mathematics

MDPI

Article

Pareto-Guided Leader Selection in the Mountain Gazelle Optimizer: A Controlled Multi-Objective Study

Preut Thanarat, Suchart Khummanee and Chatklaw Jareanpon *

Department of Computer Science, Faculty of Informatics, Mahasarakham University, Maha Sarakham 44150, Thailand; preut.t@msu.ac.th (P.T.); suchart.k@msu.ac.th (S.K.)
* Correspondence: chatklaw.j@msu.ac.th

Abstract

Where should multi-objective reasoning enter a swarm optimizer: before the search through scalarization, around the search through survival or archiving, or inside the search through Pareto-informed leader selection? We study this question on the Mountain Gazelle Optimizer (MGO) and propose a feedback-path taxonomy that separates the three roles. A census of twenty multi-objective MGO extensions finds one prior leader-mediated design and no method that assigns per-agent leaders from a Pareto archive. We develop four leader-mediated variants on one search skeleton, with a random-leader control and a leader-isolation ablation. On twenty-two benchmark problems (thirty runs each), all four variants rank ahead of five paper-based reconstructions of prior MGO designs on three indicators, and the decomposition and indicator leader rules outrank random leadership on hypervolume and IGD$^+$ under Holm correction. On fourteen held-out engineering problems, a refined indicator-guided design separates from NSGA-II, NSGA-III, MOEA/D, AGE-MOEA-II and SMS-EMOA on IGD$^+$ under returned-set scoring, and from all but SMS-EMOA on hypervolume; under evaluation-trace scoring, only from NSGA-III and MOEA/D on IGD$^+$. In sample, of the five it separates only from MOEA/D (ε-indicator, returned-set scoring). The four variants outperform the prior MGO designs as complete implementations; against established algorithms, the advantage depends on scoring.

Keywords: multi-objective optimization; leader selection; Pareto archive; architectural taxonomy; null control; budget-fair benchmarking; two-channel evaluation; Mountain Gazelle Optimizer

MSC: 90C29

Check for updates

Academic Editor: Janez Žerovnik

Received: 29 September 2026
Revised: 22 October 2026
Accepted: 29 October 2026
Published: 4 November 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Engineering and scientific optimization problems routinely impose conflicting objectives, such as cost against emissions or throughput against risk. Their solutions form a Pareto front of mutually non-dominated trade-offs [1,2]. For the non-convex, mixed-variable, constrained instances met in practice, multi-objective evolutionary algorithms (MOEAs) are the standard tool. Three strategy families organize them: Pareto-based selection [3], decomposition [4] and indicator-based selection [5]. Each family places its multi-objective reasoning inside selection.

Swarm and nature-inspired metaheuristics are usually extended to multiple objectives by a different route. Since multi-objective particle swarm optimization (MOPSO) [6], the dominant pattern keeps the single-objective search loop and replaces the

single global best with a leader drawn from an external Pareto archive. The multi-objective grey wolf optimizer (MOGWO), multi-objective multi-verse optimization (MOMVO) and the multi-objective salp swarm algorithm (MSSA) [7–9] follow this leader-mediated template. The same body of work has attracted a sustained methodological critique: metaphor-based metaheuristics often repackage established mechanics behind new imagery [10], warrant structural analysis [11], are frequently indistinguishable at scale under statistical testing [12], and overlap structurally with earlier methods in roughly half of a large survey [13]. The instrument of this paper is therefore a structural, metaphor-free taxonomy, and its experiment asks where multi-objective reasoning belongs inside a swarm optimizer.

We use the Mountain Gazelle Optimizer (MGO) [14] as the experimental vehicle. The MGO is recent and widely applied: a systematic review counts eighty-nine applications by 2025 [15]. Its leader-mediated multi-objective design space is almost unexplored. The search applies four mechanisms to each agent in each iteration; three of them (territorial solitary males, maternity herds and bachelor male herds) consult one leader, *gazmale*, and the fourth, migration to search for food, does not. In the single-objective MGO, *gazmale* is the best-fitness agent. Under conflicting objectives many solutions are mutually non-dominated and the leader is undefined. Among the twenty multi-objective MGO works in our census, all but one avoid this question by placing their multi-objective machinery outside the update equations, either as scalarization [16–18] or as archive- and survival-based designs [19–22]. The exception, a 2026 multi-project scheduling method [23], broadcasts one non-dominated solution as the guide for all agents. It is the one leader-mediated design under the criterion of Section 2.5, and none of the twenty draws per-agent leaders from a Pareto archive.

Besides this gap, the MGO suits a controlled study of leader selection for three reasons. Its leader rule can be changed while its published movement equations stay intact, because three of its four update mechanisms read one leader symbol, *gazmale*. Its separate population-survival step allows the leader rule and the survival rule to be varied one at a time (Section 3.2), which the PSO frameworks of the leader studies in Section 2.2 do not allow. Its few base-algorithm settings can be held constant across all designs. We treat the MGO as the procedure its equations define (Section 3.1) and draw no argument from its metaphor; the results concern the MGO, and the parts of the design that carry over to other leader-based swarm optimizers are named with the contributions below.

This study introduces no new metaphor-inspired optimizer and leaves the algebraic search operators of the MGO unchanged. It compares the points at which Pareto information enters the search dynamics and, under common management, isolates the contribution of the leader rule. We move multi-objective reasoning into the leader-selection pathway and retain the MGO update mechanisms, so that the effect of leader-mediated Pareto feedback can be measured under controlled conditions: one search skeleton, one evaluation budget, a null control and a leader-isolation ablation.

The main contribution is a controlled study of structured Pareto-archive leader selection in the MGO, situated within a taxonomy of where Pareto information enters the search, a design question that every multi-objective extension of a swarm optimizer faces. All four leader-mediated variants rank ahead of all five reconstructed prior MGO designs on all three primary indicators, and the decomposition and indicator leader rules outrank random archive leadership on hypervolume and IGD^{+} under Holm correction. The design template, consisting of the feedback-path taxonomy, the leader substitution, the null control and the two-channel scoring of returned sets and evaluation traces, can be applied to other leader-based swarm optimizers.

In detail, the paper makes three contributions. First (C1), the controlled study: four leader-selection strategies integrated into one base under a uniform protocol, with a null control and a leader-isolation ablation that test the contribution of leader selection under

common management. Second (C2), a refined design whose choices follow from the comparative findings. Third (C3), an out-of-sample validation of the refined design against established MOEAs on real-world-derived engineering instances. Three research questions organize the study: how the leader-mediated strategies compare with each other and with five reconstructed prior methods, four L0/L1 implementations and one L2 implementation (RQ1); whether a refined design repairs the weaknesses that this comparison exposes (RQ2); and whether the refined design holds against established MOEAs out of sample (RQ3).

The contribution is methodological, and the class of workload it concerns is common in emerging digital technologies. Smart-infrastructure scheduling, digital-twin-supported engineering design and edge-resource allocation are typically formulated as constrained multi-objective search under bounded evaluation budgets, which is the regime the protocol instantiates. The study makes no application-specific claim for these settings. Its external evidence comes from fourteen real-world-derived engineering instances (vehicle crashworthiness, structural, mechanical and marine design [24]). What it offers such applications is the design layer beneath them, namely which leader-mediation choices transfer and which fail; application-level validation is left to dedicated studies.

2. Background and the Architectural Taxonomy

2.1. Solution Strategies and MOEAs

A multi-objective problem minimizes $m \geq 2$ conflicting objectives over a feasible decision space. Its solution is the Pareto set of non-dominated points, and the image of that set is the Pareto front [1,25]. Four strategy families recur. Scalarization aggregates the objectives into a scalar and sweeps the aggregation parameters across runs; linear weighting cannot recover non-convex front regions for any choice of weights [26]. Pareto-based methods compare solutions by dominance and preserve diversity by crowding or niching, as in NSGA-II [3] and, for many objectives, the reference-point scheme of NSGA-III [27]. Decomposition recasts the problem as scalar subproblems on a weight-vector lattice and solves them cooperatively, as in MOEA/D [4]. Indicator-based methods select by contribution to a Pareto-compliant indicator, usually hypervolume, as in SMS-EMOA [5]. Each family supplies a candidate rule for choosing a leader from a Pareto archive.

2.2. Multi-Objective Swarm Extensions

MOPSO established the archive-plus-leader-selection template for swarm methods [6]: the search loop is preserved, an archive replaces the single best, and a heuristic (random, sigma, grid or hypervolume) picks the guides that enter the update equations at every step. The MOGWO, MOMVO and the MSSA [7–9] adopt this pattern and sample leaders from a Pareto archive during candidate generation. By the criterion formalized below, these algorithms are leader-mediated. The systematic comparison of leader rules has two direct precedents in particle swarm optimization (PSO). Padhye et al. [28] compared ten global-guide selection schemes in one MOPSO framework, among them random, dominance-based, crowding-distance and hypervolume-contribution guides, and found the hypervolume-contribution guide among the most effective by hypervolume and the most expensive to compute. Nebro et al. [29] compared leader-selection strategies within SMPSO, including random-leader controls under both crowding-distance and hypervolume-contribution archive management, and found best overall the configuration that combines a hypervolume-contribution tournament with contribution-based archive pruning. We apply that comparison design to the MGO. Our indicator-based result is consistent with their finding, with two differences: our indicator leader is sampled with probability proportional to its contribution, whereas theirs is chosen by a contribution tournament, and the lead of our indicator-based variant over the other full MGO variants does

not survive Holm correction (Section 4.1). The MGO also has a population-survival step, which neither PSO study contains, and we examine it separately. For the MGO, the controlled leader comparison has not been made, and one 2026 work applies a leader-mediated design (Section 2.5).

2.3. The Mountain Gazelle Optimizer

The MGO models four gazelle behaviours as position-update mechanisms applied to every agent in each iteration, pools the new candidates with the current population and keeps the best members, up to the population size, for the next iteration [14]. Three mechanisms combine population members with a leader, *gazmale*; migration to search for food performs a leader-free random move. The role of *gazmale* is the hinge of this work: it is unambiguous under a scalar fitness and undefined under a partial order. A structural audit of the MGO's own equations also exposes a component that the metaphor conceals: the young-male-herd term reduces almost surely to the mean of a random third of the population (Section 3.1). This observation is in line with the critique that motivates our structure-first stance [10,11].

2.4. Prior Multi-Objective MGO Works

The census combines two legs. The primary leg is a complete forward-citation screen of Abdollahzadeh et al. [14]: the full Scopus cited-by export of 19 July 2026, 480 records. The second leg consists of open-web keyword queries that cross the MGO name variants (MGO, MMGO, MGOA) with multi-objective terms. A keyword filter reduced the export to 27 explicit multi-objective candidates, and a second pass over the 159 remaining MGO application records without explicit multi-objective keywords flagged 61 more for hand screening. Inclusion required that the study apply the MGO itself to a problem with two or more objectives and report a multi-objective output; the similarly named Gazelle Optimization Algorithm is a different method and was excluded on sight. The census identifies twenty multi-objective MGO application works; three further records qualified on their abstracts only, are held as provisional in Supplementary Materials S1 (rows 11, 12 and 14) and are not counted here. Supplementary Materials S1 records the evidence and verdict for every screened record, including the records removed by the documented keyword filter. Table 1 classifies the nine most fully specified works. Of the twenty, eleven scalarize (L0), eight are survival- or archive-based (L1), one of those carrying reinforcement-learning meta-control as a compound label, and one, Pham et al. [23], selects its movement guide by trade-off information inside the iteration loop (Section 2.5); none draws per-agent leaders from a Pareto archive. One limit of evidence remains: three of the twenty describe multi-objective machinery but state no guide-selection rule anywhere in their equations, pseudocode or figures [30–32], and one of the three reproduces no update equations at all. For these the L1 label is provisional: the Pareto ranking and compromise-selection machinery supports L1, and its presence does not establish L2 without an explicit connection to the candidate-generation guides. The label records the machinery described and not a reconstructed implementation. For two of the three, the same silence forbids the opposite assertion, that their guides are chosen by scalar fitness. The third, Dai [30], carries an unresolved inconsistency of its own: the application equations specify a scalar objective, while the algorithm section describes multiple objectives and non-dominated solutions without connecting the two, so an equations-only assessment of the application formulation supports L0. That leader selection goes unspecified in three of the twenty works is itself evidence about the design axis this study isolates. Three of the nine works in Table 1 implement scalarization (L0). Said et al. [16] co-optimize weighted-sum weights as decision variables for superconducting energy-storage design. Pandey, Baltacı, et al. [17]

adapt the weights within the run from population-spread statistics for the construction time–cost trade-off. Pandey, Jadoun, et al. [18] optimize each objective of a virtual-power-plant schedule alone, compare the optima in a pay-off table and re-optimize one objective with the other as a constraint; the resulting trade-off is shown as a Pareto graph, but no dominance, filtering or diversity step is described.

Five of the nine implement archive- or survival-based designs (L1). Dede et al. [33] apply the adaptive-weight scheme of Pandey, Baltacı, et al. to a nineteen-activity time–cost instance and rank the reported front by crowding distance at termination. Ramachandra and Natarajan [19] pair a Pareto archive with fuzzy compromise selection for home energy management. Zheng and Wu [20] add Powell local search and a weight sweep around an unchanged MGO core. Refaat et al. [21] run the MGO among five algorithms against an external non-dominated archive for power-system resilience planning. Eirgash et al. [22] add non-dominated sorting for survival together with a Q-learning agent that toggles opposition-based learning; this Pareto-conditioned control of operator invocation is an axis orthogonal to the levels and is recorded as a compound label. Beyond the nine of Table 1, one further L1 design is worth recording for what it omits. Al-Dhaifallah et al. [32] merge the MGO with a multi-objective salp-swarm scheme in which every individual generates five candidate positions per iteration, four from the MGO's mechanisms and one from the salp chain, and dominance resolves which the individual becomes; the salp half is implemented without the archive-based food-source selection of its parent algorithm, so a partner that is leader-mediated in its own right contributes none of that machinery to the merged method. In none of these designs is the Pareto archive read to select *gazmale* during candidate generation.

Table 1. Architectural classification of the nine most fully specified multi-objective MGO works, from a forward-citation census of twenty studies (Scopus export of 19 July 2026; three further abstract-level records are held as provisional in Supplementary Materials S1). L0, scalarization; L1, survival or archive; L2, leader-mediated (Section 2.5).

Reference	Application	Multi-Objective Mechanism	Level
Said et al. (2023) [16]	SMES voltage-sag mitigation	Weighted sum, weights co-optimized	L0
Pandey, Baltacı, et al. (2026) [17]	Construction time–cost	Adaptive-weight scalarization	L0
Pandey, Jadoun, et al. (2025) [18]	Multi-area VPP scheduling	Pay-off table, constrained re-optimization, operator choice	L0
Dede et al. (2026) [33]	Construction time–cost	Adaptive weights, terminal crowding-distance ranking	L1
Ramachandra & Natarajan (2024) [19]	IoT home energy management	Pareto archive with fuzzy compromise	L1
Zheng & Wu (2025) [20]	Energy-community DSM	Weight sweep, Powell search, fuzzy compromise	L1
Refaat et al. (2025) [21]	Power-system resilience planning	External non-dominated archive	L1
Eirgash et al. (2026) [22]	Time–cost–safety-risk	NDS survival, Q-learning control of OBL	L1 + meta-control
Pham et al. (2026) [23]	Multi-project resource scheduling	In-loop fuzzy-selected non-dominated guide; NDS and crowding survival; no separate archive described	L2

2.5. The Architectural Taxonomy and the Gap

We classify a multi-objective extension by the path along which Pareto information feeds back into the search (Figure 1). At **L0 (scalarization)**, aggregation precedes the search and the operator never sees an objective vector. At **L1 (survival/archive)**, multi-objective information may govern survival, archiving or terminal compromise, including in-loop non-dominated sorting for survival, but the movement guides that enter the update equations are chosen by a rule that does not use the trade-off information, such as best scalar fitness or a fuzzy aggregate of the objectives applied without a Pareto-derived guide pool. Operationally, L1 requires a documented dominance or front-diversity step (for survival, archive maintenance, front construction or terminal selection); a scalar search with no such step is L0, and a set of outcomes only labelled a Pareto front does not establish L1. Levels describe the published method; front-recovery steps added to our reconstructions for scoring (Section 3.3) do not change them. At **L2 (leader-mediated)**, trade-off information (dominance, decomposition, diversity or an indicator) is used during the search to form the guide pool, to select guides, or to assign guides to the candidate-generation equations. An external archive, a diversity measure and per-agent guide assignment are each optional at L2; multi-objective survival or output archiving alone does not establish L2. The decisive feature is the in-loop pathway by which trade-off information forms the guide pool or selects among it, and not whether the final choice reduces to one scalar score: a fuzzy value used to pick one member of the current non-dominated set is L2, and a fuzzy aggregate used as the fitness that names the best agent keeps the guide choice outside L2, without itself establishing L1. Pareto-conditioned control of operator invocation, as in reinforcement-learning meta-control, is an orthogonal axis and is denoted by a compound label. The L0–L2 taxonomy is introduced in this study. It records where trade-off information enters and influences the search; the level numbers follow the position of that entry point along the feedback path, and no level is presumed to perform better than another. The levels are architectural categories, and the criterion is independent of performance. In what follows, *leader-mediated* and L2 are used interchangeably.

Under this criterion the prior MGO works occupy L0 or L1 with one exception. Pham et al. [23] select, inside the iteration loop, the non-dominated solution with the highest fuzzy-decision-making value as the single guide for the next generation, with non-dominated sorting and crowding-distance survival and no separate archive described (their Table 1 and Section 2.3). This is L2, and it is the one leader-mediated design in the census. MOPSO, the MOGWO, MOMVO and the MSSA are L2 by the same criterion. The L2 design studied here replaces the single best-fitness *gazmale* with per-agent leaders drawn from a Pareto archive by dominance, decomposition or indicator criteria. This changes the operational semantics of the three leader-consulting mechanisms and leaves their algebraic form intact. Archive mediation and per-agent assignment are design features of this study's variants and are not conditions of the level. The reconstruction of the single-guide design of Pham et al. enters the comparisons of Section 4 as B5.

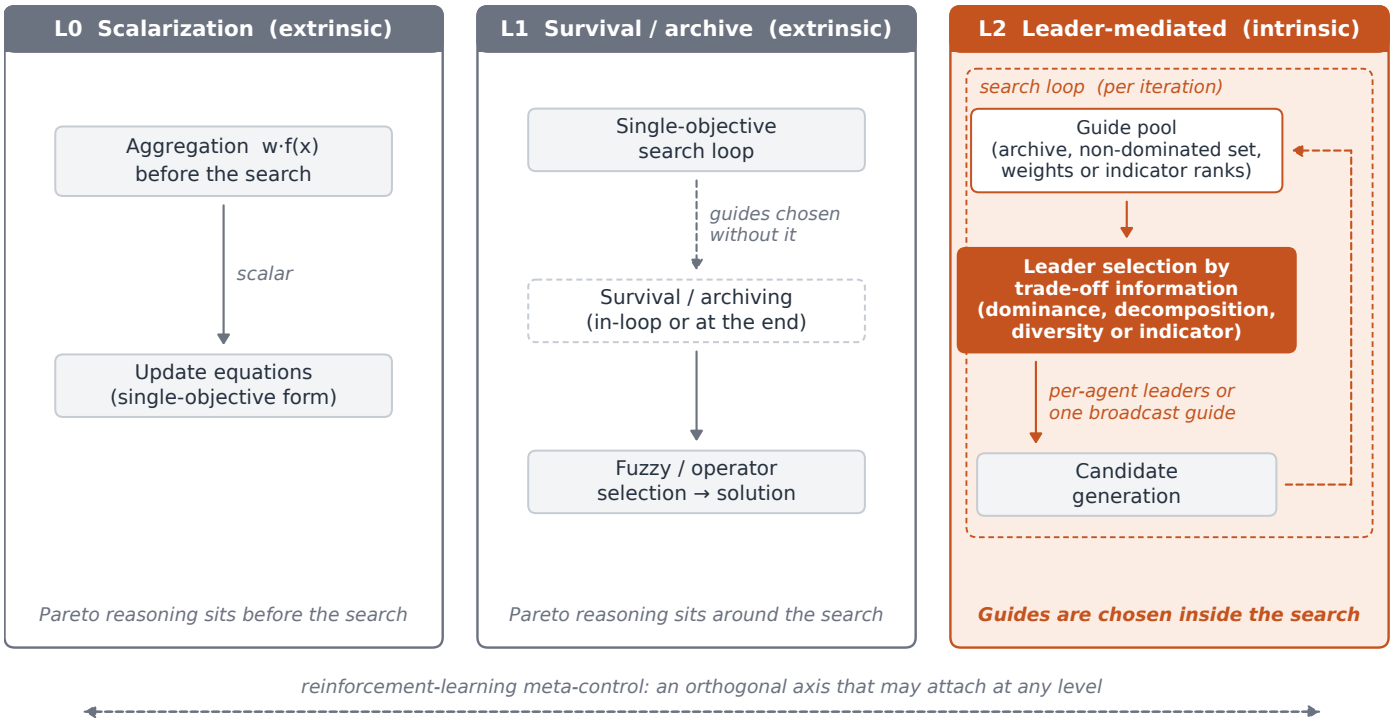


Figure 1. The proposed three-level feedback-path taxonomy: L0 scalarization, L1 survival/archive and L2 leader-mediated, with reinforcement-learning meta-control as an orthogonal axis. Solid arrows carry the flow of the search; the dashed arrow in L1 marks that the guides are chosen without the survival or archiving information; the dashed return edge in L2 marks the feedback of trade-off information into guide selection at every iteration. Of the twenty prior multi-objective MGO works, nineteen occupy L0 or L1 and one is a single-guide L2 design (Pham et al. [23]); this study addresses archive-mediated, per-agent L2 leadership.

3. Method

3.1. The L2 Substitution and Shared Skeleton

One structural modification makes the MGO leader-mediated (L2): the scalar-best leader is replaced by an archive-drawn one. Selection is performed once per gazelle, so different agents follow different non-dominated leaders within one iteration. The selected leader is reused at all four leader-injection sites: once in the territorial mechanism (Equation (5)), once in the maternity mechanism (Equation (6)), and twice in the bachelor mechanism (Equation (7), directly and inside D_i). The migration mechanism (Equation (8)) reads no leader and is unchanged.

The corrected update equations of the validated synchronous port are retained verbatim; only the binding of *leader* changes. Let t be the current iteration, T the iteration limit and D the decision-space dimension. The symbol r denotes fresh uniform scalars or vectors on $[0,1]$, $n(D)$ a standard-normal vector, and $\rho_1, \ldots, \rho_6$ independent integer multipliers drawn from $\{1,2\}$. With the iteration-decay scalar

$$a = -1 + t\left(-\tfrac{1}{T}\right), \tag{1}$$

each coefficient use draws one of four alternatives uniformly,

$$Cof \in \left\{ r(D),\ \ (a+1)+r,\ \ a\,n(D),\ \ u \odot v^2 \odot \cos\big((2r)\,u\big) \right\}, \tag{2}$$

with u, v independent standard-normal vectors and $\odot$ elementwise multiplication. The gazelle-specific stochastic factor is

$$F = n(D)\ \exp\!\left(2 - 2\tfrac{t}{T}\right), \tag{3}$$

and the young-male-herd term is

$$BH = X_{\mathrm{ra}}\,\lfloor r_1 \rfloor + M^{\mathrm{pr}}\,\lceil r_2 \rceil, \qquad M^{\mathrm{pr}} = \text{mean of } \lceil N/3 \rceil \text{ uniformly selected members}, \tag{4}$$

which reduces to M^{pr} almost surely, because $\lfloor r_1 \rfloor = 0$ and $\lceil r_2 \rceil = 1$ for continuous draws; this is the vestigial structure noted in Section 2.3. The four candidates generated for gazelle X_i are

$$\mathrm{TSM} = \mathit{leader} - \left|(\rho_1 BH - \rho_2 X_i) \odot F\right| \odot \mathit{Cof}_1, \tag{5}$$

$$\mathrm{MaH} = (BH + \mathit{Cof}_2) + (\rho_3\,\mathit{leader} - \rho_4\,X_{\mathrm{rand}}) \odot \mathit{Cof}_3, \tag{6}$$

$$\mathrm{BMH} = (X_i - D_i) + (\rho_5\,\mathit{leader} - \rho_6\,BH) \odot \mathit{Cof}_4, \qquad D_i = \left(|X_i| + |\mathit{leader}|\right)(2r - 1), \tag{7}$$

$$\mathrm{MSF} = (ub - lb) \cdot r + lb, \qquad r \sim U(0,1) \text{ a single scalar}, \tag{8}$$

each clamped to $[lb, ub]$ before evaluation. The defining L2 substitution is

$$\mathit{gazmale} \ \rightarrow\ \lambda\big(A(t),\, X_i,\, t,\, \Theta_\lambda\big), \tag{9}$$

where $A(t)$ is the bounded external Pareto archive and λ returns one archive member for gazelle i. The right-hand sides of (5)–(7), including D_i, are otherwise unchanged, and (8) reads no leader.

All L2 configurations of this study share one batch $\mu + \lambda$ skeleton (Figure 2, Algorithm 1) with population $N = 91$ and archive capacity $A_{\max} = 100$. Each iteration generates $4N$ candidates, reduces the $5N$ pool to N survivors by a population-management operator π_{pop}, and admits candidates to the archive under feasibility-first dominance, pruning to capacity by an archive-management operator π_{arch}. The population size $N = 91$ equals the number of Das–Dennis reference directions that NSGA-III and MOEA/D use in this study ($m = 2$ with $H = 90$, $m = 3$ with $H = 12$, and at $m = 4$ two layers of 56 and 35 directions; Table A2), so the leader-mediated designs, B5 and the population-based comparators share one population size, and V2 assigns one weight vector of that lattice to each gazelle (V2 ran only at $m \leq 3$). The capacity $A_{\max} = 100$ equals the common terminal capacity of evaluation-trace scoring (Section 3.4 and Table A3) and matches the archive of at most 100 leaders in the SMPSO leader study of Nebro et al. [29]. Both values were fixed before the runs, and Section 4.5 reports the sensitivity of the refined design to them. In the comparator study the archive-based designs return up to 100 points, and the established comparators and B5 up to 91; evaluation-trace scoring applies one capacity to every algorithm and so controls for this difference in returned-set size.

The archive supplying λ is fixed within an iteration (synchronous update), which gives an unambiguous intervention point; admission uses feasibility-first domination on constrained problems. On constrained problems every comparison in the skeleton follows the feasibility-first rule of Deb [34]: a feasible solution precedes an infeasible one, infeasible solutions are ordered by their total constraint violation (the sum of the positive parts of the constraint functions), and feasible solutions by dominance or by the operator's own criterion (Table A1). With initialization counted, the iteration limit under a per-run budget FE is

$$T = \left\lfloor \frac{FE - N}{4N} \right\rfloor, \tag{10}$$

so $FE = 30{,}000$ with $N = 91$ gives $T = 82$ and 29,939 counted evaluations.

Algorithm 1 Shared L2 MO-MGO skeleton (leader operator λ; management operators $\pi = (\pi_{\text{pop}}, \pi_{\text{arch}})$).

Require: objective and constraint functions on the box $[lb, ub] \subset \mathbb{R}^D$; population size N; archive capacity $A_{\max}$; evaluation budget FE, which fixes the iteration limit T by Equation (10); leader operator λ with parameters Θ_λ; management operators π_{pop} and π_{arch}
Ensure: the final archive $A(T)$, with $|A(T)| \le A_{\max}$
initialize $P(0)$, $|P(0)| = N$, uniformly within $[lb, ub]$; evaluate $P(0)$
$A(0) \leftarrow$ non-dominated members of $P(0)$
for $t = 0, \ldots, T-1$ **do**
  compute the iteration scalar a ▷ Equation (1)
  $C(t) \leftarrow \emptyset$
  for each gazelle $X_i \in P(t)$ **do**
    $leader_i \leftarrow \lambda(A(t), X_i, t, \Theta_\lambda)$ ▷ Equation (9)
    sample fresh F, BH, coefficients and integer multipliers ▷ Equations (2)–(4)
    generate TSM_i, MaH_i, BMH_i using $leader_i$; generate MSF_i ▷ Equations (5)–(8)
    clamp all four candidates to $[lb, ub]$, evaluate, add to $C(t)$
  end for
  $P(t{+}1) \leftarrow \pi_{\text{pop}}\big(P(t) \cup C(t),\, N\big)$ ▷ $5N \to N$
  $A(t{+}1) \leftarrow$ admit and dominance-purge $A(t)$ with the pool
  if $|A(t{+}1)| > A_{\max}$ **then** $A(t{+}1) \leftarrow \pi_{\text{arch}}(A(t{+}1), A_{\max})$
  end if
end for
return $A(T)$

3.2. Leader-Mediated Variants and Controls

Table 2 summarizes the four variants and the controls by leader rule, population survival and archive pruning, and names the established method whose principle each leader rule transposes. Four variants instantiate the strategy families as leader rules under the shared skeleton. **V1 (Pareto dominance)** selects by a binary tournament on non-dominated rank and crowding distance, and manages population and archive by NSGA-II sorting and crowding. **V2 (decomposition)** assigns one Das–Dennis weight vector per gazelle and selects the neighbourhood Tchebycheff-optimal archive member; it has no native external-archive policy, reuses V1's crowding pruning, and is therefore a decomposition-flavoured package and not canonical MOEA/D. **V3 (indicator-based)** samples a leader with probability proportional to its internal, adaptively normalized hypervolume contribution (ΔHV) and manages population and archive by iterative smallest-ΔHV removal. The normalization is fixed from the set under consideration: each objective is mapped from its minimum and maximum over that set to $[0, 1]$, the reference point is 1.1 per objective, and every leave-one-out term uses the same transform (Table A2); it is separate from the scoring normalization of Table A3. **V4 (mixture)** delegates each leader call to V1, V2 or V3 with equal fixed probability. V4 is a heterogeneous leader-selection strategy: it tests whether the three leader policies complement one another when they are mixed, before any adaptive switching is introduced, and the equal fixed probabilities are the uninformed choice, which adds no tuned parameter. Over the 660 V4 runs of the standard suite (7462 leader calls per run), the three rules received 33.2%, 33.4% and 33.4% of the calls (V1, V2 and V3), and no run departs from one-third by more than 1.7 percentage points. The null control **V0** draws the leader uniformly at random from the archive under V1's management, the random-leader control that Nebro et al. [29] used for SMPSO. It is the minimal leader-mediated configuration, with full in-loop Pareto machinery and no leader structure, and it is analysed separately from the primary comparison. The leader-isolation (λ) ablation compares V0

and the structured leaders under common management and isolates the contribution of the leader rule.

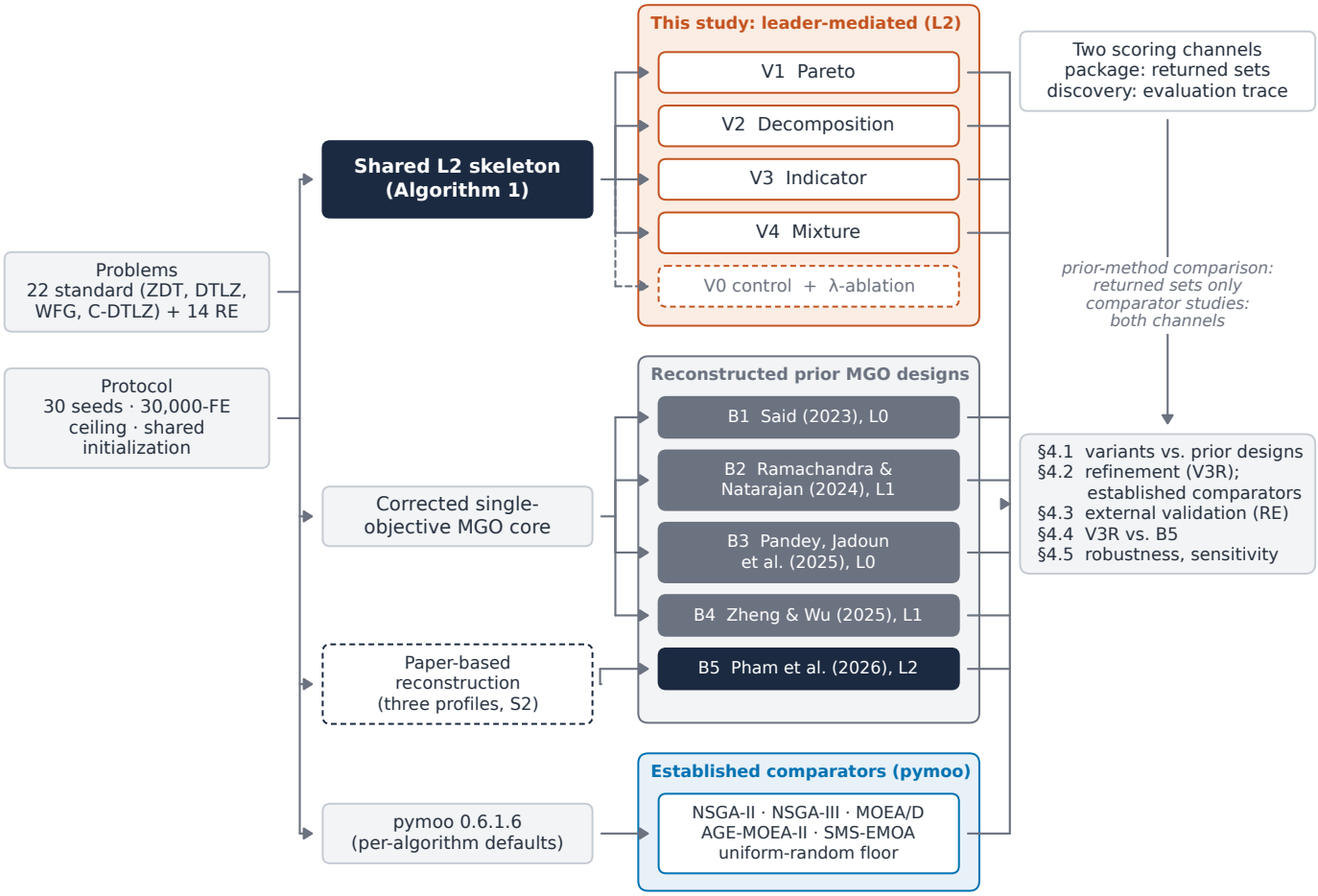


Figure 2. Study pipeline. One problem set and one protocol feed three arms: the shared L2 skeleton instantiates the four leader-mediated variants, the V0 control and the λ-ablation; the corrected single-objective MGO core hosts the reconstructed prior designs B1–B4, and a separate paper-based reconstruction hosts B5; pymoo hosts the established comparators and the random floor. The prior-method comparison (Section 4.1) uses returned-set scoring; the comparisons with the established comparators (Sections 4.2–4.4) use both returned-set and evaluation-trace scoring. The refined design V3R (Section 4.2) is composed from the findings of Section 4.1 and is therefore evaluated in sample on the standard suite and out of sample on the RE suite. Prior designs: B1, Said et al. [16]; B2, Ramachandra and Natarajan [19]; B3, Pandey, Jadoun, et al. [18]; B4, Zheng and Wu [20]; B5, Pham et al. [23].

Figure 3 summarizes the six rules on a common archive sketch; each panel also records the management pair $(\pi_{\text{pop}}, \pi_{\text{arch}})$ that completes the package.

Table 2 lists the three components of every L2 configuration: the leader rule λ, the population survival π_{pop} and the archive pruning π_{arch}. Leader selection and population survival are separate mechanisms in this design, and the configurations are chosen so that their effects can be read separately. Along the chain V0, V3-λ, V3R, V3, adjacent configurations differ in exactly one component: V0 and V3-λ differ in the leader rule, V3-λ and V3R in the archive pruning, and V3R and V3 in the population survival. The V2-λ arm plays the same role for the decomposition leader, and V1 and V4 already run under the common crowding management.

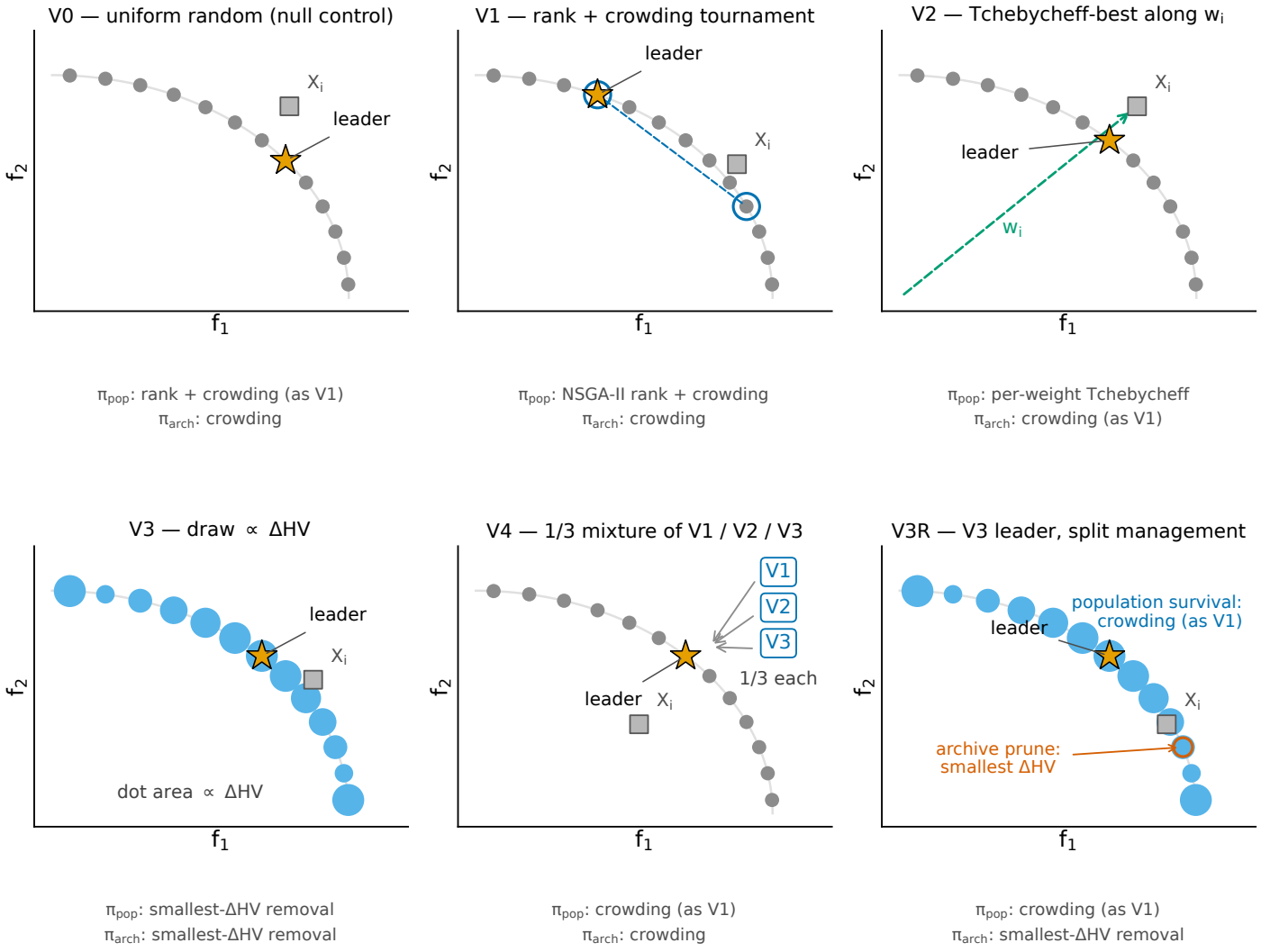


Figure 3. Leader selection from the archive under each algorithm, on a common archive sketch in a bi-objective minimization space (gray dots: archive members of $A(t)$ on a concave front; gray square: the gazelle X_i; orange star: the selected *leader*$_i$). V0 draws uniformly. V1 runs a binary tournament on non-dominated rank (blue circles: the two candidates), with ties broken by larger crowding distance. V2 takes the Tchebycheff argmin along the gazelle's weight vector w_i (green dashed arrow) among the members of its weight neighbourhood $B(i)$. V3 samples with probability proportional to the clamped internal hypervolume contribution plus an ε floor (blue dot area $\propto$ ΔHV). V4 delegates each call to λ_{V1}, λ_{V2} or λ_{V3} with equal, untuned probability. V3R retains V3's leader rule and archive pruning (vermilion circle: the archive member pruned) and restores crowding-based population survival. The footer of each panel records the management pair $(\pi_{\text{pop}}, \pi_{\text{arch}})$ that completes the package.

Computational cost. All runs obey a common 30,000-evaluation ceiling. The default L2 designs and the established comparators consume 29,939 evaluations each (Equation (10)); the achieved spends of the sweep baselines are listed in Table A2 and those of the population-size sensitivity configurations in Section 4.5. Within the matched-budget comparisons, complexity differences are therefore selection-and-management overhead. Candidate generation costs $O(N^2D)$ per iteration for every design, because each gazelle averages a random third of the herd. For V0, V1 and V4 the selection terms are fast non-dominated sorting of the $5N$ pool, $O(mN^2)$, and crowding survival, $O(mN \log N)$; V2 adds the neighbourhood Tchebycheff scans, $O(mN|A|)$ per iteration. Crowding pruning of the archive, used by V0, V1, V2 and V4, recomputes crowding after each of its r removals from the post-admission set S, at a cost of $O(r\,m|S| \log |S|)$ per iteration. The indicator-based designs are the expensive ones. In the reference formulation, each reduction step recom-

putes the leave-one-out hypervolume contributions of the surviving set S, which is the post-admission set and may exceed $A_{\max}$ before pruning, so one reduction performs up to $O(|S|^2)$ exact hypervolume evaluations: on the $5N$ pool for V3, and on the archive's post-admission set for V3R, whose population survival reverts to crowding. The executed implementation follows a guarded accelerated path at $m = 2$ with the exact reference reduction as fallback, so the reference cost is a worst case and not the cost of every run. Exact hypervolume is evaluated through pymoo's moocore backend (version disclosure in Appendix A). It is inexpensive at $m = 2$, grows steeply with m, and the study never exceeds $m = 4$. Section 4.5 and Table A4 report the measured wall-clock times. On the standard suite, which is mostly bi-objective, the leader-mediated configurations other than V3 take 4.6 to 5.7 s per run in median against 1.1 s for NSGA-III, an overhead common to them; the cost of exact hypervolume contributions shows at three and four objectives, where V3R's median on the RE suite is 50.3 and 90.1 s against 1.6 and 1.7 s for NSGA-III. These overheads would shrink to a small fraction of runtime in expensive-evaluation regimes, where the shared evaluation budget dominates.

Table 2. Components of the L2 configurations and the principle that each leader rule transposes. Crowding survival (π_{pop}) denotes NSGA-II non-dominated sorting with crowding-distance truncation of the last accepted front; crowding pruning (π_{arch}) denotes iterative removal of the archive member with the smallest crowding distance, recomputed after each removal; smallest-ΔHV denotes iterative removal of the member with the smallest hypervolume contribution, recomputed after each removal. In the chain V0, V3-λ, V3R, V3 each step changes one component (Section 4.2).

Configuration	Leader Rule λ	Principle and Source	Population π_{pop}	Archive π_{arch}	Role
V0	uniform random	none; random-leader control as in SMPSO [29]	crowding	crowding	null control
V1	rank and crowding tournament	Pareto rank and crowding, NSGA-II [3]	crowding	crowding	variant; its own λ-ablation arm
V2	neighbourhood Tchebycheff	decomposition, MOEA/D [4]	neighbourhood Tchebycheff	crowding	variant
V2-λ	neighbourhood Tchebycheff	decomposition, MOEA/D	crowding	crowding	λ-ablation arm
V3-λ	ΔHV-proportional draw	hypervolume contribution, SMS-EMOA [5]	crowding	crowding	λ-ablation arm
V3R	ΔHV-proportional draw	hypervolume contribution, SMS-EMOA	crowding	smallest-ΔHV	refined design (Section 4.2)
V3	ΔHV-proportional draw	hypervolume contribution, SMS-EMOA	smallest-ΔHV	smallest-ΔHV	variant
V4	V1, V2 or V3 rule, drawn uniformly per call	mixture of the three rules above	crowding	crowding	variant

3.3. Reconstructed Prior Designs

Five prior multi-objective MGO designs, B1–B5, are reconstructed from the algorithmic descriptions in their original publications and compared under the common protocol. They are paper-based reconstructions and are not the original software: several source studies leave implementation details unspecified or include application-specific components that do not transfer to the present benchmarks. Supplementary Materials S4 lists the implementation assumptions of B1–B4, and Supplementary Materials S2 those of B5. Four of the five, at L0 or L1, are separate wrappers around the corrected single-objective MGO core and never use an archive-drawn leader. **B1 (Said et al. [16], L0)** performs one augmented-vector run that co-optimizes the weights and recovers a front by post-hoc binning. **B2 (Ramachandra and Natarajan [19], L1)** runs the single-objective MGO driven by a

fuzzy decision-making scalar reconstructed from the paper's terminal compromise rule, and maintains a side archive that never guides the search; because the source defines no unique in-loop leader, B2 is held at L1 by explicit reconstruction, a caveat carried through the analysis. **B3 (Pandey, Jadoun, et al. [18], L0)** reconstructs an ε-constraint sweep of repeated single-objective runs; the original's terminal operator-driven compromise selection, which picks one solution after the search, is omitted because the comparison scores the returned set. B3 carries the level of its source method, L0; the non-dominated filter that our reconstruction applies to the sub-run optima before scoring is not part of that method. **B4 (Zheng and Wu [20], L1)** performs a fixed-weight sweep; the original's Powell local refinement and terminal fuzzy compromise stages are likewise omitted, so B4 represents the sweep architecture of that work and not its full pipeline. Two 2026 works are not reconstructed: the Q-learning design contributes on an axis orthogonal to leader selection, and the adaptive-weight design is described too thinly for a faithful reconstruction.

The fifth design, **B5**, is a paper-based reconstruction of the one prior leader-mediated design, Pham et al. [23]. Its update scheme is that of the published method, including its opposition-based learning and chaotic updates, so it is hosted separately from the MGO core. Matching the evaluation budget does not match the generation count: the opposition evaluations are charged to the same budget, so B5 completes fewer generations than the 82 iterations of the shared skeleton, and its schedule horizon and achieved generation count are kept distinct in the reconstruction. It spends the same 29,939 evaluations, starts from the shared initial population (to which it applies its opposition-based refinement), and returns its final non-dominated set. The published description leaves several details unresolved: the random factor in the food-search movement, the schedule horizon of its step-size and jumping-rate controls, the normalization range of its fuzzy guide selection, and its feasibility handling. A reference profile (B5) and two sensitivity profiles (B5-L, literal food-search reading; B5-C, schedule calibrated to the achieved generation count) instantiate the reconstruction under a documented assumptions ledger (Supplementary Materials S2). The reference profile enters the comparisons of Sections 4.1–4.3. Each sensitivity profile also replaces it in a repeat of the prior-method family of Section 4.1, and both join the comparator families in Section 4.4. The comparison protocol and the reconstruction were specified in writing before the production runs, and pilots validated the implementation against that specification.

Comparisons involving B1–B5 therefore evaluate the reconstructed algorithmic designs under a common evaluation budget; they do not reproduce or re-assess the numerical results reported in the original studies. A performance difference is a property of the reconstructed design under this protocol, and this qualification weighs most for the methods whose published descriptions required additional implementation assumptions.

3.4. Experimental Protocol

The standard comparison uses twenty-two problems that span convex, concave, disconnected, biased, multimodal and constrained fronts, drawn from four scalable suites, ZDT [35], DTLZ [36], WFG [37] and C-DTLZ [38]: ZDT1–4 and ZDT6; DTLZ1–4 and DTLZ7 (three objectives); WFG1–9 (two objectives); and C1-DTLZ1, C2-DTLZ2 and C3-DTLZ4 (three objectives). External validation uses fourteen unconstrained real-world engineering (RE) problems (RE21–25, RE31–37, RE41–42) vendored from the reference implementation [24]. A curated subset of the constrained real-world multi-objective suite of Kumar et al. [39] is reserved for a later study under the same protocol. We call one algorithm on one problem a *cell*. Every cell ran thirty independent seeds, each run under a common ceiling of 30,000 function evaluations, which is the fairness invariant. Thirty runs per algorithm and problem is a common convention, used also in the SMPSO leader study

of Nebro et al. [29]; the seeds estimate the median of each cell, and the problems are the units of the cross-problem tests described below. Single-run algorithms consumed 29,939 evaluations; the sweep baselines divided the budget across many short scalar sub-runs.

Confirmatory evaluation uses hypervolume (HV), IGD$^+$ [40] and the additive ε-indicator (I_ε^+) as primary indicators, with generational distance and spacing as secondary diagnostics. All indicators are computed in an objective space normalized by the ideal and nadir points of each problem's reference set, and the hypervolume reference point is 1.1 per normalized objective. On the standard suite the reference set is a fixed sample of the analytic Pareto front, against which IGD$^+$ and the ε-indicator are also measured. On the RE suite, which has no analytic fronts, the reference set is the non-dominated union of the sets scored on each channel (defined below): the returned sets on the package channel and the trace-derived sets on the discovery channel, of all participating algorithms, the three B5 profiles included (Supplementary Materials S2, Section S2.5). Table A3 gives the generators and point counts. Inference is non-parametric and is applied across problems: Friedman ranks [41], Wilcoxon signed-rank post-hoc pairs [42] under Holm correction [43], and matched-pairs rank-biserial effect sizes, following Derrac et al. [44] at $\alpha = 0.05$. The unit of analysis is the problem. Within an algorithm–problem cell the thirty seeds estimate the distribution of the indicator on that problem, and the cell enters every cross-problem test as its median over the thirty seeds. The problems are the matched units of the Friedman and Wilcoxon tests, so no seed enters a cross-problem test as a separate observation, and the bootstrap intervals resample problems as paired blocks. Seed-level tests appear only for the single-problem contrasts of Section 4.2, where the seeds are the units; they, and the two across-suite paired tests of the refined design in the same section, are single contrasts reported without multiplicity correction and lie outside the Holm families. The null control and the ablation are analysed separately from the primary comparison. Each Holm family comprises all $k(k-1)/2$ pairwise contrasts within one roster, one indicator and one channel, and the correction is applied within that family only, never pooled across indicators or channels. Failure to reject a pairwise null is reported as insufficient evidence of a difference and never as equivalence. The rosters are: the nine-algorithm prior-method family (V1–V4, B1–B5; thirty-six pairs); the comparator roster of V3, V3R, B5, the five established comparators and the random floor (thirty-six pairs on the external suite; on the standard suite MOEA/D participates only on the nineteen unconstrained problems, and the package-channel roster excludes V3); and the external-suite comparator roster of the same nine algorithms. The established comparators span the strategy families of Section 2.1 at one canonical representative each, NSGA-II (Pareto), MOEA/D (decomposition) and NSGA-III (the reference-direction line of the same family), joined by SMS-EMOA [5], which seats the indicator-based family that the best-ranked variant transposes, and by AGE-MOEA-II [45], a current front-geometry-adaptive descendant; a uniform-random search floor completes the roster.

The leader-mediated swarm precedents MOPSO, the MOGWO, MOMVO and the MSSA enter the study as architectural precedents and are not performance comparators. They instantiate leader-mediated multi-objective search on different base search dynamics, so a direct comparison with the proposed variants would confound the effect of the leader-selection architecture with differences between the base optimizers. The aim of this study is to isolate how alternative Pareto-informed leader mechanisms affect one MGO search process. The controlled comparisons therefore use variants that share one search skeleton, together with established MOEAs that represent the distinct strategy families of Section 2.1. Direct cross-swarm benchmarking against MOPSO, the MOGWO, MOMVO and the MSSA is a complementary study, and its results would not bear on the architectural claims examined here.

Section 4.4 reads the contrasts between the refined design and B5 from these comparator families and adds the two sensitivity profiles of the B5 reconstruction.

The proposed archive-based designs (V0–V4, V3R) return a bounded run archive, the pymoo comparators return the non-dominated set of their final population, and every indicator rewards returned-set content. The comparisons with the established comparators are therefore reported on two channels; the prior-method comparison of Section 4.1 uses the package channel only, because the reconstructed designs are compared as the documented reconstructed implementations, with the adaptations and omissions stated in Section 3.3. The *package channel* scores what each published pipeline returns. The *discovery channel* scores, identically for every algorithm, the feasible non-dominated set of its complete evaluation trace, captured passively at the counted-evaluator boundary, validated by a bit-identity gate, canonicalized, and capacity-truncated once at termination by a common crowding-based truncation (iterative smallest-crowding-distance removal with recomputation after each deletion, boundary points retained). We refer to package-channel scoring as *returned-set scoring* and to discovery-channel scoring as *evaluation-trace scoring*. Under evaluation-trace scoring every algorithm is scored with access to its complete evaluation trace, so no design is penalized for a parsimonious output policy. The terminal capacity is an upper bound and not a guaranteed cardinality: a trace whose feasible non-dominated set is smaller than the capacity is scored as it is. A cardinality-cap sensitivity (caps of 10–100 points under the same common truncator, with per-cap coverage reported) accompanies the discovery channel and is descriptive only. Constraint handling follows the feasibility-first Deb rule [34] throughout, in the variants' management, in the reconstructed baselines' selection scalars and in archive admission, and the locally generated reference geometry is verified against direct problem evaluation at the Pareto-set condition before use.

The two channels can rank the same runs differently because a returned set reflects the output policy as well as the search: an archive keeps non-dominated points that a final population can lose, archive-based designs can return more points, and V3 and V3R prune their archives by ΔHV, which favours hypervolume. Evaluation-trace scoring removes these reporting differences by applying one capacity and one crowding-based truncation to every algorithm's evaluated points; the effect of archive management on the search itself remains. A difference between the channels therefore indicates the output-policy effect without measuring it; on the RE suite each channel also has its own reference union.

4. Results

4.1. Comparative Study and Leader-Isolation Ablation (RQ1)

4.1.1. Architectural Stratification

Across the twenty-two problems, the Friedman omnibus rejects equal performance on every indicator ($p \approx 10^{-22}$–10^{-23}). The mean ranks (Table 3) stratify by provenance. The four proposed designs occupy the top four ranks on every indicator, and the whole proposed block outranks the whole block of prior implementations: the worst variant mean rank precedes the best prior design on HV (V1, 3.64 against B5, 4.84), on IGD$^+$ (3.68 against 4.89) and on the ε-indicator (V1, 3.48 against B2, 5.14). This comparison evaluates complete implementations across architectural levels. Because B5 is itself L2, differences between it and the proposed variants cannot be attributed to the presence of leader mediation alone; the L2-versus-L0/L1 reading applies to the B1–B4 subset. Of the twenty cross-block pairwise contrasts, fifteen reach Holm significance on HV, fourteen on IGD$^+$ and all twenty on the ε-indicator, often at near-maximal effect; within the B1–B4 subset the counts are twelve of sixteen, twelve of sixteen and sixteen of sixteen. Two qualifications apply. Part of the margin against the sweep baselines reflects their shallow constituent sub-runs under the shared ceiling: passive per-sub-run tracing of the sweeps shows every

scalar sub-run making material progress, but on the representative three-objective problem roughly three-quarters are still improving at termination, so the budget-limited reading attaches chiefly to the three-objective sweep cells. The pairs that fail to separate lie in the columns of B2, the most classification-sensitive prior work (four pairs on HV and on IGD^+), and of B5 (V3 on HV, $p_{holm} = 0.051$; V2 and V3 on IGD^+, 0.095 and 0.080). The proposed-versus-B5 conclusion does not depend on the B5 reconstruction profile. With B5 replaced by the literal food-search profile B5-L or the calibrated-schedule profile B5-C, all four proposed designs keep better mean ranks than B5 on every indicator (B5 at 5.77, 5.80 and 5.98 under B5-L and at 4.80, 4.80 and 5.07 under B5-C, on HV, IGD^+ and the ε-indicator), the cross-block counts become sixteen, sixteen and nineteen of twenty under B5-L and fifteen, fourteen and twenty under B5-C, and B2 and B5 remain unseparated on every indicator (Supplementary Materials S3, Section S3.5). Under every profile V1–V4 occupy the first four positions in the same order and B3, B4 and B1 the last three; B2 and B5 share fifth and sixth place, with B2 ahead on the ε-indicator under the reference profile and on all three indicators under B5-L.

Table 3. Mean ranks across the twenty-two standard problems, by primary indicator (lower is better), for the nine-algorithm prior-method family. Algorithms ordered by hypervolume rank. Provenance: proposed designs of this study, or reconstructed prior designs; the level of B1–B4 is the classification of Table 1, and B2's is by explicit reconstruction (Section 3.3).

Algorithm	Provenance and Rule	Level	HV	IGD^+	I_ε^+
V3	proposed, indicator leader	L2	2.14	2.18	2.34
V4	proposed, mixture leader	L2	2.68	2.73	2.70
V2	proposed, decomposition leader	L2	3.36	3.27	3.02
V1	proposed, Pareto leader	L2	3.64	3.68	3.48
B5	reconstructed prior, single broadcast guide	L2	4.84	4.89	5.30
B2	reconstructed prior, fuzzy archive	L1	4.91	4.95	5.14
B3	reconstructed prior, ε-constraint sweep	L0	7.32	7.36	7.09
B4	reconstructed prior, weight sweep	L1	7.73	7.55	7.68
B1	reconstructed prior, augmented scalarization	L0	8.39	8.39	8.25

4.1.2. Attribution: The Null Control and the λ-Ablation

The advantage follows a gradient. The null control V0, the minimal L2 configuration with in-loop Pareto retention and archive coupling, already beats B2 on sixteen of twenty-two problems on each primary indicator (rank-biserial $r = +0.61$, $+0.61$ and $+0.75$ on HV, IGD^+ and the ε-indicator). This comparison is descriptive, because V0 and B2 differ in more than those ingredients; only the λ-ablation isolates one. A reading sensitivity bounds the reconstruction risk around B2. Under the plausible domination-driven interpretation of its undefined in-loop ordering (B2-alt), which is leader-mediated under the criterion of Section 2.5 because front rank then selects the guide inside the loop, the baseline improves on the fuzzy reading on fifteen of twenty-two problems on HV (sixteen on IGD^+ and fourteen on the ε-indicator) and still loses to the random-leader control V0 on eighteen of twenty-two on HV (eighteen on IGD^+ and twenty-one on the ε-indicator). The V0–B2-alt contrast is a package comparison of two in-loop designs with different archive roles, and the block-level conclusion holds under either reconstruction choice. The λ-isolation ablation then shows that the decomposition and indicator leaders add a further, separately demonstrated increment: under common management, they outrank a uniform-random draw from the same archive on HV and IGD^+ under Holm correction ($|r| \approx 0.69$ to 0.89), the indicator leader also separates from the random draw on the ε-indicator, and the dominance leader V1 separates from it on IGD^+ only. The rank texture (Figure 4) shows a clean column split, interrupted by one dark cell in the V3 column. The null control and the

leader-isolation ablation concern the proposed shared skeleton; they are not component ablations of B5.

B5 sits between the proposed block and the L0/L1 designs. Its mean ranks are close to B2's (4.84 against 4.91 on HV, 4.89 against 4.95 on IGD^+, 5.30 against 5.14 on the ε-indicator), and the B2–B5 contrast is not separated on any indicator ($r = -0.32$, -0.29 and -0.42 in favour of B5 on HV, IGD^+ and the ε-indicator; $p_{\text{holm}} = 0.40$, 0.74 and 0.43). B5 outranks B1, B3 and B4 under Holm correction on all three indicators, and V1 and V4 outrank B5 on all three ($r = +0.79$ to $+0.98$). Table 4 gives the descriptive per-problem map. B5 wins against each proposed design on one to four of the twenty-two problems, with the same three problems recurring (DTLZ3, DTLZ4 and WFG6, on which B5 ranks second of nine), and loses on all three constrained problems to every proposed design, so the margin is not concentrated in the constrained subset. Its one failure is ZDT2, on which its median front reduces to a single extreme point (median hypervolume 0.110, the same value as B1, which is the one tie in the map). B5 returned a feasible set on every seed of every problem.

Table 4. Descriptive per-problem map of B5 against each algorithm of the nine-algorithm family: wins–losses–ties over the twenty-two standard problems, on the median of thirty seeds per cell, package channel. A tie is exact equality of the two medians; no tolerance is applied. Ties are dropped as zero differences from the paired tests of the nine-algorithm family (Supplementary Materials S3). No hypothesis test accompanies this table.

Indicator	V1	V2	V3	V4	B1	B2	B3	B4
HV	3–19–0	4–18–0	4–18–0	3–19–0	21–0–1	14–8–0	21–1–0	21–1–0
IGD^+	3–19–0	4–18–0	4–18–0	3–19–0	21–0–1	14–8–0	21–1–0	20–2–0
I_ε^+	1–21–0	3–19–0	3–19–0	1–21–0	20–1–1	15–7–0	18–4–0	20–2–0

Figure 5 shows every pairwise contrast of the family on hypervolume. All contrasts between a proposed design and B1, B3 or B4 are significant; the contrasts with B2 are not; among the proposed designs only V1–V4 separates; B5 is separated from V1, V2 and V4 and not from V3 or B2.

4.1.3. The Best-Ranked Variant and Its Localized Failure

The indicator-based V3 is the best-ranked variant on convergent descriptive grounds (Figure 6): it holds the best mean rank on all three indicators and on the two largest families (WFG and ZDT, fourteen of twenty-two problems). Its within-variant advantage does not survive Holm correction, so V3 is identified descriptively, and the choice of donor for the refinement is recorded as a design commitment. The advantage is indicator-general: V3 also leads on IGD^+ and the ε-indicator, which are unrelated to its internal selection rule, which argues against an explanation in which V3 merely optimizes its own yardstick. Against this lead stands one sharply localized design failure (Figure 7). On the multimodal ZDT4, V3's median hypervolume falls to 0.48 against $\approx$0.87 for its siblings, with the approximation set trapped on a local front. The λ-ablation shows that the collapse is a management effect: under crowding management the same ΔHV leader is free of it (V3-λ, median 0.87). The V3–V3R contrast of Section 4.2 isolates population survival, because V3R keeps V3's leader and archive pruning and restores crowding survival, and V3R is also free of the collapse (median 0.87). In the tested configuration, the responsible component is therefore V3's smallest-ΔHV *survival* rule, which on problems dense with local fronts prunes the population toward a single high-contribution cluster. On the constrained C1-DTLZ1, by contrast, V3 is the strongest algorithm in the family (median hypervolume 1.10). The mixture V4 has the second-best descriptive mean rank on all three indicators, ahead of V1 and V2 and behind V3 (Table 3). The V4–V3 difference compares complete configurations, because V4

uses crowding for population survival and archive pruning and V3 uses smallest-ΔHV removal for both (Table 2), so it does not isolate the effect of mixing the leader rules; adaptive switching among the leader rules is listed as future work in Section 6.

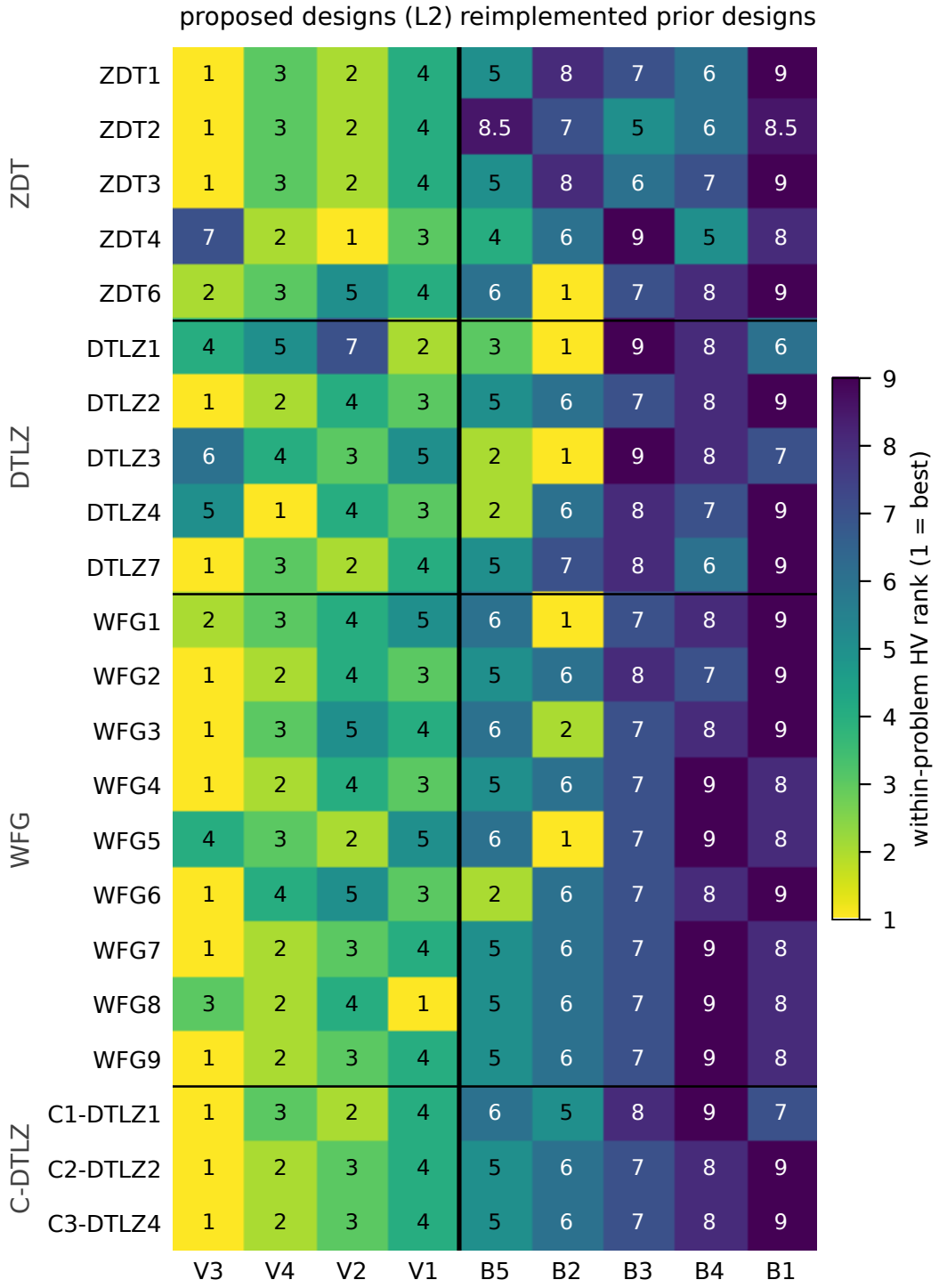


Figure 4. Rank heatmap of within-problem hypervolume rank (1 = best) across the twenty-two problems for the nine-algorithm family. Light cells are good ranks and dark cells poor ranks. The four proposed-design columns hold the best ranks; the ZDT4 cell in the V3 column is its localized design failure; B5's darkest cell is ZDT2.

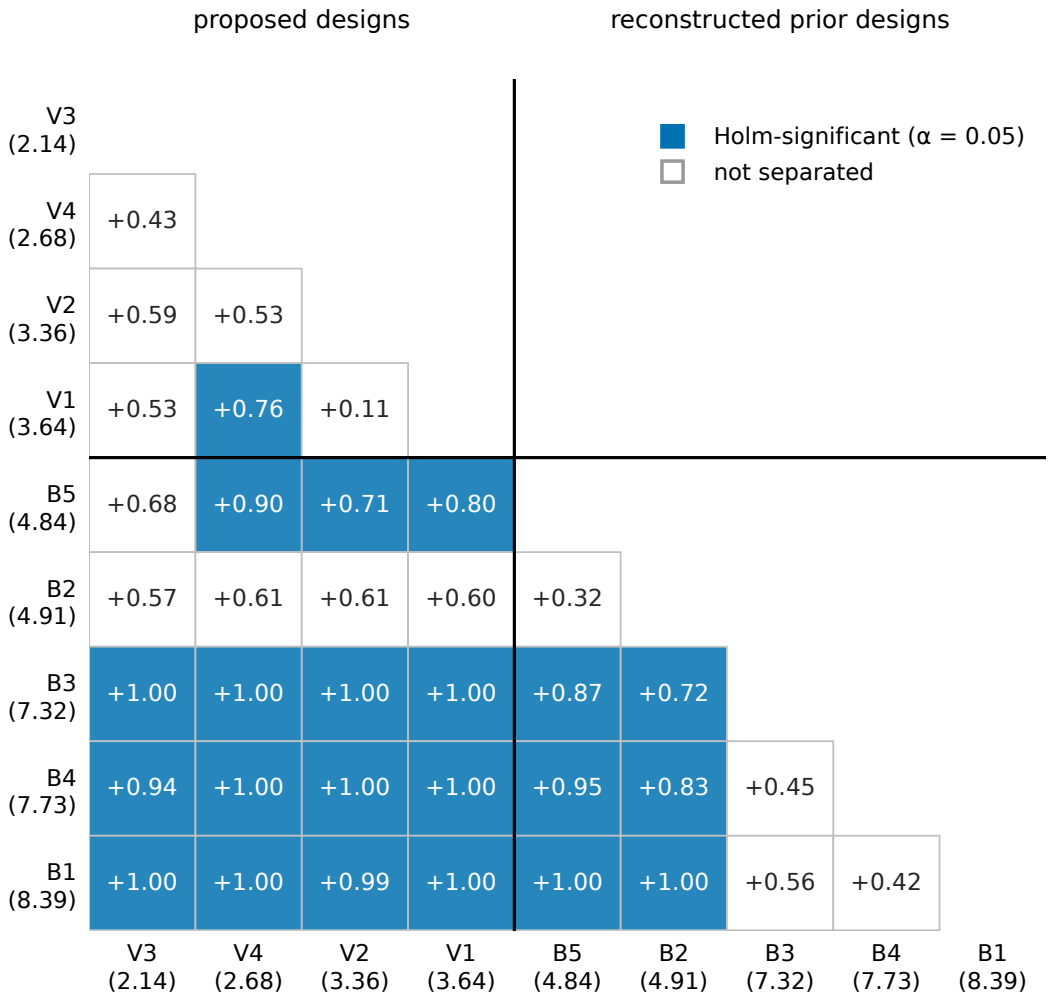


Figure 5. Pairwise contrasts of the nine-algorithm prior-method family on hypervolume across the twenty-two standard problems, algorithms ordered by mean Friedman rank (in parentheses; lower is better). Each cell gives the matched-pairs rank-biserial r of the column algorithm against the row algorithm, so a positive value favours the better-ranked algorithm; shaded cells are Holm-significant at $\alpha = 0.05$ within the family of thirty-six pairs, and white cells are not separated. The lines divide the proposed designs (V1–V4) from the reconstructed prior designs (B1–B5).

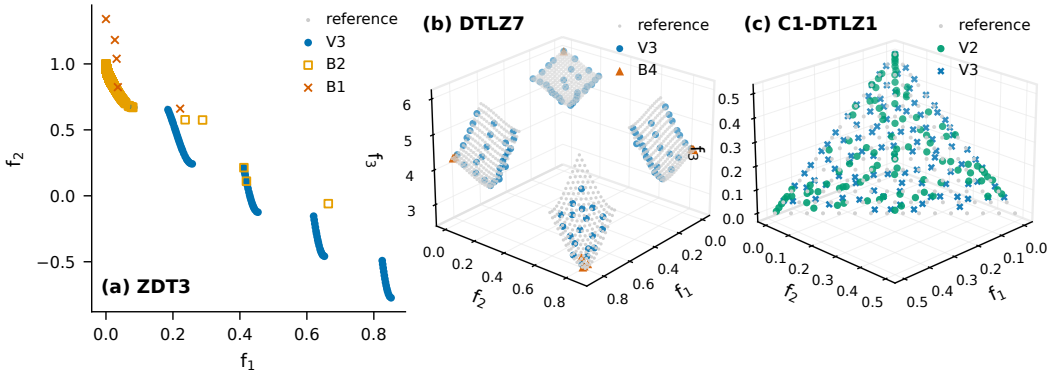


Figure 6. Representative approximation sets over the reference front, each algorithm at its median-hypervolume seed. (**a**) ZDT3, a disconnected bi-objective front, on which the leader-mediated variant V3 recovers all five Pareto segments while the L1 and L0 baselines B2 and B1 cover them only in part; (**b**) DTLZ7, a three-objective four-region front, on which V3 attains the family's best coverage against B4; (**c**) the constrained C1-DTLZ1, on which V2 and V3 both distribute across the feasible Pareto triangle. In (**a**), the grey reference points lie beneath the plotted solutions.

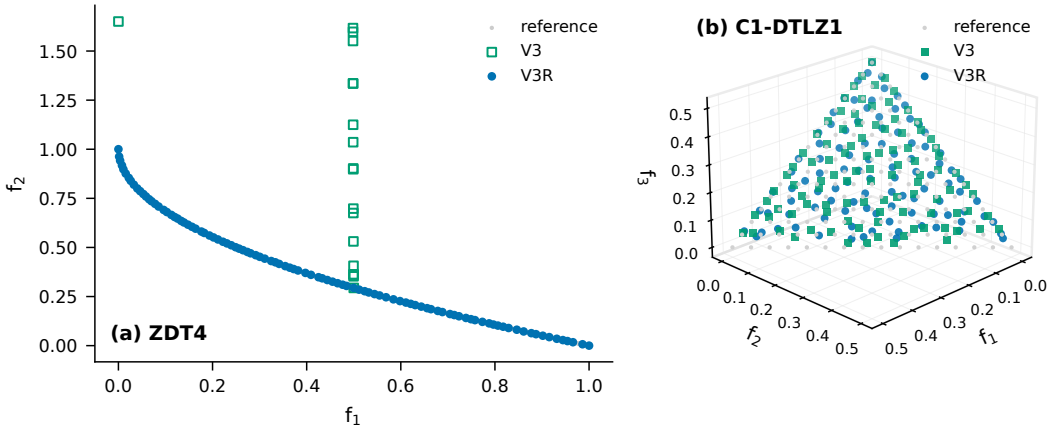


Figure 7. The targeted repair: approximation sets on ZDT4 (V3 trapped on a local front, V3R on the global front) and on C1-DTLZ1 (V3 and V3R, both fully distributed across the feasible triangle). In (**a**), the grey reference points lie beneath the V3R solutions.

4.2. Refined Design and Comparison with Established MOEAs (RQ2)

The diagnosis constrains the design space. In the tested configuration, the evidence does not implicate the ΔHV leader; it implicates the ΔHV population survival. Abandoning ΔHV management wholesale would be wasteful, because that management matches V3 on most problems and concedes on a few. The refined design **V3R** applies this as a principle, indicator pressure where quality is read and diversity pressure where the population is selected: it retains V3's ΔHV leader and ΔHV archive pruning and restores crowding-based population survival. It adds no parameter, no operator and no tuning; it is a recomposition of components already specified, evaluated once.

On the targeted failure the repair is complete (Table 5, Figure 7): median hypervolume rises from 0.484 to 0.871 on ZDT4, seed-level significant at maximal effect ($W = 0$, $p = 1.9 \times 10^{-9}$, $r = +1.00$). Against the donor, the refinement concedes small margins on the three problems where ΔHV survival is strongest, at most 0.011 median hypervolume, on C1-DTLZ1 ($r = -0.55$), DTLZ2 (-0.99) and C2-DTLZ2 (-0.82), and the across-suite paired test is even (ten wins to twelve, $r = +0.01$, $p = 0.97$). Against the crowding-managed control V3-λ, V3R ranks higher on all three indicators (seventeen problems to five on HV and IGD$^+$ and sixteen to six on the ε-indicator; $r = +0.62$, $+0.59$ and $+0.58$; $p = 0.009$, 0.014 and 0.016), which shows that the retained archive-side ΔHV pruning contributes measurable value. Read along the one-component chain of Table 2, the three steps separate the roles of the leader and of population survival. Replacing the uniform leader by the ΔHV leader under crowding management (V0 to V3-λ) moves the hypervolume mean rank in the ablation family from 3.59 to 1.50 ($r = +0.89$, $p_{\text{holm}} < 0.001$). Adding ΔHV archive pruning (V3-λ to V3R) improves seventeen of the twenty-two problems on HV. Exchanging crowding population survival for smallest-ΔHV survival (V3R to V3) leaves the across-suite paired test even and produces the ZDT4 collapse, against gains of at most 0.011 median hypervolume on three problems. Each step changes one component while the others keep their settings at that step, so each effect is measured in one context and interactions between components were not estimated. Along this chain, the advantage of V3R over the null control goes with the leader rule and the archive pruning, and population survival with the ZDT4 failure and its repair. The anytime view (Figure 8) shows the same repair dynamically: on ZDT4 the refined design escapes the local front early and holds the global level for the remaining budget, and on a problem the repair does not target, the trajectories stay close throughout.

Against the established comparators (Table 6), the two scoring channels of Section 3.4 give channel-specific in-sample answers. Under discovery scoring, no pairwise contrast

between a proposed design and an established comparator survives Holm correction. SMS-EMOA, whose survival step is driven by hypervolume contributions and is therefore aligned with the primary indicator, holds the best descriptive hypervolume mean rank (2.80); the leader-mediated designs and the remaining comparators follow within 1.6 ranks (V3R 3.59 to AGE-MOEA-II 4.34); the largest contrast between them (SMS-EMOA over V3) is $r = -0.42$ at $p_{\text{holm}} = 1.0$; NSGA-II holds the best descriptive ε-indicator rank; and apart from the random floor the only separation is V3R over B5 (5.14), on all three indicators ($r = +0.81$, $+0.79$ and $+0.88$; $p_{\text{holm}} = 0.008$, 0.011 and 0.009), a contrast the nineteen-problem roster does not reproduce. Per protocol the rest is non-separation and not equivalence. Under package scoring, the same runs scored on returned sets, SMS-EMOA (mean rank 2.07) exceeds several established comparators on hypervolume under Holm correction (over NSGA-II at $p_{\text{holm}} = 0.027$, NSGA-III at 0.004 and AGE-MOEA-II at 0.032), while SMS-EMOA and V3R (2.50) are not separated from each other ($r = -0.19$, $p_{\text{holm}} = 1.0$), ahead of NSGA-III (3.50), AGE-MOEA-II (3.93), NSGA-II (4.25) and B5 (4.86); B5 is separated from V3R ($r = +0.90$, $p_{\text{holm}} < 0.001$) and from SMS-EMOA (+0.68, 0.040). In sample, then, the leader-mediated designs are not separated from the modern comparators, AGE-MOEA-II and SMS-EMOA, under either channel, and among the established comparators V3R separates only from MOEA/D, on the ε-indicator in the nineteen-problem package family ($r = +0.78$, $p_{\text{holm}} = 0.034$). On the twenty-two-problem families no contrast of V3R with an established comparator is rejected even before correction ($|r| \leq 0.46$; smallest unadjusted $p = 0.059$). The performance advantage over the established MOEAs rests on the out-of-sample suite.

Because V3R was composed after the full standard-suite results had been observed, its improvement claim is in sample; the external suite carries the out-of-sample burden.

Table 5. The targeted repair: median hypervolume for V3, the crowding-managed reference (V3-λ) and the refined design V3R; seed-level Wilcoxon test of V3R against V3.

Problem	V3	V3-λ	V3R	V3R − V3	
				p	r
ZDT4	0.484	0.870	0.871	1.9×10^{-9}	$+1.00$
C1-DTLZ1	1.096	1.056	1.085	7.1×10^{-3}	-0.55

Table 6. Discovery-channel mean hypervolume ranks of the comparator roster (lower is better). Left: all twenty-two problems, MOEA/D not applicable on the constrained trio. Right: nineteen unconstrained problems including MOEA/D.

Algorithm	22 Problems	19 Unconstrained
SMS-EMOA	2.80	3.16
V3R	3.59	3.42
V3	3.82	3.89
NSGA-II	4.11	4.74
NSGA-III	4.32	4.79
AGE-MOEA-II	4.34	5.11
B5	5.14	5.21
MOEA/D	n/a	5.79
RANDOM	7.89	8.89

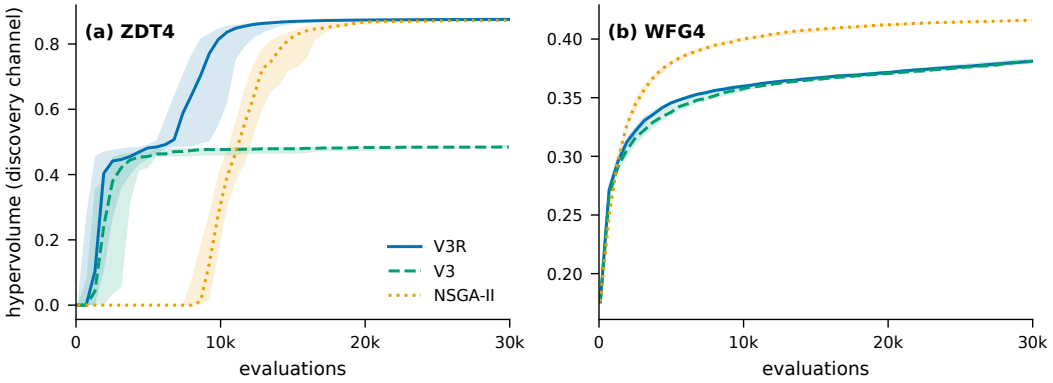


Figure 8. Anytime hypervolume on the discovery channel: the feasible non-dominated set of each run's evaluation prefix, scored in the normalized space of Table A3 (median and interquartile band over all thirty seeds). (**a**) ZDT4: V3 plateaus on a local front while V3R reaches the global-front level within roughly a third of the budget; NSGA-II registers zero hypervolume until its population passes the 1.1 reference clip, then converges to the same level. (**b**) WFG4: the three trajectories rise together, with NSGA-II holding an edge on this instance, consistent with the no-separation reading of the discovery channel.

4.3. External Validation (RQ3)

On the fourteen RE engineering problems, which played no part in any design decision, the in-sample picture reverses: the leader-mediated designs lead on both channels (Table 7). On the package channel V3 and V3R hold the two best mean ranks on all three indicators, and V3R separates under Holm correction from every established comparator on IGD^+ and from four of the five on hypervolume: from NSGA-II at $r = +0.98$ (95% bootstrap CI [+0.89, +1.00] on hypervolume, [+0.87, +1.00] on IGD^+), from NSGA-III and MOEA/D at +1.00 ([+1.00, +1.00]) and from AGE-MOEA-II at +0.94 ([+0.71, +1.00]), with $p_{\text{holm}} \leq 0.030$; from SMS-EMOA at +0.90 ([+0.62, +1.00], $p_{\text{holm}} = 0.022$) on IGD^+, while the same contrast on hypervolume (+0.81, [+0.31, +0.98]) falls short of correction within the thirty-six-pair family ($p_{\text{holm}} = 0.063$). V3 separates from the same four comparators on both indicators. On the ε-indicator the separation holds against the three canonical comparators and the floor and does not hold against the two modern additions (+0.75 and +0.78, $p_{\text{holm}} = 0.150$). The intervals are percentile bootstrap intervals from 10,000 resamples of the fourteen problems drawn as paired blocks, with the rank-biserial statistic recomputed on each resample under a fixed seed; the degenerate interval [+1.00, +1.00] is the ceiling case in which every problem favours V3R, so every resample returns the maximal statistic. The margins are visible in raw medians: on the four-objective RE41 and RE42, V3R attains 0.859 and 0.838 against 0.777 and 0.792 for SMS-EMOA, the strongest comparator, which at four objectives estimates hypervolume contributions by Monte Carlo sampling (Table A2). On the discovery channel, which controls for output policy, V3R holds the best mean rank on hypervolume (3.00) and IGD^+ (2.96), with V3 second on both (3.43 and 3.07); on the ε-indicator SMS-EMOA (3.64) edges V3R (3.71). The contrasts that survive Holm correction in the nine-member external-suite roster are on IGD^+: V3R over NSGA-III ($r = +0.90$, CI [+0.64, +1.00], $p_{\text{holm}} = 0.035$) and over MOEA/D (+0.90, [+0.64, +1.00], 0.035), with NSGA-II short of correction (+0.78, [+0.34, +1.00], 0.236). Under evaluation-trace scoring the two modern comparators are not separated from the leader-mediated designs ($r = +0.16$ to +0.64 for V3R and +0.09 to +0.62 for V3). The divergence between the channels indicates, and does not measure, the returned-set-policy component, because each channel is scored against its own constructed union.

The leading method changes with the reporting-set cap. Descriptive mean ranks of the nine-member roster over the fourteen problems on the discovery channel, with each run's discovered set capped at K points, read as follows on hypervolume: SMS-EMOA

leads at the caps of ten (4.00) and twenty (3.50), V3 at fifty (3.43) and V3R at full capacity (3.00); on IGD^+ AGE-MOEA-II leads at ten, V3 at twenty and fifty and V3R at full capacity; on the ε-indicator AGE-MOEA-II leads at ten and fifty, V3 at twenty and SMS-EMOA at full capacity. At every cap the hypervolume front is led by either a leader-mediated design or the hypervolume-contribution-driven SMS-EMOA; the hypervolume lead passes from SMS-EMOA at the two smaller caps to V3 and V3R at the two larger ones, while on the ε-indicator it returns to an established comparator at full capacity. The transition is not monotonic: V3R's hypervolume mean rank moves from 4.36 at ten to 4.71 at twenty before improving to 3.86 at fifty and 3.00 at full capacity. These are descriptive rankings, not significance findings; the full ladder and the per-cap coverage of every method are in Supplementary Materials S2 (Section S2.8). The engineering fronts are irregular, partly degenerate and shaped by folded constraint-violation objectives, a geometry for which fixed reference directions are poorly matched and on which MOEA/D falls toward the random floor. This is a statement about irregular fronts and not a claim that decomposition is inferior in general; archive-guided leadership assumes no front regularity and adapts to what the archive discovers.

Between V3R and V3 the external suite separates on neither channel. Package-channel hypervolume medians differ by at most 0.0021 with a non-significant lean toward the donor ($r = -0.45$ on HV, $p_{\text{holm}} \geq 0.95$), and on the discovery channel even that lean dissolves ($r = -0.05$ and -0.22 on HV and IGD^+, +0.41 on the ε-indicator). The repair targets a failure mode the RE suite does not contain. Per protocol this is insufficient evidence of a difference and not equivalence, with no directional cost detectable under evaluation-trace scoring.

Table 7. External suite, discovery channel: mean ranks of the nine-member roster over the fourteen RE problems (lower is better).

Algorithm	HV	IGD^+	I_ε^+
V3R	3.00	2.96	3.71
V3	3.43	3.07	4.93
SMS-EMOA	3.61	3.36	3.64
AGE-MOEA-II	3.93	3.79	4.21
NSGA-III	4.96	5.14	5.00
NSGA-II	5.21	5.32	4.93
B5	6.00	5.71	5.79
MOEA/D	6.36	7.07	5.86
RANDOM	8.50	8.57	6.93

4.4. Comparison with the Leader-Mediated Prior Design

Table 8 collects the contrasts between the refined design and B5 from the comparator families of Sections 4.2 and 4.3. B5 ranks sixth or seventh of k in every family but one, ahead of the random floor and behind the Pareto-based comparators; the exception is the nineteen-problem discovery IGD^+ family, where it shares a mean rank of 5.16 with NSGA-II and AGE-MOEA-II. On the twenty-two standard problems its mean hypervolume rank is 4.86 of seven against 2.50 for V3R and 2.07 for SMS-EMOA, and on the fourteen engineering problems it is 6.36 of nine. V3R separates from it under Holm correction on every package-channel family and indicator, with large matched-pairs effects (Table 8). On the engineering suite the effect is at the ceiling: V3R's median hypervolume is higher, and its median IGD^+ and ε-indicator lower, on every one of the fourteen problems. On the discovery channel the effects have the same sign and clear Holm correction only on the twenty-two-problem family; on the nineteen-problem family and on the engineering suite they do not, and on the engineering ε-indicator the estimate is uncertain ($r = +0.52$, CI $[-0.03, +0.90]$). Against the established comparators, B5 is never better under Holm correction than any algorithm

other than the random floor (and MOEA/D on engineering hypervolume under package scoring), while V3R, V3 and SMS-EMOA separate from it on every engineering indicator under package scoring.

The three reconstruction profiles (B5, B5-L and B5-C) do not change this picture. No difference between the reference profile and the schedule-calibrated profile survives Holm correction in any family, which does not establish their equivalence. The literal food-search profile is worse on the twenty-two-problem families, where it is separated from the other two on most indicators. Its feasibility failures concentrate on C1-DTLZ1, where it returns no feasible solution on twenty-six of thirty seeds and is scored at the worst case; the profile differences are significant on the twenty-two-problem families and not on the nineteen-problem families that omit the constrained problems. In the joint families the three profiles occupy the lower-middle positions of each roster, in an order that varies between families.

The intervals are unadjusted percentile bootstrap intervals (10,000 problem-block resamples) reported alongside Holm-adjusted p-values. Where two designs are both scored at the worst case on a problem, that block is a tie and is dropped from the paired test as any zero difference is; this rule was applied to the analysis after the runs and is documented in the supplement. Two boundaries apply. The comparison is one of complete algorithm implementations, including survival and output policies: B5 returns the non-dominated set of its final population, whereas the proposed designs return bounded archives, so the differences do not isolate the effects of archive maintenance or per-agent leadership. The families of this section are those of the comparator study and do not contain B1–B4 or V1, V2 and V4; the comparison of B5 with those designs is that of Section 4.1.

Table 8. V3R against B5 within the comparator families of Sections 4.2 and 4.3. Matched-pairs rank-biserial r (positive favours V3R) with unadjusted 95% percentile bootstrap interval and Holm-adjusted p within the family of $k(k-1)/2$ pairs; k is the roster size.

Channel, Suite (k)	HV		IGD^+		ε-Indicator	
	r [95% CI]	p_{Holm}	r [95% CI]	p_{Holm}	r [95% CI]	p_{Holm}
Package, 22 standard (7)	+0.90 [+0.68, +1.00]	<0.001	+0.88 [+0.64, +1.00]	<0.001	+0.97 [+0.86, +1.00]	<0.001
Package, 19 unconstrained (8)	+0.86 [+0.60, +1.00]	0.007	+0.84 [+0.56, +1.00]	0.011	+0.96 [+0.82, +1.00]	<0.001
Package, 14 engineering (9)	+1.00 [+1.00, +1.00]	0.004	+1.00 [+1.00, +1.00]	0.004	+1.00 [+1.00, +1.00]	0.004
Discovery, 22 standard (8)	+0.81 [+0.53, +0.97]	0.008	+0.79 [+0.49, +0.97]	0.011	+0.88 [+0.64, +0.99]	0.009
Discovery, 19 unconstrained (9)	+0.75 [+0.37, +0.95]	0.079	+0.73 [+0.34, +0.95]	0.110	+0.84 [+0.53, +0.99]	0.052
Discovery, 14 engineering (9)	+0.79 [+0.39, +1.00]	0.175	+0.81 [+0.43, +1.00]	0.121	+0.52 [−0.03, +0.90]	1.000

4.5. Cross-Class Robustness and Parameter Sensitivity

Robustness distinguishes the architectures sharply. The rosters of this subsection differ by statistic: the relative-hypervolume figures draw on the sixteen algorithms of Table 9, the stability figures on the proposed and reconstructed prior designs, the indicator-agreement check on the nine-algorithm prior-method family, and the runtime comparison of Table A4 on every algorithm of the study. Relative hypervolume is each algorithm's median divided by the best median among these sixteen algorithms, one denominator for all; the ablation configurations V0, V2-λ and V3-λ and the B2-alt reading are not included. Several methods fail almost completely on at least one problem: B1, B3, NSGA-II, MOEA/D, AGE-MOEA-II, SMS-EMOA and the random floor all touch a relative hypervolume of zero,

and NSGA-III falls to 0.137 on DTLZ3, B4 to 0.121, B5 to 0.204 on ZDT2 and B2 to 0.319. The proposed V1–V4 family and V3R never drop below 0.337 of the problem-best anywhere in the suite. Table 9 gives the median of each algorithm's relative hypervolume beside its worst case. The proposed designs have medians from 0.968 to 0.985, against 0.962 to 1.000 for the five established MOEAs, and worst cases from 0.337 to 0.342, all on DTLZ3, against at most 0.137 (NSGA-III) for the established MOEAs; the reconstructed prior designs have medians from 0.367 (B4) to 0.905 (B5). Seed-to-seed variability of the proposed designs is four- to sixteen-fold tighter than that of B1–B4 (median cell IQR 0.0035–0.0046 against 0.018–0.056; Figure 9); B5's median cell IQR of 0.0083 lies between the two. The mean-rank orderings of the nine-algorithm family agree across indicators (Spearman rank correlation 1.000 between HV and IGD^+, 0.983 between each of them and the ε-indicator).

Computational cost is the counterweight (Table A4). On the standard suite, fourteen of whose twenty-two problems are bi-objective, V3R's median wall-clock time per run is 5.72 s against 1.11 s for NSGA-III, a ratio of 5.15; V1, V2 and V4 take 4.71, 5.48 and 5.14 s, so on this suite the overhead is common to the archive-based designs. The overhead grows with the number of objectives. The ratio of V3R's median to NSGA-III's rises from 4.55 at $m = 2$ to 7.90 at $m = 3$ on the standard suite, and on the RE suite from 4.40 at $m = 2$ to 31.9 at $m = 3$ and 52.5 at $m = 4$, where V3R takes 50.3 and 90.1 s per run. At $m = 3$ V3R takes 2.79 times as long as V1, against 0.98 times at $m = 2$, which is consistent with the exact hypervolume-contribution pruning of its archive (Section 3.2). SMS-EMOA, which also selects by hypervolume contributions, takes 1.50 s per run at $m = 3$ and 105 s at $m = 4$ on the RE suite; at four objectives pymoo replaces its exact contributions by Monte Carlo estimates. V3 differs from V3R in population survival, which in V3 removes the smallest-ΔHV member of the $5N$ pool iteratively; its median is 20.7 s on the standard suite, with a ninetieth percentile of 210.44 s (236.28 s on the three-objective subset). The difference between V3 and V3R reflects that change together with its effect on the search; it is a difference of aggregate medians, and the time spent inside the survival routine was not measured. These differences would shrink to a small fraction of wall-clock cost in expensive-evaluation regimes, where objective evaluation accounts for most of it, a regime the analytic RE suite does not itself instantiate; for cheap benchmarks the overhead is a real limitation and is reported as one.

Design-side parameter sensitivity was assessed for the refined design on a full 3×3 factorial grid, archive capacity $A_{\max} \in \{50, 100, 200\}$ by population size $N \in \{50, 91, 150\}$, on six problems spanning the four standard families, thirty seeds per configuration, under the shared 30,000-evaluation ceiling (the budget arithmetic of Section 3.1 yields achieved spends of 29,850, 29,939 and 29,550 for the three population sizes). Table 10 reports the median hypervolume of every configuration. The findings read as stability. Population size is essentially inert (spread at most $\approx$0.0043 at fixed capacity). Archive capacity is the one mildly sensitive axis: it raises the median monotonically by between 0.0011 and 0.0312 from $A_{\max}$ = 50 to 200, most visibly on the three-objective problems, a direction consistent with the capacity bound acting on the returned set (the grid itself does not separate output-side truncation from any accompanying change in search dynamics). The pre-specified operating point (100, 91) sits mid-range and was fixed before any result was observed. The centre configuration, re-run as part of this grid on the pinned environment, reproduced all banked returned fronts and hypervolume values exactly, with negligible floating-point differences in some derived distance metrics.

Table 9. Relative hypervolume on the twenty-two standard problems. On each problem, an algorithm's median hypervolume over thirty seeds is divided by the best median among the sixteen algorithms listed; the table gives the median of these ratios over the problems and the smallest ratio (worst case). Values are rounded to three decimals from the stored results. SMS-EMOA's median relative hypervolume, 0.9996, rounds to 1.000; it has the best median hypervolume on ten of the twenty-two problems. MOEA/D has no cells on the three constrained problems; its ratios are taken over the other nineteen, and on the constrained problems the best median is taken over the other fifteen algorithms.

Algorithm	Median	Worst Case
V1	0.973	0.341
V2	0.968	0.342
V3	0.985	0.339
V4	0.979	0.342
V3R	0.982	0.337
B1	0.391	0
B2	0.800	0.319
B3	0.697	0
B4	0.367	0.121
B5	0.905	0.204
NSGA-II	0.973	0
NSGA-III	0.988	0.137
MOEA/D	0.962	0
AGE-MOEA-II	0.983	0
SMS-EMOA	1.000	0
Random floor	0.206	0

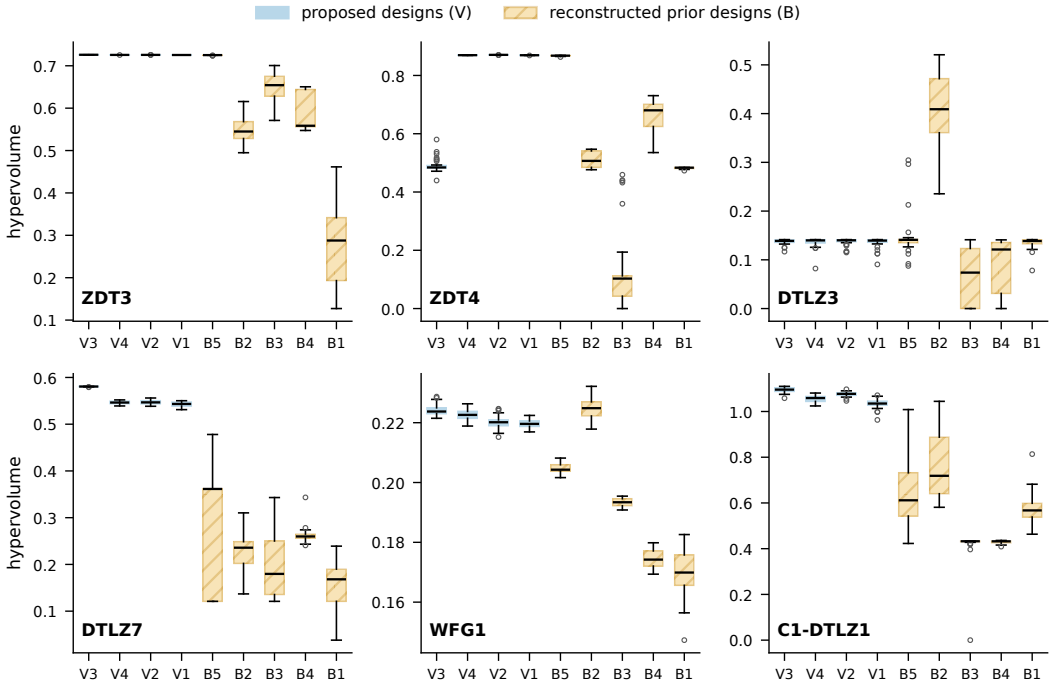


Figure 9. Seed-level hypervolume distributions on six representative problems (thirty seeds; proposed designs in blue, reconstructed prior designs in hatched orange, B5 included). The compressed variant boxes against the baselines' wide spread is the robustness pattern of Section 4.5; DTLZ3, on which the fuzzy-archive baseline B2 exceeds the variant block, is retained deliberately as the principal counter-example.

Table 10. Parameter sensitivity of the refined design V3R: median hypervolume per configuration over thirty seeds (archive capacity $A_{\max}$ × population size N; achieved evaluations per run in parentheses after N).

$A_{\max}$	N	ZDT1	ZDT4	DTLZ2	DTLZ7	WFG4	C2-DTLZ2
50	50 (29,850)	0.867	0.866	0.718	0.566	0.382	0.652
50	91 (29,939)	0.867	0.866	0.719	0.565	0.380	0.650
50	150 (29,550)	0.867	0.866	0.719	0.565	0.377	0.651
100	50 (29,850)	0.872	0.871	0.738	0.580	0.383	0.660
100	91 (29,939)	0.872	0.871	0.740	0.580	0.381	0.663
100	150 (29,550)	0.872	0.871	0.741	0.580	0.379	0.663
200	50 (29,850)	0.874	0.874	0.747	0.588	0.383	0.665
200	91 (29,939)	0.874	0.874	0.750	0.588	0.381	0.665
200	150 (29,550)	0.874	0.874	0.751	0.588	0.380	0.668

5. Discussion

Why leader mediation suits the MGO. Three of four mechanisms consult the leader, so rebinding it to an archive member reshapes three-quarters of each gazelle's candidate generation, per agent and per iteration. This exposure is wider than the single update rule of PSO or GWO; the present study observes the difference and does not isolate it as the cause. The gradient result and the ablation together connect both ingredients with the observed ordering: the random-leader control V0, which keeps in-loop Pareto retention, already beats the best-ranked L0/L1 baseline B2 on sixteen of twenty-two problems, and the decomposition and indicator leader rules add a further increment that the λ-ablation separates from the random draw on HV and IGD$^+$ under Holm correction. The dominance leader V1 separates from V0 on IGD$^+$ only ($p_{\text{holm}} = 0.034$; 0.085 on hypervolume). The deeper lesson concerns the division of labour between leadership and survival, which Table 2 separates component by component: an indicator rule that is effective for choosing leaders can be destructive as a survival rule, as on the multimodal ZDT4 in the tested configuration; separating the two, the principle behind V3R, resolves the failure and keeps the strength.

What in-loop Pareto reasoning contributes beyond prior work. Linear weighted sums cannot recover unsupported Pareto points as global minimizers, even when their weights are varied [26]; a stochastic weighted-sum search may nevertheless visit or retain such points. This limitation does not apply to scalarization in general: ε-constraint and suitably formulated Tchebycheff methods can recover non-convex portions of the front. The L0 designs reconstructed here are a co-optimized weighted sum (B1) and an ε-constraint sweep (B3). Survival-based and archive-based wrappers do use front information, for survival, archiving or the terminal compromise; what they lack under this taxonomy is explicit guide-selection feedback, so the movement guides are chosen by a rule the trade-off information does not reach. Reinforcement-learning meta-control acts orthogonally to leader selection. In the leader-mediated designs the archive selects the guides at every step. In this study that came with a robustness floor and a seed-to-seed stability that the reconstructed wrappers B1–B4 did not show; this study does not isolate whether guide-selection feedback causes that difference, because no comparison switches archive-based guidance on and off with population survival and archive management held fixed. The out-of-sample result against the established comparators is: on problems with no voice in their design, a leader-mediated design holds the best discovery-channel rank on hypervolume and IGD$^+$, under Holm correction on IGD$^+$ against NSGA-III and MOEA/D, and under returned-set scoring separates from all five established comparators on IGD$^+$, the front-geometry-adaptive AGE-MOEA-II and the hypervolume-contribution-driven SMS-EMOA included, and from all but SMS-EMOA on hypervolume; under discovery scoring it is not separated from NSGA-II, AGE-MOEA-II or SMS-EMOA. The contrast

between the two channels is consistent with an output-policy contribution to conventional comparisons. Two within-level contrasts bound this reading. In the prior-method family of Section 4.1, B5 ranks below every proposed design and is not separated from the L1 design B2. In Section 4.4 the refined design separates from B5 under Holm correction on every package-channel family and less consistently under discovery scoring. Both are comparisons of complete implementations and not ablations of the archive or of per-agent assignment, so they show that leader mediation as such does not account for the proposed designs' advantage, without attributing that advantage to any one component.

What is inherited, what is introduced, and what each part does. Because metaphor-based optimizers are routinely challenged as renamed recombinations of familiar operators, Table 11 makes the provenance of every component explicit: what is inherited from the MGO, what this study introduces, which established concept each leader rule transposes, and which comparison supports each component's contribution, with the limit of that inference. Five rows rest on controlled component evidence (the V0 control, the λ-ablation for V1 and V2, V3-λ and V3R); the archive row rests on complete-package comparisons only, and the update-mechanism row on none. Two facts distinguish the resulting designs from existing archive-based optimizers. First, the candidate generator is the MGO's four update mechanisms throughout; there is no recombination or mutation operator of MOEA lineage anywhere in the family, so archive information reaches candidate generation only through the leader binding and not through a replaced survival loop. Second, the null control shows that the distinction is functional: V0 retains the identical archive machinery and the archive-coupled leader draw, removes only the *structured choice* of which member leads, and has the worst mean rank of the λ-ablation on every indicator. The decomposition and indicator leaders separate from it on HV and IGD^+ under Holm correction, and the dominance leader on IGD^+ only, so archive coupling alone does not reproduce the gains of the structured rules where these separate.

Where the approach does not win. The evidence bounds the claim as much as it supports it. In sample, on the standard suite, SMS-EMOA, whose survival step is driven by hypervolume contributions, holds the best descriptive hypervolume and IGD^+ ranks on both channels, V3R the best package-channel ε-indicator rank and NSGA-II the best discovery-channel ε-indicator rank, tied with V3R on the nineteen unconstrained problems. Under discovery scoring no contrast between a proposed design and an established comparator is separable, and under package scoring the separations among the established comparators that survive Holm correction belong to SMS-EMOA, over several of them, and to NSGA-III over MOEA/D (ε-indicator, nineteen problems), with SMS-EMOA and V3R not separated from each other; the contrasts with B5 are the subject of Section 4.4. In sample, then, the leader-mediated designs are at best peers, apart from V3R's separation from MOEA/D in that ε-indicator family; their advantage over the established MOEAs appears out of sample, on irregular fronts, and is not a uniform superiority. Before refinement the best-ranked variant fails outright on one multimodal problem; one prior baseline exceeds the whole variant block on DTLZ3; on the ε-indicator the two modern comparators are not separated even out of sample; and the discovery channel indicates, without measuring it, a returned-set component in the package-channel margins. Wall-clock cost, finally, ranges from about five times NSGA-III's on the standard suite to about fifty times at four objectives (Table A4). These are the boundaries within which the claims hold.

Choosing among the proposed designs. For a choice among V1, V2 and V3R, the evidence supports V3R as the starting choice: it is the only proposed design evaluated out of sample, it removes V3's failure on ZDT4, and on the standard suite its median time per run is 5.7 s against 4.7 s for V1 and 5.5 s for V2; V1 and V2 were not run on the RE suite. On the RE suite V3R's median time per run under the 30,000-evaluation ceiling

is 7.7 s at two objectives and 90 s at four. On cost grounds, V3R suits problems whose evaluations are expensive relative to the optimizer's own computation: when one evaluation takes a second or more, the 29,939 evaluations of one run take more than eight hours in sequence and this overhead is small in comparison, so the choice can rest mainly on solution quality. On problems as cheap to evaluate as these benchmarks, the overhead has to be weighed against any gain in quality. The evidence holds under the 30,000-evaluation ceiling; expensive simulation-based optimization often runs with far smaller budgets and with surrogate models, a regime this study did not test. Beyond four objectives the cost of exact hypervolume contributions grows steeply, and V1 and V2, which compute no hypervolume, are the designs to evaluate there first.

Table 11. Operator correspondence: provenance of every component of the leader-mediated family, the comparison that supports its contribution, and the limit of that inference (controlled = a component ablation under common management; package = a comparison of complete implementations).

Component	Provenance	Nearest Established Concept	Supporting Comparison and Limit of Inference
Four update mechanisms (TSM, MaH, BMH, MSF)	inherited verbatim (validated port)	leader-guided swarm position updates	none: the search substrate itself, not removable
Leader binding λ (*gazmale* $\leftarrow$ archive)	introduced here for MGO	MOPSO/MOGWO leader-from-archive template	controlled (V0): a uniform draw keeps all in-loop machinery and has the worst ablation mean rank, Holm-separated from the decomposition and indicator leaders on HV and IGD^{+}; the structured choice is isolated, the binding itself is not
λ_{V1}: rank and crowding tournament	transposed	NSGA-II selection [3]	controlled (λ-ablation): the leader rule under common management; V1 over V0 is not significant on HV
λ_{V2}: neighbourhood Tchebycheff	transposed	MOEA/D [4]	controlled (V2-λ): the decomposition leader outranks the random draw under common management
λ_{V3}: draw $\propto \Delta$HV	transposed	SMS-EMOA contribution [5]; SMPSOhv contribution tournament [29]	controlled (V3-λ): the leader under crowding is free of the ZDT4 collapse, which, in the tested configuration, implicates survival and not the leader
Bounded archive, feasibility-first admission	shared infrastructure	Pareto-archive tradition [6]	package only: B1–B4 and B5 are complete implementations that differ in more than archive use; no archive-on/archive-off control was run, so the archive's own contribution is not isolated
Split management (V3R)	introduced recomposition	indicator pruning of the archive, as in SMPSOhv [29], with crowding survival of the population	controlled (V3, V3-λ, V3R): restoring crowding survival removes the ZDT4 collapse; retaining ΔHV archive pruning improves over V3-λ

Limitations. The findings apply to the instances tested. The five prior designs B1–B5 are paper-based reconstructions from published descriptions and may diverge from the originals; their assumption ledgers are in Supplementary Materials S2 (B5) and S4 (B1–B4), and under both B5 sensitivity profiles all four proposed designs keep better mean ranks than B5 on every indicator, while B2 and B5 change rank positions (Section 4.1). The study is scoped to problems of at most four objectives: archive-based leader selection at five or more changes regime in archive cardinality, hypervolume-oracle cost and leader-signal sparsity, and is reserved as a distinct question for future work. The refined design carries no convergence proof. Three internal caveats bound the claims: V3R was composed after the standard-suite results had been seen, so the external suite carries the out-of-sample burden; the RE indicators are comparative against constructed reference unions, one per

channel, and not absolute; and V3R's runtime ranges from about five times NSGA-III's on the standard suite to about fifty times at four objectives on the RE suite. The RE suite bounds the external evidence further: its fourteen problems are inexpensive to evaluate, in eleven of them the last objective is the total constraint violation, the union for RE36 has only 28 points, two problems have four objectives, and among the proposed designs only V3R and V3 were run on it. Two caveats concern the construction of the comparison. B2 is held at L1 by an explicit reconstruction, while the domination-driven reading B2-alt is L2 under the same criterion; both are classifications of our reconstructions, because the published description does not settle which its authors implemented, so empirical proximity to B2 does not resolve that ambiguity. V2 pairs a decomposition leader with crowding-based archive pruning, so conclusions involving it concern the specified package and not canonical decomposition. Two further caveats concern the census and the primary comparison. The census rests on a forward-citation screen of twenty works, corroborated by two independent surveys [15,46]; three further records qualified on their abstracts only and are held as provisional in Supplementary Materials S1, outside the count. Three of the twenty works describe multi-objective machinery without stating any guide-selection rule, so their L1 level is provisional and inferred from that machinery; none of the three states a rule that would place it at L2, but neither do they exclude one, and for one of them an equations-only assessment of the application formulation supports L0. As with any census, unindexed work cannot be excluded. The prior-method comparison remains a package comparison, in which the reconstructed designs return what their reconstructed pipelines return, with the adaptations and omissions stated in Section 3.3, and with the discovery channel applied to the comparator studies where output policies differ.

6. Conclusions and Future Work

For the Mountain Gazelle Optimizer, moving multi-objective reasoning inside the update loop is consequential on the suites and indicators evaluated. Of the prior multi-objective MGO designs, one mediates its leader in-loop, through a single broadcast guide inside a modified update scheme; the designs studied here do so by a single operational substitution of the leader that leaves every original movement equation algebraically intact, with per-agent leaders drawn from a Pareto archive. On the standard suite, the best-ranked variant's mean hypervolume rank is 2.14 against 4.84 for the strongest prior design, B5, all four proposed designs have better mean ranks than the five prior designs on each indicator, and fifteen of the twenty proposed-versus-prior contrasts are significant on hypervolume, fourteen on IGD^+ and twenty on the ε-indicator. Under evaluation-trace scoring their (V3R and V3) discovered trade-off sets show no detected separation from the established comparators in sample, where SMS-EMOA holds the best descriptive hypervolume and IGD^+ ranks. Out of sample they (V3R and V3) hold the best ranks on hypervolume and IGD^+ at the protocol's capacity, under Holm correction (V3R only) on IGD^+ against NSGA-III and MOEA/D with evaluation-trace scoring; with returned-set scoring the refined design separates from all five established comparators, AGE-MOEA-II and SMS-EMOA included, on IGD^+, and from all but SMS-EMOA on hypervolume. Against B5, the refined design also separates on every returned-set family of the comparator rosters and less consistently under evaluation-trace scoring. Across the standard suite, the proposed designs' worst-case relative hypervolume is at least 0.337, which none of the reconstructed prior designs attains. Separately, the six-problem parameter study of the refined design shows limited sensitivity to the tested population sizes and archive capacities. These results were obtained as a controlled structural experiment, with a null control, a leader-isolation ablation and a two-channel reporting protocol that scores discovered and returned trade-off sets separately, and they show, for MGO, what structured leader selection from a Pareto archive adds under

common management; the design template behind them, consisting of the feedback-path taxonomy, the leader substitution, the null control and the two-channel scoring, can be applied to other leader-based swarm optimizers. The contribution is bounded accordingly: it is a result about where Pareto reasoning resides in one optimizer family, established at two to four objectives on thirty-six benchmark problems, and not a general superiority claim over modern multi-objective evolutionary algorithms, which the in-sample evidence would not support.

Three questions follow directly. A constrained real-world comparison on the reserved subset of Kumar et al. [39] would test the feasibility-first archive under active constraints. Online adaptive switching among leader strategies is motivated by the statistically inseparable variants and their class-conditional dominance. A convergence analysis of the multiple-leader dynamics is open and non-trivial, because different agents follow different leaders within a single iteration. Reference-direction-aware archive maintenance and restart or adaptive leader scheduling, suggested by the residual weaknesses, are reserved for a subsequent design cycle.

On the evidence of this study, the multi-objective performance of the MGO depends on which Pareto mechanism selects its leaders and on how its population and archive are managed. Leader-mediated feedback places trade-off information directly in candidate generation, and the size of the proposed designs' advantage depends on the problem class and on the evaluation protocol.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/math14214043/s1, Supplementary Materials S1, literature-census protocol and per-record screening log (S1_census.pdf); Supplementary Materials S2, the B5 reconstruction of Pham et al. [23]: profiles and assumptions ledger, environment record, the external-suite reference unions, the complete Holm families of the comparator study (Sections 4.2–4.4) with bootstrap intervals, and the external-suite reporting-cap ladder with per-cap coverage (S2_extension.pdf); Supplementary Materials S3, the prior-method family: all pairwise contrasts with bootstrap intervals, the B5 profile sensitivity and an exploratory comparison of the null control V0 with B1–B4 (S3_prior_family.pdf); Supplementary Materials S4, the reconstruction ledger of B1–B4: sources, gap-fills, omitted stages and conventions (S4_baseline_ledger.pdf). References [47–69] are cited in the Supplementary Materials.

Author Contributions: Conceptualization, P.T. and C.J.; methodology, P.T.; software, P.T.; validation, P.T.; formal analysis, P.T.; investigation, P.T.; data curation, P.T.; writing—original draft preparation, P.T.; writing—review and editing, P.T., C.J. and S.K.; visualization, P.T.; supervision, C.J. and S.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research was financially supported by Mahasarakham University.

Institutional Review Board Statement: The Research Ethics Committee of Mahasarakham University reviewed and approved this study (Approval No. 451-447/2026).

Informed Consent Statement: Not applicable.

Data Availability Statement: The source code, test suite, campaign runners, analysis scripts and analysis files supporting this study are openly available in Zenodo at https://doi.org/10.5281/zenodo.23831874 (version v2.13) under the GNU General Public License v3.0. Every number reported in the article and in Supplementary Materials S2 and S3 is read from these analysis files. The raw per-cell result archives, including the evaluation traces (about 5.8 GB), are not deposited because of their size and are available from the corresponding author on reasonable request. The campaigns are deterministic and can be rerun with the archived code; bit-level replay is scoped to the pinned computational environment of Table A1.

Acknowledgments: During the preparation of this manuscript, the authors used ChatGPT (GPT-6 Astra; OpenAI) and Claude (Claude Opus 5.5; Anthropic) to improve language, grammar, readability,

and academic style; to analyse and review the study's code; and to help verify the data analysis. These tools were not used for data collection, interpretation of the results, or formulation of the scientific conclusions. The authors critically reviewed, verified, and revised all AI-assisted suggestions and take full responsibility for the manuscript's content.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A. Complete Experimental Configuration

Tables A1–A3 summarize the principal operative settings of the protocol, and Table A4 reports the measured runtimes. The released code and the locked environment (with the hypervolume-backend qualification in Table A1) are authoritative for every value not listed.

Table A1. Experimental configuration summary, part 1: algorithms and budgets. All values are fixed by the protocol before any comparison; nothing is tuned per problem.

Setting	Value
Software stack	CPython 3.14.4; numpy 2.4.4; scipy 1.17.1; pymoo 0.6.1.6; installed from the hashed lock (`requirements.lock`), with the one exception recorded in the next row
Hypervolume backend (moocore)	*Main study:* the version was not recorded in the run manifests; the lock specifies 0.3.2, whereas the installation of record on the campaign machine has carried 0.2.0 since before the campaigns, an inference from installation history and not a record. *Compatibility evidence, separate from provenance:* 0.2.0 and 0.3.2 agree to the bit on sixteen representative fronts; agreement between versions cannot identify which version ran. *B5 (Sections 4.1 and 4.4):* moocore 0.2.0 recorded in every cell (macOS 14.6, arm64)
Platform	macOS 14.6, Apple silicon (arm64); single process; bit-level replay scoped to this platform
Evaluation budget	ceiling 30,000 per run; single-run designs achieve $N + \lfloor (30000 - N)/4N \rfloor \cdot 4N = 29{,}939$ at $N = 91$ ($T = 82$ iterations of $4N$ candidates after initialization); sweep baselines spend less under the same ceiling (see the sweep row)
Population/archive	$N = 91$; $A_{\max} = 100$ bounds the post-pruning archive only (the pre-pruning admission set may transiently exceed it); feasibility-first dominance admission; per-design pruning (crowding for V0/V1/V4 and V3R's population; smallest-ΔHV for V3 and for the V3/V3R archive)
Constraint handling	feasibility-first rule of Deb [34]; total violation $CV = \sum_j \max(g_j, 0)$ over the constraints $g_j \leq 0$. Archive admission: a feasible point dominates an infeasible one, and infeasible points are compared by violation alone, so the archive holds either feasible members only or minimum-violation infeasible members only. Crowding and smallest-ΔHV survival: feasible members first, by the operator's criterion, then infeasible members by ascending violation. V2 survival: feasibility, violation and Tchebycheff value, in that order; the V2 ideal point z^* is updated from all evaluated objective vectors, infeasible ones included. B1–B4: the same rule on their scalar (Supplementary Materials S4). pymoo comparators: pymoo's feasibility-first handling; pymoo's MOEA/D rejects constrained problems, so MOEA/D is not applicable to the C-DTLZ problems
Seeds and RNG	seeds 0–29 (thirty per cell); PCG64 streams from one fixed SeedSequence contract: a shared initialization stream plus one search stream per algorithm id
Initial population	one shared uniform 91-member sample per (problem, seed), injected into every single-run native-space design and every pymoo comparator; the sweep baselines' scalar sub-runs draw their own 30-member initializations from dedicated streams
Variation operators (MGO family)	the four MGO mechanisms (TSM, MaH, BMH, MSF) of the validated synchronous port, coefficient vectors exactly as in Abdollahzadeh et al. [14]; no auxiliary crossover or mutation

Table A2. Experimental configuration summary, part 2: leader rules, comparators and sweep budgets.

Setting	Value
V2 leader rule	deterministic: for the gazelle's weight vector w_i, the archive member that minimizes $g^{\mathrm{te}}(f \mid w_i, z^*) = \max_j w_{ij}\,(\lvert f_j - z_j^*\rvert + 10^{-10})$ among the members whose best-fitting weight vector lies in the neighbourhood $B(i)$ of w_i (all members if none does); 91-vector weight lattice, neighbourhood size 10; z^* is the running ideal point
V3 leader rule	draw probability proportional to each archive member's ΔHV contribution, clamped at zero and offset by 10^{-3} of the mean contribution; uniform fallback when all contributions vanish
ΔHV normalization (V3, V3R)	for the set S under consideration (the archive $A(t)$, fixed within the iteration, for the leader rule; the pool or the post-admission archive for smallest-ΔHV removal), each objective is mapped from $[\min_S f_j, \max_S f_j]$ to $[0, 1]$, and an objective with zero range maps to 0; reference point 1.1 per objective; $\Delta\mathrm{HV}(a) = \mathrm{HV}(S) - \mathrm{HV}(S \setminus \{a\})$ by exact leave-one-out under that one transform, clamped at zero; during smallest-ΔHV removal the transform and all contributions are recomputed after each removal. This normalization is internal to the designs and separate from the scoring normalization of Table A3
Comparator configuration	pymoo 0.6.1.6 per-algorithm defaults: SBX with polynomial mutation ($\eta_m = 20$) throughout, with $\eta_c = 15$ and crossover probability 0.9 for NSGA-II, AGE-MOEA-II and SMS-EMOA, $\eta_c = 20$ and probability 1.0 for MOEA/D, and $\eta_c = 30$ and probability 1.0 for NSGA-III; MOEA/D with 20-neighbour decomposition and neighbour-mating probability 0.9; SMS-EMOA in pymoo's generational form (91 offspring per generation) with internal normalization, reference point $1 + \varepsilon$ per objective, and Monte Carlo-approximated hypervolume contributions at $m = 4$ (10,000 samples, seed 1; exact at $m \leq 3$); population 91 throughout; AGE-MOEA-II executed on the pure-Python path (CPython 3.14.4; Table A1)
Reference directions	Das–Dennis: $m = 2$, $H = 90$ (91 directions); $m = 3$, $H = 12$ (91); $m = 4$, two layers, $H = 5$ boundary (56) plus $H = 4$ inner scaled 0.5 (35), 91 in total
Sweep baselines	under the shared 30,000-evaluation ceiling, B4 and two-objective B3 run 21 scalar sub-runs of population 30 for 11 iterations each ($21 \times (30 + 11 \cdot 120) = 28{,}350$ evaluations); three-objective B3 runs 19 sub-runs of population 30 for 12 iterations ($19 \times (30 + 12 \cdot 120) = 27{,}930$); B1/B2 are single runs (Section 3.3)

Table A3. Experimental configuration summary, part 3: scoring, channels and statistics.

Setting	Value
Reference sets	standard suite: fixed analytic Pareto fronts, deduplicated and filtered to non-dominated points; ZDT1, ZDT2, ZDT4 and ZDT6: 1000 points of the closed-form front; ZDT3: pymoo's disconnected front (97 points); DTLZ1–4: 1035 points on a Das–Dennis lattice ($H = 44$); DTLZ7: 900 points from a 60×60 grid of the closed-form front; WFG1–9: closed-form fronts from 2001-point sweeps of the shape functions of Huband et al. [37] (547 points for WFG2, 2001 otherwise); C1-DTLZ1, C2-DTLZ2 and C3-DTLZ4: pymoo's fixed sets (136, 76 and 136 points). RE suite: the non-dominated union of the returned sets (package channel) or of the trace-derived sets (discovery channel) of all participating algorithms and seeds, one union per channel (Supplementary Materials S2, Section S2.5)
Normalization	all indicators in the normalized objective space $(f - z^{\mathrm{ideal}})/(z^{\mathrm{nadir}} - z^{\mathrm{ideal}})$, with z^{ideal} and z^{nadir} the per-objective minimum and maximum of the problem's reference set; zero-range coordinates divide by one
Hypervolume	exact, computed with the moocore backend (version: see Software stack); reference point 1.1 per normalized objective, which is the nadir of the reference set inflated by 10% of its range
Other indicators	IGD^+, GD and the additive ε-indicator against the problem's reference set; additive ε by its max–min–max definition, $I_\varepsilon^+(A, R) = \max_{r \in R} \min_{a \in A} \max_j (a_j - r_j)$; Schott spacing (sample s.d., ddof = 1)

Table A3. *Cont.*

Setting	Value
Scoring policy	indicators computed on the feasible non-dominated subset of each return; an empty feasible set scores worst-case and is flagged, never dropped
Discovery channel	passive trace capture at the counted-evaluator boundary; feasible non-dominated set of the complete trace in canonical order; one terminal capacity-100 truncation by iterative smallest-crowding removal; cardinality caps $\{10, 20, 50, 100\}$ descriptive only
Statistics	Friedman across problems on median-over-thirty-seeds cells; Holm-corrected Wilcoxon signed-rank post hoc within each roster–indicator–channel family; matched rank-biserial effect sizes; percentile bootstrap CIs (10,000 problem-block resamples, fixed seed); $\alpha = 0.05$

Table A4. Median wall-clock time per run in seconds (three significant figures), by suite and number of objectives m. The time is measured around the optimizer call only and excludes scoring. Standard suite: fourteen problems at $m = 2$ and eight at $m = 3$, thirty seeds each; MOEA/D ran only on the nineteen unconstrained problems. RE suite: five problems at $m = 2$, seven at $m = 3$ and two at $m = 4$; a dash marks a design that was not run on it. "All" is the median over all runs of the suite. Machine load was not controlled, so differences of about 20% or less are not interpretable. B5 was timed with its evaluation-trace recording inside the timed call, and AGE-MOEA-II ran pymoo's numba-decorated code as plain Python.

Algorithm	Standard Suite			RE Suite			
	$m = 2$	$m = 3$	All	$m = 2$	$m = 3$	$m = 4$	All
V0	5.01	3.14	4.56	–	–	–	–
V1	5.07	3.18	4.71	–	–	–	–
V2	6.01	3.55	5.48	–	–	–	–
V3	12.6	160	20.7	58.2	227	447	189
V4	5.35	3.48	5.14	–	–	–	–
V2-λ	5.48	3.49	5.13	–	–	–	–
V3-λ	5.21	3.51	5.01	–	–	–	–
V3R	4.98	8.87	5.72	7.72	50.3	90.1	34.3
B1	0.497	0.491	0.493	–	–	–	–
B2	8.74	2.03	6.09	–	–	–	–
B2-alt	2.94	1.76	2.45	–	–	–	–
B3	0.402	0.378	0.395	–	–	–	–
B4	0.402	0.381	0.398	–	–	–	–
B5	0.287	0.284	0.286	0.380	0.401	0.576	0.406
NSGA-II	0.982	0.983	0.982	1.50	1.39	1.61	1.43
NSGA-III	1.09	1.12	1.11	1.76	1.58	1.72	1.61
MOEA/D	10.5	10.6	10.5	10.3	10.9	11.0	10.8
AGE-MOEA-II	14.8	10.8	14.2	21.2	26.5	32.1	25.8
SMS-EMOA	0.931	1.13	1.00	1.17	1.50	105	1.36
Random floor	0.496	0.701	0.570	1.13	1.22	7.30	1.26

References

1. Deb, K. *Multi-Objective Optimization Using Evolutionary Algorithms*; John Wiley & Sons: Chichester, UK, 2001.
2. Coello Coello, C.A.; Lamont, G.B.; Van Veldhuizen, D.A. *Evolutionary Algorithms for Solving Multi-Objective Problems*, 2nd ed.; Genetic and Evolutionary Computation Series; Springer: New York, NY, USA, 2007. https://doi.org/10.1007/978-0-387-36797-2.
3. Deb, K.; Pratap, A.; Agarwal, S.; Meyarivan, T. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Trans. Evol. Comput.* **2002**, *6*, 182–197. https://doi.org/10.1109/4235.996017.
4. Zhang, Q.; Li, H. MOEA/D: A multiobjective evolutionary algorithm based on decomposition. *IEEE Trans. Evol. Comput.* **2007**, *11*, 712–731. https://doi.org/10.1109/TEVC.2007.892759.
5. Beume, N.; Naujoks, B.; Emmerich, M. SMS-EMOA: Multiobjective selection based on dominated hypervolume. *Eur. J. Oper. Res.* **2007**, *181*, 1653–1669. https://doi.org/10.1016/j.ejor.2006.08.008.

6. Coello Coello, C.A.; Salazar Lechuga, M. MOPSO: A proposal for multiple objective particle swarm optimization. In Proceedings of the 2002 Congress on Evolutionary Computation (CEC '02), Honolulu, HI, USA, 12–17 May 2002; Volume 2, pp. 1051–1056. https://doi.org/10.1109/CEC.2002.1004388.
7. Mirjalili, S.; Saremi, S.; Mirjalili, S.M.; Coelho, L.d.S. Multi-objective grey wolf optimizer: A novel algorithm for multi-criterion optimization. *Expert Syst. Appl.* **2016**, *47*, 106–119. https://doi.org/10.1016/j.eswa.2015.10.039.
8. Mirjalili, S.; Jangir, P.; Mirjalili, S.Z.; Saremi, S.; Trivedi, I.N. Optimization of problems with multiple objectives using the multi-verse optimization algorithm. *Knowl.-Based Syst.* **2017**, *134*, 50–71. https://doi.org/10.1016/j.knosys.2017.07.018.
9. Mirjalili, S.; Gandomi, A.H.; Mirjalili, S.Z.; Saremi, S.; Faris, H.; Mirjalili, S.M. Salp Swarm Algorithm: A bio-inspired optimizer for engineering design problems. *Adv. Eng. Softw.* **2017**, *114*, 163–191. https://doi.org/10.1016/j.advengsoft.2017.07.002.
10. Sörensen, K. Metaheuristics—The metaphor exposed. *Int. Trans. Oper. Res.* **2015**, *22*, 3–18. https://doi.org/10.1111/itor.12001.
11. Aranha, C.; Camacho Villalón, C.L.; Campelo, F.; Dorigo, M.; Ruiz, R.; Sevaux, M.; Sörensen, K.; Stützle, T. Metaphor-based metaheuristics, a call for action: The elephant in the room. *Swarm Intell.* **2022**, *16*, 1–6. https://doi.org/10.1007/s11721-021-00202-9.
12. Vermetten, D.; Doerr, C.; Wang, H.; Kononova, A.V.; Bäck, T. Large-Scale Benchmarking of Metaphor-Based Optimization Heuristics. In Proceedings of the Genetic and Evolutionary Computation Conference (GECCO '24), Melbourne, VIC, Australia, 14–18 July 2024; pp. 41–49. https://doi.org/10.1145/3638529.3654122.
13. Soto Calvo, M.; Lee, H.S. Systematic taxonomic framework of metaheuristic algorithms using hierarchical clustering and structural criteria: How novel is the novelty? *Artif. Intell. Rev.* **2025**, *59*, 61. https://doi.org/10.1007/s10462-025-11456-8.
14. Abdollahzadeh, B.; Gharehchopogh, F.S.; Khodadadi, N.; Mirjalili, S. Mountain Gazelle Optimizer: A new nature-inspired metaheuristic algorithm for global optimization problems. *Adv. Eng. Softw.* **2022**, *174*, 103282. https://doi.org/10.1016/j.advengsoft.2022.103282.
15. Anka, F.; Gharehchopogh, F.S.; Tejani, G.G.; Mousavirad, S.J. Advances in Mountain Gazelle Optimizer: A comprehensive study on its classification and applications. *Int. J. Comput. Intell. Syst.* **2025**, *18*, 247. https://doi.org/10.1007/s44196-025-00968-4.
16. Said, S.M.; Abdel-Salam, M.; Nayel, M.; Hashem, M.; Kamel, S.; Jurado, F.; Ebeed, M. Optimal design and cost of superconducting magnetic energy storage for voltage sag mitigation in a real distribution network. *J. Energy Storage* **2023**, *73*, 108864. https://doi.org/10.1016/j.est.2023.108864.
17. Pandey, M.; Baltacı, Y.; Mishra, S.; Sharma, M.; Bhattarai, S.K.; Sethi, K.C. Modified adaptive weight multi-objective mountain gazelle optimizer for construction time-cost trade-off optimization problems. *Asian J. Civ. Eng.* **2026**, *27*, 1937–1946. https://doi.org/10.1007/s42107-025-01594-3.
18. Pandey, A.K.; Jadoun, V.K.; Sabhahit, J.N.; Sharma, S. Interconnected operation and economic feasibility-based sustainable planning of virtual power plant in multi-area context. *Smart Cities* **2025**, *8*, 37. https://doi.org/10.3390/smartcities8010037.
19. Ramachandra, N.; Natarajan, R. State-of-the-art and real-time implementation of an IoT-based home energy management system for a cluster of dwellings. *Heliyon* **2024**, *10*, e35887. https://doi.org/10.1016/j.heliyon.2024.e35887.
20. Zheng, H.; Wu, Z. Enhancing economic and environmental performance of energy communities: A multi-objective optimization approach with mountain gazelle optimizer. *Sustain. Comput. Inform. Syst.* **2025**, *46*, 101098. https://doi.org/10.1016/j.suscom.2025.101098.
21. Refaat, M.M.; Al-Dhaifallah, M.; Ismail, M.M.; AbouOmar, M.S.; El Ferik, S. Dimensionality-reduced multi-objective framework for resilient power systems with electric vehicles and hydrogen-based energy storage. *Sci. Rep.* **2025**, *15*, 45150. https://doi.org/10.1038/s41598-025-33349-0.
22. Eirgash, M.A.; Tiang, J.J.; Ateş, B.; Sharma, A.; Lim, W.H. A Reinforcement Learning Guided Oppositional Mountain Gazelle Optimizer for Time–Cost–Risk Trade-Off Optimization Problems. *Buildings* **2026**, *16*, 144. https://doi.org/10.3390/buildings16010144.
23. Pham, V.H.S.; Dau, T.D.; Tran, L.A.; Nguyen Dang, N.T. A novel modified Mountain Gazelle Optimizer for enhancing resource scheduling problem in multi-project environment. *Int. J. Manag. Sci. Eng. Manag.* **2026**, 1–23. https://doi.org/10.1080/17509653.2026.2694402.
24. Tanabe, R.; Ishibuchi, H. An easy-to-use real-world multi-objective optimization problem suite. *Appl. Soft Comput.* **2020**, *89*, 106078. https://doi.org/10.1016/j.asoc.2020.106078.
25. Miettinen, K. *Nonlinear Multiobjective Optimization*; International Series in Operations Research & Management Science; Kluwer Academic Publishers: Boston, MA, USA, 1999; Volume 12. https://doi.org/10.1007/978-1-4615-5563-6.
26. Das, I.; Dennis, J.E. A closer look at drawbacks of minimizing weighted sums of objectives for Pareto set generation in multicriteria optimization problems. *Struct. Optim.* **1997**, *14*, 63–69. https://doi.org/10.1007/BF01197559.
27. Deb, K.; Jain, H. An evolutionary many-objective optimization algorithm using reference-point-based nondominated sorting approach, Part I: Solving problems with box constraints. *IEEE Trans. Evol. Comput.* **2014**, *18*, 577–601. https://doi.org/10.1109/TEVC.2013.2281535.

28. Padhye, N.; Branke, J.; Mostaghim, S. Empirical comparison of MOPSO methods: Guide selection and diversity preservation. In Proceedings of the 2009 IEEE Congress on Evolutionary Computation, Trondheim, Norway, 18–21 May 2009; pp. 2516–2523. https://doi.org/10.1109/CEC.2009.4983257.
29. Nebro, A.J.; Durillo, J.J.; Coello Coello, C.A. Analysis of leader selection strategies in a multi-objective particle swarm optimizer. In Proceedings of the 2013 IEEE Congress on Evolutionary Computation, Cancun, Mexico, 20–23 June 2013; pp. 3153–3160. https://doi.org/10.1109/CEC.2013.6557955.
30. Dai, S. Optimal energy management of multi-energy multi-microgrid networks using Mountain Gazelle Optimizer for cost and emission reduction. *Energy* **2025**, *329*, 136640. https://doi.org/10.1016/j.energy.2025.136640.
31. Chrouda, A.; Alsenani, T.R.; Alamri, S.; Agrawal, M.K.; Sadeq, A.M.; Khan, M.S.; Ooi, J.B. An innovative green methanol synthesis through natural gas–biogas integration and post-combustion CO_2 capture: An integrated 4E evaluation, life cycle assessment and machine learning-driven multi-objective optimization. *Appl. Therm. Eng.* **2026**, *300*, 131462. https://doi.org/10.1016/j.applthermaleng.2026.131462.
32. Al-Dhaifallah, M.; Refaat, M.M.; Alaas, Z.; Abdel Aleem, S.H.E.; El-kholy, E.E.; Ali, Z.M. Multi-objectives transmission expansion planning considering energy storage systems and high penetration of renewables and electric vehicles under uncertain conditions. *Energy Rep.* **2024**, *11*, 4143–4164. https://doi.org/10.1016/j.egyr.2024.03.060.
33. Dede, T.; Eirgash, M.A.; Kysiak, A.; Uçan, H.A. Enhanced multi-objective Mountain Gazelle Optimization via modified adaptive weight approach for construction time-cost trade-off problems. *Constr. Optim. Energy Potential* **2026**, *15*, 32–46. https://doi.org/10.17512/bozpe.2026.15.04.
34. Deb, K. An efficient constraint handling method for genetic algorithms. *Comput. Methods Appl. Mech. Eng.* **2000**, *186*, 311–338. https://doi.org/10.1016/S0045-7825(99)00389-8.
35. Zitzler, E.; Deb, K.; Thiele, L. Comparison of multiobjective evolutionary algorithms: Empirical results. *Evol. Comput.* **2000**, *8*, 173–195. https://doi.org/10.1162/106365600568202.
36. Deb, K.; Thiele, L.; Laumanns, M.; Zitzler, E. Scalable multi-objective optimization test problems. In Proceedings of the 2002 Congress on Evolutionary Computation (CEC '02), Honolulu, HI, USA, 12–17 May 2002; Volume 1, pp. 825–830. https://doi.org/10.1109/CEC.2002.1006980.
37. Huband, S.; Hingston, P.; Barone, L.; While, L. A review of multiobjective test problems and a scalable test problem toolkit. *IEEE Trans. Evol. Comput.* **2006**, *10*, 477–506. https://doi.org/10.1109/TEVC.2005.861417.
38. Jain, H.; Deb, K. An evolutionary many-objective optimization algorithm using reference-point based nondominated sorting approach, Part II: Handling constraints and extending to an adaptive approach. *IEEE Trans. Evol. Comput.* **2014**, *18*, 602–622. https://doi.org/10.1109/TEVC.2013.2281534.
39. Kumar, A.; Wu, G.; Ali, M.Z.; Luo, Q.; Mallipeddi, R.; Suganthan, P.N.; Das, S. A benchmark-suite of real-world constrained multi-objective optimization problems and some baseline results. *Swarm Evol. Comput.* **2021**, *67*, 100961. https://doi.org/10.1016/j.swevo.2021.100961.
40. Ishibuchi, H.; Masuda, H.; Tanigaki, Y.; Nojima, Y. Modified distance calculation in generational distance and inverted generational distance. In *Proceedings of the Evolutionary Multi-Criterion Optimization (EMO 2015)*; Gaspar-Cunha, A., Henggeler Antunes, C., Coello Coello, C.A., Eds.; Lecture Notes in Computer Science; Springer International Publishing: Cham, Switzerland, 2015; Volume 9019, pp. 110–125. https://doi.org/10.1007/978-3-319-15892-1_8.
41. Friedman, M. The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *J. Am. Stat. Assoc.* **1937**, *32*, 675–701. https://doi.org/10.1080/01621459.1937.10503522.
42. Wilcoxon, F. Individual comparisons by ranking methods. *Biom. Bull.* **1945**, *1*, 80–83. https://doi.org/10.2307/3001968.
43. Holm, S. A simple sequentially rejective multiple test procedure. *Scand. J. Stat.* **1979**, *6*, 65–70.
44. Derrac, J.; García, S.; Molina, D.; Herrera, F. A practical tutorial on the use of nonparametric statistical tests as a methodology for comparing evolutionary and swarm intelligence algorithms. *Swarm Evol. Comput.* **2011**, *1*, 3–18. https://doi.org/10.1016/j.swevo.2011.02.002.
45. Panichella, A. An improved Pareto front modeling algorithm for large-scale many-objective optimization. In *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO '22)*; ACM: New York, NY, USA, 2022; pp. 565–573. https://doi.org/10.1145/3512290.3528732.
46. Sholeh, M.N.; Elizar, E. A systematic review of the Mountain Gazelle Optimizer: Methodological variants and application domains. *Evol. Intell.* **2026**, *19*, 79. https://doi.org/10.1007/s12065-026-01196-1.
47. Abdelsattar, M.; Mesalam, A.; Fawzi, A.; Hamdan, I. Mountain Gazelle Optimizer for standalone hybrid power system design incorporating a type of incentive-based strategies. *Neural Comput. Appl.* **2024**, *36*, 6839–6853. https://doi.org/10.1007/s00521-024-09433-3.
48. Mahdad, B. Improved Mountain Gazelle Optimizer based interactive distributed strategy to solving large scale OPF. *Eng. Res. Express* **2024**, *6*, 025341. https://doi.org/10.1088/2631-8695/ad4cb7.

49. Salimon, S.A.; Fajinmi, I.O.; Adebiyi, O.W.; Atilola, O.F. Simultaneous allocation of distributed generation and shunt capacitor using multi-objective Mountain Gazelle Optimizer for technical and economic benefits. *Eng. Res. Express* **2025**, *7*, 045389. https://doi.org/10.1088/2631-8695/ae1d19.
50. Luu, N.C. Applying BIM technology and MGO algorithm to optimize construction time and cost for construction foundations categories. *Evol. Intell.* **2025**, *18*, 123. https://doi.org/10.1007/s12065-025-01104-z.
51. Pravija Raj, P.V.; Pawar, P.M.; Khedr, A.M. AI-Driven Optimization of Multi-Objective Task Scheduling in Fog-Assisted IoT Environments. In *Cognition-Based Smart Computing: Foundations, Frameworks, and Applications*; Chaudhari, S.S., Chavan, P.V., Mahalle, P., Pawar, P.M., Eds.; CRC Press: Boca Raton, FL, USA, 2026; pp. 1–18. https://doi.org/10.1201/9781003501220-1.
52. Zubair Iftikhar, M.; Imran, K.; Imran Akbar, M.; Ghafoor, S. Optimal distributed generators allocation with various load models under load growth using a meta-heuristic technique. *Renew. Energy Focus* **2024**, *49*, 100550. https://doi.org/10.1016/j.ref.2024.100550.
53. Purnomo, A.; Mustika, I.W.; Wijaya, F.D. A Diversity-Aware Fuzzy-Adaptive Mountain Gazelle Optimizer for Home Energy Management. In Proceedings of the 2025 IEEE International Conference on Communication, Networks and Satellite (ComNetSat), Padang, Indonesia, 11–13 December 2025; pp. 498–504. https://doi.org/10.1109/COMNETSAT68601.2025.11324505.
54. Pratap, A.; Tiwari, P.; Maurya, R.; Maurya, P. Adaptive genetic-enhanced Mountain Gazelle Optimization for integrated network reconfiguration, distributed generation, and capacitor bank planning in large-scale radial distribution systems. *Electr. Eng.* **2025**, *107*, 6571–6609. https://doi.org/10.1007/s00202-024-02871-1.
55. Salimon, S.A.; Fajinmi, I.O.; Adewuyi, O.B.; Pandey, A.K.; Adebiyi, O.W.; Kotb, H. Graph theory-enhanced integrated distribution network reconfiguration and distributed generation planning: A comparative techno-economic and environmental impacts analysis. *Clean. Eng. Technol.* **2024**, *22*, 100808. https://doi.org/10.1016/j.clet.2024.100808.
56. Kamel, R.M.; Hashem, M.; Ebeed, M.; Alanzi, S.S.; Abd El Hamid, A.M. Stochastic optimal planning and operation of RESs and SMESs in distribution networks for enriching their techno-economic-environmental performance under protective schemes. *Comput. Electr. Eng.* **2026**, *129*, 110817. https://doi.org/10.1016/j.compeleceng.2025.110817.
57. Salman, G.A.; Al-Bahrani, L.T. Enhancement of radial distribution network based on optimal location and sizing of photovoltaic distributed generation using Mountain Gazelle Optimizer and Eel and Grouper Optimizer. *Int. J. Smart Sens. Intell. Syst.* **2025**, *18*. https://doi.org/10.2478/ijssis-2025-0057.
58. Pham, V.H.S.; Dau, T.D.; Nguyen Dang, N.T.; Le, D.A.T.; Tran, L.A. Efficient Resource Leveling in Multi-project Scheduling Environment with an Integrated Mountain Gazelle Optimizer and Opposition-Based Learning. In *Proceedings of the 4th International Conference on Structural Health Monitoring and Engineering Structures (SHM&ES 2025), Nha Trang, Vietnam, 7–8 August 2025*; Cuong, L.T., Fantuzzi, N., Capozucca, R., Quyen, V.T.B., Khatir, S., Eds.; Lecture Notes in Civil Engineering; Springer: Cham, Switzerland, 2026; Volume 747, pp. 257–265. https://doi.org/10.1007/978-3-032-04645-1_31.
59. Bento, M.E.C. Design of a Wide-Area Power System Stabilizer resilient to permanent communication failures using bio-inspired algorithms. *Results Control Optim.* **2023**, *12*, 100258. https://doi.org/10.1016/j.rico.2023.100258.
60. Zhu, X.; Ren, Z.; Liu, X.; Yang, S.; Xu, K.; Zhu, Y. Generation and optimization of optical single-sideband signal using multi-harmonic phase modulation. *Measurement* **2026**, *269*, 120766. https://doi.org/10.1016/j.measurement.2026.120766.
61. Gao, J.; Maalla, A.; Li, X.; Zhou, X.; Lian, K. Comprehensive model for efficient microgrid operation: Addressing uncertainties and economic considerations. *Energy* **2024**, *306*, 132407. https://doi.org/10.1016/j.energy.2024.132407.
62. Turgut, O.E.; Asker, M.; Yesiloz, H.B.; Genceli, H.; AL-Rawi, M. Chaotic Mountain Gazelle Optimizer Improved by Multiple Oppositional-Based Learning Variants for Theoretical Thermal Design Optimization of Heat Exchangers Using Nanofluids. *Biomimetics* **2025**, *10*, 454. https://doi.org/10.3390/biomimetics10070454.
63. Gharehchopogh, F.S.; Rishakan, K.F. Several Improved Models of the Mountain Gazelle Optimizer for Solving Optimization Problems. *Comput. Model. Eng. Sci.* **2026**, *146*, 24. https://doi.org/10.32604/cmes.2025.073808.
64. Salimon, S.A.; Fajinmi, I.O.; Onatoyinbo, O.O.; Alimi, O.A. Allocation of Single and Multiple Multi-Type Distributed Generators in Radial Distribution Network Using Mountain Gazelle Optimizer. *Technologies* **2025**, *13*, 265. https://doi.org/10.3390/technologies13070265.
65. Odelu, P.; Shiva, C.K.; Sen, S.; Vedik, B.; Arya, Y. Hybrid Optimization for Robust Frequency Regulation in Shipboard Microgrids With Renewable Sources, Energy Storage, and Controllable Loads. *Optim. Control Appl. Methods* **2026**, *47*, 312–336. https://doi.org/10.1002/oca.70061.
66. Karkour, Y.; Daoudi, J.; Tajani, C. Using the Mountain Gazelle Optimizer to Construct Consistent Pairwise Comparison Matrices Within Acceptable Consistency Ratio Thresholds. *Math. Model. Comput.* **2026**, *13*, 395–405. https://doi.org/10.23939/mmc2026.02.395.
67. Liu, H. Predicting the influence of yoga on chronic venous insufficiency utilizing the Multi-Layer Perceptron Classifier. *J. Med. Eng. Technol.* **2025**, *49*, 141–163. https://doi.org/10.1080/03091902.2025.2471331.

68. Jameel, M.; Abouhawwash, M. Revolutionizing optimization: An innovative nutcracker optimizer for single and multi-objective problems. *Appl. Soft Comput.* **2024**, *164*, 112019. https://doi.org/10.1016/j.asoc.2024.112019.
69. Agushaka, J.O.; Ezugwu, A.E.; Abualigah, L. Gazelle optimization algorithm: a novel nature-inspired metaheuristic optimizer. *Neural Comput. Appl.* **2023**, *35*, 4099–4131. https://doi.org/10.1007/s00521-022-07854-6.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.