Geometric Particle Swarm Optimisation

Alberto Moraglio, Cecilia Di Chio, and Riccardo Poli

Department of Computer Science,
University of Essex,
Wivenhoe Park, Colchester, CO4 3SQ, UK
{amoragn,cdichi,rpoli}@essex.ac.uk

Abstract. Using a geometric framework for the interpretation of crossover of recent introduction, we show an intimate connection between particle swarm optimization (PSO) and evolutionary algorithms. This connection enables us to generalize PSO to virtually any solution representation in a natural and straightforward way. We demonstrate this for the cases of Euclidean, Manhattan and Hamming spaces.

1 Introduction

Particle swarm optimisation (PSO) [5] has traditionally been applied to continuous search spaces. Although a version of PSO for binary search spaces has been defined [4], attempts to extend PSO to richer spaces, e.g., combinatorial spaces, have not been very successful.

There are two ways of extending PSO to richer spaces. Firstly one can rethink and adapt the PSO for each new solution representation. Secondly one can define a mathematical generalisation of the notion (and motion) of particles for a general class of spaces. This second approach has the advantage that a PSO can be derived in a principled way for any search space belonging to the given class. Here we follow this approach.

In particular, we show *formally* how a general form of PSO (without the inertia term) can be obtained by using theoretical tools developed for evolutionary algorithms using geometric crossover and geometric mutation. These are representation-independent operators that generalise many pre-existing search operators for the major representations, such as binary strings [7], real vectors [7], permutations [8], syntactic trees [8] and sequences [12].

In Sec. 2, we introduce the geometric framework and introduce the notion of multi-parental geometric crossover. In Sec. 3, we recast PSO in geometric terms and generalize it to generic metric spaces. In Sec. 4, we apply these notions to the Euclidean, Manhattan and Hamming spaces. In Sec. 5, we discuss how to specialise the general PSO automatically to virtually any solution representation using geometric crossover. In Sec. 6, we present conclusions and future work.

2 Geometric Framework

Geometric operators are defined using the notions of line segment and ball. These notions and the corresponding genetic operators are well-defined once a notion

M. Ebner et al. (Eds.): EuroGP 2007, LNCS 4445, pp. 125–136, 2007.

of distance in the search space is defined. Defining search operators as functions of the search space is opposite to the standard way where the search space is seen as a function of the search operators employed [3].

2.1 Geometric Preliminaries

The terms *distance* and *metric* denote any real valued function that conforms to the axioms of identity, symmetry and triangular inequality. A simple connected graph is naturally associated to a metric space via its *path metric*: the distance between two nodes in the graph is the length of a shortest path between the nodes. Distances arising from graphs via their path metric are called *graphic distances*. Similarly, an edge-weighted graph with strictly positive weights is naturally associated to a metric space via a *weighted path metric*.

In a metric space (S, d) a *closed ball* is a set of the form $B(x; r) = \{y \in S | d(x, y) \leq r\}$ where $x \in S$ and r is a positive real number. A *line segment* is a set of the form $[x; y] = \{z \in S | d(x, z) + d(z, y) = d(x, y)\}$ where $x, y \in S$ are called extremes of the segment. Metric ball and metric segment generalise the familiar notions of ball and segment in the Euclidean space to any metric space through distance redefinition. In general, there may be more than one shortest path (*geodesic*) connecting the extremes of a metric segment; the metric segment is the union of all geodesics.

We assign a structure to the solution set by endowing it with a notion of distance d. $M = (S, d)$ is a solution *space* and $L = (M, g)$ is the corresponding *fitness landscape*.

A family $\mathcal{X}$ of subsets of a set X is called *convexity on* X if: (C1) the empty set $\emptyset$ and the universal set X are in $\mathcal{X}$, (C2) if $\mathcal{D} \subseteq \mathcal{X}$ is non-empty, then $\bigcap \mathcal{D} \in \mathcal{X}$, and (C3) if $\mathcal{D} \subseteq \mathcal{X}$ is non-empty and totally ordered by inclusion, then $\bigcup \mathcal{D} \in \mathcal{X}$. The pair $(X, \mathcal{X})$ is called *convex structure*. The members of $\mathcal{X}$ are called *convex sets*. By the axiom (C1) a subset A of X of the convex structure is included in at least one convex set, namely X. From axiom (C2), A is included in a smallest convex set, the *convex hull* of A: $co(A) = \bigcap\{C | A \subseteq C \in \mathcal{X}\}$. The convex hull of a finite set is called a *polytope*. The axiom (C3) requires *domain finiteness* of the convex hull operator: a set C is convex iff it includes $co(F)$ for each finite subset F of C. The convex hull operator applied to set of cardinality two is called *segment operator*. Given a metric space $M = (X, d)$ the segment between a and b is the set $[a, b]_d = \{z \in X | d(x, z) + d(z, y) = d(x, y)\}$. The abstract *geodetic convexity* $\mathcal{C}$ on X induced by M is obtained as follow: a subset C of X is geodetically-convex provided $[x, y]_d \subseteq C$ for all x, y in C. If co denotes the convex hull operator of $\mathcal{C}$, then $\forall a, b \in X : [a, b]_d \subseteq co\{a, b\}$. The two operators need not to be equal: there are metric spaces in which metric segments are not all convex.

2.2 Two-Parent and Multi-parent Geometric Crossover

Definition 1. *(Geometric crossover) A binary operator is a geometric crossover under the metric d if all offspring are in the segment between its parents.*

The definition is *representation-independent* and, therefore, crossover is well-defined for any representation. Being based on the notion of metric segment, *crossover is only function of the metric d* associated with the search space.

This class of operators is really broad. For example, it includes: various types of blend or line crossovers, box recombinations, and discrete recombinations [7]; homologous crossovers [7,9]; PMX, Cycle crossover and merge crossover [8]; homologous GP crossovers [11]; and several others [12,7,8,10].

We now provide the following extension:

Definition 2. *(Multi-parental geometric crossover) In a multi-parental geometric crossover, given n parents $p_1, p_2, \ldots, p_n$ their offspring are contained in the metric convex hull of the parents $co(\{p_1, p_2, \ldots, p_n\})$ for some metric d.*

Theorem 1. *(Decomposable three-parent recombination) Every recombination $RX(p_1, p_2, p_3)$ that can be decomposed as a sequence of 2-parental geometric crossovers GX and GX' under the same metric, so that $RX(p_1, p_2, p_3) = GX(GX'(p_1, p_2), p_3)$, is a three-parental geometric crossover.*

Proof. Let P be the set of parents and $co(P)$ their metric convex hull. By definition of metric convex hull, for any two points $a, b \in co(P)$ their offspring are in the convex hull $[a, b] \subseteq co(P)$. Since $P \subseteq co(P)$, any two parents $p_1, p_2 \in P$ have offspring $o_{12} \in co(P)$. Then any other parent $p_3 \in P$ when recombined with o_{12} produces offspring o_{123} in the convex hull $co(P)$. So the three-parental recombination equivalent to the sequence of geometric crossover $GX'(p_1, p_2) \to o_{12}$ and $GX(o_{12}, p_3) \to o_{123}$ is a multi-parental geometric crossover.

3 Geometric PSO

3.1 Basic, Canonical PSO Algorithm and Geometric Crossover

Consider the canonical PSO in Algorithm 1. It is well known that one can write the equation of motion of the particle without making explicit use of its velocity.

Let x be the position of a particle and v be its velocity. Let $\hat{x}$ be the current best position of the particle and let $\hat{g}$ be the global best. Let v' and v'' be the velocity of the particle and $x' = x + v$ and $x'' = x' + v'$ its position at the next two time ticks. The equation of velocity update is the linear combination: $v' = w_1 v + w_2(\hat{x} - x') + w_3(\hat{g} - x')$ where w_1, w_2 and w_3 are scalar coefficients. To eliminate velocities we substitute the identities $v = x' - x$ and $v' = x'' - x'$ in the equation of velocity update and rearrange it to obtain an equation that expresses x'' as function of x and x': $x'' = (1 + w_1 - w_2 - w_3)x' - w_1 x + w_2\hat{x} + w_3\hat{g}$.

If we set $w_1 = 0$, which corresponds to setting $\omega = 0$ (i.e., the particle has no inertia), x'' becomes independent on its position two time ticks earlier x. The equation of motion becomes:

$$x'' = (1 - w_2 - w_3)x' + w_2\hat{x} + w_3\hat{g}. \tag{1}$$

In these conditions, the main feature that allows the motion of particles is the ability to perform linear combinations of points in the search space. As we will see in the next section, we can achieve this same ability by using multiple (geometric) crossover operations. This makes it possible to obtain a generalisation of PSO to generic search spaces.

Algorithm 1. Standard PSO algorithm

for all particle i **do**
  initialise position x_i and velocity v_i
end for
while stop criteria not met **do**
  for all particle i **do**
    set personal best $\hat{x}_i$ as best position found so far by the particle
    set global best $\hat{g}$ as best position found so far by the whole swarm
  end for
  for all particle i **do**
    update velocity using equation

$$v_i(t+1) = \omega v_i(t) + \phi_1 U(0,1)(\hat{g}(t) - x_i(t)) + \phi_2 U(0,1)(\hat{x}_i(t) - x_i(t)) \quad (2)$$

    update position using equation

$$x_i(t+1) = x_i(t) + v_i(t+1) \quad (3)$$

  end for
end while

3.2 Geometric Interpretation of Linear Combinations

If $v_1, ..., v_n$ are vectors and $a_1, ..., a_n$ are scalars, then the *linear combination* of those vectors with those scalars as coefficients is : $a_1 v_1 + a_2 v_2 + a_3 v_3 + \cdots + a_n v_n$. A linear combination on n linearly independent vectors spans completely an n-dimensional space but not a higher dimensional one. So, the linear combination of three linearly independent points spans a 3-dimensional space but not a 4-dimensional one.

An *affine combination* of vectors $v_1, ..., v_n$ is a linear combination $\sum_i a_i \cdot x_i$ in which $\sum_i a_i = 1$. When a vector represents a point in space, the affine combination of 2 independent points spans completely the line passing through them; the affine combination of 3 points spans completely the plane (2D line) passing through them; increasing number of linearly independent points span completely higher dimensional "lines".

A *convex combination* is an affine combination of vectors where all coefficients are non-negative. It is called "convex combination", since, when vectors represent points in space, the set of all convex combinations constitutes the convex hull.

A special case is $n = 2$, where a point formed by the convex combination will lie on a straight line between two points. For three points, their convex hull is the triangle with the points as vertices.

Theorem 2. *In a PSO with no inertia ($\omega = 0$) and where learning rates are such that $\phi_1 + \phi_2 \leq 1$, the next position of a particle x' is within convex hull formed by its current position x, its local best $\hat{x}$ and the swarm best $\hat{g}$.*

Proof. As we have seen in Sec. 3.1, when $\omega = 0$, a particle's update equation becomes the linear combination in equation (1). Notice that this is an affine combination since the coefficients of x', $\hat{x}$ and $\hat{g}$ add up to 1. *Interestingly, this means that the new position of the particle is coplanar with x', $\hat{x}$ and $\hat{g}$.* If we restrict w_2 and w_3 to be positive and their sum to be less than 1, equation (1) becomes a convex combination. *Geometrically this means that the new position of the particle is in the convex hull formed by (or more informally, between) its previous position, its local best and the swarm best.*

In the next section, we generalize this simplified form of PSO from real vectors to generic metric spaces. Mutation will be required to extend the search beyond the convex hull.

3.3 Convex Combinations in Metric Spaces

Linear combinations are well-defined for vector spaces, algebraic structures endowed with scalar product and vectorial sum. A metric space is a set endowed with a notion of distance. The set underlying a metric space does not normally come with well-defined notions of scalar product and sum among its elements. So a linear combination of its elements is not defined. How can we then define a convex combination in a metric space? Vectors in a vector space can easily be understood as points in a metric space. However, the interpretation of scalars is not as straightforward: what do the scalar weights in a convex combination mean in a metric space?

As seen in Sec. 3.2, a convex combination is an algebraic description of a convex hull. However, even if the notion of convex combination is not defined for metric spaces, convexity in metric spaces is still well-defined through the notion of metric convex set that is a straightforward generalization of traditional convex set. Since convexity is well-defined for metric spaces, we still have hope to generalize the scalar weights of a convex combination trying to make sense of them in terms of distance.

The weight of a point in a convex combination can be seen as a measure of relative linear attraction toward its corresponding point versus attractions toward the other points of the combination. The closer the weight to one, the stronger the attraction to its corresponding point. The point resulting from a convex combination can be seen as the equilibrium point of all the attraction forces. The distance between the equilibrium point and a point of the convex combination is therefore a decreasing function of the level of attraction (weight) of the point: the stronger the attraction, the smaller its distance to the equilibrium point. This observation can be used to reinterpret the weights of a convex combination in a metric space as follows: $y = w_1x_1 + w_2x_2 + w_3x_3$ with w_1, w_2 and w_3 greater than zero and $w_1 + w_2 + w_3 = 1$ is generalized to mean that y is a point such that $d(x_1, y) \sim 1/w_1$, $d(x_2, y) \sim 1/w_2$ and $d(x_3, y) \sim 1/w_3$.

This definition is formal and valid for all metric spaces but it is non-constructive. In contrast a convex combination, not only defines a convex hull, but it tells also how to reach all its points. So, how can we actually pick a point in the convex hull respecting the above distance requirements? Geometric crossover will help us with this, as we show in the next section.

The requirements for a convex combination in a metric space are:

1. Convex weights: the weights respect the form of a convex combination: $w_1, w_2, w_3 > 0$ and $w_1 + w_2 + w_3 = 1$
2. Convexity: the convex combination operator combines x_1, x_2 and x_3 and returns a point in their metric convex hull, or simply triangle, under the metric of the space considered
3. Coherence between weights and distances: the distances to the equilibrium point are decreasing functions of their weights
4. Symmetry: the same value assigned to w_1, w_2 or w_3 has the same effect (so in a equilateral triangle, if the coefficients have all the same value, the distances to the equilibrium point are the same).

3.4 Geometric PSO Algorithm

The generic Geometric PSO algorithm is illustrated in Algorithm 2. This differs from the standard PSO (Algorithm 1) in that: there is no velocity, the equation of position update is the convex combination, there is mutation and the parameters w_1, w_2, and w_3 are non-negative and add up to one. The specific PSO for the Euclidean, Manhattan and Hamming spaces use the randomized convex combination operators described in Sec. 4 and space-specific mutations.

Algorithm 2. Geometric PSO algorithm

for all particle i **do**
    initialise position x_i at random in the search space
end for
while stop criteria not met **do**
    for all particle i **do**
        set personal best $\hat{x}_i$ as best position found so far by the particle
        set global best $\hat{g}$ as best position found so far by the whole swarm
    end for
    for all particle i **do**
        update position using a randomized convex combination

$$x_i = CX((x_i, w_1), (\hat{g}, w_2), (\hat{x}_i, w_3)) \quad (4)$$

        mutate x_i
    end for
end while

4 Geometric PSO for Specific Spaces

4.1 Euclidean Space

Geometric PSO for the Euclidean space is not an extension of the traditional PSO. We include it to show how the general notions introduced in the previous section materialize in a familiar context. The convex combination operator for the Euclidean space is the traditional convex combination that produces points in the traditional convex hull.

In Sec. 3.3, we have mentioned how to interpret the weights in a convex combination in terms of distances. In the following we show analytically how the weights of a convex combination affect the relative distances to the equilibrium point. In particular we show that the relative distances are decreasing functions of the corresponding weights.

Theorem 3. *In a convex combination, the distances to the equilibrium point are decreasing functions of the corresponding weights.*

Proof. Let a, b and c be three points in $\mathbb{R}^n$ and $x = w_a a + w_b b + w_c c$ be a convex combination. Let us now decrease w_a to $w'_a = w_a - \Delta$ such that w'_a, w'_b and w'_c still form a convex combination and that the relative proportions of w_b and w_c remain unchanged: $\frac{w'_b}{w'_c} = \frac{w_b}{w_c}$. This requires w'_b and w'_c to be $w'_b = w_b(1+\Delta/(w_b+w_c))$ and $w'_c = w_c(1+\Delta/(w_b+w_c))$. The equilibrium point for the new convex combination is $x' = (w_a - \Delta)a + w_b(1 + \Delta/(w_b + w_c))b + w_c(1 + \Delta/(w_b + w_c))c$. The distance between a and x is $|a - x| = |w_b(a - b) + w_c(a - c)|$ and the distance between a and the new equilibrium point is $|a - x'| = |w_b(1 + \Delta/(w_b + w_c))(a - b) + w_c(1 + \Delta/(w_b + w_c))(a - c)| = (1 + \Delta/(w_b + w_c))|a - x|$. So when w_a decreases ($\Delta > 0$) and w_b and w_c maintain the same relative proportions, the distance between the point a and the equilibrium point x increases ($|a - x'| > |a - x|$). Hence the distance between a and the equilibrium point is a decreasing function of w_a. For symmetry this applies to the distances between b and c and the equilibrium point: they are decreasing functions of their corresponding weights w_b and w_c, respectively.

The traditional convex combination in the Euclidean space respects the four requirements for a convex combination presented in Sec. 3.3.

4.2 Manhattan Space

In the following we first define a multi-parental recombination for the Manhattan space and then prove that it respects the four requirements for being a convex combination presented in Sec. 3.3.

Definition 3. *(Box recombination family) Given two parents a and b in $\mathbb{R}^n$, a box recombination operator returns offspring o such that $o_i \in [min(a_i, b_i), max(a_i, b_i)]$ for $i = 1 \dots n$.*

Theorem 4. *(Geometricity of box recombination) Any box recombination is geometric crossover under Manhattan distance*

Theorem 4 is an immediate consequence of the product geometric crossover theorem.

Definition 4. *(Three-parent Box recombination family) Given three parents a, b and c in* $\mathbb{R}^n$*, a box recombination operator returns offspring o such that* $o_i \in [min(a_i, b_i, c_i), max(a_i, b_i, c_i)]$ *for* $i = 1 \dots n$*.*

Theorem 5. *(Geometricity of three-parent box recombination) Any three-parent box recombination is geometric crossover under Manhattan distance.*

Proof. We prove it by showing that any multi-parent box recombination $BX(a, b, c)$ can be decomposed as a sequence of two simple box recombinations. Since simple box recombination is geometric (Theorem 4), this theorem is a simple corollary of the multi-parental geometric decomposition theorem (Theorem 1).

We will show that $o' = BX(a, b)$ followed by $BX(o', c)$ can reach any offspring $o = BX(a, b, c)$. For each i we have $o_i \in [min(a_i, b_i, c_i), max(a_i, b_i, c_i)]$. Notice that $[min(a_i, b_i), max(a_i, b_i)] \cup [min(a_i, c_i), max(a_i, c_i)] = [min(a_i, b_i, c_i), max(a_i, b_i, c_i)]$. We have two cases: (i) $o_i \in [min(a_i, b_i), max(a_i, b_i)]$ in which case o_i is reachable by the sequence $BX(a, b)_i \to o_i, BX(o, c)_i \to o_i$; (ii) $o_i \notin [min(a_i, b_i), max(a_i, b_i)]$ then it must be in $[min(a_i, c_i), max(a_i, c_i)]$ in which case o_i is reachable by the sequence $BX(a, b)_i \to a_i, BX(a, c)_i \to o_i$

Definition 5. *(Weighted multi-parent Box recombination) Given three parents a, b and c in* $\mathbb{R}^n$ *and weights* w_a*,* w_b *and* w_c*, a weighted box recombination operator returns offspring o such that* $o_i = w_{a_i} a_i + w_{b_i} b_i + w_{c_i} c_i$ *for* $i = 1 \dots n$*, where* w_{a_i}*,* w_{b_i} *and* w_{c_i} *are a convex combination of randomly perturbed weights with expected values* w_a*,* w_b *and* w_c*.*

The difference between box recombination and linear recombination (Euclidean space) is that in the latter the weights w_a, w_b and w_c are randomly perturbed only once and the same weights are used for all the dimensions, whereas the former one has a different randomly perturbed version of the weights for each dimension.

The weighted multi-parent box recombination belongs to the family of multi-parent box recombination because $o_i = w_{a_i} a_i + w_{b_i} b_i + w_{c_i} c_i \in [min(a_i, b_i, c_i), max(a_i, b_i, c_i)]$ for $i = 1 \dots n$, hence it is geometric.

Theorem 6. *(Coherence between weights and distances) In weighted multi-parent box recombination, the distances of the parents to the expected offspring are decreasing functions of the corresponding weights.*

The proof of theorem 6 is a simple variation of that of theorem 3.

In summary in this section we have introduced the weighted multi-parent box recombination and shown that it is a convex combination operator satisfying the four requirements of a metric convex combination for the Manhattan space: convex weights by definition (Definition 4), convexity (geometricity, Theorem 5), coherence (Theorem 6) and symmetry (self-evident).

4.3 Hamming Space

In the following we first define a multi-parental recombination for binary strings that is a straightforward generalization of mask-based crossover with two parents and then prove that it respects the four requirements for being a convex combination in the Hamming space presented in Sec. 3.3.

Definition 6. *(Three-parent mask-based crossover family) Given three parents* a, b *and* c *in* $\{0,1\}^n$*, generate randomly a crossover mask of length* n *with symbols from the alphabet* $\{\mathtt{a}, \mathtt{b}, \mathtt{c}\}$*. Build the offspring* o *filling each position with the bit from the parent appearing in the crossover mask at the position.*

The weights w_a, w_b and w_c of the convex combination indicate for each position in the crossover mask the probability of having the symbols a, b or c.

Theorem 7. *(Geometricity of three-parent mask-based crossover) Any three-parent mask-based crossover is geometric crossover under Hamming distance.*

Proof. We prove it by showing that any three-parent mask-based crossover can be decomposed as a sequence of two simple mask-based crossovers. Since simple mask-based crossover is geometric, this theorem is a simple corollary of the multi-parental geometric decomposition theorem (Theorem 1).

Let m_{abc} the mask to recombine a, b and c producing the offspring o. Let m_{ab} the mask obtained by substituting all occurrences of c in m_{abc} with b and m_{bc} the mask obtained by substituting all occurrences of a in m_{abc} with b. Now recombine a and b using m_{ab} obtaining b'. Then recombine b' and c using m_{bc} where the b's in the mask stand for alleles in b'. The offspring produced by the second crossover is o, so the sequence of the two simple crossovers is equivalent to the three-parent crossover. This is because the first crossover passes to the offspring the all genes it needs to take from a according to m_{abc} and the rest of the genes are all from b; the second crossover corrects those genes that should have been taken from parent c according to m_{abc} but were taken from b instead.

Theorem 8. *(Coherence between weights and distances) In weighted three-parent mask-based crossover, the distances of the parents to the expected offspring are decreasing functions of the corresponding weights.*

Proof. We want to know the expected distance from parent p_1, p_2 and p_3 and their expected offspring o as a function of the weights w_1, w_2 and w_3. To do so, we first determine, for each position in the offspring, the probability to be the same as p_1. From that then we can easily compute the expected distance between p_1 and o. We have that

$$pr\{o = p_1\} = pr\{p_1 \to o\} + pr\{p_2 \to o\} \cdot pr\{p_1|p_2\} + pr\{p_3 \to o\} \cdot pr\{p_1|p_3\} \quad (5)$$

where: $pr\{o = p_1\}$ is the probability of a bit of o at a certain position to be the same as the bit of p_1 at the same position; $pr\{p_1 \to o\}$, $pr\{p_2 \to o\}$ and $pr\{p_3 \to o\}$ are the probabilities that a bit in o is taken from parent p_1, p_2 and p_3, respectively (these coincide with the weights of the convex combination w_1, w_2 and w_3); $pr\{p_1|p_2\}$ and $pr\{p_1|p_3\}$ are the probabilities that a bit taken from p_2 or p_3 coincides with the one in p_1 at the same location. These last two probabilities equal the number of common bits in p_1 and p_2 (p_1 and p_3) over the length of the strings n. So $pr\{p_1|p_2\} = 1 - H(p_1, p_2)/n$ and $pr\{p_1|p_3\} = 1 - H(p_1, p_3)/n$ where $H(\cdot, \cdot)$ is the Hamming distance. So equation (5) becomes

$$pr\{o = p_1\} = w_1 + w_2(1 - H(p_1, p_2)/n) + w_3(1 - H(p_1, p_3)/n). \qquad (6)$$

Hence the expected distance between the parent p_1 and the offspring o is: $E(H(p_1, o)) = n \cdot (1 - pr\{o = p_1\}) = w_2 H(p_1, p_2) + w_3 H(p_1, p_3)$. Notice that this is a decreasing function of w_1 because increasing w_1 forces w_2 or w_3 to decrease since the sum of the weights is constant, hence $E(H(p_1, o))$ decreases. Analogously, $E(H(p_2, o))$ and $E(H(p_3, o))$ are decreasing functions of their weights w_2 and w_3, respectively.

In summary in this section we have introduced the weighted multi-parent mask-based crossover and shown that it is a convex combination operator satisfying the four requirements of a metric convex combination for the Hamming space: convex weights by definition (Definition 5) , convexity (geometricity, Theorem 7), coherence (Theorem 8) and symmetry (self-evident).

5 Towards a Geometric PSO for GP and Other Representations

Before looking into how we can extend geometric PSO to other solution representations, we will discuss the relation between 3-parental geometric crossover and the symmetry requirement for a convex combination.

For each of the spaces considered in section 4, we have first considered, or defined, a three-parental recombination and then we proved that it is a three-parental geometric crossover by showing that it can actually be decomposed into two sequential applications of a geometric crossover for the specific space.

However, we could have skipped altogether the *explicit definition* of a three-parental recombination. In fact to obtain the three-parental recombination we could have used two sequential applications of a known two-parental geometric crossover for the specific space. This composition is indeed a three-parental recombination, it combines three parents, and it is decomposable by construction, hence it is a three-parental geometric crossover. This, indeed, would have been simpler than the route we took.

The reason we preferred to define explicitly a three-parental recombination is that the requirement of symmetry of the convex combination is true by construction: if the roles of any two parents are swapped exchanging in the three-parental recombination both positions and respective recombination weights, the resulting recombination operator is equivalent to the original operator.

The symmetry requirement becomes harder to enforce and prove for a three-parental geometric crossover obtained by two sequential applications of a two-parental geometric crossover. We illustrate this in the following. Let us consider three parents a, b and c with positive weights w_a, w_b and w_c which add up to one. If we have a symmetric three-parental weighted geometric crossover ΔGX, the symmetry of the recombination is guaranteed by the symmetry of the operator. So, $\Delta GX((a, w_a), (b, w_b), (c, w_c))$ is equivalent to $\Delta GX((b, w_b), (a, w_a), (c, w_c))$, hence the requirement of symmetry on the weights of the convex combination holds. If we consider a three-parental recombination defined by using twice a two-parental genetic crossover GX we have:

$$\Delta GX((a, w_a), (b, w_b), (c, w_c)) = GX((GX((a, w'_a), (b, w'_b)), w_{ab}), (c, w'_c)) \quad (7)$$

with the constraint that w'_a and w'_b are positive and add up to one and w_{ab} and w'_c are positive and add up to one. It is immediate to notice the inherent asymmetry in this expression: the weights w'_a and w'_b are not directly comparable with w'_c because they are relative weights between a and b. Moreover there is the extra weight w_{ab}. This makes the requirement of symmetry problematic to meet: given the desired w_a, w_b and w_c, what values of w'_a, w'_b, w_{ab} and w'_c do we have to choose to obtain an equivalent symmetric 3-parental weighted recombination expressed as a sequence of two two-parental geometric crossovers?

For the Euclidean space, it is easy to answer this question using simple algebra: $\Delta GX = w_a \cdot a + w_b \cdot b + w_c \cdot c = (w_a + w_b)(\frac{w_a}{w_a+w_b} \cdot a + \frac{w_b}{w_a+w_b} \cdot b) + w_c \cdot c$. Since the convex combination of two points in the Euclidean space is $GX((x, w_x), (y, w_y)) = w_x \cdot x + w_y \cdot y$ and $w_x, w_y > 0$ and $w_x + w_y = 1$ then $\Delta GX((a, w_a), (b, w_b), (c, w_c)) = GX((GX((a, \frac{w_a}{w_a+w_b}), (b, \frac{w_b}{w_a+w_b})), w_a + w_b), (c, w_c))$. This question may be less straightforward to answer for other spaces, although we could use the equation above as a rule-of-thumb to map the weights of ΔGX and the weights in the sequential GX decomposition.

Where does this discussion leave us in relation to the extension of geometric PSO to other representations? We have seen that there are two alternative ways to produce a convex combination for a new representation: (i) explicitly define a symmetric three-parental recombination for the new representation and then prove its geometricity by showing that it is decomposable into a sequence of two two-parental geometric crossovers, or (ii) use twice the simple geometric crossover to produce a symmetric or nearly symmetric three-parental recombination. In this paper we used the first approach, but the second option is also very interesting because *it allows us to extended automatically geometric PSO to all representations we have geometric crossovers for, such as permutations, GP trees, variable-length sequences, to mention a few, and virtually any other complex solution representation.*

6 Conclusions and Future Work

We have extended the geometric framework with the notion of multi-parent geometric crossover that is a natural generalization of two-parental geometric crossover: offspring are in the convex hull of the parents. Then, using the geometric framework, we have shown an intimate relation between a simplified form of PSO, without the inertia term, and evolutionary algorithms. This has enabled us to generalize in a natural, rigorous and automatic way PSO for any type of search space for which a geometric crossover is known.

We have specialised the general PSO to Euclidean, Manhattan and Hamming spaces, obtaining three instances of the general PSO for the specific spaces.

In future work we will consider geometric PSO for permutation spaces and spaces of genetic programs, for which several geometric crossovers exist. We will also test the geometric PSO experimentally.

References

1. T. Bäck and D. B. Fogel and T. Michalewicz. Evolutionary Computation 1: Basic Algorithms and Operators. *Institute of Physics Publishing, 2000.*
2. M. Deza and M. Laurent. Geometry of cuts and metrics. *Springer, 1991.*
3. T. Jones. *Evolutionary Algorithms, Fitness Landscapes and Search.* PhD thesis, University of New Mexico, 1995.
4. J. Kennedy and R.C. Eberhart, *A Discrete Binary Version of the Particle Swarm Algorithm*, IEEE, 1997.
5. J. Kennedy and R.C. Eberhart, *Swarm Intelligence*, Morgan Kaufmann, 2001.
6. S. Luke and L. Spector. A Revised Comparison of Crossover and Mutation in Genetic Programming. *Proceedings of Genetic Programming Conference, 1998.*
7. A. Moraglio and R. Poli. Topological interpretation of crossover. In *Proceedings of the Genetic and Evolutionary Computation Conference*, pages 1377–1388, 2004.
8. A. Moraglio and R. Poli. Topological crossover for the permutation representation. In *GECCO 2005 Workshop on Theory of Representations*, 2005.
9. A. Moraglio and R. Poli. Product Geometric Crossover. In *Proceedings of the Parallel Problem Solving from Nature Conference*, pages 1018–1027, 2006.
10. A. Moraglio and R. Poli. Geometric Crossover for Sets, Multisets and Partitions. In *Proceedings of the Parallel Problem Solving from Nature Conference*, pages 1038–1047, 2006.
11. A. Moraglio and R. Poli. Geometric Landscape of Homologous Crossover for Syntactic Trees. In *Proceedings of the 2005 IEEE Congress on Evolutionary Computation*, pages 427–434, 2005.
12. A. Moraglio, R. Poli, and R. Seehuus. Geometric crossover for biological sequences. In *Proc. of European Conference on Genetic Programming*, pages 121–132, 2006.
13. P. M. Pardalos and M.G.C. Resende. Handbook of Applied Optimization. *Oxford University Press, 2002.*
14. N. Radcliffe. Equivalence Class Analysis of Genetic Algorithms. *Journal of Complex Systems, 1991*, vol. 5, pages 183–205.
15. M. L. J. Van de Vel. Theory of Convex Structures. *North-Holland, 1993* .