

Energy-consistent pseudopotentials for quantum Monte Carlo calculations

M. Burkatzki¹, C. Filippi² and M. Dolg¹

¹*Institute for Theoretical Chemistry, University of Cologne, Greinstr. 4, 50939 Cologne, Germany*

²*Instituut-Lorentz, Universiteit Leiden, Niels Bohrweg 2, NL-2333 CA Leiden, The Netherlands*

We present scalar-relativistic energy-consistent Hartree-Fock pseudopotentials for the main-group elements. The pseudopotentials do not exhibit a singularity at the nucleus and are therefore suitable for quantum Monte Carlo (QMC) calculations. We demonstrate their transferability through extensive benchmark calculations of atomic excitation spectra as well as molecular properties. In particular, we compute the vibrational frequencies and binding energies of 26 first- and second-row diatomic molecules using post Hartree-Fock methods, finding excellent agreement with the corresponding all-electron values. We also show that our pseudopotentials give superior accuracy than other existing pseudopotentials constructed specifically for QMC. Finally, valence basis sets of different sizes (VnZ with n=D,T,Q,5 for 1st and 2nd row, and n=D,T for 3rd to 5th row) optimized for our pseudopotentials are also presented.

I. INTRODUCTION

Pseudopotentials are widely used in quantum chemistry and condensed matter physics to increase the computational efficiency of ab-initio electronic structure calculations. The chemically inert core electrons are replaced by an angular-momentum-dependent effective potential acting on the valence electrons. This results in a smaller number of explicitly treated electrons and usually a smoother potential, significantly reducing the computational cost for heavy elements. The use of pseudopotentials is well established in density functional theory (DFT), Hartree-Fock (HF) and subsequent highly-correlated quantum chemical calculations¹.

Pseudopotentials are a very important ingredient also in quantum Monte Carlo (QMC) calculations². The scaling of fixed-node diffusion Monte Carlo (DMC) with the number of electrons N is a modest N^4 when compared to a highly-correlated quantum chemical approach such as the coupled-cluster singles and doubles method with a perturbative treatment of the triples (CCSD(T)) which scales as N^7 and is therefore limited in its applicability to small systems³⁻⁶. Unfortunately, the computational cost of DMC increases approximately^{7,8} as $Z^{5.5-6.5}$ with the atomic number Z . Therefore, the use of pseudopotentials to reduce the effective value of Z significantly improves the efficiency of DMC in dealing with heavier atoms and has made feasible in recent years the application of this technique to complex systems.

Many complete sets of ab-initio pseudopotentials are available in the literature but most of them were not generated specifically for use within QMC. For instance, DFT pseudopotentials constructed in combination with the local density and various generalized gradient approximations are readily available but it appears from others⁹ as well as personal experience that the use of DFT pseudopotentials in DMC calculations may lead to significantly larger errors than the use of HF pseudopotentials. Various pseudopotentials generated within HF are also publicly available but generally diverge as $1/r^2$ or $1/r$ at the nucleus since they were constructed for use in quantum chemistry codes without paying attention to smoothness considerations. These singularities will however result in large time-step errors in DMC if the pseudo-valence-orbitals used to construct the determinantal component of the QMC

trial wave function do not have the correct behavior close to the nucleus⁹. Since the orbitals in the trial wave function are usually taken from a DFT, a HF or a small multi-determinant correlated calculation performed with a quantum chemistry code, they are often expressed in terms of Gaussian functions which cannot properly describe the orbitals at small radii in the presence of a divergent pseudopotential. Very large Gaussian basis sets would be required to ameliorate the problem.

Greeff and Lester⁹ and Ovcharenko *et al.*¹⁰ have generated non-singular HF pseudopotentials from Be to Ne and Al to Ar for use in QMC calculations. These pseudopotentials have however several drawbacks as they only cover part of the periodic table, do not include relativistic effects, are not delivered with basis sets, and do not have a p-projector for the second-row elements. More recently, Trail and Needs^{11,12} have generated non-divergent Dirac-Fock spin-orbit averaged relativistic potentials for H to Ba and Lu to Hg. They also proposed a transformation to remove the long-range non-locality which arises from the exchange interaction in HF theory and supplied spin-orbit pseudopotentials. However, their pseudopotentials lack the corresponding basis sets, and do not include a d-projector for the transition-metal and post-transition-metal elements but treat the d-potential as local.

In this paper, we design non-singular energy-consistent scalar-relativistic HF pseudopotentials for the main-group elements, which are especially suitable for QMC calculations. The pseudopotentials are given as a compact expansion over few Gaussian functions multiplied by powers of the electron-nucleus separation and are therefore readily usable in any quantum chemistry package employed to generate the starting trial wave function for the QMC calculation. For all pseudopotentials, we also provide Gaussian basis sets of different sizes. For the 1st and 2nd row elements, the basis sets range from valence-double-zeta up to valence-quintuple-zeta (VnZ, n=D,T,Q,5) while, for the elements of the 3rd and 4th row, we provide VDZ and VTZ basis sets.

We demonstrate the accuracy and transferability of our pseudopotentials through extensive benchmark calculations. We compute the atomic excitation spectra of all elements within Hartree-Fock and the vibrational frequencies and binding energies of 26 first- and second-row diatomic molecules using post Hartree-Fock methods. We find that our pseudopotentials give excellent agreement with the corresponding all-

electron values, and are more accurate than the pseudopotentials constructed specifically for QMC by Lester and coworkers^{9,10} and by Trail and Needs^{11,12}. Finally, we test the performance of our pseudopotentials within QMC for the carbon and silicon atoms, where we obtain a comparable size of the DMC localization error but a higher efficiency than when using the Trail-Needs pseudopotentials.

The paper is organized as follows. In Section II, we describe how we generate the pseudopotentials and basis sets. In Section III, we present a series of benchmark calculations to test the accuracy and transferability of our pseudopotentials. In Section IV, we discuss the size of the localization error and the efficiency of our pseudopotentials in QMC. All equations are given in atomic units (a.u.). The pseudopotential and basis-set information is available in the supplemental material¹³ and from the authors institute homepages^{14,15}.

II. THEORY

For a thorough discussion of pseudopotential theory including relativistic effects, we refer the reader to Ref. 16 and references therein. The electronic valence Hamiltonian for a system with N_v valence electrons and N_n nuclei is given by

$$\mathcal{H}_{\text{val}} = -\frac{1}{2} \sum_i^{N_v} \Delta_i + \sum_{i < j}^{N_v} \frac{1}{r_{ij}} + \sum_i^{N_v} \sum_I^{N_n} V_{\text{pp}}^I(i), \quad (1)$$

where i and j index the electrons, and I the nuclei. $\hat{V}_{\text{pp}}^I(i)$ is an angular-momentum-dependent (semi-local) pseudopotential operator which mimics core-valence repulsion and core-valence orthogonality:

$$\hat{V}_{\text{pp}}^I(i) = V_{\text{loc}}^I(r_{Ii}) + \sum_{l=0}^{l_{\text{max}}} V_l^I(r_{Ii}) \hat{P}_l^I, \quad (2)$$

where \hat{P}_l^I denotes a projection operator on the spherical harmonics centered on the nucleus I :

$$\hat{P}_l^I = \sum_{m=-l}^l |Y_{lm}\rangle \langle Y_{lm}|. \quad (3)$$

Due to orthogonality considerations, the value of l_{max} should at least equal the highest-angular-momentum quantum number present in the core of the atom. The parameters in the local component V_{loc}^I and in the non-local channels V_l^I are determined as described below.

A. Shape- and energy-consistent pseudopotentials

The pseudopotentials presented in this work are constructed following the so-called *energy-consistent* scheme¹⁶. Since most pseudopotentials available in the literature are instead generated according to the *shape-consistent* method, we will briefly review both pseudopotential-generation schemes and illustrate the difference between the two approaches.

In the shape-consistent approach, the all-electron valence orbitals are obtained either in HF^{9-12,17-19} or in Kohn-Sham DFT²¹⁻²³ for the lowest atomic configuration and transformed into node-less pseudo-valence-orbitals which exactly match the all-electron ones beyond a chosen core radius. The pseudo-orbitals are usually constructed to have the same norm as the all-electron orbitals to ensure that the scattering properties of the pseudopotential and the full potential have the same energy variations to first order when transferred to other environments²¹. The pseudopotential is then derived either by inversion of the HF or Kohn-Sham DFT equations^{11,17,21-24} or by direct fit of the all-electron eigenvalues and orbitals beyond some core radius using a parameterized analytical pseudopotential^{9,18}. If inversion is used, the pseudopotential is obtained on a numerical grid and must be fitted to a Gaussian representation for its use in standard quantum chemistry packages. This non-linear fit is a numerically fragile procedure yielding long analytic expansions with large oscillating coefficients^{18,24}. On the other hand, a very compact analytic expression is obtained when using a direct fit to generate the pseudopotential^{9,18}.

The energy-consistent approach follows a rather different philosophy. While a shape-consistent pseudopotential is only determined by the reference atomic configuration in which it was generated and norm-conservation ensures its transferability to some degree, an energy-consistent pseudopotential is obtained by requiring that it reproduces the all-electron valence energies of a number of different configurations for the atom under consideration. A parameterized analytical form for the pseudopotential is chosen and the parameters determined by direct fit of the reference energies. It has been shown²⁵ that energy-consistent pseudopotentials yield a slightly more accurate description of the valence correlation energies than the shape-consistent pseudopotentials of Refs. 19,24.

B. Energy-consistent pseudopotential generation

In this work, the all-electron reference energies to construct our energy-consistent pseudopotentials are calculated at the scalar-relativistic, i.e. spin-orbit interaction free, Wood-Boring Hartree-Fock level of theory²⁶ within the LS coupling scheme. For atomic one-electron systems in s states, the energy-dependent Wood-Boring Hamiltonian yields the same eigenvalues as the Dirac Hamiltonian as well as the exact large components of the eigenfunctions. For other angular momenta, the J-averaged Dirac-Hartree-Fock solutions are approximated. Due to the use of a local potential approximation in the mass-velocity, Darwin and averaged spin-orbit terms, the results are approximate for many-electron atoms. The deviations in energy differences for valence excitations, ionization potentials and electron affinities for elements as heavy as gold from spin-orbit averaged Dirac-Hartree-Fock data using a Dirac-Coulomb-Hamiltonian are typically at most a few hundredths of an electron volt. The main advantage of the Wood-Boring approach is the possibility to perform calculations within the nonrelativistic coupling scheme (i.e. LS cou-

pling for atoms) which is nowadays used by almost all QMC and the majority of quantum chemical studies. Consequently, the adjustment and the subsequent use of the pseudopotential will employ the same coupling scheme. All calculations, both for the generation of the reference energies as well as the actual pseudopotential fit, are performed with a numerical finite-difference code to avoid errors introduced by a finite basis set²⁷.

The all-electron LS-state averaged scalar-relativistic HF energies are computed for a multitude of electronic configurations of the atom under consideration. These configurations include energetically low-lying single and double excitations of the valence electrons as well as the first and second cation if the atom contains enough valence electrons. The anionic ground-state is also included if the anion calculation converges. Convergence of the anion is obtained for the main-group elements to the right of the 3rd column (boron column) inclusive. For the selected configurations, the total valence reference energies are then obtained by subtracting the core energy from the all-electron energy. For example, for silicon where we use a neon core, we subtract the Si^{4+} energy from the energy of the ground- and the excited-state configurations to obtain the total valence energies for each configuration. The configurations used to generate the pseudopotentials can be found for all elements in the supplemental material¹³ and are listed for carbon and silicon in Table I and II, respectively.

We construct the pseudopotentials using the largest possible “noble” core plus a filled d-shell for the elements to the right of the transition metals. In addition, an f-shell is included for the elements Tl, Pb, Bi, Po, At, and Rn. A local potential with no Coulombic singularity is also provided for H and He. To determine the parameters in our pseudopotentials, we proceed as follows. Starting with an analytical pseudopotential guess, we iteratively solve the valence-only HF equations and optimize the pseudopotential parameters to reproduce the calculated total valence energies. This procedure yields very compact analytic pseudopotentials as discussed in the following Section.

C. Functional form of the pseudopotential

Our pseudopotentials are parameterized as a sum of Gaussian functions multiplied by powers of the electron-nucleus separation, which is the standard form accepted by any quantum chemistry package. As discussed in the introduction, the pseudopotential should have no singularities at the nucleus, so we choose as local component (Eq. 2) the following simple expression (we omit the atom index I):

$$V_{\text{loc}}(r) = -\frac{Z_{\text{eff}}}{r} + \frac{Z_{\text{eff}}}{r} \exp(-\alpha r^2) + Z_{\text{eff}}\alpha r \exp(-\beta r^2) + \gamma \exp(-\delta r^2), \quad (4)$$

where Z_{eff} is the effective charge of the nucleus and α , β , γ , and δ are variational parameters. The local component is finite and behaves quadratically for small r .

For a 1st-row element, the pseudopotential is then given by

$$\hat{V}_{\text{pp}} = V_{\text{loc}}(r) + \zeta \exp(-\eta r^2) |0\rangle \langle 0|, \quad (5)$$

where $|0\rangle \langle 0|$ denotes the projection operator with angular momentum $l = 0$ (Eq. 3), and ζ and η are two free parameters. The full s -potential should be repulsive inside the core region and become attractive at large distances as illustrated for the carbon and silicon atoms in Fig. 1. Since we desire a smooth potential, we require that the s -potential has a negative curvature at the origin by imposing the additional non-linear constraint on the variational parameters:

$$\gamma\delta + \zeta\eta > 0. \quad (6)$$

For a second-row element, we introduce a p -projector as

$$\hat{V}_{\text{pp}} = V_{\text{loc}}(r) + \zeta \exp(-\eta r^2) |0\rangle \langle 0| + \mu \exp(-\nu r^2) |1\rangle \langle 1|, \quad (7)$$

which yields two additional variational parameters, μ and ν . Similarly, we require a negative curvature for the p -potential obtaining a second constraint:

$$\gamma\delta + \mu\nu > 0. \quad (8)$$

Analogously, each further projector for rows 3 and 5 (d - and f -cores, respectively) is described by a single Gaussian function, has two additional variational parameters and yields one additional constraint. For this quadratic fitting problem, our pseudopotential-generation code²⁷ was extended to include the DONLP2 solver by Spellucci^{28–30}.

For all elements, the quality of the fit is higher than the one obtained in the construction of the singular energy-consistent pseudopotentials of Ref. 31 even though, in most cases, we employ an even larger set of reference configurations in the fit³². Consequently, our pseudopotentials will be at least as accurate and transferable as those of Ref. 31, and their good performance is demonstrated in Section III.

D. Valence basis sets

For all pseudopotentials, we present basis sets in the supplemental material¹³. The basis sets are constructed to minimize the single-reference CCSD(T) ground-state energy. Since alkaline elements only have one valence electron while at least two electrons are needed to account for correlation in the construction of the basis set, the alkaline basis sets are adjusted by minimization of the dimer energy. The same procedure is also followed for the alkaline-earth elements to avoid that the basis sets become too compact. For all the other elements, the atomic energy is used. The CCSD(T) calculations are performed using the MOLPRO program package³³.

The optimization is carried out in three steps. First, an even-tempered set of 7 to 10 primitives with s - and p -symmetry is generated and optimized to minimize a CCSD(T) calculation for the atom. Second, the primitives for each angular momentum are contracted using the orbital coefficients from

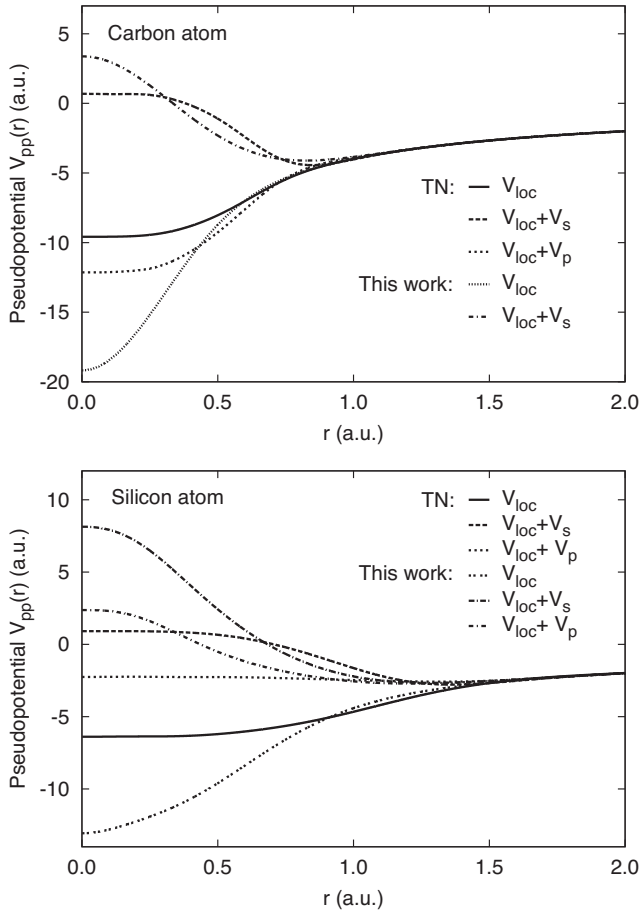


FIG. 1: Our energy-consistent non-singular pseudopotentials and the Trail-Needs (TN) pseudopotentials for the carbon (top) and the silicon (bottom) atom. All values are in a.u.

the HF reference state to yield one contracted s- and one p-function. All V_nZ ($n=D,T,Q,5$) basis sets for one element share the same contracted s- and p-function to reduce the computational demands in the basis-set generation. In a third step, uncontracted primitives are added according to the so-called “correlation consistent” scheme as proposed by Dunning *et al.*³⁴ Since the VDZ basis consists of a (2s 2p 1d) set of functions, the VTZ, VQZ and V5Z basis comprise a (3s 3p 2d 1f), a (4s 4p 3d 2f 1g), and a (5s 5p 4d 3f 2g 1h) set of functions, respectively.

III. TEST CALCULATIONS

The use of a pseudopotential yields significant computational benefits but represents an approximation which introduces a systematic error and must therefore be tested. To establish both the accuracy and transferability of our pseudopotentials, we compute atomic excitation spectra within Hartree-Fock and molecular properties employing well established correlated approaches like second order Møller-Plesset

perturbation-theory (MP2) and the CCSD(T) method. We then address the QMC specific issues of localization error and efficiency.

A. Hartree-Fock atomic excitations

Atomic excitation energies, ionization potentials and electron affinities are well established in the pseudopotential literature as a measure of the pseudopotential quality^{10,11,19,22,31}. Since our pseudopotentials are adjusted to reproduce these properties, they are likely to perform better in these atomic tests than pseudopotentials which are not energy-consistent.

In Table I, we list the errors obtained with our pseudopotential on the excitation energies, ionization potentials and electron affinities of the carbon atom with respect to the corresponding LS-state averaged scalar-relativistic Wood-Boring Hartree-Fock all-electron energies. We also present the errors obtained with the pseudopotentials of Ovcharenko *et al.*¹⁰ and with the pseudopotentials by Trail and Needs^{11,12}. In Table II, we show the same comparison for the silicon atom. All the pseudopotential and the all-electron calculations are performed with a finite-difference HF code²⁷ to avoid errors due to the use of a finite basis set.

As expected, our energy-consistent pseudopotentials perform significantly better than the other pseudopotentials in reproducing the all-electron atomic spectra of both carbon and silicon. For carbon, the mean absolute deviation (MAD) from the all-electron results is 0.002 eV for our pseudopotential, more than one order of magnitude smaller than the MAD for the other pseudopotentials. For silicon, we obtain an MAD of 0.004 eV while the Trail-Needs pseudopotential gives the higher value of 0.012 eV. Not surprisingly, the pseudopotential by Ovcharenko *et al.* fails to properly describe all the atomic configurations of silicon due to the missing p-projector in the pseudopotentials of all 2nd-row elements. For this reason, we will omit these pseudopotentials from our molecular benchmark calculations which focus on the 1st- and 2nd-row elements.

It is important to note that the pseudopotentials by Trail and Needs present a similar problem for the 3rd- and 4th-row elements where they lack a d-projector. Therefore, an analogous analysis for germanium gives an MAD of 0.021 eV for the Trail-Needs pseudopotential compared to an MAD of only 0.008 eV for our pseudopotential. The error increases for the 4th row where, for tin, we obtain an MAD of 0.044 eV when using Trail-Needs pseudopotential compared to an MAD of only 0.008 eV for our pseudopotential.

Finally, since the Trail-Needs pseudopotentials were generated using the Dirac-Coulomb and not the Wood-Boring Hamiltonian, we also compute the all-electron reference energies for carbon at the multi-configuration J-averaged Dirac-Hartree-Fock level. The MAD of the Dirac-Hartree-Fock energies from the all-electron Wood-Boring values is 0.002 eV with a maximum deviation of 0.003 eV. Therefore, the larger errors obtained in the atomic energies of Table I when using the Trail-Needs pseudopotentials are not due to the particular choice of the reference all-electron relativistic Hamiltonian.

Carbon configuration	TN	OAL	This work
$2s^2 2p^2$	0.0000	0.0000	0.0000
$2s^1 2p^3$	0.0696	0.0393	0.0018
$2s^2 2p^1$	0.0127	0.0184	0.0039
$2s^1 2p^2$	0.0451	0.0266	-0.0005
$2s^2 2p^1 3d^1$	0.0127	0.0183	0.0040
$2s^2 3d^1$	0.0487	0.0438	0.0012
$2s^2 2p^1 3s^1$	0.0135	0.0180	-0.0106
$2s^1 2p^2 3d^1$	0.0449	0.0264	-0.0005
$2s^2 3d^2$	0.0486	0.0438	0.0012
$2s^2 4d^1$	0.0493	0.0445	-0.0001
$2s^2 2p^1 4d^1$	0.0127	0.0184	0.0040
$2s^2 2p^3$	-0.0049	0.0141	-0.0011
$2s^2$	0.0499	0.0451	-0.0012
MAD	0.032	0.027	0.002

TABLE I: Errors on the HF excitation energies, ionization potentials and electron affinities for the carbon atom computed with different pseudopotentials. The errors are given with respect to the corresponding all-electron scalar-relativistic Hartree-Fock energies. TN denotes the pseudopotential by Trail and Needs and OAL those by Ovcharenko *et al.* The mean absolute deviation (MAD) is also given. All energies are in eV.

Silicon configuration	TN	OAL	This work
$3s^2 3p^2$	0.0000	0.0000	0.0000
$3s^1 3p^3$	0.0347	-0.0228	-0.0019
$3s^2 3p^1$	0.0105	0.0122	0.0070
$3s^1 3p^2$	0.0298	-0.0239	0.0037
$3s^2 3p^1 3d^1$	0.0091	0.0354	0.0074
$3s^2 3d^1$	-0.0024	0.4206	0.0012
$3s^1 3p^1 3s^1$	0.0104	0.0097	-0.0055
$3s^1 3p^2 3d^1$	0.0272	0.0115	0.0030
$3s^2 3d^2$	-0.0032	0.3986	0.0026
$3s^2 4d^1$	0.0051	0.2071	-0.0025
$3s^2 3p^1 4d^1$	0.0097	0.0260	0.0072
$3s^2 3p^3$	-0.0107	-0.0116	-0.0059
$3s^2$	0.0088	0.0114	-0.0027
MAD	0.012	0.092	0.004

TABLE II: As in Table I but for the silicon pseudopotential. All energies are in eV.

As already mentioned in Sec. II B, the atomic excitations obtained using the J-averaged Dirac-Hartree-Fock solutions are very similar to the Wood-Boring HF results even for elements as heavy as gold.

The HF atomic excitation spectra computed with our pseudopotentials can be found for all elements in the supplemental material¹³. The largest error we obtain is of 0.043 eV for the two-fold ionization potential of astatine.

B. Correlated molecular properties

We test the transferability of our pseudopotentials to correlated calculations by performing MP2 calculations of the equilibrium bond lengths and vibrational frequencies, and CCSD(T) calculations of the binding energies for the 26 1st- and 2nd-row diatomic molecules of the G2-test suite³⁵. The results are then compared to the corresponding valence-only-correlated all-electron values computed using the scalar-relativistic Douglas-Kroll-Hess³⁶⁻⁴¹ (DKH) Hamiltonian. We also compute the same molecular properties using the pseudopotentials by Trail and Needs^{11,12}.

We use basis sets of V5Z quality for our pseudopotentials and construct equivalent V5Z basis sets for the Trail-Needs pseudopotentials. To ensure a fair comparison, we follow the same procedure outlined in Section II D in constructing these large and flexible basis sets for both pseudopotentials. For the all-electron calculations, we employ the V5Z basis sets from Refs. 42–46 in a re-contracted form to be used with the DKH-Hamiltonian as described in Ref. 47. We restrict both the pseudopotential and the all-electron basis sets to g-functions as highest angular-momentum quantum number due to limitations of the program package GAUSSIAN⁴⁸ used for these calculations.

We optimize the bond lengths of the dimers at the V5Z/MP2 level of theory and determine the analytic harmonic frequencies at the same level of theory. The total energies of the dimers are then computed within V5Z/CCSD(T) on the V5Z/MP2 optimized geometries. The atomic ground-state energies are also calculated within V5Z/CCSD(T) and the binding energy obtained as difference of the energy of the dimer and the energies of the two atoms.

In Table III, we compare the calculated pseudopotential bond lengths, frequencies, and binding energies against the all-electron results. The MAD from the all-electron frequencies is 5 cm^{-1} and 10 cm^{-1} for our and Trail-Needs pseudopotentials, respectively. For both pseudopotentials, the largest error occurs for the LiF molecule: Trail-Needs pseudopotentials give an error of 36.33 cm^{-1} whereas the error for our pseudopotential is 22.71 cm^{-1} . For the binding energies, the MAD for our pseudopotentials is 0.5 kcal/mol which is less than half the value obtained with the Trail-Needs pseudopotentials. The largest error for our pseudopotentials is 1.75 kcal/mol for SiO while, for Trail-Needs pseudopotentials, is 4.60 kcal/mol for LiF.

Both sets of pseudopotentials yield very accurate bond lengths with an MAD from the all-electron values of 0.007 Å for the Trail-Needs pseudopotentials and of 0.008 Å for ours. The largest error for our and Trail-Needs pseudopotentials occurs for LiF, and it is equal to 0.031 Å and 0.036 Å, respectively. We note that deviations of the order of 0.01 Å do not have a significant impact on other properties such as the binding energy: For example, a compression of 0.013 Å from the all-electron equilibrium distance causes a change in the binding energy of the P_2 dimer of only 0.05 kcal/mol in a CCSD(T) all-electron calculation. Even in the case of LiF where we have the largest error, a compression of 0.036 Å from equilibrium raises the all-electron CCSD(T) binding

Dimer	Bond length ($\text{\AA}\cdot 10^{-2}$)		Frequency (cm^{-1})		E_{binding} (kcal/mol)	
	TN	This work	TN	This work	TN	This work
LiH	2.84	-0.77	-30.10	-12.00	2.28	0.02
BeH	-0.60	-0.50	1.70	2.85	-0.32	-0.09
CH	0.08	-0.11	-10.6	-5.61	0.55	0.21
NH	0.07	-0.04	-8.24	-5.91	0.62	0.40
OH	0.08	-0.01	-7.54	-4.90	0.52	0.27
HF	0.07	-0.01	-5.96	-4.31	0.48	0.15
HCl	-0.22	-0.51	9.75	5.81	-0.61	-0.50
Li ₂	1.55	-1.31	-6.09	-0.80	0.86	-0.33
LiF	3.51	-3.59	-36.33	-22.71	4.60	-1.28
CN	-0.04	-0.42	-27.04	-20.22	1.98	-0.52
CO	0.12	-0.34	-12.27	-2.88	1.78	-0.94
N ₂	0.09	-0.24	-13.14	-6.77	1.96	-0.24
NO	-0.01	-0.19	34.55	-8.05	2.01	0.28
O ₂	0.33	-0.01	-11.86	0.32	1.49	-0.06
F ₂	0.35	-0.08	-4.52	2.91	0.66	-0.18
Na ₂	-1.59	-1.43	-0.10	-0.51	-0.25	-0.18
Si ₂	-0.94	-1.22	4.77	1.77	-0.77	-0.64
P ₂	-0.97	-1.29	9.14	4.59	-0.83	-0.41
S ₂	-1.04	-1.45	7.38	5.41	-1.17	-0.92
Cl ₂	-0.75	-1.44	2.04	2.48	-0.45	-0.74
NaCl	-2.60	-1.50	4.12	-2.14	-1.33	0.26
SiO	-0.35	-0.97	6.88	-3.07	-2.05	-1.75
CS	-0.32	-0.86	2.36	2.59	-0.22	-1.17
SO	-0.75	-1.07	8.93	6.14	-1.14	-1.56
ClO	-0.47	-1.45	5.03	7.40	-0.12	-1.09
ClF	-0.27	-0.87	-0.31	0.30	0.29	-0.48
MAD	0.7	0.8	10	5	1.1	0.5

TABLE III: Errors on the V5Z/MP2 bond lengths, the V5Z/MP2 vibrational frequencies, and V5Z/CCSD(T) binding energies for 26 dimers using our pseudopotentials and those by Trail and Needs (TN). The errors are given with respect to the corresponding all-electron values computed at the same level of theory using the DKH Hamiltonian to account for relativistic effects.

energy by only 0.17 kcal/mol.

C. Pseudopotential performance in QMC

We address two issues specific to QMC, that is the localization error in DMC and the efficiency in VMC. For all QMC calculations, we use the CHAMP program package⁴⁹.

In the DMC method, one obtains the best energy within the fixed-node approximation, i.e. the lowest-energy state with the same zeros (nodes) as a given trial wave function. In the presence of non-local pseudopotentials, the standard DMC approach cannot be applied and the so-called ‘‘locality approximation’’ is usually employed, which approximates the non-local potential with a new effective core potential by localizing the non-local potential on the trial wave function^{8,50–52}.

	Carbon		Silicon	
	TN	This work	TN	This work
σ	0.281	0.258	0.118	0.104
τ_{corr}	1.16	1.16	1.19	1.17
T_{CPU}	1.00	1.00	1.14	1.00
$\kappa_{\text{our}}/\kappa_{\text{TN}}$	1.19		1.49	

TABLE IV: Relative efficiency $\kappa_{\text{our}}/\kappa_{\text{TN}}$ of our and Trail-Needs (TN) pseudopotentials. T_{CPU} is the computer time for a VMC run of the same length, relative to the time obtained with our pseudopotentials. σ is the root-mean-square fluctuations in a.u. and τ_{corr} the autocorrelation time of the local energy computed in a VMC run.

This new effective potential is explicitly many-body but is local and can be easily incorporated in a DMC algorithm. However, the potential depends now on the trial wave function, and the energy computed in DMC is no longer necessarily variational and depends on the quality of the trial wave function⁵³. As the trial wave function approaches the fixed-node solution obtained without the locality approximation, the DMC energy converges to the correct fixed-node energy.

To estimate the effect of the locality approximation for a given pseudopotential, we employ three different wave functions with the same determinantal component and, consequently, the same nodes, but with different Jastrow factors. We use no Jastrow factor, a two-body, and a sophisticated three-body Jastrow factor⁵⁴ whose parameters are optimized by energy minimization⁵⁵. For the determinantal component, we use a single HF determinant computed with an uncontracted even-tempered single-particle basis of 9 s- and 9 p-functions whose exponents were separately optimized for both pseudopotentials as described in Section II D.

In Fig. 2, we plot the fixed-node DMC energy (E_{FN}) of the carbon and silicon atoms for the three wave functions and for our and Trail-Needs pseudopotentials. As the wave function becomes more accurate, the difference between the variational Monte Carlo (VMC) and the fixed-node energy goes to zero, and the fixed-node energy approaches the value one would obtain without the locality approximation (it can be estimated by extrapolation as $E_{\text{FN}}^{\text{extr}}$). If the pseudopotential were local, the fixed-node DMC energy would not depend on the Jastrow factor, so the slope of the curves in Fig. 2 gives a qualitative idea of the size of the localization error for the different pseudopotentials. We find that the two sets of pseudopotentials perform comparably: For carbon, the Trail-Needs pseudopotential gives a slightly smaller localization error than our pseudopotential while the behavior of the two pseudopotentials is identical for the silicon atom.

Finally, we estimate the relative efficiency of our and Trail-Needs pseudopotentials in a VMC calculation of the total energy obtained as an average the local energy $E_L = \Psi^{-1}\mathcal{H}\Psi$ over the configurations sampled from the square of the wave function Ψ . To quantify the efficiency, we compute the ratio $\kappa_{\text{our}}/\kappa_{\text{TN}}$ with κ defined as

$$\kappa = 1/(\sigma^2 \tau_{\text{corr}} T_{\text{CPU}}), \quad (9)$$

where T_{CPU} is the computational time, σ the root-mean-

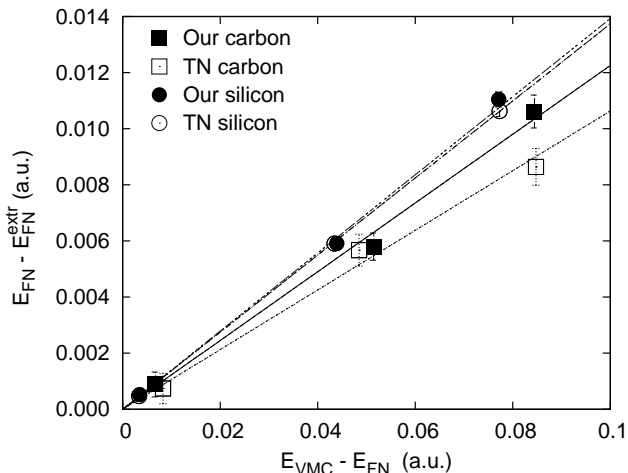


FIG. 2: Fixed-node DMC energies (E_{FN}) for the carbon and silicon pseudo-atoms computed for our and Trail-Needs (TN) pseudopotentials. For both atoms and pseudopotentials, we use as trial wave function a Hartree-Fock determinant with no Jastrow, a two-body and a three-body Jastrow factor. A more accurate wave function corresponds to a smaller difference between the VMC and the DMC fixed-node energies. $E_{\text{FN}}^{\text{extr}}$ denotes the fixed-node energy which is obtained by extrapolation for each system.

square fluctuations of the local energy, and τ_{corr} the autocorrelation time of the local energy calculated as explained in Ref. 56. A larger value of κ corresponds to a smaller product $\sigma^2 \tau_{\text{corr}} T_{\text{CPU}}$ and, therefore, a higher efficiency. The statistical error on the energy decays as $\sigma \sqrt{\tau_{\text{corr}}/M}$ with the number M of Monte Carlo samples, so a smaller $\sigma^2 \tau_{\text{corr}}$ requires a shorter VMC run to obtain the same statistical error on the energy. Then, for a fixed length of the run, one also desires a small value of T_{CPU} , which will depend on how often the non-local component of the pseudopotential must be computed or, equivalently, on how compact in space the non-local components $V_i(r)$ (Eq. 2) are.

In Table IV, we collect the results for the carbon and silicon atoms obtained with our and Trail-Needs pseudopotentials. The pseudopotentials for the two atoms are shown in Fig. 1. As trial wave function, we use a single determinant multiplied by a sophisticated three-body Jastrow factor and optimize all parameters in the orbitals and in the Jastrow factors by energy minimization. We use the same single-particle basis for the orbitals as in the calculations for the localization error. To allow a fair comparison of T_{corr} , we put all pseudopotentials on a grid instead of using their Gaussian form since the expansion in Gaussian functions of the Trail-Needs pseudopotentials is much longer than that of our pseudopotentials. For all pseudopotentials, we do not compute the non-local component if all non-local $V_i(r)$ are smaller than 10^{-4} .

We find that the autocorrelation time τ_{corr} is roughly the same for both pseudopotentials while our pseudopotentials yield smaller fluctuations of the local energy. For silicon, our pseudopotential is more compact giving a saving in computational time: The non-local components $V_s(r)$ and $V_p(r)$ of our pseudopotential are smaller than 10^{-4} beyond a radius of

2.33 Å and 2.38 Å, respectively, while the corresponding radii for Trail-Needs pseudopotential are 2.43 Å and 3.13 Å. From the ratios $\kappa_{\text{our}}/\kappa_{\text{TN}}$, we conclude that, for the carbon and silicon atoms, our pseudopotentials are respectively about 19% and 49% more efficient than Trail-Needs pseudopotentials.

IV. CONCLUSIONS

We present non-singular energy-consistent scalar-relativistic Hartree-Fock pseudopotentials for the main-group elements, and demonstrate their accuracy and transferability by performing extensive benchmarking. We compute the Hartree-Fock atomic excitation spectra of all elements, obtaining excellent agreement with the corresponding all-electron results. To test the transferability to correlated calculations, we use a test-suite of 26 first- and second-row diatomic molecules, and find that our pseudopotentials reproduce the valence-only-correlated all-electron CCSD(T) binding energies and MP2 vibrational frequencies with a mean average deviation of 0.5 kcal/mol and 5 cm^{-1} , respectively. Finally, we provide a variety of basis sets (VnZ with n=D,T,Q,5 for 1st and 2nd row, and n=D,T for 3rd to 5th row) optimized for our pseudopotentials. The basis sets and the pseudopotentials are given in a standard Gaussian representation and, to our knowledge, can be used in every standard quantum chemistry program package which supports pseudopotentials.

Non-singular pseudopotentials for QMC have been published before by two other groups but they are inferior to the pseudopotentials generated in this work. The pseudopotentials by Ovcharenko *et al.*^{9,10} do not account for relativistic effects, cover only a small part of the periodic table, are not delivered with basis sets, and do not provide a p-projector for the second-row elements. The pseudopotentials by Trail and Needs^{11,12} do account for relativistic effects and are available for a large number of elements. However, they also lack the corresponding basis sets, and their transition-metal and post-transition-metal pseudopotentials do not include a d-projector but treat the d-potential as local. In addition, we show that our pseudopotentials give a higher accuracy both when computing the Hartree-Fock excitation spectra and in the correlated calculations of the dimer binding energies and vibrational frequencies, where the average error is two times smaller than the one obtained with the Trail-Needs pseudopotentials. Finally, we test the behavior of our pseudopotentials within QMC for the carbon and silicon atoms, where we recover a comparable size of the DMC localization error but a higher efficiency than when using the Trail-Needs pseudopotentials. The pseudopotential and basis-set data is available from the supplemental material¹³ and from the authors institute homepages^{14,15}.

Future directions of this work include the generation of transition-metal pseudopotentials and the corresponding basis sets.

Acknowledgments

We thank C. Umrigar for useful discussions and a careful reading of the manuscript.

-
- ¹ W. Kutzelnigg, *Physica Scripta*, **36**, 416 (1987)
- ² W. M. C. Foulkes, L. Mitas, R. J. Needs, and G. Rajagopal, *Rev. Mod. Phys.* **73**, 33 (2001).
- ³ F. Coester, *Nucl. Phys.* **7**, 421 (1958).
- ⁴ F. Coester and H. Kümmel, *Nucl. Phys.* **17**, 477 (1960).
- ⁵ J. Cizek, *J. Chem. Phys.* **45**, 4256 (1966).
- ⁶ J. Paldus, X. Li, *Adv. Chem. Phys.* **110**, 1 (1999).
- ⁷ D. M. Ceperley, *J. Stat. Phys.* **43**, 815 (1986).
- ⁸ B. L. Hammond, P. J. Reynolds, W. A. Lester, Jr., *J. Chem. Phys.* **87**, 1130 (1987).
- ⁹ C. W. Greeff and W. A. Lester, Jr., *J. Chem. Phys.* **109**, 1607 (1998).
- ¹⁰ I. Ovcharenko, A. Aspary-Guzik, and W. A. Lester, Jr., *J. Chem. Phys.* **114**, 7790 (2001).
- ¹¹ J.R. Trail and R.J. Needs, *J. Chem. Phys.* **122**, 174109 (2005).
- ¹² J.R. Trail and R.J. Needs, *J. Chem. Phys.* **122**, 014112 (2005).
- ¹³ See EPAPS Document No. for the pseudopotential and basis-set data. This document can be reached through a direct link in the online article's HTML reference section or via EPAPS homepage (<http://www.aip.org/pubservs/epaps.html>).
- ¹⁴ Institute for Theoretical Chemistry, University of Cologne homepage: <http://www.uni-koeln.de/math-nat-fak/tcchem/>
- ¹⁵ Instituut-Lorentz, Universiteit Leiden homepage: <http://www.lorentz.leidenuniv.nl/filippi/>
- ¹⁶ M. Dolg, *Effective Core Potentials* published in *Modern Methods and Algorithms of Quantum Chemistry, 2nd ed.*, **3**, 507 (2000) (John von Neumann Institute for Computing, Jülich, Germany).
- ¹⁷ P. A. Christiansen, Y. S. Lee, and K. S. Pitzer, *J. Chem. Phys.* **71**, 4445 (1979).
- ¹⁸ W. J. Stevens, H. Basch, and M. Krauss, *J. Chem. Phys.* **81**, 6026 (1984).
- ¹⁹ W. R. Wadt and P. J. Hay, *J. Chem. Phys.* **82**, 284 (1985).
- ²⁰ G. P. Kerker, *J. Phys. C* **13**, L189 (1980).
- ²¹ D. R. Hamann, M. Schlüter, and C. Chiang, *Phys. Rev. Lett.* **43**, 1494 (1979).
- ²² G. B. Bachelet, D. R. Hamann, and M. Schlüter, *Phys. Rev. B* **26**, 4199 (1982).
- ²³ N. Troullier and J. L. Martins, *Phys. Rev. B* **43**, 1993 (1991).
- ²⁴ L. F. Pacios and P. A. Christiansen, *J. Chem. Phys.* **82**, 2664 (1985).
- ²⁵ M. Dolg, *J. Chem. Phys.* **104**, 4061 (1996).
- ²⁶ J. H. Wood and A. M. Boring, *Phys. Rev. B* **18**, 2701 (1978)
- ²⁷ Program MCHF77: C. Froese Fischer, *Comput. Phys. Commun.* **14**, 145 (1978); M. Dolg, modified version for pseudopotential and quasi-relativistic calculations.
- ²⁸ P. Spellucci, *Math. Prog.* **82**, 413 (1998)
- ²⁹ P. Spellucci, *Math. Meth. of Oper. Res.* **47**, 355 (1998).
- ³⁰ Program DONLP2: P. Spellucci, available under <ftp://ftp.mathematik.tu-darmstadt.de/pub/-department/software/opti/>
- ³¹ A. Bergner, M. Dolg, W. Küchle, H. Stoll, H. Preuss, *Mol. Phys.* **80**, 1431 (1993).
- ³² unpublished laboratory data provided by M. Dolg
- ³³ MOLPRO Version 2002.3; MOLPRO is a package of ab initio programs designed by H.-J. Werner and P. J. Knowles. The authors are R. D. Amos, A. Bernhardsson, A. Berning, P. Celani, D. L. Cooper, M. J. O. Deegan, A. J. Dobbyn, F. Eckert, C. Hampel, G. Hetzer, P. J. Knowles, T. Korona, R. Lindh, A. W. Lloyd, S. J. McNicholas, F. R. Manby, W. Meyer, M. E. Mura, A. Nicklaß, P. Palmieri, R. Pitzer, G. Rauhut, M. Schütz, U. Schumann, H. Stoll, A. J. Stone, R. Tarroni, T. Thorsteinsson, and H.-J. Werner.
- ³⁴ T.H. Dunning, Jr., K.A. Peterson, and D.E. Woon, Correlation consistent basis sets for molecular calculations, in *Encyclopedia of Computational Chemistry*, edited by P.v.R. Schleyer (Wiley & Sons, 1998).
- ³⁵ L.A. Curtiss, K. Raghavachari, G. W. Trucks, and J. A. Pople, *J. Chem. Phys.* **94**, 7221 (1991).
- ³⁶ M. Douglas and N. M. Kroll, *Ann. Phys.* **82**, 89 (1974).
- ³⁷ B. A. Heß, *Phys. Rev. A* **32**, 756 (1985).
- ³⁸ B. A. Heß, *Phys. Rev. A* **33**, 3742 (1986).
- ³⁹ G. Jansen and B. A. Heß, *Phys. Rev. A* **39**, 6016 (1989).
- ⁴⁰ R. Samzow and B. A. Heß, *Chem. Phys. Lett.* **184**, 491 (1991).
- ⁴¹ R. Samzow, B. A. Heß, and G. Jansen, *J. Chem. Phys.* **96**, 1227 (1991).
- ⁴² T. H. Dunning Jr., *J. Chem. Phys.* **90**, 1007 (1989).
- ⁴³ R. A. Kendall, T. H. Dunning Jr., and R. J. Harrison, *J. Chem. Phys.* **96**, 6796 (1992).
- ⁴⁴ D. E. Woon and T. H. Dunning Jr., *J. Chem. Phys.* **98**, 1358 (1993).
- ⁴⁵ K. A. Peterson, D. E. Woon, and T. H. Dunning Jr., *J. Chem. Phys.* **100**, 7410 (1994).
- ⁴⁶ A. Wilson, T. van Mourik, and T. H. Dunning Jr., *J. Mol. Struct. (Theochem)* **388**, 339 (1997).
- ⁴⁷ W. A. de Jong, R. J. Harrison, and D. A. Dixon, *J. Chem. Phys.* **114**, 48 (2001).
- ⁴⁸ Gaussian 03, Revision C.02, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, and J. A. Pople, Gaussian, Inc., Wallingford CT, 2004.
- ⁴⁹ CHAMP is a quantum Monte Carlo program package written by C. J. Umrigar and C. Filippi, and collaborators; <http://www.ilorentz.org/~filippi/champ.html>.
- ⁵⁰ M. M. Hurley and P. A. Christiansen, *J. Chem. Phys.* **87**, 1069 (1986).
- ⁵¹ P. A. Christiansen, *J. Chem. Phys.* **88**, 4867 (1988).

- ⁵² L. Mitas, E. L. Shirley, and D. M. Ceperley, *J. Chem. Phys.* **95**, 3467 (1991).
- ⁵³ M. Casula, C. Filippi, and S. Sorella, *Phys. Rev. Lett.* **95**, 100201 (2005).
- ⁵⁴ C. Filippi and C. J. Umrigar, *J. Chem. Phys.* **105**, 213 (1996).
As Jastrow correlation factor, we use the exponential of the sum of three fifth-order polynomials of the electron-nuclear (e-n), the electron-electron (e-e), and of pure 3-body mixed e-e and e-n distances, respectively. The Jastrow factor is adapted to deal with pseudo-atoms, and the scaling factor κ is set to 0.6 and 0.5 for carbon and silicon, respectively.
- ⁵⁵ C. J. Umrigar, J. Toulouse, C. Filippi, S. Sorella, and R. G. Hennig, *Phys. Rev. Lett.* **98**, 110201 (2007).
- ⁵⁶ C.J. Umrigar, *Phys. Rev. Lett.* **71**, 408 (1993).