

# Effect of ensemble generalization on the highest-occupied Kohn-Sham eigenvalue

Cite as: J. Chem. Phys. **143**, 104105 (2015); <https://doi.org/10.1063/1.4930119>

Submitted: 10 May 2015 . Accepted: 24 August 2015 . Published Online: 10 September 2015

Eli Kraisler, Tobias Schmidt, Stephan Kümmel, and Leeor Kronik



View Online



Export Citation



CrossMark

## ARTICLES YOU MAY BE INTERESTED IN

[Fundamental gaps with approximate density functionals: The derivative discontinuity revealed from ensemble considerations](#)

The Journal of Chemical Physics **140**, 18A540 (2014); <https://doi.org/10.1063/1.4871462>

[A self-interaction-free local hybrid functional: Accurate binding energies vis-à-vis accurate ionization potentials from Kohn-Sham eigenvalues](#)

The Journal of Chemical Physics **140**, 18A510 (2014); <https://doi.org/10.1063/1.4865942>

[Density-functional thermochemistry. III. The role of exact exchange](#)

The Journal of Chemical Physics **98**, 5648 (1993); <https://doi.org/10.1063/1.464913>

## Lock-in Amplifiers up to 600 MHz

starting at

\$6,210



Zurich  
Instruments

Watch the Video



# Effect of ensemble generalization on the highest-occupied Kohn-Sham eigenvalue

Eli Kraiser,<sup>1,a)</sup> Tobias Schmidt,<sup>2,a)</sup> Stephan Kümmel,<sup>2</sup> and Leeor Kronik<sup>1</sup>

<sup>1</sup>Department of Materials and Interfaces, Weizmann Institute of Science, Rehovoth 76100, Israel

<sup>2</sup>Theoretical Physics IV, University of Bayreuth, 95440 Bayreuth, Germany

(Received 10 May 2015; accepted 24 August 2015; published online 10 September 2015)

There are several approximations to the exchange-correlation functional in density-functional theory, which accurately predict total energy-related properties of many-electron systems, such as binding energies, bond lengths, and crystal structures. Other approximations are designed to describe potential-related processes, such as charge transfer and photoemission. However, the development of a functional which can serve the two purposes simultaneously is a long-standing challenge. Trying to address it, we employ in the current work the ensemble generalization procedure proposed by Kraiser and Kronik [Phys. Rev. Lett. **110**, 126403 (2013)]. Focusing on the prediction of the ionization potential via the highest occupied Kohn-Sham eigenvalue, we examine a variety of exchange-correlation approximations: the local spin-density approximation, semi-local generalized gradient approximations, and global and local hybrid functionals. Results for a test set of 26 diatomic molecules and single atoms are presented. We find that the aforementioned ensemble generalization systematically improves the prediction of the ionization potential, for various systems and exchange-correlation functionals, without compromising the accuracy of total energy-related properties. We specifically examine hybrid functionals. These depend on a parameter controlling the ratio of semi-local to non-local functional components. The ionization potential obtained with ensemble-generalized functionals is found to depend only weakly on the parameter value, contrary to common experience with non-generalized hybrids, thus eliminating one aspect of the so-called “parameter dilemma” of hybrid functionals. © 2015 AIP Publishing LLC. [<http://dx.doi.org/10.1063/1.4930119>]

## I. INTRODUCTION

Modern density-functional theory (DFT), based on the theoretical foundation laid by Hohenberg, Kohn, and Sham<sup>1,2</sup> in the 1960s, in principle provides an exact framework for treating the many-electron problem. Within this framework, the electron-electron interaction is expressed via the exchange-correlation (xc) energy term,  $E_{xc}[n]$ , which is a functional of the electron density  $n(\mathbf{r})$ .<sup>3–5</sup> In practice, one has to approximate this energy contribution, aiming at a numerically efficient yet accurate description of the electronic structure of various many-electron systems, such as atoms, molecules, and solids.

During the eventful history of DFT, many density functional approximations (DFAs) to the exact xc energy were developed.<sup>6,7</sup> The performance of each of them can be evaluated from a twofold perspective: On the one hand, ground-state quantities such as binding energies, bond lengths, and crystal structures are closely related to the total energy of the system and hence to the approximate xc energy  $E_{xc}$  itself. On the other hand, there also exists a range of physical properties and processes, whose description is substantially influenced by the xc potential,  $v_{xc}(\mathbf{r}) := \delta E_{xc}[n]/\delta n(\mathbf{r})$ . Prominent examples for the latter category are charge-transfer and ionization processes, as well as the description of photoemission spectra. Here, especially the

approximate interpretation of Kohn-Sham (KS) eigenvalues as a physically meaningful density of states lies at the focal point of ongoing research (see Ref. 8 and references therein).

For applications to real materials, one would wish to have a DFA with a good performance from *both* perspectives. However, this is not the case for many existing DFAs. The development of such a DFA is a long-standing challenge, as discussed below.

One of the exact relations in KS-DFT is the ionization potential (IP) theorem,  $-\varepsilon_{ho} = I$ ,<sup>9–14</sup> which relates the highest occupied (ho) KS eigenvalue,  $\varepsilon_{ho}$ , to the IP,  $I$ , i.e., the removal energy of one electron from an  $N$ -electron system. Consequently, there are two fundamentally different ways to obtain the IP in DFT: by evaluating  $-\varepsilon_{ho}$  or (with more computational effort) by calculating the energy difference of the ionized and neutral system, with  $N_0 - 1$  and  $N_0$  electrons, respectively,  $I_{\Delta SCF} = E(N_0 - 1) - E(N_0)$ . This is usually referred to as the  $\Delta$ SCF approach. Note that while the latter relies on accurate total energy values, the former relies on the *potential* to yield an accurate ho energy level.

The  $\Delta$ SCF approach, which has been extensively used since the early days of DFT,<sup>15–25</sup> usually yields an IP with a satisfactory accuracy of a few percent with respect to experiment, for atoms and small molecules, even with standard (semi-)local DFAs, such as the local spin-density approximation (LSDA)<sup>26</sup> or the generalized gradient approximation (GGA).<sup>27</sup> Comparison of  $-\varepsilon_{ho}$  to the experimental IP, however, shows poor correspondence for

<sup>a)</sup>E. Kraiser and T. Schmidt contributed equally to this work.

many DFAs. For example, with the aforementioned (semi-)local approximations, one can observe an underestimation of up to 50%, as manifested in, e.g., Refs. 21 and 28–32.

This failure has been related to various systematic shortcomings of existing functionals. First, the one-electron self-interaction problem, i.e., the fact that for many DFAs the Hartree energy  $E_H$  is not canceled by  $E_{xc}$  if evaluated on one-electron ground-state densities,<sup>22</sup> is well-known to have a large impact on the quality of KS eigenvalues.<sup>8</sup> Second, many DFAs show a potential that features an incorrect long-range asymptotic behavior, again with negative effect on the interpretation of KS eigenvalues (see, e.g., Ref. 33 and references therein). Note that while the two issues are related in a physical sense, they are not the same and their connection in the construction of reasonable DFAs is far less obvious.<sup>34</sup> Another shortcoming affecting the KS eigenvalues is the deviation of the total energy curve,  $E(N)$ , as a function of the number of electrons,  $N$ , from piecewise linearity, for fractional  $N$  (see, e.g., Refs. 14 and 35–48). In the literature, this phenomenon is sometimes referred to as many-electron self-interaction<sup>35–37,39</sup> or as a (de-)localization error.<sup>38,41,46</sup> For (semi-)local functionals, one obtains a convex energy curve rather than a straight line. Such a deviation reflects negatively also on systems with an integer  $N$ : it leads to disagreement between the IPs predicted by the  $\Delta$ SCF method and those predicted via  $-\varepsilon_{ho}$ . This happens because the slope of the energy curve (to the left) equals, according to Janak’s theorem,<sup>49</sup> the ho KS eigenvalue. Therefore, even when one is interested only in closed systems, with an integer number of electrons, it is important to tackle the problem of lack of piecewise-linearity in order to obtain a physically meaningful value for  $\varepsilon_{ho}$ .

There exist many approaches to address the aforementioned shortcomings and obtain accurate results for the IP via  $\varepsilon_{ho}$ . Self-interaction correction<sup>22</sup> schemes lead to a significant improvement in the interpretation of KS eigenvalues.<sup>50</sup> Yet, their performance for ground-state energetics is debatable.<sup>51–55</sup> Approaches that approximate directly the xc potential<sup>56–59</sup> yield eigenvalues that satisfactorily reproduce the experimental IP, due to modified long-range properties of the potential. However, for these functionals total-energy related quantities are not accessible.<sup>60–62</sup> New types of GGAs can yield significantly improved potential properties,<sup>63–65</sup> but at the cost of being less accurate for total energies. Global hybrid functionals,<sup>66–71</sup> which linearly combine (semi-)local xc energy components, with a weight  $(1 - a)$ , and exact exchange (EXX, i.e., the Fock-integral evaluated with KS orbitals), with a weight  $a$ , mitigate the one-electron self-interaction error and often yield an excellent description of properties related to the total energy, for  $a \approx 0.25$ . However, since the self-interaction is only partly canceled, the KS potential falls off too quickly in the asymptotic limit, and  $-\varepsilon_{ho}$  is typically far from describing experimental IPs.

Nevertheless, with global hybrids it is possible to find a value of  $a$  such that the global hybrid will produce a piecewise linear energy curve and therefore an improved value for  $-\varepsilon_{ho}$ . This happens because for fractional  $N$  the non-local EXX component of the global hybrid produces a concave energy curve (see, e.g., Ref. 37), while the (semi-)

local components usually cause a convex energy curve, which therefore cancel each other. However, this cancellation is achieved with values of  $a \approx 0.75$ ,<sup>72–74</sup> which in most cases significantly compromises the performance of the functional for other quantities.<sup>33,70,71,75–77</sup>

This creates what we call the “parameter dilemma”: while an accurate description of *energy*-related quantities requires a certain value for the functional’s parameter, the accurate description of *potential*-related quantities requires a different value, and there is no value that provides a satisfactory description of both.<sup>78–80</sup>

Local hybrid functionals<sup>81–83</sup> aim at preserving the good energetics of global hybrid functionals while reducing the self-interaction error by introduction of a more flexible, space-dependent mixing of (semi-)local and non-local components (see Ref. 84 for an overview and discussion). However, we recently illustrated using a specially constructed local hybrid functional, termed ISOcc in the following, that the aforementioned “parameter dilemma” persists also there.<sup>85</sup> Similarly, in range-separated hybrids (RSHs) the values of the range-separation parameter have to be different to accurately reproduce, e.g., atomization energies and ionization potentials.<sup>29</sup>

It has been recently shown<sup>86,87</sup> that an alternative way to improve the prediction of the IP via  $-\varepsilon_{ho}$  is given by employment of the ensemble approach<sup>9,88–90</sup> in KS-DFT. This approach allows for the generalization of the Hartree and xc functionals for fractional  $N$  such that the piecewise linearity behavior of the total energy is restored, to a large extent. As a result, better correspondence of  $\varepsilon_{ho}$  to the experimental IP and to the  $\Delta$ SCF value is achieved, as demonstrated for the  $H_2$  molecule and the C atom with the LSDA.

Here, we employ the ensemble generalization procedure proposed in Ref. 87 to the Hartree and common approximate xc functionals, aiming to address the aforementioned challenge of simultaneous prediction of energy-related and potential-related properties with one DFA. Focusing on the prediction of the IP via the ho KS eigenvalue, we examine a variety of xc approximations: the local spin-density approximation, semi-local generalized gradient approximations, as well as global and local hybrids. Results for a representative test set of 26 light diatomic molecules and single atoms are presented. We find that the ensemble generalization systematically improves the prediction of the IP, for a wide variety of systems and xc functionals, changing the general tendency from under- to a small overestimation, compared to experiment. This improvement is achieved without any change in total energy-related properties. For hybrids that include a parameter, the IP obtained with ensemble-generalized functionals is found to be only weakly dependent on the parameter value, contrary to common experience with non-generalized hybrids. Thus, the ensemble approach eliminates one aspect of the “parameter dilemma.”

## II. THEORETICAL BACKGROUND

For completeness, we briefly present the ensemble generalization to the approximate Hartree-exchange-correlation (Hxc) density functional, focusing on its influence on the

highest occupied KS energy level,  $\varepsilon_{\text{ho}}$ . A complete derivation can be found in Refs. 86 and 87.

First, we formally consider a system with a fractional number of electrons,  $N = N_0 - 1 + \alpha$ , where  $N_0 \in \mathbb{N}$  and  $\alpha \in [0, 1]$ , so that  $\alpha = 1$  corresponds to a neutral and  $\alpha = 0$  to a singly ionized system. Subsequently, we take the limit  $\alpha \rightarrow 1^-$ , focusing on neutral systems with an integer number of electrons.

In a landmark article, Perdew *et al.* have shown that the zero-temperature ground state of an interacting many-electron system with fractional  $N$  should be described by an ensemble state.<sup>9</sup> This state is a linear combination of the pure ground states for  $N_0 - 1$  and  $N_0$  electrons, with the classical statistical weights of  $(1 - \alpha)$  and  $\alpha$ , respectively.<sup>135,136</sup> The ground-state energy of this ensemble state has then been shown to be equal to  $E(N) = (1 - \alpha)E(N_0 - 1) + \alpha E(N_0)$ , i.e., it is a piecewise-linear function of  $N$ .<sup>9</sup> This result is a general one, applying to any many-electron system. Therefore, in principle it trivially carries over to DFT; because if the *exact* exchange-correlation functional is used, DFT-based energies must reproduce the all-electron ones.

As mentioned in the Introduction, in practice *approximate* density functionals often deviate significantly from the exact functional when trying to describe a quantum system with fractional  $N$  in KS-DFT, and it has been traditionally assumed that this is just another manifestation of the approximate nature of the functional used. However, in Ref. 87 it was pointed out that much of this deviation is due to the fact that the pure-state exchange-correlation expression is used for both integer and fractional densities, whereas the fractional KS system must itself be in an ensemble state. This happens because the number of particles in the KS system equals the number of electrons in the real, interacting system and is also fractional. Therefore, the ground state of the KS system must also be expressed as an ensemble of the  $(N_0 - 1)$ - and  $N_0$ -KS states, obtained from the same KS potential, even if one uses an approximate functional. Reference 87 therefore suggested that any approximate Hartree-exchange-correlation functional can be generalized for an ensemble ground state using ensemble state theory<sup>9,88-90</sup> (for other recent uses of the ensemble approach, see Refs. 46, 91, and 92). Performing an ensemble average of the many-electron Coulomb operator  $\hat{W} = \frac{1}{2} \sum_i \sum_{j \neq i} |\mathbf{r}_i - \mathbf{r}_j|^{-1}$  in the KS system, it has been found that the pure-state Hxc energy functional can be generalized to ensemble states in the following form:

$$E_{\text{e-Hxc}}[n^{(\alpha)}] = (1 - \alpha)E_{\text{Hxc}}[\rho_{-1}^{(\alpha)}] + \alpha E_{\text{Hxc}}[\rho_0^{(\alpha)}], \quad (1)$$

which is exact for the Hartree and exchange components and approximate for the correlation. Here, the index e- indicates that the functional is ensemble-generalized,  $E_{\text{Hxc}}$  is the pure-state Hxc functional,  $\rho_p^{(\alpha)}(\mathbf{r})$  is defined as the sum of the first  $N_0 + p$  KS orbitals squared:  $\rho_p^{(\alpha)}(\mathbf{r}) = \sum_{i=1}^{N_0+p} |\varphi_i^{(\alpha)}(\mathbf{r})|^2$ , where  $p = -1$  or  $0$ , and  $n^{(\alpha)}(\mathbf{r}) = (1 - \alpha)\rho_{-1}^{(\alpha)}(\mathbf{r}) + \alpha\rho_0^{(\alpha)}(\mathbf{r})$  is the ensemble-state electron density. When  $N$  is an integer, i.e.,  $\alpha$  assumes the value of 0 or 1, the Hxc energy reduces to that obtained from the underlying pure-state Hxc functional. Therefore, ensemble-generalization does not affect the total energy at integer  $N$ . The generalization in Eq. (1) is applicable

to any xc functional and makes the Hartree and the xc energy components *explicitly* linear in  $\alpha$ . However, there may still remain an *implicit* non-linear dependence of  $E_{\text{e-Hxc}}[n^{(\alpha)}]$  on  $\alpha$ , because the KS orbitals themselves,  $\varphi_i^{(\alpha)}(\mathbf{r})$ , and consequently  $\rho_p^{(\alpha)}(\mathbf{r})$  and  $E_{\text{Hxc}}[\rho_p^{(\alpha)}]$ , may depend on  $\alpha$ . The dependence of  $\varphi_i^{(\alpha)}(\mathbf{r})$  on  $\alpha$  arises from the fact that the KS orbitals are expected to relax as one varies  $\alpha$  from 0 (positive ion) to 1 (neutral system).<sup>44,87,92</sup>

Importantly, Eq. (1) is derived by considering the generalization of pure-state functionals to ensemble states, without assuming anything *a priori* about piecewise-linearity, because it also applies to approximate exchange-correlation functionals. Nevertheless, in Refs. 87 and 93 it has been shown that by employing Eq. (1) the energy curve  $E(N)$  satisfies the piecewise-linearity criterion much more closely, being slightly concave. The concavity is related to the above mentioned implicit non-linear dependence of the energy on  $\alpha$ . Another perspective on this approximate piecewise-linearity can be obtained from the fact that Eq. (1) can be derived, with some further approximations, from different schemes that attempt to enforce piecewise-linearity explicitly.<sup>14,44,47,48,92,94</sup>

Due to the fact that the slope of  $E(N)$  changes for all  $\alpha$ , including  $\alpha \rightarrow 1^-$ , it follows from Janak's theorem,<sup>49,137</sup> which identifies  $\partial E/\partial N$  with  $\varepsilon_{\text{ho}}$ , that the ho energy level has to change, too, even for a system with an integer  $N$ . This change is obtained in practice from an ensemble generalization of the KS potential, as explained below.

The KS potential is expressed as  $v_{\text{e-KS}}(\mathbf{r}) = v_{\text{ext}}(\mathbf{r}) + v_{\text{e-Hxc}}[n](\mathbf{r})$ , where  $v_{\text{ext}}(\mathbf{r})$  is the external potential and  $v_{\text{e-Hxc}}[n](\mathbf{r}) := \delta E_{\text{e-Hxc}}/\delta n(\mathbf{r})$  is the ensemble-generalized Hxc potential. In the limit  $\alpha \rightarrow 1^-$ , this potential reduces to a sum of two terms:  $v_{\text{e-Hxc}}[n](\mathbf{r}) = v_{\text{Hxc}}[n](\mathbf{r}) + v_0[n]$  – the usual pure-state Hxc potential,  $v_{\text{Hxc}}[n](\mathbf{r})$ , and a spatially uniform term,  $v_0[n]$ , which can be written as<sup>87</sup>

$$v_0[n] = E_{\text{Hxc}}[n] - E_{\text{Hxc}}[n - |\varphi_{\text{ho}}|^2] - \int |\varphi_{\text{ho}}(\mathbf{r})|^2 v_{\text{Hxc}}[n](\mathbf{r}) d^3r. \quad (2)$$

Here and below, the superscript  $(\alpha)$  is dropped at the limit  $\alpha \rightarrow 1^-$  for brevity. Note that the ensemble-generalized KS potential does not vanish at  $r \rightarrow \infty$ , but asymptotically approaches  $v_0[n]$ . We stress that  $v_0[n]$  is a well-defined, rather than arbitrary, potential shift. It *must* be taken into account for the ensemble-generalized functional in order for the ho KS eigenvalue to equal  $\partial E/\partial N$ , i.e., to obey Janak's theorem. Note that the shift discussed here is different from the one recently proposed by Zahariev and Levy.<sup>95</sup> As clarified in Ref. 92, in Ref. 95 the potential shift makes the energy of the KS system equal the energy of the interacting system. Here, however, the shift emerges naturally from the ensemble treatment and is essential to obtaining results that are consistent with Janak's theorem. Also note that the result above has been presented in a spin-independent form for simplicity; in practice, in spin-dependent calculations, there exist potential shifts  $v_0^\sigma$  to both spin channels  $\sigma = \uparrow, \downarrow$ . Calculating  $v_0^\sigma$  with Eq. (2), we take the ho level to be the highest occupied level in the  $\sigma$ -channel considered (noted as  $\sigma - \text{ho}$ ). In the following, however, if not stated explicitly otherwise, when mentioning the ho level we refer to the



global ho:  $\varepsilon_{\text{ho}} = \max_{\sigma} \varepsilon_{\text{ho}}^{\sigma}$ , i.e., the one of the two  $\sigma$ -ho levels which is higher in energy; the same applies for the ensemble-generalized ho level,  $\varepsilon_{\text{e-ho}}$ .

To summarize, as a result of the approximate ensemble generalization of the Hxc functional (Eq. (1)),<sup>86,87</sup> in the limit of integer  $N$  the KS potentials exhibit spatially uniform shifts  $v_0^{\sigma}$ , such that all KS eigenvalues of the same spin channel are shifted by the same value (see, e.g., Fig. 4 in Ref. 93). The KS orbitals, and as a result the density and the total energy, are not changed and remain the same as those obtained with the underlying Hxc functional. Furthermore, because all eigenvalues are shifted by the same amount, eigenvalue differences (as well as quantities based on them, e.g., in linear response time-dependent DFT<sup>96</sup>) are not affected either. Therefore, the  $\sigma$ -ho energy levels of the ensemble-generalized functional can be expressed as  $\varepsilon_{\text{e-ho}}^{\sigma} = \varepsilon_{\text{ho}}^{\sigma} + v_0^{\sigma}$ , being a sum of the  $\sigma$ -ho level that emerges from a standard KS-DFT calculation prior to the ensemble generalization and the potential shift of the relevant spin channel, calculated according to Eq. (2). Comparing both  $\varepsilon_{\text{e-ho}}$  and  $\varepsilon_{\text{ho}}$  to experimental IPs and  $-I_{\text{ASCF}}$  is the main subject of Sec. IV.

### III. COMPUTATIONAL DETAILS

We concentrate on a relatively elementary, yet chemically representative, set of systems, consisting of 18 light diatomic molecules: H<sub>2</sub>, LiH, Li<sub>2</sub>, LiF, BeH, BH, BO, BF, CH, CN, CO, NH, N<sub>2</sub>, NO, OH, O<sub>2</sub>, FH, F<sub>2</sub>, and their 8 constituent atoms. The simplicity of the systems allows us to keep computational costs low and to refrain from introducing additional sources of error, e.g., searching for an optimal geometry in systems with many degrees of freedom. At the same time, systems of single-, double-, and triple-bond molecules as well as atoms (no bonding) are included in the test set, which makes the set representative of more complicated systems, as shown in previous work (see, e.g., Refs. 27 and 85).

All calculations were performed using the program package DARSEC,<sup>97,98</sup> an all-electron code that allows for electronic structure calculations of single atoms or diatomic molecules on a real-space grid represented by prolate-spheroidal coordinates. DARSEC allows one to solve the KS equations self-consistently for explicitly density-dependent, as well as orbital-dependent, functionals. For the latter, a local, multiplicative xc potential is obtained by employing the KLI<sup>99</sup> approximation to the optimized effective potential (OEP)<sup>100–102</sup> formalism. Use of this approximation has been justified in Ref. 103 for the EXX functional and in Ref. 85 for the ISOcc local hybrid functional.

For all systems, an accuracy of 0.0005 hartree in the total energy and in the ho KS eigenvalue has been achieved by appropriately choosing the parameters of the real-space grid and by iterating the self-consistent DFT cycle. For molecules the bond length was taken from experiment.<sup>104,105</sup> Differences due to atomic relaxation were found to be insignificant.<sup>138,139</sup> The net spin of the neutral systems was also taken to be as in experiment. The spin configuration of cations (used below for calculating ionization potentials from total energy differences) was obtained by removing an electron from the highest occupied orbital of the neutral.

## IV. RESULTS

### A. Effect of the ensemble correction—O<sub>2</sub> as a prototypical case

Previous work<sup>87,93</sup> has already demonstrated that the ensemble generalization of Eq. (1) significantly reduces the deviation from the piecewise-linearity condition for the total energy, i.e., greatly diminishes the delocalization error, and as a consequence eliminates the fractional dissociation error in diatomic molecules. Here, we focus on the potential shifts (Eq. (2)) that emerge from the ensemble-generalization and their effect on the Kohn-Sham energy levels. In particular, we consider the prediction of the IP via  $\varepsilon_{\text{e-ho}}$ .

For a clear understanding of the results presented in this paper, it is of advantage to first illustrate the effect of the potential shift mechanism, given by Eq. (2), on the eigenvalue structure of a particular system with an integer number of electrons. Here, we provide a detailed presentation of a selected system—the O<sub>2</sub> molecule, computed with the Perdew-Burke-Ernzerhof (PBE) GGA<sup>27</sup> at its experimental bond length of 2.2819 bohr.

Due to its electronic ground-state configuration,  ${}^3\Sigma_g^-$ , this system must be treated in a spin-polarized formalism. Consequently, it provides an interesting example for how eigenvalues belonging to different spin channels are shifted when the corresponding ensemble potential shift,  $v_0^{\sigma}$ , is applied.

The positions of the highest occupied ( $\varepsilon_{\text{ho}}^{\sigma}$ ) and lowest unoccupied ( $\varepsilon_{\text{lu}}^{\sigma}$ ) KS eigenvalues for both spin channels are depicted in Fig. 1. The eigenvalues changed by the respective potential shift, i.e.,  $\varepsilon_{\text{e-ho}}^{\sigma} = \varepsilon_{\text{ho}}^{\sigma} + v_0^{\sigma}$  and  $\varepsilon_{\text{e-lu}}^{\sigma} = \varepsilon_{\text{lu}}^{\sigma} + v_0^{\sigma}$ , as well as the negative of the experimental IP,  $-I_{\text{exp}}$ , are also included in the figure.

It can be readily observed that the unshifted highest occupied eigenvalue of the up channel,  $\varepsilon_{\text{ho}}^{\uparrow} = -0.251$  hartree, which lies higher than its spin down counterpart, poorly reproduces the negative of the experimental IP of the O<sub>2</sub> molecule. In fact, with PBE it underestimates the experimental IP of  $I_{\text{exp}} = 0.453$  hartree<sup>105</sup> by 45%, a value that is quite typical for other systems as well. However, after application of the potential shift, the highest occupied eigenvalue is  $\varepsilon_{\text{e-ho}}^{\uparrow} = -0.526$  hartree, i.e., the experimental IP is now *overestimated* by 16%. As shown below, this is a typical result also for other systems *and* other functionals.

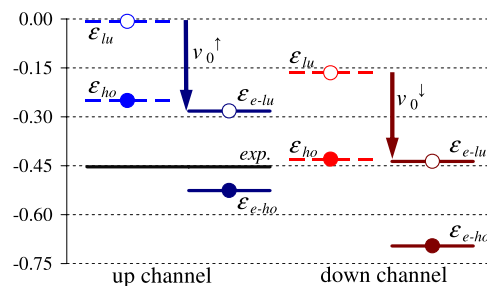


FIG. 1. Diagram of the highest occupied and lowest unoccupied KS-PBE eigenvalues of the O<sub>2</sub> molecule, for both spin channels, before and after applying the potential shifts of Eq. (2), along with the negative of the experimental IP. All values are in hartree.

From the results presented for the ensemble generalized PBE functional, the question of how other DFAs perform for the same system naturally arises. In particular, the change of the eigenvalues obtained with functionals containing a varying amount of non-local EXX is of great interest, as we know that a greater percentage of EXX already leads to a more accurate description of IPs via the highest occupied eigenvalue.

Fig. 2 provides a comparison of the unshifted and shifted  $\epsilon_{\text{ho}}$  eigenvalues to the experimental IP for  $\text{O}_2$ . Besides (semi-)local functionals such as the LSDA, PBE, and BLYP,<sup>106–108</sup> we also ensemble-generalized the global hybrid functionals B3LYP<sup>69</sup> and PBEh( $a$ )<sup>70</sup> (employed within the KS scheme using the KLI approximation), with  $a$  denoting the fixed amount of EXX combined with  $(1 - a)$  of PBE exchange and with full PBE correlation.

The unshifted eigenvalues for the three purely (semi-)local functionals (LSDA, PBE, BLYP) underestimate the IP by  $\sim 45\%$ . After ensemble generalization, we observe an overestimation by  $\sim 16\%$ . It is instructive to check to which extent this overestimation comes from errors in calculating total energy differences that are inherent to the underlying functional, and to which extent they arise from the ensemble generalization process.<sup>140</sup> Therefore, we compare the shifted and unshifted eigenvalues to  $-I_{\Delta\text{SCF}}$ . For  $\text{O}_2$  computed with PBE, one obtains  $I_{\Delta\text{SCF}} = 0.464$  hartree, which deviates from experiment by only 2.3%. With respect to this quantity, the unshifted eigenvalue yields an underestimate of 46%, while the shifted value overestimates it by 14%. We therefore realize that most of the discrepancy comes from the concavity that remains in the  $E(N)$  curve even after ensemble generalization.

In the global hybrid, PBEh( $a$ ), increasing the intrinsic amount of EXX significantly improves the correspondence of the unshifted eigenvalue to experiment. Due to the fact that the KS potential decays asymptotically more slowly with a growing value of  $a$ , the IP via  $-\epsilon_{\text{ho}}$  is very sensitive to non-local functional components included. Changing from under- to overestimation, an optimal description of  $I_{\text{exp}}$  is reached for this system with  $a \approx 0.6$ . However, for ensemble generalized DFAs the value of  $-\epsilon_{e-\text{ho}}$  systematically overestimates the IP with respect to experiment for the  $\text{O}_2$  molecule, regardless of

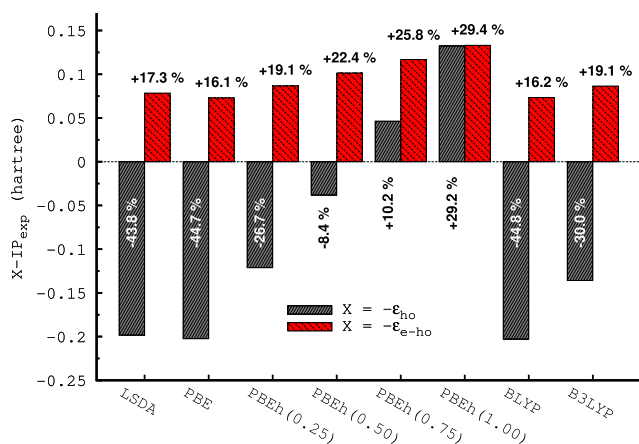


FIG. 2. Comparison of  $-\epsilon_{\text{ho}}$  and  $-\epsilon_{e-\text{ho}}$  to the experimental IP of the  $\text{O}_2$  molecule, calculated with different DFAs. The corresponding labels provide the relative deviation in percent.

the value of  $a$ , while at the same time being far less sensitive to the amount of non-locality in the functional expression. While for “plain” PBE the relative error now reads  $\sim +16\%$ , it increases to  $+29\%$  when full non-local exchange combined with PBE correlation is used. The reason for this reduced sensitivity lies in the following mechanism: while the absolute value of  $\epsilon_{\text{ho}}$  grows with increasing  $a$ , the potential shift  $v_0$  is reduced, roughly commensurately, because the Hartree+EXX functional has zero potential shift.<sup>87</sup>

## B. Evaluating the test set—A systematic study

Following the illustration of the mechanism of the potential shift for a single system, we now focus on the mean discrepancy in the evaluation of the experimental IP via shifted and unshifted KS eigenvalues, for a variety of functionals.<sup>109</sup> We use the test set of systems introduced in Sec. III as a basis for averaging.

We emphasize that the eigenvalues’ shift is expected to improve the correspondence between the negative of the  $\epsilon_{\text{ho}}$  eigenvalue and the ionization energy obtained via the  $\Delta\text{SCF}$  method, for a given DFA. We compare the shifted and unshifted eigenvalues to experiment, and not to  $\Delta\text{SCF}$  values, relying on the aforementioned fact that the  $\Delta\text{SCF}$  reliably describes systems of our test set, with small average relative errors: 3.4% for PBE and 4.2% for both the LSDA and ISOcc( $c = 0.5$ ).

We define the averaged relative error in the ionization potential,

$$\delta_{\text{IP}} = \sqrt{\frac{1}{M} \sum_{j=1}^M \left( \frac{-\epsilon^{(j)} - I_{\text{exp}}^{(j)}}{I_{\text{exp}}^{(j)}} \right)^2}. \quad (3)$$

Here, the index  $j$  runs over all systems in the test setup to the total number  $M = 26$ , and  $\epsilon$  stands either for the shifted ( $\epsilon_{e-\text{ho}}$ ) or unshifted ( $\epsilon_{\text{ho}}$ ) highest occupied KS eigenvalue.

Note that in Eq. (3) the unsigned deviation from experimental IPs is employed to avoid a misleading result of zero average relative error, which emerges when there occurs an overestimation for some systems and an underestimation for others. However, in order to be able to distinguish between systematic over- or underestimation, an additional measure is defined accordingly as

$$S = \frac{1}{M} \sum_{j=1}^M \text{sgn}(-\epsilon^{(j)} - I_{\text{exp}}^{(j)}). \quad (4)$$

While  $\delta_{\text{IP}}$  provides the mean deviation from experimental values in %, the quantity  $S$  indicates the average trend of the prediction, being naturally confined to the interval  $[-1, 1]$ . Namely, for a systematic overestimation, we obtain  $S = 1$ , and for a systematic underestimation,  $S = -1$ . Both quantities,  $\delta_{\text{IP}}$  and  $S$ , were obtained for various DFAs and their ensemble-generalized counterparts.

Fig. 3 shows the corresponding results for the LSDA, the semi-local PBE and BLYP, the global hybrid functionals B3LYP and PBEh(0.25), the EXX, and the ISOcc(0.5) local hybrid functional. Note that for EXX the results for the regular and ensemble-generalized functional coincide, because the

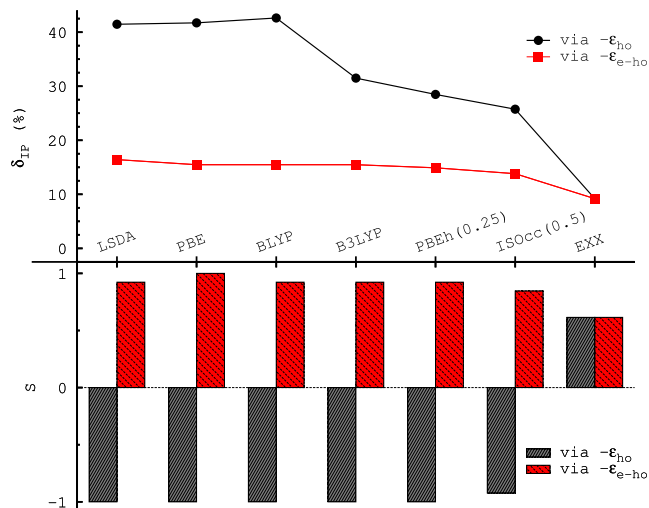


FIG. 3. Average relative error  $\delta_{IP}$  in % (upper axis) and signum function  $S$  (lower axis) for the LSDA, PBE, BLYP, B3LYP, PBE( $a = 0.25$ ), ISOcc( $c = 0.5$ ), and pure EXX (black) as well as their ensemble-generalized versions (red).

Hartree+EXX functional exhibits a zero potential shift.<sup>87,141</sup> Fig. 4 provides the corresponding results for the PBEh( $a$ ) global hybrid functional as a function of the parameter  $a$ , i.e., on various amounts of non-local EXX.<sup>142</sup> The value  $a = 0$  in this figure reproduces the PBE result. Fig. 5 depicts  $\delta_{IP}$  and  $S$  obtained with the local hybrid ISOcc. The latter functional was developed using the so-called local mixing function rather than a fixed mixing ratio of non-local and semi-local components. It contains a free parameter  $c$ , which implicitly determines the intrinsic amount of EXX included in the local hybrid. Higher  $c$  values correspond to a higher fraction of EXX being included. Consequently, for ISOcc both the quantities given by Eqs. (3) and (4) are functions of this parameter  $c$ , in analogy to the global hybrid PBEh( $a$ ).

In principle, both parameters  $a$  and  $c$ , in PBEh( $a$ ) and ISOcc( $c$ ), respectively, are free. However, as mentioned earlier, it is known that in terms of total energy related

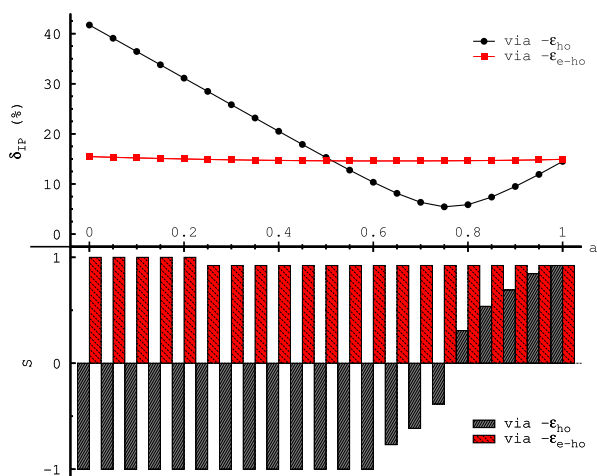


FIG. 4. Average relative error  $\delta_{IP}(a)$  in % (upper axis) and signum function  $S(a)$  (lower axis) for PBEh( $a$ ) (black) and e-PBEh( $a$ ) (red) as a function of the parameter  $a$ .

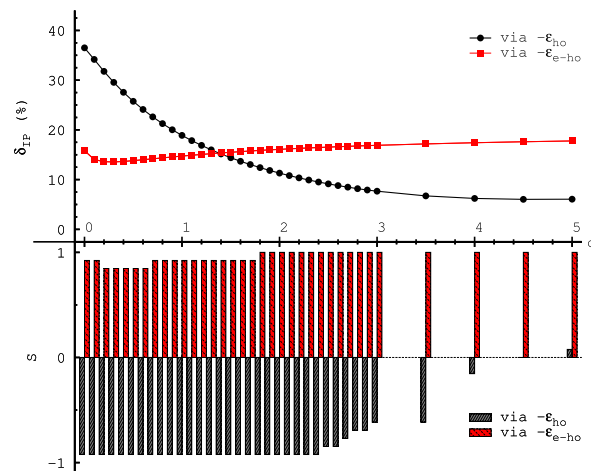


FIG. 5. Average relative error  $\delta_{IP}(c)$  in % (upper axis) and signum function  $S(c)$  (lower axis) for ISOcc( $c$ ) (black) and e-ISOcc( $c$ ) (red) as a function of the parameter  $c$ .

quantities, PBEh( $a$ ) performs best for  $a = 0.25$ , while we recently showed that for the ISOcc( $c$ ) functional, the optimal parameter value is  $c = 0.5$ . Therefore, both functionals, using their optimal respective parameters, play a special role in the following discussion, and Fig. 3 shows their performance in comparison to the other DFAs.

Fig. 3 clearly indicates that using the unshifted eigenvalues  $\epsilon_{ho}$ , the three (semi-)local functionals LSDA, PBE, and BLYP strongly and systematically underestimate  $\delta_{IP}$  by  $\approx 41\% - 43\%$ . Regarding hybrids, for the global hybrid B3LYP, we obtain an underestimation of 31%, for PBEh( $a = 0.25$ ) 28%, and for the local hybrid ISOcc( $c = 0.5$ ) 26%. The improvement of hybrids over (semi-)local functionals is explained by the fact that the non-local terms in hybrids lead to a partial cancellation of the self-interaction error and an improved behavior of the xc potential in the asymptotic limit.

If the parameters  $a$  and  $c$  are varied, Figs. 4 and 5 illustrate that when using  $\epsilon_{ho}$ , the global hybrid PBEh( $a$ ) and the local hybrid ISOcc( $c$ ) show a transition in their parameter-dependent  $S$ -function from negative to positive values. This feature clearly indicates that, for the systems studied here, it is possible to fit the corresponding functional parameter for a given system so that  $\epsilon_{ho}$  exactly gives the experimental IP. If  $a$  and  $c$  are optimized to reduce the error  $\delta_{IP}$ , we obtain an underestimation of 5% for  $a = 0.75$  in PBEh and of 6% for  $c = 4.5$  in ISOcc. Therefore, by changing the parameters  $a$  and  $c$ , we are able to strongly reduce the average error in the IP of our test set. However, this comes at a price in total energy-related quantities, as has been shown in Refs. 72 and 85, and is the subject of the so-called “parameter dilemma” presented in Sec. I.

When using the ensemble-corrected highest occupied eigenvalues  $\epsilon_{e-ho}$ , we obtain a completely different picture. First, the systematic underestimation now changes to an overestimation. All of the aforementioned functionals now show a very similar average error of  $\delta_{IP} \approx 14\% - 17\%$ , which is significantly smaller than the results from non-generalized functionals. Second, for ensemble-generalized hybrid functionals e-PBEh( $a$ ) and e-ISOcc( $c$ ), there is no

transition from an underestimation to overestimation regime, but rather a systematic overestimation of the IP, independent of the parameter value. In other words, the amount of non-locality included in the hybrid functional plays a minor role in the description of IPs via shifted KS eigenvalues, in contrast to their unshifted counterparts. This confirms that the mechanism of cancellation between the change in the potential shift of Eq. (2) and the highest occupied eigenvalue with a varying amount of non-locality is not particular to the O<sub>2</sub> molecule, but rather a systematic feature of ensemble-generalized functionals. Furthermore, in the ensemble-generalized version of DFAs, the “parameter dilemma” does not emerge: since the ensemble-generalized eigenvalues describe IPs with an accuracy almost independent of the amount of EXX included, one cannot deduce a preferred value of the parameter by minimizing  $\delta_{IP}$ . Therefore, in principle one could use the functional with the parameter optimized to describe binding processes and structural quantities and rely on the description of IPs via the shifted eigenvalues  $\varepsilon_{e-ho}$ . In this case, our results for  $\delta_{IP}$  using thermochemically optimized functionals with ensemble-generalization (such as e-B3LYP, e-PBEh( $a = 0.25$ ), and e-ISOcc( $c = 0.5$ )) indicate a clear improvement over their non-generalized counterparts.

Our results further indicate that even functionals whose xc terms were constructed on different grounds and from different perspectives, such as, for example, the PBE and BLYP functional, yield similar values of roughly  $\delta_{IP} \approx 15\%$  after applying the ensemble generalization. As even the inclusion of non-local components does not lead to significant change, one might wonder if this “natural border” of 15% is inherent to the ensemble shift mechanism *regardless* of the specific form of the respective DFA put to task. This question has been checked by varying the parameters  $\mu$  and  $\kappa$  used in the construction of the PBE exchange functional.<sup>27,110</sup> We find that for different choices of  $\mu$  and  $\kappa$  one obtains different values for the average relative error  $\delta_{IP}$ . For instance, using PBE exchange with a value of  $\mu = 1.0$  together with the original  $\kappa = 0.8401$  results in an error of  $\delta_{IP} = 20\%$  when using  $\varepsilon_{e-ho}$ , while a combination of the original  $\mu = 0.21951$  and  $\kappa = 5.0$  leads to  $\delta_{IP} = 8\%$ . From this we conclude that the ensemble-generalization as such does not lead to a fixed systematic error in the description of experimental IPs via KS eigenvalues. However, the results of this subsection suggest that after the ensemble generalization the functionals examined here have a common missing part, which causes the described discrepancy in  $\delta_{IP}$ .

Before concluding this subsection, we note that while we have focused our work on the IP of neutral atoms and diatomic molecules, IPs of ions may in principle be assessed in the same manner. In particular, the electron affinity (EA) of the neutral can be explored as the IP of the singly charged anion (barring geometrical relaxation). Unfortunately, for the atoms and very small molecules studied here, it is well-known<sup>111–113</sup> that with common semi-local approximations, negative ions of small systems may erroneously be predicted to be unstable. However, when performing calculations with finite basis sets, as in, e.g., Ref. 114, unbound states can be artificially stabilized,<sup>115</sup> because the basis set effectively confines the unbound electron to the vicinity of the neutral

system. Because the ensemble generalization discussed here does not change the total energies of systems with integer electrons (including neutrals, cations, and anions), anions that are not bound with the underlying xc functional will remain unbound even if its ensemble-generalized version is employed. Furthermore, although ensemble-generalization will generally shift the energy levels, including the unoccupied ones, the question of whether the lowest unoccupied KS orbital has a bound or unbound character will not be affected,<sup>93</sup> because orbitals are unchanged by a uniform shift of the potential.

### C. Ensemble-generalization and the *Aufbau* principle

In general, at zero temperature the energy levels in the KS system have to be occupied according to the *Aufbau* principle, i.e., the levels are occupied without “holes,” starting with the lowest ones up. In the following, we term such an occupation *proper*. An example for a proper occupation is given in Fig. 1 for the O<sub>2</sub> molecule. All calculations performed for this work, except for a few discussed below, yield proper occupation.

In spin-polarized calculations a special situation can occur, when each of the spin channels is occupied properly itself, while the system as a whole possesses a “hole” in its occupation. For example, this happens when the lu level of the  $\downarrow$ -channel appears lower than the ho level of the  $\uparrow$ -channel. An occupation of this kind is termed *proper in a broad sense*. It is emphasized here that a broad-sense-proper density obeys all the required restrictions related to a rigorous definition and differentiability of energy functionals;<sup>116,117</sup> therefore it can serve as a legitimate solution of a many-electron system. In the past, broad-sense-proper occupations have been observed in certain transition-metal and lanthanide atoms and ions in LSDA and PBE calculations,<sup>24</sup> as well as in the Li atom with the EXX<sup>118</sup> and with the exact KS potential, which has been obtained from accurate wave-function-methods based spin densities.<sup>119</sup> The latter result strengthens our understanding that a broad-sense-proper occupation is not necessarily an artifact of some DFAs, but rather is an expected result, because it may appear even with the exact functional.

In the current work we find that the ensemble generalization, by means of the potential shifts  $v_0^\sigma$ , yields broad-sense-proper results for systems which appeared strictly proper before. Figure 6 illustrates the situation for the Li

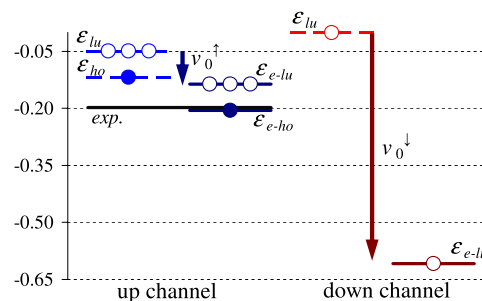


FIG. 6. Diagram of the highest occupied and lowest unoccupied KS eigenvalues of the Li atom, for both spin channels, before and after application of the potential shifts of Eq. (2), obtained within the PBE functional, along with the negative of the experimental IP. All values are in hartree. The highest occupied eigenvalue at the  $\downarrow$ -channel is lower than  $-0.65$  hartree and is therefore not shown for clarity.



atom calculated with PBE and e-PBE. Due to the fact that  $v_0^\uparrow = -0.087$  hartree, while  $v_0^\downarrow = -0.603$  hartree, the  $\downarrow$ -e-lu level appears below the  $\uparrow$ -e-ho level, causing a broad-sense-proper occupation. The significant differences in the values of the two potential shifts are associated with the different nature of the  $\sigma$ -ho orbitals: the  $\uparrow$ -ho orbital is a relatively delocalized, high-lying  $2s$  orbital, whereas the  $\downarrow$ -ho orbital is a localized, low-lying  $1s$  orbital. A similar situation is observed in the Na atom, for which we obtained a broad-sense-proper result too. To summarize, we view the appearance of broad-sense-proper occupations in the ensemble treatment of the alkaline atoms Li and Na as another feature of the exact DFT result, which has been recovered by the ensemble generalization.

#### D. The derivative discontinuity (DD) in ensemble-generalized functionals

While in the current work we are concerned primarily with the IP of atoms and molecules, it is worth discussing a related quantity—the (fundamental) gap,  $E_g$ . By definition,  $E_g = I - A$ , where  $A$  is the electron affinity, i.e., the energy gained by adding one electron to the system. As opposed to the IP, the gap of the interacting system does not equal the gap of the KS system,  $E_g^{\text{KS}} = \varepsilon_{\text{lu}} - \varepsilon_{\text{ho}}$ , even for the exact xc functional. Instead,  $E_g = E_g^{\text{KS}} + \Delta$ , where  $\Delta$  is the DD—a “jump” experienced by the KS potential when it is varied with respect to  $N$ , and  $N$  crosses an integer value.<sup>9,28,120–126</sup>

There exist several ways to find the DD. First, for finite systems it can be obtained using total energy differences,

$$\Delta_E = E(N_0 + 1) - 2E(N_0) + E(N_0 - 1) - E_g^{\text{KS}}. \quad (5)$$

Second, the DD can be obtained as suggested in Refs. 127 and 128,

$$\Delta_{\text{OEP}} = \langle \varphi_{\text{lu}} | u_{\text{xc},\text{lu}} | \varphi_{\text{lu}} \rangle - \langle \varphi_{\text{lu}} | v_{\text{xc}} | \varphi_{\text{lu}} \rangle, \quad (6)$$

where  $u_{\text{xc},i}(\mathbf{r}) := \varphi_i^{-1}(\mathbf{r}) \delta E_{\text{xc}} / \delta \varphi_i(\mathbf{r})$ , i.e., the orbital-specific xc potential of the  $i$ th orbital and  $v_{\text{xc}} := \delta E_{\text{xc}} / \delta n$ , i.e., the local xc potential, which in general has to be obtained via the OEP procedure<sup>100–102</sup> (hence the index OEP). The derivation of  $\Delta_{\text{OEP}}$  assumes the “alignment equality”  $\langle \varphi_{\text{ho}} | u_{\text{xc},\text{ho}} | \varphi_{\text{ho}} \rangle = \langle \varphi_{\text{ho}} | v_{\text{xc}} | \varphi_{\text{ho}} \rangle$ , which determines the free constant in  $v_{\text{xc}}(\mathbf{r})$  as part of the OEP procedure. Additional approaches to introduce the derivative discontinuity include Refs. 129–133 (see also Refs. 41, 46, 86, and 134 for an overview).

Finally, an approximation for the DD can be obtained from an ensemble treatment for a given underlying Hxc functional, as proposed in Ref. 86,

$$\Delta_{\text{ens}} = E_{\text{Hxc}}[n + |\varphi_{\text{lu}}|^2] - 2E_{\text{Hxc}}[n] + E_{\text{Hxc}}[n - |\varphi_{\text{ho}}|^2] + \int d^3r v_{\text{Hxc}}[n](\mathbf{r}) (|\varphi_{\text{ho}}(\mathbf{r})|^2 - |\varphi_{\text{lu}}(\mathbf{r})|^2). \quad (7)$$

The first way requires three independent self-consistent calculations of the total energy (hence the index E): of the neutral system, the cation, and the anion. In contrast, the second and third ways yield the DD from KS quantities of the neutral system only, which is an advantage when considering infinite systems.

Relying on our experience with ensemble-generalized calculations for atoms and small molecules (Ref. 87 and

this work), we expect  $\Delta_{\text{ens}}$  obtained with an approximate xc functional to be larger than  $\Delta_E$ . As has been shown in Fig. 2 of Ref. 86 (lower panel),  $\varepsilon_{\text{e-ho}}$  is obtained as being somewhat too low immediately to the left of an integer  $N$  and somewhat too high immediately to the right of it. As a result,  $\Delta_{\text{ens}}$  overestimates the true discontinuity. This overestimate is related to the residual concavity of the  $E(N)$  curve after the ensemble generalization. In the current study, we showed that the overestimate in  $\varepsilon_{\text{e-ho}}$  to the left of the integer point, which corresponds to the negative of the IP, is systematic, i.e., it happens in various systems and with different functionals. Consequently, we expect a systematic overestimate for  $\Delta_{\text{ens}}$  and the resulting  $E_g$ .

The discrepancy between  $\Delta_{\text{OEP}}$  and  $\Delta_{\text{ens}}$  has a different origin. While  $\Delta_{\text{OEP}}$  originates because the KS potentials are differently “aligned” (see above) to the left and to the right of an integer point,  $\Delta_{\text{ens}}$  comes from two sources (see Ref. 86 for detailed explanations): the first is the same as for  $\Delta_{\text{OEP}}$ ; the second is the fact that the ensemble-generalized KS potential does not approach zero at  $r \rightarrow \infty$ , but rather a constant  $v_0$  (see Eq. (2)), which is different to the left and to the right of an integer  $N$ .  $\Delta_{\text{OEP}}$  does not consider the second source described above, assuming (correctly in the context of Refs. 127 and 128) that the potentials asymptotically tend to zero. In fact, it can be analytically shown that  $\Delta_{\text{OEP}}$  is an ingredient in  $\Delta_{\text{ens}}$ , which was denoted by  $\Delta_1$  in Ref. 86.

#### V. CONCLUSIONS AND SUMMARY

In the current work, we employed the ensemble-generalization procedure<sup>87</sup> for a test set of 26 diatomic molecules and single atoms, for a variety of xc functionals. These include the LSDA, the semi-local PBE and BLYP, the global hybrids B3LYP and PBEh(a), and the local hybrid ISOcc(c). We focused on the prediction of the IP via the highest occupied KS eigenvalue,  $\varepsilon_{\text{e-ho}}$ .

We found that implementing the ensemble approach improves, on average, the correspondence of  $\varepsilon_{\text{e-ho}}$  with the experimental IP for all xc functionals considered, changing the general tendency in the IP prediction from a gross underestimation to a smaller overestimation.

For functionals that include a parameter, namely, the hybrids PBEh(a) and ISOcc(c), we observed a rather weak dependence of  $\varepsilon_{\text{e-ho}}$  on the respective functional parameter, while yielding a roughly constant overestimation to the IP, with respect to experiment. This eases the so-called “parameter dilemma”: there are no two optimal values of the functional’s parameter originating from fitting to total energy-related quantities as opposed to fitting potential-related quantities. Instead, the parameter can be determined relying on energetics only, because of its weak influence on the value of  $\varepsilon_{\text{e-ho}}$ . Indeed, the average relative error in the ionization potential,  $\delta_{\text{IP}}$ , equals approximately 15% for all ensemble-generalized xc functionals, as can be seen from Fig. 3. Surprisingly, such features of the underlying xc functional, as being local (LSDA), semi-local (PBE, BLYP), or non-local (B3LYP, PBEh, ISOcc), relying on features of the homogeneous electron gas (LSDA, PBE, PBEh, ISOcc) or not (BLYP, B3LYP), are of little relevance with respect to the

IP prediction, once the functional is used in the ensemble-generalized form. We therefore conclude that upon ensemble generalization (Eq. (1)) all the functionals we tested share the same deficiency. It is most probably related to the remaining concavity of the  $E(N)$  curve, due to the implicit dependence of the KS orbitals on  $\alpha$ . Therefore, future improvement in the IP prediction via  $\varepsilon_{e\text{-ho}}$  may be achieved via formulating a correction that will remove the remaining concavity in  $E(N)$ .

## ACKNOWLEDGMENTS

Financial support by the German-Israeli Foundation, the European Research Council, and the Lise Meitner center for computational chemistry is gratefully acknowledged. E.K. is a recipient of the Levzion scholarship. T.S. acknowledges support from the Elite Network of Bavaria (“Macromolecular Science” program).

- <sup>1</sup>P. Hohenberg and W. Kohn, *Phys. Rev. B* **136**, 864 (1964).  
<sup>2</sup>W. Kohn and L. Sham, *Phys. Rev. A* **140**, 1133 (1965).  
<sup>3</sup>R. G. Parr and W. Yang, *Density-Functional Theory of Atoms and Molecules* (Oxford University Press, New York, 1989).  
<sup>4</sup>R. Dreizler and E. K. U. Gross, *Density Functional Theory: An Approach to the Quantum Many-Body Problem* (Springer, Berlin, 1990).  
<sup>5</sup>*A Primer in Density Functional Theory*, edited by C. Fiolhais, F. Nogueira, and M. Marques (Springer-Verlag, Berlin-Heidelberg, 2003).  
<sup>6</sup>A. D. Becke, *J. Chem. Phys.* **140**, 18A301 (2014).  
<sup>7</sup>K. Burke, *J. Chem. Phys.* **136**, 150901 (2012).  
<sup>8</sup>L. Kronik and S. Kümmel, *Top. Curr. Chem.* **347**, 137 (2014).  
<sup>9</sup>J. P. Perdew, R. G. Parr, M. Levy, and J. L. J. Balduz, *Phys. Rev. Lett.* **49**, 1691 (1982).  
<sup>10</sup>M. Levy, J. P. Perdew, and V. Sahni, *Phys. Rev. A* **30**, 2745 (1984).  
<sup>11</sup>C.-O. Almbladh and U. von Barth, *Phys. Rev. B* **31**, 3231 (1985).  
<sup>12</sup>J. P. Perdew and M. Levy, *Phys. Rev. B* **56**, 16021 (1997).  
<sup>13</sup>M. K. Harbola, *Phys. Rev. B* **60**, 4545 (1999).  
<sup>14</sup>I. Dabo, A. Ferretti, and N. Marzari, *Top. Curr. Chem.* **347**, 193 (2014).  
<sup>15</sup>B. Y. Tong and L. J. Sham, *Phys. Rev.* **144**, 1 (1966).  
<sup>16</sup>V. L. Moruzzi, J. F. Janak, and A. R. Williams, *Calculated Electronic Properties of Metals* (Pergamon, New York, 1978).  
<sup>17</sup>U. von Barth and L. Hedin, *J. Phys. C: Solid State Phys.* **5**, 1629 (1973).  
<sup>18</sup>O. Gunnarsson, B. I. Lundqvist, and J. W. Wilkins, *Phys. Rev. B* **10**, 1319 (1974).  
<sup>19</sup>J. F. Janak, V. L. Moruzzi, and A. R. Williams, *Phys. Rev. B* **12**, 1257 (1975).  
<sup>20</sup>O. Gunnarsson and B. I. Lundqvist, *Phys. Rev. B* **13**, 4274 (1976).  
<sup>21</sup>O. Gunnarsson and R. O. Jones, *Phys. Scr.* **21**, 394 (1980).  
<sup>22</sup>J. P. Perdew and A. Zunger, *Phys. Rev. B* **23**, 5048 (1981).  
<sup>23</sup>S. Kotochigova, Z. H. Levine, E. L. Shirley, M. D. Stiles, and C. W. Clark, *Phys. Rev. A* **55**, 191 (1997).  
<sup>24</sup>E. Kraisler, G. Makov, and I. Kelson, *Phys. Rev. A* **82**, 042516 (2010).  
<sup>25</sup>U. Argaman, G. Makov, and E. Kraisler, *Phys. Rev. A* **88**, 042504 (2013).  
<sup>26</sup>Y. Wang and J. P. Perdew, *Phys. Rev. B* **45**, 13244 (1992).  
<sup>27</sup>J. P. Perdew, K. Burke, and M. Ernzerhof, *Phys. Rev. Lett.* **77**, 3865 (1996).  
<sup>28</sup>M. Allen and D. Tozer, *Mol. Phys.* **100**, 433 (2002).  
<sup>29</sup>E. Livshits and R. Baer, *Phys. Chem. Chem. Phys.* **9**, 2932 (2007).  
<sup>30</sup>T. Körzdörfer, S. Kümmel, and M. Mundt, *J. Chem. Phys.* **129**, 014110 (2008).  
<sup>31</sup>U. Salzner and R. Baer, *J. Chem. Phys.* **131**, 231101 (2009).  
<sup>32</sup>S. Klüpfel, P. Klüpfel, and H. Jónsson, *Phys. Rev. A* **84**, 050501 (2011).  
<sup>33</sup>L. Kronik, T. Stein, S. Refaely-Abramson, and R. Baer, *J. Chem. Theory Comput.* **8**, 1515 (2012).  
<sup>34</sup>T. Schmidt, E. Kraisler, L. Kronik, and S. Kümmel, *Phys. Chem. Chem. Phys.* **16**, 14357 (2014).  
<sup>35</sup>P. Mori-Sánchez, A. J. Cohen, and W. Yang, *J. Chem. Phys.* **125**, 201102 (2006).  
<sup>36</sup>O. A. Vydrov, G. E. Scuseria, and J. P. Perdew, *J. Chem. Phys.* **126**, 154109 (2007).  
<sup>37</sup>A. Ruzsinszky, J. P. Perdew, G. I. Csonka, O. A. Vydrov, and G. E. Scuseria, *J. Chem. Phys.* **126**, 104102 (2007).  
<sup>38</sup>A. J. Cohen, P. Mori-Sánchez, and W. Yang, *Science* **321**, 792 (2008).  
<sup>39</sup>R. Haunschild, T. M. Henderson, C. A. Jiménez-Hoyos, and G. E. Scuseria, *J. Chem. Phys.* **133**, 134116 (2010).  
<sup>40</sup>I. Dabo *et al.*, *Phys. Rev. B* **82**, 115121 (2010).  
<sup>41</sup>A. J. Cohen, P. Mori-Sánchez, and W. Yang, *Chem. Rev.* **112**, 289 (2012).  
<sup>42</sup>T. Stein, J. Autschbach, N. Govind, L. Kronik, and R. Baer, *J. Phys. Chem. Lett.* **3**, 3740 (2012).  
<sup>43</sup>S. N. Steinmann and W. Yang, *J. Chem. Phys.* **139**, 074107 (2013).  
<sup>44</sup>G. Borghi, A. Ferretti, N. L. Nguyen, I. Dabo, and N. Marzari, *Phys. Rev. B* **90**, 075135 (2014).  
<sup>45</sup>M. A. Mosquera and A. Wasserman, *Phys. Rev. A* **89**, 052506 (2014).  
<sup>46</sup>M. A. Mosquera and A. Wasserman, *Mol. Phys.* **112**, 2997 (2014).  
<sup>47</sup>N. L. Nguyen, G. Borghi, A. Ferretti, I. Dabo, and N. Marzari, *Phys. Rev. Lett.* **114**, 166405 (2015).  
<sup>48</sup>G. Borghi, C.-H. Park, N. L. Nguyen, A. Ferretti, and N. Marzari, *Phys. Rev. B* **91**, 155112 (2015).  
<sup>49</sup>J. Janak, *Phys. Rev. B* **18**, 7165 (1978).  
<sup>50</sup>T. Körzdörfer, S. Kümmel, N. Marom, and L. Kronik, *Phys. Rev. B* **79**, 201205 (2009); **82**, 129903 (2010).  
<sup>51</sup>D. Cremer, *Mol. Phys.* **99**, 1899 (2001).  
<sup>52</sup>O. A. Vydrov and G. E. Scuseria, *J. Chem. Phys.* **121**, 8187 (2004).  
<sup>53</sup>O. A. Vydrov, G. E. Scuseria, J. P. Perdew, A. Ruzsinszky, and G. I. Csonka, *J. Chem. Phys.* **124**, 94108 (2006).  
<sup>54</sup>D. Hofmann, S. Klüpfel, P. Klüpfel, and S. Kümmel, *Phys. Rev. A* **85**, 062514 (2012).  
<sup>55</sup>S. Klüpfel, P. Klüpfel, and H. Jónsson, *J. Chem. Phys.* **137**, 124102 (2012).  
<sup>56</sup>R. van Leeuwen and E. J. Baerends, *Phys. Rev. A* **49**, 2421 (1994).  
<sup>57</sup>D. J. Tozer and N. C. Handy, *J. Chem. Phys.* **109**, 10180 (1998).  
<sup>58</sup>A. D. Becke and E. R. Johnson, *J. Chem. Phys.* **124**, 221101 (2006).  
<sup>59</sup>W. Cencek and K. Szalewicz, *J. Chem. Phys.* **139**, 024104 (2013).  
<sup>60</sup>A. Karolewski, R. Armiento, and S. Kümmel, *J. Chem. Theory Comput.* **5**, 712 (2009).  
<sup>61</sup>A. P. Gaiduk, S. Chulkov, and V. N. Staroverov, *J. Chem. Theory Comput.* **5**, 699 (2009).  
<sup>62</sup>A. Karolewski, R. Armiento, and S. Kümmel, *Phys. Rev. A* **88**, 052519 (2013).  
<sup>63</sup>R. Armiento and S. Kümmel, *Phys. Rev. Lett.* **111**, 036402 (2013).  
<sup>64</sup>V. Vlček, G. Steinle-Neumann, L. Leppert, R. Armiento, and S. Kümmel, *Phys. Rev. B* **91**, 035107 (2015).  
<sup>65</sup>F. Tran, P. Blaha, M. Betzinger, and S. Blügel, *Phys. Rev. B* **91**, 165121 (2015).  
<sup>66</sup>A. D. Becke, *J. Chem. Phys.* **98**, 1372 (1993).  
<sup>67</sup>A. D. Becke, *J. Chem. Phys.* **98**, 5648 (1993).  
<sup>68</sup>J. P. Perdew, M. Ernzerhof, and K. Burke, *J. Chem. Phys.* **105**, 9982 (1996).  
<sup>69</sup>P. J. Stephens, F. J. Devlin, C. F. Chabalowski, and M. J. Frisch, *J. Phys. Chem.* **98**, 11623 (1994).  
<sup>70</sup>C. Adamo and V. Barone, *J. Chem. Phys.* **110**, 6158 (1999).  
<sup>71</sup>M. Ernzerhof and G. E. Scuseria, *J. Chem. Phys.* **110**, 5029 (1999).  
<sup>72</sup>N. Sai, P. F. Barbara, and K. Leung, *Phys. Rev. Lett.* **106**, 226403 (2011).  
<sup>73</sup>Y. Imamura, R. Kobayashi, and H. Nakai, *Chem. Phys. Lett.* **513**, 130 (2011).  
<sup>74</sup>V. Atalla, M. Yoon, F. Caruso, P. Rinke, and M. Scheffler, *Phys. Rev. B* **88**, 165122 (2013).  
<sup>75</sup>Y. Zhao and D. Truhlar, *Acc. Chem. Res.* **41**, 157 (2008).  
<sup>76</sup>N. Marom, A. Tkatchenko, M. Scheffler, and L. Kronik, *J. Chem. Theory Comput.* **6**, 81 (2010).  
<sup>77</sup>T. Körzdörfer and S. Kümmel, *Phys. Rev. B* **82**, 155206 (2010).  
<sup>78</sup>P. Verma and R. J. Bartlett, *J. Chem. Phys.* **136**, 44105 (2012).  
<sup>79</sup>P. Verma and R. J. Bartlett, *J. Chem. Phys.* **137**, 134102 (2012).  
<sup>80</sup>P. Verma and R. J. Bartlett, *J. Chem. Phys.* **140**, 18A534 (2014).  
<sup>81</sup>F. G. Cruz, K.-C. Lam, and K. Burke, *J. Phys. Chem. A* **102**, 4911 (1998).  
<sup>82</sup>J. Jaramillo, G. E. Scuseria, and M. Ernzerhof, *J. Chem. Phys.* **118**, 1068 (2003).  
<sup>83</sup>M. Kaupp, H. Bahmann, and A. V. Arbuznikov, *J. Chem. Phys.* **127**, 194102 (2007).  
<sup>84</sup>P. de Silva and C. Corminboeuf, *J. Chem. Phys.* **142**, 074112 (2015).  
<sup>85</sup>T. Schmidt, E. Kraisler, A. Makmal, L. Kronik, and S. Kümmel, *J. Chem. Phys.* **140**, 18A510 (2014).  
<sup>86</sup>E. Kraisler and L. Kronik, *J. Chem. Phys.* **140**, 18A540 (2014).  
<sup>87</sup>E. Kraisler and L. Kronik, *Phys. Rev. Lett.* **110**, 126403 (2013).  
<sup>88</sup>E. H. Lieb, *Int. J. Quantum Chem.* **24**, 243 (1983).  
<sup>89</sup>M. Levy and J. Perdew, in *Density Functional Methods in Physics*, edited by R. M. Dreizler and J. da Providência (Plenum, 1985).  
<sup>90</sup>R. van Leeuwen, *Adv. Quantum Chem.* **43**, 25 (2003).  
<sup>91</sup>T. Gould and J. F. Dobson, *J. Chem. Phys.* **138**, 14103 (2013).  
<sup>92</sup>A. Görling, *Phys. Rev. B* **91**, 245120 (2015).  
<sup>93</sup>E. Kraisler and L. Kronik, *Phys. Rev. A* **91**, 032504 (2015).

- <sup>94</sup>X. Zheng, A. J. Cohen, P. Mori-Sánchez, X. Hu, and W. Yang, *Phys. Rev. Lett.* **107**, 026403 (2011).
- <sup>95</sup>M. Levy and F. Zahariev, *Phys. Rev. Lett.* **113**, 113002 (2014).
- <sup>96</sup>M. Casida, in *Recent Advances in Density-Functional Methods*, edited by D. Chong (World Scientific, 1995), p. 155.
- <sup>97</sup>A. Makmal, S. Kümmel, and L. Kronik, *J. Chem. Theory Comput.* **5**, 1731 (2009); **7**, 2665 (2011).
- <sup>98</sup>A. Makmal, "Orbital-dependent-functionals within density functional theory: Methodology and applications," Ph.D. thesis, Weizmann Institute of Science, 2010.
- <sup>99</sup>J. Krieger, Y. Li, and G. Iafate, *Phys. Rev. A* **46**, 5453 (1992).
- <sup>100</sup>T. Grabo, T. Kreibich, and E. K. U. Gross, *Mol. Eng.* **7**, 27 (1997).
- <sup>101</sup>E. Engel and R. M. Dreizler, *Density Functional Theory - An Advanced Course* (Springer-Verlag, Berlin, Heidelberg, 2011).
- <sup>102</sup>S. Kümmel and L. Kronik, *Rev. Mod. Phys.* **80**, 3 (2008).
- <sup>103</sup>Y. Li, J. Krieger, and G. Iafate, *Phys. Rev. A* **47**, 165 (1993).
- <sup>104</sup>*CRC Handbook of Chemistry and Physics*, 92th ed., edited by D. R. Lide (CRC, London, 2011), available on <http://www.hbcnetbase.com>.
- <sup>105</sup>See <http://webbook.nist.gov> for experimental data.
- <sup>106</sup>A. D. Becke, *Phys. Rev.* **38**, 3098 (1988).
- <sup>107</sup>C. Lee, W. Yang, and R. G. Parr, *Phys. Rev. B* **37**, 785 (1988).
- <sup>108</sup>B. Miehlich, A. Savin, H. Stoll, and H. Preuss, *Chem. Phys. Lett.* **157**, 200 (1989).
- <sup>109</sup>See supplementary material at <http://dx.doi.org/10.1063/1.4930119> for detailed numerical data in tabular form.
- <sup>110</sup>L. Pedroza, A. da Silva, and K. Capelle, *Phys. Rev. B* **79**, 201106(R) (2009).
- <sup>111</sup>E. Engel, "Orbital-dependent functionals for the exchange-correlation energy: A third generation of density functionals," in *A Primer in Density Functional Theory*, Lectures in Physics Vol. 620, edited by C. Fiolhais, F. Nogueira, and M. A. Marques (Springer, 2003).
- <sup>112</sup>M.-C. Kim, E. Sim, and K. Burke, *J. Chem. Phys.* **134**, 171103 (2011).
- <sup>113</sup>O. Guliamov, L. Kronik, and J. M. L. Martin, *J. Phys. Chem. A* **111**, 2028 (2007).
- <sup>114</sup>J. M. Galbraith and H. F. Schaefer, *J. Chem. Phys.* **105**, 862 (1996).
- <sup>115</sup>N. Rösch and S. B. Trickey, *J. Chem. Phys.* **106**, 8940 (1997).
- <sup>116</sup>E. Kraisler, G. Makov, N. Argaman, and I. Kelson, *Phys. Rev. A* **80**, 032115 (2009).
- <sup>117</sup>E. Kraisler, "Density-functional calculations in atomic systems," M.Sc. thesis, Tel-Aviv University, Israel, 2010.
- <sup>118</sup>A. Makmal, S. Kümmel, and L. Kronik, *Phys. Rev. A* **83**, 062512 (2011).
- <sup>119</sup>O. V. Gritsenko and E. J. Baerends, *J. Chem. Phys.* **120**, 8364 (2004).
- <sup>120</sup>J. P. Perdew and M. Levy, *Phys. Rev. Lett.* **51**, 1884 (1983).
- <sup>121</sup>J. Perdew, in *Density Functional Methods in Physics*, edited by R. M. Dreizler and J. da Providência (Plenum, 1985).
- <sup>122</sup>R. Godby, M. Schlüter, and L. Sham, *Phys. Rev. B* **36**, 6497 (1987).
- <sup>123</sup>R. Godby, M. Schlüter, and L. Sham, *Phys. Rev. B* **37**, 10159 (1988).
- <sup>124</sup>M. K. Harbola, *Phys. Rev. A* **57**, 4253 (1998).
- <sup>125</sup>G. K.-L. Chan, *J. Chem. Phys.* **110**, 4710 (1999).
- <sup>126</sup>A. M. Teale, F. de Profijt, and D. J. Tozer, *J. Chem. Phys.* **129**, 044110 (2008).
- <sup>127</sup>J. B. Krieger, Y. Li, and G. J. Iafate, *Phys. Rev. A* **45**, 101 (1992).
- <sup>128</sup>A. Görling and M. Levy, *Phys. Rev.* **52**, 4493 (1995).
- <sup>129</sup>W. Kohn, *Phys. Rev. B* **33**, 4331 (1986).
- <sup>130</sup>X. Andrade and A. Aspuru-Guzik, *Phys. Rev. Lett.* **107**, 183002 (2011).
- <sup>131</sup>N. I. Gidopoulos and N. N. Lathiotakis, *J. Chem. Phys.* **136**, 224109 (2012).
- <sup>132</sup>J.-D. Chai and P.-T. Chen, *Phys. Rev. Lett.* **110**, 033002 (2013).
- <sup>133</sup>E. J. Baerends, O. V. Gritsenko, and R. van Meer, *Phys. Chem. Chem. Phys.* **15**, 16408 (2013).
- <sup>134</sup>I. Dabo et al., Psi-k Newsletter **119**, 1 (2013), pdf available at [http://www.psi-k.org/newsletters/News\\_119/Highlight\\_119.pdf](http://www.psi-k.org/newsletters/News_119/Highlight_119.pdf).
- <sup>135</sup>Here and below, it is assumed that the ground states of the system of interest and of its ion are not degenerate, or that the degeneracy can be lifted by applying an infinitesimal external field.
- <sup>136</sup>The fact that only contributions from the  $N_0 - 1$ - and the  $N_0$ -states are included relies on the conjecture that the series  $E(N_0)$  for  $N_0 \in \mathbb{N}$  is a convex, monotonously decreasing series. In other words, all ionization energies  $I(N_0) := E(N_0 - 1) - E(N_0)$  are positive, and higher ionizations are always larger than the lower ones:  $I(N_0 - 1) > I(N_0)$ . This conjecture, although strongly supported by experimental data, remains without proof, to the best of our knowledge.<sup>4,41,88</sup>
- <sup>137</sup>Janak's theorem<sup>49</sup> states that the  $i$ th KS eigenenergy,  $\varepsilon_i$ , equals  $\partial E / \partial f_i$ —the derivative of the total energy of the interacting system,  $E$ , with respect to the occupation of the  $i$ th level,  $f_i$ . It can be shown that with the exact xc functional the ho eigenenergy,  $\varepsilon_{\text{ho}}$ , has to equal  $\varepsilon_{\text{ho}} = \partial E / \partial \alpha = \partial E / \partial N = E(N_0) - E(N_0 - 1) = -I$ , i.e., it equals the negative of the IP.
- <sup>138</sup>For the LSDA, relaxation runs have been performed for all molecules. It was found that the experimental bond length lies within the numerical error range of the relaxed bond length in all cases. We checked that the ho and  $e$ -ho energy values for the relaxed geometries agree with the ones at experimental geometries within 0.002 hartree, except for H<sub>2</sub>, NH, and F<sub>2</sub>, where the difference reaches 0.005 hartree. For the ISOc functional, similar relaxation checks were performed, as described in Ref. 85.
- <sup>139</sup>In this context, we note that the values reported by some of us in Ref. 87 for the relaxed H<sub>2</sub> molecule, namely, the  $e$ -ho energy,  $\varepsilon_{e\text{-ho}}$ , and as a result—the fundamental gap of the ion,  $E_g$ , are slightly different upon closer observation. In fact, at the relaxed bond length of  $L = 1.45$  Bohr, these values are  $\varepsilon_{e\text{-ho}} = 0.618$  hartree = 1.236 Ry and  $E_g = 0.671$  hartree = 1.341 Ry and not 1.223 Ry and 1.320 Ry, respectively. The difference originates from retrieving the value for  $\varepsilon_{e\text{-ho}}$  directly, not relying on the chemical potential  $\mu$  calculated in the DARSEC program, with a temperature of 1K.
- <sup>140</sup>We recall that the latter does not produce a strictly piecewise linear energy curve  $E(N)$ , but there typically remains some concavity, which is attributed to the implicit dependence of  $E(N)$  on  $\alpha$  via the KS orbitals. This concavity affects the value of  $\varepsilon_{e\text{-ho}}$ . However, even in case  $E(N)$  would be exactly piecewise linear,  $\varepsilon_{e\text{-ho}}$  would reproduce  $-I_{\text{ASCF}}$  rather than the experimental IP.
- <sup>141</sup>Note that the combination of the EXX functional with the standard form for the Hartree functional results in an intrinsically ensemble-generalized functional if the ground state is described by an ensemble comprised of two pure many-electron states. This is the case throughout this work as we describe the ionization process by extracting an electron from a specific spin-channel. If the number of many-electron states is larger than two (as is the case, e.g., if both spin channels are fractionally occupied), then the EXX is not intrinsically ensemble-generalized, but an appropriate ensemble generalization, proposed in Ref. 91, is available.
- <sup>142</sup>When calculating the NH molecule with the LSDA or PBEh( $a$ ) using values of  $0 \leq a \leq 0.55$ , the global  $\varepsilon_{\text{ho}}$  and  $\varepsilon_{e\text{-ho}}$  do not belong to the same spin channel, a behavior that has not been observed in any other system in our test set.