

# MAP: an MP2 accuracy predictor for weak interactions from adiabatic connection theory

Stefan Vuckovic,<sup>1</sup> Eduardo Fabiano,<sup>2</sup> Paola Gori-Giorgi,<sup>3</sup> and Kieron Burke<sup>1</sup>

<sup>1</sup>*Department of Chemistry, University of California, Irvine, CA 92697, USA<sup>a</sup>*

<sup>2</sup>*Institute for Microelectronics and Microsystems (CNR-IMM), Via Monteroni, Campus Unisalento, 73100 Lecce, Italy*

<sup>3</sup>*Department of Theoretical Chemistry and Amsterdam Center for Multiscale Modeling, FEW, Vrije Universiteit, De Boelelaan 1083, 1081HV Amsterdam, The Netherlands*

Second order Møller-Plesset perturbation theory (MP2) approximates the exact Hartree-Fock (HF) adiabatic connection (AC) curve by a straight line. Thus by using the deviation of the exact curve from the linear behaviour, we construct an indicator for the accuracy of MP2. We then use an interpolation along the HF AC to transform the exact form of our indicator into a highly practical MP2 accuracy predictor (MAP) that comes at negligible additional computational cost. We show that this indicator is already applicable to systems that dissociate into fragments with a non-degenerate ground state, and we illustrate its usefulness by applying it to the S22 and S66 datasets.

## I. INTRODUCTION

The adiabatic connection (AC) formalism connects a single particle picture to the fully interacting system in different electronic structure theories<sup>1-12</sup>. As such, it has played an important role in the development of both density functional theory (DFT) and wavefunction theory (WFT) methods. On the DFT side, the AC provides justification and rationalization of widely popular hybrid<sup>13-15</sup> and double hybrid functionals,<sup>16-18</sup> and it has been used for the construction of other classes of density functional approximations.<sup>19-28</sup> A simple geometric construction of the AC curve has been used to obtain a lower bound to the correlation energy in DFT,<sup>24</sup> and it has been used to rationalize the amount of exact exchange in the widely used PBE0 hybrid functional.<sup>14,29</sup> On the WFT side, the Hartree-Fock (HF) AC has as weak-interaction expansion the Møller-Plesset perturbation theory<sup>30</sup>. It was also recently proposed how the AC formalism can be used to recover missing correlation energy for a broad range of multireference WFTs.<sup>12,31,32</sup>

In the present paper, we use the AC formalism to gain more insight into the performance of second-order perturbation theory and provide an indicator for its accuracy. Our construction is very simple and uses the fact that in the second-order perturbation theories (PT2; in both DFT and HF variants of the AC formalism) the AC curve is approximated by a straight line, whose slope is equal to twice the PT2 correlation energy. Thus, the two PT2 are more accurate the more linear the exact AC curve is. Following this, a remarkably simple geometric construction of the AC curves yields an indicator for the accuracy of the PT2 methods. We use an interpolation along the HF adiabatic connection formalism to transform the exact form of our indicator into a practical tool for predicting the accuracy of MP2. We show that this tool is readily applicable to systems that dissociate into

fragments with nondegenerate ground states. Applying it to the S22 and S66 datasets, we illustrate the usefulness of our indicator for predicting failures of MP2 when applied to noncovalently bonded systems.

## II. THEORY

We briefly review the basics of the AC formalism in DFT and HF theory. In either theory, we define a coupling-constant  $\lambda$  dependent Hamiltonian. In DFT, it reads as:<sup>4-6</sup>

$$\hat{H}_\lambda^{\text{DFT}} = \hat{T} + \lambda \hat{V}_{ee} + \hat{V}_\lambda^{\text{DFT}}, \quad (1)$$

where  $\hat{T}$  is the kinetic energy operator and  $\hat{V}_{ee}$  is the electron-electron repulsion operator. The  $\hat{V}_\lambda^{\text{DFT}}$  operator represents a one-body potential, which forces  $\Psi_\lambda^{\text{DFT}}$ , the ground state of eq 1, to integrate to the physical density  $\rho = \rho_1$  for all  $\lambda$  values. At  $\lambda = 1$ ,  $\hat{V}_\lambda^{\text{DFT}}$  is equal to  $\hat{V}_{\text{ext}}$  (the (nuclear) external potential). The corresponding HF AC Hamiltonian is given by (see, e.g., refs 11 and 33):

$$\hat{H}_\lambda^{\text{HF}} = \hat{T} + \hat{V}_{\text{ext}} + \lambda \hat{V}_{ee} + (1 - \lambda) (\hat{J} + \hat{K}), \quad (2)$$

where  $\hat{J} = \hat{J}[\rho^{\text{HF}}]$  and  $\hat{K} = \hat{K}[\{\phi_i^{\text{HF}}\}]$  are the standard HF Coulomb and exchange operators that depend on the HF density  $\rho^{\text{HF}}$  and occupied HF orbitals  $\phi_i^{\text{HF}}$ . They are computed once in the HF calculation for the physical system and do not depend on  $\lambda$ . A key difference between the two ACs is that the density of  $\Psi_\lambda^{\text{HF}}$  (the ground state of the Hamiltonian of eq 2) varies with  $\lambda$ , whereas the density of  $\Psi_\lambda^{\text{DFT}}$  is always forced to be that of the physical system. But at  $\lambda = 1$ ,  $\hat{H}_1^{\text{DFT}} = \hat{H}_1^{\text{HF}} = \hat{H}$ , and thus:  $\Psi_1^{\text{DFT}} = \Psi_1^{\text{HF}} = \Psi$ .

In either theory,

$$E_c = \langle \Psi | \hat{H} | \Psi \rangle - \langle \Psi_0 | \hat{H} | \Psi_0 \rangle, \quad (3)$$

and the AC formula for the correlation energy follows in

<sup>a</sup>) Electronic mail: svuckovi@uci.edu

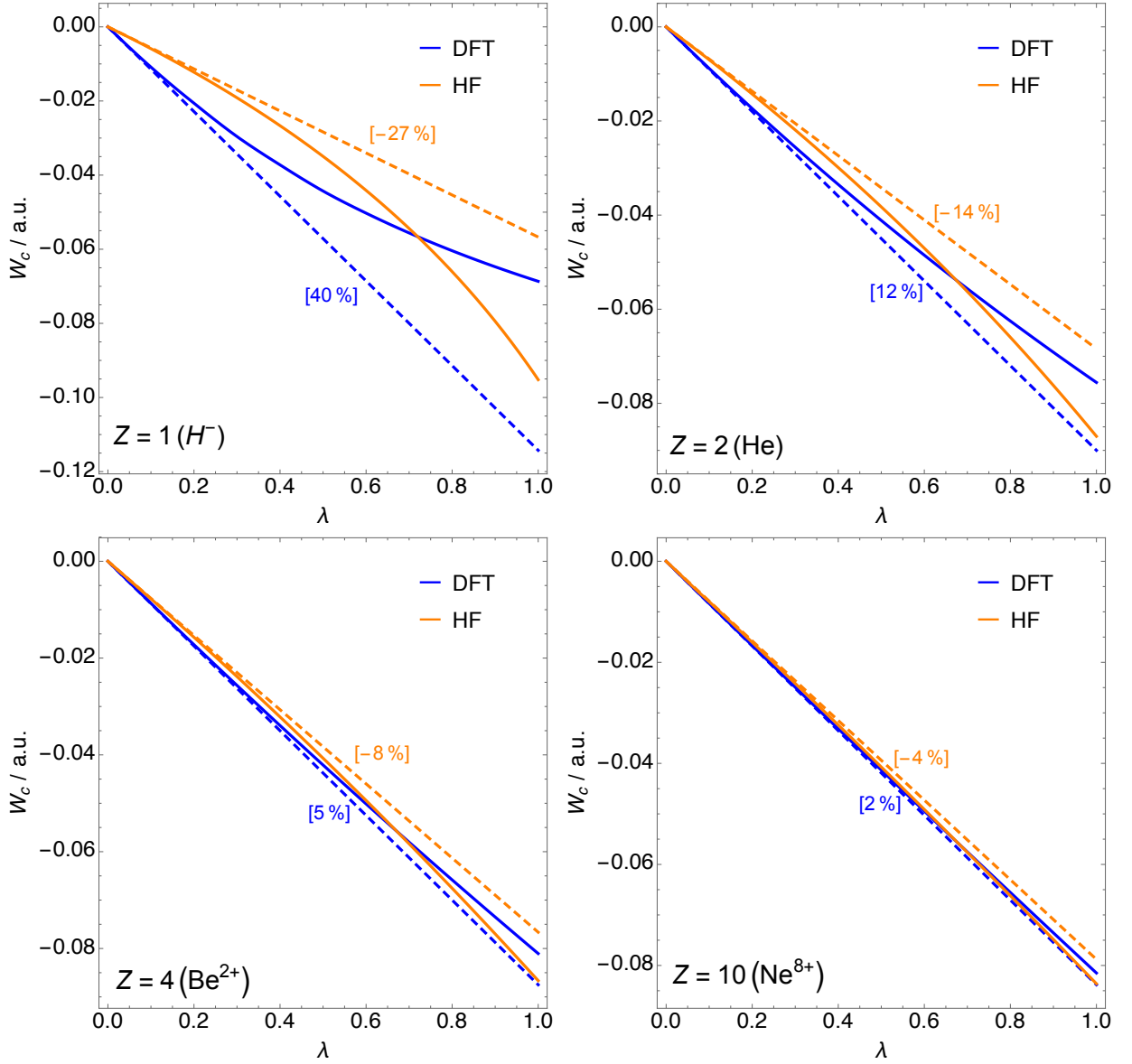


FIG. 1: DFT and HF AC curves for the selected members of the helium isoelectronic series. Dashed lines represent the AC curves from the second-order perturbation theory  $2E_c^{\text{PT}2}\lambda$ . The numbers in square brackets are the relative errors of  $E_c^{\text{PT}2}$ :  $(E_c^{\text{MP}2} - E_c^{\text{HF}})/E_c^{\text{HF}}$  in the case of HF AC, and  $(E_c^{\text{GL}2} - E_c^{\text{DFT}})/E_c^{\text{DFT}}$  in the case of DFT AC

both cases from the HellmannFeynman theorem,

$$E_c = \int_0^1 W_{c,\lambda} d\lambda. \quad (4)$$

In DFT, the underlying AC integrand  $W_{c,\lambda}$  is given by

$$W_{c,\lambda} = \langle \Psi_\lambda | \hat{V}_{\text{ee}} | \Psi_\lambda \rangle - \langle \Psi_0 | \hat{V}_{\text{ee}} | \Psi_0 \rangle \quad (\text{DFT}), \quad (5)$$

whereas its HF counterpart is

$$W_{c,\lambda} = \langle \Psi_\lambda | \hat{V}_{\text{ee}} - \hat{J} - \hat{K} | \Psi_\lambda \rangle - \langle \Psi_0 | \hat{V}_{\text{ee}} - \hat{J} - \hat{K} | \Psi_0 \rangle \quad (\text{HF}). \quad (6)$$

In DFT (eq 5),  $\Psi_0$  is the Kohn-Sham wavefunction, and in the HF AC (eq 6),  $\Psi_0$  is the HF Slater determinant,

which minimizes  $\hat{H}$ . Utilizing the expansion of  $W_{\lambda,c}^{\text{DFT}/\text{HF}}$  at small  $\lambda$  up to  $n$ -th order, we can write:

$$W_{\lambda,c}^{(n)} = \sum_{m=2}^n m E_c^{\text{PT}m} \lambda^{m-1}, \quad (7)$$

where  $E_c^{\text{PT}n}$  is the correlation energy from the  $n$ -th order perturbation theory, given by

$$E_c^{(n)} = \sum_{m=2}^n E_c^{\text{PT}m}. \quad (8)$$

Within the HF AC,  $E_c^{\text{PT}m}$  is obtained from Møller-Plesset (MP) perturbation theory (PT=MP), whereas in

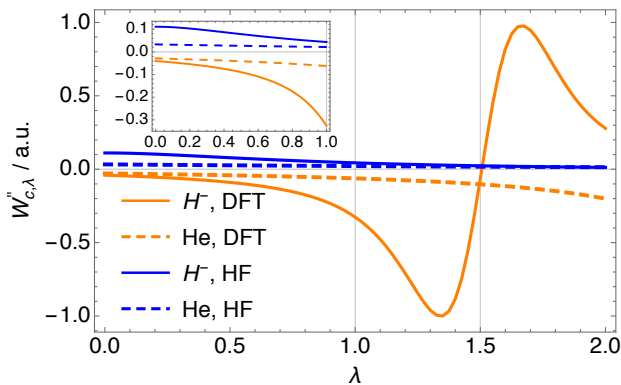


FIG. 2: The curvature of the HF and DFT AC curves,  $W''_{c,\lambda} = \partial^2 W_{c,\lambda} / \partial \lambda^2$ , for  $H^-$  and He. The inset zooms in on the region of the plots for the  $\lambda$  domain between 0 and 1. The DFT AC fitting functions used in this figure are given in supporting information

the DFT case  $E_c^{\text{PT}m}$  is obtained from Görling-Levy perturbation theory (PT=GL).<sup>34,35</sup> By truncation to second order in  $\lambda$ ,  $W_{\lambda,c}$  is approximated by a straight line:

$$W_{\lambda,c}^{\text{DFT/HF}} \approx 2E_c^{\text{PT}2} \lambda, \quad (9)$$

which sets  $E_c^{\text{DFT/HF}} \approx E_c^{\text{PT}2}$ . Both MP2 and GL2 theories are pillars of electronic structure theory, and their use is widespread in many calculations. Besides the widespread use of the MP2 method and its extensions in their standalone versions (see, e.g., ref 36 for a review), the PT2 correlation energy is also used as an ingredient for double hybrids<sup>37–39</sup>.

### III. ILLUSTRATIONS

In Figure 1, we show the AC curves in DFT and HF theories for the members of the helium isoelectronic series, namely for  $H^-$ , He,  $Be^{2+}$  and  $Ne^{8+}$ . For  $H^-$  and He, the second derivative of both  $W_{c,\lambda}^{\text{HF}}$  and  $W_{c,\lambda}^{\text{DFT}}$  (w.r.t.  $\lambda$ ) is plotted in Figure 2 for  $\lambda$  values between 0 and 2. The AC curves have been obtained from the  $\Psi_{\lambda}^{\text{HF/DFT}}$  wavefunctions at the full-CI/aug-cc-pCVTZ level.<sup>40</sup> The DFT AC curves have been taken from Refs. 41 and 42, while those of the HF AC have been obtained from the  $\Psi_{\lambda}^{\text{HF}}$  wavefunction, which we construct in the present work [the full details are given in the supporting information]. While both AC curves decrease with  $\lambda$ , that their convexity can be different is already evident from Figure 1. As it can be seen from Figure 2,  $W_{\lambda,c}^{\text{DFT}}$  is convex for both systems. In fact,  $W_{\lambda,c}^{\text{DFT}}$  is believed to be always convex (or at least piecewise convex)<sup>24</sup> and this is supported by the highly accurate numerical evidence.<sup>41–43</sup> On the other hand, we can see from Figure 2 that the convexity of  $W_{\lambda,c}^{\text{HF}}$  is not definite. For  $H^-$ ,  $W_{\lambda,c}^{\text{HF}}$  is concave up to  $\lambda \sim 1.5$  and then it becomes convex. For

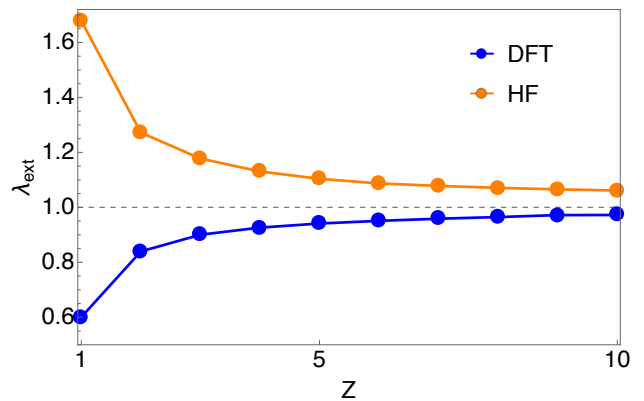


FIG. 3:  $\lambda_{\text{ext}}$  for both DFT and HF AC curves for the helium isoelectronic series as a function of nuclear charge,  $Z$

He, the convexity changes later, at  $\lambda \sim 3.4$ . In fact, although often concave at small  $\lambda$ , we know that  $W_{\lambda,c}^{\text{HF}}$  must change convexity at larger  $\lambda$ , in order to approach a finite asymptotic value<sup>33</sup>  $W_{\infty,c}^{\text{HF}}$  when  $\lambda \rightarrow \infty$ .

Staying with Figures 1, we can notice that the curvature of both DFT and HF AC curves are the strongest in the case of  $H^-$ , and then it decreases as we increase the nuclear charge,  $Z$ . Thus the relative errors in the corresponding GL2/MP2 correlation energies also decrease with  $Z$  (even though the GL2 overestimates here the magnitude of  $E_c^{\text{DFT}}$  and MP2 underestimates the magnitude of  $E_c^{\text{HF}}$  in all cases). Furthermore, the DFT and HF curves are getting closer to each other as  $Z$  increases, and for  $Ne^{8+}$  the two curves are nearly overlapping.

### IV. PRACTICAL PREDICTOR FOR THE ACCURACY OF THE MP2 THEORY WHEN APPLIED TO NONCOVALENT SYSTEMS

Utilizing that  $E_c^{\text{PT}2}$  is more accurate the more linear the exact  $W_{c,\lambda}$  is, here we use a quantity defined in ref 24 as an indicator of accuracy of the two second-order perturbation theories. This indicator is defined by:<sup>24</sup>

$$\lambda_{\text{ext}} = \frac{W_{c,1}}{W'_{c,0}}. \quad (10)$$

The  $\lambda_{\text{ext}}$  quantity is simply a value of  $\lambda$  at which the extrapolated PT line ( $W_{c,\lambda} = \lambda W'_0 = 2\lambda E_c^{\text{PT}2}$ ) reaches the  $W_{c,\lambda} = W_{c,1}$  horizontal line. As such, it represents a dimensionless measure of the curvature of ACs. For ACs convex in  $\lambda$  (within the relevant  $\lambda$  region between 0 and 1),  $\lambda_{\text{ext}}$  needs to be less than 1. For these curves, the error of PT2 vanishes as  $\lambda_{\text{ext}}$  approaches 1 (from below). Thus, when the AC curve is a straight line,  $X$  is equal to 1 and the PT2 is exact. For AC curves concave in  $\lambda$  (again within the relevant  $\lambda$  region between 0 and 1),  $\lambda_{\text{ext}}$  is greater than 1 and for these AC curves the error of PT2 also vanishes when  $\lambda_{\text{ext}}$  approaches 1 (from above).

To illustrate this, in Figure 3, we show  $\lambda_{\text{ext}}$  for the members of the helium isoelectronic series. In the case of HF ACs, the underlying  $\lambda_{\text{ext}}$  value for  $\text{H}^-$  is  $\sim 1.7$ , for He it drops to  $\sim 1.3$  and then it further decreases with  $Z$ . In the case of DFT ACs, the underlying  $\lambda_{\text{ext}}$  value for  $\text{H}^-$  is  $\sim 0.6$ , then for He it increases to  $\sim 0.8$  and then further increases with  $Z$ . In both DFT and HF case,  $\lambda_{\text{ext}}$  approaches 1 (although from different directions) as  $Z$  increases, given that both MP2 and GL2 correlation energies become exact in the  $Z$  limit of the helium isoelectronic series.<sup>35,41</sup> We can also see from Figure 3 that  $\lambda_{\text{ext}}$  pertaining to the DFT AC approaches 1 faster than its HF counterpart. This mirrors the fact the error of GL2 error decreases more quickly than that of MP2 at larger  $Z$  (Figure 1).

So far we have discussed the differences between the HF and DFT AC curves and in the remainder of this paper we focus only on HF AC aiming to provide a practical tool for predicting the accuracy of MP2. The quantity of eq 10, via  $W_{c,1}$ , requires knowledge of the fully interacting wave-function, and thus its direct use as an indicator for the accuracy of MP2 is impractical. We aim at circumventing this problem by obtaining  $W_{c,1}^{\text{HF}}$  via interpolation between the weakly and strongly interacting limits of the ACs. This idea was proposed by Seidl and co-workers in the context of the DFT AC.<sup>20,44</sup> Recent papers have also explored its use in the context of the HF AC, obtaining rather good results for interaction energies, particularly<sup>45,46</sup> (but not only<sup>47</sup>) of non-covalently bonded systems. To use this approach in the HF AC context, we employ the following SPL (after Seidl, Perdew and Levy) interpolation form<sup>44</sup>

$$W_{c,\lambda}^{\text{SPL}}(\mathbf{W}) = W_{c,\infty} \left( 1 - \left( 1 + \frac{4E_c^{\text{PT}2}\lambda}{W_{c,\infty}} \right)^{-1/2} \right), \quad (11)$$

where  $\mathbf{W} = \{W^1, \dots, W^k\}$  is the set of input ingredients from which the interpolation is built, which in this case is  $\mathbf{W} = \{W_0, W'_0, W_\infty\}$ , with  $W_0 = E_x$ ,  $W'_0 = 2E_c^{\text{PT}2}$ ,  $W_{c,\infty} = W_\infty - W_0$ . This interpolation form has been used extensively in the literature.<sup>23,24,28,48,49</sup> We should immediately remark that  $W_{c,\lambda}^{\text{SPL}}$  is always convex, and as such cannot provide a good model for the HF adiabatic connection of a given system. However, as most often in chemistry, we are interested here in interaction energies. At least for non-covalently bonded systems, interpolations like the SPL one for interaction energies work extremely well in the HF case,<sup>45,46</sup> pointing to the fact that the interaction energy HF adiabatic connection curve is probably convex, and very well modeled by the difference between two convex curves, as we are going to detail in the following.

Consider a bound system (e.g., a molecular complex)  $M$  whose individual fragments are  $F_i$ . We are interested in the interaction energy AC curve, which is given by:

$$W_{\lambda,c}^{\text{int}}(M) = W_{\lambda,c}(M) - \sum_{i=1}^N W_{\lambda,c}(F_i). \quad (12)$$

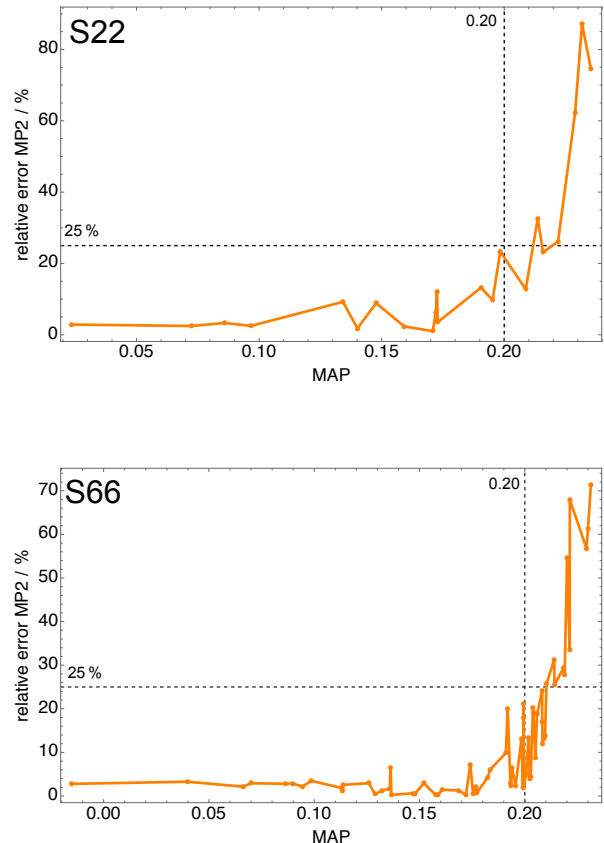


FIG. 4: The relative errors in MP2 binding energies as a function of  $\lambda_{\text{ext}}^{\text{SPL}}$  for the S22 (top panel) and S66 (lower panel) datasets

To compute  $W_{\lambda}^{\text{SPL,int}}(M)$ , we generalize the size-consistency correction of ref 46 to define:

$$W_{c,\lambda}^{\text{SPL,int}}(M) = W_{c,\lambda}^{\text{SPL}}(\mathbf{W}(M)) - W_{c,\lambda}^{\text{SPL}} \left( \sum_i^N \mathbf{W}(F_i) \right), \quad (13)$$

where  $\mathbf{W}(M)$  and  $\mathbf{W}(F_i)$  are the input ingredients of the complex and of the fragments, respectively. Equation 13 works for a system  $M$  whose fragments  $F_i$  have nondegenerate ground states, as in this case it is guaranteed that  $W_{c,\lambda}^{\text{SPL,int}}(M)$  vanishes when the distance between the fragments is set to infinity. The use of eq 13 is discussed in more details in supporting information.

To complete the model, we need  $W_\infty^{\text{HF}}$ , whose exact (fully nonlocal) form has been recently revealed,<sup>33</sup> with ongoing efforts in exploring whether this form can be actually useful for building approximations to  $E_c^{\text{HF}}$ . For practical reasons here we approximate  $W_\infty^{\text{HF}}$  with the point-charge-plus-continuum (PC) semilocal model eval-

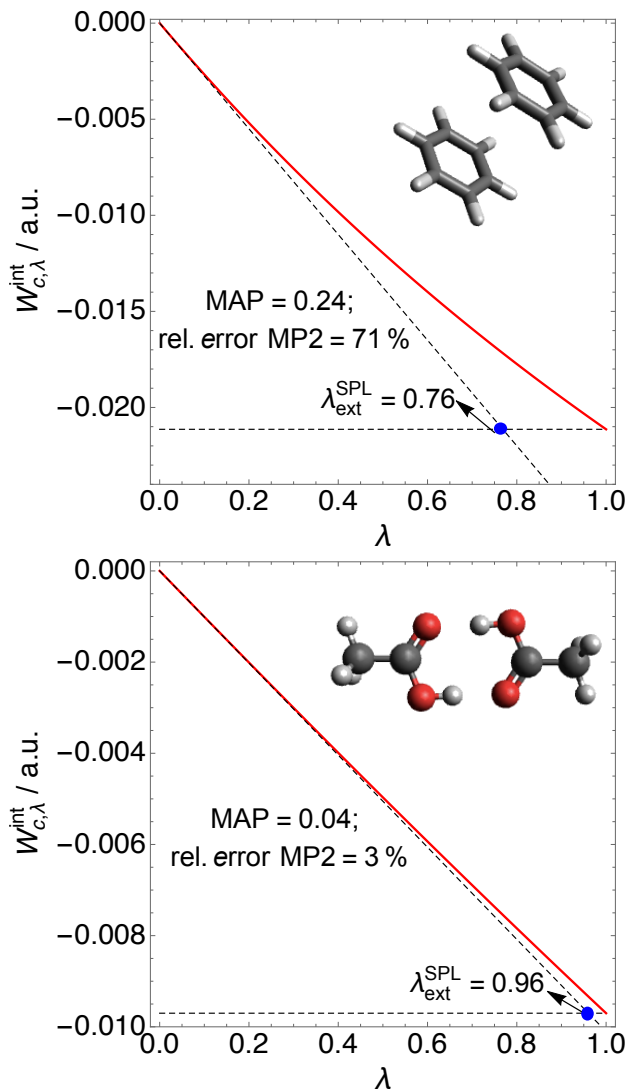


FIG. 5: The interaction AC curves obtained by the SPL interpolation via eq 11 for the benzene dimer (upper panel) and the acetic acid dimer (lower panel)

uated on  $\rho^{\text{HF}}(\mathbf{r})$ :<sup>50</sup>

$$W_{\infty}^{\text{HF}} \approx W_{\infty}^{\text{PC}}[\rho^{\text{HF}}] = \int \left[ A \rho^{\text{HF}}(\mathbf{r})^{4/3} + B \frac{|\nabla \rho^{\text{HF}}(\mathbf{r})|^2}{\rho^{\text{HF}}(\mathbf{r})^{4/3}} \right] d\mathbf{r}, \quad (14)$$

where  $A = -1.451$ ,  $B = 5.317 \times 10^{-3}$ . The correlation part of  $W_{\infty}^{\text{PC}}$  is obtained as  $W_{c,\infty}^{\text{PC}}[\rho^{\text{HF}}] = W_{\infty}^{\text{HF}}[\rho^{\text{HF}}] - E_x^{\text{HF}}$ . It has been recently shown that the combination of the PC model approximation and the SPL interpolation form of eq 13 yields rather accurate interaction energies for systems that we consider in the present work.<sup>46</sup> In Appendix A, we further discuss the use of the PC model in this context. We remark that in addition to the SPL form, other forms have been proposed in the literature (see, e.g., refs 46 and 23). However, for the systems that we consider here, the difference between the results obtained with the SPL form and other ones is very small.<sup>46</sup>

Combining eqs 10, 11 and 13 we find the  $\lambda_{\text{ext}}^{\text{SPL}}$  indicator that pertains to the interaction HF AC curve of eq 13:

$$\lambda_{\text{ext}}^{\text{SPL}} = \frac{W_{c,\lambda=1}^{\text{SPL,int}}(M)}{2E_c^{\text{MP2}}(M) - 2\sum_i^N E_c^{\text{MP2}}(F_i)}, \quad (15)$$

where  $W_{c,\lambda=1}^{\text{SPL,int}}(M)$  is given by:

$$W_{c,\lambda=1}^{\text{SPL,int}}(M) = W_{c,\infty}^{\text{PC}}(M) \left( 1 - \left( 1 + \frac{4E_c^{\text{MP2}}(M)}{W_{c,\infty}^{\text{PC}}(M)} \right)^{-1/2} \right) - \sum_i^N W_{c,\infty}^{\text{PC}}(F_i) \left( 1 - \left( 1 + \frac{4\sum_i^N E_c^{\text{MP2}}(F_i)}{\sum_i^N W_{c,\infty}^{\text{PC}}(F_i)} \right)^{-1/2} \right). \quad (16)$$

With eqs 15 and 16 we have what we need to compute  $\lambda_{\text{ext}}^{\text{SPL}}$  corresponding to the HF AC for the interaction energies of molecular complexes bonded by non-covalent interactions. The principal point of  $\lambda_{\text{ext}}^{\text{SPL}}$  is its use as an indicator for the accuracy of the MP2 theory. We define the MP2 accuracy predictor (MAP) in terms of  $\lambda_{\text{ext}}^{\text{SPL}}$  of eq 15:  $\text{MAP} = 1 - \lambda_{\text{ext}}^{\text{SPL}}$ , to make the MP2 error increase as the predictor increases. In Figure 4, we plot the relative error in the MP2 binding energies as a function of MAP for the S22 and S66 datasets.

We can observe a general trend that the MP2 errors on average decrease as MAP approaches 0 (i.e. the corresponding AC curve becomes more linear). We can also observe that when MAP is less than 0.20, the relative errors of MP2 are always below 25% (as denoted by dashed lines in the Figure 4). With even a slightly higher MAP (around 0.25), MP2 errors are skyrocketing (up to 80%) and here we encounter stacking complexes, for which MP2 failures are well-known. On the other hand, for hydrogen bonded systems MAP values approach 0, the AC interaction curves become more linear and consequently the MP2 becomes more accurate. In Figure 5, we show the benzene dimer and the acetic acid dimers AC curves obtained by eq 11, representing a situation when MP2 is accurate (the latter case) and when it is not (the former case). We also calculate MAE and MARE for the three subsets of the S66 dataset, and these are shown in Figure 6, as a function of averaged MAP pertaining to a given subset. As expected, MAP increases as we go from H-bonded complexes, over complexes classified as “others” (those bonded by a combination of dispersion and electrostatics) to complexes bonded by dispersion. This indicates that the accuracy of MP2 also increases in this order.

## V. CONCLUSIONS AND OUTLOOK

In summary, we use the AC insights to better understand and predict the accuracy of PT2 theories. We also report the highly accurate HF AC curves for the helium isoelectronic series and compare them with their DFT

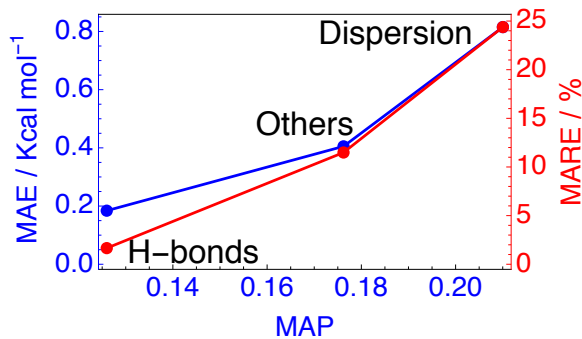


FIG. 6: MARE and MAE for the MP2 method as a function of averaged  $\lambda_{\text{ext}}^{\text{SPL}}$  for the subsets of the S66 dataset

counterparts. While the exact DFT AC curves have been studied extensively in the literature<sup>41–43,51</sup>, to the best of our knowledge the HF AC curves (Figure 1) are reported for the first time here.

We transform the exact form of our  $\lambda_{\text{ext}}$  indicator into a practical tool (MAP) for predicting the accuracy of the MP2 method for systems that dissociate into fragments with non-degenerate ground states. An important point to note about the MAP predictor is that it practically comes at no additional computational cost. Computing it by means of eqs 15 and 16 requires only (beyond the MP2 calculation itself)  $W_{\infty}^{\text{PC}}[\rho^{\text{HF}}]$ , which is easily computed from  $\rho^{\text{HF}}(\mathbf{r})$  and its gradient. This practical aspect of the MAP predictor, combined with its relevance for noncovalent interactions (NIs) and the popularity of MP2 methods for NIs, is even more useful in the light of recent findings of Furche and co-workers.<sup>52</sup> Namely, these authors have found that the performance of MP2 for NIs systematically worsens with the increase of a molecular size. Thus they advise caution when MP2 is used for calculating NIs between large molecules, given that the results can be even qualitatively wrong. This is where our MAP predictor can come into play, as it can gauge the reliability of such calculations (as shown in Figure 4.)

The MAP indicator is presently applicable to systems that dissociate into fragment with non-degenerate ground states. To address this, we will obtain exact AC curves for small covalently bonded diatomics, and then we will use these curves to get hints on how to transform the exact  $\lambda_{\text{ext}}$  into a practical indicator that also works for systems that dissociate into fragments with degenerate ground states.

In future work we will also explore the possibility of defining and analyzing the local HF AC curves, as it has been done for their DFT counterparts.<sup>24,41,53</sup> This could prove useful in using the  $\lambda_{\text{ext}}$  indicator locally (i.e. at a given point in space).

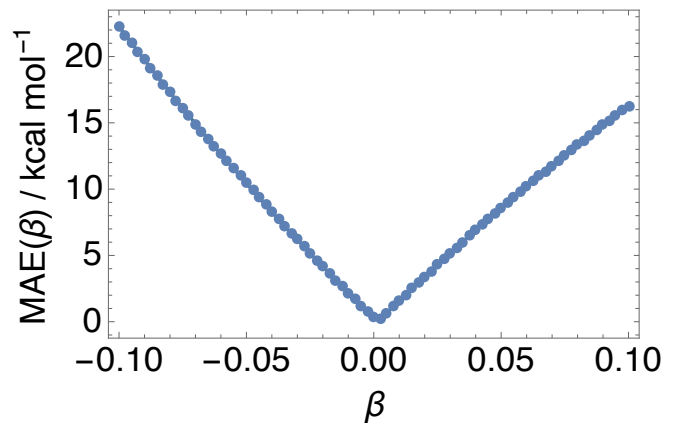


FIG. 7: MAE for the S66 dataset of the SPL interpolation scheme as a function of  $\beta$ , where we set:  $W_{\infty}^{\text{HF}}[\rho^{\text{HF}}] = W_{\infty}^{\beta}[\rho^{\text{HF}}]$  (eq A2)

## VI. ACKNOWLEDGMENTS

SV acknowledges funding from the Rubicon project (019.181EN.026), which is financed by the Netherlands Organisation for Scientific Research (NWO). KB acknowledges funding from NSF (CHE 1856165). PGG acknowledges funding from the European Research Council under H2020/ERC Consolidator Grant corr-DFT [Grant Number 648932] and from NWO under Vici grant 724.017.001.

### Appendix A: The PC model and $W_{\lambda}^{\text{HF}}$

As explained in sec IV, while  $W_{c,\lambda}^{\text{SPL,int}}(M)$  (l.h.s. of eq 13 is an accurate approximation to  $W_{\lambda,c}^{\text{int}}(M)$  (r.h.s of eq 12) for NIs,<sup>46</sup> we do not expect the SPL scheme to accurately approximate the two terms on the r.h.s. of eq 12. Comparing the size of the MP2 and CCSD(T) total energies, we expect these two terms to have a concave adiabatic connection curve. On the other hand, we expect  $W_{\lambda,c}^{\text{int}}(M)$  to be convex (given that MP2 overbinds a vast majority of S22 and S66 complexes). The SPL AC curve of eq 11 is always convex, and thus if the two terms on r.h.s. of eq 12 are concave that would be missed by the SPL interpolation. Thus, the accuracy of the  $W_{c,\lambda}^{\text{SPL,int}}(M)$  curve for NIs,<sup>46</sup> results from an error cancellation between the complex and the monomers. A similar error cancellation has been observed for the fixed-node error in Quantum Monte Carlo calculations of NIs.<sup>54</sup>

In this same light we discuss in more details the use of  $W_{\infty}^{\text{PC}}[\rho^{\text{HF}}]$  in the SPL interpolation scheme as an approximation to  $W_{\infty}^{\text{HF}}[\rho^{\text{HF}}]$ . First we note that the exact  $W_{\infty}^{\text{HF}}[\rho^{\text{HF}}]$  is expected to be much lower than  $W_{\infty}^{\text{PC}}[\rho^{\text{HF}}]$ . This is because  $W_{\infty}^{\text{PC}}[\rho]$  is energetically very close to the exact  $W_{\infty}^{\text{DFT}}[\rho]$  (see refs 8 and 48), while the following in-



equality holds<sup>33</sup>

$$W_{\infty}^{\text{HF}}[\rho^{\text{HF}}] \leq W_{\infty}^{\text{DFT}}[\rho^{\text{HF}}] + 2E_x^{\text{DFT}}[\rho^{\text{HF}}] \quad (\text{A1})$$

Thus one can even think of approximating  $W_{\infty}^{\text{HF}}$  with  $W_{\infty}^{\beta=2}$ , where

$$W_{\infty}^{\beta}[\rho^{\text{HF}}] = W_{\infty}^{\text{PC}}[\rho^{\text{HF}}] + \beta E_x^{\text{DFT}}[\rho^{\text{HF}}]. \quad (\text{A2})$$

Despite this reasoning, we show here that the use of the “bare” PC model (i.e.  $W_{\infty}^{\beta=0}[\rho^{\text{HF}}]$ ) when used in the SPL interpolation scheme gives more accurate interaction energies. We illustrate this in Figure 7, which shows the MAE for the S66 dataset of the SPL interpolation varies with  $\beta$ , when we set  $W_{\infty}^{\text{HF}}[\rho^{\text{HF}}] = W_{\infty}^{\beta}[\rho^{\text{HF}}]$ . The minimum in Figure 7 lies very close to  $\beta = 0$ . Precisely, it is at  $\beta = 0.0016$ ; (MAE = 0.27 kcal/mol) and at  $\beta = 0$  the MAE is 0.35 kcal/mol. The error rapidly increases as we go away from this minimum in either of the directions, and already at  $|\beta| > 0.01$  the error becomes huge. By exploring the exact HF AC curves in future work, we will try to better understand why the PC model works so well here (i.e. why the accuracy of the SPL interpolation is lost if a  $W_{\infty}[\rho^{\text{HF}}]$  lower or higher than  $W_{\infty}^{\text{PC}}[\rho^{\text{HF}}]$  is used). We can only make some speculative remarks at this stage. The first one is that we are constructing *de facto* an interpolation for interaction energies, and so in some way the PC model is capturing what is needed in this context. After all, we are using the difference between two convex curves to approximate a (probably) convex curve resulting from the difference between two curves that change convexity. Also, we are not using here the next leading term in the large- $\lambda$  expansion of  $W_{c,\lambda}^{\text{HF}}$ , appearing at orders  $\lambda^{-1/2}$ ,<sup>33</sup> which is positive (as in DFT),<sup>55</sup> and expected to be much larger than the DFT one, because, besides zero-point oscillations, it also contains the effect of the operator  $-\hat{K}$ .<sup>33</sup> When using only the leading term at large  $\lambda$ ,  $W_{\infty}^{\text{HF}}$  must effectively take into account also the positive (large) zero-point term. For this reason it is not so surprising that a smaller (in absolute value)  $W_{\infty}^{\text{HF}}$  gives better results than the full one. However, it remains a very intriguing fact the accuracy of the PC model in this context.

- <sup>1</sup>W. Pauli, Handbuch der Physik **24**, 162 (1933).
- <sup>2</sup>H. Hellman, Franz Deuticke, Leipzig, 285 (1937).
- <sup>3</sup>R. P. Feynman, Physical Review **56**, 340 (1939).
- <sup>4</sup>J. Harris and R. Jones, J. Phys. F **4**, 1170 (1974).
- <sup>5</sup>
- <sup>6</sup>O. Gunnarsson and B. I. Lundqvist, Phys. Rev. B **13**, 4274 (1976).
- <sup>7</sup>A. Savin, F. Colonna, and R. Pollet, Int. J. Quantum. Chem. **93**, 166 (2003).
- <sup>8</sup>S. Vuckovic, L. O. Wagner, A. Mirtschink, and P. Gori-Giorgi, J. Chem. Theory Comput. **11**, 3153 (2015).
- <sup>9</sup>Z.-F. Liu and K. Burke, Physical Review A **79**, 064503 (2009).
- <sup>10</sup>S. Vuckovic, M. Levy, and P. Gori-Giorgi, J. Chem. Phys **147**, 214107 (2017).
- <sup>11</sup>K. Pernal, Int. J. Quantum. Chem. **118**, e25462 (2018).
- <sup>12</sup>K. Pernal, Physical review letters **120**, 013001 (2018).
- <sup>13</sup>A. D. Becke, J. Chem. Phys. **98**, 5648 (1993).
- <sup>14</sup>J. P. Perdew, M. Ernzerhof, and K. Burke, J. Chem. Phys. **105**, 9982 (1996).
- <sup>15</sup>Y. Zhao, N. E. Schultz, and D. G. Truhlar, Journal of Chemical Theory and Computation **2**, 364 (2006).
- <sup>16</sup>S. Grimme, The Journal of chemical physics **124**, 034108 (2006).
- <sup>17</sup>L. Goerigk and S. Grimme, J. Chem. Theory Comput. **7**, 291 (2010).
- <sup>18</sup>K. Sharkas, J. Toulouse, and A. Savin, The Journal of chemical physics **134**, 064113 (2011).
- <sup>19</sup>M. Ernzerhof, Chem. Phys. Lett. **263**, 499 (1996).
- <sup>20</sup>M. Seidl, J. P. Perdew, and S. Kurth, Physical review letters **84**, 5070 (2000).
- <sup>21</sup>P. Mori-Sanchez, A. J. Cohen, and W. T. Yang, J. Chem. Phys. **125**, 201102 (2006).
- <sup>22</sup>A. D. Becke, J. Chem. Phys. **138**, 074109 (2013).
- <sup>23</sup>S. Vuckovic, T. J. Irons, A. Savin, A. M. Teale, and P. Gori-Giorgi, J. Chem. Theory Comput. **12**, 2598 (2016).
- <sup>24</sup>S. Vuckovic, T. J. P. Irons, L. O. Wagner, A. M. Teale, and P. Gori-Giorgi, Phys. Chem. Chem. Phys. **19**, 6169 (2017).
- <sup>25</sup>H. Bahmann, Y. Zhou, and M. Ernzerhof, J. Chem. Phys. **145**, 124104 (2016).
- <sup>26</sup>S. Vuckovic and P. Gori-Giorgi, J. Phys. Chem. Lett. **8**, 2799 (2017).
- <sup>27</sup>T. Gould and S. Vuckovic, The Journal of chemical physics **151**, 184101 (2019).
- <sup>28</sup>S. Vuckovic, Journal of chemical theory and computation **15**, 3580 (2019).
- <sup>29</sup>K. Burke, M. Ernzerhof, and J. P. Perdew, Chem. Phys. Lett. **265**, 115 (1997).
- <sup>30</sup>C. Møller and M. S. Plesset, Physical Review **46**, 618 (1934).
- <sup>31</sup>E. Pastorczak and K. Pernal, Journal of chemical theory and computation **14**, 3493 (2018).
- <sup>32</sup>K. Pernal, The Journal of chemical physics **149**, 204101 (2018).
- <sup>33</sup>M. Seidl, S. Giarrusso, S. Vuckovic, E. Fabiano, and P. Gori-Giorgi, The Journal of chemical physics **149**, 241101 (2018).
- <sup>34</sup>A. Görling and M. Levy, Physical Review B **47**, 13105 (1993).
- <sup>35</sup>A. Görling and M. Levy, Physical Review A **50**, 196 (1994).
- <sup>36</sup>D. Cremer, Wiley Interdisciplinary Reviews: Computational Molecular Science **1**, 509 (2011).
- <sup>37</sup>S. Grimme, The Journal of chemical physics **118**, 9095 (2003).
- <sup>38</sup>Y. Jung, R. C. Lochan, A. D. Dutoi, and M. Head-Gordon, The Journal of chemical physics **121**, 9793 (2004).
- <sup>39</sup>F. Neese, T. Schwabe, S. Kossmann, B. Schirmer, and S. Grimme, Journal of chemical theory and computation **5**, 3060 (2009).
- <sup>40</sup>T. H. Dunning, J. Chem. Phys. **90**, 1007 (1989).
- <sup>41</sup>S. Vuckovic, T. J. Irons, A. Savin, A. M. Teale, and P. Gori-Giorgi, J. Chem. Theory Comput. **12**, 2598 (2016).
- <sup>42</sup>A. Teale, S. Coriani, and T. Helgaker, The Journal of chemical physics **130**, 104111 (2009).
- <sup>43</sup>A. Teale, S. Coriani, and T. Helgaker, The Journal of chemical physics **132**, 164115 (2010).
- <sup>44</sup>M. Seidl, J. P. Perdew, and M. Levy, Physical Review A **59**, 51 (1999).
- <sup>45</sup>E. Fabiano, P. Gori-Giorgi, M. Seidl, and F. Della Sala, J. Chem. Theory Comput **12**, 4885 (2016).
- <sup>46</sup>S. Vuckovic, P. Gori-Giorgi, F. Della Sala, and E. Fabiano, J. Phys. Chem. Lett. (2018).
- <sup>47</sup>S. Giarrusso, P. Gori-Giorgi, F. Della Sala, and E. Fabiano, The Journal of chemical physics **148**, 134106 (2018).
- <sup>48</sup>M. Seidl, P. Gori-Giorgi, and A. Savin, Physical Review A **75**, 042511 (2007).
- <sup>49</sup>D. P. Kooi and P. Gori-Giorgi, Theoretical chemistry accounts **137**, 166 (2018).
- <sup>50</sup>M. Seidl, J. P. Perdew, and S. Kurth, Physical Review A **62**, 012502 (2000).
- <sup>51</sup>F. Colonna and A. Savin, J. Chem. Phys. **110**, 2828 (1999).
- <sup>52</sup>B. Nguyen, G. P. Chen, M. M. Agee, A. M. Burow, M. Tang, and F. Furche, **preprint ChemRxiv:10.26434/chemrxiv.11124251.v1** (2019).
- <sup>53</sup>T. J. Irons and A. M. Teale, Mol. Phys. **114**, 484 (2016).
- <sup>54</sup>M. Dubecký, L. Mitas, and P. Jurečka, Chemical Reviews **116**,

5188 (2016).  
<sup>55</sup>P. Gori-Giorgi, G. Vignale, and M. Seidl, *Journal of chemical*

*theory and computation* **5**, 743 (2009).