RESEARCH ARTICLE | SEPTEMBER 26 2023

### Electron density-based protocol to recover the interacting quantum atoms components of intermolecular binding energy 🕗

Aleksei A. Anisimov 🖾 💿 ; Ivan V. Ananyev 💿

Check for updates

J. Chem. Phys. 159, 124113 (2023) https://doi.org/10.1063/5.0167874



**CrossMark** 



APL Quantum Bridging fundamental quantum research with technological applications

Now Open for Submissions No Article Processing Charges (APCs) through 2024



**Submit Today** 



View Online

ARTICLE

ГŢJ

Export Citation

# Electron density-based protocol to recover the interacting quantum atoms components of intermolecular binding energy

Cite as: J. Chem. Phys. 159, 124113 (2023); doi: 10.1063/5.0167874 Submitted: 15 July 2023 • Accepted: 11 September 2023 • Published Online: 26 September 2023

Aleksei A. Anisimov<sup>1,2,a)</sup> D and Ivan V. Ananyev<sup>3</sup>

#### **AFFILIATIONS**

<sup>1</sup>A.N. Nesmeyanov Institute of Organoelement Compounds, Russian Academy of Sciences, Vavilova Str. 28, GSP-1, Moscow 119334, Russian Federation

<sup>2</sup>National Research University Higher School of Economics, Myasnitskaya Str. 20, Moscow 101000, Russian Federation

<sup>3</sup>N.S. Kurnakov Institute of General and Inorganic Chemistry, Russian Academy of Sciences, GSP-1, Leninsky prospect 31, Moscow 119991, Russian Federation

<sup>a)</sup>Author to whom correspondence should be addressed: anisimov.alex.a@gmail.com

#### ABSTRACT

A new approach for obtaining interacting quantum atoms-defined components of binding energy of intermolecular interactions, which bypasses the use of standard six-dimensional integrals and two-particle reduced density matrix (2-RDM) reconstruction, is proposed. To examine this approach, three datasets calculated within the density functional theory framework using the def2-TZVP basis have been explored. The first two, containing 53 weakly bound bimolecular associates and 13 molecular clusters taken from the crystal, were used in protocol refinement, and the third one containing other 20 bimolecular and three cluster systems served as a validation reference. In addition, to verify the performance of the proposed approach on an exact 2-RDM, calculations within the coupled cluster formalism were performed for part of the first set systems using the cc-pVTZ basis set. The process of optimization of the proposed parametric model is considered, and the role of various energy contributions in the formation of non-covalent interactions is discussed with regard to the obtained trends.

Published under an exclusive license by AIP Publishing. https://doi.org/10.1063/5.0167874

#### I. INTRODUCTION

Owing to the significant impact on the stability and properties of macromolecules and condensed phases, intermolecular interactions are fruitful objects of numerous studies.<sup>1-6</sup> In this respect, the role of theoretical modeling of such systems cannot be overestimated; at the present time, the solution of almost every problem in the corresponding areas of scientific research is based on either the preliminary prediction of system's behavior or a posteriori analysis of experimental data. However, the full-scale modeling of large chemical entities stabilized by intermolecular forces still remains a challenging task if the *ab initio* footing of quantum chemistry is utilized. This enables a root for the simplified theoretical models to be used in practical purposes of the description of intermolecular interactions.

The role of theoretical chemistry in this respect is to provide information on the transferability of an interaction between

different systems regarding its influence on their stability and properties. The transferability concept implies a decomposition of the whole intermolecular interaction into chemically relevant, binding contributions known as non-covalent interactions. Analyzing the strength and nature of these contributions is the ultimate goal of countless approaches that differ in both the definition of the binding interaction and the metrics used.

One of the most rigorous and yet intuitive theoretical methods that can be used to analyze intermolecular interactions is the combination of topological partitioning of 3D coordinate space7 and extracting electronic molecular Hamiltonian contributions within McWeeney's separability theory.<sup>8</sup> The topological partitioning onto atomic basins in the real space is usually performed within the Quantum Theory of Atoms in Molecules (QTAIM),<sup>9</sup> i.e., by means of the analysis of the electron density function  $\rho(\mathbf{r})$ . The QTAIM method is of a particular interest for the analysis of intermolecular interactions as it allows us to trace the most important bonding two-center interactions corresponding to so-called "privileged exchange interaction channels."<sup>10</sup> That is, among all atomic pairs from different molecules, QTAIM highlights those with the most pronounced covalent character, which ensures the directionality of an interaction<sup>11</sup> and, hence, the transferability of its structural characteristics. In its turn, McWeeney's theory combined with the integration over topological basins provides a route to decompose the energy of interaction between any two chemically meaning-ful fragments of space into the physically defined one-particle and two-particle contributions. Such a treatment is widely known as the Interacting Quantum Atoms (IQA)<sup>12,13</sup> approach and serves as the real space analog of commonly used energy decomposition schemes,<sup>14–17</sup> yet allowing us to analyze the interaction's nature in the orbital invariant manner.

Although the IQA method is successfully used in solving various tasks, including those emerged in studies of intermolecular interactions,<sup>18-22</sup> its applicability to large systems struggles from time-consuming computations. This remains a problem even when invoking approximations to the most resource-intensive twoparticle contributions; the latter can be rigorously achieved only if the exact two-particle reduced density matrix (2-RDM) is known. The approximations for 2-RDM commonly used<sup>23,24</sup> are based on the treatment of the full one-particle reduced density matrix (1-RDM) that raises two important problems. First, the accuracy of the approximations is questioned even if the high rung post-HF wavefunction theory is under consideration.<sup>25,26</sup> Second, the requirement for 1-RDM to be computed does not allow us to reduce calculation time sufficiently and prevents from using the IQA decomposition when only the trace of 1-RDM  $[\rho(r)]$  is known. This eliminates the possibility to investigate probably the most important systems with intermolecular interactions-crystalline objects. For crystals, the  $\rho(\mathbf{r})$  function and its derivatives (computed, for example, by means of periodic plane wave codes or reconstructed from high-resolution x-ray diffraction datasets) are usually the only available descriptors of electronic structure.

This paper aims to overcome these problems by providing a protocol to estimate IQA contributions into the binding energy of intermolecular interactions  $E_{bind}$  using the information about the  $\rho(\mathbf{r})$  field only. Namely, three main IQA contributions into  $E_{bind}$ (deformation energy and Coulombic and exchange-correlation in situ energies) are approximated on the basis of electronic population of intermolecular surface corresponding to bonding between topological fragments<sup>27</sup> and the widely known multipole expansion of the Coulombic term.<sup>28</sup> The performance of the suggested protocol is tested by means of the regression analysis for three model sets of associates stabilized by intermolecular interactions. The more representative set of molecular dimers (further referred to as DS) stabilized by different types of non-covalent interactions is composed of associates from the well-known S66<sup>29</sup> and X40<sup>30</sup> sets with the addition of a few systems to obtain a more uniform intermolecular energies distribution. The second set (hereinafter referred to as CS) consists of several molecular clusters cut from the corresponding optimized crystal structures and is designed to test interactions formed by a caged central molecule with its supramolecular environment, i.e., to emulate forces that stabilize packing of molecules in a condensed state, such as crystals. While these two sets were used for both tuning and testing the protocol, the third verification set (VS, 20 bimolecular systems and three molecular clusters) was utilized to provide an external validation of the protocol's performance when using the parameters fitted against the **DS** and **CS** datasets. To reduce the computational cost, the IQA contributions for both sets were approximated by means of the DFT scaling technique.<sup>31</sup> However, to verify the reliability of the proposed protocol, the exact two-particle IQA terms were calculated for some associates from **DS** using the coupled cluster wavefunction.

The discussion is structured as follows. First, the IQA-based decomposition of the intermolecular interaction energy is briefly recalled (for a more descriptive consideration of IQA theory, the reader is referred to the recent review<sup>13</sup>) and the approximations of corresponding IQA contributions are discussed. Then, the approaches used to calculate the geometry,  $\rho(\mathbf{r})$  function, and IQA terms for systems in **DS** and **CS** are detailed. Finally, the results of regression analysis invoked to optimize and verify the proposed protocol are summarized.

#### **II. THEORETICAL BACKGROUND**

#### A. Basics of the IQA method

As mentioned in Introduction, the essence of the QTAIMbased IQA method is to combine the exhaustive  $\rho(r)$ -based topological partitioning of a system onto atomic basins ( $\Omega$ ) and the reduced density matrix formalism in order to decompose the total electronic energy into all possible intra- and interatomic contributions,

$$E = \sum_{\Omega} E_{self}^{\Omega} + \frac{1}{2} \sum_{\Omega} \sum_{\Omega' \neq \Omega} E_{int}^{\Omega\Omega'}.$$
 (1)

The additivity of contributions for each  $\Omega$  enables a simple procedure to calculate the energy of any multiatomic fragment. By setting  $\Omega$  and  $\Omega'$  as unions of basins composing two disjoint multiatomic fragments A and B, Eq. (1) can be applied to calculate not only so-called self-energies of these fragments but also the potential energy of interaction between them in their *in situ* states. The latter is known<sup>32</sup> as the *in situ* interaction energy  $E_{in situ}$  and is simply equal to  $E_{int}^{AB}$  (the sum of corresponding  $E_{int}^{\Omega\Omega'}$  terms).

The RDM formalism allows us to further decompose each contribution from (1) into one- and two-particle terms. However, for practical purposes of reducing computational cost, the decomposition of  $E_{int}^{\Omega\Omega'}$  is of more interest as it allows us to elucidate the nature of an interfragment interaction by performing the IQA calculations only for associates in their bonded, *in situ* state of interest. Indeed, all terms of the  $E_{int}^{AB}$  contribution can be grouped by their dependence on 1-RDM and 2-RDM to provide corresponding classical Coulombic and exchange–correlation energies,

$$E_{\text{int}}^{\text{AB}} = V_C^{\text{AB}} + V_{xc}^{\text{AB}} = \sum_{\Omega \in \mathcal{A}} \sum_{\Omega' \in \mathcal{B}} V_C^{\Omega \Omega'} + \sum_{\Omega \in \mathcal{A}} \sum_{\Omega' \in \mathcal{B}} V_{xc}^{\Omega \Omega'}.$$
 (2)

The  $V_{C}^{AB}$  contribution is responsible for the classical electrostatic component and can be computed exactly from the corresponding charge densities,

$$V_{C}^{AB} = \sum_{\Omega \in \Lambda} \sum_{\Omega' \in B} \int_{\Omega} d\mathbf{r}_{1} \int_{\Omega'} d\mathbf{r}_{2} \\ \times \frac{\{Z_{\Omega} \delta(\mathbf{r}_{1} - \mathbf{R}_{\Omega}) - \rho(\mathbf{r}_{1})\} \cdot \{Z_{\Omega'} \delta(\mathbf{r}_{2} - \mathbf{R}_{\Omega'}) - \rho(\mathbf{r}_{2})\}}{\mathbf{r}_{12}}.$$
 (3)

In its turn, the non-classical  $V_{xc}^{AB}$  contribution in (2) formally depends on the exchange–correlation part of 2-RDM  $\rho_{xc}(r_1, r_2)$ ,

$$V_{xc}^{AB} = \sum_{\Omega \in A} \sum_{\Omega' \in B} \int_{\Omega} d\mathbf{r}_1 \int_{\Omega'} d\mathbf{r}_2 \frac{\rho_2^{XC}(\mathbf{r}_1, \mathbf{r}_2)}{\mathbf{r}_{12}}.$$
 (4)

The  $V_{xc}^{AB}$  term is known<sup>33</sup> to be connected with the delocalization index,<sup>34</sup> which basically demonstrates how much the electrons of two atoms are prone to association, thereby serving as the realspace analog of bond order.<sup>35</sup> In other words, the  $V_{xc}^{AB}$  term provides information on covalent character of the interaction between A and B.<sup>36</sup>

Combining decompositions (1) and (2), one can write the binding energy  $E_{bind}$  (measurable dissociation energy minus zero-point terms and relaxation of nuclei) in the following manner:<sup>32</sup>

$$E_{bind} = E_{def} + E_{in \ situ} = E_{def} + V_C^{AB} + V_{xc}^{AB}.$$
 (5)

In the context of intermolecular interactions, A and B correspond to the unions of basins from different molecules and  $E_{def}$  denotes the change in self-energy of these molecules (including intra- and interatomic terms) occurring upon the formation of intermolecular interaction.

The IQA analysis can be very time consuming for large systems, mostly due to high resource-intensity of two-particle and, especially, diatomic terms computation.<sup>13</sup> Some previously developed tricks<sup>27</sup> and observations will be discussed in Sec. II B with the aim to make substantial CPU time savings for the IQA analysis of intermolecular interactions based on decomposition (5).

# B. Electron density-based approximations of IQA contributions into the intermolecular binding energy

Evidently, the exchange-correlation contribution is the most problematic term of any energy decomposition. As mentioned, it requires the knowledge of 2-RDM that formally restricts the applicability of the IQA approach only to wave function theories and forces one to handle rather voluminous matrices. To avoid these inconveniences, a number of solutions were developed for both density functional theory (DFT) and wavefunction theory (WFT, post-HF) methods. For the Kohn-Sham version of DFT, the exchange-correlation IQA contributions are commonly evaluated using atomic scaling factors<sup>31,37</sup> (though the bond order density approach is also known<sup>38,39</sup>) with the  $V_{xc}^{\Omega\Omega'}$  term estimated from the Hartree-Fock exchange integrals over Kohn-Sham orbitals. In contrast to the fully additive DFT IQA schemes, the known 2-RDM approximations for the post-HF methods struggle to provide the correct value of total electronic energy and are based on the reduced density matrix functional theory (RDMFT<sup>40,41</sup>) utilizing natural orbitals and their occupation numbers.<sup>23,24</sup> Finally, the multipolar approach  $V_{xc}^{\Omega\Omega'}$  is also known with the delocalization index (still depending on 2-RDM) being the first term of the expansion.<sup>2</sup>

Anyhow, all mentioned approximations are based on the consideration of at least one-particle density matrix that prevents IQA to be used in any orbital-free description of electronic structure. Recently,<sup>27</sup> inspired by the exchange–correlation rationalization<sup>10</sup> of bonding between QTAIM topological atoms and the similarity between atomic surface terms of the hypervirial theorem<sup>42–44</sup> and delocalization indices, we proposed the following model for the interatomic  $V_{xc}^{\Omega\Omega'}$  contribution, which is solely based on the peculiarities of  $\rho(\mathbf{r})$  function:

$$V_{xc}^{\Omega\Omega'} \cong \frac{k}{|\boldsymbol{R}_{\Omega\Omega'}|^{p_2}} \oint_{S} ds(\boldsymbol{r}) \boldsymbol{R}_{\Omega\Omega'} \cdot \boldsymbol{n}(\boldsymbol{r}) \rho(\boldsymbol{r})^{p_1} + b.$$
(6)

Here, *k*, *b*, *p*<sub>1</sub>, and *p*<sub>2</sub> are the parameters obtained by a fitting procedure and  $\mathbf{R}_{\Omega\Omega'}$  is the internuclear vector; the integration is performed over the zero-flux surface of  $\rho(\mathbf{r})$  separating two topological basins. The reliability of this model was validated for diatomic systems only where it allows us to reproduce  $V_{xc}^{\Omega\Omega'}$  (at least, approximated by means of RDMFT) with sufficiently high accuracy.<sup>27</sup> Its applicability in the case of intermolecular interactions will be shown in this paper based on the following extension to the interfragment energy:

$$V_{xc}^{AB} \cong b + k \left( \sum_{S(\Omega \mid \Omega')} \frac{1}{|\boldsymbol{R}_{\Omega \Omega'}|^{p_2}} \oint_{S(\Omega \mid \Omega')} ds(\boldsymbol{r}) \boldsymbol{R}_{\Omega \Omega'} \cdot \boldsymbol{n}(\boldsymbol{r}) \rho(\boldsymbol{r})^{p_1} \right),$$
(7)

where the summation is carried out for all interatomic surfaces between atoms  $\Omega$  and  $\Omega'$  denoting QTAIM bonding contributions and composing the whole intermolecular surface. Hereafter, this sum will be referred to as *integral* for simplicity.

The next component of  $E_{bind}$  to be approximated, namely, the deformation energy  $E_{def}$ , is responsible for the rearrangements of electronic structure upon the formation of interaction and describes charge redistribution as well as changes of spin states and kinetic energies of the fragments. This term is no less problematic to evaluate due to the need of additional computations of the free, non-interacting fragments. Luckily, a possible interrelation between the deformation energy and  $V_{xc}^{AB}$  was noticed previously for intermolecular interactions (see, for instance, Ref. 45). The authors of the mentioned paper stated that these terms "virtually cancel" each other for the interactions of LP... $\pi$  type; in fact,  $V_{xc}^{AB}$  and  $E_{def}$  were found to be nearly equal in absolute value. It can be assumed that this tendency can be extended to other types of non-covalent interactions. The analysis of IQA terms for some bimolecular datasets published previously,<sup>46,47</sup> indeed, demonstrates that the  $V_{xc}^{AB} \cong -\hat{E}_{def}$  trend holds for various types of non-covalent interactions except for those with a predominantly covalent character (i.e., with a large magnitude of  $V_{xc}^{AB}$ : GeO<sub>2</sub>···Br<sup>-</sup>, GeH<sub>3</sub>F···N<sub>3</sub><sup>-</sup>, SiO<sub>2</sub>···CO, GeH<sub>4</sub>···F<sup>-</sup>,  $SiO_2 \cdots NCH$ ,  $SiH_3F \cdots N_3^-$ ,  $SiO_2 \cdots Br^-$ ,  $SiO_2 \cdots CS$ ,  $SiH_4 \cdots F^-$ ; see Fig. S1).

Keeping in mind that the most important reason for the electronic structure deformation occurred upon the formation of weak interactions (such as intermolecular ones) is the redistribution of electronic charge; the interrelation between  $V_{xc}^{AB}$  and  $E_{def}$  terms can be rationalized as the consequence of  $\rho(\mathbf{r})$  behavior. The  $V_{xc}^{\Omega\Omega'}$  terms are linked with the delocalization phenomenon, which is always a two-side coin: the delocalization of electrons between two space basins is related to the intrabasin localization of electronic population (the situation that is commonly observed upon the formation of intermolecular interactions), an increase in the delocalization

index lowers the localization one. In the same manner, according to the stationary atomic (fragment) hypervirial theorem, the emanation of a new intermolecular surface is always the reflection of changes occurring within the atomic basins. The surface integral in (6) approximating the  $V_{xc}^{\Omega\Omega'}$  terms can be seen as the measure of these rearrangements as it is linked with the basin integrals of  $\mathbf{r}_{\Omega} \cdot \nabla \rho(\mathbf{r}_{\Omega})$  (where  $\mathbf{r}_{\Omega}$  is a position vector centered on nucleus  $\Omega$ ) at constant electronic populations from the perspective of the divergence theorem, which describes a special case of the hypervirial theorem for the  $\widehat{\mathbf{r}_{\Omega}^2}$  density momentum operator.<sup>48</sup> These observations allow us to use the r.h.s. of (6) and (7) (probably, with the reparameterization to increase its accuracy) to evaluate  $E_{def}$  values for intermolecular interactions.

Finally, the electrostatic contribution is to be discussed. While it can be calculated directly from the charge density [see Eq. (3)], the computations of corresponding six-dimensional integrals are still too consuming. The most known and widely used approximation to  $V_C^{\Omega\Omega'}$  deals with the multipolar approach utilizing a binomial Taylor expansion of  $\mathbf{r}_{12}^{-1} = |\mathbf{R}_{\Omega\Omega'} - (\mathbf{r}_1 - \mathbf{r}_2)|^{-1} (|\mathbf{R}_{\Omega\Omega'}|$  is the internuclear distance). The applicability of this approximation was thoroughly studied (in particular, by the group of Popelier<sup>28,49–51</sup>) to reconstruct the atomic electrostatic potentials and the  $V_C^{\Omega\Omega'}$  energies with a less than 1 kcal/mol accuracy. In particular, a fast pseudo-convergence for the  $V_C^{\Omega\Omega'}$  expansions was observed for atomic pairs having either low values of electron density within the interatomic region or large magnitudes of  $R_{\Omega\Omega'}$  that enabled an extensive use of this approximation in studies of intermolecular interactions and even in developments of new force fields for long-range interactions.<sup>52</sup>

Although there are at least two known equivalent formulations of  $V_C^{\Omega\Omega'}$  expansions (in spherical harmonics<sup>28,53</sup> and Cartesian tensors), in this work, we focused on the Buckingham-type expansion<sup>54</sup> of electrostatic energy calculated for each pair of topological atoms,

$$V_{C}^{\Omega\Omega'} = Tq^{\Omega}q^{\Omega'} + T_{\alpha}\left(q^{\Omega}\mu_{\alpha}^{\Omega'} - q^{\Omega'}\mu_{\alpha}^{\Omega}\right) + T_{\alpha\beta}\left(\frac{1}{3}q^{\Omega}\Theta_{\alpha\beta}^{\Omega'} + \frac{1}{3}q^{\Omega'}\Theta_{\alpha\beta}^{\Omega} - \mu_{\alpha}^{\Omega}\mu_{\beta}^{\Omega'}\right) + T_{\alpha\beta\gamma}\left(\frac{1}{15}q^{\Omega}\Phi_{\alpha\beta\gamma}^{\Omega'} - \frac{1}{15}q^{\Omega'}\Phi_{\alpha\beta\gamma}^{\Omega} - \frac{1}{-\frac{1}{3}\mu_{\alpha}^{\Omega}\Theta_{\beta\gamma}^{\Omega'} + \frac{1}{3}\mu_{\alpha}^{\Omega'}\Theta_{\beta\gamma}^{\Omega}}\right) + \cdots, \qquad (8)$$

where  $q^{\Omega}, \mu^{\Omega}_{\alpha}, \Theta^{\Omega}_{\alpha\beta}, \Phi^{\Omega}_{\alpha\beta\gamma}$  are the charge, dipole, quadrupole, and octupole traceless moments of the atomic basin  $\Omega$  (defined solely in terms of moments of electron density distribution) and *T* is the corresponding symmetric tensor  $(T_{\alpha\beta\gamma\delta...} = \nabla_{\alpha}\nabla_{\beta}\nabla_{\gamma}\nabla_{\delta}...R^{-1})$ .

#### **III. COMPUTATIONAL METHODS**

### A. Models, quantum chemistry, and IQA computations

Two test sets of model species and their combinations were considered for searching, checking, and analyzing the proposed trends (see Tables S1 and S2), while the third set (Table S3) is constructed to validate the performance of the proposed protocol on systems not included in the parameter optimization procedure. The weakly bonded dimer set (referred to as **DS**) with various types of intermolecular interactions consists of systems from the well-known S66<sup>29</sup> and X40<sup>55</sup> sets, and few new systems were also added in order to obtain a more uniform distribution of intermolecular energies. Quantum chemical calculations were carried out using two methods: the PBE0 functional,<sup>56</sup> which is considered the gold standard in density functional theory,<sup>57</sup> and the coupled-cluster with single and double exitations (CCSD) theory,<sup>58</sup> which is considered the gold standard among post-HF methods.

In the former case, the calculation was carried out using the Gaussian 09 (rev D03) program,<sup>59</sup> and the standard Ahlrich's basis def2-TZVP was used.<sup>60</sup> The optimal geometry was found for each dimer with the use of the tight convergence criteria and the Grimme's dispersion correction<sup>61</sup> (Becke–Johnson damping<sup>62</sup>). All structures from **DS** correspond to the energy minima according to the normal mode calculations (ultrafine grids were used). The QTAIM and DFT IQA analysis was performed using the AIMAll program.<sup>63</sup>

In the case of CCSD calculations, the geometry optimization was performed in the Gaussian program, whereas the exact unrelaxed second-order density matrices were obtained for the resulting geometries using the MRCC program<sup>64,65</sup> at the facilities of JRC PMR IGIC RAS. Due to the increased computational cost, the analyzed set was reduced, and the cc-pVTZ<sup>66</sup> basis set was chosen. For the exact RDMs, the QTAIM and IQA computations were performed using the TWOE program.<sup>67–69</sup> This shortened set will be further referred to as **SS**.

The second set consists of so-called caged systems (further referred to as CS; see Table S2), which are represented by molecular clusters cut from the corresponding optimized crystal structures.<sup>70</sup> For these systems, the following procedures were conducted. First of all, the full crystal structure relaxation was initially performed using the CRYSTAL17<sup>71,72</sup> software at the PBE0-D3/def2TZVP<sup>60</sup> level. Two approaches were then investigated using the same level of theory in the Gaussian 09 (rev D03) program:<sup>59</sup> the most straightforward one included the generation of wavefunctions for isolated molecular clusters based on CRYSTAL17 fixed geometry data and the second one involved additional geometry optimization of each central molecule in cluster in order to obtain more reasonable  $E_{def}$ values. In both cases, the QTAIM and IQA analysis was performed by means of the AIMAll program.<sup>63</sup> Both strategies generally showed similar values and trends (Fig. S2), but for some systems, the results for the second approach were somewhat inadequate with highly overestimated values of the energy components (see Table S2)-for this reason, it was decided to focus on the first approach. The  $E_{int}^{AB}$ ,  $V_C^{AB}$ , and  $V_{xc}^{AB}$  contributions were calculated as the

The  $E_{int}^{AB}$ ,  $V_C^{AB}$ , and  $V_{xc}^{AB}$  contributions were calculated as the sums of corresponding pair energies between atoms of two different molecules (**DS** and **SS**) and between atoms of the central molecule and surrounding molecules (**CS**). Due to the unfeasibility of straightforward variational calculations of  $E_{bind}$  and  $E_{def}$  values for **CS** (as the clusters without central molecules cannot be considered as free of intermolecular interactions), the IQA-based approximation was used. Recently, the values of  $E_{bind}$  calculated using this simple approach have been found to be in rather good agreement with the crystal lattice energies  $E_{coh}$ .<sup>70</sup> Namely, the  $E_{bind}$  and  $E_{def}$  terms were estimated as the differences of corresponding energies (the total electronic one for  $E_{bind}$  and the self-energy for  $E_{def}$ ) of a central molecule in its bounded (in cluster) and isolated states, multiplied by two to scale the resulting molecular energy to the

bimolecular set level. Indeed, while the interaction energy and its components for one molecule need to be multiplied by a number of molecules in crystal to model extensive specific quantity, a proper comparison between one-molecule (central molecules from **CS**) and two-molecule (**DS**) energetic values requires the multiplication of the formers by two.

The third set (referred to as **VS**) is intended to verify the parameters obtained for **DS** and **CS**. It includes 20 systems from D1200,<sup>73</sup> SH250,<sup>74</sup> R739,<sup>75</sup> HB375,<sup>76</sup> and HB300SPX<sup>77</sup> sets as well as three caged systems (see Table S3) in order to obtain a representative range of interaction types and energies. All the quantum chemical and IQA calculations for this set were carried out in a completely analogous manner to the **DS** and **CS** systems.

#### **B.** Processing of interatomic surfaces

The calculations of the surface integrals in the r.h.s. of expression (7) were performed using a self-written script based on the modified Multiwfn program<sup>78</sup> and approximating IASes numerically (due to the absence of analytical solutions for IAS<sup>79</sup>). The standard algorithm of IAS processing was slightly changed to increase numerical integration accuracy by avoiding the widely known problem of "free space" arising for weak and non-directional noncovalent interactions. In particular, this issue was faced for many systems from the **CS** when the standard Multiwfn algorithm was used (see, for instance, Fig. S3). Note that no such problem was observed for the skeletal trajectories from AIMAII used to perform the IQA computations (see above); a trivial test on the sanity of the resulting volume integrals is passed (atomic Lagrangian values were of  $10^{-5}$  order in magnitude for all systems). The modified algorithm is as follows:

- (1) Generation of skeletal trajectories using the bisection method starting at the (3, -1) critical point of  $\rho(\mathbf{r})$ . For each surface, there were 100 trajectories of 200 points with 0.03 a.u. step (see Fig. S4).
- (2) Obtaining an integration grid by specific partitioning of sectors between two neighboring skeletal trajectories. First, the spaces between the skeletal trajectories are divided into triangles in such a way that the added section between the points has minimal length. Next, for each triangle, if the length of added section exceeds the distance between the points in the skeletal trajectories, an additional partitioning is performed, leading to objects with a smaller area—tetragons and triangle, the centers of which are the integration grid (see Fig. S5).

The accuracy of this algorithm was verified by comparison with an external standard; its role was played by "flat" integrals  $\oint_S ds(\mathbf{r})\rho(\mathbf{r})$  generated by means of the AIMAll program for the **DS**. These integrals were also calculated using the described approach, and the values turned to be quite similar (see Table S4) with the mean absolute deviation being less than 0.0011 a.u. Thus, the proposed algorithm for calculating surface integrals can be considered as sufficient for the purposes of this paper.

#### C. Optimization of protocol parameters

The optimization of parameters of the model (7) was made by an iterative scheme used in the previous work:<sup>27</sup> the parameters  $p_1$ and  $p_2$  were optimized (starting from 1.0 to 1.0, respectively) by the Broyden–Fletcher–Goldfarb–Shanno (BFGS) method<sup>80–83</sup> using the SciPy library in order to obtain the highest coefficient of determination for a linear trend with coefficients k and b. Hence, at each step, for the parameters  $p_1$  and  $p_2$ , the coefficients k and b were linearly adjusted by the least squares method and the corresponding coefficient of determination was calculated, which later served as the value of the function of arguments  $p_1$  and  $p_2$  maximized by the BFGS method. For non-linear optimization, the dimensionless value of  $5 \times 10^{-3}$  was used as the numerical gradient cutoff.

In order to obtain the Coulomb part of intermolecular binding energy without the direct six-dimensional integrations, we decided to rely on the multipole expansion in terms of Buckinghamtype Cartesian tensors,<sup>54</sup> which were calculated for QTAIM atomic basins up to and including the  $1/R_{AB}^{-7}$  terms (i.e., including all dipole–quadrupole and charge–octupole energies) via the modified Multiwfn program.<sup>78</sup>

#### IV. THE VERIFICATION OF THE PROPOSED PROTOCOL

### A. Exchange-correlation contribution into *in situ* intermolecular energy

We start from checking the performance of the proposed parametric model (7) to estimate  $V_{xc}^{AB}$  in the case of intermolecular interactions. The regression based on the optimized  $p_1$  and  $p_2$  parameters demonstrates an excellent quality for both the **DS** and **CS** (see Table I and Figs. 1 and S6). Moreover, the observed dependence is also preserved if the exact 2-RDMs are used (i.e., for the **SS**), albeit with slightly different resulting parameters. The obtained mean absolute errors (MAEs) for the model with the optimal parameters are as follows: 0.35 kcal/mol (**DS**), 1.66 kcal/mol (**CS**), and 0.35 kcal/mol (**SS**).

Two important notes should be made regarding the observed trends. First, the reason for the difference in the parameters obtained by the optimization for **DS** and **SS** (especially in values of *k*) is to be mentioned. This difference is clearly due to the approximations used to calculate interatomic exchange–correlation energies from DFT (see discussion above): the Hartree–Fock functional underestimates correlation effects. Second, the optimized parameters in the case of **SS** (Fig. 1) are somewhat different from those for diatomic systems<sup>27</sup> ( $p_1 = 1.113$  and  $p_2 = 1.094$  with k = -0.416 and b = -0.001 for diatomics;  $p_1 = 1.078$  and  $p_2 = 1.028$  with k = -0.301 and b = -0.0007 for **SS**). Although this difference is negligible especially taking into account the use of Müller<sup>23</sup> approximation for diatomics, there may be several additional reasons, including the rather small size of the **SS**. Moreover, it should be noted that the

**TABLE I.** Data for the linear trends between  $V_{xc}^{AB}$  and model (7) (at starting and optimized parameters).

	,				
Test set	$p_1$	$p_2$	k	Ь	$R^2$
DS	1.0	1.0	-0.202	0.0036	0.9829
DS	1.137	1.053	-0.389	0.0006	0.9975
CS	1.0	1.0	-0.198	0.0315	0.9338
CS	1.195	1.013	-0.430	-0.0105	0.9945
SS	1.0	1.0	-0.204	0.0019	0.9935
SS	1.078	1.028	-0.301	0.0007	0.9969

### The Journal of Chemical Physics

ARTICLE

pubs.aip.org/aip/jcp



**FIG. 1.** The obtained linear trend between minus intermolecular exchange–correlation energy  $-V_{ee\ XC}^{AB}$  and the *integral* for **DS** and **SS** (top) and for **CS** (bottom). The parameters  $p_1$  and  $p_2$  are taken from the optimization. All quantities are given in a.u.

deviation of the  $p_1$  and  $p_2$  parameters from the optimized values affects the coefficient of determination only slightly. The numerical gradient magnitude for the last steps of optimization does not exceed  $1 \times 10^{-2}$ , which indicates a gentle slope of  $R^2$  as a function of parameters near  $R^2 = 0.99$ . Nevertheless, the performance of the trend for **SS** together with its similarity with that for diatomics demonstrates the reliability of the model (7) for calculations of the  $V_{xc}^{AB}$ component for intermolecular interactions.

Finally, the results for the **CS** and **DS** show a remarkable similarity that allows us to use the average values of parameters to estimate  $V_{xc}^{AB}$  for both the isolated and crystalline systems on the level of PBE0 method accuracy. Indeed, the optimization of (7) for the combined set **DS+CS** results in  $p_1 = 1.171$  and  $p_2 = 1.044$  with the MAE value equal to 0.82 kcal/mol (see Fig. 2).

Moreover, the external validation of model (7) confirms its robustness. Namely, its use with the parameters obtained for the **DS+CS** demonstrates a preservation of the protocol performance for the **VS**: the observed coefficients of linear regression coincide with those obtained earlier, and  $R^2$  retains a high value with the MAE value equal to 1.27 kcal/mol (see Fig. 3).

#### B. Deformation energy contribution

In full agreement with the assumption that the deformation energy should be compensated by the  $V_{xc}^{AB}$  term for any weak



**FIG. 2.** The obtained linear trend between minus intermolecular exchange–correlation energy  $-V_{eo\ XC}^{\Omega\Omega'}$  and the *integral* for the combined **DS+CS** using the optimal parameters. All quantities are given in a.u.

intermolecular interaction (see above), the comparison of IQA components for **DS**, **CS**, and **SS** demonstrates excellent linear dependencies between  $V_{xc}^{AB}$  and  $E_{def}$  (see Table II and Fig. S7).



**FIG. 3.** The obtained linear trend between minus intermolecular exchange–correlation energy  $-V_{ee\ XC}^{\Omega\Omega\prime}$  and the *integral* for the **VS** using the parameters from optimization against the combined **DS+CS**. All quantities are given in a.u.

TADI E II. Data for the linear dependence

Test set	k	b	$R^2$		
DS	-1.018	-0.0004	0.9970		
CS	-1.156	-0.0134	0.9924		
SS	-0.979	-0.0005	0.9962		

v/AB

**TABLE III.** Data for linear dependencies between the parametric model (optimized against  $V_{xc}^{OO'}$ ) and  $E_{def}$ .

		$E_{def} = k \cdot Integral + b$				
Test set	$P_1$	$p_2$	k	b	$R^2$	
DS	1.137	1.053	0.396	-0.0009	0.9950	
CS	1.195	1.013	0.498	-0.0013	0.9875	
SS	1.078	1.028	0.295	-0.0012	0.9941	
DS+CS	1.171	1.044	0.459	-0.0007	0.9982	

From this, model (7) with the parameters optimized against  $V_{xc}^{AB}$  values should already be a satisfactory approximation to  $E_{def}$ . Indeed, the corresponding determination coefficients are rather high (see Table III and Fig. S8), while the MAE values are 0.49, 2.78, 0.95, and 0.48 kcal/mol for DS, CS, DS+CS, and SS, respectively.

The optimization against  $E_{def}$  expectedly results in the increasing of  $R^2$  (Table IV and Fig. S9) and lowering errors: 0.48, 1.1, 0.94, and 0.42 kcal/mol for **DS**, **CS**, **DS**+**CS**, and **SS**, respectively. It should be noted that the resulting parameters  $R^2$  and MAEs deviate only slightly from those obtained by the optimization against  $V_{xc}^{AB}$  for **DS**, **SS**, and **DS**+**CS**.

This contrasts with the **CS** where the optimal values of  $p_1$  and  $p_2$  differ significantly for the  $V_{xc}^{AB}$  and  $E_{def}$  estimations. The behavior of the optimization trajectory for  $E_{def}$  can be explained by the shortcoming of the deformation energy definition used for the caged systems (see Sec. III for details). The doubling of the change of self-energy of the central molecule is only a crude approximation to describe the deformation of the entire molecular cluster as it obviously produces higher  $E_{def}$  values due to the overestimation of energy changes of boundary molecules. Despite this problem for caged systems, the combined **DS+CS** is still of a sufficient quality that allows us to use the obtained parameters to estimate deformation contributions of  $E_{bind}$  at the DFT accuracy. Again, model

**TABLE IV.** Data for linear trends obtained by re-optimization of parameters  $p_1$  and  $p_2$  for model (7) against  $E_{def}$ .

Test set	$p_1$	$p_2$	k	Ь	$R^2$
DS	1.136	1.069	0.402	-0.0009	0.9951
CS	1.289	0.620	0.452	-0.0009	0.9912
SS DS+CS	1.115	1.039	0.354 0.507	-0.0007 -0.0006	0.9949



**FIG. 4.** The obtained linear trend between deformation energy  $E_{def}$  and the *integral* for the **VS** using the parameters from optimization for combined **DS**+**CS**. All quantities are given in a.u.

(7) preserves its performance (now for estimations of deformation energy) for the VS systems if the DS+CS parameters are used: there is good agreement between the observed and previously obtained k and b coefficients, while the  $R^2$  value is still high and the MAE value is less than 1 kcal/mol (see Fig. 4).

#### C. Electrostatic term

The last component of intermolecular binding energy  $E_{bind}$  to be approximated is the Coulomb contribution  $V_C^{AB}$ . It should be noted that the direct optimization of model (7) against this term provides unsatisfactory results (see Table V). In other words, the electronic population of topologically defined intermolecular surface cannot be used to correctly describe the Coulomb contribution into the *in situ* intermolecular energy. This is in total agreement with the known dependence of the topological bonding on not Coulombic but exchange–correlation energies.<sup>10,48</sup>

In its turn, the multipole approximation of  $V_C^{\Omega\Omega'}$  terms by expansion (8) truncated up to  $1/R_{\Omega\Omega'}^{-7}$  terms provides a reasonable alternative to the exact calculations based on Eq. (3). The trends obtained for the **DS**, **CS**, **DS+CS**, and **SS** (Fig. 5 and S10) are of an excellent quality with the corresponding MAEs being less than 0.3, 2.1, 0.6, and 0.3 kcal/mol.

**TABLE V.** Data for linear trends obtained by re-optimization of parameters  $p_1$  and  $p_2$  for model (7) against  $V_C^{\Omega\Omega'}$ .

	-				
Test set	$p_1$	$P_2$	k	b	$R^2$
DS	0.189	4.121	-0.029	0.0016	0.9467
CS	1.617	0.639	-0.688	0.0007	0.9627
SS	1.585	0.495	-0.523	0.0197	0.840
DS+CS	1.818	0.641	-1.243	0.0004	0.9100

### The Journal of Chemical Physics



**FIG. 5.** The linear trend for the combined **DS**+**CS** (top) and the **SS** (bottom) between the approximated  $V_C^{\Omega\Omega'}$  using the multipole expansion and actual  $V_C^{\Omega\Omega'}$  retrieved from the ordinary IQA procedure. All quantities are given in a.u.

## D. Total binding energy of intermolecular interaction

Given the demonstrated possibility of restoring all the main components of  $E_{bind}$ , it is of interest to verify the applicability of the approximations used in obtaining this integral quantity by simple summation of its evaluated parts. Quite expectedly, though, the resulting accuracy of this approach is rather low: the sum of negative  $E_{in \, situ}$  (= $V_C^{AB} + V_{xc}^{AB}$ ) and positive  $E_{def}$  terms, both large in absolute values, leads to the summation of errors arising in the recovery of each component (see Figs. S11–S14). This results in a significant noise level of linear trends, rather than low  $R^2$  values (not more than 0.976). Moreover, the trends can even mistake in a sign of binding energy value if the most weak interactions are under consideration.

Nevertheless, a valuable information on the role of  $E_{bind}$  components can be obtained from these trends when combining the above described  $\rho(r)$ -based approximations (having the optimized

 $p_1$  and  $p_2$  parameters) with the values obtained from the ordinary IQA procedure (Table VI, Figs. S11–S14).

For instance, the use of (4) to calculate  $V_{xc}^{AB}$  terms (even when the exact 2-RDM is known) shows no pronounced positive effect on the accuracy of  $E_{bind}$  estimations. A similar situation is observed when approximated  $E_{def}$  terms are changed over the variational ones. However, one must recall that the approximation of deformation energy based on the  $\rho(\mathbf{r})$  topology may provide unsatisfactory results for strong interactions with large covalent contributions. Finally, the approximation of  $V_C^{AB}$  can be considered as the most crucial one: the use of 6D integrations of atomic charge densities [see Eq. (3) instead of the approximation (8) lowers MAE values (less than 1 kcal/mol for DS+CS) and makes the trends pronouncedly less noisy (compare Fig. 6 and Fig. S13). Although we have to mention that this change in accuracy may be originated from the accuracy in treatment of atomic basins by the software used to expand  $V_C^{\Omega\Omega'}$  terms (see Sec. III), it agrees well with the role of electrostatic forces as the driver of intermolecular interactions. Anyway, higher-order terms of the  $V_C^{\Omega\Omega'}$  expansion can be readily calculated to increase the accuracy in restoring of the Coulombic component of  $E_{bind}$ .

Finally, the surface integral of a slightly different form (yet also counting electronic population of intermolecular surface) was suggested in our research group to estimate the binding energy and successfully tested for a number of systems from the DS.<sup>84</sup> This allows us to assume that the parametric model (7) can produce binding energy when reoptimized. Indeed, by carrying out the above procedure for systems from Ref. 84, we were able to achieve similar accuracy ( $R^2 = 0.9841$ , MAE 0.65 kcal/mol vs  $R^2 = 0.9804$  and 0.3 kcal/mol in Ref. 84). However, the associates with stronger interactions (such as Cl<sub>3</sub>COH...H<sub>2</sub>O, F<sub>3</sub>COH...H<sub>2</sub>O, H<sub>3</sub>COH...HF, H<sub>3</sub>COH...HCl, H<sub>3</sub>COH...HBr) added to generate the **DS** in the current study significantly violate the linearity of trend, even if included in the re-optimization procedure (see Fig. S15). This also suggests the inability of the proposed model (7) to properly describe  $E_{bind}$ , thus indirectly confirming that the population of intermolecular surface as a strength of bonding between topological basins is purely determined by exchange-correlation intermolecular effects.

#### E. Efficiency of the proposed protocol

In terms of efficiency, the proposed protocol has several fundamental advantages over the standard IQA procedures for estimating both the intermolecular exchange–correlation energy and the deformation energy terms, while retaining a satisfactory level of accuracy. First, the presented approach is only based on the analysis of the electron density function and does not require 2RDM or its approximations, which substantially reduces the computational cost of the method. Furthermore, the relative efficiency of the proposed protocol in terms of the necessary amount of calculations compared to standard procedures is quite obvious: while the former approach requires surface integrals, the exact  $V_{xc}^{AB}$  terms require six-dimensional volume integrals, and at least two additional variational calculations are required for the exact determination of the deformation energy, in general.

TABLE VI. Data for linear trends between the binding energies and the sum of its components obtained either by the abo	ove
approximation (marked as approx.) or by the conventional IQA procedure.	

Binding energy components used	k	b	$R^2$	MAE
DS				
$E_{bind} = V_{xc}^{AB}(approx.) + E_{def}(approx.) + V_{C}^{AB}(approx.)$	0.910	-0.0004	0.9759	0.406
$E_{bind} = V_{xc}^{AB}(approx) + E_{def}(approx) + V_C^{AB}$	0.977	-0.0001	0.9800	0.335
$E_{bind} = V_{xc}^{AB}(approx) + E_{def} + V_C^{AB}(approx)$	0.904	-0.0004	0.9716	0.474
$E_{bind} = V_{xc}^{AB} + E_{def}(approx \cdot) + V_C^{AB}(approx.)$	0.899	-0.0004	0.9585	0.513
CS				
$E_{bind} = V_{xc}^{AB}(approx.) + E_{def}(approx.) + V_{C}^{AB}(approx.)$	0.838	-0.0040	0.9008	2.925
$E_{bind} = V_{xc}^{AB}(approx) + E_{def}(approx) + V_{C}^{AB}$	0.937	-0.0020	0.9597	1.832
$E_{bind} = V_{xc}^{AB}(approx \cdot) + E_{def} + V_C^{AB}(approx.)$	0.873	-0.0032	0.8851	3.571
$E_{bind} = V_{xc}^{AB} + E_{def}(approx \cdot) + V_C^{AB}(approx.)$	0.867	-0.0033	0.9284	2.551
DS+CS				
$E_{bind} = V_{xc}^{AB}(approx.) + E_{def}(approx.) + V_{C}^{AB}(approx.)$	0.917	-0.0008	0.9391	1.119
$E_{bind} = V_{xc}^{AB}(approx \cdot) + E_{def}(approx \cdot) + V_C^{AB}$	0.981	-0.0004	0.9651	0.922
$E_{bind} = V_{xc}^{AB}(approx) + E_{def} + V_C^{AB}(approx)$	0.920	-0.0007	0.9389	1.243
$E_{bind} = V_{xc}^{AB} + E_{def}(approx \cdot) + V_C^{AB}(approx.)$	0.934	-0.0007	0.9481	0.953
SS				
$E_{bind} = V_{xc}^{AB}(approx.) + E_{def}(approx.) + V_{C}^{AB}(approx.)$	1.027	0.000 03	0.9526	0.524
$E_{bind} = V_{xc}^{AB}(approx) + E_{def}(approx) + V_{C}^{AB}$	0.982	-0.0001	0.9752	0.375
$E_{bind} = V_{xc}^{AB}(approx \cdot) + E_{def} + V_C^{AB}(approx.)$	1.038	0.0001	0.9620	0.495
$E_{bind} = V_{xc}^{AB} + E_{def}(approx \cdot) + V_C^{AB}(approx.)$	1.006	0.0001	0.9412	0.558

<sup>a</sup>Given in kcal/mol.



**FIG. 6.** Linear trend between the binding energy and the sum of the  $V_{xc}^{AB}$  and  $E_{def}$  energies obtained via the parametric model (8) and the IQA-defined  $V_{c}^{AB}$  for the combined **DS+CS**.

#### V. CONCLUSIONS

Summarizing, the  $\rho(\mathbf{r})$ -based protocol to approximately calculate IQA contributions of intermolecular binding energy is suggested and tested on the model sets comprised by associates having various types of non-covalent interactions. The protocol's verification is made based on the IQA components calculated in the DFT (PBE0) approximation as well as by the treatment of the exact 2-RDM computed at the CCSD level. The essential feature of the protocol is the ability to estimate exchange–correlation and deformation contributions from the knowledge of electron density topology and electronic population of the QTAIM-defined intermolecular surface. It can be noticed that the proposed protocol works within the paradigm of quantum chemical topology<sup>7</sup> and makes it possible to fetch out the so-called bonding contributions into the  $V_{xc}^{AB}$  and  $E_{def}$ terms.

The corresponding estimations of the intermolecular  $V_{xe}^{AE}$  term use the extension of the previously published model,<sup>27</sup> which was suggested for diatomic species and is confirmed to be accurate for intermolecular interactions in this study. The  $E_{def}$  estimations are made in a similar fashion owing to the proximity of absolute values of these two terms found in this paper for a broad range of non-covalent interactions having relatively low covalent (exchange–correlation) contributions. This proximity is arguably 15 December 2023 08:26:33

ARTICLE

linked with the interdependence of delocalization and localization of electrons at constant populations of topological basins, which is known from the theory of delocalization indices and has been recently demonstrated<sup>48</sup> for the surface integrals used to evaluate  $V_{xc}^{AB}$  in this work. Although this allows us to avoid the direct evaluation of  $E_{def}$  term, the total intercompensation of  $V_{xc}^{AB}$  and  $E_{def}$  cannot be expected in a general case: a non-zero charge transfer between molecules may occur upon the formation of even weak intermolecular interactions.

Still, in full concordance with the conventional point of view, the relation between  $V_{xc}^{AB}$  and  $E_{def}$  makes the electrostatic component to be the main driver of the intermolecular interaction formation. The corresponding  $V_{c}^{AB}$  term is of long-range character and, hence, is not easy (if possible) to approximate based on properties of bonding between topological fragments. However, the multipolar expansion used to calculate interatomic Coulomb terms for the studied systems has once again confirmed its robustness.

Here, one should note that the proposed protocol can also be used to estimate the total intermolecular binding energy as the sum of  $V_{xc}^{AB}$ ,  $E_{def}$ , and  $V_C^{AB}$  although with a relatively low accuracy (not less than 1.2 kcal/mol in average for the most representative set of species), which can be improved by using a non-approximated electrostatic term also relying only on the trace of 1-RDM.

Thus, using the combination of fitted surface integral method and multipole expansion, the main IQA components of  $E_{bind}$  can be numerically yet rather accurately reconstructed from the electron density function without resorting to six-dimensional integrals and full 1-RDM or 2-RDM computations. The approach is anticipated to be helpful when unveiling the nature of intermolecular interactions in theoretical studies of large supramolecular ensembles as well as in experimental charge density investigations of single crystals.

#### SUPPLEMENTARY MATERIAL

See the supplementary material for all the observed linear trends as well as the tables containing all data on the optimization of model (7).

#### ACKNOWLEDGMENTS

This work was supported by the Russian Science Foundation (Project No. 22-13-00238). We are grateful to the Computational Center of Novosibirsk State University (www.nusc.ru) for the generous allotment of computational time for CCSD geometry optimization. Finally, the help by Dr. I. V. Fedyanin (INEOS RAS) with the periodic calculations is gratefully acknowledged.

#### AUTHOR DECLARATIONS

#### **Conflict of Interest**

The authors have no conflicts to disclose.

#### **Author Contributions**

Aleksei A. Anisimov: Data curation (equal); Investigation (equal); Software (equal); Writing – original draft (equal); Writing – review & editing (supporting). **Ivan V. Ananyev**: Conceptualization (lead); Data curation (supporting); Funding acquisition (lead); Investigation (equal); Writing – original draft (equal); Writing – review & editing (lead).

#### DATA AVAILABILITY

The data that support the findings of this study are available within the article and its supplementary material. The software that supports the findings of this study is available from the corresponding author upon reasonable request.

#### REFERENCES

<sup>1</sup>K. Müller-Dethlefs and P. Hobza, "Noncovalent interactions: A challenge for experiment and theory," Chem. Rev. **100**(1), 143–168 (2000).

<sup>2</sup>J. Černý and P. Hobza, "Non-covalent interactions in biomacromolecules," Phys. Chem. Chem. Phys. **9**(39), 5291–5303 (2007).

<sup>3</sup>E. V. Bartashevich and V. G. Tsirelson, "Interplay between non-covalent interactions in complexes and crystals with halogen bonds," Russ. Chem. Rev. **83**(12), 1181 (2014).

<sup>4</sup>S. E. Wheeler, T. J. Seguin, Y. Guan, and A. C. Doney, "Noncovalent interactions in organocatalysis and the prospect of computational catalyst design," Acc. Chem. Res. **49**(5), 1061–1069 (2016).

<sup>5</sup>P. L. A. Popelier, "Non-covalent interactions from a quantum chemical topology perspective," J. Mol. Model. **28**(9), 276 (2022).

<sup>6</sup>R. Laplaza, F. Peccati, R. A. Boto, C. Quan, A. Carbone, J.-P. Piquemal, Y. Maday, and J. Contreras-García, "NCIPLOT and the analysis of noncovalent interactions using the reduced density gradient," WIREs Comput. Mol. Sci. 11(2), e1497 (2021).

<sup>7</sup>P. L. A. Popelier, in *The Chemical Bond II: 100 Years Old and Getting Stronger*, edited by D. M. P. Mingos (Springer International Publishing, Cham, 2016), pp. 71–117.

<sup>8</sup>R. McWeeny, in *Methods of Molecular Quantum Mechanics*, 2nd ed. (Academic Press, London, 1992).

<sup>9</sup>R. F. W. Bader, in *Atoms in Molecules: A Quantum Theory*, 1st Paperback ed. (Clarendon Press, Oxford, England, New York, 1994).

<sup>10</sup>A. M. Pendás, E. Francisco, M. A. Blanco, and C. Gatti, "Bond paths as privileged exchange channels," Chem. - Eur. J. **13**(33), 9362–9371 (2007).

<sup>11</sup>K. A. Lyssenko, "Analysis of supramolecular architectures: Beyond molecular packing diagrams," Mendeleev Commun. **22**(1), 1–7 (2012).

<sup>12</sup> M. A. Blanco, A. Martín Pendás, and E. Francisco, "Interacting quantum atoms: A correlated energy decomposition scheme based on the quantum theory of atoms in molecules," J. Chem. Theory Comput. 1(6), 1096–1109 (2005).

<sup>13</sup>J. M. Guevara-Vela, E. Francisco, T. Rocha-Rinza, and Á. M. Pendás, "Interacting quantum atoms—A review," Molecules 25(17), 4028 (2020).

<sup>14</sup>B. Jeziorski, R. Moszynski, and K. Szalewicz, "Perturbation theory approach to intermolecular potential energy surfaces of van der Waals complexes," Chem. Rev. 94(7), 1887–1930 (1994).

<sup>15</sup>P. Su and H. Li, "Energy decomposition analysis of covalent bonds and intermolecular interactions," J. Chem. Phys. 131(1), 014102 (2009).

<sup>16</sup>E. D. Glendening and A. Streitwieser, "Natural energy decomposition analysis: An energy partitioning procedure for molecular interactions with application to weak hydrogen bonding, strong ionic, and moderate donor-acceptor interactions," J. Chem. Phys. **100**(4), 2900–2909 (1994).

<sup>17</sup>K. Kitaura and K. Morokuma, "A new energy decomposition scheme for molecular interactions within the Hartree-Fock approximation," Int. J. Quantum Chem. 10(2), 325–340 (1976).

<sup>18</sup>O. A. Syzgantseva, V. Tognetti, and L. Joubert, "On the physical nature of halogen bonds: A QTAIM study," J. Phys. Chem. A 117(36), 8969–8980 (2013).

 $^{19}$  F. Sagan, R. Filas, and M. P. Mitoraj, "Non-covalent interactions in hydrogen storage materials LiN(CH<sub>3</sub>)<sub>2</sub>BH<sub>3</sub> and KN(CH<sub>3</sub>)<sub>2</sub>BH<sub>3</sub>," Crystals **6**(3), 28 (2016).

ARTICLE

<sup>20</sup> A. A. Larin, D. D. Degtyarev, I. V. Ananyev, A. N. Pivkina, and L. L. Fershtat, "Linear furoxan assemblies incorporating nitrobifuroxan scaffold: En route to new high-performance energetic materials," Chem. Eng. J. **470**, 144144 (2023).

<sup>21</sup>A. A. Larin, I. V. Ananyev, E. V. Dubasova, F. E. Teslenko, K. A. Monogarov, D. V. Khakimov, C. He, S. Pang, G. A. Gazieva, and L. L. Fershtat, "Simple and energetic: Novel combination of furoxan and 1,2,4-triazole rings in the synthesis of energetic materials," Energ. Mater. Front. 3(3), 146–153 (2022).

<sup>22</sup>A. A. Anisimov and I. V. Ananyev, "Revisiting the energy treatment of the density of molecular crystals: An interrelation between intermolecular interaction energies and changes of molecular volume," Russ. Chem. Bull. **70**(8), 1429–1437 (2021).

<sup>23</sup> A. M. K. Müller, "Explicit approximate relation between reduced two- and one-particle density matrices," Phys. Lett. A 105(9), 446–452 (1984).
 <sup>24</sup> O. Gritsenko, K. Pernal, and E. J. Baerends, "An improved density matrix func-

<sup>24</sup>O. Gritsenko, K. Pernal, and E. J. Baerends, "An improved density matrix functional by physically motivated repulsive corrections," J. Chem. Phys. **122**(20), 204102 (2005).

<sup>25</sup>M. García-Revilla, E. Francisco, A. Costales, and A. Martín Pendás, "Performance of the density matrix functional theory in the quantum theory of atoms in molecules," J. Phys. Chem. A **116**(4), 1237–1250 (2012).

<sup>26</sup>I. Cukrowski and P. M. Polestshuk, "Reliability of interacting quantum atoms (IQA) data computed from post-HF densities: Impact of the approximation used," Phys. Chem. Chem. Phys. **19**(25), 16375–16386 (2017).

<sup>27</sup>A. A. Anisimov and I. V. Ananyev, "Interatomic exchange-correlation interaction energy from a measure of quantum theory of atoms in molecules topological bonding: A diatomic case," J. Comput. Chem. **41**(25), 2213–2222 (2020).

<sup>28</sup>P. L. A. Popelier and D. S. Kosov, "Atom-atom partitioning of intramolecular and intermolecular Coulomb energy," J. Chem. Phys. **114**(15), 6539–6547 (2001).

<sup>29</sup>J. Řezáč, K. E. Riley, and P. Hobza, "S66: A well-balanced database of benchmark interaction energies relevant to biomolecular structures," J. Chem. Theory Comput. 7(8), 2427–2438 (2011).

<sup>30</sup> J. Řezáč, K. E. Riley, and P. Hobza, "Benchmark calculations of noncovalent interactions of halogenated molecules," J. Chem. Theory Comput. 8(11), 4285–4292 (2012).

 <sup>31</sup>E. Francisco, J. L. Casals-Sainz, T. Rocha-Rinza, and A. Martín Pendás, "Partitioning the DFT exchange-correlation energy in line with the interacting quantum atoms approach," Theor. Chem. Acc. 135(7), 170 (2016).
 <sup>32</sup>D. Menéndez-Crespo, A. Costales, E. Francisco, and Á. Martín Pendás, "Real-

<sup>32</sup>D. Menéndez-Crespo, A. Costales, E. Francisco, and Á. Martín Pendás, "Real-space *in situ* bond energies: Toward a consistent energetic definition of bond strength," Chem. - Eur. J. **24**(36), 9101–9112 (2018).

<sup>33</sup>A. Martín Pendás and E. Francisco, "Real space bond orders are energetic descriptors," Phys. Chem. Chem. Phys. **20**(23), 16231–16237 (2018).

<sup>34</sup> R. F. W. Bader, A. Streitwieser, A. Neuhaus, K. E. Laidig, and P. Speers, "Electron delocalization and the Fermi hole," J. Am. Chem. Soc. 118(21), 4959–4965 (1996).

<sup>35</sup>M. García-Revilla, P. L. A. Popelier, E. Francisco, and ÁM. Pendás, "Nature of chemical interactions from the profiles of electron delocalization indices," J. Chem. Theory Comput. 7(6), 1704–1711 (2011).

<sup>36</sup>M. García-Revilla, E. Francisco, P. L. A. Popelier, and A. Martín Pendás, "Domain-averaged exchange-correlation energies as a physical underpinning for chemical graphs," ChemPhysChem 14(6), 1211–1218 (2013).

<sup>37</sup>P. Maxwell, Á.M. Pendás, and P. L. A. Popelier, "Extension of the interacting quantum atoms (IQA) approach to B3LYP level density functional theory (DFT)," Phys. Chem. Chem. Phys. **18**(31), 20986–21000 (2016).

<sup>38</sup>P. Salvador and I. Mayer, "One- and two-center physical space partitioning of the energy in the density functional theory," J. Chem. Phys. **126**(23), 234113 (2007).

<sup>39</sup>M. Gimferrer, S. Danés, D. M. Andrada, and P. Salvador, "Merging the energy decomposition analysis with the interacting quantum atoms approach," J. Chem. Theory Comput. **19**(12), 3469–3485 (2023).

<sup>40</sup>T. L. Gilbert, "Hohenberg-Kohn theorem for nonlocal external potentials," Phys. Rev. B 12(6), 2111–2120 (1975).

<sup>41</sup>S. M. Valone, "Consequences of extending 1-matrix energy functionals from pure-state representable to all ensemble representable 1 matrices," J. Chem. Phys. 73(3), 001344–1349 (2008).

<sup>42</sup>J. O. Hirschfelder, "Classical and quantum mechanical hypervirial theorems," J. Chem. Phys. **33**(5), 1462–1466 (2004). <sup>43</sup>S. T. Epstein, "Coordinate invariance, the differential force law, and the divergence of the stress-energy tensor," J. Chem. Phys. **63**(8), 003573–3574 (2008).

<sup>44</sup>S. Srebrenik and R. F. W. Bader, "Sufficient conditions for fragment and regional virial theorems," J. Chem. Phys. **61**(7), 2536–2539 (1974).

<sup>45</sup>Z. Badri, C. Foroutan-Nejad, J. Kozelka, and R. Marek, "On the non-classical contribution in lone-pair– $\pi$  interaction: IQA perspective," Phys. Chem. Chem. Phys. **17**(39), 26183–26190 (2015).

<sup>46</sup>J. L. Casals-Sainz, A. C. Castro, E. Francisco, and Á.M. Pendás, "Tetrel interactions from an interacting quantum atoms perspective," Molecules 24(12), 2204 (2019).

<sup>47</sup>D. Suárez, N. Díaz, E. Francisco, and A. Martín Pendás, "Application of the interacting quantum atoms approach to the S66 and ionic-hydrogen-bond datasets for noncovalent interactions," ChemPhysChem 19(8), 973–987 (2018).

<sup>48</sup>A. A. Anisimov and I. V. Ananyev, "On the relationship between the strength of bonding between topological atoms and the exchange-correlation energy," Int. J. Quantum Chem. **123**(9), e27082 (2023).

<sup>49</sup>P. L. A. Popelier, L. Joubert, and D. S. Kosov, "Convergence of the electrostatic interaction based on topological atoms," J. Phys. Chem. A **105**(35), 8254–8261 (2001).

<sup>50</sup>D. S. Kosov and P. L. A. Popelier, "Atomic partitioning of molecular electrostatic potentials," J. Phys. Chem. A **104**(31), 7339–7345 (2000).
 <sup>51</sup>P. L. A. Popelier and M. Rafat, "The electrostatic potential generated by topo-

<sup>51</sup>P. L. A. Popelier and M. Rafat, "The electrostatic potential generated by topological atoms: A continuous multipole method leading to larger convergence regions," Chem. Phys. Lett. **376**(1-2), 148–153 (2003).

<sup>52</sup> J. C. R. Thacker, A. L. Wilson, Z. E. Hughes, M. J. Burn, P. I. Maxwell, and P. L. A. Popelier, "Towards the simulation of biomolecules: Optimisation of peptide-capped glycine using FFLUX," Mol. Simul. 44(11), 881–890 (2018).

<sup>53</sup>A. Volkov and P. Coppens, "Calculation of electrostatic interaction energies in molecular dimers from atomic multipole moments obtained by different methods of electron density partitioning," J. Comput. Chem. 25(7), 921–934 (2004).

<sup>54</sup>A. D. Buckingham, Advances in Chemical Physics (John Wiley & Sons, Ltd., 1967), pp. 107–142.

<sup>55</sup>P. Hobza, "Calculations on noncovalent interactions and databases of benchmark interaction energies," Acc. Chem. Res. **45**(4), 663–672 (2012).

<sup>56</sup>J. P. Perdew, M. Ernzerhof, and K. Burke, "Rationale for mixing exact exchange with density functional approximations," J. Chem. Phys. **105**(22), 9982–9985 (1996).

<sup>57</sup>M. G. Medvedev, I. S. Bushmarinov, J. Sun, J. P. Perdew, and K. A. Lyssenko, "Density functional theory is straying from the path toward the exact functional," *Science* **355**(6320), 49–52 (2017).

<sup>58</sup> R. J. Bartlett, C. E. Dykstra, and J. Paldus, in *Advanced Theories and Computational Approaches to the Electronic Structure of Molecules*, edited by C. E. Dykstra (Springer, Netherlands, Dordrecht, 1984), pp. 127–159.

<sup>59</sup>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. 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. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, T. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. 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, and J. B. Foresman, *Gaussian*, Gaussian, Inc., Wallingford CT, 2016.

<sup>60</sup>F. Weigend and R. Ahlrichs, "Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy," Phys. Chem. Chem. Phys. 7(18), 3297–3305 (2005).

<sup>61</sup>S. Grimme, A. Hansen, J. G. Brandenburg, and C. Bannwarth, "Dispersioncorrected mean-field electronic structure methods," Chem. Rev. **116**(9), 5105–5154 (2016).

<sup>62</sup>S. Grimme, S. Ehrlich, and L. Goerigk, "Effect of the damping function in dispersion corrected density functional theory," J. Comput. Chem. **32**(7), 1456–1465 (2011).

J. Chem. Phys. **159**, 124113 (2023); doi: 10.1063/5.0167874 Published under an exclusive license by AIP Publishing 63 T. Keith, AIMAll, TK Gristmill Software, Overland Park, KS, 2019.

<sup>64</sup> M. Kállay, P. R. Nagy, D. Mester, Z. Rolik, G. Samu, J. Csontos, J. Csóka, P. B. Szabó, L. Gyevi-Nagy, B. Hégely, I. Ladjánszki, L. Szegedy, B. Ladóczki, K. Petrov, M. Farkas, P. D. Mezei, and Á. Ganyecz, "The MRCC program system: Accurate quantum chemistry from water to proteins," J. Chem. Phys. **152**(7),074107 (2020). <sup>65</sup> M. Kállay, P. R. Nagy, D. Mester, Z. Rolik, G. Samu, J. Csontos, J. Csóka, P. B. Szabó, L. Gyevi-Nagy, B. Hégely, I. Ladjánszki, L. Szegedy, B. Ladóczki, K. Petrov, M. Farkas, P. D. Mezei, and Á. Ganyecz, MRCC, 2022.

<sup>66</sup>T. H. Dunning, "Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen," J. Chem. Phys. **90**(2), 1007–1023 (1989).

<sup>67</sup>P. M. Polestshuk, "Accurate integration over atomic regions bounded by zeroflux surfaces," J. Comput. Chem. **34**(3), 206–219 (2013).

<sup>68</sup> P. M. Polestshuk, "Ad hoc methods for accurate determination of Bader's atomic boundary," J. Chem. Phys. **139**(5), 054108 (2013).

<sup>69</sup>P. Polestshuk, "TWOE code: An efficient tool for explicit partition of coupled cluster and configuration interaction energies into atomic and diatomic contributions," J. Phys. Chem. A 125(15), 3198–3207 (2021).
<sup>70</sup>V. A. Karnoukhova, I. V. Fedyanin, E. V. Dubasova, A. A. Anisimov, and I.

<sup>10</sup>V. A. Karnoukhova, I. V. Fedyanin, E. V. Dubasova, A. A. Anisimov, and I. V. Ananyev, "Concerning virial-based estimations of strength of bonding intermolecular interactions in molecular crystals and supramolecular complexes," Mendeleev Commun. **33**(3), 353–356 (2023).

<sup>71</sup> R. Dovesi, A. Erba, R. Orlando, C. M. Zicovich-Wilson, B. Civalleri, L. Maschio, M. Rérat, S. Casassa, J. Baima, S. Salustro, and B. Kirtman, "Quantum-mechanical condensed matter simulations with CRYSTAL," WIREs Comput. Mol. Sci. 8(4), e1360 (2018).

<sup>72</sup> R. Dovesi, V. R. Saunders, C. Roetti, R. Orlando, C. M. Zicovich-Wilson, F. Pascale, B. Civalleri, K. Doll, N. M. Harrison, I. J. Bush, P. D'Arco, M. Llunell, M. Causà, Y. Noël, L. Maschio, A. Erba, M. Rerat, and S. Casassa, CRYSTAL17 User's Manual, University of Torino, 2017. <sup>73</sup>J. Řezáč, "Non-covalent interactions atlas benchmark data sets 5: London dispersion in an extended chemical space," Phys. Chem. Chem. Phys. 24(24), 14780-14793 (2022).

 $^{74}$ K. Kříž and J. Řezáč, "Non-covalent interactions atlas benchmark data sets 4:  $\sigma$ -Hole interactions," Phys. Chem. Chem. Phys. **24**(24), 14794–14804 (2022).

<sup>75</sup>K. Kříž, M. Nováček, and J. Řezáč, "Non-covalent interactions atlas benchmark data sets 3: Repulsive contacts," J. Chem. Theory Comput. **17**(3), 1548–1561 (2021).

<sup>76</sup>J. Řezáč, "Non-covalent interactions atlas benchmark data sets: Hydrogen bonding," J. Chem. Theory Comput. **16**(4), 2355–2368 (2020).

<sup>77</sup>J. Řezáč, "Non-covalent interactions atlas benchmark data sets 2: Hydrogen bonding in an extended chemical space," J. Chem. Theory Comput. 16(10), 6305–6316 (2020).

<sup>78</sup>T. Lu and F. Chen, "Multiwfn: A multifunctional wavefunction analyzer," J. Comput. Chem. **33**(5), 580–592 (2012).

<sup>79</sup> P. L. A. Popelier, "An analytical expression for interatomic surfaces in the theory of atoms in molecules," Theor. Chim. Acta 87(6), 465–476 (1994).

<sup>80</sup>C. G. Broyden, "The convergence of a class of double-rank minimization algorithms: 2. The new algorithm," IMA J. Appl. Math. 6(3), 222–231 (1970).

<sup>81</sup> R. Fletcher, "A new approach to variable metric algorithms," Comput. J. **13**(3), 317–322 (1970).

<sup>82</sup>D. Goldfarb, "A family of variable-metric methods derived by variational means," Math. Comput. 24(109), 23–26 (1970).

<sup>83</sup>D. F. Shanno, "Conditioning of quasi-Newton methods for function minimization," Math. Comput. 24(111), 647–656 (1970).

<sup>84</sup>A. Romanova, K. Lyssenko, and I. Ananyev, "Estimations of energy of noncovalent bonding from integrals over interatomic zero-flux surfaces: Correlation trends and beyond," J. Comput. Chem. **39**(21), 1607–1616 (2018).