Self-Interaction and Strong Correlation in DFTB[†]

B. Hourahine,*,[‡] S. Sanna,[§,#] B. Aradi,[#] C. Köhler,[#] Th. Niehaus,[#] and Th. Frauenheim[#]

Department of Physics, University of Strathclyde, John Anderson Building, 107 Rottenrow, Glasgow G4 0NG, United Kingdom, Theoretische Physik, Universität Paderborn, Warburger Strasse 100, D-33098 Paderborn, Germany, BCCMS, Universität Bremen, Otto-Hahn-Allee 1, 28359 Bremen, Germany

Received: January 9, 2007; In Final Form: May 22, 2007

The density functional based tight-binding (DFTB) method can benefit substantially from a number of developments in density functional theory (DFT) while also providing a simple analytical proving ground for new extensions. This contribution begins by demonstrating the variational nature of charge-self-consistent DFTB (SCC-DFTB), proving the presence of a defined ground-state in this class of methods. Because the ground state of the SCC-DFTB method itself can be qualitatively incorrect for some systems, suitable forms of the recent LDA+U functionals for SCC-DFTB are also presented. This leads to both a new semilocal self-interaction correction scheme and a new physical argument for the choice of parameters in the LDA+U method. The locality of these corrections can only partly repair the HOMO−LUMO gap and chemical potential discontinuity, hence a novel method for introducing this further physics into the method is also presented, leading to exact derivative discontinuities in this theory at low computational cost. The prototypical system NiO is used as an illustration for these developments.

I. Introduction

As is discussed in several of the contributions to this special issue, the DFTB method is intended to give a good approximation to the results of density-functional theory (DFT). Examples of its application to the solid state,[1−5] and molecular systems[6] demonstrate the versatility of the method both in its original[7] and in self-consistent charge (SCC-DFTB)[8] forms, with other contributions in this special volume presenting details of many of the recent applications and extensions to these methods. However, there are several regions of the periodic table that require particular care when investigating with a DFT-related approach (the d and f blocks being prime examples) at least in part from the spurious self-interaction present in many density functionals. Apparently innocuous systems such as negatively charged ions or open systems in contact with electron reservoirs also present challenges, again related to the self-interaction problems of DFT.

This paper is organized as follows. In section II a short proof is given that tight-binding computational schemes such as SCC-DFTB, where terms quadratic in the charge fluctuations are included in the energy, can actually possess a ground state as long as the on-site coupling terms are of greater of equal magnitude to the off site contributions. This is explicitly demonstrated for the softened Coulombic interaction of the Klopman−Ohno−Mataga−Nishimoto[9−11] model, but also holds for the more elaborate softened Coulombic interaction currently used in SCC-DFTB.

Having established that SCC-DFTB actually possesses a ground state, in section III, we present a new extension of the SCC-DFTB method to at least qualitatively treat strongly correlated systems. This leads to a new method for removing much of the local part of self-interaction, as well as presenting a new physical recipe for the choice of constants in the LDA+U method. The equivalence between a form of the pseudo-self-interaction correction[12,13] method and LDA+U is also demonstrated for the first time.

As discussed recently[14,22] some care has to be taken when the self-interaction issue is considered, as it can manifest in several ways. As discussed in section IV, even after the localized states in NiO in SCC-DFTB are corrected, an error of around half of the experimental band gap still has to be accounted for by including the derivative discontinuity. A simple method to include this contribution both to the band structure and to its effects on the energy for fractionally charged systems is then presented.

II. Presence of a Ground State in SCC-DFTB

Foulkes and Haydock[15] considered the connection between tight-binding and Kohn−Sham DFT. If in a DFT, V_{in} makes n_{out} the ground-state charge density, then as given in eq 25a of ref 15:

$$\epsilon[n_{\text{in}}]^{\text{TB}} = E[n_{\text{out}}]^{\text{DFT}} - \frac{1}{2}\int\int \frac{\delta^2 F}{\delta n^2}\bigg|_{n_{\text{in}}} \delta n\, \delta n \tag{1}$$

where, as explained in detail in the reference, the energy functionals $\epsilon[n_{\text{in}}]^{\text{TB}}$ and $E[n_{\text{out}}]^{\text{DFT}}$ depend on the input and output densities of the self-consistent process, respectively. The charge fluctuations, δn, are given as the difference between the output and input densities, $\delta n = n_{\text{out}} - n_{\text{in}}$ and abrieviated as in ref 15:

$$\int\int \frac{\delta^2 F}{\delta n^2}\bigg|_{n_{\text{in}}} \delta n\, \delta n = \int\int \frac{\delta^2 F}{\delta n(\mathbf{r})\delta n(\mathbf{r}')}\bigg|_{n_{\text{in}}} \delta n(\mathbf{r})\, \delta n(\mathbf{r}')\, d^3\mathbf{r}\, d^3\mathbf{r}' \tag{2}$$

[†] Part of the “DFTB Special Section”.

[‡] University of Strathclyde.

[§] Universität Paderborn.

[#] Universität Bremen.

10.1021/jp070173b CCC: $37.00

Foulkes and Haydock concluded that empirical tight-binding is a stationary input-only DFT. Non-SCC-DFTB[4,7] has the same stationary behavior as $\epsilon[n_{\mathrm{in}}]^{\mathrm{TB}}$, but what about self-consistent charge DFTB?[8] From the DFTB choice of expansion reference charge distribution, n_{atomic}, which becomes the input charge distribution in the final calculation, then

$$E[n_{\mathrm{out}}]^{\mathrm{SCC-DFTB}} = \epsilon[n_{\mathrm{in}}]^{\mathrm{DFTB}} + \frac{1}{2}\sum_{ab}^{\mathrm{atoms/orbitals}} \gamma_{ab}\Delta n_{\mathrm{out}}^{a}\Delta n_{\mathrm{out}}^{b} \quad (3)$$

with the intention that γ is a discretized approximation for

$$\gamma \approx \int\int \frac{\delta^2 F}{\delta n^2}\bigg|_{n_{\mathrm{atomic}}} \quad (4)$$

with a point-like model for the charge fluctuations

$$\Delta n_{\mathrm{out}}^{a} = n_{\mathrm{out}}^{a} - n_{\mathrm{atomic}}^{a} \quad (5)$$

instead of the field used in eq 1. In this expression, a is either an atom or a unique l-shell index depending on whether the model is atom or atom-shell resolved (both forms of DFTB being in use for different applications). By comparison with eq 1 this expression will also be bounded in energy from below (though not necessarily by the DFT ground state) if the condition $\Delta q_a \gamma_{ab} \Delta q_b \leqslant \Delta {q_a}^2 \gamma_{aa} + \Delta {q_b}^2 \gamma_{bb}$ is met. This holds as long as $\gamma(r)$ decays monotonically with increasing r, and $\gamma_{ab}(r=0) \leqslant \gamma_{aa}(r=0) + \gamma_{bb}(r=0)$. The (somewhat complex) current functional form of γ,[8] by construction, decays as a function of separation and also obeys the inequality for positive Hubbard U values. Similarly, the Klopman−Ohno−Mataga−Nishimoto[9−11] type of softened Coulombic interaction, of the general form

$$\gamma_{ab}(r) = \frac{1}{\sqrt{{r_{ab}}^2 + \dfrac{c}{(\gamma_{aa} + \gamma_{bb})^2}}} \quad (6)$$

also is bounded, as it decays monotonically with r, and at $r = 0$ also meets the requirement of the inequality as long as $c \geqslant 1$, as $\gamma_{ab}(r = 0) = (\gamma_{aa} + \gamma_{bb})/c$.

III. Semilocal Orbitally Dependent Functionals

Full self-interaction[17] correction or GW[18] methods can give a good description of some systems for which LDA/GGA(/SCC-DFTB) fail, alternatively hybrid functionals either using a contribution from raw Hartree−Fock[19] or including long-range screening have been applied. These methods are relatively expensive, but there are cheaper approximations to them. LDA+U[20] adds contributions from the Hubbard model to mean-field functional (this is a form of GW[20]). Alternatively, Pseudo-SIC removes the local part of self-interaction using projected atomic states.[12] Many of these method have recently been compared for the volume driven magnetic phase transition of MnO,[21] but we are not aware of a systematic comparison of these methods for other correlated oxides, such as NiO. Several of these methods, in particular the hybrid functionals, while agreeing well with experiment, usually ignore the contribution to the band gap from the derivative discontinuity.[22]

We now briefly describe the LDA+U and pseudo-SIC methods to set the context for their implementation within the semiempirical framework of SCC-DFTB. As discussed by Anisimov et al.,[20] it is natural to separate electrons into localized d or f electrons and delocalized s and p electrons. Although for the latter an orbitally independent one-electron potential (as in LDA) will suffice, a Hartree−Fock-like interaction better describes the local interactions. This is of the form $(1/2)\sum_{i\neq j} n_i n_j$, where n_i are the occupancies of the localized shells. If we assume that the Coulomb energy of the electron−electron interaction as a function of the total number of electrons $N = n_i$ is well represented by LDA (even if it gives wrong single-particle energies), then LDA already contains part of this energy. This must be subtracted from the total energy and instead replaced with a Hubbard-model-like term. As a result we get the functional[23,24]

$$E = E_{\mathrm{LDA}} - \frac{1}{2}UN(N-1) + \frac{U}{2}\sum_{i\neq j} n_i n_j \quad (7)$$

Strictly speaking, the process of subtracting the double-counting of the electron−electron interaction of strongly correlated electrons from the LDA total energy and substituting it with a Hubbard Hamiltonian-like term is not without ambiguity.

One prescription for an LDA+U can be constructed by considering uniform occupation of the correlated states, around a mean-field limit (AMF).[23] For strongly correlated systems (or in the presence of a crystal/ligand field) the limit of uniform occupancy is not correct. This has led to the suggestion of another correction in the fully localized limit (FLL).[23,24] AMF and FLL correct the mean field double-counting if the local occupation numbers are respectively all equal or alternatively only 0 or 1. Most of the modern LDA+U calculations rely on one of these two functionals, although in real materials the occupation numbers should lie between these two limits. Hence neither AMF nor FLL is strictly speaking correct for real systems; one should therefore use an interpolation between the two limits.[16] However, AMF and FLL will bracket the correct limit.

After the initial submission of this manuscript, a third form of LDA+U was proposed;[26] this is functionally similar to the pseudo-SIC correction discussed in section 3.2, though derived in a different manner.

A. LDA+U-Like Approach in SCC-DFTB. Although it has previously been suggested that for *empirical* tight-binding the effects of on-site correlation can be mimicked by an empirical adjustment of symmetry resolved on-site energies,[27] this is problematic for example for low symmetry d electron systems, or for f manifolds. In these cases, a different correction would apply, in principle, to each l and m quantum numbered atomic state, breaking the rotational symmetry of the method. Similar problems in maintaining rotational invariance occur on incorporation of self-interaction corrected or hybrid functionals into the parametrization for the non-self-consistent tight-binding Hamiltonian.

The rotationally invariant form of LDA+U[25] bypasses these problems, with a correction to the LDA potential is of the form

$$\Delta V^{\mathrm{LDA+U}}_{\mu\nu\sigma\,\mu\in l,\nu\in l} = -(U_l - J_l)(n^{\sigma}_{\mu\nu} - \mathrm{DC}[n^{\sigma}_{\mu\nu}]) \quad (8)$$

where n^{σ} is the local spin occupation matrix within a given atomic manifold, and $(U_l - J_l)$ is the screened and spherically averaged electron−electron interaction. DC[n] is the double-counting term, and the two limiting cases FLL and AMF are

$$\mathrm{DC}^{\sigma}_{\mu\nu}[n]^{\mathrm{FLL}} = \frac{1}{2}\delta_{\mu\nu} \quad (9)$$

$$\mathrm{DC}^{\sigma}_{\mu\nu}[n]^{\mathrm{AMF}} = \frac{\mathrm{Tr}(n_{\mu\nu})}{2l+1}\delta_{\mu\nu} \quad (10)$$

where l is the angular quantum number.

$(U_l - J_l)$ is usually taken to be either an adjustable parameter or the result of a constrained DFT calculation.[28] We instead present a new prescription for choosing U and J from atomic calculations on the basis of a comparison between LDA+U and a new form of the pseudo-SIC method. Because the SCC-DFTB energy aims to be a reasonable approximation to the LDA energy, it seems sensible to adopt the form of the LDA+U energy correction unchanged for SCC-DFTB.

There is then the issue of how to choose the on-site occupation matrix[29] for a nonorthogonal basis. We present results for both the on-site[30] and the dual basis[29] forms of LDA+U. In the on-site case the modification to the total energy and the SCC-DFTB Hamiltonian can be written in as functions of atomic sub-blocks of the single-particle density matrix (ρ), whereas for the dual basis case, the occupation matrix takes the form of a generalization to the Mulliken charges:

$$n^{\sigma}_{\mu\in A,\nu\in A} = \frac{1}{2}\sum_{B}\sum_{\tau\in B}(S_{\mu\tau}\rho_{\tau\nu} + \rho_{\mu\tau}S_{\tau\nu}) \tag{11}$$

where the diagonal of the occupation matrix is then basis-function-resolved Mulliken charges.

There is a third alternative, the "full" form of the populations,[29] which has some similarity to the projectors described for correcting for derivative discontinuities in section IV. However, all of the occupation matrices presented for LDA+U are atomic shell block-diagonal, unlike the full density matrix used in section IV.

B. Pseudo-SIC. Full self-interaction corrected LDA is relatively expensive; hence several cheaper approximations have appeared. The pseudo-SIC (pSIC) scheme[12] and its recent refinements[13,31] project Kohn−Sham states onto a localized (often confined) atomic basis. Because the majority of the self-interaction error is local in character this hopefully captures the majority of the error. In the later variants of the method, local orbital occupation numbers for that projected basis, p_i, are constructed for each state (i.e., $p_i = 1$ for the original method[12]). The exchange−correlation and Hartree potentials are then modified by subtracting off the local self-interaction in this basis:

$$V^{\sigma}_{\text{xc\&Hartree}}[n,m] \rightarrow V^{\sigma}_{\text{xc\&Hartree}}[n,m] - \sum_i p^{\sigma}_i V^{\sigma}_{\text{xc\&Hartree}}[n=1,m=1] \tag{12}$$

where $V_{\text{xc\&Hartree}}$ is calculated for a fully occupied ($n = 1$) and spin-polarized ($m = 1$) state. Additionally, the potential can be scaled, as discussed in some detail in ref 13,

$$V^{\sigma}_{\text{xc\&Hartree}}[n,m] \rightarrow V^{\sigma}_{\text{xc\&Hartree}}[n,m] - \alpha\sum_i p^{\sigma}_i V^{\sigma}_{\text{xc\&Hartree}}[n=1,m=1] \tag{13}$$

for example, to account for electronic relaxation on electron removal ($\alpha = 1/2$ in the work of Filippetti and Spaldin; however, this is incorrect for a system with a single electron[31]). As yet, no energy expression related by variational principle to the potential is available;[13,31] hence no expression for interatomic forces has been derived.

Interestingly, summing over the occupied states in the projected basis gives the on-diagonal parts of an occupation matrix

$$V^{\sigma}_{\text{xc\&Hartree}}[n,m] \rightarrow V^{\sigma}_{\text{xc\&Hartree}}[n,m] - \alpha\sum_i p^{\sigma}_i V^{\sigma}_{\text{xc\&Hartree}}[n=1,m=1] \tag{14}$$

$$= V^{\sigma}_{\text{xc\&Hartree}}[n,m] - \alpha V^{\sigma}_{\text{xc\&Hartree}}[n=1,m=1]\delta_{\mu\nu}n^{\sigma}_{\mu\nu} \tag{15}$$

This is not invariant to unitary transforms; however, this was rectified for the original LDA+U formulation[32] by including off-diagonal elements for each atom block.

$$V^{\sigma}_{\text{xc\&Hartree}}[n,m] \rightarrow V^{\sigma}_{\text{xc\&Hartree}}[n,m] - \alpha V^{\sigma}_{\text{xc\&Hartree}}[n=1,m=1]n^{\sigma}_{\mu\nu} \tag{16}$$

Due to the similarity in the functional of the potential to LDA+U, we can write an energy expression (which *is* directly connected to the potential) as

$$\Delta E_{\text{pSIC}} = -\alpha\sum_a\sum_{l\in a} V^{\sigma}_{\text{xc\&Hartree}}[n=1,m=1]\sum_{\mu\in l,\nu\in l} n^{\sigma}_{\mu\nu}\, n^{\sigma}_{\nu\mu} \tag{17}$$

This is then written as a type of non-double-counted LDA+U (i.e., DC[n] is absent from eq 8), without ambiguity in the choice of $(U - J)$ because the pre-factor comes from the exchange−correlation potential. Additionally, atomic forces can be derived.[29]

This has some similarity to the ASIC form of Pemmaraju et al.;[31] however, they use the "full" form of the projectors[29] and do not provide a variationally connected energy. This functional form has also been independently developed by Seo;[26] in this approach the electronic relaxation can be included in the adjustable U value instead of by use of a scaling parameter.

C. Exchange−Correlation and the Atomic Parameters. We now discuss an explicit relation between the LDA+U constants and the on-site LSDA potential. Equation 44 of Anisimov et al.[20] relates the atomic F^0 Slater integral and the exchange, J, to the LDA exchange−correlation potential for an atomic state as

$$V^{m\sigma}_{\text{LSDA}} = F^0N - \frac{1}{2}(F^0 - J) - JN_{\sigma} \tag{18}$$

Thus for the orbital occupation choices for which V_{LSDA} is calculated in pSIC for LDA ($V_{\text{LSDA}}[n=1,m=1]$, i.e., $N = N_{\sigma} = 1$) the spherically symmetric part of the exchange−correlation potential is

$$V^{\sigma}_{\text{LSDA}}[n=1,m=1] = \frac{F^0 - J}{2} \tag{19}$$

In DFTB, we use U in the SCC-DFTB correction[8] and this is related to the screened F^0 Slater integral by[20]

$$U = F^0 - W^{\text{screening}} \tag{20}$$

Similarly, the diagonal part of the atomic spin coupling matrix, W, used in spin-polarized SCC-DFTB,[33] is equivalent to $-J_l/2$ (from considering the energy of two parallel spins within the same shell). If we assume in the zeroth limit that screening in the isolated atom, for which U is calculated in this work, is small ($W \approx 0$) then for the basis that is used to expand the local states in pSIC

$$V_{m\sigma}[n=1,m=1]^{\text{LSDA}}_{l} \approx \frac{(U - J)^{\text{atomic}}_{l}}{2} \tag{21}$$

(Because in DFTB, U is calculated numerically from atomic calculations, we actually include this screening contribution.)

In this approximation, pSIC then gives a contribution to the potential of

$$\Delta V^{\text{pSIC}} = -\alpha \frac{(U-J)_l^{\text{atomic}}}{2} n^{\sigma} \qquad (22)$$

By comparison with relaxation-corrected pSIC ($\alpha = {}^1/_2$) for a system with a Hubbard gap, the same potential would be applied to the lower Hubbard band by a FLL-LDA+U contribution of

$$\Delta V^{\text{FLL}} = -\frac{(U-J)_l^{\text{atomic}}}{2}\left(n^{\sigma} - \frac{1}{2}\right) \qquad (23)$$

Because the fully occupied states in the local manifold experience a net potential shift of $-(U - J)_l^{\text{atomic}}/4$.

This suggests that LDA+U and relaxation-corrected pSIC have the same effect on the occupied band structure, and that $U - J = (U - J)_l^{\text{atomic}}/2$ is a sensible first choice for the parameters in LDA+U. This agrees with the (empirical) choice of $\approx 0.5(U - J)_l^{\text{atomic}}$ being suitable for many LDA+U applications.[16,20] Because there are different potentials for unoccupied states in pSIC and FLL−LDA+U these methods give different gaps,[31] and different total energy corrections.

All of the above corrections share the feature that they are semilocal (decaying on the length scale of the overlap matrix in the dual basis form); hence they cannot fully address effects such as the derivative discontinuity in Kohn−Sham theory.[34,35] To correct for this effect requires removing the restriction that only on-site blocks of the occupation matrix must be considered.

1. Explicit DFTB Expressions. For clarity, the explicit additions used to apply the LDA+U and pseudo-SIC corrections to spin-resolved SCC-DFTB are now given. The spin-DFTB energy expressions themselves are given for example in ref 36. The potential contributions in the three cases can be written as

$$\text{(FLL)} \qquad \Delta V_{\mu\nu}^{\sigma} = -\alpha (U-J)_l^{\text{atomic}}\left(n_{\mu\nu}^{\sigma} - \frac{1}{2}\delta_{\mu\nu}\right) \qquad (24)$$

$$\text{(pSIC)} \qquad \Delta V_{\mu\nu}^{\sigma} = -\alpha \frac{(U-J)_l^{\text{atomic}}}{2} n_{\mu\nu}^{\sigma} \qquad (25)$$

$$\text{(AMF)} \qquad \Delta V_{\mu\nu}^{\sigma} = -\alpha (U-J)_l^{\text{atomic}} \delta n_{\mu\nu}^{\sigma} \qquad (26)$$

where

$$\delta n_{\mu\nu}^{\sigma} = n_{\mu\nu}^{\sigma} - \frac{\sum_{\nu} n_{\nu\nu}^{\sigma}}{2l+1}\delta_{\mu\nu} \qquad (27)$$

μ and ν are within the same l shell of an atom. In the pSIC case, this correction is applied to all shells of all atoms, whereas for LDA+U, only the d/f states are corrected.

The additions to the total energy are

$$\Delta E^{\text{FLL}} = -\alpha \sum_{a}^{\text{atoms}} \sum_{l \in a} \frac{(U-J)_l^{\text{atomic}}}{2} \sum_{\sigma} \sum_{\mu\nu} ((n_{\mu\nu}^{\sigma})^2 - n_{\mu\nu}^{\sigma})_{\nu\mu \in l} \qquad (28)$$

$$\Delta E^{\text{pSIC}} = -\alpha \sum_{a}^{\text{atoms}} \sum_{l \in a} \frac{(U-J)_l^{\text{atomic}}}{2} \sum_{\sigma} \sum_{\mu\nu} (n_{\mu\nu}^{\sigma})_{\nu\mu \in l}^{2} \qquad (29)$$

$$\Delta E^{\text{AMF}} = -\alpha \sum_{a}^{\text{atoms}} \sum_{l \in a} \frac{(U-J)_l^{\text{atomic}}}{2} \sum_{\sigma} \sum_{\mu\nu} (\delta n_{\mu\nu}^{\sigma})_{\nu\mu \in l}^{2} \qquad (30)$$

The present results for NiO were obtained using the dual[29] basis form for the orbital occupation matrix, *n*, as given in eq 11. As discussed above, α is the scaling introduced in the work of Filippetti and Spaldin.[13]

D. NiO. To illustrate the discussion of LDA+U and pSIC within SCC-DFTB, we now present results for the antiferromagnetic II phase of NiO. The present calculations were performed with the experimental lattice constant for a four-atom FCC unit cell, using a 12^3 Monkhorst-Pack[37] *k*-point sampling grid. The *U* and *W* values used in the present work are taken from appendix C of ref 38 and are reproduced in Table 1.

LDA finds this material to be a metal, and LSDA finds a narrow gap antiferromagnet while severely underestimating the magnetic moment for each Ni atom. This material is a prototypical application for LDA+U.[20]

The density of states around the Fermi level for the three LDA+U-like methods implemented in SCC-DFTB (FLL, AMF, and pSIC) with a dual basis occupation matrix are shown in Figure 1. In the case of the two DFTB+U methods only the Ni^{3d} shell is corrected, whereas all states are treated with pSIC. The *U* and *J* parameters are obtained from the atomic values used in spin-resolved SCC-DFTB. The strength of the correction is scaled by α , between the spin-SCC-DFTB only limit at $\alpha = 0$ and the full relaxation uncorrected case at $\alpha = 1$. As discussed above, we suggest $\alpha = 0.5$ as a default first choice in accordance with the suggestion of Filippetti and Spaldin.[13] The spin only magnetic moment and band gap are shown in Figure 2.

For all three corrections, for $\alpha \gtrsim 0.2$ a gap opens up between the t_{2g} Ni^{3d} states at the valence band maximum and the O^{2p} conduction-band minimum. The unoccupied oxygen-p states remain unchanged on increasing the magnitude of the corrections, but the Ni-d states drop in energy. For the FLL correction, there is a band gap collapse at $\alpha \approx 1$, when the Ni-d collide with states around $E_F - 5$ eV, which also causes a loss of magnetism for the Ni (Figure 2). As shown in Table 2, with $\alpha = 0.5$, the magnetic moment of the two LDA+U methods is similar to experimental values, but slightly underestimates the value (there is also an additional contribution from orbital magnetism[39] which is neglected here). The pSIC results give a somewhat smaller magnetic moment. Both trends match what is observed with LSDA+U[16,39] and self-interaction corrected LSDA.[40] The obtained band gaps (E_g) are approximatetly half of the experimental values, with the pSIC case being further reduced. This again matches the behavior of their DFT counterparts. As discussed in section IV there is also a substantial contribution from including effects of derivative discontinuities. $E_g + \Delta$ is in much closer agreement with the experiment gap (this procedure should in fact overestimate optical gaps, due to the absence of excitonic contributions).

IV. Derivative Discontinuities and Gaps

At low temperatures, typical density functionals have a quadratic energy dependency as a function of total charge in open systems (this can also apply for the localized LDA+U-type corrections discussed in section III). The universal density functional, however, can show sections of straight line segments between integer numbers of particles[42] with an associated jump in the chemical potential at integer fillings. This behavior of approximate functionals is due to the failure of cancellation of

TABLE 1: Reproduction of the Atomic *U* and *W* Values (au) Taken from Ref 38

U	*s*	*p*	*d*	*W*	*d*	*p*	*s*
O	0.397	0.364	–	*p*		−0.027	
				s		−0.028	−0.032
Ni	0.256	0.227	0.487	*d*	−0.017		
				p	−0.001	−0.010	
				s	−0.003	−0.009	−0.009

the self-interactions of the Coulombic and exchange correlation contributions to the energy (For neutral systems, however, this at least allows for the possibility of a derivative discontinuity). This then has implication for the single-particle gap in Kohn−Sham-like theories.[34,35] Additionally, this has recently been shown to impact on transport properties calculated with DFT and related techniques.[43]

For the original non-self-consistent DFTB, because the band structure is independent of the number of particles, then automatically the energy shows such discontinuities on filling nondegenerate levels. Similarly, the chemical potential jumps at even filling. SCC-DFT, as with the usual DFT functionals, does not show such behavior due to the self-interaction of the long-range Hartree-like contributions.

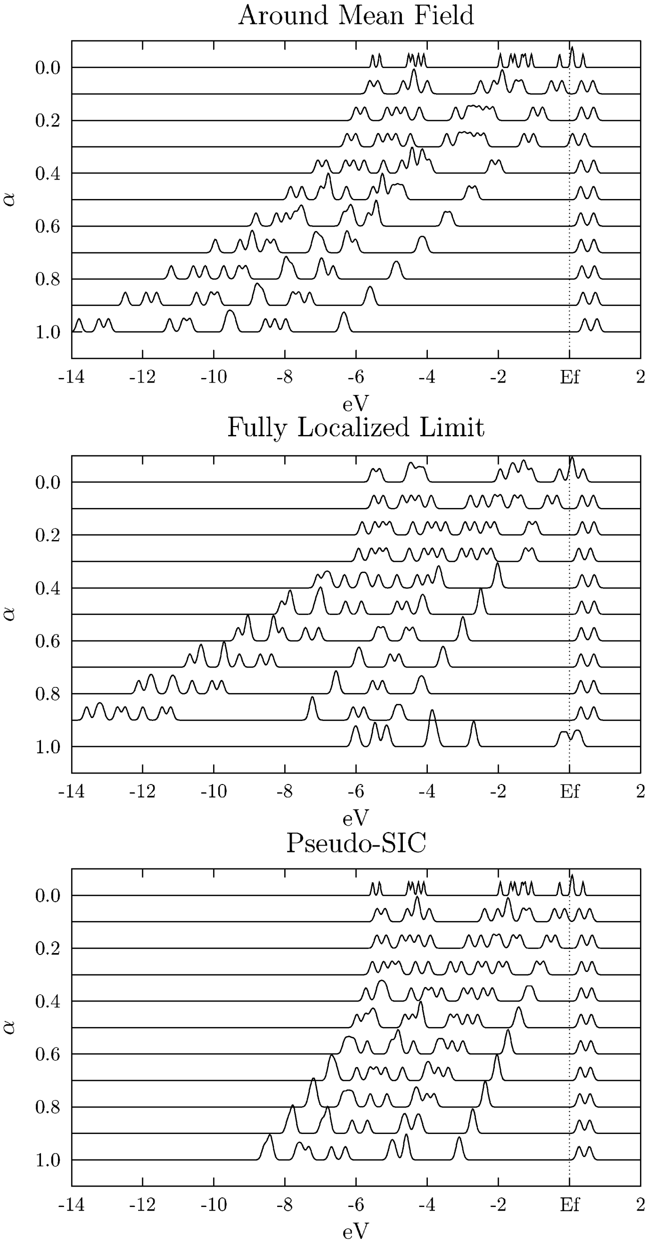

Figure 1. Density of states around the Fermi level for NiO in the AMF-II configuration using the three on-site corrections discussed in section III.

For a nondegenerate system with integer total number of particles (N), at low temperatures the reduced single-particle density matrix of the system is idempotent (i.e., it has eigenvalues, $n_i \in \{0, 1\}$, and hence the density matrix equals its own square, because for all occupation numbers $n_i = n_i^2$). For a nonorthogonal basis this requires that

$$\rho - \rho \cdot S \cdot \rho = 0 \tag{31}$$

where the matrix product uses ρ, the reduced single-particle density matrix, and S the overlap matrix of the basis in which it is represented. For a noninteger number of particles, N (and also for degenerate systems in an ensemble average state), some states of the system must instead have fractional single-particle occupation numbers, $0 < n_i < 1$, then the condition expressed in eq 31 is violated (but only for the subspace of the single-particle density matrix including such states). This relation then gives a method of projecting out a function of such fractionally occupied states. Hence, because the error in the approximate functional on going between a nondegenerate N particle system and an $N + 1$ system is normally quadratic in occupation of these fractionally occupied states, this gives a method to correct the energy to give the desired discontinuity.

One side issue is that of the spin degrees of freedom. If a spin-unpolarized approach is used, the idempotency condition will only be met for occupations equivalent to an even number of particles, hence any correction built on this approach would introduce the derivative discontinuity at even fillings (instead of all integers). Hence, as an alternative to constructing a projector using eq 31, these corrections can be written directly in terms of occupation number in a diagonal representation.

Because the approximate density functionals are usually considered to give a good estimate of the energy at integer particle numbers, a correction that leaves the energy of such systems unchanged is desirable. Choosing an energy contribution of the form

$$\Delta E_{\mathrm{DD}} = \sum_{\sigma} \frac{c(\mathcal{N})}{2} \mathrm{Tr}(S \cdot (\rho^{\sigma} - \rho^{\sigma} \cdot S \cdot \rho^{\sigma})) \tag{32}$$

for which there is no energy contribution for idempotent matrices, with $c(\mathcal{N})$ given by

$$c(\mathcal{N}) = \frac{\partial^2 E(\mathcal{N})}{\partial \mathcal{N}^2} \tag{33}$$

For a purely quadratic error in the functional, $c(\mathcal{N}) = c(N)$. In the case of systems with a HOMO−LUMO gap, there is a discontinuity present even in approximate functionals; hence a distinction should be made between derivatives going toward greater or fewer particle numbers.

The associated contribution to the potential is then

$$\Delta V_{\mathrm{DD}}^{\sigma} = c(\mathcal{N}) \left(\frac{S}{2} - S \cdot \rho^{\sigma} \cdot S \right) \tag{34}$$

which is linear in occupation number and has the form of a scissor operator between the occupied and un-occupied orbitals.

For $c(\mathcal{N})$, in the present examples this is evaluated numerically; however, it is possible to obtain this analytically, either from the second derivative of the energy as with eq 33, or as a

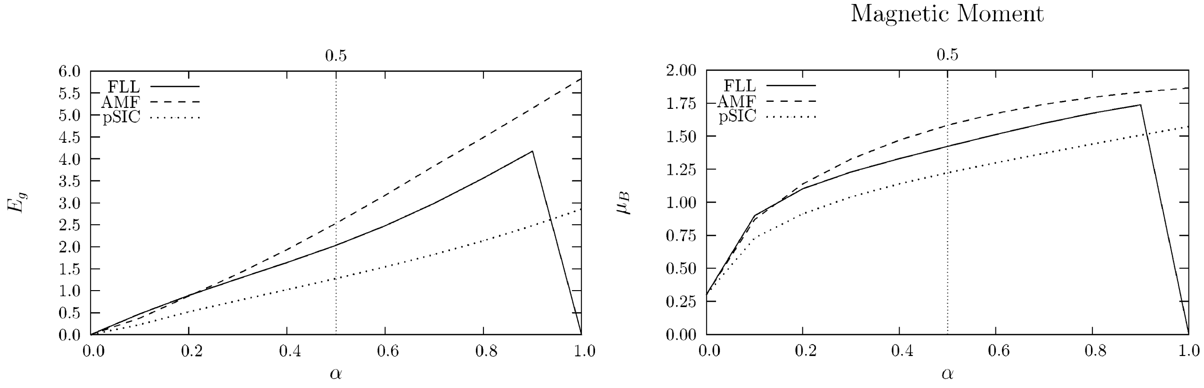

Figure 2. Band gap and spin contribution to the Ni magnetic moment in AMF-II NiO for the three on-site corrections discussed in section III.

TABLE 2: Spin Contribution to the Magnetic Moment per Ni Atom (μ_B) and Band Gap (eV) for the Antiferromagnetic Spin Configuration of NiO[a]

	μ_B	E_g	$E_g + \Delta$
SCC-DFTB	0	0.0	
spin-SCC-DFTB	0.31	0.0	
AMF	1.58	2.54	5.37
FLL	1.42	2.03	4.94
pSIC	1.22	1.28	3.81
LSDA-SIC[b]	1.156	~2.86	
LSDA+U[c]	1.71	2.73	
B3LYP[d]	1.67	4.1	
expt[e]	1.64−1.90	4.0−4.3	

[a] Results the SCC-DFTB, spin-resolved SCC-DFTB, and spin-resolved SCC-DFTB with the AMF, FLL, or pSIC contributions are presented. The orbitally dependent contributions use $\alpha = 0.5$ scaling, as discussed in the text. The results of SCC-DFTB + orbitally dependent contributions additionally include a correction for the derivative discontinuity ($E_g + \Delta$). [b] DFT results are taken from Ködderitzsch *et al.*[40] [c] DFT results are taken from Chaplygin.[39] [d] DFT results are taken from Moreira *et al.*[19] [e] Experimental data are taken from ref 41 and references therein.

TABLE 3: HOMO−LUMO Gaps (eV) for Test Molecules with CuS Contacting, Compared with the Energy Definition of the Gap, $E(N + 1) + E(N - 1) - 2E(N)$

	E based	ϵ_g (SCC-DFTB)	$\epsilon_g + \Delta$
benzene	5.134	1.285	5.358
bifenilene	4.476	1.351	4.603
trifenilene	4.119	1.406	3.808
tetrafenilene	3.893	1.385	4.132

first derivative of the HOMO level energy with respect to total charge by using Janak's theorem[44] and linear response.[45]

A. Application of the Derivative Discontinuity Correction. Functionals with the correct derivative discontinuity should have band gaps consistent with the thermodynamic charge-transfer gap, i.e., $E(N + 1) + E(N - 1) - 2E(N)$, where N is the number of electrons in the neutral state. This thermodynamic definition should match $\epsilon_{\mathrm{LUMO}} - \epsilon_{\mathrm{HOMO}}$, and this then gives a simple test for the validity of the proposed correction (if we assume relaxation effects in the charged systems are small). In principle, when comparing with optical gaps, however, one should also incorporate final state excitonic effects, but these have not been considered. Table 3 shows the thermodynamic and band-structure gaps of four small aromatic systems. These are example systems for molecular electronic devices, where two CuS contacts have replaced H atoms. The original SCC-DFTB band gaps do not agree well with the thermodynamic gaps, but applying the above scissor-like correction brings the band structure into much close agreement (but overestimating the gap in the FLL and AMF cases). Additionally, the chemical potential of the system now discontinuously jumps at integer occupation of the system, as would be required for example in a correct simulation of Coulomb blockade.

This correction has also been applied on-top of the local correlation corrections for NiO; see Table 2. This substantially improves the agreement with the observed optical gap (the correction accounts for nearly half of the band gap in this case).

V. Conclusions

This work discusses the connection between SCC-DFTB and the results of Foulkes and Haydock on the properties of tight-binding. Having demonstrated that SCC-DFTB has a well-defined ground state (bounded from below in energy), we discussed the issue of adding extra contributions to the DFTB expressions to produce ground states closer to the correct chemistry and physics of the system. For localized states, the connection between LDA+U and pseudo-self-interaction corrections is presented, giving a new prescription for a parameter free LDA+U and, for the first time, variationally connected pSIC energies and potentials. The developed methods are then tested on NiO, producing good agreement with density-functional based methods in the literature.

A further correction for the (absent) derivative discontinuity in SCC-DFTB and (many) DFT functionals is then suggested and applied to example molecular systems as well as NiO. The proposed correction improves the agreement between the single-particle levels and the thermodynamic HOMO−LUMO gap. In the case of NiO this is shown to bring the derived gap into much closer agreement with experiment.

Acknowledgment. This work was partially founded by the European Commission through the RENiBEl (Rare Earths Doped Nitrides for High Brightness Electroluminescence Emitters) Network (Contract No. HPRN-CT-2001-00297). B.H. greatfully acknowledges the Royal Society of Edinburgh BP Research Fellowship for funding. We thank Dr. Igor Chaplygin (Technische Universität Dresden) for useful discussion about LDA+U.

References and Notes

(1) Haugk, M.; Elstner, J.; Frauenheim, Th.; Staab, T. E. M.; Latham, C. D.; Jones, R.; Leipner, H. S.; Heine, T.; Seifert, G.; Sternberg, M. *Phys. Stat. Solidi B* **2000**, *217*, 473.

(2) Widany, J.; Frauenheim, Th.; Köhler, Th.; Sternberg, M.; Porezag, D.; Jungnickel, G.; Seifert, G. *Phys. Rev. B* **1996**, *53*, 4443.

(3) Elsner, J.; Jones, R.; Heggie, M. I.; Sitch, K.; Haugk, M.; Frauenheim, Th.; Oberg, S.; Briddon, P. R. *Phys. Rev. B* **1993**, *58*, 12571.

(4) Porezag, D.; Frauenheim, Th.; Köhler, Th.; Seifert, G.; Kaschner, R. *Phys. Rev. B* **1995**, *51*, 12947.

(5) Elsner, J.; Haugk, M.; Jungnickel, J.; Frauenheim, Th. *J. Mater. Chem.* **1996**, *6*, 1649.

(6) Elstner, M.; Frauenheim, Th.; Kaxiras, E.; Seifert, G.; Suhai, S. *Phys. Stat. Solidi B* **2000**, *217*, 375.

(7) Seifert, G.; Eschrig, H.; Bieger, W. *Z. Phys. Chem.* **1986**, *267*, 529.

(8) Elstner, M.; Porezag, D.; Jungnickel, G.; Elstner, J.; Haugk, M.; Frauenheim, T.; Shuai, S. and Seifert, G. *Phys. Rev. B* **1998**, *58*, 7260.

(9) Klopman, G. *J. Am. Chem.* Soc. **1964**, *86*, 4550.

(10) Ohno, K. *Theor. Chim. Acta* **1964**, *2*, 219.

(11) Mataga, N.; Nishimoto, K. *Z. Phys. Chem.* **1957**, *12*, 335.

(12) Vogl, D.; Krüger, P.; Pollmann, J. *Phys. Rev. B* **1998**, *58*, 3865; **1996**, *55*, 12838; **1998**, *58*, 3865.

(13) Filippetti, A.; Spaldin, N. A. *Phys. Rev. B* **2003**, *67*, 125109.

(14) Ruzsinszky *et al. J. Chem. Phys*. **2006**, *125*, 194112; **2007**, *126*, 104102.

(15) Foulkes, W. M.; Haydock, R. *Phys. Rev. B* **1989**, *39*, 12520.

(16) Petukhov, A. G.; Mazin, I. I.; Chioncel, L.; Lichtenstein, A. I. *Phys. Rev. B* **2003**, *67*, 153106.

(17) Perdew, J. P.; Zunger, A. *Phys. Rev. B* **1981**, *23*, 5048.

(18) Aryasetiawan, F.; Gunnarson, O. *Rep. Prog. Phys.* **1998**, *61*, 237.

(19) de Moreira, P. R.; Illas, F.; Martin, R. L. *Phys. Rev. B* **2002**, *65*, 155102.

(20) Anisimov, V. I.; Aryasetiawan, F.; Lichtenstein. A. I. *J. Phys. Condens. Matter* **1999**, *9*, 767.

(21) Kasinathan; et al. D. *Phys. Rev. B* **2006**, *74*, 195110.

(22) Mori-Sánchez, P.; Cohen, A. J.; Yang, W. *J. Chem. Phys.* **2006**, *125*, 201102.

(23) Czyżyk, M. T.; Sawatzky, G. A. *Phys. Rev. B* **1994**, *49*, 14211.

(24) Anisimov, V. I.; Solovyev, I. V.; Korotin, M. A.; Czyżyk, M. T.; Sawatzky, G. A. *Phys. Rev. B* **1993**, *48*, 16929.

(25) Dudarev, S. L.; Botton, G. A.; Savrasov, S. Y.; Humphreys, C. J.; Sutton, A. P. *Phys. Rev. B* **1998**, *57*, 1505.

(26) Seo, D.-K. To appear in *Phys. Rev. B*; cond-mat/0703503 2007.

(27) Papaconstantopoulos, D. A.; Hellberg, C. S. *Phys. Rev. Lett.* **2002**, *89*, 029701.

(28) Pickett, W. E.; Erwin, S. C.; Ethridge, E. C. *Phys. Rev. B* **1998**, *58*, 1201.

(29) Han, M. J.; Ozaki, T.; Yu, J. *Phys. Rev. B* **2006**, *73*, 045110.

(30) Eschrig, H.; Koepernik, K.; Chaplygin, I. *J. Solid State Chem.* **2003**, *176*, 482.

(31) Pemmaraju, C. D.; Archer, T.; Sanchez-Portal, D.; Sanvito, S. *Phys. Rev. B* **2007**, *75*, 045101.

(32) Liechtenstein, A. I.; Anisimov, V. I.; Zaanen, J. *Phys. Rev. B* **1995**, *52*, R5467.

(33) Köhler, C.; Seifert, G.; Frauenheim, Th. *Chem. Phys.* **2005**, *309*, 23.

(34) Perdew, P.; Levy, M. *Phys. Rev. Lett*. **1983**, *51*, 1884.

(35) Sham, J.; Schlüter, M. *Phys. Rev. Lett*. **1983**, *51*, 1888.

(36) Kohler, C.; Frauenheim, T.; Hourahine, B.; Seifert, G.; Sternberg, M. *J. Phys. Chem. A* **2007**, *111*, 5622.

(37) Monkhorst, H. J.; Pack, J. D. *Phys. Rev. B* **1976**, *13*, 5188.

(38) Köhler, C. Ph.D. Thesis, Univ. Paderborn, 2003.

(39) Chaplygin, I. Ph.D. Thesis, Tech. Univ. Dresden, 2002.

(40) Ködderitzsch, D.; Hergert, W.; Szotek, Z.; Temmerman, W. M. *Phys. Rev. B* **2003**, *68*, 125114.

(41) Anisimov, V. I.; Zaanen, J.; Andersen, O. K. *Phys. Rev. B* **1991**, *44*, 943.

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

(43) Toher, C.; Filippetti, A.; Sanvito, S.; Burke, K. *Phys. Rev. Lett.* **2005**, *95*, 146402.

(44) Janak, J. F.; *Phys. Rev. B* **1978**, *18*, 7165.

(45) Witek, H. A.; Irle, S.; Morokuma, K. *J. Chem. Phys.* **2004**, *121*, 5163.