UQ Indices, Information Inequalities & applications in chemical kinetics

Markos A. Katsoulakis

University of Massachusetts Amherst

November 16th, 2017

IPAM Workshop IV: Uncertainty Quantification for Stochastic Systems and Applications

- Georgios Arampatzis (ETH, Zurich)
- Kostis Gourgoulias (babylon)
- Eric Hall (UMass, Amherst)
- Jinchao Feng (UMass, Amherst)
- Yannis Pantazis (University of Crete)
- Pedro Vilanova (UMass, Amherst)
- Jie Wang (UMass, Amherst)
- Paul Dupuis (Applied Math, Brown)
- Petr Plechac (Math, U. of Delaware)
- Luc Rey-Bellet (Math&Stat, UMass Amherst)
- Dion Vlachos (Chem. Engineering & Energy Institute, U. of Delaware)
- Ivan Lee (Army Research Lab, Adelphi)

Outline

Introduction

Towards Scalable UQ Methods

UQ Information Inequalities

1. Fast Sensitivity Screening in Complex Reaction Networks

2. UQ in structured probabilistic models: multiple sources of uncertainty

3. Optimization and Design under model-from uncertainty

A probabilistic perspective to Uncertainty Quantification

1. Model not described by a single probability model (measure) P but rather, due to uncertainties*, by a set of probability measures:

$$\mathcal{Q} = \{Q \text{ "close" to some } P\}$$

2. Given a **Quantity of Interest (QoI)** f, can one find uncertainty bounds/guarantees :

$$\inf_{Q \in \mathcal{Q}} E_Q[f] \leq E_Q[f] \leq \sup_{Q \in \mathcal{Q}} E_Q[f] \quad \text{all } Q \text{ "close" to } P$$

3. Desirable bounds: tight, computable, scalable, ??

---

*Sources of uncertainty

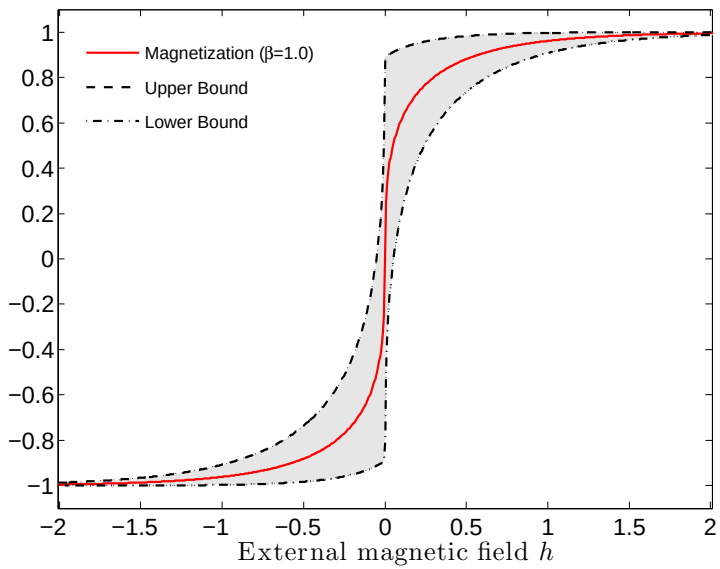
- Uncertainties in parameters, $P = P^\theta$ and $\theta \approx \theta^*$ where θ^* may be due to statistical estimation, sparse data, etc.
- The set $\mathcal{Q}$ may be determined by numerical approximations or approximate inference, such as minimizing relative entropy.
- "Model-form" Uncertainty: "U on U"

Model Bias - Probability metrics

Uncertainty bounds $\mapsto$ Model Bias:

$$I^+(f, P; \eta) := \max_{Q \in \mathcal{Q}} E_Q f - E_P f\,, \quad I^-(f, P; \eta) = \min_{Q \in \mathcal{Q}} E_Q f - E_P f$$

Example: Uncertainty in predictions[1] as function of a control parameter h.
(see optimization section)



How to build $\mathcal{Q} = \{Q \text{ "close" to } P\}$?

- Need probability metrics/divergences to discriminate between two probabilistic models P and Q:

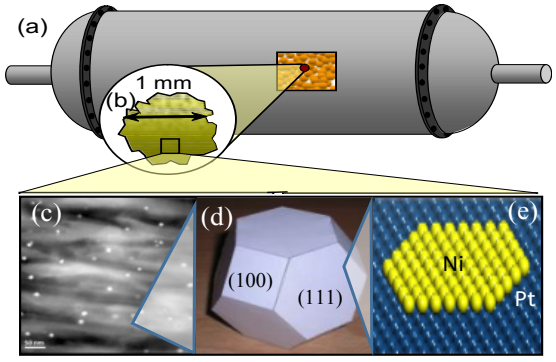
$$d(P, Q) \geq \eta \quad \text{or} \quad d(P, Q) \leq \eta \quad \eta : \text{error tolerance}$$

- **What is a suitable metric?**

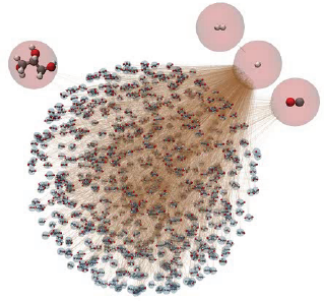
[1] K., Rey-Bellet, Wang, **J. Comp. Phys.**, (2017)

Scalable UQ and Sensitivity Analysis

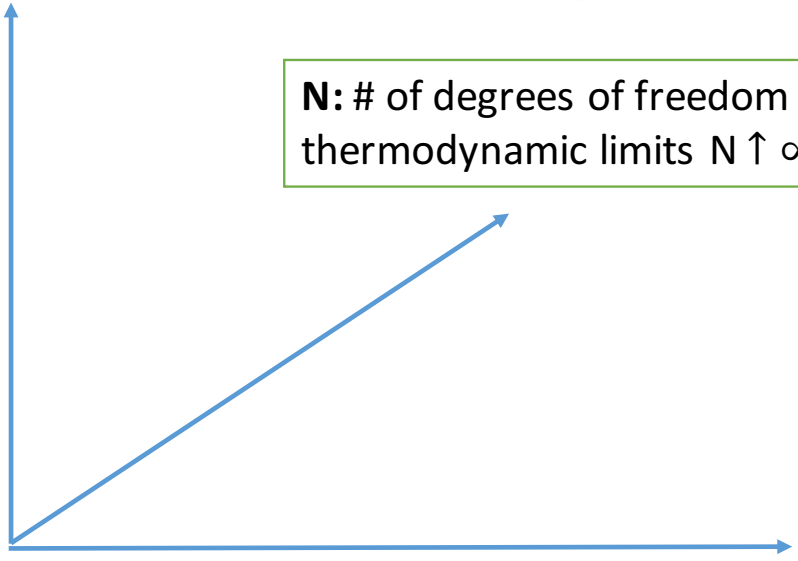
Necessary Scaling Properties for Uncertainty Quantification methods



T, N, K >>1

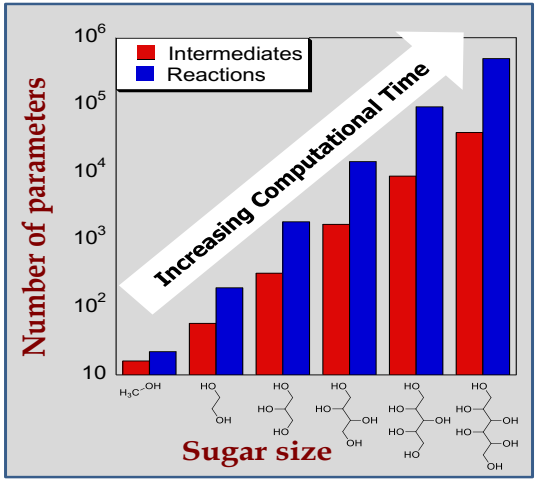


T: time, long-time integration steady state, T ↑ ∞



N: # of degrees of freedom thermodynamic limits N ↑ ∞

K: # of model parameters



Test some probability metrics/divergences

- **Kullback-Leibler Divergence** (Relative Entropy) between P and Q :

$$R(P|Q) = \int \log \frac{dP}{dQ} dP;$$

- **Hellinger distance**:

$$H(P,Q) = (\int (\sqrt{dP} - \sqrt{dQ})^2)^{1/2} = (\int (\sqrt{p} - \sqrt{q})^2 d\mu)^{1/2};$$

  - $0 \leq H^2(P,Q) \leq 2$

- χ^2 **divergence**:

$$\chi^2(P \mid\mid Q) = \int (\frac{dP}{dQ} - 1)^2 dQ.$$

- Are they applicable for high-dimensional or long-time regimes ?
- Can we control in terms of $d(P_N, Q_N)$, for $N \gg 1$

$$\mathbb{E}_{P_N}[f] - \mathbb{E}_{Q_N}[f]$$

- First test product probabilities P_N and Q_N where $N \gg 1$

$$P_N(\sigma_1, ..., \sigma_N) = \prod_{i=1}^{N} P_i(\sigma_i)\,, \qquad Q_N(\sigma_1, ..., \sigma_N) = \prod_{i=1}^{N} Q_i(\sigma_i)$$

- QoIs for $\sigma = (\sigma_1, ..., \sigma_N)$: $f(\sigma) = \frac{1}{N} \sum_{i=1}^{N} g(\sigma_i)$, sample variance, ...

Some (poor) scaling properties

- **Csiszar-Kullback-Pinsker:** $| E_{P_N}(f) - E_{Q_N}(f) | \leq \|f\|_\infty \sqrt{2R(P_N|Q_N)}$
  Chain Rule for KL: $R(P_N|Q_N) = N \times R(P|Q)$:

$$| E_{P_N}(f) - E_{Q_N}(f) | \quad \leq \quad \sqrt{N}\|f\|_\infty \sqrt{2R(P|Q)}$$

- **Hellinger:** $| E_{P_N}[f] - E_{Q_N}[f] | \leq || f ||_\infty H(P_N || Q_N)$
  However, for $P \neq Q$

$$H^2(P_N || Q_N) = 2\left(1 - (1 - \frac{H^2(P || Q)}{2})^N\right) \overset{N\to\infty}{\to} 2$$

- **Chapman Robbins:** $| E_{P_N}(f) - E_{Q_N}(f) | \leq \sqrt{Var_{P_N}(f)}\sqrt{\chi^2(Q_N || P_N)}$

- $\mathrm{Var}_{P_N}(f) = \frac{1}{N}\mathrm{Var}_P(g)$ and $\chi^2(Q_N || P_N) = \left(1 + \chi^2(Q || P)\right)^N - 1$
- Then for $P \neq Q$: $\mathrm{Var}_{P_N}(f)\chi^2(Q_N || P_N) \overset{N\to\infty}{\to} \infty$

1. $\|f\|_\infty$ can easily take large values, e.g. reaction networks;
2. When $N \gg 1$, the bounds either blow up or lose their selectivity.

[1] K., Rey-Bellet, Wang, **J. Comp. Phys.**, (2017)

Many parameters - why UQ/'systems tasks' scale poorly?

- **Example:** System of **N** ODEs

$$\dot{y} = f(y;\theta)\,, \qquad y(0) = y_0 \in \mathbb{R}^N$$

  - Goal: Perform SA on the model parameters $\theta \in \mathbb{R}^K$.

- Define sensitivity indices:

$$s_k = \frac{\partial y}{\partial \theta_k}$$

- A new system of ODEs is derived and augmented to the previous:

$$\dot{s}_k = \frac{\partial f}{\partial y} s_k + \frac{\partial f}{\partial \theta_k}, \;\; k = 1, ..., K$$

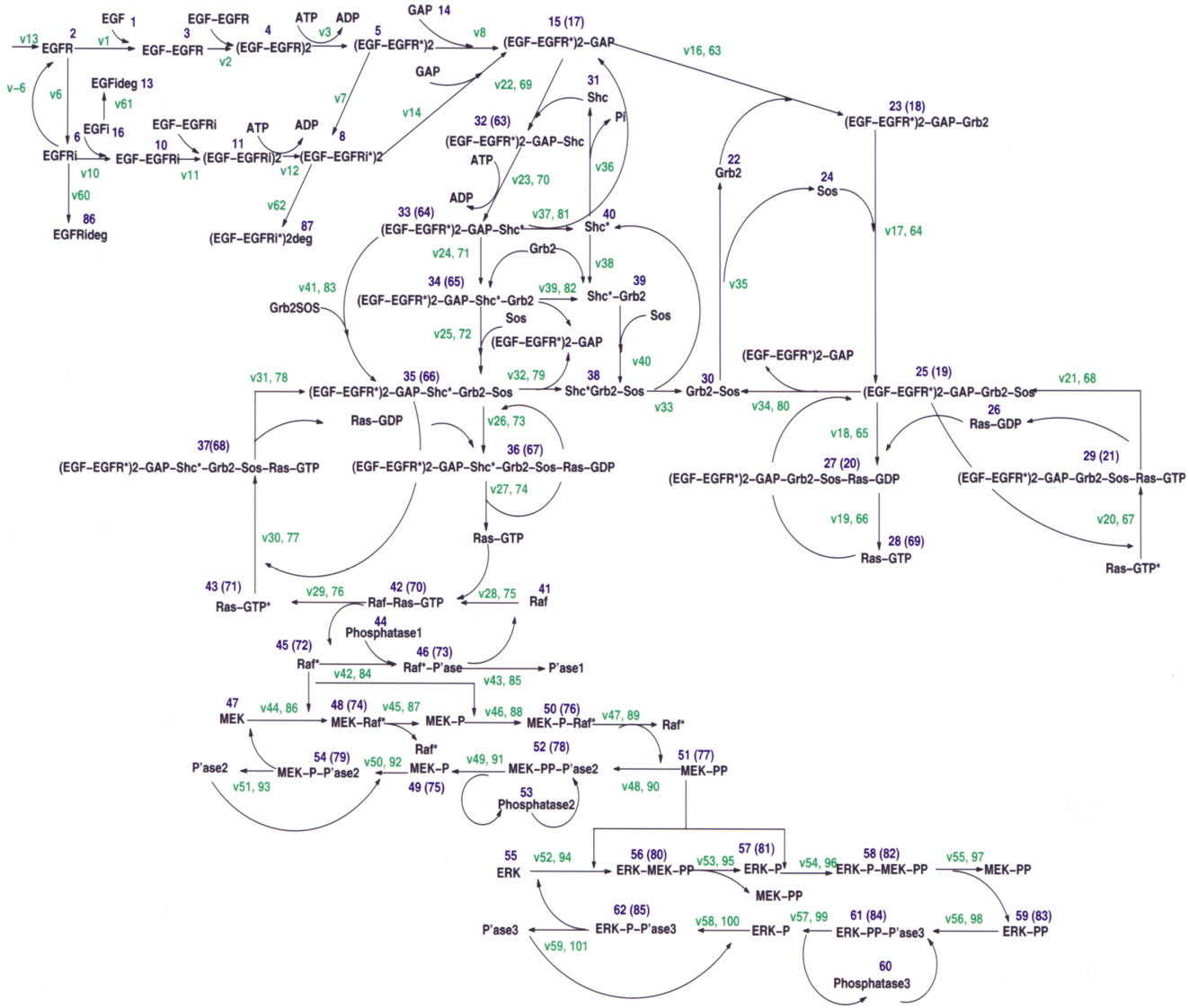
  - need to solve $K \times N$ additional equations.
  - *Long-time integration* necessary: engineering operating regimes are (usually) at steady state!

Stochastic Dynamics

have significant additional challenges: *high variance, rare events*, etc.

This challenge is common:

Similar issues can also arise in numerous network dynamics problems, e.g. in queueing systems, systems biology[1], chemical reaction networks (later), ...



[1] Schoeberl et al., **Nature Biotechnology**, 2002

A tool towards these challenges: UQ Info. Inequalities

$$\Xi_-(Q|P;f) \leq E_Q[f] - E_P[f] \leq \Xi_+(Q|P;f)\,,$$

- **A new divergence:** $\Xi_\pm(Q|P;f)$
  - $\Xi_\pm(Q|P;f) \geq 0$, (resp. $\Xi_-(Q|P;f) \leq 0$)
  - $\Xi_\pm(Q|P;f) = 0$ if and only if $P = Q$ a.s. or f is deterministic $P-$a.s.
  - Divergence contains information on QoI f, e.g.
- **Variational representation**:

$$\Xi_+(Q|P;f) = \inf_{\lambda>0}\left\{\frac{1}{\lambda}\log E_P[e^{\lambda(f-E_P[f])}] + \frac{1}{\lambda}R(Q|P)\right\}$$

- Bounds scale[1] with system size N, time T

;

[1] Dupuis, Katsoulakis, Pantazis, Plechac *SIAM JUQ* '16 Katsoulakis, Rey-Bellet, Wang, *J. Comp. Phys.* '17; earlier work: Chowdhary and Dupuis, ESAIM '13.

Some needed theory work: How to compute/estimate $\Xi_\pm$?

- Advanced **Monte Carlo** for the direct simulation of $\Xi_\pm$

  (see related Phase I work)
- **Linearization**[1]: $\Xi_+(Q|P;f) = \sqrt{\mathrm{Var}_P[f]}\sqrt{2R(Q|P)} + \mathcal{O}(R(Q|P))$
- **Concentration Inequalities**[2]**:** Find $\Phi : \mathbb{R} \to (0,\infty]$ such that

$$E_P[e^{\lambda(f-E_P[f])}] \leq \Phi(\lambda) \quad \text{for all } \lambda \in (-\lambda_0, \lambda_0)\,, \tag{1}$$

$$\text{Substitute in} \quad \Xi_+(Q|P;f) = \inf_{\lambda>0}\left\{\frac{1}{\lambda}\log E_P[e^{\lambda(f-E_P[f])}] + \frac{1}{\lambda}R(Q|P)\right\}$$

$$\text{to obtain} \quad U_+(Q|P;f) := \inf_{\lambda>0}\left\{\frac{1}{\lambda}\log \Phi(\lambda) + \frac{1}{\lambda}R(Q|P)\right\}$$

Typical examples of concentration inequalities:

1. Hoeffding, Bennett, sub-Gaussian
2. McDiarmid inequalities
3. For Markov Chains, Markov Random Fields, quantum systems, etc.

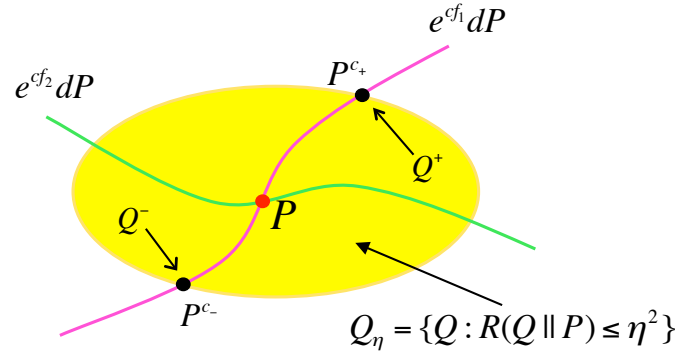
---

[1] P. Dupuis, M. K., Y. Pantazis, P. Plechac *SIAM/ASA J. of UQ* '16; H. Lam *Math. Oper. Res.* (2016)

[2] Gourgoulias, K. Rey-Bellet, Wang, *Arxiv*, (2017)

Tightness of bounds and implications: a UQ Index

- Tightness: Upper/lower bounds in family $\mathcal{Q}_\eta$ are **attained** at $Q_\pm = P^{c\pm}$
- 1-parameter tilted models $dP^c \sim e^{cf_i} dP$ corresp. to QoIs f_i.



- **1.** Define the **UQ Indices** for any QoI f around model P as the two **worst case scenarios**:

$$I^+(f, P; \eta) = \max_{Q \in \mathcal{Q}_\eta} E_Q f \tag{2}$$

and

$$I^-(f, P; \eta) = \min_{Q \in \mathcal{Q}_\eta} E_Q f \tag{3}$$

where $\mathcal{Q}_\eta = \{Q : R(Q|P) \le \eta\}$

- **2. Tightness of bounds $\rightarrow$ Finite-dimensional representation of UQ Indices:**

$$I^+(f, P; \eta) = \min_{\lambda > 0} \left[\frac{1}{\lambda} \log \int e^{\lambda f(\zeta)} P(\zeta) d\zeta + \frac{\eta}{\lambda} \right] \tag{4}$$

- **3. If the uncertainty level η is small:**

$$I^+(f, P; \eta) = E_P f + \sqrt{2 Var_P(f)} \eta^{1/2} + \frac{1}{3} \frac{\kappa_3(f)}{Var_P(f)} \eta + \mathcal{O}(\eta^{3/2}) \tag{5}$$

Example[1]: UQ Indices for nonlinear response+"beyond" in Gibbs measures, Markov random fields, etc

$$\mathcal{Q}_\eta = \{Q : R(Q|P) \leq \eta\}, \quad \eta = \text{specified uncertainty level}$$

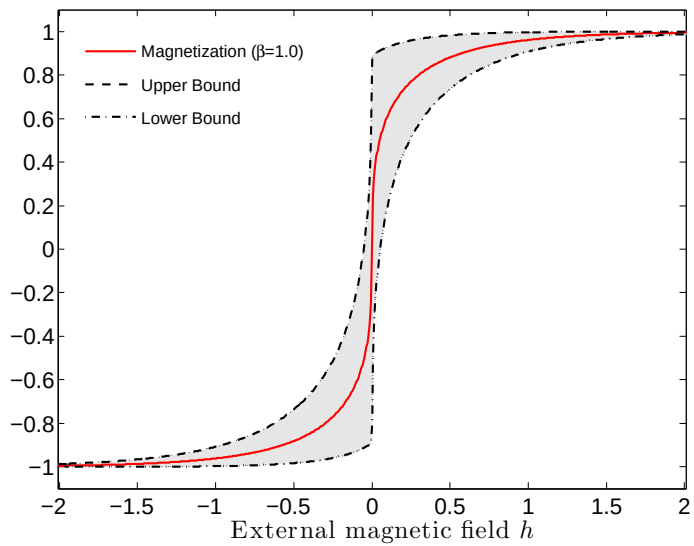
- $I^+(f, P; \eta) = \max_{Q \in \mathcal{Q}_\eta} E_Q f$ (top line)

and

- $I^-(f, P; \eta) = \min_{Q \in \mathcal{Q}_\eta} E_Q f$ (bottom)
- Red curve: $E_P f$=average lattice magnetization
- Bounds are tight[1].

Also:

- QoI depends on parameter $h \in [-2, 2]$. h: (robust) optimization parameter



[1] K., Rey-Bellet, Wang, **J. Comp. Phys.**, (2017)

How to compute/estimate $\Xi_\pm$?

- Accelerated **Monte Carlo** methods for the direct simulation of $\Xi_\pm$
  (see talks by Paul Dupuis and Petr Plechac)
- **Linearization**[1]: $\Xi_+(Q|P;f) = \sqrt{\mathrm{Var}_P[f]}\sqrt{2R(Q|P)} + \mathcal{O}(R(Q|P))$
- **Concentration Inequalities**[2]**:** Find $\Phi : \mathbb{R} \to (0,\infty]$ such that

$$E_P[e^{c(f-E_P[f])}] \leq \Phi(c) \quad \text{for all } c \in (-c_0, c_0)\,, \tag{6}$$

$$\text{Substitute in} \quad \Xi_+(Q|P;f) = \inf_{c>0}\left\{\frac{1}{c}\log E_P[e^{c(f-E_P[f])}] + \frac{1}{c}R(Q|P)\right\}$$

$$\text{to obtain} \quad U_+(Q|P;f) := \inf_{c>0}\left\{\frac{1}{c}\log \Phi(c) + \frac{1}{c}R(Q|P)\right\}$$

Typical examples of concentration inequalities:

1. Hoeffding, Bennett, sub-Gaussian
2. McDiarmid inequalities
3. For Markov Chains, Markov Random Fields, quantum systems, etc.

---

[1] P. Dupuis, M. K., Y. Pantazis, P. Plechac *SIAM/ASA J. of UQ* '16; H. Lam *Math. Oper. Res.* (2016)

[2] Gourgoulias, K. Rey-Bellet, Wang, *Arxiv*, (2017)

Sensitivity Analysis via Linearized Information Inequalities

- $P^\theta = P^\theta(X)$: parametrized family of distributions, $\theta \in \mathbb{R}^K$.
- $E_{P^\theta} f$: Quantity of Interest (QoI) where $f = f(X)$
- Perturbations in P^θ: unit vector v, $\epsilon << 1$

$$E_{P^{\theta+\epsilon v}} f - E_{P^\theta} f \leq \Xi_+(P^{\theta+\epsilon v}|P^\theta; f) \quad \mapsto \quad \text{(linearization)}$$
$$\leq \sqrt{\mathrm{Var}_P^\theta[f]}\sqrt{2R(P^{\theta+\epsilon v}|P^\theta)} + \mathcal{O}(R(P^{\theta+\epsilon v}|P^\theta))$$

$$R(P^{\theta+\epsilon v}|P^\theta) = \frac{\epsilon^2}{2} v^T \mathcal{I}_R(P^\theta) v \ + \mathcal{O}(\epsilon^3)\,.$$

$\mathcal{I}_R(P^\theta)$: Fisher Information Matrix, Hessian of KL. As $\epsilon \to 0$:

$$|\partial_v E_{P^\theta} f| \leq \sqrt{\mathrm{Var}_{P^\theta}(f)}\sqrt{v^T \mathcal{I}_R(P^\theta) v}\,.$$

1. Cramer-Rao-type bounds for QoIs $E_{P^\theta} f$
2. $\mathcal{I}_R(P^\theta)$: Information geometry, curvature at θ
3. Spectral analysis of $\mathcal{I}_R(P^\theta) \mapsto$ sensitivity analysis for the entire P^θ.

See discussion in Dupuis, K., Pantazis, Plechac *SIAM J. of UQ (2016)*

"Cramer-Rao" type sensitivity bounds for risk-sensitive functionals

1. Comparing QoIs which are **risk-sensitive** functionals, e.g. partition functions, CGFs, $\log P(A)$, etc.

$$\frac{1}{\gamma} log \int e^{\gamma g} dQ \quad \text{vs.} \quad \frac{1}{\gamma} log \int e^{\gamma g} dP$$

2. Sensitivity Index for parametrized family P^θ and likelihood ratio method (analogous to Glynn et al):

$$\nabla_\theta \log E_{P^\theta}[e^{\gamma g}] = E_{P^\theta_\gamma}[\nabla_\theta \log P^\theta] \quad \text{where} \quad dP^\theta_\gamma \propto e^{\gamma g} dP^\theta$$

3. Sensitivity Bounds:

$$v^T \nabla_\theta \log E_{P^\theta}[e^{\gamma g}] \leq \Xi_+(P^\theta_\gamma \| P^\theta; v^T \nabla_\theta \log P^\theta) := \inf_{c>0} \{ \frac{1}{c} \log E_{P^\theta} e^{c\gamma v^T \nabla_\theta \log P^\theta} + \frac{1}{c} R(P^\theta_\gamma \| P^\theta) \}$$

4. Compare to Cramer-Rao: $\gamma \to 0$, $\gamma^{-1} \log E_{P^\theta}[e^{\gamma g}] \to E_{P^\theta} g$

$$\frac{R(P^\theta_\gamma \| P^\theta)}{\gamma^2} \to \frac{Var_{P^\theta}(g)}{2}, \gamma^{-1} \log E_{P^\theta} e^{c\gamma v^T \nabla_\theta \log P^\theta} \to \text{Cov}(v^T \nabla_\theta \log P^\theta) = v^T \mathcal{I}_R(P^\theta) v$$

Sensitivity on Path Space - Scalability in $T \gg 1$

For ergodic QoIs and path measures $P_{0:T}$ in $[0, T]$:

$$f(\{X_T\}) = \frac{1}{T}\int_0^T g(X_t)dt$$

$$E_{P^{\theta+\epsilon v}_{0:T}} - E_{P^{\theta}_{0:T}} \leq \Xi_+(P^{\theta+\epsilon v}P_{0:T}|P^{\theta}_{0:T}; f) \approx \sqrt{T \cdot \mathrm{Var}_{P^{\theta}_{0:T}}[f]}\sqrt{\frac{2}{T} \cdot R(P^{\theta+\epsilon v}_{0:T}|P^{\theta}_{0:T})}$$

As $T \to \infty$:

$$T \cdot \mathrm{Var}_{P_{0:T}}[f] \to 2\int_0^\infty E_P[g(X_t)g(X_0)]dt \quad \textit{Integrated Autocorr. Time}$$

$$\frac{1}{T} \cdot R(Q_{0:T} \mid P_{0:T}) \to \mathcal{H}(Q \mid P) \quad \textit{Relative Entropy } \textbf{Rate}$$

$$\mathcal{H}(Q^\theta \mid Q^{\theta+\epsilon}) = \frac{\epsilon^2}{2} v^T \mathcal{I}_{\mathcal{H}}(Q^\theta) v + O(|\epsilon|^3) \quad \textit{Path-Fisher Info. Matrix (FIM)}$$

- How do we use all that? Are they computable?

[1] See statistical estimators for $\mathcal{H}$ in Pantazis, K., *J. Chem. Phys.* (2013)

1.Fast Sensitivity Screening for Reaction Networks

- Reaction network with:
  - N **species**, $\mathbf{S} = \{S_1, ..., S_N\}$
  - M **reactions**, $\mathbf{R} = \{R_1, ..., R_M\}$
- State: $\mathbf{X}(t) = [X_1(t), ..., X_N(t)]^T$:
  - $X_i(t)$: number of molecules of species S_i at time t
- Stoichiometry vector of j-th reaction: $\nu_{\mathbf{j}}$.
- Continuous time Markov process with countable state space:

$$\mathbb{P}\{\mathbf{X}(t+\Delta t) = x+\nu_{\mathbf{j}} \mid \mathbf{X}(t) = x\} = a_j^{\theta}(x)\Delta t + o(\Delta t),$$

- Propensities depend on parameters $\theta = (\theta_1, \ldots, \theta_K)$, e.g.

$$a_j^{\theta}(x) = \theta_j g_j(x), \quad j = 1, ..., M \quad \text{(mass action kinetics)}$$

$$a_j^{\theta}(x) = \frac{\theta_j x_A}{\theta_{j'} + x_A}, \quad \text{(Michaelis-Menten kinetics)}$$

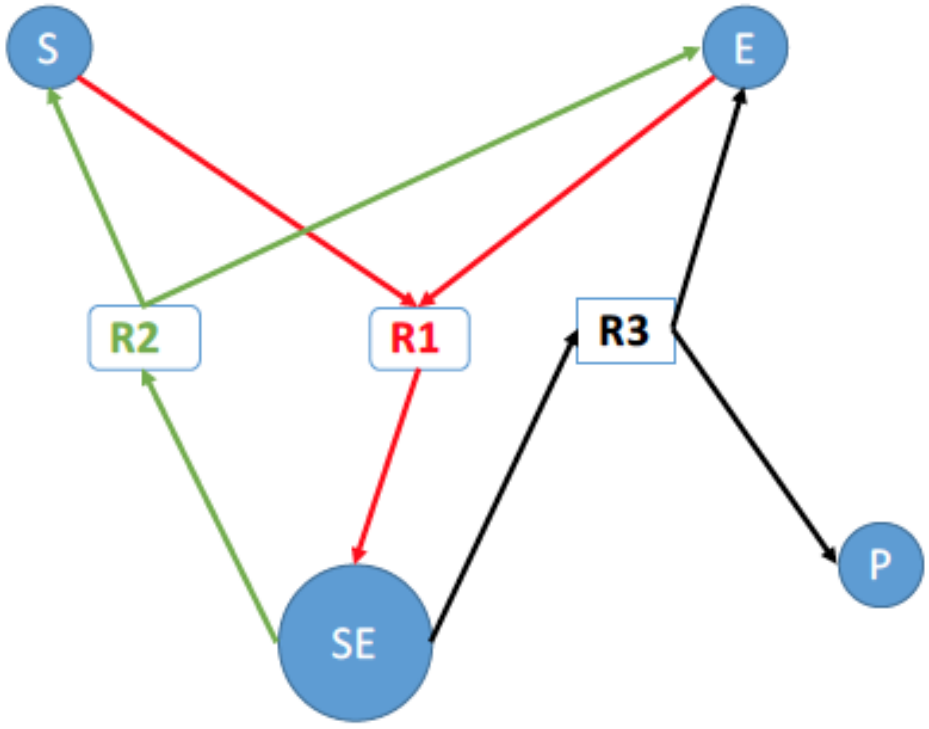
- Generator:

$$\mathcal{L}f(x) = \sum_{j=1}^{M} a_j^{\theta}(x)\left[f(x+\nu_{\mathbf{j}}) - f(x)\right],$$

A simple example

Graph of Chemical Reaction Network

Species: S, E, SE, P
Reactions: R1, R2, R3



Michaelis-Menten Reaction Network

$$\mathrm{R1}: \mathrm{S} + \mathrm{E} \rightarrow SE$$
$$\mathrm{R2}: \mathrm{SE} \rightarrow S + E$$
$$\mathrm{R3}: \mathrm{SE} \rightarrow P + E$$

Stoichiometry Matrix

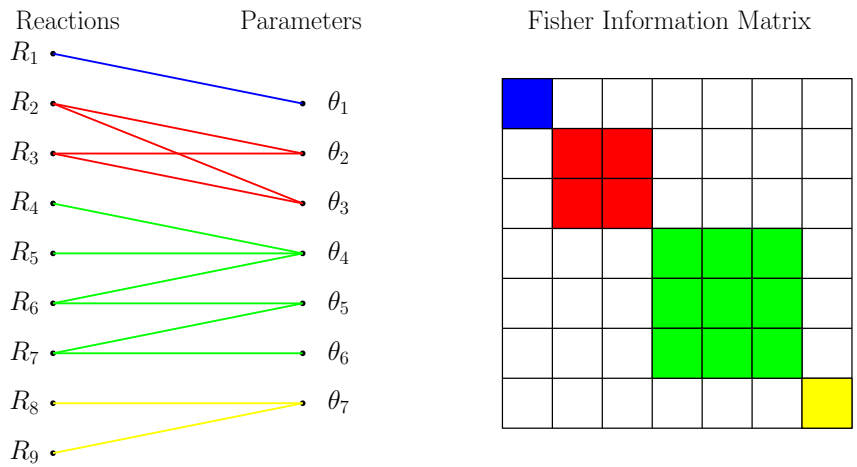
$$\nu = \begin{array}{c} \begin{matrix} \mathrm{R1} & \mathrm{R2} & \mathrm{R3} \end{matrix} \\ \begin{pmatrix} -1 & 1 & 0 \\ -1 & 1 & 1 \\ 1 & -1 & -1 \\ 0 & 0 & 1 \end{pmatrix} \end{array} \begin{matrix} \mathrm{S} \\ \mathrm{E} \\ \mathrm{SE} \\ \mathrm{P} \end{matrix}$$

Path FIM for networks: sparse structure

- **path FIM** of reaction networks: we typically have a block diagonal structure[1]

$$\mathcal{I}_H(P^\theta) = E_{\mu^\theta} \sum_{j=1}^{M} a_j^\theta(x_t) \nabla_\theta \log a_j^\theta(x_t) \nabla_\theta \log a_j^\theta(x_t)^T$$

μ^θ: steady state (unknown-needs to be sampled)



- Scalable computations - FIM **scales linearly** in the number of parameters.
- Contains key information:
  1. Graph structure
  2. Dynamics on the graph: reaction rates and their functional form

[1] Y. Pantazis, M. K. *J. Chem. Phys.* (2013); M. K., Y. Pantazis, D. Vlachos, *BMC Bioinformatics*, (2013)

How to use the UQ bounds for sensitivity screening:

- **1. Definition of UQ Indices:**

$$I^+(f, P; \eta) := \max_{R(Q||P) \leq \eta} E_Q f - E_P f$$

- **2. Representation of UQ Indices:**

$$I^+(f, P; \eta) = \inf_{c>0} \left[\frac{1}{c} \log \int e^{c\bar{f}(x)} P(dx) + \frac{\eta}{c} \right]$$

- **3. Local Sensitivity Index (non-parametric):**

$$I^+(f, P; \eta) = \max_{Q \in \mathcal{Q}_\eta} E_Q f - E_P f = \sqrt{2 Var_P(f)} \eta^{1/2} + \frac{1}{3} \frac{\kappa_3(f)}{Var_P(f)} \eta + \mathcal{O}(\eta^{3/2})$$

Non-parametric family "around" P (for some small $\eta > 0$.):

$$\mathcal{Q}_\eta = \{Q : R(Q||P) \leq \eta\}$$

Sensitivity of species vs. sensitivity of parameters

Comparing different QoIs/species f and g for fixed level of uncertainty η:

$$I^{\pm}(f, P; \eta) \approx \pm\sqrt{2 Var_P(f)}\eta^{1/2} \quad \text{vs.} \quad \pm\sqrt{2 Var_P(g)}\eta^{1/2} \approx I^{\pm}(g, P; \eta)\,.$$

To leading order, **comparison between variances/autocorrelations**:

$$Var_P(f) \quad \text{vs.} \quad Var_P(g)$$

Sensitivity of species vs. sensitivity of parameters

Comparing different QoIs/species f and g for fixed level of uncertainty η:

$$I^{\pm}(f, P; \eta) \approx \pm\sqrt{2Var_P(f)}\eta^{1/2} \quad \text{vs.} \quad \pm\sqrt{2Var_P(g)}\eta^{1/2} \approx I^{\pm}(g, P; \eta)\,.$$

To leading order, **comparison between variances/autocorrelations**:

$$Var_P(f) \quad \text{vs.} \quad Var_P(g)$$

Comparing different model perturbations η_1 and η_2 for fixed QoI f:

$$I^{\pm}(f, P; \eta_1) \approx \pm\sqrt{2Var_P(f)}\eta_1^{1/2} \quad \text{vs.} \quad \pm\sqrt{2Var_P(f)}\eta_2^{1/2} \approx I^{\pm}(f, P; \eta_2)\,.$$

Sensitivity of species vs. sensitivity of parameters

Comparing different QoIs/species f and g for fixed level of uncertainty η:

$$I^{\pm}(f, P; \eta) \approx \pm\sqrt{2Var_P(f)}\eta^{1/2} \quad \text{vs.} \quad \pm\sqrt{2Var_P(g)}\eta^{1/2} \approx I^{\pm}(g, P; \eta)\,.$$

To leading order, **comparison between variances/autocorrelations**:

$$Var_P(f) \quad \text{vs.} \quad Var_P(g)$$

Comparing different model perturbations η_1 and η_2 for fixed QoI f:

$$I^{\pm}(f, P; \eta_1) \approx \pm\sqrt{2Var_P(f)}\eta_1^{1/2} \quad \text{vs.} \quad \pm\sqrt{2Var_P(f)}\eta_2^{1/2} \approx I^{\pm}(f, P; \eta_2)\,.$$

- Parametric models and parameter space *unit vectors* $v_1, v_2 \in \mathbb{R}^k$ and $\epsilon \in \mathbb{R}$,

$$\eta_i = R(P^{\theta+\epsilon v_i}|P^{\theta}) = \frac{\epsilon^2}{2} v_i^T \mathcal{I}_{\mathcal{H}}(P^{\theta}) v_i \; + \mathcal{O}(\epsilon^3)\,, \quad i = 1, 2\,.$$

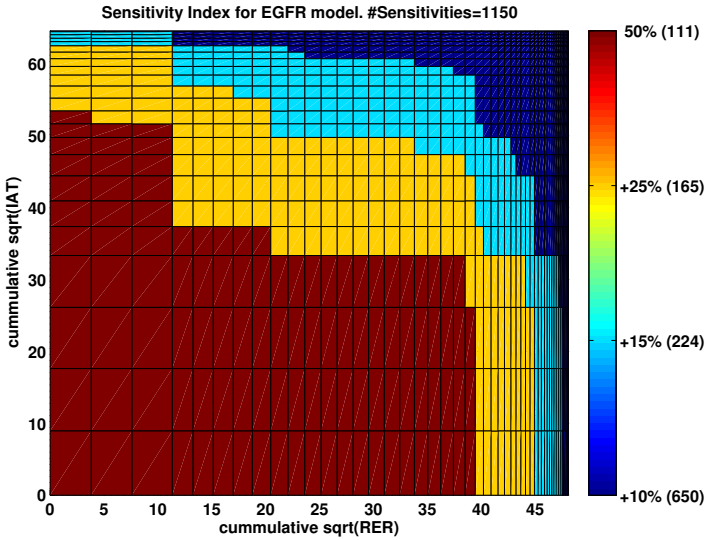
Thus

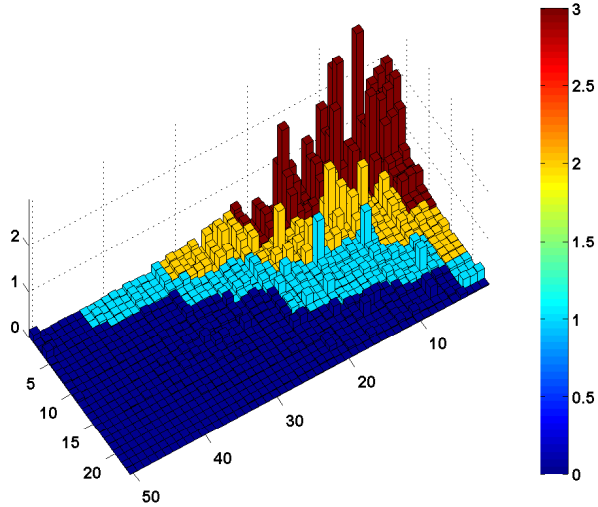
$$\sqrt{Var_{P^\theta}(f)}\sqrt{v_1^T \mathcal{I}_{\mathcal{H}}(P^{\theta}) v_1} \quad \text{vs.} \quad \sqrt{Var_{P^\theta}(f)}\sqrt{v_2^T \mathcal{I}_{\mathcal{H}}(P^{\theta}) v_2}$$

Example: Fast Sensitivity Screening - Information Metrics & QoIs

Biological network describing Epidermal Growth Factor Receptor. *[Kholodenko et.al., J. Biol. Chem., 1999]*

47 reactions, 23 species, 23 observables, 50 parameters, $23 \times 50 = 1150$ sensitivities





- Screening model sensitivities:
  $S_{F,v}(P^\theta) = \lim_{\varepsilon \to 0} \varepsilon^{-1} \left(\mathbb{E}_{P^{\theta+\varepsilon v}}[F] - \mathbb{E}_{P^\theta}[F] \right)$

$$|S_f(\theta_k)| \leq \sqrt{\mathrm{Var}_{\mu_P^\theta}(f)} \sqrt{\mathcal{I}_H(Q^\theta)_{k,k}}$$

- $\mathrm{Var}_{\mu_P^\theta}(f)$: Integrated Autocorrelation
- $\mathcal{I}_H(Q^\theta)$: Path Fisher Info. Matrix

[1] Arampatzis, Pantazis, K. *PLOS 1*, (2015); Dupuis, K., Pantazis, Plechac *SIAM UQ*, (2016)

Contrast to unsorted parameters/observables:

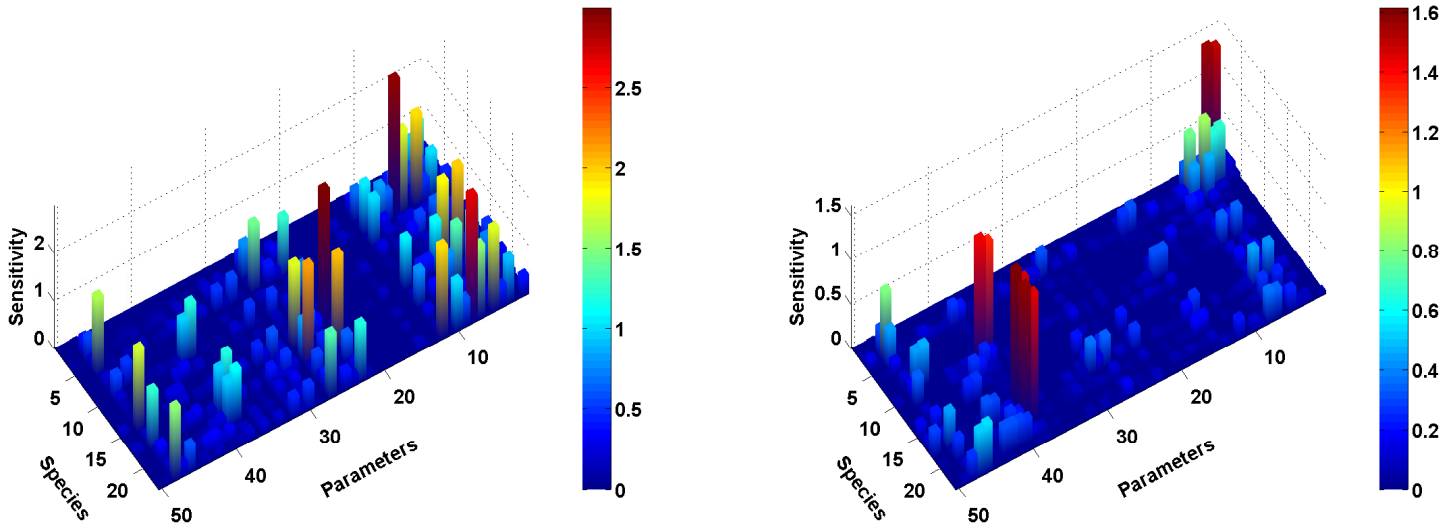


Figure 1: Unsorted Sensitivity Analysis in (a) time interval [0, 50], (b) steady state

Fast Screening: 2-stage strategy for Sensitivity Analysis:

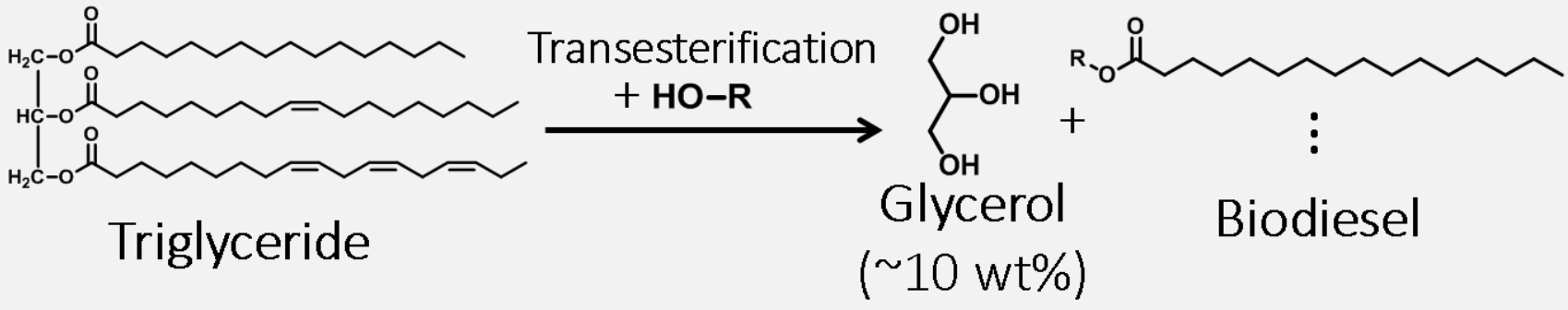
1. Screens out insensitive parameters using the UQ Information Inequalities:

$$K \mapsto K' \quad \text{parameters}$$

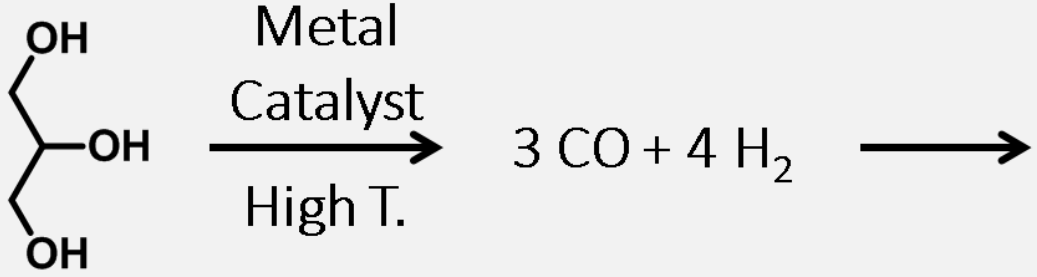
2. Accurately compute sensitivities for remaining K' parameters using exact methods: adjoint equations, likelihood ratio, stochastic coupling,...
3. In the next example: $K \approx 8,000$, $K' = 6$

A complex system: Glycerol reforming for H_2 production

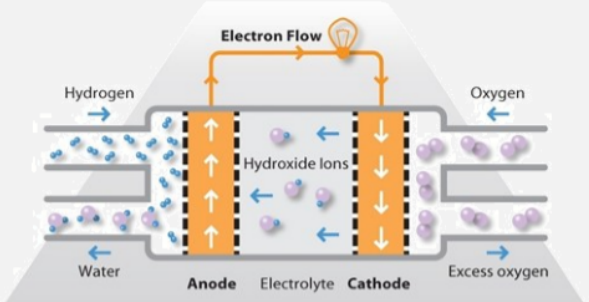
Glycerol Byproduct from Biodiesel Production



Catalytic Reforming



H_2/O_2 Fuel cell

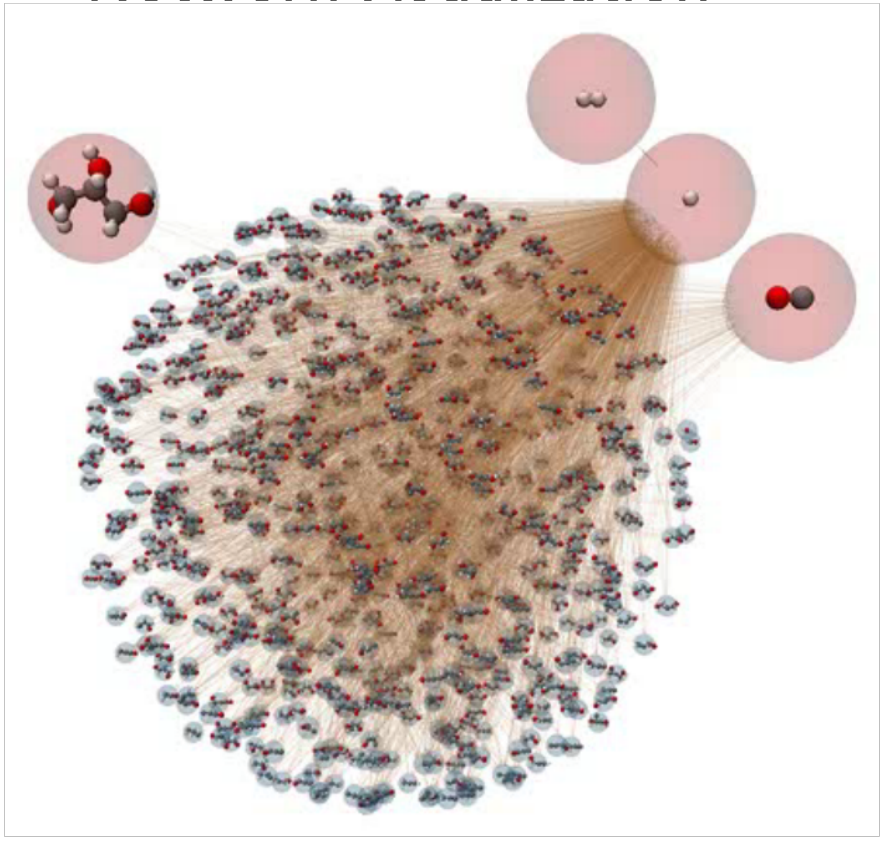


www.fuelcelltoday.com/technologies/afc

Yang *et al*. *Biotechnol. Biofuels.* **2012**, *5*, 13

Glycerol reaction network

Network Visualization



Microkinetic Model

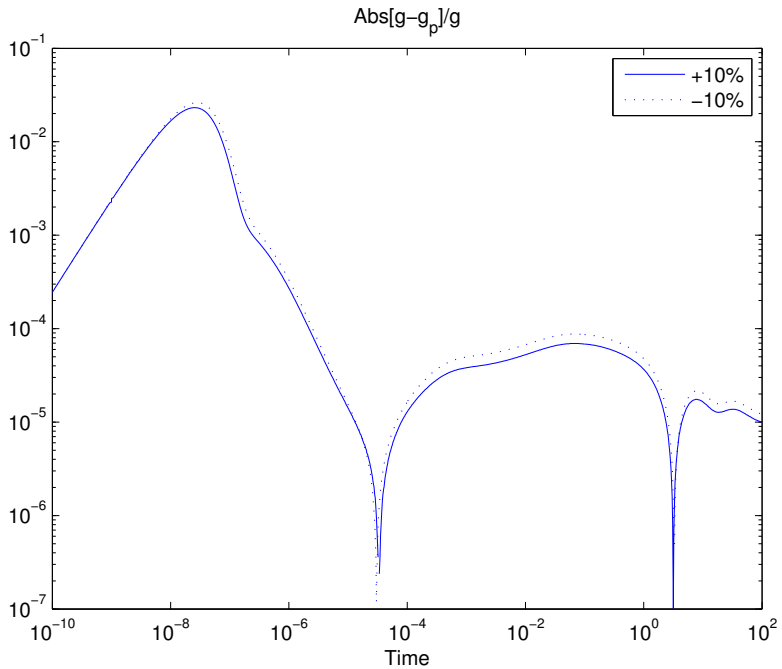
- 780 Species (780 ODEs)
- 3898 Reversible Reactions (~8000 parameters)
- Parameterized using fast screening DFT-based methods
- Running time: 139 sec. (Fortran-based ODE solver)
- Sensitivity analysis: ~6 days
- Sparse system

- Sensitivity analysis and UQ impractical realistic chemistries

Fast Sensitivity Screening: $\sim 10^4$ Parameters

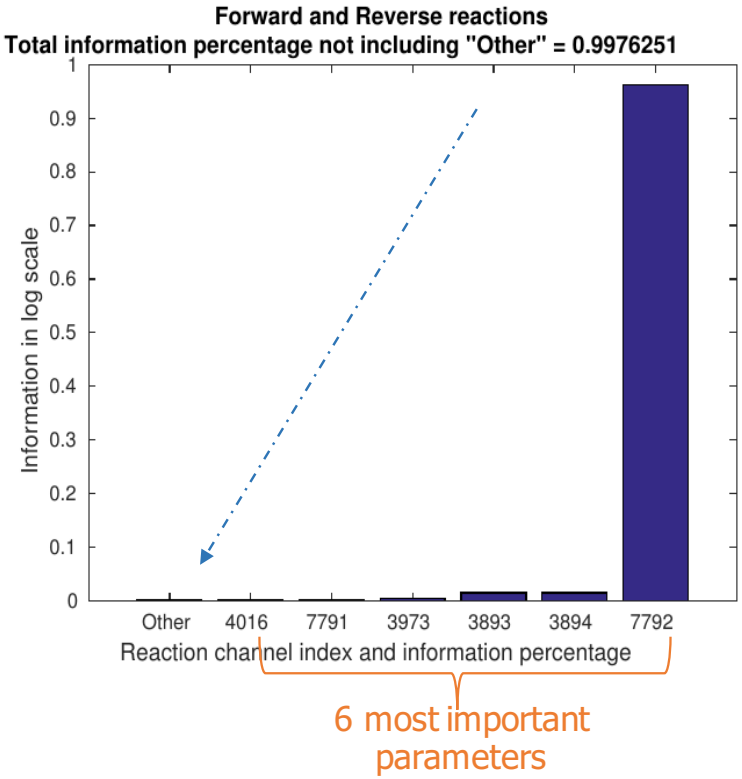
Reforming of Glycerol

Validation of Sensitivity Findings



Relative change in the observable (Glycerol concentration) by perturbing all but the most significant reaction channels (parameters) by ±10%.

Sensitivity Screening using path FIM



- Out of 7796 reaction channels, we discover 6 which accumulate more than 99% of the information (in the path-FIM).
- The SA is performed in a non-intrusive way with any existing integrator/simulator as a black-box.

Based on the reduction in the # of reaction rates, the computational acceleration is $\sim \times$ 1300.

[1] ongoing work by P. Vilanova, K., D. Vlachos

2. UQ for Bayesian (reaction) networks

- Earlier: impact of model and parameter perturbations on complex, high dimensional systems; parameter vector θ viewed as a constant that is either known or that should be inferred from data with great certainty.
- Here: consider *uncertainty* and *correlations* in the parameter vector θ in the sensitivity analysis.
- Ensembles of models (Bayesian network formulation):

$$P(x|\theta)\pi(\theta)$$

  - $P(x|\theta)$: Forward model, i.e. pdf of state $X = x$ for fixed parameter θ (reaction network, KMC, etc). For example, for ODEs:

$$P(x|\theta) = \delta(x - Y(t;\theta))\,, \quad \text{where} \quad \dot{Y} = f(Y;\theta)\,, \quad Y(0) = Y_0 \in \mathbb{R}^N\,, \theta \in \mathbb{R}^K\,.$$

  - $\theta \in \Theta$: model parameters, indexing of different models, etc.
  - $\pi(\theta)$: PDF of θ; uncertainty in knowledge of θ.
  - Correlations between the parameter sets $\theta = (\theta_1, \theta_2)$:

$$\pi(\theta_1, \theta_2) = \pi(\theta_2|\theta_1)\pi(\theta_1)\,, \qquad \pi(\theta_1)\text{: marginal}$$

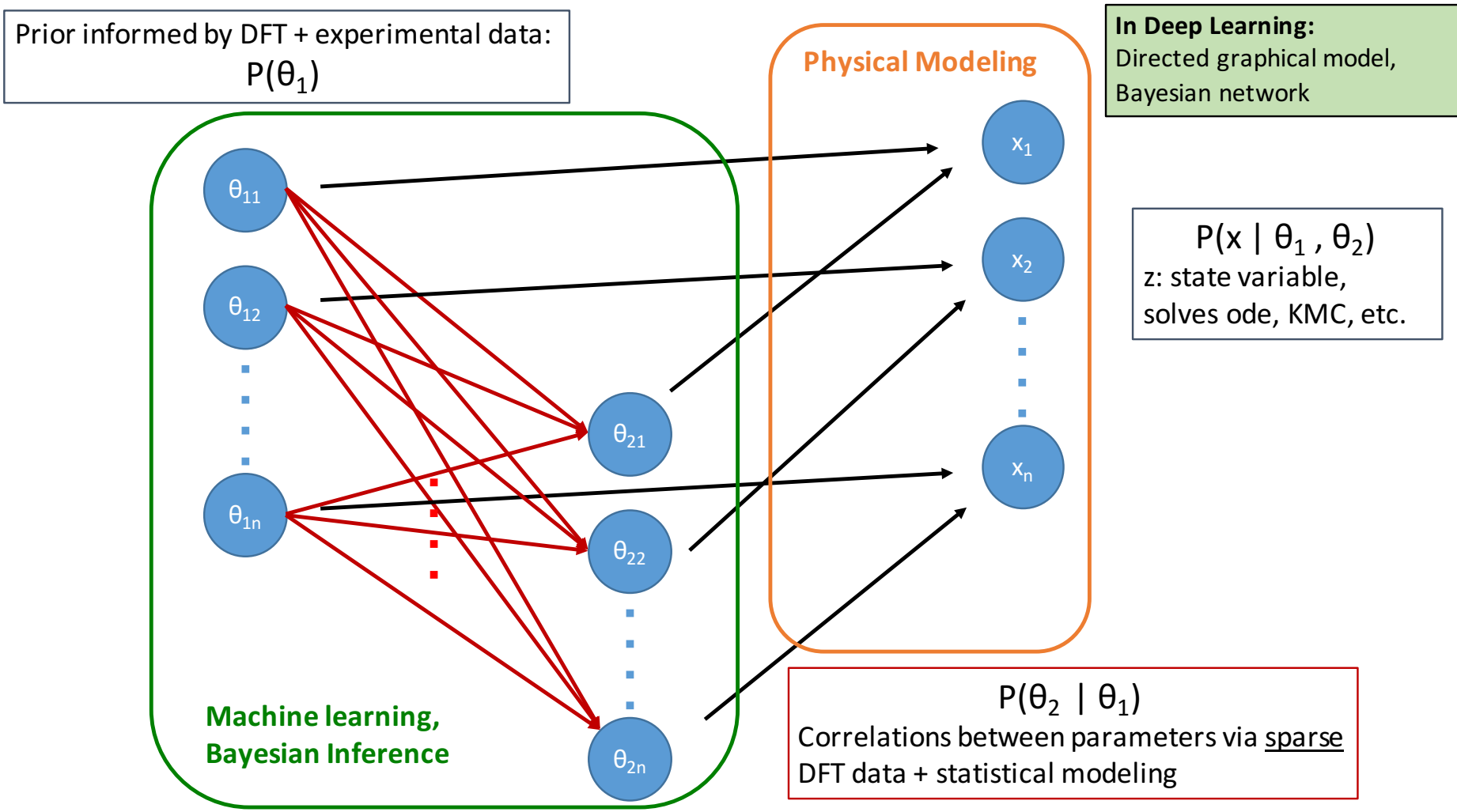
- **Model-form uncertainty:**

$$P(x|\theta)\pi(\theta) \quad \text{vs.} \quad P'(x|\theta)\pi'(\theta)$$

and respective predictions on QoIs

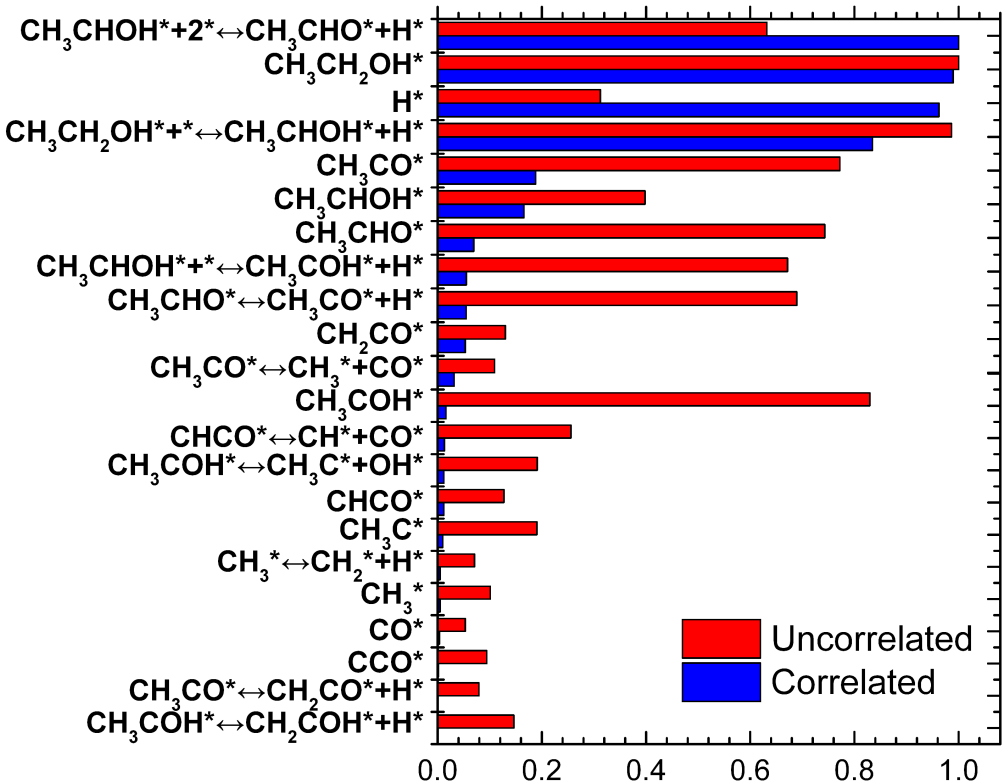
$$f = f(X)\,, \quad X \sim P(x), P'(x)$$

A structured probabilistic representation



Correlations Strongly Impact Predictions

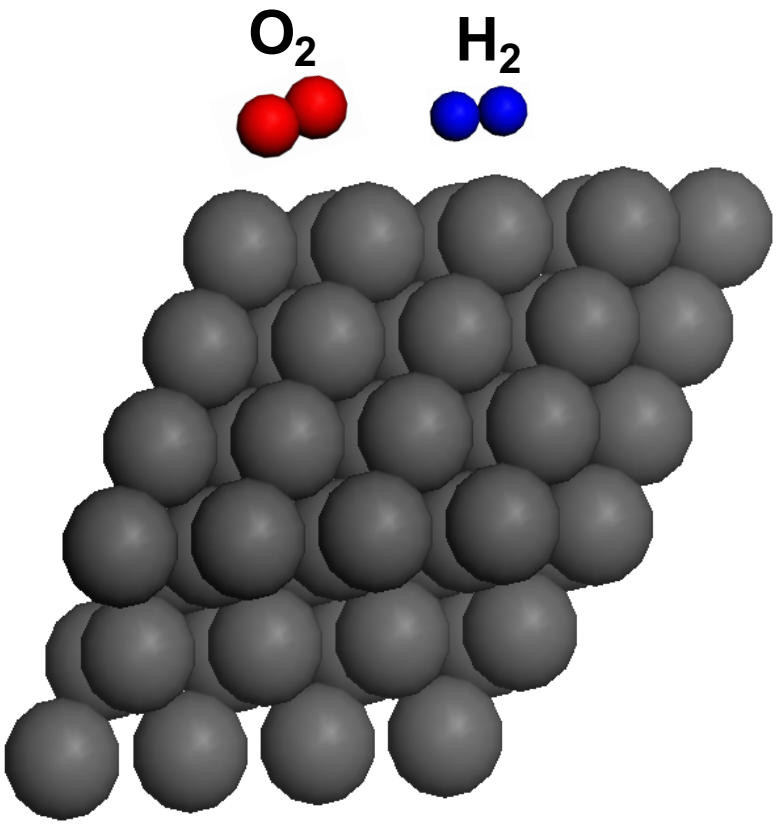
- Apply the methodology to a complex reaction network of *Ethanol Steam Reforming* on a Pt/Al_2O_3 catalyst with 67 species and 160 reactions.
- Correlations reduce significantly the number of important parameters
- Sensitivity of key parameters differ when correlations are included



[1] Sutton, Guo, Katsoulakis, and Vlachos, *Nature Chemistry*, (2016)

Example: Langmuir bimolecular adsorption model

Prototype Physical Model



Adsorption Desorption

$$\frac{d\hat{\theta}_{H^*}}{dt} = k_{H_2}^{ads} P_{H_2} (1 - \hat{\theta}_{H^*} - \hat{\theta}_{O^*})^2 - k_{H_2}^{des} \hat{\theta}_{H^*}^2$$

$$\frac{d\hat{\theta}_{O^*}}{dt} = k_{O_2}^{ads} P_{O_2} (1 - \hat{\theta}_{O^*} - \hat{\theta}_{O^*})^2 - k_{O_2}^{des} \hat{\theta}_{O^*}^2$$

$$\hat{\theta}_{H^*} = \frac{(K_{H_2} P_{H_2})^{\frac{1}{2}}}{1 + (K_{H_2} P_{H_2})^{\frac{1}{2}} + (K_{O_2} P_{O_2})^{\frac{1}{2}}}$$

$$\hat{\theta}_{O^*} = \frac{(K_{O_2} P_{O_2})^{\frac{1}{2}}}{1 + (K_{H_2} P_{H_2})^{\frac{1}{2}} + (K_{O_2} P_{O_2})^{\frac{1}{2}}}$$

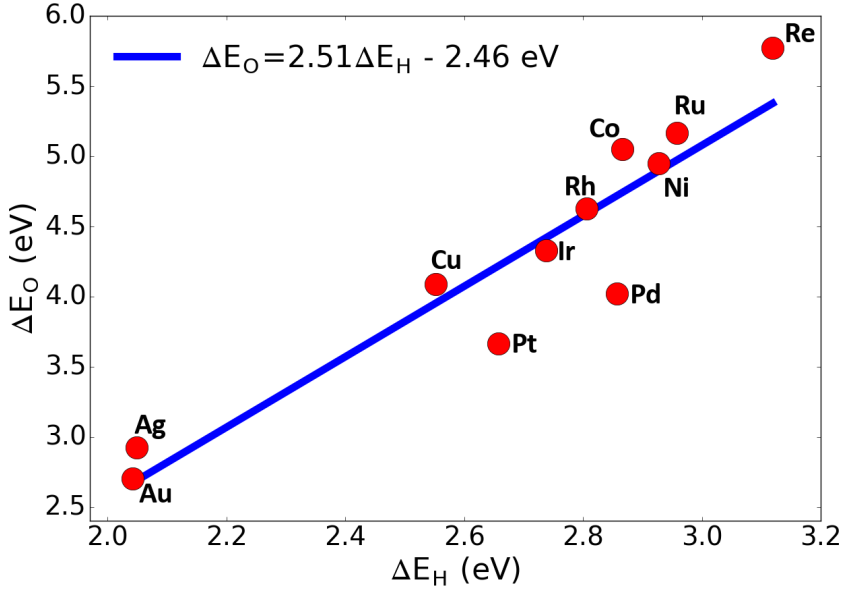
$K = k^{ads}/k^{des} = f(\Delta E_H, \Delta E_O)$
determined by DFT

[1] Feng, Lansford, Mironenko, Pourkargar, Vlachos, M. K., submitted (2017)

Prior and Correlations

$$p(\Delta E_H, \Delta E_O) = p(\Delta E_O | \Delta E_H) p(\Delta E_H)$$

correlation × prior



Prior distribution for ΔE_H

$$\mathbb{E}_H \sim Gamma(a_H, b_H)$$

$$p(\Delta E_H) = \frac{1}{b_H^{a_H}\Gamma(a_H)} \Delta E_H^{a_H - 1} \exp(-\frac{\Delta E_H}{b_H})$$

where $a_H = x_H^2/(x_H - y_H)^2$ and $b_H = (x_H - y_H)^2/x_H$

Table 1: Experimental and DFT-calculated enthalpies of adsorption for atomic oxygen and hydrogen on Pt(111).

Adsorbate	Experimental enthalpy (eV)	DFT computed enthalpy (eV)
O	3.71 ±0.07	3.68
H	2.63	2.69

Probabilistic model(s) for parameter correlations

- Data-based stochastic model for correlations between ΔE_H and ΔE_O :

$$\Delta E_O = g(\Delta E_H) = a\Delta E_H + b + \omega, \quad \omega \sim \Omega$$

  - $\Omega = 0$ (deterministic):

$$p(\Delta E_O \,|\, \Delta E_H) = \delta(\mathrm{a}\ \Delta E_H\ + \mathrm{b}\ -\ \Delta E_O)$$

  - Ω is a parametric family of distributions:

$$p(\Delta E_O \,|\, \Delta E_H) = \mathrm{p}_{\omega+1}(\Delta E_O - \mathrm{a}\ \Delta E_H\ -\mathrm{b}\ +1)$$

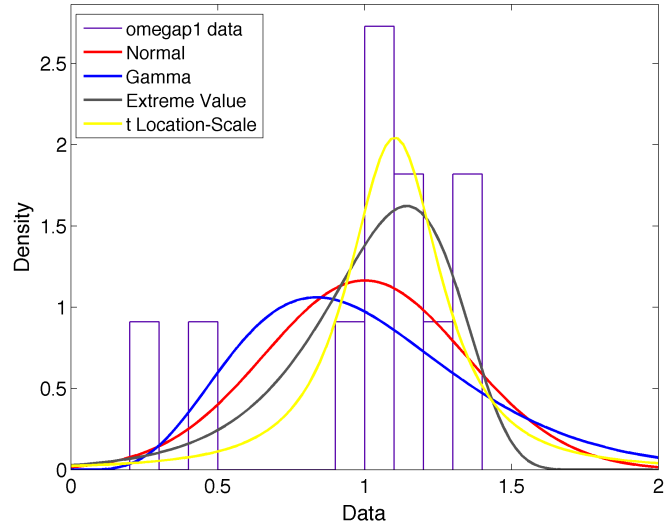
  - Ω is non-parametric models:
    - Bootstrapping from the histogram:

$$\mathrm{p_n}(\omega) = \sum_{\mathrm{i}=1}^{\mathrm{n}} \frac{\nu_i}{\mathrm{nh}} I(\omega \in B_i)$$

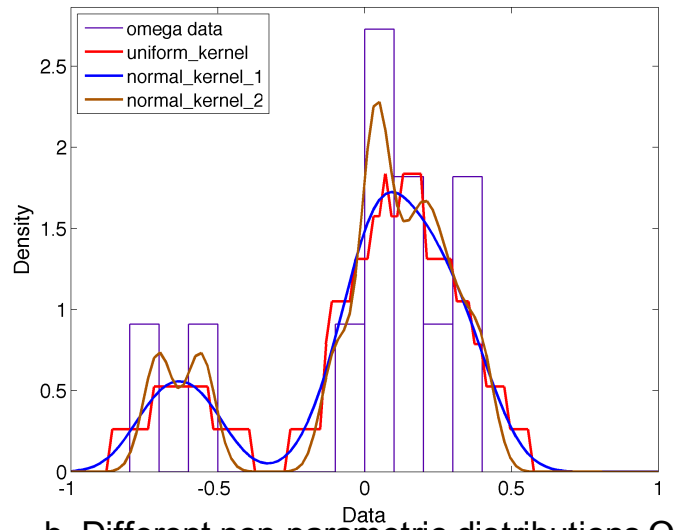
    - Kernel density estimators

$$\mathrm{p}(\omega) = \frac{1}{m} \sum_{\mathrm{i}=1}^{m} \frac{1}{\mathrm{h}} K\left(\frac{\omega - X_i}{h}\right)$$

(where K could be the density of Uniform or Normal distribution)



a. Different parametric distributions Ω



b. Different non-parametric distributions Ω

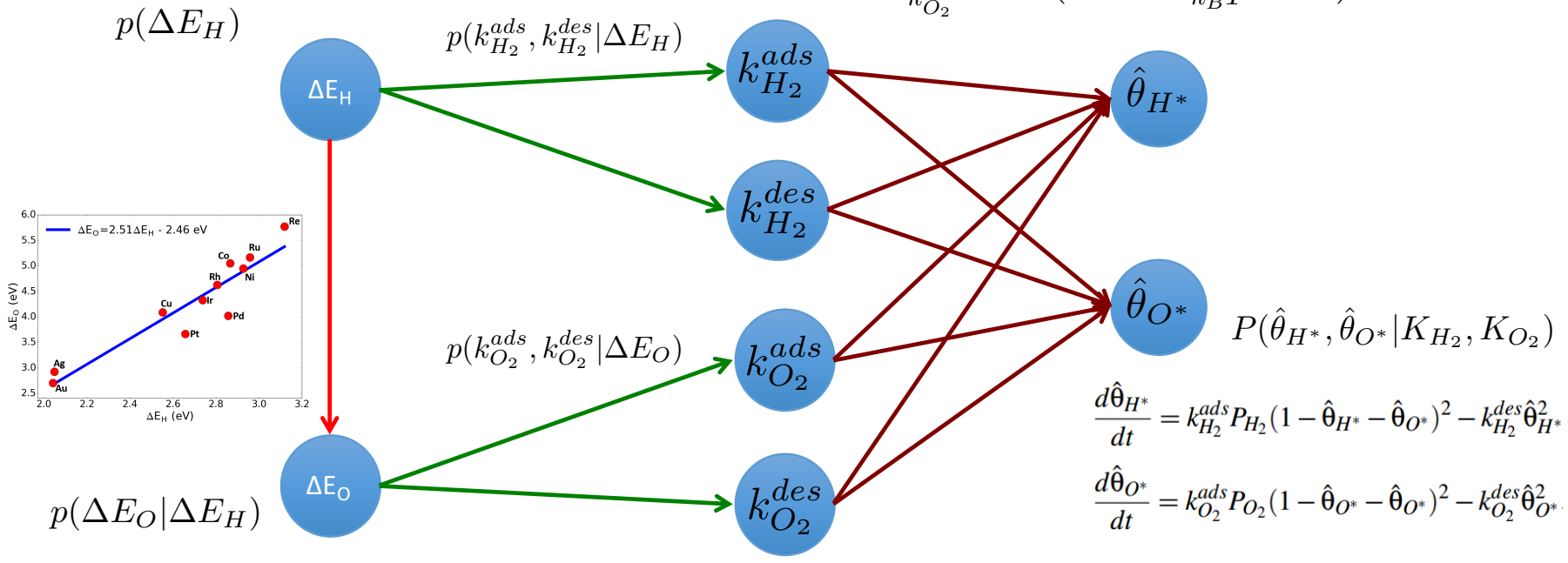
Completed Probabilistic Graphical Model formulation

Table 1: Experimental and DFT-calculated enthalpies of adsorption for atomic oxygen and hydrogen on Pt(111).

Adsorbate	Experimental enthalpy (eV)	DFT computed enthalpy (eV)
O	3.71 ±0.07	3.68
H	2.63	2.69

$$K_{H_2} = \frac{k^{ads}_{H_2}}{k^{des}_{H_2}} = \exp\Big(-\frac{\Delta G_{H_2 \to 2H^*}(\Delta E_H)}{k_B T}\Big) P_o^{-1}$$

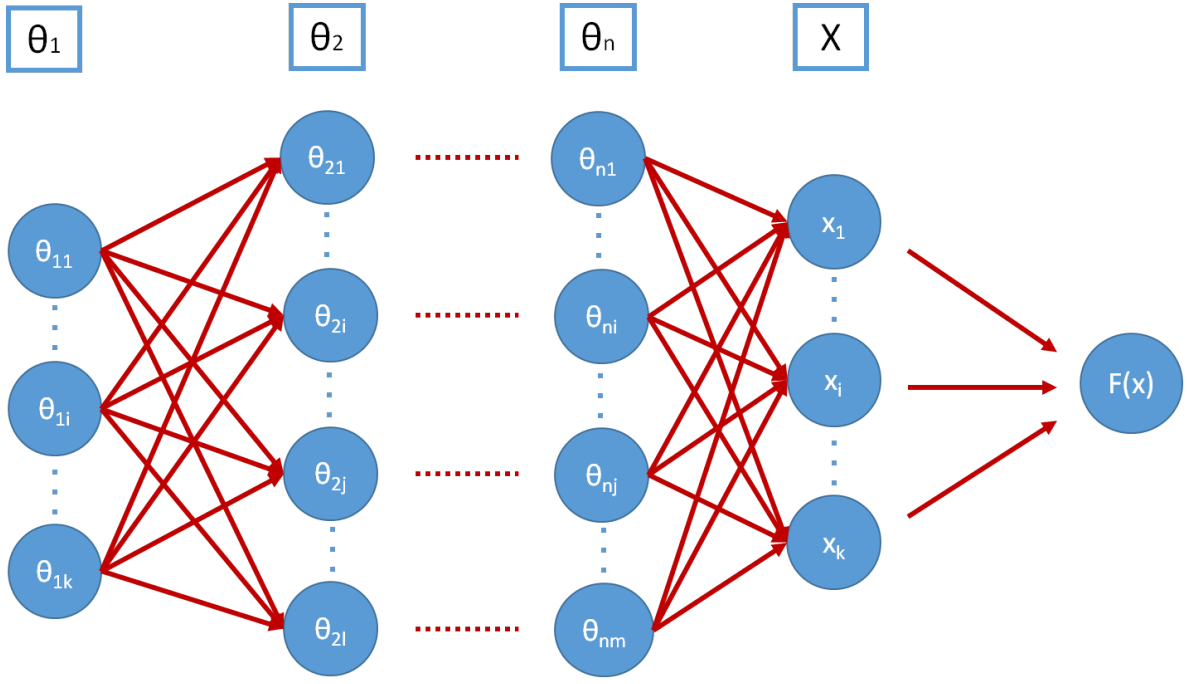
$$K_{O_2} = \frac{k^{ads}_{O_2}}{k^{des}_{O_2}} = \exp\Big(-\frac{\Delta G_{H_2 \to 2H^*}(\Delta E_O)}{k_B T}\Big) P_o^{-1}$$



$$\frac{d\hat{\theta}_{H^*}}{dt} = k^{ads}_{H_2} P_{H_2}(1 - \hat{\theta}_{H^*} - \hat{\theta}_{O^*})^2 - k^{des}_{H_2}\hat{\theta}^2_{H^*}$$

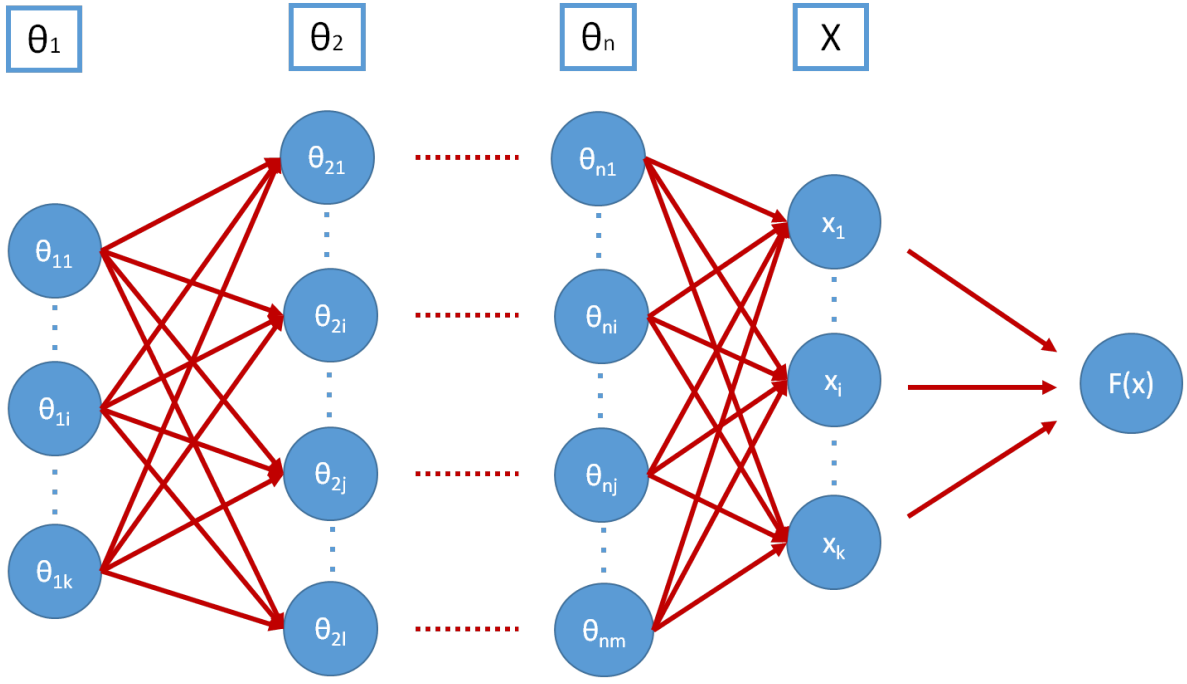
$$\frac{d\hat{\theta}_{O^*}}{dt} = k^{ads}_{O_2} P_{O_2}(1 - \hat{\theta}_{O^*} - \hat{\theta}_{O^*})^2 - k^{des}_{O_2}\hat{\theta}^2_{O^*}$$

$$P(\hat{\theta}_{H^*}, \hat{\theta}_{O^*}|k^{ads}_{H_2}, k^{ads}_{O_2}, k^{des}_{H_2}, k^{des}_{O_2}) \prod_{I=H,O} p(k^{ads}_{I_2}, k^{des}_{I_2}|\Delta E_I) p(\Delta E_O|\Delta E_H) p(\Delta E_H)$$



$$P(dx, d\theta) = \prod_{j=1}^{k} P(dx_j|\Gamma(x_j)) \Big[\prod_{i=2}^{n} \prod_{j=1}^{k_i} \pi_i(d\theta_{i,j}|\Gamma(\theta_{i,j})) \Big] \prod_{j=1}^{k_1} \pi_1(d\theta_{1,j})$$

$\Gamma(z)$: parents of node z n: number of layers k_i: number of elements in layer



$$P(dx, d\theta) = \prod_{j=1}^{k} P(dx_j|\Gamma(x_j)) \Big[\prod_{i=2}^{n} \prod_{j=1}^{k_i} \pi_i(d\theta_{i,j}|\Gamma(\theta_{i,j})) \Big] \prod_{j=1}^{k_1} \pi_1(d\theta_{1,j})$$

$\Gamma(z)$: parents of node z n: number of layers k_i: number of elements in layer

Chain rule for relative entropy: $P = p(x_2|x_1)p(x_1)$, $Q = q(x_2|x_1)q(x_1)$

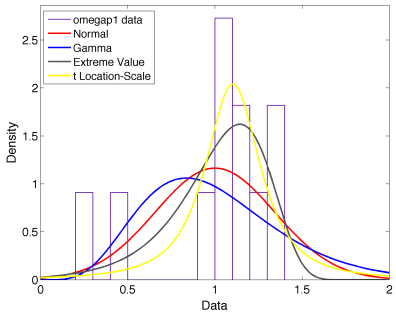
$$R(P|Q)) = E_P R(p(\cdot|x_1)|q(\cdot|x_1)) + R(p(\cdot)|q(\cdot))$$

Conditional relative entropy:

$$E_P R(p(\cdot|x_1)|q(\cdot|x_1)) := \int \Big(\int \log \frac{p(x_2|x_1)}{q(x_2|x_1)} p(x_2|x_1) dx_2 \Big) p(x_1) dx_1 \qquad (7)$$

UQ on the Graphical Model

$$P(X|\theta)\pi_2(\theta_2|\theta_1)\pi_1(\theta_1) \text{ vs } P(X|\theta)\pi_2'(\theta_2|\theta_1)\pi_1'(\theta_1)$$



Uncertainty in Correlations models

Uncertainty in DFT, experiments

Hybrid [1] UQ information inequalities for ***QoIs*** *F :*

$$\mathbb{E}_{\pi'}[F] - \mathbb{E}_{\pi}[F] \leq \inf_{c>0}\left[\frac{1}{c}\log\int e^{c\bar{F}(\theta)}\pi(d\theta) + \frac{1}{c}R(\pi'||\pi)\right]$$

$$R(\pi'(\theta_1,\theta_2)||\pi(\theta_1,\theta_2)) = \eta_1 + \eta_2$$

$$\begin{aligned}\eta_2 &:= R(\pi_2'(\theta_2|\theta_1)||\pi_2(\theta_2|\theta_1)) \\ &= \int R(\pi_2'(\cdot|\theta_1)||\pi_2(\cdot|\theta_1))\pi_1(d\theta_1) \\ &= \int\int \log\frac{\pi_2'(\theta_2|\theta_1)}{\pi_2(\theta_2|\theta_1)}\pi_2'(d\theta_2|\theta_1)\pi_1(d\theta_1)\end{aligned}$$

$$\begin{aligned}\eta_1 &:= R(\pi_1'(\theta_1)||\pi_1(\theta_1)) \\ &= \int \log\frac{\pi_1'(\theta_1)}{\pi_1(\theta_1)}\pi_1'(d\theta_1)\end{aligned}$$

Conditional Relative Entropy

Is the "suitable" quantity for UQ on PGMs: bounds for QoIs+accounts for correlations

[1] *Chowdhary and Dupuis, ESAIM (2013)*

UQ on the Graphical Model - Implementation

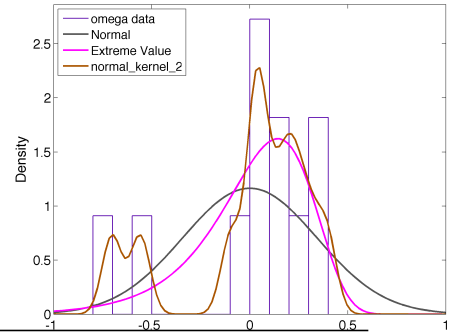
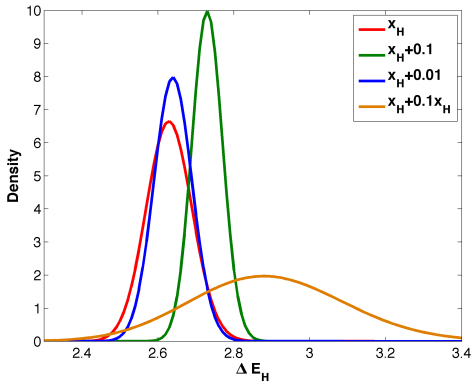


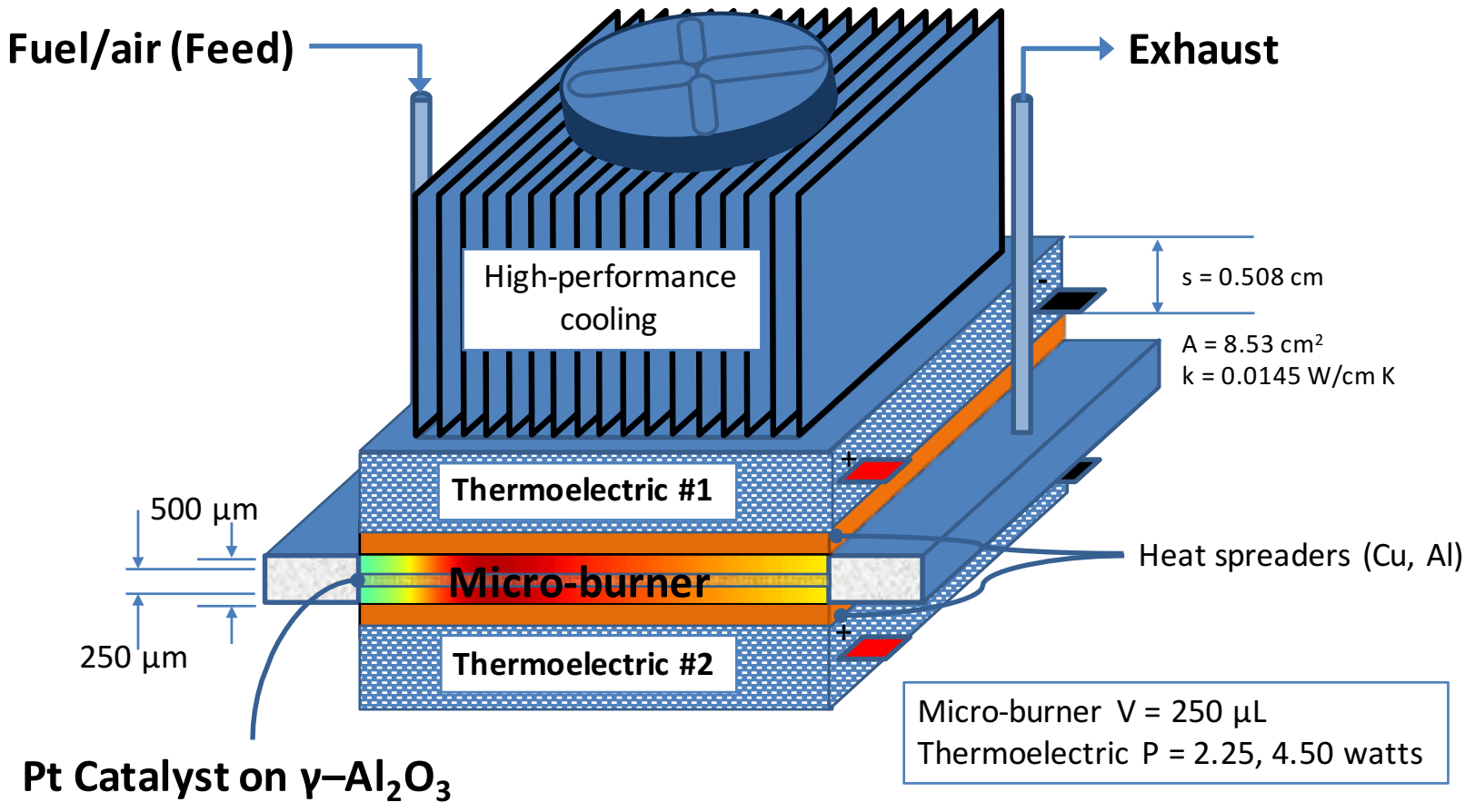
Table 1: $R(\pi'(\theta_1,\theta_2)||\pi(\theta_1,\theta_2))$ by different models

$\pi_2'(\theta_2\|\theta_1)\times\pi_1'(\theta_1)$	$\pi_2(\theta_2\|\theta_1)\times\pi_1(\theta_1)$	η_1	η_2	UQ index $I(f,\pi;\eta_1,\eta_2)$
$p(\Delta E_O)p(\Delta E_H)$	$p_n(\Delta E_O\|\Delta E_H)p(\Delta E_H)$	0	2.856	0.222
$p_{ev}(\Delta E_O\|\Delta E_H)p(\Delta E_H)$n	$p_n(\Delta E_O\|\Delta E_H)p(\Delta E_H)$	0	0.115	0.0132
$p_h(\Delta E_O\|\Delta E_H)p(\Delta E_H)$	$p_n(\Delta E_O\|\Delta E_H)p(\Delta E_H)$	0	17.751	0.973
$p_{n_{0.05}}(\Delta E_O\|\Delta E_H)p(\Delta E_H)$	$p_n(\Delta E_O\|\Delta E_H)p(\Delta E_H)$	0	17.481	0.970
$p_n(\Delta E_O\|\Delta E_H)p'_{0.01}(\Delta E_H)$	$p_n(\Delta E_O\|\Delta E_H)p(\Delta E_H)$	0.0446	0	0.00595
$p_n(\Delta E_O\|\Delta E_H)p'_{0.1}(\Delta E_H)$	$p_n(\Delta E_O\|\Delta E_H)p(\Delta E_H)$	1.504	0	0.125
$p_n(\Delta E_O\|\Delta E_H)p'_{10\%}(\Delta E_H)$	$p_n(\Delta E_O\|\Delta E_H)p(\Delta E_H)$	12.12	0	0.0141



3. Design under Uncertainty

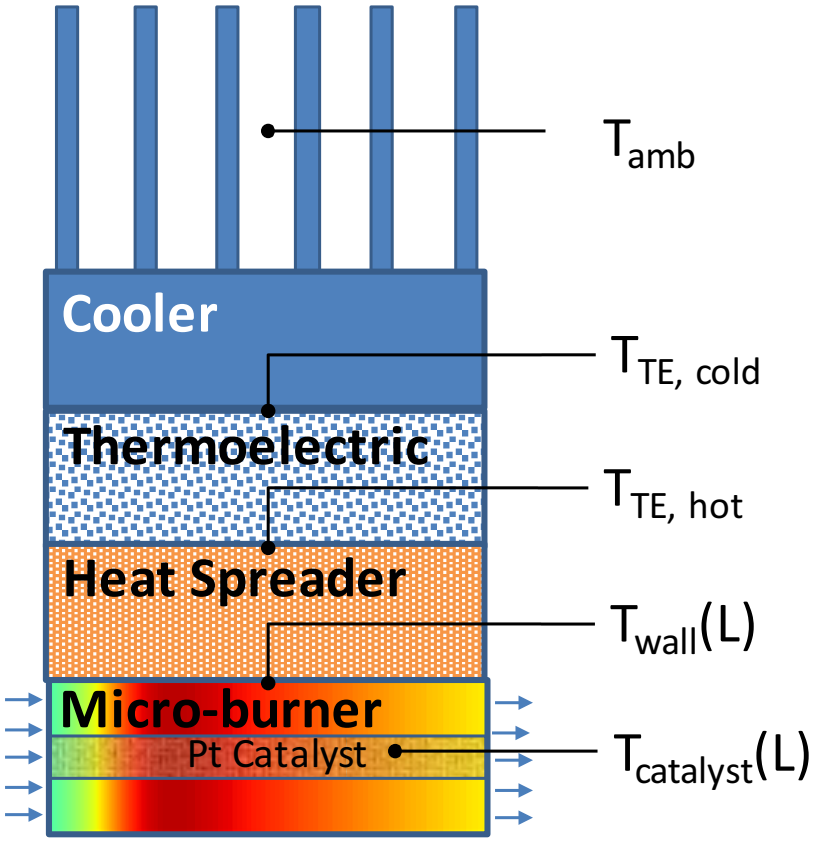
Schematic of Power Generation Device: Fuel to Electricity



Paul Dupuis, M. K, Ivan Lee, Dion Vlachos

Funded by DARPA

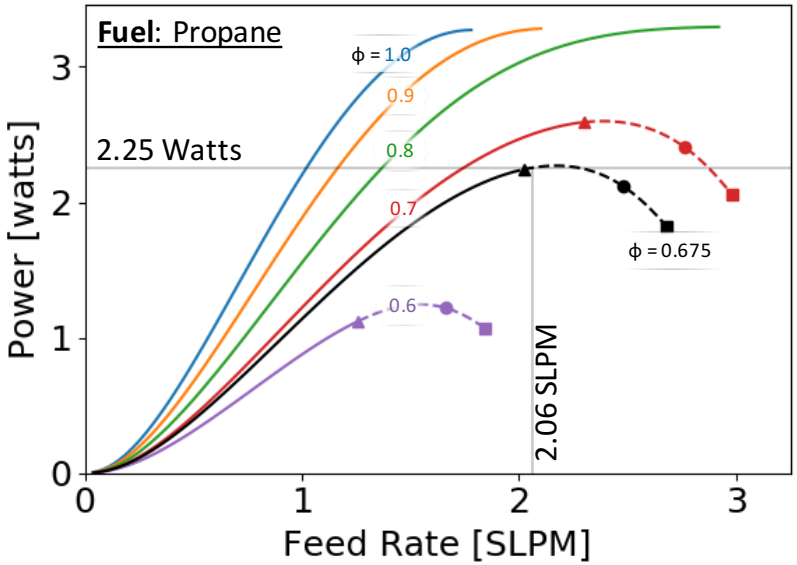
Governing equations at steady state: deterministic/stochastic system

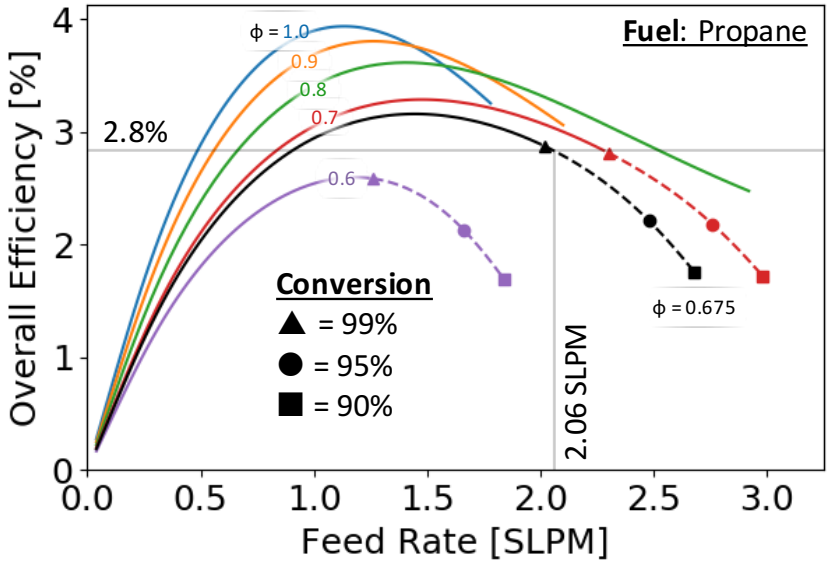


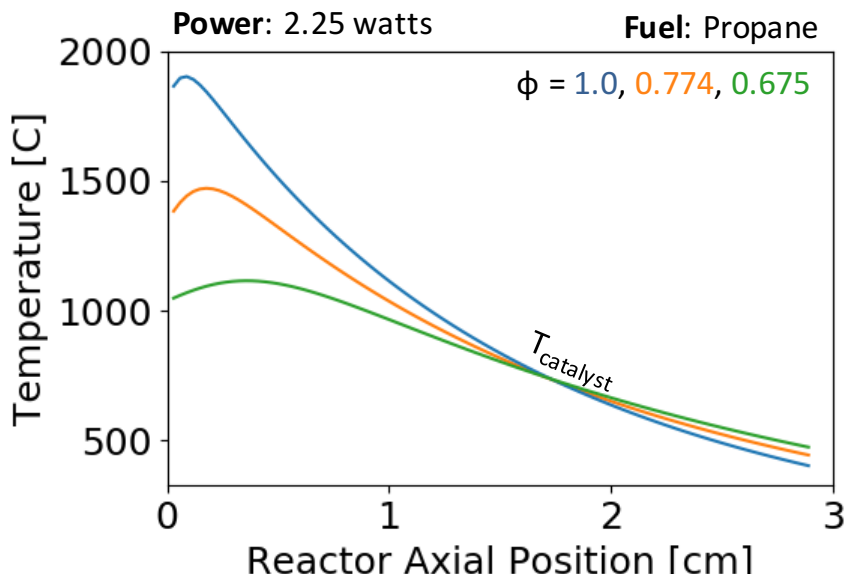
- Energy Balance
  (steady state heat equation)
- Materials Balance
- Surface Chemistry Kinetics
  (Chemical Reaction Network)

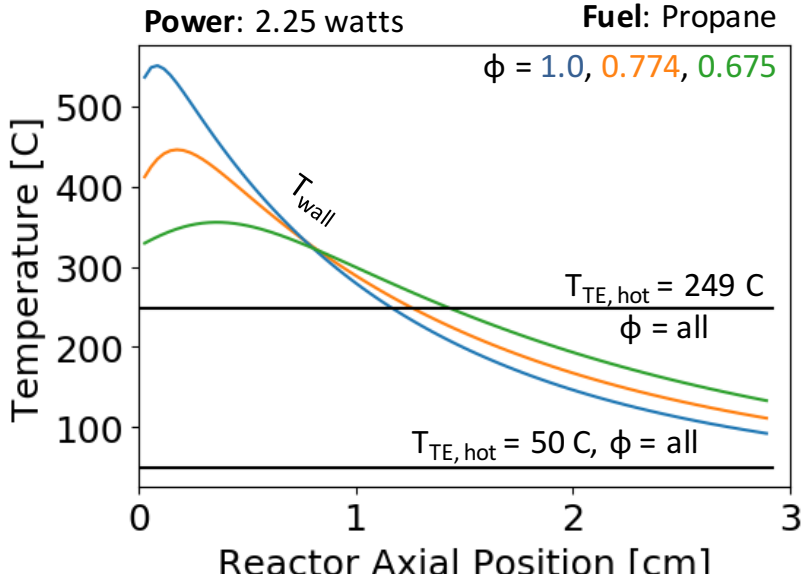
An early test for device performance:

deterministic PDE+ mean field, unoptimized









Variables

- **Spatially distributed state variables**: obtained by solving energy/mass balance and chemical kinetics:

$$z = z(L) = \left(T(L), X_{\text{fuel}}(L), X_{\text{surface}}(L)\right), \quad 0 \leq L \leq \text{Length} = L_0$$

- **Optimization parameters**

$$c = (F_{\text{fuel}}, \phi, L_0 = \text{Length})$$

- Other parameters (or mechanisms) **subject to uncertainty**:

$$\theta = (\text{reaction rates, fluctuations in flow, missing mechanisms})$$

1. θ can be a *random variable* with (partly) known distribution Q.
2. State variables z depend also on θ and c:

$$z = z(c, \theta, L) \tag{8}$$

Optimization

QoIs

1. **Power:** $Pw = Pw(c; z, \theta)$
2. **Efficiency:** $\text{Eff} = \text{Eff}(c; z, \theta)$
3. **Temperature(s)** $T_{\text{TE, hot}} = T_{\text{TE, hot}}(c; z, \theta)$ and $T_{\text{cat}} = T_{\text{cat}}(c; z, \theta)$

Optimization

$$\max_{c} \text{Eff}(c; z, \theta) = \text{Eff}(c^*; z, \theta)$$

subject to the constraints:

1. **Minimum power threshold:**

$$Pw(c; z, \theta) \geq Pw_{-} = 2.25\text{Watts}$$

2. **Maximum temperature on heat spreader and catalyst:**

$$T_{\text{TE, hot}} \leq T^{+}_{\text{TE, hot}} = 249^o C\,, \quad T_{\text{cat}} \leq T^{+}_{\text{cat}} = 1000^o C$$

The optimal c^* may also need to satisfy a lower performance bound: $\text{Eff}(c^*; z, \theta) \geq 2.8\%$

Robust Optimization under Uncertainty

- Desire **guarantees** of performance, in the presence of inherent **multiple** uncertainties that need to be modeled and accounted in the optimization.
- Rely on **robust optimization/control** framework:
  - Dupuis, James, Petersen *Math. of Control, Signals and Systems* (2000),
  - Petersen, James, Dupuis *IEEE Trans. on Auto. Control* (2000).
  - Substantial (more) recent literature in Macroeconomics, Operations Research, Finance.
- UQ information methods to account for model uncertainties.

Probabilistic Formulation

1. Assume that the parameter vector θ is a random variable or represents parameters in a structured probabilistic model (e.g. directed graphs), with PDF P:

$$\theta \sim P$$

$P =$ nominal, (surrogate, approximation, etc) of "real" model $= Q$

2. Optimized QoI:

$$E_P\Big[\text{Eff}(c; z, \theta)\Big] = E_P f(c)$$

3. Set of **alternative probabilistic models** to P, within model uncertainty level η:

$$\mathcal{Q}_\eta = \{Q : R(Q|P) \leq \eta\}$$

4. Constraints (later).

Robust Optimization Formulation

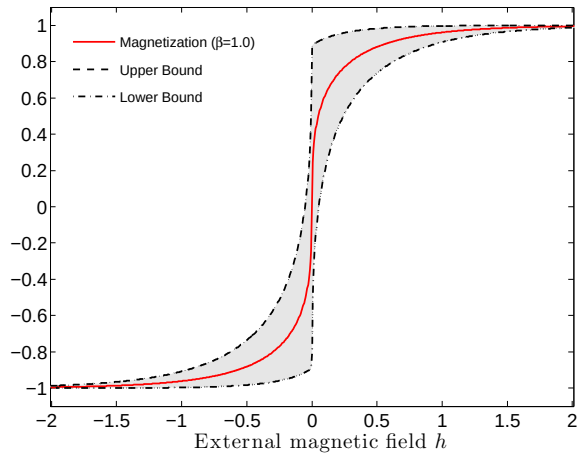
We consider the two robust (=worst case) optimization problems:

$$\max_{c} \max_{Q \in \mathcal{Q}_\eta} E_Q f(c)$$

and

$$\max_{c} \min_{Q \in \mathcal{Q}_\eta} E_Q f(c)$$

where $\mathcal{Q}_\eta = \{Q : R(Q|P) \leq \eta\}$.



Some challenges

- **#1:** How to solve the robust optimization problem given that

$$\mathcal{Q}_\eta = \{Q : R(Q|P) \leq \eta\}$$

  is infinite dimensional?

  $\downarrow$

  We use the UQ information inequalities developed earlier.

- **Non-parametric statistics** aspects: uncertainty due to (a) sparse data, or (b) lack of knowledge of the right class of parametrized models
- **# 2:** Impose **constraints** on certain QoIs e.g. minimum power, maximum temperature, but taking into account
  - *randomness*, and
  - *model uncertainty*

Robust optimization formulations

Remove the infinite dimensional uncertainty in $\mathcal{Q}_\eta$:

$$\max_c \min_{Q\in\mathcal{Q}_\eta} E_Q f(c)\Big[\mathrm{Eff}(c;z,\theta)\Big] = \max_c I^-(f,P;\eta)$$

1. $\max_c \min_{Q\in\mathcal{Q}_\eta} E_Q f(c) = \max_c \max_{\lambda<0}\left[\frac{1}{\lambda}\log\int e^{\lambda f(c,\zeta)}P(\zeta)d\zeta + \frac{\eta}{\lambda}\right]$
2. If the uncertainty level η is relatively small, linearization implies:

$$\max_c \min_{Q\in\mathcal{Q}_\eta} E_Q f(c) \approx \max_c \left[E_P f(c) - \sqrt{2Var_P(f(c))}\eta^{1/2} + \ldots\right]$$

3. Can get more practical but less sharp robust bounds using **concentration inequalities**[1]

$$\int e^{\lambda f(c;\zeta)}P(\zeta)d\zeta \le \Phi(c,\lambda)\,, \quad \lambda\in(-\lambda_0,\lambda_0)$$

for some function Φ (that needs to be identified)[1]

[1] For UQ Info. Inequalities: Gourgoulias, K. Rey-Bellet, Wang, *Arxiv*, '17.

Robust Optimization under constraints

Application constraints:

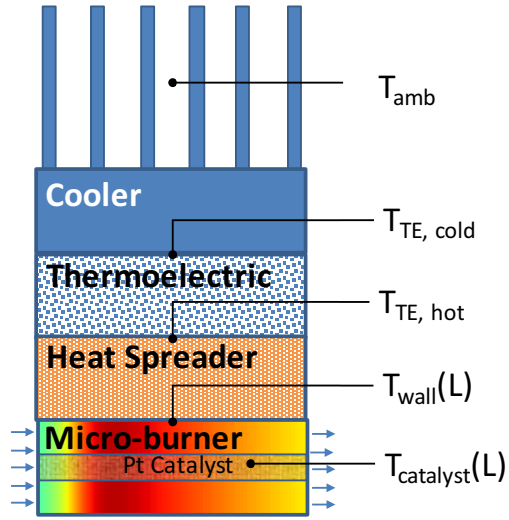
1. **Minimum power threshold:**

$$Pw(c; z, \theta) \geq Pw_{-} = 2.25\text{Watts}$$

2. **Maximum temperature on heat spreader and catalyst:**

$$T_{\text{TE, hot}} \leq T^{+}_{\text{TE, hot}} = 249^{o}C\,,$$

$$T_{\text{cat}} \leq T^{+}_{\text{cat}} = 1000^{o}C$$



Mathematical formulation:

$$E_Q h_i(X) \leq \eta_i \quad \text{(constraints)}$$

$$\max_{c} \min_{Q \in \mathcal{Q}_\eta + \text{constraints}} E_Q f(c) = \max_{c} \max_{\lambda, \lambda_i < 0} \left[\frac{1}{\lambda} \log \int e^{\lambda f(c,\zeta) + \sum_i \lambda_i h_i(\zeta)} P(\zeta) d\zeta + \frac{\eta}{\lambda} + \sum_i \lambda_i \eta_i \right]$$

Conclusions

- Information metrics for quantifying model form uncertainty: scaling properties
- Tightness control of QoIs in terms of KL divergence $\mapsto$ UQ Indices
- Fast sensitivity screening using path FIM for high dimensional Chemical Reaction Networks (CRN)
- Probabilistic Graphical Models (PGM) arise naturally in CRN and related modeling
- UQ information methods are suitable for multiple sources of uncertainty in PGMs
- Robust optimization, design and UQ.

References

- **Information Inequalities and UQ**

  Path-Space Information Bounds for Uncertainty Quantification and Sensitivity Analysis of Stochastic Dynamics, P. Dupuis, K., Y. Pantazis, P. Plechac, **SIAM/ASA JUQ** (2016).

  Scalable Information Inequalities for Uncertainty Quantification, M.K., L. Rey-Bellet, J. Wang, **J. Comp. Phys.** (2017).

  How biased is your model? Concentration Inequalities, Information and Model Bias, K. Gourgoulias, M.K., L. Rey-Bellet, and J.Wang. Arxiv (2017).

  Robust information divergences for model-form uncertainty arising from sparse data in random PDE, E.J. Hall, M.K., Arxiv (2017).

- **Correlations, UQ and Bayesian networks**

  Effects of correlated parameters and uncertainty in electronic-structure-based chemical kinetic modelling, J. E. Sutton, W. Guo, M.K. , D. G. Vlachos. **Nature Chemistry**, (2016).

  Nonparametric correlative uncertainty quantification: Application to a Langmuir bimolecular adsorption model., J. Feng; J. Lansford; A. Mironenko; D. Babaei Pourkargar; D. G. Vlachos, M.K. submitted (2017)

- **Sensitivity Analysis in Path Space**

  A Relative Entropy Rate Method for Path Space Sensitivity Analysis of Stationary Complex Stochastic Dynamics, Y. Pantazis, M. Katsoulakis, **J. Chem. Phys.** (2013).

  Parametric Sensitivity Analysis for Biochemical Reaction Networks based on Pathwise Information Theory, M. Katsoulakis, D. Vlachos, Y. Pantazis, **BMC Bioinformatics**, (2013).

  Accelerated Sensitivity Analysis in High-Dimensional Stochastic Reaction Networks, G. Arampatzis, M. K. Y. Pantazis, **PLOS 1**, (2015).

  ISAP-MATLAB package for sensitivity analysis of high-dimensional stochastic chemical networks, W. Hu, Y. Pantazis and M.K. , **J. Stat. Software**, (2018).