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

Niccolò Ricardi,* Cristina E. González-Espinoza,* and Tomasz Adam Wesołowski*

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

E-mail: Niccolo.Ricardi@unige.ch; Cristina.GonzalezEspinoza@unige.ch; Tomasz.Wesolowski@unige.ch

Abstract

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 electrionic polarisation of the environment.

1 Introduction

The formal framework of Frozen-Density Embedding Theory (FDET) provides the basis of multiscale 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] Throughout the present work, $E_{v_{AB}}^{FDET}[\Psi_A, \rho_B]$ denotes the FDET energy functional in the case of interacting N_A-electron Hamiltonian ($\hat{H}_A$) and the embedded N_A-electron wavefunction (Ψ_A). The second independent variable is a non-negative real function representing the electron density of the environment ($\rho_B(\mathbf{r})$).

$E_{v_{AB}}^{FDET}[\Psi_A, \rho_B]$ satisfies the following relations:

$$E_{v_{AB}}^{FDET}[\Psi_A^o, \rho_B] = E_{v_{AB}}^{HK}[\rho_A^o + \rho_B] \geq E_o, \tag{1}$$

where $E_{v_{AB}}^{HK}[\rho]$ is the Hohenberg-Kohn energy functional, E_o is the exact energy of the system of N_{AB} electrons in the external potential $v_{AB}(\mathbf{r})$, and $\Psi_A^o = \arg\min_{\Psi_A} E_{v_A,v_B}^{FDET}[\Psi_A, \rho_B]$.

Not clear a

Any practical implementation of the formal framework of FDET involves one approximation and one assumption of entirely different origin that affect the total energy obtained numerically. The approximation concerns the bi-functional $E_{xcT}^{nad}[\rho_A, \rho_B]$. It is defined implicitly (see Ref. 2), and its analytic form is not known. $E_{xcT}^{nad}[\rho_A, \rho_B]$ together with its derivative, is approximated by means of some analytic expression. $v_{xct}^{nad}[\rho_A, \rho_B](\mathbf{r}) = \frac{\delta E_{xct}^{nad}[\rho_A,\rho_B]}{\delta \rho_A(\mathbf{r})}$ is one of the components of the FDET embedding potential. The assumption, on the other hand, concerns the choice of the density $\rho_B(\mathbf{r})$. For the exact functional $E_{v_{AB}}^{FDET}[\Psi_A, \rho_B]$ and such $\rho_B(\mathbf{r})$ that $\forall_{\mathbf{r}}\ \rho_o(\mathbf{r}) - \rho_B(\mathbf{r}) > 0$, the equality in the right-hand-side of Eq. 1 is reached. We refer to this as the *non-negativity condition*. Depending on the choice of $\rho_B(\mathbf{r})$ the target function defined as

$$\rho_t(\mathbf{r}) \equiv \rho_o(\mathbf{r}) - \rho_B(\mathbf{r}) \tag{2}$$

might violate the non-negativity condition. In such a case, the exact ground-state energy cannot be obtained from FDET because $\rho_t(\mathbf{r})$ is not N_A-representable. FDET yields only the upper bound to the exact ground-state energy of the total system. FDET formulation of the embedding problem results in the computational gains if compared to conventional quantum mechanical treatment of all N_{AB} electrons. They originate from the fact that optimisation of Ψ_A^o involves solving a N_A-electron problem with $N_A < N_{AB}$. This comes, however, at a price. FDET can yield the upper bound of the total energy of the whole system (E_o). In multi-level simulation methods based on FDET, however, using the exact bi-functional $E_{xcT}^{nad}[\rho_A, \rho_B]$ and its functional derivative is not practical. Similarly it is not possible to verify whether the non-negativity condition is satisfied or not. Both would require solving a complete N_{AB}-electron problem. The effects due to the approximation for $E_{xcT}^{nad}[\rho_A, \rho_B]$ and the possible violation of the non-negativity condition combine leading thus to the deviations of the total energy and density derived from FDET from their exact counterparts. 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. Wesolowski and Savin analysed the errors due to approximation to $E_{xcT}^{nad}[\rho_A, \rho_B]$ for various choices for $\rho_B(\mathbf{r})$ such that the non-negativity condition was not violated in a model system for which the exact solutions of FDET are available violated.[6] TW to NR: English and scientific clarity need improvement in the sentence below.

Fux et al.,[7] demonstrated that the violation is minor and only due to numerical effects

Nicco: the effects on what ? in which systems?

The secondary objective of the present analyses is investigation of the relation between the non-negativity condition and electronic polarisation of the environment.

For each analysed quantity (density or energy) obtained from FDET, the difference from the reference obtained from the conventional calculations for the whole complex is consider as its error, which results from the choice of $\rho_B(\mathbf{r})$ and the approximation for $E_{xcT}^{nad}[\rho_A, \rho_B]$. The quantification of the relative significance of the two sources of errors in the FDET energy is

made by comparing the FDET results obtained using two extreme choices: $\rho_B(\mathbf{r}) = \rho_B^{isol}(\mathbf{r})$, where $\rho_B^{isol}(\mathbf{r})$ is the electron density of the isolated environment, and $\rho_B(\mathbf{r}) = \rho_B^{FAT}(\mathbf{r})$, where $\rho_B^{FAT}(\mathbf{r})$ denotes the density $\rho_B(\mathbf{r})$ minimising the FDET energy functional with respect to variation of both independent variables for which this functional is defined. In both cases the non-negativity condition might be violated but the principal sources of errors in the FDET energies are different. In the optimised case, the errors in density and energy can be attributed mainly to the used approximation for $E_{xcT}^{nad}[\rho_A, \rho_B]$. If the non-negativity condition is violated, it is the result of errors in the approximation used for $v_{xct}^{nad}[\rho_A, \rho_B](\mathbf{r})$. In the $\rho_B(\mathbf{r}) = \rho_B^{isol}(\mathbf{r})$ case, on the other hand the violation of the non-negativity condition result from the two factors combine. For any other choice of $\rho_B(\mathbf{r})$, the relative significance of the two factors might be different. Besides the two above choices for $\rho_B(\mathbf{r})$, other densities are also considered in order to relate the magnitude of violation of the non-negativity condition to the electronic polarisation of the environment.

The analyses are made for four weakly bound molecular clusters comprising two or more molecules. In each case, the partition of $v_{AB}(\mathbf{r}) = v_A(\mathbf{r}) + v_B(\mathbf{r})$ and $N_{AB} = N_A + N_B$ corresponds to the selected molecule and the remaining one(s), respectively. The second order Møller-Plesset perturbation theory (MP2) energies and the first-order (MP1) densities obtained for the whole cluster are used as a reference. The FDET energy corresponding to this reference is obtained using the recently derived formula for the total energy.[5] We refer to this variant of FDET as *embedded second-order Møller-Plesset* perturbation theory (FDET-MP2).

The chosen clusters were considered previously in our study of complexation induced shifts in the excitation energies[8] 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 $\rho_B(\mathbf{r})$ 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-

ground-state

negativity condition on the total energy, or *c*) both. For this reason, the present work focusses on ground-state properties. 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*) 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*) benzimidazolide bound to two formic acid molecules (PyrBnz-2HCOOH). Throughout this work, the following convention is used for the names of the systems (AAA-BBB): AAA is associated with Ψ_A and BBB with $\rho_B(\mathbf{r})$. 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 bond distance (1.36 Å) is comparable to that of its (NR what is "its covalent bond"? covalent bond (1.09 Å). In PyrBnz-2HCOOH, the embedded system (PyrBnz) is a hydrogen acceptor and the atom involved carries a significant negative charge.

2 Computational Details

2.1 Approximate FDET-based methods and their "exact" counterparts

For interpretation purposes, it is convenient to split the external potential $v_{AB}(\mathbf{r})$ into components defining the embedded species and the environment $v_{AB}(\mathbf{r}) = v_A(\mathbf{r}) + v_B(\mathbf{r})$. The embedded wavefunction and the total energy obtained from FDET do not depend on such splitting. It is convenient to consider $v_A(\mathbf{r})$ being the Coulomb potential due to the nuclei of the embedded species. 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 ($\hat{H}_B$). Using the above convention, the FDET energy functional $E^{FDET}_{v_A,v_B}[\Psi_A, \rho_B]$ or the embedded wavefunction of the Full Config-

uration Interaction form reads:

$$E^{FDET}_{v_A,v_B}[\Psi_A,\rho_B] = \langle\Psi_A|\hat{H}_A|\Psi_A\rangle + E^{HK}_{v_B}[\rho_B] + E^{elst,int}_{v_A,v_B}[\rho_A,\rho_B] + E^{nad}_{xcT}[\rho_A,\rho_B], \tag{3}$$

where $\rho_A(\mathbf{r})$ is the density corresponding to Ψ_A and all classical electrostatic contributions to the interaction energy are collected in $E^{elst,int}_{v_A,v_B}[\rho_A,\rho_B]$.

All but the $E^{nad}_{xcT}[\rho_A,\rho_B]$ and $E^{HK}_{v_B}[\rho_B]$ components of this functional have known analytic form of their dependencies on $\rho_B(\mathbf{r})$ and Ψ_A. Concerning $E^{HK}_{v_B}[\rho_B]$, its numerical evaluation is not needed in most of practical applications of FDET. A simulation that aims at obtaining such quantities that are given as expectation values of some operators evaluated at the optimal embedded wave-function, for instance. In multi-level simulations, on the other hand, evaluation of $E^{HK}_{v_B}[\rho_B]$ is replaced by the expression corresponding to the applied formalism for the environment. Concerning the $E^{nad}_{xcT}[\rho_A,\rho_B]$ term, it is approximated:

$$E^{nad}_{xcT}[\rho_A,\rho_B] \approx \tilde{E}^{nad}_{xcT}[\rho_A,\rho_B] \tag{4}$$

by means of some analytic expression (see section 2.4) making it possible to obtain analytically the corresponding component of the embedding potential.

The FDET embedding potential depends on the form of the embedded wavefunction.[2] For the embedded wavefunctions of the Full Configuration Interaction form, it reads:

$$v^{FDET}_{emb}[\rho_A,\rho_B;v_B](\mathbf{r}) = v_B(\mathbf{r}) + \int \frac{\rho_B(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}d\mathbf{r}' + \frac{\delta E^{nad}_{xct}[\rho_A,\rho_B]}{\delta\rho_A(\mathbf{r})}. \tag{5}$$

In case of embedded single determinants, the FDET energy functional comprises an additional term (correlation functional $E_c[\rho_A]$) and the embedding potential comprises an additional component (the functional for the correlation potential $v_c[\rho_A](\mathbf{r}) = \frac{\delta E_c[\rho_A]}{\delta\rho_A(\mathbf{r})}$). According to our recent formal result, using the approximation to $v_c[\rho_A](\mathbf{r})$ can be avoided even if the

embedded wavefunction is a single determinant.[5] The effect of neglecting $v_c[\rho_A](\mathbf{r})$ on the energy can be evaluated if the correlation energy and correlation induced change of electron density are available for an auxiliary N_A-electron system defined by the external potential $v'(\mathbf{r})$. The derived expression for the total energy is consistent with the Hohenberg-Kohn energy functional up to the quadratic terms in the correlation induced density change ($\Delta\rho$). The external potential in the auxiliary system comprises the potential $v_A(\mathbf{r})$ and the self-consistently obtained FDET potential given in Eq. 5. In case of FDET-MP2, the relation between the quantities available in practice and the Hohenberg-Kohn energy functional reads:

$$\begin{aligned} E^{HK}_{v_{AB}}[\rho^{MP1}_A + \rho_B] =& E^{FDET}_{v_A,v_B}[\Phi'^{HF}_A, \rho_B] + E^{c(MP2)}_{v'_A} + E^{HK}_{v_B}[\rho_B] \\ & E_k[\Delta\rho^{c(MP1)}_{v'_A}, \rho'^{HF}_A, \rho_B] + O(\Delta\rho)^2 + E^{c(MPn \text{ for } n>2)))}_{v'_A} \end{aligned} \tag{6}$$

where

$$\begin{aligned} & E_k[\Delta\rho^{c(MP1)}_{v'_A}, \rho'^{o}_A, \rho_B] = \\ & \int \Delta\rho^{c(MP1)}_{v'_A}(\mathbf{r}) \int \rho'^{HF}_A(\mathbf{r}') f^{nad}_{xcT}[\rho'^{HF}_A, \rho_B](\mathbf{r}, \mathbf{r}') d\mathbf{r}' \mathrm{d}\mathbf{r}, \end{aligned} \tag{7}$$

with

$$f^{nad}_{xcT}[\rho_A, \rho_B](\mathbf{r}, \mathbf{r}') = \frac{\delta^2 E^{nad}_{xcT}[\rho_A, \rho_B]}{\delta\rho_A(\mathbf{r})\delta\rho_A(\mathbf{r}')}, \tag{8}$$

where $\Delta\rho^{c(MP1)}_{v'_A}(\mathbf{r}) = \rho^{MP1}_{v'_A}(\mathbf{r}) - \rho^{HF}_{v'_A}(\mathbf{r})$ and the $E^{c(MPn \text{ for } n>2)))}_{v'_A}$ term, which collects all higher than second order corrections to the energy in the perturbation series. The higher order contributions to the energy ($E^{c(MPn \text{ for } n>2))}_{v'_A}$ and $O(\Delta\rho)^2$) are neglected in the numerical implementation of the above formula. The FDET-MP2 energy used to obtain the numerical

results is given by:

$$
\begin{aligned}
E^{FDET-MP2}_{v_A,v_B} &= \langle \Phi'^{HF}_A | \hat{H}_A | \Phi'^{HF}_A \rangle + E^{c(MP2)}_{v'_A} \\
&+ E^{elst,int}_{v_A,v_B}[\rho'^{HF}_A, \rho_B] + \tilde{E}^{nad}_{xcT}[\rho'^{HF}_A, \rho_B] \\
&+ \langle \Phi'^{HF}_B | \hat{H}_B | \Phi'^{HF}_B \rangle + E^{c(MP2)}_{v'_B} + \tilde{E}_k[\Delta\rho^{c(MP1)}_{v'_A}, \rho'^{HF}_A, \rho_B],
\end{aligned}
\tag{9}
$$

where a tilde in $\tilde{E}^{nad}_{xcT}[\rho_A, \rho_B]$ indicates that the analytical formula used to approximate the exact functional (see section 2.4), Φ'^{HF}_B denotes the optimal and consistent with the embedding potential embedded determinant, and $\langle \Phi'^{HF}_B | \hat{H}_B | \Phi'^{HF}_B \rangle + E^{c(MP2)}_{B,v'}$ is the MP2 energy used to approximate $E^{HK}_{v_B}[\rho_B]$ as:

$$
E^{HK}_{v_B}[\rho_B] \approx \langle \Phi'^{HF}_B | \hat{H}_B | \Phi'^{HF}_B \rangle + E^{c(MP2)}_{v'_B} \tag{10}
$$

At the dissociation limit, both MP2 reference energy and the energy obtained from Eq. 9 are the same provided the electron density of the isolated environment is used as $\rho_B(\mathbf{r})$ and $E^{nad}_{xcT}[\rho_A, \rho_B]$ is approximated by means of any semi-local functional. Any other choice for $\rho_B(\mathbf{r})$ would lead to the increase the energy leaving it as the only factor responsible for the difference between the FDET energy and the corresponding energy obtained by treating the whole complex as one quantum system. For that reason, we will refer to the difference between the energy obtained from Eq. 9 and the reference MP2 energy at any separation between the subsystems as the *error of the FDET energy* and denote this quantity as $\Delta\tilde{E}^{FDET}[\rho_B]$:

$$
\Delta\tilde{E}^{FDET}[\rho_B] = E^{FDET-MP2}_{v_A,v_B} - E^{MP2}_{v_{AB}}. \tag{11}
$$

The above notation reflects the two principal sources error in the total energy evaluated according to Eq. 9: *a*) the violation of the non-negativity condition by the used $\rho_B(\mathbf{r})$ which is indicated by the square brackets, and *b*) approximation given in Eq. 4 indicated by a tilde.

At a finite separation, the contribution to $\Delta E^{FDET}[\rho_B]$ due to the choice of $\rho_B(\mathbf{r})$, which might violate the non-negativity condition, combines with that due to $E_{xcT}^{nad}[\rho_A, \rho_B] \approx \tilde{E}_{xcT}^{nad}[\rho_A, \rho_B]$. The latter can be either positive or negative. We point out here that the error in $\tilde{E}_{xcT}^{nad}[\rho_A, \rho_B]$ affects the total energy in two ways. $\tilde{E}_{xcT}^{nad}[\rho_A, \rho_B]$ is one of the components of the total energy given in Eq. 9. The error in $\tilde{E}_{xcT}^{nad}[\rho_A, \rho_B]$ contributes directly to $\Delta E^{FDET}[\rho_B]$. The error in $\frac{\delta \tilde{E}_{xct}^{nad}[\rho_A,\rho_B]}{\delta \rho_A(\mathbf{r})}$ on the other hand leads to the error in the optimal embedded wave function (and the corresponding density). It contributes, therefore, to $\Delta \tilde{E}^{FDET}[\rho_B]$ indirectly. Comparison of Eqs. 6 and 9 makes it possible to identify also other factors which might affect $\Delta E^{FDET}[\rho_B]$ defined in Eq. 11 at finite separations. We start with the case where $\rho_B(\mathbf{r})$ is taken from MP2 calculations for the Hamiltonian $\hat{H}_B$, i.e. for $\rho_B(\mathbf{r}) = \rho_B^{isol}(\mathbf{r})$). Incomplete cancellation of higher-than second order corrections to the energy for the three Hamiltonians ($\hat{H}_A$, $\hat{H}_B$, and $\hat{H}_{AB}$) contributes also to $\Delta \tilde{E}^{FDET}[\rho_B^{isol}]$. Similarly, the incompleteness of the used basis sets in the three cases contributes also to $\Delta \tilde{E}^{FDET}[\rho_B^{isol}]$. For $\rho_B(\mathbf{r}) \neq \rho_B^{isol}(\mathbf{r})$, another factor adds up. It arises from the fact that the correlation energy used to approximate $E_{v_B}^{HK}[\rho_B]$ is evaluated not for the potential $v_B(\mathbf{r})$ but for the potential $v'_B(\mathbf{r})$, i.e., the potential defining the second auxiliary system which is the sum of $v_B(\mathbf{r})$ and the corresponding embedding potential.

Turning back to the FDET-MP2 expression for the total energy given in Eq. 9, it can be rewritten as:

$$\begin{aligned} E_{v_A,v_B}^{FDET-MP2} &= E_{v_A,v_B}^{FDET-HF} \\ &+ E_{v'_A}^{c(MP2)} + E_{v'_B}^{c(MP2)} + \tilde{E}_k[\Delta\rho_{v'_A}^{c(MP1)}, \rho_A'^{HF}, \rho_B], \end{aligned} \tag{12}$$

where

$$E^{FDET-HF}_{v_A,v_B} = \langle \Phi'^{HF}_A | \hat{H}_A | \Phi'^{HF}_A \rangle + \langle \Phi'^{HF}_B | \hat{H}_B | \Phi'^{HF}_B \rangle + E^{elst,int}_{v_A,v_B}[\rho'^{HF}_A, \rho_B] + \tilde{E}^{nad}_{xcT}[\rho'^{HF}_A, \rho_B]. \quad (13)$$

The above form of this equation makes it evident that the evaluation of the correlation energy and the correlation-induced change of the density is made *a posteriori*, i.e., after the single determinant (either Φ'^{HF}_A or Φ'^{HF}_B) is obtained. The burden of evaluation of the correlation energy is postponed till the self-consistency between the embedded single determinant and embedding potential is reached. At no additional costs, once the FDET-MP2 energies are obtained, the numerical value of $E^{FDET-HF}_{v_A,v_B}$ is also available. They are also discussed in the present work together with the corresponding reference, which is the Hartree-Fock energy of the whole complex. This choice of the reference guarantees that FDET-HF energy evaluated for $\rho_B(\mathbf{r}) = \rho^{isol}_B(\mathbf{r})$ is the same as the reference energy at infinite separation. The attribution to the possible deviation of the $E^{FDET-HF}_{v_A,v_B}$ energy from the reference Hartree-Fock energy is, however, less straightforward than in the FDET-MP2 case. Whereas the correlation effects on energy and density are completely neglected in the reference Hartree-Fock data, they are partially taken into account in $E^{FDET-HF}_{v_A,v_B}$. In FDET-HF, the correlation affects both the total energy (through $\tilde{E}^{nad}_{xcT}[\rho'^{o}_A, \rho_B]$) and the density (through the $\frac{\delta \tilde{E}^{nad}_{xcT}[\rho'^{o}_A,\rho_B]}{\delta \rho_A(\mathbf{r})}$ component of the embedding potential).

2.2 Measure of the magnitude of the violation of the non-negativity condition

Owing to the availability of the reference density obtained from the calculations for the whole complex ($\rho^{o}_{AB}(\mathbf{r})$), makes it possible to quantify the magnitude of the violation of the

non-negativity condition:

$$\forall \mathbf{r}\ \rho^t(\mathbf{r}) > 0. \tag{14}$$

by means of a parameter $M[\rho_t]$ defined as:

$$\begin{aligned} M[\rho_t] &= -\int \rho^t(\mathbf{r}) \cdot \Theta(-\rho^t(\mathbf{r}))\ \mathrm{d}\mathbf{r} \\ &= -\int\limits_{\{\rho^t<0\}} \rho^t, \end{aligned} \tag{15}$$

where Θ is the Heaviside step function. The values of M are bound between 0 and $\int \rho_B = N_B$.

For comparison purposes, the following measure of the overall error of the total density obtained from FDET is used:

$$P[\rho_A^o, \rho_B, \rho_{AB}^o] = \frac{1}{2} \cdot \int |\rho_A^o + \rho_B - \rho_{AB}^o|, \tag{16}$$

.

The $\frac{1}{2}$ in the definition of the parameter P results in the following bonds for M and P (see the Appendix):

$$M[\rho_{AB}^o - \rho_B] \leq P[\rho_A^o, \rho_B, \rho_{AB}^o] \leq \int \rho_{AB}^o = N_{AB}. \tag{17}$$

2.3 Self-consistency of the embedded wavefunction and embedding potential

The basic FDET equality given in Eq. 1, holds only if the optimal embedded wavefunction is consistent with the embedding potential (the latter depends on the embedded wavefunction). Assuring self-consistency involves solving repetitively the quantum N_A-electron problem.

Linearisation of $\tilde{E}_{xcT}^{nad}[\rho_A, \rho_B]$ in $\rho_A(\mathbf{r})$ makes it possible to avoid it.[9]

$$\tilde{E}_{xcT}^{nad}[\rho_A, \rho_B] \approx \tilde{E}_{xcT}^{nad}[\rho_A^{init}, \rho_B] + \Delta^{lin}[\rho_A, \rho_A^{init}, \rho_B], \tag{18}$$

where:

$$\Delta^{lin}[\rho_A, \rho_A^{init}, \rho_B] = \int \left(\rho_A(\mathbf{r}) - \rho_A^{init}(\mathbf{r})\right) \left.\frac{\delta \tilde{E}_{xcT}^{nad}[\rho_A, \rho_B]}{\delta \rho_A(\mathbf{r})}\right|_{\rho_A^{init}} \mathrm{d}\mathbf{r}. \tag{19}$$

Linearisation of $\tilde{E}_{xcT}^{nad}[\rho_A, \rho_B]$ was used in Ref. 9 to deal with the issue of non-orthogonality of embedded wavefunctions for different electronic states. In such a case, $\rho_A^{init}(\mathbf{r})$ corresponds to the ground-state density whereas $\rho_A(\mathbf{r})$ to one of the excited states. In the present work, Eq. 18 is used for another purpose - to accelerate obtaining the self-consistent ground-state embedded wavefunction. In calculations, in which $\rho_B(\mathbf{r})$ was optimised, the freeze-and-thaw iterative procedure was used.[10] In the subsequent iteration of this procedure, the indices A and B exchange their role in the FDET equations. Since the optimised densities $\rho_A(\mathbf{r})$ do not change significantly in the subsequent iterations, the density obtained in a given freeze-and-thaw iteration is used as $\rho_A^{init}(\mathbf{r})$ in the linearized expression for $\tilde{E}_{xcT}^{nad}[\rho_A, \rho_B]$ in the subsequent iteration. Although the embedding potential and the optimal density are not consistent at intermediate steps, the consistency is reached in the final iteration.

2.4 Numerical implementation

$E_{xcT}^{nad}[\rho_A, \rho_B]$ and its functional derivative was approximated using local-density approximation for all its components: Thomas-Fermi[11,12] for the kinetic energy, Dirac-Slater[13] for the exchange energy, and the Vosko-Wilk-Nusair[14] for correlation energy. The aug-cc-pVDZ basis set atomic basis sets were 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 Ψ_A was constructed using only atomic functions centred on the atoms defining the potential $v_A(\mathbf{r})$.

TW to NR: wording of the part below is not clear and must be improved, moreover the

format of the bibitem entries for the citations where was incorrect:

Both the freeze and thaw and the iterative optimization of ρ_A were performed with the author's own version[15] of CCJob.[16] This handled the authomatic submission of Q-Chem5.4[17] calculations.

The integration to obtain $M[\rho^{ref} - \rho_B]$ and $P[\rho_A^o, \rho_B, \rho^{ref}]$ was performed with PySCF.[18] The parsing and collecting of values was performed with the author's version[19] of CCParser[20] and CCDatabase.[21] The plotting in the figures was performed with pandas[22] and matplotlib.[23]

2.5 Explicit treatment of the electronic polarisation of $\rho_B(\mathbf{r})$ 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. At the hearth of this difficulty lies the fact that in FDET, the electronic polarisation of each subsystem is not well-defined. As long as the non-negativity condition is satisfied, any addition of density to $\rho_B(\mathbf{r})$ is not changing the FDET energy. The partition of the total density and consequently the effect of polarisation of $\rho_B(\mathbf{r})$ is, therefore, not unique. In FDET based methods, such that $E_{xcT}^{nad}[\rho_A, \rho_B] \approx \tilde{E}_{xcT}^{nad}[\rho_A, \rho_B]$ optimisation of both $\rho_A(\mathbf{r})$ and $\rho_B(\mathbf{r})$ leads usually to a unique pair. Optimisation of $\rho_B(\mathbf{r})$ corresponds to two effect 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 uniquely attributed to the electronic polarisation of $\rho_B(\mathbf{r})$.

To investigate the relation between the violation of the non-negativity condition end the electronic polarisation, the FDET results are compared three types of $\rho_B(\mathbf{r})$: *i*) $\rho_B(\mathbf{r}) =$

$\rho_B^{isol}(\mathbf{r})$ in which $\rho_B(\mathbf{r})$ is not affected by the environment, *ii*) $\rho_B(\mathbf{r}) = \rho_B^{FAT}(\mathbf{r})$, in which $\rho_B(\mathbf{r})$ is optimised, and *iii*) $\rho_B(\mathbf{r}) = \rho_B^{pp(XX)}(\mathbf{r})$, in which $\rho_B(\mathbf{r})$ is pre-polarised by an external electrostatic field specified by XX. Two pre-polarisation fields were considered. In each case, it was generated by atomic net charges localised by the atoms defining $v_A(\mathbf{r})$ as it is a common practice in multi-level simulation methods. The atomic charges were obtained either from the Mulliken population analysis of $\rho_A(\mathbf{r})$ leading to $\rho_B^{pp(MC)}$ or the ChelPG fitting to the electrostatic potential generated by $\rho_A(\mathbf{r})$ leading to $\rho_B^{pp(ChelPG)}$. For $\rho_B(\mathbf{r}) = \rho_B^{isol}(\mathbf{r})$ the electronic polarisation of $\rho_B(\mathbf{r})$ is not accounted for explicitly although. At finite separation between subsystems, however, the atomic basis sets used for $\rho_A(\mathbf{r})$ makes it possible to affect the total-density also near the atoms defining $v_B(\mathbf{r})$. In FDET terms, this can be interpreted as an implicit treatment of polarisation of the environment in the FDET. ref

The comparative analysis of the three approaches to the electronic polarisation within FDET, are made only for the monomer expansion. this is because... The pre-polarisation by net atomic charges misses the quantum effects due to the Fermi statistics of electrons which leads to unstable numerical results with increasing the used basis sets.[25,26]

3 Results and discussion

3.1 FDET results for optimised $\rho_B(\mathbf{r})$

The FDET results obtained using the supermolecular expansion and the optimised $\rho_B(\mathbf{r})$ are collected in Table 1. The FDET energies obtained with optimised $\rho_B(\mathbf{r})$ are in a very good agreement with the reference values. The largest deviations do not exceed 2.5 kcal/mol (about 10% relative error). Such good performance of the used approximation for $E_{xcT}^{nad}[\rho_A, \rho_B]$ could be expected based on our previous numerical experience[27–29] and is in line with a more completely recent benchmarking of FDET-MP2 energies.[30]

Limiting the basis set expansion to functions centred on atoms in each subsystem lowers the interaction energies reflecting thus the variational principle. Except the Uracil-3H_2O, the

Table 1: FDET-MP2 interaction energies (in kcal/mol) and the integrated negative densities (M in atomic units) obtained with freeze-and-thaw optimised $\rho_B(\mathbf{r})$ and all atomic basis sets (supermolecular expansion). The values of $\Delta\tilde{E}^{FDET}$ are given in parentheses.

	supermolecular expansion		
complex	$M[\rho^{ref} - \rho_B^{FAT}]$	$E_{int}^{FDET-MP2}[\rho_B^{FAT}]$	$E_{int}^{ref\ a}$
7HQ-2MeOH	0.007	-15.02 (2.45)	-17.47
7HQ-formate	0.007	-34.06 (2.42)	-36.48
Uracil-5H2O	0.014	-41.09 (-2.47)	-38.62
PyrBnz-2HCOOH	0.013	-35.92 (1.39)	-36.53

[a] The reference MP2 interaction energies are counterpoint corrected.

Table 2: FDET-MP2 interaction energies (in kcal/mol) and the integrated negative densities (M in atomic units) obtained with freeze-and-thaw optimised $\rho_B(\mathbf{r})$ and reduced number of atomic basis sets (monomer expansion). The values of $\Delta\tilde{E}^{FDET}$ are given in parentheses.

	monomer expansion		
complex	$M[\rho^{ref} - \rho_B^{FAT}]$	$E_{int}^{FDET-MP2}[\rho_B^{FAT}]$	$E_{int}^{ref\ a}$
7HQ-2MeOH	0.013	-13.85 (3.62)	-17.47
7HQ-formate	0.036	-28.58 (7.90)	-36.48
Uracil-5H2O	0.024	-38.39 (0.23)	-38.62
PyrBnz-2HCOOH	0.016	-33.31 (3.22)	-36.53

[a] The reference MP2 interaction energies are counterpoint corrected.

energies are closer to the MP2 reference if the basis sets centred on all atoms of the complex are included (supermolecular expansion). The largest improvement (from 7.90 kcal/mol to 2.42 kcal/mol) occurs for the 7HQ-formate case. The intermolecular bonding has arguably the largest covalent character for this complex. In all cases, the magnitude of the violation of the non-negativity condition (M) is smaller if the full sets of the atomic sets is used. In the 7HQ-formate case, including basis sets localised on all atoms in the

The numerical values of $\Delta\tilde{E}^{FDET}$ and the parameter M obtained using the optimised subsystem densities and the supermolecular expansion collected in Table 1 can be attributed to the used approximation for $E_{xcT}^{nad}[\rho_A, \rho_B] \approx \tilde{E}_{xcT}^{nad}[\rho_A, \rho_B]$. No assumption was made concerning $\rho_B(\mathbf{r})$. In the exact limit ($\tilde{E}_{xcT}^{nad}[\rho_A, \rho_B] \rightarrow E_{xcT}^{nad}[\rho_A, \rho_B]$), $M \rightarrow 0$ and $\Delta\tilde{E}^{FDET} \rightarrow 0$.

3.2 FDET results for arbitrary chosen $\rho_B(\mathbf{r})$

Tables 3 and 4 collect the results obtained for $\rho_B(\mathbf{r}) = \rho_B^{isol}(\mathbf{r})$. The optimisation of $\rho_B(\mathbf{r})$ not only reduces $\Delta\tilde{E}^{FDET}$ by about factor of three but also leads to even larger overall reduction of the parameter M if the supermolecular expansion is used. This indicates clearly that the main contributions to the FDET energies originates rather from the violation of the non-negativity condition than from the used approximation for $E_{xcT}^{nad}[\rho_A, \rho_B]$. Compared to supermolecular expansion, reduction of the number of the centers of atomic sets leads to a slight increase of $\Delta\tilde{E}^{FDET}$ and rather insignificant effect on the parameter M.

Table 3: FDET-MP2 interaction energies (in kcal/mol) and the integrated negative densities (M in atomic units) obtained with $\rho_B(\mathbf{r}) = \rho_B^{isol}(\mathbf{r})$ and all atomic basis sets (supermolecular expansion). The values of $\Delta\tilde{E}^{FDET}$ are given in parentheses.

	supermolecular expansion		
complex	$M[\rho^{ref} - \rho_B^{isol}]$	$E_{int}^{FDET-MP2}[\rho_B^{isol}]$	E_{int}^{ref} [a]
7Hq 2MeOH	0.121	-11.20 (6.27)	-17.47
7Hq formate	0.206	-23.34 (13.14)	-36.48
Uracil 5H2O	0.234	-32.47 (6.15)	-38.62
PyrBnz 2HCOOH	0.184	-27.17 (9.36)	-36.53

[a] The reference MP2 interaction energies are counterpoint corrected.

Table 4: FDET-MP2 interaction energies (in kcal/mol) and the integrated negative densities (M in atomic units) obtained with $\rho_B(\mathbf{r}) = \rho_B^{isol}(\mathbf{r})$ and reduced number of atomic basis sets (monomer expansion). The values of $\Delta\tilde{E}^{FDET}$ are given in parentheses.

	monomer expansion		
complex	$M[\rho^{ref} - \rho_B^{isol}]$	$E_{int}^{FDET-MP2}[\rho_B^{isol}]$	$E_{int}^{ref\ a}$
7Hq 2MeOH	0.123	-10.91 (6.56))	-17.47
7Hq formate	0.205	-23.23 (13.25)	-36.48
Uracil 5H2O	0.237	-32.00 (6.62)	-38.62
PyrBnz 2HCOOH	0.185	-25.98 (10.55)	-36.53

[a] The reference MP2 interaction energies are counterpoint corrected.

TW to NR: check the numbers pls the part in blue below is probably the most important message of the whole paper:

The above results allow us to identify the principal reason for rather excellent performance of FDET in evaluation of the complexation induced shifts of the vertical excitation energies in similar intermolecular complexes reported in Ref. 8 (errors in the order 0.04 eV). The numerical values of $\Delta\tilde{E}^{FDET}$ given in Table 4 were obtained using the the same approximations/assumptions as the ones used in the evaluation of the vertical excitation energies reported in Ref. 8. The errors in the energy of ground states are about one order of magnitude larger than the errors in the vertical excitation excitation energies (6 kcal/mol is equivalent to 0.26 eV).

A similar analysis to that concerning FDET-MP2 made for of the FDET-HF component of the FDET-MP2 energy (and the corresponding Hartree-Fock reference) reveals the trends in line with the ones found for FDET-MP2 (see Supporting Information) and will not be discussed further here.

The fact that optimisation of $\rho_B(\mathbf{r})$ not only reduces the errors in the FDET energies but also leads to lowering the magnitude of violation of the non-negativity condition indicates the link between electronic polarisation of $\rho_B(\mathbf{r})$. The subsequent part concerns a more detailed investigation of this link further aiming at more efficient ways to generate $\rho_B(\mathbf{r})$ than optimising by means of the freeze-and-thaw iterations. We start with noticing that

using all atomic centres (supermolecular expansion) instead of restricted number of centres (monomer expansion) results in much smaller effect on the energy than optimisation of $\rho_B(\mathbf{r})$ in the investigated complexes and with the used approximation for $E_{xcT}^{nad}[\rho_A, \rho_B]$. The subsequent analyses are, therefore, made using data obtained with the monomer expansion.

Figures 1 and 2 show the the magnitude of the violation of the non-negativity condition and the interaction energy for the two choices of $\rho_B(\mathbf{r})$ ($\rho_B^{isol}(\mathbf{r})$ and $\rho_B^{FAT}(\mathbf{r})$) discussed previously together with the data obtained with pre-polarised $\rho_B(\mathbf{r})$. Both methods to treat the electronic polarisation of $\rho_B(\mathbf{r})$ by the electric field generated by the embedded species (field generated by Muliken or ChelPG net atomic charges) result in lowering the magnitude of the violation of the non-negativity condition corresponding to and bring the value of the parameter M closer to that for $\rho_B^{FAT}(\mathbf{r})$ (see Figures 2. Pre-polarisation of $\rho_B(\mathbf{r})$ using the Mulliken net atomic charges brings the interaction energies and the parameter M slightly closer to that corresponding to $\rho_B^{FAT}(\mathbf{r})$. In one case (7HQ-2MeOH), such pre-polarisation of $\rho_B(\mathbf{r})$ results in a negligible effect. Pre-polarisation of $\rho_B(\mathbf{r})$ using the ChelPG net atomic charges leads to a considerably larger effect bringing the interaction energy and the parameter M closer to the values obtained using optimised $\rho_B^{FAT}(\mathbf{r})$. The advantage of using the ChelPG net atomic charges over the Muliken ones could be expected due to the fact that the former ones are fitted to the electrostatic potential generated by the embedded molecule.

TW to NR: I would show only counterpoint corrected data in the Figures. Counterpoint corrections to the interaction energy should be given as a table in SI.

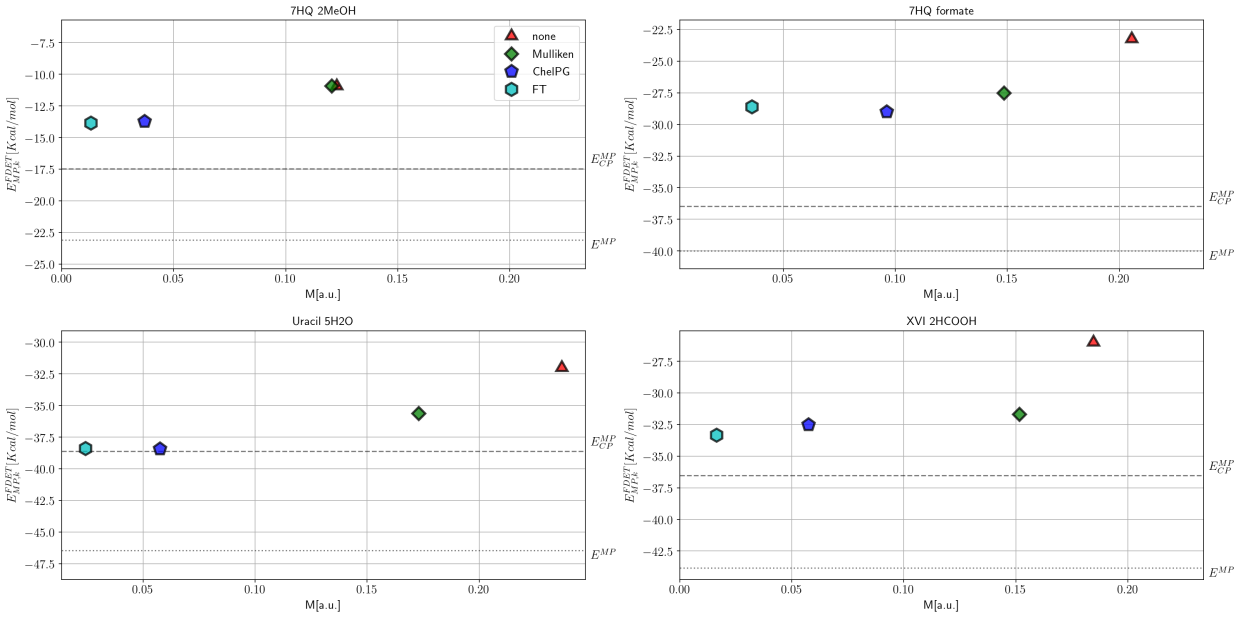

Figure 1: Integrated negative density M and the FDET-MP2 interaction energy for various choices of $\rho_B(\mathbf{r})$: a) $\rho_B^{isol}(\mathbf{r})$ (orange triangles), b) $\rho_B^{FAT}(\mathbf{r})$ (light blue hexagons), c) $\rho_B^{pp(Muliken)}(\mathbf{r})$ (green diamonds), and d) $\rho_B^{pp(ChelPG)}(\mathbf{r})$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

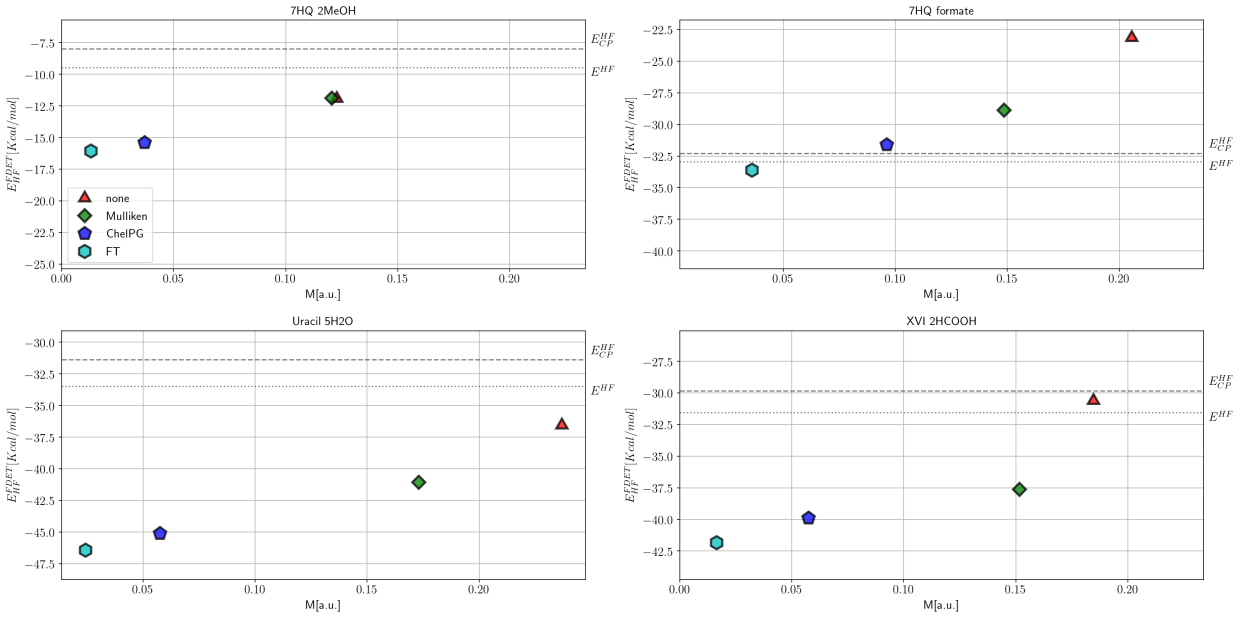

Figure 2: Integrated negative density M and the FDET-HF component of the FDET-MP2 interaction energy for various choices of $\rho_B(\mathbf{r})$: a) $\rho_B^{isol}(\mathbf{r})$ (orange triangles), b) $\rho_B^{FAT}(\mathbf{r})$ (light blue hexagons), c) $\rho_B^{pp(Muliken)}(\mathbf{r})$ (green diamonds), and d) $\rho_B^{pp(ChelPG)}(\mathbf{r})$ (dark blue pentagons). Data obtained using the monomer expansion. Horizontal lies indicate the reference interaction energy.

Data shown in Figure 3 make it possible to turn back to the key issue of the present work, namely, the estimation of the relative significance of the two factors affecting the FDET energies (approximation used for $E_{xcT}^{nad}[\rho_A, \rho_B]$ and the choice for $\rho_B(\mathbf{r})$). We note that the total density obtained with optimised $\rho_B(\mathbf{r})$ is not exact due to the used functional not only derivative of $E_{xcT}^{nad}[\rho_A, \rho_B]$ is not exact. The parameter P defined in Eq. 16 can be used as a measure of the error in density due to approximating this component of the embedding potential. Only with the exact embedding potential and the same basis sets used for FDET and reference calculation the parameter P reaches zero. In all cases, P for $\rho_B^{FAT}(\mathbf{r})$ is much smaller than for any other considered choices for $\rho_B(\mathbf{r})$. The magnitude of the drop of the value of P following the change of $\rho_B(\mathbf{r})$ from $\rho_B^{isol}(\mathbf{r})$ to $\rho_B^{FAT}(\mathbf{r})$ compared to the magnitude of the parameter P corresponding to $\rho_B^{FAT}(\mathbf{r})$ provides this the quantitative measure of the relative significance of these two factors. In all cases, the use of the $\rho_B^{isol}(\mathbf{r})$ leads to larger error in density than the error due to used $E_{xcT}^{nad}[\rho_A, \rho_B]$. The pre-polarisation of the $\rho_B^{isol}(\mathbf{r})$ by the electric field generated by the embedded species greatly reduces the significance of the second factor. Concerning the parameter M obtained for different $\rho_B(\mathbf{r})$, the above trends are even more pronounced. The change of $\rho_B(\mathbf{r})$ from $\rho_B^{isol}(\mathbf{r})$ to $\rho_B^{FAT}(\mathbf{r})$ is accompanied by a significant drop of the parameter M exceeding largely the value of M corresponding to $\rho_B^{FAT}(\mathbf{r})$. TW to NR: the origin of the graph should be at M=0 and P=0 in all cases

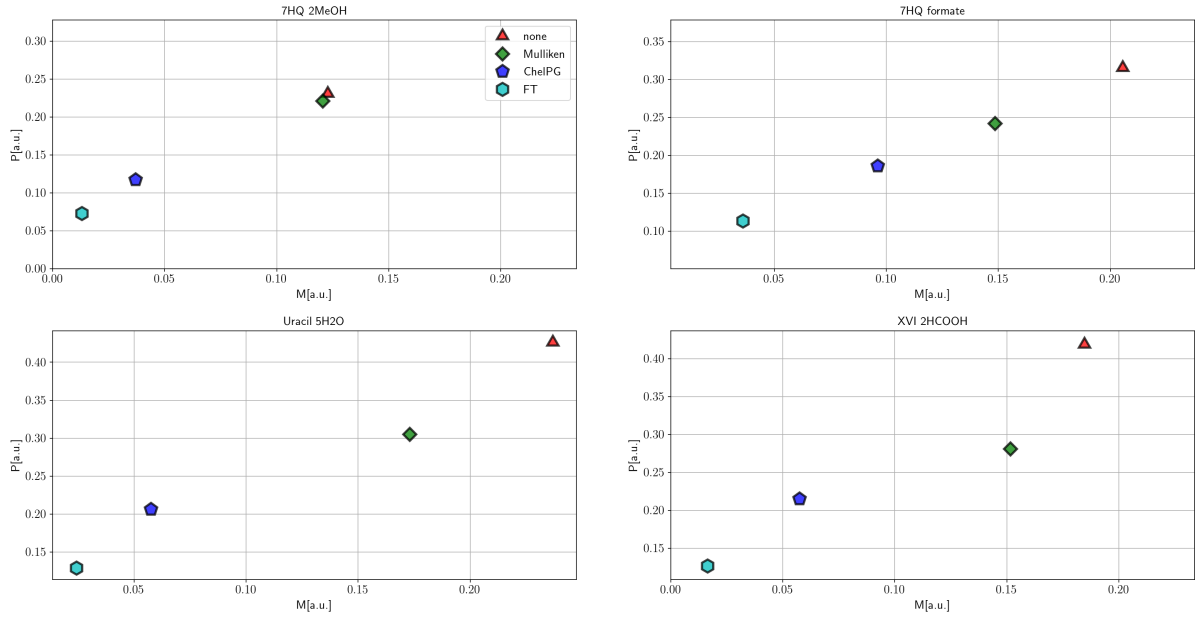

Figure 3: Integrated negative density M and the total density error P for various choices of $\rho_B(\mathbf{r})$: a) $\rho_B^{isol}(\mathbf{r})$ (orange triangles), b) $\rho_B^{FAT}(\mathbf{r})$ (light blue hexagons), c) $\rho_B^{pp(Muliken)}(\mathbf{r})$ (green diamonds), and d) $\rho_B^{pp(ChelPG)}(\mathbf{r})$ (dark blue pentagons). Data obtained using the monomer expansion.

3.3 FDET energy along the freeze-and-thaw optimisation of $\rho_B(\mathbf{r})$

The numerical procedure used to optimise $\rho_B(\mathbf{r})$ yields a series of intermediate densities $\rho_B^{FAT,n}(\mathbf{r})$, providing thus additional set of data for the analysis of the relation between the non-negativity violation and the errors in the energy. N optimisation cycles, provide $\frac{N}{2}$ intermediate densities $\rho_B^{FAT,n}(\mathbf{r})$. Moreover, the intermediate numerical results make it possible to demonstrate the numerical performance of the $\tilde{E}_{xcT}^{nad}[\rho_A, \rho_B]$. In conventional FDET calculations, obtaining the FDET-HF component of the FDET-MP2 energy involves an iterative procedure leading to the optimal single determinant which is consistent with the embedding potential. Linearisation makes it possible to eliminate these iterations. At each even-numbered iteration, i.e., for a given $\rho_B^{FAT,n}(\mathbf{r})$, the FDET-HF energy was evaluated in two ways: *i*) conventionally way, in which additional iterations were used to make the embedding potential and the embedded single determinant consistent, and *ii*) using the linearisation in which the Taylor expansion of $\tilde{E}_{xcT}^{nad}[\rho_A, \rho_B]$ starts at $\rho_A^{FAT,n-1}(\mathbf{r})$ i.e. the

embedding potential was taken from the density obtained in the previous FAT iteration and was not changed in the procedure to optimise the embedded single determinant.

The FDET-HF obtained in two ways are shown in Figure 4 for each even-numbered FAT iteration. The energies are almost indistinguishable which shows great saving potential as far as numerical implementation of the FDET is concerned. The agreement between the energies obtained in the two above ways does not mean that the dependency of the embedding potential on embedded wavefunction is numerically insignificant. On the contrary, it has to be accounted for by $\Delta^{lin}[\rho_A, \rho_A^{init}, \rho_B]$ (see Eq. 19). The magnitude of this term reaches as much as 19 kcal/mol for the 7HQ-formate complex. All the above trends are paralleled by the numerical results obtained using the monomer expansion (see the Supporting Information). TW: to be checked

The magnitude of the $\Delta^{lin}[\rho_A, \rho_A^{init}, \rho_B]$ is systematically smaller than that obtained using the supermolecular expansion which reflects the fact that additional basis sets centres on atoms in the environment results in larger variation of $\rho_A(\mathbf{r})$.

Figure 4 shows also that the FAT iterations converge rapidly. Further optimisation of $\rho_B(\mathbf{r})$ beyond the first one is not needed in practice in the case of the two typical hydrogen-bonded complexes. In the case of more strongly bound complexes (7HQ-formate and PyrBnz-2HCOOH), going beyond the first optimisation of $\rho_B(\mathbf{r})$ affects the total energy only marginally NR: maybe the numbers should be given here - absolute or relative contributions to the interaction energy for instance.

The behaviour of $M[\rho^{ref} - \rho_B^{FT,n}]$ and $P[\rho_A^o, \rho_B, \rho^{ref}]$ along the freeze-and-thaw iterations is very similar to that of the energy (see the Supporting Information).

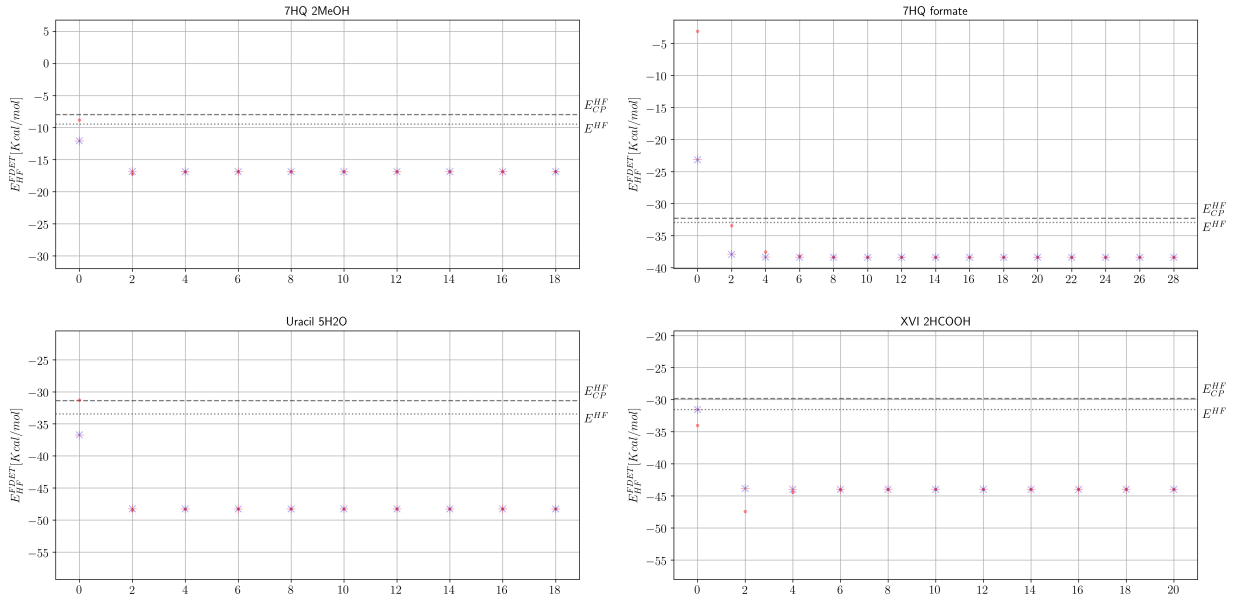

Figure 4: FDET-HF interaction energies for $\rho_B(\mathbf{r})$ obtained for even-numbered FAT iterations and with the supermolecular expansion (blue "X") and their linearised counterparts ("red +"). The energies evaluated with neglected $\Delta^{lin}[\rho_A, \rho_A^{init}, \rho_B]$ in Eq. 19 Eq. 19 (red dots) are shown as red dots.TW to NR: the figures are not good for the journal (resolution), maybe only one is to be shown and the rest to SI?

4 Conclusions

The present work provides the link between the two interpretations of the electronic polarisation of the environment in multi-level simulations. The traditional view based on theory of intermolecular interactions in which the electronic polarisation is one of the component of the energy and the FDET view where its magnitude cannot be evaluated in a unique and straightforward manner. FDET energies are affected by the choice of the electron density $\rho_B(\mathbf{r})$ associated with the environment. This choice leads to the increase of the FDET energy if the chosen $\rho_B(\mathbf{r})$ is locally larger than the exact density of the total system, i.e. if the non-negativity condition is violated. In practical calculations, the effect of the choice of $\rho_B(\mathbf{r})$, which might violate the non-negativity condition, combines with the effect due to the approximation for the functional $E_{xcT}^{nad}[\rho_A, \rho_B]$. Using various measures of the errors in the quantities derived from FDET, it is comprehensively shown that violation of the non-

negativity of the target electron density is the principal source of error in the FDET energies if the the electron density of the isolated environment is used as $\rho_B(\mathbf{r})$. The magnitude of the violation of the non-negativity condition can be greatly reduced by simple techniques to polarise $\rho_B(\mathbf{r})$ available in multi-level simulations.

Concerning the use of the density $\rho_B(\mathbf{r})$ corresponding the isolated environment ($\rho_B^{isol}(\mathbf{r})$) in the FDET based methods, several previous numerical studies showed that it leads to complexation induced shifts of vertical excitation energies of astonishing accuracy. For a representative set of weakly bound intermolecular complexes/clusters, the overall accuracy of the complexation induced shifts of the vertical excitation energy lies in the range of 0.04 eV.[8] In the present work, we looked at the origin of this accuracy by making a detailed analysis of the errors in the energy of one electronic state (ground-state) for four systems taking from the set analysed in Ref. 8. The present study shows that choosing $\rho_B(\mathbf{r}) = \rho_B^{isol}(\mathbf{r})$ leads to errors in the ground-state energy which are about one order of magnitude larger. This indicates that the errors in the energy due to violation of the non-negativity condition for different electronic states cancel to large extend.

In some calculations, the numerical implementation of the used FDET based method made it possible to apply the linearisation technique to deal with the dependency of the embedding potential. The linearisation of $E_{xcT}^{nad}[\rho_A, \rho_B]$ was originally introduced and tested to deal with the $\rho_A(\mathbf{r})$-dependency of the embedding potential in calculations of different electronic states.[9,31] It was shown - for the first time - to be equally robust for the ground-state energy resulting in great possible savings and no loss of accuracy in the total energy.

Supporting Information Available

The data is provided in csv.

- Filename: data as comma separated values. The labels are explained below:

– label: explanation

5 Appendix: relation between $P[\rho_A^o+\rho_B, \rho^{ref}]$ and $M[\rho^{ref}-\rho_B]$

We show below that:

$$M[\rho^{ref} - \rho_B] \leq P[\rho_A^o, \rho_B, \rho^{ref}] \leq \int \rho^{ref} = N_{AB}. \tag{20}$$

Due to the fact that:

$$\int \rho_A^o + \rho_B = \int \rho_o = N_{AB}, \tag{21}$$

where N_{AB} is the number of electrons of the supersystem, the integrated difference of these two densities is zero:

$$\begin{aligned} 0 &= \int \rho_A^o + \rho_B - \rho_o \\ &= \int\limits_{\rho_o<\rho_A^o+\rho_B} \rho_A^o + \rho_B - \rho_o + \int\limits_{\rho_o>\rho_A^o+\rho_B} \rho_A^o + \rho_B - \rho_o \\ &\Leftrightarrow \int\limits_{\rho_o<\rho_A^o+\rho_B} \rho_A^o + \rho_B - \rho_o = \int\limits_{\rho_o>\rho_A^o+\rho_B} \rho_o - \rho_A^o - \rho_B \end{aligned} \tag{22}$$

As a consequence, we can reformulate $P[\rho_A^o, \rho_B, \rho^{ref}]$:

$$\begin{aligned} P[\rho_A^o, \rho_B, \rho^{ref}] =& \frac{1}{2} \cdot \int |\rho_A^o + \rho_B - \rho^{ref}| \\ =& \frac{1}{2} \cdot \int\limits_{\rho_o < \rho_A^o + \rho_B} \rho_A^o + \rho_B - \rho_o + \\ & \frac{1}{2} \cdot \int\limits_{\rho_o > \rho_A^o + \rho_B} \rho_o - \rho_A^o - \rho_B \\ \overset{(22)}{=} & \int\limits_{\rho_o < \rho_A^o + \rho_B} \rho_A^o + \rho_B - \rho_o. \end{aligned} \tag{23}$$

We can then split the integration space and obtain:

$$\begin{aligned} & P[\rho_A^o, \rho_B, \rho^{ref}] = \int\limits_{\rho_o < \rho_B} \rho_A^o + \rho_B - \rho_o \\ & + \int\limits_{\rho_B \leq \rho_o < \rho_B + \rho_A^o} \rho_A^o + \rho_B - \rho_o \\ & \overset{(2)}{=} \int\limits_{\rho_o < \rho_B} \rho_A^o - \rho^t + \int\limits_{\rho_B \leq \rho_o < \rho_B + \rho_A^o} \rho_A^o + \rho_B - \rho_o \\ & \overset{(15)}{=} M[\rho^t] + \int\limits_{\rho_o < \rho_B} \rho_A^o + \\ & \int\limits_{\rho_B \leq \rho_o < \rho_B + \rho_A^o} \rho_A^o + \rho_B - \rho_o \\ & \geq M[\rho^t], \end{aligned} \tag{24}$$

because both integrals on the right hand side of Eq. 24 are non-negative. The upper bound in Eq. 20 is apparent from Eq. 23.

References

(1) Wesolowski, T. A.; Warshel, A. Frozen density functional approach for ab initio calculations of solvated molecules. *J. Phys. Chem.* **1993**, *97*, 8050–8053.

(2) Wesołowski, T. A. Embedding a multideterminantal wave function in an orbital-free environment. *Phys. Rev. A* **2008**, *77*, 012504.

(3) Pernal, K.; Wesolowski, T. A. Orbital-free effective embedding potential: Density-matrix functional theory case. *Int. J. Quantum Chem.* **2009**, *109*, 2520–2525.

(4) Wesolowski, T. A.; Shedge, S.; Zhou, X. Frozen-Density Embedding Strategy for Multilevel Simulations of Electronic Structure. *Chem. Rev.* **2015**, *115*, 5891–5928.

(5) Wesolowski, T. A. On the Correlation Potential in Frozen-Density Embedding Theory. *J. Chem. Theory Comput.* **2020**, *16*, 6880–6885.

(6) 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.

(7) 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. *J. Chem. Phys.* **2010**, *132*, 164101.

(8) 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. *Phys. Chem. Chem. Phys.* **2018**, *20*, 26053–26062.

(9) Zech, A.; Aquilante, F.; Wesolowski, T. A. Orthogonality of embedded wave functions for different states in frozen-density embedding theory. *J. Chem. Phys.* **2015**, *143*, 164106.

(10) Wesolowski, T. A.; Weber, J. Kohn-Sham equations with constrained electron density: an iterative evaluation of the ground-state electron density of interacting molecules. *Chem. Phys. Lett.* **1996**, *248*, 71–76.

(11) Thomas, L. H. The calculation of atomic fields. *Math. Proc. Cambridge Philos. Soc.* **1927**, *23*, 542.

(12) Fermi, E. Eine statistische Methode zur Bestimmung einiger Eigenschaften des Atoms und ihre Anwendung auf die Theorie des periodischen Systems der Elemente. *Z. Phys.* **1928**, *48*, 73–79.

(13) Slater, J. C. The Theory of Complex Spectra. *Phys. Rev.* **1929**, *34*, 1293–1322.

(14) Vosko, S. H.; Wilk, L.; Nusair, M. Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis. *Can. J. Phys.* **1980**, *58*, 1200–1211.

(15) Ricardi, N. CCJob. `https://gitlab.unige.ch/Niccolo.Ricardi/CCJob`, 2021.

(16) Zech, A. CCJob. `https://github.com/spectre007/CCJob`, 2021.

(17) 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 Wergifosse, 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.

(18) 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 Python-based simulations of chemistry framework. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2017**, *8*, e1340.

(19) Ricardi, N. CCParser. `https://github.com/NicoRicardi/CCParser`, 2021.

(20) Zech, A. CCParser. `https://github.com/spectre007/CCParser`, 2021.

(21) Ricardi, N.; González-Espinoza, C. E. CCDatabase. `https://github.com/NicoRicardi/CCDatabase`, 2020.

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

(23) Hunter, J. D. Matplotlib: A 2D Graphics Environment. *Comput. Sci. Eng.* **2007**, *9*, 90–95.

(24) Jeziorski, B.; Moszynski, R.; Szalewicz, K. Perturbation Theory Approach to Intermolecular Potential Energy Surfaces of van der Waals Complexes. *Chem. Rev.* **1994**, *94*, 1887–1930.

(25) 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. *J. Phys. Chem. A* **2011**, *115*, 10018–10026.

(26) 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. *J. Chem. Theory Comput.* **2011**, *7*, 213–222.

(27) Wesolowski, T. A.; Tran, F. Gradient-free and gradient-dependent approximations in the total energy bifunctional for weakly overlapping electron densities. *J. Chem. Phys.* **2003**, *118*, 2072–2080.

(28) Kevorkyants, R.; Dulak, M.; Wesolowski, T. A. Interaction energies in hydrogen-bonded systems: A testing ground for subsystem formulation of density-functional theory. *J. Chem. Phys.* **2006**, *124*, 024104.

(29) Dułak, M.; Wesołowski, T. A. Interaction energies in non-covalently bound intermolecular complexes derived using the subsystem formulation of density functional theory. *J. Mol. Model.* **2007**, *13*, 631–642.

(30) Sen, R.; González-Espinoza, C. E.; Zech, A.; Dreuw, A.; Wesolowski, T. A. Benchmark of the Extension of Frozen-Density Embedding Theory to Nonvariational Correlated Methods: The Embedded-MP2 Case. *Journal of Chemical Theory and Computation* **2021**, *17*, 4049–4062.

(31) Wesolowski, T. A. Embedding potentials for excited states of embedded species. *J. Chem. Phys.* **2014**, *140*, 18A530.

Graphical TOC Entry

TOC to be made. Probably isosurfaces of target density.