Analyzing Hypervolume Indicator Based Algorithms

ETH Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

max planck institut informatik

MOTIVATION

recent trend in multiobjective evolutionary algorithms (MOEAs):
explicit incorporation of user preferences by using indicators

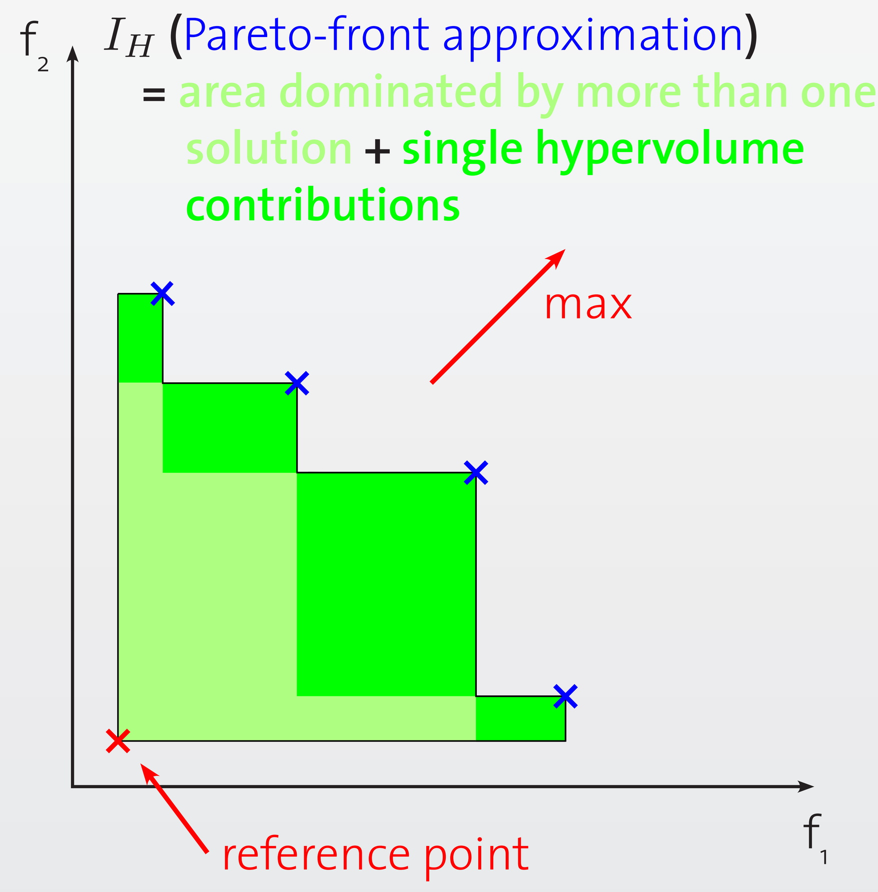
hypervolume indicator based MOEAs showed better performance in experiments than classical MOEAs

why?

Goals:
understand why hypervolume-based search is that successful
understand basic properties of hypervolume indicator

Approach:
rigorous running time analyses of a hypervolume-based MOEA for
(i) approaching the Pareto front
(ii) approximating large Pareto fronts

THE HYPERVOLUME INDICATOR

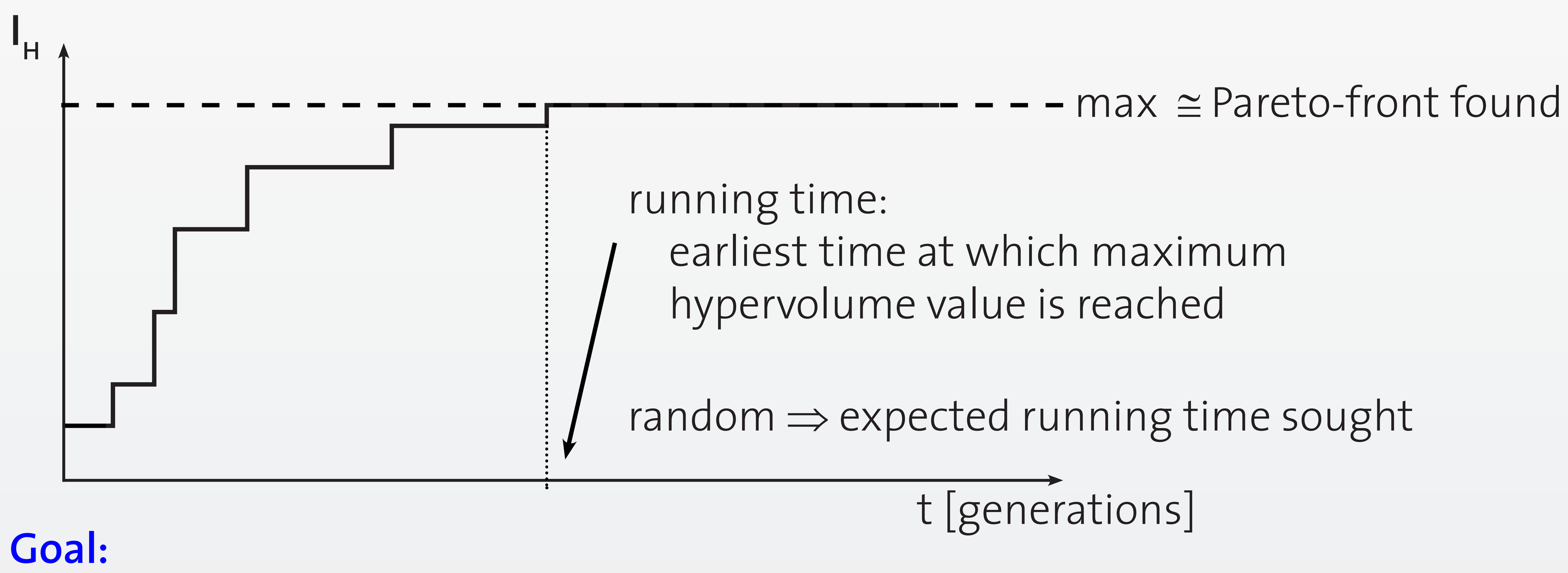


(unary) hypervolume indicator

I_H (A) = hypervolume/area of dominated part of search space between front A and reference point

Pareto-dominance compliant:
finding the Pareto front
⇔ maximizing I_H [flei2003a]

RUNNING TIME ANALYSES



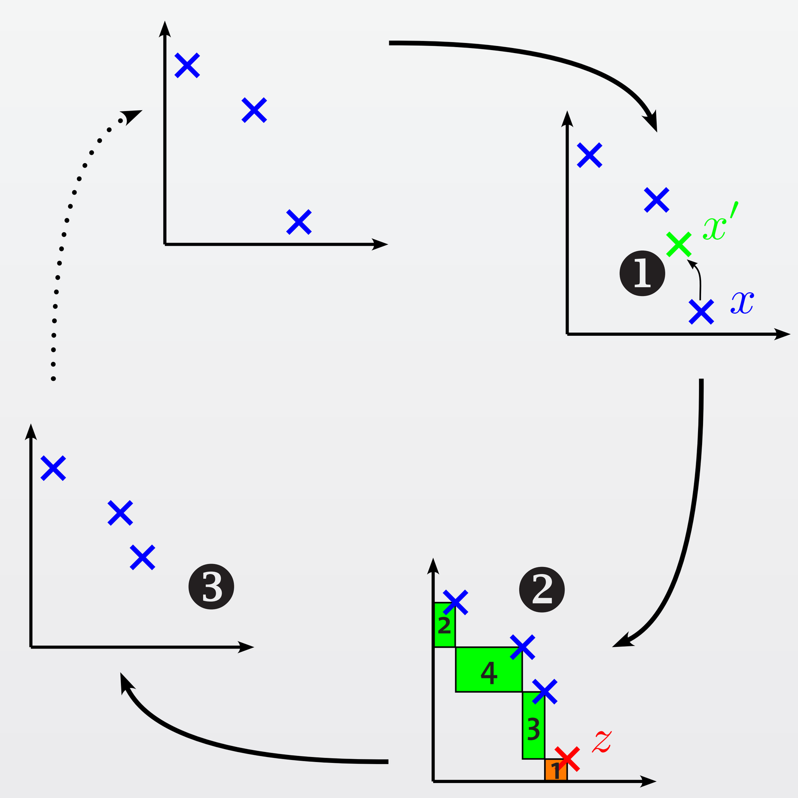
Goal:
give upper bound for expected running time until Pareto-front is reached or approximated

Here:
only 2 objectives; w.l.o.g. maximization

Ideas:
- consider no worsening in I_H
- if a set of solutions is dominated by another set ⇒ hypervolume indicator value is higher for the latter
- local improvement is possible if single point is placed optimally with respect to its neighbors

(μ+1) SIBEA

SIBEA: Simple Indicator-Based EA [zbt2007a]
(μ+1)-selection also used in SMS-EMOA [bne2007a] and MO-CMA-ES [ihr2007a]



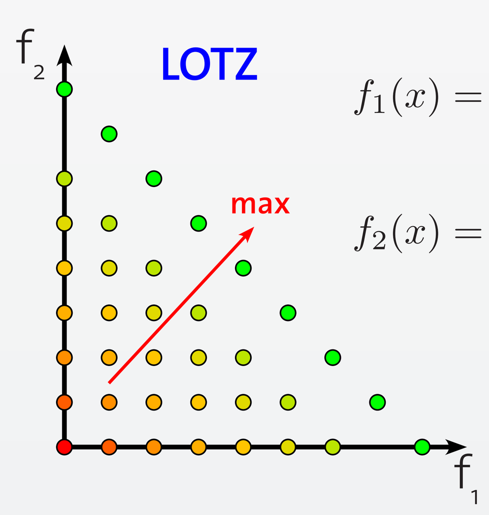
(μ+1)SIBEA

generate initial population $P \subseteq \{0,1\}^n$ at random

repeat:
❶ mutate randomly selected $x \in P$ to x' by flipping each bit of x with probability $1/n$
$P' = P \cup \{x'\}$
❷ for all solutions $x \in P$, determine the hypervolume loss $d(x) = I_H(P') - I_H(P' \setminus \{x\})$
❸ choose a $z \in P$ with smallest loss $d(z)$
$P = P' \setminus \{z\}$

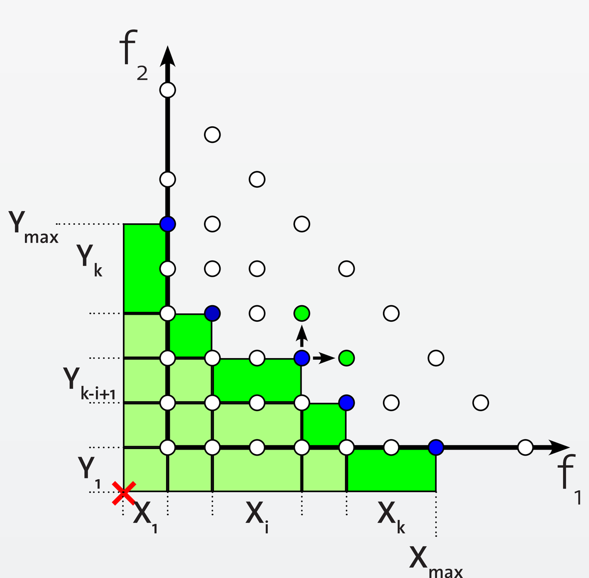
Properties:
- no worsenings of I_H over time
- duplicated solutions are removed first
- in general, no global convergence to Pareto-front! [ztb2008a]

APPROACHING THE PARETO FRONT



$$f_1(x) = \mathrm{LO}(x) = \sum_{i=1}^{n} \prod_{j=1}^{i} x_j$$

$$f_2(x) = \mathrm{TZ}(x) = \sum_{i=1}^{n} \prod_{j=i}^{n} (1 - x_j)$$

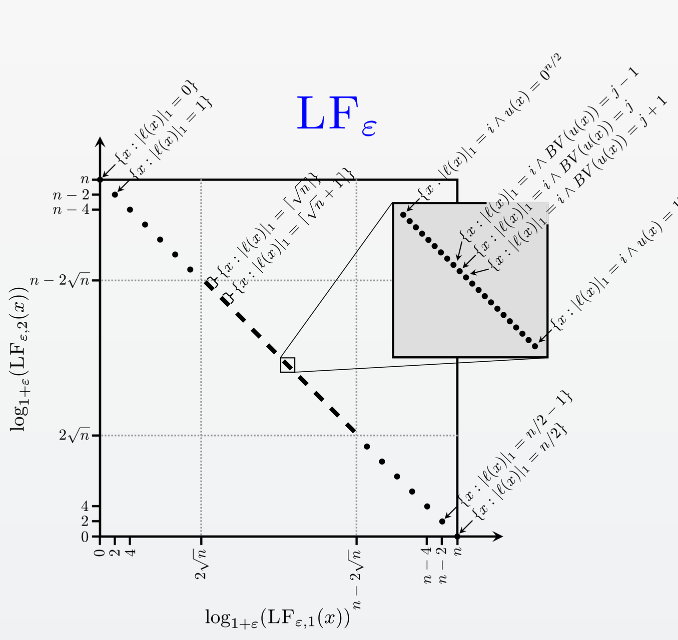


Theorem: Choosing $\mu \geq n+1$, the (μ+1)SIBEA optimizes LOTZ in $\mathcal{O}(\mu n^2)$ generations.

Sketch of Proof:
wlog, reference point is (-1,-1) and $\{x_1, \ldots, x_n\}$ are the non-dominated solutions in P
$2k$ possible mutations that increases I_H with prob. $\frac{1}{\mu} \cdot \frac{1}{n}(1-1/n)^{n-1} \geq \frac{1}{e\mu n}$ each
total increase of all mutations is at least $\max\{X_{\max}, Y_{\max}\} \geq \sqrt{X_{\max} \cdot Y_{\max}} \geq \sqrt{I_H}$
expected increase of 1 mutation is therefore $\geq \sqrt{I_H}/2k$; with Markov, the increase of I_H in $8k$ good mutations is $\sqrt{I_H}$ w.h.p.
expected running time for an increase of $\sqrt{I_H}$ is $\mathcal{O}(\frac{\mu n}{2k} \cdot 8k) = \mathcal{O}(\mu n)$
by induction, $\mathcal{O}(n)$ increases by $\sqrt{I_H}$ are sufficient to reach the front
once on the front, SIBEA needs time $\mathcal{O}(\mu n)$ to find one of the at most n non-visited Pareto-optimal points

Conclusion: For $\mu = \Theta(n)$, SIBEA is as fast as global SEMO [giel2003a] although the population contains more than one solution when approaching the front.

APPROXIMATING LARGE PARETO FRONTS

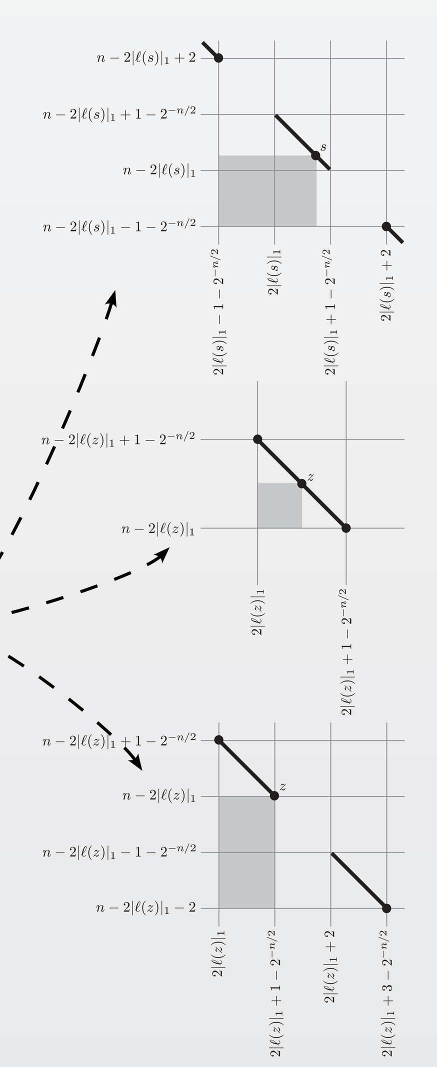


$$f_1(x) = \mathrm{LF}_{\varepsilon,1}(x) := \begin{cases} (1+\varepsilon)^{2 \cdot |\ell(x)|_1 + 2^{-n/2} \cdot \mathrm{BV}(u(x))} & \min\{|\ell(x)|_0, |\ell(x)|_1\} \geq \sqrt{n} \\ (1+\varepsilon)^{2 \cdot |\ell(x)|_1} & \text{otherwise,} \end{cases}$$

$$f_2(x) = \mathrm{LF}_{\varepsilon,2}(x) := \begin{cases} (1+\varepsilon)^{2 \cdot |\ell(x)|_0 + 2^{-n/2} \cdot \mathrm{BV}(\overline{u(x)})} & \min\{|\ell(x)|_0, |\ell(x)|_1\} \geq \sqrt{n} \\ (1+\varepsilon)^{2 \cdot |\ell(x)|_0} & \text{otherwise.} \end{cases}$$

Theorem: Choosing $\mu \geq n/2+3$, the (μ+1)SIBEA finds an ε-approximation of LF_ε within expected $\mathcal{O}(\mu n \log n)$ time.

Sketch of Proof:
wlog, reference point is ((1+ ε)[-1],(1+ ε)[-1]), and we call a solution s with $\{x \in P : |l(x)|_1 = k\} = \{s\}$ sole
we need to prove that in all cases, a sole solution stays in P
an ε-approximation is reached if for all possible k we have at least one solution with $|\ell(x)|_1 = k$ [hn2008a]
prob. to mutate to an x with $|\ell(x)|_1 = b$ is $\geq \frac{1}{\mu}\frac{\min\{b+1, n/2-b+1\}}{en}$
summing up over all possible b yields the theorem



Conclusion: Optimizing the hypervolume allows for a faster search on LF_ε without the need to adjust ε as in [hn2008a].

Dimo Brockhoff[1], Tobias Friedrich[2], and Frank Neumann[2]
[1] ETH Zurich, 8092 Zurich, Switzerland, dimo.brockhoff@tik.ee.ethz.ch
[2] Max-Planck-Institut für Informatik, 66123 Saarbrücken, Germany, firstname.lastname@mpi-inf.mpg.de