

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

Niccolò Ricardi,^{*} Cristina E. González-Espinoza,^{*} and Tomasz Adam Wesolowski^{*}

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

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

Abstract

Tentative abstract

1 Introduction

Frozen-Density Embedding Theory (FDET)^{1,2} is a multiscale method where the environment is described solely by means of its electron density $\rho_B(\mathbf{r})$ while a wavefunction Ψ_A models the system of interest. By virtue of the second Hohenberg-Kohn theorem

$$E_{v_A, v_B}^{FDET}[\Psi_A^o, \rho_B] = E_{v_{AB}}^{HK}[\rho_A^o + \rho_B] \geq E_o, \quad (1)$$

where $E_{v_A, v_B}^{FDET}[\Psi_A, \rho_B]$ is the FDET energy and

$$\Psi_A^o = \arg \min_{\Psi_A} E_{v_A, v_B}^{FDET}[\Psi_A, \rho_B]. \quad (2)$$

[can we avoid defining E^{HK} , ρ_A , and E_o ?] If we define a target density

$$\rho^t(\mathbf{r}) = \rho_o(\mathbf{r}) - \rho_B(\mathbf{r}), \quad (3)$$

it is apparent that the second part of Eq. 1 becomes an equality if and only if $\rho_A^o(\mathbf{r}) = \rho^t(\mathbf{r})$. This, in turn, is only possible if

$$\forall \mathbf{r} \quad \rho^t(\mathbf{r}) > 0. \quad (4)$$

Such condition, which we will name non-negativity of the target density, cannot be en-

forced in practical calculations, where $\rho^o(\mathbf{r})$ is not available. Hence, $E_{v_A, v_B}^{FDET}[\Psi_A^o, \rho_B] > E_o$. There have been examples in the literature where FDET was applied to cases where ρ^t is non-negative everywhere⁴ or where the violation is minor and only due to numerical effects.⁵ There have not been, although, analyses on the effect of the negativity of target densities in practical calculations.

If all the terms composing $E_{v_A, v_B}^{FDET}[\Psi_A^o, \rho_B]$ and E_o are available in their exact form, the only factor inducing an error $\Delta E = E_{v_A, v_B}^{FDET}[\Psi_A^o, \rho_B] - E_o$ is the negativity of the target density. In most practical calculations though, both the FDET and exact energies need to be approximated. The first aim of this manuscript is hence to assess if there is a relation between some measurement of the extent of the violation of the non-negativity of the target density condition (cf. Eq. 4) and $E_{v_A, v_B}^{FDET}[\Psi_A^o, \rho_B]$.

If the case where ρ_B is the density of subsystem B in absence of subsystem A is considered, which will be denoted as ρ_B^{isol} , the theory of molecular interactions indicates that such a density is not properly polarized, and would ascribe any error to such factor. Conversely, ρ_B^{isol} will generally cause violations of the condition of Eq. 4, meaning that within FDET formalism the error will be ascribed to this factor. This suggests the presence of a relation between the

negativity of the target density and polarization. Consequently, within this manuscript the relation between different treatments of polarization and the negativity of the target density will be investigated.

To this end, we analyzed the 4 systems from Ref. 6, selected because they display different strength of interaction, number of molecules in the environment, number of non-covalent interactions, and charge of the environment. Specifically, 7-hydroxyquinoline bound to 2 methanol molecules (7Hq 2MeOH) and uracil bound to 5 water molecules (Uracil 5H₂O) can be considered examples of “standard” interactions, as the groups involved in the hydrogen bonding are neutral hydrogen donors and acceptors. In the cases of 7-hydroxyquinoline bound to formate (7Hq formate) and pyridiniumyl benzimidazolide bound to 2 formic acid molecules (PyrBnz 2HCOOH) represent more peculiar cases. In the former case, the environment acts as the hydrogen acceptor and is negatively charged. Additionally, the hydrogen bond distance (1.36 Å) is comparable to that of its covalent bond (1.09 Å). We have discussed in Ref. 7 how this poses a challenge due to the large overlap of the subsystem densities. In the case of PyrBnz 2HCOOH, subsystem A acts as hydrogen acceptor and the atom involved carries a significant negative charge. Note that throughout this work the naming is given so that the system of interest always precedes the environment.

outline of sections

2 Treatment of polarization

Polarization can be treated differently by using a different external potential in the calculation of the environment density ρ_B . As previously mentioned, one option is to use ρ_B^{isol} , i.e. the density obtained with the nuclear potential $v_B(\mathbf{r})$ as the sole external potential. This constitutes the simplest approach and allows us to relate to the theory of molecular interactions. Another option is to use point-charges to represent the system of interest and add the

corresponding potential to the nuclear potential. This approach, where the point charges are obtained from the isolated density of the system of interest, is referred to as prepolarization. Two different methods were used in this work to obtain the point charges: Mulliken population analysis, leading to ρ_B^{ppMC} , and the ChelPG fitting to the electrostatic potential, leading to ρ_B^{ppEC} . Lastly, an iterative procedure known as freeze and thaw can be used. This consists in switching which subsystem density is optimized (active subsystem, A), and which one is frozen (subsystem B) using the densities from the previous iteration, until the changes in either energy or density fall below a certain threshold. The environment density at the end of such procedure will be denoted by ρ_B^{FT} .

Since FDET is based on the second Hohenberg-Kohn theorem, ρ_A must be optimized with variational methods. This implies that the freeze and thaw procedure can only be performed with variational methods. In order to ensure this, and to ensure consistency of level of theory throughout all the polarization treatment, all our environment densities were computed with the Hartree-Fock(HF) method.

3 Approximate FDET-based methods

In practical calculations, some approximation has to be chosen to obtain $E^{ref} \approx E_o$. Generally, the same approximation is used for \hat{H}_A in the embedding procedure. Furthermore, approximate functionals are used in the calculation of $E_{v_A, v_B}^{FDET}[\Psi_A, \rho_B]$. [potential??] In this manuscript, we used FDET-based second order Møller-Plesset perturbation theory (FDET-MP2), whose details will now be discussed.

The use of MP2 in FDET-based calculations poses two challenges: the fact that, as previously mentioned, FDET applies directly only to variational wavefunctions Ψ_A^{var} , and the modeling of $v_c[\rho_A]$. Ref. 8 provides an expression for the increase of total energy if a single-determinant wavefunction Φ_A is used in lieu of Ψ_A and $v_c[\rho_A]$ is neglected. Based on this, up

to the first order:

$$E_{v_{AB}}^{HK}[\rho_A^o + \rho_B] = E_{v_{A,v_B}}^{FDET}[\Phi_A^o, \rho_B] + E_{v'}^c + E_k[\Delta\rho_{v'}^c, \rho_A^o, \rho_B], \quad (5)$$

where

$$E_k[\Delta\rho_{v'}^c, \rho_A^o, \rho_B] = \int \Delta\rho_{v'}^c(\mathbf{r}) \int \rho_A^o(\mathbf{r}') f_{xcT}^{nad}[\rho_A^o, \rho_B](\mathbf{r}, \mathbf{r}') d\mathbf{r}' d\mathbf{r}, \quad (6)$$

with

$$f_{xcT}^{nad}[\rho_A, \rho_B](\mathbf{r}, \mathbf{r}') = \frac{\delta^2 E_{xcT}^{nad}[\rho_A, \rho_B]}{\delta\rho_A(\mathbf{r})\delta\rho_A(\mathbf{r}')}. \quad (7)$$

The potential v' denotes the embedding potential obtained neglecting $v_c[\rho_A]$. Φ_A^o denotes the single-determinant wavefunction optimized in such a potential, ρ_A^o its corresponding density, and lastly $E_{v'}^c$ and $\Delta\rho_{v'}^c$ its energy and density correlation contributions.

The term $E_{v_{A,v_B}}^{FDET}[\Phi_A^o, \rho_B]$ represents the FDET energy of a single determinant wavefunction optimized within v' ,

$$E_{v_{A,v_B}}^{FDET}[\Phi_A^o, \rho_B] = \langle \Phi_A^o | \hat{H}_A | \Phi_A^o \rangle + E_{v_B}^{HK}[\rho_B] + E_{v_{A,v_B}}^{elst,int}[\rho_A^o, \rho_B] + E_{xcT}^{nad}[\rho_A^o, \rho_B]. \quad (8)$$

In practical calculations the exchange, correlation, and kinetic bifunctionals within $E_{xcT}^{nad}[\rho_A^o, \rho_B]$ is approximated by $\tilde{E}_{xcT}^{nad}[\rho_A^o, \rho_B]$. In the approximate case of MP2, $E_{A,v'}^{(2)}$ is used instead of $E_{v'}^c$, and the first order density correction in v' is used to approximate $\Delta\rho_{v'}^c$. As a result, the FDET-MP2 energy reads:

$$E_{v_{A,v_B}}^{FDET-MP2}[\Phi_A^o, \rho_B] = \langle \Phi_A^o | \hat{H}_A^{HF} | \Phi_A^o \rangle + E_{v_{A,v_B}}^{elst,int}[\rho_A^o, \rho_B] + \tilde{E}_{xcT}^{nad}[\rho_A^o, \rho_B] + \tilde{E}_{v_B}^{HK}[\rho_B] + E_{A,v'}^{(2)} + E_k[\Delta\rho_{v'}^c, \rho_A^o, \rho_B], \quad (9)$$

where the Hohenberg-Kon energy of the frozen density is approximated as $\tilde{E}_{v_B}^{HK}[\rho_B]$, which in our case is an MP2-based approximation of $E_{v_B}^{HK}[\rho_B]$ depending on the specific choice of ρ_B , based on the availability of E_{B,v_B}^{HF} and $E_{B,v_B}^{(2)}$.

In the case of ρ_B^{isol} , $\tilde{E}_{v_B}^{HK}[\rho_B] = E_{B,v_B}^{HF} + E_{B,v_B}^{(2)}$, because both the HF energy and the second order correction in the correct external potential are readily available. It is worth highlighting that in such case, even if the exact correlation energy E_{B,v_B}^c could be used instead of $E_{B,v_B}^{(2)}$, this would still constitute an approximation. This is because $E_{B,v_B}^{HF} + E_{B,v_B}^{(2)} = E_{v_B}^{HK}[\rho_{B,o}]$ which is in turn a lower value than the desired $E_{v_B}^{HK}[\rho_B^{HF}]$. For prepolarized densities, E_{B,v_B}^{HF} is readily available, but $E_{B,v_B}^{(2)}$ is not. $\tilde{E}_{v_B}^{HK}[\rho_B] = E_{B,v_B}^{HF} + E_{B,\tilde{v}_B}^{(2)}$ is used, where $\tilde{v}_B = v_B + v_{PC}$ where the second term is the potential due to the point charges. For ρ_B^{FT} we chose to approximate $E_{v_B}^{HK}[\rho_B]$ with values from an embedding calculation where the active subsystem is switched again, after the freeze and thaw procedure has converged, using some terms of Eq. 9:

$$\tilde{E}_{v_{B,n}}^{HK}[\rho_{B,n}] = E_{A,n+1,v'}^{HF} + E_{A,n+1,v'}^{(2)} + E_k[\Delta\rho_{v'}^c, \rho_{A,n+1}^o, \rho_{B,n+1}]. \quad (10)$$

Since the terms B and A being refer to which system is active, regardless of the cycle, the cycle is denoted as well, with n being the cycle at which convergence was reached.

[fix] and will be compared the MP2 energy of the supersystem E_{AB}^{MP2} , used as E^{ref} .

We also considered the self-consistent part of Eq. 9, which we denoted as FDET-HF and reads:

$$E_{v_{A,v_B}}^{FDET-HF}[\Phi_A^o, \rho_B] = \langle \Phi_A^o | \hat{H}_A^{HF} | \Phi_A^o \rangle + E_{v_{A,v_B}}^{elst,int}[\rho_A^o, \rho_B] + \tilde{E}_{xcT}^{nad}[\rho_A^o, \rho_B] + \tilde{E}_{v_B}^{HF}[\rho_B], \quad (11)$$

where $\tilde{E}_{v_B}^{HF}[\rho_B]$ is the HF component of $\tilde{E}_{v_B}^{HK}[\rho_B]$.

It should be noted that FDET-HF only constitutes the Hartree-Fock component of the FDET-MP2 energy, and not a proper method of its own. The intermonomer correlation - accounted for by $\tilde{E}_{xcT}^{nad}[\rho_A^o, \rho_B]$ is not present in ρ_{AB}^{HF} and E_{AB}^{HF} . As a consequence, FDET-HF as outlined in this manuscript cannot approach the supermolecular Hartree-Fock results.

4 Orphan

The embedding potential depends on ρ_A due to $v_{xcT}^{nad}[\rho_A, \rho_B]$. In the case of freeze and thaw, self-consistency is reached. On the other hand, when ρ_B is not optimized but fixed, as in the other 3 cases in this manuscript, albeit it is technically possible to update the embedding potential at each cycle of the embedded HF self-consistent optimization,⁹ such an implementation is not always available. Two other approaches are possible. The first is an iterative procedure where at cycle i the embedding potential is obtained with $\rho_{A,i-1}^o$, until changes in energy fall below a certain threshold. The second, named linearized FDET, is to do a single calculation where $v_{xcT}^{nad}[\rho_A^{init}, \rho_B]$ uses a user-chosen ρ_A^{init} and a first order Taylor expansion to calculate the energy at the resulting density ρ_A^o . Such approach has proved extremely accurate for excitations.¹⁰ Limited analysis, yet to be published, has been performed for ground-state energies within this approximation. Due to this, except when otherwise mentioned, the iterative optimization of ρ_A has been used.

Since both freeze and thaw and the iterative optimization of ρ_A need to reach self-consistency, they must be performed with variational methods, in our case Hartree-Fock, and the second order energy correction are added afterwards, according to Eq. 9. As a consequence, ρ_{AB}^{HF} is used as the reference density ρ^{ref} , and the target density is approximated as:

$$\rho^t \approx \rho_{AB}^{HF} - \rho_B \quad (12)$$

All the calculations mentioned so far were performed with the so-called monomer expansion (ME), which consists in an expansion for each subsystem where the basis functions are centered only on its nuclei. This causes a localization effect, and prevents ρ_A to delocalize in the region of B , except when very diffuse basis functions are used. Alternatively, the so called supermolecular expansion (SE) can be used, where both subsystems use a basis set expansion where basis functions are centered on each nucleus of the supersystem. The latter allows for a larger degree of delocalization, even

when compact basis functions are used. We performed SE embedding with ρ_B^{isol} and ρ_B^{FT} , where both of these were also obtained with a supermolecular expansion.

4.1 Density descriptors

In order to investigate the relation between negativity and energy, it would be preferable to have a parameter that accounts for the extent of the violation of the condition in Eq. 4. Similarly, the total density from FDET $\rho_A^o + \rho_B$ can be compared to the reference density ρ^{ref} . For both cases, we use a norm of the type:

$$L(f) = \int_{R^3} |f| d\mathbf{r} \quad (13)$$

To estimate the negativity of ρ^t , we decided to integrate the target density only where it is negative, which is equivalent to using $f = \min(0, \rho^t)$, resulting in:

$$\begin{aligned} M[\rho_t] &= - \int \rho^t(\mathbf{r}) \cdot \Theta(-\rho^t(\mathbf{r})) d\mathbf{r} \\ &= - \int_{\{\rho^t < 0\}} \rho^t, \end{aligned} \quad (14)$$

where Θ is the Heaviside step function and the subscript under the integral sign replaces $\Theta(-\rho^t(\mathbf{r}))$ because it restricts the integration space. The resulting values are bound between 0 and $\int \rho_B = N_B$.

Conversely, the use of $f = \rho_A^o + \rho_B - \rho^{ref}$ yields:

$$P[\rho_A^o, \rho_B, \rho^{ref}] = \frac{1}{2} \cdot \int |\rho_A^o + \rho_B - \rho^{ref}|, \quad (15)$$

where the $\frac{1}{2}$ factor has been introduced to ensure that the resulting values are bound according to:

$$M[\rho^{ref} - \rho_B] \leq P[\rho_A^o, \rho_B, \rho^{ref}] \leq \int \rho^{ref} = N_{AB}. \quad (16)$$

The demonstration of the relation in Eq. 16 is available in the Appendix.

5 Computational Details

Both the freeze and thaw and the iterative optimization of ρ_A were performed with the author’s own version¹¹ of CCJob.¹² This handled the automatic submission of Q-Chem5.4¹³ calculations. All calculations were performed with the aug-cc-pVDZ basis set, and the non additive functionals and potentials were all based on the linear density approximation: Thomas-Fermi^{14,15} for the kinetic component, Dirac-Slater¹⁶ for the exchange, and Vosko-Wilk-Nusair¹⁷ for correlation. 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 version¹⁹ of CCParser²⁰ and CCDatabase.²¹ The plotting in the figures was performed with pandas²² and matplotlib.²³

6 Results and discussion

Ψ_A^o is uniquely determined by ρ_B , and consequently so are all quantities derived from Ψ_A^o . Then:

$$E^{FDET}[\rho_B] := E_{v_A, v_B}^{FDET}[\Psi_A^o, \rho_B]. \quad (17)$$

Furthermore, when analyzing numerical examples, it is often easier to compare interaction energies rather than ground-state energies. To this end, we introduce the following notation:

$$\begin{aligned} E_{int}^{ref} &:= E^{ref} - E_A^{(AB)} - E_B^{(AB)} \\ E_{int}^{FDET}[\rho_B] &:= E^{FDET}[\rho_B] - E_A^{(exp)} - E_B^{(exp)}, \end{aligned} \quad (18)$$

Where the superscript within parentheses denotes the centers of the basis set expansion, with (AB) denoting a supermolecular expansion(SE), and (exp) is (AB) for SE calculations and (A) and (B) for the two subsystems respectively for ME calculations. This ensures that both $E_{int}^{FDET}[\rho_B]$ and E_{int}^{ref} are free from basis set superposition error. For E_{int}^{ref} , this is equivalent to the use of a counterpoise correction.²⁴ Unless otherwise mentioned, all results shown concern ME, which localizes the subsystem densities by expanding it with a basis set whose functions are only centered on the subsystem’s atoms (cf.

Section 4). The numerical results will not only be influenced by the negativity of the target density, but also by the aforementioned localization effect and by the approximate nature of $E_{xcT}^{nad}[\rho_A, \rho_B]$ and $v_{xcT}^{nad}[\rho_A, \rho_B]$. We assume that the error contribution of these last two is quasi-constant, which is more likely when considering a single state of a single system where only ρ_B is changed.

Let us start by comparing the different values of integrated negative target density $M[\rho^{ref} - \rho_B]$ and the embedded Hartree-Fock component of the interaction energy $E_{int}^{FDET-HF}[\rho_B]$. This is shown in Fig. 1. Every decrease of $M[\rho^{ref} - \rho_B]$ corresponds to a decrease of $E_{int}^{FDET-HF}[\rho_B]$. It is worth mentioning that no specific relation (be it linear, quadratic, etc...) is to be expected: the relation between electron density and energy is not a straightforward algebraic expression, but an ongoing field of research. It is not even certain if a monotonic relation between $M[\rho^{ref} - \rho_B]$ and $E_{int}^{FDET-HF}[\rho_B]$, which is observed in Fig. 1, is general.

Let us now turn to the relation between polarization and the negativity of the target density. The different types of treatment of polarization reduce $M[\rho^{ref} - \rho_B]$. Prepolarization with Mulliken charges seems to lead to the smallest improvement in integrated negative target density, with almost no improvement in the case of 7Hq 2MeOH. This is consistent with the known issues of Mulliken charges, particularly when diffuse basis sets are used, as in this case. Prepolarization with ChelPG charges leads to a considerably larger improvement, which is consistent with the fact ChelPG charges are fit to the electrostatic potential. Lastly, ρ_B^{FT} leads to the lowest value of $M[\rho^{ref} - \rho_B]$. Such value is not zero because of the inexact embedding potential, in turn caused by the approximate non additive bifunctional $v_{xcT}^{nad}[\rho_A, \rho_B]$ and to the neglect of $v_c[\rho_A]$, and because of the total lack of correlation in ρ^{ref} (cf. Section 3).

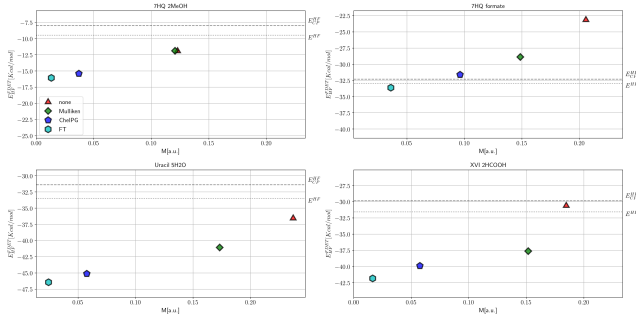


Figure 1: Integrated negative density M versus E_{HF} , for monomer expansion calculations

As one would expect, when $M[\rho^{ref} - \rho_B]$, so does $P[\rho_A^o, \rho_B, \rho^{ref}]$ (cf. Fig. 2). The value for freeze and thaw $P[\rho_A^o, \rho_B^{FT}, \rho^{ref}]$ may be considered an estimate of the quality of the approximations made to obtain the embedding potential. The relation of Eq. 16 is respected.

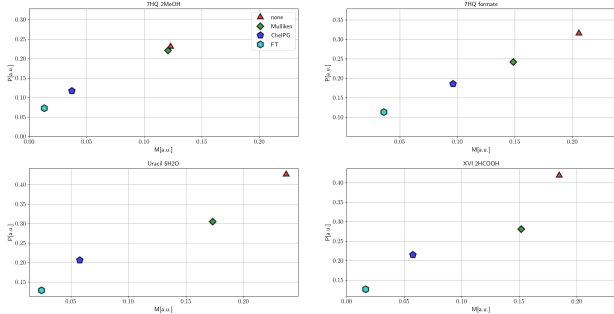


Figure 2: Integrated negative density M versus total density error P , for monomer expansion calculations

Within the results from FDET-MP2, shown in Fig. 3, lower values of $M[\rho^{ref} - \rho_B]$ lead to lower or equal energy, with only one exception. With this level of theory, although, the improvement on the energy seems to reduce in entity at lower values of $M[\rho^{ref} - \rho_B]$, as evidenced by the smaller gap $E_{int}^{FDET-MP2}[\rho_B^{ppEC}] - E_{int}^{FDET-MP2}[\rho_B^{FT}]$ which is always below 0.80 Kcal/mol, and even negative for 7Hq formate. Albeit MP2 is not a variational method, its second order energy and density corrections can be used to estimate the exact FDET energy, as outlined in subsection 3 and in Refs. 8 and 25. With such an approach, $E^{FDET-MP2}[\rho_B^{FT}]$ should be the lowest obtainable embedded energy with a chosen approximation of $v_{xcT}[\rho_A, \rho_B]$. For 7Hq formate this is

not the case. $E^{FDET-MP2}[\rho_B^{FT}]$ is, although, only 0.41 Kcal/mol larger than the lowest value ($E^{FDET-MP2}[\rho_B^{ppEC}]$). This can be ascribed to the fact that Eq. 9 is only an approximate application of Eq. 5. Specifically: a) the second order energy correction is used instead of the full correlation energy both for E_v^c , and $E_{v_B}^{HK}[\rho_B] = E_{B,v_B}^{HF} + E_{v_B}^c$, b) a different external potential is used in at least part of the calculation of the approximate Hohenberg-Kohn energy $\tilde{E}_{v_B}^{HK}$, and such approximation varies between ρ_B^{FT} and ρ_B^{ppEC} , as outlined in Section 3, c) the basis set expansion is incomplete, c) higher than first order terms in the Taylor series leading to Eq. 5 may be playing a role.

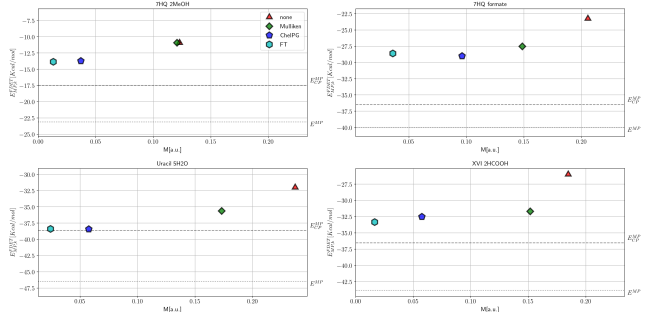


Figure 3: Integrated negative density M versus E_{MP} , for monomer expansion calculations

It is also worth mentioning that, within the monomer expansion, the FDET-MP2 interaction energies from ρ_B^{FT} are rather good, with the exception of 7Hq formate, and that the results from ρ_B^{ppEC} are very similar, as previously mentioned. Both the integrated negative densities $M[\rho^{ref} - \rho_B^{FT}]$ and the interaction energies $E_{int}^{FDET-MP2}[\rho_B^{FT}]$ improve significantly when the supermolecular expansion (SE) is used. By thus allowing full delocalization of the subsystem densities the results are improved to such an extent that the largest integrated negative target density amounts to only 0.014 electrons and the largest energy error amounts to less than 2.5 Kcal/mol (cf. Table 1).

Let us now analyze the convergence during the freeze and thaw procedure: if the optimization converges in N cycles, $\frac{N}{2}$ possible densities $\rho_B^{FT,n}$ and consequently values of $M[\rho^{ref} - \rho_B^{FT,n}]$ are obtained, and it is possible to observe the convergence of various properties and descrip-

Table 1: Monomer (ME) and supermolecular (SE) expansion results for Freeze and Thaw. Energies are in Kcal/mol and integrated densities in a.u.

	Freeze and Thaw				
	ME		SE		
	$M[\rho^{ref} - \rho_B^{FT}]$	$E_{int}^{FDET-MP2}[\rho_B^{FT}]$	$M[\rho^{ref} - \rho_B^{FT}]$	$E_{int}^{FDET-MP2}[\rho_B^{FT}]$	E^{ref}
7Hq 2MeOH	0.013	-13.85	0.007	-15.02	-17.47
7Hq formate	0.036	-28.58	0.007	-34.06	-36.48
Uracil 5H2O	0.024	-38.39	0.014	-41.09	-38.62
PyrBnz 2HCOOH	0.016	-33.31	0.013	-35.92	-36.53

tors along the procedure. Two different strategies were used to obtain at each even-numbered cycle an energy to relate to the integrated negative target density: $E_{int}^{FDET-HF}[\rho_B]$, where ρ_A is iteratively optimized as outlined in Section 4, and linearized FDET ($E_{int}^{linFDET-HF}[\rho_B]$). As mentioned in Section 4, linearized FDET is based on a first order Taylor expansion centered on a user-chosen density ρ_A^{init} . Consequently:

$$E_{xcT}^{nad}[\rho_A, \rho_B] = E_{xcT}^{nad}[\rho_A^{init}, \rho_B] + \Delta^{lin}[\rho_A, \rho_A^{init}, \rho_B], \quad (19)$$

where:

$$\Delta^{lin}[\rho_A, \rho_A^{init}, \rho_B] = \int (\rho_A(\mathbf{r}) - \rho_A^{init}(\mathbf{r})) \left. \frac{\delta E_{xcT}^{nad}[\rho_A, \rho_B]}{\delta \rho_A(\mathbf{r})} \right|_{\rho_A^{init}} d\mathbf{r}. \quad (20)$$

Since we always use $\rho_B^{FT, n-1}$ as ρ_A^{init} , $\Delta^{lin}[\rho_A, \rho_A^{init}, \rho_B]$ will be zero at the end of the freeze and thaw procedure. Fig. 4 shows $E_{int}^{FDET-HF}[\rho_B^{FT, n}]$, $E_{int}^{linFDET-HF}[\rho_B^{FT, n}]$, and $E_{int}^{linFDET-HF}[\rho_B^{FT, n}] - \Delta^{lin}[\rho_A, \rho_A^{init}, \rho_B]$ for the first ** cycles of SE freeze and thaw.

The decrease in self-consistent energy $E_{int}^{FDET-HF}[\rho_B^{FT, n}]$ from the zero-th to the second cycle brings the energy almost to the fully converged value. The behavior of $M[\rho^{ref} - \rho_B^{FT, n}]$ and $P[\rho_A^o, \rho_B, \rho^{ref}]$ is extremely similar to that of the energy, namely after 2-4 cycles values are almost equal to the fully converged ones. Because of this, even if no deviation from monotonicity was observed, figures analogous to 1 and 2 are not particularly insightful, and are made available in the Supporting Information.

The performance of linearized FDET lies outside the analysis of the negativity of the tar-

get density, but is available without performing any additional calculation. Linearized FDET appears to be notably accurate, giving results which are almost indistinguishable from the fully self-consistent ones, for all systems, including 7Hq formate, where $\Delta^{lin}[\rho_A, \rho_A^{init}, \rho_B^{FT, 0}]$ amounts to 19 Kcal/mol. This constitutes the first ground-state extension of the analysis performed for excitations in Ref. 10. Despite $E_{int}^{linFDET-HF}[\rho_B^{FT, 2}]$ being almost equal to the converged value, $\Delta^{lin}[\rho_A, \rho_A^{init}, \rho_B^{FT, 2}]$ may be non-negligible (e.g. 7Hq formate and PyrBnz 2HCOOH). All these observations apply to ME as well, and analogous figures are available in the Supporting Information.

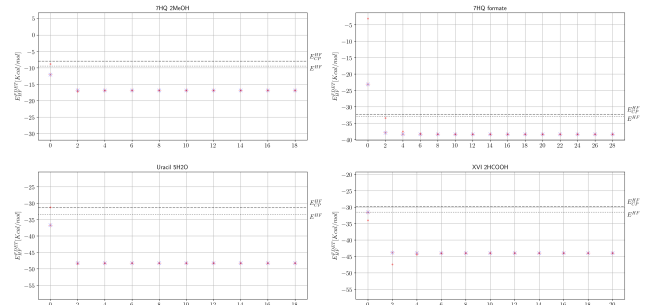


Figure 4: $E^{FDET}[\Psi_A, \rho_B; v_B]$ with fully self-consistent calculations (blue “X”), linearized FDET (red “+”), and not self consistent direct calculation (red dots).

7 Conclusions

By virtue of the second Hohenberg-Kohn theorem, if $\exists \mathbf{r}$ for which $\rho^t(\mathbf{r}) < 0$, then $E_{vAB}^{HK}[\rho_A^o + \rho_B] = E_o + \Delta E > E_o$. No exact relation is known although between the negativity of

ρ^t and ΔE . We have shown how in the selected test cases decreasing values of integrated negative target density leads to a decrease in energy. Prepolarization techniques reduce the integrated negative target density, and consequently lower the ground-state energy.

The negativity of the target density is shown to be a sizeable error contribution. As a consequence, quantifying it with a descriptor such as $M[\rho^t]$ is to be preferred when performing error analysis. Conclusions about the quality of functionals, potentials, or basis set expansions should either be drawn from calculations with the same target density or account for their difference by means of some descriptor such as $M[\rho^t]$.

Similarly, due to the energetic effect of the negativity of the target density, some type of treatment of polarization ought to be preferred when properties pertaining to a single state are concerned. To this end, we have shown how prepolarization with point charges derived from the electrostatic potential is a well performing tradeoff.

8 Perspective and outlook

If excitations are considered, the error on the excitation energy is given by the difference of the energy error in each state. Similarly, this could be compared to the difference of integrated negative target density $M[\rho^{t,ES}] - M[\rho^{t,GS}]$, with *ES* and *GS* denoting the excited- and ground-state respectively. Such differences can be expected to vary less than individual values, rendering the use of ρ_B^{isol} a more reasonable choice for excitation properties than for single-state ones. Similarly, the effect of prepolarization on $M[\rho^{t,ES}] - M[\rho^{t,GS}]$ may be expected to be harder to predict than that on each state.

Acknowledgement The authors thank funding xyz.

Supporting Information Available

The data is provided in csv.

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

– label: explanation

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

We show below that:

$$M[\rho^{ref} - \rho_B] \leq P[\rho_A^o, \rho_B, \rho^{ref}] \leq \int \rho^{ref} = N_{AB}. \quad (21)$$

Due to the fact that:

$$\int \rho_A^o + \rho_B = \int \rho_o = N_{AB}, \quad (22)$$

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

$$\begin{aligned} 0 &= \int \rho_A^o + \rho_B - \rho_o \\ &= \int_{\rho_o < \rho_A^o + \rho_B} \rho_A^o + \rho_B - \rho_o + \int_{\rho_o > \rho_A^o + \rho_B} \rho_A^o + \rho_B - \rho_o \\ &\Leftrightarrow \int_{\rho_o < \rho_A^o + \rho_B} \rho_A^o + \rho_B - \rho_o = \int_{\rho_o > \rho_A^o + \rho_B} \rho_o - \rho_A^o - \rho_B \end{aligned} \quad (23)$$

As a consequence, we can reformulate

$$P[\rho_A^o, \rho_B, \rho^{ref}]:$$

$$\begin{aligned} P[\rho_A^o, \rho_B, \rho^{ref}] &= \frac{1}{2} \cdot \int |\rho_A^o + \rho_B - \rho^{ref}| \\ &= \frac{1}{2} \cdot \int_{\rho_o < \rho_A^o + \rho_B} \rho_A^o + \rho_B - \rho_o + \\ &\quad \frac{1}{2} \cdot \int_{\rho_o > \rho_A^o + \rho_B} \rho_o - \rho_A^o - \rho_B \\ &\stackrel{(23)}{=} \int_{\rho_o < \rho_A^o + \rho_B} \rho_A^o + \rho_B - \rho_o. \end{aligned} \quad (24)$$

We can then split the integration space and obtain:

$$\begin{aligned} P[\rho_A^o, \rho_B, \rho^{ref}] &= \int_{\rho_o < \rho_B} \rho_A^o + \rho_B - \rho_o \\ &+ \int_{\rho_B \leq \rho_o < \rho_B + \rho_A^o} \rho_A^o + \rho_B - \rho_o \\ &\stackrel{(3)}{=} \int_{\rho_o < \rho_B} \rho_A^o - \rho^t + \int_{\rho_B \leq \rho_o < \rho_B + \rho_A^o} \rho_A^o + \rho_B - \rho_o \\ &\stackrel{(14)}{=} M[\rho^t] + \int_{\rho_o < \rho_B} \rho_A^o + \\ &\quad \int_{\rho_B \leq \rho_o < \rho_B + \rho_A^o} \rho_A^o + \rho_B - \rho_o \\ &\geq M[\rho^t], \end{aligned} \quad (25)$$

because both integrals on the right hand side of Eq. 25 are non-negative. The upper bound in Eq. 21 is apparent from Eq. 24.

References

- (1) Wesolowski, T. A. Embedding a multiterminantal wave function in an orbital-free environment. *Phys. Rev. A* **2008**, *77*, 012504.
- (2) Wesolowski, T. A.; Warshel, A. Frozen density functional approach for ab initio calculations of solvated molecules. *J. Phys. Chem.* **1993**, *97*, 8050–8053.
- (3) Aquilante, F.; Wesolowski, T. A. Self-consistency in frozen-density embedding theory based calculations. *J. Chem. Phys.* **2011**, *135*, 084120.
- (4) 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.
- (5) 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.
- (6) 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.
- (7) 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. *J. Chem. Theory Comput.* **2018**, *14*, 4028–4040.
- (8) Wesolowski, T. A. On the Correlation Potential in Frozen-Density Embedding Theory. *J. Chem. Theory Comput.* **2020**, *16*, 6880–6885.
- (9) Dulak, M.; Kaminski, J. W.; Wesolowski, T. A. Linearized orbital-free embedding potential in self-consistent calculations. *Int. J. Quantum Chem.* **2009**, *109*, 1886–1897.
- (10) 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.
- (11) Ricardi, N. CCJob.

- (12) Zech, A. CCJob. <https://github.com/spectre007/CCJob>, 2021.
- (13) 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.; Hernandez 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.; Khistyayev, K.; Kim, J.; Kis, G.; Klunzinger, P.; Koczor-Benda, Z.; Koh, J. H.; Kosenkov, D.; Koulalias, L.; Kowalczyk, T.; Krauter, C. M.; Kue, K.; Kunitsa, A.; Kus, T.; Ladjnszki, 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.; Man-soor, 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.; Mulli-nax, J. W.; Oosterbaan, K. J.; Paron, G.; Paul, A. C.; Paul, S. K.; Pavoevi, 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.; Ras-solov, 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.
- (14) Thomas, L. H. The calculation of atomic fields. *Math. Proc. Cambridge Philos. Soc.* **1927**, *23*, 542.
- (15) 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.

- (16) Slater, J. C. The Theory of Complex Spectra. *Phys. Rev.* **1929**, *34*, 1293–1322.
- (17) 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.
- (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 Pythonbased simulations of chemistry framework. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2017**, *8*, e1340.
- (19) Ricardi, N. CCParser.
- (20) Zech, A. CCParser.
- (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) Boys, S. F.; Bernardi, F. The calculation of small molecular interactions by the differences of separate total energies. Some procedures with reduced errors. *Molecular Physics* **1970**, *19*, 553–566.
- (25) Sen, R.; González-Espinoza, C. E.; Zech, A.; Dreuw, A.; Wesolowski, T. A. Benchmark of the Extension of Frozen-Density Embedding Theory to Non-variational Correlated Methods: The Embedded-MP2 Case. *Journal of Chemical Theory and Computation* **2021**, *17*, 4049–4062.

Graphical TOC Entry

TOC to be made. Probably isosurfaces of target density.