# Latest Progress in Optimization Software

Hans D Mittelmann

School of Mathematical and Statistical Sciences
Arizona State University

INFORMS Annual Meeting
Phoenix, AZ
16 October 2023

# Outline

# DECISION TREE FOR OPTIMIZATION SOFTWARE

## BENCHMARKS FOR OPTIMIZATION SOFTWARE

**By Hans Mittelmann (mittelmann at asu.edu)**

### END OF A BENCHMARKING ERA

**For many years our benchmarking effort had included the solvers CPLEX, Gurobi, and XPRESS. Through an action by Gurobi at the 2018 INFORMS Annual Meeting this has come to an end. IBM and FICO demanded that results for their solvers be removed and then we decided to temporarily remove those of Gurobi as well. CPLEX was in fifteen of the benchmarks, Gurobi and XPRESS in thirteen. See here for more details. In late November 2019 selected benchmarks for Gurobi were added.**
**A partial record of previous benchmarks can be obtained from this webpage and some additional older benchmarks**

# What is on top of the Benchmark Page?

- Links to history and older benchmarks
- Especially the 2018 event eliminating CPLEX and XPRESS
- Hint that full solver logfiles are provided
- Explanation why performance profiles are not used
- Link to Matt Miltenberger's visualization tool

# Visualizations of Mittelmann benchmarks

Interactive plots showing pairwise time differences for every instance and every solver

View On GitHub

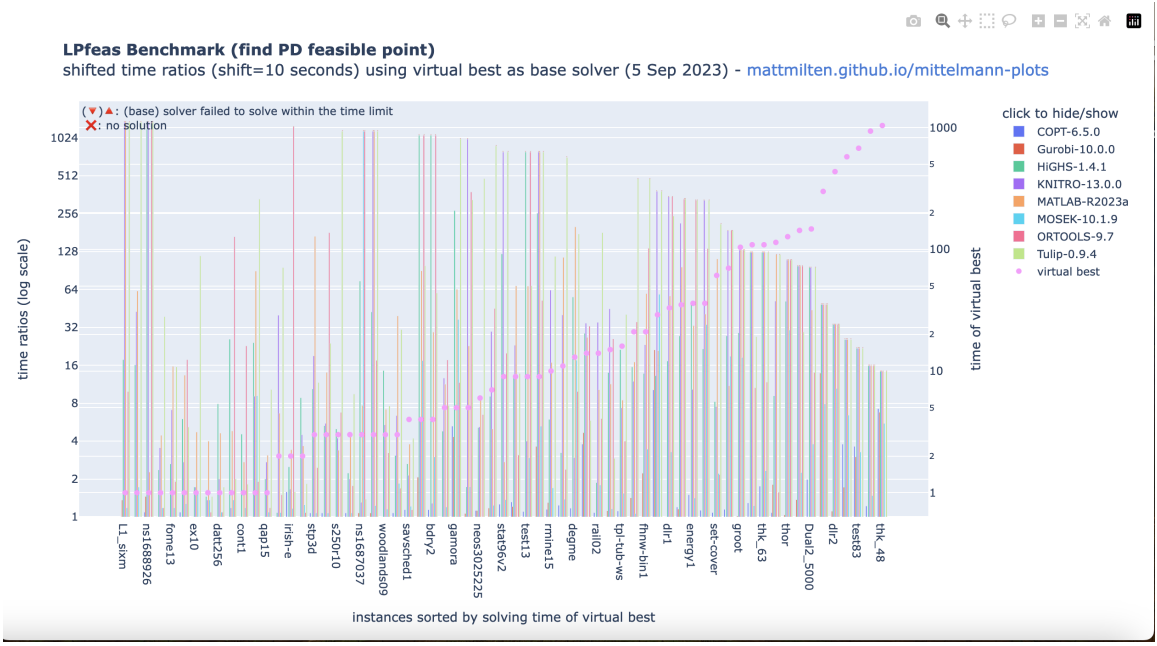


Interactive charts comparing the results of Hans Mittelmann's benchmarks. Each solver can be selected to show pairwise running time factors for every other solver in the respective benchmark. These plots should make browsing the results easier. The score (scaled shifted geometric mean) is recomputed using the reported solving times. We also calculate a "virtual best" or "ensemble" solver that picks the best performance over all solvers for every single problem instance. This might reveal how much potential the individual solvers still have. Please let me know if you have a question or if there is an error.

## LPfeas Benchmark (find PD feasible point) (5 Sep 2023)

Choose base solver for comparison:

| solver | score (as reported) | solved of 65 |
|---|---|---|
| virtual best | 0.72 | 100% |
| COPT-6.5.0 | 1.00 (1.00) | 100% |
| Gurobi-10.0.0 | 1.07 (1.07) | 100% |
| MOSEK-10.1.9 | 2.49 (2.49) | 98% |


LPfeas Benchmark (find PD feasible point)
shifted time ratios (shift=10 seconds) using virtual best as base solver (5 Sep 2023) - mattmilten.github.io/mittelmann-plots
(▼)▲: (base) solver failed to solve within the time limit
✕: no solution
time ratios (log scale)
1024
512
256
128
64
32
16
8
4
2
1
time of virtual best
1000
5
2
100
5
2
10
5
2
1
click to hide/show
COPT-6.5.0
Gurobi-10.0.0
HiGHS-1.4.1
KNITRO-13.0.0
MATLAB-R2023a
MOSEK-10.1.9
ORTOOLS-9.7
Tulip-0.9.4
virtual best
L1_sixm
ns1688926
fome13
ex10
datt256
cont1
qap15
irish-e
stp3d
s250r10
ns1687037
woodlands09
savsched1
bdry2
gamora
neos3025225
stat96v2
test13
rmine15
degme
rail02
tpl-tub-ws
fhnw-bin1
dlr1
energy1
set-cover
groot
thk_63
thor
Dual2_5000
dlr2
test83
thk_48
instances sorted by solving time of virtual best

# History and Scope of the Benchmarks

- Start about 1998
- Summary presented at INFORMS Annual starting 2002
- 22 benchmarks in these areas
  - TSP, (MI)LP, (MI)SOCP, SDP, convex/nonconvex (MI)QCP, QUBO, (MI)NLP, MPEC
- Codes evaluated: 48, of those 28 actively developed, 14 commercial
- Running times per code: up to 3 weeks

## COMBINATORIAL OPTIMIZATION

- *Concorde-TSP with different LP solvers (3-3-2023)*

## LINEAR PROGRAMMING

The Simplex and Barrier benchmarks are replaced by benchmarks that show how well solvers find a primal-dual feasible point (as barrier methods in general do) or an optimal basic solution (as simplex methods in general do). Codes do not have to use a textbook version of either method.

- *LPfeas Benchmark (find a PD feasible point) (10-4-2023)*
- *LPopt Benchmark (find optimal basic solution) (10-5-2023)*
- *Large Network-LP Benchmark (commercial vs free) (10-4-2023)*

## MIXED INTEGER LINEAR PROGRAMMING

- *MILP Benchmark - MIPLIB2017 (10-4-2023)*
- *MILP cases that are slightly pathological (10-5-2023)*
- *Infeasibility Detection for MILP Problems (10-4-2023)*

## SEMIDEFINITE/SQL PROGRAMMING

- *SQL problems from the 7th DIMACS Challenge (8-8-2002)*

☞ *MISOCP Benchmark (9-27-2023)*

**NONLINEAR PROGRAMMING**

☞ *AMPL-NLP Benchmark (1-18-2022)*

**MIXED INTEGER QPS AND QCPS**

☞ *Non-commercial convex QP Benchmark (9-16-2021)*

☞ *Binary Non-Convex QPLIB Benchmark (7-12-2023)*

☞ *Non-Convex QUBO-QPLIB Benchmark (9-20-2023)*

☞ *Discrete Non-Convex QPLIB Benchmark (non-binary) (7-14-2023)*

☞ *Continuous Non-Convex QPLIB Benchmark (7-18-2023)*

☞ *Convex Continuous QPLIB Benchmark (9-28-2023)*

☞ *Convex Discrete QPLIB Benchmark (9-27-2023)*

**MIXED INTEGER NONLINEAR PROGRAMMING**

☞ *MINLP Benchmark (7-10-2023)*

# Selected Benchmarks

These seven benchmarks follow next:

- convex: LPfeas, LPopt, MILP, SDP
- nonconvex: QUBO-QPLIB, MIQCP, MINLP
- remember: "the great watershed in optimization isn't between linearity and nonlinearity, but convexity and nonconvexity"
- Talk will report incremental progress; more substantial progress will be announced at meeting: Gurobi-Nonlinear, XPRESS-Global etc

# LPfeas Benchmark (find PD feasible point)

Top of benchmark table shown

- Total of 65 instances (16 hidden)
- Sizes up to 30m/30m/35m rows/cols/nonzeros
- Intel i7-11700K, 3.6GHz, 64GB, 15,000 secs wall clock
- Own instance selection

| 65 probs | 1.11 | 1 | 1.80 | 66.4 | 2.60 | 19.2 | 26.9 | 22.9 | 17.0 |
|---|---|---|---|---|---|---|---|---|---|
| solved | 65 | 65 | 62 | 36 | 64 | 52 | 50 | 43 | 49 |

| probs | Gurobi | COPT | MDOPT | TULIP | MOSEK | HiGHS | MATL | KNITRO | PDLP% |
|---|---|---|---|---|---|---|---|---|---|

# LPopt Benchmark (find optimal basic solution)

- Same own instance selection as in LPfeas benchmark

  COPT-7.0.0, MindOpt-1.0.0, HiGHS-1.6.0, Gurobi-10.0, OptVerse-0.2.13

  MOSEK-10.1.9, SOPLEX-6.0.0

- Intel i7-11700K, 3.6GHz, 64GB, 15,000 secs wall clock

| 65 probs | 26.1 | 1.36 | 1 | 1.89 | 3.95 | 5.61 | 17.4 | 57.0 | 87.2 | 39.8 |
|---|---|---|---|---|---|---|---|---|---|---|
| solved | 40 | 65 | 65 | 63 | 57 | 52 | 51 | 33 | 32 | 42 |

| probs | CLP | Gurob | COPT | MDOPT | OPTV | MOSEK | HiGHS | GLOP | SPLX | MATL |
|---|---|---|---|---|---|---|---|---|---|---|

# The MIPLIB2017 Benchmark Instances

- Total of 240 instances
- Sizes up to 1.5m/1m/43m rows/cols/nonzeros
- Intel i7-11700K, 3.6GHz, 64GB, 7,200 secs wall clock

| | CBC | Gurobi | COPT | SCIP | SCIPC | HiGHS | Matlab | SMOO | XSMOO | MDOPT |
|---|---|---|---|---|---|---|---|---|---|---|
| unscal | 1328 | 81.5 | 126 | 888 | 727 | 720 | 2715 | 612 | 510 | 301 |
| scaled | 16.3 | 1 | 1.54 | 10.9 | 8.92 | 8.83 | 33.3 | 7.51 | 6.26 | 3.70 |
| solved | 107 | 227 | 212 | 137 | 152 | 159 | 73 | 163 | 172 | 196 |

(X)SMOO: (X)Smoothie (FiberSCIP+HiGHS, Cplex/Soplex)

# Several SDP-codes on sparse and other SDP problems

- Total of 75 instances, Own instance selection
- Sizes up to 100k/100k/450m variables/constraints/nonzeros
- AMD Ryzen 9 5900X, 12 cores, 128GB , 40,000 secs wall clock

| | 1 | 5.21 | 3.22 | 10.5 | 5.14 | 28.9 | 7.86 | 1.44 |
|---|---|---|---|---|---|---|---|---|
| count of "a" | 6 | 5 | 2 | 17 | 13 | 2 | 11 | 12 |
| solved of 75 | 75 | 70 | 73 | 61 | 69 | 62 | 70 | 75 |

| problem | COPT | CSDP | MOSEK | SDPA | SDPT3 | SeDuMi | HDSDP | MDOPT |
|---|---|---|---|---|---|---|---|---|

"a": reduced accuracy

# Nonconvex QUBO-QPLIB Benchmark

- Total of 23 instances
- (variables,densities) from (1225,0.3) to (150,78.9)
- AMD Ryzen 9 5900X, 12 cores, 128GB , 3,600 secs wall clock
- CPLEX in BARON, SCIP, OCTERACT, QUBOWL, Gurobi in SHOT

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| mean | 1.79 | 5.46 | 2.36 | 1.74 | 1.41 | 1 | 5.40 | 1.99 |
| solved | 12 | 8 | 12 | 12 | 13 | 15 | 9 | 11 |

| prob# | BARON | SCIP | MCSPARSE | OCTERACT | GUROBI | QUBOWL | BIQBIN | SHOT |
|---|---|---|---|---|---|---|---|---|

# Discrete Non-Convex QPLIB Benchmark (non-binary)

- Total of 160 instances
- Sizes up to 36k/120k/5k variables/total constraints/nonlin constraints
- Intel Xeon E5-4657L, 48 cores, 512GB, 8 threads, 10,800 secs wall clock

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| mean | 31.2 | 7.24 | 1.19 | 37.5 | 18.7 | 3.21 | 1 |
| solved | 29 | 56 | 81 | 15 | 37 | 79 | 81 |

| prob# | ANTIGONE | BARON | SHOT$ | MINOTAUR | SCIP | OCTERACT | GUROBI$ |
|---|---|---|---|---|---|---|---|

$ Gurobi has 16 instances with constraint violations > 1e-6
SHOT has in all solved instances constraint violations < 1e-8
OCTERACT does not report.

# Mixed Integer Nonlinear Programming Benchmark

- Total of 87 instances; run through GAMS
- Sizes up to 100k/100k/5k variables/lin constraints/nonlin constraints
- Intel i7-11700K, 3.6GHz, 64GB, 7,200 secs wall clock
- Instances from MINLPLib

| | ANTIGONE | BARON | COUENNE | LINDO | OCTERACT | SCIP |
|---|---|---|---|---|---|---|
| scaled | 39.3 | 2.3 | 89.8 | 32.8 | 1.0 | 10.3 |
| solved | 53 | 77 | 24 | 42 | 87 | 64 |

Since Octeract will be removed from GAMS, it will be frozen at version 4.7.1

# Conclusions

## Some Takeaways

- Gurobi and BARON have serious competition
- SDP code development has picked up again; COPT, MindOpt, MOSEK lead
- SMOO is strong in MILP but Gurobi, COPT, MindOpt lead
- SCIP is strong in MINLP: OCTERACT, BARON, SCIP
- Global optimization is now standard but challenges remain

# THANK YOU

Questions?

Slides of talk at
http://plato.asu.edu/talks/informs2023.pdf