Ab Initio Effective One-Electron Potential Operators. II. Applications for Exchange-Repulsion Energy in Effective Fragment Potentials

Bartosz Błasiak,[1, a)] Robert W. Góra,[1] and Wojciech Bartkowiak[1]
Department of Physical and Quantum Chemistry, Faculty of Chemistry, Wrocław University of Science and Technology, Wybrzeże Wyspiańskiego 27, Wrocław 50-370, Poland

(Dated: 4 February 2020)

In Paper I, the effective one-electron potentials (OEP) method was introduced and demonstrated as an efficient approach to reduce the computational cost of evaluation of the charge-transfer interaction energy within the effective fragment potential method (EFP2) by an average factor of 20, making it no longer a bottleneck in EFP2-based simulations of complex systems. Here, the OEP technique is used to enhance computational efficiency in evaluating the exchange-repulsion EFP2 interaction energy by redefining the first-order repulsive term of Murrell et al. [Murrell et al., Proc. R. Soc. Lond. A 284, 566 (1965)] through the extended density fitting in incomplete auxiliary basis. In the proposed approach the evaluation of the kinetic energy integrals is no longer required and the computational cost can be reduced roughly by a factor of 1.5 as compared to the original EFP2 formulation.

I. INTRODUCTION

The exchange-repulsion (EXR) interaction energy usually describes the non-electrostatic repulsion between two isolated and unperturbed wavefunctions in an interacting molecular complex.[1–7] In the second generation of the effective fragment potentials (EFP2) method[8–10] the EXR energy is an important component of the total interaction energy. It is derived from the intermolecular perturbation theory of Murrell et al.[1,2] up to the second order with respect to the wavefunction overlap assuming Hartree-Fock[11] (HF) wavefunctions, and is given by[5,12,13]

$$E^{\mathrm{Ex-Rep}} \approx E^{\mathrm{Ex}} + E^{\mathrm{Rep}}(S^{-1}) + E^{\mathrm{Rep}}(S^{-2}) \,. \quad (1)$$

In the above equation, E^{Ex} is the exchange energy,[12]

$$E^{\mathrm{Ex}} \approx 2 \sum_{i\in A}^{\mathrm{Locc}} \sum_{j\in B}^{\mathrm{Locc}} \sqrt{\frac{-2\ln|S_{ij}|}{\pi}} \frac{S_{ij}^2}{r_{ij}} \,, \quad (2)$$

whereas the first- and second-order repulsion terms are accordingly[5,13]

$$E^{\mathrm{Rep}}(S^{-1}) \approx -2 \sum_{i\in A}^{\mathrm{Locc}} \sum_{j\in B}^{\mathrm{Locc}} S_{ij} \times \left\{ \sum_{k\in A}^{\mathrm{Locc}} F_{ik}^A S_{kj} + \sum_{l\in B}^{\mathrm{Locc}} F_{jl}^A S_{li} - 2T_{ij} \right\} \quad (3)$$

and

$$E^{\mathrm{Rep}}(S^{-2}) \approx 2 \sum_{i\in A}^{\mathrm{Locc}} \sum_{j\in B}^{\mathrm{Locc}} S_{ij}^2 \times \left\{ \sum_{x\in A}^{\mathrm{At}} \frac{-Z_x}{r_{xb}} + \sum_{y\in B}^{\mathrm{At}} \frac{-Z_y}{r_{ya}} + \sum_{k\in A}^{\mathrm{Locc}} \frac{2}{r_{jk}} + \sum_{l\in B}^{\mathrm{Locc}} \frac{2}{r_{il}} - r_{ij}^{-1} \right\} . \quad (4)$$

The indices i, j, k and l label the localized occupied orbitals (LMO's, denoted by 'Locc') located at their charge centroids $\mathbf{r}_{i(j)}$, whereas x and y label atomic nuclei (denoted by 'At') with atomic numbers $Z_{x(y)}$ located at $\mathbf{r}_{x(y)}$. The relative distances are defined by $r_{uw} = |\mathbf{r}_u - \mathbf{r}_w|$. S_{ij} and T_{ij} are the overlap and kinetic-energy integrals, respectively, whereas F_{ik}^X are the Fock matrix elements of unperturbed monomer X, all defined in LMO basis. It is known that evaluating Eq. (1) is remarkably efficient and gives usually very accurate estimates of the reference EXR energies according to the symmetry adapted perturbation theory[4] (SAPT). The computational cost to evaluate the exchange and second-order terms is much smaller than that of the first-order term, which requires evaluation of not only the one-electron overlap, but also the kinetic energy integrals in a space of occupied molecular orbital (MO) basis.

In this work, an alternative formulation for the first-order term is proposed based on the effective one-electron potential (OEP) operator technique developed in the preceding contribution.[14] It is shown that using this approach calculation of kinetic energy integrals can be avoided what leads to increased computational efficiency.

II. RESULTS

A theoretical foundation of the first-order repulsive term in the EFP2 model is the perturbation theory of Murrell et al.[1] in which

$$E^{\mathrm{Rep}}(S^{-1}) = -2 \sum_{i\in A}^{\mathrm{Locc}} \sum_{j\in B}^{\mathrm{Locc}} S_{ij} \Big\{ -W_{ij}^A - W_{ij}^B + \sum_{k\in A}^{\mathrm{Occ}} \left[2\langle ij|kk\rangle - \langle ik|jk\rangle \right] + \sum_{l\in B}^{\mathrm{Occ}} \left[2\langle ij|ll\rangle - \langle il|jl\rangle \right] \Big\} \quad (5)$$

a) blasiak.bartosz@gmail.com; https://www.polonez.pwr.edu.pl

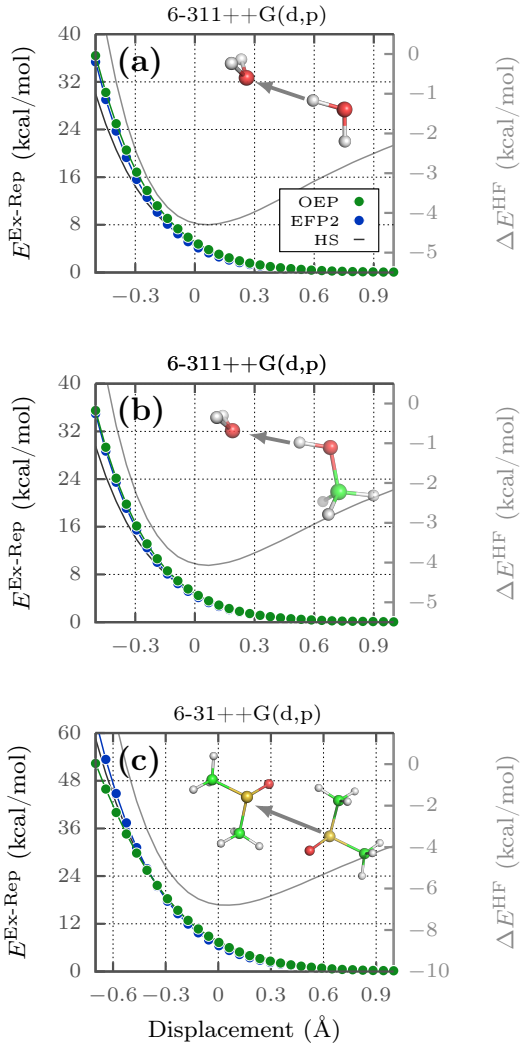

FIG. 1. **Asymptotic dependence of the exchange-repulsion interaction energy for selected bi-molecular complexes.** (a) water dimer, (b) water-methanol complex, and (c) DMSO dimer in S_i symmetry, were one molecule has been translated along the vector specified in the insets relative to initial geometry, optimized at HF/6-31+G(d,p) level of theory. The total interaction energy is also shown for comparison in light grey color in this figure. Interaction energies were obtained by using the 6-311++G(d,p) primary basis set for (a) and (b), and 6-31++G(d,p) for (c). In OEP calculations, EDF-1 scheme combined with small optimized auxiliary basis sets (cf. Supplementary Information) were used.

where $W_{ij}^A \equiv \langle i|\hat{v}_{\text{nuc}}^A|j\rangle$, $\hat{v}_{\text{nuc}}^A$ is the electrostatic potential operator due to nuclei and ERIs are defined by

$$\langle\alpha\beta|\gamma\delta\rangle \equiv \iint \frac{\phi_\alpha^*(\mathbf{r}_1)\phi_\beta(\mathbf{r}_1)\phi_\gamma^*(\mathbf{r}_2)\phi_\delta(\mathbf{r}_2)}{|\mathbf{r}_1-\mathbf{r}_2|}d\mathbf{r}_1 d\mathbf{r}_2 \ . \quad (6)$$

Now, note that the Coulomb and exchange integrals can be re-cast as follows:

$$\langle ij|kk\rangle \equiv -\sum_{\mu\in A} C_{\mu i}^A \langle\mu|\hat{v}_{kk}|j\rangle \ , \quad (7a)$$

$$\langle ik|jk\rangle \equiv -\sum_{\mu\in A} C_{\mu k}^A \langle\mu|\hat{v}_{ik}|j\rangle \ . \quad (7b)$$

In the above equations, the auxiliary potential operators are given by

$$\hat{v}_{ik} \equiv -\int d\mathbf{r}|\mathbf{r}\rangle \left[\int d\mathbf{r}' \frac{\phi_i^*(\mathbf{r}')\phi_k(\mathbf{r}')}{|\mathbf{r}'-\mathbf{r}|}\right]\langle\mathbf{r}| \ . \quad (8)$$

ERIs can be effectively eliminated from Eq. (5) by using the following prescription[14]

$$\sum_t \sum_{ij} \sum_{kl\in A} \mathcal{F}_t\left[\langle ij|k^A l^A\rangle\right] = \sum_{ij}\langle i|\hat{v}_{\text{eff}}^A|j\rangle \ . \quad (9)$$

where $\mathcal{F}_t$ is a certain and well defined functional of ERIs. Application of Eq. (9) enables to define a joint OEP operator constructed from nuclear, Coulomb and exchange contributions as

$$-W_{ij}^A + \sum_{k\in A}^{\text{Occ}} \left\{2\langle ij|kk\rangle - \langle ik|jk\rangle\right\} \equiv \sum_{\mu\in A}\langle\mu|\hat{v}_{\text{eff}}^{A[\mu i]}|j\rangle \quad (10)$$

with

$$\hat{v}_{\text{eff}}^{A[\mu i]} \equiv C_{\mu i}\hat{v}_{\text{nuc}}^A + \sum_{k\in A}^{\text{Occ}}\left[2C_{\mu i}^A\hat{v}_{kk}^A - C_{\mu k}^A\hat{v}_{ik}^A\right] \ . \quad (11)$$

On the other hand, it immediately follows that

$$\hat{v}_{\text{eff}}^{A[\mu i]}|\mu\rangle = \sum_{k\in A}^{\text{Occ}}\left\{2\hat{v}_{kk}^A|i\rangle - \hat{v}_{ik}^A|k\rangle\right\} \ . \quad (12)$$

For practical calculations, the right hand side of the above equation can be expanded in the auxiliary basis,

$$\sum_{k\in A}^{\text{Occ}}\left\{2\hat{v}_{kk}^A|i\rangle - \hat{v}_{ik}^A|k\rangle\right\} \cong \sum_{\xi\in A}^{\text{DF}} V_{\xi i}^A|\xi\rangle \ , \quad (13)$$

where the matrix $\mathbf{V}^A$ can be considered as a set of effective fragment parameters. Doing the same operations on the twin operators associated with the molecule B original theory of Murrell et. al from Eq. (5) reduces to

$$E^{\text{Rep}}(S^{-1}) \cong -2\sum_{i\in A}^{\text{Occ}}\sum_{j\in B}^{\text{Occ}} S_{ij}\left\{\sum_{\xi\in A}^{\text{DF}} V_{\xi i}^A S_{\xi j} + \sum_{\eta\in B}^{\text{DF}} V_{\eta j}^B S_{\eta i}\right\} , \quad (14)$$

where the effective fragment parameters can be obtained from the extended density fitting[14] (EDF) by

$$V_{\xi i}^X = \sum_{\eta\in X}^{\text{DF}}\sum_{\varepsilon,\zeta\in X}^{\text{RI}} \left[\mathbf{R}^{-1}\right]_{\xi\eta} R_{\eta\varepsilon}\left[\mathbf{S}^{-1}\right]_{\varepsilon\zeta}\langle\zeta|\hat{v}_{\text{eff}}^{X[\zeta i]}|i\rangle \ . \quad (15)$$

TABLE I. **CPU timings in miliseconds of exchange-repulsion single point energy calculations via various methods**[a]

	$(H_2O)_2$[b]		H_2O–$HOCH_3$[b]		$(DMSO)_2$[c]	
Hayes & Stone[d]	1.16×10^3	(4.81)	1.12×10^4	(4.72)	1.31×10^5	(7.34)
Murrell et al.[e]	2.33×10^2	(4.53)	3.87×10^3	(4.44)	2.79×10^4	(6.84)
EFP2	0.496	(4.45)	0.84	(4.44)	4.65	(6.56)
OEP/large[f]	0.884	(4.83)	1.836	(4.83)	8.95	(6.77)
OEP/mini[f]	0.324	(5.13)	0.566	(4.81)	3.00	(7.41)

[a] 1.2 GHz AMD EPYC™ 7301 16-Core Processor, calculations performed on 1 core. Exchange-repulsion energies are given in parentheses for reference (kcal/mol). Time profiling of the code performance was carried out as in Ref.[14]. See also the implementation details in the main text.
[b] Primary basis set: 6-311++G(d,p).
[c] Primary basis set: 6-31++G(d,p).
[d] Reference[3].
[e] Reference[1].
[f] Same as EFP2 but with $E^{\mathrm{Rep}}(S^{-1})$ replaced by formula in Eq. (14). Auxiliary basis sets for the EDF in $E^{\mathrm{Rep}}(S^{-1})$ from Eq. (14) are: 'large' - aug-cc-pVDZ-jkfit for $(H_2O)_2$ and H_2O–$HOCH_3$ systems as well as aug-cc-pVDZ-ri for $(DMSO)_2$ system; 'mini' - optimized small basis sets (see Supporting Information).

In this procedure, $R_{\alpha\gamma} = \iint d\mathbf{r}_1 d\mathbf{r}_2 \frac{\phi^*_\alpha(\mathbf{r}_1)\phi_\beta(\mathbf{r}_2)}{|\mathbf{r}_1-\mathbf{r}_2|}$, $\mathbf{S}$ is the matrix of overlap integrals in AO basis and the required OEP matrix elements can be calculated from

$$\left\langle\zeta\middle|\hat{v}_{\mathrm{eff}}^{X[\zeta i]}\middle|i\right\rangle = -\sum_{x\in X}^{\mathrm{At}} W_{\zeta i}^{(x)} + \sum_{\beta\gamma\delta\in X}^{\mathrm{AO}} \left\{2C^X_{\beta i}D^X_{\gamma\delta} - C^X_{\gamma i}D^X_{\beta\delta}\right\}\left\langle\zeta\beta\middle|\gamma\delta\right\rangle , \quad (16)$$

where $\mathbf{D}^X$ and $\mathbf{C}^X$ are the one-particle AO density and the LCAO-MO matrices of isolated molecule X, respectively. Eq. (14) and Eq. (3) have almost the same form, with two exceptions in the new OEP formulation: (i) only overlap integrals are needed, which are relatively easy to compute. Furthermore, there is no need to evaluate the kinetic energy integrals, as in the original EFP2 formulation, which are computationally slightly more demanding; (ii) overlap integrals need to be evaluated also between auxiliary basis set. The smaller the size of the auxiliary basis, the less expensive evaluation of the $E^{\mathrm{Rep}}(S^{-1})$ becomes. The alternative EFP2 formulation is therefore still given by Eq. (1) but with $E^{\mathrm{Rep}}(S^{-1})$ replaced by OEP-based first-order term from Eq. (14).

In our in-house plug-in to the Psi4 quantum chemistry program,[15] we implemented EXR energy models from Eq. (1) (EFP2), Eq. (14) (OEP), as well as the approximate intermolecular perturbation theory with exchange of Murrell et al.[1], and the exact EXR energy of Hayes & Stone[3], here referred to as the HS model. The latter was implemented in the dimer-centered basis set[16] (DCBS), to eliminate the basis set superposition error (BSSE) in the benchmark calculations. The Boys method[17] was used to localize molecular orbitals—the same as in the original EFP2 formulation. Note here that, since Murrell et al. theory is invariant with respect to unitary transformation of molecular orbitals, the OEP-based expression in Eq. (14) is also invariant and does not require orbital localization.

Three complexes: $(H_2O)_2$, H_2O–$HOCH_3$, and $(DMSO)_2$, were chosen as model systems to analyze the asymptotic dependence of EXR energy, which is shown in Figure 1. The reference (zero-displacement) geometries were obtained as described in Ref.[14] and the structures along the translation direction are depicted in the insets in Figure 1. Energy-optimizations were performed at the HF/6-31+G(d,p) level, as implemented in the Gaussian16 quantum chemistry program package.[22] OEP and EFP2 models correctly describe the EXR energy at all separations for the studied model systems with very similar accuracy. However, contrary to our previous application of the OEP technique[14], where approximately 20-fold speedups as compared to the EFP2 model were achieved for the evaluation of the charge-transfer energy, the CPU timings of EXR calculations in the OEP method are comparable but roughly 2 times higher than that of the EFP2 model when the usual density fitting auxiliary basis sets such as aug-cc-pVDZ-jkfit or aug-cc-pVDZ-ri are used (Table I). Computational cost of $E^{\mathrm{Rep}}(S^{-1})$ in EFP2 model is approximately $s(on^2+o^2n)+t(on^2+o^2n)+2o^3$, whereas in OEP model is $s(on^2+o^2n)+2ao^2$, both of which have comparable magnitudes. Here, s and t denote the relative costs of evaluation of the overlap and kinetic-energy one electron integrals (OEIs), whereas o, n and a is the number of LMOs, primary AOs and auxiliary AOs, respectively. It is clear that the critical parameter is the number of auxiliary basis functions a, which should be comparable with the number of occupied orbitals o as t is usually only 2–3 times larger than s. To investigate this issue more thoroughly we developed small auxiliary basis sets for H_2O, CH_3OH and DMSO, via the basis set optimization method described in Appendix A to Pa-

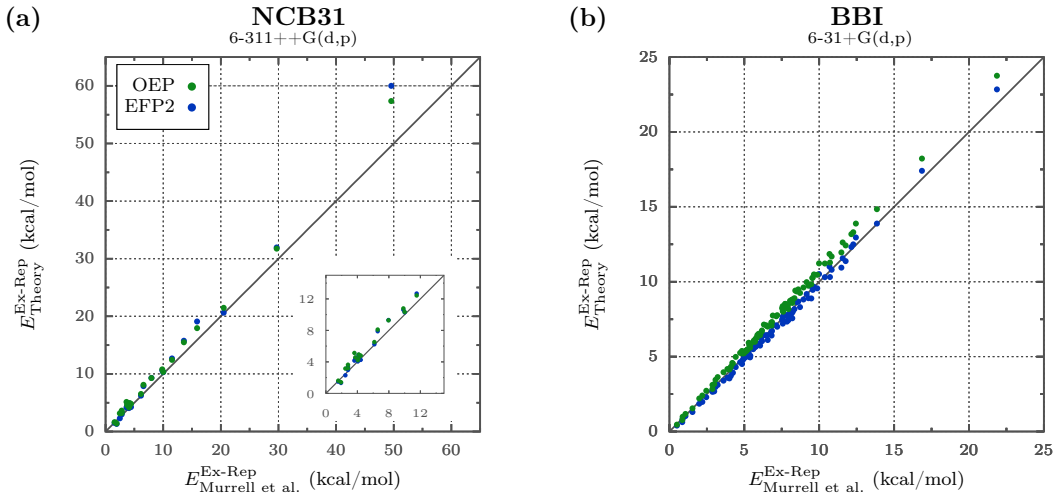

FIG. 2. **Accuracy of the OEP and EFP2 models of exchange-repulsion energy across various bi-molecular systems.** (a) NCB31 database[18–20] of non-covalent interactions and (b) BBI subset[21] of backbone-backbone interactions in proteins from the BioFragment Database. For the OEP calculations, the EDF-1 scheme with the aug-cc-pVDZ-jkfit auxiliary basis set was used.

per I in conjunction with the basin hopping global optimization algorithm[23–26] as implemented in the SciPy Python library[27] (see Supplementary Information). Applying such basis sets reduces computational effort appreciably and makes the OEP method on average 1.5 times faster than EFP2. For instance, calculation of the EXR energy in the DMSO dimer by using 6-31++G(d,p) primary basis set and 1.2 GHz AMD EPYC$^{\text{TM}}$ 7301 16-Core Processor requires about 4.6 ms and 3.0 ms when EFP2 and OEP models are used, respectively.

To investigate the accuracy of the OEP-based repulsion term across a variety of interacting systems, a selection of bi-molecular complexes from the non-covalent interactions database NCB31 developed by the Truhlar's group,[18–20] as well as the BioFragment Database subset BBI for backbone-backbone interactions in proteins of Sherrill group,[21] as implemented in the Psi4 program,[15] was utilized. The computed EXR energies were compared with the reference Murrell et al. results (Figure 2) as well as with the reference HS results (Figure S1). On average, OEP and EFP2 methods are in good agreement with the reference models and correlation R^2 coefficients are between 90–99% in all data sets. The mutual differences between EFP2 and OEP estimations of $E^{\text{Rep}}(S^{-1})$ are on the average 0.7 kcal/mol (Figure S2) which shows that the ERI elimination proposed here is accurate. Root mean square errors (RMSE) of EXR energy estimation via OEP and EFP2 models are around 2.0 kcal/mol and 2.6 kcal/mol in the NCB31/6-311++G(d,p) data set, and 0.6 kcal/mol and 0.3–1.2 kcal/mol in the BBI/6-31+G(d,p) data set, respectively. However, while the OEP and EFP2 models tend to overestimate the EXR energy as compared to the Murrell et al. reference by 7–12% (except for the EFP2 model in the BBI/6-31+G(d,p) data set that shows very good agreement), they consistently underestimate the EXR energy as compared to the HS reference by 0.7–7% (OEP) and 8–15% (EFP2). This is mostly due to BSSE, that is corrected for only in the reference calculations with the HS model. Nevertheless, OEP model is rather of roughly comparable accuracy as EFP2 model in all the systems studied. Together with the performance data of both models from Table I, OEP model consistently outperforms the EFP2 approach provided that the minimal auxiliary basis set is used.

III. SUMMARY AND A FEW CONCLUDING REMARKS

In this work, the effective one-electron potential (OEP) technique, proposed in the preceding paper for the effective elimination of electron repulsion integrals in ab initio calculations, was utilized to reduce the computational cost of evaluation of the exchange repulsion (EXR) energy in the EFP2 model. Starting from the first-order formula for EXR in Murrell et al. perturbation theory, being a foundation of the EFP2 EXR term, OEP-based expression was derived which requires evaluation of only the overlap one-electron integrals, in contrast to EFP2 approach which requires also the kinetic-energy one-electron integrals. The reported results indicate that a following model should be more efficient for calculations of the total exchange-repulsion energy in the EFP2 model: (i) E^{Ex} and $E^{\text{Rep}}(S^{-2})$ are evaluated from the original EFP2 formulae of Jensen and Gordon, i.e., from Eqs. (2) and (4), respectively; (ii) $E^{\text{Rep}}(S^{-1})$ term is evaluated from Eq. (14) and assuming a small auxiliary basis sets, optimized for each EFP2 fragment separately.

ACKNOWLEDGMENTS

This project is carried out under POLONEZ programme which has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 665778. This project is funded by National Science Centre, Poland (grant no. 2016/23/P/ST4/01720) within the POLONEZ 3 fellowship. Wroclaw Centre for Networking and Supercomputing (WCSS) is acknowledged for computational resources.

[1] J. N. Murrell, M. Randić, D. R. Williams, and H. C. Longuet-Higgins, Proc. R. Soc. Lond. A **284**, 566 (1965).
[2] P. Otto and J. Ladik, Chem. Phys. **8**, 192 (1975).
[3] I. Hayes and A. Stone, Molecular Physics **53**, 83 (1984).
[4] B. Jeziorski, R. Moszyński, and K. Szalewicz, Chem. Rev. **94**, 1887 (1994).
[5] J. H. Jensen and M. S. Gordon, Mol. Phys. **89**, 1313 (1996).
[6] A. Heßelmann, G. Jansen, and M. Schütz, J. Chem. Phys. **122**, 014103 (2005).
[7] M. Mandado and J. M. Hermida-Ramón, J. Chem. Theory Comput. **7**, 633 (2011).
[8] M. S. Gordon, Q. A. Smith, P. Xu, and L. V. Slipchenko, Annu. Rev. Phys. Chem. **64**, 553 (2013).
[9] K. A. Nguyen, R. Pachter, and P. N. Day, J. Chem. Phys. **140**, 244101 (2014).
[10] P. N. Day, J. H. Jensen, M. S. Gordon, S. P. Webb, W. J. Stevens, M. Krauss, D. Garmer, H. Basch, and D. Cohen, J. Chem. Phys. **105**, 1968 (1996).
[11] C. C. J. Roothaan, Rev. Mod. Phys. **23**, 69 (1951).
[12] J. H. Jensen, J. Chem. Phys. **104**, 7795 (1996).
[13] J. H. Jensen and M. S. Gordon, J. Chem. Phys. **108**, 4772 (1998).
[14] B. Błasiak, J. D. Bednarska, M. Chołuj, and W. Bartkowiak, J. Chem. Phys. , Submitted (2019).
[15] R. M. Parrish, L. A. Burns, D. G. A. Smith, A. C. Simmonett, A. E. DePrince, III, E. G. Hohenstein, U. Bozkaya, A. Y. Sokolov, R. Di Remigio, R. M. Richard, J. F. Gonthier, A. M. James, H. R. McAlexander, A. Kumar, M. Saitow, X. Wang, B. P. Pritchard, P. Verma, H. F. Schaefer, III, K. Patkowski, R. A. King, E. F. Valeev, F. A. Evangelista, J. M. Turney, T. D. Crawford, and C. D. Sherrill, J. Chem. Theory Comput. **13**, 3185 (2017).
[16] G. Chałasiński and M. Gutowski, Mol. Phys. **54**, 1173 (1985).
[17] S. F. Boys, Rev. Mod. Phys. **32**, 296 (1960).
[18] Y. Zhao, N. E. Schultz, and D. G. Truhlar, J. Chem. Theory Comput. **2**, 364 (2006).
[19] Y. Zhao and D. G. Truhlar, J. Chem. Theory Comput. **1**, 415 (2005).
[20] Y. Zhao, N. E. Schultz, and D. G. Truhlar, J. Chem. Phys. **123**, 161103 (2005).
[21] L. A. Burns, J. C. Faver, Z. Zheng, M. S. Marshall, D. G. A. Smith, K. Vanommeslaeghe, A. D. MacKerell, K. M. Merz, and C. D. Sherrill, J. Chem. Phys. **147**, 161727 (2017).
[22] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, and D. J. Fox, "Gaussian16 Revision C.01," (2016), gaussian Inc. Wallingford CT.
[23] D. J. Wales, Energy Landscapes, (Cambridge University Press, Cambridge, UK, 2003).
[24] D. J. Wales and H. A. Scheraga, Science **285**, 1368 (1999).
[25] D. J. Wales and J. P. K. Doye, J. Phy. Chem. A **101**, 5111 (1997).
[26] Z. Li and H. A. Scheraga, Proc. Nat. Acad. Sci. **84**, 6611 (1987).
[27] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. Jarrod Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. Carey, İ. Polat, Y. Feng, E. W. Moore, J. Vand erPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, P. van Mulbregt, and S. . . Contributors, arXiv e-prints , arXiv:1907.10121 (2019), arXiv:1907.10121 [cs.MS].

Supplementary Information
Ab Initio Effective One-Electron Potential Operators. II. Applications for Exchange-Repulsion Energy in Effective Fragment Potentials

Bartosz Błasiak,[1, a)] Robert W. Góra,[1] and Wojciech Bartkowiak[1]
Department of Physical and Quantum Chemistry, Faculty of Chemistry, Wrocław University of Science and Technology, Wybrzeże Wyspiańskiego 27, Wrocław 50-370, Poland

(Dated: 4 February 2020)

CONTENTS

I. MINIMALISTIC AUXILIARY BASIS SETS FOR WATER, METHANOL AND DMSO MOLECULES

The basis sets are given below. Note that in case of DMSO, the basis set size is smaller than the minimal basis size (3p orbital on S atoms were not necessary to obtain good fitting in the test basis set).

```
[ water ]
cartesian
# primary basis set: 6-311++G** 6D
****
H     0
S   1   1.00
          0.500              1.000000
****
O     0
S   2   1.00
        937.182              0.986162
          6.892              0.014190
S   1   1.00
         45.156              1.000000
P   2   1.00
         24.823              0.277579
          2.783              0.301334
****

[ methanol ]
cartesian
# primary basis set: 6-311++G** 6D
****
H     0
S   1   1.00
          0.539              1.000000
****
C     0
S   1   1.00
        939.683              1.000000
S   1   1.00
         30.185              1.000000
P   1   1.00
          5.832              1.000000
P   1   1.00
          0.326              1.000000
****
O     0
S   1   1.00
        731.638              1.000000
S   1   1.00
         37.387              1.000000
P   1   1.00
          9.214              1.000000
P   1   1.00
          0.368              1.000000
****

[ dmso ]
cartesian
# primary basis set: 6-31++G** 6D
****
H     0
S   1   1.00
          0.365              1.000000
****
C     0
S   1   1.00
        207.472              1.000000
S   1   1.00
         15.052              1.000000
P   1   1.00
          7.228              1.000000
****
O     0
S   1   1.00
        302.339              1.000000
S   1   1.00
         15.248              1.000000
P   1   1.00
          7.612              1.000000
****
```

a) blasiak.bartosz@gmail.com; https://www.polonez.pwr.edu.pl

```
S     0
S   1   1.00
      689.986             1.000000
S   1   1.00
      145.419             1.000000
S   1   1.00
        9.046             1.000000
P   1   1.00
       28.979             1.000000
****
```

II. ACCURACY OF OEP AND EFP2 MODELS FOR EXCHANGE-REPULSION ENERGY ESTIMATION

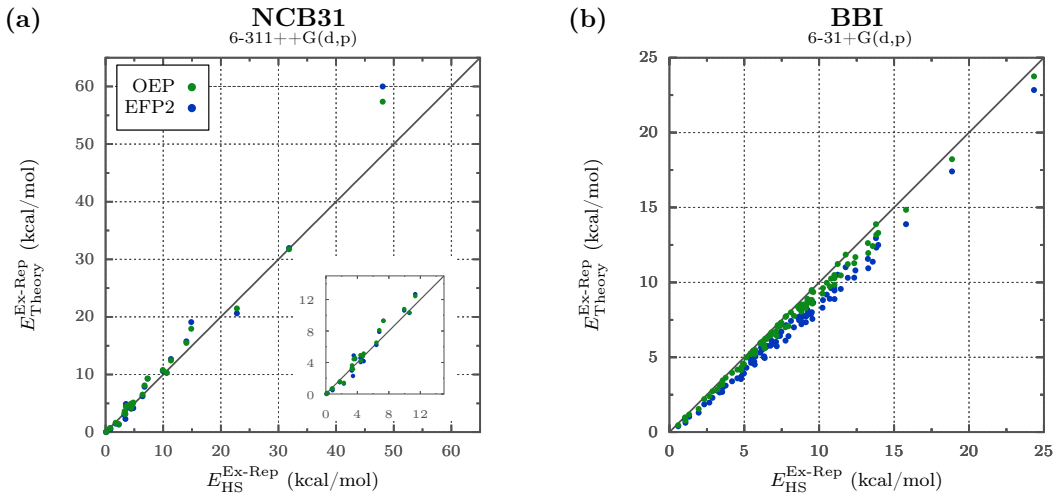

FIG. S1. **Accuracy of the OEP and EFP2 models of exchange-repulsion energy across various bi-molecular systems.** (a) NCB31 database[1–3] of non-covalent interactions and (b) BBI subset[4] of backbone-backbone interactions in proteins from the BioFragment Database. For the OEP calculations, the EDF-1 scheme with the aug-cc-pVDZ-jkfit auxiliary basis set was used.

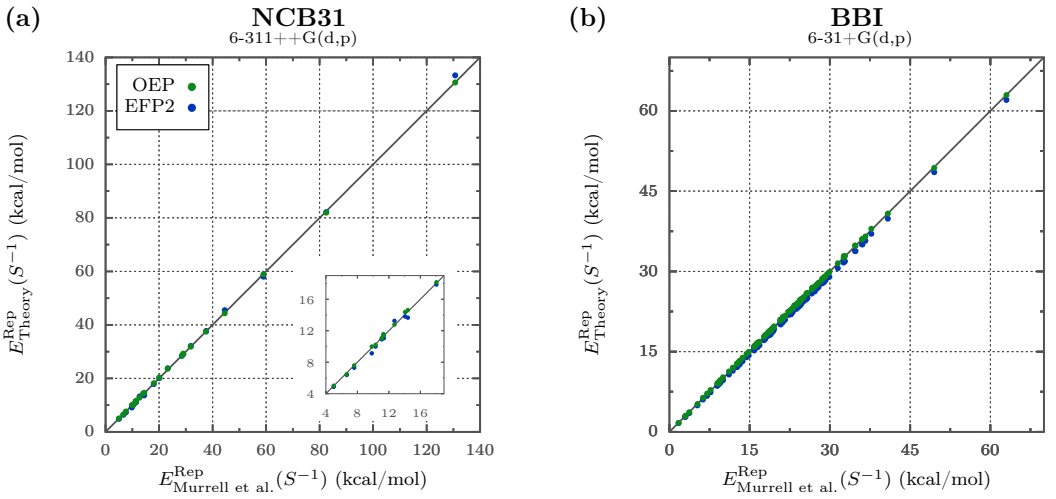

FIG. S2. **Accuracy of the OEP and EFP2 models of first-order repulsion energy across various bi-molecular systems.** (a) NCB31 database[1–3] of non-covalent interactions and (b) BBI subset[4] of backbone-backbone interactions in proteins from the BioFragment Database. For the OEP calculations, the EDF-1 scheme with the aug-cc-pVDZ-jkfit auxiliary basis set was used.

III. REFERENCES

[1]Y. Zhao, N. E. Schultz, and D. G. Truhlar, J. Chem. Theory Comput. **2**, 364 (2006).
[2]Y. Zhao and D. G. Truhlar, J. Chem. Theory Comput. **1**, 415 (2005).
[3]Y. Zhao, N. E. Schultz, and D. G. Truhlar, J. Chem. Phys. **123**, 161103 (2005).
[4]L. A. Burns, J. C. Faver, Z. Zheng, M. S. Marshall, D. G. A. Smith, K. Vanommeslaeghe, A. D. MacKerell, K. M. Merz, and C. D. Sherrill, J. Chem. Phys. **147**, 161727 (2017).