Economic Impacts of Wind Covariance Estimation on Power Grid Operations

Cosmin G. Petra

Mathematics and Computer Science Division

Argonne National Laboratory (ANL)

petra@mcs.anl.gov

Contributors: V. Zavala (ANL), E. D. Nino (V. Tech), M. Anitescu (ANL)

FERC Technical Conference on RT and Market Efficiency through Improved Software

June 23-25, 2014

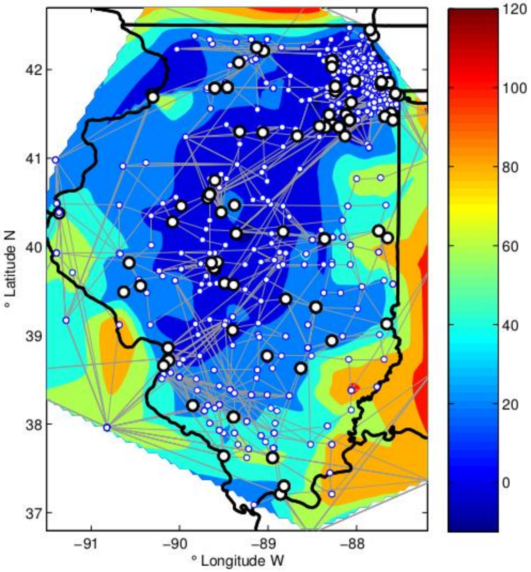
Outline

- Economic dispatch models in power grid
- Stochastic optimization models
- Wind volatility, weather forecasting and (re)sampling
- Covariance estimation and impact on optimal dispatch
  - Simple model
  - Simulation of economic dispatch for Illinois power grid
- Conclusions

Economic dispatch models

- Basis for the electricity distribution and electricity market
- Used by all Independent System Operators (ISOs) in the US.
- In the simpler form, for direct currents, is formulated as a linear programming problem

$$
\begin{aligned}
\min_{x,f} \quad & \sum_{i \in G} c_i x_i \\
\text{subj.to:} \quad & \tau_n(f) + \sum_{i \in T(n)} x_i = d_n, \forall n \in N \\
& f \in U \\
& x_i \in C_i, \forall i \in G
\end{aligned}
$$



Stochastic dispatch models

- Adoption of highly volatile renewable energy and randomness in demand requires stochastic formulations
- Cost-optimal decision in the presence of uncertain generation/demand
- Two-stage linear stochastic programming with recourse: “energy only” model (Pritchard, Zakeri, Philpott, 2010)

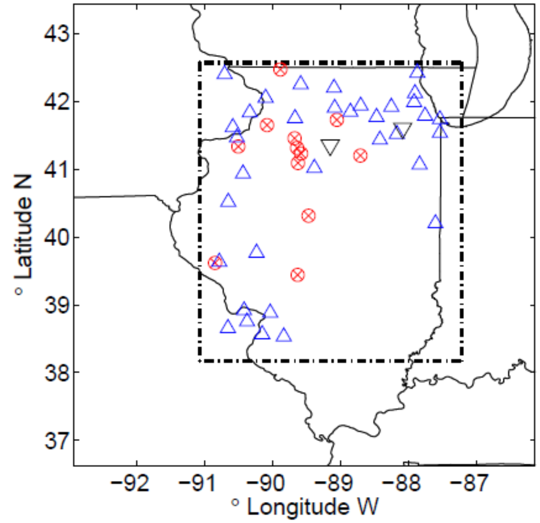
$$
\begin{aligned}
\min_{x,X(\omega),f,F(\omega)} \quad & \sum_{i\in G} c_i x_i + \mathbb{E}_\omega \sum_{i\in G} p_i |x_i - y_i(\omega)| \\
\text{subj.to:} \quad & \tau_n(f) + \sum_{i\in T(n)} x_i = d_n, \forall n \in N \\
& \tau_n(F(\omega)) + \sum_{i\in T(n)} y_i(\omega) = d_n, \forall n \in N, \omega \in \Omega \\
& f, F(\omega) \in U, \forall \omega \in \Omega \\
& x_i, y_i(\omega) \in U_i, \forall i \in G, \omega \in \Omega
\end{aligned}
$$

- Two-markets: “ahead”, decisions/prices to be taken-now and “realtime”, scenario-specific adjustments in decisions/prices.
- The model is ISO revenue adequate (no “missing money”).

Integrating wind samples in the economic dispatch model

- The probability distributions are usually not known, and sampling is used (Ω is finite in practice).

- Treatment of uncertainty in dispatch
  - **Approach 1**: Wind farms bid energy based on their own, independent forecasts
    - Wind correlation between wind farms is lost
  - **Approach 2**: Centralized wind/weather forecast at the ISO level
    - Properly accounting for correlation(s)



- Here we show that Approach 2 should be considered: ignoring or missing correlation information leads to inefficient dispatch.

Motivating example – role of correlation in dispatch

- A very simplistic model: 3 generators (of which 2 wind farms and 1 thermal), 1 demand node, no line constraints
- Power outputs of the wind farms are $W_1 \sim \mathcal{N}(w_1, \sigma_1)$ and $W_2 \sim \mathcal{N}(w_2, \sigma_2)$, and the correlation is ρ ($\rho = \mathbb{E}[(W_1 - w_1)(W_2 - w_2)]/(\sigma_1\sigma_2)$).

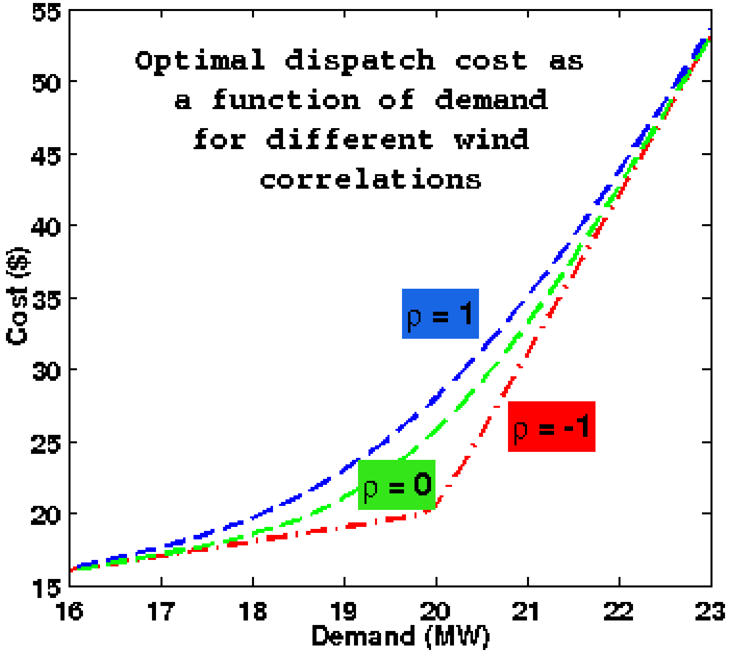
- **How does correlation affect the optimal dispatch cost?**
- The optimization problem can be solved analytically, and the (expected) optimal dispatch cost is:

$$c_d(\rho) = c_w d + (c_{th} - c_w)((d - w_1 - w_2)\Phi(d, \sigma_1^2 + 2\rho\sigma_1\sigma_2 + \sigma_2^2) + \sigma^2\phi(d, \sigma_1^2 + 2\rho\sigma_1\sigma_2 + \sigma_2^2))$$

- Here Φ and ϕ are the cumulative distribution and probability distribution functions of $W = W_1 + W_2$

- The optimal dispatch cost is an **increasing function** of the correlation ρ !

Motivating example - continued



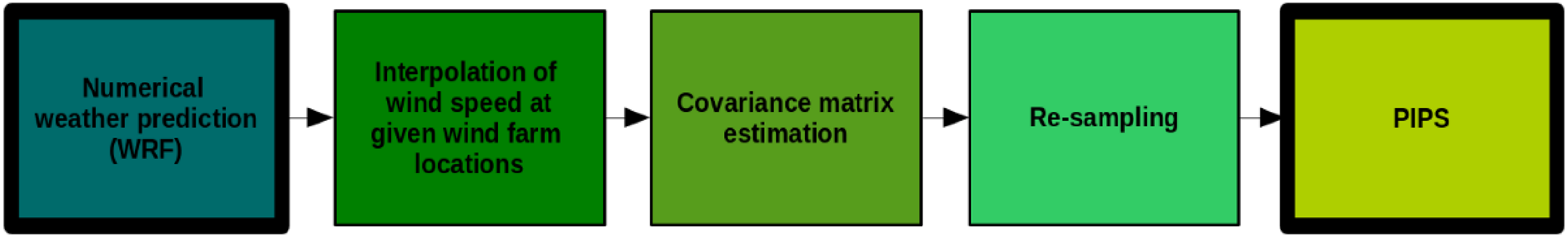
- Not accounting for positive correlation leads to an "optimistic dispatch" (more wind power is thought to be available).
- Ignoring negative correlations results in an "pessimistic dispatch" (foreseen wind is low, therefore, more thermal generation is dispatched).
- In both cases higher operating costs are obtained over time:
  - "optimistic" case potentially ends in using expensive power from the reserves to replace the wind that was predicted but not realized.
  - "pessimistic" case has higher dispatch cost since more thermal generation than necessary is dispatched.
- Also leads to arbitrage opportunities in the power market for participants that account or have better approximation of the correlation.

A framework for stochastic economic dispatch

- **What about real-world large-scale power grid systems?**

- Analytical analysis of such complex systems is virtually impossible.
- Computer simulations are needed.

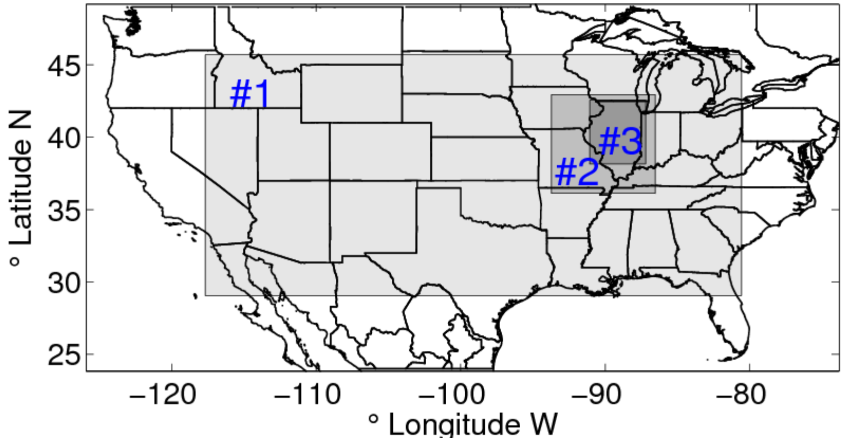
- Weather forecasting is integrated with decision making under the same computational framework.



- Wind samples using WRF, resampling using shrinkage estimators (more later).
- **PIPS (Petra et all)** - parallel optimization solver for high performance computing platforms (BG/P, BG/Q, Cray XE6, XC30, XK7).

Wind forecast

- Weather forecasting @ Argonne (E. Constantinescu)



- WRF (Weather Research and Forecasting) Model
  - Real-time grid-nested simulation using atmospheric models
  - Done on high performance computing platforms but still computationally expensive
  - Only 30 samples or less can be obtained in times compatible with operational practice

Covariance estimation

- A small number of samples may not accurately capture the uncertainty.
- We assume Gaussian distribution of wind speeds and resample to generate more samples.
- **The statistical problem**: estimate the covariance matrix Q a random p-dimensional vector based on a number of n samples
  - $X = [x_1; x_2; \cdots x_n] \in \mathbb{R}^{p \times n}$ are the samples
  - Let $\bar{x} = (x_1 + x_2 + \ldots + x_n)/n$ denote the sample mean
  - An estimator of the covariance matrix would then be

$$S = \frac{1}{n}\left[\sum_{i=1}^{n} (\mathbf{x}_i - \overline{\mathbf{x}}) \cdot (\mathbf{x}_i - \overline{\mathbf{x}})^{\mathrm{t}}\right] \in \mathbb{R}^{p \times p}$$

- Estimating covariance matrix is an issue in this situation since the number of samples (n=30) is smaller than the number of random variables (p=O(100)).

Shrinkage estimators

- Estimators of the form

$$S_\Theta = \rho_1 \cdot I + \rho_2 \cdot S$$

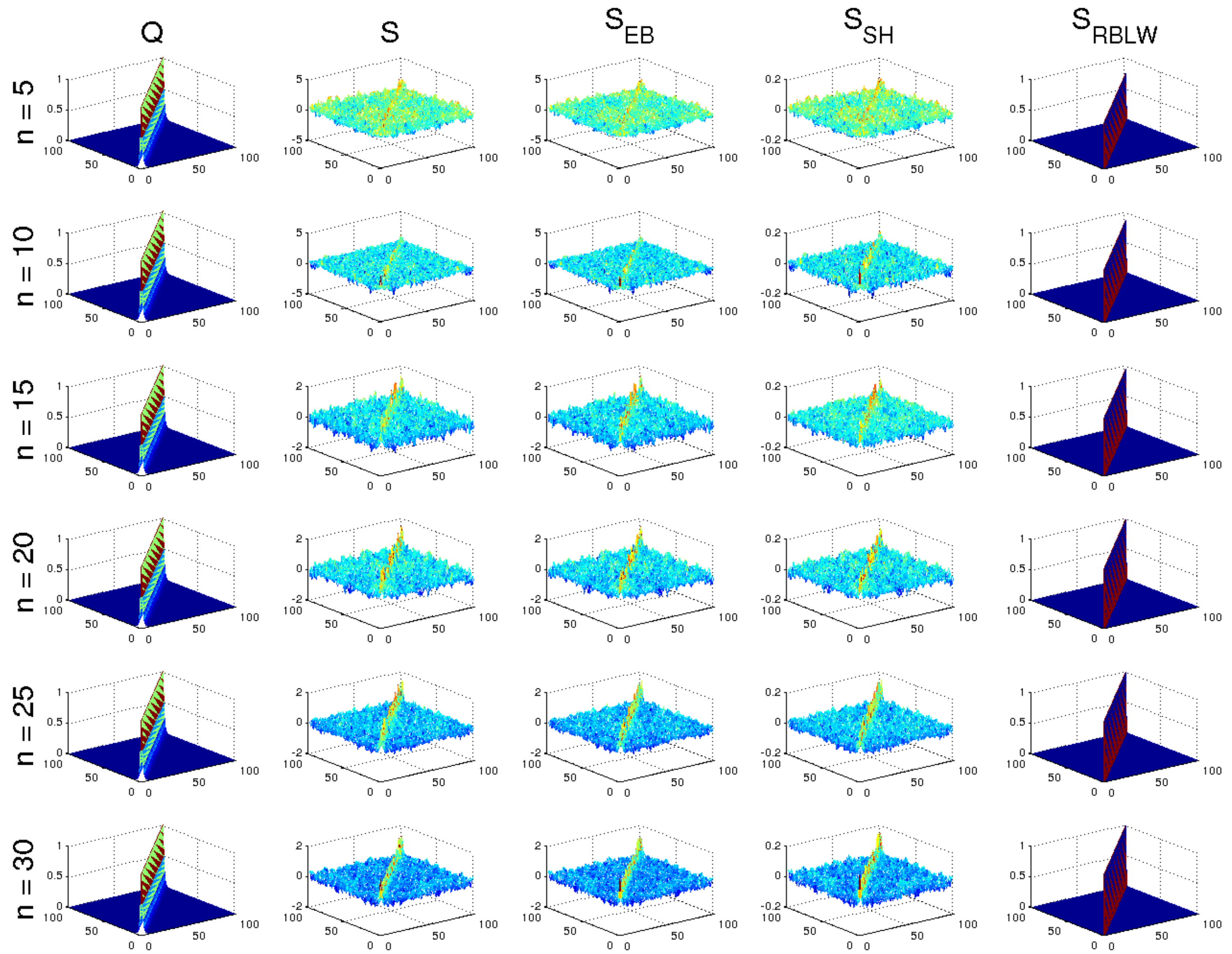
- Where the parameters are chosen so that

$$\min_{\rho_1, \rho_2} \mathbb{E}\left[\| Q - S_\Theta \| \right]$$

- Rao Blackwell Ledoit Wolf (2004) estimator

$$\widehat{\Sigma}_{RBLW} = \rho_{RBLW} \cdot I_{p \times p} + (1 - \rho_{RBLW}) \cdot S, \text{ where}$$
$$\rho_{RBLW} = \min\left(\frac{\frac{n-2}{n} \cdot \mathrm{tr}\left(S^2\right) + \mathrm{tr}^2(S)}{(n+2) \cdot \left[\mathrm{tr}(S^2) - \frac{\mathrm{tr}^2(S)}{p} \right]}, 1 \right)$$

Validation on an Autoregressive process (AR)



S_{EB} is the empirical Bayesian estimator and S_{SH} is Stein's SVD decomposition-based estimator.

Stochastic ED as (dual) block-angular LPs

Extensive form

$$
\begin{array}{rccccccccccc}
\min & c_0^T x_0 & + & c_1^T x_1 & + & c_2^T x_2 & + & \dots & + & c_N^T x_N & & \\
\text{s.t.} & A x_0 & & & & & & & & & = & b_0, \\
& T_1 x_0 & + & W_1 x_1 & & & & & & & = & b_1, \\
& T_2 x_0 & & & + & W_2 x_2 & & & & & = & b_2, \\
& \vdots & & & & & & \ddots & & & & \vdots \\
& T_N x_0 & & & & & & & + & W_N x_N & = & b_N, \\
& x_0 \geq 0, & & x_1 \geq 0, & & x_2 \geq 0, & & \dots, & & x_N \geq 0. & &
\end{array}
$$

- Easy to build practical instances having billions of decision variables and constraints

  ➔ Requires distributed memory computers
- Real-time solution needed in power grid applications

Interior-point optimization solver - PIPS

Convex quadratic problem

$$\text{Min } \frac{1}{2}x^T Qx + c^T x$$

$$\text{subj. to. } Ax = b$$

$$x \geq 0$$

IPM Linear System

$$\begin{bmatrix} Q+\Lambda & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = rhs$$

2 solves per IPM iteration
- predictor direction
- corrector direction

Multi-stage SP

Two-stage SP

nested arrow-shaped linear system (modulo a permutation)

$$\begin{bmatrix} H_1 & B_1^T & & & & & & 0 & 0 \\ B_1 & 0 & & & & & & A_1 & 0 \\ & & H_2 & B_2^T & & & & 0 & 0 \\ & & B_2 & 0 & & & & A_2 & 0 \\ & & & & \ddots & & & \vdots & \vdots \\ & & & & & H_S & B_N^T & 0 & 0 \\ & & & & & B_N & 0 & A_S & 0 \\ 0 & A_1^T & 0 & A_2^T & \dots & 0 & A_S^T & H_0 & A_0^T \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 & A_0 & 0 \end{bmatrix}$$

S is the number of scenarios

Special Structure of KKT System (Arrow-shaped)

$$\begin{bmatrix} K_1 & & & B_1 \\ & \ddots & & \vdots \\ & & K_N & B_N \\ B_1^T & \dots & B_N^T & K_0 \end{bmatrix} \begin{bmatrix} \Delta z_1 \\ \vdots \\ \Delta z_N \\ \Delta z_0 \end{bmatrix} = \begin{bmatrix} r_1 \\ \vdots \\ r_N \\ r_0 \end{bmatrix}$$

where,

$$K_i := \begin{bmatrix} \bar{Q}_i & W_i^T \\ W_i & 0 \end{bmatrix}, \qquad K_0 := \begin{bmatrix} \bar{Q} & A^T \\ A & 0 \end{bmatrix},$$

$$B_i := \begin{bmatrix} 0 & 0 \\ T_i & 0 \end{bmatrix}, \qquad i = 1, 2, \dots, N.$$

Block Elimination

$$\begin{bmatrix} K_1 & & & B_1 \\ & \ddots & & \vdots \\ & & K_N & B_N \\ B_1^T & \dots & B_N^T & K_0 \end{bmatrix} \begin{bmatrix} \Delta z_1 \\ \vdots \\ \Delta z_N \\ \Delta z_0 \end{bmatrix} = \begin{bmatrix} r_1 \\ \vdots \\ r_N \\ r_0 \end{bmatrix}$$

Multiply row *i* by $-B_i^T K_i^{-1}$ and sum all the rows to obtain

$$\left(K_0 - \sum_{i=1}^{N} B_i^T K_i^{-1} B_i \right) \Delta z_0 = r_0 - \sum_{i=1}^{N} B_i^T K_i^{-1} r_i$$

The matrix $C := K_0 - \sum_{i=1}^{N} B_i^T K_i^{-1} B_i$ is the Schur-complement of the diagonal $K_1, \dots, K_N$ block.

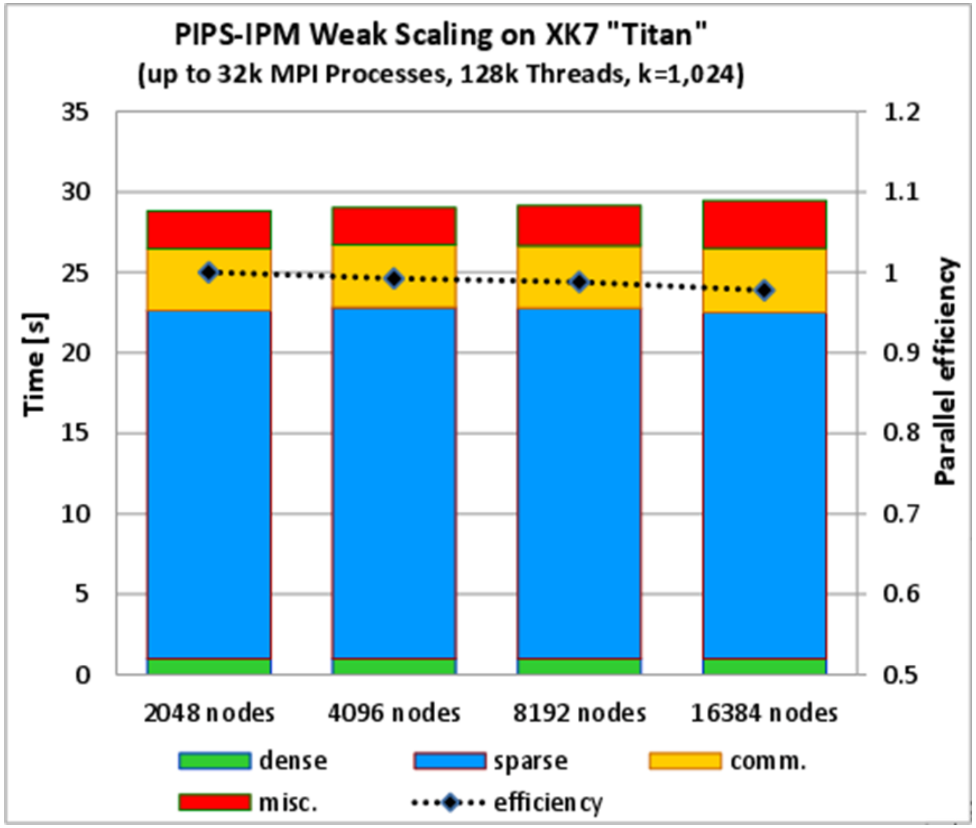
Parallel Solution Procedure for KKT System

1. Calculate $B_i^T K_i^{-1} B_i,\ i = 1, \ldots, N$ ("Compute S.C.")
2. Form $C := K_0 - \sum_{i=1}^N B_i^T K_i^{-1} B_i$ ("Form S.C.")
3. Factorize $C = L_0 D_0 L_0^T$ ("Factor S.C.")
4. Solve $\Delta z_0 = C^{-1}(r_0 - \sum_{i=1}^N B_i^T K_i^{-1} r_i)$
5. Solve $\Delta z_i = K_i^{-1}(B_i \Delta z_0 - r_i),\ i = 1, \ldots, N$

- Steps 1 and 5 trivially parallel
  - "Scenario-based decomposition"
- Extra care needed for computational bottlenecks 2, 3, and 5: multithreaded or GPU accelerated linear algebra, tuned communication, etc.
- Realtime is achieved using with an augmented incomplete factorization coupled with BiCGStab (to speed-up 1.).

PIPS performance (Petra et al., 2013)

- C++ code, MPI+OpenMP, runs on a variety of high performance computing platforms: “MIRA” IBM BG/P-Q (Argonne), “Titan” Cray XK7 (Oak Ridge), “Piz Daint” XC30 (Swiss National Computing Centre)



The largest instance has 4.08 billion decision variables and 4.12 billion constraints.

Simulations of State of Illinois' power grid

- The network consists of 2522 lines, 1908 buses, 870 demand buses, 225 generators, of which 32 are wind farms.
- Some of the wind farms are hypothetical and replace coal generators.
- Wind "installed" capacity is 17%. Adoption is around 15%.

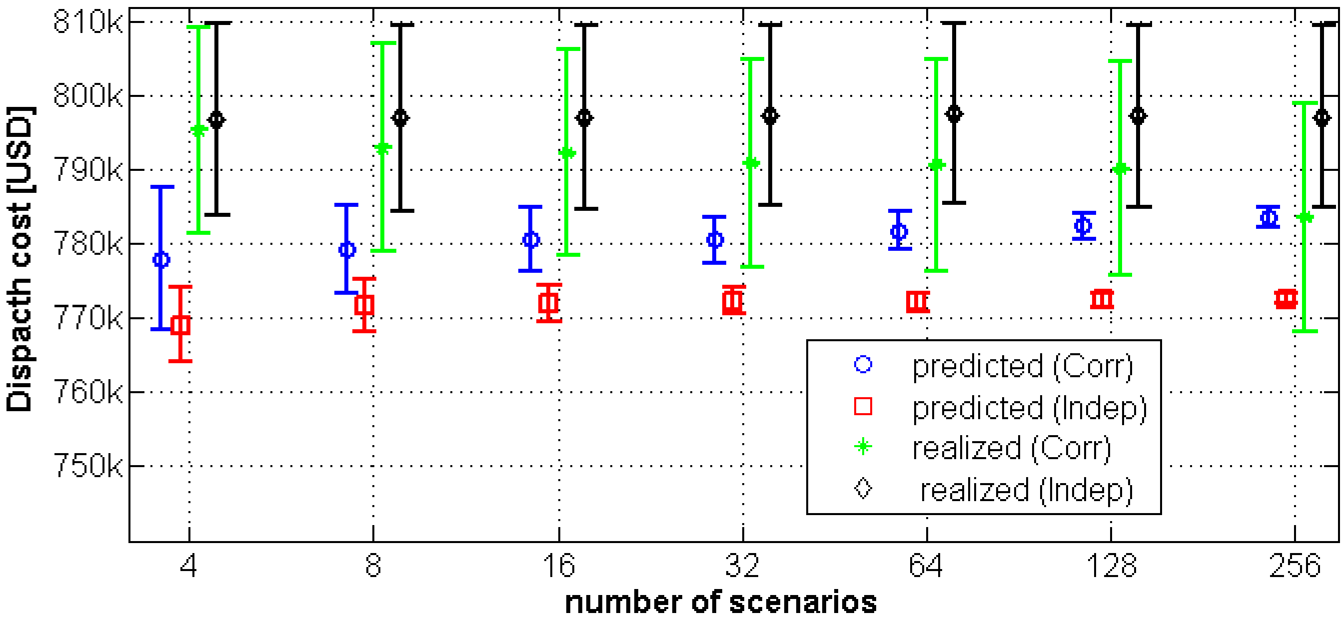
- RBLW covariance matrix ("corr.") vs diagonal covariance matrix ("indep.")
  - Dispatch cost
  - Ahead/realtime prices

- Both dispatch cost and the prices are random under the resampling scheme, therefore we compute confidence intervals.

- A problem with 256 scenarios has a little bit less than 1 million variables and 1 million constraints.

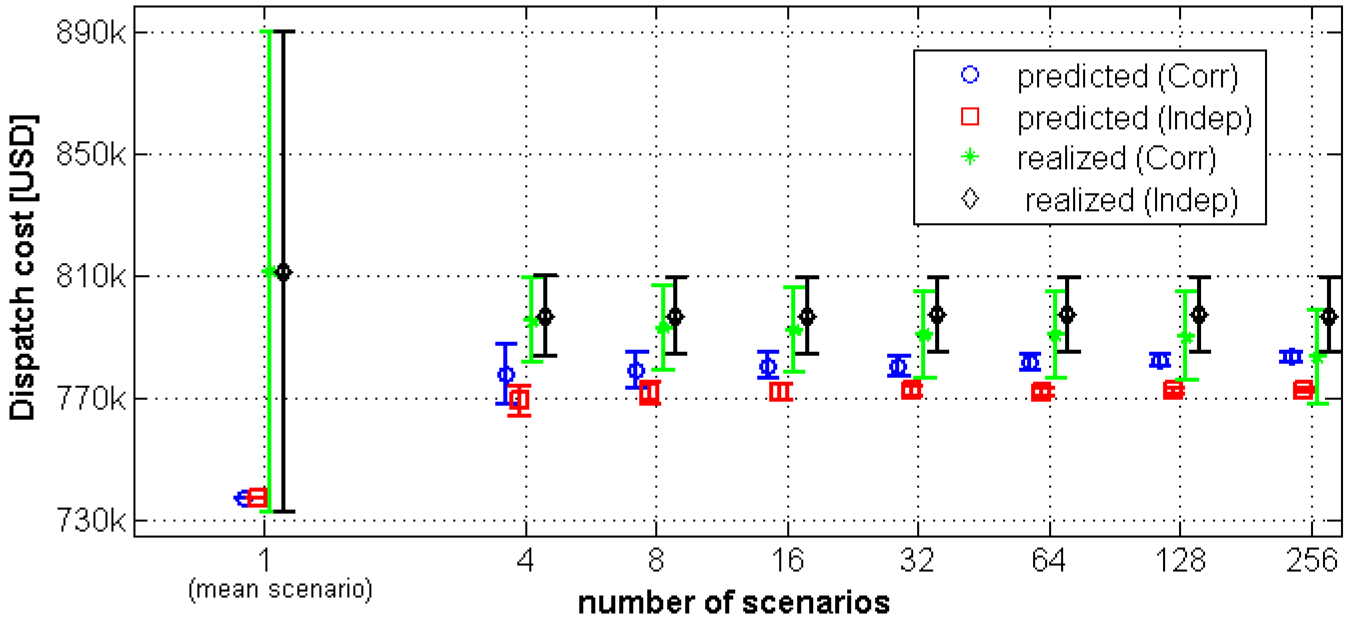
- Used Argonne's BG/P "Intrepid" and BG/Q "Mira" platforms for computing confidence intervals.

Dispatch cost – correlation vs independent resampling



- Gap of about $14,000 per hour or 1.6% ($170 million over a year).
- The gap does not seem to close as the number of scenarios increases.
- About 256 scenarios seem to offer a decent approximation (std. dev. is 0.36%)

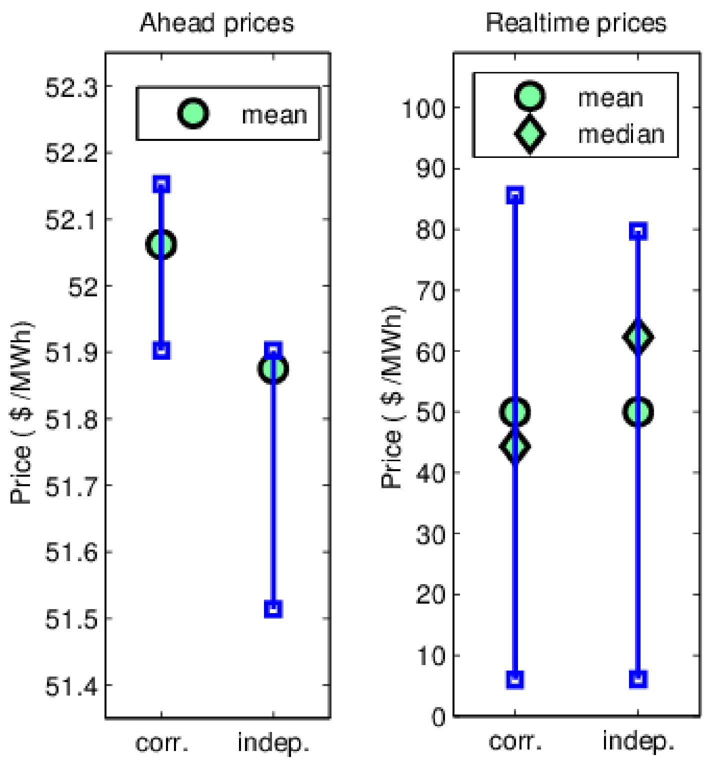
Dispatch cost – stochastic vs deterministic



Gap between stochastic and deterministic is $27,000 or 3.6%

Prices - correlation vs independent resampling

95% confidence intervals for prices (LMP) at a “typical” bus



Conclusions

- Improper correlation estimation may lead to inefficient pricing and higher dispatch costs, negatively impacting social welfare.

- We advocate for centralized weather forecasting in power grid dispatch.

- Better covariance estimation potentially leads to more efficient pricing/lower dispatch cost.

Details in **C. G. Petra, V. Zavala, E.D. Nino, M. Anitescu ,** "***Economic Impacts of Wind Covariance Estimation on Power Grid Operations***" (submitted to IEEE Power Systems)

Thank you for your attention!

Any questions?

Additional material

Empirical Bayesian estimator

- Again, n<<p

$$S_{EB} = \frac{p \cdot n - 2 \cdot n - 2}{p \cdot n^2} \cdot m_{EB} \cdot I + \frac{n}{n+1} \cdot S \in \mathbb{R}^{p \times p}, \text{ where}$$
$$m_{EB} = tr(S \cdot I)/p$$

Stein's SVD decomposition-based estimator

$$\mathbf{S} = \mathbf{U} \cdot \Sigma \cdot \mathbf{U}^{\mathrm{t}} ,$$

$$\Sigma = \mathbf{diag}\left(\sigma_1, \sigma_2, \ldots, \sigma_p\right) \in \Re^{p \times p} .$$

$$\hat{\sigma}_i = \frac{n \cdot \sigma_i}{n - p + 1 + 2 \cdot \sigma_i \cdot \sum_{k=1, k \neq i}^{p} \frac{1}{\sigma_i - \sigma_k}} , i = 1, 2, \ldots, p ,$$

$$\Lambda = \mathbf{diag}\left(\hat{\sigma}_1, \hat{\sigma}_2, \ldots, \hat{\sigma}_p\right) \in \Re^{p \times p} ,$$

$$\mathbf{S}_{\mathrm{SH}} = \mathbf{U} \cdot \Lambda \cdot \mathbf{U}^{\mathrm{t}} \in \Re^{p \times p} .$$