Data Analytics for Membrane Material Innovations

Xinhong Liu, Jonathan A. Ouimet, William A. Phillip, Alexander W. Dowling

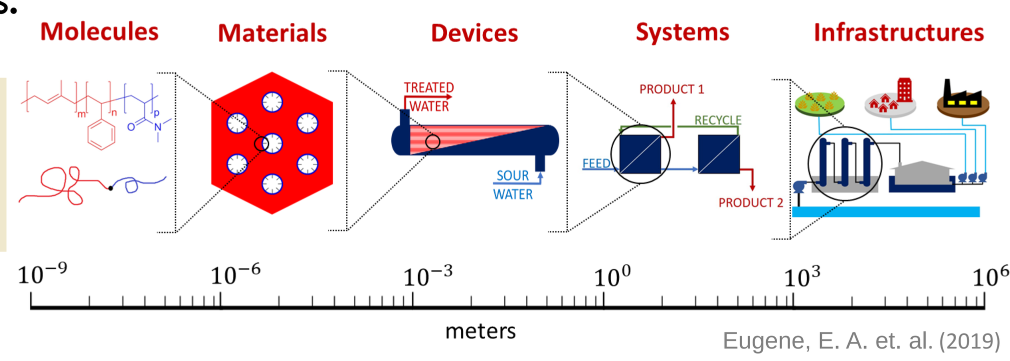
Department of Chemical and Biomolecular Engineering, University of Notre Dame

Introduction

Background

Structure-property-performance relationships provide the fundamental knowledge to guide both the inverse material design and the large-scale process design for novel membrane applications.

Molecular-to-Infrastructure Design

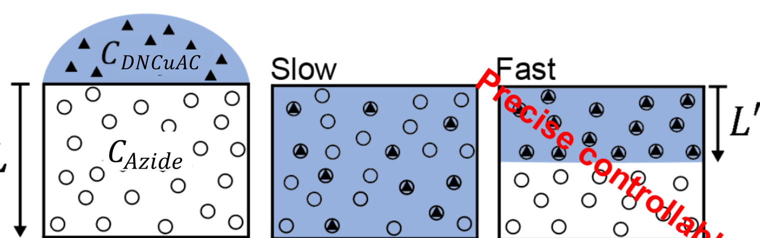


Eugene, E. A. et. al. (2019)

Purpose

How data analytics help facilitate membrane innovations?

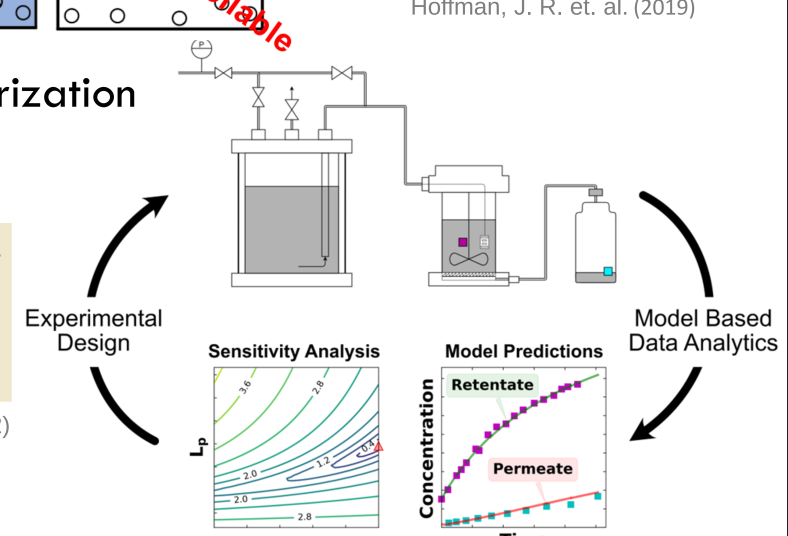
- Membrane Functionalization



Precise Control of the Structure

Hoffman, J. R. et. al. (2019)

- Membrane Characterization

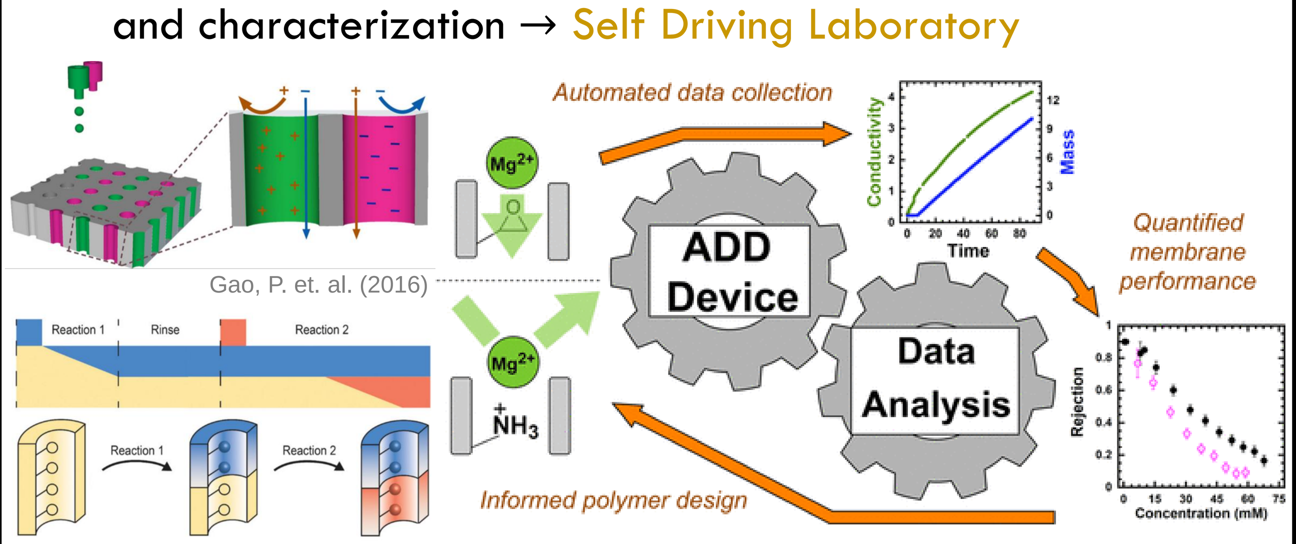


DATA: Diafiltration Apparatus for high-Throughput Analysis

Ouimet, J. A. et. al. (2022)

Future Work

- ❑ MBDoE for model discrimination, with 1000+ build from model hierarchy
- ❑ System automation for both membrane functionalization and characterization → Self Driving Laboratory



Gao, P. et. al. (2016)

Hoffman, J. R. et. al. (2020)

Muetzel, Z. W. et. al. (2022)

References & Additional Information

Reactive Inks	DATA	MBDoE	Dowling Lab	Email Contact
Liu, X., et. al. (2022)	Ouimet, J. A. et. al. (2022)	Liu, X., et. al. (2022)		

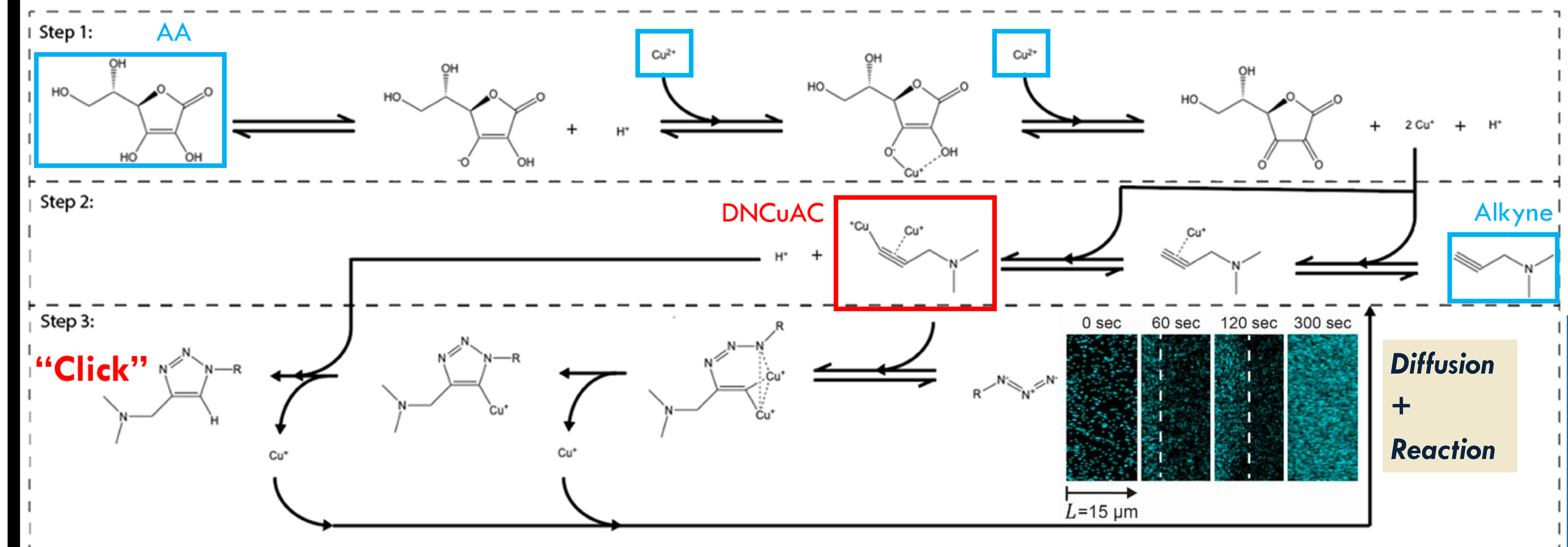
Acknowledgments

We gratefully acknowledge the support from NSF:

- Award CBET-2147605
- Award CBET-1941596
- Award 1932206

Precise Controllable Membrane Functionalization

"Click" reaction



Diffusion + Reaction

Experimentalists loaded excess harmful components (i.e., copper) to achieve a high concentration of DNCuAC for fast reaction.

Visually observe the functionalization process by Scanning Electron Microscopy−Energy Dispersive X-ray Spectrometry (SEM-EDX)

Hoffman, J. R. et. al. (2019)

Optimize reactive ink formulation

$$\min_x \; w^T x \quad \text{- Cost of ink}$$

$$s.t. \quad \theta(t_{end}) \geq \theta^*$$

- Fast conversion requirements

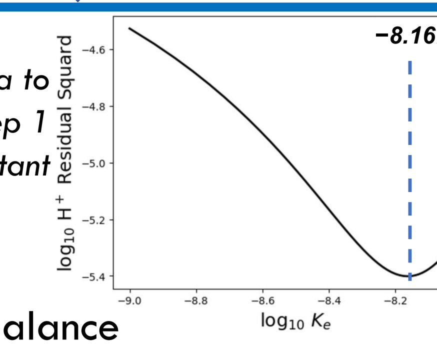
$$\theta(t_{end}) = f_3(C_{D0}, C_{A0})$$

- Step 3 model

$$C_{D0} = f_{1,2}(x)$$

- Step 1&2 model

Use pH data to identify step 1 equilibrium constant



Equations:

- Charge balance
- Reaction (main + side) equilibriums
- Component conservations

Diffusion-reaction

$$\frac{\partial C_D}{\partial t} = D\frac{\partial^2 C_D}{\partial z^2} - k_{az} C_A \cdot C_D$$

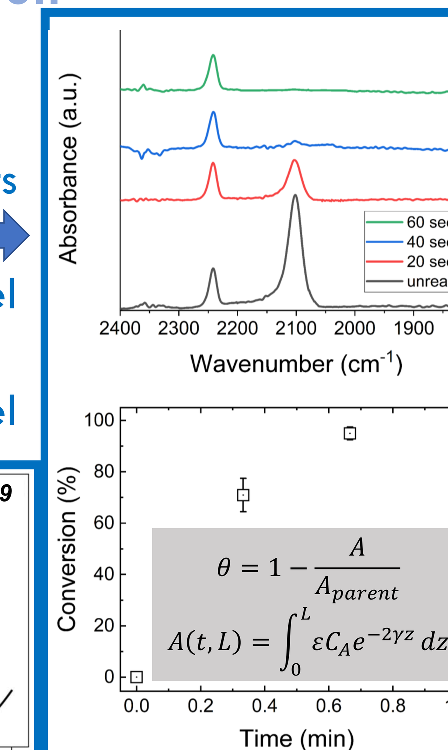
$$\frac{\partial C_A}{\partial t} = -k_{az} C_A \cdot C_D$$

I.C. $C_D(0,z) = 0, \forall z > 0$

$C_A(0,z) = C_{A0}, \forall z \geq 0$

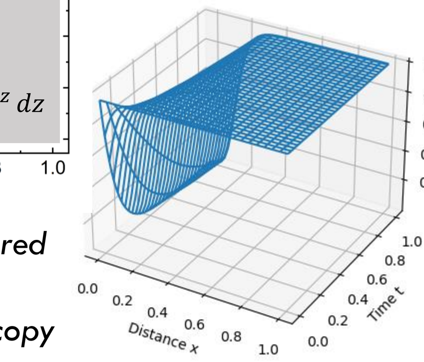
B.C. $C_D(t,0) = C_{D0}, \forall t \geq 0$

$\frac{\partial C_D}{\partial z}(t,L) = 0, \forall t \geq 0$



$$\theta = 1 - \frac{A}{A_{parent}}$$

$$A(t, L) = \int_0^L \varepsilon C_A e^{-2\gamma z} dz$$



Azide Conversion Monitered with Fourier Transform Infrared (FTIR) Spectroscopy

Hoffman, J. R. et. al. (2019)

High-Throughput Membrane Characterization Exploring Structure-Property-Performance Relationships

Diafiltration model & parameter estimation

Parameters

- L_p: hydraulic permeability
- B : solute permeability coefficient
- σ : reflection coefficient

Solution-Diffusion

Water flux across the membrane

$$J_w = L_p \Delta P - \sigma \Delta \pi$$

$$\Delta \pi = nRT(c_{in} - c_h)$$

Solute flux across the membrane

$$J_s = B(c_{in} - c_h)$$

Concentration Polarization

$$\frac{c_{in} - c_h}{c_f - c_h} = exp\left(\frac{J_w}{k}\right)$$

$$\frac{kb}{D} = 0.23\left(\frac{bc^0}{\nu}\right)^{0.57}\left(\frac{\nu}{D}\right)^{0.33}$$

DAEs	Time steps	Diafiltration reservoir	Stirred cell	Hold-up	Collecting vial
Start-up ("Lag" mode)	Hold-up volume filled	$\frac{dm_d}{dt} = 0$	$\frac{dm_f}{dt} = -A_m\rho J_w$; $\frac{d(c_f m_f)}{dt} = -c_d\frac{dm_d}{dt} - A_m\rho J_s$	$\frac{dm_h}{dt} = A_m\rho J_w$; $\frac{d(c_h m_h)}{dt} = A_m\rho J_s$	-
	No diafiltrate induced			$\frac{dm_h}{dt} = 0$; $\frac{d(c_h m_h)}{dt} = A_m\rho J_s - c_h A_m\rho J_w$	$\frac{dm_v}{dt} = A_m\rho J_w$; $\frac{d(c_v m_v)}{dt} = \frac{dm_v}{dt}c_h$
	First drop of diafiltrate	$\frac{dm_d}{dt} = -S - A_m\rho J_w$	$\frac{dm_f}{dt} = S$; $\frac{d(c_f m_f)}{dt} = -c_d\frac{dm_d}{dt} - A_m\rho J_s$		
Vial collection	First drop in vial				

$$\hat{\theta} = \arg\min_{\theta} \sum_i w_{m,i}(m_{v,i} - \hat{m}_{v,i})^2$$

Mass collected

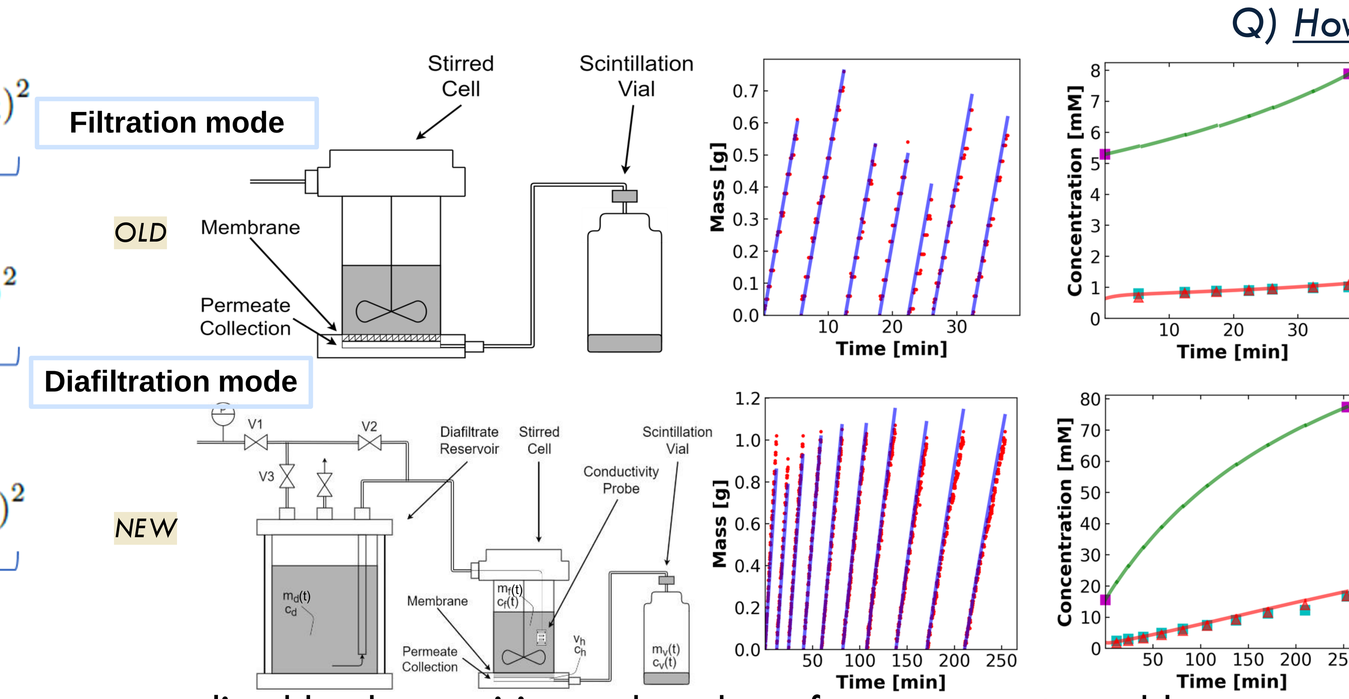
$$+ \sum_j w_{c_v,j}(c_{v,j} - \hat{c}_{v,j})^2$$

Vial concentration

$$+ \sum_k w_{c_f,k}(c_{f,k} - \hat{c}_{f,k})^2$$

Retentate concentration

s.t. DAE constraints



A) Weighted least-square normalized by the precision and number of measurements enables nonlinear parameter estimation from different type of data in dynamic systems.

Q) *How to use different types of data to calibrate the model parameters?*

Fitted parameters

L_p	B	σ
4.30 L/(m²-hr-bar)	0.67 μm/s	0.83 dimensionless
3.25 L/(m²-hr-bar)	0.31 μm/s	1 dimensionless

Fisher Information Matrix (FIM) Analyses

Q1) *Are the parameters reliable?*

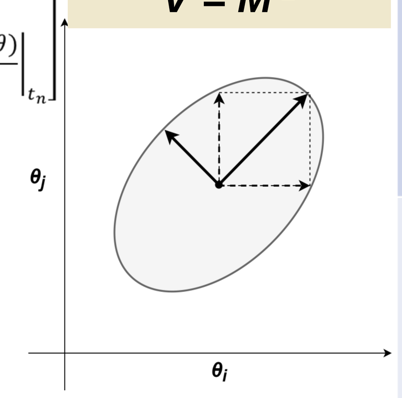
$$M = V_{\hat{\theta}}^{-1} = Q^T \Sigma^{-1} Q$$

$$Q(\theta) = \begin{bmatrix} \frac{\partial f(x_1,\theta)}{\partial \theta_1}\Big|_{t_1} & \cdots & \frac{\partial f(x_1,\theta)}{\partial \theta_m}\Big|_{t_1} \\ \vdots & \ddots & \vdots \\ \frac{\partial f(x_n,\theta)}{\partial \theta_1}\Big|_{t_n} & \cdots & \frac{\partial f(x_n,\theta)}{\partial \theta_m}\Big|_{t_n} \end{bmatrix}$$

$$\Sigma = \begin{bmatrix} \sigma_{x_1}^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_{x_n}^2 \end{bmatrix}$$

- M - Fisher information matrix
- $V_{\hat{\theta}}$ - Covariance matrix for $\hat{\theta}$
- Q - Sensitivity matrix
- Σ - Output error matrix
- $\hat{\theta}$ - Best fit parameters

Confidence Ellipse for Covariance Matrix $V = M^{-1}$

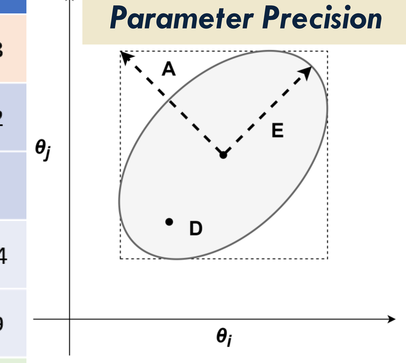


Mode	Eigenvalues of M	Eigenvectors of M		
		L_p	B	σ
Filtration	4.85E+03	0.086	0.0033	0.9963
	1.63E+05	0.002	-1	0.0032
	4.37E+07	-0.9963	-0.0017	0.086
Diafiltration	1.96E+06	-0.1405	-0.9468	-0.2894
	1.27E+08	0.9503	-0.211	0.2289
	8.18E+08	-0.2778	-0.2429	0.9294

A1) Eigen decomposition of FIM reveals diafiltration contains more information to infer the parameters, especially σ.

Q2) *Does the investment in new instrument and additional data collection worth it?*

FIM-Based Metrics for Parameter Precision

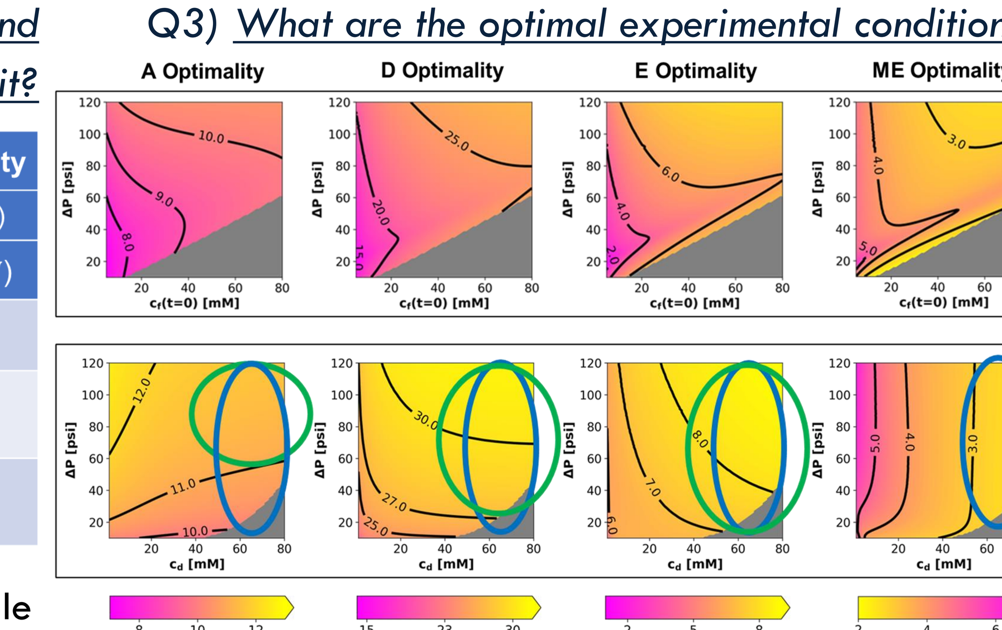


- A - Dimensions of the enclosing box
- D - Area or volume
- E - Size of major axis

min $g(V)$ ⇕ max $g(M)$	A-optimality	D-optimality	E-optimality
	trace(V)	det(V)	max eig(V)
	trace(M)	det(M)	min eig(M)
Initial	9.87E+08	9.12E+22	3.93
Incorporate New Sensor	2.06E+10 1E2↑	1.93E+24 1E2↑	9.55 143%↑
Model Start-up	1.95E+13 1E3↑	2.76E+30 1E6↑	207.91 1E2↑

A2) FIM-based metrics indicate valuable experiment modifications.

Q3) *What are the optimal experimental conditions?*



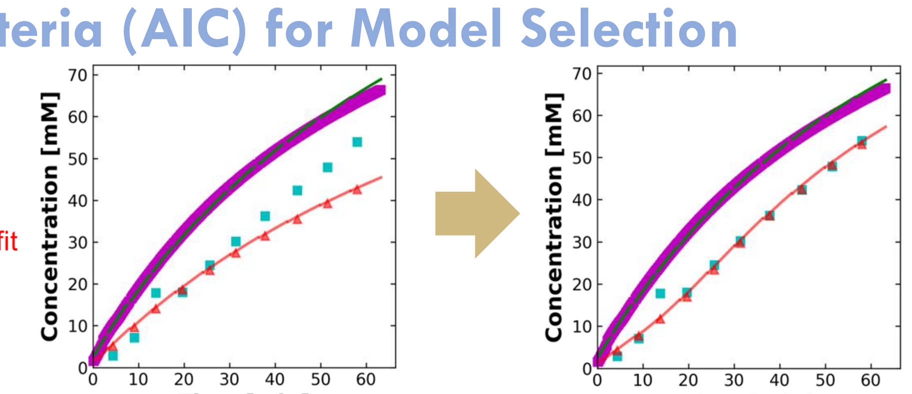
A3) Diafiltration experiments should be performed with $c_d \geq 50$ mM and $\Delta P \geq 50$ psi.

Akaike Information Criteria (AIC) for Model Selection

$$\Delta AIC = N \cdot \ln\left(\sum_i w_i (y_i - \hat{y}_i)^2\right) + 2n_p$$

Quality of fit — Penalty of overfit

- N - Number of measurements
- n_p - Number of model parameters
- y - Model predictions
- $\hat{y}$ - Measurements



$B = constant$ → $B = \sum_{i \in I} b_i c_f^{\,i}$

Predictions from Empirical Models for Charged Membrane

Q) *What are the dominant phenomena & mechanisms?*

A) For charged membrane NF270, B shows a quadratic dependence on feed side concentration.

The calibrated model is ready to be embedded in the scale-up, design, and optimization of membrane-based separation systems with high computing performance. It also can be applied to inform and filter the phenomenological models imported into model discrimination problems.

Science-Based Data Analytics Workflow