Università degli Studi "G.D'Annunzio"
Dipartimento di Scienze

Symmetry-breaking and local search: The case of LABS

Andrea Roli

December 20, 2004

Technical Report no. R-2005-002 *Research Series*

Symmetry-breaking and local search: The case of LABS

Andrea Roli

Dipartimento di Scienze
Università degli Studi "G.D'Annunzio"
Pescara – Italy
a.roli@unich.it

December 20, 2004

Abstract. We report and discuss a series of experiments in which we compare the search space of LABS induced by modeling the problem with and without symmetry-breaking constraints. Furthermore, we compare the local search effectiveness in the two cases. In most of the instances analyzed, we observe that the total basin of attraction of global optima in the model with symmetry-breaking constraints is reduced by a factor that is higher than the search space reduction factor. We also experimentally find that local search is strongly affected by the size of global optima basins of attraction. To a certain extent, this behavior can explain why symmetry-breaking constraints have negative impact on local search.

Keywords: *Local search, Symmetry breaking, Low autocorrelation binary sequence*

Contents

1 Introduction

Symmetry-breaking has been proved to be very effective when combined with complete solvers [3, 8]. This can be explained by observing that symmetry-breaking constraints considerably reduce the search space. Nevertheless, the use of symmetry-breaking constraints (hereinafter referred to as SB constraints) seem to have opposite effect on local search-based solvers, despite the search space reduction. In [6, 7] some examples of this phenomenon are reported. When the problem is modeled with SB constraints, the search cost[1] is higher than the one corresponding to the model with symmetries.

In this work, we analyze the effect of SB constraints on local search for the problem called Low auto-correlation binary sequences (LABS). We first briefly define a model of the search space explored by a local search algorithm in Section 2. In Section 3, we report results concerning the search space characteristics of small size instances of LABS. Then, in Section 4, we compare the performance of four local search algorithms on both the original model and the model in which some symmetric solutions have been cut.

2 The search graph and its main characteristics

The local search process can be viewed as an exploration of a landscape aimed at finding an optimal solution, or a *good* solution, i.e., a solution with a quality above a given threshold.[2]

We define the search space explored by a local search algorithm as a *search graph*. The topological properties of such a graph are defined upon the neighborhood structure, that generate the *neighborhood graph*.

2.1 Neighborhood graph

A *Neighborhood Graph* (NG), also called *Fitness Landscape* (FL), is defined by a triple: $\mathcal{L} = (S, \mathcal{N}, f)$, where:

- S is the set of feasible states;[3]
- $\mathcal{N}$ is the neighborhood function $\mathcal{N} : S \to 2^S$ that defines the neighborhood structure, by assigning to every $s \in S$ a set of states $\mathcal{N}(s) \subseteq S$.
- f is the objective function $f\colon\ S \to \mathbb{R}^+$

The neighborhood graph can be interpreted as a graph (see Figure 1) in which nodes are states (labeled with their objective value) and arcs represent the neighborhood relation between states.

The neighborhood function $\mathcal{N}$ implicitly defines an *operator* φ which takes a state s_1 and transforms it into another state $s_2 \in \mathcal{N}(s_1)$. Conversely, given an operator φ, it is possible to define a neighborhood of a variable $s_1 \in S$:

$$\mathcal{N}_\varphi(s_1) = \{s_2 \in S \setminus \{s_1\} \mid s_2 \text{ can be obtained by one application of } \varphi \text{ on } s_1\}$$

[1] Runtime and number of variable flips

[2] For the rest of this paper, we will suppose, without loss of generality, that the goal of the search is to find an optimal solution. Indeed, the same conclusions we will draw can be extended to a set including also good solutions.

[3] In the field of metaheuristics, feasible states are also called *solutions*.

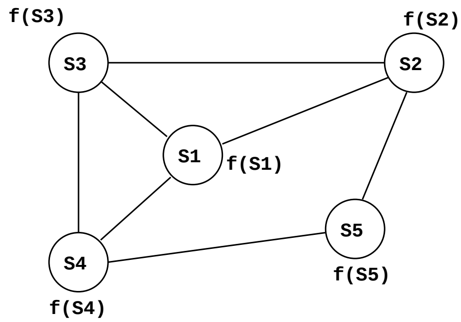


Figure 1. *Example of undirected graph representing a neighborhood graph (fitness landscape). Each node is associated with a solution s_i and its corresponding objective value $f(s_i)$. Arcs represent transition between states by means of φ. Undirected arcs correspond to symmetric neighborhood structure.*

In most of the cases, the operator is *symmetric*: if s_1 is a neighbor of s_2 then s_2 is a neighbor of s_1. In a graph representation (like the one depicted in Figure 1) undirected arcs represent symmetric neighborhood structures.

2.2 Search graph

The exploration process of local search methods can be seen as the evolution in (discrete) time of a discrete dynamical system [1, 4]. The algorithm starts from an initial state and describes a trajectory in the state space, that is defined by the neighborhood graph. The system dynamics depends on the strategy used; simple algorithms generate a trajectory composed of two parts: a *transient* phase followed by an *attractor* (a fixed point, a cycle or a complex attractor). Algorithms with advanced strategies generate more complex trajectories which can not be subdivided in those two phases.

It is useful to define the search as a walk on the neighborhood graph. In general, the choice of the next state is a function of the search history (i.e., the sequence of the previous visited states) and the iteration step. Formally: $s(t+1) = \phi(\langle s(0), s(1), \ldots, s(t)\rangle, t)$, where the function ϕ is defined on the basis of the search strategy. ϕ could also depend on some parameters and can be either deterministic or stochastic.

For instance, let us consider a deterministic version of the Iterative Improvement local search. The trajectory starts from a point s_0, exhaustively explores its neighborhood, picks the neighboring state s' with minimal objective function value[4] and, if s' is better than s_0, it moves from s_0 to s'. Then this process is repeated, until a minimum $\hat{s}$ (either local or global) is found. The trajectory does not move further and we say that the system has reached a fixed point ($\hat{s}$). The set of points from which $\hat{s}$ can be reached is the basin of attraction of $\hat{s}$.

Once we have introduced also the search strategy, the edges of the graph can be oriented and labeled with transition probabilities (whenever it is possible to evaluate them). This will lead to the definition of concepts such as basins of attraction, state reachability and graph navigation. In the following, this resulting graph will be referred to as *search graph*.

We would like to remark that, while the neighborhood graph topology is only dependent on the neighborhood structure and the problem model, the basins of attraction and other

[4]Ties are broken by enforcing a lexicographic order of states.

related search graph characteristics depend also on the particular algorithm used.

2.3 Basins of attraction

The concept of *basin of attraction* (BOA) has been introduced in the context of dynamical systems, in which it is defined referring to an *attractor*. Concerning our model of local search, we will use the concept of basin of attraction of any node of the search graph. Moreover, for this definition to be valid for any state of the search graph, we have to relax the requirement that the goal state is an attractor. Therefore, the basin of attraction will also depend on the particular termination condition of the algorithm. In the following, we will suppose to apply a termination condition such that the algorithm is stopped as soon as a stagnation condition is detected, that is when no improvements to the solutions are found after a maximum number of steps. This termination condition corresponds to the concept of steady state in dynamical systems. We will initially consider the case of deterministic systems, then we will relax this hypothesis and extend the definition to stochastic systems.

Definition Given a deterministic algorithm $\mathcal{A}$, the basin of attraction $\mathcal{B}(\mathcal{A}|s)$ of a point s, is defined as the set of states that, taken as initial states, give origin to trajectories that include point s. The cardinality of a basin of attraction represents its size (in this context, we always deal with finite spaces).

Given the set S^* of the global optima, the union of the BOA of global optima $I^* = \bigcup_{i \in S^*} \mathcal{B}(\mathcal{A}|i)$ represents the set of desirable initial states of the search. Indeed, a search starting from $s \in I^*$ will eventually find an optimal solution. Since it is usually not possible to construct an initial solution that is guaranteed to be in I^*, the ratio $rGBOA = |I^*|/|S|$ can be taken as an indicator of the probability to find an optimal solution. On the extreme case, if we start from a random solution, the probability to find a global optimum is exactly $|I^*|/|S|$. Therefore, the higher this ratio, the higher the probability of success of the algorithm.

In the case of stochastic local search, we may define a probabilistic basin of attraction, as a generalization of the previous case.

Definition Given a (stochastic) algorithm $\mathcal{A}$, the basin of attraction $\mathcal{B}(\mathcal{A}|s;p^*)$ of a point s, is defined as the set of states that, taken as initial states, give origin to trajectories that include point s *with probability* $p \geq p^*$. Also in this case, we define the union of the BOA of global optima: $I^*(p) = \bigcup_{i \in S^*} \mathcal{B}(\mathcal{A}|i;p)$. For simplicity, in the following we will write $\mathcal{B}(s;p^*)$ instead of $\mathcal{B}(\mathcal{A}|s;p^*)$ when the algorithm involved is clear from the context.

This definition includes the previous one as a special case. Indeed, if $p^* = 1$ we are interested in finding the states generating trajectories that will eventually reach to s. It is also important to note that if $p_1 > p_2$, then $\mathcal{B}(s;p_1) \subseteq \mathcal{B}(s;p_2)$.

Given a local search algorithm $\mathcal{A}$, the topology and structure of the search graph determine the effectiveness of $\mathcal{A}$. In particular, the reachability of optimal solutions is the key issue. Therefore, the characteristics of the BOA of optimal solutions are of dramatic importance.

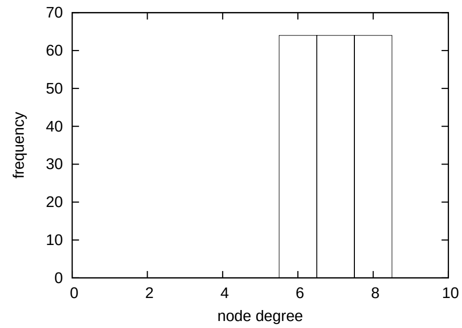


Figure 2. *Node degree frequency of the neighborhood graph in the case of $n = 10$ and the model with SB constraints.*

2.4 A conjecture on the effect of symmetry-breaking constraints

We conjecture that the main reason for SB constraints being harmful for local search has to be found in the reduction of $rGBOA$ defined on the basis of a simple iterative improvement local search. In fact, even the most complex local search algorithms incorporate a greedy heuristic which is the one that characterizes iterative improvement. Therefore, if SB constraints reduce $rGBOA$, then the more a local search is similar to iterative improvement, the more it should be affected by SB constraints. Furthermore, we should also observe that local search algorithms equipped with complex exploration strategies are less affected by SB constraints. In this work, we aim at experimentally verifying this conjecture.

3 LABS search space

LABS consists of finding an assignment to binary variables such that an energy function defined upon them is minimized. Given n binary variables $x_1, \dots, x_n$, which can assume a value in $\{-1, +1\}$, we define the k-th correlation coefficient of a complete variable assignment s $C_k(s) = \sum_{i=1}^{n-k} x_i\ x_{i+k}$, $k = 1, \dots, n-1$ and the total function to be minimized is $E(s) = \sum_{k=1}^{n-1} C_k^2(s)$.

We exhaustively explored the search space of LABS, for n ranging from 6 up to 18.[5] As neighborhood function we chose the one defined upon unitary Hamming distance, that is the most used one for problems defined over binary variables. (We emphasize that this choice determines the fundamental topological properties of the search space.) In the model with SB constraints (hereinafter referred to as $\mathcal{M}_s$, while the original model will be referred to as $\mathcal{M}$), only a subset of symmetric solutions has been cut, by enforcing constraints on the three left-most and right-most variables [5].

We first have to study how the neighborhood graph changes upon the application of SB constraints. In $\mathcal{M}$, the neighborhood graph induced by single variable flips is a hypercube in which each node is connected to other n nodes. This graph has a constant degree equal to n. The neighborhood graph associated to $\mathcal{M}_s$ is characterized by a node degree frequency that varies in a small range, around a mean value slightly small than n (see an example in Figure 2). The topological characteristics of this graph are not affecting the search, since the reachability

[5]The size limit is due to the exhaustiveness of the analysis.

Table 1. *Search space characteristics of LABS instances.*

n	feasible states		global optima		local optima		global BOA	
	no SB	with SB	no SB	with SB	no SB	with SB	no SB	with SB
6	64	12	28	5	0	0	1.0	1.0
7	128	24	4	1	24	5	0.40625	0.33333
8	256	48	16	3	8	2	0.86328	0.77083
9	512	96	24	4	84	16	0.42969	0.37500
10	1024	192	40	7	128	29	0.54590	0.45833
11	2048	384	4	1	240	52	0.03906	0.04427
12	4096	768	16	3	264	61	0.07544	0.06901
13	8192	1536	4	1	496	111	0.01831	0.01953
14	16384	3072	72	11	664	177	0.21240	0.15202
15	32768	6144	8	2	1384	326	0.01956	0.01742
16	65536	12288	32	8	1320	332	0.05037	0.04972
17	131072	24576	44	9	3092	721	0.05531	0.04073
18	262144	49152	16	2	5796	1372	0.02321	0.01068

of nodes is not significantly perturbed. Therefore, we can exclude that SB constraints in LABS affect local search by perturbing the topological properties of the neighborhood graph.

We have now also to consider the features of the search graph, which, in general, can be algorithm dependent. (This is the case for basins of attraction, while local and global optima only depend on the objective function and the neighborhood.) The search space characteristics of interest are the number of feasible states, the number of global and local optima and the value $rGBOA$. These values are reported in Table 1. The basins of attraction are defined with respect to deterministic iterative improvement. Observe that in all the cases, except for $n = 11, 13$, $rGBOA(\mathcal{M}_s) < rGBOA(\mathcal{M})$.

In Figures 3 and 4, we plotted the number of global (resp. local) optima versus n. We observe that the global optima have no apparent correlation with n, whilst the number of local optima seems to increase exponentially with n. (The relation between number of local optima and n can be fitted with a very good approximation by a line in a logarithmic plot.)

An interesting perspective of the search space can be given by plotting the ratio of the number of global (resp. local) optima to the search space size. This ratio is plotted in Figures 5 and 6. These plots show that the density of global optima decreases exponentially with n, while the density of local optima decreases much slower and, for the highest values of n, it is almost constant.[6]

In Figure 7, we plotted $rGBOA$ against n. We can note that this quantity decreases with n, approximately following a negative exponential. In Figure 8, $rGBOA$ is plotted against the number of global optima. From the plot we observe no (evident) correlation between the two quantities.

[6] The latter observation is in accordance with the plot in Figure 4 and the analysis of larger instances may enable us to model the asymptotic behavior.

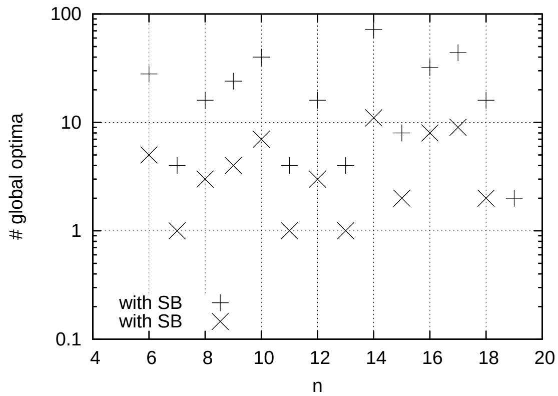


Figure 3. *Number of global optima versus instance size.*

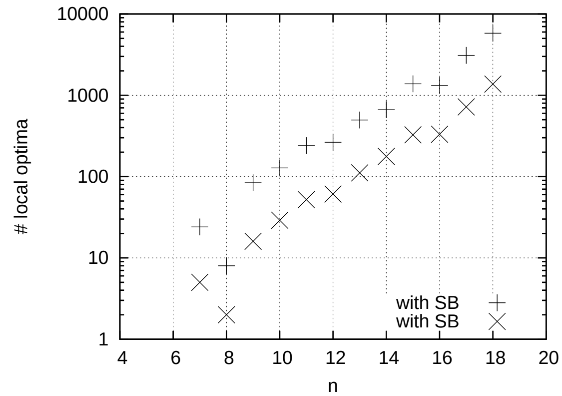


Figure 4. *Number of local optima versus instance size.*

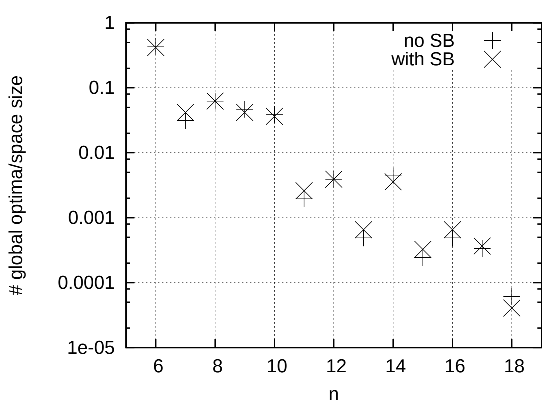


Figure 5. *Ratio of global optima w.r.t. search space size for* $n = 6, \ldots, 18$.

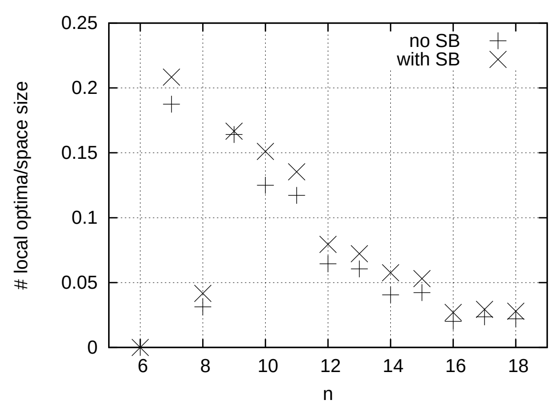


Figure 6. *Ratio of local optima w.r.t. search space size for* $n = 6, \ldots, 18$.

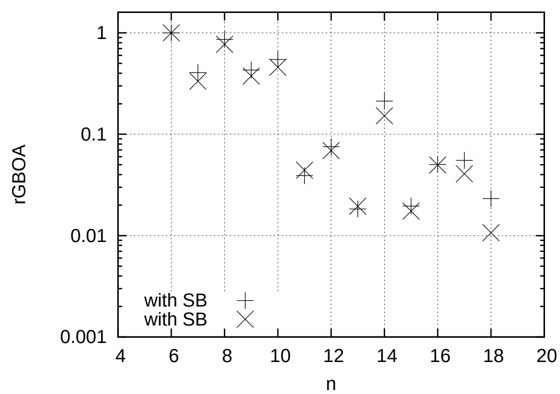


Figure 7. *Total size of the global optima basin of attraction (normalized to the number of feasible states) plotted against instance size.*

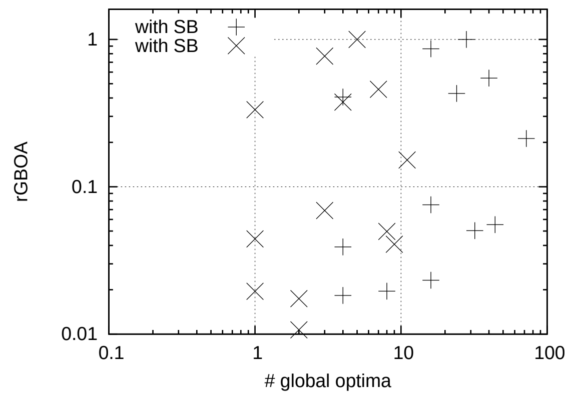


Figure 8. *Total size of the global optima basin of attraction (normalized to the number of feasible states) plotted against the number of global optima.*

4 Local search behavior

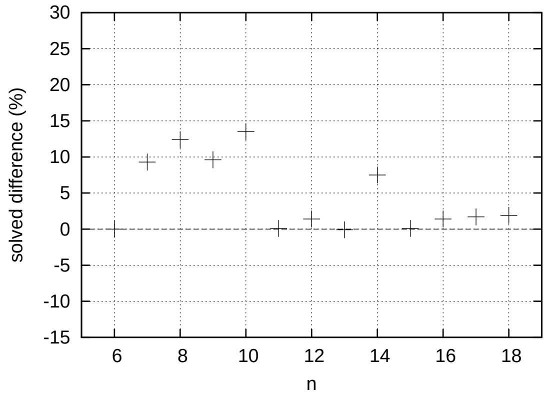
We attacked LABS (with $n = 6, \ldots, 18$) with four different local search algorithms: Best improvement with randomly broken ties (BI), First improvement with random order among neighbors (FI), Simulated annealing (SA) and Tabu search (TS).[7] From the perspective of search space exploration, the algorithms chosen exhibit a varying explorative attitude, starting from the lowest of BI to the highest of TS, while all keeping a 'greedy' character.[8] We run each algorithm on the original model and on the model with SB constraints. The algorithms are stopped after $10n$ non-improving moves. This termination condition enables us to compare the algorithms on the basis of the best solution the returned once a steady state is reached. (In the literature of metaheuristics, this state is also commonly called *stagnation*.) Table 2 gives a synoptic view of the algorithm performance in term of success ratio (out of 1000 runs).

A comparison of the performance of each algorithm on the two problem models is given in Figures 9, 10, 11, 12, in which we plotted the difference of solved instances (perc.) against n, i.e., $\Delta_{\%} = (\text{solved(noSB)}-\text{solved(SB)})/10$. Note that the performance on $\mathcal{M}$ dominates the one on $\mathcal{M}_s$ in all but the TS case.

The correlation between number of successes and $rGBOA$ is particularly interesting. From the plots in Figures 13, 14, 15 and 16, we observe that for BI and FI the number of successes is proportional to the size of the global optima basin of attraction. In the case of SA and TS, while the correlation is still observable, we note that the performance remains quite high even for low values of $rGBOA$, especially in the case of TS. The value of $rGBOA$ have been measured on the basis of deterministic best improvement, therefore it is not surprising that both BI and FI show a proportional relation between successes and fraction of states that make the search converging to a global optimum. SA performs a more effective search space

[7]The parameters of the two metaheuristics SA and TS have not been optimized. The initial temperature in SA has been set after a simple trial-and-test procedure. The tabu tenure in TS is randomly restarted each iteration in a range between 1 and $n/2$, in the spirit of robust tabu search [9].

[8]A deep discussion on this topic, involving also intensification and diversification, can be found in [2].



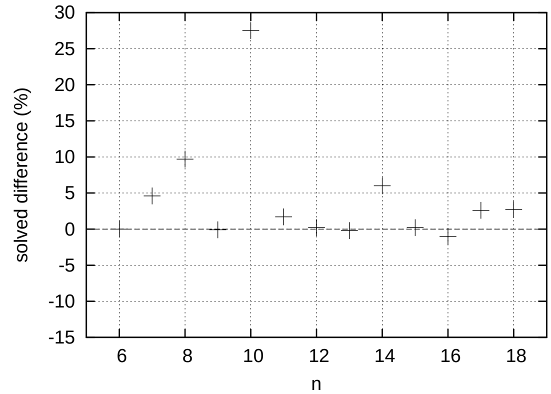
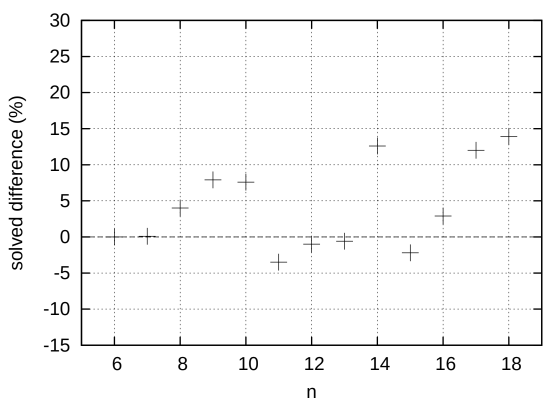


Figure 9. ***Best improvement***. *Percentage of the difference between solved instances in the original model and solved instances in the model with SB constraints.*

Figure 10. ***First improvement***. *Percentage of the difference between solved instances in the original model and solved instances in the model with SB constraints.*



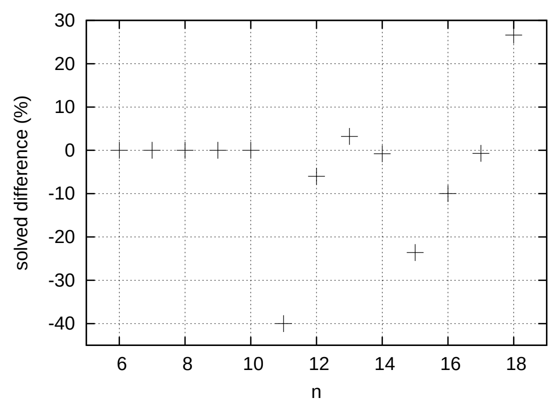


Figure 11. ***Simulated annealing***. *Percentage of the difference between solved instances in the original model and solved instances in the model with SB constraints.*

Figure 12. ***Tabu search***. *Percentage of the difference between solved instances in the original model and solved instances in the model with SB constraints.*

Table 2. *Synopsis of the number of solved instances (out of 1000 runs) of the four local search algorithms on the original model and the model with SB constraints.*

n	**BI**		**FI**		**SA**		**TS**	
	no SB	with SB	no SB	with SB	no SB	with SB	no SB	with SB
6	1000	1000	1000	1000	1000	1000	1000	1000
7	417	324	530	484	1000	999	1000	1000
8	875	751	825	728	1000	960	1000	1000
9	438	342	266	267	913	834	1000	1000
10	561	426	995	720	996	920	1000	1000
11	42	41	47	30	101	136	528	928
12	77	63	39	37	308	318	835	895
13	16	17	3	5	105	111	283	251
14	200	125	202	142	857	731	992	1000
15	25	24	22	20	157	179	363	599
16	66	52	29	39	336	307	819	919
17	60	43	71	45	508	388	909	916
18	25	6	35	8	270	131	678	412

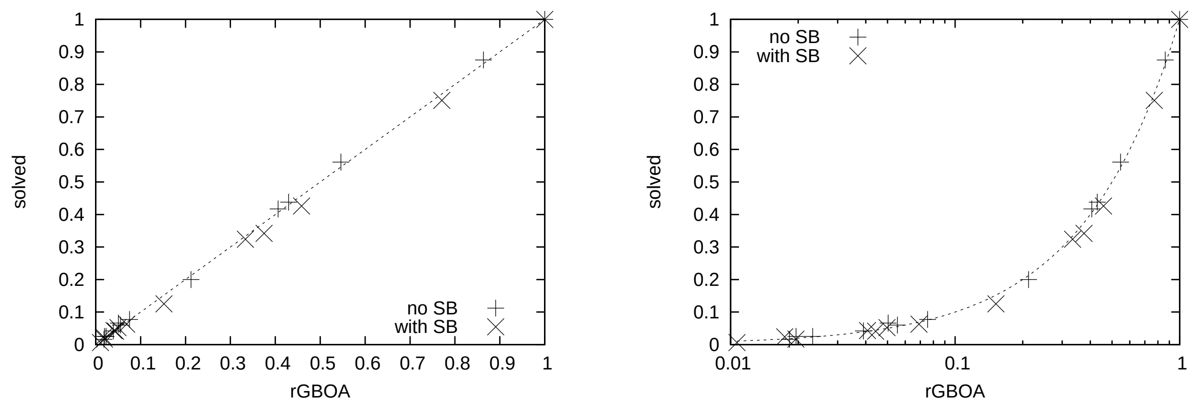


Figure 13. ***Best improvement**. Solved instances (perc.) plotted against the size of the global optima BOA. Linear scale (left) and log-scale (right).*

exploration than iterative improvement procedures, and even more TS, therefore the number of successes they achieve is much higher than that of BI and FI. It is interesting to note that the performance of both SA and TS starts to degrade (quite abruptly) when the normalized size of the global optima basin of attraction approaches a threshold value. Moreover, for TS this value is smaller than for SA, in other words, the more sophisticated an exploration strategy is, the lower the value of $rGBOA$ at which the performance starts to be strongly affected.

The relation between $rGBOA$ and local search performance can also be observed by plotting the ratio of solved instances $solved(\mathcal{M})/solved(\mathcal{M}_s)$ against the ratio of the basins of attraction $rGBOA(\mathcal{M})/rGBOA(\mathcal{M}_s)$. These graphs are plotted in Figures 17, 18, 19 and 20. As we can see, in the case of relatively simple local search procedures, such as BI and FI, the

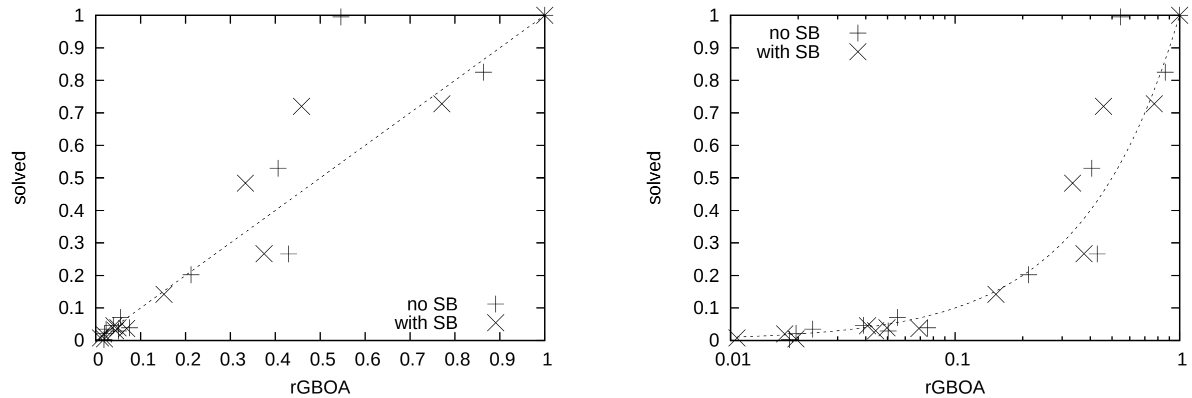


Figure 14. ***First improvement***. *Solved instances (perc.) plotted against the size of the global optima BOA. Linear scale (left) and log-scale (right).*

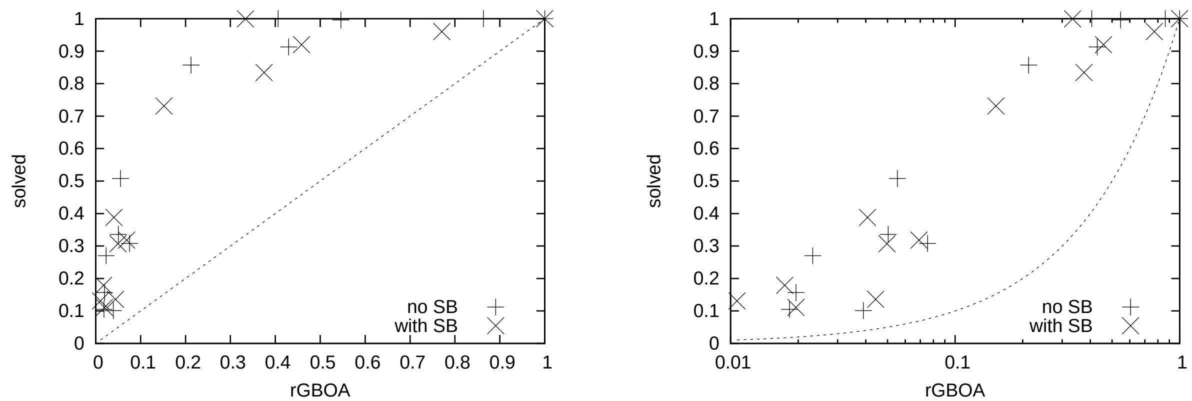


Figure 15. ***Simulated annealing***. *Solved instances (perc.) plotted against the size of the global optima BOA. Linear scale (left) and log-scale (right).*

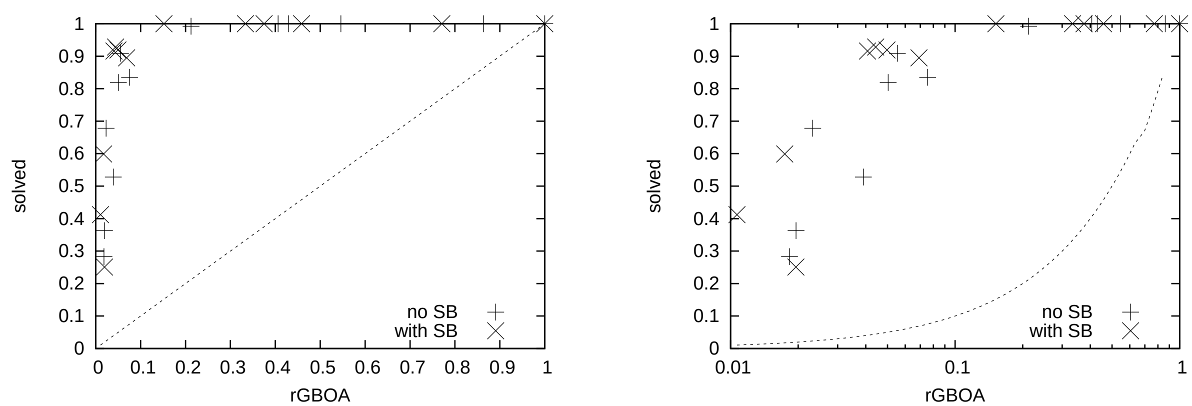


Figure 16. ***Tabu search***. *Solved instances (perc.) plotted against the size of the global optima BOA. Linear scale (left) and log-scale (right).*

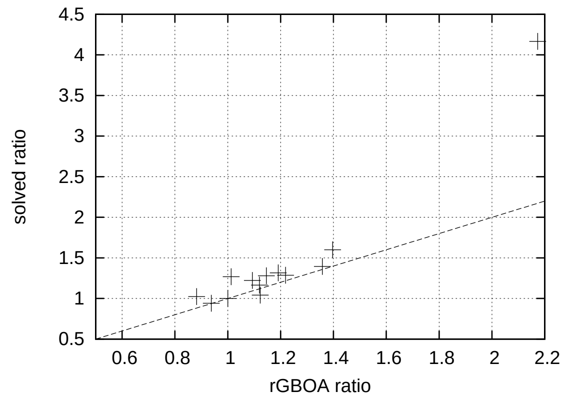


Figure 17. ***Best improvement***. *Ratio of solved instances vs. ratio of global optima basin of attraction size.*

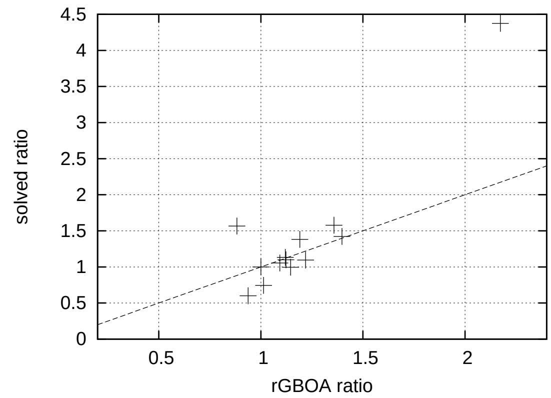


Figure 18. ***First improvement***. *Ratio of solved instances vs. ratio of global optima basin of attraction size.*

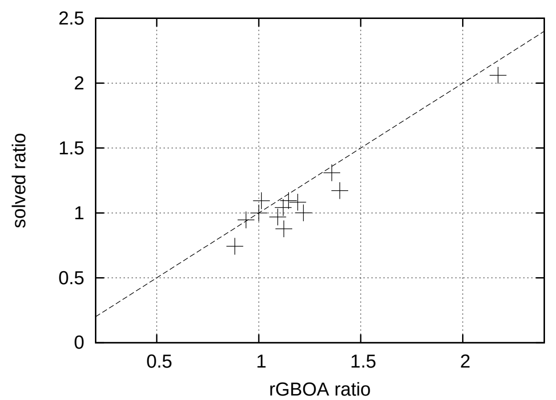


Figure 19. ***Simulated annealing***. *Ratio of solved instances vs. ratio of global optima basin of attraction size.*

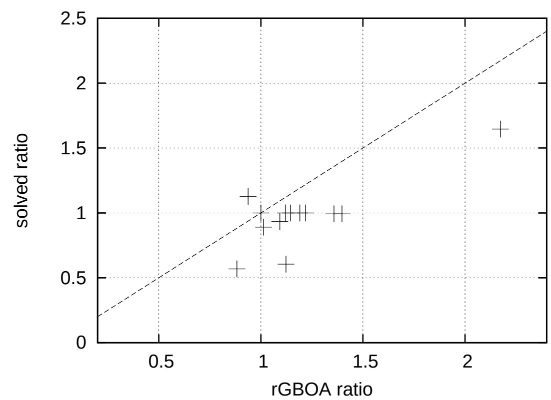


Figure 20. ***Tabu search***. *Ratio of solved instances vs. ratio of global optima basin of attraction size.*

points in the plot are approximately positioned around the line that bisects the quadrant. The cases of SA and TS diverge with respect to the previous results, providing evidence for the fact that the higher the exploration, the lower the negative impact of SB constraints.

5 Discussion

The available data are still not sufficient to draw strong conclusions on the subject, however we have experimental results to support our conjecture. First of all, it is apparent that $rGBOA$ is reduced in the model with SB constraints. Another important observation is that local search performance is strongly affected by the size of the global optima basin of attraction. This relation is in the form of a positive correlation (i.e., the smaller the BOA, the lower the performance) and it is well approximated by a linear relation in the case of simple local search algorithms (BI and FI), while it is nonlinear in the case of more complex search strategies (SA and TS). The nonlinearity of this relation plays a big role when we compare the performance of local search algorithms (in terms of success ratio). In fact, large differences in $rGBOA$ imply large deviations of the performance. But on the other side, when the difference is quite small, other factors come into play. Indeed, in some cases we can observe that the variance in the performance can not directly be attributed only to the size of global optima basin of attraction.

The reduction of $rGBOA$ and the correlation between performance and $rGBOA$ could give a first order empirical explanation of why SB constraints have been observed to be harmful for local search.

Finally, we have to note that the more complex the search strategy is, the more it could take advantage of the search space reduction, even if $rGBOA$ decreases. This issue should be investigated in detail, especially by experimenting with large size instances.

References

[1] Y. Bar–Yam. *Dynamics of Complex Systems.* Studies in nonlinearity. Addison–Wesley, 1997.

[2] C. Blum and A. Roli. Metaheuristics in combinatorial optimization: Overview and conceptual comparison. *ACM Computing Surveys*, 35(3):268–308, September 2003.

[3] James Crawford, Matthew L. Ginsberg, Eugene Luck, and Amitabha Roy. Symmetry-breaking predicates for search problems. In *KR'96: Principles of Knowledge Representation and Reasoning*, pages 148–159. Morgan Kaufmann, San Francisco, California, 1996.

[4] R. L. Devaney. *An introduction to chaotic dynamical systems.* Addison–Wesley, second edition, 1989.

[5] S. Mertens. Exhaustive search for low-autocorrelation binary sequeneces. *J. Phys. A*, 29:L473–L481, 1996.

[6] S. Prestwich. First-solution search with symmetry breaking and implied constraints. In *Proc. of CP'01 Workshop on Modelling and Problem Formulation*, 2001.

[7] S. Prestwich. Supersymmetric modeling for local search. In *Proc. of SymCon'02 Workshop on Symmetry and Constraint Satisfaction Problems*, 2002.

[8] J. F. Puget. Symmetry breaking revisited. In *Principle and Practice of Constraint Programming – CP02*, volume 2470 of *Lecture Notes in Computer Science*, pages 446–451. Springer-Verlag, 2002.

[9] E. Taillard. Robust taboo search for the quadratic assignment problem. *Parallel Computing*, 17:443–455, 1991.