JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

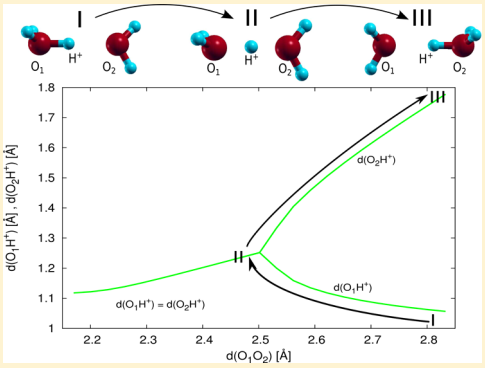
Quantum Monte Carlo Study of the Protonated Water Dimer

Mario Dagrada,*[,†] Michele Casula,[†] Antonino M. Saitta,[†] Sandro Sorella,[‡] and Francesco Mauri[†]

[†]Institut de Minéralogie, de Physique des Matériaux, et de Cosmochimie (IMPMC), Sorbonne Universités - UPMC Paris 06, UMR CNRS 7590, Muséum National d'Histoire Naturelle, IRD UMR 206, 4 Place Jussieu, 75005 Paris, France

[‡]International School for Advanced Studies (SISSA) Via Beirut 2,4 34014 Trieste, Italy and INFM Democritos National Simulation Center, Trieste, Italy

ABSTRACT: We report an extensive theoretical study of the protonated water dimer $H_5O_2^+$ (Zundel ion) by means of the highly correlated variational Monte Carlo and lattice regularized Monte Carlo approaches. This system represents the simplest model for proton transfer (PT), and a correct description of its properties is essential in order to understand the PT mechanism in more complex aqueous systems. Our Jastrow correlated AGP wave function ensures an accurate treatment of electron correlation. By exploiting the advantage of contracting the primitive basis set over atomic hybrid orbitals, we are able to limit dramatically the number of variational parameters with a systematic control on the numerical precision, a crucial ingredient in order to simulate larger systems. For both energetics and geometrical properties, our QMC results are found to be in excellent agreement with state-of-the-art coupled cluster CCSD(T) techniques. A comparison with density functional theory in the PBE approximation points to the crucial role of electron correlation for a correct description of the PT in the dimer. We prove that the QMC framework used in this work is able to resolve the tiny energy differences (~0.3 kcal/mol) and structural variations involved in proton transfer reactions. Our approach combines these features and a favorable N^4 scaling with the number of particles which paves the way to the simulation of more realistic PT models. A test calculation on a larger protonated water cluster is carried out. The QMC approach used here represents a promising candidate to provide the first high-level *ab initio* description of PT in water.



1. INTRODUCTION

Water is a key element for life. Nevertheless, its properties arising from its unique structure are not fully understood yet. From a theoretical point of view, liquid water has been investigated relying on molecular dynamics techniques such as the empirical force field methods[1−3] and *ab initio* molecular dynamics (AIMD).[4−9] For the latter technique, the atomic force fields are constructed from potential energy surfaces (PESs) directly derived from first-principles electronic structure calculations.

Calculating a PES for liquid water has been a longstanding challenge in the scientific community. The main problems are related to the description of breaking and formation of hydrogen bonds, weak dispersive interactions, and polarization effects in the network of polar molecules.

Moreover, the simulation of liquids requires a large system size, and therefore a good scalability is fundamental. Hence, the technique which has been most frequently employed is density functional theory (DFT), since it combines scalability and flexibility. However, DFT-based AIMD poorly reproduces properties such as the melting point temperature[10] and the oxygen−oxygen radial distribution function.[11]

Another major problem not yet elucidated by AIMD, is the description of proton transfer (PT) in liquid water, the so-called Grotthuss mechanism.[12−14] This process is extremely important in a wide range of biological systems, and it influences many dynamical processes in material science, biochemistry, and bioenergetics. For example, PT is one of the main mechanisms for charge transport through cell membranes;[15,16] it plays an important role in several enzymatic reactions and in photosynthesis.[17,18] Moreover, the proton has the role of mediator and promoter of many acid−base reactions in solution.[19−21]

Despite its importance, an accurate quantitative explanation of the PT underlying mechanism has not been reached yet; theoretical investigations still fail to reproduce its basic properties, such as the activation barrier and the proton diffusion coefficient.[22−26] Nevertheless, in recent years AIMD simulations have given many insights into the proton transfer mechanism.[25,27−30] In particular, they have revealed the complex nature of the process, which involves both structural and dynamical changes of the hydrogen-bonded network.[31]

Two main structural models have been proposed to explain proton hydration: Eigen et al.[32] considered the $H_9O_4^+$ complex as the PT building block in which a hydronium H_3O^+ is strongly bound to three water molecules; Zundel and Metzger[33] proposed a different reaction which involves the simpler $H_5O_2^+$ complex (the Zundel ion or protonated water dimer) as the PT core between two H_2O molecules.

Received: December 12, 2013
Published: March 26, 2014

   dx.doi.org/10.1021/ct401077x | *J. Chem. Theory Comput.* 2014, 10, 1980−1993

It has been demonstrated[27] that these two models occur only as limiting or ideal structures. Nevertheless, they constitute a perfect testing ground for the most advanced simulation techniques. In particular, $H_5O_2^+$ is the smallest system in which an excess proton is shared between water molecules; due to its simplicity, many studies with different techniques have been carried out to elucidate its structure, energetics, and vibrational properties.[24,34−40] Very recently, an accurate PES has been produced[41] with high precision coupled cluster techniques including single, double, and perturbative triple excitations (CCSD(T)).

In the present paper, we report on an extensive study of the protonated water dimer using a quantum Monte Carlo (QMC) approach based on a highly correlated variational wave function for energy and geometry calculations.

This approach has a very good scalability with the number of particles with respect to other post-HF methods. We demonstrate that it ensures the precision of ∼0.3 kcal/mol required for an accurate study of the system. These appealing targets are achieved by our Jastrow correlated antisymmetrized geminal power (JAGP) wave function, developed on an efficient contracted single particle basis set. While the JAGP provides the necessary ingredients to treat both dynamical and static correlations, the contracted basis set, based on atomic hybrid orbitals, reduces dramatically the computational cost to optimize the JAGP wave function. Thanks to these promising features, we hope that it will pave the way to a complete understanding of the physics of PT and other properties of liquid water.

The paper is organized as follows. In section 2, we introduce the computational techniques used in this work. In particular, in subsection 2.2, we describe the variational ansatz and basis set, which is the optimal compromise between accuracy and computational cost for water clusters. In section 3.1, we study the basis set and wave function convergence for the single H_2O molecule and show that we have an accuracy comparable to the latest QMC studies. Section 3.2 is devoted to the main results of our paper on the protonated water dimer. We perform structural calculations aimed at finding the ground state geometry of $H_5O_2^+$ at the JAGP level of theory. In order to resolve a zero-temperature potential energy landscape, we choose the distance between the two oxygens as a natural reaction coordinate. Along this path, we investigate how the symmetric "Zundel" configuration—namely the one with the proton evenly shared between the two oxygens—evolves to a symmetry-broken geometry, with the excess proton localized on one H_2O. We believe that this crossover is critical to understanding the physics of PT in larger water clusters. The evaluation of proton transfer static barriers at fixed oxygen−oxygen ($\overline{OO}$) distances shows that our QMC approach has a global accuracy close to the recent CCSD(T) calculations performed by Huang et al.,[41] with a discrepancy of ∼0.3 kcal/mol obtained by diffusion Monte Carlo (DMC) on the variational Monte Carlo (VMC) geometries. In subsection 3.2.5, we study the dependence of the QMC computational cost with the system size by simulating a more realistic PT model, namely the cluster composed by six water molecules and one excess proton. A comparison with analogous CCSD(T) calculations is also reported. We summarize the results and draw our conclusions in section 4.

2. COMPUTATIONAL METHODS

2.1. Density Functional Theory Calculations. In this section, we describe DFT-based calculations on the Zundel ion. They are carried out to obtain optimized starting geometries and wave functions for further quantum Monte Carlo calculations.

The DFT geometry was obtained using Car−Parrinello dynamics performed in a plane wave basis set with the QUANTUM ESPRESSO[42] suite of codes. Technical details of this calculation are reported in subsection 2.1.1.

Once the ground-state DFT geometry was found, the Kohn−Sham (KS) orbitals, used as a starting input for the QMC wave function, are determined with the DFT code of the TURBORVB package,[43] which employs a localized Gaussian basis set within the standard LDA approximation, as described in subsection 2.1.2.

2.1.1. Plane Wave DFT. With the aim of finding the optimal DFT geometry for further QMC structural relaxation (see section 2.2.2), we use the Perdew−Burke−Ernzerhof[44] (PBE) version of the generalized gradient approximation for the exchange-correlation functional, which has been widely used to describe properties of liquid water.[45] Core electrons are taken into account using a norm conserving Trouiller−Martins pseudopotential.[46] The Kohn−Sham orbitals are expanded over plane waves up to a cutoff of 37.5 Ha for the wave function and 150 Ha for the charge density in a periodic box of 30 Bohr radii (a_0) per side.

The zero-temperature geometry relaxation is performed with damped dynamics within the Car−Parrinello (CP) approach.[47] This method allows a very small force convergence threshold of $\sim 10^{-5}$ Ha/a_0 on each atomic component at the end of the relaxation. The potential energy surface of the dimer is very flat[38] and is characterized by energy variations smaller than 1 kcal/mol. Therefore, this accuracy on atomic forces is essential to providing a reliable DFT description of the system.

2.1.2. Gaussian DFT. We perform Gaussian DFT calculations with both pseudopotentials and full Coulomb electron−ion potentials. The single-particle Kohn−Sham (KS) orbitals $\{\phi_i(\mathbf{r})\}$ are used later on as an optimized starting guess for the QMC wave function, as discussed in section 2.2.1.

The one-body DFT wave function is expanded over a primitive basis of *Gaussian type orbitals* (GTOs) centered on the atom a, defined up to a normalization constant as

$$\psi^{\mathrm{GTO}}_{a,(l,m,n)}(\mathbf{r}) \propto |\mathbf{r}-\mathbf{R}_a|^l e^{-\zeta_{l,n}|\mathbf{r}-\mathbf{R}_a|^2} Z_{l,m}(\Omega_{\mathbf{r}-\mathbf{R}_a}) \quad (1)$$

with $m \in [-l,l]$ and $n \in [1,n_l]$, where $\{l,m\}$ are angular momentum quantum numbers, n_l identifies the number of Gaussians for each angular momentum shell, $Z_{l,m}(\Omega)$ are the spherical harmonics, and $\mathbf{r}$ and $\mathbf{R}_a$ are the electron and ion positions, respectively. Our DFT basis set is *even tempered*; namely for each angular momentum l, the exponents of the Gaussians are expressed as a power series: $\zeta_{l,n} = \alpha_l\beta_l^{n-1}$ for $n = 1, ..., n_l$. This allows a reduction of the number of parameters, as for each shell l the full $\zeta_{l,n}$ series is fixed just by three values: α_l, β_l, and n_l. Denoting the set of GTO quantum numbers as $\mu = \{l,m,n\}$, the DFT orbitals $\{\phi_i(\mathbf{r})\}$ expanded over the primitive basis set read

$$\phi_i(\mathbf{r}) = \sum_a^{N_{\mathrm{atoms}}} \sum_\mu^{N_{\mathrm{basis}}(a)} c^i_{a,\mu}\psi^{\mathrm{GTO}}_{a,\mu}(\mathbf{r}) \quad (2)$$

dx.doi.org/10.1021/ct401077x | J. Chem. Theory Comput. 2014, 10, 1980−1993

where N_{atoms} is the number of atoms of the system and $N_{basis}(a)$ is the size of the GTO set of the atom a. Thus, the global DFT wave function can be expressed as the determinant of the matrix $\{\hat{\phi}_i(\mathbf{x}_j)\}$:

$$\Psi_{DFT}(\mathbf{x}_1, ..., \mathbf{x}_N) = \det \hat{\phi}_i(\mathbf{x}_j) \tag{3}$$

where $\{\mathbf{x}_i = (\mathbf{r}_i\sigma_i)\}$ are the spatial and spin coordinates of the N electrons in the system, and $\hat{\phi}_i(\mathbf{x}_j) = \phi_i(\mathbf{r}_j)\delta_{\sigma_i\sigma_i}$ are spin orbitals with definite spin projection σ_i, δ being the Kronecker delta.

The *ab initio* Hamiltonian of a quantum system is characterized by the divergence of the Coulomb potential at the electron−electron and electron−ion coalescence points. When the potential energy becomes infinite, the wave function of the system must have a cusp to resolve the corresponding singularity in the Hamiltonian.[48] These singularities represent a serious issue when an approximate wave function ansatz is employed to solve the Schrödinger equation of a quantum system.

Within a DFT framework, the divergence created by electron−electron coalescence is intrinsically solved by the mean-field description of the system, which maps the interacting Hamiltonian onto one with independent electrons. Nevertheless, the electron−ion coalescence still represents a problem. One way to resolve the latter divergence is to employ a smooth pseudopotential. The pseudointeraction is chosen such that it eliminates the need of electron−ion cusps in the wave function, leading to a faster basis set convergence.

In this work, we employ an alternative route to solve this issue. We introduce an additional factor to the DFT wave function ansatz which automatically satisfies the electron−ion cusp conditions and cancels out the corresponding divergences of the Hamiltonian.

We rewrite eq 3 as

$$\tilde{\Psi}_{DFT}(\mathbf{x}_1, ..., \mathbf{x}_N) = J_1(\mathbf{r}_1, ..., \mathbf{r}_N) \det \hat{\phi}_i(\mathbf{x}_j) \tag{4}$$

where the function J_1 is called the one-body Jastrow factor, borrowed from the QMC notation. It is defined by means of a simple function $u(\mathbf{r})$ which contains only one variational parameter b:

$$u(|\mathbf{r} - \mathbf{R}|) = \frac{1 - e^{-b|\mathbf{r}-\mathbf{R}|}}{2b} \tag{5}$$

The one-body Jastrow factor J_1 then reads

$$J_1 = \exp(-\sum_{i}^{N}\sum_{j}^{N_{atoms}} (2Z_j)^{3/4} u((2Z_j)^{1/4}|\mathbf{r}_i - \mathbf{R}_j|)) \tag{6}$$

where $\mathbf{R}_j$ are the atomic coordinates corresponding to the atomic number Z_j. In the function u, the multiplicative factor $(2Z_j)^{3/4}$ is set by the electron−ion cusp conditions, while the scaling length $(2Z_j)^{1/4}$ deals with the large distance behavior given by the random phase approximation.[49]

Equation 4 has the typical functional form of the Jastrow-single determinant (JSD) variational ansatz for quantum Monte Carlo calculations. Indeed the one-body Jastrow of eq 6 was introduced by Sorella et al.[50] to provide a solution of the electron−ion cusp conditions within a QMC framework. In contrast with QMC, our $\tilde{\Psi}_{DFT}$ presents only the one-body Jastrow factor since in DFT the many-body electron−electron interaction is solved within a mean-field description.

By letting the wave function fulfill the electron−ion cusp conditions, we are able to perform calculations considering the bare Coulomb interaction for both oxygen and hydrogen atoms (*all-electron* calculations). Furthermore, at variance with conventional DFT calculations based on Gaussian localized orbitals, in our approach all overlap matrix elements involved are computed as numerical integrals over a uniform mesh. Such integrals, due to the presence of the Jastrow factor, converge very rapidly with the number of points in the mesh and thus can be evaluated with a reasonable computational effort.[51]

With the aim of reducing the computational cost, calculations are carried out also by replacing the 1s core electrons of oxygen with the pseudopotential approximation, while the hydrogen atoms are always treated with the full Coulomb potential.

The oxygen pseudopotential that we used in both DFT and QMC calculations is the Burkatzki−Filippi−Dolg (BFD) pseudopotential, introduced by Burkatzki et al.[52] It is built by means of an energy-consistent approach in order to reproduce the valence all-electron excitation energies for a number of different atomic configurations computed at the scalar relativistic Hartree−Fock (HF) level.

Gaussian DFT calculations are performed with the local density approximation (LDA) for the exchange-correlation functional. The optimal basis sets which reach the DFT complete basis set limit (CBS) are O(8s8p5d) H(5s3p) and O(9s9p6d) H(5s5p) for the pseudopotential and all-electron cases, respectively. Expanding the wave function over these basis sets, the variational parameters (α_l, β_l, n_l for each angular momentum l and the one-body Jastrow parameter b) are optimized at the DFT level of theory, by total energy minimization.

2.2. Variational Quantum Monte Carlo Calculations.

Quantum Monte Carlo refers to several numerical techniques for electronic structure calculations of quantum systems. The interest of the scientific community in these methods has grown considerably over the past three decades, and they have been successfully applied to highly correlated electronic systems, from small- to medium-size quantum chemistry systems.

The QMC techniques used in this work are wave-function-based; hence they aim at finding the many-body wave function in a representation as close as possible to the true ground state of the system. The QMC trial wave function has to satisfy several criteria: it needs to be continuous, square-integrable, and computable in a finite amount of time. Despite these restrictions, its analytic form is very flexible, and thus correlations can be introduced in a compact and efficient way.

Another appealing feature of QMC methods is their scaling with the number of particles. Independently of the technique, QMC provides an intrinsic scaling of $\sim N^3-N^4$, whereas state-of-the-art coupled cluster single and double (CCSD) is $\propto N^6$ and coupled cluster single, double, and triple (CCSD(T)) methods are $\propto N^7$. Therefore, QMC allows the treatment of larger molecular clusters with respect to other highly correlated methods.

Major efforts have been made to reduce the multiplicative scaling prefactor of QMC techniques, i.e., the cost of a single-point calculation. The research has followed two main paths. From one side, more compact trial wave functions have been developed. One of the most promising functions is the Jastrow-antisymmetrized geminal power (JAGP) used in this work. A description of the wave function ansatz is given in subsection 2.2.1.

A second way to decrease the scaling prefactor relies on improving the quantum Monte Carlo estimators. QMC total

energy calculations are in general much more efficient than those of other observables, due to the *zero-variance* property of the energy estimator. Generally speaking, estimators of other important observables, such as the charge density, do not possess this property. Some other bare estimators, such as the one for the atomic forces, could have infinite variance. Thus, they affect the overall computational efficiency of the QMC calculation. In particular, in recent years, a major improvement has been achieved in the development of an efficient estimator for nuclear forces.[53−56]

In the present work, we use a version of this estimator based on the space-warp coordinate transformation[57] and implemented with an exact infinitesimal differentiation method.[58] This allows a single sample calculation of the ionic forces with a computational effort on the order of N^3; this scaling is comparable to that of the total energy. Moreover, the infinite variance of the space-warped force estimator in the proximity of the nodal surface of the wave function is solved, as explained by Sorella and co-workers.[59,60] The latter reference provides also a nice review on the latest progress in the QMC nuclear forces evaluation. Therefore, by means of this state-of-the-art scheme, we are able to perform efficient structural optimizations at the QMC level of theory also for relatively large atoms such as oxygen.

2.2.1. Wave Function Ansatz. The typical Quantum Monte Carlo wave function is made of a symmetric bosonic factor (*Jastrow factor*) applied to an antisymmetric Fermionic part (*determinantal part*):

$$\Psi(\mathbf{x}_1, ..., \mathbf{x}_N) = J(\mathbf{r}_1, ..., \mathbf{r}_N) \times \psi_D(\mathbf{x}_1, ..., \mathbf{x}_N) \tag{7}$$

where the set $\{\mathbf{x}_i = (\mathbf{r}_i, \sigma_i)\}$ represents spatial and spin coordinates of the electrons, as in eq 3.

The Jastrow factor is a function of the electron−electron separation. It is meant to reproduce correlations of the true many-body wave function well; it contributes mainly to the description of dynamical correlation effects arising from charge fluctuations. Furthermore, it has been shown[50] that the Jastrow factor is particularly suitable in the treatment of van der Waals intermolecular forces. They also play an important role in the physics of the protonated water dimer; therefore an efficient parametrization of the Jastrow is essential. We split this factor into one-body, two-body, and three/four-body terms ($J = J_1J_2J_3$).

The one-body factor accounts for electron−ion interactions, and it has been already introduced in eqs 5 and 6. The two-body term deals with the electron−electron interactions. In complete analogy to the one-body Jastrow, it is parametrized by a simple function of the electron−electron separation:

$$u(|\mathbf{r}_i - \mathbf{r}_j|) = \frac{1 - e^{-b|\mathbf{r}_i - \mathbf{r}_j|}}{2b} \tag{8}$$

Then, it becomes

$$J_2(\mathbf{r}_1, ..., \mathbf{r}_N) = \exp\left(\sum_{i<j}^{N} u(|\mathbf{r}_i - \mathbf{r}_j|)\right) \tag{9}$$

J_1 and J_2 satisfy the Kato cusp conditions,[48] correcting the divergence of the Coulomb potential energy at electron−ion and electron−electron coalescence points, respectively.

The function u rapidly decays to a constant value as the electron−ion and electron−electron distances increase; thus the large distance behavior of correlations is described by the three/four-body Jastrow term J_3. It deals with electron−electron−ion (if $a = b$) and electron−ion−electron−ion (in the case $a \neq b$) correlation effects and represents an essential part of our wave function since it takes also into account inhomogeneous correlation; therefore we parametrize it in a richer way than the other Jastrow terms:

$$J_3(\mathbf{r}_1, ..., \mathbf{r}_N) = \exp\left(\sum_{i<j}^{N} \Phi_J(\mathbf{r}_i, \mathbf{r}_j)\right)$$

$$\Phi_J(\mathbf{r}, \mathbf{r}') = \sum_{a,b}^{N_{\text{atoms}}} \sum_{\mu,\nu}^{N_{\text{Jbasis}}} g_{\mu,\nu}^{a,b} \psi_{a,\mu}^{J}(\mathbf{r}) \psi_{b,\nu}^{J}(\mathbf{r}')$$

where N_{Jbasis} represents the number of GTOs of the primitive Jastrow basis for each atom. The Jastrow uncontracted orbitals $\psi_{a,\mu}^{J}(\mathbf{r})$ have the same form as the primitive GTO basis set $\psi_{a,\mu}^{\text{GTO}}(\mathbf{r})$ for the determinantal part in eq 1. The Jastrow primitive basis used in this work is O(3s2p1d) H(2s1p).

The choice of the Fermionic part of the wave function is more delicate. In the case of spin unpolarized systems ($N_\uparrow = N_\downarrow$) of N electrons, the Fermionic part is expressed as an antisymmetrized product of the *geminals* or pairing functions $\Phi(\mathbf{x}_i,\mathbf{x}_j)$ of the system:

$$\Psi_D(\mathbf{x}_1...\mathbf{x}_N) = \hat{A}[\Phi(\mathbf{x}_1, \mathbf{x}_2), ..., \Phi(\mathbf{x}_{N-1}, \mathbf{x}_N)] \tag{10}$$

The geminals are antisymmetric functions of two electron coordinates written as a product of a spatial symmetric part and a spin singlet:

$$\Phi(\mathbf{x}_i, \mathbf{x}_j) = \phi(\mathbf{r}_i, \mathbf{r}_j) \frac{\delta(\sigma_i, \uparrow)\delta(\sigma_j, \downarrow) - \delta(\sigma_i, \downarrow)\delta(\sigma_j, \uparrow)}{\sqrt{2}}$$

The simplest but computationally most expensive expansion of the geminal is over the *uncontracted atomic orbitals*, and it reads

$$\phi(\mathbf{r}_i, \mathbf{r}_j) = \sum_{a,b}^{N_{\text{atoms}}} \sum_{\mu,\nu}^{N_{\text{basis}}} \lambda_{\mu,\nu}^{a,b} \psi_{a,\mu}^{\text{GTO}}(\mathbf{r}_i) \psi_{b,\nu}^{\text{GTO}}(\mathbf{r}_j) \tag{11}$$

where the orbitals $\psi_{a,\mu}^{\text{GTO}}(\mathbf{r})$ have the form reported in eq 1 with $\mu = (l,m,n)$. N_{basis} represents the number of primitive GTOs per atom. The introduction of the many-body Jastrow factor J in the total QMC wave function (eq 7) allows a reduction of the primitive basis set size without a loss of accuracy. This reduction has several advantages within the QMC framework. At first, it obviously decreases the computational effort for the wave function calculation. Furthermore, it allows a more robust energy minimization. Indeed, given the fact that QMC energy derivatives are noisy, the use of a more compact primitive basis set reduces its redundancy and helps to find the global minimum more quickly, since the number of the effective directions in the Hilbert space is smaller. The size-reduced primitive basis employed for the determinantal part is O(5s5p2d) H(4s2p) and O(6s6p2d) H(3s2p) for pseudopotential and all-electron calculations, respectively. The number of variational parameters expanded over the GTO atomic orbitals is $P^{\text{AGP}} \propto N_{\text{totalbasis}} \times (N_{\text{totalbasis}} + 1)$, where $N_{\text{totalbasis}}$ is the total number of GTOs of the whole system.

In this work, we use two distinct functional forms for ψ_D. They are distinguished by the number of nonzero eigenvalues—i.e., the rank—of the matrix $\{\lambda_{\mu,\nu}^{a,b}\}$ in eq 11.

The first one is commonly referred to as single determinant (SD), and it is recovered when the $\{\lambda_{\mu,\nu}^{a,b}\}$ matrix has the lowest possible rank compatible with the number of electrons

in the system, namely when its rank is equal to $N/2$. It can be shown that it is equivalent to the Slater determinant ansatz for HF calculations. In this case, the general expansion (eq 11), after diagonalizing the matrix $\{\lambda_{\mu,\nu}^{a,b}\}$, is written as

$$\phi(\mathbf{r}_i, \mathbf{r}_j) = \sum_{k}^{N/2} \lambda_k \psi_k^{\mathrm{MO}}(\mathbf{r}_i)\psi_k^{\mathrm{MO}}(\mathbf{r}_j) \tag{12}$$

where $\psi_k^{\mathrm{MO}}(\mathbf{r}) = \sum_{a,\mu} c_{a,\mu}^{\ k} \psi_{a,\mu}^{\mathrm{GTO}}(\mathbf{r}_j)$ are *molecular orbitals* (MO). In this work, we obtain them starting from the optimized Kohn−Sham orbitals of Gaussian DFT calculations discussed in section 2.1.2. The number of variational parameters in this case is $P^{\mathrm{SD}} \propto N/2 \times N_{\mathrm{totalbasis}}$, where $N_{\mathrm{totalbasis}}$ is the number of linear coefficients for each MO, equal to the total number of GTOs. Close to the CBS limit $N_{\mathrm{total\ basis}} \gg N$, the SD wave function ensures a significant reduction in the number of parameters with respect to the fully uncontracted AGP expansion in eq 10. By multiplying the SD by the Jastrow factor, one obtains the JSD wave function, which is optimized simultaneously in both the J and SD parts.

The SD represents a particular limit of the function in eq 10. By letting the rank of $\{\lambda_{\mu,\nu}^{a,b}\}$ become greater than $N/2$, one introduces multiconfigurational states and goes beyond the single determinant representation. In this more general case, the determinantal wave function is called Antisymmetrized Geminal Power (AGP).

The AGP is the particle-conserving version of the Bardeen−Cooper−Schrieffer (BCS) wave function, and it accounts for static correlations in the system, namely those deriving from nearly degenerate energy levels. Together with the Jastrow factor, it forms the Jastrow-Antisymmetrized Geminal Power (JAGP) wave function, which represents a practical implementation of the resonating valence bonds introduced by Linus Pauling for chemical systems.[61] The JAGP wave function has been proven to be accurate in describing a wide range of strongly correlated systems,[50,62−65] and it is the second variational form we tested in our work after the JSD.

In order to benefit from the AGP ansatz without paying the cost of dealing with an excessively large number of variational parameters, we develop the AGP expansion on a contracted basis set. We use atomic hybrid orbitals, also employed by Zen et al.[60] as contractions of the primitive GTOs. One of the features of the AGP ansatz (without Jastrow) is that the geminal is directly related to the one-body density matrix of the system. To find the most effective local (atomic) basis set, we project the full one-body density matrix on its local atomic constituents, retaining only the values $\lambda_{\mu,\nu}^{a,b}$ with $a = b$ in the expansion of eq 11 and setting the other terms to zero. By diagonalizing the projected one-body density matrix, we obtain a set of local natural orbitals for each atom. These atomic orbitals are hybrid (i.e., linear combination of primitive Gaussians not restricted to a given angular momentum shell), as they describe the hybridization arising from the atomic embedding, which breaks the spherical symmetry around the nucleus. We call them *atomic hybrid orbitals*. Thanks to this construction, for each atom we keep not only the information on the local electronic structure due to the nuclear charge but also the information on the nuclear embedding in the compound, namely the impact of the environment on its electronic structure.

The atomic hybrid orbitals are analogous to the well-known natural orbitals,[66,67] which have already been successfully applied to several areas of electronic structure calculations.[68−70]

Despite their similarity, our orbitals differ from the natural orbitals since they do not require orthonormality and they are generated via simple energy optimization.

The initial AGP used to fix the hybrid basis can be determined by DFT calculations (in which case, the rank of the AGP is $N/2$ as it comes directly from an SD wave function), or by a previously optimized JAGP wave function (in which case the Jastrow factor is disregarded and only the determinantal part is taken in the one-body density matrix). Once the hybrid basis set is chosen and the AGP is expanded upon it, the geminal is

$$\phi(\mathbf{r}_i, \mathbf{r}_j) = \sum_{a,b}^{N_{\mathrm{atoms}}} \sum_{\alpha,\beta}^{N_{\mathrm{hyb}}} \tilde{\lambda}_{\alpha,\beta}^{a,b} \psi_{a,\alpha}^{\mathrm{hyb}}(\mathbf{r}_i)\psi_{b,\beta}^{\mathrm{hyb}}(\mathbf{r}_j) \tag{13}$$

where $\psi_{a,\alpha}^{\mathrm{hyb}}(\mathbf{r}) = \sum_{\mu} c_{\mu}^{\ \alpha} \psi_{a,\mu}^{\mathrm{GTO}}(\mathbf{r})$ are the contracted atomic hybrid orbitals, and $\mu = (l,m,n)$ are as before. N_{hyb} is the number of atomic hybrid orbitals required for an accurate description of each atom. After their first determination by DFT or by previous JAGP calculations, they are further optimized in the QMC energy minimization.

The number of variational parameters describing the wave function is $P^{\mathrm{hyb}} \propto N_{\mathrm{totalhyb}}^{\ 2} + N_{\mathrm{basis}} \times N_{\mathrm{totalhyb}}$, where N_{totalhyb} is the total number of hybrid Gaussians in the whole system. Since $N \approx N_{\mathrm{totalhyb}} \ll N_{\mathrm{totalbasis}}$, the hybrid orbitals represent the optimal basis set, which reduces significantly the total number of parameters in the correlated AGP framework. In order to find the best value of N_{hyb} for oxygen and hydrogen, we carried out a detailed analysis of the variational energy versus the hybrid basis size for the single water molecule, reported in section 3.1.

2.2.2. Quantum Monte Carlo Methods. All QMC calculations of this work have been carried out with the TurboRVB suite of codes.[43]

All of the variational parameters of the JSD and JAGP wave functions have been optimized by means of the stochastic reconfiguration method with a Hessian accelerator, also called the "linear method."[50,71−74] The forces on the parameters and ionic positions always have a finite variance thanks to a reweighting scheme for finite systems introduced by Zen et al.,[60] which cures the variance divergence around the nodes of the wave function.

A systematic way of improving the quality of the VMC ansatz is to perform lattice regularized diffusion Monte Carlo (LRDMC) calculations,[75,76] a projective technique which filters out the high-energy components and yields a more accurate total energy. In the case of nonlocal pseudopotentials, as in the case of oxygen, LRDMC goes beyond the locality approximation, by always providing a variational upper bound of the true ground state energy.

In the present paper, we report both VMC and LRDMC calculations of the protonated water dimer. The BFD energy-consistent pseudopotential described in section 2.1.2 has been employed to replace the oxygen-core 1s electrons. We also carried out some all-electron calculations for testing purposes. Statistical error bars are kept smaller than 0.1 kcal/mol in all final QMC calculations.

3. RESULTS AND DISCUSSION

We report the results of our QMC study of the protonated water dimer. This section is organized as follows.

Table 1. Pseudopotential Calculations[a]

wave function ansatz	VMC energies: energy E_x (Ha)	VMC energies: variance (Ha^2)	VMC energies: $E_x - E_{JSD}$ (mHa)	number of parameters: $\lambda_{\alpha,\beta}^{a,b}$	number of parameters: det orbitals	number of parameters: total
JSD (uncontracted orbitals)	−17.24821(7)	0.2655(6)	0.0	682	18	895[b]
JAGP (hybrid orbitals: 4O 1H)	−17.25013(8)	0.2635(12)	−1.91(11)	21	158	374
JAGP (hybrid orbitals: 4O 5H)	−17.25183(6)	0.2510(6)	−3.62(10)	105	238	538
JAGP (hybrid orbitals: 8O 2H)	−17.25267(7)	0.2426(18)	−4.46(10)	78	298	571
JAGP (hybrid orbitals: 8O 5H)	−17.25302(6)	0.2412(34)	−4.89(10)	171	358	724
JAGP (uncontracted orbitals)	−17.25389(6)	0.2296(5)	−5.68(10)	682	18	895

[a]VMC energies and number of parameters for the QMC wave functions used in this work of the single water molecule, fixed here at the experimental geometry. The BFD pseudopotential has been used for O. The primitive Gaussian basis set is O(5s,5p,2d) H(4s,2p) for the determinant. The number of total parameters varies depending on the type of contractions used in the determinantal part. The Jastrow functional form has been kept fixed and developed on a primitive Gaussian basis set of O(3s,2p,1d) H(2s,1p). This gives a number of 213 Jastrow parameters, divided into 184 $g_{\mu,\nu}^{a,b}$, nine ζ Gaussian exponents in the uncontracted basis, one two-body homogeneous Jastrow factor coefficient, and one parameter for the analogous one-body part. The other parameters come from the determinant and are reported in the last set of columns. [b]Here, the number of parameters is the same as that in the JAGP wave function since in the JSD ansatz we rewrite the corresponding geminal (of rank $N/2$) on the uncontracted basis in order to optimize the MOs, as explained by Marchi et al.[64]

Table 2. As Table 1, but for All-Electron Calculations[a]

wave function ansatz	VMC energies: energy E_x (Ha)	VMC energies: variance (H^2)	VMC energies: $E_x - E_{JSD}$ (mHa)	number of parameters: $\lambda_{\alpha,\beta}^{a,b}$	number of parameters: det orbitals	number of parameters: total
JSD (uncontracted orbitals)	−76.40025(8)	1.412(3)	0.0	1383	19	1819
JAGP (hybrid orbitals: 9O 2H)	−76.40504(9)	1.399(6)	−4.79(12)	91	361	870
JAGP (uncontracted orbitals)	−76.40660(7)	1.374(3)	−6.35(11)	1383	19	1819

[a]Here, we report the VMC energies and number of parameters for the all-electron QMC wave functions of the single water molecule, taken at the QMC relaxed geometry. The geometries are reported in Table 3. The primitive Gaussian basis set is O(6s,6p,2d) H(4s,2p) for the determinant. The number of total parameters varies depending on the type of contractions used in the determinantal part. The Jastrow functional form has been kept fixed and developed on a primitive Gaussian basis set of O(3s,3p,1d) H(2s,1p). This gives a number of 418 Jastrow parameters, divided into 406 $g_{\mu,\nu}^{a,b}$, 10 ζ Gaussian exponents in the uncontracted basis, one two-body homogeneous Jastrow factor coefficient, and one parameter for the analogous one-body part. The other parameters come from the determinant and are reported in the last set of columns.

The first part (section 3.1) is devoted to benchmark computations on the water molecule which aim at demonstrating the quality of our QMC ansatz.

The second part (section 3.2) presents the QMC study of the protonated water dimer. Different levels of theory are compared with QMC calculations for both geometry and energetics in order to assess the accuracy of our approach in the study of proton transfer systems.

3.1. Benchmark Calculations on the H_2O Molecule. The single water molecule has been the subject of many numerical studies based on several quantum chemistry methods. It is essential to have a good description of its structural and electronic properties in order to tackle the study of larger water clusters, as the intramolecular degrees of freedom will significantly affect the intermolecular environment due to the large water dipole moment and the strong directionality of the H bond.

Here, we report our pseudopotential and all-electron calculations on the water molecule with different wave function types and basis sets, with the aim at choosing the best compromise between accuracy and efficiency in order to transfer the most convenient ansatz to the Zundel complex and eventually to larger water clusters.

3.1.1. Pseudopotential Calculations. As reported in section 2, the BFD pseudopotential[52] has been used for oxygen, while the two hydrogens have been treated all-electron. The primitive Gaussian basis set is O(5s,5p,2d) H(4s,2p) for the determinant, while for the Jastrow factor it is O(3s,2p,1d) H(2s,1p). The latter basis set has been recently claimed to be one of the most accurate in an extensive QMC study of single molecule water properties,[60] which used the same Jastrow ansatz as ours. The inhomogeneous three- and four-body Jastrow term (eq 10) has been developed directly on the uncontracted primitive basis set, whose flexibility guarantees a very good description of dynamical correlations. For the antisymmetric part, we tested two main wave function forms, the single Slater determinant (obtained by using a geminal with rank $N/2$) and the AGP function. The difference in energy between the JSD and the JAGP wave functions, reported in Table 1, shows the size of static correlations in the system, which amounts to 5−6 mHa. This leads also to better geometrical properties, as seen in Table 3. The JAGP geometry is closer to the experiment than the JSD one, in both the OH distance and the HOH angle. The structural effects of static correlations in the water molecule have been already pointed out by Zen et al.,[60] where they were attributed mainly to a change in the local description of the oxygen atom. In order to analyze in deeper detail the role of the AGP correlations in the water molecule, we studied the natural orbital occupations coming from the diagonalization of the

Table 3. Geometrical Properties of the Global Minimum of the Water Molecule[a]

	$\overline{OH}$ (Å)	∠HOH (deg)
pseudo JSD	0.9542(4)	104.730
pseudo JAGP	0.9549(4)	104.549
all-electron JSD	0.9539(4)	105.187
all-electron JAGP	0.9557(4)	105.101
experiment[77]	0.95721(30)	104.522(50)

[a]We report a comparison between different QMC wave functions and experimental results.[77]

dx.doi.org/10.1021/ct401077x | J. Chem. Theory Comput. 2014, 10, 1980−1993

geminal. One of the appealing features of the AGP theory is that the geminal wave function is directly related to the one-body density matrix of the system (without the Jastrow factor). Its diagonalization yields the molecular natural orbitals as eigenvectors, and their weights are related to the modulus of the AGP eigenvalues which are plotted in Figure 1. This figure

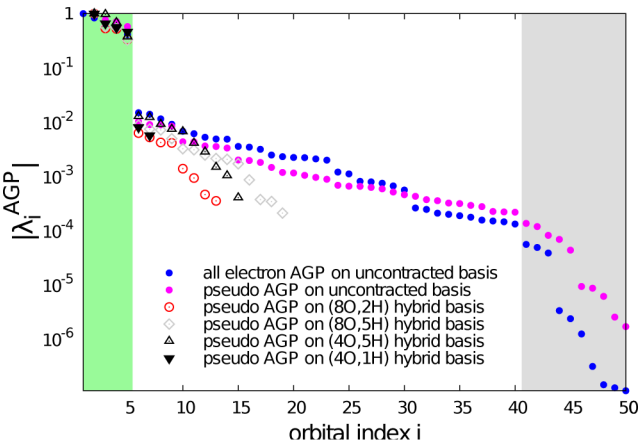


Figure 1. Semilog plot of the modulus of the AGP eigenvalues versus the molecular orbital index for different basis sets and calculations. The orbital indices always include the oxygen 1s electrons, replaced in the pseudopotential calculations. The green area represents the exactly occupied molecular orbitals in the single Slater determinant representation, with $\lambda_i = 1$ for $i \in \{1, ..., \text{HOMO}\}$ and $\lambda_i = 0$ for $i \geq$ LUMO. In the AGP representation, by diagonalizing the geminal, we obtain the corresponding molecular orbitals eigenvectors and their weights λ_i^{AGP} (eigenvalues). In the AGP, the orbitals above the HOMO are also occupied, with a weight which jumps across the HOMO to LUMO transition (in going from the green to white region). The gray area shows the MO occupation tail falling rapidly to zero in the full AGP (expanded on a primitive basis set), signaling that the MOs above that threshold start to become irrelevant to the description of the static correlations in the system.

shows that indeed the orbitals above the HOMO (highest occupied molecular orbital in the single Slater determinant representation) have a sizable weight, with a distribution which falls abruptly to zero only after the 40th orbital (gray area in the plot). This reflects the multideterminant character of the water molecule, taken into account by the AGP ansatz. We thus believe that, although the entanglement of quantum levels at the origin of static correlations can come from the oxygen atom, its impact in water has a genuine molecular character. Finally, the multideterminant AGP representation leads to a better description of the nodal surface of the true ground state, with a gain of about 2.5 mHa in the fixed node LRDMC energy with respect to the energy obtained using the JSD as a trial wave function, as reported in Table 4.

We turn now our attention to the question of how to reduce the AGP basis set in an effective way. So far, both the JSD and JAGP wave functions have been developed on the uncontracted primitive basis in order to best exploit its flexibility. Thus, the total number of variational parameters is 895 (see the last column of Table 1), quite large for a single molecule, particularly if one would like to tackle the study of larger water clusters by means of QMC techniques. The most important limitation of this approach is that the number of variational parameters corresponding to the matrix elements $\lambda_{\alpha,\beta}^{a,b}$ increases as the square of the atomic basis size; therefore

Table 4. LRDMC Energy Results Extrapolated to the Zero Lattice Space Limit[a]

	pseudo	all-electron
E_{JSD} (Ha)	−17.26280(6)	−76.42475(15)
E_{JAGP} (hybrid) (Ha)	−17.26522(6)	−76.42696(14)
E_{JAGP} (Ha)	−17.26528(6)	−76.42690(14)
$E_{\text{JAGP}} - E_{\text{JSD}}$ (mHa)	−2.48(9)	−2.15(21)
E_{JAGP}(hybrid) $- E_{\text{JSD}}$ (mHa)	−2.42(9)	−2.21(21)

[a]The LRDMC calculations are performed in the fixed-node approximation. In the last two rows, we compute the energy gain due to a better nodal description provided by the JAGP wave function with respect to the JSD one, in both uncontracted and contracted bases. The hybrid 8O 2H and 9O 2H basis sets have been used for the pseudopotential and all-electron calculations, respectively. Note that the agreement between the pseudopotential and all-electron calculations has an accuracy of the order of 0.1 mHa, despite their very different total energies. Finally, the contracted hybrid basis set gives the same LRDMC energy as the uncontracted one, proving that the nodal surface is not affected by the contraction.

this size should be minimized in order to make the approach feasible for a large number of molecules.

As explained in section 2.2.1, we define a new smaller basis set by contracting the O(5s,5p,2d)/H(4s,2p) Gaussian primitive basis via atomic natural hybrid orbitals. Each atom in the system is described by its own set of hybrid orbitals. We study how the size of the contracted hybrid basis set affects the quality of the geminal expansion. We compare it with the rigorous lowest energy limit provided by the uncontracted JAGP reference previously computed. The results are reported in Table 1. The smallest basis set which includes the 1s for H and the 2s and 2p orbitals for O, thus taking into account the 2s2p near degeneracy at the atomic O level, is the 4O 1H hybrid set (in self-explaining notation). It gives the poorest energy and variance among the hybrid basis sets considered yet is lower than the JSD ansatz. The best energy is obtained with the largest hybrid basis tried here, namely the 8O 5H set. It recovers a large fraction of static correlations, and its energy is less than 1 mHa above the uncontracted JAGP one. However, the price to pay is that the parameter reduction is weak, the total number of parameters being close to that of the full JAGP expansion (see last column of Table 1). Indeed, while the number of $\lambda_{\alpha,\beta}^{a,b}$ is still significantly lower than for the uncontracted basis set, the number of parameters in the contracted orbitals grows too much. The best compromise between efficiency, i.e., total number of variational parameters, and accuracy, i.e., good variational energy, is provided by the 8O 2H basis, as it yields a significant gain in energy with a small/moderate number of parameters. This advantage will be remarkable for a large number of atoms, since the number of variational parameters corresponding to the atomic natural hybrid orbitals grows only *linearly* with the number of atoms; on the other hand, the number of parameters corresponding to $\lambda_{\alpha,\beta}^{a,b}$ grows instead quadratically, but it remains still affordable since it is dramatically reduced by this approach (see Table 2).

Finally, we study how the AGP spectrum changes with the contracted hybrid basis sets. Figure 1 shows that, after a complete wave function optimization, the natural orbital eigenvalues magnitude of the hybrid AGP covers the 10^{-2} to 10^{-4} range of the fully uncontracted AGP expansion, except for the shortest 4O 1H basis, which clearly spans a too small Hilbert space. Moreover, we checked that the JAGP expanded on the optimal 8O 2H basis gives the same fixed node LRDMC

energy as the full JAGP, demonstrating that the nodal surface is properly described (see Table 4) even by the hybrid 8O 2H contraction. As reported in the Computational Methods, the gain in efficiency of the hybrid basis set is expected to be larger and larger as the system size increases, since the quadratic growth of the $\lambda_{\alpha,\beta}^{a,b}$ with the number of atoms depends strongly on the atomic basis size.

The combination of the BFD pseudopotential for oxygen and the JAGP ansatz together with the 8O 2H hybrid basis set, tested in the water molecule, is thus chosen to study the protonated water dimer.

3.1.2. All-Electron Calculations. To study the accuracy of the pseudopotential approximation, we also carried out all-electron calculations on the water molecule. With respect to the pseudopotential calculations, the primitive and contracted basis sets for O have been extended in order to account for the additional 1s electrons. The primitive basis set is now O(6s6p2d) and O(3s3p1d) for the determinant and Jastrow term, respectively, while the optimal contracted hybrid basis set is 9O 2H. In Table 2, we report the variational energies for different wave functions. The energy gain provided by the all-electron JAGP wave function is very close to the one in pseudopotential calculations. The substantial agreement between the two calculations is also apparent in Figure 1, where the eigenvalues of the higher energy molecular natural orbitals in the AGP behave similarly. The LRDMC energy difference between the JAGP and JSD trial wave functions coincides within the error bars with the one with pseudopotentials. The nodal contribution to the fixed-node energy is the same. The JAGP LRDMC energy is one of the best ever published in the literature. To the best of our knowledge, the best QMC energy of a single water molecule has been computed by Clark et al.[78] (−76.4368(4) Ha). Our JAGP LRDMC energy, reported in Table 4, is less than 10 mHa above this result; furthermore, it is in statistical agreement with recent computations by Lüchow and Fink[79] (−76.429(1) Ha), who used 300 determinants in the trial wave function, and with Zen et al.[60] (−76.42660 Ha), who used the same ansatz as ours in the trial. The LRDMC energy is only 11 mHa higher than the extrapolated exact result of −76.438 Ha.[80]

The all-electron calculations confirm the importance of including static correlations to have a better description of the geometry, as shown in Table 3. However, the HOH angle turns out to be less accurate than the one obtained with pseudopotentials if compared to the experiment, most probably because it is very sensitive to the basis set convergence, which is harder to reach in all-electron calculations. Another drawback of all-electron calculations is of course the larger variance for an equivalent wave function ansatz, due to the 1s electron fluctuations, as one can easily conclude from the comparison between Table 1 and Table 2.

Due to the larger primitive basis set required in all-electron calculations, the parameter reduction allowed by the hybrid basis contraction (9O 2H) has a great impact on the efficiency. The total number of parameters is reduced by almost a factor of 4 in the determinantal part of the single molecule, without any significant loss of accuracy in the JAGP total energy (see Table 2). The few all-electron calculations done on the Zundel ion are performed with this contracted basis set.

3.2. Protonated Water Dimer. The protonated water dimer represents the simplest model for proton transfer in aqueous systems. In our work, we study the energetics of $H_5O_2^+$ by choosing a suitable reaction coordinate (RC). Selecting a RC allows one to reduce the complexity of the full-dimensional PES and project it onto a single-dimensional subset which retains the desired physical features of the full hypersurface.

As modeled by our system, the proton transfer reaction takes place in three different steps. First, the excess proton is bound to one water molecule forming a $H_3O^+ + H_2O$ complex. By thermal fluctuations, the oxygen−oxygen separation can approach the optimal distance of the Zundel complex (around 2.39 Å); at this stage, the system assumes a "Zundel configuration" with the proton equally shared between the two oxygens. A further stretch of $\overline{OO}$ destabilizes the Zundel configuration, and a new $H_2O + H_3O^+$ complex is produced. The overall effect of this process is a transfer of a proton along the hydrogen bond between two oxygen atoms. This mechanism suggests the choice of the $\overline{OO}$ separation as the RC for the dimer potential energy curve.

In the following subsections, we present the geometry and energetics of the protonated water dimer. Results on the global minimum geometry are first reported. The potential energy curve is the result of a stretching of the $\overline{OO}$ distance; the behavior of the excess proton is thus investigated along the same path. Particular attention is devoted to the behavior of the $H_3O^+ + H_2O$ complex in the broken-symmetry region of the energy landscape. Proton transfer static barriers at different $\overline{OO}$ distances are calculated in order to further estimate the accuracy of our approach. All QMC calculations are performed with the JAGP wave function ansatz developed on an atomic hybrid basis set, as discussed in section 3.1.

3.2.1. Properties of the Symmetric Global Minimum. The minimum energy structure of $H_5O_2^+$ has been debated in the literature, as there are two candidates with competing energies: a C_2 symmetric structure, commonly known as the Zundel configuration, with the proton evenly shared between the two oxygens, and a C_s-Inv one with the proton slightly closer to one H_2O molecule (see Figure 2). Several *ab initio* investigations[35,38] have shown that a better treatment of electron correlation results in a change of the ground state (GS) geometry from the C_s to the C_2 configuration. Accurate highly correlated studies[36−38,41] have eventually confirmed that the global minimum is C_2 symmetric.

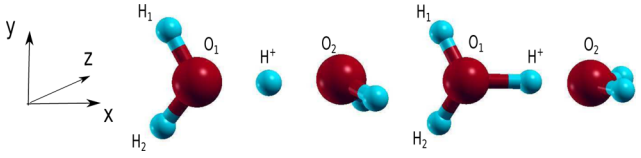


Figure 2. QMC optimized geometries for global C_2 minimum (left) and for C_s local minimum (right).

At a QMC level of theory, the $H_5O_2^+$ GS has a C_2 Zundel configuration with centrosymmetric excess protons (left-hand side of Figure 2), in agreement with previous studies. The main geometrical parameters of the global minimum are presented in Table 5, for both the pseudopotential and all-electron calculations.

When compared with CCSD(T), the QMC ground state geometries agree to within an atomic separation of 0.005 Å. The agreement in the intramolecular angles is within 0.3°, while the largest discrepancy between QMC and CCSD(T) is in the dihedral angles ω between two H_2O planes, which are related to soft vibrational modes and therefore affected by a larger

dx.doi.org/10.1021/ct401077x | J. Chem. Theory Comput. 2014, 10, 1980−1993

Table 5. Geometrical Properties (Distances in Å, Angles ∠ and Dihedral Angles ω in deg) of the C_2-Symmetry Minimum of Protonated Water Dimer, Comparison between Different Computational Methods[a]

	DFT-PBE	DFT-B3LYP[37]	QMC − with pseudo	QMC − all-electron	CCSD(T)[41]
$\overline{O_1O_2}$	2.4111		2.3847(5)	2.3905(4)	2.3864
$\overline{O_1H^+}$	1.2074	1.2172	1.1942(8)	1.1989(5)	1.1950
$\overline{H^+O_2}$	1.2074		1.1930(5)	1.1944(6)	1.1950
$\angle O_1H^+O_2$	173.661	173.6	174.71(7)	174.43(9)	173.730
$\overline{O_1H_1}$	0.9697	0.9784	0.9605(8)	0.9630(7)	0.9686
$\overline{O_1H_2}$	0.9691	0.9778	0.9650(8)	0.9628(6)	0.9682
$\angle H_1O_1H_2$	109.161		109.16(9)	109.40(7)	108.8
ω-$H_1O_1H^+O_2$	295.690		293.5	296.6	295.3
ω-$H_2O_1H^+O_2$	163.809		158.5	162.0	163.6

[a]Figure 2 (left-hand side) for atomic notation.

statistical bias during the relaxation. Nevertheless, these differences do not affect the overall description of the GS and the energetics of the system.

In the present work, also the C_s-Inv structure has been taken into account (right-hand side of Figure 2). Table 6, which

Table 6. Geometrical Properties (Distances in Å, Angles in deg) of the C_s-Inv Minimum of Protonated Water Dimer[a]

	DFT-B3LYP[37]	QMC − with pseudo	QMC − all-electron	CCSD(T)[41]
$\overline{O_1O_2}$		2.3996(6)	2.3913(3)	2.3989
$\overline{O_1H^+}$		1.1154(8)	1.1285(5)	1.1233
$\overline{H^+O_2}$	1.2507	1.2852(4)	1.2648(4)	1.2720
$\angle O_1H^+O_2$	175.4	176.5(1)	175.29(6)	175.646
$\overline{O_1H_1}$	0.9746	0.9641(7)	0.9635(4)	0.9641
$\overline{O_1H_2}$	0.9741	0.9625(4)	0.9616(5)	0.9645
$\angle H_1O_1H_2$		110.56(8)	110.66(6)	110.153

[a]See Figure 2 (right-hand side) for atomic notation.

reports the VMC optimized C_s geometries, confirms the trend seen in Table 5 for the GS, although the discrepancies in the bond lengths are slightly larger between different methods, the CCSD(T) values being in between the all-electron and pseudopotential VMC results.

In the pseudopotential calculations, the energy difference between the C_s configuration and the C_2-symmetric global minimum turns out to be 0.25(8) kcal/mol at the VMC level and 0.23(8) kcal/mol at the LRDMC. This is in satisfactory agreement with previous results carried out with Møller−Plesset perturbation theory[36] (0.28 kcal/mol) and CCSD(T)[41] (0.46 kcal/mol) techniques.

3.2.2. Stretching the $\overline{OO}$ *Distance.* From the C_2-sym global minimum, we stretch the $\overline{OO}$ distance in order to study the potential energy curve and elucidate the proton transfer character in the dimer. We verified that an analogous stretch from the C_s local minimum yields a higher energy over the entire $\overline{OO}$ range.

The structural relaxation at the VMC level for fixed $\overline{OO}$ separation requires a careful procedure due to the flatness of the PES. Starting from a PBE-DFT optimized geometry (see section 2.1.1 for technical details) and a JAGP variational wave function fully optimized in the electronic part, the atomic coordinates are relaxed with the steepest descent method, by considering them as additional variational parameters of the QMC wave function.

For the sake of comparison, we minimize a parametrized full-dimensional PES fitted from CCSD(T) calculations[41] to find the best coupled cluster estimates of energy and geometry. By means of the *downhill simplex* minimization technique, we find the configuration of lowest CCSD(T) energy at the same constrained $\overline{OO}$ distance as the corresponding QMC and DFT calculations.

In Figure 3, we plot the energy landscape along the RC for PBE-DFT, CCSD(T), VMC, and LRDMC; the latter

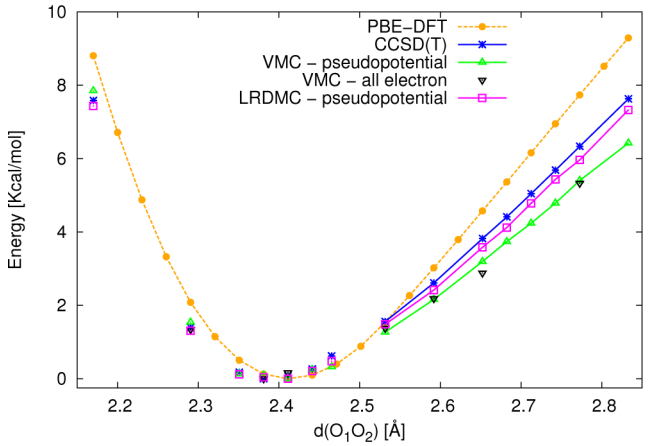


Figure 3. Potential energy curve (kcal/mol) of the protonated water dimer projected on the $\overline{OO}$ distance. Comparison between different computational methods. Structural relaxation is performed at each level of theory. Each curve has its minimum as the reference point.

computed at the VMC geometry. For the VMC technique, we also report all-electron calculations for some $\overline{OO}$ separations. We notice a good agreement among all techniques in the region at the left of the global minimum of the curve, except for a constant shift by 0.02 Å between the PBE-DFT results and the others. The PBE $\overline{OO}$ minimum is indeed located at 2.41 Å, while the minimum of the other methods turns out to be at ∼2.39 Å. Hence, this part of the energy curve is only slightly influenced by a better treatment of correlations.

On the other hand, the region at the right of the minimum, at intermediate $\overline{OO}$ distances (≥2.55 Å), displays a different behavior. The PBE-DFT overestimates the slope of the curve with respect to the most accurate techniques. LRDMC, which yields the best QMC correlation energy, shows remarkably good agreement with the state-of-the-art CCSD(T) results. In particular, all CCSD(T) values are in the range of ∼0.3 kcal/

mol, three times the statistical error of the LRDMC calculations, as shown in Figure 4. Figure 4 also shows that

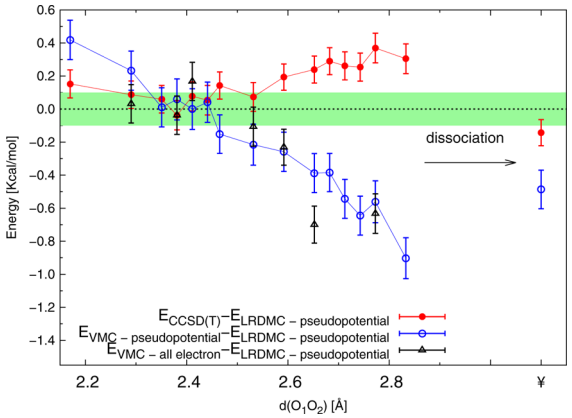


Figure 4. CCSD(T) and VMC energy landscape (kcal/mol) as a function of the $\overline{OO}$ distance (Å) with respect to the LRDMC energies (the zero of the *y* axis). Full potential VMC results are reported for few points along with dissociation energies. The green area represents the statistical error bar achieved in a typical LRDMC run, i.e. ∼0.1 kcal/mol.

pseudopotential and all-electron VMC calculations are in statistical agreement for the reported $\overline{OO}$ separations. This demonstrates the quality of the BFD pseudopotentials in the Zundel ion, verified also in Table 5 and Table 6 for the geometry of the C_2-sym and C_s-Inv minima, respectively (see section 3.2.1).

The curves in Figure 3 are obtained with the minimum energy geometry at each level of theory (except for the LRDMC, whose geometry is set at the VMC level). In order to have a more reliable comparison and avoid the bias from the use of different geometries, we carried out the same calculations employing the VMC-optimized structures for every technique. The result is reported in Figure 5. The trend displayed in Figure 3 for method-optimized geometries is remarkably enhanced when the same configuration of the dimer is considered. Away from the minimum, the PBE-DFT energies show a larger overestimation of the slope.

The slope of the dimer potential energy curve is related to the behavior of the excess proton in the system. In Figure 6,

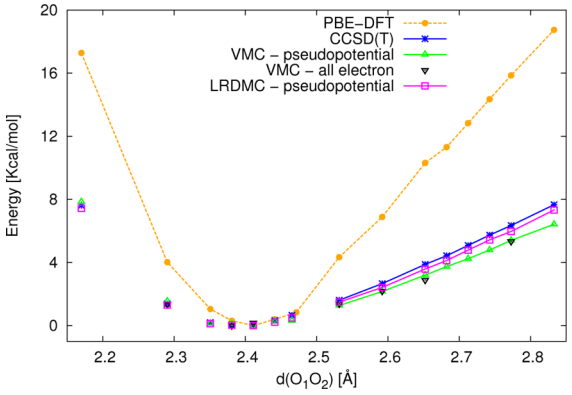


Figure 5. Protonated water dimer energy landscape (kcal/mol) as a function of $\overline{OO}$ distance (Å). All the calculations are performed with VMC-optimized geometry. The zero energy reference point corresponds to the minimum of each curve.

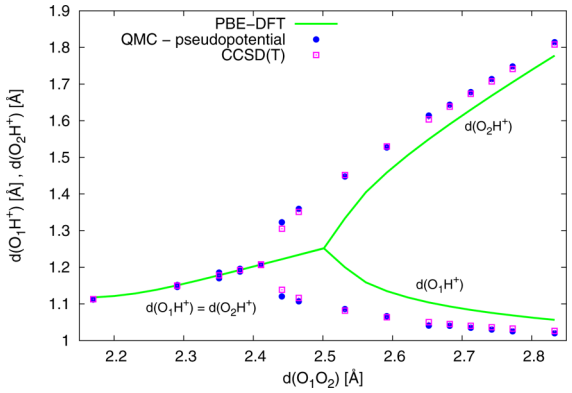


Figure 6. Separations (Å) between the two oxygens and the excess proton as a function of the reaction coordinate for different computational methods.

this property is elucidated. We report the separations between the excess proton and each of the two oxygens in the same plot: $\overline{O_1H^+}$ and $\overline{O_2H^+}$. The structure of the dimer is relaxed at different levels of theory with the same procedure carried out for the potential energy curve in Figure 3. The plot clearly shows the appearance of two distinct regimes of the dimer. One is characterized by a symmetric Zundel configuration with the proton evenly bonded to the two oxygens; from the point of view of proton transfer physics, it is basically equivalent to the GS configuration. Stretching the $\overline{OO}$ distance results in the formation of a $H_3O^+ + H_2O$ complex with the proton localized on one water molecule. These configurations belong to the *asymmetric regime* of the dimer. Within this regime, the initial C_2 point symmetry of the GS geometry is broken due to proton localization.

The distances obtained by the QMC relaxation of the atomic coordinates are in excellent agreement with the CCSD(T) calculations; in particular, the root-mean-square distance between the two data sets over the whole $\overline{OO}$ range is ∼0.007 Å for both $\overline{O_1H^+}$ and $\overline{O_2H^+}$. Electron correlation plays a key role in determining the stability of the symmetric configuration of the dimer. We define d_c as the critical $\overline{OO}$ distance at which the symmetric configuration of the excess proton is broken. The overestimation of the PES slope by PBE-DFT in Figure 3 corresponds to an overestimation of d_c by ∼+0.13 Å with respect to higher-level post-HF *ab initio* methods. Furthermore, DFT poorly describes the geometry in the broken-symmetry region close to d_c, where the discrepancy in the $\overline{OH^+}$ distances is the largest (going up to 0.15 Å).

3.2.3. Implications for More Realistic PT Models. The zero-temperature potential energy curves reported in Figure 3 and Figure 5 seem to conflict with the proton transfer mechanism discussed in the introduction of section 3.2, since the configuration with a centrosymmetric proton is energetically favored, and therefore it represents a stable state rather than a transition state between two asymmetric configurations with a localized proton.

However, this contradiction can be easily explained. Indeed, Marx et al.[27] have shown that the introduction of thermal and polarization effects due to the physical environment favors the asymmetric regime of the complex. At finite temperature, the *free energy* landscape displays a global minimum shifted toward the asymmetric regime, and the Zundel-like structure does not represent the energetically favored configuration any more.

dx.doi.org/10.1021/ct401077x | J. Chem. Theory Comput. 2014, 10, 1980−1993

Moreover, a recent experimental result showed that the average $\overline{OO}$ distance in liquid water is 2.81 Å,[81] which clearly corresponds to a symmetry-broken configuration of the dimer (see Figure 6). In any case, in order to jump from a water molecule to one of its neighbors, the proton must pass through a Zundel configuration. Correctly describing the energetics and geometry of the protonated water dimer in the symmetry-breaking transition region is therefore of paramount importance to get an accurate description of the PT in more realistic models.

As mentioned before, Figure 6 highlights that at 0 K a better treatment of electron correlation yields the asymmetric regime stably over a considerably wider range of $\overline{OO}$ distances with respect to DFT results. As a consequence, we expect the calculated proton diffusion in water to be very sensitive to the level of theory.[25−27]

3.2.4. Properties of the Symmetry-Broken Configurations. A quantity which has been extensively studied over the past years[24,25,82,83] is the *static proton transfer barrier*, i.e., the barrier that the H^+ has to overcome in order to jump from one H_2O molecule to the other at a fixed $\overline{OO}$ distance in the symmetry-broken regime. This quantity does not provide a realistic comparison with the experimental activation barrier for PT, as the $\overline{OO}$ distance will shorten during the proton hopping. Nevertheless, it is relevant in order to provide a further check of the accuracy of our QMC approach. Fixing the $\overline{OO}$ distance, the barriers are obtained as the difference between the asymmetric configuration with a localized proton and a structure with the excess proton at equal distance from the two oxygens. Calculations are performed in three representative $\overline{OO}$ separations; the results are shown in Table 7, where they are compared with existing data in the literature.

Table 7. Static Proton Transfer Barriers (kcal/mol) at Fixed $\overline{OO}$ Separations[a]

method		O−O distance		
energy	geometry	2.47 Å	2.6 Å	2.7 Å
DFT-PBE		−0.74	0.19	1.94
VMC		0.28(6)	2.99(8)	5.99(8)
LRDMC	VMC	0.37(8)	2.64(7)	5.57(7)
CCSD(T)		0.22	2.37	5.24
CCSD(T)	VMC	0.28	2.29	5.32
CCSD(T)[83]	MP2		2.08	4.85
QCISD(T)[83]	MP2		2.06	4.82
MS-EVB[25]			2.05	5.11
MP2[83]			1.77	4.39

[a]Comparison between different levels of theory at different geometries. If not specified, the structure is optimized with the same method as the corresponding energy calculation.

The first $\overline{OO}$ distance is 2.47 Å, very close to the d_c displayed by highly correlated approaches (2.43 Å). In contrast, at a DFT level, the configuration with the centrosymmetric proton is still energetically favored, as shown by Figure 6. It has been noted[38] that the PES in this region is particularly flat. Furthermore, the potential energy curve of the dimer develops on very tiny energy differences around d_c. These issues make the calculations in this region of the PES extremely delicate since the stochastic noise can considerably affect the quality of the QMC predictions.

Table 7 shows that QMC and CCSD(T) display a vanishing energy barrier on the order of 0.2−0.3 kcal/mol at a $\overline{OO}$ distance close to d_c. The height of this barrier is slightly above the attained statistical error in our typical QMC run. However, despite the very sensitive behavior of the dimer PES around d_c, the accuracy of our force minimization algorithm allows us to account for the tiny energy differences involved. Thus, it ensures the precision necessary to describe the PT physics in the dimer. As we discuss in the next section, a similar accuracy can be achieved in larger molecular clusters with a reasonable amount of computational time. Therefore, our QMC framework guarantees a reliable description of the PT physics also for more realistic models.

The other results are obtained at larger oxygen separations, further away from d_c. They confirm the general behavior already seen along the $\overline{OO}$ reaction coordinate. LRDMC and CCSD(T) results are in a agreement within 0.3 kcal/mol, whereas the VMC overestimates the barrier of about 0.6 kcal/mol. As is known from previous works, DFT substantially underestimates the barrier with respect to post-HF methods, which provide a better treatment of correlations.

Finally, let us analyze the extreme limit of the asymmetric Zundel configuration, namely when $\overline{OO} \to \infty$, with the formation of one H_2O and one hydronium. The dissociation energy D_e of $H_5O_2^+$ is computed by setting the distance between the two oxygens to 14 Å. With the CCSD(T) PES, we checked that this is already in the large distance plateau. We get a D_e of 33.02(9) kcal/mol by VMC and 32.54(8) kcal/mol by LRDMC, to be compared with the CCSD(T) value of 32.68 kcal/mol computed Huang et al.,[41] while the PBE-DFT gives D_e = 29.55 kcal/mol. The agreement between the LRDMC and CCSD(T) is impressive, while it is good for the VMC estimate. This is mainly due to the size consistency of the JAGP ansatz, obtained once the Jastrow factor is close to the complete basis set limit.[84] The VMC and CCSD(T) dissociation energies are plotted in Figure 4 with respect to the LRDMC values.

3.2.5. Test on a Larger Molecular Cluster. QMC methods present a favorable scalability with the number of particles with respect to other highly correlated approaches such as CC. With the aim of demonstrating this feature for our framework, we performed a benchmark calculation on a more realistic PT model composed of six water molecules and one excess proton, which will be the subject of a further study. In Table 8, we report a comparison of the computational time required to carry out typical VMC and LRDMC runs for different sizes of

Table 8. Total Computational Wall Time in Hours of Typical VMC and LRDMC Runs at Fixed Optimized Variational Parameters and Geometry[a]

# of water molecules	total wall time (h) on 512 CPUs	
	VMC	LRDMC
1	0.05	0.15
2	0.24	2.13
6	6.49	164.35

[a]Calculations have been performed with the TurboRVB[43] program on 32 thin nodes of the Curie HCP machine (2.7 GHz, 16 core Intel Sandy Bridge processors). The target statistical error reached is of 0.06 kcal/mol in total energies. A single water molecule, a protonated dimer, and a larger cluster of six water molecules with one excess proton are compared. The LRDMC is carried out at a lattice space of 0.125 a_0.

the protonated water cluster. These calculations are performed at fixed optimized variational parameters and geometry (*single-point calculations*).

The VMC structural relaxations are more expensive than single-point runs since they involve variational parameter optimization and atomic forces evaluation. Within our calculations, the computational cost of a typical structural relaxation run in order to obtain well-converged parameters and forces is about 5 times larger than the VMC single-point calculations presented in Table 8. Note that this cost strictly depends on the quality of the starting wave function and geometry of the system.

All calculations in Table 8 have been carried out on the HPC Curie thin nodes (2.7 GHz 16 core Intel Sandy Bridge processors), and the QMC target statistical error has been set to 0.06 kcal/mol for the total energy. By means of a simple polynomial fit on the data in Table 8, we see that the LRDMC, carried out at a lattice space of $a = 0.125a_0$, displays an almost perfect N^4 scaling. The simpler variational Monte Carlo technique shows an N^3 scaling. As the variance scales linearly with N, and the QMC all-electron move costs N^3, the resulting theoretical scaling is N^4. An apparent scaling faster than N^4 is probably due to the BLAS-based implementation of the TurboRVB program, which becomes more efficient as the system size increases, by reducing the N^4 prefactor the most. It turns out that the LRDMC is feasible in a reasonable computational time for the $6H_2O$ cluster, while the VMC is still cheap at that cluster size.

In order to contextualize the values reported in Table 8, we performed a benchmark CCSD(T) calculation of the small Zundel ion on the same machine. We use the Dunning's correlation consistent aug-cc-pVTZ basis set, the same employed by Huang et al.[41] for the fitted PES. Calculations have been done with the Gamess[85] program. A single-point run on 128 CPUs costs 2.25 min, approximately 1 order of magnitude less than the corresponding QMC calculation, since at this cluster size the large QMC prefactor dominates over its favorable scaling.

According to the theoretical N^7 scaling of CCSD(T), an analogous calculation on the six-molecule cluster would approximately have the same cost as the corresponding VMC calculation. Therefore, the six-molecule complex represents a crossing point in the relative efficiency between VMC and CCSD(T) methods; it is obvious that a further increase in the cluster size would make the QMC approach considerably favored in terms of computational demand.

Moreover, the coupled cluster estimated cost has been obtained by assuming no parallelization overhead and no constraints of memory allocation, criteria which are hard to meet in coupled cluster calculations. On the other hand, quantum Monte Carlo has less demanding memory requirements and an almost perfect parallel scaling with the number of cores.

4. CONCLUSIONS

In this paper, we presented an extensive study of the protonated water dimer by means of the VMC and LRDMC techniques.

The JAGP ansatz employed in this work implements an accurate treatment of both static and dynamical correlations. The expansion of the determinantal part over atomic hybrid orbitals ensures a drastic reduction of the number of variational parameters thus making the wave function optimization procedure efficient and robust even for large systems. The comparison with previous published QMC calculations on the single water molecule showed the quality of our wave function ansatz.

Total energy calculations were performed with the less expensive variational Monte Carlo approach and the more precise projective diffusion Monte Carlo. The powerful minimization algorithm implemented in the TurboRVB software allowed an efficient estimation of the forces acting on each atomic component with a reasonable computational cost. Hence, both energetics and geometry calculations were performed within the QMC framework.

A simple mechanism of proton transfer in the dimer has been presented and exploited to choose a suitable reaction coordinate for the potential energy curve. The energy landscape as a function of the oxygen−oxygen distance is computed and compared with density functional theory in the PBE approximation and with CCSD(T) results. LRDMC is in excellent agreement with CCSD(T) calculations (within 0.3 kcal/mol), whereas minor differences (up to about 1 kcal/mol) are shown by VMC.

A similar accuracy has been recently obtained by diffusion Monte Carlo calculations for the water dimer dissociation energy[86,87] and for the energetics of small water clusters.[88] In contrast with previous works, we are able to optimize both the electronic and structural parts at the variational Monte Carlo level, yielding highly accurate trial wave functions, while in previous works the geometry was either taken from experimental results or drawn from force field calculations.

VMC structural relaxation provides geometries remarkably close to the ones obtained by a CCSD(T) fitted PES. We show the presence of two distinct regimes of the dimer depending on the oxygen−oxygen distance: one with a centrosymmetric excess proton and the other with the proton localized on one of the water molecules. The stability of these configurations crucially depends on the level of theory; a better treatment of electron correlation results in the stability of the asymmetric proton geometry over a wider range of $\overline{OO}$ distances.

These results, together with the proton transfer static barrier and the dissociation energy D_e, show that our QMC approach has a global accuracy comparable with the state-of-the-art coupled cluster in both geometry and energetics of the dimer, with the advantage of having a better scaling with the number of particles, as demonstrated by a test simulation on a larger protonated water cluster.

These features make this approach a very promising candidate for the study of proton transfer in complex aqueous systems. In particular, we have shown that the VMC method is cheap, provides very accurate geometries, and gives a global error of less than 1 kcal/mol in the most important region for PT physics. We hope that our work will inspire further studies in this direction and pave the way for accurate highly correlated simulations of more realistic proton transfer models which will eventually shed new insights onto the PT mechanism in water.

AUTHOR INFORMATION

Corresponding Author

*E-mail: mario.dagrada@impmc.upmc.fr.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Joel M. Bowman and Xinchuan Huang for providing us with the coupled cluster potential energy surface of $H_5O_2^+$. M.C. acknowledges computational resources in the form of the GENCI grant number x2013096493.

■ REFERENCES

(1) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926−935.
(2) Sprik, M. *J. Phys. Chem.* **1991**, *95*, 2283−2291.
(3) Mahoney, M. W.; Jorgensen, W. L. *J. Chem. Phys.* **2000**, *112*, 8910−8922.
(4) Silvestrelli, P. L.; Parrinello, M. *J. Chem. Phys.* **1999**, *111*, 3572−3580.
(5) Grossman, J. C.; Schwegler, E.; Draeger, E. W.; Gygi, F.; Galli, G. *J. Chem. Phys.* **2004**, *120*, 300−311.
(6) VandeVondele, J.; Mohamed, F.; Krack, M.; Hutter, J.; Sprik, M.; Parrinello, M. *J. Chem. Phys.* **2005**, *122*, 014515.
(7) Jonchiere, R.; Seitsonen, A. P.; Ferlat, G.; Saitta, A. M.; Vuilleumier, R. *J. Chem. Phys.* **2011**, *135*, 154503.
(8) Carloni, P.; Rothlisberger, U.; Parrinello, M. *Acc. Chem. Res.* **2002**, *35*, 455−464.
(9) Friesner, R. A.; Dunietz, B. D. *Acc. Chem. Res.* **2001**, *34*, 351−358.
(10) Yoo, S.; Zeng, X. C.; Xantheas, S. S. *J. Chem. Phys.* **2009**, *130*, 221102.
(11) Lee, H.-S.; Tuckerman, M. E. *J. Chem. Phys.* **2007**, *126*, 164501.
(12) von Grotthuss, F. C. J. D. T. *Ann. Chim.* **1806**, *58*, 54−73.
(13) Agmon, N. *Chem. Phys. Lett.* **1995**, *244*, 456−462.
(14) Marx, D. *ChemPhysChem* **2006**, *7*, 1848−1870.
(15) Decoursey, T. E. *Physiol. Rev.* **2003**, *83*, 475−579.
(16) Freier, E.; Wolf, S.; Gerwert, K. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 11435−11439.
(17) Deisenhofer, J.; Epp, O.; Sinning, I.; Michel, H. *J. Mol. Biol.* **1995**, *246*, 429−457.
(18) Lancaster, C. R. D.; Michel, H.; Honig, B.; Gunner, M. R. *Biophys. J.* **1996**, *70*, 2469−2492.
(19) Eigen, M. *Angew. Chem., Int. Ed.* **1964**, *3*, 1−19.
(20) Mohammed, O. F.; Pines, D.; Dreyer, J.; Pines, E.; Nibbering, E. T. J. *Science* **2005**, *310*, 83−86.
(21) Mohammed, O. F.; Pines, D.; Pines, E.; Nibbering, E. T. *Chem. Phys.* **2007**, *341*, 240−257.
(22) Luz, Z.; Meiboom, S. *J. Am. Chem. Soc.* **1964**, *86*, 4768−4769.
(23) Noam, A. *J. Chim. Phys.* **1996**, *93*, 1714−1736.
(24) Wei, D.; Salahub, D. R. *J. Chem. Phys.* **1994**, *101*, 7633−7642.
(25) Schmitt, U. W.; Voth, G. A. *J. Chem. Phys.* **1999**, *111*, 9361−9381.
(26) Lapid, H.; Agmon, N.; Petersen, M. K.; Voth, G. A. *J. Chem. Phys.* **2005**, *122*, 014506.
(27) Marx, D.; Tuckerman, M. E.; Hutter, J.; Parrinello, M. *Nature* **1999**, *397*, 601−604.
(28) Markovitch, O.; Chen, H.; Izvekov, S.; Paesani, F.; Voth, G. A.; Agmon, N. *J. Phys. Chem. B* **2008**, *112*, 9456−9466.
(29) Marx, D.; Chandra, A.; Tuckerman, M. E. *Chem. Rev.* **2010**, *110*, 2174−2216 PMID: 20170203.
(30) Saitta, A. M.; Saija, F.; Giaquinta, P. V. *Phys. Rev. Lett.* **2012**, *108*, 207801.
(31) Hassanali, A.; Giberti, F.; Cuny, J.; Kühne, T. D.; Parrinello, M. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 13723−13728.
(32) Wicke, E.; Eigen, M.; Ackermann, H. *Z. Phys. Chem.* **1954**, *1*, 340−364.
(33) Zundel, G.; Metzger, H. *Z. Phys. Chem.* **1968**, *58*, 225−245.
(34) Tuckerman, M. E.; Marx, D.; Klein, M. L.; Parrinello, M. *Science* **1997**, *275*, 817−820.
(35) Valeev, E. F.; Schaefer, H. F. *J. Chem. Phys.* **1998**, *108*, 7197−7201.
(36) Auer, A. A.; Helgaker, T.; Klopper, W. *Phys. Chem. Chem. Phys.* **2000**, *2*, 2235−2238.
(37) Wales, D. J. *J. Chem. Phys.* **1999**, *110*, 10403−10409.
(38) Xie, Y.; Remington, R. B.; Schaefer, H. F. *J. Chem. Phys.* **1994**, *101*, 4878−4884.
(39) Headrick, J. M.; Diken, E. G.; Walters, R. S.; Hammer, N. I.; Christie, R. A.; Cui, J.; Myshakin, E. M.; Duncan, M. A.; Johnson, M. A.; Jordan, K. D. *Science* **2005**, *308*, 1765−1769.
(40) Park, M.; Shin, I.; Singh, N. J.; Kim, K. S. *J. Phys. Chem. A* **2007**, *111*, 10692−10702 PMID: 17910422.
(41) Huang, X.; Braams, B. J.; Bowman, J. M. *J. Chem. Phys.* **2005**, *122*, 044308.
(42) Giannozzi, P.; Baroni, S.; Bonini, N.; Calandra, M.; Car, R.; Cavazzoni, C.; Ceresoli, D.; Chiarotti, G. L.; Cococcioni, M.; Dabo, I.; Dal Corso, A.; de Gironcoli, S.; Fabris, S.; Fratesi, G.; Gebauer, R.; Gerstmann, U.; Gougoussis, C.; Kokalj, A.; Lazzeri, M.; Martin-Samos, L.; Marzari, N.; Mauri, F.; Mazzarello, R.; Paolini, S.; Pasquarello, A.; Paulatto, L.; Sbraccia, C.; Scandolo, S.; Sclauzero, G.; Seitsonen, A. P.; Smogunov, A.; Umari, P.; Wentzcovitch, R. M. *J. Phys.: Condens. Matter* **2009**, *21*, 395502.
(43) Sorella, S. *TurboRVB, Quantum Monte Carlo software for electronic structure calculations*. http://people.sissa.it/∼sorella/web/.
(44) Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1997**, *78*, 1396−1396.
(45) Kühne, T. D.; Krack, M.; Parrinello, M. *J. Chem. Theory Comput.* **2009**, *5*, 235−241.
(46) Troullier, N.; Martins, J. L. *Phys. Rev. B* **1991**, *43*, 1993−2006.
(47) Car, R.; Parrinello, M. *Phys. Rev. Lett.* **1985**, *55*, 2471−2474.
(48) Kato, T. *Comm. Pure Appl. Math.* **1957**, *10*, 151−177.
(49) Holzmann, M.; Ceperley, D. M.; Pierleoni, C.; Esler, K. *Phys. Rev. E* **2003**, *68*, 046707.
(50) Sorella, S.; Casula, M.; Rocca, D. *J. Chem. Phys.* **2007**, *127*, 014105.
(51) Azadi, S.; Cavazzoni, C.; Sorella, S. *Phys. Rev. B* **2010**, *82*, 125112.
(52) Burkatzki, M.; Filippi, C.; Dolg, M. *J. Chem. Phys.* **2007**, *126*, 234105.
(53) Assaraf, R.; Caffarel, M. *J. Chem. Phys.* **2003**, *119*, 10536−10552.
(54) Chiesa, S.; Ceperley, D. M.; Zhang, S. *Phys. Rev. Lett.* **2005**, *94*, 036404.
(55) Badinski, A.; Haynes, P. D.; Trail, J. R.; Needs, R. J. *J. Phys.: Condens. Matter* **2010**, *22*, 074202.
(56) Per, M. C.; Snook, I. K.; Russo, S. P. *Phys. Rev. B* **2012**, *86*, 201107.
(57) Filippi, C.; Umrigar, C. J. *Phys. Rev. B* **2000**, *61*, R16291−R16294.
(58) Sorella, S.; Capriotti, L. *J. Chem. Phys.* **2010**, *133*, 234111.
(59) Attaccalite, C.; Sorella, S. *Phys. Rev. Lett.* **2008**, *100*, 114501.
(60) Zen, A.; Luo, Y.; Sorella, S.; Guidoni, L. *J. Chem. Theory Comput.* **2013**, *9*, 4332−4350.
(61) Pauling, L. *The Nature of the Chemical Bond and the Structure of Molecules and Crystals: An Introduction to Modern Structural Chemistry*; Cornell University Press: Ithaca, NY, 1939.
(62) Casula, M.; Sorella, S. *J. Chem. Phys.* **2003**, *119*, 6500−6511.
(63) Casula, M.; Attaccalite, C.; Sorella, S. *J. Chem. Phys.* **2004**, *121*, 7110−7126.
(64) Marchi, M.; Azadi, S.; Casula, M.; Sorella, S. *J. Chem. Phys.* **2009**, *131*, 154116.
(65) Sterpone, F.; Spanu, L.; Ferraro, L.; Sorella, S.; Guidoni, L. *J. Chem. Theory Comput.* **2008**, *4*, 1428−1434.
(66) Davidson, E. R. *Rev. Mod. Phys.* **1972**, *44*, 451−464.
(67) Foster, J. P.; Weinhold, F. *J. Am. Chem. Soc.* **1980**, *102*, 7211−7218.
(68) Abrams, M. L.; Sherrill, C. D. *Chem. Phys. Lett.* **2004**, *395*, 227−232.
(69) Grüneis, A.; Booth, G. H.; Marsman, M.; Spencer, J.; Alavi, A.; Kresse, G. *J. Chem. Theory Comput.* **2011**, *7*, 2780−2785.
(70) DePrince, A. E.; Sherrill, C. D. *J. Chem. Theory Comput.* **2013**, *9*, 2687−2696.
(71) Sorella, S.; Capriotti, L. *Phys. Rev. B* **2000**, *61*, 2599−2612.
(72) Sorella, S. *Phys. Rev. B* **2001**, *64*, 024512.
(73) Sorella, S. *Phys. Rev. B* **2005**, *71*, 241103.

(74) Umrigar, C. J.; Toulouse, J.; Filippi, C.; Sorella, S.; Hennig, R. G. *Phys. Rev. Lett.* **2007**, *98*, 110201.
(75) Casula, M.; Filippi, C.; Sorella, S. *Phys. Rev. Lett.* **2005**, *95*, 100201.
(76) Casula, M.; Moroni, S.; Sorella, S.; Filippi, C. *J. Chem. Phys.* **2010**, *132*, 154113.
(77) Benedict, W. S.; Gailar, N.; Plyler, E. K. *J. Chem. Phys.* **1956**, *24*, 1139−1165.
(78) Clark, B. K.; Morales, M. A.; McMinis, J.; Kim, J.; Scuseria, G. E. *J. Chem. Phys.* **2011**, *135*, 244105.
(79) Lüchow, A.; Fink, R. F. *J. Chem. Phys.* **2000**, *113*, 8457−8463.
(80) Feller, D.; Boyle, C. M.; Davidson, E. R. *J. Chem. Phys.* **1987**, *86*, 3424−3440.
(81) Bergmann, U.; Di Cicco, A.; Wernet, P.; Principi, E.; Glatzel, P.; Nilsson, A. *J. Chem. Phys.* **2007**, *127*, 174504.
(82) Pavese, M.; Chawla, S.; Lu, D.; Lobaugh, J.; Voth, G. A. *J. Chem. Phys.* **1997**, *107*, 7428−7432.
(83) Sadhukhan, S.; Muñoz, D.; Adamo, C.; Scuseria, G. E. *Chem. Phys. Lett.* **1999**, *306*, 83−87.
(84) Neuscamman, E. *Phys. Rev. Lett.* **2012**, *109*, 203001.
(85) Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S.; Windus, T. L.; Dupuis, M.; Montgomery, J. A. *J. Comput. Chem.* **1993**, *14*, 1347−1363.
(86) Benedek, N. A.; Snook, I. K.; Towler, M. D.; Needs, R. J. *J. Chem. Phys.* **2006**, *125*, 104302.
(87) Gurtubay, I. G.; Needs, R. J. *J. Chem. Phys.* **2007**, *127*, 124306.
(88) Gillan, M. J.; Manby, F. R.; Towler, M. D.; Alfé, D. *J. Chem. Phys.* **2012**, *136*, 244105.