Molecular hydrogen adsorbed on benzene: Insights from a quantum Monte Carlo study

Todd D. Beaudet, Michele Casula, Jeongnim Kim, Sandro Sorella, and Richard M. Martin

View Online   Export Citation

ARTICLES YOU MAY BE INTERESTED IN

Weak binding between two aromatic rings: Feeling the van der Waals attraction by quantum Monte Carlo methods
The Journal of Chemical Physics **127**, 014105 (2007); https://doi.org/10.1063/1.2746035

Correlated geminal wave function for molecules: An efficient resonating valence bond approach
The Journal of Chemical Physics **121**, 7110 (2004); https://doi.org/10.1063/1.1794632

Geminal wave functions with Jastrow correlation: A first application to atoms
The Journal of Chemical Physics **119**, 6500 (2003); https://doi.org/10.1063/1.1604379

Molecular hydrogen adsorbed on benzene: Insights from a quantum Monte Carlo study

Todd D. Beaudet,[1,a)] Michele Casula,[1,b)] Jeongnim Kim,[2,3,c)] Sandro Sorella,[4,5,d)] and Richard M. Martin[1,e)]

[1]*Department of Physics, University of Illinois at Urbana-Champaign, 1110 W. Green Street, Urbana, Illinois 61801, USA*
[2]*National Center for Supercomputing Applications, 1205 W. Clark Street, Urbana, Illinois 61801, USA*
[3]*Materials Computational Center, University of Illinois at Urbana-Champaign, 104 S. Goodwin Avenue, Urbana, Illinois 61801, USA*
[4]*SISSA, International School for Advanced Studies, 34014 Trieste, Italy*
[5]*DEMOCRITOS, National Simulation Center, 34014 Trieste, Italy*

We present a quantum Monte Carlo study of the hydrogen-benzene system where binding is very weak. We demonstrate that the binding is well described at both variational Monte Carlo (VMC) and diffusion Monte Carlo (DMC) levels by a Jastrow correlated single determinant geminal wave function with an optimized compact basis set that includes diffuse orbitals. Agreement between VMC and fixed-node DMC binding energies is found to be within 0.18 mhartree, suggesting that the calculations are well converged with respect to the basis. Essentially the same binding is also found in independent DMC calculations using a different trial wave function of a more conventional Slater–Jastrow form, supporting our conclusion that the binding energy is accurate and includes all effects of correlation. We compare with previous calculations, and we discuss the physical mechanisms of the interaction, the role of diffuse basis functions, and the charge redistribution in the bond. © *2008 American Institute of Physics*. [DOI: 10.1063/1.2987716]

I. INTRODUCTION

In this paper we study the adsorption of molecular hydrogen on a benzene ring. Benzene is structurally similar to the five- or six-member rings that are characteristic building blocks of carbon nanotubes,[1] fullerenes,[2] metal-organic frameworks (MOFs),[3–5] and other related materials that have been recently studied with regard to hydrogen storage. While this system is not expected to bind hydrogen reversibly at ambient temperatures due to an expected binding energy that is much weaker than the necessary $20–40$ kJ/mol H_2 ($\sim 7–15$ mhartree H_2) needed for reversibility,[4] an accurate description of this structure is highly relevant to ongoing related research. Besides this, the hydrogen-benzene system is a good test case for theoretical predictions because of the stringent requirements to reliably determine binding energies at the desired accuracy level. Further, a careful study of this system is an important test of the transferability of empirical potentials[6,7] that have been constructed primarily from experimental data on graphitic systems. In this paper we consider only the case where the hydrogen dimer is oriented along the C_6 symmetry axis of the benzene molecule. Other papers[8,9] have found this to be the favored configuration and orientational differences are not taken into account in this work, since our main purpose is to present benchmark calculations for the most stable geometry.

There have been many previous studies of the binding of H_2 on benzene and related systems using various methods including density functional theory (DFT), Møller-Plesset second order perturbation theory (MP2), coupled cluster (CC) with single and double excitations (CCSD), and variations of these.[4,8,9] The values obtained so far for the binding energy, falling in the range of $0.4–1.9$ mhartree,[4,8,9] are very small and require a high level of accuracy of all the methods. The DFT calculations have great advantage as they are fast, scale well with system size, and can be readily converged with respect to the basis. However, the accuracy of their results is limited by the approximation on the exchange and correlation functionals, and there is no known way to systematically improve it. The many-body CC methods are the most accurate, although their applications are limited to small systems and not-so-large basis set due to poor scaling with the number of electrons and the size of the basis. Perturbation methods such as MP2 are valuable theories, with a system size scaling better than any CC method, but with intermediate accuracy. In the H_2-benzene system, perhaps the most accurate results to date have been derived from MP2 and CCSD(T) calculations by Hübner *et al.*[9] They found that the binding increases with increasing basis size in MP2 calculations, whereas it decreases as the level of the theory is improved to CCSD(T). Based on the best MP2 binding energy (1.87 mhartree) and the best CCSD(T) value (1.16 mhartree) with affordable bases, they estimated the actual binding energy to be ~ 1.5 mhartree. It should be noted that these numbers have already included basis set superposition error corrections as high as 0.36 mhartree and so rep-

resent a substantial fraction of the binding energy. Such methods necessitate carefully extrapolating the results with respect to the basis set and the level of theory. However, such extrapolations represent a difficulty due to the high computational cost of large bases, particularly in the CCSD(T) framework.

In the present work we study the hydrogen-benzene problem using quantum Monte Carlo (QMC) methods, that offer several advantages: many-body correlation effects can be explicitly included in the wave function, scaling with the system size is favorable like DFT, and calculations are variational and usually less dependent on the basis set. For a review and references to earlier work, see Ref. 10. By means of QMC, a trial correlated wave function can be optimized in the variational Monte Carlo (VMC) framework,[11,12] and its energy can be further minimized by the diffusion Monte Carlo (DMC) algorithm, which stochastically projects the optimized VMC (trial) wave function to the ground state. The only fundamental limitation is the well known "sign problem" for fermion systems, that does not allow a numerically stable calculation. Therefore, in this case, the so-called fixed node (FN) approximation is adopted by constraining the diffusion within the nodal pockets of the initial variational wave function.[10] Thus, the FN DMC method is unbiased only if the nodes of the trial wave function coincide with those of the true ground state. We addressed the issue of the FN bias in two ways. First, we used advanced QMC optimization methods[11–14] and physical principles to find a Jastrow correlated antisymmetrized geminal product[15,16] (JAGP) which gives a VMC binding energy with an accuracy comparable to the post Hartree–Fock methods. In addition, we computed the binding energy at the DMC level using the JAGP and a simpler Slater–Jastrow (SJ) form with a Perdew–Burke–Ernzerhof (PBE-DFT) optimized basis set as the trial wave function. The agreement found between them supports clearly the idea that our results are independent of the basis set and variational form, and it is a check for the accuracy of our DMC calculations against the FN approximation, since the nodes of the two wave functions are *a priori* different. This is encouraging for another reason: although the SJ trial function is not as accurate as the JAGP at the VMC level, it is more easily extended to larger systems which are important for future work. The necessary condition for that agreement is using a basis sufficiently extended in the tails. This is not surprising for a system driven by van der Waals (VdW) interactions which lead to weak binding and large equilibrium distance, as pointed out in Ref. 17, but it is a crucial point since the tails are not very important in the total energy. Our work shows that DMC can capture the correct binding as long as the basis is extended enough to allow accurate sampling of the outer regions of the molecules. This is brought out by a detailed study of the electron density changes due to binding.

The paper is organized as follows. In Sec. II we describe the QMC methods employed as well as the SJ and JAGP wave functions that serve as the variational guess in our QMC calculations. In Sec. III we discuss our results on the binding energy of the hydrogen-benzene system, where the hydrogen is oriented perpendicular to and centered over the

benzene at various molecular spacings. In Sec. IV we compare our findings to previous works. In Sec. V we discuss the physics of the hydrogen-benzene bond in terms of its electron density, by comparing the QMC and DFT-PBE results. Finally, we draw our conclusions in Sec. VI.

II. COMPUTATIONAL DETAILS

A. Wave functions

A wave function that describes a system of N identical fermions must be antisymmetric under particle exchange. To simplify the description of such a wave function, it is often useful to factor the wave function into a positive symmetric part, called the Jastrow factor, and an antisymmetric part so that a wave function can be expressed as

$$\Psi(\mathbf{x}_1, \ldots, \mathbf{x}_N) = J(\mathbf{x}_1, \ldots, \mathbf{x}_N)\Psi_{AS}(\mathbf{x}_1, \ldots, \mathbf{x}_N), \tag{1}$$

where $\mathbf{x}_i \equiv \{\mathbf{r}_i, \sigma_i\}$ is a space-spin coordinate, $J(\mathbf{x}_1, \ldots, \mathbf{x}_N)$ is the Jastrow factor, and $\Psi_{AS}(\mathbf{x}_1, \ldots, \mathbf{x}_N)$ is the antisymmetric part. The Jastrow can be further factored into one-body, two-body, three-body, and higher-body terms $(J = J_1 J_2 J_3 \cdots)$ which correspond to effective electron-ion, electron-electron, electron-electron-ion, etc., interactions.

One of the choices of trial function is to approximate the antisymmetric part as a single Slater determinant of spin orbitals. If there are no spin orbit interactions, the energy depends only on the spatial part of the wave function which can be written as a product of spin up and spin down determinants. In the unpolarized case, the spatial form is given by

$$\Psi_{AS}(\mathbf{x}_1, \ldots, \mathbf{x}_N) = \begin{vmatrix} \varphi_1(\mathbf{r}_1^\uparrow) & \cdots & \varphi_{N/2}(\mathbf{r}_1^\uparrow) \\ \vdots & \ddots & \vdots \\ \varphi_1(\mathbf{r}_{N/2}^\uparrow) & \cdots & \varphi_{N/2}(\mathbf{r}_{N/2}^\uparrow) \end{vmatrix}$$
$$\times \begin{vmatrix} \varphi_{N/2+1}(\mathbf{r}_1^\downarrow) & \cdots & \varphi_N(\mathbf{r}_1^\downarrow) \\ \vdots & \ddots & \vdots \\ \varphi_{N/2+1}(\mathbf{r}_{N/2}^\downarrow) & \cdots & \varphi_N(\mathbf{r}_{N/2}^\downarrow) \end{vmatrix}, \tag{2}$$

where each $\varphi_i(\mathbf{r})$ is a single-body space orbital and N is the total number of electrons. We used GAUSSIAN03 (Ref. 18) to derive the single-body orbitals $\varphi_i(\mathbf{r})$ via the PBE (Refs. 19 and 20) density functional with the VTZ Gaussian basis[21] modified to include diffuse functions from the aug-cc-pVTZ basis.[22]

We chose to use a very simple Jastrow factor because our goal was to improve DMC efficiency as opposed to obtaining a well converged binding curve at the VMC level. The Jastrow factor we applied to the Slater determinant is a Wagner–Mitas form[23] modified so that the electron-ion and electron-electron cusp conditions are fulfilled. The one- and two-body Jastrow terms we use are given by

$$J_1(\mathbf{R}) = \prod_{ia} \exp\left[\sum_k (b_{ak} r_{ia} + c_{ak}) v_{ak}(r_{ia}) \right] \tag{3}$$

and

$$J_2(\mathbf{R}) = \prod_{i<j} \exp\left[\sum_k (b_k r_{ij} + c_k) v_k(r_{ij})\right], \tag{4}$$

where $\mathbf{R} = \{\mathbf{r}_1, \ldots, \mathbf{r}_N\}$ specifies the N electron space coordinates, i and a index electrons and nuclei, respectively, r_{ia} and r_{ij} are electron-ion and electron-electron distances, and k indexes the expansion terms. In our work we used three terms and, when needed, one cusp term. In the above equations, $u_k(r) = (1 - z(r/r_{\text{cut}}))/(1 + \beta_k z(r/r_{\text{cut}}))$, with $z(x) = x^2(6 - 8x + 3x^2)$ and parameters b, c, β optimizable (with the exception of those that are cusp dependent). The function $z(x)$ has the properties $z(0) = z'(0) = z'(1) = 0$ and $z(1) = 1$, so that the Jastrow has a well defined cutoff at $r_{\text{cut}} = 10$ bohrs. Cusps between same spin electrons are not accounted for. This is justified because of the Pauli exclusion principle, which keeps them apart. It should be emphasized that the single-body Slater orbitals obtained from PBE-DFT are not further optimized since we would like to check the accuracy of the PBE-DFT nodes with respect to a more correlated and fully optimized wave function, such as the JAGP form described below. However, optimizing the above Jastrow is convenient as it improves the VMC energy and variance and shortens the DMC projection time, without changing the nodes.

The other trial function used in this work is the JAGP, where the antisymmetric part is a single determinant of two-body orbitals (geminals). This approach has been successfully applied in several contexts where electron correlations play a significant role. For example, the JAGP form is related to the pairing in the BCS wave function for superconductivity,[24,25] the resonating valence bond proposed by Pauling in 1939,[26] and can be used to describe strongly correlated electrons in transition metals. Recent applications in quantum chemistry include benzene,[15] the benzene dimer interacting via weak van der Waals forces,[12] and iron dimer.[27]

Since the ground state of the hydrogen-benzene system is an unpolarized spin singlet ($N^\uparrow = N^\downarrow = N/2$) the spatial part of the AGP wave function can be written as a determinant of pairing functions[28] without including unpaired orbitals, namely,

$$\Psi_{AS}(\mathbf{X}) = \begin{vmatrix} \phi(\mathbf{r}_1^\uparrow, \mathbf{r}_1^\downarrow) & \ldots & \phi(\mathbf{r}_1^\uparrow, \mathbf{r}_{N/2}^\downarrow) \\ \vdots & \ddots & \vdots \\ \phi(\mathbf{r}_{N/2}^\uparrow, \mathbf{r}_1^\downarrow) & \ldots & \phi(\mathbf{r}_{N/2}^\uparrow, \mathbf{r}_{N/2}^\downarrow) \end{vmatrix}, \tag{5}$$

where $\mathbf{X} = \{\mathbf{x}_1, \ldots, \mathbf{x}_N\}$ specifies the N electron space-spin coordinates and the paring function $\phi(\mathbf{r}_i^\uparrow, \mathbf{r}_j^\downarrow)$ can be expanded in single-body atomic orbitals so that

$$\phi(\mathbf{r}_i^\uparrow, \mathbf{r}_j^\downarrow) = \sum_{lmab} \lambda_{ab}^{lm} \varphi_{al}(\mathbf{r}_i^\uparrow) \varphi_{bm}(\mathbf{r}_j^\downarrow), \tag{6}$$

where l and m index the orbitals centered on ions a and b, respectively.

The Jastrow factor in the JAGP wave function is somewhat different from the one applied to the Slater determinant. The cusp conditions are fulfilled through the one-body J_1 and two-body J_2 Jastrow terms, written as

$$J_1(\mathbf{R}) = \prod_{ia} \exp[-(2Z_a)^{3/4} u((2Z_a)^{1/4} r_{ia})] \tag{7}$$

and

$$J_2(\mathbf{R}) = \prod_{i<j} \exp[\, u(r_{ij}))], \tag{8}$$

where $\mathbf{R}$ is an all-electron configuration, i and j are electron indices, and a is a nuclear index. The ion centers have effective charge Z_a and the function $u(x)$ satisfies the electron-ion and electron-electron cusp conditions between unlike-spin particles with $u(0) = 0$ and $u'(0) = \frac{1}{2}$. Here, $u(r) \equiv (F/2)(1 - e^{-r/F})$, where F is an optimizable parameter. In Eq. (7), the argument of u is multiplied by $(2Z_a)^{1/4}$ in order to satisfy the random phase approximation behavior at large r_{ia}.[29]

A distinguishing feature of the JAGP with respect to the simple SJ wave function is the presence of electron-electron-ion and electron-ion-electron-ion terms, conventionally referred to as three- and four-body Jastrow factors. In the JAGP wave function, they are written as the exponential of a pairing function like the one in Eq. (6), namely,

$$J_{34}(\mathbf{R}) = \prod_{ij} \exp\left[-\sum_{ablm} g_{lm}^{ab} \chi_{al}(\mathbf{r}_i^\uparrow) \chi_{bm}(\mathbf{r}_j^\downarrow)\right]. \tag{9}$$

Here g_{lm}^{ab} are optimizable parameters and $l(m)$ is an index for single-particle Gaussian orbitals χ_{al} centered on nucleus $a(b)$. The three- and four-body Jastrow terms provide for electron correlations substantially beyond the largely cusp related one- and two-body terms and are able to describe subtle effects such as van der Waals forces.[30] However, Eq. (9) does not include the three-body cusp conditions recently derived by Fournias et al.,[31] which can improve the quality of the nodes of the JAGP wave function described here. The effect of the three-body cusp conditions in the energy optimization and nodal structure is presently under investigation.

The JAGP bases are constructed from Gaussians as follows. For the AGP basis we use a contracted $(6s6p)/[2s2p]$ for the carbon atoms, $(2s2p)/[1s1p]$ for molecular hydrogen's atomic sites and a single s Gaussian for benzene's hydrogen sites. For the Jastrow we use an uncontracted $(3s2p)$ basis for the carbon sites, an uncontracted $(1s1p)$ for molecular hydrogen's atomic sites, and a single s Gaussian for benzene's hydrogen sites. Each atomic basis in the Jastrow includes a constant that generates additional electron-ion terms when multiplied by other orbitals χ_{bm} in Eq. (9). For benzene's hydrogen constituents we used just a single s Gaussian both in the AGP and Jastrow geminals, since they are not supposed to play a key role in the interaction between the hydrogen molecule and the benzene ring. The fully optimized benzene basis included in the JAGP wave function gives a quite good variational energy for aromatic rings.[15] An analysis of the basis used for the hydrogen molecule will be given in Sec. III A.

B. Methods

In setting up our Hamiltonian, we use the Born–Oppeheimer approximation, a Hartree–Fock norm conserving soft pseudopotential for the He core of carbon,[32] and the bare Coulomb potential for hydrogen and electron-electron

interactions. Our procedure is to start with a trial wave function which includes variational parameters (see Sec. II A for the forms employed in this work). We proceed to optimize its energy and variance at the VMC level using minimization methods suitable for the particular form.[11,12,33,34] The resulting analytic wave function is projected to the FN ground state using DMC methods[35,36] recently developed to yield a stable simulation and an upper bound of the ground state energy even for nonlocal pseudopotentials.

As we mentioned above, we use the full electron-nucleus Hamiltonian except for the carbon core which is replaced by a pseudopotential. This leads to a better statistics due to a narrower energy scale, a reduction in the number of optimization parameters, a more stable optimization of our JAGP wave function,[12] and a larger DMC time step needed for convergence, which results in a cheaper computational cost of the simulation. On the other hand, its drawback is that part of the fully local Coulomb potential is replaced by a nonlocal pseudopotential $V_{nonlocal}$ that is angular momentum dependent. Within the VMC framework the corresponding angular integration of the nonlocal potential remains possible since the wave function is known analytically. However, problems arise in the FN DMC because the FN ground state is given only by a stochastic sampling. A partial solution is the localization approximation, where the trial (or guiding) wave function Ψ_G is used to approximate the projected ground state so that the nonlocal pseudopotential terms can be evaluated.[10] However, numerical instabilities are introduced and the projected energy is no longer a variational upper bound of the original nonlocal FN Hamiltonian.

Our FN DMC calculations are done with either continuous or lattice regularized (LRDMC) moves both of which utilize a common means of treating the nonlocal part of the pseudopotential. In contrast to the localization approximation, we use a breakup[36] of the nonlocal potential that localizes the positive matrix elements into the branching term while treating the negative matrix elements as a nonlocal diffusion operator sampled via a *heat bath* scheme.[36] The positive and negative terms are defined by

$$V^{\pm}_{\mathbf{R}',\mathbf{R}} = 1/2(V_{\mathbf{R}',\mathbf{R}} \pm |V_{\mathbf{R}',\mathbf{R}}|), \tag{10}$$

where

$$V_{\mathbf{R}',\mathbf{R}} = \frac{\Psi_G(\mathbf{R}')}{\Psi_G(\mathbf{R})} \langle \mathbf{R}' | V_{nonlocal} | \mathbf{R} \rangle, \tag{11}$$

and $\mathbf{R}, \mathbf{R}'$ are all-electron configurations on a quadrature mesh with one electron rotated around a pseudoion.[37] The breakup corresponds to an effective Hamiltonian H^{eff}, defined as

$$\begin{aligned} H^{eff}_{\mathbf{R},\mathbf{R}} &= K + V^{eff}(\mathbf{R}), \\ H^{eff}_{\mathbf{R}',\mathbf{R}} &= \langle \mathbf{R}' | V_{nonlocal} | \mathbf{R} \rangle, \quad \text{if } V_{\mathbf{R}',\mathbf{R}} < 0, \end{aligned} \tag{12}$$

with the modified local potential $V^{eff}(\mathbf{R}) = V_{loc}(\mathbf{R}) + \Sigma_{\mathbf{R}'} V^{+}_{\mathbf{R}',\mathbf{R}}$ that includes the *sign flip* terms. The FN ground state energy of the Hamiltonian in Eq. (12) is a variational upper bound of the original nonlocal Hamiltonian.[38] Furthermore, DMC stability is improved substantially compared to

the local approximation, where the most attractive parts of the localized pseudopotential can result in a walker population "blow up." Moving the negative part of the localization into a diffusion-like term causes the walkers to be driven away from such regions.

The main difference between the DMC Hamiltonian reported in Eq. (12) and the LRDMC is the kinetic operator K, which is replaced by a discretized K^a in the LRDMC approach, and treated on the same footing as $V_{nonlocal}$. K^a is a linear combination of two discrete operators with incommensurate lattice spaces a and a' ($a' = \nu a$, with ν an irrational number >1), namely,

$$K^a = -\frac{\eta}{2}(\Delta^{a,p} + \Delta^{a',1-p}), \tag{13}$$

where $\Delta^{a,p}$ is the discretized Laplacian with mesh a and weighting function p (see Refs. 34 and 12) and $\eta = 1 + \mu a^2$ is a prefactor with the parameter μ tunable to improve the efficiency of the diffusion process. Working with two incommensurate meshes helps to sample densely the continuous space by performing discrete moves of lengths a and a'. The finest hop samples more likely regions near atomic centers while the coarser one samples more often valence regions, the result being an efficient sampling of the overall configuration space. The difference between the continuous and discretized local kinetic energies is added to $V^{eff}(\mathbf{R})$, resulting in a mesh dependent potential,

$$V^a(\mathbf{R}) = V^{eff}(\mathbf{R}) + \left[\frac{(K - K^a)\Psi_G}{\Psi_G} \right](\mathbf{R}). \tag{14}$$

The consequence is a faster convergence of the energies in the $a \to 0$ extrapolation. In spite of the discretization of K [Eq. (13)] and the redefinition of V^{eff} [Eq. (14)], the LRDMC method is equivalent to the continuous space FN DMC with Hamiltonian in Eq. (12). Indeed, in the limit of small mesh sizes a and a', the discretized Hamiltonian H^a approaches the continuous H. The usual DMC Trotter breakup results in a time step error while the LRDMC paradigm results in a space step error, but both share the same upper bound property in the zero-time-step zero-lattice-space limit and converge to the same projected FN energy.[36]

Our SJ calculations were done using continuous space DMC with QMCPACK.[39] This code provides many features that make it easy to work with SJ wave functions. The LRDMC method, available in the TURBORVB,[40] has been applied to the JAGP wave function after a full optimization of its parameters. We used two optimization procedures. For the SJ work we employed the method of conjugate gradients (CGs) introduced by Hestenes and Stiefel[33] in 1952. This is a first-derivative method that finds the minimum of a cost function (in our case a linear combination of the variance and the energy), in a number of steps significantly smaller than the standard steepest descent method, because for a quadratic cost function it converges in a finite number of iterations, at most equal to the dimension of the vector space.[41,42] We optimized ten parameters of the Jastrow functions but used the same VTZ basis set at all separations. However, the statistical noise inherent in the QMC framework limits the

applicability of our CG implementation to systems involving not too many parameters, such as our SJ optimization. The JAGP optimization, on the other hand, involves a large number (~1000) of parameters, mainly coming from the λ_{ab}^{lm} [Eq. (6)] and g_{lm}^{ab} [Eq. (9)] matrices in the AGP and Jastrow geminal expansions over the atomic basis set. Therefore, an optimization technique robust under stochastic conditions is required. For this we used the stochastic reconfiguration (SR) method recently introduced by one of us[34] in conjunction with subsequent improvements,[11–14] including Hessian acceleration which is explained in Ref. 12, that have been shown to be very efficient to minimize the variational energy.

III. RESULTS

In this section we present results for hydrogen-benzene binding where the hydrogen molecule is oriented along the C_6 symmetry axis of the benzene molecule. Previous studies[8,9] found this configuration the most stable. Here, we do not take into account other possible orientations because our goal is to check the accuracy of different QMC wave functions and provide benchmarks for the lowest energy configuration. In order to resolve its potential energy surface, we consider the system at different molecular center-of-mass separations R. In our QMC calculations we have kept the geometry of each molecule fixed and close to its experimental structure.[43] We checked the effect of relaxing the geometries at the MP2 level and found an energy lowering on the order of microhartrees, whose effect is completely negligible in this case.

We emphasize that our QMC results do not include any corrections for basis set superposition error (BSSE). The binding curves and the final results for binding energies and distances are determined from directly calculated energy differences with the largest computed distance ($R = 15$ bohrs) taken as the zero energy reference. On the other hand the results presented here for DFT calculations have included a correction; we quantified the BSSE using PBE-DFT using the VTZ basis with added diffuse functions that are described in Sec. II A. In that case, the BSSE correction was 0.39 mhartree, roughly half the corrected binding energy which was found to be 0.79 mhartree at 6.45 bohrs.

The BSSE arises due to an incomplete basis set, and it is important to point out that the magnitude of the effect is different in the various QMC methods. In VMC the BSSEs are due to the finite basis and the consequences can be understood using the same arguments as for other variational methods. In the present work, the error is greatly reduced because we fully optimize the AGP and Jastrow bases along with all exponents at each R. In DMC methods, the basis is complete in the continuous configuration space, and the only inherent limitation is the FN approximation. Since we use nodes determined with a finite basis there is necessarily some error due to superposition; however, the effect upon the final DMC energies is greatly reduced since the diffusion algorithm leads to the best possible estimate of the energy within the nodal constraint.

Our results support this analysis and justify the conclusion that the BSSE errors are negligible for the QMC calcu-

lations reported here. The good agreement between the VMC and DMC JAGP results, presented in Sec. III A, highlights that the basis set superposition bias is not relevant (smaller than the statistical error of ~0.2 mhartree) for the fully optimized basis set used in the JAGP wave function, while the agreement between the projected SJ and JAGP energies, shown in Sec. III B, suggests that the FN bias is negligible.

It should be noted that there is another possible kind of basis error that can also occur due to restrictions in the trial wave functions. If the trial functions are zero (or very small) in regions of configuration space, then the DMC calculation may not properly sample the full space. This can happen particularly in the tails of the wave functions, and it is important to ensure that the basis includes sufficiently diffuse functions so that the tail regions are properly sampled. This is especially relevant for the calculation of weak binding energies with small overlap in the tails of the molecular wave functions.

Finally, we note that a further possible source error arises through use of a pseudopotential to replace the effect of the cores of the carbon atoms. In the previous section we discussed the procedures for treating the errors due to the use of nonlocal pseudopotentials in DMC. Errors due to these and other effects of the pseudopotential should cancel in the energy differences because the effects occur mainly in the core region, which changes very little as a function of the distance between the molecules for any separation relevant to the present problem.

A. Jastrow correlated antisymmetric geminal power

We optimized the variational JAGP wave function described in Sec. II A by means of the most recent version of the SR energy minimization with Hessian acceleration,[12] as mentioned in Sec. II B. Although the basis set used here is quite compact, it turns out that the variational energies are very accurate, as we optimize also the exponents of both the determinantal and Jastrow part. For instance, the basis set for the hydrogen molecule is a $(2s2p)/[1s1p]$ Gaussian in the AGP expansion, while it is an uncontracted $(1s1p)$ Gaussian plus a constant in the Jastrow geminal. In spite of this small basis set, the variational energy of an isolated H_2 molecule is $-1.174\,077(29)$ Ha, very close the exact result $(-1.174\,475$ Ha$)$.[44] The second Gaussian in the s and p contractions of the hydrogen AGP is fairly diffuse, their exponents ranging from 0.05 to 0.1, as the distance R between the benzene molecule and the hydrogen dimer shrinks from 15 to 6 bohrs.

We found that the inclusion of the diffuse orbitals in the basis set of the hydrogen molecule is crucial for the hydrogen-benzene binding, both at the VMC and LRDMC level. On the other hand, some Gaussians related to the contracted p orbital of the benzene ring become more delocalized in the binding region. This is reasonable because the interaction is supposedly driven by the resonance between the carbon p_z and molecular hydrogen s components of the total wave function. Therefore, the minimal basis set should include diffuse orbitals on both sides. We would like to stress that the extension of those diffuse orbitals is not determined

TABLE I. LRDMC binding energy ($E(R=6)-E(R=15)$) dependence on mesh size a. The energies are reported in mhartrees and the lengths are in bohrs.

a	$E_{binding}$
0.125	1.53(24)
0.25	1.57(19)
0.5	2.07(23)

TABLE II. Fitting parameters of the Morse function [see Eq. (15)] which minimize the χ^2 of the JAGP-LRDMC and SJ-DMC data sets. Their error is computed by means of a Bayesian analysis (Ref. 45) based on the statistical distribution of the FN energy points. The energies are reported in mhartrees and the lengths are in bohrs.

	JAGP	SJ
a	0.56(7)	0.66(9)
E_b	1.53(12)	1.43(16)
R_0	6.33(15)	6.31(21)

a priori, but is found by optimizing the wave function with the necessary variational freedom.

After a full optimization of the variational wave function at several distances ($R=5$, 5.5, 6, 7, 8, 10, and 15 bohrs) we carried out VMC and LRDMC simulations to study the properties of the system, in terms of energetics and charge density distribution. The LRDMC kinetic parameter in Eq. (13) which optimizes the lattice space extrapolation is $\mu=3.2$, that allows one to work with a quite large (and highly efficient) mesh size ($a=0.25$ a.u.). Properly setting the parameters of the LRDMC effective Hamiltonian is crucial in order to speed up the simulation, and so be able to resolve the small binding energy of this system. To check the convergence of our LRDMC energies with respect to the mesh size, we computed $E(R=6)-E(R=15)$ for $a=0.125$, 0.25, and 0.5, as reported in Table I. It is apparent that the energy differences are converged within the error bar of 0.25 mhartree in the lattice space range taken into account. It is therefore accurate to work with $a=0.25$.

The results of our calculations of the VMC and LRDMC dispersion curves are presented in Fig. 1(a), which shows the energy as a function of distance R relative to the value at $R=15$ bohrs for each of the methods. There is excellent agreement between the two curves, with a difference that is
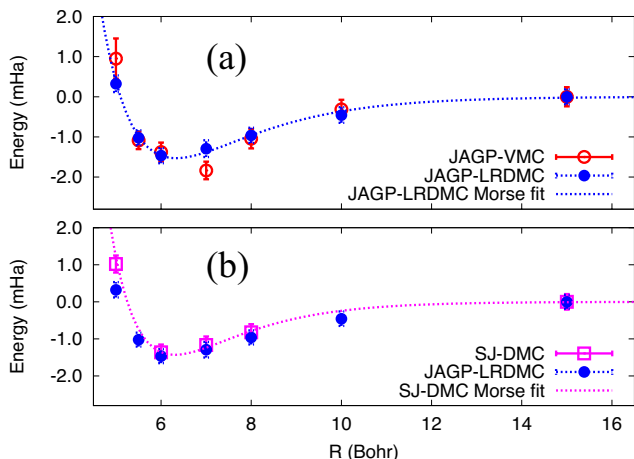
less than 0.18 mhartree for most points. Of course, the diffusion calculation leads to a lower total energy than the variational calculation in every case, but the agreement of the two methods for the energy difference supports the idea that our results are accurate and the calculated binding energy is close to the exact value.

In order to extract the values for the equilibrium distance R_0 and the binding energy E_b, we fitted our LRDMC points with the Morse function,

$$V(R) = E_\infty + E_b[e^{-2a(R-R_0)} - 2e^{-a(R-R_0)}], \tag{15}$$

where a is related to the zero point motion of the effective one dimensional potential $V(R)$ and E_∞ is chosen to be $E(R=15)$, i.e., the zero of energy. This choice is motivated by the fact that the overlap of the wave function in between the two fragments is negligible at that distance. Beyond that point the variation in $V(R)$ up to infinity is much smaller than the statistical accuracy of our points. We estimated the error on the fitting parameters by carrying out a Bayesian analysis of the fit, in a way similar to what described in Ref. 45. Our results are 6.33(15) bohrs for the equilibrium distance and 1.53(12) mhartree for the binding energy, as reported in Table II.

B. SJ trial function

At this point, it is interesting to make a comparison with a simple SJ wave function to determine whether the use of the JAGP is necessary to get the correct dispersion energy out of the FN projection. The Slater part is constructed of PBE-DFT single-body orbitals and a simple Jastrow as described in Sec. II A. Again, we chose to use a very simple Jastrow factor because our goal was only to improve DMC efficiency as opposed to obtaining a well converged binding curve at the VMC level. The Jastrow factor was optimized within the VMC framework using the conjugate gradient method,[33] as explained in Sec. II B. While the SJ variational energy is quite poor, its quality is not directly reflective of the DMC energy, which depends only on the nodes of the trial wave function.

We found that the DFT nodes are very good by carrying out DMC simulations with the nonlocal scheme described in Sec. II B. Our projection was done in time steps of $\tau=0.01$ which we found to be converged as reported in Table III. Remarkably, the DMC-SJ energies are in very good agreement with the LRDMC-JAGP data points [see Fig. 1(b)]. Indeed, the SJ fitting parameters of the Morse dispersion curve [Eq. (15)], such as binding energy, equilibrium dis-

FIG. 1. (Color online) QMC results for the dispersion energy of the hydrogen-benzene bond as a function of intermolecular distance R with zero energy difference taken at $R=15$ bohrs. (a) compares variational and the diffusion results using the correlated geminal wave function, labeled JAGP-VMC and JAGP-LRDMC. (b) compares diffusion results using two types of trial functions, the JAGP [the same as in (a)], and the SJ function labeled SJ-DMC. Morse fits of the diffusion data for the two wave functions are also plotted as continuous curves. The close agreement of all three results is a strong evidence that the binding curve is accurate and the analytic JAGP function [defined in Eqs. (5)–(9)] is a reliable representation of the fully correlated many-body valence wave function.

TABLE III. DMC binding energy ($E(R=6)-E(R=15)$) dependence on time step τ. The energy extrapolated for $r \to 0$ is within one error bar from the point at $r=0.01$. Therefore, we chose $r=0.01$ as the time step for all our DMC simulations. The energies are reported in mhartrees and the time steps are in Ha^{-1}.

τ	E_b
0.01	1.38(19)
0.02	0.93(19)
0.04	0.64(15)

FIG. 2. (Color online) Results for hydrogen-benzene binding as a function of intermolecular distance R using four theoretical methods. The JAGP-LRDMC data and Morse fit with zero binding energy taken at R =15 bohrs is shown in solid black (deepest). The PBE-DFT counterpoise corrected result using the VTZ basis (Ref. 21) plus diffuse functions from the aug-cc-pVTZ basis (Ref. 22) is shown in solid light black (green). The Crowell and Brown (Ref. 7) empirical potential (shallowest) that takes into account the bond asymmetry of the sp^2 hybridized carbon atom is shown in dotted dark gray (blue). The Mattera et al. (Ref. 6) empirical potential that seeks to reproduce the hydrogen bound states over graphite by a much simpler model is shown in dotted light gray (red).

tance, and curvature, differ from the JAGP ones by less than one error bar (Table II). This consistency between different trial wave functions signals that the FN bias is negligible and the results are well converged. Moreover, in addition to the nodes of the PBE wave function being good, the PBE binding energy is underestimated only by a factor of 2 with respect to our best value. It is notable that the PBE functional performs quite well, even though it does not include any VdW contribution. In the case of a pure VdW bond, the PBE result should be much poorer, as already pointed out by Hamel and Côté.[8] This is suggestive of a more complex binding mechanism which goes beyond the standard physisorption. We will focus on this point in Sec. V.

IV. COMPARISON TO OTHER WORK

The hydrogen-benzene system has been the subject of several theoretical works, whereas to our knowledge no direct study of this system has been carried out on the experimental side. Hydrogen adsorbed on MOFs, where benzene-like structures serve as ligands, has been studied by Rosi et al.[3] who performed inelastic neutron scattering (INS) measurements. The INS data could be related to the rotational states of hydrogen adsorbed over benzene. However, the binding sites in the MOF structure are not known with certainty, and thus it is hard to find a one-to-one correspondence between the experiment and the isolated hydrogen-benzene compound.

Given the lack of direct experimental data for this system, we compare our results with those from empirical models that are often used to estimate complex system properties, such as the hydrogen storage capabilities of carbon nanotubes and fullerene nanocages.[46,47] Here we consider two empirical models, both derived from experiments of hydrogen molecules scattered on graphite surface, carried out by Mattera et al.[6] To reproduce their data, they proposed a simple model interaction between the carbon atoms and the hydrogen dimer which depends only on the distance from the graphite layers by assuming lateral average. This model was improved later by Crowell and Brown,[7] who constructed an empirical potential based not only on the experimental scattering data but also on the polarization constants built in the VdW (6,12) potential. Their model assumes both a radial and angular dependence, which takes into account the sp^2 hybridization asymmetry of carbon atoms in graphitic and aromatic compounds. We applied these potentials to the hydrogen-benzene system by summing the terms for the six carbons taking into account distance and, for the Crowell
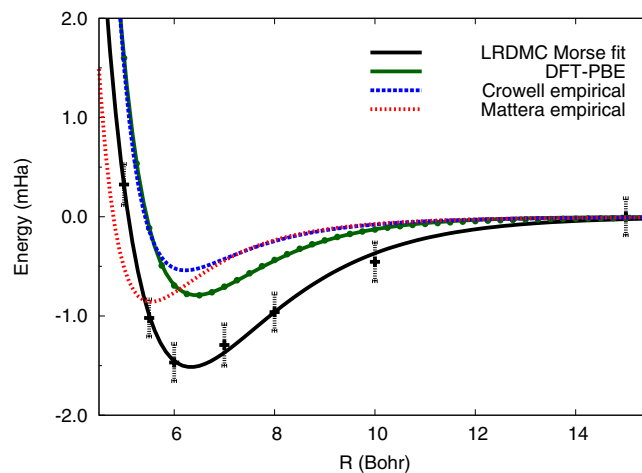
potential, the angle the hydrogen-carbon interaction makes with the benzene C_6 axis. Both empirical potentials significantly underbind the system, roughly by factors of 2 and 3, respectively, when compared to the JAGP LRDMC results (see Fig. 2). More precisely, Mattera's interaction gives a binding energy of 0.86 mhartree at 5.6 bohrs, while Crowell's gives a minimum of 0.54 mhartree at 6.2 bohrs.

Hamel and Côté[8] calculated the dispersion curves using DFT with the local density and generalized gradient approximations (LDA and GGA), where the GGA is implemented in the PBE density functional.[19,20] Their calculations used a plane wave basis with a 60 hartree cutoff. They found that the DFT-LDA gives the strongest binding (3.30 mhartree), while the DFT-PBE binding is much weaker (0.69 mhartree). This is consistent with the general overbinding of LDA and underbinding of PBE. It is also well known that DFT is not a favorable method for systems where VdW forces play an important role,[50] in those cases, MP2 and CCSD(T) can be applied with more reliability. Hamel and Côté also calculated binding curves using those theories. They found MP2/6-311+G(2df,2p) binding of 1.58 mhartree and CCSD(T)/6-31+G(d,p) binding of 0.65 mhartree.

Perhaps the most careful and accurate MP2 and CCSD(T) calculations were done by Hübner et al.[9] In order to resolve the weak interaction between hydrogen and benzene, high accuracy is required, and so a large basis set is needed to reduce both BSSE and incompleteness error which are a significant fractions of the binding energy (the BSSE was found to be as much as ~25% of the final estimated binding). On the other hand, the use of a larger basis set is limited by a poorer scaling of the calculations, particularly at the CCSD(T) level of theory, which is the most expensive. In their work, Hübner et al. optimized the binding distance us-
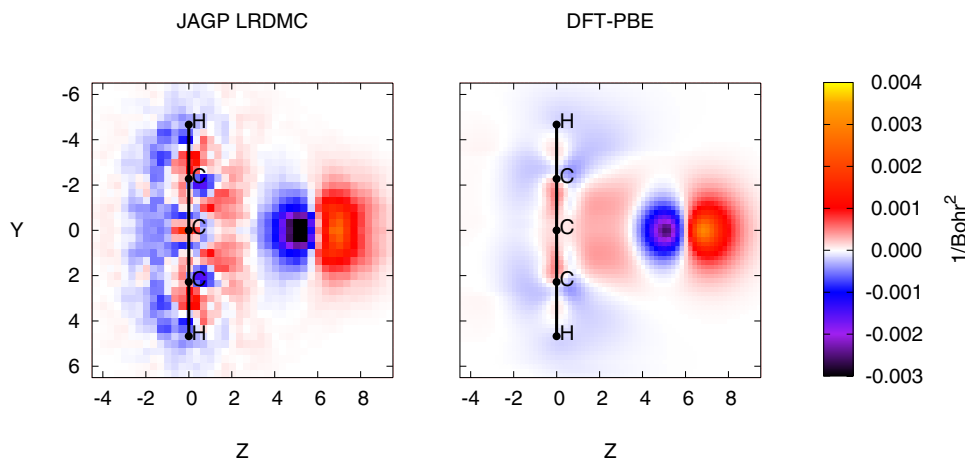
FIG. 3. (Color online) Contour plots of the difference in projected electronic charge per unit area between hydrogen-benzene separated by 6 bohrs and the isolated hydrogen and benzene using JAGP-LRDMC and PBE-DFT. The *x*-axis has been integrated over so that the charge per unit area has been projected into the *yz*-plane. (Left) The areal charge density difference is a mixed estimate of LRDMC calculations with a JAGP trial wave function. (Right) Computation is done within the PBE-DFT framework using the VTZ basis (Ref. 21) plus diffuse functions from the aug-cc-pVTZ basis (Ref. 22).

ing MP2 with the TZVPP basis. They found a center-of-mass distance of 5.80 bohrs and a binding energy of 1.47 mhartree. This geometry was then used for further MP2 and CCSD(T) calculations. The CCSD(T) method with the same TZVPP basis gives 1.17 mhartree, while the MP2 theory was pushed up to a aug-cc-pVQZ′ basis to give a binding of 1.83 mhartree, a significant increase from the TZ-VPP basis. At this point, it is possible to estimate the true binding energy by correcting the best MP2 energy with the CCSD(T)-MP2 difference obtained at the TZVPP level. This gives a value of ~1.5 mhartree, remarkably close to the JAGP LRDMC binding of 1.53 ± 0.12 mhartree, found in this work.

V. ANALYSIS OF THE BONDING

In order to investigate more deeply the physics of hydrogen adsorbed on benzene, we study the induced difference in electronic density at the equilibrium bond distance with respect to the separated fragments. For this study we compare our best DMC results to the density functional calculation using the PBE functional. The QMC densities are calculated from the optimized correlated geminal (JAGP) as a mixed estimator, which is an accurate representation of the DMC results since the diffusion calculation leads to only small changes (within the error bar) from the VMC density. The contour plot in Fig. 3 shows the difference in the calculated electron density at the separation $R=6$ bohrs. Here, the electron density of the isolated molecules has been subtracted from the combined system so that the change in charge distribution due to bonding is apparent. In this figure the benzene ring lies in the *xy* plane at $z=0$ and the hydrogen molecule is oriented along the *z* axis, with its center of mass at $z=6$ bohrs. The two dimensional plot in the *yz* plane is generated by integrating the density distribution over the *x* coordinate. As one can see, the hydrogen molecule is polarized by the electronic repulsion with the benzene cloud, which pushes the electrons to the opposite side of the molecule, leading to a static dipole moment on the hydrogen. On the other hand, the density redistribution in the benzene is non-

trivial and shows patches of charge accumulation and depletion. To catch the net effect of this redistribution, we integrated the density also over the *y* coordinate and obtained an effective linear density profile, plotted in Fig. 4. Here, it is apparent that the overall effect on the benzene is the formation of another effective dipole moment, oriented to the same direction as the static dipole moment on the hydrogen molecule, which lowers the electrostatic energy. Notice that in Fig. 4 we have plotted separately the VMC and the LRDMC mixed estimate of the densities. The close agreement supports our conclusion that the VMC wave function is very accurate not only for the energy but also for other properties such as the density.

At large distances the attractive interaction is due to VdW dispersive forces, which is included in the Monte Carlo
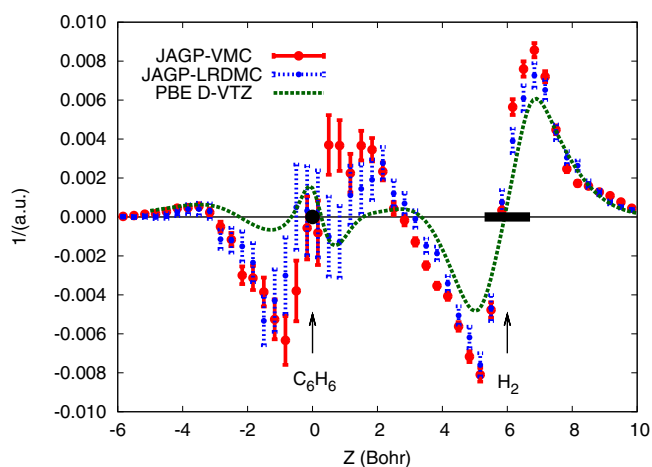

FIG. 4. (Color online) Difference in linear electronic charge density between hydrogen-benzene separated by 6 bohrs and the isolated hydrogen and benzene using three theories. The *x*- and *y*-axes have been integrated over so that the charge per unit length has been projected into the *z*-axis. The data with solid red error bars show the induced charge changes using the analytic JAGP wave function at the VMC level. The data with dotted blue error bars show the mixed estimate of the density given by the LRDMC projection of the JAGP trial wave function. The dotted green curve without error bars shows the PBE-DFT result using the VTZ basis (Ref. 21) plus diffuse functions from the aug-cc-pVTZ basis (Ref. 22).

calculations. At short distances the interaction is repulsive due to overlap of the closed shells, which would lead to density displaced outward on both the hydrogen and benzene, i.e., opposite dipoles on the two molecules. However, Figs. 3 and 4 show that the hydrogen-benzene bond is not a pure VdW interaction, since in the binding region also electrostatic effects come in with the onset of dipolar interactions that lower the charge repulsion. For comparison, density differences calculated using the PBE density functional are also shown in Figs. 3 and 4 at the separation $R=6$ bohrs. Of course, the PBE functional does not include VdW interactions so that the binding decreases too rapidly at large distance as shown in Fig. 2. Nevertheless, near the equilibrium distance the density is similar to the QMC result but with smaller magnitude of the change in density, which is consistent with the fact that the PBE functional underbinds the system. It is well known that GGA functionals such as PBE tend to underbind because they favor systems with larger gradients, whereas LDA tends to overbind molecules and solids since it favors more homogeneous systems.[42] Recent work by Langreth *et al.*[51,52] has led to improved functionals including van der Waals interactions; however, they have not been considered here.

VI. SUMMARY AND CONCLUSIONS

We have presented VMC and DMC results for the adsorption of hydrogen on a benzene ring and compared them with previous work. We used two types of variational correlated wave functions, a SJ function with DFT-PBE optimized single-body orbitals and a JAGP function fully optimized at the VMC level by means of the SR energy minimization. In this work, we have shown strong evidence that our results are very accurate since we have found essentially the same results in three independent QMC calculations: one JAGP-VMC variational simulation with no FN error and two DMC simulations based on different trial wave functions (JAGP and SJ) with possibly different nodes. The agreement among our three calculations is within $\simeq 0.2$ mhartree, which is mainly due to statistical accuracy on the QMC energies and gives an upper bound for the magnitude of underlying errors, such as the basis set incompleteness and the FN bias.

Our best estimate for the binding energy is 1.53(12) mhartree at an equilibrium distance of 6.33(15) bohrs, obtained by using the LRDMC method with the nodes of the JAGP wave function. Our result agrees well with the conclusion of Hübner *et al.*[9] who used MP2 and CCSD(T) methods and estimated the binding to be ~1.5 mhartree based on extrapolation which accounts for basis set and level of theory. The resulting binding energy is two to three times larger than those given by empirical potentials[6,7] and DFT-PBE calculations which are often employed in more complex systems, suggesting that their results could be substantially affected by this lack of accuracy. It would be interesting to extend the present work, by studying the transferability of such empirical potentials on other aromatic and graphitic structures.

We found that the JAGP wave function provides a very accurate dispersion curve for this system already at the varia-tional level. This result is remarkable because we were able to derive a compact analytic form which can be used for accurate determination of properties other than the energy by means of the VMC method with no sign problem. The JAGP wave function captures the resonating valence bonds of benzene in its geminal construction as well as the van der Waals interaction through many-body correlations in the Jastrow factor, as shown in previous work on benzene dimer.[12] The basis for both the AGP and Jastrow geminals is of compact Gaussian form that does not go beyond p-orbitals, but includes diffuse orbitals with optimized exponents.

By means of the DMC method, we also studied the hydrogen-benzene problem using a more conventional SJ wave function. The single-body orbitals included in the Slater determinant were derived using the GGA PBE density functional with a VTZ Gaussian basis[21] modified to include diffuse functions from the aug-cc-pVTZ basis,[22] as discussed in Sec. II A, which is essential for an unbiased DMC sampling of long-range VdW effects. The Slater basis set is roughly four times larger than its JAGP counterpart whereas the Jastrow factor is of minimal form, satisfying cusp conditions and improving computational scaling. Our findings suggest that for this particular problem the geminal form is not essential to get an accurate DMC energy and can be replaced with a Slater determinant and DFT optimized orbitals in the DMC calculations. While the JAGP uses a more compact basis, the SR optimization involves a large number of parameters coming from the Jastrow and AGP geminals expanded on atomic orbitals. This makes the SJ wave function with DFT-PBE single-body orbitals more desirable for DMC calculations in larger related systems.

Finally, we examined how the electronic density of the isolated molecules changes due to the bonding. The change in density, displayed in Figs. 3 and 4, shows that near the equilibrium distance there is the formation of static dipoles that can lower the electrostatic energy, indicating a bonding mechanism beyond VdW. Density functional calculations using the PBE functional lead to similar density profiles but with smaller magnitude, in agreement with the well known underbinding tendency of that functional. This means that the interaction between hydrogen and benzene is not a pure VdW effect, since it can be partially captured by a DFT-PBE formalism which does not include dispersive interactions. This also clarifies why the DFT-PBE nodes of the SJ wave functions are very good and equivalent to the JAGP nodes to predict the correct binding energy at the DMC level.

To conclude, we have reported on a detailed analysis of the hydrogen adsorption over molecular benzene by QMC methods, which are shown to be very accurate and reliable to predict the energetics and other physical properties of the system. This framework is therefore promising to study hydrogen interacting with graphitic or other aromatic compounds, particularly important for the hydrogen storage problem.

ACKNOWLEDGMENTS

[1] A. C. Dillon *et al.*, Nature (London) **386**, 377 (1997).

[2] Y. Zhao, Y.-H. Kim, A. C. Dillon, M. J. Heben, and S. B. Zhang, Phys. Rev. Lett. **94**, 155504 (2005).

[3] N. L. Rosi *et al.*, Science **300**, 1127 (2003); http://www.sciencemag.org/cgi/reprint/300/5622/1127.pdf.

[4] R. C. Lochan and M. Head-Gordon, Phys. Chem. Chem. Phys. **8**, 1357 (2006).

[5] M. Fichtner, Adv. Eng. Mater. **7**, 443 (2005).

[6] L. Mattera *et al.*, Surf. Sci. **93**, 515 (1980).

[7] A. D. Crowell and J. S. Brown, Surf. Sci. **123**, 296 (1982).

[8] S. Hamel and M. Côté, J. Chem. Phys. **121**, 12618 (2004).

[9] O. Hübner, A. Glöss, M. Fichtner, and W. Klopper, J. Phys. Chem. A **108**, 3019 (2004).

[10] W. M. C. Foulkes, L. Mitas, R. J. Needs, and G. Rajagopal, Rev. Mod. Phys. **73**, 33 (2001).

[11] C. J. Umrigar, J. Toulouse, C. Filippi, S. Sorella, and R. G. Hennig, Phys. Rev. Lett. **98**, 110201 (2007).

[12] S. Sorella, M. Casula, and D. Rocca, J. Chem. Phys. **127**, 014105 (2007).

[13] S. Sorella, Phys. Rev. B **71**, 241103 (2005).

[14] C. J. Umrigar and C. Filippi, Phys. Rev. Lett. **94**, 150201 (2005).

[15] M. Casula, C. Attaccalite, and S. Sorella, J. Chem. Phys. **121**, 7110 (2004).

[16] M. Casula and S. Sorella, J. Chem. Phys. **119**, 6500 (2003).

[17] C. Diedrich, A. Lüchow, and S. Grimme, J. Chem. Phys. **123**, 184106 (2005).

[18] M. J. Frisch *et al.*, GAUSSIAN 03, Revision C.02, Gaussian, Inc., Wallingford, CT, 2004.

[19] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).

[20] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **78**, 1396 (1997).

[21] M. Burkatzki, C. Filippi, and M. Dolg, J. Chem. Phys. **126**, 234105 (2007).

[22] R. A. Kendall, J. Thorm, H. Dunning, and R. J. Harrison, J. Chem. Phys. **96**, 6796 (1992).

[23] L. K. Wagner and L. Mitas, J. Chem. Phys. **126**, 034105 (2007).

[24] J. Bardeen, L. N. Cooper, and J. R. Schrieffer, Phys. Rev. **106**, 162 (1957).

[25] J. R. Schrieffer, *Theory of Superconductivity* (Benjamin, New York, 1964).

[26] L. C. Pauling, *The Nature of the Chemical Bond*, 3rd ed. (Cornell University Press, Ithaca, 1960).

[27] M. Casula and S. Sorella (unpublished).

[28] J. P. Bouchaud, A. Georges, and C. Lhuillier, J. Phys. (Paris) **49**, 553 (1988).

[29] M. Holzmann, D. M. Ceperley, C. Pierleoni, and K. Esler, Phys. Rev. E **68**, 046707 (2003).

[30] M. Casula, Ph.D. thesis, Scuola Internazionale Superiore di Studi Avanzati (SISSA), 2005.

[31] S. Fournais, M. Hoffmann-Ostenhof, T. Hoffmann-Ostenhof, and T. Ø. Sørensen, Commun. Math. Phys. **255**, 183 (2005).

[32] The pseudopotentials we used are norm-conserving Hartree–Fock generated by E. Shirley's code with the construction by D. Vanderbilt, Phys. Rev. B 32, 8412 (1985).

[33] M. R. Hestenes and E. Stiefel, J. Res. Natl. Bur. Stand. **49**, 409 (1952).

[34] S. Sorella, Phys. Rev. B **64**, 024512 (2001).

[35] M. Casula, C. Filippi, and S. Sorella, Phys. Rev. Lett. **95**, 100201 (2005).

[36] M. Casula, Phys. Rev. B **74**, 161102 (2006).

[37] S. Fahy, X. W. Wang, and S. G. Louie, Phys. Rev. B **42**, 3503 (1990).

[38] D. F. B. ten Haaf, H. J. M. van Bemmel, J. M. J. van Leeuwen, W. van Saarloos, and D. M. Ceperley, Phys. Rev. B **51**, 13039 (1995).

[39] J. Kim *et al.*, QMCPACK, Materials Computation Center, http://www.mcc.uiuc.edu.

[40] S. Sorella *et al.*, TURBORVB, SISSA, http://www.sissa.it.

[41] M. C. Payne, M. P. Teter, D. C. Allan, T. A. Arias, and J. D. Joannopoulos, Rev. Mod. Phys. **64**, 1045 (1992).

[42] R. M. Martin, *Electronic Structure: Basic Theory and Practical Methods* (Cambridge University Press, Cambridge, 2004).

[43] The actual bond lengths used in this work are: C–C=2.63 bohrs, H–C =2.04 bohrs, and H–H=1.40 bohrs. They are close to the best experimental and theoretical values (Refs. 48 and 49).

[44] W. Kolos and L. Wolniewicz, J. Chem. Phys. **49**, 404 (1968).

[45] L. K. Wagner, J. Phys.: Condens. Matter **19**, 343201 (2007).

[46] O. V. Pupysheva, A. A. Farajian, and B. I. Yakobson, Surf. Sci. **8**, 767 (2008).

[47] Q. Wang and J. K. Johnson, J. Chem. Phys. **110**, 577 (1999).

[48] D. Feller and D. Dixon, J. Phys. Chem. A **104**, 3048 (2000).

[49] K. P. Huber and G. Herzberg, *Molecular Spectra and Molecular Structure IV. Constants of Diatomic Molecules* (Van Nostrand Reinhold, New York, 1976).

[50] Note that in our work, we used the single-body orbitals from the PBE-DFT calculation in the SJ wave function. The DMC energies depend only on the accuracy of the nodes of the many-body wave function. The DMC calculation includes VdW attraction and other terms and the result is independent of the errors in the PBE functional for the energy.

[51] M. Dion, H. Rydberg, E. Schröder, D. C. Langreth, and B. I. Lundqvist, Phys. Rev. Lett. **92**, 246401 (2004).

[52] T. Thonhauser *et al.*, Phys. Rev. B **76**, 125112 (2007).