Exploring weight-dependent density-functional approximations for ensembles in the Hubbard dimer

Killian Deur, Laurent Mazouin, Bruno Senjean, and Emmanuel Fromager

Laboratoire de Chimie Quantique, Institut de Chimie, CNRS / Université de Strasbourg, 4 rue Blaise Pascal, 67000 Strasbourg, France

Received: date / Revised version: date

Abstract. Gross–Oliveira–Kohn density-functional theory (GOK-DFT) is an extension of DFT to excited states where the basic variable is the ensemble density, i.e. the weighted sum of ground- and excited-state densities. The ensemble energy (i.e. the weighted sum of ground- and excited-state energies) can be obtained variationally as a functional of the ensemble density. Like in DFT, the key ingredient to model in GOK-DFT is the exchange-correlation functional. Developing density-functional approximations (DFAs) for ensembles is a complicated task as both density and weight dependencies should in principle be reproduced. In a recent paper [Phys. Rev. B 95, 035120 (2017)], the authors applied exact GOK-DFT to the simple but nontrivial Hubbard dimer in order to investigate (numerically) the importance of weight dependence in the calculation of excitation energies. In this work, we derive analytical DFAs for various density and correlation regimes by means of a Legendre–Fenchel transform formalism. Both functional and density driven errors are evaluated for each DFA. Interestingly, when the ensemble exact-exchange-only functional is used, these errors can be large, in particular if the dimer is symmetric, but they cancel each other so that the excitation energies obtained by linear interpolation are always accurate, even in the strongly correlated regime.

PACS. PACS-key describing text of that key – PACS-key describing text of that key

1 Introduction

Even though the Gross–Oliveira–Kohn ensemble density-functional theory (eDFT) [1,2,3,4] for excited states is not routinely used nowadays for the computation of excited state properties, the approach has regained interest in recent years [5,6,7,8,9,10,11,12,13,14,15,16]. Indeed, eDFT stands as a potential alternative to the popular time-dependent linear response DFT for the description of charge transfer excitations, near-degeneracies and multiple electronic excitations. Note that, in addition to eDFT, other in-principle-exact time-independent extensions of DFT to excited states have been explored over the years (mostly at the formal level) by Levy and coworkers, either by considering special cases where the standard universal functional yields excited-state energies [17], or by using the external potential (or its ground-state density) as an additional variable, thus forming a bifunctional [18,19,20,21]. The latter complication can actually be overcome for Coulomb systems [22]. Note that all these formulations are state specific, i.e. one specific excited-state density is targeted. This is an important difference with eDFT where the basic variable is a state-averaged density.

In eDFT, weights are assigned to the ground and the excited states that belong to the ensemble under study. Therefore, it is in principle crucial to model, for a fixed density, the weight dependence of the ensemble exchange-correlation density-functional energy. Let us stress that, in the general formulation of the theory, the weights only need to be ordered (the largest one being assigned to the ground state). Boltzmann weights can of course be employed but it is not compulsory. Using fixed (energy-independent) weights might indeed be appealing for practical calculations [24,25]. One of the limitations of approximate eDFT is the so-called ghost-interaction error [26] which arises when the Hartree energy (which is quadratic in the density) is computed with the ensemble density (i.e. the weighted sum of ground- and excited-state densities). This issue is actually related to the proper description of the weight dependence in the exchange energy [27,28,29].

In order to investigate the weight dependence of both exchange and correlation density functionals, some of the authors have recently applied eDFT to the two-site Hubbard model [30]. Despite its simplicity, the model is actually nontrivial and can be used as a lab for testing new ideas in DFT [31,32,33,34,35]. The work presented in Ref. [30] deals with the exact theory, which means that exact ensemble correlation energies have been computed.

Corresponding author: fromagere@unistra.fr
numerically for various density and correlation regimes. Even though these results are precious for parameterizing density-functional approximations (DFAs), it is still unclear how this can actually be achieved. Following Carra-ascal et al. [31], we propose to expand the exact ensemble correlation functional around specific cases like, for example, the symmetric one. A general strategy, that we expect to be applicable not only to the Hubbard dimer but also to more realistic models, will be presented. It uses Legendre–Fenchel transforms [38] instead of the more popular Levy–Lieb constrained search formalism [37]. The paper is organized as follows. After a brief review on exact eDFT and its application to the Hubbard dimer (Sec. 2), the derivation of Taylor expansions for the exact ensemble correlation energy will be presented in several density and correlation regimes (Sec. 3). Following a summary of simple and more advanced DFAs based on the latter expansions (Sec. 4), results obtained for the total two-state singlet ensemble energy and the corresponding excitation energy are presented and discussed in Sec. 5. Conclusions are finally given in Sec. 6.

2 Theory

For sake of clarity, a brief introduction to eDFT and its application to the Hubbard dimer is given in this section. More details can be found in Ref. [30] and the references therein.

2.1 Ensemble density-functional theory for excited states

The ensemble energy is a convex combination of \(N\)-electron ground- and excited-state energies with ordered coefficients (also called weights), the largest one being assigned to the ground state. In the particular case of two states (ground and first-excited) which is considered in this work, the ensemble energy associated to the electronic Hamiltonian with local potential \(v\),

\[
\hat{H}[v] = \hat{T} + W_{ee} + \int dr \, v(r) n(r),
\]

reads

\[
E^w[v] = (1 - w) E_0[v] + w E_1[v],
\]

where \(E_0[v]\) and \(E_1[v]\) are the ground- and first-excited-state energies of \(\hat{H}[v]\), and the ensemble weight \(w\) is such that \((1 - w) \geq w \geq 0\) or, equivalently,

\[
0 \leq w \leq 1/2.
\]

In Eq. (1), \(\hat{T}\), \(W_{ee}\) and \(n(r)\) denote the kinetic energy, two-electron repulsion and density operators, respectively. Gross, Oliveira and Kohn have shown [3] that the ensemble energy is a functional of the ensemble density,

\[
n^w[v](r) = (1 - w)n_0[v](r) + w n_1[v](r),
\]

where \(n_0[v]\) and \(n_1[v]\) denote the ground- and first-excited-state densities of \(\hat{H}[v]\), and that it can be determined variationally as follows,

\[
E^w[v] = \inf_n \left\{ F^w[n] + \int dr \, v(r)n(r) \right\},
\]

where \(F^w[n]\) is the \(w\)-dependent analog of the Hohenberg–Kohn universal functional for ensembles. While it is usually expressed within the Levy–Lieb constrained-search formalism [37], which would involve two many-body wavefunctions, we will instead use a Legendre–Fenchel transform-based expression. The latter is simply obtained from Eq. (5) by considering a fixed density \(n\) and writing, for any potential \(v\), the following inequality,

\[
E^w[v] \leq F^w[n] + \int dr \, v(r)n(r),
\]

or, equivalently,

\[
F^w[n] \geq E^w[v] - \int dr \, v(r)n(r),
\]

thus leading to the final expression

\[
F^w[n] = \sup_v \left\{ E^w[v] - \int dr \, v(r)n(r) \right\}.
\]

As discussed further in the rest of this work, the latter expression has the advantage of using a single variable, namely the local potential \(v\), and will not require the use of many-body wavefunctions, which is extremely convenient for deriving density-functional approximations.

In the conventional Kohn–Sham (KS) formulation of eDFT [3], the universal ensemble functional is split into the noninteracting analog of \(F^w[n]\), namely the noninteracting ensemble kinetic energy functional \(T_s^w[n]\), and the complementary ensemble Hartree, exchange and correlation (Hxc) density-functional energies,

\[
F^w[n] = T_s^w[n] + E_H[n] + E_x^w[n] + E_c^w[n].
\]

In analogy with Eq. (8), we have

\[
T_s^w[n] = \sup_v \left\{ E_{K_S}^v[n] - \int dr \, v(r)n(r) \right\},
\]

where \(E_{K_S}^v[n]\) is the ensemble energy of \(\hat{T} + \int dr \, v(r)n(r)\). Note that, in the decomposition of Eq. (9), the conventional (weight-independent) Hartree functional is used,

\[
E_H[n] = \frac{1}{2} \int dr dr' \, \frac{n(r)n(r')}{|r - r'|},
\]

which, in practice, can induce substantial ghost interaction errors [26,27,28,29]. In the exact theory, the latter are removed by the weight-dependent exchange and correlation functionals. Regarding the exchange energy, a general expression has been derived and tested recently in
It allows for the construction of an ensemble exact exchange functional (EEXX) from the exact density (if available) or through an optimized effective potential (OEP) procedure. In the Hubbard dimer, which is studied in this work, the EEXX energy is an explicit functional of the density. We will therefore focus in the following on the weight dependence of the correlation energy.

According to Eqs. (5) and (9), for a given local external potential $v_{\text{ext}}$, the exact ensemble energy $E^w = E^w[v_{\text{ext}}]$ is obtained variationally as follows in KS-eDFT,

$$E^w = \inf_n \left\{ T^w_n[n] + E_H[n] + E^w_{\text{xc}}[n] + E^w_{\text{c}}[n] \right\} + \int dr v_{\text{ext}}(r)n(r),$$  \hspace{1cm} (12)

The ensemble non-interacting kinetic energy functional is usually expressed in terms of the KS orbitals, thus leading to the analog for ensembles of the self-consistent KS equations. This step is actually unnecessary in the Hubbard dimer since the exact analytical expression for $T^w_n[n]$ is known.

As readily seen from Eq. (2), the ensemble energy varies linearly with the ensemble weight. Consequently, the excitation energy (or optical gap) $\Omega = E^w_1[v_{\text{ext}}] - E^w_0[v_{\text{ext}}]$ can be determined either by differentiation,

$$\Omega = \frac{dE^w}{dw},$$ \hspace{1cm} (13)

or by linear interpolation [24],

$$\Omega = 2 \left( E^{w=1/2} - E^{w=0} \right).$$ \hspace{1cm} (14)

Eqs. (13) and (14) are equivalent in the exact theory. However, as clearly illustrated in the following, they will give different results, that might also be weight-dependent, when DFAs are used, as expected. Note that, by using the stationarity of the minimizing ensemble density $n^w = n^w[v_{\text{ext}}]$ in Eq. (12), we obtain from Eq. (13) the simplified in-principle-exact expression,

$$\Omega = \left[ \frac{\partial T^w_n[n]}{\partial w} + \frac{\partial E_{\text{xc}}^w[n]}{\partial w} + \frac{\partial E_{\text{c}}^w[n]}{\partial w} \right]_{n=n^w},$$ \hspace{1cm} (15)

where, according to Eq. (10), the first term on the right-hand side is nothing but the KS optical gap, and the last two terms correspond to exchange and correlation derivative discontinuity contributions.

Let us finally stress that the expression for $dE^w/dw$ given in the right-hand side of Eq. (15) remains valid when approximate functionals are used as long as the (now approximate) ensemble energy $E^w$ is calculated variationally (i.e. by minimization over densities) according to Eq. (12). This is due to the stationarity of the (now approximate) minimizing ensemble density $n^w$. As pointed out previously, in this case, $dE^w/dw$ might become $w$-dependent and therefore, for a given value of $w$, it may deviate from the slope obtained by linear interpolation (right-hand side of Eq. (14)).

### 2.2 Ensemble DFT for the two-site Hubbard model

In the two-site Hubbard model, the ab initio Hamiltonian of Eq. (1) is simplified as follows,

$$\hat{T} \rightarrow \hat{T} = -t \sum_{\sigma = \uparrow, \downarrow} \left( \hat{a}_{i\sigma}^{\dagger} \hat{a}_{i\sigma} + \hat{a}_{i\sigma}^{\dagger} \hat{a}_{i\sigma} \right),$$

$$\hat{W}_{\text{ec}} \rightarrow \hat{U} = U \sum_{i=0}^{1} \hat{a}_{i\uparrow} \hat{a}_{i\downarrow} \hat{a}_{i\downarrow}^{\dagger} \hat{a}_{i\uparrow}^{\dagger},$$

$$\int dr v(r)\hat{n}(r) = \frac{\Delta v}{2} (\hat{n}_1 - \hat{n}_0),$$ \hspace{1cm} (16)

where operators are written in second quantization and the labels 0 and 1 refer to the first and second atomic site, respectively. The density operator on site $i$ reads $\hat{n}_i = \sum_{\sigma = \uparrow, \downarrow} \hat{a}_{i\sigma} \hat{a}_{i\sigma}^{\dagger}$, as shown in Refs. [30] and [31], various correlation and density regimes can be explored by varying the three parameters of the model, namely $t$ (the hopping parameter), $U$ (the strength of the on-site two-electron repulsion) and the local potential parameter $\Delta v$ which controls the asymmetry of the model. Following Ref. [30], we will describe in the rest of this work a two-electron ensemble consisting of the ground- and first-excited singlet states of the Hubbard dimer. The exact energies $E_i$ ($i = 0, 1$), which are functions of $t$, $U$ and $\Delta v$, can be determined analytically by solving the following third-order polynomial equation [50][81][92],

$$-4t^2U + (4t^2 - U^2 + \Delta v^2) E_i + 2UE_i^3 = E_i^3. \hspace{1cm} (17)$$

In this context, a trial density consists in principle of two numbers, $n_0$ and $n_1$, which are the occupations of site 0 and 1, respectively. In the particular case of two electrons, the density can be reduced to a single occupation number $n = n_0$ since $n_1 = 2 - n_0$. Consequently, for a given external local potential $\Delta v = \Delta v_{\text{ext}}$, the exact ensemble energy $E^w = E^w(\Delta v_{\text{ext}})$ of the two-electron Hubbard dimer can be expressed as follows in KS-eDFT,

$$E^w = \inf_n \left\{ E^w_{\Delta v_{\text{ext}}}(n) \right\},$$ \hspace{1cm} (18)

where the density-functional ensemble energy to be minimized reads

$$E^w_{\Delta v_{\text{ext}}}(n) = T^w_n(n) + E_H(n) + E^w_{\text{xc}}(n) + E^w_{\text{c}}(n) + \Delta v_{\text{ext}} \times (1 - n),$$ \hspace{1cm} (19)

in analogy with the ab initio expression of Eq. (12). Note that the $t$ and $U$ dependencies of the various density-functional energy contributions have been dropped for clarity. Note also that the latter functionals are in fact functions of the occupation number $n$ that will be referred to as density in the rest of this work. As shown in Ref. [30], exact analytical expressions can be derived for all func-
tionals except the correlation one:
\[ T^w(n) = -2t \sqrt{1 - w^2} - (1 - n)^2, \]
\[ E_H(n) = U \left( 1 + (1 - n)^2 \right), \]
\[ E^w(n) = \frac{U}{2} \left[ 1 + w - \frac{(3w - 1)(1 - n)^2}{1 - w^2} \right] - E_H(n). \]
(20)
(21)
(22)

As readily seen from Eq. (20), a density \( n \) is ensemble non-interacting \( v \)-representable if
\[ |1 - n| \leq 1 - w. \]
(23)

For densities in the latter range, the exact ensemble correlation energy can be obtained numerically as follows \[30],
\[ E^w_c(n) = F^w(n) - T^w(n) - E_H(n) - E^w_c(n), \]
(24)
where, in analogy with the \( ab \) \( ini \) expression in Eq. (8),
\[ F^w(n) = \sup_{\Delta v} \left\{ E^w(\Delta v) + \Delta v \times (n - 1) \right\}. \]
(25)

Note that, for a trial potential \( \Delta v \), the ensemble energy \( E^w(\Delta v) \) is determined from Eq. (17). Obviously, for practical calculations, analytical DFAs are preferable to numerical ones. Moreover, developing a general strategy for the derivation of weight-dependent correlation functionals that might also be applicable to \( ab \) \( ini \) Hamiltonians is highly desirable. We will show in the following how explicit correlation density functionals can be constructed by expanding the Legendre–Fenchel transform of Eq. (25) in the vicinity of various density and correlation regimes.

3 Taylor expansions of the exact ensemble correlation functional

3.1 Expansion around the symmetric case

For convenience we introduce the on-site repulsion \( u = U/(2t) \), local potential \( \nu = \Delta v/(2t) \) and ensemble Legendre–Fenchel transform
\[ f^w(\delta) = F^w(1 + \delta)/(2t) \]
(26)
per unit of \( 2t \), thus leading to (see Eq. (25))
\[ f^w(\delta) = \sup_{\nu} \left\{ (1 - w)e_0(\nu) + we_1(\nu) + \nu \delta \right\} = (1 - w)e_0(\nu) + we_1(\nu) + \nu(\delta) \delta, \]
(27)
where, for given values of \( u \) and \( \nu \), the individual energies \( e_i = E_i/(2t) \equiv e_i(\nu, u) \) of the ground- \( (i = 0) \) and first-excited \( (i = 1) \) singlet states are, according to Eq. (17), solutions of
\[ -u + e_i(1 - u^2 + \nu^2) + 2u e_i^2 = e_i^3. \]
(28)

Note that the \( \nu \)-dependence of \( e_0 \) and \( e_1 \) has been dropped in Eq. (27) for clarity. In order to expand the ensemble Legendre–Fenchel transform \( F^w(\delta) \) around the symmetric \( n = 1 \) case, which is equivalent to expanding \( f^w(\delta) \) around \( \delta = 0 \),
\[ f^w(\delta) = f^w(0) + \delta \frac{d f^w(\delta)}{d \delta} \bigg|_{\delta = 0} + \frac{\delta^2}{2} \frac{d^2 f^w(\delta)}{d \delta^2} \bigg|_{\delta = 0} + O(\delta^3), \]
(29)
we need to calculate energy derivatives. Indeed, by using the stationarity of the maximizing potential \( \nu(\delta) \) in Eq. (27), we obtain
\[ \frac{d f^w(\delta)}{d \delta} = \nu(\delta), \]
(30)
thus leading to
\[ \frac{d^2 f^w(\delta)}{d \delta^2} \bigg|_{\delta = 0} = \frac{d \nu(\delta)}{d \delta} \bigg|_{\delta = 0}. \]
(31)
The latter response of the potential (to deviations in density from the symmetric case) is determined from the stationarity condition, which holds for any \( \delta \),
\[ (1 - w) \frac{\partial e_0(\nu)}{\partial \nu} \bigg|_{\nu=\nu(\delta)} + w \frac{\partial e_1(\nu)}{\partial \nu} \bigg|_{\nu=\nu(\delta)} = -\delta, \]
(32)
thus giving after differentiation with respect to \( \delta \),
\[ \frac{d \nu(\delta)}{d \delta} \bigg|_{\delta = 0} = -\left[ (1 - w) \frac{\partial^2 e_0(\nu)}{\partial \nu^2} + w \frac{\partial^2 e_1(\nu)}{\partial \nu^2} \right]_{\nu=\nu(0)}^{-1}. \]
(33)

Differentiating Eq. (28) with respect to \( \nu \) gives
\[ \frac{\partial e_i(\nu)}{\partial \nu} \times \left[ 1 - u^2 + \nu^2 + 4ue_i(\nu) - 3e_i^2(\nu) \right] = -2\nu e_i(\nu), \]
(34)
which, when combined with Eqs. (30) and (32), leads to the expected solution \[30],
\[ \nu(0) = 0 = \frac{d f^w(\delta)}{d \delta} \bigg|_{\delta = 0}. \]
(35)
Similarly, by differentiating Eq. (34) with respect to \( \nu \) and using Eq. (33), we obtain (see Appendix A)
\[ \frac{d^2 f^w(\delta)}{d \delta^2} \bigg|_{\delta = 0} = \frac{g(u)}{2 \left( 1 + u \left[ u g(u) - 1 \right] \right)}, \]
(36)
where
\[ g(u) = \left( \frac{u}{2} + \sqrt{1 + \left( \frac{u}{2} \right)^2} \right) \left[ 1 + \left( \frac{u}{2} + \sqrt{1 + \left( \frac{u}{2} \right)^2} \right)^2 \right]. \]
(37)
Turning to the ensemble correlation energy (per unit of 2t),

\[ e_c^w(\delta) = f^w(\delta) - f^w(\delta, u = 0) = e_{HS}^w(\delta), \tag{38} \]

where, according to Eq. (22),

\[ e_{HS}^w(\delta) = \frac{u}{2} \left[ 1 + w - \left( 3w - 1 \right) \delta^2 \right]/ \left( 1 - w^2 \right), \tag{39} \]

we finally obtain from Eqs. (35) and (36) the following expansion through second order in \( \delta \),

\[ e_c^w(\delta) = (1 - w) \left[ 1 - \sqrt{1 + \left( \frac{u}{2} \right)^2} \right] \]
\[ + \frac{\delta^2}{4} \left[ \frac{g(u)}{1 + w[u g(u) - 1]} - \frac{2(u + 1 - w(3u + 1))}{(1 - w)^2} \right] \]
\[ + O(\delta^3), \tag{40} \]

where we used the simplified expression

\[ f^w(0) = \frac{u(1 + w)}{2} - (1 - w) \sqrt{1 + \left( \frac{u}{2} \right)^2}, \tag{41} \]

which is deduced from Eqs. (27), (35), (A.3), and (A.4). Note that, as expected, the expansion obtained by Carvalho et al. around \( n = 1 \) for the ground-state functional (see Eq. (B.13) in Ref. 31) where \( \rho \) corresponds to our \( |\delta| \) is recovered from Eq. (40) when \( w = 0 \).

Let us finally focus on the behavior of the expansion in Eq. (40) when \( |\delta| = 1/2 \) and \( w = \frac{1}{2} - \eta \) where \( \eta \ll 1 \) (i.e. close to the equi-ensemble case and far from the symmetric case). Truncation through second order in \( \delta \) and first order in \( \eta \) gives

\[ 2e_c^{\frac{1}{2} - \eta} \left( \pm \frac{1}{2} \right) \approx \]
\[ \left[ \frac{1}{2} - \frac{1}{u} + \frac{1}{u^2} \right] + \frac{g(u)}{4(1 + u g(u))} \]
\[ + \eta \left[ 3 - 5u^2 - 2 \sqrt{1 + \left( \frac{u}{2} \right)^2} - \frac{g(u)(1 - u g(u))}{(1 + u g(u))^2} \right]. \tag{42} \]

Interestingly, the latter correlation energy expression will vary as follows in the strongly correlated limit,

\[ e_c^{\frac{1}{2} - \eta} \left( \pm \frac{1}{2} \right) \approx \frac{1}{4} - 3u\eta, \tag{43} \]

and, as readily seen, an unphysical positive result is obtained when \( \eta = 0 \). In other words, the expansion in Eq. (40) is expected to fail in practice if calculations are performed with \( w = 1/2 \) in such regimes of density and correlation. Note also that, when \( \eta > 0 \), the expansion in Eq. (43) becomes

\[ e_c^{\frac{1}{2} - \eta} \left( \pm \frac{1}{2} \right) \approx -3u\eta, \tag{44} \]

which is actually incorrect, as will be discussed further in Sec. 3.3.

### 3.2 Expansion in the weakly correlated regime

By following the same strategy as in Sec. 3.1, we will expand in this section the ensemble Legendre–Fenchel transform around \( u = 0 \) for a fixed ensemble non-interacting \( \nu \)-representable deviation \( \delta \) from the symmetric case, i.e. any deviation such that (see Eq. (23))

\[ |\delta| \leq 1 - w. \tag{45} \]

For clarity, we will make both \( \nu \)- and \( \delta \)-dependencies explicit in Eq. (27), thus leading to

\[ f^w(\delta, \nu) = \sup_{\nu} \left\{ (1 - w)e_0(\nu, u) + we_1(\nu, u) + \nu \delta \right\} \]
\[ = (1 - w)e_0 \left( \nu^w(\delta, \nu), u \right) + we_1 \left( \nu^w(\delta, \nu), u \right) + \nu \nu^w(\delta, \nu) \delta, \tag{46} \]

and the Taylor expansion

\[ f^w(\delta, 0) = f^w(\delta, 0) + u \frac{df^w(\delta, u)}{du} \bigg|_{u=0} + \frac{u^2}{2} \frac{d^2f^w(\delta, u)}{du^2} \bigg|_{u=0} + \mathcal{O}(u^3), \tag{47} \]

where, according to Eq. (20),

\[ f^w(\delta, 0) = T^w_s(1 + \delta)/(2t) \]
\[ = -\sqrt{(1 - w)^2 - \delta^2}, \tag{48} \]

with the corresponding maximizing (KS) potential \( \nu^w(\delta, 0) = \frac{\delta}{\sqrt{(1 - w)^2 - \delta^2}} \).

From the stationarity condition in Eq. (52), which holds for any \( u \) and that, for clarity, we will rewrite as follows,

\[ \left[ (1 - w) \frac{\partial e_0(\nu, u)}{\partial \nu} + w \frac{\partial e_1(\nu, u)}{\partial \nu} \right]_{\nu = \nu^w(\delta, \nu)} = -\delta, \tag{50} \]

it comes

\[ \frac{df^w(\delta, u)}{du} = \left[ (1 - w) \frac{\partial e_0(\nu, u)}{\partial \nu} + w \frac{\partial e_1(\nu, u)}{\partial \nu} \right]_{\nu = \nu^w(\delta, \nu)}, \tag{51} \]

and

\[ \frac{d^2f^w(\delta, u)}{du^2} \bigg|_{u=0} = \left[ (1 - w) \frac{\partial^2 e_0(\nu, u)}{\partial \nu^2} + w \frac{\partial^2 e_1(\nu, u)}{\partial \nu^2} \right] \bigg|_{u=0} \]
\[ + \left[ (1 - w) \frac{\partial^2 e_0(\nu, u)}{\partial \nu \partial u} + w \frac{\partial^2 e_1(\nu, u)}{\partial \nu \partial u} \right] \bigg|_{u=0}, \tag{52} \]

where the linear response of the potential \( \partial \nu^w(\delta, u)/\partial u \) is determined by differentiating Eq. (50) with respect to \( u \).
thus leading to
\[
\frac{\partial \nu^u(\delta, u)}{\partial u} = - (1 - w) \frac{\partial^2 c_0(\nu, u)}{\partial \nu \partial u} + w \frac{\partial^2 c_1(\nu, u)}{\partial \nu^2} \bigg|_{\nu = \nu^u(\delta, u)},
\]

As shown in Appendix B, simple expressions (in terms of \(\delta\) and \(w\)) can be obtained for all energy derivatives, thus showing that the exact ensemble \(H_x\) energy is recovered through first order in \(u\), as expected, while the ensemble correlation energy (obtained through second order in \(u\)) reads
\[
e_w^u(\delta, u) = \frac{1}{2} \frac{d^2 f^u(\delta, u)}{d u^2} \bigg|_{u = 0} + \mathcal{O}(u)\]
\[
= - \frac{(1 - w)^2 - \delta^2}{2} \left[ \frac{1 + \delta^2}{(1 - w)^2} \left( 3 - \frac{4(1 - 3w)^2}{(1 - w)^2} \right) \right] + \mathcal{O}(u). \tag{54}
\]

Note that, by inserting the following expansion of \(g(u)\) (see Eq. (47)) into Eq. (40),
\[
g(u) = 2(1 + u) + \frac{5}{4} u^2 + \mathcal{O}(u^3), \tag{55}
\]

or by expanding the expression in Eq. (54) through second order in \(\delta\), we recover the same expression, as expected.
\[
e_w^u(\delta, u) = \frac{(1 - w)}{8} \delta^2 \left[ \frac{8(1 - 3w)^2 - 3(1 - w)^2}{16(1 - w)^3} \right] + \mathcal{O}(u, \delta^3). \tag{56}
\]

As readily seen from Eq. (56), in this regime of correlation, the ensemble density-functional correlation energy will be concave when \(\frac{1 - 4\sqrt{3}}{60} \approx 0.16 \leq w \leq \frac{1 - 4\sqrt{3}}{60} \approx 0.45\), and convex otherwise.

### 3.3 Strongly correlated limit

Let us, for convenience, consider the Legendre–Fenchel transform in Eq. (27) per unit of \(u\),
\[
\mathcal{T}^w(\delta) = f^w(\delta)/u = \sup_{\overline{\nu}} \left\{ (1 - w)\bar{c}_0(\overline{\nu}) + w\bar{c}_1(\overline{\nu}) + \overline{\nu}\delta \right\}, \tag{57}
\]

where \(\overline{\nu} = \nu/u\) and, according to Eq. (28), the \(\overline{\nu}\)-dependent ground- and first-excited-state energies are, in the strongly correlated limit \((u \to +\infty)\), solutions of
\[
\bar{e}_0 \times (\bar{e}^2 - 1) + 2 (\bar{e}_0)^2 = (\bar{e}_0)^3, \tag{58}
\]

thus leading to
\[
\mathcal{P}_0(\overline{\nu}) = \inf \left\{ 0, 1 - |\overline{\nu}| \right\}, \quad \mathcal{P}_1(\overline{\nu}) = \sup \left\{ 0, 1 - |\overline{\nu}| \right\}. \tag{59}
\]

Therefore,
\[
\mathcal{T}^w(\delta) \xrightarrow{u \to +\infty} \sup \left\{ \mathcal{T}^w_{\leq}(\delta), \mathcal{T}^w_{\geq}(\delta) \right\}, \tag{60}
\]

where, according to Eq. (45),
\[
\mathcal{T}^w_{\leq}(\delta) = \sup_{|\overline{\nu}| \leq 1} \left\{ (1 - w) (1 - |\overline{\nu}|) + \overline{\nu}\delta \right\}
\]
\[
\sup \left\{ \delta, -\delta \right\}, \tag{61}
\]

and
\[
\mathcal{T}^w_{\geq}(\delta) = \sup_{|\overline{\nu}| \leq 1} \left\{ w (1 - |\overline{\nu}|) + \overline{\nu}\delta \right\}
\]
\[
\sup \left\{ \delta, w - \delta \right\}. \tag{62}
\]

We conclude that
\[
\mathcal{T}^w(\delta) \xrightarrow{u \to +\infty} \sup \left\{ \delta, w - \delta \right\}. \tag{63}
\]

A graphical summary of Eq. (63) is given in Fig. 1. As readily seen, the functional will return \(w\) for densities in the range \(|\delta| \leq w\), thus leading to (see Eqs. (39) and (48),
\[
e_c^w(\delta) \xrightarrow{u \to +\infty} - \frac{1}{2} \left( (1 - w) - \frac{(3w - 1)\delta^2}{(1 - w)^2} \right). \tag{64}
\]

Interestingly, the Taylor expansion of the ensemble correlation energy through second order in \(\delta\) (see Eq. (40)) is becoming exact in the strongly correlated limit for densities in the range \(|\delta| \leq w\). As readily seen from Eq. (64), in this regime of correlation, the density-functional ensemble correlation energy is concave if \(0 < w \leq 1/3\) and convex otherwise. From Eqs. (45) and (63), we finally see that, as expected, the ensemble functional equals the ground-state one for densities in the range \(w \leq |\delta| \leq 1 - w:\)
\[
\mathcal{T}^w(\delta) = \mathcal{T}^w_{|\delta| = 0}(\delta) = |\delta|. \tag{65}
\]

As a result, in this density regime, the ensemble correlation energy reads
\[
e_c^w(\delta) \xrightarrow{u \to +\infty} |\delta| - \frac{1}{2} \left( (1 + w) - \frac{(3w - 1)\delta^2}{(1 - w)^2} \right). \tag{66}
\]

In the particular case \(|\delta| = 1/2\) and \(w = \frac{1}{2} - \eta\ (\eta > 0)\) which has been considered previously in Sec. 3.1, Eq. (66) is applicable, thus leading to the following Taylor expansion through first order in \(\eta,\)
\[
e_c^{w = \frac{1}{2} - \eta}(\pm \frac{1}{2}) \xrightarrow{u \to +\infty} -2\eta u. \tag{67}
\]
then two cases must be distinguished. Either $0 \leq w \leq |\delta|$ and, in this case, Eq. (70) applies, or $|\delta| \leq w \leq 1 - |\delta|$ and then Eq. (64) applies, thus leading to

$$\frac{1}{u} \frac{\partial e_w^u(\delta)}{\partial w} \bigg|_{w=|\delta|} \rightarrow \frac{1}{2} \left[ 1 + \frac{\delta^2(1 + 3w)}{(1 - w)^3} \right]. \quad (72)$$

Note that, as readily seen from Eqs. (70) and (72), and expected from Ref. [30], for densities that fulfill the condition in Eq. (71), there is a jump in the ensemble correlation energy derivative with respect to the weight $w$ when the latter crosses $|\delta|:

$$\left[ \frac{\partial e_w^u(\delta)}{\partial w} \bigg|_{w=|\delta|} - \frac{\partial e_w^u(\delta)}{\partial w} \bigg|_{w=|\delta|} \right] \rightarrow u \rightarrow +\infty. \quad (73)$$

Let us finally consider the particular case of the ensemble ($w = 1/2$) for which the derivative of the ensemble correlation energy must be taken at $w = \frac{1}{2} - \eta$ where $\eta \rightarrow 0^+$. In the strongly correlated limit (which also corresponds to the atomic $t = 0$ limit) we should have

$$\frac{1}{u} \frac{\partial e_w^u(\delta = \pm \frac{1}{2})}{\partial w} \bigg|_{w=\frac{1}{2}^-} \rightarrow 2, \quad (74)$$

while the expression in Eq. (72) predicts the (unphysical) result

$$\frac{1}{u} \frac{\partial e_w^u(\delta = \pm \frac{1}{2})}{\partial w} \bigg|_{w=\frac{1}{2}^+} \rightarrow 3. \quad (75)$$

Note that Eqs. (74) and (75) are in agreement with Eqs. (44) and (67).

### 4 Density-functional approximations and computational details

A summary of the various DFAs that will be tested on the Hubbard dimer in Sec. [5] is given here. The simplest approximation consists in using the (weight-independent) ground-state (GS) xc functional,

$$E_{x}^w(n) \rightarrow E_{x}^{w=0}(n) \quad \text{and} \quad E_{c}^w(n) \rightarrow E_{c}^{w=0}(n). \quad (76)$$

It will be referred to as GSxc. The other approximations will all use the (weight-dependent) ensemble exact exchange functional (see Eq. (22)). The ensemble exchange-only approximation ($E_{x}^w(n) \rightarrow 0$) will be referred to as EEXX. The ensemble correlation energy will then be modeled either at the (weight-independent) ground-state level,

$$E_{c}^w(n) \rightarrow E_{c}^{w=0}(n), \quad (77)$$

giving the GSc approximation, or with weight dependent functionals. In the latter case, we will use the perturbation theory expansion through second order (PT2) in the density deviation $\delta = n - 1$ from the symmetric case
[the expansion is given in Eq. (40) and will be referred to as \(\delta\text{-PT2}\) as well as the PT2 expansion in the weakly correlated regime, i.e. around \(w = U/(2t) = 0\) [the expansion is given in Eq. (54) and will be referred to as \(u\text{-PT2}\). All calculations have been performed with \(2t = 1\). The accurate parameterization of Carrascal et al. (see Eqs. (102)-(115) in Refs. [31,40]) has been used for the ground-state correlation functional in GSxc and GSc calculations. Excitation energies have been computed within the various approximations either by differentiation (see Eq. (15)) or by linear interpolation (see Eq. (14)). In the former case, the excitation energy reads as follows, according to Eq. (20),

\[
\frac{dE^w}{dw} = \frac{2t(1 - w)}{\sqrt{(1 - w)^2 - (1 - n_w)^2}} + \frac{\partial E^w_c(n)}{\partial w} \bigg|_{n = n_w} \nonumber
\]

\[
+ \frac{\partial E^w_v(n)}{\partial w} \bigg|_{n = n_w} . \tag{78}
\]

\section{5 Results and discussion}

In practical DFT calculations, the error in the energy is not only due to the approximate functional that is employed. It also depends on the deviation from the exact result of the density obtained by the minimization in Eq. (18), which is formally equivalent to solving the ensemble KS equations self-consistently. Therefore, in the following, we will distinguish the so-called functional driven error (Sec. 5.1), which is evaluated for a fixed density \(n\), from the density driven one, which will be discussed in the rest of this section.

\subsection{5.1 DFAs and functional driven error}

Functional driven errors have already been studied in Ref. [30] for both GSc and GSxc approximations. Density functional correlation energies obtained at the \(\delta\text{-PT2}\) level (see Eq. (40)) are shown in Fig. 2. As expected, accurate correlation energies are obtained around \(n = 1\). We also observe the changes in convexity when \(w\) increases in both weakly and strongly correlation regimes, as predicted by Eqs. (54) and (64), respectively. A major drawback of the \(\delta\text{-PT2}\) approximation is that it gives a non-zero (even positive) correlation energy at the border of the \(\nu\)-representability domain, which is of course unphysical (see Appendix C). In the light of Fig. 2 it is clear that \(\delta\text{-PT2}\) should only be applied to equi-ensembles (i.e. for \(w = 1/2\)). Even though, in that case, accurate correlation energies are obtained for a larger range of densities, in particular in the strongly correlated regime, a spurious positive contribution remains when \(n = 1/2\) or \(n = 3/2\) as \(U\) increases, as expected from Eq. (43) and illustrated in Fig. 3.

Turning to the \(u\text{-PT2}\) approximation (see Eq. (54) and Fig. 1), accurate correlation energies are obtained in the weakly correlated regime for all densities, as expected. Errors become large, especially around the symmetric \(n=1\) ensemble density, as \(U\) increases. Interestingly, the equi-ensemble seems to be less affected by the overestimation of the correlation energy than ensembles where the ground state dominates (i.e. \(w \ll 1/2\)). Finally, unlike \(\delta\text{-PT2}, u\text{-PT2}\) gives by construction (see Eq. (54)) the correct correlation energy which is equal to zero as shown in Appendix C at the border of the representability domain.

\subsection{5.2 Density-functional total energy profile and minimizing densities for equi-ensembles}

This section deals with the optimization of the ensemble density through minimization of the total ensemble density-functional energy in Eq. (19). For analysis purposes, exact ensemble densities have been plotted in Fig. 5 with respect to \(U\) and \(\Delta v_{\text{ext}}\) for various values of the ensemble weight \(w\). As expected from Ref. [30], the interacting density profile \((U > 0)\) satisfies the non-interacting \(\nu\)-representability condition in Eq. (23). Density domains can clearly be distinguished and, in particular, it appears that the ensemble density undertakes critical changes around \(U/\Delta v_{\text{ext}} \approx \pm 1\) and \(\Delta v_{\text{ext}} \approx 0\), which can be summarized as follows when \(U/(2t)\) is sufficiently large,

\[
\begin{cases}
  w & -1 < \frac{U}{\Delta v_{\text{ext}}} < 0 \\
  1 - w & \infty < \frac{U}{\Delta v_{\text{ext}}} < -1 \\
  1 & \text{for} \Delta v_{\text{ext}} \approx 0 \tag{79} \\
  1 + w & 1 < \frac{U}{\Delta v_{\text{ext}}} < +\infty \\
  2 - w & 0 < \frac{U}{\Delta v_{\text{ext}}} < +1.
\end{cases}
\]

Note that, in the particular case of the equi-ensemble \((w = 1/2)\), ensemble densities will essentially be equal to \(1/2, 1\) (in the vicinity of the symmetric case) or \(3/2\).

Let us now focus on the approximate calculation of ensemble densities. Calculating the ensemble energy profiles for the set of non-interacting \(\nu\)-representable ensemble densities within all aforementioned approximations will allow us to detect possible local minima that can lead to wrong minimizing ensemble densities and convergence issues. In the exact theory, both non-interacting kinetic and xc functionals are weight-dependent so that the total density-functional energy is strictly convex. In practical calculations, however, there is no straightforward way to develop weight-dependent functionals and one has to recur to approximations such as neglecting the weight dependence, like in GSxc (see Sec. 4). In the following, we discuss what effect the neglect or the (partial) introduction of weight dependence in the xc functional has on the profile of the total ensemble energy.

A selection of peculiar and problematic cases are plotted in Fig. 6. The minimizing ensemble densities are obtained by global brute-force minimization and are plotted
Comparison is made with the exact results (solid lines) of Ref. [30].

As both the kinetic and the exact ground-state functionals are strictly convex. Nevertheless, it gives quite poor equi-ensemble energies (see Fig. [1]), which is due to the fact that, for \( w = 1/2 \), the excited state contributes to half of the ensemble energy and thus the weight dependence cannot be completely neglected. The minimizing ensemble densities are correct in the symmetric case (top panel of Fig. [6]) but as soon as \( \Delta v_{\text{ext}} \) increases they are too far off from the exact ones (see Fig. [7]). Turning to the GSX approximation (see Sec. [4]), the equi-ensemble energy profile is not strictly convex for all values of \( U \) (see Fig. [8]). On the one hand, adding the EEXX to the ground-state correlation functional yields better equi-ensemble energies than GSXc but they are still too poorly described. The minimizing ensemble densities, on the other hand, are exact in the symmetric case and in asymmetric cases where \( \Delta v_{\text{ext}} \) is sufficiently large compared to \( U \) (see Fig. [7]). However, in the intermediate case, i.e., when \( U \gtrsim \Delta v_{\text{ext}} \) (see the middle panel of Fig. [6]), the global minimum abruptly switches place with another minimum located at \( n = 1 \) and causes the discontinuity at \( U = 6 \) in the plot of the minimizing ensemble density as a function of \( U \) (see the middle panel of Fig. [7]). Let us stress that, even when GSX gives the right density by global minimization, the existence of local minima and maxima in the strongly correlated regime

Fig. 2. \( \delta \)-PT2 correlation energy (dashed lines) plotted as a function of the density for various correlation regimes and ensemble weights. Comparison is made with the exact results (solid lines) of Ref. [30].

Fig. 3. \( \delta \)-PT2 correlation energy (dashed lines) plotted as a function of \( u = U/(2t) \) for \( n = 1.5 \) and weights close to \( 1/2 \). Comparison is made with the exact results (solid lines).
whereas for ensemble energy has a unique minimizing ensemble density, top panel of Fig. 6). In the specific case where water minima appear on the ensemble energy profile (see the which the strict convexity is suppressed and two degener-

ity in the top panel of Fig. 7. Away from the symmet-

uality thanks to the hole-particle symmetry (see Appendix D). When

energy contribution.

complete neglect of weight dependence in the correlation en-

of the equi-ensemble energy profile induced by the com-

potential at

ground-state xc

due not only to the discontinuity in the ground-state xc

functionnal:

turning on the

will lead to serious convergence issues when searching for

stationary densities, which would be equivalent to solv-
ing the ensemble KS equations self-consistently. This is
due not only to the discontinuity in the ground-state xc

potential at $n = 1$ [30,31] but also to the non-convexity of the equi-ensemble energy profile induced by the com-

plete neglect of weight dependence in the correlation en-

ergy contribution.

Neglecting the correlation energy in the GSc scheme

leads to the EEXX approximation. In the latter case, the

minimization can be carried out analytically for the sym-

metric dimer (see Appendix D). When $w \leq 1/3$, the en-

semble energy has a unique minimizing ensemble density, whereas for $w > 1/3$ there is a critical value of $U$ beyond

which the strict convexity is suppressed and two degener-

ate minima appear on the ensemble energy profile (see the
top panel of Fig. 4). In the specific case where $w = 1/2$, this value is $U = 1$. This abrupt change explains why the

EEXX minimizing ensemble density exhibits a discontinuity in the top panel of Fig. 4. Away from the symmet-
ic case, the EEXX equi-ensemble energy has the correct
global minimum even though it exhibits non-convexity.

Note that, as shown in Appendix C, the equi-ensemble

EEXX energy is exact at the border of the non-interacting $\nu$-representability domain, i.e. when $n = 1/2$ or $n = 3/2$.

Turning to the weight-dependent $\delta$-PT2 correlation DFA (see Eq. (40)), the equi-ensemble energy exhibits convex-

ity in both weakly and strongly correlated regimes (see Fig. 4). Thus, unlike GSc and EEXX, the minimization scheme is robust and does not lead to discontinuities in the minimizing ensemble densities. $\delta$-PT2 is essentially exact around the symmetric case, by construction. Errors ap-
ppear in the minimizing density when $U \gg \Delta v_{\text{ext}} > 0$ (see the middle panel of Fig. 7). As in GSc, as soon as $\Delta v_{\text{ext}}$ is sufficiently large, the exact minimizing ensemble den-
sity is almost recovered. Moreover, thanks to the absence of density derivative discontinuities in the $\delta$-PT2 correla-

tional functional (see Fig. 2), self-consistent calculations of (stationary) ensemble densities in the strongly correlated regime are expected to converge smoothly, which is clearly an advantage from a practical point of view.

Let us finally discuss the performance of the $\nu$-PT2 approximation which uses a weight dependent density-

functional correlation energy based on a perturbative ex-

pansion of the exact correlation energy around $U = 0$ (see
Fig. 5. Map of the exact ensemble density plotted as a function of $U$ and $\Delta v_{\text{ext}}$ for various ensemble weights.

As expected, $u$-PT2 performs well in the weakly correlated regime. As soon as $U$ increases, it faces the same problem as Gsc and EEXX (see middle and bottom panels of Fig. 6). Indeed, the ensemble energy loses convexity (local minima and maxima appear), which leads to discontinuities in the minimizing ensemble densities (see the middle and bottom panels of Fig. 7).

In summary, including weight dependence into the ensemble correlation energy is crucial in order to obtain quantitatively good results (densities and energies) and avoid potential convergence issues when searching for stationary densities of the total energy or, equivalently, when solving the self-consistent ensemble KS equations. Furthermore, keeping only the weight dependence in the exchange part has proven to be insufficient. The best reproduction of the exact equi-ensemble energy profiles and minimizing ensemble densities is by far obtained by the $\delta$-PT2 approximation. It is valid for both the weakly and strongly correlated regime and, despite being based on an expansion around $n = 1$, it also yields decent results in the asymmetric case.

5.3 Ensemble energy derivatives

In practice, any weight in the range $0 \leq w \leq 1/2$ can in principle be used for computing the excitation energy. As argued in Sec. 5.1, we expect the equi-ensemble case ($w = 1/2$) to be the most favorable one for the DFAs discussed previously, especially $\delta$-PT2 (see Eq. (40)). We focus in this section on the calculation of approximate excitation energies by differentiation (see Eq. (78)). In order to evaluate both functional driven and total errors, results obtained with the exact and the minimizing ensemble densities are shown in Figs. 8 and 9 respectively. In addition, the difference between the two excitation energies is plotted in Fig. 10 in order to visualize the impact of density driven errors.

As shown in Fig. 9, all the DFAs using a weight independent ensemble correlation energy (namely GSxc, GSc, and EEXX) underestimate the excitation energy. Unphysical negative excitation energies are even obtained with GSc and EEXX in the strongly correlated regime, as expected [30]. In the latter regime, EEXX exhibits large density driven errors in the symmetric case only, in agreement with Sec. 5.2. In the asymmetric case, the error is purely functional driven. The opposite is observed for GSxc. Note that, at the GSc level of approximation and for $\Delta v_{\text{ext}} = 1$ (see the middle panel of Fig. 9), the excitation energy exhibits a discontinuity around $U = 6$, as expected from Sec. 5.2. Interestingly, even though GSc gives a completely wrong ensemble density in this regime of correlation, the accumulation of functional and density driven errors (see the middle panels of Figs. 8 and 10) leads to relatively
Fig. 6. Exact and approximate total equi-ensemble density-functional energies plotted for various interaction strengths and external potentials. See text for further details.

Fig. 7. Exact and approximate equi-ensemble minimizing densities plotted as a function of $U$ for various external potentials. See text for further details.

Turning to weight-dependent correlation DFAs, $u$-PT2 (see Eq. (54)) performs well only for relatively small $U$ values, as expected. The discontinuities observed for large $U$ values in asymmetric cases are induced by sudden changes in the minimizing ensemble density as $U$ increases (see Sec. 5.2 for further details). Unlike GSc, $u$-PT2 does not benefit from error cancellations in the strongly correlated regime. In the asymmetric case, the excitation energies are indeed significantly overestimated (see the middle and bottom panels of Fig. 9). Note that taking into account functional driven errors only would lead to negative excitation energies in this case (see the middle and bottom panels of Fig. 8). Thanks to (too) large additional density driven errors (see the middle and bottom panels in Fig. 10), positive excitation energies are finally obtained.

Let us now focus on the $\delta$-PT2 approximation (see Eq. (40)). It is, by construction, exact for the symmetric dimer. In the asymmetric $\Delta v_{\text{ext}} = 5$ case, however, $\delta$-PT2 overestimates the excitation energy significantly as $U$ increases. This was actually expected from Eqs. (74) and (75) since the ensemble density is, in this case, close to 3/2 (see the bottom panel of Fig. 7). Interestingly, the density driven error is substantial in this case (see the middle and bottom panels of Fig. 8). As readily seen from Eq. (78), the non-interacting kinetic energy contri-
5.4 Linear interpolation method

The linear interpolation method (LIM) [24] is an alternative to the differentiation of the ensemble energy for the extraction of excitation energies. As readily seen from Eq. (14), the latter are calculated within LIM from both ground-state and equi-ensemble energies. Since we use the accurate parameterization of Carrascal et al. [31] for the ground-state correlation functional, errors in our LIM excitation energies will exclusively originate from the ensemble xc DFA that is used. Results obtained with the exact and minimizing ensemble densities are shown in Fig. 11. In the symmetric case (top panel), $\delta$-PT2 (see Eq. (10)) is exact in all correlation regimes while $u$-PT2 (see Eq. (54)) performs well only for relatively small $U$ values, as expected. The lack of weight dependence in GSxc and GSc leads to an underestimation of the excitation energy. EEXX performs surprisingly well in this case, even though it exhibits large functional driven and density driven errors. As shown in Fig. 12, these errors cancel each other as $U$ increases. Note that the interaction derivative discontinuities around $U = 1$ in the non-interacting and Hxc ensemble energies originate from the sudden change in the minimizing ensemble density discussed previously (see the top panel of Fig. 7).

Turning to asymmetric cases (see the middle and bottom panels of Fig. 11), EEXX and exact excitation energy curves are essentially on top of each other. This is simply due to the fact that, as $U$ increases, the ensemble density becomes very close to $3/2$ so that the equi-ensemble correlation energy vanishes (see Appendix C). Note that, at the EEXX level of approximation, both functional and density driven errors are equal to zero in this case. Interestingly, inserting the exact equi-ensemble density into the GSxc functional gives relatively good results in the strongly correlated regime, as expected [30]. However, as shown in Fig. 13, large density driven errors lead to a significant underestimation of the excitation energy in this regime. Note finally that, as expected from Ref. [30], GSc systematically underestimates the excitation energy. Regarding the weight dependent correlation DFAs, $u$-PT2 performs as well as EEXX when the exact ensemble density (which is close to $3/2$) is used, as expected from Eq. (54). Unlike EEXX, $u$-PT2 suffers from significant density driven errors (see Fig. 13) when $U$ is sufficiently large, thus leading to a deterioration of the excitation energy. On the other
hand, δ-PT2 gives relatively accurate results in the same regime of density and correlation. As shown in Fig. 13 and expected from Eqs. (14), (43) and (67), a residual error (equal to -0.5 per unit of 2T) when exact densities are used is obtained as \( U \) increases, which is due to an unphysical positive correlation energy contribution. Note that the latter error is essentially functional driven.

6 Conclusions

Ensemble DFT for excited states has been applied to the two-electron Hubbard dimer. An ensemble consisting of the first two singlet states has been considered. While Ref. 30 was focusing on the exact calculation of (weight-dependent) correlation energies, the design of analytical density-functional approximations (DFAs) as well as their performance in practical (self-consistent) calculations has been investigated in this work. Simple DFAs based on the expansion of the exact ensemble Hohenberg–Kohn functional in various density and correlation regimes have been considered. Note that the use of Legendre–Fenchel transforms (rather than the usual Levy–Lieb constrained-search formalism which would involve two many-body wavefunctions) is extremely convenient since a single variable, namely the local potential difference \( \Delta v \), is needed (see Eq. 25).

As clearly shown in this simple but nontrivial model, the equi-ensemble case is the simplest one to model, as long as the ensemble density remains away from the border of the non-interacting \( \rho \)-representability domain. If so, the most reliable DFA, referred to as δ-PT2, uses an expansion through second order in the density deviation \( \delta = n - 1 \)
from the symmetric case (see Eq. (40)). It performs relatively well in both weakly and strongly correlated regimes and is, by construction, exact for the symmetric dimer.

Excitation energies have been extracted in two ways. The differentiation with respect to the ensemble weight (see Eq. (15)) does not always give satisfactory results, especially in the asymmetric strongly correlated regime. A proper description of the weight-dependent correlation density derivative discontinuities in the strongly correlated regime would be needed. Those actually appear at the border of the density domain of applicability of \( \delta \text{-PT2} \). The second approach (see Eq. (14)), namely the linear interpolation method (LIM), is much more reliable especially since it avoids the difficult task of modelling the xc derivative discontinuity. Despite a spurious positive correlation energy contribution which appears in the asymmetric case as the on-site repulsion increases, combining \( \delta \text{-PT2} \) with LIM gives relatively accurate results. Note that this spurious contribution originates from the fact that the equi-ensemble density is outside the domain of applicability of \( \delta \text{-PT2} \). Better DFAs would be obtained by interpolating the expansions obtained in the various regimes. This is left for future work.

Let us finally stress that the combination of the ensemble exact exchange-only energy (EEXX) [see Eq. (22)] with LIM yields remarkably accurate excitation energies, even in the strongly correlated regime. In the symmetric case, both functional and density driven errors are large but they actually cancel each other.

Finally, we would like to stress that the technique we used for deriving the \( \delta \text{-PT2} \) functional (which relies on the Legendre–Fenchel transform formalism) is expected to be useful also for the development of \textit{ab initio} weight-dependent exchange and correlation DFAs in finite model systems like electrons on a hypersphere [41]. Work is currently in progress in this direction.

Acknowledgments

This work was funded by the Ecole Doctorale des Sciences Chimiques 222 (Strasbourg) and the ANR (MCFUNEX project, Grant No. ANR-14-CE06-0014-01). The authors are pleased to dedicate this work to Hardy Gross on the occasion of his 65th birthday.

7 Authors contributions

All the authors were equally involved in the preparation of the manuscript (derivation of equations, implementation of the various approximations, test calculations and writing). All the authors have read and approved the final manuscript.

A Appendix: simplified expression for

\[
\frac{d^2 f_{\text{ex}}(\delta)}{d\delta^2} \bigg|_{\delta=0}
\]

By differentiating Eq. (34) with respect to \( \nu \) and taking \( \nu = \nu(0) = 0 \) it comes

\[
\frac{\partial^2 e_i(\nu)}{\partial \nu^2} \bigg|_{\nu=0} \times \left[ 1 - u^2 + 4u e_i(0) - 3e_i^2(0) \right] = -2e_i(0),
\]

since, according to Eq. (34), \( \partial e_i(\nu)/\partial \nu \bigg|_{\nu=0} = 0 \). In the particular (symmetric) case \( \nu = 0 \), Eq. (28) becomes

\[
e_i = u \quad \text{or} \quad e_i^2 - u e_i - 1 = 0,
\]

thus leading to

\[
e_0(0) = \frac{u - \sqrt{u^2 + 4}}{2}
\]

and

\[
e_1(0) = u.
\]

In the following, we will use the shorthand notation \( e = e_0(0) \) for convenience. Consequently, we obtain from Eq. (A.1) the following explicit expressions,

\[
\frac{\partial^2 e_0(\nu)}{\partial \nu^2} \bigg|_{\nu=0} = -\frac{2e}{1 - u^2 + 4ue - 3e^2},
\]

or, equivalently (see Eq. (A.2)),

\[
\frac{\partial^2 e_0(\nu)}{\partial \nu^2} \bigg|_{\nu=0} = \frac{2e}{2 + u^2 - ue},
\]

and

\[
\frac{\partial^2 e_1(\nu)}{\partial \nu^2} \bigg|_{\nu=0} = -2u.
\]
Combining Eqs. (31), (33), (A.6), and (A.7) leads to
\[
\frac{d^2 f^w(\delta)}{d\delta^2} \bigg|_{\delta=0} = \frac{1}{2} \frac{2 + u^2 - w \epsilon}{w \epsilon (2 + u^2 - w \epsilon)} - e. \tag{A.8}
\]
Finally, by using the simplified expression,
\[
2 + u^2 - w \epsilon = 1 + \left( \frac{u}{2} + \sqrt{1 + \left( \frac{u}{2} \right)^2} \right)^2,
\]
we obtain (see Eq. (37))
\[
\frac{2 + u^2 - w \epsilon}{e} = -g(u), \tag{A.10}
\]
thus leading to the final expression in Eq. (36).

**B Appendix: simplified expressions for**

\[
\frac{d f^w(\delta, u)}{d u} \bigg|_{u=0} \quad \text{and} \quad \frac{d^2 f^w(\delta, u)}{d u^2} \bigg|_{u=0}
\]

According to Eq. (28), the individual energies read as follows in the non-interacting case (\(u = 0\)),
\[
e_0(\nu, 0) = -\sqrt{1 + \nu^2},
\]
\[
e_1(\nu, 0) = 0,
\]
so that
\[
\frac{\partial e_0(\nu, u)}{\partial \nu} \bigg|_{u=0} = -\frac{\nu}{\sqrt{1 + \nu^2}},
\]
\[
\frac{\partial e_1(\nu, u)}{\partial \nu} \bigg|_{u=0} = 0,
\]
and
\[
\frac{\partial^2 e_0(\nu, u)}{\partial \nu^2} \bigg|_{u=0} = -\frac{1}{(1 + \nu^2)^{3/2}},
\]
\[
\frac{\partial^2 e_1(\nu, u)}{\partial \nu^2} \bigg|_{u=0} = 0. \tag{B.3}
\]
Moreover, we obtain the following expressions from the differentiation of Eq. (28) with respect to \(u\) and/or \(\nu\):
\[
\frac{\partial e_0(\nu, u)}{\partial u} \bigg|_{u=0} = \frac{1 - 2e^2_0(\nu, 0)}{1 + \nu^2 - 3e^2_0(\nu, 0)} = \frac{\nu^2 + 1}{\nu^2 + 1}, \tag{B.4}
\]
\[
\frac{\partial e_1(\nu, u)}{\partial u} \bigg|_{u=0} = \frac{1}{\nu^2 + 1}, \tag{B.5}
\]
\[
\frac{\partial^2 e_0(\nu, u)}{\partial u^2} \bigg|_{u=0} = \frac{2e_0(\nu, 0)}{1 + \nu^2 - 3e^2_0(\nu, 0)} \times \left[ 1 + \frac{\partial e_0(\nu, u)}{\partial u} \bigg|_{u=0} \left( 3 \frac{\partial e_0(\nu, u)}{\partial u} \bigg|_{u=0} - 4 \right) \right] = -\frac{\nu^2 + 1}{(1 + \nu^2)^{5/2}}. \tag{B.6}
\]
By inserting Eq. (49) into the latter equation we finally recover, as expected, the expression for the exact ensemble Hx energy (see Eq. (39)) per unit of \(u\):
\[
\frac{d f^w(\delta, u)}{d u} \bigg|_{u=0} = \frac{e^w_{Hx}(\delta)}{u}. \tag{B.11}
\]
Turning to the ensemble correlation energy, it comes from Eqs. (53), (B.3), (B.8), and (B.9) that
\[
\frac{\partial^w(\delta, u)}{\partial \nu} \bigg|_{u=0} = \frac{\nu(1 - 3w)}{(1 - w)^{1/2}} \left( 1 + \nu^2 + 2 \right) \bigg|_{\nu = w^u(\delta, 0)}, \tag{B.12}
\]
which, according to Eqs. (52), (B.6), and (B.7) leads to
\[
\frac{d^2 f^w(\delta, u)}{d u^2} \bigg|_{u=0} = \frac{(1 - w)}{4(1 + \nu^2)^{3/2}} \times \left( 1 + 4\nu^2 - \frac{4\nu^2(1 - 3w)^2}{(1 - w)^2} \right) \bigg|_{\nu = w^u(\delta, 0)}. \tag{B.13}
\]
Finally, by using the following relations (see Eq. (49)),
\[
\frac{1}{1 + \nu^2(\delta, 0)} = \frac{1 - w^2 - \delta^2}{(1 - w)^2},
\]
\[
\frac{1}{1 + \nu^u(\delta, 0)} = 1 + \frac{3\delta^2}{(1 - w)^2}, \tag{B.14}
\]
\[
\frac{1}{1 + \nu^2(\delta, 0)} = \frac{\delta^2}{(1 - w)^2}.
\]
we recover the expression in Eq. (54).

**C Appendix: correlation energy at the border of the \( ν \)-representability domain**

As readily seen from Eq. (23), at the border of the non-interacting \( ν \)-representability domain, the density is such that \( |n - 1| = 1 - w \) or, equivalently,

\[
 n = 1 \pm (1 - w). \tag{C.1}
\]

When \( |\Delta ν|/t \to +\infty \) and \( |\Delta ν| > U \), the ground- and first-excited state energies read as follows, according to Eq. (17),

\[
 E_0(\Delta ν) = U - |\Delta ν|,
\]

\[
 E_1(\Delta ν) = 0, \tag{C.2}
\]

and, consequently (see Eq. (33)),

\[
 \frac{\partial E_0(\Delta ν)}{\partial \Delta ν} = -\frac{\Delta ν}{|\Delta ν|},
\]

\[
 \frac{\partial E_1(\Delta ν)}{\partial \Delta ν} = 0. \tag{C.3}
\]

Thus we conclude that the stationarity condition in Eq. (32) is fulfilled for \( \delta = n - 1 = \pm (1 - w) \) when \( |\Delta ν|/t \to +\infty \) and \( \Delta ν/(n - 1) \) is positive. The resulting ensemble Legendre–Fenchel transform (see Eq. 25) reads

\[
 (1 - w) \left( U - |\Delta ν| \right) \pm \Delta ν \left( 1 - w \right) \to -\infty \quad F^w \left( 1 \pm (1 - w) \right) = U(1 - w). \tag{C.4}
\]

Since, according to Eq. (20),

\[
 T^w = 1 \pm (1 - w), \tag{C.5}
\]

it comes from Eqs. (22) and (C.4),

\[
 \left[ F^w(n) - T^w \right]_{n = 1 \pm (1 - w)} = \left[ E_\Pi(n) + E^w_\nu(n) \right]_{n = 1 \pm (1 - w)}, \tag{C.6}
\]

or, equivalently,

\[
 E^w_\nu \left( 1 \pm (1 - w) \right) = 0. \tag{C.7}
\]

**D Appendix: EEXX ensemble energy minimization in the symmetric case**

For \( |\Delta ν| = 0 \), the minimization of the (approximate) EEXX ensemble energy leads to the following equation:

\[
 (n - 1) \left[ U(1 - 3w) + \frac{2t}{(1 - w)^2 - (1 - n)^2} \right] = 0. \tag{D.1}
\]

After factoring out the obvious solution \( n = 1 \), we are left with a quadratic equation. The discriminant reads

\[
 \Delta = 4U^2(1 - 3w)^2(w - 1)^2 \left[ U^2(1 - 3w)^2 - 4t^2(w - 1)^2 \right] \tag{D.2}
\]

and is zero for \( U = 0 \) and the critical value

\[
 U_{crit} = \frac{2t(1 - w)}{3w - 1}. \tag{D.3}
\]

The second derivative of the ensemble energy functional with respect to \( n \) contains all the information about the convexity:

\[
 \frac{d^2E_{EEXX}(n)}{dn^2} = \frac{U(1 - 3w)}{(1 - w)^2} + \frac{2t(1 - w)^2}{[(1 - w)^2 - (1 - n)^2]^{3/2}}. \tag{D.4}
\]

For \( U \leq U_{crit} \), the EEXX ensemble energy is strictly convex and has exactly one unique global minimum \( (n = 1) \) whereas for \( U > U_{crit} \) the quadratic equation possesses two solutions,

\[
 n = 1 \pm \frac{\sqrt{3}}{2U^2(3w - 1)^2}, \tag{D.5}
\]

which leads to two degenerate minima. The other solution \( n = 1 \) is a maximum in this case (see Eq. (D.4) and the top panel of Fig. 6). We notice that for any \( w \leq 1/3 \), the ensemble energy within the EEXX approximation has always one global minimum, independently of \( U \). In the case of equi-ensembles \( (w = 1/2) \), there is one unique solution as long as \( U \leq 2t \).

**References**

1. A.K. Theophilou, J. Phys. C (Solid State Phys.) 12, 5419 (1979)
2. E.K.U. Gross, L.N. Oliveira, W. Kohn, Phys. Rev. A 37, 2805 (1988)
3. E.K.U. Gross, L.N. Oliveira, W. Kohn, Phys. Rev. A 37, 2809 (1988)
4. E.K.U. Gross, L.N. Oliveira, W. Kohn, Phys. Rev. A 37, 2821 (1988)
5. O. Franck, E. Fromager, Mol. Phys. 112, 1684 (2014)
6. A. Pribram-Jones, Z. Hui Yang, J. R. Traill, K. Burke, R. J. Needs, C. A. Ullrich, J. Chem. Phys. 140, 18A541 (2014)
7. Z. H. Yang, J. R. Traill, A. Pribram-Jones, K. Burke, R. J. Needs, C. A. Ullrich, Phys. Rev. A 90, 042501 (2014)
8. K. Pernal, N. I. Gidopoulos, E. Pastoreczak, Adv. Quantum Chem. (2015)
9. Z. H. Yang, A. Pribram-Jones, K. Burke, C. A. Ullrich, Phys. Rev. Lett. 119, 033003 (2017)
10. A. Nikiforov, J.A. Gamez, W. Thiel, M. Huix-Rotllant, M. Filatov, J. Chem. Phys. 141, 124122 (2014)
11. M. Filatov, WIREs Comput Mol Sci 5, 146 (2015)
12. M. Filatov, M. Huix-Rotllant, I. Burghardt, J. Chem. Phys. 142, 184104 (2015)
13. M. Filatov, F. Liu, K.S. Kim, T.J. Martínez, J. Chem. Phys. 145, 244104 (2016)
14. M. Filatov, T.J. Martínez, K.S. Kim, J. Chem. Phys. 147, 064104 (2017)
15. T. Gould, S. Pittalis, Phys. Rev. Lett. 119, 243001 (2017)
16. T. Gould, L. Kronik, S. Pittalis, arXiv preprint arXiv:1801.06314 (2018)
17. J.P. Perdew, M. Levy, Phys. Rev. B 31, 6264 (1985)
18. M. Levy, A. Nagy, Phys. Rev. Lett. 83, 4361 (1999)
19. M. Levy, A. Nagy, Phys. Rev. A 59, 1687 (1999)
20. A. Nagy, M. Levy, Phys. Rev. A 63, 052502 (2001)
21. P.W. Ayers, M. Levy, Phys. Rev. A 80, 012508 (2009)
22. P.W. Ayers, M. Levy, A. Nagy, Phys. Rev. A 85, 042518 (2012)
23. E. Pastorczak, N.I. Gidopoulos, K. Pernal, Phys. Rev. A 87, 062501 (2013)
24. B. Senjean, S. Knecht, H.J.Aa. Jensen, E. Fromager, Phys. Rev. A 92, 012518 (2015)
25. B. Senjean, E.D. Hedegård, M.M. Alam, S. Knecht, E. Fromager, Mol. Phys. 114, 968 (2016)
26. N. Gidopoulos, P. Papaconstantinou, E. Gross, Phys. Rev. Lett. 88, 033003 (2002)
27. E. Pastorczak, K. Pernal, J. Chem. Phys. 140, 18A514 (2014)
28. M.M. Alam, S. Knecht, E. Fromager, Phys. Rev. A 94, 012511 (2016)
29. M.M. Alam, K. Deur, S. Knecht, E. Fromager, J. Chem. Phys. 147, 204105 (2017)
30. K. Deur, L. Mazouin, E. Fromager, Phys. Rev. B 95, 035120 (2017)
31. D.J. Carrascal, J. Ferrer, J.C. Smith, K. Burke, J. Phys. Condens. Matter 27, 393001 (2015)
32. J.C. Smith, A. Pribram-Jones, K. Burke, Phys. Rev. B 93, 245131 (2016)
33. E. Fromager, Mol. Phys. 113, 419 (2015)
34. B. Senjean, M. Tsuchizu, V. Robert, E. Fromager, Mol. Phys. 115, 48 (2017)
35. B. Senjean, N. Nakatani, M. Tsuchizu, E. Fromager, arXiv preprint arXiv:1710.03125 (2017)
36. E.H. Lieb, Int. J. Quantum Chem. 24, 243 (1983)
37. M. Levy, Proc. Natl. Acad. Sci. 76, 6062 (1979)
38. M. Levy, Phys. Rev. A 52, R4313 (1995)
39. J. Hubbard, Proc. R. Soc. London A 276, 238 (1963)
40. D. Carrascal, J. Ferrer, J. Smith, K. Burke, J. Phys. Condens. Matter 29, 019501 (2016)
41. P.F. Loos, P.M. Gill, Phys. Rev. Lett. 103, 123008 (2009)