

Reduced density matrix functional for many-electron systems

S. Sharma,^{1,2,3,*} J. K. Dewhurst,⁴ N. N. Lathiotakis,^{5,2} and E. K. U. Gross^{2,3}

¹Fritz Haber Institute of the Max Planck Society, Faradayweg 4-6, D-14195 Berlin, Germany

²Institut für Theoretische Physik, Freie Universität Berlin, Arnimallee 14, D-14195 Berlin, Germany

³European Theoretical Spectroscopy Facility (ETSF)

⁴School of Chemistry, The University of Edinburgh, Edinburgh EH9 3JJ, United Kingdom

⁵Theoretical and Physical Chemistry Institute, The National Hellenic Research Foundation, Vass. Constantinou 48, 11635 Athens, Greece

(Received 26 October 2008; published 24 November 2008)

Reduced density matrix functional theory for the case of solids is presented and an exchange-correlation functional based on a fractional power of the density matrix is introduced. We show that compared to other functionals, this produces more accurate behavior for total energies as a function of particle number for finite systems. Moreover, it captures the correct band-gap behavior for conventional semiconductors, as well as strongly correlated Mott insulators, where a gap is obtained in the absence of any magnetic ordering.

DOI: 10.1103/PhysRevB.78.201103

PACS number(s): 71.10.-w, 71.20.Nr, 71.27.+a, 71.45.Gm

One of the most dramatic failures of the usual local-density approximation (LDA) or generalized-gradient-type approximations to the exchange-correlation (xc) functional of density-functional theory (DFT) is the incorrect prediction of a metallic ground state for the strongly correlated Mott insulators, of which transition-metal oxides (TMOs) may be considered as prototypical. For some TMOs (NiO and MnO) spin-polarized calculations do show a very small band gap (up to 95% smaller than experiments) but only *as the result of antiferromagnetic (AFM) ordering*; however, all TMOs are found to be metallic in a spin-unpolarized treatment. On the other hand, it is well known experimentally that these materials are insulating in nature even at elevated temperatures (much above the Néel temperature).¹ This indicates that the magnetic order is not the driving mechanism for the gap and is just a co-occurring phenomenon. A real challenge for any kind of *ab initio* theory then is the prediction of an insulating state for these strongly correlated materials in the absence of magnetic order. Until now the main focus of reduced density matrix functional theory (RDMFT) has been on finite systems such as atoms and molecules^{2–11} with various xc functionals,^{3–8,11–15} which are essentially modifications of the original Müller functional.² In the present work we extend RDMFT to the case of solid-state systems and introduce a functional which generates not only accurate gaps for conventional semiconductors, but demonstrates insulating behavior for Mott-type insulators in the nonmagnetic phase.

Formally, the one-body reduced density matrix γ for a pure state of N electrons is defined as (spin degrees of freedom are omitted for simplicity)

$$\gamma(\mathbf{r}, \mathbf{r}') = N \int \Psi(\mathbf{r}, \mathbf{r}_2, \dots, \mathbf{r}_N) \Psi^*(\mathbf{r}', \mathbf{r}_2, \dots, \mathbf{r}_N) d^3 r_2 \dots d^3 r_N. \quad (1)$$

Diagonalization of this matrix produces a set of natural orbitals,¹⁶ ϕ_i , and occupation numbers, n_i , leading to the spectral representation

$$\gamma(\mathbf{r}, \mathbf{r}') = \sum_i n_i \phi_i(\mathbf{r}) \phi_i^*(\mathbf{r}'), \quad (2)$$

where the necessary and sufficient conditions for ensemble N representability¹⁷ require $0 \leq n_i \leq 1$ for all i , and $\sum_i n_i = N$. In terms of γ , the total ground-state energy of the interacting system is¹⁸ (atomic units are used throughout)

$$E_V[\gamma] = \frac{-1}{2} \int \lim_{\mathbf{r} \rightarrow \mathbf{r}'} \nabla_{\mathbf{r}}^2 \gamma(\mathbf{r}, \mathbf{r}') d^3 r' + \int \rho(\mathbf{r}) V(\mathbf{r}) d^3 r + \frac{1}{2} \int \frac{\rho(\mathbf{r}) \rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3 r d^3 r' + E_{xc}[\gamma], \quad (3)$$

where $\rho(\mathbf{r}) = \gamma(\mathbf{r}, \mathbf{r})$, V is a given external potential, and E_{xc} is what we call the xc energy functional. Minimizing the total energy with $E_{xc}[\gamma] = -\frac{1}{2} \int |\gamma(\mathbf{r}, \mathbf{r}')|^2 / |\mathbf{r} - \mathbf{r}'| d^3 r d^3 r'$ is equivalent to the Hartree-Fock (HF) method. The HF functional satisfies the exact condition of the xc hole integrating to minus one; however it does not satisfy the condition of convexity, which is required by the exact functional.¹⁹ Müller proposed² a simple alternative to the HF functional in which $|\gamma(\mathbf{r}, \mathbf{r}')|^2$ is replaced by $[\gamma^\alpha(\mathbf{r}, \mathbf{r}')]^* \gamma^{1-\alpha}(\mathbf{r}, \mathbf{r}')$, where γ^α indicates the power used in the operator sense, i.e.,

$$\gamma^\alpha(\mathbf{r}, \mathbf{r}') = \sum_i n_i^\alpha \phi_i(\mathbf{r}) \phi_i^*(\mathbf{r}'). \quad (4)$$

Müller's functional satisfies both conditions²⁰ for all p with $0 < p < 1$. All studies of this functional to date have set $p = 1/2$ (in the rest of the paper we refer to this as the Müller functional). As is well known, however, this functional severely overestimates electron correlation,^{5,6,12,14} and we find that changing the value of p away from 1/2 only exacerbates this problem. Therefore in the present work we instead replace $|\gamma(\mathbf{r}, \mathbf{r}')|^2$ by $|\gamma^\alpha(\mathbf{r}, \mathbf{r}')|^2$, where $1/2 \leq \alpha \leq 1$. This simple functional form interpolates between the uncorrelated HF limit ($\alpha=1$) and the overcorrelating Müller functional ($\alpha=1/2$). Remarkably, we find that an intermediate value of α exists for which nonmagnetic TMOs are insulators, as well as accurate gaps found for a diverse set of semiconductors

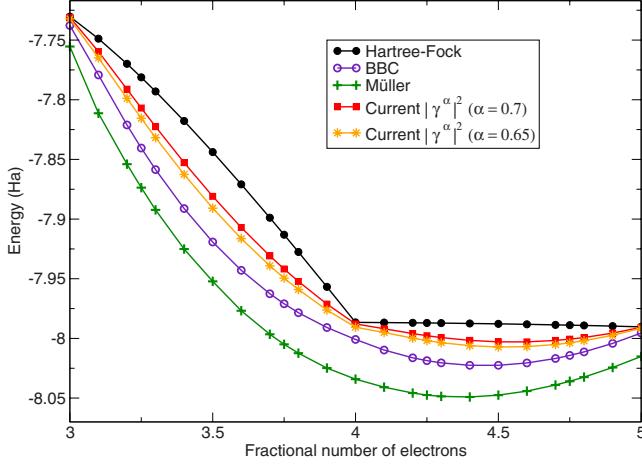


FIG. 1. (Color online) Energy for LiH molecule as a function of fractional number of excess electrons. Results are obtained using various approximations to the xc functional.

and insulators. The price to pay for this improved treatment of correlation is that this functional form fails to satisfy the xc hole condition exactly. However, all numerical tests show that it is at least convex, and in fact, as shown in Fig. 1 for the case of finite systems (Fig. 1 is a plot for a representative finite system, the LiH molecule), the variation of energy as a function of charge obtained using current functional (for $\alpha = 0.65, 0.7$) is much closer to the required straight-line behavior¹⁹ than the Müller functional or the so-called BBC functionals.⁵ This type of linear behavior has also been recently demonstrated by Cohen *et al.*²¹ for their density functional.

In solids because of the underlying lattice periodicity of the external potential, the one-body reduced density matrix for crystals has the symmetry $\gamma(\mathbf{r}+\mathbf{T}, \mathbf{r}'+\mathbf{T})=\gamma(\mathbf{r}, \mathbf{r}')$, where \mathbf{T} is a primitive translation vector, and thus the natural orbitals are also Bloch states. Since there are no Kohn-Sham-like equations to solve, a direct minimization over these natural orbitals and occupation numbers is required while maintaining the ensemble N -representability conditions. In practical terms, the natural orbitals are expanded in terms of a set of previously converged Kohn-Sham states, and optimization of the natural orbitals is performed by varying the expansion coefficients. We should stress that the use of Kohn-Sham states as a basis is merely a numerical convenience, and results do not depend upon the starting point in any way. All calculations are performed using the state-of-the-art full-potential linearized augmented plane-wave (FP-LAPW) method,²² implemented within the EXCITING code.²³

For finite systems the chemical potential is defined as $\mu(N)=dE(N)/dN$, where $E(N)$ is the ground-state energy of N electrons moving in a fixed external potential. For the exact functional, $\mu(N)$ consists of horizontal lines with possible steps at integer values of N .¹⁹ If N_0 is the electron number for which the total system is charge neutral, then the step at N_0 given by $E_g=\lim_{\eta \rightarrow 0}[\mu(N_0+\eta)-\mu(N_0-\eta)]$ represents the band gap or twice the chemical hardness. For periodic solids both E and N are infinite. Moreover, adding a finite charge η to each unit cell while keeping the external potential fixed would lead to an infinitely charged system

which is not stable. The only accessible quantity is the energy per unit volume, \tilde{E} , as a function of excess charge η per unit cell, where $\tilde{E}_{V+\delta v}(\eta)$ is the self-consistent total energy evaluated with an external potential (V) plus the Coulomb potential of a constant charge background (δv), which makes the total system charge neutral. Clearly, the quantity \tilde{E} and the chemical potential $\tilde{\mu}(\eta)=d\tilde{E}_{V+\delta v}(\eta)/d\eta$ are conceptually quite different from the corresponding quantities $E(\eta)$ and $\mu(\eta)$ of a finite system because for infinite systems the external potential is not kept fixed as η is changed. For the case of small excess charge η , the quantity \tilde{E} can be written, to second order in η , as

$$\tilde{E}_{V+\delta v}(\eta)=\tilde{E}_V(\eta)+\int \delta v(\mathbf{r})\tilde{\rho}_{N^\pm\eta}(\mathbf{r})d^3r. \quad (5)$$

The Coulomb potential from the constant charge (η) background is given by

$$\delta v(\mathbf{r})=\eta \int \frac{1}{|\mathbf{r}-\mathbf{r}'|}d^3r', \quad (6)$$

and $\tilde{\rho}_M(\mathbf{r})$ —the self-consistent M -electron density evaluated with lattice potential plus neutralizing background potential—can be expressed as

$$\tilde{\rho}_{N^\pm\eta}(\mathbf{r})=\tilde{\rho}_N(\mathbf{r})\pm\eta n^\pm(\mathbf{r}) \quad (7)$$

with $\int n^\pm(\mathbf{r})d^3r=1$. Combining Eqs. (5)–(7) we arrive at

$$\tilde{E}_{V+\delta v}(\eta)=\tilde{E}_V(\eta)\pm\eta \int \frac{\tilde{\rho}_N(\mathbf{r})\pm\eta n^\pm(\mathbf{r})}{|\mathbf{r}-\mathbf{r}'|}d^3rd^3r'. \quad (8)$$

Derivative of Eq. (8) with respect to η yields

$$\tilde{\mu}(\eta)=\mu(\eta=0^-)+\begin{cases} c_1\eta & \text{for } \eta<0 \\ E_g+c_r\eta & \text{for } \eta>0 \end{cases} \quad (9)$$

with $c_1=2\int n^-(\mathbf{r})/|\mathbf{r}-\mathbf{r}'|d^3rd^3r'$ and $c_r=2\int n^+(\mathbf{r})/|\mathbf{r}-\mathbf{r}'|d^3rd^3r'$. From Eq. (9) we can see that introduction of the background charge has no effect on the size of the band gap. Another property that is clear from this equation is that unlike for the case of finite systems $\tilde{\mu}$ does not consist of horizontal lines on either side of $\eta=0$. Although Eq. (9) is only valid for small values of η , in practice we find a linear behavior for $\tilde{\mu}(\eta)$ even for larger values of η , a fact which can be attributed to the metallic nature that the system acquires on addition or removal of charge.

In Fig. 2 we show the plot of the chemical potential versus excess electronic charge for a prototype metal (Cu), semiconductor (Si), and Mott insulator (NiO) obtained using the current functional with this value of $\alpha=0.65$. For Cu one observes a nearly linear behavior with a small negative curvature. In contrast, Si and NiO show a qualitatively different behavior in which the curvature of the chemical potential changes sign. We interpret this as the appearance of a smoothed discontinuity with a linear behavior to the left and to the right. This is strikingly different from that of Cu, which remains metallic. This smoothing of discontinuities, a consequence of approximating the xc functional, has already been noted in work on finite systems.¹⁰ Nevertheless, owing

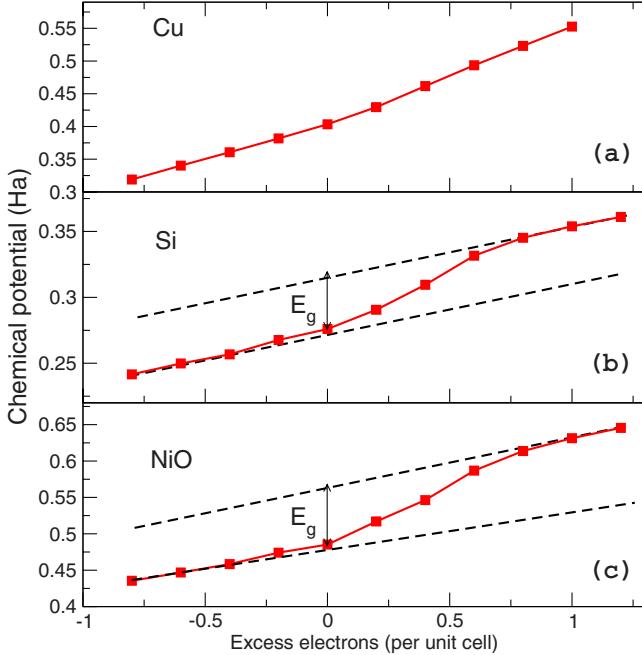


FIG. 2. (Color online) Chemical potential (in Ha) versus the excess charge per unit cell for (a) Cu, (b) Si, and (c) NiO. The results are obtained using the current functional with $\alpha=0.65$.

to the near linearity of the chemical potential on either side of zero excess charge [see Eq. (9)], the fundamental gap may be estimated¹⁰ by construction of two tangents as shown in Fig. 2. This allows a rigorous test of the functional by comparison with experiment of the fundamental gap of a wide class of materials: semiconductors, insulators, and strongly correlated systems. We emphasize that for the Müller and the BBC functionals, without self-interaction correction, all these systems turn out to be metallic.

We first consider semiconductors and ionic insulators, for which results are shown in Fig. 3, again using the current functional with $\alpha=0.65$. The agreement between the calculated gap and experiments is good, with an average deviation in the calculated gap of 9.5%, comparable to the *GW* method where if we choose the best (worst) result from the literature for each material we find an average deviation of 3% (9%). It must be stressed that this good agreement of band gaps is found for a wide range of semiconductors/insulators, with gaps ranging from 1 eV (Ge) to 14.2 eV (LiF) and the character of the material changing from predominantly covalent to ionic. We find that on increasing the value of α from 0.65 to 0.7 the agreement with experiment improves somewhat, now with an average deviation of 7.5%. In most cases gaps are overestimated with $\alpha=0.7$ and underestimated with $\alpha=0.65$. Reducing the value of α below 0.65 has deleterious consequences; already for $\alpha=0.6$ several of the semiconductors are metallic. Therefore, there exists a small range of α (0.65–0.7) for which the calculated gaps agree well with experiment.

The TMOs have fundamental gaps ranging from 4.2 eV (NiO) to 2.4 eV (CoO); however, the origin of the gap is profoundly different from conventional semiconductors/insulators. Here the gap is opened by strong Mott-Hubbard

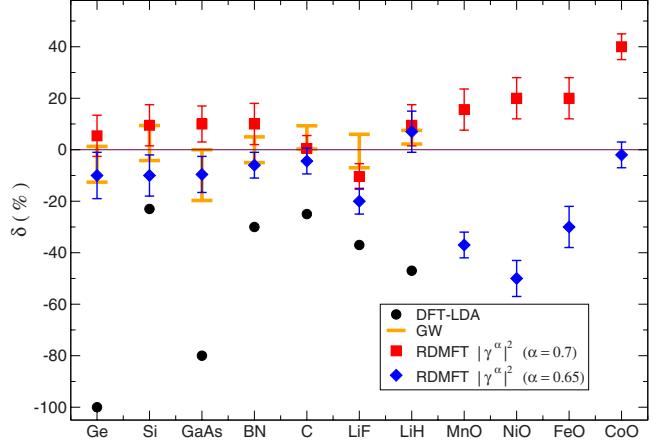


FIG. 3. (Color online) Percentage deviation (δ) of calculated band gap from experiment. DFT results (circles) are with the LDA. RDMFT values (blue/dark gray and red/gray) are obtained using the current functional with $\alpha=0.65, 0.7$. The *GW* data (orange/light gray vertical lines) are taken from Ref. 24; a line is drawn between the smallest and the largest value.

correlations and, toward the end of the 3d series, charge-transfer effects also contribute.¹ Gap formation by strong correlations is largely unaffected by the magnetic state, and so to focus exclusively on Mott-Hubbard physics we calculate these materials in the nonmagnetic state. As can be seen in Fig. 3 the current functional (for both $\alpha=0.65, 0.7$) finds TMOs to be insulators. In this case the sensitivity of the gap to α is significantly greater; however, as before, the best value of α is still in the range 0.65–0.7. We should stress that the calculation of gaps by the aforementioned tangent construction comes with, of course, an inherent inaccuracy illustrated by the error bars²⁵ in Fig. 3.

The traditional quasiparticle technique used to treat TMOs, the *GW* method, has only been applied to NiO and MnO in the presence of AFM ordering.²⁶ The reason for this lies in the fact that all DFT calculations to date lead to metallic FeO and CoO, and for the *GW* method to produce a gap starting from this metallic ground state would require a diverging self-energy. In contrast, RDMFT is a fully *ab initio* nonperturbative theory, which with the current functional and a fixed value of α , seems to capture not only the physics of conventional semiconductors/insulators, but also that of strongly correlated materials. What still remains is the method for choosing α . It is clear from the present work that the optimal value of α lies in a small range (0.65–0.70). In order to refine α any further one could do one of two things: make α a functional of density or fit α to best reproduce results for a wide range of materials.

It is clear that the current functional improves the value of the band gaps; however, it is also important that it performs well for properties already adequately described by DFT within LDA. An example of such a property is the equilibrium lattice constant. In Table I the values for Si, cubic BN, and diamond are presented. The results for the equilibrium lattice parameter obtained with RDMFT, using both Müller and current functional ($\alpha=0.7$), are as good as the values obtained using DFT. For the current functional with α

TABLE I. Equilibrium lattice parameter (in a.u.). The DFT values are obtained using the LDA. RDMFT results are obtained with the Müller and current functional with $\alpha=0.7$. In the last row the average percentage deviation of calculated results from experiments is given.

Solid	Expt.	DFT-LDA	Müller	Current
Diamond	6.74	6.68	6.78	6.75
Si	10.26	10.188	10.49	10.55
BN	6.83	6.758	6.86	6.82
		1.03	1.07	1

=0.65 the average percentage deviation of calculated results from experiments is 1.2%.

In conclusion, an extension of RDMFT to solids is presented with introduction of a functional. The values of the fundamental gap for various semiconductors and ionic insu-

lators are dramatically improved from DFT-LDA values and are in as good agreement with experiment as GW values. Furthermore, strong Mott-Hubbard correlation is captured in this approach, as calculations of the nonmagnetic TMOs show. The Müller functional appears to be unsuitable for use with solids, as it fails to produce a gap even for conventional semiconductors.

We have shown that RDMFT is a viable theory for the study of solids where many-body effects are important, and where DFT-based theories have notably failed. This opens up many future possibilities, such as the study of high- T_C superconductors in their underdoped Mott insulating phase. It should also, hopefully, stimulate efforts to develop the theory formally, including a temperature-dependent extension and a method to produce quasiparticle spectra.

We acknowledge Deutsche Forschungsgemeinschaft and the Network of Excellence NANOQUANTA (Grant No. NMP4-CT-2004-50019) for financial support.

*sangeeta.sharma@physik.fu-berlin.de

- ¹O. Tjernberg, S. Soderholm, G. Chiaia, R. Girard, U. O. Karlsson, H. Nylen, and I. Lindau, Phys. Rev. B **54**, 10245 (1996); W. Jauch and M. Reehuis, *ibid.* **70**, 195121 (2004).
- ²A. M. K. Müller, Phys. Lett. **105A**, 446 (1984).
- ³S. Goedecker and C. J. Umrigar, Phys. Rev. Lett. **81**, 866 (1998).
- ⁴M. A. Buijse and E. J. Baerends, Mol. Phys. **100**, 401 (2002).
- ⁵O. Gritsenko, K. Pernal, and E. J. Baerends, J. Chem. Phys. **122**, 204102 (2005).
- ⁶J. M. Herbert and J. E. Harriman, Chem. Phys. Lett. **382**, 142 (2003).
- ⁷G. Csányi, S. Goedecker, and T. A. Arias, Phys. Rev. A **65**, 032510 (2002).
- ⁸C. Kollmar, J. Chem. Phys. **121**, 11581 (2004).
- ⁹N. N. Lathiotakis, N. Helbig, and E. K. U. Gross, Phys. Rev. A **72**, 030501(R) (2005).
- ¹⁰N. Helbig, N. N. Lathiotakis, M. Albrecht, and E. K. U. Gross, Europhys. Lett. **77**, 67003 (2007).
- ¹¹P. Leiva and M. Piris, J. Chem. Phys. **123**, 214102 (2005).
- ¹²G. Csányi and T. A. Arias, Phys. Rev. B **61**, 7348 (2000).
- ¹³K. Pernal and J. Cioslowski, J. Chem. Phys. **120**, 5987 (2004).
- ¹⁴N. N. Lathiotakis, N. Helbig, and E. K. U. Gross, Phys. Rev. B **75**, 195120 (2007).
- ¹⁵M. Piris, Int. J. Quantum Chem. **106**, 1093 (2006).
- ¹⁶P.-O. Löwdin, Phys. Rev. **97**, 1474 (1955).

¹⁷A. Coleman, Rev. Mod. Phys. **35**, 668 (1963). This proof can be easily extended to the case of fractional number of electrons, Sharma *et al.* (unpublished).

¹⁸T. L. Gilbert, Phys. Rev. B **12**, 2111 (1975).

¹⁹J. P. Perdew, R. G. Parr, M. Levy, and J. L. Balduz, Phys. Rev. Lett. **49**, 1691 (1982).

²⁰R. L. Frank, E. Lieb, R. Seiringer, and H. Siedentop, Phys. Rev. A **76**, 052517 (2007).

²¹A. J. Cohen, P. Mori-Sánchez, and W. Yang, Phys. Rev. B **77**, 115123 (2008).

²²D. J. Singh, *Planewaves, Pseudopotentials and the LAPW Method* (Kluwer Academic, Boston, 1994).

²³J. K. Dewhurst, S. Sharma, and C. Ambrosch-Draxl (<http://exciting.sourceforge.net>).

²⁴S. Massidda, A. Continenza, M. Posternak, and A. Baldereschi, Phys. Rev. B **55**, 13494 (1997); M. Rohlfing, P. Kruger, and J. Pollmann, *ibid.* **48**, 17791 (1993); Eric L. Shirley, *ibid.* **58**, 9579 (1998); M. Shishkin and G. Kresse, *ibid.* **75**, 235102 (2007); S. Lebegue, M. Alouani, B. Arnaud, and W. E. Pickett, Europhys. Lett. **63**, 562 (2003).

²⁵Cubic splines are fitted to the chemical-potential dataset. The derivative of this spline allows the construction of a continuous range of possible left and right tangents from which the range of possible gaps may be obtained.

²⁶S. V. Faleev, M. van Schilfgaarde, and T. Kotani, Phys. Rev. Lett. **93**, 126406 (2004).