OPTIMIZATION WITHOUT ALGEBRAIC MODELS:
Algorithms, software, and applications

Nick Sahinidis
Center for Computer-Aided Process Decision-making
Department of Chemical Engineering
Carnegie Mellon University
sahinidis@cmu.edu

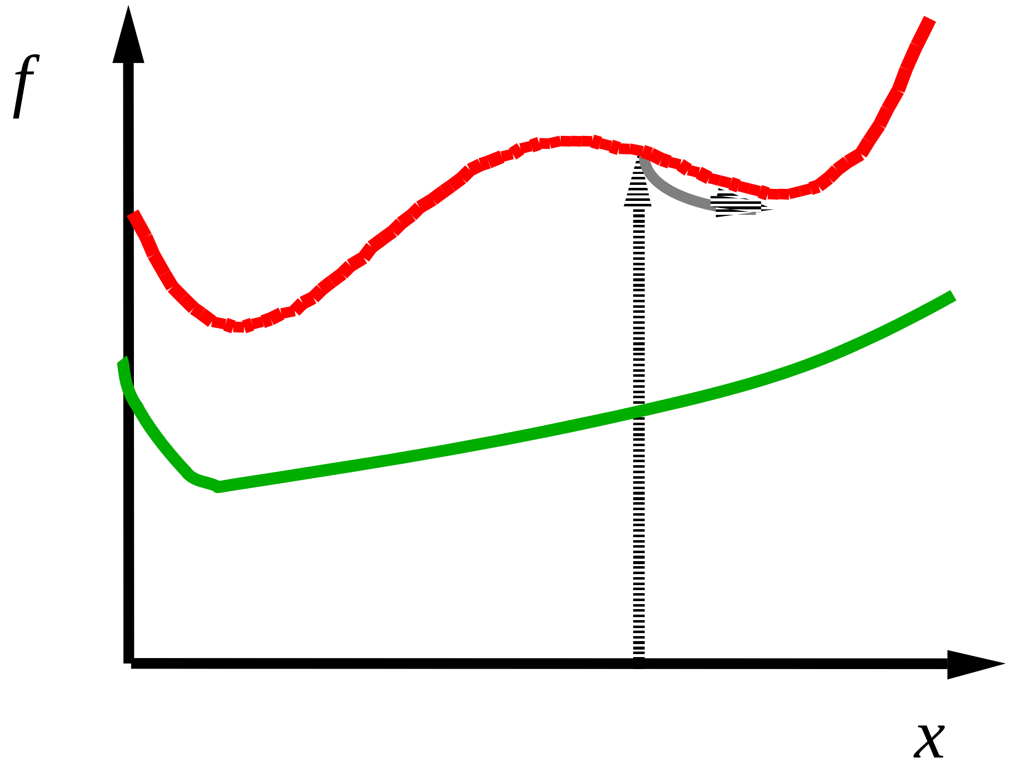
Acknowledgments:
Luis Miguel Rios
NIH, NSF and DOE/NETL

HISTORICAL DEVELOPMENTS IN OPTIMIZATION

- **300 B.C: Shortest distance from a point to a line (Euclid)**
- **1600s: Leibniz/Newton (Calculus)**
- **1847: Gradient methods (Cauchy)**
- **1875: Minimum free energy principle (Gibbs)**
- **Late 40's: Linear Optimization**
  - **Army operations; Linear objective function and constraints**
- **Late 50's: Nonlinear Optimization**
  - **Chemical process industries; Nonlinear functions**
- **60's: Integer Optimization**
  - **Discrete manufacturing; Integer variables to model discrete decisions and economies of scale**

GLOBAL OPTIMIZATION OF MIXED-INTEGER NONLINEAR OPTIMIZATION PROBLEMS

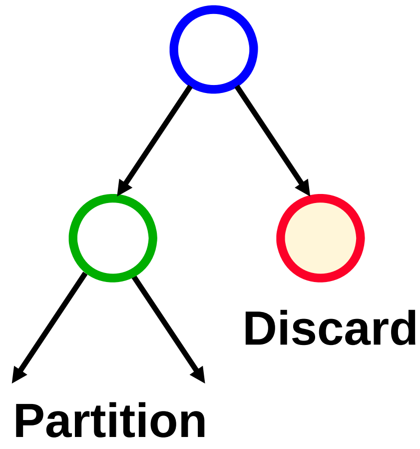
CONVEXIFICATION

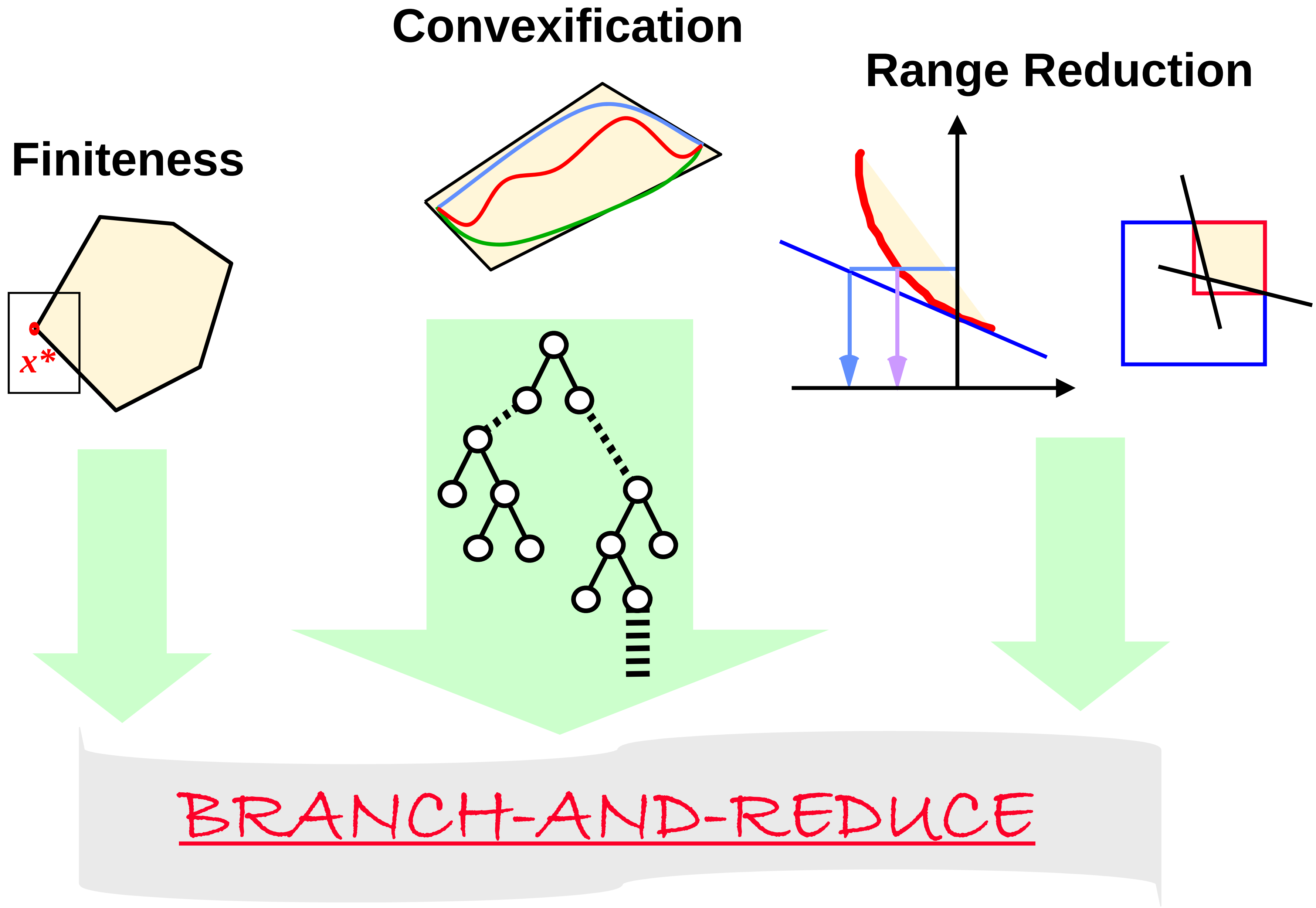


Classical optimization algorithms provide a local minimum “closest” to the starting point used

BRANCH-AND-BOUND

Search Tree





Convexification
Range Reduction
Finiteness
x*
BRANCH-AND-REDUCE

BARON SOFTWARE

- **First commercial software to offer deterministic guarantee of global optimality for multi-extremal nonlinear optimization problems**
- **Two-pronged approach to technology transfer**
  - **Commercial**
    - » **Under the modeling languages GAMS and AIMMS**
  - **Free**
    - » **Under the NEOS server for optimization**

BARON IN APPLICATIONS

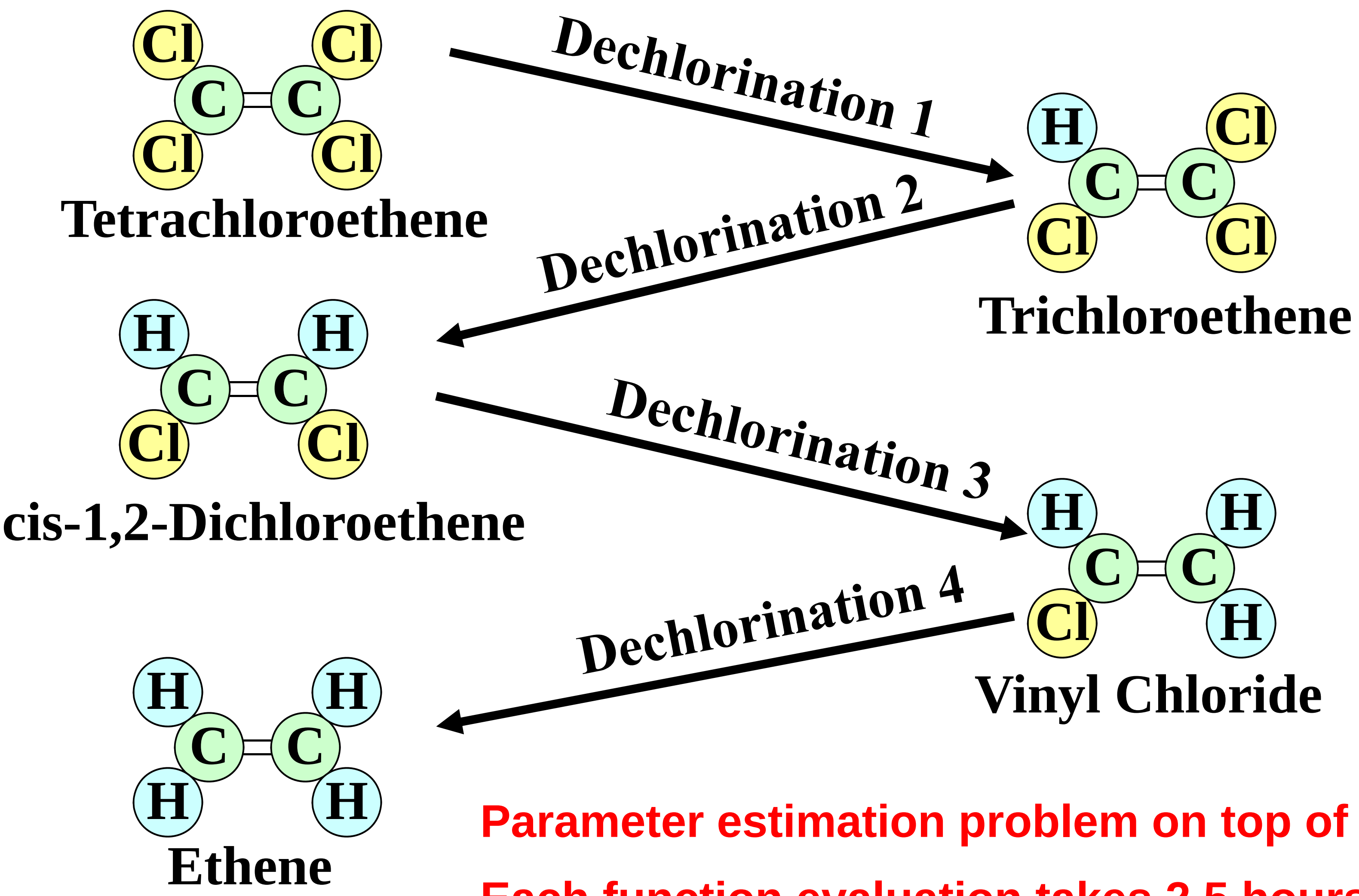
- **Development of new Runge-Kutta methods for partial differential equations**
  - **Ruuth and Spiteri, *SIAM J. Numerical Analysis,* 2004**
- **Energy policy making**
  - **Manne and Barreto, *Energy Economics,* 2004**
- **Model estimation and automatic control**
  - **Bemporand and Ljung, *Automatica,* 2004**
- **Agricultural economics**
  - **Cabrini *et al.*, *Manufacturing and Service Operations Management,* 2005**
- **Portfolio optimization for wealth-dependent risk preferences**
  - **Rios and Sahinidis, *Annals of Operations Research,* 2010**

THE ALGEBRAIC OPTIMIZATION PARADIGM

- **Algebraic models**
  - Require optimization expertise
  - Take a long time to develop
  - Often require restrictive assumptions to increase solvability
- **Practitioners do not have models when**
  - Proprietary software required for simulation
  - Optimization is required in an experimental setting

MODEL CALIBRATION

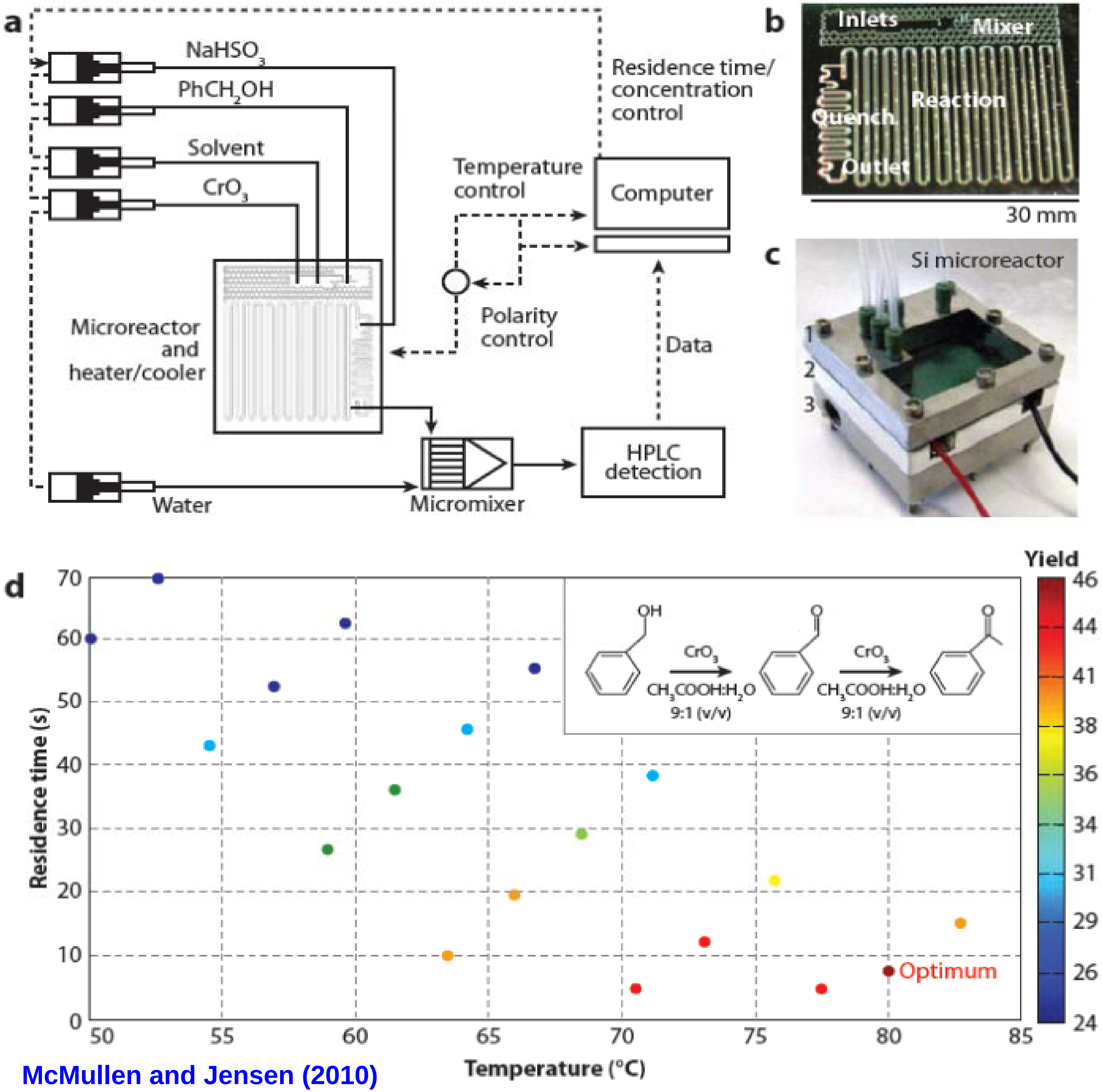
(Mugunthan and Shoemaker, 2005)



Parameter estimation problem on top of PDEs

Each function evaluation takes 2.5 hours

MICROREACTOR OPTIMIZATION

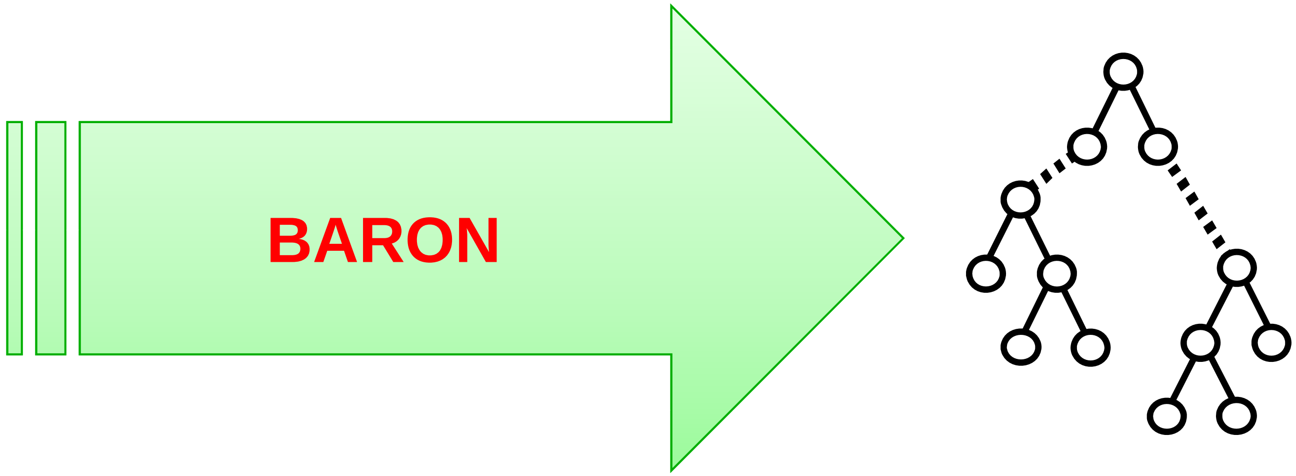


McMullen and Jensen (2010)

AUTOMATIC TUNING OF OPTIMIZATION SOFTWARE

Options

- **Relaxation**
- **Reduction**
- **Branching**



Optimize solver performance over

- **A collection of test problems**
- **During run-time**

DERIVATIVE-FREE OPTIMIZATION

- **Optimization of a function for which**
  - **derivative information is not symbolically available**
  - **derivative information is not numerically computable**
- **Studied in a variety of areas under the terms:**
  - **Black-box optimization**
  - **Simulation-based optimization**
  - **Design of experiments**
  - **Response surface methods**
  - **Active learning**

Optimization without an algebraic model

TIMELINE OF INNOVATION IN DERIVATIVE-FREE OPTIMIZATION

1960

- 1961 Hooke and Jeeves algorithm is proposed
- 1962 First simplex-based optimization algorithm
- 1965 Nelder-Mead simplex algorithm is proposed
- 1969 First use of trust-region quadratic-based models

1970

- 1973 First published monograph
- 1975 Genetic algorithms are proposed
- 1979 Hit-and-run algorithms are proposed

1980

- 1983 First use of simulated annealing in optimization
- 1989 DACE stochastic model is proposed

1990

- 1991 Convergence of multi-directional search algorithm is shown
- 1993 Ideas from Lipschitzian optimization introduced
- 1994 Geometry considerations for points in trust-region methods
- 1995 Implicit filtering and particle swarm algorithms proposed
- 1997 DACE surrogate model introduced
- 1998 First use of radial basis functions in surrogate models
- 1999 Introduction of multilevel coordinate search

2000

- 2002 First use of augmented Lagrangian in pattern search methods
- 2003 Generating set nomenclature introduced
- 2004 Incorporation of filters and simplex derivatives in pattern search
- 2009 First textbook

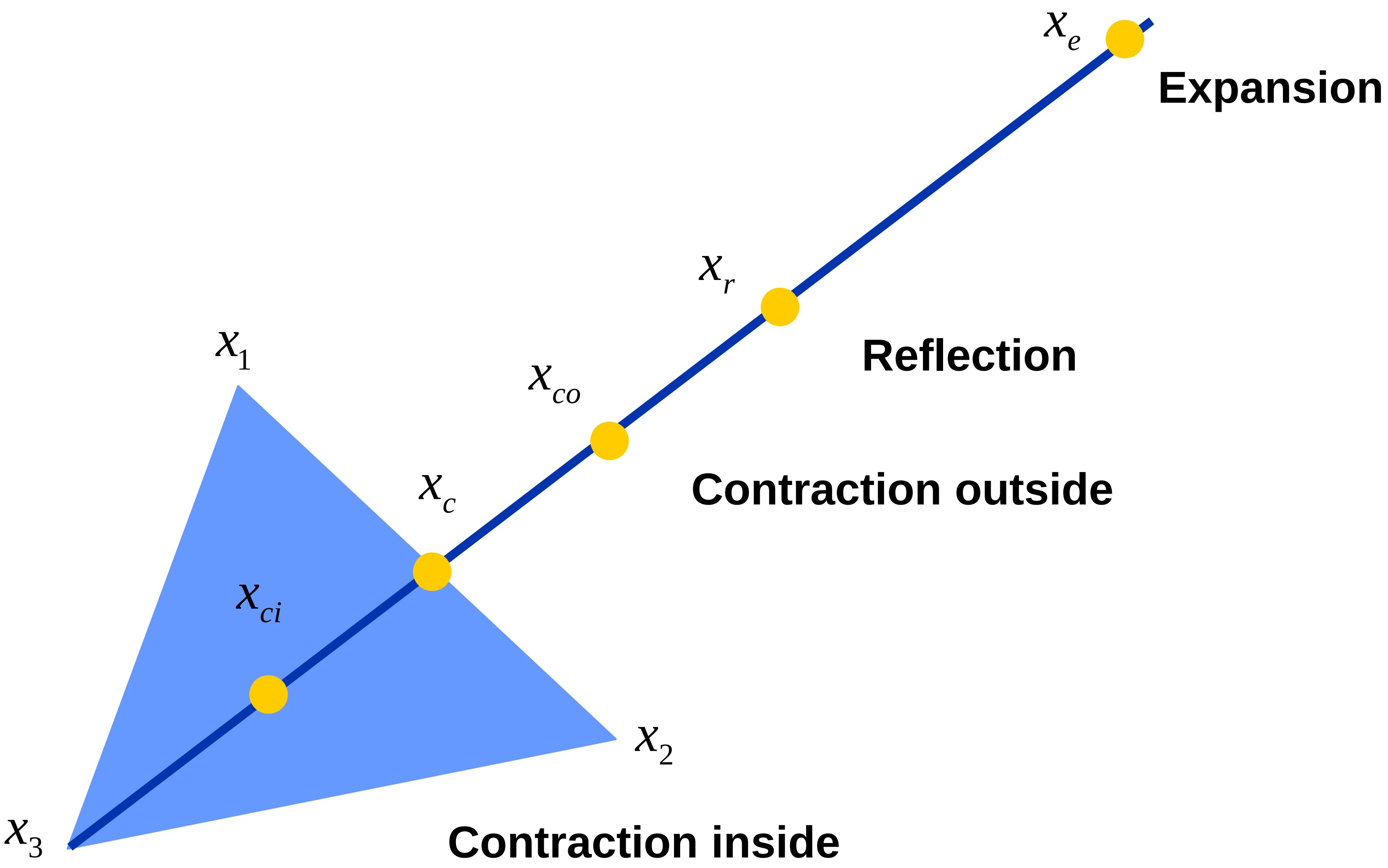
MOST CITED WORKS

Publication	Year appeared	Citations[1]
Hooke and Jeeves [64]	1961	2281
Nelder and Mead [99]	1965	13486
Brent [25]	1973	2019
Holland [58]	1975	31494
Kirkpatrick et al. [78]	1983	23053
Eberhart and Kennedy [44, 77]	1995	20369

1. From Google Scholar on 20 December 2011.

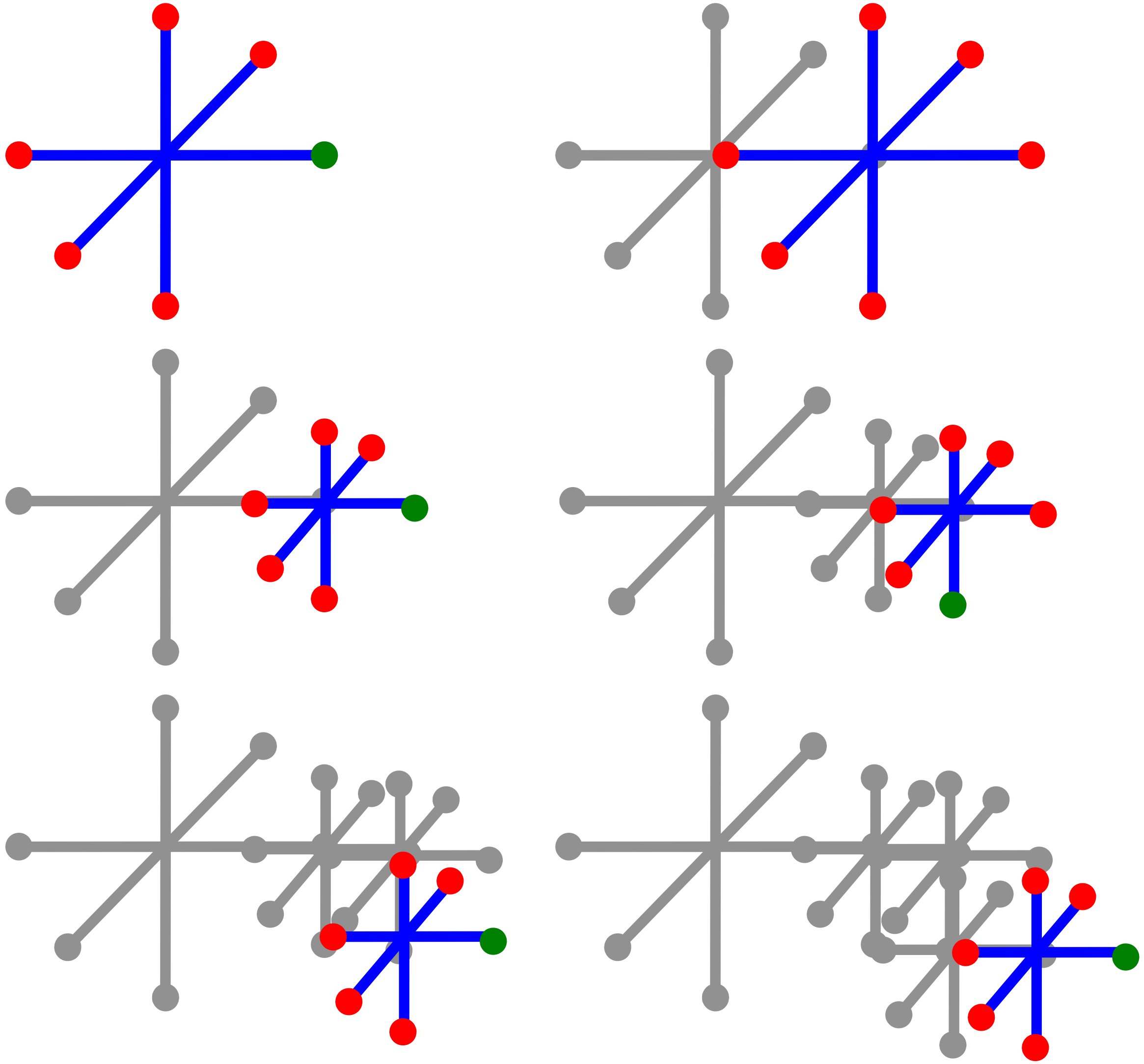
NELDER-MEAD ALGORITHM

(Nelder and Mead, 1965)



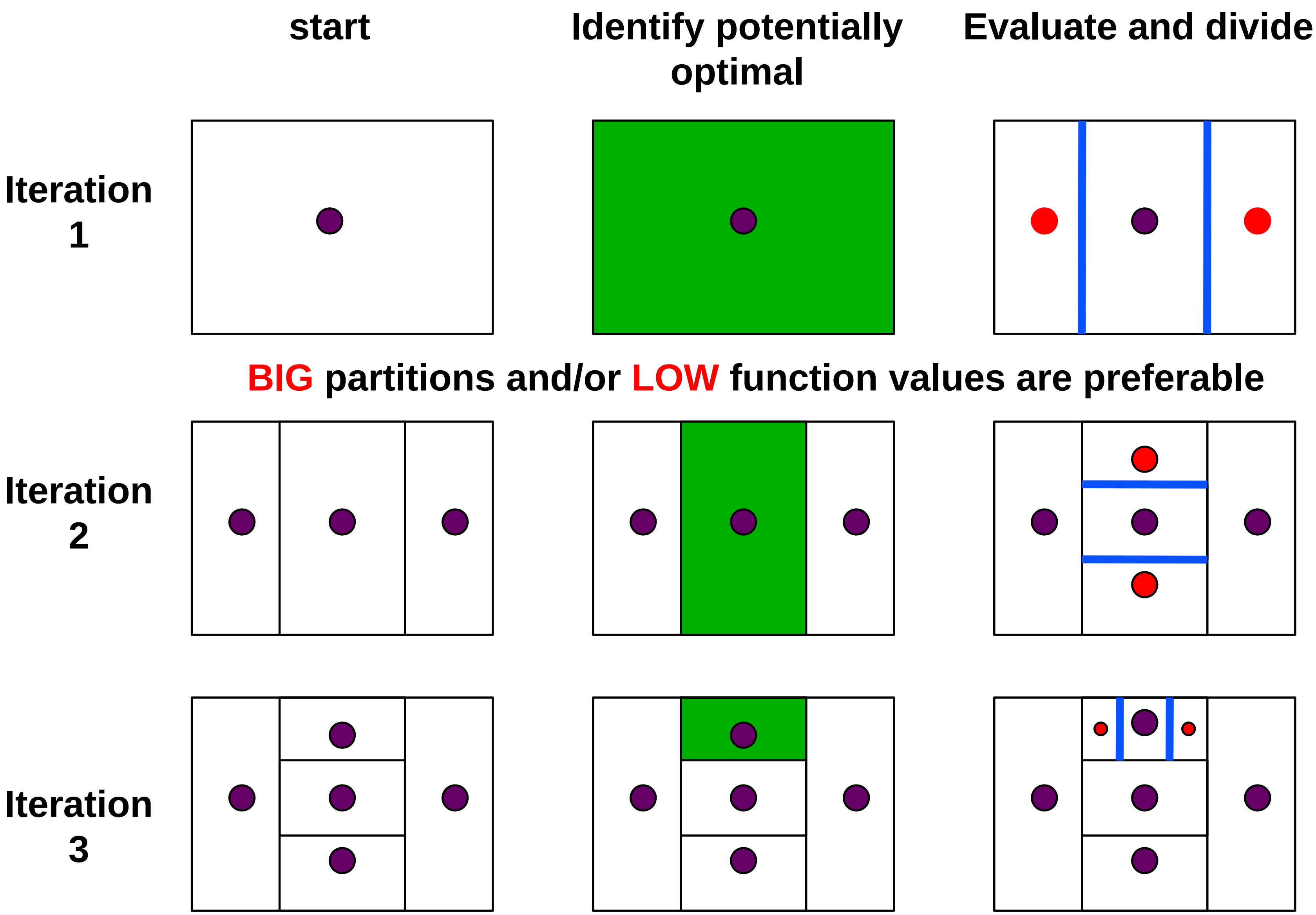
PATTERN SEARCH ALGORITHMS

(Hooke and Jeeves, 1961; Torczon, 1997)



DIRECT ALGORITHM

(Jones et al., 1993)



DERIVATIVE-FREE OPTIMIZATION ALGORITHMS

- **LOCAL SEARCH METHODS**
  - **Direct local search**
    - » **Nelder-Mead simplex algorithm**
    - » **Generalized pattern search and generating search set**
  - **Based on surrogate models**
    - » **Trust-region methods**
    - » **Implicit filtering**
- **GLOBAL SEARCH METHODS**
  - **Deterministic global search**
    - » **Lipschitzian-based partitioning**
    - » **Multilevel coordinate search**
  - **Stochastic global optimization**
    - » **Hit-and-run**
    - » **Simulated annealing**
    - » **Genetic algorithms**
    - » **Particle swarm**
  - **Based on surrogate models**
    - » **Response surface methods**
    - » **Surrogate management framework**
    - » **Branch-and-fit**

DERIVATIVE-FREE OPTIMIZATION SOFTWARE

LOCAL SEARCH

- FMINSEARCH (Nelder-Mead)
- DAKOTA PATTERN (PPS)
- HOPSPACK (PPS)
- SID-PSM (Simplex gradient PPS)
- NOMAD (MADS)
- DFO (Trust region, quadratic model)
- IMFIL (Implicit Filtering)
- BOBYQA (Trust region, quadratic model)
- NEWUOA (Trust region, quadratic model)

GLOBAL SEARCH

- DAKOTA SOLIS-WETS (Direct)
- DAKOTA DIRECT (DIRECT)
- TOMLAB GLBSOLVE (DIRECT)
- TOMLAB GLCSOLVE (DIRECT)
- MCS (Multilevel coordinate search)
- TOMLAB EGO (RSM using Kriging)
- TOMLAB RBF (RSM using RBF)
- SNOBFIT (Branch and Fit)
- TOMLAB LGO (LGO algorithm)

STOCHASTIC

- ASA (Simulated annealing)
- CMA-ES (Evolutionary algorithm)
- DAKOTA EA (Evolutionary algorithm)
- GLOBAL (Clustering - Multistart)
- PSWARM (Particle swarm)

SOLVERS COMPARED

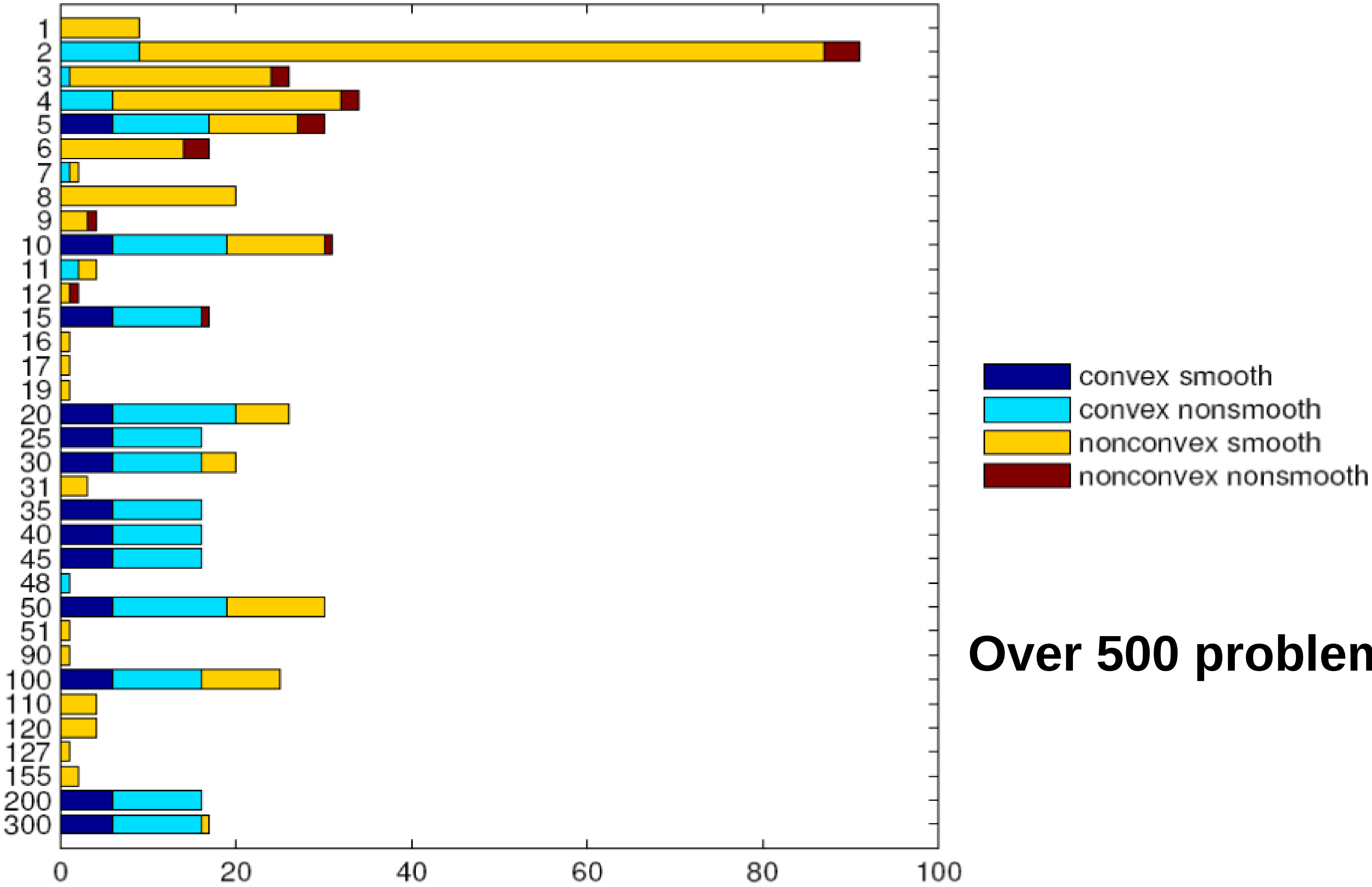
Solver	URL	Version	Language	Bounds	Constraints
ASA	www.ingber.com	26.30	C	required	no
BOBYQA	Available by email from mjdp@cam.ac.uk	2009	Fortran	required	no
CMA-ES	www.lri.fr/~hansen/cmaesintro.html	3.26beta	Matlab	optional	no
DAKOTA/DIRECT	www.cs.sandia.gov/dakota/	4.2	C++	required	yes
DAKOTA/EA	www.cs.sandia.gov/dakota/	4.2	C++	required	yes
DAKOTA/PATTERN	www.cs.sandia.gov/dakota/	4.2	C++	required	yes
DAKOTA/SOLIS-WETS	www.cs.sandia.gov/dakota/	4.2	C++	required	yes
DFO	projects.coin-or.org/Dfo	2.0	Fortran	required	yes
FMINSEARCH	www.mathworks.com	1.1.6.2	Matlab	no	no
GLOBAL	www.inf.u-szeged.hu/~csendes	1.0	Matlab	required	no
HOPSPACK	software.sandia.gov/trac/hopspack	2.0	C++	optional	yes
IMFIL	www4.ncsu.edu/~ctk/imfil.html	0.86	Matlab	required	yes
MCS	www.mat.univie.ac.at/~neum/software/mcs/	2.0	Matlab	required	no
NEWUOA	Available by email from mjdp@cam.ac.uk	2004	Fortran	no	no
NOMAD	www.gerad.ca/nomad/	3.3	C++	optional	yes
PSWARM	www.norg.uminho.pt/aivaz/pswarm/	1.3	Matlab	required	yes*
SID-PSM	www.mat.uc.pt/sid-psm/	1.1	Matlab	optional	yes*
SNOBFIT	www.mat.univie.ac.at/~neum/software/snobfit/	2.1	Matlab	required	no
TOMLAB/GLCCLUSTER	tomopt.com	7.3	Matlab	required	yes
TOMLAB/LGO	www.pinterconsulting.com/	7.3	Matlab	required	yes
TOMLAB/MULTIMIN	tomopt.com	7.3	Matlab	required	yes
TOMLAB/OQNLP	tomopt.com	7.3	Matlab	required	yes

* Handles linear constraints only.

APPROACH

- Started seven years ago
- Collected over 500 NLP benchmarks
  - Algebraic formulations; global solutions known (BARON)
- Developed unified interface to 25+ solvers
- Average-case comparisons
  - Based on median objective function value of 10 runs from randomly generated starting points
  - Solver solved problem if solution within 0.01 or 1% of optimal
- Tested all solvers with default options
- Communicated results with developers, who
  - Revised software
  - Revised algorithmic options

TEST PROBLEM CHARACTERISTICS

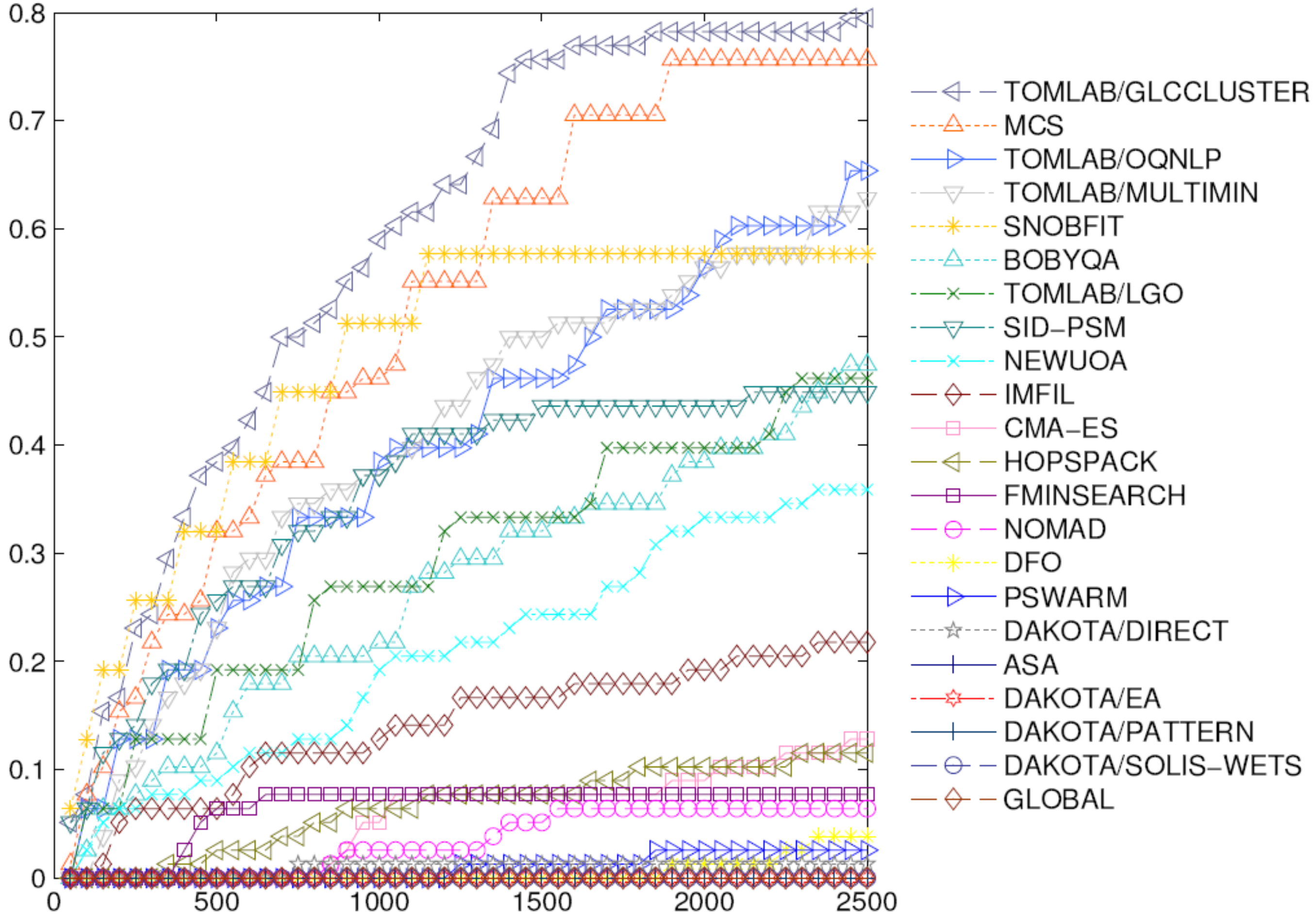


Over 500 problems

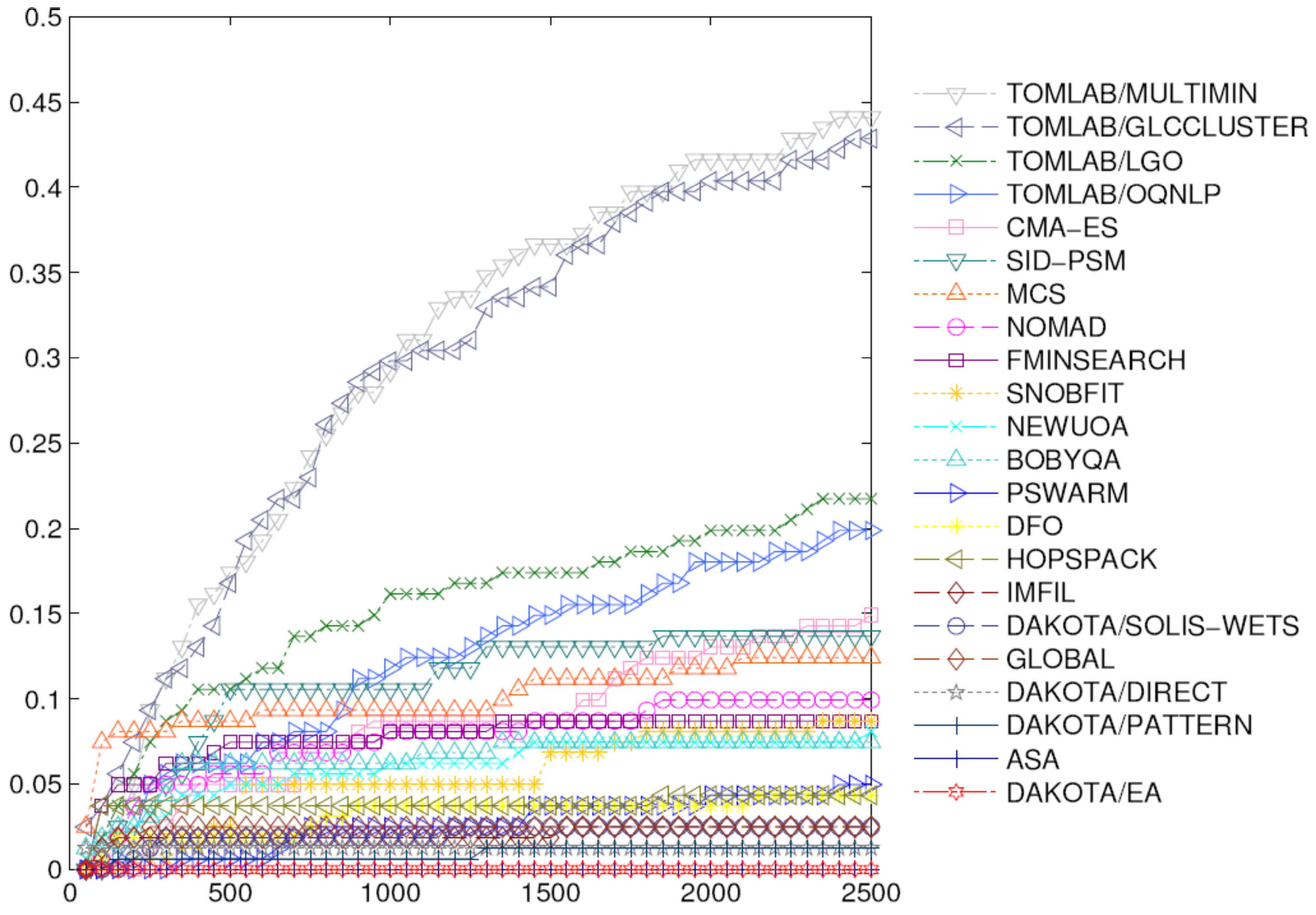
QUESTIONS ADDRESSED

- Which solvers are most likely to find near-global optima?
- Which solvers are most likely to improve starting points?
- Does quality drop significantly as problem size increases?
- Is there a minimal subset of existing solvers that would suffice to solve a large fraction of problems?

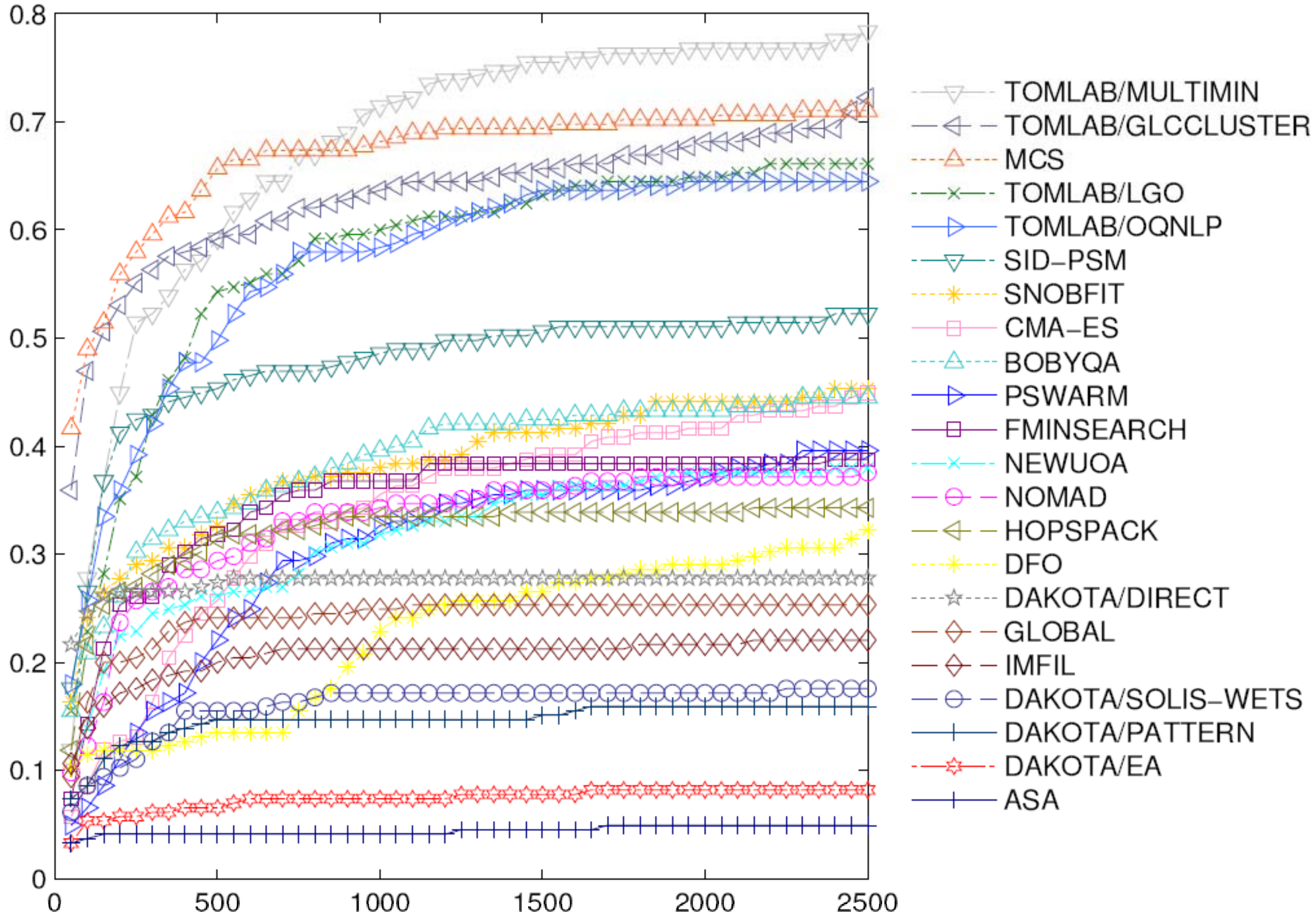
FRACTION OF PROBLEMS SOLVED: CONVEX SMOOTH



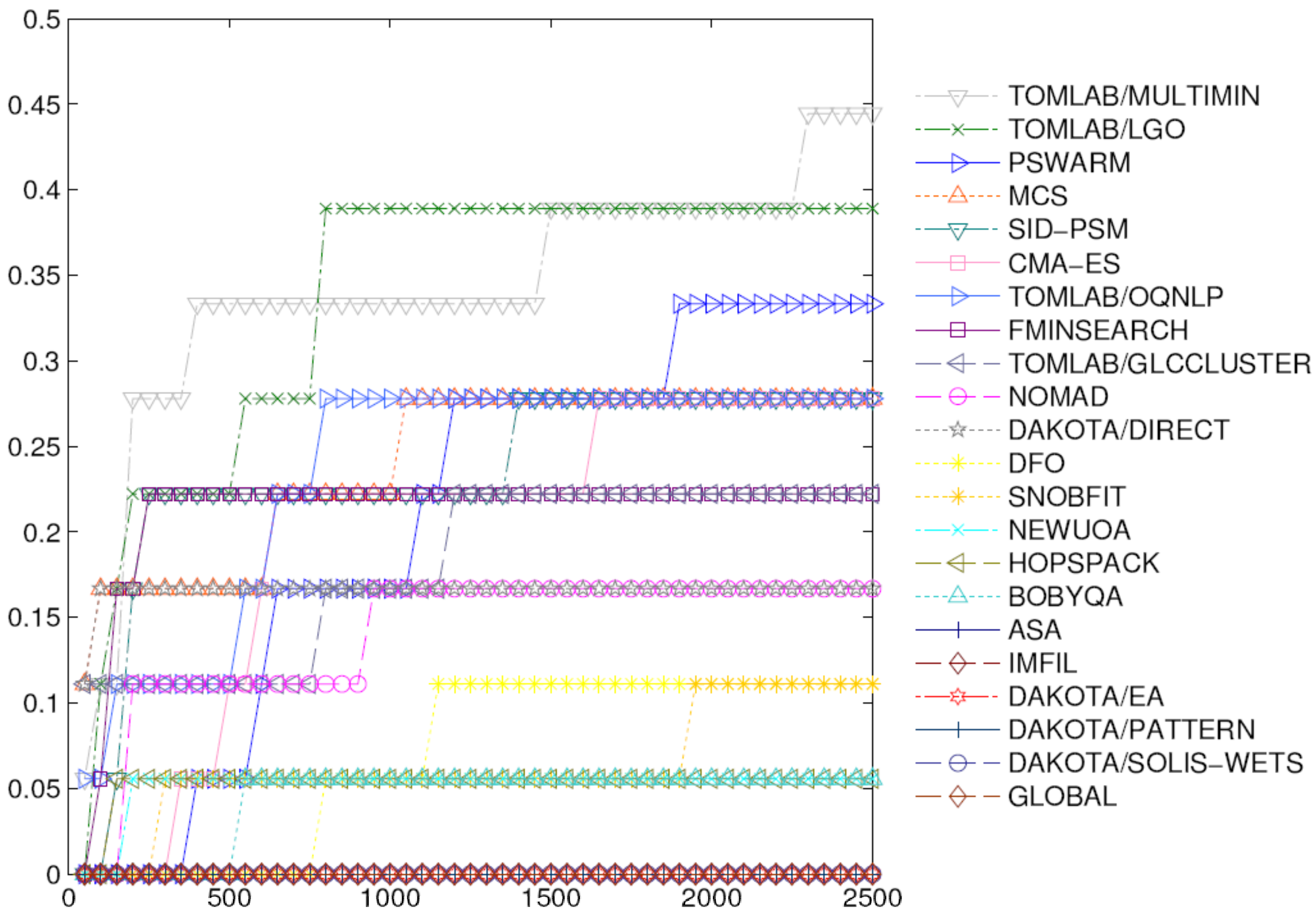
FRACTION OF PROBLEMS SOLVED: CONVEX NONSMOOTH



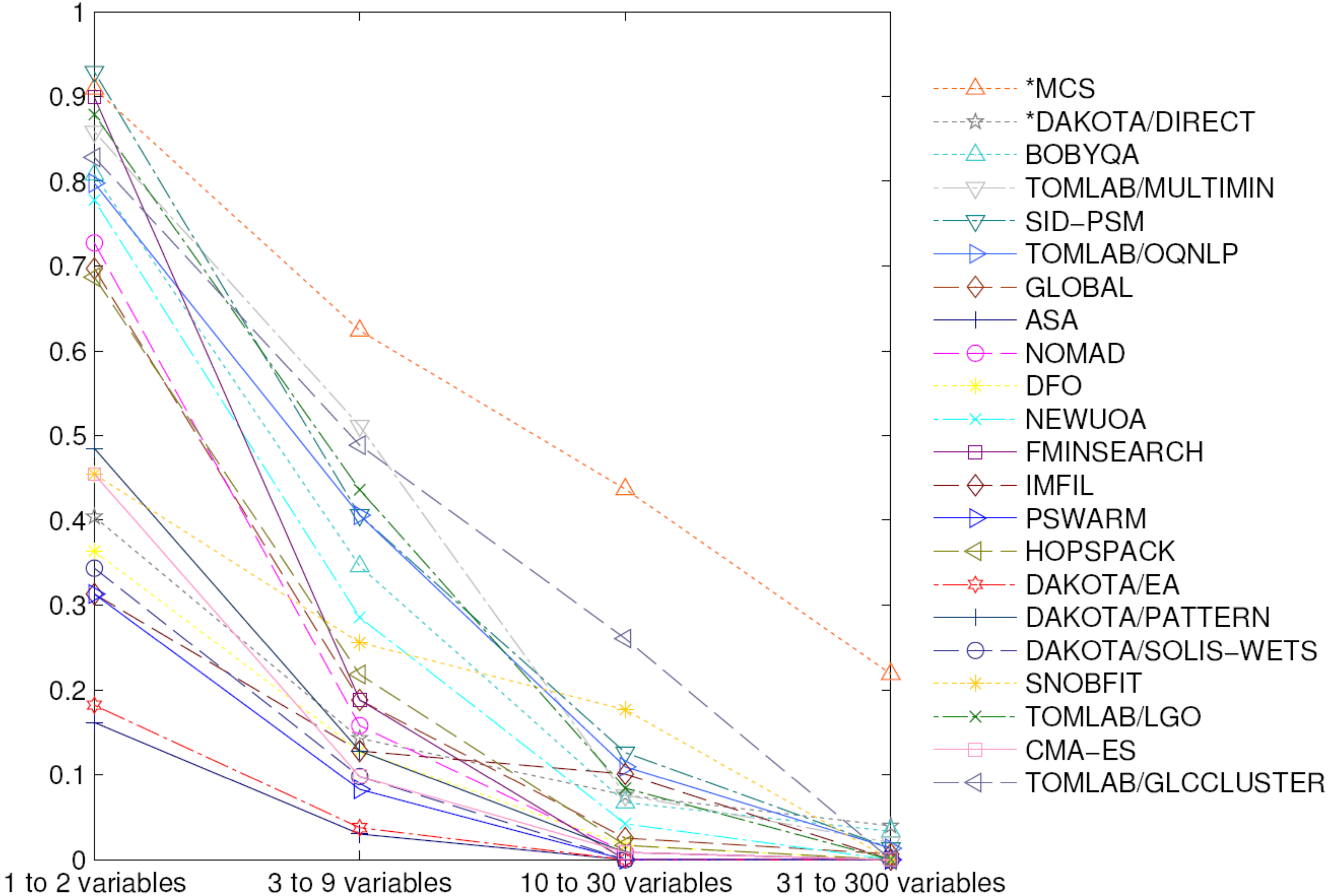
FRACTION OF PROBLEMS SOLVED: NONCONVEX SMOOTH



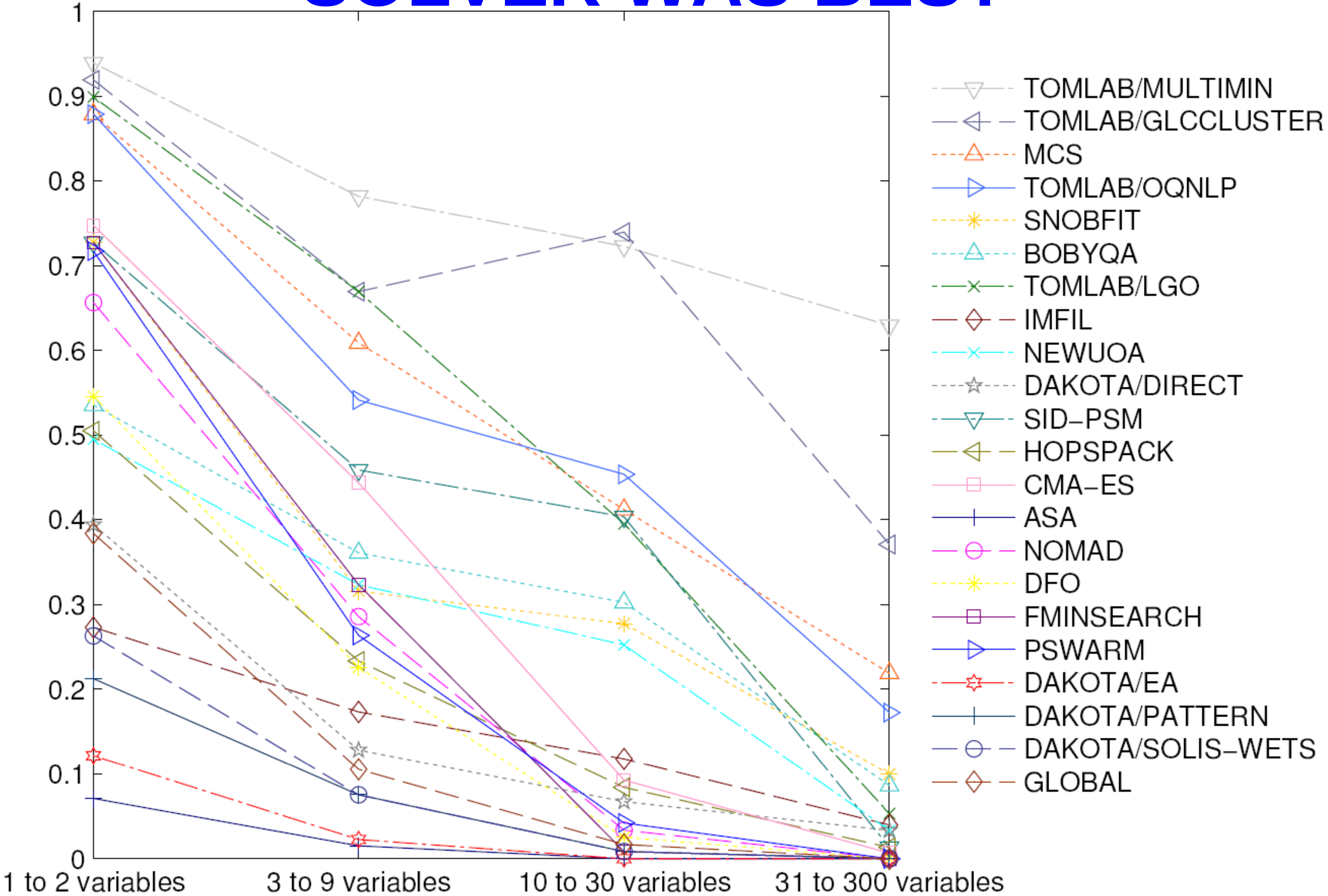
FRACTION OF PROBLEMS SOLVED: NONCONVEX NONSMOOTH



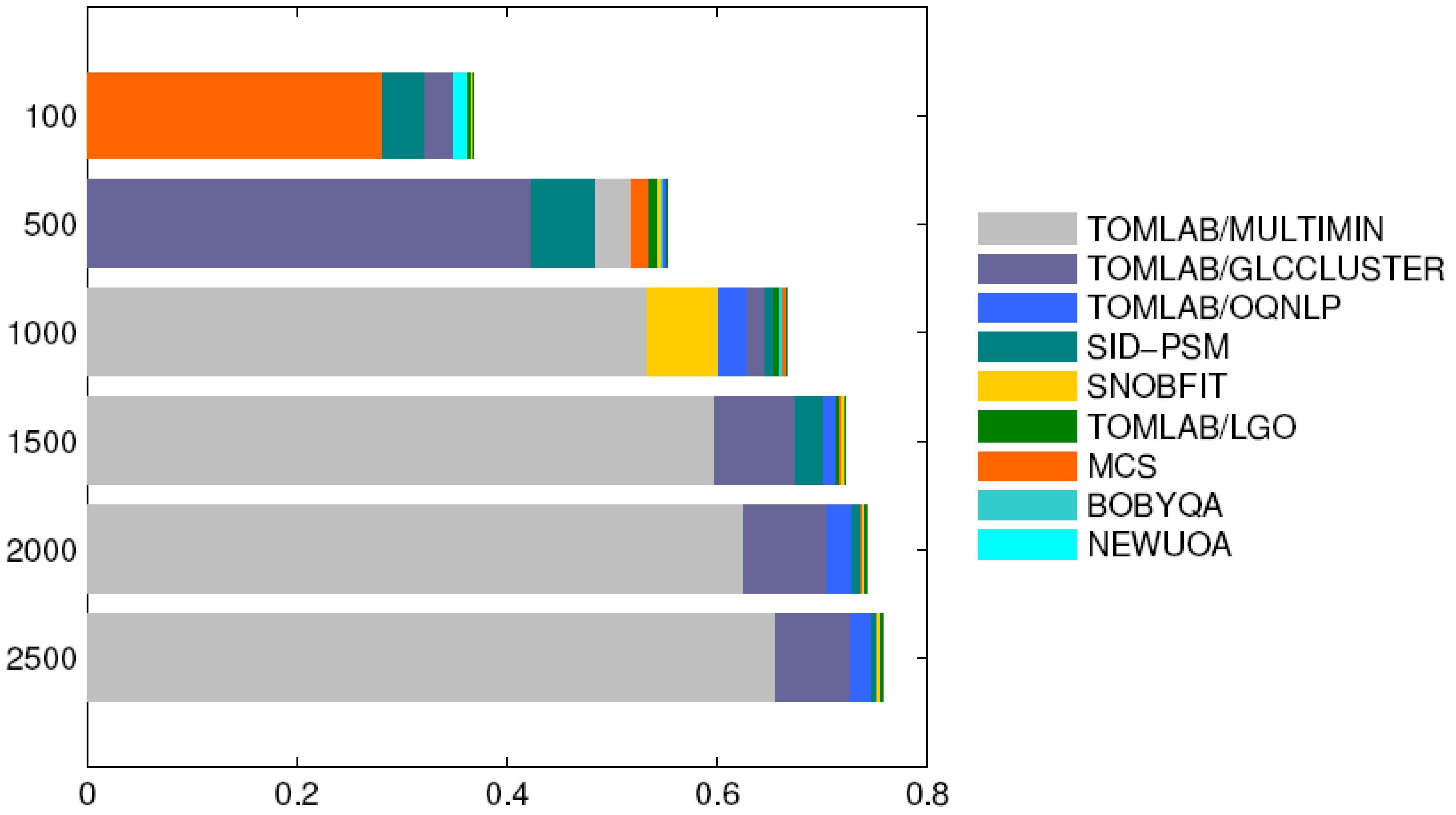
FRACTION OF PROBLEMS SOLVED AS A FUNCTION OF PROBLEM SIZE



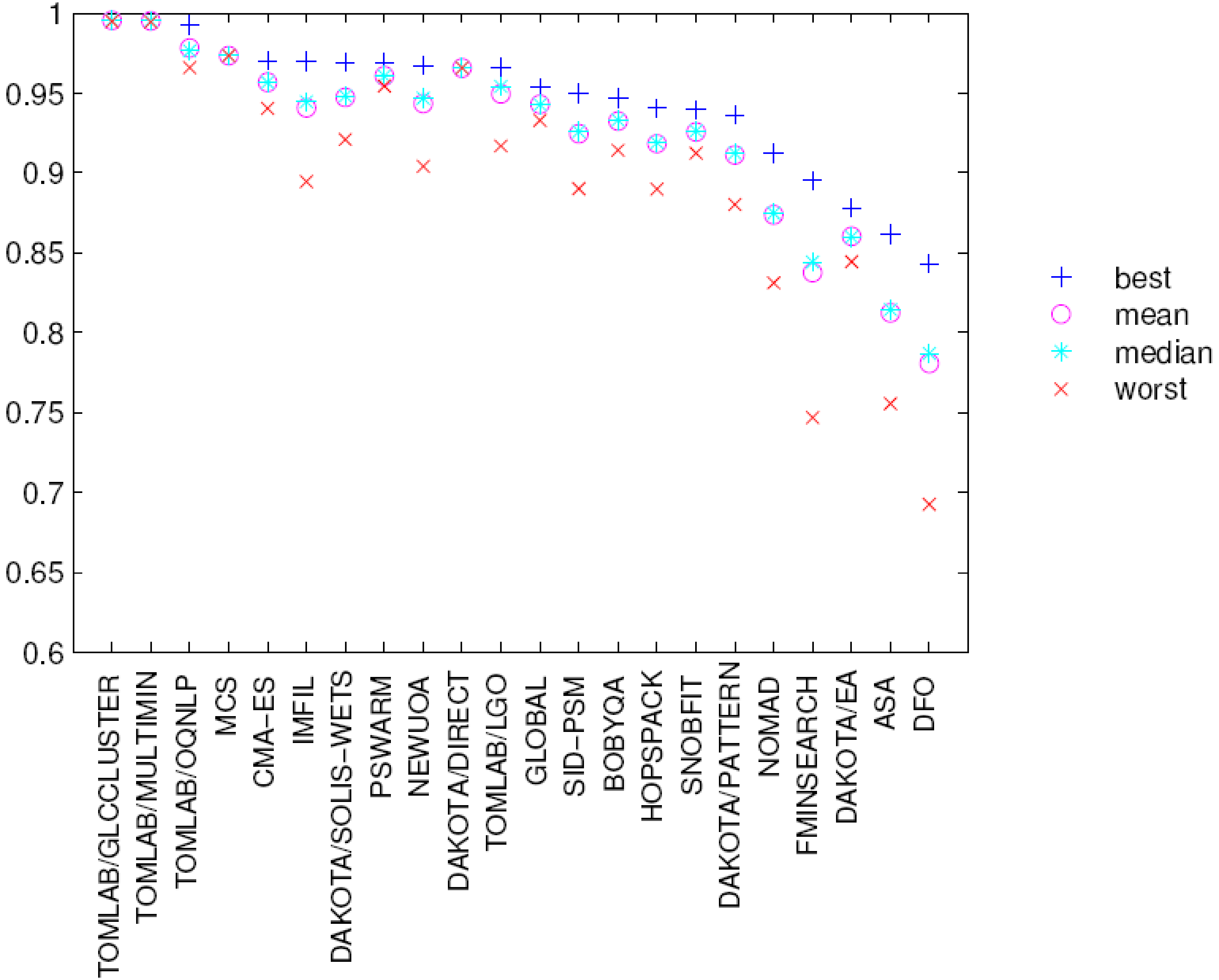
FRACTION OF PROBLEMS SOLVER WAS BEST



MINIMAL SET OF SOLVERS
ALL PROBLEMS

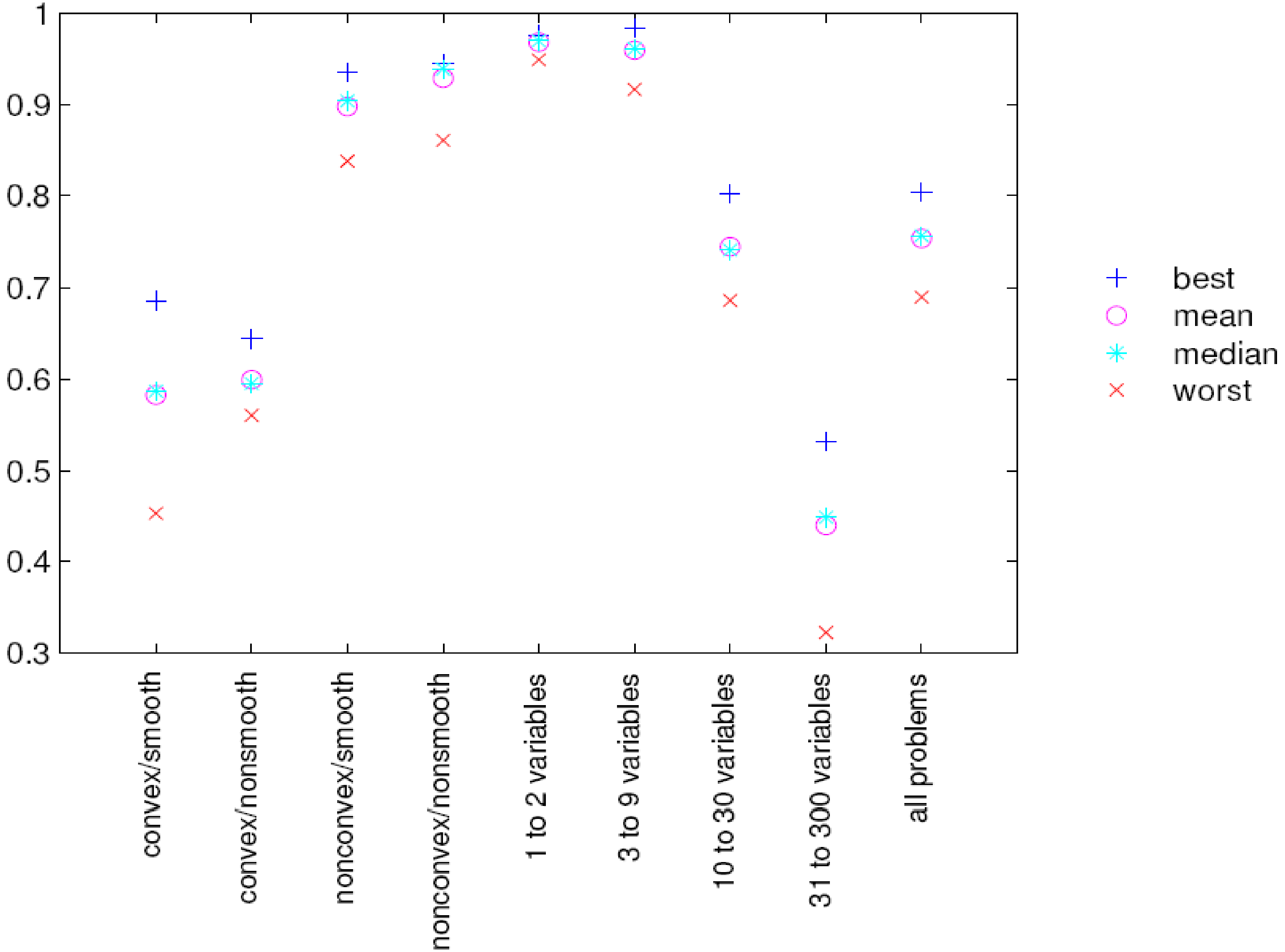


ERROR BARS



1
0.95
0.9
0.85
0.8
0.75
0.7
0.65
0.6
TOMLAB/GLCCLUSTER
TOMLAB/MULTIMIN
TOMLAB/OQNLP
MCS
CMA−ES
IMFIL
DAKOTA/SOLIS−WETS
PSWARM
NEWUOA
DAKOTA/DIRECT
TOMLAB/LGO
GLOBAL
SID−PSM
BOBYQA
HOPSPACK
SNOBFIT
DAKOTA/PATTERN
NOMAD
FMINSEARCH
DAKOTA/EA
ASA
DFO
best
mean
median
worst

ERROR BARS—NEWUOA

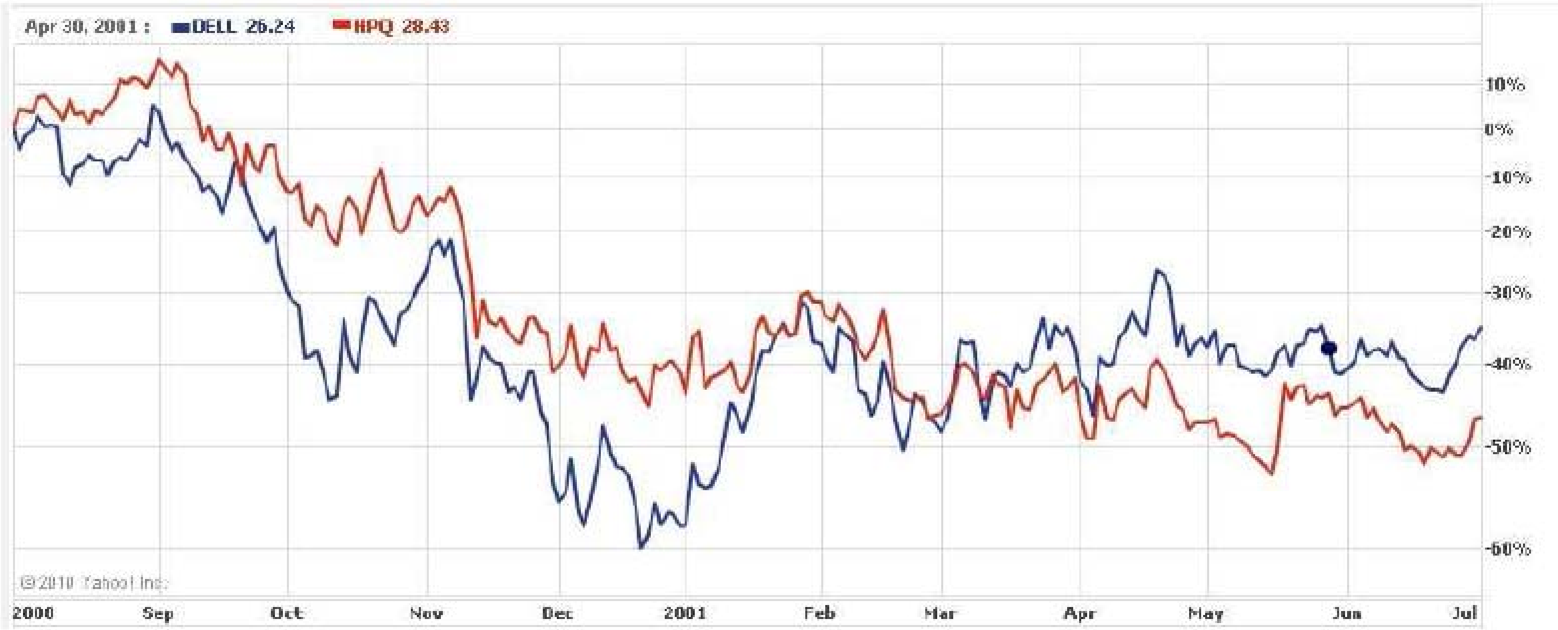


APPLICATION TO TRUE BLACK-BOX MODELS

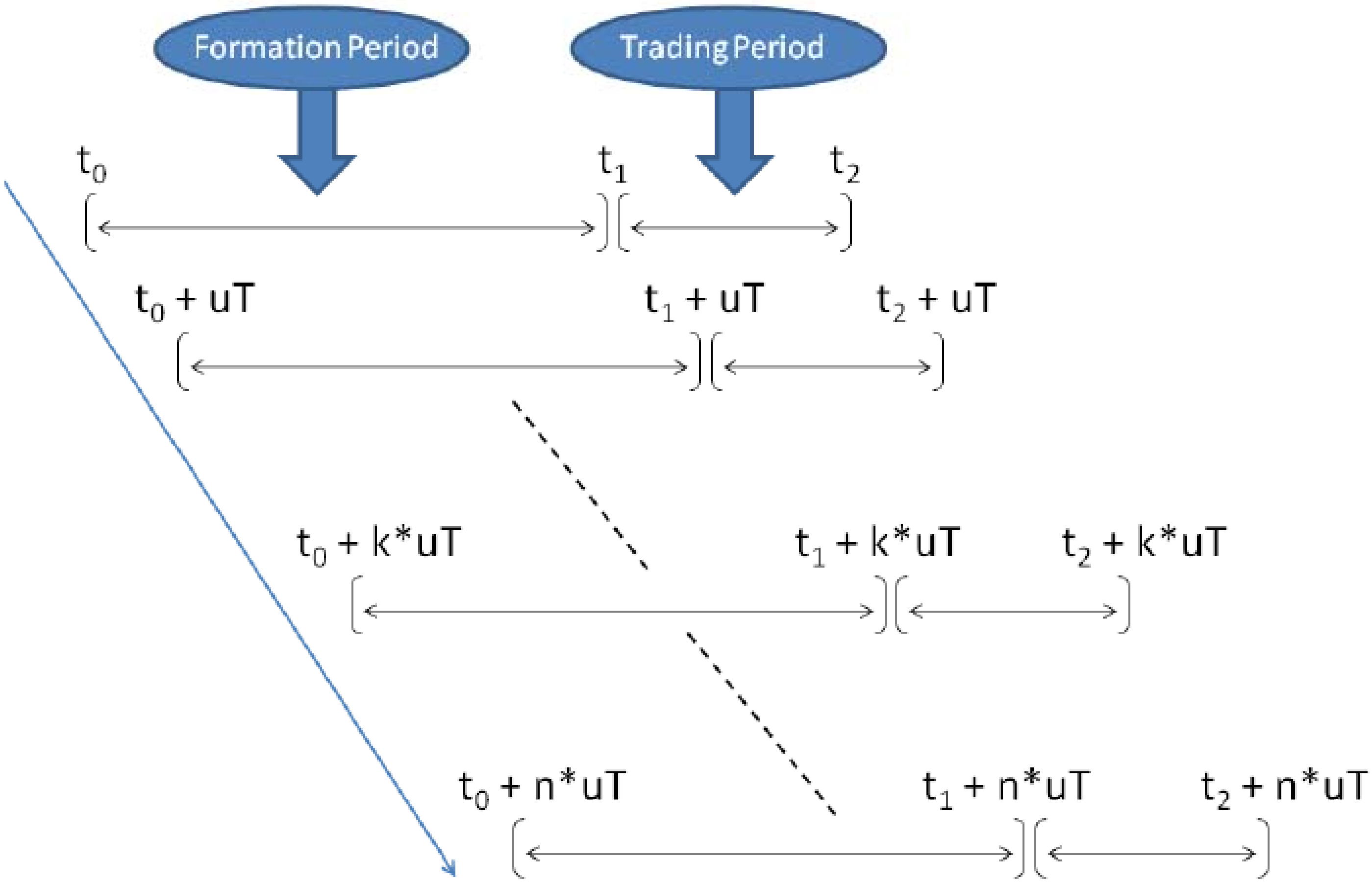
- Portfolio optimization using the omega function
  - R. Desai (MS thesis 2010)
- Pairs trading
  - Y. Zheng (MS thesis 2011)
- Protein structural alignment
  - S. Shah (PhD thesis 2011)
- Optimizing polymerase chain reaction (PCR)
  - K.-F. Chang (MS thesis 2011)

Relative solver performance on black-box models is similar to that presented for 500+ algebraic models

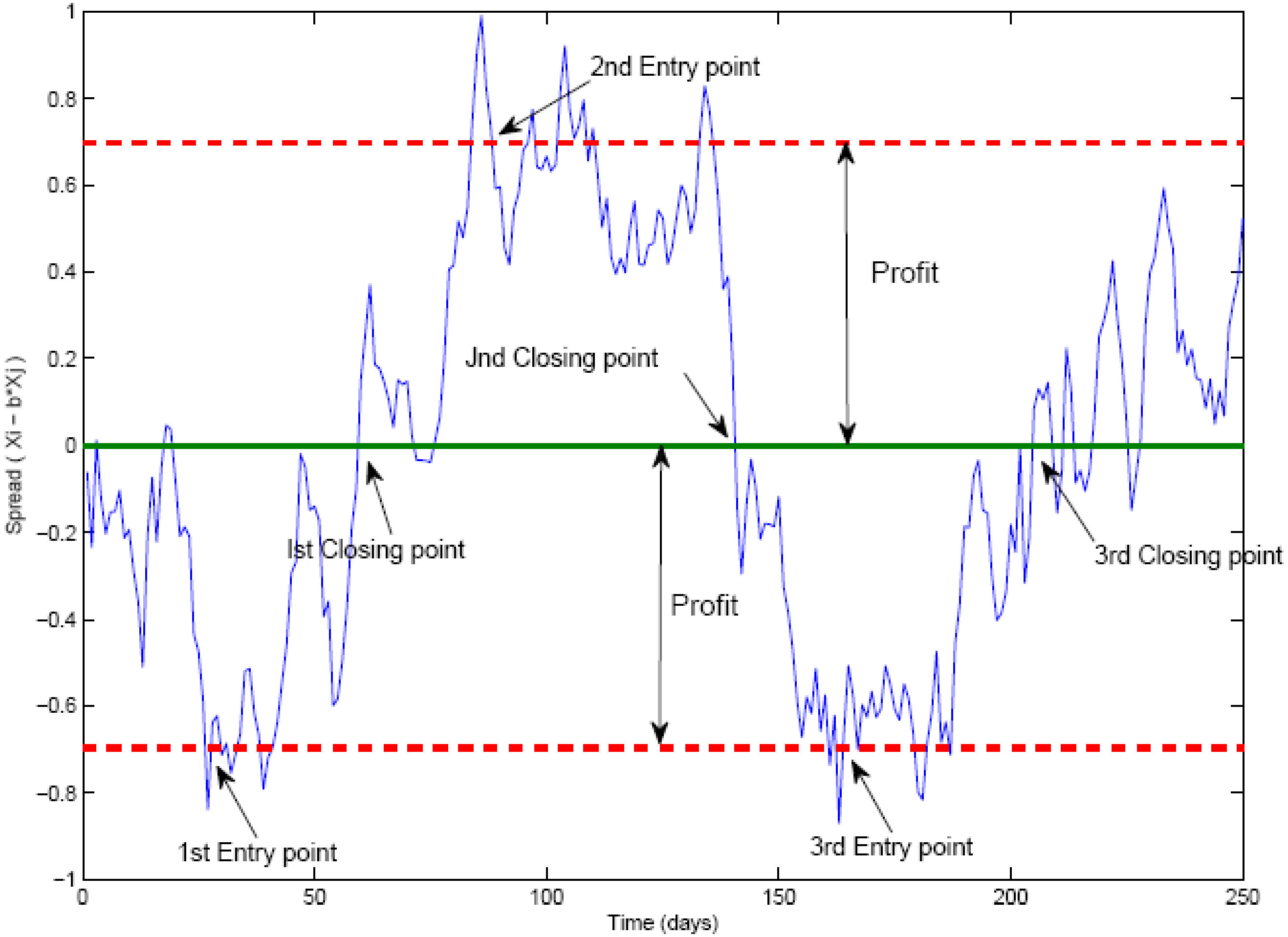
PAIRS TRADING



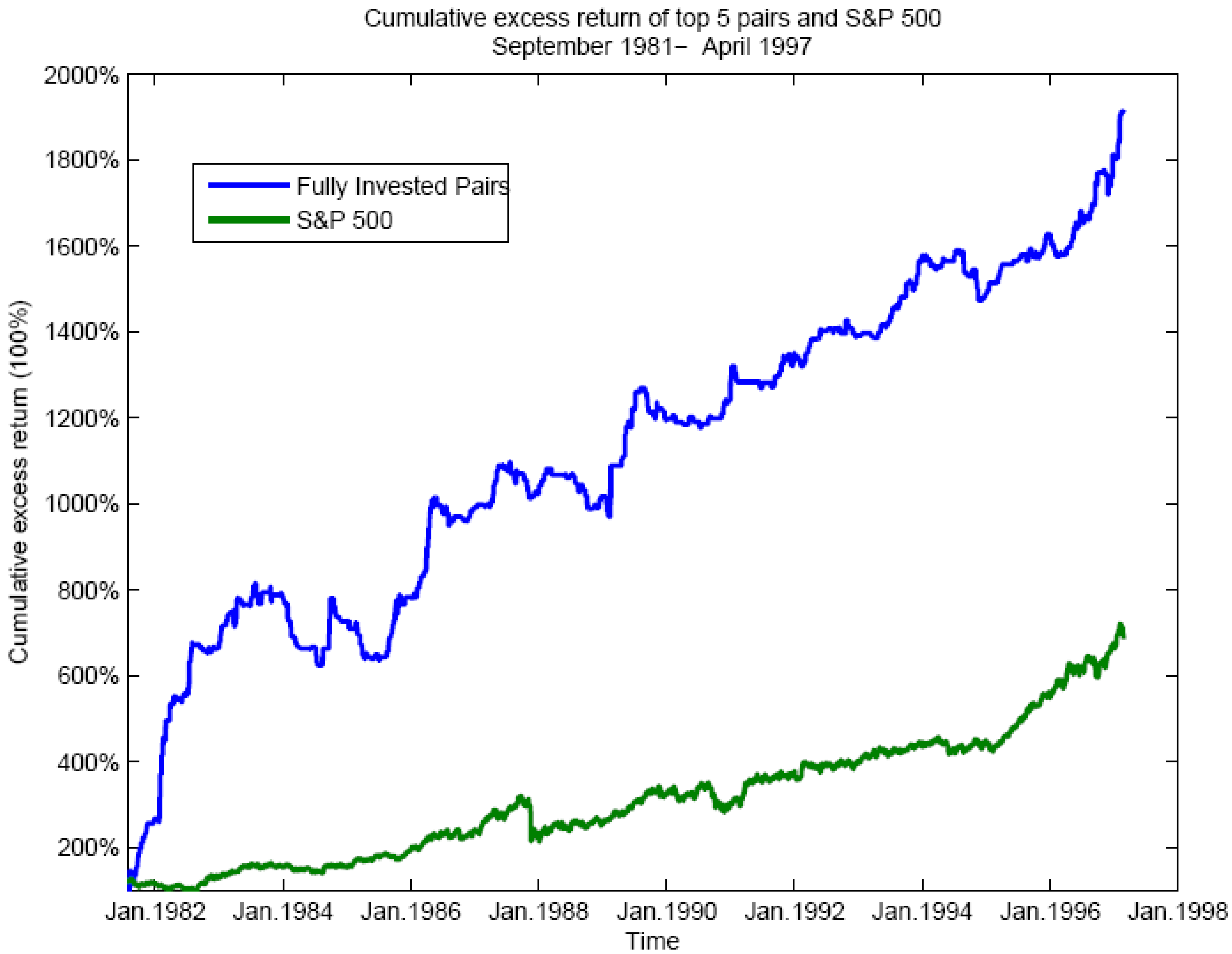
TRAINING AND TRADING



OPTIMAL TRADING THRESHOLDS



CUMULATIVE RETURNS (9/81-4/97)

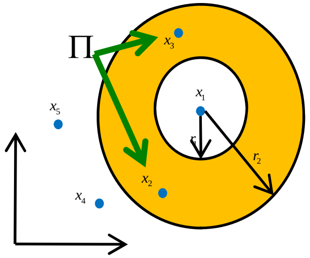


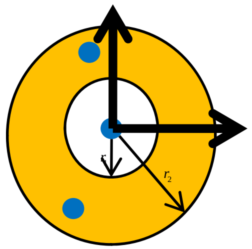
Annual return 19.5% versus 11.8% for S&P 500

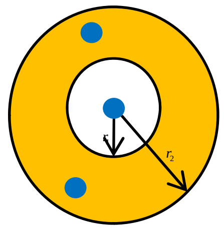
NEW ALGORITHMS

- **Simulation/experimentation is expensive**
- **Solve auxiliary algebraic models to global optimality to expedite search**
  - **Decide where to sample objective function**
    - » **Guarantee geometry**
  - **Construct surrogate models**
    - » **Higher-quality surrogates**
  - **Solve trust-region subproblems**
    - » **Escape local minima; guarantee convergence**
- **BARON is highly efficient for problems below 100 variables**

MODEL-AND-SEARCH LOCAL ALGORITHM







Collect points around current iterate

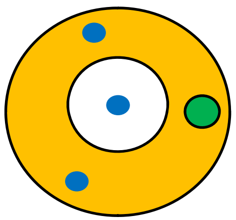
Scale and shift origin to current iterate

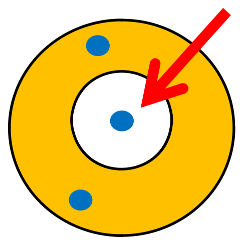
Add points: n linearly independent points

Evaluate points in specific order

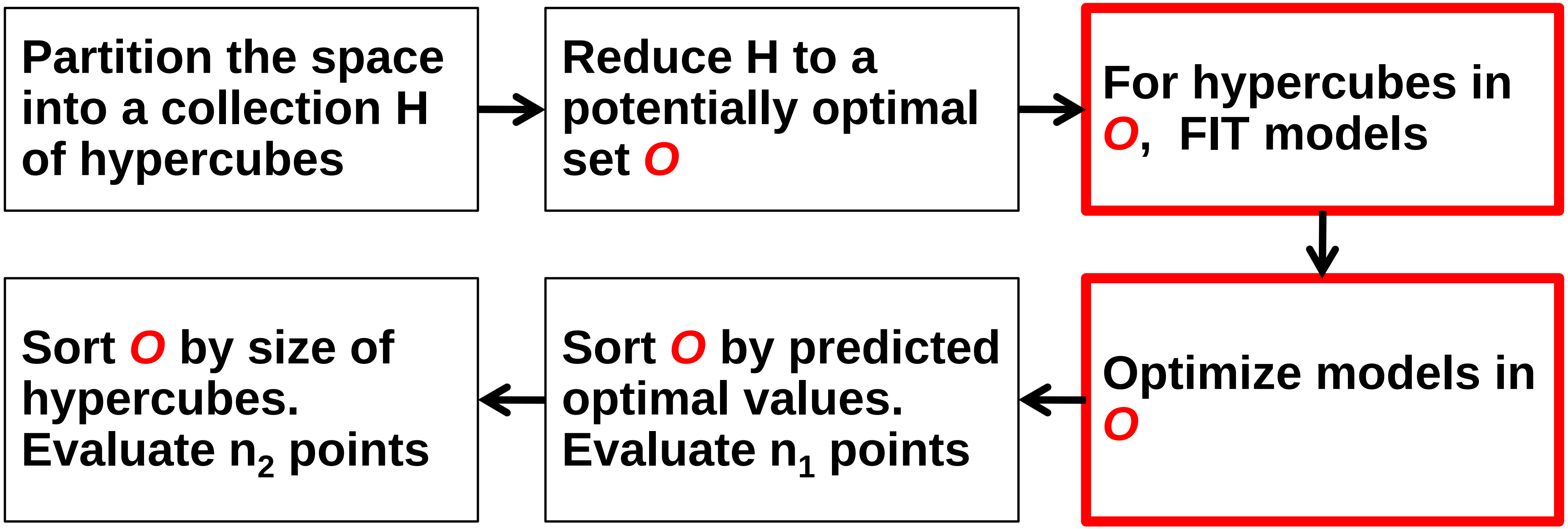
Build and optimize interpolating model

Check for positive basis. Add points if necessary

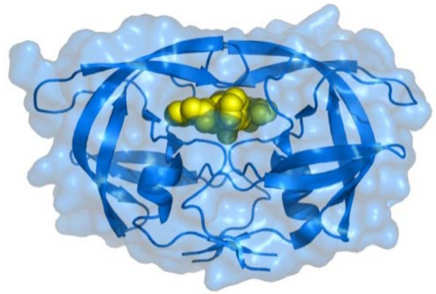




BRANCH-AND-MODEL GLOBAL ALGORITHM

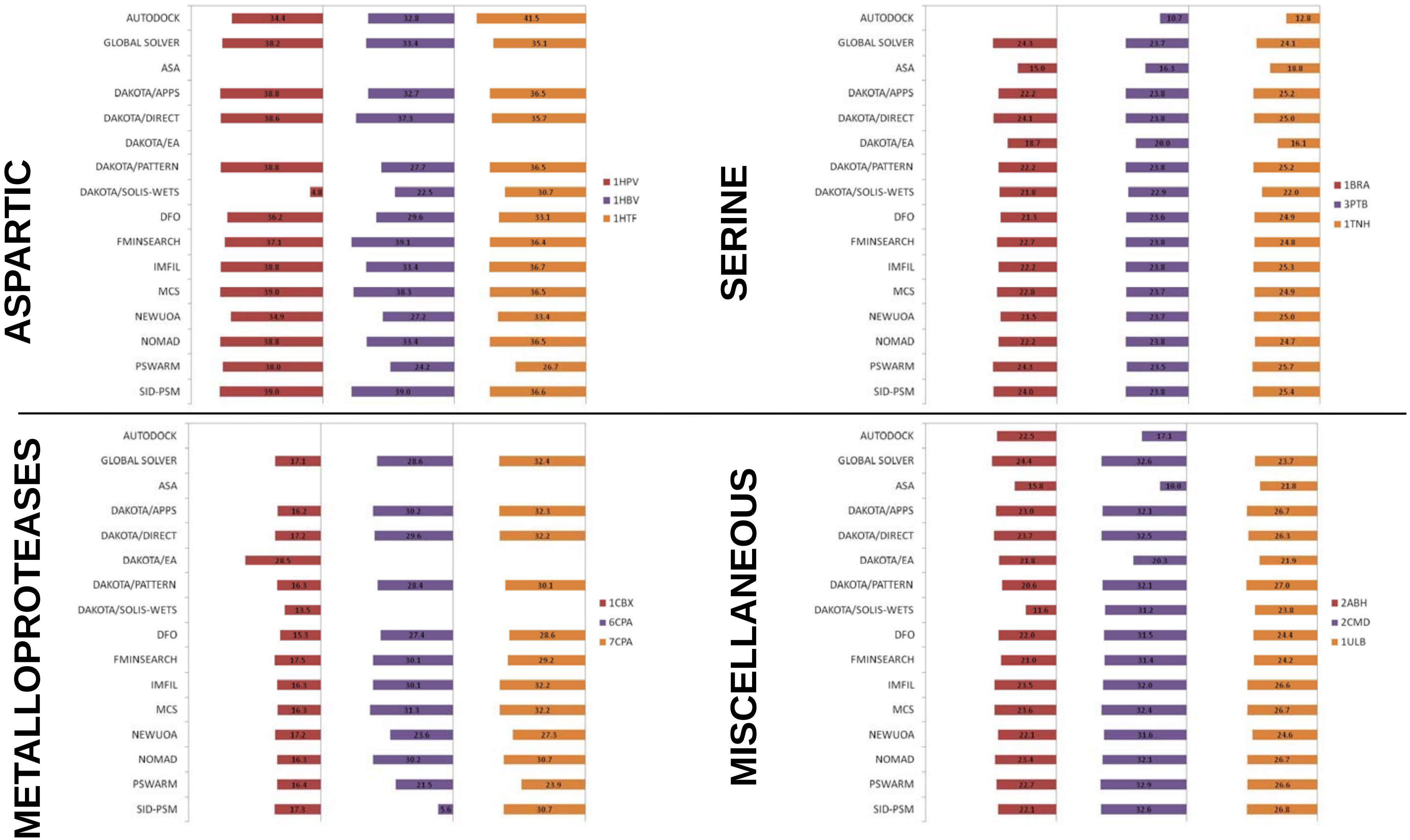


PROTEIN-LIGAND DOCKING



- **Identify binding site and pose**
- **Conformation must minimize binding free energy**
- **Docking packages**
  - **AutoDock, Gold, FlexX …**
  - **Most rely on genetic and other stochastic search algorithms**

BINDING ENERGIES

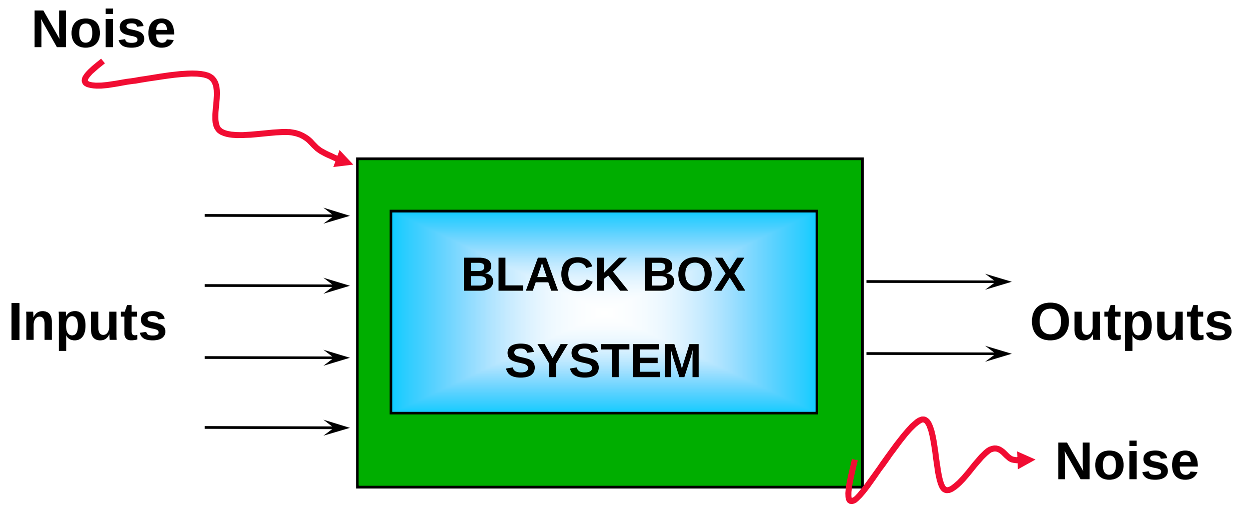


B&M outperformed AutoDock in 11 out of 12 cases, and found the best solution amongst all solvers for 3 complexes

CONCLUSIONS

- **Deterministic solvers perform better than stochastic solvers**
  - **Commercial TOMLAB solvers**
  - **Free MCS/SNOBFIT solvers**

- **Many opportunities**
  - **New algorithms and theory are needed**
  - **Applications abound**
  - **Systematic treatment of noise**



Noise
Inputs
BLACK BOX
SYSTEM
Outputs
Noise