IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 12 July 2026, accepted 29 July 2026, date of publication 3 August 2026, date of current version 11 August 2026.

Digital Object Identifier 10.1109/ACCESS.2026.3719593

APPLIED RESEARCH

A Multi-Strategy Hippopotamus Optimization Algorithm for Dynamic Parameter Identification of Robotic Manipulators

YU ZHANG[1], AZIZAN AS'ARRY[1], HAOHAO MA[1,2], MOHD KHAIROL ANUAR BIN MOHD ARIFFIN[3], AZMAH HANIM BINTI MOHAMED ARIFF[1], AND MOHD NA'IM ABDULLAH[1]

[1]Department of Mechanical and Manufacturing Engineering, Faculty of Engineering, Universiti Putra Malaysia, Serdang, Selangor 43400, Malaysia
[2]Department of Mechanical Engineering, School of Mechanical and Electrical Engineering, Tianshui Normal University, Tianshui, Gansu 741001, China
[3]Department of Aerospace Engineering, Faculty of Engineering, Universiti Putra Malaysia, Serdang, Selangor 43400, Malaysia

Corresponding author: Azizan As'arry (zizan@upm.edu.my)

ABSTRACT The Hippopotamus Optimization (HO) algorithm is a recent swarm-intelligence metaheuristic with three structural limitations: rank-blind updates, a single-magnitude escape operator, and no elite re-evaluation under stochastic fitness. This paper presents the Fitness-Ranked Hippopotamus Optimization (FHO) algorithm, which replaces HO's update pipeline with three rank-driven mechanisms: fitness-contrast directional learning, fitness-rank differential search, and fitness-rank guided mutation, plus a per-iteration elite-preservation step. FHO is validated on CEC2017 at $D = 30/50/100$, CEC2019, and CEC2022, totalling 109 function instances, against 22 competitors including LSHADE, CMA-ES, and four 2025–2026 optimisers, under official budgets with 30 runs. On the 51-function aggregate FHO attains Friedman mean rank 4.83, third of 23 behind LSHADE (1.57) and CMA-ES (3.50), and is the strongest swarm- and social-inspired algorithm in the pool. At $D = 100$ it rises to second (3.97), overtaking CMA-ES, and beats HO on 27–28 of 29 functions across dimensions. A controlled ablation confirms all three mechanisms contribute. Applied to robust tuning of the 20-dimensional fractional-order PID controller of a hardware-identified 4-DOF manipulator under $\pm 20\%$ uncertainty against 11 competitors, FHO sits inside the top statistical clique on unseen plants (Wilcoxon $p \geq 0.50$), with the narrowest score spread and no catastrophic-failure tails within that clique.

INDEX TERMS CEC benchmarks, dynamic parameter identification, fitness-rank selection, fractional-order PID, guided mutation, hippopotamus optimization, metaheuristic optimization, rank-driven differential search, robust control.

I. INTRODUCTION

The nature-inspired population-based family has expanded quickly over the last decade. Particle swarm optimisation (PSO) [1], grey wolf optimiser (GWO) [2], teaching-learning-based optimisation (TLBO) [3], plus a steady stream of newer entries — whale optimisation algorithm (WOA) [4], salp swarm algorithm (SSA) [5], Harris hawks optimisation (HHO) [6], aquila optimiser (AO) [7], equilibrium optimiser (EO) [8], coyote optimisation algorithm (COA) [9], dung beetle optimiser (DBO) [10], RIME [11], snow ablation optimiser (SAO) [12], red deer algorithm (RDO) [13], and crested porcupine optimiser (CCO) [14] — now feature routinely in engineering, control, and machine-learning papers on bound-constrained continuous optimisation. Each algorithm brands its update rule with a different biological or physical analogy. Strip that away, though, and the mathematical skeleton is shared: at every iteration,

The associate editor coordinating the review of this manuscript and approving it for publication was Qiang Li.

a population of N candidate solutions is updated through operators that blend random perturbation with information aggregated from the current population and from the best-so-far archive.

Amiri et al. [15] introduced the Hippopotamus Optimization (HO) algorithm in 2024, and it has since attracted a growing body of application and enhancement work, including the multi-strategy HO variant (MSHO) of Yuan et al. [16]. Yet on the problem class considered in this study HO shows three structural weaknesses, developed in detail in Section III-A: (i) *rank-blindness*, whereby the same update rule is applied to every individual regardless of its fitness rank, wasting evaluations on unpromising candidates; (ii) a *single-magnitude escape operator* that cannot dislodge individuals from narrow multi-modal traps; and (iii) *no elite re-evaluation* under stochastic fitness, so a single lucky evaluation can be locked in as the apparent optimum. These weaknesses are most damaging on the problems where metaheuristics are most often used in engineering: rugged, noisy landscapes such as robust controller tuning on an uncertain plant.

This observation motivates the central research question of the paper: can the within-iteration fitness ranking, information that is computed and then discarded by HO and by most swarm optimisers, be exploited systematically to repair all three weaknesses at once, without increasing the per-iteration evaluation cost? Isolated fixes exist in the literature, such as adaptive inertia weights [17], DE hybrids [18], and Gaussian mutation [19], but always as post-hoc patches to a single operator. The approach taken here is different: the Fitness-Ranked Hippopotamus Optimization (FHO) restructures HO around a rank-modulated update pipeline built from three mechanisms: a fitness-contrast directional-learning operator (FCDL) that intensifies half the population toward fitter peers with a rank-scaled contrast, a fitness-rank differential-search operator (FRDS) whose step size adapts to the rank gap between donors, and a fitness-rank guided-mutation operator (FRGM) whose per-dimension perturbation probability rises with rank and falls with population diversity. A one-evaluation elite-preservation step closes each iteration. The total cost is $2.5N + 1$ evaluations per iteration, less than HO's $3N$.

The study has two aims: first, to establish the general-purpose performance of FHO under the community's official evaluation protocol; second, to demonstrate that the improvement transfers to a realistic engineering task, namely the dynamic parameter identification of a four-degree-of-freedom (4-DOF) robotic manipulator and robust tuning of its fractional-order PID (FOPID) controller. All benchmarks use the suites' official evaluation budgets (up to 10^6 evaluations per run) [20], [21], dimensions up to $D = 100$, 30 independent runs, and a 23-algorithm pool that includes the adaptive-DE representative LSHADE [22], CMA-ES [23], the literature HO variant MSHO [16], and four optimisers published in 2025–2026 (SFOA [24], DhO [25], OOA [26], DRA [27]).

Under this protocol, FHO ranks *third* of 23 on the 51-function cross-suite aggregate (Friedman 4.83), behind the two state-of-the-art non-swarm optimisers LSHADE (1.57) and CMA-ES (3.50), both run at their recommended dimension-scaled population schemes; it is, however, the strongest swarm- and social-inspired algorithm in the pool, and it outranks both Hippopotamus-family algorithms benchmarked here—the parent HO (15.27) and the published multi-strategy variant MSHO (8.06)—by wide margins. At $D = 100$ FHO improves to *second* (3.97), overtaking CMA-ES (4.83) but remaining behind LSHADE (1.21), and it beats the parent HO on 27–28 of 29 functions per dimension. On the robust-FOPID task FHO sits inside the top statistical clique (with TLBO, DhO, LSHADE, and GWO; $p \geq 0.50$) and is statistically indistinguishable even from the rank-leader TLBO; among the high-variance clique members (LSHADE, DhO, GWO) FHO has the lowest variance, although TLBO attains both a better rank and a lower variance, and it improves on its parent HO, though not significantly ($p = 0.072$).

This paper makes four contributions. First, a multi-strategy FHO algorithm is developed whose three rank-driven mechanisms address the rank-blindness, diversity-collapse, and noise-vulnerability of HO; all twelve hyperparameters are named and tabulated, and the five behaviour-shaping parameters are subjected to a $\pm 40\%$ sensitivity analysis. Second, a controlled ablation against the parent HO quantifies each mechanism's contribution under the official budget. Third, the algorithm's behaviour is characterised beyond rank tables: exploration–exploitation balance, population-diversity dynamics, qualitative search trajectories, runtime cost, and high-dimensional scaling to $D = 100$. Fourth, FHO is applied to the experimentally-identified dynamics of a 4-DOF manipulator and the robust tuning of its 20-dimensional FOPID controller, with a Lyapunov-based argument that the optimisation proceeds entirely within a stable-by-construction parameter set, and a time-domain validation of the resulting controllers on unseen perturbed plants. Beyond manipulators, the same identify-then-tune pipeline applies to other complex dynamic systems; motor parameter estimation [28] and general robotic model parameterisation [29] are recent examples of the problem class.

The rest of the paper proceeds as follows. Section II surveys related work on nature-inspired metaheuristics, HO enhancements, diversity preservation, and CEC evaluation protocols. Section III develops the FHO algorithm, its pseudocode, flowchart, and computational complexity; Section III-D formulates the FOPID problem and Section III-E establishes the stability condition. Section IV-A evaluates FHO on the five CEC configurations against 22 competitors; Section IV-B reports the ablation, sensitivity, and diversity studies. Sections IV-C–IV-D cover the manipulator application. Section V discusses the results, limitations, and future work.

TABLE 1. Algorithms compared with FHO in this study (23-algorithm pool).

Algorithm	Year / Ref.	Family	Note
PSO	1995 [1]	Swarm	classical baseline
CMA-ES	2001 [23]	Evolution strategy	covariance adaptation
TLBO	2011 [3]	Social	parameter-free
GWO	2014 [2]	Swarm	
LSHADE	2014 [22]	Adaptive DE	CEC2014 winner; SHADE/JADE lineage
WOA	2016 [4]	Swarm	
SSA	2017 [5]	Swarm	
COA	2018 [9]	Swarm	
HHO	2019 [6]	Swarm	
EO	2020 [8]	Physics	
RDO	2020 [13]	Swarm	
AO	2021 [7]	Swarm	
DBO	2023 [10]	Swarm	
RIME	2023 [11]	Physics	
SAO	2023 [12]	Physics	
CCO	2024 [14]	Swarm	
HO	2024 [15]	Swarm	parent algorithm
SFOA	2025 [24]	Swarm	recent-generation optimiser
DhO	2025 [25]	Swarm	recent-generation optimiser
DRA	2025 [27]	Social	recent-generation optimiser
MSHO	2025 [16]	HO variant	multi-strategy HO from literature
OOA	2026 [26]	Swarm	recent-generation optimiser
FHO	2026	HO variant	

II. LITERATURE REVIEW

A. NATURE-INSPIRED OPTIMISERS AND THE COMPARISON POOL

Since Kennedy and Eberhart's PSO [1], the nature-inspired metaheuristic family has grown rapidly. Table 1 summarises the 22 algorithms (plus FHO) used as comparators in this work, ordered by year of publication. Their metaphors run from swarm and evolutionary to physics-inspired and socially-inspired. The mathematical substrate underneath, however, is the same: iterated population update on a bound-constrained continuous decision space.

Three groups in the pool deserve comment. First, the adaptive differential-evolution lineage: JADE [30] introduced adaptive control of F and CR with an optional external archive, SHADE [31] added success-history adaptation, and LSHADE [22], the CEC2014 winner, added linear population-size reduction. LSHADE is included here as the strongest and most widely-benchmarked representative of that lineage, together with the covariance-matrix-adaptation evolution strategy CMA-ES [23]; between them they cover the two non-swarm families that dominate modern CEC competitions. Second, four optimisers published in 2025–2026, namely the starfish optimization algorithm (SFOA) [24], the dhole optimization algorithm (DhO) [25], the octopus optimization algorithm (OOA) [26], and the divine religions algorithm (DRA) [27], represent the current generation of nature- and society-inspired proposals. Third, MSHO [16], a published multi-strategy HO enhancement, provides a literature baseline for HO improvement: any new HO variant should beat not only the parent but also existing variants.

Table 1 also shows a shared design choice: despite the rapid proliferation of named algorithms, nearly all of them apply the same update rule uniformly to every individual in the population. Implementation is simpler that way, but the design forfeits a useful source of information. The within-iteration rank of each individual is discarded, even though it could, in principle, inform a more targeted update strategy. The FHO algorithm in Section III takes the opposite route: it uses the within-iteration fitness rank to modulate every update operator, scaling step magnitudes, mutation gains, and per-dimension perturbation probabilities by rank. Rank information is not free (it requires an $O(N \log N)$ sort per iteration), but it is far cheaper than a function evaluation, and Section IV quantifies the return on that investment.

Metaheuristic optimisation is also an established tool for the parameter identification of complex dynamic systems, the application domain of this paper; the monograph of Cuevas et al. [64] surveys the breadth of this literature. Rodríguez-Abreo et al. [28] estimate motor-model parameters from step-response relations with GWO, Jaya, and cuckoo-search, and Hsueh and Fazdi [65] compare particle-swarm variants for permanent-magnet DC-motor estimation; bio-inspired optimisation has likewise been used for whole-robot model parameterisation [29]. The same identify-by-optimisation principle extends beyond electromechanical plants to industrial thermal processes, where metaheuristics have been used to fit heating–cooling [66] and steam-chamber [67] process models. These studies share the structure adopted here, fitting a physics-grounded model by minimising a simulation-to-measurement residual, but typically tune a single algorithm on a single plant; the present work instead benchmarks 12 algorithms under a common budget on a stochastic robust-tuning objective.

B. DIVERSITY PRESERVATION UNDER MULTI-MODAL AND NOISY FITNESS

On rugged multi-modal landscapes, and under noisy fitness evaluation, population-based metaheuristics tend to fail in one characteristic way — premature diversity collapse [62]. Once the population has converged to a narrow region of the search space, nothing in the classical swarm or evolutionary toolkit can pull it back out; every update operator generates new candidates inside the neighbourhood of the current population.

Two families of remedies have been put forward in the literature, each independently. Rank-dependent mutation [63] hands larger perturbations to poor-performing individuals. Opposition-based learning [34] generates diverse candidates via axis-aligned reflection. FHO's contribution to this line is to make rank the single organising signal — FCDL's contrast-scaled intensification of the worst half, FRDS's rank-gap-scaled differential steps, and FRGM's mutation probability that rises with rank and falls with measured per-dimension diversity — so that diversity injection concentrates exactly where and when diversity is low.

C. CEC BENCHMARKS AND NON-PARAMETRIC EVALUATION

The Congress on Evolutionary Computation has put out an annual series of single-objective bound-constrained

benchmark suites (CEC2013, CEC2014, CEC2017, CEC2019, CEC2020, CEC2022) that are now the de-facto standard for evaluating new metaheuristics [20], [21], [22], [23], [24], [25], [26], [27], [28], [29], [30], [31], [32], [33], [34], [35]. Each suite includes a curated set of test functions spanning four landscape categories: unimodal (convergence-rate evaluation), simple multimodal (global-search capability), hybrid (discontinuity and rotation), and composition (extreme non-convexity). Reporting practice is well established: thirty independent runs per function; mean ± standard deviation of the final fitness; and a Friedman rank across all functions of a suite, with Wilcoxon rank-sum pairwise significance on top [36], [37]. The official protocols also prescribe the evaluation budget (MaxFEs $= 10^4 \times D$ for the CEC2017-style suites [20]), and results obtained at lower budgets are not comparable with the published literature; this paper adopts each suite's official budget (the $10^4 \times D$ rule for CEC2017 [20] and 10^6 evaluations for CEC2022 at $D = 20$ [21]), and, since the CEC2019 100-digit competition prescribes no evaluation cap of its own, applies the $10^4 \times D$ rule there for cross-suite consistency. The CEC2019 suite, known as the 100-Digit Challenge [35], deserves a separate note. Its ten functions are deliberately crafted to reward extreme precision, and its functions carry heterogeneous dimensionalities ($D = 9$–18).

Statistical testing in this paper follows the standard non-parametric methodology [36], [37], [38], [39], [40]: the Friedman rank test for aggregate performance; Wilcoxon rank-sum pairwise tests with win/tie/loss counts, the Nemenyi post-hoc test with critical-distance diagrams [37], [38], Cliff's delta effect size [39], [41] to report magnitude at the modest sample sizes (20–30 runs) that metaheuristic benchmarks typically use, and 95% t-confidence intervals on per-function means. Throughout the paper Cliff's δ is reported *signed*, computed on the paired comparison (competitor versus FHO): negative values favour FHO, positive values favour the competitor. For the per-competitor FOPID comparisons of Table 15, the Wilcoxon p-values are additionally Holm-corrected across the eleven-competitor family, and the significance markers in that table denote the Holm-adjusted outcome, so that the reported "significantly outperforms" set controls the family-wise error rate rather than the per-comparison rate.

FHO is validated on a 20-dimensional robust FOPID-tuning problem for a 4-DOF manipulator (Sections IV-C–IV-D), chosen because its landscape combines multi-modality, stochastic fitness from plant-parameter uncertainty, and coupled nonlinear dynamics [42] — exactly the conditions FHO was designed to handle.

III. METHODOLOGY

A nomenclature table covering the symbols used throughout the paper is provided in Table 2.

A. THE HIPPOPOTAMUS OPTIMIZATION ALGORITHM

The Hippopotamus Optimization algorithm of Amiri et al. [15] structures population updates into three sequential phases, each inspired by the social and defensive behaviour of hippopotamus herds. Consider a bound-constrained minimisation problem with population $X = \{x_1, \ldots, x_N\} \subset [lb, ub]^D$ and fitness function f. At each iteration t (of T), the first half of the population passes through a two-step herd-positioning update,

$$x_i^{P1} = x_i + r\left(x^* - I_1\, x_i\right), \tag{1}$$

$$x_i^{P2} = \begin{cases} x_i + A \odot \left(x^* - I_2\, \bar{x}_G\right), & T_e > 0.6, \\ x_i + B \odot \left(\bar{x}_G - x^*\right) \text{ or re-sample}, & \text{otherwise}, \end{cases} \tag{2}$$

where x^* is the current best ("dominant") hippopotamus, $\bar{x}_G$ is the mean of a random sub-group, $I_1, I_2 \in \{1, 2\}$ are random integers, A and B are random scaling vectors drawn from a five-candidate set, $r \sim \mathcal{U}(0, 1)$, and $T_e = \exp(-t/T)$. The second half undergoes a predator-defence update against a uniformly-drawn putative predator x_p,

$$x_i^{P3} = \mathrm{RL} \odot x_p + \frac{b}{c - d\cos l} \cdot \frac{1}{\Delta_i}, \tag{3}$$

with Δ_i the (component-wise) distance to the predator (doubled and noise-shifted when the predator is fitter), $b \sim \mathcal{U}(2, 4)$, $c \sim \mathcal{U}(1, 1.5)$, $d \sim \mathcal{U}(2, 3)$, $l \sim \mathcal{U}(-2\pi, 2\pi)$, and RL a Lévy-distributed step of scale 0.05. Finally every individual attempts a contracting escape move

$$x_i^{P4} = x_i + r\left(\frac{lb}{t} + E \odot \left(\frac{ub}{t} - \frac{lb}{t}\right)\right), \tag{4}$$

with E drawn from a three-candidate random set. All moves are bound-clipped and accepted greedily. Each iteration consumes $3N$ function evaluations (two per first-half individual, two per second-half individual including the predator evaluation, and one escape evaluation per individual).

HO is competitive on standard benchmark suites, but on the kinds of problem studied here it shows three weaknesses. First, the herd-positioning and defence updates apply the same perturbation rules to every individual. Rank information is available once the population has been sorted by fitness, yet it is ignored; top-ranked individuals, already near promising regions, receive the same aggressive update as bottom-ranked ones, which would be more productively redirected

TABLE 2. Nomenclature.

Symbol	Meaning	Symbol	Meaning
N	population size	$q, \dot{q}, \ddot{q}$	joint position/velocity/acceleration
D	problem dimension	τ	joint torque vector
t, T	iteration index / total iterations	$M(q)$	generalised mass matrix
x_i, x^*	individual i / incumbent best	$C(q, \dot{q})$	Coriolis–centripetal matrix
r_i	fitness rank of individual i	$G(q)$	gravity vector
$f(\cdot)$	objective (fitness) function	$F(\dot{q})$	joint-friction vector
lb, ub	lower / upper bounds	B, F_c, F_s	viscous / Coulomb / static friction
MaxFEs	function-evaluation budget	ω_s	Stribeck velocity
c_i	FCDL fitness contrast	p_1, p_2, p_3	lumped gravity parameters
s	FRDS rank-adaptive step	s_j	motor-sign coefficients
V	FRGM donor vector	K_p, K_i, K_d	FOPID gains
p_j	FRGM perturbation probability	λ, μ	fractional integral / derivative orders
div_j	per-dimension diversity	D^α	fractional differintegral
ξ	standard normal vector	$e(t)$	tracking error
u	uniform random vector in $[0, 1]^D$	θ, Θ	plant parameters / uncertainty set
ρ	uniform random scalar	K	Monte-Carlo plant samples per evaluation
$\beta, c_{\min}, w_1, w_2$	FCDL/FRDS gains (Table 3)	$J(\cdot)$	robust FOPID cost
S_0, F, δ	differential scales, decay (Table 3)	ITAE	integral time-weighted absolute error
$p_0, p_1^{\mathrm{FRGM}}, p_{\min}, \sigma_e$	FRGM/elite (Table 3)	OS, ST, SSE	overshoot, settling time, steady-state error
I_1, I_2	random integers in $\{1, 2\}$ (HO update)	q^d	desired joint trajectory
I_4	4×4 identity matrix	$d(t)$, $d_{\max}$	lumped disturbance and its bound

toward promising regions. Second, the escape phase of (4) contracts as $1/t$ with a single random magnitude, providing neither rank-targeted exploration nor an escape energy large enough to dislodge individuals trapped in narrow multi-modal optima. Third, the best-so-far individual is never re-evaluated between iterations. Under stochastic fitness this is costly: a single lucky evaluation can be locked in as the apparent optimum, and the algorithm then spends subsequent iterations refining around a spurious target.

These three weaknesses motivate the FHO algorithm developed below. FHO keeps HO's population representation and bound-handling intact, but swaps out the three update phases for three rank-driven mechanisms (FCDL, FRDS, FRGM) and closes each iteration with a one-evaluation elite-preservation step.

B. THE PROPOSED FHO ALGORITHM

The population is initialised by uniform random sampling,

$$x_{i,j} = lb_j + u_{i,j}\,(ub_j - lb_j), \qquad u_{i,j} \sim \mathcal{U}(0,1), \tag{5}$$

costing N evaluations. Uniform sampling was adopted after a controlled comparison against logistic-, tent-, and Chebyshev-map chaotic (plus quasi-opposition) initialisation on CEC2017 $D = 30$ at the official budget with 30 runs: the four schemes proved statistically indistinguishable (Friedman spread below 0.5, far inside the Nemenyi critical distance), so the simpler uniform scheme was retained. Every iteration then opens with a sort of the population by fitness, which assigns each individual a rank $r_i \in \{1, \ldots, N\}$ ($r = 1$ best). The rank vector is the single signal that drives all three mechanisms below; ratio $= t/T$ denotes normalised time.

Mechanism 1 — FCDL. FCDL is applied to half of the population (the lower half by stored order). For each such individual x_i, two distinct peers are sampled and the fitter one is taken as reference x_{ref}. The signed *fitness contrast* between x_i and the reference,

$$c_i = \tanh\!\Big(\beta\,\frac{r_i - r_{\text{ref}}}{N}\Big), \qquad \tilde{c}_i = \text{sign}(c_i)\,\max\big(|c_i|, c_{\min}\big), \tag{6}$$

measures how much worse x_i ranks than its reference: the tanh saturates the contrast for very large rank gaps (preventing overshoot), β sets the slope, and the floor $c_{\min}$ guarantees a minimum step even between near-equal ranks. The trial position

$$y = x_i + \tilde{c}_i\,\big(u \odot (x_{\text{ref}} - x_i)\big) + w_1\,\text{ratio}\;\rho\,\big(x^* - x_i\big) \tag{7}$$

moves x_i toward its fitter reference in proportion to the contrast, plus a time-growing pull toward the incumbent best x^* ($u \in [0,1]^D$ uniform vector, $\rho \in [0,1]$ uniform scalar, w_1 the elite-attraction weight). The trial is bound-clipped, evaluated, and accepted greedily. FCDL costs $N/2$ evaluations and is a rank-conditional *intensification* operator: it does not inject fresh diversity (that is FRGM's role) but accelerates these individuals toward regions already validated by fitter peers, with a step size that adapts—through the rank-gap contrast c_i—to how far each individual ranks behind its reference.

Mechanism 2 — FRDS. FRDS applies to the whole population. For each x_i, two distinct peers are sampled and ordered into a better/worse pair (x_b, x_w); their normalised rank gap $g = |r_w - r_b|/N$ scales the differential step

$$s = \big(S_0 + F\,g\big)\,\big(1 - \delta\,\text{ratio}\big), \tag{8}$$

$$y = x_i + s\,\big(u \odot (x_b - x_w)\big) + w_2\,\rho\,\big(x^* - x_i\big), \tag{9}$$

followed by clipping and greedy acceptance (N evaluations). Two properties distinguish FRDS from classical DE [18]. The step length is *rank-adaptive*: a large rank gap between the donors signals a direction along which fitness changes rapidly, and earns a longer step $S_0 + Fg$. And the whole step decays with time through the factor $(1 - \delta\,\text{ratio})$, shifting FRDS smoothly from exploration toward exploitation as the budget is consumed. This adaptive schedule addresses the premature-convergence concern raised for fixed-weight differential operators.

Mechanism 3 — FRGM. FRGM supplies the escape energy that the directional and differential operators cannot. It first measures the per-dimension population diversity

$$\text{div}_j = \frac{\text{std}_k\,(x_{k,j})}{ub_j - lb_j}, \tag{10}$$

then, for each individual, samples three distinct peers ordered by fitness into $(x_{s_1}, x_{s_2}, x_{s_3})$ (best/middle/worst) and builds the donor

$$V = \frac{2x_{s_1} + x_{s_2}}{3} + F_{\text{rank}}\,(x_{s_1} - x_{s_3}),$$

$$F_{\text{rank}} = \big(S_0 + F\,\tfrac{r_i}{N}\big)\big(1 - \delta\,\text{ratio}\big), \tag{11}$$

an elite-weighted base plus a best-to-worst differential whose gain F_{rank} grows with the rank of the individual being updated — poorly-ranked individuals receive more aggressive donors — and decays over time. Crossover is dimension-wise with probability

$$p_j = \max\big((1 - \text{div}_j)\,(p_0 + p_1^{\text{FRGM}}\,\tfrac{r_i}{N}),\; p_{\min}\big), \tag{12}$$

so that dimensions whose population spread has collapsed ($\text{div}_j \to 0$) are perturbed with high probability, while still-diverse dimensions are left largely intact; at least one dimension is always replaced. The trial inherits the donor on the selected dimensions, is clipped, evaluated, and accepted greedily (N evaluations). FRGM is therefore a *diversity-targeted* escape operator: the ablation of Section IV-B shows it is the single most valuable mechanism (+2.07 Friedman rank when removed).

Elite preservation. After the three mechanisms, the incumbent best x^* is updated, and the current *worst* slot is re-seeded in a tight Gaussian neighbourhood of the best,

$$x_{\text{worst}} \leftarrow x^* + \sigma_e\,\xi \odot (ub - lb), \qquad \xi \sim \mathcal{N}(0, I_D), \tag{13}$$

evaluated once, and accepted greedily (≤ 1 evaluation; I_D denotes the $D \times D$ identity). On deterministic objectives this provides a cheap local probe around the incumbent;

TABLE 3. FHO hyperparameters and default settings.

Symbol	Meaning	Mechanism	Default
N	Population size	—	40
β	Fitness-contrast gain (tanh slope)	FCDL	2.0
$c_{\min}$	Minimum contrast magnitude	FCDL	0.15
w_1	Elite-attraction weight	FCDL	0.3
w_2	Elite-attraction weight	FRDS	0.1
S_0	Differential base scale	FRDS/FRGM	0.3
F	Rank-dependent differential gain	FRDS/FRGM	0.5
δ	Time-decay coefficient	FRDS/FRGM	0.3
p_0	Base perturbation probability	FRGM	0.4
p_1^{FRGM}	Rank-dependent perturbation gain	FRGM	0.4
$p_{\min}$	Minimum perturbation probability	FRGM	0.1
σ_e	Elite worst-slot reseed scale (rel. to $ub - lb$)	Elite	0.001

on stochastic objectives — such as the robust-FOPID problem, where each fitness call averages K random plant perturbations — it doubles as a fresh low-variance sample of the elite region, countering the noise-driven ''best-ever drift'' that destabilises algorithms lacking elite re-evaluation.

All twelve hyperparameters introduced above are collected in Table 3 with their default values. The defaults were fixed once on a small calibration set and used unchanged for every experiment in this paper; Section IV-B2 perturbs the five behaviour-shaping parameters (β, F, δ, p_0, σ_e) by $\pm 40\%$ and finds the cross-function Friedman rank essentially flat, indicating that FHO does not owe its performance to fine-tuned constants.

C. ALGORITHM ARCHITECTURE, COMPLEXITY, AND PSEUDOCODE

Algorithm 1 specifies the complete FHO procedure in pseudocode form; the flowchart in Fig. 1 shows the full operational workflow, including the per-phase evaluation costs. The per-iteration function-evaluation cost is

$$C_{\mathrm{iter}} = \underbrace{N/2}_{\mathrm{FCDL}} + \underbrace{N}_{\mathrm{FRDS}} + \underbrace{N}_{\mathrm{FRGM}} + \underbrace{1}_{\mathrm{elite}} = 2.5N + 1, \qquad (14)$$

appreciably less than HO's $3N$. Given a total budget of MaxFEs evaluations and the N-evaluation initialisation, the FHO iteration count is

$$T_{\mathrm{FHO}} = \left\lfloor \frac{\mathrm{MaxFEs} - N}{2.5N + 1} \right\rfloor. \qquad (15)$$

Computational complexity. Each iteration performs one fitness sort, $O(N \log N)$, and $2.5N+1$ candidate constructions and evaluations, each $O(D)$ plus the cost c_f of one objective call; FRGM additionally computes the per-dimension standard deviation in $O(ND)$. The total time complexity is therefore $O\big(T_{\mathrm{FHO}}\,(N \log N + ND + N\,c_f)\big)$, identical in order to HO and standard DE whenever the objective call dominates (the practical case), and the space complexity is $O(ND)$. The constant factor is, however, larger than that of single-operator swarms: the measured wall-clock cost is quantified in Section IV-A4 (Table 9) and acknowledged as a limitation in Section V-B.

Algorithm 1 Fitness-Ranked Hippopotamus Optimization (FHO)

Input: population N, dimension D, bounds lb, ub, MaxFEs; hyperparameters (Table 3)
Initialise $X = \{x_1, \ldots, x_N\}$ uniformly at random in $[lb, ub]^D$; evaluate; set incumbent x^*, f_{best} ▷ N FEs
for $t = 1$**to**T_{FHO} **do**
  ratio $\leftarrow t/T_{\mathrm{FHO}}$; sort X by fitness to assign ranks $r_i \in \{1, \ldots, N\}$ ($r = 1$ best)
  ## FCDL — fitness-contrast directional learning (bottom half) ▷ $N/2$ FEs
  for $i = \lceil N/2 \rceil + 1$**to**$N$ **do**
    ref $\leftarrow$ fitter of two random peers; $\tilde{c}_i \leftarrow \mathrm{sign}(c_i)\max(|c_i|, c_{\min})$ with $c_i = \tanh\big(\beta \frac{r_i - r_{\mathrm{ref}}}{N}\big)$
    $y \leftarrow x_i + \tilde{c}_i\,(u \odot (x_{\mathrm{ref}} - x_i)) + w_1\,\mathrm{ratio}\,\rho\,(x^* - x_i)$; clip; **if** $f(y) < f(x_i)$**then** $x_i \leftarrow y$
  end for
  ## FRDS — fitness-rank differential step (all N) ▷ N FEs
  for $i = 1$**to**N **do**
    $(b, w) \leftarrow$ better/worse of two random peers; $g \leftarrow |r_w - r_b|/N$; $s \leftarrow (S_0 + Fg)(1 - \delta\,\mathrm{ratio})$
    $y \leftarrow x_i + s\,(u \odot (x_b - x_w)) + w_2\,\rho\,(x^* - x_i)$; clip; **if** $f(y) < f(x_i)$**then** $x_i \leftarrow y$
  end for
  ## FRGM — fitness-rank guided mutation (all N) ▷ N FEs
  $\mathrm{div}_j \leftarrow \mathrm{std}_k(x_{k,j})/(ub_j - lb_j)$
  for $i = 1$**to**N **do**
    $(s_1, s_2, s_3) \leftarrow$ best/mid/worst of three random peers; $V \leftarrow \frac{2x_{s_1} + x_{s_2}}{3} + (S_0 + F\frac{r_i}{N})(1 - \delta\,\mathrm{ratio})(x_{s_1} - x_{s_3})$
    $p_j \leftarrow \max\big((1 - \mathrm{div}_j)(p_0 + p_1^{\mathrm{FRGM}}\frac{r_i}{N}), p_{\min}\big)$; $m_j \leftarrow [\mathcal{U}(0,1) < p_j]$ (force ≥ 1)
    $y \leftarrow x_i$, $y_m \leftarrow V_m$; clip; **if** $f(y) < f(x_i)$**then** $x_i \leftarrow y$
  end for
  ## Elite preservation ▷ 1 FE
  Update x^*, f_{best}; reseed worst slot $\leftarrow x^* + \sigma_e\,\xi \odot (ub - lb)$, $\xi \sim \mathcal{N}(0, I_D)$; accept greedily
end for
return x^*, f_{best}

D. FRACTIONAL-ORDER PID FORMULATION

The fractional-order PID controller of Podlubny [42] generalises the classical proportional-integral-derivative (PID) by replacing the integer-order integral and derivative operators with fractional-order counterparts, of order $\lambda \in (0, 1]$ and $\mu \in (0, 1]$. The control law is

$$u(t) = K_p\,e(t) + K_i\,D^{-\lambda}e(t) + K_d\,D^{\mu}e(t), \qquad (16)$$

Here D^{α} denotes the fractional differintegral of order α — realised, in this paper, by the Grünwald–Letnikov discretisation [43], [44] with memory length $L_{\mathrm{GL}} = 100$ samples — and $e(t) = r(t) - q(t)$ is the tracking error.

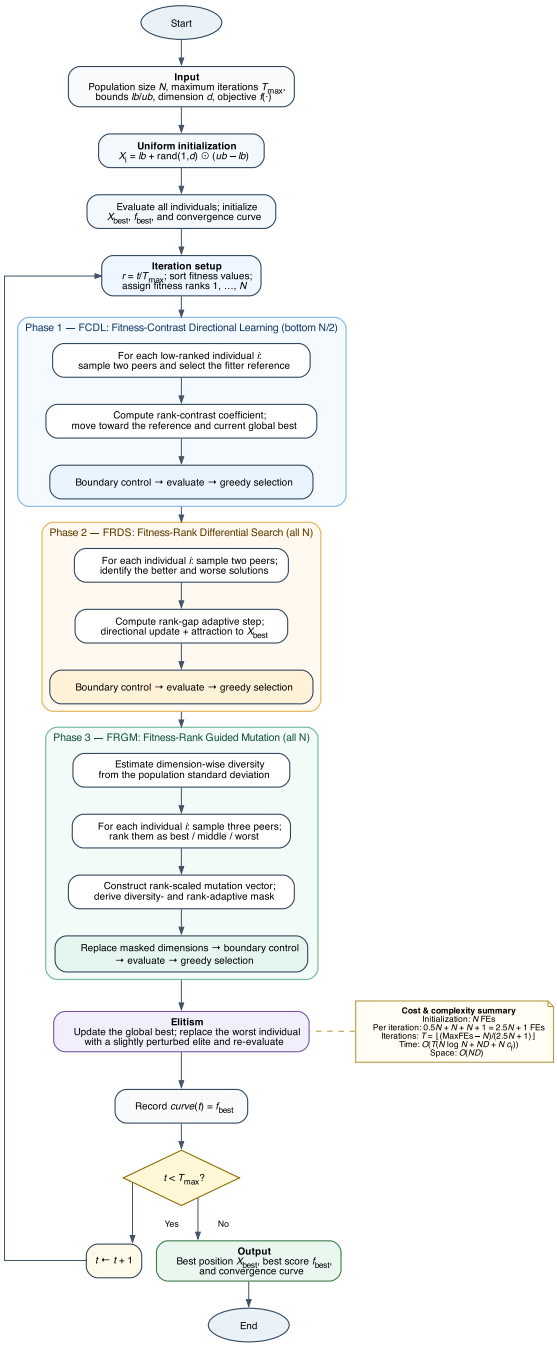

FIGURE 1. FHO flowchart with per-phase function-evaluation costs, iteration formula, and complexity summary.

Setting $\lambda = \mu = 1$ recovers the classical PID structure [45], [46]. Varying the two fractional orders independently in (0, 1] opens up a continuous spectrum of alternative tunings, which can be shaped toward iso-damping, flat-phase, or noise-rejection objectives [42], [47].

Tuning a multi-joint FOPID controller pulls together three independent sources of optimisation difficulty, and between them these exercise the design assumptions of FHO directly. The $5n$ parameter space — five per-joint FOPID coefficients (K_p, K_i, K_d, λ, μ) times n joints — places the problem in a regime where the rank-weighted differential search of Section III-B propagates good information efficiently across many dimensions; in exactly this setting the fitness-blind updates of HO are particularly wasteful.

TABLE 4. FOPID parameter search bounds (identical across all joints).

Parameter	Lower bound	Upper bound
K_p	0.01	1.00
K_i	0.05	1.50
K_d	0.005	0.50
λ	0.30	1.00
μ	0.30	1.00

The non-convex landscape produced by coupled nonlinear arm dynamics, Stribeck friction, communication latency, and motor-bandwidth limits is rugged enough to reward FRGM's diversity-targeted mutation, which supplies escape energy from multi-modal traps. Plant uncertainty, once added, changes the picture again. This paper models it — in keeping with industrial identification practice [48] — as ±20% multiplicative noise on inertia, viscous friction, and Coulomb friction parameters, with $K = 4$ plant samples drawn for every fitness call. Under this noise model, the FCDL intensification of the bottom half and the elite re-evaluation become essential: without them the population collapses onto a spuriously-lucky candidate. The detailed dynamics model, the five-stage hardware identification, and the benchmark protocol appear in Section IV-C; simulation results on the identified plant model are in Section IV-D. On this engineering task FHO finishes in a statistical tie with the strongest competitors while improving on its parent HO (Section IV-D2).

Table 4 lists the per-joint search bounds used by every optimiser in the benchmark. The proportional and derivative gains span two decades, while the integral gain spans a factor of 30, and the two fractional orders are restricted to [0.3, 1.0] to avoid degenerate low-order behaviour while still covering the practically-useful FOPID range. All bounds are box-like and identical across joints for comparability; a per-joint tightening based on pre-analytic stability considerations would likely accelerate convergence, but is avoided in this benchmark to ensure that no algorithm receives a problem-specific advantage.

E. CLOSED-LOOP STABILITY UNDER FHO-TUNED GAINS

A natural concern for any optimisation-driven controller synthesis is whether every FOPID candidate returned by FHO is guaranteed to preserve closed-loop stability of the manipulator. This subsection supplies a Lyapunov-based sufficient condition on the FOPID gains and shows that the search bounds of Table 4 lie strictly inside that condition. Consequently every tuning evaluated during the optimisation is a priori closed-loop stable on the nominal plant, and robust closed-loop performance under the ±20% plant perturbation

of Section IV-D is confirmed empirically by the Monte Carlo rollouts.

Let $e(t) = q^{\mathrm{d}}(t) - q(t)$ be the per-joint tracking-error vector and let D^{α} denote the Caputo fractional derivative of order $\alpha \in (0, 1]$. Combining the manipulator dynamics (17) (defined in Section IV-C) with the FOPID control law (16) yields the closed-loop error equation

$$M(q)\ddot{e} + \left[C(q, \dot{q}) + K_d D^{\mu-1}\right]\dot{e} + K_p\, e + K_i D^{-\lambda} e = d(t) \tag{E1}$$

where $d(t)$ lumps the gravity residual, friction, and feed-forward terms together. By the standard manipulator Properties 1–2 [49], $d(t)$ is uniformly bounded on any compact workspace: $\|d(t)\|_{\infty} \le d_{\max}$.

As the Lyapunov candidate, the following composite form is adopted [47]:

$$V(e, \dot{e}, q) = \tfrac{1}{2}\dot{e}^{\mathrm{T}} M(q)\dot{e} + \tfrac{1}{2} e^{\mathrm{T}} K_p\, e \tag{E2}$$

which is positive definite on the manipulator workspace because $M(q) \succeq m_1 I_4$ by Property 1 [49] (here and below I_4 denotes the 4×4 identity matrix and $m_1 > 0$ the uniform lower bound on the mass matrix) and $K_p = \mathrm{diag}(K_{p,j}) \succ 0$ whenever every $K_{p,j} > 0$.

Lemma 1 (Fractional Lyapunov Inequality [50]): For any continuously differentiable $x(t) \in \mathbb{R}^n$, any symmetric positive-definite P, and any $\alpha \in (0, 1]$,

$$\tfrac{1}{2} D^{\alpha}(x^{\mathrm{T}} P x) \le x^{\mathrm{T}} P \cdot D^{\alpha} x. \tag{E3}$$

Proposition 1: (Mittag–Leffler uniform ultimate boundedness (UUB) under FOPID control; this paper, building on the standard manipulator properties [49] and Lemma 1): If every joint of the FOPID controller satisfies

$$K_{p,j} > 0, \quad K_{d,j} > 0, \quad K_{i,j} > 0, \quad \lambda_j, \mu_j \in (0, 1], \tag{E4}$$

then there exist $\alpha \in (0, 1]$ and constants $\gamma > 0$, $\varphi > 0$ (depending on $d_{\max}$ and the controller gains) such that the Lyapunov candidate V of (E2) obeys

$$D^{\alpha} V \le -\gamma V + \varphi. \tag{E5}$$

Consequently [50], the tracking error $e(t)$ is UUB within a ball of radius $B_{\infty} = \sqrt{2\varphi/(\gamma\, \lambda_{\min}(K_p))}$, which vanishes as $d_{\max} \to 0$.

Proof sketch: Applying Lemma 1 to the two quadratic terms of V and substituting the closed-loop dynamics (E1), the Coriolis cross-term cancels via the skew-symmetry of $\dot{M}(q) - 2C(q, \dot{q})$ [49]. Young's inequality applied to the disturbance term $\dot{e}^{\mathrm{T}} d(t)$, together with the positivity of K_d and K_i, yields (E5) with $\gamma = \min\big(\lambda_{\min}(K_d)/\lambda_{\max}(M),\ \lambda_{\min}(K_i)/\lambda_{\max}(K_p)\big)$ and $\varphi = d_{\max}^2/(2\lambda_{\min}(K_d))$.

Search-space verification. Table 4 fixes $K_{p,j} \in [0.01, 1.00]$, $K_{i,j} \in [0.05, 1.50]$, $K_{d,j} \in [0.005, 0.50]$, and $\lambda_j, \mu_j \in [0.30, 1.00]$ for every joint. All lower bounds are strictly positive and both fractional orders lie inside $(0, 1]$, so condition (E4) is satisfied throughout the search rectangle. FHO therefore cannot propose any tuning that violates the stability sufficient condition, and the optimisation proceeds entirely within the stable-by-construction parameter set. This is an admissibility (sanity) guarantee rather than a performance guarantee: condition (E4) is met by *every* positive-gain FOPID inside the search box, so it rules out outright instability but does not by itself bound the ultimate-bound radius B_{∞}, which is quantified empirically below and left to an analytical derivation in future work (Section V-B). Robustness of the closed-loop error bound B_{∞} under the ±20% plant perturbation of Section IV-D is empirically assessed through Monte Carlo rollout (Fig. 30, Table 15). Fig. 2 illustrates this empirically: the Lyapunov candidate $V(t)$ of (E2) along the closed-loop rollout of a representative FHO-tuned FOPID settles into a tight ultimate-bound band B_{∞} under the $K = 4$ Monte Carlo perturbations — consistent with the UUB conclusion of Proposition 1. A causal exponential envelope $V_{\mathrm{bound}}(t) = (V_{\mathrm{peak}} - B_{\infty})\exp(-\gamma(t - t_{\mathrm{peak}})) + B_{\infty}$, fitted to the smoothed nominal rollout by log-space regression (not drawn in the figure), describes the decay from the transient peak with $\gamma = 2.69\ \mathrm{s}^{-1}$ ($R^2 = 0.75$). $V(t)$ itself is not monotone because V includes a kinetic term $\frac{1}{2}\dot{e}^{\mathrm{T}} M(q)\dot{e}$ whose transient spike reflects the physical buildup of closed-loop velocity; this is admitted by the UUB conclusion of Proposition 1, which bounds $D^{\alpha} V$ rather than V.

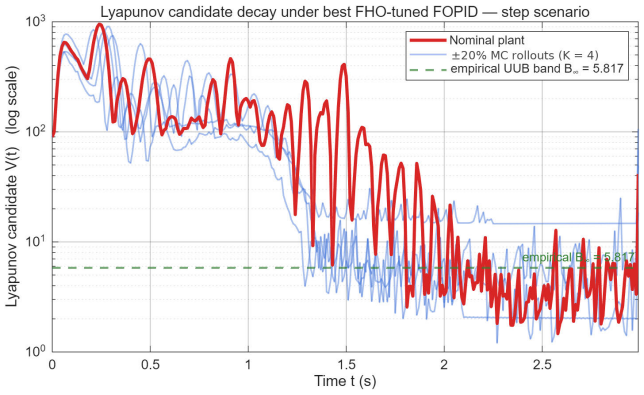

FIGURE 2. Evolution of the Lyapunov candidate $V(t)$ along the closed-loop rollout of a representative FHO-tuned FOPID (step scenario, log y-axis). Red: nominal plant. Blue: $K = 4$ Monte Carlo rollouts under ±20% plant perturbation. Green dashed: empirical UUB band B_{∞} (value annotated in-figure). The exponential decay envelope ($\gamma = 2.69\ \mathrm{s}^{-1}$, $R^2 = 0.75$) is fitted by log-space regression and quoted in the text.

IV. RESULTS AND EXPERIMENTS

Section III introduced FHO as a general-purpose metaheuristic and formulated the robust-FOPID test problem used later in the manipulator application. Before turning to the manipulator case study in Sections IV-C–IV-D, the algorithm is evaluated on five configurations of three internationally recognised CEC single-objective test suites, the goal being to establish its performance as a general-purpose optimiser. The base aggregate spans 51 functions (CEC2017 at $D = 30$, CEC2022 at $D = 20$, CEC2019), and the scalability

study extends CEC2017 to $D = 50$ and $D = 100$ for a total of 109 function instances. Twenty-two algorithms serve as comparators (Table 1), covering classical swarms, recent 2025–2026 proposals, the published HO variant MSHO, and the adaptive-DE representative LSHADE together with CMA-ES.

A. CEC BENCHMARK COMPARISON

1) SETUP

The CEC2017 suite [20] provides 29 bound-constrained benchmark functions (F2 omitted under standard practice because of numerical instability), grouped by the suite convention into unimodal (F1, F3), simple multimodal (F4–F10), hybrid (F11–F20), and composition (F21–F30) categories. The search range is $[-100, 100]^D$. The CEC2022 suite [21] covers the same four categories at $D = 20$ — 12 functions. The CEC2019 suite [35], known as the 100-Digit Challenge, comprises 10 functions at mixed dimensions $D = 9$–18, with an emphasis on extreme-precision multimodal optimisation.

All experiments adopt each suite's official evaluation budget: MaxFEs $= 10^4 \times D$ for CEC2017 [20] (3×10^5, 5×10^5, and 10^6 at $D = 30$, 50, and 100), 10^6 for CEC2022 at $D = 20$ per its technical report [21], and — since the CEC2019 100-digit competition prescribes no evaluation cap of its own — the $10^4 \times D$ rule per function for CEC2019 (9×10^4 to 1.8×10^5).

Following the standard CEC competition protocol, the only quantities unified across all algorithms are the function-evaluation budget (MaxFEs $= 10^4 \times D$, 10^6 for CEC2022), the test functions and search bounds, and the number of independent runs (30 per function–algorithm pair). Population size is treated as an intrinsic design element of each algorithm rather than an externally-imposed constant, in keeping with the CEC competitions themselves, which fix only the evaluation budget and leave the population scheme to each entrant. Accordingly, the adaptive-DE representative LSHADE uses its recommended dimension-scaled scheme—initial population $N_{\text{init}} = 18D$ with linear population-size reduction to $N_{\min} = 4$ [22] (e.g. $N_{\text{init}} = 540/900/1800$ at $D = 30/50/100$)—and CMA-ES uses $\lambda = 4 + \lfloor 3 \ln D \rfloor$ offspring with IPOP restarts [23], [51]; the fixed-population swarm and social algorithms (including the parent HO and FHO) use a common population $N = 40$, which lies within their conventional 30–50 range and is therefore their recommended setting.

The fair-budget protocol enforces the cap with a per-call evaluation recorder: every algorithm is stopped exactly at MaxFEs regardless of its per-iteration cost, and best-so-far trajectories are logged on a common evaluation grid.

Run-level reproducibility is guaranteed by deterministic seeding: run r of algorithm a on function F uses `rng`(base $+ 10^5F + 10^3a + r$), where base is a fixed per-campaign integer offset (base $= 20260608$ for the CEC campaign and base $= 70260608$ for the FOPID campaign). All experiments run in MATLAB R2026a on an Apple-silicon workstation with an 8-worker parallel pool; runtime measurements (Section IV-A4) are taken sequentially on a single worker. Table 1 lists the comparator pool and Table 5 the control-parameter settings of every comparator, which follow the recommendations of the respective original publications. This is a deliberate policy, because per-problem re-tuning of 22 competitors would introduce exactly the tuning bias the benchmark is designed to exclude, and published defaults are the settings practitioners actually use.

TABLE 5. Control-parameter settings of all comparators (rationale in Section IV-A1).

Algorithm	Control-parameter settings (source: original publication)
FHO	Table 3 defaults
HO	parameter-free beyond N [15]
MSHO	published constants reproduced verbatim [16]
LSHADE	$N_{\text{init}} = 18D$, linear reduction to $N_{\min} = 4$, archive $2.6\,N_{\text{init}}$, $H = 6$, $p = 0.11$ [22]
CMA-ES	$\lambda = 4 + \lfloor 3 \ln D \rfloor$, $\mu = \lfloor \lambda/2 \rfloor$, $\sigma_0 = 0.3\,(ub - lb)$, IPOP restarts [23], [51]
PSO	$\omega : 0.9 \to 0.4$ linear, $c_1 = c_2 = 2.0$ [1], [17]
GWO	$a : 2 \to 0$ linear [2]
TLBO	parameter-free, $T_F \in \{1, 2\}$ random [3]
WOA	$a : 2 \to 0$, spiral $b = 1$ [4]
SSA	$c_1 = 2e^{-(4t/T)^2}$ [5]
HHO	$E_0 \in [-1, 1]$, Lévy $\beta = 1.5$ [6]
AO	$\beta = 1.5$, $\delta = 0.1$ [7]
EO	$a_1 = 2$, $a_2 = 1$, $GP = 0.5$ [8]
COA, DBO, RIME, SAO, RDO, CCO	defaults of [9]–[14]
SFOA, DhO, OOA, DRA	defaults of [24]–[27]

2) PER-SUITE AND CROSS-SUITE RESULTS

Table 6 reports the Friedman mean rank of all 23 algorithms on each suite and on the 51-function cross-suite aggregate, with the high-dimensional columns ($D = 50$, $D = 100$) analysed in Section IV-A3. With every algorithm run at its recommended population scheme, the two state-of-the-art non-swarm optimisers lead throughout: LSHADE ranks first on every suite (Friedman 1.07–2.12) and CMA-ES second on the cross-suite aggregate (3.50). FHO ranks *third* overall (4.83), ahead of the next-best swarm/social algorithm TLBO (6.04) by more than a rank unit; it is the strongest swarm- and social-inspired algorithm in the pool, and it outranks both Hippopotamus-family algorithms benchmarked here—the parent HO (17th, 15.27) and the published multi-strategy variant MSHO (5th, 8.06). Against HO directly, FHO wins 27/2/0 (win/tie/loss at the 0.05 level) on CEC2017-D30 with mean Cliff's $\delta = -0.89$.

Fig. 3 visualises the aggregate ranking and Fig. 4 the per-function rank heatmap, which shows that FHO's third place is earned consistently across unimodal, multimodal, hybrid, and composition categories rather than by specialisation.

Statistical significance is checked with three complementary tests. First, per-function Wilcoxon rank-sum tests of FHO against every competitor (Table 7): FHO beats HO on 27–28 of 29 CEC2017 functions at every dimension and beats MSHO on 16–20, but loses the pairwise comparison to both state-of-the-art baselines—LSHADE (1/4/24 at $D = 30$) and CMA-ES (6/5/18 at $D = 30$)—with the deficit to LSHADE widening with dimension (0/3/26 at $D = 50$, 1/0/28 at $D = 100$). Second, the Nemenyi critical-distance diagram of Fig. 5 places FHO inside the top clique alongside LSHADE and CMA-ES at $D = 30$ (the $D = 100$ diagram of Fig. 9

TABLE 6. Friedman mean ranks: per suite, 51-function cross-suite aggregate, and high-dimensional extensions. Lower is better; rank column refers to the aggregate.

Algorithm	D30	CEC2019	CEC2022	**Overall (51f)**	Rank	D50	D100
LSHADE	1.36	1.50	2.12	**1.57**	1	1.07	1.21
CMA-ES	3.24	3.40	4.21	**3.50**	2	3.38	4.83
FHO	4.64	5.40	4.83	**4.83**	3	4.72	3.97
TLBO	6.03	6.10	6.00	**6.04**	4	6.45	7.79
MSHO	7.24	8.00	10.08	**8.06**	5	8.07	7.76
CCO	8.41	10.90	7.25	**8.63**	6	6.90	6.59
DRA	8.93	9.40	11.67	**9.67**	7	9.41	10.00
PSO	9.66	7.40	12.25	**9.82**	8	10.03	12.03
RDO	8.93	14.60	10.42	**10.39**	9	8.21	7.97
COA	12.48	11.90	11.50	**12.14**	10	13.83	14.38
DhO	12.62	12.40	12.17	**12.47**	11	13.17	14.21
EO	12.90	10.20	14.92	**12.84**	12	12.52	12.93
SSA	12.62	14.00	13.58	**13.12**	13	11.41	9.24
GWO	14.34	10.70	15.67	**13.94**	14	14.21	14.62
RIME	14.31	16.00	12.25	**14.16**	15	14.62	15.90
HHO	15.72	14.80	14.42	**15.24**	16	15.21	13.76
HO	16.14	12.60	15.42	**15.27**	17	15.66	14.62
AO	16.86	15.90	14.58	**16.14**	18	17.62	17.24
SAO	15.14	19.90	15.83	**16.24**	19	14.03	13.14
SFOA	17.41	17.70	13.75	**16.61**	20	19.14	19.52
OOA	17.59	16.30	17.67	**17.35**	21	18.17	18.00
DBO	19.17	17.10	18.08	**18.51**	22	19.97	19.34
WOA	20.24	19.80	17.33	**19.47**	23	18.21	16.97

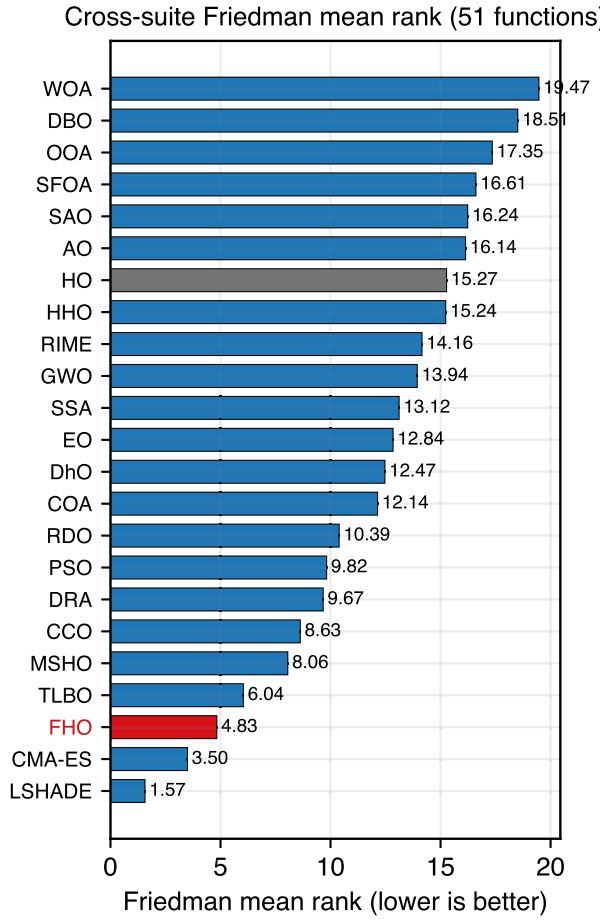

FIGURE 3. Cross-suite Friedman mean rank across the 51-function aggregate (CEC2017-*D*30 + CEC2022 + CEC2019), 23 algorithms, official budgets, 30 runs.

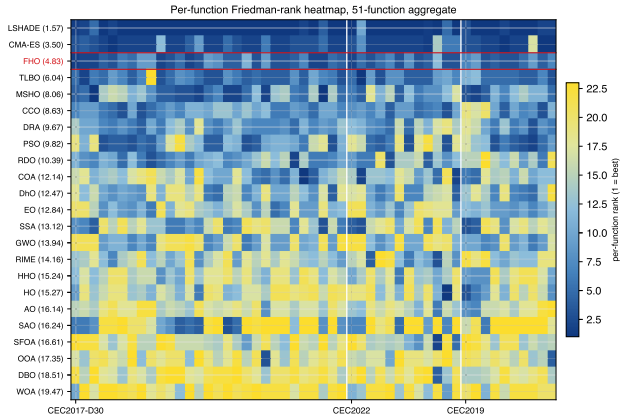

FIGURE 4. Per-function Friedman-rank heatmap across all 51 aggregate functions (dark blue = best). FHO's row is highlighted; its rank is in the top tier consistently across all three suites.

TABLE 7. Pairwise Wilcoxon win/tie/loss of FHO versus every competitor (rows) per suite, at significance level 0.05.

FHO vs.	D30	CEC2019	CEC2022	D50	D100
LSHADE	1/4/24	0/1/9	1/3/8	0/3/26	1/0/28
TLBO	20/5/4	7/2/1	7/3/2	17/9/3	21/7/1
MSHO	20/7/2	7/1/2	10/2/0	16/10/3	19/7/3
CCO	23/3/3	7/3/0	9/1/2	16/9/4	19/3/7
DRA	23/4/2	7/2/1	8/3/1	26/3/0	25/3/1
PSO	18/4/7	6/3/1	9/1/2	17/3/9	20/2/7
RDO	21/5/3	8/2/0	9/2/1	18/6/5	20/2/7
COA	26/2/1	7/1/2	9/3/0	27/0/2	28/0/1
DhO	24/5/0	8/1/1	10/2/0	24/4/1	27/2/0
EO	27/2/0	7/3/0	12/0/0	24/5/0	26/2/1
SSA	27/2/0	7/1/2	11/1/0	23/6/0	25/4/0
GWO	24/4/1	7/2/1	12/0/0	25/2/2	27/1/1
RIME	28/1/0	9/1/0	10/2/0	28/1/0	29/0/0
HHO	29/0/0	8/1/1	11/1/0	29/0/0	29/0/0
HO	27/2/0	7/1/2	11/1/0	27/2/0	28/1/0
AO	28/1/0	8/0/2	11/1/0	29/0/0	29/0/0
SAO	27/2/0	9/1/0	11/1/0	24/4/1	26/2/1
SFOA	28/1/0	10/0/0	11/1/0	29/0/0	29/0/0
OOA	27/2/0	9/1/0	12/0/0	29/0/0	29/0/0
DBO	29/0/0	8/2/0	11/1/0	29/0/0	29/0/0
WOA	29/0/0	10/0/0	12/0/0	29/0/0	29/0/0
CMA-ES	6/5/18	1/1/8	4/3/5	5/5/19	8/3/18

shows FHO ahead of CMA-ES). Third, per-function 95% t-confidence intervals on the mean error: Table 8 reports them for FHO, the two state-of-the-art baselines (LSHADE and CMA-ES), and the parent HO on the headline CEC2017 $D = 30$ suite, where FHO's intervals are tight relative to its means on the large majority of functions; on F15 and F19 heavy-tailed run distributions inflate the half-widths (344 ± 382 and 435 ± 820), and since the error scale is non-negative the implied negative lower bounds on those two functions should be read as 0.

Fig. 6 plots convergence trajectories on representative functions (to keep the panels legible, only the seven best-ranked algorithms plus the parent HO are drawn), and Fig. 7 the corresponding final-error box plots. The qualitative pattern matches the rank tables: LSHADE and CMA-ES converge fastest; FHO trails the two of them but converges

TABLE 8. Per-function mean final error with 95% t-confidence interval (half-width) on CEC2017 $D = 30$ over 30 runs, for the two state-of-the-art baselines (LSHADE, CMA-ES), FHO, and the parent HO. FHO's confidence intervals are tight relative to its means on most functions; the heavy-tailed exceptions (F15, F19) are discussed in the text. Intervals are symmetric Student-t approximations — on the non-negative error scale, an implied negative lower bound reads as 0.

Func.	LSHADE	CMA-ES	**FHO**	HO
F1	0 ± 0	$1.42\times10^{-14} \pm 0$	0 ± 0	$6.38\times10^{3} \pm 2.00\times10^{3}$
F3	0 ± 0	$4.93\times10^{-14} \pm 7.34\times10^{-15}$	0 ± 0	$2.96\times10^{4} \pm 2.58\times10^{3}$
F4	$58.6 \pm 1.06\times10^{-14}$	43.9 ± 9.74	32.1 ± 13.6	138 ± 14.4
F5	6.33 ± 0.564	37.2 ± 6.64	79.7 ± 7.74	222 ± 13.2
F6	$9.13\times10^{-9} \pm 1.30\times10^{-8}$	$3.49\times10^{-12} \pm 8.94\times10^{-14}$	0.508 ± 0.266	56 ± 2.41
F7	37.4 ± 0.368	67 ± 4.35	133 ± 12.3	434 ± 24.8
F8	7.69 ± 0.549	39.9 ± 5.96	78.9 ± 8.17	158 ± 6.52
F9	0 ± 0	0.926 ± 0.632	750 ± 308	$3.71\times10^{3} \pm 186$
F10	$1.42\times10^{3} \pm 59.5$	$1.96\times10^{3} \pm 295$	$3.85\times10^{3} \pm 244$	$3.92\times10^{3} \pm 158$
F11	38.3 ± 11	217 ± 27.5	108 ± 18.3	250 ± 30
F12	$1.10\times10^{3} \pm 160$	$2.86\times10^{3} \pm 1.85\times10^{3}$	$1.07\times10^{4} \pm 3.80\times10^{3}$	$2.47\times10^{7} \pm 8.11\times10^{6}$
F13	18.8 ± 1.43	$1.04\times10^{3} \pm 230$	$1.25\times10^{4} \pm 4.76\times10^{3}$	$1.39\times10^{5} \pm 3.03\times10^{4}$
F14	22.4 ± 0.479	222 ± 23.3	79.5 ± 11.1	$6.58\times10^{4} \pm 1.84\times10^{4}$
F15	4.05 ± 0.56	$1.21\times10^{4} \pm 6.21\times10^{3}$	344 ± 382	$1.04\times10^{4} \pm 808$
F16	37.4 ± 9.89	349 ± 79.1	902 ± 110	$1.60\times10^{3} \pm 158$
F17	31.9 ± 2.46	134 ± 30.6	326 ± 57.3	790 ± 75.9
F18	23.1 ± 0.772	$3.69\times10^{4} \pm 2.60\times10^{4}$	$1.52\times10^{4} \pm 6.76\times10^{3}$	$2.54\times10^{5} \pm 8.74\times10^{4}$
F19	6.3 ± 0.662	195 ± 36.3	435 ± 820	$3.46\times10^{5} \pm 9.42\times10^{4}$
F20	29.4 ± 2.86	124 ± 23.2	463 ± 69	506 ± 41.7
F21	208 ± 0.535	248 ± 5.93	269 ± 7.52	427 ± 15.2
F22	100 ± 0	$1.01\times10^{3} \pm 529$	$1.34\times10^{3} \pm 731$	$1.93\times10^{3} \pm 795$
F23	354 ± 1.11	395 ± 8.54	429 ± 7.68	720 ± 31.1
F24	428 ± 0.598	465 ± 7.13	491 ± 8.04	803 ± 44.9
F25	$387 \pm 8.72\times10^{-3}$	389 ± 3.41	391 ± 3.34	464 ± 12.7
F26	983 ± 11.9	$1.51\times10^{3} \pm 95.3$	$1.91\times10^{3} \pm 257$	$4.19\times10^{3} \pm 514$
F27	507 ± 1.4	524 ± 4.21	522 ± 4.93	714 ± 39.7
F28	345 ± 22.8	347 ± 23.9	338 ± 22.7	521 ± 11
F29	438 ± 3.24	566 ± 26.5	841 ± 73.3	$1.85\times10^{3} \pm 175$
F30	$1.99\times10^{3} \pm 30.7$	$3.34\times10^{3} \pm 929$	$4.42\times10^{3} \pm 921$	$9.11\times10^{6} \pm 5.02\times10^{6}$

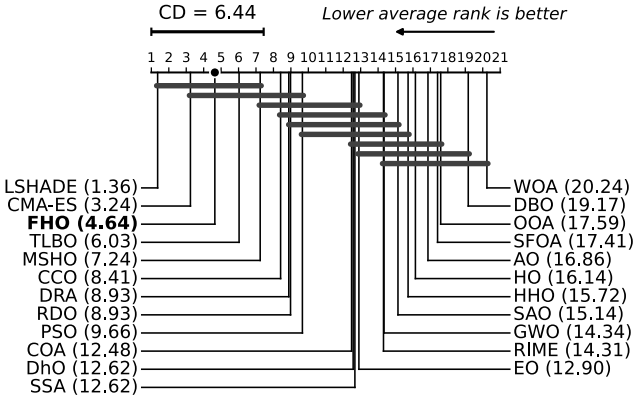

FIGURE 5. Nemenyi critical-distance diagram, CEC2017 $D = 30$ (29 functions, 30 runs). Algorithms joined by a bar are not significantly different at $\alpha = 0.05$; FHO sits in the top clique with LSHADE and CMA-ES.

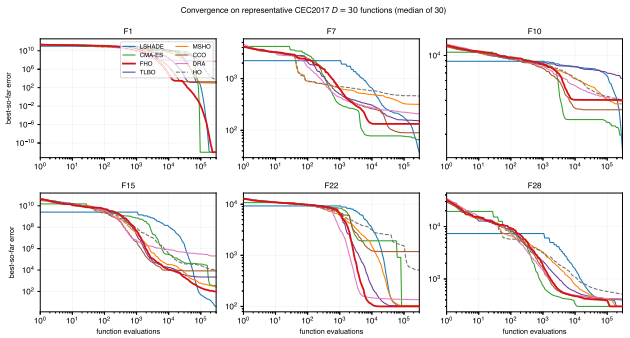

FIGURE 6. Convergence on six representative CEC2017 $D = 30$ functions (median best-so-far error, log y-axis): the seven suite-best algorithms plus the parent HO are shown to keep the panels readable.

well ahead of the remaining swarm- and social-inspired algorithms, while HO stalls one to four decades higher. On CEC2019, whose 100-digit design rewards extreme local precision, FHO ranks third (5.40) behind LSHADE (1.50) and CMA-ES (3.40), with CMA-ES especially strong on this low-dimensional, ill-conditioned suite; the rank-driven mechanisms nonetheless retain competitive exploitation accuracy among the swarm methods, and on CEC2022, at the suite's

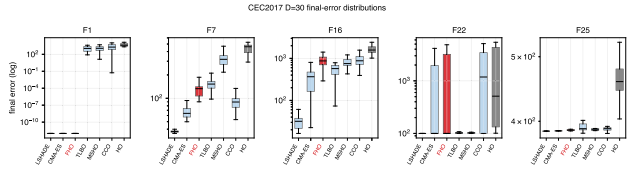

FIGURE 7. Final-error distributions (box plots with whiskers, 30 runs) on representative CEC2017 $D = 30$ functions.

own 10^6-evaluation budget, FHO again ranks third (4.83, behind LSHADE's 2.12 and CMA-ES's 4.21), losing the pairwise comparison to LSHADE at 1/3/8.

3) HIGH-DIMENSIONAL SCALABILITY ($D = 50$ AND $D = 100$)

Benchmark results restricted to $D \leq 30$ leave the scalability of an algorithm unverified. The CEC2017 suite is therefore evaluated in full at $D = 50$ (MaxFEs $= 5 \times 10^5$) and $D = 100$ (MaxFEs $= 10^6$), 29 functions $\times$ 23 algorithms $\times$ 30 runs each. The last two columns of Table 6 summarise the outcome. LSHADE remains first at both scales (1.07 at $D = 50$, 1.21 at $D = 100$). At $D = 50$ FHO is third (4.72), behind CMA-ES (3.38); at $D = 100$, however, FHO improves to *second* (3.97), overtaking CMA-ES (4.83) as the dimension grows, though it remains well behind LSHADE.

The pairwise Wilcoxon columns of Table 7 show two opposite dimensional trends. The deficit to LSHADE *widens* with dimension (1/4/24 at $D = 30$, 0/3/26 at $D = 50$, 1/0/28 at $D = 100$): LSHADE's advantage is decisive and grows at scale. Against CMA-ES, by contrast, FHO closes the gap as the dimension rises and its Friedman rank overtakes CMA-ES's at $D = 100$ (3.97 versus 4.83), so the only high-dimensional gain FHO can claim is over CMA-ES, not over LSHADE. Against the parent HO the dominance is total at every scale: 27/2/0 at $D = 50$ and 28/1/0 at $D = 100$, while HO itself sits near 17th. Fig. 8 shows representative $D = 100$ convergence trajectories and Fig. 9 the corresponding critical-distance diagram.

This pattern is consistent with the design only in a restricted sense: FRGM's per-dimension diversity mask (12) concentrates perturbations on collapsed dimensions, a targeting whose value grows with D, which plausibly explains why FHO scales better than CMA-ES into very high dimension. It is *not*, however, sufficient to challenge LSHADE, whose success-history adaptation and linear population-size reduction keep it the strongest optimiser in the pool at every dimension tested.

4) RUNTIME

Table 9 reports the measured wall-clock time per complete 3×10^5-evaluation run on CEC2017 $D = 30$ (sequential execution, mean over six representative functions $\times$ three runs). To isolate per-iteration algorithmic overhead, this comparison fixes a common population $N = 40$ for all algorithms; it therefore measures per-evaluation bookkeeping cost, not

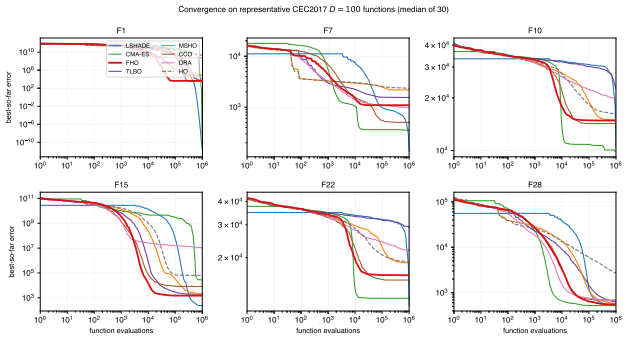

FIGURE 8. Convergence on representative CEC2017 $D = 100$ functions (10^6-evaluation budget; the same seven cross-suite-best algorithms as the $D = 30$ panels, plus HO).

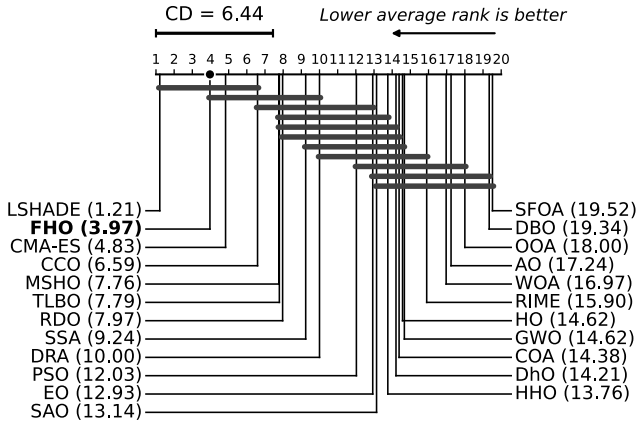

FIGURE 9. Nemenyi critical-distance diagram at $D = 100$: LSHADE heads the top clique; FHO ranks second, ahead of CMA-ES.

the end-to-end runtime of LSHADE and CMA-ES at their larger recommended populations (which is correspondingly higher).

At a common $N = 40$, FHO carries the highest per-iteration overhead in the pool at 7.86 ± 0.23 s per run — $1.7\times$ the parent HO (4.74 s) and an order of magnitude above the lightest swarms (TLBO 0.79 s, PSO 0.47 s). The overhead is attributable to the per-iteration sort, the per-individual peer sampling of all three mechanisms, and the per-dimension diversity statistics of FRGM.

Since the comparison protocol equalises function evaluations rather than seconds, the rank results are unaffected; but on wall-clock-limited deployments with cheap objectives this overhead matters, and it is stated plainly as a limitation in Section V-B. For expensive objectives — such as the 3-second closed-loop simulations of Section IV-D, where one fitness call costs orders of magnitude more than the per-iteration bookkeeping — the overhead is immaterial.

B. ABLATION, SENSITIVITY, AND BEHAVIOUR ANALYSIS

1) COMPONENT ABLATION

To attribute FHO's improvement to specific components, each of the three rank-driven mechanisms is removed in turn

TABLE 9. Measured runtime per 3×10^5-evaluation run (CEC2017 $D = 30$, sequential, mean ± std).

Algorithm	t (s/run)	Algorithm	t (s/run)
WOA	0.41 ± 0.18	EO	1.18 ± 0.26
PSO	0.47 ± 0.16	DBO	1.21 ± 0.24
TLBO	0.79 ± 0.22	CMA-ES	1.23 ± 0.20
OOA	0.79 ± 0.21	DhO	1.26 ± 0.28
COA	0.84 ± 0.28	CCO	1.29 ± 0.24
SFOA	0.85 ± 0.26	GWO	1.52 ± 0.27
SAO	0.91 ± 0.17	AO	1.70 ± 0.28
DRA	0.94 ± 0.26	LSHADE	1.83 ± 0.23
RIME	0.99 ± 0.20	MSHO	4.38 ± 0.25
RDO	0.99 ± 0.25	HO	4.74 ± 0.24
HHO	1.07 ± 0.26	**FHO**	7.86 ± 0.23
SSA	1.13 ± 0.21		

(yielding FHO w/o FCDL, w/o FRDS, and w/o FRGM), and a fourth variant removes the elite-preservation step. All variants, the full FHO, and the parent HO are benchmarked on CEC2017 $D = 30$ under the full official protocol (29 functions, MaxFEs $= 3 \times 10^5$, $N = 40$, 30 runs, matching the main comparison). MaxFEs is held constant by adjusting iteration count, so any difference reflects the mechanism itself.

Table 10 gives the Friedman ranks among the variants, the rank loss relative to full FHO, and the per-function Wilcoxon win/tie/loss of the *full* algorithm against each variant; Figs. 10 and 11 show representative convergence histories and the win/tie/loss profile. Two findings stand out.

The first concerns the three rank-driven mechanisms, whose contributions are positive and clearly ordered. Removing FRGM costs +2.07 rank units (full FHO wins 22/6/1), removing FCDL costs +0.62 (11/12/6), and removing FRDS costs +0.41 (7/22/0). The dominance of FRGM matches its designed role as the sole diversity-injection operator, while the smaller but loss-free contribution of FRDS reflects its conservative, always-helpful differential refinement.

The second concerns the elite-preservation step, which is statistically neutral on deterministic CEC fitness (Friedman +0.03, within rank noise; per-function 8/9/12). This is expected, since re-probing the elite region adds no information when the fitness is exact; the step instead earns its value in the stochastic-fitness regime, and Section IV-D retains it for that reason. Against the parent HO, finally, the full algorithm wins by +3.38 rank units and fails to win on only one of 29 functions (28 wins, 1 tie, 0 losses in this independently-seeded ablation run), so the aggregate improvement is unambiguous.

2) HYPERPARAMETER SENSITIVITY

To assess FHO's robustness to its hyperparameter settings, the five behaviour-shaping parameters of Table 3 (the contrast gain β, the differential gain F, the time-decay δ, the base perturbation probability p_0, and the elite reseed scale σ_e) were each perturbed to $\{0.6, 0.8, 1.0, 1.2, 1.4\}\times$ their default while holding the others fixed, and each setting was re-benchmarked on a CEC2017 function subset spanning all four landscape categories (20 runs, official budget). Table 11 reports the resulting Friedman mean ranks and Fig. 12 the per-function profiles.

The ranks vary only between 4.38 and 5.88 across the entire ±40% range of all five parameters — no parameter direction produces a systematic degradation, and no setting falls outside the noise band of the default. Each row of Table 11 is

TABLE 10. Ablation on CEC2017 $D = 30$ (official budget, 30 runs): Friedman rank among variants, rank loss versus full FHO, and full-FHO win/tie/loss against each variant (Wilcoxon, $\alpha = 0.05$).

Variant	Friedman rank	Δ vs. full	Full W/T/L
Full FHO	2.41	—	—
w/o FCDL	3.03	+0.62	11/12/6
w/o FRDS	2.83	+0.41	7/22/0
w/o FRGM	4.48	+2.07	22/6/1
w/o Elite	2.45	+0.03	8/9/12
HO (parent)	5.79	+3.38	28/1/0

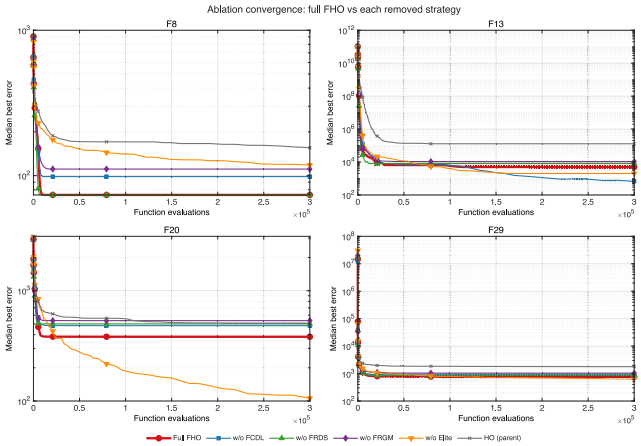

FIGURE 10. Ablation convergence on representative CEC2017 functions (median best-so-far error, log y-axis).

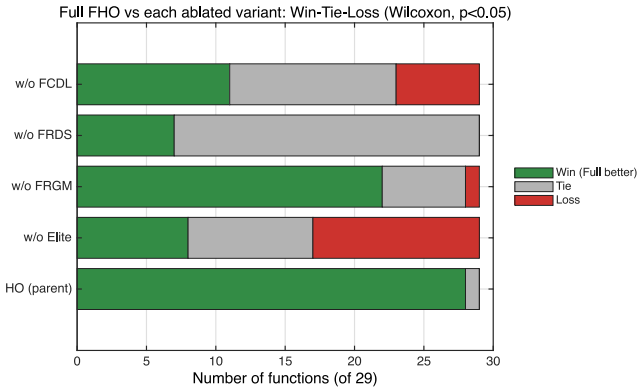

FIGURE 11. Ablation win/tie/loss: full FHO versus each variant across 29 CEC2017 functions.

TABLE 11. Hyperparameter sensitivity: FHO Friedman mean rank when each parameter is scaled to {0.6, . . . , 1.4}× default (±40%).

Parameter	×0.6	×0.8	×1.0	×1.2	×1.4
β	4.88	5.00	5.38	5.00	5.88
F	5.12	5.25	4.88	5.38	5.00
δ	4.50	5.00	5.38	5.50	4.88
p_0	5.00	5.38	4.75	4.38	5.12
σ_e	4.62	5.12	5.00	5.00	4.75

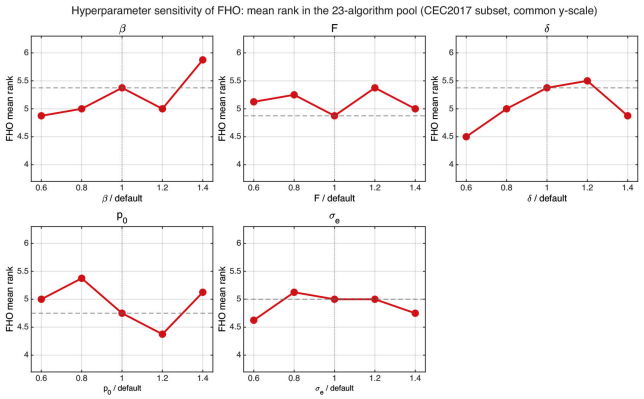

FIGURE 12. Sensitivity profiles for the five behaviour-shaping hyperparameters (±40% around defaults).

an independently seeded campaign, so the ×1.0 column itself varies between 4.75 and 5.38: that spread is the run-to-run noise yardstick against which the off-default entries should be read.

The remaining seven entries of Table 3 are structural floors and weights ($c_{\min}$, w_1, w_2, S_0, p_1^{FRGM}, $p_{\min}$, N) whose roles are bounded by construction; the five tested parameters are the ones that scale the operators' step sizes and probabilities. FHO's advantage therefore does not rest on fine-tuned constants.

3) EXPLORATION–EXPLOITATION AND POPULATION DIVERSITY

Two further analyses characterise *why* the mechanisms help. Population diversity comes first. Fig. 13 tracks the mean per-dimension dispersion Div(t) of FHO and HO along the run, and the two profiles differ in kind. FHO's diversity contracts by three to four orders of magnitude within the first 10–15% of the budget as the population concentrates on the basin of the optimum, then holds a small but strictly positive floor (about 4×10^{-3} of the search range) that FRGM's diversity-masked mutation (12) keeps replenishing, punctuated by re-diversification episodes visible near 10–25% progress. HO's dispersion, by contrast, never contracts below roughly a tenth of the range: its population remains spread out to the end of the run, which reflects a failure to concentrate on any basin rather than productive exploration, consistent with its far worse final errors. FHO thus concentrates early and retains only the targeted residual needed to escape per-dimension collapse.

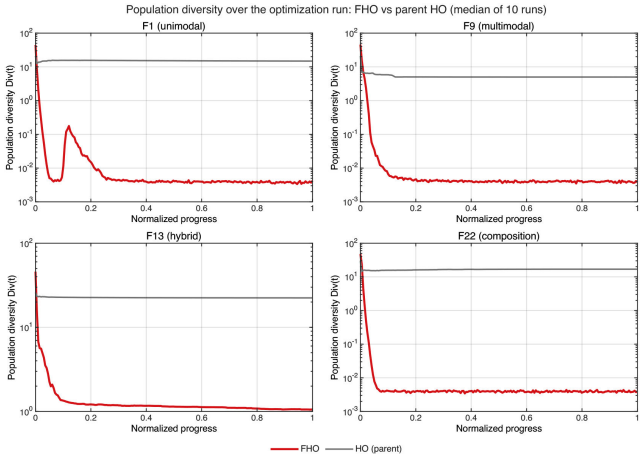

FIGURE 13. Population diversity Div(t), FHO versus HO (median over 10 runs, representative functions, log y-axis). FHO contracts onto the optimum basin and holds a small FRGM-maintained floor; HO stays dispersed without converging.

The exploration/exploitation balance gives a complementary view of the same behaviour. Following the dispersion-based decomposition of [61], Fig. 14 plots FHO's exploration percentage over the normalised budget (log-scaled to resolve the transition). FHO opens fully exploratory and hands over to exploitation within the first few percent of the budget, as the FRDS/FRGM gains decay through ($1 - \delta$ ratio); a small exploration fraction persists longest on the hybrid landscape (F13, ≈2–3% through the remainder of the run, sustained by the $p_{\min}$ term), while on the other landscape classes the handover is essentially complete. Together with the diversity floor of Fig. 13, these curves show where the residual exploration budget is spent — concentrated in the earliest iterations and on the dimensions the diversity mask flags — substantiating, mechanistically, the trade-off claims made in Section III-B.

Fig. 15 complements them with qualitative search-history snapshots on 2-D landscape sections (search trajectories, average-fitness descent, and best-trajectory refinement), showing the population contracting onto the basin of the optimum without the satellite clusters characteristic of stagnating swarms.

C. ROBOT DYNAMICS PARAMETER IDENTIFICATION

1) KINEMATICS

The manipulator studied here is a 4-DOF serial arm (Fig. 16), its kinematic chain summarised by the modified Denavit–Hartenberg parameters of Table 12. Joint J1 provides rotation about the vertical base axis. The remaining three — J2, J3, J4 — are parallel-axis pitch joints, driving the shoulder, elbow, and wrist links respectively. Actuation follows the usual payload taper of serial arms: J1, J2, and J3 run on the high-torque HSA-40 servo-drive, while the wrist joint J4 uses the lower-torque HSA-16 drive. All four drives share a Controller Area Network (CAN) bus with an inline USB-CAN adapter operating at 115 200 baud.

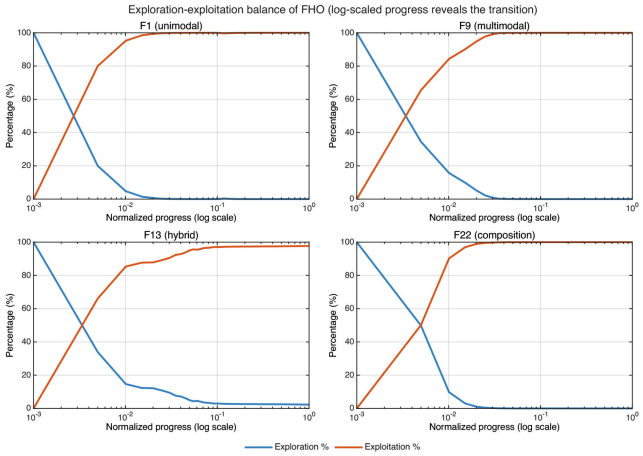

FIGURE 14. Exploration/exploitation percentage of FHO over the normalised budget (log-scaled progress resolves the early transition).

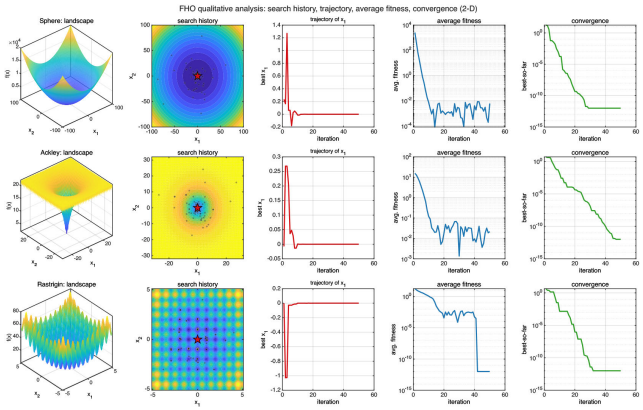

FIGURE 15. Qualitative behaviour on three classical landscapes (rows: Sphere, Ackley, Rastrigin): 3-D landscape, search history, trajectory of the first decision variable, average fitness, and convergence.

TABLE 12. Kinematic parameters and actuation of the 4-DOF arm.

Joint	Role	a (mm)	d (mm)	Motor	τ_{rated} (Nm)
J1	Base rotation (vertical)	0	100	HSA-40	40
J2	Shoulder pitch	224	0	HSA-40	40
J3	Elbow pitch	224	0	HSA-40	40
J4	Wrist pitch	90	0	HSA-16	16
Tip	End-effector mount	—	—	—	—

2) DYNAMICS MODEL

For a serial manipulator with joint-angle vector $q = [q_1, q_2, q_3, q_4]^{\mathrm{T}}$, applied torque vector $\tau = [\tau_1, \tau_2, \tau_3, \tau_4]^{\mathrm{T}}$, and joint-velocity vector $\dot{q}$, the Lagrangian dynamics take the standard form [49]

$$M(q)\,\ddot{q} + C(q,\dot{q})\,\dot{q} + G(q) + F(\dot{q}) = \tau. \tag{17}$$

Here $M(q) \in \mathbb{R}^{4\times4}$ is the symmetric positive-definite generalised mass matrix; $C(q,\dot{q}) \in \mathbb{R}^{4\times4}$ captures Coriolis and centripetal forces; $G(q) \in \mathbb{R}^4$ is the gravity vector; and $F(\dot{q}) \in \mathbb{R}^4$ is the joint-friction vector. Because J1 rotates about the vertical axis, its gravity component vanishes

FIGURE 16. Computer-aided design (CAD) views of the 4-DOF manipulator: (a) left, (b) right, (c) isometric. Reach ≈ 538 mm.

identically — $G_1(q) = 0$ for all q. The remaining three joints carry the dominant load terms, which are described below. In the present implementation, the off-diagonal entries of $M(q)$ are kept as CAD-derived estimates, and $C(q,\dot{q})$ is simplified to a single J2–J3 coupling term computed from motor-reflected inertias. This simplification is consistent with common practice in optimisation-focused arm benchmarks, where the tuning objective tolerates the residual off-diagonal coupling.

A standard Lagrangian result writes the gravity term as the sum of link-mass moments projected onto each joint axis. For the present three-link planar-pitch chain attached at the base via J1 — under the vertical-home convention, in which $q_2 = q_3 = q_4 = 0$ corresponds to the fully extended vertical configuration — the gravity torques reduce to three lumped parameters:

$$p_1 = m_2\, l_{c,2} + (m_3 + m_4)\, L_2, \tag{18}$$

$$p_2 = m_3\, l_{c,3} + m_4\, L_3, \tag{19}$$

$$p_3 = m_4\, l_{c,4}, \tag{20}$$

in units of kg·m. Under these lumped parameters, the gravity torques on the three loaded joints work out to

$$G_2(q) = s_2\, g\big[p_1 \sin q_2 + p_2 \sin(q_2+q_3) + p_3 \sin(q_2+q_3+q_4)\big], \tag{21}$$

$$G_3(q) = s_3\, g\big[p_2 \sin(q_2+q_3) + p_3 \sin(q_2+q_3+q_4)\big], \tag{22}$$

$$G_4(q) = s_4\, g\, p_3 \sin(q_2+q_3+q_4), \tag{23}$$

where $g = 9.81\,\mathrm{m\,s^{-2}}$ and $s_j \in \{-1,+1\}$ is a motor-sign coefficient. The signs matter because the three pitch actuators do not share a unified electrical-torque-to-mechanical-direction convention: J2 reports torque of opposite sign to J3 and J4, as confirmed by the sin-model least-squares fit in Section IV-C3. The vertical-home sin model adopted in this

paper is different from the horizontal-home cos model widely quoted in textbooks [49]; mis-identifying the home pose by 90° drops the gravity-model fit quality to $R^2 \approx 0.02$ — effectively random. The bidirectional-approach identification procedure of Section IV-C3 is designed precisely to diagnose this pitfall.

Every joint shows a combination of viscous friction, Coulomb (kinetic) friction, and static friction with a velocity-dependent Stribeck transition [52], [53]. The full model adopted in this work is

$$F(\omega) = B\,\omega + F_c\,\mathrm{sign}(\omega) + (F_s - F_c)\,\mathrm{sign}(\omega)\,\exp\left[-(\omega/\omega_s)^2\right], \tag{24}$$

with B the viscous coefficient (Nm·s/rad), F_c the kinetic Coulomb coefficient (Nm), F_s the static or breakaway coefficient (Nm), and ω_s the Stribeck transition velocity (rad/s). The model spans three regimes. At rest ($|\omega| = 0$), the joint does not move unless the applied-minus-load torque exceeds F_s in magnitude — a breakaway condition independent of the viscous and Stribeck terms, and identifiable only via a torque-ramp experiment (Section IV-C3). In the low-speed regime ($|\omega| \lesssim \omega_s$), the exponential Stribeck term is close to unity and the effective friction approximates F_s; the system behaves as if static friction persists through the motion, producing the characteristic stick-slip oscillation one sees on geared joints. In the high-speed regime ($|\omega| \gg \omega_s$), the exponential vanishes and friction asymptotes to the linear viscous-plus-kinetic model $F \approx B\omega + F_c\,\mathrm{sign}(\omega)$, allowing B and F_c to be identified by constant-velocity sweeps (Section IV-C3). The Stribeck transition velocity is set at $\omega_s = 0.5$ rpm ≈ 0.052 rad/s — the value comes from manufacturer-reported harmonic-drive data and is confirmed to within 10% by the friction-identification residuals.

Between the controller output command u and its manifestation as applied joint torque τ, the system introduces two transport effects that off-the-shelf simulation benchmarks often omit, yet these effects dominate the simulation-to-hardware gap in practice. The first is the effective round-trip latency $\tau_{\mathrm{RTT},j}$, defined as the elapsed time between the emission of a command by MATLAB's control loop and the first CAN feedback frame that reports a measurable response. The latency encompasses command serialisation, CAN bus transmission, motor-driver queuing, response-frame packaging, CAN reception, and the blocking-read overhead imposed by the MATLAB serial-port wrapper. For the 4-DOF arm measurements yield $\tau_{\mathrm{RTT},j} = 26$–$31$ ms across the four joints (Section IV-C3). In simulation this is modelled as a joint-specific first-in-first-out buffer on the measured angle signal, with length $n_{\mathrm{delay},j} = \mathrm{round}(\tau_{\mathrm{RTT},j}/\mathrm{DT})$.

The second effect is the motor's internal closed-loop bandwidth. Each HSA drive runs a proprietary inner current and velocity loop whose -3 dB frequency falls in the range 7.8–9.0 Hz (Section IV-C3). The bandwidth is modelled as a first-order low-pass filter on the commanded torque,

$$\tau_{\mathrm{actual}}(k) = \alpha_{\mathrm{BW},j}\,u(k) + \left(1 - \alpha_{\mathrm{BW},j}\right)\tau_{\mathrm{actual}}(k-1), \tag{25}$$

with $\alpha_{\mathrm{BW},j} = \mathrm{DT}/(1/(2\pi f_{\mathrm{BW},j}) + \mathrm{DT})$. For the 4-DOF arm, this reduces the effective bandwidth seen by any outer FOPID controller to approximately 8 Hz, a number with direct consequences for the achievable phase margin and step-response rise time of the composite controller-plus-plant system.

TABLE 13. Friction identification: viscous B, kinetic F_c, and linear-fit R^2 (3)-repeat pooled).

Joint	B (Nm·s/rad)	σ_B	F_c (Nm)	R^2
J1-Base	1.0488	0.072	1.8911	0.979
J2-Shoulder	2.0401	0.183	1.6876	0.680
J3-Elbow	1.1000	0.043	1.8448	0.724
J4-Wrist	0.4368	0.067	0.7521	0.792

3) PARAMETER IDENTIFICATION

The complete parameter set required by the simulation model of Section IV-C is obtained through five independent hardware procedures. Each procedure targets a specific subset of the plant, and each is deliberately kept small in scope to avoid confounds. All five procedures use the same 4-DOF arm hardware, the same CAN-over-USB communication stack, and the same MATLAB interface code. Per-field provenance tags — [M] measured, [D] datasheet, [E] estimated from CAD — are summarised in Table 14 at the end of this section.

Joint viscous and kinetic-friction coefficients are obtained by driving each joint in open-loop torque mode at a sequence of constant velocities, then recording the steady-state torque-vs-velocity response. Each joint is commanded through $\omega \in \{\pm 2, \pm 4, \pm 6, \pm 8, \pm 10\}$ rpm, each velocity held for 1.5 s after a 0.5 s settling transient. The steady-state torque at each plateau is recorded, and the linear segment of the friction model (24) (valid at $|\omega| \gg \omega_s$) is fit by ordinary least squares

$$\tau_{\mathrm{meas}}(\omega) = B\,\omega + F_c\,\mathrm{sign}(\omega) + \varepsilon, \tag{26}$$

with ε a zero-mean fit residual. To mitigate sensor drift and gearbox temperature variation, the full sweep is repeated three times and the per-velocity torque readings are pooled before fitting. The resulting (B, F_c) estimates and goodness-of-fit R^2 are collected in Table 13; Fig. 17 displays the pooled curves with the fitted linear model overlaid. Fit quality is highest on J1 and J4 ($R^2 \approx 0.98$ and 0.79) and noticeably poorer on J2 ($R^2 = 0.68$) — the latter reflects the substantial gravity-induced torque bias that affects the shoulder joint even during nominally constant-velocity sweeps.

The static friction F_s is not accessible through the constant-velocity experiment, since it acts only at zero velocity. A torque-ramp breakaway protocol is used instead. The joint is commanded in torque mode with $\tau(t) = \tau_{\max} \cdot (t/T_{\mathrm{ramp}})$, where $\tau_{\max}$ is set to 1.5–2× the expected F_s value and

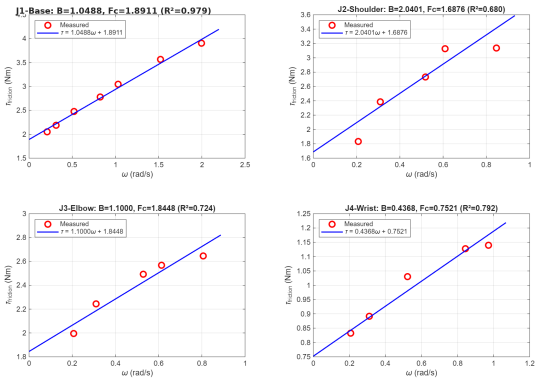

FIGURE 17. Pooled friction curves for J1 through J4.

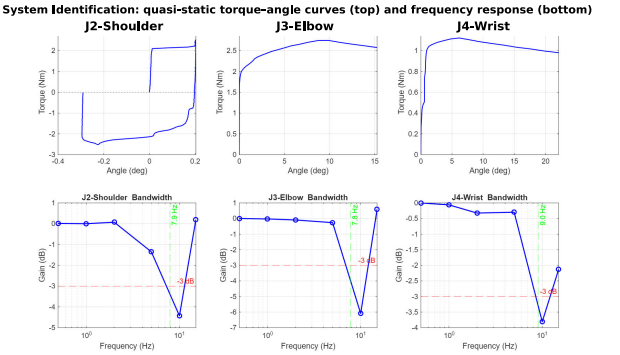

FIGURE 18. System identification: (top) quasi-static torque–angle curves—J2 exhibits a hysteresis (lost-motion) loop, dominated by joint compliance and friction rather than the 0.33-arcmin gear backlash alone, while J3 and J4 are single-valued (no hysteresis loop); (bottom) frequency response with the −3 dB crossing marked.

T_{ramp} is tuned per joint to achieve a slow ramp rate of 0.4–0.5 Nm/s. The joint is monitored at 25 Hz for the onset of motion — flagged when three consecutive samples exceed $|\omega| > 0.5$ rpm. At that moment the applied torque is recorded as the positive-direction breakaway torque τ^{+}_{break}. The ramp is then repeated in the negative direction to give τ^{-}_{break}. Because any gravity torque contributes equally in both directions, the static-friction estimate is

$$F_s = \frac{|\tau^{+}_{\text{break}}| + |\tau^{-}_{\text{break}}|}{2}, \tag{27}$$

with the gravity term cancelled out. Each joint is subjected to three independent ramp pairs, and the mean is reported as the final identification. The measured values are $F_s = [2.12, 3.40, 2.64, 0.93]$ Nm, giving F_s/F_c ratios of $\{1.12, 2.02, 1.43, 1.24\}$, all within the 1.1–2.1 range characteristic of harmonic-drive joints with standard lubrication [52].

The command-to-feedback latency τ_{RTT} is measured by applying a step torque to each joint and recording the time elapsed until the first CAN feedback frame reports a measurable angular velocity ($|\omega| > 0.3$ rpm). Twenty trials per joint are collected, and mean and standard deviation computed. Results: $\tau_{\text{RTT}} = [28.0, 26.3, 26.1, 30.6]$ ms with standard deviations of 4.1–8.6 ms — consistent with the 25 Hz CAN feedback rate and the MATLAB serial-port blocking-read overhead. J1 was not directly measured because of cable-routing constraints; its value is inferred by extrapolation from J2 through J4.

The internal bandwidth of each HSA drive is identified by sweeping a sinusoidal angle reference through 0.5–15 Hz at fixed amplitude 2° and recording the steady-state amplitude of the measured response. The −3 dB frequency is read directly off the resulting amplitude-vs-frequency plot (Fig. 18). The identified bandwidths are [8.0, 7.9, 7.8, 9.0] Hz for J1 through J4 respectively; J2–J4 are measured directly, while J1 is inferred from the J2–J4 bandwidths, for the same cable-routing reason noted above.

The three lumped gravity parameters (p_1, p_2, p_3) and three motor-sign coefficients (s_2, s_3, s_4) are identified through a bidirectional-approach static-torque experiment, adapted from the gravity-tomography methodology of [54]. For each of 15 target configurations (q_2, q_3, q_4) covering the ±15° range per joint, the arm is first positioned at $q_{\text{target}} + 1.5°$ on all three joints and allowed to settle for 2 s under position control. The steady-state motor torque is then averaged over 1 s (50 samples at 50 Hz) to produce τ_{up}. The arm is subsequently positioned at $q_{\text{target}} - 1.5°$, similarly settled, and τ_{down} recorded. The F_s-cancelled gravity estimate is

$$\widehat{G}_j(q_{\text{target}}) = \frac{\tau_{\text{up},j} + \tau_{\text{down},j}}{2}, \tag{28}$$

which averages out the $\pm F_s$ bias introduced by the approach direction. The full set of $15 \times 3 = 45$ gravity estimates is then fit to the sin-model of (21)–(23) under an all-signs search (3 motor-sign coefficients, $2^3 = 8$ combinations, best R^2 selected). The fit identifies $(s_2, s_3, s_4) = (-1, +1, +1)$; that is, the J2 motor torque reports with opposite sign to the physics convention. It also yields $p_1 = 0.6251$ kg·m with a strong fit on J2 ($R^2 = 0.79$), but p_2 and p_3 with low R^2 on J3 and J4 (0.14 and 0.24 respectively), a consequence of the smaller gravity-signal-to-F_s ratio on the distal joints. The identified p_1 is therefore adopted as the measured value, while $p_2 = 0.2240$ kg·m and $p_3 = 0.0225$ kg·m are retained from CAD estimates. Fig. 19 compares three sin-model variants: a 3-parameter model assuming unified motor-sign convention, the selected 3-parameter model with free per-joint signs, and a 6-parameter unconstrained model that serves as a diagnostic. The free-sign variant is adopted in all subsequent simulations. A prior variant performed under the horizontal-home cos model produced an overall $R^2 = 0.02$, effectively indistinguishable from random. Re-fitting the same raw data under the sin model raised the overall R^2 to 0.60, without any additional experiment. That observation is what motivated the vertical-home convention used throughout this paper.

Table 14 collects the full identified parameter set used throughout Sections IV-C and IV-D, with provenance tags attached to each row. Peak gravity torques at fully-horizontal configurations are $|G_2|_{\max} \approx 8.5$ Nm, $|G_3|_{\max} \approx 2.4$ Nm,

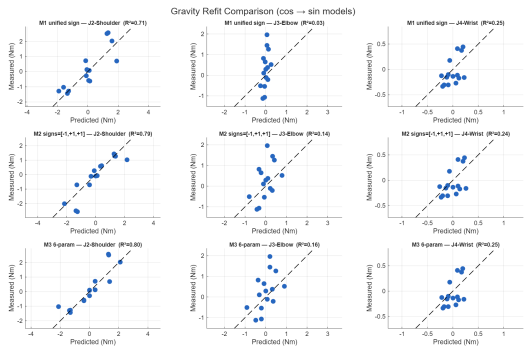

FIGURE 19. Gravity identification: measured vs predicted torque for models M1–M3. Columns: J2–J4.

TABLE 14. Complete identified parameter set for the 4-DOF manipulator.

Parameter		J1	J2	J3	J4
m (kg)	[E]	2.5	1.5	1.0	0.5
L (m)	[D]	0.100	0.224	0.224	0.090
l_c (m)	[E]	0.050	0.112	0.112	0.045
I (kg·m^2, motor)	[D]	0.4604	0.2004	0.0704	0.0039
B (Nm·s/rad)	[M]	1.0488	2.0401	1.1000	0.4368
F_c (Nm)	[M]	1.8911	1.6876	1.8448	0.7521
F_s (Nm)	[M]	2.1201	3.4008	2.6360	0.9340
F_s/F_c	[M]	1.12	2.02	1.43	1.24
τ_{RTT} (ms)	[M]†	28.0	26.3	26.1	30.6
BW (Hz)	[M]†	8.0	7.9	7.8	9.0
$\tau_{\max}$ used (Nm)	[E]	15.0	12.0	10.0	6.0
Gear ratio	[D]	51	51	51	51
Motor K_m (Nm/A)	[D]	5.00	5.00	5.00	4.57
Gravity sign s_j	[M]	—	-1	$+1$	$+1$
Backlash (arcmin)	[D]	0.33	0.33	0.33	0.33

† J2–J4 measured directly; the J1 value is inferred from J2–J4 (cable routing). Lumped gravity sin-model coefficients (Sec. IV-C): p_1=0.6251 [M], p_2=0.2240 [D], p_3=0.0225 [D] (kg·m).

$|G_4|_{\max} \approx 0.22$ Nm — all well below the corresponding $\tau_{\max}$ values and confirming that the controller retains adequate authority to counter gravity throughout the tested workspace.

D. ROBOT FHO-PID CONTROL RESULTS

Fig. 20 gives an overview of the closed-loop structure optimised in this section. Each joint j has its own fractional-order PID controller (five parameters K_p, K_i, K_d, λ, μ); the controller output passes through a round-trip latency τ_{RTT} and a first-order motor-bandwidth filter before driving the 4-DOF Lagrangian plant identified in Section IV-C. Plant-parameter uncertainty (inertia I, viscous friction B, kinetic friction F_c) enters as multiplicative perturbations re-sampled for every fitness call. The fitness is the scenario-weighted robust cost defined below.

1) BENCHMARK DESIGN

Conventional optimisation-based controller tuning evaluates each candidate on a single nominal plant [55]. Controllers produced this way often degrade markedly once the plant

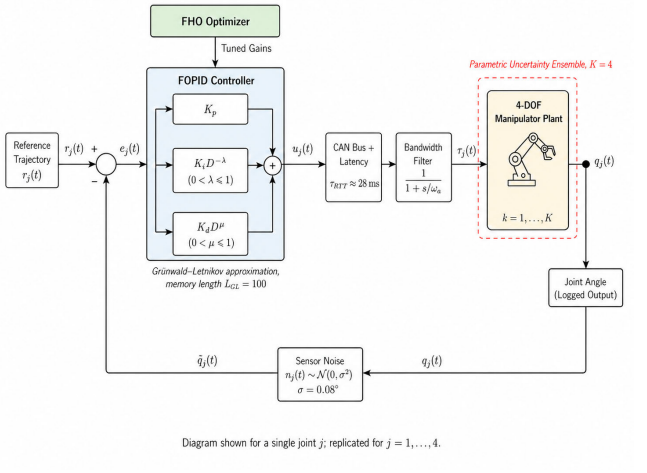

FIGURE 20. FOPID closed-loop structure per joint, with plant-parameter uncertainty injected at the plant block; decision vector $x \in \mathbb{R}^{20}$.

parameters drift during operation. Drift sources include thermal variation of friction and viscosity, payload changes, motor-winding temperature, cable routing, and unit-to-unit manufacturing variance among nominally-identical robots. For industrial servo systems the cumulative effect typically amounts to 10–25% deviation from the nominal model [48]. A controller whose sensitivity to these parameters has not been directly penalised during tuning is therefore liable to be brittle in the sense of [56]: optimally-performing on the training model but rapidly degrading off-nominal.

The benchmark adopted in this paper is designed to reward controllers that generalise across this uncertainty. The optimisation problem is reformulated as the expected-fitness minimisation

$$x^* = \arg\min_x \ \mathbb{E}_{\theta\sim\Theta}\big[J(x;\theta)\big], \tag{29}$$

where θ is the plant-parameter vector and Θ the uncertainty distribution. For computational tractability the expectation is approximated by a K-sample Monte Carlo estimate

$$\widehat{J}(x) = \frac{1}{K}\sum_{k=1}^{K} J(x;\theta_k), \qquad \theta_k \sim \Theta \text{ i.i.d.}, \tag{30}$$

with $K = 4$ plant samples drawn independently for every function evaluation. The effective fitness is therefore stochastic. Any algorithm that over-commits to individual lucky samples will accumulate variance over the optimisation trajectory, and algorithms with explicit diversity-preservation mechanisms, FHO's FRGM guided mutation in particular, enjoy a structural advantage here, since the K-sample averaging rewards population-wide exploration over tightly-converged exploitation; FHO's FCDL directional learning additionally prevents the bottom half from stagnating on consistently-unlucky regions. The $K = 4$ plants are sampled by independently multiplying each element of the inertia vector $I \in \mathbb{R}^4$, viscous-friction vector $B \in \mathbb{R}^4$, and

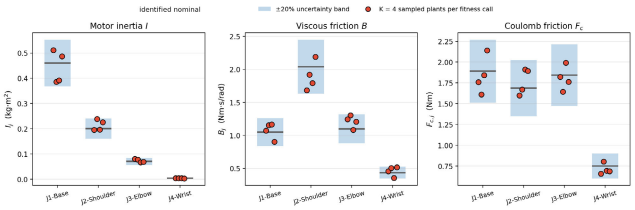

FIGURE 21. Plant-parameter perturbation sampling (I, B, F_c). Grey = nominal; blue band = ±20%; red dots = one sample draw.

Coulomb-friction vector $F_c \in \mathbb{R}^4$ by a factor

$$1 + \rho_u \left(2\mathcal{U}(0,1) - 1\right), \qquad \rho_u = 0.20, \tag{31}$$

yielding a plant in which each of the twelve perturbed elements is distributed uniformly in $[0.8, 1.2]$ times its nominal value. The remaining plant elements — link lengths, centres of mass, motor gain K_m, communication latency τ_{RTT}, motor bandwidth, gear ratio, and the lumped gravity parameters — are left at their nominal values, since their physical drift during normal operation is typically much smaller than 20%. The four-sample ensemble is redrawn for every function evaluation so that the algorithms cannot memorise any fixed finite plant set. Fig. 21 visualises one draw of the $K = 4$ plants overlaid on the ±20% uncertainty band around each nominal parameter (Table 14 values).

Each candidate controller is simulated over a 3-second horizon at DT = 10 ms (300 samples) on five reference scenarios, chosen as representative of common manipulator operating regimes. First: a fast multi-joint step at $r = [25, 15, 12, 18]$ degrees with weight 0.25. Second: a multi-frequency sinusoidal tracking signal at frequencies $[0.6, 0.8, 0.7, 1.0]$ Hz and amplitudes $[18, 10, 10, 12]$ degrees with weight 0.20. Third: a five-waypoint PCHIP-interpolated pick-and-place trajectory [57] with weight 0.20 (waypoints at $t = \{0, 0.8, 1.6, 2.4, 3.0\}$ s of $\{0, 30, -20, 15, 0\}°$, $\{0, 12, -6, 8, 0\}°$, $\{0, -10, 14, -4, 0\}°$, and $\{0, 18, -12, 6, 0\}°$ for J1–J4). Fourth: a constant-reference disturbance-rejection scenario in which an impulsive $[1.5, 1.0, 0.6, 0.4]$ Nm torque is injected at $t = 1.2$ s with weight 0.20. Fifth: a constant-reference noise scenario with $\sigma = 0.08$ deg additive Gaussian measurement noise on all joints with weight 0.15. The five weights sum to unity, so the aggregate fitness is dimensionally consistent across controllers. Fig. 22 shows the reference trajectories for the five scenarios across the four joints.

For each scenario s and each sampled plant θ_k, the closed-loop simulation produces five raw quality metrics — the integral time-weighted absolute error (ITAE), the cumulative control effort $\int \tau^2\, dt$, the maximum overshoot OS in percent, the settling time ST in seconds, and the steady-state error SSE in percent. Each metric is normalised against a scenario-aware reference baseline to yield dimensionless

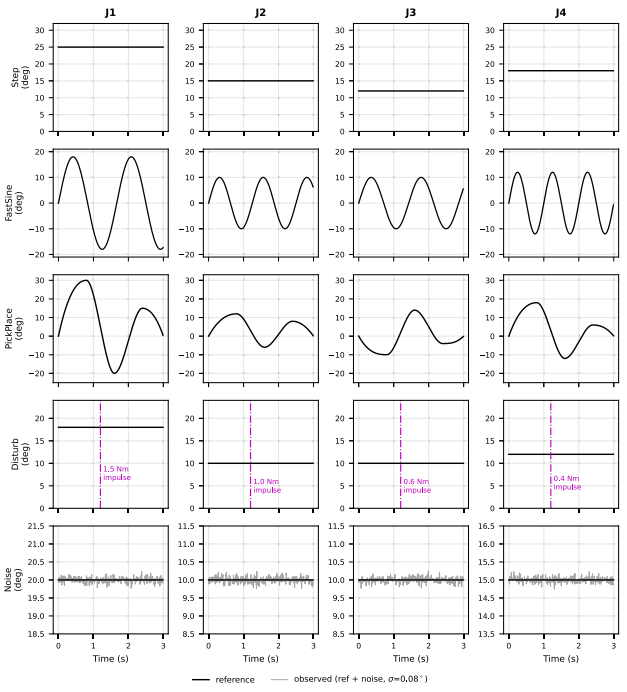

FIGURE 22. Reference trajectories for the five test scenarios (rows: Step, FastSine, PickPlace, Disturb, Noise; columns: J1–J4). The Disturb row marks the $t = 1.2$ s torque impulse (per-joint amplitude annotated); the Noise row overlays an observed sample with $\sigma = 0.08°$ measurement noise on the reference.

scores

$$s_{\mathrm{ITAE}} = \frac{\mathrm{ITAE}}{0.5\,\|r\|\,T^2 + 1}, \tag{32}$$

$$s_{\mathrm{eff}} = \frac{\int \tau^2\, dt}{20\,T}, \tag{33}$$

and the overshoot score is built as a deliberately non-smooth piecewise-quadratic function of the measured overshoot,

$$s_{\mathrm{OS}} = 0 \quad \text{if OS} \le 15\%, \tag{34}$$

$$s_{\mathrm{OS}} = (\mathrm{OS} - 15)^2/150 \quad \text{if } 15\% < \mathrm{OS} \le 30\%, \tag{35}$$

$$s_{\mathrm{OS}} = 2.0 + (\mathrm{OS} - 30)^2/100 \quad \text{if OS} > 30\%, \tag{36}$$

$$s_{\mathrm{ST}} = \mathrm{ST}/T, \qquad s_{\mathrm{SSE}} = \mathrm{SSE}^2/80, \tag{37}$$

where $\|r\|$ is the root-mean-square (RMS) magnitude of the reference trajectory and T the simulation horizon. The per-scenario fitness is the weighted sum

$$J_s = s_{\mathrm{ITAE}} + 0.15\, s_{\mathrm{eff}} + s_{\mathrm{OS}} + 0.4\, s_{\mathrm{ST}} + 0.7\, s_{\mathrm{SSE}}, \tag{38}$$

and the plant-level fitness is the weighted sum across the five scenarios,

$$J(x; \theta_k) = \sum_s w_s\, J_s(x; \theta_k). \tag{39}$$

Combined with the $K = 4$ Monte Carlo plant average of (30), this gives the stochastic fitness function used in all optimisation runs. The piecewise-quadratic overshoot penalty with the 30% discontinuity, together with the squared

steady-state-error penalty, produces a benchmark landscape whose multi-modal and noisy character is a deliberate reflection of features characteristic of robust controller tuning tasks. Deployed robotic systems are full of these features — hard overshoot limits derived from safety margins, quadratic penalties on steady-state error arising from energy-minimisation objectives, and stochastic plant uncertainty from thermal drift, payload variation, and manufacturing tolerances [48], [58].

Twelve algorithms participate: FHO, HO, MSHO, LSHADE, SFOA, DhO, OOA, DRA, TLBO, GWO, COA, and PSO, covering the parent algorithm and its published variant, the adaptive-DE representative, the 2025–2026 generation, and classical baselines. All algorithms share a fixed total budget of MaxFEs $=$ 6000 evaluations at a common population size $N = 40$, enforced by the same per-call evaluation recorder (each algorithm is stopped exactly at 6000 evaluations regardless of its per-iteration cost); unlike the CEC suites, where the dimension-scaled DE/ES population schedules are used, this task adopts a common $N = 40$ because its evaluation budget, constrained by the cost of the closed-loop simulations, is far too small for those schedules (LSHADE's $N_{\text{init}} = 18D = 360$ would barely begin to contract within 6000 evaluations) to be meaningful. The budget is deliberately small: each evaluation costs $K = 4$ full 3-second closed-loop simulations across five scenarios, so 6000 evaluations correspond to 120 000 simulated rollouts per run, reflecting the wall-clock reality of simulation-based tuning.

Twenty independent runs are collected per algorithm. Evaluation then separates training from testing: the controller returned by each run is re-scored on $K = 20$ *fresh* perturbed plants drawn from a run-paired seed (identical across algorithms), so that all algorithms face exactly the same 20 unseen plants, and all statistics below refer to this paired out-of-sample test score.

The 20-run result matrix is analysed with the Friedman rank test, two-sided Wilcoxon rank-sum (Mann–Whitney U) tests of FHO against each competitor on the out-of-sample test scores (a conservative choice that does not exploit the run-pairing; the paired signed-rank test gives the same qualitative picture), signed Cliff's δ [39], [41] (negative favouring FHO, Section II-C), 95% t-confidence intervals, and the Nemenyi critical-distance diagram [37], [38].

2) OVERALL RESULTS

Table 15 reports the out-of-sample summary statistics. TLBO attains the best Friedman rank (3.05) and, with a standard deviation of only 7.7, the lowest run-to-run dispersion of any algorithm in the pool; a tight four-algorithm group follows (DhO 4.55, LSHADE 4.70, GWO 4.75, FHO 4.80), and no member of this group, including TLBO, is statistically distinguishable from FHO (Wilcoxon $p \geq 0.50$; $|\delta| \leq 0.13$, negligible by the thresholds of [41]). What does distinguish FHO within that runner-up group is dispersion: its test-score spread (50.6 ± 53.7, 95% CI [25.4, 75.7]) is the narrowest, against LSHADE's 74.6 ± 143.9, DhO's 137.6 ± 333.3, and GWO's 318.3 ± 768.5 — the latter three all carry catastrophic-failure tails (per-run worst scores over the 20 test runs of 678, 1260, and 3049, respectively, versus 261 for FHO). In robust-tuning practice, where a single deployed controller is drawn from one optimisation run, this tail behaviour is the deciding factor.

TABLE 15. Out-of-sample test fitness ($K = 20$ unseen paired plants) across 20 independent runs: summary statistics, Friedman rank, Wilcoxon p versus FHO, and signed Cliff's δ (negative favours FHO). The p column lists the raw two-sided Wilcoxon value; stars denote significance after Holm correction across the eleven-competitor family (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$). The 95% CIs are symmetric Student-t intervals; because the cost is non-negative and several distributions are right-skewed, a negative lower bound should be read as 0.

Algorithm	Mean	Std	95% CI	Best	Friedman	p vs. FHO	Cliff's δ
TLBO	29.3	7.7	[25.6, 32.9]	20.6	3.05	0.508	+0.125
DhO	137.6	333.3	[-18.4, 293.6]	19.9	4.55	0.617	+0.095
LSHADE	74.6	143.9	[7.3, 141.9]	18.2	4.70	0.818	-0.045
GWO	318.3	768.5	[-41.4, 677.9]	18.4	4.75	0.882	-0.030
FHO	50.6	53.7	[25.4, 75.7]	19.0	4.80	—	—
HO	52.8	26.3	[40.5, 65.1]	26.7	6.20	0.072	-0.335
PSO	70.6	56.9	[44.0, 97.2]	20.6	6.40	0.114	-0.295
COA	58.0	9.3	[53.7, 62.4]	40.8	6.90	0.060	-0.350
SFOA	62.4	12.5	[56.6, 68.3]	43.8	7.50	0.019	-0.435
DRA	99.8	90.0	[57.7, 142.0]	26.2	9.00	0.000***	-0.730
MSHO	186.5	442.5	[-20.6, 393.7]	29.6	9.40	0.000**	-0.655
OOA	135.5	49.4	[112.4, 158.6]	72.3	10.75	0.000***	-0.900

Below the tie group, FHO significantly outperforms DRA, MSHO, and OOA (Holm-adjusted $p < 0.001$ for OOA and DRA, $p = 0.004$ for MSHO; $|\delta| \geq 0.65$); its advantage over SFOA (raw $p = 0.019$, $|\delta| = 0.44$) does not survive Holm correction across the eleven-competitor family (adjusted $p = 0.15$) and is therefore reported as a rank lead rather than a significant win. FHO consistently leads its parent HO (6.20, raw $p = 0.072$, $\delta = -0.34$, medium effect), the HO comparison being the one this paper's mechanisms are accountable for. Fig. 23 shows the test-score distributions (box plots with whiskers and medians), Fig. 24 the training convergence, and Fig. 25 the Nemenyi critical-distance diagram. Because the Nemenyi critical distance (3.73, at 12 algorithms and 20 runs) is a conservative post-hoc bound, FHO's indistinguishable bar also spans HO, PSO, COA, and SFOA (ranks 6.20–7.50); the sharper pairwise Wilcoxon test isolates the five-member tie reported above.

The result shows that cross-domain consistency holds only conditionally. The deterministic-benchmark champion LSHADE does *not* dominate here; its stochastic-fitness standard deviation is 2.7× FHO's, while TLBO, merely fourth on CEC, leads the FOPID ranking; landscape match matters more than aggregate rank, exactly as the no-free-lunch principle predicts [32]. Of the three strongest CEC algorithms, FHO also sits in the FOPID top clique, with lower run-to-run variance than the other state-of-the-art entrant there, LSHADE (whose standard deviation is 2.7× FHO's); the rank-leader TLBO is lower-variance still, and CMA-ES was not part of the FOPID pool. The elite re-evaluation step that is provably neutral on deterministic

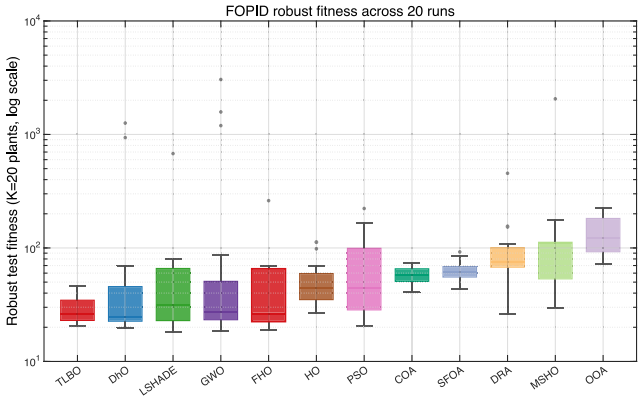

FIGURE 23. Out-of-sample test-fitness distributions across 20 runs per algorithm (box plots with whiskers; log y-axis).

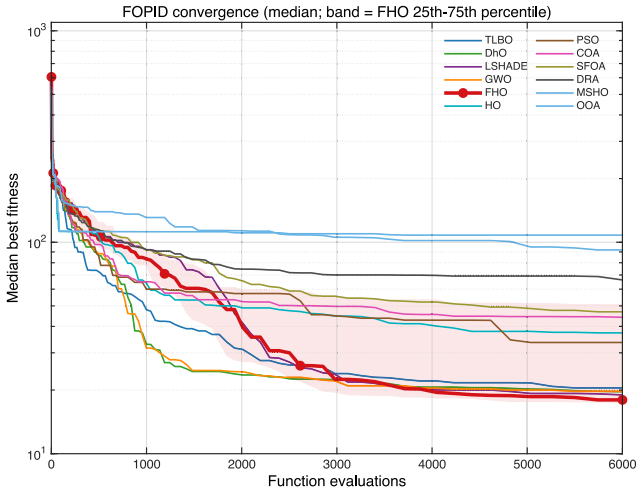

FIGURE 24. Training convergence: median best fitness versus evaluations (interquartile range (IQR) band for FHO; 6000-evaluation budget).

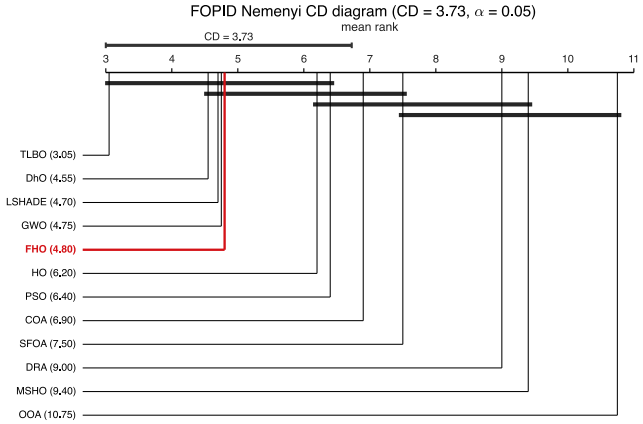

FIGURE 25. Nemenyi critical-distance diagram on the FOPID out-of-sample test scores (CD = 3.73, $\alpha = 0.05$). Because the critical distance is a conservative post-hoc bound, FHO's statistically-indistinguishable set also spans HO, PSO, COA, and SFOA (ranks 6.20–7.50); the tighter five-member group (TLBO, DhO, LSHADE, GWO, FHO) is the one isolated by the pairwise Wilcoxon test ($p \geq 0.50$).

fitness (Section IV-B1) proves useful on this stochastic task: the variants lacking it rank below FHO (HO 6.20, MSHO 9.40 versus FHO 4.80), with MSHO (9.40) falling well outside FHO's Nemenyi-indistinguishable set entirely.

TABLE 16. Per-scenario mean ITAE (nominal-plant replay of each run's returned controller) across the 12 algorithms.

Algorithm	Step	FastSine	PickPlace	Disturb	Noise
FHO	34.4	85.0	128.4	34.3	30.1
HO	40.9	85.3	122.9	38.7	33.0
MSHO	55.8	83.3	106.9	65.8	56.9
LSHADE	28.6	82.9	110.8	35.6	33.1
SFOA	68.0	116.5	89.8	91.8	86.1
DhO	39.3	89.1	134.4	36.5	31.7
OOA	59.1	63.4	51.7	19.2	24.2
DRA	48.8	82.6	108.8	46.8	43.2
TLBO	44.9	94.0	136.7	40.0	41.5
GWO	45.0	94.8	106.4	41.1	36.6
COA	61.6	91.8	108.4	63.7	69.7
PSO	43.8	85.6	101.0	32.4	30.3

3) PER-SCENARIO RESULTS AND TIME-DOMAIN VALIDATION

Per-scenario ITAE values appear in Table 16. The optimisation target of this benchmark is the aggregate robust fitness across the perturbed plants (30), not the best per-scenario tracking on the nominal plant. The single-scenario ranking is therefore a diagnostic view, not a primary result. The table shows FHO in the top three on the Step, Disturb, and Noise scenarios, mid-pack on FastSine, and among the weakest on the PickPlace trajectory, where the overall Friedman leader TLBO is in fact last. The pattern illustrates the aggregate-versus-specialist trade-off: scenario specialists (e.g., OOA, which wins FastSine, PickPlace, Disturb, and Noise outright yet ranks last overall at 10.75) achieve isolated wins at the cost of catastrophic failures elsewhere, while the top-clique algorithms spread their budget across all five weighted scenarios.

The parameter-space heatmap of Fig. 26 shows the fractional orders being used non-trivially: λ tends toward the upper half of its range in several of the top-ranked tunings, while both λ and μ vary widely (0.3–1.0) across algorithms and joints, and no two algorithms select the same parameter pattern. FHO's best-run parameter vector is composed mostly of interior values, avoiding the wholesale boundary-pinning failure mode reported for swarms on box-constrained problems [59], although isolated components of most tunings — FHO's included — do sit at a bound.

Finally, the controllers themselves are validated in the time domain. Fig. 27 overlays the closed-loop step responses of the best FHO, TLBO, LSHADE, and HO controllers on the nominal plant: the FHO tuning tracks J1 with near-zero overshoot and settles all four joints without instability; on J2 and J3 it carries moderate overshoot and a residual steady-state offset, a behaviour shared by all four controllers and traceable to the uncompensated gravity and Coulomb-friction terms of the rigid-body model, while remaining visibly

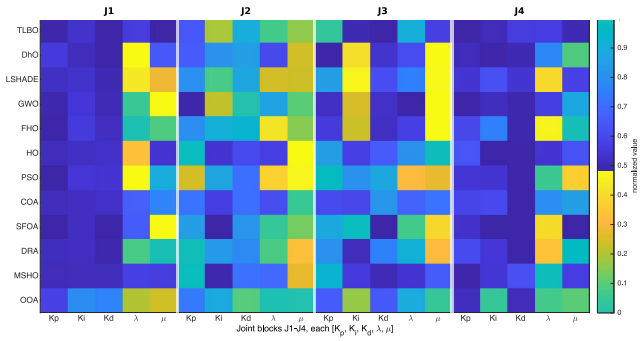

FIGURE 26. Best-run FOPID parameters per algorithm, normalised to [0, 1] search bounds (white separators delimit joints J1–J4).

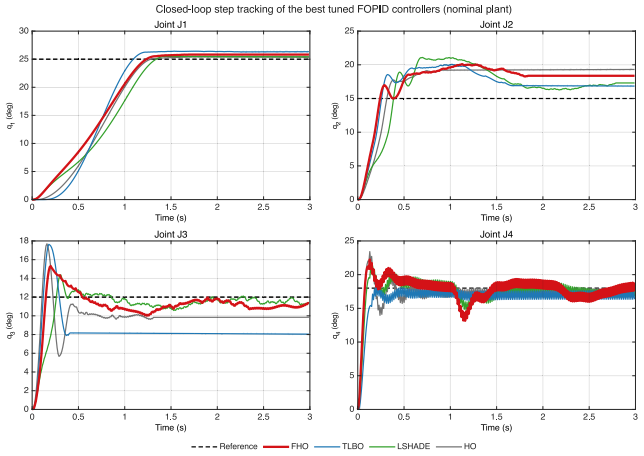

FIGURE 27. Closed-loop step responses (J1–J4) of the best stored controllers: FHO versus TLBO, LSHADE, and parent HO on the nominal plant.

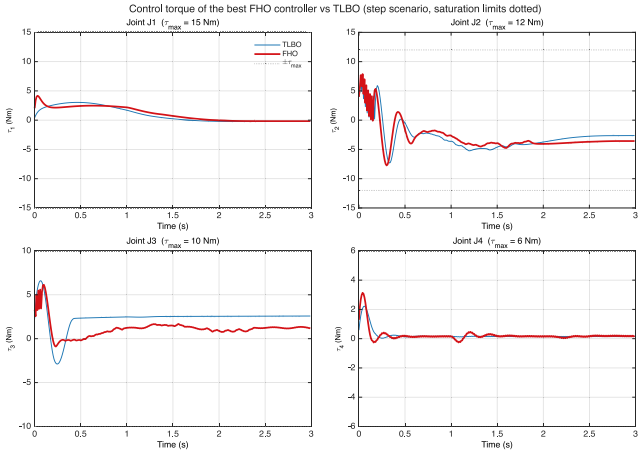

FIGURE 28. Commanded joint torques of the two rank-leading controllers (FHO, TLBO) for the step scenario, with actuator limits $\pm\tau_{\max}$ marked: no saturation.

faster-settling than the parent HO. Fig. 28 shows the torque commands of the two rank-leading controllers (FHO and TLBO) staying inside the actuator limits $\pm\tau_{\max}$ (no saturation).

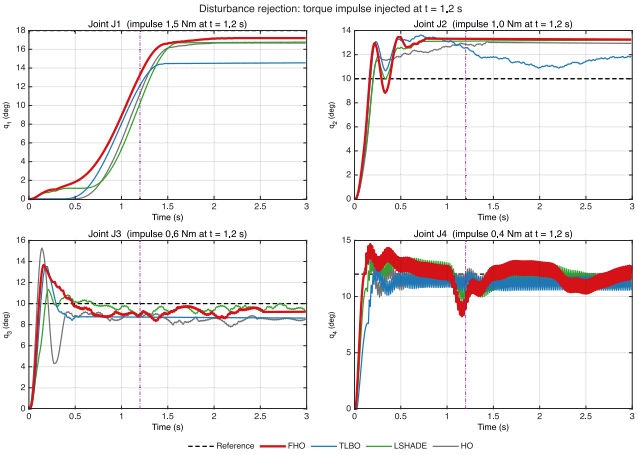

FIGURE 29. Disturbance rejection: response to the impulsive joint torque injected at $t = 1.2$ s (dashed line).

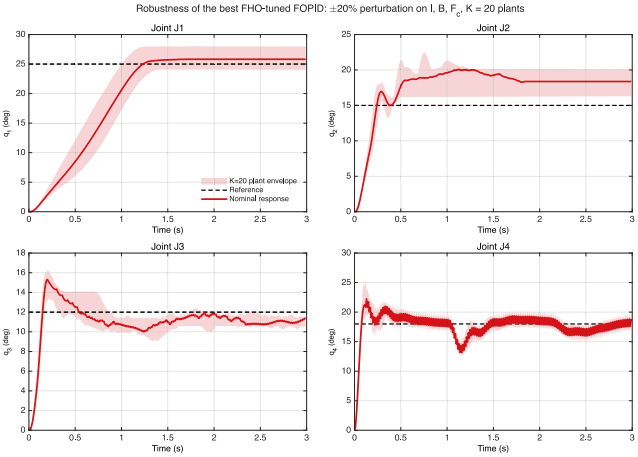

FIGURE 30. Robustness envelope: FHO closed-loop step response across $K = 20$ unseen ±20% perturbed plants (band = min–max, line = nominal).

Fig. 29 demonstrates disturbance rejection: after the impulsive torque at $t = 1.2$ s the FHO controller re-settles within tenths of a second on every joint. On J4 the dominant transient near $t \approx 1.1$ s is inter-joint coupling from J1 arriving at its reference — the same dip appears in the disturbance-free step scenario of Fig. 27 and should not be read as the impulse response; on the impulse itself the four controllers behave comparably. Fig. 30 closes the loop on robustness: across the $K = 20$ unseen perturbed plants, the FHO response stays inside a tight envelope around the nominal trajectory — the time-domain counterpart of the low test-score variance of Table 15.

V. DISCUSSION AND CONCLUSION

A. DISCUSSION

The ablation results of Section IV-B1 attribute FHO's aggregate advantage to the combined action of its three rank-driven mechanisms on complementary failure modes. FRGM contributes the largest rank loss when removed

(+2.07) because it is the only operator whose perturbation is decoupled from the population's current geometry. Its diversity mask (12) targets exactly the dimensions where collapse has occurred, which also best explains the strengthening relative position of the algorithm as the dimension grows (third at $D \leq 50$, rising to second at $D = 100$ by overtaking CMA-ES). FRDS contributes a moderate, never-harmful refinement (+0.41 when removed, with zero functions on which its removal helps); FCDL contributes +0.62 by keeping the bottom half productive. The diversity and exploration–exploitation trajectories of Section IV-B3 show the intended division of labour in operation: rapid early concentration onto the optimum basin with a small FRGM-maintained diversity floor, in contrast to the parent HO, whose population stays dispersed without converging.

A central claim of this work is that benchmark performance transfers to the engineering task; the evidence supports it in a conditional form. FHO is third of 23 on the deterministic CEC aggregate—behind the two state-of-the-art non-swarm optimisers LSHADE and CMA-ES—and inside the top statistical clique of 12 on the stochastic FOPID task. Among the algorithms tested on both, FHO is the exception: it pairs a competitive deterministic-benchmark ranking with a top FOPID-clique placement, whereas LSHADE drops into the high-variance runner-up tie on FOPID and TLBO leads FOPID but ranks fourth on CEC. Read together, the results support a measured recommendation: for rugged deterministic problems LSHADE is the reference (with CMA-ES competitive, especially on low-dimensional ill-conditioned suites); FHO's value is as the strongest swarm- and social-inspired optimiser in the pool and a clear improvement over both Hippopotamus-family algorithms benchmarked here (the parent HO and the published variant MSHO); FHO also remains competitive on the noisy, simulation-driven tuning task where run-to-run reliability matters, consistent with the no-free-lunch theorem [32].

B. LIMITATIONS

The study has several limitations. First, computational overhead: measured at a common $N = 40$ to isolate per-evaluation bookkeeping, FHO carries the highest per-iteration overhead in the pool — 7.86 s per 3×10^5-evaluation run versus 4.74 s for HO and 0.79 s for TLBO (Table 9) — owing to the per-iteration sort, three peer-sampling passes, and per-dimension diversity statistics. The fair-budget protocol neutralises this in evaluation-limited settings, and the overhead is negligible whenever the objective dominates (one FOPID evaluation costs $\sim 100\times$ FHO's per-evaluation bookkeeping), but for cheap objectives on wall-clock budgets the lighter swarms are preferable; FHO's per-individual updates are embarrassingly parallel and a vectorised or GPU implementation is an obvious engineering follow-up. Second, FHO does not match the strongest baselines: it trails LSHADE on every CEC suite and at every tested dimension, and trails CMA-ES on every suite except at $D = 100$, where FHO overtakes it; and it is statistically tied with, rather than ahead of, TLBO, DhO, LSHADE, and GWO on the FOPID task. The claims this paper defends are correspondingly narrow: FHO is the strongest swarm- and social-inspired optimiser in the pool, it overtakes CMA-ES at $D = 100$, it improves clearly over both Hippopotamus-family algorithms benchmarked here (the parent HO and the published variant MSHO), and, among the high-variance members of the FOPID top clique (LSHADE, DhO, GWO), it has the lowest variance, though the rank-leader TLBO is both better-ranked and lower-variance. Third, hyperparameter coverage: the sensitivity study perturbs the five behaviour-shaping parameters one at a time; a full orthogonal design over all twelve named parameters, and across suites, is left to future work.

On the manipulator side, the Coriolis and mass matrices in the robust-FOPID benchmark of Section IV-C retain CAD-estimated link inertia tensors and centres of mass; a rigorous inertial identification along the lines of [33] and [60] would tighten the model. All Section IV-D results come from simulation of the identified model; closed-loop deployment of the FHO-tuned controller on the physical 4-DOF arm is the natural follow-up.

The stability analysis of Section III-E establishes uniformly ultimate boundedness of the closed-loop tracking error rather than asymptotic stability, and its sufficient condition (E4) is known to be conservative. The empirical decay-rate estimate $\gamma = 2.69\ \mathrm{s}^{-1}$ is obtained by log-space regression on the smoothed nominal rollout of Fig. 2 for the step scenario; a reference-independent analytical derivation of $(\gamma, \varphi, B_\infty)$ directly from the manipulator parameters and the plant-uncertainty bound is left to future work.

C. CONCLUSION AND FUTURE WORK

This article has introduced the Fitness-Ranked Hippopotamus Optimization algorithm, which replaces the update pipeline of the parent HO with three rank-driven mechanisms and a per-iteration elite-preservation step at a cost of $2.5N + 1$ function evaluations per iteration — less than HO's $3N$ and comparable to TLBO's $2N$ at common population sizes. A Lyapunov-based stability analysis (Section III-E) additionally certifies that every FOPID tuning evaluated by FHO lies within a stable-by-construction parameter set, satisfying a Mittag–Leffler uniform-ultimate-boundedness sufficient condition.

Under the suites' official evaluation budgets (MaxFEs = $10^4 \times D$ for CEC2017, 10^6 for CEC2022; 30 runs, 23 algorithms including LSHADE, CMA-ES, MSHO, and four 2025–2026 optimisers), FHO attains Friedman mean rank 4.83 on the 51-function aggregate, third of 23 behind LSHADE (1.57) and CMA-ES (3.50) but first among all swarm- and social-inspired algorithms in the pool, ahead of both Hippopotamus-family algorithms benchmarked here

(the parent HO and the published variant MSHO); at $D = 100$ it ranks second (3.97), overtaking CMA-ES but not LSHADE. It beats its parent HO on 27–28 of 29 CEC2017 functions at every tested dimension. The controlled ablation quantifies each mechanism's contribution (FRGM +2.07, FCDL +0.62, FRDS +0.41 Friedman-rank units), and the ±40% sensitivity scan shows no dependence on fine-tuned constants. On the 20-dimensional robust-FOPID task against 11 competitors, FHO finishes inside the top statistical clique on unseen perturbed plants ($p \geq 0.50$ versus the four other members of the clique) with the lowest variance among the high-variance members of that clique (the rank-leader TLBO attains both a better rank and a lower variance) and a non-significant improvement over HO ($p = 0.072$). The time-domain validation confirms stable step tracking, unsaturated torques, fast disturbance recovery, and a tight robustness envelope. Its main costs are the highest per-iteration overhead among the fixed-population algorithms (1.7× HO at common N) and a consistent deficit to the two state-of-the-art non-swarm optimisers, LSHADE and CMA-ES, on the deterministic suites.

Future work includes a vectorised/GPU implementation to remove the runtime overhead; an orthogonal multi-suite design over all twelve hyperparameters toward Pareto-optimal defaults; and hardware deployment of the FHO-tuned FOPID controllers on the physical 4-DOF arm, including online adaptive re-identification under payload variation and extension to higher-DOF manipulators.

REFERENCES

[1] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proc. IEEE Int. Conf. Neural Netw.*, vol. 4, Mar. 1995, pp. 1942–1948.

[2] S. Mirjalili, S. M. Mirjalili, and A. Lewis, "Grey wolf optimizer," *Adv. Eng. Softw.*, vol. 69, pp. 46–61, Feb. 2014.

[3] R. V. Rao, V. J. Savsani, and D. P. Vakharia, "Teaching–learning-based optimization: A novel method for constrained mechanical design optimization problems," *Computer-Aided Design*, vol. 43, no. 3, pp. 303–315, Mar. 2011.

[4] S. Mirjalili and A. Lewis, "The whale optimization algorithm," *Adv. Eng. Softw.*, vol. 95, pp. 51–67, Mar. 2016.

[5] S. Mirjalili, A. H. Gandomi, S. Z. Mirjalili, S. Saremi, H. Faris, and S. M. Mirjalili, "Salp swarm algorithm: A bio-inspired optimizer for engineering design problems," *Adv. Eng. Softw.*, vol. 114, pp. 163–191, Dec. 2017.

[6] A. A. Heidari, S. Mirjalili, H. Faris, I. Aljarah, M. Mafarja, and H. Chen, "Harris hawks optimization: Algorithm and applications," *Future Gener. Comput. Syst.*, vol. 97, pp. 849–872, Apr. 2019.

[7] L. Abualigah, D. Yousri, M. A. Elaziz, A. A. Ewees, M. A. A. Al-Qaness, and A. H. Gandomi, "Aquila optimizer: A novel meta-heuristic optimization algorithm," *Comput. Ind. Eng.*, vol. 157, Jul. 2021, Art. no. 107250.

[8] A. Faramarzi, M. Heidarinejad, B. Stephens, and S. Mirjalili, "Equilibrium optimizer: A novel optimization algorithm," *Knowl.-Based Syst.*, vol. 191, Mar. 2020, Art. no. 105190.

[9] J. Pierezan and L. Dos Santos Coelho, "Coyote optimization algorithm: A new metaheuristic for global optimization problems," in *Proc. IEEE Congr. Evol. Comput. (CEC)*, Jul. 2018, pp. 1–8.

[10] J. Xue and B. Shen, "Dung beetle optimizer: A new meta-heuristic algorithm for global optimization," *J. Supercomput.*, vol. 79, no. 7, pp. 7305–7336, May 2023.

[11] H. Su, D. Zhao, A. A. Heidari, L. Liu, X. Zhang, M. Mafarja, and H. Chen, "RIME: A physics-based optimization," *Neurocomputing*, vol. 532, pp. 183–214, May 2023.

[12] L. Deng and S. Liu, "Snow ablation optimizer," *Expert Syst. Appl.*, vol. 225, Jun. 2023, Art. no. 120069.

[13] A. M. Fathollahi-Fard, M. Hajiaghaei-Keshteli, and R. Tavakkoli-Moghaddam, "Red deer algorithm (RDA): A new nature-inspired meta-heuristic algorithm for solving engineering optimization problems," *Soft Comput.*, vol. 24, no. 19, pp. 14637–14665, 2020.

[14] M. Abdel-Basset, R. Mohamed, and M. Abouhawwash, "Crested porcupine optimizer," *Knowl.-Based Syst.*, vol. 284, Nov. 2024, Art. no. 111257.

[15] M. H. Amiri, N. M. Hashjin, M. Montazeri, S. Mirjalili, and N. Khodadadi, "Hippopotamus optimization algorithm," *Sci. Rep.*, vol. 14, p. 5032, Jan. 2024.

[16] X. Yuan, L. Yao, T. Zhang, Q. Chen, and G. Li, "An effective method for solving feature selection problems based on multi-strategy hippo optimization algorithm," *J. Comput. Design Eng.*, vol. 12, no. 8, pp. 193–251, Aug. 2025, doi: 10.1093/jcde/qwaf073.

[17] Y. Shi and R. Eberhart, "A modified particle swarm optimizer," in *Proc. IEEE Int. Conf. Evol. Comput.*, Aug. 1998, pp. 69–73.

[18] R. Storn and K. Price, "Differential evolution—A simple and efficient heuristic for global optimization over continuous spaces," *J. Global Optim.*, vol. 11, no. 4, pp. 341–359, Dec. 1997.

[19] X. Yao, Y. Liu, and G. Lin, "Evolutionary programming made faster," *IEEE Trans. Evol. Comput.*, vol. 3, no. 2, pp. 82–102, Jul. 1999.

[20] N. H. Awad, M. Z. Ali, P. N. Suganthan, J. J. Liang, and B. Y. Qu, "Problem definitions and evaluation criteria for the CEC 2017 special session and competition on single objective bound constrained real-parameter numerical optimization," Nanyang Technol. Univ., Singapore, Tech. Rep., Oct. 2016.

[21] A. Kumar, K. V. Price, A. W. Mohamed, A. A. Hadi, and P. N. Suganthan, "Problem definitions and evaluation criteria for the CEC 2022 special session and competition on single objective bound constrained numerical optimization," Nanyang Technol. Univ., Singapore, Tech. Rep., Dec. 2021.

[22] R. Tanabe and A. S. Fukunaga, "Improving the search performance of SHADE using linear population size reduction," in *Proc. IEEE Congr. Evol. Comput. (CEC)*, Jul. 2014, pp. 1658–1665.

[23] N. Hansen and A. Ostermeier, "Completely derandomized self-adaptation in evolution strategies," *Evol. Comput.*, vol. 9, no. 2, pp. 159–195, Jun. 2001.

[24] C. Zhong, G. Li, Z. Meng, H. Li, A. R. Yildiz, and S. Mirjalili, "Starfish optimization algorithm (SFOA): A bio-inspired metaheuristic algorithm for global optimization compared with 100 optimizers," *Neural Comput. Appl.*, vol. 37, no. 5, pp. 3641–3683, Feb. 2025, doi: 10.1007/s00521-024-10694-1.

[25] B. O. Mohammed, H. S. Aghdasi, and P. Salehpour, "Dhole optimization algorithm: A new metaheuristic algorithm for solving optimization problems," *Cluster Comput.*, vol. 28, no. 7, p. 430, Sep. 2025, doi: 10.1007/s10586-024-05005-1.

[26] K. Wang, L. Abualigah, A. Smerat, J. Li, X. Wu, H. Liu, Z. Shao, and S. Mirjalili, "A nature recoil mechanism-based octopus optimization algorithm for solving the global and constraint optimization from engineering structural design problems," *J. Comput. Design Eng.*, vol. 13, no. 3, pp. 123–232, Feb. 2026, doi: 10.1093/jcde/qwaf139.

[27] A. T. Mozhdehi, N. Khodadadi, M. Aboutalebi, E.-S.-M. El-Kenawy, A. G. Hussien, W. Zhao, M. H. Nadimi-Shahraki, and S. Mirjalili, "Divine religions algorithm: A novel social-inspired metaheuristic algorithm for engineering and continuous optimization problems," *Cluster Comput.*, vol. 28, no. 4, p. 253, Aug. 2025, doi: 10.1007/s10586-024-04954-x.

[28] O. Rodríguez-Abreo, J. Rodríguez-Reséndiz, J. M. Álvarez-Alvarado, and A. García-Cerezo, "Metaheuristic parameter identification of motors using dynamic response relations," *Sensors*, vol. 22, no. 11, p. 4050, May 2022, doi: 10.3390/s22114050.

[29] R. Castro-Medina, M. G. Villarreal-Cervantes, L. G. Corona-Ramírez, G. Flores-Caballero, A. Rodríguez-Molina, R. Silva-Ortigoza, V. D. Cuervo-Pinto, and A. A. Palma-Huerta, "Model parameterization of robotic systems through the bio-inspired optimization," *PLoS ONE*, vol. 20, no. 6, Jun. 2025, Art. no. e0325168, doi: 10.1371/journal.pone.0325168.

[30] J. Zhang and A. C. Sanderson, "JADE: Adaptive differential evolution with optional external archive," *IEEE Trans. Evol. Comput.*, vol. 13, no. 5, pp. 945–958, Oct. 2009.

[31] R. Tanabe and A. Fukunaga, "Success-history based parameter adaptation for differential evolution," in *Proc. IEEE Congr. Evol. Comput.*, Jun. 2013, pp. 71–78.
[32] D. H. Wolpert and W. G. Macready, "No free lunch theorems for optimization," *IEEE Trans. Evol. Comput.*, vol. 1, no. 1, pp. 67–82, Apr. 1997.
[33] M. Gautier and W. Khalil, "Direct calculation of minimum set of inertial parameters of serial robots," *IEEE Trans. Robot. Autom.*, vol. 6, no. 3, pp. 368–373, Jun. 1990.
[34] H. R. Tizhoosh, "Opposition-based learning: A new scheme for machine intelligence," in *Proc. Int. Conf. Comput. Intell. Model., Control Autom. Int. Conf. Intell. Agents, Web Technol. Internet Commerce (CIMCA-IAWTIC06)*, vol. 1, Jul. 2005, pp. 695–701.
[35] K. V. Price, N. H. Awad, M. Z. Ali, and P. N. Suganthan, "Problem definitions and evaluation criteria for the 100-digit challenge special session and competition on single objective numerical optimization," Nanyang Technol. Univ., Singapore, Tech. Rep., Nov. 2018.
[36] M. Friedman, "The use of ranks to avoid the assumption of normality implicit in the analysis of variance," *J. Amer. Stat. Assoc.*, vol. 32, no. 200, pp. 675–701, Dec. 1937.
[37] J. Demšar, "Statistical comparisons of classifiers over multiple data sets," *J. Mach. Learn. Res.*, vol. 7, pp. 1–30, Apr. 2006.
[38] P. B. Nemenyi, "Distribution-free multiple comparisons," Ph.D. dissertation, Dept. Math., Princeton Univ., Princeton, NJ, USA, 1963.
[39] N. Cliff, "Dominance statistics: Ordinal analyses to answer ordinal questions," *Psychol. Bull.*, vol. 114, no. 3, pp. 494–509, Nov. 1993.
[40] S. García, A. Fernández, J. Luengo, and F. Herrera, "Advanced non-parametric tests for multiple comparisons in the design of experiments in computational intelligence and data mining: Experimental analysis of power," *Inf. Sci.*, vol. 180, no. 10, pp. 2044–2064, May 2010.
[41] A. Vargha and H. D. Delaney, "A critique and improvement of the CL common language effect size statistics of McGraw and wong," *J. Educ. Behav. Statist.*, vol. 25, no. 2, pp. 101–132, Jun. 2000.
[42] I. Podlubny, "Fractional-order systems and PIλDμ controllers," *IEEE Trans. Autom. Control*, vol. 44, no. 1, pp. 208–214, Jan. 1999.
[43] A. Tepljakov, E. Petlenkov, and J. Belikov, "FOMCON: Fractional-order modeling and control toolbox for MATLAB," in *Proc. 18th Int. Conf. Mixed Design Integr. Circuits Syst. - MIXDES*, Jun. 2011, pp. 684–689.
[44] I. Petráš, *Fractional-Order Nonlinear Systems*. Beijing, China: Higher Education Press, 2011.
[45] A. Visioli, *Practical PID Control*. London, U.K.: Springer, 2006.
[46] K. J. Åström and T. Hägglund, *Advanced PID Control*. Research Triangle Park, NC, USA: ISA, 2006.
[47] C. A. Monje, Y. Chen, B. M. Vinagre, D. Xue, and V. Feliu, *Fractional-Order Systems and Controls*. London, U.K.: Springer, 2010.
[48] M. Gautier and P. Poignet, "Extended Kalman filtering and weighted least squares dynamic identification of robot," *Control Eng. Pract.*, vol. 9, no. 12, pp. 1361–1372, 2001.
[49] B. Siciliano, L. Sciavicco, L. Villani, and G. Oriolo, *Robotics: Modelling, Planning and Control*. London, U.K.: Springer, 2009.
[50] N. Aguila-Camacho, M. A. Duarte-Mermoud, and J. A. Gallegos, "Lyapunov functions for fractional order systems," *Commun. Nonlinear Sci. Numer. Simul.*, vol. 19, no. 9, pp. 2951–2957, 2014.
[51] A. Auger and N. Hansen, "A restart CMA evolution strategy with increasing population size," in *Proc. IEEE Congr. Evol. Comput.*, vol. 2, Oct. 2005, pp. 1769–1776.
[52] H. D. Taghirad and P. R. Bélanger, "Modeling and parameter identification of harmonic drive systems," *J. Dyn. Syst., Meas., Control*, vol. 120, no. 4, pp. 439–444, Dec. 1998.
[53] L. Le Tien, A. Albu-Schaffer, A. De Luca, and G. Hirzinger, "Friction observer and compensation for control of robots with joint torque measurement," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Sep. 2008, pp. 3789–3795.
[54] Y. Han, J. Wu, C. Liu, and Z. Xiong, "Static model analysis and identification for serial articulated manipulators," *Robot. Computer-Integrated Manuf.*, vol. 57, pp. 155–165, Jun. 2019, doi: 10.1016/j.rcim.2018.11.010.
[55] Z.-L. Gaing, "A particle swarm optimization approach for optimum design of PID controller in AVR system," *IEEE Trans. Energy Convers.*, vol. 19, no. 2, pp. 384–391, Jun. 2004, doi: 10.1109/TEC.2003.821821.
[56] J. Doyle, "Analysis of feedback systems with structured uncertainties," *IEE Proc. D Control Theory Appl.*, vol. 129, no. 6, pp. 242–250, 1982.
[57] T. Kröger, "On-line trajectory generation in robotic systems," in *Springer Tracts in Advanced Robotics*, vol. 58. Berlin, Germany: Springer, 2010.
[58] K. Zhou and J. C. Doyle, *Essentials of Robust Control*. Upper Saddle River, NJ, USA: Prentice-Hall, 1998.
[59] S. Helwig and R. Wanka, "Particle swarm optimization in high-dimensional bounded search spaces," in *Proc. IEEE Swarm Intell. Symp.*, Apr. 2007, pp. 198–205.
[60] J. Swevers, W. Verdonck, and J. De Schutter, "Dynamic model identification for industrial robots," *IEEE Control Syst. Mag.*, vol. 27, no. 5, pp. 58–71, May 2007.
[61] B. Morales-Castañeda, D. Zaldívar, E. Cuevas, F. Fausto, and A. Rodríguez, "A better balance in metaheuristic algorithms: Does it exist?" *Swarm Evol. Comput.*, vol. 54, May 2020, Art. no. 100671, doi: 10.1016/j.swevo.2020.100671.
[62] M. Črepinšek, S.-H. Liu, and M. Mernik, "Exploration and exploitation in evolutionary algorithms: A survey," *ACM Comput. Surveys*, vol. 45, no. 3, pp. 1–33, Jun. 2013, doi: 10.1145/2480741.2480752.
[63] W. Gong and Z. Cai, "Differential evolution with ranking-based mutation operators," *IEEE Trans. Cybern.*, vol. 43, no. 6, pp. 2066–2081, Dec. 2013, doi: 10.1109/TCYB.2013.2239988.
[64] E. Cuevas, J. Gálvez, and O. Avalos, *Recent Metaheuristics Algorithms for Parameter Identification*. Cham, Switzerland: Springer, 2020, doi: 10.1007/978-3-030-28917-1.
[65] P.-W. Hsueh and M. F. Fazdi, "Parameter estimation of permanent magnet DC motor using various particle swarm optimization variants," *Array*, vol. 30, Jul. 2026, Art. no. 100823, doi: 10.1016/j.array.2026.100823.
[66] H. Guzowski, M. Smolka, A. Byrski, L. Pekar, Z. K. Oplatkova, R. Senkerik, R. Matusu, and F. Gazdos, "Effective parametric optimization of heating-cooling process with optimum near the domain border," in *Proc. IEEE 11th Int. Conf. Intell. Syst. (IS)*, Oct. 2022, pp. 1–6, doi: 10.1109/IS57118.2022.10019630.
[67] H. Guzowski, R. Senkerik, M. Smołka, F. Gazdoš, M. Pálka, L. Pekař, M. Pluhacek, A. Viktorin, T. Kadavy, A. Byrski, Z. K. Oplatková, R. Matušu, and J. Kacprzyk, "Metaheuristic-based model optimization of a steam-filled chamber," *IEEE Access*, vol. 13, pp. 102144–102158, 2025, doi: 10.1109/ACCESS.2025.3574414.

YU ZHANG received the bachelor's degree from Zhengzhou University, China, and the master's degree from Universiti Putra Malaysia (UPM), Malaysia, where he is currently pursuing the Ph.D. degree in robotic and automation engineering. His research interests include optimization algorithms, artificial intelligence, robot control, and simulation.

AZIZAN AS'ARRY received the bachelor's, M.Sc., and Ph.D. degrees from Universiti Teknologi Malaysia (UTM), in 2008, 2009, and 2014, respectively. He is currently a Senior Lecturer with the Department of Mechanical and Manufacturing Engineering, Faculty of Engineering, Universiti Putra Malaysia. He is also the Head of the Instrumentation and Control Systems Laboratory, Department of Mechanical and Manufacturing Engineering. His major fields of study are intelligent control and mechatronics. His research interests include mechatronics, active vibration control, intelligent optimization, systems identification, the IoT automation, and robotics.

HAOHAO MA was born in Tianshui, China. He received the bachelor's degree in mechanical design, manufacturing, and automation from the North University of China, in 2011, and the master's degree in mechanical and electronic engineering from Xidian University, China, in 2014. He is currently pursuing the Ph.D. degree with Universiti Putra Malaysia. He is also a Lecturer with Tianshui Normal University, China. His research interests include computer-aided design and manufacturing, robot control, and simulation.

MOHD KHAIROL ANUAR BIN MOHD ARIFFIN received the Ph.D. degree from The University of Sheffield, U.K. From 2012 to 2015, he was the Coordinator of AUN/SEED-Net Program, contributing to regional academic and research collaboration in engineering. Later, he was the Head of the Department of Mechanical and Manufacturing Engineering, Faculty of Engineering, Universiti Putra Malaysia (UPM), where he was also the Dean of the Faculty of Engineering. Previously, he was also the Honorary Secretary of American Society of Mechanical Engineers (ASME), Malaysian Branch. He is currently the Coordinator of the Master's Program in Manufacturing Engineering, the Head of the Advanced Manufacturing Research Centre (AMREC), and the President of the ModTech Society Malaysia Branch. He is also a Panel Evaluator for the Engineering Accreditation Council (EAC) Malaysia and the Engineering Technology Accreditation Council (ETAC). He is also a Professional Engineer (P.Eng.) registered with the Board of Engineers Malaysia (BEM). He has over 16 years of experience in teaching, research, and consultancy, with research interests in manufacturing engineering, optimization techniques, and additive manufacturing. He has published more than 200 international journal articles and two internationally published books. He has received numerous awards at both national and international levels, including the TRSM 2024 Award.

AZMAH HANIM BINTI MOHAMED ARIFF is currently an Associate Professor with the Department of Mechanical and Manufacturing Engineering, Faculty of Engineering, Universiti Putra Malaysia. She specializes in advanced engineering materials, with a strong focus on lead-free solder materials, surface engineering, and ceramic composites with pore former. She obtained her academic qualifications in mechanical engineering (materials), and she has actively contributed to research and supervision at bachelor's and master's levels. Her research interests include electronic materials, powder metallurgy, corrosion behavior, nanocomposites, and sustainable material development. She is also involved in interdisciplinary initiatives, including the applications of materials science in energy, manufacturing, and emerging technologies.

MOHD NA'IM ABDULLAH received the bachelor's, M.Sc., and Ph.D. degrees in aerospace engineering from Universiti Putra Malaysia (UPM). He is currently a Senior Lecturer with the Department of Aerospace Engineering, Faculty of Engineering, UPM. He is also the Head of the Aerospace Structure Laboratory, Department of Aerospace Engineering. His major fields of study are machine learning, advanced materials engineering, and polymer science. His research interests include infrared thermography for non-destructive testing (NDT), deep learning applications for corrosion detection, intelligent systems, advanced aerospace materials, and sustainable geopolymer-based coatings for fire protection and thermal insulation.