Foundations of Correlated Mutations for Integer Programming

FOGA'25 · Leiden, The Netherlands

Ofer M. **Shir**
ofersh@telhai.ac.il
Tel-Hai College
MIGAL Galilee Research Institute LTD.

Michael **Emmerich**
michael.t.m.emmerich@jyu.fi
JYVÄSKYLÄN YLIOPISTO
UNIVERSITY OF JYVÄSKYLÄ

Full Paper

Source Code

Framework: Heuristics for Integer Programming (IP)

IP's theoretical complexity recently advanced to $(\log(2n))^{O(n)}$ steps [Rothvoss-Reis'24], but heuristics dominate practice. Evolution Strategies are especially attractive solvers:

- Intrinsic **mixed-integer capabilities**
- Well-developed **self-adaptation mechanisms**
- High efficacy in handling **unbounded search spaces** in practice

Notation and Preliminaries

ES Mutation Mechanism:

$$\vec{x}_{\text{NEW}} = \vec{x}_{\text{CURR}} + \vec{z} \in \mathbb{Z}^n$$

$\vec{z}$ drawn from a multivariate distribution.

Discrete Probability:

$$\mathbf{Pr}\{z = k\} = p_k, \quad \sum_k p_k = 1.0$$

Shannon entropy (unpredictability):

$$H := -\sum_{k=-\infty}^{\infty} p_k \log_2 p_k$$

Key Metrics:

ℓ_1*-norm* (integer lattice distance):

$$\|\vec{z}\|_1 := \sum_{i=1}^{n} |z_i|$$

The expectation is the **mean step-size**

$$S := \mathbb{E}\left[\|\vec{z}\|_1\right] = \sum_{i=1}^{n} \mathbb{E}\left[|z_i|_1\right],$$

due to the stochastic independence.
For n i.i.d. variables: $S = n \cdot \mathbb{E}\left[|z_1|\right]$.

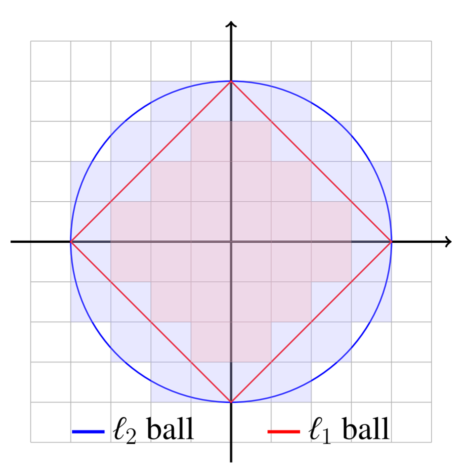
Working Hypothesis and Research Questions

Hypothesis: the ℓ_1-norm is the natural measure over the integer lattice.
The role of (truncated) Gaussianity remains unclear.
Can we leverage established ℓ_2-norm continuous results?

Research Questions:

1. **Geometry & Entropy:** How to design mutations respecting ℓ_1 geometry?
2. **Correlated ℓ_1-based mutations:** Can we construct correlated preserving dependencies?

ℓ_1 vs. ℓ_2 Geometry

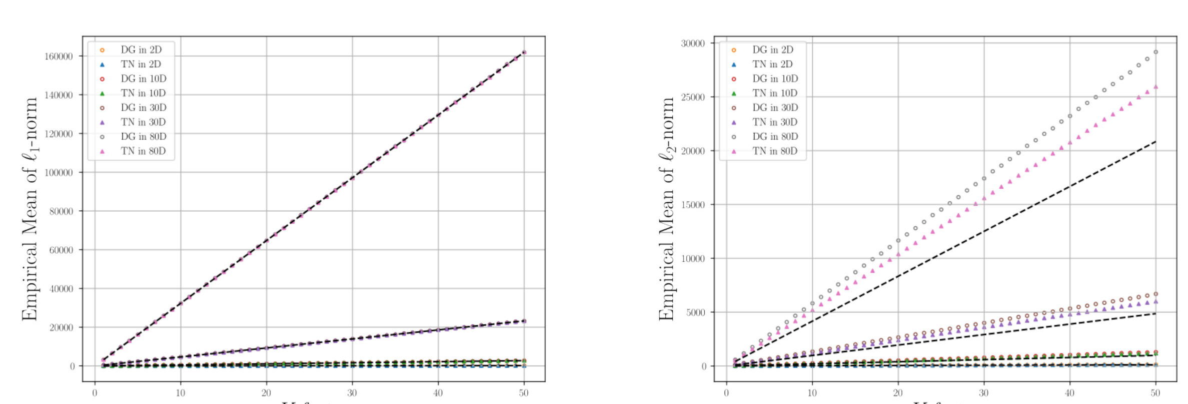


Integer Mutation Distributions

Discrete Uniform (**DU**)	Shifted Binomial (**SB**)	Truncated Normal (**TN**)	Double Geometric (**DG**)
Range: $\{-N, \ldots, N\}$	Range: $\{-N/2, \ldots, N/2\}$	Support: $\mathbb{Z}$; $z \sim \mathcal{N}(0, \sigma^2)$	Support: $\mathbb{Z}$; $z = \mathcal{G}(p) - \mathcal{G}(p)$
$\mathbf{Pr}\{X = k\} = \frac{1}{2N+1}$	$\mathbf{Pr}\{Y = k\} = \binom{N}{k+\frac{N}{2}} 2^{-N}$	$\mathbf{Pr}\{\texttt{round}(z) = k\} = \frac{1}{2}\left[\text{erf}\left(\frac{k+0.5}{\sqrt{2}\sigma}\right) - \text{erf}\left(\frac{k-0.5}{\sqrt{2}\sigma}\right)\right]$	$\mathbf{Pr}\{z = k\} = \frac{p}{2-p}(1-p)^{\|k\|}$
Step: $S_{DU} = n\frac{N(N+1)}{2N+1}$	Step: $S_{SB} \approx n\sqrt{\frac{2N}{\pi}}$	Step: $S_{TN} \approx n\sigma\sqrt{\frac{2}{\pi}}$	Step: $S_{DG} = n\frac{2(1-p)}{p(2-p)}$
Entropy: $H = \log_2(2N+1)$	Historical: First ES (1964)	Most common in IESs	ℓ_1-optimal (Rudolph)
Max entropy on range	via Galton boards	ℓ_2-based geometry	Max entropy given S

Empirical Mean of ℓ_1 vs. ℓ_2: TN & DG

Populations of randomly generated n-dimensional integer vectors with an increasing scale of individual step-sizes, that is $S_i = K \cdot i$, subject to a factor $K \in \{1, \ldots, 50\}$:



Correlated Integer Mutations via Rotations

Challenge: ℓ_1-norm-preserving rotations are unrealistic in the general case. How to correlate?

Approach: Apply Schwefel's rotations and round:

$$\vec{z}_c = \texttt{round}\left[\left(\prod_{i=1}^{n-1} \prod_{j=i+1}^{n} \mathbf{R}(\alpha_{ij})\right) \cdot \vec{z}_u\right]$$

$\vec{\alpha}$: $n(n-1)/2$-dimensional vector of angles.

Implementation:

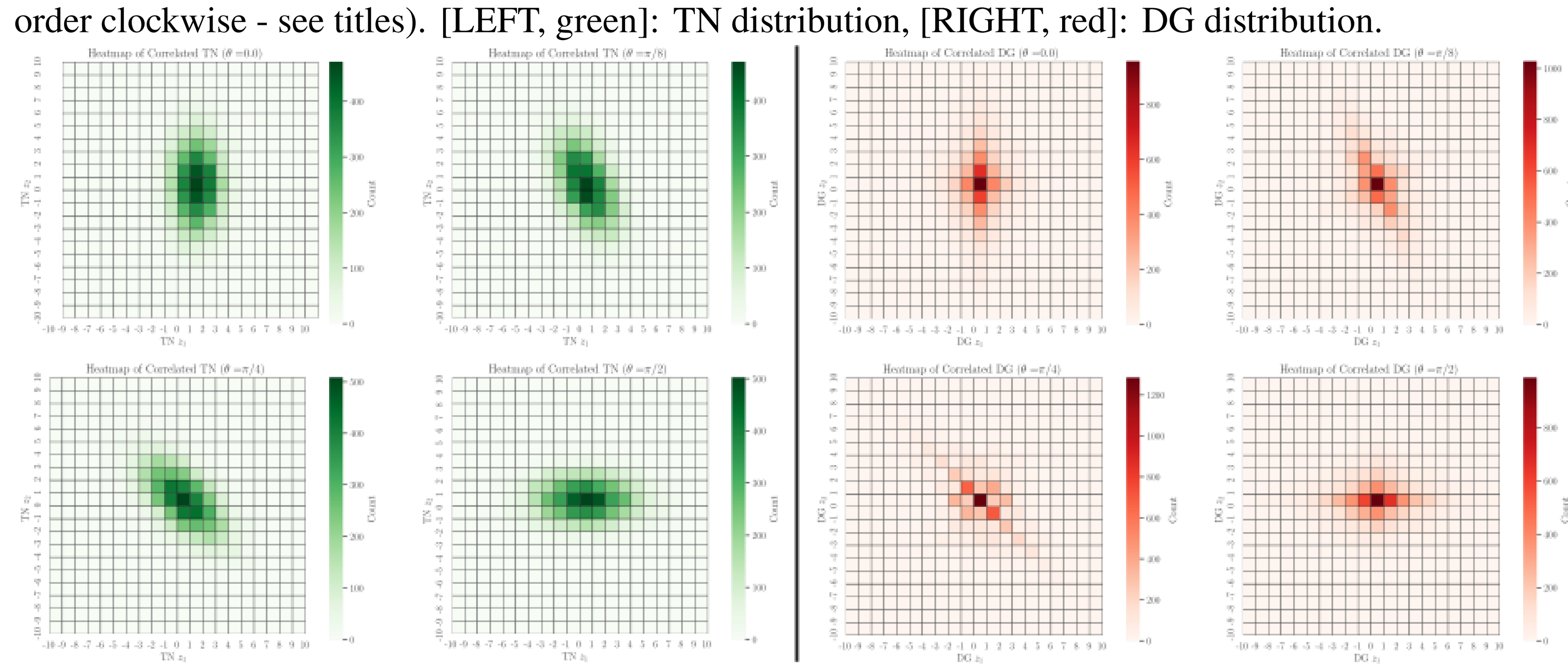
```
rotateInt (z⃗, α⃗)
  for j = 1, ..., n · (n − 1)/2 do
    z⃗ ⟵ R(α_j) z⃗
  end
  return {round (z⃗)}
```

Input: uncorrelated $\vec{z}_u$, angles $\vec{\alpha}$
Output: correlated integer $\vec{z}_c$

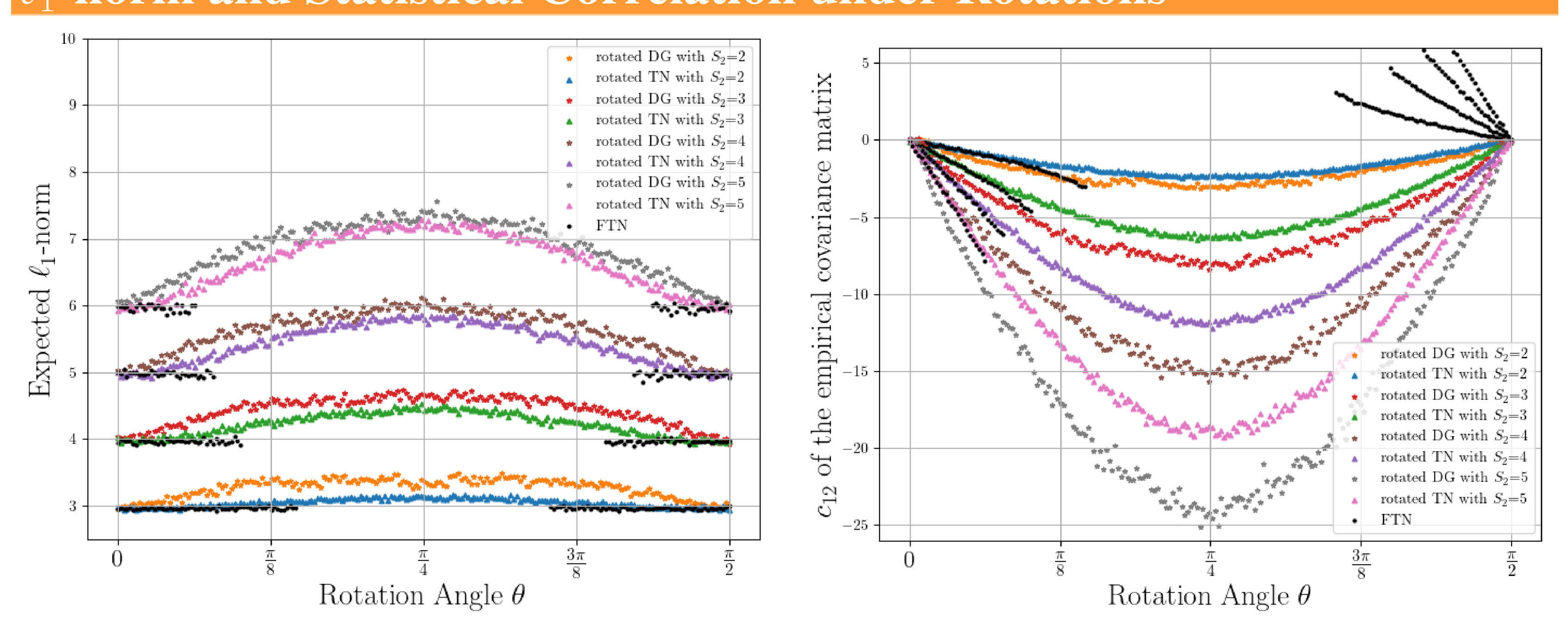
Rotation matrix $\mathbf{R}(\alpha_{k\ell})$: $r_{kk} = r_{\ell\ell} = \cos(\alpha_{k\ell})$, $r_{k\ell} = -r_{\ell k} = -\sin(\alpha_{k\ell})$ (identity otherwise)

2D Population Visualization (population size = 10^4)

Heatmaps depicting populations of rotated 2D samples with $S_1 = 1$, $S_2 = 2$ and $\theta \in \{0, \frac{\pi}{8}, \frac{\pi}{4}, \frac{\pi}{2}\}$ (set in this order clockwise - see titles). [LEFT, green]: TN distribution, [RIGHT, red]: DG distribution.



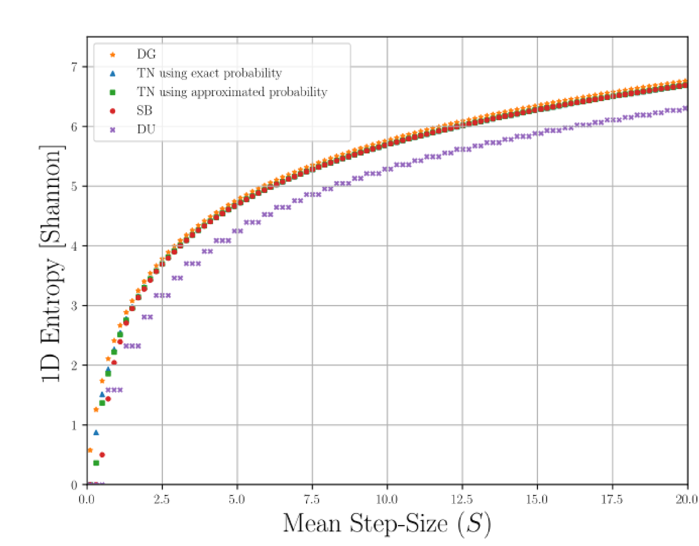
ℓ_1-norm and Statistical Correlation under Rotations



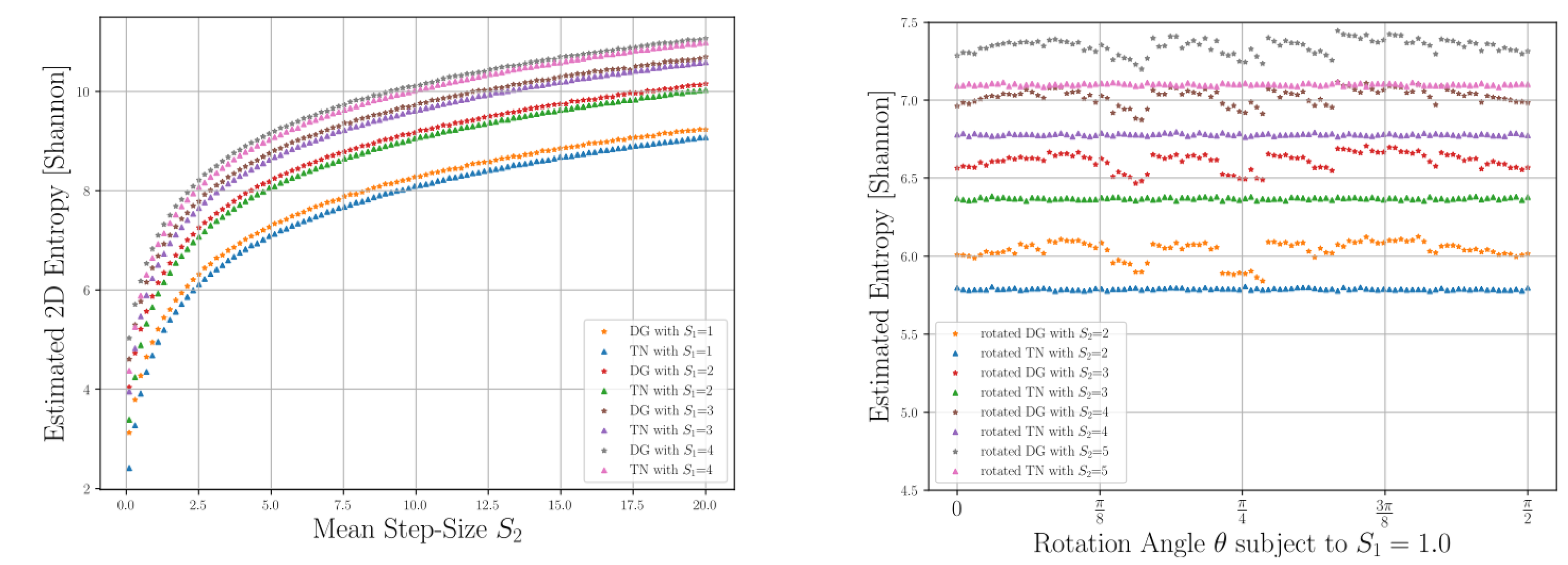
Shannon's 1D Entropy vs. Mean Step-Size: DG is the Maximizer

Key Finding: The entropy function of single-variable distributions over the spectrum of S reveals that DG achieves maximum entropy while controlling the defining step-size.

This relationship demonstrates the optimality of the DG method in terms of information-theoretic measures, numerically validating Rudolph's result.



Estimated Entropy of 2D Samples: Correlated & Uncorrelated



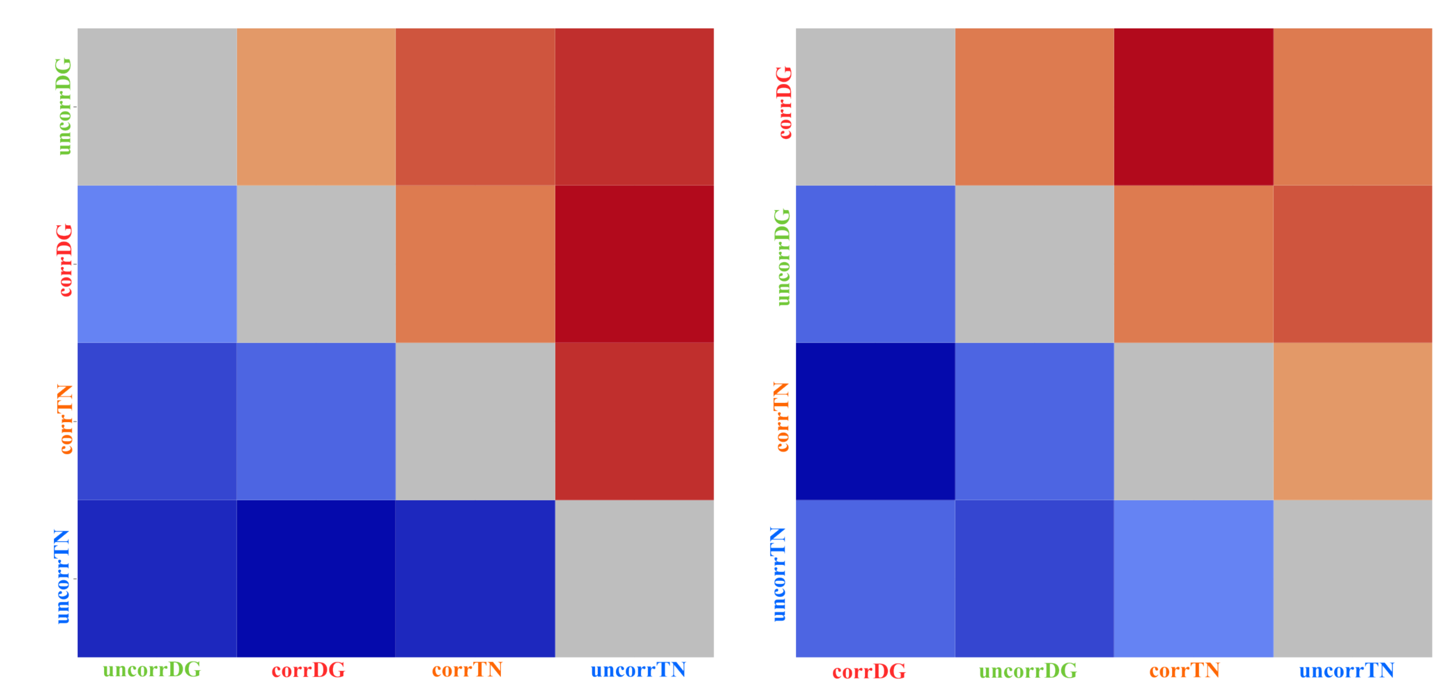
Numerical Validation: Unbounded Integer Quadratic

$$\text{minimize}_{\vec{x}} \ \frac{1}{c}\left[(\vec{x} - \vec{\xi}_0)^T \mathbf{H} (\vec{x} - \vec{\xi}_0)\right]$$
$$\text{subject to: } \vec{x} \in \mathbb{Z}^n$$

(H-1) Discus: $(H_{\text{disc}})_{11} = c$, $(H_{\text{disc}})_{ii} = 1$ $(i > 1)$
(H-2) Cigar: $(H_{\text{cigar}})_{11} = 1$, $(H_{\text{cigar}})_{ii} = c$ $(i > 1)$
(H-3) RotEllipse: $H_{\text{RE}} = \mathcal{O} H_{\text{ellipse}} \mathcal{O}^{-1}$
(H-4) HadEllipse: $H_{\text{HE}} = \mathcal{S} H_{\text{ellipse}} \mathcal{S}^{-1}$

Conditioning: $c \in \{10, 10^2, \ldots, 10^6\}$ • **Total:** 24 instances/dimension

Results of the Standard-IES per 64D



Key findings:

- `uncorrDG` dominates separable problems
- `corrDG` dominates nonseparable problems
- DG-based IESs consistently outperform TN-based

Summary and Outlook

Key Contributions:

- **Theoretical:** Established that integer optimization benefits from ℓ_1-norm symmetries rather than classical ℓ_2-invariance, motivating geometry-respecting mutation operators for discrete spaces
- **Algorithmic:** Extended DG distribution to correlated integer sampling, achieving highest entropy among tested kernels for given step lengths, thus maximizing exploratory power
- **Empirical:** Demonstrated superior convergence on IQP benchmarks, though revealing universal stagnation near optima - a phase-transition-like phenomenon requiring further investigation

Future Directions: Runtime analysis of stagnation mechanisms • Derandomized step-size adaptation (CMA-ES integration) • NK landscapes and MI testbeds • Boundary-aware mutations for constrained problems