GO Competition Challenge 2: Analysis and Lessons Learned

October 16, 2022

PNNL: Brent Eldridge, Stephen Elbert, Arun Veeramany, and Jesse Holzer

ASU: Hans Mittelmann

PNNL-SA-174880

Pacific Northwest NATIONAL LABORATORY

Partners

More information:

https://GOCompetition.Energy.gov

AIMMS
AMPL
GAMS
GUROBI OPTIMIZATION
NREL NATIONAL RENEWABLE ENERGY LABORATORY
ATM
TEXAS The University of Texas at Austin
IBM
mosek
SIEMENS Ingenuity for life
WISCONSIN UNIVERSITY OF WISCONSIN–MADISON
ASU ARIZONA STATE UNIVERSITY
Los Alamos NATIONAL LABORATORY EST. 1943
Georgia Tech

Pacific Northwest
NATIONAL LABORATORY

Grid Optimization (GO) Competition Goals

- Accelerate the development of transformational and disruptive methods for solving optimization problems related to the electric power grid
- Provide a transparent, fair, and comprehensive evaluation of new solution methods.
- Challenge 2 – September 2020 – October 2021 (4 divisions completed)
  - $2.4 million in prizes awarded to 9 teams
  - 15 teams participated in the Final Event, C1 winning teams funded by prize money
  - 11 million CPU-hours
  - Data from Georgia Tech, Texas A&M, UW-Madison

Challenge 2 Problem Formulation

- Built on Challenge 1 SCOPF problems (a minimization problem)
  - Single period ACOPF with security constraints
  - Short term operational actions – 5 to 15 minutes prior to real time
  - Use in planning – pre-determine actions that can be deployed in real time

- Outline for today:
  - Problem hardness
  - Penalties
  - Flexible demand
  - Solve time
  - Solution ensembles
  - Transmission switching
  - Use of HPC

Problem Hardness Measures

A priori method

How much potential for improvement?

$$H_{prior} = \frac{\text{relaxed upper bound cost delta}}{\text{cost of feasible ACOPF solution}}$$

A posteriori methods

How much benefit from selecting the best algorithms?

$$H_{post,0} = \frac{\text{best score} - \text{2nd score}}{\text{2nd score}}$$

$$H_{post,1} = \frac{stdev(\text{5 best scores})}{avg(\text{5 best scores})}$$

$$H_{post,2} = \frac{stdev(\text{5 best scores})}{max(\text{5 best scores})}$$

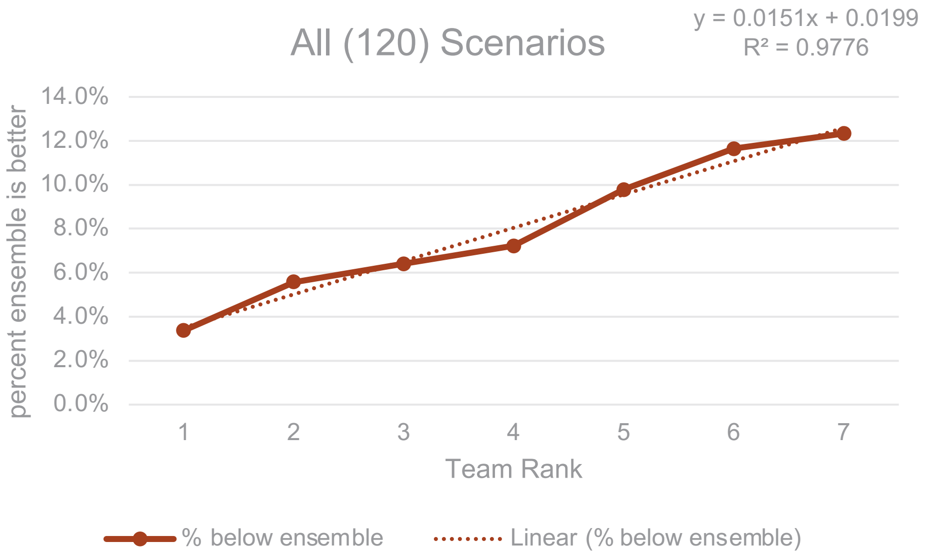
Pearson Coefficient with H_{prior}:

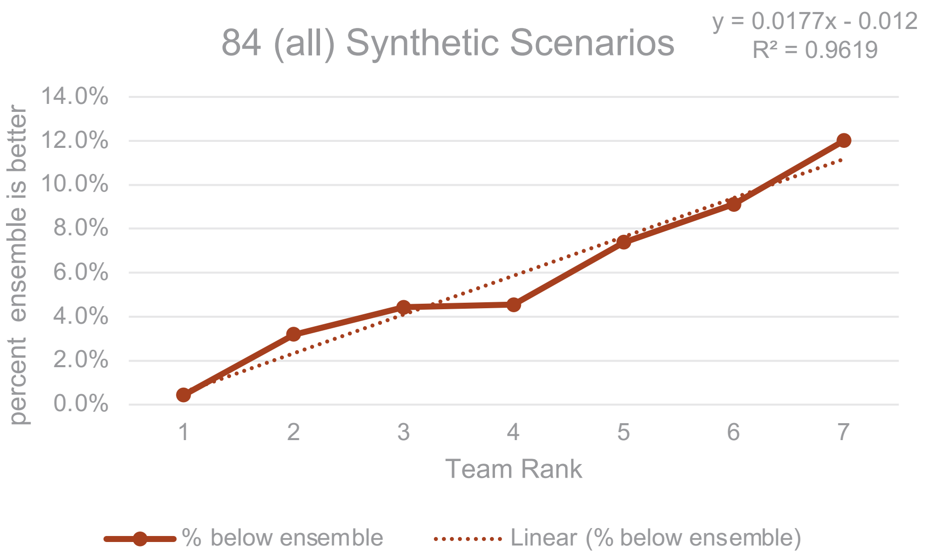
Scenarios (#)	$H_{post,0}$	$H_{post,1}$	$H_{post,2}$
All (120)	-0.1327	-0.0692	-0.0692
Synthetic (84)	-0.2055	-0.1593	-0.1608
Industry (36)	0.811	0.917	0.915

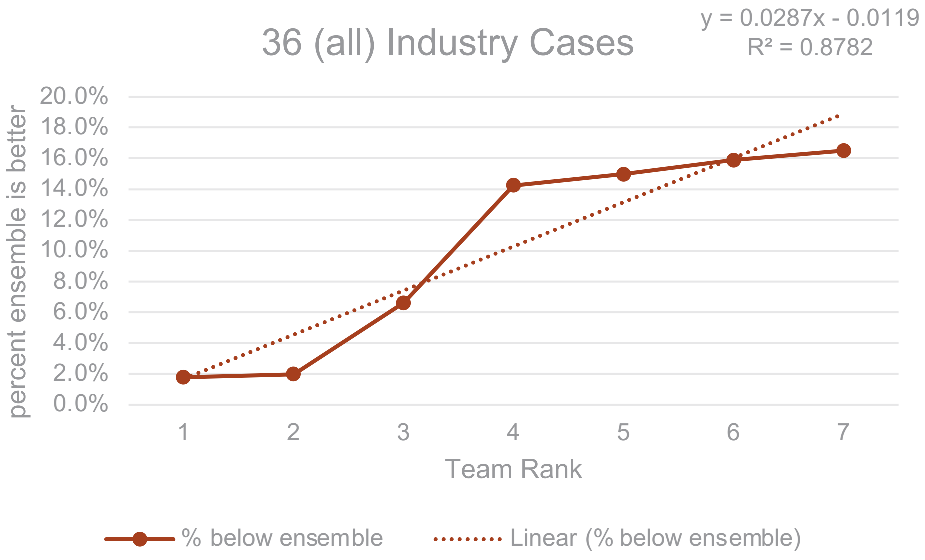
- Synthetic datasets: Correlation is 0.35-0.5 when H_{prior} is small.
- Offshoot: a priori difficulty is harder to measure when there is a large opportunity to improve the solution.

Synthetic and Industry Dataset Differences

Comparison of individual team performance to ensemble (best score from each scenario)



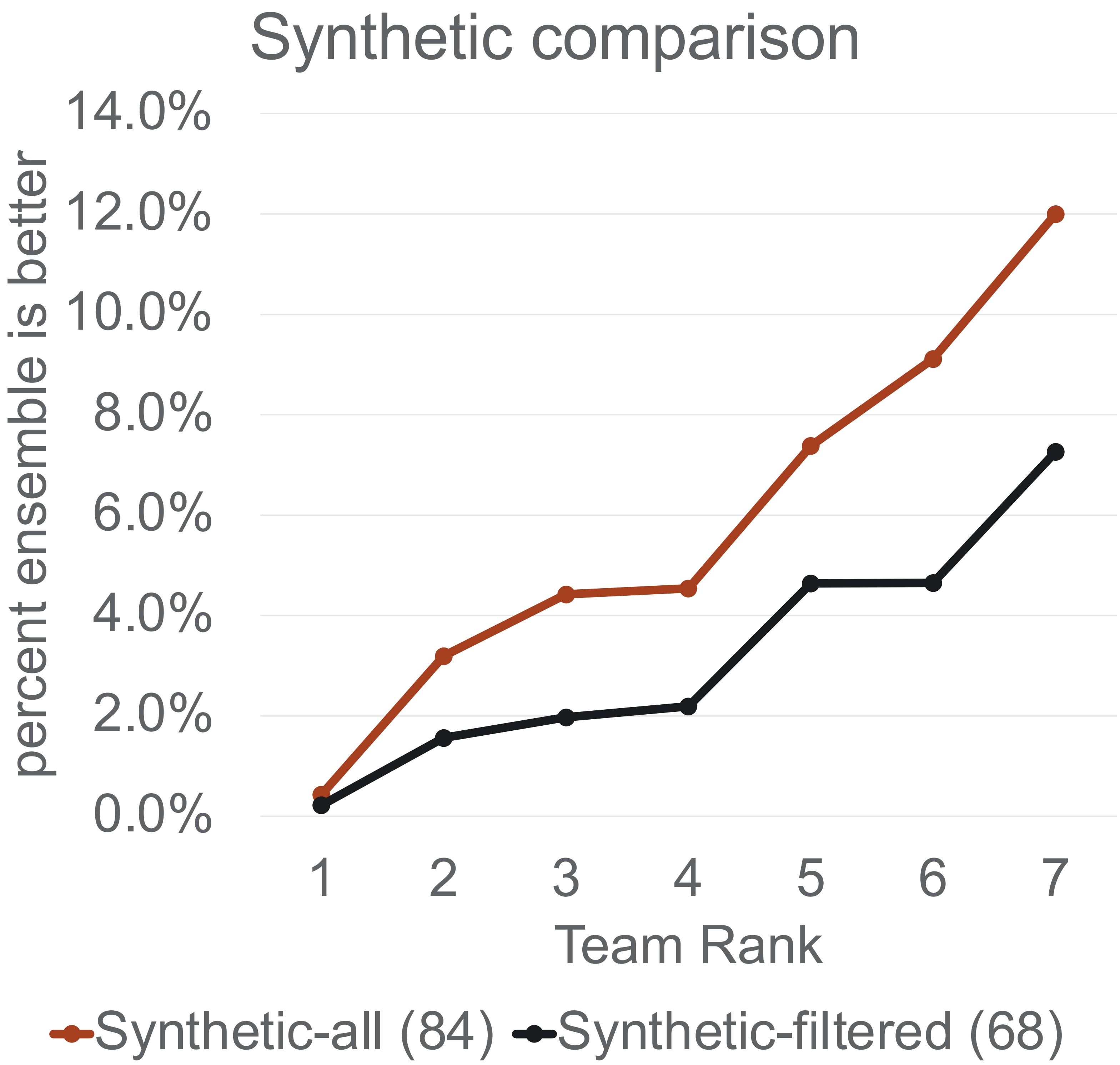




- All & synthetic networks: smooth linear score increase compared to the ensemble.
- Industry: only 3 teams within 12% of the ensemble score.
- This raises a question of whether industry case difficulties were caused by coding errors (e.g., parsing) or because the optimization problems were inherently more difficult.

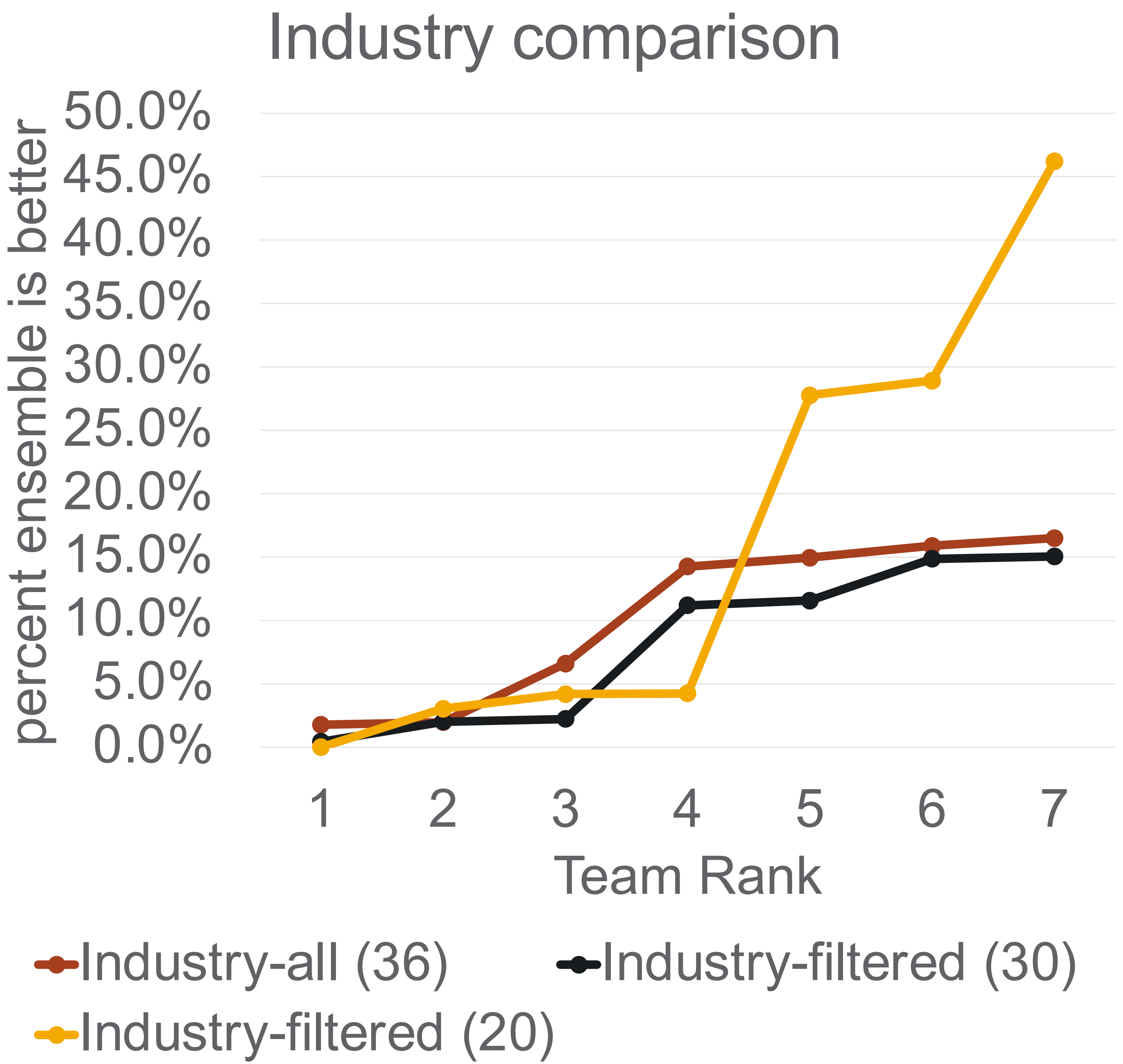
Synthetic case comparison

- Excluded 16 cases that caused failures in top-performing solvers.
- Result: more consistency between scores.
- This means that the excluded cases were difficult for all teams.



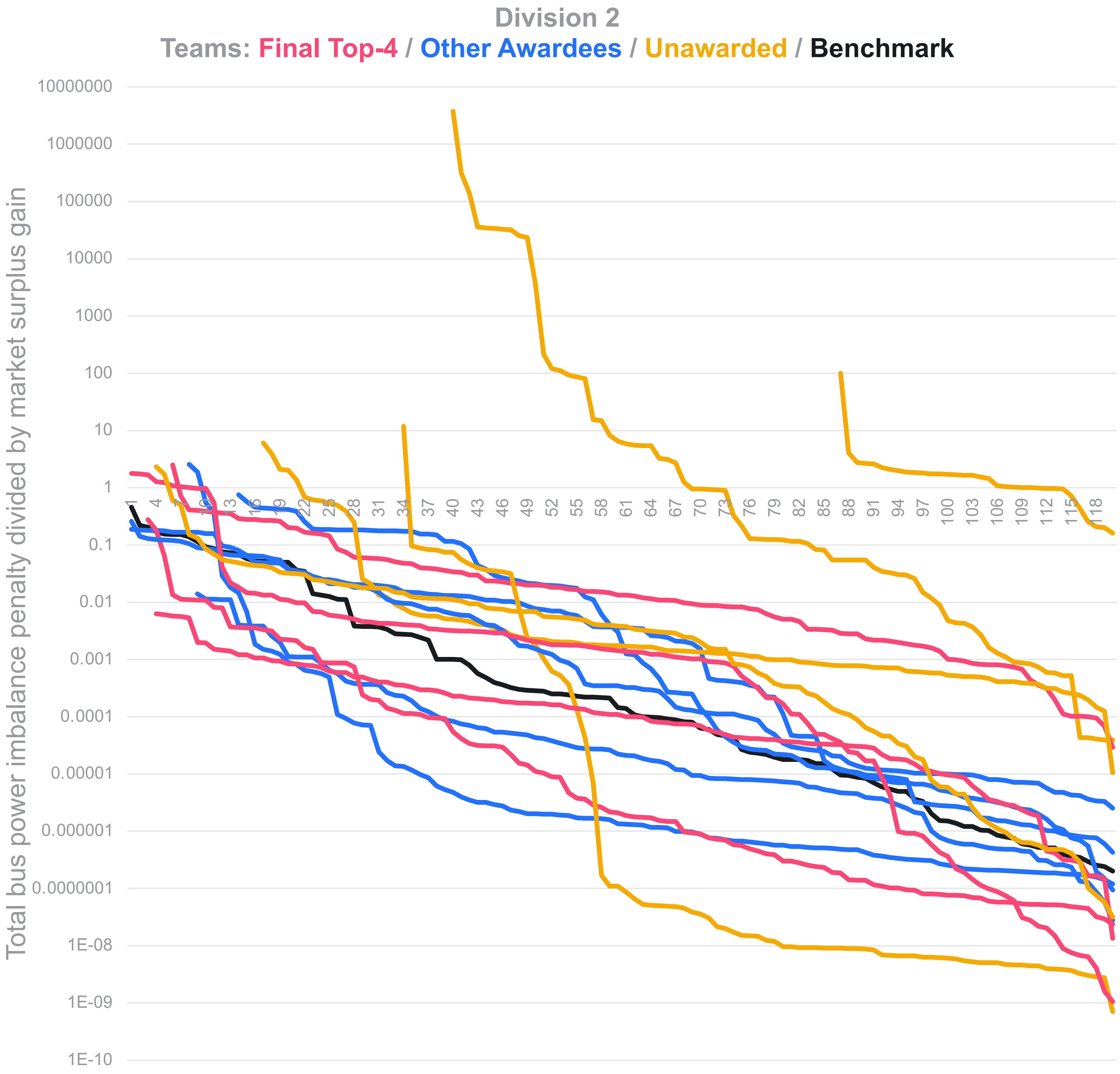
Industry case comparison

- **Filter 1**: remove 6 cases w/ failures from 4 teams
  - Curve moves lower, so all teams had difficulty.
- **Filter 2**: remove all 10 French network cases
  - Many teams perform much worse.
  - French network failures were specific to one team



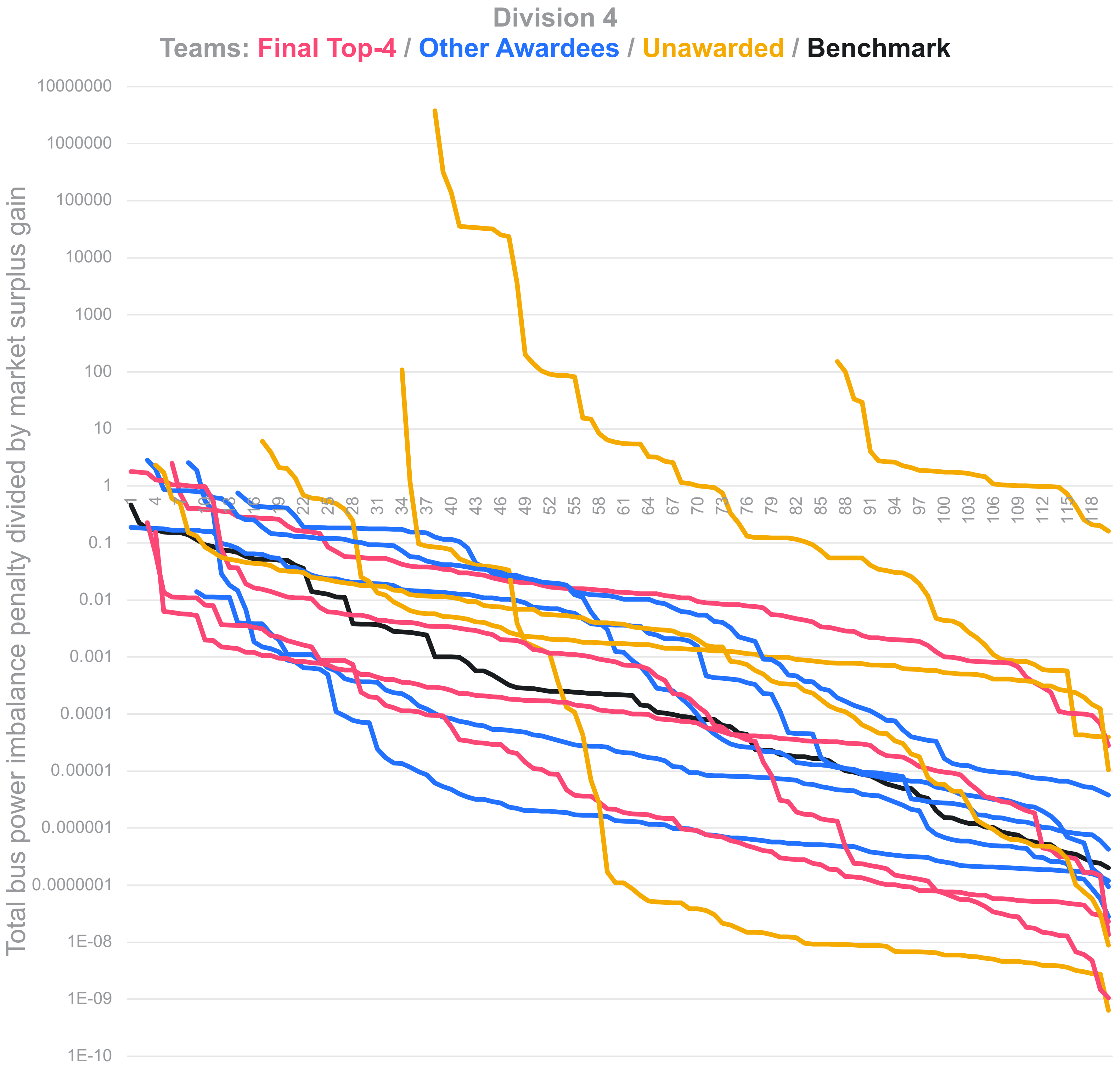
Effect of penalties

- Plotted on right: bus imbalance penalties relative to MS gain score.
  - Sorted in decreasing order
  - “Missing” data (rightward shift) for solutions replaced by MSpp
- Imbalance penalties had very little influence among top teams.
- Improving prior point solution was very important to win prize money.



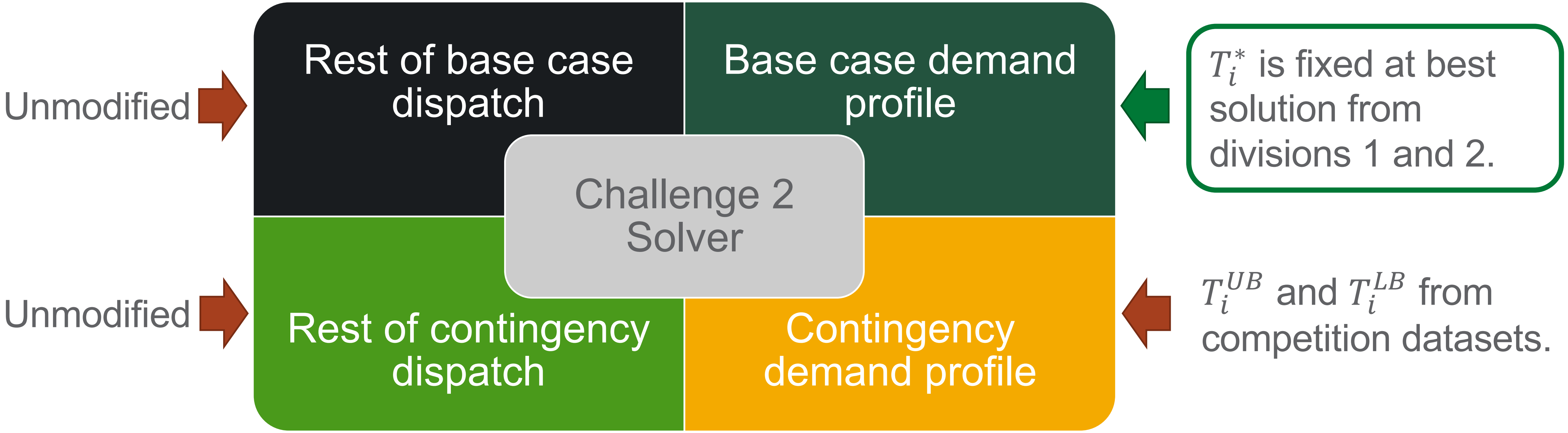
Effect of penalties

- Plotted on right: bus imbalance penalties relative to MS gain score.
  - Sorted in decreasing order
  - “Missing” data (rightward shift) for solutions replaced by MSpp
- Imbalance penalties had very little influence among top teams.
- Improving prior point solution was very important to win prize money.

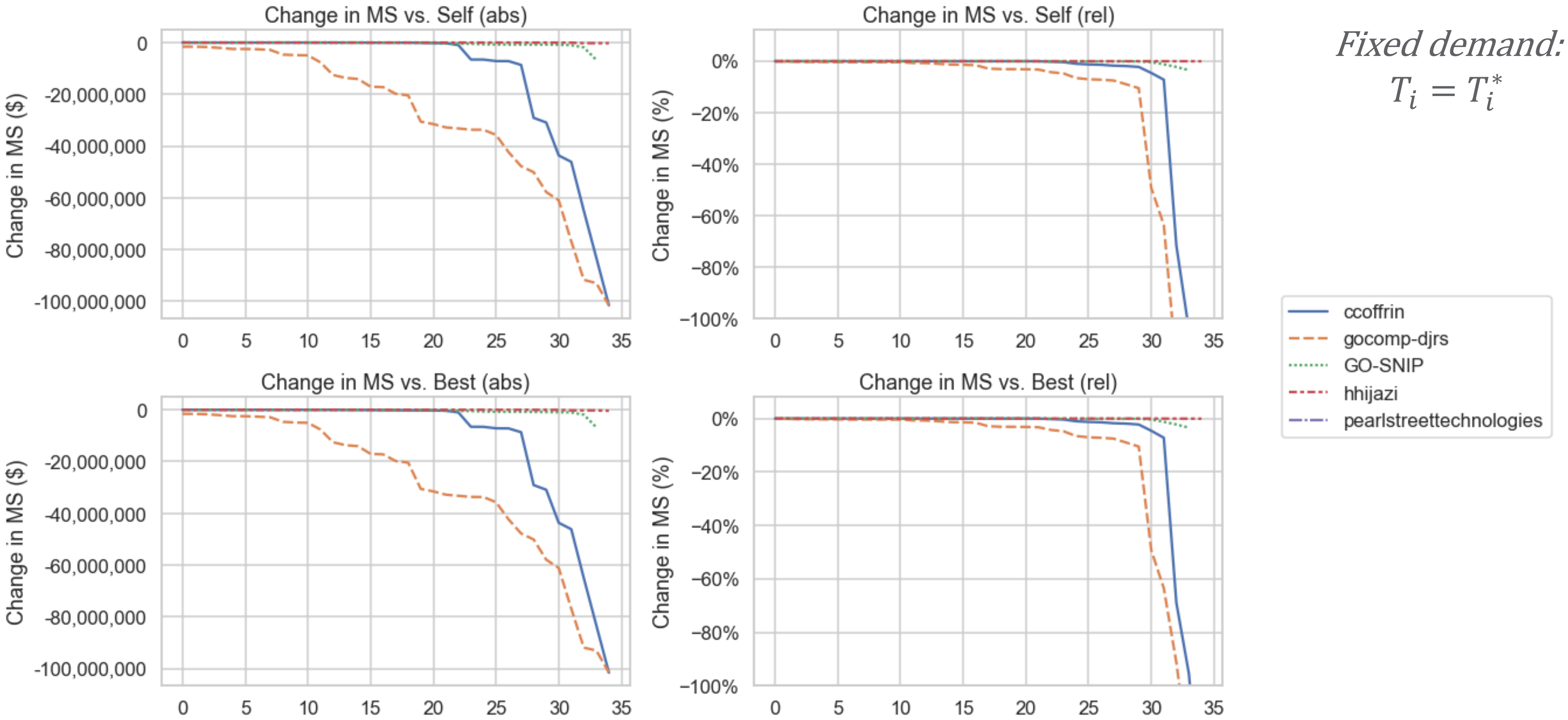


Effect of flexible demand on solution quality

- Solvers provided optimal load profiles given the value of bid-in demand.
  - Could this “optimal” load profile be fixed and then improved by other solvers?
  - Or, is load flexibility a necessary part of the optimization routines?
- Experimental set up was complicated due to Challenge 2 formulation, because the same upper and lower bounds are applied to pre- and post-contingency demand constraints.
- Load flexibility played an important role in managing contingencies, since almost all contingency constraints could be solved with post-contingency load curtailments.



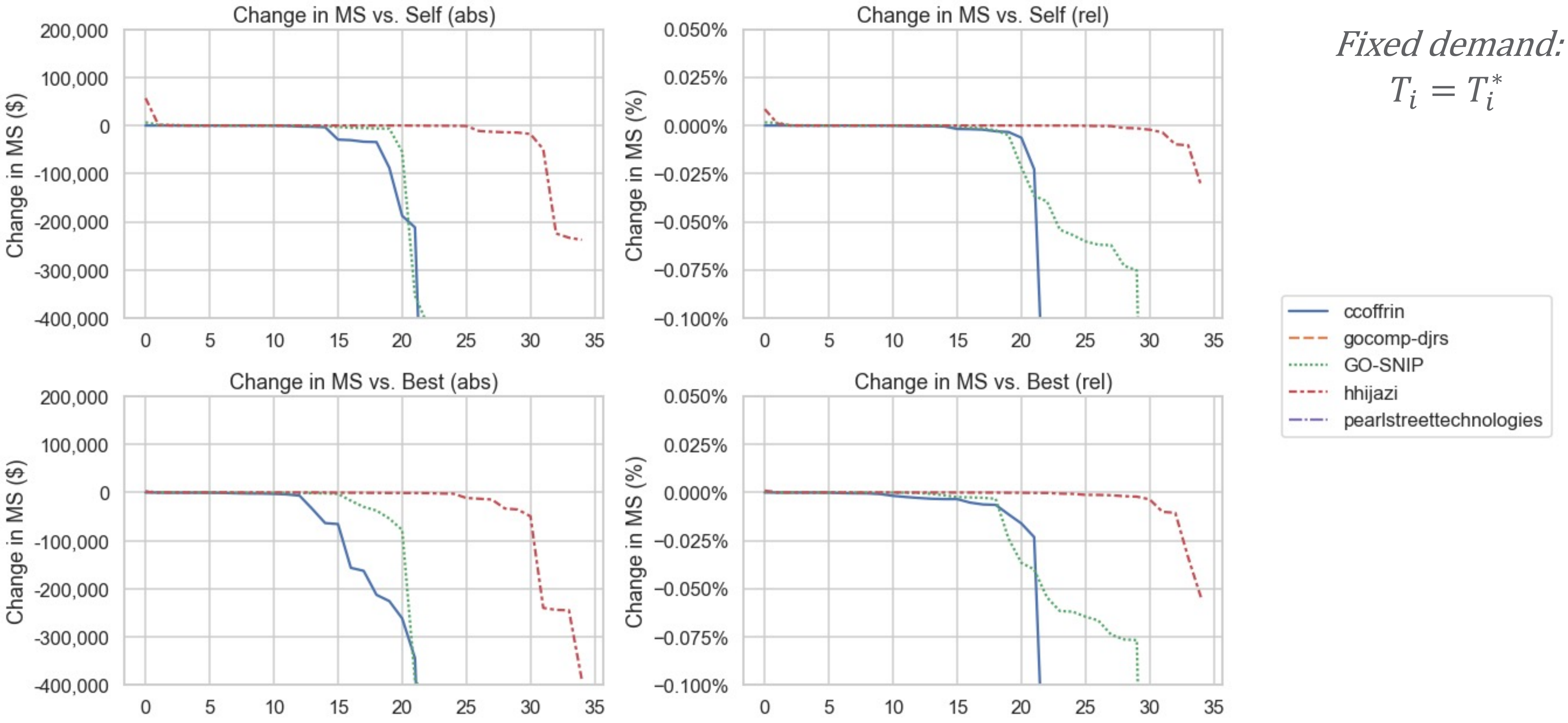
Many solvers had difficulty incorporating fixed loads during ex post analysis.



Fixed demand:

$$T_i = T_i^*$$

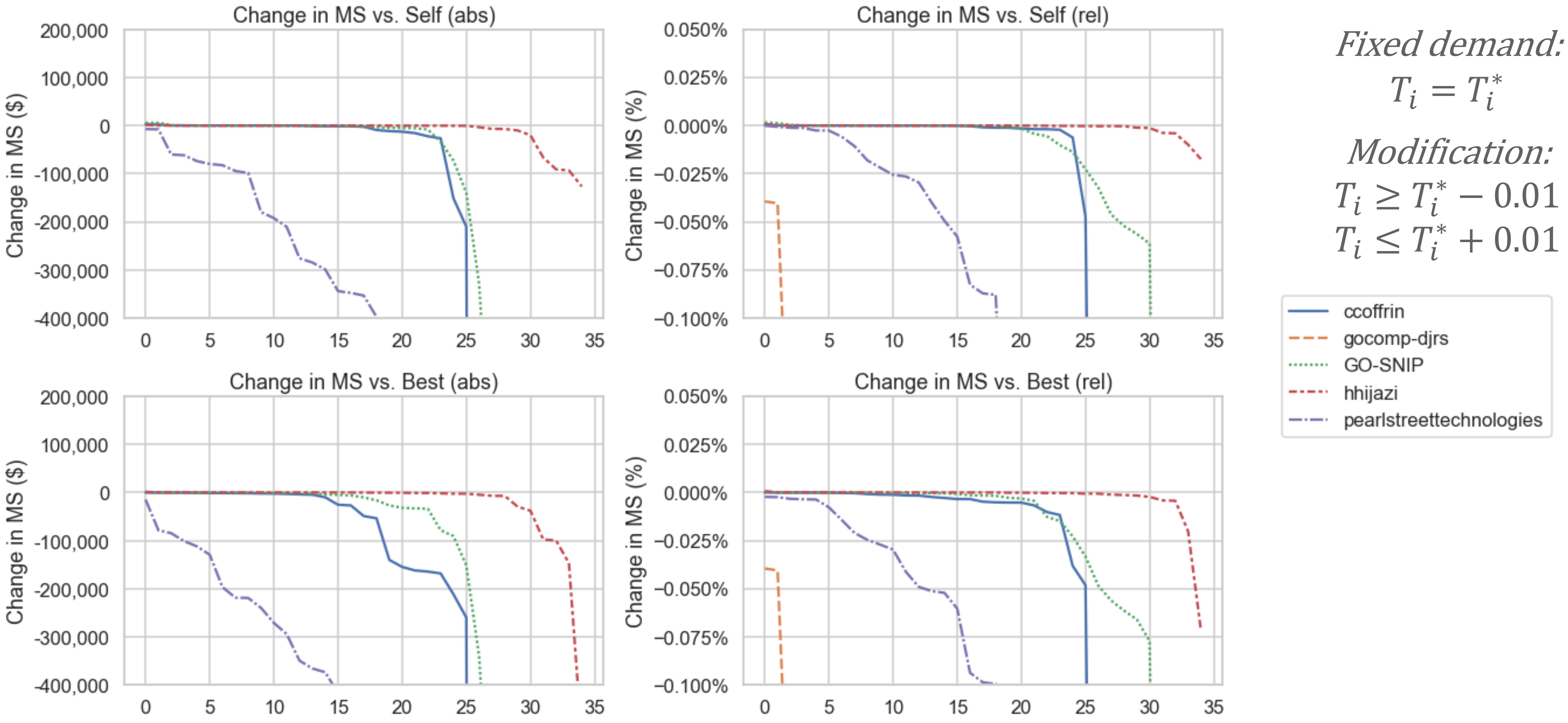
Rarely, fixed demand could improve an individual score but not the best score.



Fixed demand:

$$T_i = T_i^*$$

Modification with ± 0.01 feasibility neighborhood: Still no improvements.



Fixed demand:

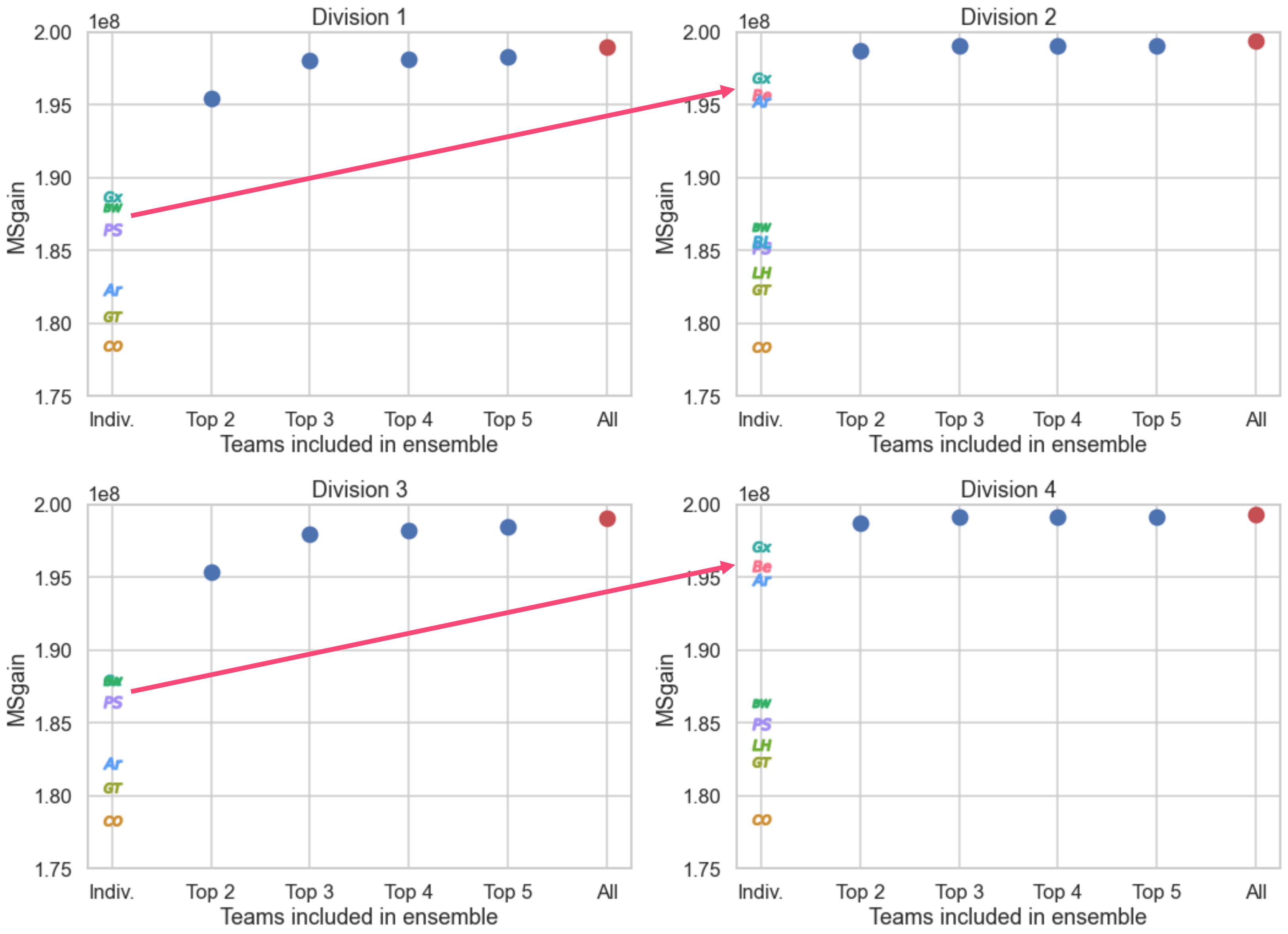
$$T_i = T_i^*$$

Modification:

$$T_i \geq T_i^* - 0.01$$
$$T_i \leq T_i^* + 0.01$$

Ensemble scores of the top 2/3/4/5 teams

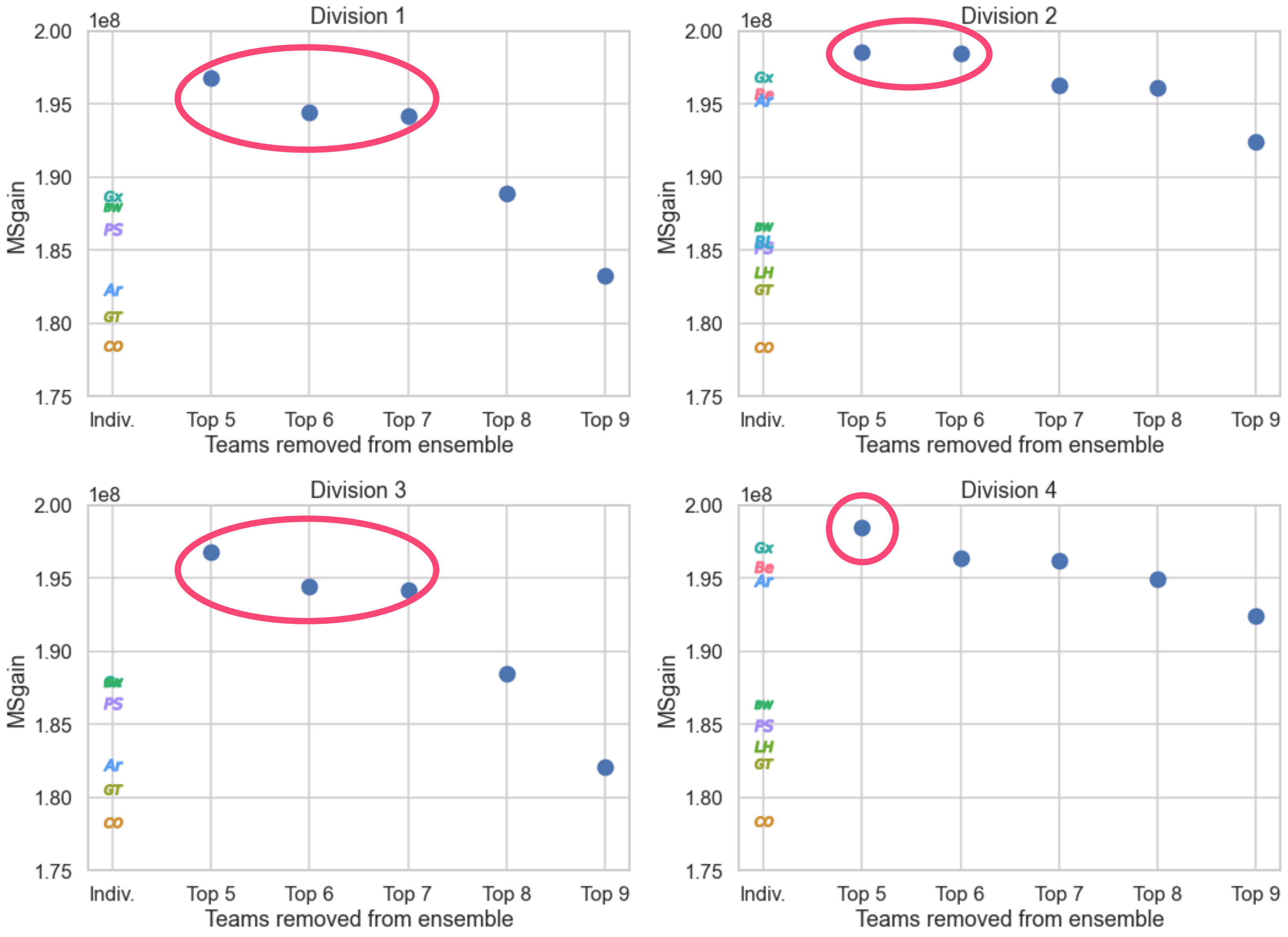
- The overall ensemble score is almost equal in each division.
- But individual teams still benefitted from additional time in Div. 2 and 4.
- Algorithm approach is more important when time is more limited.



Pacific Northwest
NATIONAL LABORATORY

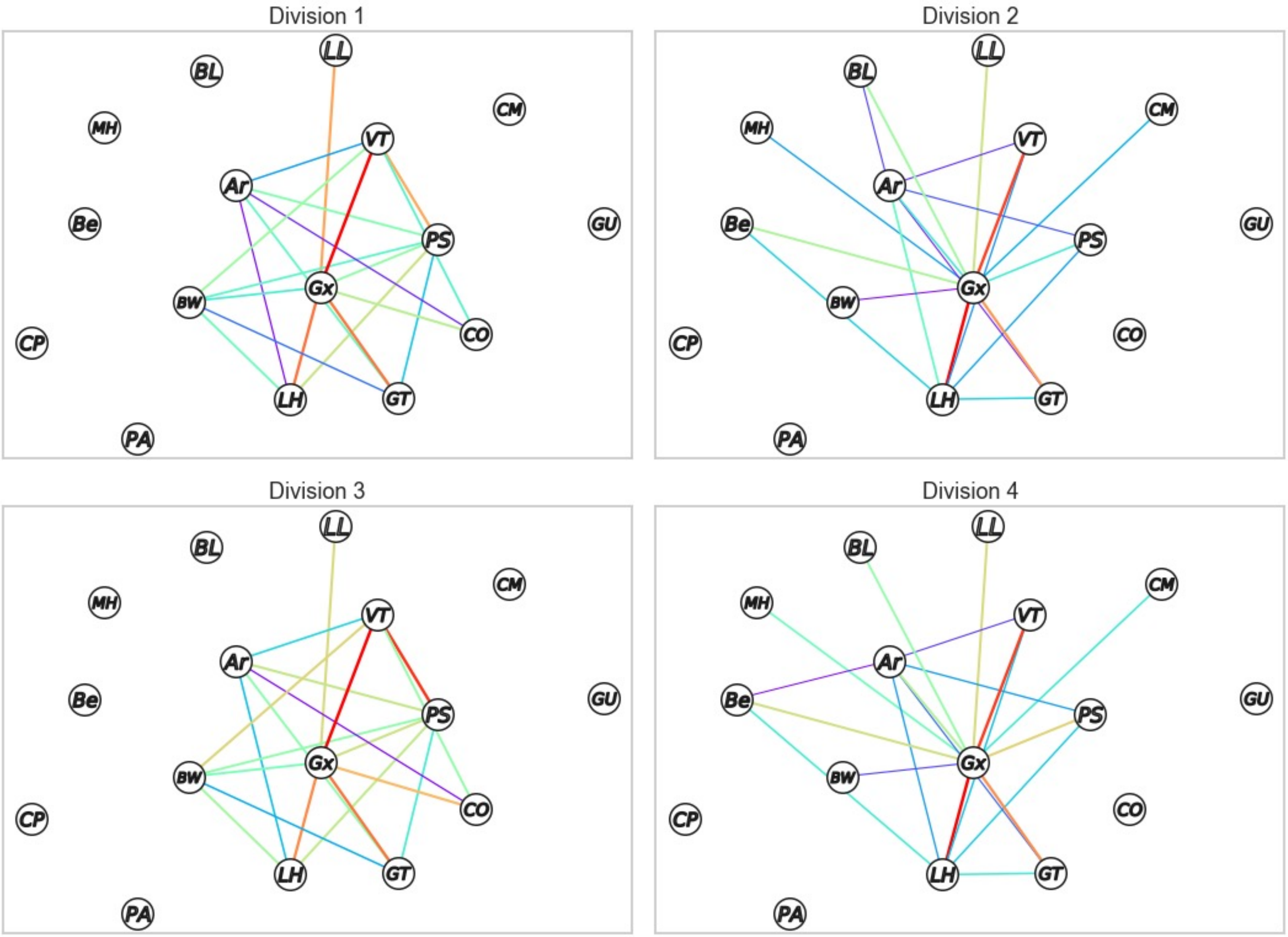
Ensemble scores with top teams removed

- An ensemble without any of the top 5 teams still would have won each division.
- Lower scoring teams were still able to find high-quality solutions, but not consistently.



Pairwise ensembles (best in parallel)

- The top-scoring teams almost always scored high on the same cases.
  - Almost no benefit to combining solutions from the top teams
- Best pairings are GravityX with one of the VaTech, GaTech, or Lehigh teams



Twenty highest-scoring pairs are colored in (red=highest).

Transmission switching results

- Overall, very few attempts
  - Only five of the 16 teams attempted transmission switching.
  - Total 205 solutions with switches, but only 18 with solution improvement > 1%.
  - High risk, low reward: ***reduced*** teams' Division 4 scores in most (57%) of the attempts.
- Most switching improvements occurred in networks with < 10,000 busses.

Team	(a) # cases with switches	(b) # base switches	(c) # contin. switches	(d) (D4-D2)/D2 > 0.01	(d ÷ a) % success
Benchmark	19	16	0	5	26.3%
Bigwood	5	3	0	1	20.0%
LBNL	90	171	19,825	7	7.8%
Artelys	88	370	14,986	5	5.7%
GERS	3	0	~40,000	0	0.0%

Impact of high-performance computing resources

- Most teams used full HPC resources.
  - 6 nodes w/ 144 cores. Exceptions:
  - Bigwood: 1 node / 24 cores
  - Colorado: 3 nodes / 72 cores
  - VaTech: 6 nodes / 24 cores
- No apparent advantage or disadvantage to HPC.

Team Name	Total Prize, FE+T3 ($k)	# nodes	# cores
GravityX	730	6	144
NU_Columbia_Artelys	530	6	144
GOT-BSI-OPF (Bigwood)	420	1	24
Pearl Street Technologies	340	6	144
Electric Stampede (Colo.)	140	3	72
GMI-GO (GaTech)	120	6	144
Monday Mornings (LBNL)	60	6	144
GO-SNIP (Lehigh)	30	6	144
Gordian Knot (VaTech)	30	6	24
ARPA-e Benchmark	N/A	6	144

Conclusions

- GO Competition depends on selecting appropriately difficult cases.
  - a priori hardness metrics are challenging, especially for cases with large potential improvement.
  - Safdarian *et al.* describes industry/synthetic difference in more detail.
- No evidence that penalties were a major influence in final scores/rankings.
- Flexible demand was necessary for solver performance.
- Shorter time limits led to more competition between solvers.
  - Bottom-ranked ensembles outperformed each individual team.
  - Best pairwise ensembles came from top individual team + mid-ranked teams.
- Few attempts at transmission switching.
  - Inclusive: problem too hard? network dependence? did cases reflect real-world?

Pacific Northwest
NATIONAL LABORATORY

Thank you