



HAL
open science

Locality and computational reliability of linear response calculations for molecular systems

Marco d'Alessandro, Luigi Genovese

► **To cite this version:**

Marco d'Alessandro, Luigi Genovese. Locality and computational reliability of linear response calculations for molecular systems. *Physical Review Materials*, 2019, 3 (2), 10.1103/PhysRevMaterials.3.023805 . hal-02138937

HAL Id: hal-02138937

<https://hal.science/hal-02138937v1>

Submitted on 22 Jul 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Locality and Computational Reliability of Linear Response Calculations for Molecular Systems

Marco D'Alessandro

*Istituto di Struttura della Materia-CNR (ISM-CNR),
Via del Fosso del Cavaliere 100, 00133 Roma, Italia*

Luigi Genovese

Laboratoire de Simulation Atomistique (LSim), Univ. Grenoble Alpes, CEA, INAC-MEM, 38000 Grenoble, France

(Dated: February 26, 2019)

By performing a critical analysis of the fundamental equations of linear-response (LR) formalism in molecules, we explore the interplay between locality of the response density operator and numerical convergence of LR-related quantities. We show that for frequencies below the first ionization potential (IP) of the system, it is possible to express the response density by employing localized states only. Above this threshold energy, such a locality property cannot be achieved. Such considerations may be transposed in terms of the molecule's excited states. We show that not all the system's excitations can be considered on equal footing. There is a discrete sector of excitations – which may also extend above IP – that can be parametrized by observable, localized states, which can be computationally expressed with high precision, provided an adequate level of completeness. We present indicators that can help to quantify such potential observable properties of an excitation, that can be evaluated in any discretization scheme. The remaining excitation modes belong to a continuum spectrum that, on the contrary, is not directly associated to observable properties and can only be effectively represented in a given computational setup. Such considerations are important not only for reproducibility of the results among different computer codes employing diverse formalisms, but also in view of providing a deeper understanding on the impact of models' approximations on the scientific outcomes of the simulation.

I. INTRODUCTION

For a theoretical approach that is based on numerical calculations, for which no analytic reference solution exists, reducing the computational uncertainty is the only possible way to shed light on the predictive power of the model. In other terms, the “accuracy” of a result with respect to experimental data may be reliably quantified only when the computational uncertainty is guaranteed to be significantly lower than the observed discrepancy. For communities employing density functional theory (DFT), the need of such a “calibration” of computer codes has led to the DELTACODES project¹. This project represents a remarkable example of precision-driven initiative where a grassroots community of code-owners struggled to extract well-defined computational results in the context of DFT calculations. Ground-state quantum mechanical quantities as lattice constants and bulk modulus of elementary crystals extracted from different codes were compared with each other, with the objective of reducing the uncertainties on such data, and in some sense, to define a *calibration* procedure for a solid-state Physics DFT code. Any other code able to extract the same quantities with the same approximations may be compared with the results of the community in order to assess its computational reliability.

The problem is even more stringent for calculations that refer to quantities *beyond* ground state, like for instance the time-dependent density functional theory (TDDFT)²⁻⁴, from which quantities like optical excitations or polarizability tensors of a given system may be

studied and put in relation with experimental data. Here also, a plethora of computational approaches exist, that employ different numerical formalism and basis sets, and it will be of great importance to identify quantities that can be “measured” in a given model, to assess and quantify their computational uncertainty.

In this paper, we focus our attention on linear-response (LR) treatments of *molecular* systems, in particular referring to LR-TDDFT calculation. We will investigate the impact that *open* (i.e., isolated) boundary conditions (BC), which are imposed to the system, will have on the computational evaluation of LR quantities. We perform analytic analysis on the equations beyond LR treatment and discuss the interplay between reproducible LR quantities and the *locality* of the related objects. Such considerations will enable us to distinguish quantities which *a priori* can be compared among different treatments - and therefore tested for reproducibility - from others which would *explicitly* depend of the employed computational treatment.

It is indeed relatively easy to have converged indicators for numerical ground-state (GS) quantities: on one side, the variational theorem guarantees that the lower the GS energy, the more precise the numerical result; on the other side, GS quantities of molecular system can be expressed only in terms of localized (i.e. bound) states. These considerations make it relatively easy to increase the basis set completeness in view of the extraction of reference results.

We will see in what follows that the situation is not as simple in the context of LR. Indeed the response density

– even for molecular systems – loses its localized behavior for high frequencies, independently on the nature of the specific system and perturbation. Furthermore, we will demonstrate that to increase the basis set completeness, delocalized oscillatory degrees of freedom cannot be excluded *a priori*, even in the localized regime.

Lastly, we will show that the reliability of the convergence criteria that can be established for LR, where no variational theorem exists – depends explicitly on the target frequency range and on the features of the numerical basis set.

The paper is organized as follows. Initially, we will inspect how the question of reproducibility may be expressed for ground state calculations of molecules. Following these guidelines, we will then move to LR equations, by first considering the equations of motion of the response density operator, given the application of a perturbing field. Then, we will consider the behavior of the “free oscillations”, i.e. the excited states of the molecule. Our considerations will prove to be useful for the understanding of the *analytic structure* of LR quantities like the linear susceptibility of the system (i.e. the reducible polarizability). We support our considerations with numerical results on simple LR-TDDFT quantities. For readability purposes, the details of each calculation are explained in Appendix C.

Although we refer, as anticipated, to TDDFT models, such considerations might also be transposed to many-body perturbation theory approaches, in the context of quasiparticle equations.

A. Long range behavior of the eigenvalue solution of the Schrödinger equation

Let us start our discussion by reviewing some well-known concepts of quantum mechanics under a perspective that will turn out to be very useful in the forthcoming discussions about linear-response quantities. In particular we revisit the impact that the isolated BC have on the solutions of the Schrödinger equation.

Consider a (one-body) wave function $|\psi\rangle$ solution of the equation $\hat{H}|\psi\rangle = \epsilon|\psi\rangle$, where we split the Hamiltonian operator into the sum of a kinetic \hat{T} and a potential $\hat{V} \equiv \hat{H} - \hat{T}$ term. In view of discussing the impact of the BC, it is instructive to transform the Schrödinger equation in a scattering problem, by writing

$$|\psi\rangle = \hat{G}_{\hat{T}}(\epsilon) |\hat{V}\psi\rangle, \quad (1)$$

where we have introduced the Green function of the Helmholtz operator $\hat{G}_{\hat{T}}(\epsilon) = (\epsilon - \hat{T})^{-1}$, resolvent of the free-particle Hamiltonian $H_s \equiv \hat{T}$.

When describing the ground state of a molecule, we suppose that the potential operator, even if non-local, can be effectively restricted to wave functions that can be considered as nonzero only on a bounded domain. In other terms, a molecule can be associated to a “scatterer”

that has a *finite* spatial extension; if a state $|\psi\rangle$ is localized, it will be the same for the ket $|\hat{V}\psi\rangle$, though perhaps in a larger region. For a computational analysis of the problem, this property is fundamental to the choice of the numerical treatment as we know that computational basis sets which are tailored to express asymptotically vanishing wave functions are better suited for the solution of the problem.

This fact is made apparent by the expression of the kernel of $\hat{G}_{\hat{T}}(\epsilon)$ in the coordinate representation. When dealing with Isolated BC, it is easy to see that the solution of Eq. (1) can be expressed by:

$$\langle \mathbf{r} | \hat{G}_{\hat{T}}(\epsilon) | \mathbf{r}' \rangle = \frac{1}{4\pi} \begin{cases} \frac{e^{-\alpha|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|}, & \text{for } \epsilon < 0 \\ \frac{e^{i\alpha|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|}, & \text{for } \epsilon \geq 0 \end{cases}, \quad (2)$$

where $\alpha = \sqrt{2|\epsilon|}$.

From the above expressions we can see that, for molecular systems, the value $\epsilon = 0$ represents a *threshold*⁵ energy that separates two different classes of solutions. Negative-energy solutions are *localized*, i.e. they exhibit bound-state behavior far away from the scatterer. These are bound states and, by means of their normalizability, they are associated to *discrete*, well-defined energies. The bound states can be of course labeled by quantum numbers which are defined by the properties of the scatterer, namely the symmetries of the potential \hat{V} . In a computational discretization of the Schrödinger equation, such states can in principle be expressed with arbitrary precision, provided that the numerical treatment offers an adequate level of completeness. Being discrete and well-identified, the bound-state energies are a property of the molecule and can be associated to *observable* quantities. Their energies can be compared among different treatments and it is in principle possible to provide reference values.

On the other hand, positive energy solutions behave differently. Even though the potential is localized the associated Helmholtz kernel is a spherical wave and its convolution with the source term provides a delocalized wavefunction $\psi(\mathbf{r})$ in the whole space. As they cannot be normalized they belong to the essential spectrum of the Hamiltonian operator and form a *continuum* of states. Outside of the support of the $|\hat{V}\psi\rangle$ the solutions of Eq. (1) can be put in bijection with the eigenstates of the free-particle Hamiltonian \hat{H}_s . For positive energies, the quantum numbers of the eigenstates are determined by the Laplacian operator, regardless of the particular features of the potential.

It is easy to see from these concepts that a computational treatment that is tailored for the discretization of bound states will not be as effective for continuum states, due to the different long-range asymptotic behavior - which in turn is a clear consequence of the BC of the problem. In a numerical treatment, where we discretize the problem with a *finite* number of degrees of freedom, we will only have access to a *pseudo-continuum* of states,

whose energy values and density of states will depend on the representation of the kinetic operator in the basis set. As a consequence, the set of eigenstates with $\epsilon > 0$ expressed in a given computational setup depends on the features of the employed basis set.

Direct numerical evidence of this problem has been discussed in⁶, where the authors analyze the effects of a finite basis set on the occupied and virtual orbitals of a molecular system and establish a clear correspondence between the orbital localization character, and the independence of its energy with respect to the basis set. Furthermore, it is clearly shown that the energies of the unbound orbitals are not stable with respect to parameters like the size of the computational domain.

II. FLUCTUATION STATES. A LOCALIZATION ARGUMENT FOR THE RESPONSE DENSITY

It is evident that the arguments presented above have a limited interest for the study of ground-state quantities of molecules. By definition, all these quantities can be expressed as a functional which *entirely* depends on bound states. As an illustration it is sufficient to recall the Hohenberg-Kohn theorem that states that the ground-state (GS) energy is a functional of the GS charge density $\hat{\rho}_0$. For molecular systems which are electronically stable (i.e. no anionic or metastable configurations), such a quantity is localized, in the sense defined in the above section. Stated otherwise, *there is no need to express efficiently any portion of the continuum spectrum* to have a reliable GS treatment of molecules.

Let us now inspect the case of LR calculations. In this framework, we consider the unperturbed Hamiltonian \hat{H}_0 and assumes that the GS density $\hat{\rho}_0$ is accessible and expressed in a given computational treatment. The linear response formalism allows us to evaluate the modification of the expectation value of a generic observable induced by a time- (or frequency-) dependent perturbing field $\hat{\Phi}(\omega)$ acting on the system. This quantity, written in the frequency domain, can be expressed through the evaluation of the *linear response functional*

$$\langle \delta \hat{O} \rangle_{\Phi}(\omega) = \text{tr} \left(\hat{\rho}'_{\Phi}(\omega) \hat{O} \right), \quad (3)$$

where \hat{O} represents the observable⁷ under inspection and we have introduced the *response density operator* $\hat{\rho}'_{\Phi}(\omega)$ that codifies the modification of $\hat{\rho}_0$ induced by the perturbation $\hat{\Phi}(\omega)$.

The response density operator is expressed as the first-order variation of $\hat{\rho}_0 = \sum_{\{p\}} |\psi_p\rangle \langle \psi_p|$, where the set $\{|\psi_p\rangle\}$ denote the occupied states, *i.e.* the subset of the bound eigenstates⁸ of \hat{H}_0 with (negative) energy ϵ_p lower than the Fermi level. Using this notation we denote the energy of the HOMO level as ϵ_h and identify the value of the first ionization potential with $\text{IP} \equiv |\epsilon_h|$.

The response density satisfies an equation of motion written in the form of a quantum Liouville operator (for

example see⁹)

$$(\omega - \hat{\mathcal{L}}) \hat{\rho}'_{\Phi}(\omega) = [\hat{\Phi}(\omega), \hat{\rho}_0]. \quad (4)$$

The Liouvillian superoperator $\hat{\mathcal{L}}$ for self-consistent systems is expressed as the sum of the unperturbed part $\hat{\mathcal{L}}_0$ plus a coupling term $\hat{\mathcal{K}}$. The action of these terms reads, respectively

$$\hat{\mathcal{L}} \hat{O} \equiv (\hat{\mathcal{L}}_0 + \hat{\mathcal{K}}) \hat{O} = [\hat{H}_0, \hat{O}] + [\hat{V}'[\hat{O}], \hat{\rho}_0], \quad (5)$$

where $\hat{V}'[\hat{O}] \equiv \int d\mathbf{r} d\mathbf{r}' \frac{\delta \hat{V}[\hat{\rho}_0]}{\delta \rho(\mathbf{r}, \mathbf{r}')} O(\mathbf{r}, \mathbf{r}')$ encodes the response of the $\hat{\rho}$ -dependent potential to a modification of the density operator. The inspection of the linear order time evolution of the eigenstates of \hat{H}_0 shows that the response density acts as a transition operator, linking the occupied and empty subspaces of \hat{H}_0 . An operator \hat{O}_{\perp} with this feature satisfies the *transverse* condition

$$\hat{O}_{\perp} \equiv \hat{\rho}_0 \hat{O} \hat{Q}_0 + \hat{Q}_0 \hat{O} \hat{\rho}_0, \quad (6)$$

where $\hat{Q}_0 = \hat{\mathbb{I}} - \hat{\rho}_0$ is the projector to the empty subspace of \hat{H}_0 (for more details see Appendix A). In other terms, the condition $\hat{\rho}'_{\Phi} = \hat{\rho}'_{\Phi \perp}$ holds.

We can therefore introduce an explicit representation of the response density parametrized as

$$\hat{\rho}'_{\Phi}(\omega) = \sum_p (|\psi_p\rangle \langle f_p^{\Phi}(-\omega)| + |f_p^{\Phi}(\omega)\rangle \langle \psi_p|). \quad (7)$$

Here we have introduced the set of ω -dependent *fluctuation states* (FS) $|f_p^{\Phi}(\omega)\rangle = \hat{\rho}'_{\Phi}(\omega) |\psi_p\rangle$, that have values in the unoccupied subspace of \hat{H}_0 . However, non-Hermiticity of $\hat{\rho}'_{\Phi}(\omega)$ implies that the fluctuation states are complex quantities.

An analysis of the Liouville equation (4) for the response density written in this fashion evidences that the equations of motion of fluctuation states are written as a modified Sternheimer equation¹⁰⁻¹²

$$[\omega - (\hat{H}_0 - \epsilon_p)] |f_p^{\Phi}(\omega)\rangle = \hat{Q}_0 (\hat{\Phi}(\omega) + \hat{V}'[\hat{\rho}'_{\Phi}(\omega)]) |\psi_p\rangle, \quad (8)$$

and we observe that the first-order Hamiltonian contains, apart from the perturbing field, a further term $\hat{V}'[\hat{\rho}'_{\Phi}]$ due to the density-dependence of \hat{H}_0 . It is easy to see that the solution of Eq. (8) can be recast in a scattering problem:

$$|f_p^{\Phi}(\omega)\rangle = \hat{G}_{\hat{T}}(\omega + \epsilon_p) \left(\hat{V} |f_p^{\Phi}(\omega)\rangle + |s_p^{\Phi}(\omega)\rangle \right), \quad (9)$$

with the state $|s_p^{\Phi}(\omega)\rangle = \hat{Q}_0 (\hat{\Phi}(\omega) + \hat{V}'[\hat{\rho}'_{\Phi}(\omega)]) |\psi_p\rangle$ as a source term.

From equation (9) we see that for each choice of p , the value $\omega = |\epsilon_p|$ represents the threshold level that governs the asymptotic behavior of the associated Helmholtz kernel. As a ulterior *stringent* hypothesis, let us now

assume that the source term $|s_p^\Phi(\omega)\rangle$ is spatially localized, *i.e.* its evaluation can be restricted to a bounded real-space domain. Even when such a locality argument holds true, for $\omega > |\epsilon_p|$, the Helmholtz kernel behaves as a spherical wave and gives rise to a $|f_p^\Phi(\omega)\rangle$ which is *delocalized* over the entire real-space domain. Such delocalization is an intrinsic consequence of the value of ω , and does not depend on the particular choice of $\hat{\Phi}$. Only for $\omega < |\epsilon_p|$ the kernel has an exponential damping factor and $|f_p^\Phi(\omega)\rangle$ exhibits a bound-state behavior in the long range, once again provided that the above locality arguments on $|s_p^\Phi(\omega)\rangle$ – hence on \hat{V}' and $\hat{\Phi}$ – are valid.

These arguments demonstrate that for frequencies above the IP threshold energy, the response density operator cannot be expressed in terms of localized states. Locality of FS (and therefore of $\hat{\rho}'_\Phi(\omega)$) can be only recovered for frequencies below such threshold.

A. Localized behavior for static perturbations

The locality property of the fluctuation state can be directly confirmed for static perturbations at $\omega = 0$. Let us consider a static perturbing operator $\hat{\Phi}_s$ that can be restricted to a finite domain. Any local operator, *e.g.* a static electric field \mathbf{F} , with associated scalar potential $\Phi_s[\mathbf{F}](\mathbf{r}) = -\mathbf{F} \cdot \mathbf{r}$, would satisfy such a property. We consider the perturbed Hamiltonian $\hat{H}_0[\hat{\rho}_{\Phi_s}] + \hat{\Phi}_s[\mathbf{F}]$, where $\hat{\rho}_{\Phi_s}$ represents the GS density of the perturbed system, and assume that the perturbation strength ($|\mathbf{F}|$ in the above example) is sufficiently small to be consistent with a LR treatment of the problem. In this case, the occupied eigenfunction $|\psi_p^{\Phi_s}\rangle$ of the perturbed Hamiltonian can be extracted with traditional GS techniques, since from the discussion around Eq. (1), they exhibit bound-state behavior.

From the linearized Schrödinger equation of the perturbed problem, it is easy to see that the FS at $\omega = 0$ can be expressed in terms of the corresponding bound state of the perturbed Hamiltonian, *i.e.* $|f_p^{\Phi_s}(0)\rangle \simeq \hat{Q}_0 |\psi_p^{\Phi_s}\rangle$. For any value of p , the locality of the $|f_p^{\Phi_s}(0)\rangle$ therefore directly stems from the bound-state behavior of $|\psi_p^{\Phi_s}\rangle$. As $\omega = 0$ is by hypothesis lower than any of the $|\epsilon_p|$, this is consistent with the above considerations.

Thanks to such observations one can express quantities like the static polarizability tensor α_{ij} (see *e.g.*¹³) from GS calculations of the perturbed Hamiltonian. By expressing the induced dipole $\langle \delta \hat{\mathbf{r}} \rangle_{\Phi_s}$ in the form Eq. (3), and identifying the static response density as $\hat{\rho}'_{\Phi_s}(0) = \hat{\rho}_{\Phi_s} - \hat{\rho}_0$, we have

$$\alpha_{ij} = -\frac{1}{F_j} \left(\text{tr}(\hat{r}_i \hat{\rho}_{\Phi_s}[F_j]) - d_i^{(\text{GS})} \right), \quad (10)$$

where $\mathbf{d}^{(\text{GS})} = \langle \mathbf{r} \rangle$ is the ground state dipole. Figure 1 illustrates this procedure for a *CO* molecule perturbed by a static electric field. Care has been taken in keeping

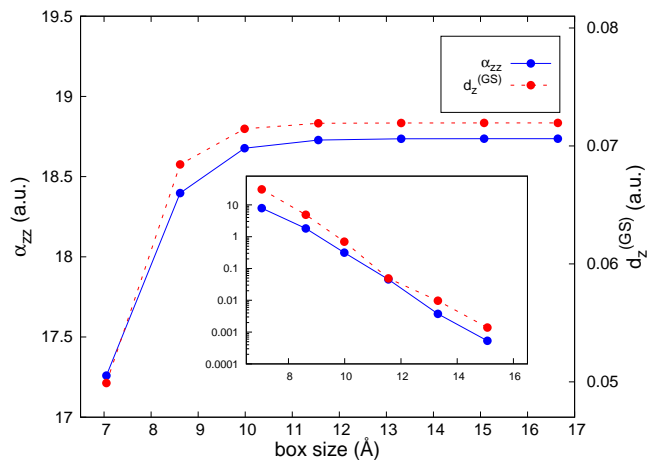


Figure 1: (Color online) Convergence of the static polarizability α_{zz} (blue continuous line) and of the ground state dipole d_z (red dotted line) of a *CO* molecule as a function of the size of the computational domain. The inset evidences the relative convergence (in percent) of the two quantities with respect to the corresponding value associated with the largest box.

the field strengths F_j small enough to preserve the validity of the linear response regime. The local character of the fluctuation states at $\omega = 0$ is probed by verifying the convergence of the static polarizability versus the size of the computational domain used to express the fluctuation states. The convergence rate is compared with the one of the static dipole of the molecule computed at zero external field. The results show that the typical sizes of $|f_p^\Phi(0)\rangle$ are analogous to those of the unperturbed KS orbitals, which is a direct proof of the bound-state behavior of the FS.

B. Localization properties at finite ω . Emergence of a strong and a weak regime

The analysis described above has shown that the building blocks of the response density $\{|f_p(\omega)\rangle\}$ exhibit bound-state long-range behavior *only* when $\omega < |\epsilon_p|$, while they behave as unbound states for values of ω above these threshold levels. This fact has interesting implications in the evaluation of LR quantities. Indeed, writing the response density in terms of FS implies that

$$\langle \delta \hat{O} \rangle_\Phi(\omega) = 2 \sum_p \langle \psi_p | \hat{O} | f_p^\Phi(\omega) \rangle, \quad (11)$$

where we have assumed that \hat{O} is a real symmetric operator and chosen a real set of occupied molecular states. Let us suppose that the observable of interest \hat{O} also can be restricted to localized domains in the same way as the potentials \hat{V} and $\hat{\Phi}$. Once again, it would be enough to consider a local operator for this condition to be met.

Two distinct regimes can be identified, on the basis of the value of ω . The first one, which we will refer to as the *below threshold* regime, is realized when ω is lower than the ionization potential $|\epsilon_h|$. In this case all the FS are below their threshold level and Eq. (11) is expressed in terms of genuinely localized quantities. At these frequencies, the linear response functional $\langle \delta \hat{O} \rangle_{\Phi}(\omega)$ may thus be evaluated with computational setups that are similar to those usually employed for GS calculations. The convergence in this regime is “strong”, in the sense that the basis functions employed enable us to express with arbitrary precision *both* the states $|\hat{O}\psi_p\rangle$ and $|f_p^{\Phi}(\omega)\rangle$. This is, of course, true even when the above states are only implicitly expressed by the employed LR treatment, as e.g. in the case of Krylov spaces generated from the states $|\hat{\Phi}(\omega)\psi_p\rangle$ ^{9,14,15}. In this regime, a computational treatment based on localized basis set is susceptible to provide a precise answer, assuming a reasonable level of completeness.

On the contrary, the computation of the linear response functional can be much more demanding in the *above threshold regime*, realized for $\omega > |\epsilon_h|$, in which fluctuation states start behaving as unbound wave functions. As by hypothesis $|\hat{O}\psi_p\rangle$ is bound-state like, *only* the scalar products of Eq. (11) can be evaluated in a localized domain. In this case a computational description of fluctuation states is only meaningful in the “weak” sense, where the bound state character of $\langle \psi_p | \hat{O} | \cdot \rangle$ behaves as a regulator. This is naturally achieved in the above mentioned Krylov space treatments, which explicitly deal with the expression of the scalar products. However, in this regime, imposing *by design* a localized behavior of the FS might reveal very dangerous in view of convergence of the results since the computational basis set has to be able to express a delocalized FS in the support domain of $\langle \psi_p | \hat{O} | \cdot \rangle$. This fact implies that the degrees of completeness of the computational basis should be increased by adding the oscillatory degrees of freedom needed to describe the FS in the relevant domain. Stated otherwise, to achieve precise results in this regime the underlying basis set has to be able to express delocalized states.

These considerations prove why, when a given computer code employs established GS numerical techniques, *it is much easier to converge LR quantities for ω below IP*. Above the IP threshold, a computational setup provides a reliable assessment of (11) only if it is able (even implicitly) to express an unbiased description of the fluctuation states in the support of $\{\langle \psi_p | \hat{O} | \cdot \rangle\}$. The non-vanishing spherical-wave behavior in the long range of $|f_p^{\Phi}(\omega)\rangle$ is therefore responsible for these convergence difficulties, especially for high frequencies. The IP threshold has an evident physical interpretation: for energies in a range higher than the ionization potential, localized and delocalized states might be coupled by the perturbation. Obtaining convergence of the observable quantities at these energies is therefore much more challenging if one employs numerical techniques which were originally tailored for GS calculations.

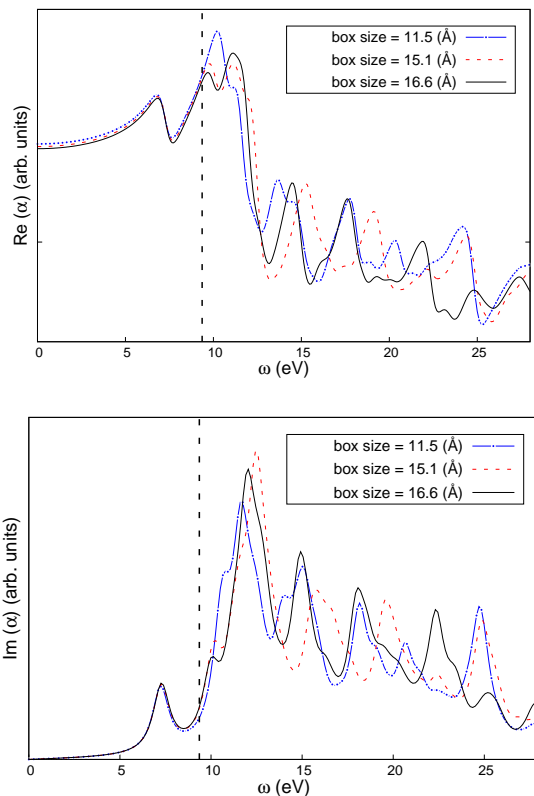


Figure 2: (Color online) LR dynamical polarizability of the *CO* molecule in computational setups with different sizes of the simulation domain. The top panel displays the real part and the bottom one the imaginary part. The value of the IP threshold is indicated by a vertical dashed line.

As an example in support of these arguments we compare the dynamical polarizability $\alpha(\omega)$ of a *CO* molecule computed in three distinct real-space computational setups, which differ in the dimension of the computational domain. Results reported in the two panels of Fig. 2 show the real and imaginary parts of $\alpha(\omega)$, respectively. In all the cases, a typical behavior emerges: the curves are almost coincident in the below threshold regime, whereas the effects of the computational setup are clearly visible for values of ω which exceed the threshold value. For these high-frequency regions, convergence might be reached for domains of much larger sizes than those presented here⁹, which are, as already noticed, more than sufficient below threshold.

III. EXCITATION OPERATORS. THE ANALYTIC STRUCTURE OF THE SUSCEPTIBILITY FUNCTIONAL

It is possible to reformulate this type of analysis by directly looking at LR quantities that are defined independently from the perturbing operators and target ob-

servable. This can be achieved by studying the *susceptibility functional*, defined implicitly through the operator equation

$$\hat{\rho}'(\omega) = (\omega - \mathcal{L})^{-1} \left[\hat{\Phi}(\omega), \hat{\rho}_0 \right] = \hat{\chi}(\omega) \hat{\Phi}(\omega). \quad (12)$$

The susceptibility (super)operator $\chi(\omega)$ is therefore closely related to the resolvent of the Liouvillian of Eq. (4), which is completely defined from the molecule Hamiltonian $\hat{H}_0[\hat{\rho}]$. The ‘‘excitation modes’’ of the molecule are defined through the *excitation operators*, satisfying

$$\hat{\mathcal{L}}\hat{E}_a = \Omega_a \hat{E}_a \quad \hat{E}_a \hat{\mathcal{L}} = \Omega_a \hat{E}_a, \quad (13)$$

together with the operator orthonormalization condition

$$\text{tr} \left(\hat{E}_a \hat{E}_b \right) = \delta_{ab}. \quad (14)$$

Excitations defined in this way may be considered as a basis of the operators that satisfy the transverse condition as in Eq. (6), and enable us to express the susceptibility as a spectral decomposition

$$\hat{\chi}(\omega) = \sum_{\{a\}} \frac{\hat{B}_a}{\omega - \Omega_a}, \quad (15)$$

where the action of the spectral (super)operator \hat{B}_a reads

$$\hat{B}_a \cdot = \hat{E}_a \text{tr} \left(\left[\hat{\rho}_0, \hat{E}_a \right] \cdot \right).$$

These formulas allow us to present an excitation-based description of Eq. (11) as follows

$$\langle \delta \hat{O} \rangle = \sum_{\{a\}} \frac{\text{tr} \left(\hat{O} \hat{B}_a \hat{\Phi}(\omega) \right)}{\omega - \Omega_a}. \quad (16)$$

A. Localization features of the excitation operators. The analytic structure of linear susceptibility

The properties of the Liouvillian superoperator $\hat{\mathcal{L}}$ (sketched in Appendix A) imply that both the left and right operator-valued eigenstates (13) satisfy the transverse condition (6) and can be parametrized as

$$\begin{aligned} \hat{E}_a &= \sum_p \left(|\phi_p^a\rangle \langle \psi_p| + |\psi_p\rangle \langle \chi_p^a| \right), \\ \hat{E}_a &= \left[\hat{\rho}_0, \hat{E}_a^t \right] = \sum_p \left(|\psi_p\rangle \langle \phi_p^a| - |\chi_p^a\rangle \langle \psi_p| \right). \end{aligned} \quad (17)$$

Each excitation ‘‘mode’’ of the system, with associated energy Ω_a , is thus described by a set of states $\{\phi_p^a, \chi_p^a\}_{p \in a}$, defined in the unoccupied subspace. These

objects represent, respectively, the state in which $|\psi_p\rangle$ is excited – or from whom it decays – when the system is subject to the *monochromatic* perturbation $\hat{\Phi}_a \equiv \left[\hat{E}_a, \hat{\rho}_0 \right]$. We denote with the shorthand $p \in a$ the fact that the sum of Eq. (17) is in general restricted only to a a -dependent subset of the occupied orbitals, depending on the particular symmetry of the excitation. Excited states satisfy the normalization condition

$$\sum_p \left(\langle \phi_p^a | \phi_p^b \rangle - \langle \chi_p^b | \chi_p^a \rangle \right) = \delta_{ab}. \quad (18)$$

By applying (13) to the above parametrization of the excitations, we obtain the following Sternheimer-type equations for the excited states:

$$\begin{aligned} \left[\Omega_a - (\hat{H}_0 - \epsilon_p) \right] |\phi_p^a\rangle &= \hat{Q}_0 \hat{V}'[\hat{E}_a] |\psi_p\rangle, \\ \langle \chi_p^a | \left[-\Omega_a - (\hat{H}_0 - \epsilon_p) \right] &= \langle \psi_p | \hat{V}'[\hat{E}_a] \hat{Q}_0. \end{aligned} \quad (19)$$

By writing these equations in a basis set of virtual states we obtain the well-known Casida’s eigenvalue equation² associated to the excitation energy Ω_a (see Appendix B). The spectrum is symmetric with respect to the inversion of the eigenvalues $\Omega_a \rightarrow -\Omega_a$ and, given a specific excitation $\{\phi_p^a, \chi_p^a\}$, the associated solution with opposite energy is described by the transposed pair $\{\chi_p^a, \phi_p^a\}$. We concentrate our analysis on the sector of positive energies, i.e. $\Omega_a > 0$.

In the same spirit as the previous sections, we present a formal solution of (19) as follows

$$\begin{aligned} |\phi_p^a\rangle &= \hat{G}_{\hat{T}}(\Omega_a + \epsilon_p) \left(\hat{V} |\phi_p^a\rangle + |s_p[\hat{E}_a]\rangle \right), \\ \langle \chi_p^a | &= \left(\langle \chi_p^a | \hat{V} + \langle s_p[\hat{E}_a] | \right) \hat{G}_{\hat{T}}(-\Omega_a + \epsilon_p), \end{aligned} \quad (20)$$

where the source terms are here defined as:

$$|s_p[\hat{O}]\rangle = \hat{Q}_0 \hat{V}'[\hat{O}] |\psi_p\rangle, \quad \langle s_p[\hat{O}] | = \langle \psi_p | \hat{V}'[\hat{O}] \hat{Q}_0.$$

First, we point out that the Helmholtz kernel associated with the $|\chi_p^a\rangle$ state contains an exponential damping factor for *all* the positive values of Ω_a . The asymptotic long-range behavior of $|\chi_p^a\rangle$ thus depends only on the locality properties of $\langle s_p[\hat{E}_a] |$, which in turn are related to the behavior of the operator $\hat{V}'[\hat{E}_a]$.

On the other hand, the long-range behavior of $\hat{G}_{\hat{T}}(\Omega_a + \epsilon_p)$ depends on Ω_a , with a threshold of value $\Omega_a = |\epsilon_p|$. For energies above this value, the behavior of the state $|\phi_p^a\rangle$ is delocalized regardless of the particular nature of the operators \hat{V} and \hat{V}' . This means that, if an excitation operator \hat{E}_a has an energy such that $\Omega_a > |\epsilon_p|$ for at least one of the $p \in a$ defining the excitation, such an operator *cannot* be parametrized by employing localized states only.

As a consequence, we deduce that a molecular excitation *may* be expressed in terms of localized states *only*

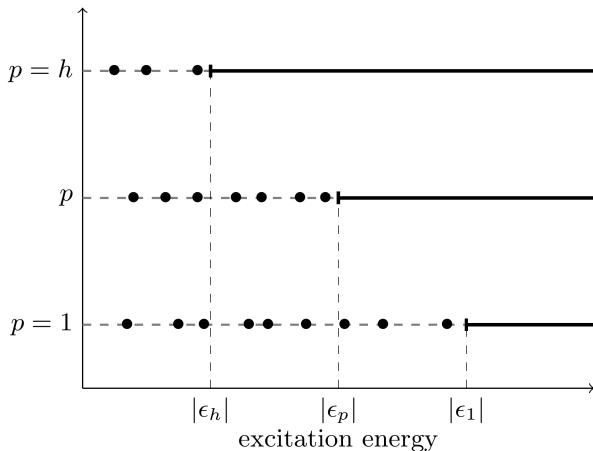


Figure 3: Excitation landscape. Excitations are split in horizontal lines according to the threshold levels p . Filled bullets refer to localized excitations whereas the thick black lines describe delocalized ones belonging a continuum spectrum.

when its energy is lower than *all* the ionization potentials of the occupied states participating to the excitation, namely, if $\Omega_a < |\epsilon_p| \forall p \in a$. This will only be true if the operator $\hat{V}'[\hat{E}_a]$ can be restricted to localized kets and bras. If any of these hypotheses does not hold, the excitations of the systems are defined via genuinely delocalized wave functions. Furthermore, excited states undergo the orthonormalization condition (18) that acts as an ulterior constraint and has the effect of determining the *quantization* of the energy levels Ω_a associated to the bound-state-like objects. For delocalized states, as in the case of the continuum states of \hat{H}_0 , such a condition has to be interpreted in a distributional sense and does not lead to quantization of Ω_a . Thus we can conclude that all the states $|\chi_p^a\rangle$ and the $|\phi_p^a\rangle$ with Ω_a below the threshold value constitute – if localized – a genuine *discrete* set, while those in the unbound energy range behave as generalized eigenvectors associated to a continuum spectrum.

The susceptibility functional has therefore a peculiar analytic structure made of multiple threshold energies, one for each of the ionization potentials associated with the occupied states. If an excitation \hat{E}_a has an energy $\Omega_a < |\epsilon_a|$ with $\epsilon_a = \max(\epsilon_p)_{p \in a}$, it may belong to a discrete spectrum; otherwise it *has* to belong to a continuum of states. The typical structure of the excitation landscape is depicted in Fig. 3, once again assuming that the perturbed potential $\hat{V}'[\hat{E}]$ gives rise to a bound state when applied to $|\psi_p\rangle$. Below the first ionization energy $|\epsilon_h|$ only localized excitations are present. Conversely, in the intermediate energy interval that extends from $|\epsilon_h|$ to the absolute value of the deepest occupied orbital, other discrete excitations may exist, and are embedded in a continuum of delocalized excitations. For energies higher than $|\epsilon_1|$ *no localized* excitations are anymore possible, regardless of the particular behavior of \hat{V}' : the excita-

tion landscape in this region contains a continuum set of delocalized states.

IV. PHYSICAL RELEVANCE OF THE EXCITATION SPECTRUM. A COMPARISON AMONG LOCALIZED AND CONTINUUM SECTORS

A. Observable features of the excitation spectrum

The analysis of the above section reveals that the excitation spectrum is composed of two distinct sectors. The first one is realized by the (finite) set of localized excitations with discrete eigenvalues. Applying the terminology introduced in section IA to the present case, we can affirm that such localized excitations behave as *observable* quantities and the associated energies possess ideal features of reproducibility: if the numerical basis employed is complete enough to represent the states $\{\phi_p^a, \chi_p^a\}$, the energy Ω_a converges to its reference value. A computational setup in which delocalized degrees of freedom are excluded by design can, in principle, describe excitations of this sector in an exact way.

The second sector contains a continuum spectrum of delocalized excitations. The associated states $|\phi_p^a\rangle$ behave as the unbound virtual orbitals analyzed in⁶ (see Sec.IA) and are deeply affected by the computational setup used to represent them. In particular, the presence of a finite computational domain implies a fictitious discretization of the excitation levels and the associated spectrum exhibits an explicit dependence on the size of the domain. Excitations belonging to this sector thus lose their intrinsic physical meaning and have to be understood simply as objects providing an effective representation of the density of states (DOS) of the Liouvillian superoperator, by means of the spectral operators \hat{B}_a of Eq. (16).

In Fig. 4 and 5b we plot the DOS of the Liouvillian for the CO and benzene molecules, for two different numerical setups of increasing simulation domain size. The figures show clearly that all the excitations of energy Ω_a lying below the molecule ionization potential (IP) have a localized behavior, as their energy does not vary with the size of the simulation domain. On the other hand, above the first ionization threshold the excitations have an energy that strongly depends on the computational setup, which is in line with the expected pseudo-continuum character of this sector. In addition, for the benzene molecule it is also possible to identify excitations of energy above IP that still possess a localized behavior, fulfilling the condition $\Omega_a < |\epsilon_a|$ [see the inset of Fig. 5b]. In Fig. 5a we collect the excitations of the benzene molecule following this criterion: the agreement with the expected analytic structure presented in Fig. 3 is remarkable.

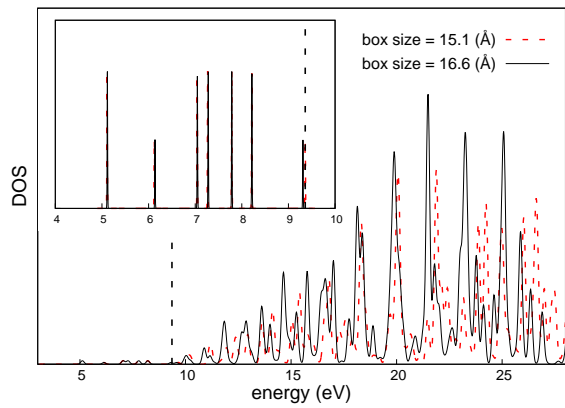


Figure 4: (Color online) DOS of CO excitation energies Ω_a . The two curves correspond to two computational domains of different sizes. The inset contains the sole contribution of localized excitations. The first IP threshold is depicted as the vertical dashed line.

B. Relative importance of discrete and continuum excitations

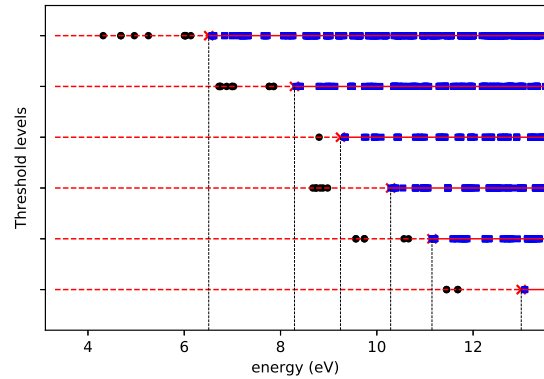
So far, we have showed with both formal arguments and numerical calculations which show that the excitation spectrum splits into two sectors with very distinct features. It is then reasonable to ask which is their relative importance and, in particular, if there is the possibility to express observable quantities, in a suitable ω range, by only employing the sector of genuinely localized excitations. Stated otherwise, can the *restriction* of $\hat{\chi}(\omega)$ to only localized excitations be useful to extract some physical quantities?

In this regard it is illustrative to apply the above arguments to the evaluation of the dynamical polarizability tensor $\alpha_{ij}(\omega)$. Employing the formalism of Eq. (16) provides

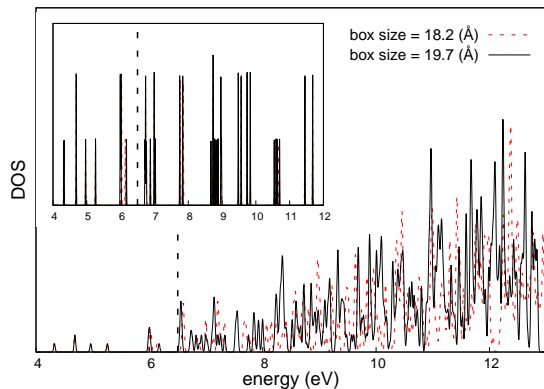
$$\text{Im}(\alpha_{ij}(\omega)) = \sum_a \text{tr}(\mathbf{r}_i \hat{B}_a \mathbf{r}_j) \delta(\omega - \Omega_a). \quad (21)$$

The imaginary part of the dynamical polarizability tensor therefore can be seen as the partial density of states of the Liouvillian projected on $\hat{\mathbf{r}}_i \hat{\mathbf{r}}_j$. For this reason, it is evident that below the *first* excitation threshold (IP), the contributions to $\text{Im}(\alpha_{ij}(\omega))$ can be expressed only in terms of localized, thus observable, excited states. On the contrary, the non-observable, basis dependent continuum sector of the excitation starts to contribute to $\text{Im}(\alpha_{ij}(\omega))$ above IP; the localized sector is not anymore sufficient to express Eq. (21) in this ω range.

To further address this point, we exploit the formal relation between response density and susceptibility functional $\hat{\rho}'(\omega) = \hat{\chi}(\omega)\hat{\Phi}(\omega)$ and make usage of the representation (7) of the response density. Pursuing this procedure provides the excitation-based representation of FS



(a) Excitation landscape



(b) DOS of Excitation energies

Figure 5: (Color online) Benzene molecule. (a): Excitation Landscape: the excitations have been separated into discrete (black circle) and continuum (blue square) sectors following the criteria of Appendix 3. (b): DOS of excitation energies Ω_a . The first IP threshold is depicted as the vertical dashed line.

as follows:

$$|f_p^{\mathbf{r}_j}(\omega)\rangle = \sum_a \frac{\langle w_p^a | \hat{\mathbf{r}}_j | \psi_p \rangle}{\omega^2 - \Omega_a^2} |w_p^a\rangle, \quad (22)$$

where we have employed the notation $|w_p^a\rangle = |\phi_p^a\rangle + |\chi_p^a\rangle$. This is the well-known equivalence of the Sternheimer formalism with the spectral representation of the LR susceptibility, that is always valid in any computational basis set provided the full spectrum of excited states (including the pseudo-continuum ones) is considered. The above equation shows that both the sectors of localized and delocalized states $|w_p^a\rangle$ contribute to the FS for *any* value of ω . However, the discussion of Sec. II has shown that the FS is a genuinely localized state below the IP threshold. The equivalence implies that in this regime the contribution of the delocalized excitations in Eq. (22) has to resum into a localized, asymptotically vanishing FS. The localization below threshold is manifest for the imaginary part of the FS, that is explicitly written as

the sum over $|w_p^a\rangle$ below IP [see Eq. (21)]. Instead, for the real part of FS, the *entire* excitation spectrum contributes to the construction of the FS, which has to be interpreted as a localized wave-packet in the basis of the pseudo-continuum excitations.

Hence, contrary to GS, we conclude that *for LR calculations, even when the response density is localized, the contribution coming from the continuum sector of the excitations can never be neglected*. In other terms, we cannot exclude that a fluctuation state, even if genuinely localized, will have a nonzero projection into the essential part of the Liouvillian spectrum, which in turn cannot be expressed by only employing localized states. This fact shows that it is *impossible* to restrict the spectral representation of $\chi(\omega)$ to only localized, discrete, *code-independent* excitations, even for representing LR results in the $\omega = 0$ limit. The basis-dependent sector of the excitation will be needed to represent the linear susceptibility in the appropriate region of space.

C. Effective representation of $\hat{\chi}(\omega)$ for localized response densities

The above arguments have shown that the only quantity which can be expressed in terms of the localized sector of the excitation spectrum is the imaginary part of the susceptibility functional $\hat{\chi}(\omega)$ in the optical regime (i.e. $\omega < \text{IP}$). The response density for a generic perturbation and ω value requires the inclusion in the Liouvillian DOS of the delocalized part of the excitation spectrum, even in view of expressing localized fluctuation states.

A given numerical treatment typically provides only an effective description of the sector of delocalized excitation which in turns gives rise to an effective representation of the susceptibility functional. It appears therefore important to identify indicators that enable us to validate the capability of $\hat{\chi}(\omega)$ to express the localized FS [and thus the complete response density $\hat{\rho}'_{\Phi}(\omega)$] below the IP threshold. For instance, the static polarizability tensor α_{ij} can be calculated as

$$\alpha_{ij} = \oint da \frac{w_{ij}^a}{\Omega_a^2}, \quad w_{ij}^a \equiv \sum_{p \in a} |\langle \psi_p | \hat{\mathbf{r}}_i | w_p^a \rangle|^2. \quad (23)$$

Such a value can be straightforwardly compared with results of Eq. (10), where the localized FS is obtained from GS techniques as described in Sec. II. This is the well-known S_{-2} sum-rule (see, e.g.,¹⁶), that expresses the value of the static polarizability in terms of the molecule's oscillator strengths. In the context of a numerical evaluation of the excitations, such sum rule can therefore be interpreted as the capability of the pseudo-continuum sector of the excitation to express the localized FS – and consequently the reference value of α_{ij} , at $\omega = 0$.

We have analyzed the fulfillment of Eq. (23) for molecules of various sizes and symmetries. In Fig. 6 we show explicitly such a comparison for the *CO* and *C₆₀*

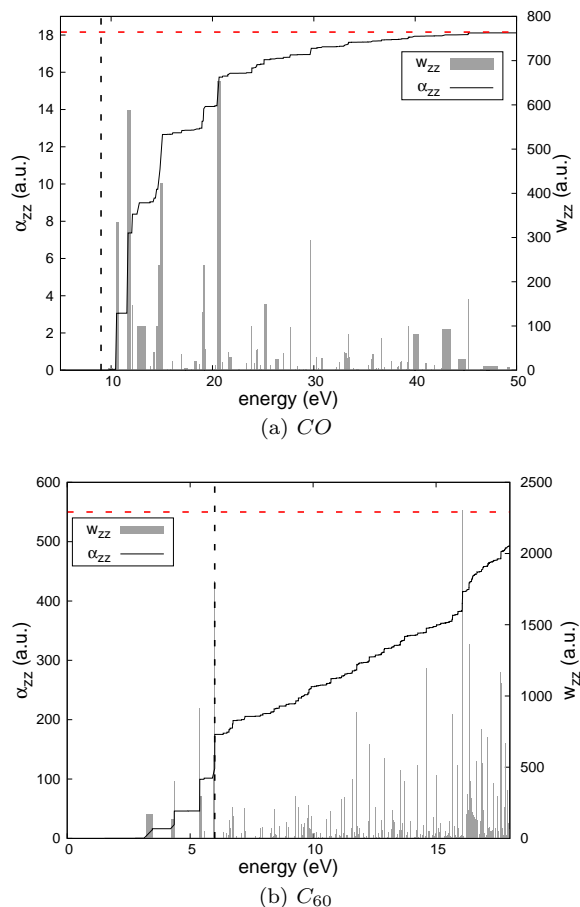


Figure 6: (Color online) Convergence of the static polarizability α_{zz} of the *CO* (panel a) and *C₆₀* (panel b) molecules, w.r.t the excitations considered in the integral of Eq. (23). The reference value obtained from Eq. (10) is represented as a horizontal red dashed line. The value of the IP energy is indicated by a vertical black dashed line. For each excitation energy, the values of the oscillator strengths w_{zz}^a of Eq. (23) are also plotted.

molecules. We notice that, in both the cases, a considerable number of pseudo-continuum excitations is required to find the correct result; the sector of the discrete excitations appears to be largely insufficient to express the reference fluctuation state. In particular, in the case of *CO*, the complete excitation set is able to reproduce the reference value of α_{zz} , but the contribution of localized sector is negligible w.r.t. the one coming from the continuum one. Instead, for the *C₆₀*, the localized sector provides a more significant contribution but a higher number of the pseudo-continuum excitations should be included in the sum of (23) to satisfy the S_{-2} sum-rule.

This scenario is confirmed by the results for the other molecules presented in Table I; varying the molecule size and symmetry does not modify the qualitative behavior of the excitation landscape (see Supplementary Material at¹⁷ for further details about the dependence of the excitation landscape on the size of the molecule). In addi-

Mol. (N_{at})	$\langle \alpha \rangle_{ref}$	$\langle \text{Eq. (23)} \rangle$	% LS	% $\Omega_a < IP$
N ₂ (2)	12.49	12.38	$1.6 \cdot 10^{-6}$	$1.6 \cdot 10^{-6}$
CO (2)	15.07	14.79	8.6	8.6
H ₂ O (3)	10.87	10.72	0.1	0.1
Benzene (12)	71.11	70.32	3.0	$7.2 \cdot 10^{-5}$
Aflatoxin (35)	N/A	226.27	15.87	14.84
C ₆₀ (60)	561.59	509.60	38.47	34.41

Table I: Comparison of the average static polarizabilities for molecules of various sizes (indicated by the number of atoms N_{at}) and symmetries, obtained from the reference calculations of Eq. (10) and from the sum over the excitations extracted from the Casida coupling matrix. The last two columns show the percentage of the contribution to Eq. (23) which come from the localized sector (LS) of the excitations, by focusing also on the LS before the first IP, where all the excited states are localized. We see that by increasing the molecules’ sizes the LS contribution increases in percentage, however being always largely insufficient; the contribution coming from the pseudocontinuum excitations can never be neglected.

tion the localized sector of the excitations alone is largely insufficient to express α_{ij} . Such considerations are independent of the number of atoms of the molecule, as the range in ω needed to achieve satisfactory agreement is largely above the last valence ionization threshold, which is an intensive quantity.

Once again, this proves that, for any numerical treatment, pseudo-continuum excitations are fundamental in view of a correct expression of the response density below IP. However, the fulfillment of S_{-2} is not a universal quality indicator of $\chi(\omega)$ as the latter may be still badly represented in the high frequency regime: it is enough to recall the strong dependency on the box size of the absorption spectra of Fig. 4. For all these computational setups, the S_{-2} sum rule is, on the contrary, well satisfied.

V. CONCLUSIONS AND PERSPECTIVES

We have performed a critical assessment of the fundamental equations that govern the linear response theory of molecular systems. We have shown that the response density for values of ω above the first IP *has* to be expressed in terms of genuinely delocalized states. This result has relevant consequences in the choice of the appropriate computational basis set to be employed for the LR treatment.

For frequencies below the IP threshold, under reasonable assumptions on the perturbing operators, the response density is genuinely localized. In this regime, it is possible to achieve convergence of the numerical result by increasing the basis completeness in the same spirit as GS calculations. More complete basis sets will pro-

vide more precise results, regardless of their asymptotic behavior. A comparison among various computer codes in view of reproducibility should therefore be performed initially in this regime.

From a computational perspective, a set of localized basis functions, which *by design* excludes long-range oscillatory behavior, may be complete enough to express $\hat{\chi}(\omega)$ *below* IP, but it will never be able to capture the entire features of the response density operator above the ionization threshold: the system’s excitations expressed in this basis are reliable *only* in the optical regime. High-energy excitations belong to a continuum sector (which should not be confused with the continuum eigenstates of the unperturbed GS Hamiltonian) and the excitation density of states for $\omega > IP$ provided in this way has to be considered as an effective representation of $\hat{\chi}(\omega)$ in order to fulfill Eq. (22). This fact has been pointed out several times (see, e.g.,^{11,12}) in the literature.

On the contrary, computational treatments that are able to express delocalized states may be adequate to describe more efficiently the excitations DOS, especially due to the capability of the basis to capture oscillatory behavior. Like in localized basis sets, excitations below threshold can be precisely calculated with the same paradigm of traditional GS calculations. However, in this computational setup, to correctly express the linear susceptibility below IP – more specifically the real part of the fluctuation state – care should be taken in considering enough delocalized excitations to guarantee fulfillment of sum-rules like Eq. (23).

For generic LR quantities, which are functionals of the linear susceptibility, the quality of $\hat{\chi}(\omega)$ will then depend on the frequency ω of interest: a given spectral representation may provide high-quality results for static $\omega = 0$ regimes, while being unable to provide convergence for high-energy absorption spectra. Absorption spectra below IP are relatively easy to converge, as they are observables which depend only on localized quantities. However, the fulfillment of sum rules like the S_{-2} is not *per se* a guarantee of the quality of high energy absorption spectra. Moreover, if the basis set employed for the evaluation of the target α_{ij} in Eq. (23) is identical to the one employed for the LR treatment, the S_{-2} is trivially satisfied in the basis. It is therefore important to have a set of reference values in the complete basis set limit.

We believe that these considerations, based on simple manipulations on the fundamental LR equations, will help to increase the reliability of LR calculations in the community, and in building test-sets that would help in the *calibration* of present and future computer codes, thereby increasing the predictivity of present-day theoretical approaches. The example of the static polarizability of molecules presents ideal features as an initial playground, as it has already been pointed out recently¹⁸. Work is ongoing in this direction.

ACKNOWLEDGMENTS

We thank Thierry Deutsch for useful comments on the manuscript and Laura Ratcliff for proofreading.

Appendix A: Transverse operators and right action of the Liouvillian superoperator

Given an unperturbed Hamiltonian and density operator, which satisfy $[\hat{H}_0, \hat{\rho}_0] = 0$, we may decompose a generic operator in two contributions, $\hat{O} = \hat{O}_\parallel + \hat{O}_\perp$, defined as:

$$\begin{aligned}\hat{O}_\parallel &\equiv \hat{\rho}_0 \hat{O} \hat{\rho}_0 + \hat{Q}_0 \hat{O} \hat{Q}_0, \\ \hat{O}_\perp &\equiv \hat{\rho}_0 \hat{O} \hat{Q}_0 + \hat{Q}_0 \hat{O} \hat{\rho}_0 = [\hat{\rho}_0, [\hat{\rho}_0, \hat{O}]],\end{aligned}\quad (\text{A1})$$

with $\hat{Q}_0 = \hat{\mathbb{I}} - \hat{\rho}_0$. The operator \hat{O}_\parallel is constructed to satisfy $[\hat{\rho}_0, \hat{O}_\parallel] = [\hat{H}_0, \hat{O}_\parallel] = 0$. Also, it is easy to verify that given two generic \hat{O}, \hat{O}' , we have $\text{tr}(\hat{O}'_0 \hat{O}_\perp) = 0$.

The projection \hat{O}_\perp is therefore the relevant part for the commutator

$$[\hat{\rho}_0, \hat{O}] = [\hat{\rho}_0, \hat{O}_\perp] = \hat{\rho}_0 \hat{O} \hat{Q}_0 - \hat{Q}_0 \hat{O} \hat{\rho}_0, \quad (\text{A2})$$

consequently the Liouvillian superoperator, which contains by definition a commutator with \hat{H}_0 and $\hat{\rho}_0$ in its unperturbed and coupling part respectively, is constructed such as $\hat{\mathcal{L}}\hat{O} = \hat{\mathcal{L}}\hat{O}_\perp$. In addition, the image operator satisfies the transverse condition $(\hat{\mathcal{L}}\hat{O})_\parallel = 0$.

Here \hat{V}' , which expresses the perturbation induced by the density dependence on ground state Hamiltonian, can be conveniently expressed by defining the *scalar coupling kernel*

$$U[\hat{O}; \hat{O}'] \equiv \int d\mathbf{r} d\mathbf{r}' \text{tr} \left(\hat{O} \frac{\delta \hat{V}[\hat{\rho}_0]}{\delta \rho(\mathbf{r}, \mathbf{r}')} \right) \text{tr}(\hat{O}'|\mathbf{r}\rangle\langle\mathbf{r}'|), \quad (\text{A3})$$

on the basis of which

$$\hat{V}'[\hat{O}] = \int d\mathbf{r} d\mathbf{r}' U[\mathbf{r}\rangle\langle\mathbf{r}'|; \hat{O}] |\mathbf{r}'\rangle\langle\mathbf{r}|. \quad (\text{A4})$$

We notice here in passing that in general $\hat{V}'[\hat{O}_\perp] \neq \hat{V}'[\hat{O}]$. We assume nonetheless that the scalar coupling kernel is symmetric, i.e. $U[\hat{O}; \hat{O}'] = U[\hat{O}'; \hat{O}]$, which is a condition generally satisfied in DFT Hamiltonians.

The action from the right of $\hat{\mathcal{L}}$ can be assessed through the equivalence $\text{tr}(\hat{O}'(\hat{\mathcal{L}}\hat{O})) = \text{tr}((\hat{O}'\hat{\mathcal{L}})\hat{O}_\perp)$, which implies that also $(\hat{O}'\hat{\mathcal{L}})_\parallel = 0$. We obtain:

$$\hat{O}\hat{\mathcal{L}} = -[\hat{H}_0, \hat{O}] + \int d\mathbf{r} d\mathbf{r}' U[[\hat{\rho}_0, \hat{O}]; |\mathbf{r}'\rangle\langle\mathbf{r}|] (|\mathbf{r}\rangle\langle\mathbf{r}'|)_\perp.$$

With this definition it is explicitly apparent that $\hat{O}\hat{\mathcal{L}} = \hat{O}_\perp\hat{\mathcal{L}}$, similarly to the left action. The action of the Liouvillian reads as

$$\begin{aligned}\hat{O}\hat{\mathcal{L}} &= -[\hat{H}_0, \hat{O}] + \left(V'[[\hat{\rho}_0, \hat{O}]] \right)_\perp = -[\hat{H}_0, \hat{O}] + \\ &+ \hat{\rho}_0 \hat{V}'[[\hat{\rho}_0, \hat{O}]] \hat{Q}_0 + \hat{Q}_0 \hat{V}'[[\hat{\rho}_0, \hat{O}]] \hat{\rho}_0.\end{aligned}\quad (\text{A5})$$

In the last equivalence we have made usage of the transverse property [see Eq. (6)] of the coupling superoperator. The above formulas together with Eq. (5) are the starting point for the solution of the eigenvalues equations (13).

Appendix B: Casida equations

We show that Casida equations are equivalent to the equations of motion for the excited states (19). To this aim we introduce an explicit basis $\{|s\rangle\}$ in the subspace of empty states so that both $|\phi_p^a\rangle$ and $\langle\chi_p^a|$ can be represented as

$$|\phi_p^a\rangle = \sum_s X_{ps}^a |s\rangle, \quad \langle\chi_p^a| = \sum_s \langle s| Y_{ps}^a. \quad (\text{B1})$$

Plugging this expansion into Eqs. (19) and projecting on a arbitrary element of the basis provides a linear systems of equations for the coefficients X_{ps}^a and Y_{ps}^a , namely

$$\begin{aligned}\sum_{s'} (H_{0ss'} - \epsilon_p \delta_{ss'}) X_{ps'}^a + \langle s| \hat{V}'[\hat{E}_a] |\psi_p\rangle &= \Omega_a X_{ps}^a, \\ -\sum_{s'} (H_{0s's} - \epsilon_p \delta_{s's}) Y_{ps'}^a - \langle\psi_p| \hat{V}'[\hat{E}_a] |s\rangle &= \Omega_a Y_{ps}^a.\end{aligned}$$

Using this basis, excitation operators are expressed as a linear combination of transition operators $\hat{e}_{ps} \equiv |s\rangle\langle\psi_p|$ and $\hat{d}_{sp} \equiv |\psi_p\rangle\langle s|$, as follows

$$\hat{E}_a = \sum_{ps} \hat{e}_{ps} X_{ps}^a + \hat{d}_{sp} Y_{ps}^a. \quad (\text{B2})$$

The contribution of the coupling operator in the equations of motion of excited states can be expressed by means of the coupling kernel (A3) as

$$\begin{aligned}\langle s| \hat{V}'[\hat{E}_a] |\psi_p\rangle &= \sum_{qs'} \left(U[\hat{d}_{sp}; \hat{e}_{qs'}] X_{qs'}^a + U[\hat{d}_{sp}; \hat{d}_{s'q}] Y_{qs'}^a \right), \\ \langle\psi_p| \hat{V}'[\hat{E}_a] |s\rangle &= \sum_{qs'} \left(U[\hat{e}_{ps}; \hat{e}_{qs'}] X_{qs'}^a + U[\hat{e}_{ps}; \hat{d}_{s'q}] Y_{qs'}^a \right),\end{aligned}$$

Given that we employ real functions both for p and s we also have $\hat{e}_{ps}(\mathbf{r}, \mathbf{r}') = \hat{d}_{sp}(\mathbf{r}', \mathbf{r})$ and the density operator is symmetric in $\mathbf{r} \leftrightarrow \mathbf{r}'$. These conditions applied together imply that $U[\hat{e}_{ps}; \hat{e}_{qs'}] = U[\hat{d}_{sp}; \hat{d}_{s'q}]$, as well as $U[\hat{d}_{sp}; \hat{e}_{qs'}] = U[\hat{e}_{ps}; \hat{d}_{s'q}]$, and so Eqs. (19) can be recasted in matrix form

$$\sum_{qs'} \begin{pmatrix} F_{ps}^{qs'} & D_{ps}^{qs'} \\ -D_{ps}^{qs'} & -F_{ps}^{qs'} \end{pmatrix} \begin{pmatrix} X_{qs'}^a \\ Y_{qs'}^a \end{pmatrix} = \Omega_a \begin{pmatrix} X_{ps}^a \\ Y_{ps}^a \end{pmatrix} \quad (\text{B3})$$

where

$$\begin{aligned} F_{ps}^{qs'} &= (H_{0ss'} - \epsilon_p \delta_{ss'}) \delta_{pq} + U[\hat{d}_{sp}; \hat{e}_{qs'}], \\ D_{ps}^{qs'} &= U[\hat{d}_{sp}; \hat{d}_{s'q}]. \end{aligned} \quad (\text{B4})$$

This is the well-known Casida equation in the basis of transitions for the excitation of energy Ω_a . The computational reliability of the eigenvalue Ω_a depends on the capability of the basis set $|s\rangle$ to fulfill of Eq. (B1). Note that this may be valid only for a subset of the solutions of the eigenproblem (B3).

Appendix C: Computational details

The illustrative calculations presented in this paper are all KS-DFT calculations at the LDA level of theory. We employed the BIGDFT code¹⁹, that makes usage of Daubechies Wavelets as computational basis set. This orthonormal basis presents optimal features in view of reproducibility of the results, and it has proved to represent precisely and explicitly a system with free as well as periodic boundary conditions (BC), without any basis set superposition error nor supercell aliasing, for free BC.

For ground-state results, the important parameters in the context of convergence of the calculations are the spacing of the wavelet grid and the size of the simulation domain: for bound-state like functions, convergence is achieved by reducing the grid spacing and increasing the simulation box size. Norm-conserving pseudopotentials (PSP) of the HGH²⁰ type are employed to remove the core electrons. Such PSP have proved to be able to provide all-electron accuracy for molecular systems.

Our LR calculation is performed with the Casida formalism²¹, where it is also important to include a sufficient number of unoccupied states in the transition operators. As pointed out in the text such states behave as plane waves: the features of wavelets enable us to represent localized and delocalized states on equal footing. For each size of the simulation domain, the number of unoccupied states used for building the Casida' eigenproblem has been increased up to convergence of the presented results.

All the details of the calculations, the *Jupyter* notebooks of the data analysis, including the ones presented in the Supplementary Material, may be found at the URL²²

1. CO

The molecule is oriented along the z axis and the z -dimension of the box is equal to 11.5, 15.1 and 16.6

Ångström. Bond length is 1.235Ångström, stretched w.r.t. the equilibrium LDA position such as to increase the value of the polarizability. 280 unoccupied states have been computed in each setup, which span an energy range of the empty KS orbitals of 30.7, 19.3 and 15.4 eV respectively. These energy values should not be confused with the excitations energy ranges reported in Figs. 2 and 4. Once again, such choices of unoccupied states ensures convergence of the excitation spectrum in the presented range.

The DOS reported in figure 4 are obtained by adding a smearing parameter of 8×10^{-2} eV whereas in the inset the shift is 5×10^{-3} eV.

2. Benzene

The molecule is oriented in the xy plain and the x -dimension of the box is equal to 15.0, 18.2 and 19.7 Ångström. 220 states have been computed in each setup and the corresponding energy range for the unoccupied states is 16.9, 11.6 and 9.7 eV.

The DOS reported in Fig. 5b are obtained by adding a complex shift of 1.3×10^{-2} eV whereas in the inset the shift is 5×10^{-3} eV.

3. Excitation Landscapes

The attribution of each excitation of the appropriate sector passes through the evaluation of the norm of the states ϕ_p^a and χ_p^a , in the form of Eq. (B1). Only states with norm greater than a given tolerance value (taken as the 1% of the total norm) are considered and the associated excitation is attributed to the first sector if the condition $\Omega_a < |\epsilon_p|$ with $\epsilon_a = \max(\epsilon_p)_{p \in a}$. This computational procedure provides results in optimum agreement with the formal arguments reported above. Indeed the excitation landscape of [for instance, Fig. (5a) for the case of benzene] evidences the presence of discrete excitations (the black bullets) embedded in a pseudo-continuum represented by the blue filled square. The comparison between the excitation DOS computed in different computational domain shows that this way of labeling the excitation is in line with the locality of the corresponding states since the corresponding energies easily converge with respect to the increase of the simulation domain.

¹ K. Lejaeghere, G. Bihlmayer, T. Björkman, P. Blaha, S. Blügel, V. Blum, D. Caliste, I. E. Castelli, S. J. Clark,

A. Dal Corso, S. de Gironcoli, T. Deutsch, J. K. Dewhurst,

- I. Di Marco, C. Draxl, M. Duřak, O. Eriksson, J. A. Flores-Livas, K. F. Garrity, L. Genovese, P. Giannozzi, M. Giantomassi, S. Goedecker, X. Gonze, O. Grranäs, E. K. U. Gross, A. Gulans, F. Gygi, D. R. Hamann, P. J. Hasnip, N. A. W. Holzwarth, D. Iuřan, D. B. Jochym, F. Jollet, D. Jones, G. Kresse, K. Koepernik, E. Küçükbenli, Y. O. Kvashnin, I. L. M. Locht, S. Lubeck, M. Marsman, N. Marzari, U. Nitzsche, L. Nordström, T. Ozaki, L. Paulatto, C. J. Pickard, W. Poelmans, M. I. J. Probert, K. Refson, M. Richter, G.-M. Rignanese, S. Saha, M. Scheffler, M. Schlipf, K. Schwarz, S. Sharma, F. Tavazza, P. Thunström, A. Tkatchenko, M. Torrent, D. Vanderbilt, M. J. van Setten, V. Van Speybroeck, J. M. Wills, J. R. Yates, G.-X. Zhang, and S. Cottenier, *Science* **351** (2016), 10.1126/science.aad3000.
- ² M. E. Casida, “Time-dependent density functional response theory for molecules,” in *Recent Advances in Density Functional Methods* (World Scientific, 1995) pp. 155–192.
- ³ E. Runge and E. K. U. Gross, *Phys. Rev. Lett.* **52**, 997 (1984).
- ⁴ G. Onida, L. Reining, and A. Rubio, *Rev. Mod. Phys.* **74**, 601 (2002).
- ⁵ The threshold value in general should correspond to asymptotic value of the potential at infinity.
- ⁶ N. M. Boffi, M. Jain, and A. Natan, *The Journal of Chemical Physics* **144**, 084104 (2016).
- ⁷ For a sake of concreteness, we are limiting our considerations to observables that do not explicitly depend on time.
- ⁸ In what follows we will assume that the set of states $\{|\psi_p\rangle\}$ are described by real functions, when projected in the \mathbf{r} -representation.
- ⁹ D. Rocca, R. Gebauer, Y. Saad, and S. Baroni, *The Journal of Chemical Physics* **128**, 154105 (2008).
- ¹⁰ G. D. Mahan, *Phys. Rev. A* **22**, 1780 (1980).
- ¹¹ H. Hübener, M. A. Pérez-Osorio, P. Ordejón, and F. Giustino, *Phys. Rev. B* **85**, 245125 (2012).
- ¹² H. Hübener and F. Giustino, *Phys. Rev. B* **89**, 085129 (2014).
- ¹³ D. S. De, S. Saha, L. Genovese, and S. Goedecker, *Phys. Rev. A* **97**, 063401 (2018).
- ¹⁴ B. Walker, A. M. Saitta, R. Gebauer, and S. Baroni, *Phys. Rev. Lett.* **96**, 113001 (2006).
- ¹⁵ J. Brabec, L. Lin, M. Shao, N. Govind, C. Yang, Y. Saad, and E. G. Ng, *Journal of Chemical Theory and Computation* **11**, 5197 (2015), <https://doi.org/10.1021/acs.jctc.5b00887>.
- ¹⁶ L. O. Wagner, Z.-h. Yang, and K. Burke, “Exact conditions and their relevance in tddft,” in *Fundamentals of Time-Dependent Density Functional Theory*, edited by M. A. Marques, N. T. Maitra, F. M. Nogueira, E. Gross, and A. Rubio (Springer Berlin Heidelberg, Berlin, Heidelberg, 2012) pp. 101–123.
- ¹⁷ URL inserted by the publisher.
- ¹⁸ D. Hait and M. Head-Gordon, *Phys. Chem. Chem. Phys.* **20**, 19800 (2018).
- ¹⁹ L. Genovese, A. Neelov, S. Goedecker, T. Deutsch, S. A. Ghasemi, A. Willand, D. Caliste, O. Zilberberg, M. Rayson, A. Bergman, and R. Schneider, *The Journal of Chemical Physics* **129**, 014109 (2008), <https://doi.org/10.1063/1.2949547>.
- ²⁰ A. Willand, Y. O. Kvashnin, L. Genovese, . Vzquez-Mayagoitia, A. K. Deb, A. Sadeghi, T. Deutsch, and S. Goedecker, *The Journal of Chemical Physics* **138**, 104109 (2013), <https://doi.org/10.1063/1.4793260>.
- ²¹ B. Natarajan, L. Genovese, M. E. Casida, T. Deutsch, O. N. Burchak, C. Philouze, and M. Y. Balakirev, *Chemical Physics* **402**, 29 (2012).
- ²² https://github.com/luigigenovese/LR-nb/blob/master/POLARIZABILITY/nb_paper/LR_Analysis.ipynb.