PURDUE UNIVERSITY

Systems Engineering and Process Safety

Can Li (Tsan)
Assistant Professor (Joined 2022)
Davidson School of Chemical Engineering

Research Group

Fundamental Research	Applications	People and projects
Mathematical Optimization under Uncertainty	Energy systems optimization, chemical process design and operation, supply chain	Asha: Distributed manufacturing of the electrified processes Kaiyu: Integrated scheduling and control Casper: Electrification of Process Heating (co-advised with Dr. Cornelius Masuku)
Machine Learning	Fault detection, speech/image recognition, computational biology, Acceleration of optimization algorithms	Chi: end-to-end learning of optimization problems Hao: Explainable ML for fault detection Yen-Chun: ML for stem cell (co-advised with Dr. Xiaoping Bao)

Mathematical Programming under Uncertainty

Deterministic Optimization Model

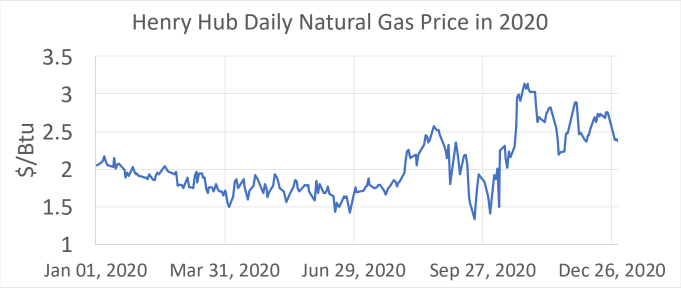
- Model **decision-making** process as an **optimization** problem

$$
\begin{aligned}
\min \quad & f(x, y, \theta) \\
s.t. \quad & g(x, y, \theta) \leq 0 \\
& h(x, y, \theta) = 0 \\
& x \in X, y \in Y
\end{aligned}
$$

- **Variables** continuous x, discrete y the capacity of process (x), whether to install a process or not (y)
- **Constraints** g, h the mass balance, to satisfy the customer demand
- **Objective** f minimize total cost
- **Parameters** θ Product demand, unit cost, thermal and kinetic properties
- The input parameters θ can be uncertain.

Sources of Parameter Uncertainty

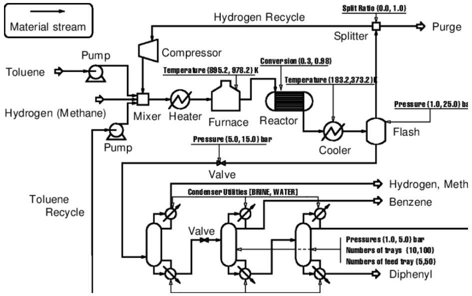
- Long-term forecasts, e.g., natural gas price



- Short-term changing conditions, e.g., extreme weather



- Real-time inaccurate measurement, e.g., temperature, pressure



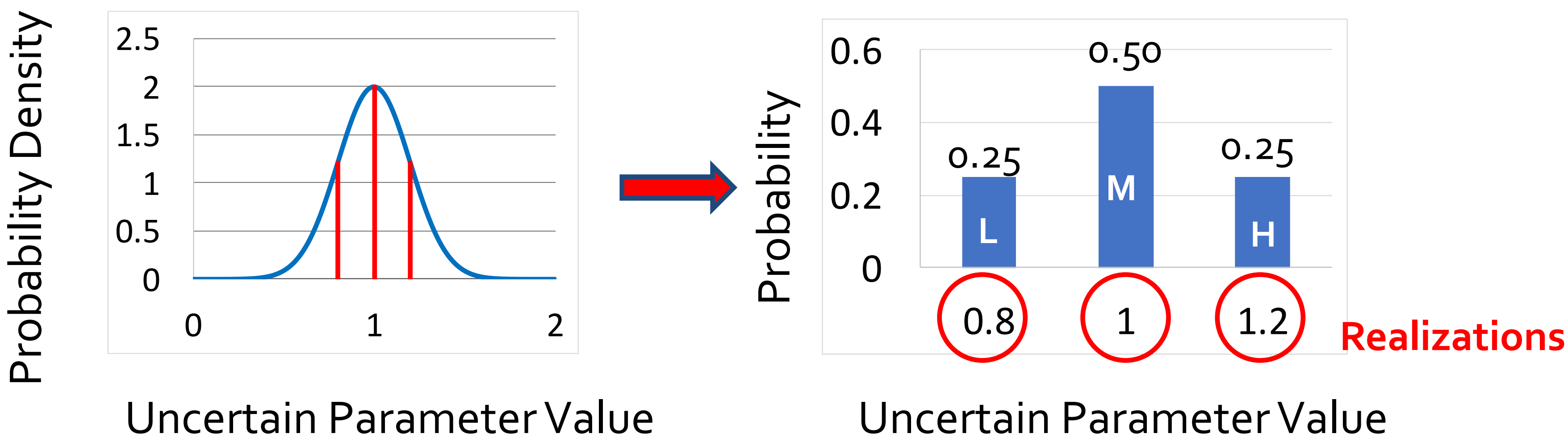
How Do We Model Uncertainty in Optimization Problems?

- Not a uniquely-defined problem
  - **Multiple** ways to hedge against uncertainty

The jungle of stochastic optimization
(credit: Warren Powell)

Stochastic Programming

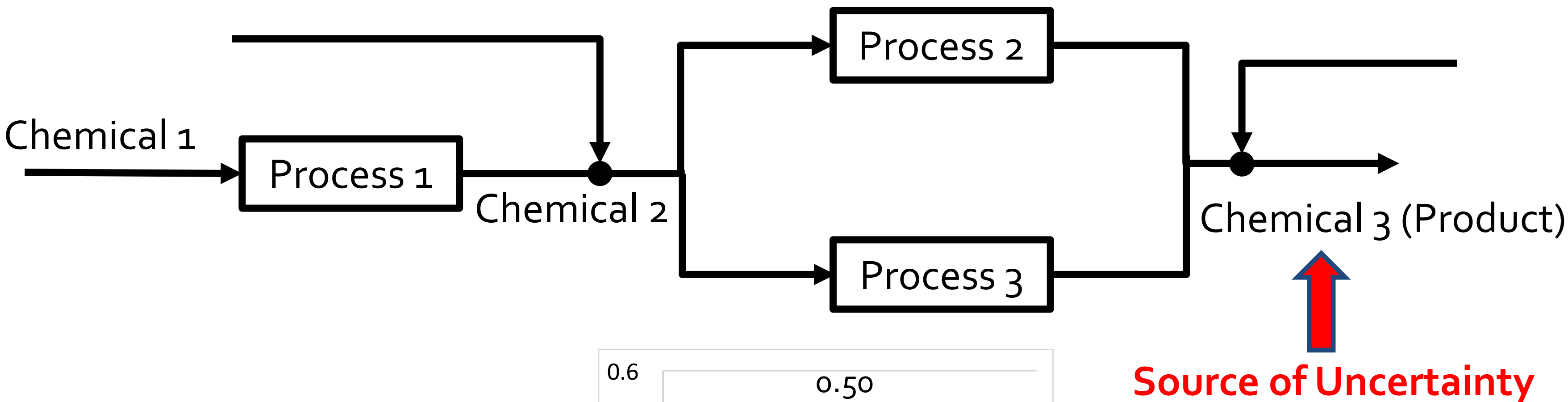
- **Stochastic programming** is a framework for modeling optimization problems that involve uncertainty (Birge and Louveaux, 2011)
- Uncertainty can be characterized by **probability distributions** known *a priori*
  - Continuous distributions → Discrete distributions

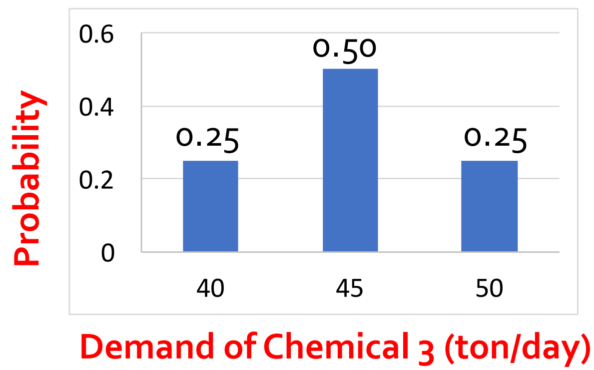


- Each realization of uncertainty parameters is called **a scenario**
- Optimize the **expected value** of the objective over all possible scenarios, i.e., **a risk-neutral approach**
- Common framework: **Two stage stochastic programming**

A Motivating Example

- **Superstructure**



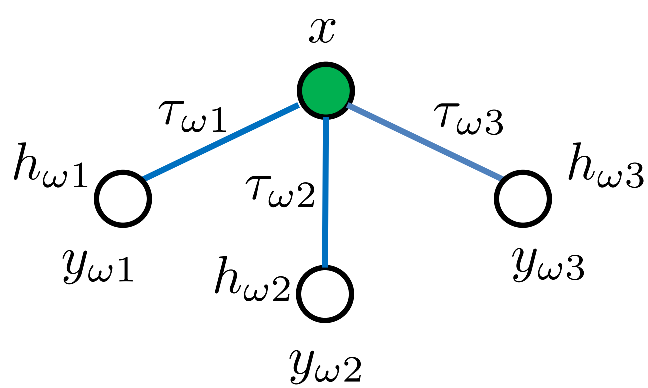


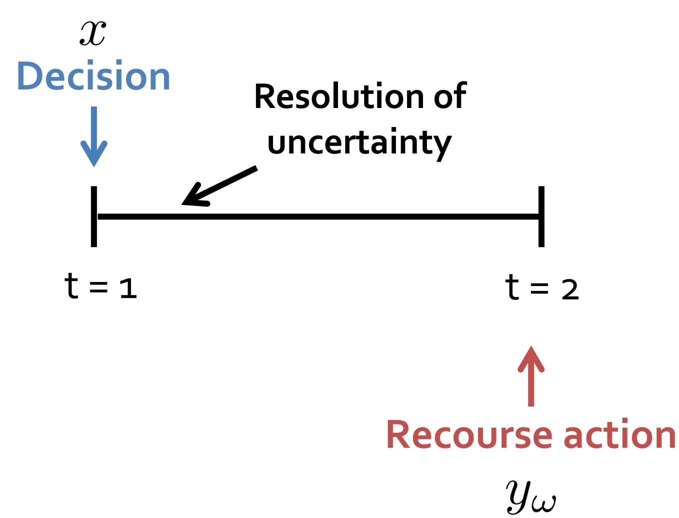
- **First stage decisions:** which process to install, the capacity of each process
- **Second stage decisions:** the mass flow rate of each stream
- **Constraints:** satisfy customer demands, mass balance

Two Stage Stochastic Programming

- First stage decisions: **Here and now**
- Second stage decisions: **Wait and see, Recourse decisions**

$$
\begin{aligned}
\min \quad & c^T x + \sum_{\omega \in \Omega} \tau_\omega d_\omega^T y_\omega \\
s.t. \quad & Ax \leq b \\
& W_\omega y_\omega \leq h_\omega - T_\omega x \qquad \forall \omega \in \Omega \\
& x \in \{x = (x1, x2) : x_1 \in \{0,1\}^n, x_2 \geq 0\} \\
& y_\omega \in \{y : y \geq 0\}
\end{aligned}
$$

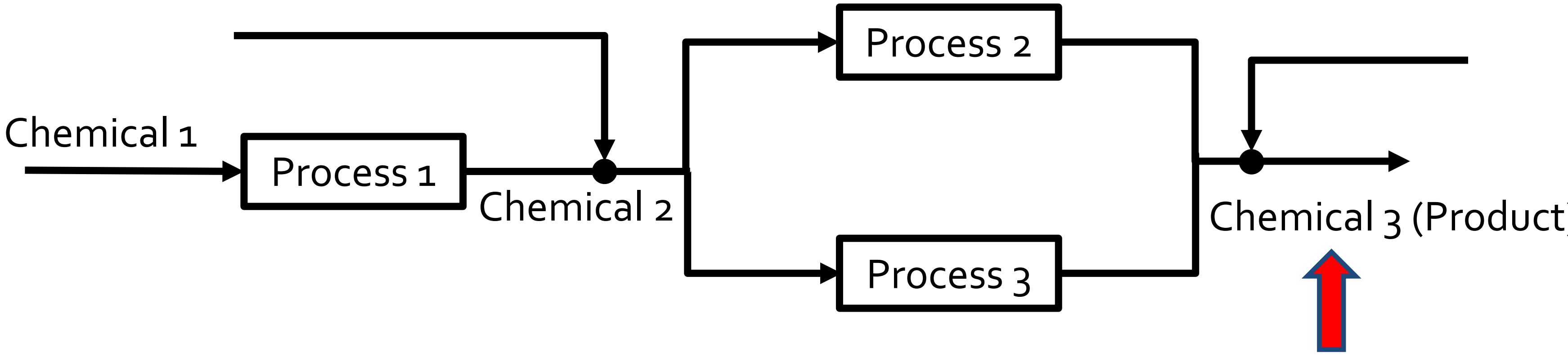




Deterministic Equivalent

Investment Planning of Process Networks

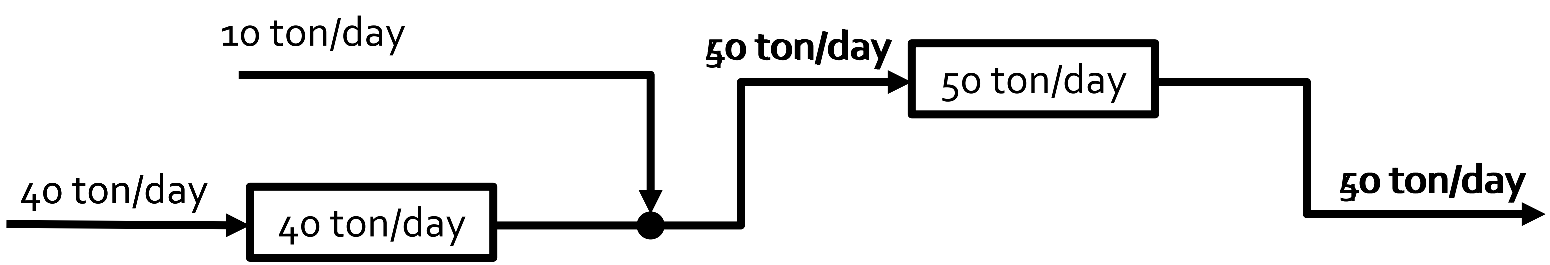
- **Superstructure**



Source of Uncertainty

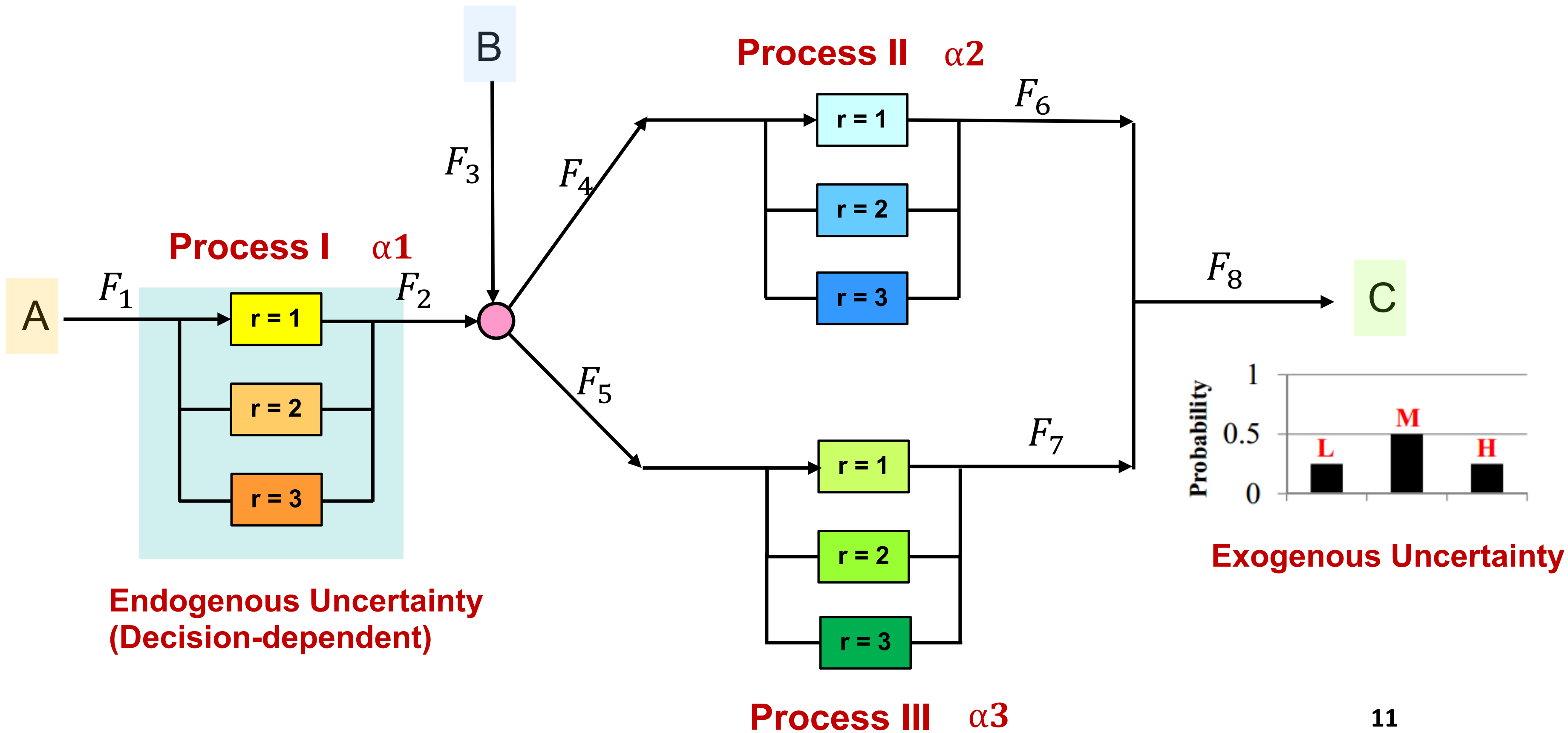
- **First stage decisions**
- **Second stage decisions**

High demand
Low demand



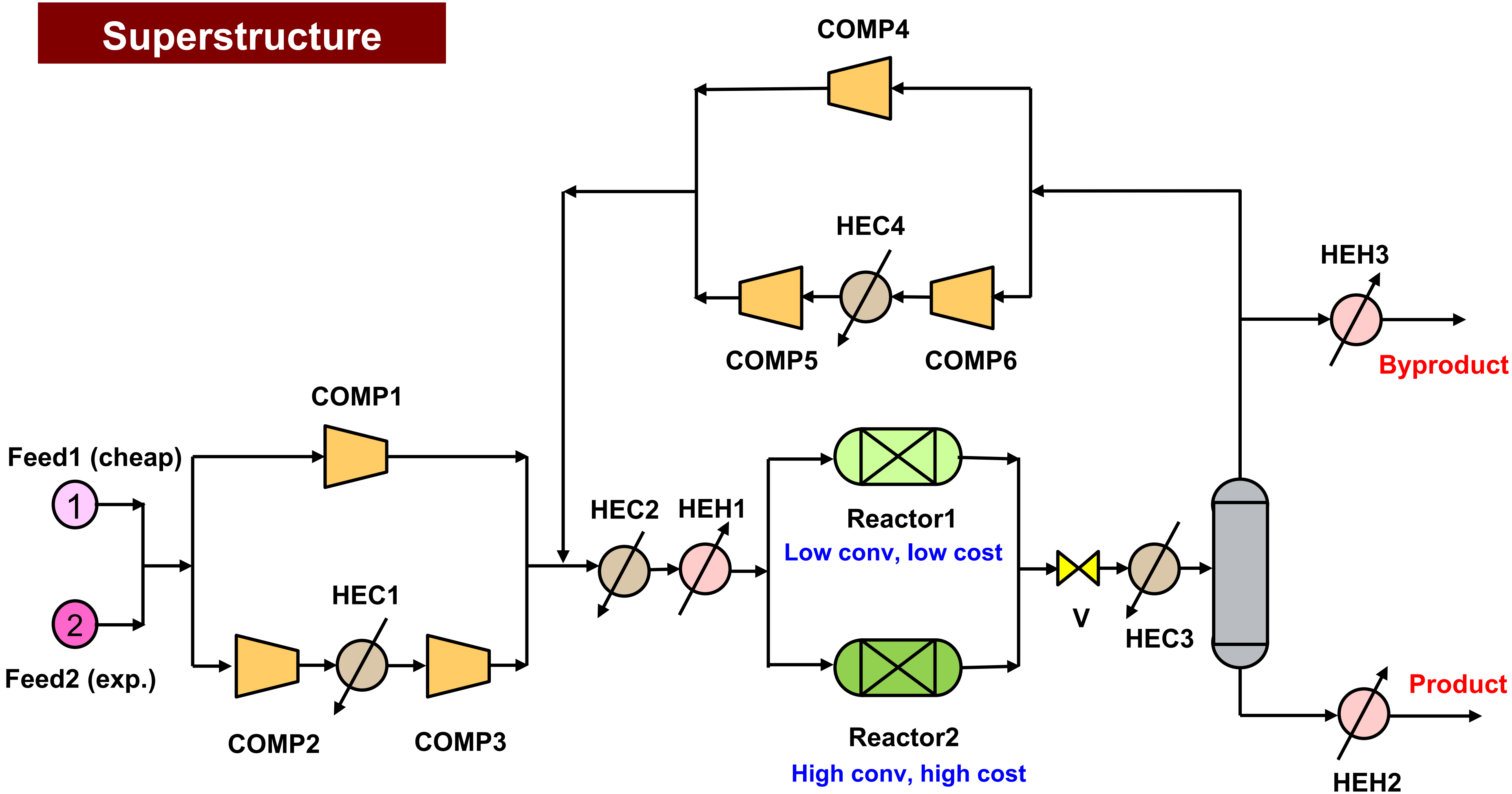
Extension to Process Reliability

- Motivation: Process units may fail
- Solution: Have backup units to improve reliability
- Trade-off: Investment cost v.s. system reliability



Industrial Methanol Synthesis

Superstructure

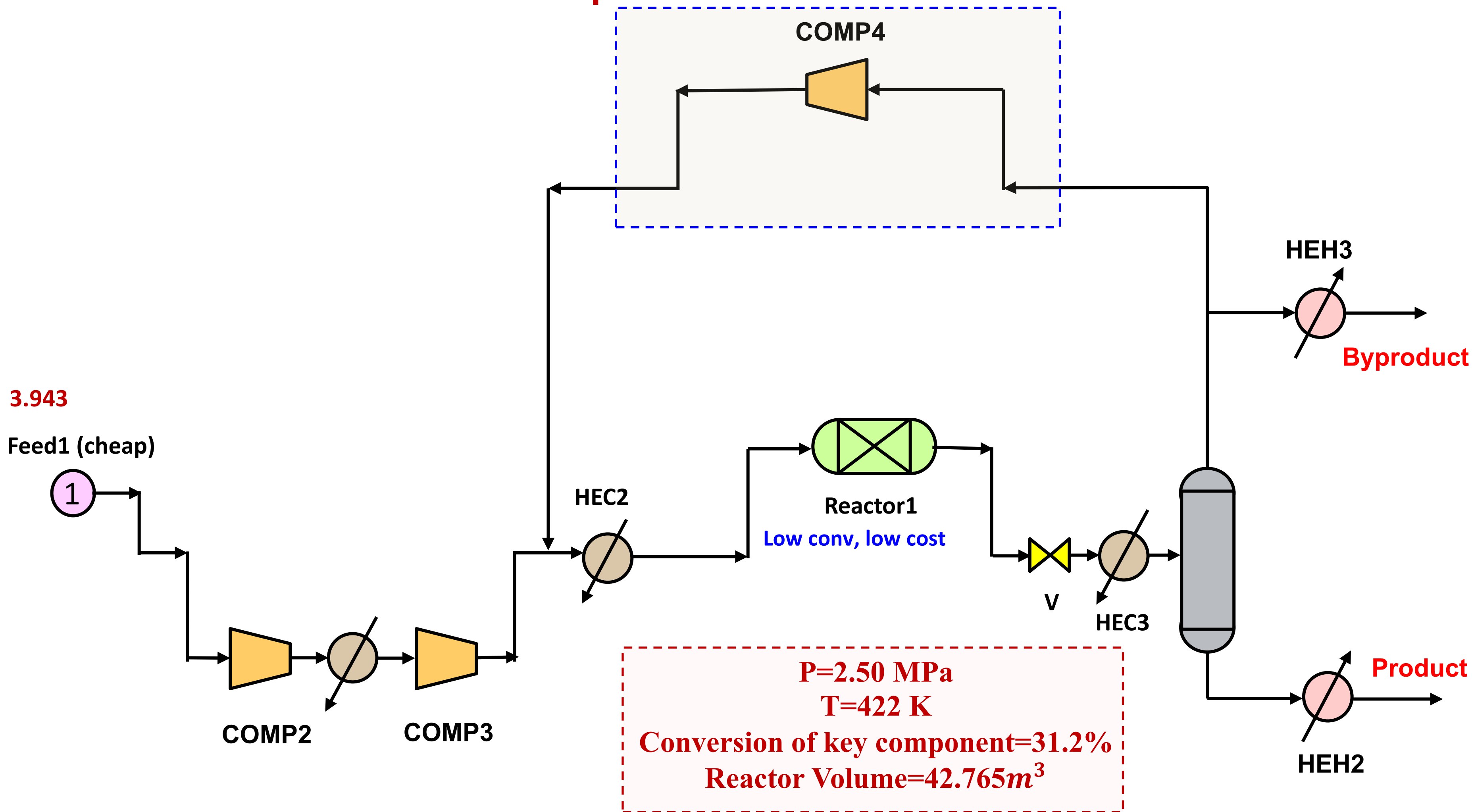


Key optimization variables in the reactors:
operating pressure and the conversion per pass

Industrial Methanol Synthesis

Deterministic Model

Profit = 4115.3749 ($1000 PER YEAR)
Too optimistic

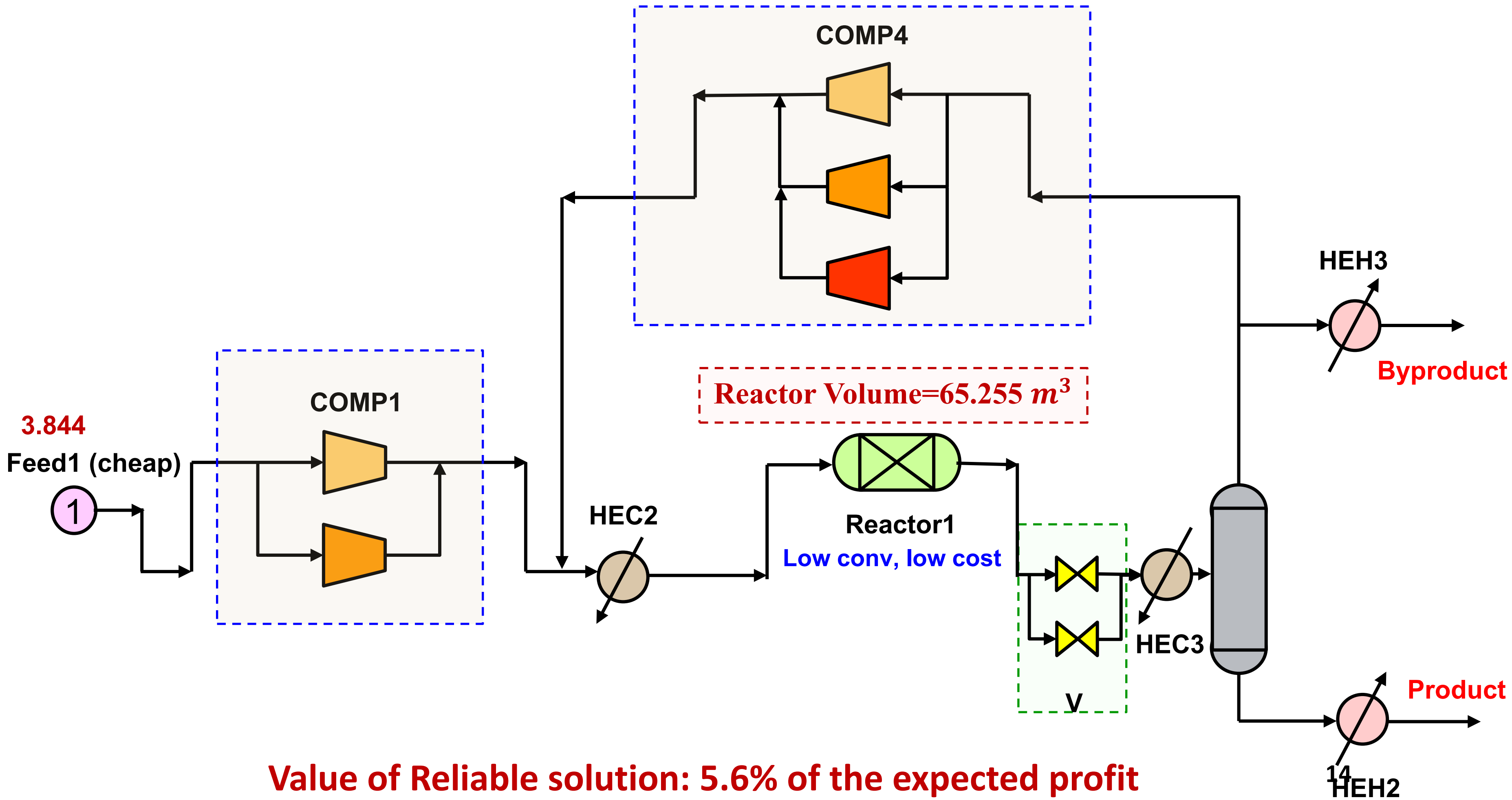


Industrial Methanol Synthesis

Stochastic Programming Model

Expected Profit = 3203.6879 ($1000 PER YEAR)

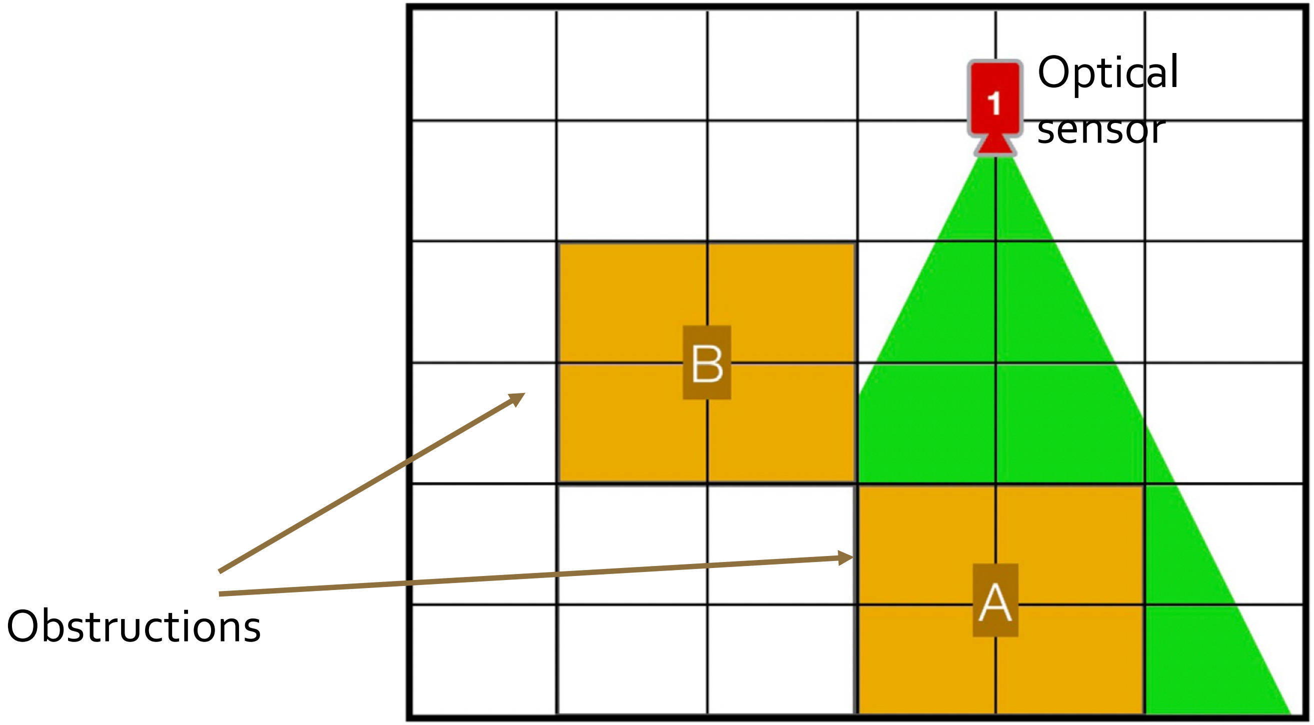
Consider demand uncertainty and reliability Simultaneously



Value of Reliable solution: 5.6% of the expected profit

Sensor Placement under Uncertainty

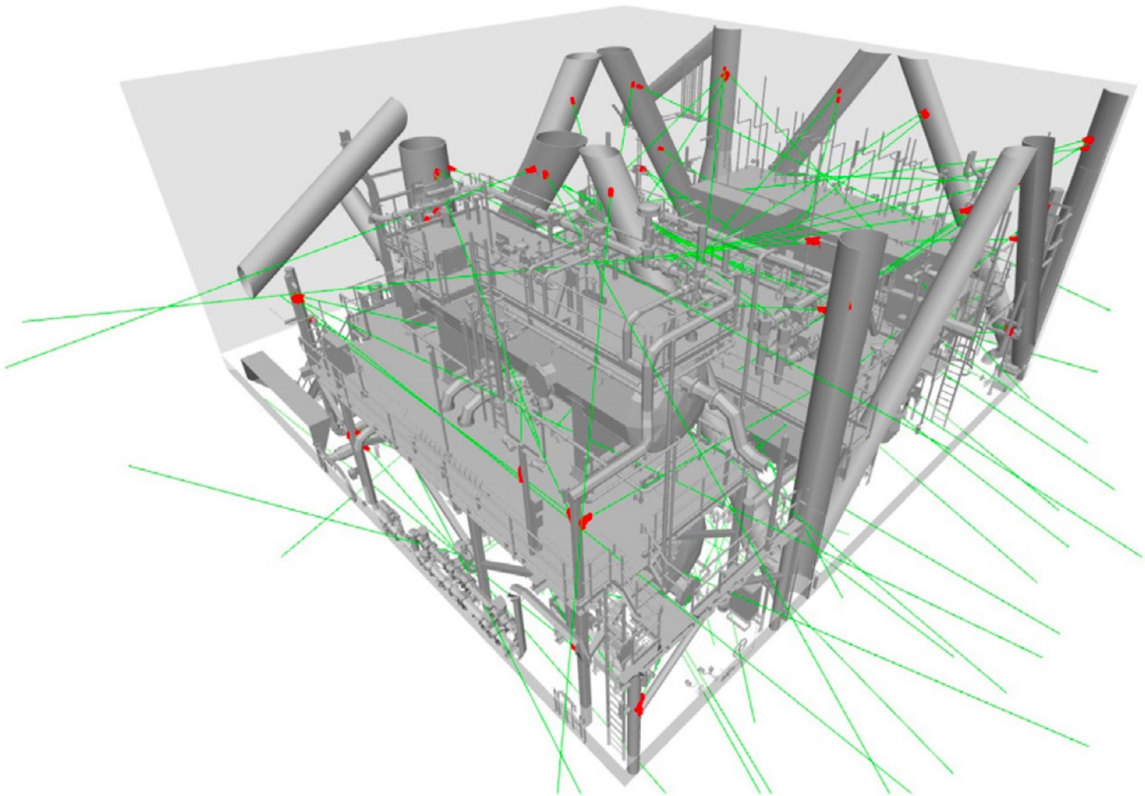
- Motivation: Determine the optimal configurations of sensors to maximize the probability of detecting safety hazards
- Flame, smoke, and heat detectors using chemical or optical sensors



Work of Prof. Carl Laird with P2SAC

Sensor Placement under Uncertainty

- Facility with 81 candidate flame detector locations (Kenexis Consulting Corporation)



T. Zhen, K.A. Klise, S. Cunningham et al. / Process Safety and Environmental Protection 132 (2019) 47–58

Mathematical Optimization for Sensor Placement

- Mixed-integer nonlinear programming (MINLP) formulation

Maximize expected coverage

$$\underset{x,\sigma}{\text{maximize}} \quad \sum_{e \in E} \sigma_e w_e$$

$$\text{subject to} \quad \sum_{l \in L} x_l \le k$$

Place at most k sensors

$$\sigma_e = 1 - 1[\prod_{l \in L_e} (1 - p_{l,e} x_l)] \qquad \forall e \in E$$

expected coverage of entity

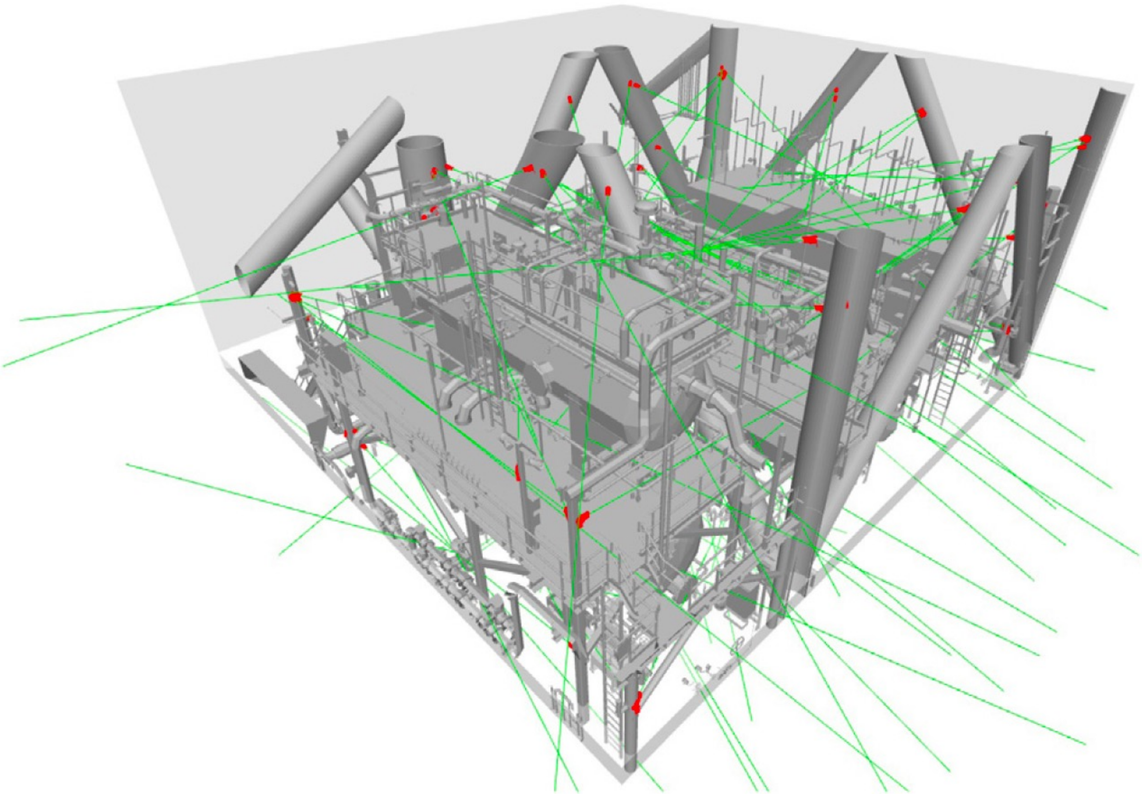
$$x_l \in \{0, 1\} \qquad \forall l \in L$$

$$0 \le \sigma_e \le 1 \qquad \forall e \in E$$

Binary variable, whether to place a sensor at location l

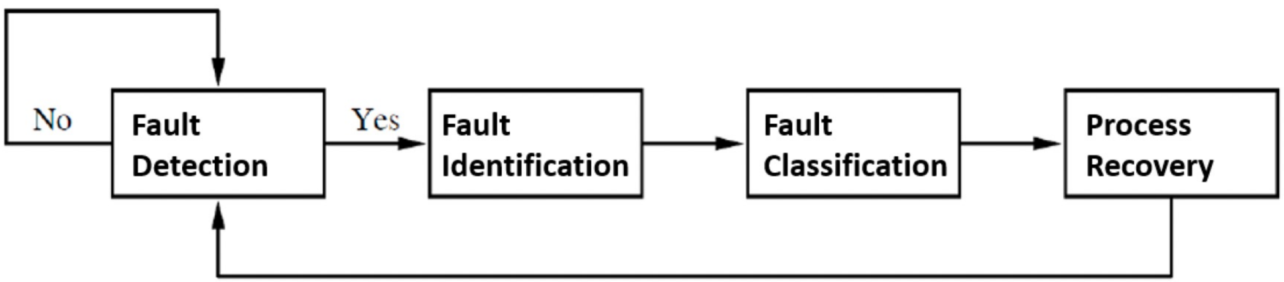
Sensor Placement under Uncertainty

- Facility with 81 candidate flame detector locations (Kenexis Consulting Corporation)
- Find the optimal configuration within 2 hours with a tailored algorithm



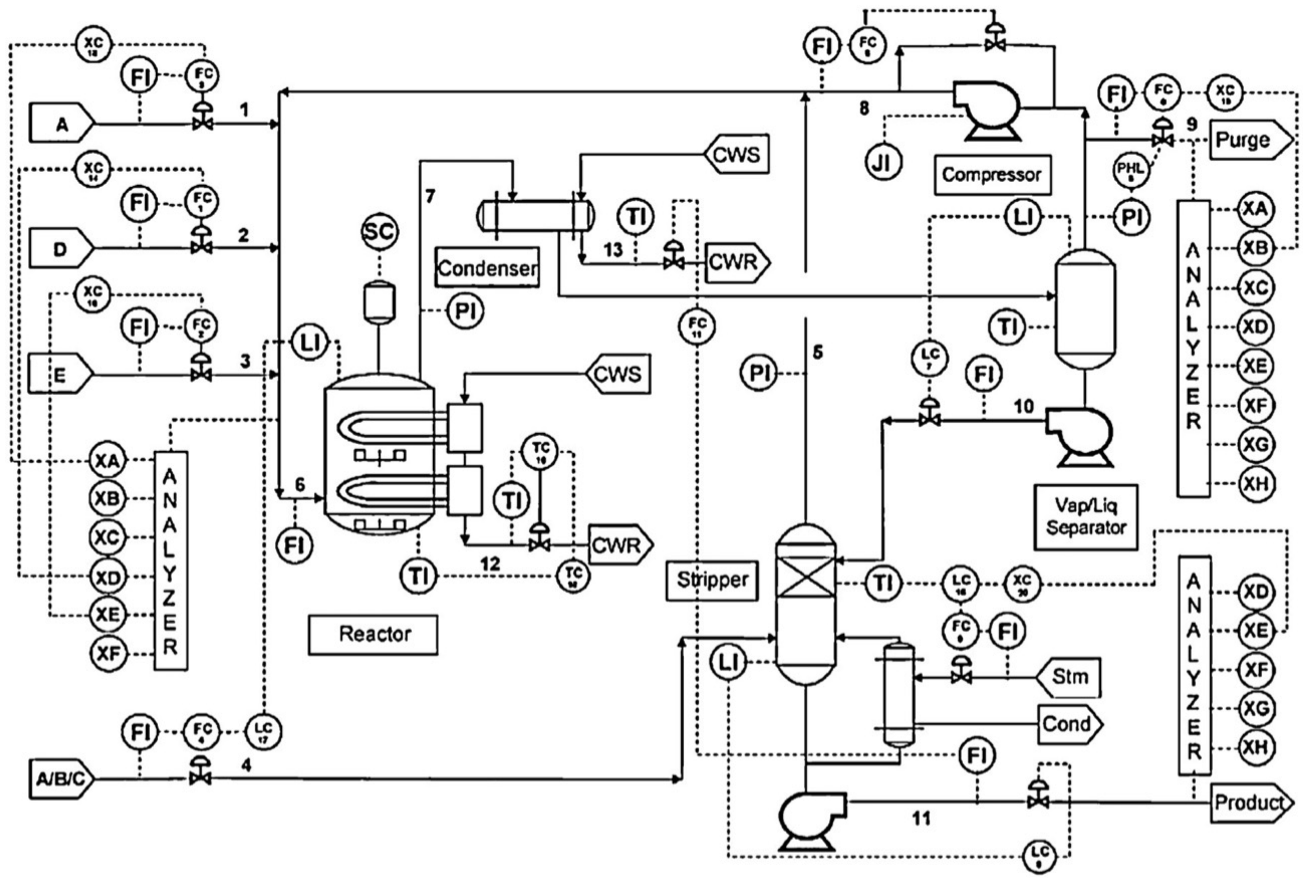
T. Zhen, K.A. Klise, S. Cunningham et al. / Process Safety and Environmental Protection 132 (2019) 47–58

Machine Learning for Process Monitoring

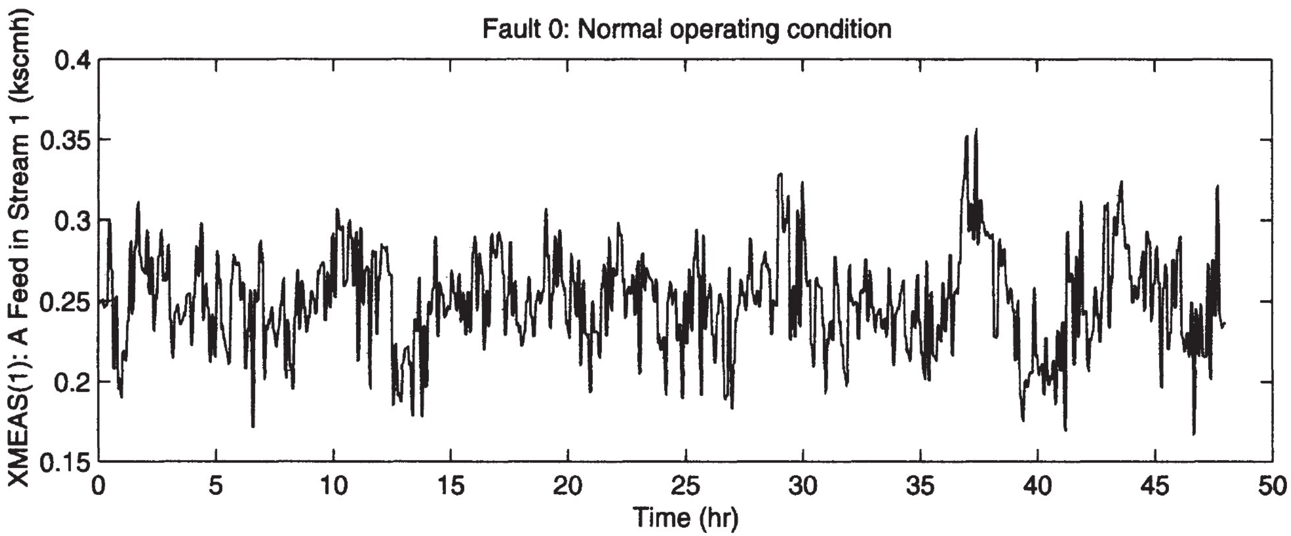


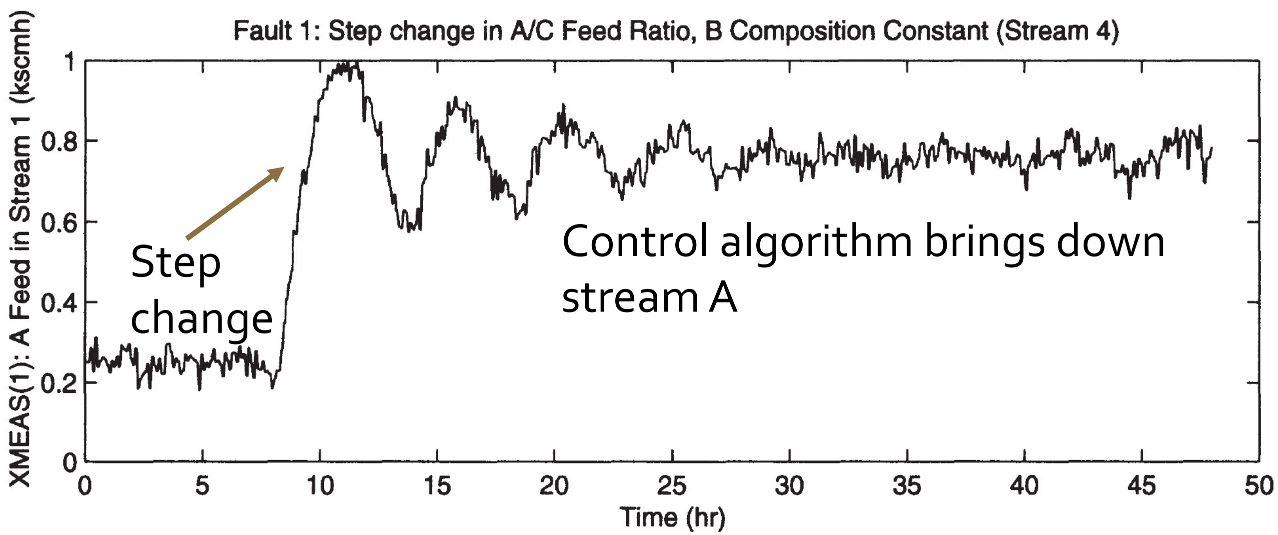
Tennessee Eastman process

- From measured state variables, perform fault detection



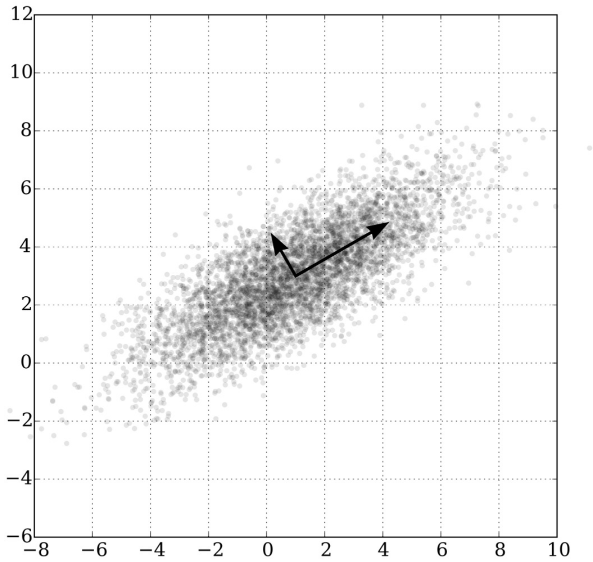
Example: Change in A/C Feed ratio

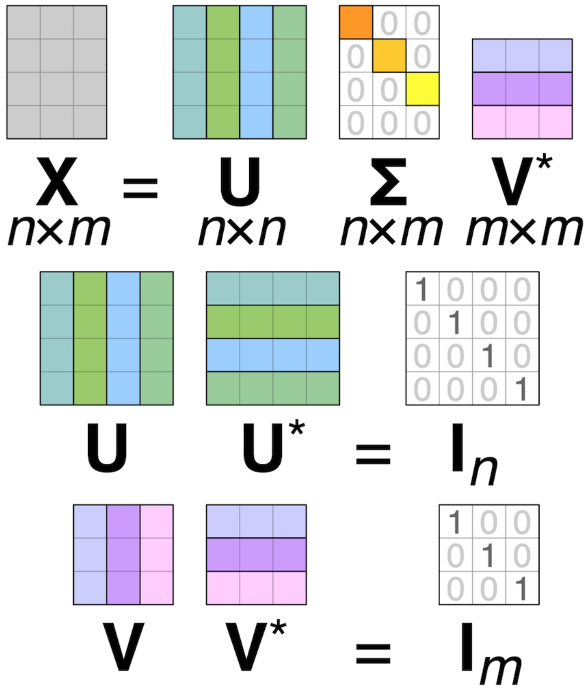




Classical Machine Learning Algorithm

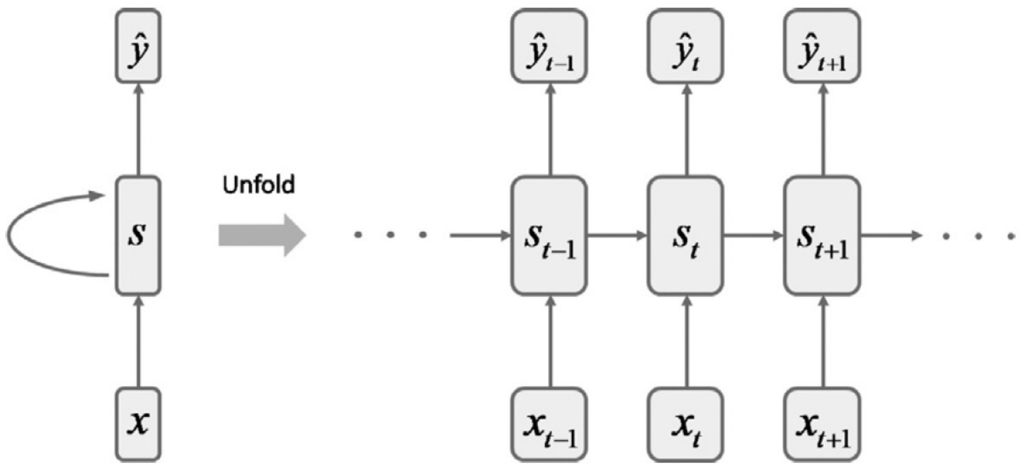
- Principal component analysis: identify the principal components where the data have the largest variance. The non-principal components are “noise”.
- Approach: singular value decomposition
- Pros: Interpretability
- Cons: Low accuracy for nonlinear processes.





Deep learning methods

- Recurrent neural network
- Pros: capable of handling nonlinearity and a high degree of spatio-temporal correlation
- Cons: hard to interpret



s_t:state

x_t: input

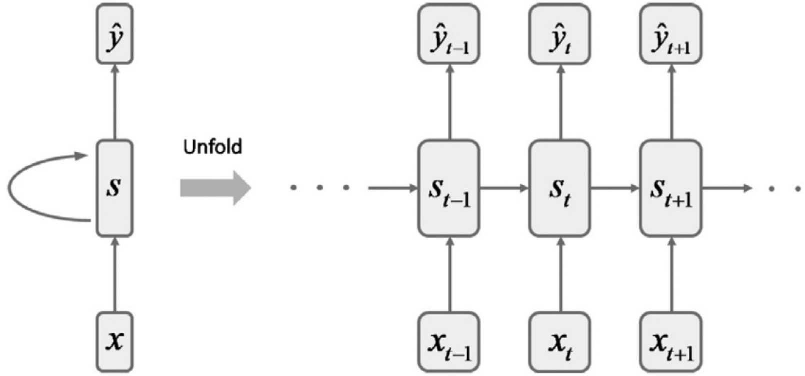
y_t: output

$$\boldsymbol{s}_t = f_s(\boldsymbol{x}_t, \boldsymbol{s}_{t-1} | \theta_s)$$
$$\hat{\boldsymbol{y}}_t = \boldsymbol{W}_y \boldsymbol{s}_t + \boldsymbol{b}_y$$

Explainable machine learning

- Ongoing research by PhD student Hao Chen

Machine learning-based models



Computationally efficient to use online
Hard to interpret

Physics-based models

$$\frac{dc}{dt} = \frac{F_0(c_0 - c)}{\pi r^2 h} - k_0 \exp\left(-\frac{E}{RT}\right) c$$

$$\frac{dT}{dt} = \frac{F_0(T_0 - T)}{\pi r^2 h} + \frac{-\Delta H}{\rho C_p} k_0 \exp\left(-\frac{E}{RT}\right) c + \frac{2U}{r \rho C_p}(T_c - T)$$

$$\frac{dh}{dt} = \frac{F_0 - F}{\pi r^2}$$

High fidelity but expensive to solve

Research Group

Fundamental Research	Applications	People and projects
Mathematical Optimization under Uncertainty	Energy systems optimization, chemical process design and operation, supply chain	Asha: Distributed manufacturing of the electrified processes Kaiyu: Integrated scheduling and control Casper: Electrification of Process Heating (co-advised with Dr. Cornelius Masuku)
Machine Learning	Fault detection, speech/image recognition, computational biology, Acceleration of optimization algorithms	Chi: end-to-end learning of optimization problems Hao: Explainable ML for fault detection Yen-Chun: ML for stem cell (co-advised with Dr. Xiaoping Bao)