

Hubbard parameters from density-functional perturbation theory

Iurii Timrov, Nicola Marzari, and Matteo Cococcioni

*Theory and Simulation of Materials (THEOS) and National Centre for Computational Design and Discovery of Novel Materials (MARVEL),
École Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland*



(Received 4 May 2018; revised manuscript received 27 July 2018; published 16 August 2018)

We present a transparent and computationally efficient approach for the first-principles calculation of Hubbard parameters from linear-response theory. This approach is based on density-functional perturbation theory and the use of monochromatic perturbations. In addition to delivering much improved efficiency, the present approach makes it straightforward to calculate automatically these Hubbard parameters for any given system, with tight numerical control on convergence and precision. The effectiveness of the method is showcased in three case studies— Cu_2O , NiO , and LiCoO_2 —and by the direct comparison with finite differences in supercell calculations.

DOI: [10.1103/PhysRevB.98.085127](https://doi.org/10.1103/PhysRevB.98.085127)

I. INTRODUCTION

The development of density-functional theory (DFT) [1,2] has allowed modeling of a broad spectrum of properties for a large variety of systems. In practical applications, DFT relies on approximations to the exchange-correlation (xc) electronic interactions, among which the local-density approximation (LDA) and the generalized-gradient approximation (GGA) are the most popular ones. Both approximations suffer from self-interaction errors (SIE), which limit the accuracy only to systems with weak and moderate electronic correlations. In systems with strongly localized electrons of d and f types, SIE in LDA and GGA are much larger, which leads to overdelocalization of these electrons and quantitative and sometimes even qualitative failures in the description of complex materials (e.g., metallic instead of insulating ground states).

Various corrective methods have been devised to deal with SIE. In particular, in DFT with hybrid functionals—such as, e.g., B3LYP [3–7] or HSE06 [8,9]—a fraction of the nonlocal Fock exchange is used for all (strongly localized and nonstrongly localized) electrons with a fraction of (semi)local exchange and a fully (semi)local correlation. This approach is computationally more expensive than DFT with (semi)local functionals, because the Fock exchange is a nonlocal integral operator which acts on Kohn-Sham (KS) wave functions. However, there has been recent progress in the development of very efficient techniques, such as the adaptively compressed exchange [10], to speed up such calculations. Another recent suggestion is the meta-GGA functional SCAN (strongly constrained and appropriately normed semilocal density functional) [11], which has shown promising results for many systems [12,13]. Having a marginal increase in the computational cost, DFT with the SCAN functional uses a xc potential which depends on KS wave functions via a kinetic energy density. However, SCAN still contains significant SIE [12].

A popular approach to alleviate SIE in DFT calculations is to use Hubbard extensions to approximate DFT energy functionals [14–16]. The rationale for this is that Hubbard corrections impose piecewise linearity in the energy functional as a function of occupations [17,18] and thus remove SIE in the Hubbard manifold both in extended systems and in molecular

ones [19]. Within this approach, often referred to as DFT + U , the Hubbard correction acts selectively on strongly localized manifolds (of d or f types, typically) through projectors on the corresponding states, while electrons on more delocalized states are treated at the level of approximate DFT. What has made this approach popular is certainly the possibility of achieving significant improvement in the description of systems with localized electrons, while maintaining a combination of simplicity and reduced computational costs. As for other methods based on the Hubbard model, for example, DFT + DMFT [20–22], the effectiveness of the functional depends critically on the value of the effective interaction parameters (i.e., the Hubbard U parameter), which represents the strength of the onsite electron-electron interactions on localized states [19,23,24]. Unfortunately, the value of these parameters is not known *a priori* and it is a common practice in literature to evaluate them semiempirically by fitting various experimental properties (when available), which prevents this method from being fully *ab initio* and from being predictive for novel materials. Most importantly, it is often forgotten that U acts on a Hubbard manifold that is defined in terms of atomic orbitals, typically taken from the atomic calculations used to generate the respective pseudopotentials, that can be constructed with different degrees of oxidation. Hence, these manifolds, and the relative U parameters, are not transferable and one should not consider U as a universal number for a given element or material (see the appendix in Ref. [25]). During the past 30 years, there has been a large effort to develop methods for the first-principles calculation of U . Among these, the constrained DFT (cDFT) approach [17,26–34], the Hartree-Fock-based approaches [35–38], and the constrained random phase approximation (cRPA) approach [39–45] are the most popular.

In this work, we present a method for the calculation of U using the reciprocal space formulation of density-functional perturbation theory (DFPT) [46,47]. Our approach, by construction, is equivalent to the linear-response cDFT (LR-cDFT) approach of Ref. [17]. Broadly, we refer to linear response as to a first-order variation in the electronic spin charge density, occupations, and potentials upon application of a perturbation. This can be done with finite differences, typically using a

supercell [17], or, as shown here, using DFPT. The DFPT approach allows us to recast an isolated perturbation in a supercell as a sum of monochromatic perturbations in its primitive unit cell; these can be computed independently, with a significant reduction of the computational cost [47]. The favorable scaling of the DFPT approach makes it suitable for incorporation in high-throughput materials screening, and the automation of the elaborate postprocessing of the results improves notably the level of user friendliness. Lastly, we note the formal similarity of the present DFPT-based algorithm for the calculation of Hubbard U with that used to evaluate the screening coefficients for Koopmans' corrections [18,48,49].

The paper is organized as follows. In Sec. II, we recall the rotationally invariant formulation of DFT+ U ; Sec. III contains a brief reminder about the linear-response (LR) calculation of U using the LR-cDFT approach of Ref. [17]; in Sec. IV, we present the DFPT approach to compute U ; Sec. V contains technical details of our calculations; Sec. VI highlights several benchmarks of this method, demonstration of the convergence of U in DFPT, and comparison of scaling of LR-cDFT and DFPT; and finally, in Sec. VII, we give our conclusion. For the sake of simplicity and clarity, in this paper we present the formulation of the DFPT method in the framework of norm-conserving pseudopotentials (PPs) and for gapped systems. The generalization of DFPT to ultrasoft (US) PPs [50], projector augmented wave (PAW) method [51], metals, and intersite interactions [52] will be provided in a following paper. Some mathematical details of the DFPT method are discussed in the appendix. Hartree atomic units are used throughout the paper.

II. DFT+ U

In this section, we briefly review the DFT+ U approach [14,15]. The total energy is defined as

$$E_{\text{DFT}+U} = E_{\text{DFT}} + E_U, \quad (1)$$

where E_{DFT} is the DFT total energy of an approximate functional and E_U is a corrective Hubbard energy term. Many suggestions for E_U have been put forward in the literature (see, e.g., Refs. [14–16,53–58]). Here we take the simplified rotationally invariant formulation by Dudarev *et al.* [16] with

$$\begin{aligned} E_U &= \frac{1}{2} \sum_{I\sigma m_1 m_2} U^I (\delta_{m_1 m_2} - n_{m_1 m_2}^{I\sigma}) n_{m_2 m_1}^{I\sigma} \\ &= \frac{1}{2} \sum_{I\sigma} U^I \text{Tr}[(\mathbf{1} - \mathbf{n}^{I\sigma}) \mathbf{n}^{I\sigma}], \end{aligned} \quad (2)$$

where I is the atomic site index, σ is the spin index, m_1 and m_2 are magnetic quantum numbers associated with a specific angular momentum, and U^I is an effective Hubbard parameter, which can be considered as the difference between the spherically averaged on-site Coulomb repulsion and on-site Hund's exchange J^I . The symbol “Tr” indicates the trace (i.e., the sum over the diagonal elements) of the matrix it acts on. The corrective Hubbard energy of Eq. (2) in its practical incarnation within DFT acts as a linearization condition for the total energy functional (as a function of populations $\mathbf{n}^{I\sigma}$ of the localized Hubbard manifold); i.e., it acts as a self-interaction correction to standard DFT [17–19]. In Eq. (2), $n_{m_1 m_2}^{I\sigma}$ are matrices which measure the occupation of localized orbitals

$\varphi_m^I(\mathbf{r}) \equiv \varphi_m^{\gamma(I)}(\mathbf{r} - \mathbf{R}_I)$ centered on the I th atomic site in the cell \mathbf{R}_I [$\gamma(I)$ is the atomic type of the I th atom] [59]. If we consider a periodic insulating crystal, these occupation matrices can be computed as

$$n_{m_1 m_2}^{I\sigma} = \sum_{\mathbf{k}} \sum_v^{N_{\text{occ}}} \langle \psi_{v\mathbf{k}\sigma}^\circ | \hat{P}_{m_2 m_1}^I | \psi_{v\mathbf{k}\sigma}^\circ \rangle, \quad (3)$$

where $\hat{P}_{m_2 m_1}^I$ is the projector on the manifold of localized (atomic) orbitals

$$\hat{P}_{m_2 m_1}^I = |\varphi_{m_2}^I\rangle\langle\varphi_{m_1}^I|. \quad (4)$$

The projector $\hat{P}_{m_2 m_1}^I$ is the key component of Hubbard-corrected functionals, since it defines the Hubbard manifold (typically localized on an atom) upon which the corrections are constructed. Note that the occupation matrices $n_{m_1 m_2}^{I\sigma}$ are real and symmetric with respect to m_1 and m_2 . In Eq. (3) and hereafter, with the superscript \circ we indicate quantities which refer to the unperturbed ground state of the system. In Eq. (3), $N_{\mathbf{k}}$ is the number of \mathbf{k} points in the first Brillouin zone, v is the electronic band index which runs over N_{occ} occupied states, and $\psi_{v\mathbf{k}\sigma}^\circ$ are the ground-state Kohn-Sham (KS) wave functions, which satisfy the orthonormality condition, $\langle \psi_{v\mathbf{k}\sigma}^\circ | \psi_{v'\mathbf{k}'\sigma'}^\circ \rangle = \delta_{vv'} \delta_{\mathbf{k}\mathbf{k}'} \delta_{\sigma\sigma'}$. The normalization in Eq. (3) and in the following is chosen according to definitions in Eqs. (A3), (A4), and (A5) in the Appendix A1. The KS wave functions are determined by solving the KS equations,

$$\hat{H}_\sigma^\circ | \psi_{v\mathbf{k}\sigma}^\circ \rangle = \varepsilon_{v\mathbf{k}\sigma}^\circ | \psi_{v\mathbf{k}\sigma}^\circ \rangle; \quad (5)$$

i.e., they are the eigenvectors of the total Hamiltonian \hat{H}_σ° which reads

$$\hat{H}_\sigma^\circ = \hat{H}_{\text{DFT},\sigma}^\circ + \hat{V}_{\text{Hub},\sigma}^\circ, \quad (6)$$

where

$$\hat{H}_{\text{DFT},\sigma}^\circ = -\frac{1}{2} \nabla^2 + \hat{V}_{\text{NL}}^\circ + \hat{V}_{\text{loc}}^\circ + \hat{V}_{\text{Hxc},\sigma}^\circ, \quad (7)$$

and

$$\hat{V}_{\text{Hub},\sigma}^\circ = \sum_{I m_1 m_2} U^I \left(\frac{\delta_{m_1 m_2}}{2} - n_{m_1 m_2}^{I\sigma} \right) \hat{P}_{m_1 m_2}^I \quad (8)$$

is the Hubbard potential [23]. $\hat{V}_{\text{Hub},\sigma}^\circ$ can be obtained from Eq. (2) by taking a functional derivative of E_U with respect to KS wave functions [and by making use of Eq. (3)]. In Eq. (7), the first term is the kinetic-energy operator, the second and the third terms are operators corresponding to the nonlocal and local parts of the pseudopotential (which represent interactions of electrons with ions), respectively, and the fourth term is the operator representing the Hartree and exchange-correlation (Hxc) potential. The Hxc potential is the sum of the Hartree and xc contributions:

$$V_{\text{Hxc},\sigma}^\circ(\mathbf{r}) = \int_V \frac{\rho^\circ(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' + V_{\text{xc},\sigma}^\circ(\mathbf{r}), \quad (9)$$

where V is the volume of the crystal, and the xc potential is defined as the functional derivative of the xc energy $E_{\text{xc}}[\rho_\sigma]$ with respect to the spin charge density ρ_σ :

$$V_{\text{xc},\sigma}^\circ(\mathbf{r}) = \left. \frac{\delta E_{\text{xc}}}{\delta \rho_\sigma(\mathbf{r})} \right|_{\rho_\sigma(\mathbf{r}) = \rho_\sigma^\circ(\mathbf{r})}. \quad (10)$$

For the ground state, the spin charge density reads

$$\rho_\sigma^\circ(\mathbf{r}) = \sum_{\mathbf{k}} \sum_v^{N_{\mathbf{k}}} |\psi_{v\mathbf{k}\sigma}^\circ(\mathbf{r})|^2, \quad (11)$$

and the total charge density results from the sum over spins, $\rho^\circ(\mathbf{r}) = \sum_\sigma \rho_\sigma^\circ(\mathbf{r})$.

III. HUBBARD U FROM LINEAR-RESPONSE CONSTRAINED DFT

As already mentioned in the introduction, in spite of the fact that the semiempirical fitting of the Hubbard parameters is still a common practice, their evaluation from *first principles* is necessary to remove any empiricism, and hence many approaches have been proposed [17,26–45]. However, a thorough comparative analysis of their formulation and performance has been rarely and partially attempted in the literature (see, e.g., Ref. [42]). It is our belief that the completion of this task should be based on a detailed comparison of screening mechanisms in calculations of the effective Hubbard parameters.

In this section, we recall basic aspects of the LR-cDFT approach of Ref. [17] for calculations of the effective Hubbard parameter U . In fact, the approach presented in this work is a DFPT-based reformulation of LR-cDFT whose results will thus represent a useful reference. In LR-cDFT, U is computed as the second derivative of the total energy of the system with respect to atomic occupations that are defined as the “on-site” trace of the occupation matrices, $n^I = \sum_\sigma \text{Tr}[\mathbf{n}^{I\sigma}] = \sum_{\sigma,m} n_{mm}^{I\sigma}$ [see Eq. (3)]. In implementations of DFT, like the present one [60,61], where a plane-wave basis set and pseudopotentials are used, atomic occupations are obtained as output quantities (i.e., constructed from the solution of the KS equations) and cannot be controlled from input. To overcome this problem, i.e., to control ground-state occupations as independent variables, the LR-cDFT method performs a Legendre transformation of the total energy. In practice, the total energy functional $E_{\text{DFT}}[\rho_\sigma]$ is augmented with a linear combination of the products of atomic occupations n^I and Lagrange multipliers $\{\lambda^I\}$. Subsequently, this functional is minimized with respect to the spin charge densities ρ_σ , which gives the ground-state energy as a function of the coefficients $\{\lambda^I\}$:

$$E(\{\lambda^I\}) = \min_{\rho_\sigma} \left\{ E_{\text{DFT}}[\rho_\sigma] + \sum_I \lambda^I n^I \right\}, \quad (12)$$

from which the total energy as a function of the ground-state atomic occupations n^I is recovered by a Legendre transformation:

$$\bar{E}(\{n^I\}) = E(\{\lambda^I\}) - \sum_I \lambda^I n^I. \quad (13)$$

This definition of the total energy [Eq. (13)] can be used to evaluate derivatives with respect to n^I . Based on these definitions, the first and second derivatives give

$$\frac{d\bar{E}}{dn^I} = -\lambda^I, \quad (14)$$

$$\frac{d^2\bar{E}}{d(n^I)^2} = -\frac{d\lambda^I}{dn^I} = -(\chi^{-1})_{II}, \quad (15)$$

where the last equality defines the response matrix χ . This matrix contains the response of atomic occupations to the potential shift used to perturb the ground state of the system [see Eq. (12)]: $\chi_{IJ} = dn^I/d\lambda^J$ (I and J being atomic indices). From linear-response theory, as detailed in Ref. [17], the Hubbard U parameter is defined as

$$U^I = (\chi_0^{-1} - \chi^{-1})_{II}, \quad (16)$$

where χ_0 is the noninteracting analog of χ . Equation (16) expresses U as the difference between two total-energy second derivatives [see Eq. (15)]: the one obtained at self-consistency ($-\chi^{-1}$) and the one evaluated at the first iteration of the perturbed run ($-\chi_0^{-1}$), meaning the curvature of the total energy of the noninteracting KS system. The subtraction of the latter term is motivated by the fact that it is not related to electron-electron interactions. According to a complementary view, Eq. (16) can be seen as the solution of the Dyson equation for χ and χ_0 , where U acts as the interaction kernel ($\chi = \chi_0 + \chi_0 U \chi$). In practice, in order to compute the variations of atomic occupations, one needs to solve modified KS equations where atomic potentials are perturbed one at a time [cf. with Eq. (5)]:

$$(\hat{H}_\sigma + \lambda^J \hat{V}_{\text{pert}}^J) |\psi_{v\mathbf{k}\sigma}\rangle = \varepsilon_{v\mathbf{k}\sigma} |\psi_{v\mathbf{k}\sigma}\rangle. \quad (17)$$

These equations can be obtained by taking a functional derivative of $E_{\text{DFT}}[\rho_\sigma] + \sum_I \lambda^I n^I$ [see Eq. (12)] with respect to $\psi_{v\mathbf{k}\sigma}^*(\mathbf{r})$. Based on Eq. (3), it is easy to show that the perturbing potential is

$$\hat{V}_{\text{pert}}^J = \sum_m \hat{P}_{mm}^J, \quad (18)$$

which is a sum of the projections on the localized atomic orbitals of the J th atom [see Eq. (4)]. In Eq. (17), \hat{H}_σ , $\psi_{v\mathbf{k}\sigma}$, and $\varepsilon_{v\mathbf{k}\sigma}$ are the Hamiltonian, the KS wave functions, and the KS energies of the perturbed system, respectively. For the inverse response matrices to correspond to atomic on-site interactions, it is necessary to make sure that these localized, neutral (i.e., not charged) perturbations do not interact with their images through the periodic-boundary conditions, which requires the use of progressively larger supercells. The final results (i.e., the effective Hubbard parameters) thus have to be converged with respect to the supercell’s size.

In practice, after solving Eq. (17) for at least two values of λ^J (centered around 0 eV, which corresponds to the unperturbed ground state), the response matrices can be easily computed using finite differences:

$$\chi_{IJ} = \frac{\Delta n^I}{\Delta \lambda^J}. \quad (19)$$

The variation in the strength λ^J of the perturbing potential should be small, for the response of the system to be in the linear regime, but large enough for the response to be significant with respect to numerical noise. Then, the full response matrix χ is constructed by perturbing one at a time the crystallographically inequivalent Hubbard manifolds of the individual Hubbard atoms. When one Hubbard atom or manifold is perturbed, the response of all the atoms is recorded, so that one column of the response matrix χ is determined. The other matrix elements (in the columns corresponding to equivalent Hubbard atoms)

are then reconstructed by symmetry. The dimension of the response matrix χ is $N_H^{\text{sc}} \times N_H^{\text{sc}}$, where N_H^{sc} is the number of Hubbard atoms in the supercell. We also point out that the change in the occupations of the non-Hubbard atoms (not explicitly included in the response matrices) contributes to the screening of the effective interaction.

The procedure described above applies to the construction of both χ and χ_0 . However, χ is computed after the self-consistent solution of Eq. (17), while χ_0 , the KS (bare) response, is computed at the first iteration, where the Hxc potential is equal to that of the unperturbed system. It is worth noting that the off-diagonal matrix elements in $(\chi_0^{-1} - \chi^{-1})_{IJ}$ represent *intersite* interactions, which are not included in the standard, on-site only DFT + U Hamiltonian. These can be important when the hybridization between strongly localized and nonlocalized states occurs. These terms are included in the so-called DFT + U + V functional [52] to which the DFPT-based calculation of Hubbard interaction parameters introduced in this work has been also extended, as will be presented elsewhere.

It is important to note that U can be computed starting from a DFT + U ground state ($U \neq 0$). This is particularly useful when U is evaluated self-consistently [62]. In fact, this goal is achieved through a series of linear-response evaluations of the Hubbard parameters based on the DFT + U ground state obtained with the U determined in the previous step; the cycle is stopped when the computed value of U is within a fixed threshold from that of the previous step. When evaluating U from a DFT + U ground state, all the equations discussed above apply after substituting E_{DFT} with $E_{\text{DFT}+U}$ in Eq. (12), and with the extra condition that the Hubbard potential in the Hamiltonian \hat{H}_σ , appearing in Eq. (17), is kept fixed to its unperturbed ground-state value:

$$\hat{H}_\sigma = \hat{H}_{\text{DFT},\sigma} + \hat{V}_{\text{Hub},\sigma}^\circ. \quad (20)$$

This is a consequence of the fact that U is defined as the curvature of the DFT part of the total energy [first term in Eq. (1)] as a function of atomic occupations [17]. Keeping $\hat{V}_{\text{Hub},\sigma}$ (the first derivative of the Hubbard energy) fixed to its unperturbed value $\hat{V}_{\text{Hub},\sigma}^\circ$ avoids the problem where the second derivative of the total energy [see Eq. (15)] contains a finite contribution from the Hubbard correction. We note that in this work we will not compute U self-consistently and we perform only one-shot calculations starting from the DFT ground state with $U = 0$.

The LR-cDFT approach is straightforward and easy to implement in electronic-structure codes. However, it is computationally expensive because of its cubic scaling with the size of the supercell. Furthermore, this approach requires us to perform a number of supercell calculations (depending on how many nonequivalent Hubbard atoms are in the system) and convergence tests with respect to the supercell size to avoid unphysical interactions between periodic images of the perturbed atoms. Even though some extrapolation techniques have been devised to speed up the convergence of U with respect to the supercell size [17], more computationally efficient algorithms are highly desirable to overcome these difficulties, especially in view of their possible deployment in highly automatized, high-throughput calculations.

IV. HUBBARD U FROM DENSITY-FUNCTIONAL PERTURBATION THEORY

In this section, we show how the LR-cDFT calculation of the effective Hubbard parameters outlined in Sec. III can be reformulated within the framework of DFPT, i.e., solving self-consistently the system of linear equations that can be obtained from perturbing to first-order Kohn-Sham equations around the ground state of the system [46,47]. Although formally an all-order perturbation theory, the limitation to first order in the perturbations makes standard DFPT a linear-response approach, and thus equivalent to LR-cDFT. It is important to remark that DFPT allows for a reciprocal space formulation of the LR problem, with wavelength-specific (i.e., momentum-specific) perturbations and responses [47]. As will be shown later, this presents the same advantages that a DFPT phonon calculation [47] has with respect to a finite-difference approach with ionic displacements frozen in the crystal structure (frozen phonons): The possibility of performing calculations in primitive unit cells (thus avoiding costly supercells), computational costs essentially uniform across the Brillouin zone, exploitation of symmetries, and a higher level of automation. The transition from LR-cDFT to DFPT will be made in two steps: first, substituting finite differences with continuous derivatives, and second, recasting perturbative calculations in supercells as series of monochromatic perturbations in primitive unit cells.

A. Hubbard U from DFPT in real space

In this section, we will show how to compute the noninteracting and interacting response matrices χ_0 and χ which are needed for the calculation of U [see Eq. (16)]. With the DFPT formalism in a “real-space” implementation, localized perturbations are applied in the same supercells used in the LR-cDFT approach, so the only difference is that the LR matrices are continuous derivatives of atomic occupations,

$$\chi_{IJ} = \sum_{\sigma,m} \frac{dn_{mm}^{I\sigma}}{d\lambda^J}, \quad (21)$$

which, based on Eq. (3), can be constructed from the LR KS wave functions directly obtained from DFPT:

$$\begin{aligned} \frac{dn_{m_1 m_2}^{I\sigma}}{d\lambda^J} &= \sum_{\mathbf{k}} \sum_v^{N_{\text{occ}}} \left[\langle \psi_{v\mathbf{k}\sigma}^\circ | \hat{P}_{m_2 m_1}^I \left| \frac{d\psi_{v\mathbf{k}\sigma}}{d\lambda^J} \right. \right] \\ &+ \left\langle \frac{d\psi_{v\mathbf{k}\sigma}}{d\lambda^J} \left| \hat{P}_{m_2 m_1}^I | \psi_{v\mathbf{k}\sigma}^\circ \right. \right\rangle. \end{aligned} \quad (22)$$

Here, J and I label the atomic sites where the perturbation is applied and whose change in occupation is being measured, respectively. Within DFPT, the LR KS wave functions $\frac{d\psi_{v\mathbf{k}\sigma}(\mathbf{r})}{d\lambda^J}$ are the solutions of the linear-response KS equations (from first-order perturbation theory):

$$\begin{aligned} &(\hat{H}_\sigma^\circ - \varepsilon_{v\mathbf{k}\sigma}^\circ) \left| \frac{d\psi_{v\mathbf{k}\sigma}}{d\lambda^J} \right\rangle \\ &= - \left(\frac{d\hat{V}_{\text{Hxc},\sigma}}{d\lambda^J} - \frac{d\varepsilon_{v\mathbf{k}\sigma}}{d\lambda^J} + \hat{V}_{\text{pert}}^J \right) | \psi_{v\mathbf{k}\sigma}^\circ \rangle. \end{aligned} \quad (23)$$

Equation (23) is obtained from Eq. (17) by a Taylor expansion in λ^J to first order of all quantities. Here, \hat{H}_σ° , $\varepsilon_{v\mathbf{k}\sigma}^\circ$, $\psi_{v\mathbf{k}\sigma}^\circ$

are, respectively, the total Hamiltonian [see Eq. (6)], the KS energies, and the KS wave functions of the system in its unperturbed ground state, while \hat{V}_{pert}^J , $\frac{d\varepsilon_{v\mathbf{k}\sigma}}{d\lambda^J}$, and $\frac{d\hat{V}_{\text{Hxc},\sigma}}{d\lambda^J}$ are the perturbing potential [see Eq. (18)] and the first-order variations of the KS energies and Hxc potential. From Eqs. (9) and (10), it is easy to see that in the coordinate representation the latter quantity can be expressed as

$$\begin{aligned} \frac{dV_{\text{Hxc},\sigma}(\mathbf{r})}{d\lambda^J} &= \int_{\mathcal{V}} \frac{1}{|\mathbf{r}-\mathbf{r}'|} \frac{d\rho(\mathbf{r}')}{d\lambda^J} d\mathbf{r}' \\ &+ \sum_{\sigma'} \int_{\mathcal{V}} \kappa_{\text{xc},\sigma\sigma'}(\mathbf{r},\mathbf{r}') \frac{d\rho_{\sigma'}(\mathbf{r}')}{d\lambda^J} d\mathbf{r}', \end{aligned} \quad (24)$$

where the first and the second terms are the LR Hartree and xc potentials, respectively, and $\kappa_{\text{xc},\sigma\sigma'}(\mathbf{r},\mathbf{r}')$ is the xc kernel

$$\kappa_{\text{xc},\sigma\sigma'}(\mathbf{r},\mathbf{r}') = \frac{\delta^2 E_{\text{xc}}}{\delta\rho_{\sigma}(\mathbf{r})\delta\rho_{\sigma'}(\mathbf{r}')} \Bigg|_{\substack{\rho_{\sigma}(\mathbf{r})=\rho_{\sigma}^{\circ}(\mathbf{r}) \\ \rho_{\sigma'}(\mathbf{r}')=\rho_{\sigma'}^{\circ}(\mathbf{r}')}}, \quad (25)$$

which is local in the local-spin-density approximation (or semilocal in the generalized-gradient approximation) [63]. In Eq. (24), $\frac{d\rho(\mathbf{r})}{d\lambda^J} = \sum_{\sigma} \frac{d\rho_{\sigma}(\mathbf{r})}{d\lambda^J}$ is the total LR charge density, where

$$\frac{d\rho_{\sigma}(\mathbf{r})}{d\lambda^J} = 2 \text{Re} \left\{ \sum_{\mathbf{k}} \sum_{v}^{N_{\text{occ}}} \psi_{v\mathbf{k}\sigma}^{\circ*}(\mathbf{r}) \frac{d\psi_{v\mathbf{k}\sigma}(\mathbf{r})}{d\lambda^J} \right\}. \quad (26)$$

In the last equation, the absence of the imaginary component is a consequence of time-reversal symmetry. It is important to note that in Eq. (23) there are no other first-order terms, because the localized orbitals $\varphi_m^J(\mathbf{r})$ are a fixed basis set. DFPT involves solving the system of linearly coupled equations (23)–(26), iteratively and self-consistently.

If the LR KS equations (23) are solved starting from the DFT+ U ground state with $U \neq 0$, then the response of the Hubbard potential $\frac{d\hat{V}_{\text{Hub},\sigma}}{d\lambda^J}$ must be set to zero, for the same reason discussed in Sec. III after Eq. (20). As a consequence, this term is not present on the right-hand side of Eq. (23).

It is well known in DFPT [47] that only the component of the LR KS wave functions obtained by projection on the empty (conduction) states manifold gives a nonzero contribution to the LR spin charge density (26) and LR occupation matrices (22). Therefore, it is convenient to work directly with the projection of Eq. (23) onto this manifold:

$$\begin{aligned} &(\hat{H}_{\sigma}^{\circ} + \alpha\hat{O}_{\sigma} - \varepsilon_{v\mathbf{k}\sigma}^{\circ}) \left| \frac{d\tilde{\psi}_{v\mathbf{k}\sigma}}{d\lambda^J} \right\rangle \\ &= -\hat{\mathcal{P}}_{\sigma} \left(\frac{d\hat{V}_{\text{Hxc},\sigma}}{d\lambda^J} + \hat{V}_{\text{pert}}^J \right) |\psi_{v\mathbf{k}\sigma}^{\circ}\rangle. \end{aligned} \quad (27)$$

In Eq. (27), \hat{O}_{σ} is the projector on the occupied states manifold, $\hat{O}_{\sigma} = \sum_{\mathbf{k}'} \sum_{v'}^{N_{\text{occ}}} |\psi_{v'\mathbf{k}'\sigma}^{\circ}\rangle \langle \psi_{v'\mathbf{k}'\sigma}^{\circ}|$, while the operator $\hat{\mathcal{P}}_{\sigma}$ is the projector on the empty states manifold, and it is expressed using the identity relation $\hat{\mathcal{P}}_{\sigma} = 1 - \hat{O}_{\sigma}$, in order to avoid sums over empty states that would typically be very slow converging with respect to the number of states [47]. In Eq. (27), $|\frac{d\tilde{\psi}_{v\mathbf{k}\sigma}}{d\lambda^J}\rangle \equiv \hat{\mathcal{P}}_{\sigma} |\frac{d\psi_{v\mathbf{k}\sigma}}{d\lambda^J}\rangle$ is the conduction component of the LR KS wave functions. The extra term $\alpha\hat{O}_{\sigma}$ on the left-hand side of Eq. (27) is not strictly required by the

projection on conduction states: Its presence serves the purpose of lifting the singularity of the operator $\hat{H}_{\sigma}^{\circ} - \varepsilon_{v\mathbf{k}\sigma}^{\circ}$ on the valence manifold, which would create numerical issues when solving it. For this purpose, the parameter α is fixed to twice the spread of the unperturbed Kohn-Sham spectrum, $\alpha = 2(\max[\varepsilon_{v\mathbf{k}\sigma}^{\circ}] - \min[\varepsilon_{v\mathbf{k}\sigma}^{\circ}])$, which makes the operator on the left-hand side of Eq. (27) nonsingular. In practice, if Eq. (27) is solved using iterative algorithms [64] and the trial solution is chosen orthogonal to the occupied states manifold, then the orthogonality is preserved during the iteration cycle and there is no need of the extra term $\alpha\hat{O}_{\sigma}$ on the left-hand side of Eq. (27) [47]. It is worth noting that the term $\frac{d\varepsilon_{v\mathbf{k}\sigma}}{d\lambda^J}$ appearing in Eq. (23) has disappeared from Eq. (27) because $\hat{\mathcal{P}}_{\sigma} |\psi_{v\mathbf{k}\sigma}^{\circ}\rangle = 0$ for valence states; this means that variations of the valence KS energies (due to the perturbation) do not contribute to the conduction component of the LR KS wave functions. The expressions for the LR occupation matrices and spin charge densities [see Eqs. (22) and (26), respectively] remain identical, if we replace $\frac{d\psi_{v\mathbf{k}\sigma}}{d\lambda^J}$ by $\frac{d\tilde{\psi}_{v\mathbf{k}\sigma}}{d\lambda^J}$ [65].

The DFPT method presented to this point, i.e., solving Eq. (27) and computing the LR occupation matrices using Eq. (22), is, by construction, exactly equivalent to the LR-cDFT method, when used on the same supercell, but offers a more automatic and elegant way of computing the Hubbard parameters from linear response, avoids the need for finite differences, and instead computes response quantities directly as analytical derivatives, with a higher control of accuracy. However, the formulation of DFPT discussed so far presents no computational advantages with respect to LR-cDFT, because most of the computational effort is due to the use of supercells. This latter aspect is, in fact, unavoidable within LR-cDFT. Within DFPT, instead, perturbations in supercells can be recast as sums over monochromatic perturbations in primitive unit cells which can be computed independently, thus leading to a significant reduction of the computational cost to calculate U , as will be discussed in detail in Sec. IV B.

B. Hubbard U from DFPT in reciprocal space

In this section, we show how the response of a system to isolated perturbations can be recast in terms of monochromatic perturbations using the DFPT formalism outlined in Sec. IV A. These perturbations can be treated within primitive unit cells and thus present a computational cost which is substantially independent of the wavelength of the perturbation considered, $\lambda = 2\pi/|\mathbf{q}|$ (differences in these costs are still possible depending on the symmetry of the perturbation, i.e., the small group of \mathbf{q}). The first part of this section will be dedicated to explain the general idea behind this reformulation; the second one will instead illustrate all the technical details of the implementation in DFPT.

In the following, we will consider for simplicity only supercells whose lattice vectors are integer multiples of those of the primitive unit cell. For supercells with lattice vectors not parallel to those of the primitive unit cell, a conventional unit cell can be usually identified (whose volume is typically a small integer multiple of that of the primitive unit cell) that can be used for DFPT calculations with a relatively small increase in the computational cost.

1. From localized perturbations in supercells to monochromatic perturbations in primitive unit cells

A supercell of size $L_1 \times L_2 \times L_3$ is defined here, having lattice vectors $\{\mathbf{A}_i\}$ that are integer multiples of those of the primitive unit cell $\{\mathbf{a}_i\}$,

$$\mathbf{A}_i = L_i \mathbf{a}_i, \quad (28)$$

where $i = 1, 2, 3$. Let $\{\mathbf{B}_i\}$ be the reciprocal lattice basis vectors of this supercell. Based on their definition, it is easy to determine the relationship with the reciprocal lattice vectors $\{\mathbf{b}_i\}$ of the primitive unit cell,

$$\mathbf{B}_i = 2\pi \frac{\mathbf{A}_j \times \mathbf{A}_k}{V} = 2\pi \frac{\mathbf{a}_j \times \mathbf{a}_k}{L_i v} = \frac{\mathbf{b}_i}{L_i}, \quad (29)$$

where we have used the fact that $V = L_1 L_2 L_3 v$, with v being the volume of the primitive unit cell. Based on Eq. (29), the reciprocal lattice vectors of the supercell

$$\mathbf{G}_{klm} = k \mathbf{B}_1 + l \mathbf{B}_2 + m \mathbf{B}_3 \quad (30)$$

can be rewritten as

$$\mathbf{G}_{klm} = \frac{k}{L_1} \mathbf{b}_1 + \frac{l}{L_2} \mathbf{b}_2 + \frac{m}{L_3} \mathbf{b}_3, \quad (31)$$

where k, l , and m are integer numbers, whose range is fixed by the kinetic-energy cutoff. We now define the quotients k', l' , and m' and the remainders \bar{k}, \bar{l} , and \bar{m} of the fractional coefficients $k/L_1, l/L_2, m/L_3$ in Eq. (31) as

$$k = k' L_1 + \bar{k}, \quad (32)$$

$$l = l' L_2 + \bar{l}, \quad (33)$$

$$m = m' L_3 + \bar{m}. \quad (34)$$

With these definitions, we can rewrite any \mathbf{G}_{klm} as follows:

$$\mathbf{G}_{klm} = \mathbf{g}_{k'l'm'} + \mathbf{q}_{\bar{k}\bar{l}\bar{m}}, \quad (35)$$

where $\mathbf{g}_{k'l'm'}$ are reciprocal lattice vectors for the primitive unit cell, i.e.,

$$\mathbf{g}_{k'l'm'} = k' \mathbf{b}_1 + l' \mathbf{b}_2 + m' \mathbf{b}_3, \quad (36)$$

and $\mathbf{q}_{\bar{k}\bar{l}\bar{m}}$ are vectors residing inside the first Brillouin zone of the primitive unit cell, i.e.,

$$\mathbf{q}_{\bar{k}\bar{l}\bar{m}} = \frac{\bar{k}}{L_1} \mathbf{b}_1 + \frac{\bar{l}}{L_2} \mathbf{b}_2 + \frac{\bar{m}}{L_3} \mathbf{b}_3, \quad (37)$$

where $0 \leq \bar{k} < L_1$, $0 \leq \bar{l} < L_2$, and $0 \leq \bar{m} < L_3$. Based on Eq. (29), it can be seen from Eq. (37) that $\mathbf{q}_{\bar{k}\bar{l}\bar{m}}$ are, in fact, reciprocal lattice vectors of the supercell.

In LR-cDFT, an isolated perturbation is achieved by choosing a sufficiently large supercell. The Fourier expansion of the corresponding perturbing potential thus contains only reciprocal lattice vectors of the supercell:

$$\begin{aligned} V(\mathbf{r}) &= \sum_{\mathbf{G}} e^{i\mathbf{G}\cdot\mathbf{r}} V(\mathbf{G}) \\ &\equiv \sum_{klm} e^{i\mathbf{G}_{klm}\cdot\mathbf{r}} V(\mathbf{G}_{klm}). \end{aligned} \quad (38)$$

Using Eq. (35), we can rewrite the potential of Eq. (38) as a sum of monochromatic perturbations whose wave vectors \mathbf{q} are defined in Eq. (37):

$$\begin{aligned} V(\mathbf{r}) &= \sum_{\bar{k}\bar{l}\bar{m}} \sum_{k'l'm'} e^{i(\mathbf{g}_{k'l'm'} + \mathbf{q}_{\bar{k}\bar{l}\bar{m}})\cdot\mathbf{r}} V(\mathbf{g}_{k'l'm'} + \mathbf{q}_{\bar{k}\bar{l}\bar{m}}) \\ &\equiv \sum_{\mathbf{q}} \sum_{\mathbf{g}} e^{i(\mathbf{g} + \mathbf{q})\cdot\mathbf{r}} V(\mathbf{g} + \mathbf{q}) \\ &= \sum_{\mathbf{q}} e^{i\mathbf{q}\cdot\mathbf{r}} \bar{V}_{\mathbf{q}}(\mathbf{r}). \end{aligned} \quad (39)$$

Equation (39) defines the lattice-periodic potential $\bar{V}_{\mathbf{q}}(\mathbf{r})$ as

$$\bar{V}_{\mathbf{q}}(\mathbf{r}) = \sum_{\mathbf{g}} e^{i\mathbf{g}\cdot\mathbf{r}} V(\mathbf{g} + \mathbf{q}). \quad (40)$$

The simple derivation above shows that the response to a perturbation with the periodicity of a supercell can be equivalently computed in the primitive unit cell as the sum of the responses to monochromatic perturbations on a grid of \mathbf{q} points defined by Eq. (37). It is important to stress that the size of the \mathbf{q} points grid is determined by the size of the supercell. In Sec. IV B 2, we show how this can be implemented in DFPT.

2. Monochromatic perturbations in DFPT

Let us now proceed with the reformulation of DFPT in terms of monochromatic perturbations in a primitive unit cell (this setting will define the meaning of ‘‘DFPT calculations’’ from this point on). For this purpose, it is necessary to re-express localized, but supercell-periodic, responses as sums of monochromatic contributions, as already done for the perturbing potential in Sec. IV B 1, and show that they can be computed one by one, solving \mathbf{q} -specific first-order perturbative equations.

Let us consider a supercell of size $L_1 \times L_2 \times L_3$ [see Sec. IV B 1] whose first Brillouin zone is sampled by $N_{\mathbf{k}}^{\text{sc}}$ \mathbf{k} points. Based on the discussion in Sec. IV B 1, the size of the supercell determines the grid of \mathbf{q} points to be used, which is $L_1 \times L_2 \times L_3$ (hence the number of \mathbf{q} points is $N_{\mathbf{q}} = L_1 L_2 L_3$). It is also important to note that the DFPT calculation requires $N_{\mathbf{k}} = N_{\mathbf{k}}^{\text{sc}} N_{\mathbf{q}}$ \mathbf{k} points to sample the first Brillouin zone of the primitive unit cell with the same accuracy. If \mathbf{R}_I describes the position of the I th atom in the supercell, then this can be represented as $\mathbf{R}_I = \mathbf{R}_l + \boldsymbol{\tau}_s$, where \mathbf{R}_l is the Bravais lattice vector of the l th primitive unit cell and $\boldsymbol{\tau}_s$ is the position of the s th atom in the l th primitive unit cell. Hence, each index I corresponds to two indices (l, s) . By taking $I = (l, s)$ and $J = (l', s')$, the interacting response matrices (21) can be written as [66]

$$\chi_{sl, s'l'} = \sum_{\sigma, m} \frac{d n_{mm}^{sl\sigma}}{d \lambda^{s'l'}}, \quad (41)$$

and a similar expression is used for the noninteracting response matrices χ_0 . By performing mathematical manipulations as explained in the appendix (a bar over the symbol of a quantity indicates its lattice-periodic part), it can be shown that the LR

occupation matrices, Eq. (22), can be expressed as

$$\frac{dn_{m_1 m_2}^{s\sigma}}{d\lambda^{s'l'}} = \frac{1}{N_{\mathbf{q}}} \sum_{\mathbf{q}} e^{i\mathbf{q}\cdot(\mathbf{R}_l - \mathbf{R}_{l'})} \Delta_{\mathbf{q}}^{s'} \bar{n}_{m_1 m_2}^{s\sigma}, \quad (42)$$

where $\Delta_{\mathbf{q}}^{s'} \bar{n}_{m_1 m_2}^{s\sigma}$ is the lattice-periodic response of the atomic occupation to a monochromatic perturbation of wave vector \mathbf{q} [67]:

$$\begin{aligned} \Delta_{\mathbf{q}}^{s'} \bar{n}_{m_1 m_2}^{s\sigma} &= \frac{1}{N_{\mathbf{k}}} \sum_{\mathbf{k}} \sum_v^{N_{\text{occ}}} \left[\langle \bar{u}_{v\mathbf{k}\sigma}^{\circ} | \hat{P}_{m_2, m_1, \mathbf{k}, \mathbf{k}+\mathbf{q}}^s | \Delta_{\mathbf{q}}^{s'} \bar{u}_{v\mathbf{k}\sigma} \rangle \right. \\ &\quad \left. + \langle \bar{u}_{v\mathbf{k}\sigma}^{\circ} | \hat{P}_{m_1, m_2, \mathbf{k}, \mathbf{k}+\mathbf{q}}^s | \Delta_{\mathbf{q}}^{s'} \bar{u}_{v\mathbf{k}\sigma} \rangle \right], \quad (43) \end{aligned}$$

N_{occ} being the number of occupied states in the primitive unit cell. Here, $\bar{u}_{v\mathbf{k}\sigma}^{\circ}(\mathbf{r})$ and $\Delta_{\mathbf{q}}^{s'} \bar{u}_{v\mathbf{k}\sigma}(\mathbf{r})$ are the lattice-periodic parts of the ground-state and LR KS wave functions, respectively, while $\hat{P}_{m_2, m_1, \mathbf{k}, \mathbf{k}+\mathbf{q}}^s$ is the lattice-periodic projector operator

$$\hat{P}_{m_2, m_1, \mathbf{k}, \mathbf{k}+\mathbf{q}}^s = |\bar{\varphi}_{m_2, \mathbf{k}}^s \rangle \langle \bar{\varphi}_{m_1, \mathbf{k}+\mathbf{q}}^s|, \quad (44)$$

constructed on the lattice-periodic components of Bloch sums of localized functions $\varphi_m^s(\mathbf{r} - \mathbf{R}_l)$ (see the appendix for more details). Equation (43) shows that the quantities $\Delta_{\mathbf{q}}^{s'} \bar{n}_{m_1 m_2}^{s\sigma}$ can be computed by knowing $\Delta_{\mathbf{q}}^{s'} \bar{u}_{v\mathbf{k}\sigma}(\mathbf{r})$. Since the perturbative problem is expanded to first order, perturbations at different wavelengths do not interact with each other, and it is possible to show that the responses $\Delta_{\mathbf{q}}^{s'} \bar{u}_{v\mathbf{k}\sigma}(\mathbf{r})$ can be obtained directly from the self-consistent solution of Eq. (27) specialized to single lattice-periodic \mathbf{q} -specific (i.e., monochromatic) perturbations [68]:

$$\begin{aligned} &(\hat{H}_{\mathbf{k}+\mathbf{q}, \sigma}^{\circ} + \alpha \hat{O}_{\mathbf{k}+\mathbf{q}, \sigma} - \varepsilon_{v\mathbf{k}\sigma}^{\circ}) | \Delta_{\mathbf{q}}^{s'} \bar{u}_{v\mathbf{k}\sigma} \rangle \\ &= -\hat{P}_{\mathbf{k}+\mathbf{q}, \sigma}^{\circ} (\Delta_{\mathbf{q}}^{s'} \hat{V}_{\text{Hxc}, \sigma} + \hat{V}_{\text{pert}, \mathbf{k}+\mathbf{q}, \mathbf{k}}^{s'}) | \bar{u}_{v\mathbf{k}\sigma}^{\circ} \rangle. \quad (45) \end{aligned}$$

Here $\hat{H}_{\mathbf{k}+\mathbf{q}, \sigma}^{\circ}$ is the lattice-periodic part of the total ground-state Hamiltonian of the system [see Eq. (A13)], and $\hat{V}_{\text{pert}, \mathbf{k}+\mathbf{q}, \mathbf{k}}^{s'}$ is the monochromatic \mathbf{q} component of the perturbing potential:

$$\hat{V}_{\text{pert}, \mathbf{k}+\mathbf{q}, \mathbf{k}}^{s'} = \sum_m \hat{P}_{m, m, \mathbf{k}+\mathbf{q}, \mathbf{k}}^{s'}, \quad (46)$$

with $\hat{P}_{m, m, \mathbf{k}+\mathbf{q}, \mathbf{k}}^{s'}$ defined in Eq. (44). $\hat{O}_{\mathbf{k}+\mathbf{q}, \sigma}$ and $\hat{P}_{\mathbf{k}+\mathbf{q}, \sigma}^{\circ}$ are the lattice-periodic parts of the projectors on the occupied- and empty-state manifolds, respectively, and $\Delta_{\mathbf{q}}^{s'} \bar{V}_{\text{Hxc}, \sigma}(\mathbf{r})$ is the lattice-periodic part of the LR Hxc potential [see Eq. (24)]:

$$\begin{aligned} \Delta_{\mathbf{q}}^{s'} \bar{V}_{\text{Hxc}, \sigma}(\mathbf{r}) &= \int_V \frac{1}{|\mathbf{r} - \mathbf{r}'|} e^{-i\mathbf{q}\cdot(\mathbf{r} - \mathbf{r}')} \Delta_{\mathbf{q}}^{s'} \bar{\rho}(\mathbf{r}') d\mathbf{r}' \\ &\quad + \sum_{\sigma'} \int_V \kappa_{\text{xc}, \sigma\sigma'}(\mathbf{r}, \mathbf{r}') e^{-i\mathbf{q}\cdot(\mathbf{r} - \mathbf{r}')} \Delta_{\mathbf{q}}^{s'} \bar{\rho}_{\sigma'}(\mathbf{r}') d\mathbf{r}', \quad (47) \end{aligned}$$

where $\Delta_{\mathbf{q}}^{s'} \bar{\rho}(\mathbf{r}) = \sum_{\sigma} \Delta_{\mathbf{q}}^{s'} \bar{\rho}_{\sigma}(\mathbf{r})$ is the total LR lattice-periodic charge density, and $\Delta_{\mathbf{q}}^{s'} \bar{\rho}_{\sigma}(\mathbf{r})$ is the lattice-periodic part of the LR spin charge density, which is

$$\Delta_{\mathbf{q}}^{s'} \bar{\rho}_{\sigma}(\mathbf{r}) = 2 \text{Re} \left\{ \frac{1}{N_{\mathbf{k}}} \sum_{\mathbf{k}} \sum_v^{N_{\text{occ}}} \bar{u}_{v\mathbf{k}\sigma}^{\circ*}(\mathbf{r}) \Delta_{\mathbf{q}}^{s'} \bar{u}_{v\mathbf{k}\sigma}(\mathbf{r}) \right\}. \quad (48)$$

Equations (41)–(48) represent the core results of this work, and their implementation is discussed in the following. It is important to note that Eqs. (45) refer to lattice-periodic functions and operators and that, for a given \mathbf{q} and a given s' , they need to be solved self-consistently and simultaneously for all \mathbf{k} , v , and σ , due to the coupling between them introduced by Eqs. (47) and (48). Equations at different \mathbf{q} and s' are, instead, decoupled from one another and the linear-response problem is decomposed into a set of independent problems that can be solved on separate computational resources, thus allowing for straightforward parallelization.

Besides better scaling (as will be discussed in Sec. VIC), DFPT is also more user-friendly than LR-cDFT, due to the fully automated postprocessing operations required for the calculation of U . As in LR-cDFT (see Sec. III), in DFPT only the crystallographically inequivalent Hubbard atoms of each type are perturbed, and these perturbations can be parallelized; In addition, the identification of the inequivalent Hubbard atoms as based on the symmetry of the system becomes automatic. In DFPT calculations, the \mathbf{q} -specific components of the response occupation matrices $\Delta_{\mathbf{q}}^{s'} \bar{n}_{m_1 m_2}^{s\sigma}$ are computed for all the Hubbard atoms s in the primitive unit cell, responding to perturbations on all the inequivalent Hubbard atoms s' [see Eq. (43)]. The full response occupation matrix is then constructed at a negligible computational cost in a postprocessing step using Eq. (42) and using the symmetry of the system in order to recover any missing elements. These postprocessing steps are fully automated, thus avoiding the cumbersome and system-specific postprocessing operations of the LR-cDFT approach.

In summary, all one needs to do in order to compute Hubbard parameters through DFPT is to solve the system of equations (45) for a grid of \mathbf{q} points which represent the folding of a supercell of desired size [see Eq. (37)] and to calculate the LR occupation matrices by performing the sum over \mathbf{q} components defined in Eq. (42).

V. TECHNICAL DETAILS

The DFPT approach introduced in Sec. IV has been implemented in the QUANTUM ESPRESSO package [60,61] and will be distributed as an open-source module in a future release. Here, we will verify the implementation of DFPT by applying it to the case of bulk NiO, Cu₂O, and LiCoO₂. We note that although we presented the DFPT formalism for the case of norm-conserving pseudopotentials, the implementation with ultrasoft pseudopotentials and PAW has also been completed and fully tested; in fact, the results presented in the following have been obtained from calculations using US PPs. In order to get some insight about the extension of DFPT to US PPs and PAW, the reader is referred to Refs. [69] and [70]. The rather complex technical details for the Hubbard case will be presented in detail in a following paper.

All calculations are performed using the plane-wave (PW) pseudopotential method and the generalized-gradient approximation (GGA) for the xc functional constructed with the PBEsol prescription [71]. US PPs are taken from Pslibrary 0.3.1 [72,73]. KS wave functions and charge-density are expanded in PWs up to a kinetic-energy cutoff of 80 and 640 Ry for NiO and Cu₂O, and 120 and 960 Ry for LiCoO₂, respectively. Convergence tests for U have shown that by using

these cutoffs one obtains U with an accuracy of better than 0.01 eV. We have chosen such high accuracy uniquely for the sake of comparing DFPT and LR-cDFT. The Brillouin zones of the different materials have been sampled with uniform unshifted \mathbf{k} and \mathbf{q} point meshes, i.e., including the Γ point, of different sizes. We have used the experimental lattice parameters $a = 4.17 \text{ \AA}$ for the rock-salt crystal structure of antiferromagnetic NiO [74], $a = 4.27 \text{ \AA}$ for the simple cubic cell of Cu_2O with space group O_h^4 (or $Pn\bar{3}m$) [75], and $a = 4.96 \text{ \AA}$, $\alpha = 32.99^\circ$, $z = 0.24$ (where z is the atomic positional parameter along the trigonal axis) for the rhombohedral cell of LiCoO_2 [76].

Consistently with the implementation of DFT + U in QUANTUM ESPRESSO, the approach presented here is based on atomic orbitals. Other choices for the localized basis set are possible (e.g., Wannier functions); the final values for the Hubbard parameters can be expected to depend on the choice of localized functions, as well as the xc functional and pseudopotentials [34]; notably, pseudopotentials constructed from all-electron atoms in different oxidation states will give rise to different Hubbard manifolds and U parameters [25]. To construct the projectors of Eq. (44) we have used atomic orbitals which were orthogonalized using Löwdin's method [77]. In the LR-cDFT calculations with finite differences, we used $\lambda^J = \pm 0.05 \text{ eV}$ [see Eq. (17)]. The LR KS equations (45) were solved using the conjugate-gradient algorithm [64] and the mixing scheme of Ref. [78] for the response Hxc potential (47) to speed up convergence.

In the following, we present the values of U computed using DFPT and LR-cDFT starting from the GGA ground state [i.e., $U = 0$ in Eq. (8)]. A self-consistent calculation of the Hubbard parameters is straightforwardly possible by iteration [34,79], but it is not pursued here.

VI. RESULTS AND DISCUSSION

A. Validation

Table I shows a comparison between the values of U computed for Cu_2O , NiO, and LiCoO_2 , using the LR-cDFT and DFPT as described in Secs. III and IV, respectively, starting from the GGA (PBEsol) ground state. The two series of calculations are set up equivalently: LR-cDFT calculations are performed with supercells of size $L_1 \times L_2 \times L_3$, and the corresponding reciprocal-space DFPT calculations are performed

TABLE I. Comparison of U (in eV) for Cu in Cu_2O , Ni in NiO, and Co in LiCoO_2 computed using LR-cDFT with supercells (SC) of different size and various \mathbf{k} point meshes, and using DFPT for a primitive unit cell with \mathbf{k} and \mathbf{q} point meshes chosen to be a folded equivalent of the LR-cDFT calculations.

| Method | \mathbf{k} -mesh | SC-size/ \mathbf{q} -mesh | Cu_2O | NiO | LiCoO_2 |
|---------|--------------------------|-----------------------------|-----------------------|-------|------------------|
| LR-cDFT | $6 \times 6 \times 6$ | $2 \times 2 \times 2$ | 11.263 | 7.895 | 7.472 |
| DFPT | $12 \times 12 \times 12$ | | 11.268 | 7.900 | 7.473 |
| LR-cDFT | $4 \times 4 \times 4$ | $3 \times 3 \times 3$ | 11.287 | 8.146 | 7.538 |
| DFPT | $12 \times 12 \times 12$ | | 11.291 | 8.149 | 7.541 |
| LR-cDFT | $3 \times 3 \times 3$ | $4 \times 4 \times 4$ | 11.295 | 8.168 | 7.548 |
| DFPT | $12 \times 12 \times 12$ | | 11.293 | 8.172 | 7.550 |

using a primitive unit cell with \mathbf{q} point meshes of equal size $L_1 \times L_2 \times L_3$. We present here the case with $L_1 = L_2 = L_3$, but we have also thoroughly tested less uniform cases with $L_1 \neq L_2 \neq L_3$. As for what concerns \mathbf{k} point meshes, the ones used in LR-cDFT are folded in the Brillouin zone of the primitive unit cell to maintain the same density of \mathbf{k} points in DFPT and achieve an equivalent representation of ground-state quantities (as, e.g., the spin charge density).

By construction, LR-cDFT and DFPT are meant to give the same value of U . As evident from Table I, the values of U obtained from LR-cDFT and DFPT agree within $5 \times 10^{-3} \text{ eV}$ for all three systems considered here, a negligible numerical difference that produces irrelevant variations in DFT + U results. In order to obtain an even closer agreement between these results, one would need to achieve even tighter convergence of the response matrices (in the present calculations they agree to better than 10^{-4} eV^{-1}) which not only incurs in significantly higher computational costs but is also difficult to achieve for LR-cDFT, for the reasons explained below.

It is important to note that the numerical precision of the DFPT approach is inherently higher than that of LR-cDFT. In DFPT, the iterative self-consistent solution of the LR KS equations (45) for a primitive unit cell is continued until the response matrices (41) are converged with the required tolerance. Conversely, in LR-cDFT the convergence of the response matrices depends on the precision of the iterative self-consistent solution of the modified KS equations (17) for a supercell. Controlling and improving this convergence in LR-cDFT is intrinsically more difficult, because of the larger simulation cells which make numerical noise more significant. Moreover, the numerical derivatives of the occupation matrices (19) are less practical as one needs to check that they are being performed within a linear response regime and, in addition, they are also more prone to the propagation and amplification of the numerical noise, affecting the occupations.

B. Convergence of U in DFPT

Figure 1 shows the convergence of U with respect to the number of \mathbf{q} points used in the reciprocal-space DFPT calculations within the primitive unit cell, which is equivalent to the size of supercell that would be needed in a LR-cDFT calculation. The figure shows results for various \mathbf{k} point meshes (corresponding to $N_{\mathbf{k}}$ \mathbf{k} points) to study also the dependence of the convergence of U with respect to this parameter. As for other quantities, the choice of the optimal \mathbf{k} and \mathbf{q} point meshes depends on the accuracy desired for U . Here we have used a very strict convergence threshold $\Delta U = 0.01 \text{ eV}$ in order to demonstrate the numerical consistency of all our approaches. In practice, however, such a high accuracy is typically not needed to study materials, even if some exceptions are known (see, e.g., Refs. [80,81]). As can be seen from Fig. 1, the converged value of U is controlled by the size of both \mathbf{k} and \mathbf{q} point meshes. Note that denser \mathbf{k} point meshes show a faster convergence with respect to the size of the \mathbf{q} point mesh.

The converged U of Cu in Cu_2O [see Fig. 1(a)] is 11.29 eV, which is obtained with optimal \mathbf{k} and \mathbf{q} point meshes equal to $4 \times 4 \times 4$ and $3 \times 3 \times 3$, respectively. In the case of NiO [see Fig. 1(b)], the U of Ni converges to a value of 8.18 eV and requires $8 \times 8 \times 8$ and $5 \times 5 \times 5$ \mathbf{k} and \mathbf{q} point meshes,

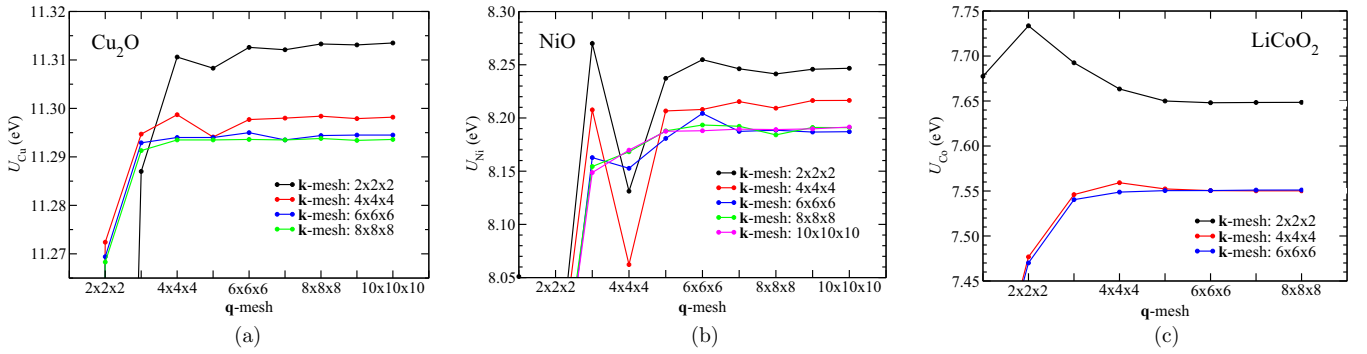


FIG. 1. Convergence of U in DFPT with respect to \mathbf{k} and \mathbf{q} point meshes for sampling of the first Brillouin zone for (a) Cu_2O , (b) NiO , and (c) LiCoO_2 .

respectively. For LiCoO_2 [see Fig. 1(c)], the converged U of Co is 7.55 eV; the calculation required \mathbf{k} and \mathbf{q} point meshes equal to $4 \times 4 \times 4$ and $3 \times 3 \times 3$, respectively. Obviously, since U is not variational with respect to the size of the \mathbf{k} and \mathbf{q} point meshes, its dependence on these parameters cannot be expected to be monotonic. Also, we remind here that these values of U are dependent on the manifold chosen and other factors (see Sec. V).

For all three systems studied here, the converged values of U are obtained using \mathbf{q} point meshes coarser than \mathbf{k} point meshes. In fact, when the \mathbf{q} point mesh is as dense or denser than the \mathbf{k} point mesh, the value of U is observed to plateau quite rapidly.

While the convergence tests shown in Fig. 1 are straightforward to perform, they might represent quite a significant computational overload to the calculations on a specific material. It is thus appropriate to extract, from the results discussed above, some indications on how dense a \mathbf{k} point mesh one has to use to obtain a reliable value of the Hubbard parameters (so that only the convergence with respect to the \mathbf{q} point mesh remains to be studied) and, specifically, how this \mathbf{k} point mesh compares with that necessary to converge other properties (e.g., total energy, forces, or stress) to typical precisions. To this purpose, first of all, we recall once again that the convergence threshold chosen here for the value of U (0.01 eV) is very strict. A change of this parameter by 0.01 eV (from the converged values) would result in a variation of total energies and stresses of 3, 4, and 5 meV/cell and 0.06, 0.20, and 0.03 Kbar for Cu_2O , NiO , and LiCoO_2 , respectively. These variations are probably smaller than the precision required for these quantities in most calculations, especially because only energy differences are meaningful. On the other hand, the \mathbf{k} point meshes determined above from the convergence study on the value of U lead to a convergence of total energies and stresses to within, respectively, 14, 1, and 7 meV/cell and 1.6, 0.5, and 0.2 Kbar for Cu_2O , NiO , and LiCoO_2 , which are probably comparable to or actually larger than the precision thresholds on these quantities needed in most cases. From Fig. 1, one can also easily see that if a larger convergence threshold is chosen for U , e.g., 0.05 eV (still quite strict), while nothing changes for LiCoO_2 , $2 \times 2 \times 2$ and $4 \times 4 \times 4$ \mathbf{k} point meshes are sufficiently dense to converge the value of this quantity for Cu_2O and NiO , respectively.

In summary, generalizing the results obtained here, reasonable levels of accuracy on U can be achieved (upon converging it with respect to the \mathbf{q} point mesh) with \mathbf{k} point meshes

as dense or coarser than needed for other properties. This can be understood since ultimately one is calculating second derivatives in energy upon charging and discharging; these are pretty robust quantities (as opposed to properties that depend only on, e.g., the states at the Fermi level). However, using coarser \mathbf{k} point meshes is not obviously convenient: Because of the “wavy” behavior of U with respect to the number of \mathbf{q} points, larger \mathbf{q} point meshes might be needed (depending on the material and the desired level of accuracy) that might offset the computational advantage (see Sec. VIC).

C. Scaling

We now discuss a rough estimate of the computational cost of the reciprocal-space DFPT approach and compare its scaling with that of the LR-cDFT approach.

DFT calculations for a cell containing N_{at} atoms and using $N_{\mathbf{k}}$ \mathbf{k} points to sample the first Brillouin zone have a computational cost that scales as the cubic power of N_{at} and linearly with $N_{\mathbf{k}}$: $T_{\text{DFT}} = AN_{\mathbf{k}}(N_{\text{at}})^3$. Here, A is a constant factor for self-consistent total energy calculations which depends on the kinetic-energy cutoff and various technical details of the implementation. The LR-cDFT approach consists of a series of calculations using supercells; hence, its computational cost can be expressed as $T_{\text{LR-cDFT}} = (A + 2B)N_{\text{pert}}N_{\mathbf{k}}^{\text{sc}}(N_{\text{at}}^{\text{sc}})^3$, where $N_{\text{at}}^{\text{sc}} = N_{\mathbf{q}}N_{\text{at}}$ is the number of atoms in the supercell, N_{pert} is the number of perturbations (i.e., the number of inequivalent Hubbard atoms which must be perturbed), and, as was discussed in Sec. IV B 2, $N_{\mathbf{k}}^{\text{sc}} = N_{\mathbf{k}}/N_{\mathbf{q}}$. The factor $2B$ accounts for the cost of two total energy perturbative calculations [with two different values of λ^J ; see Eq. (17)]. In the previous formula, $B < A$ since fewer iterations are typically needed in order to reach convergence when the calculation starts from a previously converged unperturbed total potential. Conversely, the computational cost of the DFPT approach can be expressed as $T_{\text{DFPT}} = (A + N_{\mathbf{q}}N_{\text{pert}}C)N_{\mathbf{k}}(N_{\text{at}})^3$, where the factor A accounts for the DFT calculation using a primitive unit cell in the unperturbed state and the factor C ($C > A$) represents the unitary cost of the $N_{\mathbf{q}}N_{\text{pert}}$ DFPT calculations which are necessary to compute the response matrices. This estimate neglects the computational cost of the preliminary non-self-consistent calculations of the lattice-periodic parts of the ground-state KS wave functions at $\mathbf{k} + \mathbf{q}$ points, $\vec{u}_{\mathbf{k}+\mathbf{q}\sigma}^{\circ}(\mathbf{r})$, which are needed to construct the projectors $\hat{\mathcal{O}}_{\mathbf{k}+\mathbf{q},\sigma}$ and $\hat{\mathcal{P}}_{\mathbf{k}+\mathbf{q},\sigma}$ [see Eqs. (A18) and (A21), respectively]. In fact, this cost

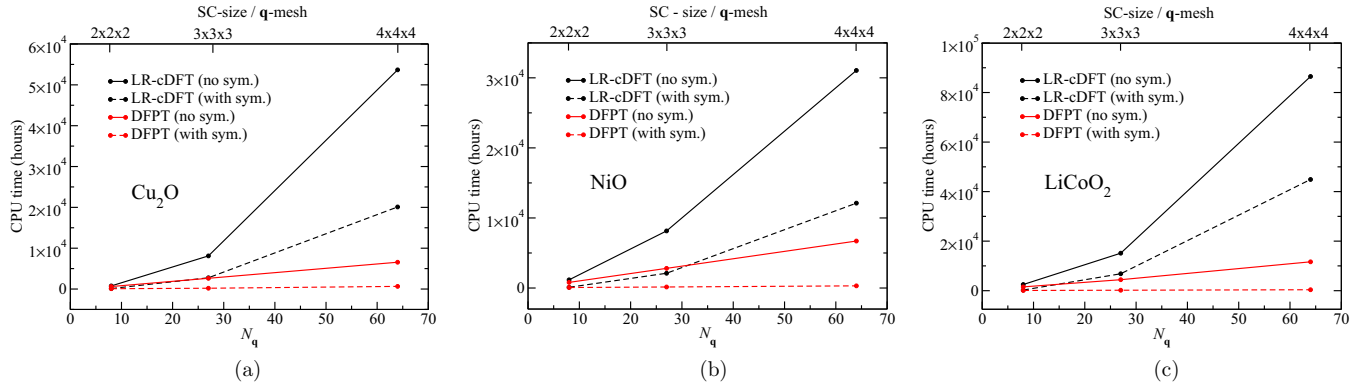


FIG. 2. Comparison of CPU time (in core hours) needed to compute U using LR-cDFT and DFPT for (a) Cu_2O , (b) NiO , and (c) LiCoO_2 . The results are shown as a function of $N_{\mathbf{q}}$ (lower horizontal axis) and the supercell (SC) size (or, equivalently, the size of the \mathbf{q} point mesh) (upper horizontal axis) for the cases with and without symmetry for \mathbf{k} (and \mathbf{q}) points. The result for DFPT with symmetry for \mathbf{q} points is represented on the same scale as without symmetry for the sake of clarity. The \mathbf{k} point meshes used for LR-cDFT and DFPT calculations are listed in Table I. Solid lines are guides for the eye.

is very small compared to that of solving self-consistent LR KS equations (45). Putting everything together, the direct comparison between the costs of LR-cDFT and DFPT gives, after simple algebraic manipulations, the following ratio:

$$\frac{T_{\text{LR-cDFT}}}{T_{\text{DFPT}}} = \frac{(A + 2B)N_{\text{pert}}N_{\mathbf{q}}^2}{A + N_{\mathbf{q}}N_{\text{pert}}C}. \quad (49)$$

Based on our experience, a rough estimate of the factors appearing in this equations is $B \approx (2/3)A$ and $C \approx 7A$. After neglecting the first term in the denominator (since $A \ll N_{\mathbf{q}}N_{\text{pert}}C$), and by using these estimates of B and C , we obtain $T_{\text{LR-cDFT}}/T_{\text{DFPT}} \approx N_{\mathbf{q}}/3$. Therefore, as the supercell size (and, consequently, the number of \mathbf{q} points in the primitive unit cell) is larger, DFPT becomes more convenient because of its better scaling with respect to LR-cDFT. We want to stress that this estimate of timings of the LR-cDFT and DFPT methods is necessarily approximate, and in practice deviations from these results might be observed because of, e.g., the size of the system [and the possible predominance of fast Fourier transforms (FFT) on diagonalization costs], the way the algorithms are implemented, the efficiency of the code in writing and reading operations to/from the disk, the efficiency of the parallelization of operations in the the total-energy and linear-response calculations, etc. An additional advantage of DFPT is the possibility to exploit the symmetry of the system, which allows us to restrict \mathbf{q} points to the irreducible wedge of the first Brillouin zone. As a consequence, the factor $N_{\mathbf{q}}$ in the denominator of Eq. (49) might be reduced while the factor $N_{\mathbf{q}}$ in the numerator remains unchanged. It should also be noted that using supercells (as in LR-cDFT) typically makes convergence more difficult (more iterations are needed) because the numerical noise tends to be larger. On the other hand, DFPT calculations in primitive unit cells can be parallelized more effectively because they are based on denser \mathbf{k} point meshes. In fact, the parallelization over \mathbf{k} points is intrinsically more flexible: Since all equations are diagonal in \mathbf{k} , \mathbf{k} -specific blocks can be solved separately by different groups of cores. This way the communications between groups of cores working on different \mathbf{k} points are limited to the instances when summations over the Brillouin zone are needed.

In order to check the comparative scaling analysis completed above, we have compared the computational costs of calculations of U with LR-cDFT and DFPT as a function of the size of supercells (or, equivalently, \mathbf{q} point meshes) with and without symmetry. The results are shown in Fig. 2. When symmetry is not used, LR-cDFT scales roughly quadratically with respect to $N_{\mathbf{q}}$ (with deviations due to reasons mentioned above), while DFPT shows a clear linear behavior, as expected. The difference in CPU time between LR-cDFT and DFPT thus increases with the size of the supercell or the size of the \mathbf{q} point mesh, respectively. When symmetry is used, we can see a pronounced decrease in the CPU time for both methods. In this case, LR-cDFT takes less CPU time because the number of \mathbf{k} points is decreased due to symmetry, while the size of the supercells is the same as when symmetry is not used; conversely, in DFPT not only the number of \mathbf{k} points is decreased but also the number of \mathbf{q} points is decreased, which leads to a decrease of CPU time by a larger factor with respect to LR-cDFT. In the case of DFPT with symmetry, for all three systems we have $N_{\mathbf{q}} = 4$, $N_{\mathbf{q}} = 6$, and $N_{\mathbf{q}} = 13$ when $2 \times 2 \times 2$, $3 \times 3 \times 3$, and $4 \times 4 \times 4$ \mathbf{q} point meshes are used, respectively (in Fig. 2 this case is shown with the same $N_{\mathbf{q}}$ as when symmetry is not used, just for the sake of clarity). The ratio of timings of LR-cDFT and DFPT both without symmetry departs our rough theoretical estimate $T_{\text{LR-cDFT}}/T_{\text{DFPT}} \approx N_{\mathbf{q}}/3$ (it is smaller by a factor of 2–5), and this is not surprising since many factors (parallelization, FFT, etc.) do influence strongly the prediction for scaling. Nevertheless, we can say that approximately $T_{\text{LR-cDFT}}/T_{\text{DFPT}} \propto N_{\mathbf{q}}$ with a prefactor which depends on many technical details of the implementation. Therefore, we can conclude that DFPT is a significantly more efficient method than LR-cDFT for calculating U , and, in particular, it offers a large computational advantage in the study of convergence (as in Sec. VIB). The most important point, though, is its robustness and automation.

VII. CONCLUSIONS

We have presented an implementation of the linear-response method for first-principles calculations of Hubbard

parameters. Based on DFPT, it avoids costly supercell calculations by summing the responses to monochromatic perturbations in primitive unit cells. We have showcased its use to compute U for several prototypical transition-metal oxides and demonstrated that the DFPT approach gives the same results as LR-cDFT, greatly decreasing the computational cost, improving the precision in determining these parameters, and making their calculation robust, automatic, and user friendly. This method opens the way for high-throughput studies of materials with strongly localized d - and/or f -type electrons.

The DFPT approach presented here is static and hence allows us to compute static Hubbard parameters. However, a generalization to the time domain (using time-dependent density-functional (perturbation) theory [82,83]) can be investigated which would allow us to obtain frequency-dependent Hubbard parameters, like in the cRPA approach. Finally, a generalization of DFPT to tackle intersite interactions has also been done and will be discussed elsewhere, together with the extension of the formalism to metals and ultrasoft pseudopotentials and the PAW method.

ACKNOWLEDGMENTS

This research was supported by the Swiss National Science Foundation (SNSF), through Grant No. 200021-179139, and its National Centre of Competence in Research (NCCR) MARVEL. Computer time was provided by CSCS (Piz Daint) through Project No. s836.

APPENDIX: LATTICE PERIODICITY

In this appendix, we discuss the details of the DFPT formalism for periodic solids. In particular, we show how linear-response quantities in a supercell with an isolated perturbation (i.e., a perturbation that has the periodicity of the supercell) can be rewritten as a sum of monochromatic \mathbf{q} components in a primitive unit cell.

1. Bloch sums of localized functions and Bloch functions

We start from setting the formalism of Bloch sums and their lattice-periodic parts.

Using the fact that $I = (l, s)$, with l being the cell index (\mathbf{R}_l pointing to the cell), and s being the atomic index (τ_s inside cell \mathbf{R}_l), so that the position of atom I is $\mathbf{R}_l + \tau_s$, we have that the localized functions are

$$\varphi_m^I(\mathbf{r}) = \varphi_m^{l,s}(\mathbf{r}) = \varphi_m^s(\mathbf{r} - \mathbf{R}_l) = \varphi_m^{\gamma(s)}(\mathbf{r} - \mathbf{R}_l - \tau_s), \quad (\text{A1})$$

where $\gamma(s)$ is the atomic type of the s th atom. In a periodic solid, we can construct Bloch sums of localized functions (A1) [84]. By doing so, we obtain

$$\tilde{\varphi}_{m,\mathbf{k}}^s(\mathbf{r}) = \frac{1}{\sqrt{N_{\mathbf{k}}}} \sum_{\mathbf{R}_l} e^{i\mathbf{k}\cdot\mathbf{R}_l} \varphi_m^s(\mathbf{r} - \mathbf{R}_l), \quad (\text{A2})$$

where summations run over the $N_{\mathbf{k}}$ primitive unit cells of the Born-von Karman supercell. Given Eq. (A2), we can define inverse Bloch sums as

$$\varphi_m^s(\mathbf{r} - \mathbf{R}_l) = \frac{1}{\sqrt{N_{\mathbf{k}}}} \sum_{\mathbf{k}} e^{-i\mathbf{k}\cdot\mathbf{R}_l} \tilde{\varphi}_{m,\mathbf{k}}^s(\mathbf{r}). \quad (\text{A3})$$

It is useful to recall the analogy of Eqs. (A2) and (A3) with those between KS wave functions and Wannier functions (see, e.g., Ref. [85]).

According to Bloch theorem, in periodic solids the ground-state KS wave functions can be written as

$$\psi_{v\mathbf{k}\sigma}(\mathbf{r}) = \frac{1}{\sqrt{N_{\mathbf{k}}}} e^{i\mathbf{k}\cdot\mathbf{r}} \bar{u}_{v\mathbf{k}\sigma}(\mathbf{r}), \quad (\text{A4})$$

where $\bar{u}_{v\mathbf{k}\sigma}(\mathbf{r})$ are the lattice-periodic parts of the ground-state KS wave functions [$\bar{u}_{v\mathbf{k}\sigma}(\mathbf{r} + \mathbf{R}_l) = \bar{u}_{v\mathbf{k}\sigma}(\mathbf{r})$]. The functions $\tilde{\varphi}_{m,\mathbf{k}}^s(\mathbf{r})$ are Bloch-like functions, as $\psi_{v\mathbf{k}\sigma}(\mathbf{r})$, and can also be expressed in the same way,

$$\tilde{\varphi}_{m,\mathbf{k}}^s(\mathbf{r}) = \frac{1}{\sqrt{N_{\mathbf{k}}}} e^{i\mathbf{k}\cdot\mathbf{r}} \tilde{\varphi}_{m,\mathbf{k}}^s(\mathbf{r}), \quad (\text{A5})$$

where $\tilde{\varphi}_{m,\mathbf{k}}^s(\mathbf{r})$ is defined as

$$\tilde{\varphi}_{m,\mathbf{k}}^s(\mathbf{r}) \equiv e^{-i\mathbf{k}\cdot\tau_s} \tilde{\varphi}_{m,\mathbf{k}}^{\gamma(s)}(\mathbf{r} - \tau_s). \quad (\text{A6})$$

The normalization factor $1/\sqrt{N_{\mathbf{k}}}$ has been chosen such that the lattice-periodic functions $\bar{u}_{v\mathbf{k}\sigma}(\mathbf{r})$ are normalized to unity in the primitive unit cell, with inner products involving them understood as integrals over one primitive unit cell, namely,

$$\begin{aligned} \langle \psi_{v\mathbf{k}\sigma} | \psi_{v'\mathbf{k}'\sigma'} \rangle &\equiv \int_V \psi_{v\mathbf{k}\sigma}^*(\mathbf{r}) \psi_{v'\mathbf{k}'\sigma'}(\mathbf{r}) d\mathbf{r} \\ &= \delta_{\mathbf{k},\mathbf{k}'} \langle \bar{u}_{v\mathbf{k}\sigma} | \bar{u}_{v'\mathbf{k}'\sigma'} \rangle, \end{aligned} \quad (\text{A7})$$

where

$$\begin{aligned} \langle \bar{u}_{v\mathbf{k}\sigma} | \bar{u}_{v'\mathbf{k}'\sigma'} \rangle &\equiv \int_v \bar{u}_{v\mathbf{k}\sigma}^*(\mathbf{r}) \bar{u}_{v'\mathbf{k}'\sigma'}(\mathbf{r}) d\mathbf{r} \\ &= \delta_{vv'} \delta_{\sigma\sigma'}, \end{aligned} \quad (\text{A8})$$

where v is the volume of the primitive unit cell and $V = v N_{\mathbf{k}}$. Here we use the notation $\langle a|b \rangle$ which stands for the inner product between a and b over the whole crystal of volume V when a and b represent $\psi_{v\mathbf{k}\sigma}(\mathbf{r})$, and the inner product over the primitive unit cell of volume v when a and b represent $\bar{u}_{v\mathbf{k}\sigma}(\mathbf{r})$.

2. Ground-state quantities

In this section, we will show how to rewrite all ground-state quantities that enter DFPT (see Sec. IV A) as the product of a lattice-periodic part and appropriate phase factors. The derivation is based on the equations of Sec. A 1.

Using Eqs. (11) and (A4), it is easy to show that the ground-state spin charge density can be written as

$$\rho_{\sigma}^{\circ}(\mathbf{r}) = \frac{1}{N_{\mathbf{k}}} \sum_{\mathbf{k}} \sum_{\mathbf{v}}^{N_{\text{occ}}} |\bar{u}_{v\mathbf{k}\sigma}^{\circ}(\mathbf{r})|^2. \quad (\text{A9})$$

Similarly, it can be shown that using Eqs. (3), (4), (A3), (A4), and (A5), the ground-state occupation matrices can be written as

$$n_{m_1 m_2}^{s\sigma} = \frac{1}{N_{\mathbf{k}}} \sum_{\mathbf{k}} \sum_{\mathbf{v}}^{N_{\text{occ}}} \langle \bar{u}_{v\mathbf{k}\sigma}^{\circ} | \hat{P}_{m_2, m_1, \mathbf{k}, \mathbf{k}}^s | \bar{u}_{v\mathbf{k}\sigma}^{\circ} \rangle, \quad (\text{A10})$$

where $\hat{P}_{m_2, m_1, \mathbf{k}, \mathbf{k}}^s$ corresponds to the $\mathbf{q} = \mathbf{0}$ case of Eq. (44). It is important to stress that the ground-state occupation

matrices (A10) do not depend on the index of the primitive unit cell, which is a consequence of the periodicity of the crystal.

The ground-state Hamiltonian, Eq. (6), in the coordinate representation can be represented as

$$H_\sigma^\circ(\mathbf{r}, \mathbf{r}') = \frac{1}{N_{\mathbf{k}}} \sum_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{r}} \bar{H}_{\mathbf{k},\sigma}^\circ(\mathbf{r}, \mathbf{r}') e^{-i\mathbf{k}\cdot\mathbf{r}'}, \quad (\text{A11})$$

where the Bloch phases and the lattice-periodic part have been separated and the latter reads

$$\bar{H}_{\mathbf{k},\sigma}^\circ(\mathbf{r}, \mathbf{r}') = \bar{H}_{\text{DFT},\mathbf{k},\sigma}^\circ(\mathbf{r}, \mathbf{r}') + \bar{V}_{\text{Hub},\mathbf{k},\sigma}^\circ(\mathbf{r}, \mathbf{r}'), \quad (\text{A12})$$

which in the operator form can be written as

$$\hat{H}_{\mathbf{k},\sigma}^\circ = \hat{H}_{\text{DFT},\mathbf{k},\sigma}^\circ + \hat{V}_{\text{Hub},\mathbf{k},\sigma}^\circ. \quad (\text{A13})$$

The first term in Eq. (A13) can be expressed as

$$\hat{H}_{\text{DFT},\mathbf{k},\sigma}^\circ = -\frac{1}{2}(\nabla + i\mathbf{k})^2 + \hat{V}_{\text{NL},\mathbf{k}}^\circ + \hat{V}_{\text{loc}}^\circ + \hat{V}_{\text{Hxc}}^\circ, \quad (\text{A14})$$

where the first term is the kinetic energy operator, $\hat{V}_{\text{NL},\mathbf{k}}^\circ$ is the lattice-periodic component of the nonlocal part of PP as in standard DFT plane-wave implementations [47,69], and the last two terms are the local part of PP and Hxc potentials. The nonlocality of $\bar{H}_{\text{DFT},\mathbf{k},\sigma}^\circ(\mathbf{r}, \mathbf{r}')$ comes from the nonlocality of $\bar{V}_{\text{NL},\mathbf{k}}^\circ(\mathbf{r}, \mathbf{r}')$. $V_{\text{Hxc}}^\circ(\mathbf{r})$ is defined in Eqs. (9) and (10), using the expression for the ground-state spin charge density (A9). The lattice-periodic part of the ground-state Hubbard potential, i.e., second term in Eq. (A13), reads [see Eq. (8)]

$$\hat{V}_{\text{Hub},\mathbf{k},\sigma}^\circ = \sum_{sm_1m_2} U^s \left(\frac{\delta_{m_1m_2}}{2} - n_{m_1m_2}^{s\sigma} \right) \hat{P}_{m_1,m_2,\mathbf{k},\mathbf{k}}^s, \quad (\text{A15})$$

where $n_{m_1m_2}^{s\sigma}$ is given by Eq. (A10), and $\hat{P}_{m_1,m_2,\mathbf{k},\mathbf{k}}^s$ can be obtained from Eq. (44) by taking $\mathbf{q} = \mathbf{0}$. As was noted for the ground-state occupation matrices [see Eq. (A10)], also $\hat{V}_{\text{Hub},\mathbf{k},\sigma}^\circ$ does not depend on the index of the primitive unit cell due to the periodicity of the crystal.

Finally, let us consider the projectors on occupied and empty states, \hat{O}_σ and \hat{P}_σ , which appear in Eq. (27). Similarly to Eq. (A11), the operator \hat{O}_σ can be written in the coordinate representation as

$$\mathcal{O}_\sigma(\mathbf{r}, \mathbf{r}') = \frac{1}{N_{\mathbf{k}}} \sum_{\mathbf{k}'} e^{i\mathbf{k}'\cdot\mathbf{r}} \bar{\mathcal{O}}_{\mathbf{k}'\sigma}(\mathbf{r}, \mathbf{r}') e^{-i\mathbf{k}'\cdot\mathbf{r}'}, \quad (\text{A16})$$

where $\bar{\mathcal{O}}_{\mathbf{k}'\sigma}(\mathbf{r}, \mathbf{r}')$ is the lattice-periodic part defined as

$$\bar{\mathcal{O}}_{\mathbf{k}'\sigma}(\mathbf{r}, \mathbf{r}') = \sum_{\nu}^{N_{\text{occ}}} \bar{u}_{\nu\mathbf{k}'\sigma}^\circ(\mathbf{r}) \bar{u}_{\nu\mathbf{k}'\sigma}^{\circ*}(\mathbf{r}'). \quad (\text{A17})$$

Using an operator notation, it is possible to express the same projector as follows:

$$\hat{\mathcal{O}}_{\mathbf{k}'\sigma} = \sum_{\nu'}^{N_{\text{occ}}} |\bar{u}_{\nu'\mathbf{k}'\sigma}^\circ\rangle \langle \bar{u}_{\nu'\mathbf{k}'\sigma}^\circ|. \quad (\text{A18})$$

In the same manner, it can be shown that the operator \hat{P}_σ can be written in the coordinate representation as

$$\mathcal{P}_\sigma(\mathbf{r}, \mathbf{r}') = \frac{1}{N_{\mathbf{k}}} \sum_{\mathbf{k}'} e^{i\mathbf{k}'\cdot\mathbf{r}} \bar{\mathcal{P}}_{\mathbf{k}'\sigma}(\mathbf{r}, \mathbf{r}') e^{-i\mathbf{k}'\cdot\mathbf{r}'}, \quad (\text{A19})$$

where the lattice-periodic part $\bar{\mathcal{P}}_{\mathbf{k}'\sigma}(\mathbf{r}, \mathbf{r}')$ can obviously be written as follows,

$$\bar{\mathcal{P}}_{\mathbf{k}'\sigma}(\mathbf{r}, \mathbf{r}') = \delta(\mathbf{r} - \mathbf{r}') - \bar{\mathcal{O}}_{\mathbf{k}'\sigma}(\mathbf{r}, \mathbf{r}'), \quad (\text{A20})$$

which translates in the operator notation as

$$\hat{\mathcal{P}}_{\mathbf{k}'\sigma} = 1 - \hat{\mathcal{O}}_{\mathbf{k}'\sigma}. \quad (\text{A21})$$

3. Decomposition of linear-response quantities into monochromatic components

In this section, we will demonstrate how a localized, neutral perturbation that has the periodicity of a supercell (chosen large enough to mimic a truly isolated perturbation) can be recast as a sum of monochromatic (\mathbf{q} -specific) perturbations in a primitive unit cell, as discussed in Sec. IV B 1. Let us start from the perturbing potential, Eq. (18). In periodic-boundary conditions, perturbations cannot be isolated. In order to model the response of occupations to an isolated perturbation, supercells must be used, whose size is chosen large enough to make interactions between periodic replicas negligible. In mathematical terms, this translates into the following approximation for the perturbing potential [Eq. (18)]:

$$\begin{aligned} V_{\text{pert}}^{s'l'}(\mathbf{r}, \mathbf{r}') &= \sum_m \varphi_m^{s'}(\mathbf{r} - \mathbf{R}_{l'}) \varphi_m^{s'*}(\mathbf{r} - \mathbf{R}_{l'}) \\ &\simeq \sum_{\mathbf{R}_{\text{sc}}} \sum_m \varphi_m^{s'}(\mathbf{r} - \mathbf{R}_{l'} - \mathbf{R}_{\text{sc}}) \\ &\quad \times \varphi_m^{s'*}(\mathbf{r} - \mathbf{R}_{l'} - \mathbf{R}_{\text{sc}}). \end{aligned} \quad (\text{A22})$$

If we approximate each of the localized functions in Eq. (A22) with the inverse Bloch sums, Eq. (A3), then the largest (nontrivial) supercell vectors $\mathbf{R}_{\text{sc}}^{\text{max}}$ on which the external sum in Eq. (A22) runs over are such that $\mathbf{R}_{\text{sc}}^{\text{max}} \cdot \Delta\mathbf{k} = 2\pi$, where $\Delta\mathbf{k}$ is the minimum spacing between \mathbf{k} points in the first Brillouin zone of the primitive unit cell. Larger \mathbf{R}_{sc} 's are, in fact, implicitly summed over, because of the periodicity of inverse Bloch sums [see Eq. (A3)]. Therefore, the external summation in Eq. (A22) runs over N_{sc} vectors \mathbf{R}_{sc} , where $N_{\text{sc}} = N_{\mathbf{k}}^{\text{sc}} = N_{\mathbf{k}}/N_{\mathbf{q}}$. Thus, using Eq. (A3), we can rewrite Eq. (A22) as follows:

$$\begin{aligned} V_{\text{pert}}^{s'l'}(\mathbf{r}, \mathbf{r}') &= \frac{1}{N_{\mathbf{k}}} \sum_{\mathbf{R}_{\text{sc}}} \sum_m \sum_{\mathbf{k}'} e^{-i\mathbf{k}'\cdot\mathbf{R}_{l'}} \bar{\varphi}_{m,\mathbf{k}'}^{s'}(\mathbf{r}) \\ &\quad \times \sum_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{R}_{l'}} \bar{\varphi}_{m,\mathbf{k}}^{s'*}(\mathbf{r}') e^{-i(\mathbf{k}'-\mathbf{k})\cdot\mathbf{R}_{\text{sc}}}. \end{aligned} \quad (\text{A23})$$

In Eq. (A23), the sum over \mathbf{R}_{sc} of the phase factor $e^{-i(\mathbf{k}'-\mathbf{k})\cdot\mathbf{R}_{\text{sc}}}$ imposes the difference between \mathbf{k}' and \mathbf{k} to be equal to reciprocal lattice vectors of the supercell, which we label as \mathbf{q} (see Sec. IV B 1). Therefore, in the summation over \mathbf{k}' only $N_{\mathbf{q}}$ terms survive, which correspond to $\mathbf{k}' = \mathbf{k} + \mathbf{q}$. Thus, we obtain

$$\begin{aligned} V_{\text{pert}}^{s'l'}(\mathbf{r}, \mathbf{r}') &= \frac{1}{N_{\mathbf{k}}} \sum_{\mathbf{R}_{\text{sc}}} \sum_m \sum_{\mathbf{q}} \sum_{\mathbf{k}} e^{-i(\mathbf{k}+\mathbf{q})\cdot\mathbf{R}_{l'}} \bar{\varphi}_{m,\mathbf{k}+\mathbf{q}}^{s'}(\mathbf{r}) \\ &\quad \times e^{i\mathbf{k}\cdot\mathbf{R}_{l'}} \bar{\varphi}_{m,\mathbf{k}}^{s'*}(\mathbf{r}') e^{-i\mathbf{q}\cdot\mathbf{R}_{\text{sc}}}. \end{aligned} \quad (\text{A24})$$

Since \mathbf{q} vectors are the reciprocal lattice vectors of the supercell, $e^{-i\mathbf{q}\cdot\mathbf{R}_{\text{sc}}} = 1$, and the summation over \mathbf{R}_{sc} gives a prefactor N_{sc} . Equation (A24) can thus be written as

$$V_{\text{pert}}^{s'l'}(\mathbf{r}, \mathbf{r}') = \frac{1}{N_{\mathbf{q}}} \sum_{\mathbf{q}} e^{-i\mathbf{q}\cdot\mathbf{R}_{l'}} \sum_m \sum_{\mathbf{k}} \tilde{\varphi}_{m,\mathbf{k}+\mathbf{q}}^{s'}(\mathbf{r}) \tilde{\varphi}_{m,\mathbf{k}}^{s'*}(\mathbf{r}). \quad (\text{A25})$$

Finally, by using Eq. (A5), we can rewrite Eq. (A25) as

$$V_{\text{pert}}^{s'l'}(\mathbf{r}, \mathbf{r}') = \frac{1}{N_{\mathbf{q}}} \sum_{\mathbf{q}} e^{-i\mathbf{q}\cdot\mathbf{R}_{l'}} V_{\text{pert},\mathbf{q}}^{s'}(\mathbf{r}, \mathbf{r}'), \quad (\text{A26})$$

where

$$V_{\text{pert},\mathbf{q}}^{s'}(\mathbf{r}, \mathbf{r}') = \frac{1}{N_{\mathbf{k}}} \sum_{\mathbf{k}} e^{i(\mathbf{k}+\mathbf{q})\cdot\mathbf{r}} \tilde{V}_{\text{pert},\mathbf{k}+\mathbf{q},\mathbf{k}}^{s'}(\mathbf{r}, \mathbf{r}') e^{-i\mathbf{k}\cdot\mathbf{r}'}. \quad (\text{A27})$$

In Eq. (A27), $\tilde{V}_{\text{pert},\mathbf{k}+\mathbf{q},\mathbf{k}}^{s'}(\mathbf{r}, \mathbf{r}')$ is the lattice-periodic part of the perturbing potential:

$$\tilde{V}_{\text{pert},\mathbf{k}+\mathbf{q},\mathbf{k}}^{s'}(\mathbf{r}, \mathbf{r}') = \sum_m \tilde{P}_{m,m,\mathbf{k}+\mathbf{q},\mathbf{k}}^{s'}(\mathbf{r}, \mathbf{r}'), \quad (\text{A28})$$

with $\tilde{P}_{m,m,\mathbf{k}+\mathbf{q},\mathbf{k}}^{s'}(\mathbf{r}, \mathbf{r}')$ representing the diagonal element of the lattice-periodic operator in the coordinate representation, which reads

$$\tilde{P}_{m_2,m_1,\mathbf{k}+\mathbf{q},\mathbf{k}}^{s'}(\mathbf{r}, \mathbf{r}') = \tilde{\varphi}_{m_2,\mathbf{k}+\mathbf{q}}^{s'}(\mathbf{r}) \tilde{\varphi}_{m_1,\mathbf{k}}^{s'*}(\mathbf{r}'). \quad (\text{A29})$$

Equation (A29) can be written in operator form [see Eq. (44)]. Equation (A26) constitutes a *decomposition* of the isolated perturbing potential into monochromatic components modulated by a phase factor corresponding to the considered wave vector \mathbf{q} [86].

Similarly to the decomposition of the perturbing potential, Eq. (A26), we can represent the LR KS wave functions of the supercell as the inverse Bloch sum of the LR KS wave functions of the primitive unit cell,

$$\frac{d\tilde{\psi}_{v\mathbf{k}\sigma}(\mathbf{r})}{d\lambda^{s'l'}} = \frac{1}{N_{\mathbf{q}}} \sum_{\mathbf{q}} e^{-i\mathbf{q}\cdot\mathbf{R}_{l'}} \Delta_{\mathbf{q}}^{s'} \tilde{\psi}_{v\mathbf{k}\sigma}(\mathbf{r}), \quad (\text{A30})$$

where $\Delta_{\mathbf{q}}^{s'} \tilde{\psi}_{v\mathbf{k}\sigma}(\mathbf{r})$ are the LR KS wave functions of the primitive unit cell corresponding to the monochromatic perturbation with wave vector \mathbf{q} , $V_{\text{pert},\mathbf{q}}^{s'}$. From (A30), we can define the Bloch sum

$$\Delta_{\mathbf{q}}^{s'} \tilde{\psi}_{v\mathbf{k}\sigma}(\mathbf{r}) = \sum_{\mathbf{R}_{l'}} e^{i\mathbf{q}\cdot\mathbf{R}_{l'}} \frac{d\tilde{\psi}_{v\mathbf{k}\sigma}(\mathbf{r})}{d\lambda^{s'l'}}. \quad (\text{A31})$$

It is important to note that the normalization factors $1/N_{\mathbf{q}}$ in Eq. (A30) and 1 in Eq. (A31) are chosen in such a way that these definitions are consistent with the normalization in Eq. (A26). The LR KS wave functions $\Delta_{\mathbf{q}}^{s'} \tilde{\psi}_{v\mathbf{k}\sigma}(\mathbf{r})$ are Bloch-like functions, and hence they can be written as

$$\Delta_{\mathbf{q}}^{s'} \tilde{\psi}_{v\mathbf{k}\sigma}(\mathbf{r}) = \frac{1}{\sqrt{N_{\mathbf{k}}}} e^{i(\mathbf{k}+\mathbf{q})\cdot\mathbf{r}} \Delta_{\mathbf{q}}^{s'} \tilde{u}_{v\mathbf{k}\sigma}(\mathbf{r}), \quad (\text{A32})$$

where $\Delta_{\mathbf{q}}^{s'} \tilde{u}_{v\mathbf{k}\sigma}(\mathbf{r})$ are the lattice-periodic functions, in agreement with the definition in Eq. (A4).

Similarly to the decomposition of the perturbing potential and of the LR KS wave functions into monochromatic \mathbf{q} components [see Eqs. (A26) and (A30), respectively], also the LR spin charge density can be decomposed into \mathbf{q} -specific terms,

$$\frac{d\rho_{\sigma}(\mathbf{r})}{d\lambda^{s'l'}} = \frac{1}{N_{\mathbf{q}}} \sum_{\mathbf{q}} e^{-i\mathbf{q}\cdot\mathbf{R}_{l'}} \Delta_{\mathbf{q}}^{s'} \tilde{\rho}_{\sigma}(\mathbf{r}), \quad (\text{A33})$$

where

$$\Delta_{\mathbf{q}}^{s'} \tilde{\rho}_{\sigma}(\mathbf{r}) = e^{i\mathbf{q}\cdot\mathbf{r}} \Delta_{\mathbf{q}}^{s'} \tilde{\rho}_{\sigma}(\mathbf{r}) \quad (\text{A34})$$

and $\Delta_{\mathbf{q}}^{s'} \tilde{\rho}_{\sigma}(\mathbf{r})$ is given by Eq. (48). Since the LR Hxc potential depends *linearly* on the variation in the spin charge density [see Eq. (24)], the decomposition in Eq. (A33) means that also the LR Hxc potential can be decomposed in a similar way:

$$\frac{dV_{\text{Hxc},\sigma}(\mathbf{r})}{d\lambda^{s'l'}} = \frac{1}{N_{\mathbf{q}}} \sum_{\mathbf{q}} e^{-i\mathbf{q}\cdot\mathbf{R}_{l'}} \Delta_{\mathbf{q}}^{s'} \tilde{V}_{\text{Hxc},\sigma}(\mathbf{r}), \quad (\text{A35})$$

where

$$\Delta_{\mathbf{q}}^{s'} \tilde{V}_{\text{Hxc},\sigma}(\mathbf{r}) = e^{i\mathbf{q}\cdot\mathbf{r}} \Delta_{\mathbf{q}}^{s'} \tilde{V}_{\text{Hxc},\sigma}(\mathbf{r}), \quad (\text{A36})$$

with $\Delta_{\mathbf{q}}^{s'} \tilde{V}_{\text{Hxc},\sigma}(\mathbf{r})$ given by Eq. (47).

-
- [1] P. Hohenberg and W. Kohn, *Phys. Rev.* **136**, B864 (1964).
[2] W. Kohn and L. Sham, *Phys. Rev.* **140**, A1133 (1965).
[3] I. de P. R. Moreira, F. Illas, and R. L. Martin, *Phys. Rev. B* **65**, 155102 (2002).
[4] F. Corà, M. Alfredsson, G. Mallia, D. Middlemiss, W. Mackrodt, R. Dovesi, and R. Orlando, *Struct. Bonding (Berlin)* **113**, 171 (2004).
[5] X.-B. Feng and N. M. Harrison, *Phys. Rev. B* **69**, 035114 (2004).
[6] M. Alfredsson, G. D. Price, C. R. A. Catlow, S. C. Parker, R. Orlando, and J. P. Brodholt, *Phys. Rev. B* **70**, 165111 (2004).
[7] F. Tran, P. Blaha, K. Schwarz, and P. Novák, *Phys. Rev. B* **74**, 155108 (2006).
[8] V. L. Chevrier, S. P. Ong, R. Armiento, M. K. Y. Chan, and G. Ceder, *Phys. Rev. B* **82**, 075122 (2010).
[9] D.-H. Seo, A. Urban, and G. Ceder, *Phys. Rev. B* **92**, 115118 (2015).
[10] L. Lin, *J. Chem. Theory Comput.* **12**, 2242 (2016).
[11] J. Sun, A. Ruzsinszky, and J. P. Perdew, *Phys. Rev. Lett.* **115**, 036402 (2015).
[12] D. A. Kitchaev, H. Peng, Y. Liu, J. Sun, J. P. Perdew, and G. Ceder, *Phys. Rev. B* **93**, 045132 (2016).
[13] Y. Hinuma, H. Hayashi, Y. Kumagai, I. Tanaka, and F. Oba, *Phys. Rev. B* **96**, 094102 (2017).
[14] V. I. Anisimov, J. Zaanen, and O. K. Andersen, *Phys. Rev. B* **44**, 943 (1991).

- [15] V. Anisimov, F. Aryasetiawan, and A. Liechtenstein, *J. Phys.: Condens. Matter* **9**, 767 (1997).
- [16] S. L. Dudarev, G. A. Botton, S. Y. Savrasov, C. J. Humphreys, and A. P. Sutton, *Phys. Rev. B* **57**, 1505 (1998).
- [17] M. Cococcioni and S. de Gironcoli, *Phys. Rev. B* **71**, 035105 (2005).
- [18] I. Dabo, A. Ferretti, N. Poilvert, Y. Li, N. Marzari, and M. Cococcioni, *Phys. Rev. B* **82**, 115121 (2010).
- [19] H. J. Kulik, M. Cococcioni, D. A. Scherlis, and N. Marzari, *Phys. Rev. Lett.* **97**, 103001 (2006).
- [20] V. Anisimov, A. Poteryaev, M. Korotin, A. Anokhin, and G. Kotliar, *J. Phys.: Condens. Matter* **9**, 7359 (1997).
- [21] A. I. Liechtenstein and M. I. Katsnelson, *Phys. Rev. B* **57**, 6884 (1998).
- [22] G. Kotliar, S. Savrasov, K. Haule, V. Oudovenko, O. Porcollet, and C. Marianetti, *Rev. Mod. Phys.* **78**, 865 (2006).
- [23] B. Himmetoglu, A. Floris, S. de Gironcoli, and M. Cococcioni, *Int. J. Quant. Chem.* **114**, 14 (2014).
- [24] G. Moynihan, G. Teobaldi, and D. O'Regan, [arXiv:1704.08076](https://arxiv.org/abs/1704.08076) (unpublished).
- [25] H. Kulik and N. Marzari, *J. Chem. Phys.* **129**, 134314 (2008).
- [26] P. H. Dederichs, S. Blügel, R. Zeller, and H. Akai, *Phys. Rev. Lett.* **53**, 2512 (1984).
- [27] A. K. McMahan, R. M. Martin, and S. Satpathy, *Phys. Rev. B* **38**, 6650 (1988).
- [28] O. Gunnarsson, O. K. Andersen, O. Jepsen, and J. Zaanen, *Phys. Rev. B* **39**, 1708 (1989).
- [29] M. S. Hybertsen, M. Schlüter, and N. E. Christensen, *Phys. Rev. B* **39**, 9028 (1989).
- [30] O. Gunnarsson, *Phys. Rev. B* **41**, 514 (1990).
- [31] W. E. Pickett, S. C. Erwin, and E. C. Ethridge, *Phys. Rev. B* **58**, 1201 (1998).
- [32] I. V. Solovyev and M. Imada, *Phys. Rev. B* **71**, 045103 (2005).
- [33] K. Nakamura, R. Arita, Y. Yoshimoto, and S. Tsuneyuki, *Phys. Rev. B* **74**, 235113 (2006).
- [34] M. Shishkin and H. Sato, *Phys. Rev. B* **93**, 085135 (2016).
- [35] N. J. Mosey and E. A. Carter, *Phys. Rev. B* **76**, 155123 (2007).
- [36] N. Mosey, P. Liao, and E. Carter, *J. Chem. Phys.* **129**, 014103 (2008).
- [37] A. N. Andriotis, R. M. Sheetz, and M. Menon, *Phys. Rev. B* **81**, 245103 (2010).
- [38] L. A. Agapito, S. Curtarolo, and M. B. Nardelli, *Phys. Rev. X* **5**, 011006 (2015).
- [39] M. Springer and F. Aryasetiawan, *Phys. Rev. B* **57**, 4364 (1998).
- [40] T. Kotani, *J. Phys.: Condens. Matter* **12**, 2413 (2000).
- [41] F. Aryasetiawan, M. Imada, A. Georges, G. Kotliar, S. Biermann, and A. I. Liechtenstein, *Phys. Rev. B* **70**, 195104 (2004).
- [42] F. Aryasetiawan, K. Karlsson, O. Jepsen, and U. Schönberger, *Phys. Rev. B* **74**, 125106 (2006).
- [43] E. Sasioglu, C. Friedrich, and S. Blügel, *Phys. Rev. B* **83**, 121101(R) (2011).
- [44] L. Vaugier, H. Jiang, and S. Biermann, *Phys. Rev. B* **86**, 165105 (2012).
- [45] B. Amadon, T. Applencourt, and F. Bruneval, *Phys. Rev. B* **89**, 125110 (2014).
- [46] S. Baroni, P. Giannozzi, and A. Testa, *Phys. Rev. Lett.* **58**, 1861 (1987).
- [47] S. Baroni, S. de Gironcoli, A. Dal Corso, and P. Giannozzi, *Rev. Mod. Phys.* **73**, 515 (2001).
- [48] G. Borghi, A. Ferretti, N. L. Nguyen, I. Dabo, and N. Marzari, *Phys. Rev. B* **90**, 075135 (2014).
- [49] N. Colonna, N. L. Nguyen, A. Ferretti, and N. Marzari, *J. Chem. Theory Comput.* **14**, 2549 (2018).
- [50] D. Vanderbilt, *Phys. Rev. B* **41**, 7892 (1990).
- [51] P. E. Blöchl, *Phys. Rev. B* **50**, 17953 (1994).
- [52] V. L. Campo Jr. and M. Cococcioni, *J. Phys.: Condens. Matter* **22**, 055602 (2010).
- [53] M. T. Czyżyk and G. A. Sawatzky, *Phys. Rev. B* **49**, 14211 (1994).
- [54] A. G. Petukhov, I. I. Mazin, L. Chioncel, and A. I. Liechtenstein, *Phys. Rev. B* **67**, 153106 (2003).
- [55] V. Anisimov, A. Kozhevnikov, M. Korotin, A. Lukoyanov, and D. Khafizullin, *J. Phys.: Condens. Matter* **19**, 106206 (2007).
- [56] V. I. Anisimov, I. V. Solovyev, M. A. Korotin, M. T. Czyżyk, and G. A. Sawatzky, *Phys. Rev. B* **48**, 16929 (1993).
- [57] A. I. Liechtenstein, V. I. Anisimov, and J. Zaanen, *Phys. Rev. B* **52**, R5467(R) (1995).
- [58] H. Eschrig, K. Koepernik, and I. Chaplygin, *J. Solid State Chem.* **176**, 482 (2003).
- [59] The localized orbitals can be chosen to be, e.g., atomiclike, orthonormalized atomiclike, or localized Wannier functions.
- [60] P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G. Chiarotti, M. Cococcioni, I. Dabo *et al.*, *J. Phys.: Condens. Matter* **21**, 395502 (2009).
- [61] P. Giannozzi, O. Andreussi, T. Brumme, O. Bunau, M. B. Nardelli, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, M. Cococcioni *et al.*, *J. Phys.: Condens. Matter* **29**, 465901 (2017).
- [62] The idea of computing U self-consistently is not new, and there are several ways to do it (see, e.g., Refs. [19,79,34] for the self-consistent calculation of U within LR-cDFT).
- [63] A. Dal Corso and S. de Gironcoli, *Phys. Rev. B* **62**, 273 (2000).
- [64] M. Payne, M. Teter, D. Allen, T. Arias, and J. Joannopoulos, *Rev. Mod. Phys.* **64**, 1045 (1992).
- [65] This can be demonstrated simply by inserting an identity operator $\hat{1} = \hat{P}_\sigma + \hat{O}_\sigma$ in Eqs. (22) and (26). By doing so, the terms containing \hat{O}_σ will cancel out due to the property $\langle \frac{d\psi_{vk\sigma}}{d\lambda^J} | \psi_{vk\sigma} \rangle + \langle \psi_{vk\sigma} | \frac{d\psi_{vk\sigma}}{d\lambda^J} \rangle = 0$ (which can be obtained by applying $\frac{d}{d\lambda^J}$ to the orthonormality condition of the KS wave functions). Therefore, in Eqs. (22) and (26) will remain only the terms containing \hat{P}_σ ; i.e., the LR occupation matrices and spin charge densities depend only on the conduction component of the LR KS wave functions, i.e., $|\frac{d\psi_{vk\sigma}}{d\lambda^J}\rangle$.
- [66] In practice, for the sake of convenience, instead of the matrix in Eq. (41), the following matrix can be used: $\chi_{s+l, s'+l'} = \sum_{\sigma, m} \frac{dn_{mm}^{s+l\sigma}}{d\lambda^{s'+l'}}$. This response matrix is obtained simply by reorganizing the elements of the response matrix of Eq. (41).
- [67] The first and the second terms in Eq. (43) look identical with the only difference in the order of the indices m_1 and m_2 . In order to write these two terms in a similar form, we have used the time-reversal symmetry.
- [68] Note that in Eq. (45) there is no index s because the solution of this equation $|\Delta_{\mathbf{q}}^{s'} \bar{u}_{vk\sigma}\rangle$ does not depend on s . This can be understood by looking at Eq. (43): Index s belongs to the operators $\hat{P}_{m_2, m_1, \mathbf{k}, \mathbf{k}+\mathbf{q}}^s$ and $\hat{P}_{m_1, m_2, \mathbf{k}, \mathbf{k}+\mathbf{q}}^s$ and not to $|\Delta_{\mathbf{q}}^{s'} \bar{u}_{vk\sigma}\rangle$. Physically this means that once the lattice-periodic parts of the linear-response Kohn-Sham wave functions $|\Delta_{\mathbf{q}}^{s'} \bar{u}_{vk\sigma}\rangle$ are computed by perturbing a specific atom s' , the response of the

- occupation matrix of all the atoms s can be obtained by simply projecting $|\Delta_{\mathbf{q}}^s \bar{u}_{\nu, \mathbf{k}\sigma}\rangle$ on their states.
- [69] A. Dal Corso, *Phys. Rev. B* **64**, 235118 (2001).
- [70] A. Dal Corso, *Phys. Rev. B* **81**, 075123 (2010).
- [71] J. P. Perdew, A. Ruzsinszky, G. I. Csonka, O. A. Vydrov, G. E. Scuseria, L. A. Constantin, X. Zhou, and K. Burke, *Phys. Rev. Lett.* **100**, 136406 (2008).
- [72] A. Dal Corso, *Comput. Mater. Sci.* **95**, 337 (2014).
- [73] Ultrasoft pseudopotentials of the Pslibrary 0.3.1 [72] are taken from the SSSP library 0.7 [www.materialscloud.org/]: Ni.pbesol-n-rrkjus_psl.0.1.UPF, O.pbesol-n-rrkjus_psl.0.1.UPF, Li.pbesol-s-rrkjus_psl.0.2.1.UPF, Co.pbesol-spn-rrkjus_psl.0.3.1.UPF, and Cu.pbesol-dn-rrkjus_psl.0.2.UPF.
- [74] L. Bartel and B. Morosin, *Phys. Rev. B* **3**, 1039 (1971).
- [75] R. Wyckoff, *Crystal Structures* (Wiley, New York, 1965), Vol. 1.
- [76] J. Akimoto, Y. Gotoh, and Y. Oosawa, *J. Solid State Chem.* **141**, 298 (1998).
- [77] I. Mayer, *Int. J. Quant. Chem.* **90**, 63 (2002).
- [78] D. D. Johnson, *Phys. Rev. B* **38**, 12807 (1988).
- [79] H. Hsu, K. Umemoto, M. Cococcioni, and R. Wentzcovitch, *Phys. Rev. B* **79**, 125124 (2009).
- [80] R. Bianco, M. Calandra, and F. Mauri, *Phys. Rev. B* **92**, 094107 (2015).
- [81] M. Hellgren, J. Baima, R. Bianco, M. Calandra, F. Mauri, and L. Wirtz, *Phys. Rev. Lett.* **119**, 176401 (2017).
- [82] E. Runge and E. K. U. Gross, *Phys. Rev. Lett.* **52**, 997 (1984).
- [83] M. A. L. Marques, N. T. Maitra, F. M. S. Nogueira, E. K. U. Gross, and A. Rubio, editors, *Fundamentals of Time-Dependent Density Functional Theory*, Lecture Notes in Physics (Springer-Verlag, Berlin, 2012), Vol. 837.
- [84] N. Ashcroft and N. Mermin, *Solid State Physics* (Holt-Saunders, Philadelphia, 1976).
- [85] N. Marzari, A. Mostofi, J. Yates, I. Souza, and D. Vanderbilt, *Rev. Mod. Phys.* **84**, 1419 (2012).
- [86] In the inverse Bloch sum (A26), the normalization factor is taken as $1/N_{\mathbf{q}}$ and not as $1/\sqrt{N_{\mathbf{q}}}$, in contrast to Eq. (A3). This is a valid choice, and it is done for later convenience.