acta physica slovaca vol. 64 No. 5, 501 – 574 **October 2014**

QUANTUM MONTE CARLO FOR NONCOVALENT INTERACTIONS: A TUTORIAL REVIEW*

Matúš Dubecký[1]
Regional Centre of Advanced Technologies and Materials, Department of Physical Chemistry, Faculty of Science, Palacký University Olomouc, tř. 17 listopadu 12, 771 46 Olomouc, Czech Republic

Received 8 March 2015, in final form 2 April 2015, accepted 3 April 2015

Theoretical predictions of noncovalent interaction energies, important e.g. in drug-design or design of hydrogen-storage materials, belong to grand-challenges of contemporary quantum chemistry. In this respect, quantum Monte Carlo (QMC) approaches based on the fixed-node diffusion Monte Carlo (FN-DMC), provide a promising alternative to the commonly used coupled-cluster (CC) methods for their benchmark accuracy, massive parallelism, and favorable scaling. The current tutorial review provides a brief state-of-the-art overview of QMC in ab initio quantum chemical calculations of noncovalent interaction energies, covering recent advances in this field: computational protocols based on FN-DMC with single Slater determinant guiding functions, range of their applicability, tradeoffs, analysis of the success of this simple approach in weakly bound molecular complexes and other related topics. The review is supplemented by an easy-to-grasp practical and detailed tutorial on calculation of molecular interaction energies using a free quantum chemical software (`GAMESS`, `QWalk`). This part thus provides an accessible starting point for the readers interested in practical calculations.

DOI: 10.2478/apsrt-2014-0005

PACS: 02.70.Ss, 03.65.Ge, 05.30.Fk, 31.25.-v, 67.10.Db, 71.10.w, 71.15.-m, 71.20.-b

KEYWORDS: Noncovalent Interactions, Electronic Structure, Quantum Monte Carlo, Correlated Electrons, Fermion Nodes

Contents

[1]E-mail address: matus.dubecky@upol.cz

*Post-processing corrections were made by the author at pages 505, 507, 515, 519, 525, 528, 530, 533, 536-7, 542-3, 546, 548, 553, 557, 560, and 562-8.

1 Introduction

The noncovalent interactions are individually weak and collectively strong interactions between molecules and/or their parts that are abundant everywhere in nature [1,2]. They play an important role in many areas, including but not limited to the stability and function of bio-macromolecules (DNA, RNA, proteins) [3,4], enzyme catalysis [5], molecular recognition [6,7], condensation of noble gases, molecular crystals [8], adsorption of molecules on crystal surfaces [9], properties of liquids, drug design [10,11] or technological applications including hydrogen storage [12,13] and nanotechnology [14]. The most important examples of noncovalent interactions include hydrogen bonding and stacking [3] but also ion-aromatic interactions [15,6], halogen bonding [16], hydrophobic interactions [17, 18] and metallophilicity [19, 20] belong to this class. While the hydrogen bonding is mostly an electrostatic effect, the stacking is dominated by the London dispersion. Compared to a typical covalent bond energy (100 kcal/mol), noncovalent interactions are much weaker and depending on their nature and size of the molecule(s), these interactions may achieve strengths as small as 0.5 kcal/mol (e.g. methane dimer) and up to 30 kcal/mol in small complexes.

For their abundance and technological importance, noncovalent interactions are widely studied. While experiments make it possible to assess their strength [21] by providing e.g. dissociation and adsorption enthalpies [22], they do not provide information on their nature. Here theory complements experiments by providing more detailed information necessary for their fundamental understanding [23]. One can e.g. build models that target specific questions or monitor/obtain quantities not directly available from experiment. Studies combining experiment and theory are therefore more often able to provide fundamental understanding of noncovalent interactions including their nature [1, 3, 24, 22, 25].

Since the nature of noncovalent interactions varies, the final interaction energies depend on a fine interplay of electrostatics, polarization, London dispersion and exchange repulsion, and because their typical scale starts on the order of 1 kcal/mol, their computations require most accurate methods, that are able to describe all contributing effects on equal footing. Only theories able to solve the stationary Schrödinger equation accurately enough are acceptable in this field. In this work we are interested in solutions to the non-relativistic stationary Schrödinger equation in Born-Oppenheimer approximation, i.e. model of electrons in the field of fixed nuclei, that suffices for most problems. Interaction energies require solutions to this equation for the composite system (molecular complex, e.g. water dimer) and its constituents (monomers, e.g. two water molecules for the dimer case). I.e. multiple independent computations (three for a water dimer example) must be performed with accuracy that provides errors not bigger than a few % in the final total energy difference [26] with respect to the exact answer. For small systems, this requirement converts to a level of subchemical accuracy (0.1 kcal/mol) [27] while in larger systems, chemical accuracy (error no bigger than 1 kcal/mol; 1 kcal/mol = 0.035 eV) is considered reliable enough. This task poses a long-standing challenge to modern computational chemistry.

Only a few ab initio many-body methods are able to reach such accuracy. These are referred to as benchmark methods. The most common and very reliable benchmark approach is the coupled-cluster (CC) theory, while a promising alternative for larger systems is the fixed-node diffusion Monte Carlo that is of our interest in this review.

In general, benchmark methods are computationally very demanding and therefore they can be applied to comparatively small molecular complexes. Since one often needs to understand

properties of large models with a large number of particles, e.g. studies of (nanoscopic) water droplets [28,29,30] or large bio-molecular models [4], expensive methods are not directly applicable. Instead, these high-accuracy methodologies are used to benchmark small representative models [31, 32, 26, 33] and such results are subsequently used as a reference for parametrization (or appropriate selection) of cheaper DFT methodologies [34, 13], semi-empirical quantum models [3,35] or classical force fields [36,37].

The most important reference method for ab initio quantum mechanical noncovalent interaction energy computations is CC singles and doubles with perturbative triples [38], known as CCSD(T). This method guarantees a desired benchmark accuracy if large-enough one-particle basis sets and/or complete basis set (CBS) extrapolations are employed [26, 39]. Nevertheless, because of the rapid growth of the CCSD(T) computational cost with a one-particle basis set size K, $\propto O(K^7)$, and even more steeply growing demands of higher-order CC approaches that may be used to assess quality of CCSD(T) itself (e.g. CCSDT(Q) [40,41]), their practical use remains limited to relatively small systems [42, 31, 43, 32, 44, 45, 46, 47]. An advent of local CC methods [48, 49] with an improved scaling may possibly change a situation and lead to applicability of CC methodologies to larger systems than possible to date.

A fixed-node diffusion Monte Carlo (FN-DMC), a member of the quantum Monte Carlo (QMC) class of stochastic approaches provides a promising alternative for its high accuracy, low-order polynomial scaling $\propto O(N^{3-4})$ [50] with the number of electrons N and its intrinsic massive parallelism [50, 51, 52]. Even though the scaling of FN-DMC is favorable, very large prefactor to the CPU cost comes from its statistical nature since error bars of expectation values proportional to $\propto \frac{1}{\sqrt{\text{steps}}}$ must be converged to an acceptable level. Benchmark noncovalent interaction energy calculations that require excessively small error bars on the order of 0.1-1 kcal/mol are therefore very expensive when it comes to large complexes (e.g. buckyball-hydrogen complex) or similar. Their benchmarking is nevertheless possible and required to assess cheaper methodologies since yet no other method is available to this end. Contemporary FN-DMC methodology is able to achieve accuracy of 0.1-0.2 kcal/mol in small trivially saturated complexes [53,54]. Since in large complexes millions of CPU-hours need to be invested in order to get meaningful data [55], further development of algorithms, identification of possible truncations in trial wave function expressions [54, 56] and understanding/ascertaining the limits of FN-DMC methodology is required.

This tutorial review presents an introduction to quantum chemistry and electronic-structure QMC, personal perspective on progress in theoretical estimation of noncovalent interaction energies by QMC limited by the work of the author, and a hands-on tutorial. The text aims to provide a convenient starting point to newcomers that want to study noncovalent interactions between closed-shell complexes by QMC.

The remaining parts of the Introduction (following subsections) define an electronic structure problem, molecular Hamiltonian, Born-Oppenheimer approximation and Slater determinants. Quantum chemical methods typically used to approach the electronic structure problem are then introduced. These include mean-field methods, Hartree-Fock (HF or self-consistent field; SCF) and density functional theory (DFT), and typically used post-HF methods with a possibility of explicit construction of a wave function: configuration interaction (CI) and multi-determinant or multi-configurational SCF (MC-SCF). A main reason for the inclusion of the mentioned methods is their possible use in FN-DMC based approaches as an input. Trial wave functions from ap-

proximate methods are required in FN-DMC to provide nodal surface boundaries (places where the wave function vanishes, $\Psi = 0$) in order to preserve antisymmetry of the simulated electronic state [50, 52]. In addition, the CC methods are briefly introduced in order to provide a possibility of direct comparison with respect to QMC, since both methods are largely used in the field of benchmark noncovalent interaction energy estimations. The rest of this review deals with QMC. A general introduction to QMC methods (Section 2) including diffusion and variational Monte Carlo is followed by a part devoted to specific considerations of QMC usage in the domain of noncovalent interactions (Section 3). The final section (Section 4) provides a detailed and easy-to-grasp hands-on tutorial on the calculation of interaction energy of water dimer by the FN-DMC-based QMC approach.

1.1 Electronic Structure Problem

In molecular electronic structure theory, we are in interested in theoretical understanding of molecules that are atomic-scale objects. Since quantum effects dominate at this scale, the quantum theory must be applied in order to obtain meaningful results. One therefore seeks for solutions to the eigenvalue problem, known as the stationary Schrödinger equation

$$\hat{H}|\Psi_i\rangle = E_i|\Psi_i\rangle, \tag{1.1}$$

that in principle provides access to all physical observables in the conservative quantum system described by the Hamiltonian $\hat{H}$, by virtue of the quantum mechanical expectation value

$$\langle O\rangle = \langle\Psi_i|\hat{O}|\Psi_i\rangle. \tag{1.2}$$

which returns mean estimate of the quantity O (identified with the operator $\hat{O}$) in a system in the state $|\Psi_i\rangle$. The $\hat{H}$ is a total energy operator that defines the set of allowed energies $\{E_i\}$ and the corresponding states $|\Psi_i\rangle$.

In the non-relativistic quantum mechanics of molecules, that mostly suffices in light main-group elements and is of our sole interest here, one defines a so-called molecular Hamiltonian that describes a system of N electrons and M nuclei (without external fields) as

$$\hat{H} = \hat{T}_e + \hat{T}_n + \hat{V}_{ee} + \hat{V}_{en} + \hat{V}_{nn}, \tag{1.3}$$

where (in atomic units, $\hbar = m_e = e = 1$)

$$\hat{T}_e = -\frac{1}{2}\sum_{i=1}^{N}\nabla_i^2 \tag{1.4}$$

is the electron kinetic energy operator,

$$\hat{T}_n = -\frac{1}{2}\sum_{A=1}^{M}\frac{\nabla_A{}^2}{m_A} \tag{1.5}$$

is the nuclear kinetic energy operator,

$$\hat{V}_{en} = -\sum_{i=1}^{N}\sum_{A=1}^{M}\frac{Z_A}{r_{iA}} \tag{1.6}$$

is the operator describing attractive Coulomb interactions of electrons and nuclei,

$$\hat{V}_{ee} = \sum_{i=1}^{N} \sum_{j>i}^{N} \frac{1}{r_{ij}} \tag{1.7}$$

describes Coulomb electronic repulsions, and,

$$\hat{V}_{nn} = \sum_{A=1}^{M} \sum_{B>A}^{M} \frac{Z_A Z_B}{R_{AB}} \tag{1.8}$$

is the nuclear Coulomb repulsion term. The m_A and Z_A is a mass and charge of a nucleus A, $r_{iA} = |\mathbf{r}_i - \mathbf{r}_A|$, $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$ and $R_{AB} = |\mathbf{R}_A - \mathbf{R}_B|$ are the corresponding electron-electron, electron-nucleus and nucleus-nucleus distances, respectively.

Solutions to the stationary Schrödinger equation with a full molecular Hamiltonian (Equation 1.3) are exceedingly complicated for the presence of non-separable interaction terms and contain fully quantum description of electrons and nuclei.

Since the nuclei are much more massive than electrons (in atomic units, $m_A \gg 1$), the complexity of the problem may be significantly reduced by the separation of electron and nuclear motion, i.e.

$$\Psi(\mathbf{R}, \mathbf{R}_n) = \Psi_e^{\mathbf{R}_n}(\mathbf{R}) \Psi_n(\mathbf{R}_n) \tag{1.9}$$

where

$$\mathbf{R} = (\mathbf{r}_1, \mathbf{r}_2, ..., \mathbf{r}_N)$$

is a $3N$ dimensional electron position vector and

$$\mathbf{R}_n = (\mathbf{R}_1, \mathbf{R}_2, ..., \mathbf{R}_M)$$

is a $3M$ dimensional nuclear position vector. Since the electrons are able to accommodate adiabatically in the field of moving nuclei and the nuclei "feel" a mean interaction with electrons, it is useful and typically safe to introduce Born-Oppenheimer (BO) approximation, where the masses of nuclei in the term $\hat{T}_n$ are formally set to infinity. I.e., in BO approximation, the nuclei are stationary and the nuclear kinetic energy is neglected. The BO Hamiltonian for fixed nuclear positions reads

$$\hat{H}^{BO} = \hat{T}_e + \hat{V}_{en} + \hat{V}_{ee} + V_{nn}. \tag{1.10}$$

Note that the V_{nn} is a constant and not an operator here. Since only the electronic terms require solutions to the eigenvalue problem, one formally rewrites the equation to the form

$$\hat{H}^{BO} = \hat{H}_e + V_{nn}. \tag{1.11}$$

and looks for the solutions to the eigenvalue problem (omitting the state label i)

$$\hat{H}_e |\Psi_e\rangle = E_e |\Psi_e\rangle \tag{1.12}$$

with an electronic Hamiltonian

$$\hat{H}_e = \hat{T}_e + \hat{V}_{en} + \hat{V}_{ee}. \tag{1.13}$$

that describes motion of electrons in the field of fixed nuclei. The Equation 1.12 is also known as the electronic structure problem that is central to quantum chemistry.

A total energy in BO approximation is finally obtained as

$$E^{BO} = E_e + V_{nn}. \tag{1.14}$$

The electronic wave functions Ψ_e and the corresponding energies E_e in BO approximation depend parametrically on the nuclear coordinates, i.e. $\Psi_e = \Psi_e^{\mathbf{R}_n}(\mathbf{R})$ and $E^{BO} \equiv E^{BO}(\mathbf{R}_n)$. That is, for each set of nuclear coordinates $\mathbf{R}_n$, one obtains a new Ψ_e and a new corresponding E^{BO}. In this way, one is able to map potential energy surfaces (PESs) of a given set of atoms. Stationary states of PES define important atomic configurations. Minima of the PES define stable isomers while the first order saddle points describe transition states. Energy differences between stationary states of PES describe thermodynamic and kinetic properties of the system of interest. For more details on BO approximation, location of stationary points and PESs in general, the reader may want to consult Ref. 57 and/or excellent textbooks, including Szabo and Ostlund [58], Wilson [59], Jensen [60] and Helgaker et al. [61]. For brevity, in the following text we relabel the BO Hamiltonian $\hat{H}^{BO} \equiv \hat{H}$, the exact Born-Oppenheimer ground-state solution $E^{BO} \equiv E$ and the corresponding electronic wave function $\Psi_e \equiv \Psi$.

The exact solutions to the electronic structure problem (Equation 1.12) are still extremely difficult for large N due to the presence of the many-body electron-electron interaction term $\hat{V}_{ee}$. Therefore, various approximations must be invoked to further simplify the problem. Nearly-exact solutions are only possible for comparatively small systems and are used for benchmark purposes [39].

In general, various approaches to solve the electronic structure problem (Equation 1.12) exist and multiple criteria of method classification may be used. One e.g. distinguishes mean-field and many-body methods, self-consistent and perturbative [61], etc. Below, we will concentrate only on methods that are directly related to the topic of this review. These include ground-state methods able to reach benchmark accuracy (CC) and methods that may provide approximate nodal surfaces ($\Psi = 0$) as an input to QMC. We will briefly discuss mean-field methods, Hartree-Fock (HF) and density functional theory (DFT) [62,63] that rely on one-electron picture. The HF lacks correlation and is thus used as a reference function in high-level methods. DFT, on the other hand, contains an approximate exchange-correlation potential. This theory is therefore much more useful and successful, but its predictivity is not guaranteed. Accuracy of DFT is frequently not satisfactory in systems where the exchange-correlation energy interplay plays a crucial role, e.g. in transition metals [64,65], transition-metal oxides [66], organo-metallic complexes [67,68, 69], bond breaking reactions with complicated open-closed shell re-couplings [70,65] and related potential energy surfaces [59]. Furthermore, DFT in the original formulation lacks dispersion and is thus unable to correctly describe non-covalently bound systems [71, 72].

The post-HF methods (Section 1.7) typically deal with an explicit construction of the wave function as a multi-determinant expansion in terms of excitations from the reference HF ground-state. Determinants serve here as a many-electron basis set (Section 1.2). In general, various

post-HF methods improve upon HF and provide various levels of accuracy ranging from intermediate to very accurate. Below, we discuss some of these approaches: configuration-interaction (CI), multi-configurational SCF (MC-SCF) and coupled cluster (CC) theory. Depending on a problem at hand and computer available, one always resorts to tradeoff.

Finally, a completely different route to solve the Equation 1.12 is provided by a class of stochastic methods known as QMC that are extensively discussed in Sections 2 and 3.

1.2 Slater Determinant

A proper fermionic wave function must be antisymmetric with respect to the permutation of any pair of particles to satisfy the Pauli exclusion principle, i.e.

$$\hat{P}_{ij}\Psi(...,\mathbf{x}_i,...,\mathbf{x}_j,...) = -\Psi(...,\mathbf{x}_j,...,\mathbf{x}_i,...), \tag{1.15}$$

where $\hat{P}_{ij}$ is a permutation operator for a pair of particles i and j. The $\mathbf{x}_i = \{\mathbf{r}_i,\sigma_i\}$ is a four-dimensional composite space-spin coordinate for the electron i.

A convenient and probably a most simple antisymmetric many-electron ansatz for N electrons is the normalized antisymmetrized product of one-particle spin-orbitals known as Slater determinant,

$$\Psi_0(\mathbf{x}_1,\mathbf{x}_2,,...,\mathbf{x}_N) = \frac{1}{\sqrt{N!}}\begin{pmatrix} \chi_1(\mathbf{x}_1) & \chi_1(\mathbf{x}_2) & \cdots & \chi_1(\mathbf{x}_N) \\ \chi_2(\mathbf{x}_1) & \chi_2(\mathbf{x}_2) & \cdots & \chi_2(\mathbf{x}_N) \\ \vdots & \vdots & \ddots & \vdots \\ \chi_N(\mathbf{x}_1) & \chi_N(\mathbf{x}_2) & \cdots & \chi_N(\mathbf{x}_N) \end{pmatrix}, \tag{1.16}$$

where each included spin-orbital χ_i, is occupied by an electron j, i.e. $\chi_i(\mathbf{x}_j)$. Each χ_i may be written as a product of spatial- and spin-dependent function, $\chi_i(\mathbf{x}_j) = \phi(\mathbf{r}_j)\omega(\sigma_j)$. Note that in the non-relativistic framework, the two-state spin coordinate is defined ad hoc so that the Pauli principle is obeyed. In quantum chemistry, one typically expands the spatial part (orbital) as a linear combination of atom-centered Gaussians (Section 1.3) and optimizes the coefficients of this expansion in a self-consistent fashion (cf. Sections 1.5 and 1.7.2).

Since the determinant contains N occupied orbitals (each electron occupies one orbital) it is a matrix of $N \times N$ elements. In practice, a Slater determinant is used in two distinct ways. Either a as a total many-electron wave function ansatz (in HF and DFT, cf. 1.5 and 1.8) or as a many-electron basis set element in post-HF methods (Section 1.7) like CI.

1.3 Algebraic Approximation and Atomic Basis Sets

Since exact analytic expressions of solutions to the many-body electronic structure problem are not known, solution must be expanded in a finite basis set. Typical theories used in quantum chemistry deal with antisymmetrized spin-orbital products, as introduced above. The spatial orbitals ϕ may be expanded numerically either on a grid, or algebraically as an expansion in a finite basis set $\{B_k\}$:

$$\phi_i(\mathbf{r}) = \sum_k C_{ik}B_k(\mathbf{r}). \tag{1.17}$$

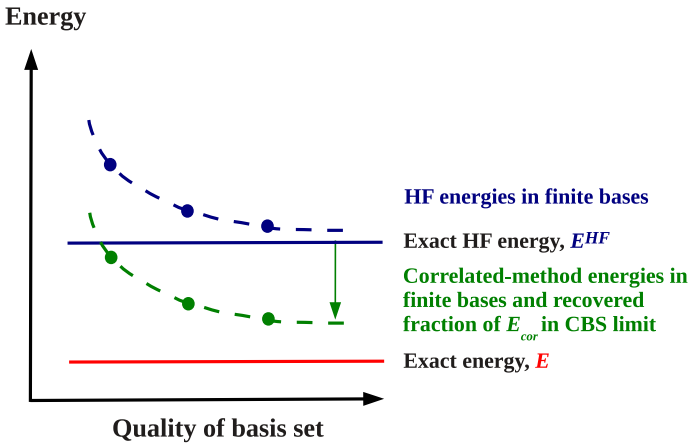

Fig. 1.1. Finite (incomplete) basis sets (circles) allow to approach limit of the given method depending on basis set flexibility. Correlated methods recover certain fraction of correlation energy (Figure 1.4) which depends on the method of choice and saturation of one-particle basis set.

Both approaches allow for a finite matrix formulation convenient for solution on a computer. Quantum chemistry typically relies on an algebraic approach. Here, simple analytic functions with analytic derivatives are preferable for the large number of integral evaluations required in practical computations. A vast part of modern quantum chemistry relies on Gaussian type atom-centered basis sets (atomic basis sets) $G_k(\mathbf{r}) \propto e^{-r^2}$ used to expand orbitals:

$$\phi_i(\mathbf{r}) = \sum_k C_{ik} G_k(\mathbf{r}). \tag{1.18}$$

In addition to possibility of analytical evaluation, Gaussians allow a systematic extension allowing complete-basis set (CBS) extrapolations [73] and the solutions expanded to Gaussians converge rapidly enough to guarantee reasonable charge distributions.

Introduction of a finite (e.g. Gaussian) basis-set inevitably introduces errors: basis-set incompleteness error and possibly also basis-set superposition error (BSSE). The former error relates to the flexibility of a given basis set expansion. Since any finite basis set does not cover the space fully, the more functions we add, the more space electrons have to avoid each other and the total energy expectation value approaches closer and closer the CBS limit of the given method (Figure 1.1). The latter type of error, BSSE, emerges when one calculates differences between the total energy of a super-system AB and its constituents $A + B$ in a small basis set. Since the tails of atom-centered Gaussians of B contribute in the region of A and vice versa, the basis set on A or B in the super system is effectively better than in the case of isolated constituents - this error vanishes in the CBS limit. The counterpoise correction where one computes constituent energies in the basis set of the super system was proposed to eliminate this type of error [74, 75], though this approach was questioned in case of atoms [76].

Despite the mentioned limitations, the Gaussian atom-centered basis sets are the most computationally efficient functions that have been developed and tabulated in various quality for

most of the atoms in the periodic table. For example, visit the EMSL basis set exchange portal at `https://bse.pnl.gov/bse/portal`. For more details and nomenclature of basis sets, the reader is referred to other sources [60, 61].

1.4 Variation Principle

The variation method is an important approach for approximate solution to the eigenvalue problems that cannot be solved exactly [58]. An example being the stationary Schrödinger equation for electrons (Equation 1.12). The variation principle implies that a quantum mechanical total energy expectation value of any approximate normalized trial wave function $|\tilde{\Psi}\rangle$ is an upper bound to the exact solution for a given state, i.e.

$$\langle\tilde{\Psi}|\hat{H}|\tilde{\Psi}\rangle = \tilde{E} \geq E \tag{1.19}$$

and the equality holds only if the trial wave function is identical with the exact solution for the given state, i.e. $\langle\tilde{\Psi}|\hat{H}|\tilde{\Psi}\rangle = E$ only if $|\tilde{\Psi}\rangle = |\Psi\rangle$. It follows that the energy $\tilde{E}$ of an approximate solution $\tilde{\Psi}$ is higher than the exact energy E. The variation method allows one to minimize the total energy expectation value with respect to parameters of a parametric trial wave function and find the lowest value of the total energy that corresponds to the variational estimate of the exact energy [58]. This problem is not easy in general, nevertheless minimization of linear parameters may be reduced to a matrix diagonalization problem.

1.5 Hartree-Fock

The Hartree-Fock (HF) method, also known as self-consistent field (SCF) is a simplest one-electron approximation that properly accounts for the permutation symmetry of fermions. HF relies on a variational optimization of a single-determinant N-electron ground-state wave function. One starts with a Slater determinant (using a shorthand where the normalization factor is omitted and only diagonal elements of the determinant are shown)

$$\Psi_0(\mathbf{x}_1, \mathbf{x}_2, , ..., \mathbf{x}_N) = |\chi_1(\mathbf{x}_1)\chi_2(\mathbf{x}_2)...\chi_N(\mathbf{x}_N)\rangle \tag{1.20}$$

and the spin-orbitals χ_i are varied to minimize the total electronic ground-state energy expectation value

$$E^{HF} = \min_{\{\chi_i\}}\langle\Psi_0|\hat{H}|\Psi_0\rangle \tag{1.21}$$

with a constraint that they remain mutually orthogonal, $\langle\chi_i|\chi_j\rangle = \delta_{ij}$, and the wave function Ψ_0 remains normalized to unity, i.e.

$$\int |\Psi_0|^2 \mathrm{d}\mathbf{x} = 1 \tag{1.22}$$

where $\mathrm{d}\mathbf{x} = \prod_{i=1}^{N} \mathrm{d}\mathbf{x}_i$. A derivation of the minimization with BO Hamiltonian leads to Hartree-Fock equations [77, 58, 78], an eigenvalue problem

$$\hat{f}(\mathbf{x}_i)\chi_k(\mathbf{x}_i) = \epsilon_i\chi_k(\mathbf{x}_i) \tag{1.23}$$

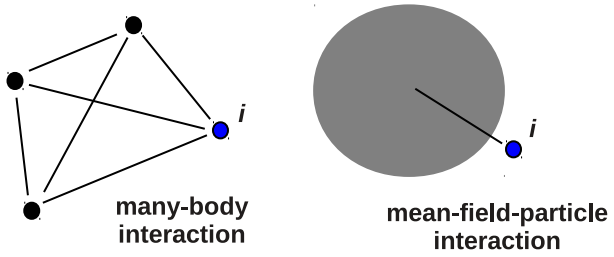

Fig. 1.2. In reality, electron i and other electrons experience many-body interactions with each other, while in a mean-field description, the electron experiences interaction with an average potential from the remaining electrons. A mean-field description thus ignores inter-particle correlations and causes effective increase of the total energy since events where particles meet at the same point, not possible in reality due to their repulsion, are included in average. Recovery of correlation lowers the total energy expectation value with respect to the mean-field, since correlations avoid electron coalescence which is very costly due to the Coulomb repulsion.

which is a set of N coupled non-linear eigenvalue equations instead of one equation for N fully interacting electrons which would be much more difficult to solve. The ϵ_is arise from the orthonormality constraints on single particle orbitals and are identified as energies required to remove the i-th electron from the system [58]. According to Koopmans, the first ionization energy of a molecule equals to the negative energy of the highest occupied molecular orbital (HOMO) [79]. The $\hat{f}(\mathbf{x}_i)$ is an effective one-electron Fock operator

$$\hat{f}(\mathbf{x}_i) = -\frac{1}{2}\nabla_i^2 - \sum_{A=1}^{M} \frac{Z_A}{r_{iA}} + v^{HF}(\mathbf{x}_i) \tag{1.24}$$

contains a mean electronic potential $v^{HF}(\mathbf{x}_i)$ acting on electron i (Figure 1.2) that reads

$$v^{HF}(\mathbf{x}_i) = \sum_j \int \mathrm{d}\mathbf{r}' \frac{|\phi_j(\mathbf{r}')|^2}{|\mathbf{r} - \mathbf{r}'|}\phi_i(\mathbf{r}) - \sum_j \delta_{\sigma_i,\sigma_j} \int \mathrm{d}\mathbf{r}' \frac{\phi_j^*(\mathbf{r}')\phi_i(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}\phi_j(\mathbf{r}). \tag{1.25}$$

The first term in this equation is known as Hartree term and is responsible for the electrostatic repulsion of electron "clouds". The second term is called exchange and has no classical analog [57]. It is a sum over all pairs of orbitals with the same spin projection. Pairs of orbitals with a different spin projections are distinguishable and do not contribute to this quantum term [78]. The effect of exchange term is to keep electrons with the same spin apart so that the Pauli principle is obeyed, giving rise to exchange hole (Fermi hole) containing unit positive charge around each electron [50]. Since the potential exerted on electron i depends on average motion of all remaining electrons, an iterative procedure must be used to solve the problem until the self-consistency is obtained. The name SCF comes from this property.

The physical interpretation of the HF solution (determinant that satisfies Equation 1.21) is that each electron moves independently in orbital χ_i under the average influence of all the other electrons (Figure 1.2) [78] and electron correlations are ignored. The determinant correctly de-

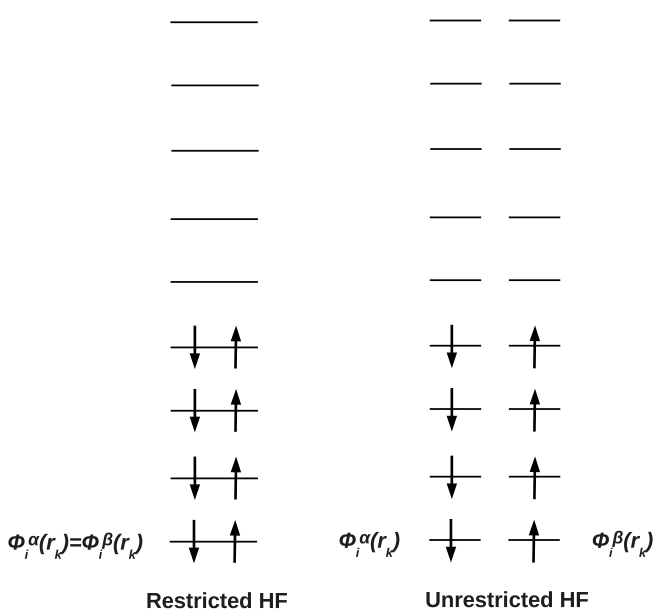

Fig. 1.3. Comparison of restricted HF (RHF) and unrestricted HF (UHF). In RHF (left), the spatial orbitals ϕ are constrained to have the same shape for both spin channels (α and β), while in UHF, this restriction is released (right).

scribes the exchange effects arising from the permutation symmetry of fermions but ignores correlation since there is an unphysical finite probability of finding a pair of electrons with opposite spin in the same region of space (see the next section).

In the non-relativistic HF, the spin state of each orbital must be specified prior to the calculation. If the determinant contains $N_\uparrow$ spin-orbitals with $\sigma_i = \uparrow$ and $N_\downarrow = N - N_\uparrow$ with $\sigma_i = \downarrow$, it is an eigenfunction of $\hat{S}_z$ operator, with eigenvalue $(N_\uparrow - N_\downarrow)/2$ [50]. Since not all states of systems of interest can be described in terms of a single closed-shell determinant, modifications must be made to the mostly used restricted Hartree-Fock (RHF) approach where the spin-orbitals are constrained to have the same spatial distribution for both spin channels (Figure 1.3, left). If the wave function is assumed to be a single determinant in which there is one or more unpaired electrons, there are two possible variants. The restricted open-shell Hartree-Fock (ROHF) [80, 81] which is similar to the RHF or the unrestricted Hartree-Fock (UHF) [82]. In ROHF, all electrons except those explicitly required to occupy open-shell orbitals populate doubly occupied orbitals as in RHF. Such wave functions are eigenfunctions of the spin operator $\hat{S}^2$, but the constraint of double occupancy on orbitals raises the total variational energy. The UHF imposes no restriction on occupancy and each spin channel has its own spatial function space (Figure 1.3, right). However, the UHF wave functions are not eigenfunctions of $\hat{S}^2$ contrary to physical states that must be eigenfunctions of $\hat{S}^2$. UHF solutions contain undesired high spin contributions often referred to as spin contaminants. On the other hand, since the contaminants typically make only a small contribution to the wave function, their effect is ignored [83]. The

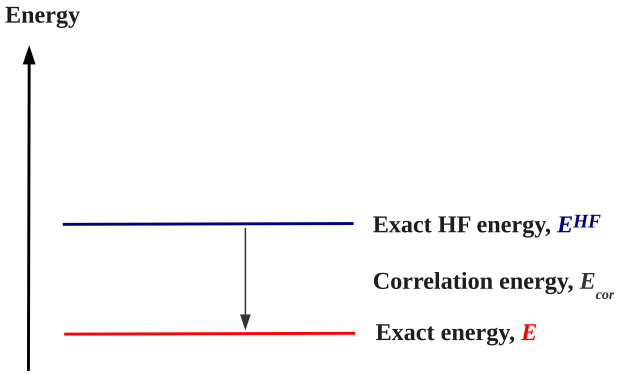

Fig. 1.4. The correlation energy is typically defined as an energy missing at a mean-field level (here HF) with respect to the exact eigenvalue of the Hamiltonian, E.

RHF is unable to describe dissociation of molecules into open shell fragments properly and gives completely wrong predictions in such cases. The UHF provides qualitatively correct picture of dissociations but the resulting potential energy curves are not accurate [58] and the spin symmetry of the wave function is lost [84]. More details on HF may be found in book of Szabo and Ostlund [58]. In order to improve upon HF, electronic correlations must be taken into account. Since the HF lacks electron correlation it is mainly used as a reference for more accurate post-HF calculations (Section 1.7) or as an input trial wave function to QMC (Section 2).

1.6 Correlation Energy

The total energy E^{HF} corresponding to the exact HF solution, i.e. solution of HF equations in the limit of complete basis set, is called a Hartree-Fock limit. Variational property implies that it is an upper bound to the exact ground-state eigenvalue E of the time-independent electronic Schrödinger equation with a molecular Hamiltonian.

The correlation energy E_{cor},

$$E_{cor} = E - E^{HF} \tag{1.26}$$

is a theoretical quantity typically defined as a difference between HF (or possibly other mean-field [85]) and exact solution to the electronic Schrödinger equation (Figure 1.4). It thus measures a deviation of mean-field from the exactness. The definition is ambiguous and vast amount of literature refers to the energy missing with respect to either restricted (Löwdin [86]) or unrestricted HF (Pople and Binkley [87]), i.e. $E^{HF} = E^{RHF}$ or $E^{HF} = E^{UHF}$. Both definitions are useful and widely used in literature depending on needs (cf. also Refs. 88 and 89) but care must be taken so that the comparisons are meaningful since the definitions are non-equivalent.

The E_{cor} is typically a small but very important fraction (typically $\sim 1\%$) of the total energy E. All quantities of interest, e.g. excitation energies, ionization potentials or dissociation and interactions energies are of the same order of magnitude as E_{cor}, and furthermore, valence electron correlation effects frequently dominate these quantities. In order to obtain theoretical quantities

comparable to experiment or meaningful predictivity, one needs to account for the correlation effects (i.e. improve upon HF) and attempt to recover a large fraction of the correlation energy.

Three factors contribute to the correlation of electronic motions [59]. The Coulomb interaction energy diverges whenever two electrons closely approach each other. In those regions of space, a two-electron density matrix (cf. e.g. Pilar [83]) vanishes and is very small nearby. This effect gives rise to a Coulomb hole which is known to have a cusp [90]. The Pauli principle requiring the antisymmetry of the total wave function with respect to the permutation of the electronic coordinates gives a second type of contribution to the total correlation energy. In a single determinant there is no correlation between electrons with opposite spin, therefore there is a finite probability of finding them at the same place. On the other hand, the probability of finding two electrons with the same spin at the same place is zero. This is known as a Fermi or statistical correlation, which gives rise to a Fermi hole. It is deeper than the Coulomb hole, but has no cusp [59]. Statistical correlation of electrons of a single determinant is not regarded as a contribution to E_{cor} [59] (by the definition; Equation 1.26). The point group symmetry of the nuclear coordinate system and the corresponding irreducible representation of the electronic wave function of the system influences the motions of electrons as well. This is a third type of effect that affects the electronic correlation [59].

On the other hand, the electron correlations may be subdivided (and often are) into static and dynamic [91, 85] or dynamical (short-range) and non-dynamical (long-range) [92]. Static correlations, also known as multi-reference character, near-degeneracy or in solid-state community as strong correlation, are present if the exact wave function is dominated by more than one configuration (cf. Section 1.7.1). Such correlations are efficiently accounted for by a small number of important determinants (or configurations) or relaxation of orthogonality constraints (as possible e.g. in generalized valence bond method [85] not discussed here). Dynamical correlations are well described by large numbers of determinants, as e.g. in CI (Section 1.7.1). The determinantal many-body bases with Gaussian orbitals as frequently used is in general not well suited to account for the dynamical correlations in many-body electronic wave functions - the total energies converge very slowly with the number of determinants since each determinant is a mean-field in nature. Therefore, combinatorial numbers (Section 1.7.1) of determinants are required to converge solutions. Inclusion of electron correlation may be accomplished in various ways. Some of the available approaches are described in the following sections (e.g. CI or DFT) and in the next chapter (QMC). Perhaps the most efficient method to account for correlation is to introduce explicit dependence of a wave function on the inter-electronic distances as e.g. in R12 [93, 94] class of methods or Jastrow explicit correlation term as used in QMC (see Section 2.5). Recovery of correlation energy is a topic central to quantum chemistry.

We note here that various aspects of correlation, useful from pedagogical point of view, may be studied in homo-atomic dimers. Here the correlations may be divided on qualitative grounds, into three different groups, to longitudinal ("left-right"), axial and angular, respectively [59]. The former is a tendency to associate electrons to respective nuclei at large separations so they are important for the right description of dissociations. The angular and axial correlation effects (discussed by Hirschfelder and Limmett [95]) become increasingly important when reducing the nuclear separation, they are related to mixing of orbital characters and symmetry type issues . For more details, cf. Ref. 59 and the references included.

1.7 Post-Hartree-Fock Methods

In this part, determinant-based post-HF methods are introduced. These include CI, MC-SCF and CC. These methods improve upon HF in a sense that they recover some unknown part of the correlation energy missing at ~~the~~ a mean-field level. Some of these methods are able to efficiently describe many-electron properties if the one-particle (Gaussian) basis sets are appropriately saturated and enough determinants are used (Section 1.2).

1.7.1 Configuration Interaction

The conceptually simplest method to describe electronic many-body effects is the configuration interaction (CI). The basic idea is to diagonalize N-electron Hamiltonian in a basis of Slater determinants or their linear combinations (rotated spin-adapted many-body basis set) known as configuration functions or configuration state functions (CSF's). The number of determinants that can be constructed for a set of given $2K$ one-electron spin-orbitals is $\binom{2K}{N}$ [58], i.e. the number of configurations increases combinatorially with the number of basis functions K.

In principle, the complete many-electron basis set leads to a full CI expansion in the form

$$|\Psi\rangle = c_0|\Psi_0\rangle + \sum_{ar} c_a^r|\Psi_a^r\rangle + \sum_{\substack{a<b\\r<s}} c_{ab}^{rs}|\Psi_{ab}^{rs}\rangle + \sum_{\substack{a<b<c\\r<s<t}} c_{abc}^{rst}|\Psi_{abc}^{rst}\rangle + \ldots, \tag{1.27}$$

from which one would obtain exact energies of the ground state and of all excited states of the system, an exact solution to the many-electron problem within a given basis set [58]. Restriction on indices (e.g. $a < b, r < s$) in the summation insures, that a given excited determinant is included in the sum only once. The possible determinants in the expansion are the ground state $|\Psi_0\rangle$ (often HF reference), singly excited states $|\Psi_a^r\rangle$ which differ from $|\Psi_0\rangle$ in having the spin-orbital χ_a replaced by χ_r, doubly excited determinants $|\Psi_{ab}^{rs}\rangle$, triply excited $|\Psi_{abc}^{rst}\rangle$, etc., up to N-tuply excited determinants (see Figure 1.5). The coefficients $c_0, c_a^r, c_{ab}^{rs}, c_{abc}^{rst}, ...$, in CI, are varied to minimize the total energy expectation value (by virtue of diagonalization).

If a single configuration dominates the solution, i.e. if one of the coefficients squared approaches $c_i^2 \to 1$, the state is called a single-reference state and the determinant i is called a reference function. On the contrary, if the solution is dominated by more determinants with important weights, we are talking about multireference states. While the full CI (FCI; in physics known as exact diagonalization) describes all correlations on equal footing, truncated CI expansions (see below), must use the whole reference space (e.g. determinants with $c_i^2 > 0.1$) in order to describe the state of interest correctly. That is, in single-reference cases, single and double excitations from one determinant suffice, whereas in multireference cases, all dominant determinants must be considered in the reference space. The latter method is known as multireference CI (MR-CI). The so called FCI-QMC method uses a stochastic sampling of FCI determinant space and provides stochastic estimates of FCI total energy within a given one-particle basis set [96].

In practice, one can handle only a finite set of N-electron trial functions. Truncated CI expansions provide an upper bound to exact solutions. For example, if single and double excitations are included, we have a CI singles and doubles (CISD) method, for which the computational cost scales as $O(N^6)$ [50]. A disadvantage of truncated CI method is single-reference character, the lack of size-extensivity and size-consistency. The size-extensivity problem means incorrect scaling of the total energy with the number of electrons, which prevents one to calculate e.g. binding

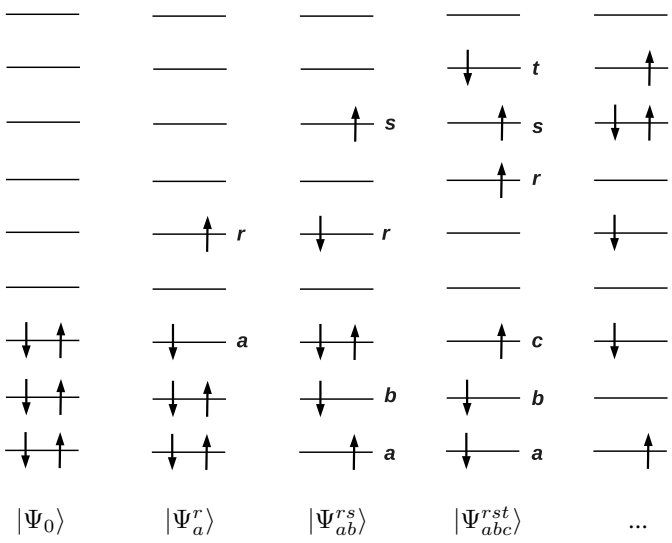

Fig. 1.5. Examples of ground-state and excited determinants (configurations) used in CI expansions (for more details, see text).

energies [57, 50, 51]. The size-consistency problem that is related to size-extensivity is a lack of separability when a total energy of the composite system $E(A+B)$ with A and B separated to infinity is not equal to the sum of the energies for isolated constituents, i.e. the methods where $E(A+B) \neq E(A)+E(B)$ are not size consistent. These failures may be in part corrected by the Davidson correction [97] and the method is then denoted as CISD+Q. Another approach that improves a size-consistency of CISD is known as quadratic CI, QCISD [98] and is also used in noncovalent interaction energy estimations (e.g. Ref. 99).

The optimal CISD expansion recovers some in general unknown fraction of electronic correlations missing at the HF level and may therefore serve as a better qualitative approximation beyond HF. The CI natural orbitals (NO) which diagonalize the first-order density matrix and the corresponding occupation numbers may help to identify the "deadwood". They allow one to select only the most populated orbitals for subsequent calculations with much shorter determinant expansions [100] for higher level theories, e.g. multi-configurational SCF (next section) and/or QMC [101]. The NO's based expansions lead to more rapid convergence with respect to the corresponding canonical orbitals that diagonalize the Fock operator [102, 103].

1.7.2 Multi-Configuration SCF

A next post-HF method, which offers compromise between accuracy and scaling, $O(N^{5-6})$, with a possibility of being size-consistent (if properly treated; cf. below) is a multi-configuration SCF (MC-SCF). It relies on the variational optimization of the expansion coefficients c_i (in CI-like fashion), simultaneously with the SCF improvement of orbitals in determinants Ψ_i in the sum

$$|\Psi^{MC-SCF}\rangle = \sum_i c_i |\Psi_i\rangle. \tag{1.28}$$

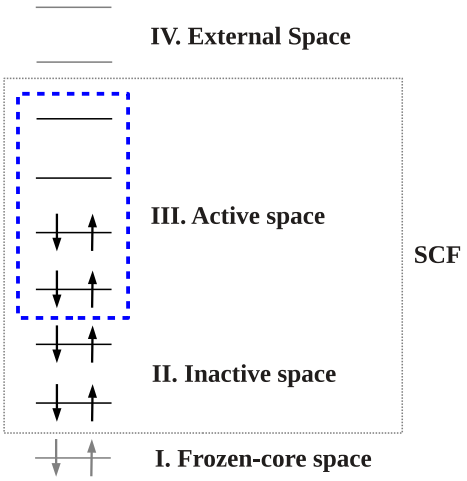

Fig. 1.6. Partitioning of orbital space typically used in MC-SCF and CAS-SCF methods (for more details, see text).

The sum has a truncated CI expansion form.

The computational effort is reduced by partitioning electrons to orbitals of interest (typically taken from previous HF, DFT or CISD calculations) to four groups according to their chemical importance (Figure 1.6):

(I) The doubly-occupied frozen orbitals which are kept intact during the calculation (typically the group is empty or only the core orbitals are included),

(II) the so-called inactive orbitals that are doubly-occupied and SCF optimized (inactive space), and

(III) the orbitals allowed to be CI-like populated (including FCI) and SCF optimized simultaneously with a space II. This space of orbitals is usually called active space.

(IV) The external orbitals (virtuals) serve as a basis for mixing with occupied orbitals in variational SCF procedure.

The size-consistent behavior of the method is achieved by including all chemically relevant orbitals in group III and allowing the maximum order of excitation to be equal to the number of electrons in the group or the number of empty orbitals in the space III (bigger wins). Of relevance are all occupied valence orbitals which play role in the studied process (e.g. bond breaking or excitation) and all corresponding correlating unoccupied orbitals from the reference. This approach is known as the complete-active-space SCF (CAS-SCF) [104]. It allows one to study (at a qualitative level) larger systems compared to the high-level methods of quantum chemistry, but the combinatorial increase of the CPU time and memory requirements are still prohibitive. In CAS-SCF, relatively small active spaces are available (up to about 16 electrons in 16 orbitals). In practical MC-SCF studies, one must often restrict the active-space and consider less electrons than should be considered and even lower the maximum allowed excitation leading to MC-SCF which is not guaranteed to be size consistent. With a proper orbital selection based on chemical intuition, MC-SCF results may be qualitatively correct and useful for interpreta-

tions. Larger CI/CAS spaces are now tractable by virtue of the density matrix renormalization approaches [105, 106, 107, 108].

The CAS-SCF [104] approach is a multi-reference mean-field in a sense that dynamical correlations are mostly omitted. Possible subsequent perturbative treatment (CAS-SCF with MP2 [109], CAS-PT2, or analog known under shorthand MCQDPT [110]) that adds a second-order perturbative correction to the missing correlation is considered to be a benchmark method in quantum chemistry for multi-reference systems of a medium size, but it is known that CAS-PT2 and its analogs overestimate electron correlation and underestimate excitation energies [111]. In addition, prohibitively large one-electron basis sets are needed to converge the results. More viable way to improve upon CAS-SCF is to use the output wave functions in subsequent QMC calculations [111, 112] (see the next section).

1.7.3 Coupled Cluster

The coupled-cluster theory is a post-HF method that relies on an exponential ansatz approximation to the solution of the electronic wave function $|\Psi\rangle$,

$$|\Psi\rangle = e^{\hat{T}}|\Psi_0\rangle. \tag{1.29}$$

Here $|\Psi_0\rangle$ is a reference wave function, typically a single Slater determinant from HF, and $\hat{T}$ is a cluster operator that produces a many-body basis set by virtue of its formal expansion

$$\hat{T} = \hat{T}_1 + \hat{T}_2 + \hat{T}_3 + ... \tag{1.30}$$

where each rank i denotes an excitation order of operator $\hat{T}_i$. For example,

$$\hat{T}_1 = \sum_a \sum_r t_a^s \hat{a}_r^\dagger \hat{a}_a, \tag{1.31}$$

$$\hat{T}_2 = \frac{1}{4} \sum_a \sum_b \sum_r \sum_s t_{ab}^{rs} \hat{a}_r^\dagger \hat{a}_s^\dagger \hat{a}_b \hat{a}_a, \tag{1.32}$$

etc. The sums contain amplitudes t and usual Fermionic creation ($\hat{a}_i^\dagger$) and annihilation operators ($\hat{a}_i$), respectively (cf. Helgaker et al. [61] for a neat introduction to fermion operators and Fock space). Each term in the sums creates one many-body basis function (determinant, cf. Section 1.7.1). Since the exponential $e^{\hat{T}}$ may be rewritten as a Taylor series,

$$e^{\hat{T}} = 1 + \hat{T} + \frac{1}{2}\hat{T}^2 + ... \tag{1.33}$$

one is able to generate various truncated levels of the CC theory. For example, CC singles doubles (CCSD) keeps $\hat{T}_1$ and $\hat{T}_2$ in the $\hat{T}$-expansion and keeps all terms that arise, i.e.

$$e^{\hat{T}_1+\hat{T}_2} = 1 + \hat{T}_1 + \hat{T}_2 + \frac{1}{2}\hat{T}_1^2 + \hat{T}_1 T_2 + \frac{1}{2}\hat{T}_2^2. \tag{1.34}$$

Note that CCSD contains certain quadruply excited configurations from the reference, and is also size consistent and size extensive, contrary to CISD.

Accuracy of CC levels goes roughly (but not necessarily) as [3]

$$\text{CCSD} < \text{CCSD(T)} < \text{CCSDT} < \text{CCSDT(Q)} < \text{CCSDTQ} < ... < \text{FCI}.$$

Depending on the order of truncation one may keep all terms that arise from up to $\hat{T}_3$ (CCSDT), up to $\hat{T}_4$ (CCSDTQ), etc. The price of the CC theory progressively grows with an increasing level of truncation (e.g. CCSD $O(K^6)$, CCSD(T) $O(K^7)$, CCSDT $O(K^8)$, CCSDTQ $O(K^{10})$ with K basis functions) and at the same time, increasing levels of CC progressively converge to exact solutions. In single-reference systems, CCSD with perturbative triples known as CCSD(T) [38] provides an excellent tradeoff between cost and accuracy. CCSD(T) in combination with reasonably saturated basis-sets is used as a high-level benchmark method (e.g. accurate prediction of noncovalent interactions [113]). The noncovalent-interaction benchmark capability of CCSD(T) extrapolated to the complete basis set limit (CCSD(T)/CBS) has been verified on a set of small noncovalent complexes by a higher-order CCSDT(Q) method [26] and in small bases also by the CCSDTQ, CCSDTQP and full CI (configuration interaction) [39]. For more details on CC method, the reader is referred to other sources [61, 114].

Note that the genuine accuracy of CCSD(T) for sizable noncovalent systems has not yet been confirmed to subchemical accuracy at a higher level (CCSDT(Q)), compared to small systems [39]. In large systems, CCSD(T)/CBS may be cross-checked against QMC able to provide accurate and independent interaction energy estimates [53, 54, 55, 115]. Note that extended basis sets (e.g. aug-cc-pVTZ) are unavoidable in CC calculations, because small bases lead to qualitatively incorrect results [26].

1.8 Density Functional Theory

The most successful modern electronic structure one-electron method, DFT [62, 63], is based on the electronic charge density $\rho(\mathbf{r})$ which is a 3-dimensional object in real space (where $\mathbf{r}$ is a 3-dimensional real-space position vector), which is much less complicated quantity compared to the $3N$-dimensional many-body wave function (spin is omitted here)

$$\Psi(\mathbf{r}_1, \mathbf{r}_2, ..., \mathbf{r}_N).$$

It might seem quite a crude an approximation, but in principle, DFT can be formally formulated exactly [62]. The existing formulations are based on the one-particle mean-field picture. The electronic charge density may be obtained from the known wave function by integration as

$$\rho(\mathbf{r}) = N \int \Psi^*(\mathbf{r}, \mathbf{r}_2, ..., \mathbf{r}_N)\Psi(\mathbf{r}, \mathbf{r}_2, ..., \mathbf{r}_N)\mathrm{d}\mathbf{r}_2...\mathrm{d}\mathbf{r}_N, \tag{1.35}$$

where N is the number of electrons and $\Psi(\mathbf{r}_1, \mathbf{r}_2, ..., \mathbf{r}_N)$ is the N-electron wave function.

Hohenberg and Kohn [62] have shown, that there exists a universal functional $F[\rho(\mathbf{r})]$ of the charge density distribution independent of the external potential $V_{\text{ext}}(\mathbf{r})$ that uniquely defines the total energy of the system of electrons

$$E = \int \mathrm{d}\mathbf{r} V_{\text{ext}}(\mathbf{r})\rho(\mathbf{r}) + F[\rho(\mathbf{r})]. \tag{1.36}$$

It achieves a minimum when the charge density $\rho(\mathbf{r})$ coincides with the true ground-state charge density distribution in the external potential $V_{\text{ext}}(\mathbf{r})$ for the non-degenerate ground-states. The functional $F[\rho(\mathbf{r})]$ contains the electron-electron interaction, and hence must contain all many-body effects [57].

In order to construct densities corresponding to fermionic wave functions, Kohn and Sham [63] proposed an ansatz for the charge density as a sum of independent one-particle orbitals ϕ_i that enter single determinant, expressed as

$$\rho(\mathbf{r}) = \sum_i |\phi_i(\mathbf{r})|^2. \tag{1.37}$$

This allows one to evaluate all terms in the equation

$$E[\rho] = T_0[\rho] + \int \mathrm{d}\mathbf{r}\rho(\mathbf{r})\Big(V(\mathbf{r}) + \frac{1}{2}U(\mathbf{r})\Big) + E_{xc}[\rho] \tag{1.38}$$

exactly, except for the exchange-correlation part $E_{xc}[\rho]$. T_0 is the kinetic energy of a system of non-interacting electrons with density $\rho(\mathbf{r})$ and $U(\mathbf{r})$ is the classical Coulomb interaction for electrons. Sum of T, V, and U, covers most of the total energy of the system, and the remaining $E_{xc}[\rho]$ is the unknown exchange-correlation (xc) energy functional, giving a small but important contribution containing E_{cor} (cf. Section 1.6). Most of the chemistry occurs via this small amount and therefore approximation to the $E_{xc}[\rho]$ plays a central role in DFT.

Minimization of Equation (1.38) constrained by normalization

$$\int \mathrm{d}\mathbf{r}\rho(\mathbf{r}) = N \tag{1.39}$$

leads to a functional, which can be minimized by method of Lagrange multipliers in order to find the optimum density $\rho(\mathbf{r})$, i.e. the optimal set of Kohn-Sham orbitals. Solutions have to be found in a self-consistent iterative manner similar to HF (Section 1.5). The effective numerical cost of DFT asymptotically scales as $O(N^{3-4})$, since the dominant step in the iterative procedure is diagonalization of the Kohn-Sham Hamiltonian matrix [57] or orthonormalization of the Kohn-Sham orbitals in iterative-type optimization methods [116].

Many approximations to $E_{xc}[\rho]$ have been proposed. The simplest local density approximation (LDA) introduced by Kohn and Sham [63] assumes that the exchange-correlation energy (and hole) behaves locally as in a homogeneous electron gas (HEG) with the same density [117] and reads

$$E_{xc}^{LDA} = \int \mathrm{d}\mathbf{r}\rho(\mathbf{r})\epsilon_{xc}^{HEG}[\rho(\mathbf{r})], \tag{1.40}$$

with $\epsilon_{xc}^{HEG}[\rho]$ typically taken from parametrized QMC calculations [57]. Further improvements such as the generalized gradient approximation (GGA) are based on the LDA and contain dependence on the semi-local gradient expansion

$$E_{xc}^{GGA} = \int \mathrm{d}\mathbf{r}(\rho_\uparrow, \rho_\downarrow, |\nabla\rho_\uparrow|, |\nabla\rho_\downarrow|, \nabla^2\rho_\uparrow, \nabla^2\rho_\downarrow). \tag{1.41}$$

This class of xc functionals is constructed so that certain sum rules and conditions are satisfied [57]. The most commonly used xc functionals are the PBE [118] and B3LYP [119]. The former is popular in the solid-state community and builds on first principles, the latter, popular in quantum chemistry is of hybrid type, i.e. a part of the exact HF exchange is mixed in to the local/semi-local DFT exchange

$$E_x^{\text{hybrid}} = (1-\alpha)E_x^{DFT} + \alpha E_x^{HF}. \tag{1.42}$$

Hybrid functionals are important in systems with localized valence electrons, such as, for instance, molecules containing transition metal atoms. In fact, in addition to the above-mentioned LDA-, GGA- and hybrid types, there exist many other xc-functionals of different flavors e.g. meta-GGA functionals [120], meta-hybrid functionals [120, 121], double-hybrids [122], dispersion-corrected DFT functionals [123, 124], etc., proposed to fix various failures of the DFT theory.

For many systems, DFT offers quite reasonable accuracy at a low computational cost, but the exact form of the exchange-correlation density functional is unknown and since there is no variational theorem among functionals there is no systematic way to improve them. The only way of improvement is a statistical approach [125] possible with an increasing level of functional sophistication (also cost grows) climbing the empirical Jacob's ladder introduced by Perdew [126]. For the reasons sketched, no guarantee exists that any existing functional will give accurate results for a new unknown system. The GGA functionals can sometimes be less accurate than LDA depending on the formulation and the system of interest. The functionals should be always tested on representative model systems before large scale production calculations [68,69, 13].

Another drawbacks of DFT include missing derivative discontinuity with respect to the number of particles [125], the self-interaction error and lack of dispersion. The self-interaction error means incomplete cancellation between the self-Coulomb term and approximate self-exchange contribution [57]. This problem may lead to unphysical properties e.g. artificial minima on the PESs [57, 65]. In systems, where DFT is not accurate enough, it is often necessary to switch to explicit many-body approaches rather than trying to incorporate many-body effects into xc-functionals in DFT, which is a mean-field method by construction [63, 57].

Since the original formulation of DFT as a mean-field lacks long-range correlations (dispersion), the method is unable to correctly describe non-covalently bound systems. For example, DFT is not able to correctly reproduce weak bonding of noble gas atoms - it often shows repulsive potential at the minimum geometry instead of binding, or minima are much weaker than the correct ones [72]. A next example being a competition between stacking and hydrogen bonding in adenine-thymine DNA base pairs. In hydrogen-bonded and stacked configurations the PBE xc functional completely fails in case of the stacked configuration while the hydrogen bound configuration is qualitatively correct [31]. Last but not least, DFT is unable to reproduce metallophilic interactions between metals that stem from weak attractive interactions between closed d shells. An example being the so-called aurophilicity - a metallophilic Au-Au interaction [19, 127]. It is of interest due to its ability to bind ligated gold centers to various motifs (e.g. chains), making it a versatile tool in materials science [19,127,20,128,129]. In general cases when the dispersion is expected to play important role, dispersion corrected DFT seems to provide a reasonable performance/cost solution [71, 72, 34, 13] while in closed-shell metal complexes with a charge transfer

and metallophilicity, range-separated hybrid functionals [130,131] possibly including dispersion corrections are expected to provide reasonable performance when properly benchmarked.

1.9 Prediction of Noncovalent Interactions

As sketched above, the calculations of noncovalent interactions may be performed at various levels of theory, starting from cheap semi-empirical methods applicable to large systems [71,35], dispersion-corrected DFT [71] suitable for medium to large systems [34], high-level methods like MP2 [109] up to benchmark methods like CCSD(T) [2] and QMC [53] that are very costly. Predictive methods involve either high-level and benchmark methods or proper parametrization for a representative class of systems of interest. In this review, we are primarily interested in benchmark calculations of interaction energies defined as

$$\Delta E = E_D - \sum_i E_i, \tag{1.43}$$

where E_D is the energy of a super system D and E_i is a total energy of its constituent i with a structure as in D, i.e. calculation of ΔE does not account for deformation. Negative ΔE indicates that the system is more stable and one must supply heat to break the system to its constituents. Of our interest are methods able to obtain ΔE with a highest possible accuracy - a benchmark accuracy.

1.10 Benchmark Accuracy

Binding energies of noncovalent complexes are much weaker than a typical covalent bond energy (100 kcal/mol). A typical range of these interactions in small molecules starts from 0.5 kcal/mol and may achieve 30 kcal/mol or more. Since combination of various effects (electrostatics, polarization, dispersion, exchange repulsion) enters the final results, theoretical treatment of noncovalent interactions requires methods of exceptional quality. Only approaches with a degree of accuracy beyond the accepted chemical accuracy (1 kcal/mol) are thus considered satisfactory enough. In small noncovalent complexes (e.g. the set A24 [26]), an ideal benchmark method should provide interaction energies that differ by no more than 0.1 kcal/mol from the exact solutions. This level is called a subchemical accuracy [27] and poses a long-standing challenge to modern computational chemistry.

The accuracy of 1 or 0.1 kcal/mol above is meant formally "per bond" (in small systems). In large systems with multiple "bonds" such accuracy is indeed not achievable and experience shows that relative energy is a good measure of accuracy. The methods able to achieve up to a few % ($<5\%$) in the final interaction energy differences may be considered as benchmark ones, whereas methods with errors below 1% are excellent and may be used to assess benchmark methods [26].

2 Quantum Monte Carlo

The quantum Monte Carlo (QMC) is a class of techniques based on random sampling that are used to solve quantum mechanical problems. Electronic structure problems related to chemistry and solid state physics are within QMC typically treated by the variational Monte-Carlo (VMC) and continuum diffusion Monte Carlo (DMC). The history of QMC began after introduction of the statistical "Monte Carlo method" for solution of multidimensional differential equations by Metropolis and Ulam in 1949 [132]. This work contains the first description of DMC and indicates that stochastic sampling allows for an efficient evaluation of high dimensional integrals that arise in computation of many problems in physics including many-body quantum mechanical expectation values (e.g. Equation 1.12). Soon after, Kalos (1962) [133] realized that the time-independent Schrödinger equation in differential form may be recast to an integral form involving Green's function which may be simulated by real-space random walks [134]. The first VMC calculations for molecular systems were published by Conroy in 1964 [135, 136, 137, 138] and DMC was well established by the work of Anderson and others [139, 140, 141, 142, 143, 144].

Methods of QMC provide highly accurate many-body treatment of quantum effects with a favorable low-order polynomial scaling on the order of $O(N^{3-4})$ [50], where N is the number of electrons. The basic formulations of VMC and fixed-node DMC (FN-DMC; see below) satisfy the variational principle with a zero-variance zero-bias property (exact wave function as a trial function produces constant energy that is an exact eigenvalue of the Hamiltonian). The functional form of trial wave functions accessible by QMC is not restricted as in quantum chemistry, since the expectation values are integrated stochastically in real-space [145]. QMC thus makes it possible to sample sophisticated many-body wave function ansätze with explicit inter-electronic dependencies efficiently. The price to be paid is that the intrinsic statistical error bar ϵ of expectation values is proportional to the inverse square root of the number of statistically independent sampling points i.e. $\epsilon \propto 1/\sqrt{\text{sampling points}}$, so the time to converge the accuracy is typically long. Nevertheless, the methods are intrinsically parallel and can efficiently run on modern supercomputers. On the other hand, the statistical nature of total energies poses a challenge in evaluation of energy derivatives (Section 2.7) in QMC [145] that are substantial e.g. in geometry optimizations and molecular dynamics relying on evaluation of accurate instantaneous forces acting on nuclei. Despite the mentioned limitations, QMC approaches are behind some of the paradigmatic results obtained for quantum systems - e.g. the correlation energy of electron gas [117], widely used in DFT. QMC methods were used to study properties of many systems [50,52] including atoms and molecules [146, 147, 70, 148, 149, 150, 151], solid-state systems [50, 152, 153, 154, 155] with up to thousands of electrons [156, 157], excited states of molecular systems [111, 158, 112, 159] and noncovalent interactions [160, 161, 53, 162, 55, 163, 164] (Section 3).

The VMC method is based on the variational principle and stochastic evaluation of multidimensional integrals. The input parametric trial wave function is optimized in a sequence of steps within a chosen functional form and minimization algorithm involving cost function (e.g. variance, total energy or their linear combination), until the total energy expectation value no longer changes within a prescribed statistical uncertainty. VMC is mostly able to recover about 70-90% of the total correlation energy, which is not enough, and thus the VMC results are typically further improved in DMC (Figure 2.1).

The DMC method is the most popular projector technique, in which a stochastic imaginary-time evolution is used to enhance the ground-state component of the best available starting trial

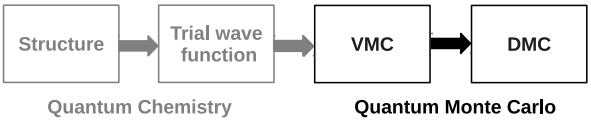

Fig. 2.1. A sketch of a typical multi-stage work-flow of a total energy single-point QMC calculation. An initial molecular model structure and determinant part of the trial wave function are obtained from quantum chemistry. The wave function is then extended by inclusion of an explicit correlation Jastrow factor and variationally optimized at the VMC level. A final production calculations are performed by the FN-DMC.

wave function [50]. The method drives the electron distribution to the bosonic ground-state, and therefore approximations must be made to treat fermions. One way to overcome this problem is to impose a constraint of a fixed nodal surface ($\Psi_T = 0$) as an infinite barrier for electrons, i.e. one imposes the fixed-node (FN) approximation. FN-DMC typically recovers 90-95% of the correlation energy [156, 165] which is very promising. Since the results depend on the trial wave function, reasonable models must be provided as an input in order to obtain high accuracy. FN-DMC is therefore combined with trial wave functions taken from cheaper methods like HF, DFT or CI supplemented by the Jastrow factor and improved in VMC. Practical QMC calculations therefore consist of a subsequent use of different methods that together provide a powerful multistage tool (Figure 2.1). The representation of a wave function as an ensemble of real space electron position vectors in QMC (Figure 2.2) results in a CBS-equivalent mode (for that the walkers may sample the whole space) and energies are thus insensitive to BSSE (Section 1.3; except for a weak dependency possible via Ψ_T) [52]. The real space representation has also another advantages, e.g. the maxima of wave functions in real space allow for a Lewis structure identification useful for chemists [166].

In the following we introduce DMC, FN approximation, importance sampling, FN-DMC, trial wave functions used in QMC and their variational improvement. In addition, a brief discussion of energy derivatives and treatment of effective-core potentials is covered. The next section is devoted to a brief review of the current status of QMC in the domain of noncovalent interactions and discusses the specific aspects of QMC approaches in this field. For an additional information, the reader may want to consult review articles [50, 51, 155, 52, 167].

2.1 Diffusion Monte Carlo

The diffusion Monte Carlo (DMC) is one of the real-space stochastic projector techniques [145] and is a key tool in electronic-structure QMC. It is based on solving the imaginary time many-body Schrödinger equation

$$-\frac{\partial\Psi(\mathbf{R},\tau)}{\partial\tau} = D\nabla^2\Psi(\mathbf{R},\tau) + (E_T - \hat{V})\Psi(\mathbf{R},\tau), \tag{2.1}$$

where Ψ is a solution and E_T is an energy offset. One may formally expand the solution in a complete set of the (unknown) Hamiltonian eigenfunctions $\{\Psi_k\}$ with the corresponding energies $\{E_k\}$,

$$\Psi(\mathbf{R},\tau) = \sum_{k=0}^{\infty} C_k\Psi_k(\mathbf{R})e^{-(E_k-E_0)\tau} = C_0\Psi_0 + \sum_{k=1}^{\infty} C_k\Psi_k(\mathbf{R})e^{-(E_k-E_0)\tau} \tag{2.2}$$

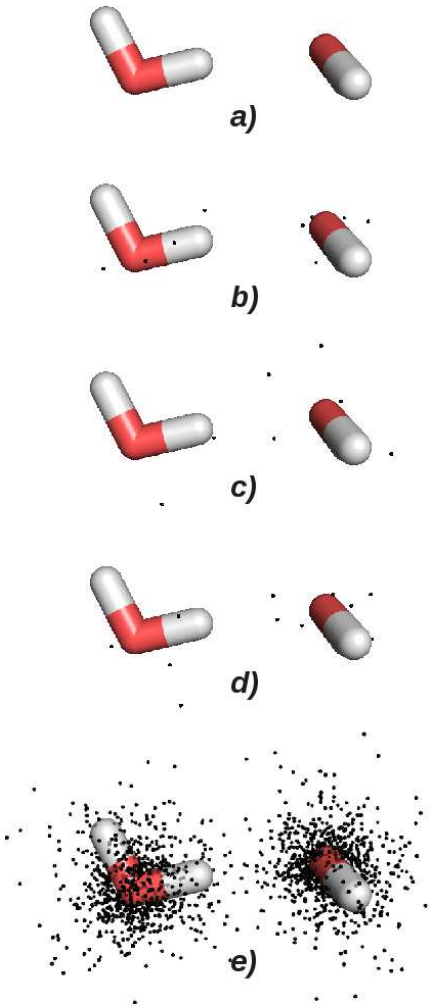

Fig. 2.2. The water dimer structure (a), some representative real-space walker snapshots that sample the electron configuration space (b,c,d), and the stochastic representation of the wave function (e) that consists of a large walker ensemble.

By setting $E_T = E_0$ (i.e. exact energy of ground-state) in Equation 2.2, in the limit of $\tau \to \infty$ the DMC imaginary-time evolution algorithm projects out the stationary ground-state component of the initial trial wave function $\Psi \to C_0\Psi_0$. Contributions of higher states in the sum are exponentially damped in τ. If Ψ is explicitly required to be orthogonal to the ground-state, asymptotic solution projects out an excited state [168]. Note that a solution of imaginary time Schrödinger equation is equivalent to the solution of stationary Schrödinger equation with the same Hamiltonian.

The RHS of Equation 2.1 is the sum of diffusion and rate terms. One may solve the equation

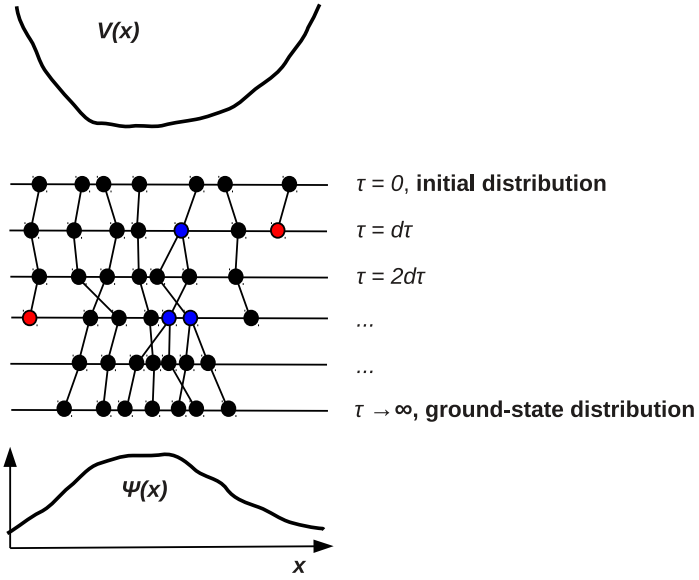

Fig. 2.3. An example of the DMC walker evolution in a one-dimensional harmonic potential well $V(x)$. The initial random positions are drawn from uniform distribution. As the imaginary-time τ propagation proceeds, the distribution converges towards the exact ground state $\Psi(x)$. The walkers disappear (die) in the regions of high potential energy (red) and multiplicate in the low-energy region near the potential minimum (blue).

by evolving ensemble of Brownian particles (Figure 2.2) in space and time in combination with birth/death process (rate) which causes walkers to disappear from regions of high potential and multiplicate around potential minima [139,50,168]. An example of imaginary-time DMC evolution of a walker ensemble in a one-dimensional harmonic quantum well is shown in Figure 2.3.

The propagator (Green's function) used to stochastically solve the stationary electronic Schrödinger equation is derived as follows. The Equation 2.1 may be recast to the integral form

$$\Psi(\mathbf{R}, \tau + \delta\tau) = \int G(\mathbf{R} \leftarrow \mathbf{R}^{'}, \delta\tau)\Psi(\mathbf{R}^{'}, \tau)\mathrm{d}\mathbf{R}^{'}. \tag{2.3}$$

where $G(\mathbf{R} \leftarrow \mathbf{R}^{'}, \delta\tau) = \langle \mathbf{R}|e^{-\delta\tau(\hat{H}-E_T)}|\mathbf{R}^{'}\rangle$ is a Green's function which contains the projection operator and satisfies Equation 2.1 as the wave function, i.e.

$$-\frac{\partial G(\mathbf{R} \leftarrow \mathbf{R}^{'}, \tau)}{\partial \tau} = (\hat{H} - E_T)G(\mathbf{R} \leftarrow \mathbf{R}^{'}, \tau), \tag{2.4}$$

with a boundary condition $G(\mathbf{R} \leftarrow \mathbf{R}^{'}, 0) = \delta(\mathbf{R} - \mathbf{R}')$. It represents the probability of moving particle from point $\mathbf{R}'$ to $\mathbf{R}$ in an imaginary time $\delta\tau$ [145,50].

Except for the special simple cases of limited interest, we do not know the exact expression of the Green's function for interacting particles. Using the Trotter-Suzuki approximation for the

exponential sum of operators $\hat{A}$ and $\hat{B}$,

$$e^{-\tau(\hat{A}+\hat{B})} = e^{-\tau\hat{B}/2}e^{-\tau\hat{A}}e^{-\tau\hat{B}/2} + O(\tau^3) \tag{2.5}$$

and choosing $\hat{A} = \hat{T}$ and $\hat{B} = \hat{V} - E_T$, the Green's function may be recast to

$$\begin{aligned} G(\mathbf{R} \leftarrow \mathbf{R}^{'}, \delta\tau) &= \langle \mathbf{R}|e^{-\delta\tau(\hat{T}+\hat{V}-E_T)}|\mathbf{R}^{'}\rangle \\ &\approx e^{-\delta\tau[\hat{V}(\mathbf{R})-E_T]/2}\langle \mathbf{R}|e^{-\delta\tau\hat{T}}|\mathbf{R}^{'}\rangle e^{-\delta\tau[\hat{V}(\mathbf{R}^{'})-E_T]/2}. \end{aligned} \tag{2.6}$$

A possible and useful form of the approximate Green's function correct for $\delta\tau \rightarrow 0$ [168] contains the known solutions to the diffusion and rate equation (known from radioactive decay) respectively

$$G(\mathbf{R} \leftarrow \mathbf{R}^{'}, \delta\tau) \approx (2\pi\delta\tau)^{-3N/2}\exp\left[-\frac{(\mathbf{R}-\mathbf{R}^{'})^2}{2\delta\tau}\right]\exp[-\delta\tau[\hat{V}(\mathbf{R}) + \hat{V}(\mathbf{R}^{'}) - 2E_T]/2] \tag{2.7}$$

and the error is proportional to $\delta\tau^3$ [50]. The second exponential factor acts as a renormalization and accomplishes the birth/death process. The population of walkers (typically about $10^3 - 10^4$) fluctuates around the desired mean as it is controlled by E_T [139,50]. More details may be found e.g. in Ref. [50].

Since $\Psi(\mathbf{R}, \tau)$ is identified as the density of walkers that must be non-negative everywhere in space, the basic DMC evolution converges to the bosonic ground-state. An additional approximation must be made in order to allow for a simulation of fermions, or in other words to overcome the so-called fermionic sign problem. One of the viable options is the fixed-node approximation. For an overview of alternative methods, we refer the reader to the book of Hammond et al. [145].

2.2 Fixed-Node Approximation

A basic idea of the fixed-node (FN) approximation is to divide the electronic configuration space by a nodal surface of an available trial wave function Ψ_T to regions (pockets) with the same sign and solve the Equation 2.1 with this constraint, which effectively bosonizes the Fermi problem [51]. One can show that the total energy bias in FN approximation is proportional to the square of the nodal displacement error [51] which is a favorable property. Let Ψ_0 be the Fermi ground state and Φ_0 the FN solution. Let us define

$$\Psi_0 = \Phi_0 + \lambda\Phi_1 + \lambda^2\Phi_2 + ... \tag{2.8}$$

where λ is the nodal perturbation and Φ_1, Φ_2, ... are the first-, second-order, etc. wave functions orthogonal to Φ_0 with a fixed-node energy E_{FN}. The first-order energy vanishes due to the orthogonality, i.e.

$$E^{(1)} = \langle \lambda\Phi_1|\hat{H}|\Phi_0\rangle + \langle \Phi_0|\hat{H}|\lambda\Phi_1\rangle = 2\lambda E_{\mathrm{FN}}\langle \Phi_1|\Phi_0\rangle = 0. \tag{2.9}$$

The first nonzero correction is of the second-order in the nodal displacement error [145]. Because of this property, the FN-DMC is expected not to be critically sensitive to the accuracy of nodal surfaces that are reasonably close to the exact solution.

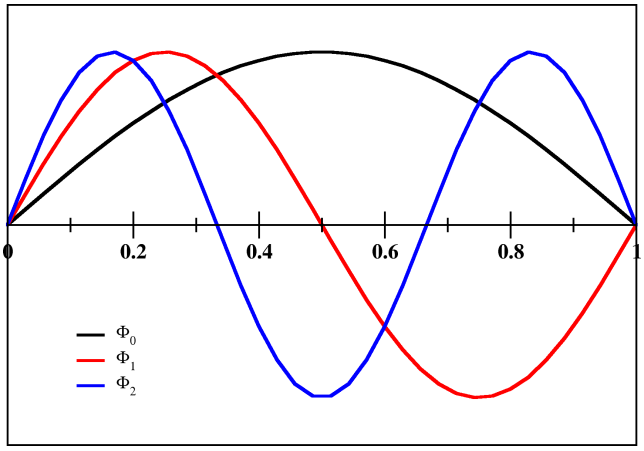

Fig. 2.4. Three lowest-energy one-dimensional particle-in-a-box states/orbitals.

Note that the nodes of the one-electron wave functions (orbitals) have nothing to do with the nodes of the antisymmetric many body wave function. Consider an example of three one-dimensional orbitals, as shown in Figure 2.4. In case of a ground state, the ϕ_0 is node-less, the orbital ϕ_1 has a node right at the center of a box, while the next higher-energy orbital ϕ_2 has two nodal points. Two-electron configuration space is easy to visualize and well demonstrates properties of the nodes of the total fermionic wave function when compared to the orbital product or bosonic case (not shown).

A two-electron product wave function, $\Psi(\mathbf{r}_1, \mathbf{r}_2)$, for two same-spin electrons (triplet pairing) that occupy one-particle states ϕ_0 and ϕ_1 reads

$$\Psi(\mathbf{r}_1, \mathbf{r}_2) = \phi_0(\mathbf{r}_1)\phi_1(\mathbf{r}_2) \tag{2.10}$$

and its graph is visualized in Figure 2.5 (top left). On the other hand, the wave function in form of an antisymmetrized product (norm omitted)

$$\Psi(\mathbf{r}_1, \mathbf{r}_2) = \phi_0(\mathbf{r}_1)\phi_1(\mathbf{r}_2) - \phi_1(\mathbf{r}_1)\phi_0(\mathbf{r}_2) \tag{2.11}$$

reveals picture as rendered in Figure 2.5 (bottom left). It indeed shows that the nodes of the antisymmetrized product lie in places where the product function and its permuted counterpart equal. That is, the nodes of an antisymmetrized product occur primarily at places where the one-particle orbitals have significant non-zero values and equal each other. In addition, two similar examples where the particle/s occupy higher states are shown in Figure 2.5 (middle, right).

2.3 Importance Sampling

A simple DMC algorithm outlined above is extremely inefficient mainly for the renormalization factor $\exp[-\delta\tau(V(\mathbf{R})+V(\mathbf{R}')-2E_T)/2)]$ present in Equation (2.7). It causes large fluctuations of the population and so does the local energy, from step to step, since the rate term diverges whenever two electrons or electron and nucleus approach too close to each other. This leads to

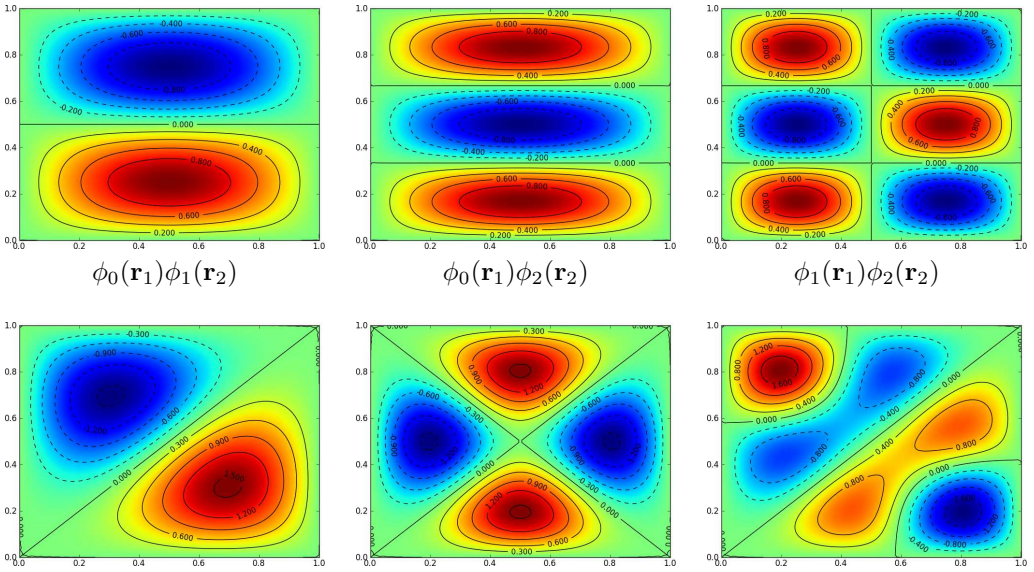

Fig. 2.5. The total spin-polarized (triplet) wave functions $\Psi(1,2)$ for two particles. The top panels show product wave functions while the bottom panels show the corresponding antisymmetrized product wave functions (the norm is omitted). The orbital indices refer to the Figure 2.4. The figure illustrates, that while in the case of products (top) the one-electron orbital nodes directly enter the total wave functions, in the antisymmetrized case (bottom), the positions of the nodes are not directly related to the nodes of one-electron orbitals.

an extremely poor statistical behavior [169] and makes the renormalization process "ill defined" [50].

The difficulties may be overcome by the importance-sampling transformation provided with help of a guiding wave function $\Psi_T(\mathbf{R})$ [170] that must be supplied as an input to the DMC method. If we consider the mixed distribution $f(\mathbf{R},\tau) = \Psi(\mathbf{R},\tau)\Psi_T(\mathbf{R})$ and substitute it into Equation 2.1 multiplied by $\Psi_T(\mathbf{R})$, we obtain equation in the form

$$-\frac{\partial f(\mathbf{R},\tau)}{\partial \tau} = -\frac{1}{2}\nabla^2 f(\mathbf{R},\tau) + \nabla.[\mathbf{v}_D(\mathbf{R})f(\mathbf{R},\tau)] + [E_L(\mathbf{R}) - E_T)]f(\mathbf{R},\tau), \quad (2.12)$$

where $\mathbf{v}_D(\mathbf{R})$ is the $3N$-dimensional drift velocity defined by

$$\mathbf{v}_D(\mathbf{R}) = \nabla \ln|\Psi_T(\mathbf{R})| = \frac{\nabla\Psi_T(\mathbf{R})}{\Psi_T(\mathbf{R})}, \quad (2.13)$$

and

$$E_L(\mathbf{R}) = \Psi_T(\mathbf{R})^{-1}\hat{H}\Psi_T(\mathbf{R}) \quad (2.14)$$

is the local energy [50]. This formulation implicitly imposes the FN approximation [139]. Evolution of the sampling in imaginary time generates the distribution $f(\mathbf{R}, \tau)$ in which the $\Psi(\mathbf{R})$ is the lowest energy stochastic wave function within the nodes as given by Ψ_T.

A consequence of the transformation is that the density of walkers is enhanced in the regions where Ψ_T is large and vice versa, since the drift velocity $\mathbf{v}_D(\mathbf{R})$ drives the walker population in the direction of increasing $|\Psi_T|$. The factor $(\hat{V}(\mathbf{R}) - E_T)$ occurring in Equation 2.4 is replaced by $(E_L(\mathbf{R}) - E_T)$ which is much smoother. Although the exact $\Psi_T = \Psi$ is unknown, divergences in the local energy may be eliminated by choosing a Ψ_T which satisfies Kato cusp conditions [171].

The importance sampling diminishes fluctuations and improves the efficiency of sampling by orders of magnitude. Without this improvement, DMC simulations involving many electrons would not be possible [50].

2.4 Fixed-Node Diffusion Monte Carlo

The DMC with a FN approximation as imposed by the importance sampling transformation, FN-DMC, is a key tool in electronic structure QMC. The quality of FN-DMC results depends on the quality of nodal boundaries of the FN-approximation, within which the FN-DMC provides a rigorous upper bound to the exact energy [51, 168, 172]. Although not exact, in practice, FN-DMC gives very accurate energies [50] and typically recovers 90-95% of the correlation energy [156, 165]. Nevertheless, the quality of the nodal surfaces is necessary prerequisite. A recent dependence of FN-error on electron density and node non-linearity [173] enables better estimation of FN-errors.

In the FN-DMC simulation (Figure 2.6), walkers are sampled from the mixed distribution $\Psi(\mathbf{R}, \tau)\Psi_T(\mathbf{R})$ which allows to share the trial nodes from Ψ_T [51]. They are evolved in a way as described in the Section 2.1. After every DMC step, the sign of the wave function is checked and every walker that crosses a node is deleted (Figure 2.6). The algorithm projects out the exact ground state of each region separately making no assumptions on the shape of the wave function [50].

In practice, the DMC simulation proceeds as follows. First, a reasonably large ensemble is generated from the distribution $|\Psi_T|^2$ using VMC. The ensemble is then evolved in FN-DMC according to the short-time approximation to the Green's function of the importance-sampled imaginary-time Schrödinger equation (Equation 2.7) with a time step chosen so as to give high acceptance ratio (>99%). Note that high acceptance implies that a single walker propagation is highly correlated from step to step - the serial correlation is avoided by taking large block of moves and only average energies per block are taken to total averages instead of taking each step. The DMC process involves biased diffusion since the importance sampling directs the walkers to parts of the configuration space where $|\Psi_T|$ is large. The better the approximate Green's function, the larger the time step that can be used without a detectable bias [174]. After a period of equilibration when excited state contributions mostly die out, walkers trace out the probability distribution $f(\mathbf{R})/\int f(\mathbf{R})\mathrm{d}\mathbf{R}$, and statistics accumulation stage may follow. The

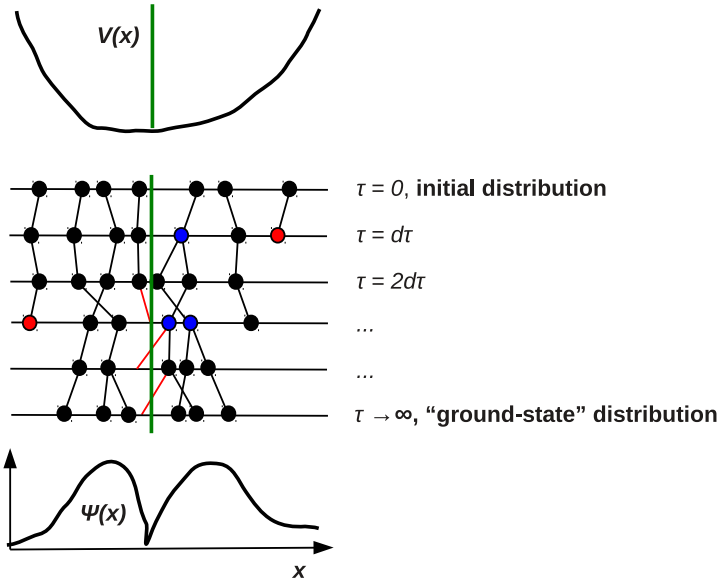

Fig. 2.6. An example of the FN-DMC walker evolution in one-dimensional harmonic potential well $V(x)$. The initial distribution of walker positions is uniform. As the imaginary-time τ propagation proceeds, the distribution converges towards the "ground-state" within the given nodal boundary, i.e. to the first excited state $\Psi_1(x)$, since the node (green) restricts evolution of the ensemble to the genuine ground-state by virtue of step-rejection (red lines). The walkers disappear in the regions of high potential energy (red), occur and multiplicate around the potential minimum (blue).

FN-DMC energy estimator is given by

$$E_{\mathrm{DMC}} = \frac{\int f(\mathbf{R}) E_L(\mathbf{R}) \mathrm{d}\mathbf{R}}{\int f(\mathbf{R}) \mathrm{d}\mathbf{R}} \tag{2.15}$$

which tends to the exact value with an increasing accuracy of the nodal surface of Ψ_T. The FN-DMC algorithm is highly accurate and satisfies the zero-variance zero-bias principle [175], i.e., the variance of the energy and its bias goes to zero as the trial wave function approaches the exact eigenstate [169, 175].

In summary, FN-DMC is a highly accurate method suitable for systems with large number of electrons allowing one to study ground- and excited states of many-electron systems.

2.5 Trial Wave Functions

Quality of trial wave functions controls the statistical sampling efficiency in QMC, and limits the accuracy: FN-DMC simulations are limited by the quality of input nodal surfaces and provides exact solutions within the given boundaries. The trial wave functions should therefore be as accurate and easy to evaluate as possible.

In post-Hartree-Fock methods (Section 1.7) that are typical for high level quantum chemistry, it is common to express many-electron wave function as a linear combination of determinants

containing Gaussian expanded orbitals which do not satisfy Kato cusp conditions [171]. The cusp conditions arise as limiting conditions due to cancellation of divergences in potential and kinetic energies in the Schrödinger equation [50] and the exact wave functions satisfy them exactly. For electrons close to the nucleus, the cusp conditions require

$$\frac{1}{\Psi}\frac{\partial \Psi}{\partial r_{iA}}\bigg|_{r_{iA}=0} = -Z_A. \tag{2.16}$$

The electron-electron cusp conditions read

$$\frac{1}{\Psi}\frac{\partial \Psi}{\partial r_{ij}}\bigg|_{r_{ij}=0} = \frac{1}{2}, \tag{2.17}$$

for anti-parallel spins, and

$$\frac{1}{\Psi}\frac{\partial \Psi}{\partial r_{ij}}\bigg|_{r_{ij}=0} = \frac{1}{4}, \tag{2.18}$$

for parallel spins [51]. Sampling of a trial wave function without cusps gives rise to large statistical fluctuations in the local energy E_L. Therefore, cusp conditions should be built into the wave function in order to prevent such a divergent behavior at the electron-electron and electron-nucleus coalescence points. Doing so improves the overall numerical stability and reduces the total variance making the computations much more feasible.

One of the compact functional forms of the trial wave function frequently used in QMC is the Slater-Jastrow [50] function (we do not discuss other ansätze in this work). It is expressed as a product of the Slater part with determinants D_n^σ for each spin channel σ, and positive definite Jastrow correlation factor $J(\mathbf{R})$ [176],

$$\Psi_T(\mathbf{R}) = e^{J(\mathbf{R})}D = \Big[\sum_n c_n D_n^{\uparrow}(\mathbf{r}_1, \mathbf{r}_2, ..., \mathbf{r}_{N_\uparrow}) D_n^{\downarrow}(\mathbf{r}_{N_\uparrow+1}, ..., \mathbf{r}_N)\Big] e^{J(\mathbf{R})}, \tag{2.19}$$

containing explicit functional dependence on electron-nucleus and electron-electron distances (cf. Section 1.1 for definitions) efficiently accounting for dynamical correlation effects [50]. The Jastrow function is not antisymmetric with respect to the exchange of electrons with opposite spins, but it has been reported to give a negligibly small spin contamination, further decreasing with an improved accuracy of the trial wave functions [177]. The reason is that exact eigenfunctions of the non-relativistic Hamiltonian commute with spin operators.

The Slater part of the wave function may contain single or sum of multiple determinants. One always looks for compact wave functions. I.e., as few determinants as able to capture the right physics.

One of the possible functional forms of the Jastrow factor, containing the electron-nucleus, electron-electron and electron-electron-nucleus explicit inter-particle correlation terms reads [51]

$$J(\{r_{ij}\}, \{r_{iA}\}, \{r_{jA}\}) = \sum_{A=1}^{M}\sum_{i=1}^{N}\chi_A(r_{iA}) + \sum_{i=1}^{N}\sum_{j>i}^{N} u(r_{ij}) \tag{2.20}$$

$$+\sum_{A=1}^{M}\sum_{i=1}^{N}\sum_{j>i}^{N} f_I(r_{iA}, r_{jA}, r_{ij}).$$

The electron-electron term reduces the magnitude of the many-electron wave function whenever two electrons come to a close proximity, i.e. it reduces the probability of finding two electrons close to each other, thus decreasing a mean electron-electron interaction energy which is equivalent to the recovery of the correlation energy beyond a mean-field. The terms containing electron-nucleus dependence allow to vary the electron density [51]. The third term improves the description of correlations and lowers the local energy fluctuations [51]. The three terms are expanded in terms of one-dimensional basis functions as

$$\chi(r_{iA}) = \sum_k c_k^{en} a_k(r) \tag{2.21}$$

$$u(r_{ij}) = k_{\text{cusp}}(r) + \sum_k c_k^{ee} b_k(r) \tag{2.22}$$

$$f(r_{iA}, r_{jA}, r_{ij}) = \sum_{klm} c_{klm}^{een}[a_k(r_{iA})a_l(r_{jA}) + a_k(r_{jA})a_l(r_{iA})]b_m(r_{ij}). \tag{2.23}$$

This functional form in combination with a basis set that satisfies cusp conditions may account for the dominant fraction of electronic correlations with a relatively small number of variational parameters [51,178,179]. The possibility to explicitly built cusps into the wave function is one of the reasons why the QMC methods are much more efficient w.r.t. methods of quantum chemistry based on Gaussian-expanded orbitals, where the wave function are expanded in an enormous number of determinants in order to mimic the cusp conditions which is not exactly possible anyway in this basis [58].

2.6 Variational Monte Carlo

The variational Monte Carlo (VMC) relies on a variational principle (Section 1.4) and the Monte Carlo evaluation of multidimensional integrals as e.g. quantum mechanical expectation values (Equation 1.2). The trial wave function Ψ_T which is a "good" approximation to the true ground-state is typically supplied as an input to VMC [50] and variationally improved. In order to stochastically integrate Equation 1.12, it is recast to

$$E[\Psi_T] = \frac{\int |\Psi_T(\mathbf{R})|^2 [\Psi_T(\mathbf{R})^{-1} \hat{H} \Psi_T(\mathbf{R})] \mathrm{d}\mathbf{R}}{\int |\Psi_T(\mathbf{R})|^2 \mathrm{d}\mathbf{R}}, \tag{2.24}$$

where $\mathbf{R} = (\mathbf{r}_1, \mathbf{r}_2, ..., \mathbf{r}_N)$ is the walker (Fig. 2.2), i.e. the set of spatial positions of all electrons in the system (complete basis set). Metropolis algorithm (Appendix A) [145, 50, 180] is used to sample a set of walkers $\{\mathbf{R}_i : i = 1, ..., K\}$ from the density

$$\frac{|\Psi_T(\mathbf{R})|^2}{\int |\Psi_T(\mathbf{R})|^2 \mathrm{d}\mathbf{R}}. \tag{2.25}$$

For each i, after equilibration period, the local energy

$$E_L = \Psi_T(\mathbf{R})^{-1}\hat{H}\Psi_T(\mathbf{R}) \tag{2.26}$$

is calculated and saved, and at the end of the run averaged over Q (Q = number of walkers $\times$ independent steps) according to

$$E_{VMC} = \frac{1}{Q}\sum_{i=1}^{Q} E_L(\mathbf{R}_i). \tag{2.27}$$

The expectation values of other operators may be evaluated in an analogous way [50].

The optimization of trial wave functions is a crucial and complicated task in QMC [181]. The quality of FN-DMC results, that are of ultimate interest in QMC methodology (Figure 2.1)), and sampling efficiency of calculations are strongly affected by the quality of input wave functions and their flexibility (in VMC).

The Jastrow parameters and coefficients of the determinantal expansion and/or the orbitals of the wave function may be indeed variationally optimized in VMC. Depending on choice, the number of variational parameters may vary in the range of a few dozens to many thousands. The optimization improves description of the correlation effects, lowers the variational energy and lowers the statistical noise (due to zero-variance zero-bias principle) leading to a lower CPU time needed to reach a certain level of statistical uncertainty in production VMC and DMC runs. Optimization of the spatial orbitals in the Slater part / determinant and/or optimization of CI expansion coefficients leads to a change of nodal surfaces leading to an improvement [169], or worsening due to the noise [112]. The result depends on size of the ensemble, optimization procedure and quality of the input.

To tackle the problem of VMC minimization of expectation values of multidimensional parametric functions in the presence of statistical noise, various cost functions have been adopted in literature to minimize the VMC expectation values, e.g. variance [182, 181], energy [183] or their linear combination [183, 184]. The energy-dominated minimization has been reported to be the most efficient for subsequent DMC simulations [185].

The energy/variance landscape for finite sample depends on the distribution of configurations. Several cycles (5-20) of configuration generation and optimization are normally carried out in VMC optimization and the wave function that gives a lowest variational energy and variance is then used for the final FN-DMC production calculations.

Although the VMC optimization can be quite powerful, the necessity to guess a functional form of the trial wave function limits its accuracy and the way to systematically improve it to the exact non-relativistic limit is very costly (e.g. adding more determinants [186] and/or more Jastrow terms) and a simple way is not known [169]. In practice, the VMC is able to recover about 70-90% of correlation effects which is not enough for high accuracy calculations. VMC is thus mainly used to produce trial wave functions required as a guiding functions of improved quality for the more powerful projector DMC technique described above.

2.7 Energy Derivatives

In addition to cost, QMC suffers from a statistical nature of its results in a sense that variance of total energy derivatives diverges in the most straightforward approach.

Here we briefly demonstrate the source of a problem. The approximate derivative of the energy E with respect to nuclear position displacement δq_i,

$$\frac{dE}{dq_i} \approx \frac{E(\delta q_i) - E_0}{\delta q_i} \tag{2.28}$$

of two statistically independent energy estimates $E(\delta q_i)$ and E_0 each with a statistical error σ_e, results in a statistical uncertainty $\sigma \propto \sigma_e/\delta q_i$. For δq_i typically ~ 0.01Å, the uncertainty in derivative σ is 100 times greater than the σ_e. The methods to overcome this difficulty have been proposed [187, 188].

A great improvement may be obtained by the correlated sampling method in which both energies of interest $E(\delta q_i)$ and E_0 are evaluated using the same ensemble of walkers, resulting in correlated statistical fluctuations (further improvement is possible by a space-warp transformation [189]). The variance of the energy difference increases with the increasing displacement as the correlation decreases. For small displacements and large samples, strong correlation is expected and in ideal case the difference would have zero statistical uncertainty [145]. The method has been successfully applied to small molecules [145] or to couple DFT molecular dynamics with FN-DMC [190].

The Hellman-Feynman theorem (HFT) [191] states that

$$\nabla_q E = \frac{\int \Psi \nabla_q \hat{H} \Psi \mathbf{dx}}{\int |\Psi|^2 \mathbf{dx}}, \tag{2.29}$$

providing a simple means to obtain energy gradients [145]. It is satisfied by the exact eigenfunctions of $\hat{H}$ and HF wave functions in the CBS limit. Any other approximations are not guaranteed to satisfy HFT and are typically not satisfactory [192]. However, HFT may be used to estimate quality of trial wave functions, since the sum of the Hellman-Feynman forces should vanish for exact solutions [145].

In QMC, the problem with HFT arises in the potential gradient term of the Hellman-Feynman estimator,

$$\nabla_q \hat{V} = \sum_{AB} \frac{Z_A Z_B}{r_{AB}^3}(\mathbf{r}_B - \mathbf{r}_A) - \sum_{iA} \frac{Z_A}{r_{iA}^3}(\mathbf{r}_i - \mathbf{r}_A), \tag{2.30}$$

since the variance of the latter term is unbounded [145]. No reasonable estimate can be obtained from sampling HFT in this form. Other promising schemes were thus proposed [188].

2.8 Effective-Core Potentials

Effective-core potentials (ECP's) [193] have been successfully used in solid-state physics and quantum chemistry for decades and are now routinely used in QMC [50, 52, 155, 194]. The idea is to replace a chemically inactive localized core electrons by an effective potential acting on valence electrons [145], which offers considerable speedup of calculations for heavy elements due to a reduced number of explicitly treated electrons. Another advantage is that by screening out the inner electrons, we deal with a numerically smaller numbers so that the calculated energy differences are less sensitive to errors coming from multi-scale electron dynamics in the full-electron problem.

The computational cost of QMC with respect to the atomic number has been estimated to scale as $Z^{5.5-6.5}$ [145, 195], i.e. far more pronounced than the scaling with the number of explicitly treated electrons ($O(N^{3-4})$). Such an unfavorable scaling is caused by large fluctuations of local energy due to strong potentials and large kinetic energies near the core. In addition, the different time/space scale of the sampled core and its high density implies that valence shells that are of interest are not sampled efficiently [51]. The most efficient way to improve this behavior is to eliminate the core electrons from the explicit calculation and use ECP's instead [50]. The CPU time of ECP calculations scales more effectively, with $Z_{eff}^{3.4}$, and Z_{eff} remains small for all elements [195]. In cases where the relativistic spin-orbit effects are negligible, the pseudo-potential approximation introduces errors typically of the same order or lower than error introduced by the FN approximation [196]. Accurate QMC results with ECP's were obtained [146, 197, 50, 52].

The typical non-local ECP angular momentum dependent operator typically used to simulate repulsion between core and valence electrons reads,

$$\hat{V}_{ECP} = \hat{V}_{loc} + \hat{V}_{nl} \tag{2.31}$$

where $\hat{V}_{loc}$ is a local part shared by all angular momentum channels, and

$$\hat{V}_{nl} = \sum_{l}^{l_{max}} \hat{V}_l(r) \sum_{m} |lm\rangle\langle lm| \tag{2.32}$$

is a non-local angular momentum lm dependent part.

The non-local contribution to VMC local energy corresponds to

$$\frac{\hat{V}_{nl}\Psi_T}{\Psi_T} = \sum_{l}^{l_{max}} \hat{V}_l(r_{iA}) \frac{2l+1}{4\pi} \int_{4\pi} P_l(\cos(\theta'_{iA})) \quad \frac{\Psi_T(\cdots, \mathbf{r}'_\mathbf{i}, \cdots)}{\Psi_T(\cdots, \mathbf{r}_\mathbf{i}, \cdots)} d\Omega'_{iA}, \tag{2.33}$$

where P_l is the Legendre polynomial of degree l. The angular integral over the spherical surface centered on atom A passing through the i-th electron [51, 146], must be evaluated numerically and its accuracy depends only on quadrature techniques used [50].

In DMC, the non-local part of the ECP gives rise to the term $\langle \mathbf{R}|\exp(-\delta\tau\hat{V}_{nl})|\mathbf{R}'\rangle$ in the Green's function, not guaranteed to be non-negative for all $\mathbf{R}$, $\mathbf{R}'$ and $\delta\tau$. Since the ensemble of walkers evolves according to equation

$$-\frac{\partial f}{\partial t} = -\frac{1}{2}\nabla^2 f + \nabla.(\mathbf{v}_D f) + \frac{(\hat{H} - E_T)\Psi_T}{\Psi_T} f + \left\{ \frac{\hat{V}_{ECP}\Psi_T}{\Psi_T} - \frac{\hat{V}_{ECP}\Psi}{\Psi} \right\} f, \tag{2.34}$$

and the sign of the walkers may fluctuate, the analogy of fermion-sign problem emerges [50]. One of the possible approaches to overcome this difficulty is use of localization approximation which corresponds to neglect of the term in curly brackets. This however requires highly accurate trial wave functions (in the quadratic-convergence sense) to keep the error proportional to $(\Psi_T - \Psi)^2$ small [146]. Such wave functions are available in practice, e.g. VMC optimized Slater-Jastrow trial wave functions were found to behave well in this respect [198, 199, 194]. A more general way to include non-local potentials in the standard DMC method without using the locality approximation has been formulated recently [200]. The algorithm includes the so-called

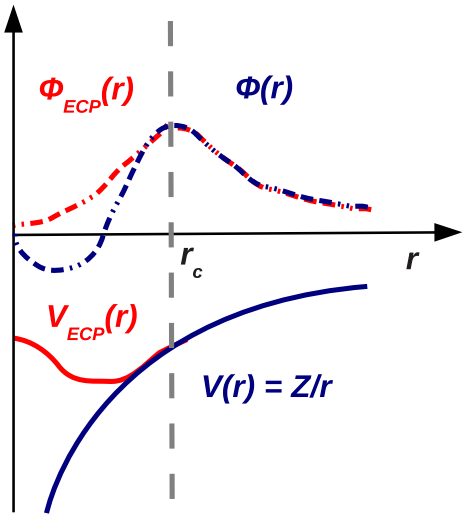

Fig. 2.7. An idea of effective-core potential (ECP) V_{ECP} (solid red line) is to replace localized and energetically separated core electrons moving in the atomic potential V (blue) by a potential that mimics their effect ($V_{ECP} = V$ and $\phi_{ECP} = \phi$) in the valence region (i.e. for $r > r_c$). In addition, for QMC it is desirable to have non-divergent ECP (at origin) as shown.

T-moves generated according to the non-local potentials. By using an effective Hamiltonian approach, stable simulations obeying variational principle are obtained.

The ECP's in QMC lead to a dramatic improvement of efficiency, but since errors are introduced, one must make tradeoffs. For example, in the case of a Fe atom, the ECP with Ne core has been reported to provide a reasonable accuracy and the acceptable statistical sampling efficiency [146].

A typical ab initio ECPs used to generate input trial wave functions diverge at the nucleus (V in Figure 2.7) [198]. The singularities lead to a large time-step bias if the pseudo-valence orbitals are not properly described close to the nucleus, which requires extremely large Gaussian basis sets [201]. It is therefore desirable to have a non-divergent ECP's for QMC calculations.

The non-singular ECP's suitable for QMC have been given by Greeff [201], Ovcharenko [202], Trail and Needs [203, 204] and Burkatzki et al. [198, 199]. A recent comparison of Trail/Needs and Burkatzki et al. ~~The~~ ECPs shows their similar performance in the main-group molecules [194].

3 Noncovalent Interactions by FN-DMC

This part is devoted to the description of specific use of FN-DMC in the domain of molecular noncovalent interactions, in light of the recent advances. The examples given are limited by the work of the author.

There are at least three important advantages of FN-DMC based QMC approaches in the field of noncovalent interactions. First, the method accurately describes the dynamic correlation effects crucial for noncovalent interactions, since it recovers all many-body correlations within the constraints given by the nodal surface of the best available trial wave function [50]. Therefore, the basis-set dependence only enters the trial wave function construction stage. In FN-DMC one generates CBS (Section 1.3) equivalent data without need of extrapolation with respect to the size of one-particle basis set. Second, the error due to the FN approximation is expected to cancel out in energy differences nearly exactly (see below for a detailed explanation). Third, a low-order polynomial scaling of order $\propto O(N^{3-4})$ with the number of electrons N is more favorable than a typical scaling of benchmark quantum chemical benchmark methods like ordinary coupled-cluster CCSD(T) or QCISD. Here we note that the situation in CC community changes for a new improved scaling methodology [48, 49]. Nevertheless, the former two advantages of QMC hold, and, in addition, a high parallelism allowing use of large supercomputers and a possible use of QMC in periodic an multireference systems (e.g. studies of noncovalent interactions between closed-shell molecule and transition metal [205]) makes it a very interesting method for benchmark studies. Further development of QMC methodology is therefore very valuable and desired.

This section continues with a brief literature review, and two important single-determinant fixed-node error cancellation-based protocols are discussed and compared. The related technical aspects are discussed more extensively in the final subsections.

3.1 Literature Overview

An important example of QMC in calculation of weak intermolecular forces in He_2 and He-LiH complexes was presented by Mella and Anderson (2003) [160]. The authors extensively discussed advantages of FN-DMC (position representation that is a complete basis set without basis-set-superposition error, size-consistency, variational property) for noncovalent interactions (FN error cancellation) and found an excellent agreement of FN-DMC vs. the state-of-the art quantum chemical calculations (CCSD(T)).

A subsequent important work by Diedrich et al. (2005) [206] reported on the use of FN-DMC in five noncovalent complexes, including methane dimer, ammonia dimer, water dimer, and two important benzene dimer conformations (T-shaped and parallel displaced). Apart from the main message (below), the work mentions an important fact that the overlap of molecular complex fragments densities is small in weakly interacting systems and therefore the errors in inter-fragment electron exchange have a small contribution to the overall accuracy. A consequence is that a product of monomer wave functions is a good approximation to the wave function of the whole complex and such an approach may benefit from the fixed-node error cancellation (Section 3.2). A primary contribution of this work lies in an extensive study of sampling errors due to the use of limited Gaussian basis sets in a trial wave function. The authors claim, that since the Gaussians fall off more rapidly than the correct Slater-type tails, local energy and

drift asymptotically diverge. Tests using TZV, aug-TZV and aug-QZV and some truncated bases show that larger bases slightly improve the performance for that they shift the region of divergence away from the sampled regions of the molecule. In methane, aug-QZV bases were found necessary to obtain reasonable results.

An extensive testing of FN-DMC on a test set containing 22 molecular complexes (S22 [31]) and a more extended study on adenine/thymine and cytosine/guanine DNA base pairs interactions was reported in 2008 by Körth et al. [161]. The FN-DMC results on the S22 set obtained with the QZV basis sets and single-determinant Slater-Jastrow guiding wave functions shown significant discrepancies with respect to the CCSD(T) reference data. Namely, the mean absolute deviation was found to be 0.68 kcal/mol and seven out of all 22 complexes revealed errors bigger than 1 kcal/mol. In some cases the relative error was exceedingly high (benzene-methane 58 %, benzene dimer parallel-displaced 40 %, benzene-HCN 24 %, benzene dimer T-shaped 37%). In light of the recent update of CCSD(T)/CBS results for the S22 set [32] that confirmed previous CCSD(T)/CBS estimates with extended basis sets, it turns out that the failure is on the side of QMC.

This fact motivated our work in this field. We asked: could we obtain better results? What is the approach to do so? Is it possible to optimize the approach so that it becomes general and easy-to-use ("black-box" type approach) for non-expert users interested to use QMC in benchmark noncovalent energy calculations? None of these questions is yet fully answered.

We were able to show that single-determinant FN-DMC based easy-to-use protocols similar to those mentioned above are able to provide errors significantly smaller than shown previously [161]. Namely, we demonstrated the ability of the FN-DMC to reach the CCSD(T)/CBS within 0.1 kcal/mol on a teaching set of six small noncovalent complexes: the dimers of ammonia, water, hydrogen fluoride, methane, ethene and ethene/ethyne complex [53]. This finding was very encouraging. A primary difference in our approach, when compared to the previous work, is a use of augmentation functions instead of higher cardinality of the basis set used, and exhaustive VMC optimization of trial wave functions. The successive procedure was subsequently tested on complexes of benzene/methane, benzene/water and T-shape benzene dimer, where the FN-DMC deviated by no more than 0.25 kcal/mol with respect to the best available CCSD(T)/CBS interaction energy estimates. More explicitly, in benzene- methane complex (S22, i.e. the same structure as used in Ref. 161), we were able to reduce the error from -0.87 kcal/mol to -0.15 kcal/mol while in the T-shaped benzene dimer from S22, the error dropped from 1 kcal/mol to 0.2 kcal/mol [53] while keeping the approach very simple. Furthermore, the FN-DMC data were away only within one σ suggesting that the remaining error may be of statistical nature (at least in part). The related relative errors (in the benzene dimer case) dropped from 58 to 10 % and from 40 to 7 %, respectively. Below, we refer to the corresponding computational scheme (protocol) as 3BJ, which indicates that the used Jastrow term contains up to three-body electron-electron-nucleus terms.

Subsequently we were interested whether there is a space for the reduction of a CPU cost. It turned out that one order of magnitude gain in CPU cost may be obtained by the reduction of the number of terms in the Jastrow factor. Namely, the Jastrow factor containing up to two-body (electron-electron and electron-nucleus) terms shown acceptable performance while significantly reducing the overall CPU cost [54] (this protocol is referred to as 2BJ; cf. below). The 2BJ protocol allowed us to extend the applicability to larger systems like coronene-H_2 complex (Figure 3.1) [13].

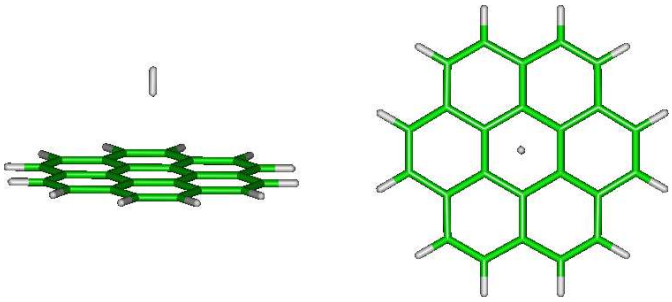

Fig. 3.1. The coronene-H_2 complex studied by QMC in Ref. 13.

The currently ongoing research suggests a further possible reduction of the size of Gaussian basis sets used to construct trial wave functions for noncovalent complexes [56]. The overall work is nevertheless still far from complete - we believe that the currently available approaches may well be used in single-point estimates of interaction in trivially saturated closed-shell complexes. In small molecules with double and triple bonds (e.g. HCN dimer [54] or formic acid [205]), residual errors remaining in QMC turned out to be non-negligible in both, 2BJ and 3BJ approaches [54]. Improvement in these cases remains as our next goal. Our general goal is to provide a transferable computational protocol based on FN-DMC that is useful for a wide class of non-covalently interacting systems. We are indeed interested in physics behind, i.e. a detailed understanding of reasons for the observed behavior of FN-DMC successes and/or failures. Our parallel project is focused on ascertaining of the limits of FN-DMC-based error-cancellation protocols including closed-closed shell and open-closed shell complexes [205].

We note that except for other more complex QMC approaches [207, 208], FN-DMC-based approaches on single-determinant trial functions (containing e.g. Kohn-Sham orbitals) have been used in many studies on small/medium [209, 161, 210, 211, 53, 115, 212, 13] and medium/large molecular systems [99, 28, 213, 214, 215, 216, 217, 55, 163, 164] or periodic systems [218, 212] where conventional correlated methods of quantum chemistry were inapplicable due to the prohibitive computational cost and/or other limitations.

3.2 Error-Cancellation Protocols

In this part, we describe, understand and compare the accurate (3BJ) and reduced (2BJ) protocols, relying on the FN-error cancellation, that were identified and tested on limited sets of noncovalent complexes by the author and colleagues in 2013 and 2014 [53, 54].

First, by the FN-error cancellation we mean that for every "reasonable" approximate method providing trial wave functions for the complex $A + B$ and its constituents A and B, we expect cancellation of FN-errors (in total energy differences) in the limit of weak interaction. I.e., we expect that in the super system, the nodes in the region of the molecule A are almost intact in the presence of the weakly interacting molecule B and vice-versa. Even though the wave functions are approximate, the FN-error of the approximation is the same in both cases, monomer and dimer alike, so the errors cancel out. This idea is illustrated and explained in the Figure 3.2. Direct inspection of nodal surfaces indeed confirms this trend (Figure 3.3) [53]. Limits of this

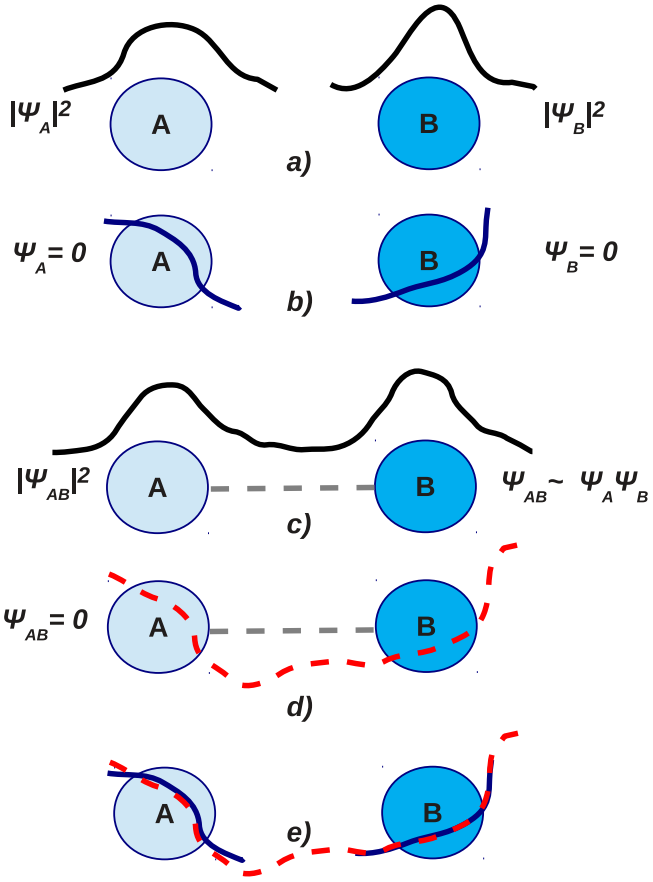

Fig. 3.2. Explanation of the FN-error cancellation concept: if the isolated monomer molecules A and B (a) with their corresponding wave functions (a) and the related nodal surfaces (b) interact weakly (c) and create a new closed quantum system $A + B$ (c), the nodal surfaces of the wave function Ψ_{AB} (d) are very similar near A and B (b vs. d) in the limit of weak interaction (e). Since there is a vanishing overlap of Ψ_A and Ψ_B, the exchange repulsion between such pairs of same-spin electrons where each resides on one of the monomers is negligible. The FN error cancellation is therefore expected in total energy differences for "reasonable" approximate wave functions Ψ_A, Ψ_B and Ψ_{AB}, which is indeed observed (Figure 3.3).

approach are currently unknown [205] except for the following.

The FN-error cancellation behavior assumes that an approximate method describes nodes of monomers and dimer on equal footing. This assumption may well work in cases where monomers are closed-shell complexes, or at least one of them is a closed-shell complex. Otherwise, we deal with a situation similar to bond-breaking (multireference character arises) and such situations are never easy to describe by cheap/approximate methods since equal footing description is hardly possible if the super system and its constituents possess various multiplicities [150].

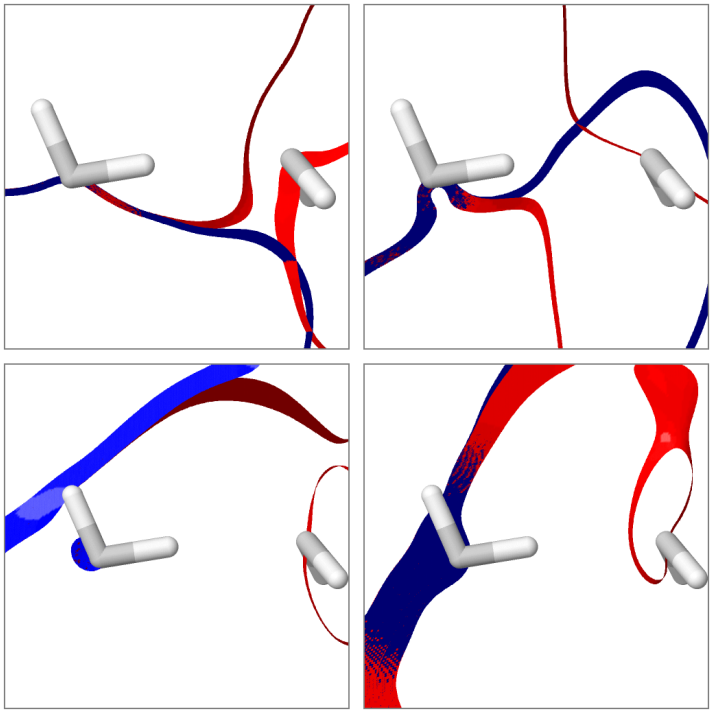

Fig. 3.3. The four representative comparisons of the cuts through the multidimensional nodal hyper-surfaces ($\Psi_T = 0$) of single-determinant B3LYP/aug-TZV wave functions for the water dimer complex (nodes in red) and one of its isolated frozen water molecules (left; nodes in blue). The produced 2D ribbon-like single-electron scans are essentially indistinguishable in the region of the monomer where the electron density achieves its maximum. These images provide thus a visual rationalization of the FN-error cancellation observed in weakly interacting molecular complexes (cf. Figure 3.2). The cuts were produced by single-electron scans of a 3D volume, keeping the remaining $N - 1$ electron positions fixed in both, dimer and monomer alike. The monomer scans used a subset of the dimer electron positions to produce equivalent surfaces. Since the nodes exhibit rather complicated folds, representative nodal cuts were performed in a cuboid squeezed in one dimension so that the corresponding plots appear as ribbon-like [53].

By the FN-error cancellation protocol, we mean a FN-DMC based multi-stage QMC calculation that does not involve nodal surface optimization at the VMC level (Figure 2.1). Rather, at the VMC level, only a Jastrow factor in the trial wave function is optimized to minimize the total energy variance but the nodal surfaces are kept intact as they come from quantum chemistry. This allows to use the protocol as is and thus makes it easy-to-use (or more "black-box"-like) rather than system-specific. At the current stage, it seems that at least for trivially saturated main-group elements, the FN-error cancellation is satisfactory however complete and systematic understanding of error cancellation in QMC energy differences is still far from complete. It is also possible that we will not be able to provide a general "black-box" protocol. As shown below, the current approaches require more extended testing and non-negligible errors are still present in

Tab. 3.1. The FN-DMC interaction energies E (kcal/mol) obtained by the 2BJ and 3BJ protocols compared to the CCSD(T)/CBS reference interaction energies E_R (kcal/mol) along with the corresponding differences Δ (kcal/mol), relative differences Δ_R (%) and statistics: mean error (ME, kcal/mol), mean unsigned error (MUE, kcal/mol) and relative unsigned error (RUE,%).

Complex	E_R	E/2BJ[c]	Δ	Δ_R	E/3BJ[d]	Δ	Δ_R
Ammonia dimer	-3.15[a]	-3.21±0.09	0.06	1.9	-3.10±0.06	-0.05	1.6
Water dimer	-5.07[a]	-5.28±0.10	0.21	4.1	-5.26±0.08	0.19	3.7
HF dimer	-4.62[b]	-4.77±0.12	0.15	3.2	-4.68±0.10	0.06	1.3
Methane dimer	-0.53[a]	-0.60±0.07	0.07	13.2	-0.44±0.05	-0.09	17.0
Ethene dimer	-1.48[a]	-1.53±0.13	0.05	3.4	-1.47±0.09	-0.01	0.7
Ethene /ethyne	-1.50[a]	-1.54±0.14	0.04	2.7	-1.56±0.08	0.06	4.0
HCN dimer	-4.75[b]	-5.09±0.08	0.34	7.2	-5.13±0.06	0.38	8.0
Methane HF	-1.65[b]	-1.29±0.07	-0.36	21.8	-1.54±0.04	-0.11	6.7
FD[f] dimer	-4.55[b]	-4.90±0.10	0.35	7.7	-4.88±0.06	0.33	7.3
Ethyne dimer	-1.52[b]	-1.74±0.07	0.22	14.5	-1.62±0.05	0.10	6.6
Ethene Ar	-0.36[b]	-0.19±0.07	-0.17	47.2	-0.22±0.04	-0.14	38.9
Benzene/H_2[c]	-1.03[c]	-1.08±0.10	0.05	4.9	-	-	-
Coronene/H_2[e]	-1.12[e]	-1.03±0.16	-0.09	8.0	-	-	-
Benzene/water	-3.29[a]	-3.21±0.15	-0.08	2.4	-3.53±0.13	0.24	7.3
Benzene/methane	-1.45[a]	-1.50±0.15	0.05	3.4	-1.30±0.13	-0.15	10.3
Benzene dimer T	-2.71[a]	-2.53±0.23	-0.18	6.6	-2.88±0.16	0.17	6.3
Adenine-Thymine S	-11.66[a]	-11.00±0.30	-0.66	5.7	-	-	
			ME:	0.00		ME:	0.07
			MUE:	0.18		MUE:	0.15
			RUE:	9.29		MUE:	8.54

[a] Ref. 32, [b] Ref. 26, [c] Ref. 53, [d] Ref. 54, [e] Ref. 13, [f] formaldehyde.

systems with multiple bonds (e.g. HCN or formaldehyde), high-symmetry (HCN, Ar containing complexes) or if the amplitude of interaction is small (CH_4) below 1 kcal/mol (Table 3.1).

We note that FN-error cancellation approaches rely on suboptimal nodes, in contrast to FN-DMC with fully VMC optimized nodal surfaces (by optimization of CI coefficients and/or spatial orbitals). The full optimizations are however performed on fixed samples in the presence of noise and such approaches must be much more carefully converged on large ensembles of walkers (that are costly) to reach the target 0.1 kcal/mol accuracy (in some cases this may be undoable) in energy differences.

The 2BJ and 3BJ protocols (or recipes) that we refer to in this part rely on the following methodology. The geometries of complexes used for testing of 2BJ and 3BJ protocols were taken from the databases of noncovalent complexes S22 [31] and A24 [26], except for the benzene/H_2 [13] complex. The ECPs with the corresponding basis sets developed by Burkatzki et al. (2007) [198] were used throughout the studies cited, with the exception of the H where a more recent version was used [219]. The full sets of augmentation functions were taken as they come from the corresponding Dunning bases [73]. Single-determinant Slater-Jastrow [176] trial

Tab. 3.2. The FN-DMC interaction energies E (kcal/mol) in the A24 set from the Ref. 54, obtained by the 2BJ protocol, compared to the CCSD(T)/CBS reference interaction energies E_R (kcal/mol) along with the corresponding differences Δ and statistics: mean error (ME, kcal/mol), mean unsigned error (MUE, kcal/mol) and relative unsigned error (RUE,%).

Label	Complex	E_R	E/2BJ	Δ
1	Water ammonia	-6.493	-6.71±0.07	-0.22
2	Water dimer	-5.006	-5.30±0.05	-0.29
3	HCN dimer	-4.745	-5.09±0.08	-0.35
4	HF dimer	-4.581	-4.88±0.05	-0.30
5	Ammonia dimer	-3.137	-3.30±0.04	-0.17
6	Methane HF	-1.654	-1.29±0.07	0.37
7	Ammonia methane	-0.765	-0.83±0.06	-0.06
8	Methane water	-0.663	-0.57±0.06	0.09
9	Formaldehyde dimer	-4.554	-4.90±0.10	-0.35
10	Water ethene	-2.557	-2.55±0.07	0.01
11	Formaldehyde ethene	-1.621	-1.70±0.08	-0.07
12	Ethyne dimer	-1.524	-1.74±0.07	-0.21
13	Ethene ammonia	-1.374	-1.45±0.07	-0.07
14	Ethene dimer	-1.09	-1.08±0.07	0.01
15	Methane ethene	-0.502	-0.54±0.06	-0.04
16	Borane methane	-1.485	-1.48±0.04	0.01
17	Methane ethane	-0.827	-0.85±0.06	-0.03
18	Methane ethane	-0.607	-0.68±0.06	-0.07
19	Methane dimer	-0.533	-0.63±0.03	-0.10
20	Methane Ar	-0.405	-0.18±0.07	0.23
21	Ethene Ar	-0.364	-0.19±0.07	0.18
22	Ethene ethyne	0.821	0.94±0.07	0.12
23	Ethene dimer	0.934	1.06±0.03	0.13
24	Ethyne dimer	1.115	1.23±0.03	0.11
			ME:	-0.05
			MUE:	0.15
			RUE:	12.1

wave functions were constructed using B3LYP or HF orbitals from `GAMESS` [220]. The used isotropic Jastrow factors [221, 176] included either the electron-electron and electron-nucleus terms (2BJ), or additional electron-electron-nucleus terms (3BJ). The Jastrows were expanded in a fixed basis set of polynomial Padé functions [51]. The parameters of the positive definite Jastrow terms were optimized by the Hessian driven Levenberg-Marquardt VMC optimization of at least 10x10 iterations (i.e. a full VMC energy calculation after each 10 optimization steps on a fixed walker population), using the variance optimization [222] or a linear combination [183] of energy (95%) and variance (5%) as a cost function. The optimized trial wave functions were subsequently used in the production FN-DMC runs performed with a time step of 0.01/0.005 a.u. The ECPs were treated within the locality approximation [146, 50] (LA, for testing purposes) or

using the T-moves scheme [200] for the treatment of ECPs beyond LA. The target walker populations ranged from 5k (for small systems) up to about 20k (for the largest system). All QMC calculations were performed using the code `QWalk` [179].

The data reported to date obtained with the 2BJ and 3BJ protocols are summarized in Table 3.1 and 3.2. The Table 3.1 compares both protocols in various systems, while the latter, Table 3.2, shows a performance of 2BJ in the whole benchmark database A24. The observations and comments on the Tables follow:

i) The performances of both protocols appear as statistically similar (Table 3.1). Note that the 2BJ is one order of magnitude faster [54] and is thus more favorable.

ii) The mean errors (ME) close to zero for both protocols suggest a statistical nature of the (part of) residual errors which may be further reduced by extensive sampling.

iii) The mean unsigned errors (MUE) close to 0.2 kcal/mol are acceptable but not ideal for sizes of systems considered. Further improvements are now underway (cf. iv).

iv) The largest errors in 3BJ protocol suggest the current limits of the method in systems with multiple-bonds: HCN dimer and formaldehyde dimer. The large errors in the remaining complexes, e.g. methane dimer and ethene-Ar complex were reduced [205] by the increased number of ECP integration points (atom-closed-shell complexes in general require this) and/or spin-restricted Jastrow factor. The errors shown for these cases thus do not relate to the failure of FN-error cancellation (as expected).

3.3 Technical Aspects

Development of methodology based on FN-DMC, that finally leads to accurate noncovalent interaction energies in extended sets of molecules and scenarios involves extensive testing and elimination of biases that affect the results. This has to be done in a stepwise manner since many parameters enter the multistage approach [223, 68]. These include: selection of Ψ_T (method, ECP and related basis set, number of determinants) and its VMC optimization (VMC cost function, number of optimization cycles, etc.) and parameters that enter the final production FN-DMC calculations (time step, walker ensemble, treatment of ECP). Here we describe a general approach to search for the protocols reported. Some useful details from the extensive tests related to the key points (as mentioned above), that helped to find/improve 2BJ and 3BJ (cf. above) approaches are included.

i) The trial wave function ansatz must be selected first (we assume that the molecular structures are available). The standard and frequently used choice that offers a good performance/cost ratio is the Slater-Jastrow [50, 52] wave function, a product of determinant(s) and an explicit correlation term (Jastrow factor; Section 2.5). This is our default choice. It is then necessary to choose an effective core potential (ECP, if any) which is "a must" (for the scaling issue vs. CPU cost - cf. Section 2.8; in noncovalent interactions, extremely small error bars make the calculations very costly) and a basis set (e.g., aug-TZV). Subsequently, the Slater determinant(s) are constructed with DFT, Hartree-Fock (HF) or post-HF orbitals. At any level, SCF procedure must be used (in truncated MC-SCF case, SCF post-processing step with truncated CI expansion is recommended), otherwise the FN-error cancellation is not guaranteed. Finally, terms included in the Jastrow factor [176] must be specified. They include, for instance, electron-electron (ee), electron-nucleus (eN), and electron-electron-nucleus (eeN) terms, which contain explicit functional dependence on inter-particle distances and thus efficiently describe dynamic correlation

effects [50].

ii) The variational (VMC) optimization step consists of the selection of the VMC cost function and a parametric optimization of Ψ_T, which may or may not include adjustment of the nodal surfaces. In general, it is possible to improve the nodes by the reoptimization of the orbitals and/or determinant expansion coefficients but in schemes relying on the FN-error cancellation, this must be avoided (Figure 3.2).

iii) The final FN-DMC ground-state projection calculations depend, in addition to the Ψ_T optimized in step ii), on the parameters of the DMC simulation. The DMC parameters include an imaginary time step, the treatment of ECPs, the target walker population/s and target error bar/s.

In general, the parameters and/or choices in all points mentioned above affect the final interaction energies obtained after the production DMC simulations as total energy differences. Since the FN-DMC generates statistically independent total-energy expectation values with associated error bars, the parameters in i) and ii) are important for that, in addition to accuracy, they modify the total energy variance which in turn determines the statistical sampling efficiency and therefore the total calculation run-time.

3.3.1 Trial Wave Functions

As already mentioned, our default choice is a Slater-Jastrow form of trial wave function ansatz. The Slater part contains only single determinant (for economic and "black-box" reasons). Since we would like to rely on the FN-error cancellation (for reasons sketched in Section 3.2), we always keep the orbital coefficients from either DFT or HF fixed. Various Jastrow explicit correlation terms were tested (see below).

3.3.2 Basis sets

The sensitivity of FN-DMC to the one-particle basis set size used to construct single Slater determinant (with B3LYP orbitals) trial wave functions is well apparent from the following example. In the ammonia dimer complex, the reference CCSD(T)/CBS interaction energy amounts to -3.15 kcal/mol [32]. The FN-DMC runs using trial wave functions expanded in non-augmented TZV and QZV basis sets resulted in interaction energies of -3.33$\pm$0.07 and -3.47$\pm$0.07 kcal/mol, whereas the aug-TZV and aug-QZV bases led to FN-DMC interaction energies of -3.10$\pm$0.06 and -3.13$\pm$0.07 kcal/mol, respectively [53, 54]. The presence of augmentation functions in Ψ_T therefore seem to be crucial, whereas an increase of the basis set cardinality beyond the TZV level seems to play a smaller role than in the standard methods of quantum chemistry. In the methane dimer and ethene dimer complexes, the aug-TZV data were found to be statistically indistinguishable from the aug-QZV data [54]. Since the aug-TZV basis set reached the reference data within 0.1 kcal/mol in the whole teaching set considered [53], we currently take is as a reference default. We note in passing that recent research suggests a further possibility of a basis set reduction [56].

3.3.3 Orbitals

Since the FN-DMC energies depend primarily on the nodal surface of the available trial wave function Ψ_T, we tested HF and B3LYP orbitals in Slater determinants. In the ammonia dimer complex [53], we found that the HF and B3LYP orbitals provide FN-DMC interaction energies that are indistinguishable within the error bars, namely -3.12±0.07 vs. -3.10±0.06 kcal/mol, and both in good agreement with the CCSD(T)/CBS reference (-3.15 kcal/mol [32]). Similar conclusions apply in the case of methane dimer, separately in both types of the schemes considered, 2BJ and 3BJ alike. On the other hand, Benedek et al. [209] reported on a difference between HF and B3LYP orbitals in case of water dimer and HF produced results in better agreement with CCSD(T) compared to B3LYP, though the error bars of 0.18 kcal/mol prevent any conclusive statements. In our calculations. the total energies from B3LYP orbitals were always found to produce variationally lower total energy expectation values [154, 224] than those from HF (e.g. in ammonia dimer by ∼0.001 a.u.). This indicates a better quality of the B3LYP nodal surfaces. We therefore use B3LYP orbitals in all calculations as a default choice. We again stress that the results are expected to depend on the choice of the orbitals only very weakly due to the favorable FN-error cancellation (Figure 3.2) that takes place in weakly interacting complexes [160, 206, 161, 53].

3.3.4 Jastrow Factor

The variations of the Jastrow term considered include a reduction of terms (2BJ) with respect to the 3BJ approach. We also tested a so-called distinct Jastrow factor including distinct parameter sets (3BJ Dis.) on the non-equivalent atoms of the same type [54]. The water dimer is an example where the 3BJ Jastrow factor was not sufficient since the corresponding interaction energy (-5.26±0.08 kcal/mol) deviated considerably from the reference (-5.07 kcal/mol [32]). In order to reach a subchemical accuracy margin (0.1 kcal/mol), a distinct 3BJ Jastrow correlation factor must be considered (3BJ Dis.,-5.15±0.08 kcal/mol). On the other hand, the use of 2BJ scheme in combination with the distinct feature (2BJ Dis.) is not recommended, since in the case we tested, it produced the value that is too off (-5.50±0.08 kcal/mol). The reason probably being an insufficient number of sampling points used in VMC optimization (cf. Ref.162). A reasonable compromise between the accuracy and cost is thus provided by the 2BJ scheme (-5.24±0.09 kcal/mol) with electron-electron and electron-nucleus terms in the Jastrow term, which is acceptable within the less tight (than our original criterion of 0.1 kcal/mol) but acceptable target error criterion, e.g. 0.2-0.3 kcal/mol.

3.3.5 Variational Improvement of Trial Wave Functions

We have compared the variance and linear combination of energy and variance (our default choice with 95% of energy and 5% of variance) as a cost function in VMC optimization [54] of the Jastrow term parameters. The test of variance cost function in ammonia dimer lowered the local energy variance, as expected, but the total FN-DMC energy remained higher (e.g. by 0.0012 a.u. in dimer), and the interaction energy produced in this way (-3.28±0.1 kcal/mol) deviated more from the reference (-3.15 kcal/mol) than the "linear combination" cost function (-3.10±0.06 kcal/mol). We thus recommend use of a large fraction of energy in the VMC-optimization cost function, in accord with the Ref. 183.

3.3.6 Treatment of Effective-Core Potentials

In order to reduce the numerical cost of our calculations, we have removed core electrons (from atoms) and simulated their effect by appropriate ECPs. In ammonia dimer, we have found that the T-moves [200] scheme used to treat ECPs in DMC produces more accurate results [54] than the locality approximation [146] (where the error reaches $\sim$0.2 kcal/mol), as expected, and we have thus combined T-moves with a short time step (0.005 a.u.).

3.3.7 Time-Step Bias in FN-DMC

Our standard conservative time-step setting of 0.005 a.u. is used in order to avoid extrapolation of energy to a zero time step [161]. We have tested the DMC time step of 0.01 a.u. in addition to our default choice [54]. In ammonia dimer, deterioration of the final interaction energies in the case of 3BJ with an increase of the time step was observed, whereas in the case of 2BJ, no apparent dependence arose. In the case of water dimer, the calculation of energies using an increased time step showed that only the 3BJ Dis. scheme is able to approach the reference in the zero time step limit. In ethene dimer, there was no significant dependence on the time step in either of the cases considered (2BJ and 3BJ). The time step of 0.005 a.u. is accurate enough for our purposes and we recommend it as a default in general cases and for s/p systems.

4 Tutorial: Interaction Energy of Water Dimer

This section is intended to provide an accessible starting point for newcomers interested in practical calculations. It provides a detailed hands-on tutorial on calculation of a single point interaction energy of a water dimer - an archetypal example of a complex stabilized by a hydrogen bond.

We intend to compute an interaction energy $\Delta E = E_D - E_{M1} - E_{M2}$ for water dimer (D) consisting of dimer-frozen monomers $M1$ and $M2$, i.e. monomers remain in atomic configuration identical to the dimer. The calculation of FN-DMC interaction energy requires multiple steps:

i) generation / setup of molecular structures;

ii) preparation of a Slater determinants for Slater-Jastrow trial wave functions (for all three configurations; D, $M1$, $M2$) which involves three single point DFT calculations that will generate one-particle orbitals;

iii) variational optimization of Jastrow term in the Slater-Jastrow trial wave function, and,

iv) FN-DMC production calculations.

Below, all these steps are described in detail.

The tutorial assumes that the reader has access to the working versions of the codes `GAMESS` and `QWalk`. Both are free of charge and may be easily compiled under GNU Linux on a notebook or cluster and tested. The respective websites (`http://www.msg.ameslab.gov/gamess/` and `http://qwalk.org`) provide users with a documentation and the packages for download.

4.1 Structure of the Water Dimer

The benchmark methods like CCSD(T) or QMC are excessively expensive. One therefore tries to avoid structure minimizations whenever possible. A typical strategy in noncovalent interaction-energy reference calculations is to use a fixed set of structures preoptimized in best available methods like MP2 [109] or dispersion-corrected DFT [71]. Any method that provides structures close to required minima is satisfactory since details near true minima are not critical in fixed-nuclei benchmarks.

In the current tutorial, we use a benchmark minimum structure of the water dimer (Figure 4.1) while the corresponding monomer constituents (Figure 4.1) are kept frozen in order to get an

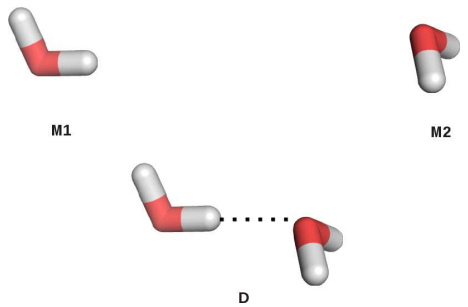

Fig. 4.1. The image of water dimer (D) and its constituent isolated frozen water molecules ($M1$, $M2$) used in the tutorial. Hydrogen bond that stabilizes the dimer is indicated by the dotted line. Pymol [225].

interaction energy,

$$\Delta E = E_D - E_{M1} - E_{M2} \tag{4.1}$$

i.e. we do not account for the structure stabilization effect in the final total energy difference.

The molecular structures are typically stored (on computer) in the so-called `xyz` files. The `.xyz` file with the geometry of water dimer from the A24 [26] database may look e.g. like:

```
6

  O   -0.066999140   0.000000000   1.494354740
  H    0.815734270   0.000000000   1.865866390
  H    0.068855100   0.000000000   0.539142770
  O    0.062547750   0.000000000  -1.422632080
  H   -0.406965400  -0.760178410  -1.771744500
  H   -0.406965400   0.760178410  -1.771744500
```

As illustrated, the `.xyz` file contains a total count of atoms (initial line, one empty line) and one line for each atom, containing its type and a three-dimensional position vector in Cartesian coordinates (in Å).

4.2 Construction of Trial Wave Functions

After preparation of the structures of interest, here water dimer and both water monomers (Figure 4.1), one needs to construct a trial electronic wave function Ψ_T required as an input to FN-DMC (cf. Section 2.4). Here we use a popular Slater-Jastrow form of trial wave function (Section 2.5) that has been proven to be very useful while simple. We need to construct both, the Slater (determinant/s) and the Jastrow parts, merge them and use a variational optimization (VMC, Section 2.6) to minimize the total energy expectation value with respect to the parameters of the Jastrow factor (next section). We do not touch parameters in the Slater part so as to keep the nodal surfaces intact (cf. Sections 2.6 and 3.2).

As discussed in the Section 2.5, we are interested in as few Slater determinants as possible. The optimum strategy, we prefer to use here, uses a single Slater determinant filled with orbitals from DFT. We prefer to use the B3LYP xc functional to produce molecular orbitals. Since the used protocol relies on FN-error cancellation, other self-consistent wave functions may be used as well, e.g. Hartree-Fock or even MC-SCF. The DFT calculations executed with the `GAMESS` code produce outputs that can be converted to a suitable form required by the `QWalk`.

The `GAMESS` code requires text-based input files, that specify all parameters of calculation. The molecular structure, method, Gaussian basis set used in the algebraic approximation (Section 1.3), and other optional parameters.

A simple header of the input file (`D.inp`) may look like:

```
$CONTRL DFTTYP=B3LYP SCFTYP=RHF RUNTYP=ENERGY
        EXETYP=RUN ISPHER=1 MULT=1 ICHARG=0 $END
$GUESS GUESS=HUCKEL $END
```

This input asks `GAMESS` to use DFT with B3LYP xc functional (`DFTTYP=B3LYP`), a spin-restricted form of the determinant (`SCFTYP=RHF`), single point calculation (`RUNTYP=ENERGY`),

i.e. a total energy on a fixed molecule (no structure optimization). It asks for the use of spherical harmonics in the one-particle basis set expansion (`ISPHER=1`) and sets up singlet spin multiplicity (`MULT=1`) and a neutral charge (`ICHARG=0`). The final deck, `$GUESS`, asks for an initial orbital guess of Hückel type [83] that is typically a good starting point in main-group molecules near equilibrium. Note that the `GAMESS` is a FORTRAN 77/90 based program that is sensitive to the position of keywords in the input. Each deck is required to begin in the second row, otherwise the program will terminate with an error indicated in the output file (`D.log` or `D.out`).

In practice, one needs a slightly more complicated form of the input to the `GAMESS`, that includes specific features required for high-accuracy studies of noncovalent interactions. An example being a tight SCF convergence required since our approach relies on the FN-error cancellation (Section 2.2):

```
 $CONTRL SCFTYP=RHF DFTTYP=B3LYP RUNTYP=ENERGY
         EXETYP=RUN ISPHER=1 MULT=1 ICHARG=0 ICUT=11
         ECP=READ $END
 $CONTRL MAXIT=200 $END
 $SCF CONV=1.0D-9 SHIFT=.T. DIIS=.T. FDIFF=.F. $END
 $GUESS GUESS=HUCKEL $END
```

That is achieved by setting `CONV=1.0D-9` in the `$SCF` deck.

After setting up the header part of the input file, one needs to add an information on geometry, basis set and possibly on ECP(s). `GAMESS` provides more options for structure and one-particle basis set input (cf. `GAMESS` manual). Here we use a format where each line of the information on structure (atom, charge, position) is explicitly followed by a basis set for the given atom. For example, such an input for two oxygen atoms 3Å far apart, with a minimum valence atom-centered Gaussian basis set may look like (symmetry point group C_1):

```
 $DATA
Oxygen dimer
C1
O 8.0 0.0 0.0 0.0
s 6 1.00
1 0.125346     0.055741
2 0.268022     0.304848
3 0.573098     0.453752
4 1.225429     0.295926
5 2.620277     0.019567
6 5.602818    -0.128627
p 6 1.00
1 0.083598     0.044958
2 0.167017     0.150175
3 0.333673     0.255999
4 0.666627     0.281879
5 1.331816     0.242835
6 2.660761     0.161134

O 8.0 0.0 0.0 3.0
s 6 1.00
```

```
1 0.125346      0.055741
2 0.268022      0.304848
3 0.573098      0.453752
4 1.225429      0.295926
5 2.620277      0.019567
6 5.602818     -0.128627
p 6 1.00
1 0.083598      0.044958
2 0.167017      0.150175
3 0.333673      0.255999
4 0.666627      0.281879
5 1.331816      0.242835
6 2.660761      0.161134

 $END
```

A basis set for each atom is entered immediately below the given atomic information. Angular momentum, number of primitives and "1.0" are provided for each contraction on the next "number of primitives" lines each providing an index, and, Gaussian exponents and the corresponding coefficients in the contraction expansion (Ref. 58). For large systems, scripts (e.g. bash/awk) or small programs are required to produce similar inputs. Another `GAMESS` syntax where the bases are kept in a separate file may also be considered (cf. `GAMESS` manual).

Finally, if ECPs are required, they must be provided in the `$ECP` deck. For example, the input for the oxygen dimer reads:

```
 $ECP
O-QMC GEN 2 1
3
6.00000000 1 9.29793903
55.78763416 3 8.86492204
-38.81978498 2 8.62925665
1
38.41914135 2 8.71924452
O-QMC
 $END
```

Each ECP contains a first line describing its internal name, the word `GEN`, the number of electrons removed from the core and `1` in that particular order. Then for each channel, a number of Gaussians used to expand ECP is provided, e.g. 3 and 1 in our case. Below follow the Gaussian parameters, coefficient, order corresponding to the angular momentum and exponent. For each species, it is enough to specify ECP only once at a place of the first occurrence and use the label (here `O-QMC`) at the relevant positions later (order of atoms matters). More details may be found in the `GAMESS` manual. The ECPs and bases are tabulated in the required format on the web for a whole periodic table (e.g. EMSL basis set exchange at `https://bse.pnl.gov/bse/portal`).

A full example of DFT input for water dimer required for this tutorial may be found in the Appendix B. Similar input files must also be prepared for monomers (`M1.inp` and `M2.inp`) in order to achieve a final result. In the current example, simple copy of the `D.inp` and deletion

of inappropriate lines (in parts with geometry, basis and ECP) is satisfactory. Successful runs of these files produce output files each containing the line `TOTAL ENERGY` = val with a nonzero value val in a.u. For our case, the runs produce three total energies (in a.u.):

```
D  -34.5214791814
M1 -17.2570591118
M2 -17.2570687677
```

First, we see that the sum of the monomer energies (-34.51413 a.u.) is higher than in the dimer, as expected. The interaction energy ΔE is negative (ΔE =-0.00735 a.u. = -4.613 kcal/mol) which is caused by the presence of hydrogen bond that stabilizes the system. The interaction energy from DFT has already a right order of magnitude. This behavior is typical for hydrogen bonding but not for the London dispersion, which is typically omitted or underestimated in standard DFT without empirical dispersion corrections.

For each input file, `D.inp`, `M1.inp` and `M2.inp`, a successful execution of the `GAMESS` code produces two files for each run (`.log` or `.out` and `.dat`). These files contain all data about the molecules and their electronic structures, including optimized orbital coefficients required for further processing. These files need to be converted to the format required by the `QWalk` software.

The utility able to convert `GAMESS` output to `QWalk` input comes with the `QWalk` package under the name `gamess2qmc` and the compiled package contains it in `src/converter` directory. Having the `GAMESS` files ready for a system of interest, e.g. `D.inp`, `D.dat`, `D.log` and the binary `gamess2qmc` file available, the conversion is executed by the command:

```
gamess2qmc D
```

in the same directory where the mentioned files are present. This command produces various files. The most important ones include:

```
D.basis
D.orb
D.slater
D.sys
```

The `.basis` file contains information on the one-particle basis sets, the `.orb` contains molecular orbital expansion coefficients in the one-electron basis set read from `.basis`, the `.slater` file contains information on determinant (or determinant expansion) including occupation numbers, and the `.sys` file contains molecular structure and optional effective core potentials. For more details, the reader should consult the `QWalk` website. This step accomplishes the construction of the determinant, or, in other words, Slater part of the trial wave function for the water dimer. Similar procedure must be performed for both monomers.

Finally, we need to construct a Jastrow (explicit correlation) term. For all systems, water dimer and both monomers, we enter the following input to the files `D.jastrow`, `M1.jastrow` and `M2.jastrow`:

```
JASTROW2
  GROUP { OPTIMIZEBASIS
    TWOBODY_SPIN {
      FREEZE
```

```
    LIKE_COEFFICIENTS { 0.25 0 }
    UNLIKE_COEFFICIENTS { 0 0.5 }
  }
  EEBASIS {
    EE
    CUTOFF_CUSP
    GAMMA 12
    CUSP 1
    CUTOFF 7.5
  }
  EEBASIS {
    EE
    CUTOFF_CUSP
    GAMMA 24
    CUSP 1
    CUTOFF 7.5
  }
}
GROUP {
  ONEBODY { OPTIMIZEBASIS
    COEFFICIENTS { O  0.0 0.0 0.0 0.0 0.0 0.0 }
    COEFFICIENTS { H  0.0 0.0 0.0 0.0 0.0 0.0 }
  }
  TWOBODY_SPIN {
    LIKE_COEFFICIENTS { 0.0 0.0 0.0 0.0 0.0 0.0 }
    UNLIKE_COEFFICIENTS { 0.0 0.0 0.0 0.0 0.0 0.0 }
  }
  EIBASIS {
    O
    POLYPADE
    RCUT 7.5
    BETA0 -0.8
    NFUNC 6
  }
  EIBASIS {
    H
    POLYPADE
    RCUT 7.5
    BETA0 -0.8
    NFUNC 6
  }
  EEBASIS {
    EE
    POLYPADE
    RCUT 7.5
    BETA0 -0.8
    NFUNC 6
  }
}
```

This Jastrow factor file contains two groups (`Group { ...}`), the first describes cusp conditions (Section 2.5), i.e. short-range electron-electron interactions, while the second part contains explicit electron-nucleus and electron-electron dependencies expanded in polynomial Padé functions (Section 2.5). All linear expansion parameters are initialized to zero whereas the nonlinear parameters [51], `GAMMA` and `BETA0` are initialized to reasonable starting values. An electron-electron-nucleus interactions are available in `qWalk` but we will not use them here for two reasons. First, for simplicity of the tutorial, and second, it has been recently shown that these terms are not critical in noncovalent interactions since high-accuracy interaction energy benchmarks (errors in A24 set of up to $\sim$0.2 kcal/mol) were obtained even with a simple form as used here [54]. For more details on syntax, please refer to the `QWalk` manual and/ or website.

The `QWalk` code requires an input for the total wave function. We want to use a product of Slater part and Jastrow terms just produced. For the dimer, this is accomplished by the use of the Slater-Jastrow wave function syntax:

```
SLATER-JASTROW
WF1 {
 include D.slater
}
WF2 {
 include D.jastrow
}
```

We can write this input to a separate file, e.g. `D.wfin` (cf. above), and point to it in the following calculations.

4.3 Variational Optimization of Trial Wave Functions

In this section, we continue the tutorial with a variational optimization of the Slater-Jastrow trial wave functions. An example of water dimer (D) is considered explicitly, but analogous procedures must be performed on monomers ($M1$, $M2$) in order to obtain three optimal wave functions, one for each structure.

As described in Refs. 53 and 54, an exhaustive optimization of the Jastrow factor is required in order to produce benchmark FN-DMC interaction energies (while retaining the nodes fixed). This may seem counter-intuitive since the FN-DMC should rely on the node locations only. Nevertheless, in the presence of ECPs this is no longer true exactly and as discussed in the Section 2.8, high-quality wave functions must be used to keep the errors from the ECP term/s small. Therefore, we recommend 10 macroiteractions of the Jastrow parameter optimization, each containing VMC run to decorrelate walkers and optimization cycle that itself contains 10 optimization steps (microiterations). An iterative optimization used in the following requires a fixed set of walkers to perform a variational optimization of Jastrow parameters. An initial set of walkers may be generated by the execution of the following input file called `D.00` here:

```
include D.sys
trialfunc { include D.wfin  }
method {
    VMC
    nconfig 128
    nstep 10
```

```
nblock 10
timestep 1.0
ndecorr 2
storeconfig D.00.cfg
}
```

This file asks `QWalk` to perform a VMC run using the molecule specified in `D.sys` with a wave function as given by `D.wfin`. The number of walkers is specified by the `nconfig 128` directive. This input setting is intended for run on 128 cores giving 16k walkers in total. On a serial machine, one should ask for `nconfig 16384` in order to get the corresponding results but that would also take much longer wall time. The product `nstep×nblock×ndecorr` dictates a total number of steps taken by VMC, while `nstep×nblock` steps are taken to averages. A VMC time step (in a.u.) is set by the `timestep 1.0` and a rule of thumb is to optimize this value (by hand) so that the average acceptance of the proposed move is about 0.5. Values that deviate much from 0.5 lead to inefficient sampling. Finally, the directive `storeconfig` forces `QWalk` to store the walkers from the last step to the external file called `D.00.cfg`. The dumped walkers are used in the subsequent optimization step. Successful execution of `D.00` by `qWalk`, i.e. execution of either

```
mpirun -np 128 qwalk D.00
```

command in the parallel MPI environment (here with 128 CPUs), or

```
qwalk D.00
```

on a serial machine, produces output file `D.00.o` and log-file `D.00.log`, and walker file `D.00.cfg`. The final part of the `.o` file may look like:

```
...
****Block 8 avg life 2.232080078 max life 25
Total acceptance 0.4319580078
accept_level0   0.4319580078 tries 40960
      total_energy     kinetic   potential      nonlocal
##0   -34.03982324  25.61833447  -62.62638788   2.968230172  (value)
&&0     1.78545844  10.79617832   12.98868051   7.725685429  (sigma)

****Block 9 avg life 2.155908203 max life 25
...
-----------------Run Ended-----------

Threw out the first 2 blocks in the averaging.
total_energy0  -34.03727464 +/-  0.001223071843 (sigma 1.785981072 )
kinetic0        25.67420271 +/-   0.01192128435 (sigma 10.84522712 )
potential0     -62.67347893 +/-   0.01681769336 (sigma 12.85949717 )
nonlocal0       2.962001576 +/-  0.008519645851 (sigma 7.494211758 )
weight0                   1 +/- 3.510833469e-17 (sigma           0 )

approximate number of independent points: 2132307.173

Maximum lifetime 28 steps
Average lifetime : 2.062632243 steps
VMC Done.
```

I.e. output for each block is printed and the final part shows the results including the total energy -34.037 ± 0.0012 a.u. and its components, and the total energy variance $\sigma = 1.786$. In principle, the total final energy should be equal to the energy of a Slater determinant filled with B3LYP orbitals (i.e. an energy of a given determinant in a HF sense, not a DFT total energy) but since we impose cusp conditions in the Jastrow, the energy may slightly differ (one can use pure Slater wave function to perform tests in suspicious cases). The exponents of the Jastrow part are set to zero in the first step, so the Jastrow term does not affect the energy. This indeed changes during the VMC optimization which lowers the energy and variance due to the recovery of electron correlation. Note that the acceptance ratio is printed along with every block, here e.g. `accept_level0 0.4319580078`. This value should be close to 0.5 and acceptable values range between 0.4 and 0.6. Therefore, the current value is acceptable.

The first optimization step and subsequent VMC run is performed with the input `D.01` to the `QWalk`:

```
include D.sys
trialfunc { D.wfin  }
method {
   OPTIMIZE2
   nconfig 128
   USE_WEIGHTS
   EXTENDED_WFOUTPUT
   MINFUNCTION MIXED
   MIXING 0.95
   iterations 10
   readconfig D.00.cfg
    }
method {
    VMC
    nconfig 128
    nstep 10
    nblock 10
    timestep 1.0
    ndecorr 2
    storeconfig D.01.cfg
    }
```

The first method `OPTIMIZE2` uses a Hessian-driven parametric optimization on a fixed sample of 16k walkers (read from `D.00.cfg`). We ask the code to use a so-called mixed optimization (`MINFUNCTION MIXED`) with a mixing parameter `MIXING 0.95` that asks for 95% of energy and 5% of variance in the VMC cost function. The optimization is immediately followed by a VMC run that uses the optimized trial wave function (the final VMC optimized wave function appears in `D.wfout`) as an input and gets an estimate of its current variational energy. The output `D.01.o` contains output from the optimization process that may look like the following (modified to fit the page):

```
energy= -34.02575097 dispersion= 1.784951283    damping 0
energy= -34.293397   dispersion= 1.098378543    damping 1.734
energy= -34.39505169 dispersion= 0.9035406262   damping 0.578
energy= -34.43944678 dispersion= 0.8434325102   damping 0.192
```

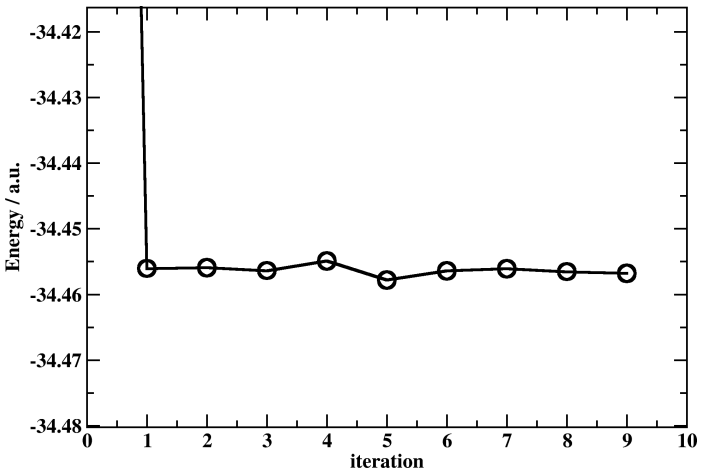

Fig. 4.2. Illustration of the total energy VMC optimization process in water dimer with a Jastrow factor containing electron-electron and electron-nucleus correlation terms (24 parameters).

```
energy= -34.45626056 dispersion= 0.8261633254   damping 0.064
energy= -34.46028562 dispersion= 0.8298016407   damping 0.021
energy= -34.46071783 dispersion= 0.8291122493   damping 0.007
energy= -34.46070401 dispersion= 0.8289127661   damping 0.042
energy= -34.46069802 dispersion= 0.8288406003   damping 0.128
energy= -34.46069799 dispersion= 0.828840127    damping 14.82
energy= -34.46069799 dispersion= 0.8288401213   damping 1187
```

Such an output allows one to monitor the process of optimization. The energy should drop in average from step to step, as well as dispersion (the value in the example is acceptable). Otherwise one needs to increase the total number of walkers and/or generate new set of initial configurations. The subsequent VMC run shows a final line:

```
total_energy0 -34.45608654 +/- 0.000630447426 (sigma 0.845195489 )
```

that may be compared to the previous value -34.037±0.0012 ($\sigma = 1.786$). The first optimization step gained as much as 0.45 a.u. from the unknown amount of the missing correlation energy and the variance dropped significantly, that implies an improved statistical sampling efficiency in the following runs. We should continue with similar iterations, until the final VMC energies (and possibly also parameters) reveal plateau (Figure 4.2 and 4.3). For 10 macroiterations, the VMC energies for this system and the current setting of Slater-Jastrow term reveal the following (or similar) output (collected VMC outputs):

```
-34.03727464 +/- 0.0012230718430 (sigma    1.7859810720 )
-34.45608654 +/- 0.0006304474260 (sigma    0.8451954890 )
-34.45595733 +/- 0.0008388553173 (sigma    0.8431546987 )
-34.45641135 +/- 0.0006066402253 (sigma    0.8400059639 )
-34.45493141 +/- 0.0004683774926 (sigma    0.8371553754 )
-34.45782164 +/- 0.0006786864839 (sigma    0.8446739063 )
-34.45644215 +/- 0.0005223909637 (sigma    0.8395436414 )
```

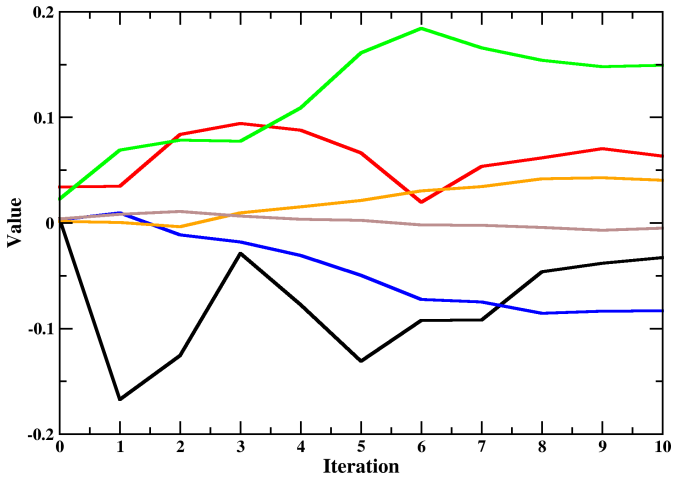

Fig. 4.3. The convergence of electron-hydrogen parameter values along the VMC optimization (macroiterations) in the water dimer. The steps correspond to the energies shown in Figure 4.2.

```
-34.45612170 +/- 0.0005251669739 (sigma    0.8365135158 )
-34.45657306 +/- 0.0005192325894 (sigma    0.8385674579 )
-34.45679469 +/- 0.0003264510037 (sigma    0.8427501227 )
```

The plot of these energies (Figure 4.2) clearly indicates that the one optimization cycle is enough to optimize the energy of a simple Jastrow factor with 24 parameters. However, the parameters need more time to relax, as illustrated in Figure 4.3 for six electron-hydrogen parameters in the Jastrow factor from the current example. In larger molecules, the total energy converges slowly. For instance, in benzene dimer, 7 steps are required to reach a start of a plateau while at least 2 more cycles were required to prove that the total energy is converged - we therefore recommend exhaustive convergence that is typically achieved by a 10×10 rule in small complexes [53] (cf. Section 4.3).

After obtaining a plateau in energy and in parameters, we can safely stop the optimization and use the corresponding final wave function as an input to FN-DMC.

4.4 Fixed-Node Diffusion Monte Carlo

A final production FN-DMC ground-state projection calculations for all three systems (D, M1, M2) may be performed using the optimized trial wave functions and this is always recommended since the VMC optimized Slater-Jastrow wave functions have much lower statistical variance of the total energy than the initial Slater determinant.

An initial thermalization run is required before each FN-DMC statistics-accumulation stage. Here we use the following input to this end, assuming a copy of the optimal wave function in file `D.wfout_final` reads:

```
include D.sys
trialfunc { include D.wfout_final }
```

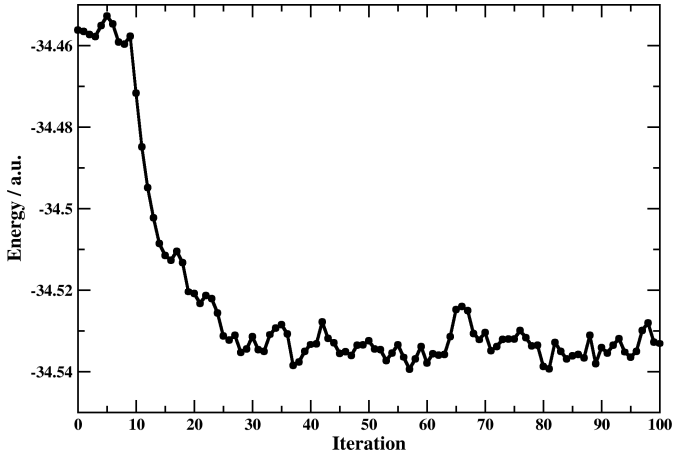

Fig. 4.4. An example of FN-DMC thermalization after initialization VMC run (first 10 steps) for the water dimer with a Slater-Jastrow trial wave function and VMC optimized Jastrow parameters (10 steps/iteration, timestep 0.005 a.u.).

```
method {
     DMC
     TMOVES
     nconfig 128
     nstep 10
     nblock 100
     timestep 0.005
     ndecorr 2
     readconfig D.09.cfg
     storeconfig D.dmc0.cfg
     }
```

As illustrated in Figure 4.4, the thermalization is apparent from the plateau in the total energy which appears after about 500 steps (50 iterations, each 10 steps long). Note that the equilibration run uses smaller number of steps than the typical ”rule of thumb“ (one block should take 1 a.u. of time to let the walkers decorrelate and take statistically independent samples to averages) to allow a better resolution of the plateau. After thermalization, we are ready for the production run. This must be long enough to guarantee that the required error bars are smaller than or equal than the desired value. An example of total energy statistical accumulation run obtained by execution of the following file

```
include D.sys
trialfunc { include D.wfout_final }
method {
     DMC
     TMOVES
     nconfig 128
     nstep 200
```

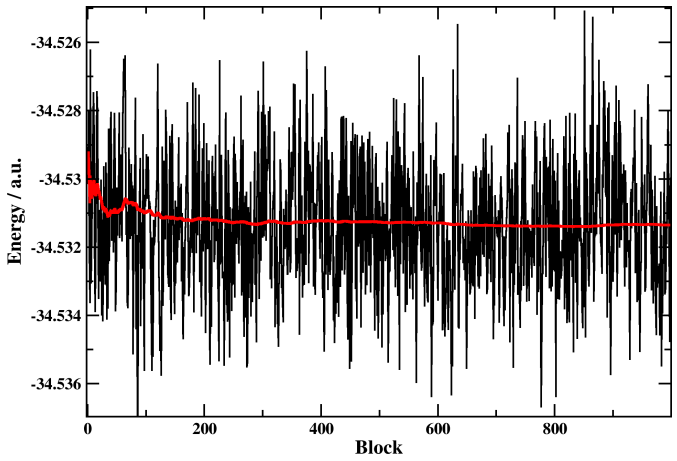

Fig. 4.5. Total energies (black) and actual average energy (red) along the production FN-DMC run (200 steps/block, timestep 0.005 a.u.) in water dimer example.

```
nblock 100
timestep 0.005
readconfig D.dmc0.cfg
storeconfig D.dmc.cfg
}
```

for the water dimer is shown in Figure 4.5. Both, instant total energies (black) and actual average (red) are shown. Note the fluctuations of mean at the beginning. We encourage the readers to always plot similar figures in order to detect a possible drift of the final expectation values that may suggest need for longer runs. Final total energies that enter the interaction energy calculation read

```
D  -34.53136 +/- 6.27E-05 (sigma    0.836 )
M1 -17.26148 +/- 4.28E-05 (sigma    0.592 )
M2 -17.26151 +/- 4.14E-05 (sigma    0.589 )
```

and therefore, our final interaction energy equals to -5.25±0.09 kcal/mol that may be compared to the CCSD(T) reference value of -5.006 kcal/mol [26]. A more accurate value may be obtained with an improved Jastrow factor [54]. Not that FN-DMC significantly improves upon the DFT value obtained above (-4.6 kcal/mol), when compared to the CCSD(T)/CBS.

The reader has now enough information to setup and run an independent interaction energy benchmark calculation based on FN-DMC. Since the methods presented are under development, we strongly encourage the readers to follow citations to this article.

Acknowledgement

The author is grateful to Petr Jurečka for the discussions, comments on the manuscript and motivation to deal with the topic of this review, Lubos Mitas and Michal Bajdich for valuable discussions and Mária Sudolská for a critical review of the manuscript and useful suggestions.

The support from the Ministry of Education, Youth and Sports of the Czech Republic (project LO1305) is gratefully acknowledged.

References

[1] P. Hobza and K. Müller-Dethlefs. *Non-Covalent Interactions*. RSC Theoretical and Computational Chemistry Series. The Royal Society of Chemistry, 2010.

[2] P. Hobza. *Acc. Chem. Res.*, 45:663–672, 2012.

[3] K. E. Riley, M. Pitoňák, P. Jurečka, and P. Hobza. *Chem. Rev.*, 110:5023–5063, 2010.

[4] J. Šponer, J. E. Šponer, A. Mládek, P. Jurečka, P. Banáš, and M. Otyepka. *Biopolymers*, 99:978, 2013.

[5] M. Juríček et al. *Nature Chem.*, 6:222–228, 2014.

[6] M. Mascal, A. Armstrong, and M. D. Bartberger. *J. Am. Chem. Soc.*, 124:6274–6276, 2002.

[7] L. M. Salonen, M. Ellermann, and F. Diederich. *Angew. Chem. Int. Ed.*, 50:4808–4842, 2011.

[8] L. Kronik and A. Tkatchenko. *Acc. Chem. Res.*, 47:3208–3216, 2014.

[9] W. Liu, A. Tkatchenko, and M. Scheffler. *Acc. Chem. Res.*, 47:3369–3377, 2014.

[10] R. Rubbiani et al. *J. Med. Chem.*, 54:8646–8657, 2011.

[11] A. Meyer et al. *Angew. Chem. Int. Ed.*, 51:8895–8899, 2012.

[12] S. Patchkovskii, J. S. Tse, S. N. Yurchenko, L. Zhechkov, T. Heine, and G. Seifert. *Proc. Nat. Acad. Sci.*, 102(30):10439–10444, 2005.

[13] M. Kocman, P. Jurečka, M. Dubecký, M. Otyepka, Y. Cho, and K. S. Kim. *Phys. Chem. Chem. Phys.*, 17:6423–6432, 2015.

[14] V. Georgakilas, M. Otyepka, A. B. Bourlinos, V. Chandra, N. Kim, K. C. Kemp, P. Hobza, R. Zbořil, and K. S. Kim. *Chem. Rev.*, 112:6156–6214, 2012.

[15] M. E. Ali and P. M. Oppeneer. *J. Phys. Chem. Lett.*, 2:939–943, 2011.

[16] M. Kolář, J. Hostaš, and P. Hobza. *Phys. Chem. Chem. Phys.*, 16:9987–9996, 2014.

[17] D. Chandler. *Nature*, 437:640–647, 2005.

[18] E. E. Meyer, K. J. Rosenberg, and J. Israelachvili. *Proc. Natl Acad. Sci. USA*, 103:15739–15746, 2006.

[19] H. Schmidbaur and A. Schier. *Chem. Soc. Rev*, 37:1931–1951, 2008.

[20] P. Pyykkö. *Chem. Soc. Rev.*, 37:1967, 2008.

[21] A. Shokri, Y. Wang, G. A. ODoherty, X.-B. Wang, and S. R. Kass. *J. Am. Chem. Soc.*, 135(47):17919–17924, 2013.

[22] P. Lazar, F. Karlický, P. Jurečka, M. Kocman, E. Otyepková, K. Šafářová, and M. Otyepka. *J. Am. Chem. Soc.*, 135(16):6372–6377, 2013.

[23] G. Chałasiński and M. M. Szczęśniak. *Chem. Rev.*, 100:4227–4252, 2000.

[24] R. E. Dawson et al. *Nature Chem.*, 2:533–538, 2010.

[25] J. C. Amicangelo, D. G. Irwin, C. J. Lee, N. C. Romano, and N. L. Saxton. *J. Phys. Chem. A*, 117(6):1336–1350, 2013.

[26] J. Řezáč and P. Hobza. *J. Chem. Theory Comput.*, 9:2151–2155, 2013.

[27] M. S. Schuurman, S. R. Muir, W. D. Allen, and H. F. Schaefer III. *J. Chem. Phys.*, 120:11586–11599, 2004.

[28] B. Santra, A. Michaelides, M. Fuchs, A. Tkatchenko, C. Filippi, and M. Scheffler. *J. Chem. Phys.*, 129:194111, 2008.

[29] A. Vítek, A. Ofiala, and R. Kalus. *Phys. Chem. Chem. Phys.*, 14:15509–15519, 2012.

[30] M. A. Morales, J. R. Gergely, J. McMinis, J. M. McMahon, J. Kim, and D. M. Ceperley. *J. Chem. Theory Comput.*, 10(6):2355–2362, 2014.

[31] P. Jurečka, J. Šponer, J. Černý, and P. Hobza. *Phys. Chem. Chem. Phys.*, 8:1985–1993, 2006.

[32] T. Takatani, E. G. Hohenstein, M. Malagoli, M. S. Marshall, and C. D. Sherrill. *J. Chem. Phys.*, 132:144104, 2010.

[33] R. Sedlák, T. Janowski, M. Pitoňák, J. Řezáč, P. Pulay, and P. Hobza. *J. Chem. Theory Comput.*, 9(8):3364–3374, 2013.

[34] J. Šponer, N. Mládek, A. Špačková, X. Cang, T. E. Cheatham III, and S. Grimme. *J. Am. Chem. Soc.*, 135:9785, 2013.

[35] J. G. Brandenburg, M. Hochheim, T. Bredow, and S. Grimme. *J. Phys. Chem. Lett.*, 5:4275–4284, 2014.

[36] S. J. Cox, M. D. Towler, D. Alfè, and A. Michaelides. *J. Chem. Phys.*, 140(17):174703, 2014.

[37] J. Šponer, P. Banáš, P. Jurečka, M. Zgarbová, P. Kührová, M. Havrila, M. Krepl, P. Stadlbauer, and M. Otyepka. *J. Phys. Chem. Lett.*, 5(10):1771–1782, 2014.

[38] K. Raghavachari and G. W. Trucks. *Chem. Phys. Lett.*, 157:479–483, 1989.

[39] L. Šimová, J. Řezáč, and P. Hobza. *J. Chem. Theory Comput.*, 9:3420–3428, 2013.

[40] J. Řezáč and P. Hobza. *J. Chem. Theory Comput.*, 10:3066–3073, 2014.

[41] D. G. A. Smith, P. Jankowski, M. Slawik, H. A. Witek, and K. Patkowski. *J. Chem. Theory Comput.*, 10:3140–3150, 2014.

[42] J. Raab, A. Antušek, S. Biskupič, and M. Urban. *Collect. Czech. Chem. Commun.*, 69:189–212, 2004.

[43] R. Podeszwa, K. Patkowski, and K. Szalewicz. *Phys. Chem. Chem. Phys.*, 12(23):5974–5979, 2010.

[44] J. Řezáč, K. E. Riley, and P. Hobza. *J. Chem. Theory Comput.*, 7(8):2427–2438, 2011.

[45] J. Brndiar and I. Štich. *J. Chem. Theory Comput.*, 8(7):2301–2309, 2012.

[46] D. G. A. Smith and K. Patkowski. *J. Chem. Theory Comput.*, 9:370–389, 2013.

[47] M. Sudolská, F. Louis, and I. Černušák. *J. Phys. Chem. A*, 118:9512–9520, 2014.

[48] J. Yang, W. Hu, D. Usvyat, D. Matthews, M. Schütz, and G.K.-L. Chan. *Science*, 345:640–643, 2014.

[49] M. Schütz, O. Masur, and D. Usvyat. *J. Chem. Phys.*, 140(24):244107, 2014.

[50] W. M. C. Foulkes, L. Mitas, R. J. Needs, and G. Rajagopal. *Rev. Mod. Phys.*, 73:33–83, 2001.

[51] M. Bajdich and L. Mitas. *Acta Phys. Slovaca*, 59:81–168, 2009.

[52] B. M. Austin, D. Y. Zubarev, and W. A. Lester. *Chem. Rev.*, 112:263–288, 2012.

[53] M. Dubecký, P. Jurečka, R. Derian, P. Hobza, M. Otyepka, and L. Mitas. *J. Chem. Theory Comput.*, 9:4287–4292, 2013.

[54] M. Dubecký, R. Derian, P. Jurečka, L. Mitas, P. Hobza, and M. Otyepka. *Phys. Chem. Chem. Phys.*, 16:20915–20923, 2014.

[55] A. Ambrosetti, D. Alfè, R. A. DiStasio, and A. Tkatchenko. *J. Phys. Chem. Lett.*, 5:849–855, 2014.

[56] M. Dubecký. *Online preprint*, page arXiv:1501.06673.

[57] I. Štich. *Acta Phys. Slovaca*, 57(1):1–176, 2007.

[58] N. S. Ostlund and A. Szabo. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*, 1996.

[59] S. Wilson. *Electron Correlation in Molecules*. Dover, 1984.

[60] F. Jensen. *Introduction to Computational Chemistry*. Wiley, 2 edition, 1999.

[61] T. Helgaker, P. Jørgensen, and J. Olsen. *Molecular Electronic-Structure Theory*, John Wiley and Sons, Ltd., 2000.

[62] P. Hohenberg and W. Kohn. *Phys. Rev. B*, 136(3B):B864, 1964.

[63] W. Kohn and L. J. Sham. *Phys. Rev.*, 140(4A):1133, 1965.
[64] F. Karlický and M. Otyepka. *Int. J. Quant. Chem.*, 114:987–992, 2014.
[65] F. Karlický, P. Lazar, M. Dubecký, and M. Otyepka. *J. Chem. Theory Comput.*, 9: 3670–3676, 2013.
[66] L. Wagner and L. Mitas. *Chem. Phys. Lett.*, 370:412–417, 2003.
[67] P. M. Polestshuk, P. I. Dem'yanov, and I. G. Ryabinkin. *J. Chem. Phys.*, 129(5):054307, 2008.
[68] L. Horváthová, M. Dubecký, L. Mitas, and I. Štich. *Phys. Rev. Lett.*, 109:053001, 2012.
[69] L. Horváthová, M. Dubecký, L. Mitas, and I. Štich. *J. Chem. Theory Comput.*, 9:390–400, 2013.
[70] M. Caffarel, R. Hernandez-Lamoneda, A. Scemama, and A. Ramirez-Solis. *Phys. Rev. Lett.*, 99(15):153001, 2007.
[71] S. Grimme, J. Anthony, S. Ehrlich, and H. Krieg. *J. Chem. Phys.*, 132:154104, 2010.
[72] J. Klimeš and A. Michaelides. *J. Chem. Phys.*, 137:120901, 2012.
[73] T. H. Dunning Jr. *J. Chem. Phys.*, 90:1007–1023, 1989.
[74] S. F. Boys and F. Bernardi. *Mol. Phys.*, 19(4):553, 1970.
[75] S. Simon, M. Duran, and J.J. Dannenberg. *J. Chem. Phys.*, 105(24):11024–11031, 1996.
[76] L. M. Mentel and E. J. Baerends. *J. Chem. Theor. Comput.*, 10:252–267, 2014.
[77] C. C. J. Roothaan. *Rev. Mod. Phys.*, 23(2):69–89, 1951.
[78] D. C. Meredith and S. E. Koonin. *Computational Physics,* Westview Press, Colorado, 1990.
[79] T. Koopmans. *Physica*, 1(1-6):104 – 113, 1934.
[80] C. C. J. Roothaan. *Rev. Mod. Phys.*, 32:179, 1960.
[81] J. A. Pople, P. A. Dobosh, and J. S. Binkley. *Mol. Phys.*, 28:1423, 1974.
[82] R. K. Nesbet and J. A. Pople. *J. Chem. Phys.*, 22:571, 1954.
[83] Pilar F. L. *Elementary Quantum Chemistry*, Dover Publications, New York, 1990.
[84] V. Bachler, G. Olbrich, F. Neese, and K. Wieghardt. *Inorg. Chem.*, 41(16):4179–4193, 2002.
[85] W. A. Goddard. *Phys. Rev.*, 157(1):81, 1967.
[86] P. O. Löwdin. *Adv. Chem. Phys.*, 2:207–322, 1959.
[87] J. A. Pople and J. S. Binkley. *Mol. Phys.*, 29:599, 1975.
[88] J. W. Hollett and P. M. W. Gill. *J. Chem. Phys.*, 134:114111, 2011.
[89] J. W. Hollett, L. K. McKemmish, and P. M. W. Gill. *J. Chem. Phys.*, 134:224103, 2011.
[90] W. A. Bingel. *Theor. Chem. Acc.*, 5(4):341, 1966.
[91] W. A. Goddard. *Phys. Rev.*, 157(1):73, 1967.
[92] O. Sinanoglu. *Adv. Chem. Phys.*, 6:315–412, 1964.
[93] J. Noga and W. Kutzelnigg. *J. Chem. Phys.*, 101:7738, 1994.
[94] M. Silkowski, M. Lesiuk, and R. Moszynski. *J. Chem. Phys.*, 142:124102, 2015.
[95] J. O. Hirschfelder and J. W. Linnett. *J. Chem. Phys.*, 18(1):130–142, 1950.
[96] G. H. Booth, A. J. W. Thom, and A. Alavi. *J. Chem. Phys.*, 131:054106, 2009.
[97] S. R. Langhoff and E. R. Davidson. *Int. J. Quant. Chem.*, 8(1):61–72, 1974.
[98] J. A. Pople, M. Head-Gordon, and K. Raghavachari. *J. Chem. Phys.*, 87(10):5968–5975, 1987.
[99] S. Grimme, C. Diedrich, and M. Körth. *Angew. Chem. Int. Ed.*, 45:625–629, 2006.
[100] J. Ivanic and K. Ruedenberg. *Theor. Chem. Acc.*, 106(5):339–351, 2001.
[101] A. Lüchow and R.F. Fink. *J. Chem. Phys.*, 113(19):8457–8463, 2000.
[102] P. O. Löwdin. *Phys. Rev.*, 97(6):1474–1489, 1955.
[103] R. S. Grev and H. F. Schaefer. *J. Chem. Phys.*, 96(9):6850–6856, 1992.

[104] B. O. Roos. *Ab initio Methods in Quantum Chemistry II.*, pages 399–444, 1987.
[105] K. Boguslawski, K. H. Marti, and M. Reiher. *J. Chem. Phys.*, 134(22):224101, 2011.
[106] Y. Kurashige, G. K.-L. Chan, and T. Yanai. *Nature Chem.*, 5:660–666, 2013.
[107] S. Sharma, K. Sivalingam, F. Neese, and G. K.-L. Chan. *Nature Chem.*, 6:927–933, 2014.
[108] S. Wouters and D. Van Neck. *Eur. Phys. J. D*, 68:272, 2014.
[109] C. Møller and M. S. Plesset. *Phys. Rev.*, 46:0618–0622, 1934.
[110] H. Nakano. *J. Chem. Phys.*, 99(10):7983–7992, 1993.
[111] F. Schautz, F. Buda, and C. Filippi. *J. Chem. Phys.*, 121(12):5836–5844, 2004.
[112] T. Bouabca, N. Ben Amor, D. Maynau, and M. Caffarel. *J. Chem. Phys.*, 130(11):114107, 2009.
[113] E. J. Carrell, C. M. Thorne, and G. S. Tschumper. *J. Chem. Phys.*, 136:014103, 2012.
[114] M. Iliaš, V. Kellö, and M. Urban. *Acta Phys. Slovaca*, 60(3):259–391, 2010.
[115] Y. S. Al-Hamdani, D. Alfè, O. A. von Lilienfeld, and A. Michaelides. *J. Chem. Phys.*, 141:18C530, 2014.
[116] R. Car and M. Parinello. *Phys. Rev. Lett.*, 55(22):2471–2474, 1985.
[117] D. M. Ceperley and B. J. Alder. *Phys. Rev. Lett.*, 45:566–569, Aug 1980.
[118] J. P. Perdew, K. Burke, and M. Ernzerhof. *Phys. Rev. Lett.*, 77:3865, 1996.
[119] N. C. Handy and A. D. Boese. *J. Chem. Phys.*, 114:5497, 2001.
[120] J. Tao, J. P. Perdew, V. N. Staroverov, and G. E. Scuseria. *Phys. Rev. Lett.*, 91:146401, 2003.
[121] Y. Zhao and D. Truhlar. *Theor. Chem. Acc.*, 120:215–241, 2008.
[122] S. Grimme and F. Neese. *J. Chem. Phys.*, 127(15):154116, 2007.
[123] C. J. Cramer and D. G. Truhlar. *Phys. Chem. Chem. Phys.*, 11:10757–10816, 2009.
[124] C. D. Sherrill. *J. Chem. Phys.*, 132(11):110902, 2010.
[125] P. Mori-Sánchez and A. J. Cohen. *Phys. Chem. Chem. Phys.*, 16:14378, 2014.
[126] J. P. Perdew and K. Schmidt. *Density Functional Theory and its Applications to Materials,* American Institute of Physics: Melville, NY, 2001.
[127] H. Schmidbaur and A. Schier. *Chem. Soc. Rev*, 41:370–412, 2012.
[128] R. F. Liu and et al. *J. Chem. Theory Comput.*, 7:2399, 2011.
[129] M. Dubecký and H. Su. *J. Phys. Chem. C*, 116:17714–17720, 2012.
[130] J. V. Koppen, M. Hapka, M. Modrzejewski, Malgorzata M. Szczesniak, and G. Chalasinski. *J. Chem. Phys.*, 140:244313, 2014.
[131] A. Otero-de-la Roza, J. D. Mallory, and E. R. Johnson. *J. Chem. Phys.*, 140:18A504, 2014.
[132] N. Metropolis and S. Ulam. *J. Am. Stat. Assoc.*, 44(247):335–341, 1949.
[133] M. H. Kalos. *Phys. Rev.*, 128:1791–1795, 1962.
[134] J.B. Anderson. *Quantum Monte Carlo: origins, development, applications*. Oxford, 2006.
[135] H. Conroy. *J. Chem. Phys.*, 41(5):1327, 1964.
[136] H. Conroy. *J. Chem. Phys.*, 41(5):1331, 1964.
[137] H. Conroy. *J. Chem. Phys.*, 41(5):1336, 1964.
[138] H. Conroy. *J. Chem. Phys.*, 41(5):1341, 1964.
[139] J.B. Anderson. *J. Chem. Phys.*, 63(4):1499–1503, 1975.
[140] J. B. Anderson. *J. Chem. Phys.*, 65(10):4121–4127, 1976.
[141] D.J. Klein and H.M. Pickett. *J. Chem. Phys.*, 64(11):4811–4812, 1976.
[142] J.B. Anderson. *Intl. J. Quant. Chem.*, 15:109–120, 1979.

[143] J.B. Anderson. *J. Chem. Phys.*, 73(8):3897–3899, 1980.
[144] J.B. Anderson. *J. Chem. Phys.*, 82(6):2662–2663, 1985.
[145] B. L. Hammond, W. A. Lester Jr., and P. J. Reynolds. World Scientific, 1994.
[146] L. Mitas, E. L. Shirley, and D. M. Ceperley. *J. Chem. Phys.*, 95(5):3467–3475, 1991.
[147] M. Caffarel, J.P. Daudey, J.L. Heully, and A. Ramirez-Solis. *J. Chem. Phys.*, 123(9):094102, 2005.
[148] M. Dubecký, R. Derian, L. Mitas, and I. Štich. *J. Chem. Phys.*, 133:244301, 2010.
[149] M. Bajdich, L. T. Murilo, R. Q. Hood, P. R. C. Kent, and F. Rebodero. *Phys. Rev. Lett.*, 104:193001, 2010.
[150] F. R. Petruzielo, J. Toulouse, and C. J. Umrigar. *J. Chem. Phys.*, 136(12):124116, 2012.
[151] J. Granatier, M. Dubecký, P. Lazar, M. Otyepka, and P. Hobza. *J. Chem. Theory Comput.*, 9:1461–1468, 2013.
[152] J. Kolorenč and L. Mitas. *Phys. Rev. B*, 75(23):235118, 2007.
[153] J. Kolorenč and L. Mitas. *Phys. Rev. Lett.*, 101(18):185502, 2008.
[154] J. Kolorenč, S. Hu, and L. Mitas. *Phys. Rev. B*, 82:115108, 2010.
[155] J. Kolorenč and L. Mitas. *Rep. Prog. Phys.*, 74:026502, 2011.
[156] N. Nemec, M. D. Towler, and R. J. Needs. *J. Chem. Phys.*, 132(3):034111, 2010.
[157] R. J. Needs, M. D. Towler, N. D. Drummond, and P. López Riós. *J. Phys.-Cond. Mat.*, 22:023201, 2010.
[158] F. Schautz and C. Filippi. *J. Chem. Phys.*, 120(23):10931–10941, 2004.
[159] A. Aspuru-Guzik, O. El Akramine, J. C. Grossman, and W. A. Jr. Lester. *J. Chem. Phys.*, 120(7):3049–3050, 2004.
[160] M. Mella and J. B. Anderson. *J. Chem. Phys.*, 119:8225–8228, 2003.
[161] M. Körth, A. Lüchow, and S. Grimme. *J. Phys. Chem. A*, 112:2104–2109, 2008.
[162] A. Zen, Y. Luo, S. Sorella, and L. Guidoni. *J. Chem. Theory Comput.*, 9(10):4332–4350, 2013.
[163] A. Benali, L. Shulenburger, N. A. Romero, J. Kim, and O. A. von Lilienfeld. *J. Chem. Theory Comput.*, 10(8):3417–3422, 2014.
[164] M. J. Deible, O. Tuguldur, and K. D. Jordan. *J. Phys. Chem. B*, 118:8257–8263, 2014.
[165] R. N. Barnett, P. J. Reynolds, and W. A. Lester. *J. Chem. Phys.*, 84(9):4992–4996, 1986.
[166] A. Lüchow. *J. Comp. Chem.*, 35:854–864, 2014.
[167] L. K. Wagner. *Int. J. Quant. Chem.*, 114:94–101, 2014.
[168] P. J. Reynolds, D. M. Ceperley, B. J. Alder, and W. A. Lester. *J. Chem. Phys.*, 77:5593–5603, 1982.
[169] M. D. Towler. *Phys. Stat. Sol. (B)*, 243:2573 – 2598, 2006.
[170] R.C. Grimm and R.G. Storer. *J. Comp. Phys.*, 7(1):134, 1971.
[171] T. Kato. *Commun. Pure App. Math*, 10(2):151–177, 1957.
[172] J.W. Moskowitz, K.E. Schmidt, M.A. Lee, and M.H. Kalos. *J. Chem. Phys.*, 77(1):349–355, 1982.
[173] K. Rasch, S. Hu, and L. Mitas. *J. Chem. Phys.*, 140:041102, 2014.
[174] M. P. Nightingale, K. J. Runge, and C. J. Umrigar. *J. Chem. Phys.*, 99:2865, 1993.
[175] R. Assaraf and M. Caffarel. *J. Chem. Phys.*, 119(20):10536–10552, 2003.
[176] R. Jastrow. *Phys. Rev.*, 98:1479–1484, 1955.
[177] C. J. Huang, C. Filippi, and C. J. Umrigar. *J. Chem. Phys.*, 108(21):8838–8847, 1998.
[178] K. E. Schmidt and J. W. Moskowitz. *J. Chem. Phys.*, 93(6):4172–4178, 1990.
[179] L. K. Wagner, M. Bajdich, and L. Mitas. *J. Comp. Phys.*, 228:3390–3404, 2009.

[180] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. *J. Chem. Phys.*, 21(6):1087–1092, 1953.

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

[182] C. J. Umrigar, K. G. Wilson, and J. W. Wilkins. *Phys. Rev. Lett.*, 60(17):1719–1722, 1988.

[183] C. J. Umrigar and C. Filippi. *Phys. Rev. Lett.*, 94:150201, 2005.

[184] M. Meierovich, A. Mushinski, and M.P. Nightingale. *J. Chem. Phys.*, 105(15):6498–6504, 1996.

[185] D. M. Ceperley. *J. Stat. Phys.*, 43(5-6):815–826, 1986.

[186] M. A. Morales, J. McMinis, B. K. Clark, J. Kim, and G. E. Scuseria. *J. Chem. Theory Comput.*, 8:2181–2188, 2012.

[187] A. Badinski, P.D. Haynes, J.R. Trail, and R.J. Needs. *J. Phys.-Cond. Mat.*, 22(7):074202, 2010.

[188] S. Moroni, S. Saccani, and C. Filippi. *J. Chem. Theory Comput.*, 10:4823–4829, 2014.

[189] C. Filippi and C. J. Umrigar. *Phys. Rev. B*, 61(24):R16291–R16294, 2000.

[190] J.C. Grossman and L. Mitas. *Phys. Rev. Lett.*, 94(5):056403, 2005.

[191] R.P. Feynman. *Phys. Rev.*, 56(4):340–343, 1939.

[192] X. Lin, H.K. Zhang, and A.M. Rappe. *J. Chem. Phys.*, 112(6):2650–2654, 2000.

[193] J.C. Philips and L. Kleinman. *Phys. Rev.*, 116:287–294, 1959.

[194] J. R. Trail and R. J. Needs. *J. Chem. Theory Comput.*, 10(5):2049–2053, 2014.

[195] B. L. Hammond, P. J. Reynolds, and W. A. Lester. *J. Chem. Phys.*, 87(2):1130–1136, 1987.

[196] L. Wagner. *Quantum Monte Carlo for Transition Metal Systems: Method Developments and Applications*, PhD. Thesis, 2006.

[197] L. Mitas. *Phys. Rev. A*, 49(6):4411–4414, 1994.

[198] M. Burkatzki, C. Filippi, and M. Dolg. *J. Chem. Phys.*, 126:234105, 2007.

[199] M. Burkatzki, C. Filippi, and M. Dolg. *J. Chem. Phys.*, 129(16):164115, 2008.

[200] M. Casula. *Phys. Rev. B*, 74:161102, 2006.

[201] C. W. Greeff and W. A. Lester. *J. Chem. Phys.*, 109(5):1607–1612, 1998.

[202] I. Ovcharenko, A. Aspuru-Guzik, and Jr. W. A. Lester. *The J. Chem. Phys.*, 114(18):7790–7794, 2001.

[203] J.R. Trail and R.J. Needs. *J. Chem. Phys.*, 122(17):174109, 2005.

[204] J.R. Trail and R.J. Needs. *J. Chem. Phys.*, 122(1):014112, 2005.

[205] M. Dubecký et al. *To be published.*

[206] C. Diedrich, A. Lüchow, and S. Grimme. *J. Chem. Phys.*, 123:184106, 2005.

[207] S. Sorella, M. Casula, and D. Rocca. *J. Chem. Phys.*, 127:014105, 2007.

[208] F. Sterpone, L. Spanu, L. Ferraro, S. Sorella, and L. Guidoni. *J. Chem. Theory Comput.*, 4:1428–1434, 2008.

[209] N. A. Benedek, I. K. Snook, M. D. Towler, and R. J. Needs. *J. Chem. Phys.*, 125:104302, 2006.

[210] M. Bajdich, F. Rebodero, and P. R. C. Kent. *Phys. Rev. B*, 82:081405, 2010.

[211] M. Körth, S. Grimme, and M. D. Towler. *J. Phys. Chem. A*, 115:11734–11739, 2011.

[212] P. Ganesh et al. *J. Chem. Theory Comput.*, 10(12):5318–5323, 2014.

[213] J. Ma, D. Alfé, A. Michaelides, and E. Wang. *J. Chem. Phys.*, 130:154303, 2009.

[214] A. Tkatchenko, D. Alfé, and K. S. Kim. *J. Chem. Theory Comput.*, 8:4317–4322, 2012.

[215] K. Hongo, M. A. Watson, R. S. Sánchez-Carrera, T. Iitaka, and A. Aspuru-Guzik. *J. Phys. Chem. Lett.*, 1(12):1789–1794, 2010.

[216] K. Hongo, T. C. Nguyen, and R. Maezono. *J. Chem. Theory Comput.*, 9:1081–1086, 2013.
[217] D. Alfè, A. P. Bartók, G. Csányi, and M. J. Gillan. *J. Chem. Phys.*, 138(22):221102, 2013.
[218] L. Shulenburger, M. P. Desjarlais, and T. R. Mattsson. *Phys. Rev. B*, 90:140104, 2014.
[219] M. Dolg and C. Filippi. *Private communication.*
[220] M. W. Schmidt, K. K. Baldridge, J. A. Boatz, and et. al. *J. Comput. Chem.*, 14:1347–1363, 1993.
[221] J. W. Moskowitz and K. E. Schmidt. *J. Chem. Phys.*, 97:3382–3385, 1992.
[222] N. D. Drummond and R. J. Needs. *Phys. Rev. B*, 72(8):085124, 2005.
[223] M. Dubecký, R. Derian, L. Horváthová, M. Allan, and I. Štich. *Phys. Chem. Chem. Phys.*, 13:20939–20945, 2011.
[224] M. C. Per, K. A. Walker, and S. P. Russo. *J. Chem. Theory Comput.*, 8:2255–2259, 2012.
[225] The PyMOL Molecular Graphics System; Version 1.3r Schrödinger; LLC. 2010.

A Metropolis Method

The Metropolis algorithm [180] allows for a straightforward stochastic sampling of an arbitrarily complex distribution function with unknown normalization. An idea is to generate a Markov chain of sampling points by moving (each) walker by the following rules [50]:

i) start the walker at a random position $\mathbf{R}$,

ii) propose a trial move to a new point $\mathbf{R}^{'}$ according to a chosen transition probability density function $T(\mathbf{R}^{'} \leftarrow \mathbf{R})$ (e.g. Green's function),

iii) accept the proposed move with a probability

$$\mathrm{Min}\Big(1, \frac{P(\mathbf{R}^{'})T(\mathbf{R}^{'} \leftarrow \mathbf{R})}{P(\mathbf{R})T(\mathbf{R} \leftarrow \mathbf{R}^{'})}\Big) \tag{A.1}$$

where P is a given guiding function (e.g. trial wave function Ψ_T). If the move is accepted, the walk continues from point $\mathbf{R}^{'}$, otherwise the point $\mathbf{R}$ is used in the next step.

iv) return to the point ii).

After the initial thermalization period, that depends on the starting point(s), the walkers sample P [50].

B DFT Input: Water Dimer

```
 $CONTRL SCFTYP=RHF DFTTYP=B3LYP RUNTYP=ENERGY
  ECP=READ EXETYP=RUN ISPHER=1 ICHARG=0 MULT=1 ICUT=11 $END
 $CONTRL MAXIT=200 $END
 $SCF CONV=1.0D-9 FDIFF=.F. SHIFT=.T. DIIS=.T. $END
 $GUESS GUESS=HUCKEL $END
 $DATA
A24-02-water_water_Cs
C1
O 8.0 -0.066999140 0.000000000 1.494354740
s 9 1.00
1 0.125346     0.055741
2 0.268022     0.304848
3 0.573098     0.453752
4 1.225429     0.295926
5 2.620277     0.019567
6 5.602818    -0.128627
7 11.980245    0.012024
8 25.616801    0.000407
9 54.775216   -0.000076
s 1 1.00
1 1.686633     1.000000
s 1 1.00
1 0.237997     1.000000
p 9 1.00
1 0.083598     0.044958
2 0.167017     0.150175
3 0.333673     0.255999
4 0.666627     0.281879
5 1.331816     0.242835
6 2.660761     0.161134
7 5.315785     0.082308
8 10.620108    0.039899
9 21.217318    0.004679
p 1 1.00
1 0.184696     1.000000
p 1 1.00
1 0.600621     1.000000
d 1 1.00
1 0.669340     1.000000
d 1 1.00
1 2.404278     1.000000
f 1 1.00
1 1.423104     1.000000
S   1
  1      0.0737600             1.0000000
P   1
  1      0.0597400             1.0000000
```

```
D   1
  1      0.2140000              1.0000000
F   1
  1      0.5000000              1.0000000

H 1.0 0.815734270 0.000000000 1.865866390
s 3 1.00
1 8.70088777  0.01416863
2 1.95552046  0.05917887
3 0.55904363  0.17101570
s 1 1.0
1 0.16492543  1.00000000
s 1 1.0
1 0.03206251  1.00000000
p 1 1.0
1 0.79007442  1.00000000
p 1 1.0
1 0.15235137  1.00000000
d 1 1.0
1 0.88381786  1.00000000
S   1
  1      0.0252600              1.0000000
P   1
  1      0.1020000              1.0000000
D   1
  1      0.2470000              1.0000000

H 1.0 0.068855100 0.000000000 0.539142770
s 3 1.00
1 8.70088777  0.01416863
2 1.95552046  0.05917887
3 0.55904363  0.17101570
s 1 1.0
1 0.16492543  1.00000000
s 1 1.0
1 0.03206251  1.00000000
p 1 1.0
1 0.79007442  1.00000000
p 1 1.0
1 0.15235137  1.00000000
d 1 1.0
1 0.88381786  1.00000000
S   1
  1      0.0252600              1.0000000
P   1
  1      0.1020000              1.0000000
D   1
  1      0.2470000              1.0000000
```

```
O 8.0 0.062547750 0.000000000 -1.422632080
s 9 1.00
1 0.125346     0.055741
2 0.268022     0.304848
3 0.573098     0.453752
4 1.225429     0.295926
5 2.620277     0.019567
6 5.602818    -0.128627
7 11.980245    0.012024
8 25.616801    0.000407
9 54.775216   -0.000076
s 1 1.00
1 1.686633     1.000000
s 1 1.00
1 0.237997     1.000000
p 9 1.00
1 0.083598     0.044958
2 0.167017     0.150175
3 0.333673     0.255999
4 0.666627     0.281879
5 1.331816     0.242835
6 2.660761     0.161134
7 5.315785     0.082308
8 10.620108    0.039899
9 21.217318    0.004679
p 1 1.00
1 0.184696     1.000000
p 1 1.00
1 0.600621     1.000000
d 1 1.00
1 0.669340     1.000000
d 1 1.00
1 2.404278     1.000000
f 1 1.00
1 1.423104     1.000000
S   1
  1      0.0737600              1.0000000
P   1
  1      0.0597400              1.0000000
D   1
  1      0.2140000              1.0000000
F   1
  1      0.5000000              1.0000000

H 1.0 -0.406965400 -0.760178410 -1.771744500
s 3 1.00
1 8.70088777  0.01416863
2 1.95552046  0.05917887
3 0.55904363  0.17101570
```

```
s 1 1.0
1 0.16492543  1.00000000
s 1 1.0
1 0.03206251  1.00000000
p 1 1.0
1 0.79007442  1.00000000
p 1 1.0
1 0.15235137  1.00000000
d 1 1.0
1 0.88381786  1.00000000
S   1
  1      0.0252600              1.0000000
P   1
  1      0.1020000              1.0000000
D   1
  1      0.2470000              1.0000000

H 1.0 -0.406965400 0.760178410 -1.771744500
s 3 1.00
1 8.70088777  0.01416863
2 1.95552046  0.05917887
3 0.55904363  0.17101570
s 1 1.0
1 0.16492543  1.00000000
s 1 1.0
1 0.03206251  1.00000000
p 1 1.0
1 0.79007442  1.00000000
p 1 1.0
1 0.15235137  1.00000000
d 1 1.0
1 0.88381786  1.00000000
S   1
  1      0.0252600              1.0000000
P   1
  1      0.1020000              1.0000000
D   1
  1      0.2470000              1.0000000

 $END
 $ECP
O-QMC GEN 2 1
3
  6.00000000 1 9.29793903
 55.78763416 3 8.86492204
-38.81978498 2 8.62925665
1
38.41914135 2 8.71924452
H-QMC  GEN 0 0
```

```
3
 1.000000000000 1 25.000000000000
25.000000000000 3 10.821821902641
-8.228005709676 2  9.368618758833
H-QMC
O-QMC
H-QMC
H-QMC
 $END
```

Dr. Matúš Dubecký graduated in 2008 at the Slovak University of Technology in Bratislava. In 2012, he obtained a PhD degree in solid state physics from Comenius University in Bratislava. The PhD study was pursued at the Institute of Physics, Slovak Academy of Sciences and focused on excited states of molecules by quantum Monte Carlo methods. He is now a postdoc at the Regional Centre of Advanced Technologies and Materials, Palacký University Olomouc, Czech Republic, and his current research interests include method development for ab initio electronic structure computations, including quantum Monte Carlo and density matrix renormalization. Applications of these methods include non-covalent interactions in molecular complexes and strongly correlated molecules/solids.