# Holistic Modeling and Analysis of Multistage Manufacturing Processes with Sparse Effective Inputs and Mixed Profile Outputs

Andi Wang

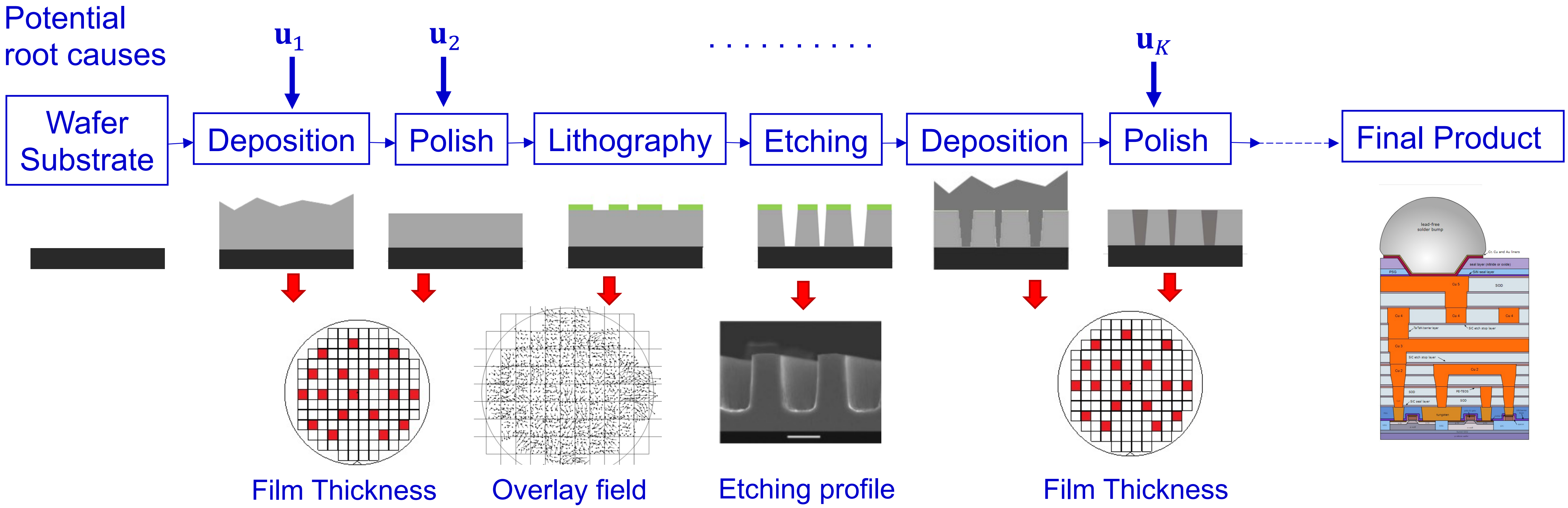
Arizona State University

*Website:* https://web.asu.edu/andi-wang *Email:* andi.wang@asu.edu

*August 6, 2021*

Wang, A., & Shi, J. (2021). Holistic modeling and analysis of multistage manufacturing processes with sparse effective inputs and mixed profile outputs. *IISE Transactions*, *53*(5), 582-596.

# Research Motivation




Multistage:

- Error propagation

Data-rich:

- Mixed profiles
- Redundant information in $\mathbf{u}_i$'s

## Problems to tackle

- Which potential root causes affect the quality measurements in the MMP?
- How those root causes lead to the variation of the quality measurements?

# Process and Data Characteristics

Cascading effects

- Potential causes may affect current and later stages

Sparsity of actual root causes

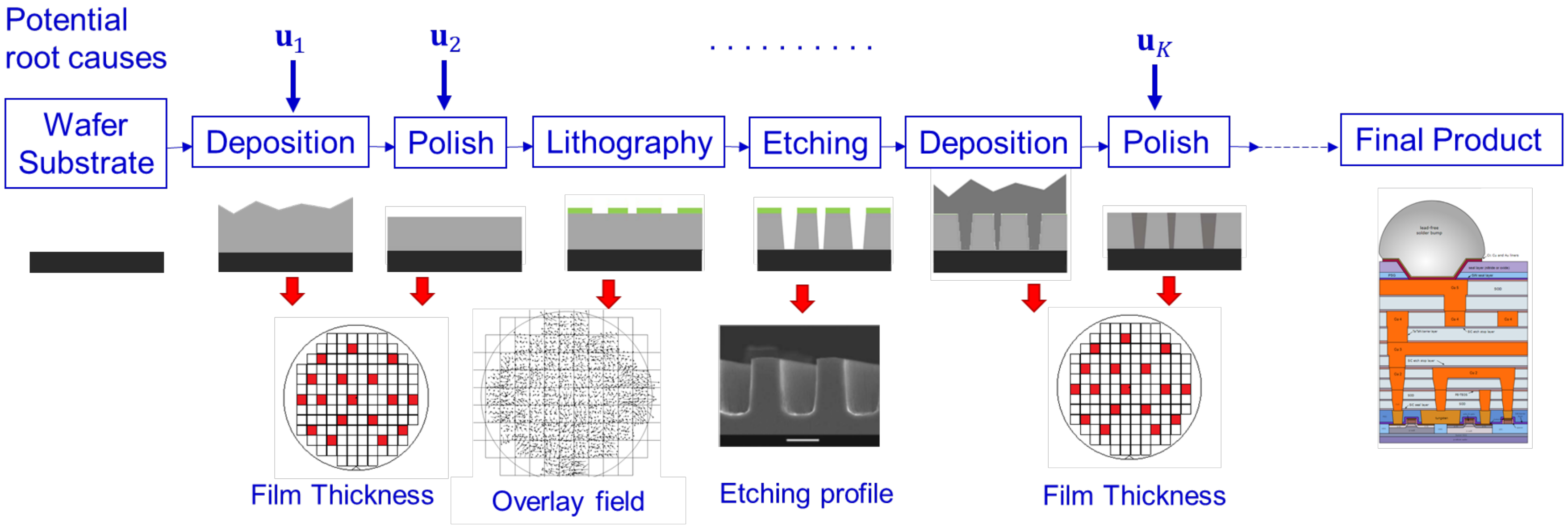
- Not all potential causes affect the quality simultaneously

Smoothness of quality measurements

- Profiles of smooth curves and images

Few underlying quality issues

- Actual root causes link to few latent quality issues and thus cause limited variation patterns of profiles



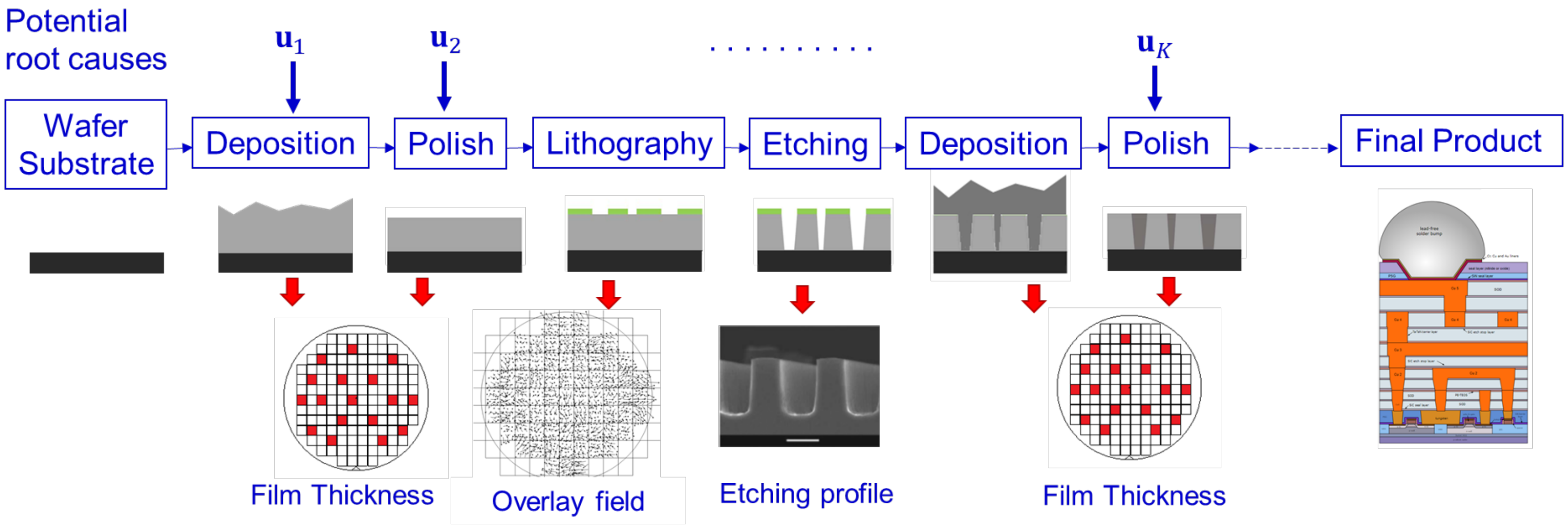
# Common Existing Approaches

**Stream-of-variation** [1,2]
"State vectors" are not well defined.

**Quality Flow Model** [3,4]
Not suitable for data-rich MMPs.

**Separate models** [e.g., 5,6]
Risk missing key features.

**Two-step approach** [e.g., 7]
May not reach consensus on the actual root causes.





[1] Shi, J. (2006). *Stream of variation modeling and analysis for multistage manufacturing processes*. CRC press.
[2] Shi, J., & Zhou, S. (2009). Quality control and improvement for multistage systems: A survey. *IISE Transactions*, *41*(9), 744-753.
[3] Ju, F., Li, J., Xiao, G., Huang, N., & Biller, S. (2013). A quality flow model in battery manufacturing systems for electric vehicles. *IEEE Transactions on Automation Science and Engineering*, *11*(1), 230-244.
[4] Ju, F., Li, J., Xiao, G., & Arinez, J. (2013). Quality flow model in automotive paint shops. *International Journal of Production Research*, *51*(21), 6470-6483.
[5] Lin, T. H., Hung, M. H., Lin, R. C., & Cheng, F. T. (2006, May). A virtual metrology scheme for predicting CVD thickness in semiconductor manufacturing. In *Proceedings 2006 IEEE International Conference on Robotics and Automation, 2006. ICRA 2006.* (pp. 1054-1059). IEEE.
[6] Ma, X., Zhao, Q., Zhang, H., Wang, Z., & Arce, G. R. (2018). Model-driven convolution neural network for inverse lithography. *Optics express*, *26*(25), 32565-32584.
[7] Moyne, J., & Iskandar, J. (2017). Big data analytics for smart manufacturing: Case studies in semiconductor manufacturing. *Processes*, *5*(3), 39.

# Research Objective

Objective:

To develop a holistic framework for the MMP in data rich environment

Which are the actual root causes that affect the process quality?

How these root causes affect the quality measurements?

How to associate the quality measures with several key variation patterns?

Unique contributions:

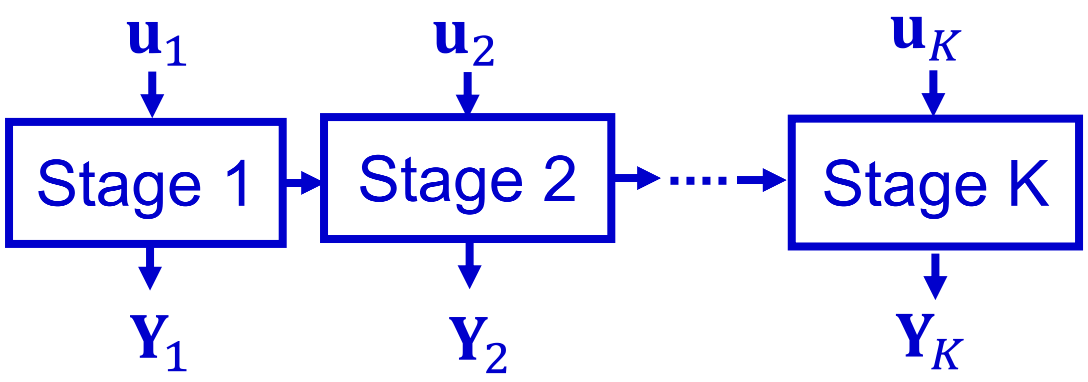
- Analyze heterogeneous quality measures and potential root causes simultaneously under a unified framework.
- Develop a solution procedure based on distributed optimization

# Proposed Process Model

Multiple potential root causes ($\mathbf{u}_i$)

$$\mathbf{u}_i = \left(u_{i1}, \dots, u_{ij}, \dots, u_{iJ_i}\right)^{\mathsf{T}}$$

Quality measurements of multiple functional signals or images ($\mathbf{Y}_k$)




Linearity assumption and cascading effect

$$\mathbf{Y}_k = \mathbf{B}_{k0} + \sum_{i=1}^{k} \sum_{j=1}^{J_i} u_{ij} \mathbf{B}_{ij,k} + \mathbf{E}_k, k = 1, \dots, K.$$

$\mathcal{B}_0 = \{\mathbf{B}_{k0} : k = 1, \dots, K\}$ Offset matrices

$\mathcal{B} = \{\mathbf{B}_{ij,k} : i = 1, \dots, k; k = 1, \dots, K; j = 1 \dots J_i\}$ Effect matrices between $\mathbf{Y}_k$ and $u_{ij}$'s

# Problem Formulation and Expected Outcomes

Linearity assumption and cascading effect

$$\mathbf{Y}_k = \mathbf{B}_{k0} + \sum_{i=1}^{k}\sum_{j=1}^{q_i} u_{ij}\mathbf{B}_{ij,k} + \mathbf{E}_k, k = 1, \dots, K.$$

Offset mat's $\mathcal{B}_0$ $=\{\mathbf{B}_{k0}\}$
Effect mat's $\mathcal{B}$ $=\{\mathbf{B}_{ij,k}\}$

By solving $\mathcal{B}, \mathcal{B}_0$, we can answer the three problems simultaneously

- Actual root causes:

$$\{(i,j): \mathbf{B}_{ij,k} = \mathbf{0}, k = i, \dots, K\}$$

- Subspace of all variation patterns for stage $k$:

$$\text{span}\ \{\mathbf{B}_{ij,k}: i = 1, \dots, k, j = 1, \dots, n_i\}$$

- How root causes $u_{ij}$ affect the quality measurements $\mathbf{Y}_k$:

Individual effect matrix $\mathbf{B}_{ij,k}$

# Problem Formulation: Magnitude of Prediction Error

$$\min_{\mathcal{B},\mathcal{B}_0} \mathcal{L}(\mathcal{B},\mathcal{B}_0) + \lambda_1 p_1(\mathcal{B},\mathcal{B}_0) + \lambda_2 p_2(\mathcal{B},\mathcal{B}_0) + \lambda_3 p_3(\mathcal{B}) + \lambda_4 p_4(\mathcal{B})$$

Offset mat's $\mathcal{B}_0 = \{\mathbf{B}_{k0}\}$
Effect mat's $\mathcal{B} = \{\mathbf{B}_{ij,k}\}$

$\Updownarrow$

Magnitude of the prediction error:

For good prediction of the quality measurements using potential causes

$$\mathcal{L}(\mathcal{B},\mathcal{B}_0) = \sum_{n=1}^{N}\sum_{k=1}^{K}\left\|\mathbf{Y}_k^{\{n\}} - \mathbf{B}_{k0} - \sum_{i=1}^{k}\sum_{j=1}^{q_i} u_{ij}^{\{n\}}\mathbf{B}_{ij,k}\right\|_F^2$$

# Problem Formulation: Smoothness

$$\min_{\mathcal{B},\mathcal{B}_0} \mathcal{L}(\mathcal{B},\mathcal{B}_0) + \lambda_1 p_1(\mathcal{B},\mathcal{B}_0) + \lambda_2 p_2(\mathcal{B},\mathcal{B}_0) + \lambda_3 p_3(\mathcal{B}) + \lambda_4 p_4(\mathcal{B})$$

Offset mat's $\mathcal{B}_0 = \{\mathbf{B}_{k0}\}$
Effect mat's $\mathcal{B} = \{\mathbf{B}_{ij,k}\}$

Quality measure as functional curves

Quality measure as images

$$p_1(\mathcal{B},\mathcal{B}_0) = \sum_{k\in\mathcal{S}} \left[ \sum_{m=0}^{m_k} \|\mathbf{D}_S \mathbf{B}_{k0}(m,:)\|_2^2 + \sum_{i=1}^{k} \sum_{j=1}^{q_i} \sum_{m=1}^{m_k} \left\|\mathbf{D}_S \mathbf{B}_{ij,k}(m,:)\right\|_2^2 \right]$$

$$p_2(\mathcal{B},\mathcal{B}_0) = \sum_{k\in\mathcal{I}} \left[ \|\mathbf{D}_I \operatorname{vec}(\mathbf{B}_{k0})\|_2^2 + \sum_{i=1}^{k} \sum_{j=1}^{q_i} \left\|\mathbf{D}_I \operatorname{vec}\left(\mathbf{B}_{ij,k}\right)\right\|_2^2 \right]$$

$\mathbf{D}_S, \mathbf{D}_I$: discretized versions of $\partial^2/\partial x^2$ and $(\partial^2/\partial x^2 + 2\partial^2/\partial x\partial y + \partial^2/\partial y^2)$ [1].

[1] Buckley, M. J. (1994). Fast computation of a discretized thin-plate smoothing spline for image data. *Biometrika*, *81*(2), 247-258.

# Problem Formulation: Sparsity of Potential Root Causes

$$\min_{\mathcal{B},\mathcal{B}_0} \mathcal{L}(\mathcal{B},\mathcal{B}_0) + \lambda_1 p_1(\mathcal{B},\mathcal{B}_0) + \lambda_2 p_2(\mathcal{B},\mathcal{B}_0) + \lambda_3 p_3(\mathcal{B}) + \lambda_4 p_4(\mathcal{B})$$

Offset mat's $\mathcal{B}_0 = \{\mathbf{B}_{k0}\}$
Effect mat's $\mathcal{B} = \{\mathbf{B}_{ij,k}\}$

Select the potential root causes

$$p_3(\mathcal{B}) = \sum_{i=1}^{K} \sum_{j=1}^{q_i} \left\| \mathbf{B}_{ij,\cdot} \right\|_2$$

$$\mathbf{B}_{ij,\cdot} = \begin{bmatrix} \mathrm{vec}(\mathbf{B}_{ij,i}) \\ \vdots \\ \mathrm{vec}(\mathbf{B}_{ij,K}) \end{bmatrix}$$

A long vector containing all elements in $\mathcal{B}$ associated with $u_{ij}$.

# Problem Formulation: Few Latent Quality Issues

$$\min_{\mathcal{B},\mathcal{B}_0} \mathcal{L}(\mathcal{B},\mathcal{B}_0) + \lambda_1 p_1(\mathcal{B},\mathcal{B}_0) + \lambda_2 p_2(\mathcal{B},\mathcal{B}_0) + \lambda_3 p_3(\mathcal{B}) + \lambda_4 p_4(\mathcal{B})$$

Offset mat's $\mathcal{B}_0 = \{\mathbf{B}_{k0}\}$
Effect mat's $\mathcal{B} = \{\mathbf{B}_{ij,k}\}$

Restrict the number of variation patterns of each stage's quality measurement

$$p_4(\mathcal{B}) = \sum_{k=1}^{K} \left\| \mathbf{B}_{\cdot\cdot,k} \right\|_*$$

$$\mathbf{B}_{\cdot\cdot,k} = \left[\operatorname{vec}\left(\mathbf{B}_{11,k}\right) \ \dots \ \operatorname{vec}\left(\mathbf{B}_{kq_k,k}\right)\right]$$

Each column is a variation pattern of stage $k$ output caused by a $u_{ij}$.

# Problem Formulation: Summary

$$\min_{\mathcal{B},\mathcal{B}_0} \mathcal{L}(\mathcal{B},\mathcal{B}_0) + \lambda_1 p_1(\mathcal{B},\mathcal{B}_0) + \lambda_2 p_2(\mathcal{B},\mathcal{B}_0) + \lambda_3 p_3(\mathcal{B}) + \lambda_4 p_4(\mathcal{B})$$

Offset mat's $\mathcal{B}_0 = \{\mathbf{B}_{k0}\}$
Effect mat's $\mathcal{B} = \{\mathbf{B}_{ij,k}\}$

$$\sum_{n=1}^{N}\sum_{k=1}^{K}\left\|\mathbf{Y}_k^{\{n\}} - \mathbf{B}_{k0} - \sum_{i=1}^{k}\sum_{j=1}^{q_i} u_{ij}^{\{n\}}\mathbf{B}_{ij,k}\right\|_F^2$$

$$+\sum_{k\in\mathcal{S}}\left[\sum_{m=0}^{m_k}\|\mathbf{D}_S\mathbf{B}_{k0}(m,:)\|_2^2 + \sum_{i=1}^{k}\sum_{j=1}^{q_i}\sum_{m=1}^{m_k}\|\mathbf{D}_S\mathbf{B}_{ij,k}(m,:)\|_2^2\right]$$

$$+\sum_{k\in\mathcal{I}}\left[\|\mathbf{D}_I\,\mathrm{vec}(\mathbf{B}_{k0})\|_2^2 + \sum_{i=1}^{k}\sum_{j=1}^{q_i}\left\|\mathbf{D}_I\,\mathrm{vec}\left(\mathbf{B}_{ij,k}\right)\right\|_2^2\right]$$

$$+\sum_{i=1}^{K}\sum_{j=1}^{q_i}\left\|\mathbf{B}_{ij,\cdot}\right\|_2 + \sum_{k=1}^{K}\left\|\mathbf{B}_{\cdot\cdot,k}\right\|_*$$

Convex formulation, but lots of parameters!

# Solution Based on ADMM Consensus Algorithm (1)

Analyze the problem:

- Summation of five terms.
- Each term has disjoint groups of coefficients.
- Each term consists of simple convex functions.

$$\sum_{n=1}^{N}\sum_{k=1}^{K}\left\|\mathbf{Y}_k^{\{n\}} - \mathbf{B}_{k0} - \sum_{i=1}^{k}\sum_{j=1}^{q_i} u_{ij}^{\{n\}}\mathbf{B}_{ij,k}\right\|_F^2$$
$$+\sum_{k\in\mathcal{S}}\left[\sum_{m=0}^{m_k}\|\mathbf{D}_S\mathbf{B}_{k0}(m,:)\|_2^2 + \sum_{i=1}^{k}\sum_{j=1}^{q_i}\sum_{m=1}^{m_k}\|\mathbf{D}_S\mathbf{B}_{ij,k}(m,:)\|_2^2\right]$$
$$+\sum_{k\in\mathcal{I}}\left[\|\mathbf{D}_I\,\mathrm{vec}(\mathbf{B}_{k0})\|_2^2 + \sum_{i=1}^{k}\sum_{j=1}^{q_i}\|\mathbf{D}_I\,\mathrm{vec}(\mathbf{B}_{ij,k})\|_2^2\right]$$
$$+\sum_{i=1}^{K}\sum_{j=1}^{q_i}\|\mathbf{B}_{ij,\cdot}\|_2 + \sum_{k=1}^{K}\|\mathbf{B}_{\cdot\cdot,k}\|_*$$

Apply a parallel ADMM consensus algorithm [1]

Key step: update groups of $\mathcal{B}, \mathcal{B}_0$ in parallel. Need to calculate

$$\mathrm{prox}_{\eta f(\cdot)}[\mathbf{x}] = \underset{\mathbf{v}}{\mathrm{argmin}}\left\{f(\mathbf{v}) + \frac{1}{2\eta}\|\mathbf{x}-\mathbf{v}\|^2\right\}$$

$f$ – quadratic loss, $\|\cdot\|_2$, $\|\cdot\|_*$, $p_S(\cdot) = \|\mathbf{D}_S\cdot\|_2^2$ and $p_I(\cdot) = \|\mathbf{D}_I\cdot\|_2^2$

[1] Parikh, N., & Boyd, S. (2014). Proximal algorithms. *Foundations and Trends in optimization*, *1*(3), 127-239.

# Solution Based on ADMM Consensus Algorithm (2)

We give the efficient algorithms for the prox operator of curve and image smoother.

Calculate $\mathrm{prox}_{\eta p_S(\cdot)}[\mathbf{x}]$:

1. $\mathbf{x}^* \leftarrow \mathrm{DCT}(\mathbf{x})$
2. $\tilde{x}_i^* \leftarrow x_i^* / \left[1 + 4\eta\left(1 - \cos\left(\frac{i-1}{d}\pi\right)\right)^2\right]$
3. $\tilde{\mathbf{x}} \leftarrow \mathrm{IDCT}(\mathbf{x}^*)$

Calculate $\mathrm{prox}_{\eta p_I(\cdot)}[\mathbf{T}]$:

1. $\mathbf{T}^* \leftarrow \mathrm{DCT2}(\mathbf{T})$
2. $\tilde{t}_{ij}^* \leftarrow t_{ij}^* / \left[1 + 4\eta\left(2 - \cos\left(\frac{i-1}{m}\pi\right) - \cos\left(\frac{j-1}{n}\pi\right)\right)^2\right]$
3. $\widetilde{\mathbf{T}} \leftarrow \mathrm{IDCT2}(\widetilde{\mathbf{T}}^*)$

Computational Complexity

$O(m^2) \rightarrow O(m \log m)$

Computational Complexity

$O(m^2 n^2) \rightarrow O\big(mn(\log m + \log n)\big)$

# Overview of the ADMM Consensus Algorithm

Initiate $\overline{\mathcal{B}} = \mathcal{B}^{(1)} = \mathcal{B}^{(2)} = \mathcal{B}^{(3)} = \mathcal{B}^{(4)} = \mathcal{U}^{(1)} = \mathcal{U}^{(2)} = \mathcal{U}^{(3)} = \mathcal{U}^{(4)} = \mathcal{O}$ of the same shape as $\mathcal{B}$. Initiate $\overline{\mathcal{B}}_0 = \mathcal{B}_0^{(1)} = \mathcal{B}_0^{(2)} = \mathcal{U}_0^{(1)} = \mathcal{U}_0^{(2)} = \mathcal{O}$ of the same shape as $\mathcal{B}_0$.

**Do:**

(1) Save $\overline{\mathcal{B}}_{0,\text{prev}} \leftarrow \overline{\mathcal{B}}_0$ and $\overline{\mathcal{B}}_{\text{prev}} \leftarrow \overline{\mathcal{B}}$.

**Until** $\max_{i=1,\dots,4} \|\overline{\mathcal{B}} - \mathcal{B}^{(i)}\|$, $\max_{i=1,2} \|\overline{\mathcal{B}}_0 - \mathcal{B}_0^{(i)}\|$, $\max_{i=1,\dots,4} \|\overline{\mathcal{B}} - \overline{\mathcal{B}}_{\text{prev}}\|$ and $\max_{i=1,2} \|\overline{\mathcal{B}}_0 - \overline{\mathcal{B}}_{0,\text{prev}}\| \le \epsilon$.

**Initialize**

4 replicates for $\mathcal{B}$

2 replicates for $\mathcal{B}_0$

**Local Computation**

All **`For`** loops (2a) – (2d) can be performed in parallel using **`prox`** operators

**Global Aggregation**

All elements in assignments can be performed in parallel

**Iterate Until Convergence**

# Simulation Study

A four-stage test-bed motivated by semiconductor manufacturing processes

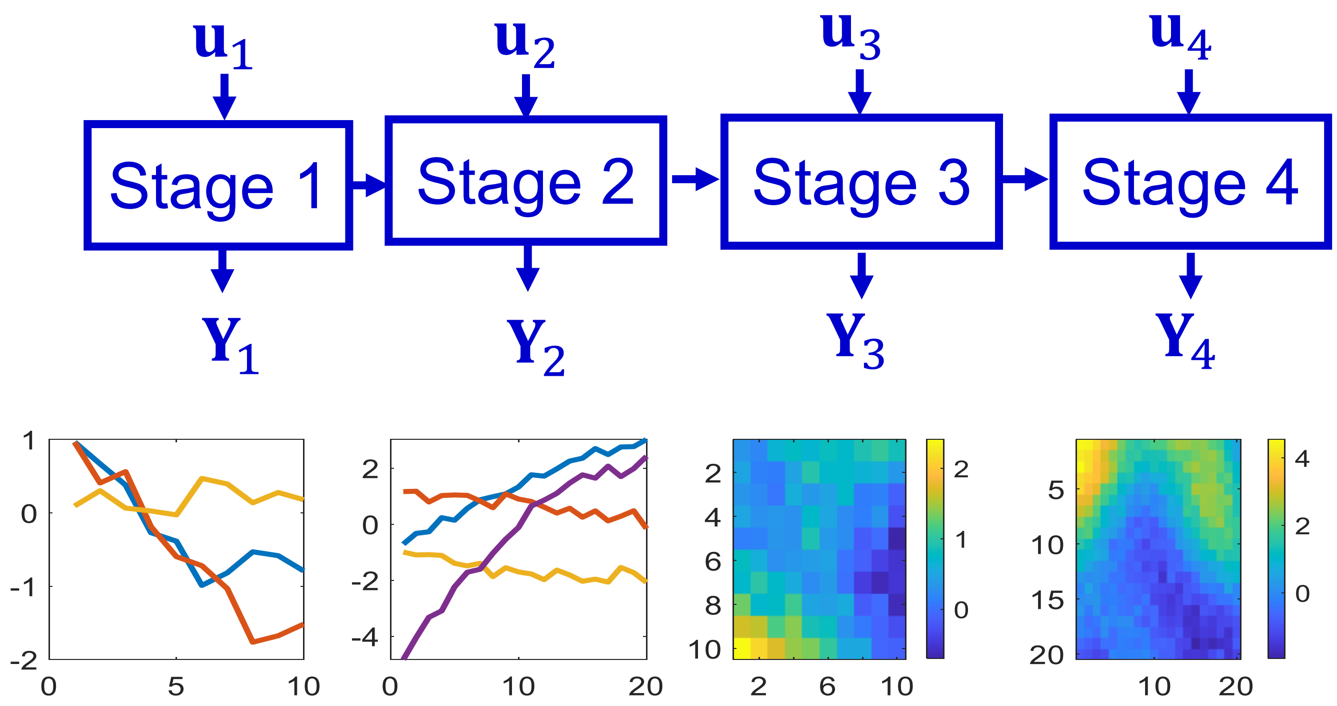
$\mathbf{u}_i = (u_{i1}, \dots, u_{i,20})$

20 potential root causes per stage

Mixed quality measures

Four Settings

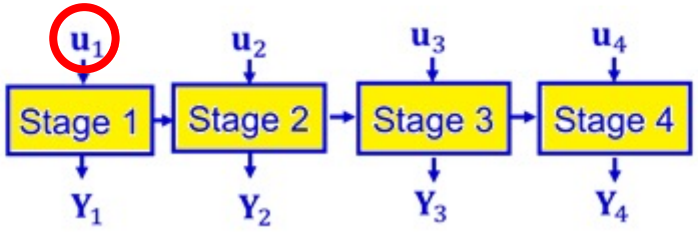
- 3 or 6 actual root causes
- 2 or 5 quality variation patterns



# Illustrating the Effect of Actual Root Causes (1)

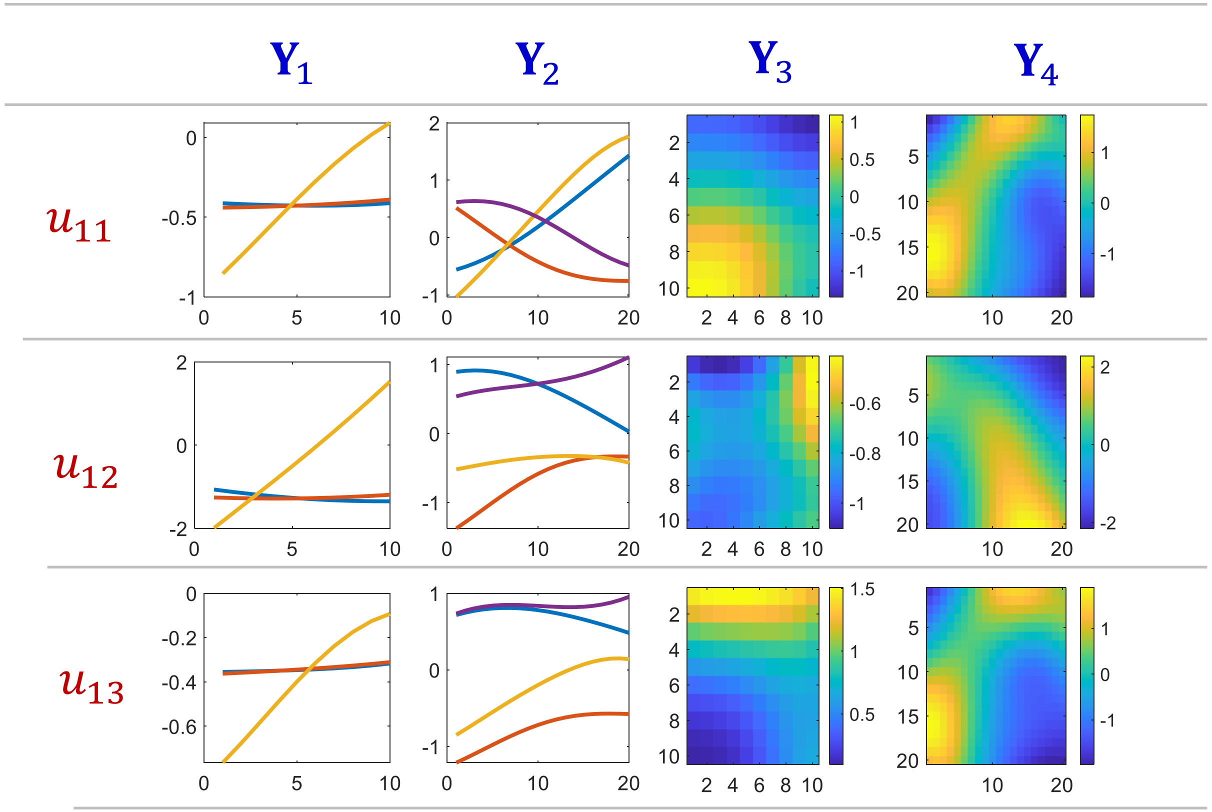
3 actual root causes

2 quality variation patterns

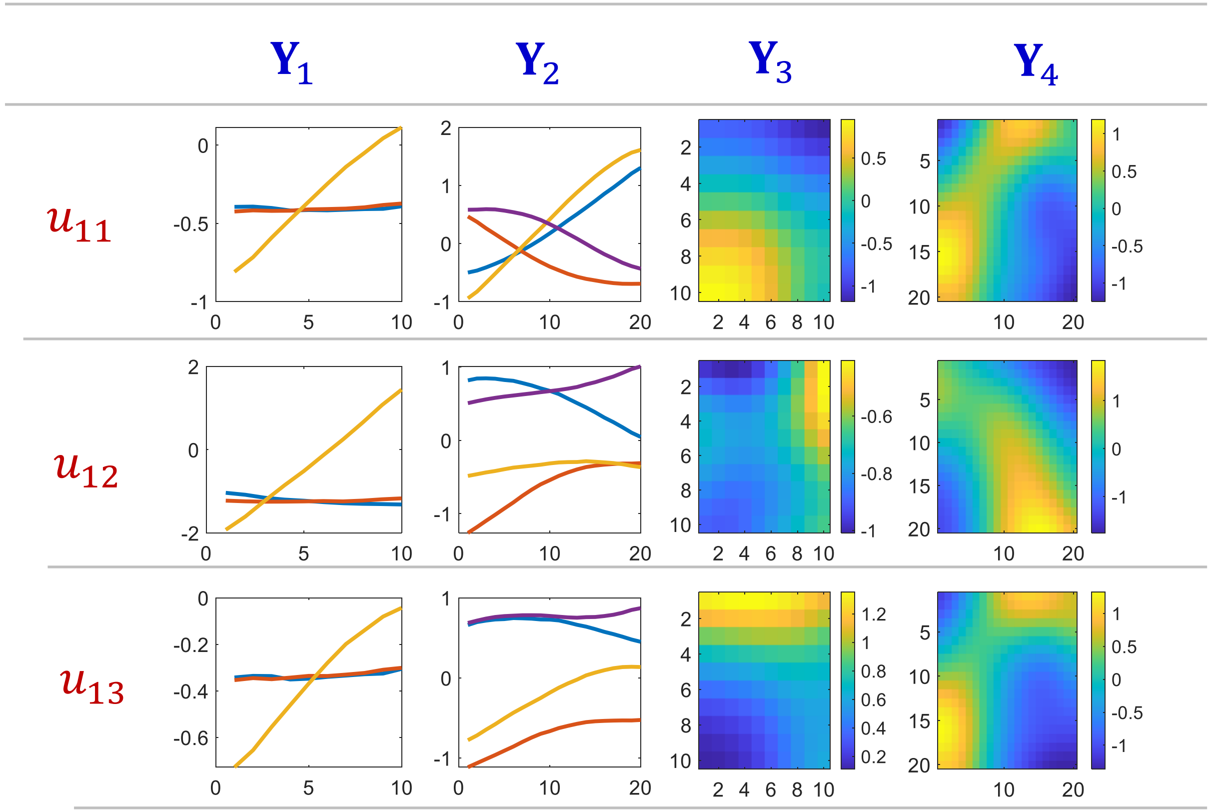



Effect of actual root causes

from stage 1




True parameters of $\mathbf{B}_{ij,k}$

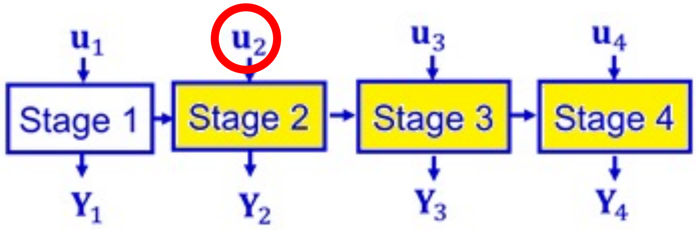



Estimated parameters of $\mathbf{B}_{ij,k}$

# Illustrating the Effect of Actual Root Causes (2)

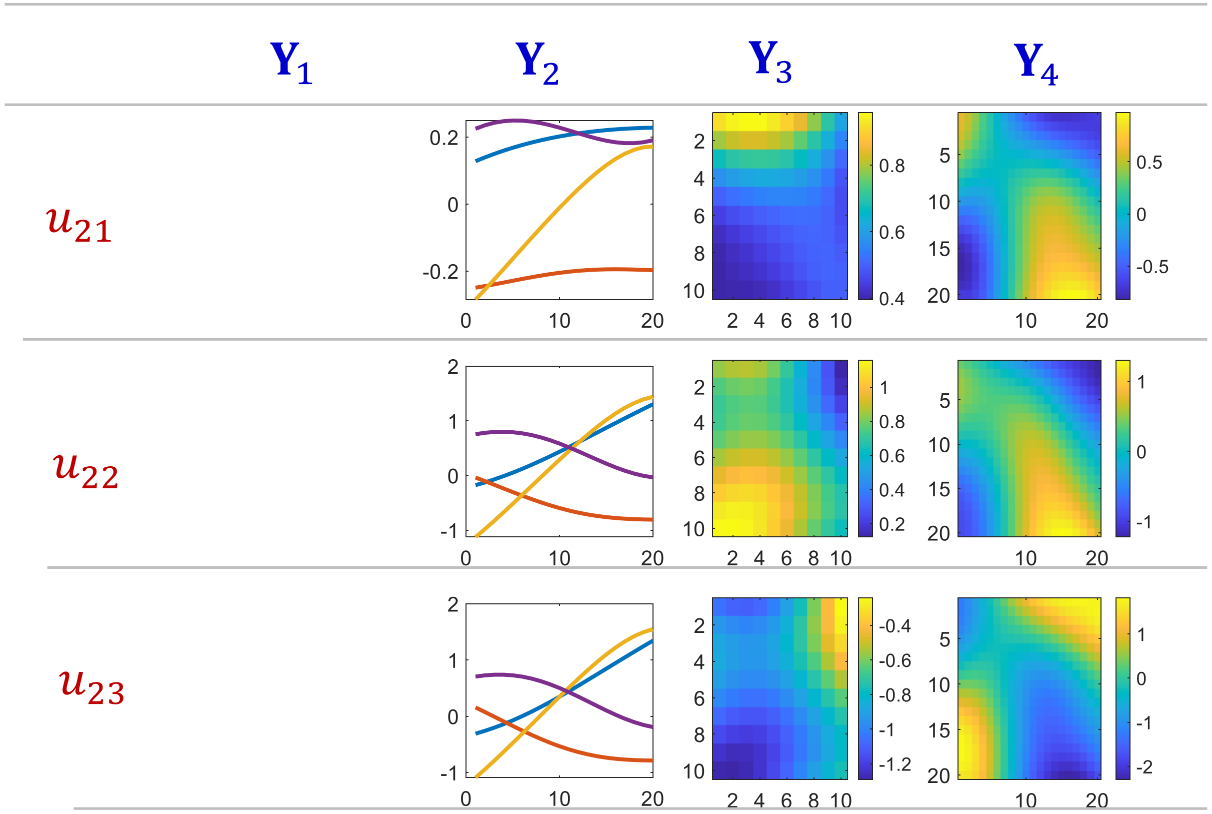
3 actual root causes

2 quality variation patterns

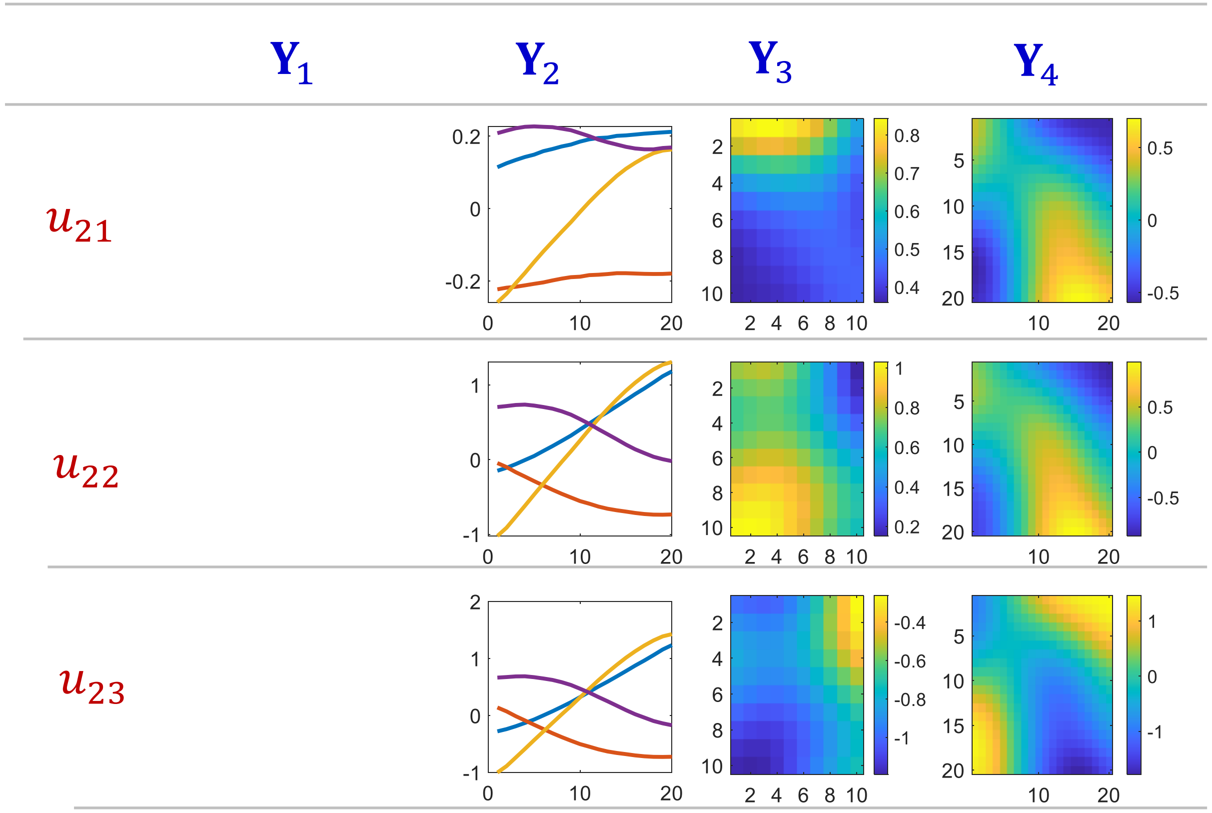



Effect of actual root causes

from stage 2




True parameters of $\mathbf{B}_{ij,k}$

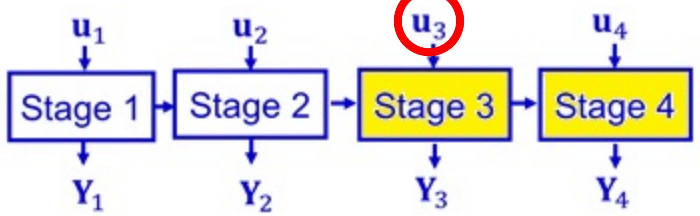



Estimated parameters of $\mathbf{B}_{ij,k}$

# Illustrating the Effect of Actual Root Causes (3)

3 actual root causes

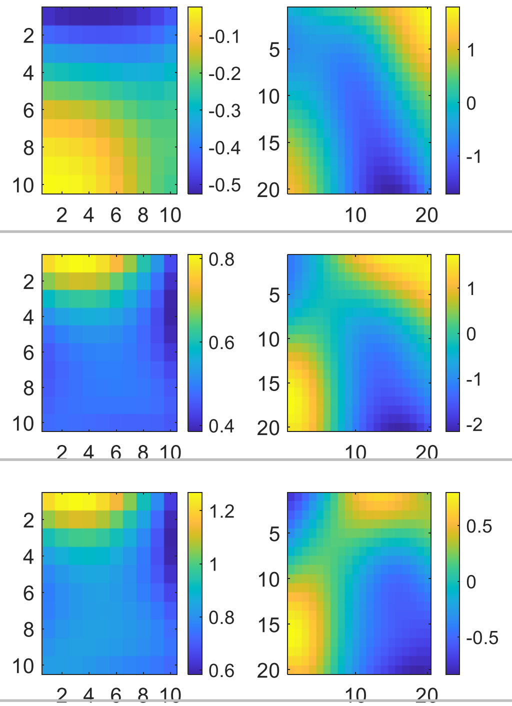
2 quality variation patterns




Effect of actual root causes

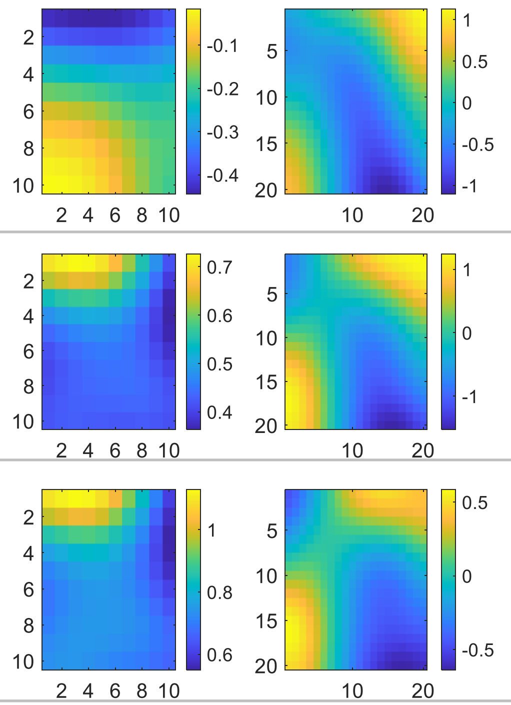
from stage 3

| | $\mathbf{Y}_1$ | $\mathbf{Y}_2$ | $\mathbf{Y}_3$ | $\mathbf{Y}_4$ |
|---|---|---|---|---|
| $u_{31}$ | | | | |
| $u_{32}$ | | | | |
| $u_{33}$ | | | | |



True parameters of $\mathbf{B}_{ij,k}$

| | $\mathbf{Y}_1$ | $\mathbf{Y}_2$ | $\mathbf{Y}_3$ | $\mathbf{Y}_4$ |
|---|---|---|---|---|
| $u_{31}$ | | | | |
| $u_{32}$ | | | | |
| $u_{33}$ | | | | |

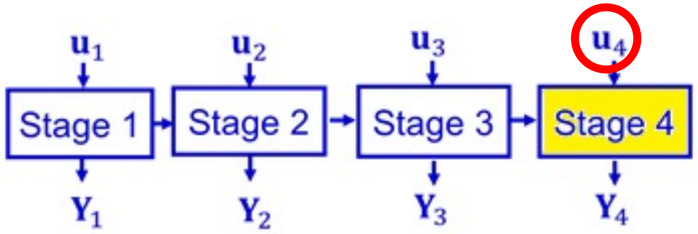


Estimated parameters of $\mathbf{B}_{ij,k}$

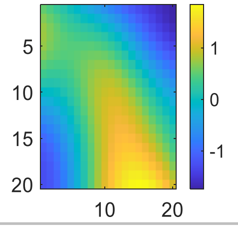
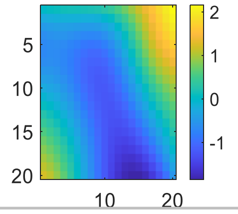
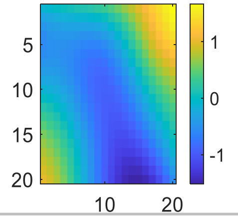
# Illustrating the Effect of Actual Root Causes (4)

3 actual root causes

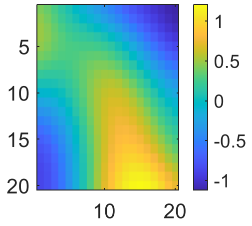
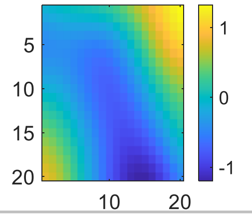
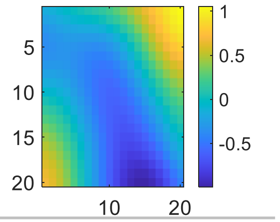
2 quality variation patterns




Effect of actual root causes from stage 4

| | $\mathbf{Y}_1$ | $\mathbf{Y}_2$ | $\mathbf{Y}_3$ | $\mathbf{Y}_4$ |
|---|---|---|---|---|
| $u_{41}$ | | | |  |
| $u_{42}$ | | | |  |
| $u_{43}$ | | | |  |

True parameters of $\mathbf{B}_{ij,k}$

| | $\mathbf{Y}_1$ | $\mathbf{Y}_2$ | $\mathbf{Y}_3$ | $\mathbf{Y}_4$ |
|---|---|---|---|---|
| $u_{41}$ | | | |  |
| $u_{42}$ | | | |  |
| $u_{43}$ | | | |  |

Estimated parameters of $\mathbf{B}_{ij,k}$

# Summary of the Findings

**Correctly identified actual root causes**

- Stage 1-3:

  All actual root causes are identified.

- Stage 4:

  No type I error (missing actual root causes).

  Small type II error (average false root causes < 1.4%).

**Correctly identified variation patterns**

- With 2 variation patterns:

  The number of variation patterns is always correctly identified.

- With 5 variation patterns:

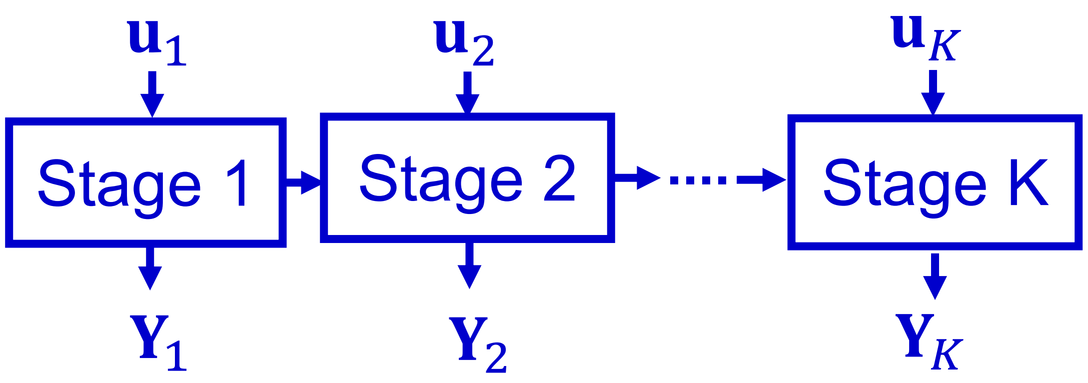
  The numbers can be identified correctly if they are linearly independent.

# Summary

- A holistic modeling and root cause diagnostic framework for data-rich MMPs.

Multiple potential root causes $\mathbf{u}_i$'s

Quality measurements of multi-signals or images $\mathbf{Y}_i$'s




✓ Identify the actual root causes

✓ Understand the root cause's effect on profile measurements

✓ Identify the subspace of underlying quality variation patterns

- First MMP analysis method based on distributed optimization.
- Extendable to more types of data by adjusting loss and penalization terms.

# About Myself


## Andi Wang

Assistant Professor

The Polytechnic School

Arizona State University

https://web.asu.edu/andi-wang

andi.wang@asu.edu

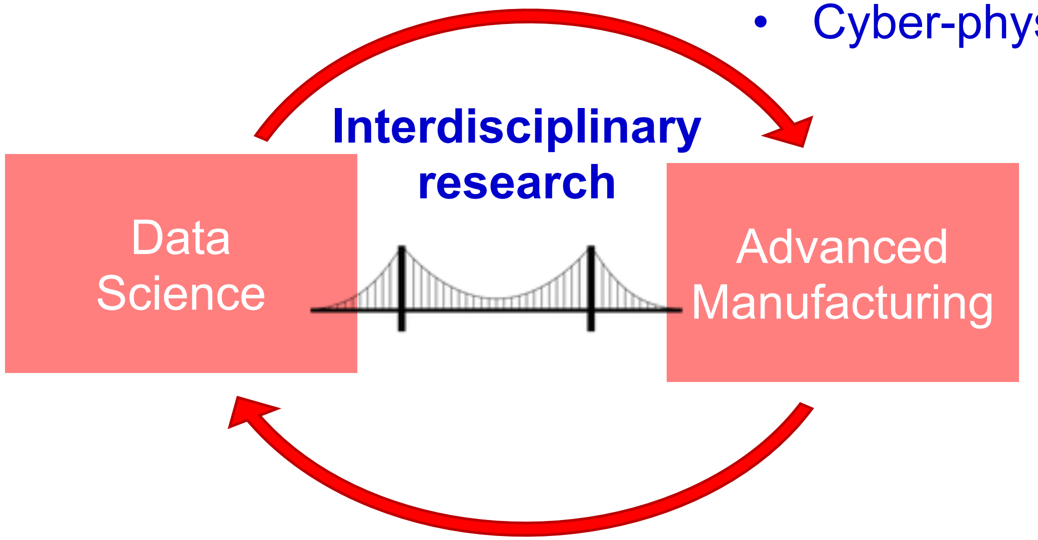

### Algorithms and Solutions

- Accurate
- Efficient
- Scalable

### Applications

**Intelligent Manufacturing Systems**

- Steel Rolling
- Semiconductor Manufacturing
- Additive Manufacturing
- Internet-of-Things
- Cyber-physical Systems




### Tools

- Machine Learning
- High-dimensional Stat
- Large Scale Optimization

### Data

- High-speed
- Massive
- Heterogeneous
- Complex-structured

### Problems

**Variation modeling and analysis**

- Monitoring / Detection
- Diagnostics & Prognostics
- Forecasting and prediction

## Call for new PhD students

- Research Area

  industrial data analysis, machine learning for engineering applications, smart manufacturing, and data-driven modeling for complex systems.

- Requirements
  - B.S. or M.S. in engineering or statistics
  - Motivated for interdisciplinary research

## Collaboration opportunities

Please contact me with data problems in engineering applications!