Benchmarking exchange-correlation functionals for hydrogen at high pressures using quantum Monte Carlo

Raymond C. Clay, III,[1,*] Jeremy Mcminis,[2] Jeffrey M. McMahon,[1] Carlo Pierleoni,[3] David M. Ceperley,[1] and Miguel A. Morales[2]

[1]*Department of Physics, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA*

[2]*Lawrence Livermore National Laboratory, Livermore, California 94550, USA*

[3]*Department of Physical and Chemical Sciences, University of L'Aquila and CNISM UdR L'Aquila, Via Vetoio, I-67010 L'Aquila, Italy*

(Received 29 January 2014; revised manuscript received 14 April 2014; published 12 May 2014)

The *ab initio* phase diagram of dense hydrogen is very sensitive to errors in the treatment of electronic correlation. Recently, it has been shown that the choice of the density functional has a large effect on the predicted location of both the liquid-liquid phase transition and the solid insulator-to-metal transition in dense hydrogen. To identify the most accurate functional for dense hydrogen applications, we systematically benchmark some of the most commonly used functionals using quantum Monte Carlo. By considering several measures of functional accuracy, we conclude that the van der Waals and hybrid functionals significantly outperform local density approximation and Perdew-Burke-Ernzerhof. We support these conclusions by analyzing the impact of functional choice on structural optimization in the molecular solid, and on the location of the liquid-liquid phase transition.

DOI: 10.1103/PhysRevB.89.184106 PACS number(s): 62.50.−p, 67.80.ff, 63.20.dk, 64.70.kt

I. INTRODUCTION

As a consequence of being the most abundant element in the universe, hydrogen can be readily observed to exist over a very large range of temperatures and pressures. Additionally, for many applications, such as inertial confinement fusion or modeling Jovian planets, the pressures and temperatures of hydrogen can vary by orders of magnitude within a given system. Understanding the behavior of such systems necessitates an accurate phase diagram over a large range of thermodynamic conditions, which by some estimates needs to be accurate to at least 1% [1].

Attaining this level of accuracy is an ongoing challenge. Despite hydrogen's simplicity, significant nuclear quantum effects (NQEs) and electronic correlation cooperate to give rise to a rich phase diagram. There are known to be four insulating molecular solid phases, an insulating molecular liquid phase, and a conducting atomic liquid phase at high temperatures [2]. Most interestingly, theory predicts the existence of a low-temperature metallic phase above 350 GPa, which could be a superconductor with an unusually high T_c or even something more exotic [3–6].

Although a rough qualitative agreement between experiment and theory has been reached for most of the hydrogen phase diagram, *ab initio* methods have struggled quantitatively in the high-pressure regime of hydrogen. This is because the relevant energy scales in high-pressure hydrogen are comparable to the errors introduced by approximating NQEs and electronic correlation. Recent simulations show that using the quasiharmonic approximation (QHA) produces noticeable quantitative differences from an exact treatment in quantities such as pair correlation functions and enthalpies of solid hydrogen structures [7]. Additionally, the approximation of electronic correlation through the use of density functionals has been shown to have a large effect on the predicted phase diagram. For example, the location of the liquid-liquid phase transition (LLPT) in hydrogen could change by as much as 160 GPa in path-integral molecular dynamics (PIMD) simulations depending on whether Perdew-Burke-Ernzerhof (PBE) or vdW-DF2 was used [8]. In the solid phase, it was found that by using either optB88-vdB [9] or vdW-DF2 [10] in PIMD, noticeably better agreement with experiment was obtained for the band gap in the $C2/c$ phase and the transition pressure between phases II and III as compared to PBE [7,11].

Recent algorithmic and computational advances allow a systematic treatment of NQEs, for example, with PIMD or path integral + generalized Langevin equation (PI + GLE) [12]. However, there have been, to date, few systematic studies to identify the most accurate functional for high-pressure hydrogen [13]. PBE has been advocated on error-cancellation grounds, but several previous studies indicate that some functionals are more accurate than the semilocal functionals. As dispersion forces are significant in solid hydrogen, it is expected that nonlocal functionals that include this effect will calculate the total and relative energies more accurately, hence, previous interest in vdW-DF2 and optB88-vdW. Additionally, hybrid functionals such as Heyd-Scuseria-Ernzerhof (HSE) are known to give better band gaps than semilocal functionals, hence the interest in these functionals to describe the insulator-to-metal transitions in high-pressure hydrogen [8]. Because of technological limitations, there are little high-quality experimental data available for pressures higher than 200 GPa, so identifying accurate functionals must be done using theoretical methods.

The purpose of this work is to systematically benchmark several types of semilocal, nonlocal van der Waals (vdW), and hybrid functionals against fixed-node projector quantum Monte Carlo (QMC). QMC is an accurate many-body method that has proven to be exceptionally accurate in the study of hydrogen, making it suitable as a reference in regions where experimental data are sparse. In Sec. II and in the Supplemental Material [14], we provide details of the methods used. In Sec. III, we quantitatively establish that nonlocal density functionals have superior energetics but poorer pressure estimation as compared with LDA and PBE,

*rcclay2@illinois.edu

justifying the results of Refs. [7,11]. We then compare how functional choice affects the predicted hydrogen structures by looking at bond lengths, QMC enthalpies, the shape of the intramolecular potential, and the location of the liquid-liquid phase transition. In Sec. IV, we draw our conclusions. We have also included Supplemental Material, covering computational details, and providing detailed tables of density functional performance [14].

II. METHOD

Gauging the accuracy of density functionals has been typically done based on experimental data, for example, the G2 test set of molecular binding energies [15]. In contrast, our method is fully *ab initio*, and tailored to dense hydrogen. We used DFT-PIMD simulations to generate various "test sets" for a variety of phases and thermodynamic conditions as described in the following. After establishing reference energies and pressures with QMC, we performed DFT calculations on our test set using up to 10 different density functionals. Various measures were used to quantify functional accuracy. Full details are given in the Supplemental Material [14]. First, we outline the quantum Monte Carlo methods and then the various test sets and error estimates.

A. Quantum Monte Carlo calculations

Several different QMC techniques were used. The quantum Monte Carlo package (QMCPACK) [16,17] was used in all diffusion Monte Carlo (DMC) and variational Monte Carlo (VMC) simulations. The Born-Oppenheimer path-integral Monte Carlo (BOPIMC) code [18–21] was used in all reptation quantum Monte Carlo (RQMC) simulations. For all solid and liquid structures considered in this work, we first performed a wave-function optimization using VMC. We used Slater-Jastrow–type wave functions with backflow transformation in all cases. The single-particle orbitals were obtained from a PBE-DFT calculation using the QUANTUM ESPRESSO [22] package. This was followed by DMC calculations, within the fixed-node approximation. For a subset of the solid configurations we use RQMC, combined with correlated sampling, to calculate energy differences as we varied the bond length of the molecules in the solid. The QMC calculations were all electron, using the bare Coulomb interaction.

We used the virial estimator to calculate pressures directly from QMC, which for the Coulomb interaction has the form $P = \frac{1}{3\Omega}(2T + V)$, where T (V) is the kinetic (potential) energy and Ω is the simulation (supercell) volume. In order to minimize systematic errors in the calculation, we rewrite the estimator as $P = \frac{1}{3\Omega}(E + T)$, where E is the ground-state energy, and use an extrapolated estimator for the kinetic energy [23] $T_{\text{extrap}} = 2T_{\text{DMC}} - T_{\text{VMC}}$, which produces a pressure that is correct to second order in the quality of the trial wave function.

Controllable errors, such as time-step error, projection time, and population bias, were reduced to be comparable to our desired statistical error, giving an accuracy of approximately 0.01 mHa/proton for energy and 0.3 GPa for the pressure estimates. Finite-size effects were handled through a combined use of twist-averaged boundary conditions and post-processing corrections, which are detailed in the Supplemental Material [14]. The only source of uncontrollable errors comes through the use of the fixed-node approximation, which we expect to be very small for hydrogen.

B. Test sets

We define a test set S as a set of M proton configurations $R_S = \{\mathbf{R}_1, \ldots, \mathbf{R}_M\}$ all at a given density, temperature, and phase (liquid/solid). To broadly classify a functional's accuracy, it is important that these test sets consist of uncorrelated but physically reasonable configurations, representative of the state of hydrogen at high pressures.

By "physically reasonable," we mean that we would like to sample configurations roughly in proportion to their likelihood of appearance under some set of considered thermodynamic conditions. This is because the most probable configurations will most strongly affect most thermodynamic averages. For example, if we wish to calculate expectation values over the canonical ensemble, we would be more interested in establishing the accuracies of configurations with large Boltzmann weights (low energies) over those with extremely small Boltzmann weights (very high energies). As we do not actually know the potential energy surface, we will end up approximating it with DFT for the purposes of sampling from thermodynamic probability distributions. This will not yield precise thermodynamic averages, but should be adequate for roughly identifying probable configurations.

By "uncorrelated" we mean that the samples are drawn in a statistically independent manner from a desired probability distribution. Specifically, the configurations are generated from a single molecular dynamics simulation. Configurations are output to the test set at well-separated times along the MD trajectory, long enough apart in time that they can be considered to be independently drawn from the Boltzmann distribution. We verified that successive samples were uncorrelated by looking at the autocorrelation function of the energy.

Given the above considerations, we use constant volume PIMD simulations based on DFT to generate our test sets. In particular, we use the PI + GLE algorithm of Ceriotti *et al.* [12], which has been shown to significantly accelerate the rate of convergence of PIMD calculations with respect to the number of imaginary-time slices. By using PIMD to generate our configurations, we ensure that we adequately sample both nuclear quantum effects and thermal fluctuations, which are critical for a proper description of hydrogen. Note that for hydrogen at 1000 K, NQE are larger than thermal effects. Care was taken to ensure that the selected configurations were statistically independent and well equilibrated.

For the molecular solid, we considered the following structures obtained in previous structure searching studies: $C2/c$ [24], $Cmca$-12 [24], $Pbcn$ [24], and $Cmca$ [25]. For each of these structures, we first obtained zero-temperature reference configurations from DFT, at pressures of $P^{\text{DF}} = 200$ and 300 GPa, by performing constant pressure structural relaxation using the vdW-DF2 functional. P^{DF} is calculated using the vdW-DF2 functional. The band gaps calculated with this functional were in good agreement with experimental measurements when nuclear quantum effects are taken into account [26]. Starting from these configurations, PIMD

simulations were performed at a temperature of $T = 200$ K, also with the vdW-DF2 functional. All the simulations in the solid, and hence the configurations in the resulting test sets, were performed with 96 protons. PIMD + GLE simulations in the solid were performed with a modified version of the Vienna *ab initio* simulation package (VASP) [27–30]. We used the projector augmented wave representation of VASP and a $3 \times 3 \times 3$ k-point grid in the simulations. For each combination of pressure and structure, we generated configuration sets with at least 20 configurations. While we refer to the configuration sets by a structure and a pressure, these are really constant density sets. In addition, the density corresponds to that of the static lattice optimized at the given electronic pressure; thermal and quantum effects are not included during the structure optimization. Because the difference in density for different structures optimized at a given pressure is small, P^{DF} is simply a convenient label unless otherwise indicated.

For liquid hydrogen, we considered configurations of 54 protons at three densities given by $r_s = 1.30$, 1.45, 1.60, where r_s is the Wigner-Seitz radius, and a temperature of $T = 1000$ K. The three densities correspond to a fully dissociated atomic liquid, a molecular liquid in the neighborhood of the liquid-liquid phase transition, and a fully molecular liquid, respectively. These test sets allow us to compare the performance of DFT functionals in different environments in the liquid. The PIMD + GLE simulations in this case were performed with a modified version of QE and the vdW-DF2 functional. We used a Troullier-Martins norm-conserving pseudopotential with a cutoff of 0.5 bohrs and a $2 \times 2 \times 2$ k-point grid. Test sets in the liquid were generated with approximately 100 configurations at each density.

Finally, we used an enlarged set of zero-temperature configurations in order to study the influence of the choice of functional on structural properties in the solid. Specifically, we study the following structures from previous structure searching studies: $C2/c$, $Cmca$-12, $Pbcn$ [24], $Cmca$ [25], and $mC24$-$C2/c$ [31]. Ground-state structures were relaxed using the PBE, vdW-DF, and vdW-DF2 functionals at $P^{\mathrm{DF}} =$ 200, 300, and 400 GPa.

For additional details on the configuration sets used in this work, see the included Supplemental Material [14] section.

C. Density functional comparison

For each configuration in a test set in the solid, we calculated its energy and pressure using the following 10 functionals: LDA, PBE [32,33], vdW-DF [34], vdW-DF2 [10], vdW-optPBE, vdW-optB88 [9], vdW-optB86B [35], Becke-Lee-Yang-Parr (BLYP) [36,37], vdW-TS [38], and HSE [39]. For the vdW-TS functional, we used PBE exchange. We used a restricted set of functionals for the liquid test sets: LDA, PBE, vdW-DF, vdW-DF2, and HSE.

For all M configurations in a given test set S, we compute the density functional error of an observable $\mathcal{A}$ as $\delta\mathcal{A}^{\mathrm{DF}} = \mathcal{A}^{\mathrm{DF}} - \mathcal{A}^{\mathrm{QMC}}$. From this, we used two general measures of error. The first is the average error which we define as

$$\langle \delta\mathcal{A}^{\mathrm{DF}} \rangle_S = \frac{1}{M} \sum_{\mathbf{R}_i \in S} \delta\mathcal{A}^{\mathrm{DF}}(\mathbf{R}_i). \tag{1}$$

We also use a more general error measurement, akin to a mean-absolute error, which we define as

$$\langle |\delta\widetilde{\mathcal{A}^{\mathrm{DF}}}| \rangle_S = \frac{1}{M} \sum_{\mathbf{R}_i \in S} |\delta\mathcal{A}^{\mathrm{DF}}(\mathbf{R}_i) - c^{\mathrm{DF}}|. \tag{2}$$

Here, c is a density-functional-dependent offset, chosen by minimizing $\langle |\delta\widetilde{\mathcal{A}^{\mathrm{DF}}}| \rangle_{S'}$ over some set S' (which does not have to equal S). In the Results section, we will devise two error measures for energies that differ in their choice of reference point, but the motivation and justification will be handled later.

III. RESULTS

A. Benchmarking

Our concern in this section will be to establish how well various functionals capture different features of the Born-Oppenheimer (BO) potential energy surface, i.e., how well a functional calculates the electronic energy as a function of the proton positions. We will facilitate this discussion by distinguishing between two concepts: "global energetics" and "local energetics." Additionally, we complement this discussion by considering how well density functionals estimate pressures.

"Global energetics" will refer to how well a given functional is capable of capturing energy differences between arbitrary physically reasonable configurations at a fixed density. This is relevant in deciding which structure from a list of different structures has the lowest energy at a fixed density.

"Local energetics" will refer to how well a functional captures energy differences between a structure and perturbations around that structure, at a fixed density. In the case of a solid where a well-defined local minimum exists in the potential energy surface, we need to describe the shape of the potential energy well as accurately as possible. This is relevant for applications involving thermal and quantum fluctuations, such as the calculation of phonons and vibrational spectra, since what is relevant are energy differences between closely related structures. Thus, the ability of a functional to accurately capture these small energy differences will affect the quality of those predictions.

The previous two definitions refer to how well different features of the BO energy surface are captured at a fixed density. When determining phase stabilities, however, it is often important to compare energies between structures at different densities. By using basic thermodynamic identities, we can relate the error in the pressure to how much the energy error changes between similar structures at slightly different densities:

$$\frac{\partial}{\partial\rho}\delta e^{\mathrm{DF}} = \rho^{-2}\delta P^{\mathrm{DF}}. \tag{3}$$

We will focus on the errors in the pressure in this discussion, appealing to the above relationship.

A difficulty occurs when we try to compare DFT energy errors relative to QMC, $\delta e^{\mathrm{DF}} = e^{\mathrm{DF}} - e^{\mathrm{QMC}}$, because the DFT energy can have a constant offset depending on both the functional and the pseudopotential. Instead, we benchmark errors in energy differences $\delta\Delta e^{\mathrm{DF}} = \Delta e^{\mathrm{DF}} - \Delta e^{\mathrm{QMC}}$, where QMC and DFT energy differences are measured relative to some reference energy. Reference energies can be chosen in

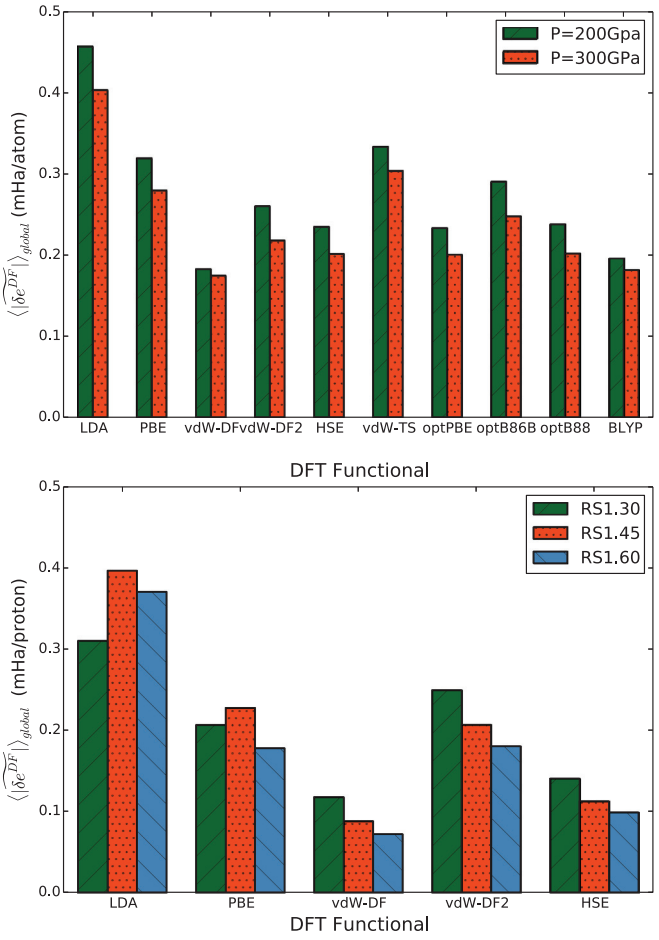

FIG. 1. (Color online) (Top) $\langle|\delta\widetilde{e^{\mathrm{DF}}(P)}|\rangle_{\mathrm{global}}$ versus functional for solid molecular test sets. Data for $P = 200$ and 300 GPa are shown in the legend. (Bottom) $\langle|\delta\widetilde{e^{\mathrm{DF}}(P)}|\rangle_{\mathrm{global}}$ versus functional for the liquid test sets. Data for $r_s = 1.30$, 1.45, and 1.60 shown.

various ways (relative to dissociated hydrogen atoms, relative to an arbitrary configuration, etc.), but we have preferred to use the methods described in the following.

1. Global energetics

To establish which functional has the best global energetic properties, we looked at the mean-absolute errors in the energy per proton. For a particular DFT pressure P^{DF}, we build a test set S' by including all the configurations consistent with that pressure, including ones with different structures. Then, for each pressure and density functional independently, we choose a reference point $c^{\mathrm{DF}}(P^{\mathrm{DF}})$ by minimizing $\langle|\delta\widetilde{e^{\mathrm{DF}}}|\rangle_{S'}$. In practice, this amounts to choosing the median of δe^{DF} on this aggregated test set S'.

Using this choice for $c^{\mathrm{DF}}(P^{\mathrm{DF}})$, we calculated $\langle|\delta\widetilde{e^{\mathrm{DF}}}|\rangle_S$ over each structure and averaged this over all structures to obtain what we denote $\langle|\delta\widetilde{e^{\mathrm{DF}}(P^{\mathrm{DF}})}|\rangle_{\mathrm{global}}$.

In the top of Fig. 1, we plot $\langle|\delta\widetilde{e^{\mathrm{DF}}(P)}|\rangle_{\mathrm{global}}$ versus functional for solid molecular hydrogen. We included results for both the $P = 200$ and the 300 GPa structures, which are marked with striped and dotted bars, respectively. Two things immediately stand out. The first is that nearly all of the hybrid and improved van der Waals functionals, excluding vdW-TS, noticeably outperform the LDA and PBE functionals. Second, the vdW-DF functional seems to have the best global energetic performance out of all functionals considered, followed by BLYP and HSE.

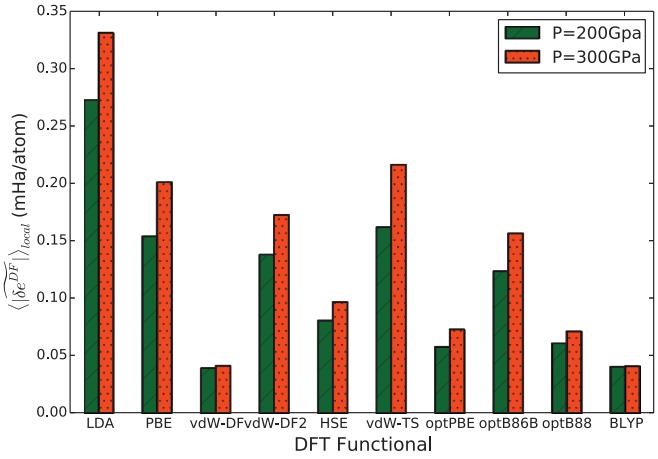

FIG. 2. (Color online) $\langle|\delta\widetilde{e^{\mathrm{DF}}(P)}|\rangle_{\mathrm{local}}$ versus functional for solid molecular test sets. Data for $P = 200$ and 300 GPa are shown in the legend.

In the bottom of Fig. 1, we show a plot of $\langle|\delta\widetilde{e^{\mathrm{DF}}(\rho)}|\rangle_{\mathrm{global}}$ for the liquid configurations. We have included data for the three densities $r_s = 1.30$, 1.45, 1.60, which are identified in the legend. Notice that as in the solids, vdW-DF is the best performing functional, although the hybrid functional HSE is a close runnerup.

Despite the vast differences in structures and densities, we see a very consistent picture regarding how accurate various functionals are in capturing global energetics. For the solid test set, we find that PBE is accurate to approximately 0.3 mHa/proton in dense hydrogen, whereas vdW-DF and HSE are good to 0.19 mHa/proton and 0.24 mHa/proton, respectively. The errors are smaller in the liquid phase, but the ordering of these functionals is the same for both cases with vdW-DF noticeably more accurate.

2. Local energetics

To measure the local energetics, we again used a shifted mean absolute error for the energy per proton, but with the reference point chosen to be specific to a given structure. For a test set S corresponding to a particular structure at pressure P, we again let the energy shift c^{DF} be chosen to minimize $\langle|\delta\widetilde{e^{\mathrm{DF}}(P)}|\rangle_S$ on the same set S. Averaging $\langle|\delta\widetilde{e^{\mathrm{DF}}(P)}|\rangle_S$ over all structures gives us a pressure-dependent measure of the local energetic errors, which we will denote $\langle|\delta\widetilde{e^{\mathrm{DF}}(P)}|\rangle_{\mathrm{local}}$. Notice that in this case, we are only concerned with relative errors between close configurations in the potential energy surface; systematic shifts between the various structures are not considered.

In Fig. 2, we show $\langle|\delta\widetilde{e^{\mathrm{DF}}(P)}|\rangle_{\mathrm{local}}$ versus density functional for solid molecular hydrogen. The results for $P = 200$ and 300 GPa are shown on the plot with dashed and dotted bars, respectively. The vdW-DF functional was the most accurate in capturing relative energy differences between

similar configurations with BLYP a close second. After these functionals, the optPBE and HSE functionals exhibited fair performance. The worst performing DF was LDA, followed by vdW-TS, and then jointly by vdW-DF2 and PBE. This same trend was observed for the global energetic performance in Fig. 1.

It is interesting to note how the magnitudes of the global energetic and local energetic errors compare. LDA and vdW-DF2 have local errors that are approximately 70% the size of their global errors, and thus experience only modest accuracy gains when considering energy differences between closely related structures. HSE and PBE perform moderately better, having local errors that are approximately 50% and and 60% of their global errors. Lastly, the vdW-DF functional, beyond having the lowest magnitude of global energetic errors, experiences a local energy error that is approximately 25% of the global errors.

In summary, we find that the van der Waals functionals were most able to calculate relative energy differences around local minima, with vdW-DF and BLYP having the smallest local energetic errors. This might have been guessed from the previous section, as these same functionals were, on average, the best for capturing large-scale energy differences. Thus, for structural relaxation, zero-point energy calculations, QMD, and other applications where location and shape of the local minimum are important, the vdW-DF functional is strongly recommended.

3. Pressures

For a test set S corresponding to a structure at particular density, we averaged $\langle \delta P^{\mathrm{DF}} \rangle_S$ over all structures at the same pressure or density to estimate the error in the pressure. The top of Fig. 3 shows $\langle \delta P^{\mathrm{DF}} \rangle$ for the solids. We see that in contrast to the local and global energetics sections, the semilocal functionals have some of the lowest pressure errors. HSE is the best performing functional in this regard. Note that the van der Waals functionals are among the worst performing functionals for the average pressures, with vdW-DF coming in just behind vdW-DF2 for highest pressure errors. These observations are also seen in the bottom of Fig. 3, which shows $\langle P^{\mathrm{DF}} \rangle$ for the liquid configurations.

We also looked at $\langle |\delta \widetilde{P^{\mathrm{DF}}}| \rangle_{\mathrm{local}}$ for the solids and $\langle |\delta \widetilde{P^{\mathrm{DF}}}| \rangle_{\mathrm{global}}$ for the liquids, defined as was done with the energy errors in the local and global energetic sections. We find that the magnitude of $\langle |\delta \widetilde{P^{\mathrm{DF}}}| \rangle_{\mathrm{local}}$ across all configurations and densities is statistically indistinguishable from the error bars of our QMC pressure estimates, indicating that the errors in the pressure are roughly independent of the configurations. Thus, the pressure errors observed were mostly functional and density-dependent constant offsets from P^{QMC}. Such was not the case for the energy.

To conclude, when it comes to capturing global and local energy differences at a fixed density, including exact exchange or van der Waals effects will generally improve the energetics of density functionals for dense hydrogen. The vdW-DF functional in particular gives noticeable improvements over PBE in capturing global energetics, and does exceptionally well for capturing local energetics. In spite of this, HSE and the semilocal functionals outperform nearly all the van der Waals functionals when it comes to correctly estimating pressures. Given how systematic the pressure errors are, one can correct the pressure of energetically favorable DFs such as vdW-DF by estimating an overall correction from either LDA or QMC. Fortunately, these errors are far more consistent than the energy errors, and so there should be some way of improving upon these functionals for future hydrogen applications.

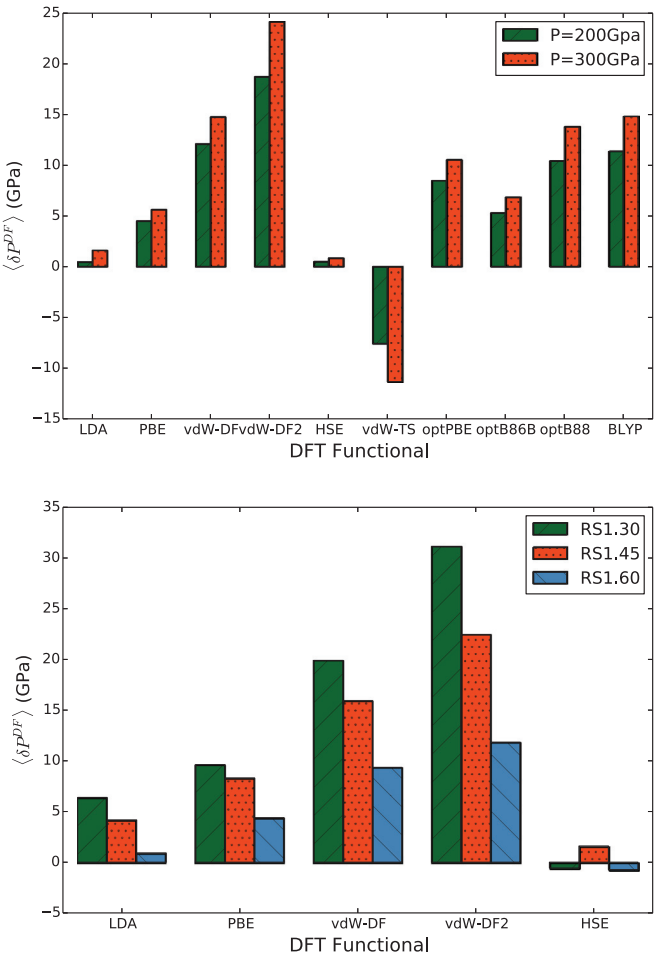

FIG. 3. (Color online) (Top) $\langle \delta P^{\mathrm{DF}} \rangle$ versus DFT functional for the molecular solids. (Bottom) $\langle \delta P^{\mathrm{DF}} \rangle$ versus DFT functional for the liquid configurations.

B. Effects of functional choice

In this section, we see how the energetic considerations of the benchmarking section relate to current problems of interest in the phase diagram of high-pressure hydrogen. Specifically, we look at how accurately different functionals predict H_2 bond lengths relative to QMC optimized structures. We also look at QMC cold curves for ground-state structures optimized with different functionals, and at the relation between the location of the LLPT and the mean absolute error of a selection of DFT functionals.

1. Bond lengths

The magnitude and pressure dependence of the bond length of the hydrogen molecule in the solid depends significantly on

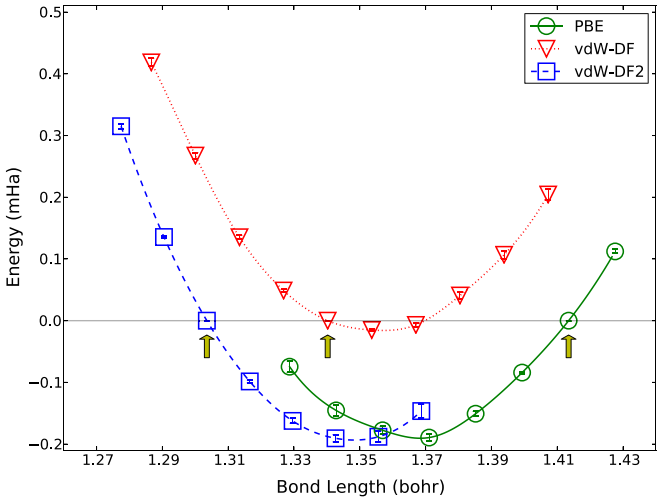

FIG. 4. (Color online) Molecular bond-length dependence of the QMC energy for the $C2/c$ structure at 200 GPa. The energy is measured with respect to the optimal bond length according to DFT. Arrows denote the equilibrium bond length predicted by DFT.

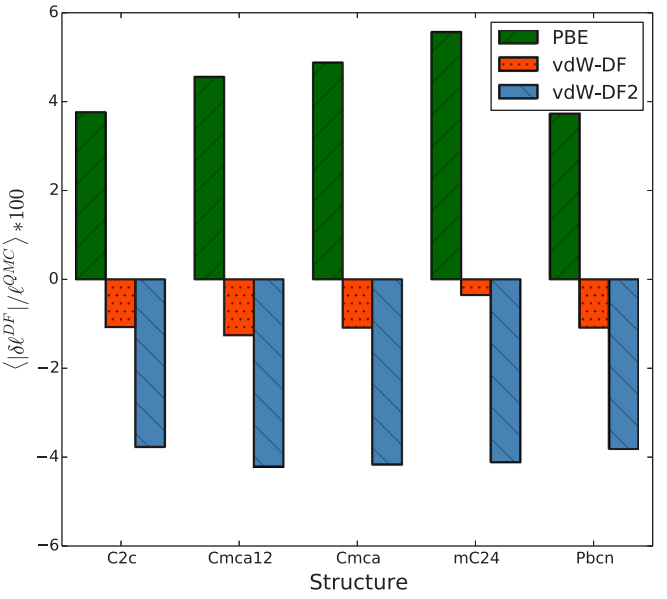

FIG. 5. (Color online) Error in the molecular bond length ℓ of selected DFT functionals, relative to QMC optimized values. The results are averaged over all pressures considered in this work since the pressure dependence of the error is small.

the DFT exchange-correlation functional. The LDA and PBE functionals predict bond lengths that are larger than those of the isolated molecule, while HSE, vdW-DF, and vdW-DF2 predict bond lengths that are smaller. In order to measure the ability of each functional to predict the correct magnitude and pressure dependence of the bond length of the molecule in the solid phase, we calculated the dependence of the energy on bond length for several of the proposed solid phases of hydrogen using QMC and compared with DFT predictions. To do this, we optimized a set of candidate structures ($C2/c$, $Cmca$, $Cmca12$, $Pbcn$, and $mC24$-$C2/c$) with the PBE, vdW-DF, and vdW-DF2 density functionals at three different pressures: 200, 300, and 400 GPa. For each combination of structure, density functional, and pressure, we calculated the dependence of the energy as a function of the molecular bond length using a QMC correlated sampling technique. In particular, we calculated the change in energy produced by scaling all the molecular bond lengths in the solid by a given fraction. Figure 4 shows an example of this procedure for the $C2/c$ structure. In this figure, the energy difference is with respect to the optimal molecular bond length according to the density functional used in the structural optimization. As described above, PBE shows a significant overestimation of the bond length, while vdW-DF2 underestimates it with a comparable magnitude. The vdW-DF functional, on the other hand, agrees well with the QMC predictions producing a structure with a negligible energy error due to the relaxation of the bond length. This relaxation energy can be significant in structures predicted by the other functionals and, since it is not guaranteed to be consistent between structures, it can significantly bias structural predictions. While the optimal bond lengths according to QMC are very similar for this structure, they depend slightly on the other structural parameters; we only relax the bond lengths in this calculation, leaving both molecular orientations and simulation cell fixed. As we discuss in the following, this produces an additional variation in the energy of the structures, making them dependent on the functional used to optimize them, even after the bond lengths have been relaxed.

Figure 5 shows the difference between the optimal QMC and DFT molecular bond length for each structure and density functional, averaged over all the pressures considered. In general, the discrepancy between QMC and DFT on the magnitude of the bond length is fairly insensitive to pressure in the range considered. Table I shows a summary of the difference in QMC energy between the optimal DFT and QMC bond lengths for all the structures, pressures, and DFT functionals considered in this work. Figure 6 shows these same data in a scatter plot, organized by P^{DF} and DF. The spread of values in the case of PBE is quite large, while the corresponding spread for vdW-DF is very small. Inaccuracies in the bond length can lead to important limitations in the predictive capabilities of DFT in this regime of the phase diagram, mainly due to the immediate proximity of this regime to a dissociation and a metal-insulator transition at higher pressures. Notice that more accurate electronic-structure calculations typically require the use of structures from more approximate methods, such as DFT. Incorrect predictions of structural properties will lead to biases in the predictions of more accurate methods. On the other hand, the ordering of structures can also be significantly biased if structural parameters are not accurate, in this case the existence of molecules with large bond lengths present in many of the proposed structures for hydrogen near metallization is put into question by these calculations since they have all been predicted using PBE, which severely overestimates bond lengths.

2. *Ground-state structures*

Although other groups have used QMC energies to calculate the cold curves of proposed solid hydrogen structures [40],

TABLE I. QMC energy difference (in mHa/proton) between the structures at the optimal DFT and QMC bond lengths. Error bars are shown in parentheses. The mean energy difference, averaged over all structures and all pressures, is shown in square brackets next to each functional name.

Structure	200 (GPa)	300 (GPa)	400 (GPa)
	PBE [0.19]		
$C2/c$	0.180(6)	0.186(9)	0.171(9)
$Cmca$	0.30(1)	0.15(2)	0.15(1)
$Cmca-12$	0.185(8)	0.24(2)	0.16(2)
$Pbcn$	0.37(1)	0.28(2)	0.23(2)
$mC24-C2/c$	0.069(4)	0.079(5)	0.102(6)
	vdW-DF [0.01]		
$C2/c$	0.018(3)	0.007(3)	0.015(3)
$Cmca$	0.011(4)	0.007(4)	0.017(4)
$Cmca-12$	0.016(3)	0.012(7)	0.023(4)
$Pbcn$	0.016(3)	0.016(3)	0.015(4)
$mC24-C2/c$	−0.005(3)	0.001(5)	0.008(3)
	vdW-DF2 [0.22]		
$C2/c$	0.197(6)	0.239(8)	0.247(7)
$Cmca$	0.25(1)	0.21(1)	0.22(1)
$Cmca-12$	0.236(6)	0.260(7)	0.24(1)
$Pbcn$	0.22(1)	0.216(8)	0.22(1)
$mC24-C2/c$	0.222(2)	0.18(1)	0.22(2)

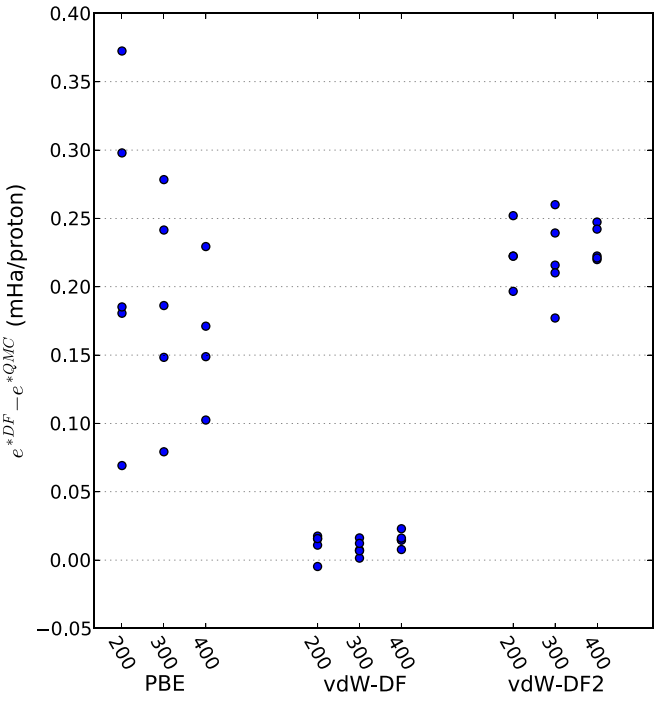

FIG. 6. (Color online) Scatter plot of $e^{*\mathrm{DF}} - e^{*\mathrm{QMC}}$ for structures optimized with different functionals and at different P^{DF}. The asterisks denote that we are measuring the QMC energy difference between a DF optimized structure and a bond-length optimized structure. The numbers on the x axis correspond to the P^{DF} in GPa at which all ground-state structures were relaxed. The names under each triplet of numbers denote the DF used in the structural optimization.

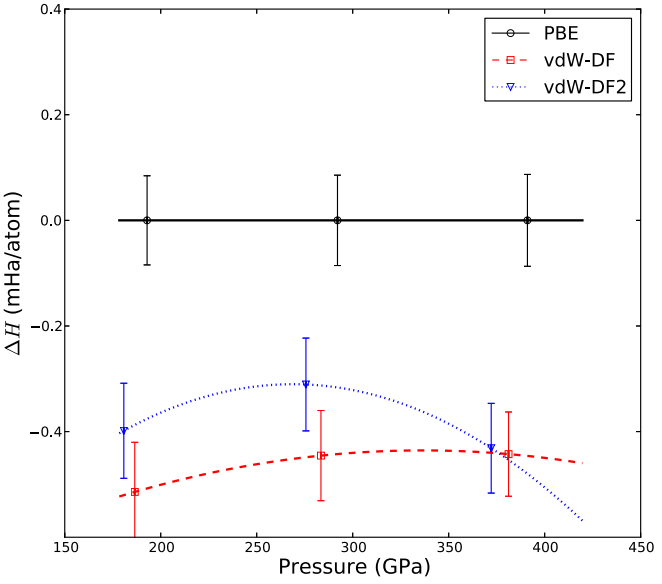

FIG. 7. (Color online) Relative enthalpies of $C2/c$ structures relaxed with either PBE, vdW-DF, or vdW-DF2.

here we use QMC energies and pressures to compare the relative enthalpies of proposed ground-state structures optimized with different functionals. In Fig. 7, we show the QMC enthalpies (relative to the enthalpies of the PBE optimized structures) of $C2/c$ structures relaxed with the PBE, vdW-DF, and vdW-DF2 functionals at $P_{\mathrm{DFT}} = 200$, 300, and 400 GPa, respectively. Structures relaxed with the vdW-DF and vdW-DF2 functionals have lower enthalpies than those generated with PBE, as we might have guessed from the relative energetics and from the previous discussion of bond lengths. Notice that the QMC pressures for all structures differ from their DFT values. The magnitude of the deviance follows the ordering we discussed above.

At the top of Fig. 8, we show the relative enthalpies for structures obtained through PBE and vdW-DF optimization. As in the previous figure, relative deviations from data for $C2/c$ structures optimized with PBE functional are shown. We see that structures optimized using the PBE and vdW-DF functionals exhibit similar qualitative features; the ordering of the ground-state structures is consistent between functionals, as are the pressure trends in the relative enthalpy curves. However, the vdW-DF structures are all lower in enthalpy than their PBE counterparts, and there are noticeable quantitative differences. For instance, the relative enthalpy differences between $Cmca$ and $mC24-C2/c$ are much larger in PBE around 200 and 400 GPa than in the vdW-DF functional.

At the bottom of Fig. 8, we compare the relative enthalpy differences between structures optimized with the vdW-DF and vdW-DF2 functionals. Note that even though the relative enthalpies for vdW-DF2 are all lower than for PBE, they always remain a little higher than for structures optimized using the vdW-DF functional. Both functionals illustrate the same qualitative trends, although note that the quantitative agreement is much greater than between vdW-DF and PBE. For instance, the relative enthalpy difference between $Cmca$ and $mC24-C2/c$ is much smaller with vdW-DF2 than PBE.

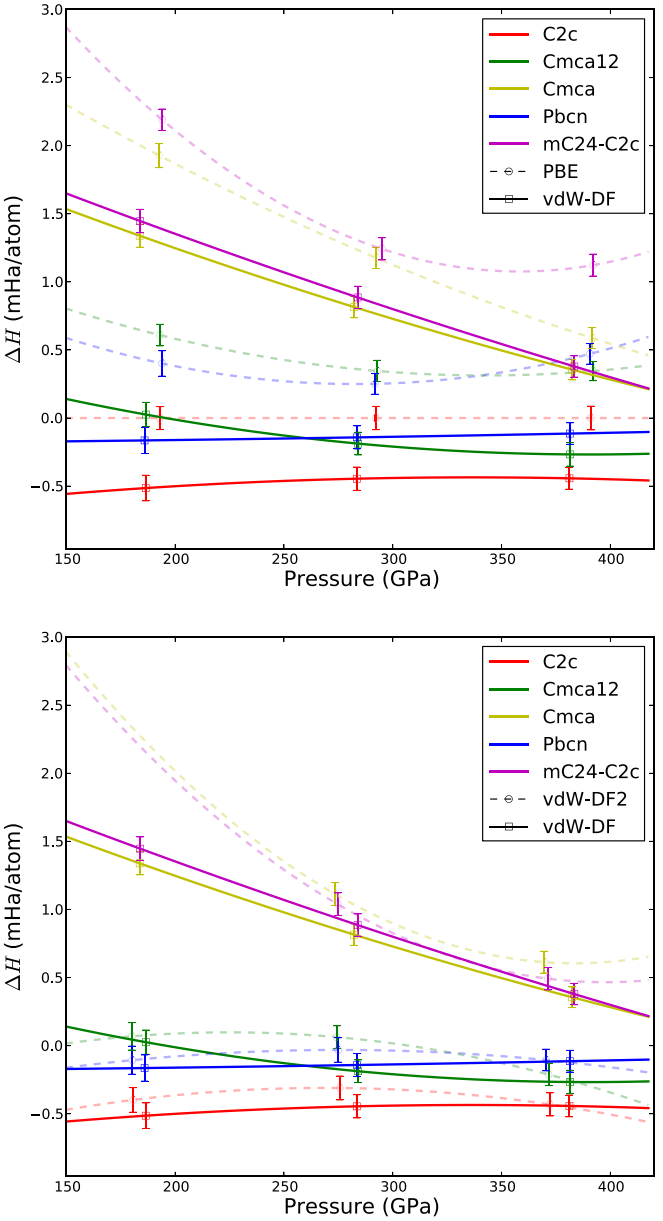

FIG. 8. (Color online) (Top) Relative enthalpies of candidate ground-state structures relaxed with either PBE or vdW-DF. Structures relaxed with the vdW-DF functional are shown with bold solid lines, and those with the PBE functional with light dashed lines. The color of the line denotes the structure. (Bottom) Comparing vdW-DF and vdW-DF2 functionals. The $P^{\mathrm{DF}} = 200$ GPa $Cmca$ and $mC24-C2/c$ structures optimized with the vdW-DF2 functional had significantly lower QMC pressures than the vdW-DF and PBE structures, and so these points are not shown but were used in the fit.

3. *Intramolecular potential*

While an accurate bond length is necessary for a precise prediction of the relative electronic enthalpy between different structures, the shape of the intramolecular potential is equally important for the accurate description of both thermal and quantum ionic components of the enthalpy. From the calculations of the bond-length dependence of the energy, we can measure the ability of each DFT functional to reproduce the intramolecular potential of the hydrogen molecule in various structures. For this purpose, similar to the RQMC calculations presented above, we calculated the bond length dependence of the electronic energy with PBE, vdW-DF, and vdW-DF2, on the same structures presented in Sec. III B 1. From the resulting energies, we obtain the curvature α of the potential at the corresponding equilibrium bond length by fitting a quadratic function. This curvature is directly related to the vibrational frequency of the molecule, which is the leading contribution to the zero-point energy due to its high frequency. Inaccuracies in the curvature of the potential will lead to systematic errors in the zero-point energies calculated with DFT.

TABLE II. Comparison of the curvature of the intramolecular potential of hydrogen between selected DFT functionals and RQMC. The table shows the percentage of difference between RQMC and DFT, calculated as $PE(\%) = (\alpha_{\mathrm{DFT}} - \alpha_{\mathrm{RQMC}})/\alpha_{\mathrm{RQMC}} * 100$, where α is the second derivative of the total energy as a function of the molecular bond length in the solid at the equilibrium distance. The relative mean absolute error, averaged over all structures and all pressures, is shown in parentheses next to each functional name.

Pressure (GPa)	$C2/c$	$Cmca$	$Cmca-12$	$Pbcn$	$mC24-C2/c$
			PBE (24.7)		
200	−26(3)	−28(5)	−24(5)	−21(4)	−34(5)
300	−21(3)	43(16)	−29(23)	−23(4)	−20(9)
400	−25(5)	−12(7)	19(12)	3(9)	−42(6)
			vdW-DF (7.3)		
200	−3(2)	−17(2)	−10(2)	−9(1)	−5(3)
300	−8(1)	−6(3)	−8(4)	−9(2)	−5(5)
400	9(5)	5(9)	0(7)	−4(6)	11(7)
			vdW-DF2 (19.2)		
200	19(5)	20(6)	34(2)	27(3)	37(6)
300	10(3)	6(7)	15(7)	16(4)	22(8)
400	12(4)	10(7)	15(6)	25(6)	37(16)

Table II and Fig. 9 show a comparison of the curvatures of the intramolecular potential between various DFT functionals and RQMC calculations. It is clear that PBE systematically underestimates the magnitude of the curvature, by an average of ∼20% over the studied pressure range. vdW-DF2, on the other hand, systematically overestimates the curvature, but by a smaller amount. In both cases, the variation with structure is large. As can be expected, the vdW-DF functional shows a good overall agreement with QMC, producing an average discrepancy of ∼5% over the entire configuration set. This is consistent with the results of Sec. III A 2 that shows that vdW-DF gives a more accurate estimate of the local potential energy surface of the solid and of the relative energies of different molecular configurations, relative to PBE and vdW-DF2. As a consequence, both thermal and zero-point components of the energies should be computed with this functional for a more predictive calculation.

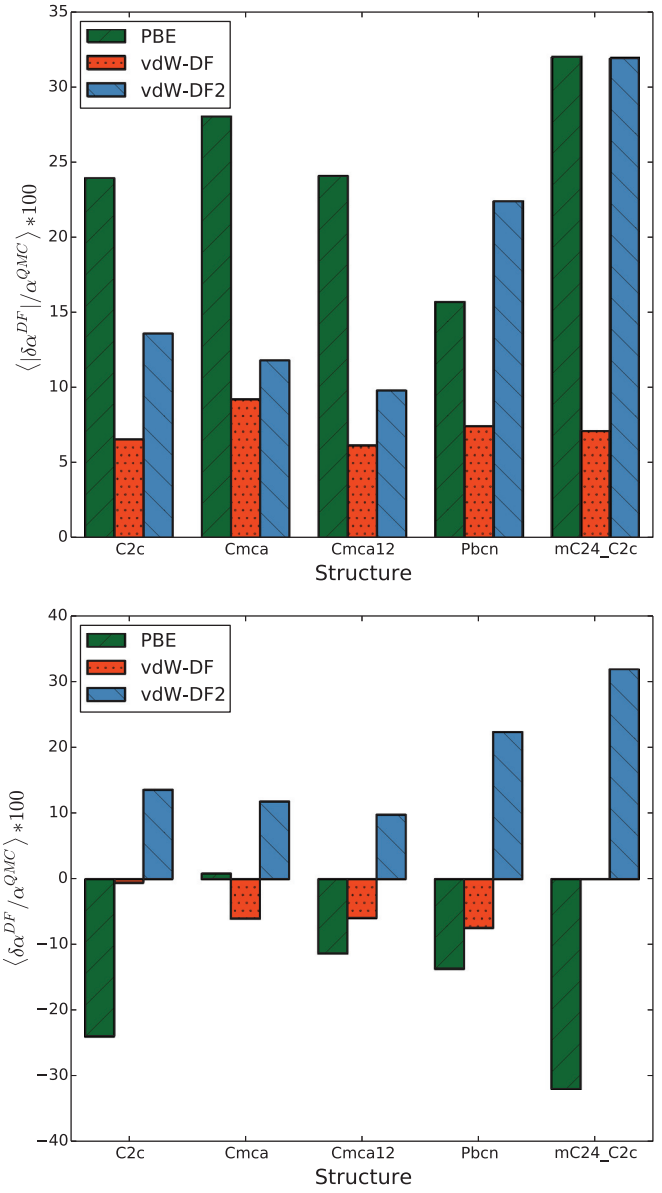

FIG. 9. (Color online) Above are two measures of the relative errors in the curvature α for different structures and functionals. The angle brackets denote averaging over all pressures. (Top) Relative mean absolute error of the curvature for each functional and structure. (Bottom) Relative mean error of the curvature.

4. *Liquid-liquid phase transition (LLPT)*

The location of the LLPT in hydrogen has been recently shown to depend significantly on both NQEs and on the treatment of electronic exchange and correlation [8]. While the inclusion of NQEs typically reduces the transition pressure by ∼60–80 GPa, different DFT functionals produce variations by as much as ∼200 GPa. In this section, we use the QMC calculations presented in Sec. III A 1 to show that there is a correlation between the location of the LLPT estimated by a given functional and the difference in the average energy between the QMC energy and a DF energy.

The predicted transition pressure will be affected by differences of errors in the two branches of the free energy

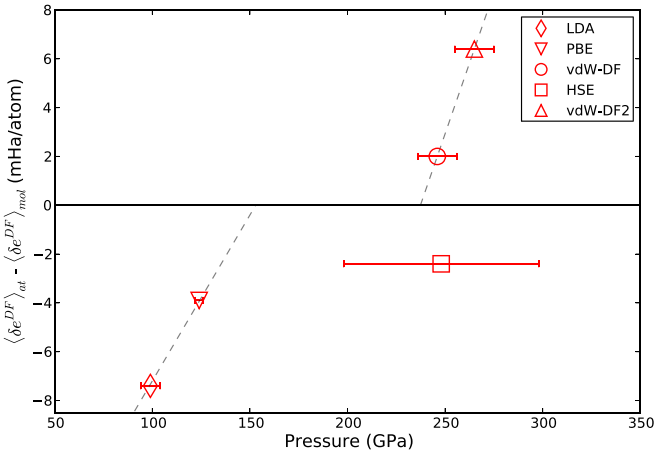

FIG. 10. (Color online) $\langle \delta e^{\mathrm{DF}} \rangle_{\mathrm{at}}$-$\langle \delta e^{\mathrm{DF}} \rangle_{\mathrm{mol}}$ for a selection of DFT functionals as a function of estimated P^{LLPT} for 54 protons at $T =$ 1000 K. The different marker styles indicate different functionals. Dashed lines are linear fits to the positive and negative data points.

$F(T,V)$ isotherm at the transition point. To estimate the principal effect of DF errors, we look at the difference in DF internal energy errors between two different densities corresponding to the atomic liquid and molecular liquid. We begin, as in the global and energetic sections of this paper, by defining a test set S' to be the aggregate of all liquid test configurations at $r_s = 1.30, 1.45, 1.60$. Choosing $\widetilde{c^{\mathrm{DF}}}$ to be the median of this aggregated set, we calculate $\langle \delta e^{\widetilde{\mathrm{DF}}}(r_s) \rangle$ for $r_s = 1.30$ and 1.60. Then, $(\langle \delta e^{\mathrm{DF}} \rangle_{\mathrm{at}} - \langle \delta e^{\mathrm{DF}} \rangle_{\mathrm{mol}})$ measures the mean energy shift between the atomic and molecular states.

We then estimate the transition pressure using several functionals; the procedure for calculating the transition pressure is given in the Supplemental Material [14]. In Sec. III A 3, we found that there is a systematic and sometimes sizable error in DF pressure estimates, which causes an additional bias of the transition pressures. We correct for this error by fitting the pressure errors to $\delta P^{\mathrm{DF}}(r_s) = a_0 + a_1 P^{\mathrm{QMC}}(r_s) + a_2 [P^{\mathrm{QMC}}(r_s)]^2$ for each functional, where the coefficients a_i are assumed to be independent of density. We then solve this equation for P^{QMC} as a function of P^{DF}, which gives us a corrected transition pressure.

In Fig. 10, we plot $\langle \delta e^{\mathrm{DF}} \rangle_{\mathrm{at}}$-$\langle \delta e^{\mathrm{DF}} \rangle_{\mathrm{mol}}$ versus the corrected transition pressure for all considered functionals. We see that the errors change sign as we go from PBE, LDA, and HSE to the van der Waals functionals. If we knew the functional relationship between the energy errors and the transition pressure, the point where that function crosses the x axis should coincide with the correct transition pressure for a 54-atom system at 1000 K. Although our data are too sparse to characterize this function, we can attempt to bound the transition. By performing linear fits on the positive and negative data points independently, we show in Fig. 10 that the transition probably lies between approximately 150 and 240 GPa. This is about a 40% reduction in the transition pressure uncertainty due to electronic correlation pointed out in previous work [8].

We have made preliminary estimates of the transition pressure using the coupled electron-ion Monte Carlo [18], a

method which treats the electrons with QMC. Using quantum protons, we estimate the transition pressure for a 54-atom system at $T = 1000$ K to be around 221 GPa; details will be published in a future work. It is reassuring to see that this estimate lies within the transition pressure error bounds obtained in this section.

IV. SUMMARY AND CONCLUSIONS

In this paper, we presented a detailed benchmark of DFT exchange-correlation functionals in high-pressure hydrogen using accurate QMC calculations. Particular care was taken to control systematic errors in the QMC calculations, including size effects, time step, twist averaging, projection times, and population control. We find that the performance of most DFT functionals depends on the property being studied. While LDA and HSE consistently produce the best pressures in both solid and liquid phases, vdW-DF is clearly superior in terms of local energy differences in the potential energy surface. HSE and vdW-DF perform equally well in terms of energetics in liquid hydrogen close to metallization, but with errors in opposite directions. In general, PBE does a rather poor job at describing the relative energies of configurations and describes quite poorly the properties of the molecular bond. This leads to a large underestimate of the metallization transition and of both thermal and zero-point energy contributions. Predictions made with this functional are much less accurate. While no functional considered was capable of accurately describing relative energy differences over a large region of the phase diagram, vdW-DF was found to produce excellent results within a given phase, particularly for the reproduction of the intramolecular potential and equilibrium bond length. This functional should be used to estimate zero-point energy, which is dominated by energy differences close to an equilibrium configuration.

ACKNOWLEDGMENTS

M.A.M. and J.M. were supported by the U. S. Department of Energy at the Lawrence Livermore National Laboratory under Contract No. DE-AC52-07NA27344. M.A.M., J.B.M., R.C., and D.M.C. were supported through the Predictive Theory and Modeling for Materials and Chemical Science program by the Basic Energy Science (BES), DOE. R.C., J.M.M., and D.M.C. were also supported by DOE Grant No. DE-NA0001789. C.P. was supported by the Italian Institute of Technology (IIT) under the SEED project Grant No. 259 SIMBEDD. Computer time was provided by the U. S. DOE INCITE program, Lawrence Livermore National Laboratory through the 7th Institutional Unclassified Computing Grand Challenge program, and PRACE Project No. 2011050781.

[1] J. McMahon, M. A. Morales, C. Pierleoni, and D. M. Ceperley, Rev. Mod. Phys. **84**, 1607 (2012).
[2] S. T. Weir, A. C. Mitchell, and W. J. Nellis, Phys. Rev. Lett. **76**, 1860 (1996).
[3] E. Babaev, A. Sudbø, and N. W. Ashcroft, Nature (London) **431**, 666 (2004).
[4] K. Moulopoulos and N. W. Ashcroft, Phys. Rev. Lett. **66**, 2915 (1991).
[5] S. A. Bonev, E. Schwegler, T. Ogitsu, and G. Galli, Nature (London) **431**, 669 (2004).
[6] J. M. McMahon and D. M. Ceperley, Phys. Rev. B **84**, 144515 (2011).
[7] M. A. Morales, J. M. McMahon, C. Pierleoni, and D. M. Ceperley, Phys. Rev. B **87**, 184107 (2013).
[8] M. A. Morales, J. M. McMahon, C. Pierleoni, and D. M. Ceperley, Phys. Rev. Lett. **110**, 065702 (2013).
[9] J. Klimeš, D. R. Bowler, and A. Michaelides, J. Phys.: Condens. Matter **22**, 022201 (2010).
[10] K. Lee, E. D. Murray, L. Kong, B. I. Lundqvist, and D. C. Langreth, Phys. Rev. B **82**, 081101 (2010).
[11] X.-Z. Li, B. Walker, M. I. J. Probert, C. J. Pickard, R. J. Needs, and A. Michaelides, J. Phys.: Condens. Matter **25**, 085402 (2013).
[12] M. Ceriotti, D. E. Manolopoulos, and M. Parrinello, J. Chem. Phys. **134**, 084104 (2011).
[13] S. Azadi and W. M. C. Foulkes, Phys. Rev. B **88**, 014115 (2013).
[14] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevB.89.184106 for computational details and tables of density functional error measures.
[15] L. A. Curtiss, K. Raghavachari, P. C. Redfern, and J. A. Pople, J. Chem. Phys. **106**, 1063 (1997).
[16] K. Esler, J. Kim, D. Ceperley, and L. Shulenburger, Comput. Sci. Eng. **14**, 40 (2012).
[17] J. Kim, K. P. Esler, J. McMinis, M. A. Morales, B. K. Clark, L. Shulenburger, and D. M. Ceperley, J. Phys.: Conf. Ser. **402**, 012008 (2012).
[18] C. Pierleoni and D. M. Ceperley, in *Computer Simulations in Condensed Matter Systems: From Materials to Chemical Biology Volume 1 SE - 18*, Vol. 703 of Lecture Notes in Physics, edited by M. Ferrario, G. Ciccotti, and K. Binder (Springer, Berlin, 2006), pp. 641–683.
[19] M. A. Morales, C. Pierleoni, and D. M. Ceperley, Phys. Rev. E **81**, 021202 (2010).
[20] M. A. Morales, C. Pierleoni, E. Schwegler, and D. M. Ceperley, Proc. Natl. Acad. Sci. USA **107**, 12799 (2010).
[21] E. Liberatore, M. A. Morales, D. M. Ceperley, and C. Pierleoni, Mol. Phys. **109**, 3029 (2011).
[22] P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G. L. Chiarotti, M. Cococcioni, I. Dabo *et al.*, J. Phys.: Condens. Matter **21**, 395502 (2009).
[23] D. M. Ceperley and M. H. Kalos, in *Monte Carlo Methods in Statistical Physics*, edited by K. Binder (Springer, Berlin, 1979), Vol. 7, p. 145.
[24] C. J. Pickard and R. J. Needs, Nat. Phys. **3**, 473 (2007).
[25] B. Edwards, N. W. Ashcroft, and T. Lenosky, Europhys. Lett. **34**, 519 (1996).
[26] M. A. Morales, S. Hamel, K. Caspersen, and E. Schwegler, Phys. Rev. B **87**, 174105 (2013).
[27] G. Kresse and J. Hafner, Phys. Rev. B **47**, 558 (1993).

[28] G. Kresse and J. Hafner, Phys. Rev. B **49**, 14251 (1994).
[29] G. Kresse and J. Furthmüller, Comput. Mater. Sci. **6**, 15 (1996).
[30] P. E. Blöchl, Phys. Rev. B **50**, 17953 (1994).
[31] H. Liu, H. Wang, and Y. Ma, J. Phys. Chem. C **116**, 9221 (2012).
[32] J. P. Perdew and A. Zunger, Phys. Rev. B **23**, 5048 (1981).
[33] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).
[34] M. Dion, H. Rydberg, E. Schröder, D. C. Langreth, and B. I. Lundqvist, Phys. Rev. Lett. **92**, 246401 (2004).
[35] J. Klimeš, D. R. Bowler, and A. Michaelides, Phys. Rev. B **83**, 195131 (2011).
[36] C. Lee, W. Yang, and R. G. Parr, Phys. Rev. B **37**, 785 (1988).
[37] A. D. Becke, Phys. Rev. A **38**, 3098 (1988).
[38] A. Tkatchenko and M. Scheffler, Phys. Rev. Lett. **102**, 073005 (2009).
[39] J. Heyd, G. E. Scuseria, and M. Ernzerhof, J. Chem. Phys. **118**, 8207 (2003).
[40] S. Azadi, W. M. C. Foulkes, and T. D. Kühne, New J. Phys. **15**, 113005 (2013).