Supplementary Materials

Evolutionary Heterogeneous Multitasking for Quality Diversity Optimization

I. Result and Analysis

A. Standard Deviations of Experimental Results Across Different Test Problems

Tables S.I, S.II, and S.III present the standard deviations of the experimental results for the HMQD-based algorithms and the corresponding single-task QD baseline algorithms over several different metrics in the CEC17M benchmark, manipulators, and higher-dimensional hexapod robot problems. The standard deviation values in Tables S.I, S.II, and S.III correspond to each entry in Tables IV, V, and XII of the main text, and are used to evaluate the performance stability of the algorithm in different tasks. It can be observed that HMQD-based algorithms exhibit low standard deviations in most tasks and metrics, indicating stable performance.

B. Comparative Analysis of Convergence Curves on Manipulator Problems

This section provides a comparison of the QD-score convergence curves for the manipulator problems, as shown in Fig. S1, aiming to illustrate the performance differences between QD algorithms and their corresponding HMQD-based variants across three groups of problems. The convergence trends on the two tasks in each group show that the HMQD framework could simultaneously solve both tasks more efficiently, resulting in a higher QD-score. Therefore, with the help of the HMQD framework, the HMQD-based variants can achieve significantly better performance than the original single-task solvers. Besides, the results indicate that when the dimensions of the source task and the target task are different, the knowledge transfer method designed is still effective.

C. Effectiveness of Knowledge Transfer Based on VAEs

The convergence curves of mean fitness obtained by CMA-ME, NVAE-CMA-ME, and HMQD-CMA-ME to solve benchmark problems are shown in Fig. S4. The HMQD-based algorithms outperform both the NVAE-based and the baseline single-task optimization algorithms on the benchmark problems, validating the effectiveness of the proposed VAE-based knowledge transfer strategy.

Moreover, the results reveal an intriguing observation that traditional implicit knowledge transfer fails to achieve competitive performance on the benchmark problems, even performing worse than the original CMA-ME algorithm. This may be attributed to two primary factors. Firstly, traditional implicit transfer does not prioritize the selection of elite solutions but instead applies crossover and mutation directly to the solution set, potentially introducing numerous invalid or even detrimental solutions, thereby degrading performance.

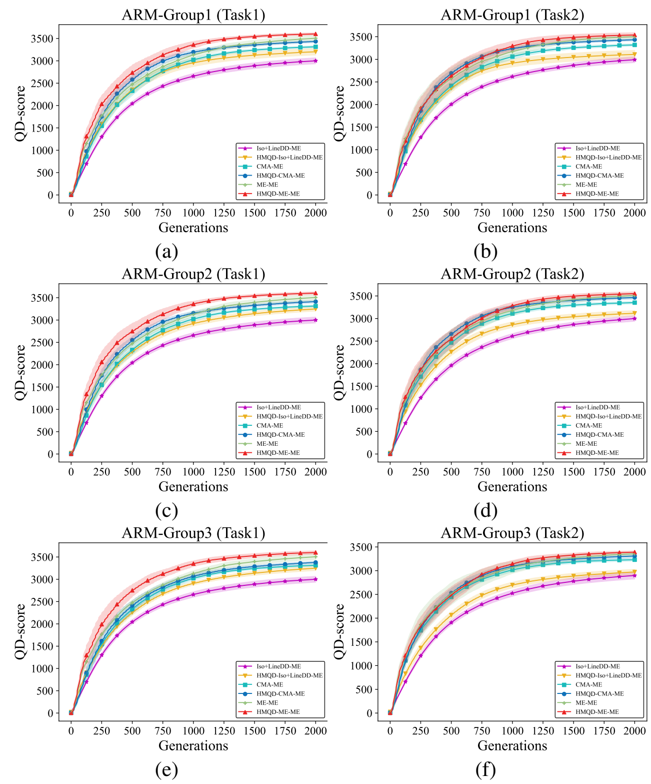

Fig. S1. Convergence curves on metric QD-score (QDS) of HMQD-Iso+LineDD-ME, HMQD-CMA-ME, HMQD-ME-ME, Iso+LineDD-ME, CMA-ME, and ME-ME algorithms in 30 repetitions of the manipulator problems. The shaded areas around the curves represent the standard deviation. (a), (b): G_1; (c), (d): G_2; (e), (f): G_3.

Secondly, it overlooks the deeper structural features or latent patterns of the solutions, limiting its effectiveness. In contrast, the proposed HMQD framework selectively filters high-quality solution sets and utilizes a VAE-based approach to learn the representations from them, effectively capturing useful latent features from the source task that are relevant to the target task. This enables a more accurate and effective knowledge transfer process.

Likewise, Fig. S3 shows the convergence curves of QD-score obtained by ME-ME, NVAE-ME-ME, and HMQD-ME-ME to solve manipulator problems. It can be observed that the HMQD framework significantly outperforms the NVAE method and the baseline of single-task optimization. Interestingly, in the manipulator experiments, NVAE exhibits a certain degree of effectiveness, possibly because when using QD algorithms to explore the reachable space of manipulators, the behavior space is dispersed, but all solutions in the archive tend to cluster into a specific region in the genotypic space. Therefore, if two manipulator tasks share a certain degree of

TABLE S.I

STANDARD DEVIATION OF EXPERIMENTAL RESULTS FOR THE HMQD-CVT-ME, HMQD-ISO+LINEDD-ME, HMQD-CMA-ME, CVT-ME, ISO+LINEDD-ME, AND CMA-ME ALGORITHMS ON THE CEC17M BENCHMARK TEST PROBLEMS OVER THE METRICS QDS, BEF, AND COV. THE VALUES IN THIS TABLE REPRESENT THE STANDARD DEVIATIONS CORRESPONDING TO EACH ENTRY IN TABLE IV OF THE MAIN TEXT.

Problem	Metrics	HMQD-CVT-ME		CVT-ME		HMQD-Iso+LineDD-ME		Iso+LineDD-ME		HMQD-CMA-ME		CMA-ME	
		Task1	Task2	Task1	Task2	Task1	Task2	Task1	Task2	Task1	Task2	Task1	Task2
CI+HS	QDS	1.33×10^2	1.37×10^5	2.30×10^3	2.19×10^6	2.11×10^2	2.19×10^5	5.75×10^2	8.42×10^5	8.70×10^1	1.08×10^5	8.91×10^1	7.61×10^4
	BEF	0.0227	31.1640	0.6353	566.8527	0.0196	29.7665	0.0488	84.6471	0.0154	25.7039	0.0481	59.0702
	COV	0.1795	0.0	0.1795	0.0	0.0	0.1795	0.0	0.1795	0.0	0.2494	0.0	0.0
CI+MS	QDS	3.36×10^2	1.13×10^5	2.30×10^3	4.13×10^5	4.21×10^2	1.83×10^5	6.54×10^2	5.31×10^5	3.34×10^2	1.12×10^5	1.69×10^2	8.40×10^4
	BEF	0.4502	30.5371	1.3106	149.9458	0.3946	29.3537	0.5289	58.3662	0.4726	17.6723	0.5457	41.8787
	COV	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1795	0.0	0.0	0.0
CI+LS	QDS	5.05×10^0	8.14×10^5	4.50×10^0	9.64×10^5	4.48×10^0	2.09×10^6	5.00×10^0	2.10×10^6	4.50×10^0	2.37×10^6	2.51×10^0	2.96×10^6
	BEF	0.0363	330.4463	0.0318	304.6215	0.0368	639.2448	0.0364	503.1734	0.0301	621.4849	0.0358	700.4650
	COV	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
PI+HS	QDS	3.25×10^5	3.09×10^6	2.04×10^6	8.44×10^6	3.06×10^5	3.86×10^6	4.73×10^5	4.97×10^6	2.29×10^5	3.82×10^5	8.77×10^4	2.58×10^5
	BEF	54.6964	603.4079	512.0042	2081.3399	36.8673	702.9388	58.7007	855.3005	81.8648	264.2997	57.0238	251.7909
	COV	0.1795	0.0	0.0	0.1795	0.1795	0.2494	0.0	0.0	0.3000	0.2494	0.1795	0.2494
PI+MS	QDS	3.68×10^2	1.62×10^{10}	3.12×10^3	1.96×10^{10}	4.70×10^2	1.61×10^{10}	7.60×10^2	1.55×10^{10}	3.37×10^2	1.56×10^{10}	2.10×10^2	1.76×10^{10}
	BEF	0.4726	2.15×10^4	1.0646	2.49×10^6	0.4271	7.31×10^3	0.4589	5.09×10^4	0.3228	9.26×10^3	0.4714	1.77×10^5
	COV	0.0	0.0	0.0	0.0	0.1795	0.0	0.0	0.0	0.0	0.0	0.0	0.0
PI+LS	QDS	7.72×10^2	2.56×10^3	6.08×10^3	1.83×10^3	2.23×10^3	3.47×10^3	1.82×10^3	3.47×10^3	4.56×10^3	2.49×10^3	1.10×10^4	2.87×10^3
	BEF	0.5264	0.7853	2.5693	0.6853	1.0983	1.1805	0.8923	1.2834	2.2982	1.0454	5.1008	0.9933
	COV	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
NI+HS	QDS	1.38×10^{10}	1.65×10^5	1.23×10^{11}	2.15×10^6	1.46×10^{10}	3.49×10^5	2.27×10^{10}	6.47×10^5	1.59×10^{10}	1.56×10^5	1.43×10^{10}	1.04×10^5
	BEF	3.77×10^4	34.6345	2.68×10^7	502.1091	2.11×10^4	29.1618	1.63×10^5	53.0772	4.42×10^4	44.4700	3.49×10^5	56.1058
	COV	0.0	0.1795	0.0	0.0	0.1795	0.2494	0.0	0.0	0.0	0.0	0.2494	0.1795
NI+MS	QDS	3.49×10^2	1.51×10^3	5.23×10^2	5.57×10^3	4.39×10^2	1.94×10^3	4.80×10^2	4.50×10^3	1.01×10^2	2.24×10^3	9.65×10^1	3.95×10^3
	BEF	0.0555	1.3478	0.1491	1.5258	0.0617	1.0934	0.0771	1.3248	0.0683	1.4650	0.0582	1.4285
	COV	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
NI+LS	QDS	1.01×10^5	1.07×10^6	4.88×10^6	1.15×10^6	2.50×10^5	2.03×10^6	2.69×10^5	2.30×10^6	8.96×10^4	8.96×10^6	8.56×10^4	2.68×10^6
	BEF	60.5430	362.5471	1510.9854	380.9641	29.3287	551.0663	48.3805	621.2823	36.2434	677.0162	41.9219	573.1455
	COV	0.0	0.0	0.7118	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

TABLE S.II

STANDARD DEVIATION OF EXPERIMENTAL RESULTS FOR THE HMQD-ISO+LINEDD-ME, HMQD-CMA-ME, HMQD-ME-ME, ISO+LINEDD-ME, CMA-ME, AND ME-ME ALGORITHMS ON REACHABLE SPACE EXPLORATION PROBLEMS OF MANIPULATORS OVER THE METRICS QDS, BEF, AND COV. THE VALUES IN THIS TABLE REPRESENT THE STANDARD DEVIATIONS CORRESPONDING TO EACH ENTRY IN TABLE V OF THE MAIN TEXT.

Group	Index	HMQD-Iso+LineDD-ME		Iso+LineDD-ME		HMQD-CMA-ME		CMA-ME		HMQD-ME-ME		ME-ME	
		Task1	Task2	Task1	Task2	Task1	Task2	Task1	Task2	Task1	Task2	Task1	Task2
G_1	QDS	61.0596	79.9025	65.6762	63.1356	21.5977	27.6176	23.9667	25.6431	25.2931	48.6991	60.1341	40.5376
	BEF	0.0062	0.0062	0.0108	0.0107	0.0037	0.0041	0.0058	0.0061	0.0032	0.0026	0.0082	0.0080
	COV	65.3982	90.6196	72.9937	81.1126	25.2110	29.4866	27.0471	26.1518	33.3341	60.1048	46.4360	43.1992
G_2	QDS	47.215	58.1030	65.6762	59.7201	28.6241	30.9255	23.9667	26.7926	27.7388	39.7812	60.1341	51.2098
	BEF	0.0077	0.0071	0.0108	0.0088	0.0028	0.0033	0.0058	0.0055	0.0008	0.0013	0.0082	0.0038
	COV	47.0619	57.7602	72.9937	64.1266	30.8257	31.0531	27.0471	34.7748	31.6806	46.7343	46.4360	58.4193
G_3	QDS	51.8567	55.1684	65.6762	72.8000	24.1441	36.0082	23.9667	27.0938	35.3340	26.3586	60.1341	34.2146
	BEF	0.0069	0.0064	0.0108	0.0097	0.0029	0.0036	0.0058	0.0053	0.0008	0.0015	0.0082	0.0088
	COV	49.7531	60.2377	72.9937	77.1601	24.5990	37.9256	27.0471	31.8552	39.4056	28.8129	46.4360	28.8038

TABLE S.III

STANDARD DEVIATION OF EXPERIMENTAL RESULTS FOR THE HMQD-ME-ME AND ME-ME ON HEXAPOD PROBLEMS OVER THE METRICS QDS, BEF, AND COV. THE VALUES IN THIS TABLE REPRESENT THE STANDARD DEVIATIONS CORRESPONDING TO EACH ENTRY IN TABLE XII OF THE MAIN TEXT.

Metrics	HMQD-ME-ME			ME-ME		
	Task1	Task2	Task3	Task1	Task2	Task3
QDS	1.2244	0.9511	2.1042	3.0200	2.9747	3.4383
BEF	0.0023	0.0028	0.0031	0.0033	0.0025	0.0038
COV	68.5601	50.7887	81.9481	46.0656	79.4754	36.2064

similarity, knowledge transfer may still be effective without additional processing in the later stages of the optimization process, as the solution space distribution becomes relatively simple.

D. Effectiveness of ARAKT Method

The convergence curves of mean fitness obtained by CMA-ME, AET-CMA-ME, GNT-CMA-ME, RDT-CMA-ME, and HMQD-CMA-ME to solve benchmark problems are shown in Fig. S4. Compared to AET, GNT, and RDT methods, HMQD effectively leverages the UCB1-based strategy to dynamically select transfer methods, ultimately achieving better convergence and performance, particularly in tasks with higher similarity (e.g., CI+HS). In CI+MS problems with relatively low similarity, it can be observed that HMQD does not perform so good in the early stages but surpasses AET, GNT, and RDT in the later stages. This phenomenon may be attributed to the early exploration phase of HMQD, during which the algorithm extensively evaluates the effectiveness of various transfer methods. Furthermore, the substantial differences between tasks likely contribute to reduced exploration efficiency in the initial stages, resulting in suboptimal optimization performance. However, as the optimization progresses, the accumulation of data and a clearer understanding of task-specific characteristics enable the adaptive mechanisms within the HMQD framework to identify and select more appropriate transfer methods with greater precision, thereby improving its performance in the later stages. Notably, as shown in Fig. S4(f), even when knowledge transfer methods appear less effective than the original CMA-ME algorithm in the low-similarity CI+LS task, the HMQD framework nevertheless achieves better performance compared to AET, GNT, and RDT.

Additionally, the QD-score curves obtained by ME-ME, AET-ME-ME, GNT-ME-ME, RDT-ME-ME, and HMQD-ME-ME to solve manipulator problems are shown in Fig. S5. As can be seen from the detail view in Fig. S5, in the manipulator experiment, HMQD-ME-ME achieved a faster convergence speed than all other variants, validating the effectiveness of the

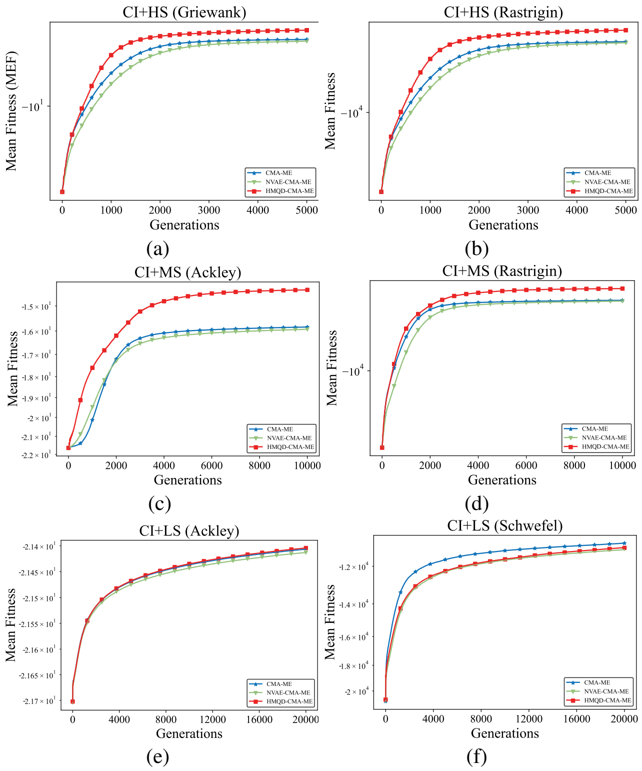

Fig. S2. Convergence curves of the HMQD-CMA-ME, NVAE-CMA-ME, and CMA-ME algorithms in 30 repetitions of the benchmark problems. (a), (b): CI+HS; (c), (d): CI+MS; (e), (f): CI+LS.

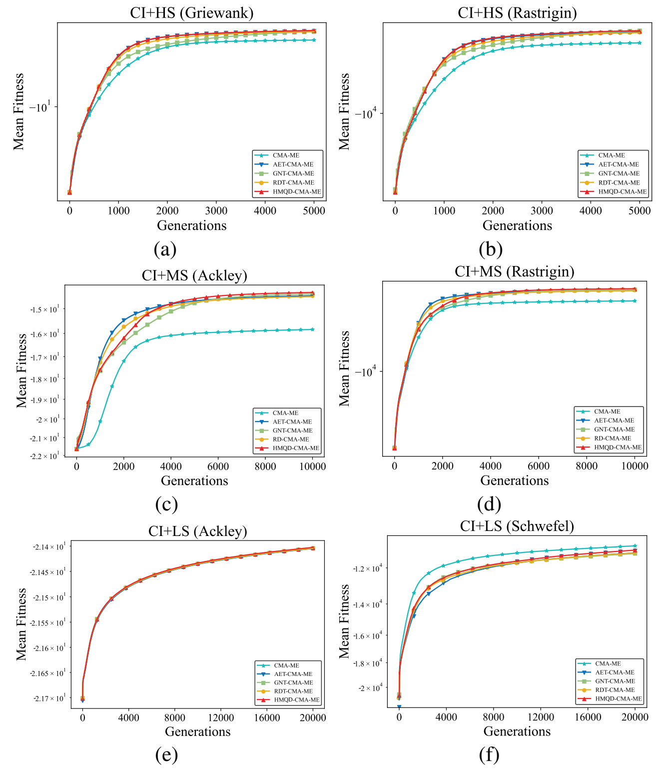

Fig. S4. Convergence curves of the HMQD-CMA-ME, AET-CMA-ME, GNT-CMA-ME, RDT-CMA-ME, and CMA-ME algorithms in 30 repetitions of the manipulator problems. (a), (b): CI+HS; (c), (d): CI+MS; (e), (f): CI+LS.

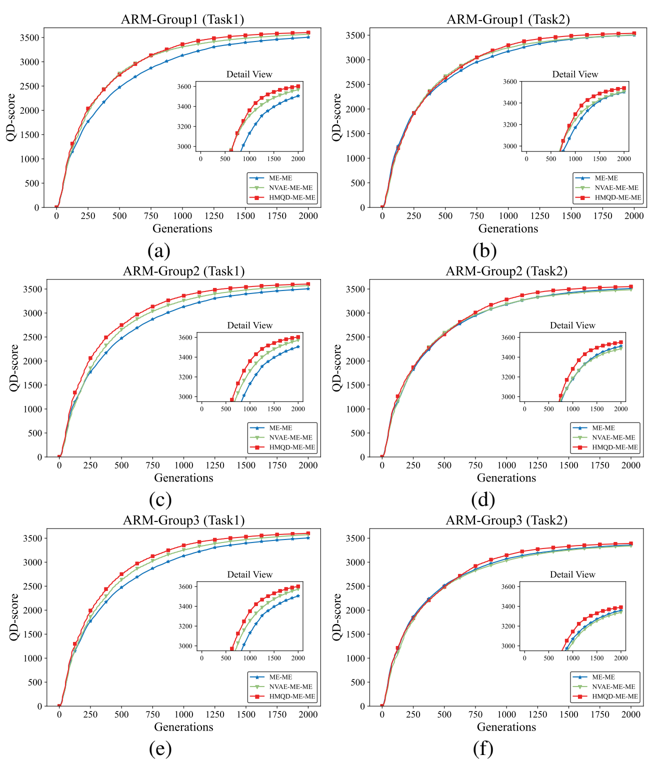

Fig. S3. Convergence curves of the HMQD-ME-ME, NVAE-ME-ME, and ME-ME algorithms in 30 repetitions of the manipulator problems. (a), (b): G_1; (c), (d): G_2; (e), (f): G_3.

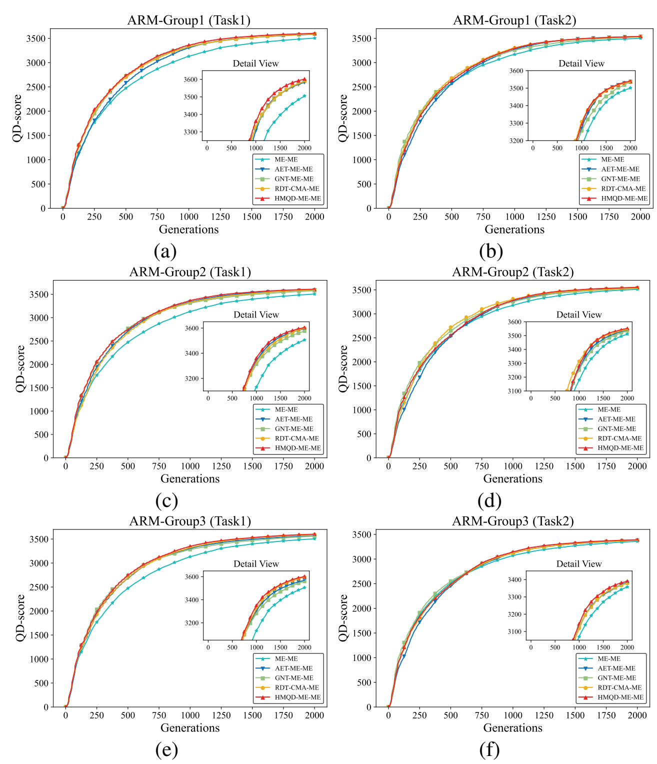

Fig. S5. Convergence curves of the HMQD-ME-ME, AET-ME-ME, GNT-ME-ME, RDT-ME-ME, and ME-ME algorithms in 30 repetitions of the manipulator problems. (a), (b): G_1; (c), (d): G_2; (e), (f): G_3.

adaptive knowledge transfer method proposed in the HMQD framework.

Another important consideration is that the GNT method can generate high-performance solutions representing the source task archive in the early stages of optimization, which is more conducive to assimilation by the target task. In contrast, the AET method randomly selects solutions from the target task as inputs to the VAE encoder, providing greater exploratory diversity. However, the preferred transfer method may vary across different convergence stages of the same task or between different tasks (i.e., some methods favor AET, while others are more suited to GNT). Therefore, the adaptive knowledge transfer strategy is crucial for enhancing the generalization ability of the proposed HMQD framework across various task scenarios.

E. Discussions on Alternative Variants for Knowledge Transfer

During the development of the HMQD framework, we have explored and preliminarily evaluated alternative approaches for knowledge transfer between tasks. This section details two alternative variants (i.e., AET-T and ESDT). While these variants show promise in certain aspects or provide valuable insights, our current AET method is found to be more consistently effective in our primary experimental setups. Nevertheless, we still agree that AET-T and ESDT remain valuable and interesting approaches for knowledge transfer. One of the strengths of the HMQD framework is its modularity, and AET-T and ESDT could certainly be integrated as an alternative transfer component within HMQD, potentially being more suitable for specific task pairings or scenarios. The descriptions of alternative variants and the experimental results are presented in the following subsections.

1) Target Oriented AutoEncoder-based Method (AET-T): The AET-T variant was conceptualized as a straightforward approach where a solution from the source task is directly processed using the VAE trained on the target task's archive. AET-T intuitively aims to adapt the source solution to the target task's learned representational space. Specifically, the AET-T transfer process begins by sampling a solution p from the source task's archive A_s. This sampled solution p is then directly reconstructed using the autoencoder (VAE) trained on the target task's archive, that is $x = M_t.decode(M_t.encode(p))$. Finally, this reconstructed solution x is evaluated on the target task. Similarly, when the dimensionality of the source task differs from that of the target task, the sampled source task solution's dimensionality is directly aligned with the target task's by truncation or concatenation.

To evaluate the effectiveness of AET-T relative to our proposed AET, we conduct experiments using the CMA-ME solver on benchmark problems with high and medium similarity (i.e., HS, MS), and the ME-ME solver on the manipulator problems. The comparative results are presented in Tables S.III and S.IV. We can observe that AET-T does not achieve better performance than the current AET on most tasks. We speculate that AET-T underperforms because directly transforming source task solutions into the new task

TABLE S.III
MEAN EXPERIMENTAL RESULTS OF THE AET-T-CMA-ME SOLVER AND AET-CMA-ME SOLVER ON CEC17M BENCHMARK TEST PROBLEMS OVER THE METRICS MEF, BEF, AND COV. BOLD INDICATES THE TOP PERFORMER IN THE PAIRWISE COMPARISON.

Problem	Metric	AET-T-CMA-ME		AET-CMA-ME	
		Task1	Task2	Task1	Task2
CI+HS	MEF	-2.8241×10^0 $(+)$	-2.3069×10^3 $(+)$	$\mathbf{-2.7921 \times 10^0}$	$\mathbf{-2.2689 \times 10^3}$
	BEF	-1.1600×10^0	-5.6659×10^2	$\mathbf{-1.0590 \times 10^0}$	$\mathbf{-4.5375 \times 10^2}$
	COV	4999.94	5000.00	5000.00	5000.00
CI+MS	MEF	-1.4731×10^1 $(+)$	-2.2490×10^3 $(+)$	$\mathbf{-1.4500 \times 10^1}$	$\mathbf{-2.2053 \times 10^3}$
	BEF	-6.9101×10^0	-5.2516×10^2	$\mathbf{-5.3488 \times 10^0}$	$\mathbf{-4.2587 \times 10^2}$
	COV	5000.00	5000.00	5000.00	5000.00
PI+HS	MEF	$\mathbf{-2.3815 \times 10^3}$ $(-)$	$\mathbf{-9.0399 \times 10^3}$ $(-)$	-2.8283×10^3	-9.4439×10^3
	BEF	$\mathbf{-5.6628 \times 10^2}$	-1.9958×10^3	-8.9620×10^2	$\mathbf{-1.8474 \times 10^3}$
	COV	4999.80	5000.00	4999.83	4999.83
PI+MS	MEF	-1.4948×10^1 $(+)$	-2.3694×10^8 $(+)$	$\mathbf{-1.4929 \times 10^1}$	$\mathbf{-2.3382 \times 10^8}$
	BEF	-7.7249×10^0	-5.6695×10^5	$\mathbf{-7.2177 \times 10^0}$	$\mathbf{-9.9831 \times 10^4}$
	COV	5000.00	5000.00	5000.00	5000.00
NI+HS	MEF	-2.4082×10^8 $(+)$	$\mathbf{-2.3364 \times 10^3}$ $(-)$	$\mathbf{-2.3774 \times 10^8}$	-2.4061×10^3
	BEF	-7.9603×10^5	-5.7576×10^2	$\mathbf{-7.0860 \times 10^5}$	$\mathbf{-5.5386 \times 10^2}$
	COV	5000.00	5000.00	5000.00	5000.00
NI+MS	MEF	$\mathbf{-3.0122 \times 10^0}$ $(-)$	-3.7794×10^0 $(+)$	-3.3063×10^0	$\mathbf{-3.7108 \times 10^1}$
	BEF	$\mathbf{-1.2650 \times 10^0}$	-2.4466×10^1	-1.4517×10^0	$\mathbf{-2.4183 \times 10^1}$
	COV	5000.00	5000.00	5000.00	5000.00

'$\approx$', '$+$', and '$-$' indicate that the comparison algorithm is similar to, significantly worse than, and significantly better than the AET-CMA-ME algorithm, respectively, based on Wilcoxon rank-sum tests. The significance level is α = 0.05.

space via the target VAE may lead to the loss of critical information or the introduction of unnecessary biases during the transfer process. This could make it difficult for the transferred solutions to fully adapt to the target task's specific demands. Therefore, AET-T has not been adopted in HMQD at this time.

TABLE S.IV
MEAN EXPERIMENTAL RESULTS OF THE AET-T-ME-ME SOLVER AND AET-ME-ME SOLVER ON REACHABLE SPACE EXPLORATION PROBLEMS OF MANIPULATORS OVER THE METRICS QDS, BEF, AND COV. BOLD INDICATES THE TOP PERFORMER IN THE PAIRWISE COMPARISON.

Group	Metric	AET-T-ME-ME		AET-ME-ME	
		Task1	Task2	Task1	Task2
G_1	QDS	3488.05 $(+)$	3519.61 $(+)$	**3589.96**	**3537.31**
	BEF	0.99496	0.99376	**0.99623**	**0.99547**
	COV	3761.43	**3756.53**	**3781.77**	3750.59
G_2	QDS	3509.69 $(+)$	**3539.67** $(\approx)$	**3594.00**	3538.64
	BEF	0.99468	0.99547	**0.99540**	**0.99548**
	COV	3776.80	**3815.47**	**3811.97**	3802.87
G_3	QDS	3534.68 $(+)$	3378.03 $(+)$	**3568.85**	**3382.81**
	BEF	0.99158	**0.99483**	**0.99243**	0.99447
	COV	3789.25	3695.30	**3802.17**	**3698.13**

'$\approx$', '$+$', and '$-$' indicate that the comparison algorithm is similar to, significantly worse than, and significantly better than the AET-ME-ME algorithm, respectively, based on Wilcoxon rank-sum tests. The significance level is α = 0.05.

2) Encoder-Source Decoder-Target Transfer Method (ESDT): The ESDT variant is designed as another strategy to handle knowledge transfer, particularly when addressing tasks with differing solution dimensionalities, by explicitly using components from both source and target VAEs. This method leverages the shared latent space z as an intermediary for transfer. Specifically, the ESDT transfer process begins by sampling a solution p from the source task's archive A_s. This solution p is then encoded into a shared latent space vector z using the source task's encoder, $M_s.encode(p)$. Subsequently, this latent vector z is decoded into a new solution x tailored for the target task t by employing the target task's decoder

$M_t.decode(z)$. Finally, the transferred solution x is evaluated on the target task.

TABLE S.V
MEAN EXPERIMENTAL RESULTS OF THE ESDT-CMA-ME SOLVER AND AET-CMA-ME SOLVER ON CEC17M BENCHMARK TEST PROBLEMS OVER THE METRICS MEF, BEF, AND COV. BOLD INDICATES THE TOP PERFORMER IN THE PAIRWISE COMPARISON.

Problem	Metric	ESDT-CMA-ME		AET-CMA-ME	
		Task1	Task2	Task1	Task2
CI+HS	MEF	-2.8370×10^{0}	-2.3390×10^{3}	$\mathbf{-2.7921 \times 10^{0}}$	$\mathbf{-2.2689 \times 10^{3}}$
	BEF	-1.1271×10^{0}	-5.5385×10^{2}	$\mathbf{-1.0590 \times 10^{0}}$	$\mathbf{-4.5375 \times 10^{2}}$
	COV	4999.97	4999.93	5000.00	5000.00
CI+MS	MEF	-1.4623×10^{1}	-2.2234×10^{3}	$\mathbf{-1.4500 \times 10^{1}}$	$\mathbf{-2.2053 \times 10^{3}}$
	BEF	-6.4895×10^{0}	-4.8165×10^{2}	$\mathbf{-5.3488 \times 10^{0}}$	$\mathbf{-4.2587 \times 10^{2}}$
	COV	5000.00	5000.00	5000.00	5000.00
PI+HS	MEF	$\mathbf{-2.3643 \times 10^{3}}$	$\mathbf{-7.3620 \times 10^{3}}$	-2.8283×10^{3}	-9.4439×10^{3}
	BEF	$\mathbf{-5.3098 \times 10^{2}}$	$\mathbf{-3.1908 \times 10^{2}}$	-8.9620×10^{2}	-1.8474×10^{3}
	COV	4999.96	5000.00	4999.83	4999.83
PI+MS	MEF	$\mathbf{-1.4651 \times 10^{1}}$	-2.3814×10^{8}	-1.4929×10^{1}	$\mathbf{-2.3382 \times 10^{8}}$
	BEF	$\mathbf{-6.2301 \times 10^{0}}$	-4.0531×10^{5}	-7.2177×10^{0}	$\mathbf{-9.9831 \times 10^{4}}$
	COV	5000.00	5000.00	5000.00	5000.00
NI+HS	MEF	-2.4177×10^{8}	-2.4651×10^{3}	$\mathbf{-2.3774 \times 10^{8}}$	$\mathbf{-2.4061 \times 10^{3}}$
	BEF	-8.1238×10^{5}	$\mathbf{-5.2911 \times 10^{2}}$	$\mathbf{-7.0860 \times 10^{5}}$	-5.5386×10^{2}
	COV	5000.00	5000.00	5000.00	5000.00
NI+MS	MEF	$\mathbf{-2.8261 \times 10^{1}}$	-3.8179×10^{1}	-3.3063×10^{0}	$\mathbf{-3.7108 \times 10^{1}}$
	BEF	$\mathbf{-1.0693 \times 10^{0}}$	$\mathbf{-2.2272 \times 10^{1}}$	-1.4517×10^{0}	-2.4183×10^{1}
	COV	5000.00	5000.00	5000.00	5000.00

'$\approx$', '$+$', and '$-$' indicate that the comparison algorithm is similar to, significantly worse than, and significantly better than the AET-CMA-ME algorithm, respectively, based on Wilcoxon rank-sum tests. The significance level is α = 0.05.

TABLE S.VI
MEAN EXPERIMENTAL RESULTS OF THE ESDT-ME-ME SOLVER AND AET-ME-ME SOLVER ON REACHABLE SPACE EXPLORATION PROBLEMS OF MANIPULATORS OVER THE METRICS QDS, BEF, AND COV. BOLD INDICATES THE TOP PERFORMER IN THE PAIRWISE COMPARISON.

Group	Metric	ESDT-ME-ME		AET-ME-ME	
		Task1	Task2	Task1	Task2
G_1	QDS	3558.04 (+)	3499.92 (+)	**3589.96**	**3537.31**
	BEF	0.99339	**0.99569**	**0.99623**	0.99547
	COV	3779.00	3749.63	**3781.77**	**3750.59**
G_2	QDS	3564.87 (+)	3512.59 (+)	**3594.00**	**3538.64**
	BEF	0.99486	0.99518	**0.99540**	**0.99548**
	COV	3806.45	3801.30	**3811.97**	**3802.87**
G_3	QDS	3523.38 (+)	3358.92 (+)	**3568.85**	**3382.81**
	BEF	0.99186	0.99416	**0.99243**	**0.99447**
	COV	3796.93	3690.13	**3802.17**	**3698.13**

'$\approx$', '$+$', and '$-$' indicate that the comparison algorithm is similar to, significantly worse than, and significantly better than the AET-ME-ME algorithm, respectively, based on Wilcoxon rank-sum tests. The significance level is α = 0.05.

For instance, preliminary experiments using the CMA-ME solver on benchmark problems with high and medium similarity (i.e., HS, MS), and the ME-ME solver on the manipulator problems are conducted. The comparative results are summarized in Tables S.V and S.VI. We can observe that ESDT also generally does not achieve better performance than the current AET on most tasks. We suspect ESDT underperforms compared to our main AET due to a potential encoder-decoder mismatch. While using a shared latent space, the source encoder's feature emphasis might not fully align with the target decoder's reconstruction capabilities for high-quality target solutions, leading to less effective or poorly adapted transfers. Therefore, ESDT has also not been adopted as the primary transfer mechanism in HMQD at this time. Therefore, how to better handle tasks with different dimensions is a direction worthy of further research in the heterogeneous multitasking framework in the immediate future.