

Density matrix functional theory vis-à-vis density functional theory

16.4.2007

Ryan Requist
Oleg Pankratov

1 Introduction

Recently, there has been renewed interest in density matrix functional theory (DMFT) [1], which is a formalism for treating the ground state of a many-electron problem. Approximations for the exchange-correlation (xc) energy functional have been found, which are fairly accurate for atoms, small molecules, and the homogeneous electron gas. The level of accuracy approaches that of computationally more demanding methods, such as the method of configuration interaction.

Formally, DMFT is similar to density functional theory (DFT). The chief difference is that in DMFT the basic variable is the one-particle reduced density matrix (referred to as the “1-matrix” in the following) instead of the density. Both the 1-matrix and the density are single-particle quantities, but the 1-matrix contains more information: it contains *all* single-particle information, e.g., the current density, kinetic energy density, magnetization density, and of course, also the particle density. The additional information content of the 1-matrix may assist in the construction of approximate xc functionals. Indeed, consider a quantity with yet more information content, the two-particle reduced density matrix (2-matrix.) The energy is a simple *linear* functional of the 2-matrix (see Sec. 3.) In DFT, on the other hand, the ground state energy is a complicated nonlinear and nonlocal functional of the density. In terms of information content, the 1-matrix is intermediate between the density and the 2-matrix. One may ask why the simple 2-matrix energy functional is not minimized directly. This cannot be done, because it is not known how to restrict the trial space to “physical” 2-matrices, that is, 2-matrices that could be obtained from the contraction of *some* many-particle wavefunction. Fortunately, it is known how to restrict the trial space to “physical” 1-matrices, and the energy can be minimized directly by searching over all trial 1-matrices.

It is important for 1-matrix functional theory that the kinetic energy is a known functional of the 1-matrix. This is not the case in DFT. The kinetic energy is not a functional of the density alone, and it must be calculated indirectly. The most successful approach is to map the many-particle problem onto a fictitious system of noninteracting particles, called the Kohn-Sham system, whose kinetic energy is readily evaluated, though it does not coincide with the exact kinetic energy. For the reasons above, the approximate xc energy functionals of DMFT have a different structure than the common approximations of DFT, such as the local density approximation (LDA) and generalized gradient approximation (GGA).

This report is structured as follows. In section 2, we review the concept of the density matrix. In section 3, we outline the foundations of DMFT. In section 4, we describe approximations that have been proposed for the energy functional in DMFT. In section 5, we survey some of the results obtained using these approximations. The conclusions and outlook are given in section 6.

2 Review of the density matrix

Depending on the context, the state of a quantum system is described by either a wavefunction or a density matrix. A system that can be described by a wavefunction is said to be in a pure state. It is not always possible to describe a system by a wavefunction. In particular, this is true of open systems and situations where the state is incompletely specified. Such a system is said to be in a mixed state and is described by a density matrix,

$$\rho(x_1, x_2, \dots, x_N; x'_1, x'_2, \dots, x'_N) = \sum_i w_i \Psi_i(x_1, x_2, \dots, x_N) \Psi_i^*(x'_1, x'_2, \dots, x'_N) \quad (1)$$

where $x = (r, \sigma)$. Expanded in a complete basis for the system $\{\phi_j\}$, the many-particle states are $\Psi_i = \sum_j c_{ij} \phi_j$. In the $\{\phi_j\}$ basis, the density matrix has elements

$$\rho_{jk} = \sum_i w_i c_{ij} c_{ik}^*. \quad (2)$$

The statistical weights w_i represent incomplete knowledge of the state. The statistical nature of the density matrix is clearly illustrated in the calculation

of averages. Consider the expectation value of an observable \hat{A} in a pure state

$$\langle \hat{A} \rangle = \sum_a a |\langle a | \Psi \rangle|^2, \quad (3)$$

where $|a\rangle$ is the eigenstate of \hat{A} with eigenvalue a . For a mixed state, the expectation value is the statistical average of the expectation values of the individual states that comprise the mixture,

$$\begin{aligned} \langle \hat{A} \rangle &= \sum_i w_i \langle \Psi_i | \hat{A} | \Psi_i \rangle \\ &= \sum_a a \langle a | \sum_i w_i |\Psi_i\rangle \langle \Psi_i| | a \rangle \\ &= \text{Tr}(\hat{\rho} \hat{A}). \end{aligned} \quad (4)$$

The last line is the familiar expression for calculating averages with respect to a mixed state with the density matrix expressed as an operator,

$$\hat{\rho} = \sum_i w_i |\Psi_i\rangle \langle \Psi_i|. \quad (5)$$

The basic variable of density matrix functional theory is the 1-matrix. The 1-matrix is a contraction of the full many-particle density matrix,

$$\rho_1(x; x') = N \int dx_2 \dots dx_N \rho(x, x_2, \dots, x_N; x', x_2, \dots, x_N) \quad (6)$$

where $\int dx = \sum_\sigma \int dr$. An equivalent definition is

$$\rho_1(x; x') = \text{Tr}(\hat{\rho} \hat{\psi}^\dagger(x') \hat{\psi}(x)), \quad (7)$$

where $\hat{\psi}^\dagger(x)$ and $\hat{\psi}(x)$ are creation and annihilation field operators that satisfy the fermionic commutation relation $\{\hat{\psi}(x), \hat{\psi}^\dagger(x')\} = \delta_{\sigma, \sigma'} \delta(r - r')$. The 1-matrix is Hermitean and may always be diagonalized. The eigenstates of the 1-matrix are called natural orbitals (NO's). The eigenvalue equation for a NO with spin σ (spin up or spin down) is

$$\int dr' \rho_1(r\sigma; r'\sigma) \chi_{i\sigma}(r') = n_{i\sigma} \chi_{i\sigma}(r). \quad (8)$$

The eigenvalue $n_{i\sigma}$ is the occupancy of the corresponding natural orbital $\chi_{i\sigma}$. Natural orbitals play a central role in DMFT.

The two-particle reduced density matrix, the 2-matrix, is defined as

$$\rho_2(x_1, x_2; x'_1, x'_2) = \binom{N}{2} \int dx_3 \dots dx_N \rho(x_1, x_2, x_3 \dots x_N; x'_1, x'_2, x_3 \dots x_N). \quad (9)$$

3 Foundations of density matrix functional theory

DMFT and DFT are different approaches to the many-electron problem with Hamiltonian

$$\hat{H} = \sum_{i=1}^N \frac{p_i^2}{2m} + \sum_{i=1}^N \int dr'_i v(r_i, r'_i) + \sum_{i,j}^{i \neq j} \frac{e^2}{|r_i - r_j|}, \quad (10)$$

where the external potential $v(r, r')$ is nonlocal in general. DFT is limited to the case of local external potential, i.e. $v(r, r') = v(r)\delta(r, r')$. The energy of a pure state Ψ is

$$\begin{aligned} E &= \langle \Psi | \hat{H} | \Psi \rangle \\ &= \text{Tr}(\hat{\rho} \hat{H}), \end{aligned} \quad (11)$$

where $\hat{\rho} = |\Psi\rangle \langle \Psi|$ is the density matrix associated with the pure state Ψ . As the Hamiltonian contains only one-particle and two-particle operators, the energy can be expressed as a functional of the 1-matrix and 2-matrix. Consider the kinetic energy operator in terms of the field operators

$$\hat{T} = \int d^3x \hat{\psi}^\dagger(x) \frac{p^2}{2m} \hat{\psi}(x). \quad (12)$$

The kinetic energy of a state Ψ is

$$\begin{aligned}
\langle \Psi | \hat{T} | \Psi \rangle &= \int d^3x \left\langle \Psi \left| \hat{\psi}^\dagger(x) \frac{p^2}{2m} \hat{\psi}(x) \right| \Psi \right\rangle \\
&= \int \int d^3x d^3x' \delta_{\sigma,\sigma'} \delta(r-r') \left\langle \Psi \left| \hat{\psi}^\dagger(x') \frac{p^2}{2m} \hat{\psi}(x) \right| \Psi \right\rangle \\
&= \int \int d^3x d^3x' \delta_{\sigma,\sigma'} \delta(r-r') \left(-\frac{1}{2m} \nabla_r^2 \right) \left\langle \Psi \left| \hat{\psi}^\dagger(x') \hat{\psi}(x) \right| \Psi \right\rangle \\
&= \int \int d^3x d^3x' \delta(x-x') \left(-\frac{1}{2m} \nabla_r^2 \right) \rho_1(x; x'), \tag{13}
\end{aligned}$$

which is a linear functional of the 1-matrix. It is also a linear functional of the 2-matrix, because the 1-matrix is itself a linear functional of the 2-matrix. Similarly, the energy of interaction with the external potential is

$$\begin{aligned}
\langle \Psi | \hat{V} | \Psi \rangle &= \int \int d^3x d^3x' \delta_{\sigma,\sigma'} v(r, r') \rho_1(x; x') \\
&= \sum_{\sigma} \int \int d^3r d^3r' v(r, r') \rho_1(r\sigma; r'\sigma). \tag{14}
\end{aligned}$$

The electron-electron interaction is a two-particle interaction. In terms of the field operators, it may be written

$$\begin{aligned}
\hat{U} &= \frac{e^2}{2} \int \int dx dx' \frac{\hat{n}(x) \hat{n}(x')}{|r-r'|} \\
&= \frac{e^2}{2} \int \int dx dx' \hat{\psi}^\dagger(x) \hat{\psi}(x) \frac{1}{|r-r'|} \hat{\psi}^\dagger(x') \hat{\psi}(x'), \tag{15}
\end{aligned}$$

where in the first line the operator $\hat{n}(x)$ replaces $n(x)$ in the classical expression for the electrostatic energy. The interaction energy is not a known functional of the 1-matrix. It is, however, a linear functional of the 2-matrix

$$\begin{aligned}
\langle \Psi | \hat{U} | \Psi \rangle &= \frac{e^2}{2} \int \int dx dx' \frac{1}{|r-r'|} \left\langle \Psi \left| \hat{\psi}^\dagger(x) \hat{\psi}^\dagger(x') \hat{\psi}(x') \hat{\psi}(x) \right| \Psi \right\rangle \\
&= \int \int dx dx' \frac{e^2}{|r-r'|} \rho_2(x, x'; x, x'). \tag{16}
\end{aligned}$$

Therefore, the energy is a linear functional of the 2-matrix. At the same time, the Hohenberg-Kohn (HK) theorem asserts that the ground state energy is a variational functional of the density alone. This is remarkable because the density is a much simpler quantity than the 1-matrix or 2-matrix.

The HK theorem was extended by Gilbert to DMFT [1], where it states that the ground state energy is a variational functional of the 1-matrix. This energy functional may be minimized by searching over all admissible trial 1-matrices. The following conditions are satisfied by an admissible 1-matrix: 1) it is Hermitean, 2) the natural orbitals are orthogonal, 3) the natural orbitals have occupancies between 0 and 1, i.e., $0 \leq n_{i\sigma} \leq 1$, and 4) the sum of the occupancies equals the number of particles. A necessary condition for an energy minimum is that the energy functional is stationary with respect to variations of the 1-matrix. Conditions 2) and 4) are enforced by Lagrange multipliers. We introduce the energy functional

$$\mathcal{E}[\chi_{i\sigma}, n_{i\sigma}] = E[\chi_{i\sigma}, n_{i\sigma}] - \mu \left(\sum_{i\sigma} n_{i\sigma} - N \right) - \sum_{i\sigma, j\sigma'} \lambda_{i\sigma, j\sigma'} (\langle \chi_{i\sigma} | \chi_{j\sigma'} \rangle - \delta_{ij} \delta_{\sigma\sigma'}), \quad (17)$$

where the occupancies are confined to lie between 0 and 1. The stationary conditions are

$$\frac{\delta \mathcal{E}}{\delta \chi_{i\sigma}} = 0, \quad \frac{\delta \mathcal{E}}{\delta \chi_{i\sigma}^*} = 0, \quad \text{and} \quad \frac{\partial \mathcal{E}}{\partial n_{i\sigma}} = 0. \quad (18)$$

As we have considered a spin-independent Hamiltonian, the natural orbitals are spin-restricted orbitals, i.e., each member of a complete and orthonormal set of spatial orbitals is multiplied by a spin factor which is spin up or spin down. Therefore, spin up and spin down orbitals are mutually orthogonal.

4 Müller approximation

It has been proved that the energy is a variational functional of the 1-matrix. However, the functional is unknown in general. As shown in the previous section, some parts of the energy are known functionals of the 1-matrix, namely the kinetic energy and interaction energy with the external potential. Only the electron-electron interaction energy needs to be approximated. In DMFT, approximations are constructed in terms of the natural orbitals and occupancies. Therefore, they have a different structure than the common approximations in DFT, such as the local density approximation (LDA) and generalized gradient approximation (GGA). Most of the approximations proposed for the electron-electron interaction energy functional $U[\chi_{i\sigma}, n_{i\sigma}]$ are

descendants of the following functional introduced by Müller [2]

$$\begin{aligned}
U_M[\chi_{i\sigma}, n_{i\sigma}] &= \sum_{i\sigma, j\sigma'} n_{i\sigma} n_{j\sigma'} \frac{e^2}{2} \int \int dr dr' \frac{\phi_{i\sigma}^*(r) \phi_{j\sigma'}^*(r') \phi_{i\sigma}(r) \phi_{j\sigma'}(r')}{|r - r'|} \\
&+ \sum_{i, j, \sigma} \sqrt{n_{i\sigma} n_{j\sigma}} \frac{e^2}{2} \int \int dr dr' \frac{\phi_{i\sigma}^*(r) \phi_{j\sigma}^*(r') \phi_{j\sigma}(r) \phi_{i\sigma}(r')}{|r - r'|} \quad (19)
\end{aligned}$$

It is a modification of the Hartree-Fock (HF) energy functional, where in the exchange term $n_{i\sigma} n_{j\sigma}$ is replaced by $\sqrt{n_{i\sigma} n_{j\sigma}}$. This modification is intended to account for correlation, which is absent, by definition, in the HF approximation. The presence of correlation is reflected in the appearance of fractional occupancies at the energy minimum, whereas in the HF approximation the occupancies are always 0 or 1. The square root factor $\sqrt{n_{i\sigma} n_{j\sigma}}$ in the exchange term arises in the following way. Recall that the interaction energy is a linear functional of the 2-matrix. Therefore, an approximation of the interaction energy implies an approximation for the 2-matrix and vice versa. Müller modified the exchange term by the replacement $n_{i\sigma} n_{j\sigma} \rightarrow (n_{i\sigma})^p (n_{j\sigma})^{1-p}$. Such a modification has the virtue that it implies an approximation for the 2-matrix that satisfies exactly the rule

$$\int dx_2 \rho_2(x_1, x_2; x'_1, x_2) = \frac{N-1}{2} \rho_1(x_1, x'_1), \quad (20)$$

which follows directly from the definitions of the 1-matrix and 2-matrix. The particular exponent $p = 1/2$ is found upon minimizing the violation of the Pauli principle. Specifically, consider the diagonal element of the two-particle density $n_2(x, x) = \rho_2(x, x; x, x)$. It should be zero owing to the Pauli principle. However, after the modification introduced by Müller, $n_2(x, x) \leq 0$. The unique exponent $p = 1/2$ results from maximizing $n_2(x, x)$. At the time of writing, most of the approximate functionals that have been proposed are related to the Müller functional.

5 Survey of applications

In this section, we review some results of the application of approximate energy functionals to three types of problems – the homogeneous electron gas (HEG), atoms, and small molecules.

5.1 Homogeneous electron gas

The homogeneous electron gas (HEG) is an important system by which to test approximate functionals, because it has been well-studied. There is a simplification regarding the application of DMFT. Due to the translational invariance of the system, the natural orbitals are plane waves. Therefore, only the occupancies of the plane waves must be varied in the minimization of the energy. The occupancies of the plane waves are described by the momentum distribution, $n(k)$. A parametrization [3] of the momentum distribution of the HEG as calculated by quantum Monte Carlo simulations is shown in Figure 1. An important property of the momentum distribution

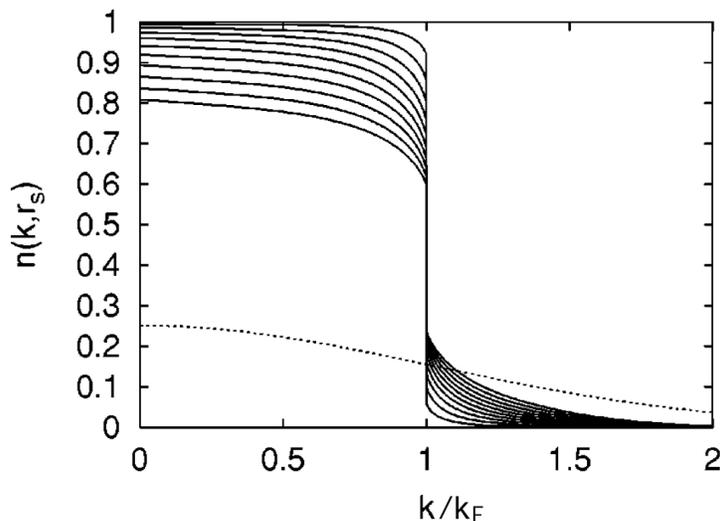


Figure 1: A parametrization of the momentum distribution of the HEG for different values of r_s . Adapted from Ref. [3]

is the smearing of occupancy over the Fermi level and the finite discontinuity at the Fermi level. These properties of the momentum distribution are consequences of the electron-electron interaction. If there were no interactions, the lowest energy plane waves would be occupied up to the Fermi level, that is, the momentum distribution would be a step function. Due to the electron-electron interaction, the relative positions of the electrons in the gas are correlated. This correlation tends to keep electrons away from each other so that around every electron there is an “exchange-correlation hole,”

representing a depletion of electron density. As an electron moves through the gas, it does not propagate as a free particle. Instead, as it moves it is followed by its xc hole and displaces other electrons. Such a bare electron surrounded by a “cloud” of displaced electrons is called a quasi-particle or “dressed” state. Due to screening, quasiparticles can be weakly interacting. Therefore, they represent the elementary low energy excitations of the electron gas. Reproducing these features of the momentum distribution is a stringent test for approximate functionals.

Figure 2 shows results [4] from the Müller functional and a closely related functional called BBC1 [5, 6] for the momentum distribution of the HEG. The momentum distribution is shown for $r_s = 1$ and $r_s = 5$. It is seen that

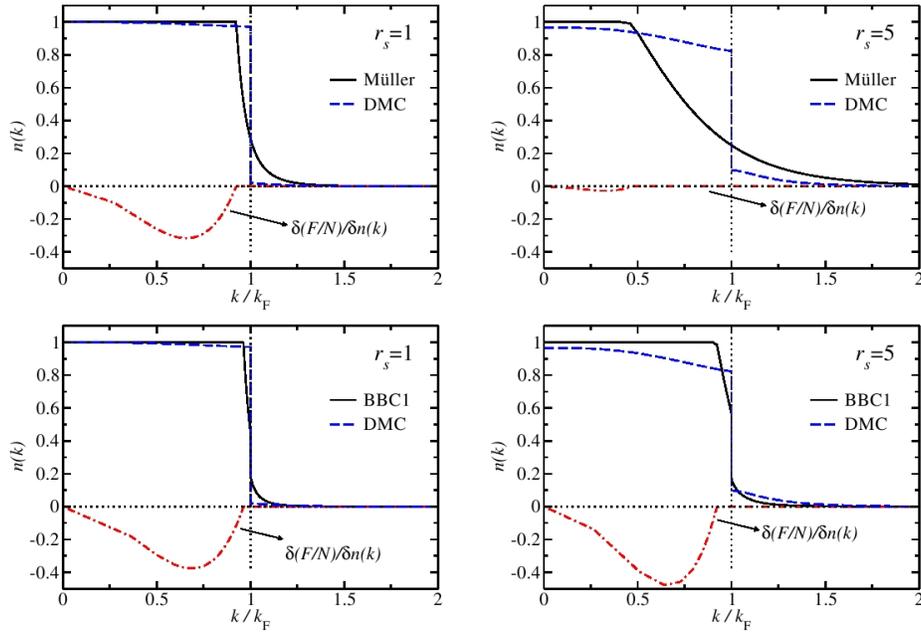


Figure 2: Approximations of the momentum distribution of the HEG. Adapted from Ref. [4]

the Müller functional overestimates the redistribution of occupancy over the Fermi level and does not reproduce the discontinuity at the Fermi level. The BBC1 functional describes better the redistribution of occupancy and does exhibit a finite discontinuity at the Fermi level.

Another quantity by which to test approximate functionals is the correlation energy of the HEG. The definition of the correlation energy is

$$E_{corr} = E - E_{HF}, \quad (21)$$

where E is the exact ground state energy and E_{HF} is the Hartree-Fock energy. In the HF approximation, the positions of parallel spin electrons are correlated due to the Pauli principle. The relative positions of antiparallel spin electrons are not correlated. The correlation energy represents the change to the HF energy upon correlation of antiparallel spin electrons and additional correlation of parallel spin electrons. The correlation energy of the HEG as predicted by various approximate functionals is shown in Figure 3. The Müller approximation overestimates the correlation energy. The

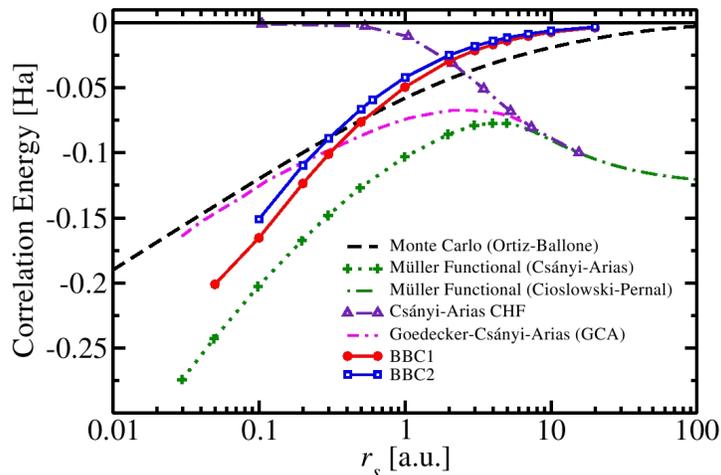


Figure 3: Correlation energy of the HEG as a function of r_s . Adapted from Ref. [4]

BBC1 approximation is fairly accurate for intermediate values of r_s . Other approximations are inaccurate in either the high density or low density limit.

5.2 Beryllium atom

The HEG is a special system because of its translation invariance. It is interesting to test approximate functionals also for finite systems. We will

examine the Beryllium atom, which has four electrons. Table 4 shows results of calculations of Beryllium using approximate functionals. Results for the

Method	Energy (a.u.)	% E_{corr}	$n(1s)$	$n(2s)$	$n(2p_z)$
RHF	-14.573 012	0.0	1.0000	1.0000	0.0000
GU	-14.653 435	85.2	1.0000	0.9574	0.0096
CH	-14.738 555	175.5	1.0000	0.7211	0.0807
CHF	-14.612 832	42.2	1.0000	0.6612	0.1077
CGA	-14.692 057	126.2	1.0000	0.6904	0.0928
HYB ^a	-14.675 633	108.8	1.0000	0.7002	0.0906
FCI	-14.644 757	76.0	0.9985	0.9061	0.0307
Exact ^b	-14.667 36	100.0	0.9982	0.9095	0.0295

Figure 4: Results for the total energy, correlation energy, and occupancies of spin-restricted natural orbitals of Beryllium. Adapted from Ref. [7]

Müller approximation are not available. However, the approximation labeled GU is the Müller functional with self-interaction terms omitted [8]. All rows represent approximate 1-matrix functionals, except those labeled RHF, FCI and exact. RHF is the restricted Hartree-Fock approximation, and FCI is full configuration interaction. In terms of the correlation energy, the GU approximation is comparable to FCI. Among the 1-matrix functionals, GU gives the most accurate values for occupation numbers.

5.3 Hydrogen Fluoride

The last example we consider is Hydrogen Fluoride, a strongly ionic molecule. Figure 5 shows the total energy as a function of the internuclear distance R . The abbreviations are the same as for the Beryllium results. The GU approximation is the most accurate 1-matrix functional over the depicted range of internuclear separation. The other approximations produce minima that are too shallow. The error is thought to be due to self-interaction. The restricted Hartree-Fock approximation, for which self-interaction is absent, gives an accurate value of the equilibrium internuclear distance.

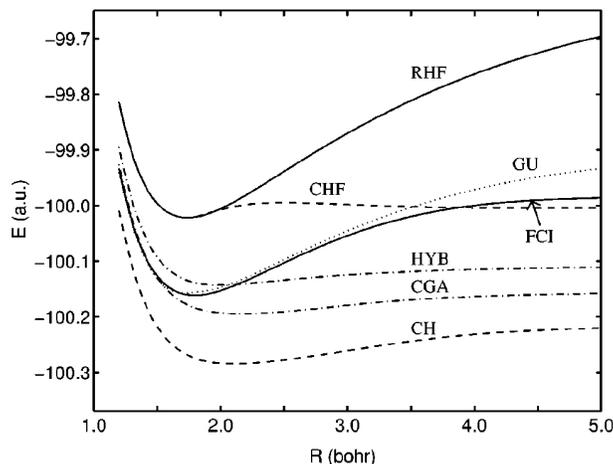


Figure 5: Energy of Hydrogen Fluoride as a function of internuclear distance. Adapted from Ref. [7]

6 Conclusions

Density matrix functional theory is an alternative approach to ground state calculations of the many-electron problem. We have summarized the formulation of the theory and compared it with density functional theory. The form of the approximations introduced in DMFT looks much different than the well-established approximations in DFT. One of the primitive approximations in DMFT is the Müller energy functional, which is an attempt to go beyond the Hartree-Fock approximation. We have reviewed some existing applications of the Müller energy functional and related approximate functionals. For reasons remarked in the Introduction, it may be hoped that further approximations will be found that are as accurate as those in DFT.

References

- [1] T. L. Gilbert, Phys. Rev. B **12**, 2111 (1975).
- [2] A. M. K. Müller, Phys. Lett. **105A**, 446 (1984).
- [3] P. Gori-Giorgi and P. Ziesche, PRB **66**, 235116 (2002).

- [4] N. N. Lathiotakis, N. Helbig, and E. K. U. Gross, cond-mat/0605531 (2006).
- [5] M. A. Buijse and E. J. Baerends, Mol. Phys. **100**, 401 (2002).
- [6] O. Gritsenko, K. Pernal, and E. J. Baerends, J. Chem. Phys. **122**, 204102 (2005).
- [7] V. N. Staroverov and G. E. Scuseria, J. Chem. Phys. **117**, 2489 (2002).
- [8] S. Goedecker and C. J. Umrigar, PRL **81**, 866 (1998).