuable even for more complex emitter-host pairs.

instantaneous transition and harmonic potentials. The displacement between the two vibrational harmonic potentials is often quantified in the dimensionless Huang-Rhys factors.

In a real system, many more than just a single vibration exist and the Huang-Rhys factors are defined per eigenmode. The latter are routinely calculated by electronic structure codes such as ORCA and the Atomic Simulation Environment (ASE). Vibrational eigenmodes u_k and the corresponding frequencies ω_l are obtained from the diagonalization of the Hessian:

$$[\mathbf{M}^{-1/2}\mathbf{K}\mathbf{M}^{-1/2}]\mathbf{u}_k = \omega_k^2\mathbf{u}_k. \tag{B1}$$

If we assume the same harmonic potential in the ground and excited state, which is often a decent approximation for the relatively rigid molecular systems investigated here, then the shift in equilibrium position $\mathbf{\Delta} = \mathbf{M}^{1/2}(\mathbf{R}_{S0} - \mathbf{R}_{S2})$ between the harmonics defines the strength of interaction. Finally, the Huang-Rhys factor $S_k = \omega_k\Delta_k^2/2$ is calculated from

$$S_k \equiv \frac{\omega_k}{2}((\mathbf{R}_{S0} - \mathbf{R}_{S1}) \cdot \sqrt{\mathbf{M}} \cdot \mathbf{u}_k)^2. \tag{B2}$$

We then compare the frequency-weighted Huang-Rhys factors $\omega_k^2 S_k/\omega_{\max}^2$, with $\omega_{\max}$ the maximal vibrational frequency, against our alternative figures of merit. The frequency-weighted Huang-Rhys factors can be seen as effective system-bath coupling strength [72] and are therefore more intuitive from a quantum-optical perspective. Normalization with the highest mode frequency $\omega_{\max}$ is chosen to keep the results unitless.

In practice, special attention has to be given to the normalization of the eigenmodes, especially their mass weighting. While the eigenmodes provided by ASE are normalized modes weighted by the inverse of the mass $(\sqrt{\mathbf{M}})^{-1} \cdot \frac{\mathbf{u}_k}{|\mathbf{u}_k|}$, ORCA provides mass-weighted modes that are normalized $\frac{(\sqrt{\mathbf{M}})^{-1}\cdot\mathbf{u}_k}{|(\sqrt{\mathbf{M}})^{-1}\cdot\mathbf{u}_k|}$. In order to utilize the normal modes in the basis-set expansion and the calculation of the Huang-Rhys factors, we require normalized modes in the form $\sqrt{\mathbf{M}} \cdot \frac{\mathbf{u}_k}{|\mathbf{u}_k|}$. Therefore, we reweight the modes obtained with ASE with $\mathbf{M}$ and the modes obtained from ORCA with $\frac{\mathbf{M}}{|\sqrt{\mathbf{M}}|}$. Figures 9(a) and 9(b) demonstrate the overall agreement between full TDDFT using ORCA and the usage of MACE-OFF with ASE. The normal modes $|\nu_i\rangle$ are obtained from ASE by taking the exported modes and weighting them with $\sqrt{\mathbf{M}}$; i.e., $|\nu_i\rangle$ provide an orthonormal basis.

We recently became aware of Ref. [79] that explores the usage of foundation machine learning models for the prediction of vibronic coupling in point-defect systems. While not overlapping with the objectives of this work, it further demonstrates the utility of foundation models for optical characterization.

[1] I. Aharonovich, D. Englund, and M. Toth, Solid-state single-photon emitters, Nat. Photon. **10**, 631 (2016).
[2] A. Aspuru-Guzik and P. Walther, Photonic quantum simulators, Nat. Phys. **8**, 285 (2012).
[3] M. L. Chan, T. J. Bell, L. A. Pettersson, S. X. Chen, P. Yard, A. S. Sørensen, and S. Paesani, Tailoring fusion-based photonic quantum computing schemes to quantum emitters, PRX Quantum **6**, 020304 (2025).
[4] P. A. Morris, R. S. Aspden, J. E. C. Bell, R. W. Boyd, and M. J. Padgett, Imaging with a small number of photons, Nat. Commun. **6**, 5913 (2015).
[5] D. G. Baranov, C. Schäfer, and M. V. Gorkunov, Toward molecular chiral polaritons, ACS Photon. **10**, 2440 (2023).
[6] J. Fojt, P. Erhart, and C. Schäfer, Controlling plasmonic catalysis via strong coupling with electromagnetic resonators, Nano Lett. **24**, 11913 (2024).

[7] C. Toninelli, I. Gerhardt, A. S. Clark, A. Reserbat-Plantey, S. Götzinger, Z. Ristanović, M. Colautti, P. Lombardi, K. D. Major, I. Deperasińska, W. H. Pernice, F. H. L. Koppens, B. Kozankiewicz, A. Gourdon, V. Sandoghdar, and M. Orrit, Single organic molecules for photonic quantum technologies, Nat. Mater. **20**, 1615 (2021).

[8] N. Tomm, A. Javadi, N. O. Antoniadis, D. Najer, M. C. Löbl, A. R. Korsch, R. Schott, S. R. Valentin, A. D. Wieck, A. Ludwig, and R. J. Warburton, A bright and fast source of coherent single photons, Nat. Nanotechnol. **16**, 399 (2021).

[9] J. Nobakht, A. Pscherer, J. Renger, S. Götzinger, and V. Sandoghdar, Hybridization of molecules via a common photonic mode, Proc. Natl. Acad. Sci. USA **122**, e2505161122 (2025).

[10] Y. Wang, H. Xu, X. Deng, T. C. H. Liew, S. Ghosh, and Q. Xiong, Topological single-photon emission from quantum emitter chains, npj Quantum Inf. **10**, 13 (2024).

[11] M. Rezai, J. Wrachtrup, and I. Gerhardt, Polarization-entangled photon pairs from a single molecule, Optica **6**, 34 (2019).

[12] J.-B. Trebbia, Q. Deplano, P. Tamarat, and B. Lounis, Tailoring the superradiant and subradiant nature of two coherently coupled quantum emitters, Nat. Commun. **13**, 2962 (2022).

[13] C. M. Lange, E. Daggett, V. Walther, L. Huang, and J. D. Hood, Superradiant and subradiant states in lifetime-limited organic molecules through laser-induced tuning, Nat. Phys. **20**, 836 (2024).

[14] Y.-S. Park, S. Guo, N. S. Makarov, and V. I. Klimov, Room temperature single-photon emission from individual perovskite quantum dots, ACS Nano **9**, 10386 (2015).

[15] S. Pezzagna, D. Rogalla, D. Wildanger, J. Meijer, and A. Zaitsev, Creation and nature of optical centres in diamond for single-photon emission—Overview and critical remarks, New J. Phys. **13**, 035024 (2011).

[16] S. Gupta, W. Wu, S. Huang, and B. I. Yakobson, Single-photon emission from two-dimensional materials, to a brighter future, J. Phys. Chem. Lett. **14**, 3274 (2023).

[17] A. A. L. Nicolet, C. Hofmann, M. A. Kol'chenko, B. Kozankiewicz, and M. Orrit, Single dibenzoterrylene molecules in an anthracene crystal: Spectroscopy and photophysics, ChemPhysChem **8**, 1215 (2007).

[18] S. Adhikari, R. Smit, and M. Orrit, Future paths in cryogenic single-molecule fluorescence spectroscopy, J. Phys. Chem. C **128**, 3 (2024).

[19] B. Gurlek and D. Wang, Small but large: Single organic molecules as hybrid platforms for quantum technologies, Phys. Rev. Res. **7**, 021001 (2025).

[20] D. Suárez-Forero, M. Jalali Mehrabad, C. Vega, A. González-Tudela, and M. Hafezi, Chiral quantum optics: Recent developments and future directions, PRX Quantum **6**, 020101 (2025).

[21] M. Quirós, S. Gražulis, S. Girdzijauskaitė, A. Merkys, and A. Vaitkus, Using SMILES strings for the description of chemical connectivity in the Crystallography Open Database, J. Chem. Inf. **10**, 23 (2018).

[22] D. Weininger, SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules, J. Chem. Inf. Comput. Sci. **28**, 31 (1988).

[23] N. Schneider, R. A. Sayle, and G. A. Landrum, Get your atoms in order—An open-source implementation of a novel and robust molecular canonicalization algorithm, J. Chem. Inf. Model. **55**, 2111 (2015).

[24] A. P. Bartók, R. Kondor, and G. Csányi, On representing chemical environments, Phys. Rev. B **87**, 184115 (2013).

[25] See Supplemental Material at http://link.aps.org/supplemental/10.1103/11bh-ywcc for results of our microscopic analysis plotted over average SOAP similarities, a detailed algorithm to the embedding strategy, further details to the two possible DBT structures and indications for their potential presence in recent experiments (including Ref. [80]), and a visualization of all investigated molecules.

[26] S. De, A. P. Bartók, G. Csányi, and M. Ceriotti, Comparing molecules and solids across structural and alchemical space, Phys. Chem. Chem. Phys. **18**, 13754 (2016).

[27] G. Landrum, *et al.*, RDKit: Open-source cheminformatics, Zenodo (2025), doi:10.5281/zenodo.18797641.

[28] H. L. Morgan, The generation of a unique machine description for chemical structures—A technique developed at Chemical Abstracts Service, J. Chem. Doc. **5**, 107 (1965).

[29] D. Rogers and M. Hahn, Extended-connectivity fingerprints, J. Chem. Inf. Model. **50**, 742 (2010).

[30] D. Bajusz, A. Rácz, and K. Héberger, Why is Tanimoto index an appropriate choice for fingerprint-based similarity calculations? J. Chem. Inf. **7**, 20 (2015).

[31] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and É. Duchesnay, Scikit-learn: Machine learning in Python, J. Mach. Learn. Res. **12**, 2825 (2011).

[32] W. Kohn and L. J. Sham, Self-consistent equations including exchange and correlation effects, Phys. Rev. **140**, A1133 (1965).

[33] F. Neese, Software update: The ORCA program system—Version 5.0, WIREs Comput. Mol. Sci. **12**, e1606 (2022).

[34] C. Lee, W. Yang, and R. G. Parr, Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density, Phys. Rev. B **37**, 785 (1988).

[35] A. D. Becke, Density-functional thermochemistry. III. The role of exact exchange, J. Chem. Phys. **98**, 5648 (1993).

[36] E. Caldeweyher, S. Ehlert, A. Hansen, H. Neugebauer, S. Spicher, C. Bannwarth, and S. Grimme, A generally applicable atomic-charge dependent London dispersion correction, J. Chem. Phys. **150**, 154122 (2019).

[37] B3LYP provides adequately precise predictions for conjugated systems, is widely employed, and remains computationally affordable. Dispersive interactions are relevant to predict the correct curvature of relaxed geometries and the def2-TZVPD basis is a safe yet cost-efficient choice for this level of theory. Predictions for DBT vertical absorption energies are much closer to experiment (785 nm) using B3LYP (675 nm) than using, e.g., wB97X-D4 (556 nm). A discussion on the example of DBT can be found in Table 2 of the Supplemental Material [25].

[38] M. E. Casida, Time-dependent density functional response theory of molecular systems: Theory, computational methods, and functionals, Theor. Comput. Chem. **4**, 391 (1996).

[39] D. P. Kovács, J. H. Moore, N. J. Browning, I. Batatia, J. T. Horton, Y. Pu, V. Kapil, W. C. Witt, I.-B. Magdău, D. J. Cole, and G. Csányi, MACE-OFF23: Transferable machine learning

force fields for organic molecules, J. Am. Chem. Soc. **147**, 17598 (2025).

[40] A fully rigorous treatment would require explicit modeling of crystal growth with embedded molecule and continuous NPT cooling with quantum-mechanical corrections at low temperatures. To this date, this remains computationally prohibitive if repeated for a large number of candidates. We have therefore decided to relax the experimentally observed low-temperature phase. Although adiabatic cooling may trap metastable configurations, our tests indicate that plausible variations in embedding site lead only to modest quantitative changes in vibronic coupling (e.g., 0.0061 versus 0.0072 for two DBT configurations), without affecting the main conclusions.

[41] Y.-T. Huang, S. R. Kavanagh, D. O. Scanlon, A. Walsh, and R. L. Z. Hoye, Perovskite-inspired materials for photovoltaics and beyond—From design to devices, Nanotechnology **32**, 132004 (2021).

[42] N. Glaser, A. Baiardi, A. Z. Lieberherr, and M. Reiher, Vibrational entanglement through the lens of quantum information measures, J. Phys. Chem. Lett. **15**, 6958 (2024).

[43] The mode-overlap approximates the matrix element $\langle i|V_{\text{vib-bare phonon}}|\bar{j}_{\text{bare phonon}}\rangle \approx \langle i|j\rangle$ in the rate obtained from Fermi's golden rule. From the point of view of the vibrational configuration, the electronic transition appears instantaneous and any change in polarization will quickly displace local vibrational and global phononic modes, broadening the typical deltalike transition probability.

[44] A. Alkauskas, B. B. Buckley, D. D. Awschalom, and C. G. Van de Walle, First-principles theory of the luminescence lineshape for the triplet transition in diamond NV centres, New J. Phys. **16**, 073026 (2014).

[45] G. Kreiza, T. Javorskis, E. Orentas, and K. Kazlauskas, Boosting reverse intersystem crossing of TADF emitter through molecular geometry adaptation to crystalline host, Adv. Opt. Mater. **12**, 2400024 (2024).

[46] W. Li, J. Zhang, Z. Gao, J. Qi, and D. Ding, Advancing biomedical applications via manipulating intersystem crossing, Coord. Chem. Rev. **471**, 214754 (2022).

[47] A. Nicolet, M. A. Kol'chenko, B. Kozankiewicz, and M. Orrit, Intermolecular intersystem crossing in single-molecule spectroscopy: Terrylene in anthracene crystal, J. Chem. Phys. **124**, 164711 (2006).

[48] M. A. Kol'chenko, B. Kozankiewicz, A. Nicolet, and M. Orrit, Intersystem crossing mechanisms and single molecule fluorescence: Terrylene in anthracene crystals, Opt. Spectrosc. **98**, 681 (2005).

[49] D. Beljonne, Z. Shuai, G. Pourtois, and J. L. Bredas, Spin-orbit coupling and intersystem crossing in conjugated polymers: A configuration interaction description, J. Phys. Chem. A **105**, 3899 (2001).

[50] G. Maggiora, M. Vogt, D. Stumpfe, and J. Bajorath, Molecular similarity in medicinal chemistry: Miniperspective, J. Med. Chem. **57**, 3186 (2014).

[51] N. C. for Biotechnology Information, PubChem compound summary for CID 135991, tribenzo(de,h,kl)naphtho(1,2,3,4-rst)pentaphene, PubChem (2025), https://pubchem.ncbi.nlm.nih.gov/compound/135991, retrieved 2 October 2025.

[52] Y. Nakakuki, T. Hirose, H. Sotome, H. Miyasaka, and K. Matsuda, Hexa-*peri*-hexabenzo[7]helicene: Homogeneously π-extended helicene as a primary substructure of helically twisted chiral graphenes, J. Am. Chem. Soc. **140**, 4317 (2018).

[53] A. Izuoka, K. Wakui, T. Fukuda, N. Sato, and T. Sugawara, Refined molecular structure of tetrabenzo[*de*,*hi*,*op*,*st*]pentacene, Acta Crystallogr. Sect. C **48**, 900 (1992).

[54] N. R. Verhart, P. Navarro, S. Faez, and M. Orrit, Intersystem crossing rates of single perylene molecules in ortho-dichlorobenzene, Phys. Chem. Chem. Phys. **18**, 17655 (2016).

[55] M. Vacha and T. Tani, Single-molecule spectroscopy of benzodiphenanthrobisanthene in a Shpolskii matrix, J. Phys. Chem. A **101**, 5027 (1997).

[56] Y.-C. Hsieh, T.-C. Wu, J.-Y. Li, Y.-T. Chen, M.-Y. Kuo, P.-T. Chou, and Y.-T. Wu, Dinaphthozethrene and diindenozethrene: Synthesis, structural analysis, and properties, Org. Lett. **18**, 1868 (2016).

[57] T. V. V. Ramakrishna and P. R. Sharp, Carbon-carbon bond formation and cleavage in the dimerization of a nickelacycle, Organometallics **23**, 3079 (2004).

[58] Y. Yoshida, S. Tango, K. Isomura, Y. Nakamura, H. Kishida, T. Koretsune, M. Sakata, Y. Nakano, H. Yamochi, and G. Saito, Charge-transfer complexes based on C_{2v}-symmetric benzo[*ghi*]perylene: Comparison of their dynamic and electronic properties with those of D_{6h}-symmetric coronene, Mater. Chem. Front. **2**, 1165 (2018).

[59] A. A. Ruch, S. Handa, F. Kong, V. N. Nesterov, D. R. Pahls, T. R. Cundari, and L. M. Slaughter, Competing amination and C–H arylation pathways in Pd/xantphos-catalyzed transformations of binaphthyl triflates: Switchable routes to chiral amines and helicene derivatives, Org. Biomol. Chem. **14**, 8123 (2016).

[60] T. Y. Latychevskaia, A. Renn, and U. P. Wild, Higher-order Stark effect on single-molecules, Chem. Phys. **282**, 109 (2002).

[61] O. C. Musgrave and J. M. S. Skakle, Diacenaphtheno[7, 8 − *b* : 7^1, 8^1-*d*]thiophene, Acta Crystallogr. E: Struct. Rep. Online **60**, o359 (2004).

[62] J. Carreras, M. Patil, W. Thiel, and M. Alcarazo, Exploiting the π-acceptor properties of carbene-stabilized phosphorus centered trications $[L_3P]^{3+}$: Applications in Pt(II) catalysis, J. Am. Chem. Soc. **134**, 16753 (2012).

[63] T. Hatakeyama, S. Hashimoto, S. Seki, and M. Nakamura, Synthesis of BN-fused polycyclic aromatics via tandem intramolecular electrophilic arene borylation, J. Am. Chem. Soc. **133**, 18614 (2011).

[64] Y.-T. Tao, S. Pola, S. Kumar, and M. M. Islam, Synthesis and characterization of contorted pentabenzo-fused coronenes as semiconducting materials, J. Org. Chem. **82**, 8067 (2017).

[65] P. Walla, F. Jelezko, P. Tamarat, B. Lounis, and M. Orrit, Perylene in biphenyl and anthracene crystals: An example of the influence of the host on single-molecule signals, Chem. Phys. **233**, 117 (1998).

[66] M. Rauhut, B. Roberts, D. Maulding, W. Bergmark, and R. Coleman, Infrared liquid-phase chemiluminescence from reactions of bis (2, 4, 6-trichlorophenyl) oxalate, hydrogen peroxide, and infrared fluorescent compounds, J. Org. Chem. **40**, 330 (1975).

[67] D. Hallett, A. P. Foster, D. Whittaker, M. S. Skolnick, and L. R. Wilson, Engineering chiral light–matter interactions in a waveguide-coupled nanocavity, ACS Photon. **9**, 706 (2022).

[68] C. Sayrin, C. Junge, R. Mitsch, B. Albrecht, D. O'Shea, P. Schneeweiss, J. Volz, and A. Rauschenbeutel, Nanophotonic

optical isolator controlled by the internal state of cold atoms, Phys. Rev. X **5**, 041036 (2015).

[69] I. Söllner, S. Mahmoodian, S. L. Hansen, L. Midolo, A. Javadi, G. Kiršanskė, T. Pregnolato, H. El-Ella, E. H. Lee, J. D. Song, *et al.*, Deterministic photon–emitter coupling in chiral photonic circuits, Nat. Nanotechnol. **10**, 775 (2015).

[70] C. Zhang, H. Hu, C. Ma, Y. Li, X. Wang, D. Li, A. Movsesyan, Z. Wang, A. Govorov, Q. Gan, *et al.*, Quantum plasmonics pushes chiral sensing limit to single molecules: A paradigm for chiral biodetections, Nat. Commun. **15**, 2 (2024).

[71] C. K. Williams and C. E. Rasmussen, *Gaussian Processes for Machine Learning* (MIT Press, Cambridge, MA, 2006), Vol. 2.

[72] K. Müller, K. Luoma, and C. Schäfer, A hierarchical approach to quantum many-body systems in structured environments, arXiv:2405.05093.

[73] F. Lindel, D. Lentrodt, S. Y. Buhmann, and C. Schäfer, Quantized embedding approaches for collective strong coupling—Connecting *ab initio* and macroscopic QED to simple models in polaritonics, J. Chem. Phys. **161**, 154111 (2024).

[74] P. Lodahl, S. Mahmoodian, S. Stobbe, A. Rauschenbeutel, P. Schneeweiss, J. Volz, H. Pichler, and P. Zoller, Chiral quantum optics, Nature (London) **541**, 473 (2017).

[75] J. Snoek, H. Larochelle, and R. P. Adams, Practical Bayesian optimization of machine learning algorithms, Adv. Neural Inf. Process. Syst. **2**, 2951 (2012).

[76] E. Judge, M. Azzouzi, A. M. Mroz, A. D. R. Chanona, and K. E. Jelfs, Applying multi-fidelity Bayesian optimization in chemistry: Open challenges and major considerations, arXiv:2409.07190.

[77] E. K. Öhman, D. Wang, R. M. Geilhufe, and C. Schäfer, Prediction of molecular single-photon emitters: A materials-modeling approach, Zenodo (2026), http://doi.org/10.5281/zenodo.18924429.

[78] J. Zirkelbach, M. Mirzaei, I. Deperasińska, B. Kozankiewicz, B. Gurlek, A. Shkarin, T. Utikal, S. Götzinger, and V. Sandoghdar, High-resolution vibronic spectroscopy of a single molecule embedded in a crystal, J. Chem. Phys. **156**, 104301 (2022).

[79] K. Sharma, A. Loew, H. Wang, F. A. Nilsson, M. Jain, M. A. L. Marques, and K. S. Thygesen, Accelerating point defect photo-emission calculations with machine learning interatomic potentials, npj Comput. Mater. **11**, 334 (2025).

[80] K. G. Schadler, C. Ciancico, S. Pazzagli, P. Lombardi, A. Bachtold, C. Toninelli, A. Reserbat-Plantey, and F. H. Koppens, Electrical control of lifetime-limited quantum emitters using 2D materials, Nano Lett. **19**, 3789 (2019).