# Towards Multi-objective Mixed Integer Evolution Strategies

**Koen van der Blom**, Kaifeng Yang, Thomas Bäck & Michael Emmerich

21-09-2018

Universiteit Leiden
The Netherlands

# Motivation

- Mixed Integer Evolution Strategy [Li et al 2013]
  - Real, integer, categorical
  - Single objective

- Existing multi-objective techniques
  - Weight space decomposition (MILP) [Przybylski et al 2010]
  - Enhanced Directed Search (EDS) [Laredo 2015]
  - Zigzag [Wang 2013, Wang 2015]
  - No distinction between integer and categorical!

- Extend MIES for multiple objectives

# Evolution Strategies

- Mimic evolution for optimisation
  - Parents generate offspring (recombination)
  - Offspring add additional variation (mutation)

- Optimal mutation strength?
  - Changes over time…

- Step size adaptation
  - Same evolutionary mechanisms!

# $(\mu + \lambda)$ Evolution Strategy [Schwefel 1981]

- $\mu$ parents generate $\lambda$ offspring
  - Update decision variables
  - Update mutation probabilities
- Select the best from $\mu \cup \lambda$
- Repeat

$t \leftarrow 0$;
$P_t \leftarrow$ init(); // $P_t \in \mathbb{S}^\mu$ : Set of solutions
**while** $t < t_{max}$ **do**
  // Generate $\lambda$ solutions by (stochastic) variation operators
  $Q_t \leftarrow$ generate($P_t$);
  evaluate($Q_t$);
  $P_{t+1} \leftarrow$ select($Q_t \cup P_t$); // Rank and select $\mu$ best
  $t \leftarrow t + 1$;
**end**

# Mixed Integer Evolution Strategy

[Li et al 2013]

$\tau \leftarrow \frac{1}{\sqrt{2_{n_r}}}, \tau' \leftarrow \frac{1}{\sqrt{2\sqrt{n_r}}}$;
$\sigma' = \sigma \exp(\tau N(0,1))$;
**foreach** $i \in \{1, \ldots, n_r\}$ **do**
  $r'_i \leftarrow r_i + \sigma' N(0,1)$;
  $r'_i \leftarrow T_{r_i^{min}, r_i^{max}}(r'_i)$; `// interval boundary treatment`
**end**

→ **Continuous Variables** (Normal Distribution)

$\tau \leftarrow \frac{1}{\sqrt{2_{n_z}}}, \tau' \leftarrow \frac{1}{\sqrt{2\sqrt{n_z}}}$;
$\varsigma' = \max(1, \varsigma \exp(\tau N(0,1)))$;
**foreach** $i \in \{1, \ldots, n_z\}$ **do**
  $u_1 \leftarrow U(0,1), u_2 \leftarrow U(0,1), \psi \leftarrow 1 - (\varsigma'/n_z)\left(1 + \sqrt{1 + (\frac{\varsigma'}{n_z})^2}\right)^{-1}$;
  $G_1 \leftarrow \left\lfloor \frac{\ln(1-u_1)}{\ln(1-\psi)} \right\rfloor, G_2 \leftarrow \left\lfloor \frac{\ln(1-u_2)}{\ln(1-\psi)} \right\rfloor$;
  $z'_i \leftarrow z_i + G_1 - G_2$;
  $z'_i \leftarrow T_{z_i^{min}, z_i^{max}}(z'_i)$; `// interval boundary treatment`
**end**

→ **Integer Variables** (Geometrical Distribution)

$\tau \leftarrow \frac{1}{\sqrt{2_{n_d}}}, \tau' \leftarrow \frac{1}{\sqrt{2\sqrt{n_d}}}$;
$p' = \frac{1}{1 + \frac{1-p}{p}\exp(-\tau N(0,1))}$;
$p' = T_{1/n_d, 0.5}(p')$;
**foreach** $i \in \{1, \ldots, n_d\}$ **do**
  **if** $U(0,1) < p'$ **then**
    choose a new element uniformly distributed out of $D_i \setminus d_i$;
  **end**
**end**

→ **Categorical Variables**

# Mutation operators

- Properties
  - Scalability (scale step size)
  - Asymmetry (maximal entropy, avoid bias)
  - Infinite support (every solution is reachable)

- Example
  - Mutation of integer variables
    [Rudolph 1994]
  - Difference of two
    Geometric distributions

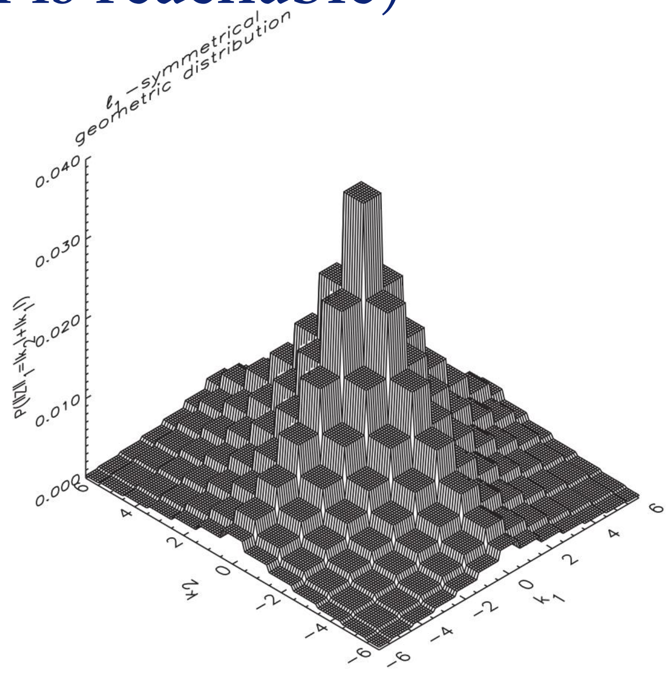
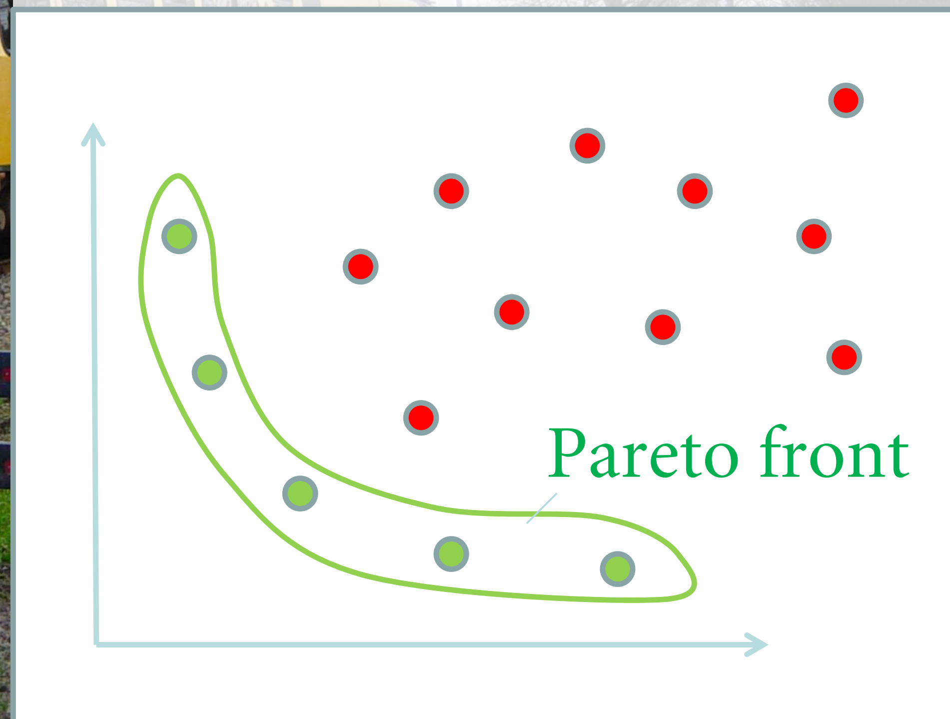



Image from [Li et al 2013]

# Multi-objective optimisation


## Train Routing

- Time Min
- Price Min



# SMS-EMOA [Emmerich et al 2005]

- Optimise hypervolume indicator
- Rank solutions
  - Non-dominated sorting
  - Hypervolume contribution

$P_0 \leftarrow$ init(); `// Initialise random start population of` $\mu$ `individuals`
$t \leftarrow 0$;
**while** *stop critirum not reached* **do**
  $q_{t+1} \leftarrow$ generate($P_t$); `// Generate one offspring by variation operators`
  $Q \leftarrow P_t \cup \{q_{t+1}\}$ ;
  $\{\mathcal{R}_1, \ldots, \mathcal{R}_I\} \leftarrow$ fast-nondominated-sort($Q$); `// All` $I$ `non-dominated fronts of` $Q$
  $r \leftarrow \text{argmin}_{s \in \mathcal{R}_I}[\Delta_{\mathcal{S}}(s, \mathcal{R}_I)]$; `// Detect element of` $\mathcal{R}_I$ `with lowest` $\Delta_{\mathcal{S}}(s, \mathcal{R}_I)$
  $P_{t+1} \leftarrow Q \setminus \{r\}$; `// Eliminate detected element`
  $t \leftarrow t + 1$;
**end**

# Hypervolume indicator

- Measure the dominated region

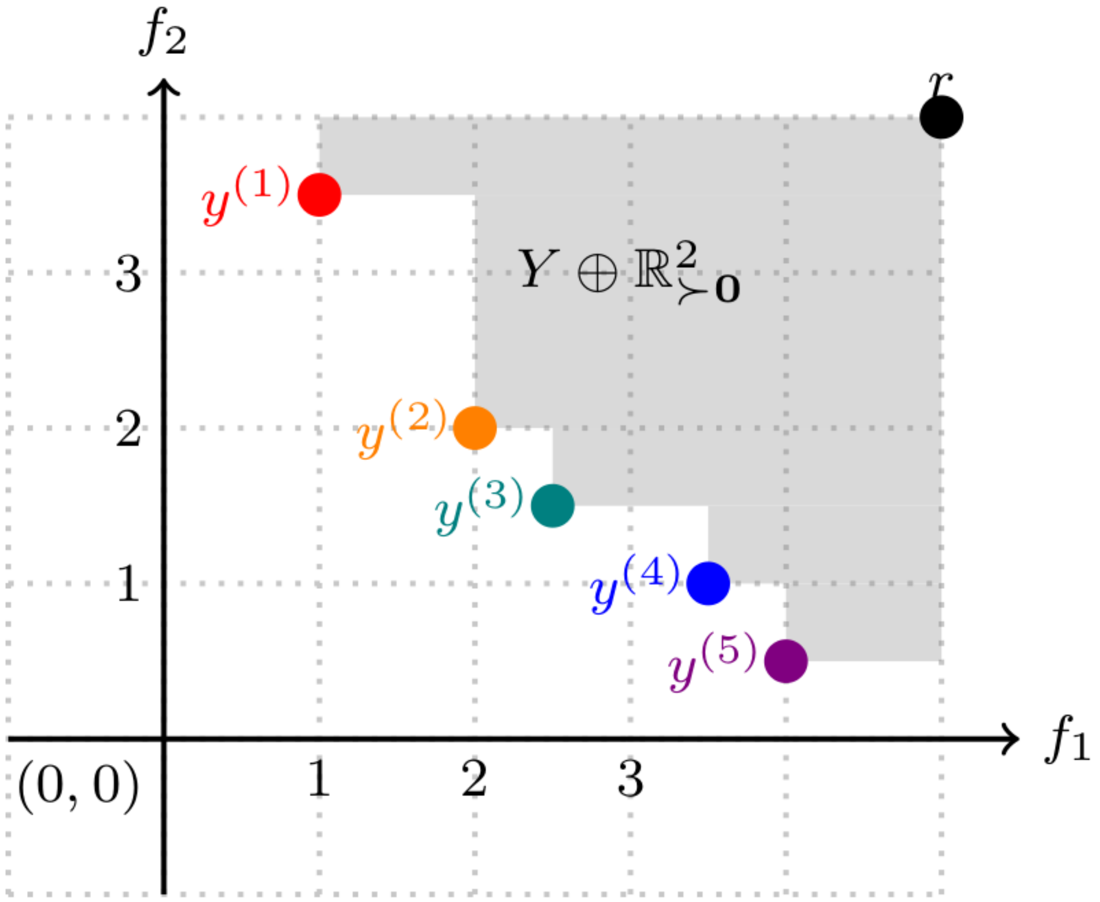



Image from [Emmerich+Deutz 2018]

# *S*-metric (hypervolume) selection

- Hypervolume contribution

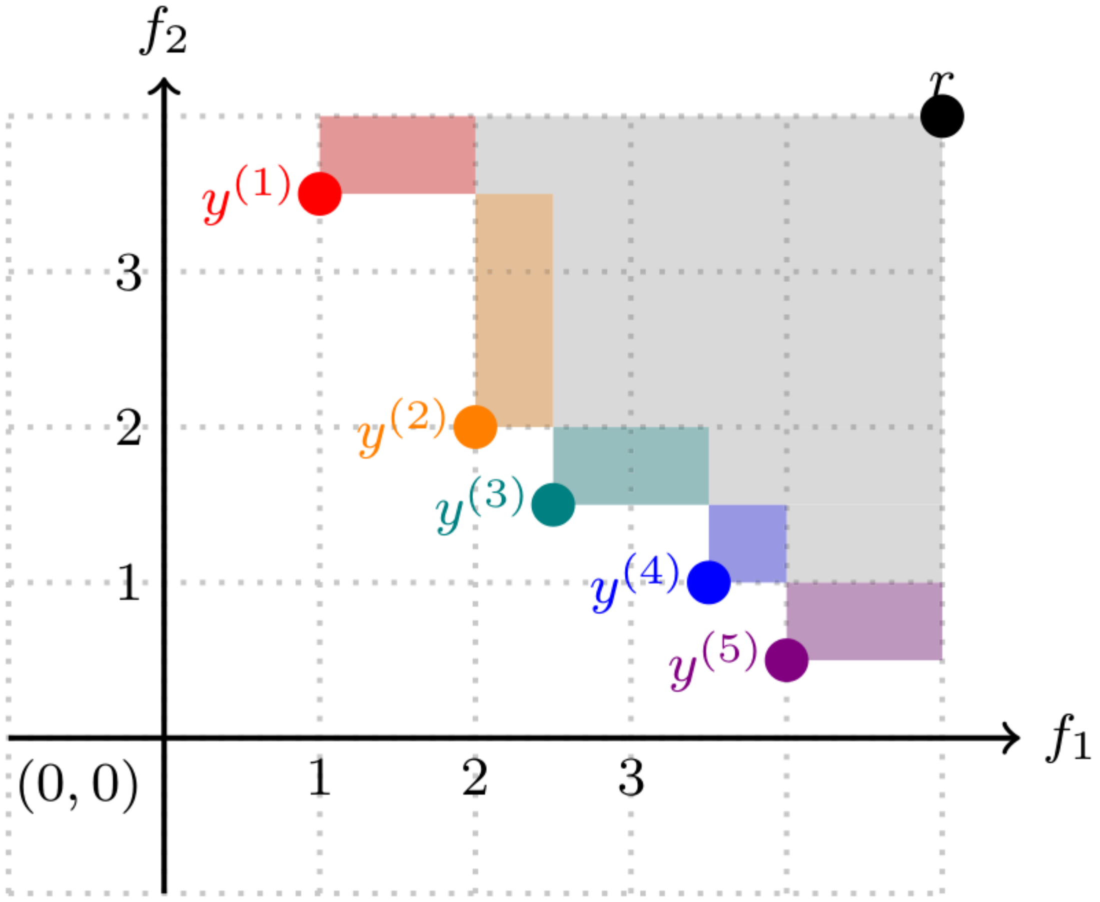
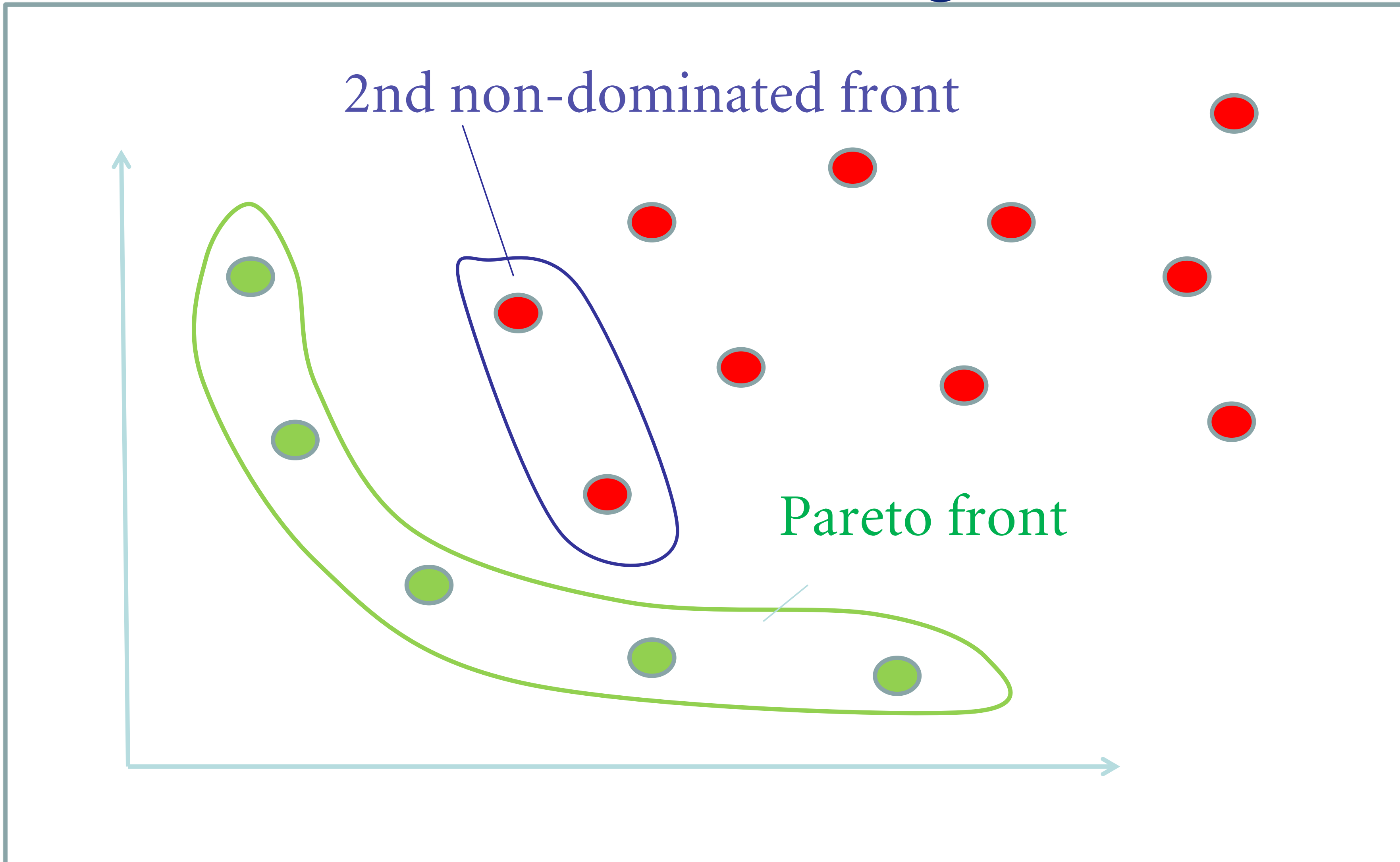



Image from [Emmerich+Deutz 2018]

# Non-dominated sorting

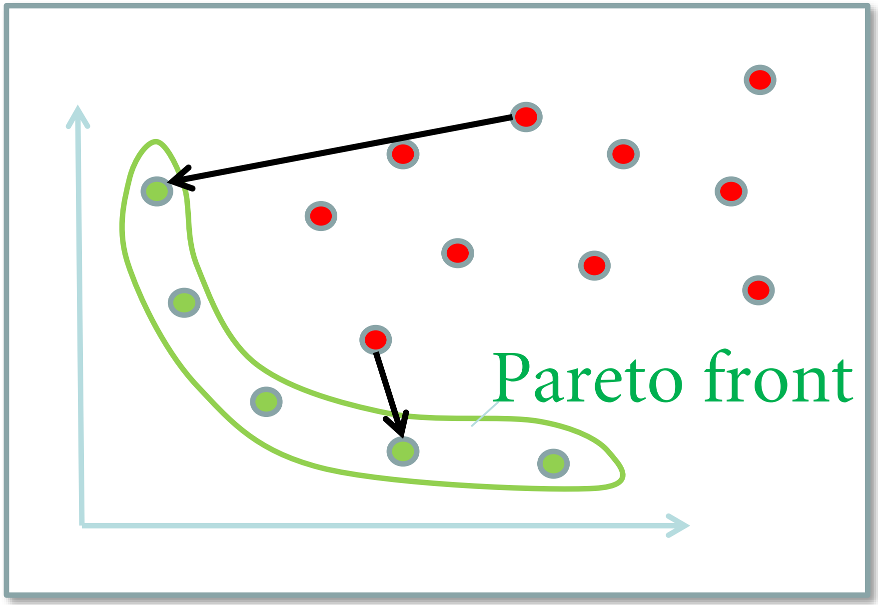


# Multi-Objective MIES

- Canonical MIES operators
- $S$-metric selection
- Non-dominated sorting
- $(\mu + 1)$ strategy
  - Always select the $\mu$ best
  - (HV never decreases)

# Alternative MO-MIES algorithms

- Mutation only
  - Best results without recombination [Wessing et al 2017]
  - Different optimal step size for different directions




- Mutation tournament
  - Greater selection pressure

# Scalable Test Problems

- Multi-sphere

$$f_{sphere_1}(\mathbf{r},\mathbf{z},\mathbf{d}) = \sum_{i=1}^{n_r} r_i^2 + \sum_{i=1}^{n_z} z_i^2 + \sum_{i=1}^{n_d} d_i^2 \rightarrow min$$

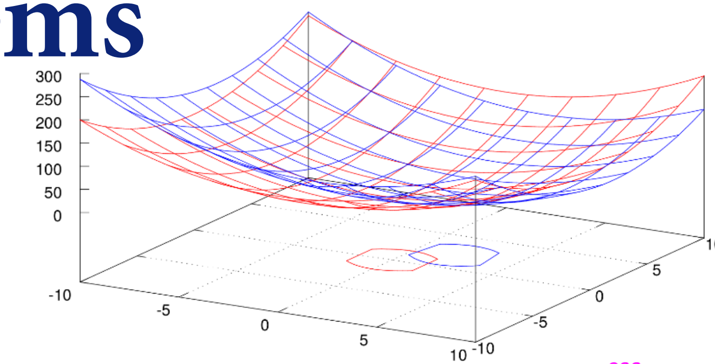
$$f_{sphere_2}(\mathbf{r},\mathbf{z},\mathbf{d}) = \sum_{i=1}^{n_r} (r_i-2)^2 + \sum_{i=1}^{n_z} (z_i-2)^2 + \sum_{i=1}^{n_d} (d_i-2)^2 \rightarrow min$$

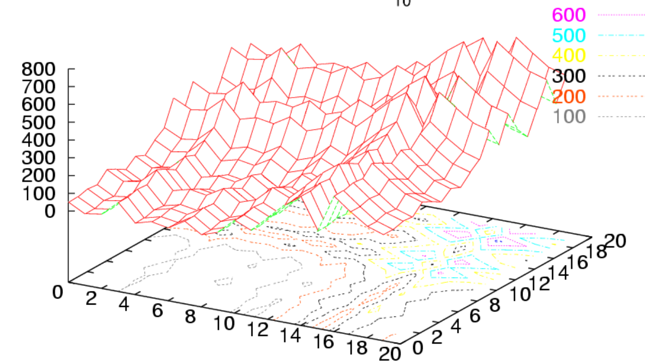
- Multi-barrier

$$f_{barrier_1}(\mathbf{r},\mathbf{z},\mathbf{d}) = \sum_{i=1}^{n_r} \left(r_i^2 + \theta \sin(r_i)^2\right) + \sum_{i=1}^{n_z} A\,[z_i]^2 + \sum_{i=1}^{n_d} B_i\,[d_i]^2 \rightarrow min$$

$$f_{barrier_2}(\mathbf{r},\mathbf{z},\mathbf{d}) = \sum_{i=1}^{n_r} \left((r_i-2)^2 + \theta \sin(r_i-2)^2\right) + \sum_{i=1}^{n_z} (A\,[z_i]-2)^2 + \sum_{i=1}^{n_d} (B_i\,[d_i]-2)^2 \rightarrow min$$

- Optical filter
  - Layers (on/off)
  - Thickness per layer



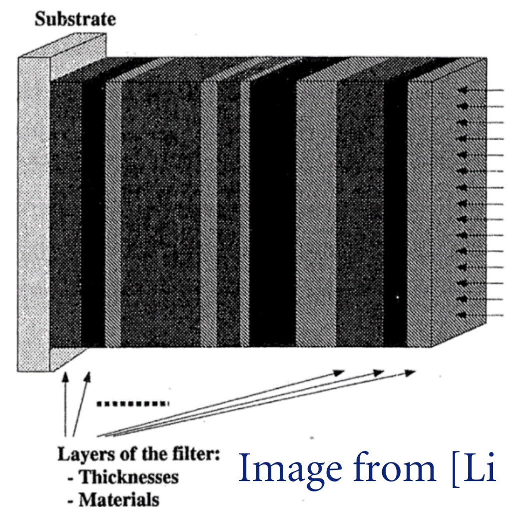



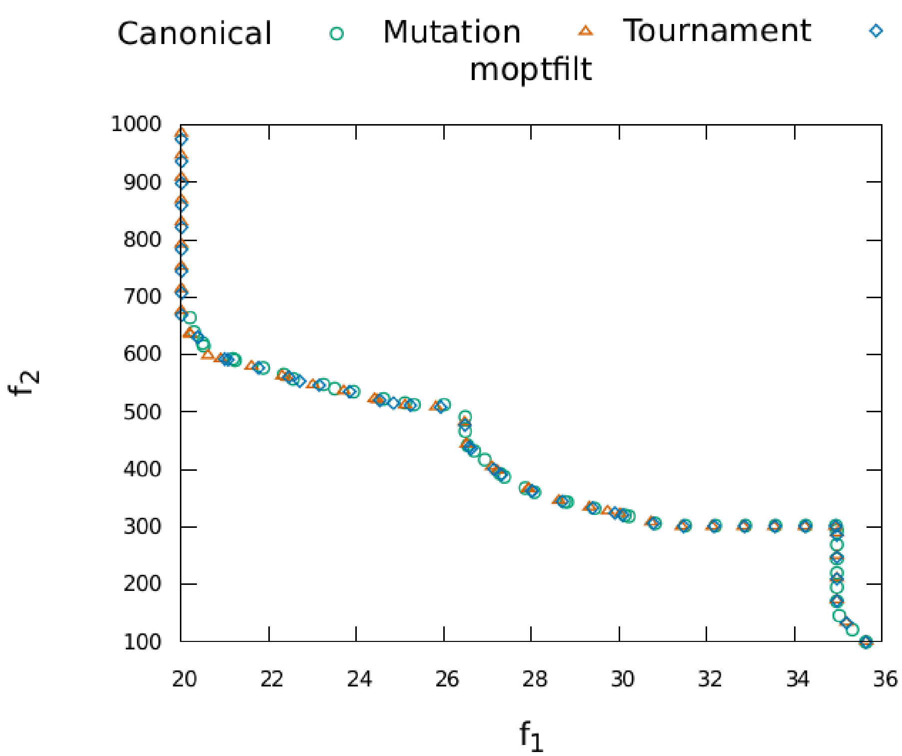

Image from [Li et al 2013]

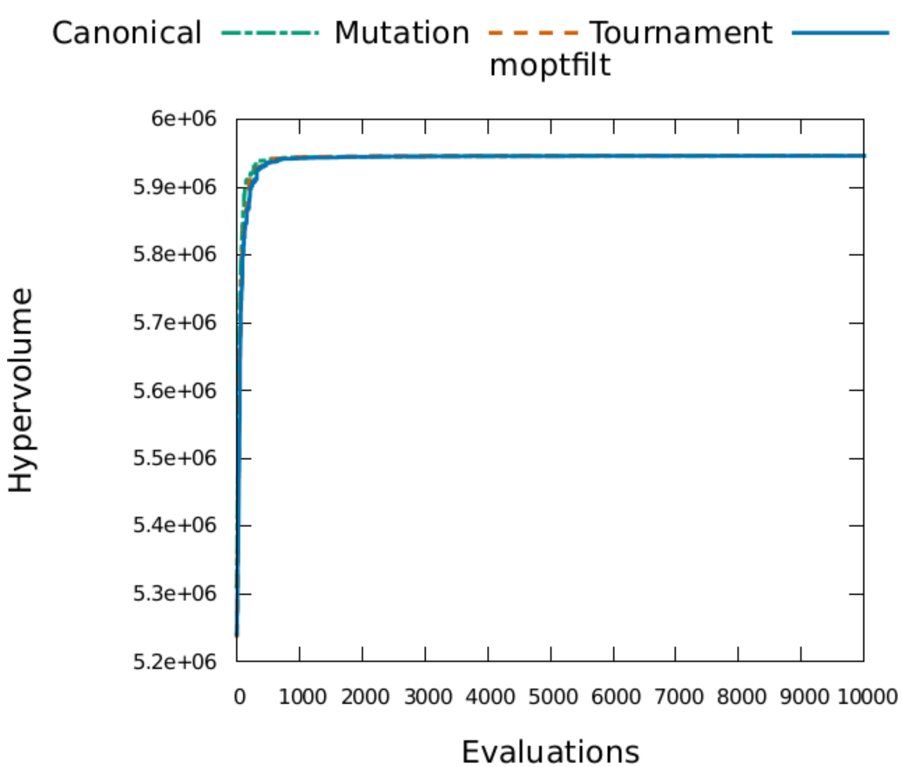
# Experimental setup

- 10,000 evaluations
- 25 repetitions

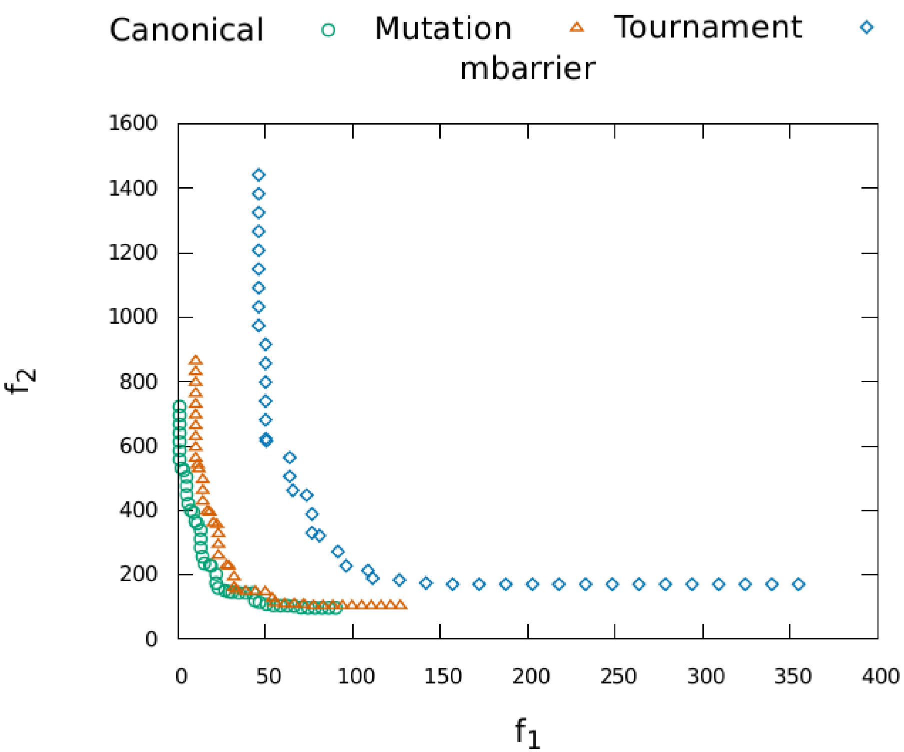
| | $n_r$ | range | $n_z$ | range | $n_d$ | range |
|---|---|---|---|---|---|---|
| $f_{sphere}$ | 5 | $[0, 20]$ | 5 | $[0, 20]$ | 5 | $[0, 20]$ |
| $f_{barrier}$ | 5 | $[0, 20]$ | 5 | $[0, 20]$ | 5 | $[0, 20]$ |
| $f_{optfilt}$ | 11 | $[0, 1]$ | N/A | N/A | 11 | $\{0, 1\}$ |

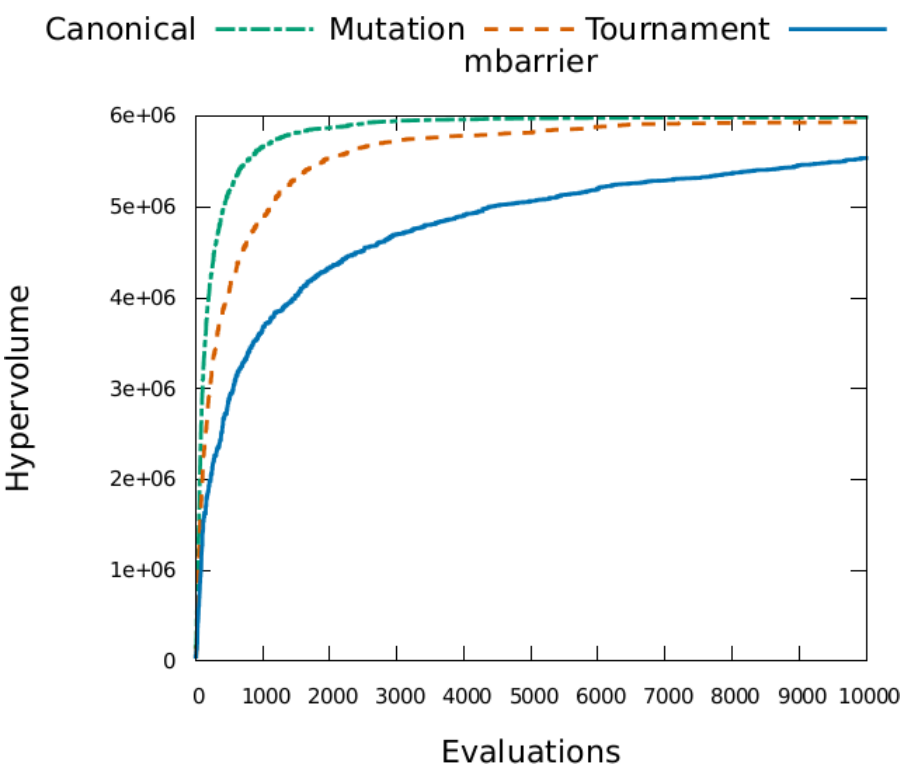
# Optical filter convergence



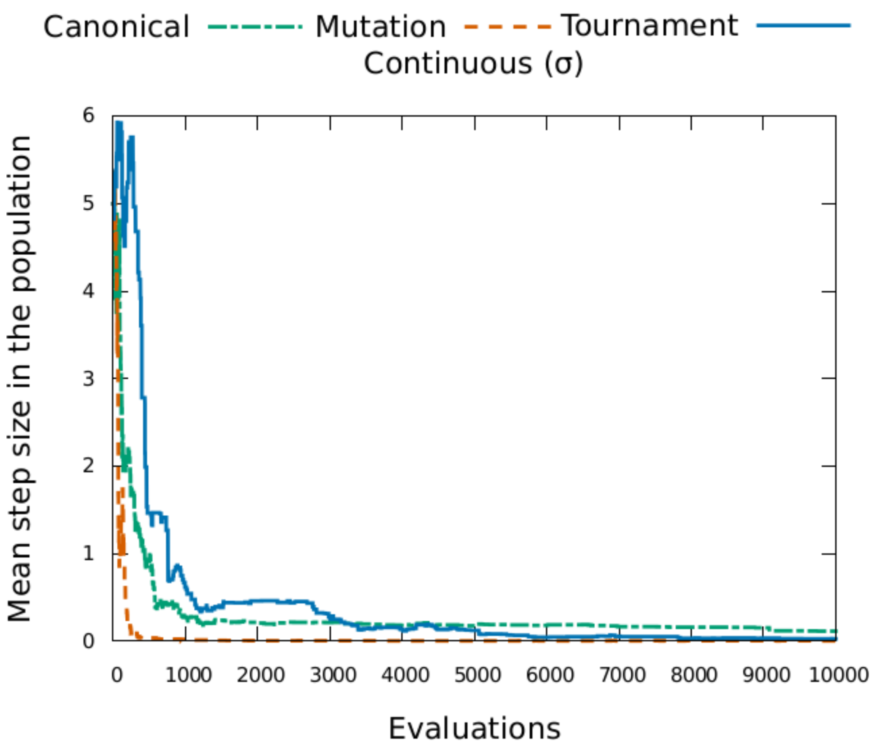



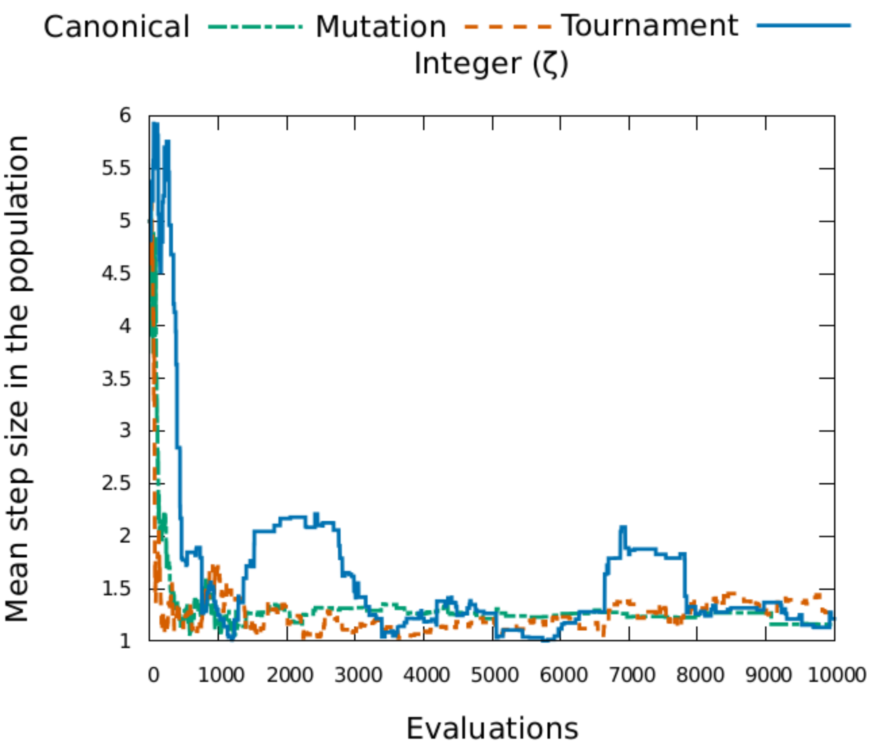
# Barrier convergence



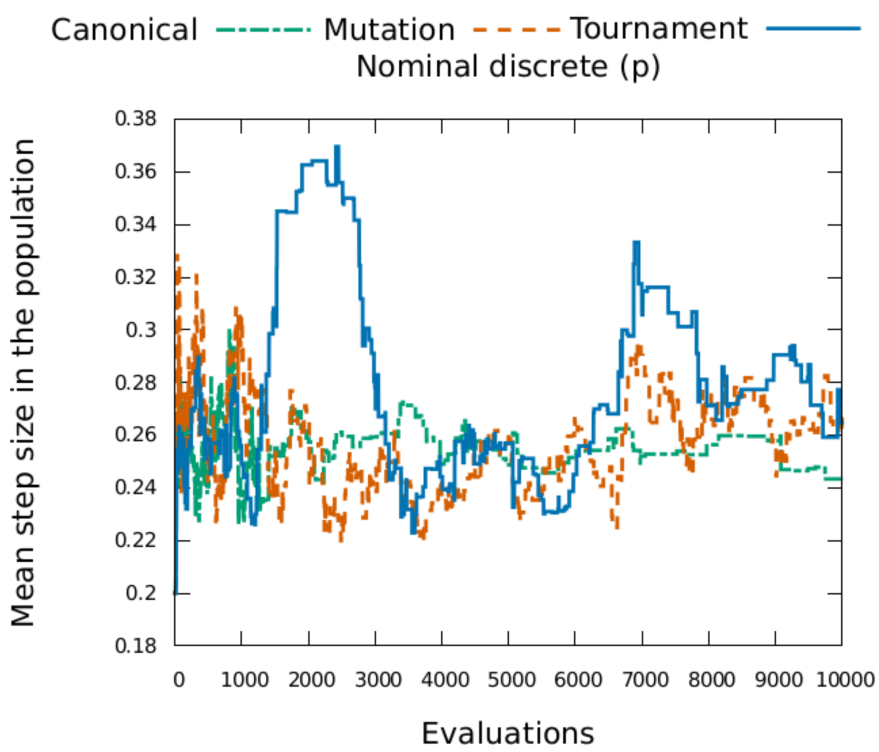



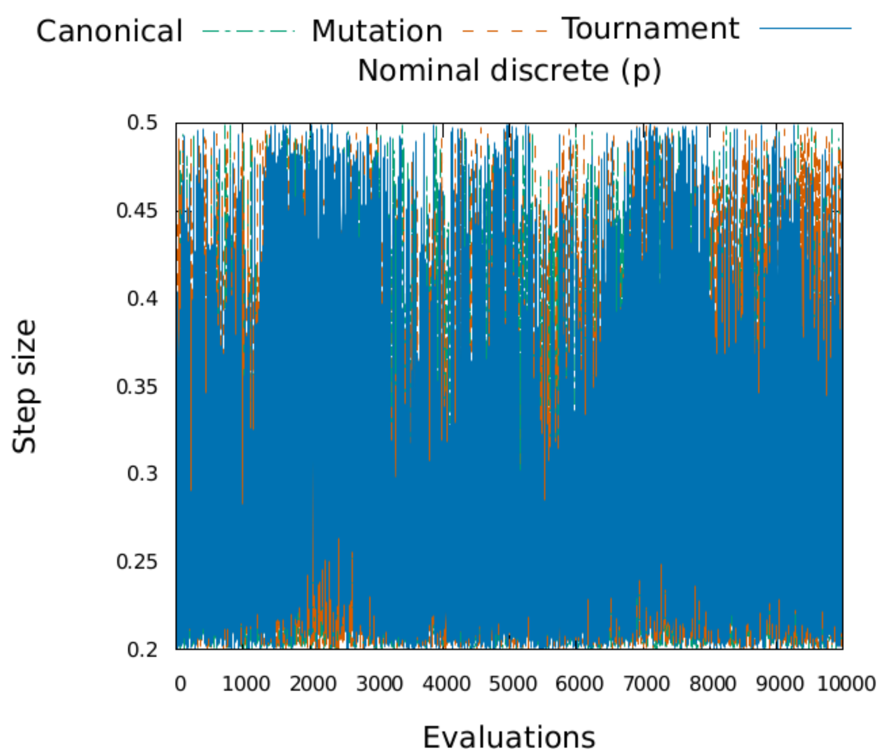
# Step size adaptation (multisphere)



# Step size adaptation – Categorical



# Future work

- Improve categorical step size adaptation
- Investigate recombination behaviour
  - Why does it work?
  - When will it not work?
- Introduce multi-objective recombination?
- Investigate integer step size adaptation
  - Can we prevent regressive behaviour?

# Summary

- **Goal:**
  - Extend the MIES algorithm for the multi-objective case
- **Plan:**
  - Evaluate MIES + SMS-EMOA (= MOMIES)
  - Evaluate mutation only variant
  - Evaluate mutation tournament variant
- **Result:**
  - Best performance for canonical MOMIES
  - Step size in continuous and integer space adapts quite well
  - Chaotic step size behaviour in categorical space
- **Future:**
  - Improve categorial step size adaptation
  - Investigate recombination behaviour

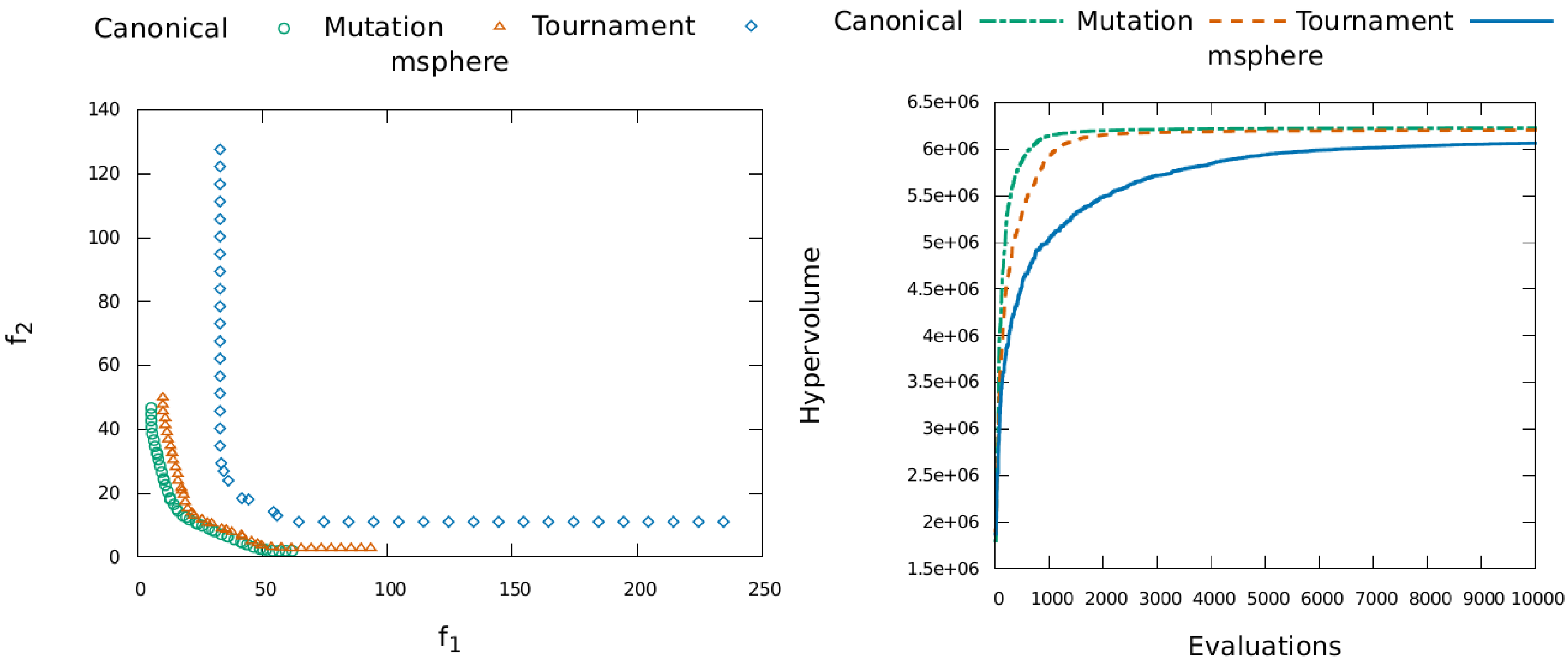
# References I


- [**Emmerich et al 2005**] M. Emmerich, N. Beume, and B. Naujoks, “An EMO algorithm using the hypervolume measure as selection criterion,” in Evolutionary Multi-Criterion Optimization, edited by C. A. Coello Coello, A. Hernández Aguirre, and E. Zitzler (Springer Berlin Heidelberg, Berlin, Heidelberg, 2005), pp. 62–76.
- [**Emmerich+Deutz 2018**] M. T. M. Emmerich and A. H. Deutz, Natural Computing 17, 585–609Sep (2018).
- [**Laredo 2015**] D. Laredo Razo, EDS: A Continuation Method for Mixed-Integer Multi-objective Optimization Problems, Master’s thesis, CINVESTAV-IPN, Mexico City (2015).
- [**Li et al 2013**] R. Li, M. T. M. Emmerich, J. Eggermont, T. Bäck, M. Schütz, J. Dijkstra, and J. H. C. Reiber, Evolutionary computation 21, 29–64 (2013).
- [**Przybylski 2010**] A. Przybylski, X. Gandibleux, and M. Ehrgott, INFORMS Journal on Computing 22, 371–386 (2010).
- [**Rudolph 1994**] G. Rudolph, “An evolutionary algorithm for integer programming,” in Parallel Problem Solving from Nature — PPSN III, edited by Y. Davidor, H.-P. Schwefel, and R. Männer (Springer Berlin Heidelberg, Berlin, Heidelberg, 1994), pp. 139–148.
- [**Schwefel 1981**] Hans-Paul Schwefel. Numerical Optimization of Computer Models. John Wiley & Sons, Inc., New York, NY, USA, 1981.
- [**Wang 2015**] H. Wang, Computers & Operations Research 61, 100–09 (2015).

# References II


- [**Wessing et al 2017**] S. Wessing, R. Pink, K. Brandenbusch, and G. Rudolph, "Toward step-size adaptation in evolutionary multi-objective optimization," in Evolutionary Multi-Criterion Optimization, edited by H. Trautmann, G. Rudolph, K. Klamroth, O. Schütze, M. Wiecek, Y. Jin, and C. Grimme (Springer International Publishing, Cham, 2017), pp. 670–684.

# Sphere convergence



# Step size adaptation