N-representability of the target density in Frozen-Density Embedding Theory based methods: numerical significance and its relation to electronic polarisation

Niccolò Ricardi,[1, a)] Cristina E. González-Espinoza,[1, b)] and Tomasz Adam Wesołowski[1, c)]

Department of Physical Chemistry, University of Geneva, Geneva (Switzerland)

(Dated: 20 June 2022)

The accuracy of any observable derived from multi-scale simulations based on Frozen-Density Embedding Theory (FDET) is affected by two inseparable factors: *i*) the approximation for the $E_{xcT}^{nad}[\rho_A, \rho_B]$ component of the FDET energy functional and *ii*) the choice of the density $\rho_B(\mathbf{r})$ for which the FDET eigenvalue equation for the embedded wavefunction is solved. A procedure is proposed to estimate the relative the significance of these two factors. Numerical examples are given for four weakly bound intermolecular complexes. It is shown that the violation of the non-negativity condition is the principal source of error in the FDET energy if ρ_B is the density of the isolated environment, i.e. is generated without taking into account the interactions with the embedded species. Reduction of both the magnitude of the violation of the non-negativity condition and the error in the FDET energy can be pragmatically achieved by means of the explicit treatment of the electronic polarisation of the environment.

[a)]Electronic mail: Niccolo.Ricardi@unige.ch

[b)]Electronic mail: Cristina.GonzalezEspinoza@unige.ch

[c)]Electronic mail: Tomasz.Wesolowski@unige.ch

I. INTRODUCTION

The formal framework of Frozen-Density Embedding Theory (FDET) provides the basis of multi-scale/multi-level simulation methods that use a multiplicative embedding operator (see the review in Ref. 1). The self-consistent expressions for the functional of the total energy and the corresponding embedding potential are available for various possible quantum descriptors used for the embedded N_A-electron system[2–5]. In FDET, the total energy of a system comprising N_{tot} electrons in the external potential v is given as a bi-functional ($E_v^{FDET}[\Psi_A, \rho_B]$) depending on two types of variables: *i*) ρ_B - a non-negative function integrating to $N_B < N_{tot}$ and *ii*) Ψ_A - an N_A-electron wavefunction. It is convenient to refer to Ψ_A as the *embedded wavefunction* and to ρ_B as the *density of the environment*. By construction, the $E_v^{FDET}[\Psi_A, \rho_B]$ is consistent with the the Hohenberg-Kohn energy functional ($E_v^{HK}[\rho]$) known in the density-functional theory[6] formulation of quantum N-electron problem (see Eq. 2 below). Throughout this work, $E_v^{HK}[\rho]$ denotes the Hohenberg-Kohn functional with v specifying the external potential. Any multi-level simulation based on the formal framework of FDET hinges on two types of approximations/assumptions:

I: the used approximation for the $E_{xcT}^{nad}[\rho_A, \rho_B]$ component of $E_v^{FDET}[\Psi_A, \rho_B]$ and

II: the choice for ρ_B.

The effect on the energy due to approximating the exact bi-functional $E_{xcT}^{nad}[\rho_A, \rho_B]$ and the corresponding functional derivative ($\frac{\delta E_{xcT}^{nad}[\rho_A,\rho_B]}{\delta\rho(\mathbf{r})}$) by means of some analytic expression ($\tilde{E}_{xcT}^{nad}[\rho_A, \rho_B]$) can be either positive and negative. The choice of ρ_B, on the other hand, affects always the total energy in the same way. As a consequence of the second Hohenberg-Kohn theorem, the optimal energy evaluated by means of $E_v^{FDET}[\Psi_A, \rho_B]$ is the upper bond of the ground-state energy of the total system (E_v^o). The magnitude of the difference between the upper bound and E_v^o depends on ρ_B. To reach zero, one of the necessary conditions is that the difference between the ground-state density of the whole system (ρ_v^o) and ρ_B is non-negative.

In FDET based multi-level simulations, ρ_B is generated using other methods than those used to optimise Ψ_A. The effect on the energy due to both factors (**I** and **II**) combine and their relative significance cannot, therefore, be determined in a straightforward matter. Concerning **II**, we define the *target density* to be the difference $\rho_v^{o(ref)} - \rho_B$, where $\rho_v^{o(ref)}$

is the density of the total system considered as a reference. In case of the exact theory, $\rho_v^{o(ref)}$ is the exact ground-state density of the whole system. For the sake of numerical analysis, $\rho_v^{o(ref)}$ can be the density of the same system obtained with some method from the Quantum Chemistry toolbox, in which the electron-electron correlation is treated in the same way as in the FDET based calculations. The necessary condition that the FDET energy and density are equal to the corresponding references is that the target density $\rho_v^{o(ref)} - \rho_B$ is N-representable (with $N = N_A = N_{tot} - N_B$). This condition is satisfied if $\rho_v^{o(ref)} - \rho_B$ is non-negative on every measurable volume element. Other conditions are of technical (the use of the same basis sets for instance) or mathematical (v-representability of $\rho_v^{o(ref)} - \rho_B$) nature.

In the present work, we propose a numerical procedure to estimate the relative importance of these two effects. The principal questions addressed are: *What is the effect on energy if* $\rho_v^{o(ref)} - \rho_B$ *is locally negative? How does this effect compare to that due to* $E_{xcT}^{nad}[\rho_A, \rho_B] \approx \tilde{E}_{xcT}^{nad}[\rho_A, \rho_B]$*?* These issues have not been studied systematically in the literature. The condition of non-negativity of $\rho_v^{o(ref)} - \rho_B$ cannot be verified in a multi-level simulation because it would require *a priori* knowledge of $\rho_v^{o(ref)}$. For a small model system, for which a given conventional quantum chemistry method can be used to provide $E_v^{o(ref)}$ and $\rho_v^{o(ref)}$, it is possible to obtain answers to the above questions. The analysis of the numerical data is made using the second-order Møller-Plesset results as the reference and the corresponding formal framework based on the extension of FDET for non-variational treatment of electron-electron correlation given in Ref. 5.

The chosen clusters were considered previously in our study of complexation-induced shifts in the excitation energies[7] which showed a remarkably good performance of the used FDET-based method despite the fact that the density of the isolated molecule(s) belonging to the environment was used as ρ_B for each electronic excited state. This could be the result of: *a*) fortuitous cancellations of errors due to the violation of the non-negativity condition in different electronic states, *b*) numerical insignificance of the violation of the non-negativity condition on the total energy, or *c*) both. For this reason, the present work focusses on one electronic state - the ground-state. The complexes selected for the present study display different strength of interaction, number of molecules in the environment, number of non-covalent interactions, and the electric charge of the environment: *i*) *cis*-7-hydroxyquinoline bound to two methanol molecules (7HQ-2MeOH), *ii*) uracil bound to five water molecules

(uracil-5H_2O), *iii*) 7-hydroxyquinoline bound to formate (7HQ-formate), and *iv*) pyridiniumyl benzimidazolide bound to two formic acid molecules (PyrBnz-2HCOOH). 7HQ-2MeOH and uracil-5H_2O are typical hydrogen bonded complexes involving neutral donor and acceptor molecules. 7HQ-formate and PyrBnz-2HCOOH) represent more peculiar cases. In 7HQ-formate, the environment acts as the hydrogen acceptor and is negatively charged. Moreover, the hydrogen is almost shared between the *cis*-7-hydroxyquinoline and the formate: the bond legth of the hydroxy group is 1.09 Å while the hydrogen bond length is 1.36 Å. In PyrBnz-2HCOOH, the embedded system (PyrBnz) is a hydrogen acceptor and the atom involved carries a significant negative charge.

II. METHODS

A. FDET and its extension for non-variational methods

For interpretation purposes, it is convenient to split the external potential $v(\mathbf{r})$ into components $v(\mathbf{r}) = v_A(\mathbf{r}) + v_B(\mathbf{r})$. The embedded wavefunction and the total energy obtained from FDET do not depend on such splitting. Once $v(\mathbf{r})$ is split, $v_A(\mathbf{r})$ defines an N_A-electron Hamiltonian ($\hat{H}_A$) whereas $v_B(\mathbf{r})$ defines a N_B-electron Hamiltonian ($\hat{H}_B$). Using the above notation, the FDET energy functional reads:

$$E_v^{FDET}[\Psi_A, \rho_B] = \langle \Psi_A | \hat{H}_A | \Psi_A \rangle + E_{v_B}^{HK}[\rho_B] + E_{v_A,v_B}^{elst,int}[\rho_A, \rho_B] + E_{xcT}^{nad}[\rho_A, \rho_B], \qquad (1)$$

where *i*) $\rho_A(\mathbf{r}) = \langle \Psi_A | \sum_{i=1}^{N_A} \delta(\mathbf{r}_i - \mathbf{r}) | \Psi_A \rangle$, *ii*) $E_{v_A,v_B}^{elst,int}[\rho_A, \rho_B]$ collects all classical electrostatic contributions to the interaction energy, and *iii*) $E_{xcT}^{nad}[\rho_A, \rho_B]$ - is the bi-functional of the non-additive kinetic and exchange-correlation energies. Its constrained-search definition is given in the original work[3] and also, for the sake of completeness - in the Suplementary Material.

The FDET energy functional satisfies the following relation:

$$\min_{\Psi_A \to N_A} E_v^{FDET}[\Psi_A, \rho_B] = E_v^{FDET}[\Psi_A^o, \rho_B] = E_v^{HK}[\rho_A^o + \rho_B] \geq E_v^o = E_v^{HK}[\rho_v^o], \qquad (2)$$

where $\rho_A^o(\mathbf{r}) = \langle \Psi_A^o | \sum_{i=1}^{N_A} \delta(\mathbf{r}_i - \mathbf{r}) | \Psi_A^o \rangle$, and where E_v^o and ρ_v^o are the ground-state energy and density for the N_{tot}-electron system defined by the potential v. The last equality in the above equation is due to the Hohenberg-Kohn theorems whereas the central one is the basic formula of FDET relating the FDET energy bi-functional to the universal Hohenberg-Kohn functional.

If the density $\rho_v^o - \rho_B$ is v-representable, ρ_A^o defined in Eq. 2 can be obtained from the Euler-Lagrange equation:

$$\left(\hat{H}_A + \hat{v}_{emb}^{FDET}[\rho_A^o, \rho_B; v_B]\right) \Psi_A^o = \lambda^o \Psi_A^o \tag{3}$$

where $v_{emb}^{FDET}[\rho_A, \rho_B; v_B]$ is the FDET embedding potential:

$$v_{emb}^{FDET}[\rho_A, \rho_B; v_B](\mathbf{r}) = v_B(\mathbf{r}) + \int \frac{\rho_B(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' + \frac{\delta E_{xct}^{nad}[\rho_A, \rho_B]}{\delta \rho_A(\mathbf{r})}. \tag{4}$$

The ρ_A-dependency of $v_{xct}^{nad}[\rho_A, \rho_B](\mathbf{r})$ leads to two particular features of Eq. 3: *i*) solving it in practice involves an iterative procedure leading to self-consistency between Φ_A and the embedding potential and *ii*) the Lagrange multiplier λ^o does not represent the energy.

For an embedded single determinant, the FDET energy functional includes also the $E_c[\rho_A]$ term (correlation functional) assuring that Eq. 2 is satisfied. The constrained search definition of $E_c[\rho_A]$ is given in the Supplementary Material. In the present work, we use, however, not the original formulation of FDET, in which the embedded wavefunction is obtained variationally, but the extension of FDET for non-variational methods[5]. This extension provides the expression for the total energy, that is consistent with the Hohenberg-Kohn theorems even though the embedded wavefunction has a single-determinant form AND the $E_c[\rho_A]$ component of the total energy functional is neglected. The fact that the embedded wavefunction Ψ_A in Eq. 3 has a single-determinant form is indicated by the used notation (Φ instead of Ψ). The potential

$$v'_A(\mathbf{r}) = v_A(\mathbf{r}) + v_{emb}^{FDET}[\rho'_A, \rho_B; v_B](\mathbf{r}), \tag{5}$$

where $\rho'_A(\mathbf{r}) = \langle \Phi'_A | \sum_{i=1}^{N_A} \delta(\mathbf{r}_i - \mathbf{r}) | \Phi'_A \rangle$, defines an auxiliary N_A-electron system. The use of a prime ($'$) in the notation (Φ'_A, ρ'_A, and v'_A) indicates that the embedding potential does not include the functional derivative of $E_c[\rho_A]$. To indicate that the optimal density differs in such case from the density ρ_A^o defined in Eq. 2, it is denoted with ρ'_A. The density ρ_A^o corresponds to the optimal and consistent with the embedding potential embedded wavefunction of the Full Configuration Interaction form (or alternatively to the optimal single determinant consistent with the embedding potential in which the correlation potential is exact). According to the recently derived formula relating the quantities available in methods treating the correlation energy of embedded electrons non-variationally to the Hohenberg-Kohn

energy functional (Eq. 38 in Ref. 5):

$$E_v^{HK}[\rho_A^o + \rho_B] = E_v^{FDET}[\Phi'_A, \rho_B] + E^c_{v'_A} - E_k[\Delta\rho^c_{v'_A}, \rho'_A, \rho_B] + O(\Delta\rho)^2, \tag{6}$$

where,

$$E_k[\Delta\rho^c_{v'_A}, \rho'_A, \rho_B] = \int \rho'_A(\mathbf{r}) \int \Delta\rho^c_{v'_A}(\mathbf{r}') f^{nad}_{xcT}[\rho'_A, \rho_B](\mathbf{r}, \mathbf{r}') d\mathbf{r}' \mathrm{d}\mathbf{r}, \tag{7}$$

$$\Delta\rho^c_{v'_A}(\mathbf{r}) = \rho^o_{v'_A}(\mathbf{r}) - \rho'_{v'_A}(\mathbf{r}) \tag{8}$$

$$f^{nad}_{xcT}[\rho_A, \rho_B](\mathbf{r}, \mathbf{r}') = \frac{\delta^2 E^{nad}_{xcT}[\rho_A, \rho_B]}{\delta\rho_A(\mathbf{r})\delta\rho_A(\mathbf{r}')}. \tag{9}$$

$\rho^o_A(\mathbf{r})$ in the left-hand side is the optimal correlated density defined in Eq. 2. The right-hand-side of Eq. 6 is used to approximate $E_v^{HK}[\rho_A^o + \rho_B]$. Any post-Hartree-Fock method, applied to the potential v'_A (cf. Eq. 5), yields the necessary terms: *i*) the optimal single determinant (Φ'_A), *ii*) the corresponding density (ρ'_A), *iii*) the correlation energy ($E^c_{v'_A}$), and *iv*) the change of the density due to correlation ($\Delta\rho^c_{v'_A}(\mathbf{r})$).

For $\Delta\rho$ representing the correlation-induced changes of the density in the exact and in the auxiliary system, $O(\Delta\rho)^2$ collects all second- and higher-order contributions to the energy. It is worthwhile to underline that, whereas the basic FDET equality (Eq. 2) is the equation for functionals, Eq. 6 is an equation for numbers (energy and its components). It holds only for a particular density ρ'_A. As far as the $E_{v_B}^{HK}[\rho_B]$ component of $E_v^{FDET}[\Psi_A, \rho_B]$ is concerned, its treatment depends on the method used to generate ρ_B and will be given below.

For a given $\rho_B(\mathbf{r})$, the FDET interaction energy is given by:

$$E_{int}^{FDET(\rho_B)} = E_v^{HK}[\rho_A^o + \rho_B] - E_{v_A}^{HK}[\rho_A^{isol}] - E_{v_B}^{HK}[\rho_B^{isol}], \tag{10}$$

where ρ_X^{isol} denotes the density of the isolated subsystem X.

Using for $E_v^{HK}[\rho_A^o + \rho_B]$ the right-hand side of Eq. 6 with neglected $O(\Delta\rho)^2$, results in:

$$E_{int}^{FDET(\rho_B)} = \langle\Phi'_A|\hat{H}_{v_A}|\Phi'_A\rangle + E^c_{v'_A} + E^{elst,int}_{v_A,v_B}[\rho'_A, \rho_B] + E^{nad}_{xcT}[\rho'_A, \rho_B] - E_k[\Delta\rho^c_{v'_A}, \rho'_A, \rho_B] - E_{v_A}^{HK}[\rho_A^{isol}] + E_{v_B}^{HK}[\rho_B] - E_{v_B}^{HK}[\rho_B^{isol}] \tag{11}$$

In the above equation, the quantities: Φ'_A, ρ'_A, $E^c_{v'_A}$, and $\Delta\rho^c_{v'_A}$ implicitly depend on ρ_B, through v'_A. For the sake of conciseness, this dependency is not indicated explicitly.

B. FDET interaction energy

The following sub-sections concerns application of Eq. 11 for different choices of ρ_B.

1. *FDET interaction energy for* $\rho_B = \rho_B^{isol}$

If the environment is modelled by means of its isolated density $\rho_B^{isol}(\mathbf{r})$, the last two terms in Eq. 11 are equal and thus cancel out leading to:

$$\begin{aligned} E_{int}^{FDET(\rho_B^{isol})} &= \langle \Phi'_A | \hat{H}_{v_A} | \Phi'_A \rangle + E^c_{v'_A} + E^{elst,int}_{v_A,v_B}[\rho'_A, \rho_B^{isol}] \\ &+ E^{nad}_{xcT}[\rho'_A, \rho_B^{isol}] - E_k[\Delta\rho^c_{v'_A}, \rho'_A, \rho_B^{isol}] - \langle \Phi_A | \hat{H}_{v_A} | \Phi_A \rangle - E^c_{v_A} \end{aligned} \tag{12}$$

where the exact relation $E^o_{v_A} = E^{HK}_{v_A}[\rho_A^{isol}] = \langle \Phi_A | \hat{H}_{v_A} | \Phi_A \rangle + E^c_{v_A}$ was used for the energy of the isolated subsystem A.

2. *FDET interaction energy for* $\rho_B^{v"}$ *being the ground-state Hartree-Fock density for some potential* $v"$

If the environment density is obtained as Hartree-Fock solution for an external potential $v''(\mathbf{r})$ other than the nuclear potential of subsystem B ($v_B(\mathbf{r})$), the numerical value of $E^{HK}_{v_B}[\rho_B^{v"}] - E^{HK}_{v_B}[\rho_B^{isol}] \geq 0$ contributes to the interaction energy. It is approximated as:

$$E^{HK}_{v_B}[\rho_B^{v"}] = \langle \Phi"_B | \hat{H}_{v_B} | \Phi"_B \rangle + E^c[\rho_B^{v"}] \approx \langle \Phi"_B | \hat{H}_{v_B} | \Phi"_B \rangle + E^c_{v"}, \tag{13}$$

where $\Phi"_B$ is the optimal determinant yielding $\rho_B^{v"}$.

Using the above approximation in Eq. 11 leads to:

$$\begin{aligned} E_{int}^{FDET(\rho_B^{v"})} &= \langle \Phi'_A | \hat{H}_{v_A} | \Phi'_A \rangle + E^c_{v'_A} + \langle \Phi"_B | \hat{H}_{v_B} | \Phi"_B \rangle + E^c_{v"} \\ &+ E^{elst,int}_{v_A,v_B}[\rho'_A, \rho_B^{v"}] + E^{nad}_{xcT}[\rho'_A, \rho_B^{v"}] - E_k[\Delta\rho^c_{v'_A}, \rho'_A, \rho_B^{v"}] \\ &- \langle \Phi_A | \hat{H}_{v_A} | \Phi_A \rangle - E^c_{v_A} - \langle \Phi_B | \hat{H}_{v_B} | \Phi_B \rangle - E^c_{v_B} \end{aligned} \tag{14}$$

where Φ_X is the Hartree-Fock wavefunction and $E^c_{v_X}$ denotes the correlation energy in the system defined by the potential v_X ($X = A$ or B).

3. *FDET interaction energy for optimised* ρ_B

Optimisation of both ρ_A and ρ_B proceeds by performing an iterative cycle (*freeze-and-thaw*, FAT) in which the subsystem A and B exchange their roles in all FDET equations[8]. In case of freezing ρ_A, ρ_B is represented be means of an embedded N_B electron single determinant (Φ'_B) that is optimal for a given ρ_A. Let us denote the quantities obtained at the end of such an optimisation with: v'^{FAT}_A, Φ'^{FAT}_A, ρ'^{FAT}_A, $E^c_{v'^{FAT}_A}$, and $\Delta\rho^c_{v'^{FAT}_A}$ for the subsystem A and v'^{FAT}_B Φ'^{FAT}_B, ρ'^{FAT}_B, $E^c_{v'^{FAT}_B}$, and $\Delta\rho^c_{v'^{FAT}_A}$ for the subsystem B. Using this notation, Eq. 14 reads:

$$\begin{aligned} E_{int}^{FDET(\rho'^{FAT}_B)} &= \langle\Phi'^{FAT}_A|\hat{H}_{v_A}|\Phi'^{FAT}_A\rangle + E^c_{v'^{FAT}_A} + \langle\Phi'^{FAT}_B|\hat{H}_{v_B}|\Phi'^{FAT}_B\rangle + E^c_{v'^{FAT}_B} \\ &+ E^{elst,int}_{v_A,v_B}[\rho'^{FAT}_A,\rho'^{FAT}_B] + E^{nad}_{xcT}[\rho'^{FAT}_A,\rho'^{FAT}_B] - E_k[\Delta\rho^c_{v'^{FAT}_A},\rho'^{FAT}_A,\rho^{FAT}_B] \\ &- \langle\Phi_A|\hat{H}_{v_A}|\Phi_A\rangle - E^c_{v_A} - \langle\Phi_B|\hat{H}_{v_B}|\Phi_B\rangle - E^c_{v_B} \end{aligned} \tag{15}$$

If the subsystems A and B exchange their roles in FDET equations (subsequent *freeze-and-thaw* iteration), the corresponding expression reads:

$$\begin{aligned} E_{int}^{FDET(\rho'^{FAT}_A)} &= \langle\Phi'^{FAT}_B|\hat{H}_{v_B}|\Phi'^{FAT}_B\rangle + E^c_{v'^{FAT}_B} + \langle\Phi'^{FAT}_A|\hat{H}_{v_A}|\Phi'^{FAT}_A\rangle + E^c_{v'^{FAT}_A} \\ &+ E^{elst,int}_{v_B,v_A}[\rho'^{FAT}_B,\rho'^{FAT}_A] + E^{nad}_{xcT}[\rho'^{FAT}_B,\rho'^{FAT}_A] - E_k[\Delta\rho^c_{v'^{FAT}_B},\rho'^{FAT}_B,\rho^{FAT}_A] \\ &- \langle\Phi_B|\hat{H}_{v_B}|\Phi_B\rangle - E^c_{v_B} - \langle\Phi_A|\hat{H}_{v_A}|\Phi_A\rangle - E^c_{v_A} \end{aligned} \tag{16}$$

Following Eq. 6, both above expressions yield $E^{HK}_v[\rho'^{FAT}_A + \rho'^{FAT}_B]$ up to the second order terms ($O(\Delta\rho)^2$). Adding Eqs. 15 and 16 and dividing the sum by two yields an expression for the interaction energy that is symmetric upon exchange A and B.

$$\begin{aligned} E_{int}^{FDET(FAT)} &= \langle\Phi'^{FAT}_B|\hat{H}_{v_B}|\Phi'^{FAT}_B\rangle + E^c_{v'^{FAT}_B} + \langle\Phi'^{FAT}_A|\hat{H}_{v_A}|\Phi'^{FAT}_A\rangle + E^c_{v'^{FAT}_A} \\ &+ E^{elst,int}_{v_B,v_A}[\rho'^{FAT}_B,\rho'^{FAT}_A] + E^{nad}_{xcT}[\rho'^{FAT}_B,\rho'^{FAT}_A] \\ &- \frac{1}{2}\left(E_k[\Delta\rho^c_{v'^{FAT}_A},\rho'^{FAT}_B,\rho^{FAT}_A] + E_k[\Delta\rho^c_{v'^{FAT}_B},\rho'^{FAT}_A,\rho^{FAT}_B]\right) \\ &- \langle\Phi_B|\hat{H}_{v_B}|\Phi_B\rangle - E^c_{v_B} - \langle\Phi_A|\hat{H}_{v_A}|\Phi_A\rangle - E^c_{v_A} \end{aligned} \tag{17}$$

For exact $E^{nad}_{xcT}[\rho_A,\rho_B]$ and exact E^c_v, all three equations Eqs. 15, 16, and 17 yield the same energy. In practical calculations, Eqs. 15 and 16 may yield different numerical results due to the following factors: *i*) the approximation: $E^{nad}_{xcT}[\rho_A,\rho_B] \approx \tilde{E}^{nad}_{xcT}[\rho_A,\rho_B]$, *ii*) approximate treatment of the correlation energy E^c_v, *iii*) the incompleteness of the used basis sets, *iv*) the difference in magnitude of the second order contributions $O(\Delta\rho)^2$(cf. Eq. 6). The

symmetrised expression for the interaction energy (Eq. 17) uniquely defines the interaction energy for practical calculations.

C. Measures of the density errors

As a measure of violation of the non-negativity condition by a given density $\rho_B(\mathbf{r})$, the parameter $M[\rho_B(\mathbf{r}) - \rho_v^{o(ref)}(\mathbf{r})]$ defined as:

$$\begin{aligned} f(\mathbf{r}) &= \rho_B(\mathbf{r}) - \rho_v^{o(ref)}(\mathbf{r}) \\ M[f] &= \int f(\mathbf{r}) \cdot \Theta(f)\ \mathrm{d}\mathbf{r}, \end{aligned} \tag{18}$$

where Θ is the Heaviside step function, is used.

M is bound as $0 \leq M \leq N_B$. At the lower bound ($M = 0$), the inequality in Eq. 2 is reached (subject of the condition of N_A-representability of $\rho_v^{o(ref)} - \rho_B$). If, additionally, $\rho_v^{o(ref)} - \rho_B$ is v-representable, the exact solution of Eq. 3 yields this density. At the upper bound ($M = N_B$), the densities $\rho_B(\mathbf{r})$ and $\rho_v^{o(ref)}(\mathbf{r})$ do not overlap at any volume element.

The parameter $P[\rho_A^{Eq.3} + \rho_B - \rho_v^{o(ref)}]$ is used as a measure of the total density obtained from Eq. 3 for a given ρ_B. It is defined as:

$$\begin{aligned} g(\mathbf{r}) &= \rho_A^{Eq.3}(\mathbf{r}) + \rho_B(\mathbf{r}) - \rho_v^{o(ref)}(\mathbf{r}) \\ P[g] &= \frac{1}{2} \int |g(\mathbf{r})| d\mathbf{r}, \end{aligned} \tag{19}$$

The factor $\frac{1}{2}$ in the definition of $P[f]$ results in the the following relation between M and P (see the Supplementary Material):

$$M[\rho_B - \rho_v^{o(ref)}] \leq P[\rho_A^{Eq.3} + \rho_B - \rho_v^{o(ref)}] \leq N_{tot}. \tag{20}$$

The upper bound corresponds to non-overlapping $\rho_A^{Eq.3}(\mathbf{r}) + \rho_B(\mathbf{r})$ and $\rho_v^{o(ref)}(\mathbf{r})$, i.e. the situation not to be encountered in practice.

In order to discuss the significance of various approximations on the density obtained from FDET, the magnitude of the total interaction induced change of the electron-density defined as

$$P_{cmpl} = P[\rho_A^{isol} + \rho_B^{isol} - \rho_v^{o(ref)}], \tag{21}$$

is given for each considered system.

Since the upper bounds of the parameters M and P depend on the system, these parameters are always compared to P_{cmpl} in *the same* system.

D. Computational Details

The interaction energy was evaluated according to Eqs. 12, 14, and 17 applicable for the corresponding choices for $\rho_B(\mathbf{r})$ using the following approximations:

$$E^c_{v'_X} \approx E^{(2)}_{v'_X} \tag{22}$$

$$\Delta\rho^c_{v'_X}(\mathbf{r}) \approx \rho^{MP1}_{v'_X}(\mathbf{r}) - \rho'_X,(\mathbf{r}) \tag{23}$$

where $X = A$ or B, and $E^{(2)}_{v'_X}$ denotes the second-order Møller-Plesset energy correction.

For the sake of brevity, the methods based on the energy expressions given in this section and the approximations used in this work are referred to as FDET-MP2.

All ρ_B densities were obtained from single determinants. The parameters M, P, and P_{cmpl}, are evaluated taking the Hartree-Fock density as $\rho_v^{o(ref)}$ in Eqs. 18 and 19. The reference interaction energy E^{ref}_{int} was calculated with MP2 theory with counterpoise correction.

$E^{nad}_{xcT}[\rho_A, \rho_B]$ and its functional derivative were approximated using local-density approximation (LDA) for all its components: Thomas-Fermi[9,10] for the kinetic energy, Dirac-Slater[11] for the exchange energy, and the parametrisation taken from Ref. 12 of the correlation energy in the uniform electron gas[13]:

$$E^{nad}_{xcT}[\rho_A, \rho_B] \approx \tilde{E}^{nad(LDA)}_{xcT}[\rho_A, \rho_B] \tag{24}$$

$$v^{nad}_{xct}[\rho_A, \rho_B](\mathbf{r}) \approx \frac{\delta\tilde{E}^{nad(LDA)}_{xct}[\rho_A, \rho_B]}{\delta\rho_A(\mathbf{r})} \tag{25}$$

The used approximation for $E^{nad}_{xcT}[\rho_A, \rho_B]$ neglects the correlation component of the FDET embedding potential needed for embedded single determinants[3]. The relation given in Eq. 6 applies therefore. For a given ρ_B, the self-consistent quantities needed in the right-hand side of Eq. 6: optimal embedded single determinant (Φ'_A) and the potential defining the auxiliary system (v'), were obtained by means of an iterative procedure involving repetitive solution of Eq. 3. Maximum five to six iterations are needed to converge the total energy within 10^{-9} Hartree. The *freeze-and-thaw* optimisation of ρ_B required 7 to 14 iterations to converge to the same threshold, where one FAT iteration consists of two density optimisations by means of Eq. 3: one where the density of one subsystem is optimised and the environment is frozen, and the subsequent one, where the roles are switched. Note that the evaluation of the correlation energy is needed only after obtaining the final (consistent with the embedding potential) embedded single determinants. The above iterative procedures

were performed using the author's version[14] of CCParser[15] and CCDatabase[16] handling the automatic submission of Q-Chem5.4[17] calculations, parsing, and collecting the results.

The numerical results discussed in this work are obtained using *aug-cc-pVDZ* atomic basis sets. For the sake of comparisons, the discussed quantities have been also obtained using other basis sets (*cc-pVDZ* and *cc-pVTZ*) and they are available in the Supplementary Information. In FDET, two types of expansions were applied: *supermolecular expansion*, in which atomic functions localised on all atoms of the system were used, or *monomer expansion*, in which the Φ_A was constructed using only atomic functions centred on atoms defining the potential v_A whereas Φ_B was constructed using only atomic functions centred on atoms defining the potential v_B.

The geometries of the investigated complexes are reported in Ref. 18. Throughout this work, the following convention is used for relating Eq. 3 and to the names of the complexes/clusters: if the system name is $AAA-BBB$, the first part (AAA) is associated with Φ_A and v_A whereas BBB with ρ_B and v_B. In case of optimisation of ρ_B, Eq. 3 is solved iteratively for both subsystems. In subsequent calculations the indices A and B exchange and the order does not matter. The *freeze-and-thaw* optimisation always starts and ends with the index B being attributed to subsystem BBB.

The parameters M and P were evaluated using the grid integration implemented in PySCF[19], which is based on Becke[20]-Lebedev[21] grids. The plots were prepared using the Python modules: pandas[22] and matplotlib[23].

E. Explicit treatment of the electronic polarisation of ρ_B in FDET

The interpretation of the electronic polarisation of the environment and its effect on the energy is different in FDET and in the theories of intermolecular interactions such as the Symmetry-Adapted Perturbation Theory[24]. Whereas it enters as a separate component of the energy in the latter case, its identification in FDET is less straightforward. This is due to the fact that in FDET, the electronic polarisation of each subsystem is not well-defined because it is not unique[25,26]. If the non-negativity condition for $\rho_v^{o(ref)} - \rho_B$ is satisfied, any variation of ρ_B is not changing the FDET energy. In FDET-based methods, on the other hand, $E_{xcT}^{nad}[\rho_A, \rho_B] \approx \tilde{E}_{xcT}^{nad}[\rho_A, \rho_B]$. Even though $\tilde{E}_{xcT}^{nad}[\rho_A, \rho_B]$ is symmetric with respect to exchanging A with B, the errors in the potentials $\tilde{v}_{xcT}^{nad}[\rho_A, \rho_B](\mathbf{r})$ and $\tilde{v}_{xcT}^{nad}[\rho_B, \rho_A](\mathbf{r})$ are not

the same. As a result, the unique pair ρ_A and ρ_B is usually obtained in the *freeze-and-thaw* calculations.[27,28] Optimisation of ρ_B corresponds thus to two effects which cannot be separated, the possibility to reduce the magnitude of the violation of the non-negativity condition, and lowering the total energy given by an approximated expression differing from the exact one by $\tilde{E}_{xcT}^{nad}[\rho_A, \rho_B] - E_{xcT}^{nad}[\rho_A, \rho_B]$. None of them can be attributed to the electronic polarisation of ρ_B only.

The FDET calculations for $\rho_B = \rho_B^{isol}$, on the other hand, take into account the electronic polarisation implicitly if the two densities $\rho_A(\mathbf{r})$ and $\rho_B(\mathbf{r})$ do overlap. Even if ρ_B is frozen, the total density near the subsystem B can be modified by the intermolecular interactions due to its ρ_A component. Since $\rho_A(\mathbf{r})$ cannot be negative, the implicit treatment of the polarisation cannot be expected to be complete. The effect of such treatment on the energy depends on the basis set used which in turn determines the amount of possible overlap between $\rho_A(\mathbf{r})$ and $\rho_B(\mathbf{r})$. Moreover, the optimised density ρ_A depends critically on the used $\tilde{v}_{xcT}^{nad}[\rho_A, \rho_B](\mathbf{r})$.

In view of the above observations, it impossible to attribute the differences between the results obtained using $\rho_B = \rho_B^{FAT}$ and $\rho_B = \rho_B^{isol}$ exclusively to the electronic polarisation of the environment. To relate these differences to the electronic polarisation of the environment an intermediate technique to generate ρ_B was used. The effect of intermolecular interactions on ρ_B was taken into account explicitly by means of polarising it by the electric field generated the isolated subsystem A. The field was approximated using the net-atomic charges corresponding to the isolated subsystem A. We refer to ρ_B obtained in this way as a pre-polarised density to reflect the fact that the polarising field is generated by isolated subsystem A. The used net atomic charges are fitted to the electric potential generated by ρ_A^{isol} are obtained using the ChelPG method[29]. The pre-polarised density obtained in this way is denoted with $\rho_B^{pp(ChelPG)}$. The electric field generated by the ChelPG charges has the same long-distance limit behaviour as the FDET embedding potential. For the sake of comparison, the charges derived from the Mulliken population analysis[30] were also used. The corresponding densities are denoted with $\rho_B^{pp(MC)}$. For pre-polarised ρ_B, the interaction energy is given by Eq. 14.

The comparative analysis of the three approaches to the electronic polarisation within FDET are made only for the *monomer expansion*. The pre-polarisation by net atomic charges misses the quantum effects due to the Fermi statistics of electrons, which leads to

unstable numerical results if the basis set is such that the diffusion of electrons towards the environment is possible.[31,32]

III. RESULTS AND DISCUSSION

A. FDET results with optimised ρ_B

The FDET interaction energies obtained using the optimised ρ_B and the complete set of atomic basis sets (*supermolecular expansion*) collected in Table I compare well with the reference interaction energies. The largest deviation from the reference energy occurs for 7HQ-2MeOH (3.20 kcal/mol representing the relative error of about 18% relative error). For the remaining three complexes, the absolute errors are smaller and the relative errors do not exceed 9%. Such good performance of the used approximation for $E_{xcT}^{nad}[\rho_A, \rho_B]$ could be expected based on our previous numerical experience[28,33–35]. The numerical values of $\delta E_{int}^{FAT} = E_{int}^{FDET(FAT)} - E_{int}^{ref}$ can be attributed mainly to the used approximation for $E_{xcT}^{nad}[\rho_A, \rho_B]$. ρ_B was optimised and the same basis set was used in FDET-MP2 and in the reference supermolecular calculations (MP2). Other factors contributing to the deviations from the reference energies are due to incompleteness of the basis sets and to the higher than second-order contributions to the correlation energy. These effects might not completely cancel each other in FDET and in the reference calculations.

The non-zero values of the parameters M and P, on the other hand, are mainly due to the approximation used for $v_{xcT}^{nad}[\rho_A, \rho_B]$. The smaller they are, the better is $\tilde{v}_{xcT}^{nad}[\rho_A, \rho_B](\mathbf{r})$. The numerical values of M, P , and δE_{int} given in Table I are considered in the subsequent discussion as the reference residual due to $v_{xcT}^{nad}[\rho_A, \rho_B](\mathbf{r}) \approx \tilde{v}_{xcT}^{nad}[\rho_A, \rho_B](\mathbf{r})$ and will be compared to the values obtained upon adding additional approximations concerning ρ_B.

Some analyses in the present work are made using a reduced set of atomic basis sets (*monomer expansion*). This constitutes an additional contribution to the deviations of FDET interaction energies from the reference ones, the extent of which can be assessed by comparison of Table I and Table II.

Limiting the basis set expansion to functions centred on one subsystem increases the interaction energies reflecting thus the variational principle (the total energy decreases). Except for uracil-3H_2O, the energies obtained with basis sets centred on all atoms of the

complex	$P^{[a]}$	$P^{[b]}_{cmpl}$	$M^{[c]}$	δE_{int} [d]	$E^{FDET(FAT)}_{int}$ [e]	E^{ref}_{int} [f]
7HQ-2MeOH	0.059	0.362	0.007	3.20	-14.27	-17.47
7HQ-formate	0.066	0.670	0.007	3.16	-33.33	-36.48
uracil-5H_2O	0.114	0.633	0.014	-1.07	-39.69	-38.62
PyrBnz-2HCOOH	0.104	0.600	0.013	1.64	-34.89	-36.53

[a] $P = P[\rho_v^{o(ref)} - \rho_A^{FDET(FAT)} - \rho_B^{FDET(FAT)}]$ with $P[\rho]$ defined in Eq. 19

[b] $P_{cmpl} = P[\rho_A^{isol} + \rho_B^{isol} - \rho_v^{o(ref)}]$ (cf. Eq. 21), with $P[\rho]$ defined in Eq. 19

[c] $M = M[\rho_v^{o(ref)} - \rho_B^{FDET(FAT)}]$ with $M[\rho]$ defined in Eq. 18

[d] $\delta E_{int} = E_{int}^{FDET(FAT)} - E_{int}^{ref}$

[e] $E_{int}^{FDET(FAT)}$ is defined in Eq. 17

[f] The reference MP2 interaction energies (E_{int}^{ref}) are counterpoise corrected.

TABLE I. Deviations of the FDET-MP2 results from the reference data. In FDET, *freeze-and-thaw* optimised ρ_B and the reduced set of atomic basis sets (*supermolecular expansion*) are used. Density measures M and P are given in atomic units, energies in kcal/mol.

complex are closer to the MP2 reference. In the *monomer expansion* case, the total density obtained as a sum of the two subsystem densities does not include products of basis functions localised in different subsystems. The largest effect on energy due to neglecting of such terms occurs for 7HQ-formate case (the error increases from 3.16 to 8.46 kcal/mol) - arguably the most covalently bound complex among the four considered in this work, as evidenced by its hydrogen bond length (cf. Section I). In all cases, the magnitude of the violation of the non-negativity condition (M) is smaller if the *supermolecular expansion* is used.

In practice, for the multi-scale numerical simulations, optimising ρ_B by means of the *freeze-and-thaw* iterations is rather not realistic. FDET-based simulations hinge on the possibility to generate ρ_B in less costly than optimising the embedded wavefunction. The following sections concern such applications of FDET. To this end, the the deviations from the reference data on energy and density will be analysed in presence of additional assumptions concerning ρ_B obtained in alternative ways.

complex	$P^{[a]}$	$P^{[b]}_{cmpl}$	$M^{[c]}$	δE_{int} [d]	$E^{FDET(FAT)}_{int}$ [e]	E^{ref}_{int} [f]
7HQ-2MeOH	0.073	0.368	0.013	4.29	-13.18	-17.47
7HQ-formate	0.114	0.673	0.036	8.46	-28.03	-36.48
uracil-5H_2O	0.129	0.643	0.024	1.44	-37.18	-38.62
PyrBnz-2HCOOH	0.127	0.606	0.016	4.11	-32.42	-36.53

[a] $P = P[\rho_v^{o(ref)} - \rho_A^{FDET(FAT)} - \rho_B^{FDET(FAT)}]$ with $P[\rho]$ defined in Eq. 19

[b] $P_{cmpl} = P[\rho_A^{isol} + \rho_B^{isol} - \rho_v^{o(ref)}]$ (cf. Eq. 21), with $P[\rho]$ defined in Eq. 19

[c] $M = M[\rho_v^{o(ref)} - \rho_B^{FDET(FAT)}]$ with $M[\rho]$ defined in Eq. 18

[d] $\delta E_{int} = E_{int}^{FDET(FAT)} - E_{int}^{ref}$

[e] $E_{int}^{FDET(FAT)}$ is defined in Eq. 17

[f] The reference MP2 interaction energies (E_{int}^{ref}) are counterpoise corrected.

TABLE II. Deviations of the FDET-MP2 results from the reference data. In FDET, *freeze-and-thaw* optimised ρ_B and the reduced set of atomic basis sets (*monomer expansion*) are used. Density measures M and P are given in atomic units, energies in kcal/mol.

B. FDET results without optimisation of ρ_B: $\rho_B = \rho_B^{isol}$

We start with the observation that the basis set restriction - i.e. using the *monomer* rather than the *supermolecular expansion* - has a much smaller effect on both densities and energies than the use of a non-optimised density - ρ_B^{isol} rather than ρ_B^{FAT} - as evidenced by the comparison of the values in Tables I, II, and III. The errors in energy (δE_{int}) and in density (P) increase by a factor of about 2 to 3, whereas the parameter M, measuring the extent of the violation of the non-negativity of the target density associated to the specific ρ_B increases one order of magnitude more (see Tables I and Tables III).

For any choice for ρ_B, the ratio $r[\rho_B]$ defined as:

$$r[\rho_B] = \frac{P[\rho_v^{o(ref)} - \rho_A^{FDET(\rho_B)} - \rho_B]}{P[\rho_v^{o(ref)} - \rho_A^{FDET(FAT)} - \rho_B^{FDET(FAT)}]} \tag{26}$$

provides a quantitive measure of the relative significance of two factors affecting the total density obtained in FDET: a) the approximation $v_{xcT}^{nad}[\rho_A, \rho_B] \approx \tilde{v}_{xcT}^{nad}[\rho_A, \rho_B]$ and b) the arbitrary choice of the procedure to generate ρ_B that might lead to such ρ_B that $\rho_v^{o(ref)} - \rho_B$

complex	$P^{[a]}$	$P_{cmpl}^{[b]}$	$M^{[c]}$	δE_{int} [d]	$E_{int}^{FDET(\rho_B=\rho_B^{isol})}$	[e] E_{int}^{ref} [f]
7HQ-2MeOH	0.228	0.362	0.121	6.49	-10.98	-17.47
7HQ-formate	0.316	0.670	0.206	13.45	-23.04	-36.48
uracil-5H_2O	0.426	0.633	0.234	6.56	-32.06	-38.62
PyrBnz-2HCOOH	0.416	0.600	0.184	9.58	-26.95	-36.53

[a] $P = P[\rho_v^{o(ref)} - \rho_A^{FDET(\rho_B=\rho_B^{isol})} - \rho_B^{isol}]$ with $P[\rho]$ defined in Eq. 19

[b] $P_{cmpl} = P[\rho_A^{isol} + \rho_B^{isol} - \rho_v^{o(ref)}]$ (cf. Eq. 21), with $P[\rho]$ defined in Eq. 19

[c] $M = M[\rho_v^{o(ref)} - \rho_B^{isol}]$ with $M[\rho]$ defined in Eq. 18

[d] $\delta E_{int} = E_{int}^{FDET(\rho_B=\rho_B^{isol})} - E_{int}^{ref}$

[e] $E_{int}^{FDET(\rho_B=\rho_B^{isol})}$ is defined in Eq. 12

[f] The reference MP2 interaction energies (E_{int}^{ref}) are counterpoise corrected.

TABLE III. Deviations of the FDET-MP2 results from the reference data. In FDET, $\rho_B = \rho_B^{isol}$ and the complete set of atomic basis sets (*supermolecular expansion*) are used. Density measures M and P are given in atomic units, energies in kcal/mol.

violates the non-negativity condition. If the *supermolecular expansion* is used, the denominator determines the error in the density due to the first factor whereas the numerator arises from the combination of the two of them. In all considered cases, $r[\rho_B^{isol}]$ is at least 3.

This observations indicate clearly that the errors of the FDET results originate mainly due to the violation of the non-negativity condition rather than due to the choice of the used approximation for $E_{xcT}^{nad}[\rho_A, \rho_B]$ and $v_{xcT}^{nad}[\rho_A, \rho_B](\mathbf{r})$.

We note also that despite a significant increase of the magnitude of the violation of the non-negativity condition in the $\rho_B = \rho_B^{isol}$ case (by at least one order of magnitude), the errors of global properties of the system such as the total density (measured by the parameter P) and the interaction energy (measured by δE_{int}) are affected less. This is probably due to the variational character of densities obtained in FDET. The error in energy due to this violation by the chosen ρ_B is compensated by means of the variationally obtained ρ_A.

Similarly to what observed for densities, comparing the FDET energies obtained using $\rho_B = \rho_B^{isol}$ to those obtained using $\rho_B = \rho_B^{FAT}$ shows clearly that the choice $\rho_B = \rho_B^{isol}$

complex	$P^{[a]}$	$P_{cmpl}^{[b]}$	$M^{[c]}$	δE_{int} [d]	$E_{int}^{FDET(\rho_B=\rho_B^{isol})}$	[e] E_{int}^{ref} [f]
7HQ-2MeOH	0.231	0.368	0.123	6.77	-10.70	-17.47
7HQ-formate	0.316	0.673	0.205	13.56	-22.93	-36.48
uracil-5H_2O	0.427	0.643	0.237	6.99	-31.63	-38.62
PyrBnz-2HCOOH	0.419	0.606	0.185	10.74	-25.80	-36.53

[a] $P = P[\rho_v^{o(ref)} - \rho_A^{FDET(\rho_B=\rho_B^{isol})} - \rho_B^{isol}]$ with $P[\rho]$ defined in Eq. 19

[b] $P_{cmpl} = P[\rho_A^{isol} + \rho_B^{isol} - \rho_v^{o(ref)}]$ (cf. Eq. 21), with $P[\rho]$ defined in Eq. 19

[c] $M = M[\rho_v^{o(ref)} - \rho_B^{isol}]$ with $M[\rho]$ defined in Eq. 18

[d] $\delta E_{int} = E_{int}^{FDET(\rho_B=\rho_B^{isol})} - E_{int}^{ref}$

[e] $E_{int}^{FDET(\rho_B=\rho_B^{isol})}$ is defined in Eq. 12

[f] The reference MP2 interaction energies (E_{int}^{ref}) are counterpoise corrected.

TABLE IV. Deviations of the FDET-MP2 results from the reference data. In FDET, $\rho_B = \rho_B^{isol}$ and the reduced set of atomic basis sets (*monomer expansion*) are used. Density measures M and P are given in atomic units, energies in kcal/mol.

introduces additional errors in energies. In the FDET terms, the increase of the error is due to a stronger violation of the non-additivity condition for $\rho_v^{o(ref)} - \rho_B$ in the former case.

The analyses presented in the present work were inspired by our previous studies concerning the FDET-based simulations of complexation-induced shifts of the excitation energies in intermolecular complexes.

As shown in Ref. 7, the $\rho_B = \rho_B^{isol}$ approximation in FDET leads to an astonishingly good overall accuracy of the complexation-induced shifts of the vertical valence excitation energies (errors in the order 0.04 eV). This magnitude of the error in energy is much smaller than the errors reported in Tables III and IV obtained using the same choice for ρ_B. According to the present analysis, such a choice for ρ_B leads to an error in the ground state energy of at least 6.5 kcal/mol (equivalent to 0.26 eV). The present analysis shows that the violation of the non-additivity condition is a major source of error in the ground-state energy if $\rho_B = \rho_B^{isol}$. Since this error is non-negative regardless of the system (see Eq. 2), the much better performance of the $\rho_B = \rho_B^{isol}$ approximation for excitation energies[7], than for the interaction energies

has its origin in a systematic compensation of these non-negative contributions to the total energy in two electronic states. This systematic compensation was observed not only for the lowest excitations but also for the higher ones.

C. FDET results with pre-polarized ρ_B

The fact that optimisation of ρ_B not only reduces the errors in the FDET energies but also leads to lowering of the magnitude of violation of the non-negativity condition indicates the link between electronic polarisation of ρ_B. The subsequent part concerns this link and aims at more efficient ways to generate ρ_B than optimising by means of the *freeze-and-thaw* iterations. We start with pointing out that using the *monomer expansion* instead of the *supermolecular expansion* results in a much smaller effect on the energy than the optimisation of ρ_B in the investigated complexes and with the used $\tilde{E}^{nad}_{xcT}[\rho_A, \rho_B]$. Only the FDET obtained using the *monomer expansion* are, therefore, discussed in this section. Figure 1 shows the results discussed in the previous sections but also data obtained using other choices for ρ_B. Pre-polarisation of ρ_B by the field generated by the ChelPG charges corresponding to ρ_A^{isol} leads to lowering of the parameter M and brings it closer to the value of M for the optimised ρ_B. Lower values of M, i.e. less significant violation of the non-negativity condition, correspond also to improved energies. The FDET energies obtained using the pre-polarised ρ_B and optimised ρ_B agree within about 1 kcal/mol. Using the Mulliken charge representation of ρ_A^{isol} does not lead to similar - desired - effects. The advantage of using the ChelPG over the Mulliken charges in order to polarise ρ_B could be expected, as the former are fitted to the electrostatic potential generated by the molecule.

Figure 2 shows that the pre-polarisation of ρ_B using the ChelPG representation of ρ_A^{isol} results also in a significant improvement in the total density.

IV. CONCLUSIONS

The present work shows comprehensively that the error in the energy and the density due to the used semi-local approximations for $E^{nad}_{xcT}[\rho_A, \rho_B]$ is significantly smaller than the error in these quantities due to the use of the density of the isolated environment as ρ_B in FDET. This fact lies at the origin of a remarkable performance of FDET in predicting

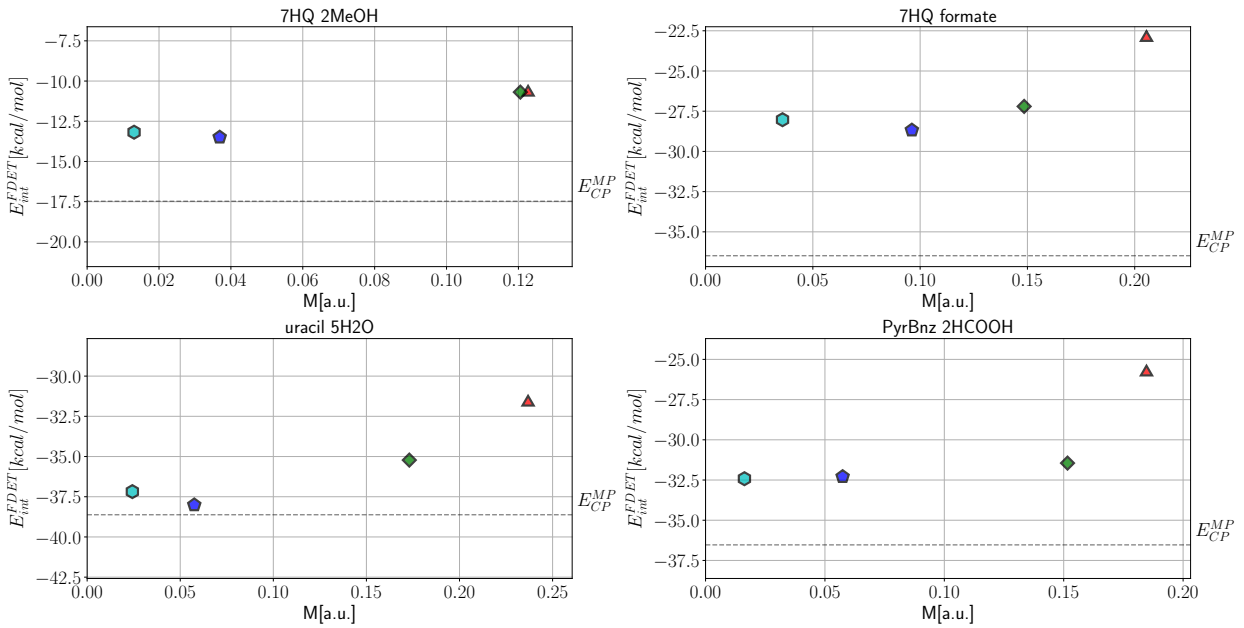

FIG. 1. Integrated negative density M and the FDET-MP2 interaction energy for various choices of ρ_B: a) ρ_B^{isol} (orange triangles), b) ρ_B^{FAT} (light blue hexagons), c) $\rho_B^{pp(Mulliken)}$ (green diamonds), and d) $\rho_B^{pp(ChelPG)}$ (dark blue pentagons). Data obtained using the *monomer expansion*. Horizontal lies indicate the reference interaction energy.

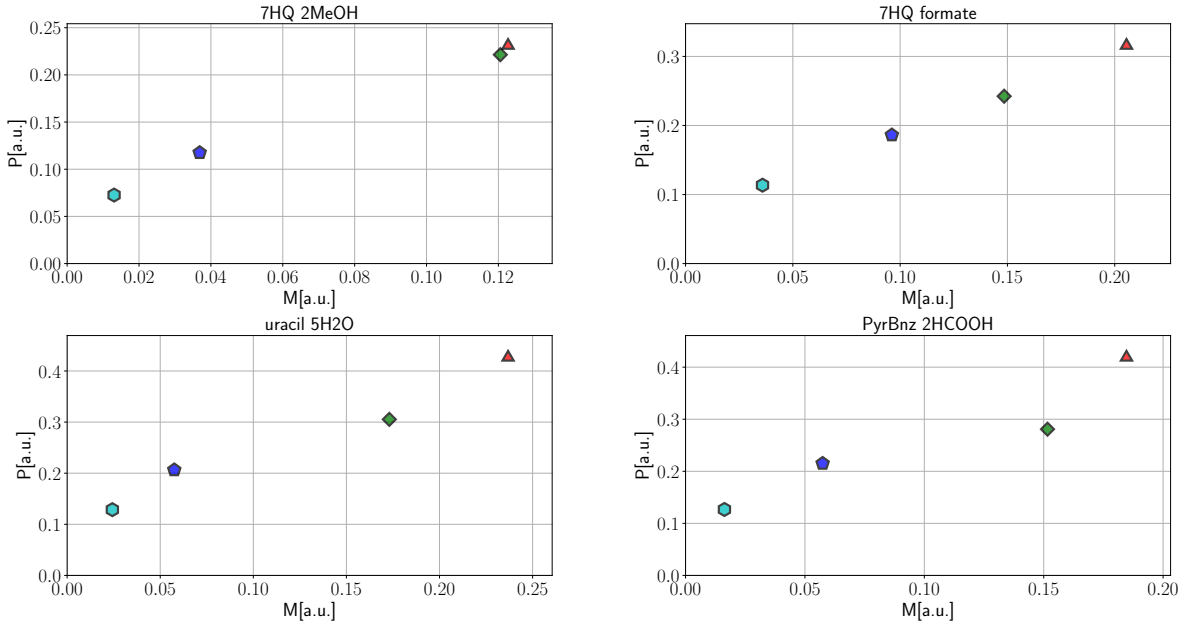

FIG. 2. Integrated negative density M and the total density error P for various choices of ρ_B: a) ρ_B^{isol} (orange triangles), b) ρ_B^{FAT} (light blue hexagons), c) $\rho_B^{pp(Mulliken)}$ (green diamonds), and d) $\rho_B^{pp(ChelPG)}$ (dark blue pentagons). Data obtained using the *monomer expansion*.

complexation-induced excitation energy shifts if the density of the isolated environment is used as ρ_B[7]. Whereas, the approximation to $E_{xcT}^{nad}[\rho_A, \rho_B]$ can result in either positive or negative energy error for a given electronic state, the energy error due to use of the isolated environment density is always non-negative (cf. Eq. 2). The origin of the good description of vertical excitation energies lies, therefore, in a systematic cancellation of these non-negative contributions for each state.

For the energy of one state, on the other hand, there is no other component that could compensate the error due to violation of this condition. In practical calculations, verification of the condition of the non-negativity of $\rho_v^{o(ref)} - \rho_B$ is not possible. This would require *a priori* knowledge of the total density. The analyses in the present work provided a link between the violation of the condition of the non-negativity of $\rho_v^{o(ref)} - \rho_B$ and the effect of electronic polarisation of ρ_B by the embedded species. Due to this link, the magnitude of the violation of the non-negativity condition of $\rho_v^{o(ref)} - \rho_B$ can be significantly reduced in practice.

SUPPLEMENTARY MATERIAL

The Supplementary Material includes, definitions of the density functionals used in FDET, the proof of Eq. 20, the counterpoise corrections for E_{int}^{ref}, the values of $E_k[\Delta\rho_{v'_X}^c, \rho'_X, \rho_Y]$, data obtained using other basis sets than the ones used in the main paper, and data plotted in the Figures. Data available in article or supplementary material.

REFERENCES

[1]T. A. Wesolowski, S. Shedge, and X. Zhou, Chemical Reviews **115**, 5891 (2015).

[2]T. A. Wesolowski and A. Warshel, The Journal of Physical Chemistry **97**, 8050 (1993).

[3]T. A. Wesołowski, Physical Review A **77**, 012504 (2008).

[4]K. Pernal and T. A. Wesolowski, International Journal of Quantum Chemistry **109**, 2520 (2009).

[5]T. A. Wesolowski, Journal of Chemical Theory and Computation **16**, 6880 (2020).

[6]P. Hohenberg and W. Kohn, Physical Review **136**, B864 (1964).

[7]N. Ricardi, A. Zech, Y. Gimbal-Zofka, and T. A. Wesolowski, Physical Chemistry Chemical Physics **20**, 26053 (2018).

[8]T. A. Wesolowski and J. Weber, Chemical Physics Letters **248**, 71 (1996).

[9]L. H. Thomas, Mathematical Proceedings of the Cambridge Philosophical Society **23**, 542 (1927).

[10]E. Fermi, Zeitschrift für Physik **48**, 73 (1928).

[11]J. C. Slater, Physical Review **34**, 1293 (1929).

[12]S. H. Vosko, L. Wilk, and M. Nusair, Canadian Journal of Physics **58**, 1200 (1980).

[13]D. M. Ceperley and B. J. Alder, Physical Review Letters **45**, 566 (1980).

[14]N. Ricardi, "Ccparser," `https://github.com/NicoRicardi/CCParser` (2021).

[15]A. Zech, "Ccparser," `https://github.com/spectre007/CCParser` (2021).

[16]N. Ricardi and C. E. González-Espinoza, "Ccdatabase," `https://github.com/NicoRicardi/CCDatabase` (2020).

[17]E. Epifanovsky, A. T. B. Gilbert, X. Feng, J. Lee, Y. Mao, N. Mardirossian, P. Pokhilko, A. F. White, M. P. Coons, A. L. Dempwolff, Z. Gan, D. Hait, P. R. Horn, L. D. Jacobson, I. Kaliman, J. Kussmann, A. W. Lange, K. U. Lao, D. S. Levine, J. Liu, S. C. McKenzie, A. F. Morrison, K. D. Nanda, F. Plasser, D. R. Rehn, M. L. Vidal, Z.-Q. You, Y. Zhu, B. Alam, B. J. Albrecht, A. Aldossary, E. Alguire, J. H. Andersen, V. Athavale, D. Barton, K. Begam, A. Behn, N. Bellonzi, Y. A. Bernard, E. J. Berquist, H. G. A. Burton, A. Carreras, K. Carter-Fenk, R. Chakraborty, A. D. Chien, K. D. Closser, V. Cofer-Shabica, S. Dasgupta, M. de Wergifosse, J. Deng, M. Diedenhofen, H. Do, S. Ehlert, P.-T. Fang, S. Fatehi, Q. Feng, T. Friedhoff, J. Gayvert, Q. Ge, G. Gidofalvi, M. Goldey, J. Gomes, C. E. González-Espinoza, S. Gulania, A. O. Gunina, M. W. D. Hanson-Heine, P. H. P. Harbach, A. Hauser, M. F. Herbst, M. Hernández Vera, M. Hodecker, Z. C. Holden, S. Houck, X. Huang, K. Hui, B. C. Huynh, M. Ivanov, A. Jász, H. Ji, H. Jiang, B. Kaduk, S. Kähler, K. Khistyaev, J. Kim, G. Kis, P. Klunzinger, Z. Koczor-Benda, J. H. Koh, D. Kosenkov, L. Koulias, T. Kowalczyk, C. M. Krauter, K. Kue, A. Kunitsa, T. Kus, I. Ladjánszki, A. Landau, K. V. Lawler, D. Lefrancois, S. Lehtola, R. R. Li, Y.-P. Li, J. Liang, M. Liebenthal, H.-H. Lin, Y.-S. Lin, F. Liu, K.-Y. Liu, M. Loipersberger, A. Luenser, A. Manjanath, P. Manohar, E. Mansoor, S. F. Manzer, S.-P. Mao, A. V. Marenich, T. Markovich, S. Mason, S. A. Maurer, P. F. McLaughlin, M. F. S. J. Menger, J.-M. Mewes, S. A. Mewes, P. Morgante, J. W. Mullinax, K. J. Oosterbaan, G. Paran, A. C. Paul, S. K.

Paul, F. Pavošević, Z. Pei, S. Prager, E. I. Proynov, A. Rák, E. Ramos-Cordoba, B. Rana, A. E. Rask, A. Rettig, R. M. Richard, F. Rob, E. Rossomme, T. Scheele, M. Scheurer, M. Schneider, N. Sergueev, S. M. Sharada, W. Skomorowski, D. W. Small, C. J. Stein, Y.-C. Su, E. J. Sundstrom, Z. Tao, J. Thirman, G. J. Tornai, T. Tsuchimochi, N. M. Tubman, S. P. Veccham, O. Vydrov, J. Wenzel, J. Witte, A. Yamada, K. Yao, S. Yeganeh, S. R. Yost, A. Zech, I. Y. Zhang, X. Zhang, Y. Zhang, D. Zuev, A. Aspuru-Guzik, A. T. Bell, N. A. Besley, K. B. Bravaya, B. R. Brooks, D. Casanova, J.-D. Chai, S. Coriani, C. J. Cramer, G. Cserey, A. E. DePrince, R. A. DiStasio, A. Dreuw, B. D. Dunietz, T. R. Furlani, W. A. Goddard, S. Hammes-Schiffer, T. Head-Gordon, W. J. Hehre, C.-P. Hsu, T.-C. Jagau, Y. Jung, A. Klamt, J. Kong, D. S. Lambrecht, W. Liang, N. J. Mayhall, C. W. McCurdy, J. B. Neaton, C. Ochsenfeld, J. A. Parkhill, R. Peverati, V. A. Rassolov, Y. Shao, L. V. Slipchenko, T. Stauch, R. P. Steele, J. E. Subotnik, A. J. W. Thom, A. Tkatchenko, D. G. Truhlar, T. Van Voorhis, T. A. Wesolowski, K. B. Whaley, H. L. Woodcock, P. M. Zimmerman, S. Faraji, P. M. W. Gill, M. Head-Gordon, J. M. Herbert, and A. I. Krylov, The Journal of Chemical Physics **155**, 084801 (2021), https://doi.org/10.1063/5.0055522.

[18]A. Zech, N. Ricardi, S. Prager, A. Dreuw, and T. A. Wesolowski, Journal of Chemical Theory and Computation **14**, 4028 (2018).

[19]Q. Sun, T. C. Berkelbach, N. S. Blunt, G. H. Booth, S. Guo, Z. Li, J. Liu, J. D. McClain, E. R. Sayfutyarova, S. Sharma, S. Wouters, and G. K. Chan, Wiley Interdisciplinary Reviews: Computational Molecular Science **8**, e1340 (2017), https://onlinelibrary.wiley.com/doi/pdf/10.1002/wcms.1340.

[20]A. D. Becke, The Journal of Chemical Physics **88**, 2547 (1988), https://doi.org/10.1063/1.454033.

[21]V. I. Lebedev and D. N. Laikov, Doklady Mathematics **59**, 477 (1999).

[22]Wes McKinney, in *Proceedings of the 9th Python in Science Conference*, edited by Stéfan van der Walt and Jarrod Millman (2010) pp. 56 – 61.

[23]J. D. Hunter, Computing in Science & Engineering **9**, 90 (2007).

[24]B. Jeziorski, R. Moszynski, and K. Szalewicz, Chemical Reviews **94**, 1887 (1994).

[25]A. Savin and T. A. Wesolowski, in *Advances in the Theory of Atomic and Molecular Systems: Conceptual and Computational Advances in Quantum Chemistry*, Progress in Theoretical Chemistry and Physics, Vol. 19, edited by P. Piecuch, J. Maruani, G. Delgado-Barrio, and S. Wilson (Springer Netherlands, 2009) pp. 311–326.

[26]M. Humbert-Droz, X. Zhou, S. V. Shedge, and T. A. Wesolowski, Theoretical Chemistry Accounts **133**, 1405 (2014).

[27]T. A. Wesolowski, The Journal of Chemical Physics **106**, 8516 (1997).

[28]M. Dułak and T. A. Wesołowski, Journal of Molecular Modeling **13**, 631 (2007).

[29]C. M. Breneman and K. B. Wiberg, Journal of Computational Chemistry **11**, 361 (1990).

[30]R. S. Mulliken, The Journal of Chemical Physics **23**, 1833 (1955).

[31]G. Fradelos and T. A. Wesołowski, The Journal of Physical Chemistry A **115**, 10018 (2011).

[32]G. Fradelos and T. A. Wesolowski, Journal of Chemical Theory and Computation **7**, 213 (2011).

[33]T. A. Wesolowski and F. Tran, The Journal of Chemical Physics **118**, 2072 (2003).

[34]R. Kevorkyants, M. Dulak, and T. A. Wesolowski, The Journal of Chemical Physics **124**, 024104 (2006).

[35]R. Sen, C. E. González-Espinoza, A. Zech, A. Dreuw, and T. A. Wesolowski, Journal of Chemical Theory and Computation **17**, 4049 (2021).