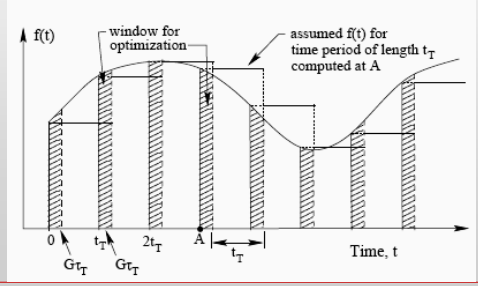
Lecture 11: Dynamic and Bilevel Optimization

- Dynamic single and multi-objective optimization
- Bilevel single and multi-objective optimization
- Omni-optimizer

Dynamic Optimization

Deb, Rao, Karthik (EMO 2007)

- Assume a statis in problem for a time step
- Find a critical frequency of change

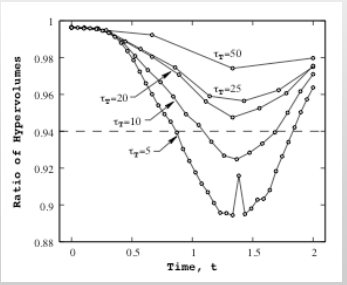


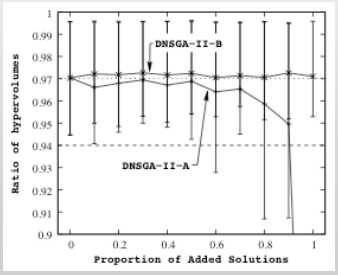
Dynamic NSGA-II

- DNSGA-IIA
  - Random solutions are introduced whenever there is a change in the problem
- DNSGA-IIB
  - Mutated solutions are introduced whenever there is a change in the problem
- Investigate if there exist a stable range of proportion η for which solutions are acceptable

Simulation Results on FDA2

- Both DNSGA-IIs are okay
- τ=20 is fine, but not τ=10 wrt hypervolume ≥0.94

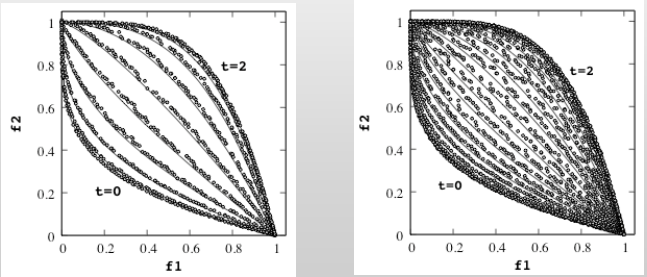




Dynamic PO fronts

- τ=20 and τ=10

too much error



A Hydro-Thermal Power Dispatch Problem

- Thermal power generation P^s_{sm} and hydroelectric power generation P^h_{hm} are variables

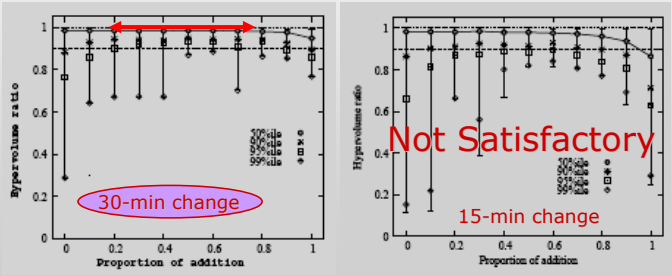
$$\begin{aligned}
\text{Minimize} \quad & f_1(\mathbf{x}) = \sum_{m=1}^{M}\sum_{s=1}^{N_s} t_m[a_s + b_s P^s_{sm} + c_s (P^s_{sm})^2 + |d_s \sin(e_s(P^s_{s,\min} - P^s_{sm}))|], \\
\text{Minimize} \quad & f_2(\mathbf{x}) = \sum_{m=1}^{M}\sum_{s=1}^{N_s} t_m[\alpha_s + \beta_s P^s_{sm} + \gamma_s (P^s_{sm})^2 + \eta_s \exp(\delta_s P^s_{sm})], \\
\text{subject to} \quad & \sum_{s=1}^{N_s} P^s_{sm} + \sum_{h=1}^{N_h} P^h_{hm} - P_{Dm} - P_{Lm} = 0, \quad m = 1,2\ldots,M, \\
& \sum_{m=1}^{M} t_m(a_{0h} + a_{1h} P^h_{hm} + a_{2h}(P^h_{hm})^2) - W_h = 0, \quad h = 1,2,\ldots,N_h. \\
& P^h_{h,\min} \le P_{hm} \le P^h_{h,\max}, \quad h = 1,2,\ldots,N_h, m = 1,2,\ldots,M, \\
& P^s_{s,\min} \le P_{sm} \le P^s_{s,\max}, \quad s = 1,2,\ldots,N_s, m = 1,2,\ldots,M.
\end{aligned}$$

- Minimize Cost (non-differentiable) and NOx emission
- Power balance and water head limits
- Known power demand P_{Dm} with time m

Dynamic Multi-Objective Hydro-Thermal Power Scheduling

- Addition of random or mutated points at changes
- 30-min change found satisfactory

EMO-2007



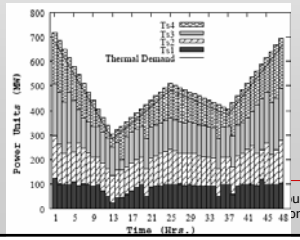
Automatic Decision-Making

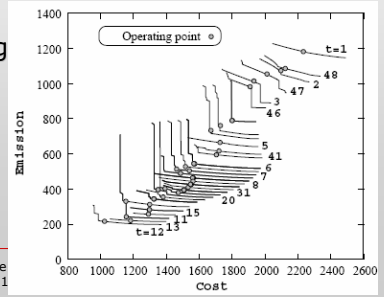
- Choose a solution with pseudo-weight

$$w_1(\mathbf{x}) = \frac{(f_1^{\max} - f_1(\mathbf{x}))/(f_1^{\max} - f_1^{\min})}{(f_1^{\max} - f_1(\mathbf{x}))/(f_1^{\max} - f_1^{\min}) + (f_2^{\max} - f_2(\mathbf{x}))/(f_2^{\max} - f_2^{\min})}$$

closer to w_1=0.5

- Online decision making

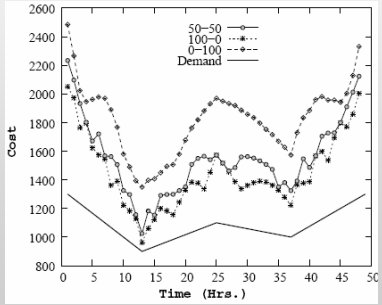




Dynamic Variation of Cost

- Effect of weight on cost

Case	Cost	Emission
50-50%	74239.07	25314.44
100-0%	69354.73	27689.08
0-100%	87196.50	23916.09



What is Bilevel Optimization?

- Two levels of optimization tasks
  - Upper level: (xu,xl)
  - Lower level: (xl)
- An upper level feasible solution must be an optimal lower level solution: (xu, xl*(xu))

$$\begin{aligned} \min._{(\mathbf{x}_u,\mathbf{x}_l)} \quad & \mathbf{F}(\mathbf{x}) = (F_1(\mathbf{x}),\ldots,F_M(\mathbf{x})), \\ \text{s.t.} \quad & \mathbf{x}_l \in \text{argmin}_{(\mathbf{x}_l)} \{\mathbf{f}(\mathbf{x}) = (f_1(\mathbf{x}),\ldots,f_m(\mathbf{x})) \mid \\ & \quad \mathbf{g}(\mathbf{x}) \geq \mathbf{0}, \mathbf{h}(\mathbf{x}) = \mathbf{0}\}, \\ & \mathbf{G}(\mathbf{x}) \geq \mathbf{0}, \mathbf{H}(\mathbf{x}) = \mathbf{0}, \\ & x_i^{(L)} \leq x_i \leq x_i^{(U)}, \quad i = 1,\ldots,n. \end{aligned}$$

Single Versus Multi-Objective BP Problems

- Single-objective BP
  - One F1 and one f1
  - Usually one target solution xu* and xl*
  - A coevolutionary approach can be devised and convergence can be expected
- Multi-objective BP
  - Usually multiple solutions xl* for each xu*
  - Find and maintain many solutions for each xu
  - Not an easy matter

Test Problem 1

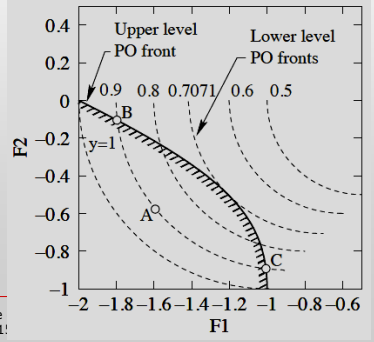
(Eichfelder, 2007)

$$\min \mathbf{F}(\mathbf{x}) = \begin{Bmatrix} x_1 - y \\ x_2 \end{Bmatrix}, \qquad \textbf{y upper, (x1,x2) lower}$$

$$\text{st } (x_1,x_2) \in \text{argmin}_{(x_1,x_2)} \left\{ \begin{matrix} \mathbf{f}(\mathbf{x}) = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\ g_1(\mathbf{x}) = y^2 - x_1^2 - x_2^2 \geq 0 \end{matrix} \right\},$$

$$G_1(\mathbf{x}) = 1 + x_1 + x_2 \geq 0,$$
$$-1 \leq x_1, x_2 \leq 1, \quad 0 \leq y \leq 1.$$

- Dashed lines are lower level PO fronts
  - Depends on y
- Upper level PO front on constraint G1
  - Maximum two solutions from each y
  - Not all y in upper PO front
- Solutions below PO front exist



Bilevel Problems in Practice

- Often appears from functional feasibility
  - Stability, equilibrium, solution to a set of PDEs
  - Ideally, lower level task must implement above
- Lower level is bypassed by approximation or by using direct simplified solution principles
  - Due to lack of suitable solution techniques
- *Stackelberg* games: Leader-follower
  - Leader must be restricted to follower's decisions
  - Follower must respect leader's decisions



Past Studies in Bilevel Programming

- Plenty of studies in single-objective BP, not much in multi-objective BP
- Optimality theory difficult
  - KKT conditions involve second derivative of lower level objectives and constraints
  - Dempe et al. (2006) developed KKT conditions
  - Impractical to implement (abstract terms)
- Fliege and Vincent (2006): BP->4-obj MOP
  - Involves derivates and unclear of extensions to multi-objective and higher level problems



Past BP Algorithms

- A common approach: Nested optimization
  - For every xu, lower level problem is solved completely
  - Computationally expensive, extension to multi-objective BP impractical
- KKT conditions of lower level problem used as constraints (Herskovits et al., 2000)
  - Lagrange multiplers increase # of decision variables
  - Constrained search space
  - Applicable to differentiable problems only
- Intelligent exhaustive upper level search
  - Subdivision approach on upper level and numerical optimization in lower level (Eichfelder, 2007)



Evolutionary Algorithms for Solving BP

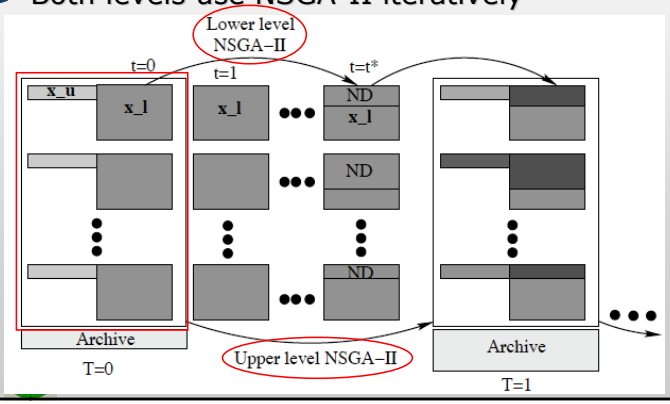
- Not a single study on multi-objective BP
- Upper GA, lower classical: Nested approach
  - Upper GA, lower LP (Mathieu et al., 1994)
  - Upper GA, lower Frank-Wolfe (Yin, 2000)
- Only Upper level with KKT of lower level
  - Hecheng and Wang, 2007; Li and Wang, 2007, others
- Coevolutionary approach (Oduguwa and Roy, 2002)
  - Populations on xu and xl co-evolve, cross-linking pops
  - Excellent idea for 1-obj BP, may not be suitable for multi-obj BP
- PSO in both levels (Li et al., 2006, Halter and Mostaghim, 2006)
  - Linear problems handled, their extensions are unclear



Proposed BLEMO

- Population structure
- Lower and upper level NSGA-II
- Archiving

- Both levels use NSGA-II iteratively



BLEMO Details

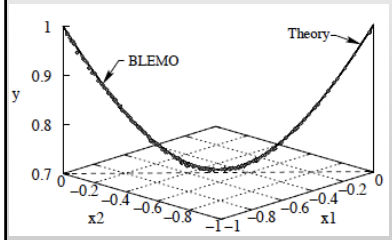
- Initial population creation
  - Follow population structure: for each xu, Nl different xl -> Nu randomly created solutions
- Upper level NSGA-II:
  - Tournament with (ND_u, CD_u) from (F,G,H)
  - Recombination, mutation to create xu^1
  - Nl different xl created from P_t -> (xu^1, xl)
- Lower level NSGA-II on subpopulations: (xu^1 is fixed)
  - Tournament with (ND_l, CD_l) from (f,g,h), second parent from archive
  - Recombination, mutation as usual
  - Run till Tl generations
- All new solutions after lower level NSGA-II in Q_t
- Combine P_t with Q_t, prefer ND_u=ND_l=1 to choose P_(t+1)
- Update archive
- Run till Tu generations

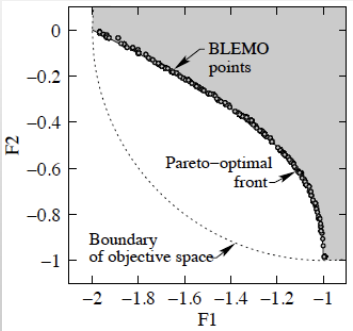
BLEMO Features:
Different from a Nested Procedure

- A population of xu is initialized
  - Lower level NSGA-II attempts to move towards their PO fronts
- Based on intermediate information, xu population is updated to move towards feasible and PO region of upper level problem
  - Again, lower level NSGA-II attempts to move their own PO fronts
  - And so on …
- Towards the end xu does not change much, but xl gets towards Pareto-optimality with iterations

BLEMO Results on Problem 1

- Nu=400, Tu=200
- Nl=40, Tl=40
- Near PO solutions found

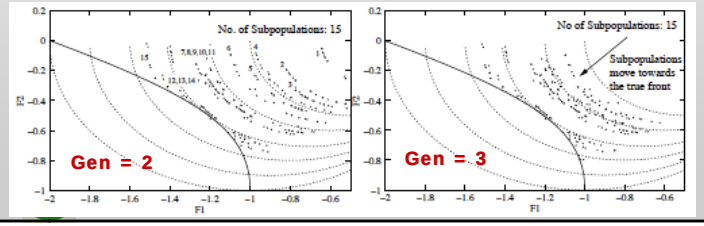




Lower Versus Upper Level Optimizations

- Identical func. eval.
- Large Tl means nested with less importance on upper level
- Small Tl means less care on lower level
- Tl=40 seems adequate

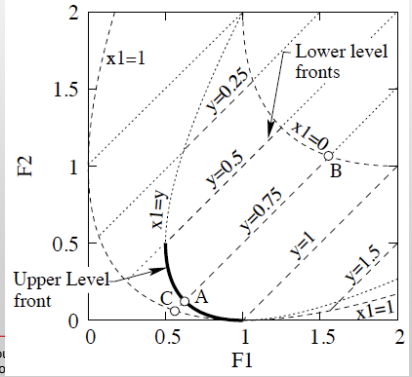
T_l	T_u	Hypervolume
20	391	0.29851
40	200	**0.30268**
100	81	0.29716
200	41	0.28358
400	20	0.23796



Test Problem 3

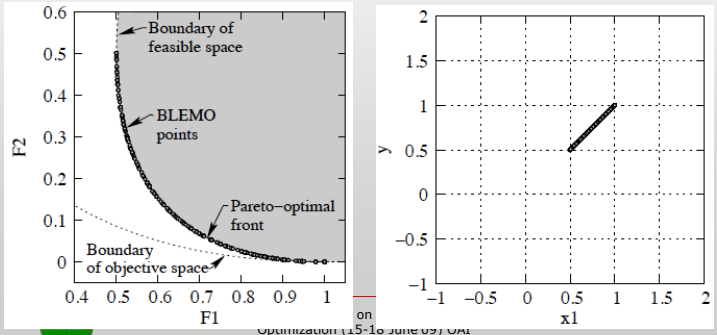
$$\text{minimize } \mathrm{F}(\mathrm{x}) = \begin{Bmatrix} (x_1-1)^2+x_2^2+y^2 \\ (x_1-1)^2+x_2^2+(y-1)^2 \end{Bmatrix},$$
$$\text{subject to } (x_1,x_2) \in \text{argmin}_{(x_1,x_2)} \left\{ \mathrm{f}(\mathrm{x}) = \begin{pmatrix} x_1^2+x_2^2 \\ (x_1-y)^2+x_2^2 \end{pmatrix} \right\},$$
$$-1 \le x_1, x_2, y \le 2.$$

- Developed in this study
- Only one point from a lower level PO front is on upper level PO front
- Scalable to higher dimensional problems
- Problem 4 (15 var.)



BLEMO Results on Problem 3

- True PO solutions are found
- Similar results on Problem 4

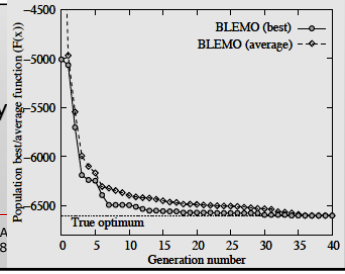


Test Problem 5 (Single-Objective)

$$\text{minimize } F(\mathrm{x}) = -(200-x_1-x_2)(x_1+x_3)-(160-x_2-x_4)(x_2+x_4),$$
$$\text{subject to}$$
$$\mathrm{x}_l \in \text{argmin}_{(\mathrm{x}_l)} \{\mathrm{f}(\mathrm{x}) = (x_1-4)^2+(x_2-13)^2+(x_3-35)^2+(x_4-2)^2\}$$
$$g_1(\mathrm{x}) = 0.4x_1+0.7x_2 \le y_1,\ g_2(\mathrm{x}) = 0.6x_1+0.3x_2 \le y_2,$$
$$g_3(\mathrm{x}) = 0.4x_3+0.7x_4 \le y_3,\ g_4(\mathrm{x}) = 0.6x_3+0.3x_4 \le y_4\},$$
$$G(\mathrm{x}) = y_1+y_2+y_3+y_4 \le 40,$$
$$0 \le y_1 \le 10,\ 0 \le y_2 \le 5,\ 0 \le y_3 \le 15, 0 \le y_4 \le 20,$$
$$0 \le x_1 \le 20,\ 0 \le x_2 \le 20,\ 0 \le x_3 \le 40,\ 0 \le x_4 \le 40.$$

(Colson, 2002)

- Known optimum:
  $\mathrm{x}_u^* = (7.36, 3.55, 11.64, 17.45)^T$
  $\mathrm{x}_l^* = (0.91, 10, 29.09, 0)^T$
  $F(\mathrm{x}^*) = -6600.0$ and $f(\mathrm{x}^*) = 57.48$.
- Identical solution obtained by BLEMO
- BLEMO scales to 1-obj problems as well

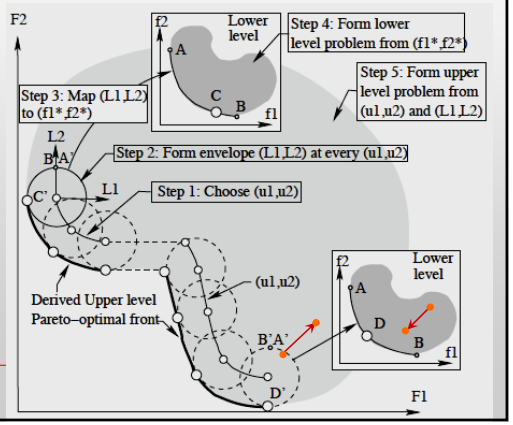


Extensions to This Study

- Difficult test problems suggested (Deb and Sinha, CEC-2009)
  - Allows an adequate test to BP algorithms
- Self-adaptive and local search based BLEMO suggested (Deb and Sinha, KanGAL Report 2009001)
  - Much faster approach
  - No additional parameter

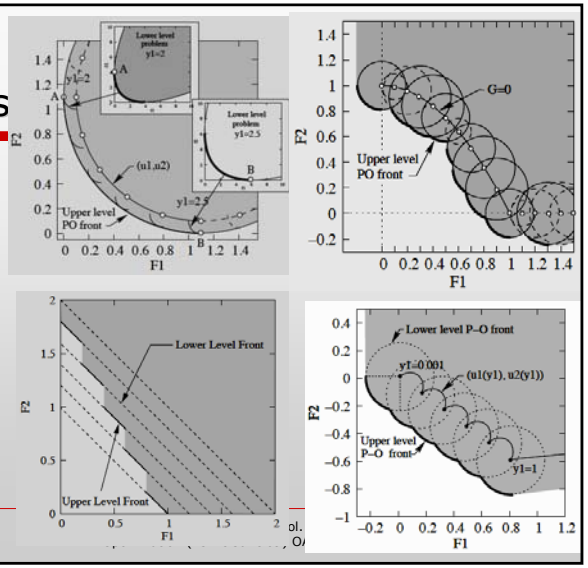
Test Problem Construction Principle

- Difficulties identified
- Bottom-up approach
- Five-step procedure
- *Conflict between lower and upper levels*



Some Problems

- Scalable in terms of variables and objectives
- Controllable difficulties

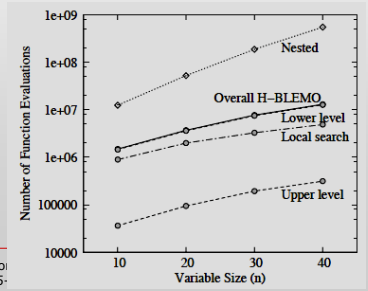


Hybrid, Self-Adaptive BLEMO

- Subpopulation size not fixed
  - Based on current population with respect to archive
- All parameters (Nl, Tl, Tu) are self-adapted
- Lower level best solutions are improved by a local search
  - Achievement scalarizing function method
- Parametric study on Nu suggests Nu=20n

Hallmark of Hybrid BLEMO

- Handle larger variables (<= 40), scalable performance
- One to two orders of magnitude less function evaluations
- Interactive lower and upper level NSGA-II quicker
- Proposed method better than nested

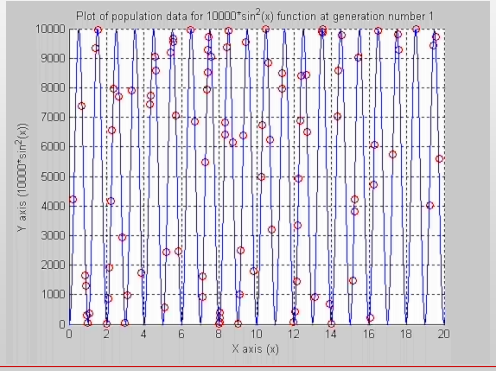


Omni-Optimizer:
Motivation from Computation

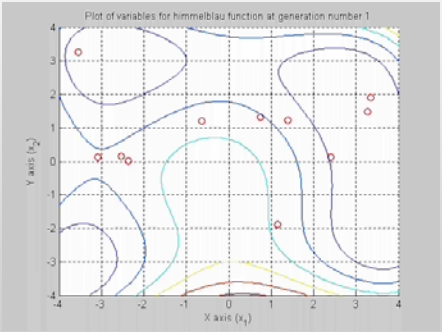
- One algorithm for multiple and single-objective problem solving (Deb and Tiwari, 2005)
  - Single objective is a degenerate case

```
Add a few integers:
begin
   print 'enter number of integers to be added', read n
   sum = 0
   for i = 1 to n
      print 'enter integer i', read a[i]
      sum = sum + a[i]
   print 'Sum =' sum
end
```

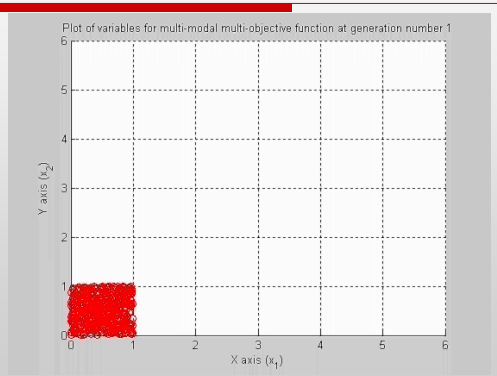

$10^4 Sin^2(x)$



Himmelblau's Function



Multi-Modal Multi-Objective Problem



Conclusions

- Multiobjective bilevel programming received lukewarm interest by EA researchers
- Population approach of EA makes tremendous potential
- Nested nature of the problem makes the task computationally expensive
- Presented a viable interacting population approach using NSGA-II
- Ongoing extensions to this original study show promise for EMO
- More efforts are called for