Negativity of the target density in practical Frozen-Density Embedding Theory based calculations

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: 8 April 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 bi-functional $E_{xcT}^{nad}[\rho_A, \rho_B]$ representing the non-additivity of density functionals for the exchange-, correlation-, and kinetic energies and *ii*) the possible violation of the non-negativity condition of the target density for a given density associated with the environment $\rho_B(\mathbf{r})$.

The relative significance of theses two factors is investigated for four representative weakly bound intermolecular clusters and various choices for $\rho_B(\mathbf{r})$. It is shown that the violation of the non-negativity condition of the target density is the principal source of error in the FDET energy if $\rho_B(\mathbf{r})$ corresponds to the isolated environment. 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 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. The self-consistent expressions for the functional for the total energy and the corresponding functional for embedding potential are available for various possible quantum descriptors of the embedded species[1–5]. The environment is described by means of the electron density ρ_B whereas the embedded species by means of the embedded N_A-electron wavefunction (Ψ_A). The total energy of a system is given by the functional $E_{v_{AB}}^{FDET}[\Psi_A, \rho_B]$ which is closely related to 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).

Any multi-level simulation based on the formal framework of FDET hinges on two types of approximations/assumptions. One concerns the used approximation for one of the components of $E_{v_{AB}}^{FDET}[\Psi_A, \rho_B]$ the other concerns the approach to generate ρ_B. In FDET, multi-level means that ρ_B is generated using other methods than those used to optimise Ψ_A. The errors due to these two types of approximations combine and their relative significance cannot be determined in a straightforward matter. The approximation used for $E_{v_{AB}}^{FDET}[\Psi_A, \rho_B]$ leads to errors in the electron density obtained and in errors in energy. The latter can be either positive and negative. The errors due to ρ_B are always non-negative, due to the second Hohenberg-Kohn theorem. The optimal electron density of the embedded species obtained in FDET, ρ_A^o, is always v-representable. As a result, if ρ_B is larger than the exact ground-state density of the whole system (ρ_{AB}^o) on some volume element, the sum of ρ_A and ρ_B cannot be equal to ρ_{AB}^o. Hence, FDET can only provide the upper bound of the exact ground-state anergy of the whole system. The necessary condition to reach the ground-state energy in FDET calculations is the non-negativity of the difference $\rho_{AB}^o - \rho_B$ on all non-zero volume elements. We will refer to this difference as the *target density*. The condition of non-negativity of $\rho_{AB}^o - \rho_B$ cannot be verified in practice because it requires the *a priori* knowledge of ρ_{AB}^o.

The principal question addressed in the present work is: *Does violation of the non-negativity condition matter in practice?* The issue has not been studied systematically in the literature. There are examples where ρ_B was produced in such a way to not violate the negativity condition. Wesolowski and Savin analysed the errors due to approximations

used for $E_{v_{AB}}^{FDET}[\Psi_A, \rho_B]$ in a model system for which the exact solutions of FDET are available. Only environment densities such that $\rho_{AB}^o - \rho_B \geq 0$ were considered.[7] Fux et al.[8] used localised Kohn-Sham orbitals from a supermolecular calculation, selecting only those localised on the environment. This led to environment densities whose violation of the non-negativity condition was small and only due to numerical effects. Nonetheless, the authors are not aware of any report of systematic analysis of the effect of sizeable violation extents.

The present work reports the result of a computational experiment on four intermolecular clusters, in which the relative significance of the two factors affecting the FDET results is estimated. In such an experiment the density and the energy of the whole cluster ρ_{AB}^o can be evaluated with an adequate method of quantum chemistry and used as a reference for various choices made for ρ_B. In particular, the relation between the violation of the condition of non-negativity of $\rho_{AB}^o - \rho_B$ and the electronic polarisation of the environment due to the interactions with the embedded species is investigated.

The chosen clusters were considered previously in our study of complexation induced shifts in the excitation energies[9] 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

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

$$E^{FDET}_{v_A,v_B}[\Psi_A,\rho_B] = \langle\Psi_A|\hat{H}_A|\Psi_A\rangle + E^{HK}_{v_B}[\rho_B] + E^{elst,int}_{v_A,v_B}[\rho_A,\rho_B] + E^{nad}_{xcT}[\rho_A,\rho_B], \tag{1}$$

where *i)* $\rho_A(\mathbf{r}) = \langle\Psi_A|\sum_{i=1}^{N_A}\delta(\mathbf{r}_i-\mathbf{r})|\Psi_A\rangle$, *ii)* $E^{elst,int}_{v_A,v_B}[\rho_A,\rho_B]$ collects all classical electrostatic contributions to the interaction energy, and *iii)* $E^{nad}_{xcT}[\rho_A,\rho_B]$ - the definition of this bifunctional depends on the form of the embedded wavefunction[2].

The FDET energy functional is defined to satisfy the following relation:

$$\min_{\Psi_A\to N_A} E^{FDET}_{v_A,v_B}[\Psi_A,\rho_B] = E^{FDET}_{v_A,v_B}[\Psi^o_A,\rho_B] = E^{HK}_{v_{AB}}[\rho^o_A+\rho_B] \geq E^o_{v_{AB}}, \tag{2}$$

where $\rho^o_A(\mathbf{r}) = \langle\Psi^o_A|\sum_{i=1}^{N_A}\delta(\mathbf{r}_i-\mathbf{r})|\Psi^o_A\rangle$, $\int\rho_B(\mathbf{r})d\mathbf{r} = N_B$, and $E^o_{v_{AB}}$ is the ground-state energy of the N_A+N_B-electron system defined by the potential $v_A(\mathbf{r})+v_B(\mathbf{r})$.

If the density $\rho^o_{AB}-\rho_B$ is v-representable, ρ^o_A defined in Eq. 2 can be obtained from the Euler-Lagrange equation:

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

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

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

The term $\frac{\delta E^{nad}_{xct}[\rho_A,\rho_B]}{\delta\rho_A(\mathbf{r})} \equiv v^{nad}_{xct}[\rho_A,\rho_B](\mathbf{r})$ depends on the form of the embedded wavefunction[2]. For a single determinant, the FDET expression for $E^{nad}_{xct}[\rho_A,\rho_B]$ includes also the correlation functional $E_c[\rho_A]$ assuring that the optimal embedded density satisfies Eq. 2. In the present work, $E^{nad}_{xct}[\rho_A,\rho_B]$ always denotes the bifunctional corresponding to the embedded

wavefunction of the Full Configuration Interaction form[2], i.e. $E_c[\rho_A]$ is not included. 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 involves an iterative procedure leading to self-consistency[10] and *ii*) the Lagrange multiplier λ^o does not represent the energy.[11]

We consider the case where the wavefunction Ψ_A in Eq. 3 has a single-determinant form, which we denote with the letter Φ. 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 the simultaneous use of a single determinant wavefunction and the neglect of $E_c[\rho_A]$ and its functional derivative. We bring the reader's attention that ρ'_A always denotes the density obtained from a self-consistent single determinant.

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):

$$\begin{aligned} E_{v_{AB}}^{HK}[\rho_A^o + \rho_B] = E_{v_A, v_B}^{FDET}[\Phi'_A, \rho_B] + E_{v'_A}^c \\ - E_k[\Delta\rho_{v'_A}^c, \rho'_A, \rho_B] + O(\Delta\rho)^2, \end{aligned} \tag{6}$$

where,

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

$$\Delta\rho_{v'_A}^c(\mathbf{r}) = \rho_{v'_A}^o(\mathbf{r}) - \rho'_{v'_A}(\mathbf{r}) \tag{8}$$

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

$\rho_A^o(\mathbf{r})$ 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 $E_{v_{AB}}^{HK}[\rho_A^o + \rho_B]$. 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 ($E_{v'_A}^c$), and *iv*) the change of the density due to correlation ($\Delta\rho_{v'_A}^c(\mathbf{r})$).

For $\Delta\rho$ representing the correlation-induced changes of energy in the exact and in the auxiliary system, $O(\Delta\rho)^2$ collects all second- and higher-order contributions to the energy. As far as the $E^{HK}_{v_B}[\rho_B]$ component of $E^{FDET}_{v_A,v_B}[\Psi_A,\rho_B]$ 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:

$$E^{FDET(\rho_B)}_{int} = E^{HK}_{v_{AB}}[\rho^o_A+\rho_B] - E^{HK}_{v_A}[\rho^{isol}_A] - E^{HK}_{v_B}[\rho^{isol}_B], \tag{10}$$

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

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

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

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 $\rho_B = \rho^{isol}_B$

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

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

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

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

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 $E_{v_B}^{HK}[\rho_B^{v"}] - E_{v_B}^{HK}[\rho_B^{isol}] \geq 0$ contributes to the interaction energy. It is approximated as:

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

where $\Phi"_B$ is the optimal determinant yielding $\rho_B^{v"}$ (not the Hartree-Fock wavefunction for N_B electrons in the potential $v_B(\mathbf{r})$!).

Using the above approximation in Eq. 11 leads to:

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

where Φ_X is the Hartree-Fock wavefunction and $E_{v_X}^c$ 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[12]. 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: $v_A'^{FAT}$, $\Phi_A'^{FAT}$, $\rho_A'^{FAT}$, $E_{v_A'^{FAT}}^c$, and $\Delta\rho_{v_A'^{FAT}}^c$ for the subsystem A and $v_B'^{FAT}$ $\Phi_B'^{FAT}$, $\rho_B'^{FAT}$, $E_{v_B'^{FAT}}^c$, and $\Delta\rho_{v_A'^{FAT}}^c$ for the subsystem B. Using this notation, Eq. 14 reads:

$$\begin{aligned} E_{int}^{FDET(\rho_B'^{FAT})} &= \langle \Phi_A'^{FAT} | \hat{H}_{v_A} | \Phi_A'^{FAT} \rangle + E_{v_A'^{FAT}}^c + \langle \Phi_B'^{FAT} | \hat{H}_{v_B} | \Phi_B'^{FAT} \rangle + E_{v_B'^{FAT}}^c \\ &+ E_{v_A,v_B}^{elst,int}[\rho_A'^{FAT}, \rho_B'^{FAT}] + E_{xcT}^{nad}[\rho_A'^{FAT}, \rho_B'^{FAT}] - E_k[\Delta\rho_{v_A'^{FAT}}^c, \rho_A'^{FAT}, \rho_B^{FAT}] \\ &- \langle \Phi_A | \hat{H}_{v_A} | \Phi_A \rangle - E_{v_A}^c - \langle \Phi_B | \hat{H}_{v_B} | \Phi_B \rangle - E_{v_B}^c \end{aligned} \tag{15}$$

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

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

Following Eq. 6, both above expressions yield $E^{HK}_{v_{AB}}[\rho_A'^{FAT} + \rho_B'^{FAT}]$ 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.

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

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 errors in density

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

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

where Θ is the Heaviside step function, is used.

M is bound $0 \leq M \leq N_B$. At the lower bound ($M = 0$), the inequality in Eq. 2 is reached (subject of the condition of N_A-representability of $\rho^o_{AB} - \rho_B$. If, additionally, $\rho^o_{AB} - \rho_B$ is v-representable, the exact solution of Eq. 3 yields this density. At the upper bound ($M = N_B$), the densities ρ_B and $\rho^o_{AB} - \rho_B$ do not overlap.

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

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

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):

$$M[\rho_B - \rho^o_{AB}] \leq P[\rho_A^{Eq.3} + \rho_B - \rho^o_{AB}] \leq N_{AB}. \tag{20}$$

The upper bound is almost impossible to reach in practical calculations, as it corresponds to the case where the overlap of $\rho_A^{Eq.3} + \rho_B$ and ρ^o_{AB} is zero. A more insightful term of comparison can be obtained by comparing the sum of isolated fragments to the reference density by means of $P[\rho_A^{isol} + \rho_B^{isol} - \rho^o_{AB}]$. In the interest of compactness, we denote this as

$$P_{cmpl} = P[\rho_A^{isol} + \rho_B^{isol} - \rho^o_{AB}] \tag{21}$$

as it actually constitutes the integrated density change upon complexation and corresponds to the density error if no optimisation of the active subsystem were performed, which in turn relates to the strength of the intersubsystem interaction.

M and P are used here to discuss different choices of $\rho_B(\mathbf{r})$ 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 Slater determinants. For consistency, so was the reference density ρ^{ref}, used to evaluate M and P parameters. The reference interaction energy E_{int}^{ref} was calculated with MP2 theory with counterpoise correction.

$E_{xcT}^{nad}[\rho_A, \rho_B]$ and its functional derivative were approximated using local-density approximation (LDA) for all its components: Thomas-Fermi[13,14] for the kinetic energy, Dirac-Slater[15] for the exchange energy, and the Vosko-Wilk-Nusair[16] for correlation energy:

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

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

The used approximation neglects the correlation component of the embedding potential applicable for single determinants[2], representing thus the case for which the relation given in Eq. 6 applies. 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 system of interest is active and the enviroment is frozen, and 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) single determinant. The above iterative procedures were performed using the author's version[17] of CCParser[18] and CCDatabase[19] handling the automatic submission of Q-Chem5.4[20] calculations, parsing, and collecting the results.

The aug-cc-pVDZ atomic basis sets was used in all calculations. 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. 21. Throughout this work, the following convention is used for relating Eq. 3 and to the names of the com-

plexes/clusters: if the system name is AAA-BBB, AAA is associated with Ψ_A and BBB with ρ_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.

Parameters M and P were evaluated using the grid integration implemented in PySCF[22], which is based on Becke[23]-Lebedev[24] grids. The plots were prepared using the Python modules: pandas[25] and matplotlib[26].

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[27]. 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.[28] The partition of the total density is not unique.[29], so, as long as the non-negativity condition for $\rho^o_{AB} - \rho_B$ is satisfied, any variation of ρ_B is not changing the FDET energy. In FDET-based methods, on the other hand, $E^{nad}_{xcT}[\rho_A, \rho_B] \approx \tilde{E}^{nad}_{xcT}[\rho_A, \rho_B]$. Even though $\tilde{E}^{nad}_{xcT}[\rho_A, \rho_B]$ is symmetric with respect to exchanging A with B, the errors in the potentials $\tilde{v}^{nad}_{xcT}[\rho_A, \rho_B](\mathbf{r})$ and $\tilde{v}^{nad}_{xcT}[\rho_B, \rho_A](\mathbf{r})$ are not the same. As a result, the unique pair ρ_A and ρ_B is usually obtained in the *freeze-and-thaw* calculations.[4,30,31] 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 $\tilde{E}^{nad}_{xcT}[\rho_A, \rho_B] - E^{nad}_{xcT}[\rho_A, \rho_B]$. None of them can be uniquely attributed to the electronic polarisation of ρ_B.

The FDET calculations for $\rho_B = \rho^{isol}_B$, 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 ρ_A is always positive, such implicit treatment of the polarisation cannot be expected to be complete. This implicit treatment depends on the basis set used which in turn determines the amount of possible overlap between ρ_A and ρ_B. 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[32]. 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[33] 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.[34,35]

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 $E_{xcT}^{nad}[\rho_A, \rho_B]$

complex	$P^{[a]}$	$P_{cmpl}^{[b]}$	$M^{[c]}$	δE_{int} [d]	$E_{int}^{FDET(FAT)}$ [e]	E_{int}^{ref} [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] $M = M[\rho^{ref} - \rho_B^{FDET(FAT)}]$ with $M[\rho]$ defined in Eq. 18

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

[c] $P = P[\rho^{ref} - \rho_{tot}^{FDET(FAT)}]$ with $P[\rho]$ defined in Eq. 19

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

[e] $E_{int}^{FDET(FAT)}$ is given 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.

could be expected based on our previous numerical experience[31,36,37]. 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

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

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

[c] $P = P[\rho^{ref} - \rho_{tot}^{FDET(FAT)}]$ with $P[\rho]$ defined in Eq. 19

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

[e] $E_{int}^{FDET(FAT)}$ is given 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.

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 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. Such simulations hinge on the possibility to

generate ρ_B in a simpler/less costly way. The following sections concern such applications of FDET. To this end, the the deviations from the reference data will be analysed in presence of additional assumptions concerning ρ_B obtained in alternative ways. In particular, the effect of these additional assumptions on the parameter M will be discussed.

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:

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

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 $\rho_{AB}^o - \rho_B$ 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, $f[\rho_B^{isol}]$ 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.

complex	$P^{[a]}$	$P^{[b]}_{cmpl}$	$M^{[c]}$	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] $M = M[\rho^{ref} - \rho_B^{isol}]$ with $M[\rho]$ defined in Eq. 18

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

[c] $P = P[\rho^{ref} - \rho_{tot}^{FDET(\rho_B=\rho_B^{isol})}]$ with $P[\rho]$ defined in Eq. 19

[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 given 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

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

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 $\rho_B = \rho_B^{isol}$ 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 $\rho_{AB}^{o} - \rho_B$ 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. 9, the $\rho_B = \rho_B^{isol}$ approximation in FDET leads to an astonishingly good overall accuracy of the complexation induced shifts of the vertical 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).

complex	$P^{[a]}$	$P^{[b]}_{cmpl}$	$M^{[c]}$	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] $M = M[\rho^{ref} - \rho_B^{isol}]$ with $M[\rho]$ defined in Eq. 18

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

[c] $P = P[\rho^{ref} - \rho_{tot}^{FDET(\rho_B=\rho_B^{isol})}]$ with $P[\rho]$ defined in Eq. 19

[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 given 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.

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[9], 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.

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 $\tilde{E}_{xcT}^{nad}[\rho_A, \rho_B]$. Only

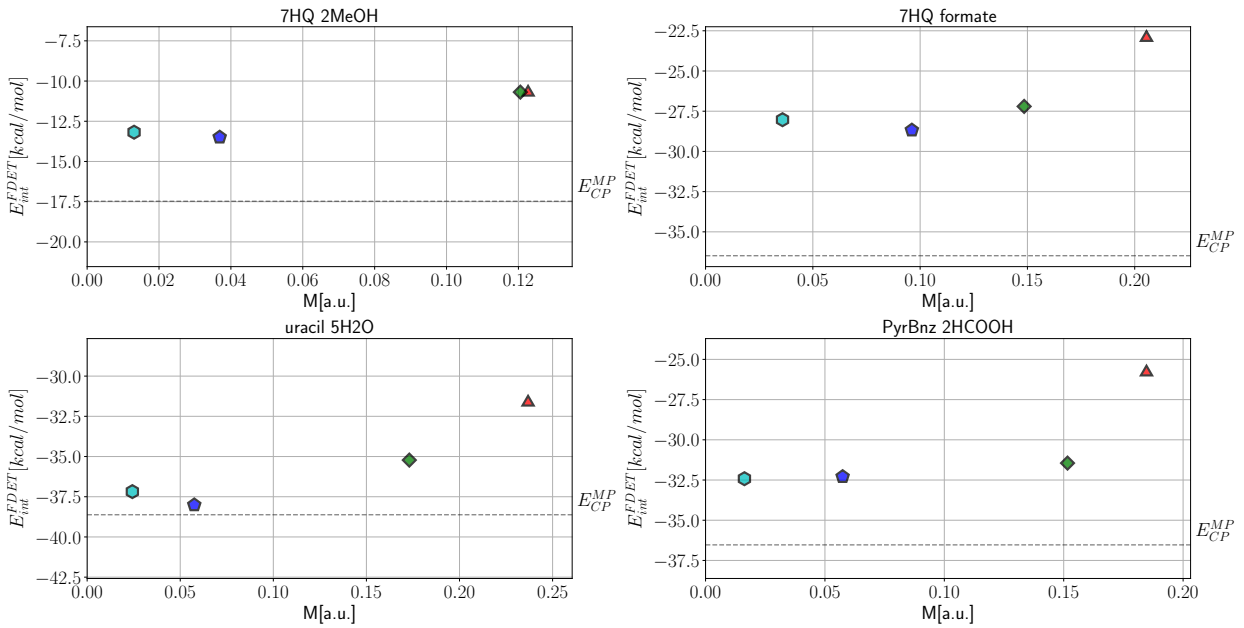

FIG. 1. Integrated negative density M and the FDET-MP2 interaction energy for various choices of ρ_B: a) ρ_B^{isol} (orange triangles), b) ρ_B^{FAT} (light blue hexagons), c) $\rho_B^{pp(Mulliken)}$ (green diamonds), and d) $\rho_B^{pp(ChelPG)}$ (dark blue pentagons). Data obtained using the *monomer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

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 1kcal/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.

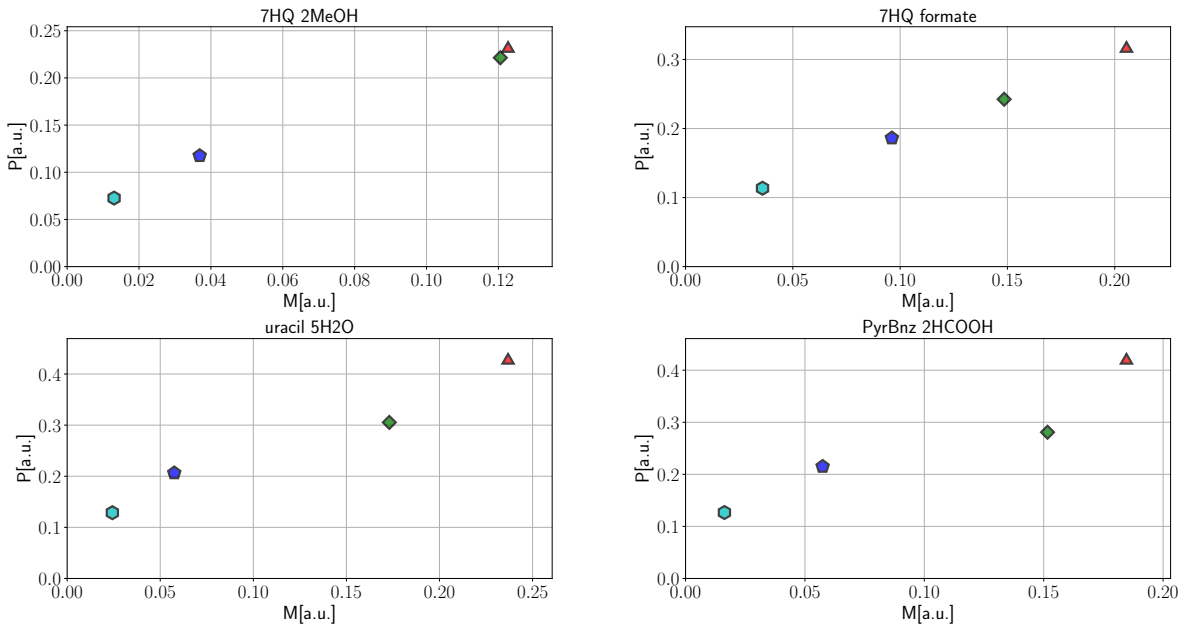

FIG. 2. Integrated negative density M and the total density error P for various choices of ρ_B: a) ρ_B^{isol} (orange triangles), b) ρ_B^{FAT} (light blue hexagons), c) $\rho_B^{pp(Mulliken)}$ (green diamonds), and d) $\rho_B^{pp(ChelPG)}$ (dark blue pentagons). Data obtained using the *monomer expansion*.

IV. CONCLUSIONS

The present work shows comprehensively that the error in the energy and the density due to the used semi-local approximations for $E_{xcT}^{nad}[\rho_A, \rho_B]$ 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 such FDET-based calculations if such density is used[9]. Whereas the approximation to $E_{xcT}^{nad}[\rho_A, \rho_B]$ can result in positive or negative energy error for a given electronic state, the energy error due to such a choice of ρ_B 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 $\rho_{AB}^o - \rho_B$ is not possible. This would require *a priori* knowledge of the total density. The present work provided a link between the violation of the condition of the non-negativity of $\rho_{AB}^o - \rho_B$ 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 $\rho^o_{AB} - \rho_B$ can be significantly reduced in practice.

SUPPLEMENTARY MATERIAL

*** The Supplementary Material includes the proof of Eq. 20, the magnitude of the counterpoise corrections for E^{ref}_{int}, the values of $E_k[\Delta\rho^c_{v'_X}, \rho'_X, \rho_Y]$ for the different calculations, and the data plotted in the Figures. Any additional data is made available on request from the authors.

REFERENCES

[1]Wesolowski, T. A.; Warshel, A. Frozen density functional approach for ab initio calculations of solvated molecules. *The Journal of Physical Chemistry* **1993**, *97*, 8050–8053.

[2]Wesołowski, T. A. Embedding a multideterminantal wave function in an orbital-free environment. *Physical Review A* **2008**, *77*, 012504.

[3]Pernal, K.; Wesolowski, T. A. Orbital-free effective embedding potential: Density-matrix functional theory case. *International Journal of Quantum Chemistry* **2009**, *109*, 2520–2525.

[4]Wesolowski, T. A.; Shedge, S.; Zhou, X. Frozen-Density Embedding Strategy for Multilevel Simulations of Electronic Structure. *Chemical Reviews* **2015**, *115*, 5891–5928.

[5]Wesolowski, T. A. On the Correlation Potential in Frozen-Density Embedding Theory. *Journal of Chemical Theory and Computation* **2020**, *16*, 6880–6885.

[6]Hohenberg, P.; Kohn, W. Inhomogeneous Electron Gas. *Physical Review* **1964**, *136*, B864–B871.

[7]Wesolowski, T. A.; Savin, A. In *Recent Progress in Orbital-Free Density Functional Theory*; Wesolowski, T. A., Wang, Y. A., Eds.; Recent Advances in Computational Chemistry; World Scientific: Singapore, 2013; Vol. 6; pp 275–295.

[8]Fux, S.; Jacob, C. R.; Neugebauer, J.; Visscher, L.; Reiher, M. Accurate frozen-density embedding potentials as a first step towards a subsystem description of covalent bonds. *The Journal of Chemical Physics* **2010**, *132*, 164101.

[9]Ricardi, N.; Zech, A.; Gimbal-Zofka, Y.; Wesolowski, T. A. Explicit vs. implicit electronic polarisation of environment of an embedded chromophore in frozen-density embedding theory. *Physical Chemistry Chemical Physics* **2018**, *20*, 26053–26062.

[10]Dułak, M.; Kaminski, J. W.; Wesołowski, T. A. Linearized orbital-free embedding potential in self-consistent calculations. *International Journal of Quantum Chemistry* **2009**, *109*, 1886–1897.

[11]Zech, A.; Aquilante, F.; Wesolowski, T. A. Homogeneity properties of the embedding potential in frozen-density embedding theory. *Molecular Physics* **2016**, *114*, 1199–1206.

[12]Wesolowski, T. A.; Weber, J. Kohn-Sham equations with constrained electron density: an iterative evaluation of the ground-state electron density of interacting molecules. *Chemical Physics Letters* **1996**, *248*, 71–76.

[13]Thomas, L. H. The calculation of atomic fields. *Mathematical Proceedings of the Cambridge Philosophical Society* **1927**, *23*, 542.

[14]Fermi, E. Eine statistische Methode zur Bestimmung einiger Eigenschaften des Atoms und ihre Anwendung auf die Theorie des periodischen Systems der Elemente. *Zeitschrift für Physik* **1928**, *48*, 73–79.

[15]Slater, J. C. The Theory of Complex Spectra. *Physical Review* **1929**, *34*, 1293–1322.

[16]Vosko, S. H.; Wilk, L.; Nusair, M. Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis. *Canadian Journal of Physics* **1980**, *58*, 1200–1211.

[17]Ricardi, N. CCParser.

[18]Zech, A. CCParser.

[19]Ricardi, N.; González-Espinoza, C. E. CCDatabase. `https://github.com/NicoRicardi/CCDatabase`, 2020.

[20]Epifanovsky, E.; Gilbert, A. T. B.; Feng, X.; Lee, J.; Mao, Y.; Mardirossian, N.; Pokhilko, P.; White, A. F.; Coons, M. P.; Dempwolff, A. L.; Gan, Z.; Hait, D.; Horn, P. R.; Jacobson, L. D.; Kaliman, I.; Kussmann, J.; Lange, A. W.; Lao, K. U.; Levine, D. S.; Liu, J.; McKenzie, S. C.; Morrison, A. F.; Nanda, K. D.; Plasser, F.; Rehn, D. R.; Vidal, M. L.; You, Z.-Q.; Zhu, Y.; Alam, B.; Albrecht, B. J.; Aldossary, A.; Alguire, E.; Andersen, J. H.; Athavale, V.; Barton, D.; Begam, K.; Behn, A.; Bellonzi, N.; Bernard, Y. A.; Berquist, E. J.; Burton, H. G. A.; Carreras, A.; Carter-Fenk, K.; Chakraborty, R.; Chien, A. D.; Closser, K. D.; Cofer-Shabica, V.; Dasgupta, S.; de Wergi-

fosse, M.; Deng, J.; Diedenhofen, M.; Do, H.; Ehlert, S.; Fang, P.-T.; Fatehi, S.; Feng, Q.; Friedhoff, T.; Gayvert, J.; Ge, Q.; Gidofalvi, G.; Goldey, M.; Gomes, J.; González-Espinoza, C. E.; Gulania, S.; Gunina, A. O.; Hanson-Heine, M. W. D.; Harbach, P. H. P.; Hauser, A.; Herbst, M. F.; Hernández Vera, M.; Hodecker, M.; Holden, Z. C.; Houck, S.; Huang, X.; Hui, K.; Huynh, B. C.; Ivanov, M.; Jász, A.; Ji, H.; Jiang, H.; Kaduk, B.; Kähler, S.; Khistyaev, K.; Kim, J.; Kis, G.; Klunzinger, P.; Koczor-Benda, Z.; Koh, J. H.; Kosenkov, D.; Koulias, L.; Kowalczyk, T.; Krauter, C. M.; Kue, K.; Kunitsa, A.; Kus, T.; Ladjánszki, I.; Landau, A.; Lawler, K. V.; Lefrancois, D.; Lehtola, S.; Li, R. R.; Li, Y.-P.; Liang, J.; Liebenthal, M.; Lin, H.-H.; Lin, Y.-S.; Liu, F.; Liu, K.-Y.; Loipersberger, M.; Luenser, A.; Manjanath, A.; Manohar, P.; Mansoor, E.; Manzer, S. F.; Mao, S.-P.; Marenich, A. V.; Markovich, T.; Mason, S.; Maurer, S. A.; McLaughlin, P. F.; Menger, M. F. S. J.; Mewes, J.-M.; Mewes, S. A.; Morgante, P.; Mullinax, J. W.; Oosterbaan, K. J.; Paran, G.; Paul, A. C.; Paul, S. K.; Pavošević, F.; Pei, Z.; Prager, S.; Proynov, E. I.; Rák, A.; Ramos-Cordoba, E.; Rana, B.; Rask, A. E.; Rettig, A.; Richard, R. M.; Rob, F.; Rossomme, E.; Scheele, T.; Scheurer, M.; Schneider, M.; Sergueev, N.; Sharada, S. M.; Skomorowski, W.; Small, D. W.; Stein, C. J.; Su, Y.-C.; Sundstrom, E. J.; Tao, Z.; Thirman, J.; Tornai, G. J.; Tsuchimochi, T.; Tubman, N. M.; Veccham, S. P.; Vydrov, O.; Wenzel, J.; Witte, J.; Yamada, A.; Yao, K.; Yeganeh, S.; Yost, S. R.; Zech, A.; Zhang, I. Y.; Zhang, X.; Zhang, Y.; Zuev, D.; Aspuru-Guzik, A.; Bell, A. T.; Besley, N. A.; Bravaya, K. B.; Brooks, B. R.; Casanova, D.; Chai, J.-D.; Coriani, S.; Cramer, C. J.; Cserey, G.; DePrince, A. E.; DiStasio, R. A.; Dreuw, A.; Dunietz, B. D.; Furlani, T. R.; Goddard, W. A.; Hammes-Schiffer, S.; Head-Gordon, T.; Hehre, W. J.; Hsu, C.-P.; Jagau, T.-C.; Jung, Y.; Klamt, A.; Kong, J.; Lambrecht, D. S.; Liang, W.; Mayhall, N. J.; McCurdy, C. W.; Neaton, J. B.; Ochsenfeld, C.; Parkhill, J. A.; Peverati, R.; Rassolov, V. A.; Shao, Y.; Slipchenko, L. V.; Stauch, T.; Steele, R. P.; Subotnik, J. E.; Thom, A. J. W.; Tkatchenko, A.; Truhlar, D. G.; Van Voorhis, T.; Wesolowski, T. A.; Whaley, K. B.; Woodcock, H. L.; Zimmerman, P. M.; Faraji, S.; Gill, P. M. W.; Head-Gordon, M.; Herbert, J. M.; Krylov, A. I. Software for the frontiers of quantum chemistry: An overview of developments in the Q-Chem 5 package. *The Journal of Chemical Physics* **2021**, *155*, 084801.

[21]Zech, A.; Ricardi, N.; Prager, S.; Dreuw, A.; Wesolowski, T. A. Benchmark of Excitation Energy Shifts from Frozen-Density Embedding Theory: Introduction of a Density-Overlap-

Based Applicability Threshold. *Journal of Chemical Theory and Computation* **2018**, *14*, 4028–4040.

[22]Sun, Q.; Berkelbach, T. C.; Blunt, N. S.; Booth, G. H.; Guo, S.; Li, Z.; Liu, J.; McClain, J. D.; Sayfutyarova, E. R.; Sharma, S.; Wouters, S.; Chan, G. K. PySCF: the Pythonbased simulations of chemistry framework. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2017**, *8*, e1340.

[23]Becke, A. D. A multicenter numerical integration scheme for polyatomic molecules. *The Journal of Chemical Physics* **1988**, *88*, 2547–2553.

[24]Lebedev, V. I.; Laikov, D. N. A quadrature formula for the sphere of the 131st algebraic order of accuracy. *Doklady Mathematics* **1999**, *59*, 477–481.

[25]Wes McKinney, Data Structures for Statistical Computing in Python. Proceedings of the 9th Python in Science Conference. 2010; pp 56 – 61.

[26]Hunter, J. D. Matplotlib: A 2D Graphics Environment. *Computing in Science & Engineering* **2007**, *9*, 90–95.

[27]Jeziorski, B.; Moszynski, R.; Szalewicz, K. Perturbation Theory Approach to Intermolecular Potential Energy Surfaces of van der Waals Complexes. *Chemical Reviews* **1994**, *94*, 1887–1930.

[28]Humbert-Droz, M.; Zhou, X.; Shedge, S. V.; Wesolowski, T. A. How to choose the frozen density in Frozen-Density Embedding Theory-based numerical simulations of local excitations? *Theoretical Chemistry Accounts* **2014**, *133*, 1405.

[29]Savin, A.; Wesolowski, T. A. *Prog. Theor. Chem. Phys.*; 2009; Vol. 19; pp 311–326.

[30]Wesolowski, T. A. Density functional theory with approximate kinetic energy functionals applied to hydrogen bonds. *The Journal of Chemical Physics* **1997**, *106*, 8516–8526.

[31]Dułak, M.; Wesołowski, T. A. Interaction energies in non-covalently bound intermolecular complexes derived using the subsystem formulation of density functional theory. *Journal of Molecular Modeling* **2007**, *13*, 631–642.

[32]Breneman, C. M.; Wiberg, K. B. Determining atom-centered monopoles from molecular electrostatic potentials. The need for high sampling density in formamide conformational analysis. *Journal of Computational Chemistry* **1990**, *11*, 361–373.

[33]Mulliken, R. S. Electronic Population Analysis on LCAO-MO Wave Functions.I. *The Journal of Chemical Physics 23*, 1833–1840.

[34]Fradelos, G.; Wesołowski, T. A. Importance of the Intermolecular Pauli Repulsion in Embedding Calculations for Molecular Properties: The Case of Excitation Energies for a Chromophore in Hydrogen-Bonded Environments. *The Journal of Physical Chemistry A* **2011**, *115*, 10018–10026.

[35]Fradelos, G.; Wesolowski, T. A. The Importance of Going beyond Coulombic Potential in Embedding Calculations for Molecular Properties: The Case of Iso-G for Biliverdin in Protein-Like Environment. *Journal of Chemical Theory and Computation* **2011**, *7*, 213–222.

[36]Wesolowski, T. A.; Tran, F. Gradient-free and gradient-dependent approximations in the total energy bifunctional for weakly overlapping electron densities. *The Journal of Chemical Physics* **2003**, *118*, 2072–2080.

[37]Kevorkyants, R.; Dulak, M.; Wesolowski, T. A. Interaction energies in hydrogen-bonded systems: A testing ground for subsystem formulation of density-functional theory. *The Journal of Chemical Physics* **2006**, *124*, 024104.