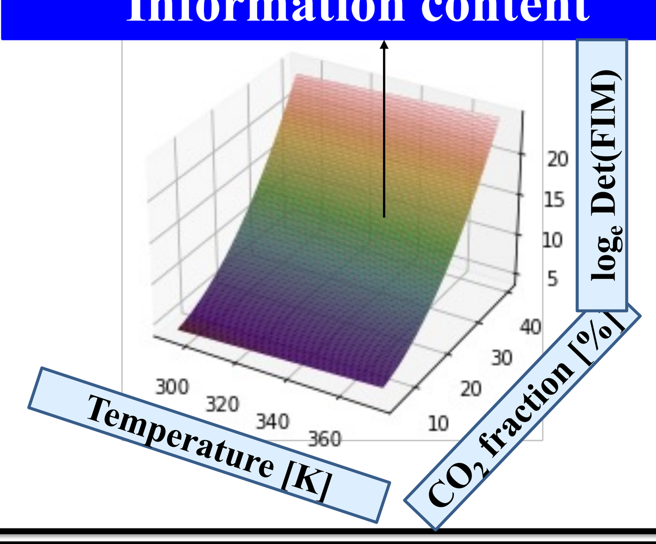
Characterizing the Pareto Optimal Trade-off between Science-Based Information Content and Measurements Cost

Pyomo.DoE Code

Jialu Wang, Alexander W. Dowling, University of Notre Dame

Science-Based Design of Experiments (SBDoE) Identifies the Most Informative Experiment

Information content



Experiments should be designed with a statistic design criterion!

- **What, where,** and **when** to measure?
- How many experiments?

First-principle model

Gradient-based optimization

1

Two Types of Measurements (Sensors)

Dynamic-cost

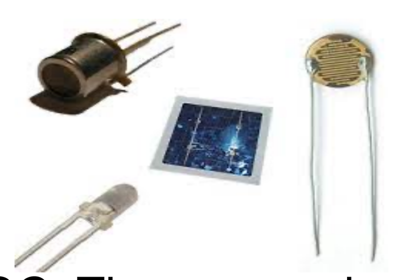
- Installation cost + cost per measurement
- Takes money or human resources at every timepoint

Static-cost (Special case)

- Installation cost + $~0 cost per timepoint
- If chosen, data at all timepoints are available

[1]

Manual measurement

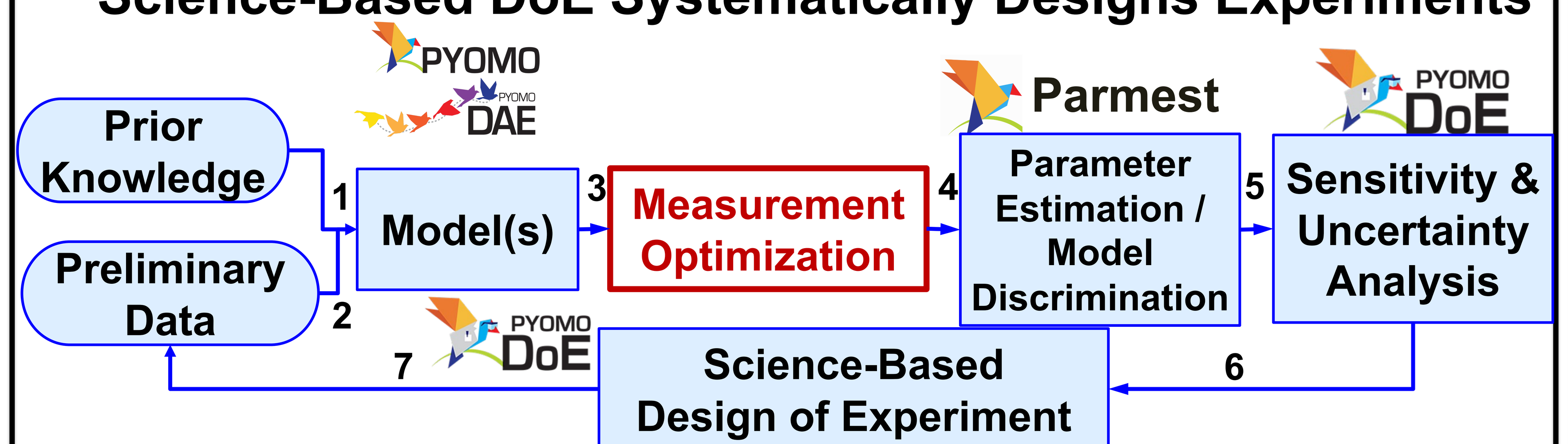


[2]

GC, Thermocouple…

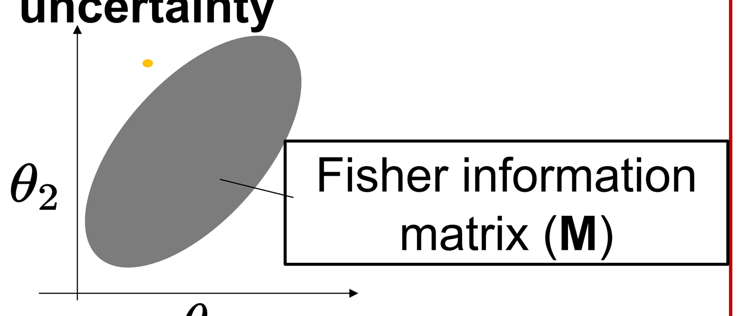
2

Science-Based DoE Systematically Designs Experiments



SBDoE maximizes Fisher Information Matrix

Aim: Minimize parameter uncertainty



SBDoE Optimization Problem

$$\max_{d} \quad \psi(\mathbf{M}+\mathbf{M}_0)$$
$$\text{s.t.} \quad \mathbf{M} \approx \mathbf{Q}^T \boldsymbol{\Sigma}_y^{-1} \mathbf{Q}$$

Model Equations (DAEs, PDAEs…)

1st order Sensitivity Matrix

$$\mathbf{Q}(\boldsymbol{d}) = \begin{bmatrix} \frac{\partial y_1}{\partial \theta_1}|_{t_1} & \cdots & \frac{\partial y_1}{\partial \theta_{N_p}}|_{t_1} \\ \vdots & \ddots & \vdots \\ \frac{\partial y_{N_y}}{\partial \theta_1}|_{t_r} & \cdots & \frac{\partial y_{N_y}}{\partial \theta_{N_p}}|_{t_r} \end{bmatrix}$$

Error Covariance Matrix

$$\boldsymbol{\Sigma}_y = \begin{bmatrix} \sigma^2(y_{1,t_1}, y_{1,t_1}) & \cdots & \sigma^2(y_{1,t_1}, y_{N,t_N}) \\ \vdots & \ddots & \vdots \\ \sigma^2(y_{N,t_N}, y_{1,t_1}) & \cdots & \sigma^2(y_{N,t_N}, y_{N,t_N}) \end{bmatrix}$$

$\boldsymbol{y}$: measurements
$\boldsymbol{\theta}$: parameters
$\boldsymbol{d}$: design variables

3

Measurement Optimization Framework

$$\max_{x} \quad \psi(\mathbf{M}+\mathbf{M}_0)$$
$$\text{s.t.} \quad \mathbf{M}_{i,j}$$

Goal: Maximize determinant or trace

$x \in \chi$: measurements, discrete decisions

$$= \sum^{S}\sum^{S} \mathbf{M}_{s,s',i,j} \cdot x_{s,s'} \quad \text{Static-Static}$$

$$+ \sum_{d}^{D}\sum_{d'}^{D}\sum_{t}^{t_d}\sum_{t'}^{t_{d'}} \mathbf{M}_{(d,t),(d',t'),i,j} \cdot x_{(d,t),(d',t')} \quad \text{Dynamic-Dynamic}$$

$$+ \sum^{S}\sum^{D}\sum^{t_d} \mathbf{M}_{s,(d,t),i,j} \cdot x_{s,(d,t)} \quad \text{Dynamic-Static}$$

D: Dynamic-cost measures

S: Static-cost measures

Continuous design relaxation guarantees with maximization

A-, D-optimality are concave

Convex Hull

$$0 \le \bar{x}_k \le 1$$
$$\bar{x}_{k,k'} \le \bar{x}_k$$
$$\bar{x}_{k,k'} \le \bar{x}_{k'}$$
$$\bar{x}_k + \bar{x}_{k'} - 1 \le \bar{x}_{k,k'}$$

Relaxed to only linear constraints

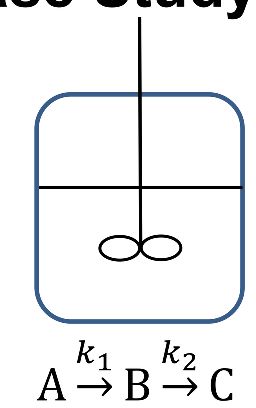
Budgets

$$\sum_{k \in K} c_k \cdot \bar{x}_k \le B$$
$$\sum_{d \in D}\sum_{t \in t_d} \bar{x}_{d,t} \le L_{total}$$
$$\sum_{t \in t_d} \bar{x}_{d,t} \le L_d, \forall d \in D$$
$$\sum_{t'}^{t'-t<T_{int}} \bar{x}_{d,t} \le 1, \forall d \in D, \forall t \in \boldsymbol{t}_d$$

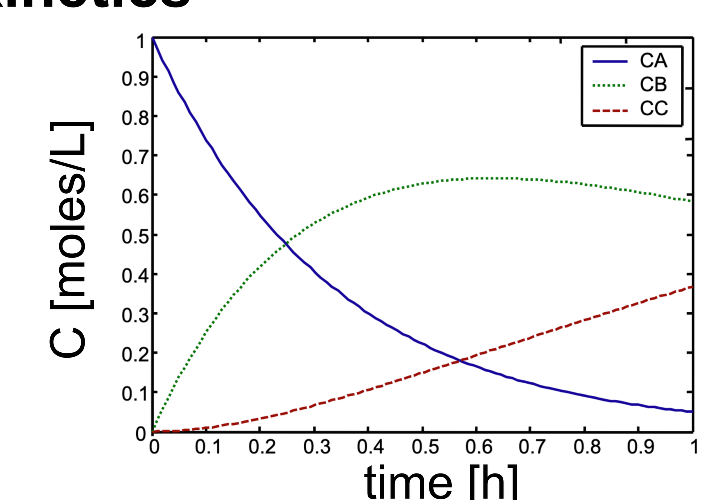
4

Case Study: Kinetics

$$k_1 = A_1 e^{-\frac{E_1}{RT}}$$
$$k_2 = A_2 e^{-\frac{E_2}{RT}}$$
$$\frac{dC_A}{dt} = r_A = -k_1 C_A$$
$$\frac{dC_B}{dt} = r_B = k_1 C_A - k_2 C_B$$
$$C_C = C_{A0} - C_A - C_B$$
$$C_B(t_0) = 0$$
$$C_C(t_0) = 0$$



$\boldsymbol{\varphi} = [C_{A0}, T(t)]$



- 4 unknown parameters (4 × 4 FIM)
- Algebraic variables
- Differential variables
  - 3 measurements
  - 8 timepoints / measurement

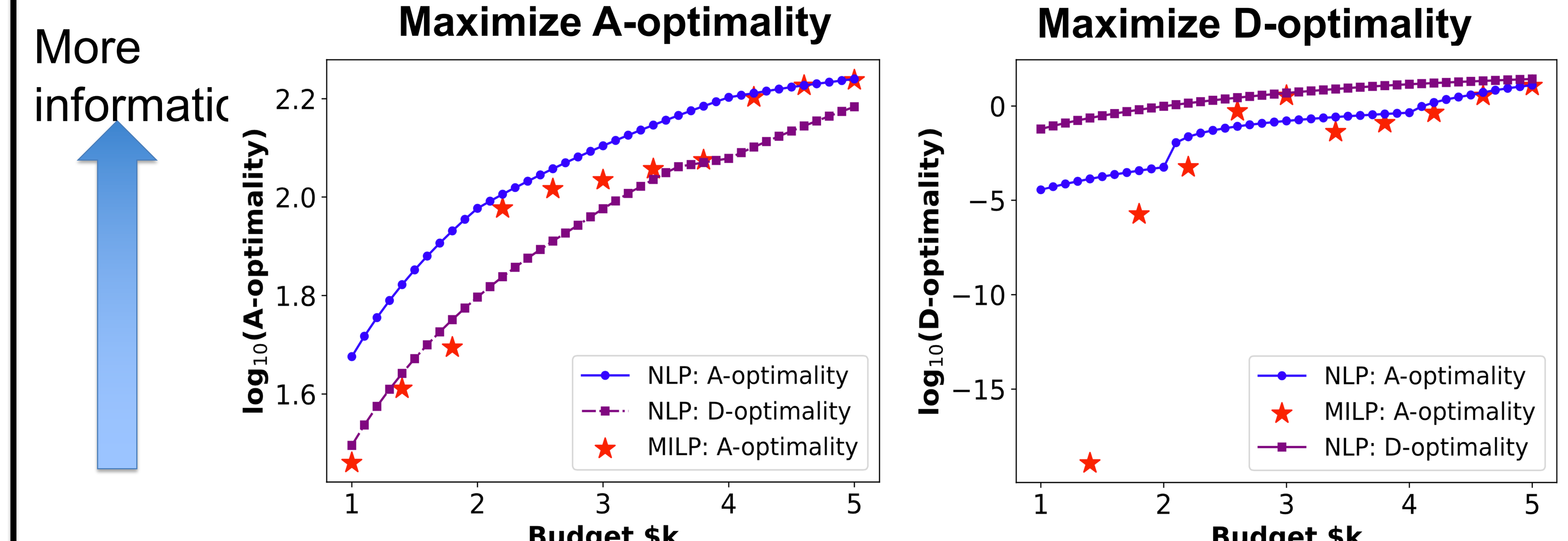
Measurements Cost

Name/Cost	Installation($)	Measure($/time)
C_A Static	2000	0
C_B Static	2000	0
C_C Static	2000	0
C_A Dynamic	200	400
C_B Dynamic	200	400
C_C Dynamic	200	400

5

A- and D-optimality Select Different Measurements

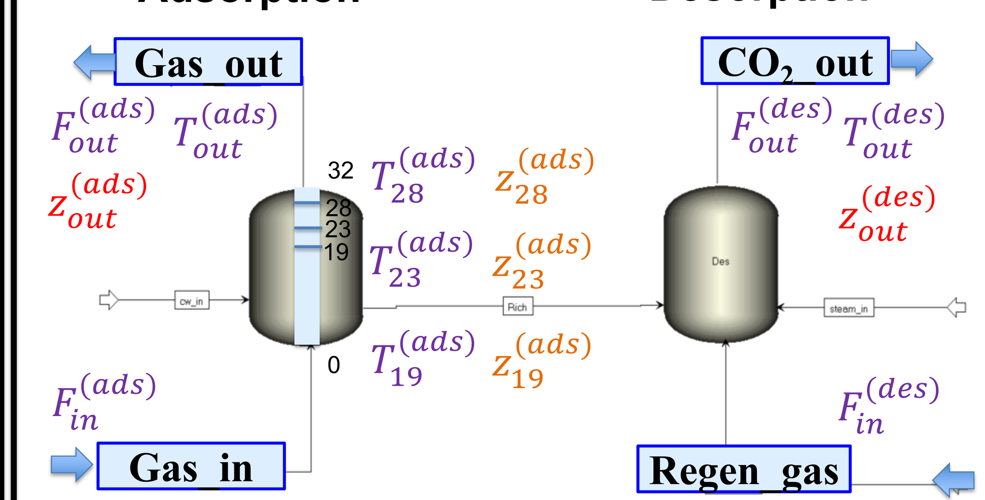
More information



6

Acknowledgements: The authors graciously acknowledge funding from the U.S. Department of Energy, Office of Fossil Energy and Carbon Management, through the Carbon Capture Program.

Case Study: Rotary Bed System (PDAE)

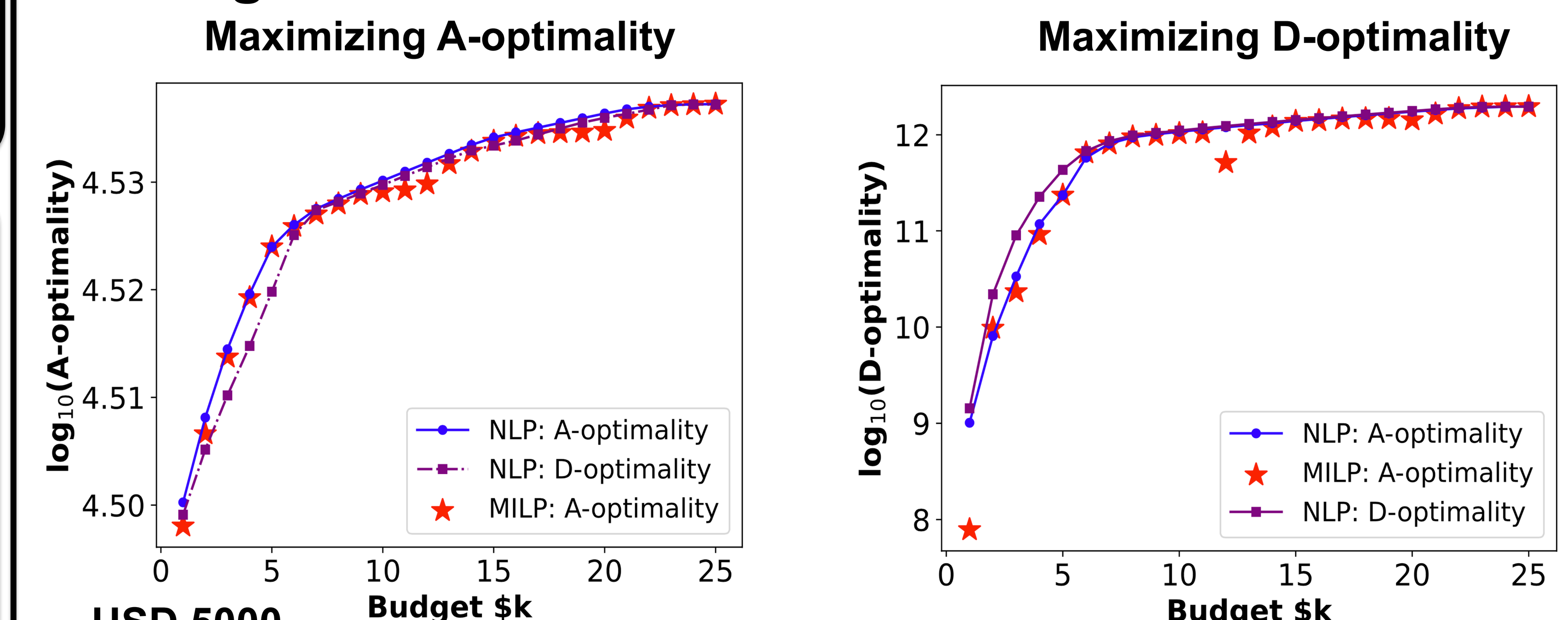


MILP for max A-optimality	NLP for max D-optimality
Variables:	Variables:
12 continuous	404 continuous
381 binary	
Constraints:	Constraints:
12 equality	23 equality
1,096 inequality	1,096 inequality

- 5 parameters (5 × 5 FIM)
- 110 timepoints / measurement
- 9 static-cost measurements
- 3 dynamic-cost measurements
- 2 either static- or dynamic-cost

7

Adding Measurements Increase the Information Content



USD 5000

Static-cost: $T_{out}^{(ads)}, T_{out}^{(des)}, T_{19}^{(ads)}, F_{out}^{(des)}, F_{in}^{(ads)}, F_{in}^{(des)}$

- Low budget: Static-cost measurements provide more information with lower price

USD 11000

Static-cost: $T_{out}^{(ads)}, T_{out}^{(des)}, F_{out}^{(ads)}, F_{out}^{(des)}, F_{in}^{(ads)}, F_{in}^{(des)}, T_{19}^{(ads)}, T_{23}^{(ads)}, T_{28}^{(ads)}$

Dynamic-cost:

$z_{out}^{(ads)}$ at 124, 134, 180, 190, 200 [min]

$z_{out}^{(des)}$ at 2, 12, 22, 32, 42 [min]

$z_{23}^{(ads)}$ at 158, 210, 220 [min]

- With more budget: More measurements are chosen
- Dynamic-cost measurements have a minimum 10 minutes interval for human sampling

USD 12000

Static-cost: $T_{out}^{(ads)}, T_{out}^{(des)}, F_{out}^{(des)}, F_{in}^{(ads)}, F_{in}^{(des)}, T_{19}^{(ads)}, z_{out}^{(ads)}$

- As budget increases, $z_{out}^{(ads)}$ is chosen as static-cost, instead of dynamic-cost before

8

Take-Aways

- Measurement optimization maximizes information content within budget while easily incorporating practical constraints.
- It can be done before building the experimental system, therefore improving data acquisition efficiency with low costs.

9

Reference: Wang, J., & Dowling, A. W. (2022). Pyomo. DOE: An open-source package for model-based design of experiments in Python. *AIChE Journal*, *68*(12), e17813.

Kusumo, K. P., Kuriyan, K., Vaidyaraman, S., García-Muñoz, S., Shah, N., & Chachuat, B. (2022). Risk mitigation in model-based experiment design: A continuous-effort approach to optimal campaigns. *Computers & Chemical Engineering*, *159*, 107680.

[1] Figure from Adobe Stock, [2] Figure from Wikipedia

CCSI² Carbon Capture Simulation for Industry Impact