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

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

*Department of Physical Chemistry, University of Geneva,*

*Geneva (Switzerland)*

(Dated: 16 June 2022)

---

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

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

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

# S1. CONSTRAINED SEARCH DEFINITIONS OF DENSITY FUNCTIONALS USED IN FDET

## A. Functionals in Kohn-Sham formulation of DFT

$$T_s[\rho] = \min_{\Phi \longrightarrow \rho} \left\langle \Phi|\hat{T}|\Phi \right\rangle = \left\langle \Phi^o[\rho]|\hat{T}|\Phi^o[\rho] \right\rangle = T_s[\rho] \tag{S1}$$

$$E_{xc}[\rho] = \min_{\Psi \longrightarrow \rho} \left\langle \Psi|\hat{T} + \hat{V}^{ee}|\Psi \right\rangle - T_s[\rho] - J[\rho] \tag{S2}$$
$$= \left\langle \Psi^o[\rho]|\hat{T} + \hat{V}^{ee}|\Psi^o[\rho] \right\rangle - T_s[\rho] - J[\rho],$$

where $\hat{V}_{ee} = \sum_{i=1}^{N} \sum_{j>i}^{N} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}$ is the electron-electron repulsion operator, $J[\rho] = \frac{1}{2} \int \int \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}|} d\mathbf{r}' d\mathbf{r}$, $\Phi$ and $\Psi$ are $N$-electron wave-functions of the single-determinant or Full Configuration Interaction form, respectively.

## B. Functionals used in FDET

For a) embedded wavefunctions of the Full Configuration Interaction form, b) embedded non-interacting reference (Kohn-Sham-like) system, and c) embedded one-particle spinless reduced density matrix, FDET uses the non-additive kinectic-exchange-correlation bifunctional defined as:

$$E_{xct}^{nad}[\rho_A, \rho_B] \equiv E_{xc}[\rho_A + \rho_B] - E_{xc}[\rho_A] - E_{xc}[\rho_B] \tag{S3}$$
$$+ T_s[\rho_A + \rho_B] - T_s[\rho_A] - T_s[\rho_B].$$

For embedded wavefunctions of truncated Configuration Interaction form, FDET uses another functional:

$$E_{xct}^{nad(truncCI)}[\rho_A, \rho_B] \equiv E_{xct}^{nad}[\rho_A, \rho_B] + E_c^{truncCI}[\rho] \leq E_{xct}^{nad}[\rho_A, \rho_B], \tag{S4}$$

where

$$E_c^{truncCI}[\rho] = \min_{\Psi \longrightarrow \rho} \left\langle \Psi|\hat{T} + \hat{V}^{ee}|\Psi \right\rangle - \min_{\Psi^{truncCI} \longrightarrow \rho} \left\langle \Psi^{truncCI}|\hat{T} + \hat{V}^{ee}|\Psi^{truncCI} \right\rangle \tag{S5}$$
$$= \left\langle \Psi^o[\rho]|\hat{T} + \hat{V}^{ee}|\Psi^o[\rho] \right\rangle - \left\langle \Psi^{o(truncCI)}[\rho]|\hat{T} + \hat{V}^{ee}|\Psi^{o(truncCI)}[\rho] \right\rangle,$$

If the embedded wavefunction has a single-determinant form:

$$E_{xct}^{nad(SD)}[\rho_A, \rho_B] \equiv E_{xct}^{nad}[\rho_A, \rho_B] + E_c[\rho] \leq E_{xct}^{nad}[\rho_A, \rho_B], \tag{S6}$$

where

$$E_c[\rho] = \min_{\Psi \longrightarrow \rho} \left\langle \Psi | \hat{T} + \hat{V}^{ee} | \Psi \right\rangle - \min_{\Phi \longrightarrow \rho} \left\langle \Phi | \hat{T} + \hat{V}^{ee} | \Phi \right\rangle \tag{S7}$$
$$= \left\langle \Psi^o[\rho] | \hat{T} + \hat{V}^{ee} | \Psi^o[\rho] \right\rangle - \left\langle \Phi^o[\rho] | \hat{T} + \hat{V}^{ee} | \Phi^o[\rho] \right\rangle,$$

is the correlation functional defined also in the Kohn-Sham formulation of DFT.

Using variationally obtained embedded single determinant AND neglecting $E_c[\rho]$ leads to the violation of the basic FDET equality relating FDET energy functional to the Hohenberg-Kohn functional (Eq. 2) leading to the increase of the energy. The energy consistent with the Hohenberg-Kohn functional correcting this increase is given in Eq. 6.

## S2.   PROOF THAT $M[\rho_B - \rho_v] \leq P[\rho_A + \rho_B - \rho_v^o]$

We start with the obvious equality,

$$\int \rho_A + \rho_B - \rho_v^o = 0, \tag{S8}$$

from which it follows that:

$$\int\limits_{\rho_v^o < \rho_A + \rho_B} \rho_A + \rho_B - \rho_v^o = \int\limits_{\rho_v^o > \rho_A + \rho_B} \rho_v^o - \rho_A - \rho_B. \tag{S9}$$

The above relation used in the definition of $P$ leads to:

$$P[\rho_A + \rho_B - \rho_v^o] = \frac{1}{2} \int |\rho_A + \rho_B - \rho_v^o| \tag{S10}$$
$$= \frac{1}{2} \int\limits_{\rho_v^o < \rho_A + \rho_B} \rho_A + \rho_B - \rho_v^o +$$
$$\frac{1}{2} \int\limits_{\rho_v^o > \rho_A + \rho_B} \rho_A + \rho_B - \rho_v^o$$
$$= \int\limits_{\rho_v^o < \rho_A + \rho_B} \rho_A + \rho_B - \rho_v^o.$$

Splitting the domain of integration of the final integral above leads to:

$$P[\rho_A + \rho_B - \rho_v^o] = \int\limits_{\rho_v^o < \rho_B} \rho_A + \rho_B - \rho_v^o + \tag{S11}$$

$$\int\limits_{\rho_B \leq \rho_v^o < \rho_B + \rho_A} \rho_A + \rho_B - \rho_v^o$$

$$= \int\limits_{\rho_v^o < \rho_B} \rho_B - \rho_v^o + \tag{S12}$$

$$+ \int\limits_{\rho_v^o < \rho_B} \rho_A + \int\limits_{\rho_B \leq \rho_v^o < \rho_B + \rho_A} \rho_A + \rho_B - \rho_v^o$$

The first integral in the right-hand-side of the equation above is equal to $M[\rho_v^o - \rho_B]$ whereas the second and third are non-negative. As a result:

$$P[\rho_A + \rho_B - \rho_v^o] \geq M[\rho_v^o - \rho_B], \tag{S13}$$

which ends the proof.

## S3. DATA

### A. cc-pVDZ

| complex | $\rho_B$ | $E_k[\Delta\rho^c_{v'_A}, \rho'_A, \rho_B]$ [a] | $E_k[\Delta\rho^c_{v'_B}, \rho'_B, \rho_A]$ [a] | $E^{FDET(\rho_B)}_{int}$ [b] | $E^{ref}_{int}$ [c] | $\Delta_{CP}$ [d] | $M$ [e] | $P$ [f] | $P_{cmpl}$ [g] |
|---|---|---|---|---|---|---|---|---|---|
| 7HQ-2MeOH | $\rho^{isol}_B$ | 0.063 | | -11.200 | -14.482 | 11.159 | 0.108 | 0.216 | 0.350 |
| 7HQ-2MeOH | $\rho^{FAT}_B$ | 0.053 | 0.096 | -14.216 | -14.482 | 11.159 | 0.007 | 0.066 | 0.350 |
| 7HQ-formate | $\rho^{isol}_B$ | -0.009 | | -25.717 | -38.148 | 11.902 | 0.176 | 0.285 | 0.648 |
| 7HQ-formate | $\rho^{FAT}_B$ | -0.036 | 0.033 | -36.933 | -38.148 | 11.902 | 0.007 | 0.070 | 0.648 |
| uracil-5H$_2$O | $\rho^{isol}_B$ | 0.166 | | -32.320 | -34.396 | 23.111 | 0.203 | 0.394 | 0.600 |
| uracil-5H$_2$O | $\rho^{FAT}_B$ | 0.147 | 0.162 | -39.794 | -34.396 | 23.111 | 0.014 | 0.126 | 0.600 |
| PyrBnz-2HCOOH | $\rho^{isol}_B$ | 0.130 | | -26.794 | -32.300 | 13.907 | 0.175 | 0.406 | 0.583 |
| PyrBnz-2HCOOH | $\rho^{FAT}_B$ | 0.103 | 0.108 | -33.852 | -32.300 | 13.907 | 0.012 | 0.110 | 0.583 |

[a] defined in Eq. 7 in the manuscript

[b] defined in Eq. 12, and Eq. 17 in the manuscript for $\rho^{isol}_B$ and $\rho^{FAT}_B$ respectively.

[c] Counterpoise corrected, i.e. $E^{ref}_{int} = E^{(AB)}_v - E^{(AB)}_{v_A} - E^{(AB)}_{v_B}$, where all values are obtained with MP2, the subscript denotes the subsystem, and the superscript denotes the centres involved in the basis set expansion.

[d] Counterpoise correction: $\Delta_{CP} = E^{(A)}_{v_A} + E^{(B)}_{v_B} - E^{(AB)}_{v_A} - E^{(AB)}_{v_B}$, where all values are obtained with MP2, the subscript denotes the subsystem, and the superscript denotes the centres involved in the basis set expansion.

[e] $M = M[\rho^{o(ref)}_v - \rho^{FDET(FAT)}_B]$ with $M[\rho]$ defined in Eq. 18 in the manuscript

[f] $P_{cmpl} = P[\rho^{isol}_A + \rho^{isol}_B - \rho^{o(ref)}_v]$ (cf. Eq. 21), with $P[\rho]$ defined in Eq. 19 in the manuscript

[g] $P = P[\rho^{o(ref)}_v - \rho^{FDET(FAT)}_{tot}]$ with $P[\rho]$ defined in Eq. 19 in the manuscript

[h] $E^{FDET(FAT)}_{int}$ is given in Eq. 17 in the manuscript

TABLE S1. Supplementary data for results obtained with the *supermolecular expansion* and the cc-pVDZ basis set. Energies are given in kcal/mol and quantities related to electron densities are given in atomic units.

| complex | $\rho_B$ | $E_k[\Delta\rho^c_{v'_A},\rho'_A,\rho_B]$ [a] | $E_k[\Delta\rho^c_{v'_B},\rho'_B,\rho_A]$ [a] | $E_{int}^{FDET(\rho_B)}$ [b] | $E_{int}^{ref}$ [c] | $\Delta_{CP}$ [d] | $M$ [e] | $P$ [f] | $P_{cmpl}$ [g] |
|---|---|---|---|---|---|---|---|---|---|
| 7HQ-2MeOH | $\rho_B^{isol}$ | 0.023 | | -10.724 | -14.482 | 11.159 | 0.131 | 0.228 | 0.354 |
| 7HQ-2MeOH | $\rho_B^{pp(Mulliken)}$ | 0.021 | | -12.063 | -14.482 | 11.159 | 0.091 | 0.171 | 0.354 |
| 7HQ-2MeOH | $\rho_B^{pp(ChelPG)}$ | 0.021 | | -12.343 | -14.482 | 11.159 | 0.063 | 0.137 | 0.354 |
| 7HQ-2MeOH | $\rho_B^{FAT}$ | 0.020 | 0.036 | -12.062 | -14.482 | 11.159 | 0.045 | 0.103 | 0.354 |
| 7HQ-formate | $\rho_B^{isol}$ | -0.021 | | -27.189 | -38.148 | 11.902 | 0.226 | 0.306 | 0.645 |
| 7HQ-formate | $\rho_B^{pp(Mulliken)}$ | -0.021 | | -29.673 | -38.148 | 11.902 | 0.181 | 0.253 | 0.645 |
| 7HQ-formate | $\rho_B^{pp(ChelPG)}$ | -0.021 | | -30.350 | -38.148 | 11.902 | 0.154 | 0.222 | 0.645 |
| 7HQ-formate | $\rho_B^{FAT}$ | -0.021 | 0.040 | -29.917 | -38.148 | 11.902 | 0.116 | 0.177 | 0.645 |
| uracil-5H$_2$O | $\rho_B^{isol}$ | 0.068 | | -30.102 | -34.396 | 23.111 | 0.264 | 0.413 | 0.605 |
| uracil-5H$_2$O | $\rho_B^{pp(Mulliken)}$ | 0.063 | | -33.053 | -34.396 | 23.111 | 0.183 | 0.303 | 0.605 |
| uracil-5H$_2$O | $\rho_B^{pp(ChelPG)}$ | 0.061 | | -33.678 | -34.396 | 23.111 | 0.144 | 0.254 | 0.605 |
| uracil-5H$_2$O | $\rho_B^{FAT}$ | 0.057 | 0.051 | -32.892 | -34.396 | 23.111 | 0.120 | 0.206 | 0.605 |
| PyrBnz-2HCOOH | $\rho_B^{isol}$ | 0.088 | | -24.896 | -32.300 | 13.907 | 0.204 | 0.422 | 0.582 |
| PyrBnz-2HCOOH | $\rho_B^{pp(Mulliken)}$ | 0.079 | | -28.615 | -32.300 | 13.907 | 0.106 | 0.265 | 0.582 |
| PyrBnz-2HCOOH | $\rho_B^{pp(ChelPG)}$ | 0.077 | | -29.135 | -32.300 | 13.907 | 0.090 | 0.233 | 0.582 |
| PyrBnz-2HCOOH | $\rho_B^{FAT}$ | 0.071 | 0.025 | -28.491 | -32.300 | 13.907 | 0.060 | 0.162 | 0.582 |

[a] defined in Eq. 7 in the manuscript

[b] defined in Eq. 12, and Eq. 17 in the manuscript for $\rho_B^{isol}$ and $\rho_B^{FAT}$ respectively, and in Eq. 14 for $\rho_B^{pp(Mulliken)}$ and $\rho_B^{pp(ChelPG)}$

[c] Counterpoise corrected, i.e. $E_{int}^{ref} = E_v^{(AB)} - E_{v_A}^{(AB)} - E_{v_B}^{(AB)}$, where all values are obtained with MP2, the subscript denotes the subsystem, and the superscript denotes the centres involved in the basis set expansion.

[d] Counterpoise correction: $\Delta_{CP} = E_{v_A}^{(A)} + E_{v_B}^{(B)} - E_{v_A}^{(AB)} - E_{v_B}^{(AB)}$, where all values are obtained with MP2, the subscript denotes the subsystem, and the superscript denotes the centres involved in the basis set expansion.

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

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

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

[h] $E_{int}^{FDET(FAT)}$ is given in Eq. 17 in the manuscript

TABLE S2. Supplementary data for results obtained with the *monomer expansion* and the cc-pVDZ basis set. Energies are given in kcal/mol and quantities related to electron densities are given in atomic units.
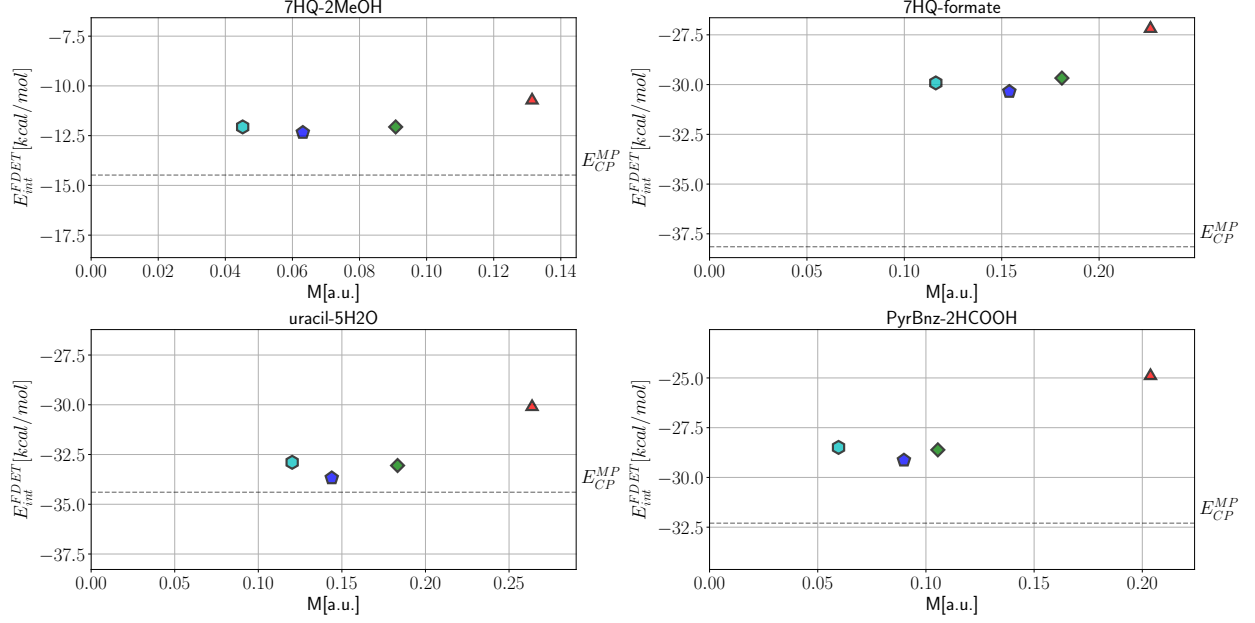
FIG. S1. Integrated negative density $M$ and the FDET-MP2 interaction energy for various choices of $\rho_B$: a) $\rho_B^{isol}$ (orange triangles), b) $\rho_B^{FAT}$ (light blue hexagons), c) $\rho_B^{pp(Mulliken)}$ (green diamonds), and d) $\rho_B^{pp(ChelPG)}$ (dark blue pentagons). Data obtained using the *monomer expansion* and the cc-pVDZ basis set. Horizontal lies indicate the reference interaction energy.

Figure S4 shows that the pre-polarisation of $\rho_B$ using the ChelPG representation of $\rho_A^{isol}$ results also in a significant improvement in the total density.
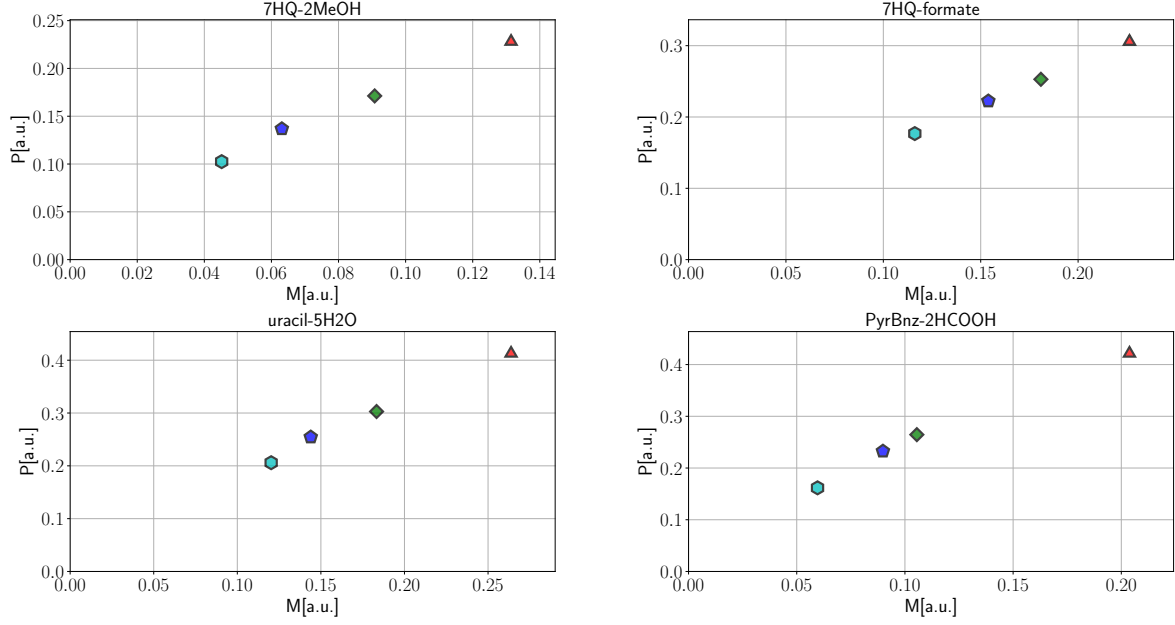
FIG. S2. Integrated negative density $M$ and the total density error $P$ for various choices of $\rho_B$: a) $\rho_B^{isol}$ (orange triangles), b) $\rho_B^{FAT}$ (light blue hexagons), c) $\rho_B^{pp(Mulliken)}$ (green diamonds), and d) $\rho_B^{pp(ChelPG)}$ (dark blue pentagons). Data obtained using the *monomer expansion* and the cc-pVDZ basis set.

## B.  aug-cc-pVDZ

| complex | $\rho_B$ | $E_k[\Delta\rho^c_{v'_A},\rho'_A,\rho_B]$ [a] | $E_k[\Delta\rho^c_{v'_B},\rho'_B,\rho_A]$ [a] | $E_{int}^{FDET(\rho_B)}$ [b] | $E_{int}^{ref}$ [c] | $\Delta_{CP}$ [d] | $M$ [e] | $P$ [f] | $P_{cmpl}$ [g] |
|---|---|---|---|---|---|---|---|---|---|
| 7HQ-2MeOH | $\rho_B^{isol}$ | 0.092 | | -10.983 | -17.473 | 5.632 | 0.121 | 0.228 | 0.362 |
| 7HQ-2MeOH | $\rho_B^{FAT}$ | 0.081 | 0.127 | -14.269 | -17.473 | 5.632 | 0.007 | 0.059 | 0.362 |
| 7HQ-formate | $\rho_B^{isol}$ | 0.020 | | -23.037 | -36.485 | 3.522 | 0.206 | 0.316 | 0.670 |
| 7HQ-formate | $\rho_B^{FAT}$ | -0.008 | 0.068 | -33.325 | -36.485 | 3.522 | 0.007 | 0.066 | 0.670 |
| uracil-5H$_2$O | $\rho_B^{isol}$ | 0.224 | | -32.065 | -38.620 | 7.839 | 0.234 | 0.426 | 0.633 |
| uracil-5H$_2$O | $\rho_B^{FAT}$ | 0.204 | 0.236 | -39.695 | -38.620 | 7.839 | 0.014 | 0.114 | 0.633 |
| PyrBnz-2HCOOH | $\rho_B^{isol}$ | 0.160 | | -26.948 | -36.532 | 7.318 | 0.184 | 0.416 | 0.600 |
| PyrBnz-2HCOOH | $\rho_B^{FAT}$ | 0.131 | 0.164 | -34.894 | -36.532 | 7.318 | 0.013 | 0.104 | 0.600 |

[a] defined in Eq. 7 in the manuscript

[b] defined in Eq. 12, and Eq. 17 in the manuscript for $\rho_B^{isol}$ and $\rho_B^{FAT}$ respectively.

[c] Counterpoise corrected, i.e. $E_{int}^{ref} = E_v^{(AB)} - E_{v_A}^{(AB)} - E_{v_B}^{(AB)}$, where all values are obtained with MP2, the subscript denotes the subsystem, and the superscript denotes the centres involved in the basis set expansion.

[d] Counterpoise correction: $\Delta_{CP} = E_{v_A}^{(A)} + E_{v_B}^{(B)} - E_{v_A}^{(AB)} - E_{v_B}^{(AB)}$, where all values are obtained with MP2, the subscript denotes the subsystem, and the superscript denotes the centres involved in the basis set expansion.

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

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

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

[h] $E_{int}^{FDET(FAT)}$ is given in Eq. 17 in the manuscript

TABLE S3. Supplementary data for results obtained with the *supermolecular expansion* and the aug-cc-pVDZ basis set. Energies are given in kcal/mol and quantities related to electron densities are given in atomic units.

| complex | $\rho_B$ | $E_k[\Delta\rho^c_{v'_A},\rho'_A,\rho_B]$ [a] | $E_k[\Delta\rho^c_{v'_B},\rho'_B,\rho_A]$ [a] | $E_{int}^{FDET(\rho_B)}$ [b] | $E_{int}^{ref}$ [c] | $\Delta_{CP}$ [d] | $M$ [e] | $P$ [f] | $P_{cmpl}$ [g] |
|---|---|---|---|---|---|---|---|---|---|
| 7HQ-2MeOH | $\rho_B^{isol}$ | 0.087 | | -10.702 | -17.473 | 5.632 | 0.123 | 0.231 | 0.368 |
| 7HQ-2MeOH | $\rho_B^{pp(Mulliken)}$ | 0.084 | | -10.688 | -17.473 | 5.632 | 0.121 | 0.221 | 0.368 |
| 7HQ-2MeOH | $\rho_B^{pp(ChelPG)}$ | 0.084 | | -13.492 | -17.473 | 5.632 | 0.037 | 0.118 | 0.368 |
| 7HQ-2MeOH | $\rho_B^{FAT}$ | 0.081 | 0.128 | -13.181 | -17.473 | 5.632 | 0.013 | 0.073 | 0.368 |
| 7HQ-formate | $\rho_B^{isol}$ | 0.021 | | -22.927 | -36.485 | 3.522 | 0.205 | 0.316 | 0.673 |
| 7HQ-formate | $\rho_B^{pp(Mulliken)}$ | 0.016 | | -27.202 | -36.485 | 3.522 | 0.148 | 0.242 | 0.673 |
| 7HQ-formate | $\rho_B^{pp(ChelPG)}$ | 0.015 | | -28.678 | -36.485 | 3.522 | 0.096 | 0.186 | 0.673 |
| 7HQ-formate | $\rho_B^{FAT}$ | 0.011 | 0.117 | -28.028 | -36.485 | 3.522 | 0.036 | 0.114 | 0.673 |
| uracil-5H$_2$O | $\rho_B^{isol}$ | 0.212 | | -31.629 | -38.620 | 7.839 | 0.237 | 0.427 | 0.643 |
| uracil-5H$_2$O | $\rho_B^{pp(Mulliken)}$ | 0.201 | | -35.221 | -38.620 | 7.839 | 0.173 | 0.305 | 0.643 |
| uracil-5H$_2$O | $\rho_B^{pp(ChelPG)}$ | 0.205 | | -38.010 | -38.620 | 7.839 | 0.058 | 0.206 | 0.643 |
| uracil-5H$_2$O | $\rho_B^{FAT}$ | 0.196 | 0.234 | -37.182 | -38.620 | 7.839 | 0.024 | 0.129 | 0.643 |
| PyrBnz-2HCOOH | $\rho_B^{isol}$ | 0.161 | | -25.795 | -36.532 | 7.318 | 0.185 | 0.419 | 0.606 |
| PyrBnz-2HCOOH | $\rho_B^{pp(Mulliken)}$ | 0.141 | | -31.448 | -36.532 | 7.318 | 0.152 | 0.281 | 0.606 |
| PyrBnz-2HCOOH | $\rho_B^{pp(ChelPG)}$ | 0.146 | | -32.291 | -36.532 | 7.318 | 0.057 | 0.215 | 0.606 |
| PyrBnz-2HCOOH | $\rho_B^{FAT}$ | 0.139 | 0.164 | -32.418 | -36.532 | 7.318 | 0.016 | 0.127 | 0.606 |

[a] defined in Eq. 7 in the manuscript

[b] defined in Eq. 12, and Eq. 17 in the manuscript for $\rho_B^{isol}$ and $\rho_B^{FAT}$ respectively, and in Eq. 14 for $\rho_B^{pp(Mulliken)}$ and $\rho_B^{pp(ChelPG)}$

[c] Counterpoise corrected, i.e. $E_{int}^{ref} = E_v^{(AB)} - E_{v_A}^{(AB)} - E_{v_B}^{(AB)}$, where all values are obtained with MP2, the subscript denotes the subsystem, and the superscript denotes the centres involved in the basis set expansion.

[d] Counterpoise correction: $\Delta_{CP} = E_{v_A}^{(A)} + E_{v_B}^{(B)} - E_{v_A}^{(AB)} - E_{v_B}^{(AB)}$, where all values are obtained with MP2, the subscript denotes the subsystem, and the superscript denotes the centres involved in the basis set expansion.

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

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

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

[h] $E_{int}^{FDET(FAT)}$ is given in Eq. 17 in the manuscript

TABLE S4. Supplementary data for results obtained with the *monomer expansion* and the aug-cc-pVDZ basis set. Energies are given in kcal/mol and quantities related to electron densities are given in atomic units.

## C.  cc-pVTZ

| complex | $\rho_B$ | $E_k[\Delta\rho^c_{v'_A},\rho'_A,\rho_B]$ [a] | $E_k[\Delta\rho^c_{v'_B},\rho'_B,\rho_A]$ [a] | $E_{int}^{FDET(\rho_B)}$ [b] | $E_{int}^{ref}$ [c] | $\Delta_{CP}$ [d] | $M$ [e] | $P$ [f] | $P_{cmpl}$ [g] |
|---|---|---|---|---|---|---|---|---|---|
| 7HQ-2MeOH | $\rho_B^{isol}$ | 0.040 | | -10.716 | -17.858 | 5.681 | 0.125 | 0.226 | 0.358 |
| 7HQ-2MeOH | $\rho_B^{pp(Mulliken)}$ | 0.038 | | -12.684 | -17.858 | 5.681 | 0.071 | 0.153 | 0.358 |
| 7HQ-2MeOH | $\rho_B^{pp(ChelPG)}$ | 0.037 | | -13.254 | -17.858 | 5.681 | 0.045 | 0.119 | 0.358 |
| 7HQ-2MeOH | $\rho_B^{FAT}$ | 0.035 | 0.068 | -13.132 | -17.858 | 5.681 | 0.021 | 0.074 | 0.358 |
| 7HQ-formate | $\rho_B^{isol}$ | 0.002 | | -25.786 | -39.521 | 6.011 | 0.214 | 0.310 | 0.661 |
| 7HQ-formate | $\rho_B^{pp(Mulliken)}$ | 0.003 | | -29.829 | -39.521 | 6.011 | 0.145 | 0.230 | 0.661 |
| 7HQ-formate | $\rho_B^{pp(ChelPG)}$ | 0.003 | | -30.768 | -39.521 | 6.011 | 0.116 | 0.196 | 0.661 |
| 7HQ-formate | $\rho_B^{FAT}$ | 0.003 | 0.064 | -30.578 | -39.521 | 6.011 | 0.057 | 0.127 | 0.661 |
| uracil-5H$_2$O | $\rho_B^{isol}$ | 0.119 | | -32.283 | -40.078 | 11.802 | 0.247 | 0.413 | 0.617 |
| uracil-5H$_2$O | $\rho_B^{pp(Mulliken)}$ | 0.112 | | -37.341 | -40.078 | 11.802 | 0.123 | 0.249 | 0.617 |
| uracil-5H$_2$O | $\rho_B^{pp(ChelPG)}$ | 0.111 | | -37.944 | -40.078 | 11.802 | 0.091 | 0.212 | 0.617 |
| uracil-5H$_2$O | $\rho_B^{FAT}$ | 0.104 | 0.099 | -37.489 | -40.078 | 11.802 | 0.054 | 0.139 | 0.617 |

[a] defined in Eq. 7 in the manuscript

[b] defined in Eq. 12, and Eq. 17 in the manuscript for $\rho_B^{isol}$ and $\rho_B^{FAT}$ respectively, and in Eq. 14 for $\rho_B^{pp(Mulliken)}$ and $\rho_B^{pp(ChelPG)}$

[c] Counterpoise corrected, i.e. $E_{int}^{ref} = E_v^{(AB)} - E_{v_A}^{(AB)} - E_{v_B}^{(AB)}$, where all values are obtained with MP2, the subscript denotes the subsystem, and the superscript denotes the centres involved in the basis set expansion.

[d] Counterpoise correction: $\Delta_{CP} = E_{v_A}^{(A)} + E_{v_B}^{(B)} - E_{v_A}^{(AB)} - E_{v_B}^{(AB)}$, where all values are obtained with MP2, the subscript denotes the subsystem, and the superscript denotes the centres involved in the basis set expansion.

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

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

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

[h] $E_{int}^{FDET(FAT)}$ is given in Eq. 17 in the manuscript

TABLE S5. Supplementary data for results obtained with the *monomer expansion* and the cc-pVTZ basis set. Energies are given in kcal/mol and quantities related to electron densities are given in atomic units.
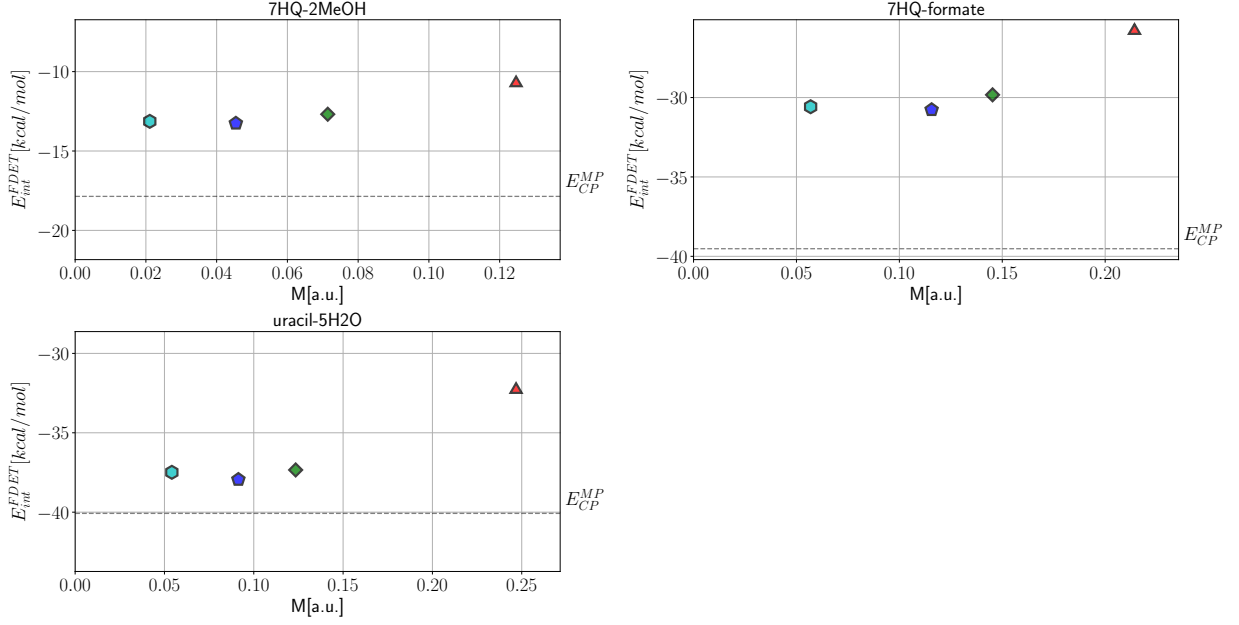
FIG. S3. Integrated negative density $M$ and the FDET-MP2 interaction energy for various choices of $\rho_B$: a) $\rho_B^{isol}$ (orange triangles), b) $\rho_B^{FAT}$ (light blue hexagons), c) $\rho_B^{pp(Mulliken)}$ (green diamonds), and d) $\rho_B^{pp(ChelPG)}$ (dark blue pentagons). Data obtained using the *monomer expansion* and the cc-pVTZ basis set. Horizontal lies indicate the reference interaction energy.

Figure S4 shows that the pre-polarisation of $\rho_B$ using the ChelPG representation of $\rho_A^{isol}$ results also in a significant improvement in the total density.
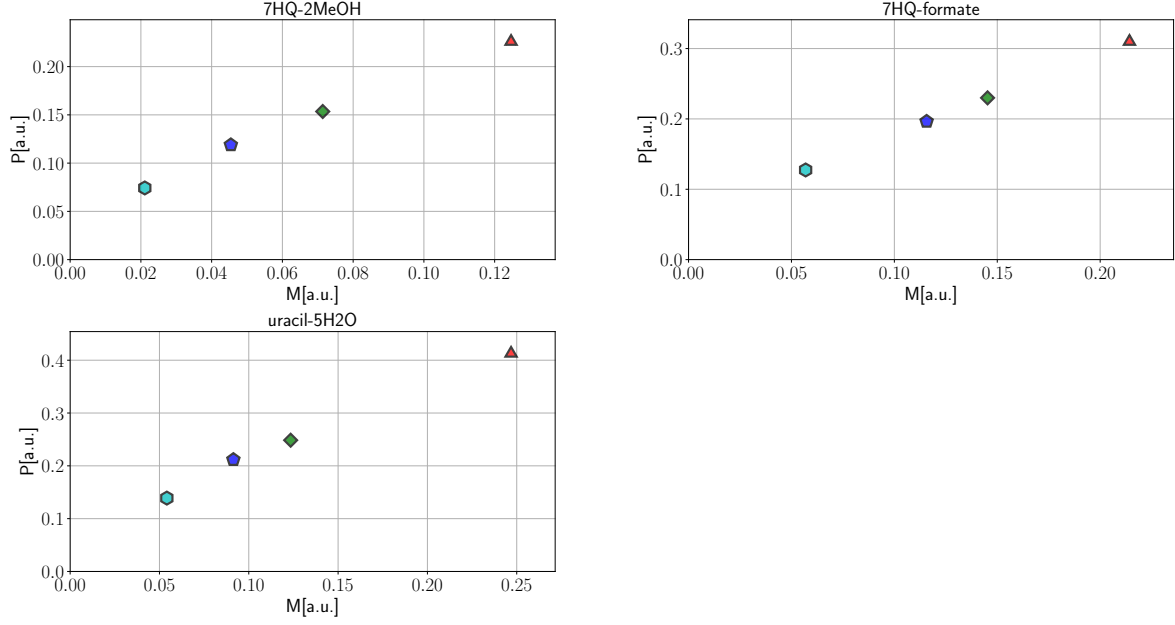
FIG. S4. Integrated negative density $M$ and the total density error $P$ for various choices of $\rho_B$: a) $\rho_B^{isol}$ (orange triangles), b) $\rho_B^{FAT}$ (light blue hexagons), c) $\rho_B^{pp(Mulliken)}$ (green diamonds), and d) $\rho_B^{pp(ChelPG)}$ (dark blue pentagons). Data obtained using the *monomer expansion* and the cc-pVTZ basis set.