Computer Physics Communications 185 (2014) 1639–1646

Contents lists available at ScienceDirect

Computer Physics Communications

journal homepage: www.elsevier.com/locate/cpc

Efficient implementation and application of the artificial bee colony algorithm to low-dimensional optimization problems

Guido Falk von Rudorff[a], Christoph Wehmeyer[a,b], Daniel Sebastiani[a,b,*]

[a] Dahlem Center for Complex Quantum Systems, Freie Universität Berlin, Arnimallee 14, 14195 Berlin, Germany

[b] Institute of Chemistry, Martin-Luther-Universität Halle-Wittenberg, von-Danckelmann-Platz 4, 06120 Halle, Germany

ARTICLE INFO

Article history:
Received 23 October 2013
Received in revised form
6 January 2014
Accepted 7 March 2014
Available online 18 March 2014

Keywords:
Swarm-intelligence
Molecular clusters
Parallelization

ABSTRACT

We adapt a swarm-intelligence-based optimization method (the artificial bee colony algorithm, ABC) to enhance its parallel scaling properties and to improve the escaping behavior from deep local minima. Specifically, we apply the approach to the geometry optimization of Lennard-Jones clusters. We illustrate the performance and the scaling properties of the parallelization scheme for several system sizes (5–20 particles). Our main findings are specific recommendations for ranges of the parameters of the ABC algorithm which yield maximal performance for Lennard-Jones clusters and Morse clusters. The suggested parameter ranges for these different interaction potentials turn out to be very similar; thus, we believe that our reported values are fairly general for the ABC algorithm applied to chemical optimization problems.

1. Introduction

Structure optimization has been a matter of interest for chemical physics and physical chemistry [1–7], as the equilibrium structure of a system is most important for describing the physical and chemical properties [2]. In particular, molecular clusters have been studied in detail [3,8–18]. Even for the most simple interaction potentials, the potential energy surface (PES) for clusters with more than four atoms is highly complicated and exhibits multiple non-connected minima. As the global minima of these cluster types cannot be found analytically, various methods have been developed and applied: artificial bee colony [19–22], basin-hopping [23,12,24], dynamic lattice search [25], genetic algorithm [26,27], Monte Carlo search [28], parallel tempering [29], particle swarm optimization [21,30], simulated annealing [21,31–33], swarm intelligence [34], tabu search [35–37] and minima hopping [38], which can be considered to be the most common approach for cluster optimization.

The artificial bee colony (ABC) algorithm has been used for various optimization problems from chemical physics [39,40], engineering [41,42], and computer science [43–46]. Recently, we presented a modification of the artificial bee colony (ABC) algorithm for the optimization of molecular geometries [47].

The main idea of the ABC algorithm is to search the PES using a swarm intelligence approach. Several replications of a random conformation, called foragers, are placed on the PES and search for local minima. The computational resources are distributed dynamically among these foragers based on their relative energies, such that foragers in energetically favorable regions get more computational resources. This stochastic element is crucial for the performance and efficiency of the ABC approach.

When applied to clusters, the scaling of the ABC algorithm is exponential w.r.t. to the cluster size. As large clusters are more interesting, we want to push the limit of the computationally accessible search space. With increasing cluster size, the optimization is only feasible with nearly optimal values for the internal parameters of the algorithm, so an understanding of the parameters of the ABC algorithm is vitally important. Therefore, we analyze the reaction of the algorithm on parameter changes and derive a set of valid parameter ranges that both helps ensuring convergence in the first place and yields good performance. For the parameter determination, we used the Lennard-Jones potential

$$E_{\mathrm{LJ}} = 4\varepsilon \sum_{i<j} \left[\left(\frac{\sigma}{r_{ij}} \right)^{12} - \left(\frac{\sigma}{r_{ij}} \right)^{6} \right] \tag{1}$$

in reduced units, that is distances in σ and energies in ε. For evaluation, we optimized cluster structures for both the Lennard-Jones potential and the Morse potential with its range-determining parameter a

$$E_{\mathrm{M}} = \varepsilon \sum_{i<j} [\exp(2a(1 - r_{ij}/r_0)) - 2\exp(a(1 - r_{ij}/r_0))]. \tag{2}$$

* Corresponding author at: Dahlem Center for Complex Quantum Systems, Freie Universität Berlin, Arnimallee 14, 14195 Berlin, Germany. Tel.: +49 3037447445.
E-mail address: daniel.sebastiani@chemie.uni-halle.de (D. Sebastiani).

http://dx.doi.org/10.1016/j.cpc.2014.03.006
0010-4655/

As the global minima are well-known for the Lennard-Jones and Morse potentials, it is easily possible to define a convergence criterion based on known reference solutions. For applying the ABC algorithm to problems without any known solution, we derive two intrinsic convergence criteria from extensive statistics.

For larger clusters or more complex problems, the computational costs increase quickly, so an efficient parallelization strategy is mandatory. As neither of the previous versions [47,19] of the ABC algorithm is parallelized, this manuscript reports two parallelization strategies and evaluate their impact on the convergence behavior.

The potential energy surface often exhibits narrow funnels that are difficult to find by MD at finite temperatures. These funnels may be of particular interest like, for example, misfolded non-equilibrium structures in protein folding. For finite temperatures, the Maxwell–Boltzmann-distribution in principle requires the knowledge of all geometries, especially those with a minimum energy such that the according Maxwell–Boltzmann factors can be calculated. Therefore, having different methods at hand for finding minimal energy conformations is desirable. In this paper, we use molecular clusters with phenomenological potentials in order to evaluate and tune the artificial bee colony algorithm.

The paper is divided into three main sections: Methods, Results and Discussion. Each of which itself is divided into three subsections on the targets of this paper: the suggested parameter ranges, convergence criteria, and parallelization strategies. Finally, the information on these separate parts is evaluated together with respect to both limitations and capabilities.

2. Methods

2.1. Working principles

The ABC algorithm basically tries to find a point with lower energy on the PES of a cluster by deriving candidate conformations from the point with the lowest energy found so far, with an additional stochastic component. Up to now, the ABC algorithm has been tested on both purely mathematical test functions [48] and cluster structures [47]. In both cases, the costs for the evaluation of each candidate conformation are very small. In this paper, we measure the computational costs by counting the number of energy calculations for a given conformation, that is, the number of single point calculations (SPC).

The total optimization process is divided into cycles [20]. Each cycle begins with an employee bee phase and ends with an onlooker bee phase, optionally followed by a scout bee phase. The employee bees hold a candidate conformation and know about the best solution of all candidate conformation they tested so far, whereas the O onlooker bees have access to the information of all employed bees. In the employee phase, each employee bee derives a new candidate conformation either by generating a totally random one or by choosing a point in the search space that is inside a hypersphere centered around the current conformation of the employed bee. In general, only if the new conformation has a lower energy than the previous one, the employed bee adapts the new position. If the employed bee discards a new position multiple times in a row, the radius of the hypersphere, the current size level, is halved. In the case an employed bee found no better position for a certain period, it becomes a scout bee and is forcibly set to a random point on the PES, where it becomes an employed bee again. This scout step is important for the sampling efficiency of the algorithm. After all employed bees have finished, the onlooker phase starts: each onlooker bee selects one employed bee based on a probability distribution defined by the energies of the current worker conformations. In this way, employed bees with a lower energy are preferred over those with higher energies. Each selected employed bee searches for a new position again. One employed bee can be selected by multiple onlooker bees. After all onlookers have completed, the cycle is finished.

In our implementation, we use implicit onlookers. This means that the onlooker bees are modeled by allowing the employed bees to perform another random local search step. In order to stress this difference, we use the expression *worker* for the foragers, disregarding whether they are currently used as employed bee, onlooker bee or scout bee during the optimization process. In other words, the number of workers W denotes the number of different positions on the PES that is kept in memory at any time.

During the onlooker phase, the algorithm allows each worker i of the W workers to perform another $S(W_i)$ steps. The distribution among the threads in the parallelized versions is given by $S(W_i, T_j)$ where T_j denotes the jth thread from t threads in total, so that

$$S(W_i) = \sum_{j=1}^{t} S(W_i, T_j) \tag{3}$$

$$S(T_i) = \sum_{j=1}^{O} S(W_j, T_i). \tag{4}$$

We assume t to be less than or equal to the number of available parallel compute nodes.

2.2. Suggested parameter ranges

In principle, an optimal set of values for the free parameters of the algorithm can be obtained for any cluster type (for example Lennard-Jones [49], Morse [50], Tersoff [51,52], and TIP5P [53]) and cluster size [54,55]. However, these settings would only be of interest for this particular application and would not be usable in order to cope with other optimization problems. Therefore, we focus on describing ranges for the algorithmic parameters. These ranges are further referred to as suggested parameter ranges (SPR) and are defined by two criteria: at least 90% of all N runs converge and the average costs are not higher than the threshold of 110% of the minimal average costs for this cluster. Although this tight criterion introduces some special values like the data for $e = 17$ in Fig. 3, it is helpful to determine the actual optimal parameter values and supports the transferability of the suggested parameter ranges.

Determining these SPR is an optimization problem itself. Therefore, we have chosen to apply the ABC algorithm to find these parameter ranges. For the target quantity that is to be optimized, we average the computational cost for reproducing the known optimal conformation over $N \simeq 500$ runs with a specific parameter set. Regarding the computational costs, Lennard-Jones clusters with five to twenty atoms are used as reference solutions [56]. Using the reduced unit representation of the Lennard-Jones potential guarantees validity of the results for all parametrizations of the Lennard-Jones potential.

Starting from the best parameter set for each cluster size e, five out of the six parameters are kept constant and the sixth one is used for the rasterization of the local minimum in parameter space. Again, each data point is calculated by averaging over N optimization runs. The resulting data is smoothed in order to be able to define a minimum from the noised information.

Fig. 1 illustrates the procedure of determining the SPR. First of all, a parameter set has to be found by the master optimization runs. Varying the number of onlookers leads to the raw data. As each of the data points may vary in spite of the averaging, the raw data is smoothed by calculating a Bézier curve, which in turn is used for SPR determination.