A DEEP REINFORCEMENT LEARNING-ASSISTED OPERATOR SELECTION FOR LARGE-SCALE MULTI-OBJECTIVE OPTIMIZATION

Anonymous authors
Paper under double-blind review

ABSTRACT

Large-scale multi-objective optimization problems (LSMOPs) pose significant challenges to conventional multi-objective evolutionary algorithms due to the exponential search space from the high-dimensional decision variables. Recently, adaptive operator selection by deep reinforcement learning (DRL) has become a prominent trend for solving LSMOPs. However, existing methods rely too heavily on conventional evolutionary operators and neglect specialized operators designed for large-scale problems, severely jeopardizing their performance. To address this limitation, this paper proposes a multi-objective evolutionary algorithm with DRL-assisted operator selection for LSMOPs, termed DRLOS-LSMOEA. First, a heterogeneous operator pool is constructed by integrating four conventional evolutionary operators with four large-scale dedicated operators, which substantially enriches the action space of the DRL agent. Second, a compact state representation built upon convergence and diversity is employed, alongside an adaptively normalized reward function. This design captures evolutionary dynamics and alleviates the curse of dimensionality, enabling effective training of the DRL model. Furthermore, a dimension control mechanism is proposed to dynamically regulate the decision variables that conventional evolutionary operators can conduct on in the later stage of evolution, thereby adaptively transitioning the search focus from global exploration to local exploitation. Extensive experiments on the LSMOP and UF benchmark suites demonstrate that DRLOS-LSMOEA significantly outperforms six state-of-the-art baselines across various test instances. The source code is available at `https://github.com/XJFCD/DRLOS_LSMOEA`.

1 INTRODUCTION

Optimization problems, such as traffic scheduling(Guo et al., 2024b), power system optimization(Zheng et al., 2015), wireless sensor network deployment (Cao et al., 2018), and machine learning feature selection(Bidgoli & Rahnamayan, 2024), typically involve multiple conflicting objectives. Scaling these real-world applications dramatically increases the number of decision variables, leading to large-scale multi-objective optimization problems (LSMOPs). The exponentially increased search space presents significant challenges for conventional evolutionary algorithms to identify high-quality Pareto fronts (PF) within limited computational budgets(Zhou et al., 2024). Consequently, developing efficient and scalable methods for LSMOPs has become a crucial challenge in evolutionary computation.

To tackle this challenge, numerous large-scale multi-objective evolutionary algorithms (LSMOEAs) that employ divide-and-conquer-based variable grouping, dimensionality reduction via problem transformation, or novel search strategies (Tian et al., 2021) have been proposed. Although these methods alleviate computational overhead by modifying conventional operators or evolutionary workflows, they are heavily dependent on static mechanisms. Their failure to proactively adapt search strategies according to real-time evolutionary feedback (Case & Lehre, 2020) inherently constrains their optimization performance.

Recently, integrating deep reinforcement learning (DRL) into evolutionary computation to achieve adaptive optimization has become a prominent trend (Wang et al., 2024b). DRL can be primarily

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

used for parameter tuning (Guo et al., 2025; Zong et al., 2025), operator selection (Yi et al., 2023; Tian et al., 2023; Ming et al., 2024), and strategy adaptation (Guo et al., 2024a; Ma et al., 2025) in evolving search process. Existing DRL-assisted operator selection approaches for large-scale multi-objective optimization generally model the operator selection procedure as a state-action-reward pipeline. Within this framework, the DRL agent learns to select suitable operators from a predefined operator pool based on real-time evolutionary feedback (Cao & Wang, 2026; Gao et al., 2026). Nevertheless, a major drawback of such methods lies in their heavy dependence on conventional evolutionary operators, namely, crossover and mutation strategies adopted by algorithms such as GA (Deb et al., 2007) and DE (Li & Zhang, 2009). They overlook operators specially tailored for large-scale multi-objective optimization, which substantially degrades their performance when addressing LSMOPs.

To address this limitation, this paper proposes a multi-objective evolutionary algorithm with DRL-assisted operator selection tailored for LSMOPs, termed DRLOS-LSMOEA. To enrich the action space of the DRL agent, we integrate multiple operators dedicated to large-scale multi-objective optimization into the operator pool. Furthermore, we devise a compact state representation as well as an adaptively normalized reward function to stabilize the training of DRL. In addition, a dimension control mechanism is designed to regulate the decision variables that conventional evolutionary operators can conduct on. The main contributions of this paper are summarized as follows.

(1) A heterogeneous operator pool comprising four conventional evolutionary operators and four large-scale dedicated operators is constructed to alleviate the poor efficiency of conventional operators for large-scale multi-objective optimization. Specifically, the conventional evolutionary operators include the simulated binary crossover (SBX) (Deb et al., 2002), polynomial mutation (PM) (Deb et al., 2002), sequential execution of SBX and PM (GA) (Deb et al., 2007), and the differential mutation (DE) (Li & Zhang, 2009). Owing to their versatility, they serve to maintain population diversity. In contrast, the four novel large-scale dedicated operators wrap the full offspring generation pipelines from four representative LSMOEAs, i.e., LMOEA-DS (Qin et al., 2021), LMOCSO (Tian et al., 2019), LMOEA-LGCS (Wang et al., 2024a), and MOPSO (Coello & Lechuga, 2002). They deliver structural guidance and strengthen convergence performance, facilitating effective navigation across high-dimensional decision spaces. To the best of our knowledge, we are the first to treat the offspring-generation procedures of these classic LSMOEAs as independent operators and assemble them into one unified operator pool.

(2) A compact state representation and an adaptively normalized reward function are designed to accurately capture evolutionary dynamics while mitigating the curse of dimensionality. Specifically, we employ a compact state representation based on convergence and diversity, alongside a tailored reward function using z-score normalized improvements to provide stable and informative feedback. Such a design guarantees robust convergence of the DRL model throughout the optimization procedure.

(3) A dimension control mechanism (DCM) is designed to further balance global exploration and local exploitation. It dynamically regulates the decision variables that conventional operators can conduct on in later stages of evolution. By concentrating computational resources on specific variables, DCM transitions the search focus from global exploration to local exploitation, thereby improving optimization efficiency. Comprehensive experiments on the LSMOP (Cheng et al., 2017) and UF (Zhang et al., 2008) benchmark suites demonstrate that DRLOS-LSMOEA achieves markedly superior performance compared with six state-of-the-art baseline algorithms across diverse test instances.

2 RL-Assisted Operator Selection

In evolutionary computation, evolutionary operators perform very differently both across different problems and different evolution stages (Wang et al., 2025). No single operator works the best on all problems and evolutionary stages. Consequently, operator selection presents the critical research issue in designing effective MOEAs.

In recent years, reinforcement learning (RL) has emerged as a promising approach for operator selection in MOEAs. It learns the behavioral patterns of operators by utilizing a discrete Q-table that records the cumulative reward of each operator under each state (Sutton et al., 1998). In contrast, deep

reinforcement learning (DRL) approximates the value function via deep Q-networks, thus making it better suited for continuous and high-dimensional state spaces (Mnih et al., 2015). Consequently, DRL has become a mainstream paradigm in operator selection. By interacting with the evolutionary environment, DRL agents can learn optimal policies to maximize cumulative rewards, offering a dynamic alternative to static heuristics.

Early efforts focused on applying DRL to operator selection for MOPs. Tian et al. (2023) proposed MOEA/D-DQN, which employs a deep Q-network for individual-level operator selection by encoding parent decision variables and weight vectors as states and taking neighborhood fitness improvements as rewards. However, as the decision space scales up, the growing dimensionality of the state space greatly impedes DRL convergence within limited evolutionary generations. This paradigm was subsequently extended to constrained MOPs (CMOPs). DRLOS-EMCMO (Ming et al., 2024) incorporates convergence, diversity, and feasibility into the state space to adaptively activate candidate operators. Nevertheless, the operator pool is still restricted to conventional evolutionary operators, leading to a narrow action space that fails to cope with the diverse search requirements of large-scale optimization.

Recently, DRL has been explored for operator selection in LSMOPs. Cao & Wang (2026) proposed RLDS-LSMOEA, a dual-space evolution framework in which an RL mechanism dynamically selects search strategies based on problem landscapes and evolutionary stages. Gao et al. (2026) developed LSDMC, a cooperative multi-agent RL framework for dynamic modular configuration responding to optimization trajectories. Although these methods verify the potential of DRL in large-scale optimization, they predominantly work within homogeneous and narrow action spaces, that is, the pool of candidate operators available to the RL agent remains limited, which severely restricts DRL's overall capabilities.

3 Proposed Method

3.1 Main Framework of DRLOS-LSMOEA

As illustrated in Figure 1, the workflow of DRLOS-LSMOEA consists of three major processes: operator selection, operator execution, and model update. The primary task of the operator selection process is to select operators to be executed in the current generation. At the initial stage, operators are chosen randomly because the DQN network has not been trained. Following a period of experience replay and accumulation, the trained DQN network is utilized to make operator-selection decisions. The core task of the operator execution process is to apply the chosen operators. During early evolution, selected operators perform search across the entire decision space. In the late evolutionary stage, whenever the selected operator is conventional, its search is confined within a subspace so as to boost local optimization performance. Finally, the model update process carries out the update of states, reward signals, and the DQN network parameters.

3.2 DRL-Assisted Operator Selection

In deep reinforcement learning-assisted evolutionary optimization, state representation plays a critical role in enabling effective learning. An ideal state vector ought to be sufficiently expressive to capture population distribution and the evolutionary dynamics, while remaining compact and robust against stochastic noise. Nevertheless, constructing such states for LSMOPs remains challenging. The high dimensionality of the decision space and complex interdependencies among variables make directly observing the population state computationally intractable and yield low-quality information. Accordingly, the DRL agent demands a sophisticated abstraction mechanism to perceive evolutionary dynamics via low-dimensional yet highly informative features. In what follows, we elaborate on the design of our state, action, and reward functions specifically tailored for LSMOPs.

3.2.1 State

To characterize the population distribution and evolutionary status while filtering redundant noise, DRLOS-LSMOEA adopts population convergence and diversity as its primary state features. These two metrics reflect the progression of the search procedure and the coverage of the solution set, respectively.

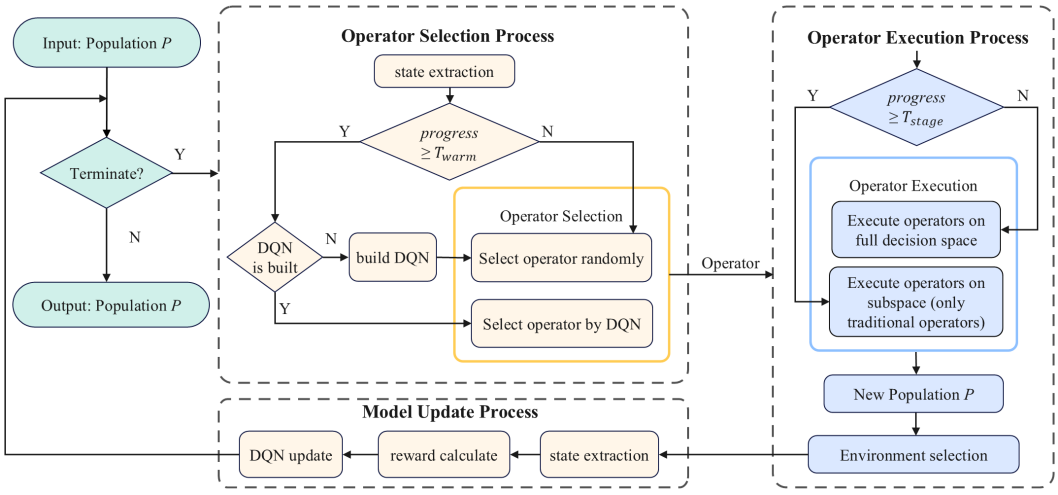


Figure 1: Framework of DRLOS-LSMOEA

The convergence feature ($conv$) is quantified using the average objective function values over the entire population. In high-dimensional fitness landscapes, the differences in objective values among individuals shrink substantially as the population converges. Meanwhile, disparate value scales across distinct objectives render raw objective values scale-dependent. To capture subtle performance gains effectively, we perform a logarithmic transformation on the average objective values. This transformation enhances sensitivity near the PF and compresses large-magnitude objective values far away from it. Mathematically, $conv$ is defined as:

$$conv = \ln\left(\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{M} f_j(\boldsymbol{x}_i) + \delta\right) \tag{1}$$

where $f_j(\boldsymbol{x}_i)$ denotes the j-th objective value of the i-th individual, and $\delta = 2^{-52}$ (i.e., the relative floating-point accuracy for double-precision arithmetic in MATLAB, approximately 2.22×10^{-16}) is a small constant introduced to avoid numerical instability. Since all objectives are to be minimized, a decreasing $conv$ value indicates progress toward the optimal PF.

The diversity metric $divs$ characterizes the distribution breadth of the population in the objective space. To achieve robustness against extreme outliers, it is defined as the sum of the inter-percentile ranges between the 10th and 90th percentiles over all M objectives:

$$divs = \sum_{j=1}^{M}[Q_j(90) - Q_j(10)] \tag{2}$$

where $Q_j(p)$ represents the p-th percentile of the j-th objective values $\{f_j(\boldsymbol{x}_i)\}_{i=1}^{N}$ for the entire population. By focusing on the central 80% of the distribution, this metric reliably captures the population spread while mitigating the impact of extreme outliers. A larger $divs$ value corresponds to higher population diversity. Combining these two metrics, the final state vector at generation t is formulated as:

$$\boldsymbol{s}_t = [conv_t, divs_t] \tag{3}$$

3.2.2 Action

Within our framework, the action space is defined as a set of candidate evolutionary operators:

$$\boldsymbol{O} = \{o_1, o_2, \ldots, o_k\} \tag{4}$$

where o_i denotes the i-th evolutionary operator and $k = 8$ denotes the total number of candidate operators. To balance search efficiency in large-scale optimization scenarios, the operator pool incorporates both conventional evolutionary and large-scale dedicated operators, as summarized in

Table S.1 of the Appendix. Conventional evolutionary operators (e.g., SBX, DE) produce isotropic perturbations that benefit population diversity and help avoid local optima, yet they may be affected by the curse of dimensionality.

In contrast, large-scale dedicated operators exploit problem structural properties or dimensionality-reduction techniques to prioritize convergence for LSMOPs, at the cost of potentially higher computational overhead and degraded population diversity. By integrating these operators with complementary behaviors into the unified action space $\boldsymbol{O}$, the DRL agent can dynamically adjust the search strategy in response to diverse problem characteristics and evolutionary stages.

3.2.3 REWARD

The reward function quantifies the instantaneous feedback obtained upon selecting a dedicated operator, steering the DRL agent toward optimal decisions. To stabilize training amid fluctuating reward magnitudes arising from different search stages and problem instances, we introduce an adaptive reward mechanism that combines sliding-window z-score normalization with a hyperbolic tangent function:

$$r_t = 0.7 \cdot \tanh\left(\frac{\Delta conv}{\sigma_c}\right) + 0.3 \cdot \tanh\left(\frac{\Delta divs}{\sigma_d}\right) \tag{5}$$

where $\Delta conv = conv_t - conv_{t+1}$ and $\Delta divs = divs_{t+1} - divs_t$ denote the improvements in convergence and diversity at generation t, respectively. σ_c and σ_d correspond to the standard deviations of $\Delta conv$ and $\Delta divs$ that are computed over a sliding window comprising recent generations. The $\tanh$ function compresses normalized values into the range $[-1, 1]$, thereby avoiding abrupt reward spikes. Assigning a higher weight of 0.7 to convergence while treating diversity as an auxiliary objective with weight 0.3, this adaptive reward yields a robust policy with strong generalization across varied fitness landscapes, mitigating overfitting toward specific problem scales. Empirical analyses of the reward distribution and the resulting operator selection probabilities are provided in Section C.3 of the Appendix.

3.3 DIMENSION CONTROL MECHANISM

Generally, evolutionary operators operating over the full decision space introduce isotropic perturbations, which confer substantial advantages in the early stages of LSMOEAs. Nevertheless, in later stages where the population converges toward promising regions, persistent perturbation across the entire search space incurs excessive stochastic noise, triggering population instability and convergence stagnation. This problem becomes exacerbated as the dimensionality of the decision space scales up. Moreover, conventional evolutionary operators (GA, DE, SBX, and PM) are inherently built upon neighborhood-based perturbation mechanisms. In high-dimensional spaces, however, neighborhood regions turn exceedingly large and sparse, which substantially impairs search performance.

To address this issue, we incorporate the dimension control mechanism (DCM), which enables the execution of conventional operators in two distinct modes. Specifically, the DCM uniformly partitions the decision variables into G mutually exclusive groups firstly. At the early stage of evolution, conventional operators are applied to all variables to promote global exploration. At the late stage, conventional operators are restricted to perform on subspaces formed by $K = \max(1, \text{round}(\alpha \cdot G))$ variable groups to guarantee the local exploitation. The α is the activation ratio and we set $\alpha = 0.3$ and $G = 6$ in this work. This strategy enables conventional operators to flexibly switch between global exploration and local exploitation according to the evolutionary stage, thereby effectively balancing the exploration–exploitation trade-off. The complete procedure of DCM is outlined in Algorithm 2 of the Appendix.

4 EXPERIMENTAL STUDIES

4.1 EXPERIMENTAL SETTINGS

Benchmark suites. To evaluate the performance of DRLOS-LSMOEA, we employ two widely used benchmark suites, LSMOP (Cheng et al., 2017) and UF (Zhang et al., 2008). For both suites, the number of objectives M is set to 2 and 3, while the decision variables D varies over 500, 1000, 2000, 5000, and 8000. Specifically, the LSMOP suite includes nine problems with challenging

properties such as non-uniform variable grouping, diverse correlations, and mixed separability. These characteristics demand algorithms to simultaneously converge to different PF regions. The UF suite comprises seven 2-objective (UF1-UF7) and three 3-objective (UF8-UF10) problems with complex interdependencies among variables. Both benchmarks present substantial optimization challenges for rigorously examining the scalability and efficiency of LSMOEAs.

Baselines. We compare DRLOS-LSMOEA against six state-of-the-art baselines, namely, DGEA (He et al., 2020), FDV (Yang et al., 2021), LERD (He et al., 2022), LMOCSO (Tian et al., 2019), LMOEA-DS (Qin et al., 2021), and MOEA/D-DQN (Tian et al., 2023). These competitors cover the major research paradigms for LSMOPs. Specifically, DGEA adopts state-guided directional search, FDV leverages fuzzy-based decision variable decomposition, LERD follows the divide-and-conquer grouping paradigm, LMOCSO corresponds to a population-driven competitive swarm optimizer, LMOEA-DS utilizes boundary-directed sampling, and MOEA/D-DQN incorporates DQN-guided operator selection.

Parameter Settings. All experiments are conducted on the PlatEMO platform (Tian et al., 2017). The population size N is set to 300, and the maximum function evaluations ($maxFE$) is fixed to 10^5 for all test instances. To guarantee a fair comparison, algorithm-specific parameters of the baseline algorithms follow the configurations reported in their original publications. For the DRL agent, a two-layer fully connected Q-network with 40 hidden units and ReLU activation is employed. The hyperparameters ε, γ, and replay buffer size are set to 0.95, 0.9, and 200, respectively. The batch size is set to 20 and the network is updated every 20 generations to reduce the computational overhead, while the warm-up phase occupies the first 20% of the evaluation budget and the stage transition takes place at 30% of the budget. Sensitivity analyses regarding the update frequency, batch size, warm-up threshold, stage transition threshold, group number G, and group activation ratio α are provided in Figs. S.3- S.7 and Tables S.11- S.22 of the Appendix, and are further elaborated in Section C.4. Parameters for individual evolutionary operators are summarized in Table S.2 of the Appendix.

Metrics. Inverted Generational Distance (IGD) (Zhou et al., 2006) and Hypervolume (HV) (While et al., 2011) are used for performance quantification. IGD computes the average distance from the true PF to the approximated solution set. 10000 reference points are uniformly sampled from the true PF. A lower IGD value reflects better overall performance in terms of both convergence and diversity. HV evaluates the volume of the objective space dominated by the obtained non-dominated solution set with respect to the reference point $\boldsymbol{r} = (1.1, 1.1, \ldots, 1.1)^\top$. A higher HV value indicates better convergence to and coverage of the PF.

4.2 Experimental Results and Discussion

Each algorithm is run 20 independent times for every test instance, and the Wilcoxon rank-sum test is adopted at a significance level of 0.05 to validate the statistical significance of performance differences. In the statistical tables, the symbols "+", "−", and "=" indicate that the compared algorithm achieves significantly better, worse, or comparable performance relative to the proposed DRLOS-LSMOEA, respectively.

4.2.1 Results on LSMOP

2-objective LSMOP. Table 1 summarizes the IGD results of all competing algorithms on the 2-objective LSMOP1-LSMOP9 test instances, and the corresponding HV results are listed in Table S.3 of the Appendix. As observed in Table 1, DRLOS-LSMOEA shows prominent competitiveness and significantly outperforms other algorithms on most LSMOP instances. Specifically, DRLOS-LSMOEA obtains statistically better IGD values on 44, 37, 44, 44, 41, and 40 out of 45 test instances when compared with DGEA, FDV, LERD, LMOCSO, LMOEA-DS, and MOEA/D-DQN, respectively. Consistent observations can be drawn from the HV results in Table S.3, where DRLOS-LSMOEA retains statistically significant superiority in the overwhelming majority of test cases. Specifically, it outperforms DGEA, LMOCSO, and LMOEA-DS on 35 out of 45 instances, and defeats LERD and MOEA/D-DQN on 34 instances, while defeating FDV on 29 instances. These results verify that the proposed framework delivers superior scalability and robustness with increasing decision-space dimensionality.

3-objective LSMOP. The IGD and HV values obtained by the competing algorithms on three-objective LSMOP problems are summarized in Tables S.4 and S.5 of the Appendix, respectively.

Table 1: IGD results of all algorithms on 2-objective LSMOPs

Problem	D	DGEA	FDV	LERD	LMOCSO	LMOEA-DS	MOEA/D-DQN	DRLOS-LSMOEA
LSMOP1	500	4.9360e-1 (2.00e-1) -	1.7042e-1 (1.64e-2) =	9.3451e-1 (9.42e-2) -	7.4641e-1 (1.00e-1) -	2.9937e-1 (1.09e-2) -	7.9753e-1 (1.13e-1) -	**1.6627e-1 (5.70e-2)**
	1000	4.8268e-1 (1.97e-1) -	2.8576e-1 (2.50e-2) -	1.4030e+0 (1.20e-1) -	1.4685e+0 (1.08e-1) -	3.1944e-1 (1.07e-2) -	1.7428e+0 (3.01e-1) -	**1.5907e-1 (5.22e-2)**
	2000	6.6022e-1 (3.01e-1) -	4.3554e-1 (1.22e-1) -	1.4864e+0 (1.22e-1) -	1.8360e+0 (1.24e-1) -	3.2848e-1 (9.72e-3) -	2.5955e+0 (2.64e-1) -	**1.6144e-1 (6.00e-2)**
	5000	6.0499e-1 (3.46e-1) -	6.2363e-1 (1.49e-1) -	1.6318e+0 (1.52e-1) -	2.0599e+0 (1.08e-1) -	3.3411e-1 (1.02e-2) -	3.2083e+0 (3.65e-1) -	**1.6605e-1 (5.49e-2)**
	8000	6.0001e-1 (3.46e-1) -	6.5197e-1 (2.02e-1) -	1.5144e+0 (1.78e-1) -	2.1225e+0 (1.01e-1) -	3.4039e-1 (1.02e-2) -	3.0340e+0 (4.11e-1) -	**1.5976e-1 (3.61e-2)**
LSMOP2	500	1.6367e-2 (2.91e-3) -	2.5033e-2 (6.74e-3) -	3.3202e-2 (8.57e-3) -	4.0788e-2 (3.92e-4) -	1.2580e-2 (3.36e-4) -	4.6141e-2 (1.06e-3) -	**4.3299e-3 (7.44e-4)**
	1000	1.0935e-2 (5.42e-3) -	1.6316e-2 (3.34e-3) -	1.6116e-2 (4.25e-3) -	2.2411e-2 (2.70e-4) -	7.0250e-3 (1.77e-4) -	2.5538e-2 (6.70e-4) -	**2.8453e-3 (3.62e-4)**
	2000	6.6145e-3 (2.37e-3) -	8.5474e-3 (2.54e-3) -	1.0841e-2 (5.89e-4) -	1.2193e-2 (2.25e-4) -	4.1669e-3 (1.42e-4) -	1.3718e-2 (3.39e-4) -	**2.0896e-3 (1.34e-4)**
	5000	4.5887e-3 (1.50e-3) -	4.5778e-3 (4.34e-4) -	5.3894e-3 (2.03e-4) -	5.7569e-3 (7.18e-5) -	2.4477e-3 (7.80e-5) -	6.4426e-3 (1.29e-4) -	**1.6625e-3 (9.73e-5)**
	8000	3.9532e-3 (1.21e-3) -	3.0854e-3 (3.83e-4) -	4.0379e-3 (1.64e-4) -	4.0837e-3 (4.46e-5) -	2.0683e-3 (4.88e-5) -	4.5582e-3 (1.60e-4) -	**1.5761e-3 (6.36e-5)**
LSMOP3	500	3.3431e+0 (6.13e+0) =	2.1460e+0 (2.35e+0) -	2.2358e+2 (6.67e+2) -	3.7658e+0 (3.39e+0) =	1.5662e+0 (6.75e-4) -	**1.0191e+0 (6.95e-2) +**	1.5627e+0 (4.20e-3)
	1000	3.4474e+0 (6.43e+0) -	**1.4535e+0 (9.58e-2) +**	1.2016e+1 (4.05e+0) -	6.2744e+0 (4.18e+0) -	1.5736e+0 (9.69e-4) -	3.3404e+0 (5.84e-1) -	1.5666e+0 (6.55e-3)
	2000	2.5445e+0 (2.50e+0) -	2.1302e+0 (6.62e-2) -	1.2821e+1 (1.50e+0) -	1.2391e+1 (3.06e+0) -	1.5770e+0 (9.29e-4) -	9.7648e+0 (1.16e+0) -	**1.5688e+0 (8.12e-3)**
	5000	3.9576e+0 (3.77e+0) -	4.0950e+0 (1.92e-1) -	1.3300e+1 (1.08e+0) -	1.9888e+1 (5.66e+0) -	1.5791e+0 (1.20e-3) -	1.9639e+1 (2.25e+0) -	**1.5733e+0 (4.67e-3)**
	8000	3.3988e+0 (3.73e+0) -	5.7332e+0 (2.03e-1) -	1.3617e+1 (1.10e+0) -	2.0268e+1 (2.60e+0) -	1.5801e+0 (1.39e-3) -	2.4666e+1 (1.22e+0) -	**1.5732e+0 (6.58e-3)**
LSMOP4	500	5.0974e-2 (4.64e-3) -	3.9679e-2 (2.05e-3) -	5.2335e-2 (5.39e-3) -	8.4057e-2 (1.27e-3) -	3.9447e-2 (5.72e-4) -	7.2805e-2 (4.03e-3) -	**2.1180e-2 (4.18e-3)**
	1000	2.6353e-2 (2.29e-3) -	2.5844e-2 (1.42e-3) -	3.0603e-2 (2.66e-3) -	5.0383e-2 (3.43e-4) -	2.1433e-2 (2.98e-4) -	4.9909e-2 (1.86e-3) -	**1.3549e-2 (1.60e-3)**
	2000	1.4600e-2 (9.72e-4) -	1.4368e-2 (4.50e-4) -	1.7127e-2 (1.09e-3) -	2.8753e-2 (9.52e-5) -	1.1469e-2 (1.76e-4) -	2.9800e-2 (6.45e-4) -	**7.3519e-3 (6.89e-4)**
	5000	6.5189e-3 (5.60e-4) -	6.4712e-3 (2.42e-4) -	7.9790e-3 (6.05e-4) -	1.3860e-2 (3.69e-5) -	5.2927e-3 (8.22e-5) -	1.4574e-2 (1.41e-4) -	**3.9211e-3 (2.89e-4)**
	8000	4.8143e-3 (6.48e-4) -	4.3457e-3 (1.51e-4) -	5.6378e-3 (3.62e-4) -	9.8427e-3 (2.77e-5) -	3.8115e-3 (7.00e-5) -	1.0325e-2 (1.08e-4) -	**2.8881e-3 (2.54e-4)**
LSMOP5	500	6.0159e+0 (6.85e-1) -	**2.1316e-1 (1.46e-1) +**	2.4266e+0 (4.84e-1) -	1.5713e+0 (1.92e-1) -	7.4202e-1 (8.67e-5) =	1.6471e+0 (2.92e-1) -	6.3878e-1 (1.49e-1)
	1000	6.4605e+0 (1.07e+0) -	1.0086e+0 (3.14e-1) -	4.1391e+0 (7.80e-1) -	3.6339e+0 (5.55e-1) -	7.4202e-1 (4.64e-5) -	3.7914e+0 (5.95e-1) -	**6.9313e-1 (1.29e-1)**
	2000	6.5999e+0 (1.29e+0) -	1.9430e+0 (5.37e-1) -	4.8974e+0 (8.43e-1) -	4.7399e+0 (5.06e-1) -	7.4203e-1 (5.97e-5) =	5.1749e+0 (7.36e-1) -	**6.8549e-1 (1.48e-1)**
	5000	5.7961e+0 (1.37e+0) -	1.9496e+0 (4.83e-1) -	4.8379e+0 (8.43e-1) -	5.1493e+0 (4.60e-1) -	7.4202e-1 (7.28e-5) -	5.9580e+0 (7.37e-1) -	**7.2992e-1 (3.59e-2)**
	8000	5.9026e+0 (1.61e+0) -	2.1118e+0 (5.15e-1) -	5.0079e+0 (5.46e-1) -	5.2287e+0 (5.44e-1) -	7.4199e-1 (1.60e-4) -	7.2466e+0 (1.41e+0) -	**6.8775e-1 (1.47e-1)**
LSMOP6	500	7.0058e-1 (1.24e-1) -	7.7547e-1 (1.37e-3) -	7.6208e-1 (2.67e-2) -	7.7333e-1 (4.15e-2) -	3.1780e-1 (2.05e-3) -	7.7673e-1 (1.40e-2) -	**2.5170e-1 (7.36e-2)**
	1000	6.4955e-1 (1.64e-1) -	7.5978e-1 (1.49e-4) -	7.1242e-1 (4.96e-2) -	7.5195e-1 (6.15e-2) -	3.1380e-1 (9.09e-3) -	7.5626e-1 (4.33e-2) -	**2.8013e-1 (4.36e-2)**
	2000	6.0773e-1 (2.01e-1) -	7.5062e-1 (2.45e-5) -	6.8669e-1 (9.77e-2) -	7.5565e-1 (8.39e-4) -	3.0891e-1 (7.60e-4) -	6.6540e-1 (1.43e-1) -	**2.5967e-1 (5.46e-2)**
	5000	6.8037e-1 (1.49e-1) -	7.4534e-1 (1.85e-6) -	7.4649e-1 (4.35e-4) -	7.4741e-1 (1.45e-4) -	3.0700e-1 (5.27e-5) -	6.0557e-1 (1.70e-1) -	**2.2744e-1 (6.50e-2)**
	8000	6.8743e-1 (1.41e-1) -	7.4409e-1 (4.90e-7) -	7.4481e-1 (1.82e-4) -	7.4540e-1 (1.51e-4) -	3.0691e-1 (1.38e-4) -	6.1918e-1 (1.76e-1) -	**2.4910e-1 (5.97e-2)**
LSMOP7	500	4.9501e+3 (9.22e+2) -	1.5123e+0 (1.29e-2) -	3.5476e+0 (3.47e-1) -	1.6407e+2 (5.71e+1) -	1.5042e+0 (1.68e-3) -	3.3328e+0 (5.57e-1) -	**1.5026e+0 (8.73e-4)**
	1000	6.3353e+3 (9.66e+2) -	1.5716e+0 (2.52e-1) -	5.8385e+0 (6.83e-1) -	1.8519e+3 (4.71e+2) -	1.5117e+0 (9.59e-4) -	5.0804e+0 (9.11e-1) -	**1.5107e+0 (4.03e-4)**
	2000	6.9549e+3 (2.59e+3) -	8.9716e+1 (3.23e+2) -	3.7203e+1 (2.05e+1) -	4.2289e+3 (1.08e+3) -	1.5147e+0 (5.53e-4) -	9.4661e+1 (1.09e+2) -	**1.5141e+0 (4.71e-4)**
	5000	6.7468e+3 (1.74e+3) -	2.6771e+2 (7.67e+2) -	1.1202e+3 (5.51e+2) -	6.2552e+3 (1.61e+3) -	1.5167e+0 (8.67e-4) -	2.4536e+3 (8.17e+2) -	**1.4879e+0 (1.26e-1)**
	8000	6.5252e+3 (2.69e+3) -	3.6240e+2 (6.10e+2) -	2.4494e+3 (8.98e+2) -	6.9700e+3 (1.95e+3) -	1.5169e+0 (6.08e-4) -	7.8034e+3 (3.25e+3) -	**1.5165e+0 (4.00e-4)**
LSMOP8	500	4.4207e+0 (6.89e-1) -	**6.7234e-2 (9.53e-3) +**	2.7894e-1 (1.01e-1) +	1.2392e+0 (1.58e-1) -	6.3850e-1 (7.39e-2) =	1.2437e-1 (8.24e-2) +	4.9898e-1 (2.10e-1)
	1000	4.4283e+0 (1.26e+0) -	**2.2539e-1 (3.46e-2) +**	1.1488e+0 (4.52e-1) -	2.7862e+0 (3.69e-1) -	7.4199e-1 (1.07e-4) -	4.6789e-1 (7.45e-3) =	5.1317e-1 (2.05e-1)
	2000	4.5045e+0 (1.66e+0) -	**4.9153e-1 (6.56e-2) =**	2.0407e+0 (3.27e-1) -	3.7426e+0 (4.37e-1) -	7.4201e-1 (8.15e-5) =	7.3174e-1 (1.09e-1) -	5.3692e-1 (2.11e-1)
	5000	4.4523e+0 (1.26e+0) -	1.1473e+0 (7.35e-1) -	2.1707e+0 (7.29e-1) -	4.1318e+0 (3.41e-1) -	7.4202e-1 (9.50e-5) -	2.7809e+0 (7.50e-1) -	**4.0872e-1 (2.13e-1)**
	8000	4.1461e+0 (1.93e+0) -	1.2497e+0 (6.95e-1) -	2.2437e+0 (4.13e-1) -	4.3293e+0 (4.29e-1) -	7.4202e-1 (6.46e-5) -	3.7312e+0 (7.32e-1) -	**5.0793e-1 (1.90e-1)**
LSMOP9	500	8.2806e+0 (2.18e+0) -	**1.3158e-1 (2.18e-2) +**	8.9059e-1 (5.42e-2) -	8.2586e-1 (1.79e-1) -	6.1361e-1 (1.39e-1) -	2.9780e-1 (5.51e-2) +	4.5978e-1 (1.92e-1)
	1000	1.0969e+1 (2.70e+0) -	**1.6852e-1 (9.46e-2) +**	1.0767e+0 (1.63e-1) -	3.0957e+0 (7.44e-1) -	5.8790e-1 (6.42e-2) -	5.5626e-1 (4.43e-2) =	4.7902e-1 (1.45e-1)
	2000	1.3233e+1 (3.28e+0) -	9.8722e-1 (5.91e-1) -	3.0275e+0 (1.65e+0) -	6.2327e+0 (1.22e+0) -	5.8999e-1 (3.18e-2) -	3.6596e+0 (1.31e+0) -	**4.7058e-1 (1.54e-1)**
	5000	1.3180e+1 (2.21e+0) -	3.8334e+0 (1.90e+0) -	5.5876e+0 (2.66e+0) -	8.2106e+0 (1.44e+0) -	5.6714e-1 (2.31e-2) -	1.0599e+1 (3.17e+0) -	**4.2479e-1 (1.71e-1)**
	8000	1.4627e+1 (3.60e+0) -	5.8659e+0 (1.87e+0) -	7.0571e+0 (2.07e+0) -	9.3638e+0 (1.88e+0) -	5.7208e-1 (5.44e-2) -	1.1613e+1 (2.60e+0) -	**4.0674e-1 (1.58e-1)**
+/-/=		0/44/1	6/37/2	1/44/0	0/44/1	0/41/4	3/40/2	

Compared to the 2-objective cases, 3-objective LSMOP optimization is considerably more challenging, which stems from the more intricate geometry of the PF and aggravated conflicts among objectives in higher-dimensional spaces. Even so, the proposed DRLOS-LSMOEA achieves competitive and stable performance on most test instances. For the IGD metric, DRLOS-LSMOEA yields significantly better performance than DGEA, FDV, LERD, LMOCSO, LMOEA-DS, and MOEA/D-DQN in 41, 37, 44, 42, 34, and 44 out of the 45 test instances, respectively. The HV results reported in Table S.5 further verify such statistical superiority over most peer baseline algorithms.

Pareto front and convergence trend. For visual comparison, Figure 2 visualizes the final Pareto fronts (a–g) and the average IGD convergence trajectories (h) of all algorithms on the 3-objective LSMOP8 instance with 1000 decision variables. As can be observed, DRLOS-LSMOEA achieves faster convergence and superior overall performance in terms of both convergence and diversity throughout the entire evolutionary process. Notably, it maintains a steady decline in IGD even during late evolutionary stages without premature stagnation. This benefit may arise from the dimension control mechanism, which mitigates stochastic noise in high-dimensional search landscapes by restricting the searching scope of conventional operators in local areas. In contrast, most baseline algorithms either become trapped in local optima (yielding solutions clustered within limited subregions) or suffer from slow convergence accompanied by substantial diversity degradation.

4.2.2 Results on UF

The IGD and HV comparisons on the UF suite are summarized in Tables S.6 and S.7 of the Appendix, respectively. As can be seen, the proposed DRLOS-LSMOEA maintains superiority over the other competing algorithms across most test instances. Specifically, in terms of IGD, DRLOS-LSMOEA achieves significantly better results on 50, 46, 45, 50, 42, and 50 out of 50 test instances when compared with DGEA, FDV, LERD, LMOCSO, LMOEA-DS, and MOEA/D-DQN, respectively. Furthermore, as listed in Table S.7, DRLOS-LSMOEA outperforms these peer algorithms on 40, 39, 37, 40, 39, and 40 instances regarding HV. These outcomes further validate its capability to effectively handle the intricate structural characteristics of the UF suite.

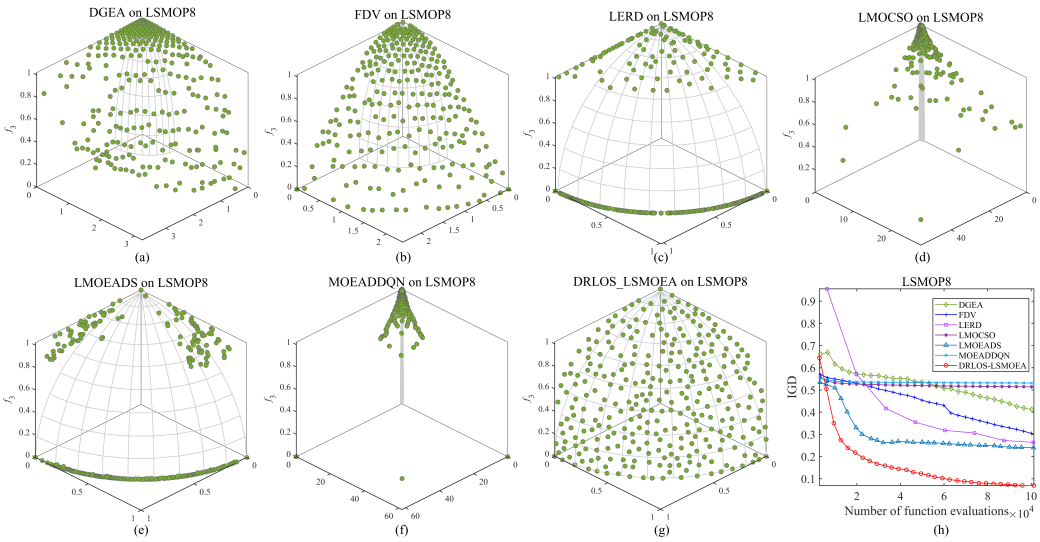


Figure 2: (a)–(g) Final solutions obtained by all algorithms on 3-objective LSMOP8 with $D = 1000$. (h) IGD convergence curves of all algorithms on 3-objective LSMOP8 with $D = 1000$

Table 2: Summary of IGD comparisons of all algorithms on LSMOP and UF benchmarks

Problem	M	DGEA	FDV	LERD	LMOCSO	LMOEA-DS	MOEA/D-DQN
LSMOP	2	0/44/1	6/37/2	1/44/0	0/44/1	0/41/4	3/40/2
	3	2/41/2	6/37/2	1/44/0	3/42/0	3/34/8	1/44/0
UF	2	0/35/0	3/31/1	0/35/0	0/35/0	0/35/0	0/35/0
	3	0/15/0	0/15/0	5/10/0	0/15/0	3/7/5	0/15/0
$+/-/=$		2/135/3	15/120/5	7/133/0	3/136/1	6/117/17	4/134/2

4.2.3 Overall Comparisons and Analyses

To comprehensively evaluate DRLOS-LSMOEA, Tables 2 and S.8 in the Appendix summarizes the overall statistical IGD and HV results across all 140 instances from the LSMOP and UF benchmarks, respectively. The row $(+/-/=)$ clearly shows that DRLOS-LSMOEA outperforms competing algorithms on the vast majority of test instances. Collectively, these quantitative and visual comparisons demonstrate its outstanding convergence, diversity, and robustness across diverse problems. Nevertheless, its reliance on a fixed operator pool and stochastic group masks slightly sacrifices search intensity on certain problem landscapes (e.g., LSMOP8 and LSMOP9).

4.3 Ablation Studies

This section conducts two ablation studies to systematically quantify the contributions of core components in DRLOS-LSMOEA. The first ablation experiment validates the necessity of the heterogeneous operator pool and the DCM by comparing DRLOS-LSMOEA with its two simplified variants. The second study investigates the influence of alternative grouping strategies embedded in the DCM. All experimental settings are identical with Section 4.1.

To verify the contributions of the heterogeneous operator pool and the DCM, two variants of DRLOS-LSMOEA are developed: DRLOS-LSMOEA-w/o LS and DRLOS-LSMOEA-w/o DCM. The former removes all large-scale dedicated operators from the operator pool while preserving the rest of the mechanisms, so as to evaluate the necessity of specialized high-dimensional search operators. By contrast, DRLOS-LSMOEA-w/o DCM disables the DCM module, which forces conventional evolutionary operators to operate over the full decision space during the entire evolutionary procedure.

Table 3 summarizes the IGD results of DRLOS-LSMOEA and its two variants. Detailed IGD values are provided in Tables S.23- S.28 of the Appendix. Specifically, DRLOS-LSMOEA-w/o LS suffers systematic performance degradation across all 140 instances (0/140/0), which confirms the indispensable role of large-scale dedicated operators. Likewise, DRLOS-LSMOEA-w/o DCM underperforms the DRLOS-LSMOEA on the vast majority of instances (4/87/49). This observation verifies that the DCM effectively mitigates random perturbations introduced by conventional evolutionary operators, and hence stabilizes both the evolutionary trajectory and the training of the DQN policy. Collectively, these empirical findings validate the individual effectiveness and necessity of the two core components.

Table 3: Summary of IGD comparisons on all test problem sets

Problem	DRLOS-LSMOEA-w/o LS (+/−/=)	DRLOS-LSMOEA-w/o DCM (+/−/=)
2-objective LSMOP	0/45/0	1/27/17
3-objective LSMOP	0/45/0	1/24/20
UF	0/50/0	2/36/12
ALL	0/140/0	4/87/49

To evaluate the impact of the variable grouping strategy within the DCM, we compare three grouping approaches, namely, random grouping, K-means clustering-based grouping, and our default uniform grouping. The random grouping scheme stochastically selects a subset of $\alpha \cdot D$ variables in each generation. For the K-means strategy, decision variables are clustered based on the Manhattan (cityblock) distance, and the partition is refreshed every 20 generations by re-clustering the current population. Upon clustering completion, variable groups are randomly chosen until the cumulative dimension reaches $\alpha \cdot D$. This ablation is conducted on the 2-objective LSMOPs with $D = 500$.

The resulting IGD and HV metrics are provided in Tables S.29- S.30 of the Appendix. As observed, uniform grouping achieves significantly better IGD values than random grouping on 6 out of 9 instances, with no cases of inferior performance. Meanwhile, K-means yields no statistically significant IGD improvement over uniform grouping across all test instances and performs substantially worse on 4 instances. The HV metrics consistently corroborate these observations.

5 Conclusions

This paper proposes DRLOS-LSMOEA, a deep reinforcement learning-assisted multi-objective evolutionary algorithm tailored for large-scale multi-objective optimization problems (LSMOPs). It constructs a heterogeneous operator pool that combines conventional and large-scale dedicated operators to substantially enrich the action space of the DRL agent, employs a compact state representation and an adaptive reward function to stabilize DRL training, and introduces a dimension control mechanism to smoothly transition from global exploration to local exploitation. These designs collectively enhance search efficiency and convergence in high-dimensional spaces.

Extensive experiments on the LSMOP and UF benchmarks demonstrate that DRLOS-LSMOEA significantly outperforms six state-of-the-art algorithms in both IGD and HV metrics, validating the effectiveness of its core components. Future work will focus on more adaptive operator pools to better respond to complex problem structures, incorporating variable correlation analysis into the grouping strategy to more precisely control the variable subsets on which different operators can act during different stages of evolution, and extending the framework to constrained, expensive, and multi-task optimization scenarios.

AI use statement

During the research, writing and preparation of this paper, no generative artificial intelligence tools were used.

REPRODUCIBILITY STATEMENT

The complete implementation of DRLOS-LSMOEA is described in Section 3, where the overall procedure and the dimension control mechanism (DCM) are summarized in Algorithm 1 and Algorithm 2 of the Appendix, respectively, and the composition and parameter configurations of the heterogeneous operator pool are specified in Tables S.1 and S.2 of the Appendix. The full experimental protocol—including the benchmark suites and tested problem dimensions, the parameter settings of DRLOS-LSMOEA and of the compared baselines, and the evaluation metrics—is presented in Section 4.1, while the number of independent runs per test instance and the statistical test used for the comparisons are described at the beginning of Section 4. The complete mean and standard deviation results, parameter sensitivity analyses (Section C.4), and ablation studies, are provided in the Appendix.

REFERENCES

Luis Miguel Antonio and Carlos A Coello Coello. Use of cooperative coevolution for solving large scale multiobjective optimization problems. In *2013 IEEE congress on evolutionary computation*, pp. 2758–2765. IEEE, 2013.

Azam Asilian Bidgoli and Shahryar Rahnamayan. Large-scale multi-objective feature selection: A multi-phase search space shrinking approach. *arXiv preprint arXiv:2410.21293*, 2024.

Bin Cao and Zhenyu Wang. Reinforcement learning guided dual-space evolutionary algorithm for large-scale multiobjective optimization. In *2026 6th International Symposium on Intelligent Robotics and Systems (ISoIRS)*, pp. 1–5. IEEE, 2026.

Bin Cao, Jianwei Zhao, Po Yang, Zhihan Lv, Xin Liu, Xinyuan Kang, Shan Yang, Kai Kang, and Amjad Anvari-Moghaddam. Distributed parallel cooperative coevolutionary multi-objective large-scale immune algorithm for deployment of wireless sensor networks. *Future Generation Computer Systems*, 82:256–267, May 2018.

Brendan Case and Per Kristian Lehre. Self-adaptation in nonelitist evolutionary algorithms on discrete problems with unknown structure. *IEEE Transactions on Evolutionary Computation*, 24 (4):650–663, 2020.

Ran Cheng, Yaochu Jin, Markus Olhofer, and Bernhard sendhoff. Test problems for large-scale multiobjective and many-objective optimization. *IEEE Transactions on Cybernetics*, 47(12): 4108–4121, December 2017.

CA Coello Coello and Maximino Salazar Lechuga. Mopso: A proposal for multiple objective particle swarm optimization. In *Proceedings of the 2002 Congress on Evolutionary Computation. CEC'02 (Cat. No. 02TH8600)*, volume 2, pp. 1051–1056. IEEE, 2002.

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, April 2002.

Kalyanmoy Deb, Karthik Sindhya, and Tatsuya Okabe. Self-adaptive simulated binary crossover for real-parameter optimization. In *Proceedings of the 9th annual conference on Genetic and evolutionary computation*, pp. 1187–1194. ACM, July 2007.

Kailai Gao, Cuie Yang, and Jinliang Ding. Multi-agent reinforcement learning-based dynamic modular configuration for large-scale multi-objective optimization. *IEEE Transactions on Emerging Topics in Computational Intelligence*, pp. 1–15, 2026.

Hongshu Guo, Yining Ma, Zeyuan Ma, Jiacheng Chen, Xinglin Zhang, Zhiguang Cao, Jun Zhang, and Yue-Jiao Gong. Deep reinforcement learning for dynamic algorithm selection: A proof-of-principle study on differential evolution. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 54 (7):4247–4259, July 2024a.

Hongshu Guo, Sijie Ma, Zechuan Huang, Yuzhi Hu, Zeyuan Ma, Xinglin Zhang, and Yue-Jiao Gong. Reinforcement learning-based self-adaptive differential evolution through automated landscape feature learning. In *Proceedings of the Genetic and Evolutionary Computation Conference*, pp. 1117–1126, 2025.

Tong Guo, Yi Mei, Ke Tang, and Wenbo Du. Cooperative co-evolution for large-scale multiobjective air traffic flow management. *IEEE Transactions on Evolutionary Computation*, 28(6):1644–1658, December 2024b.

Cheng He, Ran Cheng, and Danial Yazdani. Adaptive offspring generation for evolutionary large-scale multiobjective optimization. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 52(2):786–798, 2020.

Cheng He, Ran Cheng, Lianghao Li, Kay Chen Tan, and Yaochu Jin. Large-scale multiobjective optimization via reformulated decision variable analysis. *IEEE transactions on evolutionary computation*, 28(1):47–61, 2022.

Hui Li and Qingfu Zhang. Multiobjective optimization problems with complicated pareto sets, moea/d and nsga-ii. *IEEE Transactions on Evolutionary Computation*, 13(2):284–302, April 2009.

Ruochen Liu, Rui Ren, Jin Liu, and Jing Liu. A clustering and dimensionality reduction based evolutionary algorithm for large-scale multi-objective problems. *Applied Soft Computing*, 89: 106120, 2020.

Xiaoliang Ma, Xiaodong Li, Qingfu Zhang, Ke Tang, Zhengping Liang, Weixin Xie, and Zexuan Zhu. A survey on cooperative co-evolutionary algorithms. *IEEE Transactions on Evolutionary Computation*, 23(3):421–441, 2018.

Yuhang Ma, Bo Shen, and Anqi Pan. Competitive multitasking sampling with deep reinforcement learning for surrogate-assisted constrained multiobjective optimization. *IEEE Transactions on Evolutionary Computation*, pp. 1–1, 2025.

Fei Ming, Wenyin Gong, Ling Wang, and Yaochu Jin. Constrained multi-objective optimization with deep reinforcement learning assisted operator selection. *IEEE/CAA Journal of Automatica Sinica*, 11(4):919–931, April 2024.

Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Andrei A Rusu, Joel Veness, Marc G Bellemare, Alex Graves, Martin Riedmiller, Andreas K Fidjeland, Georg Ostrovski, et al. Human-level control through deep reinforcement learning. *nature*, 518(7540):529–533, 2015.

Douglas C Montgomery. *Design and analysis of experiments, Minitab manual*. John wiley & sons, 2010.

Mohammad Nabi Omidvar, Xiaodong Li, Yi Mei, and Xin Yao. Cooperative co-evolution with differential grouping for large scale optimization. *IEEE Transactions on evolutionary computation*, 18(3):378–393, 2013.

Mitchell A Potter and Kenneth A De Jong. A cooperative coevolutionary approach to function optimization. In *International conference on parallel problem solving from nature*, pp. 249–257. Springer, 1994.

Shufen Qin, Chaoli Sun, Yaochu Jin, Ying Tan, and Jonathan Fieldsend. Large-scale evolutionary multiobjective optimization assisted by directed sampling. *IEEE transactions on evolutionary computation*, 25(4):724–738, 2021.

Shuai Shao, Ye Tian, Yajie Zhang, and Xingyi Zhang. Knowledge learning-based dimensionality reduction for solving large-scale sparse multiobjective optimization problems. *IEEE Transactions on Cybernetics*, 55(7):3471–3484, 2025.

Richard S Sutton, Andrew G Barto, and Andrew Barto. *Reinforcement learning: An introduction*. MIT press Cambridge, 1998.

Ye Tian, Ran Cheng, Xingyi Zhang, and Yaochu Jin. Platemo: A matlab platform for evolutionary multi-objective optimization. *IEEE Computational Intelligence Magazine*, 12(4):73–87, 2017.

Ye Tian, Xiutao Zheng, Xingyi Zhang, and Yaochu Jin. Efficient large-scale multiobjective optimization based on a competitive swarm optimizer. *IEEE Transactions on Cybernetics*, 50(8): 3696–3708, 2019.

Ye Tian, Langchun Si, Xingyi Zhang, Ran Cheng, Cheng He, Kay Chen Tan, and Yaochu Jin. Evolutionary large-scale multi-objective optimization: A survey. *ACM Computing Surveys*, 54(8): 1–34, October 2021. ISSN 1557-7341.

Ye Tian, Xiaopeng Li, Haiping Ma, Xingyi Zhang, Kay Chen Tan, and Yaochu Jin. Deep reinforcement learning based adaptive operator selection for evolutionary multi-objective optimization. *IEEE Transactions on Emerging Topics in Computational Intelligence*, 7(4):1051–1064, August 2023.

Haofan Wang, Li Chen, Xingxing Hao, Rong Qu, Wei Zhou, Dekui Wang, and Wei Liu. Learning-guided cross-sampling for large-scale evolutionary multi-objective optimization. *Swarm and Evolutionary Computation*, 91:101763, 2024a.

Jia Wang, Yongfeng Li, Yeming Yang, Lingjie Li, Qiuzhen Lin, and Ling Wang. Multiobjective multitasking optimization with decomposition-based adaptive knowledge transfer. *IEEE Transactions on Evolutionary Computation*, pp. 1–1, 2025.

Xu Wang, Sen Wang, Xingxing Liang, Dawei Zhao, Jincai Huang, Xin Xu, Bin Dai, and Qiguang Miao. Deep reinforcement learning: A survey. *IEEE Transactions on Neural Networks and Learning Systems*, 35(4):5064–5078, April 2024b.

Lyndon While, Lucas Bradstreet, and Luigi Barone. A fast way of calculating exact hypervolumes. *IEEE Transactions on Evolutionary Computation*, 16(1):86–95, 2011. ISSN 1089-7782.

Xu Yang, Juan Zou, Shengxiang Yang, Jinhua Zheng, and Yuan Liu. A fuzzy decision variables framework for large-scale multiobjective optimization. *IEEE Transactions on Evolutionary Computation*, 27(3):445–459, 2021. ISSN 1089-7782.

Wenjie Yi, Rong Qu, Licheng Jiao, and Ben Niu. Automated design of metaheuristics using reinforcement learning within a novel general search framework. *IEEE Transactions on Evolutionary Computation*, 27(4):1072—1084, August 2023. ISSN 1089-778X, 1089-778X, 1941-0026.

Qingfu Zhang, Aimin Zhou, Shizheng Zhao, Ponnuthurai Nagaratnam Suganthan, Wudong Liu, Santosh Tiwari, et al. Multiobjective optimization test instances for the cec 2009 special session and competition. *University of Essex, Colchester, UK and Nanyang technological University, Singapore, special session on performance assessment of multi-objective optimization algorithms, technical report*, 264:1–30, 2008.

Xingyi Zhang, Ye Tian, Ran Cheng, and Yaochu Jin. A decision variable clustering-based evolutionary algorithm for large-scale many-objective optimization. *IEEE Transactions on evolutionary Computation*, 22(1):97–112, 2016. ISSN 1089-7782.

J.H. Zheng, J.J. Chen, Q.H. Wu, and Z.X. Jing. Multi-objective optimization and decision making for power dispatch of a large-scale integrated energy system with distributed dhcs embedded. *Applied Energy*, 154:369–379, September 2015.

Aimin Zhou, Yaochu Jin, Qingfu Zhang, Bernhard Sendhoff, and Edward Tsang. Combining model-based and genetics-based offspring generation for multi-objective optimization using a convergence criterion. In *2006 IEEE international conference on evolutionary computation*, pp. 892–899, 2006.

MengChu Zhou, Meiji Cui, Dian Xu, Shuwei Zhu, Ziyan Zhao, and Abdullah Abusorrah. Evolutionary optimization methods for high-dimensional expensive problems: A survey. *IEEE/CAA Journal of Automatica Sinica*, 11(5):1092–1105, 2024.

Heiner Zille, Hisao Ishibuchi, Sanaz Mostaghim, and Yusuke Nojima. Weighted optimization framework for large-scale multi-objective optimization. In *Proceedings of the 2016 on genetic and evolutionary computation conference companion*, pp. 83–84, 2016.

Eckart Zitzler, Marco Laumanns, and Lothar Thiele. Spea2: Improving the strength pareto evolutionary algorithm. *TIK report*, 103, 2001.

Mengjie Zong, Jingjing Li, Zhengjun Wang, and Wanrong Tan. An multi-agent reinforcement learning-based adaptive multi-objective evolutionary algorithm. In *2025 7th International Conference on Data-driven Optimization of Complex Systems (DOCS)*, pp. 184–189. IEEE, 2025.