The Research of Evolutionary Computation in Medical and Engineering Applications

by

Zhenyu Lei

A dissertation

submitted to the Graduate School of Science and Engineering for Education

in Partial Fulfillment of the Requirements

for the Degree of

Doctor of Engineering

Supervisor: Prof. Zheng Tang

Associate Supervisor: Associate Prof. Shangce Gao

University of Toyama

Gofuku 3190, Toyama-shi, Toyama 930-8555 Japan

2022

(Submitted December 13, 2022)

Acknowledgements

Time is running out, and my doctoral career is coming to an end. In the past three years, I have gained deep friendships and learned a lot of professional knowledge. I would like to express my gratitude to all those who helped me during my research and study period.

First of all, I would like to thank my supervisor Prof. Zheng Tang who led me into the colorful and meaningful field of artificial intelligence, for his support and encouragement. During my study, his valuable advice brought me a lot of design inspiration and ideas. Without his valuable advice, this study will not be so smooth.

Then, to Associate Professor Shangce Gao, who guided me into particular research and taught me a conscientious and responsible attitude in work. He provides a large number of pertinent comments and additional resources in medical and engineering application research. Without his thoughtful help, I cannot obtain my current academic performance.

Next, to all the members of the Artificial Intelligence Laboratory in the University of Toyama. Thank them for their care and help in both study and life during my doctorial period. Hereby, I hope that everyone will live happily and everything will go well.

Finally, I would like to thank my family for all of their support and encouragement throughout my study process. Their unconditional love gives me the courage to face setbacks and overcome difficulties.

Abstract

Optimization problems widely exist now in human living, scientific research, and industry. With their development, the complexity of optimization problems exponentially increases. Meanwhile, traditional exact methods are unable to provide satisfactory solutions due to their non-linearity and non-convexity. Then, evolutionary computation inspired by natural biology is proposed to handle the problems. It includes different algorithms such as genetic algorithm (GA), particle swarm optimization (PSO), gravitational search algorithm (GSA), and differential evolution (DE), and has been successfully used to solve real-world applications (e.g., medical and engineering applications). However, these applications are usually non-linearity, non-differentiable, and multi-peaks resulting in evolutionary algorithms still suffering from the issues of low performance, local optima, and premature convergence. Researchers balance between the exploitation and exploration of algorithms to improve their performance via new operator factors, self-adaptive parameters, and new learning schemes. In this thesis, I propose evolutionary algorithms to solve real-world problems (medical and engineering applications). I proposed a many-objective algorithm to solve protein structure prediction problem. I also proposed an improved genetic learning particle swarm optimizer to optimize wind farm layout optimization problems. These are introduced as follows.

(1) Protein structure prediction (PSP) problems are a major biocomputing challenge, owing to its scientific intrinsic that assists researchers to understand the relationship between amino acid sequences and protein structures, and to study the function of proteins. Although computational resources increased substantially over the last decade, a complete solution to PSP problems by computational methods has not yet been obtained. Using only one energy function is insufficient to characterize proteins because of their complexity. Diverse protein energy functions and evolutionary computation algorithms have been extensively studied to assist in the prediction of protein structures in different ways. Such algorithms are able to provide a better protein with less computational

resources requirement than deep learning methods. For the first time, this study proposes a many-objective protein structure prediction (MaOPSP) problem with four types of objectives to alleviate the impact of imprecise energy functions for predicting protein structures. A many-objective evolutionary algorithm (MaOEA) is utilized to solve MaOPSP. The proposed method is compared with existing methods by examining thirty-four proteins. An analysis of the objectives demonstrates that our generated conformations are more reasonable than those generated by single/multi-objective optimization methods. Experimental results indicate that solving a PSP problem as an MaOPSP problem with four objectives yields better protein structure predictions, in terms of both accuracy and efficiency.

(2) The wind farm layout optimization (WFLO) problem optimizes the location of wind turbines in a wind farm to reduce the wake effect and improve maximum power generation. Traditional mathematical methods cannot provide a satisfactory solution for a wind farm due to the high complexity of the problem. Therefore, mete-heuristic algorithms have been used to optimize it. Genetic algorithms (GA) have been widely used and obtained success in WFLO problems. However, GA still suffers from the issues of insufficient optimization efficiency. In this study, a genetic learning particle swarm optimization with an adaptive strategy, termed AGPSO, is proposed to optimize WFLO problems. The strategy adaptively adjusts the location of the worst turbine to improve the conversion efficiency of a wind farm. Four wind scenarios, including single wind speed with single wind direction, single wind speed with uniform multiple wind directions, single wind speed with nonuniform multiple directions, and multiple wind speeds with multiple wind directions, are utilized to verify the effectiveness of AGPSO and the effect of different wind scenarios for it. Twelve constraints and three different scales are used to further verify the robustness of AGPSO and the effect of wind turbine location on WFLO problems. Extensive experimental results demonstrate that AGPSO performs significantly better than other eight state-of-the-art competitors in terms of conversion efficiency under different wind farms, wind scenarios, and constraints.

The thesis is organized as follows. Chapter 1 describes a many-objective protein structure prediction problem and proposes a many-objective algorithm to solve it. Chapter 2 presents an adaptive replacement strategy-incorporated particle swarm optimizer for WFLO problems.

Contents

List of Figures

List of Tables

Chapter 1

MO4: A Many-Objective Evolutionary Algorithm for Protein Structure Prediction

1.1 Introduction

Proteins are organic macromolecules composed of different types of 20 amino acids and play an important role in most biological processes such as biochemical reactions, growth, structural support, and transporting nutrients. Different protein structures perform various biological functions; therefore, understanding a protein structure has been an important task [1]. At present, the structural information of proteins is mainly obtained by using three experimental methods: X-ray crystallography [2], nuclear magnetic resonance [3], and cryo-electron microscopy [4]. However, these methods are expensive and time-consuming. Two alternative computational methods, template-based modeling (TBM) and template-free modeling (FM), have been used to solve protein structure prediction (PSP) problems.

TBM is effective at constructing a protein model and can provide structural information of two-thirds of the protein families [5]. It models the protein structure based on already-determined protein structures stored in the Protein Data Bank (PDB). This approach exploits the fact that many sequences are similar to a template taken from the PDB. Generally, templates of proteins selected by single/multiple-sequence alignment methods, e.g., basic local alignment search tool (BLAST) [6], use template-based mutations, insertions, and deletions to refine a target template. In recent years, the approaches have been developed to allow rapid protein model construction [7, 8]. Although TBM can generate proteins with high accuracy, its limitation is that an unreliable protein may be

generated if no homologous structure is found in the PDB[9]. In contrast, FM models a protein structure without using the global homology information in the PDB, which allows algorithms to use fragment assemblies, secondary predictions, and statistical information [10]. Because FM seeks a protein structure from its primary sequence without relying on a homologous structure, it has the ability to predict novel protein structures. However, fragment assembly [9] uses extensive homology information to assist in protein structure prediction, thus limiting the in-depth investigation of novel protein structures. Two well-known protein structure prediction methods, namely Rosetta [11] and QUARK [12], use Monte Carlo simulations with fragment assemblies to model a protein structure. Deep-learning techniques have been widely used in PSP[13]. Most deep-learning techniques [14, 1] predict protein structures by learning all publicly available structures in PDB, which require high computational resources. For example, Alphafold2 [15] requires several weeks to train a deep neural network on a platform with 16 TPUv3s. It is widely accepted that newly discovered proteins cannot be predicted by deep learning since it generally requires large template information. To predict protein structures without homology information and reduce the computational burden, fragment assembly is not considered in our work.

According to Anfinsen's thermodynamic hypothesis [16] that structural information is encoded in an amino-acid sequence and that the energy of a protein is lowest in its native state, PSP can be regarded as a search optimization problem that aims to find the lowest-energy conformation. However, it involves searching in a very high-dimensional conformational space and has been classified as an NP-hard problem in computational theory [17]. FM models protein structures by minimizing the protein energy function in the conformational space. Therefore, large spaces require an efficient conformational search strategy. Monte Carlo simulations [18] and molecular dynamics [19] are the two main methods used for searching a protein conformational space [20].

A novel approach to predict protein structures is to search the conformational space by using evolutionary algorithms (EAs), which are efficient and have been applied to solve complex large-scale optimization problems[21, 22]. Researchers have already used them to solve the PSP problem without fragment-assembly techniques [23]. Cutello et al. proposed a two-objective evolutionary algorithm (including two functions, i.e., the bond and non-bond energies provided by Chemistry at HARvard Macromolecular Mechanics (CHARMM)) with a secondary-structure constraint that reduces the protein conformational space to solve the PSP problem [24]. Venske et al. proposed an

adaptive decomposition-based two-objective differential evolution algorithm for PSP [25]. Gao et al. utilized the solvent-accessible surface area (SASA) as one objective and combined it with bond and non-bond energy functions to define a three-objective PSP problem [26]. They proved the ability of SASA to improve the accuracy and efficiency in solving PSP problems. Song et al. used archive information to adjust the local protein structure and utilized a three-objective evolutionary strategy to search the protein conformational space [27]. Then, they developed it by using a three-objective particle-swarm optimization with two archives and three objectives: the dDFIRE, bond, and non-bond energies [28]. Zhang et al. proposed two selection strategies (i.e., secondary-structure-based and contact-based selection strategies) in a differential evolution algorithm to guide the search in conformation space [29]. They used diverse energy functions to measure the conformation in single-/multi-objective evolutionary algorithms.

The precision of an energy function is essential for solving PSP problems because it quantitatively reflects the physical interaction of proteins [30]. Energy functions are classified as physics-based and knowledge-based. The former category calculates the interatomic interaction based on quantum mechanics and elementary Coulomb, which requires calculating large complex equations. Therefore, alternative physics-based force fields, including AMBER [31], CHARMM [32], GROMOS96 [33], and OPLS [34], use experimental and quantum mechanical data to calculate bond lengths, torsion angles, and van der Waals and electrostatic interactions. A knowledge-based energy function uses the statistical regularities of known protein structures in the PDB library. This includes contact [35], orientation-dependent [36], and distance-dependent[37] potentials. Some well-known functions, such as RAPDF [38], KBP [37], DFIRE [39], DOPE [40], OPUS-PSP[41], ROSETTA [42], and RWplus [36] potentials, have been adapted successfully for PSP problems [43], protein folding [44], protein-protein docking [45], and protein structure refinement [46].

Generally, the key to solving a PSP problem using FM is fourfold. First, an effective representation of the protein is necessary to reduce the conformational space. Second, an efficient conformation search strategy is essential to ensure the efficacy of search for the native structure within a very large conformational space. Third, accurate energy functions measure conformation stability. Fourth, an effective decision-making method is required to select a better conformation in decoy archives. Thus, the PSP problem can be regarded as a single/multi-objective optimization problem that can be solved by multi-objective evolutionary algorithms (MOEA)[47]. Different types of

energy functions, such as distance-based, origination-based, bond, and non-bond, have been used to evaluate the stability of predicted protein structures [48]. These energy functions can limit the conformational space in different ways to improve the prediction accuracy. Thus, it is expected that PSP can involve different types of energy functions to enhance the accuracy of a predicted protein structure. However, because all the aforementioned energy functions may be inaccurate, using a single-objective optimization method usually yields a malformed protein structure [49]. Although multi-objective optimization methods (with 2 or 3 objectives) can provide better protein structures than single-objective methods do[50], they have a higher risk of generating unreasonable protein structures (e.g., overlap structures and flat structures with high error) than many-objective methods due to their weak conflicting relationship and selection pressure [26]. Therefore, the PSP problem should be formulated as a many-objective protein structure prediction problem (MaOPSP). By doing so, diverse aspects of the predicted structure can be measured, resulting in a more accurate prediction than obtained by the state of the art.

To the best of our knowledge, based on the aforementioned considerations, the PSP problem is treated as an MaOPSP problem for the first time in this study. The selection pressure provided by multi/many-objective optimization algorithms decreases as the number of objectives increases [51], which dramatically deteriorates their effectiveness. To address this issue, a many-objective evolutionary algorithm (MaOEA) is adopted to perform MaOPSP. In addition, a decision-making method based on clustering is used to select the final conformation.

The main contributions of this study are as follows:

1) For the first time, an MaOPSP is formulated to solve a PSP problem. It includes four objectives: bond and non-bond potentials provided by CHARMM, a SASA, and RWplus. Experimental results support our hypothesis that more optimization objectives are needed in the conformation search to improve the prediction accuracy of a protein structure.

2) A modified four-objective evolutionary algorithm (MO4) is used to search a protein conformational space. The results indicate that it can operate on a very large conformational space and provide a way to improve the prediction accuracy by using suitable operators specially designed for MaOPSP.

3) The objectives are analyzed, and the results demonstrate that a method of using many ob-

jectives is better than that with a single/multi-objective because the former can yield more accurate conformations than the latter owing to the constraint arising from the conflicting objectives.

Section 1.2 describes the PSP problem. Section 1.3 introduces the MO4 method. Section 1.4 presents a performance comparison between MO4 and other state-of-the-art methods. Finally, Section 1.5 concludes this study.

1.2 Protein Structure Prediction Problem

This section introduces research motivation and presents an efficient protein representation involving torsion angles to reduce the conformational space. Four energy functions are introduced to improve the prediction accuracy of protein structures.

1.2.1 Representation of Proteins

In MaOPSP, an effective representation of a protein is essential for reducing the conformational space. In this study, the search space is reduced by using torsion angles to represent protein structures, as shown in Fig. 1.1. There are three main chain torsion angles: around the $-N - C_\alpha-$ bond (Φ), around the $-C_\alpha - C-$ bond (Ψ), and around the $-C - N-$ bond (Ω) and side-chain torsion angles, χ_i ($i = 1, 2, 3, 4$). Side-chain angles also influence the structure and stability of proteins. Their number depends on the residue type (Table 1.1). Ω denotes the peptide bond that links two amino-acid residues and is set to 180° because the peptide bond has a partial double bond and tends to be planar. The conformational space remains very large even after using torsion angles to reduce it. To reduce it further, the torsion angles are constrained by using secondary-structure information. These constraints are shown in Table 1.2. Moreover, side-chain angles can be constrained based on the backbone-dependent rotamer library to reduce the space[52].

1.2.2 Protein Energy Function

In this study, the PSP problem is regarded as a many-objective optimization problem with four objectives, i.e., two physics-based functions of CHARMM force fields, a SASA, and a knowledge-

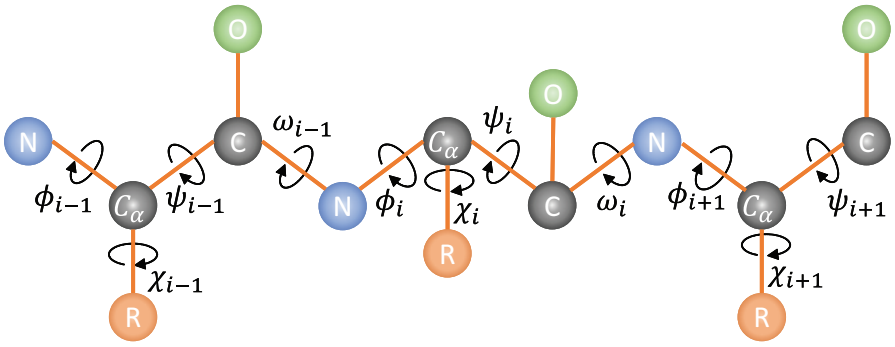


Figure 1.1: The illustration of torsion angles.

Table 1.1: Number of side-chain torsion angles in each residue.

Residues	Number of side-chain torsion angles
ALA, GLY, PRO	–
CYS, SER, THR, VAL	χ_1
ASP, ASN, HIS, LEU, PHE, ILE, TRP, TYR	χ_1, χ_2
MET, GLN, GLU	χ_1, χ_2, χ_3
ARG, LYS	χ_1, χ_2, χ_3, χ_4

based function RWplus.

1.2.2.1 CHARMM force field

The CHARMM22 force field is used to calculate bond and non-bond energy values in this work. It is a widely used force field in biological molecular simulation[53]. It calculates internal energy (bond) and external energy (non-bond), which is defined as follows:

$$E_C = E_B + E_{B'}. \tag{1.1}$$

The bond energy term includes bond stretches, bond angles, dihedral, improper angles, and

Table 1.2: Secondary structure constraint.

Secondary structure	Range of Φ	Range of Ψ
H (α-helix)	$[-67°, -47°]$	$[-57°, -37°]$
B (β-bridge)	$[-130°, -110°]$	$[110°, 130°]$
E (β-strand)	$[-130°, -110°]$	$[110°, 130°]$
G (3-10-helix)	$[-59°, -39°]$	$[-36°, 16°]$
I (pi-helix)	$[-67°, -47°]$	$[-80°, -60°]$
T (turn)	$[-180°, 180°]$	$[-180°, 180°]$
S (bend)	$[-180°, 180°]$	$[-180°, 180°]$
U (undefined)	$[-180°, 180°]$	$[-180°, 180°]$

Urey-Bradley terms, i.e.,

$$\begin{aligned} E_B \;=\; & \sum_{stretches} k_b(b-b_0)^2 \;+\; \sum_{angles} k_\theta(\theta-\theta_0)^2 \;+ \\ & \sum_{dihedrals} k_\phi[1+cos(n\phi-\delta)] \;+ \\ & \sum_{improper} k_\omega(\omega-\omega_0)^2 \;+\; \sum_{Urey-Bradley} k_u(u-u_0)^2, \end{aligned} \tag{1.2}$$

where k_b, b, and b_0 are the bond force constant, bond length, and equilibrium bond length, respectively. k_θ, θ, and θ_0 are the angle force constant, valence angle, and equilibrium angle, respectively. k_ϕ, n, ϕ, and δ are the dihedral force constant, the multiplicity of functions, torsion angle, and phase angle, respectively. k_ω, ω, and ω_0 are the improper force constant, improper angle, and equilibrium improper angle, respectively. k_u, u, and u_0 are the Urey–Bradley force constant, distance between atoms separated by two covalent bond, and equilibrium distance, respectively.

The non-bond energy term consists of Van-der-Waals and electrostatic terms, i.e.,

$$\begin{aligned} E_{B'} \;=\; & \sum_{Van-der-Waals} \varepsilon_{ij}\left[(\frac{R_{ij}}{r_{ij}})^{12} - (\frac{R_{ij}}{r_{ij}})^{6}\right] + \\ & \sum_{electrostatic} \frac{q_i q_j}{e \cdot r_{ij}}, \end{aligned} \tag{1.3}$$

where ε_{ij} is the Lennard Jones well-depth, R_{ij} is the minimum interaction distance, r_{ij} is the distance between two atoms i and j, q_i is the atomic charges, and e is the dielectric constant. Specifically, ε_{ij} and R_{ij} depend on atom types and are obtained via combination rules [32].

1.2.2.2 Solvent-accessible surface area

The SASA means calculates the surface area of atoms in solvent. It has been used to improve the accuracy of protein structure prediction [54]. It is used to restrict the surface of proteins in MaOPSP. Pyrosetta [55] is used to calculate a precise SASA with the atomic radius of 1.5Å for a given full-atomic structure in our work.

1.2.2.3 RWplus

The knowledge-based energy functions have been successfully applied in a PSP problem. More effective energy function's can undoubtedly increase prediction accuracy. This work uses a new atomic statistical potential function called RWplus that consists of a pair-wise distance-dependent and side-chain orientation-dependent energy terms [36], i.e.,

$$E_R = \sum_{\alpha,\beta} \bar{\mu}(\alpha,\beta,R) + 0.1 \sum_{A,B} \delta(A,B)\bar{\mu}(A,B,O_{AB}), \tag{1.4}$$

The pair-wise distance-dependent energy term is derived from Boltzmann's law [56]:

$$\sum_{\alpha,\beta} \bar{\mu}(\alpha,\beta,R) = -kTln\frac{p_D(\alpha,\beta,R)}{\bar{p}_D(\alpha,\beta,R)} \approx -kTln\frac{N_D(\alpha,\beta,R)}{\bar{N}_D(\alpha,\beta,R)}, \tag{1.5}$$

where k is the Boltzmann constant, T is the Kelvin temperature, and R is the distance between atom α and β. $p_D(\alpha,\beta,R)$ and $N_D(\alpha,\beta,R)$ are observed probability and number of atom pairs (α,β), respectively. $\bar{p}_D(\alpha,\beta,R)$ and $\bar{N}_D(\alpha,\beta,R)$ are the expected probability and number of atom pairs (α,β).

The orientation-dependent energy term [57] is described as:

$$\begin{aligned} \delta(A,B)\sum_{A,B} \bar{\mu}(A,B,O_{AB}) &= \delta(A,B)\cdot -kTln\frac{p_D(A,B,O_{AB})}{\bar{p}_D(A,B,O_{AB})} \\ &\approx \delta(A,B)\cdot -kTln\frac{N_D(A,B,O_{AB})}{\bar{N}_D(A,B,O_{AB})}, \end{aligned} \tag{1.6}$$

where $\delta(A,B)$ is set to one when vector pairs A and B are in contact, O_{AB} is the relative orientation between vector types A and B, $p_D(A,B,O_{AB})$ and $N_D(A,B,O_{AB})$ are the observed probability and

number of vector pair (A, B), and $\bar{p}_D(A, B, O_{AB})$ and $\bar{N}_D(A, B, O_{AB})$ are the expected probability and number of vector pair (A, B), respectively.

1.2.3 Many-objective Protein Structure Prediction Problem

In the last decade, a large number of approaches were researched for the PSP problem. They solve the PSP problem as a single/multi-objective optimization problem. We for the first time treat the PSP as MaOPSP to be solved by MaOEA. The proposed MaOPSP use four energy functions as objectives: two energy functions of the CHARMM force field, a SASA term, and a knowledge-based energy function RWplus. In addition, proteins are constructed by the torsion angles that reduce conformation space.

A multi-objective optimization problem includes more than one objective function is formulated as follows:

$$\begin{aligned} &\text{Minimiz } f(p) = \{f_1(p), f_2(p), ..., f_M(p)\} \\ &\text{subject to } p \in P, \end{aligned} \tag{1.7}$$

where p is the decision vector in the decision space P, $f_i(p),\ i = (1, 2, 3, ..., M)$ is the ith objective function where M represents the number of objective functions. When $M > 3$, it is known as an many-objective optimization problem (MaOOP). In MaOPSP, $P = \{p_1, p_2, ..., p_N\}$ represents a set of proteins. Each protein $p_k,\ (k = 1, 2, 3, ..., N)$ is described by dihedral angles, i.e.,

$$\begin{aligned} &p_k = \{\phi_1, \psi_1, X_1, \phi_2, \psi_2, X_2, ..., \phi_S, \psi_S, X_S\} \\ &X_S = \{\chi_{S_1}, \chi_{S_2}, ..., \chi_{S_r}\}, \\ &\text{subject to:} \\ &\quad \phi_1 \in \text{Range}(\phi) \text{ in Table 1.2} \\ &\quad \psi_1 \in \text{Range}(\psi) \text{ in Table 1.2} \\ &\quad -180^\circ < \chi_1 \leq 180^\circ, \end{aligned} \tag{1.8}$$

where S is the length of an amino acid, S_r is the number of side-chains of residues. Table 1.1 shows the number of side-chain of each type residues. ϕ and ψ can be constrained by secondary structures.

Table 1.2 presents the boundary constraint of dihedral angles. $f_i(p)$ is the different energy functions of proteins in Section 1.2.2. The Pareto dominance is usually used to compare the solution in MaOOP. Given two proteins p_1 and p_2, p_1 dominates p_2 (denote as $p_1 \prec p_2$), subject to:

$$\begin{aligned} &1)\ \forall i \in \{1, 2, ..., k\},\ f_i(x) \leq f_i(y);\ \ and \\ &2)\ \exists j \in \{1, 2, ..., k\},\ f_j(x) < f_j(y). \end{aligned} \tag{1.9}$$

If p_1 dominates p_2, it indicates that the p_1 is better than p_2 in at least one item of energy functions. A better solution corresponds to a better structure.

A solution $p^* \in P$ is Pareto-optimal if there dose not exist solution $p \in P$ that dominates it. The set of all Pareto-optimal solutions are named Pareto-optimal set (P_s), i.e.,

$$P_S = \{p^* \in P \mid \nexists\, y \in P,\ y \prec p^*\}. \tag{1.10}$$

P_S corresponding objective vectors from a Pareto front (P_F), i.e.,

$$P_F = \{f(x) \mid x \in P_S\}. \tag{1.11}$$

The goal of solving MaOOP is to obtain an approximation set that contains solutions as close as possible to the real P_F and as diverse as possible in the objective space. In MaOPSP, non-dominance conformations are added into an archive to construct P_S.

1.3 Four-objective Evolutionary Algorithm

This section introduces MO4 to solve an MaOPSP problem. Two selection strategies are introduced to select a better individual to pass into the next generation and update a conformation archive. A decision-making method is introduced for choosing a conformation from the archive.

Traditional multi-objective algorithms, such as the NSGA II [58] and the SPEA2 [59], cannot obtain a desired solution because the selection pressure decreases with the number of objectives increase in MaOOP [60]. To address this problem, various strategies have been proposed for improving the search performance. These include the decomposition [61], indicator [62], preference [63],

and diversity-based [64] methods. The native protein has many different conformations, and its energy landscape resembles a very complicated funnel [65]. MaOEA needs high population diversity to explore the conformational space. Therefore, we modify MaOEA based on two selection strategies to solve MaOPSP and name this MO4. Its main improvements over the original MaOEA are as follows: First, a variable mutation rate is introduced to achieve a better adaptation of MaOPSP. It enhances the effectiveness of the exploration of the conformational space in the early stage of a search process, while refining a protein structure in the later stage. Second, it saves the generated conformations in an archive, and a novel updating strategy is introduced to ensure the diversity of conformations. Third, a decision-making method [66] is used to select the best conformation from the archive via clustering. Fig. 1.2 shows the MO4 process. First, a population is constructed according to an amino-acid sequence. Second, individuals selected from the population are added into a mating pool, and individuals in the pool generate new offspring through mutation and crossover. Third, a new population is selected from the combined population of parents and offspring by environmental selection. The population updates an archive by using the non-dominance rank. Finally, the best protein structure is determined from the archive by employing a desertion-making method. The pseudo code of MO4 is presented in Algorithm 1. Those of mating selection, mutation, environmental selection, and update archive are shown in section 1.3.3.

Algorithm 1: The pseudo-code of MO4.

Input: Amino acid sequence AA
Generate the secondary structure by PSPRED.
Initialization Population P by the secondary structure.
while $Iter<MaxIter$ **do**
  P^m = MatingSelection(P)
  $P^{'}$ = Mutation(P^m)
  P = $BoundaryConstraint(P)$
  $Q = P \cup P^{'}$
  P = EnvironmentalSelection(Q)
  A = UpdateArchive(P, A)
end
The best conformation is selected from the archive by the decision-making method.
Return: The best protein.

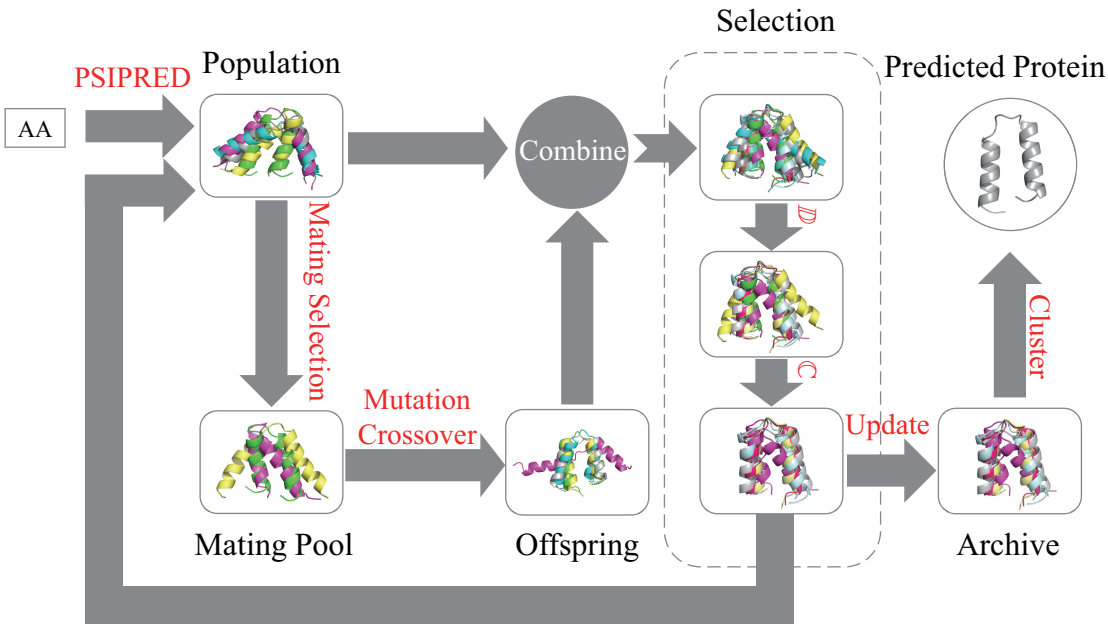


Figure 1.2: The illustration of the MO4.

1.3.1 Initialization Population

In the initialization process, a highly accurate method, denoted PSIPRED [67], is used to predict the secondary protein structure. It incorporates two feed-forward neural networks to analyze the output of PSI-BLAST [6] (http://bioinf.cs.ucl.ac.uk/psipred/). The secondary structure results of proteins restrict the population boundary. The dihedral angles of proteins are randomly constructed using Eq. (1.8). Angles ϕ, ψ, and X for each residue are generated by a uniform distribution in the boundary constraint. The initialization population, P, consists of N randomly generated conformations. In Fig. 1.2, part of the *Population* exhibits the structure of the initialization conformations.

1.3.2 Mating Selection

Mating selection plays an essential role in determining the convergence performance of the algorithm. It aims to select a mating population, P^m, from the parent population, P, to produce offspring to address the selection pressure. P^m includes better conformations in the parent population. Pareto dominance lacks the capability of measuring the convergence performance in MaOOP. Therefore, the favorable convergence function based on the Chebyshev function and favorable weight is utilized

to measure the convergence performance[68], i.e.,

$$\mathbb{C}(p) = \max_{1 \le i \le M} \left\{ \omega_{p,i} \left| f_i(p) - z_i^* \right| \right\}, \tag{1.12}$$

where M is the number of objectives. $z_i^* = (z_1^*, z_2^*, ..., z_M^*)$ is the ideal point of the problem. It is defined as the smallest objective values of the current iteration since the real optimal point is unknown in the MaOPSP problem. $\omega_{p,i}$ is the favorable weight of conformation p, defined as

$$\omega_{p,i} = \begin{cases} 0, & \text{if } f_i(p) = z_i^* \\ 1, & \text{if } f_i(p) \neq z_i^* \text{ and } \exists m \ f_m(p) = z_i^* \\ \frac{1}{f_i(p) - z_i^*} \left[\sum_{m=1}^{M} \frac{1}{f_m(p) - z_m^*} \right]^{-1}, & \text{otherwise.} \end{cases} \tag{1.13}$$

$\mathbb{C}$ indicates the convergence performance. A small value of $\mathbb{C}$ indicates that the conformation converges better because no conformation is closer to the ideal point than the best conformation of the current population. Using a binary tournament selection selects solutions to produce an offspring in mating selection. The dominant one of two randomly selected solutions is inserted into the mating pool. If both conformations are non-dominant, the one with the smaller $\mathbb{C}$ value is put into the mating pool. The better conformations of the population are selected to generate an offspring by mating selection. Otherwise, a randomly selected conformation enters the pool. After the mating selection, a polynomial mutation operator [69] and a simulated binary crossover operator [70] are used to evolve the selected mating population with a boundary constraint.

Specifically, algorithm needs a larger mutation rate to search the conformational space in the early stage of the process. By contrast, the conformations need a small mutation rate to refine the structure in the later stage. Therefore, the mutation rate is a variable defined as:

$$M_p = e^{\frac{-t}{4T}}, \tag{1.14}$$

where t is the tth iteration, and T is the maximum iteration count.

1.3.3 Mutation and Crossover

In MO4, a polynomial mutation operator and a simulated binary crossover operator are used to evolve the selected mating population. The former is defined as follows:

$$\begin{aligned} P^{'} &= P^{m} + \delta(\hat{P} - \check{P}) \\ \hat{P} &= \{\hat{\phi}_1, \hat{\psi}_1, \hat{X}_1, ..., \hat{\phi}_S, \hat{\psi}_S, \hat{X}_S\} \\ \check{P} &= \{\check{\phi}_1, \check{\psi}_1, \check{X}_1, ..., \check{\phi}_S, \check{\psi}_S, \check{X}_S\}, \end{aligned} \tag{1.15}$$

where P^m and $P^{'}$ denote the mating population and its offspring, respectively. $\hat{P}$ and $\check{P}$ are the upper and lower boundaries of P^m, respectively. $\hat{\phi}$, $\hat{\psi}$ and $\hat{X}$ represent the upper boundaries of torsion angles, respectively. $\check{\phi}$, $\check{\psi}$ and $\check{X}$ represent the lower boundaries of torsion angles, respectively. δ is described as:

$$\delta = \begin{cases} [2r + (1-2r)(1-\delta_1)^{\eta_m + 1}]^{\frac{1}{\eta_m+1}}, & \text{if } r \leq 0.5 \\ 1 - [2(1-r) + 2(r-0.5)(1-\delta_2)^{\eta_m + 1}]^{\frac{1}{\eta_m+1}}, & \text{otherwise,} \end{cases} \tag{1.16}$$

where r is a random number in interval $(0, 1)$, η_m is a positive real number, and δ_1 and δ_2 are defined as follows:

$$\begin{aligned} \delta_1 &= (P^m - \check{P})/(\hat{P} - \check{P}) \\ \delta_2 &= (\hat{P} - P^m)/(\hat{P} - \check{P}). \end{aligned} \tag{1.17}$$

Specifically, conformations need a larger mutation rate to search the conformational space in the early stage of the process. By contrast, the conformation need a small mutation rate to refine the structure in the later stage. Therefore, the mutation rate is a variable defined as:

$$M_p = e^{\frac{-t}{4T}}, \tag{1.18}$$

where t is the tth iteration, and T is the maximum iteration count.

After mutating, the simulated binary crossover of two conformations, p_1^m and p_2^m, in the mating

population, P^m, is defined as:

$$\begin{aligned} p_1^{'} &= 0.5[(1+\beta)p_1^m + (1-\beta)p_2^m] \\ p_2^{'} &= 0.5[(1-\beta)p_1^m + (1+\beta)p_2^m], \end{aligned} \tag{1.19}$$

where $p_1^{'}$ and $p_2^{'}$ denote the two conformations of the offspring, produced via the simulated binary crossover. Then, β is calculated as:

$$\beta = \begin{cases} (2r)^{\frac{1}{1+\eta_c}}, & r < 0.5 \\ (\frac{1}{2-2r})^{\frac{1}{1+\eta_c}}, & \text{otherwise}, \end{cases} \tag{1.20}$$

where η_c is the distribution index of the crossover.

Offspring $P^{'}$ is generated by the mutation and crossover operators. Conformation update their structures by modifying dihedral angles using the two operators. In addition, the algorithm uses environmental selection based on direction diversity and favorable convergence to improve its convergence performance for alleviating the selection pressure.

A boundary constraint is used to restrain the population, i.e.,

$$P = \begin{cases} \hat{P}, & P > \hat{P} \\ \check{P}, & P < \check{P} \end{cases} \tag{1.21}$$

where $\hat{P}$ and $\check{P}$ are the upper and lower boundaries of P^m, respectively.

In addiction, the pseudo-code of mating selection, environmental selection, projection, directional diversity selection and favorable convergence selection are shown in Algorithms 2, 3, 4, 5 and 6, respectively.

1.3.4 Environmental Selection

The environmental selection aims to select N conformations from parent population P and offspring $P^{'}$ to generate the next population, P_{next}, for the next iteration. However, almost all solutions are distributed on an identical Pareto front for MaOOP, which means that the weak selection pressure cannot support algorithms to find better solutions [71]. The environmental selection based on

Algorithm 2: Mating Selection.

Input: Population P, mating pool size N
$P^m = \emptyset$
while $|P^{'} < N|$ **do**
  randomly select two solutions x, y from P;
  if $x \prec y$ **then**
    $P^m = P^m \cup x$
  end
  else if $y \prec x$ **then**
    $P^m = P^m \cup y$
  end
  else if $\mathbb{C}(x) < \mathbb{C}(y)$ **then**
    $P^m = P^m \cup x$
  end
  else if $\mathbb{C}(y) < \mathbb{C}(x)$ **then**
    $P^m = P^m \cup y$
  end
  else if $rand(0, 1) < 0.5$ **then**
    $P^m = P^m \cup x$
  end
  else
    $P^m = P^m \cup y$
  end
end
Return: mating population P^m

Pareto dominance ranking only fails to ensure convergence performance. Hence, MO4 uses favorable convergence and directional diversity [71] to improve the selection pressure for finding better conformations.

In this study, a directional diversity function is used to measure the degree of diversity of the population and combined with the $\mathbb{C}$ function to implement environmental selection. First, the parent population and offspring are classified into different fronts $(F_1, F_2, ...)$ by non-dominated sorting. Iteratively, these fronts insert the next population, P_{next}, if its size is less than N. When the size of the union of the ith front and P_{next} equals N, the union is the final next population (i.e., $P_{next} = F_i \cup P_{next}$). If the size of $F_i \cup P_{next}$ is greater than N, a direction-based method and a convergence-based method select conformations from F_i and add them to P_{next} until the size of P_{next} equals N. P_{next} is projected onto a hyperplane to estimate the directional diversity of the solutions. Subsequently, L candidate conformations with a better directional diversity are selected from F_i

Algorithm 3: The pseudo-code of environmental selection.

Input: Population size N, the union of the parent population and the offspring Q
$(F_1, F_2, ...) = nondominatedsorting(Q)$
$P = \emptyset, i = 1$
while $|P \cup F_i| < N$ **do**
  $P = P \cup F_i,\ i = i + 1$
end
if $|P \cup F_i| = N$ **then**
  $P = P \cup F_i$
end
else
  $(\tilde{P}, \tilde{F}_i) = projection(P, F_i)$
  while $|P| < N$ **do**
    $(R, \tilde{R}) = direction\_based\_selection(F_i, \tilde{P}, \tilde{F}_i)$
    $(r, \tilde{r}) = convergence\_based\_selection(R, \tilde{R})$
    $P = P \cup r, \tilde{R} = \tilde{R} \cup \tilde{r}$
    $F_i = F_i - r, \tilde{F}_i = \tilde{F}_i - \tilde{r}$
  end
end
Return: P

Algorithm 4: The pseudo-code of the projection.

Input: Population P, the Pareto Front F_i.
Calculate the ideal point $z_{min} = \{f_1^{min}(x), ..., f_M^{min}(x)\}$.
Calculate the translated objective $f^{'}(x) = f(x) - z_{min}$.
Calculate the extreme points $z_i^{extreme}, i = 1, 2, ..., M$ by ASF.
Calculate the intercept $a_i, i = 1, 2, ..., M$.
Generate the normalize objective $f(x)$.
Project the normalization objective onto hyperplane.
Return: $\widetilde{P}, \widetilde{F}_i$

Algorithm 5: The pseudo-code of the direction based selection.

Input: Current Front F_i, projection points $\widetilde{P}$, $\widetilde{F_i}$, the neighborhood size k, candidate solution size L

if $\widetilde{P} = \emptyset$ **then**

  $k = min\{k, (F_i|\}$

  Calculate the distance of the projection point $\widetilde{x}$ to its k nearest points in the $\widetilde{F_i}$

end

else

  $k = min\{k, |\widetilde{P}|\}$

  Calculate the distance of the projection point $\widetilde{x}$ to its k nearest points in the $\widetilde{P}$

end

Calculate the diversity density of solutions.

$L = min\{L, |\widetilde{F_i}|\}$

Select L candidate solutions from F_i with the k smallest nearest distances.

Return: Candidate conformations R, their projection $\widetilde{R}$

Algorithm 6: The pseudo-code of the convergence based selection.

Input: Candidate conformations R and their projection points $\widetilde{R}$

Calculate the favorable convergence performance for R.

Select a **conformation** r From R with a probability proportional to its convergence performance.

Return: r

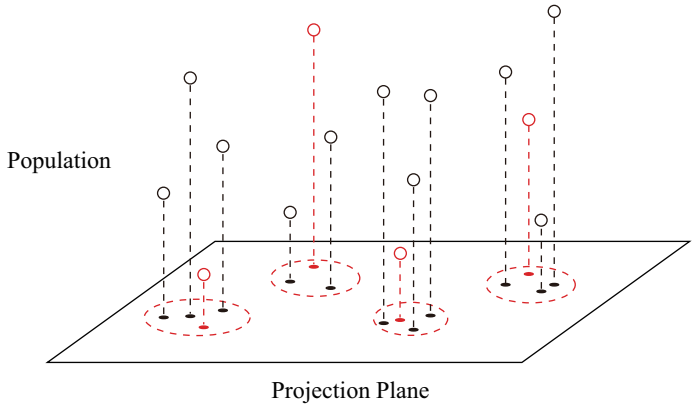


Figure 1.3: The illustration of directional diversity.

by direction-based selection. The illustration of directional diversity is exhibited in Fig. 1.3. The individual is projected onto a plane. The dense region located in the projection plane indicates that most individuals converge to this region, which suggests that such convergence direction has a high probability of finding better solutions. The directional diversity selection strategy selects L individuals with the smallest K neighborhood distance distributed in these dense regions and puts them into the next population to improve population diversity. Therefore, this strategy not only ensures population diversity but also directs a promising convergence direction. A conformation is selected by convergence-based selection from these L candidate conformations and is added to P_{next} until the size of P_{next} equals N. The pseudo-code of the environmental selection is provided in Algorithm 3.

1.3.4.1 Projection

In the projection, the objective functions are normalized to obtain a better estimate of the directional diversity of conformations because different objectives have different ranges [72]. First, the ideal point is determined by the minimum value of each objective of the union of P_{next} and F_i, i.e.,:

$$z_{min} = \{f_1^{min}(p), ..., f_M^{min}(p)\},\ p \in P_{next} \cup F_i, \tag{1.22}$$

To make the ideal point become the origin, objective value $f(p)$ is transformed into $f^{'}(p) = f(p) - z_{min}$. The extreme point is calculated by using an achievement scalarizing function with

weight vector w:

$$\hat{z}_i = \underset{p \in P_{next} \cup F_j}{argmin} \ \mathrm{ASF}(p, w) = \underset{p \in P_{next} \cup F_i}{argmin} \{ \underset{1 \le m \le M}{max} (\frac{f_m(p)}{w_m^i}) \}, \tag{1.23}$$

where $w^i = (w_1^i, ..., w_m^i)$, $w_m^i = 10^{-6}, m \neq i$ and $w_i^i = 1$. These extreme points form a hyperplane, and intercept a_i of an objective can be calculated using analytic geometry. Finally, the normalized objective is described as:

$$f_i(p) = \frac{f(p)^{'}}{a_i - \hat{z}_i}. \tag{1.24}$$

The normalized objective is projected onto a hyperplane to produce projection points, which represent the direction of a conformation in the objective space. The pseudo-code of the projection is presented in Algorithm 4.

1.3.4.2 Direction-based selection

Direction-based selection aims to select L candidate conformations with better directional diversity from Pareto front F_i. The directional diversity estimates the diversity density of conformations by using the projection points and the K-nearest neighborhood distance. $\mathbb{D}$ of conformation p_{next} in F_i is calculated as the distance from $\widetilde{p}_{next}$ to the K nearest projection points in $\widetilde{P}_{next}$, i.e.,:

$$\mathbb{D}(p) = \sum_{k=1}^{K} \frac{1}{d_k^{\widetilde{p}_{next}}}. \tag{1.25}$$

In the directional diversity selection, two parameters, K and L, impact the calculation of the directional diversity. K indicates the number of nearest conformations used to estimate the diversity density. L expresses how much convergence information is used to perform the environmental selection. The sizes of $\widetilde{P}_{next}$ and F_i determine these two parameter values. When $\widetilde{P}_{next} = \emptyset$, the distance from $\widetilde{p}_{next}$ to its K-nearest projection point in F_i is calculated to select conformations. If the size of F_i is less than K, K is set to the size of F_i. When $\widetilde{P}_{next} \neq \emptyset$ and the size of $\widetilde{P}_{next}$ is less than K, K is set to the size of $\widetilde{P}_{next}$. In addition, when the size of F_i is less than L, L is set to the size

of F_i. The pseudo-code for the directional diversity selection, through which conformations with small distances are selected, is given in Algorithm 5. In Fig. 1.2, the *Selection* section shows the modification of conformations by $\mathbb{D}$ selection.

1.3.4.3 Convergence-based selection

Convergence-based selection selects the conformation with the most promising convergence from these L candidate conformations with a good $\mathbb{D}$. In this procedure, function $\mathbb{C}$ measures the convergence performance. A selection probability that is proportional to the value of $\mathbb{C}$ ensures that any conformation has a chance to enter the next generation. The pseudo-code for the convergence-based selection is given in Algorithm 6.

1.3.5 Archive-updating Strategy and Decision-making Method

In MO4, we propose an archive-updating strategy to update P_S of conformations selected by the environmental selection are added to the archive. A non-dominated sorting method [73] is used to ensure that the conformations in the archive are non-dominant. The archive-updating strategy is described as follows:

a) Combine the current population and the archive, and select unique conformations from the combination.

b) Assign conformations to different Pareto fronts $F = \{F_1, F_2, ...\}$ via non-dominated sorting.

c) Add conformations of the first Pareto front into an archive to save P_S of conformations.

d) If the size of the updated archive is larger than its maximum size, delete conformations with a high crowding distance from the archive.

The deletion strategy based on crowding distance [74] produces a greater dispersion of conformations in the Pareto front, to ensure population diversity.

In addition, a decision-making method, MUFOLD-CL [66], is performed to select the final conformation from the conformation archive by clustering. It uses two scores, i.e., Dscore1 and

Table 1.3: The parameter settings of MO4.

Parameters	Value
Population Size N_P	50
Crossover Rate p_c	1
Crossover Distribution Index η_c	20
Mutation Distribution Index η_m	20
Number of Nearest Neighbors K	5
Number of Candidate Solutions L	3
Maximum Number of Iterations $MaxIter$	3000

Dscore2, to measure the difference and similarity of the two models, respectively. Dscore1 is correlated with the root-mean-square deviation (RMSD), and Dscore2 is highly correlated with template-modeling scoring (TM) .

1.4 Experimental Results and Discussion

In this section, four criteria are presented to assess the performance of MO4. MO4 is compared with other state-of-the-art methods to demonstrate its prediction performance on thirty-four proteins. The structures of the predicted and native proteins are illustrated to highlight their difference. The conflict among the objectives is analyzed to verify the rationality of using many-objective methods. The Pareto front and energy value of the proteins are given to demonstrate how many-objective methods can generate more reasonable and fruitful proteins than single/multi-objective methods.

1.4.1 Experimental Setup

MO4 uses some external programs to calculate the energy of conformations. The CHARMM22 force field is used to calculate bond and non-bond energy, the SASA energy is calculated by Pyrosetta, and an executable program RWplus is utilized to calculate the knowledge-based energy. The parameter settings of MO4 are listed in Table 1.3. Its performance is tested on thirty-four representative proteins in PDB. These proteins are representative ones that are selected from three different types of proteins, i.e., α, β, and α/β ones. The details of proteins are listed in Table 1.4. All experiments are implemented using Python parallel processing on a workstation with Intel(R) Xeon(R) Silver 4110 CPU@2.1GHz and 80GB RAM.

Table 1.4: The details of thirty-four proteins.

PDB ID	Type	Length	PDB ID	Type	Length
1AB1	α/β	46	1UTG	α	70
1AIL	α	70	1ZDD	α	34
1BDD	α	60	2F4K	α	35
1CRN	α/β	46	2GB1	α/β	56
1DFN	β	30	2JUC	α	55
1E0G	α/β	48	2JZQ	α/β	57
1E0M	β	37	2KDL	α	56
1ENH	α	54	2M7T	$\alpha\beta$	33
1F7M	β	46	2MR9	α	44
1G26	β	31	2P5K	α/β	64
1I6C	β	39	2P6J	α	52
1IGD	α/β	61	2P81	α	44
1K36	β	46	2PMR	α	76
1MSI	α/β	66	3DF8	α/β	109
1Q2K	α/β	31	3NRW	α	104
1ROP	α	56	3P7K	α	45
1SXD	α	91	3V1A	α	48

1.4.2 Performance Criteria

In this study, RMSD [75], global distance test-total score (GDT) [76], TM [77], and distance matrix error (DME)[24] are adopted to evaluate the similarity between the predicted and native protein structures.

RMSD calculates the difference between the predicted and native protein structures:

$$\mathrm{RMSD}_{(a,b)} = \sqrt{\frac{\sum_{i=1}^{N} d_i^2}{N}}, \tag{1.26}$$

where a and b represent two protein structures by the Kabsch rotation matrix [75], d_i is the atomic distance between them, and N is the number of matched atoms. A smaller RMSD value represents a better structure.

GDT measures the similarity between two protein structures:

$$\mathrm{GDT} = \frac{100(C_1 + C_2 + C_3 + C_4)}{M}, \tag{1.27}$$

where C_1, C_2, C_3, and C_4 represent the numbers of aligned residues at distances of $\theta/4$, $\theta/2$, θ, and

2θ (θ = 4Å), respectively. M is the total number of residues. GDT lies within interval $(0, 100)$. A larger value implies a greater similarity structure.

TM measures the topological similarity between the predicted and native structures, i.e.,

$$\text{TM} = \frac{1}{L}\sum_{i=1}^{L_{align}} \frac{1}{1 + (\frac{d_i}{d_0})^2}, \tag{1.28}$$

where L is the length of the predicted protein, L_{align} is the number of aligned residues, d_i is the distance between the ith pair residues in the two structures, and d_0 is a scale used to normalize the match difference. TM ranges in interval $(0, 1]$. A higher value indicates a more robust topological similarity.

DME measures the similarity of inter-atomic distances, i.e.,

$$\text{DME} = \frac{\sqrt{\sum_{i=1}^{N}\sum_{j=1}^{N}(d_{ij}^a - d_{ij}^b)^2}}{N}, \tag{1.29}$$

where d_{ij}^a and d_{ij}^b represent the distance between atom i and j in protein structures a and b, respectively.

1.4.3 Prediction Results for Proteins

MO4 uses the secondary structure predicted by PSIPRED to reduce the conformational space. The experimental results records bRMSD, bGDT, bTM, bDME, RMSD, GDT, TM and DME values of MO4. The first four values represent the RMSD, GDT, TM and DME for the best conformation in the archive, respectively. The last four values represent the quality of the final solution selected by a decision-making method based on clustering. The detailed experimental results are described in Table 1.16. Those results imply that MO4 yields a protein structure with RMSD of less than 3Å for structures with PID 1ROP, 1ZDD, 2MR9, 3P7K, and 3V1A and RMSD of less than 6Å for 25 proteins. The average value of bRMSD equals 5.18Å, which reflects the strong ability of MO4 to predict protein structures. The average value of bTM is greater than 0.4, which implies that the predicted and native **structures** have similar topological structures. The average value of bGDT is 54.73, which means that predicted and native **structures** are similar in terms of the aligned residues.

Table 1.5: The energy results of the MO4 on thirty-four proteins.

PDB ID		1AB1	1AII	1BDD	1CRN	1DFN	1E0G	1E0M	1ENH	1F7M	1G26
Bond	Native	2.01E+03	7.11E+02	8.28E+02	5.70E+02	2.93E+02	6.99E+02	4.26E+02	5.27E+02	5.73E+02	5.42E+02
	Predicted	5.09E+02	7.70E+02	6.78E+02	5.29E+02	4.11E+02	5.44E+02	5.14E+02	6.17E+02	7.64E+02	3.50E+02
Non-bond	Native	-7.03E+02	-1.99E+03	-2.96E+02	-7.13E+02	-7.88E+02	-4.36E+02	2.85E+00	-1.82E+03	-3.62E+02	-6.93E+01
	Predicted	9.37E+12	2.87E+19	5.41E+03	1.17E+09	1.86E+08	6.72E+11	4.50E+03	2.80E+10	3.83E+13	1.14E+17
SASA	Native	3.21E+03	4.95E+03	4.42E+03	2.96E+03	2.56E+03	3.91E+03	3.55E+03	4.13E+03	3.69E+03	2.91E+03
	Predicted	3.57E+03	6.91E+03	5.87E+03	3.83E+03	3.76E+03	4.63E+03	5.50E+03	4.47E+03	2.99E+03	2.47E+03
RWplus	Native	-7.01E+03	-1.17E+04	-7.24E+03	-6.59E+03	-3.75E+03	-6.05E+03	-3.65E+03	-9.27E+03	-4.57E+03	-3.37E+03
	Predicted	-2.94E+03	-8.10E+03	-6.14E+03	-3.23E+03	-2.08E+03	-2.87E+03	-3.28E+03	-4.34E+03	9.10E+02	-1.08E+02
PDB ID		1I6C	1IGD	1K36	1MSI	1Q2K	1ROP	1SXD	1UTG	1ZDD	2F4K
Bond	Native	8.21E+02	7.02E+02	1.16E+03	7.06E+02	3.09E+02	4.68E+02	1.07E+03	9.48E+02	2.71E+02	5.35E+03
	Predicted	5.34E+02	6.74E+02	6.27E+02	7.55E+02	2.98E+02	6.48E+02	1.00E+03	8.27E+02	4.04E+02	3.55E+02
Non-bond	Native	-9.15E+02	-3.48E+02	-4.28E+01	-1.13E+03	3.08E+02	-9.98E+02	-8.66E+02	-1.69E+03	-1.43E+03	4.95E+02
	Predicted	1.01E+11	7.00E+13	2.74E+09	1.22E+17	1.88E+12	-2.46E+02	3.46E+10	3.18E+09	5.91E+04	6.31E+06
SASA	Native	3.64E+03	4.18E+03	4.07E+03	2.77E+03	4.64E+03	3.60E+03	5.19E+03	5.51E+03	3.02E+03	3.01E+03
	Predicted	3.91E+03	5.48E+03	5.33E+03	4.81E+03	3.28E+03	7.69E+03	8.81E+03	4.88E+03	3.39E+03	3.37E+03
RWplus	Native	-3.47E+03	-8.83E+03	-4.94E+03	-3.47E+03	-8.71E+03	-1.09E+04	-1.08E+04	-1.83E+04	-4.60E+03	-4.58E+03
	Predicted	2.09E+03	-3.60E+03	-4.05E+03	-2.42E+03	-1.58E+03	-6.52E+03	-9.83E+03	-3.92E+03	-4.06E+03	-2.88E+03
PDB ID		2GB1	2JUC	2JZQ	2KDL	2M7T	2MR9	2P5K	2P6J	2P81	2PMR
Bond	Native	8.68E+02	7.82E+02	7.48E+02	6.55E+02	3.53E+02	5.08E+02	5.47E+02	4.90E+02	5.40E+02	6.65E+02
	Predicted	6.07E+02	5.90E+02	5.92E+02	5.75E+02	3.24E+02	4.58E+02	7.00E+02	5.57E+02	5.06E+02	7.47E+02
Non-bond	Native	-1.55E+02	2.63E+02	1.18E+01	2.37E+02	-2.81E+02	1.56E+02	-4.94E+02	-1.70E+03	-9.86E+02	3.25E+04
	Predicted	6.19E+12	1.47E+05	1.64E+06	4.70E+12	3.15E+10	8.78E+08	6.49E+02	8.41E+09	7.48E+04	3.30E+11
SASA	Native	3.86E+03	3.81E+03	3.76E+03	5.09E+03	2.64E+03	3.29E+03	4.36E+03	4.69E+03	5.20E+03	5.18E+03
	Predicted	6.68E+03	6.39E+03	5.95E+03	5.25E+03	2.97E+03	3.71E+03	8.34E+03	6.13E+03	4.82E+03	5.52E+03
RWplus	Native	-7.59E+03	-8.14E+03	-8.54E+03	-7.15E+03	-3.25E+03	-5.88E+03	-1.04E+04	-8.17E+03	-4.81E+03	-1.44E+04
	Predicted	-1.58E+03	-5.49E+03	-4.68E+03	-2.84E+03	-1.16E+03	-3.40E+03	-5.28E+03	-5.58E+03	-4.52E+03	-7.60E+03
PDB ID		3DF8	3NRW	3P7K	3V1A						
Bond	Native	7.99E+03	3.64E+03	3.67E+02	3.09E+03						
	Predicted	1.18E+03	1.47E+03	4.63E+02	4.96E+02						
Non-bond	Native	-2.42E+03	-3.10E+03	-5.82E+02	-7.52E+02						
	Predicted	1.14E+11	5.28E+12	2.93E+07	2.49E+07						
SASA	Native	8.07E+03	6.51E+03	4.98E+03	4.05E+03						
	Predicted	9.52E+03	7.36E+03	5.03E+03	4.28E+03						
RWplus	Native	-1.85E+04	-1.94E+04	-6.14E+03	-8.92E+03						
	Predicted	-7.35E+03	1.05E+03	-4.12E+03	-5.95E+03						

Fig. 1.4 shows the convergence plot of MO4. The horizontal axis represents the number of iterations. The vertical axis represents the normalized objective value of the four objectives, to highlight their variations more clearly. In this figure, the bond, non-bond, and RWplus energies decrease with increasing iterations. In contrast, the SASA energy remains relatively constant. The protein SASA always decreases, resulting in the collapse of the proteins. Therefore, the SASA objective maintains the rationality of predicted structure. This means that MO4 is capable of generating reasonable proteins structures.

Four proteins are selected to evaluate the search performance of MO4. Each monomer of dimer protein COLE1 ROP (PDB ID: 1ROP) includes 56 residues and forms two-α structures. The disulfide-stabilized mini protein (PDB ID: 1ZDD) consists of 34 residues and forms two helices.

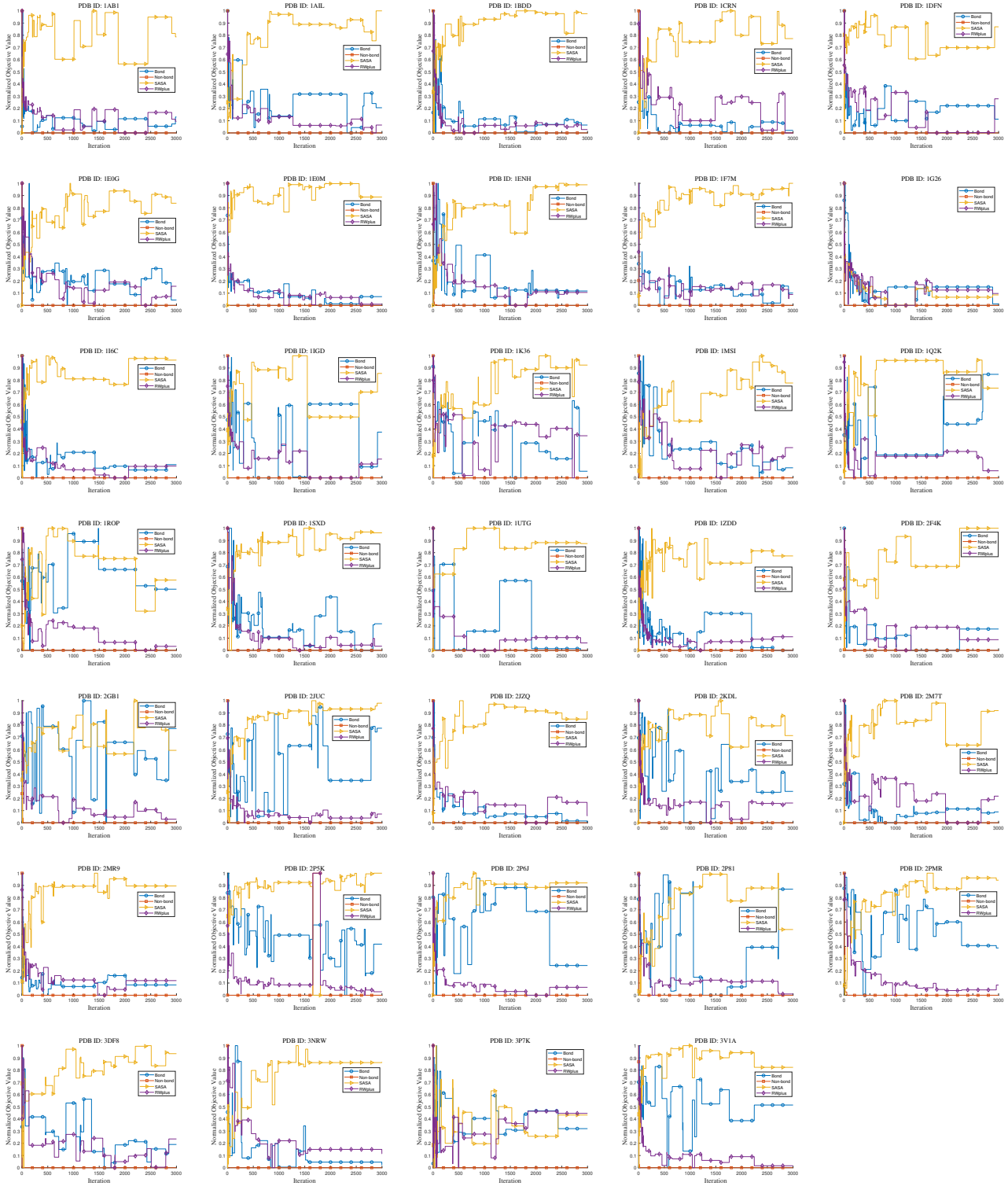


Figure 1.4: The convergence plot of four objectives of all proteins.

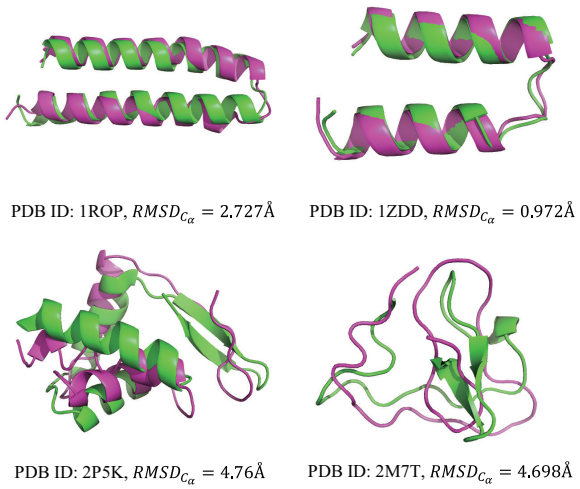


Figure 1.5: Comparison between prediction and native structures.

Table 1.6: The Wilcoxon and Friedman test results of parameter settings.

Parameters	bRMSD (Rank)	bGDT (Rank)	bTM (Rank)	bDME (Rank)	RMSD (Rank)	GDT (Rank)	TM (Rank)	DME (Rank)	Avg. Rank
K=1, L=1	8.11E-07 (8)	2.09E-07 (8)	4.06E-05 (7)	1.46E-06 (9)	1.15E-04 (7)	3.32E-04 (6)	1.82E-05 (7)	2.82E-06 (8)	7.50
K=1, L=2	7.12E-05 (3)	1.31E-04 (2)	2.59E-03 (5)	9.97E-06 (6)	8.02E-04 (5)	2.59E-03 (7)	2.21E-03 (2)	1.82E-04 (4)	4.25
K=1, L=3	4.43E-07 (6)	1.16E-05 (3)	4.80E-03 (2)	1.35E-05 (5)	5.70E-04 (3)	6.43E-03 (2)	3.06E-06 (6)	1.25E-05 (6)	4.12
K=2, L=1	8.55E-06 (7)	4.57E-06 (6)	1.40E-04 (8)	9.23E-06 (7)	7.77E-05 (8)	1.50E-04 (9)	1.46E-05 (8)	4.95E-06 (7)	7.50
K=2, L=2	1.25E-05 (4)	2.40E-06 (7)	7.15E-04 (6)	2.45E-05 (3)	4.36E-05 (6)	2.07E-04 (5)	3.45E-04 (4)	1.60E-04 (5)	5.00
K=2, L=3	3.32E-06 (5)	3.02E-05 (5)	1.46E-02 (4)	2.45E-05 (4)	6.74E-04 (4)	1.78E-03 (3)	4.68E-05 (5)	2.21E-04 (3)	4.12
K=5, L=1	3.72E-07 (9)	5.29E-07 (9)	2.04E-06 (9)	9.61E-07 (8)	4.39E-06 (9)	4.70E-04 (8)	8.83E-07 (9)	7.45E-07 (9)	8.75
K=5, L=2	1.23E-04 (2)	3.52E-05 (4)	1.23E-02 (3)	5.39E-05 (2)	6.03E-03 (2)	1.24E-03 (4)	4.70E-04 (3)	5.56E-03 (2)	2.75
K=5, L=3	– (1)	– (1)	– (1)	– (1)	– (1)	– (1)	– (1)	– (1)	1

The crystal structure of the N-terminal domain of AhrC (PDB ID: 2P5K) consists of 63 residues and forms three α-helices and two β-strands. The engineered cystine knot protein 2.5D (PDB ID: 2M7T) consists of 37 residues, including one helix and one strand. Fig. 1.5 depicts the predicted (purple) and native (green) protein structures (PDB IDs: 1ROP, 1ZDD, 2P5K, and 2M7T). This figure shows that MO4 generates a highly accurate protein structure, in terms of RMSD being less than 5Å. Specifically, MO4 is clearly accurate with regard to the α-helix structure. Fig 1.6 details the protein structure of all thirty-four proteins, together with its running time.

Table 1.5 lists the energies of the native and predicted structures. Clearly, MO4 can converge to a smaller value than the native structure for proteins 1BDD, 1E0G, 1I6C, 1IGD, and 1K36. This means that a structure with a small energy is not better than a large structure with a large energy, owing to the inaccurate energy function in PSP problem. Therefore, a novel way of predicting structures is to treat it as MaOOP, which utilizes different conflicting energy functions to alleviate the impact of any inaccurate energy functions. In MO4, the bond and SASA energies can converge

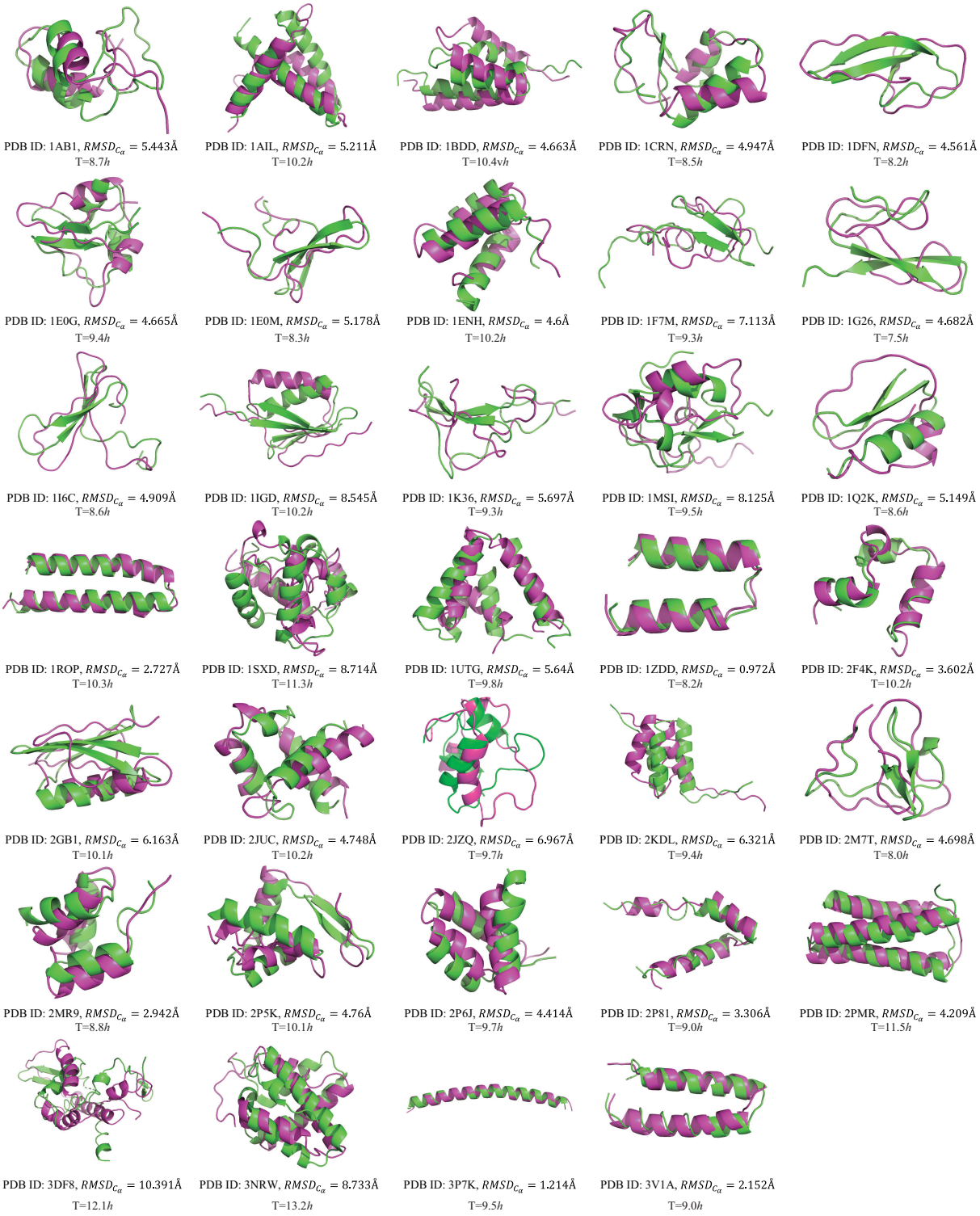


Figure 1.6: Comparison of the structure obtained by MO4 and native structure.

Table 1.7: The Wilcoxon and Friedman test results of objective analysis.

Parameters	bRMSD (Rank)	bGDT (Rank)	bTM (Rank)	bDME (Rank)	RMSD (Rank)	GDT (Rank)	TM (Rank)	DME (Rank)	Avg. Rank
$O_1 + O_2$	1.83E-07 (11)	1.83E-07 (11)	5.28E-07 (11)	1.83E-07 (11)	1.83E-07 (11)	1.83E-07 (11)	2.86E-07 (11)	1.83E-07 (11)	11.00
$O_1 + O_3$	1.66E-02 (4)	5.43E-05 (7)	1.98E-03 (9)	8.04E-04 (9)	4.76E-01 (2)	3.96E-03 (7)	3.20E-03 (8)	1.85E-01 (5)	6.38
$O_1 + O_4$	9.36E-05 (5)	4.13E-03 (3)	5.30E-03 (4)	4.95E-02 (3)	1.88E-03 (5)	1.64E-02 (3)	8.57E-02 (3)	1.63E-01 (3)	3.62
$O_2 + O_3$	1.82E-05 (8)	2.10E-06 (9)	3.04E-04 (8)	3.78E-05 (8)	1.57E-05 (9)	6.57E-05 (9)	2.46E-03 (7)	5.79E-06 (9)	8.38
$O_2 + O_4$	5.78E-05 (7)	7.36E-06 (6)	9.58E-04 (6)	1.81E-02 (4)	3.32E-06 (8)	9.80E-04 (4)	1.12E-02 (6)	2.04E-06 (8)	6.12
$O_3 + O_4$	5.89E-02 (2)	1.37E-02 (2)	1.47E-01 (2)	4.09E-01 (2)	4.15E-01 (3)	1.14E-01 (2)	1.50E-01 (2)	8.61E-01 (1)	2.00
$O_1 + O_2 + O_3$	1.59E-06 (9)	1.35E-06 (10)	4.68E-05 (10)	5.02E-05 (7)	3.60E-06 (7)	2.11E-05 (8)	1.16E-05 (9)	6.26E-06 (7)	8.38
$O_1 + O_2 + O_4$	2.00E-07 (10)	7.91E-06 (8)	1.27E-03 (7)	1.05E-06 (10)	3.41E-07 (10)	1.20E-06 (10)	3.45E-06 (10)	2.61E-07 (10)	9.38
$O_1 + O_3 + O_4$	3.45E-04 (3)	5.02E-05 (5)	6.13E-03 (3)	7.78E-03 (5)	1.40E-02 (4)	1.45E-03 (6)	7.08E-03 (4)	2.89E-01 (4)	4.25
$O_2 + O_3 + O_4$	4.95E-06 (6)	2.49E-06 (4)	1.68E-03 (5)	8.04E-04 (6)	1.07E-04 (6)	1.20E-03 (5)	4.80E-03 (5)	1.14E-03 (6)	5.38
MO4	– (1)	– (1)	– (1)	– (1)	– (1)	– (1)	– (1)	– (2)	1.13

to smaller values than in the native structures because they include some bond stretches, bond angles, dihedral angles, and improper angles. In a prediction process, protein structure can fold into a smaller volume, resulting in a decrease in bond angles, interatomic distances, and the structure surface area. However, the non-bond and RWplus energies increase with decreasing distance. Hence, these objectives are constrained.

1.4.4 Analysis of Preset Parameters and Objectives

MO4 has two important parameters K and L that affect its performance. They determine which conformations can be selected into the next population. K means the radius of dense region, and L represents the number of conformation that are selected from the a dense region. A parameter sensitivity analysis is performed to find the best parameter settings, and the Wilcoxon signed-rank [78] and Friedman tests are conducted to analyze the significant differences among different combinations of parameters. If the p-value of the Wilcoxon signed-rank test is less than 0.05, then the comparison between the best one and the current setting has a significant difference. The value of K and L varies in $\{1, 2, 5\}$ and $\{1, 2, 3\}$, respectively. The p-values of Wilcoxon results are summarized in Table 1.6, where the rank values obtained by the Friedman test are listed in brackets. From this table, it can be verified that the setting of $K = 5$ and $L = 3$ is the best, which is consistent with the suggestion in [71]. The experimental details of parameter analysis are shown in Tables 1.8 - 1.16.

In addition, regarding four objectives (i.e., bond objective, non-bond objective, SASA objective and Rwplus), an ablation study is performed to validate whether the four-objective methods are better than two/three-objective ones. The comparative experiment has 10 combinations ($C_4^2 + C_4^3 =$

Table 1.8: The experimental results of $K = 1, L = 1$.

PDB ID	bRMSD	bGDT	bTM	RMSD	GDT	TM	bDME	DME
1AB1	6.38	50.54	0.36	9.13	40.22	0.31	3.13	4.91
1AIL	5.75	47.50	0.41	11.76	37.14	0.30	2.68	5.77
1BDD	6.17	50.83	0.43	9.05	45.42	0.36	2.72	4.54
1CRN	6.87	48.37	0.34	8.23	42.39	0.30	2.67	4.39
1DFN	5.33	51.67	0.27	9.70	40.83	0.19	2.86	5.47
1E0G	4.99	51.04	0.41	9.19	41.15	0.30	2.35	4.21
1E0M	6.43	45.27	0.26	7.91	37.16	0.23	2.59	4.66
1ENH	5.72	48.15	0.40	10.68	39.81	0.32	3.17	4.93
1F7M	7.71	36.96	0.27	10.09	36.96	0.26	3.64	4.67
1G26	5.81	49.19	0.24	8.17	38.71	0.19	2.55	4.16
1I6C	6.28	42.31	0.25	10.86	35.90	0.19	3.00	6.17
1IGD	8.20	41.39	0.36	14.19	39.34	0.35	3.67	6.40
1K36	8.22	35.87	0.25	9.28	30.43	0.22	3.53	4.69
1MSI	8.38	30.68	0.25	16.52	22.73	0.19	3.67	8.12
1Q2K	3.89	57.26	0.34	6.77	45.97	0.29	2.17	4.10
1ROP	2.86	68.75	0.56	5.02	61.16	0.44	1.05	2.74
1SXD	11.16	31.87	0.32	13.43	25.00	0.23	5.09	7.06
1UTG	5.76	46.79	0.44	9.85	41.79	0.37	2.95	4.71
1ZDD	2.64	74.26	0.51	4.44	66.18	0.46	1.41	1.90
2F4K	4.14	60.61	0.35	5.29	53.03	0.30	1.79	2.22
2GB1	7.82	44.20	0.37	12.97	38.39	0.31	3.58	5.67
2JUC	5.57	49.55	0.38	5.89	43.18	0.30	2.64	3.13
2JZQ	7.54	39.47	0.32	10.77	34.21	0.27	3.64	4.91
2KDL	8.73	44.20	0.37	13.31	41.96	0.33	3.44	6.65
2M7T	5.03	48.48	0.27	8.16	37.12	0.21	2.76	3.28
2MR9	4.89	56.82	0.42	7.57	44.32	0.31	2.21	3.03
2P5K	6.67	46.43	0.37	9.42	36.51	0.29	3.29	6.39
2P6J	5.27	54.33	0.45	5.45	50.00	0.38	2.70	2.92
2P81	4.91	57.95	0.44	6.64	48.30	0.36	2.38	3.15
2PMR	5.12	52.63	0.47	10.27	44.74	0.38	2.35	4.38
3DF8	10.42	26.90	0.27	15.93	23.81	0.23	5.25	9.32
3NRW	9.93	33.82	0.32	18.59	27.21	0.26	4.87	12.42
3P7K	1.25	90.00	0.77	4.25	65.56	0.44	0.38	0.85
3V1A	2.48	81.25	0.68	4.03	70.31	0.53	1.20	1.69
Avg.	6.13	49.86	0.38	9.49	41.97	0.31	2.86	4.81

Table 1.9: The experimental results of $K = 1, L = 2$.

PDB ID	bRMSD	bGDT	bTM	RMSD	GDT	TM	bDME	DME
1AB1	6.21	49.46	0.38	8.53	42.39	0.32	2.56	4.38
1AIL	5.84	45.71	0.39	8.19	39.64	0.32	2.89	2.98
1BDD	4.24	59.58	0.49	6.52	50.83	0.41	1.95	2.18
1CRN	5.34	50.54	0.34	7.57	41.85	0.30	2.71	3.87
1DFN	5.78	47.50	0.25	8.96	35.83	0.17	3.16	4.67
1E0G	5.61	50.52	0.37	8.05	38.54	0.28	2.62	4.05
1E0M	4.81	51.35	0.30	8.86	39.19	0.24	2.22	3.72
1ENH	5.64	47.69	0.37	7.40	44.44	0.31	3.08	3.40
1F7M	6.36	40.22	0.28	11.11	28.26	0.19	2.98	5.83
1G26	4.59	52.42	0.26	6.37	45.97	0.23	2.36	2.68
1I6C	6.59	45.51	0.26	7.51	38.46	0.23	3.23	4.25
1IGD	6.08	47.95	0.38	11.60	42.62	0.34	2.89	5.33
1K36	8.11	38.59	0.28	11.87	29.89	0.19	3.92	7.06
1MSI	9.30	32.95	0.27	10.22	25.76	0.20	3.91	4.97
1Q2K	5.89	58.06	0.33	8.87	46.77	0.28	2.69	4.63
1ROP	3.22	67.41	0.54	4.43	61.16	0.49	1.10	1.86
1SXD	8.91	35.16	0.33	13.83	29.40	0.27	3.86	6.80
1UTG	6.87	48.57	0.44	10.24	42.50	0.36	3.11	4.12
1ZDD	1.82	87.50	0.68	2.28	77.94	0.52	0.81	0.99
2F4K	4.70	56.82	0.35	5.57	53.03	0.33	1.99	2.70
2GB1	7.87	45.09	0.38	13.02	38.39	0.32	3.31	5.72
2JUC	5.38	49.55	0.41	6.57	43.18	0.32	2.95	3.64
2JZQ	6.10	49.12	0.39	9.17	39.04	0.30	3.04	3.84
2KDL	8.40	46.43	0.34	10.47	40.18	0.30	3.27	4.81
2M7T	5.40	43.94	0.26	7.80	36.36	0.17	2.85	3.70
2MR9	3.27	67.05	0.50	6.27	50.00	0.34	1.81	2.54
2P5K	7.06	48.02	0.42	8.51	37.70	0.28	3.55	4.41
2P6J	5.37	53.85	0.41	6.92	42.31	0.33	2.73	3.62
2P81	4.33	57.95	0.40	6.23	48.86	0.35	2.13	2.92
2PMR	5.18	50.66	0.48	11.64	41.78	0.34	2.40	6.76
3DF8	10.36	28.33	0.30	12.64	23.81	0.25	4.91	7.13
3NRW	9.05	35.54	0.36	14.83	31.86	0.32	4.03	7.59
3P7K	1.47	89.44	0.77	3.64	70.56	0.50	0.47	1.07
3V1A	2.19	77.60	0.65	4.03	63.54	0.45	1.11	1.90
Avg.	5.80	51.65	0.39	8.52	43.00	0.31	2.72	4.12

Table 1.10: The experimental results of $K = 1, L = 3$.

PDB ID	bRMSD	bGDT	bTM	RMSD	GDT	TM	bDME	DME
1AB1	6.07	51.09	0.36	9.59	42.39	0.32	2.89	4.63
1AIL	6.04	45.71	0.43	9.17	42.50	0.36	2.76	3.66
1BDD	6.44	56.67	0.46	8.46	42.08	0.35	2.63	4.56
1CRN	5.81	49.46	0.38	8.50	42.93	0.34	3.00	4.16
1DFN	5.23	50.00	0.24	7.92	39.17	0.19	2.83	4.27
1E0G	6.05	53.12	0.44	9.71	43.75	0.32	3.53	5.21
1E0M	5.79	45.27	0.30	8.68	37.84	0.23	2.99	5.03
1ENH	4.98	57.41	0.47	6.86	44.44	0.32	2.58	3.13
1F7M	7.50	40.76	0.28	10.13	33.70	0.25	3.43	4.20
1G26	6.09	49.19	0.30	8.29	40.32	0.22	2.81	3.98
1I6C	6.16	42.95	0.28	9.71	35.26	0.20	2.69	4.56
1IGD	8.86	46.72	0.40	13.45	35.66	0.31	4.57	5.96
1K36	6.91	40.22	0.26	7.53	39.67	0.22	3.47	3.86
1MSI	8.96	29.55	0.23	15.54	25.38	0.22	4.04	8.72
1Q2K	4.38	59.68	0.32	6.61	46.77	0.28	2.34	3.92
1ROP	3.30	67.86	0.54	5.07	58.04	0.43	1.40	2.27
1SXD	9.50	34.07	0.34	11.92	29.40	0.29	4.41	5.26
1UTG	5.43	51.79	0.47	9.31	43.57	0.38	2.90	4.33
1ZDD	1.94	80.15	0.55	3.87	66.91	0.43	0.90	1.63
2F4K	4.81	53.79	0.32	6.12	47.73	0.29	2.08	2.63
2GB1	6.45	52.68	0.43	10.23	45.09	0.38	2.70	5.21
2JUC	5.94	48.64	0.40	7.53	38.64	0.28	2.77	3.97
2JZQ	6.94	38.60	0.31	9.46	32.46	0.23	3.33	4.47
2KDL	7.49	44.64	0.36	12.75	40.18	0.32	3.67	6.45
2M7T	5.74	46.97	0.23	7.07	35.61	0.17	2.96	3.65
2MR9	4.01	62.50	0.44	6.52	50.57	0.34	2.05	3.18
2P5K	6.34	47.62	0.39	8.51	40.08	0.33	2.95	4.75
2P6J	4.73	56.73	0.42	5.24	53.37	0.42	2.31	2.99
2P81	3.59	64.20	0.47	4.46	56.82	0.37	1.98	2.50
2PMR	6.38	52.30	0.48	10.75	43.42	0.37	2.58	6.89
3DF8	11.74	27.38	0.28	16.63	22.62	0.20	5.27	8.50
3NRW	10.30	34.07	0.34	15.35	27.45	0.28	4.48	8.93
3P7K	1.41	86.67	0.74	3.10	70.00	0.49	0.48	0.72
3V1A	1.89	80.73	0.70	2.88	69.27	0.50	0.97	1.44
Avg.	5.98	51.45	0.39	8.73	43.03	0.31	2.85	4.40

Table 1.11: The experimental results of $K = 2, L = 1$.

PDB ID	bRMSD	bGDT	bTM	RMSD	GDT	TM	bDME	DME
1AB1	5.89	51.09	0.36	9.43	42.39	0.33	2.77	5.21
1AIL	6.64	43.93	0.41	9.56	42.14	0.34	3.00	4.69
1BDD	6.04	56.25	0.47	9.56	41.67	0.33	2.55	3.88
1CRN	6.09	49.46	0.36	9.47	39.13	0.31	2.43	4.40
1DFN	5.43	52.50	0.25	7.60	39.17	0.18	2.88	3.55
1E0G	5.18	50.00	0.38	7.65	46.88	0.34	2.71	4.79
1E0M	4.80	50.68	0.29	6.84	41.22	0.22	2.42	3.07
1ENH	4.25	57.41	0.43	8.56	44.44	0.36	2.34	4.20
1F7M	7.04	36.41	0.26	11.70	29.89	0.20	3.56	6.10
1G26	5.63	50.81	0.29	7.55	40.32	0.21	2.58	3.99
1I6C	6.42	38.46	0.24	10.45	32.05	0.19	3.43	5.18
1IGD	9.74	42.62	0.39	12.56	35.25	0.31	4.67	6.47
1K36	7.44	39.13	0.27	11.51	30.98	0.21	3.37	5.15
1MSI	7.59	31.82	0.26	12.47	23.11	0.19	3.91	7.24
1Q2K	5.26	54.84	0.32	6.86	47.58	0.29	2.60	3.76
1ROP	3.77	63.84	0.49	5.45	54.91	0.41	1.37	2.16
1SXD	8.55	33.24	0.31	15.67	23.63	0.22	4.27	10.09
1UTG	6.37	49.64	0.46	9.86	41.07	0.37	3.14	4.00
1ZDD	2.71	72.06	0.50	4.47	62.50	0.43	1.30	2.16
2F4K	4.16	61.36	0.39	5.37	53.03	0.32	1.84	2.21
2GB1	6.48	43.30	0.36	11.25	34.38	0.28	3.12	6.40
2JUC	6.31	45.91	0.39	10.34	36.36	0.26	3.10	5.01
2JZQ	5.84	45.18	0.33	7.06	37.28	0.26	2.54	4.39
2KDL	9.05	43.75	0.36	14.21	38.84	0.33	3.92	7.91
2M7T	5.34	48.48	0.24	7.76	34.85	0.19	2.84	3.82
2MR9	4.46	62.50	0.44	7.21	46.59	0.36	2.00	2.62
2P5K	7.53	44.84	0.37	8.97	33.33	0.29	3.76	5.28
2P6J	5.23	54.81	0.40	5.31	51.92	0.38	2.83	2.92
2P81	3.82	65.91	0.50	5.55	60.80	0.46	1.74	2.53
2PMR	4.74	55.26	0.46	7.42	40.79	0.35	2.35	4.23
3DF8	11.18	29.76	0.28	17.66	23.81	0.21	5.63	10.55
3NRW	11.54	30.88	0.32	19.91	25.98	0.25	4.75	11.33
3P7K	1.36	91.11	0.79	2.58	77.78	0.64	0.51	0.82
3V1A	3.26	69.79	0.54	4.07	63.54	0.48	1.22	1.73
Avg.	6.03	50.50	0.38	9.17	41.69	0.31	2.87	4.76

Table 1.12: The experimental results of $K = 2, L = 2$.

PDB ID	bRMSD	bGDT	bTM	RMSD	GDT	TM	bDME	DME
1AB1	5.73	52.72	0.41	8.09	44.02	0.33	2.84	4.49
1AIL	5.25	48.21	0.39	8.36	38.57	0.31	2.55	3.83
1BDD	5.23	52.92	0.45	7.27	43.75	0.35	2.25	3.69
1CRN	5.90	48.91	0.36	8.07	45.65	0.31	2.78	3.85
1DFN	6.10	45.83	0.24	11.15	35.83	0.17	3.18	7.66
1E0G	5.52	50.00	0.42	8.23	44.27	0.31	2.62	4.08
1E0M	5.00	49.32	0.31	5.55	47.30	0.24	2.29	3.18
1ENH	5.48	50.46	0.41	7.01	43.98	0.34	2.87	3.43
1F7M	6.72	39.13	0.29	8.45	28.80	0.21	3.17	4.50
1G26	5.06	50.00	0.26	7.50	41.13	0.21	2.48	4.07
1I6C	6.51	45.51	0.27	8.30	33.33	0.19	3.14	4.18
1IGD	8.24	42.21	0.37	10.70	36.07	0.31	3.88	4.52
1K36	7.81	38.04	0.26	8.95	33.15	0.20	3.25	4.23
1MSI	8.49	32.58	0.28	11.61	24.24	0.19	3.39	5.89
1Q2K	5.23	57.26	0.38	7.30	46.77	0.30	2.33	3.30
1ROP	3.56	64.29	0.50	5.28	57.14	0.45	1.31	2.08
1SXD	9.07	32.97	0.32	11.41	29.67	0.28	4.57	6.28
1UTG	5.00	52.86	0.45	10.75	41.79	0.40	2.51	4.73
1ZDD	2.55	74.26	0.47	4.60	62.50	0.44	1.34	2.00
2F4K	4.50	55.30	0.33	5.33	50.76	0.29	2.01	2.57
2GB1	6.98	48.66	0.39	10.56	40.18	0.31	3.07	5.91
2JUC	5.53	49.09	0.36	7.01	40.45	0.30	2.51	3.18
2JZQ	6.05	42.98	0.32	9.01	34.65	0.28	3.27	4.31
2KDL	8.94	44.64	0.36	13.33	38.39	0.31	3.92	6.52
2M7T	5.29	48.48	0.26	7.19	40.91	0.22	2.65	4.10
2MR9	4.50	55.68	0.38	5.90	48.86	0.34	2.30	2.30
2P5K	6.48	46.03	0.42	9.48	37.70	0.30	2.96	4.75
2P6J	5.44	49.04	0.43	6.30	46.15	0.28	2.52	3.05
2P81	4.17	63.64	0.44	6.16	50.00	0.35	2.08	2.82
2PMR	5.96	49.67	0.47	13.08	42.76	0.37	2.79	6.78
3DF8	11.67	26.90	0.26	18.08	21.43	0.21	4.66	9.25
3NRW	10.74	32.35	0.33	13.85	28.19	0.28	4.49	6.19
3P7K	0.92	96.67	0.88	3.43	67.22	0.44	0.40	1.04
3V1A	3.03	71.35	0.56	3.93	65.10	0.47	1.15	1.56
Avg.	5.96	50.23	0.38	8.56	42.08	0.30	2.75	4.24

Table 1.13: The experimental results of $K = 2, L = 3$.

PDB ID	bRMSD	bGDT	bTM	RMSD	GDT	TM	bDME	DME
1AB1	6.13	47.28	0.36	8.93	40.76	0.32	2.78	4.21
1AIL	6.42	43.93	0.40	9.22	38.93	0.32	3.10	4.67
1BDD	4.80	55.00	0.45	8.09	42.08	0.32	2.08	2.73
1CRN	6.20	47.28	0.35	7.24	44.02	0.30	2.36	3.59
1DFN	7.49	42.50	0.21	12.58	37.50	0.17	3.60	8.56
1E0G	5.55	52.08	0.40	7.96	41.67	0.30	2.88	3.58
1E0M	5.66	47.97	0.33	8.69	37.84	0.25	2.63	2.65
1ENH	5.72	51.85	0.44	7.91	44.91	0.37	2.70	4.11
1F7M	6.93	37.50	0.27	9.01	32.61	0.21	3.31	5.07
1G26	5.49	48.39	0.26	7.42	40.32	0.22	2.85	4.14
1I6C	6.16	42.31	0.24	9.92	34.62	0.20	2.90	4.54
1IGD	7.67	43.03	0.36	10.64	37.70	0.32	3.66	6.25
1K36	7.30	39.13	0.28	9.47	34.78	0.21	3.44	4.13
1MSI	8.94	31.44	0.26	12.53	26.14	0.22	4.06	5.89
1Q2K	4.79	58.87	0.35	5.49	51.61	0.32	2.00	3.19
1ROP	3.81	64.29	0.53	5.07	56.70	0.41	1.19	2.49
1SXD	9.55	33.24	0.33	13.73	26.10	0.25	4.24	6.61
1UTG	5.05	53.93	0.46	8.63	42.86	0.38	2.65	4.07
1ZDD	2.06	83.09	0.58	3.73	66.91	0.46	1.03	1.41
2F4K	4.00	62.12	0.41	4.73	53.79	0.34	2.01	2.46
2GB1	7.35	45.98	0.38	10.44	36.61	0.31	3.61	4.76
2JUC	5.20	47.27	0.39	8.22	38.18	0.28	2.62	3.79
2JZQ	6.77	46.49	0.39	9.61	37.28	0.27	3.19	4.73
2KDL	7.36	47.77	0.37	10.23	40.18	0.32	3.28	4.09
2M7T	6.18	43.94	0.23	8.46	35.61	0.19	2.88	5.02
2MR9	4.27	61.36	0.46	6.74	48.30	0.35	2.13	2.99
2P5K	7.58	42.86	0.38	11.60	37.70	0.30	3.20	6.12
2P6J	4.65	55.77	0.41	5.54	50.00	0.36	2.38	2.79
2P81	4.64	59.09	0.43	6.09	54.55	0.39	2.04	2.29
2PMR	4.67	56.91	0.55	6.31	47.70	0.37	2.71	4.05
3DF8	11.81	27.14	0.26	16.71	21.19	0.19	5.75	7.23
3NRW	10.32	35.78	0.36	16.39	30.15	0.27	4.63	8.59
3P7K	1.31	88.33	0.75	3.66	71.11	0.51	0.45	0.80
3V1A	2.32	81.77	0.70	2.97	72.40	0.61	1.11	1.51
Avg.	6.00	50.76	0.39	8.65	42.73	0.31	2.81	4.21

Table 1.14: The experimental results of $K = 5, L = 1$.

PDB ID	bRMSD	bGDT	bTM	RMSD	GDT	TM	bDME	DME
1AB1	6.85	47.83	0.35	8.58	40.76	0.33	2.85	4.33
1AIL	6.70	45.71	0.39	10.13	40.36	0.33	3.10	6.33
1BDD	4.62	53.33	0.41	10.17	43.75	0.34	2.83	5.47
1CRN	5.70	45.65	0.34	8.15	40.22	0.30	2.49	4.61
1DFN	5.11	55.00	0.25	7.58	40.00	0.21	2.28	4.03
1E0G	8.39	45.31	0.32	12.79	34.38	0.25	3.70	6.97
1E0M	5.98	45.95	0.27	8.89	35.81	0.23	2.47	4.50
1ENH	6.58	49.54	0.40	11.53	40.28	0.34	3.20	6.10
1F7M	7.97	40.22	0.30	8.93	31.52	0.20	3.43	5.17
1G26	5.44	50.81	0.29	6.93	46.77	0.29	2.65	4.22
1I6C	6.46	49.36	0.29	9.22	31.41	0.18	2.80	4.80
1IGD	8.27	40.98	0.35	13.09	34.84	0.30	3.88	6.00
1K36	8.11	38.04	0.28	10.04	32.61	0.22	3.85	4.92
1MSI	10.10	31.44	0.24	14.76	24.24	0.20	4.30	8.09
1Q2K	5.28	57.26	0.35	8.03	47.58	0.30	2.15	4.59
1ROP	2.78	71.88	0.56	3.82	62.95	0.45	1.28	1.95
1SXD	9.26	32.42	0.33	14.06	28.02	0.29	4.71	7.72
1UTG	6.49	48.21	0.44	8.91	41.79	0.39	2.88	4.42
1ZDD	3.17	69.12	0.46	5.46	58.09	0.43	1.22	2.47
2F4K	3.63	61.36	0.35	4.86	55.30	0.33	1.88	2.03
2GB1	8.49	44.20	0.36	13.00	36.61	0.31	3.74	7.28
2JUC	5.70	47.73	0.38	8.31	38.64	0.31	2.85	4.77
2JZQ	7.91	38.16	0.31	10.93	32.89	0.25	3.97	7.61
2KDL	8.50	43.75	0.36	13.16	36.61	0.31	3.73	7.04
2M7T	5.73	45.45	0.23	8.96	34.09	0.19	3.04	3.98
2MR9	4.91	54.55	0.40	6.63	47.16	0.31	2.23	2.78
2P5K	7.37	41.27	0.36	10.10	33.73	0.27	3.41	5.51
2P6J	5.11	54.33	0.42	9.46	43.75	0.35	2.72	3.67
2P81	3.53	63.07	0.44	6.46	50.00	0.36	1.99	2.62
2PMR	5.11	51.64	0.50	8.06	42.11	0.38	2.35	5.32
3DF8	10.76	28.33	0.27	13.00	22.38	0.24	5.48	7.69
3NRW	9.28	32.35	0.33	15.85	27.94	0.25	4.25	11.25
3P7K	1.06	91.67	0.82	3.38	71.11	0.51	0.42	0.85
3V1A	3.75	66.15	0.52	5.28	60.42	0.42	1.38	2.07
Avg.	6.30	49.47	0.37	9.37	40.83	0.30	2.93	5.04

Table 1.15: The experimental results of $K = 5, L = 2$.

PDB ID	bRMSD	bGDT	bTM	RMSD	GDT	TM	bDME	DME
1AB1	6.05	50.54	0.38	8.57	42.93	0.32	2.95	4.25
1AIL	5.77	50.71	0.43	7.65	39.64	0.32	2.48	3.28
1BDD	4.69	60.00	0.53	7.09	45.83	0.37	2.24	3.27
1CRN	5.04	53.26	0.37	8.14	42.39	0.31	2.57	3.22
1DFN	5.49	51.67	0.25	7.11	44.17	0.18	2.44	3.41
1E0G	5.90	48.96	0.37	10.51	40.62	0.25	2.99	5.93
1E0M	5.05	49.32	0.29	6.29	47.30	0.22	2.57	3.13
1ENH	5.16	54.63	0.42	7.19	53.70	0.39	2.60	3.72
1F7M	7.40	38.04	0.27	9.33	29.89	0.20	3.32	4.58
1G26	5.70	45.97	0.25	7.19	42.74	0.23	2.67	3.31
1I6C	5.49	45.51	0.28	8.97	35.26	0.20	2.96	3.86
1IGD	7.39	52.87	0.43	11.02	45.49	0.35	3.46	5.34
1K36	7.43	36.96	0.27	10.24	29.35	0.19	3.58	5.07
1MSI	7.25	33.33	0.27	9.09	31.06	0.25	3.41	3.95
1Q2K	4.23	58.87	0.36	7.01	48.39	0.30	1.84	3.91
1ROP	2.59	69.64	0.58	5.37	56.70	0.46	1.09	1.91
1SXD	9.04	35.71	0.35	12.14	28.02	0.26	3.96	4.64
1UTG	5.85	46.07	0.43	9.04	38.93	0.35	3.18	4.26
1ZDD	3.13	69.85	0.49	4.64	62.50	0.43	1.41	2.16
2F4K	4.05	57.58	0.35	5.52	50.00	0.28	2.18	2.47
2GB1	8.05	43.75	0.36	10.28	35.71	0.30	2.83	5.23
2JUC	4.87	51.36	0.42	7.31	40.91	0.31	2.43	2.97
2JZQ	5.84	43.86	0.33	7.92	34.21	0.27	3.35	3.96
2KDL	7.14	46.88	0.36	10.80	41.96	0.31	3.39	5.16
2M7T	4.91	47.73	0.23	5.43	46.97	0.21	2.54	2.75
2MR9	4.25	61.93	0.49	5.59	50.00	0.35	2.10	2.48
2P5K	6.54	46.43	0.39	9.42	46.03	0.37	3.16	4.57
2P6J	4.91	52.40	0.42	6.23	49.04	0.35	2.56	3.51
2P81	4.81	56.82	0.42	6.47	51.14	0.36	2.14	2.56
2PMR	6.30	48.36	0.46	10.55	41.78	0.35	2.71	6.39
3DF8	10.86	27.86	0.28	13.87	23.57	0.25	4.80	6.74
3NRW	9.44	34.31	0.37	12.94	32.11	0.31	4.41	7.09
3P7K	1.42	89.44	0.79	3.05	73.89	0.55	0.46	0.85
3V1A	2.41	78.65	0.66	3.71	67.19	0.52	1.17	1.81
Avg.	5.72	51.16	0.39	8.11	43.81	0.31	2.70	3.87

Table 1.16: The experimental results of the MO4 (with $K = 5, L = 3$).

PDB ID	bRMSD	bGDT	bTM	RMSD	GDT	TM	bDME	DME
1AB1	5.44	55.98	0.41	6.82	48.37	0.32	2.55	3.59
1AIL	5.20	51.79	0.46	6.07	44.29	0.38	2.73	3.21
1BDD	4.21	55.00	0.44	6.98	41.67	0.33	1.98	3.32
1CRN	4.95	57.61	0.43	6.79	50.00	0.37	2.46	3.30
1DFN	4.56	55.83	0.29	8.33	41.67	0.20	2.36	4.24
1E0G	4.67	50.52	0.37	8.88	40.10	0.31	2.39	4.33
1E0M	5.18	50.00	0.31	7.52	39.19	0.24	2.37	3.05
1ENH	4.60	56.48	0.40	4.60	56.48	0.40	2.38	2.38
1F7M	7.11	45.11	0.34	10.09	29.89	0.20	3.33	5.42
1G26	4.68	56.45	0.32	5.89	51.61	0.27	2.31	2.91
1I6C	4.91	49.36	0.29	6.93	40.38	0.23	2.30	3.06
1IGD	8.54	47.95	0.40	9.71	40.16	0.35	3.92	5.76
1K36	5.70	44.57	0.28	9.04	36.41	0.22	2.83	3.96
1MSI	8.12	32.95	0.28	10.38	25.38	0.21	4.27	5.30
1Q2K	4.33	62.10	0.37	6.50	49.19	0.30	1.73	3.23
1ROP	2.73	71.88	0.58	2.73	71.88	0.58	1.04	1.04
1SXD	8.54	33.52	0.31	9.35	28.85	0.26	3.98	4.77
1UTG	5.64	47.86	0.42	7.91	37.86	0.33	3.03	3.24
1ZDD	0.97	94.85	0.78	1.87	77.94	0.47	0.52	1.20
2F4K	3.60	61.36	0.34	5.38	53.03	0.30	1.79	2.28
2GB1	6.16	49.11	0.39	7.62	45.09	0.33	2.66	3.41
2JUC	4.75	53.18	0.38	7.70	45.91	0.38	2.18	2.85
2JZQ	6.97	41.67	0.31	7.87	35.96	0.27	3.43	4.11
2KDL	6.32	47.32	0.37	7.12	41.07	0.32	3.03	3.89
2M7T	4.88	50.76	0.25	4.88	50.76	0.21	2.78	2.83
2MR9	2.94	67.05	0.46	5.19	51.70	0.36	1.73	2.64
2P5K	4.76	50.79	0.37	8.89	44.84	0.35	2.68	4.90
2P6J	4.41	58.17	0.48	6.20	50.96	0.40	2.55	3.85
2P81	3.31	71.02	0.58	4.67	57.95	0.40	1.66	2.18
2PMR	4.21	57.57	0.54	7.74	46.05	0.40	1.99	4.32
3DF8	8.95	34.52	0.37	13.21	21.90	0.22	4.57	5.36
3NRW	8.49	34.80	0.36	12.51	28.68	0.29	4.17	5.77
3P7K	1.21	96.11	0.89	1.50	91.11	0.81	0.30	0.49
3V1A	2.15	81.77	0.73	3.72	70.31	0.53	1.00	1.58
Avg.	5.09	55.15	0.42	7.08	46.67	0.34	2.50	3.46

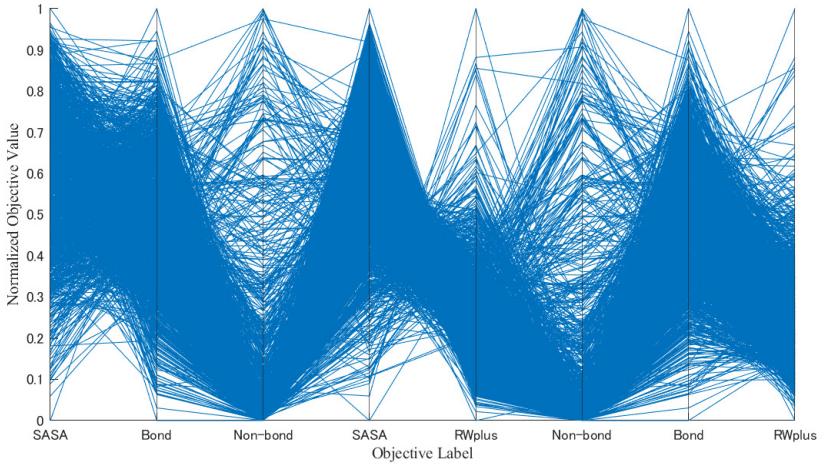


Figure 1.7: The parallel coordinates plot for protein 2P5K.

10). It is worth noting that a large number of studies have demonstrated already that multi-objective methods significantly outperform single-objective ones [79]. Hence, we just need to compare MO4 with the two and three-objective variants. The Wilcoxon and Friedman test results of objective analysis are shown in Table 1.7, where O_1, O_2, O_3, and O_4 represent the Bond, Non-bond, SASA, and Rwplus energy, respectively. From it, we can claim that MO4 can generate more accurate structures than the others in terms of seven performance metrics, and ranks the second on the metric DME. The reason is that the carefully selected four objectives can precisely evaluate and restrict the conformation in different aspects. The details of experimental results of objective analysis are listed in Tables 1.17 − 1.26.

1.4.5 Conflicting Relationship of Four Objectives

The MaOEA optimizes more than three conflicting objectives simultaneously. This feature is typical in MaOOP. A parallel coordinate plot is used to analyze the conflicting relationship among objectives[80]. The objective labels are located along the horizontal axis. The vertical axis means the normalized objective value. An objective vector is displayed by joining the performance levels in all adjacent objectives by straight lines. For the adjacent objectives, the lines represent the conflicting relationship between them. If the lines are cross, the adjacent objectives are conflicting. Otherwise, they are non-conflicting. To directly exhibit the conflicting relationship between any two

Table 1.17: The experimental results of $O_1 + O_2$.

PDB ID	bRMSD	bGDT	bTM	RMSD	GDT	TM	bDME	DME
1AB1	7.65	45.65	0.33	9.61	40.22	0.31	3.95	5.15
1AIL	8.61	41.79	0.36	13.77	36.43	0.31	4.10	7.48
1BDD	9.39	42.92	0.34	14.07	39.17	0.29	5.07	8.96
1CRN	8.07	44.02	0.33	9.92	39.67	0.30	3.49	6.11
1DFN	8.96	41.67	0.24	13.63	30.00	0.20	6.21	10.27
1E0G	8.21	44.79	0.33	13.32	35.42	0.28	3.79	8.59
1E0M	7.42	39.19	0.22	12.51	31.08	0.18	3.78	8.46
1ENH	8.43	46.76	0.38	12.54	35.65	0.29	3.64	7.82
1F7M	7.83	35.87	0.24	11.81	27.17	0.18	4.23	6.66
1G26	6.67	43.55	0.27	10.71	36.29	0.22	3.43	5.96
1I6C	8.43	37.82	0.24	12.75	29.49	0.19	4.21	8.27
1IGD	10.59	43.44	0.36	13.86	36.07	0.33	6.08	7.20
1K36	7.55	35.33	0.28	12.00	29.89	0.22	4.10	7.40
1MSI	10.92	27.65	0.23	17.57	21.97	0.19	5.46	9.48
1Q2K	4.71	54.84	0.34	8.34	42.74	0.29	2.74	5.20
1ROP	5.11	58.04	0.48	6.26	53.57	0.45	2.46	3.59
1SXD	12.26	28.02	0.27	17.67	22.80	0.21	6.97	12.31
1UTG	9.15	43.57	0.38	14.31	32.50	0.28	4.36	9.82
1ZDD	4.23	64.71	0.47	6.03	60.29	0.42	1.71	3.86
2F4K	4.87	57.58	0.36	6.09	49.24	0.31	2.21	3.47
2GB1	11.59	44.20	0.40	16.12	38.39	0.32	6.56	8.45
2JUC	8.72	40.91	0.31	12.91	32.27	0.26	4.23	7.24
2JZQ	8.28	40.35	0.30	10.13	30.70	0.25	5.64	7.20
2KDL	10.16	41.96	0.32	12.85	37.05	0.30	4.52	7.05
2M7T	6.28	44.70	0.25	8.54	36.36	0.18	3.35	5.09
2MR9	5.23	51.70	0.36	12.29	36.93	0.28	3.61	7.37
2P5K	8.50	38.10	0.32	15.10	31.75	0.26	3.94	9.06
2P6J	8.84	48.56	0.39	11.23	42.31	0.33	4.93	6.56
2P81	4.95	57.39	0.38	8.90	42.61	0.34	2.26	5.21
2PMR	9.12	43.42	0.40	14.91	36.84	0.34	6.34	10.09
3DF8	14.59	28.10	0.27	18.82	20.71	0.20	6.96	11.58
3NRW	14.10	30.88	0.30	22.00	25.74	0.24	8.53	15.39
3P7K	1.55	87.22	0.76	3.51	73.89	0.53	0.50	0.72
3V1A	3.05	66.67	0.52	5.03	60.94	0.49	1.87	2.22
Avg.	8.06	45.33	0.35	12.03	37.53	0.29	4.27	7.33

Table 1.18: The experimental results of $O_1 + O_3$.

PDB ID	bRMSD	bGDT	bTM	RMSD	GDT	TM	bDME	DME
1AB1	6.66	45.11	0.32	7.32	40.76	0.29	3.18	4.21
1AIL	6.87	44.29	0.38	8.22	35.00	0.31	3.10	4.44
1BDD	5.98	45.83	0.38	8.27	39.17	0.30	2.86	4.37
1CRN	5.57	48.91	0.35	7.07	42.39	0.30	3.10	3.78
1DFN	4.04	55.00	0.28	5.54	49.17	0.24	1.90	2.52
1E0G	6.07	44.27	0.32	8.98	35.42	0.23	3.03	3.92
1E0M	4.47	54.73	0.31	5.92	47.30	0.25	2.27	2.76
1ENH	5.22	50.46	0.41	7.96	40.28	0.32	2.88	3.84
1F7M	6.05	41.30	0.31	8.51	34.78	0.21	3.43	4.20
1G26	5.96	44.35	0.24	7.14	37.10	0.22	3.85	4.48
1I6C	4.83	50.64	0.29	6.55	40.38	0.24	2.37	3.06
1IGD	5.99	47.13	0.39	8.53	37.30	0.31	3.14	4.14
1K36	6.26	43.48	0.29	7.48	35.87	0.22	3.19	4.45
1MSI	7.40	32.95	0.28	9.47	28.79	0.23	4.02	4.36
1Q2K	5.18	53.23	0.34	5.33	53.23	0.32	2.98	3.59
1ROP	3.23	69.20	0.54	4.01	58.48	0.44	1.39	1.95
1SXD	7.23	35.71	0.34	9.08	28.85	0.26	3.69	4.21
1UTG	5.92	47.86	0.44	6.39	45.36	0.36	2.87	3.51
1ZDD	2.95	70.59	0.46	3.86	61.76	0.35	1.44	2.04
2F4K	4.65	59.09	0.41	5.61	52.27	0.31	2.67	2.94
2GB1	6.12	47.32	0.38	8.36	41.96	0.35	3.17	3.35
2JUC	5.79	47.27	0.38	7.51	39.09	0.30	2.71	3.95
2JZQ	6.19	42.11	0.34	7.50	36.84	0.30	3.33	3.53
2KDL	4.96	51.79	0.41	7.14	41.96	0.34	2.22	3.19
2M7T	4.94	47.73	0.23	6.62	38.64	0.17	2.88	3.40
2MR9	5.08	53.98	0.37	6.21	46.59	0.30	2.65	3.35
2P5K	6.00	44.44	0.38	7.81	36.11	0.29	3.04	3.97
2P6J	5.45	47.60	0.36	6.17	41.83	0.28	2.84	3.18
2P81	7.11	48.86	0.32	8.25	42.61	0.30	5.27	6.33
2PMR	4.56	51.97	0.47	6.32	43.75	0.38	2.43	2.72
3DF8	9.85	28.10	0.31	10.92	24.29	0.24	4.95	6.65
3NRW	6.70	41.18	0.43	8.74	37.50	0.38	3.61	4.20
3P7K	5.34	49.44	0.32	5.91	43.33	0.26	3.85	4.54
3V1A	4.34	54.69	0.37	4.89	50.52	0.37	3.00	3.00
Avg.	5.68	48.25	0.36	7.16	41.43	0.29	3.04	3.77

Table 1.19: The experimental results of $O_1 + O_4$.

PDB ID	bRMSD	bGDT	bTM	RMSD	GDT	TM	bDME	DME
1AB1	6.67	47.28	0.36	7.76	42.93	0.34	2.86	3.60
1AIL	5.31	52.14	0.45	6.74	45.71	0.39	2.20	2.82
1BDD	9.89	42.92	0.35	10.78	42.08	0.34	3.10	3.64
1CRN	5.57	47.83	0.32	6.33	44.02	0.28	2.67	3.00
1DFN	4.62	52.50	0.25	5.68	50.83	0.21	2.08	2.94
1E0G	5.28	54.69	0.39	9.89	44.79	0.32	2.35	4.42
1E0M	5.13	50.68	0.29	7.20	39.19	0.21	2.61	4.12
1ENH	6.51	45.37	0.33	8.99	42.59	0.30	2.95	3.26
1F7M	7.55	34.78	0.22	9.43	27.17	0.21	3.66	4.24
1G26	5.01	54.03	0.29	6.26	45.97	0.24	2.16	3.27
1I6C	5.32	47.44	0.28	7.66	37.82	0.21	2.84	4.22
1IGD	7.80	44.67	0.37	12.22	37.70	0.32	3.82	5.22
1K36	7.64	36.96	0.25	8.55	30.43	0.19	3.55	4.53
1MSI	8.96	34.09	0.28	12.31	28.41	0.22	3.79	5.28
1Q2K	5.80	52.42	0.31	7.03	41.13	0.28	2.97	3.45
1ROP	3.67	66.96	0.54	5.87	56.70	0.47	1.26	1.72
1SXD	8.04	37.09	0.39	12.31	30.22	0.27	3.77	5.00
1UTG	6.64	44.29	0.36	7.20	41.07	0.28	2.59	3.01
1ZDD	1.57	90.44	0.75	3.01	72.79	0.49	0.74	1.48
2F4K	6.04	51.52	0.31	6.81	48.48	0.30	2.53	2.81
2GB1	7.09	48.21	0.39	11.15	40.62	0.32	3.43	5.91
2JUC	6.16	44.55	0.37	8.34	36.82	0.28	2.69	3.32
2JZQ	6.24	44.74	0.34	8.50	36.84	0.30	3.13	4.45
2KDL	7.03	51.34	0.40	9.02	42.86	0.32	2.74	3.71
2M7T	5.61	47.73	0.22	6.67	43.18	0.20	2.54	3.26
2MR9	3.49	65.91	0.48	5.48	55.68	0.39	1.79	2.62
2P5K	5.03	50.40	0.40	9.81	39.68	0.33	2.58	2.91
2P6J	4.35	60.58	0.50	6.79	45.19	0.30	2.33	3.01
2P81	3.51	69.89	0.55	4.69	63.07	0.50	1.75	2.27
2PMR	4.08	60.20	0.54	6.91	43.75	0.39	2.03	3.25
3DF8	10.28	32.62	0.37	10.77	25.00	0.27	4.71	5.43
3NRW	8.85	39.71	0.38	13.34	30.39	0.30	4.61	7.24
3P7K	0.64	98.89	0.92	1.11	93.89	0.83	0.24	0.32
3V1A	2.35	75.52	0.63	3.44	67.71	0.52	1.19	1.41
Avg.	5.82	52.31	0.40	7.88	44.55	0.33	2.65	3.56

Table 1.20: The experimental results of $O_2 + O_3$.

PDB ID	bRMSD	bGDT	bTM	RMSD	GDT	TM	bDME	DME
1AB1	5.80	52.17	0.36	7.02	43.48	0.33	2.82	3.61
1AIL	6.16	46.43	0.39	8.79	40.71	0.32	2.89	4.85
1BDD	4.29	59.58	0.49	4.93	50.42	0.45	2.09	2.77
1CRN	6.19	53.80	0.40	9.13	40.76	0.30	2.79	4.73
1DFN	4.04	55.83	0.28	5.79	46.67	0.21	2.34	2.98
1E0G	7.04	47.40	0.33	11.61	36.46	0.25	3.23	5.88
1E0M	5.40	47.97	0.30	6.10	42.57	0.25	2.73	3.29
1ENH	6.09	49.54	0.43	9.02	39.35	0.32	2.94	4.22
1F7M	7.03	36.96	0.26	10.62	29.89	0.23	3.29	5.59
1G26	5.04	50.00	0.27	7.84	40.32	0.20	2.53	4.24
1I6C	5.86	42.95	0.25	9.47	33.33	0.19	3.30	4.70
1IGD	5.90	47.13	0.37	14.18	34.43	0.31	3.04	6.91
1K36	8.22	36.96	0.24	10.80	34.78	0.24	4.00	5.06
1MSI	8.87	32.58	0.28	12.51	25.00	0.19	3.74	6.12
1Q2K	5.53	54.03	0.35	8.00	45.97	0.29	2.57	4.21
1ROP	3.24	66.52	0.54	4.67	59.82	0.44	1.15	2.29
1SXD	9.53	32.14	0.31	12.40	24.73	0.23	4.06	6.62
1UTG	5.96	45.71	0.40	8.81	36.79	0.33	3.02	3.85
1ZDD	2.27	78.68	0.54	3.86	67.65	0.46	1.13	1.74
2F4K	3.76	63.64	0.40	5.26	53.79	0.36	1.95	2.39
2GB1	8.78	44.64	0.38	13.44	38.84	0.33	4.05	7.30
2JUC	5.75	44.55	0.36	8.00	40.91	0.33	2.85	4.15
2JZQ	6.84	39.47	0.30	10.17	35.53	0.25	3.75	5.67
2KDL	8.63	44.20	0.36	12.03	37.95	0.33	3.13	5.58
2M7T	5.37	46.97	0.26	8.79	34.85	0.18	2.75	3.52
2MR9	5.13	56.82	0.46	9.04	42.61	0.33	2.29	3.85
2P5K	7.73	44.84	0.41	9.75	35.71	0.30	3.48	5.34
2P6J	4.73	53.37	0.44	6.10	46.63	0.34	2.47	3.26
2P81	4.23	63.64	0.47	5.17	55.11	0.39	2.02	2.76
2PMR	6.66	49.01	0.46	8.10	42.11	0.37	3.02	4.37
3DF8	9.78	27.62	0.28	14.83	25.48	0.26	4.81	7.25
3NRW	8.19	34.07	0.35	15.98	26.72	0.28	4.25	10.20
3P7K	2.42	80.00	0.66	4.17	66.11	0.44	0.86	1.22
3V1A	2.91	75.52	0.57	4.26	61.46	0.48	1.39	1.73
Avg.	5.98	50.14	0.38	8.84	41.67	0.31	2.84	4.48

Table 1.21: The experimental results of $O_2 + O_4$.

PDB ID	bRMSD	bGDT	bTM	RMSD	GDT	TM	bDME	DME
1AB1	4.53	53.80	0.38	6.99	44.57	0.31	2.74	3.61
1AIL	5.09	49.29	0.41	10.30	42.86	0.35	2.41	4.25
1BDD	3.78	58.75	0.49	8.58	45.42	0.40	2.08	2.90
1CRN	5.81	54.35	0.40	8.36	47.83	0.35	2.15	3.48
1DFN	5.37	49.17	0.22	7.28	41.67	0.18	2.52	3.73
1E0G	6.97	47.92	0.36	9.77	38.54	0.26	3.25	6.07
1E0M	6.03	44.59	0.26	6.29	42.57	0.22	2.67	3.34
1ENH	5.78	53.24	0.44	8.88	44.44	0.35	2.60	4.71
1F7M	7.68	37.50	0.26	11.75	29.35	0.19	3.29	5.45
1G26	5.19	51.61	0.26	7.50	39.52	0.21	2.23	3.70
1I6C	6.71	41.67	0.24	8.48	33.97	0.20	3.28	4.54
1IGD	7.82	43.44	0.39	12.98	35.66	0.30	3.41	6.79
1K36	8.20	36.41	0.24	12.34	29.35	0.18	3.34	5.76
1MSI	9.48	32.58	0.26	13.27	25.38	0.21	3.70	5.33
1Q2K	5.50	52.42	0.34	7.05	45.16	0.30	2.70	3.77
1ROP	3.23	67.41	0.55	4.17	63.39	0.46	1.22	1.77
1SXD	8.86	36.81	0.37	12.72	29.95	0.28	4.38	7.01
1UTG	5.71	48.21	0.46	8.22	44.64	0.39	2.69	3.64
1ZDD	1.93	88.24	0.69	2.85	74.26	0.53	0.84	1.21
2F4K	4.32	58.33	0.37	5.65	51.52	0.30	2.00	2.43
2GB1	8.57	45.09	0.39	12.06	37.95	0.32	4.25	7.01
2JUC	5.38	49.09	0.37	10.58	40.00	0.31	3.09	4.72
2JZQ	6.24	42.11	0.31	8.10	38.16	0.25	3.01	4.36
2KDL	9.42	43.30	0.32	11.52	38.39	0.30	3.80	5.65
2M7T	5.08	46.21	0.24	5.72	46.21	0.19	2.41	3.40
2MR9	3.31	63.07	0.44	5.56	52.84	0.36	1.78	2.69
2P5K	6.55	46.83	0.39	9.94	39.29	0.29	3.16	5.56
2P6J	4.32	59.62	0.46	8.91	50.00	0.38	2.34	5.62
2P81	3.16	71.02	0.55	5.40	57.95	0.43	1.80	2.63
2PMR	7.44	53.62	0.52	12.90	39.47	0.34	2.94	7.04
3DF8	10.43	29.05	0.28	12.88	25.48	0.25	4.35	6.42
3NRW	9.87	35.78	0.35	13.95	30.39	0.32	4.99	8.66
3P7K	1.52	86.11	0.74	3.89	65.00	0.45	0.48	0.87
3V1A	2.62	74.48	0.62	4.02	65.62	0.51	1.10	1.78
Avg.	5.94	51.50	0.39	8.79	43.44	0.31	2.74	4.41

Table 1.22: The experimental results of O_3 + O_4.

PDB ID	bRMSD	bGDT	bTM	RMSD	GDT	TM	bDME	DME
1AB1	5.56	49.46	0.37	7.66	43.48	0.35	2.32	2.82
1AIL	6.47	47.50	0.36	9.21	35.00	0.29	3.36	4.09
1BDD	4.49	63.75	0.55	6.58	51.25	0.44	2.06	2.70
1CRN	5.37	52.72	0.39	7.87	42.93	0.32	2.58	3.43
1DFN	4.40	56.67	0.28	5.44	49.17	0.25	1.99	2.41
1E0G	5.39	50.52	0.37	5.39	46.88	0.30	2.51	2.81
1E0M	5.09	49.32	0.29	7.03	42.57	0.23	2.27	3.08
1ENH	3.61	62.96	0.50	6.85	45.37	0.34	2.03	2.92
1F7M	6.03	41.85	0.28	8.20	33.15	0.20	3.01	4.03
1G26	5.00	48.39	0.27	7.01	42.74	0.20	2.60	3.45
1I6C	6.49	44.23	0.27	7.33	37.18	0.21	3.09	3.90
1IGD	6.24	47.13	0.39	6.66	47.13	0.37	3.22	3.70
1K36	7.05	40.76	0.28	9.00	33.70	0.23	3.39	4.07
1MSI	7.03	32.95	0.28	9.92	27.27	0.20	3.32	4.10
1Q2K	4.68	54.84	0.33	5.77	48.39	0.28	2.18	2.72
1ROP	3.72	60.71	0.51	4.73	52.68	0.41	1.61	2.57
1SXD	8.29	37.64	0.39	10.30	30.77	0.28	3.68	4.40
1UTG	5.33	54.29	0.48	8.61	40.00	0.32	2.22	4.05
1ZDD	2.33	76.47	0.51	3.79	65.44	0.48	1.18	1.83
2F4K	4.05	60.61	0.37	4.74	56.06	0.31	2.00	2.40
2GB1	6.03	49.55	0.39	8.60	42.41	0.35	3.03	3.30
2JUC	5.91	54.55	0.47	8.05	40.45	0.33	2.08	2.64
2JZQ	6.06	44.74	0.34	8.33	40.79	0.30	2.65	3.58
2KDL	4.55	54.91	0.47	5.18	51.34	0.40	2.10	2.78
2M7T	4.76	49.24	0.27	6.46	39.39	0.19	2.40	2.75
2MR9	4.05	61.36	0.46	6.75	50.57	0.36	2.09	2.53
2P5K	6.23	47.22	0.41	7.93	38.10	0.30	3.06	4.25
2P6J	5.50	51.92	0.38	6.09	44.23	0.32	2.45	2.86
2P81	4.90	63.07	0.46	8.26	46.02	0.35	2.46	3.48
2PMR	4.73	55.26	0.52	4.73	55.26	0.45	2.34	2.59
3DF8	9.80	27.86	0.31	11.16	24.05	0.24	4.63	5.39
3NRW	7.32	39.71	0.42	9.44	37.50	0.36	3.56	4.49
3P7K	1.81	83.89	0.71	4.11	58.89	0.39	0.51	2.56
3V1A	2.22	74.48	0.63	3.50	64.06	0.50	1.20	1.83
Avg.	5.31	52.66	0.40	7.08	44.24	0.32	2.51	3.25

Table 1.23: The experimental results of $O_1 + O_2 + O_3$.

PDB ID	bRMSD	bGDT	bTM	RMSD	GDT	TM	bDME	DME
1AB1	6.20	48.37	0.35	7.64	41.30	0.30	2.80	3.77
1AIL	5.93	46.07	0.40	7.75	41.43	0.35	2.70	3.57
1BDD	7.42	43.75	0.36	9.93	38.33	0.33	2.89	5.01
1CRN	5.92	51.09	0.38	7.14	45.65	0.35	2.91	3.94
1DFN	6.11	46.67	0.24	13.09	39.17	0.21	2.96	8.94
1E0G	5.95	49.48	0.37	9.25	40.62	0.28	2.61	4.21
1E0M	5.52	50.68	0.28	7.84	42.57	0.24	2.59	3.81
1ENH	6.18	52.31	0.41	9.03	43.52	0.33	2.93	4.08
1F7M	6.93	36.96	0.24	8.88	28.26	0.20	3.79	5.30
1G26	5.51	49.19	0.29	7.08	41.13	0.24	2.19	3.45
1I6C	5.94	43.59	0.26	7.48	33.97	0.20	2.79	3.52
1IGD	8.64	42.62	0.39	10.73	40.57	0.36	3.97	5.66
1K36	7.78	35.87	0.26	9.68	30.98	0.20	3.59	4.53
1MSI	9.44	29.92	0.25	12.63	25.38	0.20	3.95	5.74
1Q2K	5.27	56.45	0.37	7.25	46.77	0.29	2.37	3.55
1ROP	3.44	66.52	0.52	5.16	54.91	0.42	1.29	2.21
1SXD	9.53	32.97	0.32	13.70	27.75	0.25	4.25	6.50
1UTG	5.72	48.93	0.45	9.30	38.57	0.32	2.87	4.93
1ZDD	2.93	72.06	0.48	5.37	58.09	0.42	1.35	2.80
2F4K	4.21	60.61	0.36	5.23	50.76	0.27	2.12	2.68
2GB1	8.49	41.52	0.36	13.02	37.95	0.31	3.80	6.83
2JUC	3.77	54.55	0.42	7.05	41.82	0.29	2.15	3.76
2JZQ	7.66	38.16	0.30	9.95	35.09	0.29	3.59	6.45
2KDL	7.15	44.20	0.35	12.30	40.18	0.30	3.24	5.26
2M7T	5.49	46.21	0.24	6.00	40.91	0.18	2.62	2.62
2MR9	4.75	58.52	0.40	7.20	44.89	0.30	2.33	3.62
2P5K	7.11	46.03	0.37	9.51	37.70	0.30	2.85	4.97
2P6J	4.12	57.21	0.46	5.64	50.00	0.37	2.15	3.02
2P81	4.44	57.39	0.40	7.43	48.30	0.35	2.05	2.64
2PMR	6.52	50.33	0.48	9.22	42.76	0.37	2.27	5.84
3DF8	11.31	27.14	0.27	14.83	23.33	0.21	5.04	7.16
3NRW	10.36	36.52	0.36	14.79	31.62	0.30	4.00	7.00
3P7K	1.59	86.11	0.72	3.16	78.33	0.59	0.48	0.68
3V1A	2.75	78.12	0.68	3.69	70.31	0.54	1.17	1.39
Avg.	6.18	49.59	0.38	8.76	42.14	0.31	2.78	4.39

Table 1.24: The experimental results of $O_1 + O_2 + O_4$.

PDB ID	bRMSD	bGDT	bTM	RMSD	GDT	TM	bDME	DME
1AB1	6.45	47.83	0.37	8.53	41.85	0.29	2.82	3.68
1AIL	4.89	53.21	0.47	10.84	43.57	0.40	2.45	4.37
1BDD	5.62	54.58	0.43	7.97	40.83	0.32	2.45	4.08
1CRN	6.68	48.37	0.35	8.70	40.76	0.31	2.95	4.07
1DFN	5.99	48.33	0.24	8.02	40.83	0.21	3.33	4.12
1E0G	6.64	43.23	0.29	10.95	33.33	0.21	3.74	7.12
1E0M	5.78	47.97	0.28	8.36	35.81	0.19	2.90	4.27
1ENH	7.17	53.24	0.45	12.83	45.37	0.36	3.18	6.17
1F7M	8.82	34.78	0.23	13.08	27.17	0.21	4.11	6.62
1G26	6.54	48.39	0.23	10.49	35.48	0.19	3.30	5.47
1I6C	7.61	36.54	0.22	9.73	32.69	0.20	3.90	4.41
1IGD	9.62	43.03	0.37	14.51	34.84	0.30	5.43	7.19
1K36	8.61	38.59	0.30	10.85	32.07	0.23	3.84	4.83
1MSI	9.77	33.33	0.29	15.74	23.48	0.19	4.55	7.20
1Q2K	6.29	53.23	0.32	9.29	40.32	0.28	2.87	5.17
1ROP	3.07	70.09	0.59	4.51	57.14	0.44	1.06	1.83
1SXD	10.07	32.69	0.34	12.96	26.92	0.24	5.07	7.03
1UTG	7.99	45.00	0.42	10.78	38.93	0.34	2.67	4.68
1ZDD	2.01	86.76	0.66	5.02	58.82	0.43	1.13	2.69
2F4K	4.03	62.88	0.37	4.03	57.58	0.29	2.05	2.19
2GB1	8.21	51.34	0.42	13.58	40.18	0.34	4.04	8.43
2JUC	6.37	46.36	0.34	11.27	41.36	0.28	3.07	6.48
2JZQ	7.51	40.35	0.29	10.50	31.14	0.23	3.46	5.54
2KDL	8.87	46.43	0.38	15.02	39.29	0.30	4.48	7.57
2M7T	6.05	44.70	0.24	8.19	37.88	0.19	3.05	4.59
2MR9	3.94	61.36	0.41	10.80	40.91	0.29	2.44	5.29
2P5K	8.50	43.65	0.36	11.54	35.32	0.26	3.64	7.72
2P6J	4.87	59.62	0.49	8.33	45.67	0.36	2.66	4.91
2P81	4.91	55.68	0.39	7.26	45.45	0.35	2.43	3.52
2PMR	7.66	46.38	0.45	10.80	37.50	0.35	3.64	6.86
3DF8	11.18	27.62	0.29	15.72	21.90	0.21	4.78	7.70
3NRW	11.49	37.25	0.37	18.84	26.72	0.26	5.89	10.66
3P7K	1.62	90.00	0.79	2.78	73.33	0.54	0.53	0.77
3V1A	2.61	78.12	0.66	4.67	60.42	0.48	1.39	2.24
Avg.	6.69	50.32	0.39	10.19	40.14	0.30	3.21	5.28

Table 1.25: The experimental results of $O_1 + O_3 + O_4$.

PDB ID	bRMSD	bGDT	bTM	RMSD	GDT	TM	bDME	DME
1AB1	5.71	51.63	0.36	7.00	42.39	0.31	2.61	2.92
1AIL	5.88	46.79	0.42	8.43	36.43	0.33	2.58	3.98
1BDD	4.78	55.42	0.48	6.31	44.17	0.35	2.27	3.53
1CRN	5.04	55.43	0.44	7.12	44.02	0.35	2.68	3.50
1DFN	5.21	52.50	0.25	9.82	35.83	0.20	2.95	5.64
1E0G	5.66	49.48	0.34	6.97	42.19	0.26	2.83	3.08
1E0M	5.07	50.68	0.32	6.99	43.24	0.23	2.30	2.78
1ENH	5.03	53.24	0.46	7.65	42.13	0.33	2.73	3.32
1F7M	6.28	40.76	0.27	8.98	31.52	0.21	2.89	4.75
1G26	5.20	50.81	0.26	7.59	37.10	0.19	2.62	3.35
1I6C	5.16	47.44	0.32	7.98	40.38	0.25	2.49	3.61
1IGD	5.93	47.13	0.38	8.89	42.21	0.36	3.03	3.64
1K36	6.42	41.85	0.28	9.09	32.61	0.23	3.18	4.48
1MSI	7.87	35.61	0.30	9.59	29.55	0.25	3.55	4.19
1Q2K	5.05	53.23	0.32	6.54	47.58	0.30	2.68	3.47
1ROP	3.39	67.41	0.54	4.09	57.59	0.47	1.38	1.65
1SXD	9.51	31.04	0.31	11.24	28.85	0.25	4.22	4.93
1UTG	7.04	41.79	0.37	8.00	33.93	0.29	2.86	3.28
1ZDD	1.81	83.82	0.61	4.09	66.18	0.45	0.90	1.85
2F4K	3.79	66.67	0.38	5.16	53.03	0.31	1.91	2.44
2GB1	7.45	42.41	0.34	9.89	36.61	0.28	3.80	5.15
2JUC	5.51	47.27	0.39	7.89	38.18	0.31	2.64	3.13
2JZQ	6.31	45.18	0.36	8.88	39.04	0.31	3.18	3.96
2KDL	6.22	45.09	0.38	7.81	41.96	0.33	2.74	3.38
2M7T	5.53	45.45	0.23	6.38	39.39	0.18	2.71	2.99
2MR9	3.96	65.34	0.45	6.38	48.86	0.33	1.95	2.90
2P5K	6.52	45.24	0.37	7.74	45.24	0.37	2.90	3.13
2P6J	4.62	55.29	0.39	6.38	50.48	0.37	2.45	3.08
2P81	4.24	60.80	0.44	5.74	56.25	0.34	2.48	2.94
2PMR	3.66	61.51	0.58	7.07	43.09	0.41	2.00	3.34
3DF8	10.74	25.24	0.28	12.16	22.86	0.24	4.66	4.99
3NRW	10.53	33.58	0.34	12.76	28.19	0.29	4.26	5.55
3P7K	1.70	87.22	0.76	3.84	61.11	0.44	0.54	1.93
3V1A	2.61	69.79	0.57	3.65	57.29	0.39	1.46	2.47
Avg.	5.57	51.53	0.39	7.59	42.34	0.31	2.66	3.51

Table 1.26: The experimental results of $O_2 + O_3 + O_4$.

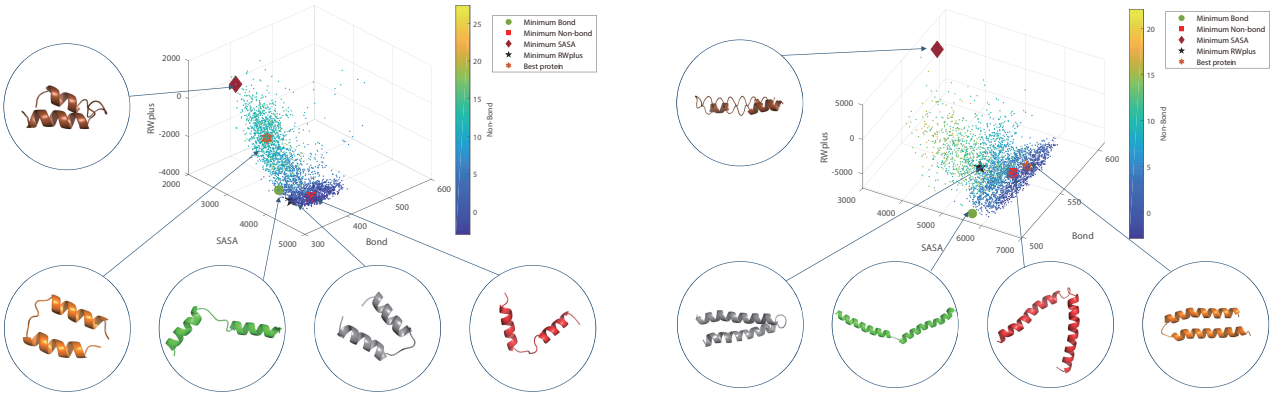
PDB ID	bRMSD	bGDT	bTM	RMSD	GDT	TM	bDME	DME
1AB1	5.52	50.54	0.37	8.96	40.22	0.34	2.88	3.94
1AIL	5.04	50.36	0.40	10.02	37.50	0.29	2.88	4.86
1BDD	4.96	55.83	0.48	5.61	50.00	0.40	2.09	2.51
1CRN	5.25	51.63	0.38	7.94	44.02	0.34	2.64	3.65
1DFN	4.75	51.67	0.25	6.77	44.17	0.20	2.08	3.52
1E0G	5.44	48.44	0.35	7.61	39.58	0.27	2.50	4.15
1E0M	5.22	49.32	0.31	6.38	41.22	0.22	2.20	2.75
1ENH	5.30	57.41	0.49	5.92	50.00	0.41	2.66	3.06
1F7M	7.22	37.50	0.26	9.44	30.43	0.20	3.43	4.48
1G26	5.24	51.61	0.26	6.26	45.16	0.21	2.59	2.98
1I6C	5.82	47.44	0.28	8.87	35.26	0.23	2.87	4.07
1IGD	9.11	45.90	0.42	10.85	40.98	0.36	4.01	5.46
1K36	7.04	38.04	0.23	11.18	29.35	0.18	3.54	5.12
1MSI	8.34	31.82	0.27	11.10	26.89	0.21	3.73	4.80
1Q2K	4.20	58.87	0.35	6.94	52.42	0.29	1.97	3.37
1ROP	3.50	67.41	0.55	4.82	60.27	0.45	1.21	2.44
1SXD	9.31	32.14	0.32	12.57	27.47	0.25	4.64	6.26
1UTG	5.31	48.21	0.41	7.40	41.07	0.34	2.66	4.37
1ZDD	2.19	77.94	0.52	3.76	66.18	0.46	0.97	1.82
2F4K	4.62	56.06	0.34	5.50	50.76	0.31	2.11	2.80
2GB1	7.50	50.45	0.41	10.48	42.86	0.34	2.87	6.50
2JUC	4.52	51.82	0.39	6.59	42.27	0.28	2.31	3.17
2JZQ	6.87	40.35	0.32	11.72	33.33	0.24	3.31	6.65
2KDL	7.93	45.54	0.35	10.30	39.29	0.31	2.90	5.39
2M7T	5.54	46.97	0.24	6.46	43.18	0.24	2.65	3.38
2MR9	3.62	67.05	0.49	7.58	50.57	0.35	1.75	2.21
2P5K	7.02	47.22	0.40	8.21	42.46	0.36	3.44	4.62
2P6J	4.89	57.69	0.46	8.62	44.71	0.35	2.72	3.90
2P81	3.97	59.09	0.41	6.29	51.14	0.36	1.93	2.62
2PMR	6.61	49.34	0.46	11.65	43.75	0.36	2.86	6.88
3DF8	11.89	27.62	0.28	14.88	23.10	0.21	5.08	8.57
3NRW	10.02	35.78	0.34	17.60	29.66	0.28	4.43	9.40
3P7K	1.84	81.67	0.66	3.75	72.78	0.55	0.50	1.10
3V1A	2.24	80.21	0.70	3.91	69.27	0.52	1.05	1.38
Avg.	5.82	51.44	0.39	8.41	43.57	0.32	2.69	4.18

objectives, the horizontal axis is extended to eight. Fig. 1.7 shows the parallel coordinate plot of a representative protein 2P5K. Similar results can be obtained for other proteins. From the figure, it is clearly observed that the lines of any adjacent objectives are cross. The conflicting relationship among the four objectives is thus clearly revealed via Fig. 1.7. From it, it is can observe that any two objectives among four are conflicting.

It is worth noting that the real Pareto front of conformations is unknown since the used energy functions are inaccurate. Fig. 1.8 shows the computed Pareto fronts of proteins 1ZDD, 1ROP, 2M7T, and 2P5K. In it, the x, y, and z-axes, respectively, represent the bond, SASA, and RWplus energy values, and the color of points shown in a logarithmic scale indicates the non-bond energy value. The figure shows a wide distribution of conformations, which implies that MO4 generates fruitful conformations. Additionally, four structure with the minimal value on each objective and the most accurate structure predicted by MO4 are illustrated in circles, to exhibit the convergence tendency of different objectives. These structures clearly show that different objectives favor different structures. SASA favors a tight structure, whereas the bond objective favors a flat one. The non-bond objective yields a stable one, but it may provide a malformed structure, as shown for 2P5K (the structure with the minimal non-bond objective value emerges a malformed structure in comparison with the native one). Similarly, RWplus may also generate a malformed structure, as shown for 1ROP (the structure with the minimal RWplus objective value emerges the interfering structure). These results repeatedly suggest that a single objective may cause the generation of unreasonable structure. Thus, the conflicting relationship among objectives alleviates the dilemma of generating a malformed structure. In contrast, many-objective methods can construct more fruitful conformations than single-objective ones because of the constraints arisen from conflicting objectives in different aspects. It is notable that, although MO4 also generates some odd structures, these structures located at the edge of the Pareto front might be eliminated by the clustering method when the final conformation are selected.

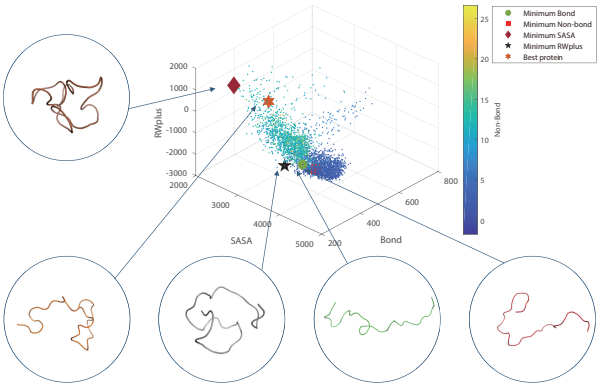
1.4.6 Comparison with Other Methods

A comparison of MO4 with six state-of-the-art methods, including I-PASE[24], APL-GA[81], MO3[26], AIMOEA[27], MOPSO[28], and MODE[82] is preformed to further demonstrate that the way of PSP problems regarded as MaOPSP problems is better than the one regarded as a multi-objective

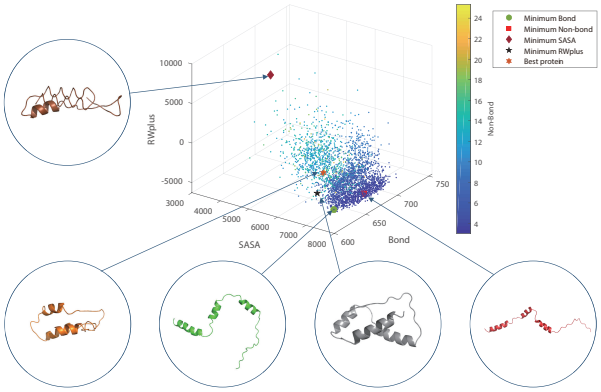


(a) The Pareto Front of four-objective of 1ZDD.

(b) The Pareto Front of four-objective of 1ROP.



(c) The Pareto Front of four-objective of 2M7T.



(d) The Pareto Front of four-objective of 2P5K.

Figure 1.8: The Pareto Front of conformations 1ZDD, 1ROP, 2M7T, and 2P5K.

PSP problem. These methods are:

1) I-PASE, for the first time, models PSP as a multi-objective problem by including bond and non-bond energies to predict protein structures.

2) APL-GA uses the probability distribution of angles to reduce the conformational space in a PSP problem to improve knowledge-based prediction accuracy.

3) MO3, for the first time, incorporates SASA as an objective to compensate bond and non-bond energies in a PSP problem. It yields a better conformation than other single-objective optimization methods.

4) AIMOEA reuses archived information by saving a Pareto set of conformations to improve the search performance. It replaces the local conformation structure to search for a more suitable structure.

Table 1.27: The experimental results of bRMSD of MO4 and other methods.

	MO4	MOPSO	AIMOES	MO3	APL-GA	I-PAES
p-value	–	3.86E-02	5.72E-03	8.53E-05	1.22E-04	6.25E-02

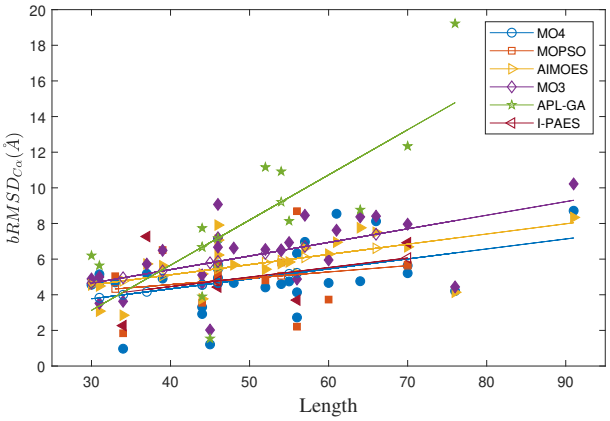


Figure 1.9: The bRMSD regression lines of MO4 and other methods.

5) MOPSO uses multi-objective particle swarm optimizer with crowding distance to solve a PSP problem. It consists of bond, non-bond, and dDFIRE [83] energy functions.

6) MODE decomposes the RWplus energy into distance and orientation-dependent terms to solve a PSP problem using a differential evolution algorithm.

Table 1.27 shows the Wilcoxon test results of MO4 and its peers. It clearly indicates that MO4 performs significantly better than MOPSO, AIMOEA, MO3, and APL-GA in terms of bRMSD. The detailed experimental results are summarized in Table 1.28. Those results show that MO4 yields the best results on 20 out of 34 proteins. In addition, Fig. 1.10 qualitatively illustrates the difference between MO4 and other methods in terms of bRMSD. The solid symbols represent the predicted structures, and the line represents the linear regression between the length and RMSD for each method. The difficulty of predicting protein structure increases with protein length. The slope of the linear regression reflects the ability of each algorithm to predict protein structures. A smaller slope signifies a stronger ability to predict the structure of long proteins. The intercept of the linear regression denotes the prediction accuracy of an algorithm. A lower intercept indicates higher prediction accuracy. This figure clearly implies that MO4 can predict the structure of long proteins better than the other methods and has higher prediction accuracy.

Similarly, Table 1.29 lists the p-values of Wilcoxon test on the RMSD of MO4 and other state-of-

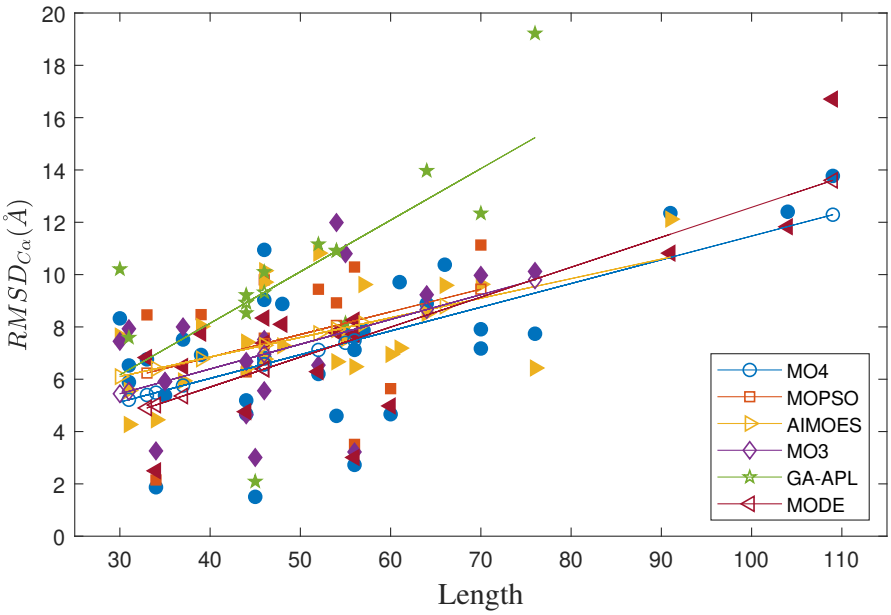


Figure 1.10: The RMSD regression lines of MO4 and other methods.

the-art methods. The detailed RMSD values are exhibited in Table 1.30. RMSD reflects the average capability of algorithms to predict protein structures. The Wilcoxon test results indicate that MO4 is significantly better than its competitors in term of RMSD, suggesting that the average capability of MO4 is better than that of other methods for most proteins. In addition, the linear regression is illustrated in Fig. 1.10, showing a smaller slope and a smaller intercept for MO4 compared to its competitors. This indicates the superior performance of MO4 for predicting longer proteins which is never seen in prior work.

Furthermore, we compare MO4 with three state-of-the-art MaOEAs (i.e., SDR [84], IT [85], and IGD [86]) in MaOPSP to demonstrate its superior performance.

1) SDR uses a strengthened dominance relation to balance convergence and diversity for many-objective problems.

2) IT embeds two independent stages to address the convergence and diversity for many-objective problems.

3) IGD is a widely used performance indicator for many-objective problems. It selects the solutions with favorable convergence and diversity.

Table 1.31 shows Wilcoxon and Friedman test results of comparison between MO4 and these

many-objective methods. It is clearly observed that MO4 is significantly better than its three peers on all eight performance metrics. The details of experimental results are shown in Tables 1.32 -1.34.

1.5 Conclusion

In the study, an MaOPSP approach is proposed for the first time to solve a PSP problem. It includes four objectives: the bond and non-bond energies, SASA, and RWplus. A diversity-based and convergence-based MaOEA is introduced for searching the large protein conformational space. Besides, an effective protein representation, torsion angles with secondary structure constraint, is restricted to reduce the large conformational space. The PSIPRED method yields highly accurate secondary structure. The generated fruitful conformations are saved in an archive, and a decision-making method selects the final conformation from the archive. The proposed MO4 is tested on thirty-four proteins and is compared with six state-of-the-art methods. Experimental results indicate that this method can generate more fruitful protein structures than its peers with single/multiple objectives. The study, therefore, supports the use of four-objective predictions. Additionally, in comparison with other state-of-the-art MaOEAs, MO4 can generate better solutions for PSP than its peers, which demonstrates its superior performance.

The results of comparative experiments among different number objectives tell that the combination of four objective functions is better than other combinations, and further prove the superiority of MO4. Conflict among the four objectives is also analyzed to verify the effectiveness of the proposed many-objective approach. The Pareto front of the four objectives clearly provides the distribution and structural difference of conformations. This suggests that the method involving many-objective functions, rather than a single/multi-objective, provides more reasonable structures than existing methods. Our energy analysis result indicates that current energy functions cannot really measure the conformations because a lower energy conformation may represent an odd structure. The odd structures are usually located at the edge of Pareto front. These structures can be eliminated via a decision-making method. MO4 can reduce the risk of generating malformation structures since it can restrict the conformational space in different aspects by conflicting energy functions.

As future work, we intend to test the method a wider range of proteins to verify its performance. The optimal mutation and selection strategies for the PSP problem should be researched to achieve

the better sampling of a large conformational space. More accurate energy functions should be added to further improve the prediction performance of MaOPSP. A suitable combination of energy functions is needed to generate reasonable and fruitful conformations. In addition, bioinformatics should be exploited to restrict a conformational space. Other methods, involving contact maps and fragment assemblies, can be used to improve prediction accuracy.

Table 1.28: The experimental results of bRMSD of MO4 and other methods.

PDB ID	MO4	MOPSO	AIMOES	MO3	APL-GA	I-PAES
1AB1	**5.44**	5.72	6.23	6.67	5.53	–
1AIl	**5.21**	–	6.69	7.96	12.34	–
1BDD	4.66	**3.73**	6.28	5.95	–	–
1CRN	4.95	5.16	–	–	–	**4.43**
1DFN	4.56	–	**4.68**	4.89	6.2	–
1E0G	**4.67**	–	5.68	6.62	–	–
1E0M	**5.18**	–	5.73	5.72	–	7.27
1ENH	**4.60**	5.81	5.75	6.52	10.92	–
1F7M	7.11	–	7.91	**7.2**	–	–
1G26	4.68	–	**4.49**	5.0	–	–
1I6C	**4.91**	6.51	5.63	6.49	–	–
1IGD	8.55	–	**6.95**	7.62	–	–
1K36	**5.70**	–	7.07	9.07	–	–
1MSI	8.13	–	**7.51**	8.41	–	–
1Q2K	5.15	–	**3.08**	3.52	5.64	–
1ROP	2.73	**2.21**	–	–	–	3.7
1SXD	**8.71**	–	8.34	10.22	–	–
1UTG	**5.64**	5.68	–	–	–	6.92
1ZDD	**0.97**	1.84	2.85	3.63	–	2.27
2GB1	**4.19**	–	5.19	4.88	–	–
2JUC	**4.75**	–	5.84	6.93	8.14	–
2JZQ	6.97	–	**6.62**	8.46	–	–
2KDL	**6.32**	8.69	–	–	–	–
2M7T	**4.70**	5.04	–	–	–	–
2MR9	**2.94**	–	5.17	5.11	7.74	–
2P5K	**4.76**	–	7.76	8.37	8.77	–
2P6J	**4.41**	4.81	5.43	6.54	11.16	–
2P81	**3.31**	3.55	3.77	4.64	3.9	–
2PMR	4.21	–	**4.14**	4.44	19.22	–
3P7K	**1.21**	–	–	2.02	1.54	–
3V1A	**2.15**	–	2.32	3.01	9.79	–
p-value	–	3.86E-02	5.72E-03	8.53E-05	1.22E-04	6.25E-02

Table 1.29: The Wilcoxon results of RMSD of MO4 and other methods.

	MO4	MOPSO	AIMOES	MO3	APL-GA	MODE
p-value	–	2.44E-04	2.77E-01	9.34E-03	1.22E-04	2.10E-01

Table 1.30: The experimental results of RMSD of MO4 and other methods.

PDB ID	MO4	MOPSO	AIMOES	MO3	APL-GA	MODE
1AB1	6.82	9.8	**6.77**	7.52	10.1	7.38
1AIl	**7.18**	–	9.63	9.97	12.34	–
1BDD	**4.66**	5.64	6.95	–	–	4.98
1CRN	6.79	7.57	–	**5.56**	–	–
1DFN	8.33	–	7.65	**7.45**	10.21	–
1E0G	8.88	–	**7.2**8	–	–	8.1
1E0M	7.52	–	**5.94**	8.0	–	6.49
1ENH	**4.60**	8.92	6.67	11.99	10.92	7.8
1F7M	10.95	–	9.71	–	–	–
1G26	5.89	–	**5.5**7	–	–	–
1I6C	**6.93**	8.47	8.02	–	–	7.76
1IGD	9.71	–	**7.19**	–	–	–
1K36	9.04	–	10.15	–	–	**8.34**
1MSI	10.38	–	**9.59**	–	–	–
1Q2K	6.53	–	**4.27**	7.93	7.59	–
1ROP	**2.73**	3.51	–	3.22	–	3.01
1SXD	12.35	–	12.12	–	–	**10.82**
1UTG	**7.91**	11.13	–	–	–	–
1ZDD	**1.87**	2.15	4.45	3.26	–	2.5
2F4K	**5.38**	–	–	5.91	–	–
2GB1	7.62	–	**6.48**	–	–	8.26
2JUC	7.70	–	**7.54**	10.8	8.14	–
2JZQ	**7.87**	–	9.62	–	–	–
2KDL	**7.12**	10.29	–	–	–	7.72
2M7T	**6.76**	8.46	–	–	–	6.83
2MR9	**5.19**	–	7.42	6.68	9.22	–
2P5K	8.89	–	**8.52**	9.23	13.97	–
2P6J	**6.20**	9.44	10.82	6.54	11.16	6.29
2p81	4.67	**6.28**	6.43	4.64	8.53	4.76
2PMR	7.74	–	**6.43**	10.12	19.22	–
3DF8	**13.77**	–	–	–	–	16.71
3NRW	12.40	–	–	–	–	**11.83**
3P7K	**1.51**	–	–	3.01	2.09	–
3V1A	3.72	–	3.96	**2.23**	9.79	–
p-value	–	2.44E-04	2.77E-01	9.34E-03	1.22E-04	2.10E-01

Table 1.31: The Wilcoxon and Friedman test results of three many-objective methods.

Parameters	IGD (Rank)	IT (Rank)	SDR (Rank)	MO4 (Rank)
bRMSD	1.83E-07 (2)	1.83E-07 (4)	1.83E-07 (3)	– (1)
bGDT	1.83E-07 (2)	1.83E-07 (3)	1.83E-07 (4)	– (1)
bTM	1.83E-07 (2)	1.83E-07 (3)	2.19E-07 (4)	– (1)
bDME	1.83E-07 (2)	1.83E-07 (4)	1.83E-07 (3)	– (1)
RMSD	4.84E-07 (2)	5.28E-07 (4)	1.24E-06 (3)	– (1)
GDT	2.21E-06 (2)	4.37E-07 (4)	1.96E-05 (3)	– (1)
TM	1.16E-05 (2)	5.75E-07 (4)	8.75E-05 (4)	– (1)
DME	4.68E-05 (2)	1.83E-07 (4)	3.28E-05 (3)	– (1)
Avg. Rank	2	3.75	3.375	1

Table 1.32: The experimental results of IGD.

PDB ID	bRMSD	bGDT	bTM	RMSD	GDT	TM	bDME	DME
1AB1	7.89	40.22	0.31	8.70	40.22	0.30	3.90	3.90
1AIL	7.79	37.86	0.30	10.60	37.86	0.28	3.78	4.04
1BDD	7.72	44.17	0.38	8.29	42.92	0.38	3.19	3.92
1CRN	7.34	42.39	0.33	7.34	41.85	0.33	2.54	2.54
1DFN	9.29	40.83	0.26	10.86	40.00	0.21	6.18	7.05
1E0G	11.14	36.98	0.25	11.14	36.98	0.25	4.77	4.77
1E0M	9.26	35.14	0.19	12.02	29.73	0.18	4.93	5.85
1ENH	7.11	44.44	0.36	9.54	44.44	0.32	3.34	3.34
1F7M	9.17	36.41	0.24	10.49	32.07	0.21	4.18	4.18
1G26	7.79	41.94	0.22	7.79	41.94	0.20	3.76	3.76
1I6C	7.23	35.90	0.22	9.96	33.33	0.22	3.66	5.39
1IGD	13.53	36.48	0.31	13.53	34.02	0.30	6.21	6.21
1K36	8.63	33.70	0.21	8.63	33.70	0.21	4.06	4.06
1MSI	11.66	26.14	0.21	11.66	26.14	0.21	5.00	5.00
1Q2K	5.42	50.00	0.29	5.42	50.00	0.29	2.63	2.63
1ROP	5.65	56.25	0.47	5.77	56.25	0.44	2.08	2.08
1SXD	11.43	28.02	0.26	11.73	26.65	0.25	4.73	4.73
1UTG	8.34	43.57	0.37	8.75	41.43	0.37	3.18	4.53
1ZDD	4.21	66.18	0.40	5.17	59.56	0.39	1.86	2.09
2F4K	4.58	53.79	0.30	5.94	47.73	0.30	2.56	3.18
2GB1	10.83	40.62	0.32	10.83	40.62	0.32	4.63	4.63
2JUC	9.05	40.45	0.32	9.05	40.45	0.29	3.63	3.63
2JZQ	8.73	33.77	0.26	9.21	33.77	0.26	3.84	4.27
2KDL	7.70	39.29	0.31	9.95	37.95	0.30	4.89	5.42
2M7T	6.50	40.91	0.20	8.18	37.88	0.18	3.08	4.39
2MR9	6.75	48.30	0.32	6.81	45.45	0.31	2.65	2.65
2P5K	9.15	42.46	0.34	9.19	35.32	0.30	5.02	5.02
2P6J	7.27	47.60	0.37	7.27	41.83	0.31	3.38	3.79
2P81	5.75	51.70	0.37	7.11	49.43	0.37	2.70	3.33
2PMR	9.98	40.79	0.37	11.39	40.13	0.35	4.80	4.80
3DF8	14.44	23.33	0.22	14.44	23.33	0.22	6.03	6.03
3NRW	13.59	28.68	0.27	15.24	26.47	0.26	6.31	6.31
3P7K	4.26	70.56	0.54	5.61	61.67	0.48	1.02	1.88
3V1A	3.46	68.75	0.50	5.00	59.38	0.40	1.91	2.54
Avg.	8.31	42.58	0.31	9.19	40.31	0.29	3.84	4.17

Table 1.33: The experimental results of IT.

PDB ID	bRMSD	bGDT	bTM	RMSD	GDT	TM	bDME	DME
1AB1	7.90	42.93	0.30	8.03	39.13	0.30	4.05	4.05
1AIL	12.15	40.36	0.36	12.81	37.50	0.30	7.80	8.02
1BDD	14.06	36.25	0.31	14.06	35.42	0.31	8.57	9.68
1CRN	9.32	40.76	0.30	9.95	40.22	0.30	6.01	6.08
1DFN	12.72	40.83	0.18	12.83	38.33	0.18	8.76	8.79
1E0G	12.40	38.02	0.25	12.92	36.98	0.25	7.52	7.52
1E0M	5.84	46.62	0.25	7.55	39.19	0.25	3.00	4.06
1ENH	10.58	40.28	0.33	10.58	40.28	0.33	5.55	5.55
1F7M	8.28	30.98	0.21	8.28	29.89	0.20	5.07	5.53
1G26	7.15	39.52	0.21	7.15	39.52	0.21	3.79	3.88
1I6C	7.15	37.18	0.23	8.79	34.62	0.19	3.06	4.42
1IGD	14.09	36.48	0.31	14.45	33.61	0.30	8.51	8.51
1K36	9.36	31.52	0.23	9.36	31.52	0.20	4.29	4.55
1MSI	22.87	22.35	0.18	24.58	20.45	0.16	17.51	19.44
1Q2K	8.17	49.19	0.31	8.23	45.16	0.31	4.18	4.46
1ROP	4.19	63.84	0.51	4.71	62.05	0.47	1.13	1.64
1SXD	25.90	26.92	0.25	26.70	25.55	0.22	20.43	20.65
1UTG	18.00	33.57	0.28	18.54	31.07	0.28	13.98	14.56
1ZDD	2.98	72.06	0.50	2.98	72.06	0.49	1.47	1.71
2F4K	5.15	53.79	0.30	5.15	53.79	0.28	2.69	2.86
2GB1	14.87	37.95	0.33	16.09	34.38	0.29	9.19	10.70
2JUC	10.55	37.27	0.29	13.12	35.00	0.26	6.78	9.11
2JZQ	10.30	34.21	0.27	12.11	32.89	0.25	6.11	6.11
2KDL	18.77	35.27	0.31	19.81	35.27	0.30	14.32	14.79
2M7T	6.49	43.18	0.21	6.75	43.18	0.21	3.33	3.33
2MR9	9.96	44.89	0.31	10.33	44.89	0.30	5.74	6.05
2P5K	19.09	37.70	0.32	20.16	32.54	0.28	13.32	14.27
2P6J	8.79	43.27	0.29	9.31	37.02	0.28	4.93	5.12
2P81	12.15	40.91	0.34	12.82	39.77	0.34	7.80	8.19
2PMR	20.39	39.47	0.37	20.71	39.47	0.37	14.08	14.57
3DF8	32.43	20.00	0.17	32.43	19.52	0.17	22.67	22.75
3NRW	35.16	31.13	0.30	36.59	23.77	0.23	28.66	30.11
3P7K	1.53	88.89	0.77	2.37	78.89	0.62	0.53	0.59
3V1A	3.96	66.67	0.49	4.15	64.58	0.49	1.85	1.85
Avg.	12.43	41.89	0.31	13.07	39.63	0.29	8.14	8.63

Table 1.34: The experimental results of SDR.

PDB ID	bRMSD	bGDT	bTM	RMSD	GDT	TM	bDME	DME
1AB1	18.31	26.09	0.19	19.64	26.09	0.18	13.23	14.35
1AIL	12.41	38.21	0.28	13.21	36.79	0.28	6.82	6.93
1BDD	21.04	20.83	0.15	21.52	20.42	0.15	14.95	15.17
1CRN	5.55	49.46	0.31	5.76	47.83	0.30	2.80	2.97
1DFN	5.78	52.50	0.29	6.68	50.83	0.29	2.55	2.82
1E0G	9.71	35.42	0.23	10.31	34.38	0.23	3.54	3.65
1E0M	9.32	37.16	0.22	9.93	35.81	0.22	4.43	4.50
1ENH	7.33	43.98	0.31	7.64	42.59	0.31	3.14	3.27
1F7M	10.79	30.98	0.20	10.79	30.98	0.20	5.26	5.32
1G26	7.81	39.52	0.20	7.99	37.10	0.20	3.87	4.11
1I6C	7.07	42.31	0.26	7.18	41.03	0.25	3.26	3.41
1IGD	10.51	36.89	0.30	11.69	36.48	0.30	5.09	5.29
1K36	9.20	28.80	0.19	9.34	26.09	0.16	4.30	4.42
1MSI	12.20	31.06	0.22	12.42	28.03	0.21	4.77	5.52
1Q2K	8.37	40.32	0.26	8.97	39.52	0.26	3.83	3.99
1ROP	3.59	62.05	0.43	3.73	59.82	0.42	1.89	2.15
1SXD	11.74	28.02	0.25	12.17	26.10	0.24	5.08	5.26
1UTG	10.78	38.21	0.31	10.87	37.14	0.30	4.31	4.41
1ZDD	2.13	86.03	0.66	2.36	79.41	0.56	0.80	0.94
2F4K	4.27	59.85	0.34	4.86	56.06	0.33	1.97	2.46
2GB1	14.25	35.71	0.29	14.49	35.71	0.29	7.40	8.16
2JUC	10.72	41.36	0.32	11.11	40.91	0.30	3.85	4.00
2JZQ	11.33	31.14	0.21	11.80	30.70	0.21	4.64	4.64
2KDL	12.27	41.96	0.35	12.58	39.73	0.35	6.56	6.65
2M7T	7.44	37.88	0.18	7.73	33.33	0.16	2.99	3.17
2MR9	7.55	45.45	0.32	7.87	42.61	0.31	2.40	2.53
2P5K	13.13	34.52	0.28	13.56	33.73	0.28	6.37	6.75
2P6J	8.81	26.44	0.17	9.31	24.52	0.16	5.50	5.50
2P81	7.58	46.59	0.37	8.80	45.45	0.36	2.59	3.31
2PMR	11.72	36.51	0.32	11.96	36.18	0.32	5.63	5.74
3DF8	15.20	22.86	0.22	20.06	21.67	0.21	5.65	10.25
3NRW	15.58	29.41	0.30	15.58	29.41	0.30	9.91	9.91
3P7K	2.59	74.44	0.60	3.38	66.11	0.47	1.08	1.93
3V1A	3.87	62.50	0.50	4.46	60.94	0.46	1.49	1.72
Avg.	9.70	41.01	0.30	10.29	39.22	0.28	4.76	5.15

Chapter 2

An Adaptive Replacement Strategy-incorporated Particle Swarm Optimizer for Wind Farm Layout Optimization

2.1 Introduction

Global climate change and global warming are the main concern problems for sustainable advancement, human survival, and economic development. Natural disasters annually increase due to climate change, resulting in human habitat damages and economic losses [87]. Human activities, such as energy generation based on fossil fuels and industrial production, increase CO_2 emissions, which is the main reason for climate change. Thus, green renewable energy has been developed to reduce the consumption of fossil fuels to alleviate climate change, including solar photovoltaic, hydroelectric, and wind power generation. The wind power generation is expected up to at the least 18% of global power by 2050, according to International Energy Agency [88]. Therefore, wind power generation has been rapidly developed, and wind energy has become one of the most potent green renewable energies. To improve outputs of wind power generation, researchers have investigated from different aspects such as site selection [89], wind turbine model [90], wake effect model [91], wind speed forecast [92], and wind farm layout optimization (WFLO) [93]. WFLO plays an important role in improving power generation because wind turbines are more stable, efficient, and controllable with manufacturing development.

The overall power generation of a wind farm is significantly less than its total rated power. The

primary factor is that the wake effect [94] reduces the input speed of a wind turbine in the downstream, resulting in a decrease in its efficiency. One main challenge for maximizing wind power generation outputs is optimizing the layout of wind turbines to reduce the wake effect. Jensens' wake effect model [95] is widely used to address WFLO problems. The model predicts the energy content of a wind farm to better estimate its power generation ability [96]. There are two construction patterns for WFLO problems: continuous and discrete models. The continuous model allows a wind turbine to locate anywhere on a wind farm. In the discrete model, a wind farm is divided into discrete grids, and wind turbines are positioned at the center of the grids. The continuous model has more capable of obtaining a promising solution than the discrete one. In contrast, the discrete model significantly simplifies WFLO computation [97]. Various researchers adopt a discrete model to handle WFLO problems [98, 99, 100]. However, WFLO problems are non-convex and NP-hard problems [100]. Traditional optimization methods can not provide a satisfactory solution. Meta-heuristic algorithms inspired by nature biology have obtained success in various real-world problems [101, 102, 103, 104], such as solar photovoltaic generation [105], protein structure prediction [106], and discrete optimization [107]. They include genetic algorithm (GA), particle swarm optimization (PSO), gravitational search algorithm (GSA), differential evolution (DE), and so on. Therefore, researchers adopt meta-heuristic algorithms as alternative methods to deal with WFLO problems, where GA is used in 75% of WFLO studies [97]. Mosetti *et al.* [108] for the first time used GA to optimize a WFLO problem with Jensen wake model. They divided a wind farm into 10×10 square grids to locate wind turbines and used three different wind scenarios (i.e., single direction and wind speed, multiple directions and single wind speed, and multiple directions and wind speeds) to verify the performance of the proposed model. Emami *et al.* [109] proposed a matrix binary coding and a novel objective function to solve WFLO problems. Gonzalez-Longatt *et al.* [110] treated the WFLO problem as a traveling salesman problem (TSP) and optimized it via a modified GA. Zeng *et al.* [111] proposed a reserved operator to improve the performance of GA for WFLO problems. Chen *et al.* [112] analyzed the effect of wind turbines with different hub heights for power generation in a small wind farm. They demonstrated that a wind farm using wind turbines with different hub heights obtained more power generation outputs than that using the same height wind turbines. Chen *et al.* [113] adopted a binary-real coding GA to solve WFLO problems. Chen *et al.* [114] modeled a multi-objective WFLO problem and optimized it via a multi-objective GA. Gao

et al. [115] analyzed the effect of different wake models for power generation. They constructed a WFLO problem by using a two-dimensional Jensen-Gaussian wake model and optimized it via a multi-population GA. The numerical experiment results demonstrated two-dimensional Jensen-Gaussian wake model was better than the original Jensen model. Parada *et al.* [116] constructed a WFLO problem with constraints by using the Jensen-Gaussian model and optimized it via a GA. Abdelsalam *et al.* [117] utilized a local search strategy to improve the performance of GA for WFLO problems. Yang *et al.* [118] proposed a GA-based boolean code to optimize WFLO problems. Ju *et al.* [99] used support vector regression to enhance the performance of GA on WFLO problems. Bai *et al.* [100] used Monte Carlo tree search to enhance the exploitation capability of GA for improving the power conversion efficiency.

Researchers have sufficiently developed various GAs to improve power conversion efficiency. They proposed several specific operators according to the characteristics of WFLO problems to enhance the performance of GAs. They also verified the effectiveness of GAs on different wake models with varying wind scenarios and constraints. However, GA suffers from the issues of insufficient optimization efficiency and have been shown to inferior to many other meta-heuristic algorithms, such as DE, PSO, GSA, etc., on many specific optimization problems [119, 120, 121]. This motivates researchers to use other algorithms instead of GA to solve WFLO problems. Beşkirli *et al.* [122] proposed binary artificial algae algorithm to optimize the grid-based WFLO problem. The numerical experiment results show that it gains better results than GA. Wang *et al.* [123] proposed DE with a new coding mechanism to optimize WFLO problems. Long *et al.* [124] proposed an adaptive DE with a data-driven surrogate model. The results showed its superior performance. Pillai *et al.* [125] implemented a comparison between GA and PSO, which verified that PSO gained better results than GA within a smaller time cost. The above studies show the potential of different evolutionary intelligence algorithms on WFLO problems. The outputs of power can further be improved by using more efficient algorithms. Therefore, in this study, WFLO problem is modeled by using a discrete model with Jensen wake effect model and twelve location constraints. Then, a genetic learning particle swarm optimization with an adaptive replacement strategy, named AGPSO, is proposed to optimize it. Meanwhile, a new index integer coding is used to reduce the dimension of problems so that it accelerates the optimization process. To verify its performance, AGPSO is compared with other eight state-of-the-art evolutionary algorithms under four wind scenarios.

Nomenclature

ALGSA	Aggregative learning gravitational search algorithm	GSA	Gravitational search algorithm
AGPSO	Adaptive replacement strategy-incorporated particle swarm optimizer	GWO	Grey wolf optimizer
AGA	Adaptive genetic algorithm	HGSA	Hierarchical gravitational search algorithm
BSA	Bird swarm algorithm	IWO	Invasive weed optimization
CJADE	Chaotic local search-based differential evolution	PSO	Particle swarm optimization
CLPSO	Comprehensive learning particle swarm optimizer	SHADE	Success-history based parameter adaptation differential evolution
DE	Differential evolution	SUGGA	Support vector regression guided genetic algorithm
GA	Genetic algorithm	TSP	Traveling salesman problem
GLPSO	Genetic learning particle swarm optimization	WFLO	Wind farm layout optimization

The contributions of this study are summarized as follows:

(1) For the first time, we propose a new adaptive replacement strategy in PSO to replace the worst turbine in the population adaptively.

(2) We propose a specific velocity update strategy according to the characteristic of WFLO problems. The numerical experiment results show that the strategy better addresses a discrete problem than the original one.

(3) A new integer coding-based turbine index is proposed to reduce the dimension of problems and executing time.

(4) An extensive comparison between AGPSO and eight meta-heuristic algorithms under four wind scenarios with twelve location constraints is implemented, and the numerical experiment results verify its superiority.

The reminder of this study is organized as follows: Section 2.2 describes the WFLO problem, Jensen wake model, and location constraints. Section 2.3 introduces the proposed AGPSO algorithm. Wind scenarios and numerical experiment results are shown in Section 2.4. Finally, Section 2.5 concludes this study.

2.2 Wind Farm Layout Optimization Problem

In this section, an objective function of WFLO problems is given, including the total construction cost and power generation outputs of a wind farm. A classic wake effect model is introduced in detail. The power curve model is described.

2.2.1 Optimization Objective

The objective function of WFLO problem is to find an optimal wind farm layout for wind turbines, which reduces the wake effect to maximize the power generation of a wind farm under a wind scenario with a minimum construction cost. Therefore, the objective of WFLO problems $\mathcal{B}$ is presented as follows:

$$\mathcal{B} = \min \frac{C(N)}{\mathcal{P}(N, \mathcal{L})} \tag{2.1}$$

where N is the total number of wind turbines in a wind farm. $\mathcal{L}$ is wind turbines layout of the wind farm. $C(N)$ is total cost of the wind farm. $\mathcal{P}(N, \mathcal{L})$ is total power generation outputs of the wind farm.

Mosetti [108] assumes that the cost of a single turbine is 1, then maximum reduction 1/3 could be achieved when n is large enough. The total cost of a wind farm is calculated as follows:

$$C(N) = N\left(\frac{2}{3} + \frac{1}{3}e^{-0.00174N^2}\right) \tag{2.2}$$

When the wind turbine number is constant, a wind farm's total cost is also fixed. Therefore, in this study, the objective of WFLO problems becomes maximizing the power generation outputs, i.e.,

$$\mathcal{B} = \max \mathcal{P}(N, \mathcal{L}) = \sum_{i=1}^{N} \sum_{v,\theta} p(v, \theta) P_i(v, \theta, \mathcal{L}) \tag{2.3}$$

where v and θ are the wind speed and direction, respectively. $p(v, \theta)$ is the probability distribution of wind speed v and direction θ, $P_i(v, \theta, \mathcal{L})$ is the power generation outputs of wind turbine i in layout $\mathcal{L}$ under wind speed v and wind direction θ. $\mathcal{L}$ is to decision variables of the problem. A conversion efficiency η is given for quantificationally showing the power generation performance of a wind

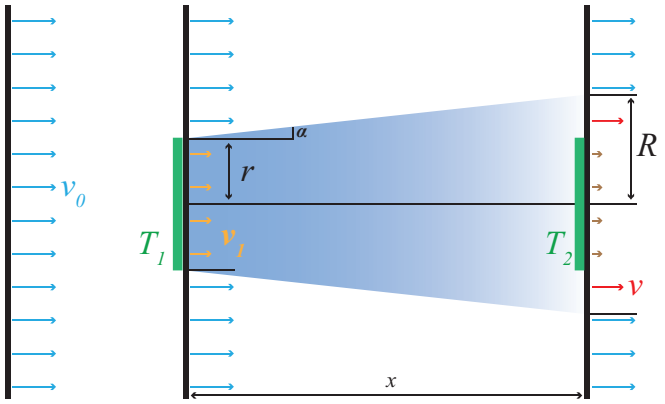


Figure 2.1: The illustration of wake effect model.

farm, i.e.,

$$\eta = \frac{\mathcal{P}(N, \mathcal{L})}{N \sum_{v,\theta} p(v, \theta) P(v, \theta)} \tag{2.4}$$

where $\mathcal{P}(N, \mathcal{L})$ is total outputs of power generation for layout $\mathcal{L}$, $P(v, \theta)$ is the rated power of a wind turbine under wind speed v and direction θ without the wake effect.

2.2.2 Wake Effect Model

Assume a wind turbine T_2 is located in the downwind of a wind turbine T_1. Wake effect affects the power generation of T_2, which refers to, when wind passes T_1, the wind speed of T_1 in downwind is discounted because of its energy is absorbed by T_1. An illustration of the wake effect model is given in Fig. 2.1.

In this study, a classic linear wake model proposed by Jensen [95] is used to quantify the wake effect. Under an assumption that momentum is conserved [126], wake wind speed v is described as follows:

$$\pi r^2 v_1 + \pi(R^2 - r^2) v_0 = \pi R^2 v \tag{2.5}$$

where v_0 is a free wind speed, v_1 is the wind speed after passing a wind turbine T_i. In general, v_1 is approximately one-third of v_0 [95]. r and R are the rotor radius of a wind turbine and the wake radius, respectively.

Solving to Eq. (2.5), wake velocity v is described as follows:

$$v = v_0 \left[1 - \tfrac{2}{3}(\tfrac{r}{R})^2 \right] \tag{2.6}$$

Due to that Jensen model assumes the momentum is conserved, then, the wake radius is proportional to a distance d_y between wind turbines T_1 and T_2, i.e.,

$$R = r + \alpha d_y \tag{2.7}$$

where α is the entrainment constant, given by [118],

$$\alpha = \frac{0.5}{\ln(\frac{h}{z_0})} \tag{2.8}$$

where h is the height of a wind turbine, z_0 is the surface roughness of the wind farm ground. When the ground is sand, z_0 is between 0.2mm and 0.3mm [127].

In real-application, a wind turbine is actually affected by multiple wind turbines [128]. Then actual wake velocity is calculated by:

$$v = v_0 \left[1 - \sqrt{\textstyle\sum_i^N (1 - \frac{v_i}{v_0})^2 \sqrt{\frac{A_i}{A_r}}} \right] \tag{2.9}$$

where v_i is the wake velocity under effect of the ith wind turbine, A_i is an overlap area between the ith upwind wind turbine wake and the downwind wind turbine, A_r is the surface area of the downwind turbine, N is the number of wind turbines. According to the law of cosines, A_i it is calculated by:

$$A_i = R^2 \left(\beta - \frac{\sin(2\beta)}{2} \right) + r^2 \left(\gamma - \frac{\sin(2\gamma)}{2} \right) \tag{2.10}$$

where β and γ are described as follows:

$$\begin{aligned} \beta &= \frac{R^2 + d_x^2 - r^2}{2Rd_x} \\ \gamma &= \frac{r^2 + d_x^2 - R^2}{2rd_x} \end{aligned} \tag{2.11}$$

where d_x is the distance between wake center and rotor center.

2.2.3 Power Curve Model

Power curve model [129] is used to measure the relation between power generation outputs $\mathcal{P}(v)$ and wind speed v, i.e.,

$$\mathcal{P}(v) = \begin{cases} 0, & v < 2 \\ 0.3v^3, & 2 \leq v < 12.8 \\ 629.1, & 12.8 \leq v \leq 18 \\ 0, & v > 18 \end{cases} \tag{2.12}$$

where the wind speed unit is m/s and the output power unit is kW.

2.3 Optimization Algorithm For Wind Farm Layout Problem

In this section, the conventional genetic learning particle swarm optimization (GLPSO) is introduced, and a modified GLPSO with an adaptive replacement strategy, named AGPSO, is described in detail, including its integer cording, particular boundary constraint, and adaptive replacement strategy.

2.3.1 Conventional Genetic Learning Particle Swarm Optimization

The conventional particle swarm optimization (PSO) updates the position of particles via previous best positions of each particle (*pbest*) and the global best position of particles (*gbest*) for finding an optimal solution [130]. GLPSO uses *pbest* and *gbest* to generate elite offspring by a genetic learning scheme, which provides a guidance for particles to improve the diversity of population [131]. It includes initialization, elite generation, particle update, and particle selection. In initialization, particles, *pbest*, and *gbest* are randomly generated in a boundary. The elite generation generates elite offspring by using genetic crossover, mutation, and selection. The crossover operation conducts an elite offspring by using *pbest* and *gbest*. The mutation operation mutates each dimension of the offspring with a mutation probability. The promising offspring are reserved by the genetic selection

operation. Then, the reserved offspring guide the position update of particles. Finally, particle selection chooses better particles for the next generation according to particles' fitness. Let X_i represent an individual, its velocity V_i, its *pbest* P_i, elite offspring O_i, and the *gbest* G of the population are shown as follows:

$$\begin{aligned} X_i &= \{x_i^1, x_i^2, x_i^3, ..., x_i^D\} \\ V_i &= \{v_i^1, v_i^2, v_i^3, ..., v_i^D\} \\ P_i &= \{p_i^1, p_i^2, p_i^3, ..., p_i^D\} \\ O_i &= \{o_i^1, o_i^2, o_i^3, ..., o_i^D\} \\ G &= \{g_1, g_2, g_3, ..., g_D\} \end{aligned} \tag{2.13}$$

where D is the dimension of problems. In WFLO problem, a particle represents a wind farm layout. The dth dimension of a particle represents the dth wind turbine. Therefore, x_i^d denotes the location of the dth wind turbine in the ith layout. Similar to x_i^d, o_i^d and p_i^d denote the location of elite and *pbest*, respectively. Besides, WFLO is a discrete maximum optimization problem. The particle with higher fitness is a better solution.

For a particle X_i, the dth dimension of its elite offspring is conducted by its *pbest* and *gbest* by the crossover operation, i.e.,

$$o_i^d = \begin{cases} r \cdot p_i^d + (1 - r) \cdot g_d, & \text{if } f(X_i) > f(X_k) \\ x_k^d, & \text{otherwise} \end{cases} \tag{2.14}$$

where r is a random uniformly distributed number in an interval $(0, 1)$, k is a random integer in a set $\{1, 2, ..., M\}$, M is the size of population, and $f(\cdot)$ calculates the fitness of particles.

In mutation operation, the dth dimension of the offspring is mutated with a mutation probability, i.e.,

$$o_i^d = rand(lb^d, up^d), \text{ if } r < p_m \tag{2.15}$$

where $rand(lb^d, up^d)$ generates a random number in the interval (lb^d, up^d), p_m is the mutation probability, lb^d and up^d are the lower and upper boundaries of the dth dimension, respectively.

After crossover and mutation operations, the promising particles are chosen to construct the elite

offspring by genetic selection. Then, they guide the position update of current particles, i.e.,

$$O_i = \begin{cases} O_i, & \text{if } f(P_i) < f(O_i) \\ P_i, & \text{otherwise} \end{cases} \tag{2.16}$$

For a particle dropped into a local optimum (i.e., it does not gain a better fitness in successive s iterations), a tournament selection is adopted to update its *pbest*. The tournament size is proportional to the size of population M. It is set to 20% of M. Therefore, when a particle dropped into a local optimum, 20%M individuals are randomly selected from population, where a particle with the best fitness is used to replace its *pbest*.

Finally, the velocity and position of particles is updated by the elite offspring, i.e.,

$$\begin{aligned} v_i^d &= \omega \cdot v_i^d + c \cdot r \cdot (o_i^d - x_i^d) \\ x_i^d &= v_i^d + x_i^d \end{aligned} \tag{2.17}$$

where ω denotes an inertia weight, and c is accelerate coefficients.

2.3.2 Modified GLPSO with Adaptive Replacement Strategy

Due to WFLO problem is a real-world discrete optimization problem, a modified GLPSO with a specific adaptive replacement strategy, named AGPSO, is proposed for handling the problem. A new integer coding strategy, a particular boundary constraint, and a velocity update strategy are proposed for solving it. The coding strategy reduces the dimension of the problem to reduce executing time. The boundary constraint is designed to control location of a wind turbine. In the velocity update step of particles, the inertia is removed to address the problem better. Besides, the adaptive replacement strategy adjusts the location of the worst wind turbine in the wind farm to improve the diversity of population and performance of the algorithm.

2.3.2.1 Integer Coding and Transitions

Previous studies generally use binary coding to model a wind farm where the location of wind turbines is 1, which means that the dimension is equal to the scale of the wind farm. For a 12×12

wind farm, the dimension of the problem adopted by the binary coding is 144. The number of wind turbines is much smaller than the size of the wind farm, resulting in the population adopted binary coding is a sparse matrix, which causes difficulty in optimizing it. Therefore, an integer coding strategy is proposed to reduce the dimension of the problem. When the number of wind turbines is fixed, the location of wind turbines is used to code so that the dimension of the problem is reduced to the number of wind turbines in the WFLO problem. For example, there are ten wind turbines in a 12×12 wind farm. The index of wind farms is drawn in Fig. 2.2. The dth dimension of a particle represents the dth wind turbine in the farm. The location index is randomly generated in an integer interval $[1, 144]$, i.e.,

$$X_i = \{5, 25, 37, 43, 49, 51, 82, 84, 98, 100\} \tag{2.18}$$

For calculating the fitness of layouts, the location index of wind turbines is transformed from the integer coding into a 2-dimension coordinate. The x and y axes are defined in intervals $[0, I_c]$ and $[0, I_r]$, where I_c and I_r are the number of column and row of wind farm includes, respectively. For the dth wind turbine in the ith layout, it locates the center of a grid, its coordinates represents (x_i^d, y_i^d), i.e.,

$$\begin{aligned} x_i^d &= \left(i - I_r \cdot \left\lfloor \frac{i-1}{I_r} \right\rfloor - 0.5\right) \cdot W_t \\ y_i^d &= \left(\left\lfloor \frac{i-1}{I_r} \right\rfloor + 0.5\right) \cdot W_t \end{aligned} \tag{2.19}$$

where W_t is the grid width.

Additionally, there are different wind directions that affect power generation. To actually measuring the power generation of a wind turbine, its coordinates are rotated by the wind direction θ, an actual location coordinates $(x_i^{d\theta}, y_i^{d\theta})$ are calculated by:

$$\begin{bmatrix} x_i^{d\theta} \\ y_i^{d\theta} \end{bmatrix} = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} x_i^d \\ y_i^d \end{bmatrix} \tag{2.20}$$

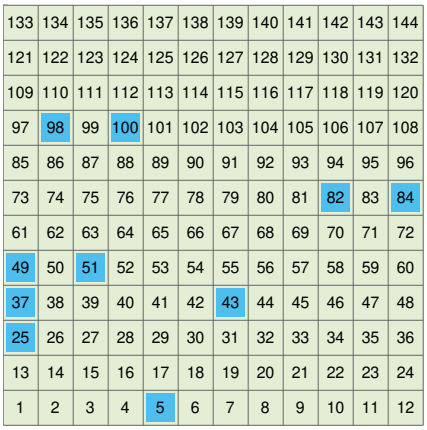


Figure 2.2: An illustration of the proposed integer coding.

2.3.2.2 Velocity Update and Boundary Constraint

The inertia of particles is removed during the velocity update of particles for the discrete WFLO problem. Therefore, Eq. (2.17) is modified as follows:

$$\begin{aligned} v_i^d &= c \cdot r \cdot (o_i^d - x_i^d) \\ x_i^d &= v_i^d + x_i^d \end{aligned} \tag{2.21}$$

In WFLO problem, a particular boundary constraint is proposed to restrict the location of wind turbines. It ensures that a wind turbine is placed in a wind farm, no multiple wind turbines located at the same place, and wind turbines are not positioned at a constraint location, formulated as:

$$x_i^d = randi(1, I_c * I_r), \text{ if } (x_i^d \notin WF) \,\&\, (x_i^d \neq x_i^k) \,\&\, (x_i^d \notin CL) \tag{2.22}$$

where *randi* randomly generates an integer in the interval $[1, I_c * I_r]$, $WF = \{1, 2, 3, ..., I_c * I_r\}$ represents all indices of a wind farm, $k = (1, 2, 3, ..., D)$ is the dimension of the problem. CL denotes the constraint locations where the wind turbine can not be placed, because there are locations that cannot locate wind turbines, such as rivers, forests, and so on, in the real application.

2.3.2.3 Adaptive Replacement Strategy

The adaptive replacement strategy replaces the worst wind turbine that generates the lowest power outputs to improve the diversity of the population and the algorithm's performance. A novel eval-

uation is proposed to measure the power generation capability of a location. For all wind turbines located at the *i*th location, their total power outputs are regarded as the power generation capability of the location. The power outputs of each wind turbine at the *i*th location are recorded and accumulated with iterations. A better location has higher accumulated power outputs. The novel evaluation empirically presents the advantage of each location. Therefore, the location with high power outputs is selected to replace the worst turbine. Before the velocity update, the worst wind turbine x_i^w in individual X_i is replaced by a tournament selection of 20%-top locations, formulated as:

$$\begin{aligned} & x_i^w = argsort(PR_i)[randi(1, 20\% * I_c * I_r)] \\ & \text{subject to} \\ & x_i^w \notin WF,\ x_i^w \neq x_i^k,\ x_i^w \notin CL \end{aligned} \tag{2.23}$$

where $argsort(PR_i)$ obtains the index of PR_i sorted by descending order. PR_i is accumulated power outputs at each location of the *i*th individual, represented by:

$$PR_i = \{P_i^1, P_i^2, P_i^3, ..., P_i^{I_c * I_r}\} \tag{2.24}$$

where P_i^1 denotes the accumulated power outputs of the first location in the *i*th individual.

The pseudo-code of AGPSO is described in Algorithm 7. Firstly, population, *pbest*, *gbest*, and velocity are initialized during the initialization step, where the population is randomly generated with integer coding, and the initial velocity of individuals is zeros. Then, the genetic learning strategy is used to update *pbest*. The tournament selection is used to improve the diversity of the population. Finally, the population and velocity are updated by using *pbest*.

2.4 Numerical Experimental Results

In the section, the experiment setups are described. The comparison between AGPSO and twelve algorithms (i.e., AGA [132], SUGGA[99], SHADE [133], CJADE [134], ALGSA [135], HGSA [136], CLPSO [137], GLPSO [131], BSA [138], IWO [139], and GWO [140]) is implemented to verify its performance under four wind scenarios and twelve constraints. AGA and SUGGA are two state-of-the-art algorithms for WFLO problems. CJADE and SHADE are two DE algorithms. AL-

Algorithm 7: The pseudo-code of AGPSO.

```
//Initialization
for i = 1 : M do
    Randomly generated X_i
    Generated velocity V_i = 0
    Evaluate f(X_i)
    Generation pbetst P_i = X_i
end
Selection the gbest G particle from pbest.
while iter < MaxIter do
    // Generation Elite Individuals
    for i = 1:M do
        // Elit Crossover
        for d = 1:D do
            k = randi(1, M)
            if f(P_i) > f(P_k) then
                o_i^d = r · p_i^d + (1 − r) · g_d
            else
                o_i^d = p_i^d
            end
        end
        // Elite Mutation
        for d = 1:D do
            if rand(0, 1) < p_m then
                o_i^d = rand(lb^d, up^d)
            end
        end
        // Elite Selection
        if f(P_i) > f(O_i) then
            O_i = P_i
        end
        if f(O_i) is not get better in s iteraions then
            Select O_j by 20%M tournament to replace O_i.
        end
    end
    Adaptively replace the worst turbine.
    // Velocity update
    V_i = c · r · (O_i − X_i)
    X_i' = X_i + V_i
    iter = +1
    // Update population
    if f(X_i') > f(X_i) then
        X_i = X_i'
    end
    Updating the gbest and pbest.
end
return X_i, f(X_i)
```

GSA and HGSA are two state-of-the-art GSA algorithms. CLPSO and GLPSO are two improved PSOs. BSA, IWO, and GWO are three state-of-the-art swarm optimization algorithm. The latter eight algorithms are used to verify the performance of different types of algorithms on WFLO problems. Finally, the best wind farm layouts, convergence graphs, and box-whisker plots are given to show the performance of algorithms.

2.4.1 Experiment Setup

The parameters of algorithms, wind farms, and wind turbines are described. The population size and maximum iterations of all algorithms are set to 120 and 200, respectively. Each experiment runs individually 51 times to gain the statistical results. The remainder parameters of the algorithms are shown in Table 2.1. The parameters of wind turbines are shown in Table 2.2. In this study, four wind scenarios (WS) are used to verify the performance of algorithms. They include single wind speed with wind direction (WS1), single speed with uniform multiple directions (WS2), single speed with nonuniform multiple directions (WS3), and multiple wind speeds with various directions scenarios (WS4). Their details are introduced in Table 2.3. The wind rose of wind scenarios is drawn in Fig. 2.3. Besides, twelve location constraints (i.e., $L_1 - L_{12}$) are used to verify the effect of different constraints for power generation and performance of algorithms, where $L_1 - L_6$ remove 24 available locations, $L_7 - L_{12}$ remove 12 available locations. L_0 represents an optimization problem without constraint. The illustration of constraints is shown in Fig. 2.4. The number of wind turbines in a wind farm affects the difficulty of WFLO problems. A larger number indicates WFLO problem is more difficult to be optimized. Therefore, WFLO problems, including 15, 20, and 25 wind turbines, are used to verify the robustness and performance of algorithms, respectively. All experiments are implemented by using MATLAB 2022a and conducted on a PC with a 3.90 GHz Intel (R) Core (TM) i9-12900K CPU and 32GB of RAM.

2.4.2 Analysis of Inertia Weight

The inertia weight is one of PSO's key parameters. It plays an important role in balancing the exploration and exploitation of algorithms [141]. A large value of inertia weight represents that algorithms explore the solution space, while a small value means that algorithms perform a local

Table 2.1: The parameters of different algorithms.

Algorihtms	Parameters
AGPSO	$c = 1.49618, s = 7, p_m = 0.01$
GLPSO	$c = 1.49618, s = 7, \omega = 0.7298, p_m = 0.01$
CLPSO	$c_1 = c_2 = 1.496, gap_m = 7$
AGA	$p_e = 0.2, p_c = 0.6, p_m = 0.1, p_r = 0.5$
SUGGA	$p_e = 0.2, p_c = 0.6, p_m = 0.1, p_r = 0.5$
SHADE	$sf = 0.5, cr = 0.5, m_s = 5, A_r = 1.4$
CJADE	$sf = 0.5, cr = 0.5, m_s = 5, c = 0.1, p = 0.05$
ALGSA	$G_0 = 100, \alpha = 20, limit = 2, p = 0.5$
HGSA	$G_0 = 100, L = 100, w_1(t) = 1 - t^6/T^6, w_2(t) = t^6/T^6$

Table 2.2: The parameters of wind turbines.

Parameters	Values
Hub height (m)	80
rotor diameter (m)	77
Surface Roughness	2.5E-4

search. In real-world applications, different problems require different exploration and exploitation strategies. Therefore, the analysis of inertia weight is implemented to select the best setup for WFLO problems. Four values of inertia weight (i.e., 1, 0.7, 0.3, and 0) are adopted to verify the effect on performance. The analysis results are shown in Table 2.4. The results exhibit that the zero inertia weight is significantly better than the others. Then, the inertia weight of AGPSO is removed to improve the exploitation of algorithms. It also indicates that an effective local search has the capability of improving algorithms' performance for WFLO problems.

2.4.3 Comparison Results under WS1

The experimental results of WS1 that is single speed ($v = 13\ m/s$) with a single direction ($\theta = 0°$) are shown in Table 2.5. The Wilcoxon signed-rank test [142, 143] is used to measure the signifi-

Table 2.3: The distribution of wind scenarios.

	Speed	0°	30°	60°	90°	120°	150°	180°	210°	240°	270°	300°	330°
WS1	$13\ m/s$	100%	–	–	–	–	–	–	–	–	–	–	–
WS2	$13\ m/s$	25%	–	–	25%	–	–	25%	–	–	25%	–	–
WS3	$13\ m/s$	20%	–	30%	–	20%	–	10%	–	10%	–	10%	–
WS4	$7\ m/s$	0.83%	0.83%	0.83%	0.83%	0.83%	0.83%	0.83%	0.83%	0.83%	0.83%	0.83%	0.83%
	$10\ m/s$	1.67%	1.67%	1.67%	1.67%	1.67%	1.67%	1.67%	1.67%	1.67%	1.67%	1.67%	1.67%
	$13\ m/s$	5.83%	5.83%	5.83%	5.83%	5.83%	5.83%	5.83%	5.83%	5.83%	5.83%	5.83%	5.83%

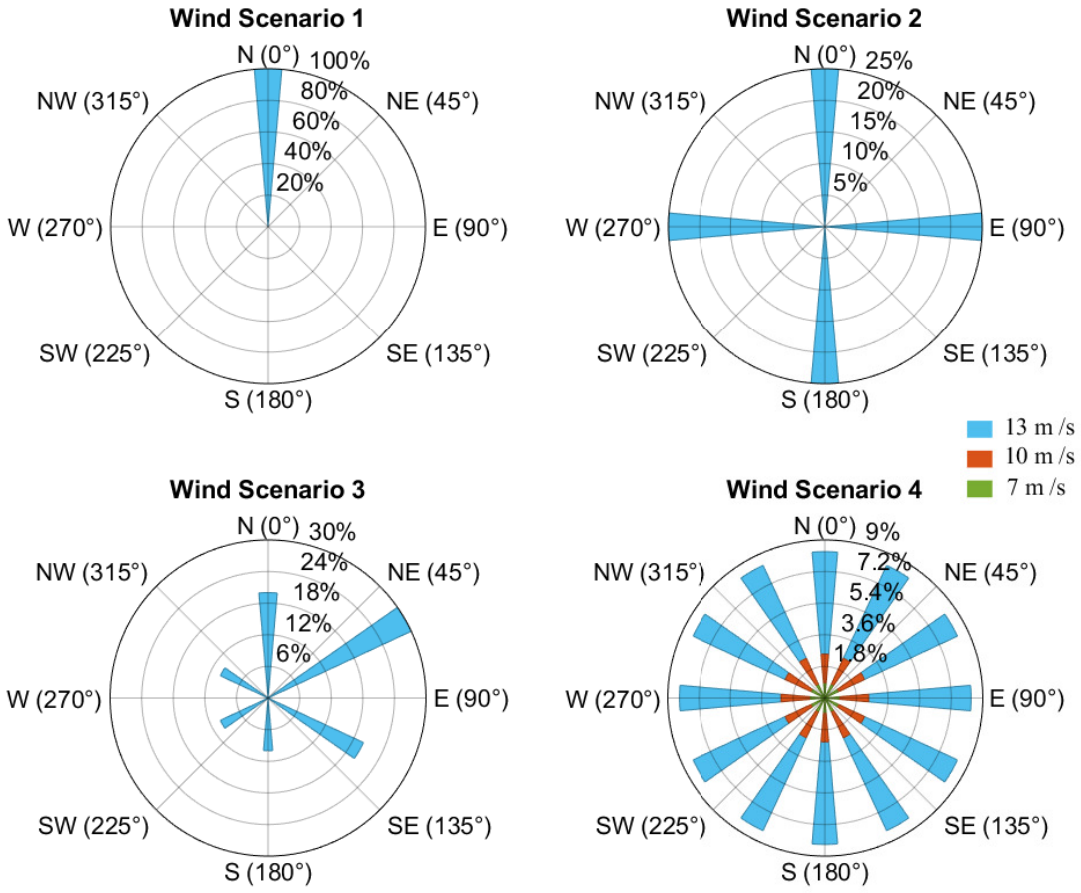


Figure 2.3: Wind rose of wind distributions.

cant difference between AGPSO and its peers. If p-value is less than 0.05, AGPSO is significantly better than its peers. Table 2.5 lists the mean and standard deviation of conversion efficiency of 51 independent experiments under different turbine numbers and constraints, where the average value refers to the average conversion efficiency of all constraints on the same wind scenario and the same number of turbines. From this table, PSO-type algorithms generally outperform other types, where GSA is the worst algorithm. It is worth noticing that those state-of-the-art DE and GSA algorithms do not use any knowledge included in the characteristics of the WFLO problem. Therefore, they are worse than AGA and SUGGA. Comparing with BSA, IWO, and GWO, AGPSO significantly outperforms them, which shows the superiority of AGPSO in swarm algorithms. Promisingly, AGPSO considers the characteristics of WFLO problems and uses particular operators to optimize the prob-

Figure 2.4: The illustration of constraints.

lem, then, it gets the best performance among all algorithms. The experimental results demonstrate that the adaptive replacement strategy has the capability of optimizing WFLO problems. Besides, the different constraints affect the conversion efficiency of a wind farm. From this table, it can be also observed that there is a difference among constraints. But comparing results between $L_1 - L_6$ and $L_7 - L_{12}$, constraints have a more negligible effect on efficiency under the simple wind scenario. The wind farm has the maximum power generation on constraint L_5 under WS1. L_5 makes turbines in the wind farm far away from other turbines, which means that empirical constraint has the capability of helping algorithms optimize the layout under a simple wind scenario. The best wind farm layouts with 25 wind turbines of all algorithms under WS1 are shown in Fig. 2.7. It is observed that the wind turbines should be located in the direction perpendicular to the wind direction. Be-

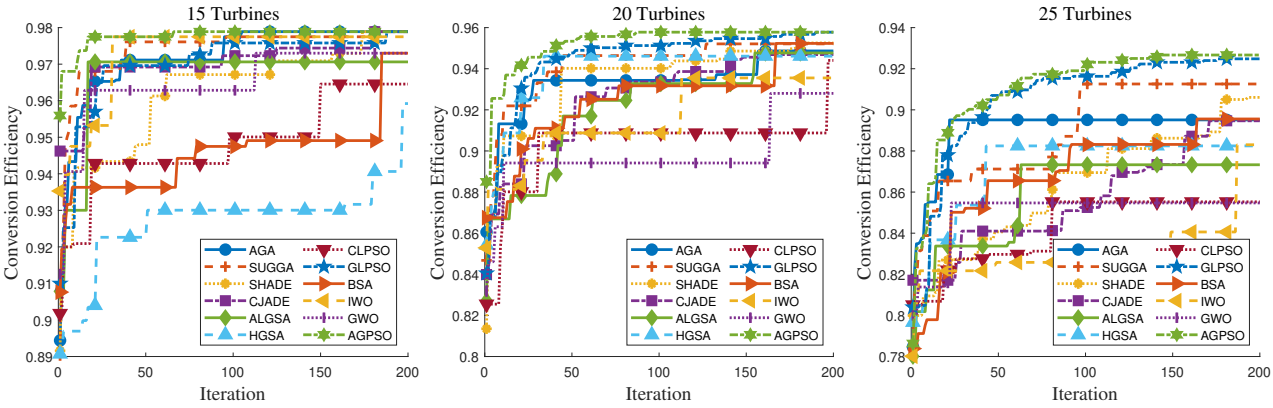


Figure 2.5: The convergence graphs of conversion efficiency under WS1.

sides, the distance between two turbines parallel to the wind direction should be as considerable as possible. This accords with the aerodynamic reasoning of wind turbine wake, in which the wake effect decreases with the distance between turbines increases [126, 144]. The results of AGPSO are better than its peers, which indicates that AGPSO has the capability of alleviating the wake effect and provides layouts satisfying aerodynamic reasoning [145]. For all constraints, AGPSO gains the best average results of 97.42%, 94.97%, and 89.92% conversion efficiency in 15, 20, and 25 turbines under WS1, respectively. In addition, the statistical test results further indicate that AGPSO is significantly better than its competitors. The convergence graphs and box-whisker plots are given in Figs. 2.5 and 2.6, respectively. From convergence graphs, it is noticed that AGPSO quickly converges to a better solution than its competitors. The convergence speed decreases as the number of turbines increases, which means that the complexity of WFLO problem increases with the turbines number increases. The box-whisker plots show that AGPSO has a more stable solution distribution than its competitors.

2.4.4 Comparison Results under WS2

WS2 is a more complex wind scenario than WS1, which possesses single wind speed (13 m/s) with four uniform wind directions ($0°, 90°, 180°,$ and $270°$) with the same probability 0.25. The experimental results of WS2 are listed in Table 2.6. From this table, AGPSO outperforms other algorithms under WS2 in terms of conversion efficiency. For this complex WS2, CJADE and SHADE are better than the state-of-the-art method SUGGA under several conditions, which indicates that DE-type al-

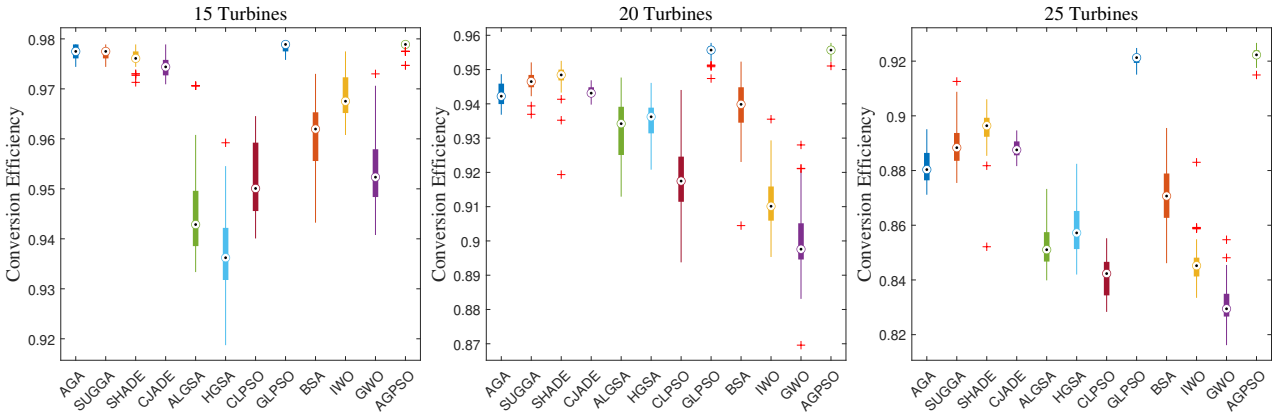


Figure 2.6: The box-whisker plots of conversion efficiency under WS1.

gorithms have the potential to handle complex WFLO problems. GSA-type algorithms are still the worst ones. In contrast to GLPSO, AGPSO has the best results on most constraints, which further demonstrates that the proposed strategy plays an important role in optimizing WFLO problems. In all constraints, L_0 has higher conversion efficiency than other constraints, which indicates that the empirical guidance can not improve the performance of algorithms for the complex WS2. The best layout with 25 turbines and the constraint L_0 under WS2 is illustrated in Fig. 2.10. Comparing results between $L_1 - L_6$ and $L_7 - L_{12}$, the constraints with high available locations have higher power generation than those with small available locations. Hence, the constraints have a high effect on the performance of algorithms for WFLO problems. For all constraints, AGPSO gets the best average conversion efficiency results of 96.31%, 89.92%, and 82.90% in 15, 20, and 25 turbines under WS2, respectively. The statistical test results mean that AGPSO is significantly better than its competitors under WS2. The results demonstrate that AGPSO has the capability of addressing the uniform multiple directions. The convergence graphs and box-whisker plots are drawn in Figs. 2.8 and 2.9, respectively. The convergence graphs and box-whisker plots indicate that AGPSO has a higher convergence speed and more stable solution than other algorithms under the complex WS2. In contrast to WS1, it is observed that the conversion efficiency of WS2 is less than that of WS1, suggesting that complex WS raises the difficulty of optimization and results in a reduction of the conversion efficiency obtained by algorithms.

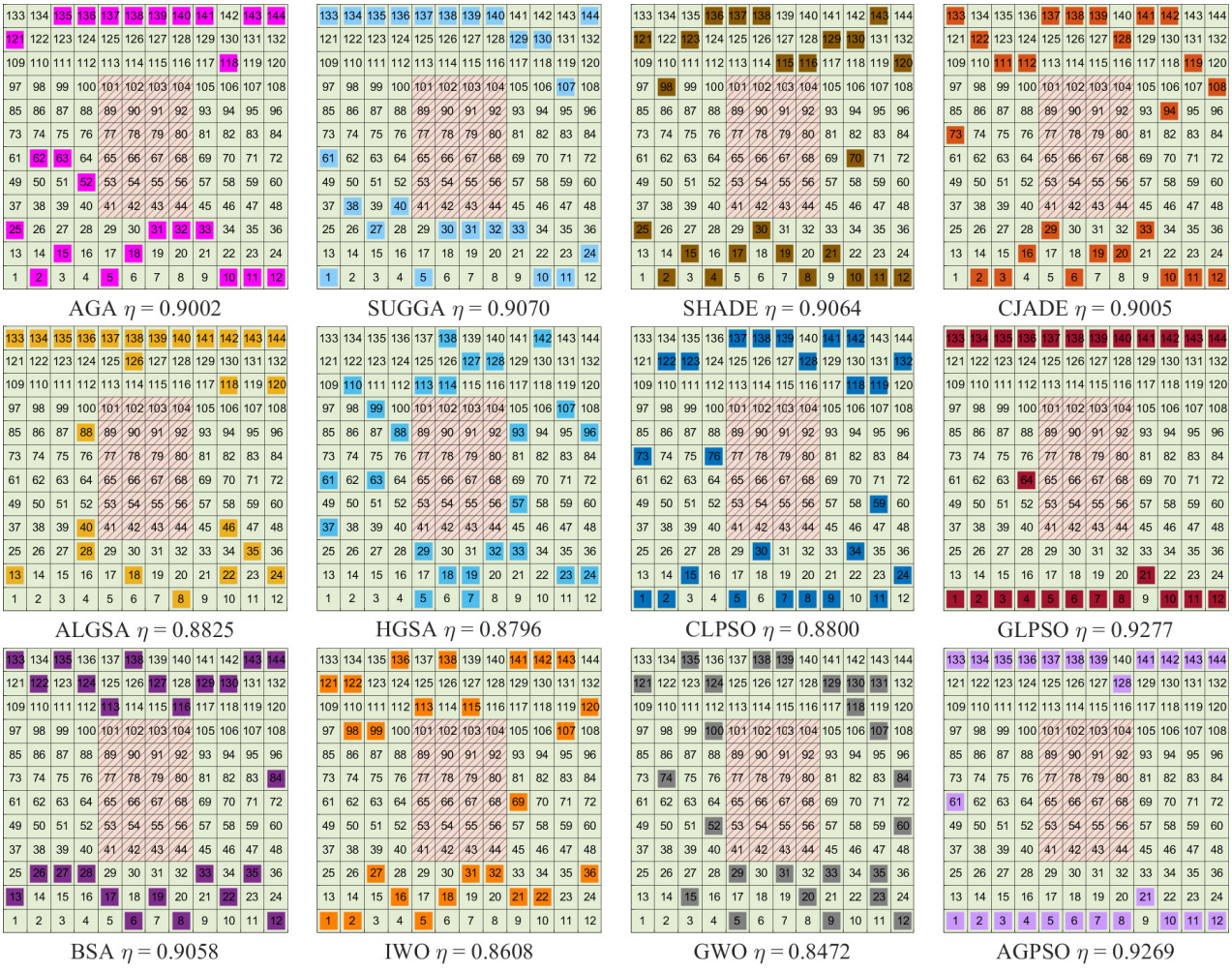


Figure 2.7: The best layout of 25 wind turbines with constrains L_5 under WS1.

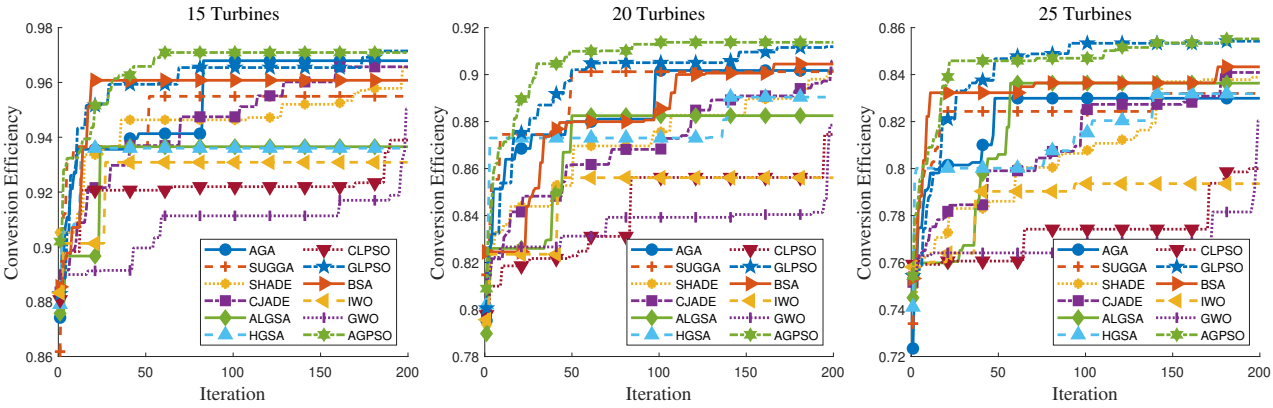


Figure 2.8: The convergence graphs of conversion efficiency under WS2.

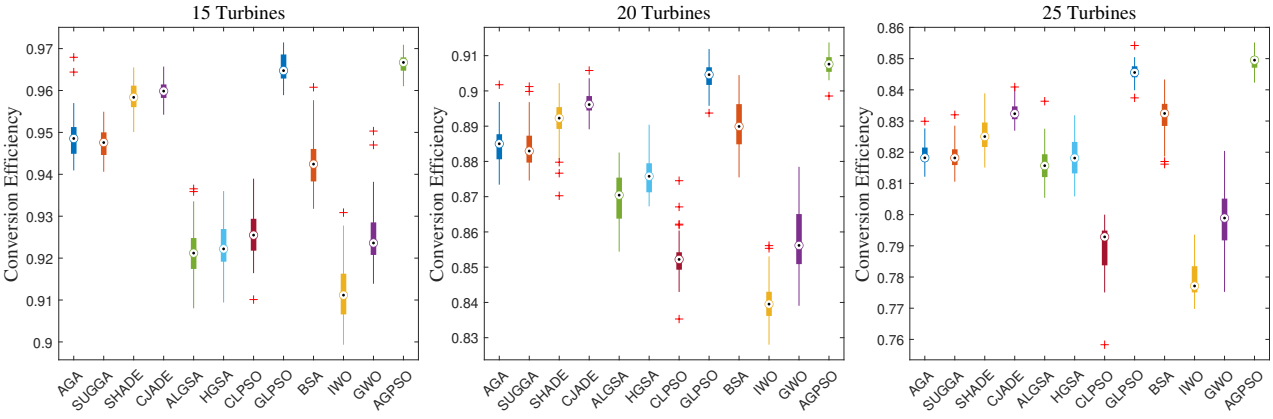
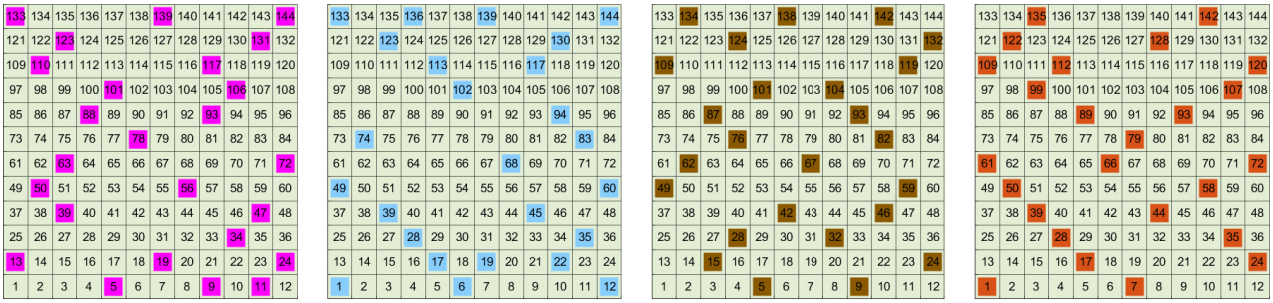


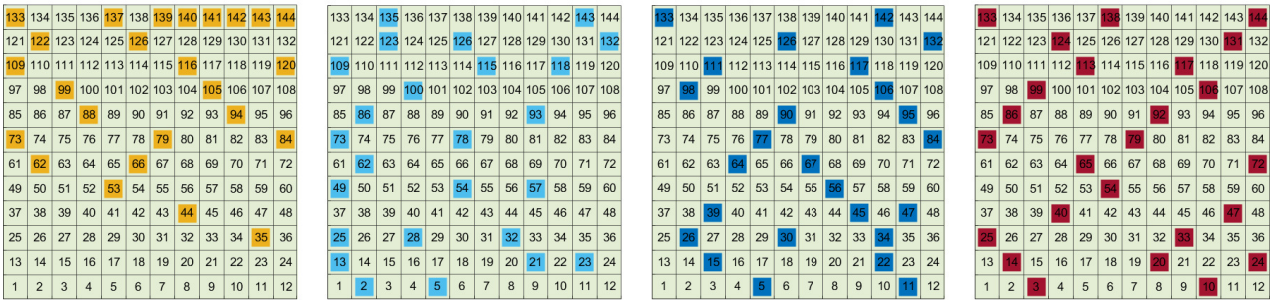
Figure 2.9: The box-whisker plots plot of conversion efficiency under WS2.

2.4.5 Comparison Results under WS3

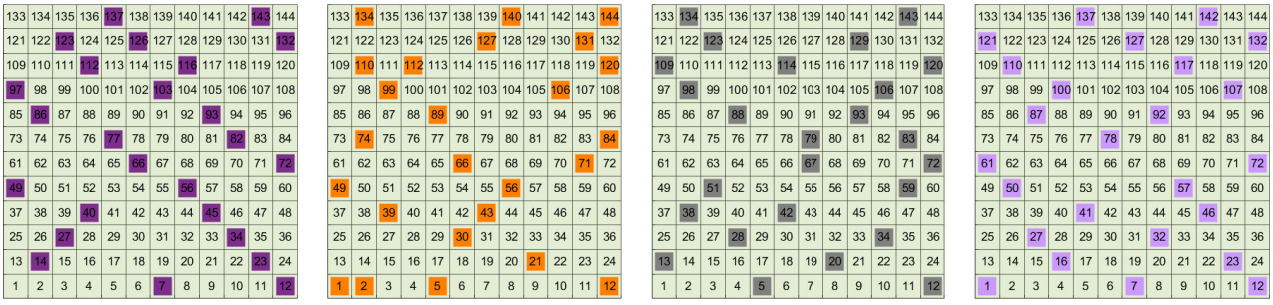
WS3 has single wind speed (13 m/s) and six wind directions (0°, 60°, 120°, 180°, 240°, and 300° with probability 0.2, 0.3, 0.2, 0.1, 0.1,and 0.1, respectively). Its numerical experiment results are shown in Table 2.7. AGPSO obtains the best mean and standard deviation in most numerical experiments among algorithms under WS3. The comparison among all algorithms with particular operators that are designed by using the characteristics of WFLO problems indicates that DE-type algorithms are better for PSO and GSA types in WFLO problems under WS3. The algorithms with particular operators are better than those without particular operators, which means that the characteristics of problems can help algorithms optimize problems. The results also show that AGPSO can provide better guidance than AGA and SUGGA for WFLO problems. For all constraints, AGPSO has the best average efficiency and gets the conversion efficiency of 99.13%, 97.97%, and 95.39% in 15,



AGA $\eta = 0.8299$ SUGGA $\eta = 0.8320$ SHADE $\eta = 0.8389$ CJADE $\eta = 0.8409$



ALGSA $\eta = 0.8363$ HGSA $\eta = 0.8319$ CLPSO $\eta = 0.8000$ GLPSO $\eta = 0.8542$



BSA $\eta = 0.8433$ IWO $\eta = 0.7936$ GWO $\eta = 0.8204$ AGPSO $\eta = 0.8552$

Figure 2.10: The best layout of 25 wind turbines with constrains L_0 under WS2.

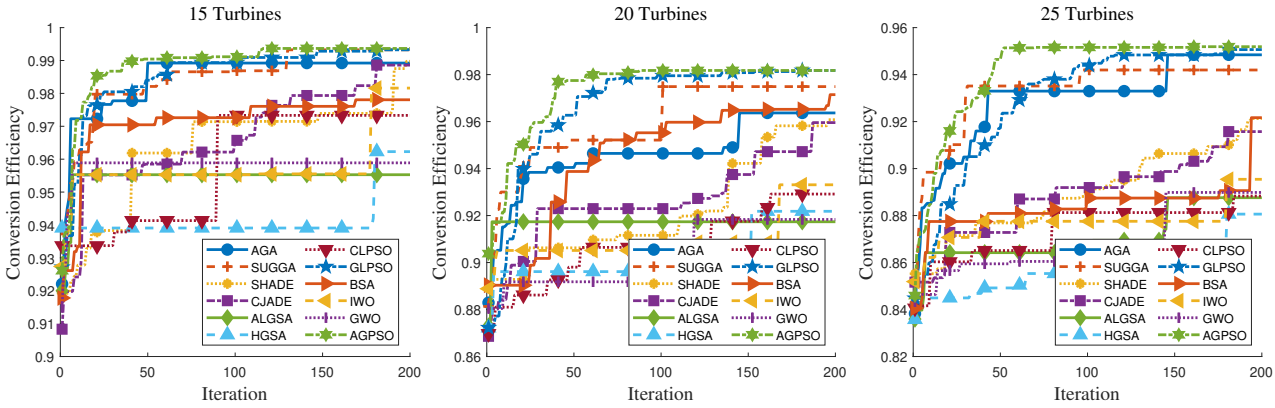


Figure 2.11: The convergence graphs of conversion efficiency under WS3.

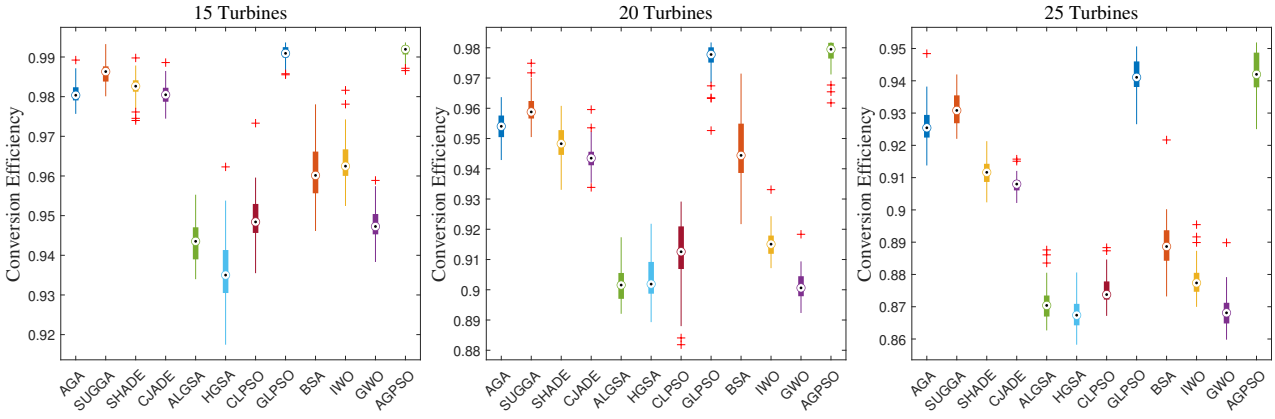
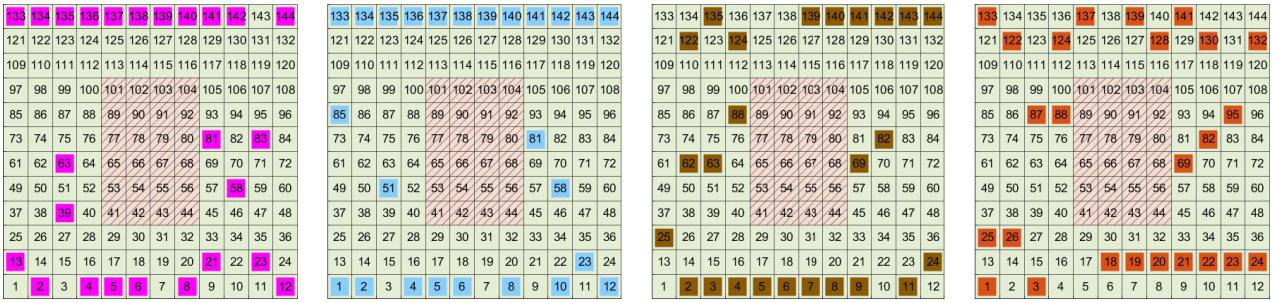
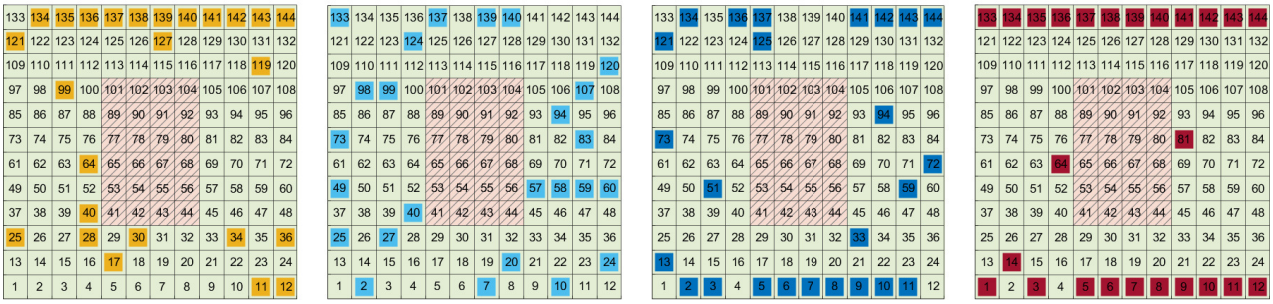


Figure 2.12: The box-whisker plots of conversion efficiency under WS3.

20, and 25 turbines under WS3, respectively. The comparison among all constraints further demonstrates that the constraints affect efficiency. A more crowded wind farm means the optimization complexity is higher. The Wilcoxon signed-rank test results demonstrate that AGPSO is superior to its peers. The convergence graphs and box-whisker plots are also given in Figs. 2.11 and 2.12, respectively. The best wind farm layout with 25 wind turbines and the constraint L_5 under WS3 is shown in Fig. 2.13. The convergence graphs imply that AGPSO has the capability of handling complex WS3. The box-whisker plots show that AGPSO provides higher quality solutions than other algorithms. In contrast to WS2, the wind direction distribution of WS3 is nonuniform. In this case, the total conversion efficiency is higher than the uniform WS2, which inspires that the WFLO problem should consider the wind scenario. Therefore, wind direction's effect is larger than wind speed's effect.



AGA $\eta = 0.9487$ SUGGA $\eta = 0.9576$ SHADE $\eta = 0.9429$ CJADE $\eta = 0.9381$



ALGSA $\eta = 0.9241$ HGSA $\eta = 0.9257$ CLPSO $\eta = 0.9112$ GLPSO $\eta = 0.9729$

BSA $\eta = 0.9560$ IWO $\eta = 0.8972$ GWO $\eta = 0.9039$ AGPSO $\eta = 0.9769$

Figure 2.13: The best layout of 25 wind turbines with constrains L_5 under WS3.

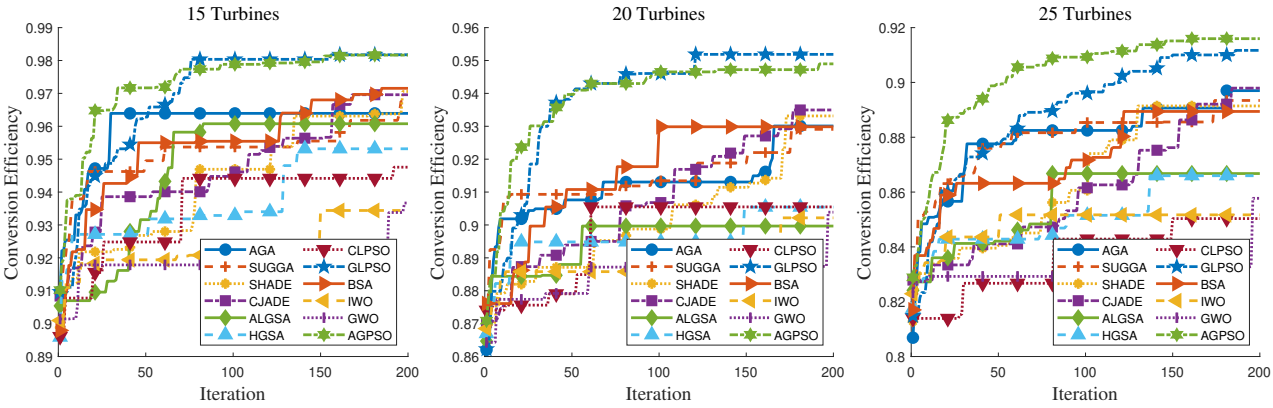


Figure 2.14: The convergence graphs of conversion efficiency under WS4.

2.4.6 Comparison Results under WS4

The above numerical experiments exhibit the performance of all algorithms under the single wind speed condition. AGPSO significantly outperforms other algorithms under such scenarios. To further verify the performance of algorithms under multiple wind speeds, a more complex wind scenario with three wind speeds and twelve directions, i.e., WS4, is used. The mean and standard deviation of conversion efficiency for WS4 are summarized in Table 2.8. AGPSO still obtains over 90% conversion efficiency for different turbine numbers and constraints under the complex WS4. For all constraints, it gains the best average value of 97.29%, 94.16%, and 90.75% in 15, 20, and 25 turbines among all algorithms, respectively. Similar to WS1, WS2, and WS3, the statistical test results indicate that AGPSO is significantly better than its competitors under WS4. The convergence graphs and box-whisker plots are shown in Figs. 2.14 and 2.15, respectively. The best wind farm layout with 25 turbines under WS4 is shown in Fig 2.16. The numerical experiment results demonstrate that AGPSO has the capability of handling multiple speeds and multiple directions. The above numerical experiments demonstrate that the adaptive replacement strategy possesses the ability to improve the performance of AGPSO under different wind scenarios.

2.5 Conclusions

This study proposes an adaptive replacement strategy-incorporated genetic learning particle swarm optimization (AGPSO) to optimize WFLO problems. The proposed algorithm is compared with

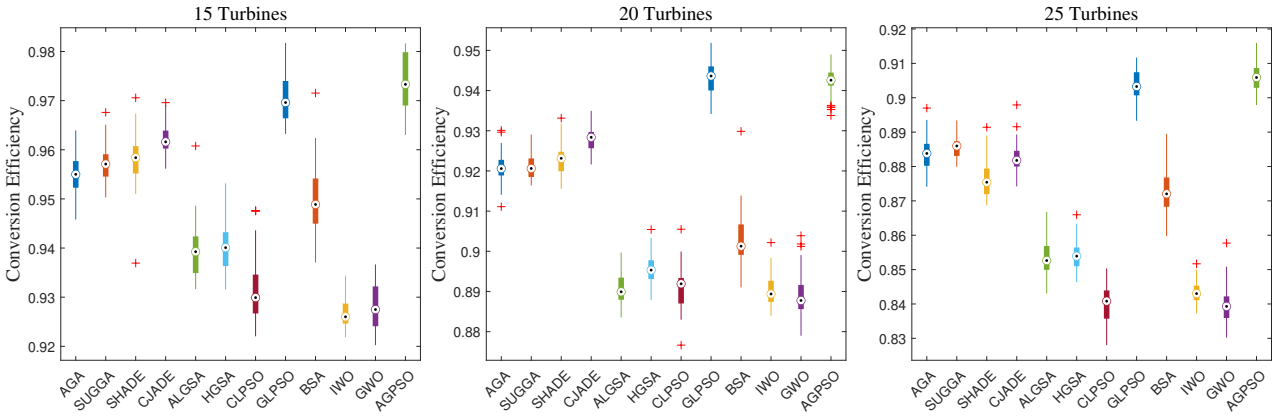


Figure 2.15: The box-whisker plots of conversion efficiency under WS4.

twelve algorithms, including two state-of-the-art methods for WFLO problems and nine representative evolutionary algorithms. Four wind scenarios, including single wind speed with single wind direction, single wind speed with uniform multiple wind directions, single wind speed with nonuniform multiple wind directions, and multiple wind speed with multi wind directions ones, are used to test the ability of handle different complexities of algorithms. Besides, twelve constraints are considered to verify the performance of all algorithms. The numerical experiment results indicate that AGPSO is significantly better than its competitors. AGPSO gains the best average of 97.42%, 94.97%, and 89.92% conversion efficiency under a single speed and direction in maximizing the objective function of WFLO problems. Under a complex wind scenario, AGPSO still obtains over 90% conversion efficiency on a wind farm with 15, 20, and 25 turbines.

General remarks and remarkable findings of this study are summarized as follows: In AGPSO, the proposed operators are designed by using the characteristics of WFLO problems. They can enhance the performance of AGPSO. The experimental results of constraints indicate that empirical rules can help algorithms improve their performance for a small wind farm with a simple wind scenario condition. Wind turbines should be located in the direction perpendicular to the wind direction, and the distance between two turbines parallel to the wind direction is as considerable as possible. Nevertheless, constraints have a high effect on power generation under complex wind scenarios. In general, a wind farm with high available locations has higher power generation ability than that with small available locations. Besides, the number of wind turbines in a wind farm is discussed to verify the effect on power generation. The complexity of WFLO problems increases

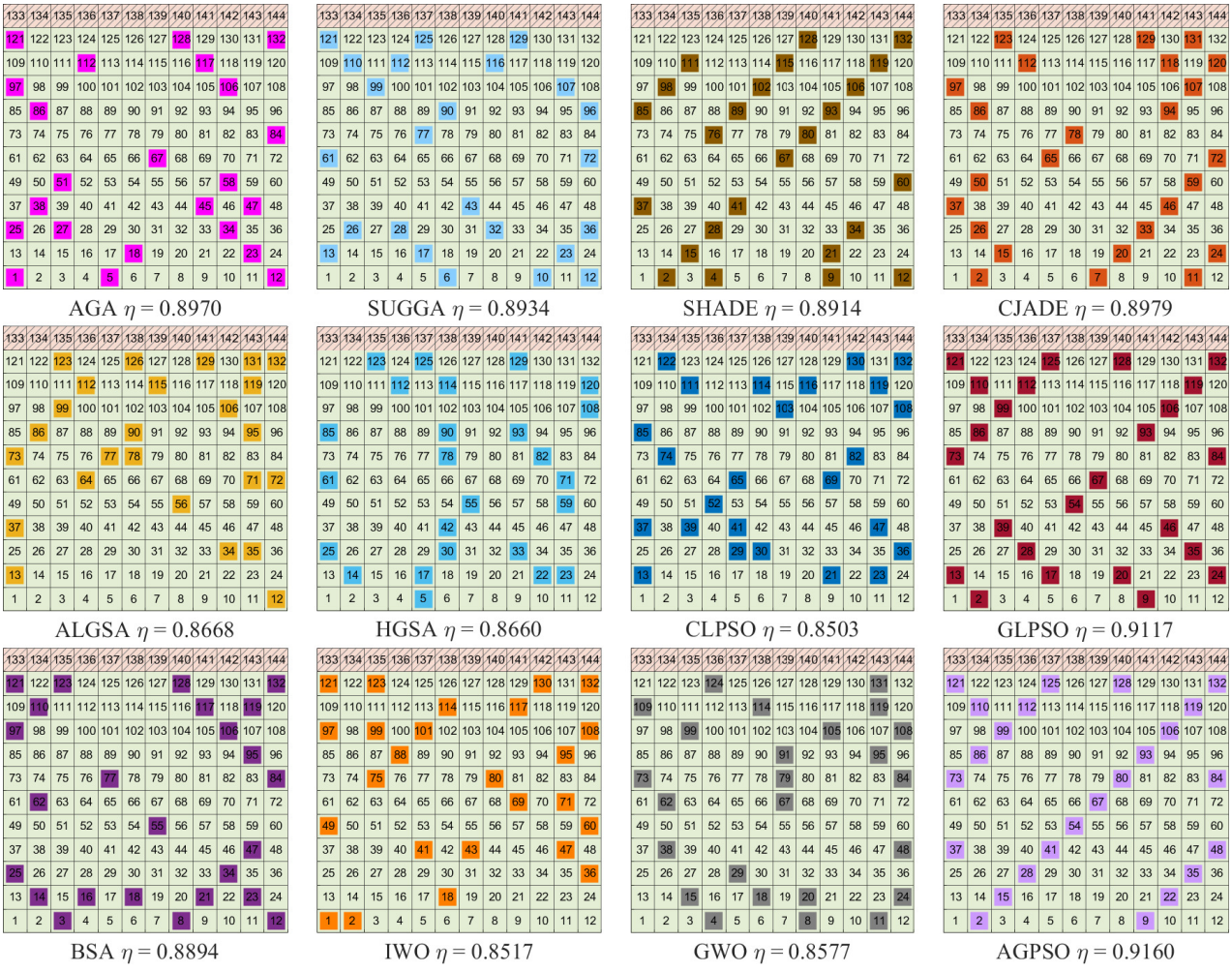


Figure 2.16: The best layout of 25 wind turbines with constrains L_7 under WS4.

as the number of turbines increases. Meanwhile, the complex wind scenario also increases the complexity. However, the wind direction has a more significant effect than wind speed on WFLO problems. In future work, as an important and challenging task, we plan to design operators that include more wind-related knowledge to further improve the performance of algorithms.

Table 2.4: The analysis of inertia weight of AGPSO under four wind scenarios.

Parameter		$\omega = 1$	$\omega = 0.5$	$\omega = 0.3$	$\omega = 0$	$\omega = 1$	$\omega = 0.5$	$\omega = 0.3$	$\omega = 0$	$\omega = 1$	$\omega = 0.5$	$\omega = 0.3$	$\omega = 0$
Turbine		15				20				25			
	L0	97.85(0.07)	97.85(0.08)	97.83(0.10)	**97.87(0.07)**	**95.48(0.32)**	95.35(0.26)	95.45(0.22)	95.48(0.32)	92.13(0.28)	92.15(0.29)	92.14(0.28)	**92.21(0.28)**
	L1	96.93(0.08)	96.90(0.14)	**96.94(0.07)**	96.94(0.07)	93.61(0.26)	93.61(0.23)	93.61(0.24)	**93.62(0.23)**	89.52(0.36)	89.59(0.30)	89.57(0.38)	**89.60(0.33)**
	L2	97.87(0.06)	97.87(0.05)	97.86(0.06)	**97.88(0.04)**	95.50(0.15)	95.54(0.16)	95.54(0.13)	**95.56(0.14)**	92.24(0.20)	92.24(0.24)	92.24(0.22)	**92.26(0.24)**
	L3	**96.37(0.12)**	96.31(0.19)	96.33(0.16)	96.37(0.12)	94.18(0.23)	94.12(0.23)	94.20(0.28)	**94.27(0.25)**	85.01(0.22)	85.03(0.20)	85.03(0.17)	**85.05(0.21)**
	L4	96.39(0.11)	96.33(0.18)	96.33(0.15)	**96.39(0.12)**	**94.24(0.27)**	94.13(0.29)	94.23(0.22)	94.24(0.27)	85.07(0.18)	85.03(0.29)	84.96(0.22)	**85.10(0.16)**
	L5	**97.87(0.05)**	97.86(0.06)	97.87(0.05)	97.87(0.05)	95.59(0.15)	**95.55(0.17)**	95.58(0.16)	95.55(0.17)	92.36(0.21)	92.34(0.22)	92.34(0.24)	**92.40(0.22)**
WS1	L6	97.83(0.10)	97.83(0.08)	97.82(0.10)	**97.86(0.09)**	95.37(0.37)	95.38(0.30)	95.43(0.28)	**95.44(0.28)**	88.46(0.44)	88.50(0.41)	**88.52(0.40)**	88.52(0.40)
	L7	97.40(0.10)	97.44(0.08)	97.44(0.07)	**97.44(0.07)**	94.62(0.22)	94.58(0.21)	**94.61(0.17)**	94.61(0.17)	90.95(0.31)	91.01(0.27)	90.94(0.32)	**91.06(0.22)**
	L8	97.84(0.08)	97.85(0.08)	97.85(0.08)	**97.86(0.06)**	95.45(0.21)	95.48(0.17)	**95.46(0.20)**	95.46(0.20)	92.21(0.25)	92.23(0.23)	92.18(0.29)	**92.24(0.23)**
	L9	97.09(0.13)	97.10(0.12)	97.08(0.12)	**97.14(0.07)**	94.77(0.26)	94.79(0.25)	94.76(0.31)	**94.84(0.23)**	88.64(0.20)	88.62(0.24)	88.65(0.22)	**88.70(0.17)**
	L10	97.10(0.09)	97.05(0.14)	97.10(0.12)	**97.12(0.10)**	94.84(0.23)	94.81(0.23)	94.79(0.30)	**94.89(0.22)**	88.57(0.22)	**88.63(0.29)**	88.61(0.19)	88.63(0.29)
	L11	97.85(0.07)	97.86(0.06)	97.85(0.08)	**97.86(0.05)**	95.50(0.18)	95.48(0.21)	**95.52(0.18)**	95.52(0.18)	**92.28(0.24)**	92.31(0.23)	92.28(0.24)	92.28(0.24)
	L12	97.80(0.10)	97.81(0.12)	97.75(0.17)	**97.82(0.10)**	**95.07(0.30)**	95.11(0.27)	95.02(0.42)	95.07(0.30)	90.81(0.26)	90.90(0.34)	90.85(0.29)	**90.93(0.24)**
p-vlaue		6.10E-04	2.44E-04	4.88E-04	–	1.08E-01	1.64E-02	9.55E-02	–	1.71E-03	1.33E-02	2.32E-03	–
	L0	**95.48(0.32)**	95.35(0.26)	95.45(0.22)	95.48(0.32)	92.13(0.28)	92.15(0.29)	92.14(0.28)	**92.21(0.28)**	97.07(0.23)	97.08(0.25)	97.02(0.23)	**97.29(0.17)**
	L1	93.61(0.26)	93.61(0.23)	93.61(0.24)	**93.62(0.23)**	89.52(0.36)	89.59(0.30)	89.57(0.38)	**89.60(0.33)**	94.95(0.33)	94.97(0.29)	95.04(0.28)	**95.09(0.26)**
	L2	95.50(0.15)	95.54(0.16)	95.54(0.13)	**95.56(0.14)**	92.24(0.20)	92.24(0.24)	92.24(0.22)	**92.26(0.24)**	95.53(0.29)	95.64(0.24)	95.58(0.27)	**95.74(0.22)**
	L3	94.18(0.23)	94.12(0.23)	94.20(0.28)	**94.27(0.25)**	85.01(0.22)	85.03(0.20)	85.03(0.17)	**85.05(0.21)**	94.95(0.33)	94.94(0.39)	94.95(0.36)	**95.06(0.23)**
	L4	**94.24(0.27)**	94.13(0.29)	94.23(0.22)	94.24(0.27)	85.07(0.18)	85.03(0.29)	84.96(0.22)	**85.10(0.16)**	95.45(0.24)	95.53(0.29)	95.54(0.30)	**95.67(0.21)**
	L5	95.59(0.15)	**95.55(0.17)**	95.58(0.16)	95.55(0.17)	92.36(0.21)	92.34(0.22)	92.34(0.24)	**92.40(0.22)**	96.49(0.30)	96.46(0.37)	96.52(0.31)	**96.66(0.23)**
WS2	L6	95.37(0.37)	95.38(0.30)	95.43(0.28)	**95.44(0.28)**	88.46(0.44)	88.50(0.41)	**88.52(0.40)**	88.52(0.40)	96.67(0.20)	96.69(0.17)	96.73(0.19)	**96.83(0.11)**
	L7	94.62(0.22)	94.58(0.21)	**94.61(0.17)**	94.61(0.17)	90.95(0.31)	91.01(0.27)	90.94(0.32)	**91.06(0.22)**	96.10(0.31)	96.09(0.31)	96.15(0.30)	**96.27(0.21)**
	L8	95.45(0.21)	95.48(0.17)	**95.46(0.20)**	95.46(0.20)	92.21(0.25)	92.23(0.23)	92.18(0.29)	**92.24(0.23)**	96.46(0.24)	96.36(0.28)	96.42(0.23)	**96.59(0.22)**
	L9	94.77(0.26)	94.79(0.25)	94.76(0.31)	**94.84(0.23)**	88.64(0.20)	88.62(0.24)	88.65(0.22)	**88.70(0.17)**	96.14(0.25)	96.08(0.30)	96.13(0.24)	**96.27(0.21)**
	L10	94.84(0.23)	94.81(0.23)	94.79(0.30)	**94.89(0.22)**	88.57(0.22)	**88.63(0.29)**	88.61(0.19)	88.63(0.29)	96.36(0.28)	96.48(0.27)	96.35(0.27)	**96.55(0.17)**
	L11	95.50(0.18)	95.48(0.21)	**95.52(0.18)**	95.52(0.18)	**92.28(0.24)**	92.31(0.23)	92.28(0.24)	92.28(0.24)	96.93(0.27)	96.92(0.26)	96.94(0.29)	**97.09(0.29)**
	L12	**95.07(0.30)**	95.11(0.27)	95.02(0.42)	95.07(0.30)	90.81(0.26)	90.90(0.34)	90.85(0.29)	**90.93(0.24)**	96.77(0.20)	96.81(0.17)	96.80(0.16)	**96.94(0.17)**
p-vlaue		1.08E-01	1.64E-02	9.55E-02	–	1.71E-03	1.33E-02	2.32E-03	–	1.22E-04	1.22E-04	1.22E-04	–
	L0	92.13(0.28)	92.15(0.29)	92.14(0.28)	**92.21(0.28)**	97.07(0.23)	97.08(0.25)	97.02(0.23)	**97.29(0.17)**	91.24(0.38)	91.21(0.39)	91.30(0.36)	**91.64(0.36)**
	L1	89.52(0.36)	89.59(0.30)	89.57(0.38)	**89.60(0.33)**	94.95(0.33)	94.97(0.29)	95.04(0.28)	**95.09(0.26)**	87.51(0.33)	87.43(0.32)	87.47(0.33)	**87.77(0.29)**
	L2	92.24(0.20)	92.24(0.24)	92.24(0.22)	**92.26(0.24)**	95.53(0.29)	95.64(0.24)	95.58(0.27)	**95.74(0.22)**	89.35(0.29)	89.24(0.22)	89.34(0.29)	**89.48(0.27)**
	L3	85.01(0.22)	85.03(0.20)	85.03(0.17)	**85.05(0.21)**	94.95(0.33)	94.94(0.39)	94.95(0.36)	**95.06(0.23)**	87.46(0.26)	87.47(0.33)	87.48(0.32)	**87.59(0.27)**
	L4	85.07(0.18)	85.03(0.29)	84.96(0.22)	**85.10(0.16)**	95.45(0.24)	95.53(0.29)	95.54(0.30)	**95.67(0.21)**	89.28(0.23)	89.29(0.27)	89.31(0.22)	**89.38(0.21)**
	L5	92.36(0.21)	92.34(0.22)	92.34(0.24)	**92.40(0.22)**	96.49(0.30)	96.46(0.37)	96.52(0.31)	**96.66(0.23)**	89.90(0.36)	89.83(0.33)	89.83(0.35)	**89.90(0.31)**
WS3	L6	88.46(0.44)	88.50(0.41)	**88.52(0.40)**	88.52(0.40)	96.67(0.20)	96.69(0.17)	96.73(0.19)	**96.83(0.11)**	89.89(0.35)	89.98(0.37)	89.92(0.33)	**90.18(0.38)**
	L7	90.95(0.31)	91.01(0.27)	90.94(0.32)	**91.06(0.22)**	96.10(0.31)	96.09(0.31)	96.15(0.30)	**96.27(0.21)**	89.61(0.40)	89.48(0.34)	89.51(0.40)	**89.97(0.29)**
	L8	92.21(0.25)	92.23(0.23)	92.18(0.29)	**92.24(0.23)**	96.46(0.24)	96.36(0.28)	96.42(0.23)	**96.59(0.22)**	90.41(0.36)	90.40(0.33)	90.43(0.38)	**90.75(0.30)**
	L9	88.64(0.20)	88.62(0.24)	88.65(0.22)	**88.70(0.17)**	96.14(0.25)	96.08(0.30)	96.13(0.24)	**96.27(0.21)**	89.56(0.37)	89.51(0.32)	89.67(0.38)	**89.75(0.41)**
	L10	88.57(0.22)	**88.63(0.29)**	88.61(0.19)	88.63(0.29)	96.36(0.28)	96.48(0.27)	96.35(0.27)	**96.55(0.17)**	90.40(0.37)	90.40(0.36)	90.49(0.32)	**90.62(0.42)**
	L11	**92.28(0.24)**	92.31(0.23)	92.28(0.24)	92.28(0.24)	96.93(0.27)	96.92(0.26)	96.94(0.29)	**97.09(0.29)**	90.64(0.36)	90.71(0.30)	90.63(0.35)	**90.89(0.25)**
	L12	90.81(0.26)	90.90(0.34)	90.85(0.29)	**90.93(0.24)**	96.77(0.20)	96.81(0.17)	96.80(0.16)	**96.94(0.17)**	90.77(0.41)	90.72(0.38)	90.85(0.43)	**91.01(0.41)**
p-vlaue		1.71E-03	1.33E-02	2.32E-03	–	1.22E-04	1.22E-04	1.22E-04	–	1.22E-04	1.22E-04	1.22E-04	–
	L0	97.07(0.23)	97.08(0.25)	97.02(0.23)	**97.29(0.17)**	91.24(0.38)	91.21(0.39)	91.30(0.36)	**91.64(0.36)**	84.51(0.31)	84.60(0.35)	84.61(0.30)	**84.90(0.29)**
	L1	94.95(0.33)	94.97(0.29)	95.04(0.28)	**95.09(0.26)**	87.51(0.33)	87.43(0.32)	87.47(0.33)	**87.77(0.29)**	79.72(0.35)	79.71(0.32)	79.72(0.28)	**80.01(0.31)**
	L2	95.53(0.29)	95.64(0.24)	95.58(0.27)	**95.74(0.22)**	89.35(0.29)	89.24(0.22)	89.34(0.29)	**89.48(0.27)**	82.35(0.29)	82.27(0.26)	82.27(0.25)	**82.44(0.22)**
	L3	94.95(0.33)	94.94(0.39)	94.95(0.36)	**95.06(0.23)**	87.46(0.26)	87.47(0.33)	87.48(0.32)	**87.59(0.27)**	79.74(0.31)	79.78(0.34)	**79.80(0.32)**	79.80(0.32)
	L4	95.45(0.24)	95.53(0.29)	95.54(0.30)	**95.67(0.21)**	89.28(0.23)	89.29(0.27)	89.31(0.22)	**89.38(0.21)**	82.37(0.29)	82.27(0.26)	82.31(0.30)	**82.39(0.25)**
	L5	96.49(0.30)	96.46(0.37)	96.52(0.31)	**96.66(0.23)**	89.90(0.36)	89.83(0.33)	89.83(0.35)	**89.90(0.31)**	**83.57(0.34)**	83.54(0.36)	83.47(0.29)	83.57(0.34)
WS4	L6	96.67(0.20)	96.69(0.17)	96.73(0.19)	**96.83(0.11)**	89.89(0.35)	89.98(0.37)	89.92(0.33)	**90.18(0.38)**	82.89(0.33)	82.98(0.39)	83.05(0.29)	**83.16(0.32)**
	L7	96.10(0.31)	96.09(0.31)	96.15(0.30)	**96.27(0.21)**	89.61(0.40)	89.48(0.34)	89.51(0.40)	**89.97(0.29)**	82.32(0.28)	82.37(0.32)	82.37(0.26)	**82.66(0.28)**
	L8	96.46(0.24)	96.36(0.28)	96.42(0.23)	**96.59(0.22)**	90.41(0.36)	90.40(0.33)	90.43(0.38)	**90.75(0.30)**	83.67(0.25)	83.50(0.27)	83.57(0.27)	**83.85(0.28)**
	L9	96.14(0.25)	96.08(0.30)	96.13(0.24)	**96.27(0.21)**	89.56(0.37)	89.51(0.32)	89.67(0.38)	**89.75(0.41)**	82.38(0.35)	82.30(0.37)	82.31(0.31)	**82.51(0.33)**
	L10	96.36(0.28)	96.48(0.27)	96.35(0.27)	**96.55(0.17)**	90.40(0.37)	90.40(0.36)	90.49(0.32)	**90.62(0.42)**	83.49(0.32)	83.53(0.30)	83.54(0.29)	**83.70(0.36)**
	L11	96.93(0.27)	96.92(0.26)	96.94(0.29)	**97.09(0.29)**	90.64(0.36)	90.71(0.30)	90.63(0.35)	**90.89(0.25)**	84.32(0.29)	84.31(0.29)	84.35(0.27)	**84.41(0.26)**
	L12	96.77(0.20)	96.81(0.17)	96.80(0.16)	**96.94(0.17)**	90.77(0.41)	90.72(0.38)	90.85(0.43)	**91.01(0.41)**	84.09(0.31)	84.04(0.33)	84.07(0.35)	**84.35(0.27)**
p-vlaue		1.22E-04	1.22E-04	1.22E-04	–	1.22E-04	1.22E-04	1.22E-04	–	2.44E-04	1.22E-04	2.44E-04	–

Table 2.5: Conversion efficiency of all algorithms under WS1.

Turbine		AGA	SUGGA	SHADE	CJADE	ALGSA	HGSA	CLPSO	GLPSO	IWO	BSA	GWO	AGPSO
	L0	97.75(0.09)	97.79(0.09)	97.82(0.08)	97.63(0.10)	96.21(0.33)	96.25(0.60)	95.89(0.66)	97.78(0.13)	97.09(0.31)	97.12(0.42)	96.08(0.58)	**97.87(0.07)**
	L1	96.90(0.10)	96.88(0.11)	96.92(0.08)	96.76(0.10)	95.23(0.13)	95.20(0.51)	95.03(0.60)	96.91(0.09)	96.18(0.30)	96.20(0.47)	95.22(0.60)	**96.94(0.07)**
	L2	97.82(0.08)	97.85(0.06)	97.88(0.04)	97.78(0.07)	97.23(0.16)	97.29(0.17)	97.23(0.24)	97.86(0.06)	97.43(0.26)	97.44(0.30)	96.81(0.33)	**97.88(0.04)**
	L3	95.82(0.22)	95.86(0.20)	95.98(0.20)	95.47(0.24)	93.52(0.93)	93.45(0.71)	93.70(0.69)	96.33(0.16)	94.67(0.45)	94.72(0.81)	93.21(0.79)	**96.37(0.12)**
	L4	95.89(0.19)	95.86(0.25)	95.86(0.26)	95.36(0.22)	92.91(0.68)	92.92(0.77)	93.76(0.68)	96.32(0.14)	94.67(0.46)	94.43(0.67)	92.94(0.72)	**96.39(0.12)**
	L5	97.82(0.09)	97.85(0.07)	97.86(0.06)	97.85(0.06)	97.56(0.13)	97.55(0.12)	97.34(0.27)	97.86(0.08)	97.29(0.24)	97.62(0.21)	96.76(0.39)	**97.87(0.05)**
15	L6	97.73(0.12)	97.71(0.11)	97.62(0.15)	97.41(0.18)	94.54(0.99)	93.75(0.88)	95.22(0.71)	97.81(0.10)	96.84(0.45)	96.09(0.67)	95.37(0.72)	**97.86(0.09)**
	L7	97.37(0.11)	97.36(0.12)	97.39(0.13)	97.25(0.12)	95.88(0.44)	95.78(0.48)	95.51(0.74)	97.40(0.10)	96.71(0.29)	96.61(0.54)	95.59(0.68)	**97.44(0.07)**
	L8	97.76(0.10)	97.81(0.09)	97.83(0.11)	97.75(0.09)	96.99(0.29)	97.12(0.21)	96.61(0.56)	97.80(0.10)	97.36(0.24)	97.31(0.34)	96.39(0.41)	**97.86(0.06)**
	L9	96.85(0.15)	96.91(0.16)	96.91(0.13)	96.60(0.17)	94.68(0.63)	94.47(0.60)	95.16(0.67)	97.06(0.13)	95.95(0.32)	96.04(0.54)	94.89(0.63)	**97.14(0.07)**
	L10	96.85(0.15)	96.91(0.14)	96.94(0.11)	96.61(0.13)	94.52(0.61)	94.54(0.51)	95.01(0.70)	97.07(0.14)	95.98(0.40)	95.82(0.54)	94.62(0.59)	**97.12(0.10)**
	L11	97.79(0.08)	97.85(0.06)	97.86(0.05)	97.79(0.09)	97.42(0.13)	97.18(0.22)	97.04(0.38)	97.84(0.09)	97.13(0.31)	97.45(0.29)	96.25(0.62)	**97.86(0.05)**
	L12	97.65(0.13)	97.73(0.13)	97.60(0.28)	97.35(0.17)	95.21(0.60)	94.92(0.65)	95.52(0.54)	97.76(0.18)	96.98(0.30)	96.40(0.60)	95.61(0.64)	**97.82(0.10)**
Average		97.23(0.12)	97.26(0.12)	97.27(0.13)	97.05(0.13)	95.53(0.47)	95.41(0.50)	95.62(0.57)	97.37(0.12)	96.48(0.33)	96.40(0.49)	95.36(0.59)	**97.42(0.08)**
p-vlaue		1.22E-04	1.22E-04	1.22E-03	1.22E-04	1.22E-04	1.22E-04	1.22E-04	1.22E-04	1.22E-04	1.22E-04	1.22E-04	–
	L0	93.49(0.45)	94.24(0.39)	93.61(0.31)	92.94(0.34)	89.54(0.61)	89.19(0.91)	89.30(1.11)	95.32(0.26)	90.77(0.67)	91.91(1.12)	89.02(0.81)	**95.48(0.32)**
	L1	91.91(0.64)	92.67(0.55)	91.77(0.51)	90.96(0.38)	87.00(0.61)	86.91(1.18)	87.01(1.03)	93.49(0.22)	88.77(0.80)	90.16(1.28)	86.82(0.87)	**93.62(0.23)**
	L2	93.98(0.35)	94.57(0.23)	94.61(0.30)	94.18(0.23)	92.54(0.53)	92.90(0.55)	91.44(1.09)	95.46(0.19)	91.84(0.79)	93.10(0.95)	90.42(0.68)	**95.56(0.14)**
	L3	93.09(0.70)	94.22(0.65)	90.27(0.70)	89.36(0.45)	84.80(0.87)	85.14(0.93)	86.17(1.28)	94.04(0.31)	86.54(0.91)	87.12(1.23)	84.51(0.90)	**94.27(0.25)**
	L4	92.93(0.87)	93.42(1.02)	90.16(0.84)	89.22(0.49)	84.27(0.84)	84.14(0.87)	85.45(0.91)	93.95(0.40)	86.32(0.78)	86.90(1.01)	84.37(1.04)	**94.24(0.27)**
	L5	94.26(0.33)	94.65(0.30)	94.77(0.51)	94.34(0.19)	93.17(0.89)	93.50(0.61)	91.91(0.98)	95.51(0.21)	91.17(0.80)	93.87(0.85)	89.97(1.13)	**95.55(0.17)**
20	L6	92.19(0.64)	93.25(0.68)	91.56(0.65)	90.87(0.41)	85.69(0.91)	84.57(1.06)	86.04(1.07)	95.26(0.32)	89.14(0.82)	88.00(1.07)	86.72(0.79)	**95.44(0.28)**
	L7	92.78(0.43)	93.55(0.44)	92.79(0.52)	91.95(0.34)	88.54(0.78)	87.91(1.01)	87.79(1.12)	94.53(0.27)	89.88(0.62)	91.30(1.13)	88.05(0.95)	**94.61(0.17)**
	L8	93.77(0.45)	94.36(0.34)	94.27(0.48)	93.65(0.24)	91.40(0.69)	91.54(0.72)	90.63(0.93)	95.43(0.25)	91.20(0.63)	92.69(0.97)	89.56(0.75)	**95.46(0.20)**
	L9	93.30(0.54)	94.18(0.42)	92.19(0.57)	91.40(0.41)	87.02(0.94)	86.94(0.77)	88.00(1.09)	94.63(0.28)	88.76(0.75)	89.70(0.85)	86.66(0.90)	**94.84(0.23)**
	L10	93.21(0.62)	93.87(0.57)	92.07(0.80)	91.33(0.38)	86.73(0.86)	86.65(0.89)	87.36(1.01)	94.67(0.31)	88.88(0.88)	89.10(0.97)	86.57(0.92)	**94.89(0.22)**
	L11	93.84(0.40)	94.40(0.37)	94.42(0.47)	93.89(0.23)	92.70(0.92)	92.10(0.61)	90.80(0.89)	95.43(0.21)	91.04(0.75)	93.05(1.03)	89.51(1.05)	**95.52(0.18)**
	L12	92.73(0.50)	93.61(0.50)	92.66(0.49)	91.73(0.41)	87.01(0.95)	86.24(1.14)	87.47(1.09)	94.99(0.30)	90.05(0.70)	89.72(1.30)	88.10(1.01)	**95.07(0.30)**
Average		93.19(0.53)	93.92(0.50)	92.70(0.55)	91.99(0.35)	88.49(0.80)	88.29(0.86)	88.41(1.05)	94.82(0.27)	89.57(0.76)	90.51(1.06)	87.72(0.91)	**94.97(0.23)**
p-vlaue		1.22E-04	1.22E-04	1.22E-04	1.22E-04	1.22E-04	1.22E-04	1.22E-04	1.22E-04	1.22E-04	1.22E-04	1.22E-04	–
	L0	87.57(0.61)	88.46(0.64)	87.59(0.60)	86.94(0.47)	82.40(0.85)	82.13(1.19)	82.01(0.99)	91.94(0.32)	83.48(0.61)	85.75(1.16)	81.27(0.82)	**92.21(0.28)**
	L1	85.34(0.82)	86.17(0.90)	84.77(0.74)	84.14(0.58)	79.34(0.71)	78.88(1.06)	79.70(1.32)	89.33(0.38)	80.77(0.63)	83.52(1.53)	78.82(0.81)	**89.60(0.33)**
	L2	88.15(0.64)	88.90(0.77)	89.50(0.80)	88.79(0.35)	85.29(0.83)	85.83(0.89)	84.13(0.64)	92.08(0.24)	84.58(0.83)	87.07(1.22)	83.15(0.76)	**92.26(0.24)**
	L3	86.95(0.94)	85.05(0.21)	81.88(0.57)	81.16(0.39)	76.63(0.90)	77.07(0.72)	77.26(0.75)	84.86(0.25)	77.85(0.73)	79.32(1.17)	76.08(0.84)	**85.05(0.21)**
	L4	86.62(1.13)	85.10(0.16)	81.64(0.76)	80.95(0.38)	76.04(0.76)	76.14(0.82)	76.95(0.77)	84.94(0.22)	77.99(0.63)	78.60(1.04)	75.92(0.70)	**85.10(0.16)**
	L5	88.28(0.61)	88.98(0.69)	89.73(0.55)	89.14(0.43)	85.90(0.88)	86.17(0.70)	85.15(1.28)	92.21(0.27)	84.02(0.79)	88.19(1.10)	82.87(1.00)	**92.40(0.22)**
25	L6	85.18(0.89)	86.59(0.87)	84.26(0.58)	83.69(0.42)	77.78(0.83)	76.89(1.00)	78.16(1.17)	88.38(0.39)	80.73(0.79)	80.14(1.03)	78.59(0.83)	**88.52(0.40)**
	L7	86.49(0.73)	87.24(0.81)	86.18(0.69)	85.54(0.49)	80.77(0.80)	80.71(1.11)	80.89(1.16)	90.71(0.35)	82.17(0.81)	84.65(1.33)	80.11(0.84)	**91.06(0.22)**
	L8	87.82(0.62)	88.64(0.63)	88.81(0.53)	87.99(0.52)	84.42(0.89)	84.60(0.82)	83.67(0.89)	92.05(0.27)	84.08(0.71)	86.56(1.34)	82.46(0.75)	**92.24(0.23)**
	L9	87.07(0.69)	88.49(0.86)	84.72(1.09)	84.43(0.53)	79.31(0.95)	79.43(0.95)	80.31(0.99)	88.49(0.23)	81.02(0.69)	82.35(1.41)	79.03(1.05)	**88.70(0.17)**
	L10	87.10(0.85)	87.86(1.03)	85.11(0.50)	84.18(0.44)	79.09(0.83)	79.11(0.86)	80.02(0.89)	88.45(0.31)	81.02(0.68)	81.98(1.01)	79.02(0.82)	**88.63(0.29)**
	L11	87.99(0.75)	88.78(0.70)	88.63(1.00)	88.14(0.47)	85.20(0.93)	85.12(0.92)	84.04(1.15)	92.03(0.31)	83.92(0.91)	87.06(1.43)	82.29(0.97)	**92.28(0.24)**
	L12	86.36(0.74)	87.19(0.51)	85.76(0.72)	85.21(0.48)	79.35(0.93)	78.53(1.06)	79.52(1.03)	90.51(0.44)	82.16(0.64)	82.25(1.40)	80.01(0.81)	**90.93(0.24)**
Average		86.99(0.77)	87.87(0.79)	86.04(0.70)	85.41(0.46)	80.89(0.85)	80.81(0.93)	80.91(1.00)	89.69(0.31)	81.83(0.73)	83.65(1.24)	79.97(0.85)	**89.92(0.25)**
p-vlaue		1.22E-03	3.05E-03	1.22E-04	1.22E-04	1.22E-04	1.22E-04	1.22E-04	1.22E-04	1.22E-04	1.22E-04	1.22E-04	–

Table 2.6: Conversion efficiency of all algorithms under WS2.

Turbine		AGA	SUGGA	SHADE	CJADE	ALGSA	HGSA	CLPSO	GLPSO	IWO	BSA	GWO	AGPSO
	L0	95.48(0.39)	95.38(0.41)	96.52(0.32)	96.74(0.20)	95.65(0.39)	95.54(0.30)	94.51(0.70)	97.14(0.20)	92.00(0.63)	96.27(0.67)	94.30(0.95)	**97.29(0.17)**
	L1	95.04(0.73)	95.09(0.26)	94.33(0.45)	94.55(0.26)	93.37(0.55)	93.60(0.48)	91.72(0.78)	94.94(0.33)	89.52(0.82)	94.02(0.68)	92.02(1.05)	**95.09(0.26)**
	L2	94.89(0.62)	94.77(0.54)	95.14(0.34)	95.40(0.22)	93.03(0.61)	93.16(0.68)	92.95(0.88)	95.59(0.24)	90.71(0.69)	94.69(0.46)	92.53(0.98)	**95.74(0.22)**
	L3	95.18(0.68)	95.06(0.23)	94.31(0.53)	94.57(0.27)	91.87(0.79)	92.80(0.50)	90.77(0.72)	94.85(0.29)	89.64(0.76)	93.57(0.56)	91.17(0.82)	**95.06(0.23)**
	L4	94.99(0.62)	94.90(0.55)	95.06(0.33)	95.25(0.18)	92.09(0.59)	93.02(0.52)	91.63(0.77)	95.53(0.24)	91.51(0.62)	93.67(0.69)	91.81(0.72)	**95.67(0.21)**
	L5	94.86(0.54)	94.73(0.38)	95.85(0.35)	95.98(0.25)	92.15(0.66)	92.28(0.61)	92.55(0.57)	96.55(0.35)	91.17(0.65)	94.26(0.61)	92.53(0.76)	**96.66(0.23)**
15	L6	95.46(0.45)	95.33(0.33)	96.23(0.32)	96.42(0.20)	91.55(0.82)	92.35(0.76)	92.04(0.91)	96.71(0.18)	91.31(0.69)	95.39(0.81)	92.62(1.05)	**96.83(0.11)**
	L7	95.43(0.53)	95.55(0.35)	95.45(0.42)	95.71(0.25)	94.57(0.42)	94.86(0.45)	93.00(0.70)	96.13(0.24)	90.77(0.71)	95.28(0.47)	93.28(1.00)	**96.27(0.21)**
	L8	95.18(0.62)	95.10(0.48)	95.96(0.33)	96.08(0.21)	94.40(0.51)	94.64(0.60)	93.54(0.78)	96.43(0.23)	91.42(0.73)	95.45(0.56)	93.50(1.05)	**96.59(0.22)**
	L9	95.24(0.54)	95.46(0.48)	95.52(0.44)	95.74(0.28)	93.95(0.59)	94.54(0.48)	92.59(0.67)	96.04(0.27)	90.85(0.67)	95.19(0.55)	92.70(0.95)	**96.27(0.21)**
	L10	95.19(0.40)	95.23(0.50)	95.78(0.56)	96.09(0.26)	93.89(0.57)	94.71(0.41)	92.58(0.65)	96.44(0.24)	92.15(0.59)	95.18(0.56)	92.88(0.65)	**96.55(0.17)**
	L11	95.28(0.35)	95.15(0.39)	96.44(0.33)	96.59(0.27)	93.68(0.65)	93.98(0.50)	93.29(0.69)	96.92(0.33)	91.63(0.57)	95.57(0.57)	93.39(0.76)	**97.09(0.29)**
	L12	95.24(0.39)	95.45(0.34)	96.30(0.41)	96.47(0.20)	94.26(0.64)	93.96(0.51)	93.15(0.80)	96.81(0.17)	92.04(0.66)	95.70(0.77)	93.61(1.15)	**96.94(0.17)**
Average		95.19(0.53)	95.21(0.45)	95.61(0.40)	95.81(0.24)	93.42(0.60)	93.80(0.52)	92.64(0.74)	96.16(0.25)	91.13(0.68)	94.94(0.61)	92.79(0.91)	**96.31(0.21)**
p-vlaue		3.66E-04	6.10E-04	1.22E-04	1.22E-04	1.22E-04	1.22E-04	1.22E-04	1.22E-04	1.22E-04	1.22E-04	1.22E-04	–
	L0	88.83(0.52)	88.76(0.40)	89.75(0.60)	90.35(0.36)	88.67(0.46)	88.83(0.45)	86.27(0.63)	91.21(0.28)	84.73(0.63)	89.81(0.71)	87.19(1.17)	**91.64(0.36)**
	L1	87.56(0.83)	87.77(0.29)	85.91(0.61)	86.71(0.29)	84.74(0.54)	85.08(0.57)	83.06(0.81)	87.50(0.29)	81.36(0.56)	86.14(0.85)	83.33(1.02)	**87.77(0.29)**
	L2	88.27(0.69)	88.17(0.81)	88.10(0.55)	88.65(0.32)	85.62(0.62)	85.69(0.64)	84.56(0.79)	89.40(0.25)	83.12(0.61)	87.91(0.59)	84.14(1.06)	**89.48(0.27)**
	L3	88.36(0.91)	87.59(0.27)	86.12(0.61)	86.68(0.39)	82.96(0.66)	84.06(0.59)	82.09(0.57)	87.51(0.35)	81.22(0.59)	85.24(0.72)	82.69(0.83)	**87.59(0.27)**
	L4	88.37(0.67)	88.09(0.64)	88.20(0.57)	88.52(0.29)	84.52(0.64)	85.55(0.54)	83.93(0.68)	89.33(0.27)	83.20(0.58)	86.78(0.70)	84.06(0.81)	**89.38(0.21)**
	L5	88.20(0.47)	87.97(0.55)	88.77(0.32)	89.02(0.30)	84.38(0.59)	84.77(0.86)	84.80(0.63)	89.80(0.36)	83.95(0.60)	87.10(0.72)	84.83(0.60)	**89.90(0.31)**
20	L6	88.33(0.45)	88.38(0.49)	88.39(0.59)	88.75(0.34)	83.59(0.78)	84.19(0.77)	83.53(0.77)	89.91(0.37)	83.37(0.65)	87.10(0.86)	84.02(0.88)	**90.18(0.38)**
	L7	88.58(0.53)	88.76(0.53)	88.04(0.59)	88.70(0.32)	86.90(0.65)	87.00(0.51)	84.39(0.74)	89.59(0.32)	83.21(0.71)	88.09(0.68)	85.36(1.01)	**89.97(0.29)**
	L8	88.48(0.56)	88.42(0.63)	89.17(0.64)	89.67(0.34)	86.96(0.71)	87.59(0.55)	85.25(0.64)	90.40(0.36)	83.99(0.61)	89.00(0.76)	85.76(1.01)	**90.75(0.30)**
	L9	88.56(0.52)	88.57(0.57)	87.97(0.81)	88.61(0.30)	86.07(0.56)	86.74(0.43)	84.21(0.68)	89.56(0.34)	83.09(0.61)	87.80(0.78)	85.15(1.11)	**89.75(0.41)**
	L10	88.55(0.53)	88.41(0.55)	88.94(0.72)	89.54(0.31)	86.69(0.68)	87.79(0.59)	84.94(0.83)	90.38(0.38)	84.00(0.65)	88.55(0.79)	85.52(1.03)	**90.62(0.42)**
	L11	88.44(0.41)	88.54(0.47)	89.33(0.74)	90.01(0.35)	86.18(0.67)	86.85(0.54)	85.66(0.60)	90.70(0.33)	84.39(0.49)	88.83(0.78)	86.09(1.05)	**90.89(0.25)**
	L12	88.77(0.47)	88.70(0.44)	89.02(0.69)	89.52(0.39)	85.94(0.49)	85.83(0.66)	84.94(0.85)	90.73(0.43)	84.40(0.68)	88.80(0.74)	85.69(1.03)	**91.01(0.41)**
Average		88.41(0.58)	88.40(0.57)	88.29(0.62)	88.82(0.33)	85.63(0.62)	86.15(0.59)	84.43(0.71)	89.69(0.33)	83.39(0.62)	87.78(0.74)	84.91(0.97)	**89.92(0.32)**
p-vlaue		3.66E-04	6.10E-04	1.22E-04	1.22E-04	1.22E-04	1.22E-04	1.22E-04	1.22E-04	1.22E-04	1.22E-04	1.22E-04	–
	L0	81.92(0.40)	81.86(0.41)	82.56(0.50)	83.27(0.32)	81.56(0.58)	81.80(0.64)	78.93(0.81)	84.56(0.29)	77.95(0.63)	83.17(0.64)	79.89(0.97)	**84.90(0.29)**
	L1	80.48(1.00)	80.01(0.31)	77.94(0.59)	78.57(0.37)	77.34(0.44)	77.21(0.43)	74.23(0.65)	79.73(0.31)	73.79(0.46)	78.43(0.66)	75.49(0.90)	**80.01(0.31)**
	L2	81.44(0.55)	81.06(0.82)	80.65(0.61)	81.17(0.38)	77.99(0.59)	77.75(0.65)	76.80(0.66)	82.23(0.27)	76.00(0.55)	80.81(0.86)	76.77(0.92)	**82.44(0.22)**
	L3	81.26(0.80)	79.80(0.32)	77.95(0.73)	78.66(0.30)	75.10(0.55)	75.90(0.53)	74.18(0.77)	79.86(0.32)	73.92(0.54)	77.22(0.70)	74.89(0.96)	**79.80(0.32)**
	L4	81.73(0.55)	81.51(0.69)	80.66(0.64)	81.14(0.29)	77.36(0.68)	78.11(0.53)	76.12(0.59)	82.26(0.29)	75.98(0.50)	79.51(0.71)	76.70(0.68)	**82.39(0.25)**
	L5	81.63(0.52)	81.32(0.46)	81.89(0.51)	82.42(0.32)	77.51(0.67)	77.74(0.77)	77.57(0.50)	83.53(0.30)	77.32(0.58)	80.43(0.88)	77.81(0.63)	**83.57(0.34)**
25	L6	81.35(0.54)	81.50(0.53)	80.86(0.71)	81.28(0.47)	75.48(0.63)	76.06(0.67)	75.71(0.79)	82.94(0.44)	76.14(0.71)	79.29(0.99)	76.16(0.82)	**83.16(0.32)**
	L7	81.50(0.55)	81.55(0.59)	80.47(0.55)	81.12(0.32)	79.83(0.42)	79.78(0.46)	76.99(0.70)	82.34(0.31)	76.08(0.48)	80.99(0.75)	78.18(1.08)	**82.66(0.28)**
	L8	81.81(0.48)	81.50(0.55)	81.88(0.53)	82.28(0.28)	80.02(0.55)	79.96(0.56)	77.88(0.59)	83.58(0.29)	77.07(0.65)	82.25(0.80)	78.59(1.10)	**83.85(0.28)**
	L9	81.73(0.69)	81.74(0.55)	80.33(0.75)	81.20(0.30)	78.33(0.58)	78.76(0.49)	76.32(0.51)	82.36(0.22)	76.00(0.64)	80.57(0.64)	77.39(0.89)	**82.51(0.33)**
	L10	81.75(0.59)	81.69(0.53)	81.61(0.65)	82.36(0.39)	79.18(0.60)	80.00(0.59)	77.67(0.67)	83.56(0.33)	77.09(0.43)	81.36(0.66)	78.41(0.91)	**83.70(0.36)**
	L11	81.86(0.38)	81.67(0.39)	82.57(0.61)	83.12(0.29)	79.20(0.64)	79.61(0.43)	78.25(0.53)	84.34(0.27)	77.70(0.56)	82.47(0.74)	78.98(0.95)	**84.41(0.26)**
	L12	81.88(0.47)	81.90(0.39)	81.65(0.80)	82.40(0.41)	78.95(0.53)	78.68(0.46)	77.18(0.77)	84.07(0.34)	77.44(0.58)	81.60(0.92)	78.27(0.89)	**84.35(0.27)**
Average		81.57(0.58)	81.50(0.56)	80.85(0.63)	81.46(0.34)	78.30(0.57)	78.57(0.55)	76.76(0.66)	82.72(0.31)	76.34(0.56)	80.62(0.77)	77.50(0.90)	**82.90(0.30)**
p-vlaue		2.32E-03	3.05E-03	1.22E-04	1.22E-04	1.22E-04	1.22E-04	1.22E-04	3.66E-04	1.22E-04	1.22E-04	1.22E-04	–

Table 2.7: Conversion efficiency of all algorithms under WS3.

Turbine		AGA	SUGGA	SHADE	CJADE	ALGSA	HGSA	CLPSO	GLPSO	IWO	BSA	GWO	AGPSO
	L0	98.63(0.24)	98.97(0.18)	98.84(0.25)	98.65(0.19)	95.87(0.49)	95.99(0.64)	96.18(0.73)	99.26(0.07)	96.94(0.46)	98.35(0.53)	95.78(0.51)	**99.30(0.07)**
	L1	98.28(0.35)	98.54(0.27)	98.37(0.34)	98.02(0.30)	95.15(0.45)	95.12(0.63)	95.37(1.01)	98.95(0.14)	96.26(0.51)	98.08(0.59)	95.02(0.51)	**98.97(0.13)**
	L2	98.80(0.23)	99.11(0.12)	99.07(0.34)	99.05(0.12)	97.74(0.49)	97.36(0.41)	97.21(0.65)	99.27(0.10)	97.46(0.38)	98.47(0.56)	96.57(0.53)	**99.27(0.10)**
	L3	98.40(0.37)	98.82(0.12)	98.23(0.34)	97.93(0.25)	95.31(0.72)	94.88(0.57)	95.69(0.72)	98.75(0.16)	95.90(0.44)	97.14(0.63)	94.89(0.59)	**98.82(0.12)**
	L4	98.29(0.38)	98.74(0.25)	98.29(0.19)	97.95(0.20)	95.36(0.43)	95.08(0.50)	95.77(0.54)	98.77(0.14)	96.33(0.33)	96.78(0.64)	95.32(0.60)	**98.83(0.10)**
	L5	98.81(0.23)	99.09(0.14)	99.14(0.20)	99.05(0.15)	98.84(0.33)	97.96(0.45)	97.31(0.71)	99.29(0.07)	97.20(0.49)	98.80(0.56)	96.64(0.58)	**99.29(0.07)**
15	L6	98.08(0.28)	98.60(0.28)	98.25(0.32)	98.06(0.28)	94.36(0.56)	93.64(0.84)	94.94(0.63)	99.07(0.22)	96.34(0.58)	96.14(0.79)	94.79(0.41)	**99.16(0.18)**
	L7	98.57(0.30)	98.77(0.24)	98.60(0.42)	98.43(0.24)	95.57(0.52)	95.70(0.76)	96.09(0.78)	99.13(0.10)	96.62(0.48)	98.43(0.50)	95.42(0.60)	**99.15(0.08)**
	L8	98.74(0.29)	99.05(0.17)	99.01(0.27)	98.91(0.17)	97.01(0.57)	97.02(0.38)	96.89(0.49)	99.27(0.08)	97.26(0.40)	98.39(0.59)	96.11(0.53)	**99.27(0.08)**
	L9	98.54(0.36)	99.00(0.22)	98.55(0.22)	98.39(0.22)	95.68(0.64)	95.34(0.49)	96.02(0.96)	99.02(0.11)	96.52(0.36)	97.74(0.63)	95.38(0.57)	**99.07(0.09)**
	L10	98.48(0.30)	98.82(0.22)	98.54(0.24)	98.31(0.21)	95.81(0.53)	95.37(0.55)	95.87(0.54)	99.02(0.11)	96.69(0.45)	97.41(0.68)	95.46(0.40)	**99.06(0.08)**
	L11	98.73(0.24)	99.04(0.18)	99.02(0.25)	98.92(0.17)	98.34(0.40)	97.09(0.43)	96.95(0.73)	99.29(0.08)	97.02(0.38)	98.63(0.59)	96.18(0.48)	**99.31(0.06)**
	L12	98.16(0.34)	98.59(0.29)	98.41(0.29)	98.22(0.28)	94.98(0.62)	94.27(0.54)	95.05(0.67)	99.14(0.13)	96.43(0.49)	96.61(0.72)	95.27(0.55)	**99.19(0.14)**
Average		98.50(0.30)	98.87(0.22)	98.64(0.28)	98.45(0.21)	96.15(0.52)	95.76(0.55)	96.10(0.70)	99.09(0.11)	96.69(0.44)	97.77(0.62)	95.60(0.53)	**99.13(0.10)**
p-vlaue		1.22E-04	8.54E-04	1.22E-04	1.22E-04	1.22E-04	1.22E-04	1.22E-04	8.54E-04	1.22E-04	1.22E-04	1.22E-04	–
	L0	96.35(0.43)	97.11(0.42)	95.97(0.40)	95.61(0.35)	92.07(0.54)	91.65(0.75)	92.03(0.91)	98.17(0.45)	92.89(0.41)	95.47(0.91)	91.51(0.66)	**98.34(0.30)**
	L1	95.42(0.51)	95.99(0.51)	94.80(0.58)	94.40(0.47)	90.17(0.55)	90.33(0.75)	91.22(1.17)	97.62(0.57)	91.54(0.51)	94.62(1.14)	90.14(0.48)	**97.82(0.44)**
	L2	96.67(0.43)	97.34(0.43)	96.59(0.49)	96.31(0.32)	93.72(0.59)	93.64(0.60)	93.36(0.55)	98.28(0.35)	93.50(0.52)	95.51(0.94)	92.41(0.60)	**98.47(0.19)**
	L3	95.79(0.58)	96.95(0.51)	94.55(0.62)	94.24(0.37)	90.81(0.70)	90.58(0.61)	90.79(0.83)	97.08(0.54)	91.32(0.60)	93.53(0.84)	89.94(0.50)	**97.40(0.47)**
	L4	95.59(0.47)	96.28(0.68)	94.92(0.48)	94.56(0.25)	91.15(0.49)	90.71(0.50)	91.82(0.46)	97.42(0.44)	92.02(0.48)	92.90(0.71)	90.80(0.48)	**97.59(0.48)**
	L5	96.58(0.51)	97.33(0.48)	96.56(0.72)	96.25(0.30)	94.63(0.57)	94.29(0.52)	93.69(0.50)	98.33(0.28)	93.03(0.45)	96.12(0.95)	92.25(0.49)	**98.49(0.16)**
20	L6	95.26(0.55)	96.17(0.47)	94.42(0.47)	94.18(0.37)	89.64(0.61)	88.64(0.57)	90.38(0.75)	96.86(0.47)	91.46(0.44)	91.99(0.87)	89.85(0.50)	**97.04(0.61)**
	L7	95.96(0.48)	96.56(0.49)	95.41(0.56)	95.05(0.40)	91.35(0.54)	91.16(0.74)	92.07(0.87)	97.98(0.37)	92.20(0.41)	94.92(1.00)	90.82(0.57)	**98.13(0.30)**
	L8	96.54(0.40)	97.19(0.48)	96.28(0.53)	96.02(0.38)	92.93(0.51)	92.95(0.65)	93.10(0.79)	98.16(0.36)	93.17(0.52)	95.74(1.06)	91.89(0.50)	**98.38(0.30)**
	L9	96.12(0.42)	97.10(0.42)	95.36(0.49)	95.01(0.37)	91.42(0.68)	90.92(0.79)	91.79(0.78)	97.78(0.49)	92.14(0.43)	94.31(0.94)	90.71(0.57)	**98.05(0.39)**
	L10	95.92(0.44)	96.59(0.50)	95.28(0.56)	95.16(0.34)	91.60(0.59)	91.16(0.62)	92.20(0.68)	97.88(0.48)	92.44(0.46)	93.64(0.95)	91.23(0.57)	**98.10(0.37)**
	L11	96.50(0.46)	97.29(0.52)	96.28(0.52)	96.04(0.29)	94.10(0.58)	93.34(0.66)	93.39(0.68)	98.27(0.32)	92.96(0.46)	95.72(0.92)	91.93(0.51)	**98.36(0.33)**
	L12	95.17(0.35)	95.87(0.50)	94.70(0.57)	94.45(0.34)	90.18(0.54)	89.78(0.61)	90.45(0.77)	97.23(0.50)	92.04(0.56)	92.48(0.79)	90.53(0.66)	**97.39(0.43)**
Average		95.99(0.46)	96.75(0.49)	95.47(0.54)	95.18(0.35)	91.83(0.58)	91.47(0.64)	92.02(0.75)	97.77(0.43)	92.36(0.48)	94.38(0.93)	91.08(0.55)	**97.97(0.37)**
p-vlaue		1.22E-04	1.22E-04	1.22E-04	1.22E-04	1.22E-04	1.22E-04	1.22E-04	1.22E-04	1.22E-04	1.22E-04	1.22E-04	–
	L0	93.77(0.37)	94.57(0.48)	92.29(0.60)	91.79(0.28)	87.81(0.65)	87.69(0.87)	88.38(0.88)	96.08(0.91)	88.88(0.44)	91.97(1.09)	87.27(0.49)	**96.23(0.67)**
	L1	92.08(0.60)	92.80(0.35)	90.70(0.42)	90.22(0.40)	85.85(0.61)	85.80(0.80)	87.25(0.98)	95.10(0.96)	87.18(0.63)	90.37(1.15)	85.62(0.56)	**95.53(0.94)**
	L2	93.89(0.41)	94.57(0.50)	92.93(0.55)	92.66(0.42)	89.53(0.52)	89.52(0.64)	89.54(0.72)	95.92(0.82)	89.39(0.49)	91.98(1.07)	88.41(0.48)	**96.32(0.73)**
	L3	92.60(0.61)	93.92(0.62)	90.56(0.57)	89.99(0.34)	86.24(0.64)	86.04(0.41)	86.88(0.59)	94.19(0.77)	86.82(0.49)	89.25(1.00)	85.46(0.51)	**94.57(0.69)**
	L4	92.59(0.66)	93.13(0.54)	91.15(0.46)	90.79(0.29)	87.10(0.55)	86.80(0.52)	87.52(0.49)	94.14(0.57)	87.82(0.50)	88.89(0.83)	86.83(0.55)	**94.26(0.62)**
	L5	93.67(0.50)	94.62(0.40)	92.97(0.53)	92.50(0.35)	90.53(0.58)	90.92(0.56)	90.15(0.68)	95.97(0.83)	89.06(0.30)	92.58(1.23)	88.34(0.59)	**96.10(0.97)**
25	L6	92.14(0.65)	93.14(0.62)	90.27(0.52)	90.13(0.30)	84.93(0.59)	84.48(0.82)	85.66(0.65)	93.58(0.56)	87.15(0.60)	87.65(0.95)	85.38(0.58)	**93.73(0.60)**
	L7	93.09(0.54)	93.76(0.61)	91.52(0.62)	91.20(0.47)	86.79(0.45)	86.78(0.63)	87.35(0.82)	95.66(0.73)	87.97(0.50)	91.17(1.08)	86.48(0.51)	**95.84(0.82)**
	L8	93.82(0.44)	94.62(0.45)	92.63(0.59)	92.38(0.39)	88.83(0.55)	88.82(0.62)	89.30(0.71)	95.97(0.80)	89.04(0.59)	92.05(1.09)	87.82(0.54)	**96.39(0.79)**
	L9	93.26(0.60)	94.41(0.57)	91.35(0.69)	91.10(0.38)	87.34(0.67)	86.67(0.55)	87.85(0.78)	95.04(0.72)	87.92(0.54)	90.48(1.16)	86.46(0.57)	**95.45(0.82)**
	L10	93.22(0.58)	93.99(0.53)	91.60(0.49)	91.28(0.35)	87.43(0.54)	87.11(0.52)	88.13(0.75)	94.99(0.66)	88.41(0.61)	90.09(0.96)	86.89(0.57)	**95.41(0.67)**
	L11	93.67(0.50)	94.58(0.46)	92.66(0.76)	92.43(0.41)	89.88(0.69)	89.44(0.52)	89.39(0.56)	96.14(0.65)	89.01(0.46)	92.44(1.15)	87.94(0.63)	**96.26(0.69)**
	L12	91.96(0.60)	92.69(0.61)	90.67(0.51)	90.56(0.39)	86.00(0.63)	85.16(0.56)	86.21(0.68)	93.86(0.48)	87.53(0.51)	88.17(0.72)	86.07(0.52)	**94.02(0.50)**
Average		93.06(0.54)	93.91(0.52)	91.64(0.56)	91.31(0.37)	87.56(0.59)	87.33(0.62)	87.97(0.72)	95.13(0.73)	88.17(0.51)	90.54(1.04)	86.84(0.55)	**95.39(0.73)**
p-vlaue		1.22E-04	1.22E-04	1.22E-04	1.22E-04	1.22E-04	1.22E-04	1.22E-04	1.22E-04	1.22E-04	1.22E-04	1.22E-04	–

Table 2.8: Conversion efficiency of algorithms under WS4.

Turbine		AGA	SUGGA	SHADE	CJADE	ALGSA	HGSA	CLPSO	GLPSO	IWO	BSA	GWO	AGPSO
	L0	95.40(0.30)	95.48(0.29)	96.32(0.46)	96.56(0.23)	94.31(0.43)	94.60(0.43)	93.48(0.50)	97.51(0.38)	93.17(0.35)	95.50(0.52)	93.23(0.47)	**97.65(0.44)**
	L1	95.14(0.39)	95.48(0.42)	95.26(0.43)	95.56(0.35)	93.12(0.47)	93.39(0.52)	92.45(0.49)	96.47(0.60)	91.91(0.31)	94.57(0.66)	92.07(0.48)	**96.80(0.69)**
	L2	95.50(0.32)	95.44(0.29)	96.32(0.29)	96.49(0.21)	93.99(0.37)	94.42(0.36)	93.70(0.42)	97.15(0.25)	93.17(0.39)	95.38(0.63)	93.42(0.56)	**97.24(0.23)**
	L3	95.21(0.41)	95.46(0.40)	95.24(0.43)	95.56(0.29)	92.26(0.52)	92.88(0.41)	92.29(0.49)	96.62(0.67)	91.98(0.40)	93.18(0.61)	91.95(0.49)	**96.93(0.64)**
	L4	95.53(0.35)	95.53(0.32)	96.21(0.27)	96.31(0.26)	93.37(0.49)	93.80(0.44)	93.17(0.43)	97.15(0.28)	93.21(0.37)	94.32(0.53)	93.14(0.44)	**97.25(0.26)**
	L5	95.46(0.37)	95.35(0.30)	96.18(0.33)	96.46(0.26)	93.51(0.38)	93.86(0.43)	93.65(0.33)	97.12(0.31)	93.32(0.34)	94.66(0.57)	93.43(0.43)	**97.13(0.30)**
15	L6	95.21(0.35)	95.39(0.41)	95.69(0.40)	95.94(0.38)	91.74(0.53)	91.80(0.44)	91.97(0.38)	96.99(0.51)	92.32(0.44)	93.52(0.82)	92.03(0.37)	**97.25(0.52)**
	L7	95.53(0.41)	95.71(0.36)	95.79(0.49)	96.21(0.30)	93.92(0.56)	93.99(0.47)	93.13(0.65)	97.05(0.50)	92.66(0.28)	94.95(0.67)	92.78(0.49)	**97.38(0.55)**
	L8	95.46(0.31)	95.44(0.31)	96.30(0.32)	96.56(0.30)	94.27(0.43)	94.54(0.37)	93.61(0.45)	97.19(0.32)	93.19(0.36)	95.55(0.61)	93.35(0.47)	**97.43(0.34)**
	L9	95.46(0.36)	95.51(0.40)	95.85(0.37)	96.13(0.39)	93.13(0.44)	93.72(0.40)	92.75(0.34)	97.09(0.53)	92.77(0.42)	94.28(0.73)	92.68(0.60)	**97.32(0.43)**
	L10	95.48(0.28)	95.52(0.27)	96.16(0.39)	96.44(0.29)	93.70(0.48)	94.01(0.47)	93.60(0.67)	97.24(0.28)	93.10(0.28)	94.73(0.73)	93.17(0.55)	**97.41(0.31)**
	L11	95.45(0.31)	95.47(0.29)	96.26(0.34)	96.48(0.21)	93.73(0.38)	94.37(0.30)	93.59(0.45)	97.19(0.30)	93.22(0.32)	94.99(0.59)	93.46(0.47)	**97.44(0.29)**
	L12	95.49(0.33)	95.62(0.28)	96.07(0.43)	96.29(0.31)	93.39(0.69)	93.15(0.52)	92.65(0.49)	97.22(0.47)	92.72(0.38)	94.49(0.82)	92.70(0.62)	**97.47(0.39)**
Average		95.41(0.35)	95.49(0.33)	95.97(0.38)	96.23(0.29)	93.42(0.47)	93.73(0.43)	93.08(0.47)	97.08(0.42)	92.83(0.36)	94.62(0.65)	92.88(0.49)	**97.29(0.42)**
p-vlaue		1.22E-04	1.22E-04	1.22E-04	1.22E-04	1.22E-04	1.22E-04	1.22E-04	1.22E-04	1.22E-04	1.22E-04	1.22E-04	–
	L0	92.09(0.48)	92.13(0.41)	92.49(0.40)	92.93(0.45)	90.18(0.43)	90.37(0.46)	88.83(0.69)	94.76(0.66)	89.02(0.35)	91.73(0.89)	88.92(0.56)	**94.88(0.63)**
	L1	91.38(0.54)	91.68(0.45)	90.87(0.53)	91.30(0.40)	88.39(0.52)	88.69(0.59)	87.35(0.51)	93.09(0.48)	87.62(0.43)	90.25(0.65)	87.24(0.51)	**93.09(0.48)**
	L2	91.99(0.29)	92.02(0.39)	92.41(0.42)	92.93(0.39)	89.93(0.53)	90.00(0.39)	89.20(0.35)	94.20(0.39)	89.06(0.42)	91.81(0.73)	89.02(0.53)	**94.35(0.35)**
	L3	91.51(0.45)	91.75(0.41)	90.95(0.41)	91.29(0.33)	87.48(0.43)	87.99(0.37)	87.51(0.38)	93.12(0.51)	87.35(0.36)	88.40(0.51)	87.16(0.44)	**93.12(0.51)**
	L4	92.09(0.38)	92.08(0.31)	92.30(0.40)	92.79(0.29)	89.06(0.36)	89.54(0.38)	89.08(0.48)	94.23(0.34)	89.03(0.40)	90.27(0.67)	88.91(0.53)	**94.23(0.34)**
	L5	92.05(0.33)	91.96(0.27)	92.50(0.38)	92.93(0.25)	89.49(0.56)	89.71(0.35)	89.30(0.38)	94.14(0.35)	89.26(0.28)	90.88(0.64)	89.22(0.38)	**94.24(0.27)**
20	L6	91.51(0.45)	91.76(0.42)	91.20(0.49)	91.55(0.39)	87.03(0.61)	87.11(0.45)	86.90(0.51)	93.67(0.55)	87.58(0.35)	88.73(0.72)	87.19(0.52)	**93.67(0.55)**
	L7	91.92(0.44)	92.09(0.37)	91.67(0.45)	92.18(0.32)	89.45(0.45)	89.54(0.41)	88.34(0.45)	93.85(0.51)	88.29(0.29)	90.98(0.81)	88.01(0.47)	**94.19(0.75)**
	L8	92.02(0.33)	91.95(0.38)	92.38(0.31)	92.80(0.32)	90.00(0.38)	90.25(0.40)	89.31(0.40)	94.42(0.36)	89.05(0.36)	91.70(0.76)	88.90(0.52)	**94.67(0.35)**
	L9	91.79(0.42)	91.98(0.39)	91.64(0.33)	92.15(0.36)	88.66(0.44)	89.11(0.44)	87.96(0.48)	93.94(0.49)	88.39(0.34)	89.84(0.70)	87.99(0.44)	**94.12(0.68)**
	L10	92.02(0.43)	92.06(0.30)	92.38(0.35)	92.83(0.34)	89.33(0.39)	89.97(0.39)	89.00(0.45)	94.42(0.42)	89.03(0.34)	90.94(0.80)	88.93(0.59)	**94.49(0.46)**
	L11	92.02(0.34)	91.98(0.35)	92.45(0.41)	92.87(0.33)	89.79(0.53)	90.24(0.37)	89.53(0.48)	94.37(0.33)	89.22(0.35)	91.09(0.80)	89.22(0.51)	**94.50(0.34)**
	L12	91.98(0.34)	92.21(0.39)	91.95(0.41)	92.24(0.33)	88.57(0.53)	88.26(0.46)	87.72(0.60)	94.27(0.62)	88.37(0.46)	90.20(0.94)	88.10(0.58)	**94.46(0.54)**
Average		91.87(0.40)	91.97(0.37)	91.94(0.41)	92.37(0.35)	89.03(0.47)	89.29(0.42)	88.46(0.47)	94.05(0.49)	88.56(0.36)	90.52(0.74)	88.37(0.51)	**94.16(0.48)**
p-vlaue		1.22E-04	1.22E-04	1.22E-04	1.22E-04	1.22E-04	1.22E-04	1.22E-04	5.25E-03	1.22E-04	1.22E-04	1.22E-04	–
	L0	88.62(0.36)	88.70(0.33)	88.47(0.50)	89.02(0.35)	86.41(0.65)	86.31(0.50)	85.07(0.55)	91.41(0.52)	85.26(0.36)	88.36(0.79)	84.79(0.47)	**91.45(0.49)**
	L1	87.65(0.52)	88.07(0.47)	86.56(0.35)	87.20(0.40)	84.17(0.55)	84.24(0.50)	82.89(0.43)	89.27(0.52)	83.24(0.33)	86.21(0.88)	82.86(0.46)	**89.48(0.34)**
	L2	88.73(0.34)	88.66(0.34)	88.68(0.44)	89.22(0.26)	85.81(0.44)	85.79(0.34)	85.23(0.39)	90.96(0.41)	85.23(0.44)	87.78(0.83)	84.88(0.38)	**91.10(0.25)**
	L3	87.88(0.49)	88.29(0.52)	86.54(0.41)	87.22(0.34)	83.07(0.38)	83.76(0.47)	83.24(0.52)	89.34(0.56)	83.18(0.31)	84.34(0.70)	82.80(0.41)	**89.34(0.56)**
	L4	88.72(0.33)	88.68(0.30)	88.51(0.45)	89.05(0.29)	85.33(0.40)	85.52(0.28)	85.03(0.40)	90.96(0.31)	85.15(0.34)	86.53(0.64)	84.83(0.44)	**90.96(0.31)**
	L5	88.77(0.31)	88.75(0.38)	88.77(0.40)	89.31(0.32)	85.57(0.34)	85.91(0.28)	85.61(0.32)	91.19(0.43)	85.51(0.31)	87.34(0.73)	85.28(0.43)	**91.32(0.38)**
25	L6	87.78(0.55)	88.09(0.54)	86.68(0.53)	87.39(0.40)	82.43(0.68)	82.45(0.56)	82.37(0.55)	89.99(0.76)	83.22(0.35)	84.16(0.74)	82.62(0.56)	**90.23(0.68)**
	L7	88.39(0.46)	88.58(0.33)	87.63(0.54)	88.26(0.44)	85.34(0.57)	85.40(0.42)	84.02(0.52)	90.37(0.46)	84.31(0.31)	87.30(0.71)	83.94(0.56)	**90.59(0.41)**
	L8	88.63(0.28)	88.69(0.34)	88.52(0.38)	89.13(0.33)	86.11(0.47)	86.23(0.43)	85.01(0.58)	91.21(0.37)	85.21(0.32)	87.88(0.64)	84.83(0.42)	**91.26(0.39)**
	L9	88.27(0.43)	88.58(0.41)	87.61(0.38)	88.19(0.41)	84.39(0.35)	84.88(0.35)	84.17(0.51)	90.44(0.47)	84.37(0.36)	85.75(0.78)	83.96(0.55)	**90.62(0.45)**
	L10	88.71(0.33)	88.71(0.32)	88.54(0.43)	89.03(0.31)	85.54(0.41)	85.90(0.42)	85.04(0.42)	91.10(0.40)	85.13(0.39)	86.71(0.75)	84.84(0.52)	**91.23(0.43)**
	L11	88.71(0.23)	88.73(0.33)	88.64(0.41)	89.26(0.27)	85.79(0.40)	86.22(0.31)	85.22(0.26)	91.20(0.36)	85.48(0.38)	87.48(0.72)	85.04(0.43)	**91.20(0.36)**
	L12	88.57(0.37)	88.65(0.30)	87.85(0.48)	88.24(0.38)	84.47(0.87)	84.12(0.61)	83.64(0.48)	90.91(0.55)	84.18(0.44)	85.97(0.94)	83.68(0.50)	**91.01(0.55)**
Average		88.42(0.39)	88.55(0.38)	87.92(0.44)	88.50(0.35)	84.96(0.50)	85.13(0.42)	84.35(0.46)	90.66(0.46)	84.57(0.36)	86.60(0.76)	84.18(0.47)	**90.75(0.43)**
p-vlaue		1.22E-04	1.22E-04	1.22E-04	1.22E-04	1.22E-04	1.22E-04	1.22E-04	5.25E-03	1.22E-04	1.22E-04	1.22E-04	–

Bibliography

[1] Andrew W Senior, Richard Evans, John Jumper, James Kirkpatrick, Laurent Sifre, Tim Green, Chongli Qin, Augustin Žídek, Alexander WR Nelson, Alex Bridgland, et al. Improved protein structure prediction using potentials from deep learning. *Nature*, 577(7792):706–710, 2020.

[2] Friedrich Schotte, Manho Lim, Timothy A Jackson, Aleksandr V Smirnov, Jayashree Soman, John S Olson, George N Phillips, Michael Wulff, and Philip A Anfinrud. Watching a protein as it functions with 150-ps time-resolved X-ray crystallography. *Science*, 300(5627):1944–1947, 2003.

[3] Kurt Wuthrich. Protein structure determination in solution by nuclear magnetic resonance spectroscopy. *Science*, 243(4887):45–50, 1989.

[4] Rafael Fernandez-Leiro and Sjors HW Scheres. Unravelling biological macromolecules with cryo-electron microscopy. *Nature*, 537(7620):339–346, 2016.

[5] Sergey Ovchinnikov, Hahnbeom Park, Neha Varghese, Po-Ssu Huang, Georgios A Pavlopoulos, David E Kim, Hetunandan Kamisetty, Nikos C Kyrpides, and David Baker. Protein structure determination using metagenome sequence data. *Science*, 355(6322):294–298, 2017.

[6] Stephen F Altschul, Thomas L Madden, Alejandro A Schäffer, Jinghui Zhang, Zheng Zhang, Webb Miller, and David J Lipman. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research*, 25(17):3389–3402, 1997.

[7] Jianyi Yang, Renxiang Yan, Ambrish Roy, Dong Xu, Jonathan Poisson, and Yang Zhang. The I-TASSER suite: protein structure and function prediction. *Nature Methods*, 12(1):7–8, 2015.

[8] Benjamin Webb and Andrej Sali. Protein structure modeling with MODELLER. In *Protein Structure Prediction*, pages 1–15. Springer, 2014.

[9] Brian Kuhlman and Philip Bradley. Advances in protein structure prediction and design. *Nature Reviews Molecular Cell Biology*, 20(11):681–697, 2019.

[10] Glennie Helles. A comparative study of the reported performance of ab initio protein structure prediction algorithms. *Journal of the Royal Society Interface*, 5(21):387–396, 2008.

[11] Rebecca F Alford, Andrew Leaver-Fay, Jeliazko R Jeliazkov, Matthew J O'Meara, Frank P DiMaio, Hahnbeom Park, Maxim V Shapovalov, P Douglas Renfrew, Vikram K Mulligan, Kalli Kappel, et al. The Rosetta all-atom energy function for macromolecular modeling and design. *Journal of Chemical Theory and Computation*, 13(6):3031–3048, 2017.

[12] Dong Xu and Yang Zhang. Ab initio protein structure assembly using continuous structure fragments and optimized knowledge-based force field. *Proteins: Structure, Function, and Bioinformatics*, 80(7):1715–1735, 2012.

[13] Mirko Torrisi, Gianluca Pollastri, and Quan Le. Deep learning methods in protein structure prediction. *Computational and Structural Biotechnology Journal*, 18:1301–1310, 2020.

[14] Yang Li, Jun Hu, Chengxin Zhang, Dong-Jun Yu, and Yang Zhang. ResPRE: high-accuracy protein contact prediction by coupling precision matrix with deep residual neural networks. *Bioinformatics*, 35(22):4647–4655, 2019.

[15] John Jumper, Richard Evans, Alexander Pritzel, Tim Green, Michael Figurnov, Olaf Ronneberger, Kathryn Tunyasuvunakool, Russ Bates, Augustin Žídek, Anna Potapenko, et al. Highly accurate protein structure prediction with AlphaFold. *Nature*, 596(7873):583–589, 2021.

[16] Christian B Anfinsen. Principles that govern the folding of protein chains. *Science*, 181(4096):223–230, 1973.

[17] Ron Unger and John Moult. Finding the lowest free energy conformation of a protein is an

NP-hard problem: proof and implications. *Bulletin of Mathematical Biology*, 55(6):1183–1198, 1993.

[18] Zhenqin Li and Harold A Scheraga. Monte Carlo-minimization approach to the multiple-minima problem in protein folding. *Proceedings of the National Academy of Sciences*, 84(19):6611–6615, 1987.

[19] Yuji Sugita and Yuko Okamoto. Replica-exchange molecular dynamics method for protein folding. *Chemical Physics Letters*, 314(1-2):141–151, 1999.

[20] Jooyoung Lee, Peter L Freddolino, and Yang Zhang. Ab initio protein structure prediction. In *From Protein Structure to Function with Bioinformatics*, pages 3–35. Springer, 2017.

[21] Kaizhou Gao, Zhiguang Cao, Le Zhang, Zhenghua Chen, Yuyan Han, and Quanke Pan. A review on swarm intelligence and evolutionary algorithms for solving flexible job shop scheduling problems. *IEEE/CAA Journal of Automatica Sinica*, 6(4):904–916, 2019.

[22] Zhiming Lv, Linqing Wang, Zhongyang Han, Jun Zhao, and Wei Wang. Surrogate-assisted particle swarm optimization algorithm with pareto active learning for expensive multi-objective optimization. *IEEE/CAA Journal of Automatica Sinica*, 6(3):838–849, 2019.

[23] Di Wang, Ling Geng, Yu-Jun Zhao, Yang Yang, Yan Huang, Yang Zhang, and Hong-Bin Shen. Artificial intelligence-based multi-objective optimization protocol for protein structure refinement. *Bioinformatics*, 36(2):437–448, 2020.

[24] Vincenzo Cutello, Giuseppe Narzisi, and Giuseppe Nicosia. A multi-objective evolutionary approach to the protein structure prediction problem. *Journal of The Royal Society Interface*, 3(6):139–151, 2006.

[25] Sandra M Venske, Richard A Gonçalves, Elaine M Benelli, and Myriam R Delgado. ADEMO/D: An adaptive differential evolution for protein structure prediction problem. *Expert Systems with Applications*, 56:209–226, 2016.

[26] Shangce Gao, Shuangbao Song, Jiujun Cheng, Yuki Todo, and Mengchu Zhou. Incorporation of solvent effect into multi-objective evolutionary algorithm for improved protein

structure prediction. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 15(4):1365–1378, 2018.

[27] Shuangbao Song, Shangce Gao, Xingqian Chen, Dongbao Jia, Xiaoxiao Qian, and Yuki Todo. AIMOES: Archive information assisted multi-objective evolutionary strategy for ab initio protein structure prediction. *Knowledge-Based Systems*, 146:58–72, 2018.

[28] Shuangbao Song, Junkai Ji, Xingqian Chen, Shangce Gao, Zheng Tang, and Yuki Todo. Adoption of an improved PSO to explore a compound multi-objective energy function in protein structure prediction. *Applied Soft Computing*, 72:539–551, 2018.

[29] Guijun Zhang, Laifa Ma, Xiaoqi Wang, and Xiaogen Zhou. Secondary structure and contact guided differential evolution for protein structure prediction. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 2018.

[30] Yuedong Yang and Yaoqi Zhou. Ab initio folding of terminal segments with secondary structures reveals the fine difference between two closely related all-atom statistical energy functions. *Protein Science*, 17(7):1212–1219, 2008.

[31] Yong Duan and Peter A Kollman. Pathways to a protein folding intermediate observed in a 1-microsecond simulation in aqueous solution. *Science*, 282(5389):740–744, 1998.

[32] Bernard R Brooks, Charles L Brooks III, Alexander D Mackerell Jr, Lennart Nilsson, Robert J Petrella, Benoît Roux, Youngdo Won, Georgios Archontis, Christian Bartels, Stefan Boresch, et al. CHARMM: the biomolecular simulation program. *Journal of Computational Chemistry*, 30(10):1545–1614, 2009.

[33] Lukas D Schuler, Xavier Daura, and Wilfred F Van Gunsteren. An improved GROMOS96 force field for aliphatic hydrocarbons in the condensed phase. *Journal of Computational Chemistry*, 22(11):1205–1218, 2001.

[34] William L Jorgensen, David S Maxwell, and Julian Tirado-Rives. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *Journal of the American Chemical Society*, 118(45):11225–11236, 1996.

[35] Chao Zhang, George Vasmatzis, James L Cornette, and Charles DeLisi. Determination of atomic desolvation energies from the structures of crystallized proteins. *Journal of Molecular Biology*, 267(3):707–726, 1997.

[36] Jian Zhang and Yang Zhang. A novel side-chain orientation dependent potential derived from random-walk reference state for protein fold selection and structure prediction. *PloS One*, 5(10):e15386, 2010.

[37] Hui Lu and Jeffrey Skolnick. A distance-dependent atomic knowledge-based potential for improved protein structure selection. *Proteins: Structure, Function, and Bioinformatics*, 44(3):223–232, 2001.

[38] Ram Samudrala and John Moult. An all-atom distance-dependent conditional probability discriminatory function for protein structure prediction. *Journal of Molecular Biology*, 275(5):895–916, 1998.

[39] Hongyi Zhou and Yaoqi Zhou. Distance-scaled, finite ideal-gas reference state improves structure-derived potentials of mean force for structure selection and stability prediction. *Protein Science*, 11(11):2714–2726, 2002.

[40] Min-yi Shen and Andrej Sali. Statistical potential for assessment and prediction of protein structures. *Protein Science*, 15(11):2507–2524, 2006.

[41] Jianpeng Ma. Explicit orientation dependence in empirical potentials and its significance to side-chain modeling. *Accounts of Chemical Research*, 42(8):1087–1096, 2009.

[42] Kim T Simons, Charles Kooperberg, Enoch Huang, and David Baker. Assembly of protein tertiary structures from fragments with similar local sequences using simulated annealing and Bayesian scoring functions. *Journal of Molecular Biology*, 268(1):209–225, 1997.

[43] José C Calvo, Julio Ortega, and Mancia Anguita. PITAGORAS-PSP: Including domain knowledge in a multi-objective approach for protein structure prediction. *Neurocomputing*, 74(16):2675–2682, 2011.

[44] Ivan Getov, Marharyta Petukh, and Emil Alexov. SAAFEC: predicting the effect of single point mutations on protein folding free energy using a knowledge-modified MM/PBSA approach. *International Journal of Molecular Sciences*, 17(4):512, 2016.

[45] Sheng-You Huang and Xiaoqin Zou. An iterative knowledge-based scoring function for protein–protein recognition. *Proteins: Structure, Function, and Bioinformatics*, 72(2):557–579, 2008.

[46] João PGLM Rodrigues, Michael Levitt, and Gaurav Chopra. KoBaMIN: a knowledge-based minimization web server for protein structure refinement. *Nucleic Acids Research*, 40(W1):W323–W328, 2012.

[47] Vikas Nanda, Sandeep V Belure, and Ofer M Shir. Searching for the pareto frontier in multi-objective protein design. *Biophysical reviews*, 9(4):339–344, 2017.

[48] Xiao-Gen Zhou, Chun-Xiang Peng, Jun Liu, Yang Zhang, and Gui-Jun Zhang. Underestimation-assisted global-local cooperative differential evolution and the application to protein structure prediction. *IEEE Transactions on Evolutionary Computation*, 24(3):536–550, 2019.

[49] Md Kamrul Islam and Madhu Chetty. Clustered memetic algorithm with local heuristics for ab initio protein structure prediction. *IEEE Transactions on Evolutionary Computation*, 17(4):558–576, 2012.

[50] Peter J Fleming, Robin C Purshouse, and Robert J Lygoe. Many-objective optimization: An engineering design perspective. In *International Conference on Evolutionary Multi-criterion Optimization*, pages 14–32. Springer, 2005.

[51] Bingdong Li, Ke Tang, Jinlong Li, and Xin Yao. Stochastic ranking algorithm for many-objective optimization based on multiple indicators. *IEEE Transactions on Evolutionary Computation*, 20(6):924–938, 2016.

[52] Roland L Dunbrack Jr and Fred E Cohen. Bayesian statistical analysis of protein side-chain rotamer preferences. *Protein Science*, 6(8):1661–1681, 1997.

[53] Bernard Brooks and Martin Karplus. Harmonic dynamics of proteins: normal modes and fluctuations in bovine pancreatic trypsin inhibitor. *Proceedings of the National Academy of Sciences*, 80(21):6571–6575, 1983.

[54] Rhys Heffernan, Kuldip Paliwal, James Lyons, Abdollah Dehzangi, Alok Sharma, Jihua Wang, Abdul Sattar, Yuedong Yang, and Yaoqi Zhou. Improving prediction of secondary structure, local backbone angles and solvent accessible surface area of proteins by iterative deep learning. *Scientific Reports*, 5(1):1–11, 2015.

[55] Sidhartha Chaudhury, Sergey Lyskov, and Jeffrey J Gray. PyRosetta: a script-based interface for implementing molecular modeling algorithms using Rosetta. *Bioinformatics*, 26(5):689–691, 2010.

[56] Manfred J Sippl. Calculation of conformational ensembles from potentials of mena force: an approach to the knowledge-based prediction of local structures in globular proteins. *Journal of Molecular Biology*, 213(4):859–883, 1990.

[57] Gang Xu, Tianqi Ma, Tianwu Zang, Weitao Sun, Qinghua Wang, and Jianpeng Ma. OPUS-DOSP: a distance-and orientation-dependent all-atom potential derived from side-chain packing. *Journal of Molecular Biology*, 429(20):3113–3120, 2017.

[58] Kalyanmoy Deb, Amrit Pratap, Sameer Agarwal, and TAMT Meyarivan. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation*, 6(2):182–197, 2002.

[59] Eckart Zitzler, Marco Laumanns, and Lothar Thiele. SPEA2: Improving the strength pareto evolutionary algorithm. *TIK Report*, 103, 2001.

[60] Shouyong Jiang and Shengxiang Yang. A strength Pareto evolutionary algorithm based on reference direction for multiobjective and many-objective optimization. *IEEE Transactions on Evolutionary Computation*, 21(3):329–346, 2017.

[61] Hisao Ishibuchi, Yu Setoguchi, Hiroyuki Masuda, and Yusuke Nojima. Performance of decomposition-based many-objective algorithms strongly depends on Pareto front shapes. *IEEE Transactions on Evolutionary Computation*, 21(2):169–190, 2016.

[62] Siwei Jiang, Jie Zhang, Yew-Soon Ong, Allan N Zhang, and Puay Siew Tan. A simple and fast hypervolume indicator-based multiobjective evolutionary algorithm. *IEEE Transactions on Cybernetics*, 45(10):2202–2213, 2014.

[63] Ke Li, Kalyanmoy Deb, and Xin Yao. R-metric: Evaluating the performance of preference-based evolutionary multiobjective optimization using reference points. *IEEE Transactions on Evolutionary Computation*, 22(6):821–835, 2017.

[64] Salem F Adra and Peter J Fleming. Diversity management in evolutionary many-objective optimization. *IEEE Transactions on Evolutionary Computation*, 15(2):183–195, 2010.

[65] Hans Frauenfelder and Daan Thorn Leeson. The energy landscape in non-biological and biological molecules. *Nature Structural Biology*, 5(9):757–759, 1998.

[66] Jingfen Zhang and Dong Xu. Fast algorithm for population-based protein structural model analysis. *Proteomics*, 13(2):221–229, 2013.

[67] Liam J McGuffin, Kevin Bryson, and David T Jones. The PSIPRED protein structure prediction server. *Bioinformatics*, 16(4):404–405, 2000.

[68] Banu Soylu and Murat Koksalan. A favorable weight-based evolutionary algorithm for multiple criteria problems. *IEEE Transactions on Evolutionary Computation*, 14(2):191–205, 2009.

[69] Kalyanmoy Deb and Mayank Goyal. A combined genetic adaptive search (GeneAS) for engineering design. *Computer Science and Informatics*, 26:30–45, 1996.

[70] Kalyanmoy Deb and Ram Bhushan Agrawal. Simulated binary crossover for continuous search space. *Complex Systems*, 9(2):115–148, 1995.

[71] Jixiang Cheng, Gary G Yen, and Gexiang Zhang. A many-objective evolutionary algorithm with enhanced mating and environmental selections. *IEEE Transactions on Evolutionary Computation*, 19(4):592–605, 2015.

[72] Kalyanmoy Deb and Himanshu Jain. An evolutionary many-objective optimization algorithm using reference-point-based nondominated sorting approach, part i: solving problems with box constraints. *IEEE Transactions on Evolutionary Computation*, 18(4):577–601, 2013.

[73] Xingyi Zhang, Ye Tian, Ran Cheng, and Yaochu Jin. An efficient approach to nondominated sorting for evolutionary multiobjective optimization. *IEEE Transactions on Evolutionary Computation*, 19(2):201–213, 2014.

[74] Carlo R Raquel and Prospero C Naval Jr. An effective use of crowding distance in multiobjective particle swarm optimization. In *Proceedings of the 7th Annual Conference on Genetic and Evolutionary Computation*, pages 257–264, 2005.

[75] Wolfgang Kabsch. A solution for the best rotation to relate two sets of vectors. *Acta Crystallographica Section A: Crystal Physics, Diffraction, Theoretical and General Crystallography*, 32(5):922–923, 1976.

[76] Adam Zemla. LGA: a method for finding 3D similarities in protein structures. *Nucleic Acids Research*, 31(13):3370–3374, 2003.

[77] Yang Zhang and Jeffrey Skolnick. Scoring function for automated assessment of protein structure template quality. *Proteins: Structure, Function, and Bioinformatics*, 57(4):702–710, 2004.

[78] Joaquín Derrac, Salvador García, Daniel Molina, and Francisco Herrera. A practical tutorial on the use of nonparametric statistical tests as a methodology for comparing evolutionary and swarm intelligence algorithms. *Swarm and Evolutionary Computation*, 1(1):3–18, 2011.

[79] Vincenzo Cutello, Giuseppe Narzisi, and Giuseppe Nicosia. Computational studies of peptide and protein structure prediction problems via multiobjective evolutionary algorithms. In *Multiobjective Problem Solving from Nature*, pages 93–114. Springer, 2008.

[80] Robin C Purshouse and Peter J Fleming. Conflict, harmony, and independence: Relationships in evolutionary multi-criterion optimisation. In *International Conference on Evolutionary Multi-Criterion Optimization*, pages 16–30. Springer, 2003.

[81] Bruno Borguesan, Mariel Barbachan e Silva, Bruno Grisci, Mario Inostroza-Ponta, and Márcio Dorn. APL: An angle probability list to improve knowledge-based metaheuristics for the three-dimensional protein structure prediction. *Computational Biology and Chemistry*, 59:142–157, 2015.

[82] Xingqian Chen, Shuangbao Song, Junkai Ji, Zheng Tang, and Yuki Todo. Incorporating a multiobjective knowledge-based energy function into differential evolution for protein structure prediction. *Information Sciences*, 540:69–88, 2020.

[83] Yuedong Yang and Yaoqi Zhou. Specific interactions for ab initio folding of protein terminal regions with secondary structures. *Proteins: Structure, Function, and Bioinformatics*, 72(2):793–803, 2008.

[84] Ye Tian, Ran Cheng, Xingyi Zhang, Yansen Su, and Yaochu Jin. A strengthened dominance relation considering convergence and diversity for evolutionary many-objective optimization. *IEEE Transactions on Evolutionary Computation*, 23(2):331–345, 2018.

[85] Yanan Sun, Bing Xue, Mengjie Zhang, and Gary G Yen. A new two-stage evolutionary algorithm for many-objective optimization. *IEEE Transactions on Evolutionary Computation*, 23(5):748–761, 2018.

[86] Yanan Sun, Gary G Yen, and Zhang Yi. IGD indicator-based evolutionary algorithm for many-objective optimization problems. *IEEE Transactions on Evolutionary Computation*, 23(2):173–187, 2018.

[87] Samer Fawzy, Ahmed I Osman, John Doran, and David W Rooney. Strategies for mitigation of climate change: a review. *Environmental Chemistry Letters*, 18(6):2069–2094, 2020.

[88] Soraida Aguilar Vargas, Gheisa Roberta Telles Esteves, Paula Medina Maçaira, Bruno Quaresma Bastos, Fernando Luiz Cyrino Oliveira, and Reinaldo Castro Souza. Wind power generation: A review and a research agenda. *Journal of Cleaner Production*, 218:850–870, 2019.

[89] Temitope Raphael Ayodele, ASO Ogunjuyigbe, O Odigie, and Josiah L Munda. A multi-criteria gis based model for wind farm site selection using interval type-2 fuzzy analytic hierarchy process: The case study of nigeria. *Applied Energy*, 228:1853–1869, 2018.

[90] Brian Hand, Ger Kelly, and Andrew Cashman. Aerodynamic design and performance parameters of a lift-type vertical axis wind turbine: A comprehensive review. *Renewable and Sustainable Energy Reviews*, 139:110699, 2021.

[91] Ryan Nash, Reza Nouri, and Ahmad Vasel-Be-Hagh. Wind turbine wake control strategies: A review and concept proposal. *Energy Conversion and Management*, 245:114581, 2021.

[92] Xiaolei Liu, Zi Lin, and Ziming Feng. Short-term offshore wind speed forecast by seasonal arima-a comparison against gru and lstm. *Energy*, 227:120492, 2021.

[93] Sohail R Reddy. Wind farm layout optimization (WindFLO): An advanced framework for fast wind farm analysis and optimization. *Applied Energy*, 269:115090, 2020.

[94] Francisco González-Longatt, P Wall, and V Terzija. Wake effect in wind farm performance: Steady-state and dynamic behavior. *Renewable Energy*, 39(1):329–338, 2012.

[95] Niels Otto Jensen. *A note on wind generator interaction*, volume 2411. Citeseer, 1983.

[96] Xiaoxia Gao, Yue Li, Fei Zhao, and Haiying Sun. Comparisons of the accuracy of different wake models in wind farm layout optimization. *Energy Exploration & Exploitation*, 38(5):1725–1741, 2020.

[97] Giovanni Gualtieri. Comparative analysis and improvement of grid-based wind farm layout optimization. *Energy Conversion and Management*, 208:112593, 2020.

[98] Huan Long, Zijun Zhang, Zhe Song, and Andrew Kusiak. Formulation and analysis of grid and coordinate models for planning wind farm layouts. *IEEE Access*, 5:1810–1819, 2017.

[99] Xinglong Ju, Feng Liu, Li Wang, and Wei-Jen Lee. Wind farm layout optimization based on support vector regression guided genetic algorithm with consideration of participation among landowners. *Energy Conversion and Management*, 196:1267–1281, 2019.

[100] Fangyun Bai, Xinglong Ju, Shouyi Wang, Wenyong Zhou, and Feng Liu. Wind farm layout optimization using adaptive evolutionary algorithm with monte carlo tree search reinforcement learning. *Energy Conversion and Management*, 252:115047, 2022.

[101] Yirui Wang, Yang Yu, Shuyang Cao, Xingyi Zhang, and Shangce Gao. A review of applications of artificial intelligent algorithms in wind farms. *Artificial Intelligence Review*, 53(5):3447–3500, 2020.

[102] Yang Yu, Zhenyu Lei, Yirui Wang, Tengfei Zhang, Chen Peng, and Shangce Gao. Improving dendritic neuron model with dynamic scale-free network-based differential evolution. *IEEE/CAA Journal of Automatica Sinica*, 9(1):99–110, 2022.

[103] Yang Yu, Shangce Gao, Yirui Wang, and Yuki Todo. Global optimum-based search differential evolution. *IEEE/CAA Journal of Automatica Sinica*, 6(2):379–394, 2018.

[104] Yurui Wang, Shangce Gao, MengChu Zhou, and Yang Yu. A multi-layered gravitational search algorithm for function optimization and real-world problems. *IEEE/CAA Journal of Automatica Sinica*, 8(1):94–109, 2021.

[105] Shangce Gao, Kaiyu Wang, Sichen Tao, Ting Jin, Hongwei Dai, and Jiujun Cheng. A state-of-the-art differential evolution algorithm for parameter estimation of solar photovoltaic models. *Energy Conversion and Management*, 230:113784, 2021.

[106] Zhenyu Lei, Shangce Gao, Zhiming Zhang, MengChu Zhou, and Jiujun Cheng. MO4: A many-objective evolutionary algorithm for protein structure prediction. *IEEE Transactions on Evolutionary Computation*, 26(3):417–430, 2021.

[107] Xiao Yang, Zonghui Cai, Ting Jin, Zheng Tang, and Shangce Gao. A three-phase search approach with dynamic population size for solving the maximally diverse grouping problem. *European Journal of Operational Research*, 302(3):925–953, 2022.

[108] GPCDB Mosetti, Carlo Poloni, and Bruno Diviacco. Optimization of wind turbine positioning in large windfarms by means of a genetic algorithm. *Journal of Wind Engineering and Industrial Aerodynamics*, 51(1):105–116, 1994.

[109] Alireza Emami and Pirooz Noghreh. New approach on optimization in placement of wind turbines within wind farm by genetic algorithms. *Renewable Energy*, 35(7):1559–1564, 2010.

[110] Francisco M Gonzalez-Longatt, Peter Wall, Pawel Regulski, and Vladimir Terzija. Optimal electric network design for a large offshore wind farm based on a modified genetic algorithm approach. *IEEE Systems Journal*, 6(1):164–172, 2011.

[111] Xiang-Jun Zeng, Jin Tao, Ping Zhang, Hui Pan, and Yuan-Yuan Wang. Reactive power optimization of wind farm based on improved genetic algorithm. *Energy Procedia*, 14:1362–1367, 2012.

[112] Ying Chen, Hua Li, Kai Jin, and Qing Song. Wind farm layout optimization using genetic algorithm with different hub height wind turbines. *Energy Conversion and Management*, 70:56–65, 2013.

[113] K Chen, MX Song, and X Zhang. Binary-real coding genetic algorithm for wind turbine positioning in wind farm. *Journal of Renewable and Sustainable Energy*, 6(5):053115, 2014.

[114] Ying Chen, Hua Li, Bang He, Pengcheng Wang, and Kai Jin. Multi-objective genetic algorithm based innovative wind farm layout optimization method. *Energy Conversion and Management*, 105:1318–1327, 2015.

[115] Xiaoxia Gao, Hongxing Yang, and Lin Lu. Optimization of wind turbine layout position in a wind farm using a newly-developed two-dimensional wake model. *Applied Energy*, 174:192–200, 2016.

[116] Leandro Parada, Carlos Herrera, Paulo Flores, and Victor Parada. Wind farm layout optimization using a gaussian-based wake model. *Renewable Energy*, 107:531–541, 2017.

[117] Ali M Abdelsalam and MA El-Shorbagy. Optimization of wind turbines siting in a wind farm using genetic algorithm based local search. *Renewable Energy*, 123:748–755, 2018.

[118] Qingshan Yang, Jianxiao Hu, and Siu-Seong Law. Optimization of wind farm layout with modified genetic algorithm based on boolean code. *Journal of Wind Engineering and Industrial Aerodynamics*, 181:61–68, 2018.

[119] Laith Abualigah, Mohamed Abd Elaziz, Ahmad M Khasawneh, Mohammad Alshinwan, Rehab Ali Ibrahim, Mohammed AA Al-qaness, Seyedali Mirjalili, Putra Sumari, and Amir H Gandomi. Meta-heuristic optimization algorithms for solving real-world mechanical engineering design problems: a comprehensive survey, applications, comparative analysis, and results. *Neural Computing and Applications*, 34:4081–4110, 2022.

[120] Jun Tang, Gang Liu, and Qingtao Pan. A review on representative swarm intelligence algorithms for solving optimization problems: Applications and trends. *IEEE/CAA Journal of Automatica Sinica*, 8(10):1627–1643, 2021.

[121] Zhi-Hui Zhan, Lin Shi, Kay Chen Tan, and Jun Zhang. A survey on evolutionary computation for complex continuous optimization. *Artificial Intelligence Review*, 55(1):59–110, 2022.

[122] Mehmet Beşkirli, İsmail Koç, Hüseyin Haklı, and Halife Kodaz. A new optimization algorithm for solving wind turbine placement problem: Binary artificial algae algorithm. *Renewable Energy*, 121:301–308, 2018.

[123] Yong Wang, Hao Liu, Huan Long, Zijun Zhang, and Shengxiang Yang. Differential evolution with a new encoding mechanism for optimizing wind farm layout. *IEEE Transactions on Industrial Informatics*, 14(3):1040–1054, 2017.

[124] Huan Long, Peikun Li, and Wei Gu. A data-driven evolutionary algorithm for wind farm layout optimization. *Energy*, 208:118310, 2020.

[125] Ajit C Pillai, John Chick, Lars Johanning, Mahdi Khorasanchi, and Sami Barbouchi. Comparison of offshore wind farm layout optimization using a genetic algorithm and a particle swarm optimizer. In *International Conference on Offshore Mechanics and Arctic Engineering*, volume 6, page V006T09A033. American Society of Mechanical Engineers, 2016.

[126] Rabia Shakoor, Mohammad Yusri Hassan, Abdur Raheem, and Yuan-Kang Wu. Wake effect modeling: A review of wind farm layout optimization using Jensen's model. *Renewable and Sustainable Energy Reviews*, 58:1048–1059, 2016.

[127] Fernando D Bianchi, Hernan De Battista, and Ricardo J Mantz. *Wind turbine control systems: principles, modelling and gain scheduling design*, volume 19. Springer, 2007.

[128] I Katic, Jørgen Højstrup, and Niels Otto Jensen. A simple model for cluster efficiency. In *European wind energy association conference and exhibition*, volume 1, pages 407–410. A. Raguzzi Rome, Italy, 1986.

[129] Bryony L. Du Pont and Jonathan Cagan. An extended pattern search approach to wind farm layout optimization. *Journal of Mechanical Design*, 134(8):081002, 2012.

[130] Ahmed AA Esmin, Rodrigo A Coelho, and Stan Matwin. A review on particle swarm optimization algorithm and its variants to clustering high-dimensional data. *Artificial Intelligence Review*, 44(1):23–45, 2015.

[131] Yue-Jiao Gong, Jing-Jing Li, Yicong Zhou, Yun Li, Henry Shu-Hung Chung, Yu-Hui Shi, and Jun Zhang. Genetic learning particle swarm optimization. *IEEE Transactions on Cybernetics*, 46(10):2277–2290, 2015.

[132] Xinglong Ju and Feng Liu. Wind farm layout optimization using self-informed genetic algorithm with information guided exploitation. *Applied Energy*, 248:429–445, 2019.

[133] Ryoji Tanabe and Alex Fukunaga. Success-history based parameter adaptation for differential evolution. In *2013 IEEE Congress on Evolutionary Computation*, pages 71–78. IEEE, 2013.

[134] Shangce Gao, Yang Yu, Yirui Wang, Jiahai Wang, Jiujun Cheng, and MengChu Zhou. Chaotic local search-based differential evolution algorithms for optimization. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 51(6):3954–3967, 2019.

[135] Zhenyu Lei, Shangce Gao, Shubham Gupta, Jiujun Cheng, and Gang Yang. An aggregative learning gravitational search algorithm with self-adaptive gravitational constants. *Expert Systems with Applications*, 152:113396, 2020.

[136] Yirui Wang, Yang Yu, Shangce Gao, Haiyu Pan, and Gang Yang. A hierarchical gravitational search algorithm with an effective gravitational constant. *Swarm and Evolutionary Computation*, 46:118–139, 2019.

[137] Jing J Liang, A Kai Qin, Ponnuthurai N Suganthan, and S Baskar. Comprehensive learning

particle swarm optimizer for global optimization of multimodal functions. *IEEE Transactions on Evolutionary Computation*, 10(3):281–295, 2006.

[138] Xian-Bing Meng, Xiao Zhi Gao, Lihua Lu, Yu Liu, and Hengzhen Zhang. A new bio-inspired optimisation algorithm: Bird swarm algorithm. *Journal of Experimental & Theoretical Artificial Intelligence*, 28(4):673–687, 2016.

[139] Shaya Karimkashi and Ahmed A Kishk. Invasive weed optimization and its features in electromagnetics. *IEEE Transactions on Antennas and Propagation*, 58(4):1269–1278, 2010.

[140] Seyedali Mirjalili, Seyed Mohammad Mirjalili, and Andrew Lewis. Grey wolf optimizer. *Advances in Engineering Software*, 69:46–61, 2014.

[141] Ahmad Nickabadi, Mohammad Mehdi Ebadzadeh, and Reza Safabakhsh. A novel particle swarm optimization algorithm with adaptive inertia weight. *Applied Soft Computing*, 11(4):3658–3670, 2011.

[142] Edmund A Gehan. A generalized wilcoxon test for comparing arbitrarily singly-censored samples. *Biometrika*, 52(1-2):203–224, 1965.

[143] Jacinto Carrasco, Salvador García, MM Rueda, Swagatam Das, and Francisco Herrera. Recent trends in the use of statistical tests for comparing swarm and evolutionary computing algorithms: Practical guidelines and a critical review. *Swarm and Evolutionary Computation*, 54:100665, 2020.

[144] LJ Vermeer, Jens Nørkær Sørensen, and Antonio Crespo. Wind turbine wake aerodynamics. *Progress in Aerospace Sciences*, 39(6-7):467–510, 2003.

[145] Tawatchai Kunakote, Numchoak Sabangban, Sumit Kumar, Ghanshyam G Tejani, Natee Panagant, Nantiwat Pholdee, Sujin Bureerat, and Ali R Yildiz. Comparative performance of twelve metaheuristics for wind farm layout optimisation. *Archives of Computational Methods in Engineering*, 29(1):717–730, 2022.