An evolutionary algorithm for the prediction of crystal structures

Alessandro Fadda[1,*] and Giuseppe Fadda[2,†]

[1]*Laboratoire d'Informatique de Paris 6 (LIP6), 4, place Jussieu, F-75005 Paris, France*
[2]*Dipartimento di Metodi e Modelli Matematici per le Scienze Applicate, Università degli Studi di Padova,*
Via Trieste, 63, I-35121 Padova, Italy
and Dipartimento di Fisica, Università degli Studi di Cagliari–Cittadella Universitaria,
S.P. Monserrato-Sestu km 0.700, I-09042 Monserrato (Cagliari), Italy

An evolutionary programming algorithm based on the exploration of the space of the conjugacy classes of crystal structures through symmetry-preserving and symmetry-breaking mutations is described. Global minima as well as saddle points can be found. Two examples of application of the algorithm are given.

I. INTRODUCTION

With the development of computing power and the parallel progress of theoretical and numerical methods in materials science, it has become possible to conduct computer experiments on solids with unprecedented accuracy when compared to experiment. It is but a small logical step to use these numerical tools to study the properties of new crystalline solids, so far not encountered in nature; finding them is the objective of "crystal structure prediction."

In the last 10 years, some research avenues for such predictions have been explored, most notably through data mining,[1,2] cluster expansion methods,[3,4] simulated annealing,[5] and evolutionary algorithms (EAs),[6–8] such as USPEX;[9–12] see Ref. [13] for a recent review. The authors of USPEX have obtained notable successes in predicting new and interesting structures, such as a high-pressure phase for boron[14] and two potentially ultrahard phases of the WN_2 compound.[15] However, in all the cited works, the search space was not unambiguously defined, and the intrinsic symmetry of crystals was not used to its fullest extent.

In a series of papers,[16–19] the concept of conjugacy classes, which was used in crystallography to derive the 14 Bravais lattices,[20,21] has been extended to general crystals, i.e., Bravais lattices plus a polyatomic basis.[22,23] This has led to the classification of all possible crystalline structures into conjugacy classes, and thus to a precise definition of the search space for new structures in terms of crystalline symmetry.

The remaining ingredient is a way to explore this space; evolutionary algorithms[24–26] are especially well suited for that task. In the present case, individuals are crystals regrouped into a population which evolves according to a Darwinian paradigm. The present approach is a fully predictive method,[13] as the only required input is the chemical composition.

The paper is organized as follows: we first present some basic concepts on evolutionary programming (EP) (Sec. II) and conjugacy classes (Sec. III), then describe the algorithm (Sec. IV). Finally, in Sec. V, two examples of application are given.

II. EVOLUTIONARY ALGORITHMS

EAs are stochastic algorithms inspired by the natural evolution of a population. From the point of view of optimiza-tion, these methods are able to find a global optimum and can be used when standard approaches are not applicable.

The problem of evolutionary optimization is to find the minimum (or maximum) of the fitness function f, which is a function from the search space S onto $\mathbb{R}$. f reflects how well a point of S (also called an individual) solves the task at hand: in other words, its degree of adaptation. A set of s points of S, called a population, is randomly created, all fitnesses are evaluated and then the evolution takes place in terms of a succession of generations: selection, recombination, mutation, fitness evaluation, and replacement phases are applied.

Selection. s individuals are chosen to generate the offsprings. Many selection processes can be used, either deterministic or stochastic but always based on the fitness of the individuals. Depending on the scheme used, some individuals can be selected more than once and a new population is created.

Evolution operators. For each selected individual, a stochastic evolution operator is applied according to some user-supplied probabilities. The purpose of mutations and crossovers is to explore the search space by applying small variations or by exchanging parts. This, with the selection mechanism, enables generation after generation to find better and better individuals.

Evaluation. The fitnesses of the new individuals (offsprings) are computed.

Replacement. The new population for the next generation is chosen among the parents and the offsprings.

The present work is based on an EP algorithm,[24] one of the major possible evolutionary approaches; the others are genetic algorithms[25,27] (GAs), evolution strategies,[26,28,29] and genetic programming.[30] The motivations of this choice are given in Sec. IV below.

EP differs from the other algorithms in some ways and different points of view can be adopted. EP works directly on the representation of an individual as a whole and there is no mapping from some genotype space (representation space) to the phenotype space (behavioral space) where the fitness is evaluated. Evolution in EP can be seen at the level of the species and not at the level of the individuals of that species. EPs generally do not see in an individual any building blocks which could be exchanged with other individuals;[31] thus no crossover operator is present in EPs, in opposition to GAs

which derive their strength mainly from this operator according to the schema theorem.[25,27] Another characteristic of EPs is the mutation operator: the severity of mutations applied to an individual is defined as a function of its fitness and is often reduced as the global optimum is approached. On the contrary, mutation rate in GA does not depend on the individual to which the mutation is applied and is generally low. Finally, the EP replacement scheme is *elitist*, i.e., the global optimum of the fitness cannot decrease.

III. CRYSTALLOGRAPHIC CONJUGACY CLASSES

By definition, a *conjugacy class* $\mathcal{C}$ in a group G is a set of subgroups H_i of G such that for any pair H_j, H_k belonging to $\mathcal{C}$, there exists $g \in G$ which maps any element h_j of H_j into an element h_k of H_k: $h_k = g \cdot h_j \cdot g^{-1}$, for all $h_k \in H_k$. H_j and H_k are then said to be *conjugate* groups in G. Conjugacy classes are equivalence classes in the mathematical sense.

Let $\mathcal{L}$ be a three-dimensional simple lattice, that is an infinite three-dimensional periodic set of points such that any two points M and M' of $\mathcal{L}$ are related by a translation

$$\overrightarrow{MM'} = \sum_{i=1}^{3} n_i \vec{a}_i, \quad n_i \in \mathbb{Z},$$

the $\vec{a}_i$ forming the lattice vectors. These vectors can always be taken to form a primitive basis, i.e., a basis which generates a cell with the smallest possible volume.[22,23] If L is the set of all the linear symmetry operations leaving $\mathcal{L}$ invariant and expressed in the primitive basis $\vec{a}_i$, it can be proved that L is a finite subgroup of the group $GL(3,\mathbb{Z})$ of invertible integral matrices. Furthermore, the various groups L belong to 14 conjugacy classes of $GL(3,\mathbb{Z})$ (also called arithmetic classes): these correspond to the 14 Bravais lattices. Further details can be found for instance in Refs. [21], [32], and [33].

In Refs. [16–19], this concept has been generalized to all crystalline structures; these can be seen equivalently as the union of a simple lattice plus a polyatomic basis composed of n atoms in the primitive cell,[22,23] or as the union of n identical Bravais lattices $\mathcal{L}_i$ each occupied by one atom of the polyatomic basis.[19] Special subgroups $\Gamma(2+n,\mathbb{Z})$ of the groups $GL(2+n,\mathbb{Z})$ are considered instead of $GL(3,\mathbb{Z})$, and special finite subgroups of $\Gamma(2+n,\mathbb{Z})$, called *lattice groups*, form conjugacy classes; any crystal structure can be mapped into one of these lattice groups. In the following, two crystals are said to be equivalent if their lattice groups are conjugate in $\Gamma(2+n,\mathbb{Z})$.

For instance, to the diamond structure corresponds a conjugacy class in $\Gamma(4,\mathbb{Z})$; the finite group of 48 elements generated by the following matrices:[34]

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ \bar{1} & \bar{1} & \bar{1} & \bar{1} \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 1 & 0 \\ \bar{1} & \bar{1} & \bar{1} & \bar{1} \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

$$\begin{pmatrix} \bar{1} & 0 & 0 & 0 \\ 0 & \bar{1} & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & \bar{1} \end{pmatrix}, \quad \begin{pmatrix} \bar{1} & 0 & 0 & 0 \\ 0 & \bar{1} & 0 & 0 \\ 0 & 0 & \bar{1} & 0 \\ 0 & 0 & 0 & \bar{1} \end{pmatrix}$$

is an element of the "diamond" class. It should be noted that this group is isomorphic to the point group $m\bar{3}m$ of the diamond structure, and that it is free of metric considerations, i.e., diamond C, Si, and Ge all have this group (or a conjugate one, see below) as lattice group.

It can be proved that lattice groups form a more precise classification than space groups;[19,35] in the latter reference, the lattice groups of the hexagonal diamond, hexagonal graphite A9, and α-La A3′ structures have been given explicitly and shown not to be conjugate groups, even though the space group is $P6_3/mmc$ in the three cases.

For crystals with a single crystallographic orbit (i.e., one Wyckoff position[21]), lattice groups and lattice complexes[36,37] are conjectured to give the same number of conjugacy classes:[35] diamond falls in this case. However, the generalization of lattice complexes to crystals having several Wyckoff positions (multiregular point systems, see Refs. [38] and [39]) is problematic, see Ref. [35] for a detailed discussion.

In practice, the description of a structure by its space group and Wyckoff positions fixes unambiguously the conjugacy class. Indeed, such a description determines a lattice group H_1 of this class and all the other elements $H_2, \ldots, H_n$ of the same class can be deduced by conjugacy. Conversely however a conjugacy class may correspond to several equivalent structures having one or more of the groups H_i as lattice groups and different but related Wyckoff positions. For instance, a crystal with space group $Fd\bar{3}m$ and Wyckoff position $8a$ defines the diamond conjugacy class but vice versa this class corresponds to structures with either position $8a$ or position $8b$ of space group $Fd\bar{3}m$: indeed, they both describe the same diamond crystal, one being shifted from the other by $\frac{1}{2} \frac{1}{2} \frac{1}{2}$.

The importance of lattice-group classification stems from its ability to define a whole class of equivalent crystalline structures *without ambiguities*, as opposed to space-group classification. As seen in the examples, space groups alone cannot be used to discriminate structures, and merely adding the set of Wyckoff positions is not sufficient either: there may exist a number $n > 1$ of equivalent crystals in a given conjugacy class, n generally increasing with the crystal complexity,[18] but such equivalences cannot be determined easily (if at all) without lattice groups.

IV. ALGORITHM

A point of the search space S is a crystalline structure, fully defined by its conjugacy class, its chemical composition, and its structural parameters, i.e., those of the cell and of the crystallographic orbits.

The only required input is the chemical formula of the crystal, for instance ZrO_2 or Al_2O_3. The individuals of the initial population are determined randomly among all pos-

sible structures having this formula; they may or may not belong to the same conjugacy classes.

The choice of one of the evolutionary algorithms depends on S. In our representation space, no semantically meaningful crossover operator can be designed: we can exchange parts between compatible crystals only but this has not the same meaning as and is not the crossover defined in the schema theorem.[25,27,31] Moreover, it is clear that the best way to explore S is to modify in some manner the cell and/or the Wyckoff orbits. The strength of such mutations is basically a function of the crystal itself, through its fitness, the internal forces, and/or the pressure of the crystal at hand. For these reasons, EP is the natural approach.

Furthermore, the discussion on crossover when it is not clearly defined has been going on for a long time in the EA community;[40,41] this is however not the scope of the present paper. This "crossover" could act as a large-step mutation and its efficiency might be highly problem-dependent.[42] It may be interesting to use such operator and compare the results.

Our algorithm follows the general EP scheme outlined in Sec. II: (1) the population is randomly created; (2) the fitnesses are evaluated; (3) based on these fitnesses, the selection scheme is applied. Different choices of selection schemes are possible, even using no selection schemes at all: in this case, every individual in the population generates one offspring; (4) for each selected individual, a mutation is chosen among all the possible ones according to some user-supplied probabilities and is applied immediately. This produces the offspring; (5) the fitnesses of the offsprings are evaluated; (6) the population of the next generation is chosen among the parents and the offsprings; and (7) back to step 3.

A. Mutations

We recall that the purpose of the mutations is to explore the search space S by applying small variations. Indeed, such changes must be fine-grained to correctly explore S.

We define two types of mutations: (1) *symmetry-preserving* mutations modify the cell or internal parameters of a crystal but leaves its conjugacy class unchanged: the symmetry of the crystal is preserved and (2) *symmetry-breaking* mutations alter the conjugacy class through either a change in one or several Wyckoff positions or by applying a phase transition leading to a subgroup or supergroup.

1. Symmetry-preserving mutations

To this category belong all mutations changing the cell or the internal parameters or both but not the conjugacy class of the crystal. These mutations are used to reach the equilibrium values (at the target pressure p_{target}) of the structural parameters; the information obtained on the interatomic forces and stress tensors may be used to improve the crystal.

(a) *Lattice mutation.* Depending on the crystal system and the computed pressure, this mutation changes the lattice parameters in order to obtain a crystal from the same system. For example, a cubic cell has one degree of freedom (the lattice parameter a) and only a is changed. Three strategies can be applied.

The first is to integrate the equation $K = -V \partial p / \partial V$, with $K = K_0 + K_0' p$, where K_0 is the bulk modulus and K_0' its derivative with respect to pressure (both at $p = 0$), previously known and/or calculated during the evolution using the Birch-Murnaghan[43–45] (or any other) equation of state.

The second strategy, especially adapted to anisotropic crystals, is to use Hooke's law $\sigma = C \cdot \epsilon$, where σ and ϵ are the stress and strain tensors respectively, in order to identify the elastic moduli C (actually, only the components which can be derived from symmetry-preserving deformations) and then to calculate the strain tensor necessary to obtain a pressure closer to p_{target}.

In both cases, no additional fitness computations are needed since all crystals from previous generations belonging to the same conjugacy class can be used to evaluate the elastic constants.

The third strategy consists in changing randomly the cell parameters according to the crystal system by adding a zero-mean Gaussian pertubation.

(b) *Mutation of the parameters of the Wyckoff positions.* This mutation changes the values of the parameters of some or all the crystallographic orbits. The severity of this mutation is based on the corresponding force F acting on one of the atoms of the considered orbit. We add a zero-mean Gaussian pertubation with a vector deviation based on F to adjust the atomic position and complete the changes to the other atoms belonging to the same orbit by symmetry. One could see this mutation as some kind of "stochastic molecular-dynamics step."

2. Symmetry-breaking mutations

In order to explore the space of conjugacy classes, a second type of mutation, which changes the symmetry of the crystal, is required. These mutations can either leave the number of atoms unchanged or can change it by an integral factor. To the first case belong (i) mutations which changes the Wyckoff positions compatibly with the number of atoms and (ii) phase transitions at the Γ point; transitions at the other points belong to the second case.

An example of mutation (i) is given in Sec. V B, Fig. 4: Al orbits $2a$ and $2b$ (for a total of four atoms) are changed into orbit $4c$ while O orbit $6d$ is changed into $6e$. Examples of the other mutations are found in Sec. V, Figs. 2 and 4.

Presently, phase transitions are done through a chain of maximal subgroups or minimal supergroups[46–49] and are chosen randomly depending on the directed graph $G = (V, E)$ comprising a set V of vertices—the space groups—together with a set E of edges—the phase transitions available to these space groups. Such a random path in the graph G is a Markov chain: the probability distribution for the system at the next step only depends on the current state of the system.

B. Initial population

The choice of the initial population is *a priori* made randomly among all possible structures having a given chemical formula; as noted above, they may or may not belong to the same conjugacy classes. However a bias can be introduced.

Simple crystals generally belong to the cubic and hexagonal systems but not always: for instance gallium and uranium are both orthorhombic at room conditions. Furthermore, as phase transitions are used to explore the search space, it might be more efficient to start from high-symmetry structures, as more subgroups are then available; however, such choice is not mandatory, as reverse transitions (from subgroup to supergroup) are also available, enabling to climb the "symmetry ladder" nevertheless.

C. Fitness function

The fitness function reflects how well an individual solves the optimization problem. For crystal structure prediction the thermodynamic potential, i.e., the free enthalpy, is the natural choice of fitness in order to determine the ground-state polymorph at given conditions of temperature and pressure. Each evaluation of the fitness function consists exactly, for first-principles computations, in a single self-consistent Kohn-Sham cycle[50] or for empirical models, in a single calculation of the potential energy.

D. Motivations and discussion

In this section, S represents the search space. The space S, as it is based on the definition of a crystal by its chemical composition, conjugacy class, and structural parameters, satisfies the elementary prerequisites of a search space, that is any point in it defines a crystal uniquely. This is the basic motivation for the use of conjugacy classes of lattice groups. Therefore, once equilibrium has been reached for a given crystal, it has been reached for all the equivalent crystals as well. Consequently, this representation allows to classify previously explored regions of S. Such information can be used during the evolution to prevent further exploration of these regions, as equilibrium has already been reached there, while allowing to visit them again, as potentially interesting symmetry-breaking mutations might still be possible.

All the mutations (except for the change of orbits) have a direct physical meaning, such as deformations and relaxations of internal parameters or phase transitions; this may prove to be an advantage for a deeper understanding of the relationship between the various phases found during the evolution.

The algorithm is capable of reaching saddle points, which have no special status in our approach; they cannot be reached by unconstrained (i.e., non symmetry-preserving) relaxation however, see Ref. [13] for a discussion. An example, cubic zirconia, is given in Sec. [V]; this illustrates the importance of such unstable solutions, as cubic zirconia has found many applications when doped, see for instance Refs. [51] and [52] and references therein.

Finally, we do not (but may in the future) use deterministic methods especially regarding the relaxation of the cell and internal parameters. Indeed, we started first to design a simple yet not simplistic canonical EP algorithm; the results in the next section show that this approach is feasible. The exploration of the search space S is done with the mutations; they are applied following a given set of probabilities and the evolution of the whole population is only guided by these probabilities and by the selection/replacement mechanism. For example, in our evolution scheme, a symmetry-breaking mutation can occur before an equilibrium value is reached as illustrated in Figs. [3] and [4] below: during the evolution of the best solution, two transitions were applied on a nonrelaxed structure, and this proved to be an important move to reach corundum. A last remark regards the computational point of view: a deterministic relaxation using n iteration steps to converge is equivalent to n single evaluations of the fitness function.

V. TEST RESULTS

We have implemented the preceding algorithm in a computer software. The present section describes two well-known examples used as testbeds for the program.

The enthalpy has been computed with QUANTUM-ESPRESSO (Ref. [53]) (ZrO_2) and GULP (Refs. [54]–[57]) (Al_2O_3). The irreducible representations of space groups are labeled according to Miller and Love.[58,59]
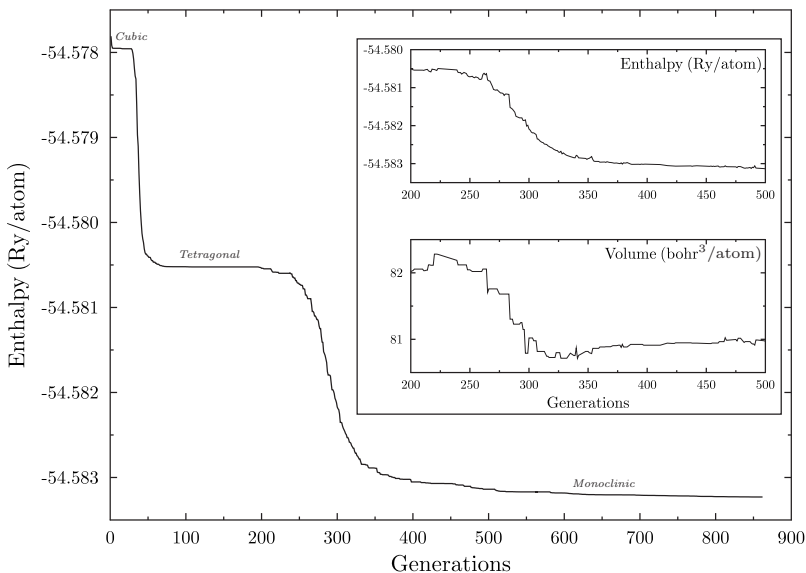

FIG. 1. (Main panel) Evolution of the absolute minimum of enthalpy with the number of generations for zirconia. Notice the alternance of rapid evolution (large variations in enthalpy) with periods of exploration (reduced variation in enthalpy). The plateaus from left to right correspond to the cubic, tetragonal, and monoclinic phases, respectively. (Inset) Evolution of the enthalpy (upper panel) and volume (lower panel) of the best individual at the current generation.
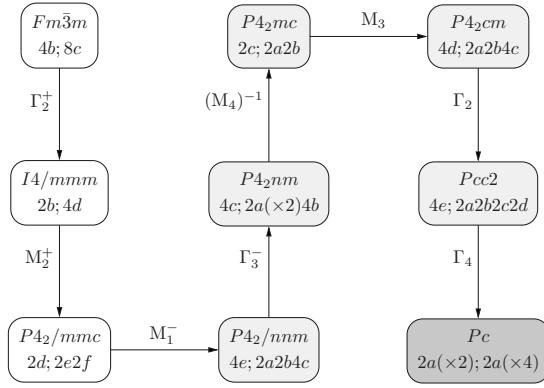
FIG. 2. Computed chain of mutations leading from the cubic $Fm\bar{3}m$ to the monoclinic Pc; the mutations are all phase transitions. The space group and the Wyckoff positions (Zr before the semicolon then O) are shown as well as the irreducible representations responsible of the phase transitions; a $^{-1}$ superscript refers to a reverse transition, i.e., from subgroup to supergroup. The background color distinguishes the families of structures, see text: white corresponds to the cubic $Fm\bar{3}m$, light gray to the tetragonal $P4_2/nmc$, and dark gray to the monoclinic $P2_1/c$.

A. Zirconia

Zirconia (zirconium dioxide ZrO_2) exists under five known forms, with three at ambient pressure: the monoclinic $P2_1/c$ (baddeleyite), stable up to 1200–1500 K,[60] the tetragonal $P4_2/nmc$ stable up to 2650 K, and the cubic $Fm\bar{3}m$ (fluorite) stable above; details can be found for instance in Refs. 51, 52, and 61. Although these phases are all in a group-subgroup relationship, the monoclinic polymorph is quite difficult to reach by transitions through maximal subgroups, one of the intermediate phases being unstable,[62] and it has many degrees of freedom (four for the cell and nine for the Wyckoff positions).

Vanderbilt ultrasoft potentials were used, based on scalar-relativistic computations and on the Perdew-Burke-Ernzerhof[63] parametrization of the generalized-

gradient approximation (GGA). The chosen electronic configuration for zirconium includes the fourth shell.

The initial population is composed of cubic and hexagonal crystals. The evolution of the best fitness (that is, the lowest enthalpy observed so far during the evolution) drawn in Fig. 1 shows that three known phases of zirconia were explored. The cubic phase was found first after a few generations, followed by the tetragonal phase just before the 30th generation, and finally the monoclinic phase after generation 200.

The convergence rate to pass from one phase to the next decreases as the number of degrees of freedom increases. The convergence of the cubic phase required the optimization of one lattice parameter only and took four generations. The tetragonal phase (mainly $P4_2mc$, see below and Fig. 2) was found by optimizing five parameters (the two lengths a and c, and one parameter for each of the three Wyckoff positions) during 15–20 generations. Finally, the evolution to the monoclinic phase represented by Pc was the result of the optimization of 22 parameters (4 for the lattice parameters a, b, c, and β, and 18 for the six orbits) and took more than 300 generations to complete. However, 80% of the gain in fitness occurred during approximately 70 generations and corresponds to the drop of volume as shown in the inset of Fig. 1.

The computed chain of groups, for the run illustrated in Fig. 1, is given in Fig. 2. There are three families: the structures with the $I4/mmm$ and $P4_2/mmc$ groups are slight deformations of the cubic fluorite structure $Fm\bar{3}m$. The other structures but the last one (light gray background in Fig. 2) are all close to the $P4_2/nmc$ polymorph; for instance, the $P4_2cm$ is identical to it up to a tolerance in contravariant (i.e., crystallographic) coordinates of 1/1000. Finally, the Pc structure, which starts relatively far from the baddeleyite $P2_1/c$, evolves toward it until convergence (see Fig. 1); the forces were converged at around 10^{-4} Ry/a.u. and the pressure at 10^{-5} GPa. The cubic structure is a saddle point in the direction of atomic displacements corresponding to the X_2^- irreducible representation,[64] as was already observed in Ref. 65, Fig. 5.
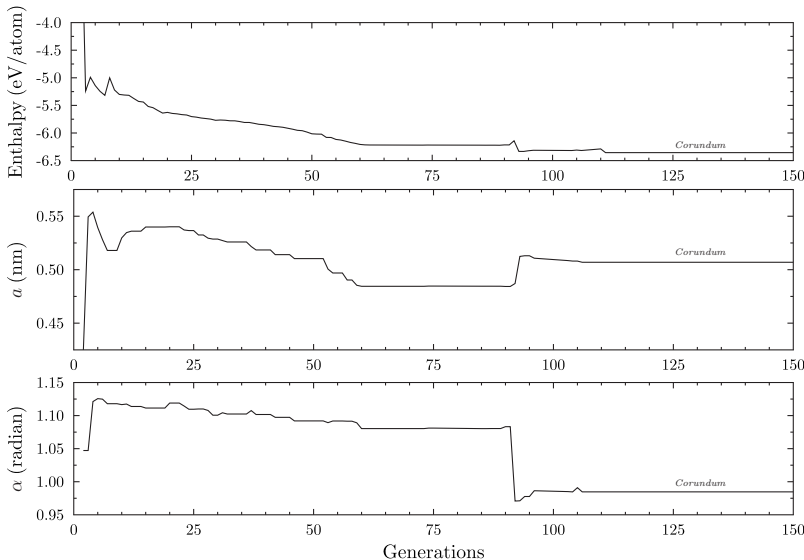

FIG. 3. Evolution with the number of generations of the enthalpy per atom (top), the lattice parameter a (middle), and the rhombohedral angle α (bottom) for alumina. The rightmost plateau corresponds to the corundum polymorph.
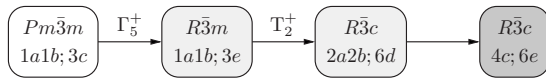
FIG. 4. Computed chain of mutations leading to the corundum structure; see key to Fig. 2. The last mutation is a change in Wyckoff positions. The white background corresponds to a cubic structure, light gray to rhombohedral crystals, and dark gray to the corundum structure.

B. Alumina

Alumina Al_2O_3 exists under a variety of forms (see for instance Refs. 66 and 67 and references therein), but the most common is corundum, with space group $R\bar{3}c$, $a=0.51280$ nm, $\alpha=0.96494$ rad, and Wyckoff positions Al ($4c$, with $x=0.35214$) and O ($6e$, $x=0.9435$) (data from Ref. 68); this crystalline structure is that of sapphire and ruby. The functional form of the interactions used for the computations is the empirical Streitz-Mintmire variable-charge potential;[69] it has been found to yield close results to first-principles calculations, even for the newly found phases.

A typical run leading to the corundum structure is illustrated in Figs. 3 and 4. This evolution started from a random population of cubic, hexagonal, and rhombohedral crystals. The corundum structure emerged at generation 110 and was the result of two transitions: first from a cubic crystal $Pm\bar{3}m$ and then a change in orbit (see Fig. 4) before a metastable structure was reached. A mutation of the rhombohedral angle α followed by the adaptation of the a cell parameter mostly completed the evolution (Fig. 3).

Another phase was also found, represented in Fig. 5 [$a=0.3022$ nm, $c=1.123$ nm, Al ($2a;2c$), O ($2b;4f$, $z=0.06411$)]; an experimentally observed polymorph[70] [$a=0.3112$ nm, $c=0.4988$ nm, Al ($2b$); O ($2c;2a$), the orbit $2a$ having a fractional occupancy of $\frac{1}{2}$] is also represented. Both have the same space group ($P6_3/mmc$) and the same ringlike structure consisting of alternate Al and O ions. The difference lies in the stacking pattern of these rings: a simple O-Al-O trimer for the observed structure, a more complex pattern reminiscent of corundum for the newly found one.

The analysis of the latter with first-principles computations (GGA ultrasoft potentials) has revealed a large difference in ground-state energy (about $+87$ meV/atom) with respect to corundum; its evaluated bulk modulus is 210 GPa,
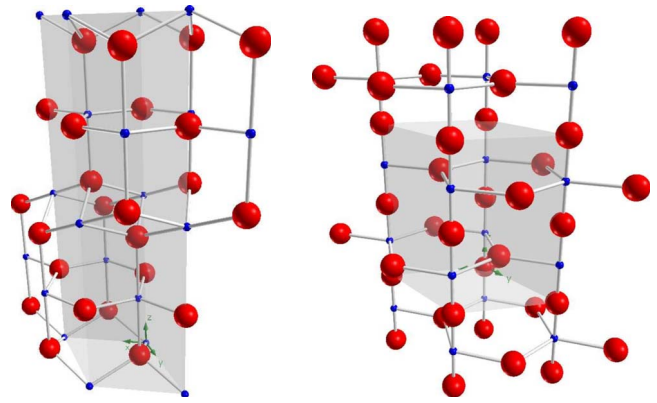

FIG. 5. (Color online) Two $P6_3/mmc$ polymorphs of alumina; the unit cell is in gray, Al^{3+} ions are in blue (black), O^{2-} ions in red (light gray), and are represented at a quarter scale. Ionic radii are taken from Ref. 71. On the left, the structure found with the EP algorithm, on the right a recently observed polymorph (Ref. 70) obtained from high-temperature reduction.

with $C_{11}=370$ and $C_{33}=380$ GPa [corundum has, respectively, 252 GPa, 497 GPa, and 499 GPa (Ref. 71)]. As corundum, this structure is an insulator with a gap estimated at 5.3 eV.

VI. CONCLUSIONS

We have presented an evolutionary programming algorithm for the prediction of crystal structures acting on the search space of crystals defined by their crystallographic conjugacy class and their cell and orbit parameters. This space is explored with physically transparent, stochastic mutations; not using unconstrained relaxation enables us to reach saddle points as well, the importance of which is illustrated by the technological relevance of cubic zirconia. The method is fully predictive, as it requires only the chemical composition as input.

A software implementing the algorithm has been written and tested in several well-known cases; we have presented two of them, and shown the capacity of the present approach to explore the search space and find the known stable structure as well as (for zirconia) the other known polymorphs, among which the unstable cubic phase.

---

[1] A. N. Kolmogorov and S. Curtarolo, Phys. Rev. B **73**, 180501 (2006).

[2] S. Curtarolo, D. Morgan, K. Persson, J. Rodgers, and G. Ceder, Phys. Rev. Lett. **91**, 135503 (2003).

[3] V. Blum and A. Zunger, Phys. Rev. B **72**, 020104 (2005).

[4] G. Trimarchi, P. Graf, and A. Zunger, Phys. Rev. B **74**, 014204 (2006).

[5] K. Doll, J. C. Schön, and M. Jansen, Phys. Rev. B **78**, 144110 (2008).

[6] N. L. Abraham and M. I. J. Probert, Phys. Rev. B **73**, 224104 (2006).

[7] N. L. Abraham and M. I. J. Probert, Phys. Rev. B **77**, 134117 (2008).

[8] G. Trimarchi and A. Zunger, Phys. Rev. B **75**, 104113 (2007).

[9] C. W. Glass, A. R. Oganov, and N. Hansen, Comput. Phys. Commun. **175**, 713 (2006).

[10] A. R. Oganov and C. W. Glass, J. Chem. Phys. **124**, 244704 (2006).

[11] A. R. Oganov, C. W. Glass, and S. Ono, Earth Planet. Sci. Lett. **241**, 95 (2006).

[12] A. R. Oganov and C. W. Glass, J. Phys.: Condens. Matter **20**, 064210 (2008).

[13] S. M. Woodley and R. Catlow, Nature Mater. **7**, 937 (2008).

[14] A. R. Oganov, J. Chen, C. Gatti, Y. Ma, Y. Ma, C. W. Glass, Z. Liu, T. Yu, O. O. Kurakevych, and V. L. Solozhenko, Nature (London) **457**, 863 (2009).

[15] H. Wang, Q. Li, Y. Li, Y. Xu, T. Cui, A. R. Oganov, and Y. Ma, Phys. Rev. B **79**, 132109 (2009).

[16] M. Pitteri, J. Elast. **14**, 175 (1984).

[17] M. Pitteri, J. Elast. **15**, 3 (1985).

[18] M. Hosoya, Bull. College of Science, Univ. of the Ryukyus **41**, 31 (1986).

[19] M. Pitteri and G. Zanzotto, Acta Crystallogr., Sect. A: Found. Crystallogr. **54**, 359 (1998).

[20] A. Bravais, J. Ec. Polytech. (Paris) **19**, 1 (1850).

[21] *International Tables for Crystallography, Volume A: Space-group symmetry*, edited by Th. Hahn (Springer, for the International Union of Crystallography, Dordrecht, 2006).

[22] N. W. Ashcroft and N. D. Mermin, *Solid-State Physics* (Saunders/Harcourt College, Fort Worth, Philadelphia, 1976).

[23] Ch. Kittel, *Introduction to Solid State Physics*, 8th ed. (Wiley, New York, 2005).

[24] L. J. Fogel, A. J. Owens, and M. J. Walsh, *Artificial Intelligence Through Simulated Evolution* (Wiley, New York, 1966).

[25] J. Holland, *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence* (University of Michigan Press, Ann Harbor, 1975).

[26] I. Rechenberg, *Evolutionsstrategie: Optimierung technischer Systeme nach Prinzipien des biologischen Evolution* (Fromman-Holzboog Verlag, Stuttgart, 1973).

[27] D. E. Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning* (Addison-Wesley, Reading, MA, 1989).

[28] H.-P. Schwefel, *Numerical Optimisation of Computer Models* (Wiley, New York, 1995).

[29] T. Bäck, *Evolutionary Algorithms in Theory and Practice* (Oxford University Press, New York, 1995).

[30] J. Koza, *Genetic Programing: On the Programing of Computers by Means of Natural Evolution* (MIT Press, Cambridge, MA, 1994).

[31] D. B. Fogel, *Evolutionary Computation: Toward a New Philosophy of Machine Intelligence*, IEEE Press Series on Computational Intelligence, 3rd ed. (Wiley-IEEE Press, Hoboken, NJ, 2005).

[32] M. Pitteri and G. Zanzotto, Acta Crystallogr., Sect. A: Found. Crystallogr. **52**, 830 (1996).

[33] M. Pitteri and G. Zanzotto, *Continuum Models for Phase Transitions and Twinning in Crystals* (Chapman and Hall, London, 2002).

[34] G. Fadda and G. Zanzotto, Int. J. Non-Linear Mech. **36**, 527 (2001).

[35] G. Fadda and G. Zanzotto, Acta Crystallogr., Sect. A: Found. Crystallogr. **57**, 492 (2001).

[36] E. Koch and W. Fischer, Acta Crystallogr., Sect. A: Found. Crystallogr. **31**, 88 (1975).

[37] W. Fischer and E. Koch, in *International Tables for Crystallography, Volume A: Space-group symmetry*, edited by Th. Hahn (Springer, for the International Union of Crystallography,

Dordrecht, 2006), Part 14.

[38] P. Engel, *Geometric Crystallography* (Kluwer Academic Publishers, Dordrecht, 1986).

[39] N. P. Dolbilin, J. C. Lagarias, and M. Senechal, Discrete Comput. Geom. **20**, 477 (1998).

[40] L. J. Eshelman, R. A. Caruana, and J. D. Schaffer, *Proceedings of the Third International Conference on Genetic Algorithms* (Morgan Kaufmann, San Francisco, 1989), pp. 10–19.

[41] D. Fogel and L. Stayton, BioSystems **32**, 171 (1994).

[42] T. Jones, *Proceedings of the Sixth International Conference on Genetic Algorithms* (Morgan Kaufmann, San Francisco, 1995), pp. 78–80.

[43] F. D. Murnaghan, Am. J. Math. **49**, 235 (1937).

[44] F. Birch, Phys. Rev. **71**, 809 (1947).

[45] F. Birch, J. Geophys. Res. **91**, 4949 (1986).

[46] *International Tables for Crystallography, Volume A1: Symmetry relations between space groups*, edited by H. Wondratschek and U. Müller (Kluwer Academic Publishers, for the International Union of Crystallography, Dordrecht, Boston, London, 2004).

[47] Bilbao crystallographic server, http://www.cryst.ehu.es

[48] M. I. Aroyo, J. M. Perez-Mato, C. Capillas, E. Kroumova, S. Ivantchev, G. Madariaga, A. Kirov, and H. Wondratschek, Z. Kristallogr. **221**, 15 (2006).

[49] M. I. Aroyo, A. Kirov, C. Capillas, J. M. Perez-Mato, and H. Wondratschek, Acta Crystallogr., Sect. A: Found. Crystallogr. **62**, 115 (2006).

[50] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).

[51] *Zirconia Engineering Ceramics*, edited by E. H. Kisi (Trans Tech, Ütikon-Zürich, 1998).

[52] O. A. Graeve, in *Ceramic and Glass Materials*, edited by J. F. Shackleford and R. H. Doremus (Springer, New York, 2008), Chap. 10.

[53] QUANTUM-ESPRESSO is a community project for high-quality quantum-simulation software, based on density-functional theory, and coordinated by Paolo Giannozzi. See http://www.quantum-espresso.org and http://www.pwscf.org.

[54] https://www.ivec.org/gulp/

[55] J. D. Gale, J. Chem. Soc., Faraday Trans. **93**, 629 (1997).

[56] J. D. Gale, Philos. Mag. B **73**, 3 (1996).

[57] J. D. Gale and A. L. Rohl, Mol. Simul. **29**, 291 (2003).

[58] S. C. Miller and W. F. Love, *Tables of Irreducible Representations of Space Groups and Co-representations of Magnetic Space Groups* (Pruett Press, Boulder, CO, 1967).

[59] A. P. Cracknell, B. L. Davies, S. C. Miller, and W. F. Love, *General Introduction and Tables of Irreducible Representations of Space Groups*, Kronecker Product Tables Vol. 1 (IFI/Plenum, New York, 1979).

[60] X. Luo, W. Zhou, S. V. Ushakov, A. Navrotsky, and A. A. Demkov, Phys. Rev. B **80**, 134119 (2009).

[61] J. M. Léger, P. E. Tomaszewski, A. Atouf, and A. S. Pereira, Phys. Rev. B **47**, 14075 (1993).

[62] D. Simeone, G. Baldinozzi, D. Gosset, M. Dutheil, A. Bulou, and T. Hansen, Phys. Rev. B **67**, 064111 (2003).

[63] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).

[64] K. Negita, Acta Metall. **37**, 313 (1989).

[65] S. Fabris, A. T. Paxton, and M. W. Finnis, Phys. Rev. B **61**, 6617 (2000).

[66] P. N. de Aza, in *Ceramics and Composite Materials*, edited by B. M. Caruta (Nova Science, New York, 2006), Chap. 3.

[67] R. H. Doremus, in *Ceramic and Glass Materials*, edited by J. F. Shackleford and R. H. Doremus (Springer, New York, 2008), Chap. 1.

[68] M. Oetzel and G. Heger, J. Appl. Crystallogr. **32**, 799 (1999).

[69] F. H. Streitz and J. W. Mintmire, Phys. Rev. B **50**, 11996 (1994).

[70] A. J. Dan'ko, M. A. Rom, N. S. Sidelnikova, S. V. Nizhank-ovskiy, A. T. Budnikov, Yu. Grin', and K. S. Kaltaev, Crystallogr. Rep. **53**, 1112 (2008).

[71] *CRC Handbook of Chemistry and Physics*, 87th ed., edited by David R. Lide (The Chemical Rubber Company Press, Boca Raton, FL, 2006).