Linear and nonlinear optics in composite systems: From diagrammatic modeling to applications

Thomas Noblet, Bertrand Busson

Linear and nonlinear optics in composite systems: From diagrammatic modeling to applications

Thomas Noblet[1] and Bertrand Busson[2]

[1)] *GRASP-Biophotonics, CESAM, University of Liege, Institute of Physics, Allée du 6 août 17, 4000 Liège, Belgium.*[a)]

[2)] *Université Paris-Saclay, CNRS, Institut de Chimie Physique, UMR 8000, 91405 Orsay, France.*

(Dated: 15 May 2024)

A bipartite system is defined as two microscopic entities able to exchange energy. When excited by light, the complete optical response functions at first (polarizabilities) and second orders (first hyperpolarizabilities) of such a system are determined using the diagrammatic theory of optics. The generality of the method is ensured by the free choice of light-matter and matter-matter interaction Hamiltonians and by the arbitrary number of quanta involved in the energy exchange. In the dipolar approximation, the optical response functions of the system (i.e., of the interacting entities) are linked to the responses of the interaction-free entities by transfer matrices. These universal matrices identically modify the optical response functions at all orders in the electromagnetic field, allowing the implementation of matter-matter interactions in higher-order processes like stimulated or spontaneous Raman scattering and four wave mixing. This formalism is then applied to various composite systems: dimers, multimers and lattices of nanoparticles and molecules, dense molecular layers, and substrate-induced image dipoles.

[a)]Electronic mail: t.noblet@uliege.be

I. INTRODUCTION

Sum-Frequency Generation (SFG) is often employed as a vibrational or electronic spectroscopy dedicated to extracting information on molecules at an interface, typically a molecular monolayer on a substrate. Such a system may look simple but it actually falls into the category of composite systems. Even if the molecular contributions to SFG constitute of course the signals of interest, such a system encompasses other contributions, some of them rather obvious (e.g., the substrate or a liquid environment) and some less easy to detect (e.g., neighboring molecules in the monolayer or an optical window[1]). History of infrared-visible Sum-Frequency Generation spectroscopy shows that, most of the times, the experimental response of complex systems may not be modeled using the response functions of single objects.[2] As the contributions of the other components of the system interfere to produce the measured SFG intensity, instead of simply adding up as is the case for linear optical processes, it is not straightforward to subtract the undesired signals to focus on the sought-after ones, and sometimes even impossible to unambiguously assign their origins. Representative examples include interference of vibrationally resonant SFG with a nonresonant signal from the substrate,[3–5] discrimination between surface and bulk SFG contributions,[6,7] influence of molecular density and packing on the vibrational response,[8,9] and modification of the molecular response due to interaction with nanoparticles.[10] Experimental strategies have been elaborated to minimize these issues (e.g., background suppression,[11] heterodyne spectroscopy,[12] phase rotation[13]) or, conversely, to benefit from them.[14–17]

In this context, we would like to emphasize the pivotal role played by Yuen-Ron Shen in nonlinear optics and surface science. He has initiated many breakthroughs in a wide range of scientific fields, and opened tracks ultimately followed by many groups in the world. Just in the first half of his career, Y.R. Shen had already investigated the origins of the separation between molecular resonant and substrate nonresonant SFG contributions,[3] the existence of interaction contributions in the nonlinear response,[18] the theory of higher-order nonlinear processes (like four wave mixing,[19,20] stimulated Raman and Brillouin scatterings,[21,22] Coherent Stokes Raman Scattering[23]), the importance of higher-order multipolar terms in the nonlinear response[24] and their influence on the surface vs. bulk response ratio,[25,26] the dipole-dipole coupling in a molecular monolayer and its influence on the optical response,[27] and the local field effects and their potential enhancement properties.[28] Given the diversity

of these optical processes and their spectroscopic applications, they have been supported by specific theoretical frames and mathematical formalisms, sometimes with few or only indirect connections between them. It is however possible to merge all of them into a single common approach establishing rules for the determination of the optical responses of composite systems at all orders in the electromagnetic fields and of the multipolar expansion.

In this perspective, we have established a diagrammatic theory of linear and nonlinear optics dedicated to composite systems, whose foundations and principles have been developed in a first paper.[29] As recalled below, response functions of single objects are represented by loops in Feynman diagrams. We have then introduced interactions between these loops, assimilating any composite system to a collection of sub-units connected by vector bosons. In a second paper, a first application was provided in details for the iconic, but rather simple, molecule-nanoparticle system.[30] In a third paper, we have also illustrated the versatility of the method by expanding it beyond the electric dipolar coupling, up to magnetic and quadrupolar terms of the multipolar expansion.[31] To complete the construction of this theoretical approach, here we aim at proving its exhaustiveness. In a first step, we describe the generic first- and second-order response functions, unrestricted in terms of interaction Hamiltonians and number of vector bosons. These functions encompass all the possible optical responses of bipartite systems. In a second step, we consider the example of dipole-dipole interactions to show that the response function of any sub-unit (e.g., a molecule or a nanoparticle) coupled with a given partner (of same or different nature) is related to its bare response function (i.e., when not coupled with the partner) through universal transfer matrices. In other words, these matrices transform the response functions of isolated objects into their dressed counterparts, taking into account the interactions within the bipartite system. These transfer matrices appear to be universal indeed as they identically modify all response functions at higher orders, accounting also for third-order nonlinear optics and beyond. Applications to various systems like molecular monolayers, dimers and lattices of particles, and functionalized substrates show that this structured approach allows to recover and extend, under a unifying formalism, most of the results elaborated in the literature in diverse fields of chemical physics.

II. BIPARTITE DIAGRAMS FOR LINEAR AND NONLINEAR RESPONSES

A. Complete enumeration of 2-loop diagrams

We define a bipartite system as composed of two entities (or subsystems) interacting with light and with each other. In the following, one is considered as the entity of interest, therefore referred to as the "main", while the other one is seen as its "partner". In the previous papers establishing the diagrammatic theory of linear and nonlinear optics,[29,30] the main was supposed to be a molecule, and the partner some inorganic entity (typically a nanoparticle). Here we address the general frame where the main and the partner refer to any kind of entity: molecule, nanostructure, or even a solid substrate. Various examples are provided in Part VI. Each entity is then described by fermion propagators (associated to electronic, vibrational or rotational states, according to the properties we aim to account for) whose interactions with light involve photons, hence considered as the inputs and outputs of the system. At first order in the electromagnetic field, the response function is a usual polarizability $\boldsymbol{\alpha}(\omega)$ where ω is the frequency of incoming and outgoing photons. The generic response function at second-order is the hyperpolarizability $\boldsymbol{\beta}(\omega_1, \omega_2)$ representing the Sum-Frequency Generation (SFG) process: two incoming light beams (ω_1, ω_2) generate an output at $\omega_3 = \omega_1 + \omega_2$. Under minimal modifications, it may also account for other second-order processes like Second Harmonic Generation (SHG, for $\omega_1 = \omega_2$), Difference-Frequency Generation (DFG, for negative ω_2), and optical rectification (for $\omega_3 = 0$).[32,33] Furthermore, in the specific context of composite systems, the interactions between the two subunits are taken into account: they are driven by one or more interaction Hamiltonians $\mathcal{H}_{\text{int}}$ defining the coupling constants which connect the fermion propagators of each entity (see the detailed discussion in Ref. 31).

We refer the reader to the original papers (Ref. 29 and 30) that establish the Feynman-Matsubara formalism used for the calculation of optical response functions and provide the twelve Feynman rules to follow in order to draw and calculate the associated loop diagrams. We recall that the diagrams are straightforwardly calculated using imaginary frequencies, while real frequencies and transition linewidths are recovered at the end of the calculation by analytic continuity. In order to access the complete response function of a composite system,

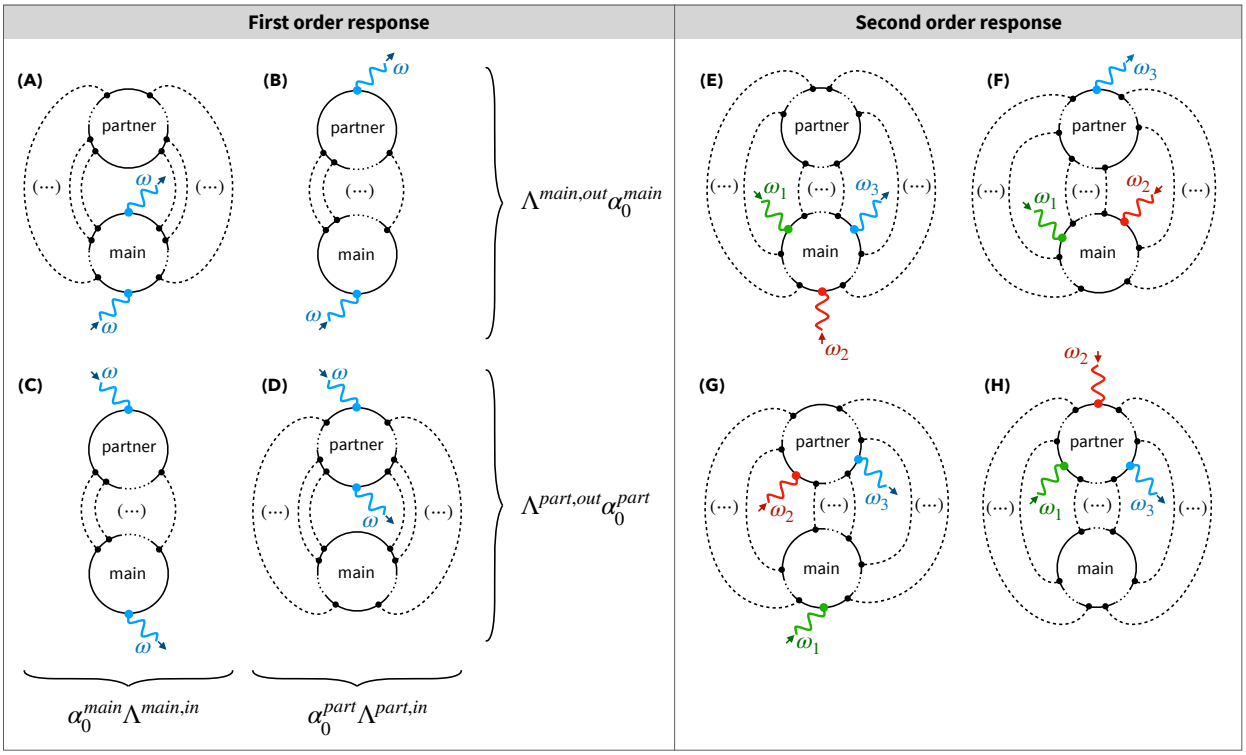

FIG. 1. Feynman diagrams generating the optical response functions of the bipartite system {main+partner}: (A−D) at first order (linear polarizability); (E−H) at second order (first hyperpolarizability). The wavy lines represent photons (light), the dashed lines depict interaction bosons, and the plain lines are fermion propagators. The symbol "(...)" indicates an arbitrary number of interaction processes. As detailed in the text, the complete set of diagrams (for given interaction Hamiltonians) is obtained by frequency filling and vertex permutation. In this instance, diagrams (E−H) illustrate only one of the possibilities to assign the photon frequencies to the wavy lines.

all possible and topologically distinct diagrams must be drawn and their contributions added up.

In a diagram, each subsystem is represented by a closed loop, in which a Matsubara frequency circulates. Such a loop consists in a sequence of fermion propagators connected by vertices representing as many interaction points (with light or with the other entity). When the two subsystems do not interact with each other, their bare response functions are called $\boldsymbol{\alpha}_0^{\mathrm{main/part}}(\imath\omega)$ and $\boldsymbol{\beta}_0^{\mathrm{main/part}}(\imath\omega_1,\imath\omega_2)$. Figure 1 gives the complete list of diagrams that it is possible to draw for a bipartite system, representing the first-order [polarizability $\boldsymbol{\alpha}(\imath\omega)$] and second-order [first hyperpolarizability $\boldsymbol{\beta}(\imath\omega_1,\imath\omega_2)$] response functions of the composite system. At this stage, the natures of the interactions between the main and the partner are arbitrary. As a consequence, the energy exchanged between them is also arbitrary and

is represented by virtual bosons connecting the loops and conveying imaginary frequencies. Without any approximation, Figure 1 describes the full and complete optical response of the system, which derives from the sum of all the possible ways to fill these generic diagrams with Matsubara frequencies, quantum states labels, coupling constants and photon frequencies. Especially, the number of matter-matter vertices (i.e., connecting the two loops) runs from 1 to infinity, and all the possibilities of vertex sequencing (i.e., the order in which they connect to each loop) must be considered. It is worth noting that, for the second-order response, considering all the vertex sequencing possibilities includes permutations of the three photon frequencies ω_1, ω_2 and ω_3. This is the first difference between the diagrams representing first- and second-order response functions. The second and most important difference lies in the description of the optical process itself. Following Feynman rule #5, the constitutive relationship $\imath\omega_3 = \imath\omega_1 + \imath\omega_2$ must appear once, and only once, in the whole second-order diagram (i.e., at only one vertex) in order to comply with the order of the nonlinear process. The corresponding "nonlinear vertex" may belong to the main or the partner loop, thus unambiguously ascribing the nonlinear process to one of the subsystems, whether the photon vertices are located on this loop or not. The diagrams of Figure 1 (E, F, G, H) may be assigned to the nonlinear response of the main (resp. the partner) inside the bipartite system, depending on which loop the nonlinear vertex is set on. They are therefore straightforwardly related to the nonlinear response $\boldsymbol{\beta}_0^{\text{main}}$ (resp. $\boldsymbol{\beta}_0^{\text{part}}$) of the same entity when considered alone. In the following, we assume that the nonlinear vertex is set on the main loop, unless otherwise noted. On the contrary, there is no such nonlinear vertex as for the first-order response and, as we will see in Part IV, it is not possible to unambiguously ascribe the linear response function of the system to one of its subsystems.

Explicit enumeration and, all the more, calculation of all these diagrams is an impossible task. However, it is possible to simplify and classify them through their splitting into elementary response functions. We explain below the hypotheses which lead to such a simplification and allow to completely skip the actual calculation process.

B. Essential hypotheses

Considering the infinite number of diagrams to add up and, for each diagram, the variable number of Matsubara frequencies over which a sum has to be performed, we introduce a few

hypotheses in order to simplify the diagrams:

- Light-matter interaction involves, in principle, a slightly different phase retardation for the electromagnetic wave at each subsystem because of the spatial separation between the main and the partner. In the diagrams, each loop should thus be assigned the phase factor deduced from the propagation of incoming and outgoing light beams at the location of the corresponding entity (e.g., $e^{\imath \boldsymbol{k}(\omega_i)\cdot \boldsymbol{R}_{\mathrm{main}}}$ where $\boldsymbol{k}$ is the wavevector of light and $\boldsymbol{R}_{\mathrm{main}}$ the position of the main). Of course, only the relative positions of the entities will contribute to a global phase shift attributed to the diagram in the end. For instance, these global phase shifts are all different for the diagrams in Fig. 1. However, they are usually neglected in interacting systems, even if they may become important for large systems,[34] when they are at the origin of specific effects like a chiral response[31,35] or of specific properties related to phase matching in some processes described by the second-order response function.[1,36,37] We do not take these phases into account in the following, but they can be explicitly reintroduced at the end of the diagrammatic calculation.[31]

- As defined in the frame of the multipolar expansion of the electric potential energy, the usual optical response functions (i.e., $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$) are often considered at the dipolar level of approximation. Even though other quadrupolar and magnetic response functions are diagrammatically computable and sometimes emerge as the leading ones,[31] we focus here on the dipolar description of light-matter interactions. Hence, the optical transition associated to any vertex making the junction between one photon and two fermion propagators is driven by the dipole moment $\boldsymbol{p}^{\mathrm{main/part}}$ of the entity.

- The fact that the dipolar interaction between the two entities is necessarily involved ensures the existence of at least one interaction Hamiltonian mediated by real bosons (more precisely, vector photons in this case). These vector photons (represented by dashed lines in the diagrams) can only convey the imaginary frequencies associated with the input and output photons (e.g., $\imath\omega_1$, $\imath\omega_2$ and $\imath\omega_3$ for second-order diagrams) and are not assigned a Matsubara frequency[30] (in Part III B, the role of bosonic loops driven by Matsubara frequencies will be discussed). Here again, the dipole-dipole coupling constitutes a first approximation of matter-matter electromagnetic interactions: other higher multipolar contributions may be added under a similar formalism, as has

been described in a previous paper.[31] In the following parts, all interaction vertices and coupling constants will then be described in terms of electric dipoles and electric fields. The interaction Hamiltonian corresponds in this case to the classical dipolar energy transfer between the main and the partner: $\langle \mathcal{H}_{\rm int} \rangle = -\boldsymbol{p}^{\rm main} \cdot \boldsymbol{E}(\omega, \boldsymbol{R})$, where $\boldsymbol{R}$ is the relative position between the two dipoles, and $\boldsymbol{E}(\omega, \boldsymbol{R}) = -\boldsymbol{W}^{\rm ee}(\imath\omega, \boldsymbol{R})\, \boldsymbol{p}^{\rm part}$ is the electric field created by the partner at the main. Defining $\hat{\boldsymbol{R}} = \boldsymbol{R}/|\boldsymbol{R}|$, the coupling matrix $\boldsymbol{W}^{\rm ee}$ is given by[38,39]

$$W^{\rm ee}_{lh}(0, \boldsymbol{R}) = \frac{\delta_{lh} - 3\hat{R}_l \hat{R}_h}{4\pi\varepsilon_0 |\boldsymbol{R}|^3}, \tag{1}$$

in the electrostatic approximation, and

$$W^{\rm ee}_{lh}(\imath\omega, \boldsymbol{R}) = \frac{e^{\imath\omega|\boldsymbol{R}|/c}}{4\pi\varepsilon_0 |\boldsymbol{R}|^3} \left[\delta_{lh} \left(1 - \imath \frac{\omega|\boldsymbol{R}|}{c} - \frac{\omega^2 |\boldsymbol{R}|^2}{c^2} \right) - \hat{R}_l \hat{R}_h \left(3 - 3\imath \frac{\omega|\boldsymbol{R}|}{c} - \frac{\omega^2 |\boldsymbol{R}|^2}{c^2} \right) \right] \tag{2}$$

when the electrodynamic terms are included. In the following, $\boldsymbol{W}^{\rm ee}$ will be noted $\boldsymbol{W}$ for simplicity.

- In this framework, the two entities are qualified as "uncoupled" when they are not coupled through the interaction Hamiltonian $\mathcal{H}_{\rm int}$ while being still considered in presence of each other. For instance, the bare linear polarizability $\boldsymbol{\alpha}_0^{\rm main}$ of the main corresponds to its linear response function computed in presence of its partner (i.e., taking into account the influence of the partner on the quantum states of the main) but without including the dipole-dipole coupling between the two ($\mathcal{H}_{\rm int} = \boldsymbol{0}$). In other words, the quantum states of each subsystem are supposed to take into account the existence of the other entity, but not through the interaction Hamiltonians involved in the diagrams. As an illustration, let us consider a molecule covalently bound to a nanoparticle. The states of the molecule engaged in the composite system are modified (compared with those of the same isolated molecule) because the molecule and the nanoparticle share part of their electron densities in the covalent bond. Within the bare response functions $\boldsymbol{\alpha}_0^{\rm mol}$ and $\boldsymbol{\beta}_0^{\rm mol}$ of the adsorbed molecule, the chemical hybridization is taken into account whereas its dipolar coupling with the nanoparticle is not. In the diagrammatric computation of the "coupled" entities, we then consider the effect of $\mathcal{H}_{\rm int}$ on their optical response functions but not on their quantum states: the

response functions of the "coupled"-molecule ($\boldsymbol{\alpha}^{\mathrm{mol}}$, $\boldsymbol{\beta}^{\mathrm{mol}}$) are obtained by summing over the same quantum states as those of the bare response functions ($\boldsymbol{\alpha}_0^{\mathrm{mol}}$, $\boldsymbol{\beta}_0^{\mathrm{mol}}$).

- Last, the multiple successive exchanges of bosons between the two subsystems imply that some fermion propagators convey identical frequencies on the same loop. On a mathematical point of view, the response function of the whole system therefore admits poles whose multiplicities correspond to the numbers of redundant frequencies conveyed on the two loops, respectively. As explained in the Supplementary Material of Ref. 30 and in the next Part, it is then possible to split each loop into elementary loops, provided that the temperature is assumed to vanish. We consider that it is the case in the following.

C. Simplification of the diagrams

The previous hypotheses have dramatic consequences as they allow us to apply the rules for diagram splitting and factorization (first elaborated in Ref. 30 and detailed in Appendix A). It is worth noting that only the skeletons of the diagrams have been drawn in Figure 1. Their frequency filling (consisting in assigning a frequency to every input/output photon, then to all fermion and boson propagators) follows from the Feynman rule #5 (i.e., energy conservation at each vertex and uniqueness of the nonlinear vertex).

In the case of first-order diagrams (Fig. 1, A−D), loop splitting leads to two topological subfamilies, depending on the locations of the input and output photons. Within diagrams (A) and (D), the two photons interact with the same loop, implying an even number V of vector bosons. As a result, the two loops exhibit V and $(V+2)$ vertices, respectively. Within diagrams (B) and (C), each photon interacts with a single loop, implying an odd number V of vector bosons. Both loops thus exhibit $(V+1)$ vertices. In all cases, the two loops encompass an even number of vertices and can be split into products of bare $\boldsymbol{\alpha}_0(\imath\omega)$ functions (Fig. 7A).

In the case of second-order diagrams (Fig. 1, E−H), the splitting is conditioned by the location of the nonlinear vertex. On the one hand, the loop which hosts the nonlinear interaction necessarily follows the structure of Figure 7B and splits into a product of (i) several $\boldsymbol{\alpha}_0$ functions (depending on the number of vector bosons) and (ii) a single $\boldsymbol{\beta}_0$ function. On the other hand, the second loop splits into a product of $\boldsymbol{\alpha}_0$ functions. Two examples of a

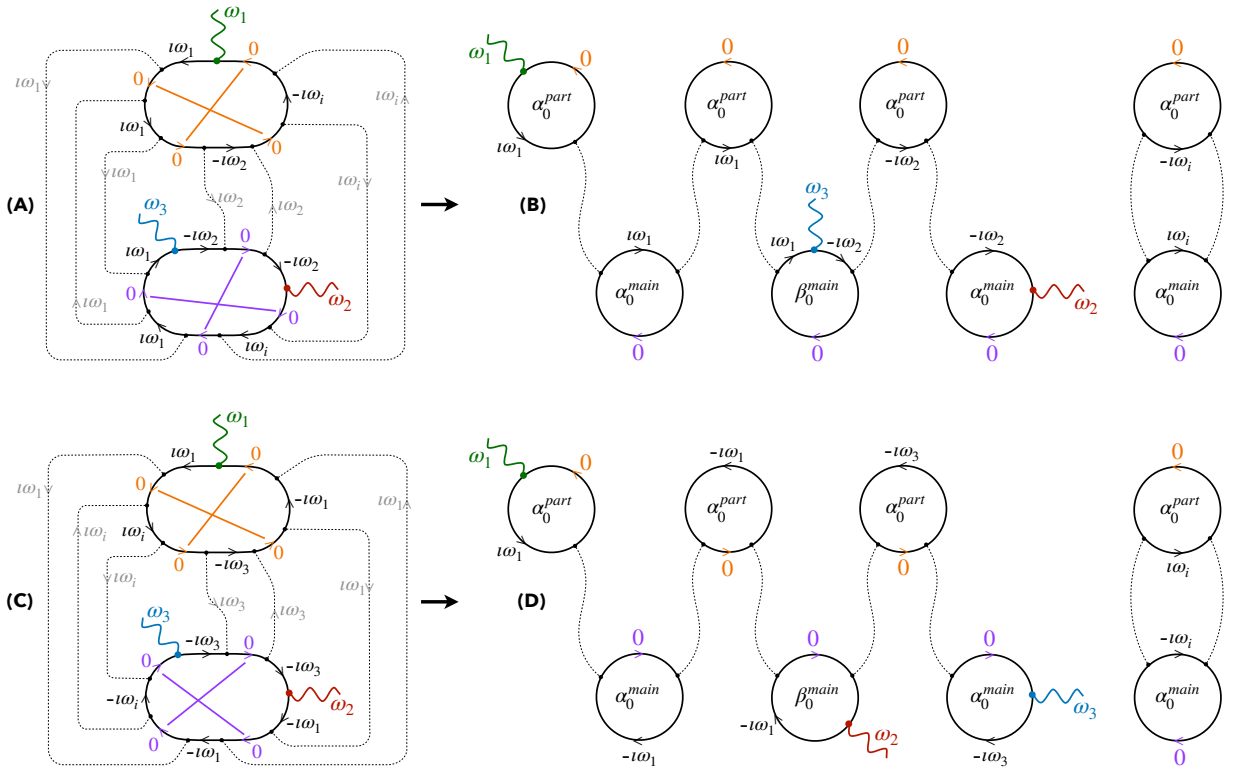

FIG. 2. Example of diagram splitting. The diagrams (A) and (C) share a common skeleton (number of vertices, light-matter interactions) but differ according to their frequency filling. The orange and purple lines link the propagators carrying identical frequencies (i.e., zero frequencies, here) and materialize the lines along which the loops can be split (see Appendix A). Each sector on the original loops generates a polarizability response function $\boldsymbol{\alpha}_0$ if it contains two vertices, or leads to the hyperpolarizability $\boldsymbol{\beta}_0$ if it contains three vertices. The simplified diagrams (B) and (D), obtained after splitting, account for different optical processes. Note that each diagram admits three versions according to the value of ω_i, for $i = 1$, 2 or 3.

second-order response function are provided in Figure 2 to illustrate the splitting procedure. We note that the natures of the various $\boldsymbol{\alpha}_0^{\text{main}}$, $\boldsymbol{\alpha}_0^{\text{part}}$ and $\boldsymbol{\beta}_0^{\text{main}}$ response functions involved in these diagrams depend on the natures of their vertices, defined by the corresponding light-matter and matter-matter Hamiltonians. For processes mixing electric dipole (e) and higher-order multipolar interactions [e.g., magnetic (m) and electric quadrupolar (Q)], different kinds of response functions (e.g., $\boldsymbol{\alpha}_0^{\text{em}}$ or $\boldsymbol{\alpha}_0^{\text{Qe}}$) must be considered.[31] Here again we restrict ourselves to dipolar approximation and assume that $\boldsymbol{\alpha}_0 = \boldsymbol{\alpha}_0^{\text{ee}}$ for all occurrences.

After the loop splitting, the original diagrams in Fig. 1 (A−D) have transformed into linear chains made of the bare polarizabilities of the main and the partner, hence connected by the exchange of bosons. Those in Fig. 1 (E−H) have transformed into three connected chains (each made of alternating $\boldsymbol{\alpha}_0^{\text{main}}$ and $\boldsymbol{\alpha}_0^{\text{part}}$ functions linked by vector bosons) meeting

at the main loop through $\boldsymbol{\beta}_0^{\text{main}}(\imath\omega_1, \imath\omega_2)$. As an important consequence, there is no need to write down or even calculate the diagrams: their expression can be directly read by following the energy flux on their drawing, translating into matrix products, from right to left (see Ref. 30 and 31, and the next Parts). The number of possible sequences is countable because boson exchanges may only occur between the main and the partner, not between an entity and itself.

III. SECOND-ORDER RESPONSE: SFG HYPERPOLARIZABILITY

After diagram splitting, the second-order response functions are paradoxically the easiest to understand because the "propagation" of the input ($\imath\omega_1$, $\imath\omega_2$) and output ($\imath\omega_3$) frequencies can be separately tracked on the diagrams. From the structure of the partner loop, we know that the three paths followed by the energy flow at each frequency are distinct and meet at the hyperpolarizability loop $\boldsymbol{\beta}_0^{\text{main}}$. Once split, each diagram then consists in a combination of elementary processes which can be gathered in three families, as pictured in Figure 3. We describe them using an incoming photon as an example (for the outgoing photon $\imath\omega_3$, the logical sequences are reversed). The first family, $\boldsymbol{P}_n^{\text{part}}$ (where the superscript indicates which entity interacts with the photon), accounts for partner-mediated interactions between light and the main entity: one photon interacts with the partner and its energy is transferred to the main, where it takes part in the SFG process. The complexity of such a mechanism, indexed by n, corresponds to the number of round-trip energy exchanges between the main and the partner. For the second family, $\boldsymbol{M}_n^{\text{main}}$, the photon directly interacts with the main, then a back and forth energy exchange with the partner happens before the photon takes part in the SFG process, involving a total of $(n+1)$ round-trip exchanges between the two subsystems. In this process, the photon does not interact with the partner. As for the third family, $\mathbf{C}_n$, the physics is quite different from the two previous ones because it involves closed circuits of energy exchange between the main and its partner [with $(n+1)$ back and forth steps], without any correlation with the SFG process itself. For this reason, we separate below the study of the in-line processes (belonging to the families $\boldsymbol{P}_n$ and $\boldsymbol{M}_n$) from the circular ones ($\boldsymbol{C}_n$).

In Fig. 2 (B), for example, frequency $\imath\omega_1$ follows a $\boldsymbol{P}_1^{\text{part}}$ process while frequency $\imath\omega_2$ undergoes an $\boldsymbol{M}_0^{\text{main}}$ process. In addition, one of the three frequencies is involved in a $\mathbf{C}_0$

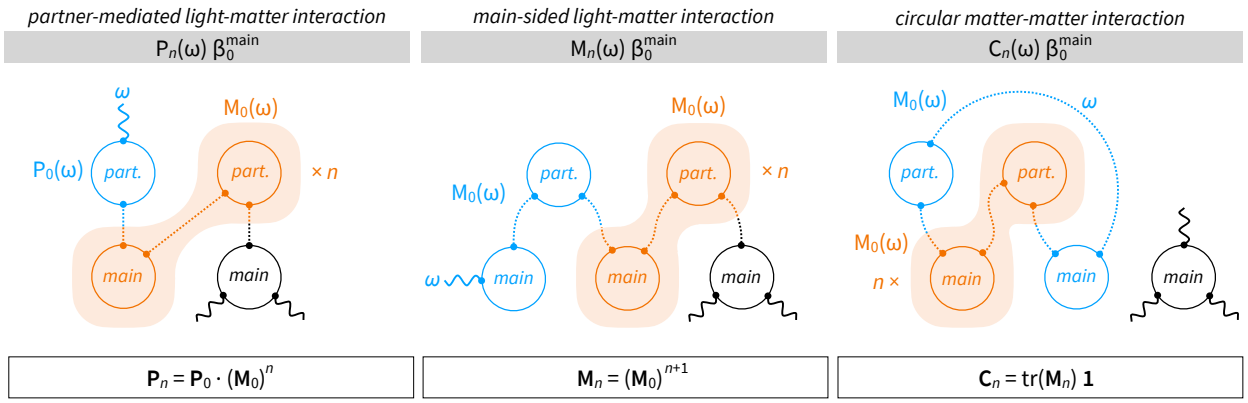

FIG. 3. The three families of subprocesses arising from the splitting of a generic bipartite diagram. The first member of each family ($\boldsymbol{P}_0^{\mathrm{part}}$, $\boldsymbol{M}_0^{\mathrm{main}}$, $\boldsymbol{C}_0$), i.e., corresponding to the lowest complexity $n = 0$, is drawn in blue. The other members ($\boldsymbol{P}_n^{\mathrm{part}}$, $\boldsymbol{M}_n^{\mathrm{main}}$, $\boldsymbol{C}_n$) involve in addition a periodic pattern (in orange) repeated n times, identical for the three families.

process. The diagram of Fig. 2 (D) consists for its part in a combination of $\boldsymbol{P}_1^{\mathrm{part}}$ processes for $\imath\omega_1$, $\boldsymbol{M}_0^{\mathrm{main}}$ processes for $\imath\omega_3$ and $\boldsymbol{C}_0$ processes for one of the three frequencies. We note that, if the nonlinear vertex is set on the partner loop in Fig. 1, the roles of the main and the partner are swapped, so the associated split diagrams encompass the partner hyperpolarizability $\boldsymbol{\beta}_0^{\mathrm{part}}$ modified by the symmetric processes $\boldsymbol{P}_n^{\mathrm{main}}$ and $\boldsymbol{M}_n^{\mathrm{part}}$, whereas $\boldsymbol{C}_n$ processes do not change.

A. In-line processes $\boldsymbol{P}_n$ and $\boldsymbol{M}_n$

As we see in Fig. 3, $\boldsymbol{P}_0^{\mathrm{part}}$ and $\boldsymbol{M}_0^{\mathrm{main}}$ are the building blocks of the three families of subprocesses. For incoming beams $(\imath\omega_1, \imath\omega_2)$, the vertex sequence in $\boldsymbol{P}_0^{\mathrm{part,in}}$ involves a first interaction with the partner (via $\boldsymbol{\alpha}_0^{\mathrm{part}}$) followed by an energy transfer to the main (via $\boldsymbol{W}$). Recalling that the corresponding matrices are introduced from right to left,[29,30] we get

$$\boldsymbol{P}_0^{\mathrm{part,in}}(\omega_{1,2}) = -\boldsymbol{W}(\imath\omega_{1,2})\boldsymbol{\alpha}_0^{\mathrm{part}}(\imath\omega_{1,2}). \tag{3}$$

Conversely, for the outgoing beam $\imath\omega_3$, the energy is first transferred from the main to the partner (via $\boldsymbol{W}$), followed by radiation of the partner at $\imath\omega_3$ (via $\boldsymbol{\alpha}_0^{\mathrm{part}}$). We have thus

$$\boldsymbol{P}_0^{\mathrm{part,out}}(\omega_3) = -\boldsymbol{\alpha}_0^{\mathrm{part}}(\imath\omega_3)\boldsymbol{W}(\imath\omega_3). \tag{4}$$

The corresponding $\boldsymbol{P}_0^{\text{part}}$-modified response functions can be written

$$\beta_{ijk}^{\text{main}}(\imath\omega_1, \imath\omega_2) = \sum_{j'} \left[\beta_0^{\text{main}}(\imath\omega_1, \imath\omega_2)\right]_{ij'k} \left[P_0^{\text{part,in}}(\omega_1)\right]_{j'j} \text{ for } \omega_1, \tag{5}$$

$$\beta_{ijk}^{\text{main}}(\imath\omega_1, \imath\omega_2) = \sum_{k'} \left[\beta_0^{\text{main}}(\imath\omega_1, \imath\omega_2)\right]_{ijk'} \left[P_0^{\text{part,in}}(\omega_2)\right]_{k'k} \text{ for } \omega_2, \tag{6}$$

$$\beta_{ijk}^{\text{main}}(\imath\omega_1, \imath\omega_2) = \sum_{i'} \left[P_0^{\text{part,out}}(\omega_3)\right]_{ii'} \left[\beta_0^{\text{main}}(\imath\omega_1, \imath\omega_2)\right]_{i'jk} \text{ for } \omega_3. \tag{7}$$

In other words, a tensor contraction between the $\boldsymbol{P}_0$ matrices and the corresponding index of the hyperpolarizability tensor is performed. In these equations, the calculated hyperpolarizability $\boldsymbol{\beta}$ is attributed to the main entity because it is the place where the nonlinear process (characterized by $\imath\omega_3 = \imath\omega_1 + \imath\omega_2$) precisely happens. In terms of classical physics, $\boldsymbol{P}_0^{\text{part}}(\omega_1)$ and $\boldsymbol{P}_0^{\text{part}}(\omega_2)$ reflect the fact that light interacts with the partner instead of interacting with the main, and creates a local dipole radiating an electric field which is used by the main to participate in the nonlinear process. When $\boldsymbol{P}_0^{\text{part}}(\omega_3)$ is involved, the output of the nonlinear process is a dipole located at the main, radiating this time an electric field to the partner which, in turn, converts it into the source dipole of SFG radiation.

In a similar way, the matrices describing the $\boldsymbol{M}_0^{\text{main}}$ processes are

$$\boldsymbol{M}_0^{\text{main,in}}(\omega_{1,2}) = \boldsymbol{P}_0^{\text{part,in}}(\omega_{1,2})\, \boldsymbol{P}_0^{\text{main,in}}(\omega_{1,2}) \tag{8}$$

$$\boldsymbol{M}_0^{\text{main,out}}(\omega_3) = \boldsymbol{P}_0^{\text{main,out}}(\omega_3)\, \boldsymbol{P}_0^{\text{part,out}}(\omega_3). \tag{9}$$

In the case where all the matrices $\boldsymbol{P}_0^{\text{main,in}}$ involved in the chains (resp. $\boldsymbol{P}_0^{\text{main,out}}$, $\boldsymbol{P}_0^{\text{part,in}}$, $\boldsymbol{P}_0^{\text{part,out}}$) are identical (i.e., for a single interaction Hamiltonian, chosen here as the dipole-dipole Hamiltonian), the generic processes are given, for all $n \geqslant 0$, by

$$\boldsymbol{P}_n^{\text{part,in}}(\omega_{1,2}) = \left[\boldsymbol{P}_0^{\text{part,in}}(\omega_{1,2})\, \boldsymbol{P}_0^{\text{main,in}}(\omega_{1,2})\right]^n \boldsymbol{P}_0^{\text{part,in}}(\omega_{1,2}) \tag{10}$$

$$\boldsymbol{P}_n^{\text{part,out}}(\omega_3) = \boldsymbol{P}_0^{\text{part,out}}(\omega_3) \left[\boldsymbol{P}_0^{\text{main,out}}(\omega_3)\, \boldsymbol{P}_0^{\text{part,out}}(\omega_3)\right]^n \tag{11}$$

$$\boldsymbol{M}_n^{\text{main,in}}(\omega_{1,2}) = \left[\boldsymbol{P}_0^{\text{part,in}}(\omega_{1,2})\, \boldsymbol{P}_0^{\text{main,in}}(\omega_{1,2})\right]^{n+1} \tag{12}$$

$$\boldsymbol{M}_n^{\text{main,out}}(\omega_3) = \left[\boldsymbol{P}_0^{\text{main,out}}(\omega_3)\, \boldsymbol{P}_0^{part,out}(\omega_3)\right]^{n+1}. \tag{13}$$

All these processes have their classical interpretation for which the energy flows between the subsystems through successive conversions from electric fields into dipoles (via polarizabili-

ties $\boldsymbol{\alpha}_0$) and from dipoles into radiated electric fields (via matrices $\boldsymbol{W}$).

The complete second-order response function $\boldsymbol{\beta}^{\text{main}}$ of the main (when coupled to the partner) is obtained by adding up the contributions of all these processes, for n running from zero to infinity. We get

$$\beta_{ijk}^{\text{main}} = \sum_{i'j'k'} \left[\Lambda^{\text{main,out}}(\omega_3)\right]_{ii'} \left[\beta_0^{\text{main}}(\imath\omega_1, \imath\omega_2)\right]_{i'j'k'} \left[\Lambda^{\text{main,in}}(\omega_1)\right]_{j'j} \left[\Lambda^{\text{main,in}}(\omega_2)\right]_{k'k} \tag{14}$$

with

$$\boldsymbol{\Lambda}^{\text{main,in}}(\omega_{1,2}) = \mathbf{1} + \sum_{n=0}^{\infty} \left[\boldsymbol{P}_n^{\text{part,in}}(\omega_{1,2}) + \boldsymbol{M}_n^{\text{main,in}}(\omega_{1,2})\right], \tag{15}$$

$$\boldsymbol{\Lambda}^{\text{main,out}}(\omega_3) = \mathbf{1} + \sum_{n=0}^{\infty} \left[\boldsymbol{P}_n^{\text{part,out}}(\omega_3) + \boldsymbol{M}_n^{\text{main,out}}(\omega_3)\right]. \tag{16}$$

At this stage, we have not taken the circular processes $\boldsymbol{C}_n$ into account (see Part III B), but we have included all the processes in line with one of the three light-matter interactions. Given Eqs. (10)−(13), the $\boldsymbol{\Lambda}$ matrices can be actually expressed as geometric series of matrices:

$$\boldsymbol{\Lambda}^{\text{main,in}}(\omega_{1,2}) = \sum_{n=0}^{\infty} \left[\boldsymbol{M}_0^{\text{main,in}}(\omega_{1,2})\right]^n \left[\mathbf{1} + \boldsymbol{P}_0^{\text{part,in}}(\omega_{1,2})\right], \tag{17}$$

$$\boldsymbol{\Lambda}^{\text{main,out}}(\omega_3) = \left[\mathbf{1} + \boldsymbol{P}_0^{\text{part,out}}(\omega_3)\right] \sum_{n=0}^{\infty} \left[\boldsymbol{M}_0^{\text{main,out}}(\omega_3)\right]^n, \tag{18}$$

whose common ratios are $\boldsymbol{M}_0^{\text{main,in}} = \boldsymbol{W}\,\boldsymbol{\alpha}_0^{\text{part}}\,\boldsymbol{W}\,\boldsymbol{\alpha}_0^{\text{main}}$ and $\boldsymbol{M}_0^{\text{main,out}} = \boldsymbol{\alpha}_0^{\text{main}}\,\boldsymbol{W}\,\boldsymbol{\alpha}_0^{\text{part}}\,\boldsymbol{W}$, respectively. It is known that such matrix series converge when the modulus of the highest eigenvalue of their common ratio is smaller than unity. As a rule of thumb, the order of magnitude of the polarizability is given by $4\pi\varepsilon_0 d^3$, where d is the characteristic size of the entity. Besides, the order of magnitude of $\boldsymbol{W}$ follows $1/4\pi\varepsilon_0|\boldsymbol{R}|^3$, where the interdistance $|\boldsymbol{R}|$ is evaluated from center to center and is consequently bigger than $d_{\text{main}} + d_{\text{part}}$. Henceforth, for non-resonant polarizabilities:

$$\sup|\boldsymbol{M}_0| \sim \left(\frac{d_{\text{main,part}}}{d_{\text{main}} + d_{\text{part}}}\right)^6 < 1, \tag{19}$$

and the series are convergent. Of course, one may argue that convergence is not ensured

anymore when one of the polarizabilities becomes resonant. For example, this is indeed the case for molecules at vibrational or electronic resonances, metallic nanoparticle at surface plasmon resonance, or quantum dot at excitonic resonance. However, the polarizabilities involved in the equations are calculated at imaginary frequencies ($\imath\omega_1$, $\imath\omega_2$ or $\imath\omega_3$) and do not exhibit any resonance on the imaginary axis. Consequently, we consider these series as convergent for all kinds of system. The $\mathbf{\Lambda}$ matrices therefore become

$$\mathbf{\Lambda}^{\mathrm{main,in}}(\omega_{1,2}) = \left[\mathbf{1} - \boldsymbol{P}_0^{\mathrm{part,in}}(\omega_{1,2})\boldsymbol{P}_0^{\mathrm{main,in}}(\omega_{1,2})\right]^{-1}\left[\mathbf{1} + \boldsymbol{P}_0^{\mathrm{part,in}}(\omega_{1,2})\right], \tag{20}$$

$$\mathbf{\Lambda}^{\mathrm{main,out}}(\omega_3) = \left[\mathbf{1} + \boldsymbol{P}_0^{\mathrm{part,out}}(\omega_3)\right]\left[\mathbf{1} - \boldsymbol{P}_0^{\mathrm{main,out}}(\omega_3)\boldsymbol{P}_0^{\mathrm{part,out}}(\omega_3)\right]^{-1}. \tag{21}$$

In the particular case of a homo-dimer, hence made of identical main and partner, we have $(\mathbf{1} - \boldsymbol{M}_0) = (\mathbf{1} - \boldsymbol{P}_0^2) = (\mathbf{1} + \boldsymbol{P}_0)(\mathbf{1} - \boldsymbol{P}_0)$ and the $\mathbf{\Lambda}$ matrices simply read

$$\mathbf{\Lambda}^{\mathrm{dimer,in}}(\omega_{1,2}) = \left[\mathbf{1} - \boldsymbol{P}_0^{\mathrm{in}}(\omega_{1,2})\right]^{-1}, \tag{22}$$

$$\mathbf{\Lambda}^{\mathrm{dimer,out}}(\omega_3) = \left[\mathbf{1} - \boldsymbol{P}_0^{\mathrm{out}}(\omega_3)\right]^{-1}. \tag{23}$$

B. Circular processes C_n

As shown in Figure 2, diagram splitting also generates circular processes of type C_n. For the simpler case of C_0, we have

$$\begin{aligned} C_0(\omega_i) &= \sum_{j,k,h,l} W_{jl}(\imath\omega_i)[\alpha_0^{\mathrm{part}}(\imath\omega_i)]_{lh} W_{hk}(\imath\omega_i)[\alpha_0^{\mathrm{main}}(\imath\omega_i)]_{kj} \\ &= \mathrm{tr}[\boldsymbol{P}_0^{\mathrm{part}}(\omega_i)\boldsymbol{P}_0^{\mathrm{main}}(\omega_i)] = \mathrm{tr}[\boldsymbol{P}_0^{\mathrm{main}}(\omega_i)\boldsymbol{P}_0^{\mathrm{part}}(\omega_i)] = \mathrm{tr}[\boldsymbol{M}_0(\omega_i)] \end{aligned} \tag{24}$$

for $i = 1, 2, 3$ and indifferently for $\boldsymbol{M}_0 = \boldsymbol{M}_0^{\mathrm{part}}$ or $\boldsymbol{M}_0^{\mathrm{main}}$, where the trace operator accounts for the circular property of the diagram. Process C_0 introduces a multiplier which scales any original diagram (i.e., representing $\boldsymbol{\beta}^{\mathrm{main}}$ or $\boldsymbol{\beta}^{\mathrm{part}}$) by a universal quantity. In the same way, the generic process C_n reads

$$C_n(\omega_i) = \mathrm{tr}\{[\boldsymbol{P}_0^{\mathrm{part}}(\omega_i)\boldsymbol{P}_0^{\mathrm{main}}(\omega_i)]^{n+1}\} = \mathrm{tr}\{[\boldsymbol{M}_0(\omega_i)]^{n+1}\}. \tag{25}$$

Through diagram splitting, more and more parallel C_n processes are generated (for each of the three optical frequencies) as and when the amount of vector bosons increases. For

example, we note that C_1 is different from $(C_0)^2$ because of the trace operator. Indeed, for $(C_0)^2$, the two $\boldsymbol{M}_0$ processes are associated to two distinct 2-loop closed diagrams and are independent from each other while, for C_1, the two $\boldsymbol{M}_0$ processes follow each other within the same 4-loop closed diagrams. In the former case, each $\boldsymbol{M}_0$ operator loops over a single polarization state [$j \to j$ as in Eq. (24) and $k \to k$] whereas, in the latter case, one $\boldsymbol{M}_0$ transforms polarization j into k while the other one turns k back into j. Taking into account all possible combinations of C_n processes (in addition to the in-line processes), the generic hyperpolarizability tensor of the main entity coupled with its partner is eventually given by

$$\beta_{ijk}^{\text{main}} = \sum_{i'j'k'} \left[\widetilde{\Lambda}^{\text{main,out}}(\omega_3)\right]_{ii'} \left[\beta_0^{\text{main}}(\imath\omega_1, \imath\omega_2)\right]_{i'j'k'} \left[\widetilde{\Lambda}^{\text{main,in}}(\omega_1)\right]_{j'j} \left[\widetilde{\Lambda}^{\text{main,in}}(\omega_2)\right]_{k'k}, \quad (26)$$

where

$$\widetilde{\boldsymbol{\Lambda}}^{\text{main}}(\omega_i) = \boldsymbol{\Lambda}^{\text{main}}(\omega_i) \prod_{n=0}^{\infty} \sum_{m=0}^{\infty} [C_n(\omega_i)]^m = \boldsymbol{\Lambda}^{\text{main}}(\omega_i) \prod_{n=0}^{\infty} \frac{1}{1 - C_n(\omega_i)} \quad (27)$$

for both "in" and "out" frequencies. In a concise notation, we can write $\widetilde{\boldsymbol{\Lambda}}^{\text{main}}(\omega_i) = \boldsymbol{\Lambda}^{\text{main}}(\omega_i) L(\omega_i)$. Given that $\sup |\boldsymbol{M}_0| < 1$, Eq. (19), the convergence of the infinite product is ensured by Eq. (25), which implies that $C_n \to 0$ when $n \to \infty$.

Even if they seem independent of the other linear or nonlinear processes occurring in the diagram, processes C_n exist only as companion processes of an in-line process because they originate in the splitting of a common mother loop (as in Fig. 2) and do not stand alone. This is coherent with the fact that they represent a scaling factor of the response function defined by all the in-line processes. In Fig. 4A, this is illustrated in the simplest case of a C_0 process co-existing with the bare hyperpolarizability $\boldsymbol{\beta}_0^{\text{main}}$. The latter translates the occurrence of SFG and implies the propagation of excited states within the system: in the classical picture, oscillating dipoles are locally generated at the optical frequencies ω_1, ω_2 and ω_3. But each dipole is not only involved in the SFG mechanism. The diagram of Fig. 4A teaches us that the as-generated dipoles may participate in back and forth energy exchanges between the main and the partner in parallel with SFG. Such a C_0 process only requires the prior excitation of oscillating dipoles within the system (carried out by the $\boldsymbol{\beta}_0^{\text{main}}$ loop), and is not topologically connected to the $\boldsymbol{\beta}_0^{\text{main}}$ loop through vector bosons (there is no radiated electric field between the two processes). This explains why it is possible to have an arbitrary

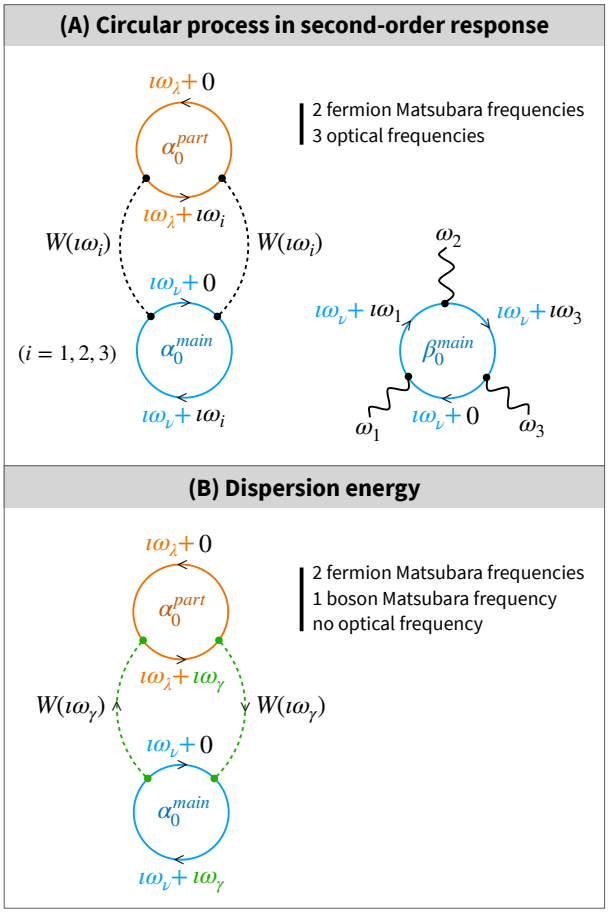

FIG. 4. Interpretation of circular processes. (A) Split diagram associated to the response function $\boldsymbol{\beta}^{\text{main}}(\imath\omega_1, \imath\omega_2) = C_0(\omega_i)\boldsymbol{\beta}_0^{\text{main}}(\imath\omega_1, \imath\omega_2)$, with $i = 1, 2$ or 3. The dashed lines represent vector bosons (real photons) conveying the optical frequency ω_i. (B) Feynman diagram associated to the dispersion energy. The dashed lines are virtual bosons described by an additional Matsubara frequency $\imath\omega_\gamma$.

number of such processes implemented in a diagram. Higher-complexity C_n contributions involve the same back and forth energy flux as C_0, repeated n times with additional flips of the polarization of light at each iteration.

Still, it is possible to draw a Feynman diagram representing something that looks like the C_0 process but without any in-line companion process (Fig. 4B). The calculation of this diagram follows from the same Feynman rules as those established for the optical response functions. We note that there are two interactions ($N_v = 2$) but no photons ($N_p = 0$), with the consequence that the value of the exchanged frequency is not dictated by an external input. As explained in Part II B, this means that the vector bosons remain virtual and, in line with the Feynman-Matsubara formalism, the frequencies conveyed by matrices $\boldsymbol{W}$ are

boson Matsubara frequencies (instead of optical frequencies). Finally, the limit of vanishing temperature is not required here because no diagram splitting is performed. The diagram is calculated accordingly in Appendix B, and leads to

$$U = -kT \sum_{\gamma} \mathrm{tr}\left[\boldsymbol{\alpha}^{\mathrm{main}}(\imath\omega_\gamma)\boldsymbol{W}(\imath\omega_\gamma)\boldsymbol{\alpha}^{\mathrm{part}}(\imath\omega_\gamma)\boldsymbol{W}(\imath\omega_\gamma)\right]. \tag{28}$$

The diagram represents this time an energy instead of a dimensionless scale factor. This comes from the additional $1/b$ prefactor associated to the boson Matsubara frequency $\imath\omega_\gamma$. A more usual formulation is obtained for vanishing temperature (Appendix B), allowing to recover a continuous description[40,41]:

$$U = -\frac{\hbar}{2\pi} \int_{-\infty}^{\infty} \mathrm{tr}\left[\boldsymbol{\alpha}^{\mathrm{main}}(\imath\omega)\boldsymbol{W}(\imath\omega)\boldsymbol{\alpha}^{\mathrm{part}}(\imath\omega)\boldsymbol{W}(\imath\omega)\right] \mathrm{d}\omega. \tag{29}$$

The energy U is negative, corresponding to an attractive force. For $\boldsymbol{W}$ describing the electric dipolar Hamiltonian in the electrostatic [Eq. (1)] or electrodynamic form [Eq. (2)], we recognize[42] the definition of the interaction energy giving birth to the London component of the van der Waals forces (i.e., the dispersion forces) as the attractive interaction between instantaneous electric dipoles. In fact, Eq. (28) coincides with the discrete MacLachlan formulation of the dispersion free energy at finite temperature,[42,43] obtained through a direct calculation of the interaction energy. Even so, our approach goes beyond MacLachlan's, allowing us to calculate the dispersion energies of various bipartite systems directly in imaginary frequencies and with the help of only one Feynman-Matsubara diagram (in the vein of the work of Dzyaloshinskii *et al.*[44]). In particular, when the electrostatic version of $\boldsymbol{W}$ is considered together with isotropic polarizabilities,[45–47] the original London formula and the $|\boldsymbol{R}|^{-6}$ law for the dispersion energy between two atoms are recovered.[42,48] For long distances, the electrodynamic version of $\boldsymbol{W}$ is required to account for retardation effects, and Eq. (29) leads to the Casimir and Polder results with an $|\boldsymbol{R}|^{-7}$ dependence.[45,47] Recent works aim at estimating the relative amplitudes of long range dispersion energy derived from Eq. (29) and dipolar energy transfer for macromolecules in solutions.[49] This dipolar energy transfer supposes that one of the molecules is established in a vibrational excited state, through thermal activation for example, but not after excitation by light. It may be seen as the thermally-induced counterpart of the $\boldsymbol{P}_0$ process. Finally, as explained in Part VI E, the

same equations as above may be used to describe the dispersion forces between the main entity and a dielectric substrate playing the role of the partner,[45,46] which demonstrates the versatility of our approach.

In this frame, we can now compare the C_0 contribution to optical response functions (Fig. 4A) with the dispersion energy (Fig. 4B). They encompass the same diagram (or subdiagram) but do not represent the same quantities. From Eq. (29), we see that the dimensionless circular energy transfer between the two subsystems at energy $\hbar\omega$ is integrated over all Matsubara energies with a unit energy density to generate the dispersion energy. In a C_0 process, the Matsubara energy in the bosonic loop is not arbitrary anymore as it is controlled by the main response function, resulting in an energy density $\delta(\omega - \omega_i)$, with $i = 1, 2$ or 3. We recover the essence of the circular energy transfer in C_n processes as dimensionless quantities. The interpretation of the C_0 process follows: it represents a forced dispersion energy transfer between the two partners, driven by the optical response functions, i.e., under excitation by light. In other words, C_0 quantifies the dispersion energy transfer between optically excited states.

IV. FIRST-ORDER POLARIZABILITY

A. Specificity of the first-order response

For the first-order response, there is only one frequency involved in the diagrams, hence ascribed to the input and output photons, and to all vector bosons. As a consequence of this indistinguishability and the absence of a nonlinear vertex, the optical process cannot be unambiguously attributed to one of the two loops. The full first-order response of the system is the sum of the four diagrams in Fig. 1 (A−D). While it is possible to associate diagram (A) to the first-order response of the main and diagram (D) to that of the partner, the assignments of diagrams (B) and (C) are a matter of perspective: it depends whether we aim to describe how each entity influences the way (i) the incoming photon interacts with the main (resp. the partner) or (ii) the outgoing photon is scattered by the main (resp. the partner). For instance, if we group diagrams (A) and (B) in Figure 1, we get a description of the partner-modified output response of the main (since the input photon is connected to the main loop in both diagrams). Instead, if we group diagrams (B) and (D), we get a

description of the main-modified input response of the partner (since the output photon is connected to the partner loop). Henceforth, grouping the diagrams according to the lines (A-B) and (C-D) in Fig. 1 ensures the uniqueness of the interaction of the incoming photon with one of the entities, whereas grouping the diagrams according to the columns (A-C) and (B-D) ensures the uniqueness of the entity responsible for the emission of the outgoing photon. As a result, there are two ways to define a response function for the main entity inside the bipartite system. We first introduce the "input" polarizability from diagrams (A)+(C) as the response function of the system for which the outgoing photon is emitted by the main, taking into account all the possible in-line modifications of the input side (hence the name). After diagram splitting (we drop the ω-dependence and focus on the in-line processes only), we obtain

$$\begin{aligned}
\boldsymbol{\alpha}^{\text{main,in}} &= \boldsymbol{\alpha}^{(A)} + \boldsymbol{\alpha}^{(C)} \\
&= \boldsymbol{\alpha}_0^{\text{main}} \sum_{n=0}^{\infty} \left[\boldsymbol{M}_0^{\text{main,in}}\right]^n + \boldsymbol{\alpha}_0^{\text{main}} \sum_{n=0}^{\infty} \left[\boldsymbol{M}_0^{\text{main,in}}\right]^n \boldsymbol{P}_0^{\text{part,in}} \\
&= \boldsymbol{\alpha}_0^{\text{main}} \sum_{n=0}^{\infty} \left[\boldsymbol{P}_0^{\text{part,in}} \boldsymbol{P}_0^{\text{main,in}}\right]^n \left[\mathbf{1} + \boldsymbol{P}_0^{\text{part,in}}\right] \qquad (30) \\
&= \boldsymbol{\alpha}_0^{\text{main}} \left[\mathbf{1} - \boldsymbol{P}_0^{\text{part,in}} \boldsymbol{P}_0^{\text{main,in}}\right]^{-1} \left[\mathbf{1} + \boldsymbol{P}_0^{\text{part,in}}\right]. \qquad (31)
\end{aligned}$$

In the same way, from diagrams (A)+(B), we define the "output" polarizability of the main entity in the system:

$$\begin{aligned}
\boldsymbol{\alpha}^{\text{main,out}} &= \left[\mathbf{1} + \boldsymbol{P}_0^{\text{part,out}}\right] \sum_{n=0}^{\infty} \left[\boldsymbol{P}_0^{\text{main,out}} \boldsymbol{P}_0^{\text{part,out}}\right]^n \boldsymbol{\alpha}_0^{\text{main}} \qquad (32) \\
&= \left[\mathbf{1} + \boldsymbol{P}_0^{\text{part,out}}\right] \left[\mathbf{1} - \boldsymbol{P}_0^{\text{main,out}} \boldsymbol{P}_0^{\text{part,out}}\right]^{-1} \boldsymbol{\alpha}_0^{\text{main}}. \qquad (33)
\end{aligned}$$

In other words:

$$\boldsymbol{\alpha}^{\text{main,out}} = \boldsymbol{\Lambda}^{\text{main,out}} \boldsymbol{\alpha}_0^{\text{main}} \text{ and } \boldsymbol{\alpha}^{\text{main,in}} = \boldsymbol{\alpha}_0^{\text{main}} \boldsymbol{\Lambda}^{\text{main,in}} \qquad (34)$$

and, for the complete polarizability of the bipartite system:

$$\boldsymbol{\alpha}^{\text{syst}} = \boldsymbol{\alpha}^{\text{main,in}} + \boldsymbol{\alpha}^{\text{part,in}} = \boldsymbol{\alpha}^{\text{main,out}} + \boldsymbol{\alpha}^{\text{part,out}}. \qquad (35)$$

When the circular processes C_n are taken into account, Eq. (34) and (35) are still valid provided that $\boldsymbol{\Lambda}$ is replaced by $\widetilde{\boldsymbol{\Lambda}}$ as in Eq. (27). In the dimer case, we note that all diagrams on Fig. 1(A−D) appear twice ($\boldsymbol{\alpha}^{\text{main,in}} = \boldsymbol{\alpha}^{\text{part,in}}$ and $\boldsymbol{\alpha}^{\text{main,out}} = \boldsymbol{\alpha}^{\text{part,out}}$), and only one occurrence must be taken into account. It is easily checked that it becomes possible to define this time a first order response function of the main, which coincides in fact with the response function of the dimeric system because

$$\boldsymbol{\alpha}^{\text{dimer}} = \boldsymbol{\alpha}^{\text{main,in}} = \boldsymbol{\alpha}_0^{\text{main}} \boldsymbol{\Lambda}^{\text{dimer,in}} = \boldsymbol{\alpha}^{\text{main,out}} = \boldsymbol{\Lambda}^{\text{dimer,out}} \boldsymbol{\alpha}_0^{\text{main}}. \tag{36}$$

B. Self-consistent equivalence

In the literature, the equations above are usually determined thanks to a self-consistent approach in the frame of the dipolar approximation of classical electromagnetism. For a system excited by an electromagnetic wave described by the electric field $\boldsymbol{E}_0$, the response for each subsystem consists in a dipole proportional to the local field $\boldsymbol{E}_{\text{loc}}$: $\boldsymbol{p}^{\text{main}} = \boldsymbol{\alpha}_0^{\text{main}} \boldsymbol{E}_{\text{loc}}^{\text{main}}$. The local field is the sum of the external field $\boldsymbol{E}_0$ and the electric field $-\boldsymbol{W}\boldsymbol{p}^{\text{part}}$ created by the dipole of the other entity:

$$\boldsymbol{p}^{\text{main}} = \boldsymbol{\alpha}_0^{\text{main}} [\boldsymbol{E}_0 - \boldsymbol{W}\boldsymbol{p}^{\text{part}}] \tag{37}$$

$$\boldsymbol{p}^{\text{part}} = \boldsymbol{\alpha}_0^{\text{part}} [\boldsymbol{E}_0 - \boldsymbol{W}\boldsymbol{p}^{\text{main}}]. \tag{38}$$

We deduce that:

$$\boldsymbol{p}^{\text{main}} = \boldsymbol{\alpha}_0^{\text{main}} [\mathbf{1} - \boldsymbol{W}\boldsymbol{\alpha}_0^{\text{part}}] \boldsymbol{E}_0 + \boldsymbol{\alpha}_0^{\text{main}} \boldsymbol{W} \boldsymbol{\alpha}_0^{\text{part}} \boldsymbol{W} \boldsymbol{p}^{\text{main}} \tag{39}$$

and

$$\boldsymbol{p}^{\text{main}} = [\mathbf{1} - \boldsymbol{P}_0^{\text{main,out}} \boldsymbol{P}_0^{\text{part,out}}]^{-1} \boldsymbol{\alpha}_0^{\text{main}} [\mathbf{1} + \boldsymbol{P}_0^{\text{part,in}}] \boldsymbol{E}_0. \tag{40}$$

It is easy to check that

$$\boldsymbol{P}_0^{\text{main,out}} \boldsymbol{P}_0^{\text{part,out}} \boldsymbol{\alpha}_0^{\text{main}} = \boldsymbol{\alpha}_0^{\text{main}} \boldsymbol{P}_0^{\text{part,in}} \boldsymbol{P}_0^{\text{main,in}}, \tag{41}$$

leading to the equality

$$[\mathbf{1} - \boldsymbol{P}_0^{\text{main,out}} \boldsymbol{P}_0^{\text{part,out}}]^{-1} \boldsymbol{\alpha}_0^{\text{main}} = \boldsymbol{\alpha}_0^{\text{main}} [\mathbf{1} - \boldsymbol{P}_0^{\text{part,in}} \boldsymbol{P}_0^{\text{main,in}}]^{-1}. \tag{42}$$

Given Eq. (40), it is eventually possible to write $\boldsymbol{p}^{\text{main}} = \boldsymbol{\alpha}^{\text{main,in}} \boldsymbol{E}_0$ with $\boldsymbol{\alpha}^{\text{main,in}} = \boldsymbol{\alpha}_0^{\text{main}} \boldsymbol{\Lambda}^{\text{main,in}}$ and to recover Eq. (31). In this description, the inclusion of the main entity into the bipartite system leads to a modification of its polarizability into an equivalent response function $\boldsymbol{\alpha}^{\text{main,in}}$ taking the influence of the partner into account. We note that Eq. (37) and (38) assume that only the input of the response function of each entity is modified by the presence of the other subsystem, hence the equivalence with $\boldsymbol{\alpha}^{\text{main,in}}$.

Conversely, let us call $\boldsymbol{\alpha}^{\text{main,out}}$ and $\boldsymbol{\alpha}^{\text{part,out}}$ the classical response functions of the system for which light interacts first with the main and the partner, respectively. The response driven by $\boldsymbol{\alpha}^{\text{main,out}}$ is the sum of (i) the bare response of the main and (ii) the part of the system response initiated at the main and emitted at the partner:

$$\boldsymbol{\alpha}^{\text{main,out}} \boldsymbol{E}_0 = \boldsymbol{\alpha}_0^{\text{main}} \boldsymbol{E}_0 - \boldsymbol{\alpha}^{\text{part,out}} \boldsymbol{W} \boldsymbol{\alpha}_0^{\text{main}} \boldsymbol{E}_0. \tag{43}$$

It leads to the coupled equations

$$\boldsymbol{\alpha}^{\text{main,out}} = \boldsymbol{\alpha}_0^{\text{main}} - \boldsymbol{\alpha}^{\text{part,out}} \boldsymbol{W} \boldsymbol{\alpha}_0^{\text{main}}, \tag{44}$$

$$\boldsymbol{\alpha}^{\text{part,out}} = \boldsymbol{\alpha}_0^{\text{part}} - \boldsymbol{\alpha}^{\text{main,out}} \boldsymbol{W} \boldsymbol{\alpha}_0^{\text{part}}, \tag{45}$$

which resolve into

$$\boldsymbol{\alpha}^{\text{main,out}} = [\mathbf{1} + \boldsymbol{P}_0^{\text{part,out}}] \boldsymbol{\alpha}_0^{\text{main}} [\mathbf{1} - \boldsymbol{P}_0^{\text{part,in}} \boldsymbol{P}_0^{\text{main,in}}]^{-1}. \tag{46}$$

With the help of Eq. (42), we recover Eq. (33).

There is indeed a full equivalence between the self-consistent classical approach and the Feynman diagrammatic method (under convergence condition and restricted to the in-line processes). Both approaches show the impossibility to unambiguously define the polarizability of the main entity inside the bipartite system, except in the dimer case or when matrices $\boldsymbol{\alpha}_0$ and $\boldsymbol{W}$ commute. On the one hand, the diagrammatic method is cumulative as all processes are enumerated then added up in order to produce the complete response, in

the form of an infinite series which converges to an integrated form. On the other hand, the self-consistent approach describes the interaction of the main with a partner (and vice versa) for which the cumulative process has already been implicitly integrated. Their equivalence is more easily seen in the dimer case: for identical main and partner, we have

$$\boldsymbol{p}^{\text{main}} = \boldsymbol{\alpha}^{\text{main}}\boldsymbol{E}_0 = \boldsymbol{\alpha}_0^{\text{main}}[\boldsymbol{E}_0 - \boldsymbol{W}(\boldsymbol{\alpha}^{\text{main}}\boldsymbol{E}_0)], \tag{47}$$

which directly gives the new polarizability of the main in the system,

$$\boldsymbol{\alpha}^{\text{main}} = [\mathbf{1} + \boldsymbol{\alpha}_0^{\text{main}}\boldsymbol{W}]^{-1}\boldsymbol{\alpha}_0^{\text{main}}, \tag{48}$$

but also its cumulative form

$$\boldsymbol{\alpha}^{\text{main}} = [\mathbf{1} - \boldsymbol{\alpha}_0^{\text{main}}\boldsymbol{W} + \boldsymbol{\alpha}_0^{\text{main}}\boldsymbol{W}\boldsymbol{\alpha}_0^{\text{main}}\boldsymbol{W} - \boldsymbol{\alpha}_0^{\text{main}}\boldsymbol{W}\boldsymbol{\alpha}_0^{\text{main}}\boldsymbol{W}\boldsymbol{\alpha}_0^{\text{main}}\boldsymbol{W} + \cdots]\,\boldsymbol{\alpha}_0^{\text{main}} \tag{49}$$

by successive iterations. It is worth noting that the conjunction of Eqs. (48) and (49) implicitly ensures the convergence condition $\sup|\boldsymbol{\alpha}_0^{\text{main}}\boldsymbol{W}| < 1$ through the series expansion $(\mathbf{1}+\boldsymbol{X})^{-1} = \sum(-1)^n\boldsymbol{X}^n$.

Interestingly, Eq. (47) is equivalent to a Dyson equation[50] for the first-order response function: the effective (or dressed) response function, calculated by enumeration and sum of the Feynman diagrams, is also obtained by the interaction of the bare main object with the dressed version of itself. In some cases, this raises the problem of the convergence of the infinite series leading to the effective response function, leading to renormalization procedures. In the simple version presented here, the series are convergent and both approaches match. When the partner differs from the main, we see from Eq. (17) and (18) that the concept of dressed objects also applies, as the dressed main polarizability is obtained by the interaction of the bare main polarizability with the dressed version of the partner polarizability (and vice versa):

$$\begin{aligned}\boldsymbol{\alpha}^{\text{main,out}} &= \boldsymbol{\Lambda}^{\text{main,out}}\boldsymbol{\alpha}_0^{\text{main}} = \left[\mathbf{1} - \boldsymbol{\alpha}^{\text{part,out}}\boldsymbol{W}\right]\boldsymbol{\alpha}_0^{\text{main}} \\ \boldsymbol{\alpha}^{\text{main,in}} &= \boldsymbol{\alpha}_0^{\text{main}}\boldsymbol{\Lambda}^{\text{main,in}} = \boldsymbol{\alpha}_0^{\text{main}}\left[\mathbf{1} - \boldsymbol{W}\boldsymbol{\alpha}^{\text{part,in}}\right].\end{aligned} \tag{50}$$

The self-consistent procedure takes a simple form in the electric dipole regime whereas the

Feynman diagram description applies in principle to all types of light-matter and matter-matter interactions [(m) and (Q) types, for example]. Furthermore, the circular diagrams C_n are not included in the self-consistent approach, even if they could be added in a second step.

V. PROPERTIES OF THE TRANSFER MATRICES Λ

A. Universality of Λ matrices

The specific case of second-order optical response is representative of all nonlinear optics: the transfer matrix $\widetilde{\boldsymbol{\Lambda}}$ is involved in the transformation of any bare response function into its dressed counterpart, whatever the order of the nonlinear process. Considering $\boldsymbol{f}_0^{(N)}$ as an N^{th}-order bare response function: $\boldsymbol{f}_0^{(N)}$ is a rank-$(N+1)$ tensor, with $\boldsymbol{f}_0^{(1)} = \boldsymbol{\alpha}_0$, $\boldsymbol{f}_0^{(2)} = \boldsymbol{\beta}_0$, and so on. Tensor $\boldsymbol{f}_0^{(N)}$ depends on N input frequencies $\{\omega_k\}_{k=1}^N$ and each photon ω_k can interact with one of the entities (i) directly [$\boldsymbol{\Lambda}(\omega_k) = \mathbf{1} + \cdots$], (ii) through a $\boldsymbol{P}_n$-type process [$\boldsymbol{\Lambda}(\omega_k) = \mathbf{1} + \sum \boldsymbol{P}_n(\omega_k) + \cdots$], or (iii) through an $\boldsymbol{M}_n$-type process [$\boldsymbol{\Lambda}(\omega_k) = \mathbf{1} + \sum \boldsymbol{P}_n(\omega_k) + \sum \boldsymbol{M}_n(\omega_k)$], with (iv) a possible C_n-type companion process [$\widetilde{\boldsymbol{\Lambda}}(\omega_k) = L(\omega_k)\boldsymbol{\Lambda}(\omega_k)$]. It is worth noting that the $C_n(\omega_k)$ processes scale indeed all the optical response functions as multipliers and their total contribution translates into a global renormalization factor $L(\omega_k) = \prod_{n=1}^{\infty} [1 - C_n(\omega_k)]^{-1}$ [see Eq. (27)] for each optical frequency.

As the $\widetilde{\boldsymbol{\Lambda}}(\omega_k)$ matrices gather all the processes which possibly affect the optical response of the main when coupled with a partner, these matrices are universal in the frame of the hypotheses listed in Part II B. For $N \geqslant 2$, the dressed response function $\boldsymbol{f}^{(N)}$ relates to its bare counterpart $\boldsymbol{f}_0^{(N)}$ through:

$$\begin{aligned} f^{(N)}_{i,j_1\cdots j_N} &= \sum_{i' j'_1 \cdots j'_N} \widetilde{\Lambda}^{\text{out}}_{ii'}(\omega_{N+1})\; f^{(N)}_{0;\, i', j'_1 \cdots j'_N} \prod_{k=1}^{N} \widetilde{\Lambda}^{\text{in}}_{j'_k j_k}(\omega_k) \\ &= \left(\sum_{i' j'_1 \cdots j'_N} \Lambda^{\text{out}}_{ii'}(\omega_{N+1})\; f^{(N)}_{0;\, i', j'_1 \cdots j'_N} \prod_{k=1}^{N} \Lambda^{\text{in}}_{j'_k j_k}(\omega_k) \right) \prod_{k=1}^{N+1} L(\omega_k), \end{aligned} \tag{51}$$

where $\omega_{N+1} = \omega_1 + \cdots + \omega_N$ is the output frequency. Some applications of this formula to third-order optical processes are provided in Part VI G. This generalizes Eq. (26), which indeed involves a contraction of the tensor quantifying the bare response function $\boldsymbol{\beta}_0$ with

$\widetilde{\mathbf{\Lambda}}$ matrices evaluated at the photon frequencies, specifying the "in" and "out" formalism in line with the input and output property of each photon. Moreover, as mentioned above, the contribution of the C_n processes does not take part in the tensor structure: they translate into a scaling factor $\prod_{k=1}^{N+1} L(\omega_k)$ which has no equivalent in the literature. For that reason, this factor is dropped in the following applications, but its physical meaning and relevance will have to be examined in future works.

B. Dimers, multimers and infinite lattices

When the composite system is made of identical entities, the transfer matrices take specific forms. According to Eq. (14), (22), (23) and (36), the $\mathbf{\Lambda}$ matrices and the response functions are quite simple for a dimer. Generalization to arbitrary multimers is conceivable by involving as many loops as the number of monomers in the diagrams. The difficulty lies in the fact that the coefficients of the interaction matrix $\boldsymbol{W}(\boldsymbol{R})$ depend on the distance and mutual orientations of the entities in the system. Describing a multimer as a finite collection of N identical entities located at sites $\{s_i\}_{i=1}^{N}$, all the loops in the diagram may now exchange energy with the other loops in any sequence: sites s_i and s_j interact through the coupling matrix $\boldsymbol{W}^{s_i,s_j}$. Defining $\boldsymbol{R}_i$ and $\boldsymbol{R}_j$ as the positions of sites s_i and s_j, respectively, the coefficients of $\boldsymbol{W}^{s_i,s_j}$ are obtained by replacing $\boldsymbol{R}$ by $\boldsymbol{R}_{i,j} = \boldsymbol{R}_i - \boldsymbol{R}_j$ in Eq. (2). The diagram splitting is performed in the same way as in the dimer case because all bosons carry the same energy whatever the associated value of $\boldsymbol{W}^{s_i,s_j}$. However, even for the first-order polarizability, counting and calculating the diagrams becomes an exhausting task this time because, contrary to the dimer case, the symmetry between the entities is broken: they do not play identical roles anymore in the system and cannot be interchanged. This also makes the self-consistent approach hardly tractable in the general case,[51] even if averaging over the positions and symmetry simplifications may help.[52] This is why the nearest neighbor approximation is commonly used for multimers, as discussed in Part V C.

Paradoxically, the situation becomes simpler when the identical entities build up an infinite 1D, 2D or 3D lattice instead of a finite multimer. The diagram encompasses in principle an infinite number of identical loops, interacting by energy exchange in all possible manners, i.e., involving any of the possible position-dependent coupling matrices $\boldsymbol{W}^{s_i,s_j}$. Given that such interactions only depend on the relative positions, we may fix the origin at

one site s_i and set $\boldsymbol{R}_i = \boldsymbol{0}$ without loss of generality. It becomes possible this way to draw an equivalent diagram with only two loops (Fig. 6A): the main site is set at the origin and the partner represents any of the infinite number of entities in the system (except the main itself). The main interacts indifferently with any other entity in the lattice (all represented by the partner), so interactions involve all possible matrices $\boldsymbol{W}^{0,s_j} \equiv \boldsymbol{W}^{s_j}$ covering all sites $\{s_j\}_{j\neq i}$, in all possible sequences and an arbitrary number of times. After loop splitting, the response function of the system thus follows the tracks of the dimer: choosing for example the "out" formulation in Eq. (36), we replace $\boldsymbol{P}_0^{\text{out}}$ by $\boldsymbol{P}_0^{\text{out},s_j} = -\boldsymbol{\alpha}_0\boldsymbol{W}^{s_j}$ for site s_j in Eq. (23), with $\boldsymbol{W}^{s_j} = \boldsymbol{W}(\boldsymbol{R}_j)$, to get

$$\boldsymbol{\Lambda}^{\text{lattice,out}} = \sum_{n=0}^{\infty} \sum_{\substack{s_1,s_2,\cdots,\\ s_n\in\{s\}}} \boldsymbol{P}_0^{\text{out},s_1} \cdots \boldsymbol{P}_0^{\text{out},s_n}, \tag{52}$$

where $\{s\}$ stands for the set of all sites different from the main. Considering that the matrices $\boldsymbol{P}_0^{\text{out},s_j}$ do not commute, the generic term in the second sum corresponds to the development of

$$\Big(\sum_{s_j\in\{s\}} \boldsymbol{P}_0^{\text{out},s_j} \Big)^n. \tag{53}$$

Under the convergence hypothesis (i.e., $\sup|\sum \boldsymbol{P}_0^{\text{out},s_j}| < 1$), the matrix $\boldsymbol{\Lambda}^{\text{lattice,out}}$ becomes

$$\boldsymbol{\Lambda}^{\text{lattice,out}} = \sum_{n=0}^{\infty} \Big(\sum_{s_j\in\{s\}} \boldsymbol{P}_0^{\text{out},s_j} \Big)^n = \Big[1 - \sum_{s_j\in\{s\}} \boldsymbol{P}_0^{\text{out},s_j} \Big]^{-1} = \Big[1 + \boldsymbol{\alpha}_0 \sum_{s_j\in\{s\}} \boldsymbol{W}^{s_j} \Big]^{-1} \tag{54}$$

and we have, for the polarizability of a molecule inside the lattice,

$$\boldsymbol{\alpha}^{\text{lattice}} = \boldsymbol{\Lambda}^{\text{lattice,out}}\boldsymbol{\alpha}_0 = \Big[1 + \boldsymbol{\alpha}_0 \sum_{s_j\in\{s\}} \boldsymbol{W}^{s_j} \Big]^{-1} \boldsymbol{\alpha}_0. \tag{55}$$

Similarly,

$$\boldsymbol{\Lambda}^{\text{lattice,in}} = \Big[1 + \Big(\sum_{s_j\in\{s\}} \boldsymbol{W}^{s_j} \Big) \boldsymbol{\alpha}_0 \Big]^{-1}. \tag{56}$$

It is worth noting that these $\boldsymbol{\Lambda}^{\text{lattice}}(\omega)$, like all $\boldsymbol{\Lambda}(\omega)$ matrices, apply to all orders of optical response functions. Hence, the expressions of Eq. (55) and (56) may be plugged, for example, into Eq. (14) to determine the hyperpolarizability of the main inside the lattice. Applications

range from molecular monolayers (Part. VI D) to chains and lattices of nanoparticles in the dipolar approximation.[53,54]

C. The nearest neighbor approximation

When finite multimers are considered, halfway between dimers and infinite lattices, the task of drawing and counting the diagrams becomes possible when the nearest neighbor approximation[55] is enforced. In this scheme, each monomer can exchange energy with its two nearest neighbors only, provided the definition of such nearest neighbors is not ambiguous (e.g., in chains of entities). It relies on the fact that the coupling matrices $\boldsymbol{W}$ depend on a high power of the inverse distances between entities (typically $|\boldsymbol{R}|^{-3}$), so that the influence of the next-to-nearest neighbors can be neglected. We illustrate the way to handle Feynman diagrams in this context for a generic trimer, before generalizing to a multimer.

1. Generic trimer

The generic trimer of distinct entities $\{1,2,3\}$ is represented in Fig. 5A. Each entity $k=1,2,3$ is characterized by its bare response functions $\boldsymbol{\alpha}_0^{(k)}$ and $\boldsymbol{\beta}_0^{(k)}$. The nearest neighbor approximation is acceptable when the distance D_{13} is "significantly bigger" than the two others. In particular, the angle θ_{13} is supposed to remain close to zero, leading to some additional screening of the interaction between monomers 1 and 3. In this frame, after diagram splitting, the sequence of energy transfers in the Feynman diagrams can be counted and sorted. They involve the two interaction matrices $\boldsymbol{W}^{12}=\boldsymbol{W}^{21}\equiv\boldsymbol{W}(\boldsymbol{R}_1-\boldsymbol{R}_2)$ and $\boldsymbol{W}^{23}=\boldsymbol{W}^{32}\equiv\boldsymbol{W}(\boldsymbol{R}_2-\boldsymbol{R}_3)$. An example is given in Fig. 5B and C for an incoming photon interacting at entity 1 and an outgoing photon at entity 2. Considering the input part of the response functions (we do not detail the "out" matrices, which follow the same derivation), it is then possible to define four elementary processes $\boldsymbol{P}_0^{1,2}$, $\boldsymbol{P}_0^{2,1}$, $\boldsymbol{P}_0^{3,2}$ and $\boldsymbol{P}_0^{2,3}$, following the generic form $\boldsymbol{P}_0^{j,k}=-\boldsymbol{W}^{jk}\boldsymbol{\alpha}_0^{(j)}$ (we drop the "in" supercripts for clarity).

Henceforth, solving the trimer problem consists in determining the three transfer matrices $\boldsymbol{\Lambda}^{(k)}$ associated with each entity: the response functions of the trimer are completely determined from Eq. (34) [i.e., $\boldsymbol{\alpha}^{(k),\mathrm{in}}=\boldsymbol{\alpha}_0^{(k)}\boldsymbol{\Lambda}^{(k),\mathrm{in}}$] at first order, and Eq. (14) at second order (provided the "out" contributions are calculated in the same way). As derived in

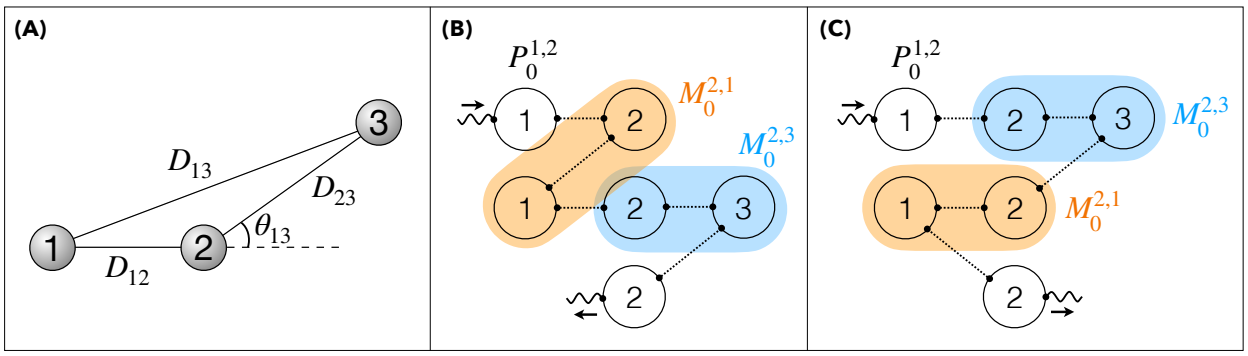

FIG. 5. (A) Geometry of a generic trimer. (B, C) Two sequences of energy exchange involved in matrix $\mathbf{\Lambda}^{(2),\mathrm{in}}$ showing the difference between $\boldsymbol{M}_0^{2,3}\boldsymbol{M}_0^{2,1}\boldsymbol{P}_0^{1,2}$ and $\boldsymbol{M}_0^{2,1}\boldsymbol{M}_0^{2,3}\boldsymbol{P}_0^{1,2}$ (cf. Appendix C).

Appendix C, the $\mathbf{\Lambda}^{(k),\mathrm{in}}$ matrices may be either explicitly calculated or determined as the solutions of the linear system

$$\begin{cases} \mathbf{\Lambda}^{(1),\mathrm{in}} - \boldsymbol{P}_0^{2,1}\,\mathbf{\Lambda}^{(2),\mathrm{in}} = \mathbf{1} \\ -\boldsymbol{P}_0^{1,2}\,\mathbf{\Lambda}^{(1),\mathrm{in}} + \mathbf{\Lambda}^{(2),\mathrm{in}} - \boldsymbol{P}_0^{3,2}\,\mathbf{\Lambda}^{(3),\mathrm{in}} = \mathbf{1} \\ - \boldsymbol{P}_0^{2,3}\,\mathbf{\Lambda}^{(2),\mathrm{in}} + \mathbf{\Lambda}^{(3),\mathrm{in}} = \mathbf{1} \end{cases} \tag{57}$$

composed of equations with matrix variables and matrix coefficients.

2. *Generic multimer*

In a multimer involving N entities, each entity k interacts with its two neighbors $(k-1)$ and $(k+1)$ analogously to the trimer case. Matrices $\mathbf{\Lambda}^{(k),\mathrm{in}}$ are therefore solutions of a system of N linear equations with matrix coefficients $\boldsymbol{P}_0^{j,k} = -\boldsymbol{W}^{jk}\boldsymbol{\alpha}_0^{(j)}$:

$$\begin{cases} \mathbf{\Lambda}^{(1),\mathrm{in}} - \boldsymbol{P}_0^{2,1}\,\mathbf{\Lambda}^{(2),\mathrm{in}} = \mathbf{1} \\ -\boldsymbol{P}_0^{1,2}\,\mathbf{\Lambda}^{(1),\mathrm{in}} + \mathbf{\Lambda}^{(2),\mathrm{in}} - \boldsymbol{P}_0^{3,2}\,\mathbf{\Lambda}^{(3),\mathrm{in}} = \mathbf{1} \\ (\cdots) \\ -\boldsymbol{P}_0^{N-2,N-1}\,\mathbf{\Lambda}^{(N-2),\mathrm{in}} + \mathbf{\Lambda}^{(N-1),\mathrm{in}} - \boldsymbol{P}_0^{N,N-1}\,\mathbf{\Lambda}^{(N),\mathrm{in}} = \mathbf{1} \\ - \boldsymbol{P}_0^{N-1,N}\,\mathbf{\Lambda}^{(N-1),\mathrm{in}} + \mathbf{\Lambda}^{(N),\mathrm{in}} = \mathbf{1} \end{cases} . \tag{58}$$

As shown in Part VI, the equations may substantially simplify according to the symmetries characterizing the geometric distribution of the entities (which have consequences on $\boldsymbol{W}$ coefficients) and their polarizabilities. The example of an N-multimer of identical molecules

adopting a linear structure is detailed in Part VI C.

VI. APPLICATIONS

In this Part, we show how the general equations of the linear and nonlinear response functions in a bipartite system or an infinite lattice apply to various situations and cover a wide range of systems. We do not include the circular processes C_n (and the normalization they induce) in the analysis in order to stick to the previous descriptions found in the literature.

A. Experimental data analysis

The present theoretical analysis provides a generic description of the response functions of the bipartite system (or the infinite lattice), and of each entity inside it, in terms of their bare response functions. It is interesting to note that the transformation from $\boldsymbol{\alpha}_0$ to $\boldsymbol{\alpha}$ involves the same $\boldsymbol{\Lambda}$ matrices as the transformation of $\boldsymbol{\beta}_0$ into $\boldsymbol{\beta}$. When the coupling has major consequences on the system, as is the case for a dimer because coupling modifies the intrinsic resonance energies of the monomers,[34,56,57] these modifications appear identically in both the first- and second-order response functions.[8,58]

There are essentially two ways to use the formulas in Eq. (26) and (34). The first one consists in modeling the bipartite system (the relative positions and orientations of the subsystems, and the bare response functions $\boldsymbol{\alpha}_0$ and $\boldsymbol{\beta}_0$ for both) in order to calculate the $\boldsymbol{P}_0$ and $\widetilde{\boldsymbol{\Lambda}}$ matrices. This way, it is possible to separately model the effect of interactions on the first- and second-order response functions, and to compare it with experiments. The second approach establishes an "all-experimental" analysis of data in bipartite systems, in the same way as was described for doubly-resonant SFG.[58,59] In this procedure, analysis of SFG data requires to first record and analyze the first-order response of the system, usually easier to measure (e.g., by absorption spectroscopy and Rayleigh scattering) than the second-order. The matrices $\widetilde{\boldsymbol{\Lambda}}^{\text{main}}$ at frequencies ω_1, ω_2 and ω_3 are then extracted from $\boldsymbol{\alpha}^{\text{main}}(\omega)$ using Eq. (34). In a second step, plugging them into Eq. (26) allows to link $\boldsymbol{\beta}^{\text{main}}$ to $\boldsymbol{\beta}_0^{\text{main}}$ without the need to model the details of the system.

Even if the "all-experimental" procedure seems attractive, one should not forget that sev-

eral issues need be tackled in order to implement it: (i) in the first-order response function, we have seen that main and partner contributions are entangled in the complete $\boldsymbol{\alpha}^{\text{syst}}$ function. Separation of $\boldsymbol{\alpha}^{\text{main,out}}$ or $\boldsymbol{\alpha}^{\text{main,in}}$ from $\boldsymbol{\alpha}^{\text{syst}}$ requires a fine knowledge of the system and the relative contributions of both subsystems. This point becomes easier when experimental conditions make it possible to neglect (or quantify) the $\boldsymbol{\alpha}_0^{\text{part}}$ contribution in Fig. 1A; (ii) for an oriented system, a polarization analysis is required in order to access several components of $\boldsymbol{\alpha}^{\text{syst}}$ and $\widetilde{\boldsymbol{\Lambda}}^{\text{main}}$; (iii) knowledge of $\boldsymbol{\alpha}_0^{\text{main}}$ is required. Here, the bare response functions of each subsystem may be approximated by the response functions of the isolated entities, which can be measured or estimated by *ab initio* calculations; (iv) finally, contributions to the SFG signal from $\boldsymbol{\beta}^{\text{main}}$ and $\boldsymbol{\beta}^{\text{part}}$ must be separated. In the literature, it is usually assumed that the SFG outputs of the main and partner response functions are separable in the experimental data. This is the essence of expressions like $\boldsymbol{\chi}^{(2)} = \boldsymbol{\chi}^{(2)}_{\text{mol}} + \boldsymbol{\chi}^{(2)}_{\text{sub}}$, or rather $\boldsymbol{\chi}^{(2)} = \boldsymbol{\chi}^{(2)}_{\text{mol}} + \boldsymbol{\chi}^{(2)}_{\text{NR}}$, often encountered in SFG publications,[60–62] where $\boldsymbol{\chi}^{(2)}$ functions represent macroscopic averages of the $\boldsymbol{\beta}$ in a {molecule + substrate} system, and "NR" stands for non-resonant. As we have seen, it is indeed possible to define second order quantities in the system which unambiguously describe the response of the main and the partner, respectively, by setting the nonlinear vertex on either of them. However, experimentally, these response functions coherently produce SFG photons adding up in phase within the measured signal (i.e., interfering in the measured intensities $I_{SFG} \propto |\boldsymbol{\chi}|^2$). When recorded together, $\boldsymbol{\chi}^{(2)}_{\text{mol}}$ and $\boldsymbol{\chi}^{(2)}_{\text{NR}}$ cannot be straightforwardly separated. Usually, their separation is achieved through the analysis of their resonances (i.e., by spectroscopic measurements), which are supposed to belong to different ranges of the electromagnetic spectrum. Practically, for IR-vis SFG, an infrared scan followed by curve fitting allows to separate the molecular response, vibrationnally-resonant, from the substrate, typically free of resonance in this range (hence the name "non-resonant"). Other methods like background-suppression[11] techniques directly lead to the main (molecular) response alone.

B. Dimers of nanoparticles and molecules

1. *Scalar polarizabilities: the case of nanospheres*

As established in Parts III and IV, the relationships between the dressed and bare response functions involve products of matrices. In the simplest case of isotropic nanoparticles (i.e., nanospheres with radius a and dielectric function ε), the bare polarizability $\boldsymbol{\alpha}_0 = \alpha_0 \mathbf{1}$ is scalar and the relationships with its dressed counterparts transform into much simpler equations. For instance, in the quasistatic approximation, we have

$$\alpha_0 = 4\pi\varepsilon_0 a^3 \frac{\varepsilon - 1}{\varepsilon + 2} \tag{59}$$

and size-dependent corrections to this formula exist.[63] Considering a homodimer of such nanoparticles, we can calculate the modifications of their optical responses as a function of the gap between them. It is known that the general optical description of a dimer of particles requires an infinite multipolar analysis, far beyond the dipolar and even quadrupolar and magnetic approximations.[64] Here we focus on a dimer for which the interdistance is big enough to neglect all higher order terms except the dipolar one. Considering the characteristic size of the system, the complete electrodynamic description of $\boldsymbol{W}$ is required. The dimer is defined by the center-to-center distance D along axis $\hat{\boldsymbol{x}}$ in a $(\hat{\boldsymbol{x}}, \hat{\boldsymbol{y}}, \hat{\boldsymbol{z}})$ frame, where D also quantifies the distance between the dipoles associated with each sphere. In this frame, the matrices $\boldsymbol{W}$ and $\boldsymbol{\Lambda}^{\text{dimer,in}} = \boldsymbol{\Lambda}^{\text{dimer,out}}$ are also diagonal, with

$$\begin{aligned} W_{xx} &= \frac{e^{\imath\omega D/c}}{4\pi\varepsilon_0 D^3}[-2 + 2\imath\omega D/c], \\ W_{yy} = W_{zz} &= \frac{e^{\imath\omega D/c}}{4\pi\varepsilon_0 D^3}[1 - \imath\omega D/c - \omega^2 D^2/c^2]. \end{aligned} \tag{60}$$

From Eq. (23), we then obtain the diagonal relationships

$$\begin{aligned} \alpha_{xx} &= \frac{\alpha_0}{1 + \alpha_0 W_{xx}} = \frac{\alpha_0}{1 - \dfrac{\alpha_0 e^{\imath\omega D/c}}{2\pi\varepsilon_0 D^3}\left[1 - \dfrac{\imath\omega D}{c}\right]}, \\ \alpha_{yy/zz} &= \frac{\alpha_0}{1 + \alpha_0 W_{zz}} = \frac{\alpha_0}{1 + \dfrac{\alpha_0 e^{\imath\omega D/c}}{4\pi\varepsilon_0 D^3}\left[1 - \dfrac{\imath\omega D}{c} - \dfrac{\omega^2 D^2}{c^2}\right]}, \end{aligned} \tag{61}$$

where we recover the formulas from the literature.[65,66] This may be extended to other particle shapes (e.g., cubes, nanorods) following the methods described in the next Parts, provided that the relative orientations of the particles are taken into account and their bare polarizabilities determined by either experiments, modeling or calculation.

In the same line, the polarizability of a binary system of spheres with different polarizabilities (e.g., different radii or materials) can be calculated from the general equations (20), (21) and (35). Dealing with scalar polarizabilities and diagonal matrices, there is again no difference between "in" and "out" quantities, and the polarizability of the system may unambiguously be split into a distinct polarizability for each sphere:

$$\begin{aligned}\alpha_{xx}^{\text{main}} &= \frac{1-\alpha_0^{\text{part}}W_{xx}}{1-\alpha_0^{\text{main}}\alpha_0^{\text{part}}(W_{xx})^2}\,\alpha_0^{\text{main}},\\ \alpha_{yy/zz}^{\text{main}} &= \frac{1-\alpha_0^{\text{part}}W_{zz}}{1-\alpha_0^{\text{main}}\alpha_0^{\text{part}}(W_{zz})^2}\,\alpha_0^{\text{main}},\end{aligned} \tag{62}$$

where W_{xx} and W_{zz} are still given by Eq. (60), or their electrostatic limits.[53,67] Equation (62) generalizes the results obtained by A. Pinchuk *et al.* in the case of a dimer of nanospheres.[66]

As for the SFG process, the bare hyperpolarizability of spheres vanishes for symmetry reasons at the dipolar level of approximation but may be relevant at the quadrupolar (Q) and magnetic (m) levels.[68,69] The various contributions to the sphere hyperpolarizability inside a dimer are then determined from Eq. (14) and involve, in the $(\hat{\boldsymbol{x}},\hat{\boldsymbol{y}},\hat{\boldsymbol{z}})$ frame, products of functions involving the same denominators as in Eq. (61) evaluated at the three frequencies $\omega_1,\omega_2,\omega_3$. However, its calculation requires a development beyond the dipolar coupling and involves hybrid response functions like $\boldsymbol{\beta}_0^{\text{Qee}}$ or $\boldsymbol{\beta}_0^{\text{eem}}$, together with coupling constants of the (em) or (Qe) types, to name a few.[31]

2. Diagonal polarizabilities: the case of molecules

Given their anisotropic microscopic structure, molecules do not exhibit scalar polarizabilities. However, depending on their symmetries, it is often possible to find a diagonal representation of $\boldsymbol{\alpha}_0$. It is then possible to generalize Eq. (61) and (62) to the cases of molecular dimers, as soon as the bare polarizabilities of the two molecules are diagonal within the same basis $(\hat{\boldsymbol{x}},\hat{\boldsymbol{y}},\hat{\boldsymbol{z}})$, with $\hat{\boldsymbol{x}}$ aligned with their intermolecular direction. On the one hand, Eq. (61) can be used to describe a molecular homodimer, provided that α_0 is

replaced by the appropriate component among $(\alpha_0)_{xx}$, $(\alpha_0)_{yy}$ and $(\alpha_0)_{zz}$. It is worth noting that the electrodynamic terms in $\boldsymbol{W}$ may be dropped in this case to keep the electrostatic contribution only. On the other hand, Eq. (62) can be applied to a molecular heterodimer by replacing α_0^{main} and α_0^{part} by their xx, yy and zz components.

Contrary to isotropic nanoparticles whose hyperpolarizabilities vanish at the dipolar level due to centrosymmetry, the second-order response of a molecular binary system is dominated by the dipolar contribution. Following Eq. (14), the hyperpolarizabilities are then given by

$$\beta_{ijk} = \Lambda_{ii}^{\text{homo}}(\omega_3)\Lambda_{jj}^{\text{homo}}(\omega_1)\Lambda_{kk}^{\text{homo}}(\omega_2) \cdot (\beta_0)_{ijk} \tag{63}$$

and

$$\beta_{ijk}^{\text{main}} = \Lambda_{ii}^{\text{hetero}}(\omega_3)\Lambda_{jj}^{\text{hetero}}(\omega_1)\Lambda_{kk}^{\text{hetero}}(\omega_2) \cdot (\beta_0)_{ijk}, \tag{64}$$

for the homo- and the heterodimer cases, respectively, with

$$\begin{aligned} \Lambda_{ii}^{\text{homo}} &= [1 + (\alpha_0)_{ii} W_{ii}]^{-1} \\ \Lambda_{ii}^{\text{hetero}} &= \frac{1 - (\alpha_0^{\text{part}})_{ii} W_{ii}}{1 - (\alpha_0^{\text{main}})_{ii} (\alpha_0^{\text{part}})_{ii} (W_{ii})^2}. \end{aligned} \tag{65}$$

3. Resonant polarizabilities: frequency shifts

When the entities involved in the dimer have a resonant polarizability, they may absorb light and show enhanced scattering properties for specific frequencies of light. This is typically the case with molecular chromophores, plasmonic nanoparticles and semiconductor quantum dots in the UV-visible range, or molecular vibration modes in the infrared range. The existence of a coupling between the two entities usually induces a shift in these resonance frequencies, as evidenced by optical spectroscopy.[34,57,70] The amplitude of this frequency shift is implicitly included in the previous equations, still it does not explicitly show up.

To understand the origin of such frequency shifts, we consider a homodimer along the $\hat{\boldsymbol{x}}$ axis made of entities with a single optical resonance frequency ω_0 (e.g., allowed molecular electronic transition, surface plasmon resonance, electron-hole exciton). The bare optical

polarizability $\boldsymbol{\alpha}_0(\imath\omega)$ of each monomer flows from the general expression[29]

$$(\alpha_0)_{ij} = \sum_{m,n} \frac{\rho(\omega_m)}{\hbar} \left(\frac{p^i_{nm}\, p^j_{mn}}{\imath\omega + \omega_{nm}} - \frac{p^j_{nm}\, p^i_{mn}}{\imath\omega - \omega_{nm}} \right), \tag{66}$$

where p^i_{nm} is the transition dipole moment along i between states m and n, and the optical resonances appear after analytic continuity from $\imath\omega \pm \omega_{nm}$ to $\omega \pm \omega_{nm} + \imath\Gamma_{nm}$. Upon the common hypothesis of strong coupling,[71] the exchange of virtual photons is supposed faster than the dissipative processes,[72] leading to neglect the widths of the optical transitions as compared with the interaction frequencies. The monomer polarizability is then reduced to its resonant contribution, with m as the ground state $|g\rangle$ and n as the excited state $|e\rangle$:

$$(\alpha_0)_{ij} = \frac{1}{\hbar} \frac{p^i_{ge}\, p^j_{eg}}{\omega_0 - \imath\omega}. \tag{67}$$

For a transition dipole moment aligned along $\hat{\boldsymbol{x}}$, $\hat{\boldsymbol{y}}$ or $\hat{\boldsymbol{z}}$, we plug this value in Eq. (61) and get, for the dimer:

$$\alpha_{ii} = \frac{(\alpha_0)_{ii}}{1 + (\alpha_0)_{ii} W_{ii}} = \frac{1}{\hbar} \frac{p^i_{eg}\, p^i_{ge}}{\omega_0 + \dfrac{p^i_{eg}\, W_{ii}\, p^i_{ge}}{\hbar} - \imath\omega}. \tag{68}$$

The quantity $p^i_{eg}\, W_{ii}\, p^i_{ge} \equiv \hbar\omega^{\text{int}}_{ii}$ is the dipole-dipole coupling constant between the excited states of the two monomers, quantifying the possibility for a $(g \to e)$ excitation to jump from one monomer to the other and become delocalized over the whole dimer.[73] From Eq. (68), we see that the optical resonance frequency is shifted by ω^{int}_{ii}. In the electrostatic limit, from Eq. (60), $W_{xx} = -2/4\pi\varepsilon_0 D^3$ is negative, whereas $W_{zz} = 1/4\pi\varepsilon_0 D^3$ is positive. Thus, the resonance frequency is redshifted for longitudinal coupling ($\omega_- = \omega_0 - |\omega^{\text{int}}_{xx}|$), and blueshifted for transverse coupling ($\omega_+ = \omega_0 + \omega^{\text{int}}_{zz}$). This behavior is well-known as a consequence of plasmonic coupling between two nanostructures.[51,74] For molecular chromophores, this accounts for the ω^{int}_{xx}-redshift of the resonance in J-type dimers and the ω^{int}_{zz}-blueshift in H-type dimers, both originating from excitonic coupling[73,75,76] ($2\hbar\omega^{\text{int}}_{ii}$ is then called the Davydov splitting).

Most of the times, the molecules are not aligned and the matrices $\boldsymbol{\alpha}_0$ differ for the main and the partner. It is true for homodimers and, all the more, for heterodimers. In principle,

the new resonance frequencies of the system are determined by finding the poles of the inverse matrix in Eq. (20) or (21) and selecting those which do not match the roots of $\mathbf{1}+\boldsymbol{P}_0^{\mathrm{part}}$. However, it is more efficient to turn back to Eq. (17) or (18). Taking the example of the "out" part of the response:

$$\boldsymbol{\alpha}^{(1),\mathrm{out}} = \left[\mathbf{1}+\boldsymbol{P}_0^{(2),\mathrm{out}}\right]\sum_{n=0}^{\infty}\left[\boldsymbol{\alpha}_0^{(1)}\boldsymbol{W}\boldsymbol{\alpha}_0^{(2)}\boldsymbol{W}\right]^n\boldsymbol{\alpha}_0^{(1)}, \tag{69}$$

where the labels "1" and "2" stand for the main and the partner, respectively. Assuming that the bare polarizabilities of both entities follow Eq. (67) with resonance frequencies ω_1 and ω_2, respectively, it is possible to determine the frequency shift $\omega_{12}^{\mathrm{int}}$ through the explicit expansion of the power series in Eq. (69). The derivation is performed in Appendix D. For a generic dimer whose geometry is characterized by the three angles $\theta_{12}=(\widehat{\boldsymbol{p}_1,\boldsymbol{p}_2})$, $\theta_1=(\widehat{\boldsymbol{p}_1,\hat{\boldsymbol{x}}})$ and $\theta_2=(\widehat{\boldsymbol{p}_2,\hat{\boldsymbol{x}}})$, an orientation factor is then introduced in the Davydov splitting,[70,73,77]

$$\hbar\omega_{12}^{\mathrm{int}} = \frac{|\boldsymbol{p}_1|\,|\boldsymbol{p}_2|}{4\pi\varepsilon_0 D^3}\left(\cos\theta_{12}-3\cos\theta_1\cos\theta_2\right), \tag{70}$$

and the two new resonance frequencies are given by

$$\omega_\pm = \frac{\omega_1+\omega_2}{2} \pm \frac{\sqrt{(\omega_1-\omega_2)^2+4(\omega_{12}^{\mathrm{int}})^2}}{2}. \tag{71}$$

Both transitions are optically allowed, as expected from the literature,[77] except for a homodimer with parallel molecules[73] (in this case, only ω_+ remains optically active, as above; see Appendix D for details).

All these simple examples provide ready-to-use expressions for the frequency shifts induced by the homo- or hetero-dimerization processes in molecular or nanoparticle systems.[78] They also show that the resonance frequencies of binary systems, modified with respect to those of the bare entities, show up indeed when the energy dispersion of the optical response functions is analyzed, albeit buried into the $\boldsymbol{\Lambda}$ matrix description of the interaction. While the dispersion properties are intrinsic to $\boldsymbol{\Lambda}$, they are actually shared by the response functions at all orders. It is easy to see that, for the outgoing component of the second-order

response function, Eq. (69) transforms into

$$\beta_{ijk}^{(1)} = \sum_{i'} \Big(\big[\mathbf{1} + \boldsymbol{P}_0^{2,\mathrm{out}} \big] \sum_{n=0}^{\infty} \Big[\boldsymbol{\alpha}_0^{(1)} \boldsymbol{W} \boldsymbol{\alpha}_0^{(2)} \boldsymbol{W} \Big]^n \Big)_{ii'} (\beta_0^{(1)})_{i'jk}. \tag{72}$$

Introducing the $\beta_0^{(1)}$ response function in terms of dipolar transition moments,[29,33] we may apply the same procedure as in Eq. (D2) to recover an expression analogous to Eq. (D4) in terms of $\boldsymbol{\beta}$ components. This proves that the resonance frequencies of the dimer, given by Eq. (71), are the same at all orders of the linear and nonlinear response, whatever the optical process. As we show below, this may be generalized to a multimeric system with an arbitrary number of units.

C. Molecular excitons in linear multimers

Excitonic coupling in molecular aggregates is attracting a renewed interest because such objects experience extreme delocalization of optical excitation,[79] which can be strongly coupled to cavity modes[72,80] or surface plasmons,[81] and show interesting charge transfer properties as organic semiconductors.[82] As reported in the literature, it is possible to calculate the electronic structure of a molecular aggregate composed of N identical monomers in the nearest neighbor approximation in order to subsequently determine its optical response.[55] However, within our approach, the knowledge of the electronic structure of the whole aggregate is not required. Here we show that it is indeed possible to determine the optically active electronic transitions involved in linear and nonlinear optics in such a multimer directly from the diagrammatic theory.

We start from the results of Part V C 2. When the N entities are identical (with a diagonal polarizability $\boldsymbol{\alpha}_0$) and aligned with a regular spacing D on axis $\hat{\boldsymbol{x}}$, all matrices are diagonal and equations along $\hat{\boldsymbol{x}}$ and $\hat{\boldsymbol{z}}$ separate. In addition, "in" and "out" processes are identical: matrices $\boldsymbol{P}_0^{j,k}$ are all equal to a single $\boldsymbol{P}_0$ with $(P_0)_{xx} = -(\alpha_0)_{xx} W_{xx}$ and $(P_0)_{zz} = -(\alpha_0)_{zz} W_{zz}$. Using Eq. (67) and setting again $\omega_{ii}^{\mathrm{int}} = p_{eg}^i \, W_{ii} \, p_{ge}^i / \hbar$ for $i \in \{x, z\}$, the matrix system in Eq. (58) transforms into two independent systems of linear equations

for the xx and zz components of $\mathbf{\Lambda}^{(i)}$, with the same structure:

$$\begin{cases} (\omega_0 - \imath\omega) & \Lambda_{ii}^{(1)} & + & \omega_{ii}^{\text{int}} & \Lambda_{ii}^{(2)} & & & = & (\omega_0 - \imath\omega) \\ \omega_{ii}^{\text{int}} & \Lambda_{ii}^{(1)} & + & (\omega_0 - \imath\omega) & \Lambda_{ii}^{(2)} & + & \omega_{ii}^{\text{int}} \ \Lambda_{ii}^{(3)} & = & (\omega_0 - \imath\omega) \\ & & & (\cdots) & & & & & \\ \omega_{ii}^{\text{int}} & \Lambda_{ii}^{(N-2)} & + & (\omega_0 - \imath\omega) & \Lambda_{ii}^{(N-1)} & + & \omega_{ii}^{int} \ \Lambda_{ii}^{(N)} & = & (\omega_0 - \imath\omega) \\ & & + & \omega_{ii}^{\text{int}} & \Lambda_{ii}^{(N-1)} & + & (\omega_0 - \imath\omega) \ \Lambda_{ii}^{(N)} & = & (\omega_0 - \imath\omega) \end{cases} . \tag{73}$$

Such a system can be written as $\boldsymbol{T}_{ii}^{(N)}(\Lambda_{ii}^{(k)})_{k=1}^{N} = (\omega_0 - \imath\omega)\,\mathbf{1}$, where $(\Lambda_{ii}^{(k)})_{k=1}^{N}$ is the vector of the $\Lambda_{ii}^{(k)}$ components (i.e., the solutions of the system) and $\boldsymbol{T}_{ii}^{(N)}$ is a tridiagonal Toeplitz matrix.[83] In order to extract the new resonance frequencies ω_k of the multimer, it is not required to solve the system, but simply to look for the poles of the solutions (i.e., the values of $\imath\omega$ which cancel their denominator). They are given by the roots of the determinant of matrix $\boldsymbol{T}_{ii}^{(N)}$:[83]

$$\omega_k = \omega_0 + 2\omega_{ii}^{\text{int}} \cos\left(\frac{k\pi}{N+1}\right) \tag{74}$$

with $k \in \{1, \cdots, N\}$. We recover in this way the fact that the energy levels are regularly spread on both sides of ω_0 with a maximum shift of $\pm 2\omega_{ii}^{\text{int}}$, with a decreasing activity for increasing k. For symmetry reasons, only half of them are optically active, for odd values of k,[55] as can be recovered by implementing the symmetry condition $\Lambda_{ii}^{(k)} = \Lambda_{ii}^{(N+1-k)}$ in Eq. (73). In line with the classical approaches reported in the literature, the diagrammatic method thus enables us to describe the optical properties of multimers. For instance, considering the limit case of an infinite polymer (i.e., $N \to \infty$), we find that there is only one mode allowed at $\omega_1 = \omega_0 + 2\omega_{ii}^{\text{int}}$, giving rise to superradiance.[76,79,84] It is interesting to compare this result (for an infinite multimer in the nearest neighbor approximation) with the optical response of an infinite 1D lattice, which takes all the site-to-site interactions into account. From Eq. (55), we see that there is indeed only one optically active resonance frequency for the system, satisfying $1 + (\alpha_0)_{ii} \sum_s W_{ii}^s = 0$. This is logical as all entities are now equivalent in the system and share a common dressed polarizability. This leads to a resonance at $\omega_0 + \omega_{ii}^{\text{int}} \sum_{-\infty}^{+\infty} 1/|k|^3 \approx \omega_0 + 2.4\,\omega_{ii}^{\text{int}}$, meaning that the nearest neighbor approximation misses about 20% of the interaction as compared to the complete calculation.[85]

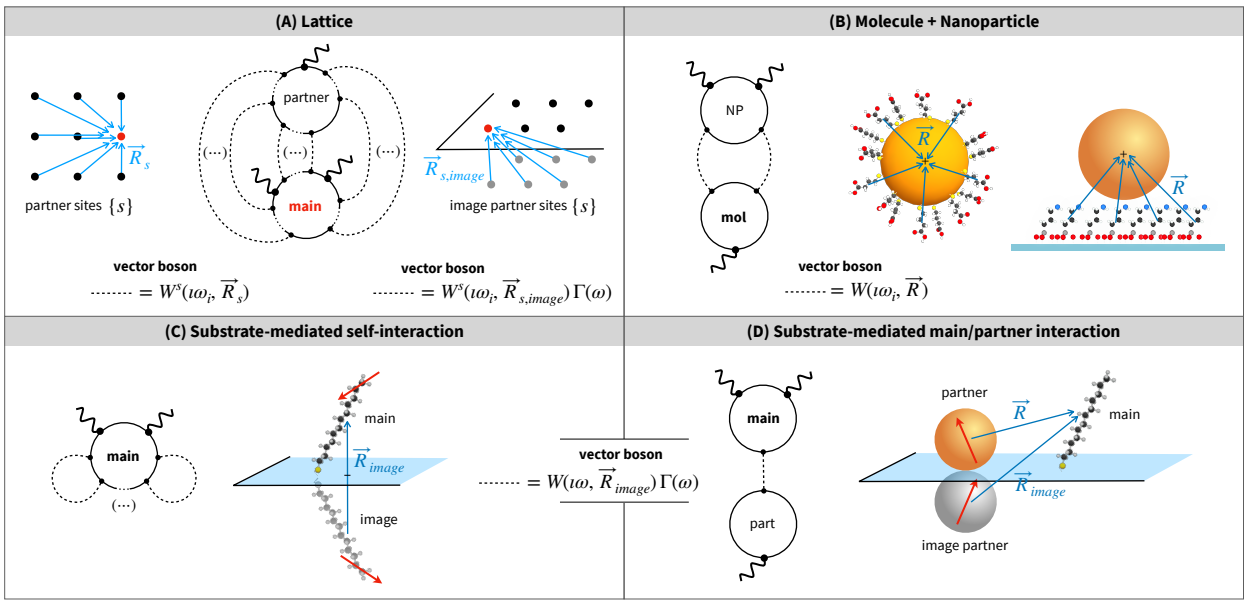

FIG. 6. Examples of bipartite and multipartite systems. (A) Lattice system; second-order response. (B) {molecule + nanoparticle} system; second-order response. (C) Self-image in {main + substrate} system; first-order response. (D) $\boldsymbol{P}_0$ process in {molecule + partner + substrate} system. The symbol "(...)" represents exchange of an infinite number of bosons. The values of the coupling matrices are indicated on the panels.

D. Molecular monolayer

A typical case of a lattice geometry is a molecular monolayer adsorbed on a substrate. Eq. (55) directly applies to provide the lattice-modified polarizability. Our matrix description allows to tackle the very general case, involving especially components of the response functions parallel and perpendicular to the surface. In most practical cases, the lattice geometry is simplified to extract single components. On a metal, for example, only the z components of the response functions (i.e., perpendicular to the surface) are considered, whereas the $\boldsymbol{W}^s$ coupling matrix involves positions $\boldsymbol{R}_s$ limited by construction to the (x, y) plane. In this case, $(\alpha_0)_{zz}$ is the only relevant component and Eq. (55) transforms into

$$\alpha_{zz}^{\text{lattice}} = \frac{(\alpha_0)_{zz}}{1 + (\alpha_0)_{zz} \sum_{\{s\}} W_{zz}^s}, \tag{75}$$

where W_{zz}^s is usually approximated in the literature by its electrostatic value $W_{zz}^s = 1/4\pi\varepsilon_0 |\boldsymbol{R}_s|^3$, allowing the explicit analytical or numerical computation of $\sum W_{zz}^s$ for elementary lattices.[86–88] In the case of vibrational spectroscopy, this model accounts for vibrational

frequency shifts by delocalization of the vibrations inside a dense molecular monolayer of adsorbates.[34] For uniaxial molecules oriented along $\hat{\boldsymbol{z}}$ [i.e., $(\alpha_0)_{zz} = \alpha_0$], we recover the result of Ref. 89. Generalization to uniaxial tilted molecules follows straightforwardly by defining θ as the molecular tilt angle with respect to the normal direction $\hat{\boldsymbol{z}}$, and by using $(\alpha_0)_{zz} = \alpha_0 \cos^2\theta$.[9] These results identically represent the first-order response of an array of nanoparticles on a metal, provided that the dipolar approximation is enforced. When the molecular layer encompasses two different molecules (represented by two different polarizabilities) or an incomplete monolayer, several options exist to calculate the polarizability in the lattice.[34,90,91] According to the coherent potential approach,[34] the loops are assigned an effective polarizability as a weighted average of the two molecular polarizabilities (or of the molecular polarizability and zero for an incomplete monolayer). This can be recovered in our description [Eq. (52)] by introducing two families of interchangeable $\boldsymbol{P}_0$ processes corresponding to the weighted molecular $\boldsymbol{P}_0^{\mathrm{mol},1}$ and $\boldsymbol{P}_0^{\mathrm{mol},2}$. As these two new operators share common $\boldsymbol{W}^s$ and differ only through their polarizabilities, the $\boldsymbol{\Lambda}^{\mathrm{lattice}}$ matrices in Eq. (55) and (56) are finally modified by replacing $\boldsymbol{\alpha}_0$ by the weighted sum of the polarizabilities of the two entities present in the lattice.

In the same way, the hyperpolarizability component β_{zzz} is obtained from Eq. (14):

$$\beta_{zzz}^{\mathrm{lattice}} = \frac{(\beta_0)_{zzz}}{[1+\alpha_0^{zz}(\omega_1)W^{\Sigma}]\,[1+\alpha_0^{zz}(\omega_2)W^{\Sigma}]\,[1+\alpha_0^{zz}(\omega_3)W^{\Sigma}]}. \tag{76}$$

Here, $(\alpha_0)_{zz}$ is written as α_0^{zz} for legibility, and W^{Σ} stands for $\sum_{\{s\}} W_{zz}^s$. Hence, Eq. (76) coincides with the results of Refs. 8 and 9, and provides in addition a generalization to incomplete or mixed monolayers and tilted molecules through the same modifications of the $\boldsymbol{\Lambda}$ factors as in the first-order case.[9] In these systems, the influence of the substrate is usually important and should be taken into account as explained in Part VI F.

E. Molecule coupled with a nanoparticle

A molecule grafted onto a nanoparticle constitutes the archetype of a strongly asymmetric bipartite system, this significant asymmetry allowing us to simplify the equations. On the one hand, the molecular polarizability ($\boldsymbol{\alpha}_0^{\mathrm{mol}}$) can be considered as much smaller than the nanoparticle polarizability ($\boldsymbol{\alpha}_0^{\mathrm{NP}}$). On the other hand, the long center-to-center distance

between the two entities (given the size of the particle) implies that $\boldsymbol{W}$ remains smaller than in the molecule-molecule systems. Consequently, $\boldsymbol{P}_0^{\mathrm{mol,out}}$ and $\boldsymbol{P}_0^{\mathrm{mol,in}}$ can be considered small quantities as compared to $\mathbf{1}$. We deduce the following relations at the lowest order of approximation:

$$\begin{aligned}
&\boldsymbol{\Lambda}^{\mathrm{mol,in}} \approx \mathbf{1} + \boldsymbol{P}_0^{\mathrm{NP,in}} \text{ and } \boldsymbol{\Lambda}^{\mathrm{mol,out}} \approx \mathbf{1} + \boldsymbol{P}_0^{\mathrm{NP,out}}, \\
&\boldsymbol{\Lambda}^{\mathrm{NP,in}} \approx \boldsymbol{\Lambda}^{\mathrm{NP,out}} \approx \mathbf{1}.
\end{aligned} \tag{77}$$

For the nanoparticle, $\boldsymbol{\alpha}^{\mathrm{NP}} \approx \boldsymbol{\alpha}_0^{\mathrm{NP}}$ and $\boldsymbol{\beta}^{\mathrm{NP}} \approx \boldsymbol{\beta}_0^{\mathrm{NP}}$. In other words, the response functions of the particle are unaffected by the dipole-dipole interaction with the molecule located in its vicinity. This approximation is very convenient to model the optical properties of decorated nanoparticles by use of simple models adapted to bare particles (e.g., quasistatic or electrostatic approximation[92]). It also makes the "all-experimental" data analysis, as desribed in Part VI A, rather straightforward in such systems. As for the molecule, we find that its hyperpolarizability becomes

$$\beta_{ijk}^{\mathrm{mol}} = \sum_{i'j'k'} \left[1 - \alpha_0^{\mathrm{NP}} W(\omega_3)\right]_{ii'} \left[\beta_0^{\mathrm{mol}}\right]_{i'j'k'} \left[1 - W\alpha_0^{\mathrm{NP}}(\omega_1)\right]_{j'j} \left[1 - W\alpha_0^{\mathrm{NP}}(\omega_2)\right]_{k'k}, \tag{78}$$

where the frequency arguments have been dropped in $\boldsymbol{\beta}$ functions for clarity. This result matches the original calculation of the modification of the molecular hyperpolarizability by dipolar coupling with the nanoparticle, performed by classical local field[93] and diagrammatic[30] methods. The $\boldsymbol{P}_0^{\mathrm{NP}}$ processes describe the indirect excitation of the molecule through the direct excitation of the nanoparticle followed by a transfer to the molecule (Fig. 6B). The consequence is the well-known amplification of the molecular SFG signals by surface plasmon resonances (for a plasmonic particle[15]) or excitons (in quantum dots[17]).

The next order of approximation in matrices $\boldsymbol{\Lambda}$ involves an additional energy exchange between the partner and the molecule ($\boldsymbol{P}_0^{\mathrm{NP,in}}\boldsymbol{P}_0^{\mathrm{mol,in}}$ and $\boldsymbol{P}_0^{\mathrm{mol,out}}\boldsymbol{P}_0^{\mathrm{NP,out}}$), so that the molecule is the only place where the light-matter interaction takes place. It is reasonable to consider that these processes are negligible when compared to the ones described by Eq. (78) because the dominant optical process lies in the excitation of the nanoparticle by light combined with a molecule-NP energy transfer—not in the dipolar feedback from the molecule to itself. Moreover, discarding the higher order terms of $\boldsymbol{\Lambda}$ amounts to considering that, when several

molecules are present (i.e., when the nanoparticle is decorated with molecules), there is no nanoparticle-mediated dipole-dipole coupling between them: the individual responses of all molecules simply add up in the system without interference from the particle.[93] The direct dipolar coupling between molecules at the surface of the particle may still contribute, and should be treated in the same way as in Part VI D. Finally, the simple expressions of the $\mathbf{\Lambda}$ matrices in Eq. (77) and (78) allow to easily implement the magnetic and quadrupolar contributions to the interaction Hamiltonians in such systems, as was done in Ref. 31.

F. Substrate as an additional partner

1. Diagrammatic description of the image dipoles

We examine here the systems in which the partner is a plane substrate (i.e., the surface of a semi-infinite bulk). Such a partner looks different in nature from the previous examples because its optical response does not take the form of a single dipole at a specific location. However, the optical response of a bulk (considered as a polarizable dielectric medium) is interpreted as the creation of a bulk density of elementary dipoles, which sum up to create a macroscopic polarization from which the dielectric function and refractive index of the bulk follow. The {main + substrate} system may require a fine-grained analysis at the atomic scale in order to precisely model their interaction. Here, we limit ourselves to the dipolar description of the system, which may be viewed as a coarse-grained approximation. Essentially, we have[94] $\boldsymbol{P} = \varepsilon_0 \boldsymbol{\chi} \boldsymbol{E}$ where $\boldsymbol{E}$ is the macroscopic field. $\boldsymbol{P} = N\langle \boldsymbol{p} \rangle$ is the macroscopic polarization, defined as the volume density of elementary dipoles, and $\boldsymbol{\chi} = \boldsymbol{\varepsilon} - \mathbf{1}$ is the susceptibility tensor, related to the first-order response function $\boldsymbol{\alpha}$, giving birth to the elementary dipoles, through $\varepsilon_0 \boldsymbol{\chi} = N\boldsymbol{\alpha}$ (when local field effects are neglected). Local field effects may be incorporated through the Clausius-Mossotti equation, for example.

The susceptibility tensor $\boldsymbol{\chi}$ acts as a macroscopic response function for the substrate, which could be represented by its own loop in the diagrams. Nevertheless, a substrate partner cannot be handled in the same way as a point-like partner because its optical properties mostly derive from a macroscopic integration of its microscopic properties. Instead, we will include its effects through a new kind of vector bosons, embodying the influence of the substrate on the main-main and main-partner interactions. In other words, the substrate

is not associated to a loop but incorporated into the diagrams through substrate-mediated vector bosons. This approach excludes the possibility of describing the direct light excitation of the substrate, but it is actually not our will because it is easily described in a classical point of view by reflectivity properties of the substrate and Fresnel coefficients for the main.[95–97] It is more interesting to focus on the diagrams where the substrate does not interact directly with light, but only through dipolar energy exchange with the other entities (main and partner), as described in Figure 6C and D. The substrate becomes in this case an intermediate during an energy exchange between the main and a partner, or between the main and itself. In the example of Fig. 6C, the presence of the substrate introduces new vector bosons which can indeed be connected on the same loop, hence accounting for a substrate-mediated self-interaction. On a mechanistic point of view, a first interaction of light with the main generates a dipolar response which is transferred to the substrate, then conveyed back to the main where it modifies its response function. Here, we recognize the phenomenon classically known as the "image dipole": the dipole initially created at the main modifies the optical response of the system in the same way as would do its own image created inside the substrate at the same distance from the surface. The reader may refer to textbooks[94] in order to review the power and limits of this description, which is not restricted to the dipole case but may extend to a multipolar description. As the phenomenon is easily modeled using simple electromagnetic analogs, we can directly implement it within diagrams. The image, contrary to an ordinary partner, does not have a proper $\boldsymbol{\alpha}_0$ response function because its polarizability is proportional to the polarizability of the main. Noting $\hat{\boldsymbol{z}}$ the axis perpendicular to the substrate surface in an $(\hat{\boldsymbol{x}}, \hat{\boldsymbol{y}}, \hat{\boldsymbol{z}})$ frame, the image reverses the direction of dipoles parallel to $\hat{\boldsymbol{x}}$ and $\hat{\boldsymbol{y}}$, but not along $\hat{\boldsymbol{z}}$.[94] When all matrices are expressed in this frame, we have

$$\boldsymbol{\alpha}^{\text{image}} = \frac{\varepsilon_{\text{sub}} - 1}{\varepsilon_{\text{sub}} + 1} \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \boldsymbol{\alpha}^{\text{main}}, \equiv \boldsymbol{\Gamma}\, \boldsymbol{\alpha}^{\text{main}} \tag{79}$$

where ε_{sub} is the dielectric function of the substrate.

Recalling that the image dipole is not a real entity but a convenient construction mimicking the dipolar behavior outside the substrate, it cannot interact directly with light. In the

classical point of view, one of the entities composing the system (the main, for instance) interacts with light, which results in the creation of its image dipole. This image then interacts by dipolar coupling with all the entities, i.e., the main itself and all its partners. Thus, the subsystems may directly exchange energy between themselves (as was the case in all previous examples) or they may involve an intermediate step through the substrate, as represented by the image. In the diagrammatic point of view, this means that two entities A and B can interact through two kinds of vector bosons: (i) direct interactions are described by the coupling matrix $\boldsymbol{W}$, defining $\mathcal{H}_{\text{int}} = \boldsymbol{p}_A \cdot \boldsymbol{W}\boldsymbol{p}_B$ and prohibiting self-interactions ($A \neq B$), and (ii) substrate-mediated interactions are described by $\boldsymbol{W\Gamma}$, defining $\mathcal{H}'_{\text{int}} = \boldsymbol{p}_A \cdot \boldsymbol{W\Gamma}\boldsymbol{p}_B$ and allowing self-interactions (when $A = B$). Matrix $\boldsymbol{W}$ is evaluated this time at the position $\boldsymbol{R}_{\text{image}}$ of the image, that is inside the substrate, and matrix $\boldsymbol{\Gamma}$ represents the "propagation" of the initial dipole (vertex on the main loop) to its image, whose dipole moment reads $p_i^{\text{image}} = \sum_j \Gamma_{ij} p_j^{\text{main}}$. As all roundtrips through the substrate involve the same frequency, the splitting of the main loop is recovered in the case of substrate-mediated interactions.

2. Self-interaction in {main + substrate} systems

Considering a {main + substrate} system, we can read the diagrams as usual by following the energy exchanges to get,

$$\boldsymbol{\alpha}^{\text{main+image}} = \boldsymbol{\alpha}_0^{\text{main}} - \boldsymbol{\alpha}_0^{\text{main}}\boldsymbol{W\Gamma}\boldsymbol{\alpha}_0^{\text{main}} + \boldsymbol{\alpha}_0^{\text{main}}\boldsymbol{W\Gamma}\boldsymbol{\alpha}_0^{\text{main}}\boldsymbol{W\Gamma}\boldsymbol{\alpha}_0^{\text{main}} - (\cdots) \tag{80}$$

$$= \boldsymbol{\alpha}_0^{\text{main}} \left[1 + \boldsymbol{W\Gamma}\boldsymbol{\alpha}_0^{\text{main}}\right]^{-1} = \left[1 + \boldsymbol{\alpha}_0^{\text{main}}\boldsymbol{W\Gamma}\right]^{-1} \boldsymbol{\alpha}_0^{\text{main}} \tag{81}$$

after summing all the diagrams (i.e., for an increasing number of substrate-modified vector bosons). We recover the expected equality between "in" and "out" descriptions. As above, the self-consistent calculation, i.e., interaction of the bare polarizability with the dressed version of itself ($\boldsymbol{\alpha}^{\text{main+image}} = \boldsymbol{\alpha}_0 - \boldsymbol{\alpha}_0\boldsymbol{W\Gamma}\boldsymbol{\alpha}^{\text{main+image}}$), directly gives the same result. The polarizability $\boldsymbol{\alpha}^{\text{main+image}}$ is sometimes called the renormalized polarizability.[98]

A direct application follows in the modification of the polarizability of a spherical nanoparticle (with radius a) located above a substrate (Fig. 6C) at a distance $D/2$. As in the case of a dimer of particles, the calculations above apply in such a system as long as all higher multipolar orders can be neglected, that is for rather small particles (for which the quasistatic ap-

proximation is enforced) and for rather long distances between the sphere and the substrate. Using Eq. (1) for $\boldsymbol{W}$ (Eq. (2) is handled in the same way), we have $W_{xx} = W_{yy} = 1/4\pi\varepsilon_0 D^3$ and $W_{zz} = -2/4\pi\varepsilon_0 D^3$, so $\boldsymbol{W\Gamma}$ is diagonal. The polarizability of the sphere being isotropic following Eq. (59), we obtain

$$\begin{aligned}
\alpha_{xx}^{\text{NP+substrate}} &= \frac{\alpha_0}{1 - \dfrac{\alpha_0}{4\pi\varepsilon_0 D^3}\dfrac{\varepsilon_{\text{sub}} - 1}{\varepsilon_{\text{sub}} + 1}}, \\
\alpha_{zz}^{\text{NP+substrate}} &= \frac{\alpha_0}{1 - \dfrac{2\alpha_0}{4\pi\varepsilon_0 D^3}\dfrac{\varepsilon_{\text{sub}} - 1}{\varepsilon_{\text{sub}} + 1}},
\end{aligned} \tag{82}$$

which matches the results of the literature.[99] When a molecular layer is adsorbed on the substrate, the dipolar coupling with the molecules does not modify the response function of the particle (according to Part VI E) and we can use this polarizability to estimate the modifications of the molecular response in the {molecule + NP + substrate} due to the presence of the particle and its image.[93] Of course, as this description does not include all higher order multipolar terms, it cannot accurately describe the intense hot spot effects in these systems as are used in SHINERS[100] or SHINE-SFG[101] experiments. Still, it may account for the energy shifts in the resonant parts of the response functions of the subsystems as a consequence of the presence of the nanoparticle image.

3. The {main + partner + substrate} systems

For {main + partner} systems, the presence of the substrate (i.e., of their images) opens a second pathway to transfer energy from the main to the partner in a $\boldsymbol{P}_0$ process. As sketched in Fig. 6D, this new process adds to the original ones. At first-order in $\boldsymbol{P}_0$:

$$\begin{aligned}
\boldsymbol{\alpha}^{\text{main,out}} &= \left[\mathbf{1} - \boldsymbol{\alpha}_0^{\text{part}}\boldsymbol{W}(\boldsymbol{R}) - \boldsymbol{\alpha}_0^{\text{part}}\boldsymbol{W}(\boldsymbol{R}_{\text{image}})\,\boldsymbol{\Gamma}\right]\boldsymbol{\alpha}_0^{\text{main}} \\
&\equiv \left[\mathbf{1} + \boldsymbol{P}_0^{\text{part,out}} + \boldsymbol{P}_{0,\text{img}}^{\text{part,out}}\right]\boldsymbol{\alpha}_0^{\text{main}},
\end{aligned} \tag{83}$$

$$\begin{aligned}
\boldsymbol{\alpha}^{\text{main,in}} &= \boldsymbol{\alpha}_0^{\text{main}}\left[\mathbf{1} - \boldsymbol{W}(\boldsymbol{R})\,\boldsymbol{\alpha}_0^{\text{part}} - \boldsymbol{W}(\boldsymbol{R}_{\text{image}})\,\boldsymbol{\Gamma}\,\boldsymbol{\alpha}_0^{\text{part}}\right] \\
&\equiv \boldsymbol{\alpha}_0^{\text{main}}\left[\mathbf{1} + \boldsymbol{P}_0^{\text{part,in}} + \boldsymbol{P}_{0,\text{img}}^{\text{part,in}}\right],
\end{aligned} \tag{84}$$

where $\boldsymbol{R}_{\text{image}} = \boldsymbol{R} - D\hat{\boldsymbol{z}}$ is the vector linking the image to the partner. It is easy to see that all bosons linking main and partner in a full diagram (Fig. 1) may be replaced, in a

companion diagram, by an energy transfer through the image. As a consequence, the $\boldsymbol{\Lambda}$ matrices are modified in the following way:

$$\begin{aligned}
\boldsymbol{\Lambda}^{\text{main,in}} &= \Big[\mathbf{1} - (\boldsymbol{P}_0^{\text{part,in}} + \boldsymbol{P}_{0,\text{img}}^{\text{part,in}})(\boldsymbol{P}_0^{\text{main,in}} + \boldsymbol{P}_{0,\text{img}}^{\text{main,in}})\Big]^{-1} \Big[\mathbf{1} + \boldsymbol{P}_0^{\text{part,in}} + \boldsymbol{P}_{0,\text{img}}^{\text{part,in}}\Big], \qquad (85)\\
\boldsymbol{\Lambda}^{\text{main,out}} &= \big[\mathbf{1} + \boldsymbol{P}_0^{\text{part,out}} + \boldsymbol{P}_{0,\text{img}}^{\text{part,out}}\big] \Big[\mathbf{1} - (\boldsymbol{P}_0^{\text{part,in}} + \boldsymbol{P}_{0,\text{img}}^{\text{part,out}})(\boldsymbol{P}_0^{\text{main,out}} + \boldsymbol{P}_{0,\text{img}}^{\text{main,out}})\Big]^{-1},
\end{aligned}$$

for a generic bipartite diagram [cf. Eq. (20) and (21)], and:

$$\begin{aligned}
\boldsymbol{\Lambda}^{\text{lattice,out}} &= \left[1 + \boldsymbol{\alpha}_0 \Big(\sum_{\{s\}} \boldsymbol{W}^s(\boldsymbol{R}_s) + \boldsymbol{W}^{s,\text{image}}(\boldsymbol{R}_{s,\text{image}})\boldsymbol{\Gamma} \Big) \right]^{-1}\\
\boldsymbol{\Lambda}^{\text{lattice,in}} &= \left[1 + \Big(\sum_{\{s\}} \boldsymbol{W}^s(\boldsymbol{R}_s) + \boldsymbol{W}^{s,\text{image}}(\boldsymbol{R}_{s,\text{image}})\boldsymbol{\Gamma} \Big) \boldsymbol{\alpha}_0 \right]^{-1} \qquad (86)
\end{aligned}$$

for a lattice on the substrate [cf. Eq. (54) and (56)]. As seen above, we could also include the self-image of the main in Eq. (85) and (86). This contribution may induce an important shift on the vibrational frequencies in a molecular layer.[102] On the other hand, it has been suggested to treat the main and its image as one object, which polarizability is directly extracted from experimental data.[89] For a molecular lattice or a thin array of nanoparticles,[103] the calculation of infinite sums involving $\boldsymbol{W}$ coefficients is more complicated, in comparison with Part VI D, due to the z component of $\boldsymbol{R}_{s,\text{image}}$ (whereas $\boldsymbol{R}_s$ is in-plane). For square lattices, the summation may be performed as a function of the lattice parameters (i.e., lattice constant, distance D between the dipoles and their images, dielectric function ε_{sub}) in order to simulate the first-[86,104] and second-order[8,9,27] response functions, and to evaluate the frequency shifts[34,90] induced by molecular packing in such lattices. This is commonly restricted to the z-components in the literature, but our matrix formalism makes it possible to take all components into account, including those parallel to the surface.

G. Application to higher-order processes

Equation (51) gives the general procedure to implement the modifications of any bare response function of a system when coupled with a partner, whatever the order of the optical response. In the case of a four-wave mixing, which is a third-order process driven by the

second hyperpolarizability $\gamma(\imath\omega_1, \imath\omega_2, \imath\omega_3)$, Equation (51) leads to

$$\gamma_{ijkl}^{\text{main}} = \sum_{i'j'k'l'} \left[\Lambda^{\text{main,out}}(\omega_4)\right]_{ii'} \left[\gamma_0^{\text{main}}\right]_{i'j'k'l'} \left[\Lambda^{\text{main,in}}(\omega_1)\right]_{j'j} \left[\Lambda^{\text{main,in}}(\omega_2)\right]_{k'k} \left[\Lambda^{\text{main,in}}(\omega_3)\right]_{l'l}, \tag{87}$$

where $\omega_4 = \omega_1 + \omega_2 + \omega_3$ is the output frequency. This is the expected result for the IR-IR-visible process when modified by the dipole-dipole coupling inside a molecular monolayer on a substrate.[8] The $\mathbf{\Lambda}$ matrix may also be introduced to account for the dressing of other third-order response functions describing stimulated Raman scattering[21] or coherent anti-Stokes Raman scattering,[32] for example.

Among higher-order processes, we may also cite spontaneous Raman scattering. In practical applications, the experimental intensities of spontaneous Raman processes are efficiently modeled through an inelastic first-order polarizability $\boldsymbol{\alpha}^{\text{Raman}}(\omega_P, \omega_S)$ where ω_P and ω_S stand for the pump and Stokes frequencies, respectively. However, such a quantity, where the initial and final states differ, is not a response function *per se* and cannot be drawn as a loop in a Feynman-Matsubara diagram. It is still possible to account for its modifications in a bipartite system using the formalism developed here. After analytic continuity to real frequencies, we identify the modified Raman polarizability in the bipartite system with

$$\boldsymbol{\alpha}^{\text{Raman,main}}(\omega_P, \omega_S) = \mathbf{\Lambda}^{\text{main,out}}(\omega_S)\, \boldsymbol{\alpha}_0^{\text{Raman,main}}(\omega_P, \omega_S)\, \mathbf{\Lambda}^{\text{main,in}}(\omega_P). \tag{88}$$

Such a quantity efficiently describes the optical enhancement of the Raman process in Surface Enhanced Raman Spectroscopy (SERS) due to the influence of a partner nanostructure. In this particular case, one may simplify $\mathbf{\Lambda}$ as in Part VI E[105] or use more sophisticated models.[106] We note that the $\mathbf{\Lambda}$ factor appears twice here, once per involved photon, leading to the well-known intensity enhancement factor[107,108] $|f(\omega_P)f(\omega_S)|^2$ where f is a measure of the leading contribution to $\mathbf{\Lambda}$. It is interesting to compare this inelastic (Raman) polarizability to the elastic (Rayleigh) case in Part IV, for which we have seen that the $\mathbf{\Lambda}$ matrix is involved only once [Eq. (34)]. From a diagrammatic point of view, there is no paradox here as this is due to the counting of the different diagrams. By definition, plugging all $\boldsymbol{P}_n$ and $\boldsymbol{M}_n$ processes at the input (ω_P) and output (ω_S) sequences of the Raman polarizability does not generate identical diagrams. Conversely for the elastic polarizability, the input and output processes are identical and they precisely involve the $\boldsymbol{\alpha}_0(\omega)$ response function they modify.

It is easy to see that plugging arbitrary numbers of $\boldsymbol{P}_n$ processes on the input and output sides creates an infinite number of redundant diagrams, which must be discarded.

VII. CONCLUSION

In the same way that the many-body problem in solid-state physics is tackled by considering pair interactions among a large number of entities, bipartite systems may be considered in optics as the elementary units of higher-complexity composite systems. Studying the structural link between the bare response functions of two entities and the response function of their combination (through a given interaction Hamiltonian) is the first step towards the understanding of the linear and nonlinear optical responses of large composite systems. Even if such 2-body systems may look simple, their complexity lies in the number of vector bosons they exchange, i.e., the number of pair interaction processes. Pushing the diagrammatic theory of optics to its limits, we have demonstrated that, under soft hypotheses, any bipartite system described by a 2-loop diagram and involving an arbitrary number of vector bosons can be analyzed through loop splitting into elementary processes. Focusing on first- and second-order responses in the dipolar approximation, summing all the possible elementary processes has led us to define universal transfer matrices $\boldsymbol{\Lambda}$ which can be used to transform the bare response functions of the subsystems (when not coupled) into their dressed analogs (when coupled). This description in terms of $\boldsymbol{\Lambda}$ matrices is operative at any order, for example to quantify specific enhancements or resonance frequency shifts due to the interaction between a main entity and its partner. They may be plugged "as is" in most systems or limited to their first few terms for rapidly vanishing series. We have also shown that generalization to multipartite systems is possible in the case of an infinite lattice of partners and when the image dipoles are taken into account. This proves that the bipartite case is indeed a way to tackle large composite systems. We have illustrated the fact that this general theory, looking complex at first sight, is rather straightforwardly applied to standard systems like nanoparticle-nanoparticle, molecule-nanoparticle and molecule-molecule dimers, or infinite lattices, in order to produce an extended and more general version of the dressed response functions in these systems than usually provided.

The theory also produces a new kind of elementary circular processes which correct the dressed response functions by a scale factor. To our knowledge, these C_n processes have

never been described in the literature, although they share the same formal representation as the coupling between instantaneous dipoles, well-known as London forces. While the London process is independent of any excitation by light, the C_n processes only exist because the system interacts with light, hence their interpretation as light-driven circular energy exchanges. Their existence is postulated but hard to demonstrate because they cannot be easily separated from the $\mathbf{\Lambda}$ contributions in experimental data.

In a broader perspective, we believe that additional theoretical considerations, numerical computations and experimental results will challenge and improve the robustness of this general theory. On the experimental side, it is possible to benchmark its predictions with original and well-controlled systems by analyzing the modifications induced by the $\mathbf{\Lambda}$ matrices, and to look for smart ways to test the relevance and the meaning of the C_n terms as well as their potential competition with $\mathbf{\Lambda}$ contributions. On the theoretical side, we anticipate future developments and look forward to making this theory more operative and easily factored at arbitrary temperature, to adapting it to more complex light-matter interactions (at any order of a Mie theory, for example), and to implementing more subtle matter-matter interactions accounting for the inner structure of the system.

DATA AVAILABILITY

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

Appendix A: Rules for diagram splitting and factorization

Due to back and forth energy transfers between the main and the partner, the fermion propagators associated to each loop may convey identical frequencies.[30] Given that each loop is assigned a fermion Matsubara frequency ω_ν and that each propagator is assigned an imaginary-time Green's function

$$\mathfrak{G}_{nn}(\imath\omega_\nu + \imath\omega) = \frac{1}{\imath\omega_\nu - (\omega_n - \imath\omega)} \tag{A1}$$

related to the quantum state $|n\rangle$, any loop characterized by same-frequency propagators is described by a product of as many same-frequency Green's functions. In the simple case of

a loop involving two same-frequency propagators associated to states n and m, the response function encompasses the quantity $\sum_{n,m} \mathfrak{G}_{nn}(\imath\omega_\nu+\imath\omega)\mathfrak{G}_{mm}(\imath\omega_\nu+\imath\omega)$. In the limit of vanishing temperature,[30] this sum is dominated by the term for which $n = m$, hence characterized by a double pole $\omega_\nu = \omega_n - \imath\omega = \omega_m - \imath\omega$. Mathematically, the response function of the loop is then factorized into a product of two elementary response functions (i.e., whose orders are both smaller than that of the initial function) and, diagrammatically, the initial loop splits into two child loops with a lower number of vertices. By mathematical induction, this property extends to loops with an arbitrary number of same-frequency propagators, leading to as many child loops. Figure 7 illustrates this process of loop splitting. The response functions of odd order factorize into products of linear $\boldsymbol{\alpha}$ functions, while those of even order factorize into products of linear $\boldsymbol{\alpha}$ functions with one $\boldsymbol{\beta}$ function. For $n \geqslant 1$:

$$f^{(2n+1)}_{\{i_k\}_{1\leqslant k\leqslant 2n+2}}(\{\imath\omega_j, -\imath\omega_j\}_{j=1}^{n}, \imath\omega_{n+1}) = (-b)^n\, \alpha_{i_1 i_{2n+2}}(\imath\omega_{n+1}) \cdot \prod_{k=1}^{n} \alpha_{i_{2k+1} i_{2k}}(\imath\omega_k) \tag{A2}$$

and, for $n \geqslant 2$:

$$f^{(2n)}_{\{i_k\}_{1\leqslant k\leqslant 2n+1}}(\{\imath\omega_j, -\imath\omega_j\}_{j=1}^{n-1}, \imath\omega_n, \imath\omega_{n+1}) = (-b)^{n-1}\, \beta_{i_1 i_{2n} i_{2n+1}}(\imath\omega_n, \imath\omega_{n+1}) \cdot \prod_{k=1}^{n-1} \alpha_{i_{2k+1} i_{2k}}(\imath\omega_k) \tag{A3}$$

The splitting of a loop is then possible if, and only if, the associated response function encompasses an alternated sequence of frequencies $\{\imath\omega_j, -\imath\omega_j\}$ among its arguments, which is always the case under the hypotheses of Part II B.

Appendix B: Calculation of diagram in Fig. 4B

Following the Feynman rules, the diagram of Fig. 4B leads to:

$$\begin{aligned} U = -\frac{1}{b^3\hbar^4} \sum_{\gamma} \sum_{\lambda,\nu} \sum_{m,m',n,n'} \sum_{i,i',j,j'} & p^i_{nm} p^j_{mn} \mathfrak{G}_{nn}(\imath\omega_\nu) \mathfrak{G}_{mm}(\imath\omega_\nu + \imath\omega_\gamma) W_{jj'}(\imath\omega_\gamma) \\ & \times p^{i'}_{n'm'} p^{j'}_{m'n'} \mathfrak{G}_{m'm'}(\imath\omega_\lambda) \mathfrak{G}_{n'n'}(\imath\omega_\lambda + \imath\omega_\gamma) W_{i'i}(\imath\omega_\gamma) \end{aligned} \tag{B1}$$

where levels $\{m, m', n, n'\}$ are the states associated to the four propagators, ω_γ is a bosonic Matsubara frequency ($\imath\omega_\gamma = 2\imath\gamma\pi/\hbar b$ with $\gamma \in \mathbb{Z}$), whereas ω_λ and ω_ν are fermionic Matsub-

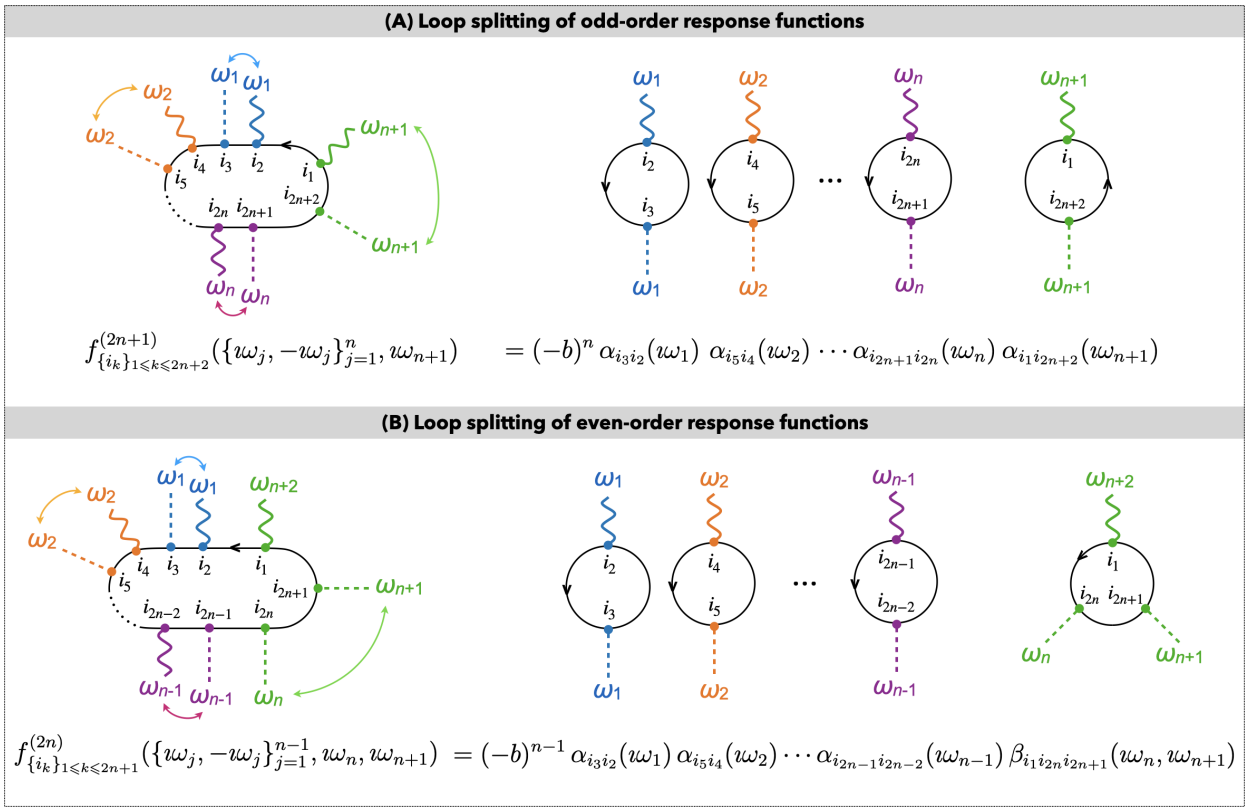

FIG. 7. Illustration of diagram splitting for (A) an odd-order response function and (B) an even-order response function. Contrary to the previous figures, curvy and dashed lines represent here for clarity the direction of the energy flux: curvy lines depict incoming photons/bosons, whereas dashed lines represent outgoing photons/bosons. The double arrows indicate that all the permutations of identical frequencies must be considered to recover the complete response functions of the child loops.

ara frequencies [e.g., $\imath\omega_\lambda = \imath(2\lambda+1)\pi/\hbar b$ with $\lambda\in\mathbb{Z}$], $\{i,i',j,j'\}$ are photon polarizations. When reorganizing the terms, we get

$$U = -\frac{1}{b}\sum_{i,i',j,j'}\sum_{\gamma}\left[\frac{-1}{b\hbar^2}\sum_{m,n}\sum_{\nu}p^{i}_{nm}p^{j}_{mn}\mathfrak{G}_{nn}(\imath\omega_\nu)\mathfrak{G}_{mm}(\imath\omega_\nu+\imath\omega_\gamma)\right]W_{jj'}(\imath\omega_\gamma)$$
$$\times\left[\frac{-1}{b\hbar^2}\sum_{m',n'}\sum_{\lambda}p^{i'}_{n'm'}p^{j'}_{m'n'}\mathfrak{G}_{m'm'}(\imath\omega_\lambda)\mathfrak{G}_{n'n'}(\imath\omega_\lambda+\imath\omega_\gamma)\right]W_{i'i}(\imath\omega_\gamma). \tag{B2}$$

In the square brackets, we recognize[29] the expression of the first-order polarizabilities $\boldsymbol{\alpha}$, leading to

$$U = -\frac{1}{b}\sum_{i,i',j,j'}\sum_{\gamma}\alpha^{\text{main}}_{ij}(\imath\omega_\gamma)W_{jj'}(\imath\omega_\gamma)\alpha^{\text{part}}_{j'i'}(\imath\omega_\gamma)W_{i'i}(\imath\omega_\gamma) \tag{B3}$$

or, equivalently,

$$U = -kT \sum_{\gamma} \mathrm{tr}\left[\boldsymbol{\alpha}^{\mathrm{main}}(\imath\omega_\gamma)\boldsymbol{W}(\imath\omega_\gamma)\boldsymbol{\alpha}^{\mathrm{part}}(\imath\omega_\gamma)\boldsymbol{W}(\imath\omega_\gamma)\right]. \tag{B4}$$

Recalling that $\imath\omega_\gamma = 2\imath\gamma\pi kT/\hbar$, we see that Eq. (B4) converges towards an integral for vanishing temperature as $2\pi kT/\hbar \sum_\gamma f(\imath\omega_\gamma) \to \int_{-\infty}^{\infty} f(\imath\omega)\mathrm{d}\omega$, leading to

$$U = -\frac{\hbar}{2\pi} \int_{-\infty}^{\infty} \mathrm{tr}\left[\boldsymbol{\alpha}^{\mathrm{main}}(\imath\omega)\boldsymbol{W}(\imath\omega)\boldsymbol{\alpha}^{\mathrm{part}}(\imath\omega)\boldsymbol{W}(\imath\omega)\right] d\omega. \tag{B5}$$

Appendix C: Derivation of trimer's equations

In Part V C 1, we consider the case of a trimer under the nearest neighbor approximation. Only two interaction matrices are involved, $\boldsymbol{W}^{12} = \boldsymbol{W}^{21}$ and $\boldsymbol{W}^{23} = \boldsymbol{W}^{32}$, so we define the four elementary processes $\boldsymbol{P}_0^{1,2} = -\boldsymbol{W}^{12}\boldsymbol{\alpha}_0^{(1)}$, $\boldsymbol{P}_0^{2,1} = -\boldsymbol{W}^{12}\boldsymbol{\alpha}_0^{(2)}$, $\boldsymbol{P}_0^{2,3} = -\boldsymbol{W}^{23}\boldsymbol{\alpha}_0^{(2)}$ and $\boldsymbol{P}_0^{3,2} = -\boldsymbol{W}^{23}\boldsymbol{\alpha}_0^{(3)}$ (and conversely for the equivalent output quantities). For each process $\boldsymbol{P}_0^{j,k}$, the photon interacts first with entity j, which interacts in turn with entity k. Hence, the back and forth processes $\boldsymbol{M}_0$ are defined accordingly: $\boldsymbol{M}_0^{1,2} = \boldsymbol{P}_0^{2,1}\boldsymbol{P}_0^{1,2}$, $\boldsymbol{M}_0^{2,1} = \boldsymbol{P}_0^{1,2}\boldsymbol{P}_0^{2,1}$, $\boldsymbol{M}_0^{2,3} = \boldsymbol{P}_0^{3,2}\boldsymbol{P}_0^{2,3}$ and $\boldsymbol{M}_0^{3,2} = \boldsymbol{P}_0^{2,3}\boldsymbol{P}_0^{3,2}$. The input part of the response function of entity k is modified by matrix $\boldsymbol{\Lambda}^{(k),\mathrm{in}}$, which must be deduced from the diagrams. As summarized in Part V A, the $\boldsymbol{\Lambda}^{(k),\mathrm{in}}$ matrix may be decomposed as a sum of $\mathbf{1}$, $\boldsymbol{P}_n^{(k)}$-type and $\boldsymbol{M}_n^{(k)}$-type processes.

We start with the central entity 2. The $\boldsymbol{P}_n^{(2)}$-type processes imply a photon interaction with entities 1 or 3 and a transfer to entity 2 (hence matrix $\boldsymbol{P}_0^{1,2} + \boldsymbol{P}_0^{3,2}$), followed by n back-and-forth energy exchanges with either entity 1 or 3 (i.e., driven by matrix $\boldsymbol{M}_0^{2,1}$ or $\boldsymbol{M}_0^{2,3}$). As for the $\boldsymbol{M}_n^{(2)}$-type processes, they involve a series of $(n+1)$ $\boldsymbol{M}_0^{2,1}$ and $\boldsymbol{M}_0^{2,3}$ processes, i.e., a product of the corresponding matrices, in arbitrary order. Considering that processes $\boldsymbol{M}_0^{2,1}$ and $\boldsymbol{M}_0^{2,3}$ do not commute, as illustrated in Fig. 5B and C, we deduce that the $\boldsymbol{\Lambda}^{(2),\mathrm{in}}$ matrix may be written as

$$\begin{aligned}\boldsymbol{\Lambda}^{(2),\mathrm{in}} &= \mathbf{1} + \sum_{n=0}^{\infty}\left[\boldsymbol{P}_n^{(2)} + \boldsymbol{M}_n^{(2)}\right] = \sum_{n=0}^{\infty}\left(\boldsymbol{M}_0^{2,1} + \boldsymbol{M}_0^{2,3}\right)^n \left(\mathbf{1} + \boldsymbol{P}_0^{1,2} + \boldsymbol{P}_0^{3,2}\right) \\ &= \left[1 - (\boldsymbol{M}_0^{2,1} + \boldsymbol{M}_0^{2,3})\right]^{-1} (\mathbf{1} + \boldsymbol{P}_0^{1,2} + \boldsymbol{P}_0^{3,2}).\end{aligned} \tag{C1}$$

Considering now entity 1, matrix $\mathbf{\Lambda}^{(1),\mathrm{in}}$ groups all the processes involving energy transfers towards entity 1 before it emits the outgoing photon. Besides the trivial process described by matrix $\mathbf{1}$, all other processes imply as their last step an energy transfer from entity 2 through a $\boldsymbol{P}_0^{2,1}$ process. In other words, we have

$$\mathbf{\Lambda}^{(1),\mathrm{in}} = \mathbf{1} + \boldsymbol{P}_0^{2,1}\mathbf{\Lambda}^{(2),\mathrm{in}} \tag{C2}$$

and

$$\mathbf{\Lambda}^{(3),\mathrm{in}} = \mathbf{1} + \boldsymbol{P}_0^{2,3}\mathbf{\Lambda}^{(2),\mathrm{in}}. \tag{C3}$$

To go further, we see that $\mathbf{\Lambda}^{(2),\mathrm{in}}$ may also be decomposed into a sum of three matrices, namely $\mathbf{1}$ for the trivial process, $\boldsymbol{P}_0^{1,2}\mathbf{\Lambda}^{(1),\mathrm{in}}$ for all processes involving an energy transfer from 1 to 2 (as their last step) and, symmetrically, $\boldsymbol{P}_0^{3,2}\mathbf{\Lambda}^{(3),\mathrm{in}}$ for the energy transfer from 3 to 2. This means that the $\mathbf{\Lambda}^{(k),\mathrm{in}}$ matrices are the solutions of the system given by Eq. (57).

Appendix D: Frequency shifts for heterodimers

Here, we aim to determine the values of the splitting frequency $\omega_{12}^{\mathrm{int}}$ and frequency shifts $\omega_{+/-}$ in the case of a heterodimer made of entities (labeled as "1" and "2") whose dipole moments are not aligned. We assume that their bare polarizabilities exhibit single resonances at ω_1 and ω_2, respectively:

$$(\alpha_0^{(1/2)})_{ij} = \frac{1}{\hbar}\frac{p_{1/2}^i\, p_{1/2}^j}{\omega_{1/2} - \imath\omega}. \tag{D1}$$

In Eq. (69), the (ij) component of the n^{th} term of the infinite sum thus reads

$$p_1^i \sum_{k,l,m,q\ldots z} \left(\frac{1}{\hbar}\right)^{2n+1} \frac{p_1^k W^{kl} p_2^l}{\omega_1 - \imath\omega}\,\frac{p_2^m W^{mq} p_1^q}{\omega_2 - \imath\omega}\,\frac{p_1^r W^{rs} p_2^s}{\omega_1 - \imath\omega}\,(\cdots)\,\frac{p_2^y W^{yz} p_1^z}{\omega_1 - \imath\omega}\,p_1^j. \tag{D2}$$

Quantity $\sum_{k,l} p_1^k W^{kl} p_2^l = \sum_{m,q} p_2^m W^{mq} p_1^q = \boldsymbol{p}_1 \cdot \boldsymbol{W}\boldsymbol{p}_2 = \hbar\omega_{12}^{\mathrm{int}}$ is the dipolar coupling constant between the states excited by $\boldsymbol{p}_1$ and $\boldsymbol{p}_2$. The generic term of the power series becomes

$$\frac{1}{\hbar}\frac{p_1^i\, p_1^j}{\omega_1 - \imath\omega}\frac{(\omega_{12}^{\mathrm{int}})^{2n}}{(\omega_1 - \imath\omega)^n(\omega_2 - \imath\omega)^n} = (\alpha_0^{(1)})_{ij}\frac{(\omega_{12}^{\mathrm{int}})^{2n}}{(\omega_1 - \imath\omega)^n(\omega_2 - \imath\omega)^n}, \tag{D3}$$

so that

$$\begin{aligned}\boldsymbol{\alpha}^{(1),\mathrm{out}} &= \left[\mathbf{1} + \boldsymbol{P}_0^{(2),\mathrm{out}}\right] \boldsymbol{\alpha}_0^{(1)} \sum_{n=0}^{\infty} \left[\frac{(\omega_{12}^{\mathrm{int}})^2}{(\omega_1 - \imath\omega)(\omega_2 - \imath\omega)}\right]^n \\ &= \left[\mathbf{1} + \boldsymbol{P}_0^{(2),\mathrm{out}}\right] \boldsymbol{\alpha}_0^{(1)} \frac{(\omega_1 - \imath\omega)(\omega_2 - \imath\omega)}{(\omega_1 - \imath\omega)(\omega_2 - \imath\omega) - (\omega_{12}^{\mathrm{int}})^2}. \end{aligned} \tag{D4}$$

As $\hbar\omega_{12}^{\mathrm{int}} = \boldsymbol{p}_1 \cdot \boldsymbol{W}\boldsymbol{p}_2$ (and given the definitions of the angles in Part VI B 3), we have:

$$\begin{aligned}\hbar\omega_{12}^{\mathrm{int}} &= \frac{1}{4\pi\varepsilon_0 D^3} \left[\boldsymbol{p}_1 \cdot \boldsymbol{p}_2 - 3(\boldsymbol{p}_1 \cdot \hat{\boldsymbol{x}})(\boldsymbol{p}_2 \cdot \hat{\boldsymbol{x}})\right] \\ &= \frac{|\boldsymbol{p}_1|\,|\boldsymbol{p}_2|}{4\pi\varepsilon_0 D^3} \left(\cos\theta_{12} - 3\cos\theta_1 \cos\theta_2\right), \end{aligned} \tag{D5}$$

which amounts to introducing an orientation factor in the Davydov splitting.[70,73,77] The two poles of Eq. (D4) are found by solving the second-order equation cancelling the denominator. For identical molecules ($\omega_1 = \omega_2 = \omega_0$), we get $\omega_\pm = \omega_0 \pm \omega_{12}^{\mathrm{int}}$. Both modes are optically active, except when the molecules adopt an identical geometry: in this case, ω_- is also a root of the numerator and only ω_+ is optically active (as discussed for homodimers). For a heterodimer, we obtain Eq. (71).

REFERENCES

[1]B. Busson, J. Chem. Phys. **159**, 034707 (2023).

[2]J. Hunt, P. Guyot-Sionnest, and Y. Shen, Chem. Phys. Lett. **133**, 189 (1987).

[3]R. Superfine, P. Guyot-Sionnest, J. Hunt, C. Kao, and Y. Shen, Surf. Sci. **200**, L445 (1988).

[4]B. Busson and A. Tadjeddine, J. Phys. Chem. C **113**, 21895 (2009).

[5]C. Humbert, O. Pluchery, E. Lacaze, A. Tadjeddine, and B. Busson, Phys. Chem. Chem. Phys. **14**, 280 (2012).

[6]H. Held, A. I. Lvovsky, X. Wei, and Y. R. Shen, Phys. Rev. B **66**, 205110 (2002).

[7]V. Ostroverkhov, G. A. Waychunas, and Y. R. Shen, Phys. Rev. Lett. **94**, 046102 (2005).

[8]M. Cho, C. Hess, and M. Bonn, Phys. Rev. B **65**, 205423 (2002).

[9]F. Vidal, B. Busson, A. Tadjeddine, and A. Peremans, J. Chem. Phys. **119**, 12492 (2003).

[10]A. N. Bordenyuk, C. Weeraman, A. Yatawara, H. D. Jayathilake, I. Stiopkin, Y. Liu, and A. V. Benderskii, J. Phys. Chem. C **111**, 8925 (2007).

[11]A. Lagutchev, S. A. Hambir, and D. D. Dlott, J. Phys. Chem. C **111**, 13645 (2007).

[12]I. V. Stiopkin, H. D. Jayathilake, A. N. Bordenyuk, and A. V. Benderskii, J. Am. Chem. Soc. **130**, 2271 (2008).

[13]Y. Li, B. Xiang, and W. Xiong, J. Chem. Phys. **150**, 114706 (2019).

[14]A. Le Rille and A. Tadjeddine, J. Electroanal. Chem. **467**, 238 (1999).

[15]L. Dalstein, C. Humbert, M. Ben Haddada, S. Boujday, G. Barbillon, and B. Busson, J. Phys. Chem. Lett. **10**, 7706 (2019).

[16]L. Fu, S.-L. Chen, and H.-F. Wang, J. Phys. Chem. B **120**, 1579 (2016).

[17]T. Noblet, L. Dreesen, S. Boujday, C. Méthivier, B. Busson, A. Tadjeddine, and C. Humbert, Commun. Chem. **1**, 76 (2018).

[18]Y. R. Shen, Appl. Phys. A **59**, 541 (1994).

[19]J.-L. Oudar and Y. R. Shen, Phys. Rev. A **22**, 1141 (1980).

[20]P. Ye and Y. R. Shen, Phys. Rev. A **25**, 2183 (1982).

[21]N. Bloembergen and Y. R. Shen, Phys. Rev. Lett. **12**, 504 (1964).

[22]Y. R. Shen and N. Bloembergen, Phys. Rev. **137**, A1787 (1965).

[23]Y. R. Shen, Phys. Rev. A **45**, 446 (1992).

[24]P. Guyot-Sionnest and Y. R. Shen, Phys. Rev. B **38**, 7985 (1988).

[25]P. Guyot-Sionnest, W. Chen, and Y. R. Shen, Phys. Rev. B **33**, 8254 (1986).

[26]P. Guyot-Sionnest and Y. R. Shen, Phys. Rev. B **35**, 4420 (1987).

[27]P. Ye and Y. R. Shen, Phys. Rev. B **28**, 4288 (1983).

[28]C. K. Chen, A. R. B. de Castro, and Y. R. Shen, Phys. Rev. Lett. **46**, 145 (1981).

[29]T. Noblet, B. Busson, and C. Humbert, Phys. Rev. A **104**, 063504 (2021).

[30]T. Noblet and B. Busson, Phys. Rev. B **105**, 205420 (2022).

[31]T. Noblet and B. Busson, J. Chem. Phys. **160**, 024704 (2024).

[32]Y. R. Shen, *The Principles of Nonlinear Optics* (Wiley, New York, USA, 1984).

[33]R. W. Boyd, *Nonlinear optics* (Academic Press, San Diego, CA, USA, 2003).

[34]B. N. J. Persson and R. Ryberg, Phys. Rev. B **24**, 6954 (1981).

[35]J. G. Kirkwood, J. Chem. Phys. **5**, 479 (1937).

[36]G. Gonella, C. Lütgebaucks, A. G. F. de Beer, and S. Roke, J. Phys. Chem. C **120**, 9165 (2016).

[37]J. Choi, J. Lee, M. Makarem, S. Huang, and S. H. Kim, J. Phys. Chem. B **126**, 6629 (2022).

[38]D. L. Andrews and B. S. Sherborne, J. Chem. Phys. **86**, 4011 (1987).

[39]D. Craig and T. Thirunamachandran, Chem. Phys. **135**, 37 (1989).

[40]A. D. McLachlan and H. C. Longuet-Higgins, Proc. Roy. Soc. London A **271**, 387 (1963).

[41]A. D. McLachlan, Mol. Phys. **6**, 423 (1963).

[42]J. N. Israelachvili, in *Intermolecular and Surface Forces (Third Edition)*, edited by J. N. Israelachvili (Academic Press, San Diego, 2011) third edition ed., pp. 107–132.

[43]A. D. McLachlan and H. C. Longuet-Higgins, Proc. Roy. Soc. London A **274**, 80 (1963).

[44]I. Dzyaloshinskii, E. Lifshitz, and L. Pitaevskii, Adv. Phys. **10**, 165 (1961).

[45]H. B. G. Casimir and D. Polder, Phys. Rev. **73**, 360 (1948).

[46]A. D. McLachlan, Mol. Phys. **7**, 381 (1964).

[47]A. D. McLachlan, Discuss. Faraday Soc. **40**, 239 (1965).

[48]F. London, Trans. Faraday Soc. **33**, 8 (1937).

[49]K. Sasihithlu and G. D. Scholes, J. Phys. Chem. B **128**, 1205 (2024).

[50]G. D. Mahan, *Many-particle physics*, 2nd ed. (Plenum Press, 1990).

[51]U. Kreibig and M. Vollmer, *Optical Properties of Metal Clusters* (Springer, Berlin, Germany, 1995).

[52]J. I. Gersten and A. Nitzan, Phys. Rev. B **29**, 3852 (1984).

[53]M. Quinten, *Optical Properties of Nanoparticle Systems* (Wiley, Weinheim, Germany, 2011).

[54]P. Bobbert and J. Vlieger, Physica A **147**, 115 (1987).

[55]A. S. Davydov, *Theory of Molecular Excitons* (McGraw-Hill, New York, NY, USA, 1962).

[56]D. Marinica, A. Kazansky, P. Nordlander, J. Aizpurua, and A. G. Borisov, Nano Lett. **12**, 1333 (2012).

[57]M. Chapman, M. Mullen, E. Novoa-Ortega, M. Alhasani, J. F. Elman, and W. B. Euler, J. Phys. Chem. C **120**, 8289 (2016).

[58]B. Busson, M. Farhat, P.-J. Nini-Teunda, S. Roy, T. Jarisz, and D. K. Hore, J. Chem. Phys. **154**, 224704 (2021).

[59]B. Busson, J. Chem. Phys. **156**, 204704 (2022).

[60]X. D. Zhu, H. Suhr, and Y. R. Shen, Phys. Rev. B **35**, 3047 (1987).

[61]A. D. Curtis, S. B. Reynolds, A. R. Calchera, and J. E. Patterson, J. Phys. Chem. Lett. **1**, 2435 (2010).

[62]P. A. Covert and D. K. Hore, J. Phys. Chem. C **119**, 271 (2015).

[63]M. Meier and A. Wokaun, Opt. Lett. **8**, 581 (1983).

[64]D. W. Mackowski, Proc. R. Soc. A Math. Phys. Eng. Sci. **433**, 599 (1991).

[65]N. S. Mueller, S. Heeg, and S. Reich, Phys. Rev. A **94**, 023813 (2016).

[66]A. Pinchuk and G. Schatz, Nanotechnology **16**, 2209 (2005).

[67]V. V. Gozhenko, L. G. Grechko, and K. W. Whites, Phys. Rev. B **68**, 125422 (2003).

[68]V. L. Brudny, B. S. Mendoza, and W. Luis Mochán, Phys. Rev. B **62**, 11152 (2000).

[69]J. I. Dadap, J. Shan, and T. F. Heinz, J. Opt. Soc. Am. B **21**, 1328 (2004).

[70]C. Tabor, D. Van Haute, and M. A. El-Sayed, ACS Nano **3**, 3670 (2009).

[71]W. T. Simpson and D. L. Peterson, J. Chem. Phys. **26**, 588 (1957).

[72]K. Joseph, B. de Waal, S. A. H. Jansen, J. J. B. van der Tol, G. Vantomme, and E. W. Meijer, J. Am. Chem. Soc. **146**, 12130 (2024).

[73]M. Kasha, H. R. Rawls, and M. A. El-Bayoumi, Pure Appl. Chem. **11**, 371 (1965).

[74]P. K. Jain, S. Eustis, and M. A. El-Sayed, J. Phys. Chem. B **110**, 18243 (2006).

[75]T. Kikteva, D. Star, Z. Zhao, T. L. Baisley, and G. W. Leach, J. Phys. Chem. B **103**, 1124 (1999).

[76]N. J. Hestand and F. C. Spano, Chem. Rev. **118**, 7069 (2018).

[77]B. Z. Packard, D. D. Toptygin, A. Komoriya, and L. Brand, J. Phys. Chem. B **102**, 752 (1998).

[78]S. Gomrok, B. K. Eldridge, E. A. Chaffin, J. W. Barr, X. Huang, T. B. Hoang, and Y. Wang, J. Chem. Phys. **160**, 144706 (2024).

[79]R. Holzinger, N. S. Bassler, H. Ritsch, and C. Genes, J. Phys. Chem. A **0**, null (2024).

[80]F. J. Garcia-Vidal, C. Ciuti, and T. W. Ebbesen, Science **373**, eabd0336 (2021).

[81]J. Bellessa, C. Bonnand, J. C. Plenet, and J. Mugnier, Phys. Rev. Lett. **93**, 036404 (2004).

[82]K. Hubenko, A. Kusber, M. Naumann, B. Büchner, and M. Knupfer, J. Chem. Phys. **160**, 144708 (2024).

[83]S. Noschese, L. Pasquini, and L. Reichel, Numer. Linear Algebra Appl. **20**, 302 (2013).

[84]S. De Boer and D. A. Wiersma, Chem. Phys. Lett. **165**, 45 (1990).

[85]E. McRae and M. Kasha, in *Physical Processes in Radiation Biology*, edited by L. Augenstein, R. Mason, and B. Rosenberg (Academic Press, 1964) pp. 23–42.

[86]A. Bagchi, R. G. Barrera, and R. Fuchs, Phys. Rev. B **25**, 7086 (1982).

[87]J. Topping and S. Chapman, Proc. Royal Soc. London. A **114**, 67 (1927).

[88]B. L. Maschhoff and J. P. Cowin, J. Chem. Phys. **101**, 8138 (1994).

[89]B. Persson and A. Liebsch, Surf. Sci. **110**, 356 (1981).

[90]G. D. Mahan and A. A. Lucas, J. Chem. Phys. **68**, 1344 (2008).

[91]M. W. Severson, C. Stuhlmann, I. Villegas, and M. J. Weaver, J. Chem. Phys. **103**, 9832 (1995).

[92]C. F. Bohren and D. R. Huffman, *Absorption and scattering of light by small particles* (Wiley, 1983).

[93]B. Busson and L. Dalstein, J. Phys. Chem. C **123**, 26597 (2019).

[94]J. D. Jackson, ed., *Classical Electrodynamics*, 3rd ed. (Wiley, 1999).

[95]J. D. E. McIntyre and D. E. Aspnes, Surf. Sci. **24**, 417 (1971).

[96]T. F. Heinz, in *Nonlinear Surf. Electromagn. Phenom.*, edited by H. E. Ponath and G. I. Stegeman (Elsevier, Amsterdam, 1991) Chap. 5, pp. 353–416.

[97]B. Busson, J. Chem. Phys. **159**, 034705 (2023).

[98]R. Fuchs and R. G. Barrera, Phys. Rev. B **24**, 2940 (1981).

[99]M. Wind, J. Vlieger, and D. Bedeaux, Physica A **141**, 33 (1987).

[100]J. F. Li, Y. F. Huang, Y. Ding, Z. L. Yang, S. B. Li, X. S. Zhou, F. R. Fan, W. Zhang, Z. Y. Zhou, D. Y. Wu, B. Ren, Z. L. Wang, and Z. Q. Tian, Nature **464**, 392 (2010).

[101]Y. He, H. Ren, E.-M. You, P. M. Radjenovic, S.-G. Sun, Z.-Q. Tian, J.-F. Li, and Z. Wang, Phys. Rev. Lett. **125**, 047401 (2020).

[102]M. Scheffler, Surf. Sci. **81**, 562 (1979).

[103]V. V. Truong and G. D. Scott, J. Opt. Soc. Am. **67**, 502 (1977).

[104]R. G. Barrera, M. del Castillo-Mussot, G. Monsivais, P. Villaseor, and W. L. Mochán, Phys. Rev. B **43**, 13819 (1991).

[105]J. Creighton, Surf. Sci. **124**, 209 (1983).

[106]C. A. Murray and S. Bodoff, Phys. Rev. Lett. **52**, 2273 (1984).

[107]A. Wokaun, J. P. Gordon, and P. F. Liao, Phys. Rev. Lett. **48**, 957 (1982).

[108]E. J. Zeman and G. C. Schatz, J. Phys. Chem. **91**, 634 (1987).