Exploring molecular flexible docking problems with evolutionary algorithm

Yi Fu[1,2]
1. School of Internet of Things Engineering,
Wuxi City College of Vocational Technology
2. WuXi Research Center for
Environmental Science & Engineering
Wuxi, China
fuyi@wxcu.edu.cn

Ji Zhao[1,2]
1. School of Internet of Things Engineering,
Wuxi City College of Vocational Technology
2. WuXi Research Center for
Environmental Science & Engineering
Wuxi, China
zhaoji@wxcu.edu.cn

***Abstract*—In quantum-behaved particle swarm optimization algorithm (QPSO), the mean best position is employed as a global attractor to attract the particles for searching solutions globally. To improve the QPSO global convergence performance further, this article proposes an improved QPSO algorithm based on joint modeling of individual particles evolutionary process (IEQPSO). To ascertain the effectiveness of the proposed variant of the QPSO algorithm, several benchmark test functions have been considered. The experimental results show the superiority of the proposed approach on benchmark test functions. To assess the effectiveness and feasibility of the proposed method on real problems, it was used for the energy optimization of molecular docking, and compared with the classical Lamarckian genetic (LGA) algorithm. The numerical results reveal the reliability of the proposed approach for sampling conformation under high flexibility rotalable bonds.**

Keywords- evolutionary algorithm; objective optimization; joint modeling; molecular docking

I. INTRODUCTION

Among the many evolutionary algorithms, quantum-behaved particle swarm optimization algorithm (QPSO) is more efficient than many conventional algorithms. Sun et al. [1] proposed QPSO which is based on a quantum delta potential model. It is inspired by quantum mechanics and trajectory analysis of PSO [2]. The signifant difference between QPSO and PSO is QPSO introduced the mean best position into the algorithm. The QPSO algorithm has been shown to successfully solve a wide range of optimization problems. Moreover many efficient strategies also have been proposed to improve the algorithm performance [3-5]. In this article, to improve the performance of QPSO on highly dimension complex problems, we propose an improved QPSO based on joint modeling of individual particles evolutionary process (IEQPSO) algorithm to improve the global search capability and the convergence performance of the optimization process. The benchmark functions were used to test the performance of IEQPSO algorithm.

The ultimate aim of optimization algorithm is to solve the reality problems. In this article, we use the IEQPSO to optimize energy search function of molecular docking. Molecular docking programs are one of the most widely used methods in drug development [6]. They can be used for predicting the correct geometry of a ligand-protein complex, for identification of novel lead compounds and preliminary lead optimization [7] and for investigating mechanisms of action of biologically active compounds [8]. The energy search of molecular docking is a complex combinatorial optimization problem.

The rest of the paper is organized as follows: in section 2 describes the the proposed method, section 3 presents the experimental results and section 4 contains the conclusion.

II. ALGORITHM

In a QPSO with M individuals, each individual is treated as a spin-less one moving in the N-dimensional quantum space. The mean best position (*mbest*) C_n is defined by the average of the personal best (*pbest*) positions of all particles.

$$C_n = (1/M)\sum_{i=1}^{M} P_{i,n} \qquad (1)$$

for i = 1,2,...,M, where $P_{i,n}$ is the personal best (*pbest*) position of particle i.

From Eq. (1), the mean best position is the average on the personal best position of all particles, which means that each particle is considered equal and has the same influence on the value of mean best position. In fact, it is not properly to consider each member equal. Here, we consider the evolutionary process of each particle into the mean best position for the QPSO. The evolutionary process of each particle is based on the joint modeling between the fitness value of the global best solution $f(G_n)$ and the value of the personal best position of the current particle $f(P_{i,n})$, and given by Eq. (2):

$$\alpha_{i,n} = \left| \frac{f(G_n)}{f(P_{i,n})} \right| \qquad (2)$$

From Eq.(2), not only is this reflects the individual particle evolutionary rate, but it also helpful for us to gain information about the importance of particles in population when they are evolving. Then, by exerting the above parameter control, the mean best position equation is as follows:

$$C_n = (1/M)\sum_{i=1}^{M} \alpha_{i,n} \cdot P_{i,n} \qquad (3)$$

978-1-5386-7445-1/18/$31.00 ©2018 IEEE
DOI 10.1109/DCABES.2018.00069

III. EXPERIMENTAL RESULTS AND ANALYSIS

A. *Performance evaluation and comparison on benchmark functions*

To evaluate the overall performance of the proposed IEQPSO on function optimization, the algorithm was tested on eight well-known benchmark functions (Table 1). For performance comparison, PSO with inertia weight (PSO-In), the original QPSO algorithm and the QPSO with weighted mean best position (WQPSO) [10] were also tested by the benchmark functions. Each algorithm ran 50 times on each problem using 40 particles to search the global best fitness value with each run executed for 2000 iterations. In PSO-In, the inertia weight is 0.9 ~ 0.4, two coefficients c1 and c2 are set 2. In QPSO and its variant, the contraction-expansion coefficient is set 1.0 ~ 0.5. The mean best fitness value and standard deviation out of 50 runs of each algorithm on each problem is presented in Table 2. The standard deviation values are reflected in brackets.

TAB.1. Benchmark functions.

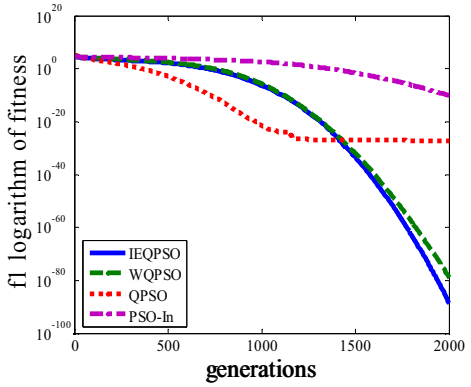
Functions	Formulations	Feasible solution space
Sphere function	$f_1(x)=\sum_{i=1}^{n} x_i^2$	(-100, 100)
Schwefel's 222 function	$f_2(x)=\sum_{i=1}^{n}\lvert x_i\rvert+\prod_{i=1}^{n}\lvert x_i\rvert$	(-10, 10)
DeJong function	$f_3(x)=\sum_{i=1}^{n} i x_i^4$	(-100, 100)
Rosenbrock function	$f_4(x)=\sum_{i=1}^{n-1}(100(x_{i+1}-x_i^2)^2+(x_i-1)^2)$	(-30, 30)
Ackley function	$f_5(x)=20+e-20e^{\frac{1}{5}\sqrt{\frac{1}{n}\sum_{i=1}^{n}x_i^2}}-e^{\frac{1}{n}\sum_{i=1}^{n}\cos(2\pi x_i)}$	(-32, 32)
Rastrigin function	$f_6(x)=\sum_{i=1}^{n}(x_i^2-10\cos(2\pi x_i)+10)$	(-5.12, 5.12)
Schaffer function	$f_7(x)=0.5+\frac{(\sin(\sqrt{x^2+y^2}))^2-0.5}{(1.0+0.001(x^2+y^2))^2}$	(-2.048, 2.048)
Griewank function	$f_8(x)=\frac{1}{4000}\sum_{i=1}^{n}x_i^2-\prod_{i=1}^{n}\cos(\frac{x_i}{\sqrt{i}})+1$	(-600, 600)

For Shifted Sphere Function (f1), all the algorithms could get a solution near the global optimal solution, but IEQPSO got better results than other methods. The results for shifted Schwefel's 222 problem (f2) show that, IEQPSO could find global optimum successfully and exhibited its great search ability in comparison to the PSO-In and QPSO. For DeJong function (f3), the WQPSO algorithm outperformed the other methods. For benchmark f4, Rosenbrock function is multi-modal function, which is too hard to be solved for many optimization problems, IEQPSO showed a little better performance among all the algorithms and its obvious advantage over other algorithms is the standard deviation values. The fifth benchmark function f5 is Ackley function with global optimum on bounds. Results indicate that all the QPSO-based methods showed a better performances for this problem than PSO. For function f6, Rastrigrin function, the IEQPSO algorithm yielded the best result. In this function, the difference between QPSO and PSO was not significant. Results obtained for f7, Schaffer problem, IEQPSO had the significant advantage over the other methods. The results for Shifted Rotated Griewank's Function without Bounds (f8) suggest that IEQPSO was able to find the better solution than others.

TAB.2. Results for the benchmark functions.

Functions	PSO-In	QPSO	WQPSO	IEQPSO
f1	1.192430e-010	2.606978e-029	1.550875e-079	**9.525978e-086**
	(2.572647e-010)	(1.123051e-028)	(4.496400e-079)	**(4.203196e-085)**
f2	3.206959e-008	4.624757e-018	2.408737e-049	**1.669145e-052**
	(3.499209e-008)	(1.591423e-017)	(3.648797e-049)	**(3.084726e-052)**
f3	8.048631e-010	6.201183e-037	**2.146062e-109**	4.094578e-092
	(3.613984e-009)	(3.341504e-036)	**(1.037800e-108)**	(2.636037e-091)
f4	7.814176e+001	4.378334e+001	6.730166e+001	**2.361887e+001**
	(4.647823e+001)	(3.173342e+001)	(4.943229e+001)	**(2.212292e-001)**
f5	2.038785e-005	1.541008e-009	**1.540960e-009**	1.540978e-009
	(1.017194e-004)	(3.804749e-014)	**(8.437228e-016)**	(8.220215e-015)
f6	3.285426e+001	2.152986e+001	9.381218e+001	**1.783571e-021**
	(8.221160e+000)	(5.855429e+000)	(2.941288e+001)	**(1.354936e-020)**
f7	3.886364e-004	5.830047e-004	3.902530e-004	**5.5511e-018**
	(1.903922e-003)	(2.307388e-003)	(1.903624e-003)	**(3.4528e-016)**
f8	1.312913e-002	6.597326e-003	6.254397e-003	**8.881784e-018**
	(1.435475e-002)	(8.014635e-003)	(1.140907e-002)	**(6.217249e-017)**

Figure 1 presents the convergence process of the algorithms on each benchmark function. It is shown that the IEQPSO algorithm had the better convergence property than their competitors during the later stage of iteration except for the functions f3 and f5. According to the results above in the tables and figures, it can be found that IEQPSO was able to find the solution of higher quality for the function compared to the other methods in most cases.



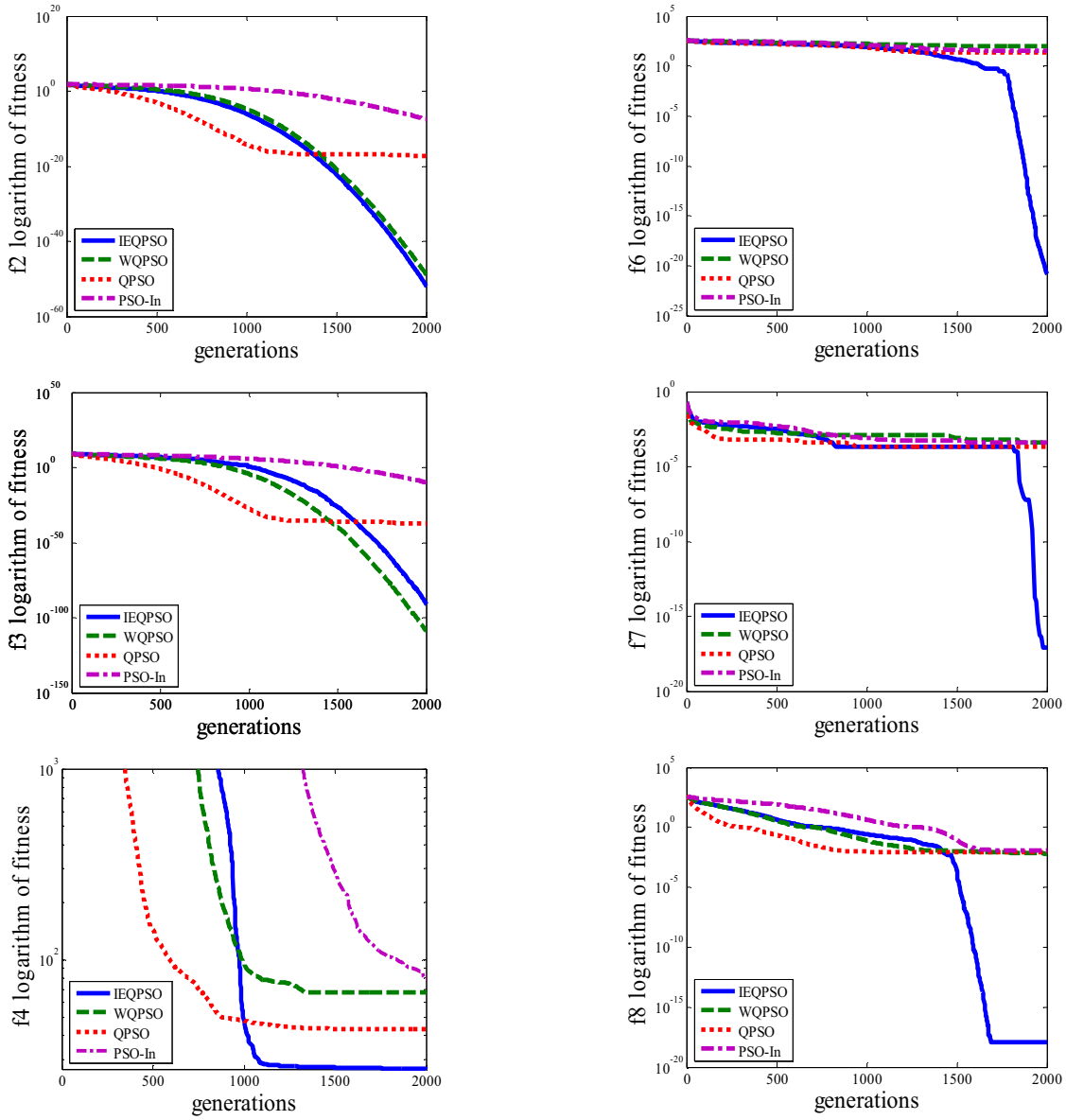


FIG.1 a-h. Convergence Process of the algorithms on each benchmark problem. (a) Shifted Sphere function (f1). (b) Schwefel's 222 function (f2). (c) DeJong function (f3). (d) Rosenbrock function (f4). (e) Ackley function (f5). (f) Rastrigrin function (f6). (g) Schaffer function (f7). (h) Griewank function (f8).

B. Molecular docking energy function

A semi-empirical free energy scoring function was used to evaluate a docked conformation in AutoDock 4 [10]. The force field consists of six pair-wise evaluations (V) and an estimate of the conformational entropy lost upon binding (ΔS_{conf}):

$$\Delta G=\left(V_{bound}^{L-L}-V_{unbound}^{L-L}\right)+\left(V_{bound}^{P-P}-V_{unbound}^{P-P}\right)+\left(V_{bound}^{P-L}-V_{unbound}^{P-L}+\Delta S_{conf}\right) \quad (4)$$

where P refers to the "protein" and L refers to the "ligand" in a protein-ligand docking calculation.

Each of the pair-wise energetic terms is defined by the following energy: a Lennard-Jones 12-6 Van der Waals interaction, a 12-10 hydrogen bond potential, a coulombic electrostatic potential and a desolvation potential.

$$V=W_{vdw}\sum_{i,j}\left(\frac{A_{ij}}{r_{ij}^{12}}-\frac{B_{ij}}{r_{ij}^{6}}\right)+W_{hbond}\sum_{i,j}E(t)\left(\frac{C_{ij}}{r_{ij}^{12}}-\frac{D_{ij}}{r_{ij}^{10}}\right)+W_{elec}\sum_{i,j}\frac{q_iq_j}{\varepsilon(r_{ij})r_{ij}}+W_{sol}\sum_{i,j}\left[S_iV_j+S_jV_i\right]e^{\left(-r_{ij}^2/2\sigma^2\right)} \quad (5)$$

In molecular docking, ligands with torsions >10 were highly flexible. Here, for molecular docking test cases, we choice 1b58 and 1ivq which have 19 and 16 torsions, respectively. Molecular docking software AutoDock4 was used, the results of docking were comppared with LGA which is the main algorithm for AutoDock4. The docking energy and the corresponding RMSD value for IEQPSO and LGA were shown in Table 3 and Figure 2.

TAB.3. Comparison of the lowest docking energy and the corresponding RMSD value of IEQPSO and LGA.

PDB	Torsions	IEQPSO		LGA	
		Energy	RMSD	Energy	RMSD
1b58	19	-19.68	1.63	-12.74	2.00
1ivq	16	-14.63	1.71	-7.25	1.63

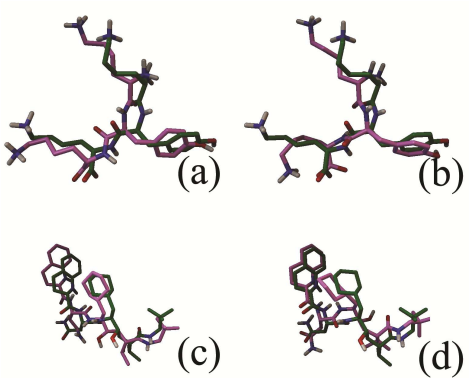


FIG.2. a-d. Docking examples of two complexes 1b58 ((a) IEQPSO, (b) LGA.) and 1ivq ((c) IEQPSO, (d) LGA.). The color of the predicted pose is pink and the crystal pose is green.

IV. CONCLUSION

This article describes an improved quantum behaved particle swarm optimization and its application in molecular docking energy optimization problem. To assess the performance of the IEQPSO, eight test cases consisting of the unimodal functions and multimodal functions have been taken for comparison. Statistical measures calculated for the optimization algorithms show that the IEQPSO algorithm significantly improves the search ability and convergence precision than other compared algorithms. For molecular docking test cases 1b58 and 1ivq, the results demonstrate that the proposed IEQPSO is feasibility for high-dimensional flexibility docking problem.

ACKNOWLEDGMENT

This work is supported by the National Natural Science Foundation of China under Grant Nos. 61502203, the Natural Science Foundation of Jiangsu Province under Grant Nos. BK20150122, the Natural Science Research Project of Jiangsu Province University under Grant Nos. 17KJB520039.

REFERENCES

[1] Sun J, Feng B, Xu W. Particle swarm optimization with particles having quantum behavior. IEEE Congress on Evolutionary Computation. 2004; 1: 325-331.

[2] Clerc M, Kennedy J. The particle swarm-explosion, stability, and convergence in a multidimensional complex space. IEEE T. Evolut.Comput. 2002; 6(1): 58-73.

[3] Lin L, Guo F, Xie X, Luo B. Novel adaptive hybrid rule network based on TS fuzzy rules using an improved quantum-behaved particle swarm optimization. Neurocomputing. 2015; 149: 1003-1013.

[4] Liu T, Jiao L, Ma W, Shang R. Quantum-behaved particle swarm optimization with collaborative attractors for nonlinear numerical problems. Commun. Nonlinear Sci. 2016; 44 : 167-183.

[5] Xue T, Li R, Tokgo M, Ri J, Han G. Trajectory planning for autonomous mobile robot using a hybrid improved QPSO algorithm. Soft Comput. 2015; 1-17.

[6] Di Muzio E, Toti D, Polticelli F. DockingApp: a user friendly interface for facilitated docking simulations with AutoDock Vina. Journal of Computer-Aided Molecular Design, 2017, 31(2): 213-218..

[7] Kawada H, Ebiike H, Tsukazaki M, et al. Lead optimization of a dihydropyrrolopyrimidine inhibitor against phosphoinositide 3-kinase (PI3K) to improve the phenol glucuronic acid conjugation. Bioorg. Med. Chem. Lett. 2013; 23: 673-678.

[8] Sengupta S, Obiorah I, Maximov PY, Curpan R, Jordan VC. Molecular mechanism of action of bisphenol and bisphenol A mediated by oestrogenreceptor alpha in growth and apoptosis of breast cancer cells. Br. J. Pharmacol. 2013; 169:167-178.

[9] Morris GM, Huey R, Lindstrom W, et al. AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility. J. Comput. Chem. 2009; 30(16): 2785-2791.

[10] Xi M, Sun J, Xu W. An improved quantum-behaved particle swarm optimization algorithm with weighted mean best position. Appl. Math. Comput. 2008; 205: 751-759.