Electronic Supplementary Information for

Introducing DDEC6 atomic population analysis: part 3. Comprehensive method to compute bond orders

Thomas A. Manz
Department of Chemical & Materials Engineering, New Mexico State University, Las Cruces, NM 88003-8001

email: tmanz@nmsu.edu

CONTENTS:

S1. Derivation of the contact-exchange formula

In this article, a capital letter (e.g., A, B, etc.) will be used to represent an atom in the reference unit cell while atoms anywhere in the material will be represented by small letter indices (e.g., j). In a periodic material, atom j is labeled by $(B, \ell_1, \ell_2, \ell_3)$, where B is an atom in the reference unit cell and ℓ_1, ℓ_2, and ℓ_3 are the translation whole numbers along the lattice vectors $\vec{v}_1$, $\vec{v}_2$, and $\vec{v}_3$, respectively, to give the nuclear position

$$\vec{R}_j = \vec{R}_B + \ell_1 \vec{v}_1 + \ell_2 \vec{v}_2 + \ell_3 \vec{v}_3 \qquad (S1)$$

where $\vec{R}_B$ is the nuclear position of atom B. For a non-periodic material, the vectors $\vec{v}_1$, $\vec{v}_2$, and $\vec{v}_3$ can be chosen to define any parallelpiped completely enclosing the electron distribution. In non-periodic materials, $\ell_1 = \ell_2 = \ell_3 = 0$ for all atoms. In materials having one periodic dimension, $\ell_2 = \ell_3 = 0$ for all atoms. In materials having two periodic dimensions, $\ell_3 = 0$ for all atoms.

$$\vec{r}_j = \vec{r} - \vec{R}_j \tag{S2}$$

is the vector from the nuclear position of atom j (i.e., $\vec{R}_j$) to position $\vec{r}$. $r_j = \left\|\vec{r}_j\right\|$ is the distance from atom j's nuclear position to position $\vec{r}$. The condition $j \neq A$ means that $\vec{R}_j \neq \vec{R}_A$. Summing over A (or B) means summation over all atoms in the unit cell. Summing over j (or i) means summation over all atoms in the material. Summing over $j \neq A$ means summation over all atoms in the material having $\vec{R}_j \neq \vec{R}_A$.

If $\rho^{DX}_{A,j}(\vec{r}, \vec{r}')$ is the probability density that an electron at position $\vec{r}$ and assigned to atom A dressed-exchanges with an electron at position $\vec{r}'$ and assigned to atom $j \neq A$, then the bond order between atoms A and $j \neq A$ is given by

$$B_{A,j} = \oint \oint \left(\rho^{DX}_{A,j}(\vec{r}, \vec{r}') + \rho^{DX}_{A,j}(\vec{r}', \vec{r}) \right) d^3\vec{r} d^3\vec{r}' = 2 \oint \oint \rho^{DX}_{A,j}(\vec{r}, \vec{r}') d^3\vec{r} d^3\vec{r}' \tag{S3}$$

$$B_{A,A} = \oint \oint \rho^{DX}_{A,A}(\vec{r}, \vec{r}') d^3\vec{r} d^3\vec{r}' \tag{S4}$$

The right-most side of Eq. (S3) arises because $\vec{r}$ and $\vec{r}'$ are interchangeable dummy integration variables. The probability density that an electron at position $\vec{r}$ dressed-exchanges with an electron at position $\vec{r}'$ is

$$\rho^{DX}(\vec{r}, \vec{r}') = \sum_i \sum_j \rho^{DX}_{i,j}(\vec{r}, \vec{r}') \tag{S5}$$

where summations over i and j mean over all atoms in the material (not just in the reference unit cell). By definition,

$$\rho^{DX}(\vec{r}, \vec{r}') = \rho^{DX}(\vec{r}', \vec{r}) \tag{S6}$$

Integrating over all positions $\vec{r}'$ yields the electron density at position $\vec{r}$

$$\rho(\vec{r}) = \oint \rho^{DX}(\vec{r}, \vec{r}') d^3\vec{r}' \tag{S7}$$

The electron dressed-exchange can be re-written in terms of a dressed-exchange hole

$$\rho^{DX\_hole}(\vec{r}, \vec{r}') = \rho^{DX}(\vec{r}, \vec{r}') / \rho(\vec{r}) \tag{S8}$$

Substituting Eq. (S8) into (S7) yields the result that the dressed-exchange hole excludes exactly one electron:

$$\oint \rho^{DX\_hole}(\vec{r}, \vec{r}') d^3\vec{r}' = 1 \tag{S9}$$

To construct a comprehensive bond order, the first step is to express the atomic exchange propensity (i.e., the tendency of each atom to exchange electrons with other atoms) as functionals of $\left\{\rho(\vec{r}), \vec{m}(\vec{r})\right\}$, where $\vec{m}(\vec{r})$ is the spin magnetization density vector that handles either collinear or non-collinear magnetism. To each atom j is assigned an atomic electron density $\rho_j(\vec{r}_j)$ and an atomic spin magnetization density vector $\vec{m}_j(\vec{r}_j)$ that are combined to form the four-vector

$$\vec{\rho}_j(\vec{r}_j) = \left(\rho_j(\vec{r}_j), \vec{m}_j(\vec{r}_j) \right) \tag{S10}$$

and its spherical average at a distance r_j from atom j's nuclear center

$$\vec{\rho}_{j}^{\mathrm{avg}}\left(r_{j}\right)=\left(\rho_{j}^{\mathrm{avg}}\left(r_{j}\right),\vec{m}_{j}^{\mathrm{avg}}\left(r_{j}\right)\right) \tag{S11}$$

Due to electron exchange over the exchange hole and the orbiting of electrons in circuitous motions around and between atoms, the atomic exchange propensity must simultaneously resemble: (a) a weighted average of $\vec{\rho}_{j}\left(\vec{r}_{j}\right)$ over the exchange hole and (b) a weighted average of $\vec{\rho}_{j}^{\mathrm{avg}}\left(r_{j}\right)$ over the exchange hole. Here, this is called the "confluence of atomic exchange propensities". From the basic identity

$$\oint\left(\left\|\delta\vec{\rho}_{j}\left(\vec{r}_{j}\right)\right\|^{2}-\left\|\delta\vec{\rho}_{j}^{\mathrm{avg}}\left(r_{j}\right)\right\|^{2}\right)d^{3}\vec{r}=\oint\left\|\delta\vec{\rho}_{j}\left(\vec{r}_{j}\right)-\delta\vec{\rho}_{j}^{\mathrm{avg}}\left(r_{j}\right)\right\|^{2}d^{3}\vec{r}\geq 0 \tag{S12}$$

it follows that $\vec{\rho}_{j}\left(\vec{r}_{j}\right)$ is more sensitive than $\vec{\rho}_{j}^{\mathrm{avg}}\left(r_{j}\right)$ to changes in the basis set, exchange-correlation theory, and charge and spin partitioning algorithm. Therefore, more stable results will occur if the atomic exchange propensity is based on $\vec{\rho}_{j}^{\mathrm{avg}}\left(r_{j}\right)$ rather than on $\vec{\rho}_{j}\left(\vec{r}_{j}\right)$.

As a hypothetical exercise, we can ask what would happen if the (modified) exchange hole were completely localized. This corresponds to a hypothetical integration over a (modified) exchange hole of negligible radius that excludes exactly one electron. In such case,

$$\rho^{X,0}\left(\vec{r},\vec{r}'\right)=\sqrt{\rho\left(\vec{r}\right)\rho\left(\vec{r}'\right)}\,\delta^{\mathrm{dirac}}\left(\vec{r}-\vec{r}'\right) \tag{S13}$$

The superscript '0' in Eq. (S13) refers to this being (modified) exchange for a completely localized exchange hole (aka 'contact exchange'). As discussed in the previous paragraph, the atomic exchange propensities should be based on $\left\{\vec{\rho}_{j}^{\mathrm{avg}}\left(r_{j}\right)\right\}$. Thus for $j\neq A$

$$\rho_{A,j}^{X,0}\left(\vec{r},\vec{r}'\right)=\frac{\vec{\rho}_{A}^{\mathrm{avg}}\left(r_{A}\right)\bullet\vec{\rho}_{j}^{\mathrm{avg}}\left(r_{j}'\right)}{\vec{\rho}^{\mathrm{avg}}\left(\vec{r}\right)\bullet\vec{\rho}^{\mathrm{avg}}\left(\vec{r}'\right)}\rho^{X,0}\left(\vec{r},\vec{r}'\right) \tag{S14}$$

$$\rho_{A,j}^{DX}\left(\vec{r},\vec{r}'\right)=\frac{\vec{\rho}_{A}^{\mathrm{avg}}\left(r_{A}\right)\bullet\vec{\rho}_{j}^{\mathrm{avg}}\left(r_{j}'\right)}{\vec{\rho}^{\mathrm{avg}}\left(\vec{r}\right)\bullet\vec{\rho}^{\mathrm{avg}}\left(\vec{r}'\right)}\rho^{DX}\left(\vec{r},\vec{r}'\right) \tag{S15}$$

where

$$\vec{\rho}^{\mathrm{avg}}\left(\vec{r}\right)=\sum_{j}\vec{\rho}_{j}^{\mathrm{avg}}\left(r_{j}\right) \tag{S16}$$

Substituting Eq. (S14) into

$$\rho_{A}\left(\vec{r}_{A}\right)=\sum_{j}\oint\rho_{A,j}^{DX}\left(\vec{r},\vec{r}'\right)d^{3}\vec{r}'=\sum_{j}\oint\rho_{A,j}^{X,0}\left(\vec{r},\vec{r}'\right)d^{3}\vec{r}' \tag{S17}$$

implies

$$\rho_{A,A}^{X,0}\left(\vec{r},\vec{r}'\right)=\sqrt{\rho_{A}\left(\vec{r}_{A}\right)\rho_{A}\left(\vec{r}_{A}'\right)}\,\delta^{\mathrm{dirac}}\left(\vec{r}-\vec{r}'\right)-\sum_{j\neq A}\frac{\vec{\rho}_{A}^{\mathrm{avg}}\left(r_{A}\right)\bullet\vec{\rho}_{j}^{\mathrm{avg}}\left(r_{j}'\right)}{\vec{\rho}^{\mathrm{avg}}\left(\vec{r}\right)\bullet\vec{\rho}^{\mathrm{avg}}\left(\vec{r}'\right)}\rho^{X,0}\left(\vec{r},\vec{r}'\right) \tag{S18}$$

In analogy to Eq. (S3) and (S4), for $j\neq A$:

$$CE_{A,j}=\oint\oint\left(\rho_{A,j}^{X,0}\left(\vec{r},\vec{r}'\right)+\rho_{A,j}^{X,0}\left(\vec{r}',\vec{r}\right)\right)d^{3}\vec{r}d^{3}\vec{r}'=2\oint\oint\rho_{A,j}^{X,0}\left(\vec{r},\vec{r}'\right)d^{3}\vec{r}d^{3}\vec{r}' \tag{S19}$$

$$CE_{A,A}=\oint\oint\rho_{A,A}^{X,0}\left(\vec{r},\vec{r}'\right)d^{3}\vec{r}d^{3}\vec{r}' \tag{S20}$$

For $j \neq A$, substituting Eq. (S14) into (S19) gives

$$\mathrm{CE}_{A,j} = 2\oint \frac{\vec{\rho}_A^{\,avg}\left(r_A\right)\bullet\vec{\rho}_j^{\,avg}\left(r_j\right)}{\vec{\rho}^{\,avg}\left(\vec{r}\right)\bullet\vec{\rho}^{\,avg}\left(\vec{r}\right)}\rho\left(\vec{r}\right)d^3\vec{r} \qquad \text{(S21)}$$

Substituting Eq. (S18) into (S20) gives

$$\mathrm{CE}_{A,A} = N_A - \tfrac{1}{2}\mathrm{SCE}_A \qquad \text{(S22)}$$

where the number of electrons assigned to atom A is

$$N_A = \oint \rho_A\left(\vec{r}_A\right)d^3\vec{r}_A \qquad \text{(S23)}$$

The sum of contact exchanges (SCE) for atom A is

$$\mathrm{SCE}_A = \sum_{j\neq A}\mathrm{CE}_{A,j} = 2\oint \frac{\vec{\rho}_A^{\,avg}\left(r_A\right)\bullet\left(\vec{\rho}^{\,avg}\left(\vec{r}\right)-\vec{\rho}_A^{\,avg}\left(r_A\right)\right)}{\vec{\rho}^{\,avg}\left(\vec{r}\right)\bullet\vec{\rho}^{\,avg}\left(\vec{r}\right)}\rho\left(\vec{r}\right)d^3\vec{r}_A \qquad \text{(S24)}$$

In practice, SCE_A is computed via the integral on the right-hand side of Eq. (S24) within the cutoff radius around atom A.

Without loss of generality, the effects of exchange hole delocalization on the bond order can be parameterized by the ratio $\Phi_{A,j} = \mathrm{BO}_{A,j}/\mathrm{CE}_{A,j}$. The following sections derive relations for computing the bond-order-to-contact-exchange ratio, $\Phi_{A,j}$.

S2. Derivation of the lower bound on the bond-order-to-contact-exchange ratio

Without loss of generality, we can divide a material into two sets of atoms, S_I and S_II, such that $S\_I \cap S\_II = \varnothing$ and $S\_I \cup S\_II$ contains all atoms in the material (not just in the reference unit cell). Then,

$$\rho_{S\_I}^{avg}\left(\vec{r}\right) = \sum_{i\in S\_I}\rho_i^{avg}\left(r_i\right) = \rho^{avg}\left(\vec{r}\right) - \rho_{S\_II}^{avg}\left(\vec{r}\right) \qquad \text{(S25)}$$

First, we consider the case of non-magnetic materials. For these, Eq. (S15) simplifies to

$$\rho_{i,j}^{DX}\left(\vec{r},\vec{r}'\right) = \frac{\rho_i^{avg}\left(r_i\right)}{\rho^{avg}\left(\vec{r}\right)}\frac{\rho_j^{avg}\left(r_j'\right)}{\rho^{avg}\left(\vec{r}'\right)}\rho^{DX}\left(\vec{r},\vec{r}'\right) \quad \text{for } i \neq j \qquad \text{(S26)}$$

Substituting Eq. (S8) into (S26) and summing over all of the atoms $i \in S\_I$ and $j \in S\_II$ gives

$$\rho_{S\_I,S\_II}^{DX}\left(\vec{r},\vec{r}'\right) = \frac{\rho_{S\_I}^{avg}\left(\vec{r}\right)}{\rho^{avg}\left(\vec{r}\right)}\frac{\rho_{S\_II}^{avg}\left(\vec{r}'\right)}{\rho^{avg}\left(\vec{r}'\right)}\rho\left(\vec{r}\right)\rho^{DX\_hole}\left(\vec{r},\vec{r}'\right) \qquad \text{(S27)}$$

Substituting Eq. (S27) into (S3) gives the bond order between set_I and set_II as

$$B_{S\_I,S\_II} = 2\oint\oint \frac{\rho_{S\_I}^{avg}\left(\vec{r}\right)}{\rho^{avg}\left(\vec{r}\right)}\frac{\rho_{S\_II}^{avg}\left(\vec{r}'\right)}{\rho^{avg}\left(\vec{r}'\right)}\rho\left(\vec{r}\right)\rho^{DX-hole}\left(\vec{r},\vec{r}'\right)d^3\vec{r}d^3\vec{r}' \qquad \text{(S28)}$$

Following Eq. (S21), the contact exchange between set_I and set_II is

$$\mathrm{CE}_{S\_I,S\_II} = 2\oint \frac{\rho_{S\_I}^{avg}\left(\vec{r}\right)}{\rho^{avg}\left(\vec{r}\right)}\frac{\rho_{S\_II}^{avg}\left(\vec{r}\right)}{\rho^{avg}\left(\vec{r}\right)}\rho\left(\vec{r}\right)d^3\vec{r} \qquad \text{(S29)}$$

The normalization condition (S9) means that Eq. (S28) can be re-written as

$$\mathrm{B}_{\mathrm{S\_I,S\_II}} = 2\oint \frac{\rho^{\mathrm{avg}}_{\mathrm{S\_I}}(\vec{\mathrm{r}})}{\rho^{\mathrm{avg}}(\vec{\mathrm{r}})} \left\langle \frac{\rho^{\mathrm{avg}}_{\mathrm{S\_II}}(\vec{\mathrm{r}}')}{\rho^{\mathrm{avg}}(\vec{\mathrm{r}}')} \right\rangle_{\mathrm{DX\_hole}(\vec{\mathrm{r}})} \rho(\vec{\mathrm{r}})\, \mathrm{d}^3\vec{\mathrm{r}} \qquad \text{(S30)}$$

where

$$\left\langle \frac{\rho^{\mathrm{avg}}_{\mathrm{S\_II}}(\vec{\mathrm{r}}')}{\rho^{\mathrm{avg}}(\vec{\mathrm{r}}')} \right\rangle_{\mathrm{DX\_hole}(\vec{\mathrm{r}})} = \oint \frac{\rho^{\mathrm{avg}}_{\mathrm{S\_II}}(\vec{\mathrm{r}}')}{\rho^{\mathrm{avg}}(\vec{\mathrm{r}}')} \rho^{\mathrm{DX\_hole}}(\vec{\mathrm{r}},\vec{\mathrm{r}}')\, \mathrm{d}^3\vec{\mathrm{r}}' \qquad \text{(S31)}$$

is the weighted average of $\rho^{\mathrm{avg}}_{\mathrm{S\_II}}(\vec{\mathrm{r}}')/\rho^{\mathrm{avg}}(\vec{\mathrm{r}}')$ over the dressed-exchange hole centered around position $\vec{\mathrm{r}}$. Substituting

$$\tau(\vec{\mathrm{r}}) = \frac{\rho^{\mathrm{avg}}_{\mathrm{S\_I}}(\vec{\mathrm{r}})}{\rho^{\mathrm{avg}}(\vec{\mathrm{r}})} \qquad \text{(S32)}$$

gives

$$\Phi_{\mathrm{S\_I,S\_II}} = \frac{\mathrm{BO}_{\mathrm{S\_I,S\_II}}}{\mathrm{CE}_{\mathrm{S\_I,S\_II}}} = \frac{2\oint \tau(\vec{\mathrm{r}})\left(1-\langle\tau\rangle_{\mathrm{DX\_hole}(\vec{\mathrm{r}})}\right)\rho(\vec{\mathrm{r}})\,\mathrm{d}^3\vec{\mathrm{r}}}{2\oint \tau(\vec{\mathrm{r}})\left(1-\tau(\vec{\mathrm{r}})\right)\rho(\vec{\mathrm{r}})\,\mathrm{d}^3\vec{\mathrm{r}}} \qquad \text{(S33)}$$

Defining

$$\langle\tau\rangle_{\mathrm{DX\_hole}(\vec{\mathrm{r}})} = \oint \tau(\vec{\mathrm{r}}')\rho^{\mathrm{DX\_hole}}(\vec{\mathrm{r}},\vec{\mathrm{r}}')\,\mathrm{d}^3\vec{\mathrm{r}}' \qquad \text{(S34)}$$

$$\Delta(\vec{\mathrm{r}}) = \tau(\vec{\mathrm{r}}) - \langle\tau\rangle_{\mathrm{DX\_hole}(\vec{\mathrm{r}})} \qquad \text{(S35)}$$

gives

$$\Phi_{\mathrm{S\_I,S\_II}} = \frac{\oint\left(\langle\tau\rangle_{\mathrm{DX\_hole}(\vec{\mathrm{r}})} + \Delta(\vec{\mathrm{r}})\right)\left(1-\langle\tau\rangle_{\mathrm{DX\_hole}(\vec{\mathrm{r}})}\right)\rho(\vec{\mathrm{r}})\,\mathrm{d}^3\vec{\mathrm{r}}}{\oint\left(\langle\tau\rangle_{\mathrm{DX\_hole}(\vec{\mathrm{r}})} + \Delta(\vec{\mathrm{r}})\right)\left(1-\langle\tau\rangle_{\mathrm{DX\_hole}(\vec{\mathrm{r}})} - \Delta(\vec{\mathrm{r}})\right)\rho(\vec{\mathrm{r}})\,\mathrm{d}^3\vec{\mathrm{r}}} \qquad \text{(S36)}$$

which expands to

$$\Phi_{\mathrm{S\_I,S\_II}} = \frac{\oint\left(\langle\tau\rangle_{\mathrm{DX\_hole}(\vec{\mathrm{r}})} - \langle\tau\rangle_{\mathrm{DX\_hole}(\vec{\mathrm{r}})}^{\;2} + \Delta(\vec{\mathrm{r}}) - \Delta(\vec{\mathrm{r}})\langle\tau\rangle_{\mathrm{DX\_hole}(\vec{\mathrm{r}})}\right)\rho(\vec{\mathrm{r}})\,\mathrm{d}^3\vec{\mathrm{r}}}{\oint\left(\langle\tau\rangle_{\mathrm{DX\_hole}(\vec{\mathrm{r}})} - \langle\tau\rangle_{\mathrm{DX\_hole}(\vec{\mathrm{r}})}^{\;2} - 2\Delta(\vec{\mathrm{r}})\langle\tau\rangle_{\mathrm{DX\_hole}(\vec{\mathrm{r}})} + \Delta(\vec{\mathrm{r}}) - \left(\Delta(\vec{\mathrm{r}})\right)^2\right)\rho(\vec{\mathrm{r}})\,\mathrm{d}^3\vec{\mathrm{r}}} \qquad \text{(S37)}$$

Since $\rho^{\mathrm{avg}}_{\mathrm{S\_I}}(\vec{\mathrm{r}}) \le \rho^{\mathrm{avg}}(\vec{\mathrm{r}})$, it directly follows from Eqs. (S32), (S34), and (S35) that

$$-1 < \Delta(\vec{\mathrm{r}}) < 1 \qquad \text{(S38)}$$

Because $\langle\tau\rangle_{\mathrm{DX\_hole}(\vec{\mathrm{r}})}$ is a weighted average of $\tau(\vec{\mathrm{r}}')$ over the dressed-exchange hole, $\langle\tau\rangle_{\mathrm{DX\_hole}(\vec{\mathrm{r}})}$ fluctuates more slowly as a function of spatial position than $\tau(\vec{\mathrm{r}})$. As proved in Eq. (S57) of the next section, $\Delta(\vec{\mathrm{r}})$ fluctuates between positive and negative values with a density weighted average value of exactly zero. Therefore, in Eq. (S37) the terms linear in $\Delta(\vec{\mathrm{r}})$ will tend to approximately average out over space. This yields

$$\Phi_{\mathrm{S\_I,S\_II}} \approx \frac{\oint\left(\left\langle\tau\right\rangle_{\mathrm{DX\_hole}(\vec{r})} - \left\langle\tau\right\rangle_{\mathrm{DX\_hole}(\vec{r})}{}^{2}\right)\rho(\vec{r})\,d^3\vec{r}}{\oint\left(\left\langle\tau\right\rangle_{\mathrm{DX\_hole}(\vec{r})} - \left\langle\tau\right\rangle_{\mathrm{DX\_hole}(\vec{r})}{}^{2} - \left(\Delta(\vec{r})\right)^2\right)\rho(\vec{r})\,d^3\vec{r}} \geq 1 \tag{S39}$$

The above analysis holds for any chosen sets S_I and S_II that sum to the whole system. Choosing S_I to be a single atom and S_II to be all of the other atoms in the system, it immediately follows that

$$\frac{\mathrm{SBO}_A}{\mathrm{SCE}_A} \gtrsim 1 \tag{S40}$$

for each atom in the material. We may reconstitute S_I and S_II to form any binary division of the entire system. Eq. (S39) will hold for any chosen set S_I if

$$\frac{B_{A,j}}{\mathrm{CE}_{A,j}} \gtrsim 1 \tag{S41}$$

holds for each bond in the material. Therefore, we impose constraint (S41) to ensure Eq. (S39) holds for every possible set S_I.

Finally, we must consider the effects of magnetism. To meet the condition that assigned bond orders for nearly degenerate components of a spin multiplet (e.g., $S_z=0$ and $S_z=\pm1$ components of a triplet state) are nearly equal: (i) the assigned $\left\{\rho_A\left(\vec{r}_A\right)\right\}$ should be a functional of $\rho\left(\vec{r}\right)$ alone without depending on $\vec{m}\left(\vec{r}\right)$ and (ii) the assigned $\vec{m}_A\left(\vec{r}_A\right)$ should resemble proportional spin partitioning as defined by

$$\vec{m}_A^{\mathrm{prop}}\left(\vec{r}_A\right) \approx \vec{m}\left(\vec{r}\right)\rho_A\left(\vec{r}_A\right)/\rho\left(\vec{r}\right) \tag{S42}$$

For proportional spin partitioning,

$$\frac{\vec{\rho}_A^{\mathrm{prop}}\left(\vec{r}_A\right)\bullet\vec{\rho}_j^{\mathrm{prop}}\left(\vec{r}_j'\right)}{\vec{\rho}^{\mathrm{prop}}\left(\vec{r}\right)\bullet\vec{\rho}^{\mathrm{prop}}\left(\vec{r}'\right)} = \frac{\rho_A\left(\vec{r}_A\right)\rho_j\left(\vec{r}_j'\right)}{\rho\left(\vec{r}\right)\rho\left(\vec{r}'\right)} \tag{S43}$$

As described in Section S1 above, $\vec{\rho}_A\left(\vec{r}_A\right)$ should be optimized to resemble $\vec{\rho}_A^{\mathrm{avg}}\left(r_A\right)$ to fulfill the confluence of atomic exchange propensities. Thus,

$$\frac{\vec{\rho}_A^{\mathrm{avg}}\left(r_A\right)\bullet\vec{\rho}_j^{\mathrm{avg}}\left(r_j'\right)}{\vec{\rho}^{\mathrm{avg}}\left(\vec{r}\right)\bullet\vec{\rho}^{\mathrm{avg}}\left(\vec{r}'\right)} \approx \frac{\vec{\rho}_A\left(\vec{r}_A\right)\bullet\vec{\rho}_j\left(\vec{r}_j'\right)}{\vec{\rho}\left(\vec{r}\right)\bullet\vec{\rho}\left(\vec{r}'\right)} \approx \frac{\vec{\rho}_A^{\mathrm{prop}}\left(\vec{r}_A\right)\bullet\vec{\rho}_j^{\mathrm{prop}}\left(\vec{r}_j'\right)}{\vec{\rho}^{\mathrm{prop}}\left(\vec{r}\right)\bullet\vec{\rho}^{\mathrm{prop}}\left(\vec{r}'\right)} = \frac{\rho_A\left(\vec{r}_A\right)\rho_j\left(\vec{r}_j'\right)}{\rho\left(\vec{r}\right)\rho\left(\vec{r}'\right)} \approx \frac{\rho_A^{\mathrm{avg}}\left(r_A\right)\rho_j^{\mathrm{avg}}\left(r_j'\right)}{\rho^{\mathrm{avg}}\left(\vec{r}\right)\rho^{\mathrm{avg}}\left(\vec{r}'\right)} \tag{S44}$$

The right-most side of Eq. (S44) is the exchange propensity factor for a non-magnetic system, while the left-most side is that for a magnetic system. Because a charge and spin partitioning optimized to satisfy the confluence relations in Eq. (S44) has similar exchange propensity factors for magnetic and non-magnetic systems, it follows that the lower bound on $\Phi_{A,j}$ is equivalent for magnetic and non-magnetic systems.

Given the confluence relations (Eq. (S44)), why are the atomic exchange propensities used herein based on $\left\{\vec{\rho}_A^{\mathrm{avg}}\left(r_A\right)\right\}$ rather than on $\left\{\vec{\rho}_A\left(\vec{r}_A\right)\right\}$, $\left\{\rho_A\left(\vec{r}_A\right)\right\}$, or $\left\{\rho_A^{\mathrm{avg}}\left(r_A\right)\right\}$? First, as proved by Eq. (S12), using $\left\{\vec{\rho}_A^{\mathrm{avg}}\left(r_A\right)\right\}$ instead of $\left\{\vec{\rho}_A\left(\vec{r}_A\right)\right\}$ reduces the sensitivity of computed bond orders to the choice of basis sets, exchange-correlation method, and charge and spin partitioning algorithm. Second, since electron exchange depends on particle spin, the atomic exchange propensity should also be spin dependent. The importance of this can be understood by considering a non-optimal charge and spin partitioning that breaks the confluence relations in Eq. (S44). For such non-optimal partitioning, the

$\left\{\vec{\rho}_{\mathrm{A}}^{\mathrm{avg}}\left(\mathrm{r}_{\mathrm{A}}\right)\right\}$, $\left\{\rho_{\mathrm{A}}^{\mathrm{avg}}\left(\mathrm{r}_{\mathrm{A}}\right)\right\}$, and $\left\{\rho_{\mathrm{A}}\left(\vec{\mathrm{r}}_{\mathrm{A}}\right)\right\}$ sets would not produce similar atomic exchange propensities. Under such circumstances, $\left\{\vec{\rho}_{\mathrm{A}}^{\mathrm{avg}}\left(\mathrm{r}_{\mathrm{A}}\right)\right\}$ can be understood as the more fundamental atomic exchange propensity, because it incorporates particle spin effects essential to describing electron exchange while $\left\{\rho_{\mathrm{A}}^{\mathrm{avg}}\left(\mathrm{r}_{\mathrm{A}}\right)\right\}$ and $\left\{\rho_{\mathrm{A}}\left(\vec{\mathrm{r}}_{\mathrm{A}}\right)\right\}$ do not include any particle spin effects. Even when the confluence relations in Eq. (S44) are satisfied, it is desirable to retain use of the more fundamental atomic exchange propensity $\left\{\vec{\rho}_{\mathrm{A}}^{\mathrm{avg}}\left(\mathrm{r}_{\mathrm{A}}\right)\right\}$ instead of the less fundamental substitutes $\left\{\rho_{\mathrm{A}}^{\mathrm{avg}}\left(\mathrm{r}_{\mathrm{A}}\right)\right\}$ or $\left\{\rho_{\mathrm{A}}\left(\vec{\mathrm{r}}_{\mathrm{A}}\right)\right\}$.

S3. Derivation of the upper bound on the bond-order-to-contact-exchange ratio

As described in the previous section, we can divide a material into two sets of atoms, S_I and S_II, such that $\mathrm{S\_I} \cap \mathrm{S\_II} = \varnothing$ and $\mathrm{S\_I} \cup \mathrm{S\_II}$ contains all atoms in the material (not just in the reference unit cell). We further divide space into two distinct regions, such that Region I (R_I) corresponds to $\rho_{\mathrm{S\_I}}^{\mathrm{avg}}\left(\vec{\mathrm{r}}\right)/\rho^{\mathrm{avg}}\left(\vec{\mathrm{r}}\right) \geq ½$ and Region II (R_II) corresponds to $\rho_{\mathrm{S\_I}}^{\mathrm{avg}}\left(\vec{\mathrm{r}}\right)/\rho^{\mathrm{avg}}\left(\vec{\mathrm{r}}\right) < ½$. For a non-magnetic system,

$$\mathrm{CE}_{\mathrm{S\_I,S\_II}}^{\mathrm{Region}} = 2\int_{\mathrm{Region}} \frac{\rho_{\mathrm{S\_I}}^{\mathrm{avg}}\left(\vec{\mathrm{r}}\right)}{\rho^{\mathrm{avg}}\left(\vec{\mathrm{r}}\right)}\frac{\rho_{\mathrm{S\_II}}^{\mathrm{avg}}\left(\vec{\mathrm{r}}\right)}{\rho^{\mathrm{avg}}\left(\vec{\mathrm{r}}\right)}\rho\left(\vec{\mathrm{r}}\right)d^3\vec{\mathrm{r}} \tag{S45}$$

$$\mathrm{B}_{\mathrm{S\_I,S\_II}}^{\mathrm{Region}} = 2\int_{\mathrm{Region}} \frac{\rho_{\mathrm{S\_I}}^{\mathrm{avg}}\left(\vec{\mathrm{r}}\right)}{\rho^{\mathrm{avg}}\left(\vec{\mathrm{r}}\right)}\left\langle\frac{\rho_{\mathrm{S\_II}}^{\mathrm{avg}}\left(\vec{\mathrm{r}}'\right)}{\rho^{\mathrm{avg}}\left(\vec{\mathrm{r}}'\right)}\right\rangle_{\mathrm{DX\_hole}\left(\vec{\mathrm{r}}\right)}\rho\left(\vec{\mathrm{r}}\right)d^3\vec{\mathrm{r}} \tag{S46}$$

where Region = R_I or R_II.

Expanding

$$\mathrm{B}_{\mathrm{S\_I,S\_II}} = \mathrm{B}_{\mathrm{S\_I,S\_II}}^{\mathrm{R\_I}} + \mathrm{B}_{\mathrm{S\_I,S\_II}}^{\mathrm{R\_II}} \tag{S47}$$

$$\mathrm{CE}_{\mathrm{S\_I,S\_II}} = \mathrm{CE}_{\mathrm{S\_I,S\_II}}^{\mathrm{R\_I}} + \mathrm{CE}_{\mathrm{S\_I,S\_II}}^{\mathrm{R\_II}} \tag{S48}$$

$\mathrm{B}_{\mathrm{S\_I,S\_II}}^{\mathrm{R\_I}}$ can be re-written as

$$\begin{aligned}\mathrm{B}_{\mathrm{S\_I,S\_II}}^{\mathrm{R\_I}} &= 2\int_{\mathrm{R\_I}}\left(1-\frac{\rho_{\mathrm{S\_II}}^{\mathrm{avg}}\left(\vec{\mathrm{r}}\right)}{\rho^{\mathrm{avg}}\left(\vec{\mathrm{r}}\right)}\right)\left(1-\left\langle\frac{\rho_{\mathrm{S\_I}}^{\mathrm{avg}}\left(\vec{\mathrm{r}}'\right)}{\rho^{\mathrm{avg}}\left(\vec{\mathrm{r}}'\right)}\right\rangle_{\mathrm{DX\_hole}\left(\vec{\mathrm{r}}\right)}\right)\rho\left(\vec{\mathrm{r}}\right)d^3\vec{\mathrm{r}} \\ &= 2\int_{\mathrm{R\_I}}\left(\frac{\rho_{\mathrm{S\_II}}^{\mathrm{avg}}\left(\vec{\mathrm{r}}\right)}{\rho^{\mathrm{avg}}\left(\vec{\mathrm{r}}\right)}\left\langle\frac{\rho_{\mathrm{S\_I}}^{\mathrm{avg}}\left(\vec{\mathrm{r}}'\right)}{\rho^{\mathrm{avg}}\left(\vec{\mathrm{r}}'\right)}\right\rangle_{\mathrm{DX\_hole}\left(\vec{\mathrm{r}}\right)} + \Delta\left(\vec{\mathrm{r}}\right)\right)\rho\left(\vec{\mathrm{r}}\right)d^3\vec{\mathrm{r}}\end{aligned} \tag{S49}$$

where $\Delta\left(\vec{\mathrm{r}}\right)$ is given by Eq. (S35). Therefore,

$$\begin{aligned}\mathrm{B}_{\mathrm{S\_I,S\_II}} &= 2\int_{\mathrm{R\_I}}\frac{\rho_{\mathrm{S\_II}}^{\mathrm{avg}}\left(\vec{\mathrm{r}}\right)}{\rho^{\mathrm{avg}}\left(\vec{\mathrm{r}}\right)}\left\langle\frac{\rho_{\mathrm{S\_I}}^{\mathrm{avg}}\left(\vec{\mathrm{r}}'\right)}{\rho^{\mathrm{avg}}\left(\vec{\mathrm{r}}'\right)}\right\rangle_{\mathrm{DX\_hole}\left(\vec{\mathrm{r}}\right)}\rho\left(\vec{\mathrm{r}}\right)d^3\vec{\mathrm{r}} \\ &+2\int_{\mathrm{R\_II}}\frac{\rho_{\mathrm{S\_I}}^{\mathrm{avg}}\left(\vec{\mathrm{r}}\right)}{\rho^{\mathrm{avg}}\left(\vec{\mathrm{r}}\right)}\left\langle\frac{\rho_{\mathrm{S\_II}}^{\mathrm{avg}}\left(\vec{\mathrm{r}}'\right)}{\rho^{\mathrm{avg}}\left(\vec{\mathrm{r}}'\right)}\right\rangle_{\mathrm{DX\_hole}\left(\vec{\mathrm{r}}\right)}\rho\left(\vec{\mathrm{r}}\right)d^3\vec{\mathrm{r}} + 2\int_{\mathrm{R\_I}}\Delta\left(\vec{\mathrm{r}}\right)\rho\left(\vec{\mathrm{r}}\right)d^3\vec{\mathrm{r}}\end{aligned} \tag{S50}$$

Defining

$$\mathrm{term\_1} = 2\int\limits_{\mathrm{R\_I}} \frac{\rho_{\mathrm{S\_II}}^{\mathrm{avg}}(\vec{r})}{\rho^{\mathrm{avg}}(\vec{r})} \left\langle \frac{\rho_{\mathrm{S\_I}}^{\mathrm{avg}}(\vec{r}')}{\rho^{\mathrm{avg}}(\vec{r}')} \right\rangle_{\mathrm{DX\_hole}(\vec{r})} \rho(\vec{r})\,d^3\vec{r} \tag{S51}$$

$$\mathrm{term\_2} = 2\int\limits_{\mathrm{R\_II}} \frac{\rho_{\mathrm{S\_I}}^{\mathrm{avg}}(\vec{r})}{\rho^{\mathrm{avg}}(\vec{r})} \left\langle \frac{\rho_{\mathrm{S\_II}}^{\mathrm{avg}}(\vec{r}')}{\rho^{\mathrm{avg}}(\vec{r}')} \right\rangle_{\mathrm{DX\_hole}(\vec{r})} \rho(\vec{r})\,d^3\vec{r} \tag{S52}$$

$$\mathrm{term\_3} = 2\int\limits_{\mathrm{R\_I}} \Delta(\vec{r})\rho(\vec{r})\,d^3\vec{r} \tag{S53}$$

gives

$$B_{\mathrm{S\_I,S\_II}} = \mathrm{term\_1} + \mathrm{term\_2} + \mathrm{term\_3} \tag{S54}$$

From Eq. (S3), $B_{\mathrm{A,j}}$ will be computed to be identical regardless of whether we take the electron for atom A to be at position $\vec{r}$ and the electron for atom j to be at position $\vec{r}'$ or if we take the electron for atom A to be at position $\vec{r}'$ and the electron for atom j to be at position $\vec{r}$. Therefore, the value of $B_{\mathrm{S\_I,S\_II}}$ will be unchanged if we take the position of electron for S_II to be at $\vec{r}$ instead of $\vec{r}'$, and the position of electron for S_I to be at $\vec{r}'$ instead of $\vec{r}$. After analogous steps as described above, we finally arrive to

$$\begin{aligned} B_{\mathrm{S\_I,S\_II}} &= 2\int\limits_{\mathrm{R\_I}} \frac{\rho_{\mathrm{S\_II}}^{\mathrm{avg}}(\vec{r})}{\rho^{\mathrm{avg}}(\vec{r})} \left\langle \frac{\rho_{\mathrm{S\_I}}^{\mathrm{avg}}(\vec{r}')}{\rho^{\mathrm{avg}}(\vec{r}')} \right\rangle_{\mathrm{DX\_hole}(\vec{r})} \rho(\vec{r})\,d^3\vec{r} \\ &+2\int\limits_{\mathrm{R\_II}} \frac{\rho_{\mathrm{S\_I}}^{\mathrm{avg}}(\vec{r})}{\rho^{\mathrm{avg}}(\vec{r})} \left\langle \frac{\rho_{\mathrm{S\_II}}^{\mathrm{avg}}(\vec{r}')}{\rho^{\mathrm{avg}}(\vec{r}')} \right\rangle_{\mathrm{DX\_hole}(\vec{r})} \rho(\vec{r})\,d^3\vec{r} - 2\int\limits_{\mathrm{R\_II}} \Delta(\vec{r})\rho(\vec{r})\,d^3\vec{r} \end{aligned} \tag{S55}$$

Comparing Eq. (S50) to (S55) gives

$$\mathrm{term\_3} = -2\int\limits_{\mathrm{R\_II}} \Delta(\vec{r})\rho(\vec{r})\,d^3\vec{r} \tag{S56}$$

Combining Eqs. (S53) and (S56), the density weighted value of $\Delta(\vec{r})$ averages to zero over all space:

$$\oint \Delta(\vec{r})\rho(\vec{r})\,d^3\vec{r} = 0 \tag{S57}$$

Since $\rho_{\mathrm{S\_I}}^{\mathrm{avg}}(\vec{r}) / \rho^{\mathrm{avg}}(\vec{r}) \geq ½$ in Region I, and

$$0 \leq \left\langle \frac{\rho_{\mathrm{S\_I}}^{\mathrm{avg}}(\vec{r}')}{\rho^{\mathrm{avg}}(\vec{r}')} \right\rangle_{\mathrm{DX\_hole}(\vec{r})} \leq 1 \tag{S58}$$

it directly follows that

$$\frac{\mathrm{term\_1}}{\mathrm{CE}_{\mathrm{S\_I,S\_II}}^{\mathrm{R\_I}}} = \frac{2\int\limits_{\mathrm{R\_I}} \frac{\rho_{\mathrm{S\_II}}^{\mathrm{avg}}(\vec{r})}{\rho^{\mathrm{avg}}(\vec{r})} \left\langle \frac{\rho_{\mathrm{S\_I}}^{\mathrm{avg}}(\vec{r}')}{\rho^{\mathrm{avg}}(\vec{r}')} \right\rangle_{\mathrm{DX\_hole}(\vec{r})} \rho(\vec{r})\,d^3\vec{r}}{2\int\limits_{\mathrm{R\_I}} \frac{\rho_{\mathrm{S\_I}}^{\mathrm{avg}}(\vec{r})}{\rho^{\mathrm{avg}}(\vec{r})} \frac{\rho_{\mathrm{S\_II}}^{\mathrm{avg}}(\vec{r})}{\rho^{\mathrm{avg}}(\vec{r})} \rho(\vec{r})\,d^3\vec{r}} \leq 2 \tag{S59}$$

Since $\rho^{\text{avg}}_{\text{S\_II}}(\vec{r})/\rho^{\text{avg}}(\vec{r}) > ½$ in Region II, and

$$0 \le \left\langle \frac{\rho^{\text{avg}}_{\text{S\_II}}(\vec{r}')}{\rho^{\text{avg}}(\vec{r}')} \right\rangle_{\text{DX\_hole}(\vec{r})} \le 1 \tag{S60}$$

it directly follows that

$$\frac{\text{term\_2}}{\text{CE}^{\text{R\_II}}_{\text{S\_I,S\_II}}} = \frac{2\displaystyle\int_{\text{R\_II}} \frac{\rho^{\text{avg}}_{\text{S\_I}}(\vec{r})}{\rho^{\text{avg}}(\vec{r})} \left\langle \frac{\rho^{\text{avg}}_{\text{S\_II}}(\vec{r}')}{\rho^{\text{avg}}(\vec{r}')} \right\rangle_{\text{DX\_hole}(\vec{r})} \rho(\vec{r})\, d^3\vec{r}}{2\displaystyle\int_{\text{R\_II}} \frac{\rho^{\text{avg}}_{\text{S\_I}}(\vec{r})}{\rho^{\text{avg}}(\vec{r})} \frac{\rho^{\text{avg}}_{\text{S\_II}}(\vec{r})}{\rho^{\text{avg}}(\vec{r})} \rho(\vec{r})\, d^3\vec{r}} < 2 \tag{S61}$$

Therefore,

$$\frac{\text{B}_{\text{S\_I,S\_II}} - \text{term\_3}}{\text{CE}_{\text{S\_I,S\_II}}} = \frac{\text{term\_1} + \text{term\_2}}{\text{CE}^{\text{R\_I}}_{\text{S\_I,S\_II}} + \text{CE}^{\text{R\_II}}_{\text{S\_I,S\_II}}} \le 2 \tag{S62}$$

The bond-order-to-contact-exchange ratio is

$$\Phi_{\text{S\_I,S\_II}} = \frac{\text{B}_{\text{S\_I,S\_II}}}{\text{CE}_{\text{S\_I,S\_II}}} = \frac{\text{term\_1} + \text{term\_2} + \text{term\_3}}{\text{CE}^{\text{R\_I}}_{\text{S\_I,S\_II}} + \text{CE}^{\text{R\_II}}_{\text{S\_I,S\_II}}} \tag{S63}$$

Examining Eqs. (S35) and (S59), the largest ratio of integrands (i.e., a value of 2) in Eq. (S59) occurs where $\Delta(\vec{r}) \Rightarrow -½$. The $\Delta(\vec{r})$ term will be positive in Region I when $\frac{\rho^{\text{avg}}_{\text{S\_I}}(\vec{r})}{\rho^{\text{avg}}(\vec{r})} > \left\langle \frac{\rho^{\text{avg}}_{\text{S\_I}}(\vec{r}')}{\rho^{\text{avg}}(\vec{r}')} \right\rangle_{\text{DX\_hole}(\vec{r})}$. Examining Eq. (S59), $\Delta(\vec{r}) > 0$ in Region I suppresses the ratio of integrands in term_1 to $\text{CE}^{\text{R\_I}}_{\text{S\_I,S\_II}}$. Moreover, the larger the value of $\Delta(\vec{r}) > 0$ in Region I, the more the ratio of these integrands is suppressed. Analogously, if we evaluate term_3 as an integral over Region II (see Eq. (S56)) instead of Region I, then the ratio of term_2 to $\text{CE}^{\text{R\_II}}_{\text{S\_I,S\_II}}$ must be suppressed whenever the contribution to term_3 in Region II is positive. Therefore, increasing the value of term_3 leads to a partial cancellation of effects in the bond-order-to-contact-exchange ratio of Eq. (S63). Specifically, when term_3 is larger, then the ratios of term_1 to $\text{CE}^{\text{R\_I}}_{\text{S\_I,S\_II}}$ and term_2 to $\text{CE}^{\text{R\_II}}_{\text{S\_I,S\_II}}$ are suppressed.

Computational results for dozens of diverse real systems indicated there is an overall constant approximate upper bound on the ratio $\Phi_{\text{S\_I,S\_II}}$. From Eq. (S62),

$$\lim_{\text{term\_3}\to 0} \Phi_{\text{S\_I,S\_II}} \le 2 \tag{S64}$$

Considering the partial cancellation of effects between term_3 and (term_1 + term_2), we thus have an approximate upper bound

$$\Phi_{\text{S\_I,S\_II}} \lesssim 2 \tag{S65}$$

Alternatively, Eq. (S65) can be justified using purely theoretical arguments. Inserting Eqs. (S6)–(S9) into (S30) gives

$$\text{B}_{\text{S\_I,S\_II}} = 2\oint\oint \frac{\rho^{\text{avg}}_{\text{S\_I}}(\vec{r})}{\rho^{\text{avg}}(\vec{r})} \left\langle \frac{\rho^{\text{avg}}_{\text{S\_II}}(\vec{r}')}{\rho^{\text{avg}}(\vec{r}')} \right\rangle_{\text{DX\_hole}(\vec{r})} \rho(\vec{r}'')\rho^{\text{DX\_hole}}(\vec{r}'',\vec{r})\, d^3\vec{r}\, d^3\vec{r}'' \tag{S66}$$

Similarly,

$$\mathrm{CE}_{\mathrm{S\_I,S\_II}} = 2\oint\oint \frac{\rho^{\mathrm{avg}}_{\mathrm{S\_I}}(\vec{r})}{\rho^{\mathrm{avg}}(\vec{r})}\frac{\rho^{\mathrm{avg}}_{\mathrm{S\_II}}(\vec{r})}{\rho^{\mathrm{avg}}(\vec{r})}\rho(\vec{r}'')\rho^{\mathrm{DX\_hole}}(\vec{r}'',\vec{r})\,d^3\vec{r}\,d^3\vec{r}'' \tag{S67}$$

Integrating over $\vec{r}$ yields

$$\mathrm{B}_{\mathrm{S\_I,S\_II}} = 2\oint \left\langle \frac{\rho^{\mathrm{avg}}_{\mathrm{S\_I}}(\vec{r})}{\rho^{\mathrm{avg}}(\vec{r})}\left\langle \frac{\rho^{\mathrm{avg}}_{\mathrm{S\_II}}(\vec{r}')}{\rho^{\mathrm{avg}}(\vec{r}')}\right\rangle_{\mathrm{DX\_hole}(\vec{r})}\right\rangle_{\mathrm{DX\_hole}(\vec{r}'')}\rho(\vec{r}'')\,d^3\vec{r}'' \tag{S68}$$

$$\mathrm{CE}_{\mathrm{S\_I,S\_II}} = 2\oint \left\langle \frac{\rho^{\mathrm{avg}}_{\mathrm{S\_I}}(\vec{r})}{\rho^{\mathrm{avg}}(\vec{r})}\frac{\rho^{\mathrm{avg}}_{\mathrm{S\_II}}(\vec{r})}{\rho^{\mathrm{avg}}(\vec{r})}\right\rangle_{\mathrm{DX\_hole}(\vec{r}'')}\rho(\vec{r}'')\,d^3\vec{r}'' \tag{S69}$$

Because

$$\tfrac{1}{2} \le \max\left(\frac{\rho^{\mathrm{avg}}_{\mathrm{S\_I}}(\vec{r})}{\rho^{\mathrm{avg}}(\vec{r})},\frac{\rho^{\mathrm{avg}}_{\mathrm{S\_II}}(\vec{r})}{\rho^{\mathrm{avg}}(\vec{r})}\right) \le 1 \tag{S70}$$

most typical continuous distributions yield

$$\left\langle \frac{\rho^{\mathrm{avg}}_{\mathrm{S\_I}}(\vec{r})}{\rho^{\mathrm{avg}}(\vec{r})}\frac{\rho^{\mathrm{avg}}_{\mathrm{S\_II}}(\vec{r})}{\rho^{\mathrm{avg}}(\vec{r})}\right\rangle_{\mathrm{DX\_hole}(\vec{r}'')} \gtrsim \tfrac{1}{2}\min\left(\left\langle \frac{\rho^{\mathrm{avg}}_{\mathrm{S\_I}}(\vec{r})}{\rho^{\mathrm{avg}}(\vec{r})}\right\rangle_{\mathrm{DX\_hole}(\vec{r}'')},\left\langle \frac{\rho^{\mathrm{avg}}_{\mathrm{S\_II}}(\vec{r})}{\rho^{\mathrm{avg}}(\vec{r})}\right\rangle_{\mathrm{DX\_hole}(\vec{r}'')}\right) \tag{S71}$$

$$\left\langle \frac{\rho^{\mathrm{avg}}_{\mathrm{S\_I}}(\vec{r})}{\rho^{\mathrm{avg}}(\vec{r})}\left\langle \frac{\rho^{\mathrm{avg}}_{\mathrm{S\_II}}(\vec{r}')}{\rho^{\mathrm{avg}}(\vec{r}')}\right\rangle_{\mathrm{DX\_hole}(\vec{r})}\right\rangle_{\mathrm{DX\_hole}(\vec{r}'')} \lesssim \min\left(\left\langle \frac{\rho^{\mathrm{avg}}_{\mathrm{S\_I}}(\vec{r})}{\rho^{\mathrm{avg}}(\vec{r})}\right\rangle_{\mathrm{DX\_hole}(\vec{r}'')},\left\langle \frac{\rho^{\mathrm{avg}}_{\mathrm{S\_II}}(\vec{r})}{\rho^{\mathrm{avg}}(\vec{r})}\right\rangle_{\mathrm{DX\_hole}(\vec{r}'')}\right) \tag{S72}$$

Comparing Eqs. (S68), (S69), (S71) and (S72) yields Eq. (S65) again.

The above analysis holds for any chosen sets S_I and S_II that sum to the whole system. Choosing S_I to be a single atom and S_II to be all of the other atoms in the system, it immediately follows that

$$\frac{\mathrm{SBO}_{\mathrm{A}}}{\mathrm{SCE}_{\mathrm{A}}} \lesssim 2 \tag{S73}$$

for each atom in the material. We may reconstitute S_I and S_II to form any binary division of the entire system. Eq. (S65) will hold for any chosen set S_I if

$$\frac{\mathrm{B}_{\mathrm{A,j}}}{\mathrm{CE}_{\mathrm{A,j}}} \lesssim 2 \tag{S74}$$

holds for each bond in the material. Therefore, we impose constraint (S74) to ensure Eq. (S65) holds for every possible set S_I.

Finally, we must consider the effects of magnetism. As explained in the previous section, a charge and spin partitioning method used to compute these bond orders should be optimized to satisfy the confluence relations in Eq. (S44). Because a charge and spin partitioning optimized to satisfy the confluence relations in Eq. (S44) has similar exchange propensity factors for magnetic and non-magnetic systems, it follows that the upper bound on $\Phi_{\mathrm{A,j}}$ is equivalent for magnetic and non-magnetic systems.

S4. Derivation of the bond-order-to-contact-exchange ratio for a buried tail

As described in Section S2, we can divide a material into two sets of atoms, S_I and S_II, such that $\mathrm{S\_I} \cap \mathrm{S\_II} = \varnothing$ and $\mathrm{S\_I} \cup \mathrm{S\_II}$ contains all atoms in the material (not just in the reference unit cell). As before, we first consider non-magnetic systems.

We define the buried tail of S_I to be the region of space for which

$$\left\langle \frac{\rho_{\mathrm{S\_II}}^{\mathrm{avg}}(\vec{r}')}{\rho^{\mathrm{avg}}(\vec{r}')} \right\rangle_{\mathrm{DX\_hole}(\vec{r})}^{\mathrm{tail}} \approx 1 \qquad \text{(S75)}$$

$$0 \le \left\langle \frac{\rho_{\mathrm{S\_I}}^{\mathrm{avg}}(\vec{r}')}{\rho^{\mathrm{avg}}(\vec{r}')} \right\rangle_{\mathrm{DX\_hole}(\vec{r})}^{\mathrm{tail}} << 1 \qquad \text{(S76)}$$

Since $\rho_{\mathrm{S\_II}}^{\mathrm{avg}}(\vec{r}) / \rho^{\mathrm{avg}}(\vec{r})$ varies smoothly and continuously, Eqs. (S75) and (S76) can be true only if

$$\frac{\rho_{\mathrm{S\_II}}^{\mathrm{avg}}(\vec{r})}{\rho^{\mathrm{avg}}(\vec{r})} \approx 1 \qquad \text{(S77)}$$

$$0 \le \frac{\rho_{\mathrm{S\_I}}^{\mathrm{avg}}(\vec{r})}{\rho^{\mathrm{avg}}(\vec{r})} \ll 1 \qquad \text{(S78)}$$

within the buried tail of S_I. From Eqs. (S29) and (S30), the local ratio of bond-order-to-contact-exchange integrands in the buried tail is therefore

$$\frac{2 \frac{\rho_{\mathrm{S\_I}}^{\mathrm{avg}}(\vec{r})}{\rho^{\mathrm{avg}}(\vec{r})} \overbrace{\left\langle \frac{\rho_{\mathrm{S\_II}}^{\mathrm{avg}}(\vec{r}')}{\rho^{\mathrm{avg}}(\vec{r}')} \right\rangle_{\mathrm{DX\_hole}(\vec{r})}}^{\sim 1} \rho(\vec{r})}{2 \frac{\rho_{\mathrm{S\_I}}^{\mathrm{avg}}(\vec{r})}{\rho^{\mathrm{avg}}(\vec{r})} \underbrace{\frac{\rho_{\mathrm{S\_II}}^{\mathrm{avg}}(\vec{r})}{\rho^{\mathrm{avg}}(\vec{r})}}_{\sim 1} \rho(\vec{r})} \approx 1 \qquad \text{(S79)}$$

The above analysis holds for any chosen sets S_I and S_II that sum to the whole system. Choosing S_I to be a single atom and S_II to be all of the other atoms in the system, it immediately follows that the ratio of $\mathrm{SBO_A}$ integrand contributions to $\mathrm{SCE_A}$ integrand contributions is ~1 in the buried tail of atom A. We may reconstitute S_I and S_II to form any binary division of the entire system. Eq. (S79) will hold for any chosen set S_I if

$$\frac{B_{A,j}^{\mathrm{tail}}}{CE_{A,j}^{\mathrm{tail}}} \approx 1 \qquad \text{(S80)}$$

holds for each bond in the material, where $B_{A,j}^{\mathrm{tail}}$ and $CE_{A,j}^{\mathrm{tail}}$ are the contributions to $B_{A,j}$ and $CE_{A,j}$, respectively, within the buried tail of atom A. Therefore, we impose constraint (S80) to ensure Eq. (S79) holds for every possible set S_I.

Finally, we must consider the effects of magnetism. As explained above, a charge and spin partitioning method used to compute these bond orders should be optimized to satisfy the confluence relations in Eq. (S44). Because a charge and spin partitioning optimized to satisfy the confluence relations in Eq. (S44) has similar exchange propensity factors for magnetic and non-magnetic systems, it follows that Eqs. (S79) and (S80) hold for both magnetic and non-magnetic systems.

S5. Constraint on density-derived localization index

Substituting Eqs. (S15) and (S8) into (S3) gives the following fundamental expression for bond order:

$$B_{A,j} = 2\oint\oint \frac{\vec{\rho}_A^{\,avg}(r_A)\bullet\vec{\rho}_j^{\,avg}(r_j')}{\vec{\rho}^{\,avg}(\vec{r})\bullet\vec{\rho}^{\,avg}(\vec{r}')}\rho^{DX}(\vec{r},\vec{r}')d^3\vec{r}d^3\vec{r}' = 2\oint\oint \frac{\vec{\rho}_A^{\,avg}(r_A)\bullet\vec{\rho}_j^{\,avg}(r_j')}{\vec{\rho}^{\,avg}(\vec{r})\bullet\vec{\rho}^{\,avg}(\vec{r}')}\rho(\vec{r})\rho^{DX\_hole}(\vec{r},\vec{r}')d^3\vec{r}d^3\vec{r}' \quad (S81)$$

Based on Eq. (S81), one might propose to explicitly construct and explicitly integrate dressed exchange holes in materials. However, closer analysis shows this is neither necessary nor preferable. Because $\rho^{DX\_hole}(\vec{r},\vec{r}')$ explicitly depends on six spatial coordinates (i.e., $\vec{r}$ and $\vec{r}'$), its explicit construction and integration would be six-dimensional and computationally expensive. From Eqs. (S41) and (S74), $CE_{A,j}$ accounts for between 50% and 100% of the $B_{A,j}$ value. As shown in Eq. (S21), computing $CE_{A,j}$ requires only a three-dimensional integration over $\vec{r}$. Thus, by using three-dimensional integrals instead of a six-dimensional integral we can manifestly calculate the majority of factors contributing to $B_{A,j}$.

As shown in Eq. (S9), $\rho^{DX\_hole}(\vec{r},\vec{r}')$ is a normalized probability distribution that integrates over $\vec{r}'$ to equal one. Therefore, integrating over $\vec{r}'$ in Eq. (S81) averages the integrand over the dressed exchange hole. Only the remaining three integration dimensions contribute beyond property averaging. This is fundamentally why we can approximate the six-dimensional integral in Eq. (S81) using three-dimensional integrals.

The key question is how to construct a set of three-dimensional integrals that accurately approximates the six-dimensional integral in Eq. (S81). To do this, we can estimate the six-dimensional integral as a series expansion of terms involving only three-dimensional integrals. Constraints and scaling relations incorporated into this series expansion will keep it well-behaved so that it never deviates far from the true value of the six-dimensional integral in Eq. (S81).

The remainder of this section studies one such constraint. Because the density-derived localization index, $B_{A,A}$, represents the number of electrons localized on atom A, it must be non-negative to be physically meaningful. Here I show the constraint on $B_{A,A}$ should even be stronger than simply imposing $B_{A,A} \geq 0$. From Eq. (S4),

$$B_{A,A} = \oint\oint \rho_{A,A}^{DX}(\vec{r},\vec{r}')d^3\vec{r}d^3\vec{r}' = N_A - ½SBO_A \quad (S82)$$

is the integrated probability that an electron at position $\vec{r}$ and assigned to atom A dressed-exchanges with an electron at position $\vec{r}'$ and assigned to atom A. Using Eq. (S30) with S_I = atom A and S_II = all other atoms in the material, Eq. (S82) simplifies to

$$B_{A,A} \approx \underbrace{\oint \frac{\rho_A(\vec{r}_A)}{\rho(\vec{r})}\rho(\vec{r})d^3\vec{r}}_{N_A} - \underbrace{\oint \frac{\rho_A^{avg}(r_A)}{\rho^{avg}(\vec{r})}\left\langle 1 - \frac{\rho_A^{avg}(r_A')}{\rho^{avg}(\vec{r}')}\right\rangle_{DX\_hole(\vec{r})}\rho(\vec{r})d^3\vec{r}}_{\approx ½SBO_A} \quad (S83)$$

Substituting the confluence of atomic exchange propensities, $\frac{\rho_A^{avg}(r_A)}{\rho^{avg}(\vec{r})} \approx \frac{\rho_A(\vec{r}_A)}{\rho(\vec{r})}$, from Eq. (S44) into the integral for N_A in Eq. (S83) and simplifying gives

$$B_{A,A} \approx \oint \frac{\rho_A^{avg}(r_A)}{\rho^{avg}(\vec{r})} \left\langle \frac{\rho_A^{avg}(r_A')}{\rho^{avg}(\vec{r}')} \right\rangle_{DX\_hole(\vec{r})} \rho(\vec{r}) d^3\vec{r} \geq 0 \tag{S84}$$

It directly follows from Eq. (S84) that $B_{A,A} > 0$ when $N_A > 0$.

Following Eqs. (S22) and (S21), a similar expansion can be performed for the self-contact exchange, $CE_{A,A}$:

$$CE_{A,A} \approx \underbrace{\oint \frac{\rho_A(\vec{r}_A)}{\rho(\vec{r})} \rho(\vec{r}) d^3\vec{r}}_{N_A} - \underbrace{\oint \frac{\vec{\rho}_A^{avg}(r_A) \bullet \left(\vec{\rho}^{avg}(\vec{r}) - \vec{\rho}_A^{avg}(r_A)\right)}{\vec{\rho}^{avg}(\vec{r}) \bullet \vec{\rho}^{avg}(\vec{r})} \rho(\vec{r}) d^3\vec{r}}_{½SCE_A} \tag{S85}$$

Substituting the confluence of atomic exchange propensities, $\vec{\rho}_A^{avg}(r_A) \approx \frac{\rho_A(\vec{r}_A)}{\rho(\vec{r})} \vec{\rho}^{avg}(\vec{r}) \approx \frac{\rho_A^{avg}(r_A)}{\rho^{avg}(\vec{r})} \vec{\rho}^{avg}(\vec{r})$, from Eq. (S44) into the integral for SCE_A in Eq. (S85) and simplifying gives

$$CE_{A,A} \approx \oint \left(\frac{\rho_A(\vec{r}_A)}{\rho(\vec{r})} \right)^2 \rho(\vec{r}) d^3\vec{r} \approx \oint \left(\frac{\rho_A^{avg}(r_A)}{\rho^{avg}(\vec{r})} \right)^2 \rho(\vec{r}) d^3\vec{r} \geq 0 \tag{S86}$$

It directly follows from Eq. (S86) that $CE_{A,A} > 0$ when $N_A > 0$.

The integrands ratio for $B_{A,A}$ in Eq. (S84) to $CE_{A,A}$ in Eq. (S86) is $\left\langle \frac{\rho_A^{avg}(r_A')}{\rho^{avg}(\vec{r}')} \right\rangle_{DX\_hole(\vec{r})} \Big/ \left(\frac{\rho_A^{avg}(r_A)}{\rho^{avg}(\vec{r})} \right)$. This is analogous to the integrands ratio $\left\langle \frac{\rho_{S\_II}^{avg}(\vec{r}')}{\rho^{avg}(\vec{r}')} \right\rangle_{DX\_hole(\vec{r})} \Big/ \left(\frac{\rho_{S\_II}^{avg}(\vec{r})}{\rho^{avg}(\vec{r})} \right)$ for SBO_A in Eq. (S30) to SCE_A in Eq. (S29), when taking S_I = atom A and S_II = all other atoms in the material. This implies we can formulate a lower and upper bound on the ratio of $B_{A,A}$ to $CE_{A,A}$ similar to that we did for $B_{A,j}$ to $CE_{A,j}$ when $j \neq A$. As shown by Eqs. (S41) and (S74), $2CE_{A,j} \gtrsim B_{A,j} \gtrsim CE_{A,j}$ for $j \neq A$. This means $B_{A,j}$ and $CE_{A,j}$ are within a factor of 2 different from each other. The relationship between the bond order and the contact exchange will thus be most meaningful if we also require that $B_{A,A}$ and the self-contact exchange, $CE_{A,A}$, are within a factor of 2 different from each other. Because $CE_{A,A} \geq B_{A,A}$, this means

$$2B_{A,A} \geq CE_{A,A} \geq B_{A,A} \geq 0 \tag{S87}$$

This constraint ensures the behavior of $B_{A,A}$ is well controlled. First, when $N_A > 0$, this ensures $B_{A,A} > 0$ because $CE_{A,A} > 0$. Second, it ensures the ratio $CE_{A,A} / B_{A,A}$ does not become too large.

S6. Derivation of the comprehensive bond order equation

Following the lower and upper bounds derived above, the bond order can be written as

$$B_{A,j} = CE_{A,j} + \Lambda_{A,j} \tag{S88}$$

$$0 \leq \Lambda_{A,j} \leq CE_{A,j} \tag{S89}$$

where $\Lambda_{A,j}$ accounts for the delocalization of the dressed-exchange hole that occurs because the dressed-exchange hole around each position $\vec{r}$ is in fact not a Dirac delta function but has a non-negligible spatial extent. $\Lambda_{A,j}$ should depend only on three-dimensional integrals of functions of the set of atomic exchange propensity factors and $\rho(\vec{r})$:

$$\Lambda_{A,j} = \text{function}\left(\left\{\oint \text{functions}\left(\left\{\frac{\vec{\rho}_i^{avg}(r_i)\bullet\vec{\rho}_k^{avg}(r_k)}{\vec{\rho}^{avg}(\vec{r})\bullet\vec{\rho}^{avg}(\vec{r})}\right\}\right)\rho(\vec{r})d^3\vec{r}\right\}\right) \tag{S90}$$

In principle, the integrands of the functions in Eq. (S90) can simultaneously depend on every member of the set $\left\{\left(\vec{\rho}_i^{avg}(r_i)\bullet\vec{\rho}_k^{avg}(r_k)\right)\Big/\left(\vec{\rho}^{avg}(\vec{r})\bullet\vec{\rho}^{avg}(\vec{r})\right)\right\}$ for any or all possible pairs (i,k). However, terms of the form

$$\oint\left(\frac{\vec{\rho}_i^{avg}(r_i)\bullet\vec{\rho}_j^{avg}(r_j)}{\vec{\rho}^{avg}(\vec{r})\bullet\vec{\rho}^{avg}(\vec{r})}\right)\left(\frac{\vec{\rho}_j^{avg}(r_j)\bullet\vec{\rho}_k^{avg}(r_k)}{\vec{\rho}^{avg}(\vec{r})\bullet\vec{\rho}^{avg}(\vec{r})}\right)\rho(\vec{r})d^3\vec{r} \Rightarrow \text{expensive} \tag{S91}$$

will be computationally expensive to integrate, because the number of three-body or higher-order terms is large in extended materials. Since our goal is to develop a computationally efficient method, we therefore restrict the allowed integrands to two-body terms. Thus, functions like Eq. (S91) are not used, while functions like

$$\left(\oint\left(\frac{\vec{\rho}_i^{avg}(r_i)\bullet\vec{\rho}_j^{avg}(r_j)}{\vec{\rho}^{avg}(\vec{r})\bullet\vec{\rho}^{avg}(\vec{r})}\right)\rho(\vec{r})d^3\vec{r}\right)\left(\oint\left(\frac{\vec{\rho}_j^{avg}(r_j)\bullet\vec{\rho}_k^{avg}(r_k)}{\vec{\rho}^{avg}(\vec{r})\bullet\vec{\rho}^{avg}(\vec{r})}\right)\rho(\vec{r})d^3\vec{r}\right) \Rightarrow \text{inexpensive} \tag{S92}$$

$$\left(\oint\left(\frac{\vec{\rho}_i^{avg}(r_i)\bullet\vec{\rho}_j^{avg}(r_j)}{\vec{\rho}^{avg}(\vec{r})\bullet\vec{\rho}^{avg}(\vec{r})}\right)^2\rho(\vec{r})d^3\vec{r}\right) \Rightarrow \text{inexpensive} \tag{S93}$$

are allowed.

These functions can be divided into types. Functions of the first type, called pairwise terms, describe an effect on $\Lambda_{A,j}$ in terms of the ratio $\left(\vec{\rho}_A^{avg}(r_A)\bullet\vec{\rho}_j^{avg}(r_j)\right)\Big/\left(\vec{\rho}^{avg}(\vec{r})\bullet\vec{\rho}^{avg}(\vec{r})\right)$ for the same atom pair. Functions of the second type, called coordination number terms, describe an effect on $\Lambda_{A,j}$ in terms of the ratios $\left(\vec{\rho}_A^{avg}(r_A)\bullet\vec{\rho}_k^{avg}(r_k)\right)\Big/\left(\vec{\rho}^{avg}(\vec{r})\bullet\vec{\rho}^{avg}(\vec{r})\right)$ and $\left(\vec{\rho}_i^{avg}(r_i)\bullet\vec{\rho}_j^{avg}(r_j)\right)\Big/\left(\vec{\rho}^{avg}(\vec{r})\bullet\vec{\rho}^{avg}(\vec{r})\right)$ for atom pairs that share at least one atom with the pair (A,j) . Since $\Lambda_{A,j}$ has two indices (not three), the third index must be eliminated through summation. Therefore, terms of the second type can only enter the bond order equation through sums over the third index. As one example, terms of the form

$$\sum_k\left(\oint\left(\frac{\vec{\rho}_A^{avg}(r_A)\bullet\vec{\rho}_k^{avg}(r_k)}{\vec{\rho}^{avg}(\vec{r})\bullet\vec{\rho}^{avg}(\vec{r})}\right)\rho(\vec{r})d^3\vec{r}\right)^2 \text{ and } \sum_k\left(\oint\left(\frac{\vec{\rho}_j^{avg}(r_j)\bullet\vec{\rho}_k^{avg}(r_k)}{\vec{\rho}^{avg}(\vec{r})\bullet\vec{\rho}^{avg}(\vec{r})}\right)\rho(\vec{r})d^3\vec{r}\right)^2 \tag{S94}$$

are allowed in $\Lambda_{A,j}$. Functions of the third type depend on atom pairs that contain neither A nor j. Because these terms contain neither A nor j, they must be secondary effects. Since our goal is to derive the simplest possible bond order equation that describes the primary bond order effects, functions of the third type must be omitted in $\Lambda_{\mathrm{A,j}}$.

For the following reasons, $\Lambda_{\mathrm{A,j}}$ can be approximately factored into a product of three functions:

$$\Lambda_{\mathrm{A,j}} \approx \chi_{\mathrm{A,j}}^{\mathrm{coord\_num}} \chi_{\mathrm{A,j}}^{\mathrm{pairwise}} \chi_{\mathrm{A,j}}^{\mathrm{constraint}} \tag{S95}$$

where $\chi_{\mathrm{A,j}}^{\mathrm{pairwise}}$ describes the type one (i.e., pairwise) functions, $\chi_{\mathrm{A,j}}^{\mathrm{coord\_num}}$ describes the type two (i.e., coordination number) functions, and $\chi_{\mathrm{A,j}}^{\mathrm{constraint}}$ enforces constraint (S87) on $\mathrm{B_{A,A}}$ and $\mathrm{B_{j,j}}$. We now consider two separate cases:

(1) *Case 1*: If the constraint on $\mathrm{B_{A,A}}$ or $\mathrm{B_{j,j}}$ is binding, then $\Lambda_{\mathrm{A,j}}$ equals the constrained value and $0 < \chi_{\mathrm{A,j}}^{\mathrm{constraint}} < 1$.

(2) *Case 2*: When this constraint is non-binding, then $\chi_{\mathrm{A,j}}^{\mathrm{constraint}} = 1$ and $\Lambda_{\mathrm{A,j}} \approx \chi_{\mathrm{A,j}}^{\mathrm{coord\_num}} \chi_{\mathrm{A,j}}^{\mathrm{pairwise}}$. For a polyatomic molecule, we could hypothetically compute what $\Lambda_{\mathrm{A,j}}$ would be if all of the type two (aka 'coordination number') terms were neglected. Let us call this quantity $\chi_{\mathrm{A,j}}^{\mathrm{pairwise}}$, because it contains only the type one (i.e., pairwise) terms that depend on the pair $(\mathrm{A,j})$. Then, let us define the ratio

$$\chi_{\mathrm{A,j}}^{\mathrm{ratio}} = \Lambda_{\mathrm{A,j}} / \chi_{\mathrm{A,j}}^{\mathrm{pairwise}} \tag{S96}$$

Since a diatomic molecule has only the pair $(\mathrm{A,j})$ and no other atom pairs, manifestly $\Lambda_{\mathrm{A,j}} = \chi_{\mathrm{A,j}}^{\mathrm{pairwise}}$ and $\chi_{\mathrm{A,j}}^{\mathrm{ratio}} = 1$ for any diatomic molecule. Since $\chi_{\mathrm{A,j}}^{\mathrm{ratio}}$ must have the same value of 1 for all diatomic molecules, manifestly it cannot have any primary dependence on the type one (i.e., pairwise) functions. Thus, the leading order deviation of $\chi_{\mathrm{A,j}}^{\mathrm{ratio}}$ from one could hypothetically depend on either (i) type two functions or (ii) products of type one and type two functions that become zero for diatomic molecules. Since the product of type two and type one functions will become negligible as $\mathrm{CE_{A,j}} \to 0$, for sufficiently small $\mathrm{CE_{A,j}}$ the (i) type two functions will form the dominant deviation of $\chi_{\mathrm{A,j}}^{\mathrm{ratio}}$ from one. Therefore, the leading order deviation of $\chi_{\mathrm{A,j}}^{\mathrm{ratio}}$ from one can be approximated by type two functions alone. Thus,

$$\chi_{\mathrm{A,j}}^{\mathrm{ratio}} \approx \chi_{\mathrm{A,j}}^{\mathrm{coord\_num}} \tag{S97}$$

where $\chi_{\mathrm{A,j}}^{\mathrm{coord\_num}}$ depends on type two functions alone. Note that $\chi_{\mathrm{A,j}}^{\mathrm{pairwise}} = 0$ only if atoms A and j have no density overlap, in which case $\Lambda_{\mathrm{A,j}} = 0$ and $\chi_{\mathrm{A,j}}^{\mathrm{coord\_num}}$ remains finite. Thus, $\Lambda_{\mathrm{A,j}}$, $\chi_{\mathrm{A,j}}^{\mathrm{pairwise}}$, and $\chi_{\mathrm{A,j}}^{\mathrm{coord\_num}}$ always have finite, non-negative values. As derived below, $0 < \chi_{\mathrm{A,j}}^{\mathrm{coord\_num}} \le 1$.

We now consider the form of $\chi_{\mathrm{A,j}}^{\mathrm{pairwise}}$. Following Eqs. (S90) and (S95), $\chi_{\mathrm{A,j}}^{\mathrm{pairwise}}$ is allowed to contain sums or products of terms having the forms

$$\oint \left(\frac{\vec{\rho}_{A}^{avg}\left(r_A\right) \bullet \vec{\rho}_{j}^{avg}\left(r_j\right)}{\vec{\rho}^{avg}\left(\vec{r}\right) \bullet \vec{\rho}^{avg}\left(\vec{r}\right)} \right)^{n} \rho\left(\vec{r}\right) d^3\vec{r} \tag{S98}$$

To proceed further, a parameterization is constructed that allows the correct order of these terms to be derived. For a non-magnetic material, define

$$f = \frac{\rho_A^{avg}\left(r_A\right) + \rho_j^{avg}\left(r_j\right)}{\rho^{avg}\left(\vec{r}\right)} \tag{S99}$$

$$s = \frac{\rho_A^{avg}\left(r_A\right)}{\rho_A^{avg}\left(r_A\right) + \rho_j^{avg}\left(r_j\right)} - \left\langle \frac{\rho_A^{avg}\left(r'_A\right)}{\rho_A^{avg}\left(r'_A\right) + \rho_j^{avg}\left(r'_j\right)} \right\rangle_{hole\left(\vec{r}\right)} \tag{S100}$$

To derive the term order, we can consider the special example where: (a) the spatial dependence of $\rho\left(\vec{r}\right)$ and $\left(\rho_A^{avg}\left(r_A\right) + \rho_j^{avg}\left(r_j\right)\right)$ are neglected, (b) A and j have equal average densities within the exchange hole, and (c) the parameter s is uniformly distributed over the range

$$-\tfrac{1}{2}f \le s \le \tfrac{1}{2}f \tag{S101}$$

For f = 1, we would expect s to vary from -½ (where $\rho_A^{avg}\left(r_A\right) = 0$) to ½ (where $\rho_A^{avg}\left(r_A\right) = \rho^{avg}\left(\vec{r}\right)$) over the exchange hole. When $f < 1$, this means other atoms contribute electron density to the exchange hole, and this makes the actual exchange hole in the polyatomic material smaller than the exchange hole that would hypothetically result from $\left(\rho_A^{avg}\left(r_A\right) + \rho_j^{avg}\left(r_j\right)\right)$ alone in a diatomic system neglecting the other atoms. The smaller exchange hole size means s will vary less over this smaller spatial region. Taking the range of $-\tfrac{1}{2} \le s \le \tfrac{1}{2}$ over the hypothetical $\left(\rho_A^{avg}\left(r_A\right) + \rho_j^{avg}\left(r_j\right)\right)$ exchange hole that would result for a diatomic system neglecting the other atoms, this implies an s range only f as large (Eq. (S101)) over the actual exchange hole in the polyatomic material. While we do not intend to apply the severe restrictions (a), (b), and (c) generally, such a restricted case should have the same term order as the general case, thereby allowing us to infer the term order of the general case from the restricted case. For this example, the integrand of the contact exchange is

$$\frac{\rho_A^{avg}\left(r_A\right)\rho_j^{avg}\left(r_j\right)}{\rho^{avg}\left(\vec{r}\right)\rho^{avg}\left(\vec{r}\right)}\rho\left(\vec{r}\right) = \left(\frac{\rho_A^{avg}\left(r_A\right) + \rho_j^{avg}\left(r_j\right)}{\rho^{avg}\left(\vec{r}\right)} \right)^2 \frac{\rho_A^{avg}\left(r_A\right)\rho_j^{avg}\left(r_j\right)}{\left(\rho_A^{avg}\left(r_A\right) + \rho_j^{avg}\left(r_j\right)\right)^2}\rho\left(\vec{r}\right) \Rightarrow f^2\left(\tfrac{1}{2}+s\right)\left(\tfrac{1}{2}-s\right)\rho \tag{S102}$$

Thus,

$$\chi_{A,j}^{pairwise} = \left(\frac{B_{A,j}}{CE_{A,j}} - 1 \right) CE_{A,j} \approx \left(\frac{2\int_{-\frac{1}{2}f}^{\frac{1}{2}f} f^2\left(\tfrac{1}{2}+s\right) \overbrace{\tfrac{1}{2}}^{\left\langle \frac{\rho_j^{avg}}{\rho_A^{avg}+\rho_j^{avg}} \right\rangle_{hole}} \rho ds}{2\int_{-\frac{1}{2}f}^{\frac{1}{2}f} f^2 \underbrace{\left(\tfrac{1}{2}+s\right)}_{\frac{\rho_A^{avg}}{\rho_A^{avg}+\rho_j^{avg}}} \underbrace{\left(\tfrac{1}{2}-s\right)}_{\frac{\rho_j^{avg}}{\rho_A^{avg}+\rho_j^{avg}}} \rho ds} - 1 \right) CE_{A,j} \tag{S103}$$

The subtraction of 1 in Eq. (S103) occurs, because we are computing only the bond order contribution in excess of the contact exchange. The integrals in Eq. (S103) are easy to evaluate analytically and yield

$$\chi_{A,j}^{\text{pairwise}} \approx \left(\frac{3}{3-f^2}-1\right)\text{CE}_{A,j} \tag{S104}$$

Since

$$\text{CE}_{A,j} \propto 2\int_{-\frac{1}{2}f}^{\frac{1}{2}f} f^2 \underbrace{\left(\tfrac{1}{2}+s\right)}_{\frac{\rho_A^{\text{avg}}}{\rho_A^{\text{avg}}+\rho_j^{\text{avg}}}}\underbrace{\left(\tfrac{1}{2}-s\right)}_{\frac{\rho_j^{\text{avg}}}{\rho_A^{\text{avg}}+\rho_j^{\text{avg}}}} ds = \frac{\left(3-f^2\right)f^3}{6} \tag{S105}$$

this gives

$$\chi_{A,j}^{\text{pairwise}} \propto \left(\frac{f^2}{3-f^2}\right)\frac{\left(3-f^2\right)f^3}{6} = \frac{f^5}{6} \tag{S106}$$

Thus, the leading order term in $\chi_{A,j}^{\text{pairwise}}$ is proportional to f^5. For the restricted conditions listed above,

$$\oint\left(\frac{\vec{\rho}_A^{\text{avg}}\left(r_A\right)\bullet\vec{\rho}_j^{\text{avg}}\left(r_j\right)}{\vec{\rho}^{\text{avg}}\left(\vec{r}\right)\bullet\vec{\rho}^{\text{avg}}\left(\vec{r}\right)}\right)^2\rho\left(\vec{r}\right)d^3\vec{r} \Rightarrow \int_{-\frac{1}{2}f}^{\frac{1}{2}f} f^4\left(\tfrac{1}{2}+s\right)^2\left(\tfrac{1}{2}-s\right)^2\rho ds \tag{S107}$$

Evaluating the integral analytically,

$$\int_{-\frac{1}{2}f}^{\frac{1}{2}f} f^2\left(\tfrac{1}{2}+s\right)^2\left(\tfrac{1}{2}-s\right)^2 ds = \frac{\left(15-10\cdot f^2+3\cdot f^4\right)}{240}f^5 \tag{S108}$$

From Eq. (S105), it follows that

$$\left(\text{CE}_{A,j}\right)^2 \propto \frac{\left(3-f^2\right)^2}{36}f^6 \tag{S109}$$

Thus, expanding $\chi_{A,j}^{\text{pairwise}}$ as a power series in f gives the two leading order terms (i.e., 5th and 6th order in f) as the linear combination

$$\Omega_{A,j} = K_1\oint\left(\frac{\vec{\rho}_A^{\text{avg}}\left(r_A\right)\bullet\vec{\rho}_j^{\text{avg}}\left(r_j\right)}{\vec{\rho}^{\text{avg}}\left(\vec{r}\right)\bullet\vec{\rho}^{\text{avg}}\left(\vec{r}\right)}\right)^2\rho\left(\vec{r}\right)d^3\vec{r} + K_2\left(\text{CE}_{A,j}\right)^2 \tag{S110}$$

$$\chi_{A,j}^{\text{pairwise}} = \min\left(\Omega_{A,j},\text{CE}_{A,j}\right) \tag{S111}$$

Because the second term of Eq. (S110) is quadratic in $\text{CE}_{A,j}$, its contribution is negligible when $\text{CE}_{A,j}$ is small and becomes important for large $\text{CE}_{A,j}$. Physically, this means the first term of Eq. (S110) accounts for $\chi_{A,j}^{\text{pairwise}}$ when $\text{CE}_{A,j}$ is small while the second term is a correction for large $\text{CE}_{A,j}$. Manifestly, $\text{CE}_{A,j}$ is large for a multiple order bond and smaller for single (or lower order) bonds. For a single (or lower order) bond, the number of electrons comprising the bond is normally less than or equal to the number of electrons in the exchange hole. For a multiple order bond, the number of electrons comprising the bond is more than the number of electrons in the exchange hole. Thus, slightly different $\chi_{A,j}^{\text{pairwise}}/\text{CE}_{A,j}$ ratios are required for single (and lower order) bonds compared to multiple order bonds. The second term in Eq. (S110) thus accounts for the multiple-bond-order effects. Because the second term in Eq. (S110) is quadratic in $\text{CE}_{A,j}$, it is necessary to impose Eq. (S111) to ensure $\chi_{A,j}^{\text{pairwise}}/\text{CE}_{A,j} \leq 1$, which guarantees that Eq. (S74) is always satisfied.

The specific value of K_1 is derived as follows. For single (or lower order) bonds,

$$\frac{\chi^{\mathrm{pairwise}}_{\mathrm{A,j}}}{\mathrm{CE}_{\mathrm{A,j}}} \approx \frac{\Omega_{\mathrm{A,j}}}{\mathrm{CE}_{\mathrm{A,j}}} \approx \mathrm{K}_1 \frac{\oint \left(\frac{\vec{\rho}^{\mathrm{avg}}_{\mathrm{A}}(\mathrm{r_A}) \bullet \vec{\rho}^{\mathrm{avg}}_{\mathrm{j}}(\mathrm{r_j})}{\vec{\rho}^{\mathrm{avg}}(\vec{\mathrm{r}}) \bullet \vec{\rho}^{\mathrm{avg}}(\vec{\mathrm{r}})} \right)^2 \rho(\vec{\mathrm{r}}) d^3\vec{\mathrm{r}}}{2\oint \left(\frac{\vec{\rho}^{\mathrm{avg}}_{\mathrm{A}}(\mathrm{r_A}) \bullet \vec{\rho}^{\mathrm{avg}}_{\mathrm{j}}(\mathrm{r_j})}{\vec{\rho}^{\mathrm{avg}}(\vec{\mathrm{r}}) \bullet \vec{\rho}^{\mathrm{avg}}(\vec{\mathrm{r}})} \right) \rho(\vec{\mathrm{r}}) d^3\vec{\mathrm{r}}} \tag{S112}$$

For a non-magnetic diatomic molecule (DM), define a parameter ζ as

$$0 \le \zeta = \frac{\rho^{\mathrm{avg}}_{\mathrm{A}}(\mathrm{r_A})}{\rho^{\mathrm{avg}}(\vec{\mathrm{r}})} \le 1 \tag{S113}$$

For a non-magnetic diatomic molecule, the right side of Eq. (S112) can be re-written as

$$\mathrm{K}_1 \frac{\oint \left(\frac{\vec{\rho}^{\mathrm{avg}}_{\mathrm{A}}(\mathrm{r_A}) \bullet \vec{\rho}^{\mathrm{avg}}_{\mathrm{j}}(\mathrm{r_j})}{\vec{\rho}^{\mathrm{avg}}(\vec{\mathrm{r}}) \bullet \vec{\rho}^{\mathrm{avg}}(\vec{\mathrm{r}})} \right)^2 \rho(\vec{\mathrm{r}}) d^3\vec{\mathrm{r}}}{2\oint \left(\frac{\vec{\rho}^{\mathrm{avg}}_{\mathrm{A}}(\mathrm{r_A}) \bullet \vec{\rho}^{\mathrm{avg}}_{\mathrm{j}}(\mathrm{r_j})}{\vec{\rho}^{\mathrm{avg}}(\vec{\mathrm{r}}) \bullet \vec{\rho}^{\mathrm{avg}}(\vec{\mathrm{r}})} \right) \rho(\vec{\mathrm{r}}) d^3\vec{\mathrm{r}}} = \mathrm{K}_1 \frac{\int_0^1 \left(\zeta(1-\zeta)\right)^2 \rho_{\mathrm{avg}}(\zeta) \mathrm{P}(\zeta) d\zeta}{2\int_0^1 \zeta(1-\zeta) \rho_{\mathrm{avg}}(\zeta) \mathrm{P}(\zeta) d\zeta} \tag{S114}$$

where $\rho_{\mathrm{avg}}(\zeta)$ is the average value of $\rho(\vec{\mathrm{r}})$ over positions having the same value of the parameter ζ and $\mathrm{P}(\zeta)$ is the probability distribution as a function of the parameter ζ. An analytic approximation of Eq. (S114) is needed. Because the product $\rho_{\mathrm{avg}}(\zeta)\mathrm{P}(\zeta)$ appears in both the numerator's integrand and the denominator's integrand in Eq. (S114), its effects on the numerator and denominator approximately cancel to yield:

$$\frac{\int_0^1 \left(\zeta(1-\zeta)\right)^2 \rho_{\mathrm{avg}}(\zeta) \mathrm{P}(\zeta) d\zeta}{\int_0^1 \zeta(1-\zeta) \rho_{\mathrm{avg}}(\zeta) \mathrm{P}(\zeta) d\zeta} \approx \frac{\int_0^1 \left(\zeta(1-\zeta)\right)^2 d\zeta}{\int_0^1 \zeta(1-\zeta) d\zeta} \tag{S115}$$

This allows Eq. (S112) to be re-written as

$$\frac{\chi^{\mathrm{pairwise}}_{\mathrm{A,j}}}{\mathrm{CE}^{\mathrm{DM}}_{\mathrm{A,j}}} \approx \mathrm{K}_1 \frac{\int_0^1 \left(\zeta(1-\zeta)\right)^2 d\zeta}{2\int_0^1 \zeta(1-\zeta) d\zeta} = \mathrm{K}_1 \frac{\left(\frac{1}{3} - \frac{2}{4} + \frac{1}{5}\right)}{2(1/6)} = \frac{\mathrm{K}_1}{10} \tag{S116}$$

Wheatley and Gopal[S1] empirically fit the correlation

$$\frac{\mathrm{B}_{\mathrm{A,j}}}{\oint \frac{\rho^{\mathrm{avg}}_{\mathrm{A}}(\mathrm{r_A}) \rho^{\mathrm{avg}}_{\mathrm{j}}(\mathrm{r_j})}{\left(\rho^{\mathrm{avg}}(\vec{\mathrm{r}})\right)^2} \rho(\vec{\mathrm{r}}) d^3\vec{\mathrm{r}}} \approx 3.32 \approx 3\frac{1}{3} \tag{S117}$$

for a set of diatomic and small polyatomic molecules. (Using a uniformly distributed exchange hole model, Manz and Sholl theoretically derived the ratio in Eq. (S117) should approximately equal 3.[S2] Because a uniformly distributed exchange hole model is not entirely accurate, their model underestimates the optimal value of ~10/3 by 10%.) For a non-magnetic system, the denominator of the left side of Eq. (S117) is ½$CE_{A,j}$. Therefore,

$$\frac{\chi_{A,j}^{pairwise}}{CE_{A,j}^{DM}} \approx \left(\frac{B_{A,j}^{pairwise}}{CE_{A,j}^{DM}} - 1\right) \approx \left(\frac{1}{2}\left(3\frac{1}{3}\right) - 1\right) = \frac{2}{3} \qquad \text{(S118)}$$

Comparing Eqs. (S116) and (S118) yields

$$K_1 \approx 20/3 \qquad \text{(S119)}$$

Although this value was derived for non-magnetic diatomic molecules, the polyatomic scaling relation in Eqs. (S104)–(S110) and the confluence relation in Eq. (S44) mean it will also apply to magnetic polyatomic materials when the charge and spin partitioning method is chosen to satisfy Eq. (S44).

The value of K_2 can be derived by noting it corrects multiple-order bonds. Since the number of bonding electrons exceeds the number of electrons in the exchange hole for multiple-order bonds, the $\chi_{A,j}^{pairwise}/CE_{A,j}$ ratio is higher for multiple-order bonds than for single (or lower order) bonds. The goal is to choose K_2 so the computed bond orders for molecules such as O_2 and N_2 approximately match their conventional values of 2 and 3, respectively. For O_2 and N_2, the ratio of conventional bond order to CCSD/DDEC6 contact exchange is 2/1.179=1.70 and 3/1.632=1.84, respectively. Notice the ratio is higher for the triply bonded N_2 than for the doubly bonded O_2; at this rate, it will approach the upper bound of 2 for an approximately quadruple bond. This suggests $\chi_{A,j}^{diatomic}/CE_{A,j}$ nearly saturates its upper bound value of one for quadruple and higher-order bonds. Hence,

$$K_2 \approx \frac{\overset{\frac{\chi_{A,j}^{pairwise}}{CE_{A,j}}\text{ upper bound}}{\overbrace{1}} - \overset{K_1\text{ term ratio}}{\overbrace{\frac{2}{3}}}}{\underbrace{2}_{\text{CE for quadruple bond}}} = \frac{1}{6} \qquad \text{(S120)}$$

Next, we turn our attention to $\chi_{A,j}^{coord\_num}$ that accounts for the coordination number effects. First, we consider why coordination number affects the bond-order-to-contact-exchange ratio. The bonding electrons travel circuitous paths around and between the atomic nuclei. When the coordination number is one, the bonding electrons spend a significant fraction of their time in the region of space between the two nuclei, but they also spend a significant fraction of their time near the surface of one atom or the other (i.e., near the material's van der Waals surface). When the coordination number is large (e.g., 12), then the atoms are buried and the bonding electrons always spend their time in the spaces between various atomic nuclei. Because the spaces between atomic nuclei is where atoms predominantly overlap, a unit of bonding electron density between atomic nuclei contributes more to the overlap population and contact exchange than a unit of bonding electron density near the material's van der Waals surface. Thus, for a given bond order the corresponding contact exchange will be higher when the coordination number is higher (e.g., the atom is buried) than when the coordination number is lower (e.g., the atom has lots of area on the material's van der Waals surface). Consequently, the bond-order-to-contact-exchange ratio is lower when the coordination number is higher. The term $\chi_{A,j}^{coord\_num}$ must be designed to estimate this effect.

$\chi_{A,j}^{coord\_num}$ should depend on higher than first-order powers of the increase in effective coordination number. First, consider a triatomic molecule BAC in which B and C are directly bonded to A, but C is not directly bonded to B. The bonding electrons populate the AB and AC bonds. A bonding electron in the AB bond will perform circuitous paths around atoms A and B. While this electron is in the space between nuclei A and C, this bonding electron will also contribute to the overlap population and contact exchange for bond AC. Similarly, while an electron in the AC bond is performing circuitous paths around atoms A and C, it will also contribute to the overlap population and contact exchange for bond AB when it is in the space between nuclei A and B. Second, consider an atom A that is directly bonded to coord_num other atoms. As bonding electrons perform circuitous paths around atoms A and the other atoms, each bonding electron from one bond (attached to atom A) has the chance to accidently contribute to the overlap population and contact exchange of (coord_num – 1) other bonds (attached to atom A). If the probabilities for a bonding electron to accidentally contribute to each of these bonds were equal and independent of the coordination number, then this would imply a $\chi_{A,j}^{coord\_num}$ whose leading order deviation from 1 would be linear in (coord_num – 1). However, as the coordination number of atom A increases, each atom directly bonded to atom A will occupy a smaller solid angle around atom A, resulting in an increased probability for a bonding electron (attached to atom A) to accidentally contribute to the overlap population or contact exchange of a neighboring bond (attached to atom A). Consequently, the leading order deviation of $\chi_{A,j}^{coord\_num}$ from 1 decreases faster than linear in (coord_num – 1). This can be mathematically implemented by choosing a functional form for $\chi_{A,j}^{coord\_num}$ that depends on higher than first-order powers of (coord_num – 1).

This coordination number effect can be parameterized by the leading order dimensionless group that can be constructed from the contact exchange components for atom A:

$$C_A = \left(SCE_A\right)^2 \Big/ \sum_{i \neq A} \left(CE_{A,i}\right)^2 \tag{S121}$$

C_A is the contact-exchange-weighted coordination number. For a set of equal $\{CE_{A,i}\}$, C_A will return the number of such values. For example, in methane all C-H bonds are equivalent so $C_A = 4$ for the carbon atom in methane. Similarly, $C_A = 1$ for all diatomic molecules. In fcc and hcp pure element crystals, C_A will be near the ideal value of 12.

To proceed further, we need a smooth switching function (i.e., a sigmoid function) that (i) returns $\chi_{A,j}^{coord\_num} = 1$ when $C_A = C_j = 1$, (ii) returns $\chi_{A,j}^{coord\_num} \to 0$ in the limit $C_A, C_j \to \infty$, and (iii) depends on higher than first-order powers of (coord_num – 1). Note that $\left(C_A + C_j - 2\right)\big/2$ is the average of (coord_num – 1) for atoms A and j. A simple smooth sigmoid function that switches between limits (i) and (ii) while depending on higher than first-order powers of $\left(C_A + C_j - 2\right)\big/2$ is

$$\chi_{A,j}^{coord\_num} = 1 - \left(\tanh\left(\left(C_A + C_j - 2\right)\big/K_3\right)\right)^2 \tag{S122}$$

The rate of switching is controlled by the value of K_3.

The value of K_3 can be derived by comparing computed SBOs to chemically sensible SBOs for reference materials. Because a carbon atom has four valence electrons to share, the chemically expected

SBOs for C atoms in most organic materials is ~4. Analogously, one expects a SBO of ~4 for Si atoms in pure Si solid and zeolites. Since alkali metals have one electron to share, the chemically expected SBO for a Li/Na/K atom in pure Li/Na/K solid is ~1. Since Hf has four valence electrons, the SBO of pure Hf solid is expected to be $\lesssim 4$. For pure Cu and Au solids, a SBO of ~3 is chemically expected. Computational tests using Eq. (S124) showed the value $\mathrm{K}_3 = 26$ achieves a reasonable compromise. With this value, the corresponding SBOs are 3.99 (graphite), 3.69 (diamond), 3.83–4.24 (C in 1ETM biomolecule), 3.91–4.12 (C in $Fe_4C_{40}H_{52}N_4O_{12}$ single molecule magnet), 3.79 (silicon solid), 1.03 (Li solid), 0.94 (Na solid), 0.87 (K solid), 4.16 (Hf solid), 4.03–4.13 (Si in natrolite), 3.10 (Cu solid), and 3.41 (Au solid). A substantially larger K_3 will cause the SBOs in the pure Au and Hf solids to be even higher than optimal. A substantially smaller K_3 will cause the SBOs in diamond, silicon solid, and K solid to be even lower than optimal. Also, the average DDEC6 bond order minus heuristic bond order was -0.01 for the 3d and 5d pure transition metal solids (excluding Mn), which confirms $\mathrm{K}_3 = 26$ is a balanced value. In summary, this gives the following parameter values:

$$\left(\mathrm{K}_1, \mathrm{K}_2, \mathrm{K}_3\right) = \left(\frac{20}{3}, \frac{1}{6}, 26\right) \tag{S123}$$

Putting all of the above results together gives the comprehensive bond order equation:

$$\mathrm{B}_{\mathrm{A,j}} = \mathrm{CE}_{\mathrm{A,j}} + \chi_{\mathrm{A,j}}^{\mathrm{coord\_num}} \chi_{\mathrm{A,j}}^{\mathrm{pairwise}} \chi_{\mathrm{A,j}}^{\mathrm{constraint}} = \mathrm{CE}_{\mathrm{A,j}} + \Lambda_{\mathrm{A,j}} \tag{S124}$$

where $\mathrm{CE}_{\mathrm{A,j}}$ is from Eq. (S21), $\chi_{\mathrm{A,j}}^{\mathrm{coord\_num}}$ is from Eq. (S122), $\chi_{\mathrm{A,j}}^{\mathrm{pairwise}}$ is from Eq. (S111), and

$$\chi_{\mathrm{A,j}}^{\mathrm{constraint}} = \min\left(1, \frac{\mathrm{CE}_{\mathrm{A,A}}}{\sum_{\mathrm{i \neq A}} \chi_{\mathrm{A,i}}^{\mathrm{coord\_num}} \chi_{\mathrm{A,i}}^{\mathrm{pairwise}}}, \frac{\mathrm{CE}_{\mathrm{B,B}}}{\sum_{\mathrm{i \neq B}} \chi_{\mathrm{B,i}}^{\mathrm{coord\_num}} \chi_{\mathrm{B,i}}^{\mathrm{pairwise}}}\right) \tag{S125}$$

$\mathrm{CE}_{\mathrm{A,A}}$ is from Eq. (S22). As explained in Section S1 above, atom j is labeled by $\left(\mathrm{B}, \ell_1, \ell_2, \ell_3\right)$. Summing Eq. (S124) over all atoms $\mathrm{j} \neq \mathrm{A}$ yields

$$\mathrm{SBO}_{\mathrm{A}} \leq \mathrm{SCE}_{\mathrm{A}} + \mathrm{CE}_{\mathrm{A,A}} = \mathrm{SCE}_{\mathrm{A}} + \left(\mathrm{N}_{\mathrm{A}} - ½\mathrm{SCE}_{\mathrm{A}}\right) \tag{S126}$$

Rearranging Eq. (S126) gives

$$\mathrm{N}_{\mathrm{A}} - ½\mathrm{SCE}_{\mathrm{A}} \leq 2\mathrm{N}_{\mathrm{A}} - \mathrm{SBO}_{\mathrm{A}} \tag{S127}$$

which is

$$\mathrm{CE}_{\mathrm{A,A}} \leq 2\mathrm{B}_{\mathrm{A,A}} \tag{S128}$$

Since $\mathrm{B}_{\mathrm{A,j}} \geq \mathrm{CE}_{\mathrm{A,j}}$ (Eq. (S41)), it also follow that

$$\mathrm{N}_{\mathrm{A}} - ½\mathrm{SBO}_{\mathrm{A}} \leq \mathrm{N}_{\mathrm{A}} - ½\mathrm{SCE}_{\mathrm{A}} \tag{S129}$$

which is

$$\mathrm{B}_{\mathrm{A,A}} \leq \mathrm{CE}_{\mathrm{A,A}} \tag{S130}$$

Together, Eqs. (S86), (S128), and (S130) yield Eq. (S87). Therefore, the comprehensive bond order equation fulfills the proper constraints on the ratios $1 \leq \mathrm{CE}_{\mathrm{A,A}} / \mathrm{B}_{\mathrm{A,A}} \leq 2$ and $1 \leq \mathrm{B}_{\mathrm{A,j}} / \mathrm{CE}_{\mathrm{A,j}} \leq 2$. As explained above, Eq. (S124) is unique, because it has the simplest mathematical form capable of accurately describing the primary bond order effects.

S7. Identification of atom pairs to include in the bond pair matrix

At the beginning of bond order analysis, a bond pair matrix was constructed that listed all translation symmetry unique pairs of atoms that might potentially have a bond order equal to or exceeding the bond print threshold.

To construct a set of translation symmetry unique atom pairs, the first atom can be constrained to reside in the reference unit cell. The second atom is allowed to reside anywhere in the material. Thus, an atom pair (A,j) can be specified as $(\mathrm{A},\mathrm{B},\ell_1,\ell_2,\ell_3)$ where A is the number of the first atom (in the reference unit cell), B is the number of the second atom (in the reference unit cell), and (ℓ_1,ℓ_2,ℓ_3) are the translation whole numbers of atom B that yield atom j. For example, $(\ell_1,\ell_2,\ell_3)=(0,0,0)$ means that atom j resides in the reference unit cell, while $(\ell_1,\ell_2,\ell_3)=(2,0,-3)$ means that atom j is formed by translating atom B twice the first lattice vector plus -3 times the third lattice vector. Because the pair (A,B) is equivalent to the pair (B,A), only atom pairs with $\mathrm{B}\geq\mathrm{A}$ are listed in the bond pair matrix. Because the pairs $(\mathrm{A},\mathrm{A},\ell_1,\ell_2,\ell_3)$, $(\mathrm{A},\mathrm{A},0,\ell_2,\ell_3)$, and $(\mathrm{A},\mathrm{A},0,0,\ell_3)$ are translations of the pairs $(\mathrm{A},\mathrm{A},-\ell_1,-\ell_2,-\ell_3)$, $(\mathrm{A},\mathrm{A},0,-\ell_2,-\ell_3)$, and $(\mathrm{A},\mathrm{A},0,0,-\ell_3)$, respectively, the pairs $(\mathrm{A},\mathrm{A},(\ell_1<0),\ell_2,\ell_3)$, $(\mathrm{A},\mathrm{A},0,(\ell_2<0),\ell_3)$, and $(\mathrm{A},\mathrm{A},0,0,(\ell_3<0))$ are not listed in the bond pair matrix. Also, $(\mathrm{A},\mathrm{A},0,0,0)$ is not included in the bond pair matrix.

DDEC6 partitioning uses a cutoff radius to achieve efficient linearly scaling computational cost as the number of atoms in the unit cell increases.[S3] The reasons for this are now briefly restated. During DDEC6 partitioning, spatial integrations must be performed for each atom.[S3] Depending on how the grid points are constructed, the number of grid points in the unit cell will tend to increase proportional to its volume or its number of atoms. Integrating each atom over all grid points in the unit cell would thus lead to a quadratically scaling computational cost as the unit cell volume or its number of atoms increases. Because the electron density of an atom decays approximately exponentially, the electron density assigned to an atom is negligible beyond some radius value. To optimize computational efficiency, integrations over electron and spin density functions can be performed for $0\leq \mathrm{r}_\mathrm{A}\leq \mathrm{cutoff\_radius}$, where cutoff_radius is chosen such that the electron density of atom A is negligible for $\mathrm{r}_\mathrm{A}>\mathrm{cutoff\_radius}$.[S3] This produces efficient linearly scaling computational cost as the number of atoms in the unit cell increases, because the integration volume per atom is kept fixed and the number of integrations increases proportional to the number of atoms.[S3]

5 Å is a nearly optimal value for the $\mathrm{cutoff\_radius}$.[S3] Francium is the most diffuse neutral element in the periodic table. Fitting the logarithm of electron density of an isolated neutral francium atom versus radius (from 3.779 to 9.449 bohr, which corresponds to 2 to 5 Å) yielded $\ln\left(\rho_{\mathrm{Fr}}^{\mathrm{avg}}(\mathrm{r}_\mathrm{A})\right)\approx -0.8005\mathrm{r}_\mathrm{A}-2.9378$ where $\rho_{\mathrm{Fr}}^{\mathrm{avg}}(\mathrm{r}_\mathrm{A})$ is in electrons/bohr^3 and r_A is in bohr. (This electron density was computed using the PW91 exchange-correlation functional near the complete basis set limit.) Integrating $4\pi \mathrm{r}_\mathrm{A}{}^2\rho_{\mathrm{Fr}}^{\mathrm{avg}}(\mathrm{r}_\mathrm{A})$ from r_A = 9.449 bohr (corresponds to 5 Å) to infinity yields 0.05 electrons. Since the CHARGEMOL program rescales the total number of electrons for the material to be exactly

correct, this does not impact the overall net charge of the system.[S4] Instead, it only impacts the quantities assigned to individual atoms in the material. Thus, as a worst case scenario using cutoff_radius = 5 Å could change some assigned atomic properties (e.g., net atomic charge, atomic spin moment, sum of bond orders, etc.) by a few hundredths of an electron. It can further be shown that integration errors will usually be much smaller than 0.01 electrons. For buried atoms, the electron density tail is constrained to decay no slower than $\ln\left(\rho_A^{avg}\left(r_A\right)\right) \propto -1.75 r_A$. Integrating $4\pi {r_A}^2 \exp\left(-1.75 r_A\right)$ from r_A = 3.779 to 9.449 bohr gives 4.8×10^{-5} electrons. Because the integrands in the contact exchange and other terms in the comprehensive bond order equation are positive semi-definite, the cutoff_radius (which neglects part of the space over which the integrand is integrated) can only lead to an under-estimation of the bond order and never an over-estimation. Therefore, it is impossible for a truly tiny bond order (e.g., 0.001) to be accidentally computed as being substantially larger (e.g., 0.01). Consequently, the computed DDEC6 bond orders and SBOs using cutoff_radius = 5 Å will always be reasonably close to their values in the limit cutoff_radius approaches infinity. In summary, cutoff_radius = 5 Å consistently achieves excellent computational cost and precision.

A series of criteria were applied to avoid atom pairs that could not possibly generate a bond order exceeding the bond print threshold. First, only pairs of atoms whose distance between nuclei, $d = \left\| \vec{R}_j - \vec{R}_A \right\|$, is ≤ 10 Å (i.e., twice the cutoff radius) need to be considered, because atoms further apart will have zero density overlap and hence zero bond order.

If the pair $\left(A, j\right)$ passed the first test, then a second test was performed. Second, for each atom we define the number of electrons outside a radius $r_A \le \text{cutoff\_radius}$ as

$$N_A^{outside}\left(r_A\right) = \int_{r_A}^{\text{cutoff\_radius}} 4\pi\left(r'_A\right)^2 \rho_A^{avg}\left(r'_A\right) dr'_A \qquad \text{(S131)}$$

From Eq. (S21), it directly follows that

$$CE_{A,j} < 2\left(N_A^{outside}\left(r_A\right) + N_j^{outside}\left(d - r_A\right)\right) \qquad \text{(S132)}$$

Therefore, the atom pair (A, j) was not included in the bond pair matrix if

$$\min_{0 \le r_A \le d} 2\left(N_A^{outside}\left(r_A\right) + N_j^{outside}\left(d - r_A\right)\right) < \text{bond\_print\_threshold} \qquad \text{(S133)}$$

In practice, Eq. (S133) is evaluated on the series of radial shells corresponding to $0 \le r_A \le d$ to find its minimum value.

If the pair $\left(A, j\right)$ passed the second test, then a third test was performed. Third, for each atom we define

$$\rho^{avg}\_\min{}_j\left(r_j\right) = \min_{\left\|\vec{r} - \vec{R}_j\right\| = r_j} \rho^{avg}\left(\vec{r}\right) \qquad \text{(S134)}$$

For the pair of atoms (A, j), we can set up a pair-wise cylindrical coordinate system by defining a z'-axis that passes through the nuclear positions of atoms A and j, and a h' radial coordinate and θ' angular coordinate relative to this z'-axis. This affords a simplification, because the distances to atoms A and j from position $\vec{r}$ depend only on the values of coordinates h' and z' and are independent of θ'. In this coordinate system,

$$r_A = +\sqrt{h'^2 + z'^2} \tag{S135}$$

$$r_j = +\sqrt{h'^2 + (d - z')^2} \tag{S136}$$

where the nuclear position of atom A forms the origin of this coordinate system. The minimum value which $\rho^{avg}(\vec{r})$ can take for $\vec{r}$ described by pair-wise cylindrical coordinates (h', θ', z') for any possible value of θ' is no less than

$$\text{min\_}\rho^{avg}(h', z') = \max\left(\rho^{avg}\text{\_min}_A(r_A), \rho^{avg}\text{\_min}_j(r_j), \left(\rho_A^{avg}(r_A) + \rho_j^{avg}(r_j)\right)\right) \tag{S137}$$

Thus, it directly follows that

$$4\oint \frac{\rho_A^{avg}(r_A)\rho_j^{avg}(r_j)}{\rho^{avg}(\vec{r})} d^3\vec{r} \le 8\pi \int_{d-\text{cutoff\_radius}}^{\text{cutoff\_radius}} \int_0^{\sqrt{\text{cutoff\_radius}^2 - \max\left(z'^2, (d-z')^2\right)}} h' \frac{\rho_A^{avg}(r_A)\rho_j^{avg}(r_j)}{\text{min\_}\rho^{avg}(h', z')} dh'dz' \tag{S138}$$

where the integration limits on the right side of Eq. (S138) include the set of positions $\vec{r}$ that are within cutoff_radius of atom A that are also within cutoff_radius of atom j. The inequality occurs, because the denominator of the integrand on the left side of Eq. (S138) is no less than the denominator of the integrand on the right side. In cylindrical coordinates (h', θ', z'), the differential volume element is $dV = h'dh'd\theta'dz'$ which upon integrating over $0 \le \theta' \le 2\pi$ yields $dV = 2\pi h'dh'dz'$. Hence, the factor of 4 on the left side of Eq. (S138) becomes $8\pi h'$ on the right side. In practice, the integral on the right side of Eq. (S138) was evaluated using quadrature employing discrete radial shells. It follows directly from the confluence of atomic exchange propensities (Eq. (S44)) that

$$2CE_{A,j} \approx 4\oint \frac{\rho_A^{avg}(r_A)\rho_j^{avg}(r_j)}{\rho^{avg}(\vec{r})} d^3\vec{r} \tag{S139}$$

Together with Eq. (S74), this implies (A, j) can be excluded from the bond pair matrix whenever the right side of Eq. (S138) falls below bond_print_threshold. Only translation symmetry unique atom pairs passing all three tests were included in the bond pair matrix.

S8. Identification of parallelpiped enclosing the relevant integration volume for each bond pair

For bond order integration terms, the relevant integration volume consists of those positions $\vec{r}$ that are simultaneously closer than cutoff_radius to atoms A and j. A point $\vec{r}$ closer than cutoff_radius to atom A or j but not to the other atom need not be included. For convenience, it is easiest to determine and store a parallelpiped enclosing the relevant integration volume. (As a side effect, this parallelpiped also encloses some points outside the cutoff_radius of atom A or j; these irrelevant points do not contribute to the bond order integration terms.)

The fractional coordinates of the eight corners of this parallelpiped are defined as (a) $(\lambda_1^{min}, \lambda_2^{min}, \lambda_3^{min})$, (b) $(\lambda_1^{min}, \lambda_2^{min}, \lambda_3^{max})$, (c) $(\lambda_1^{min}, \lambda_2^{max}, \lambda_3^{min})$, (d) $(\lambda_1^{min}, \lambda_2^{max}, \lambda_3^{max})$, (e) $(\lambda_1^{max}, \lambda_2^{min}, \lambda_3^{min})$, (f) $(\lambda_1^{max}, \lambda_2^{min}, \lambda_3^{max})$, (g) $(\lambda_1^{max}, \lambda_2^{max}, \lambda_3^{min})$, (h) $(\lambda_1^{max}, \lambda_2^{max}, \lambda_3^{max})$. The fractional coordinates $(\lambda_1, \lambda_2, \lambda_3)$ are related to the Cartesian displacement vector $\vec{X}$ by

$$\vec{X} = \lambda_1\vec{v}_1 + \lambda_2\vec{v}_2 + \lambda_3\vec{v}_3 \tag{S140}$$

where $\vec{v}_1$, $\vec{v}_2$, and $\vec{v}_3$ are the first, second, and third lattice vectors, respectively. First, we compute a reference parallelpiped defined by the eight corners $\left(\pm\lambda_1^{\text{delta}}, \pm\lambda_2^{\text{delta}}, \pm\lambda_3^{\text{delta}}\right)$ that completely encloses a sphere of radius cutoff_radius:

$$\lambda_1^{\text{delta}} = \frac{\text{cutoff\_radius}}{\left\|\vec{v}_1\right\| \min\left(\sqrt{1 - \frac{\left(\vec{v}_1 \bullet \vec{v}_2\right)^2}{\left\|\vec{v}_1\right\|^2 \left\|\vec{v}_2\right\|^2}}, \sqrt{1 - \frac{\left(\vec{v}_1 \bullet \vec{v}_3\right)^2}{\left\|\vec{v}_1\right\|^2 \left\|\vec{v}_3\right\|^2}} \right)} \tag{S141}$$

$$\lambda_2^{\text{delta}} = \frac{\text{cutoff\_radius}}{\left\|\vec{v}_2\right\| \min\left(\sqrt{1 - \frac{\left(\vec{v}_1 \bullet \vec{v}_2\right)^2}{\left\|\vec{v}_1\right\|^2 \left\|\vec{v}_2\right\|^2}}, \sqrt{1 - \frac{\left(\vec{v}_2 \bullet \vec{v}_3\right)^2}{\left\|\vec{v}_2\right\|^2 \left\|\vec{v}_3\right\|^2}} \right)} \tag{S142}$$

$$\lambda_3^{\text{delta}} = \frac{\text{cutoff\_radius}}{\left\|\vec{v}_3\right\| \min\left(\sqrt{1 - \frac{\left(\vec{v}_1 \bullet \vec{v}_3\right)^2}{\left\|\vec{v}_1\right\|^2 \left\|\vec{v}_3\right\|^2}}, \sqrt{1 - \frac{\left(\vec{v}_2 \bullet \vec{v}_3\right)^2}{\left\|\vec{v}_2\right\|^2 \left\|\vec{v}_3\right\|^2}} \right)} \tag{S143}$$

The min function in the denominator of Eqs. (S141)–(S143) accounts for the general situation in which some or all of the lattice vectors may not be perpendicular, and also works when they are perpendicular. Let the fractional coordinates of atoms A and j be denoted by $\left(\lambda_1^{\text{A}}, \lambda_2^{\text{A}}, \lambda_3^{\text{A}}\right)$ and

$$\left(\lambda_1^{\text{j}}, \lambda_2^{\text{j}}, \lambda_3^{\text{j}}\right) = \left(\lambda_1^{\text{B}}, \lambda_2^{\text{B}}, \lambda_3^{\text{B}}\right) + \left(\ell_1, \ell_2, \ell_3\right) \tag{S144}$$

Then, the bond midpoint in fractional coordinates is

$$\left(\lambda_1^{\text{mid}}, \lambda_2^{\text{mid}}, \lambda_3^{\text{mid}}\right) = \left(\left(\lambda_1^{\text{A}}, \lambda_2^{\text{A}}, \lambda_3^{\text{A}}\right) + \left(\lambda_1^{\text{j}}, \lambda_2^{\text{j}}, \lambda_3^{\text{j}}\right)\right) \Big/ 2 \tag{S145}$$

The corners of the enclosing parallelpiped are defined by

$$\lambda_1^{\text{min}} = \max\left(\left(-\lambda_1^{\text{delta}} + \max\left(\lambda_1^{\text{A}}, \lambda_1^{\text{j}}\right)\right), \left(-\lambda_1^{\text{delta}} \sqrt{1 - \left(\frac{d}{2 * \text{cutoff\_radius}}\right)^2} + \lambda_1^{\text{mid}} \right) \right) \tag{S146}$$

$$\lambda_1^{\text{max}} = \min\left(\left(\lambda_1^{\text{delta}} + \min\left(\lambda_1^{\text{A}}, \lambda_1^{\text{j}}\right)\right), \left(\lambda_1^{\text{delta}} \sqrt{1 - \left(\frac{d}{2 * \text{cutoff\_radius}}\right)^2} + \lambda_1^{\text{mid}} \right) \right) \tag{S147}$$

$$\lambda_2^{\text{min}} = \max\left(\left(-\lambda_2^{\text{delta}} + \max\left(\lambda_2^{\text{A}}, \lambda_2^{\text{j}}\right)\right), \left(-\lambda_2^{\text{delta}} \sqrt{1 - \left(\frac{d}{2 * \text{cutoff\_radius}}\right)^2} + \lambda_2^{\text{mid}} \right) \right) \tag{S148}$$

$$\lambda_2^{\text{max}} = \min\left(\left(\lambda_2^{\text{delta}} + \min\left(\lambda_2^{\text{A}}, \lambda_2^{\text{j}}\right)\right), \left(\lambda_2^{\text{delta}} \sqrt{1 - \left(\frac{d}{2 * \text{cutoff\_radius}}\right)^2} + \lambda_2^{\text{mid}} \right) \right) \tag{S149}$$

$$\lambda_3^{\text{min}} = \max\left(\left(-\lambda_3^{\text{delta}} + \max\left(\lambda_3^{\text{A}}, \lambda_3^{\text{j}}\right)\right), \left(-\lambda_3^{\text{delta}} \sqrt{1 - \left(\frac{d}{2 * \text{cutoff\_radius}}\right)^2} + \lambda_3^{\text{mid}} \right) \right) \tag{S150}$$

$$\lambda_3^{\max} = \min\left(\left(\lambda_3^{\text{delta}} + \min\left(\lambda_3^{A}, \lambda_3^{j}\right)\right), \left(\lambda_3^{\text{delta}}\sqrt{1-\left(\frac{d}{2*\text{cutoff\_radius}}\right)^2} + \lambda_3^{\text{mid}}\right)\right) \qquad \text{(S151)}$$

In practice, the above equations were implemented using an integration grid with integration points uniformly spaced along the three lattice vector directions. The enclosing parallepiped was chosen such that it always enclosed all of the relevant integration volume (i.e., rounding of the parallelpiped corners to a whole grid point was always performed to increase rather than decrease the parallelpiped volume). If another type of integration grid were to be chosen, it might be convenient to define the enclosing integration volume differently, but still in such a way that it encloses all positions $\vec{r}$ simultaneously closer than cutoff_radius to atoms A and j (i.e., $\left\{r_A \le \text{cutoff\_radius}\right\} \cap \left\{r_j \le \text{cutoff\_radius}\right\}$).

S9. Quantification of 3 factors leading to computed bond order less than 0.5 for optimized $[H_2]^+$

The optimized H_2 molecule has a formal bond order of exactly one. If the electron density of the optimized $[H_2]^+$ molecule were simply half that of the optimized H_2 molecule, then its formal bond order would equal 0.5 exactly. However, the electron density of the optimized $[H_2]^+$ molecule is not half that of the optimized H_2 molecule. There are three effects contributing to the computed DDEC6 bond order of less than 0.5 for the optimized $[H_2]^+$ molecule:

- In the $[H_2]^+$ molecule, each H atom carries a NAC = 0.5. Electrostatic repulsion between these partially charged atoms causes the $[H_2]^+$ molecule to have a longer optimized bond length (1.06 Å) compared to the optimized bond length (0.74 Å) for neutral H_2 in which each H atom has NAC = 0. The contact exchange and overlap population of $[H_2]^+$ decrease with longer bond length. Specifically, the contact exchange decreases from 0.246 (0.74 Å) to 0.189 (1.06 Å), and the overlap population decreases from 0.247 (0.74 Å) to 0.192 (1.06 Å). This contact exchange lowering due to bond elongation is theoretically predicted to reduce the computed $[H_2]^+$ bond order by $\left(\frac{5(0.246-0.189)}{3}+\frac{0.246^2-0.189^2}{6}\right)\Big/\left(\frac{5(0.246)}{3}+\frac{0.246^2}{6}\right)=23.7\%$. The actual reduction in DDEC6 bond order for $[H_2]^+$ is 24.8 % from 0.399 (0.74 Å) to 0.300 (1.06 Å).
- Even the computed contact exchange and overlap population of $[H_2]^+$ for a bond length of 0.74 Å are less than 0.5 times those of H_2 at the same bond length. At 0.74 Å, the contact exchange (overlap population) of 0.246 (0.247) for $[H_2]^+$ is 55.9% (55.6%) lower than 0.558 (0.555) for H_2. This contact exchange lowering is theoretically predicted to reduce the $[H_2]^+$ bond order computed via the comprehensive bond order equation by $\left(\frac{5(0.558-0.246)}{3}+\frac{0.558^2-0.246^2}{6}\right)\Big/\left(\frac{5(0.558)}{3}+\frac{0.558^2}{6}\right)=57.2\%$. At 0.74 Å, the actual reduction in DDEC6 bond order is 57.5% from 0.938 for neutral H_2 to 0.399 for $[H_2]^+$.
- Even the DDEC6 bond order of 0.938 for neutral H_2 at its optimized bond length (0.74 Å) is less than one. Since this optimized bond length and its electron distribution obtained via CCSD/def2QZVPPD is nearly exact, the computed bond order's 6.2% error arises not from the quantum chemistry calculation but rather via error due to the comprehensive bond order equation itself.

The above analysis can be used to rigorously track the decrease from formal bond order of exactly one for optimized neutral H_2 molecule to the computed DDEC6 bond order of optimized $[H_2]^+$. This process corresponds to the above three effects in reverse order: step 1 (formal bond order to DDEC6 bond order for neutral H_2 at optimized bond length of 0.74 Å), step 2 (DDEC6 bond order for neutral H_2 to $[H_2]^+$ at 0.74 Å), and step 3 (DDEC6 bond order for $[H_2]^+$ at 0.74 Å to its optimized bond length of 1.06 Å). Specifically, 1*(1-0.062)*(1-0.575)*(1-0.248)=0.300, which equals the computed DDEC6 bond order for the $[H_2]^+$ molecule at its equilibrium bond length.

Using this analysis framework, we can separate those portions of the bond order for optimized $[H_2]^+$ that are due to real physical effects from those that are due to computational artifacts (i.e., accuracy limitation of the comprehensive bond order equation). Specifically, step 1 (formal bond order to DDEC6 bond order for neutral H_2 at optimized bond length of 0.74 Å) represents an accuracy limitation of comprehensive bond order equation (aka 'modeling error') and is not a real physical effect. Step 2 (DDEC6 bond order for H_2 to $[H_2]^+$ at 0.74 Å) is partly a real physical effect and partly a computational limitation. Because bond order quantifies electron dressed-exchange between two atoms, and electrons of the opposite spin do not exchange, the one-electron $[H_2]^+$ molecule should have a bond order exactly half that of the two-electron H_2 molecule if the electron distribution of $[H_2]^+$ were exactly one-half that of H_2. Thus, the portion of step 2 that is due to a real physical effect can be estimated by combining this effect with the bond order reduction due to relaxation of the $[H_2]^+$ electron distribution. This bond order reduction due to relaxation of the $[H_2]^+$ electron distribution at 0.74 Å can be estimated as $\left(\frac{5(0.279-0.246)}{3}+\frac{0.279^2-0.246^2}{6}\right)\Big/\left(\frac{5(0.279)}{3}+\frac{0.279^2}{6}\right)=12.1\%$, where 0.279 = 0.558/2 is one-half the contact exchange for H_2 at 0.74 Å. Hence, the bond order reduction in step 2 that is plausibly due to real physical effects can be estimated as (1-0.50)(1-0.121) = 56.1%. Since the actual bond order reduction in step 2 is 57.5%, this implies only 1.4% reduction in step 2 is necessarily due to computational limitations (aka 'modeling error'). Because the contact exchange reduction from 0.279 for one-half H_2 to 0.246 for $[H_2]^+$ at 0.74 Å is sensitive to the charge partitioning method (e.g., DDEC6), it cannot be determined with certainty which portion of this is a real physical relaxation of the electron cloud and which portion is due to limitations of the charge partitioning method. Therefore, it is possible as a worst case scenario that the entire difference between the 57.5% observed bond order reduction and the 50.0% reduction if the electron distribution of $[H_2]^+$ were exactly one-half that of H_2 is due to computational limitations: 57.5% - 50.0% = 7.5%. Thus, between 1.4% (lower bound) and 7.5% (upper bound) bond order reduction in step 2 is due to computational limitations (aka 'modeling error'). Finally, step 3 (bond elongation of $[H_2]^+$ from 0.74 to 1.06 Å) is a real physical effect and not any kind of error.

Accordingly, my best estimate of the 'best' bond order for the $[H_2]^+$ molecule at its equilibrium bond length is found by combining the real physical effects. This gives a lower bound of 1*(1-0.561)*(1-0.248) = 0.330 and an upper bound of 1*(1-0.500)*(1-0.248) = 0.376 on the 'best' bond order. This suggests the computed DDEC6 bond order is 0.030 to 0.076 lower than the 'best' bond order for this molecule. This is a tolerable error.

References

S1. R. J. Wheatley and A. A. Gopal, *Phys. Chem. Chem. Phys.*, 2012, **14**, 2087-2091.

S2. T. A. Manz and D. S. Sholl, in *Computational Catalysis*, eds. A. Asthagiri and M. Janik, RSC, Cambridge, UK, 2014, pp. 192-222.

S3. T. A. Manz and N. Gabaldon Limas, *RSC Adv.*, 2016, **6**, 47771-47801.

S4. T. A. Manz and N. Gabaldon Limas, Chargemol program for performing DDEC analysis, http://ddec.sourceforge.net/, accessed August 2017.