# De-Risking Solutions to Optimization Problems

Daniel Bienstock[1] and **Blake Sisson**[1]

[1]Columbia University

## METHODOLOGY

**Nominal problem**

$$\begin{aligned} \min \quad & c(x) \\ s.t. \quad & x \in P \end{aligned}$$

- from practical domains
- feat. integrality, nonlinearity, nonconvexity

Risk: $x \rightarrow$ risk exposure $\Phi(x)$

- different units than cost $c(x)$.
- can incorporate uncertainty.

**Practical stakeholders' perspective**

business goals (nominal problem) > risk containment.

$\Rightarrow$ *corrective* goal: find new solution $\hat{x} \in P$ with
- $\Phi(\hat{x}) << \Phi(x^*) \leftarrow$ much less risk,
- $c(\hat{x}) \gtrapprox c(x^*) \leftarrow$ only slightly more cost.

$\Rightarrow$ instead of $\min_{x \in P} c(x) \; s.t. \; \Phi(x) \leq U$, use:

$$\begin{aligned} \min \quad & c(x) + \Theta\Phi(x) \\ s.t. \quad & x \in P \end{aligned}$$

**Our approach**

$$\begin{aligned} \min \quad & c(x) + \Theta\phi_L \\ s.t. \quad & x \in P \\ & \phi_L \geq \max_{z \in \mathcal{Z}} \max_{i \in I} \phi_i(x|z) \end{aligned}$$

- $\Theta > 0$ risk-aversion parameter.
- $I$: risky "features"
- $\mathcal{Z}$: uncertainty parameterization

Solve probe efficient frontier via cutting-planes.

## ALGORITHM

**SOFTMAX-ADVERSARIAL**

**Inputs:** $\Theta > 0$; $\alpha > 0$; $\Delta, \Delta' > 0$.
Set $t = 0$, and initialize MASTER problem

$$\min \; c(x) + \Theta\,\phi_L \quad \text{s.t.} \quad x \in P, \;\; \phi_L \geq 0.$$

**while** $t < t^{\max}$ **do**:
  Solve MASTER problem; obtain optimal solution $(x^t, \phi_L^t)$.
  Compute $z^t \in \mathcal{Z}$ by solving (within tolerance $\Delta'$)

$$\max_{z \in \mathcal{Z}} \ln \sum_{i \in I} e^{\alpha\phi_i(x^t|z)}$$

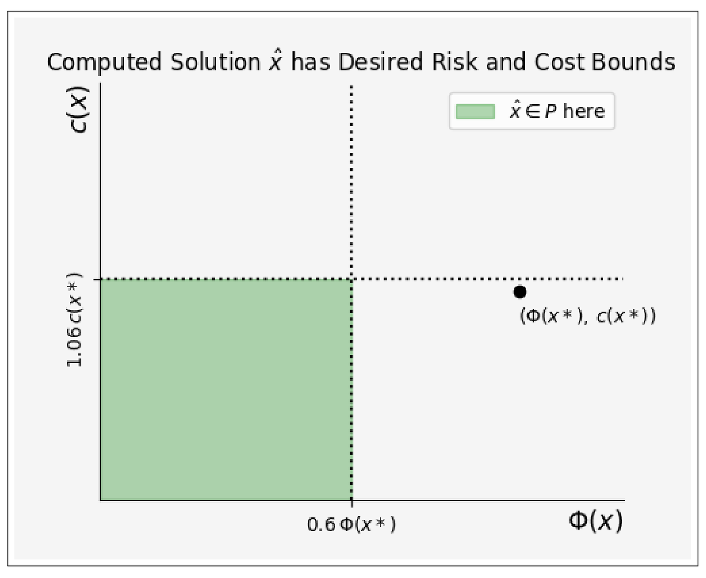
  **if** $\phi_{\max}^t \doteq \max_{i \in I} \phi_i(x^t|z^t) \leq \phi_L^t + \Delta$ **then STOP**.
  **else**
    Define $\pi_i^t = \dfrac{e^{\alpha\phi_i(x^t|z^t)}}{\sum_{j \in I} e^{\alpha\phi_j(x^t|z^t)}}$ for $i \in I$.
    Add cut $\phi_L \geq \sum_{i \in I} \pi_i^t \phi_i(x|z^t)$ to MASTER problem.
  **end if**
  $t \leftarrow t + 1$.
**end while**

**Theorem**: Algorithm terminates finitely or we reach an iteration $t$ with

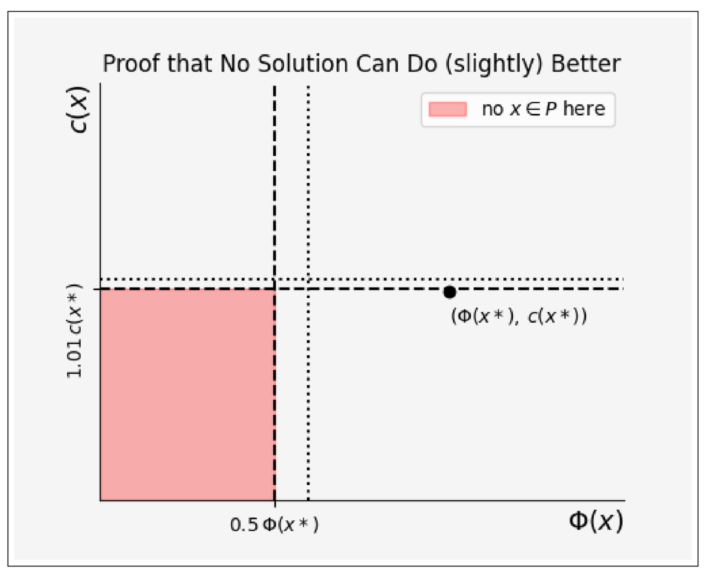
$$\phi_{\max}(x^t|z^t) \leq \lambda\phi_{\max}(x^0|z^0)$$

- $\lambda \in (0, 1)$ preselected.
- Requires $\phi_i$ uniformly continuous over $\mathcal{Z}$.
- Lemma: For non-terminal $t$, new cut is violated by at least $\Delta/4$.
- Convergence rate? TBD

**But wait, there's more!** We can provide guarantees every run*.




or

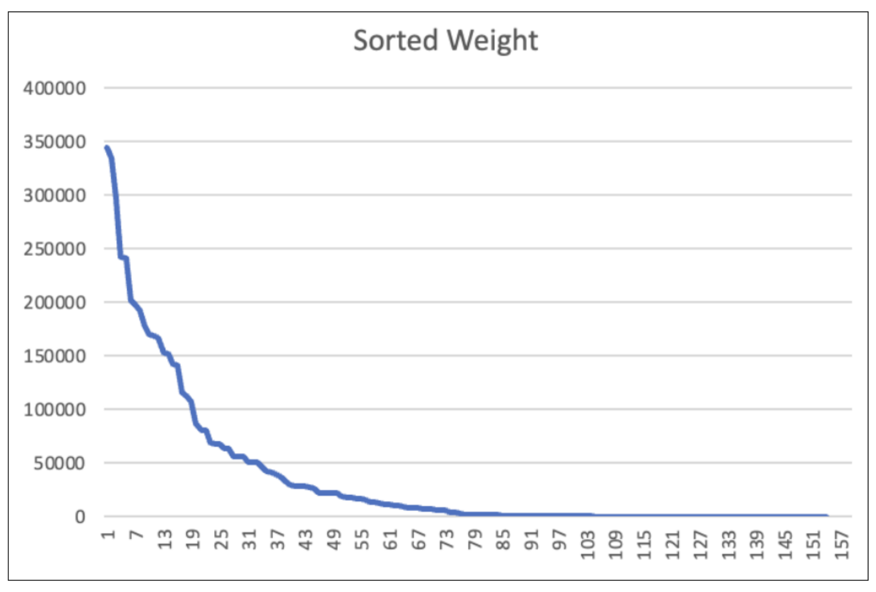



*Value of bounds tuned via $\Theta$.

## EXPERIMENTATION

### Logistics MIP with Stakeholders[1]

- 44 time periods, 17 locations, 224 links, 94 commodity types, 6 vehicles types.
- 400k+ variables (5k+ integral), 100k+ constraints.
  - $M_{\ell,\ell',t,j,h}$: # items of type $j$ shipped $\ell \rightarrow \ell'$ at time $t$ using vehicle of type $h$.




**Risk**: shipments can be interdicted.

- $I$: set of link, time-periods
- $\mathcal{Z}$: disruption budget
- $\phi_i(x|z)$: cost increase of link $i$ for logistics plan $x$ under disruption budget $z$.

**Implementation**:

- MASTER problem: Gurobi for MIPs and relaxations.
- Adversarial problem: gradient descent (ADAM, Mutlistarts), heuristics.

| Iteration | Risk metric | Cost | Cuts added | Time (s) |
|---|---|---|---|---|
| 0 | 66724 | 1095 | | 25.62 |
| 1 | 53992 | 1179 | 132 | 28.84 |
| 2 | 39969 | 1153 | 196 | 30.59 |
| 3 | 50437 | 1153 | 216 | 30.61 |
| 4 | 50385 | 1153 | 229 | 30.77 |
| 5 | 44181 | 1156 | 234 | 29.13 |
| 6 | 44181 | 1156 | 234 | 28.05 |
| 7 | 50386 | 1156 | 235 | 27.74 |
| 8 | 44181 | 1186 | 235 | 91.65 |

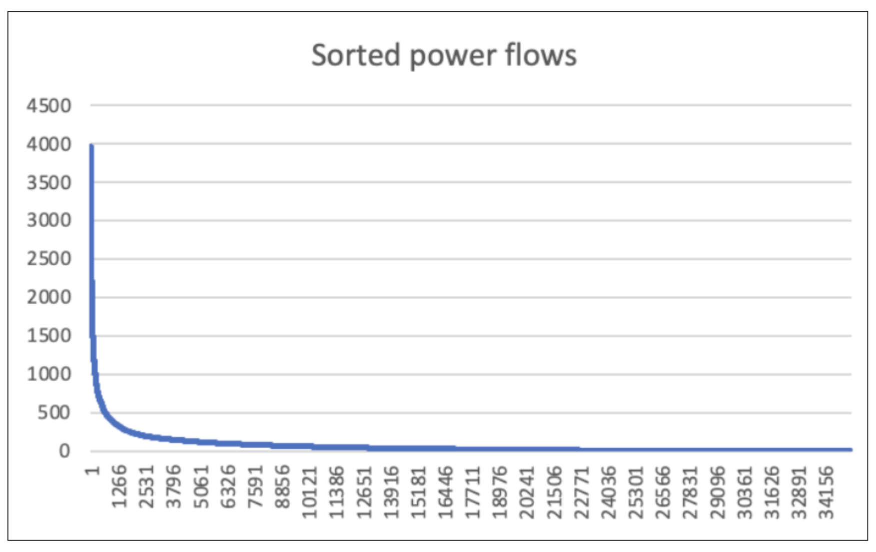
**Results**: 33% decrease in risk, 8.3% increase in cost.

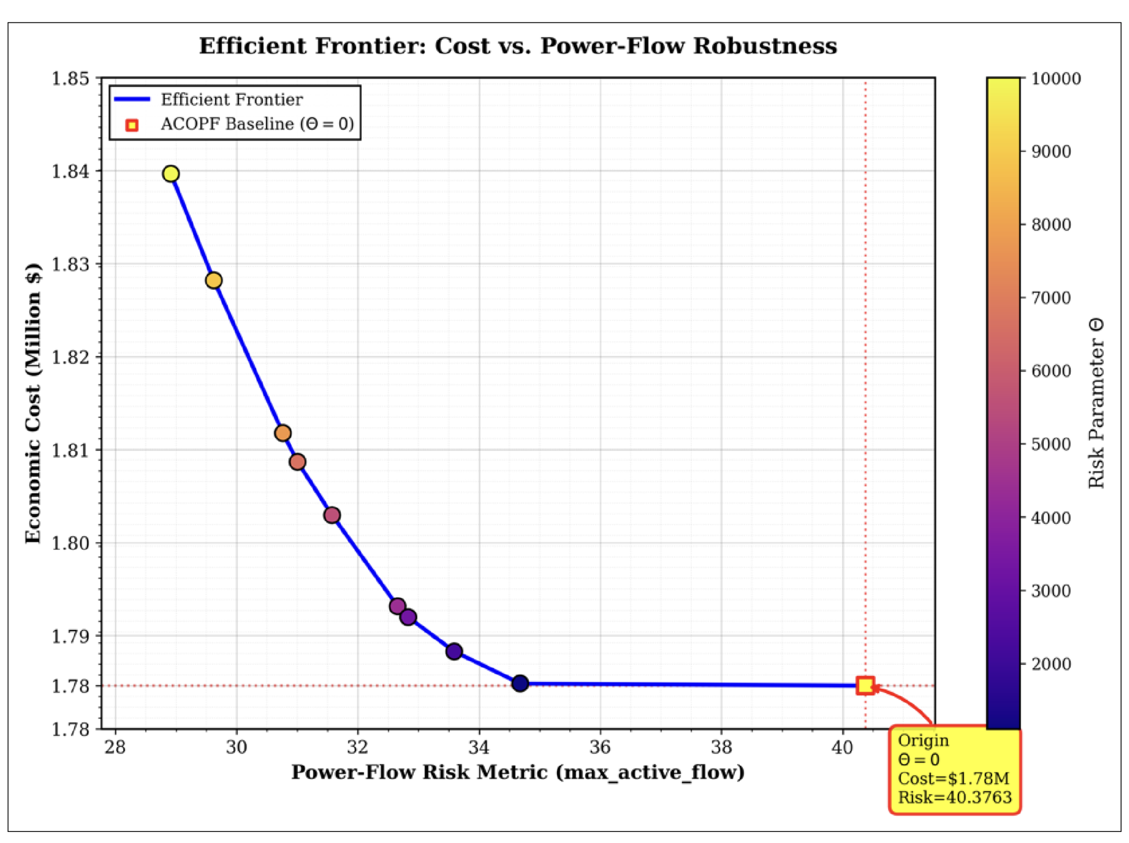
[1]J. Klib, D. Bienstock, A. Newman, Columbia University and Colorado School of Mines.

### Power Grids: case pglib_opf_case30000_goc[2]

- Minimize grid operating cost while meeting demand and obeying laws of physics.
- Nonlinear and nonconvex.
- 30k buses, 35k branches, 3.5k generators
  $\Rightarrow$ 220k variables and constraints.

**Risk**: line temperature: correlated with magnitude of power flow




**Implementation**[3]

- 5 its. per $\Theta$.
- ~20 greedy cuts per iteration.
- LP approximation during cut procedure.
- ~20s solve time per iteration.

[2]C. Coffrin et al., Los Alamos National Lab.
[3]R. Akinwonmi and A Newman, Colorado School of Mines.