Theory of Population-based Evolutionary Algorithms on Complex Fitness Landscapes[1]

Per Kristian Lehre[2] **Duc-Cuong Dang**

School of Computer Science, University of Birmingham, United Kingdom
Algorithms for Intelligent Systems, University of Passau, Germany

PPSN 2026
University of Trento, Italy
August–September 2026

UNIVERSITY OF BIRMINGHAM

UKRI Engineering and Physical Sciences Research Council

UNIVERSITÄT PASSAU

SCAN ME

[1] Main results are from our joint work (Algorithmica, 2026) and the papers with Anton Eremeev, Andre Opris, and Dirk Sudholt

[2] Supported by a Turing AI Acceleration Fellowship (EPSRC grant ref EP/V025562/1)

Outline

Fitness Landscape Metaphor (Sewall Wright, 1932)

THE ROLES OF MUTATION, INBREEDING, CROSSBREEDING AND SELECTION IN EVOLUTION

Sewall Wright, University of Chicago, Chicago, Illinois

The enormous importance of biparental reproduction as a factor in evolution was brought out a good many years ago by East. The observed properties of gene mutation—fortuitous in origin, infrequent in occurrence and deleterious when not negligible in effect—seem about as unfavorable as possible for ar
ever, a limited
ried by the sp
through which
Estimates of
range from 10
in Drosophila c
compared with
of all type gen
cies. Judging
ganisms which

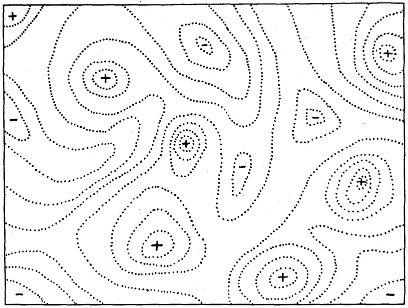
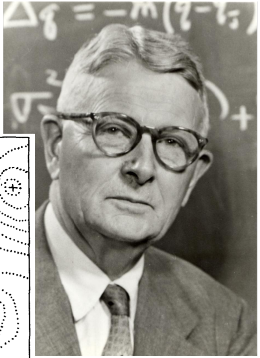


Figure 2.—Diagrammatic representation of the field of gene combinations in two dimensions instead of many thousands. Dotted lines represent contours with respect to adaptiveness.



"The problem of evolution as I see it is that of a mechanism by which the species may continually find its way from lower to higher peaks in such a field." (Wright, 1932)

Runtime Analysis (special case [3])

Algorithm (1+1) EA

Sample x_0 uniformly at random from $\{0,1\}^n$.
for $t = 0$ to ∞ **do**
  Make a copy of x_t and call it x', and flip each bit of x' with probability $1/n$
  if $f(x') \geq f(x_t)$ **then**
    $x_{t+1} := x'$
  else
    $x_{t+1} := x_t$

Runtime of (1+1) EA on function f

- $T_f := \min\{t \mid x_t \text{ is optimal }\}$

Worst-case expected runtime of (1+1) EA on problem class F

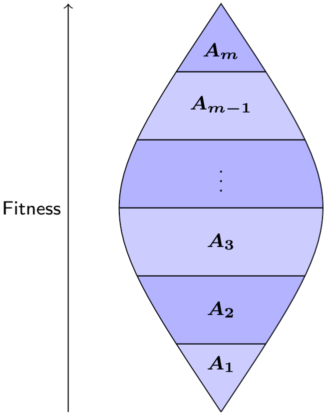
- $T_F := \max_{f \in F} \mathbf{E}[T_f]$

Runtime Analysis

- How does the structure of F influence T_F?

[3] The (1+1) EA is here just used as a simple introductory example. The rest of the tutorial will consider standard population-based EAs.

Runtime of (1+1) EA on Simple Fitness Landscapes[4]



If search space can be partioned into levels $(A_1, A_2, \dots, A_m)$ such that

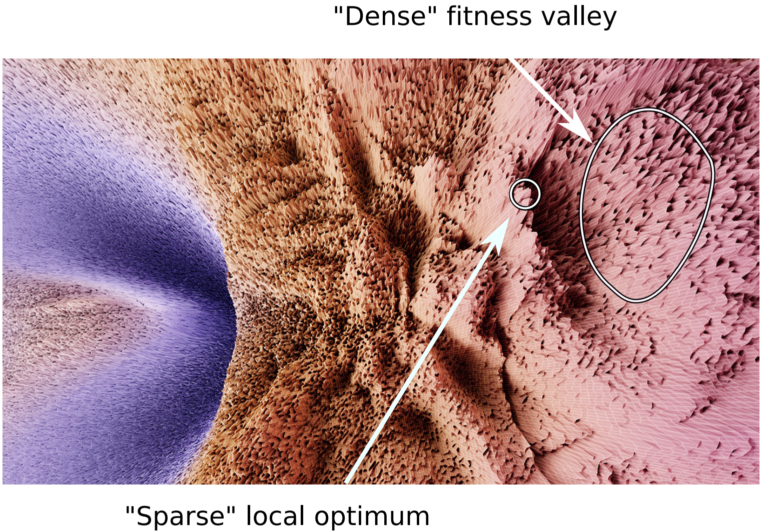
- f is monotone with respect to the levels, i.e.,
  if $x \in A_i$ and $y \in A_j$ where $i < j$, then $f(x) < f(y)$
- small number of levels with global optima in A_m, i.e.,
  $m \in \text{poly}(n)$ and for all $x \in A_m$ and $y \in \{0,1\}^n$, $f(x) \geq f(y)$
- every search point close to a higher level, i.e.,
  for all $x \in A_j$, there exists $y \in A_{j+1} \cup \cdots \cup A_m$ such that $H(x, y) = O(1)$.

then the expected time to reach the last level (optimum) is

$$\begin{aligned}\mathbf{E}\left[T_f\right] &\leq \sum_{j=1}^{m-1} \text{time to leave level } j \\ &\leq \sum_{j=1}^{m-1} \frac{1}{\text{prob. of leaving level } j} \\ &\leq \text{poly}(n)\end{aligned}$$

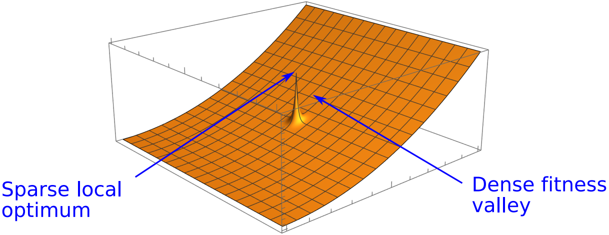
[4] This is essentially the artificial fitness level technique Wegener [2002].

Fitness Landscape Intuition[5]

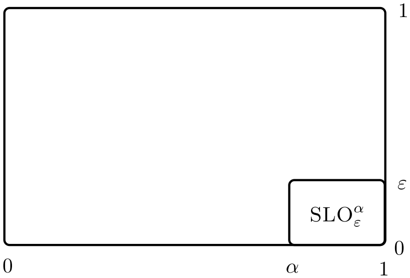


[5] Illustration from https://losslandscape.com/

Sparse Local Optima, Dense Fitness Valleys



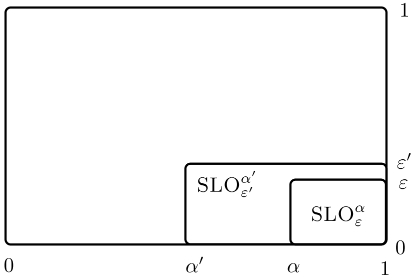
$\mathrm{SLO}_{\varepsilon}^{\alpha}$ hierarchy of problems



For $\boldsymbol{\alpha}, \boldsymbol{\varepsilon} \in [\mathbf{0}, \mathbf{1}]$, $\mathrm{SLO}_{\boldsymbol{\varepsilon}}^{\boldsymbol{\alpha}}$ is the subset of pseudo-Boolean functions with

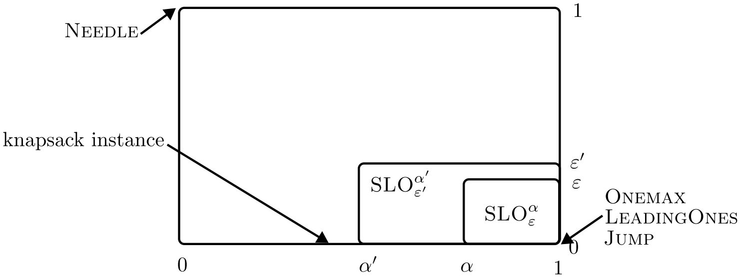
- Density of fitness valleys $\geq \boldsymbol{\alpha}$ ($\boldsymbol{\alpha} = \mathbf{1} \implies$ maximally dense/easy problem)
- Sparsity of local optima $\leq \boldsymbol{\varepsilon}$ ($\boldsymbol{\varepsilon} = \mathbf{0} \implies$ no local optima)

$\mathrm{SLO}_{\varepsilon}^{\alpha}$ hierarchy of problems

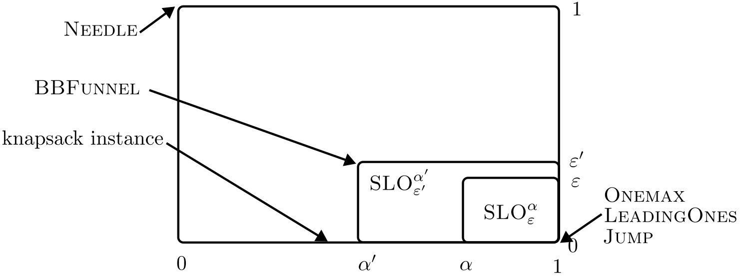


- $\mathrm{SLO}_{\mathbf{1}}^{\mathbf{0}}$ is the set of all pseudo-Boolean functions.

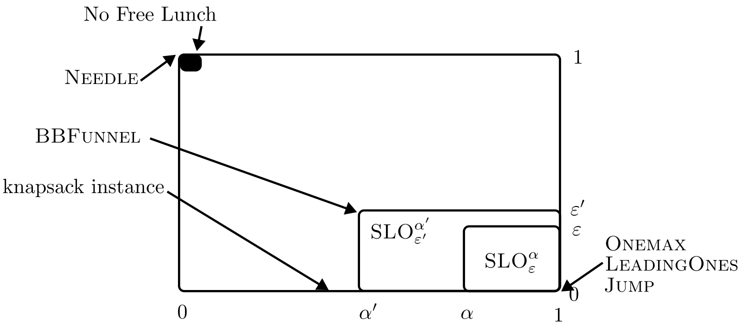
$\mathrm{SLO}_{\varepsilon}^{\alpha}$ hierarchy of problems



$\mathrm{SLO}_{\varepsilon}^{\alpha}$ hierarchy of problems



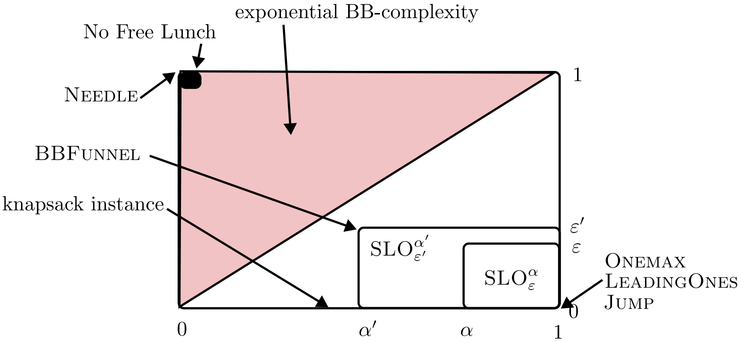
$\mathrm{SLO}^{\alpha}_{\varepsilon}$ hierarchy of problems



Theorem (No Free Lunch problem classes)

For any function class F closed under permutations, $F \subset \mathrm{SLO}^{1/n}_{1} \setminus \mathrm{SLO}^{c/n}_{1}$ for some constant $c > 1$.

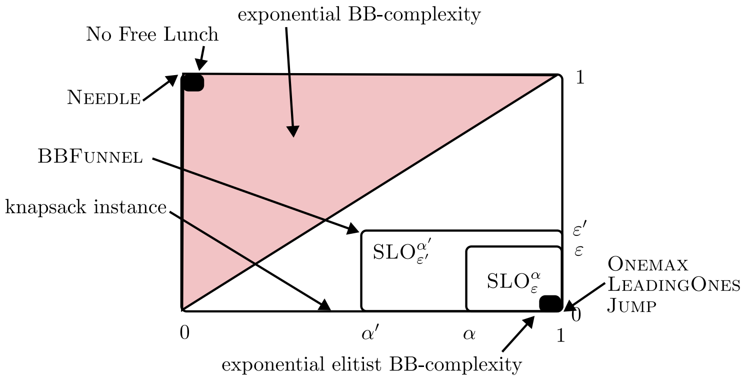
$\mathrm{SLO}_{\varepsilon}^{\alpha}$ hierarchy of problems



Theorem

For $\varepsilon > \alpha + \delta + n/3$, the unrestricted black-box complexity of $\mathrm{SLO}_{\varepsilon}^{\alpha}$ is at least $2^{n\delta/2-1}$.

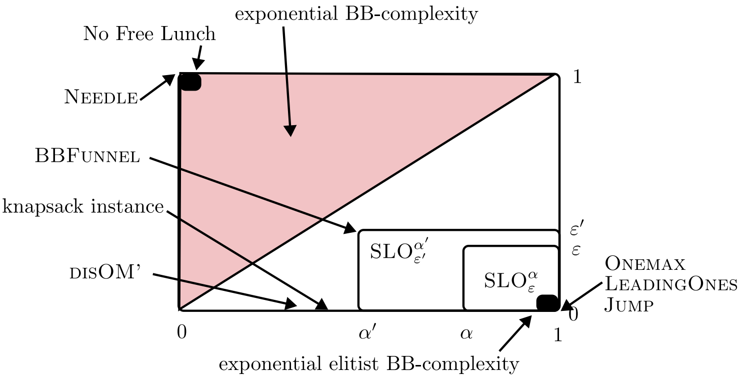
$\mathrm{SLO}_{\varepsilon}^{\alpha}$ hierarchy of problems



Theorem

For any $\delta > 0$, the $(\mu + \lambda)$ elitist black-box complexity of $\mathrm{SLO}_{\delta}^{1-\delta}$ is $e^{\Omega(n)}$.

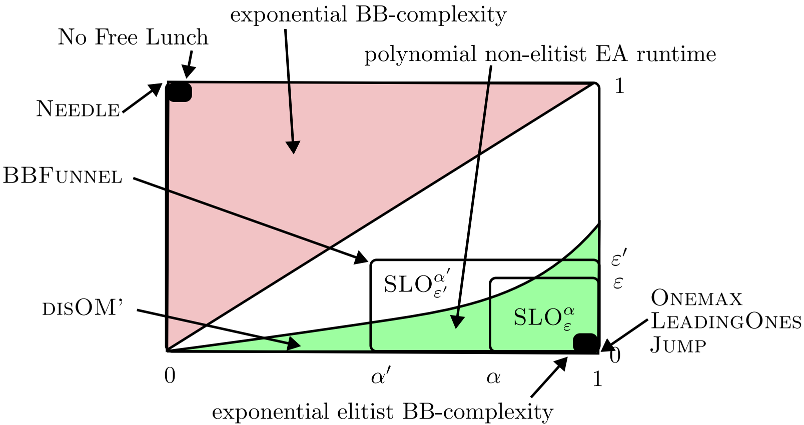
$\mathrm{SLO}_{\varepsilon}^{\alpha}$ hierarchy of problems



Theorem (ONEMAX with frozen noise)

For $\mathbf{0} < \boldsymbol{\delta} < \boldsymbol{\alpha} \leq \mathbf{1}$ *and* $\boldsymbol{p} < \left(\frac{\boldsymbol{\delta}}{\mathbf{2e}}\right)^{(\mathbf{1}+\boldsymbol{\alpha}+\boldsymbol{\delta})/\boldsymbol{\delta}}$, $\mathrm{DISOM}_{\boldsymbol{d},\boldsymbol{p}} \in \mathrm{SLO}_{\boldsymbol{\delta}}^{\boldsymbol{\alpha}-\boldsymbol{\delta}}$.

$\mathrm{SLO}_{\varepsilon}^{\alpha}$ hierarchy of problems



Black Box Model[6]

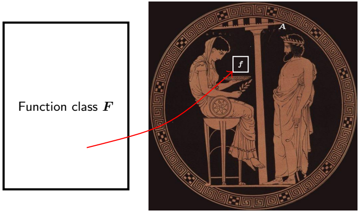


Photo: E. Gerhard (1846).

Droste, Jansen, and Wegener [2006]

[6] Also called query complexity model.

Black Box Model[6]

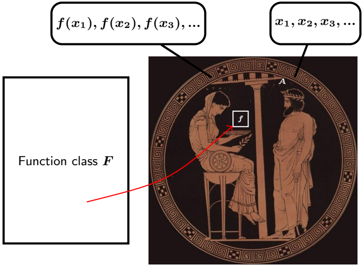


Photo: E. Gerhard (1846).

Droste et al. [2006]

[6] Also called query complexity model.

Black Box Model[6]

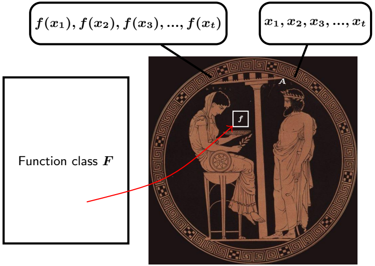


Photo: E. Gerhard (1846).

Droste et al. [2006]

- **Runtime** of algorithm A on f

$$T_{A,f} := \min_{t \in \mathbb{N}} \{t \mid \forall y \; f(x_t) \geq f(y)\}$$

- **Worst case** expected runtime

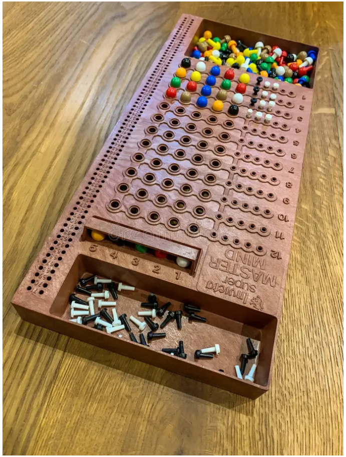
$$T_{A,\mathcal{F}} := \max_{f \in \mathcal{F}} \mathrm{E}\left[T_{A,f}\right]$$

- **Black box complexity** of $\mathcal{F}$

$$T_{\mathcal{F}} := \min_{A} T_{A,\mathcal{F}}$$

[6] Also called query complexity model.

Black-Box Complexity : Mastermind is Easy



$$\textsc{OneMax} = \{f_z | z \in \{0,1\}^n\} \text{ where } f_z(x) = \sum_{i=1}^{n} x_i \oplus z_i$$

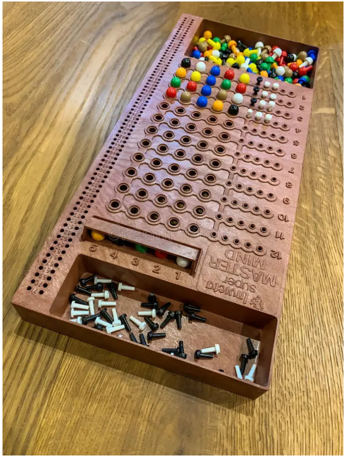
Theorem (Erdős and Renyi [1963])

$$T_{\textsc{OneMax}} = \Theta(n/\log n)$$

Proof idea

- Upper bound: Query $n/\log n$ bitstrings sampled u.a.r., whp, only one consistent z.
- Lower bound: Need to identify one among 2^n bitstrings, each query gives $O(\log(n))$ bits of information.

Black-Box Complexity : Needle in the Haystack is Hard



$$\textsc{Needle} = \{f_z | z \in \{0,1\}^n\} \text{ where } f_z(x) = \prod_{i=1}^{n} x_i \oplus z_i$$

Theorem (Droste et al. [2006])

$$T_{\textsc{Needle}} = (2^n + 1)/2$$

Proof idea

- Lower bound: Yao's minmax principle.
- Upper bound: Enumerate bitstrings in random order.

No Free Lunch Theorem

Definition

A function class $\mathcal{F} \subseteq \{f : \{0,1\}^n \to \mathbb{R}\}$ is closed under permutation iff

- for all $f \in \mathcal{F}$, for all permutations $\sigma : \{0,1\}^n \to \{0,1\}^n$, $f \circ \sigma \in \mathcal{F}$.

Theorem ([Wolpert and Macready, 1997, Droste, Jansen, and Wegener, 2002])

If $\mathcal{F}$ is closed under permutation, then for all black-box algorithms A and B,

$$\mathbf{E}_{f\sim\mathrm{Unif}(\mathcal{F})}\left[T_{A,f}\right] = \mathbf{E}_{f\sim\mathrm{Unif}(\mathcal{F})}\left[T_{B,f}\right].$$

Black Box Optimisation[7]

- What characterises the "easy" problems, i.e., those with polynomial black-box complexity?
- Are there "general" black-box algorithms which have polynomial runtime on broad classes of easy problems?

[7] A related, and possibly better understood question, is the query complexity of Boolean functions.

SLO_0^1 – The easy part of the hierarchy

A function $f : \{0,1\}^n \to \mathbb{R}$ belongs to SLO_0^1 iff there exists a partition $(A_1, A_2, \ldots, A_m)$ of $\{0,1\}^n$ such that

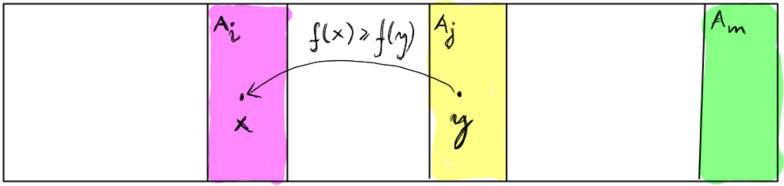
1. The last level A_m contains the globally optimal search points.
2. The Hamming-distance to a higher level is a constant wrt n.
3. Fitness increases monotonically according to the levels.

SLO_0^1 – The easy part of the hierarchy

A function $f : \{0,1\}^n \to \mathbb{R}$ belongs to SLO_0^1 iff there exists a partition $(A_1, A_2, \ldots, A_m)$ of $\{0,1\}^n$ such that

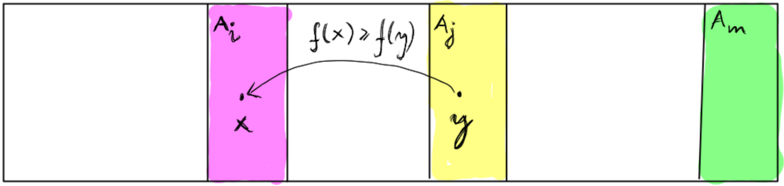
1. The last level A_m contains the globally optimal search points.
   $f(A_m) = \max_x f(x)$
2. The Hamming-distance to a higher level is a constant wrt n.
   for all $j \leq m-1$, for all $x \in A_j, H(x, A_{\geq j+1}) = O(1)$
3. Fitness increases monotonically according to the levels.
   for all $x \in A_i$ and $y \in A_j$ with $i < j$, $f(x) < f(y)$

Deceptive Pairs



$\neg$**3.** Fitness decreases with the levels:

Deceptive Pairs



¬3. Fitness decreases with the levels: $(A_i, A_j)_{i<j}$ is *deceptive* if $\exists x \in A_i$ and $\exists y \in A_j$: $f(x) \geq f(y)$.

$\mathrm{SLO}_{\varepsilon}^{\alpha}$– Incorporating deceptive levels

A function $f : \{0, 1\}^n \to \mathbb{R}$ belongs to $\mathrm{SLO}_{\varepsilon}^{\alpha}$ iff there exists a partition $(A_1, A_2, \ldots, A_m)$ of $\{0, 1\}^n$ such that

1. The last level A_m contains the globally optimal search points.

2. The Hamming-distance to a higher level is a constant wrt n.

3. Local optima are ε-sparse

4. Fitness valleys are α-dense

$\mathrm{SLO}_{\varepsilon}^{\alpha}$– Incorporating deceptive levels

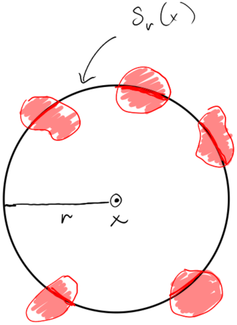
A function $f : \{0,1\}^n \to \mathbb{R}$ belongs to $\mathrm{SLO}_{\varepsilon}^{\alpha}$ iff there exists a partition $(A_1, A_2, \ldots, A_m)$ of $\{0,1\}^n$ such that

1. The last level A_m contains the globally optimal search points.
   $f(A_m) = \max_x f(x)$
2. The Hamming-distance to a higher level is a constant wrt n.
   for all $j \leq m-1$, for all $x \in A_j, H(x, A_{\geq j+1}) = O(1)$

If $(A_{i_1}, A_{j_i}), \ldots, (A_{i_u}, A_{j_u})$ are f-deceptive pairs then

3. Local optima are ε-sparse
   $\cup_{v=1}^{u} A_{i_v}$ is (ε, r)-sparse
4. Fitness valleys are α-dense
   A_{j_v} is α-dense with respect to $A_{\geq j_v}$ for all v

Sparsity

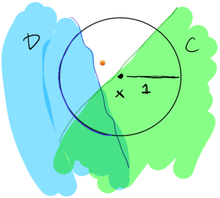


Definition
B is (ε, r)-*sparse* if and only if
for all $q \in [r]$ and for all x

$$|S_q(x) \cap B| \leq \varepsilon \cdot \binom{n}{q}$$

where $y \in S_q(x) \iff H(x, y) = q$.

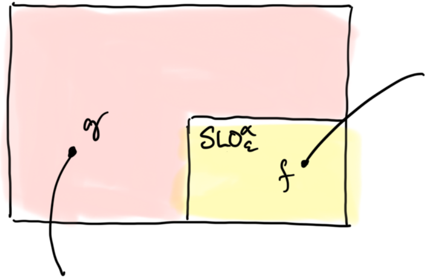
Density



Definition
C is α-*dense* wrt D if and only if for all $x \in C$

$$|S_1(x) \cap D| \geq \alpha \cdot \binom{n}{1}$$

Upper vs Lower bounds



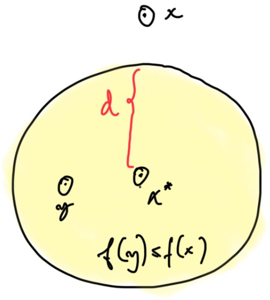
UPPER BOUNDS

Prove $f \in SLO^\alpha_\varepsilon$ by constructing a partition satisfying SLO^α_ε conditions

LOWER BOUNDS

Prove $g \notin SLO^\alpha_\varepsilon$ by showing that every partition does not satisfy SLO^α_ε-conditions

Lower bound example



Lemma SLO17

Let f be any fitness function with a unique global optimum x^*.
If there exists a search point x and $d \in \omega(1)$ such that

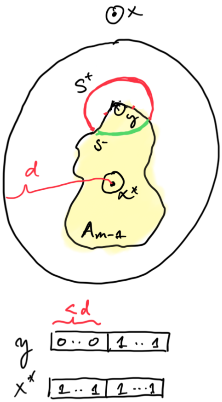
- $H(x, x^*) \geq d$
- $f(y) \leq f(x)$ for all $y \neq x^*$ with $H(x^*, y) < d$

then $f \notin \mathrm{SLO}_1^{d/n}$.

Corollary

- $\textsc{Needle} \notin \mathrm{SLO}_1^{d}$ *for any* $d = \omega(1/n)$
- $\textsc{Jump}_k \notin \mathrm{SLO}_1^{k/n}$ *for any* $k = \omega(1)$.

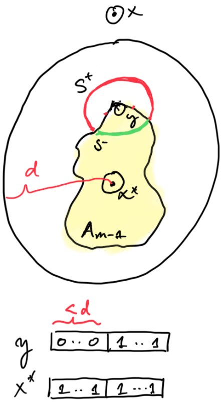
Proof of Lower Bound



Let $\boldsymbol{A} = (\boldsymbol{A}_1, \ldots, \boldsymbol{A}_m)$ be any partition which satisfies the SLO_1^{α}-conditions wrt $\boldsymbol{f}$

1. $\boldsymbol{x}^*$ is a unique global opt. $\implies \boldsymbol{A}_m = \{\boldsymbol{x}^*\}$ by SLO cond 1.

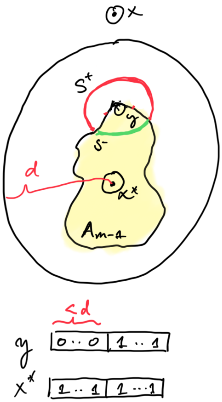
Proof of Lower Bound



Let $\boldsymbol{A} = (\boldsymbol{A_1}, \ldots, \boldsymbol{A_m})$ be any partition which satisfies the SLO_1^{α}-conditions wrt $\boldsymbol{f}$

1. $\boldsymbol{x^*}$ is a unique global opt. $\implies \boldsymbol{A_m} = \{\boldsymbol{x^*}\}$ by SLO cond 1.
2. Choose any $\boldsymbol{y} \in \boldsymbol{A_{m-1}}$ which maximises the Hamming distance to $\boldsymbol{x^*}$.
   By SLO cond. 2, $\boldsymbol{H(y, x^*) = O(1) < d}$.

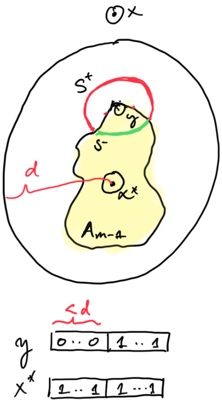
Proof of Lower Bound



Let $A = (A_1, \ldots, A_m)$ be any partition which satisfies the SLO_1^{α}-conditions wrt f

1. x^* is a unique global opt. $\implies A_m = \{x^*\}$ by SLO cond 1.
2. Choose any $y \in A_{m-1}$ which maximises the Hamming distance to x^*.
   By SLO cond. 2, $H(y, x^*) = O(1) < d$.
3. Claim 1: (A_j, A_{m-1}) is an f-deceptive pair for some $j < m - 2$
   - By assumption $H(x^*, x) \notin O(1)$, so $x \in A_j$ for some $j < m - 1$ by SLO cond 2.
   - We have $x \in A_j$, $y \in A_{m-1}$ and $f(y) \leq f(x)$, so (A_j, A_{m-2}) is f-deceptive.

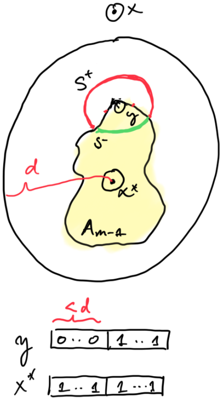
Proof of Lower Bound



Let $A = (A_1, \ldots, A_m)$ be any partition which satisfies the SLO_1^{α}-conditions wrt f

1. x^* is a unique global opt. $\implies A_m = \{x^*\}$ by SLO cond 1.
2. Choose any $y \in A_{m-1}$ which maximises the Hamming distance to x^*.
   By SLO cond. 2, $H(y, x^*) = O(1) < d$.
3. Claim 1: (A_j, A_{m-1}) is an f-deceptive pair for some $j < m - 2$
   - By assumption $H(x^*, x) \notin O(1)$, so $x \in A_j$ for some $j < m - 1$ by SLO cond 2.
   - We have $x \in A_j$, $y \in A_{m-1}$ and $f(y) \leq f(x)$, so (A_j, A_{m-2}) is f-deceptive.
4. Claim 2: $|S_1(y) \cap A_{\geq m-1}| \leq H(y, x^*) < d$
   - Partition the set $S_1(y)$ of Hamming-neighbours of y into two subsets
     - S^- are bitstrings with Hamming-distance $H(y, x^*) - 1$ to x^*
     - S^+ are bitstrings with Hamming-distance $H(y, x^*) + 1$ to x^*, and hence $S^+ \cap A_{\geq m-1} = \emptyset$
   - It follows $|S_1(y) \cap A_{\geq m-1}| = |S^- \cap A_{\geq m-1}| \leq |S^-| = H(y, x^*) < d$

Proof of Lower Bound



Let $A = (A_1, \ldots, A_m)$ be any partition which satisfies the SLO_1^α-conditions wrt f

1. x^* is a unique global opt. $\implies A_m = \{x^*\}$ by SLO cond 1.
2. Choose any $y \in A_{m-1}$ which maximises the Hamming distance to x^*.
   By SLO cond. 2, $H(y, x^*) = O(1) < d$.
3. Claim 1: (A_j, A_{m-1}) is an f-deceptive pair for some $j < m-2$
   - By assumption $H(x^*, x) \notin O(1)$, so $x \in A_j$ for some $j < m-1$ by SLO cond 2.
   - We have $x \in A_j$, $y \in A_{m-1}$ and $f(y) \leq f(x)$, so (A_j, A_{m-2}) is f-deceptive.
4. Claim 2: $|S_1(y) \cap A_{\geq m-1}| \leq H(y, x^*) < d$
   - Partition the set $S_1(y)$ of Hamming-neighbours of y into two subsets
     - S^- are bitstrings with Hamming-distance $H(y, x^*) - 1$ to x^*
     - S^+ are bitstrings with Hamming-distance $H(y, x^*) + 1$ to x^*, and hence $S^+ \cap A_{\geq m-1} = \emptyset$
   - It follows $|S_1(y) \cap A_{\geq m-1}| = |S^- \cap A_{\geq m-1}| \leq |S^-| = H(y, x^*) < d$
5. By Claim 1, SLO cond 2, and Claim 2 $\alpha n \leq |S_1(y) \cap A_{\geq m-1}| < d.$

Unrestricted Black Box Complexity

Corollary SLO30

If $\alpha < 1 - 3/n$ and $\varepsilon > \alpha + 3/n$, then the unrestricted black-box complexity of $\mathrm{SLO}^{\alpha}_{\varepsilon,n-1}$ is $2^{\Omega(n(\varepsilon-\alpha))}$.

$\Longrightarrow$ If local optima sparsity ε is larger than fitness valley density α,
then the fitness landscape is dominated by deceptive regions.
In this case, all algorithms inefficient on at least some functions in $\mathrm{SLO}^{\alpha}_{\varepsilon}$.

$(\mu + \lambda)$-Elitist Black-Box Complexity

$(\mu + \lambda)$-Elitist Black-Box Model [Doerr and Lengler, 2016]

- algorithm only remembers fittest $\boldsymbol{\mu}$ search points seen so far, produces $\boldsymbol{\lambda}$ new search points in each round
- covers a wide range of elitist EAs, such as $(\boldsymbol{\mu} + \boldsymbol{\lambda})$ EA.

$(\mu + \lambda)$-Elitist Black-Box Complexity

$(\mu + \lambda)$-Elitist Black-Box Model [Doerr and Lengler, 2016]

- algorithm only remembers fittest μ search points seen so far, produces λ new search points in each round
- covers a wide range of elitist EAs, such as $(\mu + \lambda)$ EA.

Theorem
For any constant $\alpha \in (0, 1)$, $\varepsilon \in \omega(1/n)$ and $\mu, \lambda \in \mathrm{poly}(n)$, the elitist $(\mu+\lambda)$ black-box complexity of $\mathrm{SLO}_{\varepsilon}^{\alpha}$ is $e^{\Omega(n)}$.

Theorem
For $\chi \notin (\ln(\lambda/\mu) - \delta, \ln(\lambda/\mu) + \delta)$, then (μ, λ) EA has exp. runtime on $\mathrm{SLO}_{\varepsilon}^{\alpha}$.

$(\mu + \lambda)$-Elitist Black-Box Complexity

$(\mu + \lambda)$-Elitist Black-Box Model [Doerr and Lengler, 2016]

- algorithm only remembers fittest μ search points seen so far, produces λ new search points in each round
- covers a wide range of elitist EAs, such as $(\mu + \lambda)$ EA.

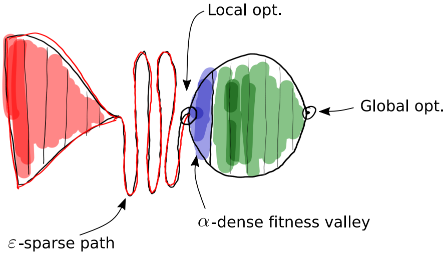
Theorem

For any constant $\alpha \in (0,1)$, $\varepsilon \in \omega(1/n)$ and $\mu, \lambda \in \text{poly}(n)$, the elitist $(\mu+\lambda)$ black-box complexity of $\mathrm{SLO}_{\varepsilon}^{\alpha}$ is $e^{\Omega(n)}$.

Theorem

For $\chi \notin (\ln(\lambda/\mu) - \delta, \ln(\lambda/\mu) + \delta)$, then (μ, λ) EA has exp. runtime on $\mathrm{SLO}_{\varepsilon}^{\alpha}$.

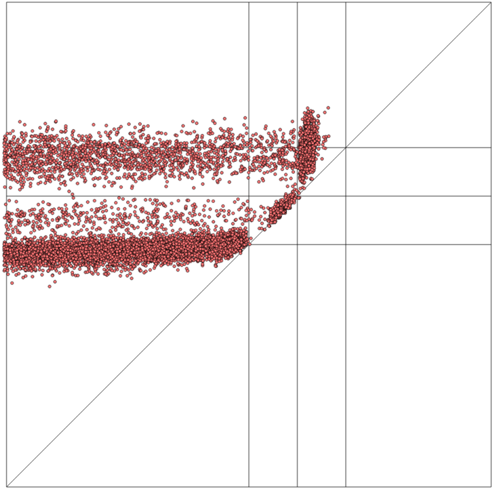
Proof.



- For any constants $\alpha, \varepsilon \in (0,1)$ $\textsc{BBFunnel} \in \mathrm{SLO}_{\varepsilon}^{\alpha}$.
- Any $(\mu + \lambda)$ BB EA needs exp. time on $\textsc{BBFunnel}$
- (μ, λ) EA has exp. runtime when $\chi \geq \ln(\lambda/\mu) + \delta$
- (μ, λ) EA behaves as $(\mu + \lambda)$ EA when $\chi \leq \ln(\lambda/\mu) - \delta$

$\square$

Web-based Simulation on BBFunnel



https://www.cs.bham.ac.uk/~lehrepk/aaai2021/

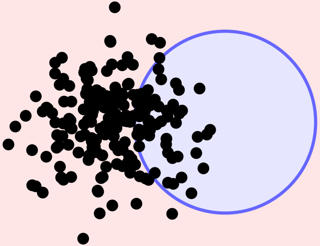
Population Process on the Boolean Hypercube

$$\mathcal{X} = \{0, 1\}^n$$

$$P_0 \sim \mathrm{Unif}\left(\mathcal{X}^{\lambda}\right)$$

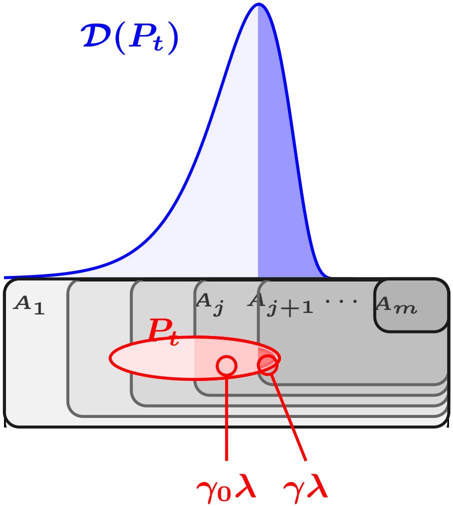
$$\forall t \in \mathbb{N}, P_t = (x_1, \ldots, x_\lambda)$$

where $x_i \sim \mathcal{D}_n(P_{t-1}, f(P_{t-1}))$ i.i.d.



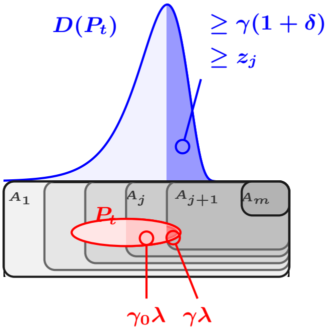
- Process induced by Markov kernel $\mathcal{D}_n : \mathcal{X}^{\lambda} \times \mathbb{R}^{\lambda} \to \Delta(\mathcal{X})$, e.g., selection and mutation in an evolutionary algorithm.

Level-Based Theorem



$$\mathcal{X} = A_{\geq 1} \supset A_{\geq 2} \supset \cdots \supset A_{\geq m-1} \supset A_{\geq m} = B$$

Level-Based Theorem



Theorem ([Corus, Dang, Eremeev, and Lehre, 2018])

If for any level $j < m$ and population P where

$$|P \cap A_{\geq j}| \geq \gamma_0 \lambda > |P \cap A_{\geq j+1}| =: \gamma\lambda$$

an individual $y \sim \mathcal{D}(P)$ is in A_{j+1} with

$$\Pr\left(y \in A_{\geq j+1}\right) \geq \max\left\{\gamma(1+\delta), z_j\right\}$$

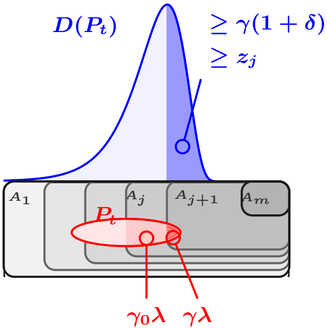
and the population size λ is at least

$$\lambda \geq \left(\frac{4}{\gamma_0 \delta^2}\right) \ln\left(\frac{128m}{z_* \delta^2}\right) \quad \text{where } z_* := \min_j z_j$$

then level A_m is reached in expected time

$$\left(\frac{8}{\delta^2}\right) \sum_{j=1}^{m-1} \left(\lambda \ln\left(\frac{6\delta\lambda}{4 + z_j \delta\lambda}\right) + \frac{1}{z_j}\right)$$

Level-Based Theorem



Theorem ([Corus et al., 2018])

If for any level $j < m$ and population P where

$$|P \cap A_{\geq j}| \geq \gamma_0 \lambda > |P \cap A_{\geq j+1}| =: \gamma\lambda$$

an individual $y \sim \mathcal{D}(P)$ is in A_{j+1} with

$$\Pr\left(y \in A_{\geq j+1}\right) \geq \max\left\{\gamma(1+\delta), z_j\right\}$$

and the population size λ is at least

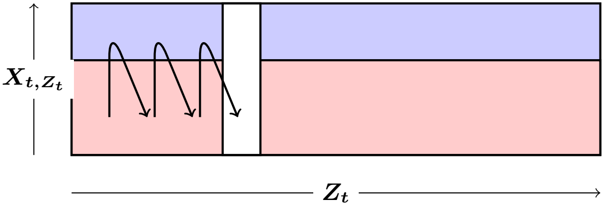
$$\lambda \geq \left(\frac{4}{\gamma_0 \delta^2}\right) \ln\left(\frac{128m}{z_* \delta^2}\right) \quad \textit{where } z_* := \min_j z_j$$

then level A_m is reached in expected time

$$\left(\frac{8}{\delta^2}\right) \sum_{j=1}^{m-1} \left(\lambda \ln\left(\frac{6\delta\lambda}{4 + z_j \delta\lambda}\right) + \frac{1}{z_j}\right)$$

$$< \left(\frac{8}{\delta^2}\right) \left(m\lambda \ln(2\delta\lambda) + \sum_{j=1}^{m-1} \frac{1}{z_j}\right)$$

Proof Idea



Population state captured by

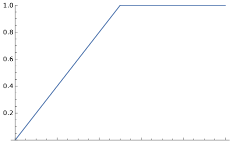
- Z_t - "current" level
- $X_{t,j}$ - # inividuals beyond j-th level

The following process has drift $\mathbf{E}[Y_t - Y_{t+1} \mid Y_t] = \Omega(\delta^2)$

$$Y_t := (m - Z_t) \ln\left(\frac{1+\lambda}{1+X_{t,Z_t}}\right)$$

Truncation Selection

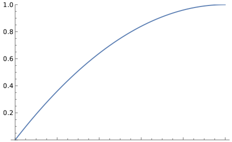
Sample uniformly at random among best μ/λ fraction of population



$$\beta(\gamma) = \min(\frac{\gamma\lambda}{\mu}, 1)$$

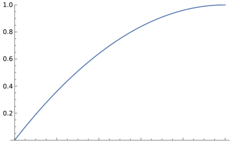
Tournament Selection

Select best among k uniformly selected individuals



$$\beta(\gamma) = 1 - (1-\gamma)^k$$

Linear Ranking Selection



$$\beta(\gamma) = \int_0^\gamma \eta(1-2x) + 2x dx$$
$$= \gamma(\eta(1-\gamma) + \gamma)$$

Algorithm Non-elitist EA with Ranking-Selection

Require: Fitness function $f : \{0,1\}^n \rightarrow \mathbb{R}$
Require: Population size $\lambda \in \mathbb{N}$ and mutation rate $\chi \in (0,n)$
Require: Ranking-based selection mechanism with cdf. β

```
Sample initial population P_0(1), ..., P_0(λ) uniformly at random.
for t = 0, 1, 2, ... until termination condition is met do
    Sort P_t by f such that f(P_t(1)) ≥ ··· ≥ f(P_t(λ)).
    for i = 1 to λ do
        Select parent index I such that Pr(I ≤ γλ) = β(γ).
        Obtain P_{t+1}(i) from P_t(I) by mutating each bit independently with prob χ/n.
```

Expected Polynomial Runtime on $\mathrm{SLO}_{\varepsilon}^{\alpha}$[8]

Theorem

If there exist constants $\psi_0, \gamma_0, \alpha \in (0,1)$, with $\psi_0 \geq \gamma_0$ and $\psi_0 + 2\gamma_0 \leq 1$, and $\delta \in (0,1), \varepsilon' \in (0, \min(\delta, 2/3)]$ with $1/\varepsilon' \in \mathrm{poly}(n)$, and $\varepsilon \in (0, \psi_0)$, and $r \in [\lfloor n/2 \rfloor]$ such that Algorithm 1 with the bitwise mutation operator with rate $\chi \in (0, n/2]$ satisfying $1/\chi \in \mathrm{poly}(n)$ and $(e\chi/r)^r \leq \varepsilon \left(1 - \frac{\chi}{n}\right)^n$, and an f-monotone selection mechanism with selection probability function β, and population size λ satisfying

(SM0) $\beta(0, \gamma) \leq \frac{\gamma(1-\varepsilon') - \varepsilon}{\left(1 - \frac{\chi}{n}\right)^n}$ *for all* $\gamma \in [\psi_0, 1]$,

(SM2a) $\beta(0, \gamma) \geq \frac{\gamma(1+\delta)}{\left(1 - \frac{\chi}{n}\right)^n}$ *for all* $\gamma \in (0, \gamma_0]$,

(SM2b) $\beta(\psi_0 + \gamma_0, \psi_0 + \gamma_0 + \gamma) \geq \frac{\gamma(1+\delta)}{\left(1 - \frac{\chi}{n}\right)^n (1 + \alpha\chi)}$ *for all* $\gamma \in (0, \gamma_0]$,

(SM3) $\frac{c}{(\varepsilon')^2} \ln(n) \leq \lambda \in \mathrm{poly}(n)$ *for a sufficiently large constant* c

then Algorithm 1 has expected polynomial runtime on $\mathrm{SLO}_{\varepsilon}^{\alpha}$.

[8] We define $\beta(x, y) := \beta(y) - \beta(x)$

Polynomial runtime of non-elitist EAs on $\mathrm{SLO}_{\varepsilon}^{\alpha}$

Theorem

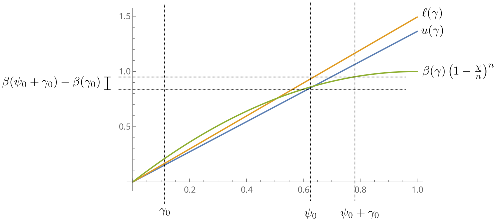
If $\lambda \in \Omega(\log(n)) \cap \mathrm{poly}(n)$ *and there exist constants* $\psi_0, \gamma_0, \delta \in (0,1)$ *st.*

$$\beta(\gamma)\left(1-\frac{\chi}{n}\right)^n \geq \gamma(1+\delta) =: \ell(\gamma), \qquad \forall \gamma \in (0, \gamma_0$$

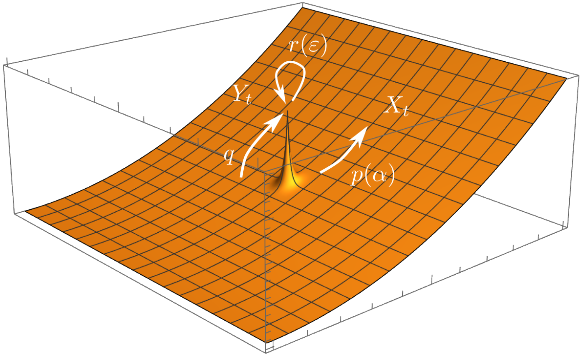
$$\beta(\gamma)\left(1-\frac{\chi}{n}\right)^n \leq \gamma(1-\varepsilon') - \varepsilon =: u(\gamma), \qquad \forall \gamma \in [\psi_0, 1$$

$$(\beta(\psi_0+\gamma_0+\gamma) - \beta(\psi_0+\gamma_0))\left(1-\frac{\chi}{n}\right)^n (1+\alpha\chi) \geq \gamma(1+\delta) \qquad \forall \gamma \in (0,\gamma_0], \psi \in [0, \psi_0$$

then the non-elitist EA optimises $\mathrm{SLO}_{\varepsilon}^{\alpha}$ *in* **expected polynomial time**.



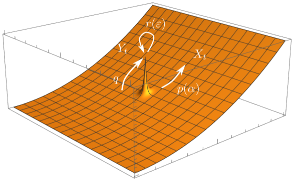
Dynamics of Peak Individuals (Y_t) vs Fitness Valley Individuals (X_t)

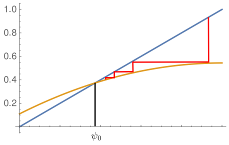


Dynamics of Peak Individuals (Y_t) vs Fitness Valley Individuals (X_t)

$$\Pr\left(y \in \text{peak}\right) < \beta(Y_t/\lambda)r(\varepsilon) + q$$

$$\Pr\left(y \in \text{fitness valley}\right) \geq \left(\beta\left(\frac{Y_t + X_t}{\lambda}\right) - \beta\left(\frac{Y_t}{\lambda}\right)\right) p(\alpha)$$





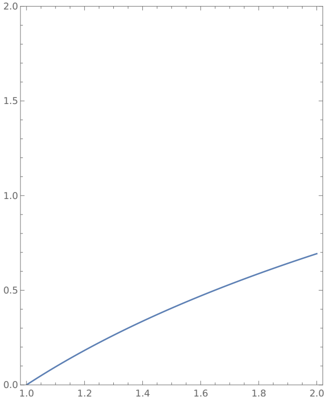
Lemma
There exists $\psi_0 \in (0,1)$ such that

$$\Pr\left(Y_t \geq \max\left\{\psi_0\lambda, \frac{Y_0}{e^{\Omega(t)}}\right\}\right) \leq te^{-\Omega(\lambda)}$$

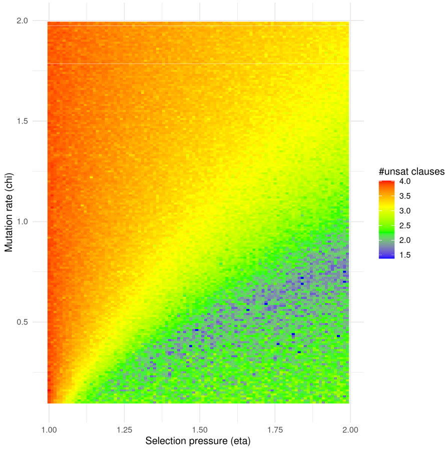
Lemma
If $Y_t \leq \psi\lambda$, $X_t \leq \gamma_0\lambda$, and p_0 "large"

$$\Pr\left(y \in \textit{fitness valley}\right) \geq \frac{X_t}{\lambda}(1+\delta)$$

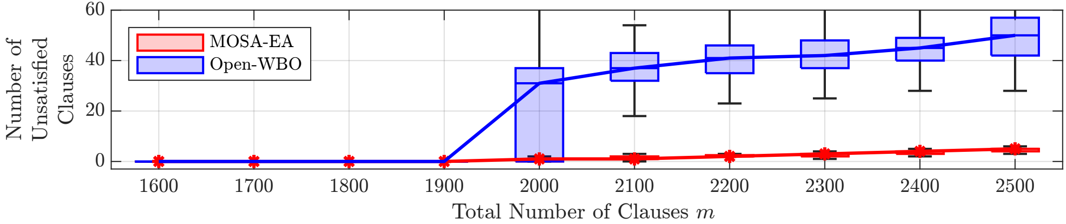
Experimental Comparison: Random 5-MAXSAT $(n = 100, m = 2500)$



$$\chi = \ln(\eta(1-\delta))$$

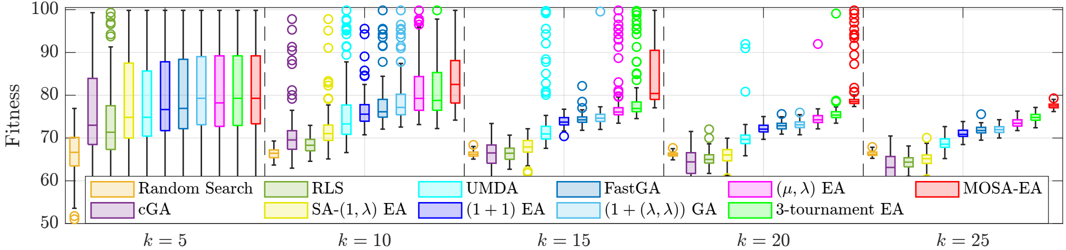


MOSA-EA vs OpenWBO[9] on random 5-MAXSAT ($n = 100$)



[9]Winner of MAXSAT competitions in 2014, 2015, 2016, 2017.

MOSA-EA on NK-Landscape ($n = 100$, 10^8 fevals)



Spiky (or frozen noise) $\textsc{Onemax}$

Let $\boldsymbol{B \sim \mathcal{V}(n,p)}$, i.e., $\mathbf{Pr}\left(\boldsymbol{x \in B}\right) = \boldsymbol{p}$ for all $\boldsymbol{x}$, then

$$\textsc{DisOM}_{\boldsymbol{d,p}}(\boldsymbol{x}) = \begin{cases} \textsc{Onemax}(\boldsymbol{x}) + \boldsymbol{d} & \text{if } \boldsymbol{x \in B} \\ \textsc{Onemax}(\boldsymbol{x}) & \text{otherwise} \end{cases}.$$

Spiky (or frozen noise) OneMax

Let $B \sim \mathcal{V}(n, p)$, i.e., $\mathbf{Pr}(x \in B) = p$ for all x, then

$$\textsc{DisOM}_{d,p}(x) = \begin{cases} \textsc{OneMax}(x) + d & \text{if } x \in B \\ \textsc{OneMax}(x) & \text{otherwise} \end{cases}.$$

Theorem (Jorritsma et al. [2023])

(1,λ) EA a factor of $1/p$ faster than (1+λ) EA

- *in finding $n - k^*$ fitness target*
- *assuming $p \in \omega(1/(n \log n)) \cap n^{-\Omega(1)}$ and $k^* \in O(n^c)$.*

Spiky (or frozen noise) OneMax

Some issues with $\textsc{DisOM}_{d,p}$

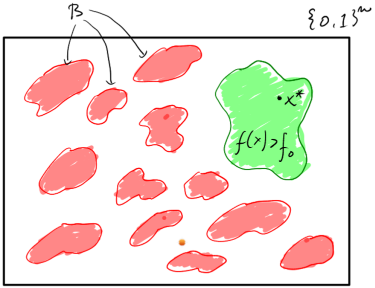
- Optimum of $\textsc{DisOM}_{d,p}$ not necessarily all ones bitstring.
- Fitness in global optimum is random.

Our version: for some $f_0 < n$, let $B \sim \mathcal{V}(n, p)$, then

$$\textsc{DisOM'}_{d,p}(x) = \min(f_0, \textsc{DisOM}_{d,p}(x))$$

- The all ones bitstring is the global optimum with fitness n.

Robustness to Perturbations

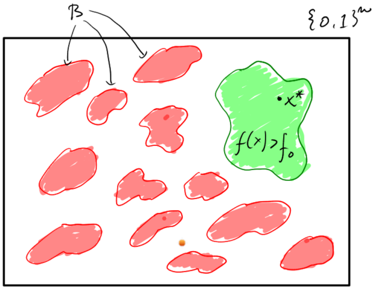


Definition

For $f : \{0,1\}^n \to \mathbb{R}$, $f_0 < \max_x f(x)$, and $B \subset \{0,1\}^n$, define $\mathcal{G}(f, f_0, B)$ to be the set of functions g such that

- $\forall x \notin B$, $g(x) = f(x)$
- $\forall x \in B$, $g(x) \leq f_0$

Robustness to Perturbations



Theorem

For all (δ, r)-sparse B and $f \in \mathrm{SLO}^{\alpha}_{\varepsilon}$ with $f_0 < \max_x f(x)$

$$\mathcal{G}(f, f_0, B) \subseteq \mathrm{SLO}^{\alpha-\delta}_{\varepsilon+\delta, r}$$

assuming that for all levels j with $\min_{x \in A_j} f(x) \leq f_0$

- A_j *is* α*-dense with respect to* $A_{\geq j}$
- $H(x, A_{\geq j+1} \setminus B) = O(1)$ *for all* $x \in A_j \setminus B$

Corollary

Let $f \sim \textsc{DisOM'}_{d,p}$ *for some const.* p, *then w.o.p.,*
the non-elitist EA with 3-tmt selection, and mutation rate $\chi = 1.09812$, *and population size* $\lambda = \Omega(\log n)$
(or any non-elitist EA satisfying the conditions of Theorem XX)
has expected runtime $O(n\lambda)$ *on* f.

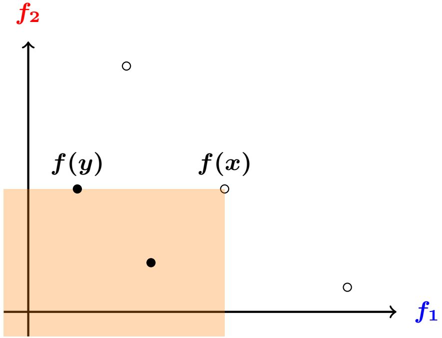
Proof idea

- $\mathcal{V}(n,p)$ for $p \leq \left(\frac{\delta}{2e}\right)^{1+\frac{1+\alpha}{\delta}}$ is (δ, r)-sparse with probability $1 - ne^{-\Omega(\alpha n)}$
- $\textsc{DisOM}_{d,p}{}' \in \mathcal{G}(\textsc{Onemax}, (1-\alpha)n, \mathcal{V}(n,p))$
- Result follows from Theorem XX

Multi-Objective Optimisation

Maximising $f := (f_1, f_2)$

x dominates y (i.e. $x \succ y$) $\Leftrightarrow f_1(x) \geq f_1(y), f_2(x) \geq f_2(y),$ and at least one inequality is strict.

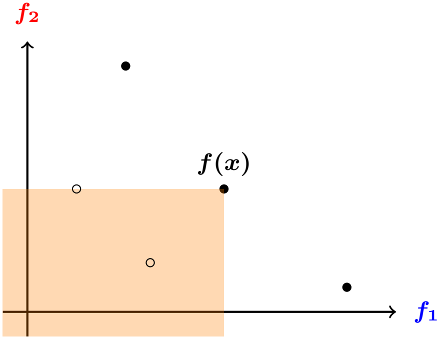


Multi-Objective Optimisation

Maximising $\boldsymbol{f} := (\boldsymbol{f_1}, \boldsymbol{f_2})$

$\boldsymbol{x}$ dominates $\boldsymbol{y}$ (i.e. $\boldsymbol{x} \succ \boldsymbol{y}$) $\Leftrightarrow \boldsymbol{f_1}(\boldsymbol{x}) \geq \boldsymbol{f_1}(\boldsymbol{y}), \boldsymbol{f_2}(\boldsymbol{x}) \geq \boldsymbol{f_2}(\boldsymbol{y})$, and at least one inequality is strict.

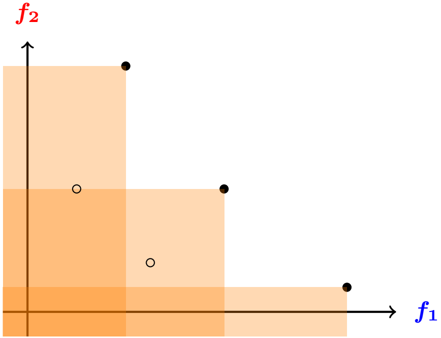
$$\mathbf{ND}_{\boldsymbol{f}}\,(\boldsymbol{x}) := \{\boldsymbol{y} \mid \neg(\boldsymbol{x} \succ \boldsymbol{y})\}$$



Multi-Objective Optimisation

Maximising $f := (f_1, f_2)$

x dominates y (i.e. $x \succ y$) $\Leftrightarrow f_1(x) \geq f_1(y), f_2(x) \geq f_2(y),$ and at least one inequality is strict.



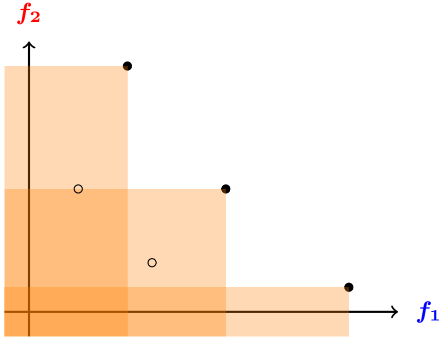
$\mathrm{ND}_f\,(x) := \{y \mid \neg(x \succ y)\}$
$F_f := f(\bigcap_{x\in\{0,1\}^n} \mathrm{ND}_f\,(x))$ is the **Pareto front**.

Goal: Find a population P that can cover F_f.
(Reduce P to a set $\longrightarrow$ we have the **Pareto set** P)

Multi-Objective Optimisation

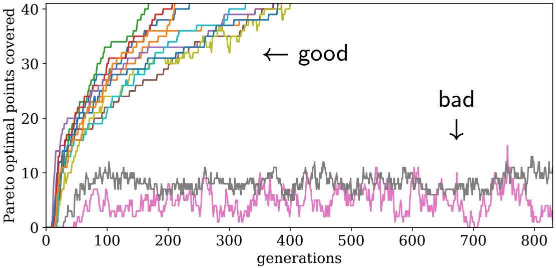
Maximising $f := (f_1, f_2)$

x dominates y (i.e. $x \succ y$) $\Leftrightarrow$ $f_1(x) \geq f_1(y)$, $f_2(x) \geq f_2(y)$, and at least one inequality is strict.



$\mathrm{ND}_f(x) := \{y \mid \neg(x \succ y)\}$
$F_f := f(\bigcap_{x \in \{0,1\}^n} \mathrm{ND}_f(x))$ is the **Pareto front**.

Goal: Find a population P that can cover F_f.
(Reduce P to a set $\longrightarrow$ we have the **Pareto set** P)
Bad algorithms can lose the coverage or take long time.

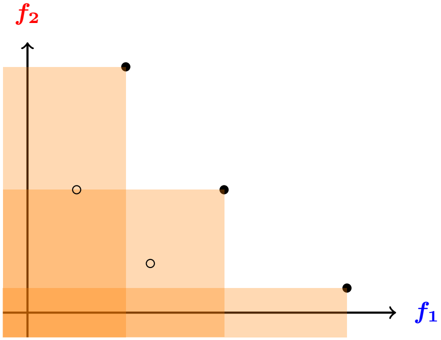


Multi-Objective Optimisation

Maximising $f := (f_1, f_2)$

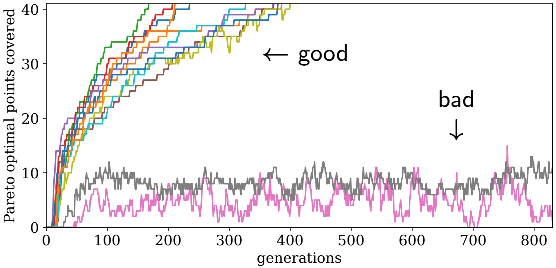
x dominates y (i.e. $x \succ y$) $\Leftrightarrow$ $f_1(x) \geq f_1(y)$, $f_2(x) \geq f_2(y)$, and at least one inequality is strict.

Hard to compare sol. as #(obj) increases



$\mathrm{ND}_f(x) := \{y \mid \neg(x \succ y)\}$
$F_f := f(\bigcap_{x \in \{0,1\}^n} \mathrm{ND}_f(x))$ is the **Pareto front**.

Goal: Find a population P that can cover F_f.
(Reduce P to a set $\longrightarrow$ we have the **Pareto set** P)
Bad algorithms can lose the coverage or take long time.



Multi-Objective Evolutionary Algorithms

Simple algorithms in early theory studies:

- (G)SEMO [Laumanns, Thiele, and Zitzler, 2004, Giel and Lehre, 2010] (300+ citations)
  Create one offspring in each generation, then keep all non-dominated solutions and avoid duplicate fitness.

Nowadays, we can analyse practical algorithms (seminal work [Zheng, Liu, and Doerr, 2022]):

- SPEA2 [Zitzler, Laumanns, and Thiele, 2001] (10100+ citations)
- NSGA-II [Deb, Pratap, Agarwal, and Meyarivan, 2002] (56000+ citations)
- SMS-EMOA [Beume, Naujoks, and Emmerich, 2007] (2100+ citations)
- MOEA/D [Zhang and Li, 2007] (8800+ citations)
- NSGA-III [Deb and Jain, 2014] (5500+ citations)

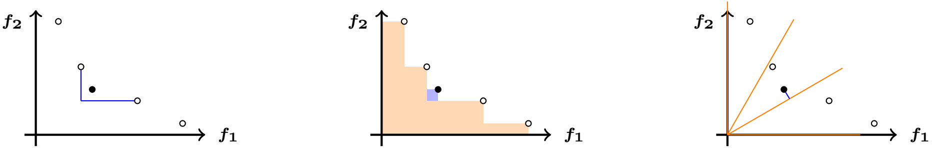
Multi-Objective Evolutionary Algorithms

Simple algorithms in early theory studies:

- (G)SEMO [Laumanns et al., 2004, Giel and Lehre, 2010]
  Create one offspring in each generation, then keep all non-dominated solutions and avoid duplicate fitness.

Nowadays, we can analyse practical algorithms (seminal work [Zheng et al., 2022]):

- SPEA2 [Zitzler et al., 2001]
- NSGA-II [Deb et al., 2002]
- SMS-EMOA [Beume et al., 2007]
- MOEA/D [Zhang and Li, 2007]
- NSGA-III [Deb and Jain, 2014]
  They often use some density estimation complementing dominance relations to make decisions.



(a) Crowding distance **(b)** Hypervolume contribution **(c)** Distance to reference rays

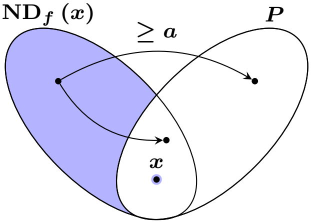
Why? What are the drawbacks of GSEMO-like algorithms?

Sparsity and Isolation

$f : \{0,1\}^n \to \mathbb{R}^d$ is **(a,b)-Pareto-sparse** if for all Pareto sets P [Dang, Opris, and Sudholt, 2026a]:

- **(C1)** For every $x, y \in P$ with $f(x) \neq f(y)$: $\mathrm{H}(y, \{x\} \cup (\mathrm{ND}_f(x) \setminus P)]) \geq a$.
- **(C2)** For every $v \in f(P)$: $|\{y \in \{0,1\}^n \mid f(y) = v\}| \leq b$.

C1 models some obstacles for GSEMO as it only keeps non-dominated solutions
$\implies$ *covering every new point on F_f requires to overcome a Hamming distance at least a.*



Sparsity and Isolation

$f\colon\{0,1\}^n \to \mathbb{R}^d$ is **(a,b)-Pareto-sparse** if for all Pareto sets P [Dang et al., 2026a]:

(C1) For every $x, y \in P$ with $f(x) \neq f(y)$: $\mathrm{H}(y, \{x\} \cup (\mathrm{ND}_f(x) \setminus P)]) \geq a$.

(C2) For every $v \in f(P)$: $|\{y \in \{0,1\}^n \mid f(y) = v\}| \leq b$.

C1 models some obstacles for GSEMO as it only keeps non-dominated solutions
$\implies$ *covering every new point on F_f requires to overcome a Hamming distance at least a.*

Sparsity requires (C1) to hold for every $u, v \in F_f$, here $u = f(x), v = f(y)$ in the condition.
$\implies$ **(a,b)-Pareto-isolated** relaxes this, requiring (C1) to hold for only one $u \in F_f$ (see paper).

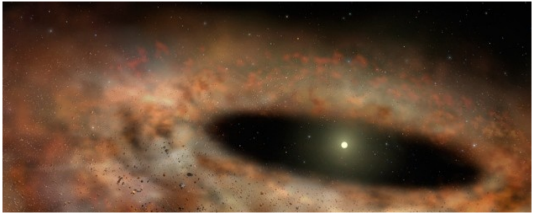


Photo: R. Skibba published in Nature, 2017.

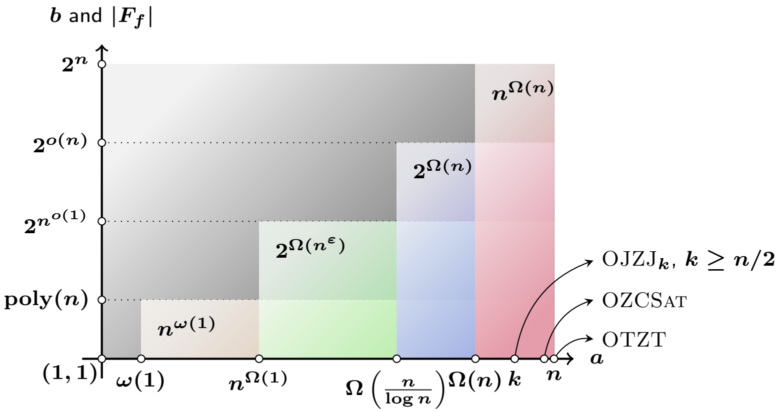
Hierarchy

Lemma

1. *Every function* $f\colon \{0,1\}^n \to \mathbb{R}^d$ *is* $(1,2^n)$*-Pareto-sparse and* $(1,2^n)$*-Pareto-isolated.*
2. f *is* (a,b)*-Pareto-sparse* $\Longrightarrow$ f *is* (a',b')*-Pareto-sparse for any* $a' \leq a$ *and any* $b' \geq b$*.*
3. f *is* (a,b)*-Pareto-isolated* $\Longrightarrow$ f *is* (a',b')*-Pareto-isolated for any* $a' \leq a$ *and any* $b' \geq b$*.*
4. f *is* (a,b)*-Pareto-sparse* $\Longrightarrow$ f *is* (a,b)*-Pareto-isolated.*

Examining the set of all Pareto optimal solutions $\bigcap_{x\in\{0,1\}^n} \mathrm{ND}_f(x)$ allows computing a, b exactly.

Problem Hardness for GSEMO



Theorem

GSEMO requires $H_{|F_f|-1} n^a / b$ expected fitness eval. to optimise any (a,b)-Pareto-sparse[10] function.
NSGA-II/III, SMS-EMOA can solve OTZT in expected $\mathcal{O}(n \log n)$ with dup. elimination.

[10] For (a,b)-Pareto-sparse, the harmonic factor $H_{|F_f|-1}$ is simply removed.

Relying Solely on Dominance Relations is Futile

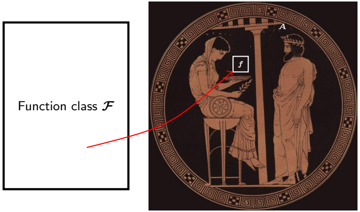


Photo: E. Gerhard (1846).

Relying Solely on Dominance Relations is Futile

For every $1 \leq i, j \leq t$, can tell one of

1. $x_i \succ x_j$
2. $x_j \succ x_i$
3. $f(x_i) = f(x_j)$
4. none of the above

is TRUE

$x_1, x_2, x_3, \ldots, x_t$

Function class $\mathcal{F}$

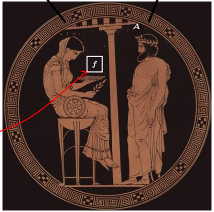


Photo: E. Gerhard (1846).

Relying Solely on Dominance Relations is Futile

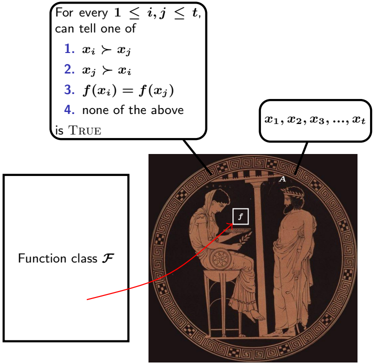


Photo: E. Gerhard (1846).

$\exists$ a function class called $\textsc{OZLinTwoOpt} =: \mathcal{F}$,

Theorem ([Dang et al., 2026b])

The BB complexity[a] on $\mathcal{F}$ of all algo. relying solely on dominance rel. to make decisions is $\geq 2^{n-3} + 1/2$.

NSGA-II/III, SMS-EMOA solve[b] any $f \in \mathcal{F}$ in $\mathcal{O}\left(n^2\right)$.

---

[a] Only to find any Pareto-optimal solution.

[b] Stronger meaning, to find a first Pareto solution then cover the front.

Relying Solely on Dominance Relations is Futile

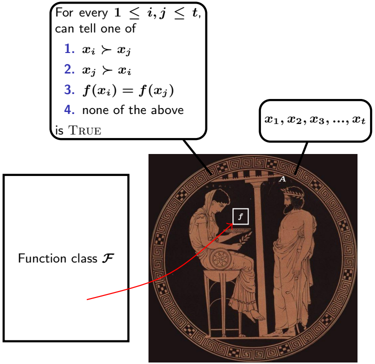


Photo: E. Gerhard (1846).

$\exists$ a function class called $\mathrm{OZL{\small IN}T{\small WO}O{\small PT}} =: \mathcal{F}$,

Theorem ([Dang et al., 2026b])

The BB complexity[a] on $\mathcal{F}$ of all algo. relying solely on dominance rel. to make decisions is $\ge 2^{n-3} + 1/2$.

NSGA-II/III, SMS-EMOA solve[b] any $f \in \mathcal{F}$ in $\mathcal{O}\left(n^2\right)$.

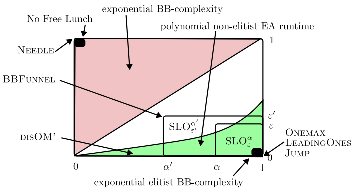
[a] Only to find any Pareto-optimal solution.

[b] Stronger meaning, to find a first Pareto solution then cover the front.

Facts and insights:

- Every $f \in \mathcal{F}$ is $(1{,}1)$-Pareto-sparse; Easy to sample f using any continuous distribution.
- Practical algorithms without their density estimations fit the black-box model.
- Algorithms in this model fail because of a lack of search focus to reach the Pareto front.

Conclusion



- Introduced problem hierarchy $\mathrm{SLO}^{\boldsymbol{\alpha}}_{\boldsymbol{\varepsilon}}$ containing all pseudo-Boolean functions
- Easy functions have low $\boldsymbol{\varepsilon}$ and high $\boldsymbol{\alpha}$
  - $\textsc{OneMax}, \textsc{LeadingOnes}, \textsc{Jump} \in \mathrm{SLO}^{\mathbf{1}}_{\mathbf{0}}$
- Hard functions have high $\boldsymbol{\varepsilon}$ and low $\boldsymbol{\alpha}$
  - Exponential BB complexity when $\boldsymbol{\varepsilon} > \boldsymbol{\alpha}$
  - Classes satisfying NFL in $\mathrm{SLO}^{\boldsymbol{\alpha}}_{\boldsymbol{\varepsilon}}$ with $\boldsymbol{\varepsilon} \approx \mathbf{1}$, $\boldsymbol{\alpha} \approx \mathbf{0}$
- Benefit of Non-Elitism
  - Elitist BB complexity is exponential for $\boldsymbol{\alpha} < \mathbf{1} - \boldsymbol{\delta}$ and $\boldsymbol{\varepsilon} > \boldsymbol{\delta}$
  - Sufficient conditions for polynomial runtime of non-elitist EAs Non-linear selection CDF, appropriate sel.pres and mut rate
  - $\textsc{DisOM'}_{\boldsymbol{d,p}}$ for $\boldsymbol{p} \in \boldsymbol{\Omega}(\mathbf{1})$ can be optimised in $\boldsymbol{O}(\boldsymbol{n} \log \boldsymbol{n})$
- Drawback of keeping only non-dominated solutions
  - $(\boldsymbol{a},\boldsymbol{b})$-Pareto-sparse modelling the obstacles in navigating the Pareto front
  - Hard functions for GSEMO have high $\boldsymbol{a}$, $|\boldsymbol{F_f}|$ and low $\boldsymbol{b}$
- Lack of search focus when relying solely on dominance to make decisions

Outlook

Open questions

- What happens in the "white region" of the SLO-hierarchy?
  - Can we design better algorithms, increasing the green region?
  - Can we prove more precise lower bounds, increasing the pink region?
- Suppose[11] f belongs to both $\mathrm{SLO}_{\varepsilon}^{\alpha}$ and $\mathrm{SLO}_{\varepsilon'}^{\alpha'}$ where $\alpha < \alpha'$ and $\varepsilon < \varepsilon'$, does f also belong to $\mathrm{SLO}_{\varepsilon'}^{\alpha}$?
- SLO for a general search operator? On different search spaces?
- Rigorous analysis of $\mathrm{OJZJ}_k, k \geq n/2$, and $\mathrm{OZCS}\textsc{at}$?

Outlook in this PPSN (both on Monday)

- Landscape characteristics for Dominating Set & Vertex Colouring. (Session 2)
- Weakness of density estimation of SPEA2, OTZT versus $\mathrm{OZCS}\textsc{at}$. (Session 3)

And your questions?

[11] This problem was suggested to us by Jon Rowe.

References I

Nicola Beume, Boris Naujoks, and Michael T. M. Emmerich. SMS-EMOA: multiobjective selection based on dominated hypervolume. *European Journal of Operational Research*, 181(3):1653–1669, 2007.

Dogan Corus, Duc-Cuong Dang, Anton V. Eremeev, and Per Kristian Lehre. Level-Based Analysis of Genetic Algorithms and Other Search Processes. *IEEE Transactions on Evolutionary Computation*, 22(5):707–719, October 2018. ISSN 1089-778X, 1089-778X, 1941-0026. doi: 10.1109/TEVC.2017.2753538. URL https://ieeexplore.ieee.org/document/8039236/.

Duc-Cuong Dang, Andre Opris, and Dirk Sudholt. Runtime analysis of functions where widely used evolutionary multi-objective algorithms beat simple ones. *ACM Trans. Evol. Learn. Optim.*, 6(2):11:1–11:33, 2026a.

Duc-Cuong Dang, Andre Opris, and Dirk Sudholt. Why dominance is not enough: Lessons learned from practical. *Algorithmica*, page (To appear), 2026b.

Kalyanmoy Deb and Himanshu Jain. An evolutionary many-objective optimization algorithm using reference-point-based nondominated sorting approach, part i: Solving problems with box constraints. *IEEE Transactions on Evolutionary Computation*, 18(4):577–601, 2014.

Kalyanmoy Deb, Amrit Pratap, Sameer Agarwal, and T. Meyarivan. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation*, 6(2):182–197, 2002.

Carola Doerr and Johannes Lengler. Introducing Elitist Black-Box Models: When Does Elitist Behavior Weaken the Performance of Evolutionary Algorithms? *Evolutionary Computation*, 25(4):587–606, October 2016. ISSN 1063-6560. doi: 10.1162/evco_a_00195. URL https://doi.org/10.1162/evco_a_00195. Publisher: MIT Press.

Stefan Droste, Thomas Jansen, and Ingo Wegener. Optimization with randomized search heuristics–the (A)NFL theorem, realistic scenarios, and difficult functions. *Theoretical Computer Science*, 287(1):131–144, September 2002. ISSN 0304-3975. doi: 10.1016/S0304-3975(02)00094-4. URL http://www.sciencedirect.com/science/article/pii/S0304397502000944.

Stefan Droste, Thomas Jansen, and Ingo Wegener. Upper and lower bounds for randomized search heuristics in black-box optimization. *Theory of Computing Systems*, 39(4):525–544, 2006.

Paul Erdős and Alfred Renyi. On two problems of information theory. *A Magyar Tudományos Akadémia Matematikai Kutató Intézetének Közleményei*, 8:229–243, 1963.

Oliver Giel and Per Kristian Lehre. On the effect of populations in evolutionary multi-objective optimisation. *Evolutionary Computation*, 18(3):335–356, 2010.

References II

Joost Jorritsma, Johannes Lengler, and Dirk Sudholt. Comma Selection Outperforms Plus Selection on OneMax with Randomly Planted Optima. In *Proceedings of the Genetic and Evolutionary Computation Conference*, GECCO '23, pages 1602–1610, New York, NY, USA, July 2023. Association for Computing Machinery.

Marco Laumanns, Lothar Thiele, and Eckart Zitzler. Running time analysis of multiobjective evolutionary algorithms on pseudo-boolean functions. *IEEE Transactions on Evolutionary Computation*, 8(2):170–182, 2004.

Ingo Wegener. Methods for the analysis of evolutionary algorithms on pseudo-boolean functions. In Ruhul Sarker, Masoud Mohammadian, and Xin Yao, editors, *Evolutionary Optimization*, pages 349–369. Kluwer, 2002.

D. H. Wolpert and W. G. Macready. No free lunch theorems for optimization. *IEEE Transactions on Evolutionary Computation*, 1(1):67–82, 1997.

Qingfu Zhang and Hui Li. MOEA/D: A multiobjective evolutionary algorithm based on decomposition. *IEEE Transactions on Evolutionary Computation*, 11(6):712–731, 2007.

Shengyu Zhang. Tight Bounds for Randomized and Quantum Local Search. *SIAM Journal on Computing*, 39(3):948–977, January 2009. ISSN 0097-5397, 1095-7111. doi: 10.1137/06066775X. URL `http://epubs.siam.org/doi/10.1137/06066775X`.

Weijie Zheng, Yufei Liu, and Benjamin Doerr. A first mathematical runtime analysis of the non-dominated sorting genetic algorithm II (NSGA-II). In *Proceedings of the AAAI Conference on Artificial Intelligence, AAAI 2022*, pages 10408–10416. AAAI Press, 2022.

Eckart Zitzler, Marco Laumanns, and Lothar Thiele. SPEA2: Improving the strength pareto evolutionary algorithm. Technical report, ETH Zurich, Computer Engineering and Networks Laboratory, 2001.