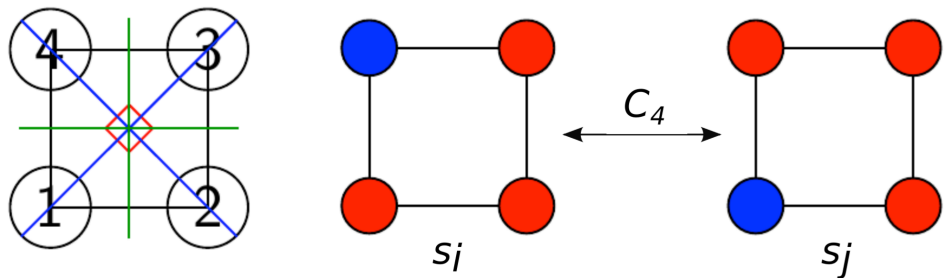## The symmetry-adapted Monte Carlo

The symmetry-adapted Monte Carlo (SA-MC) takes advantage of symmetry to perform a uniform at random sampling of the configurational space (CS) corresponding to a given cell and a given composition of a solid solution or disordered crystalline compound. The original formulation was developed by Dixon and Wilf (1983) to find independent graphs. In the following, we will briefly outline the ideas sustaining the SA-MC. To do so, we must switch from the classical mineralogical perspective on symmetry (in the Euclidian space) to a combinatorial one. When possible, we will make links to examples that should help the reader.

Elements (operations) of symmetry (point or space) groups permute points in such a way that the set of points remains unchanged. The points that are exchanged are symmetry equivalent. Consider the set S of mappings of a set D of points on a set R of elements. D contains $|D|$ points, R contains $|R|$ elements (e.g., atoms, colors, etc.). A natural consequence of the existence of a symmetry group G acting on D is that the $|R|^{|D|}$ possible mappings in S are split in symmetry independent classes (SICs) of configurations. Mappings s belonging to one SIC share the same symmetry (not the same symmetry elements). The group H of each s is a sub-group of G. These s can be exchanged by a symmetry operation $\{g|g \in G, g \notin H\}$. Recall that symmetry groups act on every type of set. For example, the action of the group on itself produces conjugacy classes (CC). The group action on S is the Pòlyà's action defined by:

$$(g \cdot s)(x) = s(g^{-1}x) \qquad \begin{cases} g \in G \\ x \in D \end{cases}$$

Let us give an example. Consider four points $D = \{1,2,3,4\}$ that are equivalent by the C_{4v} group and a mapping from this set on a set of two colors, $R = \{r, b\}$. Then, 16 (2^4) different mappings exist, among which $s_i = (r_1, r_2, r_3, b_4)$ and $s_j = (r_1, b_2, r_3, r_4)$, see the figure below.

Because the points are equivalent, it is $s_i \equiv s_j$. The two configurations are related by the operation C_4, so they belong to the same SIC. The number of SICs can be calculated

considering the symmetry group as isomorphous to a group of permutation (Cayley's theorem). In our case, the 4-fold axis permutes $1 \to 2, 2 \to 3, 3 \to 4$ and $4 \to 1$. The 2-fold axis, instead, moves $1 \to 3, 3 \to 1$ and $2 \to 4, 4 \to 2$. We say that the 4-fold axis partitions D in one cycle of length 4, while the 2-fold axis splits D in 2 cycles of length 2. The number of cycles of an operation g acting on D, $|Cyc_{D(g)}|$, gives the number of SICs, $|\Delta(S)|$, through Pòlyà's formula:

$$|\Delta(S)| = \frac{1}{|G|} \sum_{g \in G} |R|^{|Cyc_{D(g)}|}$$

The number of configurations (multiplicity) in a SIC is $|G|/|H|$ (Lagrange's theorem). Note that this formula is equivalent to the one providing the number of orientation states for twin crystals (Hahn and Klapper, 2010). Therefore, the lower the symmetry, the larger the number of configurations, which implies that a direct sampling of the CS favors low symmetry classes. To overcome this problem, a change in the probability of reaching different SICs is required.

Operations belonging to the same CC share the same cycle structure (number and length of the cycles). Then, Pòlyà's formula can be factorized by CC:

$$|\Delta(S)| = \frac{1}{|G|} \sum_{j=1}^{|C|} |C_j| |R|^{|Cyc_{D(g_j)}|}$$

where $|C|$ is the number of CC and g_j is a representative of C_j. From this expression it follows:

$$\sum_{j=1}^{|C|} \frac{|C_j| |R|^{|Cyc_{D(g_j)}|}}{|\Delta(S)||G|} = 1$$

which defines a probability distribution on the set of CC, $C = \{C_1 \dots C_{|C|}\}$:

$$Prob(C_j) = \frac{|C_j| |R|^{|Cyc_{D(g_j)}|}}{|\Delta(S)||G|}, \qquad j = 1 \dots |C|$$

Choosing a conjugacy class C_j with the probability $Prob(C_j)$, picking a $g \in C_j$, and constructing a configuration in which every element d of a cycle of g is mapped on the same element $r \in R$, a configuration whose group H contains g is obtained. In this process, the probability of reaching a given SIC depends on C_j, so it is bayesian. As shown by Dixon and Wilf (1983), by repeating this process, each SIC is obtained with the same probability $1/|\Delta(S)|$, independently of its symmetry.

Dixon, J.D. and Wilf, H.S. (1983) The random selection of unlabeled graphs, Journal of Algorithms, 4, 205–13

Hahn, Th. and Klapper, H. (2010) Twinning in crystals, in International Tables for Crystallography, Vol. D, Editor: Authier, A., p. 393-448, Publisher: Wiley