



CrossMark  
click for updates

## Review

**Cite this article:** Elstner M, Seifert G. 2014  
Density functional tight binding. *Phil. Trans. R.  
Soc. A* **372**: 20120483.  
<http://dx.doi.org/10.1098/rsta.2012.0483>

One contribution of 8 to a Theme Issue  
'Density functional theory across chemistry,  
physics and biology'.

### Subject Areas:

computational chemistry, chemical physics

### Keywords:

density-functional theory, density functional  
tight binding, semi-empirical methods

### Author for correspondence:

Marcus Elstner  
e-mail: [marcus.elstner@kit.edu](mailto:marcus.elstner@kit.edu)

# Density functional tight binding

Marcus Elstner<sup>1</sup> and Gotthard Seifert<sup>2</sup>

<sup>1</sup>Physical Chemistry, Karlsruhe Institute of Technology,  
Kaiserstrasse 12, Karlsruhe 76131, Germany

<sup>2</sup>Physical Chemistry, TU Dresden, Mommsenstrasse 13,  
Dresden 01062, Germany

This paper reviews the basic principles of the density-functional tight-binding (DFTB) method, which is based on density-functional theory as formulated by Hohenberg, Kohn and Sham (KS-DFT). DFTB consists of a series of models that are derived from a Taylor series expansion of the KS-DFT total energy. In the lowest order (DFTB1), densities and potentials are written as superpositions of atomic densities and potentials. The Kohn–Sham orbitals are then expanded to a set of localized atom-centred functions, which are obtained for spherical symmetric spin-unpolarized neutral atoms self-consistently. The whole Hamilton and overlap matrices contain one- and two-centre contributions only. Therefore, they can be calculated and tabulated in advance as functions of the distance between atomic pairs. The second contributions to DFTB1, the DFT double counting terms, are summarized together with nuclear repulsion energy terms and can be rewritten as the sum of pairwise repulsive terms. The second-order (DFTB2) and third-order (DFTB3) terms in the energy expansion correspond to a self-consistent representation, where the deviation of the ground-state density from the reference density is represented by charge monopoles only. This leads to a computationally efficient representation in terms of atomic charges (Mulliken), chemical hardness (Hubbard) parameters and scaled Coulomb laws. Therefore, no additional adjustable parameters enter the DFTB2 and DFTB3 formalism. The handling of parameters, the efficiency, the performance and extensions of DFTB are briefly discussed.

## 1. Introduction

Density-functional theory (DFT) within the orbital formulation by Kohn & Sham (KS) [1] is a very powerful

method for the calculation of physical and chemical properties of molecules, clusters and condensed matter systems. It is a very successful complement to the wave function-based methods, which build upon Hartree–Fock (HF) theory to include electron correlation.

Although efficient DFT-KS computer codes and the access to increasingly powerful computers allow one to study systems with remarkable size and complexity, there is demand for even faster but approximate methods to reach much larger system sizes, extended time scales in molecular dynamics simulations and to treat a multitude of structures of one molecular system. Classical force field methods (molecular mechanics, MM), on the other hand, are several orders of magnitude faster than DFT(HF), allowing one to treat hundreds of thousands to millions of atoms and to follow their dynamics beyond the nanosecond time scale. Such classical force fields contain a large number of empirical parameters and can be very accurate for the system they have been parametrized for, but may suffer from a limited transferability, from the unavailability of accurate force field parameters for a number of elements, and the missing quantum effects in such types of simulations.

Semi-empirical (SE) quantum chemical methods are located in the middle between DFT(HF) and MM methods, being roughly three orders of magnitude faster than DFT(HF) and slower than MM, containing a number of element-specific parameters, more than DFT methods, which are often parametrized nowadays as well, but significantly less than that in MM. These parameters require a reasonable effort in their determination and are sufficiently transferable, but most importantly, the SE methods can describe nearly all quantum effects accessible by DFT or HF and post-HF methods with few exceptions. One developmental line started from HF by neglecting and parametrizing the interaction integrals and resulted in the very successful neglect of diatomic differential overlap (NDDO) type of approaches, the most well known called MNDO, AM1 and PM3. Research along this line has been very active over the years, and several improved methods have been proposed, such as PM6 [2], PDDG-PM3 [3] or OMx [4], to name a few.

The empirical tight-binding (TB) [5] methods represent the counterpart to the SE quantum-chemical methods in the condensed phase with similar limitations. It was shown in the 1980s that a ‘TB’ scheme, i.e. the application of limited atomic orbital (AO) basis-set representation, can be used as the basis for a nearly parameter-free approximate treatment within the KS-DFT [6]. These early ideas have been further developed during the past 20 years, resulting in a density-functional theory-based tight-binding (DFTB) method. The DFTB method combines the computational efficiency of the SE ‘TB’ methods; however its parametrization is less empirical and more straightforward to conduct because it is basically related to a number of DFT calculations. In contrast to the NDDO-type methods, which start from HF and introduce correlation effects solely through parametrization, electron correlation in DFTB is included from the beginning owing to its descent from DFT. On the other hand, this strategy imports also all the DFT failures resulting from the use of approximate exchange correlation functionals.

The DFTB method has been applied to study large molecules (e.g. biomolecules), clusters, nanostructures and condensed-matter systems with a wide range of elements. The purpose of this paper is to describe the main ideas of DFTB and its potential, illustrated by a few selected applications. Finally, its conceptual limitations as well as extensions are discussed.

## 2. Theoretical approach

DFTB can be derived from a Taylor series expansion of the KS density functional total energy [1] around a properly chosen reference density  $\rho(r)$ . Instead of finding the electron density  $\rho(r)$  that minimizes the energy a reference density  $\rho^0$  is assumed which is perturbed by some density fluctuation,  $\rho(r) = \rho^0(r) + \delta\rho(r)$ . The exchange–correlation energy functional is then expanded in

a Taylor series up to the third order and the total energy can be written as

$$\begin{aligned}
 E^{\text{DFTB3}}[\rho^0 + \delta\rho] &= \frac{1}{2} \sum_{ab} \frac{Z_a Z_b}{R_{ab}} - \frac{1}{2} \iint \frac{\rho^0(\mathbf{r})\rho^0(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}' - \int V^{\text{xc}}[\rho^0]\rho^0(\mathbf{r}) d\mathbf{r} + E^{\text{xc}}[\rho^0] \\
 &+ \sum_i n_i \langle \psi_i | \hat{H}[\rho^0] | \psi_i \rangle + \frac{1}{2} \iint \left( \frac{1}{|\mathbf{r} - \mathbf{r}'|} + \frac{\delta^2 E^{\text{xc}}[\rho]}{\delta\rho(\mathbf{r})\delta\rho(\mathbf{r}')} \Big|_{\rho^0} \right) \delta\rho(\mathbf{r})\delta\rho(\mathbf{r}') d\mathbf{r} d\mathbf{r}' \\
 &+ \frac{1}{6} \iiint \frac{\delta^3 E^{\text{xc}}[\rho]}{\delta\rho(\mathbf{r})\delta\rho(\mathbf{r}')\delta\rho(\mathbf{r}'')} \Big|_{\rho^0} \delta\rho(\mathbf{r})\delta\rho(\mathbf{r}')\delta\rho(\mathbf{r}'') d\mathbf{r} d\mathbf{r}' d\mathbf{r}'' \\
 &= E^0[\rho_0] + E^1[\rho_0, \delta\rho] + E^2[\rho_0, (\delta\rho)^2] + E^3[\rho_0, (\delta\rho)^3]. \tag{2.1}
 \end{aligned}$$

Depending on the inclusion of terms from this expansion, different models appear, which are built successively on top of each other. Over the years, several DFTB models have been implemented, starting from the first-order non-self-consistent DFTB1 [7,8] (originally called DFTB), the second-order DFTB2 (originally called SCC-DFTB) [9] and the recent extension to the third-order DFTB3 [10–13].

### (a) DFTB1

The DFTB1 method takes the first two contributions of equation (2.1),  $E^0[\rho_0]$  and  $E^1[\rho_0, \delta\rho]$ , into account.

DFTB is based on a linear combination of atomic orbital (LCAO) ansatz of the KS orbitals

$$\psi_i = \sum_{\mu} c_{\mu i} \phi_{\mu}. \tag{2.2}$$

The AOs are obtained from DFT calculations of the corresponding atoms. The basis is restricted to the valence shell of the atoms. Such restriction requires an orthogonalization to the core orbitals of the atoms, which can be achieved by a Schmidt orthogonalization

$$|\phi_{\mu}\rangle = |\phi_{\mu}\rangle - \sum_{b \neq a} \sum_{\kappa} |\phi_{\kappa}^b\rangle \langle \phi_{\kappa}^b | \phi_{\mu}\rangle, \quad \mu \in \{a\}, \tag{2.3}$$

where  $|\phi_{\mu}\rangle$  is the valence AO  $\mu$  at atom  $a$  and  $|\phi_{\kappa}^b\rangle$  is a core orbital at atom  $b$ , as obtained from the corresponding atomic calculations. This orthogonalization is only necessary for the case  $a \neq b$ , because the AOs of the atomic DFT calculation represent an orthonormal set of functions. The LCAO ansatz leads to a general eigenvalue problem of the form

$$\sum_{\nu} c_{\nu i} (H_{\mu\nu}^0 - \epsilon_i S_{\mu\nu}) = 0 \quad \text{with } \nu \in b \text{ and } \forall a, \mu \in a, \tag{2.4}$$

with the Hamiltonian matrix elements  $H_{\mu\nu}^0$  and the overlap matrix elements  $S_{\mu\nu}$ . Considering the orthogonality construction (2.3) and writing the effective KS potential  $V_{\text{eff}}$  as a superposition of atomic-like potentials ( $V_{\text{eff}} \approx \sum_c V_{\text{eff}}^c$ ), one can write

$$(\phi_{\mu} | V_{\text{eff}}^c | \phi_{\nu}) \rightarrow (\phi_{\mu} | \left[ V_{\text{eff}}^c - \sum_{\kappa_c} |\phi_{\kappa_c}^c\rangle \epsilon_{\kappa_c} \langle \phi_{\kappa_c}^c | \right] | \phi_{\nu}) \quad \text{with } \nu \in b \text{ and } \mu \in a, \tag{2.5}$$

where  $\epsilon_{\kappa_c}$  is the energy of a core state  $\kappa$  at centre  $c$ . Together with its core correction term the potential can be interpreted as a pseudo-potential. The pseudo-potential appears only in the three-centre terms of the Hamiltonian matrix elements ( $a \neq b \neq c$ ) and in the 'diagonal terms'  $\nu, \mu \in a$

with  $c \neq a$ , whereas the ‘full’ potential appears in all the other terms

$$H_{\mu\nu}^0 = \left( \phi_\mu | -\frac{1}{2} \nabla^2 + \sum_j V_{\text{eff}}^c | \phi_\nu \right) - \sum_c \sum_{\kappa_c} (\phi_\mu | \phi_{\kappa_c}^c) \epsilon_{\kappa_c} (\phi_{\kappa_c}^c | \phi_\nu) \quad \mu, \nu \notin \{c\}. \quad (2.6)$$

The neglect of the three-centre terms and pseudo-potential contributions leads to the two-centre approximation for the Hamilton matrix elements

$$H_{\mu\nu}^0 = (\phi_\mu | -\frac{1}{2} \nabla^2 + V_{\text{eff}}^a + V_{\text{eff}}^b | \phi_\nu), \quad \mu \in \{a\}, \nu \in \{b\}. \quad (2.7)$$

The approximations lead to the same structure of the secular equations as in (non-orthogonal) TB schemes or the extended Hückel method; however all matrix elements are calculated within DFT, because the AO basis set  $\phi_\mu$  is computed by solving the DFT-KS equations for atoms. Pure AOs would be too diffuse for a minimal AO basis set; therefore, the atomic KS equations are usually solved applying an additional (harmonic) potential to the atomic KS equations, which leads to ‘compressed’ AOs or optimized AOs, as introduced by Eschrig [14]

$$\left[ -\frac{1}{2} \nabla^2 + v^{\text{eff}}[\rho^{\text{atom}}] + \left( \frac{r}{r_0} \right)^2 \right] \phi_\mu = \epsilon_\mu \phi_\mu. \quad (2.8)$$

The confinement radius is taken to be roughly two times the covalent radius of the atom [8]. The potentials  $V_{\text{eff}}^a = V_{\text{eff}}(\rho_a)$  contain the electron density of the neutral atoms  $a$ . This density is usually also determined from atomic KS equations, in the same way as the AO basis set. However, a slightly different confinement radius can be chosen for these initial densities  $\rho_a$ . Therefore, two parameters per element have to be determined at this stage, the confinement radius for the AO basis,  $r_0$ , and the confinement radius for the initial density,  $r_0^{\text{d}}$ .

Now, with AO basis and initial density determined, the KS equations can be solved (equation (2.4)) leading to the energy ( $n_i$ : occupation number of KS orbital  $i$ )

$$E^1 = \sum_i n_i \sum_{\mu\nu} c_\mu^i c_\nu^i H_{\mu\nu}^0 = \sum_i n_i \epsilon_i. \quad (2.9)$$

This is the electronic energy of the DFTB method, i.e. the sum of the occupied KS energies. To get an expression for the total energy, the  $E^0$  term has to be approximated.  $E^0[\rho_0]$  consists of the DFT ‘double counting’ contributions in the first line of equation (2.1) and depends only on the reference density  $\rho_0$ , i.e. the superposition of the neutral atomic densities. This means that this term is not dependent on the specific chemical environment, i.e. it can be determined for a ‘reference system’ and then be applied to other molecules in different chemical environments. This is the key to transferability of the parameters. In DFTB, this term is approximated by a sum of pair potentials called repulsive energy term [15]

$$E^0[\rho_0] \approx E_{\text{rep}} = \frac{1}{2} \sum_{ab} V_{ab}^{\text{rep}}, \quad (2.10)$$

which is either determined by comparison with DFT calculations [8] or fitted to empirical data [16]. The total energy then reads

$$E^{\text{DFTB1}} = \sum_i n_i \epsilon_i + \frac{1}{2} \sum_{ab} V_{ab}^{\text{rep}}. \quad (2.11)$$

Forces are readily computed as derivatives of the total energy, which allows very efficient geometry optimization or molecular dynamics simulations for molecular systems. This results, on the one hand, from the minimal basis representation and on the other hand from the efficient handling of the matrix elements. The  $H_{\mu\nu}$  and  $S_{\mu\nu}$  are computed once and stored, i.e. they do not have to be computed during the runtime of the DFTB program and they are read in from tables. Further, DFTB1 is a non-self-consistent TB method, i.e. the KS equations are solved only once. This makes DFTB1 about 5–10 times faster than DFTB2 and DFTB3, where a self-consistent solution of the eigenvalue problem has to be computed. DFTB1 is suitable for systems where

the charge transfer between the atoms is small, typically homonuclear systems or systems with atoms of similar electronegativity. Therefore, DFTB1 is very well suited, for example, for the description of hydrocarbons, because the higher order terms are small here. On the other hand, DFTB1 can also treat systems where a complete charge transfer between the atoms occurs, as e.g. in NaCl [10]. For systems exhibiting a delicate charge balance, higher order terms have to be considered.

## (b) DFTB2 and DFTB3

DFTB2 and DFTB3 approximate the  $E^2$  and  $E^3$  terms in equation (2.1) further. First, the density fluctuations are written as a superposition of atomic contributions

$$\delta\rho = \sum_a \delta\rho_a \quad (2.12)$$

and second, the atomic-like density fluctuations are expanded in a multipole expansion, however, only keeping the monopole term

$$\delta\rho_a = \Delta q_a F_{00}^a Y_{00}. \quad (2.13)$$

Practically, the second-order integral (third line in equation (2.1)) is evaluated assuming an exponentially decaying charge density

$$\delta\rho_a \approx \Delta q_a \frac{\tau_a^3}{8\pi} e^{-\tau_a |r-R_a|} \quad (2.14)$$

and neglecting the term  $\partial^2 E_{xc}/\partial\rho(r)\partial\rho(r')$  for the moment, i.e. only the Hartree integral is analytically evaluated. This leads to an analytical function  $\gamma_{ab}$ , and the second-order term finally reads (for  $a \neq b$ )

$$E^2(\tau_a, \tau_b, R_{ab}) = \frac{1}{2} \sum_{ab} \Delta q_a \Delta q_b \gamma_{ab}(\tau_a, \tau_b, R_{ab}). \quad (2.15)$$

The Hartree term therefore describes the interaction of the charge density fluctuations  $\delta\rho_a$  and  $\delta\rho_b$ , which reduces to a Coulomb interaction of partial charges  $\Delta q_a$  and  $\Delta q_b$  for large distances, i.e.  $\gamma_{ab}$  approaches  $1/R_{ab}$  for large distances.

For short distances, one has to deal with the electron–electron interaction within one atom, and here the exchange–correlation contributions become important. In this case,  $E^2$  describes the electron–electron interaction within one atom, and  $E^2$  becomes proportional to  $\Delta q_a^2$  and  $\gamma_{ab}$  describes the effective on-site electron–electron interaction. The exponent  $\tau_a$  is chosen such that the on-site value of the  $\gamma$ -function properly describes the atomic chemical hardness (or alternatively the Hubbard parameter as calculated from DFT) and thus implicitly takes into account the exchange–correlation contribution to the second-order term. The chemical hardness of an atom, or the Hubbard parameter, is the second derivative of the energy with respect to the charge density, i.e. exactly the third term in equation (2.1). Then, the Hartree energy can be evaluated analytically, and after introducing and approximating for the  $E_{xc}$  contribution this term boils down to an effective interaction between atomic charges, computed using the Mulliken scheme.

For  $R_{ab} = 0$ , we find  $E^2 = \frac{1}{2} \sum_a \Delta q_a^2 \gamma_{aa}$  and the function  $\gamma_{aa}$  is simply the second derivative of the total energy with respect to the charge of an atom, i.e. it equals the so-called Hubbard parameter  $U_a$ , which is twice the chemical hardness of an atom [9]. From the functional form of  $\gamma_{ab}$ , one finds that the Slater exponent of an atom  $a$  is related to the Hubbard parameter as

$$\tau_a = \frac{16}{5} U_a, \quad (2.16)$$

i.e. the width of the atomic charge density is inversely proportional to its chemical hardness. This relation is intuitive in that more diffuse atoms (or anions) have a smaller chemical hardness. A problem, however, arises because the function  $\gamma_{ab}$  encodes a constant proportionality, which turns out not to be valid across the periodic table. In fact, at least every row of the periodic table should

be treated slightly differently. The deviation is most pronounced between hydrogen and the first-row elements; therefore a modified  $\gamma_{ab}^h$  has been proposed [10,17]. This modified function can be applied within DFTB2, but it is not the 'default' option. Within DFTB3, it has become default and therefore is a key ingredient of the DFTB3 methodology [11–13].

For  $E^3$ , the same approximations are introduced as for  $E^2$ . Originally, only the diagonal terms were included [10,11,17]. The third-order terms describe the change of the chemical hardness of an atom with its charge state [10], i.e. in the third order, a new parameter is introduced, the chemical hardness derivative  $U_a^d$ . This parameter can be computed from DFT or fitted in order to improve the performance of the model. A function  $\Gamma_{ab}$  results as the derivative of the  $\gamma$ -function with respect to charge by introducing the Hubbard derivative parameter. In that way, the third-order terms within SCC-DFTB can be rather seen as a robust way to introduce the charge dependence capturing some deficiencies of problematic approximations within the second-order formalism, namely, the small size of the pseudo-AO basis as well as the very simplified density fluctuation scheme.

With all these approximations, the SCC-DFTB total energy in the third order is given by

$$E^{\text{DFTB3}} = \sum_{iab} \sum_{\mu \in a} \sum_{\nu \in b} n_i c_{\mu i} c_{\nu i} H_{\mu\nu}^0 + \frac{1}{2} \sum_{ab} \Delta q_a \Delta q_b \gamma_{ab}^h + \frac{1}{3} \sum_{ab} \Delta q_a^2 \Delta q_b \Gamma_{ab} + \frac{1}{2} \sum_{ab} V_{ab}^{\text{rep}}. \quad (2.17)$$

The derivative of this expression with respect to the molecular orbital coefficients leads to the corresponding KS equations

$$\sum_{\nu} c_{\nu i} (H_{\mu\nu} - \epsilon_i S_{\mu\nu}) = 0 \quad \text{with } \nu \in b \text{ and } \forall a, \mu \in a \quad (2.18)$$

and

$$H_{\mu\nu} = H_{\mu\nu}^0 + S_{\mu\nu} \sum_c \Delta q_c \left( \frac{1}{2} (\gamma_{ac}^h + \gamma_{bc}^h) + \frac{1}{3} (\Delta q_a \Gamma_{ac} + \Delta q_b \Gamma_{bc}) + \frac{\Delta q_c}{6} (\Gamma_{ca} + \Gamma_{cb}) \right), \quad (2.19)$$

where  $S_{\mu\nu}$  is the overlap matrix. The Hamilton matrix elements depend on the Mulliken charges  $q_a$  ( $\Delta q_a = q_a - Z_a$ ) which in turn depend on the molecular orbital coefficients  $c_{\mu i}$ , and thus these equations have to be solved self-consistently.

### (c) Parameters and efficiency

To parametrize the various DFTB models, one first has to determine four parameters per atom type. These are the confinement radii for the AOs  $\phi_{\mu}$ , which is called  $r_0$ , the confinement radii for the atomic densities  $\rho_0^a$ , which is called  $r_0^d$ , the atomic Hubbard parameters  $U_a$  and its derivative  $U_a^d$ , which both can be computed from DFT. The repulsive potentials are two-body contributions, therefore they are much more involved and some effort has to be invested in the parametrization process. Typically, on determining for every pair of atoms the total energy  $E^{\text{tot}}$  versus bond distance  $R$  for specific bonds using a high level DFT or *ab initio* method determines ( $E^1 + E^2 + E^3$ ) also versus bond distance, and then computes  $E_{\text{rep}}$  as

$$E_{\text{rep}}(R) = E^{\text{tot}} - (E^1 + E^2 + E^3). \quad (2.20)$$

The determination of  $r_0$  and  $r_0^d$  is an empirical procedure and can be quite involved [13,16], while  $U_a$  and  $U_a^d$  can be easily computed in principle. For the modified function  $\gamma^h$  one additional parameter appears, which is fitted to reproduce the water dimer binding energy [13].

The substantial advantage of using DFTB2 is its time/performance efficiency. Table 1 presents benchmark calculations for the CPU-time of a single-point energy calculation.<sup>1</sup>

DFTB2 is at least 250 times faster than RI-PBE and more than 1000 times faster than B3LYP. This acceleration is mainly owing to two factors: first the usage of a minimal basis set within

<sup>1</sup>All calculations were carried out on a single processor of a standard desktop PC. For SCC-DFTB, the DFTB+ code (<http://www.dftb.org>) was used. The DFT values were obtained using the TURBOMOLE program package [18]. For the PBE calculations, the resolution of the identity (RI) integral evaluation has been used [19]. As basis set for the DFT methods, we chose 6-31G(d) which is a rather small basis set for practical use.

**Table 1.** Calculation time in seconds for various molecules with DFT and DFTB2.

molecule	$n^a$	SCC-DFTB	RI-PBE <sup>b</sup>	B3LYP <sup>b,c</sup>
C <sub>60</sub> <sup>d</sup>	60	1	1112	9398
(Ala) <sub>20</sub> <sup>e</sup>	212	12	3418	27 605
(H <sub>2</sub> O) <sub>48</sub> <sup>f</sup>	144	3	769	3466
(H <sub>2</sub> O) <sub>123</sub> <sup>f</sup>	369	15	5488	30 822

<sup>a</sup>Number of atoms.<sup>b</sup>Basis set 6-31G(d).<sup>c</sup>B3LYP\_Gaussian keyword in TURBOMOLE.<sup>d</sup>Buckminsterfullerene.<sup>e</sup>Polyalanine in  $\alpha$ -helical form and including capping groups.<sup>f</sup>Water cluster.

SCC-DFTB and second the tabulation and neglect of integrals. For the water cluster (H<sub>2</sub>O)<sub>48</sub>, for example,  $N = 288$  basis function is needed for a minimal basis set and  $N = 864$  basis function for the 6-31G(d) basis set. The time-limiting step for obtaining the energy with all methods discussed here is a matrix-diagonalization, which scales with  $N^3$ . Thus, an acceleration just from using the minimal basis of a factor of 27 is achieved. The remaining factor is owing to the tabulation and neglect of integrals; in this example this factor is roughly 10 and 40 for comparison with RI-PBE and B3LYP, respectively.

#### (d) Extensions

Within DFT, the correct description of the weak van der Waals interactions, the London dispersion interactions, are not included in the DFTB approach as formulated above. However, the van der Waals interactions may be included in the DFTB approach *a posteriori*, which means that the van der Waals contributions are calculated afterwards and the energy  $E_{\text{disp}}$  is added to the total energy. The first implementation [20,21] showed a very promising performance, stimulating the combination with empirical dispersion also on the full DFT level [22–25]. Recently, Grimme has parametrized his D3 [25] correction for DFTB3, leading to an excellent performance of DFTB3-D3 for a large set of hydrogen- and van der Waals-bonded molecules.

As is well known, the popular DFT functionals fall short in a proper handling of strong electron correlation effects in molecules or solids. Practical ways to cure the problems related with electron localization are the LDA+U approach and the application of self-interaction correction schemes. Both have been implemented into the DFTB methodology [26]. The proper description of the correct balance between charge delocalization and polarization, for example in charge transfer complexes, is also a challenge for DFT. Rapacioli *et al.* [27] adapted recently a configuration interaction method, based on constrained DFT calculations, into the DFTB approach which allows one to investigate charge resonances in molecular complexes and to describe the proper dissociation behaviour.

Electronic excitations can be described using time-dependent density-functional response theory. The corresponding approach within the DFTB scheme has been put forward by Niehaus *et al.* [28]. The time-dependent density-functional approach was also adapted for the description of non-adiabatic processes [29] and later on combined with Tully's stochastic surface hopping algorithm [30].

For heavy atoms, a scalar relativistic treatment is used [31] and the spin-orbit coupling can also be considered [32].

In combination with the non-equilibrium Green's function technique, the DFTB approach was also extended for the calculation of electron transport properties on the molecular scale [33].

Although the DFTB approach is about two to three orders of magnitude faster than the DFT one, for large systems (more than about 1000 atoms) the scaling behaviour in the diagonalization

methods for the solution of the general eigenvalue problems becomes notable. Therefore, the so-called order- $N$  scaling schemes [34,35] or the Car–Parrinello treatment [36] have been implemented within DFTB. For calculations of even larger systems, a combination (QM/MM) of the DFTB method, as a quantum mechanical (QM) method, with empirical force fields (MM) has also been realized [37–39]. This scheme has further been extended to include also a continuum electrostatics environment in the DFTB/MM-GSBP scheme [40].

### 3. Performance and tests

DFTB is a set of parametrized computational models which require a careful parametrization before application. This has a number of implications. First, DFTB models are based on DFT; therefore these models also inherit most of the well-known DFT shortcomings. As the Hamilton matrix is computed using DFT-GGA (with PBE functional), the problem of self-interaction is also found in DFTB. This results in a number of problems, showing up in a problematic description of the KS energy spectrum, the overpolarizability problem, etc. The problematic description of van der Waals complexes was recognized early and a correction has been proposed (see above), which has later on also been adopted for full self-consistent DFT calculations [22,23]. Second, on top of DFT, the above-described approximations are introduced. This results in a limited flexibility of the computational scheme, leading to a restricted transferability of the parameters. It has been shown that for certain properties, the accuracy of DFT can be achieved. This holds, for example, for geometries and molecular geometries, or for geometries and vibrational frequencies. However, no parametrization could be derived which approaches DFT accuracy for both energies and vibrational frequencies simultaneously [13,16]. Therefore, different parametrizations have been proposed.

DFTB requires two sorts of parameters. First, there are the electronic parameters concerning the wave function and density confinement, which enter the  $E^1$  term, and the Hubbard parameters and its derivatives, which are usually computed from DFT and are therefore easy to determine. The confinement radii are a little more involved, because they can be chosen to optimize molecular properties (e.g. [13,16]). However, these parameters allow only a fine-tuning, i.e. they cannot change the performance of the models drastically. Second, there are the repulsive potential parameters constituting  $E^0$ . These are fitted to theoretical or experimental data and this fitting process is crucial for DFTB accuracy. However, only bonds are affected (no bond angles, dihedrals, etc.), i.e. the bond energies, bond distances and stretch vibrational frequencies can be modified.

DFTB has not been parametrized for the whole periodic table (for available parameters, see [www.dftb.org](http://www.dftb.org)). Basically, there are parameters available for organic systems and for inorganic systems, which are not all compatible to date. For materials science applications, there exist parameter sets also for a number of element combinations (parameter set matsci—for references, see [www.dftb.org](http://www.dftb.org)), which have been applied to several inorganic systems, based on the DFTB1 approximation (e.g. [41,42] and references therein). DFTB2 has been particularly parametrized for O, N, C, H [9], S [43] and Mg [44], a parameter set called mio. These ONCH parameters have been tested quite intensely with respect to their performance for atomization and reaction energies, molecular geometries, vibrational frequencies, ionization potentials and electron affinities, dipole moments and for non-covalent interactions [45–47]. DFTB2 performs very well for reaction energies [45], while for heats of formation other SE methods, in particular the newly developed OMx and PDDG-PM3 methods, turned out to be slightly superior.

Concerning geometries, the DFTB2 performance is close to that of full DFT-GGA, while for vibrational geometries larger deviations with respect to DFT have been reported. A special parametrization for vibrational frequencies turned out to be very successful leading also to a very accurate description for these type of properties [48]. A deeper analysis showed that there is indeed an optimization conflict: it is not possible within DFTB2 (and also DFTB3) to achieve simultaneously an accuracy comparable to that of DFT. Therefore, special parametrizations for vibrational frequencies have been proposed for DFTB2 [16] and DFTB3 [13].



Ionization potentials and electron affinities for small molecules are not very accurate using small basis set methods. In contrast to the NDO-type SE methods, these properties do not enter the DFTB parametrization process. Therefore, these values may be less accurate and a careful testing is required before the application to a new class of molecules. Dipole moments in DFTB are computed using a Mulliken population analysis, leading to an underestimation of these properties. Using parametrized charges similar to those supplied by the CM3 scheme, a huge improvement can be found [49,50]. However, the derivatives of charges and polarizabilities with respect to nuclear coordinates, as needed in order to compute IR and Raman intensities, cannot be improved by such simple extension as CM3 charges or the chemical potential equilibration extension [50,51].

The main shortcomings of DFTB2 are its overbinding, its performance for hydrogen-bonded complexes and charged species, which show up, for example, in the proton affinity values. Here, the extension to the third order and the modified function  $\gamma_{ab}^h$  has led to a major improvement. In the new DFTB3 parametrization [13], the overbinding has been removed leading to an excellent performance for reaction energies and a clear improvement for non-bonded interactions. Note that DFTB3/3OB is even better than DFT-GGA when only a small, double zeta-type, basis set is applied [13]. In particular hydrogen bonding energies, proton affinities and proton transfer barriers are very well described.

## 4. Conclusion

The DFTB models allow one now to treat a broad range of systems (molecules, clusters and solids) and their computational efficiency enables one to consider systems with up to several thousand atoms. Applications have been described including elements all over the periodic table; however a systematic parametrization across the periodic table is still missing.

With the realization of the third-order expansion in the total energy (DFTB3) an excellent performance with respect to accuracy and computational efficiency for organic and biomolecules could be reached being about two to three orders faster than DFT-GGA methods with small double zeta-type basis sets, depending on the implementation and system studied. In many cases, DFTB reaches the DFT accuracy, however not for all properties within one parametrization. The computational efficiency comes at a certain price, which has been described in this review. The derivation from DFT-GGA makes DFTB inherit the well-known DFT problems, which can be remedied only partially. In addition to this, the use of a minimal basis set, the limited flexibility owing to the monopole approximation at DFTB2/DFTB3 level and the neglect of three-centre integrals mark approximations which lead to a deterioration of the accuracy when compared with full DFT calculations. Unfortunately, there is really no free lunch.

## References

1. Kohn W, Sham LJ. 1965 Self-consistent equations including exchange and correlation effects. *Phys. Rev.* **140**, A1133–A1138. (doi:10.1103/PhysRev.140.A1133)
2. Stewart J. 2007 Optimization of parameters for semiempirical methods V: modification of NDDO approximations and application to 70 elements. *J. Mol. Model.* **13**, 1173–1213. (doi:10.1007/s00894-007-0233-4)
3. Repasky MP, Chandrasekhar J, Jorgensen WL. 2002 PDDG/PM3 and PDDG/MNDO: improved semiempirical methods. *J. Comput. Chem.* **23**, 1601–1622. (doi:10.1002/jcc.10162)
4. Kolb M, Thiel W. 1993 Beyond the MNDO model: methodical considerations and numerical results. *J. Comput. Chem.* **14**, 775–789. (doi:10.1002/jcc.540140704)
5. Goringe C, Bowler D, Hernandez E. 1997 Tight-binding modelling of materials. *Rep. Prog. Phys.* **60**, 1447–1512. (doi:10.1088/0034-4885/60/12/001)
6. Seifert G, Eschrig H, Bieger W. 1986 Eine approximative Variante des LCAO –  $X_\alpha$  Verfahrens. *Z. Phys. Chem.* **267**, 529–539.
7. Seifert G, Porezag D, Frauenheim T. 1996 Calculations of molecules, clusters, and solids with a simplified LCAO-DFT-LDA scheme. *Int. J. Quant. Chem.* **58**, 185–192. (doi:10.1002/(SICI)1097-461X(1996)58:2<185::AID-QUA7>3.0.CO;2-U)

8. Porezag D, Frauenheim T, Köhler T, Seifert G, Kaschner R. 1995 Construction of tight-binding-like potentials on the basis of density-functional theory: application to carbon. *Phys. Rev. B* **51**, 12 947–12 957. (doi:10.1103/PhysRevB.51.12947)
9. Elstner M, Porezag D, Jungnickel G, Elsner J, Haugk M, Frauenheim T, Suhai S, Seifert G. 1998 Self-consistent-charge density-functional tight-binding method for simulations of complex materials properties. *Phys. Rev. B* **58**, 7260–7268. (doi:10.1103/PhysRevB.58.7260)
10. Elstner M. 2007 SCC-DFTB: what is the proper degree of self-consistency. *J. Phys. Chem. A* **111**, 5614–5621. (doi:10.1021/jp071338j)
11. Yang Y, Yu H, York D, Cui Q, Elstner M. 2007 Extension of the self-consistent-charge density-functional tight-binding method: third-order expansion of the density functional theory total energy and introduction of a modified effective coulomb interaction. *J. Phys. Chem. A* **111**, 10 861–10 873. (doi:10.1021/jp074167r)
12. Gaus M, Cui Q, Elstner M. 2011 DFTB3: extension of the self-consistent-charge density-functional tight-binding method (SCC-DFTB). *J. Chem. Theory Comput.* **7**, 931–948. (doi:10.1021/ct100684s)
13. Gaus M, Goetz A, Elstner M. 2012 Parametrization and benchmark of DFTB3 for organic molecules. *J. Chem. Theory Comput.* **9**, 338–364. (doi:10.1021/ct300849w)
14. Eschrig H. 1989 *Optimized LCAO method and electronic structure of extended systems*. Berlin, Germany: Springer.
15. Seifert G, Joswig J-O. 2012 Density-functional tight binding: an approximate density-functional theory method. *WIREs Comput. Mol. Sci.* **2**, 456–465. (doi:10.1002/wcms.1094)
16. Gaus M, Chou C-P, Wittek H, Elstner M. 2009 Automatized parametrization of SCC-DFTB repulsive potentials: application to hydrocarbons. *J. Phys. Chem. A* **113**, 11 866–11 881. (doi:10.1021/jp902973m)
17. Elstner M. 2006 The SCC-DFTB method and its application to biological systems. *Theor. Chem. Acc.* **116**, 316–325. (doi:10.1007/s00214-005-0066-0)
18. TURBOMOLE V6.1 2009 A development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989–2007, TURBOMOLE GmbH, since 2007 (<http://www.turbomole.com>).
19. Ahlrichs R. 2004 Efficient evaluation of three-center two-electron integrals over Gaussian functions. *Phys. Chem. Chem. Phys.* **6**, 5119–5121. (doi:10.1039/B413539C)
20. Elstner M, Hobza P, Frauenheim T, Suhai S, Kaxiras E. 2001 Hydrogen bonding and stacking interactions of nucleic acid base pairs: a density functional-theory based treatment. *J. Chem. Phys.* **114**, 5149–5155. (doi:10.1063/1.1329889)
21. Zhechkov L, Heine T, Patchkovski S, Seifert G, Duarte H. 2005 An efficient a posteriori treatment for dispersion interaction in density functional-based tight binding. *J. Chem. Theory Comput.* **1**, 841–847. (doi:10.1021/ct050065y)
22. Wu Q, Yang W. 2002 Empirical correction to density functional theory for van der Waals interactions. *J. Chem. Phys.* **116**, 515–524. (doi:10.1063/1.1424928)
23. Grimme S. 2004 Accurate description of van der Waals complexes by density functional theory including empirical corrections. *J. Comput. Chem.* **25**, 1463–1473. (doi:10.1002/jcc.20078)
24. Grimme S. 2006 Semiempirical GGA-type density functional constructed with a long-range dispersion correction. *J. Comput. Chem.* **27**, 1787–1799. (doi:10.1002/jcc.20495)
25. Grimme S, Antony J, Ehrlich S, Krieg H. 2010 A consistent and accurate *ab initio* parametrization of density functional dispersion correction (DFT-D) for the 94 elements H–Pu. *J. Chem. Phys.* **132**, 154104. (doi:10.1063/1.3382344)
26. Hourahine B, Sanna S, Aradi B, Köhler C, Niehaus T, Frauenheim T. 2007 Self-interaction and strong correlation in DFTB. *J. Phys. Chem. A* **111**, 5671–5677. (doi:10.1021/jp070173b)
27. Rapacioli M, Spiegelman F, Scemama A, Mirtschink A. 2011 Modeling charge resonance in cationic molecular clusters: combining DFT-tight binding with configuration interaction. *J. Chem. Theory Comput.* **7**, 44–55. (doi:10.1021/ct100412f)
28. Niehaus TA, Suhai S, Della Sala F, Lugli P, Elstner M, Seifert G, Frauenheim T. 2001 Tight-binding approach to time-dependent density-functional response theory. *Phys. Rev. B* **63**, 085108. (doi:10.1103/PhysRevB.63.085108)
29. Frauenheim T *et al.* 2002 Atomistic simulations of complex materials: ground-state and excited-state properties. *J. Phys. Condens. Matter* **14**, 3015–3047. (doi:10.1088/0953-8984/14/11/313)

30. Mitric R, Werner U, Wohlgemuth M, Seifert G, Bonacic-Koutecky V. 2009 Nonadiabatic dynamics within time-dependent density functional tight binding method. *J. Phys. Chem. A* **113**, 12 700–12 705. (doi:10.1021/jp905600w)
31. Heera V, Seifert G, Ziesche P. 1984 A semi-relativistic variant of the scattered-wave  $X\alpha$  method. *J. Phys. B* **17**, 519–530. (doi:10.1088/0022-3700/17/4/008)
32. Kohler C, Frauenheim T, Hourahine B, Seifert G, Sternberg M. 2007 Treatment of collinear and noncollinear electron spin within an approximate density functional based method. *J. Phys. Chem. A* **111**, 5622–5629. (doi:10.1021/jp068802p)
33. Di Carlo A, Gheorghe M, Lugli P, Sternberg M, Seifert G, Frauenheim T. 2002 Theoretical tools for transport in molecular nanostructures. *Physica B* **314**, 86–90. (doi:10.1016/S0921-4526(01)01445-4)
34. Stephan U. 2000 Comparison of the convergence properties of linear-scaling electronic-structure schemes for nonorthogonal bases. *Phys. Rev. B* **62**, 16 412–16 424. (doi:10.1103/PhysRevB.62.16412)
35. Sternberg M, Galli G, Frauenheim T. 1999 NOON: a non-orthogonal localised orbital order-N method. *Comput. Phys. Commun.* **118**, 200–212. (doi:10.1016/S0010-4655(99)00202-7)
36. Rapacioli M, Barthel R, Heine T, Seifert G. 2007 Car–Parrinello treatment for an approximate density-functional theory method. *J. Chem. Phys.* **126**, 124103. (doi:10.1063/1.2566510)
37. Han W, Elstner M, Jalkanen K, Frauenheim T, Suhai S. 2000 Hybrid SCC-DFTB/molecular mechanical studies of H-bonded systems and of N-acetyl-(L-Ala)<sub>n</sub> N'-methylamide helices in water solution. *Int. J. Quant. Chem.* **78**, 459–479. (doi:10.1002/(SICI)1097-461X(2000)78:6<459::AID-QUA7>3.0.CO;2-R)
38. Cui Q, Elstner M, Kaxiras E, Frauenheim T, Karplus M. 2001 A QM/MM implementation of the self-consistent charge density functional tight binding (SCC-DFTB) method. *J. Phys. Chem. B* **105**, 569–585. (doi:10.1021/jp0029109)
39. Saebra G, Walker RC, Elstner M, Case DA, Roitberg A. 2007 Implementation of the SCC-DFTB method for hybrid QM/MM simulations within the Amber molecular dynamics package. *J. Chem. Phys. A* **111**, 5655–5664. (doi:10.1021/jp0700711)
40. Riccardi D *et al.* 2006 Development of effective quantum mechanical/molecular mechanical (QM/MM) methods for complex biological processes. *J. Phys. Chem. B* **110**, 6458–6469. (doi:10.1021/jp056361o)
41. Frenzel J, Joswig J, Sarkar P, Seifert G, Springborg M. 2005 The effects of organisation, embedding and surfactants on the properties of cadmium chalcogenide (CdS, CdSe and CdS/CdSe) semiconductor nanoparticles. *Eur. J. Inorg. Chem.* **2005**, 3585–3596. (doi:10.1002/ejic.200500390)
42. Lorenz T, Teich D, Joswig J-O, Seifert G. 2012 Theoretical study of the mechanical behavior of individual TiS<sub>2</sub> and MoS<sub>2</sub> nanotubes. *J. Phys. Chem. C* **116**, 11 714–11 721. (doi:10.1021/jp300709w)
43. Niehaus TA, Elstner M, Frauenheim T, Suhai S. 2001 Application of an approximate density-functional method to sulfur containing compounds. *J. Mol. Struct. THEOCHEM* **541**, 185–194. (doi:10.1016/S0166-1280(00)00762-4)
44. Elstner M, Cui Q, Munih P, Kaxiras E, Frauenheim T, Karplus M. 2003 Modeling zinc in biomolecules with the self consistent charge-density functional tight binding (SCC-DFTB) method: applications to structural and energetic analysis. *J. Comput. Chem.* **24**, 565–581. (doi:10.1002/jcc.10201)
45. Krüger T, Elstner M, Schiffels P, Frauenheim T. 2005 Validation of the density-functional based tight-binding approximation method for the calculation of reaction energies and other data. *J. Chem. Phys.* **122**, 114110. (doi:10.1063/1.1871913)
46. Sattelmeyer KW, Tirado-Rives J, Jorgensen WL. 2006 Comparison of SCC-DFTB and NDDO-based semiempirical molecular orbital methods for organic molecules. *J. Phys. Chem. A* **110**, 13 551–13 559. (doi:10.1021/jp064544k)
47. Otte N, Scholten M, Thiel W. 2007 Looking at self-consistent-charge density functional tight binding from a semiempirical perspective. *J. Phys. Chem. A* **111**, 5751–5755. (doi:10.1021/jp0700130)
48. Małolepsza E, Witek HA, Morokuma K. 2005 Accurate vibrational frequencies using the self-consistent-charge density-functional tight-binding method. *Chem. Phys. Lett.* **412**, 237–243. (doi:10.1016/j.cplett.2005.06.093)

49. Kalinowski JA, Lesyng B, Thompson JD, Cramer CJ, Truhlar DG. 2004 Class IV charge model for the self-consistent charge density-functional tight-binding method. *J. Phys. Chem. A* **108**, 2545–2549. (doi:10.1021/jp037288+)
50. Kaminski S, Gaus M, Elstner M. 2012 Improved electronic properties from third-order SCC-DFTB with cost efficient post-SCF extensions. *J. Phys. Chem. A* **116**, 11 927–11 937. (doi:10.1021/jp307264f)
51. Kaminski S, Giese TJ, Gaus M, York DM, Elstner M. 2012 Extended polarization in third-order SCC-DFTB from chemical-potential equalization. *J. Phys. Chem. A* **116**, 9131–9141. (doi:10.1021/jp306239c)