

# LDA+U Method: Ver. 1.0

Taisuke Ozaki, RCIS, JAIST

August 15, 2007

## 1 Total energy

In conjunction with on-site terms of the unrestricted Hartree-Fock theory, the total energy of a LDA+U method [1] within the collinear spin treatment could be defined by

$$E_{\text{LDA+U}} = E_{\text{LDA}} + E_{\text{U}} \quad (1)$$

with

$$\begin{aligned} E_{\text{U}} &= \frac{1}{2} \sum_{\sigma} \sum_i \sum_p \sum_l U_{ipl} \left[ \text{Tr}(N_{ipl}^{\sigma}) - \text{Tr}(N_{ipl}^{\sigma} N_{ipl}^{\sigma}) \right], \\ &= \frac{1}{2} \sum_{\sigma} \sum_s U_s \left[ \text{Tr}(N_s^{\sigma}) - \text{Tr}(N_s^{\sigma} N_s^{\sigma}) \right], \end{aligned} \quad (2)$$

where  $i$  is a site index,  $l$  an angular momentum quantum number,  $p$  a multiplicity number of radial basis functions,  $\sigma$  a spin index, and  $s$  an organized index of  $(ipl)$ .  $N$  is an diagonalized occupation matrix.  $U$  is the effective Coulomb electron-electron interaction energy. Considering the rotational invariance of total energy with respect to each subshell  $s$ , Eq. (2) can be transformed as follows:

$$\begin{aligned} E_{\text{U}} &= \frac{1}{2} \sum_{\sigma} \sum_s U_s \left[ \text{Tr}(A_s N_s^{\sigma} A_s^{\dagger}) - \text{Tr}(A_s N_s^{\sigma} A_s^{\dagger} A_s N_s^{\sigma} A_s^{\dagger}) \right], \\ &= \frac{1}{2} \sum_{\sigma} \sum_s U_s \left[ \text{Tr}(n_s^{\sigma}) - \text{Tr}(n_s^{\sigma} n_s^{\sigma}) \right], \\ &= \frac{1}{2} \sum_{\sigma} \sum_s U_s \left[ \sum_m n_{smm}^{\sigma} - \sum_{m,m'} n_{smm'}^{\sigma} n_{sm'm}^{\sigma} \right]. \end{aligned} \quad (3)$$

In the Eq. (3), although off-diagonal occupation terms in each subshell  $s$  are taken into account, however, those between subshells are neglected. This treatment is consistent with their rotational invariant functional by Dudarev *et al.* [2], and is a simple extension of the rotational invariant functional for the case that a different U-value is given for each basis orbital indexed with  $s \equiv (ipl)$ . In this simple extension, we can not only include multiple d-orbitals as basis set, but also can easily derive the force on atoms in a simple form as discussed later on.

The  $E_{\text{LDA+U}}$  can be expressed in terms of the Kohn-Sham eigenenergies  $\varepsilon_{\nu}^{\sigma}$  as follows:

$$\begin{aligned} E_{\text{LDA+U}} &= E_{\text{LDA}} + E_{\text{U}}, \\ &= E_{\text{band}} + \left[ E_{\text{ee}} + E_{\text{cc}} + E_{\text{xc}} - \sum_{\sigma,\nu} \langle \psi_{\nu}^{\sigma} | \hat{v}_{\text{LDA}}^{\sigma} | \psi_{\nu}^{\sigma} \rangle \right] + \left[ E_{\text{U}} - \sum_{\sigma,\nu} \langle \psi_{\nu}^{\sigma} | \hat{v}_{\text{U}}^{\sigma} | \psi_{\nu}^{\sigma} \rangle \right], \end{aligned}$$

$$\begin{aligned}
&= E_{\text{band}} + \Delta E_{\text{LDA}} + \frac{1}{2} \sum_{\sigma} \sum_s U_s \sum_{m,m'} n_{smm'}^{\sigma} n_{sm'm}^{\sigma}, \\
&= E_{\text{band}} + \Delta E_{\text{LDA}} + \Delta E_{\text{U}},
\end{aligned} \tag{4}$$

where  $\Delta E_{\text{LDA}}$  and  $\Delta E_{\text{U}}$  are the double counting corrections of LDA- and U-energies, respectively.

## 2 Occupation number

The occupation number  $n$  may be defined by

$$n_{smm'}^{\sigma} = \sum_{\nu} f_{\nu} \langle \psi_{\nu}^{\sigma} | \hat{n}_{smm'}^{\sigma} | \psi_{\nu}^{\sigma} \rangle, \tag{5}$$

where, to count the occupation number  $n$ , we define three occupation number operators given by *on-site*

$$\hat{n}_{smm'}^{\sigma} = |s\tilde{m}\sigma\rangle \langle s\tilde{m}'\sigma|, \tag{6}$$

*full*

$$\hat{n}_{smm'}^{\sigma} = |sm\sigma\rangle \langle sm'\sigma|, \tag{7}$$

*dual*

$$\hat{n}_{smm'}^{\sigma} = \frac{1}{2} \left( |s\tilde{m}\sigma\rangle \langle sm'\sigma| + |sm\sigma\rangle \langle s\tilde{m}'\sigma| \right), \tag{8}$$

where  $|s\tilde{m}\sigma\rangle$  is the dual orbital of a original non-orthogonal basis orbital  $|sm\sigma\rangle$ , and is defined by

$$|s\tilde{m}\sigma\rangle = \sum_{s'm'} S_{sm,s'm'}^{-1} |s'm'\sigma\rangle \tag{9}$$

with the overlap matrix  $S$  between non-orthogonal basis orbitals. Then, the following bi-orthogonal relation is verified:

$$\langle s\tilde{m}\sigma | s'm'\sigma \rangle = \delta_{sm\sigma, s'm'\sigma}. \tag{10}$$

The *on-site* and *full* occupation number operators have been proposed by Eschrig *et al.* [3] and Pickett *et al.* [4], respectively. It is noted that these definitions do not satisfy a sum rule that the trace of the occupation number matrix is equivalent to the total number of electrons, while only the *dual* occupation number operator fulfills the sum rule as follows:

$$\sum_{\sigma} \text{Tr}(n^{\sigma}) = \sum_{\sigma} \frac{1}{2} \{ \text{Tr}(S\rho^{\sigma}) + \text{Tr}(\rho^{\sigma}S) \} = N_{\text{ele}}, \tag{11}$$

where  $\rho^{\sigma}$  is the density matrix defined by

$$\begin{aligned}
\rho_{sm,s'm'}^{\sigma} &= \sum_{\nu} f_{\nu} \langle \psi_{\nu}^{\sigma} | \hat{\rho}_{sm,s'm'}^{\sigma} | \psi_{\nu}^{\sigma} \rangle, \\
&= \sum_{\nu} f_{\nu} c_{s\tilde{m}}^{\sigma,*} c_{s'm'}^{\sigma}
\end{aligned} \tag{12}$$

with a density operator:

$$\hat{\rho}_{sm,s'm'}^{\sigma} = |s\tilde{m}\sigma\rangle \langle s'm'\sigma|. \tag{13}$$

The notes limit the discussion to non-Bloch wave functions for simplicity, but the extension is straightforward. For three definition of occupation number operators, *on-site*, *full*, and *dual*, the occupation numbers are given by

*on-site*

$$n_{smm'}^\sigma = \rho_{sm,sm'}^\sigma, \quad (14)$$

*full*

$$n_{smm'}^\sigma = \sum_{tn,t'n'} \rho_{tn,t'n'}^\sigma S_{tn,sm} S_{sm',t'n'}, \quad (15)$$

*dual*

$$n_{smm'}^\sigma = \frac{1}{2} \sum_{tn} \left\{ S_{sm',tn} \rho_{tn,sm}^\sigma + \rho_{sm',tn}^\sigma S_{tn,sm} \right\}. \quad (16)$$

### 3 Effective potential

The derivative of the total energy Eq. (1) with respect to LCAO coefficient  $c_{\nu,tn}^\sigma$  is given by

$$\begin{aligned} \frac{\partial E_{\text{LDA+U}}}{\partial c_{\nu,tn}^{\sigma,*}} &= \frac{\partial E_{\text{LDA}}}{\partial c_{\nu,tn}^{\sigma,*}} + \frac{\partial E_{\text{U}}}{\partial c_{\nu,tn}^{\sigma,*}}, \\ &= \frac{\partial E_{\text{LDA}}}{\partial c_{\nu,tn}^{\sigma,*}} + \sum_{smm'} \frac{\partial E_{\text{U}}}{\partial n_{smm'}^\sigma} \frac{\partial n_{smm'}^\sigma}{\partial c_{\nu,tn}^{\sigma,*}} \\ &= \frac{\partial E_{\text{LDA}}}{\partial c_{\nu,tn}^{\sigma,*}} + \sum_{smm'} U_s \left( \frac{1}{2} \delta_{mm'} - n_{smm'}^\sigma \right) \frac{\partial n_{smm'}^\sigma}{\partial c_{\nu,tn}^{\sigma,*}} \\ &= \frac{\partial E_{\text{LDA}}}{\partial c_{\nu,tn}^{\sigma,*}} + \sum_{smm'} v_{\text{U},smm'}^\sigma \frac{\partial n_{smm'}^\sigma}{\partial c_{\nu,tn}^{\sigma,*}} \end{aligned} \quad (17)$$

with

*on-site*

$$\frac{\partial n_{smm'}^\sigma}{\partial c_{\nu,tn}^{\sigma,*}} = \delta_{st} \delta_{mn} c_{\nu,sm'}^\sigma \quad (18)$$

*full*

$$\frac{\partial n_{smm'}^\sigma}{\partial c_{\nu,tn}^{\sigma,*}} = \sum_{t'n'} S_{tn,sm} S_{sm',t'n'} c_{\nu,t'n'}^\sigma \quad (19)$$

*dual*

$$\frac{\partial n_{smm'}^\sigma}{\partial c_{\nu,tn}^{\sigma,*}} = \frac{1}{2} \left\{ \delta_{st} \delta_{mn} \sum_{t'n'} S_{sm',t'n'} c_{\nu,t'n'}^\sigma + c_{\nu,sm'}^\sigma S_{tn,sm} \right\} \quad (20)$$

Substituting Eqs. (18)-(20) for the second term of Eq. (17), we see

*on-site*

$$\sum_{smm'} v_{\text{U},smm'}^\sigma \frac{\partial n_{smm'}^\sigma}{\partial c_{\nu,tn}^{\sigma,*}} = \sum_{t'n'} \langle tn\sigma | \left[ \sum_{smm'} |s\tilde{m}\sigma\rangle v_{\text{U},smm'}^\sigma \langle s\tilde{m}\sigma | \right] |t'n'\sigma\rangle c_{\nu,t'n'}^\sigma, \quad (21)$$

*full*

$$\sum_{smm'} v_{\text{U},smm'}^\sigma \frac{\partial n_{smm'}^\sigma}{\partial c_{\nu,tn}^{\sigma,*}} = \sum_{t'n'} \langle tn\sigma | \left[ \sum_{smm'} |sm\sigma\rangle v_{\text{U},smm'}^\sigma \langle sm'\sigma | \right] |t'n'\sigma\rangle c_{\nu,t'n'}^\sigma, \quad (22)$$

*dual*

$$\sum_{smm'} v_{\text{U},smm'}^\sigma \frac{\partial n_{smm'}^\sigma}{\partial c_{\nu,tn}^{\sigma,*}} = \sum_{t'n'} \langle tn\sigma | \frac{1}{2} \sum_{smm'} \left[ |s\tilde{m}\sigma\rangle v_{\text{U},smm'}^\sigma \langle sm'\sigma | + |sm\sigma\rangle v_{\text{U},smm'}^\sigma \langle s\tilde{m}'\sigma | \right] |t'n'\sigma\rangle c_{\nu,t'n'}^\sigma, \quad (23)$$

Therefore, the effective projector potentials  $\hat{v}_{\text{U}}^\sigma$  can be expressed by

*on-site*

$$\hat{v}_{\text{U}}^\sigma = \sum_{smm'} |s\tilde{m}\sigma\rangle v_{\text{U},smm'}^\sigma \langle s\tilde{m}'\sigma |, \quad (24)$$

*full*

$$\hat{v}_{\text{U}}^\sigma = \sum_{smm'} |sm\sigma\rangle v_{\text{U},smm'}^\sigma \langle sm'\sigma |, \quad (25)$$

*dual*

$$\hat{v}_{\text{U}}^\sigma = \frac{1}{2} \sum_{smm'} \left[ |s\tilde{m}\sigma\rangle v_{\text{U},smm'}^\sigma \langle sm'\sigma | + |sm\sigma\rangle v_{\text{U},smm'}^\sigma \langle s\tilde{m}'\sigma | \right]. \quad (26)$$

It is clear that the effective potentials of *on-site* and *full* are Hermitian. Also, it is verified that the effective potential of *dual* is Hermitian as follows:

$$\langle tn\sigma | \hat{v}_{\text{U}}^\sigma | t'n'\sigma \rangle = \frac{1}{2} \sum_{m'} v_{\text{U},tnm'}^\sigma S_{tm',t'n'} + \frac{1}{2} \sum_m S_{tn,t'm} v_{\text{U},t'mn'}^\sigma, \quad (27)$$

$$= \langle t'n'\sigma | \hat{v}_{\text{U}}^\sigma | tn\sigma \rangle. \quad (28)$$

It should be noted that in the full and dual the  $v_{\text{U}}^\sigma$  of the site  $i$  can affect the different sites by the projector potentials by Eqs. (25) and (26) because of the overlap.

## 4 Force on atom

The derivative of the total energy with respect to atomic coordinates  $\tau_k$  consists of two contributions:

$$\frac{\partial E_{\text{LDA+U}}}{\partial \tau_k} = \frac{\partial E_{\text{LDA}}}{\partial \tau_k} + \frac{\partial E_{\text{U}}}{\partial \tau_k}. \quad (29)$$

The first term can be evaluated in the same way as in the LDA. The second term is given by

$$\begin{aligned}
\frac{\partial E_U}{\partial \tau_k} &= \sum_{\sigma, smm'} \frac{\partial E_U}{\partial n_{smm'}^\sigma} \frac{\partial n_{smm'}^\sigma}{\partial \tau_k}, \\
&= \sum_{\sigma, smm'} v_{U, smm'}^\sigma \frac{\partial n_{smm'}^\sigma}{\partial \tau_k} \\
&= \sum_{\sigma, \nu} \sum_{tn, t'n'} \left\{ \frac{\partial c_{\nu, tn}^{\sigma, *}}{\partial \tau_k} \langle tn\sigma | \hat{v}_U^\sigma | t'n'\sigma \rangle c_{\nu, t'n'}^\sigma + c_{\nu, tn}^{\sigma, *} \langle tn\sigma | \hat{v}_U^\sigma | t'n'\sigma \rangle \frac{\partial c_{\nu, t'n'}^\sigma}{\partial \tau_k} + c_{\nu, tn}^{\sigma, *} c_{\nu, t'n'}^\sigma \frac{\partial \langle tn\sigma | \hat{v}_U^\sigma | t'n'\sigma \rangle}{\partial \tau_k} \right\}. \quad (30)
\end{aligned}$$

Considering  $Hc_\nu = \varepsilon_\nu Sc_\nu$  and  $c^\dagger Sc = I$ , the first and second terms in Eq. (30) can be transformed into derivatives of the overlap matrix. The third term in Eq. (30) is analytically differentiated, since it contains just two-center integrals.

## 5 Enhancement of orbital polarization

The LDA+U functional can possess multiple stationary points due to the degree of freedom in the configuration space of occupation ratio for degenerate orbitals. If electrons are occupied with a nearly same occupancy ratio in degenerate orbitals at the first stage of SCF steps, the final electronic state often converges a stationary minimum with non-orbital polarization after the SCF iteration. Also, it is often likely that electrons are disproportionately occupied in some of degenerate orbitals due to the exchange interaction, which is so-called 'orbital polarization'. As an example of the multiple minima, we can point out a cobalt oxide (CoO) bulk in which d-orbitals of the cobalt atom are split to  $t_{2g}$  and  $e_g$  states, and the five of seven d-electrons are occupied in  $t_{2g}$  and  $e_g$  states of the majority spin, and remaining two d-electrons are occupied in the  $t_{2g}$  state of the minority spin. Then, it depends on the initial occupancy ratios for the  $t_{2g}$  states of the minority spin how the remaining two d-electrons are occupied in three  $t_{2g}$  states. If the initial occupancy ratios are uniform, we may arrive at the non-orbital polarized state. In fact, unless any special treatment is considered for the initial occupancy ratios, we see the non-orbital polarized state of the CoO bulk. In order to explore the degree of freedom for the orbital occupation, therefore, it is needed to develop a general method which explicitly induces the orbital polarization. To induce the orbital polarization, a polarized redistribution scheme is proposed as follows:

$$\text{diagonalize} \quad d_s^\sigma = V^\dagger n_s^\sigma V \quad d_s^\sigma : \text{ascending order} \quad (31)$$

$$\text{summation} \quad D = \sum_{m=1}^{2l+1} d_{sm}^\sigma \quad (32)$$

$$\begin{aligned} \text{redistribution} \quad & d'_{2l+1} = 1, \\ & d'_{2l} = 1, \\ & \dots, \\ & d'_m = D - (2l + 1 - m), \\ & d'_{m-1} = 0, \dots \end{aligned} \quad (33)$$

$$\text{where } D = \sum_m d'_m \quad (34)$$

$$\text{back transform} \quad n'^\sigma_s = V d'^\sigma_m V^\dagger \quad (35)$$

After diagonalizing each subshell matrix consisting of occupation numbers, we introduce a polarized redistribution scheme given by Eq. (33) while keeping Eq. (34). Then, by a back transformation Eq. (35), we can obtain a polarized occupation matrix for each subshell. This polarized redistribution scheme is applied during the first few SCF steps, and then no modification is made during subsequent SCF steps. This proposed scheme maybe applicable to a general case: any crystal field, any number of electrons in the subshell, and any orbitals: p,d,f,...

## 6 Orbital optimization within LDA+U

In the orbital optimization within LDA+U, let us assume that the effective U-potential in the LDA+U method is applied to the primitive basis orbital  $\chi$  instead of the optimized basis orbital  $\phi$ , which is more natural in a physical sense than the opposite assumption.

A Kohn-Sham (KS) orbital  $\psi_\mu$  in the orbital optimization method is expressed by a linear combination of primitive orbitals  $\chi$ :

$$\begin{aligned}
|\psi_\nu^\sigma\rangle &= \sum_{i\alpha} c_{\mu,i\alpha}^\sigma |\phi_{i\alpha}\rangle, \\
&= \sum_{i\alpha} c_{\mu,i\alpha}^\sigma \left\{ \sum_q a_{i\alpha q} |\chi_{i\eta}\rangle \right\}, \\
&= \sum_{i\alpha} \sum_q c_{\mu,i\alpha}^\sigma a_{i\alpha q} |\chi_{i\eta}\rangle, \\
&= \sum_{i\eta} \left\{ \sum_p c_{\mu,i\alpha}^\sigma a_{i\alpha q} \right\} |\chi_{i\eta}\rangle, \\
&= \sum_{i\eta} b_{\mu,i\eta}^\sigma |\chi_{i\eta}\rangle, \tag{36}
\end{aligned}$$

where  $\alpha \equiv (plm)$ ,  $\eta \equiv (qlm)$ ,  $c$  and  $b$  are LCAO coefficients for contracted and primitive orbitals, respectively, and  $a$  contraction coefficients. For simplicity we consider a non-Bloch expression of the one-particle wave functions, but the extension of the below description to Bloch wave functions is straightforward. Assuming that the occupation number operators defined by Eqs. (6)-(8) are constructed by the primitive orbitals, we have the occupation numbers for the on-site, full, and dual given by

*on-site*

$$n_{smm'}^\sigma = \varrho_{sm,s'm'}^\sigma, \tag{37}$$

*full*

$$n_{smm'}^\sigma = \sum_{tn,t'n'} \varrho_{tn,t'n'}^\sigma S_{tn,sm} S_{sm',t'n'}, \tag{38}$$

*dual*

$$n_{smm'}^\sigma = \frac{1}{2} \sum_{tn} \left\{ S_{sm',tn} \varrho_{tn,sm}^\sigma + \varrho_{sm',tn}^\sigma S_{tn,sm} \right\}, \tag{39}$$

where  $\varrho^\sigma$  is the primitive density matrix defined by

$$\begin{aligned}
\varrho_{sm,s'm'}^\sigma &= \sum_\nu f_\nu \langle \psi_\nu^\sigma | \hat{\varrho}_{sm,s'm'}^\sigma | \psi_\nu^\sigma \rangle, \\
&= \sum_\nu f_\nu b_{sm}^{\sigma,*} b_{s'm'}^\sigma \tag{40}
\end{aligned}$$

with a primitive density operator:

$$\hat{\varrho}_{sm,s'm'}^\sigma = |\tilde{\chi}_{sm\sigma}\rangle \langle \tilde{\chi}_{s'm'\sigma}|. \tag{41}$$

Moreover, by defining a contracted density operator:

$$\hat{\rho}_{sm,s'm'}^\sigma = |\tilde{\phi}_{sm\sigma}\rangle \langle \tilde{\phi}_{s'm'\sigma}|, \tag{42}$$

we have the contracted density matrix  $\rho^\sigma$  given by

$$\begin{aligned}
\rho_{sm,s'm'}^\sigma &= \sum_\nu f_\nu \langle \psi_\nu^\sigma | \hat{\rho}_{sm,s'm'}^\sigma | \psi_\nu^\sigma \rangle, \\
&= \sum_\nu f_\nu c_{sm}^{\sigma,*} c_{s'm'}^\sigma. \tag{43}
\end{aligned}$$

Then, the primitive density matrix  $\varrho$  is written by the contracted density matrix  $\rho$  as follows:

$$\varrho_{iqlm,i'q'l'm'}^{\sigma} = \sum_{p,p'} a_{iplmq} a_{i'p'l'm'q'} \rho_{iplm,i'p'l'm'}^{\sigma}. \quad (44)$$

Considering the variation of the total energy Eq. (1) with respect to  $b$ , we find the effective potentials of the LDA+U method with respect to the primitive basis orbital. They are given by the same expression as Eqs. (24)-(26), while the occupation number is given by Eqs. (37)-(39). After the Hamiltonian matrix with respect to the primitive basis orbital  $\chi$  is constructed, it is transformed to that of the optimized basis orbital  $\phi$  as follows:

$$\langle \phi_{iplm} | \hat{H} | \phi_{i'p'l'm'} \rangle = \sum_{q,q'} a_{iplmq} a_{i'p'l'm'q'} \langle \chi_{iqlm} | \hat{H} | \chi_{i'q'l'm'} \rangle. \quad (45)$$

The Hamiltonian matrix with respect to the contracted basis orbital is diagonalized. The procedure is summarized as follows:

1. diagonalize the contracted Hamiltonian  $\langle \phi_{iplm} | \hat{H} | \phi_{i'p'l'm'} \rangle$
2. calculate the contracted density matrix by Eq. (43)
3. calculate the primitive density matrix by Eq. (44)
4. calculate the occupation number by Eq. (37), (38), or (39)
5. construct the Hamiltonian by Eq. (24), (25), or (26)
6. contract the Hamiltonian by Eq. (45)
7. return 1

Although the optimization procedure of the contracted coefficients  $a$  is not discussed here, it can be easily verified that the same procedure as in the LDA method is derived. Thus, the orbital optimization can be performed within the LDA+U method as well as the LDA method.

## References

- [1] M. J. Han, T. Ozaki, and J. Yu, Phys. Rev. B **73**, 045110 (2006).
- [2] S. L. Dudarev, G. A. Botton, S. Y. Savrasov, C. J. Humphreys, and A. P. Sutton, Phys. Rev. B **57**, 1505 (1998).
- [3] H. Eschrig, K. Koepernik, and I. Chaplygin, J. Solid State Chem. **176**, 482 (2003).
- [4] W. E. Pickett, S. C. Erwin, E. C. Ethridge, Phys. Rev. B **58**, 1201 (1998).