

# Constraints

## 1. Chiral Indices (Coming soon)

The total and local n-linear indices, as defined in their “standard” form in the **QuBiLS-MIDAS** module not codify information about the chiral properties of the molecular structure. In order to solve this problem, we introduced a *trigonometric 3D-chirality correction factor* in molecular vector  $\bar{x}$ . In this sense, a chirality molecular vector is obtained ( $^* \bar{x}$ ), where the components of  $\bar{x}$  (for instance, Pauling electronegativity ( $x_A$ ) of the atom A) are substituted by the following term [ $x_A + \sin ((\omega_A+4\Delta)\pi/2)$ ].

The *trigonometric 3D-chirality correction factor* uses a dummy variable  $\omega_A$ , and an integer parameter  $\Delta$ , where:

$\omega_A = 1$  and  $\Delta$  is an odd number when A has R (rectus), E (entgegen), or *a* (axial) notation according to Cahn-Ingold-Prelog rules

$\omega_A = 0$  and  $\Delta$  is an even number, if A does not have 3D specific environment

$\omega_A = -1$  and  $\Delta$  is an odd number when A has S (sinister), Z (zusammen), or *e* (equatorial) notation according to Cahn-Ingold-Prelog rules

Thus, this 3D-chirality factor  $\sin ((\omega_A+4\Delta)\pi/2)$  takes different values in order to codify specific stereo-chemical information such as chirality, Z/E isomerism, and so on. This factor therefore takes values in the following order  $1 > 0 > -1$  for atoms that have specific 3D environments. The chemical idea here is not that the attraction of electrons by an atom depends on their chirality, because experience shows that chirality does not change the electronegativities of atoms in the molecule over an isotropic environment in an observable way. A very interesting point is that the present 3D-chiral (2.5) descriptor reduces to simple n-linear indices for molecules without specific 3D characteristics because  $\sin (0+4\Delta)\pi/2 = 0$ , being  $\Delta$  zero or any even number.

## 2. N-tuples

For **Duplex option (2)**, the index calculations are performed over vertex pairs  $i$  and  $j$ . Here, the  $k^{\text{th}}$  *spatial (dis-)similarity matrix* of the geometric molecular structure,  $\mathbb{G}^k$ , is used like matrix forms. For duplex relation, *bilinear*, *linear* and *quadratic indices* can be obtained. However, a generalization of these indices can be obtained using *n-linear maps*. That is to say, for **3-uples**

**option**, the index calculations are performed over vertices  $i$ ,  $j$  and  $k$  using the  $k^{\text{th}}$  *three-tuples spatial (dis-)similarity matrix*,  $\mathbb{G}\mathbb{T}^k$ ; while for **4-uples option** the index calculations are performed over vertices  $i$ ,  $j$ ,  $k$  and  $l$  employing the  $k^{\text{th}}$  *four-tuples spatial (dis-)similarity matrix*,  $\mathbb{G}\mathbb{Q}^k$ . For ternary and quaternary matrices, previously mentioned, several matrices types are used, such as:

- **3C**: All the elements of the 3-tuples matrix, with coordinates  $(i, j, k)$ , are considered.
- **3nC**: Only elements of the triple matrix that satisfy the condition that all the 3 coordinates are different are selected.
- **4C**: All the elements of the quadruple matrix, with coordinates  $(i, j, k, z)$ , are considered.
- **4nC**: Only elements of the quadruple matrix that satisfy the condition that all the 4 coordinates are different are considered