# 1 Extended Hückel Theory Matrix for GFN2-xTB

$$
\begin{aligned}
H^{EHT}_{\mu\nu} &= \frac{1}{2} K^{ll'}_{AB} S_{\mu\nu} (H_{\mu\mu} + H_{\nu\nu}) \\
&\cdot X(EN_A, EN_B) \\
&\cdot \Pi(R_{AB}, l, l') \\
&\cdot Y(\zeta^A_l, \zeta^B_{l'}), \forall \mu \in l(A), \nu \in l'(B)
\end{aligned} \tag{1}
$$

where $\mu$ and $\nu$ are AO indecies, $l$ and $l'$ index shells. Both AO's are associated with an atom labled A and B. $K^{ll'}_{AB}$ is a element and shell specific fitted constant however, in GFN2 it only depends on the shells. $S_{\mu\nu} = \langle \phi_\mu | \phi_\nu \rangle$ is just the overlap of the orbitals. In GFN2 $H_{\kappa\kappa} = h^l_A - \delta h^l_{CN'_A} CN'_A$ where $CN'_A$ is the modified GFN2-type Coordinate Number for the element of atom A.

$$
\begin{aligned}
CN'_A = &\sum_{B \neq A}^{N_{\text{atoms}}} (1 + e^{-10(4(R_{A,\text{cov}} + R_{B,\text{cov}})/3R_{AB} - 1)})^{-1} \\
&\times (1 + e^{-20(4(R_{A,\text{cov}} + R_{B,\text{cov}} + 2)/3R_{AB} - 1)})^{-1}
\end{aligned} \tag{2}
$$

$h^l_A$ and $\delta h^l_{CN'_A}$ are both fitted constants. $EN_A$ is the electronegativity of the element of atom A, given in the original `xtb` code.

$$X(EN_A, EN_B) = 1 + k_{EN} \Delta EN^2_{AB} \tag{3}$$

$$k_{EN} = 0.02 \text{ in GFN2} \tag{4}$$

$$\Delta EN^2_{AB} = (EN_A - EN_B)^2 \tag{5}$$

$$\Pi(R_{AB}, l, l') = \left(1 + k^{\text{poly}}_{A,l} \left(\frac{R_{AB}}{R_{\text{cov},AB}}\right)^{\frac{1}{2}}\right) \left(1 + k^{\text{poly}}_{B,l'} \left(\frac{R_{AB}}{R_{\text{cov},AB}}\right)^{\frac{1}{2}}\right) \tag{6}$$

$R_{\text{cov},AB}$ are the summed covalent radii ($R_{\text{cov},A} + R_{\text{cov},B}$), e.g. $R_{\text{cov},H} = 0.32$, $R_{\text{cov},C} = 0.75$ are given in the original `xtb` code. $k^{\text{poly}}_{A,l}$ and $k^{\text{poly}}_{B,l'}$ are element and shell specific constants.

$$Y(\zeta^A_l, \zeta^B_{l'}) = \left(\frac{2\sqrt{\zeta^A_l \zeta^B_{l'}}}{\zeta^A_l + \zeta^B_{l'}}\right)^{\frac{1}{2}} \tag{7}$$

Here, $\zeta^A_l$ are the STO exponents of the GFN2-xTB AO basis.
Slater Type Orbitals are defined as such:

$$\chi_{\zeta,n,l,m}(r, \theta, \varphi) = N Y_{l,m}(\theta, \varphi) r^{n-1} e^{-\zeta r} \tag{8}$$

N is a normalisation constant, Y are spherical harmonic funtions, n, l, m are the quantum numbers for the AO. $r, \theta, \varphi$ are polar 3D coordinates. $\zeta$ determines the radial extent of the STO, a large value gives rise to a function that is "tight" around the nucleus and a small value gives a more "diffuse" function. This $\zeta$ is the one mentioned in the Y term of $E_{EHT}$ and is a value fitted when constructing the basis set, thus it is given to us.

# 2 Fock Matrix for GFN2-xTB

$$\begin{aligned} F_{\mu\nu}^{GFN2-xTB} = H_{\mu\nu}^{EHT} + F_{\mu\nu}^{IES+IXC} + F_{\mu\nu}^{AES} + F_{\mu\nu}^{AXC} + F_{\mu\nu}^{D4}, \\ \forall \mu \in A, \nu \in B \end{aligned} \tag{9}$$

## 2.1 Isotropic Electrostatic and Exchange-correlation contribution

$$\begin{aligned} F_{\mu\nu}^{IES+IXC} &= -\frac{1}{2} S_{\mu\nu} \sum_C \sum_{l''} (\gamma_{AC,ll''} + \gamma_{BC,l'l''}) q_{C,l''} \\ &- \frac{1}{2} S_{\mu\nu} (q_{A,l}^2 \Gamma_{A,l} + q_{B,l'}^2 \Gamma_{B,l'}) \end{aligned} \tag{10}$$

$l, l', l''$ being the angular momenta of the orbitals $\mu, \nu$ and each of C's orbitals.

$$\Gamma_{A,l} = K_l^{\Gamma} \Gamma_A \tag{11}$$

$K_l^{\Gamma}$ is a shell specific constant common for all elements and $\Gamma_A$ is an element specific constant.

$$\gamma_{AB,ll'} = \frac{1}{\sqrt{R_{AB}^2 + \eta_{AB,ll'}^{-2}}} \tag{12}$$

$$\eta_{AB,ll'} = \frac{1}{2} \left[ \eta_A (1 + k_A^l) + \eta_B (1 + k_B^{l'}) \right] \tag{13}$$

$q_l$ is a partial Mulliken charge. $\eta_A$ and $\eta_B$ are element-specific fit parameters, while $k_A^l$ and $k_B^{l'}$ are element-specific scaling factors for the individual shells ($k_A^l = 0$ when $l = 0$).

$$GAP_A = \sum_{l \in A} q_{A,l} \tag{14}$$

$$q_{A,l} = \sum_{l' \in B} P_{ll'} S_{ll'} = GOP_l \tag{15}$$

## 2.2 Anisotropic Electrostatic and Exchange-correlation contribution

$$\begin{aligned}
F_{\mu\nu}^{AES} + F_{\mu\nu}^{AXC} &= \frac{1}{2} S_{\mu\nu} \left[ V_S(\boldsymbol{R}_B) + V_S(\boldsymbol{R}_C) \right] \\
&+ \frac{1}{2} \boldsymbol{D}_{\mu\nu}^T \left[ \boldsymbol{V}_D(\boldsymbol{R}_B) + \boldsymbol{V}_D(\boldsymbol{R}_C) \right] \\
&+ \frac{1}{2} \sum_{\alpha,\beta \in \{x,y,z\}} Q_{\mu\nu}^{\alpha\beta} \left[ V_Q^{\alpha\beta}(\boldsymbol{R}_B) + V_Q^{\alpha\beta}(\boldsymbol{R}_C) \right]
\end{aligned} \tag{16}$$

$$\boldsymbol{D}_{\mu\nu}^T = \begin{pmatrix} D_{\mu\nu}^x & D_{\mu\nu}^y & D_{\mu\nu}^z \end{pmatrix} \tag{17}$$

$$\tag{18}$$

$$\begin{aligned}
V_S(\boldsymbol{R}_C) = \sum_A \Bigg\{ & \boldsymbol{R}_C^T \Bigg[ f_5(R_{AC}) \boldsymbol{\mu}_A R_{AC}^2 - \boldsymbol{R}_{AC} 3 f_5(R_{AC}) (\boldsymbol{\mu}_A^T \boldsymbol{R}_{AC}^2) \\
& - f_3(R_{AC}) q_A \boldsymbol{R}_A C \Bigg] - f_5(R_{AC}) \boldsymbol{R}_{AC}^T \boldsymbol{\Theta}_A \boldsymbol{R}_{AC} - f_3(R_{AC}) \boldsymbol{\mu}_A^T \boldsymbol{R}_{AC} \\
& + q_A f_5(R_{AC}) \frac{1}{2} \boldsymbol{R}_C^2 \boldsymbol{R}_{AC}^2 - \frac{3}{2} q_A f_5(R_{AC}) \sum_{\alpha\beta} \alpha_{AB} \beta_{AB} \alpha_C \beta_C \Bigg\} \\
& + 2 f_{XC}^{\mu_C} \boldsymbol{R}_C^T \boldsymbol{\mu}_C - f_{XC}^{\Theta_C} \boldsymbol{R}_C^T \Bigg[ 3\boldsymbol{\Theta}_C - \mathrm{Tr}(\boldsymbol{\Theta}_C) \boldsymbol{I} \Bigg] \boldsymbol{R}_C
\end{aligned} \tag{19}$$

$$\begin{aligned}
V_D(\boldsymbol{R}_C) = \sum_A \Bigg[ & \boldsymbol{R}_{AC} 3 f_5(R_{AC}) (\boldsymbol{\mu}_A^T \boldsymbol{R}_{AC}) - f_5(R_{AC}) \boldsymbol{\mu}_A R_{AC}^2 + f_3(R_{AC}) q_A \boldsymbol{R}_{AC} \\
& - q_A f_5(R_{AC}) \boldsymbol{R}_C R_{AC}^2 + 3 q_A f_5(R_{AC}) \boldsymbol{R}_{AC} \sum_\alpha \alpha_C \alpha_{AC} \Bigg] \\
& - 2 f_{XC}^{\mu_C} \boldsymbol{\mu}_C - 2 f_{XC}^{\Theta_C} \Bigg[ 3\boldsymbol{\Theta}_C - \mathrm{Tr}(\boldsymbol{\Theta}_C) \boldsymbol{I} \Bigg] \boldsymbol{R}_C
\end{aligned} \tag{20}$$

$$\begin{aligned}
V_Q^{\alpha\beta}(\boldsymbol{R}_C) &= - \sum_A q_A f_5(R_{AC}) \left[ \frac{3}{2} \alpha_{AC} \beta_{AC} - \frac{1}{2} R_{AB}^2 \right] \\
&- f_{XC}^{\Theta_C} \left[ 3 \boldsymbol{\Theta}_C^{\alpha\beta} - \delta_{\alpha\beta} \sum_\alpha \boldsymbol{\Theta}_C^{\alpha\alpha} \right]
\end{aligned} \tag{21}$$

$\boldsymbol{\mu}_A$ is the cumulative atomic dipole moment of atom A and $\boldsymbol{\Theta}_A$ is the corresponding traceless quadrupole moment. Traceless simply means that the sum of the diagonal elements is 0. The curly braces and brackets are used in the same way as normal parenthesis for showing order of operations. $q_A$ is the atomic charge of atom A.

$$\Theta_A^{\alpha\beta} = \frac{3}{2}\theta_A^{\alpha\beta} - \frac{\delta_{\alpha\beta}}{2}\left(\theta_A^{xx} + \theta_A^{yy} + \theta_A^{zz}\right) \tag{22}$$

$$\theta_A^{\alpha\beta} = \sum_{l'\in A}\sum_{l} P_l\left(\alpha_A D_{ll'}^{\beta} + \beta_A D_{ll'}^{\alpha} - \alpha_A\beta_A S_{ll'} - Q_{ll'}^{\alpha\beta}\right) \tag{23}$$

$$q_A = Z_A - GAP_A \tag{24}$$

$$\mu_A^{\alpha} = \sum_{l'\in A}\sum_{l} P_{l'l}\left(\alpha_A S_{l'l} - D_{l'l}^{\alpha}\right) \tag{25}$$

$$D_{ll'}^{\alpha} = \langle\phi_l|\alpha_i|\phi_{l'}\rangle = \alpha_i S_{ll'} \tag{26}$$

$$Q_{ll'}^{\alpha\beta} = \langle\phi_l|\alpha_i\beta_i|\phi_{l'}\rangle = \beta_i D_{ll'}^{\alpha} \tag{27}$$

$\alpha$ and $\beta$ are Cartesian components labled $(x, y, z)^T$ with atom A being centered in $\boldsymbol{R}_A = (x_i, y_i, z_i)^T$ where i is a form of pointer/label dereferencing. $\delta_{\alpha\beta}$ is just the delta function, i.e is is 1 if $\alpha$ and $\beta$ are the same label and 0 otherwise, this serves to include the term only for the diagonal.

$$\boldsymbol{\Theta}_A = \begin{pmatrix} \Theta_A^{xx} & \Theta_A^{xy} & \Theta_A^{xz} \\ \Theta_A^{yx} & \Theta_A^{yy} & \Theta_A^{yz} \\ \Theta_A^{zx} & \Theta_A^{zy} & \Theta_A^{zz} \end{pmatrix} \tag{28}$$

$$\boldsymbol{\mu}_A = \begin{pmatrix} \mu_A^x & \mu_A^y & \mu_A^z \end{pmatrix}^T \tag{29}$$

$$\boldsymbol{R}_{AB} = \boldsymbol{R}_A - \boldsymbol{R}_B \tag{30}$$

$$R_{AB} = \sqrt{(\boldsymbol{R}_{AB}^x)^2 + (\boldsymbol{R}_{AB}^y)^2 + (\boldsymbol{R}_{AB}^z)^2} \tag{31}$$

$$f_n(R_{AB}) = \frac{f_{damp}(a_n, R_{AB})}{R_{AB}^n} = \frac{1}{R_{AB}^n}\frac{1}{1 + 6\left(\frac{R_0^{AB}}{R_{AB}}\right)^{a_n}} \tag{32}$$

$$R_0^{AB} = 0.5(R_0^{A'} + R_0^{B'}) \tag{33}$$

$$R_0^{A'} = \begin{cases} R_0^A + \frac{R_{max} - R_0^A}{1 + exp[-4(CN_A' - N_{val} - \Delta_{val})]} & \text{if } N_{val} \text{ is given} \\ 5.0 \text{ bohrs} & \text{otherwise} \end{cases} \tag{34}$$

$$R_{max} = 5.0 \text{ bohrs} \tag{35}$$

$$\Delta_{val} = 1.2 \tag{36}$$

$R_0^A$ is a fitted value for 12 elements and 5.0 for the rest. $a_n$ are adjusted global parameters. Where $f_{XC}^{\mu_A}$ and $f_{XC}^{\Theta_A}$ are fitted values.

## 2.3 Dispersion contribution

$$F_{\mu\nu}^{D4} = -\frac{1}{2}S_{\mu\nu}(d_A + d_B), \forall\mu \in A, \nu \in B \tag{37}$$

$$\begin{aligned} d_A = &\sum_r^{N_{A,ref}} \frac{\partial\xi_A^r(q_A, q_{A,r})}{\partial q_A} \sum_B \sum_s^{N_{B,ref}} \sum_{n=6,8} \\ &W_A^r(CN_{cov}^A, CN_{cov}^{A,r}) W_B^s(CN_{cov}^B, CN_{cov}^{B,s}) \xi_B^s(q_B, q_{B,s}) \times \\ &s_n \frac{C_n^{AB,ref}}{R_{AB}^n} f_n^{damp,BJ}(R_{AB}) \end{aligned} \tag{38}$$

The dispersion coefficient for two reference atoms $C_n^{AB,\mathrm{ref}}$ is evaluated at the reference points, i.e., for $q_A = q_r$, $q_B = q_s$, $CN_{\mathrm{cov}}^A = CN_{\mathrm{cov}}^r$, and $CN_{\mathrm{cov}}^B = CN_{\mathrm{cov}}^s$.

The Gaussian weighting for each reference system is given by:

$$W_A^r(CN_{cov}^A, CN_{cov}^{A,r}) = \sum_{j=1}^{N_{gauss}} \frac{1}{\mathcal{N}} \exp\left[-6j \cdot (CN_{cov}^A - CN_{cov}^{A,r})^2\right] \tag{39}$$

with

$$\sum_r^{N_{A,ref}} W_A^r(CN_{cov}^A, CN_{cov}^{A,r}) = 1 \tag{40}$$

$\mathcal{N}$ is a normalization constant.
The number of Gaussian function per reference system $N_{gauss}$ is mostly one, but equal to three for $CN_{cov}^{A,r} = 0$ and reference systems with similar coordination number.

$C_6^{AB}$ is the pairwise dipole-dipole dispersion coefficients calculated by numerical integration via the Casimir-Polder relation.

$$C_6^{AB} = \frac{3}{\pi} \sum_j w_j \overline{\alpha}_A(i\omega_j, q_A, CN_{cov}^A) \overline{\alpha}_B(i\omega_j, q_B, CN_{cov}^B) \tag{41}$$

$w_j$ are the integration weights, which are derived from a trapeziodal partitioning between the grid points $j (j \in [1, 23])$.
The isotropically averaged, dynamic dipole-dipole polarizabilites $\overline{\alpha}$ at the $j$th imaginary frequency $i\omega_j$ are obtained from the self-consistent D4 model; i.e., they are depending on the covalent coordination number and are also charge dependent.

$$\overline{\alpha}_A(i\omega_j, q_A, CN_{cov}^A) = \sum_r^{N_{A,ref}} \xi_A^r(q_A, q_{A,r}) \overline{\alpha}_{A,r}(i\omega_j, q_{A,r}, CN_{cov}^{A,r}) W_A^r(CN_{cov}^A, CN_{cov}^{A,r}) \tag{42}$$

The charge-dependency is included via the empirical scaling function $\xi_A^r$.

$$\xi_A^r(q_A, q_{A,r}) = \exp\left[3\left\{1 - \exp\left[4\eta_A\left(1 - \frac{Z_A^{eff} + q_{A,r}}{Z_A^{eff} + q_A}\right)\right]\right\}\right] \tag{43}$$

where $\eta_A$ is the chemical hardness taken from ref 98.
$Z_A^{eff}$ is the effective nuclear charge of atom A.

$C_8^{AB}$ is calculated recursively from the lowest order $C_6^{AB}$ coefficients.

$$C_8^{AB} = 3C_6^{AB}\sqrt{\mathcal{Q}^A\mathcal{Q}^B} \tag{44}$$

$$\mathcal{Q}^A = s_{42}\sqrt{Z^A}\frac{\langle r^4\rangle^A}{\langle r^2\rangle^A} \tag{45}$$

$\sqrt{Z^A}$ is the ad hoc nuclear charge dependent factor.

From the original xTB program we can see that $s_{42}$ is 0.5, and $Z^A$ is the atomic number of A.

$$\sqrt{0.5\left(\frac{r^4}{r^2}\sqrt{Z^A}\right)} \tag{46}$$

$\langle r^4\rangle$ and $\langle r^2\rangle$ are simple multipole-type expectation values derived from atomic densities which are averaged geometrically to get the pair coefficients.

$CN_{cov}^A$ is the covalent coordination number for atom A.

$q$ is the atomic charge, so $q_A$ is the atomic charge for atom A.

The scaling parameters in the dispersion model are:

$$a1 = 0.52 \quad | \quad a2 = 5.0 \quad | \quad s6 = 1.0 \quad | \quad s8 = 2.7$$

BJ = Becke-Johnson

$$f_n^{damp,BJ}(R_{AB}) = \frac{R_{AB}^n}{R_{AB}^n + (a_1 \times R_{AB}^{crit} + a_2)^6} \tag{47}$$

$$R_{AB}^{crit} = \sqrt{\frac{C_8^{AB}}{C_6^{AB}}} \tag{48}$$

$$f_9^{damp,zero}(R_{AB}, R_{AC}, R_{BC}) = \left(1 + 6\left(\sqrt{\frac{R_{AB}^{crit}R_{BC}^{crit}R_{CA}^{crit}}{R_{AB}R_{BC}R_{CA}}}\right)^{16}\right)^{-1} \tag{49}$$

# 3 Total Energy for GFN2-xTB

$$\begin{aligned}E_{GFN2-xTB} &= E_{rep}^{(0)} + E_{disp}^{(0,1,2)} + E_{EHT}^{(1)} + E_{IES+IXC}^{(2)} + E_{AES+AXC}^{(2)} + E_{IES+IXC}^{(3)} \\ &= E_{rep} + E_{disp}^{D4'} + E_{EHT} + E_{\gamma} + E_{AES} + E_{AXC} + E_{\Gamma}^{GFN2} \end{aligned} \tag{50}$$

## 3.1 Repulsion Energy

$$E_{rep} = \frac{1}{2}\sum_{A,B} \frac{Z_A^{eff} Z_B^{eff}}{R_{AB}} e^{-\sqrt{a_A a_B}(R_{AB})^{(k_f)}} \tag{51}$$

$$k_f = \begin{cases} 1 & if A, B \in \{\mathrm{H}, \mathrm{He}\} \\ \frac{3}{2} & otherwise \end{cases} \tag{52}$$

$Z^{eff}$ and $a$ are variables fitted for each element. A,B are the labels of atoms. Since we only have C and H in our systems we can simplify this quite a bit in code. $R_{AB}$ is the distance between the A and B atoms.

## 3.2 Extended Hückel Theory Energy

$$E_{EHT} = \sum_{\mu\nu} P_{\mu\nu} H_{\mu\nu}^{EHT} \tag{53}$$

$$P_{\mu\nu} = P_{\mu\nu}^0 + \delta P_{\mu\nu} \tag{54}$$

$$P^0 = \sum_A P_A^0 \tag{55}$$

$$\delta P_{\mu\nu} = ?? \quad \text{comes from the iteration, can be skipped for now} \tag{56}$$

Where $P_A^0$ is the neutral atomic reference density of A. This is known as Superposition of Atomic Densities or SAD.

## 3.3 Isotropic electrostatic and Exchange-correlation energy

### 3.3.1 Second order

$$E_{\gamma} = \frac{1}{2}\sum_{A,B}^{N_{atoms}} \sum_{l \in A} \sum_{l' \in B} q_{A,l} q_{B,l'} \gamma_{AB,ll'} \tag{57}$$

### 3.3.2 Third order

$$E_{\Gamma}^{GFN2} = \frac{1}{3}\sum_{A}^{N_{atoms}} \sum_{l \in A} (q_{A,l})^3 \Gamma_{A,l} \tag{58}$$

### 3.4 Anisotropic electrostatic energy

$$\begin{aligned} E_{AES} &= E_{q\mu} + E_{q\Theta} + E_{\mu\mu} \\ &= \frac{1}{2}\sum_{A,B}\{f_3(R_{AB})[q_A(\boldsymbol{\mu}_B^T\boldsymbol{R}_{BA}) + q_B(\boldsymbol{\mu}_A^T\boldsymbol{R}_{AB})] \\ &\quad + f_5(R_{AB})[q_A\boldsymbol{R}_{AB}^T\boldsymbol{\Theta}_B\boldsymbol{R}_{AB} + q_B\boldsymbol{R}_{AB}^T\boldsymbol{\Theta}_A\boldsymbol{R}_{AB} \\ &\quad - 3(\boldsymbol{\mu}_A^T\boldsymbol{R}_{AB})(\boldsymbol{\mu}_B^T\boldsymbol{R}_{AB}) + (\boldsymbol{\mu}_A^T\boldsymbol{\mu}_B)R_{AB}^2]\} \end{aligned} \tag{59}$$

### 3.5 Anisotropic XC energy

$$E_{AXC} = \sum_A (f_{XC}^{\mu_A}|\boldsymbol{\mu}_A|^2 + f_{XC}^{\Theta_A}||\boldsymbol{\Theta}_A||^2) \tag{60}$$

What norms are these?

### 3.6 Dispersion Energy

$$E_{disp}^{D4'} = -\sum_{A>B}\sum_{n=6,8} s_n \frac{C_n^{AB}(q_A, CN_{cov}^A, q_B, CN_{cov}^B)}{R_{AB}^n} f_{damp,BJ}^{(n)}(R_{AB})$$
$$- s_9 \sum_{A>B>C} \frac{(3cos(\theta_{ABC})cos(\theta_{BCA})cos(\theta_{CAB})+1)C_9^{ABC}(CN_{cov}^A, CN_{cov}^B, CN_{cov}^C)}{(R_{AB}R_{AC}R_{BC})^3} \quad (61)$$
$$\times f_{damp,zero}^{(9)}(R_{AB}, R_{AC}, R_{BC}).$$

The term in the second line is the three-body Axilrod– Teller–Muto (ATM) (What is this??????) term and the last line is the corresponding zero-damping function for this term.

The damping and scaling parameters in the dispersion model are:

$$s6 = 1.0 \quad | \quad s8 = 2.7 \quad | \quad s9 = 5.0$$

$C_9^{ABC}$ is the triple-dipole constant[1]:

$$C_9^{ABC} = \frac{3}{\pi}\int_0^\infty \alpha^A(i\omega)\alpha^B(i\omega)\alpha^C(i\omega)d\omega \tag{62}$$

The three-body contribution is typically $< 5-10\%$ of $E_{disp}$, so it is small enough that we can reasonably approximate the coefficients by a geometric mean as[1]:

$$C_9^{ABC} \approx -\sqrt{C_6^{AB}C_6^{AC}C_6^{BC}} \tag{63}$$

$\theta_{ABC}$ is the angle between the two edges going from B to the other two atoms. $\theta_{BCA}$ is the angle between the edges going from C to the other two and so on.

### 3.7 SAD - Superposition of Atomic Densities

The superposition of atomic densities(SAD) is an approach to obtain a good approximation of a collection of atoms, to be used as an initial guess for solving the self-consistent field(SCF) equation.

As originally implemented in DISCO, the molecular electron density can be obtained by adding the densities of all the constituting atoms.

This is how we get the density matrix for an isolated atom? equation 15 from: (https://sci-hub.box/10.1002/jcc.540030314)

$$D_{ij} = \sum_a^{occ} c_{ia}c_{ja} \tag{64}$$

[1]https://www.researchgate.net/publication/43347348_A_Consistent_and_Accurate_Ab_Initio_Parametrization_of_Density_F D_for_the_94_Elements_H-Pu

To get the coefficients we need to solve SCF for each atom? this is supposedly cheap, but idk how to do it. (https://sci-hub.box/10.1002/jcc.20393) Though the math for Direct SCF Approach is given in this paper at equation 10: (https://sci-hub.box/10.1002/jcc.540030314). This is probably how.

The SAD method is then the sum of all of these?

Equation 2 in the GFN2 paper talks about "superposition of (neutral) atomic reference densities". Is this relevant?

Direct SCF Approach

$$\begin{aligned}\Delta F_{ab} =&(c_{ia}c_{jb} + c_{ja}c_{ib})\\ &\Delta F_{ij} + (c_{ia}c_{kb} + c_{ka}c_{ib})\\ &\Delta F_{ik} + (c_{ia}c_{lb} + c_{la}c_{ib})\\ &\Delta F_{il} + (c_{ja}c_{kb} + c_{ka}c_{jb})\\ &\Delta F_{jk} + (c_{ja}c_{lb} + c_{la}c_{jb})\\ &\Delta F_{jl} + (c_{ka}c_{lb} + c_{la}c_{kb})\Delta F_{kl}\\ &= l_{ijkl}(4E^{ab}_{ij}D_{kl} + 4D_{ij}E^{ab}_{kl} - E^{ab}_{ik}D_{jl} - D_{ik}E^{ab}_{jl} - E^{ab}_{il}D_{jk} - D_{il}E^{ab}_{jk})\end{aligned} \tag{65}$$

where

$$E^{ab}_{ij} = c_{ia}c_{jb} + c_{ja}c_{ib} \tag{66}$$

Equation 18 from (https://sci-hub.box/https://doi.org/10.1021/acs.chemrev.5b00584) uses $\rho_0$ which is the superposition of neutral atom densities:

$$\rho_0 = \sum_A \rho_0^A \tag{67}$$