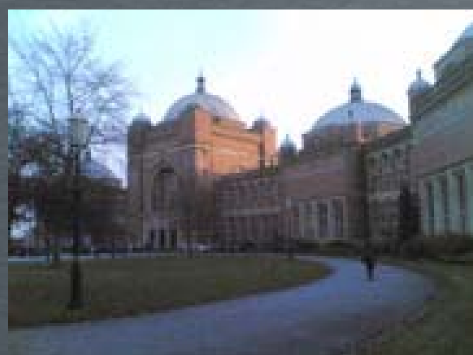
Powders & 'Peer-Pressure': Pitfalls & Progress

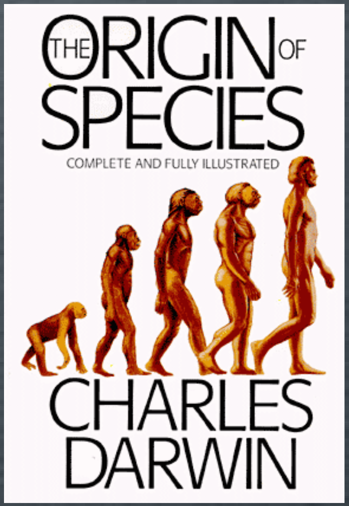


Maryjane Tremayne

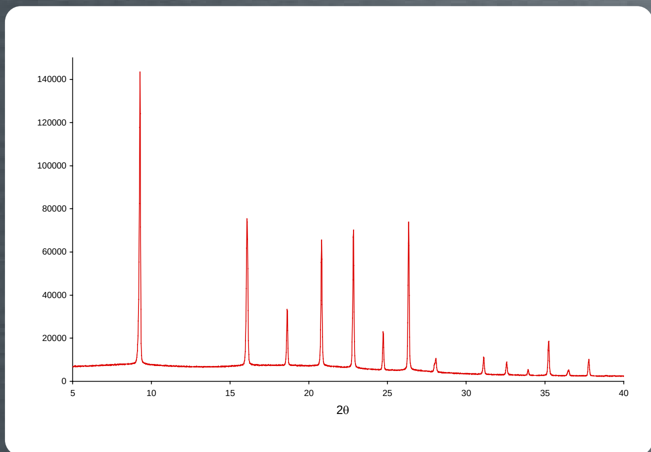
School of Chemistry, University of Birmingham, Edgbaston, Birmingham, UK.

Cultural Differential Evolution

Using **social** evolution
to guide **biological** evolution



Direct Space Methods for SDPD



Intensities

Structure solution

Predict trial structure

Compare simulated & exp data, R_{wp}

Global optimisation locates best structure

Optimum structure solution

Positions

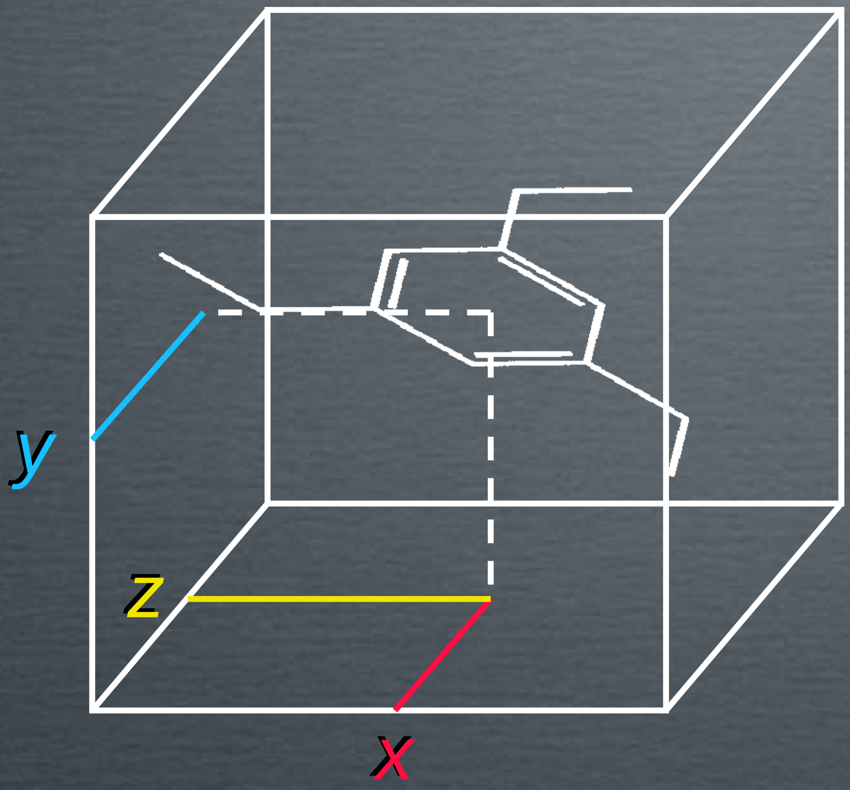
Indexing

Unit cell

Rietveld refinement

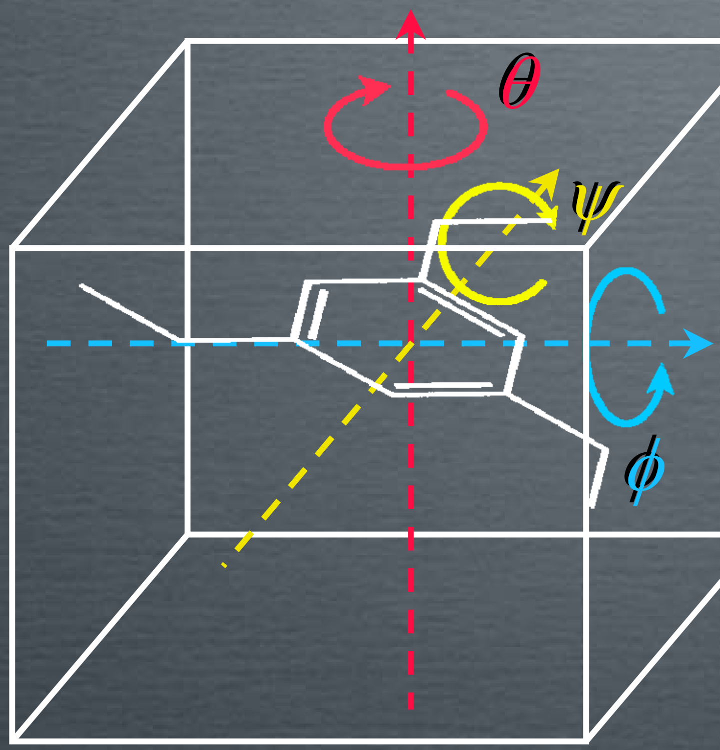
Final crystal structure

Structural Model



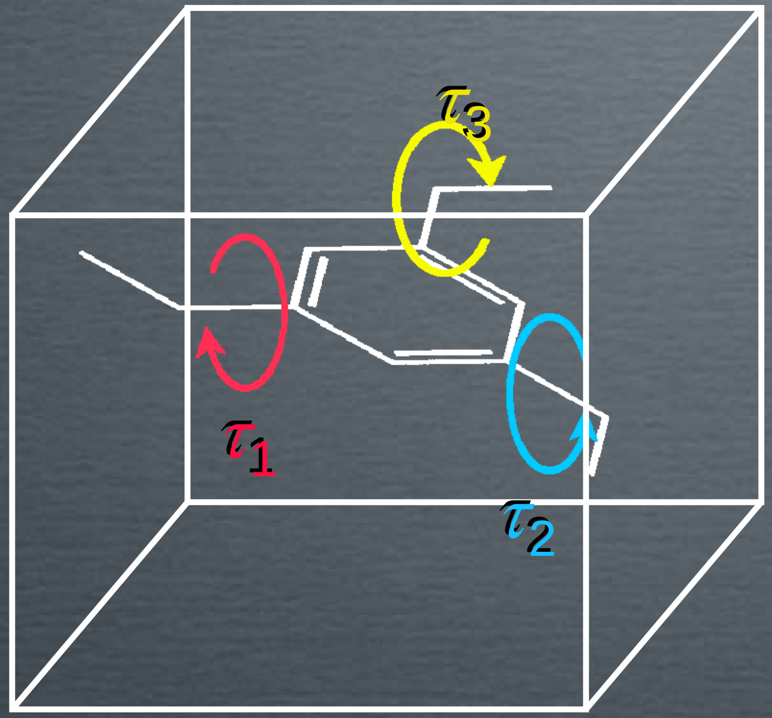
- Position: x, y, z
- Orientation: θ, ϕ, ψ
- Molecular conformation: $\tau_1 \ldots \tau_n$

Structural Model



- Position: x, y, z
- Orientation: θ, ϕ, ψ
- Molecular conformation: $\tau_1 \ldots \tau_n$

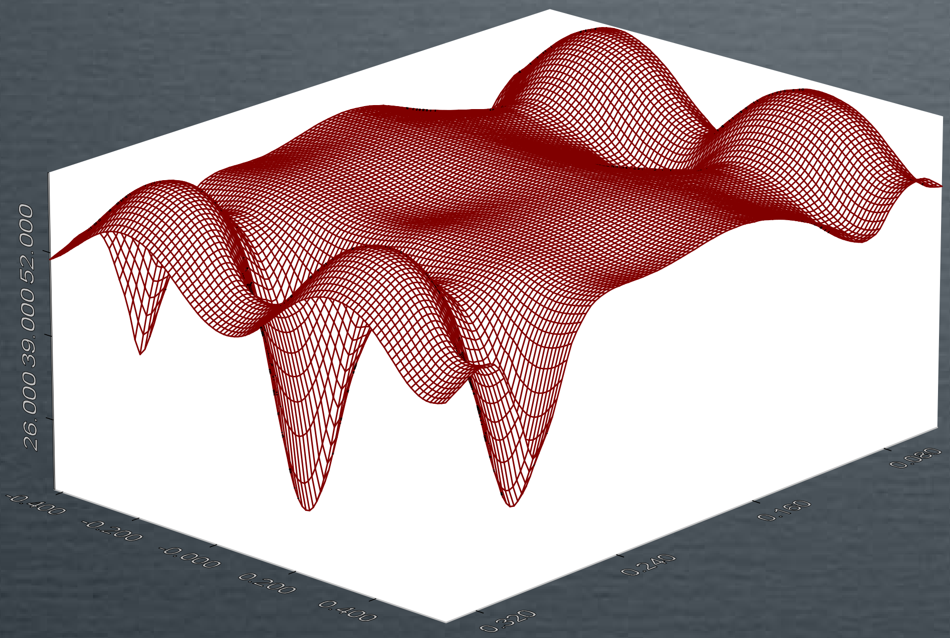
Structural Model



- Position:
  x, y, z
- Orientation:
  θ, ϕ, ψ
- Molecular conformation:
  $\tau_1 \ldots \tau_n$

Global Optimisation

- Multi-dimensional search problem: $(x,y,z,\theta,\phi,\gamma,\tau_1 \ldots \tau_n)$ per molecule or fragment
- Single objective – *minimum in R_{wp}*



- simulated annealing
- Monte Carlo
- parallel tempering
- hybrid MC/MD

Evolutionary Algorithms in SDPD

- Evolutionary algorithms - population of trial structures; mating, mutation and natural selection until global minimum found
- Each member of the population defined by genetic code $(x,y,z)[0\text{-}1](\theta,\phi,\gamma,\tau_1\ldots\ldots\tau_n)[0\text{-}360]$

→ genetic algorithms

Kariuki et al., *Chem Phys Lett.*, 280, 189, (1997)
Harris, Johnston & Kariuki, *Acta Cryst.*, A54, 632, (1998)

→ differential evolution

Seaton & Tremayne, *Chem Commun*, 880, (2002)
Tremayne, Seaton & Glidewell, *Acta Cryst.*, B58, 823, (2002)
Chong, Seaton, Kariuki & Tremayne, *Acta Cryst.*, B62, 862, (2006)

Differential Evolution (DE)

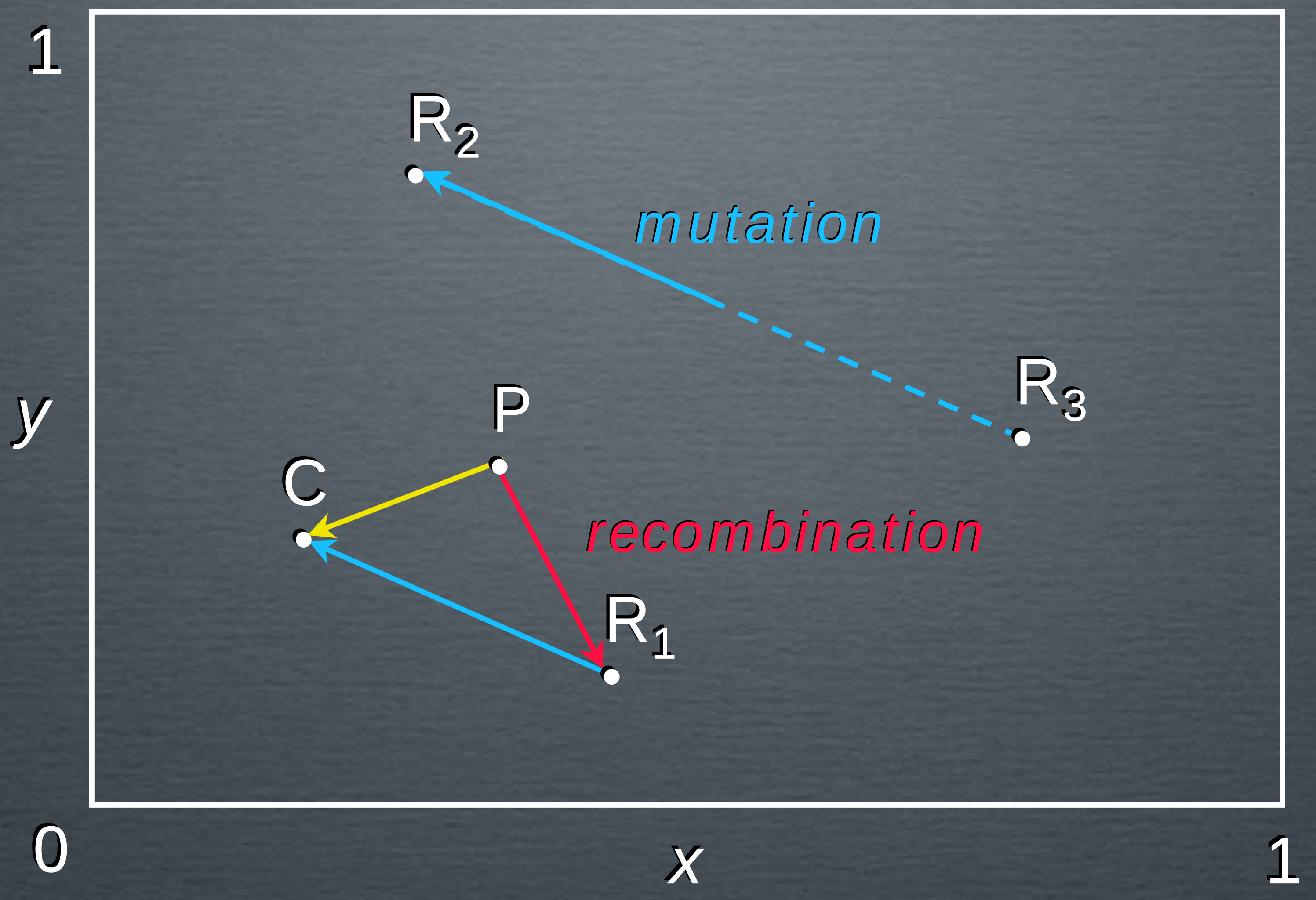
- Mating and mutation in one step:

$$\text{Trial} = \text{Parent} + K(\text{Random}_1 - \text{Parent}) + F(\text{Random}_2 - \text{Random}_3)$$

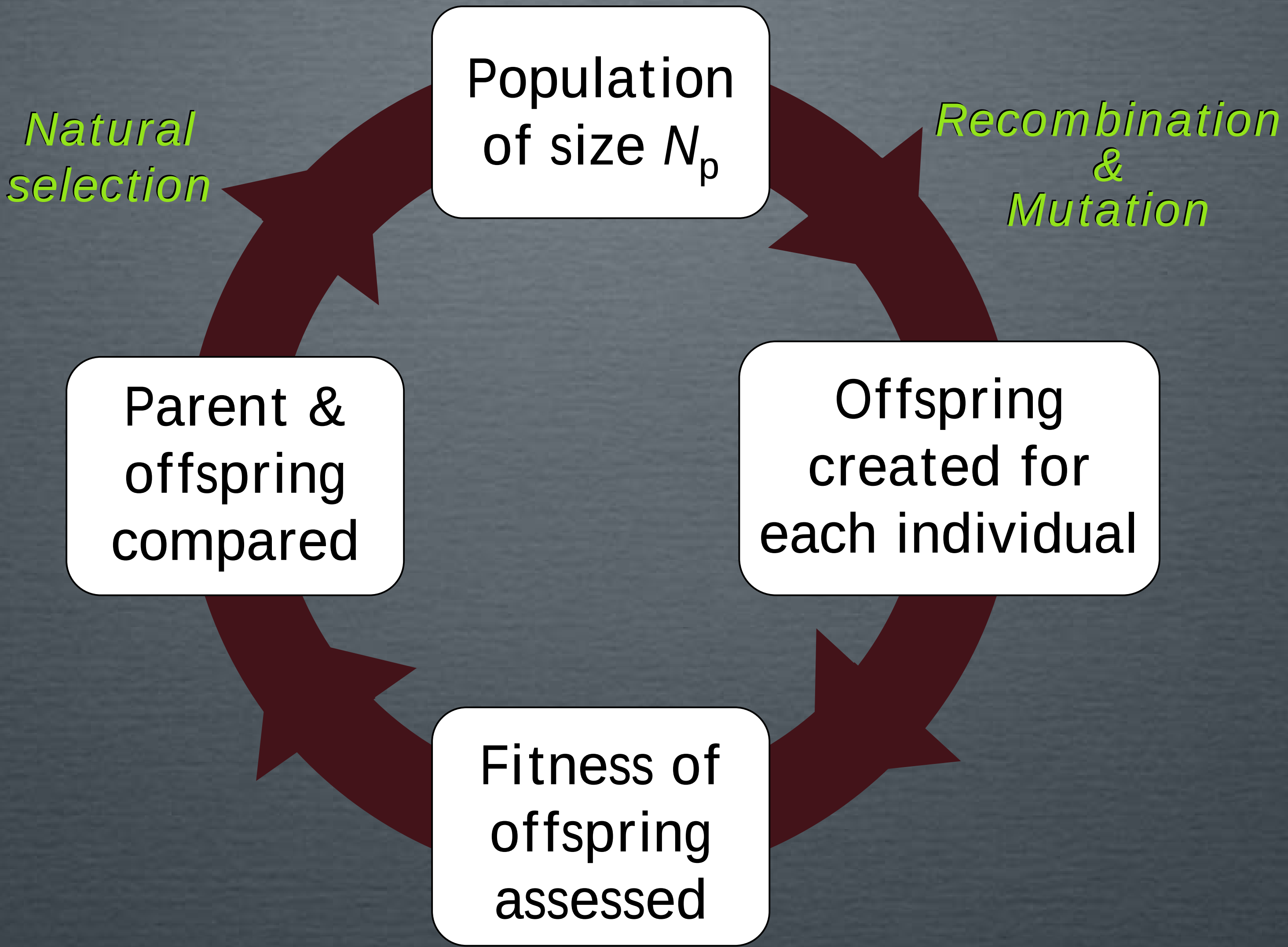
- 3 control parameters: N_p, K and F
- No mutants needed
- Best of child/parent added to population; deterministic selection – fast convergence
- New members used within the generation

Price, *New Ideas in Optimization*, McGraw-Hill, London, UK, 77, (1999)

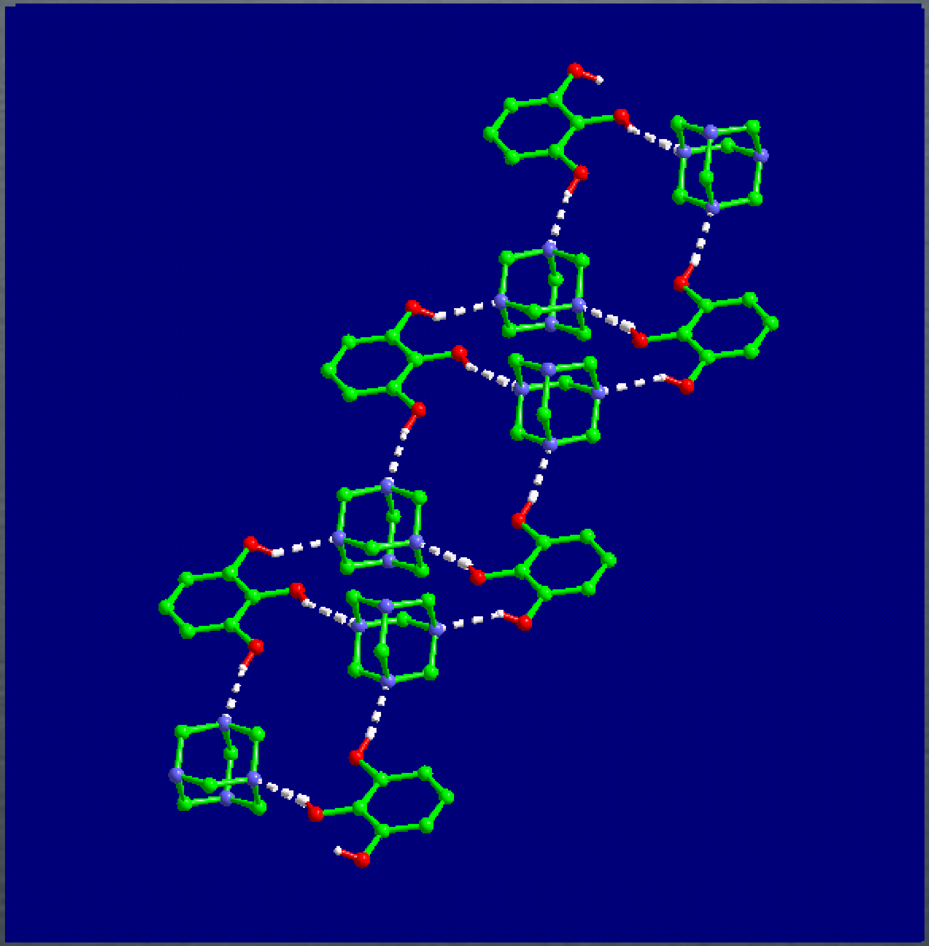
$$\text{Trial} = \text{Parent} + K(\text{Random}_1 - \text{Parent}) + F(\text{Random}_2 - \text{Random}_3)$$



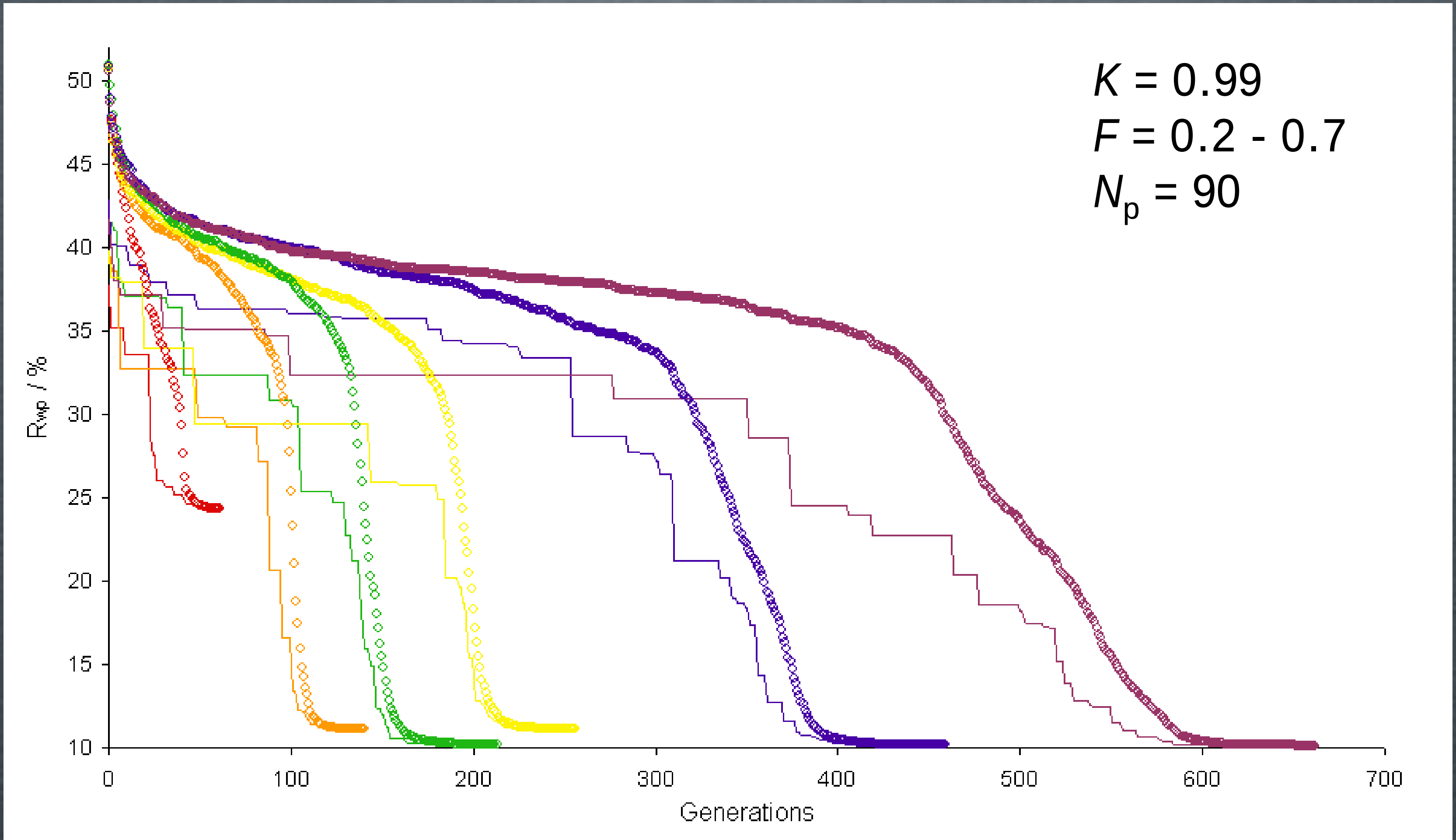
DE Calculation Cycle



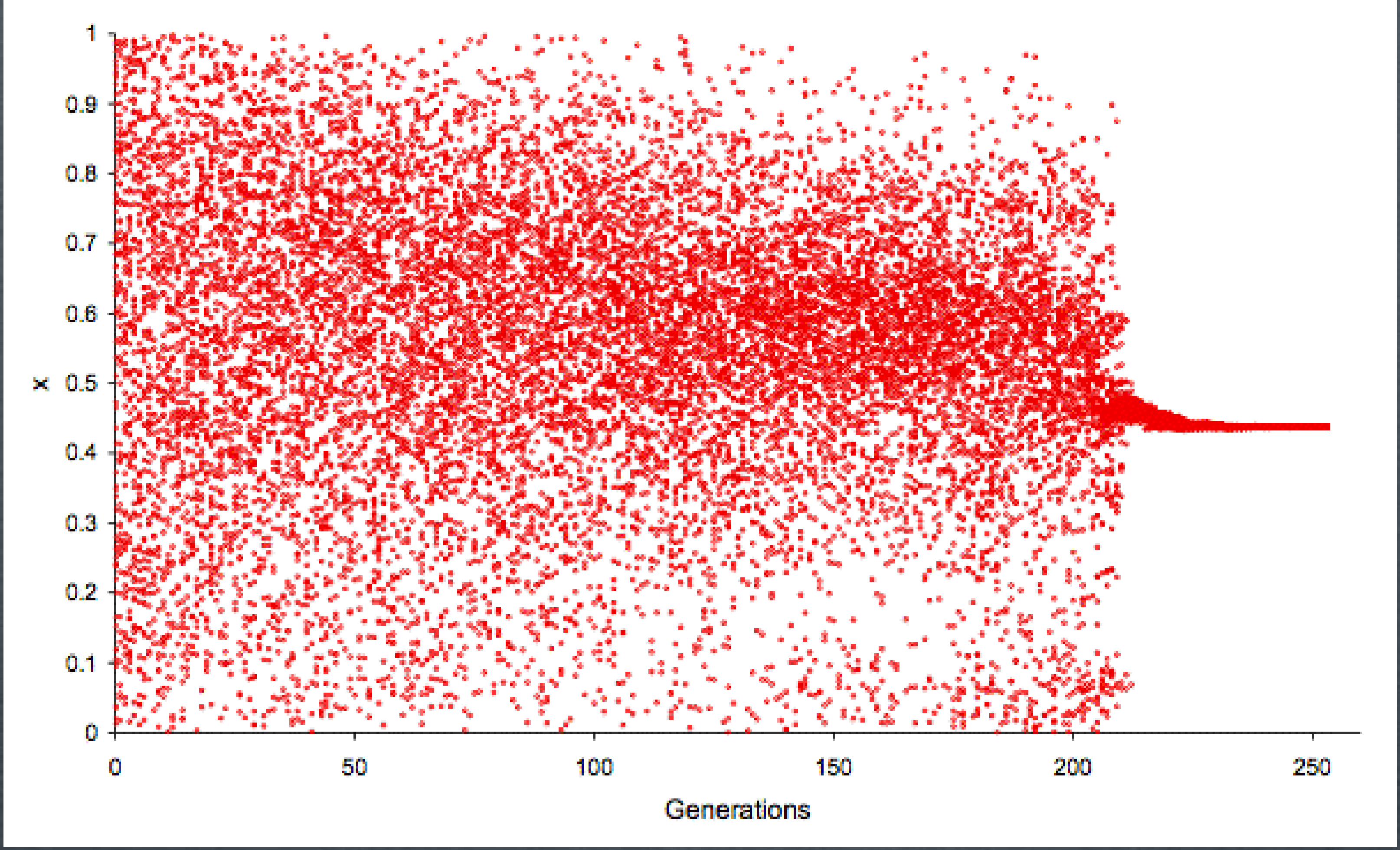
Structures Solved Using DE



Effect of Mutation Rate *F*



Evolution of Offspring Parameters



Differential Evolution

- Models biological evolution based on genetic inheritance
- Genetic evolution is a slow process
- Individuals behave independently
- *Population space*

Cultural Evolution

- Based on human social evolution
- Society adapts to change at a faster rate
- Models behavioural traits of all individuals
- *Belief space*

Reynolds, *New Ideas in Optimization*, McGraw-Hill, London, UK, 367, (1999)

Cultural Differential Evolution

- A ‘dual inheritance guidance’ system
- Uses information gained by previous generations to guide optimisation and influence adaption of population
- *Belief space* used to ‘prune’ *population space*

⬇

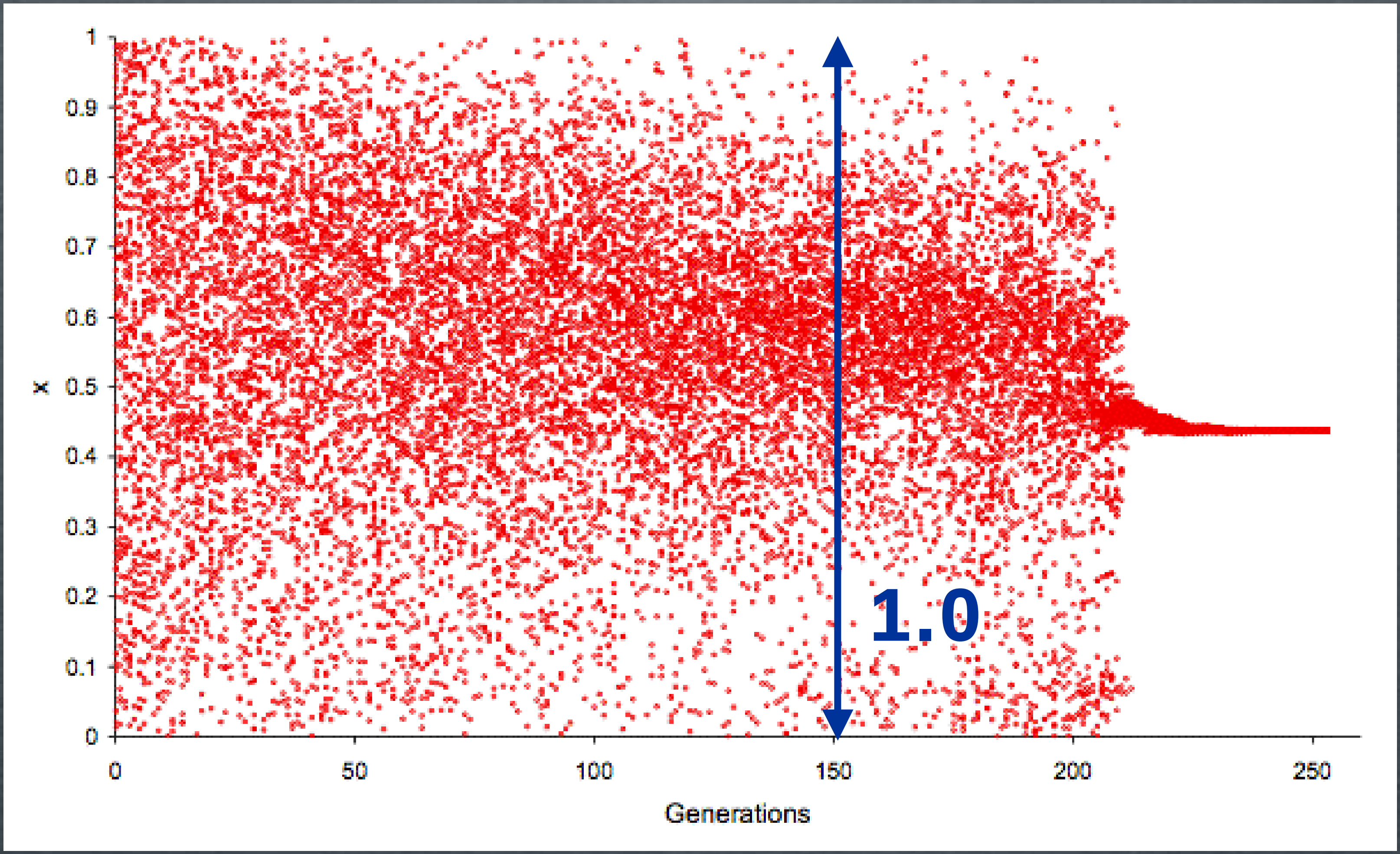
Implemented using ‘**dynamic**’ DE boundaries

DE Boundary Conditions

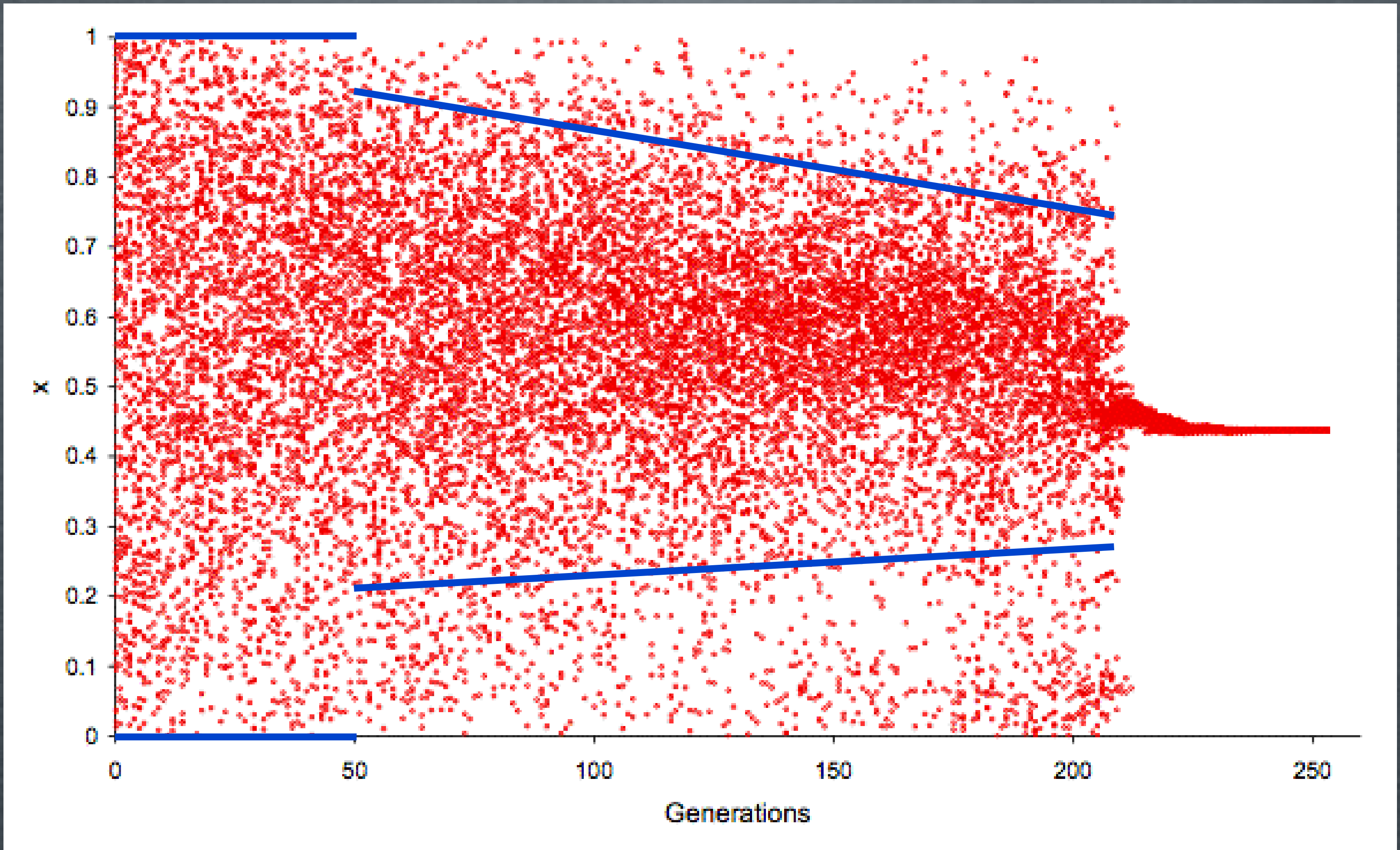
- Each parameter in genetic code ($x,y,z,\theta,\phi,\gamma,\tau_1......\tau_n$) has associated high & low bounds
- Child reset at boundary/parent median
- Implementation of **dynamic** boundaries, updated with each generation

→ Cultural Differential Evolution (CDE)

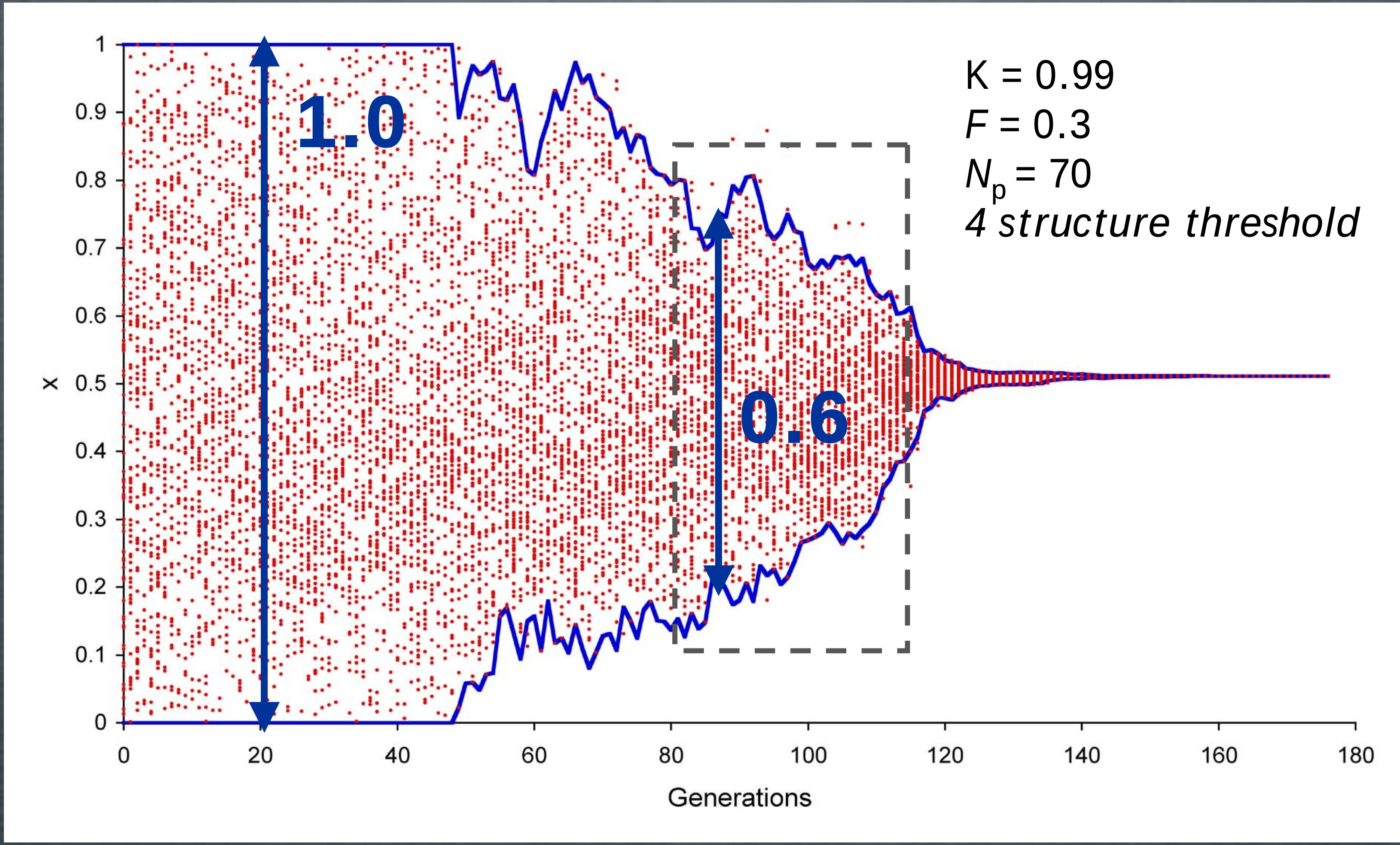
Use of Static Boundaries



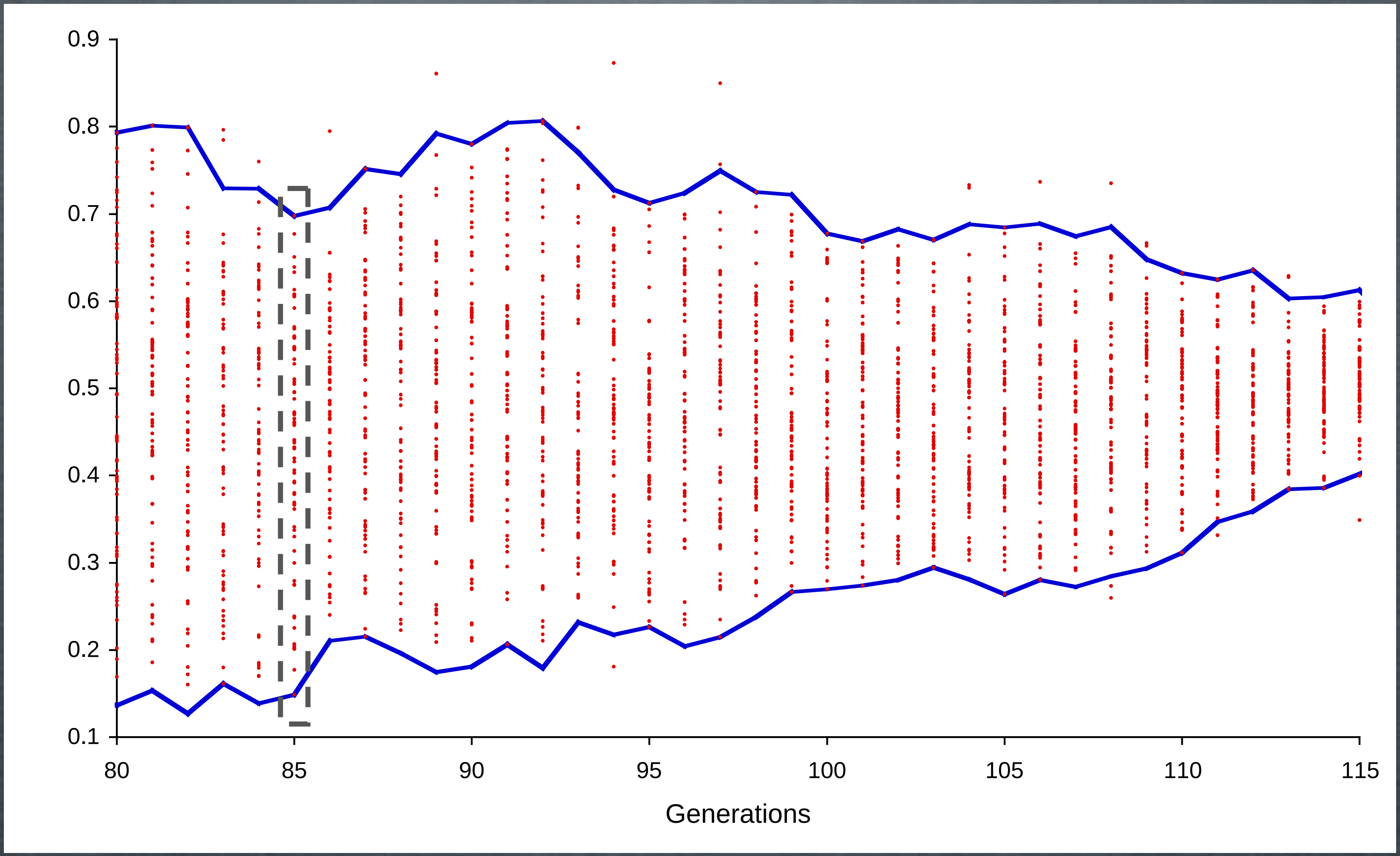
Use of Static Boundaries



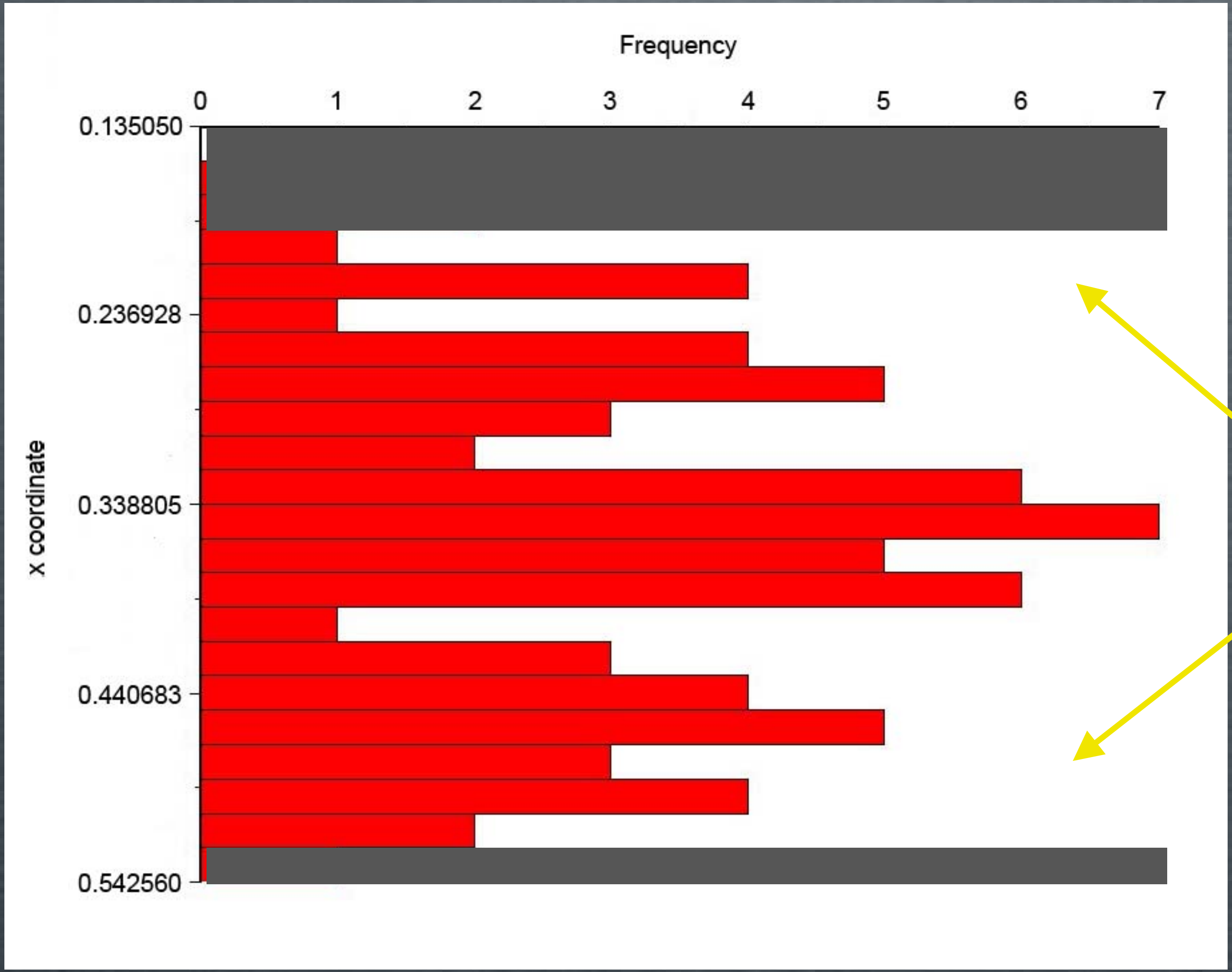
Effect of Dynamic Boundaries



Effect of Dynamic Boundaries

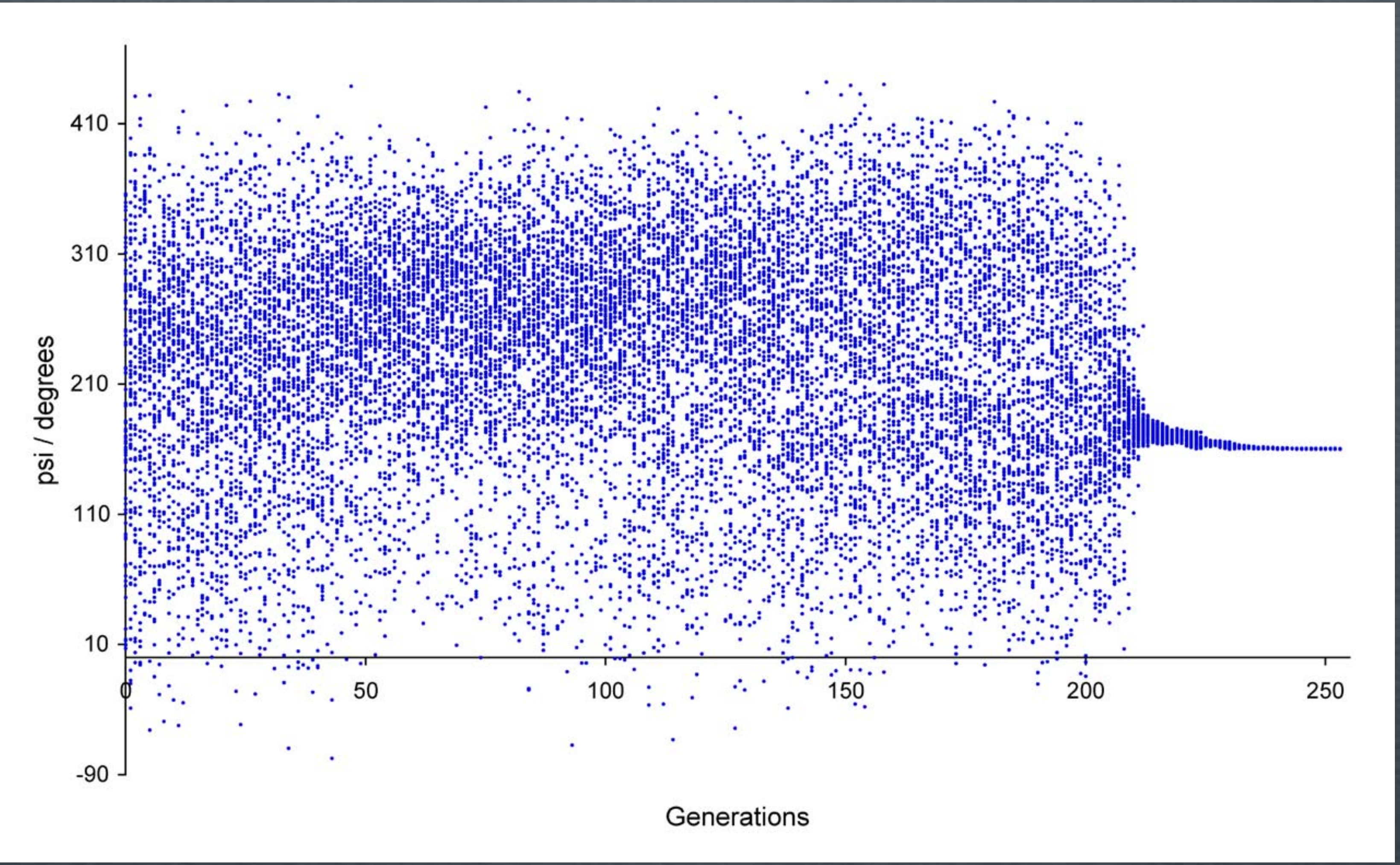


Implementation of Dynamic Boundaries

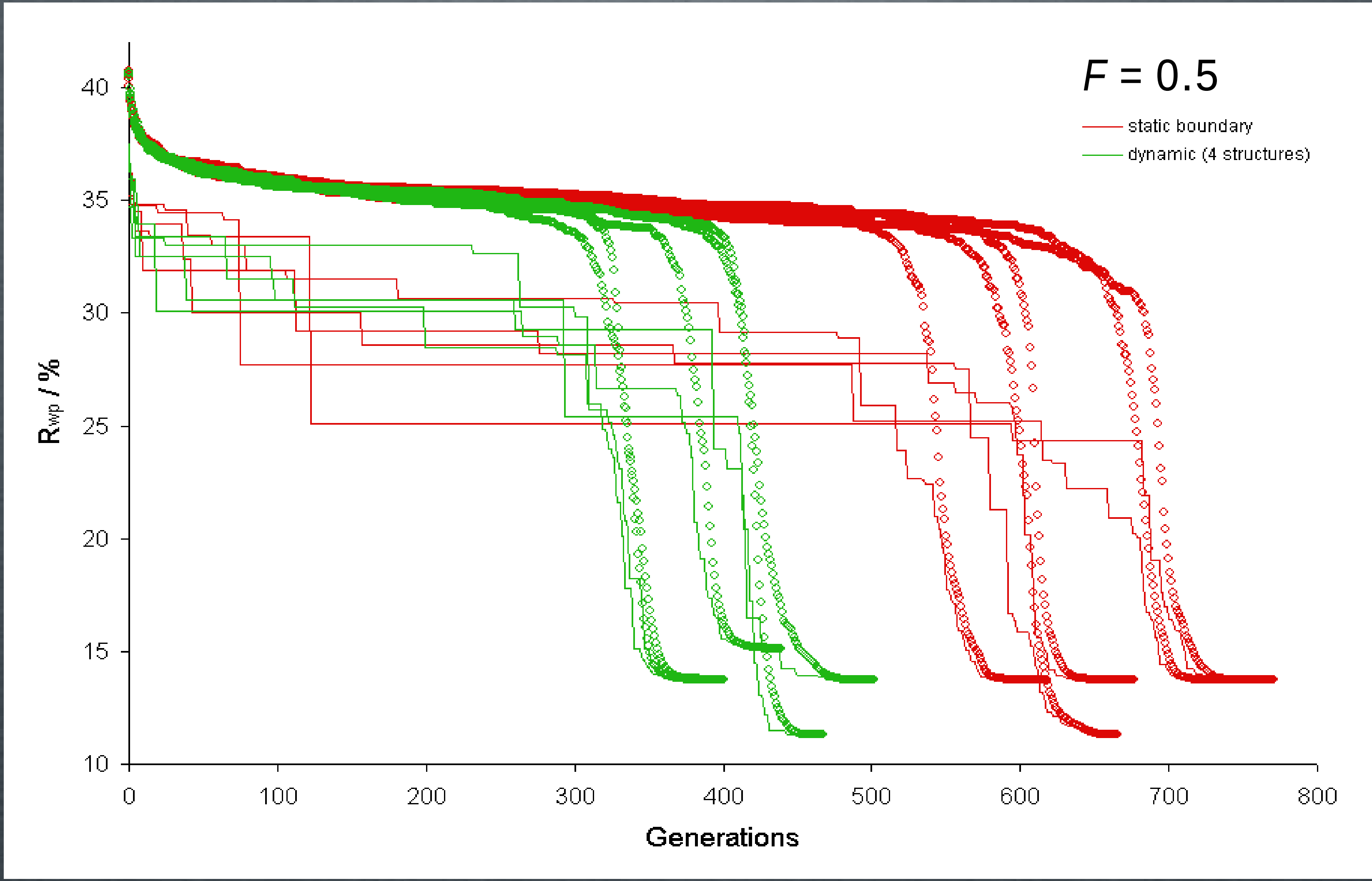


4 structure threshold (+1 bin)

Importance of ‘fully’ dynamic boundaries - allowing *expansion* and *contraction*



Effect on Convergence



Rate of Convergence

$N_p = 70$

Number of structures in threshold

	static	1	2	3	4	5	8
F = 0.3	207	171	210		203	228	
F = 0.4	408	371	373	392	288	284	190
F = 0.5	695	690	656	571	440	422	200
F = 0.6	762	1028	975	845	616	486	

successful solutions

0
1
2
3
4
5

Convergence efficiency (%)

$N_p = 70$

Number of structures in threshold

	static	1	2	3	4	5	8
F = 0.3	207	17	-1		2	-10	
F = 0.4	408	9	9	4	19	30	53
F = 0.5	695	1	6	18	37	39	71
F = 0.6	762	-35	-28	-11	19	36	

successful solutions

red	0
orange	1
yellow	2
green	3
teal	4
blue	5

105

	static	3	4	5	6	7	8
0.3	211	229	211	204	286	197	221
0.4	488	456	472	404	442	280	297
0.5	946	997	781	563	575	496	404
0.6	1365	1211	1295	982	756	629	554

140

	static	3	4	5	8	9	10	11
0.3	355	305	332	268	228	205	224	262
0.4	642	634	587	551	377	357	383	307
0.5	1125	1096	1264	1194	678	525	527	438
0.6	1403	1489	1124		1105	784	654	540

105

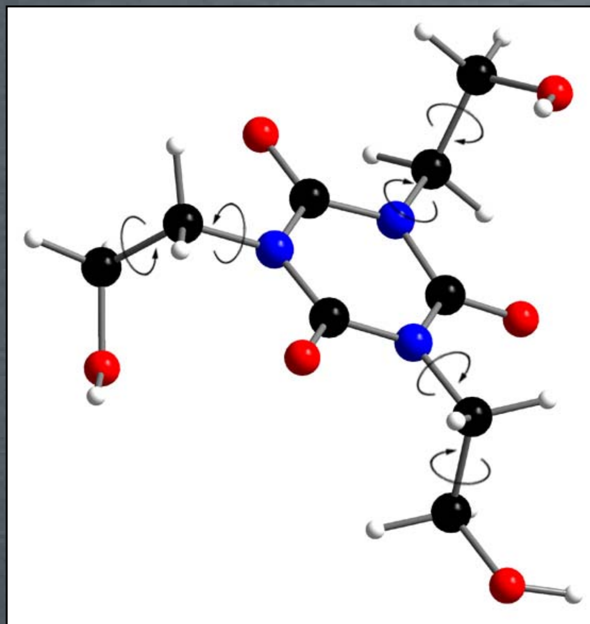
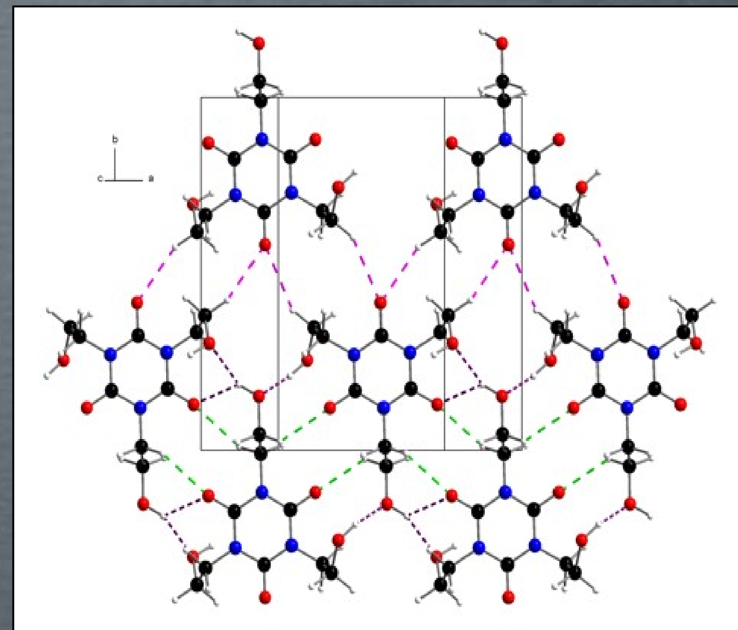
	static	3	4	5	6	7	8
0.3	211	-9	0	3	-36	7	-5
0.4	488	7	3	17	9	43	39
0.5	946	-5	17	40	39	48	57
0.6	1365	11	5	28	45	54	59

(%)

140

	static	3	4	5	8	9	10	11
0.3	355	14	6	25	36	42	37	26
0.4	642	1	9	14	41	44	40	52
0.5	1125	3	-12	-6	40	53	53	61
0.6	1403	-6	20		21	44	53	62

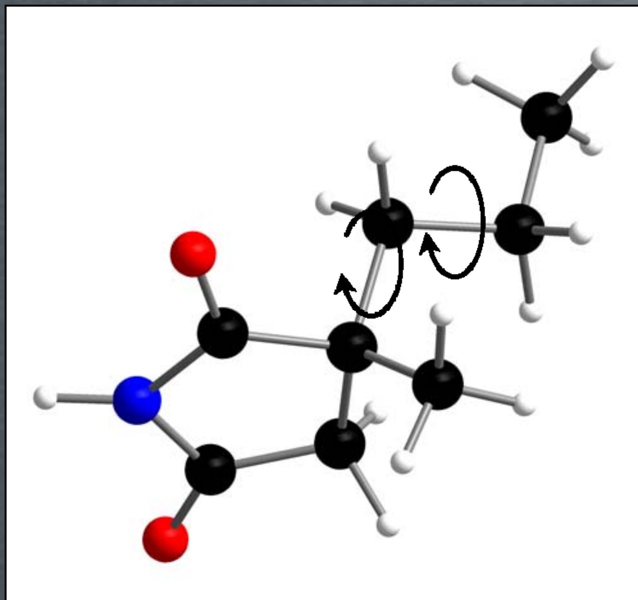
1,3,5-tris-(2-hydroxyethyl) cyanuric acid

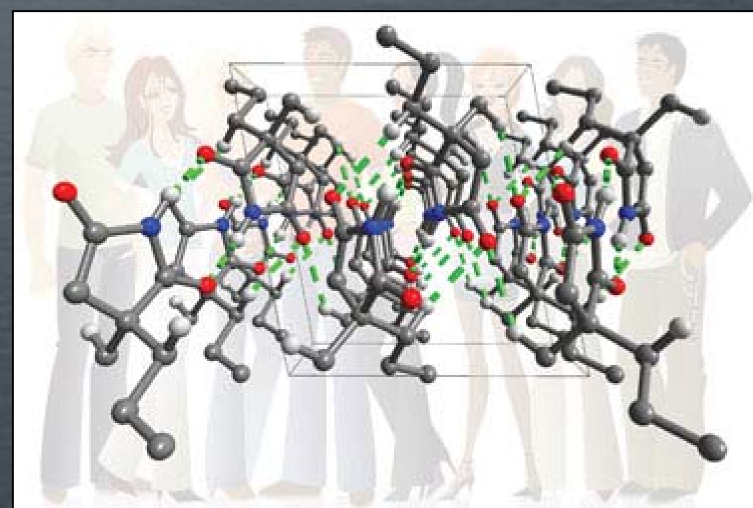
Chong, Seaton, Kariuki & Tremayne, *Acta Cryst* B58, 823, (2006)

- CDE parameters: K=0.99, F=0.5, N_p=120
  4 structure threshold
- Convergence: DE → 860, CDE → 653

α-Methyl-α-Propyl Succinimide



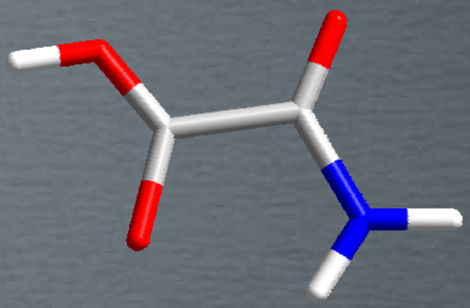
- CDE model:
  $(x,y,z,\theta,\phi,\gamma,\tau_1,\tau_2)$
- CDE parameters:
  $K=0.99$, $F=0.4$, $N_p=80$
  4 structure threshold
  CDE 461, DE 988



- CDE parameters:
  $K=0.99$, $F=0.3$, $N_p=80$
  2 structure threshold
  CDE 163, DE failed

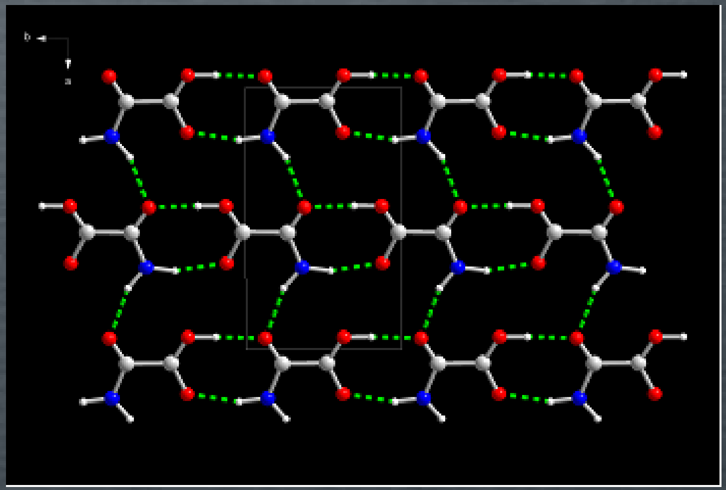
Pitfalls….SDPD….Data Limitations?

Oxamic Acid



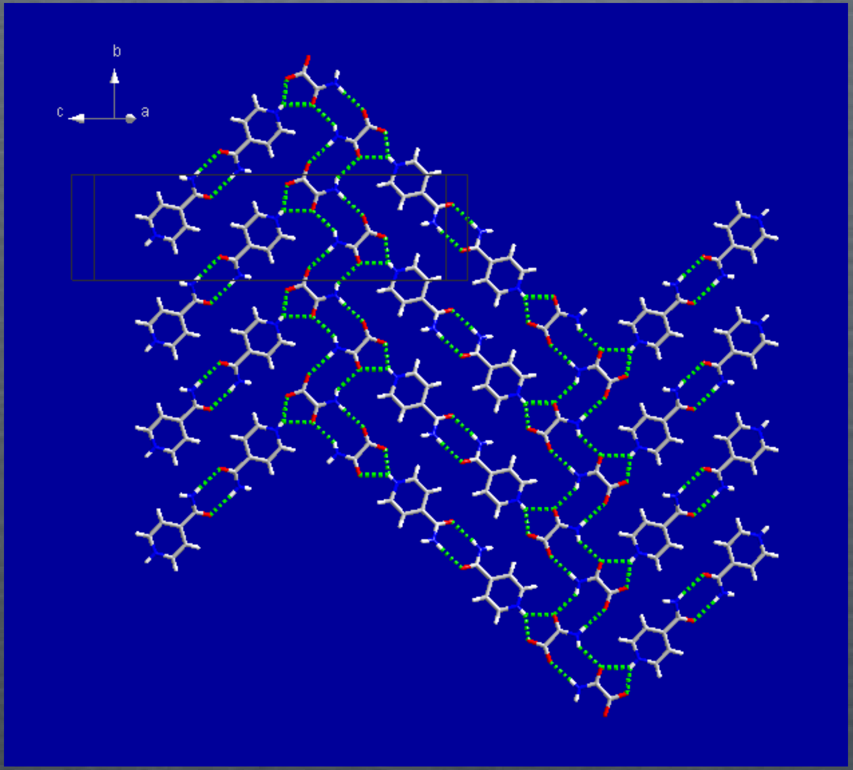
- Structure determined from lab PXRD (DE)
- Synchrotron PXRD data to distinguish C=O

- Deuteration gives new polymorph!
- Structure now confirmed



Pitfalls.....SDPD.....Data Limitations?

Oxamic Acid with ... Isonicotinamide



(1/1) salt adduct → Single-crystal structure

Pitfalls.....SDPD.....Data Limitations?

Oxamic Acid with ... Nicotinamide

- Poor quality crystals
- Structure determination from lab PXRD (CDE)

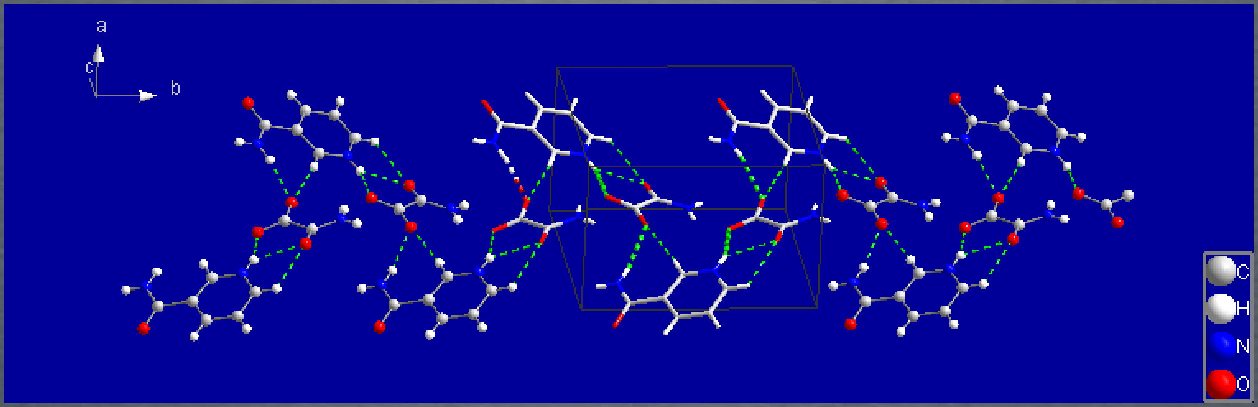
⬇

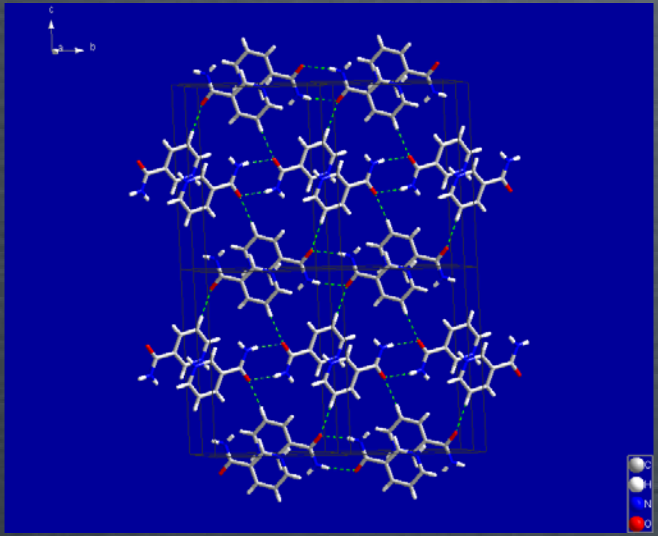
- Solid state IR – salt or cocrystal
- GC & EA – stoichiometry (NMR not conclusive)

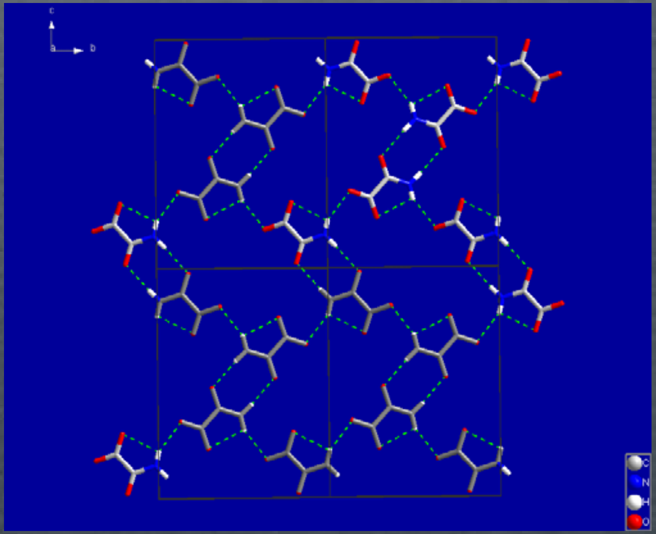
(1/1) salt adduct: distinct layers of components
but no oxamic acid H bond network

Pitfalls….SDPD….Data Limitations?

- Better crystal growth → Single-crystal structure



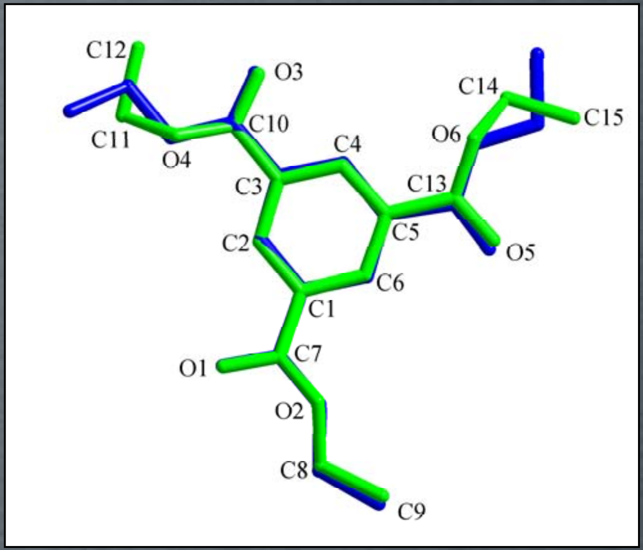




Oxamic acid *not* resolved in SDPD structure

Pitfalls.....Preferred Orientation?

Triethyl-1,3,5-benzene tricarboxylate

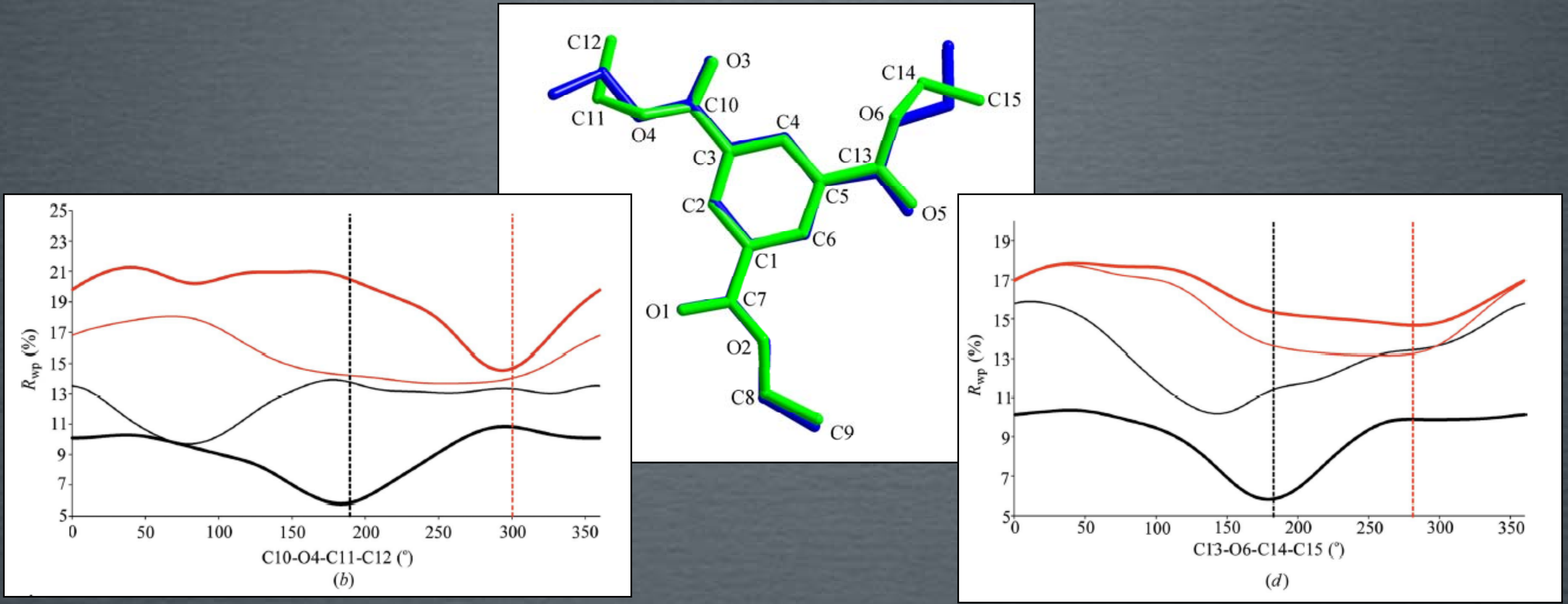


- Structure determined from lab PXRD (DE)
- but ... conformational change with refinement

Is structure solution correct?

Chong, Seaton, Kariuki & Tremayne, *Acta Cryst* B62,864 (2006)

Pitfalls….Preferred Orientation?



- Structure solution minimum differs from refinement minimum
- Effect of po parameter (& model relaxation)

Chong, Seaton, Kariuki & Tremayne, *Acta Cryst* B62,864 (2006)

Conclusions

- *Dynamic* boundaries improve search efficiency by up to ~50%.
- 'Dual inheritance' system applicable to evolutionary algorithms used in other fields?
- Optimisation of CDE related to % structure threshold?

Chong & Tremayne, *Chem Comm* 4078, (2006)

Chemistry World, 3(11), 27, 2006; Chemical Science, 11, 2006
Nature: Research Highlights 443, 375, 2006; www.SpectroscopyNOW.com
Materials Research Society Bulletin, 31, 967, 2006

Acknowledgements

Dr Colin Seaton

Dr Samantha Chong

Dr Harry Chana

Adam Cowell

Tuba Mahmud

Liana Vella-Zarb

Duncan Bell

Dr Benson Kariuki

Dr Elizabeth Shotton

Dr Robert Lancaster

The Royal Society

EPSRC, ICDD

GSK, Pfizer, MSD

UNIVERSITY OF BIRMINGHAM

cposs

THE ROYAL SOCIETY