SAND2013-2586P

Exceptional service in the national interest

Sandia National Laboratories

Stochastic Optimization and Energy Applications: Operations Unit Commitment and Generation Expansion Planning

Dr. Jean-Paul Watson

Discrete Math and Complex Systems Department

Sandia National Laboratories, New Mexico

Talk Goals

1. To convey the conceptual basics of stochastic optimization

2. To convince you that stochastic optimization *should be* pervasive in power systems operations and planning

3. To convince you that significant recent and upcoming algorithmic advances have made stochastic optimization practical in practice

4. To provide a brief survey of some research projects presently being conduced to support advanced algorithms for core power systems problems

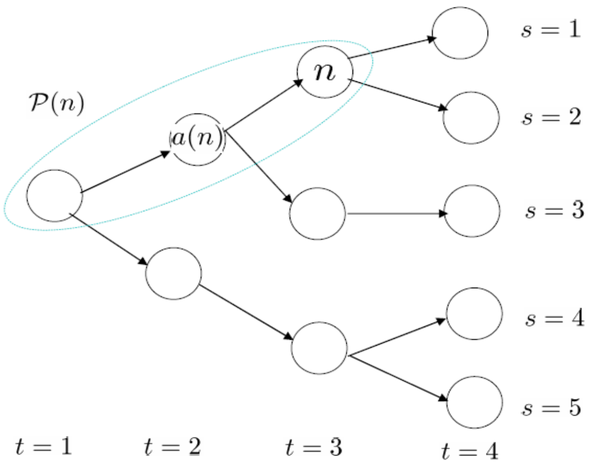
Deterministic vs. Stochastic Optimization

- Deterministic Mixed-Integer Programming (MIP)
  - The workhorse of (rigorous) Operations Research

$$
\begin{aligned}
\min \quad & \boldsymbol{c}'\boldsymbol{x} + \boldsymbol{h}'\boldsymbol{y} \\
\text{s.t.} \quad & \boldsymbol{A}\boldsymbol{x} + \boldsymbol{B}\boldsymbol{y} \leq \boldsymbol{b} \\
& \boldsymbol{x} \in Z^n_+ (\boldsymbol{x} \geq 0, \boldsymbol{x} \text{ integer}) \\
& \boldsymbol{y} \in R^n_+ (\boldsymbol{y} \geq 0)
\end{aligned}
$$

  - Approximable for most real-world problems (NP-Hard)
- Stochastic Mixed-Integer Programming (SMIP)
  - SMIP = MIP + <u>*uncertainty*</u> + <u>*recourse*</u>

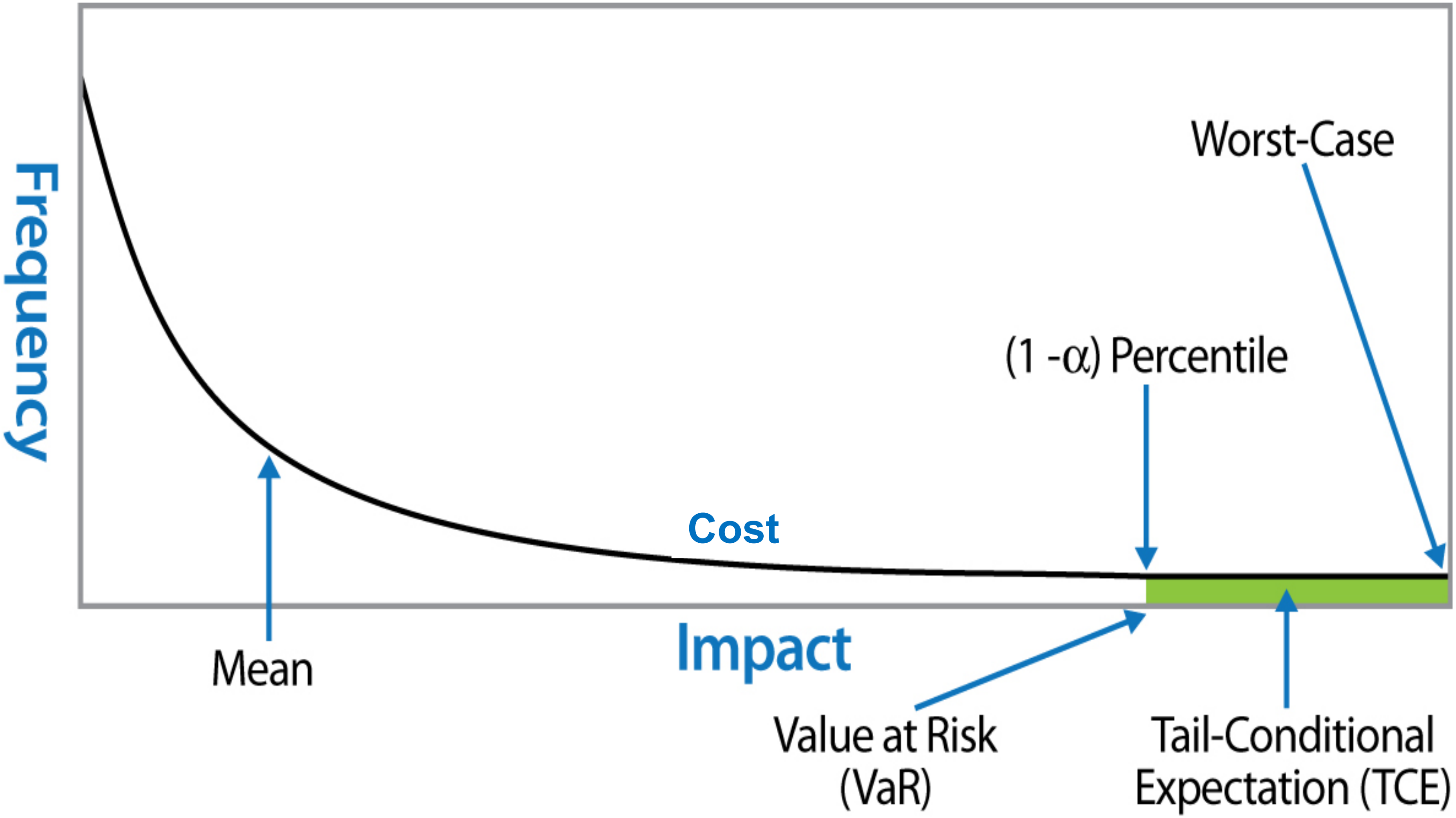
$$
\begin{aligned}
\min \quad & f(\mathbf{x}) = \mathbf{c}^T\mathbf{x} + \mathbb{E}[Q(\mathbf{x}, \omega)] \\
\text{s.t.} \quad & A\mathbf{x} \geq \mathbf{b}, \quad \mathbf{x} \in \mathbb{R}^{n_1 - p_1}_+ \times \mathbb{Z}^{p_1}_+ \\
Q(\mathbf{x}, \omega) = \quad & \min \quad \mathbf{q}(\omega)^T\mathbf{y} \\
& \text{s.t.} \quad W\mathbf{y} \geq \mathbf{h}(\omega) - T(\omega)\mathbf{x} \\
& \mathbf{y} \in \mathbb{R}^{n_2 - p_2}_+ \times \mathbb{Z}^{p_2}_+
\end{aligned}
$$



  - Still NP-Hard, but far more difficult than MIP in practice

On Risk Measures

With stochastic optimization, there is a distribution of costs – and the associated need to select a risk measure...



Stochastic Mixed-Integer Programming: The Algorithm Landscape

- The Extensive Form or Deterministic Equivalent
  - Write down the full variable and constraint set for all scenarios
  - Write down, either implicitly or explicitly, non-anticipativity constraints
  - *Attempt* to solve with a commercial MIP solver
    - Great if it works, but often doesn't due to memory or time limits
- Time-stage or "vertical" decomposition
  - Benders / L-shaped methods (including nested extensions)
  - Pros: Well-known, exact, easy for (some) 2-stage, parallelizable
  - Cons: Master problem bloating, multi-stage difficulties
- Scenario-based or "horizontal" decomposition
  - Progressive hedging / Dual decomposition
  - Pros: Inherently multi-stage, parallelizable, leverages specialized MIP solvers
  - Cons: Heuristic (depending on algorithm), parameter tuning
- *Development of general multi-stage SMIP solvers is an open research area*

Optimization and Power Systems

- Assertion # 1
  - *Almost all power systems operations and planning decision problems with 5 minutes or greater look-ahead are mixed-integer optimization problems*
  - Support: MIP solvers are used widely by ISOs, every day
  - Support: Scan IEEE TPS over the last decade
- Assertion # 2
  - *All of these problems are really stochastic, but are not treated that way because of modeling and scalability issues*
  - Support: Scheduling with renewables is inherently uncertain
  - Support: The uncertainty is there even without renewables!
  - Support: Future budgets and costs are uncertain in expansion problems

So Why Isn't Stochastic Optimization Deployed in Power Systems Contexts?

- Modeling is significantly more complex
  - Stochastic process models, multi-stage decisions
  - Need significant expertise in both optimization *and* statistics
- This is part of the issue, but only part
- The real reason is that stochastic optimization problems are in general exceptionally difficult to solve
  - Solve time are far from those required for operations problems, and don't even approach the turn-around times required for planning
- But:
  - Advances in *decomposition algorithms* over the past two decades, coupled with parallel computing, have changed the landscape!
  - Advances in *statistical inference* can dramatically simplify the computational costs associated with these problems

- *Improved Power Systems Operations using Advanced Stochastic Optimization*
  - *With UC Davis, IA State, Alstom Grid, 3-TIER, and ISO-NE*
  - *Funded under the ARPA-e GENI (Green Energy Network Integration) program*

Project Team

- Sandia National Laboratories
  - Ross Guttromson, MS, MBA, PE
  - Jean-Paul Watson, PhD
  - Cesar Silva Monroy, PhD
- Iowa State University
  - Sarah Ryan, PhD
  - Leigh Tesfatsion, PhD
  - Dionysios Aliprantis, PhD
- Alstom Grid
  - Kwok Cheung, PhD
- UC Davis
  - Roger Wets, PhD
  - David Woodruff, PhD
- ISO New England
  - Eugene Litvinov, PhD

External Advisors

- Eugene Litvinov, ISO-NE
  - Chair
- Richard O'Neill, FERC
- Ralph Masiello, KEMA
- David Morton, UT Austin

Project Goals

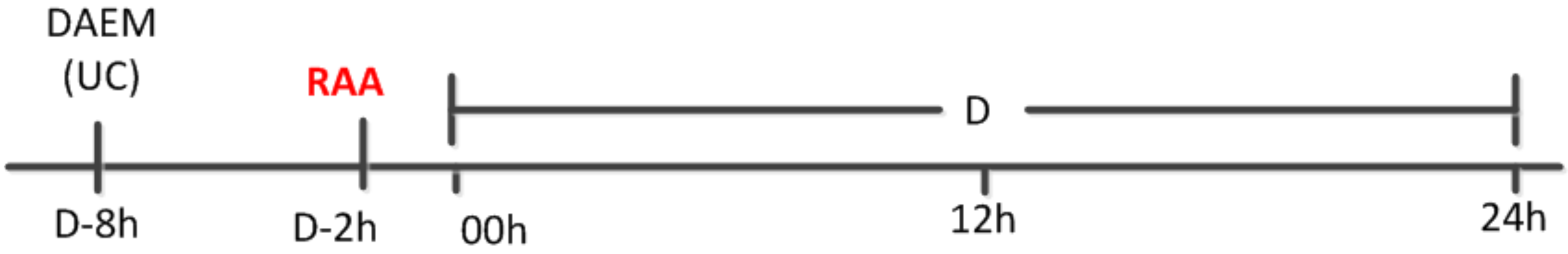
- Execute <u>*stochastic*</u> unit commitment (UC) ***at scale***, ***on real-world data sets***
  - Stochastic UC state-of-the-art is very limited (tens to low hundreds of units)
  - Our solution must ultimately be useable by an ISO
- Produce solutions ***in tractable run-times***, ***with error bounds***
  - Parallel scenario-based decomposition
    - For both upper and lower bounding (Progressive Hedging and Dual Decomp.)
  - Quantification of uncertainty
    - Rigorous confidence intervals on solution cost
- Employ high-accuracy stochastic process models
  - Leveraged to achieve computational tractability while maintaining solution quality and robustness
- Demonstrate ***cost savings on an ISO-scale system*** at ***high renewables penetration levels***

Day-Ahead Unit Commitment (SCUC D-8h)

- Day-Ahead Energy Market (DAEM or DAM)
- Clears **demand bids** and **supply offers** at 1600h on the day prior to the operating day
- Produces:
  - Hourly schedules for the next operating day for market participants (i.e., generation and demand)
  - Hourly interchange schedules
  - Hourly day-ahead Locational Marginal Prices (LMPs)

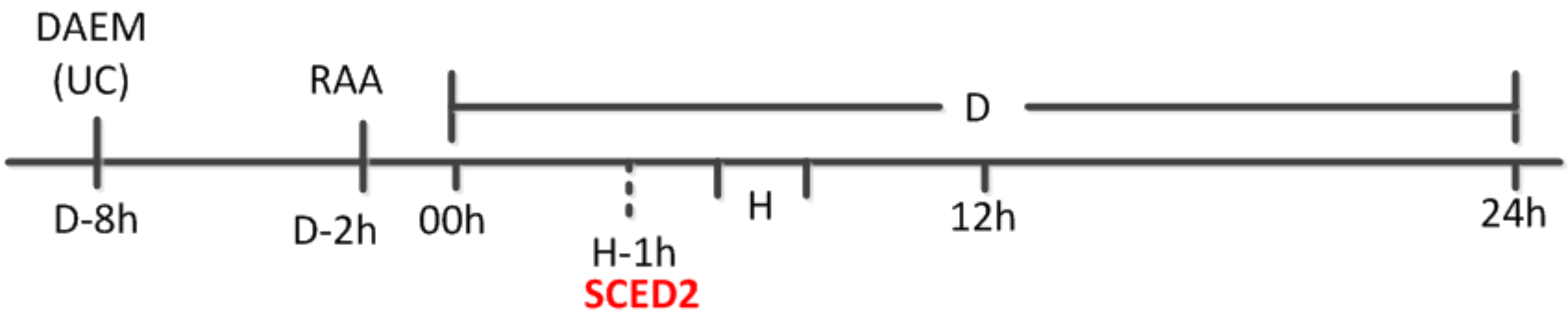
Reliability Unit Commitment - RUC (SCUC D-2h)

- Reliability Assessment (Reserve Adequacy Analysis - RAA)
- Minimize additional start-up and no load costs to provide sufficient capacity to satisfy the forecasted load plus the **operating and replacement reserve requirements**
- Clears ISO **forecasted load** at 2200h
- DAM commitments are respected
- Produces:
  - Additional commitments
  - Updated generator dispatch points



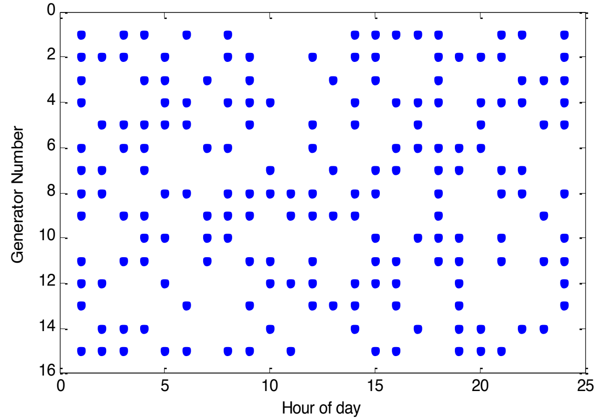
Look-Ahead SCUC (H-1h)

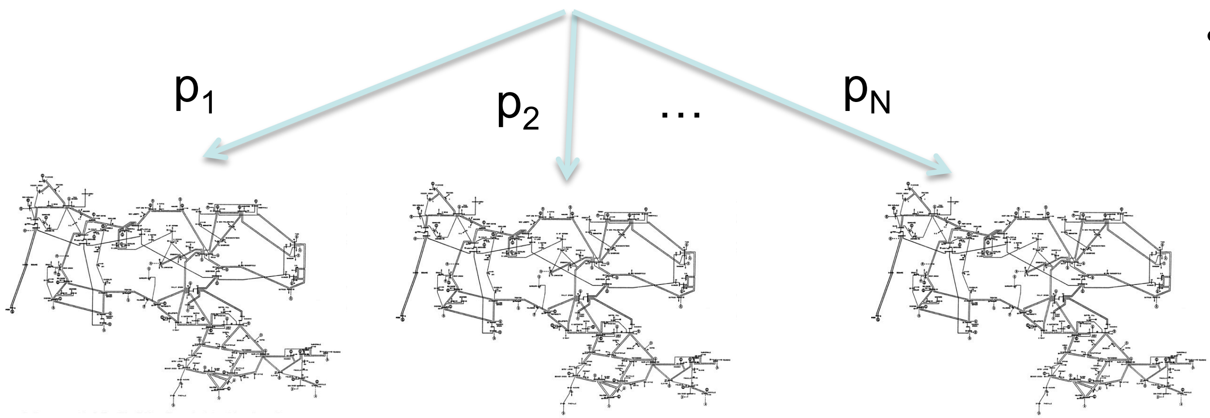
- SCED with ability to bring online fast start resources
- Intended to meet intra-hour reserve requirements
- Leverages updated load and variable generation forecasts
- Produces
  - Generator set points
  - Commitment of fast start units



The General Structure of a Stochastic Unit Commitment Optimization Model

Objective: Minimize expected cost





Scenario 1 Scenario 2 … Scenario N

First stage variables:

- Unit On / Off

Nature resolves uncertainty

- Renewables output
- Forced outages

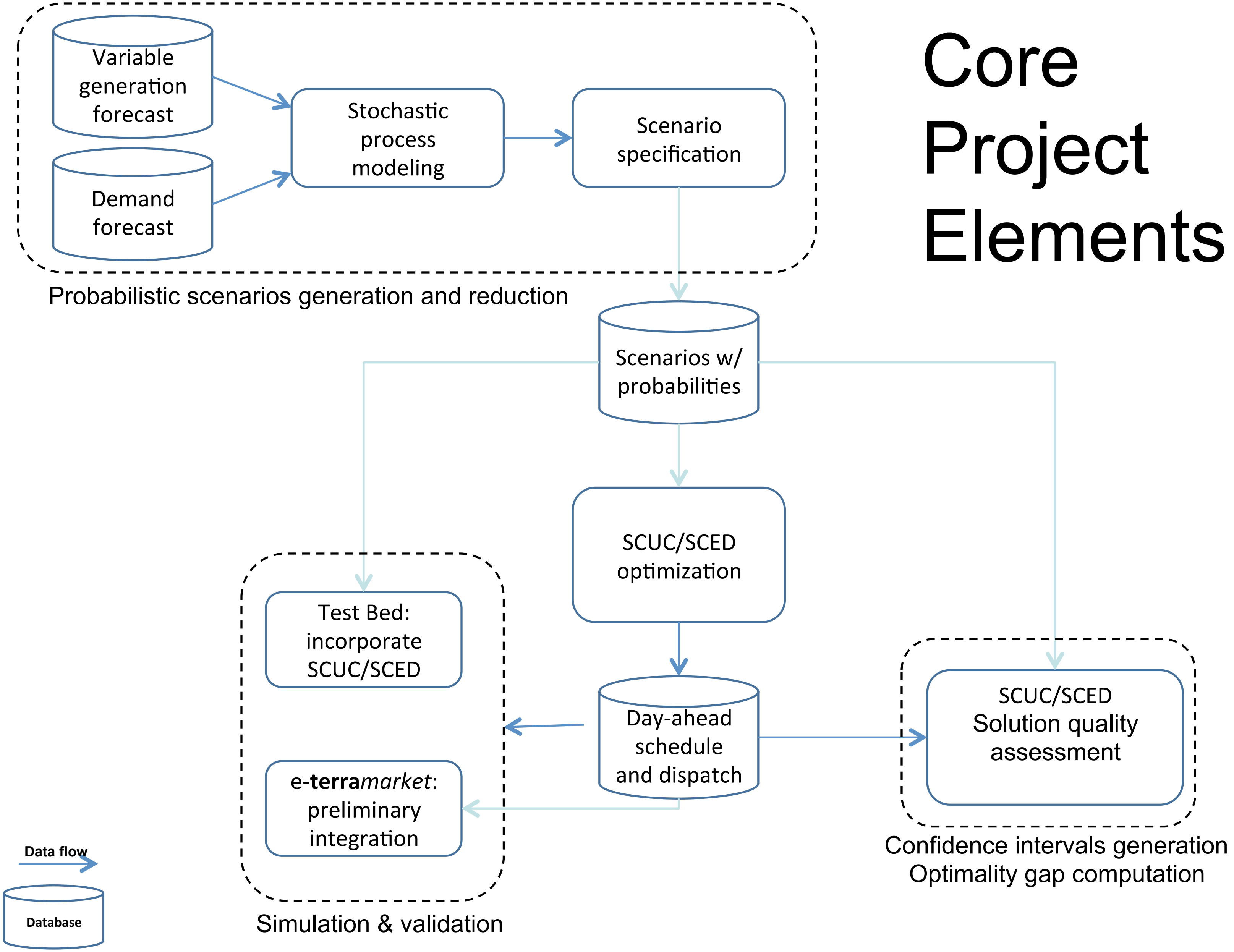
Second stage variables (*per time period*):

- Generation levels
- Power flows
- Voltage angles
- …

Uncertainty in DAM, RUC, and LA SCUC Stochastic Programming Models

- Reliability Unit Commitment
  - Renewables generator output, load, forced (unplanned) outages
  - Fewer binaries than DAM, long time horizon, many scenarios
- Look-Ahead Unit Commitment
  - Similar to Reliability Unit Commitment
  - Fewer binaries than RUC, short time horizon, few scenarios
- Day-Ahead Unit Commitment
  - In contrast to RUC and LA-SCUC, an ISO can't really make direct use of a stochastic UC in the DAM without changing DAM procedures
  - With our partners, we are exploring alternative models and experimenting with procedures that incorporate stochastic models
  - We are eager to discuss ideas offline

Core Project Elements



Impact of scenarios on decisions

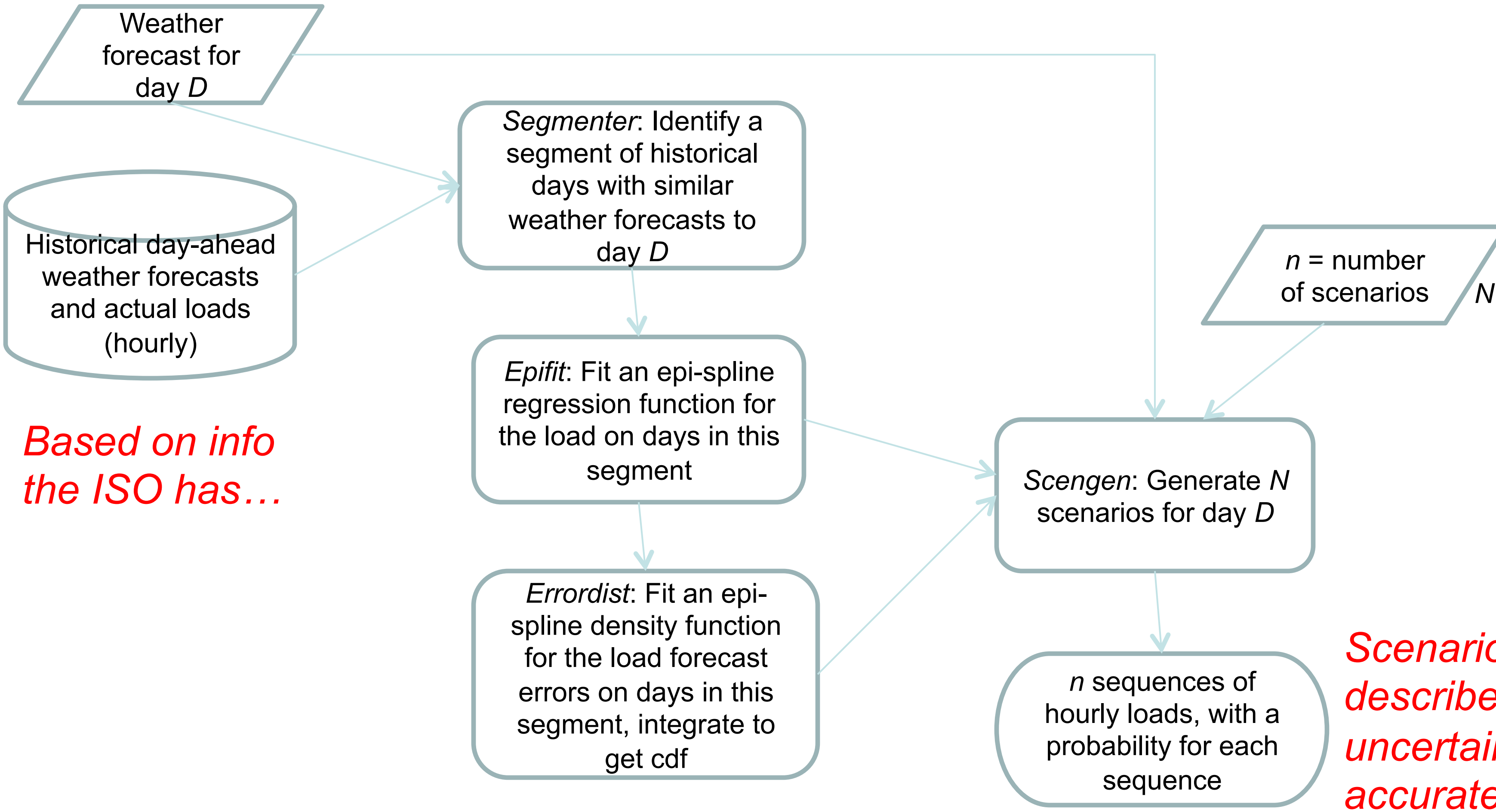
Too "narrow"

- Optimization fails to account for actual risks
- Too few low-cost units committed
  - Cost: Start up additional high-cost units
  - Reliability: Shed load
- Must include a sample of high-impact, low-probability events

Too "wide"

- Optimization result is too risk-averse
- Too many low-cost units committed
  - Cost: Excessive no-load cost of committed units
  - Environmental: spill variable generation
- No better than existing "rule-of-thumb" reserve rules

Overview of process to generate load scenarios on day *D*-1 for day *D*



Based on info the ISO has…

Scenarios describe uncertainty accurately

Scenario-Based Decomposition via Progressive Hedging (PH)

1. $k := 0$
2. For all $s \in \mathcal{S}$, $x_s^{(k)} := \text{argmin}_x (c \cdot x + f_s \cdot y_s) : (x, y_s) \in \mathcal{Q}_s$
3. $\bar{x}^k := (\sum_{s \in \mathcal{S}} p_s \, d_s x_s^{(k)}) / \sum_{s \in \mathcal{S}} p_s \, d_s$
4. For all $s \in \mathcal{S}$, $w_s^{(k)} := \rho(x_s^{(k)} - \bar{x}^{(k)})$
5. $k := k + 1$
6. For all $s \in \mathcal{S}$, $x_s^{(k)} := \text{argmin}_x (c \cdot x + w_s^{(k-1)} x + \rho/2 \left\| x - \bar{x}^{(k-1)} \right\|^2 + f_s \cdot y_s)$ $: (x, y_s) \in \mathcal{Q}_s$
7. $\bar{x}^{(k)} := (\sum_{s \in \mathcal{S}} p_s \, d_s x_s^{(k)}) / \sum_{s \in \mathcal{S}} p_s \, d_s$
8. For all $s \in \mathcal{S}$, $w_s^{(k)} := w_s^{(k-1)} + \rho \left(x_s^{(k)} - \bar{x}^{(k)} \right)$
9. $g^{(k)} := \frac{(1-\alpha)|\mathcal{S}|}{\sum_{s \in \mathcal{S}} p_s \, d_s} \sum_{s \in \mathcal{S}} \left\| x^{(k)} - \bar{x}^{(k)} \right\|$
10. If $g^{(k)} < \epsilon$, then go to step 5. Otherwise, terminate.

Rockafellar and Wets (1991)

Progressive Hedging: Some Algorithmic Issues and their Resolution

- We are dealing with mixed-integer programs
  - So we have to deal with the possibility of cycling and other manifestations of non-convergence
  - See: *Progressive Hedging Innovations for a Class of Stochastic Mixed-Integer Resource Allocation Problems*, J.P. Watson and D.L. Woodruff, Computational Management Science, Vol. 8, No. 4, 2011

- What about good values for that pesky ρ parameter?
  - Poor or ad-hoc values of ρ can lead to atrocious performance
  - The good news in unit commitment
    - We have a lot of information concerning the cost of using a generator
    - Cost-proportional rho is a known, effective strategy in Progressive Hedging
  - Also see Computational Management Science paper indicated above

Progressive Hedging: Parallelization and Bundling

- Progressive Hedging is, at least conceptually, easily parallelized
  - Scenario sub-problem solves are clearly independent
  - Advantage over Benders, in that “bloat” is distributed
    - Critical in low-memory-per-node cluster environments
  - Parallel efficiency drops rapidly as the number of processors increases
    - But: *Relaxing barrier synchronization does not impact PH convergence*

- Why just one scenario per processor?
  - Bundling: Creating miniature “extensive forms” from multiple scenarios
    - Diverse or homogeneous scenario bundles?
  - Empirically results in a large reduction in total number of PH iterations
    - Growth in sub-problem cost *must* be mitigated by drop in iteration count
    - In practice, mitigation is enabled by cross-iteration warm starts

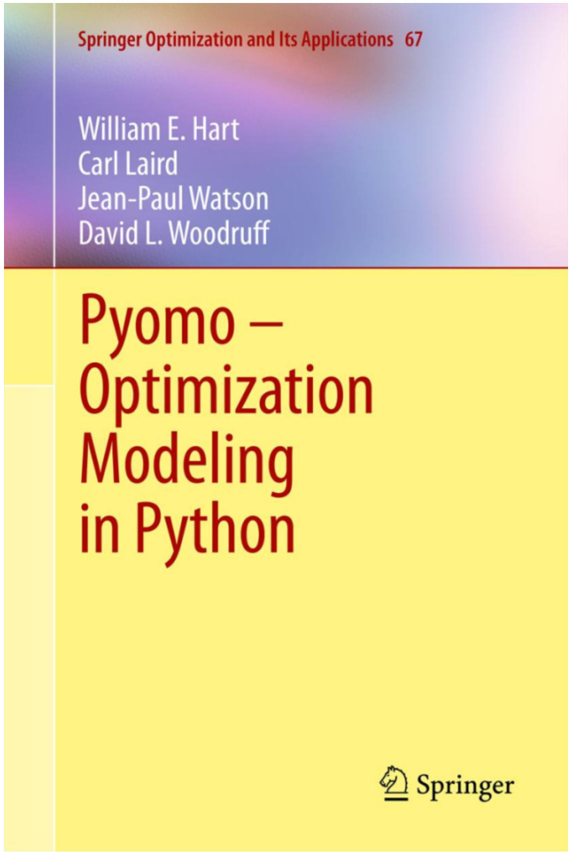
Illustrative Results: WECC-240

- Test instance
  - Modified WECC-240 instance for reliability unit commitment
  - Stochastic demand, 100 scenarios
- Extensive form
  - CPLEX, after 1 day of CPU on a 16-core workstation
    - No feasible incumbent solution
- PH, 20 iterations, post-EF solve - serial
  - ~14 hours, 2.5% optimality gap
- PH, 20 iterations, post-EF solve – parallel
  - ~15 minutes, 2.5% optimality gap
- PH, 20 iterations, post-EF solve – parallel with bundling
  - ~15 minutes, 1.5% optimality gap

Our Software Environment: Coopr

Sandia National Laboratories

- Project homepage
  - http://software.sandia.gov/coopr
- "The Book"



- Mathematical Programming Computation papers
  - Pyomo: Modeling and Solving Mathematical Programs in Python (Vol. 3, No. 3, 2011)
  - PySP: Modeling and Solving Stochastic Programs in Python (Vol. 4, No. 2, 2012)

Our Hardware Environments

- Our objective is to run on commodity clusters
  - Utilities don't have, and don't want, supercomputers
  - But they do or might have multi-hundred node clusters

- Sandia Red Sky (Unclassified Segment) – 39th fastest on TOP500
  - Sun X6275 blades
  - 2816 dual socket / quad core nodes (22,528 cores)
    - 2.93 GHz Nehalem X5570 processors
    - 12 GB RAM per compute node (1.5 GB per core) << IMPORTANT!
  - For us, the interconnection is largely irrelevant
  - Red Hat Linux (RHEL 5)

- Sandia Red Mesa (with NREL)
  - Similar to Red Sky, but dedicated for energy research

A Few Words on UC Test Instances

- From the academic literature
  - Hand-constructed instances (Silva Monroy)
  - Textbook instances (Wood and Wollenberg)
  - RUC test literature – 10 and 100 generator instances
  - Simplified CAISO+WECC 240-bus test case
- From FERC
  - PJM-inspired / anonymized large-scale DAM UC and RUC instance
- From Alstom Grid
  - 70-bus test instance, used in development and testing of e-*terra*market
- From ISO-NE
  - Eastern Interconnection Planning Model-based instances

More Words on UC Test Instances

- What baseline deterministic UC model is best?
  - Carrion and Arroyo
  - Traditional three-binary generator state representation
  - Ostrowski et al.
  - …
- Lesson Learned #1
  - Not all models are correct, and all papers have unreported bugs
- Lesson Learned #2
  - Performance is *highly* dependent upon the test case
- Lesson Learned #3
  - Existing UC test cases are *really* bad
- Lesson Learned #4
  - Validating UC models is a highly non-trivial exercise

Conclusions

- Stochastic unit commitment has been studied in the literature
  - Indications are that it holds promise
  - Computational challenges have prevented industrial adoption
  - Far easier on paper and in academia than in practice…
- Our objective is to develop scalable solutions to stochastic unit commitment
  - In tractable run-times
  - On ISO-scale systems
  - To demonstrate (or not) both practical deployment ability and cost savings
  - Using reasonable, high-accuracy stochastic process models
- We are happy to talk to:
  - ISOs, vendors, labs, and academics working toward related goals

- *Toward “Productizing” APRA-e Stochastic Unit Commitment Software Libraries*

Sandia National Laboratories

Advancing Solar Deployment on Transmission Systems

Abraham Ellis (PI), SNL

SunShot
U.S. Department of Energy

Open Tools for Operational Analysis

- Scope
  - Develop, document and disseminate a new, open-source tool that uses stochastic mixed-integer optimization techniques to study integration cost and system reliability for projected high solar deployment scenarios
- Approach
  - Customize SCUC modules based on existing optimization software, with interfaces to facilitate efficient data entry and parametric analyses
  - Validate tool functionality on a reference case
  - Make toolkit (including documentation and case studies) available through an open-source distribution model
- Impact
  - By definition, high pen PV means higher uncertainty, more stark cost-reliability tradeoff. Stochastic tools are best suited to fully assess impacts.
  - Access to a new tool that allows for more rigorous analysis of costs and reliability implications associated with high penetration solar scenarios

- *Stochastic Generation Expansion*
  - *With UC Davis and IA State*
  - *Funded by Sandia LDRD and DOE/OS*

Stochastic Generation Expansion

- Two-stage stochastic (mixed-integer) programming formulation
- Objective:

$$\min \sum_y \left(\frac{\sum_g \left(c_g m_g^{\max} \boldsymbol{U}_{g,y}\right) + \sum_s \pi_s \left(\sum_{t \in T_y} \left(\sum_g \left(l_{g,t,s} \boldsymbol{L}_{g,t,s}\right) + p_u \boldsymbol{E}_{t,s}\right)\right)}{(1+r)^y} \right) \quad (1)$$

- Constraints:

$$\sum_g \boldsymbol{L}_{g,t,s} + \boldsymbol{E}_{t,s} = d_{t,s} \quad \forall t,s \quad (2)$$

$$\boldsymbol{L}_{g,t,s} \le n_g^{max} (u_g + \sum_{y \le Y_t} \boldsymbol{U}_{g,y}) \quad \forall g,t,s \quad (3)$$

$$\sum_y \boldsymbol{U}_{g,y} \le u_g^{\max} \quad \forall g \quad (4)$$

Jin and Ryan (2010)

A MISO-Inspired Test Instance

- Uncertain parameters
  - Fuel price
  - Demand

- Scenario tree structure
  - 10 time periods
  - Branching factor of 3
    - Total of 3^9 = 19,683 scenarios

- Extensive form is very difficult to solve at this scale
  - >> weeks of CPU time
  - Even for 1000 scenarios, run-time is on the order of hours
  - Benders yields significant run-time reductions, but still requires days

Scenario Sampling: How Many is Enough?

- Discretization of the scenario tree is "standard" in stochastic programming
  - Often, no mention of solution or objective stability
  - Let alone rigorous statistical hypothesis-testing of stability
  - *Don't trust anyone who doesn't show you a confidence interval!*

- Various approaches / alternatives in the literature
  - Sample Average Approximation
  - Multiple Replication Procedure

- Formal question we are concerned with
  - What is the probability that $\widehat{x}$' s objective function value is suboptimal by more than α%?
  - But making due with a fixed set or "universe" or scenarios

The Multiple Replication Procedure (Mak, Morton, Wood 1999)

MRP:

Input: Value $\alpha \in (0,1)$ (e.g., $\alpha = 0.05$), sample size n, replication size n_g, and a candidate solution $\hat{x} \in X$.

Output: Approximate $(1-\alpha)$-level confidence interval on $\mu_{\hat{x}}$.

1. For $k = 1, 2, \ldots, n_g$:
   1.1. Sample i.i.d. observations $\xi^{k1}, \xi^{k2}, \ldots, \xi^{kn}$ from the distribution of ξ.
   1.2. Solve (SP_n) using $\xi^{k1}, \xi^{k2}, \ldots, \xi^{kn}$ to obtain x_n^{k*}.
   1.3. Calculate $G_n^k(\hat{x}) = n^{-1} \sum_{j=1}^{n} (f(\hat{x}, \xi^{kj}) - f(x_n^{k*}, \xi^{kj}))$.
2. Calculate gap estimate and sample variance by

$$\bar{G}_n(n_g) = \frac{1}{n_g} \sum_{k=1}^{n_g} G_n^k(\hat{x}) \quad \text{and} \quad s_G^2(n_g) = \frac{1}{n_g - 1} \sum_{k=1}^{n_g} (G_n^k(\hat{x}) - \bar{G}_n(n_g))^2.$$

3. Let $\epsilon_g = t_{n_g-1,\alpha} s_G(n_g)/\sqrt{n_g}$, and output the one-sided CI on $\mu_{\hat{x}}$,

$$[0, \bar{G}_n(n_g) + \epsilon_g].$$

From Bayraksan and Morton (2009) – Assessing Solution Quality in Stochastic Programs Via Sampling

MRP Results: Expected Cost Minimization

- (Some) MRP results:

Number Samples	Number Groups	Solution Cost	95% CI Width
70	10	17.4B	655M
140	10	17.4B	164M
280	10	17.5B	151M
420	10	17.4B	198M
560	10	17.4B	249M

- Key observations
  - Remarkably stable results with relatively few scenarios
  - Full-scale scenario tree likely not needed for expected cost metric

- Miscellaneous
  - A potential issue is the stability of the MRP results
  - Reporting *maximum* values above across 5 MRP replications

Conditional Value-at-Risk: Formulation

- Conditional Value-at-Risk can be expressed as an expected value computation, e.g., as follows (Schultz and Tiedemann 2006)

Proposition 5.1. *Assume that μ is discrete with finitely many scenarios $h_1, \dots, h_J$ and corresponding probabilities $\pi_1, \dots, \pi_J$. Let $\alpha \in (0, 1)$. Then the stochastic program*

$$\min\{Q_{CVaR_\alpha}(x) \ : \ x \in X\} \tag{11}$$

can be equivalently restated as

$$\begin{aligned} \min_{x,y,y',v,\eta} \Big\{ \eta + \frac{1}{1-\alpha} \sum_{j=1}^{J} \pi_j v_j \ : \ & W y_j + W' y'_j = h_j - T x, \\ & v_j \geq c^\top x + q^\top y_j + q'^\top y'_j - \eta, \\ & x \in X, \quad \eta \in \mathbb{R}, \quad y_j \in \mathbb{Z}_+^{\bar{m}}, \\ & y'_j \in \mathbb{R}_+^{m'}, \ v_j \in \mathbb{R}_+, \quad j = 1, \dots, J \Big\}. \end{aligned} \tag{12}$$

MRP Results: Conditional Value-at-Risk Minimization

- (Some) MRP results:

Number Samples	Number Groups	Solution Cost	95% CI Width
70	10	28.0B	2.1B
140	10	25.7B	1.8B
280	10	24.5B	4.2B
420	10	24.6B	2.3B
560	10	24.0B	3.8B

- Key observations
  - As expected, CVaR results are less stable than expected cost results
  - Not as variable as those seen in other domains => less tail?

- Miscellaneous
  - Much more variability across MRP replications
  - Again reporting maximum values above across 5 MRP replications

Conclusions and Future Research Directions

- We have generic software available for computing confidence intervals on the optimal objective function values for stochastic (mixed-integer) programs
  - Straightforward application to a grid generation expansion problem
  - Analytic results for both expected cost and Conditional Value-at-Risk
  - Practical implications for future development and experimentation

- In-Progress and/or Near-Term Extensions
  - Number of samples required for tight CVaR confidence intervals
  - Multi-stage versus two-stage comparisons
  - Solution stability analysis
  - Consideration of chance constraints to model service level thresholds

- *A Brief Wrap-Up…*

Talk Goals, Revisited

1. To convey the conceptual basics of stochastic optimization
2. To convince you that stochastic optimization should be pervasive in power systems operations and planning
3. To convince you that significant recent and upcoming algorithmic advances have made stochastic optimization practical in practice
4. To provide a brief survey of some research projects presently being conduced to support advanced algorithms for core power systems problems

(Some) References

- Modeling and Solving a Large-Scale Generation Expansion Planning Problem Under Uncertainty
  - S. Jin, S.M. Ryan, J.P. Watson, and D.L Woodruff
  - Energy Systems (2011), Volume 2, Issue 3-4
- Toward Scalable, Parallel Progressive Hedging for Stochastic Unit Commitment
  - S.M. Ryan, R.J.B. Wets, D.L. Woodruff, C. Silva-Monroy, and J.P. Watson
  - Proceedings of the 2013 IEEE PES General Meeting
- A New Approximation Method for Generating Day-Ahead Load Scenarios
  - Y. Feng, D. Gade, S.M. Ryan, J.P. Watson, R.J.B. Wets, and D.L. Woodruff
  - Proceedings of the 2013 IEEE PES General Meeting

QUESTIONS