

MEMO from: TW

to: NR+CGE

date: March 29, 2022

subject: The "core" part on the publication on the importance of the violation of the  $\forall_{\mathbf{r}} \rho_B(\mathbf{r}) \leq \rho_o(\mathbf{r})$  condition

Comments:

- Section 2.7

It is the "submission ready presentation of the formalism" Please read carefully the formulas. Typos might still be there.

- Sections 3.1-3.2

They are also "submission ready"

- Section 3.3 – *Good order, content + Just question about FDET-HF energy*

I had no time to work on it so it is practically the same section as the one I prepared in January. There might be some inconsistencies in the notation, typos, etc. Do not pay much attention to it. Look at the contents of the presentation.

- How we proceed: On Wednesday I am completely booked. Thursday might be the day I will have to deal with "current matters" but on Friday I would like to collect your comments (hand-written on this document or specified in a e-mails). If we need to discuss anything it can also be done on Thursday afternoon (difficult for me) or on Friday. I will collect your comment and "reinsert" the present "core" into the draft of a complete manuscript.

## 2.6 Key elements of FDET

This section is not the part of the manuscript. It is given for the sake of completeness of the formulas in the subsequent sections of this MEMO

### 2.6.1 FDET energy functional

$$E_{v_A, v_B}^{FDET}[\Psi_A, \rho_B] = \langle \Psi_A | \hat{H}_A | \Psi_A \rangle + E_{v_B}^{HK}[\rho_B] + E_{v_A, v_B}^{elst, int}[\rho_A, \rho_B] + E_{xcT}^{nad}[\rho_A, \rho_B], \quad (20)$$

where *i*)  $\rho_A(\mathbf{r}) = \langle \Psi_A | \sum_{i=1}^{N_A} \delta(\mathbf{r}_i - \mathbf{r}) | \Psi_A \rangle$ , *ii*)  $E_{v_A, v_B}^{elst, int}[\rho_A, \rho_B]$  collects all classical electrostatic contributions to the interaction energy, and *iii*)  $E_{xcT}^{nad}[\rho_A, \rho_B]$  is defined in Ref. 2, *iv*)  $v_A(\mathbf{r})$  and  $v_B(\mathbf{r})$  are two external potentials, and *v*)  $\hat{H}_A$  is a  $N_A$ -electron Hamiltonian defined by the potential  $v_A(\mathbf{r})$ .

### 2.6.2 Relation between the FDET energy functional ( $E_{v_A, v_B}^{FDET}[\Psi_A, \rho_B]$ ) to the Hohenberg-Kohn energy functional ( $E_{v_{AB}}^{HK}[\rho]$ )

$$\min_{\Psi_A \rightarrow N_A} E_{v_A, v_B}^{FDET}[\Psi_A, \rho_B] = E_{v_A, v_B}^{FDET}[\Psi_A^o, \rho_B] = E_{v_A + v_B}^{HK}[\rho_A^o + \rho_B] \geq E_{v_A + v_B}^o, \quad (21)$$

where  $\rho_A^o(\mathbf{r}) = \langle \Psi_A^o | \sum_{i=1}^{N_A} \delta(\mathbf{r}_i - \mathbf{r}) | \Psi_A^o \rangle$ ,  $\int \rho_B(\mathbf{r}) d\mathbf{r} = N_B$ , and  $E_{v_A + v_B}^o$  is the ground-state energy of the  $N_A + N_B$ -electron system defined by the potential  $v_A(\mathbf{r}) + v_B(\mathbf{r})$ .

### 2.6.3 FDET embedding potential

$$v_{emb}^{FDET}[\rho_A, \rho_B; v_B](\mathbf{r}) = v_B(\mathbf{r}) + \int \frac{\rho_B(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' + \frac{\delta E_{xcT}^{nad}[\rho_A, \rho_B]}{\delta \rho_A(\mathbf{r})}. \quad (22)$$

### 2.6.4 Equation for the stationary embedded wavefunction

$$\left( \hat{H}_A + \hat{v}_{emb}^{FDET}[\rho_A^o, \rho_B; v_B] \right) \Psi_A^o = \lambda^o \Psi_A^o \quad (23)$$

### 2.6.5 Auxiliary $N_A$ -electron system

It is defined by the potential  $v'_A(\mathbf{r})$ :

$$v'_A(\mathbf{r}) = v_A(\mathbf{r}) + v_{emb}^{FDET}[\rho'_A, \rho_B; v_B](\mathbf{r}) \quad (24)$$

where  $\rho'_A(\mathbf{r})$  is the electron density obtained from Eq. 23 if *i*) the wavefunction has the single-determinant form and *ii*) the  $v_c[\rho_A](\mathbf{r})$  component of  $\frac{\delta E_{xct}^{nad}[\rho_A, \rho_B]}{\delta \rho_A(\mathbf{r})}$  is neglected. This component assures that the optimal embedded wavefunction of a single-determinant form satisfies the basic FDET equality (Eq. 21).<sup>2</sup>

## 2.7 FDET energy using <sup>a</sup>non-variational embedded wavefunction

According to the recently derived formula relating the quantities available in methods treating the correlation energy of embedded electrons non-variationally to the Hohenberg-Kohn energy functional (Eq. 38 in Ref. 5):

$$\begin{aligned} E_{v_{AB}}^{HK}[\rho_A^o + \rho_B] &= E_{v_A, v_B}^{FDET}[\Phi_A'^{HF}, \rho_B] + E_{v_A}^c + E_{v_B}^{HK}[\rho_B] \\ &- E_k[\Delta\rho_{v_A}^c, \rho_A'^{HF}, \rho_B] + O(\Delta\rho)^2, \end{aligned} \quad (25)$$

where

$$E_k[\Delta\rho_{v_A}^c, \rho_A'^{HF}, \rho_B] = \int \rho_A'^{HF}(\mathbf{r}) \int \Delta\rho_{v_A}^c(\mathbf{r}') f_{xcT}^{nad}[\rho_A'^{HF}, \rho_B](\mathbf{r}, \mathbf{r}') d\mathbf{r}' d\mathbf{r}, \quad (26)$$

$$\Delta\rho_{v_A}^c(\mathbf{r}) = \rho_{v_A}^o(\mathbf{r}) - \rho_{v_A}'^{HF}(\mathbf{r}) \quad (27)$$

with

$$f_{xcT}^{nad}[\rho_A, \rho_B](\mathbf{r}, \mathbf{r}') = \frac{\delta^2 E_{xcT}^{nad}[\rho_A, \rho_B]}{\delta\rho_A(\mathbf{r})\delta\rho_A(\mathbf{r}')}, \quad (28)$$

where  $\rho_A^o(\mathbf{r})$  is the density <sup>1</sup> defined in Eq. 21,  $E_{v_A, v_B}^{FDET}[\Psi_A', \rho_B]$  is the FDET energy functional given in Eq. 20, whereas the following quantities correspond to an auxiliary system of  $N_A$  electrons in the potential  $v_A'(\mathbf{r})$  defined in Eq. 24:  $\Phi_A'^{HF}$  the optimal single determinant,  $\rho_A'^{HF}(\mathbf{r})$  is the corresponding density,  $\rho_{v_A}^o(\mathbf{r})$  is the exact ground-state electron density, and  $E_{v_A}^c$  is  $v_A'(\mathbf{r})$  the correlation energy. For a given  $\rho_B(\mathbf{r})$ , the FDET interaction energy is given by:

$$E_{int}^{FDET(\rho_B)} = E_{v_{AB}}^{HK}[\rho_A^o + \rho_B] - E_{v_A}^{HK}[\rho_A^{isol}] - E_{v_B}^{HK}[\rho_B^{isol}] \quad (29)$$

---

<sup>1</sup> $\rho_A^o(\mathbf{r})$  is the exact, i.e., correlated electron density. **This footnote will be removed.** I give it here to bring it to Nicco' and Cristina's attention. o↙

Using for  $E_{v_{AB}}^{HK}[\rho_A^o + \rho_B]$  for the right-hand side of Eq. 25 with neglected  $O(\Delta\rho)^2$ , results in:

$$\begin{aligned} E_{int}^{FDET(\rho_B)} &= \langle \Phi_A'^{HF} | \hat{H}_{v_A} | \Phi_A'^{HF} \rangle + E_{v'_A}^c + E_{v_A, v_B}^{elst, int}[\rho_A'^{HF}, \rho_B] + E_{xcT}^{nad}[\rho_A'^{HF}, \rho_B] \\ &- E_k[\Delta\rho_{v'_A}^c, \rho_A'^{HF}, \rho_B] - E_{v_A}^{HK}[\rho_A^{isol}] + E_{v_B}^{HK}[\rho_B] - E_{v_B}^{HK}[\rho_B^{isol}] \end{aligned} \quad (30)$$

In the above equation, each of the quantities:  $\Phi_A'^{HF}$ ,  $\rho_A'^{HF}$ ,  $E_{v'_A}^c$ , and  $\Delta\rho_{v'_A}^c$  depend on  $\rho_B(\mathbf{r})$ . For the clarity of the notation, this dependency is not indicated explicitly. It is implicitly given through  $v'_A$  that depends on  $\rho_B(\mathbf{r})$  (see Eq. 22).

The subsequent sub-sections concerns application of Eq. 30 for different choices for  $\rho_B(\mathbf{r})$ .

### 2.7.1 FDET interaction energy for $\rho_B(\mathbf{r}) = \rho_B^{isol}(\mathbf{r})$

~~FDET interaction energy for  $\rho_B(\mathbf{r}) = \rho_B^{isol}(\mathbf{r})$~~

In this case, the last two terms in Eq. 30 are equal, leading to:

$$\begin{aligned} E_{int}^{FDET(\rho_B^{isol})} &= \langle \Phi_A'^{HF} | \hat{H}_{v_A} | \Phi_A'^{HF} \rangle + E_{v'_A}^c + E_{v_A, v_B}^{elst, int}[\rho_A'^{HF}, \rho_B^{isol}] \\ &+ E_{xcT}^{nad}[\rho_A'^{HF}, \rho_B^{isol}] - E_k[\Delta\rho_{v'_A}^c, \rho_A'^{HF}, \rho_B^{isol}] - \langle \Phi_A'^{HF} | \hat{H}_{v_A} | \Phi_A'^{HF} \rangle - E_{v_A}^c \end{aligned} \quad (31)$$

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

### 2.7.2 FDET interaction energy for $\rho_B^{v''}(\mathbf{r})$ being the ground-state Hartree-Fock density for some potential $v''(\mathbf{r})$

In this case, the numerical value of  $E_{v_B}^{HK}[\rho_B^{v''}] - E_{v_B}^{HK}[\rho_B^{isol}] \geq 0$  contributes to the interaction energy. It approximated as:

$$E_{v_B}^{HK}[\rho_B^{v''}] = \langle \Phi_B''^{HF} | \hat{H}_{v_B} | \Phi_B''^{HF} \rangle + E_{v_B}^c[\rho_B^{v''}] \approx \langle \Phi_B''^{HF} | \hat{H}_{v_B} | \Phi_B''^{HF} \rangle + E_{v_B}^c, \quad (32)$$

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

Using the above approximation in Eq. 30 leads to:

$$\begin{aligned}
E_{int}^{FDET(\rho_B^{v''})} &= \langle \Phi_A'^{HF} | \hat{H}_{v_A} | \Phi_A'^{HF} \rangle + E_{v_A}^c + \langle \Phi_B'^{HF} | \hat{H}_{v_B} | \Phi_B'^{HF} \rangle + E_{v''}^c \quad * \\
&+ E_{v_A, v_B}^{elst, int}[\rho_A'^{HF}, \rho_B^{v''}] + E_{xcT}^{nad}[\rho_A'^{HF}, \rho_B^{v''}] - E_k[\Delta\rho_{v_A}^c, \rho_A'^{HF}, \rho_B^{v''}] \\
&- \langle \Phi_A^{HF} | \hat{H}_{v_A} | \Phi_A^{HF} \rangle - E_{v_A}^c - \langle \Phi_B^{HF} | \hat{H}_{v_B} | \Phi_B^{HF} \rangle - E_{v_B}^c
\end{aligned} \tag{33}$$

where  $\Phi_X^{HF}$  is the Hartree-Fock wavefunction and  $E_{v_X}^c$  denotes the correlation energy in the system defined by the potential  $v_X$  ( $X = A$  or  $B$ ).

### 2.7.3 FDET interaction energy for optimised $\rho_B(\mathbf{r})$


Optimisation of both  $\rho_A(\mathbf{r})$  and  $\rho_B(\mathbf{r})$  proceeds by performing an iterative cycle (*freeze-and-thaw, FAT*) in which the subsystem  $A$  and  $B$  exchange their roles in all FDET equations.<sup>10</sup> In case of freezing  $\rho_A(\mathbf{r})$ ,  $\rho_B$  is represented by means of an embedded  $N_B$  electron single determinant ( $\Phi_B'$ ) that is optimal for a given  $\rho_A$ . Let us denote the quantities obtained at the end of such an optimisation with:  $v_A'^{FAT}$ ,  $\Phi_A'^{FAT}$ ,  $\rho_A'^{FAT}$ ,  $E_{v_A'^{FAT}}^c$ , and  $\Delta\rho_{v_A'^{FAT}}^c$  for the subsystem  $A$  and  $v_B'^{FAT}$ ,  $\Phi_B'^{FAT}$ ,  $\rho_B'^{FAT}$ ,  $E_{v_B'^{FAT}}^c$ , and  $\Delta\rho_{v_B'^{FAT}}^c$  for the subsystem  $B$ . Using this notation, Eq. 33 reads:

$$\begin{aligned}
E_{int}^{FDET(\rho_B'^{FAT})} &= \langle \Phi_A'^{FAT} | \hat{H}_{v_A} | \Phi_A'^{FAT} \rangle + E_{v_A'^{FAT}}^c + \langle \Phi_B'^{FAT} | \hat{H}_{v_B} | \Phi_B'^{FAT} \rangle + E_{v_B'^{FAT}}^c \\
&+ E_{v_A, v_B}^{elst, int}[\rho_A'^{FAT}, \rho_B'^{FAT}] + E_{xcT}^{nad}[\rho_A'^{FAT}, \rho_B'^{FAT}] - E_k[\Delta\rho_{v_A'^{FAT}}^c, \rho_A'^{FAT}, \rho_B'^{FAT}] \\
&- \langle \Phi_A^{HF} | \hat{H}_{v_A} | \Phi_A^{HF} \rangle - E_{v_A}^c - \langle \Phi_B^{HF} | \hat{H}_{v_B} | \Phi_B^{HF} \rangle - E_{v_B}^c
\end{aligned} \tag{34}$$

If the subsystems  $A$  and  $B$  exchange their roles in FDET equations (subsequent *freeze-*

and-thaw iteration), the corresponding expression reads:

$$\begin{aligned}
E_{int}^{FDET(\rho_A'^{FAT})} &= \langle \Phi_B'^{FAT} | \hat{H}_{v_B} | \Phi_B'^{FAT} \rangle + E_{v_B'^{FAT}}^c + \langle \Phi_A'^{FAT} | \hat{H}_{v_A} | \Phi_A'^{FAT} \rangle + E_{v_A'^{FAT}}^c \quad (35) \\
&+ E_{v_B, v_A}^{elst, int}[\rho_B'^{FAT}, \rho_A'^{FAT}] + E_{xcT}^{nad}[\rho_B'^{FAT}, \rho_A'^{FAT}] - E_k[\Delta \rho_{v_B'^{FAT}}^c, \rho_B'^{FAT}, \rho_A'^{FAT}] \\
&- \langle \Phi_B^{HF} | \hat{H}_{v_B} | \Phi_B^{HF} \rangle - E_{v_B}^c - \langle \Phi_A^{HF} | \hat{H}_{v_A} | \Phi_A^{HF} \rangle - E_{v_A}^c
\end{aligned}$$

Following Eq. 25, both above expressions yield  $E_{v_{AB}}^{HK}[\rho_A'^{FAT} + \rho_B'^{FAT}]$  up to the second order terms ( $O(\Delta\rho)^2$ ). Adding Eqs. 34 and 35 and dividing the sum by two yields an expression for the interaction energy that is symmetric upon exchange  $A$  and  $B$ . 

$$\begin{aligned}
E_{int}^{FDET(FAT)} &= \langle \Phi_B'^{FAT} | \hat{H}_{v_B} | \Phi_B'^{FAT} \rangle + E_{v_B'^{FAT}}^c + \langle \Phi_A'^{FAT} | \hat{H}_{v_A} | \Phi_A'^{FAT} \rangle + E_{v_A'^{FAT}}^c \quad (36) \\
&+ E_{v_B, v_A}^{elst, int}[\rho_B'^{FAT}, \rho_A'^{FAT}] + E_{xcT}^{nad}[\rho_B'^{FAT}, \rho_A'^{FAT}] \\
&- \frac{1}{2} \left( E_k[\Delta \rho_{v_A'^{FAT}}^c, \rho_B'^{FAT}, \rho_A'^{FAT}] + E_k[\Delta \rho_{v_B'^{FAT}}^c, \rho_A'^{FAT}, \rho_B'^{FAT}] \right) \\
&- \langle \Phi_B^{HF} | \hat{H}_{v_B} | \Phi_B^{HF} \rangle - E_{v_B}^c - \langle \Phi_A^{HF} | \hat{H}_{v_A} | \Phi_A^{HF} \rangle - E_{v_A}^c
\end{aligned}$$

The above symmetrisation of the expression for the interaction energy is not needed for exact quantities. In practical calculations, Eqs. 34 and 35 might yield different numerical results due to the following factors: *i*) the approximation is used for the functional  $v_{xcT}^{nad}[\rho_A, \rho_B](\mathbf{r})$ , *ii*) approximate treatment of the correlation energy, and *iii*) the use of incomplete basis sets. The final expression assures that the interaction energy is defined in a unique way even if the approximations are used.

## 2.8 FDET with MP2 treatment of correlation for embedded $N_A$ electrons

The interaction energy was evaluated according to Eqs. 31, 33, and 36 applicable for the corresponding choices for  $\rho_B(\mathbf{r})$  using the following approximations:

$$E_{xcT}^{nad}[\rho_A, \rho_B] \approx \tilde{E}_{xcT}^{nad}[\rho_A, \rho_B] \quad (37)$$

$$E_v^c \approx E_c^{MP2} \quad (38)$$

$$\Delta\rho_v^c(\mathbf{r}) \approx \rho_v^{MP1}(\mathbf{r}) - \rho_v^{HF}(\mathbf{r}) \quad (39)$$

For the sake of compactness of the used notation, the methods based on the energy expressions given in this section and the used approximations are referred to as *FDET – MP2* in this work. Other details of the calculations are given in the Computational Details section.

*fine with me.*



## 3 Results and discussion

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

The FDET interaction energies obtained using the optimised  $\rho_B(\mathbf{r})$  and the complete set of atomic basis sets (supermolecular expansion) collected in Table 1 compare well with the reference interaction energies. 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.<sup>27-29</sup> The numerical values of

$$\delta E_{int}^{FAT} = E_{int}^{FDET(FAT)} - E_{int}^{ref} \quad (40)$$

can be attributed mainly to the used approximation for  $E_{xcT}^{nad}[\rho_A, \rho_B]$ .  $\rho_B(\mathbf{r})$  was optimised and the same basis set was used in FDET (embedded MP2) and the reference supermolecular calculations (MP2). Other factors contributing to these deviations include the errors due to the incompleteness of the basis sets and to the higher than second-order contributions to the correlation energy. These effects might not completely cancel each other in FDET and in the reference calculations. The non-zero values of the parameters  $P$  and  $M$ , on the other hand, arise from the used approximation to  $v_{xcT}^{nad}[\rho_A, \rho_B](\mathbf{r})$ . The closer they are to zero the better is the used  $\tilde{v}_{xcT}^{nad}[\rho_A, \rho_B](\mathbf{r})$ . The numerical values of  $P$ ,  $M$ , and  $\delta E_{int}$  given in Table 1 are considered in the subsequent discussion as the reference residual to be compared with the values obtained using additionally approximations especially the ones concerning  $\rho_B(\mathbf{r})$ .

Some analyses in the present work are made using a reduced set of atomic basis sets (monomer expansion). This adds an additional factor contributing to the deviations from the reference MP2 shown in Table 1. For the sake of meaningful comparisons, the corresponding residual deviations are collected in Table 2.

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<sub>2</sub>O,

*increases*  
*or do you mean the error ??*

Table 1: Deviations of the FDET-MP2 results from the reference data. In FDET, freeze-and-thaw optimised  $\rho_B(\mathbf{r})$  and the reduced set of atomic basis sets (supermolecular expansion) are used. Density measures  $M$  and  $P$  are given in atomic units, energies in kcal/mol.

complex	$P^a$	$M^b$	$\delta E_{int}^c$	$E_{int}^{FDET(FAT)}^d$	$E_{int}^{ref}^e$
7HQ-2MeOH	0.059	0.007	2.45	-15.02	-17.47
7HQ-formate	0.066	0.007	2.42	-34.06	-36.48
Uracil-5H2O	0.114	0.014	-2.47	-41.09	-38.62
PyrBnz-2HCOOH	0.104	0.013	1.39	-35.92	-36.53

$$^a M = M[\rho^{ref} - \rho_B^{FDET(FAT)}] \text{ with } M[\rho] \text{ defined in Eq. 15}$$

$$^b P = P[\rho^{ref} - \rho_{tot}^{FDET(FAT)}] \text{ with } P[\rho] \text{ defined in Eq. 16}$$

$$^c \delta E_{int} = E_{int}^{FDET(FAT)} - E_{int}^{ref}$$

$$^d E_{int}^{FDET(FAT)} \text{ is given in Eq. 36}$$

$$^e \text{ The reference MP2 interaction energies } (E_{int}^{ref}) \text{ are counterpoint corrected.}$$

Table 2: Deviations of the FDET-MP2 results from the reference data. In FDET, freeze-and-thaw optimised  $\rho_B(\mathbf{r})$  and the reduced set of atomic basis sets (monomer expansion) are used. Density measures  $M$  and  $P$  are given in atomic units, energies in kcal/mol.

complex	$P^a$	$M^b$	$\delta E_{int}^c$	$E_{int}^{FDET(FAT)}^d$	$E_{int}^{ref}^e$
7HQ-2MeOH	0.073	0.013	3.62	-13.85	-17.47
7HQ-formate	0.114	0.036	7.90	-28.58	-36.48
Uracil-5H2O	0.129	0.024	0.23	-38.93	-38.62
PyrBnz-2HCOOH	0.127	0.016	3.22	-33.31	-36.53

$$^a M = M[\rho^{ref} - \rho_B^{FDET(FAT)}] \text{ with } M[\rho] \text{ defined in Eq. 15}$$

$$^b P = P[\rho^{ref} - \rho_{tot}^{FDET(FAT)}] \text{ with } P[\rho] \text{ defined in Eq. 16}$$

$$^c \delta E_{int} = E_{int}^{FDET(FAT)} - E_{int}^{ref}$$

$$^d E_{int}^{FDET(FAT)} \text{ is given in Eq. 36}$$

$$^e \text{ The reference MP2 interaction energies } (E_{int}^{ref}) \text{ are counterpoint corrected.}$$

is this correct? labels are reversed?

if the basis sets of all atoms are included

the energies are closer to the MP2 reference if the basis sets centred on all atoms of the complex are included (supermolecular expansion). The largest numerical effect on energy due to such reduction of the number of the basis sets used is found in the 7HQ-formate case (the error increases from 2.42 to 7.90 kcal/mol).<sup>or</sup> 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 set~~s~~ of the atomic sets is used.

In practice, For the multi-scale numerical simulations practice, optimising  $\rho_B(\mathbf{r})$  by means of the freeze-and-thaw optimisation is rather not realistic. Such simulations hinge on the possibility to generate  $\rho_B(\mathbf{r})$  in a simpler/less costly way. The subsequent sections concern such uses of FDET. To this end, the ~~the~~ deviations from the reference data will be analysed in presence of additional assumptions concerning  $\rho_B(\mathbf{r})$  obtained in alternative ways.

### 3.2 FDET results without optimisation of $\rho_B$ : $\rho_B(\mathbf{r}) = \rho_B^{isol}(\mathbf{r})$

We start with  $\rho_B(\mathbf{r})$  obtained from calculations for isolated subsystem B ( $\rho_B(\mathbf{r}) = \rho_B^{isol}(\mathbf{r})$ ) collected in Tables 3 and 4 .

Table 3: Deviations of the FDET-MP2 results from the reference data. In FDET,  $\rho_B(\mathbf{r}) = \rho_B^{isol}(\mathbf{r})$  and the complete set of atomic basis sets (supermolecular expansion) are used. Density measures  $M$  and  $P$  are given in atomic units, energies in kcal/mol

<sup>e</sup>  $E_{int}^{FDET(\rho_B=\rho_B^{isol})}$  is given in Eq. 31.

complex	$P^a$	$M^b$	$\delta E_{int}^c$	$E_{int}^{FDET(\rho_B=\rho_B^{isol})}$	$d E_{int}^{ref}{}^e$
7HQ-2MeOH	0.228	0.121	6.27	-11.20	-17.47
7HQ-formate	0.316	0.206	13.14	-23.34	-36.48
Uracil-5H2O	0.426	0.234	6.15	-32.47	-38.62
PyrBnz-2HCOOH	0.416	0.184	9.36	-27.17	-36.53

<sup>a</sup>  $M = M[\rho^{ref} - \rho_B^{isol}]$  with  $M[\rho]$  defined in Eq. 15

<sup>b</sup>  $P = P[\rho^{ref} - \rho_{tot}^{FDET(\rho_B=\rho_B^{isol})}]$  with  $P[\rho]$  defined in Eq. 16

<sup>c</sup>  $\delta E_{int} = E_{int}^{FDET(\rho_B=\rho_B^{isol})} - E_{int}^{ref}$

<sup>d</sup>  $E_{int}^{FDET(\rho_B=\rho_B^{isol})}$  is given in Eq. 31

<sup>e</sup> The reference MP2 interaction energies ( $E_{int}^{ref}$ ) are counterpoint corrected.

We start comparing the FDET results obtained using optimised and not optimised  $\rho_B(\mathbf{r})$

Table 4: Deviations of the FDET-MP2 results from the reference data. In FDET,  $\rho_B(\mathbf{r}) = \rho_B^{isol}(\mathbf{r})$  and the reduced set of atomic basis sets (monomer expansion) are used. Density measures  $M$  and  $P$  are given in atomic units, energies in kcal/mol.

complex	$P^a$	$M^b$	$\delta E_{int}^c$	$E_{int}^{FDET(\rho_B=\rho_B^{isol})}{}^d$	$E_{int}^{ref}{}^e$
7HQ-2MeOH	0.231	0.123	6.56	-10.91	-17.47
7HQ-formate	0.316	0.205	13.25	-23.23	-36.48
Uracil-5H <sub>2</sub> O	0.427	0.237	6.62	-32.00	-38.62
PyrBnz-2HCOOH	0.419	0.185	10.55	-25.98	-36.53

$$^a M = M[\rho^{ref} - \rho_B^{isol}] \text{ with } M[\rho] \text{ defined in Eq. 15}$$

$$^b P = P[\rho^{ref} - \rho_{tot}^{FDET(\rho_B=\rho_B^{isol})}] \text{ with } P[\rho] \text{ defined in Eq. 16}$$

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

$$^d E_{int}^{FDET(\rho_B=\rho_B^{isol})} \text{ is given in Eq. 31}$$

$$^e \text{ The reference MP2 interaction energies } (E_{int}^{ref}) \text{ are counterpoint corrected.}$$

with the observation that neglecting the basis sets localised on a "partner" molecule does not affect significantly the deviations in density and in energy reported in Tables 3 and 4. On the other hand the numerical values of the deviations increase significantly if  $\rho_B(\mathbf{r})$  is not optimised. Deviations of the energy and the parameter  $P$  increase usually by a factor of about 2 to 3, whereas the  $P$  measuring the extend of the violation of the non-additivity of the target density by the used  $\rho_B(\mathbf{r})$  increases one order of magnitude more (see Tables 1 and Tables 3). This observations indicate clearly that the errors of the FDET results originate mainly due to the violation of the non-additivity condition and not to the choice of the used approximation for  $E_{xcT}^{nad}[\rho_A, \rho_B]$  and  $v_{xcT}^{nad}[\rho_A, \rho_B](\mathbf{r})$ .

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

### 3.3 Relation between non-negativity violation and the polarisation of $\rho_B(\mathbf{r})$

The fact that optimisation of  $\rho_B(\mathbf{r})$  not only reduces the errors in the FDET energies but also leads to lowering of the magnitude of violation of the non-negativity condition indicates the link between electronic polarisation of  $\rho_B(\mathbf{r})$ . The subsequent part concerns this link and aims 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 a much smaller effect on the energy than the optimisation of  $\rho_B(\mathbf{r})$  in the investigated complexes and with the used approximation for  $E_{xcT}^{nad}[\rho_A, \rho_B]$ . The FDET results discussed below were obtained using the monomer expansion, therefore.

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 polarise  $\rho_B(\mathbf{r})$  by the electric field generated by the embedded species (field generated by Mulliken or ChelPG net atomic charges) result in lowering the magnitude of the violation of the non-negativity condition. ~~They bring the value of the parameter  $M$  closer to that for  $\rho_B^{FAT}(\mathbf{r})$~~  (see Figure 2). 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 over the Mulliken charge representation of  $\rho_A(\mathbf{r})$  in order to polarise  $\rho_B(\mathbf{r})$  could be expected. The former ones are fitted to the electrostatic potential generated by the embedded molecule and represent better the electric field at long range.

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

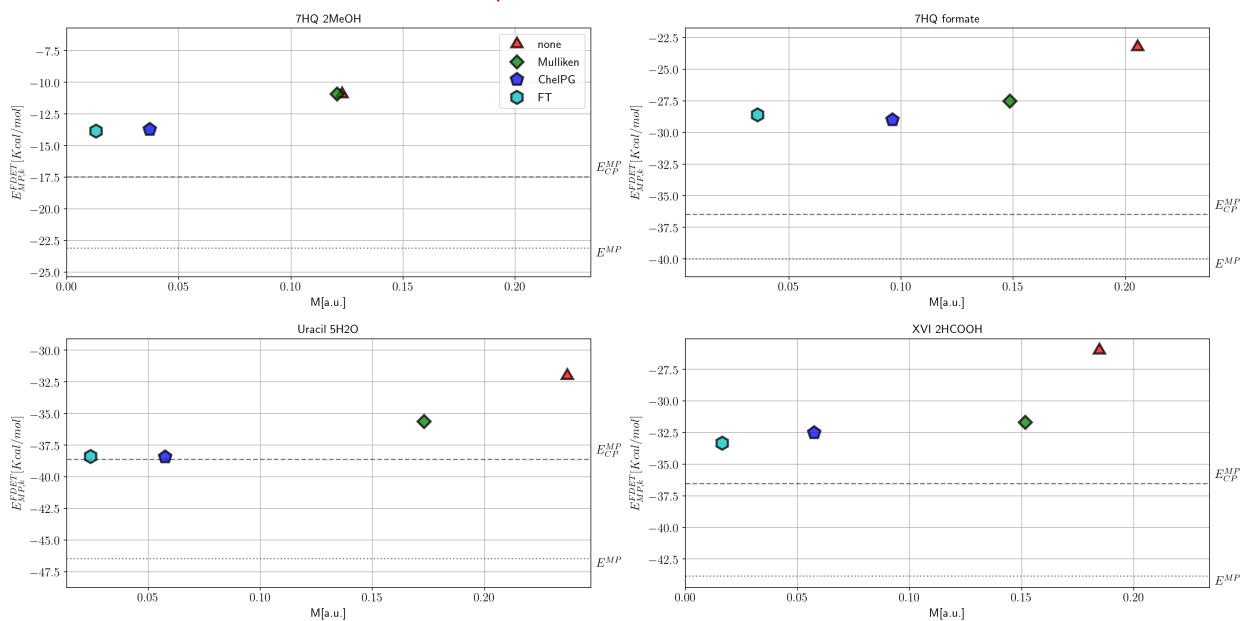


Figure 1: Integrated negative density  $M$  and the FDET-MP2 interaction energy for various choices of  $\rho_B(\mathbf{r})$ : a)  $\rho_B^{isol}(\mathbf{r})$  (orange triangles), b)  $\rho_B^{FAT}(\mathbf{r})$  (light blue hexagons), c)  $\rho_B^{pp(Muliken)}(\mathbf{r})$  (green diamonds), and d) 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

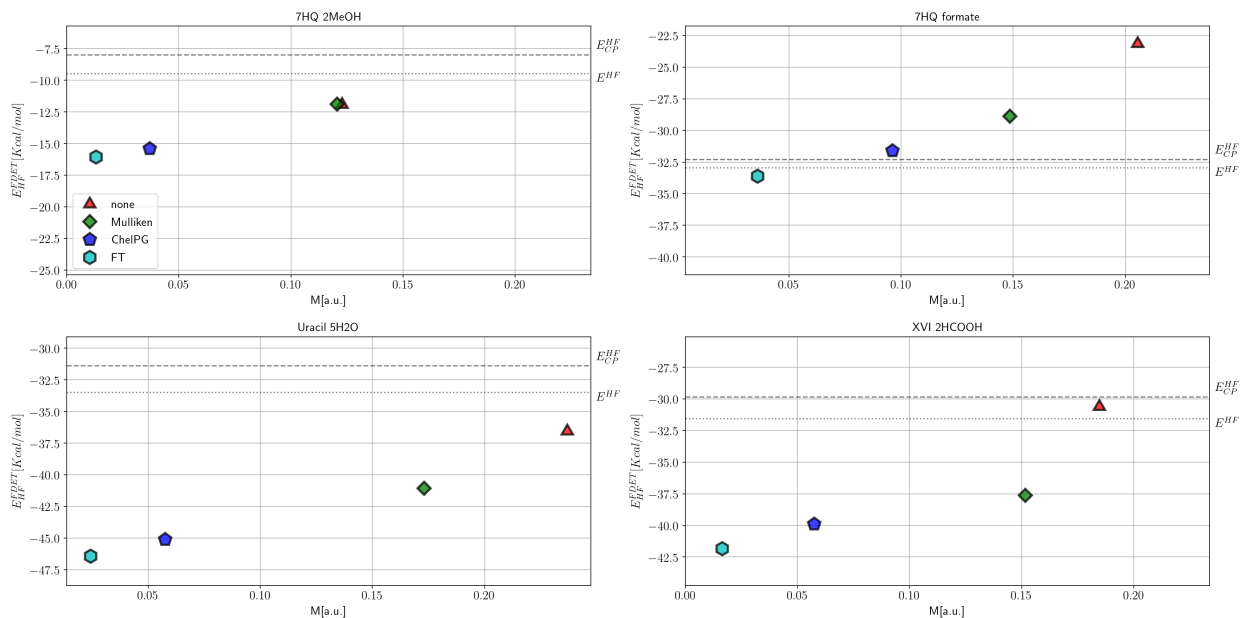


Figure 2: Integrated negative density  $M$  and the FDET-HF component of the FDET-MP2 interaction energy for various choices of  $\rho_B(\mathbf{r})$ : a)  $\rho_B^{isol}(\mathbf{r})$  (orange triangles), b)  $\rho_B^{FAT}(\mathbf{r})$  (light blue hexagons), c)  $\rho_B^{pp(Muliken)}(\mathbf{r})$  (green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

Do we want to discuss this energy??, it has not been presented

TW to NR: the origin of the graph should be at M=0 and P=0 in all cases

*larger font, markers*

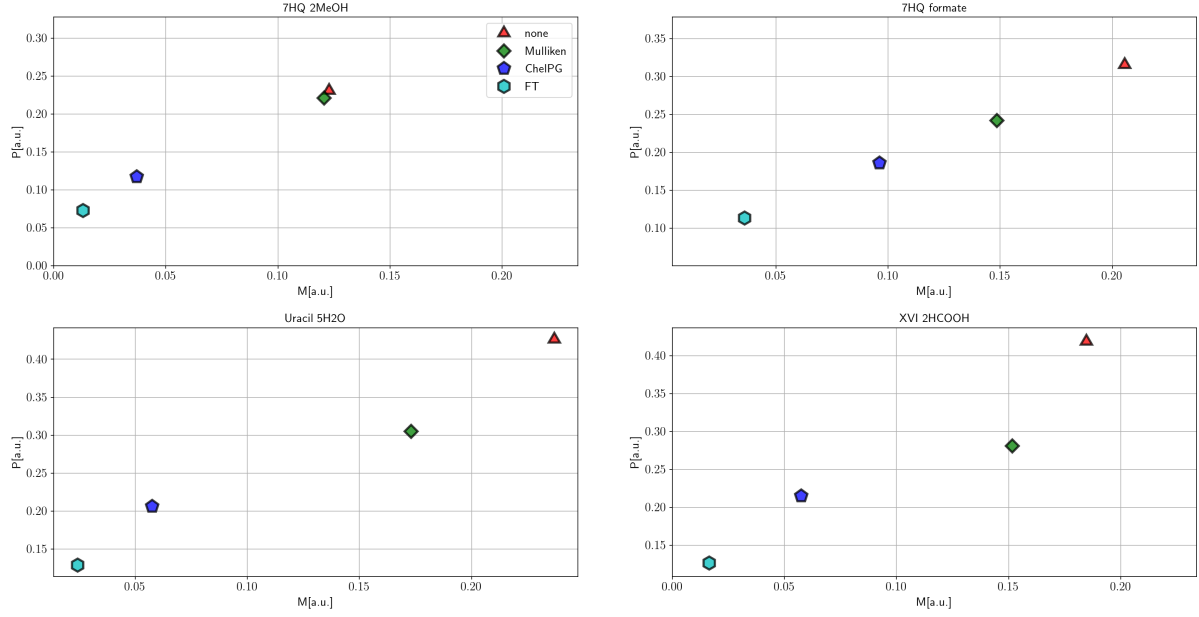


Figure 3: Integrated negative density  $M$  and the total density error  $P$  for various choices of  $\rho_B(\mathbf{r})$ : a)  $\rho_B^{isol}(\mathbf{r})$  (orange triangles), b)  $\rho_B^{FAT}(\mathbf{r})$  (light blue hexagons), c)  $\rho_B^{pp(Mulliken)}(\mathbf{r})$  (green diamonds), and d)  $\rho_B^{pp(ChelPG)}(\mathbf{r})$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