A Unifying View on Recombination Spaces and Abstract Convex Evolutionary Search

Marcos Diez García and Alberto Moraglio

Department of Computer Science, University of Exeter, Exeter EX4 4QF, UK

Abstract Previous work proposed to unify an algebraic theory of fitness landscapes and a geometric framework of evolutionary algorithms (EAs). One of the main goals behind this unification is to develop an analytical method that verifies if a problem's landscape belongs to certain abstract convex landscape classes, where certain recombination-based EAs (without mutation) have polynomial runtime performance. This paper advances such unification by showing that: (a) crossovers can be formally classified according to geometric or algebraic axiomatic properties; and (b) the population behaviour induced by certain crossovers in recombination-based EAs can be formalised in the geometric and algebraic theories. These results make a significant contribution to the basis of an integrated geometric-algebraic framework with which analyse recombination spaces and recombination based EAs.

Keywords: Abstract convex landscape · Abstract convex search · Convex hull closure · Geometric crossover · Recombination P-structure

1 Introduction

An open challenge in evolutionary computing is to identify (without running any experiments) if a fitness landscape class is well matched with an evolutionary algorithm (EA) class, where good performance guarantees can be provided. Previous work [1] laid the foundations for an analytical fitness landscape method that determines if a problem's landscape belongs to certain abstract convex landscape classes, where certain recombination-based EAs exhibit polynomial runtime [2]. The theories we began unifying, namely a general geometric framework (GF) of EAs [3] and an algebraic theory of fitness landscapes known as elementary landscapes (EL) [4], complement each other towards such challenge even if they originally pursue distinct aims.

On the one hand, GF naturally frames (geometric) mutation and recombination operators within the same search space structure when viewed as a metric space. Among the many benefits of this approach we find: having a clear way of comparing mutation and crossover in a fitness landscape, the formal design

and analysis of EAs across many representations (real vectors, sets, permutations, syntactic trees, etc.), and generalisations of particle swarm optimisation and genetic programming with important applications [3]. Besides, there has been a recent development of general runtime analysis in GF [2]. Particularly, geometric-crossover EAs (without mutation) are proved to find optimal solutions in polynomial time by restricting to classes of problems with certain abstract convex landscapes. However, GF lacks analytic tools to tell whether a given landscape is abstractly convex.

On the other hand, EL provides algebraic means to analyse combinatorial landscapes. Some of them, called 'elementary', have received special attention because: they satisfy desirable conditions related to smoothness (i.e. correlation between fitness values) and structure of local optima, and they are present in important NP-complete problems (e.g. travelling salesman, graph colouring, weight partitioning, etc.). However, the algebraic structures known as P-structures that model crossover are not so readily intuitive as geometric crossovers in GF [5]. Furthermore, studies on population-based search algorithms using P-structures are scarce [6], and no runtime analysis has been done in this algebraic context.

Recombination landscapes are not essentially different in both theories and can be unified, meaning that EL can be applied consistently to analyse certain landscapes arising in GF [1]. This paper extends their unification, by showing that the algebraic approach of EL admits dual formalisations of two fundamental aspects about crossovers and their behaviour in GF: classification of geometric crossovers and abstract convex search (accomplished by geometric-crossover EAs). The main contributions are:

1. Proving the existence of two broad classes of crossovers: those which are recombination P-structures, including geometric crossovers [1], and those which are neither geometric nor recombination P-structures. This justifies that a unification of GF and EL would not be an 'empty' theory (i.e. a theory encompassing all recombination operators).
2. A formalisation of abstract convex search in EL, describing the population-behaviour of recombination P-structure EAs; a class that is larger than complete geometric-crossover EAs. It also provides a characterisation of population-based EAs less restrictive than the existing one in EL.

This paper is organised as follows. Section 2 introduces recombination and abstract convex search in GF. Section 3 introduces recombination in EL and other concepts necessary to unify abstract convex search. Section 4 presents the main results. Section 5 summarises this work and suggests future research.

2 Recombination in the Geometric Framework

Moraglio proposed a general geometric framework (GF) of EAs, independent of the problem and representation of solutions, solely based on an axiomatic definition of distance across metric spaces [3]. In GF, the search space of configurations (representing candidate solutions) is defined as a metric space (X, d),

with configuration set X and metric $d : X \times X \to [0, \infty)$. Thus, a metric space constitutes the structure of a fitness landscape (X, d, f) with an arbitrary real-valued fitness function $f : X \to \mathbb{R}$. This geometric view of the search space allows mutation[1] and crossover operators to be conceived within the same space structure, by means of metric balls and metric segments respectively.

Definition 1 (Metric ball and segment [7]). Let (X, d) be any metric space. A *closed ball* centred at point $x \in X$ with radius $r \in \mathbb{R}_{\geq 0}$ is defined as $B_d[x; r] := \{y \in X \mid d(x, y) \leq r\}$. A *geodesic interval* or *metric segment* is defined as $[x; y]_d := \{z \in X \mid d(x, z) + d(z, y) = d(x, y)\}$, where $x, y \in X$ are called the *extremes* of the segment and $d(x, y)$ its *length*.

Many kinds of crossovers can be modelled in GF (e.g. multi-parental or probabilistic). For notational simplicity, we assume crossovers taking two parents and returning one offspring; and, denote by $\mathrm{Im}[\xi(\cdot, \cdot)]$ the set of all offspring produced with non-zero probability by a crossover ξ, called the *image set*.

Definition 2 ([Complete] geometric crossover [3]). Let (X, d) be any metric space. A crossover operator ξ is a *geometric crossover* for (X, d), if all offspring belong to the parents' metric segment. That is, $\forall x, y \in X : \mathrm{Im}[\xi(x, y)] \subseteq [x; y]_d$. If $\forall x, y \in X : \mathrm{Im}[\xi(x, y)] = [x; y]_d$, then it is called *complete geometric crossover.*

Example 1 (Uniform geometric crossover). Let a Hamming metric space $(\mathcal{H}_q^n, d_{\mathrm{H}})$ be defined by a set $\mathcal{H}_q^n := \{0, 1, \ldots, q-1\}^n$ of n-dimensional Hamming sequences on a q-ary alphabet for $q \geq 2$, and the Hamming metric d_{H}. Traditional uniform crossover, denoted $\mathrm{UNIFORM}(x, y)$ for any parents $x, y \in \mathcal{H}_q^n$, is an example of a crossover that is geometric and also complete. For instance, consider parents $100, 001 \in \mathcal{H}_2^3$. Then, we have $\mathrm{Im}[\mathrm{UNIFORM}(100, 001)] = [100; 001]_{d_{\mathrm{H}}}$. The same holds for other pair of parents as well [3].

Example 2 (Non-complete geometric crossover). Consider the traditional one-point crossover on Hamming sequences, but returning only a single offspring with its 'head' coming from the first parent and its 'tail' from the second parent: $\mathrm{ONEPOINT}(i) : \mathcal{H}_q^n \times \mathcal{H}_q^n \to \mathcal{H}_q^n, (x, y) \mapsto z = (x_1, \ldots, x_{i-1}, y_i, \ldots, y_n)$; where z is offspring of parents x and y, at crossover point $i: 1 \leq i \leq n$. Notice for example, $\mathrm{ONEPOINT}(2)(000, 111) = 011$ is not generated by $\mathrm{ONEPOINT}(i)(111, 000)$ for any $1 \leq i \leq 3$. Therefore, it is not a complete geometric crossover, because it is asymmetric, and still is geometric since offspring belong to parents segment [3].

However, some crossovers are not geometric. That is, for any metric on a given set X, there always exist some offspring that lie outside the metric segment of some parents. See Sect. 4.1 for examples of non-geometric crossovers.

Proposition 1 (Existence of non-geometric crossovers [3,8]). The set of non-geometric crossovers $\overline{\mathcal{GX}}$ is not empty. Therefore, recombination operators split into two proper classes: geometric $\mathcal{GX}$ and non-geometric $\overline{\mathcal{GX}}$. The class of complete geometric crossovers, $\mathcal{GX}$-complete, is a proper subclass of $\mathcal{GX}$.

---

[1] Geometric mutation is defined by requiring that the mutated offspring belongs to the d-metric ball of its single parent.

To prove that a crossover is geometric, it suffices to find a distance where Definition 2 holds. Proving non-geometricity is harder, as we would have to test arbitrarily many distances before reaching a conclusion, which is infeasible. Fortunately, we can circumvent this obstacle using the axiomatic inbreeding properties fulfilled by all geometric crossovers regardless of the metric space [8]: if we find a single counterexample (i.e. instance of a metric space) where any of these properties fails, then such crossover is automatically not geometric for any metric. Proposition 2 states one of the inbreeding properties that we use later for our results in Sect. 4.

Proposition 2 (Convergence property of geom. crossovers [8]). If χ is a geometric crossover and z is a child of x and y, then the recombination $\chi(x, z)$ cannot produce the other parent y, unless $z = y$.

Achieving generality in a theory may come at the expense of depth, only being able to assert trivial or tautological statements. Proposition 1 justifies that this is not the case in GF, because it is not a theory of all crossovers, rather it restricts to geometric crossovers only. Moreover, this crossover class is relevant, since many crossovers across different representations are geometric, and polynomial runtime guarantees are possible for geometric-crossover EAs in problem classes with certain abstract convex fitness landscapes [3,2]. Essentially, the abstract convex nature of landscapes is what enables good performance of geometric-crossover EAs.

2.1 Abstract Convexity

To formalise the abstract convex search performed by geometric-crossover EAs (Sect. 2.2), a more general notion of convexity than that of traditional Euclidean spaces is required. A natural choice would be geodesic convexity, since GF understands recombination in terms of geodesic intervals.

Definition 3 (Geodesic convex space [7]). Let (X, d) be a metric space. A point $z \in X$ is *geodesically between* two points $x, y \in X$, if and only if $z \in [x; y]_d$. Then, a subset $C \subseteq X$ is *geodesically convex* if and only if all points geodesically between any two points x and y in C are also in C, that is $\forall x, y \in C : [x; y]_d \subseteq C$. The family $\mathcal{C}$ of all geodesically convex subsets C on X is called the *geodesic convexity* of (X, d), and the pair $(X, \mathcal{C})$ a *geodesically convex space.*

Geodesic convexity is not equivalent to convexity (Definition 4); the latter being more abstract, where metrics are undefined and the concepts of neighbourhood or nearness are stated in terms of (convex hull) closure operators in a set-theoretical fashion. We explain next why geodesic convexity is problematic in our context.

Definition 4 (Convexity space [7]). A family $\mathcal{C}$ of subsets of a set X is called a *convexity* on X if: (I) the empty set $\emptyset$ and the universal set X are in $\mathcal{C}$; (II) $\mathcal{C}$ is stable under arbitrary intersections, that is if $\mathcal{D} \subseteq \mathcal{C}$, then $\bigcap \mathcal{D} \in \mathcal{C}$; and, (III) $\mathcal{C}$ is stable under arbitrary nested unions, that is if $\mathcal{D} \subseteq \mathcal{C}$ is totally ordered

by inclusion, then $\bigcup \mathcal{D} \in \mathcal{C}$. The pair $(X, \mathcal{C})$ is called *convexity space*; and the members of $\mathcal{C}$ are called *convex sets*.

Definition 5 (Convex hull closure [7]). Let $(X, \mathcal{C})$ be a convexity space. The *convex hull closure* of a set A, $co(A) := \bigcap \{C \mid A \subseteq C \in \mathcal{C}\}$, is the intersection of all convex supersets $C \in \mathcal{C}$, or equally the smallest convex set, that includes A.

Abstract convex search relies on convexity of sets, and precisely the problem with geodesic convexity is that it cannot guarantee across representations (i.e. metric spaces) that a population set would be a convex set as in Definition 4. This is because metric segments do not share the convexity properties of convex hulls. In other words, metric segments need not be convex hulls: $[x; y]_d \subseteq co(\{x, y\})$ for all $x, y \in X$ [7]. Convexity spaces where equality holds (i.e. geodesic convex sets equal the convex hull of their extreme points) are called *convex geometries* [9].

2.2 Abstract Convex Search

GF postulates that the core behaviour of (recombination-based) EAs can be described axiomatically across representations and problems, and, ultimately, the differences between EAs stem from the representation of solutions and adequacy of search operators [3,10,2]. To support this claim, GF proved that the main operations of geometric-crossover EAs (i.e. selection, geometric crossover and replacement) produce new individuals that always belong to the convex hull of their ancestors; regardless of the representation, problem, geometric crossover, offspring probability distribution imposed on the geometric crossover, selection and replacement strategies [3,10]. This relationship between parents and offspring via convex hulls is the key property characterising the population behaviour of any geometric-crossover EA through generations (Fig. 1).

Proposition 3 (Abstract convex search [3]). Let (X, d) be a metric space and $P_t \subseteq X$ a population at generation t. For any EA doing the cycle selection-geometric crossover-replacement: $co(P_0) \supseteq co(P_1) \supseteq \cdots \supseteq co(P_t) \supseteq co(P_{t+1})$.
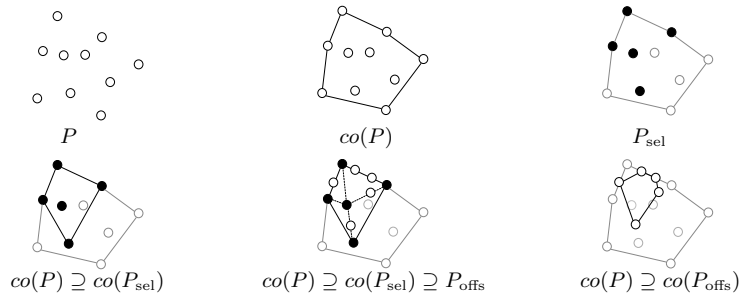

Fig. 1. Abstract convex search in two-dimensional Euclidean space: relationship between parent P, mating pool after selection P_{sel} and offspring P_{offs} populations, and their respective convex hulls $co(P)$, $co(P_{sel})$ and $co(P_{offs})$.

3 Recombination in Elementary Landscapes Theory

Stadler proposed an algebraic theory of fitness landscapes known as elementary landscapes (EL) [4]. If the geometric approach of GF (Sect. 2) allows us to seamlessly understand mutation and recombination structures in the same space, in EL the situation is more complicated. Originally, EL defines the search space structure $(X, \mathcal{N})$ in terms of (mutation) neighbourhoods $\mathcal{N} : X \to X^k$ mapping an individual to some $k > 0$ possible neighbours (e.g. bit-flip mutation). This naturally leads to formalise the search space as a graph, with the structure given by the adjacency matrix. To model recombination, EL uses an algebraic notion of interval (Definition 6) that captures general recombination structures, not just genetic crossovers [11,12].

Definition 6 (Recombination P-structure [5]). Let X be a non-empty finite set of configurations, $\mathcal{P}(X)$ its power set and an operator $\mathcal{R} : X \times X \to \mathcal{P}(X)$ mapping a pair of parents into a set of possible offspring. We call $(X, \mathcal{R})$ a *recombination P-structure* if $\forall x, y, z \in X$: (I) fix-point: $\mathcal{R}(x, x) = \{x\}$; (II) symmetry: $\mathcal{R}(x, y) = \mathcal{R}(y, x)$; (III) null-recombination: $\{x, y\} \subseteq \mathcal{R}(x, y)$; and (IV) size–monotonicity: if $z \in \mathcal{R}(x, y)$, then $|\mathcal{R}(x, z)| \leq |\mathcal{R}(x, y)|$.

Example 3 (Uniform recombination P-structure). Let parents $x, y \in \mathcal{H}_q^n$ be arbitrary Hamming sequences. Then, $\mathcal{R}_\Omega(x, y) := \{(z_1, \ldots, z_n) \mid z_i \in \{x_i, y_i\}, 1 \leq i \leq n\}$ is a recombination P-structure [11]. For instance, consider $x := 100$ and $y := 001$ in $\mathcal{H}_2^3$, then we have offspring: $\mathcal{R}_\Omega(100, 001) = \{100, 001, 000, 101\}$.

Example 4 (One-point recombination P-structure). Let parents $x, y \in \mathcal{H}_q^n$ be arbitrary Hamming sequences. Recall the single-offspring one-point crossover $\text{OnePoint}(i) : \mathcal{H}_q^n \times \mathcal{H}_q^n \to \mathcal{H}_q^n$, $(x, y) \mapsto z = (x_1, \ldots, x_{i-1}, y_i, \ldots, y_n)$, $1 \leq i \leq n$, from Example 2. Then, the one-point recombination P-structure is $\mathcal{R}_1(x, y) := \bigcup_{1 \leq i \leq n} \{\text{OnePoint}(i)(x, y), \text{OnePoint}(i)(y, x)\}$ [12]. For example, $\mathcal{R}_1(000, 111) = \bigcup_{1 \leq i \leq 3} \{\text{OnePoint}(i)(000, 111), \text{OnePoint}(i)(111, 000)\} = \{0, 1\}^3 \setminus \{010, 101\}$.

However, recombination, acting on pairs of parents, does not have a natural interpretation in graphs. Instead, [12,5] propose hypegraphs (Definition 7), where vertices remain as in graphs but edges are generalised to represent the offspring sets produced by recombining pairs of parents. Figure 2 shows an example for three-dimensional binary sequences and the uniform recombination P-structure $\mathcal{R}_\Omega$, other cases are difficult to visualise clearly.

Definition 7 (Hypergraph [5]). Let $(X, \mathcal{R})$ be a recombination P-structure. Its *hypergraph* $(X, \mathscr{E})$ has vertex set X and hyperedge set $\mathscr{E} := \{\mathcal{R}(x, y) \mid \mathcal{R}(x, y) \neq \emptyset \wedge x, y \in X\}$, where offspring $\mathcal{R}(x, y)$ represent a hyperedge.

Despite recombination search spaces are then formalised using hypergraphs, they are comparable (i.e. topologically equivalent) to mutation spaces on graphs under certain conditions [5].

Figure 2 also illustrates via hyperedges of $\mathcal{R}_\Omega$ hypergraph the following coincidence: offspring produced by the uniform recombination P-structure $\mathcal{R}_\Omega$ and Hamming segments are equivalent [11], and, independently, [10,13] show it for Hamming segments, convex hulls and traditional schema [14]. For instance, the schema `0**` corresponding to the hyperedge $\{000, 001, 010, 011\}$ is obtained by $\mathcal{R}_\Omega(010, 001)$, $[010; 001]_{d_H}$ or $co(\{010, 001\})$.
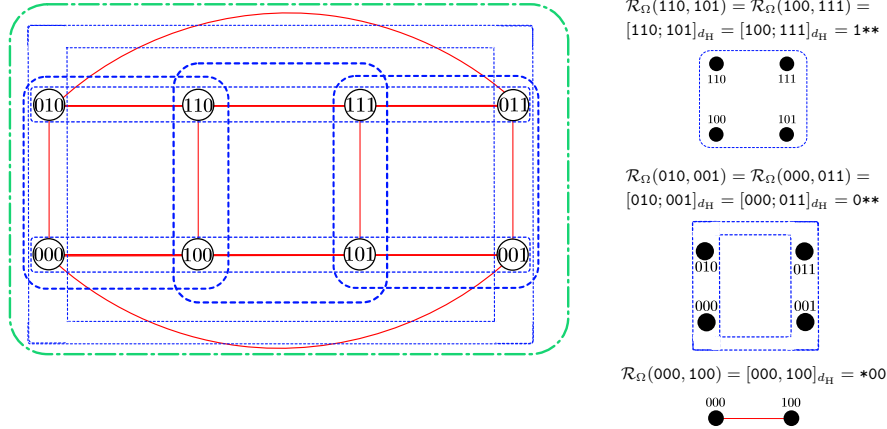

$\mathcal{R}_\Omega(110, 101) = \mathcal{R}_\Omega(100, 111) =$
$[110; 101]_{d_H} = [100; 111]_{d_H} = 1**$

$\mathcal{R}_\Omega(010, 001) = \mathcal{R}_\Omega(000, 011) =$
$[010; 001]_{d_H} = [000; 011]_{d_H} = 0**$

$\mathcal{R}_\Omega(000, 100) = [000, 100]_{d_H} = *00$

Fig. 2. Hypergraph of the uniform recombination P-structure $\mathcal{R}_\Omega$ in three-dimensional binary Hamming space: vertices are 'circles', hyperedges are both 'solid' and 'dotted-enclosing lines' (left). Enclosed vertices define hyperedges, not the 'lines shapes' or their colours. Examples of specific hyperedges, offspring of $\mathcal{R}_\Omega$ and uniform geometric crossover, and corresponding schema (right).

3.1 Crossover Random Walk

Discrete-time dynamical systems may be used to model the behaviour of recombination P-structure EAs, but this is often analytically impractical for realistic EAs [6]. Instead, the model adopted in EL is a crossover random walk, a variation on the simple random walk based on mutation moves: start with a prescribed 'father' x_0 in a population, mate it with a 'mother' y_0 randomly sampled from the population to produce a 'child' z_1, then repeat the cycle using the child as father in the next recombination (Fig. 3) [4,5,6].
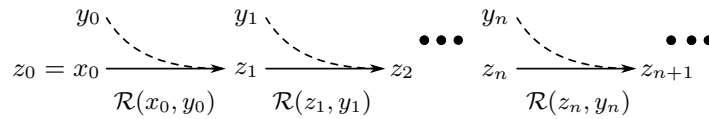

Fig. 3. Schematic representation of a crossover-walk over time.

In this manner, the sequence of fathers (or children) is a crossover random walk. The main reason [4,5,6] use this model is that the generalised adjacency matrices of hypergraphs, associated to recombination P-structures, can be

seen as transition matrices of a Markov process. That is the population search behaviour induced by recombination P-structures can be studied in terms of Markov processes. However, [6] makes two important assumptions that cause the model not to be general. First, it is assumed that there exists none or little statistical interdependence between genes (i.e. epistasis), and low selection genetic pressure; otherwise, the analysis of population behaviour is impractical. Secondly, it is formalised exclusively for traditional string crossovers (e.g. one-point or uniform), not all crossovers that can be described by general recombination P-structures. Generalising this model remains as an open problem.

3.2 Bridge to the Geometric Framework

Figure 2 hints at a close relationship between geometric crossovers and recombination P-structures. This is not coincidental. Indeed, complete geometric crossovers are also recombination P-structures (Proposition 4). Note it does not hold for general geometric crossovers, as [1] unwittingly claimed. For example, any geometric crossover where offspring never include the parents will not fulfil the null-recombination axiom (Definition 6), also the crossover in Example 2 is geometric but not a recombination P-structure due to asymmetry.

Proposition 4 ($\mathcal{GX}$-complete crossovers are recomb. P-structures [1]). Let (X, d) be any metric space with finite set X, and any crossover $\xi(x,y) :=$ $[x;y]_d$ for arbitrary parents $x, y \in X$. Then, $\forall x, y \in X$ and any metric d, (X, ξ) is a recombination P-structure.

Besides, recombination P-structures can be used to define closures that are equivalent to convex hulls [11,12]. One may attempt to define one by taking the union over all pairs of points for a given set, and collect their intervals given by a P-structure (Fig. 4b). Precisely, this is the closure $cl_{\mathcal{R}}(S) := \bigcup_{x,y \in S} \mathcal{R}(x,y)$, where $S \subseteq X$ for a general recombination P-structure $(X, \mathcal{R})$ [12]. However, it does not equal the convex hull (Fig. 4a) because, unlike convex hulls, the closure $cl_{\mathcal{R}}$ is not idempotent [15,7]. In general $cl_{\mathcal{R}}(cl_{\mathcal{R}}(S)) \neq cl_{\mathcal{R}}(S)$, compare Fig. 4c and Fig. 4b. See Fig. 5 for an example on the one-point recombination P-structure $\mathcal{R}_1$. Since $cl_{\mathcal{R}}$ is not idempotent, [12] suggests to define a recursive version, for some natural number $k \in \mathbb{N}_0$, as follows:

$$cl_{\mathcal{R}}^k(S) := \begin{cases} S & \text{if } k = 0, \\ cl_{\mathcal{R}}(S) = \bigcup_{x,y \in S} \mathcal{R}(x,y) & \text{if } k = 1, \\ cl_{\mathcal{R}}^{k-1}\left(cl_{\mathcal{R}}(S)\right) & \text{if } k \geq 2 \ . \end{cases} \tag{1}$$

The smallest k that makes $cl_{\mathcal{R}}^k$ idempotent is the *closure iteration number*[2], denoted $cin(S)$, of the subset $S \subseteq X$: $cin(S) := \min_{k \in \mathbb{N}_0} cl_{\mathcal{R}}^{k+1}(S) = cl_{\mathcal{R}}^k(S)$; and

---

[2] Certain crossover families (e.g. one-point and masked crossovers) have logarithmic upper-bounds for the closure iteration number, with increasing dimension of the search space [16]. In short, idempotency does not require excessively many iterations.

the *maximum closure iteration number* of X is $\mathrm{cin}^*(X) := \max_{S \subseteq X} \mathrm{cin}(S)$ [17].
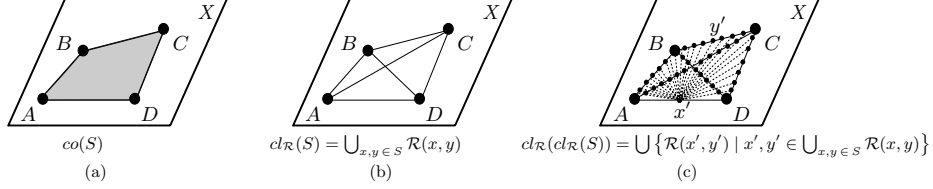

Fig. 4. Schematic comparison of (a) the convex hull closure, shaded in 'grey'; (b) recombination P-structure closure $cl_{\mathcal{R}}$, in 'solid lines'; and, (c) the twice-iterated $cl_{\mathcal{R}}^2$, in 'solid and dotted lines'. For a given subset $S := \{A, B, C, D\}$ of some space X.

Closure Iteration Numbers
(iterations until idempotency)

$$\mathrm{cin}_{\mathcal{R}_1}(\{000, 111\}) = 2$$

$$\mathrm{cin}_{\mathcal{R}_\Omega}(\{000, 111\}) = 1$$

$$cl_{\mathcal{R}_1}^1(\{000, 111\}) \cup \{010, 101\} = cl_{\mathcal{R}_1}^2(\{000, 111\}) = cl_{\mathcal{R}_\Omega}^1(\{000, 111\}) = \{0, 1\}^3$$
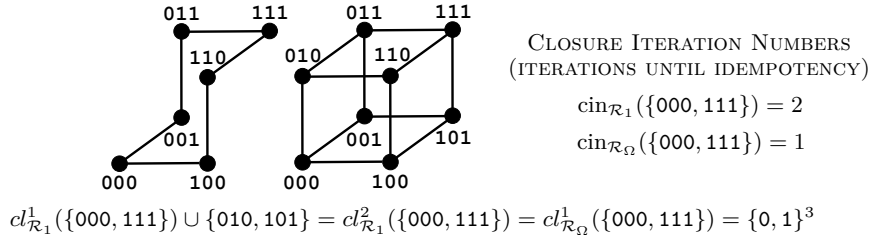
Fig. 5. Space covered on the hypercube, with set of parents $\{000, 111\}$, by: one-point recombination closure with one iteration $cl_{\mathcal{R}_1}^1$ (left), uniform recombination closure $cl_{\mathcal{R}_\Omega}^1$ and one-point twice-iterated $cl_{\mathcal{R}_1}^2$ (center). Super-indices indicate number of closure iterations, and sub-indices the recombination P-structure chosen.

However, what [12] does not explicitly notice is that $cl_{\mathcal{R}}^k$ (Eq. 1) is an equivalent definition of the convex hull closure, not just for recombination P-structures but more generally for 'relaxed' recombination P-structures satisfying only the null-recombination and symmetry axioms (also known as interval spaces).

Proposition 5 (Convex hull $\equiv$ recomb. P-structure closure [17,7]). Let $(X, \mathcal{R})$ be a recombination P-structure. Then, for any subset $S \subseteq X$ and any natural number $k \geq \mathrm{cin}(S)$: $co(S) = cl_{\mathcal{R}}^k(S)$.

4   Main Results

This section contains the two main contributions of this paper. Section 4.1 shows that the geometric versus non-geometric crossovers classification can be expanded by incorporating new recombination P-structures classes. Section 4.2 formalises abstract convex search for recombination P-structure EAs.

4.1   Axiomatic Classification of Crossovers

Proposition 1 justifies that GF is not a futile theory encompassing all crossovers, because not all of them are geometric. Since recombination P-structures may be

a larger class than complete geometric crossovers, as Proposition 4 suggests, three non-trivial relevant questions that we consider here are: (Q1) *do recombination P-structures encompass all crossovers, thus making GF-ELs unification futile*; (Q2) *is there any crossover that is a recombination P-structure but not geometric*; and, (Q3) *how recombination P-structures and geometric crossover classes compare to each other*. We tackle them in order as follows: Lemmas 1 and 2 answer question (Q1) negatively; Theorem 1 together with Lemma 5 and Theorem 2 answer question (Q2) positively; and, Theorem 3 for question (Q3).

Let us begin (Q1) by proving that Koza subtree swap crossover, which was proved to be non-geometric [8], is not a recombination P-structure.

Lemma 1. Koza subtree swap crossover is not a recombination P-structure.

Proof. It is not geometric since it does not fulfil the purity inbreeding property (i.e. recombining a parent with itself is the parent itself) [8]. Therefore, it does not fulfil the fix-point axiom of recombination P-structures $\mathcal{R}(x, x) = \{x\}$. □

Davis order crossover is non-geometric [8]. Now, we prove it is not a recombination P-structure, because it fails the symmetry axiom (Example 5); and, neither for a symmetric version (Definition 9), because it is not size-monotonic (Example 6).

Definition 8 (Davis order crossover [18]). Let individuals $x, y, z \in S_n$ be permutations of length n of a finite set S. Pick at random loci i and j such that $1 \leq i \leq j \leq n$, as indices delimiting the crossover section[3]. Then, from the first parent x (the 'cutter') copy into offspring positions $z_i, \ldots, z_j$ the crossover section $x_i, \ldots, x_j$. Finally, from left to right, fill in order the remaining positions of z with the genes of the second parent y (the 'filler') except those that have been copied. We denote this crossover, parametrised with crossover section indices i and j, by $\mathrm{DAVIS}(i, j) : S_n \times S_n \to S_n$, $(x, y) \mapsto z = \mathrm{DAVIS}(i, j)(x, y)$.

Example 5 (Davis order crossover is asymmetric). Let parents $x := 312$ (the 'cutter') and $y := 123$ (the 'filler'). Consider offspring 2$\underline{1}$3 $= \mathrm{DAVIS}(2, 2)(3\underline{1}2, 123)$, where $\underline{1}$ is the crossover section. Notice that 213 cannot be generated by $\mathrm{DAVIS}(i, j)(123, 312)$ for any crossover section indices i and j, because it is not possible to generate 2 at the first position. Either the second parent places 3 or 1 as fillers, or the first parent places 1 with the crossover section. Therefore, $\mathrm{DAVIS}(i, j)(312, 123) \neq \mathrm{DAVIS}(i, j)(123, 312)$, for $1 \leq i \leq j \leq 3$.

Example 4 demonstrated that if a crossover is asymmetric, it does not mean there are no (symmetric) recombination P-structures based on it. In fact, that example shows how symmetry is enforced on recombination P-structures for traditional genetic crossovers (e.g. two-point or uniform crossovers) [12,5]. Likewise, we may define a symmetric Davis order crossover.

---

[3] Notice that a crossover section with all parent's genes is also a valid contiguous section.

Definition 9 (Symmetric Davis order crossover). Let arbitrary parents $x, y \in S_n$ be permutations of length n of a finite set S. Then, s-DAVIS$(x, y) :=$ $\bigcup_{1 \leq i \leq j \leq n} \{$DAVIS$(i, j)(x, y),$ DAVIS$(i, j)(y, x)\}$ is what we call *symmetric Davis order crossover*; producing all possible offspring of parents x and y, over all possible crossover sections.

Example 6 (Symmetric Davis order crossover is not size-monotonic). Using the same parents as in the previous example, we have s-DAVIS$(312, 123)$ $= \{312, 123, 321, 132, 213\}$. That is $|$s-DAVIS$(312, 123)| = 5$. Recombining now parent $x := 312$ with offspring $z := 213$ we have s-DAVIS$(312, 213) = \{231\} \cup$ s-DAVIS$(312, 123)$. Therefore, $|$s-DAVIS$(312, 213)| = 6 > 5 = |$s-DAVIS$(312, 123)|$, so s-DAVIS is not size-monotonic.

Lemma 2. Neither Davis order crossover (DAVIS) nor its symmetric version (s-DAVIS) are recombination P-structures.

Proof. Examples 5 and 6 provide counterexamples respectively. □

Therefore, the answer to (Q1) is negative by Lemmas 1 and 2, because GF-EL unification is not a theory encompassing all crossovers. Regarding question (Q2), we show next two examples of crossovers that are not geometric but are recombination P-structures, at least in Hamming graphs: the all-paths crossover (Definition 10), and the intersecting-balls crossover (Definition 11). Note that these crossovers are not restricted to a particular representation, like UNIFORM crossover on strings, so one must be specified when proving whether they are geometric crossovers or recombination P-structures: choose a graph for all-paths, or a metric space for intersecting-balls.

Definition 10 (All-paths crossover [11]). Let G be a connected graph with vertex set $V(G)$. Let also $x, y, z \in V(G)$. Then, we define ALLPATHS$(x, y) := \{z \mid z$ lies on any x-y path in $G\}$, as the *all-paths crossover* that returns offspring z lying on any path between parents x and y.

Example 7 (All-paths crossover in hypercubes). Let $(\mathcal{H}_2^3, d_H)$ be the Hamming metric space of the three-dimensional hypercube graph. If parents are identical, offspring equal parents: ALLPATHS$(001, 001) = 001$. But, if the parents are different, then offspring are the whole vertex set: ALLPATHS$(000, 001) = \{0, 1\}^3$.

Lemma 3. Let $x, y \in \mathcal{H}_q^n$ be distinct vertices of a Hamming graph. Then, ALLPATHS(x, y) returns all sequences $\mathcal{H}_q^n$ as offspring.

Proof. We show constructively that all vertices in the Hamming graph are visited by at least one path between x and y. List all vertices $(v_1, v_2, \ldots, v_{q^n})$ in $\mathcal{H}_q^n$ as a Grey code cyclic sequence; where two consecutive vertices are at Hamming distance one, $d_H(v_i, v_{i+1}) = 1$, and cyclic because the same holds for the first and last vertices $d_H(v_1, v_{q^n}) = 1$. Since $x, y \in \mathcal{H}_q^n$, and $x \neq y$, then there exist distinct $v_i, v_j \in \mathcal{H}_q^n$ such that $(v_1, \ldots, v_i = x, \ldots, v_j = y, \ldots, v_{q^n})$. Then, $p_1 := (v_i = x, v_{i+1}, \ldots, v_{j-1}, v_j = y)$ and $p_2 := (v_j = y, v_{j+1}, \ldots, v_{q^n}, v_1, \ldots, v_{i-1}, v_i = x)$

are disjoint paths between x and y, because: both p_1 and p_2 start or end at x and y, each pair of consecutive vertices is an edge since they are at Hamming distance one, and are independent because the only common vertices are the start or end vertices. Therefore, any vertex in $\mathcal{H}_q^n$ is visited by p_1 or p_2. □

Theorem 1. Let $(\mathcal{H}_q^n, d_\mathrm{H})$ be the Hamming metric space of an arbitrary Hamming graph. Then, ALLPATHS(x, y) is not a geometric crossover, but it is a recombination P-structure, for any parents $x, y \in \mathcal{H}_q^n$.

Proof. Even if ALLPATHS and UNIFORM coincide in some trivial cases like $n = 1$ and $n = 2$ for $q = 2$, where the Hamming graph is isomorphic to the complete graph K_2 and cycle graph C_4 respectively, ALLPATHS cannot be geometric because the convergence property (Proposition 2) does not hold in general, as it can be deduced from Example 7. Now, let us prove that ALLPATHS is a recombination P-structure:

(I) *Fix-point*: ALLPATHS$(x, x) = \{x\}$. It holds because the only path that can begin and end at offspring x is x itself. Otherwise, cycles (i.e. loops) can be formed, and ALLPATHS only returns paths not cycles.

(II) *Symmetry*: ALLPATHS$(x, y) = $ ALLPATHS(y, x). It holds because ALLPATHS is an interval function, which by definition are symmetric [11].

(III) *Null-recombination*: $\{x, y\} \subseteq$ ALLPATHS(x, y). It holds because ALLPATHS is an interval function, and all intervals include their extremes, here parents x and y, as offspring [11].

(IV) *Size-monotonicity*: if z is offspring of ALLPATHS(x, y), then $|$ALLPATHS$(x, z)| \leq |$ALLPATHS$(x, y)|$. If $x = y$, it is trivial because there is only one offspring: $z = x = y$. Assuming $x \neq y$, we have two cases. If $z = x$, then $|$ALLPATHS$(x, x)| = 1 \leq |$ALLPATHS$(x, y)|$. For $z \neq x$, note that ALLPATHS returns offspring in all possible paths between parents and the Hamming graph is connected, so all offspring (vertices) are visited by at least one path (Lemma 3). Therefore, $|$ALLPATHS$(x, z)| = q^n \leq q^n = |$ALLPATHS$(x, y)|$, where q^n is the number of vertices of a Hamming graph. □

Definition 11 (Intersecting-balls crossover). Let (X, d) be any metric space and arbitrary parents $x, y \in X$. Then, we define the *intersecting-balls crossover* as BALLS$(d) : X \times X \to S \subseteq X$, $(x, y) \mapsto B_d[x; d(x, y)] \cap B_d[y; d(y, x)]$, parametrised with metric d.

Lemma 4 is necessary to justify that [19] used an equivalent definition of the intersecting-balls crossover, which they suggested to be geometric when in fact it is not necessarily geometric (Lemma 5).

Lemma 4. Rothlauf's crossover $R(x, y) := \{z \mid \max\{d(x, z), d(z, y)\} \leq d(x, y)\}$ is identical to BALLS$(d)(x, y)$, for any metric space (X, d) and any $x, y, z \in X$.

Proof. Follows directly: $z \in R(x, y) \iff d(x, z) \leq d(x, y)$ and $d(z, y) \leq d(x, y) \iff z \in (B_d[x; d(x, y)] \cap B_d[y; d(y, x)]) \iff z \in$ BALLS$(d)(x, y)$. □

Definition 12 (Hamming intersecting-balls crossover). Let $(\mathcal{H}_q^n, d_{\mathrm{H}})$ be any Hamming metric space and any $x, y \in \mathcal{H}_q^n$. Then, the *Hamming intersecting-balls crossover* is $\mathrm{BALLS}(d_{\mathrm{H}})(x, y) := B_{d_{\mathrm{H}}}[x; d_{\mathrm{H}}(x, y)] \cap B_{d_{\mathrm{H}}}[y; d_{\mathrm{H}}(y, x)]$.

Lemma 5. Let $(\mathcal{H}_q^n, d_{\mathrm{H}})$ be any Hamming metric space. Then, $\mathrm{BALLS}(d_{\mathrm{H}})$ is not a geometric crossover.

Proof. We prove, by counterexample, that $\mathrm{BALLS}(d_{\mathrm{H}})$ does not fulfil in all cases the convergence inbreeding property (Proposition 2). Consider parents $x := \mathtt{010}$ and $y := \mathtt{100}$ in $\mathcal{H}_2^3$. Then, we have offspring $\mathtt{001} \in \mathrm{BALLS}(d_{\mathrm{H}})(x, y)$, but $y = \mathtt{100} \in \mathrm{BALLS}(d_{\mathrm{H}})(x, \mathtt{001})$. Therefore, $\mathrm{BALLS}(d_{\mathrm{H}})$ is not geometric. $\quad\square$

Proving that $\mathrm{BALLS}(d_{\mathrm{H}})$ is a recombination P-structure, in Theorem 2, requires some extra care. Compared to the fix-point, symmetry and null-recombination axioms, size-monotonicity is more complicated and we divide its proof into three major steps. Let us briefly explain the overall idea. Size-monotonicity states: if $z \in \mathrm{BALLS}(d_{\mathrm{H}})(x, y)$, then $|\mathrm{BALLS}(d_{\mathrm{H}})(x, z)| \leq |\mathrm{BALLS}(d_{\mathrm{H}})(x, y)|$. The main difficulty is that in general $\mathrm{BALLS}(d_{\mathrm{H}})(x, z) \nsubseteq \mathrm{BALLS}(d_{\mathrm{H}})(x, y)$. However, we note that if additionally $z \in [x; y]_{d_{\mathrm{H}}}$, then inclusion does hold. This is one of the steps, proved in Lemma 6 for general metric spaces. Since, in principle, z may be anywhere in the intersection of the two balls given by $\mathrm{BALLS}(d_{\mathrm{H}})(x, y)$, the role of the other two steps is to ensure that each z can always be mapped via an automorphism to a corresponding $z' \in [x; y]_{d_{\mathrm{H}}}$; so that the size-monotonicity of $|\mathrm{BALLS}(d_{\mathrm{H}})(x, z)|$ can be proved from that of $|\mathrm{BALLS}(d_{\mathrm{H}})(x, z')|$.

Lemma 6. Let (X, d) be any metric space. Then, for all $x, y \in X$ it holds: if $z \in [x; y]_d$, then $\mathrm{BALLS}(d)(x, z) \subseteq \mathrm{BALLS}(d)(x, y)$.

Proof. Note first that if $z \in [x; y]_d$, then naturally $z \in \mathrm{BALLS}(d)(x, y)$. We need to prove that if $s \in \mathrm{BALLS}(d)(x, z)$, then $s \in \mathrm{BALLS}(d)(x, y)$. By Definition 11 we know:

$$z \in \mathrm{BALLS}(d)(x, y) \iff \big(d(x, z) \leq d(x, y)\big) \wedge \big(d(y, z) \leq d(x, y)\big) \ , \quad (2)$$

$$s \in \mathrm{BALLS}(d)(x, z) \iff \big(d(x, s) \leq d(x, z)\big) \wedge \big(d(z, s) \leq d(x, z)\big) \ . \quad (3)$$

Hence, $d(x, s) \leq d(x, z) \leq d(x, y)$, and therefore $s \in B_d[x; d(x, y)]$. To complete the proof, we need to show $s \in B_d[y; d(y, x)]$ as well. That is, $d(y, s) \leq d(x, y)$ holds. From the assumption $z \in [x; y]_d$, we know

$$d(x, y) = d(x, z) + d(z, y) \ , \quad (4)$$

and from the triangle inequality of metric spaces

$$d(y, s) \leq d(y, z) + d(z, s) \ . \quad (5)$$

Now, combining Eqs. 4 and 5 we see that

$$d(y, s) \leq d(y, z) + d(z, s) \leq d(y, z) + d(x, z) = d(x, y) \ , \quad (6)$$

because $d(z, s) \leq d(x, z)$ from Eq. 3, and symmetry of distances in metric spaces. Hence, from Eq. 6 it holds $s \in B_d[y; d(y, x)]$, so $s \in \mathrm{BALLS}(d)(x, y)$. Therefore, we conclude that if $s \in \mathrm{BALLS}(d)(x, z)$, then $s \in \mathrm{BALLS}(d)(x, y)$. $\quad\square$

Theorem 2. Let $(\mathcal{H}_q^n, d_\mathrm{H})$ be the Hamming metric space of an arbitrary Hamming graph. Then, $(\mathcal{H}_q^n, \mathrm{BALLS}(d_\mathrm{H}))$ is a recombination P-structure.

Proof. (I) *Fix-point*: $\mathrm{BALLS}(d_\mathrm{H})(x, x) = \{x\}$. In metric spaces $d_\mathrm{H}(x, y) = 0$ iff $x = y$. Therefore, $\mathrm{BALLS}(d_\mathrm{H})(x, x) = B_{d_\mathrm{H}}[x; d_\mathrm{H}(x, x)] \cap B_{d_\mathrm{H}}[x; d_\mathrm{H}(x, x)] = B_{d_\mathrm{H}}[x; 0] \cap B_{d_\mathrm{H}}[x; 0] = \{x\} \cap \{x\} = \{x\}$.

(II) *Symmetry*: $\mathrm{BALLS}(d_\mathrm{H})(x, y) = \mathrm{BALLS}(d_\mathrm{H})(y, x)$. Follows from Definition 12 and commutativity of intersection: $A \cap B = B \cap A$.

(III) *Null-recombination*: $\{x, y\} \subseteq \mathrm{BALLS}(d_\mathrm{H})(x, y)$. Since $d_\mathrm{H}(x, x) = 0 \leq d_\mathrm{H}(x, y)$, we know $x \in B_{d_\mathrm{H}}[x; d_\mathrm{H}(x, y)]$. Since $d_\mathrm{H}(x, y) = d_\mathrm{H}(x, y)$, we know as well $y \in B_{d_\mathrm{H}}[x; d_\mathrm{H}(x, y)]$. Therefore, $\{x, y\} \subseteq B_{d_\mathrm{H}}[x; d_\mathrm{H}(x, y)]$. By symmetry, $\{x, y\} \subseteq B_{d_\mathrm{H}}[y; d_\mathrm{H}(y, x)]$. So $\{x, y\} \subseteq \big(B_{d_\mathrm{H}}[x; d_\mathrm{H}(x, y)] \cap B_{d_\mathrm{H}}[y; d_\mathrm{H}(y, x)] \big) = \mathrm{BALLS}(d_\mathrm{H})(x, y)$.

(IV) *Size-monotonicity*: if z is offspring of $\mathrm{BALLS}(d_\mathrm{H})(x, y)$, then $|\mathrm{BALLS}(d_\mathrm{H})(x, z)| \leq |\mathrm{BALLS}(d_\mathrm{H})(x, y)|$. The proof has three parts:

(a) We prove that: $\forall x, y \in \mathcal{H}_q^n$, if $z \in \mathrm{BALLS}(d_\mathrm{H})(x, y)$, then $\exists z' \in [x; y]_{d_\mathrm{H}}$ such that $d_\mathrm{H}(x, z) = d_\mathrm{H}(x, z')$. Assuming that $z \in \mathrm{BALLS}(d_\mathrm{H})(x, y)$, we know $d_\mathrm{H}(x, z) \leq d_\mathrm{H}(x, y)$. Because the Hamming graph is connected and undirected, there exists at least one path between x and y, in particular shortest-paths $[x; y]_{d_\mathrm{H}}$. Because it is unweighted, for each integer distance unit between x and y there is a corresponding $z' \in [x; y]_{d_\mathrm{H}}$. Since $d_\mathrm{H}(x, z) \leq d_\mathrm{H}(x, y)$, it is always possible to find $z' \in [x; y]_{d_\mathrm{H}}$ such that $d_\mathrm{H}(x, z) = d_\mathrm{H}(x, z')$.

(b) From Lemma 6 follows that if $z' \in [x; y]_{d_\mathrm{H}}$, then $|\mathrm{BALLS}(d_\mathrm{H})(x, z')| \leq |\mathrm{BALLS}(d_\mathrm{H})(x, y)|$.

(c) We prove that: for all $x, y, z, z' \in \mathcal{H}_q^n$, if $d_\mathrm{H}(x, z) = d_\mathrm{H}(x, z')$, then $|\mathrm{BALLS}(d_\mathrm{H})(x, z)| = |\mathrm{BALLS}(d_\mathrm{H})(x, z')|$. It follows from the fact that the Hamming graph is distance-transitive [20]. This means that if $d_\mathrm{H}(x, z) = d_\mathrm{H}(x, z')$, then there exists an automorphism $\phi : \mathcal{H}_q^n \to \mathcal{H}_q^n$, where $\phi(x) = x$ and $\phi(z) = z'$. Every automorphism ϕ is an isomorphism by definition, and in graphs they are just a relabelling of vertices, therefore whatever the cardinality of $\mathrm{BALLS}(x, z)$ is, it is preserved under ϕ. That is, $|\mathrm{BALLS}(d_\mathrm{H})(x, z)| = \big| \mathrm{BALLS}(d_\mathrm{H})\big(\phi(x), \phi(z)\big) \big| = |\mathrm{BALLS}(d_\mathrm{H})(x, z')|$.

The proof size-monotonicity follows now easily. From (a) we know that for each $z \in \mathrm{BALLS}(d_\mathrm{H})(x, y)$ we can always find a $z' \in [x; y]_{d_\mathrm{H}}$, where $d_\mathrm{H}(x, z) = d_\mathrm{H}(x, z')$. But this means, using (b) and (c), that we have $|\mathrm{BALLS}(d_\mathrm{H})(x, z)| = |\mathrm{BALLS}(d_\mathrm{H})(x, z')| \leq |\mathrm{BALLS}(d_\mathrm{H})(x, y)|$. $\qquad\square$

Using the previous results we now compare recombination P-structures with geometric crossover classes (Q3).

Theorem 3 (Crossover classes). Let $\mathcal{RP}$ be the set of recombination P-structures (resp. $\overline{\mathcal{RP}}$ for crossovers that are not recombination P-structures), $\mathcal{GX}$ be the geometric crossovers, and $\mathcal{GX}$-complete be the complete geometric crossovers. It holds that: (a) $\overline{\mathcal{RP}} \neq \emptyset$; (b) $\mathcal{GX} \neq \mathcal{RP}$; (c) $\mathcal{RP} \cap \overline{\mathcal{GX}} \neq \emptyset$; (d) $\mathcal{GX}$-complete $\subset \mathcal{RP}$; and (e) $\overline{\mathcal{RP}} \subset \overline{\mathcal{GX}}$-complete.

Proof. (a) Lemmas 1 and 2 prove, respectively, that neither Koza subtree swap nor (symmetric) Davis order crossovers are recombination P-structures.

(b) Example 2 shows that the single-offspring one-point crossover (ONEPOINT) is geometric, but it is not a recombination P-structure due to its asymmetry.

(c) In Hamming graphs, Theorem 1 proves it for ALLPATHS crossover, also Lemma 5 and Theorem 2 prove it for BALLS(d_H) crossover.

(d) Follows from Proposition 4, and (c) makes the inclusion strict: complete geometric crossovers are a subclass of geometric crossovers (Proposition 1).

(e) Let $\mathcal{U}$ be the universal set of crossovers, where $\mathcal{GX}$-complete, $\mathcal{RP} \subset \mathcal{U}$. Then, using (d) and taking complements, $\overline{\mathcal{RP}} \subset \overline{\mathcal{GX}\text{-complete}}$ follows. $\square$
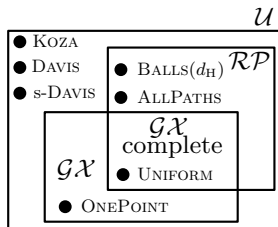
Fig. 6. Classification within the universal set $\mathcal{U}$ of crossovers by geometric crossovers $\mathcal{GX}$ and recombination P-structures $\mathcal{RP}$, with class members examples: all-paths (ALLPATHS), Hamming intersecting-balls (BALLS(d_H)), Davis order (DAVIS), symmetric Davis order (s-DAVIS), Koza subtree swap (KOZA), single-offspring one-point (ONEPOINT) and uniform (UNIFORM) crossovers.

The main purpose of Theorem 3, summarised in Fig. 6, is to start developing a unified and systematic understanding of crossovers within the geometric GF (Sect. 2) and algebraic EL (Sect. 3) theories. By relaxing the axioms of recombination P-structures or properties of geometric crossovers, new formal crossover subclasses can be incorporated, which can give us deeper insight not only into the similarities and differences in behaviour of crossovers, but also about the theories themselves. For instance, EL landscape analysis is possible when a crossover belongs to $\mathcal{GX}$-complete or $\mathcal{RP}$, but not necessarily for all those in $\overline{\mathcal{GX}\text{-complete}}$ (e.g. ONEPOINT) [1]. It should be clear that the focus of this classification is not on the practical value of crossovers (e.g. KOZA is useful, however $\mathcal{GX}$ and $\mathcal{RP}$ miss it), but what crossovers can and cannot be studied within GF and ELs.

4.2 Algebraic Abstract Convex Search

We switch our attention from recombination P-structures themselves (Sect. 4.1) to characterising the population-behaviour of EAs using them. Specifically, we prove that any recombination P-structure EA, using any selection and replacement strategies but no mutation, exhibits an abstract convex search behaviour analogue to Proposition 3.

Theorem 4 (Algebraic abstract convex search). Let $(X, \mathcal{R})$ be a recombination P-structure and $P_t \subseteq X$ the population at generation t. Then, there

exists a maximum closure iteration number $m := \text{cin}^*(X)$ such that: $cl_{\mathcal{R}}^m(P_0) \supseteq cl_{\mathcal{R}}^m(P_1) \supseteq \cdots \supseteq cl_{\mathcal{R}}^m(P_t) \supseteq cl_{\mathcal{R}}^m(P_{t+1})$; for any EA repeating the cycle: selection, recombination $\mathcal{R}$ and replacement.

Proof. We need to prove that: (a) such maximum closure iteration number exists; and, (b) the closure of recombination P-structures produces a nested chain of inclusions of population sets.

(a) From Proposition 5 we know that for any population $P_t \subseteq X$ there exists a natural number $k \geq \text{cin}(P_t)$, such that $co(P_t) = cl_{\mathcal{R}}^k(P_t)$. Since all populations P_t are subsets of X, the maximum closure iteration number $m := \text{cin}^*(X)$ exists. By definition (see below Eq. 1), $\text{cin}^*(X) \geq \text{cin}(P_t)$ for all populations P_t. Therefore, $co(P_t) = cl_{\mathcal{R}}^m(P_t)$ for all populations P_t.
(b) Since (a) proves that there exists a maximum closure iteration number such that $co(P_t) = cl_{\mathcal{R}}^m(P_t)$ for any P_t, we know also that $co(P_t) \supseteq co(P_{t+1})$ if and only if $cl_{\mathcal{R}}^m(P_t) \supseteq cl_{\mathcal{R}}^m(P_{t+1})$ for any P_t. Therefore, a chain of inclusions is formed, provided that selection and replacement can only return a subset of the convex hull of a population [10]. □

Note that Proposition 3 and Theorem 4 become equivalent for complete geometric crossovers in finite metric spaces, since they are a subclass of recombination P-structures (Theorem 3). Table 1 illustrates it for the uniform recombination P-structure $\mathcal{R}_\Omega$, which is a complete geometric crossover, whose maximum closure iteration number is one for all populations P: $co(P) = cl_{\mathcal{R}_\Omega}^1(P)$ [11].

It remains as future work to find whether 'degenerate' cases like Example 8 are the only kind of algebraic abstract convex search for crossovers in the class $\mathcal{RP} \cap \overline{\mathcal{GX}}$ (i.e. non-geometric crossovers that are recombination P-structures).

Example 8 (Abstract convex search of all-paths crossover). ALLPATHS convex search is 'degenerate' because, as long as the current population contains different parents, the vertex set of the Hamming graph is returned as offspring; until other mechanism like replacement or selection returns a population with the same individual. For instance, if $P_1 := \{001, 010, 111\}$, the following chain may be generated: $cl_{\text{ALLPATHS}}^1(P_1) = \{0, 1\}^3 \supseteq \{0, 1\}^3 \supseteq \ldots \supseteq \{0, 1\}^3 \supseteq 001$. Notice that the maximum closure iteration number of ALLPATHS is one, because ALLPATHS generates at once the whole vertex set or a singleton set of the Hamming graph, which are convex sets [17].

Theorem 4 has significant implications for GF and EL. First, it shows that the convex search of complete geometric-crossover EAs can be phrased algebraically in EL via recombination P-structures, in a way that remains consistent with GF. Secondly, that the abstract convex search behaviour of some non-geometric crossovers, for example the ALLPATHS crossover (Definition 10), can be formally justified. Finally, we now have a proper characterisation of population behaviour for recombination P-structures EAs, which is less restrictive than the previous crossover-walk model (Sect. 3.1) because it is defined for general recombination P-structures not just string crossovers, and no assumptions are made on selection
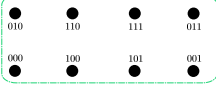
Parents population	Parents closure (and their schema)	Offspring population	Hyperedge of parents closure
$P_1 = \{001,\ 010,\ 111\}$	$cl^1_{\mathcal{R}_\Omega}(P_1)=\{0,1\}^3$ (***)	$\{010,\ 100,\ 110\}$	(graph: 010 110 111 011 / 000 100 101 001)
$P_2 = \{010,\ 100,\ 110\}$	$cl^1_{\mathcal{R}_\Omega}(P_2)=\{0,1\}^2\times\{0\}$ (**0)	$\{100,\ 000,\ 100\}$	(graph: 010 110 / 000 100)
$P_3 = \{100,\ 000,\ 100\}$	$cl^1_{\mathcal{R}_\Omega}(P_3)=\{000,100\}$ (*00)	$\{100,\ 100,\ 000\}$	(graph: 000 — 100)
$P_4 = \{100,\ 100,\ 000\}$	$cl^1_{\mathcal{R}_\Omega}(P_4)=\{000,100\}$ (*00)	$\{100,\ 100,\ 100\}$	(graph: 000 — 100)
$P_5 = \{100,\ 100,\ 100\}$	$cl^1_{\mathcal{R}_\Omega}(P_5) = 100$		(graph: 100)

Table 1. A small trace of abstract convex search for the three-dimensional binary Hamming space using uniform recombination with initial population $\{001, 010, 111\}$, population size equal three, generational replacement and no selection.

pressure or genes interdependence (epistasis). Moreover, it has a clear geometrical interpretation (Sect. 2.2) and GF has demonstrated that is analytically useful to study recombination-based EAs [2].

As a final remark, some of our results (e.g. Proposition 4 or Theorem 4) can be further generalised to more abstract algebraic spaces like interval spaces [7], by relaxing the axioms of recombination P-structures and keeping only the nullary and symmetric ones (Definition 6). It turns out that Proposition 5, which is key for Theorem 4, remains true [7]. However, our aims here have not required such degree of generality.

5 Summary and Future Work

Previous work [1] established the foundations for developing a unified approach to analyse fitness landscapes, based on an algebraic theory of fitness landscapes (Sect. 3) and a geometric framework of EAs (Sect. 2). Aiming for a solution to an open challenge in evolutionary computing: to determine a priori for a given problem if its fitness landscape matches a class of EAs where good performance can be guaranteed across representations. However, to use EL as an analytic tool, we needed to show that recombination landscapes seen from EL and GF can be unified, that is their underlying space structures are not essentially different.

This paper expands the unification between GF and EL to eventually develop an integrated GF-EL framework; where EAs can be analysed and designed across representations (GF), but also it is possible to know analytically if

good performance can be guaranteed for a given problem landscape (EL). First, we augmented the classification of crossovers in GF with algebraic crossover classes from EL (Sect. 4.1). Then, we formalised the abstract convex population behaviour of recombination P-structure EAs, as an extension of complete geometric-crossover EAs, which provides a more general characterisation of population behaviour than the existing model in EL (Sect. 4.2).

As future work, we consider the following questions. First, to augment the classification of crossovers with new classes or members and compare them. Secondly, to study the relation between abstract convex spaces (e.g. convex geometries [9]) and crossovers, by looking at how the presence or absence of convexity influences crossover behaviour. Thirdly, to export the runtime analysis of geometric-crossover EAs [2] to recombination P-structure EAs; and, conversely, use the ideas presented here to derive new insights about runtime analysis of geometric-crossover EAs. Finally, it would be interesting to integrate mutation within GF-EL unification to understand more realistic EAs using mutation and recombination operators.

References

1. García, M.D., Moraglio, A.: Bridging Elementary Landscapes and a Geometric Theory of Evolutionary Algorithms: First Steps. In: Auger, A., Fonseca, C.M., Lourenço, N., Machado, P., Paquete, L., Whitley, D. (eds.) Proceedings of the 15th International Conference on Parallel Problem Solving from Nature – PPSN XV, Lecture Notes in Computer Science, vol. 11102, pp. 194–206. Springer (2018)
2. Moraglio, A., Sudholt, D.: Principled Design and Runtime Analysis of Abstract Convex Evolutionary Search. Evolutionary Computation 25(2), 205–236 (2017), MIT Press
3. Moraglio, A.: Towards a Geometric Unification of Evolutionary Algorithms. Doctoral thesis, University of Essex, Essex, UK (November 2007)
4. Stadler, P.F.: Towards a Theory of Landscapes. In: López-Peña, R., Waelbroeck, H., Capovilla, R., García-Pelayo, R., Zertuche, F. (eds.) Complex Systems and Binary Networks, Lecture Notes in Physics, vol. 461, pp. 78–163. Springer (1995)
5. Stadler, P.F., Wagner, G.P.: Algebraic Theory of Recombination Spaces. Evolutionary Computation 5(3), 241–275 (1998), MIT Press
6. Stadler, P.F., Seitz, R., Wagner, G.P.: Population Dependent Fourier Decomposition of Fitness Landscapes over Recombination Spaces: Evolvability of Complex Characters. Bulletin of Mathematical Biology 62(3), 399–428 (2000), Society for Mathematical Biology
7. van de Vel, M.L.J.: Theory of Convex Structures. North-Holland Mathematical Library, North-Holland (1993)
8. Moraglio, A., Poli, R.: Inbreeding Properties of Geometric Crossover and Nongeometric Recombinations. In: Proceedings of the 9th International Conference on Foundations of Genetic Algorithms. pp. 1–14. FOGA '07, Springer, Mexico City, Mexico (2007)
9. Farber, M., Jamison, R.E.: Convexity in Graphs and Hypergraphs. SIAM Journal on Algebraic Discrete Methods 66(3), 231–247 (1986), Society for Industrial and Applied Mathematics

10. Moraglio, A.: Abstract Convex Evolutionary Search. In: Proceedings of the 11th Workshop on Foundations of Genetic Algorithms. pp. 151–162. FOGA '11, ACM, Schwarzenberg, Austria (2011)
11. Changat, M., Narasimha-Shenoi, P.G., Nezhad, F.H., Kovše, M., Mohandas, S., Ramachandran, A., Stadler, P.F.: Topological Representation of the Transit Sets of k-Point Crossover Operators. arXiv:1712.09022 (2017)
12. Gitchoff, P., Wagner, G.P.: Recombination Induced Hypergraphs: A New Approach to Mutation-Recombination Isomorphism. Complexity 2(1), 37–43 (1996), John Wiley & Sons, Inc.
13. Moraglio, A., Togelius, J., Silva, S.: Geometric Differential Evolution for Combinatorial and Programs Spaces. Evolutionary Computation 21(4), 591–624 (2013), MIT Press
14. Holland, J.H.: Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence. Complex Adaptive Systems, MIT Press, reprint edn. (1992)
15. Stadler, B.M.R., Stadler, P.F.: Generalized Topological Spaces in Evolutionary Theory and Combinatorial Chemistry. Journal of Chemical Information and Computer Sciences 42(3), 577–585 (2002), American Chemical Society
16. Mitavskiy, B.: Crossover Invariant Subsets of the Search Space for Evolutionary Algorithms. Evolutionary Computation 12(1), 19–46 (2004), MIT Press
17. Pelayo, I.M.: Geodesic Convexity in Graphs. SpringerBriefs in Mathematics, Springer Science & Business Media, 1st edn. (2013)
18. Davis, L.: Applying Adaptive Algorithms to Epistatic Domains. In: Proceedings of the 9th International Joint Conference on Artificial Intelligence–Volume 1. pp. 151–162. IJCAI'85, Morgan Kaufmann Publishers Inc. (1985)
19. Thorhauer, A., Rothlauf, F.: On the Bias of Syntactic Geometric Recombination in Genetic Programming and Grammatical Evolution. In: Proceedings of the 2015 Annual Conference on Genetic and Evolutionary Computation. pp. 1103–1110. GECCO '15, ACM, New York, NY, USA (2015)
20. Godsil, C.: Algebraic Combinatorics. Chapman Hall/CRC Mathematics Series, CRC Press, 1st edn. (1993)