

Sensitivity Analysis, Uncertainty Quantification and Inference in Stochastic Dynamics

Yannis Pantazis

University of Crete

IPAM, January, 2016

Sensitivity Analysis in Path-space

Introduction

Mathematical Theory

Well-mixed reaction networks

Molecular dynamics

Uncertainty Quantification and Sensitivity Bounds

Typical observables

Rare Events

Inference in Pathway Signalling

Joint work with

- ▶ Markos Katsoulakis (University of Massachusetts, Amherst)
- ▶ Paul Dupuis (Brown University)
- ▶ Vagelis Harmandaris (University of Crete)
- ▶ Petr Plechac (University of Delaware)
- ▶ Luc Rey-Bellet (University of Massachusetts, Amherst)
- ▶ Dion Vlachos (Chem. Engineering, University of Delaware)
- ▶ Georgios Arampatzis (Comp. Science & Engineering, ETH)
- ▶ William Hu (University of Massachusetts, Amherst)
- ▶ Tasos Tsourtis (University of Crete)
- ▶ Ioannis Tsamardinos (Comp. Science, University of Crete)

Outline

Sensitivity Analysis in Path-space

Introduction

Mathematical Theory

Well-mixed reaction networks

Molecular dynamics

Uncertainty Quantification and Sensitivity Bounds

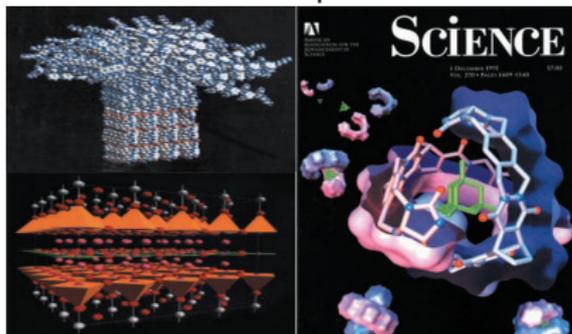
Typical observables

Rare Events

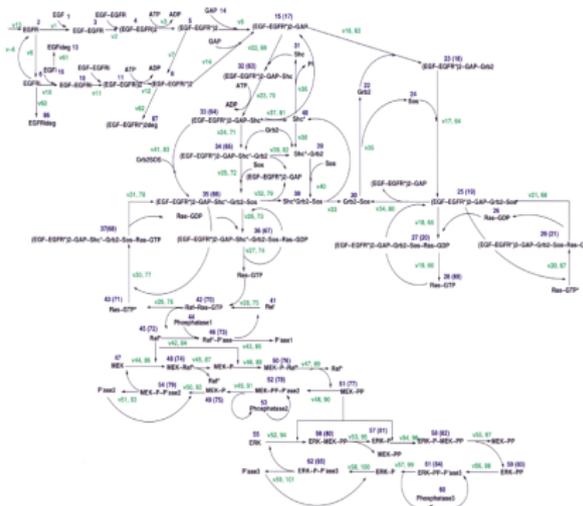
Inference in Pathway Signalling

Sensitivity analysis - Motivation

Complex Systems, DOE Office of Science Report

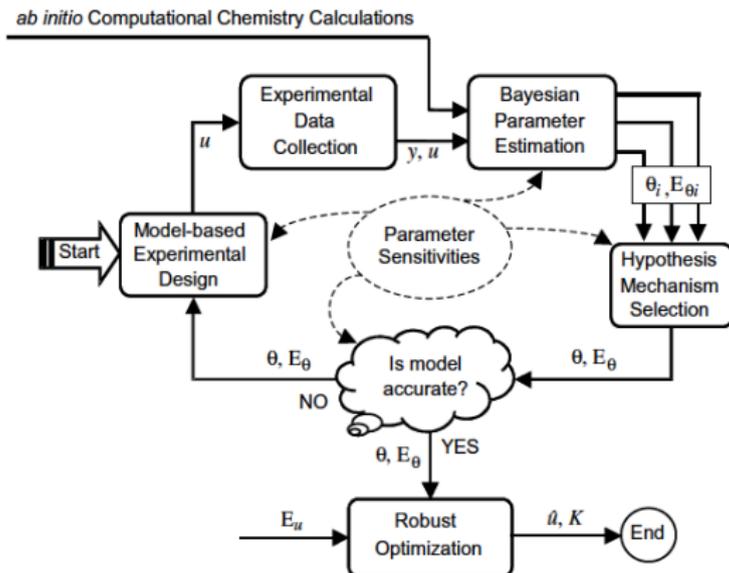


“Computational modeling of the dynamics of the MAP kinase cascade activated by surface and internalized EGF receptors”, Schoeberl et al., Nature Biotechnology, 2002



Sensitivity analysis - Motivation

- **Sensitivity Analysis:** Impact in the output caused by perturbations in the input.



Perspectives on the design and control of multiscale systems, Braatz et al., 2006

Sensitivity analysis - Related concepts & Applications

- ▶ **Robustness:** The persistence of a system to a desirable state under the perturbations of external conditions.
- ▶ **Identifiability:** The ability of the experimental data and/or the model to estimate the model parameters with high fidelity.
- ▶ **Reliability:** To ensure that the performance of a system meets some pre-specified target with a given probability.
- ▶ **Optimal experimental design:** Construct experiments so that the parameters can be estimated from the resulting experimental data with the highest statistical quality. **Fisher information matrix.**
- ▶ **Uncertainty quantification:** The science of quantitative characterization and reduction of uncertainties.

Sensitivity analysis - General

- ▶ **Global sensitivity analysis:** Study the effect of an input parameter to the output under a wide range of values or under a specified distribution.
- ▶ **Local sensitivity analysis:** Fix input parameters to a value and study the effect of small perturbations around the fixed parameter to the output. Local SA is typically a prerequisite for global SA.
 - ▶ **Observable-based SA.** Let $\{x_t\}$ is a stochastic process depending on $\theta \in \mathbb{R}$ and f is an observable. Compute

$$S_{P_t^\theta, f}(\theta) = \frac{d}{d\theta} \mathbb{E}_{P_t^\theta} [f(x)]$$

- ▶ **Density-based SA.** Compare unperturbed and perturbed densities utilizing a metric or a divergence or compute the **Fisher information matrix**.

Sensitivity analysis - Challenges

- ▶ Perform SA on **stochastic** systems.
- ▶ Perform SA for computational models with a **very large number of parameters**.
- ▶ Tackling with **non-equilibrium** and/or **out-of-equilibrium** stochastic systems.
- ▶ Infer sensitivity information about the **dynamics** of a stochastic process.
- ▶ Perform SA **fast** and **reliable**.
 - ▶ **Reduction of variance** of statistical estimators.

Preliminaries

- ▶ Let $\{x_m\}_{m \in \mathbb{Z}_+}$ be a *discrete-time Markov chain* at **steady states**.
- ▶ The path distribution from 0 to T is given by

$$Q_{0,T}^\theta(x_0, \dots, x_T) = \mu^\theta(x_0) p^\theta(x_0, x_1) \cdots p^\theta(x_{T-1}, x_T)$$

where

- ▶ $\theta \in \mathbb{R}^K$ is the parameter vector.
- ▶ $\mu^\theta(x)$ is the stationary distribution.
- ▶ $p^\theta(x, x')$ is the transition probability function.

Relative entropy in path-space

- ▶ Suggest performing parameter sensitivity analysis by comparing the path distributions utilizing **relative entropy**.
- ▶ **Definition:** The **path relative entropy** of $Q_{0,T}^\theta$ w.r.t. $Q_{0,T}^{\theta+\epsilon}$ is

$$\mathcal{R} \left(Q_{0,T}^\theta \parallel Q_{0,T}^{\theta+\epsilon} \right) := \int \log \left(\frac{dQ_{0,T}^\theta}{dQ_{0,T}^{\theta+\epsilon}} \right) dQ_{0,T}^\theta$$

- ▶ **Properties:** (i) $\mathcal{R} \left(Q_{0,T}^\theta \parallel Q_{0,T}^{\theta+\epsilon} \right) \geq 0$ and
- (ii) $\mathcal{R} \left(Q_{0,T}^\theta \parallel Q_{0,T}^{\theta+\epsilon} \right) = 0$ iff $Q_{0,T}^\theta = Q_{0,T}^{\theta+\epsilon}$
- (iii) $\mathcal{R} \left(Q_{0,T}^\theta \parallel Q_{0,T}^{\theta+\epsilon} \right) \leq \mathcal{R} \left(Q_{0,T+1}^\theta \parallel Q_{0,T+1}^{\theta+\epsilon} \right)$

Relative entropy rate & path FIM

- ▶ The **relative entropy rate (RER)** is defined as

$$\mathcal{H}(Q^\theta \parallel Q^{\theta+\epsilon}) = \lim_{T \rightarrow \infty} \frac{1}{T} \mathcal{R}(Q_{0,T}^\theta \parallel Q_{0,T}^{\theta+\epsilon})$$

- ▶ Easy to show that

$$\mathcal{H}(Q^\theta \parallel Q^{\theta+\epsilon}) = \mathbb{E}_{\mu^\theta} \left[\int p^\theta(x, x') \log \frac{p^\theta(x, x')}{p^{\theta+\epsilon}(x, x')} dx' \right]$$

Relative entropy rate & path FIM

- ▶ The **relative entropy rate (RER)** is defined as

$$\mathcal{H}(Q^\theta \parallel Q^{\theta+\epsilon}) = \lim_{T \rightarrow \infty} \frac{1}{T} \mathcal{R}(Q_{0,T}^\theta \parallel Q_{0,T}^{\theta+\epsilon})$$

- ▶ Easy to show that

$$\mathcal{H}(Q^\theta \parallel Q^{\theta+\epsilon}) = \mathbb{E}_{\mu^\theta} \left[\int p^\theta(x, x') \log \frac{p^\theta(x, x')}{p^{\theta+\epsilon}(x, x')} dx' \right]$$

- ▶ Under smoothness assumption on θ , it holds that

$$\mathcal{H}(Q^\theta \parallel Q^{\theta+\epsilon}) = \frac{1}{2} \epsilon^T \mathcal{I}_{\mathcal{H}}(Q^\theta) \epsilon + O(|\epsilon|^3)$$

- ▶ where **path Fisher information matrix** is defined as

$$\mathcal{I}_{\mathcal{H}}(Q^\theta) = \mathbb{E}_{\mu^\theta} \left[\int p^\theta(x, x') \nabla_\theta \log p^\theta(x, x') \nabla_\theta \log p^\theta(x, x')^T dx' \right]$$

What do we gain?

- ▶ **RER** contains information not only for the invariant measure but also for the stationary dynamics such as metastable dynamics, time correlations, etc..
- ▶ No need for explicit knowledge of invariant measure.
- ▶ **RER** is an observable of known test function \Rightarrow tractable and statistical estimators can provide easily and efficiently its value.
- ▶ Different ϵ 's \Rightarrow SA at different directions.

- ▶ **FIM** constitutes a **derivative-free** sensitivity analysis method.
- ▶ Robustness on parameter perturbations as well as parameter identifiability can be inferred from the **FIM**.
 - ▶ Wider **FIM** implies a more robust model.
 - ▶ Steeper **FIM** implies more identifiable parameters.

Well-mixed reaction networks

- ▶ A **continuous-time jump Markov process** is fully determined by the **transition rates**, $c(x, x')$.
 - ▶ The total rate is defined by $\lambda(x) = \sum_{x'} c(x, x')$.

- ▶ **Relative entropy rate:**

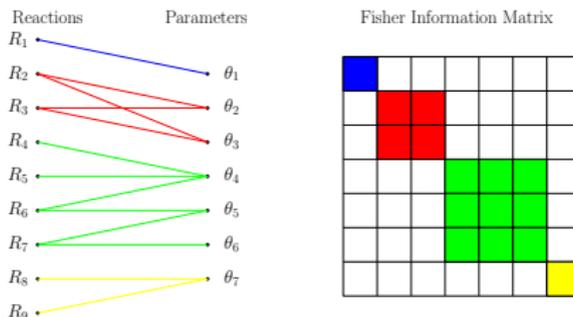
$$\mathcal{H}(Q^\theta \parallel Q^{\theta+\epsilon}) = \mathbb{E}_{\mu^\theta} \left[\sum_{x'} c^\theta(x, x') \log \frac{c^\theta(x, x')}{c^{\theta+\epsilon}(x, x')} - (\lambda^\theta(x) - \lambda^{\theta+\epsilon}(x)) \right]$$

- ▶ Derived using Girsanov formula for $dQ_{0,T}^\theta / dQ_{0,T}^{\theta+\epsilon}$.
- ▶ **Path Fisher information matrix:**

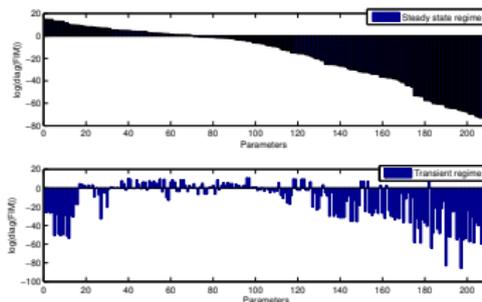
$$\mathcal{I}_{\mathcal{H}}(Q^\theta) := \mathbb{E}_{\mu^\theta} \left[\sum_{x'} c^\theta(x, x') \nabla_\theta \log c^\theta(x, x') \nabla_\theta \log c^\theta(x, x')^T \right]$$

- ▶ “Parametric Sensitivity Analysis for Biochemical Reaction Networks based on Pathwise Information Theory”, Y.P. - M. Katsoulakis - D. Vlachos, BMC Bioinformatics (2013).

Path FIM - Epidermal Growth Factor Receptor



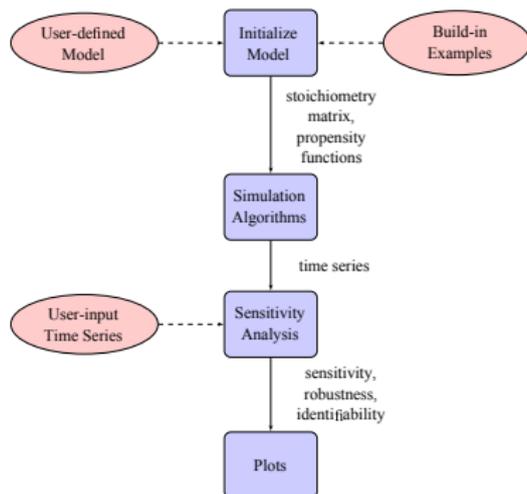
Left panel: The graph representation of the dependencies between the reactions (left column) and the model parameters (right column). Right panel: The corresponding block-diagonal structure of the FIM.



Diagonal elements of the path FIM computed at the steady state regime (upper) and at the transient regime (lower) for EGFR. Parameter sensitivities differ by orders of magnitude; known as "sloppiness" in biology.

Info-theoretic Sensitivity Analysis in Path-space (ISAP)

- ▶ “ISAP-MATLAB package for sensitivity analysis of high-dimensional stochastic chemical networks”, submitted to J. of Statistical Software, *W. Hu, Y.P. and M. Katsoulakis*.
- ▶ ISAP: performs simulation and sensitivity analysis for reaction networks.
- ▶ <http://people.math.umass.edu/pantazis/source/ISAP.zip>



Langevin process - Molecular Dynamics

- ▶ A system of stochastic differential equations with N particles:

$$\begin{cases} dq_t = M^{-1}p_t dt \\ dp_t = F^\theta(q_t)dt - \gamma M^{-1}p_t dt + \sigma dW_t, \end{cases}$$

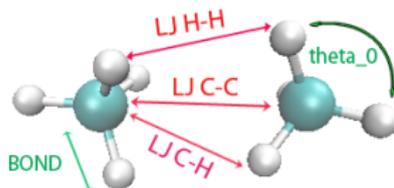
- ▶ **Relative entropy rate:**

$$\mathcal{H}(Q^\theta \| Q^{\theta+\epsilon}) = \frac{1}{2} \mathbb{E}_{\mu^\theta} [(F^{\theta+\epsilon}(q) - F^\theta(q))^T (\sigma \sigma^T)^{-1} (F^{\theta+\epsilon}(q) - F^\theta(q))]$$

- ▶ **Path Fisher information matrix:**

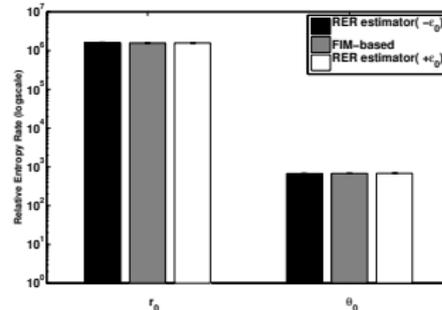
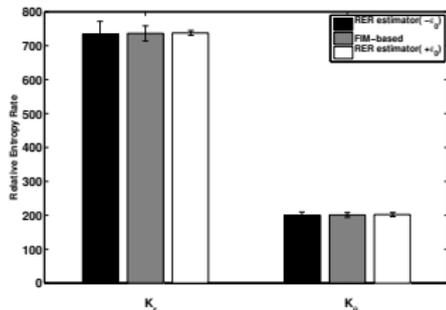
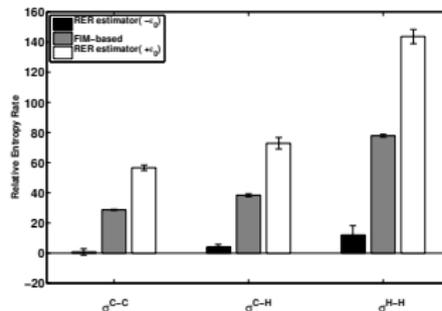
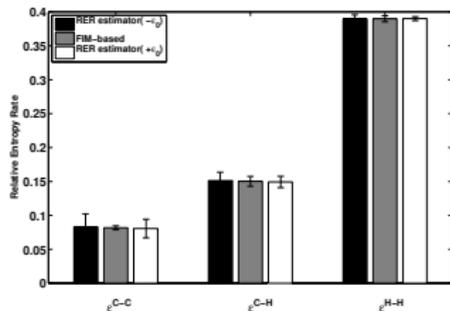
$$\mathcal{I}_{\mathcal{H}}(Q^\theta) = \mathbb{E}_{\mu^\theta} [\nabla_\theta F^\theta(q)^T (\sigma \sigma^T)^{-1} \nabla_\theta F^\theta(q)],$$

- ▶ "Sensitivity Analysis for Stochastic Molecular Systems using Information Metrics", A. Tsourtis, Y.P., M. Katsoulakis and V. Harmandaris, JCP.
- ▶ Methane - CH_4



CH4 - RER and path FIM

Per particle RER and FIM-approximated RER comparison with 5% perturbation in all the parameters. The parameters are grouped according to their order of magnitude.



Outline

Sensitivity Analysis in Path-space

Introduction

Mathematical Theory

Well-mixed reaction networks

Molecular dynamics

Uncertainty Quantification and Sensitivity Bounds

Typical observables

Rare Events

Inference in Pathway Signalling

Connection with observable-based SA

- ▶ **Question:** Are there any links between **density-based** approaches (relative entropy, FIM) and **observable-based** approaches ($S_{p^\theta, f}(\theta) = \partial_\theta \mathbb{E}_{p^\theta}[f]$)?

Connection with observable-based SA

- ▶ **Question:** Are there any links between **density-based** approaches (relative entropy, FIM) and **observable-based** approaches ($S_{P^\theta, f}(\theta) = \partial_\theta \mathbb{E}_{P^\theta}[f]$)?
- ▶ Based on Pinsker (or Csiszar-Kullback-Pinsker) inequality:

$$|\mathbb{E}_Q[f] - \mathbb{E}_P[f]| \leq \|f\|_\infty \sqrt{2\mathcal{R}(Q \| P)}$$

- ▶ Q : “true” probabilistic model
- ▶ P : “nominal” or “reference” model

Connection with observable-based SA

- ▶ **Question:** Are there any links between **density-based** approaches (relative entropy, FIM) and **observable-based** approaches ($S_{P_\theta, f}(\theta) = \partial_\theta \mathbb{E}_{P_\theta}[f]$)?
- ▶ Based on Pinsker (or Csiszar-Kullback-Pinsker) inequality:

$$|\mathbb{E}_Q[f] - \mathbb{E}_P[f]| \leq \|f\|_\infty \sqrt{2\mathcal{R}(Q \| P)}$$

- ▶ Q : “true” probabilistic model
- ▶ P : “nominal” or “reference” model
- ▶ For any bounded f with finite $\Lambda(c)$ around 0, **UQ information inequalities** (*Chowdhary & Dupuis, ESAIM '13, Li & Xiu, '12, SIAM Sci. Comp.*):

$$\sup_{c>0} -\frac{\Lambda(-c) + \mathcal{R}(Q \| P)}{c} \leq \mathbb{E}_Q[f] - \mathbb{E}_P[f] \leq \inf_{c>0} \frac{\Lambda(c) + \mathcal{R}(Q \| P)}{c}$$

- ▶ $\Lambda(c) := \log \mathbb{E}_P[\exp\{c(f - \mathbb{E}_P[f])\}]$: **cumulant generating function**.
- ▶ Based on the **Donsker-Varadhan variational formula** for relative entropy.

Uncertainty Quantification Bounds

- ▶ Theorem: Define

$$\Xi_+(Q \parallel P; f) := \inf_{c>0} \frac{\Lambda(c) + \mathcal{R}(Q \parallel P)}{c}$$

Then, $\Xi_+(Q \parallel P; f)$ is a **goal-oriented divergence** that couples the observable f and the relative entropy.

- ▶ That means

- (i) $\Xi_+(Q \parallel P; f) \geq 0$
- (ii) $\Xi_+(Q \parallel P; f) = 0$ iff $Q = P$ or $f = \mathbb{E}[f]$ P -a.s.

Uncertainty Quantification Bounds

- ▶ Theorem: Explicit representations of Ξ_+ :

$$\begin{aligned}\Xi_+(Q \| P; f) &= \tilde{\Lambda}^{-1}(\mathcal{R}(Q \| P)) \\ &= \Lambda'(\Phi^{-1}(\mathcal{R}(Q \| P))) \\ &= \sqrt{2\text{Var}_P(f)}\sqrt{\mathcal{R}(Q \| P)} + O(\mathcal{R}(Q \| P))\end{aligned}$$

- ▶ $\tilde{\Lambda}(k) = \sup_{c>0} \{ck - \tilde{\Lambda}(c)\}$: Legendre transform of $\Lambda(\cdot)$.
- ▶ $\Phi(c) := \Lambda(c) + c\Lambda'(c)$ is related with the Legendre transform of $\Lambda(\cdot)$.
- ▶ Relative entropy **controls** the uncertainty.

Sensitivity Bounds

- ▶ **Sensitivity bound** for observable f :

$$|S_{\mu^\theta, f}(\theta)| \leq \sqrt{\text{Var}_{\mu^\theta}(f)} \sqrt{\mathcal{I}(\mu^\theta)}$$

- ▶ Obtained by setting $Q = \mu^{\theta+\epsilon}$, $P = \mu^\theta$, divide by ϵ and send $\epsilon \rightarrow 0$.
- ▶ Related to the generalized Cramer-Rao theorem (Estimation theory).
- ▶ **Intergovernmental panel for climate change** uses as sensitivity index:

$$\frac{|S_{\mu^\theta, f}(\theta)|}{\text{Std}_{\mu^\theta}(f)} \leq \sqrt{\mathcal{I}(\mu^\theta)}$$

- ▶ But, in non-equilibrium systems, μ^θ is usually **not** known.

Sensitivity Bounds

- ▶ **Sensitivity bound** for **time-averaged** observable

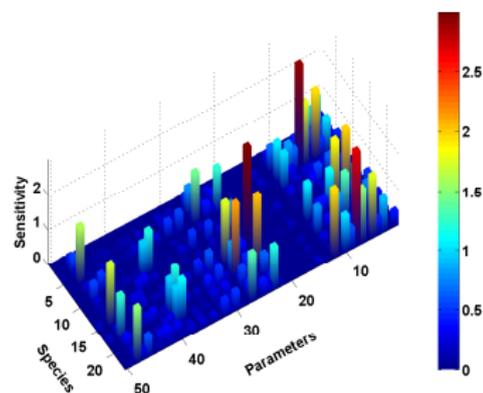
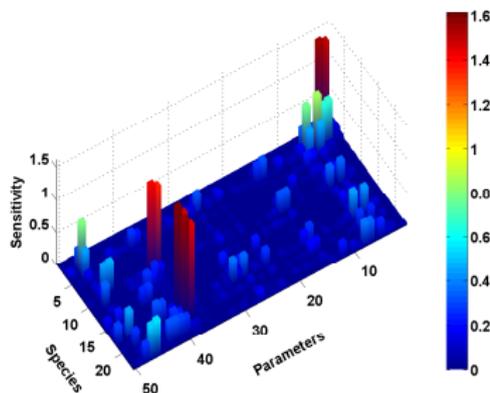
$$F(\mathbf{x}) = \frac{1}{T} \int_0^T f(x_t) dt:$$

$$|S_{Q^\theta, F}(\theta)| \leq \sqrt{\tau_{\mu^\theta}(f)} \sqrt{\mathcal{I}_{\mathcal{H}}(Q^\theta)}$$

- ▶ $\tau_{\mu^\theta}(f) := \int_{-\infty}^{\infty} \langle f(x_t), f(x_0) \rangle_{\mu^\theta} dt$ is the integrated autocorrelation time.
- ▶ More general theory for transient, long-time, infinite time and spatially-extended regimes.
- ▶ “Path-space information bounds for uncertainty quantification and sensitivity analysis of stochastic dynamics”, P. Dupuis - M. Katsoulakis - Y.P. - P. Plechac, SIAM J. on UQ, 2016

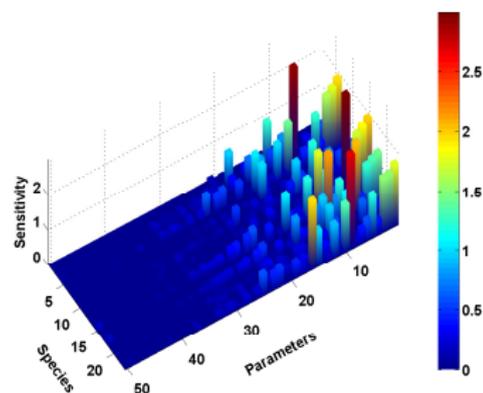
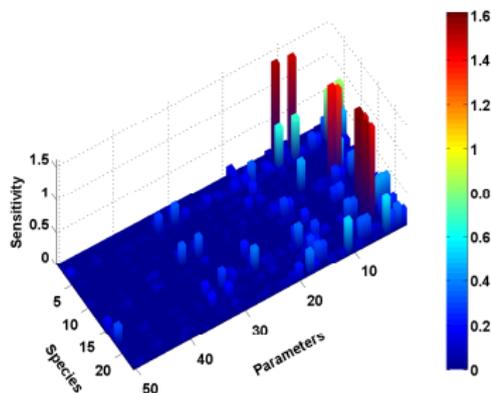
Sensitivity Bound - Validation

- ▶ Another biological model that describes Epidermal Growth Factor Receptor. [Kholodenko et.al., J. Biol. Chem., 1999]
- ▶ 23 species, 47 reactions, 50 parameters, $23 \times 50 = 1150$ sensitivity indices ($S_{ij} = \partial_{\theta_j} \mathbb{E}[X_i]$).



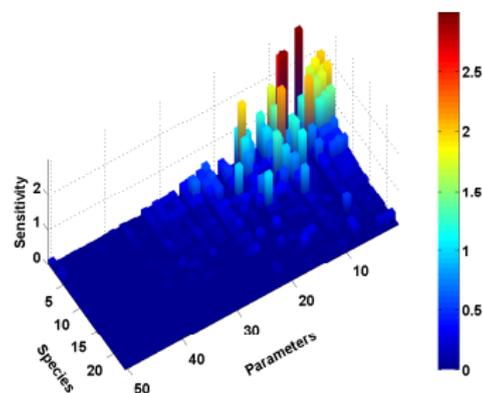
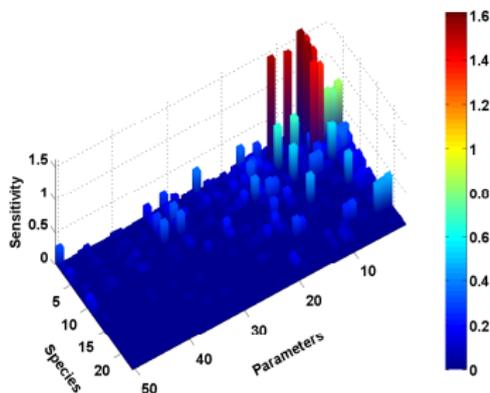
Sensitivity Bound - Validation

- ▶ Another biological model that describes Epidermal Growth Factor Receptor. [Kholodenko et.al., J. Biol. Chem., 1999]
- ▶ 23 species, 47 reactions, 50 parameters, $23 \times 50 = 1150$ sensitivity indices ($S_{ij} = \partial_{\theta_j} \mathbb{E}[X_i]$).



Sensitivity Bound - Validation

- ▶ Another biological model that describes Epidermal Growth Factor Receptor. [Kholodenko et.al., J. Biol. Chem., 1999]
- ▶ 23 species, 47 reactions, 50 parameters, $23 \times 50 = 1150$ sensitivity indices ($S_{ij} = \partial_{\theta_j} \mathbb{E}[X_i]$).



Sensitivity analysis for rare events

- ▶ An event A is **rare** if $P(A) = \mathbb{E}_P[\chi_A] \ll 1$. We usually consider $\log P(A)$.
 - ▶ Applications to: Reliability analysis, queueing theory, operation research, insurance, statistical mechanics.
- ▶ Sensitivity analysis for **rare events**:

$$S_A(P^\theta) := \partial_\theta \log P^\theta(A) = \frac{\partial_\theta P^\theta(A)}{P^\theta(A)}$$

- ▶ Extremely few previous studies on this area!
- ▶ Relative entropy is not the most appropriate divergence.

Sensitivity analysis for rare events

- ▶ An event A is **rare** if $P(A) = \mathbb{E}_P[\chi_A] \ll 1$. We usually consider $\log P(A)$.
 - ▶ Applications to: Reliability analysis, queueing theory, operation research, insurance, statistical mechanics.
- ▶ Sensitivity analysis for **rare events**:

$$S_A(P^\theta) := \partial_\theta \log P^\theta(A) = \frac{\partial_\theta P^\theta(A)}{P^\theta(A)}$$

- ▶ Extremely few previous studies on this area!
 - ▶ Relative entropy is not the most appropriate divergence.
- ▶ **Rare event bounds** based on **Renyi divergence** (*Atar, Chowdhary and Dupuis*, SIAM UQ '14):

$$\frac{1}{\alpha-1} \log Q(A) - \frac{1}{\alpha} \mathcal{R}_\alpha(Q \| P) \leq \frac{1}{\alpha} \log P(A) \leq \frac{1}{\alpha+1} \log Q(A) + \frac{1}{\alpha+1} \mathcal{R}_{\alpha+1}(P \| Q)$$

- ▶ **Renyi divergence**:

$$\mathcal{R}_\alpha(Q \| P) := \frac{1}{\alpha-1} \log \mathbb{E}_P \left[\left(\frac{dQ}{dP} \right)^\alpha \right] = \frac{1}{\alpha-1} \log \mathbb{E}_P \left[e^{\alpha \log \frac{dQ}{dP}} \right].$$

Sensitivity analysis for rare events

► **UQ bound:**

$$\sup_{\alpha > 0} -\frac{H(-\alpha) - \log P(A)}{\alpha} \leq \log Q(A) - \log P(A) \leq \inf_{\alpha > 0} \frac{H(\alpha) - \log P(A)}{\alpha}$$

- $H(\alpha) := \log \mathbb{E}_P[\exp\{\alpha \log \frac{dQ}{dP}\}]$: **cumulant generating function** of $\log \frac{dP}{dQ}$.

Sensitivity analysis for rare events

- ▶ UQ bound:

$$\sup_{\alpha > 0} -\frac{H(-\alpha) - \log P(A)}{\alpha} \leq \log Q(A) - \log P(A) \leq \inf_{\alpha > 0} \frac{H(\alpha) - \log P(A)}{\alpha}$$

- ▶ $H(\alpha) := \log \mathbb{E}_P[\exp\{\alpha \log \frac{dQ}{dP}\}]$: cumulant generating function of $\log \frac{dP}{dQ}$.
- ▶ Does these bounds remind you something?

Sensitivity analysis for rare events

► **UQ bound:**

$$\sup_{\alpha > 0} -\frac{H(-\alpha) - \log P(A)}{\alpha} \leq \log Q(A) - \log P(A) \leq \inf_{\alpha > 0} \frac{H(\alpha) - \log P(A)}{\alpha}$$

► $H(\alpha) := \log \mathbb{E}_P[\exp\{\alpha \log \frac{dQ}{dP}\}]$: **cumulant generating function** of $\log \frac{dP}{dQ}$.

► **Does these bounds remind you something?**

$$\sup_{c > 0} -\frac{\Lambda(-c) + \mathcal{R}(Q \parallel P)}{c} \leq \mathbb{E}_Q[f] - \mathbb{E}_P[f] \leq \inf_{c > 0} \frac{\Lambda(c) + \mathcal{R}(Q \parallel P)}{c}$$

Sensitivity analysis for rare events

- ▶ **Sensitivity bounds** ($Q = P^{\theta+\epsilon}$, $P = P^\theta$ and $\alpha = \frac{1}{\epsilon}(\alpha_0 + O(\epsilon))$):

$$\sup_{\alpha>0} -\frac{\bar{H}(-\alpha) - \log P^\theta(A)}{\alpha} \leq S_A(P^\theta) \leq \inf_{\alpha>0} \frac{\bar{H}(\alpha) - \log P^\theta(A)}{\alpha}$$

- ▶ $\bar{H}(\alpha) := \log \mathbb{E}_P[\exp\{\alpha \partial_\theta \log P^\theta\}]$: **cumulant generating function** of $\partial_\theta \log P^\theta$.

Sensitivity analysis for rare events

- ▶ **Sensitivity bounds** ($Q = P^{\theta+\epsilon}$, $P = P^\theta$ and $\alpha = \frac{1}{\epsilon}(\alpha_0 + O(\epsilon))$):

$$\sup_{\alpha>0} -\frac{\bar{H}(-\alpha) - \log P^\theta(A)}{\alpha} \leq S_A(P^\theta) \leq \inf_{\alpha>0} \frac{\bar{H}(\alpha) - \log P^\theta(A)}{\alpha}$$

- ▶ $\bar{H}(\alpha) := \log \mathbb{E}_P[\exp\{\alpha \partial_\theta \log P^\theta\}]$: **cumulant generating function** of $\partial_\theta \log P^\theta$.
- ▶ Extensions to bound rate function derivatives (Large Deviations, Moderate Deviations, etc).
- ▶ “Sensitivity Bounds for Rare Events based on Renyi Divergence”, *P. Dupuis, M. Katsoulakis, Y.P. and L. Rey-Bellet*

Outline

Sensitivity Analysis in Path-space

Introduction

Mathematical Theory

Well-mixed reaction networks

Molecular dynamics

Uncertainty Quantification and Sensitivity Bounds

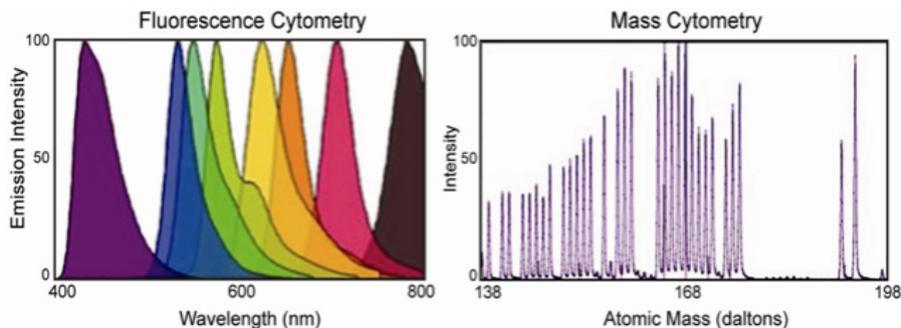
Typical observables

Rare Events

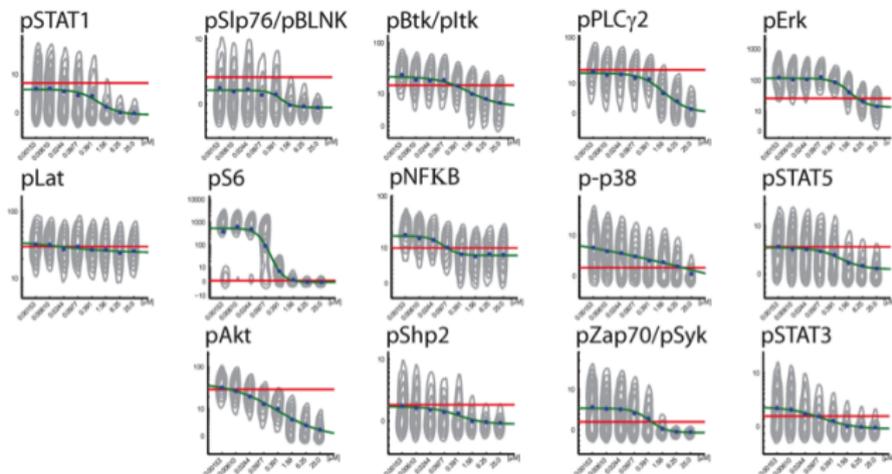
Inference in Pathway Signalling

Mass cytometry

- ▶ CYTOF: Single-cell measurements
- ▶ Nolan group at Stanford, 2010-2012
- ▶ Can measure 40-60 proteins simultaneously compared to 3-8 of flow cytometry



Mass cytometry - Data



14 phosphorylation sites per cell type

“Multiplexed mass cytometry profiling of cellular states perturbed by small-molecule regulators”, Bodenmiller et al., Nature Biotechnology, 2012

Dynamical model

- ▶ System of ODEs:

$$\dot{x} = A\psi(x) + u, \quad x(0) = x_0$$

- ▶ $x = x(t) \in \mathbb{R}^N$: State (or variable) vector
 - ▶ $A \in \mathbb{R}^{N \times Q}$: Parameter matrix
 - ▶ $\psi : \mathbb{R}^N \rightarrow \mathbb{R}^Q$: Vector-valued vector function
 - ▶ $u = u(t) : \mathbb{R} \rightarrow \mathbb{R}^N$: intervention vector function
-
- ▶ **Goal:** Given measurements and $\psi(\cdot)$ estimate A .
 - ▶ Structure estimation \iff finding the zeros of matrix A .
 - ▶ Parameter estimation \iff estimate the non-zero elements of A .

Inference approach

Ingredients:

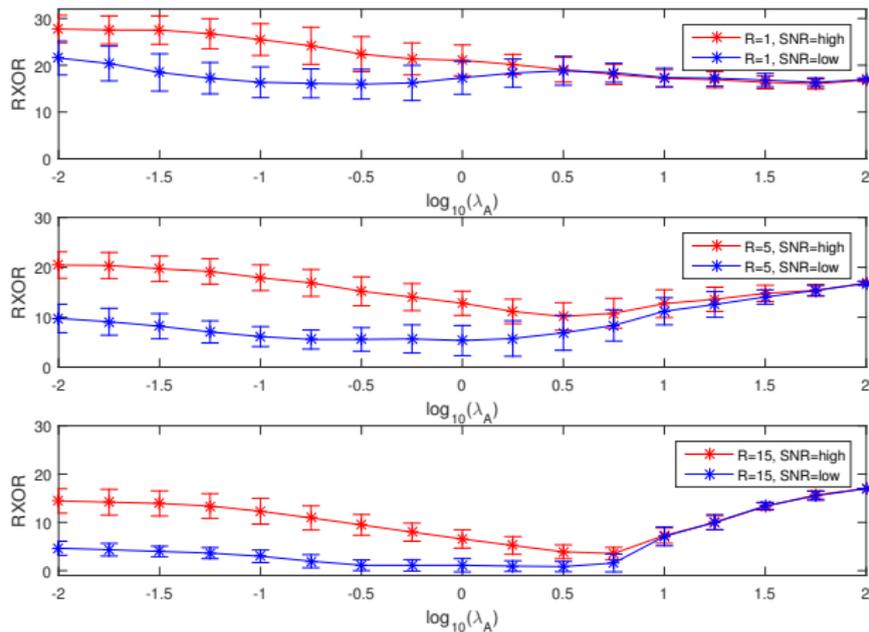
- ▶ Handle **sparse sampling** using **Collocation method** (ie, time-series interpolation).
- ▶ Handle **multiple interventions** using **Multiple shooting**.
- ▶ Add ODE equations as **soft constraints**.
- ▶ Deal with **sparsity** of A using **LASSO**-type cost functional.
 - ▶ LASSO: Least Absolute Shrinkage and Selection Operator
 - ▶ λ_A : weight that controls the sparsity.
- ▶ Overall optimize a **nonlinear** cost functional.

ABCDEFGG - An example

Reaction	Propensity	Rate constant (k_i)
$A \rightarrow B$	$k_1 X_A$	0.01
$B \rightarrow C$	$k_2 X_B$	0.04
$B \rightarrow D$	$k_3 X_B$	0.1
$C \rightarrow E$	$k_4 X_C$	0.1
$C \rightarrow F$	$k_5 X_C$	0.01
$D \rightarrow C$	$k_6 X_D$	0.2
$E \rightarrow D$	$k_7 X_E$	0.4
$E \rightarrow G$	$k_8 X_E$	0.01
$F \rightarrow G$	$k_9 X_F$	0.01
$G \rightarrow A$	$k_{10} X_G$	0.08

- ▶ Matrix A has 17 nonzero elements out of 49.

ABCDEFG - Error



References

1. *A Relative Entropy Rate Method for Path Space Sensitivity Analysis of Stationary Complex Stochastic Dynamics*, M. Katsoulakis - Y.P., J. of Chem. Phys. (2013).
2. *Parametric Sensitivity Analysis for Biochemical Reaction Networks based on Pathwise Information Theory*, M. Katsoulakis - D. Vlachos - Y.P., BMC Bioinformatics (2013).
3. *Pathwise Sensitivity Analysis in Transient Regimes*, G. Arampatzis - M. Katsoulakis - Y.P., RMMC proceedings (2015)
4. *Parametric Sensitivity Analysis for Stochastic Molecular Systems using Pathwise Information Metrics*, A. Tsourtis - V. Harmandaris - M. Katsoulakis - Y.P. (accepted)
5. *Accelerated Sensitivity Analysis in High-Dimensional Stochastic Reaction Networks*, G. Arampatzis - M. Katsoulakis - Y.P. (2015)
6. *ISAP-MATLAB package for sensitivity analysis of high-dimensional stochastic chemical networks*, W. Hu - M. Katsoulakis - Y.P. (submitted)
7. *Path-space information bounds for uncertainty quantification and sensitivity analysis of stochastic dynamics*, P. Dupuis - M. Katsoulakis - P. Plechac - Y.P., SIAM J. on UQ (2016)
8. *Sensitivity Bounds for Rare Events based on Renyi Divergence*, P. Dupuis - M. Katsoulakis - L. Rey-Bellet - Y.P. (in preparation)

The end

THANK YOU!!!