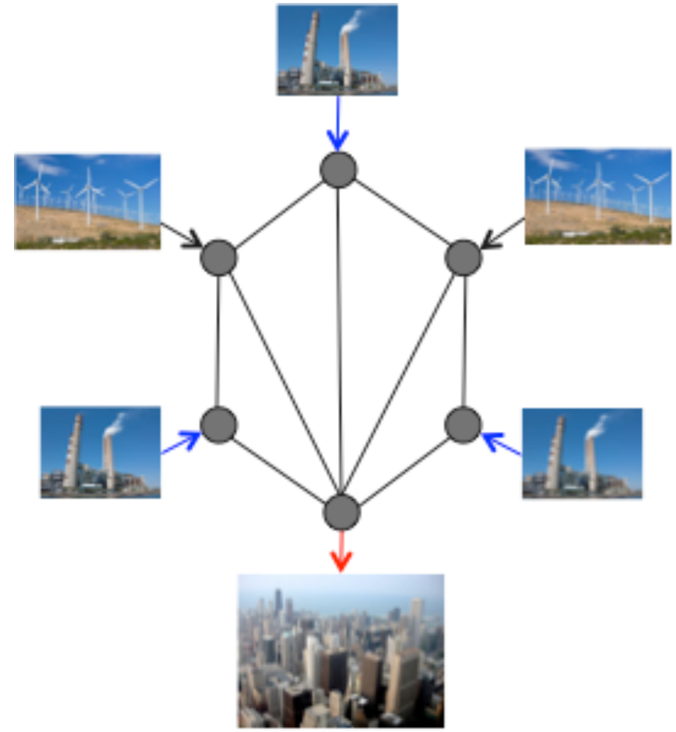
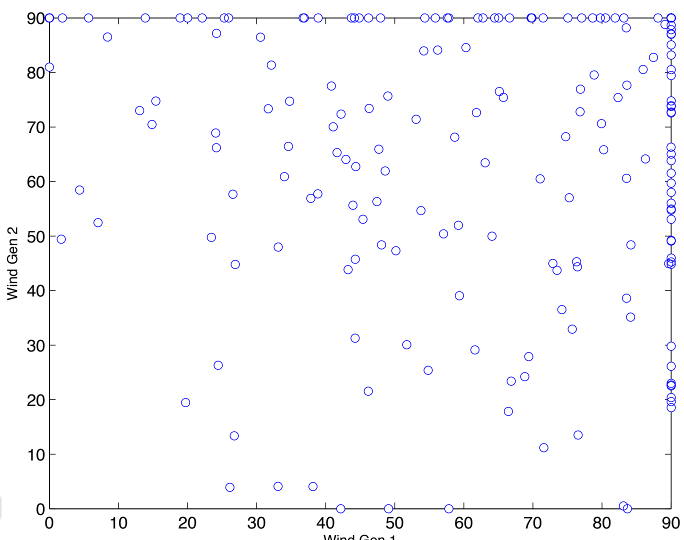
Clustering-Based Strategies for Stochastic Programs

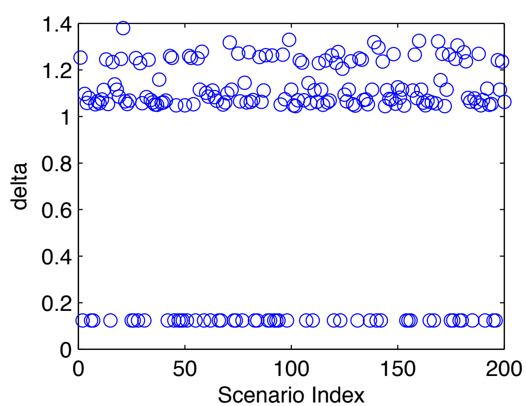
Victor M. Zavala
Assistant Computational Mathematician
Mathematics and Computer Science Division
Argonne National Laboratory
Fellow
Computation Institute
University of Chicago

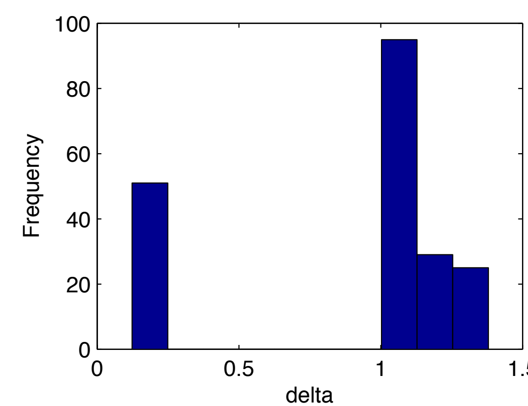
Scenario Clustering

$$
\begin{aligned}
\min \quad \mathbb{E} \Bigg[& \sum_{t\in\mathcal{T}} \sum_{i\in\mathcal{G}} \alpha_i^g g_{i,t} - \sum_{t\in\mathcal{T}} \sum_{i\in\mathcal{D}} \alpha_i^d d_{i,t} \\
& + \sum_{t\in\mathcal{T}} \sum_{i\in\mathcal{G}} \alpha_i^{g,+} (G_{i,t}(\omega) - g_{i,t})_+ - \alpha_i^{g,-} (G_{i,t}(\omega) - g_{i,t})_- \\
& - \sum_{t\in\mathcal{T}} \sum_{i\in\mathcal{D}} \alpha_i^{d,+} (D_{i,t}(\omega) - d_{i,t})_+ + \alpha_i^{d,-} (D_{i,t}(\omega) - d_{i,t})_- \Bigg] \\
\text{s.t.} \quad & \sum_{i\in\mathcal{L}_j^{rec}} f_{i,t} - \sum_{i\in\mathcal{L}_j^{snd}} f_{i,t} + \sum_{i\in\mathcal{G}_j} g_{i,t} - \sum_{i\in\mathcal{D}_j} d_{i,t} = 0, \qquad t\in\mathcal{T}, j\in\mathcal{B} \\
& \sum_{i\in\mathcal{L}_j^{rec}} F_{i,t}(\omega) - \sum_{i\in\mathcal{L}_j^{snd}} F_{i,t}(\omega) + \sum_{i\in\mathcal{G}_j} G_{i,t}(\omega) - \sum_{i\in\mathcal{D}_j} D_{i,t}(\omega) = 0, \qquad t\in\mathcal{T}, j\in\mathcal{B} \\
& -\bar{f}_{i,t} \le f_{i,t} \le \bar{f}_{i,t}, \qquad t\in\mathcal{T}, i\in\mathcal{L} \\
& 0 \le g_{i,t} \le \bar{g}_{i,t}, \qquad t\in\mathcal{T}, i\in\mathcal{G} \\
& 0 \le d_{i,t} \le \bar{d}_{i,t}, \qquad t\in\mathcal{T}, i\in\mathcal{D} \\
& -\bar{F}_{i,t}(\omega) \le F_{i,t}(\omega) \le \bar{F}_{i,t}(\omega), \qquad t\in\mathcal{T}, i\in\mathcal{L} \\
& 0 \le G_{i,t}(\omega) \le \bar{G}_{i,t}(\omega), \qquad t\in\mathcal{T}, i\in\mathcal{G} \\
& 0 \le D_{i,t}(\omega) \le \bar{D}_{i,t}(\omega), \qquad t\in\mathcal{T}, i\in\mathcal{D}
\end{aligned}
$$









Setting and Motivation

Two-Stage Convex Stochastic Program

$$\min \left(\frac{1}{2}y_0^T Q_0 y_0 + d_0^T y_0\right) + S^{-1}\sum_{s\in\mathcal{S}}\left(\frac{1}{2}y_s^T Q_s y_s + d_s^T y_s\right)$$

$$\begin{aligned}
\text{s.t.} \quad W_0 y_0 &= b_0, \ (\lambda_0) \\
T_s y_0 + W_s y_s &= b_s, \ \ (\lambda_s), \ s\in\mathcal{S} \\
y_0 &\geq 0, \quad (\nu_0) \\
y_s &\geq 0, \quad (\nu_s), \ s\in\mathcal{S}.
\end{aligned}$$

Interior Point : KKT Conditions

$$\begin{aligned}
\nabla_{y_0}\mathcal{L} = 0 &= Q_0 y_0 + d_0 + A_0^T\lambda_0 - \nu_0 + S^{-1}\sum T_s^T\lambda_s \\
\nabla_{y_s}\mathcal{L} = 0 &= Q_s y_s + d_s + W_s^T\lambda_s - \nu_s, \ s\in\mathcal{S} \\
\nabla_{\lambda_0}\mathcal{L} = 0 &= W_0 y_0 - b_0 \\
\nabla_{\lambda_s}\mathcal{L} = 0 &= T_s y_0 + W_s y_s - b_s, \ s\in\mathcal{S} \\
0 &= Y_0 V_0 e - \mu \\
0 &= Y_s V_s e - \mu, \ s\in\mathcal{S},
\end{aligned}$$

Questions:

- **How to Address Scenario Complexity?**
- **How to Address First-Stage Complexity?**

Interior Point : Newton Step & Block Representation

$$\begin{bmatrix} K_1 & & & & B_1 \\ & K_2 & & & B_2 \\ & & \ddots & & \vdots \\ & & & K_S & B_S \\ B_1^T & B_2^T & \dots & B_S^T & K_0 \end{bmatrix} \begin{bmatrix} \Delta w_1 \\ \Delta w_2 \\ \vdots \\ \Delta w_S \\ \Delta w_0 \end{bmatrix} = \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_S \\ r_0 \end{bmatrix},$$

$$\begin{bmatrix} K_{\mathcal{S}} & B_{\mathcal{S}} \\ B_{\mathcal{S}}^T & K_0 \end{bmatrix} \begin{bmatrix} \Delta w_{\mathcal{S}} \\ \Delta w_0 \end{bmatrix} = \begin{bmatrix} r_{\mathcal{S}} \\ r_0 \end{bmatrix}.$$

Addressing Scenario Complexity

Outside-the-Solver Scenario Clustering (Aggregation, Compression)

- **Key:** Cluster "Similar" Scenarios Based on Data Realizations
- *Birge 1985, Shetty 1987, Gondzio 2004, Sen 2005, Roemisch 2009, Sagastizabal 2010 ...*
- **Advantages**
  - Eliminates Redundancies and is Non-Intrusive
- **Issues**
  - Cannot Guarantee Convergence/Feasibility of Original Problem (e.g, Only Bounds)
  - Refining Clusters Iteratively Has Limited Efficiency (e.g., Reuse Basis, Warm-start)
  - Compression Expensive (e.g., Matrix Aggregation, Tree Exploration, Needs Distribution)

Inside-the-Solver Scenario Elimination by Monte-Carlo Sampling

- **Key:** Use Small Sample Set to Compute Step or Precondition *Nocedal 2011,2012, Anitescu 2011*

$$\begin{bmatrix} K_{\mathcal{S}} & B_{\mathcal{S}} \\ B_{\mathcal{S}}^T & K_0 \end{bmatrix} \begin{bmatrix} \Delta w_{\mathcal{S}} \\ \Delta w_0 \end{bmatrix} = \begin{bmatrix} r_{\mathcal{S}} \\ r_0 \end{bmatrix}.$$

$$\begin{bmatrix} K_{\mathcal{R}} & B_{\mathcal{R}} \\ B_{\mathcal{R}}^T & K_0 \end{bmatrix} \begin{bmatrix} \Delta w_{\mathcal{R}} \\ \Delta w_0 \end{bmatrix} = \begin{bmatrix} r_{\mathcal{E}} \\ r_0 \end{bmatrix}.$$

$$\Delta w_{\mathcal{E}} = K_{\mathcal{E}}^{-1}(r_{\mathcal{E}} - K_0 \Delta w_0).$$

Inside-the-Solver Scenario Elimination by Numerical Thresholds

- **Key:** Some Scenarios Have "Less" Impact on First-stage (e.g., "Inactive") *Tits 2006, Grothey 2012*
- **Advantages**
  - Adaptive (Efficient) Along Search
  - Convergence to Original Problem
- **Issues:**
  - Numerical Thresholds Needs Activity Resolution
  - Monte Carlo Sampling Can Induce Large Residuals (Ignore an "Active" Scenario)
  - Limited Compression (Does Not Eliminate "Large" Contributions & Redundancies)

Inside-the-Solver Scenario Clustering (Proposal)

$$(K_0 - \sum W_s^T K_s^{-1} W_s)\Delta w_0 = r_0 - \sum W_s^T K_s^{-1} r_s$$

Schur System

Key Idea:

- **Cluster Scenarios Inducing "Similar" Actions Along First-Stage Step**

$$\gamma_s(\Delta w_0) := W_s^T K_s^{-1} W_s \Delta w_0, \quad s \in \mathcal{S}$$

- **Find Scenario Partition That Approximately Minimizes Distortion Metric**

$$J(r_{s,i}, \bar{\gamma}_i, \gamma_s(u)) := \sum_{s \in \mathcal{S}} \sum_{i \in \mathcal{C}} r_{s,i} \|\bar{\gamma}_i - \gamma_s(u)\|$$

- **Heuristic: Applying a Clustering Scheme to Vectors** $\gamma_s(u), s \in \mathcal{S}$ **and Taking** $u = \Delta w_0^{k-1}$

Residual Characterization:

- **Clustering Leads to the Compressed Schur System**

$$(K_0 - \Omega_R B_{\mathcal{R}}^T K_{\mathcal{R}}^{-1} B_{\mathcal{R}})\Delta w_0 = r_0 - B_{\mathcal{S}}^T K_{\mathcal{S}}^{-1} r_{\mathcal{S}}$$
$$\bar{Z}\Delta w_0 = r_Z$$

Cluster Weights

$$\Omega_R := \text{diag}(\omega_r I \,|\, r \in \mathcal{R})$$

- **Approximate Step Induces Residual on Full-Space Schur System:**

$$\begin{aligned}\delta_Z(\Delta w_0) &= Z\Delta w_0 - r_Z \\ &= (Z - \bar{Z})\Delta w_0 - r_Z + \bar{Z}\Delta w_0 \\ &= (Z - \bar{Z})\Delta w_0\end{aligned}$$

- **Result I:** $\|\delta_Z(\Delta w_0)\| \leq J(r_{s,i}, \bar{\gamma}_i, \gamma_s(\Delta w_0))$ **Minimize Residual by Minimizing Dist. Metric**
- **Result II: Superlinear Convergence Possible without Full Scenario Representation**

Addressing First-Stage Complexity: Sparse Compressed KKT Systems

Key: **Avoid Schur Complement Formation/Factorization (Dense)**

Result IIIa: First-Stage Step Delivered by Compressed Schur System is Equivalent to that of Sparse Compressed System:

$$\begin{bmatrix} \Omega_{\mathcal{R}}^{-1} K_{\mathcal{R}} & B_{\mathcal{R}} \\ B_{\mathcal{R}}^T & K_0 \end{bmatrix} \begin{bmatrix} \Delta w_{\mathcal{R}} \\ \Delta w_0 \end{bmatrix} = \begin{bmatrix} \Omega_{\mathcal{R}}^{-1} r_{\mathcal{R}} \\ r_0 - B_{\mathcal{E}}^T K_{\mathcal{E}}^{-1} r_{\mathcal{E}} \end{bmatrix},$$

Result IIIb: First-Stage Step Delivered is Equivalent to that of Perturbed Full-Space System:

$$\begin{bmatrix} K_{\mathcal{S}} & B_{\mathcal{S}} \\ B_{\mathcal{S}}^T & K_0 + P_{\mathcal{R}} \end{bmatrix} \begin{bmatrix} \Delta w_{\mathcal{S}} \\ \Delta w_0 \end{bmatrix} = \begin{bmatrix} r_{\mathcal{S}} \\ r_0 \end{bmatrix}$$

$$\Phi_w \Delta w = -\Phi \text{ (KKT Residual)}$$

$$\begin{aligned} P_{\mathcal{R}} &:= \Omega_{\mathcal{R}} B_{\mathcal{R}}^T K_{\mathcal{R}}^{-1} B_{\mathcal{R}} - B_{\mathcal{S}}^T K_{\mathcal{S}}^{-1} B_{\mathcal{S}} \\ &= (\bar{Z} - Z). \end{aligned}$$

Result IV: Full-Space Residual Induced by Compressed Sparse System:

$$\begin{aligned} \delta(\Delta w) &= \Phi_w \Delta w + \Phi \\ &= (\Phi_w - \bar{\Phi}_w)\Delta w - \Phi + \bar{\Phi}_w \Delta w \\ &= (\Phi_w - \bar{\Phi}_w)\Delta w \\ &= \begin{bmatrix} 0 & 0 \\ 0 & -P_{\mathcal{R}} \end{bmatrix} \begin{bmatrix} \Delta w_{\mathcal{S}} \\ \Delta w_0 \end{bmatrix} \\ &= \begin{bmatrix} 0 \\ (Z - \bar{Z})\Delta w_0 \end{bmatrix}. \end{aligned}$$

Inexact Interior Point Setting

Merit Function $\Xi(w^k) := \|\nabla_y \mathcal{L}(w^k)\| + \|\nabla_\lambda \mathcal{L}(w^k)\| + \|Y^k V^k e\|.$

Algorithm IP-CLUSTER. Given $\epsilon > 0$, η, τ, C^{min}, and ΔC, DO:

1. Compute $\Xi(w^k)$ and ℓ^k.
2. IF $\|\Xi(w^k)\| \leq \epsilon$ TERMINATE. Otherwise, CONTINUE.
3. Set number of clusters $C^k := C^{min}$.
4. Compute cluster information $\mathcal{R}^k$ and $\omega^k_{\mathcal{R}}$ using Δw^{k-1}.
5. Compute step Δw^k using sparse compressed system.
6. Determine maximum step size $\alpha \geq (0,1]$ satisfying $\Pi(w^k + \alpha \Delta w^k) \geq (1-\tau)\Pi w^k$ and compute trial iterate $w^k_+ \leftarrow w^k + \alpha \Delta w^k$.
7. IF $\Xi(w^k_+) \leq \eta\, \Xi(w^k)$ AND $\|\delta^k\| \leq \ell^k \Phi(w^k)$ ACCEPT trial step $w^{k+1} \leftarrow w^k_+$, UPDATE $\mu^{k+1} \leftarrow \mu^k + \Delta\mu^k$, SET $k \leftarrow k+1$, and RETURN TO 1). Otherwise, RETURN to 3), and set $C^k \leftarrow \min\left\{S, C^k + \Delta C\right\}$.

Clustering for Preconditioning

Use Sparse Compressed System to Precondition Full-Space System

- **Result V: The Preconditioning Error Delivered by <u>Sparse Compressed System</u> Satisfies**

$$
\begin{aligned}
y - y^* &= \Phi_w^{-1} P y \\
&= \Phi_w^{-1}(\Phi_w - \bar{\Phi}_w) y \\
&= (I - \Phi_w^{-1}\bar{\Phi}_w) y.
\end{aligned}
$$

$$y^* = \Phi_w^{-1} u \quad \textbf{Exact Preconditioning}$$

$$y = \bar{\Phi}_w^{-1} u \quad \textbf{Inexact Preconditioning}$$

- **Result VI: The Eigenvalues of the Full-Space and Perturbed System Satisfy**

$$
\begin{aligned}
\sum_k (\lambda_k(\Phi_w) - \lambda_k(\bar{\Phi}_w))^2 &\leq \|\Phi_w - \bar{\Phi}_w\|_F^2 \\
&= \|P_{\mathcal{R}}\|_F^2 \\
&= \|Z - \bar{Z}\|_F
\end{aligned}
$$

- **Can Make Preconditioner Asymptotically Exact by <u>Increasing Number of Clusters</u>**

- **Enables Sparse Preconditioning as Opposed to Dense Schur Preconditioning *(Petra&Anitescu, 2011)***

Multi-Level Preconditioning

- **Clustering Can be Embedded in a Multi-Level Scheme**
- **Cluster Until Lower Level is Factorizable (e.g., 100 → 50 → 25)**
- **Apply Iterative Solver at Upper Levels (Never Factorize Matrix with Large Number of Scenarios)**

DEFINE Multi-Level Clusters $C_0 < C_1 < C_{N_{lev}} = S$ and Sets $\mathcal{R}_0$, $\mathcal{R}_1$, ..., $\mathcal{R}_{N_{lev}}$.

DEFINE Routine $\Delta w^j = \text{MULTILEVEL}(j, \mathcal{R})$:
IF $j = 1$

- $[L, U] = FACT(\mathcal{R}_0)$
- $\Delta w^j = QMR(Matrix = \mathcal{R}_1, Precond = [L, U])$

ELSE

- $\Delta w^j = QMR(Matrix = \mathcal{R}_j, Precond = \text{MULTILEVEL}(j-1, \mathcal{R}))$

END

SOLVE Recovered at Level N_{Lev}:

- $\Delta w^k = \text{MULTILEVEL}(N_{lev}, \mathcal{R})$

Numerical Results - CUTEr

- **Add Stochastic Perturbations to Deterministic QPs**

$$\min \frac{1}{2}y^T Q y + d^T y, \text{ s.t. } Ay = b, \ y \geq 0.$$

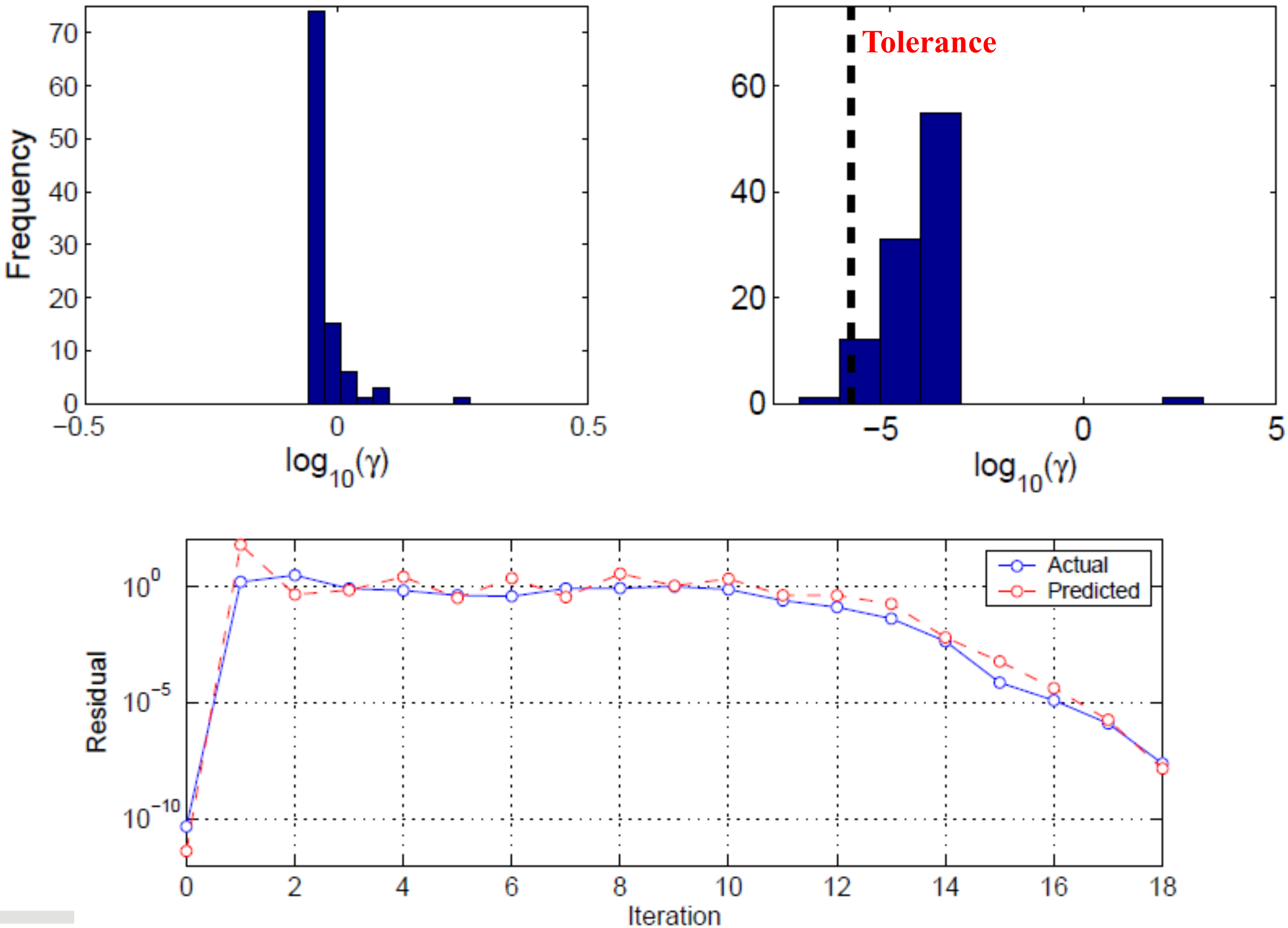
$$\begin{aligned} \min\ & e^T y_0 + S^{-1} \sum_{s \in \mathcal{S}} \frac{1}{2} y_s^T Q y_s + d^T y_s \\ \text{s.t.}\ & A y_s = b_s, \ s \in \mathcal{S} \\ & y_s + y_0 \geq 0, \ s \in \mathcal{S} \\ & y_0 \geq 0. \end{aligned}$$

- **Algorithmic Behavior (Inexact Newton) – Compression Rate of 75%**

	Exact			Inexact					
k	φ^k	Φ^k	$\lvert\mathcal{R}\rvert^k$	φ^k	Φ^k	$\lVert\delta^k\rVert$	$\lVert\delta_p^k\rVert$	$\lVert P_{\mathcal{R}}^k\rVert_F$	$\lvert\mathcal{R}^k\rvert$
1	1.29E+02	5.03E+04	100	1.29E+02	5.03E+04	4.69E-11	4.03E-12	1.01E-13	25
5	6.78E+03	1.02E+01	100	6.62E+03	9.90E+00	6.69E-01	2.63E+00	2.89E-02	25
10	4.94E+03	4.34E-01	100	4.92E+03	9.93E-01	1.00E+00	1.09E+00	3.68E-01	25
15	4.77E+03	4.75E-05	100	4.77E+03	4.42E-02	4.39E-03	6.48E-03	8.98E-01	25
16	4.77E+03	2.31E-06	100	4.77E+03	4.97E-03	7.34E-05	5.98E-04	6.35E-02	25
17	4.77E+03	1.81E-08		4.77E+03	8.25E-05	1.26E-05	4.21E-05	1.28E-01	25
18				4.77E+03	1.26E-05	1.27E-06	1.84E-06	1.95E-01	25
19				4.77E+03	1.27E-06	2.39E-08	1.47E-08	1.39E-02	25
20				4.77E+03	2.39E-08				

Numerical Results - CUTEr

- **Residual Behavior (First and Final Iteration)**



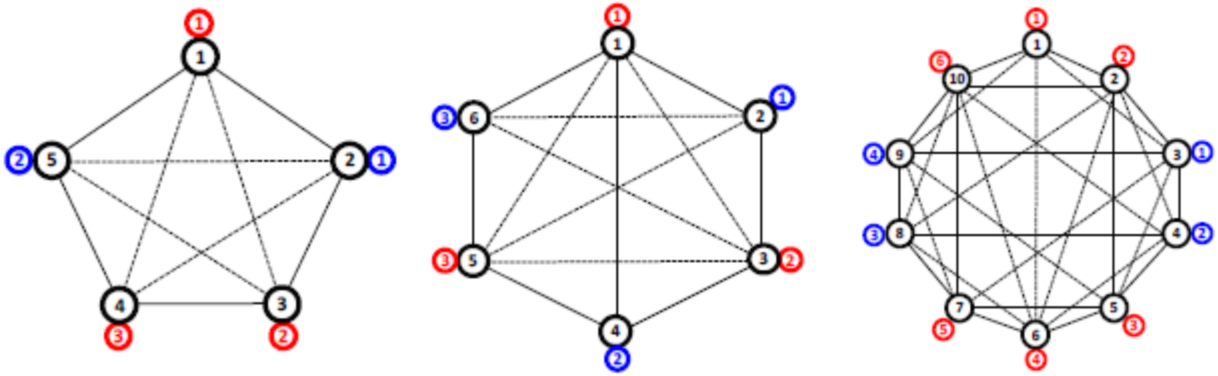
Numerical Results - CUTEr

- Number of Iterations for Problem Set (Inexact Newton) – <u>50% Compression Rates</u>

Problem	S	n	Exact	Inexact
QPTEST	100	505	17	17
ZECEVIC	100	606	20	20
HS76	100	707	19	19
GBD	100	1017	19	19
HS53	100	1110	14	14
LANDS	100	1204	23	24
LOTSCH	100	1212	17	20
QAFIRO	100	5151	34	38
HS118	100	5959	37	38
DUAL1	50	8670	17	21
DUAL2	50	9792	16	18

- Matlab's Hierarchical Clustering (Default Options)
- K-Means Unstable (Highly Dependent on Initial Guess)

Numerical Results – Network Expansion



$$\begin{aligned}
\text{min } & d_g^T \Delta g + d_p^T \Delta p + S^{-1} \sum_{s \in \mathcal{S}} c_g^T g_s \\
\text{s.t. } & \mathbf{P} p_s + \mathbf{G} g_s = \mathbf{D} d_s, \ \ s \in \mathcal{S} \\
& p_L \le p_s + \Delta p \le p_U, \ \ s \in \mathcal{S} \\
& g_L \le g_s + \Delta g \le g_U, \ \ s \in \mathcal{S} \\
& \Delta g \ge 0, \ \ \Delta p \ge 0.
\end{aligned}$$

- **Number of Iterations as Function of Compression Rates – <u>100 Total Scenarios</u>**

Rate %	$n_R = 5$	$n_R = 6$	$n_R = 10$
50	13	13	16
60	14	14	17
70	13	15	18
80	13	15	24
90	16	17	-

Numerical Results – Linderoth/Wright/Shapiro

- **Test <u>Effectiveness of Preconditioner</u> Using Scenario Clustering**
- **Compare Against Scenario Elimination and No Preconditioning**
- **“Large” Problems : Limit of 50 Scenarios in Current Implementation**
- **QMR Tolerance of 1e-6**

Problem	Strategy	n	$\lvert\mathcal{E}\rvert$	IP Iterations	Total QMR It	Avg. QMR It/IP
DUAL1	Unpreconditioned	8,670	-	17	13,462	791.9
DUAL1	Elimination	8,670	25	17	108	6.3
DUAL1	Clustering	8,670	25	17	46	2.7
DUAL1	Elimination	8,670	35	17	162	9.5
DUAL1	Clustering	8,670	35	17	58	3.4
20TERM	Clustering	35,389	25	33	156	4.7
SSN	Clustering	38,263	25	54	287	5.3

Observations:

- **Clustering 2-3 Times More Effective Than Elimination**
- **Compression Rates of 50% Achievable**
- **Preconditioning Enables Scaling to Larger Problems (Inexact Newton More Sensitive)**

Multi-Level Preconditioning

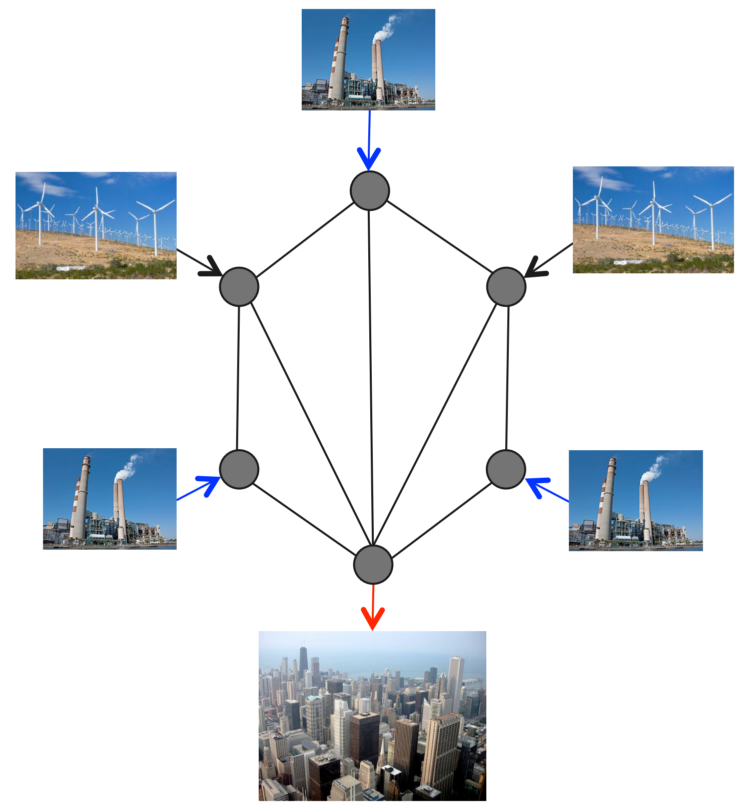
- Clustering Can be Embedded in a Multi-Level Scheme
- Cluster Until Lower Level is Factorizable (e.g., 100 → 50 → 25)
- Apply Iterative Solver at Upper Levels (Never Factorize Matrix with Large Number of Scenarios)

Test Case:

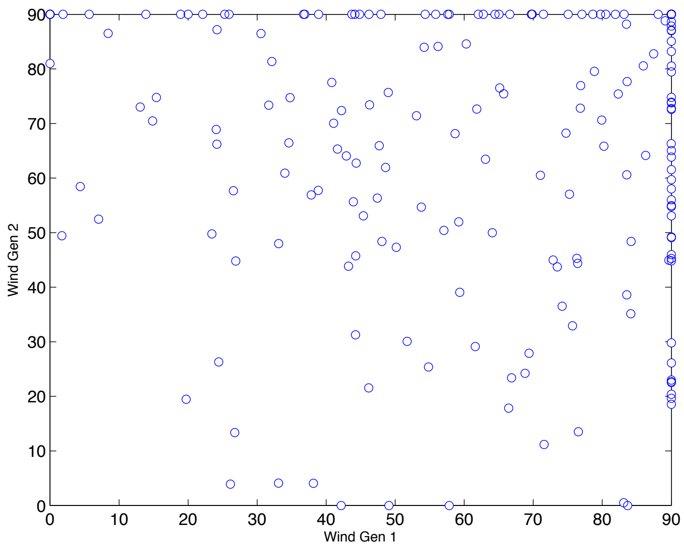
- **Stochastic Day-Ahead Market Clearing,** *Philpott, 2009*
- **One Load, 3 Thermal Generators**
- **Two Wind Power Suppliers with Uncertain Capacity**

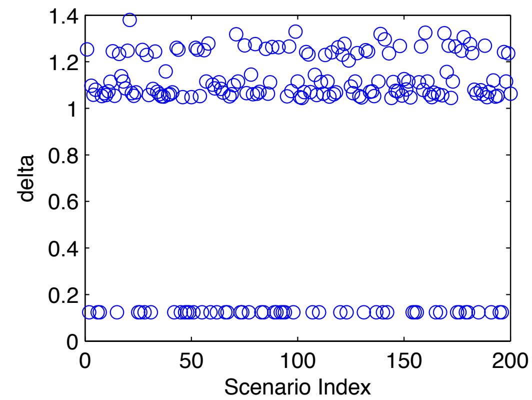
$$
\begin{aligned}
\min \quad \mathbb{E}\Bigg[& \sum_{t\in\mathcal{T}}\sum_{i\in\mathcal{G}} \alpha_i^g g_{i,t} - \sum_{t\in\mathcal{T}}\sum_{i\in\mathcal{D}} \alpha_i^d d_{i,t} \\
& + \sum_{t\in\mathcal{T}}\sum_{i\in\mathcal{G}} \alpha_i^{g,+}(G_{i,t}(\omega) - g_{i,t})_+ - \alpha_i^{g,-}(G_{i,t}(\omega) - g_{i,t})_- \\
& - \sum_{t\in\mathcal{T}}\sum_{i\in\mathcal{D}} \alpha_i^{d,+}(D_{i,t}(\omega) - d_{i,t})_+ + \alpha_i^{d,-}(D_{i,t}(\omega) - d_{i,t})_- \Bigg]
\end{aligned}
$$

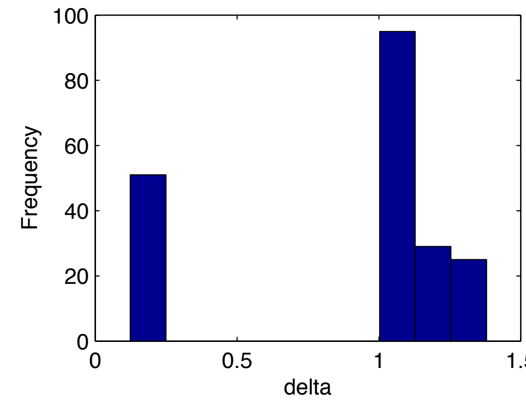
$$
\begin{aligned}
\text{s.t.} \quad & \sum_{i\in\mathcal{L}_j^{rec}} f_{i,t} - \sum_{i\in\mathcal{L}_j^{snd}} f_{i,t} + \sum_{i\in\mathcal{G}_j} g_{i,t} - \sum_{i\in\mathcal{D}_j} d_{i,t} = 0, \qquad t\in\mathcal{T}, j\in\mathcal{B} \\
& \sum_{i\in\mathcal{L}_j^{rec}} F_{i,t}(\omega) - \sum_{i\in\mathcal{L}_j^{snd}} F_{i,t}(\omega) + \sum_{i\in\mathcal{G}_j} G_{i,t}(\omega) - \sum_{i\in\mathcal{D}_j} D_{i,t}(\omega) = 0, \qquad t\in\mathcal{T}, j\in\mathcal{B} \\
& -\bar{f}_{i,t} \le f_{i,t} \le \bar{f}_{i,t}, \qquad t\in\mathcal{T}, i\in\mathcal{L} \\
& 0 \le g_{i,t} \le \bar{g}_{i,t}, \qquad t\in\mathcal{T}, i\in\mathcal{G} \\
& 0 \le d_{i,t} \le \bar{d}_{i,t}, \qquad t\in\mathcal{T}, i\in\mathcal{D} \\
& -\bar{F}_{i,t}(\omega) \le F_{i,t}(\omega) \le \bar{F}_{i,t}(\omega), \qquad t\in\mathcal{T}, i\in\mathcal{L} \\
& 0 \le G_{i,t}(\omega) \le \bar{G}_{i,t}(\omega), \qquad t\in\mathcal{T}, i\in\mathcal{G} \\
& 0 \le D_{i,t}(\omega) \le \bar{D}_{i,t}(\omega), \qquad t\in\mathcal{T}, i\in\mathcal{D}
\end{aligned}
$$



Multi-Level Preconditioning








```
Scenarios ======= 200
Variables ======= 5815
Equality Cons === 2606
Inequality Cons = 8829

it=1  obj=+1.386402e+03 phi=9.410558e+07 mu=1.000000e+06 res=+4.323098e-10 resp=+5.448376e-12 alp=+9.883401e-01 ls=0 |R1|=100 |Rn|=25 itla1=0
it=2  obj=-5.081532e+04 phi=5.559107e+06 mu=4.607473e+04 res=+7.033943e-11 resp=+7.578714e+01 alp=+9.646375e-01 ls=0 |R1|=100 |Rn|=25 itla1=2
it=3  obj=-5.365412e+04 phi=3.022808e+05 mu=2.659006e+03 res=+5.529156e-11 resp=+3.819969e-02 alp=+9.738508e-01 ls=0 |R1|=100 |Rn|=25 itla1=2
it=4  obj=-6.929561e+04 phi=8.812747e+03 mu=7.553232e+01 res=+5.301997e-13 resp=+1.816535e-04 alp=+6.885685e-01 ls=0 |R1|=100 |Rn|=25 itla1=2
it=5  obj=-2.054391e+05 phi=3.274310e+03 mu=3.088757e+01 res=+1.214272e-12 resp=+7.965202e-04 alp=+6.819124e-01 ls=0 |R1|=100 |Rn|=25 itla1=2
it=6  obj=-2.290472e+05 phi=1.287461e+03 mu=1.244863e+01 res=+5.171225e-13 resp=+1.139653e-02 alp=+8.856827e-01 ls=0 |R1|=100 |Rn|=25 itla1=2
it=7  obj=-2.467091e+05 phi=2.293253e+02 mu=2.157364e+00 res=+1.173057e-12 resp=+1.696195e-01 alp=+7.608979e-01 ls=0 |R1|=100 |Rn|=25 itla1=3
it=8  obj=-2.515922e+05 phi=7.429333e+01 mu=7.312822e-01 res=+2.200660e-11 resp=+4.038713e-01 alp=+6.942325e-01 ls=0 |R1|=100 |Rn|=25 itla1=4
it=9  obj=-2.530856e+05 phi=2.461851e+01 mu=2.435109e-01 res=+6.800266e-12 resp=+7.584001e-01 alp=+7.654397e-01 ls=0 |R1|=100 |Rn|=25 itla1=4
it=10 obj=-2.535357e+05 phi=7.259198e+00 mu=7.275386e-02 res=+3.700855e-11 resp=+1.195685e-01 alp=+9.033399e-01 ls=0 |R1|=100 |Rn|=25 itla1=4
it=11 obj=-2.536935e+05 phi=1.169468e+00 mu=1.182384e-02 res=+2.655279e-11 resp=+6.002192e-03 alp=+8.911769e-01 ls=0 |R1|=100 |Rn|=25 itla1=3
it=12 obj=-2.537202e+05 phi=2.051300e-01 mu=2.086365e-03 res=+5.877189e-12 resp=+4.325772e-03 alp=+7.258686e-01 ls=0 |R1|=100 |Rn|=25 itla1=4
it=13 obj=-2.537240e+05 phi=7.266160e-02 mu=7.433331e-04 res=+9.644704e-11 resp=+1.454257e-03 alp=+9.766710e-01 ls=0 |R1|=100 |Rn|=25 itla1=3
it=14 obj=-2.537261e+05 phi=2.246443e-03 mu=2.257112e-05 res=+3.747434e-11 resp=+2.088672e-03 alp=+9.899638e-01 ls=0 |R1|=100 |Rn|=25 itla1=2
it=15 obj=-2.537261e+05 phi=3.079853e-05 mu=2.994731e-07 res=+8.463215e-11 resp=+1.055652e-03 alp=+9.999968e-01 ls=0 |R1|=100 |Rn|=25 itla1=3
it=16 obj=-2.537261e+05 phi=2.327275e-10 mu=9.686669e-13

Elapsed Time = 1.109498e+01
```

200 Total Scenarios and 17,200 Variables in KKT System

Lowest Level Factorizes KKT System with 25 Scenarios and 2,150 Variables (Compression of 87.5%)

Conclusions and Future Work

- **Do Not Cluster Scenarios Based on Data, Cluster Based on Influence on Solution**
- **Presented Strategies to Enable Inside-The-Solver Scenario Clustering**
  - **Inexact Newton and Preconditioning**
  - **Adaptive and Convergence to Original Problem**
  - **Eliminates Scenarios with Strong and Weak Contributions (Redundancies)**
  - **Do Not Require Distributional Information**
  - **Superlinear Convergence and Preconditioning via Interior Point Setting**
- **Large Compression Rates (As Large as 90% in Certain Problems)**
- **Clustering More Effective than Elimination and Sampling**

Opportunities

- **Explore Different Clustering Techniques and Libraries (Big Research Area)**
- **Adaptive Sampling + Clustering**
- **Networks Compression**
- **Parallel Implementation**