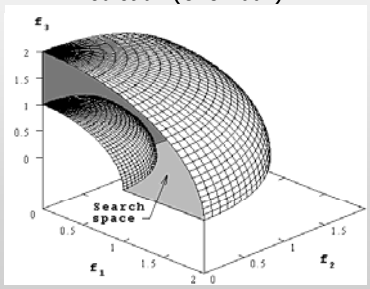
Lecture 8:
Other Aspects of EMO

- Test problem design
- Hybrid EMO-Local-Search Methods
- EMO with theoretical convergence property
- Performance metrics in EMO
- Visualization of PO front
- A case study on hydro-thermal power dispatch problem

Scalable Test Problems

Deb et al. (CEC-2002)

- **Step 1** Define Pareto-optimal front mathematically
- **Step 2** Build the objective search space using it
- **Step 3** Map variable space to objective space
- Scalable DTLZ problems suggested

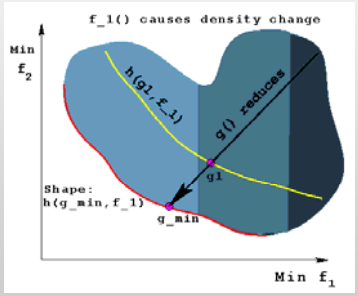


Two-Objective Test Problems

- Pareto-optimal front is controllable and known
- ZDT problems:

$$\text{Min.} \quad f_1(\mathbf{x}) = f_1(\mathbf{x}_I),$$
$$\text{Min.} \quad f_2(\mathbf{x}) = g(\mathbf{x}_{II})h(f_1, g).$$

- Choose $f_1()$, $g()$ and $h()$ to introduce various difficulties



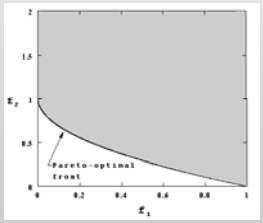
Zitzler-Deb-Thiele's Test Problems

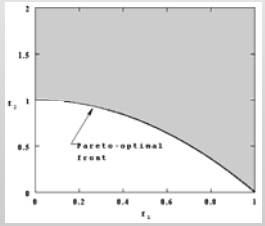
ZDT1

$$f_1(\mathbf{x}) = x_1,$$
$$g(\mathbf{x}) = 1+\frac{9}{n-1}\sum_{i=2}^{n} x_i,$$
$$h(f_1, g) = 1-\sqrt{f_1/g}.$$

ZDT2

$$f_1(\mathbf{x}) = x_1,$$
$$g(\mathbf{x}) = 1+\frac{9}{n-1}\sum_{i=2}^{n} x_i,$$
$$h(f_1, g) = 1-(f_1/g)^2.$$

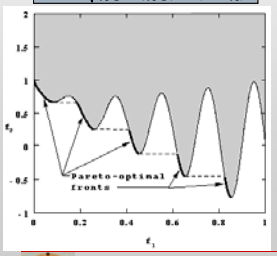




Zitzler-Deb-Thiele's Test Problems

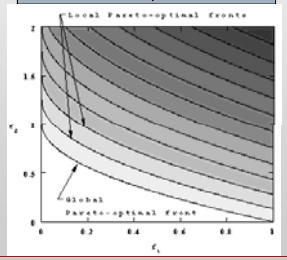
ZDT3

$$f_1 = x_1,$$
$$g = 1 + \frac{9}{n-1}\sum_{i=2}^{n} x_i,$$
$$h = 1 - \sqrt{f_1/g} - (f_1/g)\sin(10\pi f_1).$$



ZDT4

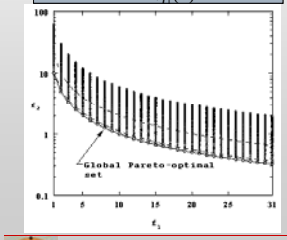
$$f_1 = x_1,$$
$$g = 10n - 9 + \sum_{i=2}^{n}(x_i^2 - 10\cos(4\pi x_i)),$$
$$h = 1 - \sqrt{f_1/g}.$$



Zitzler-Deb-Thiele's Test Problems

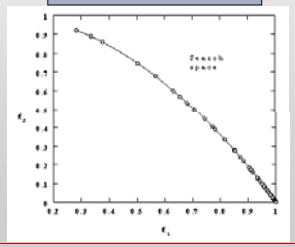
ZDT5

$$f_1 = 1 + u(x_1),\ g = \sum_{i=2}^{11} v(u(x_i))$$
$$v = \begin{cases} 2 + u(x_i) & \text{if } u(x_i) < 5, \\ 1 & \text{if } u(x_i) = 5, \end{cases}$$
$$h = 1/f_1(\mathbf{x})$$



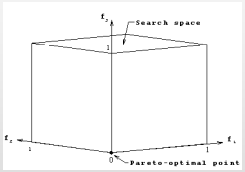
ZDT6

$$f_1 = 1 - \exp(-4x_1)\sin^6(6\pi x_1),$$
$$g = 1 + 9\left[(\sum_{i=2}^{10} x_i)/9\right]^{0.25},$$
$$h = 1 - (f_1/g)^2.$$

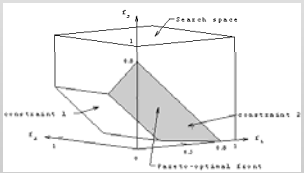


Constraint Surface Approach

- Step 1: Define a rectangular hyper-box
- Step 2: Chop off regions using constraints



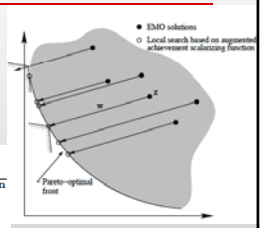
- Adv: Easy to construct
- Disadv: Difficult to define Pareto-optimal front



Hybrid EMO (Karthik, Deb & Miettinen, 2008)

- Improve EMO's convergence properties
- Introduce local search based on achievement scalarization function occasionally

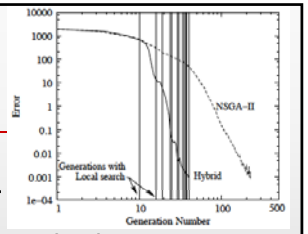
$$\text{Minimize}_{\mathbf{x}} \max_{i=1}^{M} \frac{f_i(\mathbf{X}) - z_i}{f_i^{\max} - f_i^{\min}} + \rho \sum_{j=1}^{M} \frac{f_j(\mathbf{X}) - z_j}{f_j^{\max} - f_j^{\min}}$$
$$\text{subject to } \mathbf{x} \in \mathcal{S},$$



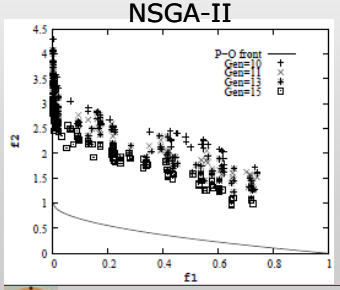
- ASF does not allow converging to weak PO points
- Periodic application of local search
  - Allow good solutions to mix well with population

Hybrid EMO (cont.)

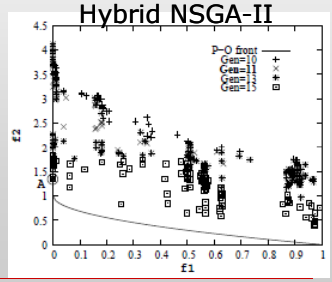
- Difference is clear
- Local search makes EMO faster



NSGA-II



Hybrid NSGA-II



Hybrid EMO

Function Evaluations

Test Problem	Original NSGA-II Best	Original NSGA-II Median	Original NSGA-II Worst	Hybrid NSGA-II Best	Hybrid NSGA-II Median	Hybrid NSGA-II Worst
ZDT1	19,400	20,900	24,600	**4,665**	7,554	8,580
ZDT2	20,600	21,700	23,200	**4,826**	6,351	7,198
ZDT3	21,900	23,500	26,300	**10,736**	16,731	22,137
ZDT4	27,100	34,500	60,300	**6,003**	7,658	14,479
3-DTLZ1	59,800	76,200	97,200	**48,352**	60,323	76,890
3-DTLZ2	38,200	54,000	93,800	**32,611**	61,508	69,650
4-DTLZ2	52,200	68,200	120,800	**45,005**	61,879	114,899

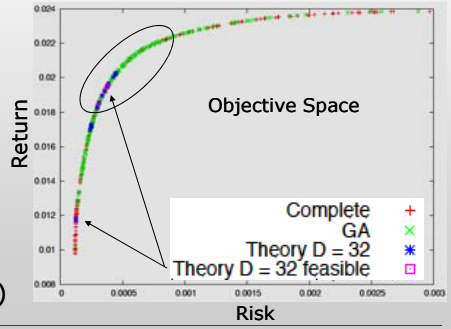
Diversity Measure (smaller, the better)

Test Problem	Original NSGA-II Best	Original NSGA-II Median	Original NSGA-II Worst	Hybrid NSGA-II Best	Hybrid NSGA-II Median	Hybrid NSGA-II Worst
ZDT1	0.0043	0.0047	0.0054	**0.0034**	0.0042	0.1630
ZDT2	0.0044	0.0053	0.0064	**0.0037**	0.0070	0.0499
ZDT3	0.0012	0.0016	0.0023	**0.0007**	0.0009	0.0010
ZDT4	0.0042	0.0047	0.0055	0.0042	0.1135	0.2851
3-DTLZ1	0.0196	0.0341	0.0403	**0.0187**	0.0224	0.0296
3-DTLZ2	0.0000	0.0000	0.0000	0.0000	0.0001	0.0030
4-DTLZ2	0.0049	0.0066	0.0086	**0.0048**	0.0067	0.0085

Case Study: Portfolio Optimization

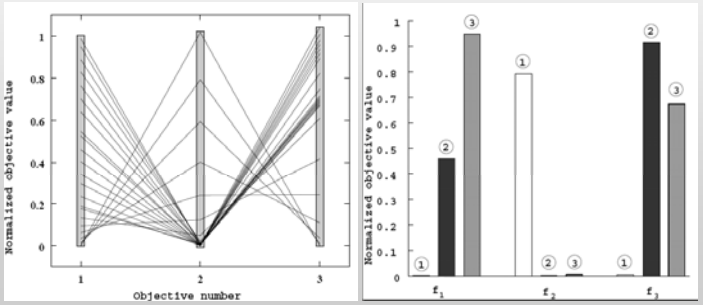
$$\min \{x^T \Sigma x\}$$
$$\max \{\mu^T x\}$$
$$s.t. \quad 1^T x = 1$$
$$x_i = 0 \text{ or } x_i \in [\alpha, \omega] \quad \alpha \geq 0$$

- Minimize risk, maximize return and Card(x)=D
- Complete: Assuming any x
- GA: x=0 or [α,w] and fixed D=32
- Theory D=32: Complete with D=32
- Theory D=32 feasible: D=32 & feasible (9 solns.)

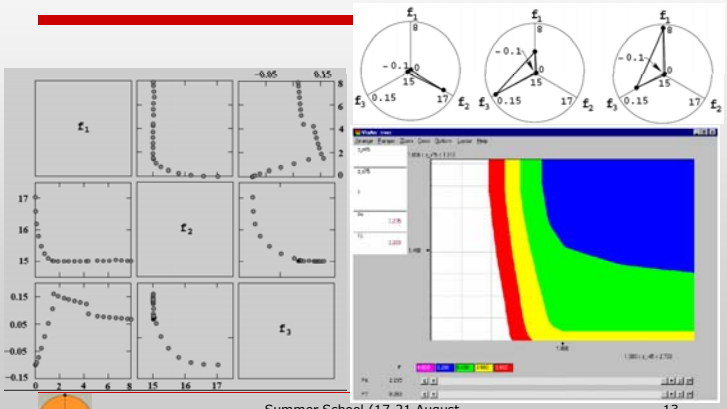


Visualization of Solutions

Value Path and Bar Diagrams

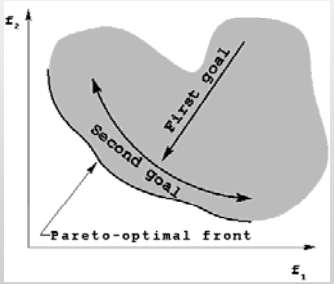


Scatter Plots, Star Coordinate Diagrams & Decision Maps



Performance Metrics

- A recent study by Zitzler et al. (2003) suggests at least *M* metrics
- Two essential metrics (functionally)
  - Convergence measure
  - Diversity measure



Metrics for Convergence

- Error ratio:

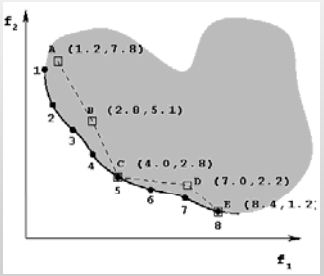
$$ER = \frac{\sum_{i=1}^{|Q|} e_i}{|Q|}$$

- Set Coverage:

$$C(A,B) = \frac{|\{b \in B \mid \exists a \in A : a \circ b\}|}{|B|}$$

- Generational distance:

$$GD = \frac{(\sum_{i=1}^{|Q|} d_i^p)^{1/p}}{|Q|}$$



Metrics for Diversity

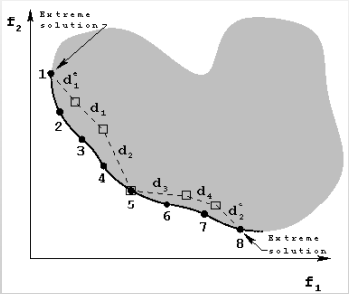
- Spacing:

$$S = \sqrt{\frac{1}{|Q|}\sum_{i=1}^{|Q|}(d_i - \bar{d})^2}$$

- Spread:

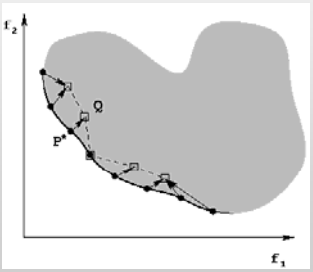
$$\Delta = \frac{\sum_{m=1}^{M} d_m^e + \sum_{i=1}^{|Q|} |d_i - \bar{d}|}{\sum_{m=1}^{M} d_m^e + |Q|\,\bar{d}}$$

- Chi-square like deviation measure

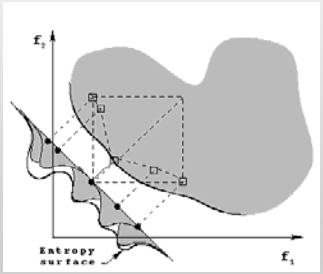


Metrics for Diversity (Cont.)

Distance from P*

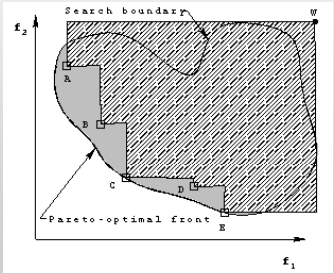


Entropy Measure

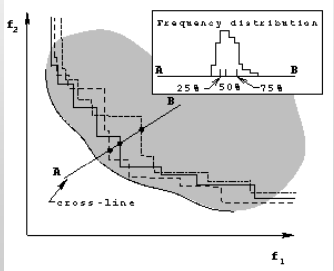


Metrics for Convergence and Diversity

Hypervolume



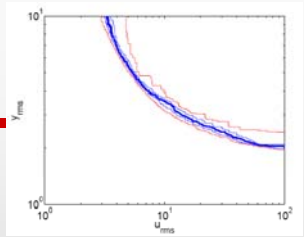
Attainment Surface Method

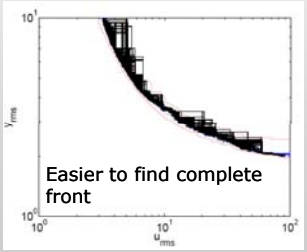


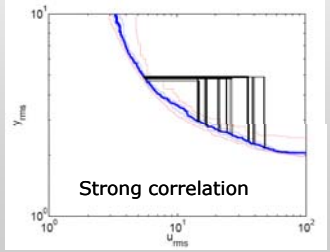
Attainment Surface Methods

- Positive and negative relationships

Fonseca (2005; 2007)



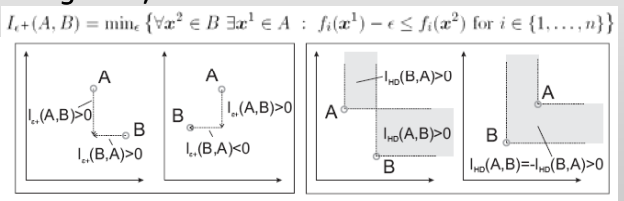




Indicator-Based EMO

- Zitzler et al. (2004)
- Move point A so as to make weakly-dominated with B
- Positive if reduction in hypervolume
- Negative, otherwise

$$I_{\epsilon+}(A,B) = \min_{\epsilon} \{\forall \boldsymbol{x}^2 \in B \; \exists \boldsymbol{x}^1 \in A \; : \; f_i(\boldsymbol{x}^1) - \epsilon \leq f_i(\boldsymbol{x}^2) \text{ for } i \in \{1,\ldots,n\}\}$$



Evolutionary Optimization and Theory

- Discussions restricted to real-parameter problems
- Two pertinent questions:
  - Are EA solutions theoretically optimal?
  - Does EA gets its solutions closer to theoretically optimum?
- EAs since 1980s, but above is not addressed adequately
- Answers may attract our theoretical counterparts to EAs



KKT Necessary Conditions

A solution ***x*** is optimum,
if following conditions are true:

- At ***x***, grad of *g* (for active constraints) and grad of *h* are linearly independent
- ***x*** satisfies following two conditions:

$$1)\ \sum_{i=1}^{M}\lambda_i\nabla f_i(\mathbf{x})+\sum_{j=1}^{J}u_j\nabla g_j(\mathbf{x})+\sum_{k=1}^{K}v_k\nabla h_k(\mathbf{x})=0,$$
$$2)\ u_j g_j(\mathbf{x})=0 \text{ for } j=1,\ldots,J.$$

- ***λ*** non-negative, non-zero vector
- ***u*** non-negative vector



Practicalities of Using KKT conditions

- Gradients may not exist
  - Numerical gradient computation
  - Subdifferential approach can be used
- Computing ***x**** from KKT conditions is difficult
  - Non-linear root-finding problem
  - Unknowns: (**x, λ, u**), usually avoided
- But KKT conditions can be checked for a ***x***
  - Linear system of equations to be solved
  - Unknowns: (**λ, u**)



Checking KKT Conditions

- Given an EA solution ***x*** in R^n
- # of active *g*(***x***) and *h*(***x***) ≤ *n*
  - Constraint qualification is usually met
- # of active *g*(***x***) and *h*(***x***) > *n*
  - Check grad *g* and *h* for linear independance
- Consider only active constraints *g*
  - *Takes care of condition 2*



Remaining Tasks

- Find ($\boldsymbol{\lambda}$, $\mathbf{u}$)$\geq$**0** and $\boldsymbol{\lambda}\neq$**0** for which

$$\sum_{i=1}^{M}\lambda_i\nabla f_i(\mathbf{z}^{(i)})+\sum_{j=1}^{J}u_j\nabla g_j(\mathbf{z}^{(i)})+\sum_{k=1}^{K}v_k\nabla h_k(\mathbf{z}^{(i)})=0.$$

- Assume, $\Sigma\lambda_i$=1 and substitute $\lambda_M=1-\sum_{i=1}^{M-1}\lambda_i$
- Above equations: *Ay=b*

$$\begin{aligned} A &= ((\nabla f_1-\nabla f_M),\ldots,(\nabla f_{M-1}-\nabla f_M),\\ &\quad \nabla g_1,\ldots,\nabla g_J,\nabla h_1,\ldots,\nabla h_k), \quad \text{(n x p), p=M+J+K-1}\\ b &= -\nabla f_M, \quad \text{(n x 1)}\\ y &= (\lambda_1,\ldots,\lambda_{M-1},u_1,\ldots,u_J,v_1,\ldots,v_K)^T. \quad \text{(p x 1)} \end{aligned}$$

Three Scenarios

- n>p: More equations than unknowns
  - Least square solution: $\bar{y}=(A^TA)^{-1}(A^Tb)$
  - Define an error measure: $\tilde{e}=\|b-A\bar{y}\|/\|b\|$
  - If error is close to zero, ***x*** is a KKT point
- n<p: Less equations than unknowns
  - Pose as an optimization problem:

$$\begin{aligned} &\text{Minimize} \quad e(\mathbf{y})=\|b-A\mathbf{y}\|/\|b\|,\\ &\text{Subject to} \quad \mathbf{y}\geq 0, \end{aligned}$$

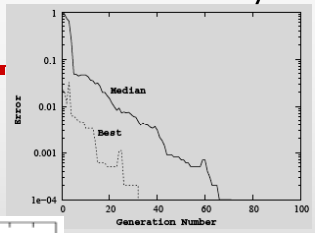
  - If *e** is close to zero, ***x*** is a KKT point
- n=p: # of equations and unknowns are same
  - Compute: $\bar{y}=A^{-1}b$ $e(\bar{y})=\|b-A\bar{y}\|/\|b\|$
  - If *e* is close to zero, ***x*** is a KKT point

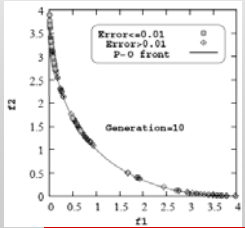
Simulation Results

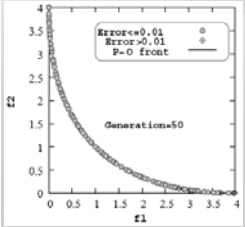
$$\begin{aligned} &\text{Minimize} \quad f_1(x,y)=x^2+y^2,\\ &\text{Minimize} \quad f_2(x,y)=(x-2)^2+y^2. \end{aligned}$$

- More and more solutions become KKT points with generation

Error History



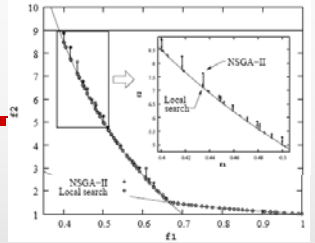




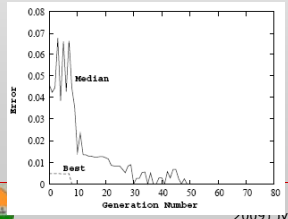
- Can potentially indicate which regions are difficult to converge
- Termination crietrion

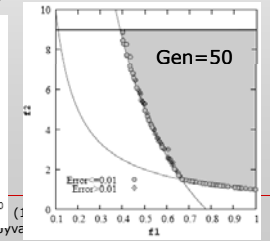
A Constrained Problem

$$\begin{aligned} &\text{Minimize} \quad f_1(\mathbf{x})=x_1,\\ &\text{Minimize} \quad f_2(\mathbf{x})=\tfrac{1+x_2}{x_1},\\ &\text{subject to} \quad g_1(\mathbf{x})\equiv x_2+9x_1\geq 6,\\ &\qquad\qquad g_2(\mathbf{x})\equiv -x_2+9x_1\geq 1,\\ &\qquad\qquad 0.1\leq x_1\leq 1, \quad 0\leq x_2\leq 5. \end{aligned}$$



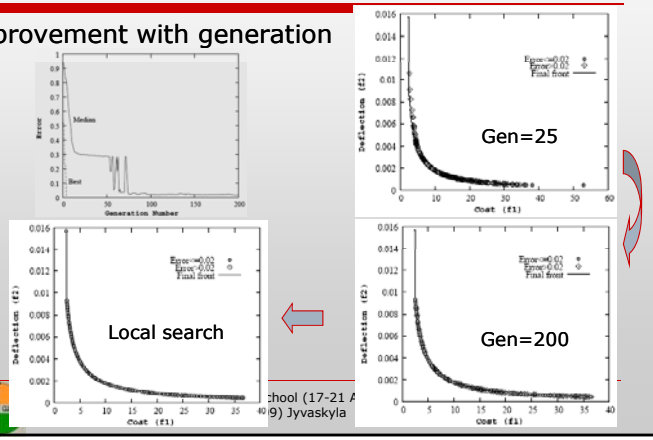
- Constrained part is difficult to converge
- Local search after 50th generation makes zero error





Welded Beam Design

Improvement with generation



Welded Beam Design (cont.)

- All solutions have zero or small error
- Largrange multipliers are non-negative
- Minimum cost solution is relatively difficult to find

h	ℓ	t	b	f_1	f_2	Error	λ_1	v_1	v_3	v_4	v_5	v_6
0.244	6.214	8.298	0.244	2.381	0.016	3.780e-04	1.054e-01	0.0627	0.0437	0.000	0.145	
0.235	5.178	10.000	0.235	2.486	0.009	0.000e+00	5.891e-03	0.0071	0.0035	1.0e-06	0.022	
0.310	3.606	10.000	0.310	3.003	0.007	0.000e+00	3.193e-03	0.0036	0.0013		0.016	
0.438	2.350	10.000	0.438	3.942	0.005	0.000e+00	1.541e-03	0.0018	0.0003		0.011	
0.653	1.470	10.000	0.714	6.009	0.003	0.000e+00	5.780e-04	0.0008			0.007	
0.857	1.075	10.000	1.263	10.033	0.002	0.000e+00	1.900e-04	0.0003			0.004	
1.139	0.777	10.000	2.235	17.006	0.001	0.000e+00	6.200e-05	0.0001			0.002	
1.341	0.644	10.000	3.070	22.911	0.001	0.000e+00	3.300e-05	7.7e-05			0.001	
1.558	0.542	10.000	4.088	30.055	0.001	0.000e+00	1.900e-05	4.9e-05			0.001	
1.223	0.751	10.000	5.000	36.723	0.000	0.000e+00	0.000e+00				0.001	0.0004

Conclusions

- Hybrid EMOs are fast and reliable
  - Using local search methods
  - Using theoretical optimality conditions
- Performance metrics are important for termination
- Credible applications using various optimization methods
  - Way forward for practical applications

NSGA-II Code Demonstration

- http://www.iitk.ac.in/kangal/soft.htm
- nsga2.c: An EMO procedure