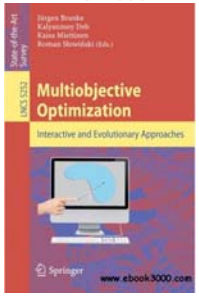
Lecture 10:
EMO and Decision Making

- Coupling EMO with MCDM techniques
  - Cone domination for partial frontier
  - Reference point approach
  - Reference direction approach
  - 'Light beam' search search
- Finding knees
- Estimating nadir point
- Interactive EMO (I-EMO)
- Recent developments

EMO and Decision-Making

- Finding a P-O set (using EMO) is half the story
- How to choose one preferred solution (MCDM)
- First EMO, then MCDM <- Early concept
- First MCDM, then EMO to find preferred solution(s)
  - Reference point based EMO, light beam search etc. (Deb & students, 2006, 2007; Molina et al., 2006)
- EMO and MCDM interactively
  - I-MODE (Deb and Chaudhuri, EMO-07)
  - Recent emphasis (Branke et al., EMO-09, Koksalan et al., 2009, Deb et al., 2009)

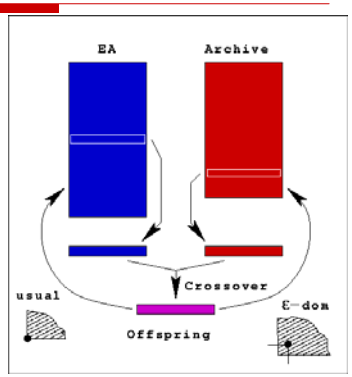
A new book



ε-MOEA: Using ε-Dominance

- EA and archive populations evolve
- One EA and one archive member are mated
- Archive update using ε-dominance
- EA update using usual dominance

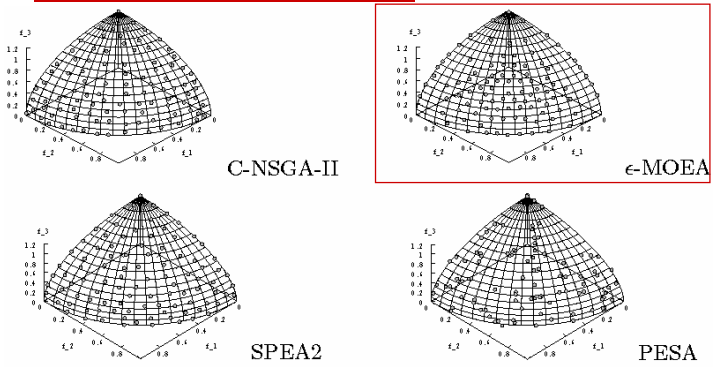
Deb, Mohan & Mishra (ECJ-2005)



Comparative Study on Three-Obj. DTLZ Functions

EMO	Convergence measure		Sparsity		Time (sec)	
	Average	Std. Dev.	Average	Std. Dev.	Average	Std. Dev.
DTLZ2						
NSGA-II	0.0137186	0.0020145	0.931111	0.0124474	17.16	0.196
C-NSGA-II	0.0107455	0.0008424	**0.999778**	0.0004968	7837.42	81.254
PESA	**0.0106292**	0.0025483	0.945778	0.0309657	88.01	12.901
SPEA2	0.0126622	0.0009540	0.998889	0.0007855	2164.42	19.858
ε-MOEA	0.0108443	0.0002823	0.999104	0.0009316	**2.01**	0.032
DTLZ3						
NSGA-II	0.0149156	0.01028	0.839228	0.02961	136.45	31.080
C-NSGA-II	0.0202315	0.00898	0.995521	0.00613	24046.03	4690.032
PESA	0.0130633	0.00449	0.722296	0.02785	89.49	12.527
SPEA2	0.0122429	0.00194	**0.999771**	0.00031	9080.81	963.723
ε-MOEA	**0.0122190**	0.00223	0.993207	0.00974	**9.42**	2.180
DTLZ5						
NSGA-II	0.00208342	11.976e-05	0.953778	0.00992	11.49	0.036
C-NSGA-II	0.00256138	30.905e-05	0.996667	0.00314	1689.16	81.365
PESA	0.00094626	11.427e-05	0.772110	0.02269	53.27	11.836
SPEA2	0.00197846	16.437e-05	**1.000000**	0.00000	633.60	14.082
ε-MOEA	**0.000953623**	4.892e-05	0.980867	0.01279	**1.45**	0.051

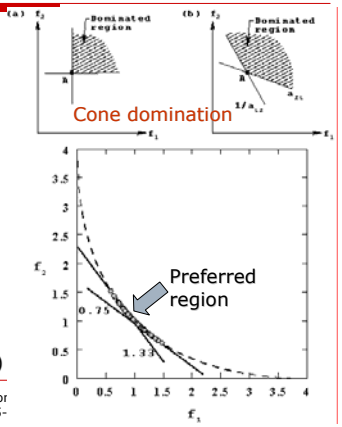
Test Problem DTLZ2



Finding a Partial Pareto Frontier: Cone Dominance

- Using a DM's preference (not a solution but a region)
- Guided domination principle: Biased niching approach
- Weighted domination approach

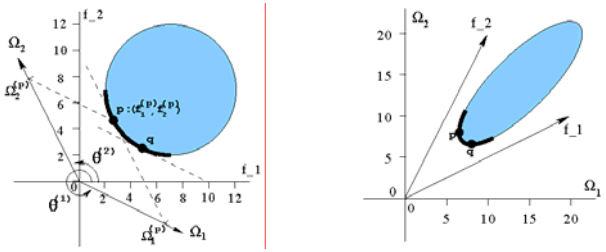
(Branke et al., 2001; Deb, 2001)



Distributed Computing of Pareto-Optimal Set

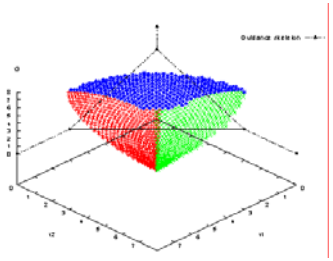
Deb, Zope & Jain (EMO-2003)

- Guided domination concept to search different parts of Pareto-optimal region
- Distributed computing of different parts

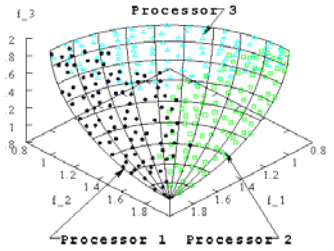


Distributed computing: A Three-Objective Problem

- Spatial computing, not temporal



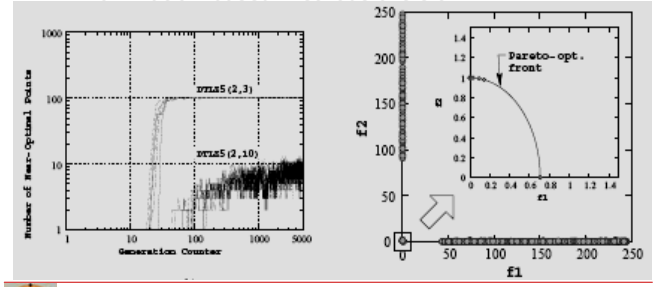
Theory



NSGA-II Simulations

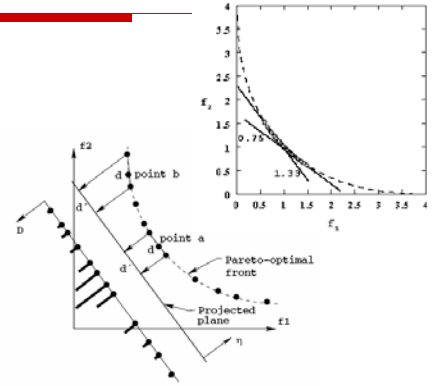
Preference-Based EMO

- EMO (NSGA-II) not efficient for many objectives
  - Large number of points needed
  - Domination-based methods are slow
  - Visualization difficulties



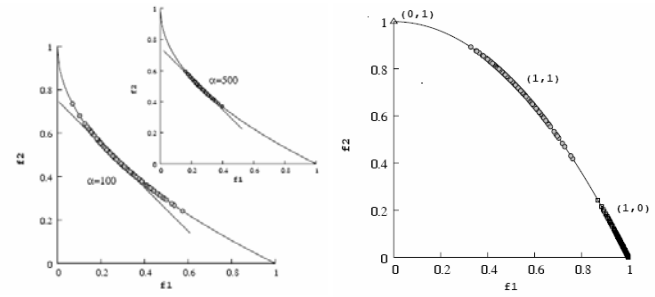
EMO for a Biased Distribution

- Choose a hyper-plane
- Project points on it
- Compute two distances: d and d'
- Compute D=d(d'/d)^a
- Point b has small D
- Point a has large D



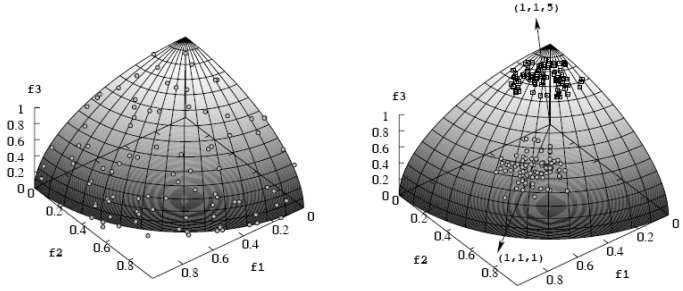
Biased Distribution in NSGA-II

ZDT1 ZDT2: a=100



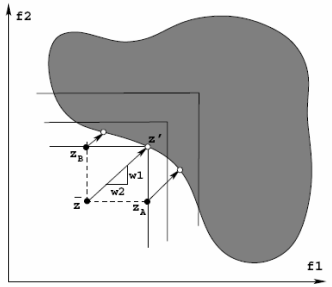
Biased NSGA-II (cont.)

Three-objective Problems: a=0 and a=500



Reference Point Based EMO Approach

- Wierzbicki, 1980
- A P-O solution closer to a reference point
  - Multiple runs
  - Too structured
- Extend for EMO
  - Multiple reference points in one run
  - A distribution of solutions around each reference point



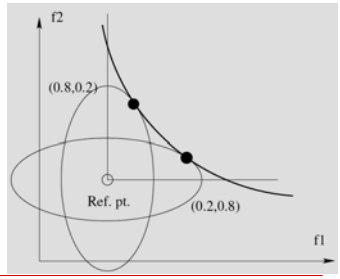
Reference Point Based NSGA-II

- Modify crowding distance operator
- Multiple reference points given
- For each RP, rank members in increasing Euclidean distance in objective space
- Prefer members with smaller ranks
- To maintain a spread, use a clearing idea
  - Pick a solution randomly
  - Clear all ε-neighboring solutions (high rank)
  - Continue with un-cleared solutions

Add Flexibility Through Weights

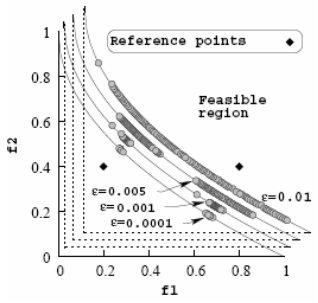
- Calculate a weighted Euclidean distance
- By changing weights, different regions can be found
- Idea should work in non-convex problems as well

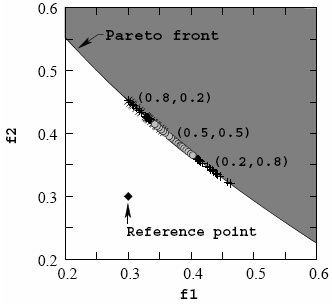
$$d_{ij} = \sqrt{\sum_{i=1}^{M} w_i \left(\frac{f_i(\mathbf{x}) - \overline{z}_i}{f_i^{\max} - f_i^{\min}} \right)^2}$$



Simulation Studies: Effect of ε and weights on ZDT1

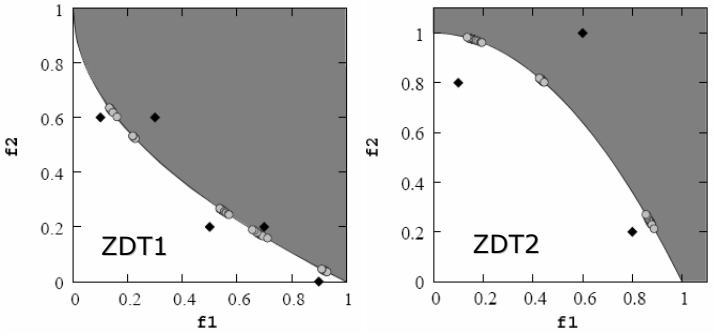
- Larger ε, more spread
- Region depends on weights





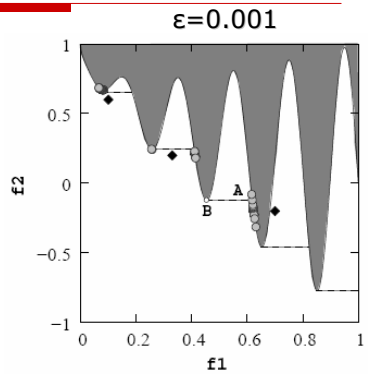
Effect of Convexity (ZDT1 vs. ZDT2)

ε=0.001, w=(0.5,0.5)



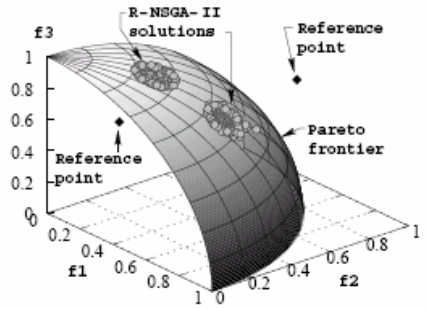
Effect on ZDT3

- For a RP, P-O points on two non-overlapping fronts found
- Versatile method
- May find non-P-O solutions
  - B dominates A
  - Reason: No real motivation to find B



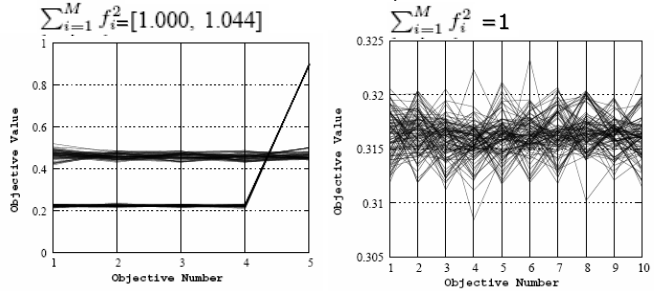
Three-Objective Problems

- DTLZ2 with two RPs (ε=0.01)



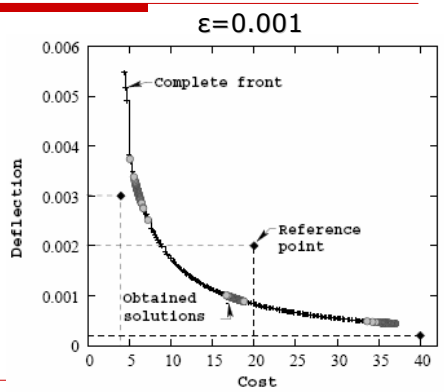
Higher Objectives

- Five objectives: (0.5,…,0.5) and (0.2,…0.2,0.8)
- 10 objectives: (0.25,…,0.25) $f_i=1/\sqrt{10}=0.316$



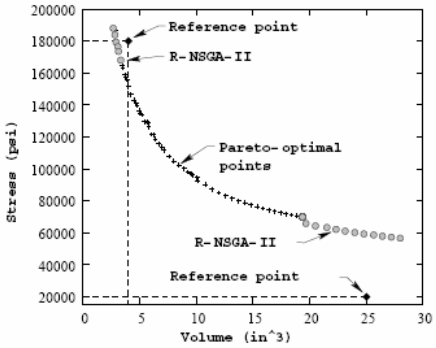
Welded Beam Design Problem

- 4 constraints
- Quicker
- More points near regions of interest
- RPs:
  - (4,0.0030)
  - (20,0.0020)
  - (40,0.0002)



Spring Design Problem

- 8 constraints
- Mixed-discrete variables
- P-O front is discrete
- ε=0.001
- RPs:
  - (4,180000)
  - (25,10000)

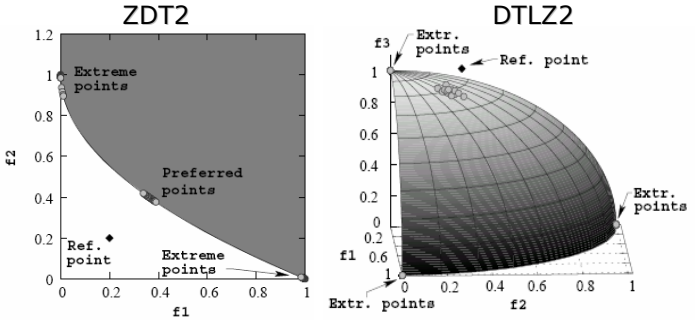


Modified Approach

- In addition to preferred points, find the extreme points
  - Provides an idea of extent of front
- Add 2M reference points with extreme population members
  - Forces extreme solutions to exist
- Helpful in maintaining diversity
  - Helps in single RP cases

Simulation with Modified Approach

Extreme points are found as well



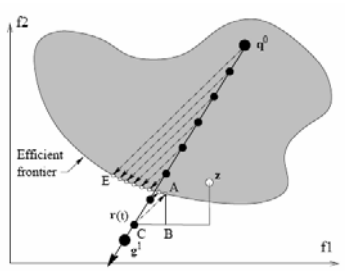
Reference Direction Based MCDM Approach

Korhonen and Laakso, 1986

- Choose a point q^0 and another point g^1
- Get direction $d^k=g^k-q^{k-1}$
- Solve for various t

$$\begin{aligned}\text{Minimize}_{\mathbf{z}} \quad & s(\mathbf{z},\mathbf{r},\mathbf{w}) = \max_{i\in I}(z_i - r_i)/w_i,\\ \text{subject to} \quad & \mathbf{r}(t) = \mathbf{q}^{k-1} + t\mathbf{d}^k.\end{aligned}$$

- Choose most preferred solution (q^k) and another point g^{k+1}



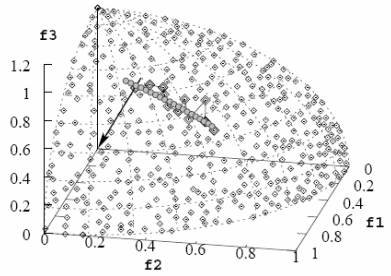
A way to get to a single preferred solution

Reference Direction Method

- A projection of reference direction on the Pareto-optimal frontier
- Provides an idea of how solutions on the P-O front move while moving along reference direction (Pareto race)
- Korhonen and his students derived analytical expressions on linear problems
- AHP process is used to determine reference direction (Korhonen and Karaivanova, 1999)
  - Further extensions exist

Reference Direction Based EMO

- Simple-minded approach
  - Find EMO solutions first and then choose best solution for each *t*
  - Not an accurate approach
- EMO cannot find too many points
- Finding unwanted points is a waste



Proposed RD-NSGA-II

- *r(t)* on reference direction are marked
- For each *t,* achievement scalarizing function *s(z,r(t),w)* is computed for all population members *z* and the smallest one is put in the first front
- These solutions are discounted and above step is repeated to construct the second front and so on
- For multiple reference directions, follow above for each $r_i(t)$

Simulation Studies: Test Problems

- ZDT2 with standard parameter setting with w=(0.5,0.5)
- $q^0=(\alpha_1,\beta_1)$, $g^1=(\alpha_2,\beta_2)$
- $\gamma=(\beta_1-\beta_2)/(\alpha_1-\alpha_2)$
- Ranges on f1:

For $\gamma > 1$:

$$\max\left\{0, \frac{-1+\sqrt{5-4(\beta_1-\alpha_1)}}{2}\right\} \le f_1 \le \min\left\{1, \frac{-1+\sqrt{5-4(\beta_2-\alpha_2)}}{2}\right\}$$

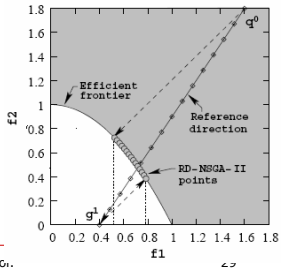
For $\gamma < 1$:

$$\max\left\{0, \frac{-1+\sqrt{5-4(\beta_2-\alpha_2)}}{2}\right\} \le f_1 \le \min\left\{1, \frac{-1+\sqrt{5-4(\beta_1-\alpha_1)}}{2}\right\}$$

For $\gamma = 1$:

$$f_1 = \frac{-1+\sqrt{5-4(\beta_1-\alpha_1)}}{2}.$$

- Example: $q^0=(1.6,1.8)$, $g^1=(0.4,0.0)$
- $\gamma=1.5$
- $0.525 \le f_1 \le 0.785$

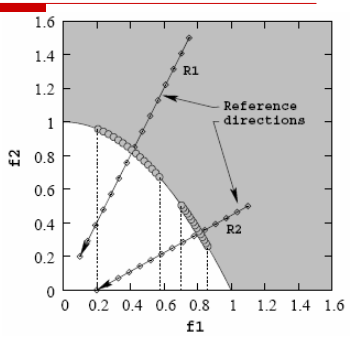


N=100, maxgen=500

Two Reference Directions

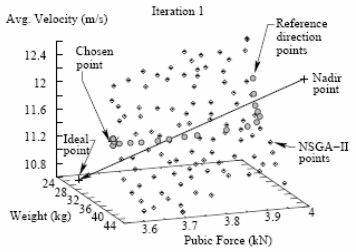
	(α_1,β_1)	(α_2,β_2)
R_1	$(0.75,1.50)^T$	$(0.10,0.20)^T$
R_2	$(1.10,0.50)^T$	$(0.20,0.00)^T$

- First RD: f1 range (0.207, 0.572)
- Second RD: f1 range (0.704, 0.860)
- Matches quite well with theory



Car Side Impact Problem

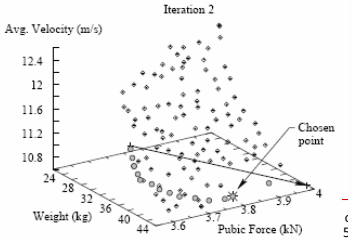
- Three objectives, 10 constraints, seven variables
- Utility: Min. avg. force at abdomen and pubic area
- q^0=nadir point=(42.69,4.00,12.44)
- g^1=ideal point=(24.37,3.59,10.61)
- 25 points -> q^1=(35.95,3.56,11.53)

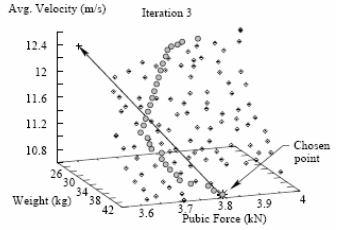


- Avg. vel (f_3) smaller
- g^2=(42.69,4.00,10.61)
- RD: q^2 to g^2
- 15 points

Second and Third Iterations

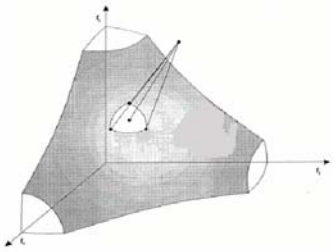
- q^2=(40.98,3.81,10.61), f_3 is reduced
- Reduce f_1 and f_2: g^3=(24.37,3.59,12.44)
- q^3=(40.92,3.81,10.61)
- Consider q^2 and q^3 close and terminate
- Declare x-vector





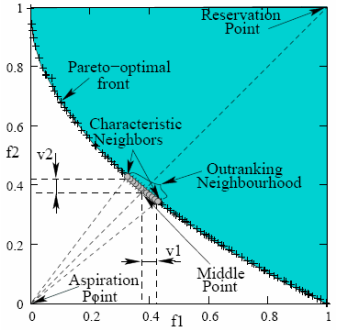
Light Beam Search
(Jaszkiewicz and Slowinski, 1999)

- User supplies aspiration and reservation points
- *Middle* point found using aspiration point as reference point
- Then, local outranking relation to obtain non-dominated points
  - User defines three thresholds per objective



LBM Basics

- Veto threshold *v*
- *a* outranks *b (aSb)*
  - In objectives for which a is smaller than *b*, difference is not larger than *v*
- Restricts non-dominated solutions within a neighborhood
- Flip side: specify *v*



Proposed LBS-NSGA-II

- Supplied parameters:
  - Aspiration (z^r) and reservation (z^v) points
  - Veto threshold v and diversity index ε
- For each non-dominated front member
  - Compute

$$d = \max\{\lambda_j(f_j - z_j^r)\} + \rho\sum_{j=1}^{M}(f_j - z_j^r) \text{ where } \lambda_j = \frac{1}{z_j^v - z_j^r}$$

  - Pop. member with smallest *d* is middle pt (z^c)
  - Augmented achievement scalarizing fn. avoids weak P-O points
  - $\rho = 10^{-6}$

Proposed LBS-NSGA-II (cont.)

- Solutions which outrank z^c compute

$$\delta = \max(f_j - z_j^c), \quad j = 1, \ldots, M.$$

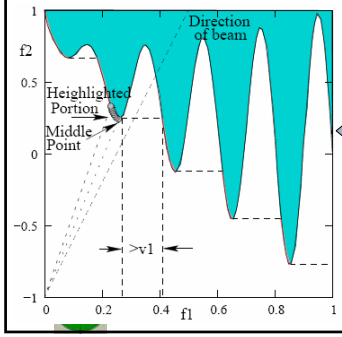
- Smaller δ is assigned a better rank
- Two solutions within ε are assigned worse rank
- Solutions which do not outrank z^c are assigned worse ranks
- Multiple beams: Use the best rank from all beams

Simulation Studies: Test Problems

N=100, maxgen=500

- ZDT1 with standard parameter setting and v=(0.05,0.05), ε=0.01

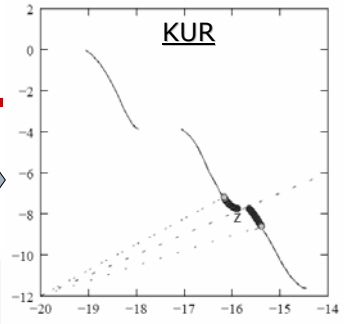


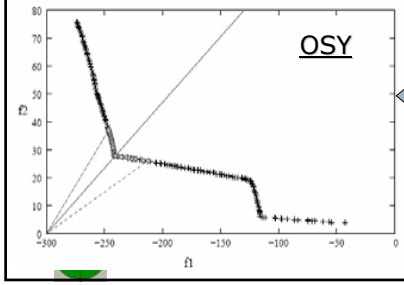


- ZDT3 with same parameters
- Right of middle pt are P-O points

Test Problems KUR and OSY

- Both sides with discontinuities obtained
- *v*=(0.5,0.8), ε=0.01

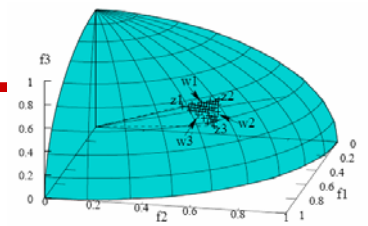




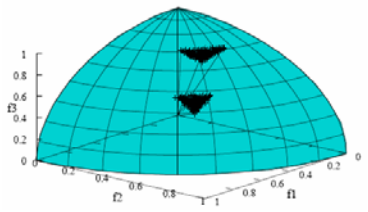
- Both sides of a knee
- *v*=(30,10), ε=0.01

Three-Objective DTLZ2

- Ideal and nadir
- Middle point = (0.577,0.577,0.577)
- *v*=(0.05,0.05,0.05)
- Theory: Triangular shape
- ε=0.025



- Two light beams with ε=0.010
  - More dense points

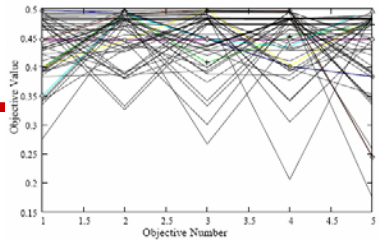


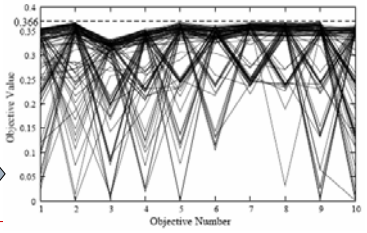
Five and Ten-Obj. DTLZ2 Problems

- Ideal and nadir
- Middle pt., $m_i=1/\sqrt{5}$
- v_i=0.05, ε=0.04
- Solutions are around 0.447
- Worst value does not exceed 0.447+0.05

10 Objectives:

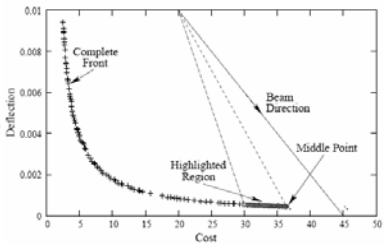
- $m_i=1/\sqrt{10}=0.316$
- Worst ≤ 0.366





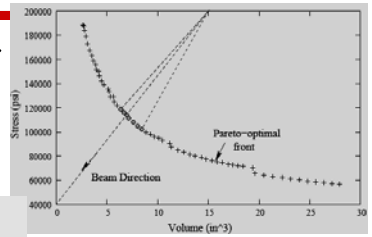
Welded Beam Design

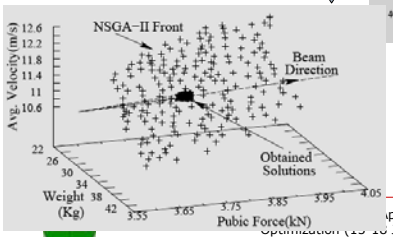
- Two objectives
- V=(5,0.0001)
- ε=0.01
- Interestingly, middle point is not along beam direction



Spring Design and Car side impact problems

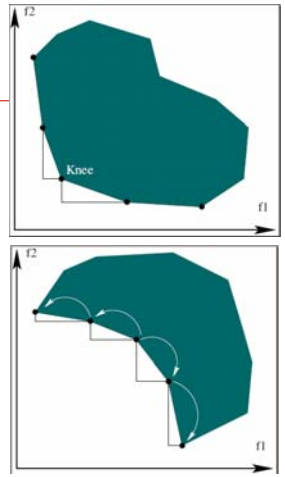
- Spring design
  - Discrete problem
- Car side impact
  - Three objectives





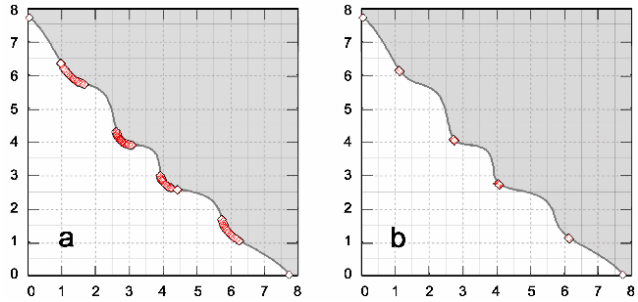
Knee Based EMO

- Find only the knee or near-knee solutions
- Knees are important solutions
  - Not much motivation to move out from knees
  - A large gain for a small loss in any pair of objectives
- Non-convex front
  - No knee point
  - Extreme solutions are attractors



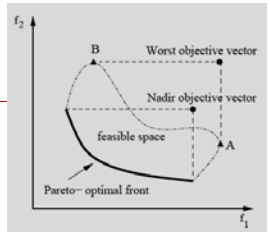
Finding Knee Solutions

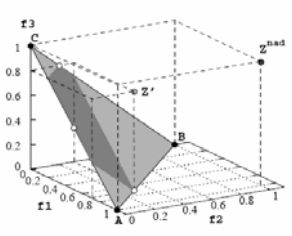
Branke et al. (2004) for more details



Nadir Point and EMO

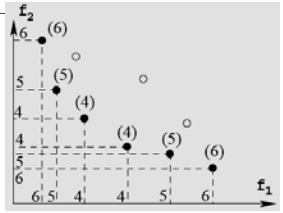
- Important for knowing range and normalization of objectives
- Difficult to find using classical method
  - Pay-off table method does not work





EMO for Finding Nadir Point

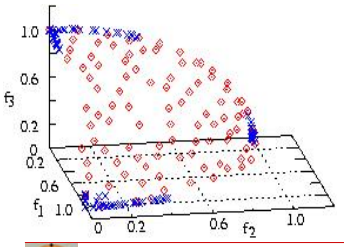
- Emphasize only extreme points
- M≤3 find complete front, else use extremized crowded NSGA-II

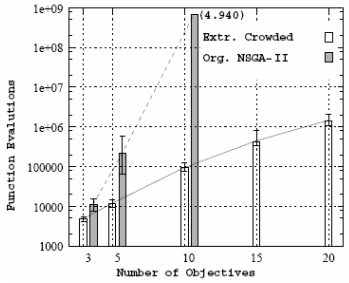


Test Problem	Pop. size	Number of generations								
		NSGA-II			Worst crowd. NSGA-II			Extr. crowd. NSGA-II		
		Best	Median	Worst	Best	Median	Worst	Best	Median	Worst
ZDT1	100	32	44	80	31	41	45	31	42	48
ZDT2	100	45	56	83	44	72	97	34	67	93
ZDT3	100	33	40	55	33	40	113	28	36	45
ZDT4	100	176	197	257	148	201	224	165	191	219
ZDT6	100	137	151	161	125	130	143	126	132	135

EMO for Finding Nadir Point (cont.)

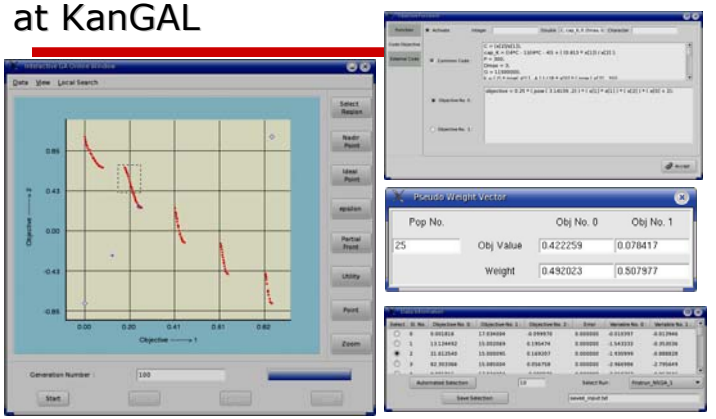
- DTLZ problems extended up to 20 objectives



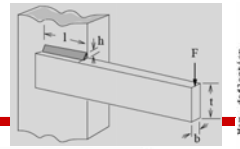


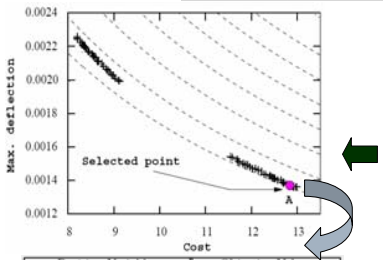
I-MODE Software Developed at KanGAL

Deb and Chaudhuri, EMO-2007

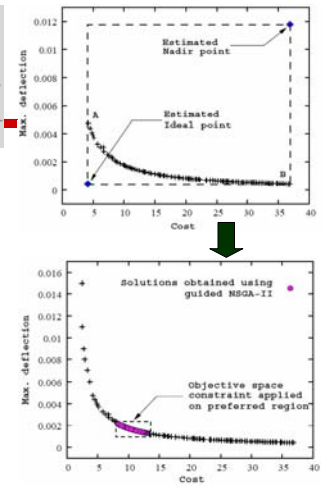


Example:





Decision Variables				Objective Values	
h (in)	l (in)	t (in)	b (in)	Cost	Max. Deflection (in)
0.917	1.009	9.856	1.672	12.838	0.001371

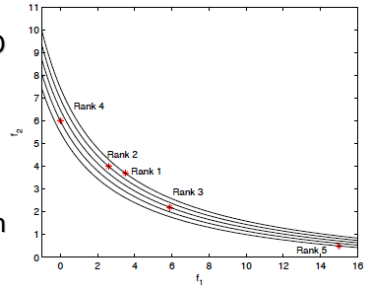


Recent Developments on EMO and Decision-Making

- Interactive EMO+MCDM approaches
- Start with EMO
- Involve decision maker for preferences
- Preferred EMO
- Repeat the above till no further improvement
- Branke et al., 2009, Deb et al., 2009

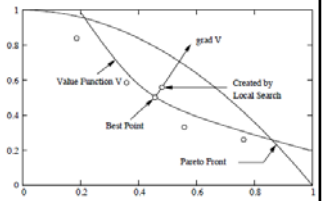
Interactive Preference-Based EMO (IP-EMO) (Deb et al., 2009)

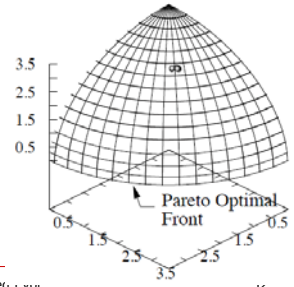
- Cluster a few (say 5) solutions from an EMO generation
- Ask the DM to rank
- Develop an utility function based on preferred rank
- Use the utility function to proceed with EMO
- Repeat the process after every T generations



IP-EMO (cont.)

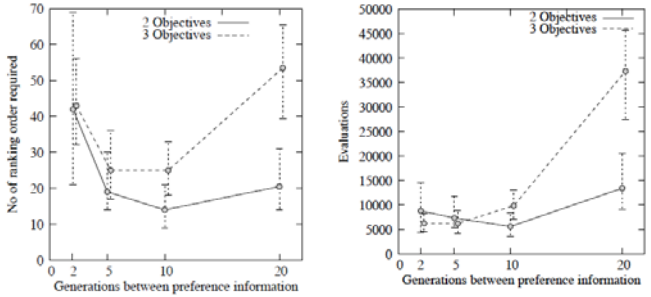
- Termination based on improvements, if any, via a local search
  - Guarantees convergence to a PO point
- Results on DTLZ2





Optimal Generation Gap

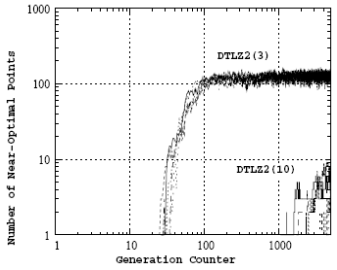
- Between 5-10 generations seem good



EMOs for Many Objectives

- Up to 4 objectives, scaling up is okay
- Popsize=200, maxgen=5000
- Only 4% points are near P-O front for 10-obj
- Both convergence and diversity suffer
- Are EMOs useless for >4 objectives?

(Deb and Saxena, CEC-06)



Two High-Dimensional Scenarios for EMO's Niche

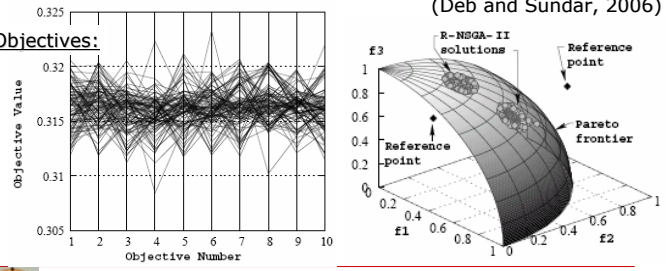
- A preferred P-O region of interest, not the entire P-O front
  - A few points may still be fine
  - Dominance concept works again
  - Visualization still a problem
- Certain structured high-dimensional problems
  - With redundant objectives
  - P-O front is lower-dimensional
  - Often found in practice

Reference Point Based EMO

- Ranking based on closeness to each reference point
- Clearing within each niche with ε
- 10 objectives: (0.25,…,0.25) $f_i=1/\sqrt{10}=0.316$

(Deb and Sundar, 2006)

10 Objectives:

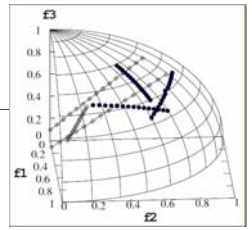


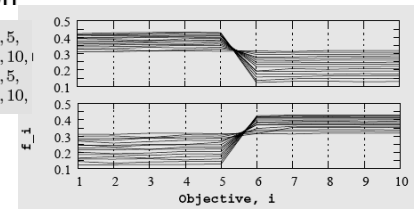
Reference Direction Based EMO

- For each point on a reference direction, rank solns. based on [illegible] function

10 Objectives:

$$R_1: \begin{cases} q_i^0 = 0.4, g_i^1 = 0, & i = 1, \ldots, 5, \\ q_i^0 = 0.1, g_i^1 = 0, & i = 6, \ldots, 10, \end{cases}$$
$$R_2: \begin{cases} q_i^0 = 0.1, g_i^1 = 0, & i = 1, \ldots, 5, \\ q_i^0 = 0.4, g_i^1 = 0, & i = 6, \ldots, 10, \end{cases}$$





(Deb and Kumar, 2006)

PCA Based NSGA-II Procedure

CEC-06

Step 1: Set an iteration counter $t = 0$ and initial set of objectives $\mathbb{I}_0 = \{1, 2, \ldots, M\}$.
Step 2: Initialize a random population for all objectives in the set $\mathbb{I}_t$, run an EMO, and obtain a population P_t.
Step 3: Perform a PCA analysis on P_t using $\mathbb{I}_t$ to choose a reduced set of objectives $\mathbb{I}_{t+1}$ using the predefined TC. Steps of the PCA analysis are as follows:
1. Compute the correlation matrix.
2. Compute eigenvalues and eigenvectors and choose non-redundant objectives using the procedure discussed in Sections 6.1 and 6.1.
3. Reduce the number of objectives further, if possible, by using the correlation coefficients of the non-redundant objectives found in item 2 above, using the procedure discussed in Section 6.1.

Step 4: If $\mathbb{I}_{t+1} = \mathbb{I}_t$, stop and declare the obtained front. Else set $t = t + 1$ and go to Step 2.

Three main Components

I. Compute Correlation Matrix

II. Interpret Eigen-values/vectors

III. Interpret Further Reduced Correlation Matrix

Only this would differ for C-PCA-NSGA-II, MVU-NSGA-II

Structured Problems

- Identify redundant objectives
- EMO+PCA in iterations

Iter.1

Iter. 1 : PCA-1 (58.83 % variance)	f_7	f_{10}
PCA-2 (28.26 % variance)	f_1	
PCA-3 (06.53 % variance)		f_8
PCA-4 (03.27 % variance)		f_8

10-objective DTLZ5 problem

	f_1	f_7	f_8	f_{10}
f_1	+	-	+	-
f_7	-	+	+	-
f_8	+	+	+	-
f_{10}	-	-	-	+

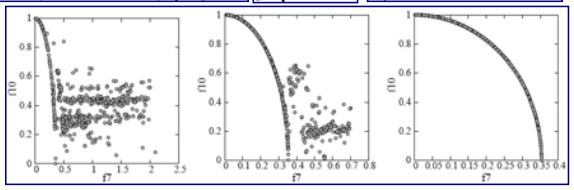
Iter.2

Iter. 2 : PCA-1 (94.58 % variance)	f_7	f_{10}
PCA-2 (4.28 % variance)	f_8	

	f_7	f_8	f_{10}
f_7	+	+	-
f_8	+	+	-
f_{10}	-	-	+

	e1:0.9458	e2:0.0428	
f_7	+0.543	-0.275	c7=0.5253
f_8	+0.457	+0.672	c8=0.4610
	PCA1	PCA2	

Saxena and Deb (CEC-2006, EMO-2007, CEC-2007)



Limitations of PCA-NSGA-II ….(i)

DTLZ5(2,10)

Iter. 1	f_1 f_7 f_8 f_{10}
Iter. 2	f_7 f_{10}
Iter. 3	f_7 f_{10}

Why f7

DTLZ5(2,20)

Iter. 1	f_2, f_5, f_{10}, f_{12}, f_{16} f_{18} f_{20}
Iter. 2	f_5 f_{18} f_{20}
Iter. 3	f_{18} f_{20}
Iter. 4	f_{18} f_{20}

Why f18

DTLZ5(2,30)

Iter. 1	f_1 f_8 f_{30}
Iter. 2	f_8 f_{30}
Iter. 3	f_8 f_{30}

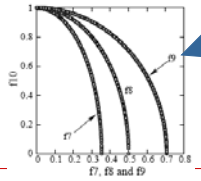
Why f8

DTLZ5(2,50)

Iter. 1	f_{13} f_{50}
Iter. 2	f_{13} f_{50}

Why f13

True dimensionality is 2 (last and any one of the rest objectives)
To this extent, results are acceptable
But
PCA-NSGA-II was expected to declare those objectives as critical, which apart from being in conflict with each other also account for variance larger than those declared redundant.



f9, f10 here in particular,
In general last two objectives must come up as critical for DTLZ5(2,M) problems

Kernel PCA

- **Given:** a data set $\mathbf{X}$ of size $M \times N$
- **Select a particular kernel form and compute the Kernal Gram Matrix 'K', given by:** $K_{ij} = \kappa(\mathbf{x}_i, \mathbf{x}_j)$ ($\mathbf{x}_i$ representing columns of $\mathbf{X}$)
- **'K', representing the covariance matrix in the feature space turns to be of dimension 'N*N'**

In context of EMO, for a 50-obj. (M) problem with pop. size, say 10,000 (N), the eigen-decomposition of 'K' will be computationally-punitively costly.

Further, the interpretation module (of eigen-values/vectors) of PCA-NSGA-II is designed for a covariance matrix of size : 'M*M'

Hence, employment of Kernel PCA in its natural form is not viable

Correntropy PCA : (C-PCA) …(i)

- **Based on generalized correlation function 'Correntropy' (Xu et al, 2006): V(x,y)**

❖ Basis for non-linear PCA

- Its positive definite, as a positive-definite kernel is an argument to the Expectation $\mathbf{V}(\mathbf{x},\mathbf{y}) = E[\kappa(\mathbf{x},\mathbf{y})]$
- By 'Characterization of Kernel Theorem' : there must exist a unique RKHS (Reproducing Kernel Hilbert Space) associated with the correntropy function, say $\langle \Pi(\mathbf{x}), \Pi(\mathbf{y}) \rangle = \mathbf{V}(\mathbf{x},\mathbf{y})$

- **Unlike in K-PCA, where data is transformed into feature space sample by sample; here in C-PCA, data is mapped component wise**

Summary of Results

	Learned Kernel		Polynomial Kernel			Gaussian Kernel		
	K = M-1	√M	n = 3	2	1	n = 3	2	1
DTLZ5(2,5)	✓	✓	✓	✓	✓	✓	✓	✓
DTLZ5(2,10)	✓	✓	✓	✓	✓	✓	✓	?
DTLZ5(2,20)	✓	✓	✓	✓	✓	✓	?	?
DTLZ5(2,30)	✓	?	✓	?	✓	✓	✓	?
DTLZ5(2,50)	✓	✓	✓	✓	✓	✓	✓	✓

K=M-1 is always found successful: non-linear dimensionality reduction without having to ponder over choice of a kernel form and its parameters

Conclusions

- EMO must be coupled with decision-making aides for a complete problem solving task
- Borrow ideas from MCDM literature
- Combination with EMO is tricky
- Interactive methods are of current interest
- EMO for large number of objectives
  - Aim for preferred solutions
  - Hope to find redundant objectives