# Can Spin-Component Scaled MP2 Achieve kJ/mol Accuracy for Cohesive Energies of Molecular Crystals?

Yu Hsuan Liang, Hong-Zhou Ye,* and Timothy C. Berkelbach*

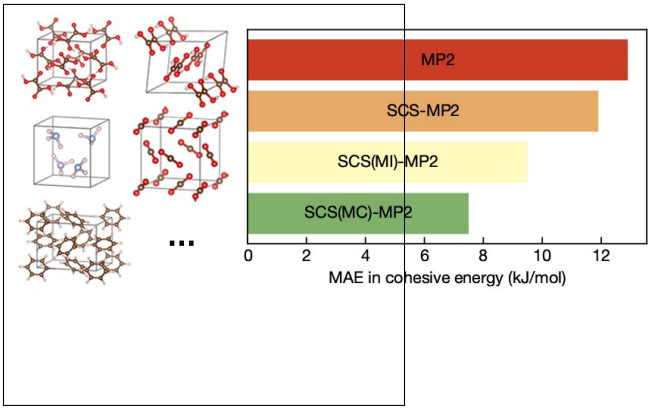
*Department of Chemistry, Columbia University, New York, NY 10027 USA*

E-mail: hzyechem@gmail.com; t.berkelbach@columbia.edu

**Abstract**

Attaining kJ/mol accuracy in cohesive energy for molecular crystals is a persistent challenge in computational materials science. In this study, we evaluate second-order Møller-Plesset perturbation theory (MP2) and its spin-component scaled models for calculating cohesive energies for 23 molecular crystals (X23 dataset). Using periodic boundary conditions and Brillouin zone sampling, we converge results to the thermodynamic and complete basis set limits, achieving an accuracy of about 2 kJ/mol (0.5 kcal/mol), which is rarely achieved in previous MP2 calculations for molecular crystals. When compared to experimental data, our results have a mean absolute error of 12.9 kJ/mol, comparable to Density Functional Theory (DFT) with the PBE functional and TS dispersion correction. By separately scaling the opposite-spin and same-spin correlation energy components, using predetermined parameters, we reduce the mean absolute error to 9.5 kJ/mol. Further fine-tuning of these scaling parameters specifically for the X23 dataset brings the mean absolute error down to 7.5 kJ/mol.

# TOC Graphic



Molecular crystals are periodic arrangements of molecules bound by weak, noncovalent interactions. They play an important role in the pharmaceutical industry, and their unique electronic and optical properties are of interest for the development of molecular optoelectronics[1]. The small energy scale of their interactions gives rise to polymorphism, where one molecule crystallizes into more than one structure, which have different properties such as their stability or solubility[1,2]. The prediction of relative stabilities of competing crystal structures is a computational grand challenge due to the chemical nature and magnitude of the competing molecular interactions, and a successful methodology would have major impact on basic and applied research.

Density functional theory (DFT)[3,4] is among the most popular methods for predicting material properties from first-principles due to its low cost and widely available implementation in popular software packages. However, DFT with the commonly employed semi-local[5] or hybrid functionals[6–8] does not accurately capture non-covalent interactions such as dispersion, which are essential for molecular crystals. Efforts to address this deficiency include the empirical '-D' corrections of Grimme[9], the exchange-hole dipole moment (XDM) method of Becke and Johnson[10], and the van der Waals $C_6$ model of Tkatchenko and Scheffler[11].

An alternative to DFT is a wave-function based approach, which naturally includes dispersion interactions at the post-Hartree-Fock (HF) level and which is, in principle, systematically improvable. Among such methods, second-order Møller-Plesset perturbation theory (MP2)[12] is the simplest correlation method and its accuracy can be significantly improved by the separate scaling of the spin components of its correlation energy (at least, for molecules)[9,13–15]. Despite the increasing use of periodic MP2 due to its relatively low cost, there are few systematic and complete reports on its performance[16–23], especially in the computationally demanding complete basis set (CBS)[24–29] limit and thermodynamic limit (TDL)[30–32].

Recently, our group reported such a study for semiconductors and insulators with strong covalent or ionic bonds[33], finding that MP2 yields reasonable predictions for lattice con-

stants, bulk moduli, and cohesive energies. For example, the mean absolute error of the cohesive energy, compared to experimental values, was 22 kJ/mol; with separate scaling of the spin components, the error was reduced to 6 kJ/mol, which is better than that of good DFT functionals like PBEsol or SCAN. In these strongly bound solids, the HF cohesive energy is qualitatively correct and constitutes about 60–70% of the MP2 cohesive energy. In contrast, in many weakly bound molecular crystals, the HF cohesive energy is qualitatively incorrect—predicting the wrong sign or only a small fraction of the true cohesive energy—and thus molecular crystals provide a more challenging test of approximate theories of electron correlation.

Recently, our group reported a periodic MP2 study of the benzene crystal[34], showing promising results and addressing potential basis set incompleteness and finite-size errors in literature MP2 values. We found that MP2 overestimated the magnitude of the cohesive energy by about 18 kJ/mol and, again, that spin scaling could reduce the error to 3–5 kJ/mol. Here, we extend this work and report tightly converged periodic MP2 calculations of the cohesive energies of the 23 molecular crystals contained in the X23 dataset of Reilly and Tkatchenko[35], which is an extension of the X16 dataset[36] and built on the C21 dataset of Otero-de-la-Roza and Johnson[37].

The crystals in the X23 dataset exhibit diverse bonding types, including dipole interactions, pure dispersion, and hydrogen bonding. For select molecular crystals, coupled-cluster theory with single, double, and perturbative triple excitations [CCSD(T)] has achieved kJ/mol accuracy in calculated cohesive energies[38,39], but its high cost—scaling as $N^7$ with system size $N$—precludes routine application to systems with large unit cells. Fifteen years ago, DiStasio and Head-Gordon targeted CCSD(T) accuracy with MP2 cost by optimizing the MP2 spin-scaling coefficients specifically for intermolecular interaction energies [SCS(MI)-MP2][15] using the S22 dataset[40] of non-covalently bonded model complexes. In the CBS limit, spin scaling reduced the mean absolute error with respect to CCSD(T) from 3.3 kJ/mol to less than 1 kJ/mol and reduced the maximum error from 12 kJ/mol to

2 kJ/mol. In this work, we test the transferability of this and other spin-scaling prescriptions for molecular crystals, whose many-body and long-range interactions may not be reflected in the dimers included in the S22 dataset.

Our calculations were performed using PySCF[41,42] with the all-electron cc-pV$X$Z ($X$=D,T,Q) basis sets[43] and periodic Gaussian density fitting of electron repulsion integrals[44–46] with corresponding JKFIT auxiliary basis sets. Core electrons were kept frozen during MP2 calculations. For a fixed $k$-point mesh containing $N_k$ points sampled uniformly from the Brillouin zone, the MP2 correlation energy per unit cell is $E^{(2)} = E^{(2)}_{\rm os} + E^{(2)}_{\rm ss}$, with the opposite-spin and same-spin components,

$$E^{(2)}_{\rm os} = -\frac{1}{N_k^3} \sum_{\boldsymbol{k}_i\boldsymbol{k}_a\boldsymbol{k}_j\boldsymbol{k}_b}{}' \sum_{iajb} T^{a\boldsymbol{k}_a,b\boldsymbol{k}_b}_{i\boldsymbol{k}_i,j\boldsymbol{k}_j} (i^{\boldsymbol{k}_i} a^{\boldsymbol{k}_a} | j^{\boldsymbol{k}_j} b^{\boldsymbol{k}_b}) \tag{1a}$$

$$E^{(2)}_{\rm ss} = -\frac{1}{N_k^3} \sum_{\boldsymbol{k}_i\boldsymbol{k}_a\boldsymbol{k}_j\boldsymbol{k}_b}{}' \sum_{iajb} [T^{a\boldsymbol{k}_a,b\boldsymbol{k}_b}_{i\boldsymbol{k}_i,j\boldsymbol{k}_j} - T^{b\boldsymbol{k}_b,a\boldsymbol{k}_a}_{i\boldsymbol{k}_i,j\boldsymbol{k}_j}] (i^{\boldsymbol{k}_i} a^{\boldsymbol{k}_a} | j^{\boldsymbol{k}_j} b^{\boldsymbol{k}_b}), \tag{1b}$$

where

$$T^{a\boldsymbol{k}_a,b\boldsymbol{k}_b}_{i\boldsymbol{k}_i,j\boldsymbol{k}_j} = \frac{(i^{\boldsymbol{k}_1} a^{\boldsymbol{k}_2} | j^{\boldsymbol{k}_3} b^{\boldsymbol{k}_4})^*}{\varepsilon^{\boldsymbol{k}_a}_a - \varepsilon^{\boldsymbol{k}_i}_i + \varepsilon^{\boldsymbol{k}_b}_b - \varepsilon^{\boldsymbol{k}_j}_j}, \tag{2}$$

$(i^{\boldsymbol{k}_1} a^{\boldsymbol{k}_2} | j^{\boldsymbol{k}_3} b^{\boldsymbol{k}_4})$ are electron repulsion integrals, and $\varepsilon^{\boldsymbol{k}_i}_i$ are HF orbital energies. The primed summation indicates conservation of crystal momentum such that $\boldsymbol{k}_i + \boldsymbol{k}_j - \boldsymbol{k}_a - \boldsymbol{k}_b = \boldsymbol{G}$, where $\boldsymbol{G}$ is a reciprocal lattice vector.

For each $N_k$, we have verified that the HF energy is converged at the QZ level and we extrapolate the MP2 correlation energy. Specifically, we perform a series of MP2 calculations with increasing $X$, and we extrapolate the correlation energy to the CBS limit using the two-point $X^{-3}$ form[8,47,48],

$$E^{(2)}(N_k, \text{CBS}) = \frac{X^3 E^{(2)}(N_k, X\text{Z}) - Y^3 E^{(2)}(N_k, Y\text{Z})}{X^3 - Y^3}, \tag{3}$$

where $X = 3, 4$ (i.e., TZ and QZ). The CBS HF and MP2 energies are then extrapolated to

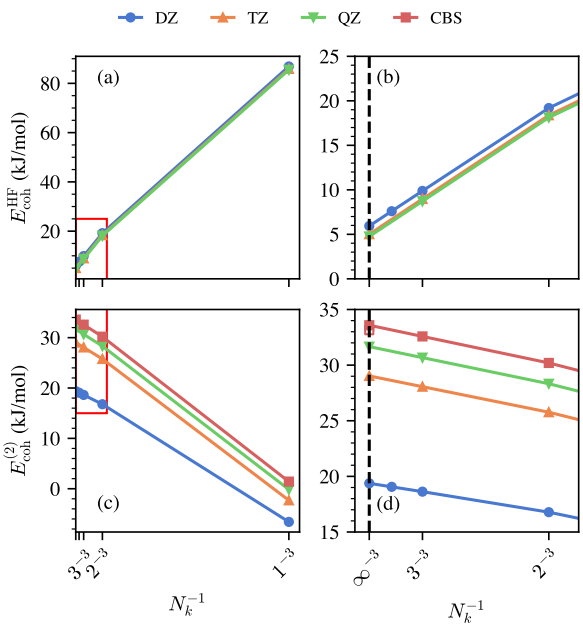



Figure 1: Thermodynamic limit convergence of energies of the ammonia crystal using different basis sets. HF cohesive energies and MP2 correlation contributions to the cohesive energies are shown in panels (a) and (c), respectively, with zoomed-in views shown in (b) and (d). Hollow square in (d) indicates the estimate obtained from the composite correction scheme described in the text, which is accurate to about 0.5 kJ/mol.

the TDL assuming finite-size errors that decay as $N_k^{-1}$,

$$E(\mathrm{TDL}, \mathrm{CBS}) = \frac{N_{k,1}E(N_{k,1}, \mathrm{CBS}) - N_{k,2}E(N_{k,2}, \mathrm{CBS})}{N_{k,1} - N_{k,2}} \tag{4}$$

which is consistent with our use of a Madelung constant correction for the integrable divergence in the HF exchange[8,47,48].

The $k$-point meshes used for extrapolation are chosen commensurate with the shape of the unit cell. For example, rhombohedral cells with lattice parameters $a = b = c$ suggest meshes like $N_k = 111, 222, 333$, while hexagonal cells with $a = b > c$ suggests meshes like $N_k = 221, 332, 443$ (we write $N_k = abc$ as shorthand for $a \times b \times c$). Optimal $k$-point mesh pairs are determined by exploration of several appropriate $k$-point mesh pairs with a cheap minimal basis set.

An example of the CBS and TDL extrapolations is shown in figure 1 for the ammonia

crystal. In this particular example, we chose $N_k = 222, 333$ as the pair of $k$-point meshes for production calculations, as it provides a small error when compared to results obtained with larger meshes but smaller basis sets; for example, in figure 1, we also present results with $N_k = 444$ in the DZ basis set, showing that extrapolation with $N_k = 222, 333$ is in good agreement with that from $N_k = 333, 444$.

Large systems in the X23 dataset required special attention because their high computational costs precludes study of appropriately large basis sets and $k$-point meshes. For example, the largest unit cell in the dataset, pyrazole, contains 80 atoms and 5872 basis functions at the QZ level. Therefore, for large systems like these, we employed a composite correction scheme to estimate the TDL/CBS limit,

$$E(\mathrm{TDL}, \mathrm{CBS}) \approx E(\mathrm{TDL}, \mathrm{TZ}) + \left[E(N_k, \mathrm{CBS}) - E(N_k, \mathrm{TZ})\right], \tag{5}$$

the accuracy of which is illustrated in figure 1 for ammonia and explored in more details for other crystals in the Supporting Information. We use periodic density fitting, and the storage requirement for 3-index Coulomb integrals is $\mathcal{O}(N_k^2 n_{\mathrm{aux}} n_{\mathrm{AO}}^2)$. For large systems demanding over 2000 GB of storage, exceeding the available disk space, we used the integral-direct algorithm implemented in PySCF[34,42]. In the Supporting Information, we provide further details on our TDL and CBS extrapolations, including the composite correction; based on testing, we conclude the uncertainty in our final numbers is around 2 kJ/mol.

Finally, we report the counterpoise corrected cohesive energy,

$$-E_{\mathrm{coh}} = \frac{E_{\mathrm{cell}}}{N_{\mathrm{mol}}} - E_{\mathrm{mol+ghost}}^{\mathrm{crystal}} + \left(E_{\mathrm{mol}}^{\mathrm{crystal}} - E_{\mathrm{mol}}^{\mathrm{gas}}\right) \tag{6}$$

where $E_{\mathrm{cell}}$ is the total crystal energy per cell, $E_{\mathrm{mol+ghost}}^{\mathrm{crystal}}$ is the energy of the molecule in its crystal geometry including basis functions from its first nearest-neighbor shell of ghost atoms, and the final term in parentheses (a molecular relaxation energy) is the energy difference between the molecule in the crystal geometry and its most stable gas phase geometry (without

**Table 1: Comparison of the cohesive energy (kJ/mol) of selected molecular crystals at the MP2 level of theory, including local MP2 (LMP2) and the hybrid QM/MM many-body interaction (HMBI) model[51]. The MBE results from Ref.[52] include the monomer relaxation energy from our own calculations as explained in the main text.**

| | ammonia | $CO_2$ | benzene |
|---|---|---|---|
| this work (PBE-TS structs.) | 38.3 | 27.8 | 76.6 |
| this work (CSD structs.) | 33.6 | 28.6 | 72.8 |
| MBE(2B)-MP2/CBS[52] | 33.4 | 28.9 | 72.3 |
| LMP2/CBS[16] | 35.6 | 29.8 | – |
| HMBI-MP2/CBS[51] | 39.3 | 29.1 | 61.6 |
| MP2/CBS[18] | 33.9 | 26.1 | 58.7 |
| MP2/CBS[53] | 35.1 | 29.4 | – |
| LMP2/p-aug-6-31G(d,p)[19] | 34.1 | 22.7 | 57.5 |

any ghost atoms). We note our sign convention is such that the cohesive energy $E_{\text{coh}}$ is the (positive) energy required to dissociate the crystal into its constituent molecules.

All geometries used are obtained from the original X23 paper[35], where all geometries (including cell parameters for crystals) were optimized with DFT using the PBE functional and the TS dispersion correction (PBE-TS)[11]. Because these geometries are the ones presented in the original X23 paper, we believe they are best situated to serve as canonical reference geometries for benchmarking and comparative studies. Moreover, previous reviews on molecular crystals have shown that PBE-TS yields the best geometries across different DFT optimization protocols[1,49], although another work has shown that energy differences obtained with different DFT optimized geometries are minimal[21]. Importantly, our own testing (see below) indicates that the use of geometries taken directly from X-ray diffraction can change the cohesive energy by 5 kJ/mol or more. Therefore, we support the use of geometries that are optimized by DFT using any reasonable dispersion-corrected functional, although further testing would be valuable. We compare our calculated cohesive energies to the revised X23 reference values[50], which were obtained by correcting experimental sublimation enthalpies for temperature-dependent vibrational contributions, including thermal expansion.

Before comparing to experiment and assessing the accuracy of spin-component scaling, we first pause to note that—in contrast to periodic, canonical MP2—the many-body expansion (MBE) has been one of the most popular formalisms for correlated calculations of molecular crystals[52–54]. Although convergence of MBE calculations is nontrivial[38,55,56], a recent paper by Sargent et al. reported carefully converged cohesive energies of most of the molecular crystals in the X23 dataset using MBE with only two-body contributions [MBE(2B)][52]. This work used structures taken directly from the Cambridge Structural Database (CSD) and did not calculate the one-body relaxation energy, so we have calculated this quantity ourselves for the same crystal geometries, to facilitate comparison.

In table 1, we present a comparison of the cohesive energies of three commonly studied molecular crystals and find that the agreement between our own periodic calculations and the MBE calculations from Ref.[52] is excellent. When we use PBE-TS structures (as we do throughout the rest of this work), the differences are 4.9, 1.1, and 4.4 kJ/mol, for ammonia, carbon dioxide, and benzene, respectively. When we repeat our calculations with the same CSD structures, the difference is reduced to 0.2, 0.3, and 0.5 kJ/mol. However, upon extending this comparison to the rest of the X23 dataset, we find that some of the CSD structures used in Ref.[52] have extremely large molecular relaxation energies, yielding abnormally bad cohesive energies. We thus suggest that unoptimized CSD geometries should be used with caution, and postpone a detailed comparison with MBE to future work. We conclude that kJ/mol agreement is possible with tightly converged MBE or periodic calculations, but that alternative geometries can cause differences of about 5 kJ/mol or more. To demonstrate the challenge of kJ/mol agreement, in the bottom half of table 1, we also show MP2 cohesive energies from a few other reports in the literature, which commonly differ by 5–15 kJ/mol.

Having shown an example of our own convergence in figure 1 and demonstrated good agreement (sub kJ/mol) with similarly converged MBE results, we now turn to a comparison with experimental cohesive energies. In figure 2, we present the error of the converged MP2 cohesive energy (with PBE-TS geometries) compared to the corrected experimental

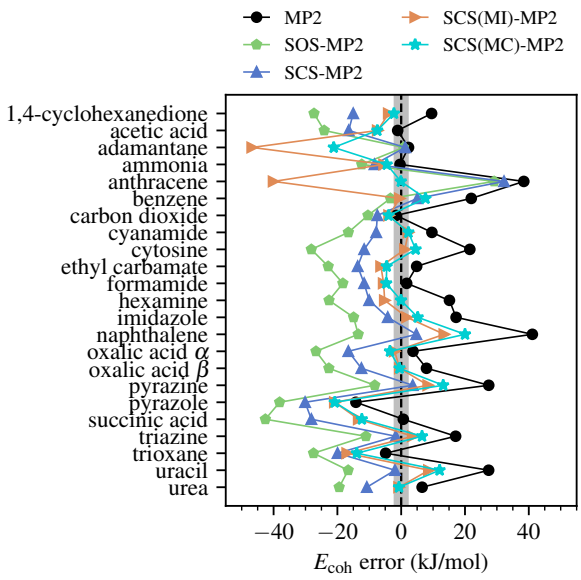



Figure 2: Error in the cohesive energy (error $= E_{\mathrm{coh}}^{\mathrm{calc}} - E_{\mathrm{coh}}^{\mathrm{exp}}$) of the molecular crystals in the X23 dataset using MP2 and four modified MP2 models: the scaled-opposite spin (SOS) model[14], the spin component-scaled (SCS) model[13], the SCS-molecular interaction [SCS(MI)] model[15], and the SCS-molecular crystal [SCS(MC)] model developed in this work. The grey shaded area of ±2 kJ/mol reflects our estimated uncertainty in our calculated numbers due to errors in extrapolations and composite corrections.

cohesive energy[50] for all 23 molecular crystals. We see that MP2 overestimates the cohesive energy, which is a well documented behavior of MP2[15,33] due to an inaccurate description of dispersion interactions and intermolecular binding energies. Indeed, the magnitudes of the error are roughly correlated with the molecular polarizabilities, as we show in figure S4. For example, the two largest MP2 errors are for naphthalene and anthracene, which also have the two largest molecular polarizabilities.

Previous benchmark studies using molecular data sets[57–59] have suggested that using smaller basis sets can improve the performance of MP2 via error cancellation. In Table S1, we provide performance statistics of MP2 in the DZ, TZ, and QZ basis set on the X23 dataset. Although the mean signed error is improved in small basis sets (as low as −4.1 kJ/mol in DZ), the mean absolute error is slightly worse. As discussed in the introduction, a more systematic route toward improving the performance of MP2 at the same cost is though spin-component scaling[13], wherein the correlation energy is calculated as $E_{\mathrm{c}}^{\mathrm{SCS}} = c_{\mathrm{os}} E_{\mathrm{os}}^{(2)} + c_{\mathrm{ss}} E_{\mathrm{ss}}^{(2)}$

and $c_{\rm os}, c_{\rm ss}$ have been previously optimized on training data. In figure 2, we show results from three different spin scaling prescriptions developed previously: the scaled-opposite spin (SOS) model[14] ($c_{\rm ss} = 0$, $c_{\rm os} = 1.3$), the spin-component-scaled (SCS) model[13] ($c_{\rm ss} = 1/3$, $c_{\rm os} = 6/5$), and the SCS-molecular interaction [SCS(MI)] model[15] ($c_{\rm ss} = 1.29$, $c_{\rm os} = 0.4$). We see that all three spin scaling prescriptions correct for the overbinding tendency of MP2. The magnitude of the corrections displays an overall trend, SOS-MP2 > SCS-MP2 > SCS(MI)-MP2, and is typically too large, resulting in predicted cohesive energies that are, on average, too small compared to experiment.

In table 2, we collect performance statistics of MP2 and the spin scaling variants, along with those of DFT results with various dispersion corrections. The errors are classifed into We see that SCS(MI)-MP2[15] delivers the best performance, with a mean absolute error (MAE) of 9.5 kJ/mol and a mean signed error (MSE) of −6.1 kJ/mol over the entire dataset. Remarkably, SCS(MI)-MP2[15] has a MAE of only 5.4 kJ/mol for the hydrogen-bonded molecular crystals, indicating that purely dispersion bound complexes are the most difficult. Indeed, with pure MP2, we find the largest errors for anthracene (49.1 kJ/mol) and naphthalene (39.0 kJ/mol), reflecting the challenge of capturing their $\pi$-electron based dispersion interactions with second-order perturbation theory, which is consistent with previous studies of molecular interactions[60,61]. For dispersion bound crystals, we find that all MP2-based methods have a MAE of about 10–20 kJ/mol.

In contrast, the improved performance for hydrogen-bonded crystals can likely be attributed to the ability of HF to capture some amount of polarization and electrostatics. Rather remarkably, DFT with the PBE functional and dispersion corrections largely outperforms MP2-based methods. Specifically, we see that PBE-TS[11] is slightly worse than MP2-based methods, whereas PBE-D3 and PBE-MBD are significantly better, exhibiting average errors of only 3-5 kJ/mol.

In figure 2, we see two prominent outliers, adamantane and anthracene, which do not follow the trend seen for the other molecular crystals. In particular, for these crystals, the cor-

rection from both SOS-MP2 and SCS-MP2 nearly vanishes, while that from SCS(MI)-MP2 is abnormally large and negative (i.e., the magnitude of the cohesive energy is significantly underestimated by more than 40 kJ/mol).

This behavior can be traced back to the degree to which a spin-component scaling prescription conserves the magnitude of the MP2 correlation energy (note that energy differences are much smaller than the total correlation energy). Specifically, we define the ratio

$$\alpha = \frac{E^{(2)}_{\mathrm{SCS}}}{E^{(2)}} = \frac{c_{\mathrm{ss}} + c_{\mathrm{os}}\gamma}{1+\gamma} \tag{7}$$

where $\gamma = E_{\mathrm{os}}/E_{\mathrm{ss}}$, and $\gamma$ is typically 3–3.5 (see Ref.[13] and figure S6). Then a theory that approximately conserves the magnitude of the MP2 correlation energy has $\alpha \approx 1$, which implies $c_{\mathrm{ss}} + 3c_{\mathrm{os}} \approx 4$; this is satisfied for the SOS and SCS models, but not for the SCS(MI) model, which has $\alpha \approx 0.6$. The SCS(MI) correlation energies are not accurate (far too small), but the method is optimized for energy differences, which can still be accurate. As detailed in section V in the SI, in the presence of spin scaling, two terms dominate the correction to the MP2 cohesive energy: one term proportional to $(\alpha - 1)E^{(2)}_{\mathrm{coh}}$ with $E^{(2)}_{\mathrm{coh}}$ the correlation part of $E_{\mathrm{coh}}$, and the other term proportional to $(c_{\mathrm{ss}} - c_{\mathrm{os}})\delta$ with $\delta = \gamma_{\mathrm{mol}} - \gamma_{\mathrm{cell}}$. Empirically, we observe $\delta \approx 0.1$ for most systems in the X23 set (figure S6). For SCS-MP2 and SOS-MP2, with $\alpha \approx 1$, the first term vanishes and the second term dominates the correction, which is negative because $c_{\mathrm{ss}} < c_{\mathrm{os}}$, and the cohesive energy is correctly reduced. By contrast, for SCS(MI)-MP2, where $\alpha \approx 0.6$, both correction terms are nonzero but with opposite sign, so they partially cancel and render the net correction smaller than the other two prescriptions. However, when $\delta \approx 0$, as is accidentally the case for the two outliers, adamantane and anthracene (figure S6), the net correction from both SCS-MP2 and SOS-MP2 vanishes, while that from SCS(MI)-MP2 is uncompensated and abnormally large, consistent with our observation from figure 2.

Approximate conservation of correlation energy has been widely regarded as a good

practice for mixing different energy components, with the most famous example being hybrid DFT[63]. The accidental vanishing of $\delta$ for the two outliers discussed above is thus unfortunate, as the analysis suggests that any spin-component scaling prescription that approximately conserves the MP2 correlation energy is doomed to show little improvement over the bare MP2 cohesive energy. When $\delta = 0$, nonzero corrections can be obtained by a theory that disregards the conservation of the MP2 correlation energy, but such a theory is at risk of uncompensated, abnormally large corrections as seen in SCS(MI)-MP2.

With all these observations in mind, we assess the extent to which further improvement is possible by a reoptimization of the spin-scaling parameters. In figure 3, we plot the cohesive energy MAE as a function of the scaling parameters and identify $c_{\mathrm{ss}} = 0.99$, $c_{\mathrm{os}} = 0.76$ to be optimal for the X23 set; we call this new prescription SCS for molecular crystals [SCS(MC)]. As shown in figure 2 for all crystals, the cohesive energies predicted by SCS(MC)-MP2 are very similar to those of SCS(MI)-MP2 except for the two outliers, where the significant underestimation by SCS(MI)-MP2 is largely corrected, resulting in an improved MAE (MSE) of 7.5 kJ/mol ($-1.3$ kJ/mol) as seen in table 2. In figure 3, we also show the region of parameters that conserve the MP2 energy (i.e., those with $\alpha \approx 1$ or with $c_{\mathrm{ss}} + 3c_{\mathrm{os}} \approx 4$). We see that the new SCS(MC) parameters do not violate this constraint nearly as much as those of SCS(MI), and empirically we find $\alpha \approx 0.8$. This compromise allows for reasonable correlation energy conservation while maintaining the freedom to correct the MP2 cohesive energy even when $\delta = 0$.

In summary, we have presented cohesive energies of 23 molecular crystals at the MP2 level of theory, with careful attention paid to basis set and finite-size errors. For our chosen geometries, we believe our results are converged to better than 2 kJ/mol. This presentation of converged MP2 cohesive energies for the entire X23 dataset is one of the most important parts of this work. While our results provide a negative answer to the question posed in the title, we believe that such a definitive answer required the extensive work carried out here. Studying only a few molecular crystals or analyzing unconverged numbers could lead one to

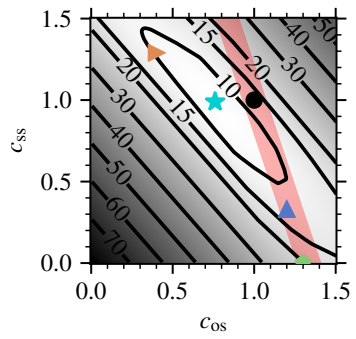



Figure 3: MAE (kJ/mol) for the spin-component scaled MP2 cohesive energy of the X23 dataset. Indicated points correspond to the scaling parameters for MP2 (black dot), SOS-MP2 (green pentagon), SCS-MP2 (blue triangle), SCS(MI)-MP2 (orange triangle), and our new SCS(MC)-MP2 (cyan star). The red-shaded area highlights the range of spin scaling parameters that approximately conserve the MP2 correlation energy.

believe the answer is "yes"; only by studying an extensive dataset like X23 with carefully converged numbers can we be sure that the answer is "no," and that alternative methods must be pursued for improved accuracy. We hope that these results will serve as useful benchmarks in future applications of wavefunction methods for molecular crystals and other solids, as well as the foundation to periodic double-hybrids, where the scope of them has been rather limited to small systems.[64–66] Importantly, we have emphasized that the cohesive energy of different but reasonable geometries can differ by 5 kJ/mol or more, which presents a challenge for precise comparisons between calculated or experimental values.

With regards to performance, we have demonstrated the well-known trend of MP2 to overbind, resulting in an overestimation of the cohesive energy by about 10–20 kJ/mol on average. Separate scaling of spin components approximately halves this error, making predictions with 5–10 kJ/mol accuracy possible. Although we proposed a new spin scaling prescription, we note that optimization against experimental values is imperfect, for reasons discussed throughout the text, and the degree to which this new model is transferable to other problems is unknown. Although we have demonstrated that kJ/mol accuracy cannot be reliably obtained within the family of spin-component scaled MP2 methods, we anticipate applications of regularized MP2[67] or double-hybrid DFT[64,66,68] as possible avenues towards

kJ/mol accuracy without the cost of coupled-cluster theory[38,39].

## Supporting Information Available

The Supporting Information is available free of charge at [publisher inserts link].

Cohesive energies for all 23 crystals at all levels of theory discussed, basis set and thermodynamic limit convergence testing, and discussion on the spin scaling correction to the cohesive energy.

## Notes

The authors declare no competing financial interest.

## Acknowledgement

We thank Caroline Sargent and Prof. David Sherrill for useful discussions and for sharing MP2/CBS data from Ref.[52]. This work was supported by the National Science Foundation under Grant Nos. CHE-1848369 and OAC-1931321. We acknowledge computing resources from the Flatiron Institute Scientific Computing Center and Columbia University's Shared Research Computing Facility project, which is supported by NIH Research Facility Improvement Grant 1G20RR030893-01, and associated funds from the New York State Empire State Development, Division of Science Technology and Innovation (NYSTAR) Contract C090171, both awarded April 15, 2010.

**Table 2: Error statistics (kJ/mol) of our MP2 results and DFT results for the X23 dataset, including mean absolute error (MAE) and mean signed error (MSE) compared to experimental values. In the final two columns, we included the classification by Cutini et al.[19], separating performance on crystals dominated by hydrogen bonding (HB) and dispersion (disp) interactions.**

| Theory | MAE | MSE | MAE (HB) | MAE (disp) |
|---|---|---|---|---|
| MP2 | 12.9 | 11.3 | 5.2 | 18.3 |
| SCS-MP2 | 11.9 | -7.9 | 14.3 | 10.5 |
| SOS-MP2 | 19.9 | -17.3 | 23.5 | 15.3 |
| SCS(MI)-MP2 | 9.5 | -6.1 | 5.4 | 15.0 |
| SCS(MC)-MP2 | 7.5 | -1.3 | 4.1 | 9.8 |
| PBE-D3[35] | 4.6 | 2.9 | 7.1 | 2.6 |
| PBE-TS[35] | 13.0 | 12.7 | 10.5 | 16.0 |
| PBE-MBD[62] | 4.5 | 3.1 | 5.0 | 3.9 |

# References


(1) Beran, G. J. O. Modeling Polymorphic Molecular Crystals with Electronic Structure Theory. *Chem. Rev.* **2016**, *116*, 5567–5613.

(2) Marom, N.; DiStasio Jr., R. A.; Atalla, V.; Levchenko, S.; Reilly, A. M.; Chelikowsky, J. R.; Leiserowitz, L.; Tkatchenko, A. Many-Body Dispersion Interactions in Molecular Crystal Polymorphism. *Angew. Chem. Int. Ed. Engl.* **2013**, *52*, 6629–6632.

(3) Kohn, W.; Sham, L. J. Self-Consistent Equations Including Exchange and Correlation Effects. *Phys. Rev.* **1965**, *140*, A1133–A1138.

(4) Hohenberg, P.; Kohn, W. Inhomogeneous Electron Gas. *Phys. Rev.* **1964**, *136*, B864–B871.

(5) Kristyán, S.; Pulay, P. Can (semi)local density functional theory account for the London dispersion forces? *Chem. Phys. Lett.* **1994**, *229*, 175–180.

(6) Kozuch, S.; Martin, J. M. L. DSD-PBEP86: in search of the best double-hybrid DFT with spin-component scaled MP2 and dispersion corrections. *Phys. Chem. Chem. Phys.* **2011**, *13*, 20104–20107.

(7) Heyd, J.; Scuseria, G. E.; Ernzerhof, M. Hybrid functionals based on a screened Coulomb potential. *J. Chem. Phys.* **2003**, *118*, 8207–8215.

(8) Broqvist, P.; Alkauskas, A.; Pasquarello, A. Hybrid-functional calculations with plane-wave basis sets: Effect of singularity correction on total energies, energy eigenvalues, and defect energy levels. *Phys. Rev. B* **2009**, *80*, 085114.

(9) Grimme, S. Accurate description of van der Waals complexes by density functional theory including empirical corrections. *J. Comput. Chem.* **2004**, *25*, 1463–1473.

(10) Becke, A. D.; Johnson, E. R. Exchange-hole dipole moment and the dispersion interaction. *J. Chem. Phys.* **2005**, *122*, 154104.

(11) Tkatchenko, A.; Scheffler, M. Accurate Molecular Van Der Waals Interactions from Ground-State Electron Density and Free-Atom Reference Data. *Phys. Rev. Lett.* **2009**, *102*, 073005.

(12) Møller, C.; Plesset, M. S. Note on an Approximation Treatment for Many-Electron Systems. *Phys. Rev.* **1934**, *46*, 618–622.

(13) Grimme, S. Improved second-order Møller–Plesset perturbation theory by separate scaling of parallel- and antiparallel-spin pair correlation energies. *J. Chem. Phys.* **2003**, *118*, 9095–9102.

(14) Jung, Y.; Lochan, R. C.; Dutoi, A. D.; Head-Gordon, M. Scaled opposite-spin second order Møller–Plesset correlation energy: An economical electronic structure method. *J. Chem. Phys.* **2004**, *121*, 9793–9802.

(15) Distasio JR., R. A.; Head-Gordon, M. Optimized spin-component scaled second-order Møller-Plesset perturbation theory for intermolecular interaction energies. *Mol. Phys.* **2007**, *105*, 1073–1083.

(16) Maschio, L.; Usvyat, D.; Schütz, M.; Civalleri, B. Periodic local Møller–Plesset second order perturbation theory method applied to molecular crystals: Study of solid NH3 and CO2 using extended basis sets. *J. Chem. Phys.* **2010**, *132*, 134706.

(17) Maschio, L.; Civalleri, B.; Ugliengo, P.; Gavezzotti, A. Intermolecular Interaction Energies in Molecular Crystals: Comparison and Agreement of Localized Møller–Plesset 2, Dispersion-Corrected Density Functional, and Classical Empirical Two-Body Calculations. *J. Phys. Chem. A* **2011**, *115*, 11179–11186.

(18) Del Ben, M.; Hutter, J.; VandeVondele, J. Second-Order Møller–Plesset Perturbation Theory in the Condensed Phase: An Efficient and Massively Parallel Gaussian and Plane Waves Approach. *J. Chem. Theory Comput.* **2012**, *8*, 4177–4188.

(19) Cutini, M.; Civalleri, B.; Corno, M.; Orlando, R.; Brandenburg, J. G.; Maschio, L.; Ugliengo, P. Assessment of Different Quantum Mechanical Methods for the Prediction of Structure and Cohesive Energy of Molecular Crystals. *J. Chem. Theory Comput.* **2016**, *12*, 3340–3352.

(20) Thomas, S. P.; Spackman, P. R.; Jayatilaka, D.; Spackman, M. A. Accurate Lattice Energies for Molecular Crystals from Experimental Crystal Structures. *J. Chem. Theory Comput.* **2018**, *14*, 1614–1623.

(21) Klimeš, J. Lattice energies of molecular solids from the random phase approximation with singles corrections. *J. Chem. Phys.* **2016**, *145*, 094506.

(22) Pham, K. N.; Modrzejewski, M.; Klimeš, J. Assessment of random phase approximation and second-order Møller–Plesset perturbation theory for many-body interactions in solid ethane, ethylene, and acetylene. *J. Chem. Phys.* **2023**, *158*, 144119.

(23) Müller, C.; Usvyat, D. Incrementally Corrected Periodic Local MP2 Calculations: I. The Cohesive Energy of Molecular Crystals. *J. Chem. Theory Comput.* **2013**, *9*, 5590–5598.

(24) Shepherd, J. J.; Grüneis, A.; Booth, G. H.; Kresse, G.; Alavi, A. Convergence of many-body wave-function expansions using a plane-wave basis: From homogeneous electron gas to solid state systems. *Phys. Rev. B* **2012**, *86*, 035111.

(25) Booth, G. H.; Tsatsoulis, T.; Chan, G. K.-L.; Grüneis, A. From plane waves to local Gaussians for the simulation of correlated periodic systems. *J. Chem. Phys.* **2016**, *145*, 084111.

(26) Callahan, J. M.; Lange, M. F.; Berkelbach, T. C. Dynamical correlation energy of metals in large basis sets from downfolding and composite approaches. *J. Chem. Phys.* **2021**, *154*, 211105.

(27) Lee, J.; Feng, X.; Cunha, L. A.; Gonthier, J. F.; Epifanovsky, E.; Head-Gordon, M. Approaching the basis set limit in Gaussian-orbital-based periodic calculations with transferability: Performance of pure density functionals for simple semiconductors. *J. Chem. Phys.* **2021**, *155*, 164102.

(28) Ye, H.-Z.; Berkelbach, T. C. Correlation-Consistent Gaussian Basis Sets for Solids Made Simple. *J. Chem. Theory Comput.* **2022**, *18*, 1595–1606.

(29) Marsman, M.; Grüneis, A.; Paier, J.; Kresse, G. Second-order Møller–Plesset perturbation theory applied to extended systems. I. Within the projector-augmented-wave formalism using a plane wave basis set. *J. Chem. Phys.* **2009**, *130*, 184103.

(30) Gruber, T.; Liao, K.; Tsatsoulis, T. Applying the Coupled-Cluster Ansatz to Solids and Surfaces in the Thermodynamic Limit. *Phys. Rev. X* **2018**, *8*, 021043.

(31) Neufeld, V. A.; Ye, H.-Z.; Berkelbach, T. C. Ground-State Properties of Metallic Solids from Ab Initio Coupled-Cluster Theory. *J. Phys. Chem. Lett.* **2022**, *13*, 7497–7503.

(32) Grüneis, A.; Marsman, M.; Kresse, G. Second-order Møller–Plesset perturbation theory applied to extended systems. II. Structural and energetic properties. *J. Chem. Phys.* **2010**, *133*, 074107.

(33) Goldzak, T.; Wang, X.; Ye, H.-Z.; Berkelbach, T. C. Accurate thermochemistry of covalent and ionic solids from spin-component-scaled MP2. *J. Chem. Phys.* **2022**, *157*, 174112.

(34) Bintrim, S. J.; Berkelbach, T. C.; Ye, H.-Z. Integral-Direct Hartree–Fock and Møller–Plesset Perturbation Theory for Periodic Systems with Density Fitting: Application to the Benzene Crystal. *J. Chem. Theory Comput.* **2022**, *18*, 5374–5381.

(35) Reilly, A. M.; Tkatchenko, A. Understanding the role of vibrations, exact exchange, and

many-body van der Waals interactions in the cohesive properties of molecular crystals. *J. Chem. Phys.* **2013**, *139*, 024705.

(36) Reilly, A. M.; Tkatchenko, A. Seamless and Accurate Modeling of Organic Molecular Materials. *J. Phys. Chem. Lett.* **2013**, *4*, 1028–1033.

(37) Otero-de-la Roza, A.; Johnson, E. R. A benchmark for non-covalent interactions in solids. *J. Chem. Phys.* **2012**, *137*, 054103.

(38) Yang, J.; Hu, W.; Usvyat, D.; Matthews, D.; Schütz, M.; Chan, G. K.-L. Ab initio determination of the crystalline benzene lattice energy to sub-kilojoule/mole accuracy. *Science* **2014**, *345*, 640–643.

(39) Borca, C. H.; Glick, Z. L.; Metcalf, D. P.; Burns, L. A.; Sherrill, C. D. Benchmark coupled-cluster lattice energy of crystalline benzene and assessment of multi-level approximations in the many-body expansion. *J. Chem. Phys.* **2023**, *158*, 234102.

(40) Jurečka, P.; Šponer, J.; Černý, J.; Hobza, P. Benchmark database of accurate (MP2 and CCSD(T) complete basis set limit) interaction energies of small model complexes, DNA base pairs, and amino acid pairs. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1985–1993.

(41) Sun, Q.; Zhang, X.; Banerjee, S.; Bao, P.; Barbry, M.; Blunt, N. S.; Bogdanov, N. A.; Booth, G. H.; Chen, J.; Cui, Z.-H. et al. Recent developments in the PySCF program package. *J. Chem. Phys.* **2020**, *153*, 024109.

(42) Sun, Q.; Berkelbach, T. C.; Blunt, N. S.; Booth, G. H.; Guo, S.; Li, Z.; Liu, J.; McClain, J. D.; Sayfutyarova, E. R.; Sharma, S. et al. PySCF: the Python-based simulations of chemistry framework. *WIREs Comput. Mol. Sci.* **2018**, *8*, e1340.

(43) Dunning, T. H., Jr. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007–1023.

(44) Sun, Q.; Berkelbach, T. C.; McClain, J. D.; Chan, G. K.-L. Gaussian and plane-wave mixed density fitting for periodic systems. *J. Chem. Phys.* **2017**, *147*, 164119.

(45) Ye, H.-Z.; Berkelbach, T. C. Fast periodic Gaussian density fitting by range separation. *J. Chem. Phys.* **2021**, *154*, 131104.

(46) Ye, H.-Z.; Berkelbach, T. C. Tight distance-dependent estimators for screening two-center and three-center short-range Coulomb integrals over Gaussian basis functions. *J. Chem. Phys.* **2021**, *155*, 124106.

(47) Paier, J.; Marsman, M.; Hummer, K.; Kresse, G.; Gerber, I. C.; Ángyán, J. G. Screened hybrid density functionals applied to solids. *J. Chem. Phys.* **2006**, *124*, 154709.

(48) Sundararaman, R.; Arias, T. A. Regularization of the Coulomb singularity in exact exchange by Wigner-Seitz truncated interactions: Towards chemical accuracy in non-trivial systems. *Phys. Rev. B* **2013**, *87*, 165122.

(49) Hoja, J.; Reilly, A. M.; Tkatchenko, A. First-principles modeling of molecular crystals: structures and stabilities, temperature and pressure. *WIREs Comput. Mol. Sci.* **2017**, *7*, e1294.

(50) Dolgonos, G. A.; Hoja, J.; Boese, A. D. Revised values for the X23 benchmark set of molecular crystals. *Phys. Chem. Chem. Phys.* **2019**, *21*, 24333–24344.

(51) Wen, S.; Beran, G. J. O. Accurate Molecular Crystal Lattice Energies from a Fragment QM/MM Approach with On-the-Fly Ab Initio Force Field Parametrization. *J. Chem. Theory Comput.* **2011**, *7*, 3733–3742.

(52) Sargent, C. T.; Metcalf, D. P.; Glick, Z. L.; Borca, C. H.; Sherrill, C. D. Benchmarking two-body contributions to crystal lattice energies and a range-dependent assessment of approximate methods. *J. Chem. Phys.* **2023**, *158*, 054112.

(53) Hofierka, J.; Klimeš, J. Binding energies of molecular solids from fragment and periodic approaches. *Electron. Struct.* **2021**, *3*, 034010.

(54) Metcalf, D. P.; Smith, A.; Glick, Z. L.; Sherrill, C. D. Range-dependence of two-body intermolecular interactions and their energy components in molecular crystals. *J. Chem. Phys.* **2022**, *157*, 084503.

(55) Xie, Y.; Glick, Z. L.; Sherrill, C. D. Assessment of three-body dispersion models against coupled-cluster benchmarks for crystalline benzene, carbon dioxide, and triazine. *J. Chem. Phys.* **2023**, *158*, 094110.

(56) Kennedy, M. R.; McDonald, A. R.; DePrince, A. E.; Marshall, M. S.; Podeszwa, R.; Sherrill, C. D. Communication: Resolving the three-body contribution to the lattice energy of crystalline benzene: Benchmark results from coupled-cluster theory. *J. Chem. Phys.* **2014**, *140*, 121104.

(57) Parker, T. M.; Burns, L. A.; Parrish, R. M.; Ryno, A. G.; Sherrill, C. D. Levels of symmetry adapted perturbation theory (SAPT). I. Efficiency and performance for interaction energies. *J. Chem. Phys.* **2014**, *140*, 094106.

(58) Lao, K. U.; Herbert, J. M. Atomic Orbital Implementation of Extended Symmetry-Adapted Perturbation Theory (XSAPT) and Benchmark Calculations for Large Supramolecular Complexes. *J. Chem. Theory Comput.* **2018**, *14*, 2955–2978.

(59) Schriber, J.; Cheney, D.; Sherrill, C. D. Levels of symmetry adapted perturbation theory (SAPT). II. Convergence of interaction energy components. *ChemRxiv* **2023**,

(60) Tsuzuki, S.; Uchimaru, T.; Matsumura, K.; Mikami, M.; Tanabe, K. Effects of the higher electron correlation correction on the calculated intermolecular interaction energies of benzene and naphthalene dimers: comparison between MP2 and CCSD(T) calculations. *Chem. Phys. Lett.* **2000**, *319*, 547–554.

(61) Nguyen, B. D.; Chen, G. P.; Agee, M. M.; Burow, A. M.; Tang, M. P.; Furche, F. Divergence of Many-Body Perturbation Theory for Noncovalent Interactions of Large Molecules. *J. Chem. Theory Comput.* **2020**, *16*, 2258–2273.

(62) Mortazavi, M.; Brandenburg, J. G.; Maurer, R. J.; Tkatchenko, A. Structure and Stability of Molecular Crystals with Many-Body Dispersion-Inclusive Density Functional Tight Binding. *J. Phys. Chem. Lett.* **2018**, *9*, 399–405.

(63) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. Ab Initio Calculation of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional Force Fields. *J. Chem. Phys.* **1994**, *98*, 11623–11627.

(64) Sharkas, K.; Toulouse, J.; Maschio, L.; Civalleri, B. Double-hybrid density-functional theory applied to molecular crystals. *J. Chem. Phys.* **2014**, *141*, 044105.

(65) Sansone, G.; Civalleri, B.; Usvyat, D.; Toulouse, J.; Sharkas, K.; Maschio, L. Range-separated double-hybrid density-functional theory applied to periodic systems. *J. Chem. Phys.* **2015**, *143*, 102811.

(66) Stein, F.; Hutter, J.; Rybkin, V. V. Double-Hybrid DFT Functionals for the Condensed Phase: Gaussian and Plane Waves Implementation and Evaluation. *Molecules (Basel, Switzerland)* **2020**, *25*, 5174.

(67) Shee, J.; Loipersberger, M.; Rettig, A.; Lee, J.; Head-Gordon, M. Regularized Second-Order Møller–Plesset Theory: A More Accurate Alternative to Conventional MP2 for Noncovalent Interactions and Transition Metal Thermochemistry for the Same Computational Cost. *J. Phys. Chem. Lett.* **2021**, *12*, 12084–12097.

(68) Wang, Y.; Li, Y.; Chen, J.; Zhang, I. Y.; Xu, X. Doubly Hybrid Functionals Close to Chemical Accuracy for Both Finite and Extended Systems: Implementation and Test of XYG3 and XYGJ-OS. *JACS Au* **2021**, *1*, 543–549.