Real-time control of network physical structures to bypass complexity: Optimization, Stochastics and Structure Recognition

D. Bienstock, J. Blanchet, V. Goyal and G. Iyengar

Columbia University, Stanford University

April 2019
Today: control enforces separation by time domain
e.g. in power grids: governor reaction \((10^{-3} \text{ sec})\), AGC (sec), OPF (mins)

Opportunity: fast sensors, algorithms
Challenges: “smart” loads, complex noise

**Research Goals:**
- Avoid separation in time domain
- Quickly recognize system structure.
- Quickly detect intrusion. **Time frame**: seconds or very few minutes
Modern power grid control: a three-tier system

1. **Top tier: OPF (Optimal Power Flow).**
   Compute generator outputs so as to safely meet projected demand levels over the next time window. Run every five minutes.

   Note: uses **estimated demands**.

   Ideally, a **QCQP** (quadratically constrained quadratic program).

   But computationally complex, plus noisy data, so normally a linearly-constrained relaxation gets used.

2. **Shorter time frame: AGC (area generation control) (seconds), inertia and governor action (much less than 1 second).**

3. “**Frequency response**” (primarily inertia) corrects, in real-time, misestimations inherent in OPF.

4. AGC supports OPF by partly correcting inadequate generation.
Traditional generation
Control centers, RTUs, PMUs, state estimation
Control centers, RTUs, PMUs

- Control center performs a regulatory and economic role
- Sensors report to control center
- Control center issues commands to (in particular) smaller generators
- Sensors: RTUs (old), PMUs (new – and more expensive)
- RTUs report once every four seconds
- PMUs report once every four seconds
- PMUs report 30 to 100 times a second
- PMUs report (AC) voltage and current (plus more ...)
- Anecdotal: PMUs overwhelming human operators
- But PMUs are the way of the future
State estimation (very abridged)

A data-driven procedure to estimate relevant grid parameters

- Even with PMUs, data can be “complex”
- Statistical procedure: “state estimation” (at control center)

DC power flow equations:

\[ B\theta = P^g - P^d \]

- Sensors provide information that fit some of the \( \theta, P^d, (P^g?) \) parameters
- State estimation: least squares procedure to estimate the rest, plus more
AC-OPF problem

\[\text{Min} \sum_k C_k(P^g_k) \quad (1a)\]

Subject to:

\[\sum_{km \in \delta(k)} V_k y^*_m (V^*_k - V^*_m) = (P^g_k - P^d_k) + j(Q^g_k - Q^d_k), \quad \forall k \quad (1b)\]

\[P^\text{min}_k \leq P^g_k \leq P^\text{max}_k, \quad Q^\text{min}_k \leq Q^g_k \leq Q^\text{max}_k, \quad V^\text{min}_k \leq |V_k| \leq V^\text{max}_k \quad \forall k \quad (1c)\]

\[|\theta_k - \theta_m| \leq \theta^\text{max}_{km} \quad \forall km \quad (1d)\]

\[|V_k y^*_m (V^*_k - V^*_m)| \leq L_{km} \quad \forall km. \quad (1e)\]

Here, at a bus (node) \( k \), \( V_k = |V_k|e^{i\theta_k} \) = voltage of \( k \); \( j = \sqrt{-1} \).

What is the best solver?
AC-OPF problem

Min \sum_k C_k(P^g_k) \quad (1a)

Subject to:

\sum_{km \in \delta(k)} V_k y_{km}^* (V_k^* - V_m^*) \quad \begin{cases} \geq (P^g_k - P^d_k) + j(Q^g_k - Q^d_k), & \forall k \end{cases} \quad (1b)

P_{k}^{\min} \leq P^g_k \leq P_{k}^{\max}, \quad Q_{k}^{\min} \leq Q^g_k \leq Q_{k}^{\max}, \quad V_{k}^{\min} \leq |V_k| \leq V_{k}^{\max} \quad \forall k \quad (1c)

|\theta_k - \theta_m| \leq \theta_{km}^{\max} \quad \forall km \quad (1d)

|V_k y_{km}^* (V_k^* - V_m^*)| \leq L_{km} \quad \forall km. \quad (1f)

Here, at a bus (node) k, \( V_k = |V_k| e^{j \theta_k} = \text{voltage of } k; j = \sqrt{-1}. \)

What is the best solver? King of the hill: log barrier methods (IPOPT and KNITRO, others)
Let’s expand on this

- Log-barrier methods do not guarantee ...
Let’s expand on this

- Log-barrier methods do not guarantee ... anything.

- But if they converge, they converge to a local optimum (of the log-barrier function).

- **But:** they use Newton’s method. On ACOPF, convergence is usually quite quick, to a solution of excellent quality.

- How do we know? Jabr’s SOCP relaxation proves it.

- Does not mean that Jabr’s relaxation is also very good? No, because it operates in a different space – does not yield a usable solution.
What happens when there is a generation/load mismatch
What happens when there is a generation/load mismatch

Frequency response:

\[ \omega \approx -c \Delta P \]
What happens when there is a generation/load mismatch

Frequency response:

mismatch $\Delta P$
What happens when there is a generation/load mismatch

Frequency response:

mismatch $\Delta P \Rightarrow$ frequency change $\Delta \omega \approx -c \Delta P$
AGC, primary and secondary response (simplified!, abridged!)

Suppose generation vs loads balance spontaneously changes (i.e. a net imbalance)?

- AC frequency changes proportionally (to first order) near-instantaneously
AGC, primary and secondary response (simplified!, abridged!)

Suppose generation vs loads balance spontaneously changes (i.e. a net imbalance)?

- AC frequency changes proportionally (to first order) near-instantaneously
- **Primary response.** (very quick) Inertia in generators contributes electrical energy to the system
AGC, primary and secondary response (simplified!, abridged!)

Suppose generation vs loads balance spontaneously changes (i.e. a net imbalance)?

- AC frequency changes proportionally (to first order) near-instantaneously
- **Primary response.** (very quick) Inertia in generators contributes electrical energy to the system
- **Secondary response (AGC).** (seconds) If estimated generation shortfall $= \Delta P$. Then:
AGC, primary and secondary response (simplified!, abridged!)

Suppose generation vs loads balance spontaneously changes (i.e. a net imbalance)?

- AC frequency changes proportionally (to first order) near-instantaneously
- **Primary response.** (very quick) Inertia in generators contributes electrical energy to the system
- **Secondary response (AGC).** (seconds) If estimated generation shortfall $= \Delta P$. Then:
  - Each generator $g$ changes output by $\alpha_g$ (and $\sum_g \alpha_g = 1$)
- **OPF.** (minutes) Reset large generator output to minimize cost using estimations of loads.
What is the future?

- Shorten the control loop – run ACOPF more frequently.
- Technical question: how do we correct a solution to a QCQP, i.e. how do we correct the solution to (quickly) go to a local optimum? Can we patch a logarithmic barrier method?
- Blurs the line between OPF and AGC. More reactors respond more frequently to signals.
- Need to watch our for noise in the data.
- However part of the goal is to reduce variability in system operation.
- Make computations variance-aware. Question: How do we estimate covariance matrices quickly, under changing stochastic conditions?
- Recognize unusual data conditions.
Three research problems

1. **First-order** methods to patch (correct) log-barrier algorithms for QCQPs

2. **Variance-aware** optimal power flow problems

3. **Real time** estimation of covariance (streaming data, variable stochastics)

4. **Application** to intrusion detection
“Cyber-physical” attacks on power grids

An adversary carries out a physical alteration of a grid (example: disconnecting a power line).

This is coordinated with a modification of sensor signals – a hack.

The goal is to disguise, or keep completely hidden, the nature of the attack and its likely consequences.

Power industry: it will never happen (“we would know what happened”).

Really?

BBGI (Columbia University, Stanford University)

SR Physics

April 2019 14 / 31
“Cyber-physical” attacks on power grids

Fact or fiction?

1. An adversary carries out a physical alteration of a grid (example: disconnecting a power line)
“Cyber-physical” attacks on power grids

Fact or fiction?

1. An adversary carries out a physical alteration of a grid (example: disconnecting a power line)

2. This is coordinated with a modification of sensor signals – a hack
“Cyber-physical” attacks on power grids

Fact or fiction?

1. An adversary carries out a physical alteration of a grid (example: disconnecting a power line)

2. This is coordinated with a modification of sensor signals – a hack

3. The goal is to disguise, or keep completely hidden, the nature of the attack and its likely consequences
“Cyber-physical” attacks on power grids

Fact or fiction?

1. An adversary carries out a physical alteration of a grid (example: disconnecting a power line)
2. This is coordinated with a modification of sensor signals – a hack
3. The goal is to disguise, or keep completely hidden, the nature of the attack and its likely consequences
4. Power industry: it will never happen (“we would know what happened”)

“Cyber-physical” attacks on power grids

Fact or fiction?

1. An adversary carries out a physical alteration of a grid (example: disconnecting a power line)

2. This is coordinated with a modification of sensor signals – a hack

3. The goal is to disguise, or keep completely hidden, the nature of the attack and its likely consequences

4. Power industry: it will never happen (“we would know what happened”)

5. Really?
More detail:

- An attacker causes physical changes in the network:
More detail:

- An attacker causes physical changes in the network:
  - In particular **load** changes (no generator changes)
  - Possibly also **line disconnections**
More detail:

- An attacker causes physical changes in the network:
  - In particular **load** changes (no generator changes)
  - Possibly also **line disconnections**

- Attacker also hacks the signal flow: the output of some sensors is altered

- Goal of the attacker is twofold:
  1. Hide the location of the attack and even the fact that an attack happened
  2. **Cause line overloads that remain hidden**
An attacker causes physical changes in the network:

- In particular **load** changes (no generator changes)
- Possibly also **line disconnections**

Attacker also hacks the signal flow: the output of some sensors is altered

Goal of the attacker is twofold:

1. **Hide the location of the attack and even the fact that an attack happened**
2. **Cause line overloads that remain hidden**

Attacker expects **full PMU deployment**.

Nonlinear optimization task solved by attacker!
A large-scale example: case2746wp

(2746 buses, 3514 lines, 520 generators, 25GW total load)
A large-scale example: case2746wp

(2746 buses, 3514 lines, 520 generators, 25GW total load)

Undetectable attack with strong overloads on branches:

(1361, 1141):
\[ \| \text{reported flow} \| = 109, \| \text{true flow} \| = 229, \text{limit} = 114 \]

(1138, 1141):
\[ \| \text{reported flow} \| = 98, \| \text{true flow} \| = 209, \text{limit} = 114 \]

Net load change: 135 MW (\(< 0.5\%\)) of total load
Another example, from case1354pegase
Non-static attack: follow-up

A blind spot in prior work?
Non-static attack: follow-up

A blind spot in prior work? **Notation:** $\mathcal{A} =$ attacked zone

**“Noisy-data attack”**

After the attack, for any bus in $\mathcal{A}$ the attacker reports (at time $t$) a complex voltage value

$$\tilde{V}_k(t) = V^R_k + \nu_k(t)$$

Here, $\nu_k(t)$ is *random*, with

$$\mathbb{E}(\nu_k(t)) = 0,$$
Non-static attack: follow-up

A blind spot in prior work? **Notation:** \( \mathcal{A} = \) attacked zone

**“Noisy-data attack”**

After the attack, for any bus in \( \mathcal{A} \) the attacker reports (at time \( t \)) a complex voltage value

\[
\tilde{V}_k(t) = V_k^R + \nu_k(t)
\]

Here, \( \nu_k(t) \) is random, with

\[
E(\nu_k(t)) = 0,
\]

(consistent with zero expected load change)
Non-static attack: follow-up

A blind spot in prior work? **Notation:** $\mathcal{A} =$ attacked zone

**“Noisy-data attack”**

After the attack, for any bus in $\mathcal{A}$ the attacker reports (at time $t$) a complex voltage value

$$\tilde{V}_k(t) = V_k^R + \nu_k(t)$$

Here, $\nu_k(t)$ is *random*, with

$$E(\nu_k(t)) = 0,$$

(consistent with zero expected load change)
Non-static attack: follow-up

A blind spot in prior work? **Notation:** $\mathcal{A} =$ attacked zone

**“Noisy-data attack”**

After the attack, for any bus in $\mathcal{A}$ the attacker reports (at time $t$) a complex voltage value

$$\tilde{V}_k(t) = V_k^R + \nu_k(t)$$

Here, $\nu_k(t)$ is *random*, with

$$\mathbb{E}(\nu_k(t)) = 0,$$

(consistent with zero expected load change)

and?
Non-static attack: follow-up

A blind spot in prior work? **Notation:** $\mathcal{A} = \text{attacked zone}$

**“Noisy-data attack”**

After the attack, for any bus in $\mathcal{A}$ the attacker reports (at time $t$) a complex voltage value

$$\tilde{V}_k(t) = V_k^R + \nu_k(t)$$

Here, $\nu_k(t)$ is *random*, with

$$E(\nu_k(t)) = 0,$$

(consistent with zero expected load change)

and? what else?
Non-static attack: follow-up

A blind spot in prior work? **Notation:** \( \mathcal{A} \) = attacked zone

**“Noisy-data attack”**

After the attack, for any bus in \( \mathcal{A} \) the attacker reports (at time \( t \)) a complex voltage value

\[
\tilde{V}_k(t) = V^R_k + \nu_k(t)
\]

Here, \( \nu_k(t) \) is random, with

\[
E(\nu_k(t)) = 0,
\]

(consistent with zero expected load change)

and? what else?
Defense, 0

- Defender is likely to know that “something” happened (and quickly). But sensor data is noisy and “something” may be inconsequential.
Defense, 0

- Defender is likely to know that “something” happened (and quickly). But sensor data is noisy and “something” may be inconsequential.

- We want a defensive action that is nearly implementable in terms of today’s grid operation.
Defense, 0

- Defender is likely to know that “something” happened (and quickly). But sensor data is noisy and “something” may be inconsequential.

- We want a defensive action that is nearly implementable in terms of today’s grid operation.

- Should not lead to false positives.

Solution: change the power quantities in a way that the attacker cannot anticipate, and identify inconsistent signals. How? A solution: change generator output by a random injection that yields a valid power flow solution ("AGC-lite" plus redispatch).
Defender is likely to know that “something” happened (and quickly). But sensor data is noisy and “something” may be inconsequential.

We want a defensive action that is nearly implementable in terms of today’s grid operation.

Should not lead to false positives.

Solution: change the power quantities in a way that the attacker cannot anticipate, and identify inconsistent signals. How?
Defense, 0

- Defender is likely to know that “something” happened (and quickly). But sensor data is noisy and “something” may be inconsequential.

- We want a defensive action that is nearly implementable in terms of today’s grid operation.

- Should not lead to false positives.

- **Solution:** change the power quantities in a way that the attacker cannot anticipate, and identify inconsistent signals. How?

- A solution: change generator output by a random injection that yields a valid power flow solution (“AGC-lite” plus redispatch).
Defense, 0’ (optimization problem)

Following attack, and in suspicion of an attack

- Defender only has access to **reported** data, which is accurate in the non-attacked zone. But the defender **does not** know the attacked zone.
Defense, 0’ (optimization problem)

Following attack, and in suspicion of an attack

- Defender only has access to **reported** data, which is accurate in the non-attacked zone. But the defender **does not** know the attacked zone.

- *(repeatedly)* Defender chooses a random subset of the AGC-responding generators, and

Defender computes a random power flow injection at those generators. Defender seeks to cause large, random changes to phase angles. Attacker cannot anticipate this random action, leading to inconsistencies in falsified data.
Defense, 0’ (optimization problem)

Following attack, and in suspicion of an attack

- Defender only has access to reported data, which is accurate in the non-attacked zone. But the defender does not know the attacked zone.

- (repeatedly) Defender chooses a random subset of the AGC-responding generators, and

- Defender computes a random power flow injection at those generators.

- Defender seeks to cause large, random changes to phase angles. Attacker cannot anticipate this random action, leading to inconsistencies in falsified data.
Random defense (simple example):

Assumes that there is a set of **safe** or **trusted** buses.
Random defense (simple example):

Assumes that there is a set of safe or trusted buses.

Repeat:

1. Pick a random pair \( s, t \) of trusted buses.
2. Pick a random value \( \Gamma > 0 \) (random under some distribution)
3. Cause an injection of \( \Gamma \) at \( s \) and \( -\Gamma \) at \( t \).
Random defense (simple example):

Assumes that there is a set of safe or trusted buses.

Repeat:

1. Pick a random pair $s, t$ of trusted buses.
2. Pick a random value $\Gamma > 0$ (random under some distribution)
3. Cause an injection of $\Gamma$ at $s$ and $-\Gamma$ at $t$.

**Theorem 1:** with high probability, the phase angle of every bus will be shifted, repeatedly.
Random defense (simple example):

Assumes that there is a set of safe or trusted buses.

Repeat:

1. Pick a random pair $s, t$ of trusted buses.
2. Pick a random value $\Gamma > 0$ (random under some distribution)
3. Cause an injection of $\Gamma$ at $s$ and $-\Gamma$ at $t$.

**Theorem 1:** with high probability, the phase angle of every bus will be shifted, repeatedly.

**Theorem 2:** The rank of the covariance matrix will be increased.
Random defense (simple example):

Assumes that there is a set of safe or trusted buses.

Repeat:

1. Pick a random pair $s, t$ of trusted buses.
2. Pick a random value $\Gamma > 0$ (random under some distribution)
3. Cause an injection of $\Gamma$ at $s$ and $-\Gamma$ at $t$.

Theorem 1: with high probability, the phase angle of every bus will be shifted, repeatedly.

Theorem 2: The rank of the covariance matrix will be increased.
“Noisy-data attack”

After the attack, for any bus in $\mathcal{A}$ the attacker reports (at time $t$) a complex voltage value

$$\tilde{V}_k(t) = V_k^R + \nu_k(t)$$

Here, $\nu_k(t)$ is random, with

$$\mathbb{E}(\nu_k(t)) = 0,$$
“Noisy-data attack”

After the attack, for any bus in $\mathcal{A}$ the attacker reports (at time $t$) a complex voltage value

$$\tilde{V}_k(t) = V_k^R + \nu_k(t)$$

Here, $\nu_k(t)$ is random, with

$$\mathbb{E}(\nu_k(t)) = 0,$$

(consistent with zero expected load change)
“Noisy-data attack”

After the attack, for any bus in $\mathcal{A}$ the attacker reports (at time $t$) a complex voltage value

$$\tilde{V}_k(t) = V_k^R + \nu_k(t)$$

Here, $\nu_k(t)$ is random, with

$$\mathbb{E}(\nu_k(t)) = 0,$$

(consistent with zero expected load change)
“Noisy-data attack”

After the attack, for any bus in $A$ the attacker reports (at time $t$) a complex voltage value

$$\tilde{V}_k(t) = V_k^R + \nu_k(t)$$

Here, $\nu_k(t)$ is random, with

$$E(\nu_k(t)) = 0,$$

(consistent with zero expected load change)

and?

---

BBGI (Columbia University, Stanford University)
SR Physics
April 2019 22 / 31
“Noisy-data attack”

After the attack, for any bus in $\mathcal{A}$ the attacker reports (at time $t$) a complex voltage value

$$\tilde{V}_k(t) = V_k^R + \nu_k(t)$$

Here, $\nu_k(t)$ is random, with

$$E(\nu_k(t)) = 0,$$

(consistent with zero expected load change)

and? what else?
“Noisy-data attack”

After the attack, for any bus in $\mathcal{A}$ the attacker reports (at time $t$) a complex voltage value

$$\tilde{V}_k(t) = V_k^R + \nu_k(t)$$

Here, $\nu_k(t)$ is random, with

$$\mathbb{E}(\nu_k(t)) = 0,$$

(consistent with zero expected load change)

and? what else?

$\rightarrow$ covariance of $\nu(t)$ should be make sense
PMU fun

We have data from an industrial partner:

- 240 PMUs
- 2 years of reported data
- 28 TB
- Soon, 500 PMUs and higher detail
Covariances matrices of PMU data have low rank!!
Covariances matrices of PMU data have low rank!!

Example: 50 PMUs, Voltage Angle, one minute

<table>
<thead>
<tr>
<th>Scaled Eigenvalue</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>4</td>
</tr>
<tr>
<td>5</td>
</tr>
<tr>
<td>6</td>
</tr>
<tr>
<td>7</td>
</tr>
<tr>
<td>8</td>
</tr>
<tr>
<td>9</td>
</tr>
</tbody>
</table>

| 1.000 | 0.078 | 0.012 | 0.009 | 0.007 | 0.004 | 0.003 | 0.002 | 0.001 |
Covariances matrices of PMU data have **low rank!!**

<table>
<thead>
<tr>
<th>Scaled Eigenvalue</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.000</td>
</tr>
<tr>
<td>2</td>
<td>0.618</td>
</tr>
<tr>
<td>3</td>
<td>0.061</td>
</tr>
<tr>
<td>4</td>
<td>0.023</td>
</tr>
<tr>
<td>5</td>
<td>0.017</td>
</tr>
<tr>
<td>6</td>
<td>0.010</td>
</tr>
<tr>
<td>7</td>
<td>0.008</td>
</tr>
<tr>
<td>8</td>
<td>0.004</td>
</tr>
<tr>
<td>9</td>
<td>0.004</td>
</tr>
<tr>
<td>10</td>
<td>0.002</td>
</tr>
</tbody>
</table>

What else? All data seems to be light-tailed.

BBGI (Columbia University, Stanford University)
SR Physics
April 2019 25 / 31
Covariances matrices of PMU data have **low rank!!**

**Example: 100 PMUs, voltage magnitude, five minutes**

<table>
<thead>
<tr>
<th></th>
<th>Scaled Eigenvalue</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.000</td>
</tr>
<tr>
<td>2</td>
<td>0.618</td>
</tr>
<tr>
<td>3</td>
<td>0.061</td>
</tr>
<tr>
<td>4</td>
<td>0.023</td>
</tr>
<tr>
<td>5</td>
<td>0.017</td>
</tr>
<tr>
<td>6</td>
<td>0.010</td>
</tr>
<tr>
<td>7</td>
<td>0.008</td>
</tr>
<tr>
<td>8</td>
<td>0.004</td>
</tr>
<tr>
<td>9</td>
<td>0.004</td>
</tr>
<tr>
<td>10</td>
<td>0.002</td>
</tr>
</tbody>
</table>

What else? All data seems to be light-tailed.
Covariances matrices of PMU data have **low rank!!**

**Example: 100 PMUs, voltage magnitude, five minutes**

<table>
<thead>
<tr>
<th></th>
<th>Scaled Eigenvalue</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.000</td>
</tr>
<tr>
<td>2</td>
<td>0.618</td>
</tr>
<tr>
<td>3</td>
<td>0.061</td>
</tr>
<tr>
<td>4</td>
<td>0.023</td>
</tr>
<tr>
<td>5</td>
<td>0.017</td>
</tr>
<tr>
<td>6</td>
<td>0.010</td>
</tr>
<tr>
<td>7</td>
<td>0.008</td>
</tr>
<tr>
<td>8</td>
<td>0.004</td>
</tr>
<tr>
<td>9</td>
<td>0.004</td>
</tr>
<tr>
<td>10</td>
<td>0.002</td>
</tr>
</tbody>
</table>

What else? All data seems to be **light**-tailed
Learning variances

**Theorem.** (Co)variance of time series can be learned

- In real time
- In streaming fashion
- Under evolving stochasticity

Want to estimate covariance of a time series $x_t$, under streaming conditions and changing stochastics

Change happens at time $t = 0$, must learn by time $t = T$. 

Spiked covariance model:

$$x_t = A_t z_t + \omega_t$$

$A_t$ is a small $r \times r$ matrix (is small)
$z_t$ is a light-tailed random variable (gaussian or sub-gaussian)
$\omega_t$ is "noise", small, uncorrelated with $z_t$

Changing stochastics:

$$\|A_t A_t^\top - A_{t-1} A_{t-1}^\top\|_2 \leq \gamma, \forall t = 2 : T$$

"Learning" covariance means: estimating the eigenspace of the covariance

Theorem: optimal error estimation grows like $O(\gamma^{1/3})$.

Adaptation of noisy power method.
Want to estimate covariance of a time series \( x_t \), under streaming conditions and changing stochastics.

Change happens at time \( t = 0 \), must learn by time \( t = T \).

**Spiked covariance** model: \( x_t = A_t z_t + \omega_t \)
Want to estimate covariance of a time series $x_t$, under streaming conditions and changing stochastics

Change happens at time $t = 0$, must learn by time $t = T$.

**Spiked covariance** model: $x_t = A_t z_t + \omega_t$

- $A_t$ a $n \times r$ matrix ($r$ is small)
- $z_t$ a light-tailed random variable (gaussian or sub-gaussian)
- $\omega_t$ is “noise”, small, uncorrelated with $z_t$

Changing stochastics: $\|A_t A_t^T - A_{t-1} A_{t-1}^T\|_2 \leq \gamma, \forall t = 2 : T$
Want to estimate covariance of a time series \( x_t \), under streaming conditions and changing stochastics.

Change happens at time \( t = 0 \), must learn by time \( t = T \).

**Spiked covariance** model: \( x_t = A_t z_t + \omega_t \)

- \( A_t \) a \( n \times r \) matrix (\( r \) is small)
- \( z_t \) a light-tailed random variable (gaussian or sub-gaussian)
- \( \omega_t \) is “noise”, small, uncorrelated with \( z_t \)

Changing stochastics: \( \|A_t A_t^T - A_{t-1} A_{t-1}^T\|_2 \leq \gamma, \forall t = 2 : T \)

“Learning” covariance means: estimating the **eigenspace** of the covariance.
Want to estimate covariance of a time series $x_t$, under streaming conditions and changing stochastics.

Change happens at time $t = 0$, must learn by time $t = T$.

**Spiked covariance** model: $x_t = A_t z_t + \omega_t$

- $A_t$ a $n \times r$ matrix ( $r$ is small)
- $z_t$ a light-tailed random variable (gaussian or sub-gaussian)
- $\omega_t$ is “noise”, small, uncorrelated with $z_t$

Changing stochastics: $\| A_t A_t^\top - A_{t-1} A_{t-1}^\top \|_2 \leq \gamma$, $\forall \ t = 2 : T$

“Learning” covariance means: estimating the **eigenspace** of the covariance

**Theorem:** optimal error estimation grows like $O(\gamma^{1/3})$.

Adaptation of noisy power method.
Covariance defense

- Under **whatever** assumptions, the attacker will produce a time series for e.g. phase angles.
Covariance defense

- Under **whatever** assumptions, the attacker will produce a time series for e.g. phase angles.
- Assume covariance of phase angles is learned by the defender.
Covariance defense

- Under **whatever** assumptions, the attacker will produce a time series for e.g. phase angles.
- Assume covariance of phase angles is learned by the defender
- (Assume of low rank)
Covariance defense

- Under **whatever** assumptions, the attacker will produce a time series for e.g. phase angles.
- Assume covariance of phase angles is learned by the defender
- (Assume of low rank)
- Defender chooses *random generator injections so as to significantly change covariance of phase angles*
Covariance defense

- Under **whatever** assumptions, the attacker will produce a time series for e.g. phase angles.

- Assume covariance of phase angles is learned by the defender

- (Assume of low rank)

- Defender chooses *random generator injections* so as to **significantly change covariance of phase angles**

- Attacker is caught with pants down

- Note: this is an adaptation of **AGC**.
Covariance defense (technical, abridged)

- Let $\Omega =$ covariance of *observed* voltage phase angles
Covariance defense (technical, abridged)

- Let \( \Omega \) = covariance of observed voltage phase angles.

- Let \( w_1, w_2, \ldots, w_r \) = eigenvectors with large enough eigenvalues.

Theorem: there is a random set of power injections (by generators) that results in covariance of phase angles \( \approx \Omega + \lambda vv^T \) where \( \lambda > 0 \).
Covariance defense (technical, abridged)

- Let \( \Omega \) = covariance of observed voltage phase angles
- Let \( w_1, w_2, \ldots, w_r \) = eigenvectors with large enough eigenvalues. \( r \ll n \) (number of buses)

Theorem: there is a random set of power injections (by generators) that results in covariance of phase angles \( \approx \Omega + \lambda vv^T \) where \( \lambda > 0 \)

On case 2746, \( \approx 10 \) vectors \( v \) cover all buses. (Dense null space vector computation: LP heuristic)
Covariance defense (technical, abridged)

- Let $\Omega = \text{covariance of observed voltage phase angles}$
- Let $w_1, w_2, \ldots, w_r = \text{eigenvectors with large enough eigenvalues}$. 
  $r \ll n$ (number of buses)
- Defender chooses vector $v \in \mathbb{R}^n$ with:
  
  $w_i^T v = 0$ for $1 \leq i \leq r$ (plus other conditions)
Covariance defense (technical, abridged)

- Let $\Omega =$ covariance of observed voltage phase angles

- Let $w_1, w_2, \ldots, w_r =$ eigenvectors with large enough eigenvalues. $r \ll n$ (number of buses)

- Defender chooses vector $v \in \mathbb{R}^n$ with:
  
  $w_i^T v = 0$ for $1 \leq i \leq r$ (plus other conditions)

- **Theorem:** there is a random set of power injections (by generators) that results in covariance of phase angles

  $\approx \Omega + \lambda vv^T$
Covariance defense (technical, abridged)

- Let $\Omega = \text{covariance of observed voltage phase angles}$

- Let $w_1, w_2, \ldots, w_r = \text{eigenvectors with large enough eigenvalues.}$
  \[ r \ll n \] (number of buses)

- Defender chooses vector $v \in \mathbb{R}^n$ with:
  \[ w_i^T v = 0 \text{ for } 1 \leq i \leq r \] (plus other conditions)

- **Theorem:** there is a random set of power injections (by generators) that results in covariance of phase angles
  \[ \approx \Omega + \lambda vv^T \text{ where } \lambda > 0 \]
Covariance defense (technical, abridged)

- Let $\Omega = \text{covariance of observed voltage phase angles}$

- Let $w_1, w_2, \ldots, w_r = \text{eigenvectors with large enough eigenvalues.}$  
  $r \ll n \ (\text{number of buses})$

- Defender chooses vector $v \in \mathbb{R}^n$ with:
  
  $w_i^T v = 0 \ \text{for} \ 1 \leq i \leq r \ \text{(plus other conditions)}$

- **Theorem:** there is a random set of power injections (by generators) that results in covariance of phase angles
  
  $\approx \Omega + \lambda vv^T \ \text{where} \ \lambda > 0$

- On case2746wp, $\approx 10$ vectors $v$ cover all buses.  
  (Dense null space vector computation: LP heuristic)
Covariance defense (technical, less abridged)

1. Let $\Omega = \text{covariance of observed voltage phase angles}$

2. Let $w_1, w_2, \ldots, w_r = \text{eigenvectors with large enough eigenvalues}$. 

Theorem: there is a random set of power injections (by generators) that results in covariance of phase angles $\approx \Omega + \lambda vv^T$ where $\lambda > 0$.

On case2746wp, there is a single vector $v$ that covers all buses.

Theorem: if $v_1, v_2 \in \text{subspace $S$}$, then $\exists \infty$ many $v \in S$ with $\text{support}(v) = \text{support}(v_1) \cup \text{support}(v_2)$. 

BBGI (Columbia University, Stanford University)  
SR Physics  
April 2019 30 / 31
Covariance defense (technical, less abridged)

1. Let $\Omega = \text{covariance of observed voltage phase angles}$

2. Let $w_1, w_2, \ldots, w_r = \text{eigenvectors with large enough eigenvalues.}$
   $r \ll n$ (number of buses)

3. Defender chooses vector $v \in \mathbb{R}^n$ with:
   
   $w_i^T v = 0$ for $1 \leq i \leq r$ and $[Bv]_i = 0$ for all non-generator $i$

4. **Theorem:** There is a random set of power injections (by generators) that results in covariance of phase angles
   
   $\approx \Omega + \lambda vv^T$
Covariance defense (technical, less abridged)

1. Let $\Omega = $ covariance of observed voltage phase angles.

2. Let $w_1, w_2, \ldots, w_r = $ eigenvectors with large enough eigenvalues.
   \[ r \ll n \] (number of buses)

3. Defender chooses vector $v \in \mathbb{R}^n$ with:
   \[ w_i^T v = 0 \text{ for } 1 \leq i \leq r \text{ and } [Bv]_i = 0 \text{ for all non-generator } i \]

4. **Theorem**: there is a random set of power injections (by generators) that results in covariance of phase angles
   \[ \approx \Omega + \lambda vv^T \] where $\lambda > 0$

5. On case2746wp, there is a single vector $v$ that covers all buses.
Covariance defense (technical, less abridged)

1. Let $\Omega = \text{covariance of observed voltage phase angles}$

2. Let $w_1, w_2, \ldots, w_r = \text{eigenvectors with large enough eigenvalues.}$ $r \ll n \ (\text{number of buses})$

3. Defender chooses vector $v \in \mathbb{R}^n$ with:
   
   $w_i^Tv = 0$ for $1 \leq i \leq r$ and $[Bv]_i = 0$ for all non-generator $i$

4. **Theorem:** there is a random set of power injections (by generators) that results in covariance of phase angles

   $\approx \Omega + \lambda vv^T$ where $\lambda > 0$

5. On case \texttt{2746wp}, there is a \textbf{single} vector $v$ that covers all buses.

   **Theorem:** if $v^1, v^2 \in \text{subspace } S$, then $\exists \infty \text{ many } v \in S$ with

   \[ \text{support}(v) = \text{support}(v^1) \cup \text{support}(v^2) \]
Publications and conference talks:

- Full paper on streaming variance estimation in preparation. (Initial version: NIPS TSW ’18)

- Full paper on variance-aware OPF (PSCC 2018), submitted.

- Full paper on detecting power grid attacks (SIAM Network Science ’18, ’19), submitted.
  https://arxiv.org/abs/1807.06707

- Full paper on sensor analysis methods (Powertech ’19, to appear).
  https://arxiv.org/abs/1811.07139