EMPIRICAL-MEASURE SEARCH STATES FOR QUALITY-DIVERSITY OPTIMIZATION

Anonymous authors
Paper under double-blind review

ABSTRACT

Quality-Diversity (QD) optimization seeks collections of solutions that are both high-performing and behaviorally diverse. Many QD methods use an archive both to retain solutions and to guide subsequent search through archive improvement. Useful search, however, does not always look like immediate improvement. When evaluated candidates fail to improve the archive, they may not be retained, so search may stop pursuing those directions, leaving valuable regions undiscovered. We therefore need to separate search-state representation from result retention, allowing past search effort to continue guiding where to search even without immediate archive improvement. We introduce Quality-Diversity Empirical Measure (QDEM), which separates search-state representation from result retention, target-relative valuation, and search control. Under a fixed reference projection, QDEM recursively maintains a fixed-dimensional state of cumulative target support from all valid evaluations, guiding search toward under-supported targets. Our experiments demonstrate that QDEM outperforms state-of-the-art methods on a broad range of standard QD benchmarks and high-dimensional visual reference tasks, while remaining competitive where quality improvement already guides search effectively. In robot morphology evolution, QDEM allows lower-fitness intermediate designs rejected by the archive to remain part of the search process, enabling controller adaptation on the new morphology to recover and exceed pre-change performance.

1 INTRODUCTION

Quality-Diversity (QD) optimization aims to discover collections of solutions that are both high-performing and behaviorally diverse (Cully et al., 2015; Pugh et al., 2016). QD has been successfully applied to a wide range of domains, including robot locomotion and design (Cully et al., 2015; Duarte et al., 2018), game level and scenario generation (Gravina et al., 2019; Bhatt et al., 2022), protein design (Boige et al., 2023), and image generation (Fontaine & Nikolaidis, 2021).

Classical methods such as MAP-Elites partition the behavior space into cells and retain the highest-quality solution in each as an elite. Archived elites generate offspring, and successful offspring update the archive, which in turn guides subsequent search (Mouret & Clune, 2015). A prominent line of subsequent work uses adaptive emitters that update their search distributions according to archive progress (Fontaine et al., 2020; Fontaine & Nikolaidis, 2023; Tjanaka et al., 2026). Across these archive-improvement approaches, search is reinforced through progress assessed against retained elites, adaptive thresholds, or learned discount values.

Useful search, however, does not always produce immediate archive progress. Reaching a valuable region may require a prolonged sequence of evaluations that appear unproductive, or even a temporary drop in quality. This limitation has already been recognized in QD: strict elitism may block lower-fitness stepping stones needed to reach better solutions (Pugh et al., 2016), and in morphology–controller coevolution, promising body changes may initially reduce performance and be discarded before the controller adapts (Mertan & Cheney, 2025). If such directions are abandoned before their value emerges, valuable regions beyond them may remain undiscovered.

This difficulty arises in part because the archive couples two roles: retaining results and guiding subsequent search. Evaluations rejected by the archive are absent from the retained repertoire, even

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.

though they still record where evaluation effort has been spent. This motivates separating search-state representation from result retention, so that rejected evaluations can still influence future search without being retained as results. To make this history usable, search effort should be represented relative to behavioral locations. We use prescribed behavioral targets, specified for example by reference data as in QDDM (Tjanaka et al., 2026), to represent past search effort and guide search toward under-supported targets.

We introduce Quality-Diversity Empirical Measure (QDEM), which explicitly distinguishes four components: result retention, search-state representation, target-relative valuation, and search control. Every valid evaluation, including those rejected by the archive, contributes support to behavioral targets. For a fixed reference set and projection, the resulting state is fixed-dimensional, updates recursively, and exactly preserves cumulative target support without storing the full evaluation history. QDEM combines task quality with target-relative search value, giving candidates that support under-supported targets search credit.

Our work makes three contributions. First, we formulate search-state representation as a distinct design dimension in QD and introduce QDEM, separating it from result retention, target-relative valuation, and search control. Second, we characterize QDEM's search state as an exact, recursive, fixed-dimensional representation of cumulative target support from all valid evaluations. Third, we demonstrate strong performance on standard and high-dimensional visual QD benchmarks, and sustained morphology–controller adaptation through lower-fitness intermediates.

The remainder of the paper is organized as follows. Section 2 provides background, and Section 3 reviews related work. Section 4 develops QDEM. Section 5 reports the experiments. Sections 6 and 7 discuss the results and conclude. Implementation details are provided in the appendix.

2 Background

2.1 Quality-Diversity Optimization

Quality-Diversity (QD) optimization seeks a repertoire of high-quality solutions with diverse behaviors under a predefined behavior measure. Let $x \in \mathcal{X}$ be a candidate solution, $q : \mathcal{X} \to \mathbb{R}$ the quality objective, and $m : \mathcal{X} \to \mathcal{Z} \subseteq \mathbb{R}^d$ denote the behavior measure. For each reachable $z \in \mathcal{Z}$, the standard objective is to solve $\max_{x \in \mathcal{X}} q(x)$ subject to $m(x) = z$ (Cully & Demiris, 2018).

2.2 Reference-Defined Targets

Quality Diversity with Datasets of Measures (QDDM) uses reference examples, such as images, to specify desired behavioral locations (Tjanaka et al., 2026). Given a behavior measure m, we write these targets as a finite reference set $\mathcal{R} = \{r_1, \ldots, r_M\}$ in representation space $\mathcal{Z}$. Each r_j specifies a target location of interest, with local neighborhood $\mathcal{N}(r_j)$, and the corresponding objective is $x_j^* \in \arg\max_x q(x)$ subject to $m(x) \in \mathcal{N}(r_j)$.

2.3 Archive-Coupled QD Search

A central line of QD methods uses the archive both to retain solutions and to define progress signals for subsequent search. MAP-Elites retains the highest-quality elite in each behavior cell and samples archived elites for variation (Mouret & Clune, 2015). For a candidate x, let $e = \mathrm{cell}(m(x))$ denote its behavior cell and $A(e)$ the retained elite in that cell. CMA-ME introduces adaptive emitters. Its improvement emitter prioritizes new cells and ranks improvements in occupied cells by $\Delta(x) = q(x) - q(A(e))$ to update its search distribution (Fontaine et al., 2020). CMA-MAE replaces direct comparison with the elite by an adaptive cell-wise threshold t_e, defining progress as $\Delta(x) = q(x) - t_e$, so candidates can provide positive progress even without becoming new elites (Fontaine & Nikolaidis, 2023). DMS further generalizes this idea with a learned continuous discount model $q_b : \mathcal{Z} \to \mathbb{R}$, which predicts the quality level used to assess progress at each behavior (Tjanaka et al., 2026).

These methods differ in how they define and propagate archive progress. However, the retained archive records what survives, not how evaluation effort has been distributed. Archive-rejected evaluations leave the retained repertoire unchanged, even though they consume evaluation budget.

Let $A(H)$ denote the archive produced by history H, and let $n_e(H)$ count the evaluations assigned to cell e. Two equal-length histories may therefore satisfy

$$A(H) = A(H'), \qquad n_e(H) \neq n_e(H').$$

For example, histories can allocate rejected evaluations to different cells yet retain identical elites. Thus, the archive cannot recover how search effort was distributed.

3 Related Work

History-based and diversity-driven search. A broad class of diversity-driven methods uses observed behaviors to guide future search. Novelty search rewards behavioral difference from past and current samples, while novelty search with local competition (NSLC) adds local quality comparisons (Lehman & Stanley, 2011a;b). MEMES maintains a bounded novelty archive, DDS estimates behavioral density from a bounded sample buffer while keeping a separate passive result archive, and Dominated Novelty Search (DNS) evaluates candidates relative to fitter individuals in the current population–offspring set (Flageat et al., 2024; Lee et al., 2024; Bahlous-Boldi et al., 2025). These methods define novelty, density, or local competition relative to sampled behaviors. QDEM instead exactly accumulates support from all valid evaluations under a fixed reference projection and uses the resulting state for target-relative valuation.

Reference-based and continuous QD. DMS learns a continuous archive-relative discount model to guide search, and Tjanaka et al. (2026) also introduce QDDM, where reference data specify desired behavioral locations. Differentiable QD exploits gradients of quality and behavior functions, while Soft QD defines a continuous quality–coverage objective whose differentiable lower bound is optimized by SQUAD (Fontaine & Nikolaidis, 2021; Hedayatian & Nikolaidis, 2026). QDEM gives references a different role by using them to index a persistent search state that records accumulated evaluation support. This state can be valued relative to the desired target allocation to guide black-box search independently of result retention. Appendix E discusses further connections.

4 Quality-Diversity Empirical Measure

QDEM comprises four components: result retention, search-state representation, target valuation, and search control. Let $\mathcal{A}_t$, $\mathbf{S}_t$, and E_t denote the result archive, support state, and generator state. The behavioral target $\mathcal{T} = (\mathcal{R}, \pi)$ comprises fixed references $\mathcal{R} = \{r_1, \ldots, r_M\}$ and desired support shares $\pi_j > 0$, with $\sum_j \pi_j = 1$. The main benchmarks use equal shares, $\pi_j = 1/M$; Appendix C.8 tests unequal target shares.

For a batch $\mathcal{B}_t$ of valid new evaluations, with quality $q(x)$ and behavior $m(x)$, the four components are

$$\begin{aligned} \text{Result retention:} \quad & \mathcal{A}_{t+1} = \mathrm{Retain}(\mathcal{A}_t, \mathcal{B}_t), \\ \text{Search state:} \quad & \mathbf{S}_{t+1} = F_R(\mathbf{S}_t, \mathcal{B}_t), \\ \text{Target valuation:} \quad & v_t(x) = \mathcal{V}_{t,\mathcal{T}}(\mathbf{S}_t, m(x)), \\ \text{Search control:} \quad & E_{t+1} = \mathrm{Update}(E_t, \mathcal{B}_t, q, v_t). \end{aligned} \tag{1}$$

The generator forms its comparison set from new evaluations and any candidates stored in E_t, then uses the quality and search-value scores to update search.

4.1 Search-State Representation

Suppose N_t valid evaluations have produced behaviors $z_i = m(x_i)$. The Dirac measure δ_{z_i} records one evaluation at its observed behavior. Fixed projection weights $w_j(z) \geq 0$, with $\sum_j w_j(z) = 1$, distribute this unit of support across the references. This gives

$$\mathcal{M}_t = \sum_{i=1}^{N_t} \delta_{z_i}, \qquad S_{t,j} = \int w_j(z)\, \mathcal{M}_t(dz) = \sum_{i=1}^{N_t} w_j(z_i). \tag{2}$$

With nearest-reference assignment, the entire unit goes to the closest target, while soft assignment distributes it across nearby targets (Appendix A.1). QDEM accumulates these contributions in $\mathbf{S}_t =$

$(S_{t,1}, \ldots, S_{t,M})$, where $S_{t,j}$ records the cumulative support assigned to reference r_j. Because every evaluation contributes one unit in total, $\sum_j S_{t,j} = N_t$, so $\mathbf{S}_t$ records both the total search effort and how that effort is distributed across targets.

Proposition 1 (Exact recursive accumulation). For fixed references and projection, Eq. (2) admits the update

$$\mathbf{S}_{t+1} = F_R(\mathbf{S}_t, \mathcal{B}_t) = \mathbf{S}_t + \sum_{x \in \mathcal{B}_t} w(m(x)), \qquad \mathbf{S}_0 = 0. \tag{3}$$

Each valid evaluation contributes to the state. Thus, $\mathbf{S}_t$ can be updated from the current state and the new batch alone, without storing the full evaluation history (Appendix A.2). This state has $O(M)$ memory complexity, independent of the number of evaluations.

To characterize exactly what information $\mathbf{S}_t$ preserves, let $v \in \mathbb{R}^M$ assign a weight to each reference. For a history H_t, define the corresponding weighted target-support query as

$$T_v(H_t) = \sum_{i=1}^{N_t} v^\top w(z_i). \tag{4}$$

Choosing v as an indicator vector recovers the cumulative support assigned to any target or group of targets.

Proposition 2 (Exact target-support representation). For a fixed reference projection,

$$\begin{aligned} T_v(H) &= v^\top \mathbf{S}(H), && \forall v \in \mathbb{R}^M, \\ \mathbf{S}(H) = \mathbf{S}(H') &\iff T_v(H) = T_v(H'), && \forall v \in \mathbb{R}^M. \end{aligned} \tag{5}$$

Thus, $\mathbf{S}(H)$ is an exact M-dimensional representation of all weighted target-support queries under the fixed projection.

4.2 Target-Relative Valuation

Given the search state $\mathbf{S}_t$, QDEM next determines where additional search effort is most valuable relative to the desired target allocation. The ratio $S_{t,j}/\pi_j$ measures the support received by target j relative to its desired share. For an increasing, continuously differentiable, and strictly concave function $\psi_t : [0, \infty) \to \mathbb{R}$, define

$$U_t(\mathbf{S}; \mathcal{T}) = \sum_{j=1}^{M} \pi_j \psi_t\left(\frac{S_j}{\pi_j}\right). \tag{6}$$

Proposition 3 (Target calibration). For a given ψ_t and total support N, every $\mathbf{S} \geq 0$ with $\sum_j S_j = N$ satisfies

$$U_t(\mathbf{S}; \mathcal{T}) \leq \psi_t(N), \qquad U_t(\mathbf{S}; \mathcal{T}) = \psi_t(N) \iff \mathbf{S} = N\pi. \tag{7}$$

Thus, for a fixed amount of total search effort, U_t is uniquely maximized when support matches the desired target allocation (Appendix A.4).

Concavity makes the marginal value of additional support decrease as relative support accumulates. For a candidate behavior z, QDEM defines its Marginal Support Gain (MSG) as

$$\mathcal{V}_{t,\mathcal{T}}(\mathbf{S}_t, z) = \mathrm{MSG}_t(z) = \sum_{j=1}^{M} \psi_t'\left(\frac{S_{t,j}}{\pi_j}\right) w_j(z). \tag{8}$$

MSG therefore gives more search credit to candidates that contribute to under-supported targets. With ψ_t fixed during scoring, MSG is the directional derivative of U_t along the candidate contribution $w(z)$ (Appendix A.5).

In the main experiments, we use equal target shares and marginal weights $\exp[-(S_{t,j}/\tau_t)^p]$, with $p > 0$ and default $p = 1$. The scale $\tau_t > 0$ adapts to the current support distribution and remains fixed while candidates are scored (Appendix A.6).

4.3 Search Control and Result Retention

QDEM combines MSG with task quality through

$$a_t(x) = q(x) + \lambda \, \mathrm{Rank01}_{\mathcal{C}_t}\big(v_t(x)\big), \qquad v_t(x) = \mathrm{MSG}_t(m(x)). \tag{9}$$

Rank01 maps MSG ranks in the comparison set $\mathcal{C}_t$ to $[0, 1]$, giving the support term a fixed scale while leaving quality unchanged. With this normalization, $\lambda \geq 0$ controls the quality–support trade-off and bounds how much quality loss support can compensate. The generator uses a_t to update its sampling distribution or select survivors.

All candidates are scored against the same pre-update state, after which new valid evaluations are added jointly. In our experiments, the QD archive is used only to retain and report results for comparison with prior methods and does not enter QDEM scoring or generator updates. Appendix A.7 and Algorithm A1 give the full procedure.

5 Experiments

We evaluate QDEM from four perspectives. A controlled valley benchmark tests whether search can continue through temporary drops in quality. Standard numerical benchmarks examine when accumulated evaluation effort provides useful guidance across objective landscapes and behavior dimensions. High-dimensional visual tasks test the approach with reference-defined behavioral targets. Finally, morphology–controller coevolution examines whether lower-fitness intermediate designs can remain in search long enough for subsequent adaptation.

5.1 Experimental Setup

Controlled continuation. We use a three-dimensional, two-cell landscape to test search through temporary drops in quality (Section 5.2).

Standard numerical benchmarks. We follow eight Linear Projection (LP) tasks from DMS: Sphere 2D/10D/20D/50D, Rastrigin 2D/10D, and Flat 2D/10D (Tjanaka et al., 2026). Each candidate is a 100-dimensional vector. After the benchmark coordinate transform, its variables are divided into equal contiguous groups and summed to form a d-dimensional behavior descriptor. The same projection is paired with a unimodal Sphere objective, a multimodal Rastrigin objective, or constant quality in Flat, varying behavior dimensionality and quality structure while keeping the decision-space dimension fixed.

High-dimensional visual targets. We use the three reference-defined visual benchmarks from DMS (Tjanaka et al., 2026). TA-MNIST and TA-F-MNIST optimize arrangements of 30 grayscale triangles. Each candidate renders a 28×28 image, whose 784 pixel values form the behavior representation, while quality measures similarity to the corresponding dataset targets. Hiker optimizes a 512-dimensional StyleGAN3 latent code to generate faces. Behavior is represented by a 512-dimensional CLIP image embedding and compared with landscape references. Hiker quality follows the released DMS implementation with the prompt "A photo of the face of a hiker."

Robot morphology evolution. We follow the EvoGym Walker-v0 morphology-controller coevolution setting of Mertan & Cheney (2025). A solution specifies a 10×10 voxel body and controller, and quality is locomotion return. Active-voxel count and existing-voxel count define the behavior descriptor, yielding 55 physically feasible niches in the reporting archive.

Baselines and reporting. The controlled task compares MAP-Elites, CMA-ME, CMA-MAE, and DMS. Standard numerical tasks compare MAP-Elites, CMA-MAE, DDS, and DMS. Following the DMS evaluation protocol, the visual tasks use the same suite without DDS because of the cost of KDE in these high-dimensional spaces (Tjanaka et al., 2026). For the robot task, we follow the MAP-Elites-based morphology-evolution setting of Mertan & Cheney (2025) and use MAP-Elites as the reference baseline for this targeted continuation test. For the standard and visual tasks, QD Score sums elite quality in the reporting archive and Coverage is the occupied fraction. Appendix B gives exact task definitions, evaluation budgets, random seeds, and method-specific parameters.

5.2 Controlled Continuation under Temporary Quality Decreases

This benchmark isolates whether search can continue when reaching a better region requires a temporary decrease in quality. For $x = (s, z_1, z_2) \in [0, 8] \times \mathbb{R}^2$, quality is $q(x) = q_{\text{center}}(s) - 0.01(z_1^2 + z_2^2)$, where

$$q_{\text{center}}(s) = \begin{cases} 0.30, & 0 \le s < 1, \\ 0.30 - 0.06(s-1), & 1 \le s \le 3, \\ 0.18 + 0.06(s-3), & 3 < s \le 5, \\ 0.30 + 0.02(s-5), & 5 < s \le 8. \end{cases} \tag{10}$$

Search starts at $(1, 0, 0)$, before quality drops to the valley at $s = 3$ and then rises to a higher peak at $s = 8$. The result cells are $C_0 : s \le 6$ and $C_1 : s > 6$. Because the starting point and the valley lie in the same result cell C_0, intermediate solutions in the valley do not replace the starting elite until their quality recovers above its level. QDEM uses two soft references at the result-cell midpoints $s = 3, 7$. All methods receive the same evaluation budget (Appendix B.2).

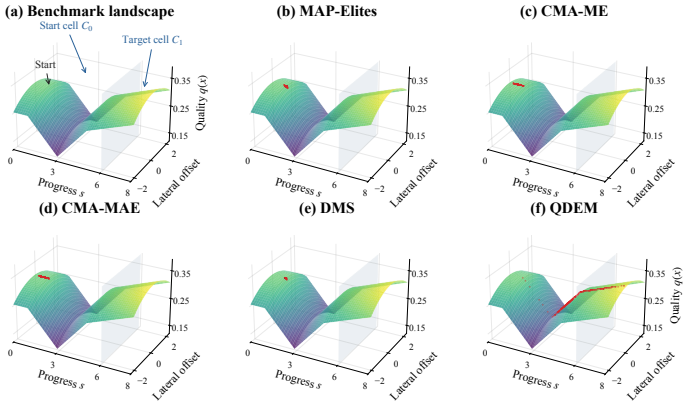


Figure 1: Two-cell valley continuation. (a) The plane at $s = 6$ separates the two result cells. (b)–(f) Red points show the mean s-coordinate of the 36 candidates evaluated in each generation for one seed, projected onto the centerline landscape.

Table 1 distinguishes forward proposals from sustained continuation. Forward proposals reports the mean fraction of candidates sampled ahead of the current search position. Valley crossing and Higher peak count runs that enter C_1 and find a solution there exceeding the starting quality, respectively. Furthest s reports the mean $\pm$ sample SD of the maximum evaluated s. Final σ_s reports the mean terminal search scale along s.

Table 1: Controlled valley results over five seeds.

Method	**Forward proposals**	**Valley crossing**	**Higher peak**	**Furthest s**	**Final σ_s**
MAP-Elites	49.9%	0/5	0/5	1.991 ± 0.071	---†
CMA-ME	50.4%	0/5	0/5	1.967 ± 0.252	0.0399
CMA-MAE	50.4%	0/5	0/5	1.967 ± 0.252	0.0399
DMS	46.8%	0/5	0/5	1.698 ± 0.121	0.0018
QDEM	---‡	**5/5**	**5/5**	$\mathbf{7.999 \pm 0.000}$	---‡

Note: † MAP-Elites has no adaptive search scale. ‡ QDEM reaches $s \simeq 8$ in every run, so its forward-proposal rate and final scale are omitted because they include post-discovery search.

The four baselines propose forward moves roughly half the time but never cross the valley. Their mean furthest progress ranges from $s = 1.70$ to 1.99, well before the valley floor at $s = 3$, showing that occasional forward proposals do not translate into sustained continuation. QDEM crosses the valley in all five runs, reaches the higher peak in all five, and attains quality 0.3587 ± 0.0010, close to the optimum of 0.36. Soft assignment gives evaluations within the same result cell different contributions to the two targets, allowing progress toward the less-supported target to receive search credit even before those evaluations improve the archive.

More seriously, failed continuation feeds back into the adaptive search itself. The final search scale along s contracts to 0.0399 for CMA-ME and CMA-MAE and 0.0018 for DMS. At these scales,

proposals become increasingly local, making further movement through the valley unlikely and effectively trapping search near the starting peak.

5.3 Generalization across Standard QD Tasks

We next test this mechanism on the eight standard black-box tasks from DMS (Tjanaka et al., 2026). These tasks provide a useful contrast. Rastrigin has a multimodal quality landscape, Flat has no quality gradient, and Sphere has a unimodal objective for which local quality improvement is usually aligned with useful search directions. If QDEM mainly helps when immediate quality progress is an incomplete guide to search, its gains should be larger on Rastrigin and Flat and smaller on Sphere.

Table 2: QD Score and Coverage on the standard QD benchmarks over 20 seeds. The best mean in each column is bold.

(a) Rastrigin and Flat

Method	Rastrigin 2D		Rastrigin 10D		Flat 2D		Flat 10D	
	QD	Cov.	QD	Cov.	QD	Cov.	QD	Cov.
QDEM	**5,867.9**	**92.34%**	**5,328.6**	88.02%	**8,218.5**	**82.19%**	**8,433.3**	**84.33%**
DMS	5,724.6	91.48%	5,129.0	**88.16%**	8,013.1	80.13%	7,904.6	79.05%
CMA-MAE	5,270.6	80.02%	156.0	1.78%	5,593.4	55.93%	1,472.5	14.72%
DDS	2,614.7	72.50%	2,848.0	55.95%	7,187.6	71.88%	4,967.9	49.68%
MAP-Elites	3,171.9	48.24%	532.3	7.59%	4,350.9	43.51%	122.2	1.22%

(b) Sphere

Method	Sphere 2D		Sphere 10D		Sphere 20D		Sphere 50D	
	QD	Cov.	QD	Cov.	QD	Cov.	QD	Cov.
QDEM	**7,123.9**	94.83%	**7,005.6**	**92.25%**	6,879.2	90.08%	6,676.6	82.42%
DMS	7,015.6	**96.35%**	6,387.8	88.93%	**7,351.1**	**95.91%**	**6,985.2**	**86.90%**
CMA-MAE	6,330.2	80.99%	715.5	7.46%	802.5	8.25%	2,329.8	24.23%
DDS	3,320.9	71.53%	3,724.5	56.61%	6,985.9	84.25%	6,370.1	74.28%
MAP-Elites	4,162.0	50.74%	201.1	2.07%	3,282.3	35.59%	3,887.7	41.07%

QDEM performs strongest overall on Rastrigin and Flat, achieving the highest QD Score on all four tasks and the highest or nearly highest coverage throughout. On Rastrigin, this shows that cumulative support remains useful when search must balance quality improvement with repertoire expansion on a multimodal landscape. On Flat, where quality provides no search signal, the broader coverage directly reflects more effective allocation of evaluation effort. Sphere shows a different pattern. QDEM remains competitive, while DMS performs better in the higher-dimensional settings. Because Sphere already provides a smooth quality gradient, immediate quality improvement can itself guide search effectively.

Appendices C.2 and C.3 test the contributions of cumulative history and candidate-specific MSG to coverage and repertoire quality.

5.4 High-Dimensional Visual Reference Tasks

Beyond the controlled and standard benchmarks, we evaluate QDEM on three high-dimensional visual-reference tasks from DMS (Tjanaka et al., 2026). All methods share the benchmark objectives, reporting references, and evaluation protocol described in Section 5.1 and Appendix B.4.

Table 3: Visual-reference results over five seeds. The best mean in each column is bold. Table D3 gives mean $\pm$ sample SD for every method.

Method	TA-MNIST		TA-F-MNIST		Hiker	
	QD	Cov.	QD	Cov.	QD	Cov.
QDEM	**962.0**	99.98%	**827.5**	**85.02%**	**428.34**	**8.30%**
DMS	953.1	**100.00%**	710.4	73.20%	121.98	3.00%
CMA-MAE	954.2	99.44%	617.4	63.06%	−25.36	1.41%
MAP-Elites	940.7	98.26%	535.9	55.02%	−19,295.61	5.90%

QDEM achieves the highest QD Score on all three tasks. TA-MNIST coverage is nearly saturated for all methods, while on the more challenging TA-F-MNIST and Hiker tasks, QDEM improves both QD Score and Coverage over DMS. On Hiker, coverage increases from 3.00% to 8.30%. Appendix C.1 shows that this gain accompanies less concentrated effort, more discoveries in each equal-budget interval, and higher final QD Score. MAP-Elites reaches more references than DMS but has a negative QD Score, showing that reference coverage alone does not ensure high-quality solutions. Figure 2 qualitatively shows QDEM producing varied faces with more consistently complete facial structure, alongside its broader reference coverage in Table 3.



Figure 2: Hiker examples showing the first three archive rows for each method from the same seed. Faces are paired with their assigned landscape references. Appendix D.3 gives expanded examples.

5.5 Search State and Morphology–Controller Adaptation

Finally, we test whether the same continuation mechanism appears in morphology–controller coevolution. Standard MAP-Elites accepts offspring whose target niche is empty or whose fitness exceeds its incumbent elite (Mouret & Clune, 2015). QDEM achieves a 28.9% higher mean QD Score than MAP-Elites while both fill all 55 feasible niches, indicating higher repertoire quality at full coverage (Table 4). Its advantage over the matched Quality-only ablation further supports the contribution of support-guided search under the same population mechanism.

Figure 3 shows three forms of recovery after a body mutation causes a temporary fitness drop. In (a), the new morphology moves to a different archive cell and recovers over several controller adaptations, with fitness fluctuating along the lineage. In (b), recovery unfolds over multiple controller updates while remaining within the same result cell. In (c), a larger initial drop within the same cell

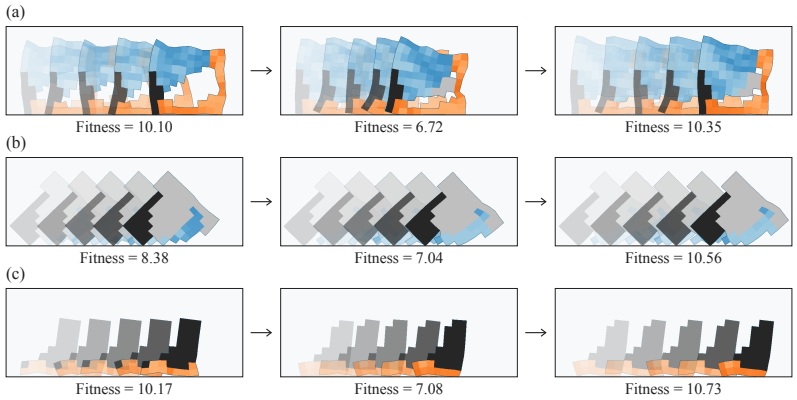


Figure 3: Morphology–controller recovery in QDEM. Rows show (a) cross-cell, (b) within-cell, and (c) rapid within-cell recovery after a large fitness drop. Columns show parents, archive-rejected body mutants retained by QDEM, and descendants recovered through controller mutation. Fitness is averaged over five rollouts. Overlaid poses are independent replays. Arrows show ancestry and may skip intermediate descendants.

Table 4: Final robot performance (mean ± sample SD over five seeds). Coverage counts occupied cells out of 55.

Method	QD Score ↑	Best fitness ↑	Coverage ↑
QDEM	**525.2 ± 64.4**	**11.80 ± 0.39**	55.0 ± 0.0
MAP-Elites	407.4 ± 37.4	11.25 ± 0.38	55.0 ± 0.0
Quality-only	92.9 ± 16.6	9.18 ± 0.91	39.4 ± 4.5

is recovered after a single controller mutation. In all three cases, the body-changing intermediate is rejected by the result archive but retained in QDEM's search population. The morphology then remains unchanged while controller adaptation produces a descendant that improves the archive. These trajectories connect the controlled-valley effect to morphology search, where poor intermediate results can still support continued adaptation.

Appendix C.4 shows that this effect extends beyond the illustrated examples. Across five seeds, 2,180 archive-rejected body mutants recover the pre-change parent's fitness within 100 generations, including 633 through controller-only adaptation on the unchanged new body. Of these recovered intermediates, 1,400 would have been excluded under quality-only ranking when they entered the search population, showing that support broadly changes which intermediates remain available for continued search.

6 Discussion

Our results establish search-state representation as a distinct design dimension in QD. QDEM is especially useful when immediate quality improvement provides incomplete search guidance, while its advantage narrows when quality itself provides a reliable search direction. The same-archive construction shows that identical retained solutions can correspond to different support states and candidate rankings (Appendix A.3). History and MSG ablations support the value of cumulative support and candidate-specific scoring (Appendices C.2 and C.3). On Hiker, QDEM discovers more regions within the same evaluation budget and achieves higher final QD Score (Appendix C.1).

The robot study shows how search state shapes the process of adaptation itself. A body change can be poor as an immediate result yet valuable as a substrate for subsequent controller adaptation (Cheney et al., 2018). Support-guided selection keeps some archive-rejected morphologies in search long enough for this value to emerge. Keeping different intermediate solutions available for continued search preserves different possibilities for subsequent adaptation. This suggests a broader view of QD in which diversity lies not only in the outcome, but also in the search process.

Separating search state from result retention also makes the search process more modular. QDEM separates what search remembers, how that state is valued, how it guides future search, and what results are retained. Archive contents are not required for QDEM scoring or generator updates, making the archive a mechanism for retaining results rather than the state that drives search. With references and projection fixed, the same support state can also be re-valued under different target shares or valuation profiles without reconstructing the evaluation history. Appendix C.8 shows that changing only the target shares redirects evaluation effort while keeping the reference geometry and generator configuration fixed. This shifts attention from only what QD retains to what the search process itself should remember. Once these roles are separated, QD can ask which aspects of search experience should be represented, how that state should adapt across targets or tasks, and what search values should be derived from it. The empirical-measure state and MSG studied here provide one answer to these questions, while richer search states, alternative valuations, different controllers, and alternative forms of result retention remain open directions.

7 Conclusion

We introduced QDEM, which separates search-state representation from result retention, target-relative valuation, and search control. Under a fixed reference projection, its fixed-dimensional state recursively and exactly preserves cumulative target support from all valid evaluations and uses this support to guide subsequent search. Across numerical, visual, and robotic tasks, QDEM shows that archive-rejected evaluations can still guide exploration and sustain adaptation through temporary quality losses. These results establish search-state representation as a distinct design dimension in QD.

AI USE STATEMENT

In this work, we used generative AI tools to assist with translation and language editing, code editing and debugging, drafting and refining mathematical proofs, and literature discovery. The underlying mathematical ideas, theorem statements, and proof strategies were developed by the authors. We did not use generative AI tools to generate the research questions, methodology, experimental design, or scientific conclusions, and the remaining disclosure categories are not applicable to this work. We also used generative AI tools for manuscript and LaTeX formatting. All AI-assisted work was reviewed by the authors. AI-assisted proof writing was independently checked, code modifications used in the experiments were tested and verified for correctness, and cited literature was verified against the original sources. We take responsibility for the final content of this work, including text, claims, and artifacts produced with the aid of generative AI.

REPRODUCIBILITY STATEMENT

The formulation and algorithmic details of QDEM are provided in the main text. Complete derivations and proofs are provided in the appendix. The appendix also documents the experimental domains, baseline configurations, hyperparameters, random seeds, evaluation protocols, and implementation details for the reported experiments. An anonymized code package containing the source code and experiment configurations is provided as supplementary material to facilitate reproduction of the results. The code will be publicly released upon acceptance.

REFERENCES

Ryan Bahlous-Boldi, Maxence Faldor, Luca Grillotti, Hannah Janmohamed, Lisa Coiffard, Lee Spector, and Antoine Cully. Dominated novelty search: Rethinking local competition in quality-diversity. In *Proceedings of the Genetic and Evolutionary Computation Conference*, pp. 104–112. ACM, 2025. doi: 10.1145/3712256.3726310.

Marc G. Bellemare, Sriram Srinivasan, Georg Ostrovski, Tom Schaul, David Saxton, and Rémi Munos. Unifying count-based exploration and intrinsic motivation. In *Advances in Neural Information Processing Systems*, volume 29, 2016. URL https://proceedings.neurips.cc/paper/2016/hash/afda332245e2af431fb7b672a68b659d-Abstract.html.

Jagdeep Singh Bhatia, Holly Jackson, Yunsheng Tian, Jie Xu, and Wojciech Matusik. Evolution Gym: A large-scale benchmark for evolving soft robots. In *Advances in Neural Information Processing Systems*, volume 34, pp. 2201–2214, 2021. URL https://proceedings.neurips.cc/paper/2021/hash/118921efba23fc329e6560b27861f0c2-Abstract.html.

Varun Bhatt, Bryon Tjanaka, Matthew C. Fontaine, and Stefanos Nikolaidis. Deep surrogate assisted generation of environments. In *Advances in Neural Information Processing Systems*, volume 35, pp. 37762–37777, 2022. URL https://proceedings.neurips.cc/paper_files/paper/2022/hash/f649556471416b35e60ae0de7c1e3619-Abstract-Conference.html.

Raphaël Boige, Guillaume Richard, Jérémie Donà, Thomas Pierrot, and Antoine Cully. Gradient-informed quality diversity for the illumination of discrete spaces. In *Proceedings of the Genetic and Evolutionary Computation Conference*, pp. 119–128. ACM, 2023. doi: 10.1145/3583131.3590407.

Nick Cheney, Josh Bongard, Vytas SunSpiral, and Hod Lipson. Scalable co-optimization of morphology and control in embodied machines. *Journal of the Royal Society Interface*, 15(143): 20170937, 2018. doi: 10.1098/rsif.2017.0937.

Alexandre Coninx and Stéphane Doncieux. Younger is better: A simple and efficient selection strategy for MAP-Elites. In *Proceedings of the Genetic and Evolutionary Computation Conference Companion*, pp. 87–88. ACM, 2021. doi: 10.1145/3449726.3459493.

Roger Creus Castanyer, Joshua Romoff, and Glen Berseth. Improving intrinsic exploration by creating stationary objectives. In *International Conference on Learning Representations*, 2024. URL https://proceedings.iclr.cc/paper_files/paper/2024/hash/7058bc192a37f5e5a57398887b05f6f6-Abstract-Conference.html.

Antoine Cully. Autonomous skill discovery with quality-diversity and unsupervised descriptors. In *Proceedings of the Genetic and Evolutionary Computation Conference*, pp. 81–89. ACM, 2019. doi: 10.1145/3321707.3321804.

Antoine Cully and Yiannis Demiris. Quality and diversity optimization: A unifying modular framework. *IEEE Transactions on Evolutionary Computation*, 22(2):245–259, 2018. doi: 10.1109/tevc.2017.2704781.

Antoine Cully, Jeff Clune, Danesh Tarapore, and Jean-Baptiste Mouret. Robots that can adapt like animals. *Nature*, 521(7553):503–507, 2015. doi: 10.1038/nature14422.

Li Ding, Jenny Zhang, Jeff Clune, Lee Spector, and Joel Lehman. Quality diversity through human feedback: Towards open-ended diversity-driven optimization. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 11072–11090. PMLR, 2024. URL https://proceedings.mlr.press/v235/ding24h.html.

Miguel Duarte, Jorge Gomes, Sancho Moura Oliveira, and Anders Lyhne Christensen. Evolution of repertoire-based control for robots with complex locomotor systems. *IEEE Transactions on Evolutionary Computation*, 22(2):314–328, 2018. doi: 10.1109/tevc.2017.2722101.

Manon Flageat, Bryan Lim, and Antoine Cully. Enhancing MAP-Elites with multiple parallel evolution strategies. In *Proceedings of the Genetic and Evolutionary Computation Conference*, pp. 1082–1090. ACM, 2024. doi: 10.1145/3638529.3654089.

Matthew C. Fontaine and Stefanos Nikolaidis. Differentiable quality diversity. In *Advances in Neural Information Processing Systems*, volume 34, pp. 10040–10052, 2021. URL https://papers.nips.cc/paper_files/paper/2021/hash/532923f11ac97d3e7cb0130315b067dc-Abstract.html.

Matthew C. Fontaine and Stefanos Nikolaidis. Covariance matrix adaptation MAP-annealing. In *Proceedings of the Genetic and Evolutionary Computation Conference*, pp. 456–465. ACM, 2023. doi: 10.1145/3583131.3590389.

Matthew C. Fontaine, Julian Togelius, Stefanos Nikolaidis, and Amy K. Hoover. Covariance matrix adaptation for the rapid illumination of behavior space. In *Proceedings of the Genetic and Evolutionary Computation Conference*, pp. 94–102. ACM, 2020. doi: 10.1145/3377930.3390232.

Matthew C. Fontaine, Ruilin Liu, Ahmed Khalifa, Jignesh Modi, Julian Togelius, Amy K. Hoover, and Stefanos Nikolaidis. Illuminating Mario scenes in the latent space of a generative adversarial network. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 35, pp. 5922–5930, 2021. doi: 10.1609/aaai.v35i7.16740.

Daniele Gravina, Ahmed Khalifa, Antonios Liapis, Julian Togelius, and Georgios N. Yannakakis. Procedural content generation through quality diversity. In *2019 IEEE Conference on Games (CoG)*, pp. 1–8. IEEE, 2019. doi: 10.1109/cig.2019.8848053.

Luca Grillotti and Antoine Cully. Relevance-guided unsupervised discovery of abilities with quality-diversity algorithms. In *Proceedings of the Genetic and Evolutionary Computation Conference*, pp. 77–85. ACM, 2022. doi: 10.1145/3512290.3528837.

Nikolaus Hansen. The CMA evolution strategy: A tutorial. *arXiv preprint arXiv:1604.00772*, 2016. URL https://arxiv.org/abs/1604.00772.

Saeed Hedayatian and Stefanos Nikolaidis. Soft quality-diversity optimization. In *International Conference on Learning Representations*, 2026. URL https://proceedings.iclr.cc/paper_files/paper/2026/hash/79eb0ed60c0660fc2fbe4a2bf75cf4cc-Abstract-Conference.html.

Tero Karras, Miika Aittala, Samuli Laine, Erik Härkönen, Janne Hellsten, Jaakko Lehtinen, and Timo Aila. Alias-free generative adversarial networks. In *Advances in Neural Information Processing Systems*, volume 34, pp. 852–863, 2021. URL https://proceedings.neurips.cc/paper/2021/hash/076ccd93ad68be51f23707988e934906-Abstract.html.

Sam Kriegman, Nick Cheney, and Josh Bongard. How morphological development can guide evolution. *Scientific Reports*, 8(1):13934, 2018. doi: 10.1038/s41598-018-31868-7.

Yann LeCun, Léon Bottou, Yoshua Bengio, and Patrick Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, 1998. doi: 10.1109/5.726791.

David H. Lee, Anishalakshmi V. Palaparthi, Matthew C. Fontaine, Bryon Tjanaka, and Stefanos Nikolaidis. Density descent for diversity optimization. In *Proceedings of the Genetic and Evolutionary Computation Conference*, pp. 674–682. ACM, 2024. doi: 10.1145/3638529.3654001.

Joel Lehman and Kenneth O. Stanley. Abandoning objectives: Evolution through the search for novelty alone. *Evolutionary Computation*, 19(2):189–223, 2011a. doi: 10.1162/evco_a_00025.

Joel Lehman and Kenneth O. Stanley. Evolving a diversity of virtual creatures through novelty search and local competition. In *Proceedings of the Genetic and Evolutionary Computation Conference*, pp. 211–218. ACM, 2011b. doi: 10.1145/2001576.2001606.

George Mathew and Igor Mezić. Metrics for ergodicity and design of ergodic dynamics for multi-agent systems. *Physica D: Nonlinear Phenomena*, 240(4–5):432–442, 2011. doi: 10.1016/j.physd.2010.10.010.

Alican Mertan and Nick Cheney. Controller distillation reduces fragile brain-body co-adaptation and enables migrations in MAP-Elites. In *Proceedings of the Genetic and Evolutionary Computation Conference*, pp. 158–166. ACM, 2025. doi: 10.1145/3712256.3726368.

Jean-Baptiste Mouret and Jeff Clune. Illuminating search spaces by mapping elites. *arXiv preprint arXiv:1504.04909*, 2015. URL https://arxiv.org/abs/1504.04909.

Jørgen Nordmoen, Frank Veenstra, Kai Olav Ellefsen, and Kyrre Glette. MAP-Elites enables powerful stepping stones and diversity for modular robotics. *Frontiers in Robotics and AI*, 8:639173, 2021. doi: 10.3389/frobt.2021.639173.

Justin K. Pugh, Lisa B. Soros, and Kenneth O. Stanley. Quality diversity: A new frontier for evolutionary computation. *Frontiers in Robotics and AI*, 3:40, 2016. doi: 10.3389/frobt.2016.00040.

Alec Radford, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry, Amanda Askell, Pamela Mishkin, Jack Clark, Gretchen Krueger, and Ilya Sutskever. Learning transferable visual models from natural language supervision. In *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pp. 8748–8763. PMLR, 2021. URL https://proceedings.mlr.press/v139/radford21a.html.

Konstantinos Sfikas, Antonios Liapis, and Georgios N. Yannakakis. Monte Carlo Elites: Quality-diversity selection as a multi-armed bandit problem. In *Proceedings of the Genetic and Evolutionary Computation Conference*, pp. 180–188. ACM, 2021. doi: 10.1145/3449639.3459321.

Amit Sinha, Matthieu Geist, and Aditya Mahajan. Periodic agent-state based Q-learning for POMDPs. In *Advances in Neural Information Processing Systems*, volume 37, pp. 62123–62159, 2024. URL https://proceedings.neurips.cc/paper_files/paper/2024/hash/71b52a5b3fe2e9303433a174b60e160d-Abstract-Conference.html.

Ivan Skorokhodov, Grigorii Sotnikov, and Mohamed Elhoseiny. Aligning latent and image spaces to connect the unconnectable. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 14144–14153, 2021. URL https://openaccess.thecvf.com/content/ICCV2021/html/Skorokhodov_Aligning_Latent_and_Image_Spaces_To_Connect_the_Unconnectable_ICCV_2021_paper.html.

Jayakumar Subramanian, Amit Sinha, Raihan Seraj, and Aditya Mahajan. Approximate information state for approximate planning and reinforcement learning in partially observed systems. *Journal of Machine Learning Research*, 23(12):1–83, 2022. URL https://jmlr.org/papers/v23/20-1165.html.

Bryon Tjanaka, Henry Chen, Matthew C. Fontaine, and Stefanos Nikolaidis. Discount model search for quality diversity optimization in high-dimensional measure spaces. In *International Conference on Learning Representations*, 2026. URL https://openreview.net/forum?id=m6Hv0yZO3n.

Vassilis Vassiliades, Konstantinos Chatzilygeroudis, and Jean-Baptiste Mouret. Using centroidal Voronoi tessellations to scale up the multidimensional archive of phenotypic elites algorithm. *IEEE Transactions on Evolutionary Computation*, 22(4):623–630, 2018. doi: 10.1109/tevc.2017.2735550.

Han Xiao, Kashif Rasul, and Roland Vollgraf. Fashion-MNIST: A novel image dataset for benchmarking machine learning algorithms. *arXiv preprint arXiv:1708.07747*, 2017. URL https://arxiv.org/abs/1708.07747.

A Theoretical and Method Details

We prove what the support state preserves and how it can be updated from new evaluations. We then show how this state determines candidate scores and give the resulting search procedure.

A.1 Local Reference Projection

Let $\mathcal{R} = \{r_1, \dots, r_M\}$ be distinct references and let $\mathcal{N}_k(z)$ contain the k nearest references to behavior z, with deterministic tie handling. For the controlled, numerical, and visual tasks, Gaussian weights with bandwidth $h > 0$ give

$$\widetilde{w}_j(z) = \exp\left(-\frac{\|z - r_j\|_2^2}{2h^2}\right) \mathbf{1}\{r_j \in \mathcal{N}_k(z)\}, \qquad w_j(z) = \frac{\widetilde{w}_j(z)}{\sum_\ell \widetilde{w}_\ell(z)}. \tag{A.1}$$

Thus each evaluation contributes one unit of support, shared among nearby targets. For hard assignment, $w_j(z)$ is one for the assigned reference and zero for all others. In both cases, $K_R(z, \cdot) = \sum_j w_j(z)\delta_{r_j}$ expresses one evaluation's contribution as a measure on the references. The references and projection remain fixed during a run. Descriptor preprocessing and bandwidth calculation are specified in Appendix B.1.

A.2 Exact and Recursive Representation of Target Support

On the measurable behavior space $(\mathcal{Z}, \mathcal{B}_\mathcal{Z})$, let $\mathcal{M}_t = \sum_{i=1}^{N_t} \delta_{z_i}$. The fixed probability kernel K_R gives the cumulative search state $\mathcal{S}_t = \mathcal{M}_t K_R = \sum_j S_{t,j}\delta_{r_j}$. For $N_t > 0$, the normalized allocation is $\eta_t = \mathcal{S}_t/N_t$, with $\eta_{t,j} = S_{t,j}/N_t$. It describes realized evaluation effort over references.

Write $\boldsymbol{\eta}_t = (\eta_{t,1}, \dots, \eta_{t,M})$ for the coordinate vector of η_t. For a valid batch $\mathcal{B}_t$ of size $b_t = |\mathcal{B}_t|$, the cumulative and normalized updates are

$$\begin{aligned} \mathbf{S}_{t+1} &= \mathbf{S}_t + \sum_{x \in \mathcal{B}_t} w(m(x)), \qquad N_{t+1} = N_t + b_t, \\ \boldsymbol{\eta}_{t+1} &= \frac{N_t \boldsymbol{\eta}_t + \sum_{x \in \mathcal{B}_t} w(m(x))}{N_t + b_t}. \end{aligned} \tag{A.2}$$

The cumulative update applies from initialization onward. Each new evaluation contributes once. Selecting a stored individual leaves support unchanged. Since $\mathbf{1}^\top \mathbf{S}_t = N_t$, the cumulative vector records both total effort and its allocation. For $N_t > 0$, it is equivalent to $(N_t, \boldsymbol{\eta}_t)$, whereas normalized allocation alone loses the total count. At initialization, QDEM uses $\mathbf{S}_0 = 0$ and equal marginal weights.

For a finite counting measure $\mathcal{M}$, let $\mathbf{S}(\mathcal{M}) = \int w \, d\mathcal{M}$. Linearity gives

$$\mathbf{S}(\mathcal{M} + \mathcal{M}') = \mathbf{S}(\mathcal{M}) + \mathbf{S}(\mathcal{M}'), \qquad \mathbf{1}^\top \mathbf{S}(\mathcal{M}) = \mathcal{M}(\mathcal{Z}). \tag{A.3}$$

The additive update proves Proposition 1. Reordering the same evaluation events leaves the final state unchanged, although it can affect intermediate states and subsequent proposals.

Target-support exactness. For any function $v : \mathcal{R} \to \mathbb{R}$,

$$\int v \, d\mathcal{S}_t = \sum_{i=1}^{N_t} \int v \, dK_R(z_i, \cdot) = \sum_j v(r_j) S_{t,j}. \tag{A.4}$$

Identifying $v : \mathcal{R} \to \mathbb{R}$ with the vector $(v(r_1), \dots, v(r_M))$, Eq. (A.4) gives $T_v(H) = v^\top \mathbf{S}(H)$. If $\mathbf{S}(H) = \mathbf{S}(H')$, all these totals coincide. Conversely, choosing a coordinate indicator for each r_j recovers $S_j(H) = S_j(H')$. This proves Proposition 2. Any representation that determines all the totals T_v must therefore also determine $\mathbf{S}(H)$. Together with the additive update in Eq. (A.3), this gives an exact online representation using M support coordinates, independent of the number of evaluation events.

With the same target shares and utility function, identical support states give identical utility and MSG for every candidate. This also holds when the utility scale is computed from support (Appendix A.6). Candidate scoring therefore requires no reconstruction of past evaluations. Subsequent proposals also depend on the generator state (Appendix A.7).

A.3 Same Archive, Different Search States

Construction and ranking consequence. Consider two references with equal target shares and a result archive whose two cells coincide with their hard-assignment regions. If their marginal weights are given by a common strictly decreasing function of support, equal-length histories can induce the same elite archive and opposite rankings of two fixed candidate behaviors.

Proof. Start both histories with the same two evaluations, one in each result cell, each of quality $q_\star$. Their support is $(1, 1)$. In history H, append $n \geq 1$ evaluations assigned to the second reference, all strictly below $q_\star$. In H', append n such evaluations assigned to the first. All additions are rejected by the result archive. Both histories have length $N = n + 2$ and the same retained individuals, but

$$\mathbf{S}(H) = (1, n + 1), \qquad \mathbf{S}(H') = (n + 1, 1). \tag{A.5}$$

Fix two candidates with projections e_1 and e_2. With strictly decreasing marginal weight function g, their MSG values are $g(1), g(n + 1)$ under H and $g(n + 1), g(1)$ under H', giving opposite strict orderings. The adaptive scales used here are permutation-invariant functions of the positive supports, so these two states also yield the same scale. This proves the claim. □

The two histories identify information lost under elite retention: evaluations with no archive improvement still change target support and subsequent candidate credit. When both candidates are compared together, equal quality and $\lambda > 0$ also reverse their ordering under a_t. Appendices C.2 and C.3 evaluate cumulative history and candidate-specific MSG in complete search runs.

A.4 Target Calibration and Divergence Interpretation

Let $\psi_t : [0, \infty) \to \mathbb{R}$ be increasing, continuously differentiable, and strictly concave, and let $\nu_\pi = \sum_j \pi_j \delta_{r_j}$ have positive reference shares summing to one. For $\mathcal{S} = \sum_j S_j \delta_{r_j}$, write $U_t(\mathcal{S})$ for the coordinate expression $U_t(\mathbf{S}; \mathcal{T})$ in Eq. (6). Its measure form is

$$U_t(\mathcal{S}) = \int \psi_t\left(\frac{d\mathcal{S}}{d\nu_\pi}\right) d\nu_\pi = \sum_j \pi_j \psi_t(S_j/\pi_j). \tag{A.6}$$

Derivatives at zero are right derivatives. The utility function is kept unchanged throughout each comparison.

Proposition A.1 (Target-measure calibration) *For every nonnegative support measure $\mathcal{S}$ on $\mathcal{R}$ with total support N,*

$$U_t(\mathcal{S}) \leq \psi_t(N), \qquad U_t(\mathcal{S}) = \psi_t(N) \iff \mathcal{S} = N\nu_\pi.$$

Thus $N\nu_\pi$ uniquely maximizes the functional among measures with the same total support.

Proof. Weighted Jensen's inequality gives

$$U_t(\mathcal{S}) = \sum_j \pi_j \psi_t\left(\frac{S_j}{\pi_j}\right) \leq \psi_t\left(\sum_j S_j\right) = \psi_t(N). \tag{A.7}$$

Strict concavity and positive reference shares make equality equivalent to $S_j/\pi_j = N$ for every j, or $\mathcal{S} = N\nu_\pi$. For $N = 0$, the zero measure is the only admissible state. □ The attainable allocations depend on the reference kernel and candidate generator.

For a divergence interpretation at total support $N > 0$, let $\eta = \mathcal{S}/N$ and define

$$f_{t,N}(u) = \psi_t(N) + N\psi_t'(N)(u - 1) - \psi_t(Nu). \tag{A.8}$$

This function is convex, nonnegative, and zero at $u = 1$. Since η and ν_π are probability measures, the affine term integrates to zero, giving

$$D_{f_{t,N}}(\eta \| \nu_\pi) = \int f_{t,N}\left(\frac{d\eta}{d\nu_\pi}\right) d\nu_\pi = \psi_t(N) - U_t(\mathcal{S}). \tag{A.9}$$

Thus the divergence measures the utility gap from the desired allocation at fixed total support. Its definition changes when either total support or the utility scale changes.

A.5 MSG AND THE VALUE OF ADDITIONAL SUPPORT

Proposition A.2 (Variational characterization of MSG) *Fix the current cumulative state $\mathcal{S}_t$ and utility function ψ_t. A candidate with behavior z contributes $K_R(z,\cdot)$, and*

$$\frac{d}{d\epsilon} U_t(\mathcal{S}_t + \epsilon K_R(z,\cdot))\Big|_{0^+} = \sum_j \psi_t'\left(\frac{S_{t,j}}{\pi_j}\right) w_j(z) = \mathrm{MSG}_t(z).$$

Thus MSG is the first-order value of adding that candidate's support to the cumulative state.

Proof. In coordinates, the path has entries $S_{t,j} + \epsilon w_j(z)$. Differentiating the finite sum in Eq. (A.6) gives

$$\frac{d}{d\epsilon} \sum_j \pi_j \psi_t\left(\frac{S_{t,j} + \epsilon w_j(z)}{\pi_j}\right)\Bigg|_{0^+} = \sum_j \psi_t'\left(\frac{S_{t,j}}{\pi_j}\right) w_j(z).$$

The factors π_j cancel, establishing Eq. (8). □ Equivalently,

$$\mathrm{MSG}_t(z) = \int_{\mathcal{R}} \psi_t'\left(\frac{d\mathcal{S}_t}{d\nu_\pi}(r)\right) K_R(z, dr).$$

The same projection weights accumulate target support and combine target-level marginal values into a candidate score.

For a step $\epsilon \geq 0$, let $\Delta_t(\epsilon, z) = U_t(\mathcal{S}_t + \epsilon K_R(z,\cdot)) - U_t(\mathcal{S}_t)$, using the current utility function throughout. Concavity gives

$$\epsilon \sum_j \psi_t'\left(\frac{S_{t,j} + \epsilon w_j(z)}{\pi_j}\right) w_j(z) \leq \Delta_t(\epsilon, z) \leq \epsilon\, \mathrm{MSG}_t(z). \tag{A.10}$$

One additional evaluation corresponds to $\epsilon = 1$. Thus MSG is the marginal value of the added support and bounds its finite-step utility gain from above.

A.6 EXPONENTIAL MARGINAL WEIGHTS

For $p > 0$ and $\tau_t > 0$, choose

$$\psi_t(u) = \int_0^u \exp\left[-\left(\frac{v}{M\tau_t}\right)^p\right] dv, \qquad \psi_t'(u) = \exp\left[-\left(\frac{u}{M\tau_t}\right)^p\right]. \tag{A.11}$$

The derivative is positive and strictly decreasing, making ψ_t increasing and strictly concave. For $p = 1$, $\psi_t(u) = M\tau_t(1 - e^{-u/(M\tau_t)})$. With the equal reference shares used in the main benchmarks, $\pi_j = 1/M$, Eq. (8) becomes

$$\frac{S_{t,j}/\pi_j}{M\tau_t} = \frac{S_{t,j}}{\tau_t}, \qquad \mathrm{MSG}_t(z) = \sum_j \exp[-(S_{t,j}/\tau_t)^p] w_j(z). \tag{A.12}$$

The scale is computed from the pre-update state and kept unchanged when comparing candidate contributions. At $\mathbf{S}_0 = 0$, every reference has marginal weight one, so every valid behavior receives MSG equal to one.

Task-specific scale choices are given in Appendix B.1. Appendix C.8 applies Eq. (A.11) with unequal target shares.

A.7 Rank Calibration and Generator Interface

Let $\mathcal{C}_t$ be a comparison set of size L_t and $s_t(x) = \mathrm{MSG}_t(m(x))$. If $L_t > 1$ and at least two values differ, define

$$r_t(x) = \mathrm{Rank01}_{\mathcal{C}_t}(s_t(x)) = \frac{\mathrm{rank}_{\mathrm{mid}}(s_t(x)) - 1}{L_t - 1}. \tag{A.13}$$

Ranks are ascending and one-based, with tied values sharing their average rank. Set $r_t(x) = 0$ when all scores tie or $L_t \leq 1$. Rank01 maps MSG ranks to a common scale for combination with raw quality.

Algorithm A1 gives QDEM search with equal target shares and the default shape $p = 1$. The generator determines the comparison set and how scores a_t update its proposals (Appendix B). Every new valid evaluation enters the state once, after scoring. The result archive separately applies its retention rule to evaluated candidates and returns the repertoire. QDEM scoring and generator updates operate independently of archive contents.

Algorithm A1 QDEM Search

Input: Evaluator, reference projection w, generator E_0, weight λ, budget
$\mathbf{S}_0 \leftarrow 0$, $N_0 \leftarrow 0$, $t \leftarrow 0$
while evaluation budget remains **do**
  $\mathcal{B}_t \leftarrow \mathrm{Propose}(E_t)$
  Evaluate $\mathcal{B}_t$; keep valid evaluations with quality $q(x)$ and behavior $m(x)$
  $\mathcal{C}_t \leftarrow \mathrm{ComparisonSet}(E_t, \mathcal{B}_t)$
  $\tau_t \leftarrow \mathrm{SupportScale}(\mathbf{S}_t)$ ▷ Appendix A.6
  $s_t(x) \leftarrow \sum_j e^{-S_{t,j}/\tau_t} w_j(m(x))$, for $x \in \mathcal{C}_t$
  $r_t(x) \leftarrow \mathrm{Rank01}_{\mathcal{C}_t}(s_t(x))$, for $x \in \mathcal{C}_t$
  $a_t(x) \leftarrow q(x) + \lambda r_t(x)$, for $x \in \mathcal{C}_t$
  $E_{t+1} \leftarrow \mathrm{UpdateGenerator}(E_t, \mathcal{C}_t, a_t)$
  $\mathbf{S}_{t+1} \leftarrow \mathbf{S}_t + \sum_{x \in \mathcal{B}_t} w(m(x))$
  $N_{t+1} \leftarrow N_t + |\mathcal{B}_t|$
  $t \leftarrow t + 1$
end while
return $E_t, \mathbf{S}_t$

For $\lambda > 0$, a lower-quality y is preferred to x under a_t precisely when

$$q(x) - q(y) < \lambda[r_t(y) - r_t(x)]. \tag{A.14}$$

Support can therefore offset a quality disadvantage of at most λ. The generator determines how favored candidates shape later search, while the result archive applies its own retention rule.

Adding support increases utility when the utility function is fixed. Across updates, task quality and allocation divergence need not improve monotonically as the support scale and generator state change.

B Experimental Setup and Implementation

This appendix gives the shared and task-specific settings for Section 5.

B.1 Shared Protocol and Experimental Suite

Within each task, methods use the same objective, behavior representation, reporting archive, evaluation budget, and seed set. Tables 2 and 3 report means. Appendix D gives the complete mean $\pm$ sample standard deviation statistics. Learning curves show the mean with a ± 1 sample-standard-deviation band. Standard comparisons use 20 seeds (42–61), while the visual comparisons and ablations use five seeds (42–46).

Table B1: Experimental suite and evaluation budgets.

Experiment	Tasks	Seeds	Budget per run
Controlled continuation	3D two-cell valley	42--46	10,800 batch evaluations + 1 initialization
Standard numerical	Sphere 2/10/20/50D; Rastrigin 2/10D; Flat 2/10D	42--61	5.4M evaluations
Visual targets	TA-MNIST; TA-F-MNIST; Hiker	42--46	1.8M (TA); 360k (Hiker)
Robot morphology	EvoGym Walker-v0	42--46	10,000 generations; 20 offspring/generation

Table B2: QDEM support settings and emitter defaults. Task-specific settings follow below, with robots in Appendix B.5.

Parameter	Setting
History	All valid evaluations
Batch update	Score the batch first; add the full batch to history afterward
Local projection (controlled/numerical/visual)	Gaussian weights on k nearest references; unit-normalized support
Marginal weight	Exponential diminishing-return weight
Support scale (controlled/numerical/visual)	Median positive support; floor 10^{-8}
Marginal weight shape	1.0
Emitter selection (numerical/visual)	μ-selection
Restart memory (numerical/visual)	22,500 recent evaluations plus 7,500 highest-quality evaluations; uniform sampling from their union

The numerical and visual implementations use CMA-ES (Hansen, 2016) to update their Gaussian search distributions.

CMA-ES settings. At restart, QDEM initializes the mean by uniformly sampling the memory union in Table B2, counting duplicate entries once. Covariance, evolution paths, and step size reset, while cumulative support is preserved. "Basic" denotes termination checks for ill-conditioned covariance, vanishing search scale, or negligible ranking differences. Numeric intervals specify additional periodic restarts in emitter updates.

Reference projection and bandwidth. Distances in Eq. (A.1) are Euclidean distances in the coordinates used by the support state. Numerical descriptors and the valley progress coordinate are mapped linearly from their specified ranges to $[-1, 1]$. TA pixels use $p \mapsto 2p - 1$. Hiker uses unit-normalized CLIP embeddings directly. All tasks use full-dimensional descriptors.

The tables report neighborhood size k. Numerical and visual tasks use $k_h = k$ to compute bandwidth, except Hiker ($k_h = 4096$). With $d_{k_h}(r)$ the distance to the k_hth nearest reference other than r, we set

$$h = \max\{\text{median}_{r \in \mathcal{Q}}\, d_{k_h}(r),\ 0.08\}\,, \tag{B.1}$$

where $\mathcal{Q}$ contains all references when $M \leq 4096$ and otherwise 4096 evenly spaced rows of the reference array. The bandwidth is fixed throughout each run. Valley sets $h = 0.32$ directly.

Support scale and comparison sets. The scale in Table B2 uses 10^{-8} at initialization. Robots instead use mean positive support with floor 10^{-6} and initial value 1. MSG ranks are computed over the full scheduler batch (540 numerical, 180 TA, and 36 Hiker or valley candidates). Each CMA-ES emitter then updates from its own candidates. Robot selection compares the complete previous survivor pool and new offspring, as specified in Appendix B.5. The support state remains unchanged during each comparison.

B.2 Controlled Continuation Benchmark

The controlled task uses $x = (s, z_1, z_2) \in [0, 8] \times \mathbb{R}^2$ and the quality function in Section 5.2, with centerline profile given by Equation (10). Table B3 specifies initialization, references, and budget. All methods share one initialization evaluation in addition to the 300 batches. Reporting assigns ties at $s = 6$ to C_0. A forward proposal has s above the optimizer mean before the batch update, or above the preceding batch mean for MAP-Elites. Section 5.2 defines the remaining metrics.

Table B3: Controlled-valley search settings.

Item	Setting
Decision space	3D: progress coordinate s and two lateral coordinates
Initialization	(1, 0, 0)
Bounds	$0 \leq$ s ≤ 8; lateral coordinates unbounded
Result archive	Nearest-site assignment to $s = 4, 8$: $C_0 : s \leq 6$, $C_1 : s > 6$
QDEM references	Cell-interval midpoints $s = 3, 7$, encoded as -0.25,0.75 by $z = s/4 - 1$, as for candidates
QDEM kernel	Normalized Gaussian weights over both references; $h = 0.32$ in encoded coordinates
Quality-support coefficient	$\lambda = 0.50$
Initial search scale	$\sigma_0 = 0.25$
Batch and iterations	36 candidates $\times$ 300 iterations
Seeds	42--46

Table B4: Controlled-valley baseline settings.

Method	Search/update signal	Key settings
MAP-Elites	Gaussian mutation from archive elites	Batch 36; $\sigma = 0.25$
CMA-ME	One CMA-ES emitter ranked by archive improvement	Batch 36; $\sigma_0 = 0.25$; μ-selection
CMA-MAE	One CMA-ES emitter ranked by threshold improvement	Batch 36; $\sigma_0 = 0.25$; μ-selection; threshold learning rate 0.10
DMS	One CMA-ES emitter ranked by learned continuous discount	Batch 36; $\sigma_0 = 0.25$; MLP 1--64--64--1, ReLU; Adam 10^{-3}; batch 32; $\leq$5 epochs/iteration

For all CMA-ES methods, no periodic restart occurs within the evaluation budget. Standard CMA-ES termination checks remain enabled.

B.3 Standard Numerical Benchmarks

All eight tasks use a 100-dimensional decision vector. The optimizer has no hard box constraint. Before behavior projection, each coordinate is transformed as

$$c(u) = \begin{cases} u, & |u| \leq 5.12, \\ 5.12/u, & \text{otherwise}, \end{cases} \tag{B.2}$$

The transformed coordinates are split into contiguous blocks and summed to form the behavior descriptor:

$$m_j(x) = \sum_{i=(j-1)b+1}^{jb} c(x_i), \qquad b = 100/d_m, \quad j = 1, \ldots, d_m. \tag{B.3}$$

Sphere and Rastrigin shift the optimum to 2.048 in every decision coordinate and normalize quality to one at the optimum. Sphere uses

$$q_S(x) = 1 - \frac{\sum_i (x_i - a)^2}{D(-5.12 - a)^2}. \tag{B.4}$$

Rastrigin uses

$$g_R(x) = 10D + \sum_i \left[(x_i - a)^2 - 10\cos\big(2\pi(x_i - a)\big) \right]. \tag{B.5}$$

$$q_R(x) = 1 - \frac{g_R(x)}{g_R(-5.12\,\mathbf{1})}. \tag{B.6}$$

Flat assigns constant quality 1.

Reference construction. For each behavior dimension d_m, 100,000 points are sampled uniformly from $[-512/d_m, 512/d_m]^{d_m}$ using the pyribs CVT sampling routine with seed 42. QDEM uses these samples, mapped to $[-1, 1]^{d_m}$, as equal-weight search references. They are shared across tasks of the same dimension and across run seeds. The reporting archive uses a 100×100 grid in 2D and 10,000 CVT cells otherwise (Vassiliades et al., 2018). Its cells are separate from the search references.

Table B5: QDEM settings for the eight standard numerical tasks.

Task	d	Refs	Neighbors k	λ	σ_0	Restart	Result archive
Sphere 2D	2	100k	256	.35	.5	basic	100×100 grid
Sphere 10D	10	100k	256	.35	.5	100	CVT-10k
Sphere 20D	20	100k	256	.35	.5	100	CVT-10k
Sphere 50D	50	100k	256	.35	.5	100	CVT-10k
Rastrigin 2D	2	100k	256	.35	.5	basic	100×100 grid
Rastrigin 10D	10	100k	256	.35	.5	100	CVT-10k
Flat 2D	2	100k	256	.35	.5	100	100×100 grid
Flat 10D	10	100k	256	.35	.5	100	CVT-10k

QDEM uses 15 emitters with batch size 36 for 10,000 iterations.

Table B6: Standard-task baseline settings.

Method	Generator and score	Key parameters
MAP-Elites	Gaussian mutation from result-archive elites	One emitter; batch 540; σ = 0.5; no restart
CMA-MAE	15 CMA-ES emitters ranked by threshold improvement	Batch 36/emitter; σ_0 = 0.5; μ-selection; threshold learning rate 0.01; basic in 2D, interval 100 otherwise
DDS	15 CMA-ES emitters ranked by negative KDE density	Batch 36/emitter; σ_0 = 1.5; μ-selection; buffer 10,000; basic restart; KDE h = 25.6/5.12/2.56/1.024 for 2/10/20/50D
DMS	15 CMA-ES emitters ranked by objective minus learned discount	Batch 36/emitter; σ_0 = 0.5; μ-selection; discount learning rate 0.10; basic in 2D, interval 100 otherwise

CMA-MAE restarts from a uniformly sampled solution in its threshold archive. DDS and DMS restart from uniformly sampled result-archive elites. They use the same CMA-ES termination checks described in Appendix B.1, in addition to their listed periodic restart rules.

DMS uses an MLP with layers d–128–128–1 and ReLU activations. Adam uses learning rate 10^{-3} with β = (0.9, 0.999). Training runs for at most five epochs per iteration with minibatch size 32 and early stopping at MSE equal to 0.05 times the task objective range. Measures are normalized to

$[-1, 1]$, discount targets to $[0, 1]$, and training starts from 1,000 archive points plus 100 empty-space points.

B.4 Visual Target Tasks

TA-MNIST and TA-F-MNIST use MNIST (LeCun et al., 1998) and Fashion-MNIST (Xiao et al., 2017), respectively, and parameterize 30 grayscale triangles. Each triangle has six vertex coordinates, one grayscale value, and one alpha value, giving 240 parameters. Images are rendered once at 28×28 on a black background with alpha scale 0.5 and no quantization. Behavior is the flattened 784-dimensional image, and quality is one minus the MSE to the nearest of 1,000 target images.

TA reference construction. Each dataset uses 1,000 images sampled without replacement from its 10,000-image test split (seed 42) for both quality evaluation and reporting cells. Pixel intensities lie in $[0, 1]$. QDEM uses the full test split as 10,000 equal-weight search references, transformed by $p \mapsto 2p - 1$. These reference sets remain fixed across runs.

Hiker objective. Hiker uses the StyleGAN3 checkpoint stylegan3-t-ffhqu-256x256.pkl (Karras et al., 2021) and the prompt "A photo of the face of a hiker." The 512-dimensional standardized code y is repeated across L synthesis layers, giving $\widetilde{y} = \mathbf{1}_L y^\top$ and $W(y) = \widetilde{y} \odot \sigma_w + \mathbf{1}_L w_{\text{avg}}^\top$. The mean and element-wise latent scale are estimated from 10,000 random StyleGAN samples. Let $e(y)$ and t be unit-normalized CLIP image and prompt embeddings (Radford et al., 2021). The prompt and latent losses are

$$L_{\text{prompt}}(y) = 2 \arcsin^2\left(\frac{\|e(y) - t\|_2}{2}\right), \qquad L_{\text{latent}}(y) = \left[\max(\|\widetilde{y}\|_F - r_0, 0)\right]^2, \tag{B.7}$$

where r_0 is 0.35 times the mean Frobenius norm of the standardized latent arrays in those 10,000 samples. Following the released DMS implementation, quality is

$$q_{\text{face}} = 1 - \tfrac{1}{2}(L_{\text{prompt}} + 0.2 L_{\text{latent}}), \quad d_{\text{ref}}(y) = \min_j [1 - e(y)^\top r_j], \quad q_{\text{ref}} = \tfrac{1}{2}(1 + d_{\text{ref}}), \quad q_{\text{Hiker}} = \tfrac{1}{2}(q_{\text{face}} + q_{\text{ref}}). \tag{B.8}$$

Hiker reference construction. All methods use a fixed subset of 10,000 LHQ images (Skorokhodov et al., 2021), sampled without replacement and encoded with CLIP ViT-B/32 into 512-dimensional embeddings normalized to unit length. The same embeddings define search references, the reference-distance quality term, and reporting centroids. Reporting uses nearest-reference cosine distance. QDEM's soft weights use Euclidean distance on the normalized embeddings.

Table B7: QDEM settings for the visual target tasks.

Task	Repr.	Refs	Neighbors k	λ	Emitters $\times$ batch	Iters.	Result archive
TA-MNIST	784D pixels	10k	64	.05	5×36	10k	1,000 targets
TA-F-MNIST	784D pixels	10k	64	.05	5×36	10k	1,000 targets
Hiker	512D CLIP	10k	160	.50	1×36	10k	CVT-10k

Table B8: Visual-task search scales and learning rates. The three CMA-ES methods share the listed initial scale; MAP-Elites uses a fixed Gaussian mutation scale.

Task	σ_0: QDEM / CMA-MAE / DMS	CMA-MAE lr	DMS lr	ME σ
TA-MNIST	0.10 / 0.10 / 0.10	0.001	0.001	0.10
TA-F-MNIST	0.10 / 0.10 / 0.10	0.10	0.10	0.10
Hiker	0.02 / 0.02 / 0.02	0.01	1.0	0.10

TA methods use 180 evaluations per iteration and Hiker uses 36. QDEM uses basic CMA-ES restarts on all three visual tasks. CMA-MAE and DMS add a restart every 50 updates on TA and use

basic restarts on Hiker. Their restart centers come from the threshold archive and result archive, respectively.

B.5 Robot Morphology Settings

We use EvoGym Walker-v0 (Bhatia et al., 2021) and the morphology–controller implementation of Mertan & Cheney (2025), without controller distillation. Table B9 gives the shared evaluation parameters and QDEM search settings.

Morphology, controller, and evaluation. A CPPN assigns empty, passive, or actuated material to each voxel. Bodies must be connected and contain at least three actuated voxels. Invalid or unchanged body mutations are resampled. Each voxel supplies five material indicators, volume, and two velocity features to the centralized MLP. Its output is $\mathrm{sigmoid}(h) - 0.4$, applied only to actuated voxels. Each offspring mutates either the body with the upstream CPPN operators or the controller with Gaussian parameter noise, retaining the other component. Fitness averages five rollout returns under the reward convention in Table B9. Controllers evolve with their bodies, without an inner training loop.

Reporting and search state. For active and existing voxel counts (a, e), reporting uses cell $(\min(\lfloor a/10 \rfloor, 9), \min(\lfloor e/10 \rfloor, 9))$. The constraint $a \leq e$ gives 55 feasible cells. The archive retains the highest fitness per cell, and QD Score sums these values. QDEM accumulates support separately using indices $(\min(\lfloor a/5 \rfloor, 19), \min(\lfloor e/5 \rfloor, 19))$. Every valid evaluation contributes once, regardless of archive acceptance.

Search population and Quality-only. QDEM starts from 20 random valid individuals and maintains a pool $\mathcal{P}_t$ of up to 60 survivors. Each generation samples 20 parents uniformly from this pool without replacement (with replacement if fewer are available) and produces 20 offspring $\mathcal{O}_t$. Selection compares the entire previous survivor pool with the new offspring:

$$\mathcal{C}_t = \mathrm{Unique}(\mathcal{P}_t \cup \mathcal{O}_t), \qquad |\mathcal{P}_t| \leq 60, \quad |\mathcal{O}_t| = 20.$$

Here Unique counts each individual once. MSG is ranked over $\mathcal{C}_t$, and the top $\min(60, |\mathcal{C}_t|)$ candidates under $q + 6r$ form the next pool. Selection uses the pre-update support state, after which only new evaluations add support. Quality-only uses the same population, sampling, variation, selection capacity, and evaluation protocol with $\lambda = 0$. MAP-Elites samples and retains archive elites.

Table B9: Robot evaluation and search settings.

Setting	Value
Environment and body	Legacy EvoGym Walker-v0; CPPN; 10 × 10 voxels
Controller	800--32--100 MLP; ReLU hidden layer
Mutation	Body/controller: 0.5/0.5; controller noise SD 0.1
Evaluation	Five rollouts; at most 500 decisions; action repeat 4
Evaluation noise	Observation and action noise SD 0.01
Reward convention	−0.05 per decision; +6.0 per rollout
Result archive	10 × 10; 55 feasible cells; strict elite replacement
History state	20 × 20 hard assignment; one unit per evaluation
Search pool / offspring	Up to 60 survivors / 20 evaluated offspring per generation
Quality / support coefficient	Raw fitness / λ = 6.0
Selection	Top 60 from the deduplicated previous survivor pool and new offspring
Replay and restarts	Disabled
Budget	10,000 generations including initialization; 200,000 candidates; 1,000,000 rollouts
Randomness	Run seeds 42--46; environment seed 17

The robot figure uses saved bodies and controllers replayed at a common spatial scale. Reported fitness comes from the original five-rollout evaluations.

C SEARCH-STATE ANALYSES AND ABLATIONS

These analyses examine evaluation allocation, cumulative history, candidate-specific MSG, and morphology–controller recovery. They also examine sensitivity to the quality–support balance, reference configuration, and choice of valuation, followed by allocation under unequal target shares. Unless noted otherwise, results use seeds 42–46 and the task budgets in Appendix B.

C.1 ALLOCATION DYNAMICS ON HIKER

This analysis relates evaluation allocation to the discoveries and repertoire quality obtained from a fixed Hiker budget. QDEM and DMS use the same 10,000 LHQ reporting references and 360,000 evaluations per run, with seeds 42–46. Each evaluated image is assigned to its nearest reference by cosine distance, $j = \arg\min_\ell [1 - e(x)^\top r_\ell]$, so reference j and reporting cell j identify the same region.

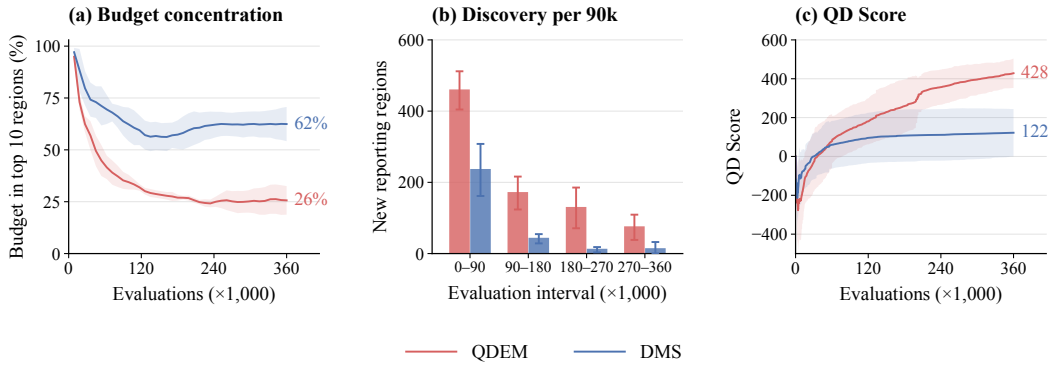


Figure C1: Hiker allocation and the results obtained from the evaluation budget. (a) Cumulative budget share assigned to the ten most-visited reporting regions. (b) Newly occupied reporting regions in each of four consecutive 90,000-evaluation intervals. (c) QD Score throughout the run. Curves and bars show means over five seeds. Bands and error bars show ± 1 sample SD, clipped to feasible bounds in (a,b). QD Score retains negative values without clipping. Endpoint labels are rounded means.

Allocation and outcomes. Budget concentration is the sum of the ten largest cumulative region visit counts divided by the evaluations used so far, recorded every 9,000 evaluations. It describes how concentrated the search is. To assess what the budget achieves, we divide the full run into four consecutive intervals of 90,000 evaluations. If $C(n)$ is the number of occupied reporting cells after n evaluations, interval b contributes $C(90{,}000b) - C(90{,}000(b-1))$ new regions, for $b = 1, \ldots, 4$ and $C(0) = 0$. We also report QD Score at every batch, summing retained quality over occupied cells.

Results. QDEM allocates less budget to the most-visited regions and discovers more new regions in every equal-budget interval than DMS (Figure C1). It sustains discovery through the final quarter and achieves a higher final QD Score. These results show more effective use of the same Hiker evaluation budget.

C.2 PERSISTENT EVALUATION HISTORY

This comparison tests whether recent or high-quality evaluations can replace cumulative history. The variants differ only in the evaluations contributing to support. All score against the state available before the current iteration. A batch contains 540 evaluations on numerical tasks, 180 on TA tasks, or 36 on Hiker.

Recent 100 batches. Support contains the latest 100 completed scheduler batches across all emitters, up to 54,000 numerical, 18,000 TA, or 3,600 Hiker evaluations.

Top-quality memory. Support is computed from the 7,500 evaluations currently retained in the independent top-quality buffer (Appendix B.1).

Previous batch only. Scoring uses support from the preceding completed batch, with an empty state for the first batch.

Table C1: Coverage and QD Score with different history sources. Results are mean ± sample SD over seeds 42--46, using 5.4M numerical, 1.8M TA, or 360k Hiker evaluations per run.

(a) Coverage

Task	Full history	Recent 100 batches	Top-quality memory	Previous batch only
Rastrigin 2D	93.34 ± 0.55%	73.07 ± 0.98%	29.96 ± 2.95%	34.66 ± 0.66%
Rastrigin 10D	88.03 ± 1.08%	44.73 ± 2.52%	38.56 ± 1.69%	46.31 ± 1.18%
Flat 2D	82.13 ± 0.73%	62.43 ± 2.57%	23.92 ± 0.57%	29.39 ± 0.74%
Flat 10D	84.44 ± 1.11%	56.19 ± 0.41%	58.81 ± 1.04%	53.53 ± 0.53%
Sphere 2D	94.65 ± 0.55%	74.60 ± 0.96%	21.78 ± 1.05%	38.07 ± 0.36%
Sphere 10D	92.52 ± 0.64%	47.11 ± 1.87%	41.07 ± 2.65%	44.12 ± 0.96%
Sphere 20D	90.19 ± 0.39%	84.45 ± 0.77%	67.63 ± 3.20%	76.04 ± 1.80%
Sphere 50D	82.48 ± 0.24%	81.43 ± 0.27%	68.35 ± 0.72%	69.71 ± 0.54%
TA (MNIST)	99.98 ± 0.04%	99.18 ± 0.28%	67.50 ± 10.71%	50.76 ± 5.17%
TA (F-MNIST)	85.02 ± 0.62%	73.30 ± 0.80%	39.30 ± 5.41%	27.86 ± 1.67%
Hiker	8.30 ± 0.36%	3.40 ± 0.57%	5.26 ± 0.80%	0.88 ± 0.14%

(b) QD Score

Task	Full history	Recent 100 batches	Top-quality memory	Previous batch only
Rastrigin 2D	5,916.6 ± 36.1	4,581.8 ± 32.0	2,306.9 ± 218.6	2,482.5 ± 48.2
Rastrigin 10D	5,333.8 ± 93.8	2,668.5 ± 132.5	2,617.1 ± 113.1	3,054.6 ± 78.0
Flat 2D	8,212.8 ± 72.9	6,242.6 ± 256.9	2,391.8 ± 56.8	2,938.8 ± 73.9
Flat 10D	8,444.2 ± 110.6	5,618.6 ± 40.7	5,880.8 ± 104.3	5,352.8 ± 53.1
Sphere 2D	7,125.0 ± 17.4	5,785.1 ± 55.1	2,031.1 ± 112.3	3,300.4 ± 33.0
Sphere 10D	7,033.4 ± 65.7	3,506.5 ± 152.1	3,429.0 ± 220.4	3,623.4 ± 85.4
Sphere 20D	6,898.8 ± 41.7	6,464.7 ± 69.3	5,670.4 ± 268.3	6,367.9 ± 184.2
Sphere 50D	6,685.8 ± 31.0	6,763.5 ± 14.7	5,930.6 ± 58.4	6,006.5 ± 70.9
TA (MNIST)	962.0 ± 0.5	948.6 ± 2.5	641.9 ± 100.6	487.2 ± 48.7
TA (F-MNIST)	827.5 ± 5.9	710.8 ± 8.0	381.9 ± 53.0	272.3 ± 16.3
Hiker	428.34 ± 75.37	114.52 ± 93.54	204.11 ± 142.58	−49.89 ± 104.37

Full history gives the highest mean Coverage on every task. Restricting support to recent batches, top-quality evaluations, or the previous batch reduces coverage, supporting the use of cumulative evaluation history.

C.3 CANDIDATE-SPECIFIC MSG

This comparison tests whether accumulated support must be used to value each candidate's own target contributions. No MSG removes the support term while leaving the remaining algorithm unchanged. Shuffled MSG permutes the normalized MSG ranks within each emitter's 36 candidates before combining them with quality. It preserves the available rank values at each update and leaves the history-accumulation rule unchanged.

Table C2: Coverage and QD Score when MSG is removed or its candidate assignment is shuffled. Results are mean ± sample SD over seeds 42--46, using 5.4M numerical, 1.8M TA, or 360k Hiker evaluations per run.

(a) Coverage

Task	Full MSG	No MSG	Shuffled MSG
Rastrigin 2D	93.34 ± 0.55%	3.59 ± 0.13%	4.19 ± 0.31%
Rastrigin 10D	88.03 ± 1.08%	0.61 ± 0.04%	1.12 ± 0.08%
Flat 2D	82.13 ± 0.73%	3.68 ± 0.09%	6.25 ± 0.60%
Flat 10D	84.44 ± 1.11%	1.60 ± 0.09%	2.41 ± 0.31%
Sphere 2D	94.65 ± 0.55%	3.17 ± 0.10%	4.34 ± 0.15%
Sphere 10D	92.52 ± 0.64%	0.46 ± 0.02%	0.69 ± 0.03%
Sphere 20D	90.19 ± 0.39%	0.96 ± 0.04%	4.16 ± 0.37%
Sphere 50D	82.48 ± 0.24%	4.36 ± 0.08%	15.08 ± 0.80%
TA (MNIST)	99.98 ± 0.04%	9.74 ± 3.57%	10.08 ± 4.09%
TA (F-MNIST)	85.02 ± 0.62%	2.22 ± 0.62%	2.28 ± 0.61%
Hiker	8.30 ± 0.36%	0.89 ± 0.17%	0.56 ± 0.13%

(b) QD Score

Task	Full MSG	No MSG	Shuffled MSG
Rastrigin 2D	5,916.6 ± 36.1	312.0 ± 11.5	351.0 ± 26.5
Rastrigin 10D	5,333.8 ± 93.8	51.7 ± 3.3	89.3 ± 6.0
Flat 2D	8,212.8 ± 72.9	368.2 ± 9.5	625.4 ± 59.8
Flat 10D	8,444.2 ± 110.6	159.8 ± 9.3	241.0 ± 31.1
Sphere 2D	7,125.0 ± 17.4	306.2 ± 9.2	416.1 ± 14.3
Sphere 10D	7,033.4 ± 65.7	44.5 ± 2.1	66.4 ± 2.9
Sphere 20D	6,898.8 ± 41.7	91.7 ± 3.9	383.1 ± 32.8
Sphere 50D	6,685.8 ± 31.0	412.5 ± 7.3	1,392.4 ± 70.2
TA (MNIST)	962.0 ± 0.5	92.3 ± 33.4	95.5 ± 38.5
TA (F-MNIST)	827.5 ± 5.9	21.5 ± 5.9	22.1 ± 5.9
Hiker	428.34 ± 75.37	−110.08 ± 173.40	−92.46 ± 93.83

Removing MSG or shuffling its candidate assignments sharply reduces Coverage, supporting the role of candidate-specific credit based on reference contribution and accumulated support.

C.4 Morphology–Controller Recovery

This analysis quantifies recovery and the role of support in the five QDEM runs in Table 4, using recorded lineages.

Recovery across seeds. We consider body mutants whose phenotype changes, whose fitness drops by at least one point relative to the direct parent, and which are rejected by the result archive but retained in the search population. Mutants must be born by generation 9,900 to allow a complete 100-generation observation window. Recovery means a descendant reaches or exceeds the pre-change parent's fitness within that window. Controller-only recovery requires all subsequent mutations along the path to change only the controller while preserving the new body.

Table C3: Recovery of retained, archive-rejected morphological intermediates within 100 generations. Lineage events can share ancestry. The five seeds are the independent experimental repeats.

Seed	Eligible intermediates	Recovered	Controller-only recovery
42	3,234	655	214
43	1,715	344	119
44	1,977	369	134
45	2,919	377	77
46	3,080	435	89
Total	12,925	2,180	633

Support changes continuation decisions. Of the 2,180 recovered intermediates, 1,400 would have been excluded by quality-only ranking when they entered the search population. This comparison keeps the observed candidate set and survivor capacity fixed. Table 4 evaluates Quality-only over complete runs.

Recovery within one search reference. A seed-42 lineage reaches $10.162 \rightarrow 7.436 \rightarrow 9.106 \rightarrow 10.567$ while remaining at 81 active and 82 existing voxels, within the same reporting cell and search reference. After the initial body change, only the controller mutates. The intermediate falls below the archive incumbent (10.377) and the quality-only selection cutoff, yet QDEM retains it and its descendant improves the archive.

C.5 Quality–Support Balance

The acquisition combines task quality with search value from cumulative support. Tables C4–C5 examine how this balance affects coverage and repertoire quality, using the 100k-reference configuration in Appendix B.3. Only the coefficient λ changes.

Table C4: Coverage across quality--support weights (mean ± sample SD over five seeds).

Task	$\lambda = 0.10$	$\lambda = 0.20$	$\lambda = 0.35$	$\lambda = 0.50$	$\lambda = 1.00$
Rastrigin 2D	42.83 ± 2.83%	63.45 ± 4.32%	93.34 ± 0.55%	94.32 ± 0.56%	92.24 ± 0.62%
Rastrigin 10D	28.02 ± 1.69%	73.18 ± 3.32%	88.03 ± 1.08%	89.28 ± 0.59%	88.79 ± 1.14%
Flat 2D	82.13 ± 0.73%	82.13 ± 0.73%	82.13 ± 0.73%	82.13 ± 0.73%	82.13 ± 0.73%
Flat 10D	84.44 ± 1.11%	84.44 ± 1.11%	84.44 ± 1.11%	84.44 ± 1.11%	84.44 ± 1.11%
Sphere 2D	52.33 ± 1.70%	77.35 ± 2.19%	94.65 ± 0.55%	95.79 ± 0.32%	93.08 ± 1.70%
Sphere 10D	38.32 ± 1.70%	90.77 ± 0.49%	92.52 ± 0.64%	91.28 ± 0.73%	88.86 ± 0.48%
Sphere 20D	48.95 ± 2.05%	89.10 ± 0.33%	90.19 ± 0.39%	89.83 ± 0.12%	89.10 ± 0.44%
Sphere 50D	59.58 ± 2.00%	81.12 ± 0.43%	82.48 ± 0.24%	82.62 ± 0.15%	82.84 ± 0.36%

Table C5: QD Score across quality--support weights (mean ± sample SD over five seeds).

Task	$\lambda = 0.10$	$\lambda = 0.20$	$\lambda = 0.35$	$\lambda = 0.50$	$\lambda = 1.00$
Rastrigin 2D	3,508.6 ± 190.7	4,760.6 ± 244.0	5,916.6 ± 36.1	5,798.6 ± 21.3	5,380.7 ± 42.4
Rastrigin 10D	1,959.5 ± 133.8	4,748.2 ± 194.5	5,333.8 ± 93.8	5,229.2 ± 36.1	4,913.0 ± 81.6
Flat 2D	8,212.8 ± 72.9	8,212.8 ± 72.9	8,212.8 ± 72.9	8,212.8 ± 72.9	8,212.8 ± 72.9
Flat 10D	8,444.2 ± 110.6	8,444.2 ± 110.6	8,444.2 ± 110.6	8,444.2 ± 110.6	8,444.2 ± 110.6
Sphere 2D	4,719.3 ± 128.4	6,381.4 ± 127.1	7,125.0 ± 17.4	7,022.2 ± 19.7	6,602.0 ± 102.8
Sphere 10D	3,301.2 ± 158.1	7,306.4 ± 52.3	7,033.4 ± 65.7	6,720.0 ± 78.8	6,212.2 ± 86.5
Sphere 20D	4,177.5 ± 184.1	7,165.1 ± 22.4	6,898.8 ± 41.7	6,685.4 ± 20.9	6,326.7 ± 72.7
Sphere 50D	5,206.9 ± 179.9	6,799.9 ± 38.7	6,685.8 ± 31.0	6,628.0 ± 19.1	6,571.8 ± 45.7

Flat gives identical results because its constant quality leaves candidate ordering determined by MSG for every positive λ. On Sphere and Rastrigin, increasing support weight above $\lambda = 0.10$ improves coverage, while the largest weight gives lower QD Scores than intermediate values. These results illustrate the trade-off between broader exploration and retained quality.

C.6 Reference Configuration

The reference configuration determines how evaluation effort is represented across targets. Tables C6–C7 compare nested reference sets with proportional neighborhoods at $\lambda = 0.35$. The bandwidth uses the same neighbor index, $k_h = k$ (Appendix B.1), so the comparison examines reference count and neighborhood size jointly.

Table C6: Coverage across reference configurations (mean ± sample SD over five seeds).

Task	**12.5k** $k = 32$	**25k** $k = 64$	**50k** $k = 128$	**100k** $k = 256$
Rastrigin 2D	88.87 ± 1.99%	90.57 ± 1.31%	91.60 ± 0.85%	93.34 ± 0.55%
Rastrigin 10D	68.79 ± 0.90%	79.03 ± 0.95%	85.00 ± 0.45%	88.03 ± 1.08%
Flat 2D	81.76 ± 0.91%	82.06 ± 0.96%	82.14 ± 0.75%	82.13 ± 0.73%
Flat 10D	77.80 ± 1.08%	80.87 ± 0.77%	82.04 ± 1.11%	84.44 ± 1.11%
Sphere 2D	94.21 ± 0.94%	94.20 ± 1.40%	94.47 ± 0.96%	94.65 ± 0.55%
Sphere 10D	76.40 ± 1.13%	85.84 ± 0.48%	90.33 ± 0.34%	92.52 ± 0.64%
Sphere 20D	79.26 ± 0.54%	84.83 ± 0.35%	88.27 ± 0.32%	90.19 ± 0.39%
Sphere 50D	72.75 ± 0.41%	76.80 ± 0.30%	79.67 ± 0.26%	82.48 ± 0.24%

Table C7: QD Score across reference configurations (mean ± sample SD over five seeds).

Task	**12.5k** $k = 32$	**25k** $k = 64$	**50k** $k = 128$	**100k** $k = 256$
Rastrigin 2D	5,756.5 ± 62.2	5,813.2 ± 76.0	5,791.5 ± 57.0	5,916.6 ± 36.1
Rastrigin 10D	4,416.4 ± 61.0	4,922.3 ± 72.1	5,195.6 ± 25.0	5,333.8 ± 93.8
Flat 2D	8,175.6 ± 91.3	8,205.8 ± 95.6	8,213.6 ± 75.0	8,212.8 ± 72.9
Flat 10D	7,780.4 ± 108.3	8,087.4 ± 76.9	8,204.4 ± 111.0	8,444.2 ± 110.6
Sphere 2D	7,097.1 ± 22.4	7,096.1 ± 44.0	7,109.3 ± 30.3	7,125.0 ± 17.4
Sphere 10D	6,085.3 ± 93.1	6,653.9 ± 18.3	6,909.3 ± 38.8	7,033.4 ± 65.7
Sphere 20D	6,425.3 ± 37.3	6,677.2 ± 47.5	6,825.9 ± 43.7	6,898.8 ± 41.7
Sphere 50D	6,143.4 ± 36.5	6,378.1 ± 32.5	6,523.4 ± 25.7	6,685.8 ± 31.0

Larger reference sets and neighborhoods improve both metrics on Sphere 10D/20D/50D, Rastrigin 10D, and Flat 10D, whereas Flat 2D changes little. The benefit of a larger reference configuration is therefore task-dependent.

C.7 Alternative Target-Relative Valuations

QDEM separates the accumulated support representation from the function that assigns it search value. To examine this separation, we compare exponential $g(u) = e^{-u}$, reciprocal $g(u) = (1 + u)^{-1}$, and linear-hinge $g(u) = \max(1 - u, 0)$ marginal weights, where $u = S_{t,j}/\tau_t$. The projection, accumulation rule, support-scale rule, generator, and quality coefficient remain unchanged.

Table C8: Coverage under alternative marginal valuations (mean ± sample SD over five seeds).

Task	Exponential	Reciprocal	Linear hinge
Rastrigin 2D	93.34 ± 0.55%	91.71 ± 1.38%	94.36 ± 0.77%
Rastrigin 10D	88.03 ± 1.08%	85.51 ± 0.34%	35.47 ± 6.10%
Flat 2D	82.13 ± 0.73%	82.39 ± 0.35%	82.90 ± 0.75%
Flat 10D	84.44 ± 1.11%	84.55 ± 1.07%	38.15 ± 2.32%
Sphere 2D	94.65 ± 0.55%	94.84 ± 0.54%	96.18 ± 0.73%
Sphere 10D	92.52 ± 0.64%	91.74 ± 0.58%	41.38 ± 3.04%
Sphere 20D	90.19 ± 0.39%	88.52 ± 0.32%	61.34 ± 3.83%
Sphere 50D	82.48 ± 0.24%	81.74 ± 0.37%	67.59 ± 1.92%
TA-MNIST	99.98 ± 0.04%	100.00 ± 0.00%	96.80 ± 2.48%
TA-F-MNIST	85.02 ± 0.62%	83.80 ± 1.06%	60.48 ± 8.95%
Hiker	8.30 ± 0.36%	7.51 ± 1.40%	9.29 ± 0.91%

Table C9: QD Score under alternative marginal valuations (mean ± sample SD over five seeds).

Task	Exponential	Reciprocal	Linear hinge
Rastrigin 2D	5,916.6 ± 36.1	5,837.2 ± 62.9	5,881.4 ± 38.5
Rastrigin 10D	5,333.8 ± 93.8	5,169.0 ± 30.3	2,137.4 ± 322.5
Flat 2D	8,212.8 ± 72.9	8,238.6 ± 35.1	8,290.4 ± 74.8
Flat 10D	8,444.2 ± 110.6	8,455.4 ± 107.0	3,815.4 ± 232.3
Sphere 2D	7,125.0 ± 17.4	7,126.0 ± 20.3	7,090.9 ± 26.7
Sphere 10D	7,033.4 ± 65.7	6,981.9 ± 63.3	3,052.9 ± 199.6
Sphere 20D	6,898.8 ± 41.7	6,792.6 ± 44.7	4,579.8 ± 299.7
Sphere 50D	6,685.8 ± 31.0	6,627.3 ± 44.4	5,323.1 ± 187.1
TA-MNIST	962.0 ± 0.5	962.4 ± 0.3	927.0 ± 25.1
TA-F-MNIST	827.5 ± 5.9	815.8 ± 10.3	588.2 ± 86.7
Hiker	428.3 ± 75.4	317.7 ± 190.8	477.9 ± 140.1

Exponential and reciprocal weights give similar results on most tasks, showing that the same state representation supports more than one effective valuation. Linear-hinge weights give higher mean Coverage and QD Score on Hiker but substantially reduce both metrics on several higher-dimensional and TA tasks. Their mixed results show how valuation changes search behavior while leaving the state representation intact.

C.8 Evaluation Allocation under Non-uniform Targets

We vary target shares on standard Flat 2D (Appendix B.3), retaining its 100-dimensional decisions, constant quality $q(x) = 1$, and $M = 100{,}000$ fixed search references. Other settings are unchanged: $k = k_h = 256$, $h = 0.08$ in normalized descriptor coordinates, $\lambda = 0.35$, 15 CMA-ES emitters with 36 candidates each, $\sigma_0 = 0.5$. Each condition uses seeds 42–46 and 5.4M evaluations per run.

Target shares. There are $N_L = 50{,}032$ references with first coordinate below zero and $N_R = 49{,}968$ in the other half. For left-half target mass $\rho \in \{0.25, 0.75\}$, set $\pi_j = \rho/N_L$ on the left and $\pi_j = (1 - \rho)/N_R$ on the right. The uniform condition keeps $\pi_j = 1/M$, giving an exact left-half mass of 50.032%, displayed as 50% in the table and figure. Each run keeps its target fixed. Using Eq. (A.11),

$$\mathrm{MSG}_t(z) = \sum_j \exp\left[-\frac{S_{t,j}}{M\pi_j\tau_t}\right] w_j(z),$$

where τ_t remains the median positive raw support with floor 10^{-8}. All valid evaluations contribute unit support. Only π changes. There is no target-dependent proposal sampling or allocation quota.

Results. Changing the left target from 25% to 75% shifts the fraction of all evaluations with first descriptor coordinate below zero by 52.20 ± 1.11 percentage points across paired seeds, with similar coverage (Table C10). Figure C2 retains boundary under-sampling and interface concentration. The experiment demonstrates allocation responsive to fixed targets, without establishing exact spatial matching.

Table C10: Non-uniform target response on standard Flat 2D (mean $\pm$ sample SD over five seeds). Allocation counts all 5.4M evaluations per run. Coverage uses the fixed 100×100 reporting archive.

Left target mass	**Evaluations in left half (%)**	**Coverage (%)**
25%	24.04 ± 0.94	81.63 ± 0.70
50%	50.14 ± 1.33	82.13 ± 0.73
75%	76.25 ± 1.33	81.75 ± 0.93

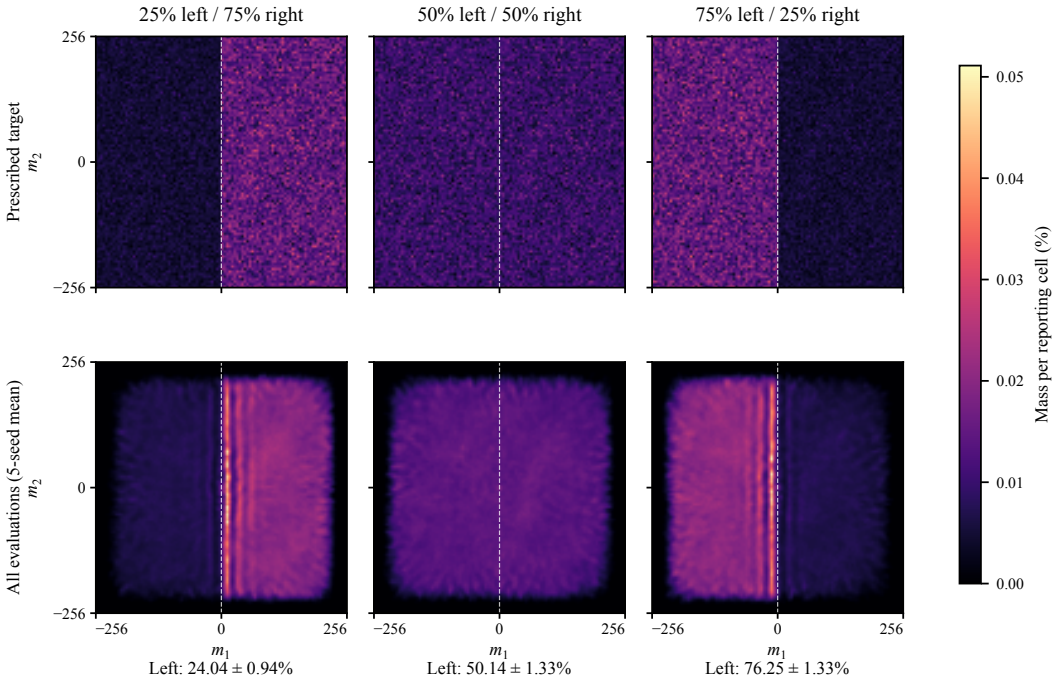


Figure C2: Target mass (top) and evaluation allocation (bottom) on Flat 2D. Bottom panels average per-run normalized histograms over five seeds, each with 5.4M evaluations. All panels use the 100×100 reporting grid, separate from search references, and a shared linear color scale without smoothing or clipping. Dashed lines divide left and right. Labels give left evaluation fractions (mean $\pm$ sample SD).

D Additional Experimental Results

This appendix complements the main comparisons with full statistics, learning curves, and expanded visual examples. The intervention and sensitivity results are grouped in Appendix C.

D.1 Standard-Benchmark Statistics and Learning Dynamics

Tables D1–D2 report final mean ± sample SD. Figures D1–D2 show the corresponding learning curves.

Table D1: Coverage on the eight standard benchmarks (mean ± sample SD over 20 seeds).

(a) Rastrigin and Flat

Method	Rastrigin 2D	Rastrigin 10D	Flat 2D	Flat 10D
QDEM	92.34 ± 1.69%	88.02 ± 0.71%	82.19 ± 0.73%	84.33 ± 0.93%
DMS	91.48 ± 2.13%	88.16 ± 0.66%	80.13 ± 2.10%	79.05 ± 1.52%
CMA-MAE	80.02 ± 0.33%	1.78 ± 0.20%	55.93 ± 3.33%	14.72 ± 1.38%
DDS	72.50 ± 4.02%	55.95 ± 1.92%	71.88 ± 3.22%	49.68 ± 1.95%
MAP-Elites	48.24 ± 0.35%	7.59 ± 0.97%	43.51 ± 0.73%	1.22 ± 0.68%

(b) Sphere

Method	2D	10D	20D	50D
QDEM	94.83 ± 0.78%	92.25 ± 0.43%	90.08 ± 0.34%	82.42 ± 0.25%
DMS	96.35 ± 1.94%	88.93 ± 0.64%	95.91 ± 0.27%	86.90 ± 0.20%
CMA-MAE	80.99 ± 0.32%	7.46 ± 1.39%	8.25 ± 1.43%	24.23 ± 0.93%
DDS	71.53 ± 3.53%	56.61 ± 1.55%	84.25 ± 0.23%	74.28 ± 0.24%
MAP-Elites	50.74 ± 0.37%	2.07 ± 0.63%	35.59 ± 2.41%	41.07 ± 0.73%

Table D2: QD Score on the eight standard benchmarks (mean ± sample SD over 20 seeds).

(a) Rastrigin and Flat

Method	Rastrigin 2D	Rastrigin 10D	Flat 2D	Flat 10D
QDEM	5,867.9 ± 73.3	5,328.6 ± 58.9	8,218.5 ± 73.1	8,433.3 ± 92.8
DMS	5,724.6 ± 86.0	5,129.0 ± 57.7	8,013.1 ± 210.4	7,904.6 ± 152.2
CMA-MAE	5,270.6 ± 21.9	156.0 ± 17.0	5,593.4 ± 332.6	1,472.5 ± 138.4
DDS	2,614.7 ± 346.9	2,848.0 ± 109.6	7,187.6 ± 322.0	4,967.9 ± 195.1
MAP-Elites	3,171.9 ± 21.5	532.3 ± 64.3	4,350.9 ± 73.2	122.2 ± 67.7

(b) Sphere

Method	2D	10D	20D	50D
QDEM	7,123.9 ± 21.2	7,005.6 ± 44.9	6,879.2 ± 37.1	6,676.6 ± 26.2
DMS	7,015.6 ± 82.0	6,387.8 ± 61.0	7,351.1 ± 107.4	6,985.2 ± 33.6
CMA-MAE	6,330.2 ± 21.5	715.5 ± 128.6	802.5 ± 135.3	2,329.8 ± 86.6
DDS	3,320.9 ± 594.4	3,724.5 ± 122.6	6,985.9 ± 20.9	6,370.1 ± 20.1
MAP-Elites	4,162.0 ± 27.2	201.1 ± 60.0	3,282.3 ± 209.4	3,887.7 ± 67.0

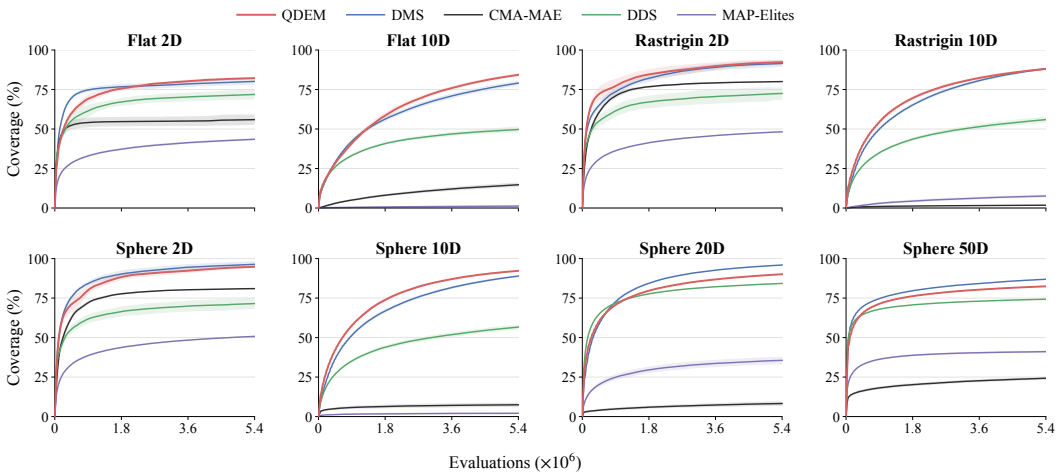


Figure D1: Coverage learning curves on the eight standard benchmarks. Solid lines show the mean across the 20 seeds. Shaded regions show ± 1 sample standard deviation.

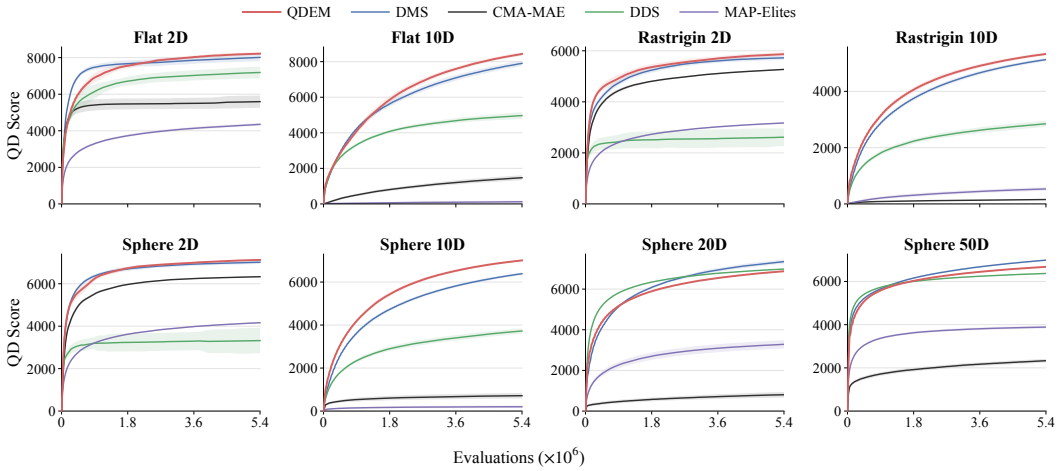


Figure D2: QD Score learning curves on the eight standard benchmarks. Solid lines show the mean across the 20 seeds. Shaded regions show ± 1 sample standard deviation.

The curves show sustained exploration on Rastrigin and Flat and the task-dependent advantage of DMS on higher-dimensional Sphere.

D.2 VISUAL-TASK STATISTICS

Table D3 gives the complete statistics underlying Table 3. Each entry uses five runs with seeds 42–46 and 10,000 iterations (1.8M evaluations for TA and 360k for Hiker).

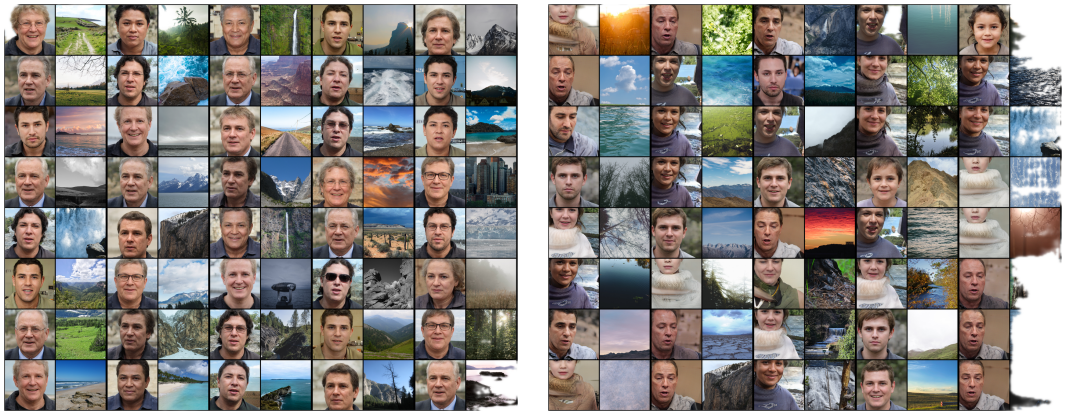
Table D3: Visual-task results: mean ± sample SD over five seeds. Coverage is in percent.

Task	**Method**	**QD Score**	**Coverage (%)**
TA-MNIST	QDEM	961.98 ± 0.48	99.98 ± 0.04
	DMS	953.13 ± 2.13	100.00 ± 0.00
	CMA-MAE	954.24 ± 1.87	99.44 ± 0.26
	MAP-Elites	940.71 ± 3.01	98.26 ± 0.34
TA-F-MNIST	QDEM	827.49 ± 5.94	85.02 ± 0.62
	DMS	710.43 ± 8.11	73.20 ± 0.85
	CMA-MAE	617.44 ± 12.87	63.06 ± 1.33
	MAP-Elites	535.92 ± 20.72	55.02 ± 2.16
Hiker	QDEM	428.34 ± 75.37	8.30 ± 0.36
	DMS	121.98 ± 122.08	3.00 ± 0.77
	CMA-MAE	-25.36 ± 128.32	1.41 ± 0.23
	MAP-Elites	$-19{,}295.61 \pm 6{,}793.34$	5.90 ± 1.48

D.3 Expanded Hiker Qualitative Comparison

Figure D3 expands Figure 2 using the final archive of seed 42. Each method samples up to 40 occupied cells uniformly without replacement and pairs each selected face with its assigned landscape reference. Displayed references vary with each method's occupied cells. Table D3 reports quality and coverage across seeds.

(a) QDEM **(b) DMS**



(c) CMA-MAE **(d) MAP-Elites**

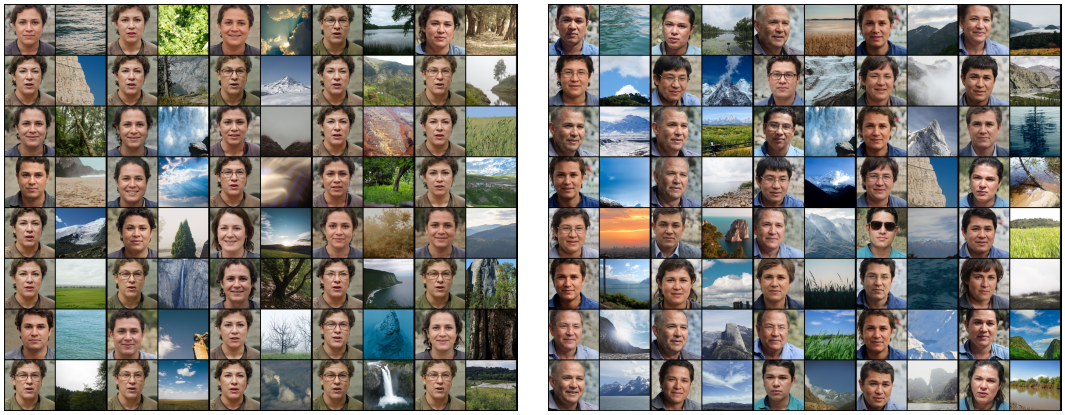


Figure D3: Expanded Hiker samples from the final archive of seed 42, using the same sampling rule for all methods. Each generated face is paired with its landscape reference.

E Extended Related Work

Selection mechanisms and repertoire containers. The modular QD framework separates containers from selection operators and discusses using different containers for exploration and the returned repertoire (Cully & Demiris, 2018). Curiosity scores parents through their offspring's success in entering a container (Cully & Demiris, 2018). Youth-based selection favors recent additions (Coninx & Doncieux, 2021), and Monte Carlo Elites balances selection counts with offspring survival through bandit rules (Sfikas et al., 2021). These statistics allocate reproduction among available parents. QDEM addresses the complementary question of what evaluation experience should persist independently of parent selection and result retention. Its fixed-dimensional support state accumulates target contributions from all valid evaluations.

Behavior representations and preferences. CVT-MAP-Elites controls repertoire size through a fixed number of Voronoi regions (Vassiliades et al., 2018). AURORA learns behavioral descriptors from observations (Cully, 2019). RUDA adjusts learned descriptor distances using downstream-task relevance (Grillotti & Cully, 2022), while QDHF learns diversity metrics from human similarity judgments (Ding et al., 2024). Behavior descriptors also determine which distinctions define diversity when searching generative latent spaces for game content (Fontaine et al., 2021). These works shape the representation or relevance of behavioral diversity. Given a representation and fixed projection, QDEM records cumulative evaluation effort at its targets in a fixed-dimensional support state.

Recursive state and history-dependent exploration. Agent-state methods recursively summarize observation–action history (Sinha et al., 2024), while information-state approaches connect history compression to evaluation and prediction requirements (Subramanian et al., 2022). This perspective motivates defining which quantities a representation must recover. QDEM specifies cumulative target-support totals and establishes their exact recovery under a fixed projection (Appendix A.2). In reinforcement learning, count-based and pseudo-count bonuses use visitation to guide exploration (Bellemare et al., 2016), and SOFE augments the state representation with sufficient statistics of exploration bonuses to create stationary exploration objectives (Creus Castanyer et al., 2024). Ergodic exploration compares time spent across locations with a prescribed spatial distribution (Mathew & Mezić, 2011). QDEM applies this perspective to black-box search by representing evaluation contributions to behavioral targets separately from the generator state.

Morphology–controller adaptation. Passive genetic diversity retains some lower-fitness individuals alongside elites to preserve potential stepping stones (Pugh et al., 2016). Morphological innovation protection temporarily reduces selection pressure after body changes to allow controller readaptation (Cheney et al., 2018). Development can expose useful morphologies and promote robustness to controller changes (Kriegman et al., 2018). Lineage analysis links diverse ancestors to successful modular robots (Nordmoen et al., 2021), while controller distillation improves transfer between morphologies and enables niche migrations (Mertan & Cheney, 2025). In QDEM, accumulated target support changes the selection value of intermediate designs, so archive-rejected morphologies can remain in the search population and produce descendants.