

Analysis of Locality in Hybrid Evolutionary Cluster Optimization

Francisco B. Pereira, *Member IEEE*, Jorge M. C. Marques, Tiago Leitão, and Jorge Tavares, *Student Member, IEEE*

Abstract—State of the art algorithms for cluster geometry optimization rely on hybrid approaches that combine the global exploration performed by evolutionary methods with local search procedures. These methods use derivative information to discover the nearest local optimum. In this paper we analyze the locality properties of this approach to gain insight on the algorithm's strengths and weaknesses and to determine the role played by each of its components. Results show that there are important differences in what concerns the locality of different mutation operators commonly used in this problem.

I. INTRODUCTION

Atomic and molecular clusters may have from a few to a huge number of aggregated particles (atoms or molecules). The function describing the interactions among those particles is designated as Potential Energy Surface (PES) and plays a central role in Chemistry. Then, exploring the energy landscapes of such a multi-dimensional PES has long been a desiderata but also a very hard task to achieve. In particular, finding the arrangement of the cluster's particles that corresponds to the lowest energy (global minimum) has been a key-step of the research in this scientific domain.

Since the early 1990's Evolutionary Algorithms (EAs) have been increasingly applied to several problems from the Chemistry/Biochemistry area. Cluster geometry optimization is an example of one of these problems [1], [2], [3]. Nearly all the approaches rely on hybrid algorithms combining EAs with local methods that use first order derivative information to guide search into the nearest local optimum.

Locality is an important requisite to ensure the efficiency of search and it has been widely studied by the evolutionary computation community [4], [5], [6], [7], [8], [9]. In general terms, this property indicates that small variations in the genotype space, usually originated by mutation, imply small variations in the phenotype space [6]. A locally strong search algorithm is able to efficiently explore the neighborhood of the current solutions. When this condition is not satisfied, the exploration performed by the EA is inefficient and tends to resemble random search.

The main goal of this paper is to perform an empirical

This work was supported by Fundação para a Ciência e Tecnologia, Portugal, under grant POSC/EIA/55951/2004.

Francisco B. Pereira, Tiago Leitão and Jorge Tavares are with the Centro de Informática e Sistemas da Universidade de Coimbra, Polo II, 3030 Coimbra, Portugal (e-mails: xico@dei.uc.pt, tleitao@student.dei.uc.pt, jast@dei.uc.pt). Francisco B. Pereira is also with the Instituto Superior de Engenharia de Coimbra, Quinta da Nora, 3030-199 Coimbra, Portugal.

Jorge M. C. Marques is with the Depart. de Química da Universidade de Coimbra, 3004-535 Coimbra, Portugal (e-mail: qtmarque@ci.uc.pt).

study on the hybrid algorithm that is usually adopted for cluster optimization. Locality measures will be used for the analysis and the framework proposed by Raidl and Gottlieb [6] is extended to deal with a real valued representation where a gradient driven local method is employed. Two distance measures, suitable for the adopted representation, are used: fitness based and structural distance.

Results allow us to gain some insight about the degree of locality induced by different mutation operators. They also clarify the importance of each one of the methods that coexist in the hybrid approach, helping to explain why local search is essential to ensure that good solutions are found. The search space defined for this problem is highly irregular and even small modifications performed by genetic operators in the structure of an individual lead to large phenotypic changes. An EA working alone is not able to overcome these difficulties. However, the addition of a gradient-driven local search method enhances locality and allows the hybrid algorithm to efficiently explore the search space. To our knowledge, this is the first locality study conducted on hybrid approaches used on cluster geometry optimization. Results help to provide a better understanding of the role played by each one of the components of the algorithm, which may be important for future applications of EAs to similar problems from the chemistry area.

The structure of the paper is the following: in section II we briefly describe Morse clusters. In section III we present the hybrid algorithm used in the experiments. Optimization results are presented and discussed in section IV. In section V we present a detailed locality analysis of this problem and, finally, section VI gathers the main conclusions.

II. MORSE CLUSTERS

Morse clusters are considered a benchmark for testing the performance of new methods for cluster structure optimization. The energy of such a cluster is represented by the N -particle pair-wise additive potential [10] defined as

$$V_{Morse} = \varepsilon \sum_i \sum_{j>i}^N [e^{-2\beta(r_{ij}-r_0)} - 2e^{-\beta(r_{ij}-r_0)}] \quad (1)$$

where the variable r_{ij} is the Cartesian distance between atoms i and j in the cluster structure. The bond dissociation energy ε , the equilibrium bond length r_0 and the range exponent of the potential β are parameters defined for each individual pair-wise Morse interaction. Usually, these are assumed to be constant for all interactions in a cluster formed by only one type of atoms. The potential of equation 1 is a scaled version [1] of the Morse function with non-atom-specific

interactions, where ε and r_0 have been both set to 1 and β has been fixed at 14, which corresponds to a short-range interaction. Global minima search is particularly challenging for short-range Morse clusters, since they have many local minima and a "noisy" PES [11]. This simplified potential has already been studied by other authors [1], [12], [13], and the minima are well established [14] for many values of N .

III. EAS FOR MORSE CLUSTER OPTIMIZATION

EAs have been used since 1993 for cluster geometry optimization. A comprehensive review of these efforts, including an outline of state-of-the art applications, can be found at [15]. In what concerns the application of EAs to Morse clusters, the most important works are from Johnston and collaborators [3], [13]. In our analysis we adopt an experimental model similar to the one used by these researchers. Its main components have been proposed and evaluated by different teams [2], [3], [16], [17].

A. Chromosome Representation and Evaluation

For aggregates with N atoms, a solution is composed by $3 \times N$ real values specifying the Cartesian coordinates of each one of the particles. Zeiri proposed this representation in 1995 [17] and, since then, it has become the most widely used in this context [2], [3]. The coordinate values range between 0 and λ . We set λ to $N^{1/3}$. This interval ensures that the aggregate volume scales correctly with N [13]. To evaluate an individual we just have to calculate its potential energy. Equation 1 is used to assign fitness.

B. Population Model and Genetic Operators

A generational model is adopted and the standard set of variation operators is used: crossover and mutation.

Cut and splice crossover, proposed by Deaven and Ho [16], is directly applied to the cluster structure and several authors claim that it enhances the performance of the algorithm [2], [13]. The purpose of this operator is to exchange sub-clusters between parents. More specifically, it selects a cutting plane (parallel to xy) for each parent and creates descendants by combining complementary parts.

Two mutation operators are tested in this paper: Sigma mutation and Flip mutation. We consider that mutation is performed on atoms, i.e., when applied it modifies the value of the three coordinates. The first operator is an evolutionary strategy (ES) like mutation and acts in the following way: when undergoing mutation, the new value v_{new} for each one of the three coordinates of an atom (x, y, z) is obtained from the old value v_{old} in the following way:

$$v_{\text{new}} = v_{\text{old}} + \sigma \times N(0,1) \quad (2)$$

where $N(0, 1)$ represents a random value sampled from a standard Normal distribution and σ is a parameter from the algorithm. The new value must be between 0 and λ .

Flip mutation works in the following way: when applied to an atom, it assigns new random values to each one of its coordinates, i.e., it moves this atom to a random location.

C. Local Optimization

Local optimization is performed with the Broyden-Fletcher-Goldfarb-Shanno limited memory quasi-Newton method (L-BFGS) of Liu and Nocedal [18], [19]. L-BFGS is a powerful optimization technique that aims to combine the modest storage and computational requirements of conjugate gradient methods with the superlinear convergence exhibited by full memory quasi Newton methods.

L-BFGS is applied to every generated individual. During local search, the maximum number of iterations that can be performed is specified by a parameter of the algorithm, the Local Search Length (LSL). However, L-BFGS stops as soon as it finds a local optimum, so the effective number of iterations can be smaller than the value specified by LSL.

IV. REVIEW OF RESULTS

Before starting the locality analysis, we present some experimental results to confirm the competence of the hybrid approach. The algorithm is applied to search for the optimal geometry of Morse clusters ranging from 19 to 50 atoms. The settings of the algorithm are the following: Evaluations: 1,000,000 (clusters between 19 and 29 atoms), 2,000,000 (clusters between 30 and 39 atoms), 3,000,000 (clusters between 40 and 50 atoms); Population Size: 100; Elitist Strategy; Tournament Selection with tourney size 10; Mutation operators: {Sigma, Flip}; $\sigma = \{0.01 \times \lambda, 0.1 \times \lambda\}$; Mutation rate: 0.05; Cut and splice crossover with rate: 0.7; LSL: 200; δ : 0.5. δ specifies the minimum distance between atoms and is used when generating the initial population or when applying genetic operators. It is useful as it prevents the generation of clusters with atoms that are excessively close to each other. Each iteration performed by L-BFGS counts as one evaluation, so the computational effort of the hybrid and non-hybrid algorithms is nearly equivalent. The initial population is randomly generated and for every set of parameters we performed 30 runs. When appropriate, statistical significance of the results is accessed with a t-test (level of significance 0.01).

In table I we present an overview of the achieved results for three sets of experiments. They distinguish one from another by the used mutation operator: {Sigma0.01, Sigma0.1, Flip}¹. Column Best shows the potential energy of the best solution found and column MBF presents the mean best fitness (average of the best fitness over the 30 runs). Column Gap presents the distance between the best solution found and the putative optimum value. A general overview of the results reveals that the hybrid approach is effective, confirming previous studies [3], [13]. It was able to find all known best solutions, including the one for the cluster with 30 atoms that was not reported in the initial paper by Roberts et al. [13], even though it was later mentioned in another

¹ The σ value used in Sigma mutation is proportional to λ to ensure that its effect is comparable for clusters of different size. In the text we will adopt the simplified notation Sigma0.01 instead of Sigma0.01 $\times\lambda$.

paper by the same authors [3].

Individually, experiments with Sigma0.01 mutation failed to find the optimum 6 times, whereas experiments with Sigma0.1 mutation were unable to find the optimum 3 times and experiments with Flip mutation did not find the global best solution for 9 clusters. Anyway, it can be seen that all gaps are tiny. For all clusters sizes and for all settings, differences between the best solution found and the MBF are small, providing evidence that the approach is robust. Also, it shows good scalability, because the gaps do not increase as the cluster size grows (with the possible exception of experiments with Flip mutation).

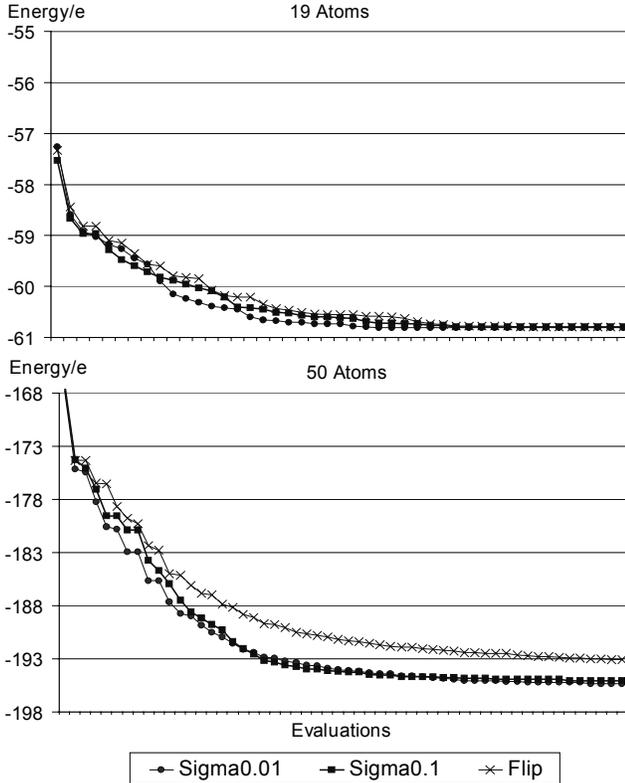


Figure 1: Evolution of MBF for instances with 19 and 50 atoms

A more detailed analysis reveals that there are some differences in what concerns results achieved by different mutation operators. First of all, results obtained by Sigma0.01 and Sigma0.1 can be considered as similar and there is never a statistical significant difference between them. In contrast, experiments with Flip mutation reveal a less consistent behavior. For small clusters (up to 30 atoms), the achieved results are analogous to the ones obtained by Sigma mutation. As the clusters grow in size, its performance starts to deteriorate. For clusters with more than 36 atoms, in most times it fails to find the optimum. Also, for aggregates with more than 30 atoms, the MBF attained by Flip mutation is always of inferior quality than those achieved by Sigma0.01 or Sigma0.1. To confirm this trend, there are some differences that are statistically significant (entries in bold in table I identify a situation where there is a significant difference between the MBFs achieved by experiments using

Sigma mutation and Flip mutation). Charts from figure 1 show the evolution of MBF along the generations for two optimization instances: a) for 19 atoms and b) for 50 atoms. It is visible that, in the smallest instance, differences in performance are very small. On the contrary, for the largest instance, there is a noticeable difference.

V. LOCALITY ANALYSIS

The largest cluster ($N=50$) was selected to perform all the tests concerning the locality analysis. A new value, $\sigma=0.25\times\lambda$, is added to the two previously used. Sigma0.25 will provide insight on the effect of σ in the performance of the algorithm. The empirical analysis is complemented with some additional experimental results.

A. Related Work

Many approaches were proposed to estimate the behavior of EAs. Some of these techniques adopt measures that are, to some extent, similar to the locality property. In this section we highlight the most relevant ones.

The concept of fitness landscapes, originally proposed by Wright [20], establishes a connection between solution candidates and their fitness values and it has been widely used to predict EAs performance. Jones and Forrest proposed fitness distance correlation as a way to determine the relation between fitness and distance to the optimum [21]. If fitness values increase as distance to the optimum decreases, then search is expected to be easy for an EA [22].

An alternative way to analyze the fitness landscape is to determine its ruggedness. Some autocorrelation measures help to determine how rugged a landscape is. Weinberger [23] proposed the adoption of autocorrelation functions to measure the correlation of all points in the search space at a given distance. Another possibility is to perform random walks and determine the correlation length, a value that directly reflects the ruggedness of the landscape [22].

Sendhoff et al. studied the conditions for strong causality on EAs [9]. A search process is said to be locally strongly causal if small variations in the genotype space imply small variations in the phenotype space. In the above-mentioned work [9], variations in genotypes are caused by mutation.

B. Definitions

When performing studies with an evolutionary framework it is usual to consider two spaces: the genotype space Φ_g and the phenotype space Φ_p . Genetic operators work on Φ_g , whereas the fitness function f is applied to solutions from Φ_p : $f: \Phi_p \rightarrow \mathcal{R}$. A direct representation is adopted in this paper. Since there is not a maturation or decoder function, genetic operators are directly applied to phenotypes. This way, it is not necessary to perform an explicit distinction between the two spaces and, from now on, we will refer to individuals or phenotypes to designate points from the search space.

To calculate the similarity between two individuals from Φ_p , a phenotypic distance has to be defined. This measure

captures the semantic difference between two solutions and is directly related to the problem being solved. We determine phenotypic distance in two ways:

Fitness based distance: determining the fitness distance between two phenotypes A , B is straightforward:

$$d_{fit}(A, B) = |f(A) - f(B)| \quad (3)$$

In cluster optimization, it calculates the difference between the energy potential values of two solutions.

Structural distance: According to equation 1, the basic features that influence the quality of a N -atom cluster are the $N \times (N-1)/2$ interactions occurring between particles forming the aggregate. The interaction between atoms i and j depends only on the distance r_{ij} between them. We implement a simple method to approximate the structural shape of a cluster. First, all the $N \times (N-1)/2$ distances between atoms are calculated. Then, they are separated into several sets according to its values. We consider 10 sets S_i . The limits for each S_i , $i=1, \dots, 10$, are defined as follows:

$$\left[\frac{i-1}{10} \times \mu, \frac{i}{10} \times \mu \right], i = 1, \dots, 10 \quad (4)$$

where μ is the maximum distance between two atoms. Considering the parameter λ , μ is equal to $\sqrt{3 \times \lambda^2}$. Structural distance captures the dissimilarity between two clusters A and B in what concerns the distances between all pairs of atoms. It is measured in the following way:

$$d_{struct}(A, B) = \frac{1}{10} \sum_{i=1}^{10} |\#S_i(A) - \#S_i(B)| \quad (5)$$

where $\#S_i(A)$ (likewise, $\#S_i(B)$) is the cardinality of subset S_i for cluster A (likewise, for cluster B).

C. Mutation Innovation

To analyze the effect of mutation on locality we adopt the innovation measure proposed by Raidl and Gottlieb [6]. The distance between the individuals involved in a mutation is used to predict the effect of the application this operator. Let X be a solution and X^m the result of applying mutation to X . The mutation innovation MI is measured as follows:

$$MI = dist(X, X^m) \quad (6)$$

Distance can be calculated using either fitness based or structural distance. MI illustrates how much innovation the mutation operator introduces, i.e., it aims to determine how much this operator modifies the semantic properties of an individual. Locality is directly related to this measure. The application of a locally strong operator implies a small modification in the phenotype on an individual (i.e., small distances between the two involved solutions). On the contrary, operators with weak locality allow large jumps in the search space, complicating the task of the search algorithm. To determine the MI , 1000 random individuals were generated and then, a sequence of mutations was applied to each one of them. In each one of the 1000 mutation series, distance is measured between the original individual and the solution created after $k \in \{1, 2, 3, 4, 5, 10, 25, 50, 100\}$ successive mutation steps. In conformity to

the adopted optimization framework, local search is considered as part of this genetic operator, i.e., L-BFGS is applied after each mutation and distance is measured using the solution that results from this operation.

1) Fitness Based Distance

To simplify the analysis, distances between the original solution and the successive mutants are grouped in different sets. Given a d_{fit} fitness distance between two solutions, set G_i to which d_{fit} is assigned, is determined in the following way: **G0**: $0 \leq d_{fit} < 1$; **G1**: $1 \leq d_{fit} < 5$; **G2**: $5 \leq d_{fit} < 10$; **G3**: $10 \leq d_{fit} < 20$; **G4**: $20 \leq d_{fit} < 30$; **G5**: $30 \leq d_{fit} < 50$; **G6**: $50 \leq d_{fit} < 100$; **G7**: $100 \leq d_{fit} < 250$; **G8**: $250 \leq d_{fit} < 500$; **G9**: $500 \leq d_{fit}$.

The specific values that were selected to determine intervals are arbitrary. The relevant information to obtain here is the distribution of the fitness distances through the sets. Situations where values tend to be assigned to higher order sets (i.e., large variations), suggest that the locality is low. In charts from figure 2 we present, for the considered number of mutations, the distances between the original solutions and the mutants iteratively generated. In each column (corresponding to a given mutation step), we show the distribution of the 1000 distances for the 10 G_i sets. We present 4 charts, one for each operator used. It is clear that there are important differences in what concerns the locality of mutation. Experiments combining L-BFGS with Sigma0.01 exhibit the highest locality and, even after 1000 steps, nearly all fitness distances belong to sets G0-G2 (more than 50% are in cluster G0). This shows that there is still a clear relation, maybe even excessive, to the departure point. On the contrary, experiments with Sigma0.25 and Flip mutation evidence low locality. The modification they induce is substantial (large displacement of one atom strongly modifies the fitness of the solution) and local search is not able to do an appropriate repair. Sigma0.1 is between these two scenarios. In the beginning it shows signs of reasonable locality, but after some steps it approaches the distribution exhibited by Sigma0.25 and Flip.

2) Structural Distance

Analysis of MI with structural distance confirms the conclusions from the previous section. In figure 3 we show the structural distance for the same operators and settings. Once again, it is clear that Sigma0.01 has a high locality, Sigma0.25 and Flip have both low locality and Sigma0.1 is between the two extremes. Results with both phenotypic distances reveal that, as σ increases, the effect on the locality of Sigma mutation tends approach that of Flip mutation.

3) Additional Tests

To complement our study, we performed some additional experiments, with an extended set of mutation rates: $\{0.01, 0.05, 0.1\}$. Furthermore, results using Sigma0.25 are also provided. A summary of results can be found in table II. If we combine the information obtained with locality analysis and these new results, it is possible to infer some conclusions related to mutation rate sensitivity and search competence.

TABLE I: OPTIMIZATION RESULTS FOR MORSE CLUSTERS BETWEEN 19 AND 50 ATOMS

N	SIGMA 0.01			SIGMA 0.1			FLIP		
	Best	MBF	Gap (%)	Best	MBF	Gap (%)	Best	MBF	Gap (%)
19	-60.812	-60.812	0.00	-60.812	-60.790	0.00	-60.812	-60.790	0.00
20	-64.792	-64.740	0.00	-64.792	-64.700	0.00	-64.792	-64.720	0.00
21	-68.783	-68.520	0.00	-68.783	-68.610	0.00	-68.783	-68.660	0.00
22	-72.792	-72.600	0.00	-72.792	-72.680	0.00	-72.792	-72.760	0.00
23	-77.302	-77.060	0.00	-77.302	-77.040	0.00	-77.302	-77.120	0.00
24	-81.309	-81.250	0.00	-81.309	-81.150	0.00	-81.309	-81.230	0.00
25	-85.477	-85.210	0.00	-85.477	-85.200	0.00	-85.477	-85.240	0.00
26	-90.211	-89.650	0.00	-90.211	-89.810	0.00	-90.211	-89.860	0.00
27	-94.220	-93.900	0.00	-94.220	-93.970	0.00	-94.220	-93.960	0.00
28	-98.332	-97.870	0.00	-98.332	-97.920	0.00	-98.332	-98.070	0.00
29	-102.774	-102.220	0.00	-102.774	-102.260	0.00	-102.774	-102.230	0.00
30	-106.836	-106.320	0.00	-106.765	-106.440	0.07	-106.836	-106.330	0.00
31	-111.761	-111.030	0.00	-111.761	-111.010	0.00	-111.761	-110.810	0.00
32	-115.767	-115.340	0.00	-115.767	-115.090	0.00	-115.767	-115.070	0.00
33	-120.741	-119.800	0.00	-120.741	-119.840	0.00	-120.741	-119.340	0.00
34	-124.748	-124.050	0.00	-124.748	-123.870	0.00	-124.748	-123.600	0.00
35	-129.737	-128.560	0.00	-129.737	-128.330	0.00	-129.737	-128.260	0.00
36	-133.745	-133.050	0.00	-133.745	-132.660	0.00	-133.745	-132.460	0.00
37	-138.708	-137.380	0.00	-138.708	-137.840	0.00	-138.690	-136.830	0.01
38	-142.715	-141.470	1.11	-144.321	-141.960	0.00	-142.715	-141.160	1.11
39	-148.327	-145.500	0.00	-148.327	-146.200	0.00	-148.327	-145.100	0.00
40	-152.334	-150.600	0.00	-152.334	-150.460	0.00	-151.704	-149.610	0.41
41	-156.633	-154.960	0.00	-156.631	-154.390	0.05	-156.631	-153.960	0.00
42	-160.641	-159.300	0.00	-160.641	-159.170	0.00	-160.366	-158.330	0.17
43	-165.373	-163.600	0.16	-165.635	-163.680	0.00	-165.372	-162.640	0.16
44	-169.520	-167.940	0.07	-169.642	-167.830	0.00	-169.642	-167.040	0.00
45	-174.386	-171.970	0.07	-174.512	-172.510	0.00	-173.410	-170.630	0.63
46	-178.416	-177.100	0.06	-178.519	-176.880	0.00	-178.517	-176.150	0.00
47	-183.508	-181.380	0.00	-183.411	-180.880	0.05	-182.652	-179.870	0.47
48	-188.889	-186.260	0.00	-188.889	-185.780	0.00	-187.430	-184.040	0.77
49	-192.898	-190.220	0.00	-192.898	-190.040	0.00	-191.951	-188.930	0.49
50	-197.910	-195.070	0.28	-198.456	-195.330	0.00	-195.900	-193.040	1.29

TABLE II: INFLUENCE OF MUTATION IN THE OPTIMIZATION OF THE MORSE CLUSTER WITH 50 ATOMS

Mutation rate	SIGMA 0.01		SIGMA 0.1		SIGMA 0.25		FLIP	
	Best	MBF	Best	MBF	Best	MBF	Best	MBF
0.01	-198.456	-195.656	-198.456	-194.505	-197.910	-193.995	-198.456	-195.236
0.05	-197.910	-195.070	-198.456	-195.330	-197.910	-193.464	-195.900	-193.040
0.1	-197.910	-195.246	-198.456	-194.971	-195.350	-191.784	-195.440	-189.613

TABLE III: INFLUENCE OF LOCAL SEARCH IN THE OPTIMIZATION OF THE MORSE CLUSTER WITH 50 ATOMS

	SIGMA 0.01		SIGMA 0.1		SIGMA 0.25		FLIP	
	Best	MBF	Best	MBF	Best	MBF	Best	MBF
Hybrid LSL200	-197.910	-195.070	-198.456	-195.330	-197.910	-193.464	-195.900	-193.040
Hybrid LSL20	-191.990	-180.914	-194.800	-183.390	-188.740	-178.810	-190.470	-178.882
EC alone	-78.340	-56.843	-68.420	-54.217	-86.220	-57.495	-75.670	-54.565

When Sigma0.01 or Sigma 0.1 are used, modifying the mutation rate does not lead to noteworthy changes in the results. The t-test shows that there are not significant differences between MBFs when changing the mutation rate from 1% to 5% and then to 10%. Therefore, the hybrid algorithm with high locality (resulting from a combination of a restrained atoms displacement followed by local search) is fairly insensitive to the variation of the mutation rate. This is a relevant effect, as it simplifies the task of parameter setting.

In contrast, it is possible to detect important differences in results obtained by Sigma0.25 and Flip mutation. For both cases, the MBF values steadily decrease as the mutation rate increases. When Sigma0.25 is used, the difference between MBFs from experiments with 5% and 10% mutation rates is statistically significant. In experiments with Flip mutation, all differences are statistically significant. Situations with low locality are therefore highly sensitive to variations in the mutation rate. This is not an unexpected outcome, as the

previous locality analysis showed that modifications performed by Sigma0.25 or Flip cause a considerable modification in the cluster properties and subsequent local optimization might not be able to recover and obtain a new solution that is phenotypically close to the original one.

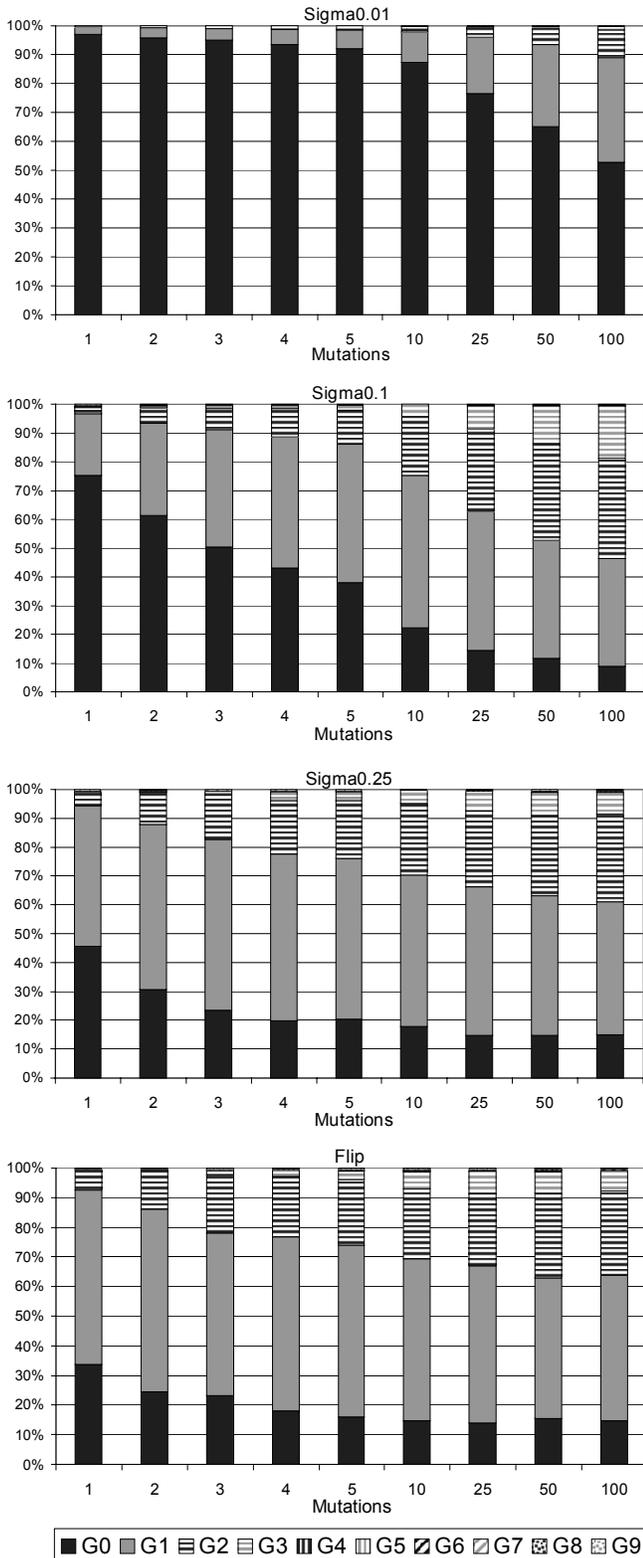


Figure 2: Distribution of fitness distances between the original solutions and the mutants iteratively generated (LSL=200).

Increasing the mutation rate magnifies this problem because mutation is applied several times before local optimization acts. These results show that, in situations where the application of mutation operators with weak locality is necessary, they should be kept at a low rate.

Finally, data from table II confirms that mutation operators with weak locality achieve worse results. With just one exception (Flip with 1% mutation rate), the MBF from experiments using Sigma0.25 or Flip is always of inferior quality than that from tests performed with Sigma0.01 or Sigma0.1. For experiments with 5% and 10% mutation rates, differences are statistically significant.

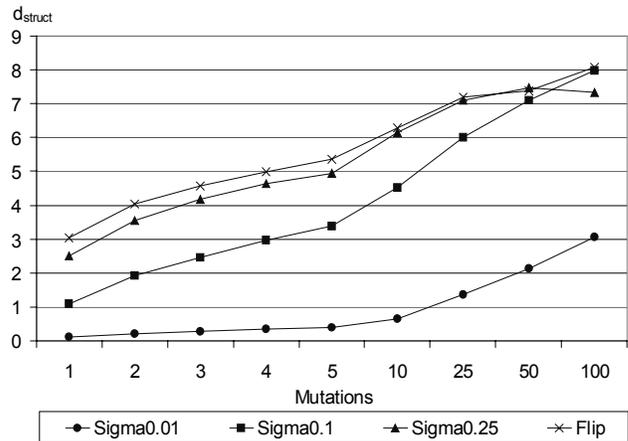


Figure 3: Average structural distances between the original solutions and the mutants iteratively generated (LSL=200).

D. L-BFGS Impact on Search

To demonstrate the influence of L-BFGS, two extra experiments were performed: one without local search and another with local search length LSL=20. Tests were performed with 5% mutation rate and results are shown in table III. Original values obtained with LSL=200 are also presented to simplify the comparison. A brief perusal of the results reveals the importance of local optimization. When alone, the performance of the EA is extremely poor. Allowing 20 steps of local optimization enables the algorithm to discover solutions that are much better than those achieved by evolution alone. Anyway, the results are still below the ones that were obtained by the original hybrid algorithm. Differences are always statistically significant.

A brief analysis of locality helps to justify the results. We will just present results for two mutations operators: Sigma0.1 and Flip. The same trend is visible for other operators. We recalculated the MI using the framework described in section V-C, with one exception: the maximum number of steps performed by local search is different. In charts a) and b) from figure 4, fitness based distance is measured in situations where L-BFGS is not applied. Charts c) and d) present results from experiments where L-BFGS is allowed to perform a maximum of 20 steps after each mutation. Chart from figure 5 presents the structural distance for these 4 situations. Results from charts 4a) and 4b)

illustrate that the locality is extremely low. After 4 mutations, nearly all fitness distances belong to G9. Moreover, all operators exhibit the same behavior confirming that even small changes in the location of an atom lead to major modifications in the cluster potential energy. This result suggests that the changes performed by the EA are too disruptive. As the locality is low, the algorithm is unable to exploit information in a useful way. Local optimization is therefore essential to make the landscape less irregular. When the MI is assessed with the structural distance measure, there are some differences between these two experiments. The structural locality of Sigma0.1 is always higher than that of Flip mutation. This may be explained by the fact that the average structural modification performed by the second operator is large, when compared to the one carried out by Sigma0.1. The difference in results obtained with the two distance measures confirms that even small structural changes can lead to large fitness modifications, making the task of the optimization very hard. In their work, Sendhoff et al. conclude that ES mutation guarantee strongly causal optimization procedures [9]. Results presented here show that this kind of mutation might not be enough when the search space is extremely rugged. Further actions (such as the addition of a local search method) might be required to ensure high locality.

Results from charts 4c) and 4d) show an increase in locality, when compared to the previous situation. In any case, the limited number of local steps is not enough and the phenotypic changes are still large. This explains the results presented in table III: experiments with LSL=20 improve the performance of the EA alone, but are still behind the results achieved by tests with LSL=200. Structural distance achieves similar results in both tests (without local optimization and with LSL=20). This suggests that, after mutation, a limited number of local steps might not be enough to reorganize the cluster in a promising way. Anyway, the increase in performance shows that limited local search is important, at least to solve some problems that might be created by mutation, such as placing two atoms too close.

Even though results from table III show that L-BFGS is crucial to achieve good results, it is nevertheless important to refer that the EA is an essential component of the hybrid approach. This is confirmed by results obtained in a final experiment using just the local search method. To ensure a fair comparison, 30 runs were done and, in each one of them, the same number of evaluations was performed. During each run, when L-BFGS converged, it was restarted in a different point of the search space. This experiment achieved the following results: Best solution found: -188.632; MBF: -183.240. It is clear that, in this problem, L-BFGS is more effective than the EA. Anyway, it was the combined application of both methods that achieved the best performance. L-BFGS is essential to perform local exploration and fine-tuning of solutions, ensuring that a high locality is maintained, but the global exploration carried out

by the EA is also mandatory to avoid premature convergence and to perform an efficient sampling of the search space.

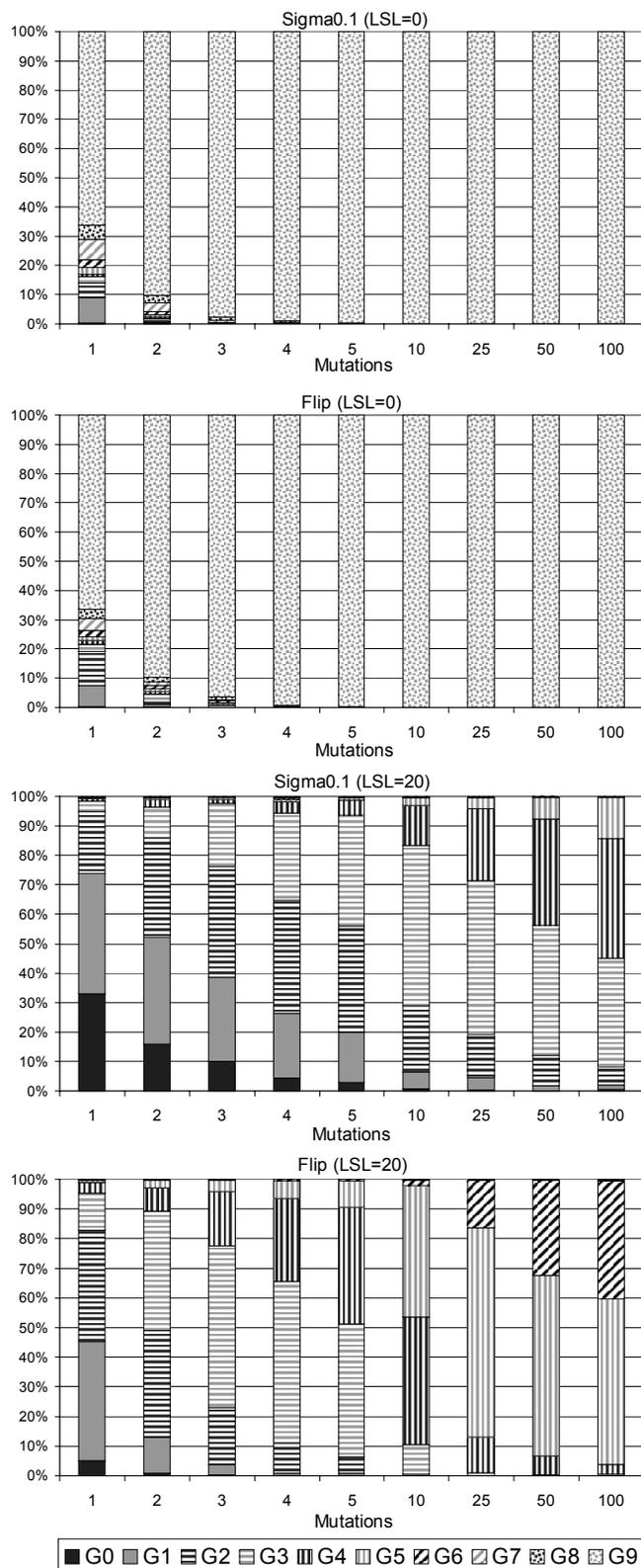


Figure 4: Distribution of fitness distances between the original solutions and the mutants iteratively generated.

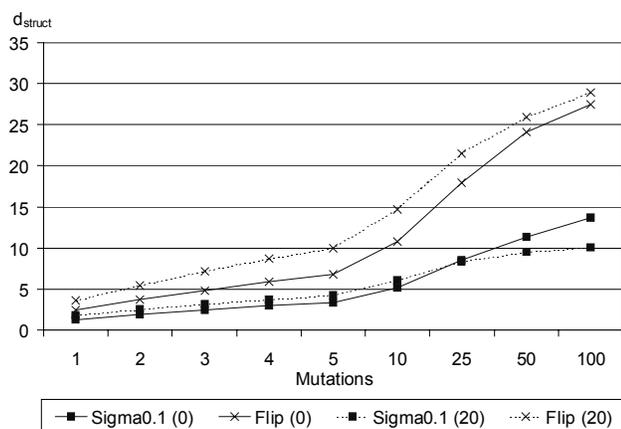


Figure 5: Average structural distances between the original solutions and the mutants iteratively generated. Values in brackets designate the LSL.

VI. CONCLUSIONS

In this paper we studied the locality properties of the hybrid evolutionary algorithm usually applied in cluster geometry optimization. Two distance measures, required to determine the semantic differences between individuals, were used to perform an extensive analysis. Results confirm that high locality is important to improve performance. In the described approach, high locality is the result of the combined application of little disruptive mutation operators and subsequent local optimization. A useful outcome from this work is that it advocates how to tailor the hybrid algorithm to maximize performance when searching for clusters with the lowest possible energy:

1. Sigma mutation is an appropriate variation operator, but a small standard deviation must be selected to ensure that a strong locality is maintained.

2. Both components of the hybrid algorithm are essential. Furthermore, the number of iterations performed by the local search should be large enough to guarantee that the modifications performed by mutation do not lead to large variations in the phenotype space.

Locality analysis conducted in this paper was mainly based on mutation walks performed on the search space. As future work, we will extend this study to crossover. We also intend to conduct this analysis directly on runs performed by the hybrid algorithm to study if the interactions occurring between selection and variation operators significantly modify the empirical results that were presented here.

ACKNOWLEDGMENT

We are grateful to the John von Neumann Institut für Computing, Jülich, for the provision of supercomputer time on the IBM Regatta p690+ (Project EPG00).

REFERENCES

[1] J. P. K. Doye and D. J. Wales. Structural Consequences of the Range of the Interatomic Potential. A Menagerie of Clusters. *J. Chem. Soc. Faraday Trans.* 93, 4233-4243, 1997.

[2] B. Hartke. Application of Evolutionary Algorithms to Global Cluster Geometry Optimization, In R. L. Johnston (Ed.), *Applications of Evolutionary Computation in Chemistry, Structure and Bonding*, 110, 33-53, 2004.

[3] R. L. Johnston. Evolving Better Nanoparticles: Genetic Algorithms for Optimising Cluster Geometries, *Dalton Transactions*, 4193-4207, 2003.

[4] J. Gottlieb and C. Eckert. A Comparison of Two Representations for the Fixed Charge Transportation Problem, In M. Schoenauer et al. (Eds.), *Parallel Problem Solving from Nature (PPSN VI)*, 345-354, Springer-Verlag LNCS, 2000.

[5] J. Gottlieb and G. Raidl. Characterizing Locality in Decoder-Based EAs for the Multidimensional Knapsack Problem, In C. Fonlupt et al. (Eds.), *Artificial Evolution: Fourth European Conference*, 38-52, Springer-Verlag LNCS, 1999.

[6] G. Raidl and J. Gottlieb. Empirical Analysis of Locality, Heritability and Heuristic Bias in Evolutionary Algorithms: A Case Study for the Multidimensional Knapsack Problem. *Evolutionary Computation*, Vol. 13, No. 4, 441-475, 2005.

[7] F. Rothlauf and D. Goldberg. Prüfernnumbers and Genetic Algorithms: A Lesson on Hoe the Low Locality on an Encoding Can Harm the Performance of Gas, In M. Schoeneauer et al. (Eds.), *Parallel Problem Solving from Nature PPSN VI*, 395-404, 2000.

[8] F. Rothlauf. On the Locality of Representations, In E. Cantú-Paz et al. (Eds.), *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO-2003)*, Part II, 1608-1609, 2003.

[9] B. Sendhoff, M. Kreutz and W. Seelen. A Condition for the Genotype-Phenotype Mapping: Causality. In T. Bäck (Ed.), *Proceedings of the 7th International Conference on Genetic Algorithms (ICGA-97)*, 73-80, 1997.

[10] P. Morse. Diatomic Molecules According to the Wave Mechanics. II. Vibrational Levels. *Phys. Rev.*, 34, 57-64, 1929.

[11] J. P. K. Doye, R. Leary, M. Locatelli and F. Schoen. Global Optimization of Morse Clusters by Potential Energy Transformations, *Inform. Journal on Computing*, 16, 371-379, 2004.

[12] F. Manby, R. L. Johnston and C. Roberts. Predatory Genetic Algorithms. *Commun. Math. Comput. Chem.* 38, 111-122, 1998.

[13] C. Roberts, R. L. Johnston and N. Wilson (2000). A Genetic Algorithm for the Structural Optimization of Morse Clusters. *Theor. Chem. Acc.*, 104, 123-130, 2000.

[14] D. J. Wales et al. The Cambridge Cluster Database, URL: <http://www-wales.ch.cam.ac.uk/CCD.html>, accessed on January 2006.

[15] B. Hartke. Global Geometry Optimization of Atomic and molecular Clusters by Genetic Algorithms, In L. Spector et al. (Eds.), *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO-2001)*, 1284-1291.

[16] D. M. Deaven and K. Ho. Molecular Geometry Optimization with a Genetic Algorithm, *Phys. Rev. Lett.* 75, 288-291, 1995.

[17] Y. Zeiri. Prediction of the Lowest Energy Structure of Clusters Using a Genetic Algorithm, *Phys. Rev.*, 51, 2769-2772, 1995.

[18] D. C. Liu and J. Nocedal. On the Limited Memory Method for Large Scale Optimization, *Mathematical Programming B*, 45, 503-528, 1989.

[19] J. Nocedal. Large Scale Unconstrained Optimization, In A. Watson and I. Duff (Eds.), *The State of the Art in Numerical Analysis*, 311-338, 1997.

[20] S. Wright. The Roles of Mutation, Inbreeding, Crossbreeding and Selection in Evolution. In *Proceedings of the 6th International Conference on Genetics*, Vol. 1, 356-366, 1932.

[21] T. Jones and S. Forrest. Fitness Distance Correlation as a Measure of Problem Difficulty for Genetic Algorithms. In L. Eshelman (Ed.), *Proceedings of the 6th International Conference on Genetic Algorithms (ICGA-95)*, 184-192, 1995.

[22] P. Merz. *Memetic Algorithms for Combinatorial Optimization Problems: Fitness Landscapes and Effective Search Strategies*. Ph.D. Thesis, Department of Electrical Engineering and Computer Science, University of Siegen, Germany, 2000.

[23] E. D. Weinberger. Correlated and Uncorrelated Fitness Landscapes and How to Tell the Difference, *Biological Cybernetics*, Vol. 63, 325-336, 1990.