Stochastic PDEs, Sparse Approximations, and Compressive Sampling

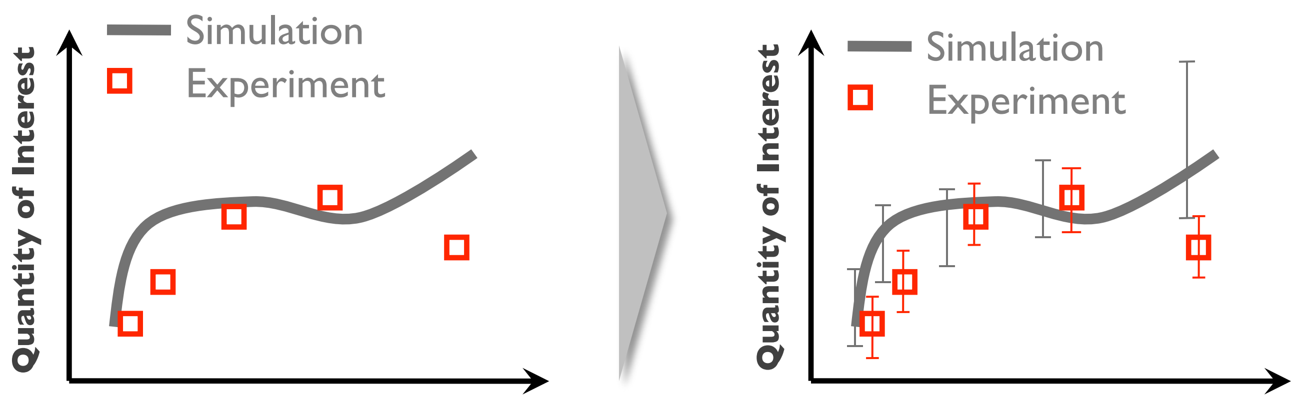
Alireza Doostan

University of Colorado Boulder
Aerospace Engineering Sciences
doostan@colorado.edu

DOE Applied Math Program Meeting, Reston, VA 2011

Motivation – Predictive simulation of engineering systems

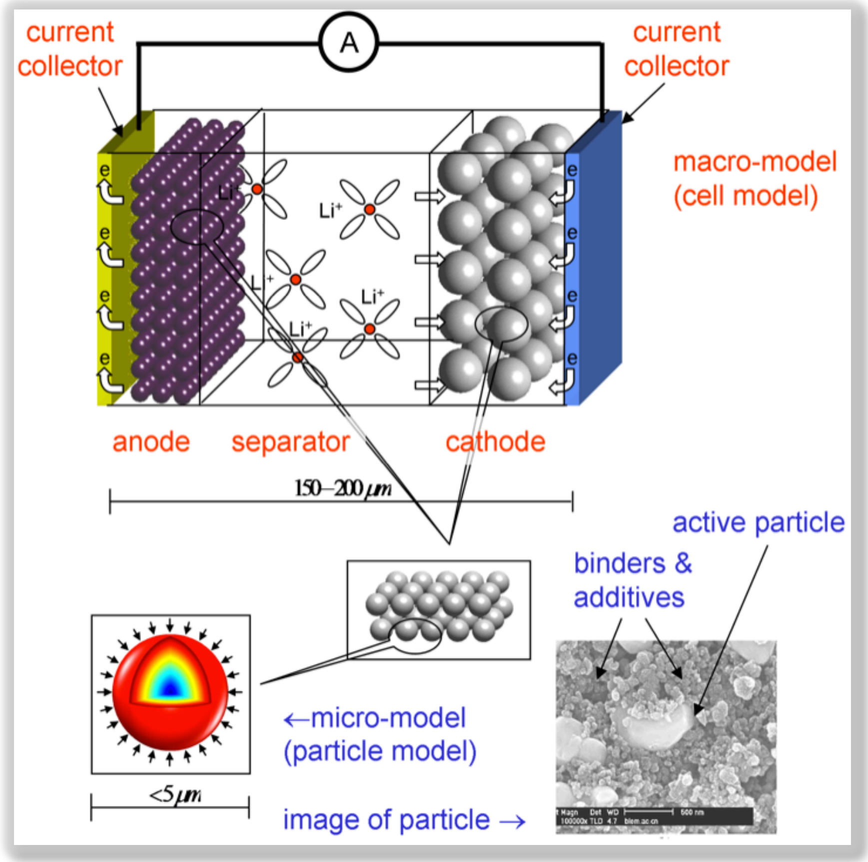
- Data-driven representation of uncertainty
  - Model parameters or structure
- Propagation of uncertainties
- Model Verification and Validation (V&V)
- Certification and uncertainty management



Challenges for complex systems

- Multiple physics
- Multiple length/time scales
- Limited and noisy data
    - sub-scale model calibration
- Large number of uncertain variables
    - uncertainty propagation
- Expensive forward solves
- Verification
- …

Lithium-ion battery cell



From deterministic to stochastic PDEs

Probabilistic approach:

- Define an abstract probability space $(\Omega, \mathcal{A}, \mathcal{P})$
- Represent data using random variables $\boldsymbol{y}(\omega) : \Omega \to \mathbb{R}^d$
  - model parameters/structure, initial conditions, boundary conditions, …

Stochastic PDEs:

$$\boldsymbol{\mathcal{L}}(\boldsymbol{x}, t, \boldsymbol{y}(\omega); u) = 0, \quad (\boldsymbol{x}, t) \in \mathcal{D} \times [0, T]$$

B.C. : $$\boldsymbol{\mathcal{B}}(\boldsymbol{x}, t, \boldsymbol{y}(\omega); u) = 0, \quad (\boldsymbol{x}, t) \in \partial\mathcal{D} \times [0, T]$$

I.C. : $$\boldsymbol{\mathcal{I}}(\boldsymbol{x}, 0, \boldsymbol{y}(\omega); u) = 0, \quad \boldsymbol{x} \in \mathcal{D}$$

- Solution is also stochastic: $u = u(\boldsymbol{x}, t, \boldsymbol{y}(\omega))$

From stochastic to parametric solution

Finite dimensional uncertainty:

$$\boldsymbol{y}(\omega) = (y_1(\omega), \ldots, y_d(\omega)) \in \mathbb{R}^d, \quad d < \infty$$

$$y_i(\omega) : \Omega \to \Gamma_i \subseteq \mathbb{R}$$ independent with known distribution functions

Parametric solution:

$$u(\boldsymbol{x}, t, y_1, \ldots, y_d) : \bar{\mathcal{D}} \times [0, T] \times \prod_{i=1}^{d} \Gamma_i \longrightarrow \mathbb{R}$$

- A parametric problem in higher dimensions
- Challenge is when d is large
- Many ideas from high-dimensional function approximation apply

A wish list for complex systems

Random inputs

$\boldsymbol{y} \sim \mathcal{P}_{\boldsymbol{y}}$

Complex
PDE System

Random output

$u(\boldsymbol{y})$

An ideal approach:

- Sampling-based (non-intrusive)
  - Legacy codes
- Fewest possible simulations
- Fast convergence
- …

Key to success:

- Exploit solution structures
  - Anisotropy
  - Low-rank
  - Sparsity in some basis
  - …

Polynomial chaos approximation

[Ghanem & Spanos 91, Xiu & Karniadakis 02, …]

Multi-dimensional spectral approximation of finite-variance $u(\boldsymbol{y})$:

$$u_p(\boldsymbol{y}) = \sum_{i=1}^{P} c_i \psi_i(\boldsymbol{y}) \xrightarrow{m.s.} u(\boldsymbol{y}) \quad \text{as} \quad p \to \infty$$

Tensor-product basis: $\psi_i(\boldsymbol{y}) = \prod_{k=1}^{d} \phi_{i_k}(y_k), \quad i_1 + \cdots + i_d \leq p$

Number of basis: $P = \dfrac{(p+d)!}{p!d!}$

Ortho-normal basis: $\int \phi_{i_k}(y_k)\phi_{j_k}(y_k)\mathcal{P}_{y_k} dy_k = \delta_{ij} \longrightarrow \int \psi_i(\boldsymbol{y})\psi_j(\boldsymbol{y}) d\mathcal{P}_{\boldsymbol{y}} = \delta_{ij}$

Chaos coefficients: $c_i = \int u(\boldsymbol{y})\psi_i(\boldsymbol{y}) d\mathcal{P}_{\boldsymbol{y}}$

Askey scheme:

$y_k \sim$ uniform	$\phi_{i_k}(y_k)$	Legendre polynomials
$y_k \sim$ Gaussian	$\phi_{i_k}(y_k)$	Hermite polynomials

A bottleneck: Curse-of-dimensionality

- **+** Fast convergence: If $u(\boldsymbol{y})$ is sufficiently smooth w.r.t. $\boldsymbol{y}$
- **-** Number of unknown coefficients:

$$P = \frac{(p+d)!}{p!d!}$$

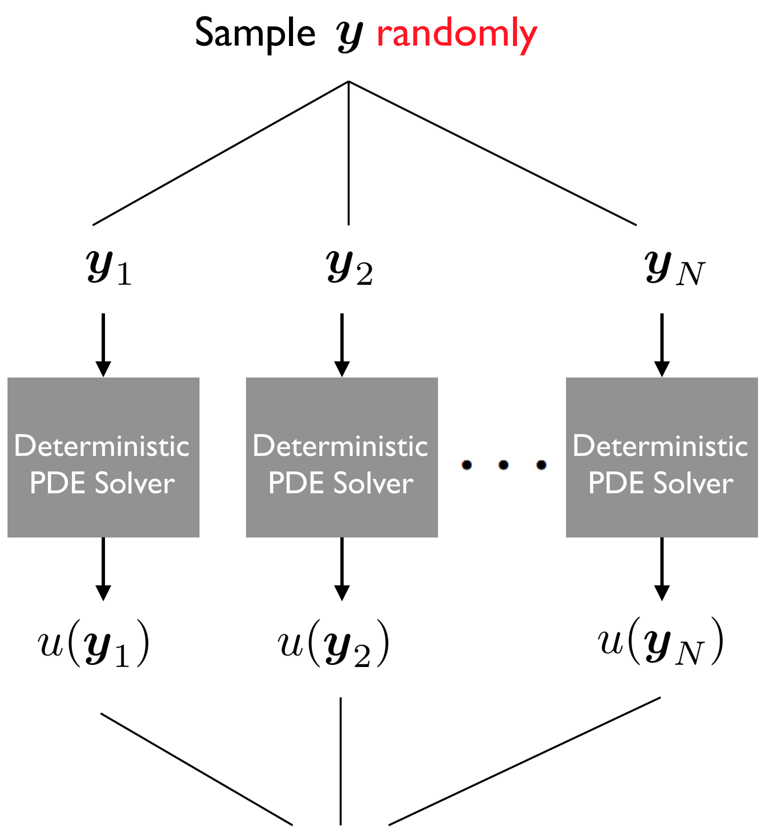
Exponential in d

$d = 40$	$p = 2$	$p = 3$	$p = 4$
P	861	$12,341$	$135,751$

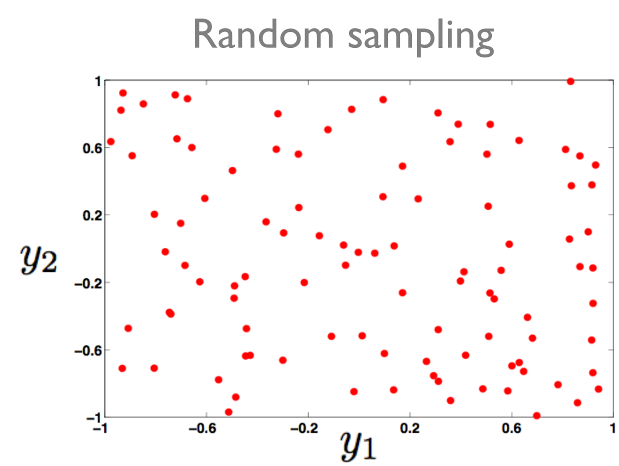
Curse-of-dimensionality: Exponential growth of computational complexity

- Intrusive (Galerkin projection) approaches
- Non-intrusive (sampling) approaches

Random sampling: Monte Carlo



$$c_i \approx \frac{1}{N} \sum_{n=1}^{N} u(\boldsymbol{y}_n)\psi_i(\boldsymbol{y}_n)$$



+ "No" curse-of-dimensionality

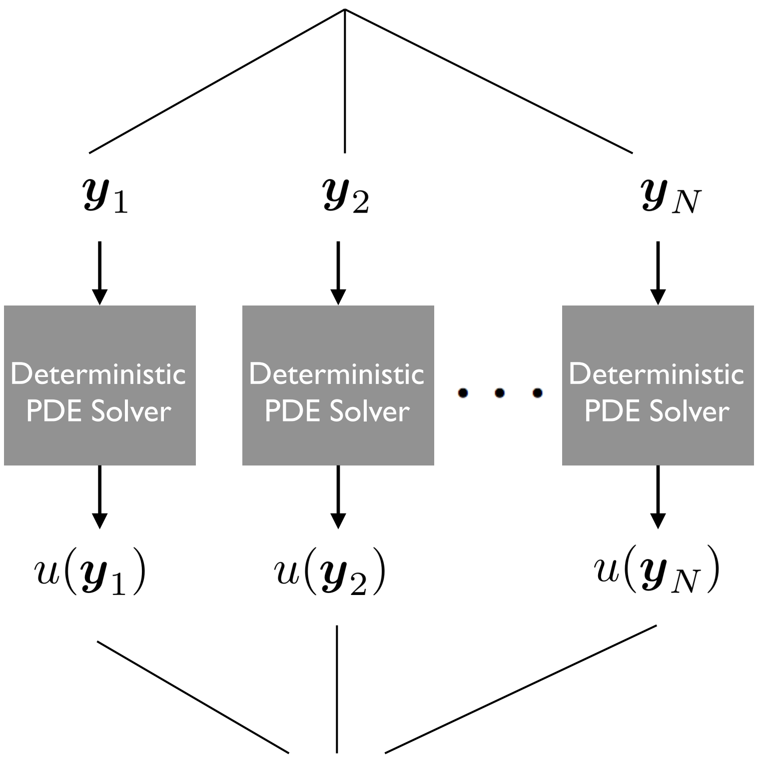
$$\mathcal{C}_d = \mathcal{O}(\mathcal{C}_1)$$

- Slow convergence

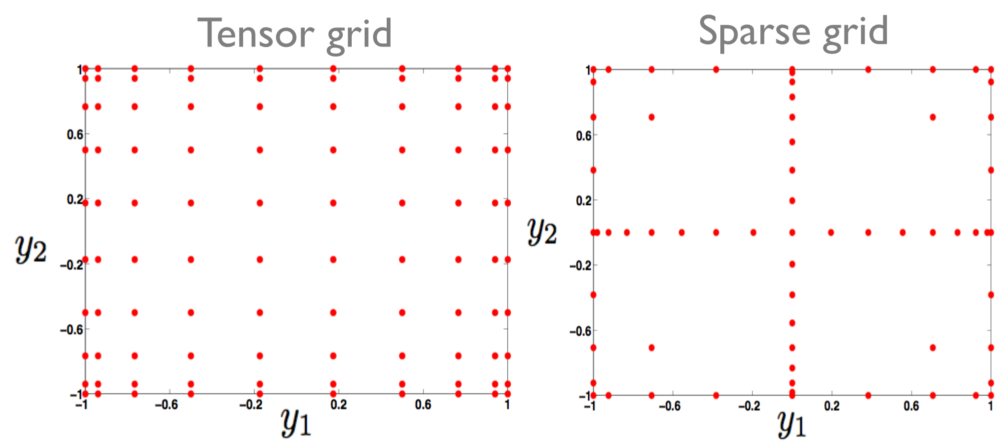
Pseudo-spectral on sparse grids – Stochastic collocation

[Xiu & Hesthaven 05, Babuska et al. 07, Ganapathysubramanian & Zabaras 07, …]

Sample $\boldsymbol{y}$ on sparse grids



$$c_i \approx \sum_{n \in \mathcal{G}} u(\boldsymbol{y}_n)\psi_i(\boldsymbol{y}_n)w_n$$



+ Reduces curse-of-dimensionality compare to tensor-product grids

- Curse of dimensionality remains

$$N \approx 2^p P, \quad d \gg 1$$

exact interpolation

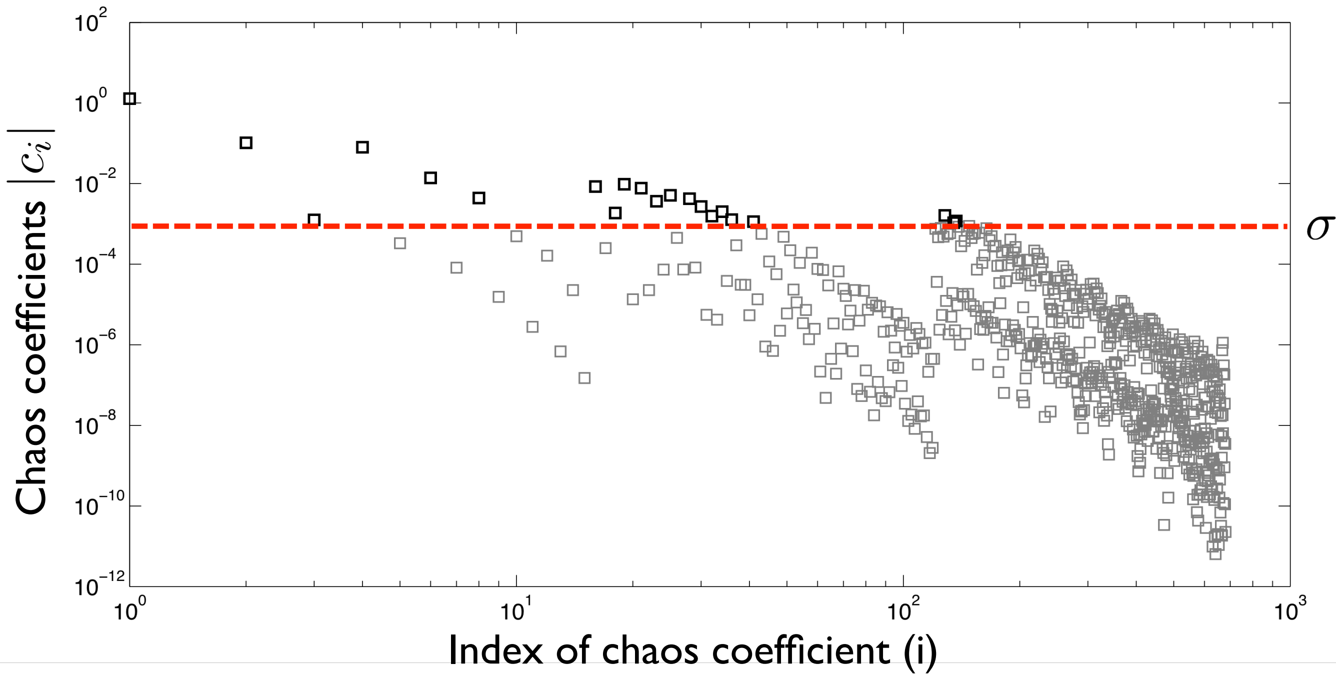
Question: What if many of the PC coefficients are negligible?

Sparsity of solution w.r.t. PC basis

$u(\boldsymbol{y})$ is sparse if it has a PC expansion with small number of important coefficients:

$$u_p(\boldsymbol{y}) = \sum_{i=1}^{P} c_i \psi_i(\boldsymbol{y})$$

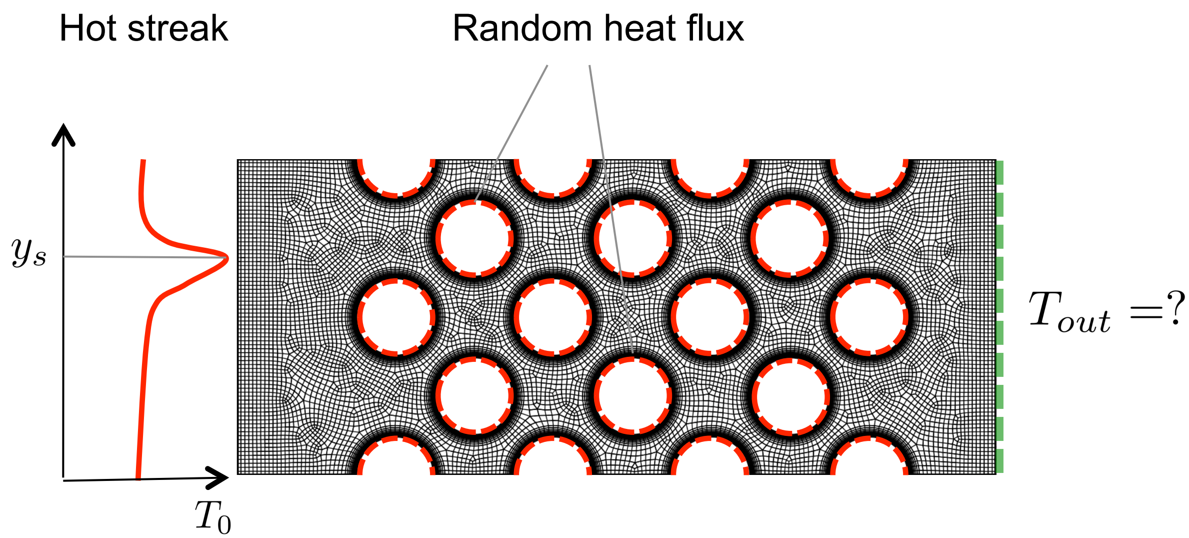
$$S = \#\{i : c_i > \sigma\} \ll P$$



The surprise:

A sparse solution can be approximated using $N \approx \alpha S \ll P$ samples!

Example – Heat transfer in a complex geometry



Reynolds-averaged Navier-Stokes:

$$
\begin{cases}
\dfrac{\partial U_i}{\partial x_i} = 0 \\
U_j \dfrac{\partial U_i}{\partial x_j} = \dfrac{\partial}{\partial x_j}\left[(\nu(T) + \nu_t)\dfrac{\partial U_i}{\partial x_j}\right] - \dfrac{1}{\rho}\dfrac{\partial P}{\partial x_i} \\
U_j \dfrac{\partial T}{\partial x_j} = \dfrac{\partial}{\partial x_j}\left[\left(\dfrac{\nu(T)}{P_r} + \dfrac{\nu_t}{P_{r_t}}\right)\dfrac{\partial U_i}{\partial x_j}\right]
\end{cases}
$$

Random B.C.'s

$Re = 500,000$

Sources of uncertainty:

- Heat flux on the cylinder wall (14 r.v's)

  $\frac{\partial T}{\partial n}|_{\Gamma_i} \sim U \quad i.i.d.$ c.o.v = %14.43

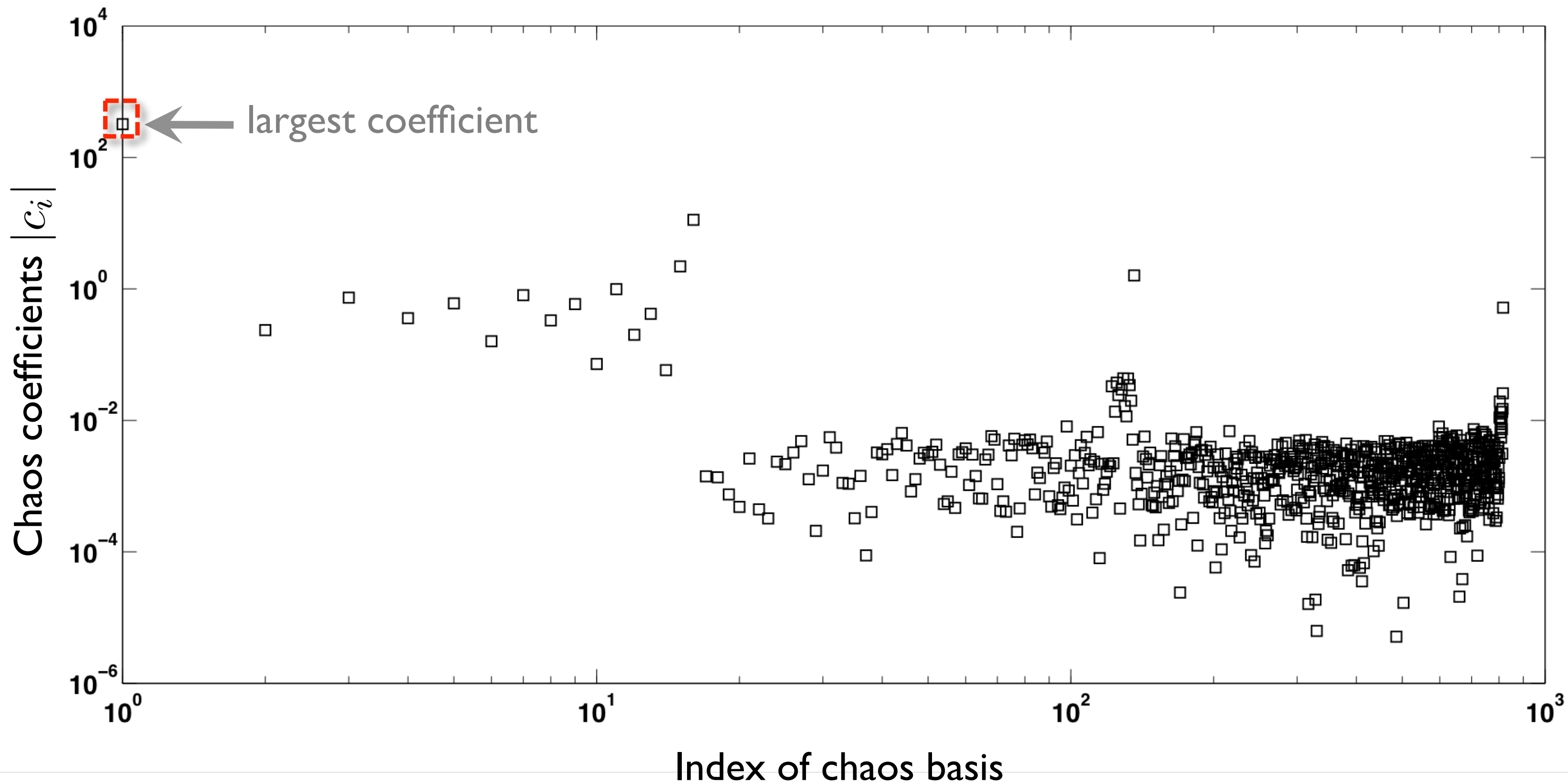
- Location of the hot streak at inflow (1 r.v.)

  $y_s \sim U$ c.o.v = %57.74

$d = 15$

Legendre PC expansion of temperature is sparse

Chaos coefficients of temperature at the outflow midpoint



$p = 3$ $\quad$ $d = 15$ $\quad$ $P = 861$ $\quad$ $S = 15$ $\quad$ Error in variance: $\mathcal{O}(10^{-4})$

A compressive sampling/sensing approach

Compressive sampling/sensing

- Geophysics
- Signal processing
- Imaging
- Statistics
- …

Our effort parallels works of:

- Claerbout
- Logan
- Donoho
- Candes
- Romberg
- Tao
- DeVore
- …

Problem setup – What are we after?

Given $N \ll P$ random samples (non-adapted):

$$\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_N \sim U[-1,1]^d$$

$$u(\boldsymbol{y}_1), u(\boldsymbol{y}_2), \ldots, u(\boldsymbol{y}_N)$$

reconstruct the S-sparse Legendre polynomial chaos expansion

$$u_p(\boldsymbol{y}) = \sum_{i=1}^{P} c_i \psi_i(\boldsymbol{y}) \qquad \|\boldsymbol{c}\|_0 = S \ll P \qquad u(\boldsymbol{y}) \in L_2([-1,1]^d)$$

Investigate approximation property:

- As $N \ll P$ stability/convergence?

Discrete representation: A matrix formulation

An underdetermined linear system:

$$
N \begin{bmatrix} \psi_1(\boldsymbol{y}_1) & \psi_2(\boldsymbol{y}_1) & \cdots & \psi_P(\boldsymbol{y}_1) \\ \vdots & \vdots & \vdots & \vdots \\ \psi_1(\boldsymbol{y}_N) & \psi_2(\boldsymbol{y}_N) & \cdots & \psi_P(\boldsymbol{y}_N) \end{bmatrix} \times \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_P \end{bmatrix} = \begin{bmatrix} u(\boldsymbol{y}_1) \\ u(\boldsymbol{y}_2) \\ \vdots \\ u(\boldsymbol{y}_N) \end{bmatrix} + \begin{bmatrix} \epsilon(\boldsymbol{y}_1) \\ \epsilon(\boldsymbol{y}_2) \\ \vdots \\ \epsilon(\boldsymbol{y}_N) \end{bmatrix} \qquad N \ll P
$$

P

truncation error

$$
\boldsymbol{\Psi} \times \boldsymbol{c} = \boldsymbol{u} + \boldsymbol{\epsilon}
$$

- $\|\epsilon\|_2 \leq \delta$
- δ has to be estimated (e.g. statistically)

Some observations:

- This is an ill-posed problem
- It has infinitely many solutions
- Requires further constrains on solution $\boldsymbol{c}$

But we know that $\boldsymbol{c}$ is sparse!

ℓ_0-minimization – Sparsest approximation

Main idea: Among all possible solutions find the one with minimum number of non-zeros:

$$(P_{0,\delta}) : \qquad \min_{\boldsymbol{c}} \|\boldsymbol{c}\|_0 \quad \text{subject to} \quad \|\boldsymbol{\Psi c} - \boldsymbol{u}\|_2 \leq \delta$$

where $\|\boldsymbol{c}\|_0 = \#\{i : c_i \neq 0\}$

- The solution is not always unique (for $\delta = 0$)!
- It is an NP-hard problem!

A heuristic:

- Convex relaxation via ℓ_1-minimization: Basis Pursuit Denoising (BPDN)

ℓ_1-minimization/Basis Pursuit Denoising (BPDN)

Main idea: Use the convex relaxation

$$(P_{1,\delta}): \qquad \min_{\boldsymbol{c}} \|\boldsymbol{c}\|_1 \quad \text{subject to} \quad \|\boldsymbol{\Psi}\boldsymbol{c} - \boldsymbol{u}\|_2 \leq \delta$$

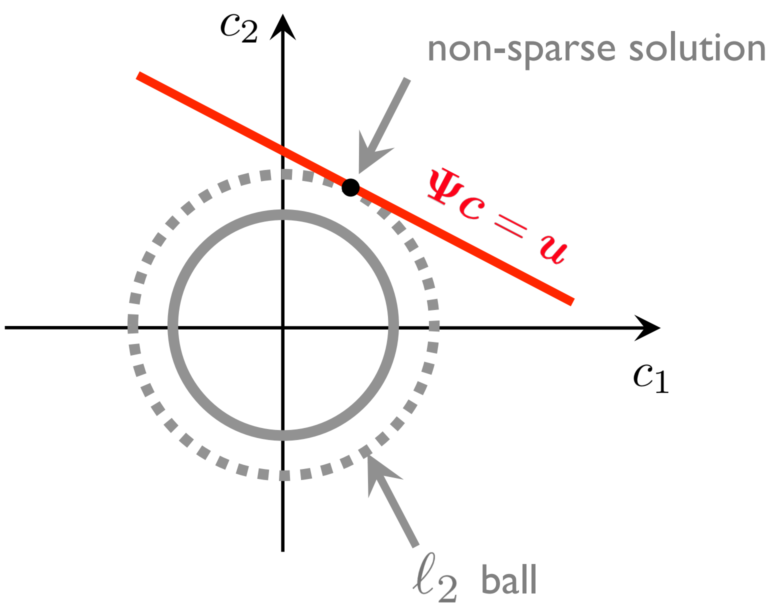
where $\|\boldsymbol{c}\|_1 = \sum_{i=1}^{P} |c_i|$

- For sufficiently sparse coefficients and with some conditions on $\boldsymbol{\Psi}$:
  - $(P_{1,\delta})$ and $(P_{0,\delta})$ share the same solution (for $\delta = 0$)
  - The solution is unique (for $\delta = 0$)
- Quadratic programming solvers:
  - Techniques such as: active set, projected gradient, interior-point continuation, etc.
  - In this work: SPGL1 with complexity $\mathcal{O}(P \ln P)$

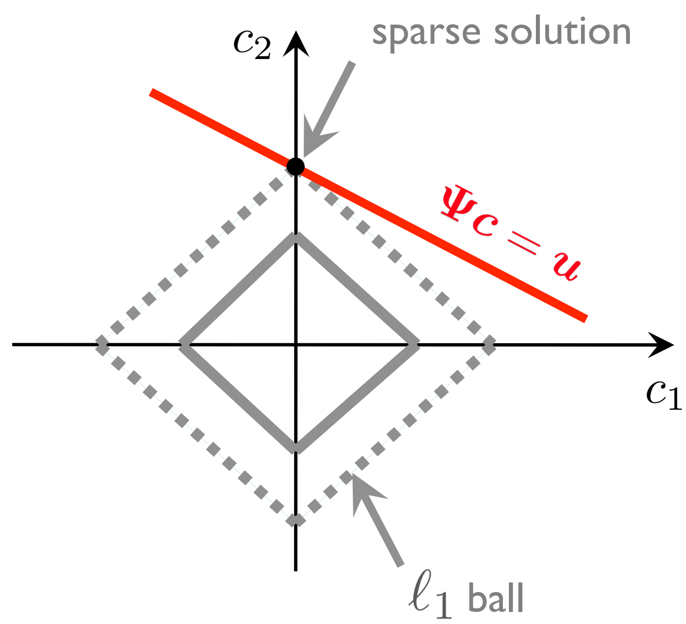
Why ℓ_1-norm promotes sparsity?

A geometric interpretation

minimum ℓ_2-norm solution



minimum ℓ_1-norm solution



Example – Elliptic stochastic differential equation

$$\begin{cases} -\nabla \cdot (a(x,\boldsymbol{y})\nabla u(x,\boldsymbol{y})) = f(x) & x \in [0,1] \\ u|_{x=0} = u|_{x=1} = 0 \end{cases}$$

Uncertain diffusion:

$$a(x,\boldsymbol{y}) = \bar{a}(x) + \sigma_a \sum_{k=1}^{d} \sqrt{\lambda_k}\phi_k(x) y_k$$

$y_k \quad i.i.d. \quad U[-1,1]$

Number of random variables: $d = 40$

$$C_{aa}(x_1,x_2) = \exp\left[-\frac{(x_1-x_2)^2}{l_c^2}\right]$$

$$l_c = 1/14 \quad \bar{a}(x) = 0.1 \quad \sigma_a = 0.021$$

Solution is sparse in Legendre chaos if:

- Covariance is piecewise analytic [Bieri & Schwab 09]
  - Smooth eigenfunctions
  - Fast decaying eigenvalues
  - e.g. Gaussian kernel

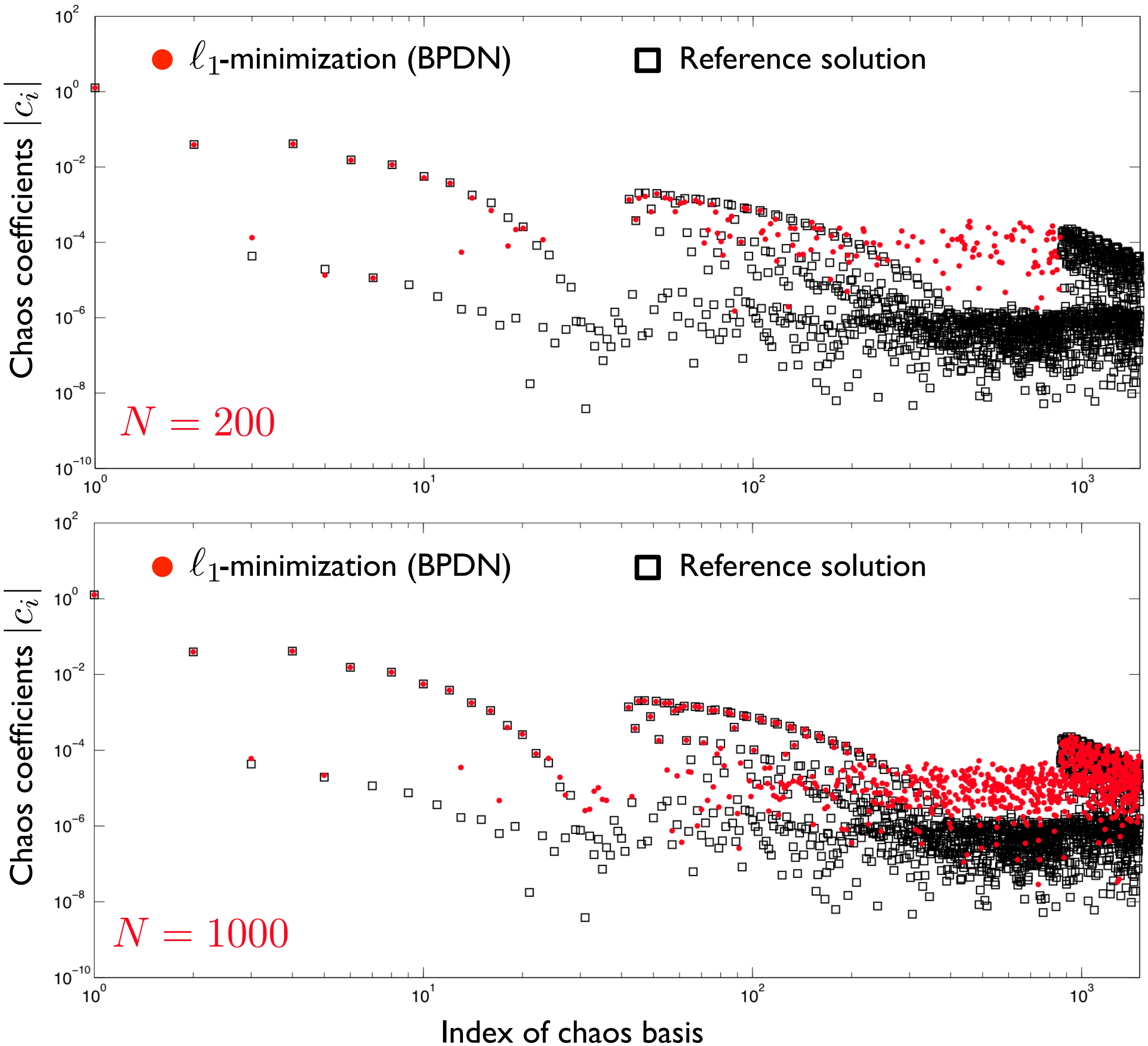
Approximation of chaos coefficients

At $x = 0.5$

$p = 3$

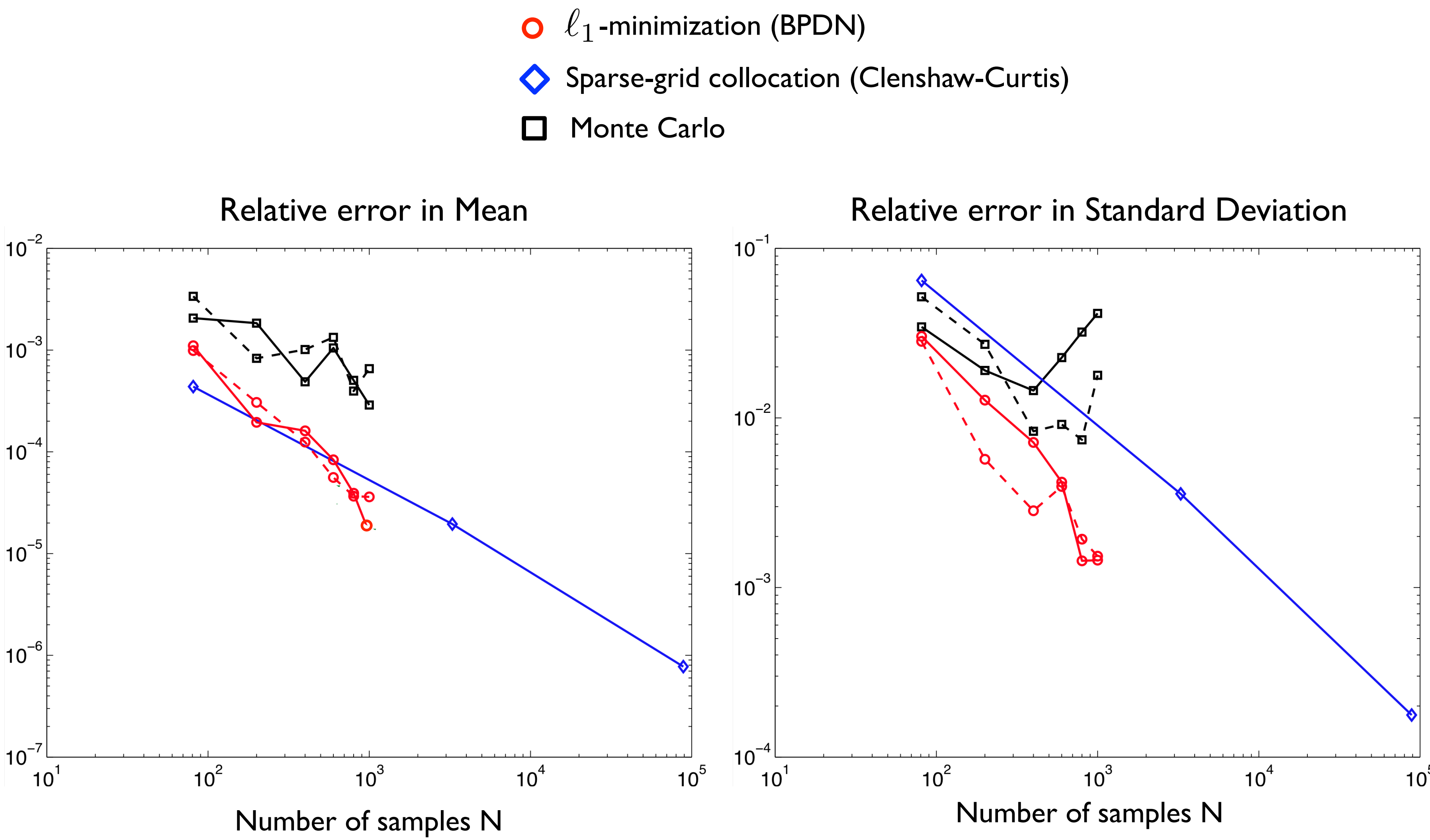
$d = 40$

$P = 1500$

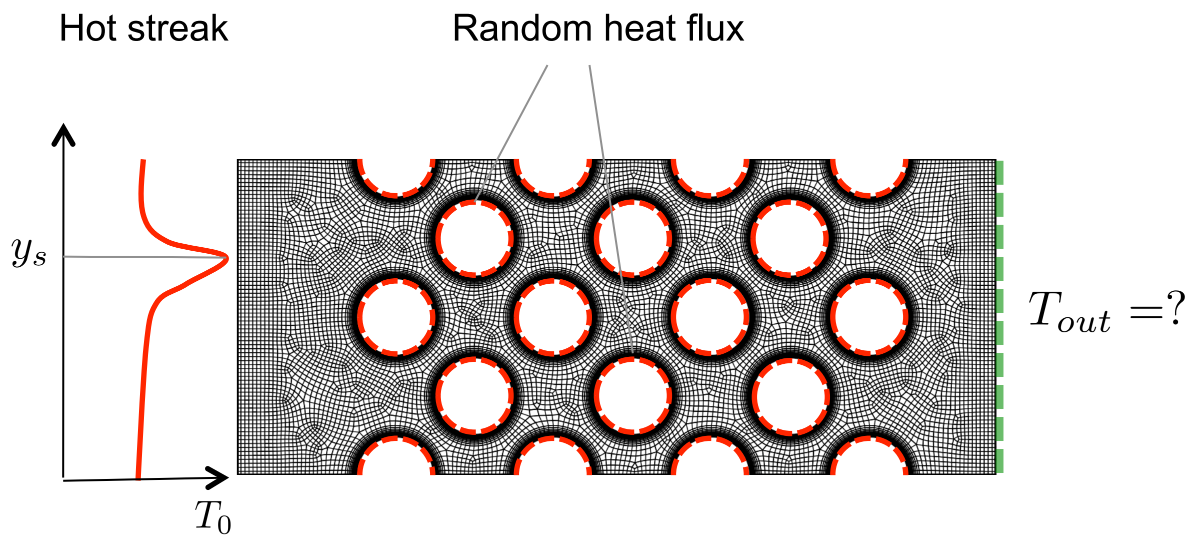


Convergence of solution statistics

At $x = 0.5$



Example – Heat transfer in a complex geometry



Reynolds-averaged Navier-Stokes:

$$\frac{\partial U_i}{\partial x_i} = 0$$

$$U_j \frac{\partial U_i}{\partial x_j} = \frac{\partial}{\partial x_j}\left[(\nu(T) + \nu_t)\frac{\partial U_i}{\partial x_j}\right] - \frac{1}{\rho}\frac{\partial P}{\partial x_i}$$

$$U_j \frac{\partial T}{\partial x_j} = \frac{\partial}{\partial x_j}\left[\left(\frac{\nu(T)}{P_r} + \frac{\nu_t}{P_{r_t}}\right)\frac{\partial U_i}{\partial x_j}\right]$$

Random B.C.'s

$Re = 500,000$

Sources of uncertainty:

- Heat flux on the cylinder wall (14 r.v's)

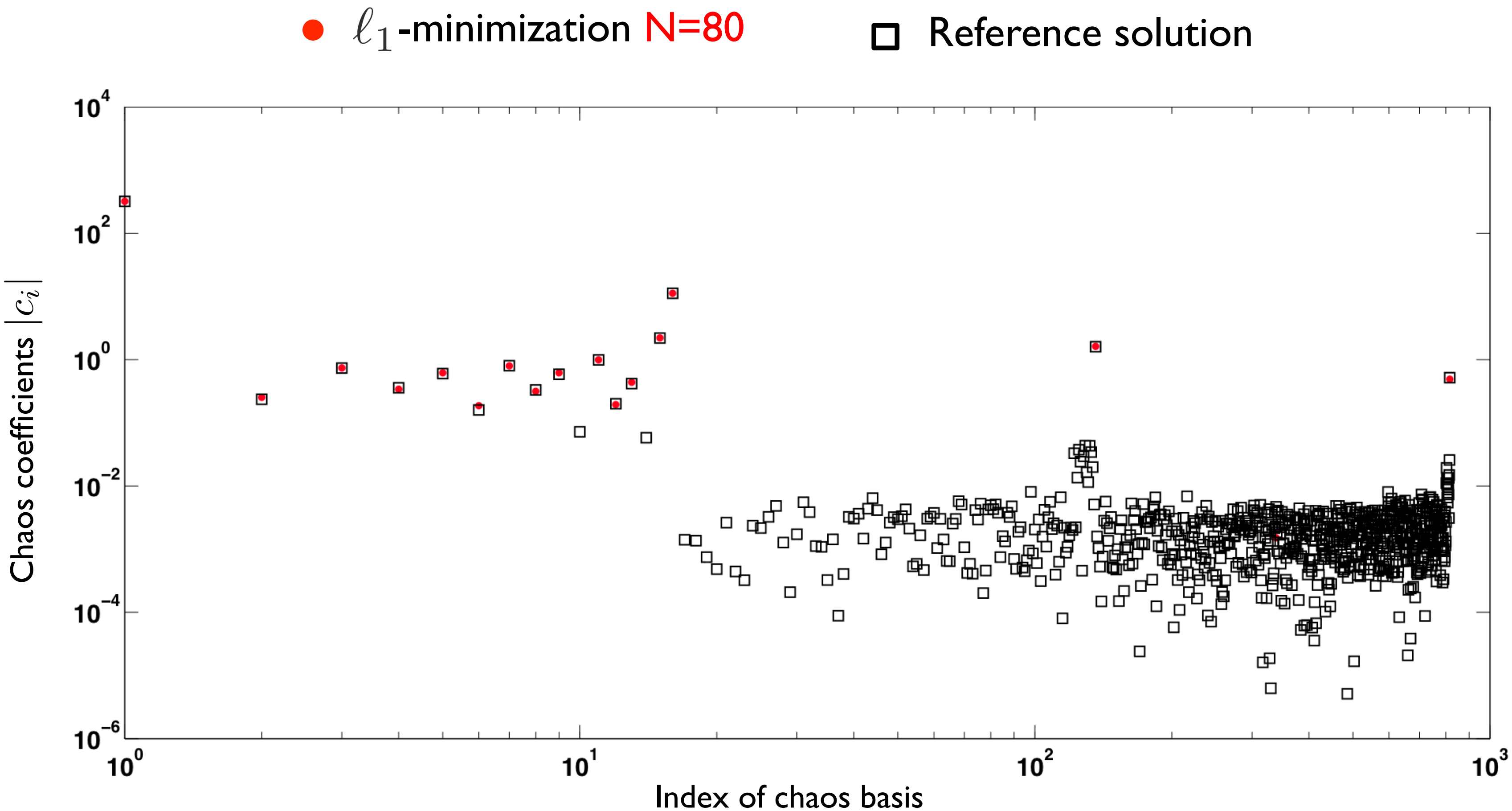
  $\frac{\partial T}{\partial n}|_{\Gamma_i} \sim U \quad i.i.d.$ c.o.v = %14.43

- Location of the hot streak at inflow (1 r.v.)

  $y_s \sim U$ c.o.v = %57.74

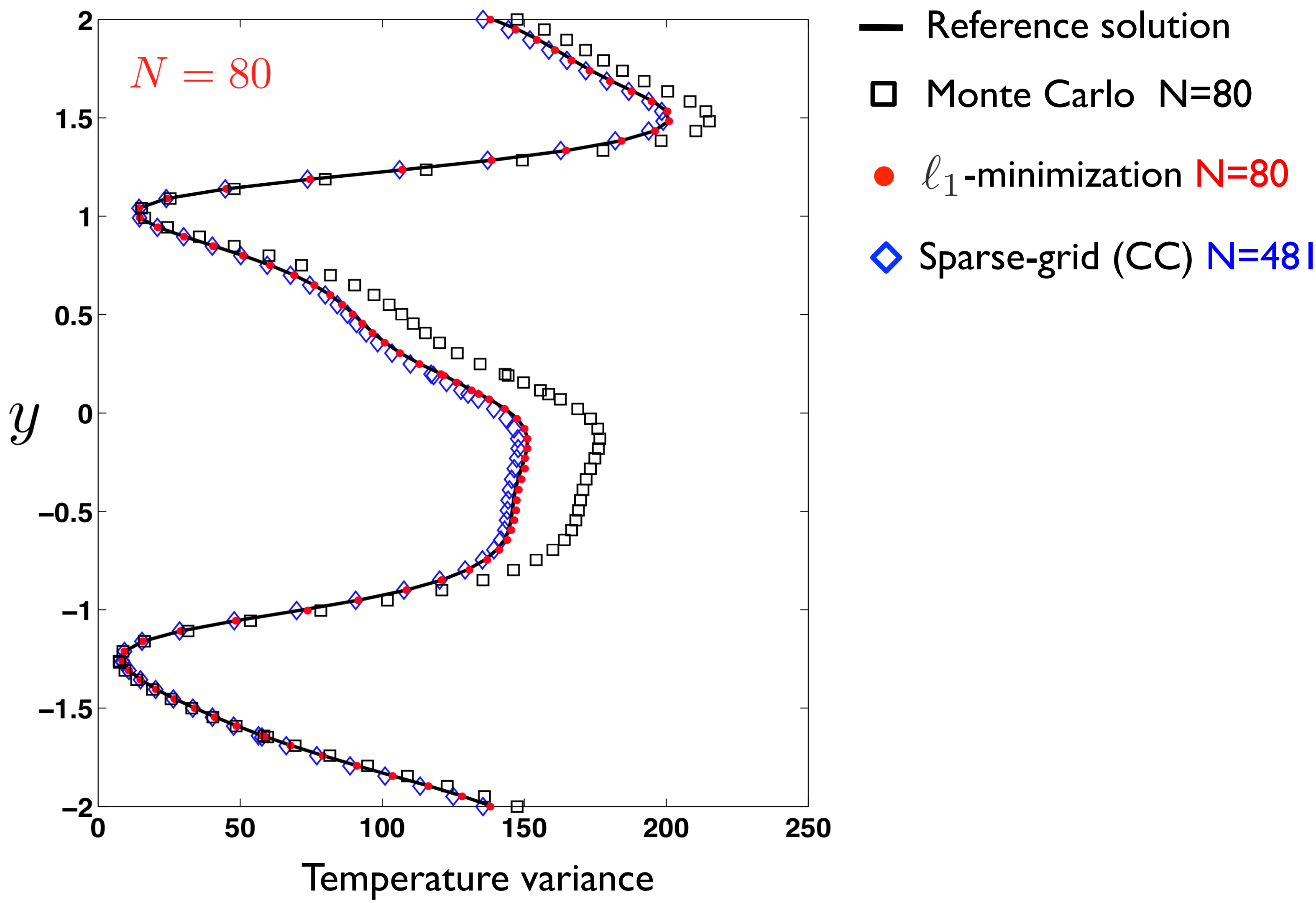
$d = 15$

Chaos coefficients of temperature at the outflow midpoint



$p = 3$ $\quad d = 15$ $\quad P = 861$ $\quad S = 15$ $\quad N = 80$

Convergence of the outflow temperature variance



Ingredients of a successful compressive sampling

1 When columns of $\mathbf{\Psi}$ are "nearly" orthogonal = small mutual coherence:

[Donoho et al. 06]

$$\mu(\mathbf{\Psi}) := \max_{j \neq k} \frac{|\boldsymbol{\psi}_j^T \boldsymbol{\psi}_k|}{\|\boldsymbol{\psi}_j\|_2 \|\boldsymbol{\psi}_k\|_2}$$

$$\mathbf{\Psi} = \begin{array}{c} \left[\begin{array}{ccccc} \cdots & \boxed{\begin{array}{c} \psi_j(\boldsymbol{y}_1) \\ \vdots \\ \psi_j(\boldsymbol{y}_N) \end{array}} & \cdots & \boxed{\begin{array}{c} \psi_k(\boldsymbol{y}_1) \\ \vdots \\ \psi_k(\boldsymbol{y}_N) \end{array}} & \cdots \\ \vdots & & \vdots & & \vdots \\ \cdots & & \cdots & & \cdots \end{array} \right] \\ \qquad\quad \boldsymbol{\psi}_j \qquad\qquad\quad \boldsymbol{\psi}_k \end{array}$$

2 When coefficient vector is "sufficiently" sparse: [Donoho et al. 06]

$$\|\boldsymbol{c}\|_0 < (1 + 1/\mu(\mathbf{\Psi}))\,/4$$

This is a pessimistic bound!

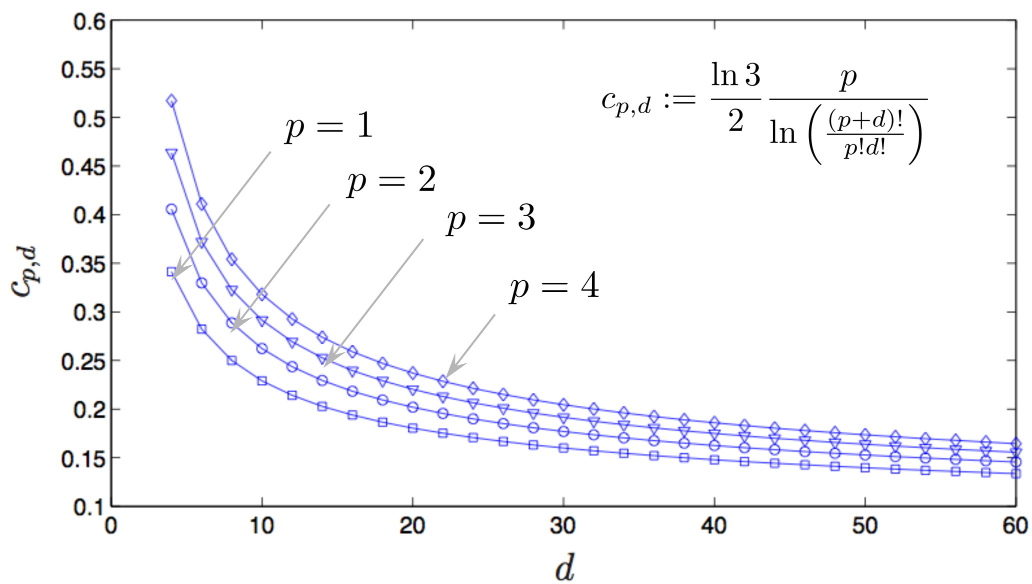
Bound on mutual coherence - Legendre PC basis

Theorem: [Doostan et al., 09; Doostan & Owhadi, 11]

As a result of the concentration of measure phenomenon on empirical correlation of PC basis:

$$Prob\left[\mu(\mathbf{\Psi}) \geq \frac{r}{1-r}\right] \leq 4P^{2-2\zeta}$$

$$0 \leq r^2 = \frac{4\zeta P^{4c_{p,d}}(\ln P)}{N} \leq \frac{1}{4} \qquad \zeta > 1$$



ℓ_1-minimization is stable for Legendre PC expansion

Theorem (General stability of BPDN): [Doostan & Owhadi, 11]

Let $u(\boldsymbol{y})$ be any essentially bounded function of i.i.d. random variables $\boldsymbol{y} = (y_1, \cdots, y_d)$ uniformly distributed on $\Gamma := [-1,1]^d$. Assume there exists:

$$u_p^0(\boldsymbol{y}) = \sum_{i=1}^{P} c_i^0 \psi_i(\boldsymbol{y}) \quad \text{such that:} \quad \begin{cases} \|\boldsymbol{c}^0\|_0 = S \\ \left\|u - u_p^0\right\|_{L^\infty(\Gamma)} \le \epsilon \end{cases}$$

Then using

Spars-grid

$$\boxed{N \ge c_1 P^{4c_{p,d}} (\ln P) S} \longleftrightarrow \boxed{N \approx 2^p P}$$

random realization of solution:

$$\boxed{\left\|u - u_p^{1,\delta}\right\|_{L^2(\Gamma)} \le c_2 \epsilon + c_3 \frac{\delta}{\sqrt{N}}}$$

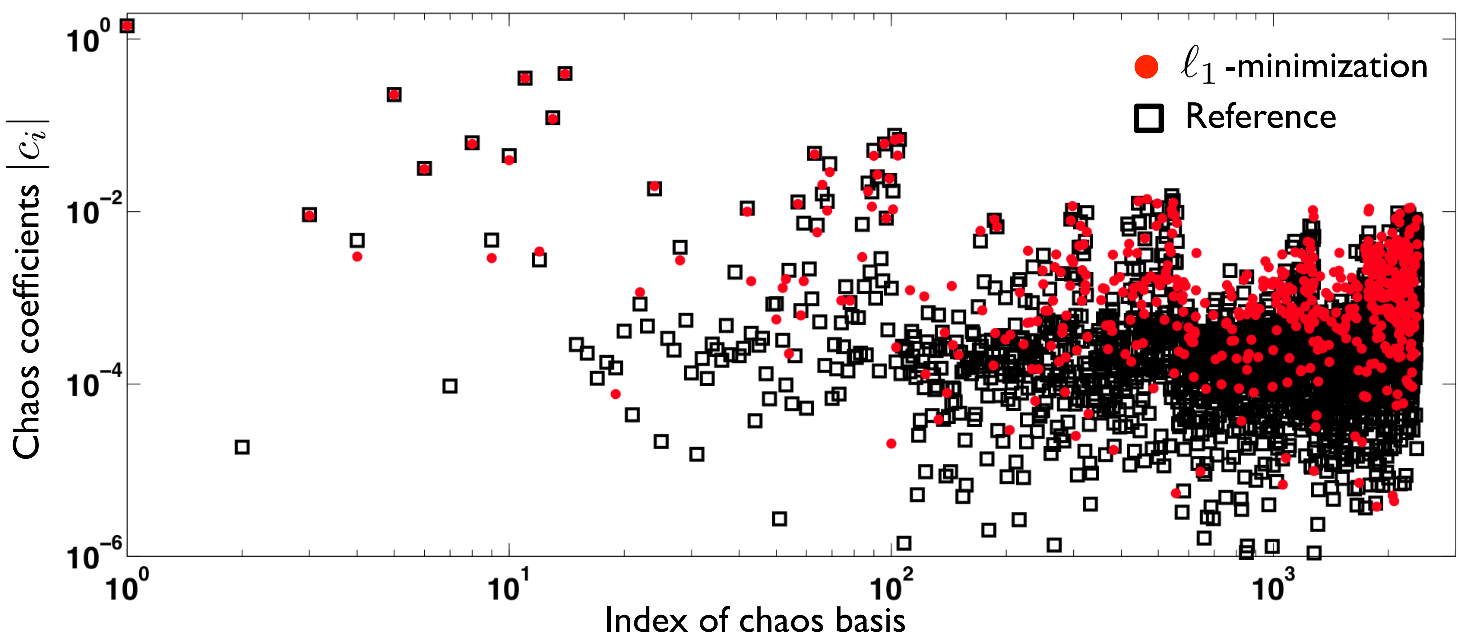
with overwhelming probability.

Hydrogen Oxidation in Supercritical Water

Effect of parametric uncertainties on species concentration

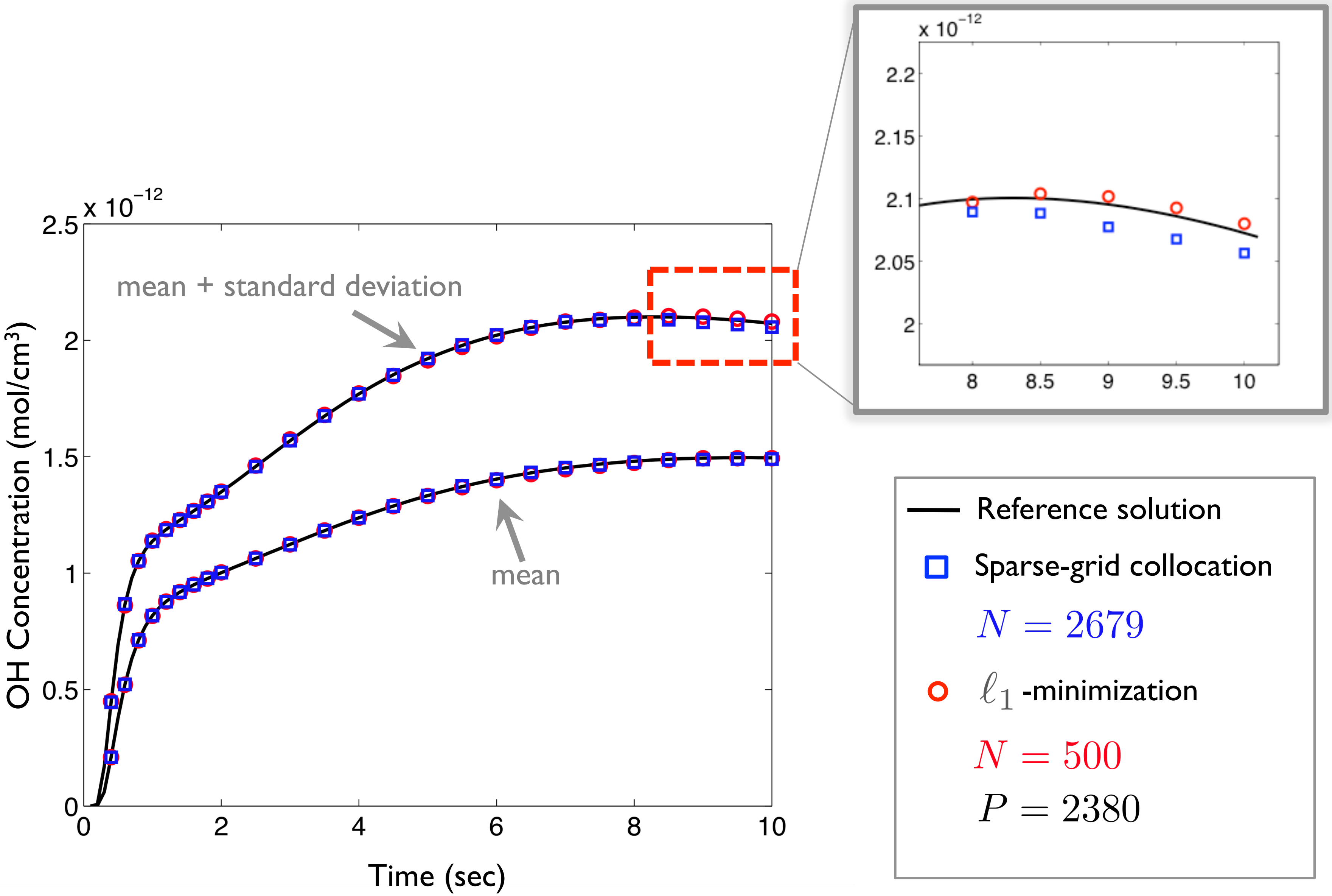
- System of stiff nonlinear ODEs
- Uncertain reaction rates: 8 independent lognormals
- Uncertain enthalpies of formation: 5 independent Gaussians
- Prior work: [Phenix et al. 98, Reagan et al. 02/04, Le Maitre et al. 04/07, Najm et al. 09, Alexanderian et al. 11]

Hermite expansion of OH concentration ($t = 6.5\ sec$)



$$d = 13,\ p = 4,\ P = 2380,\ N = 500$$

Statistics of OH concentration



Outlook

Design of sampling strategy:

- **Question:** How to optimally choose $\{\boldsymbol{y}_i\}_{i=1}^{N}$ for a given N ?
- **A possibility:** Bayesian formulation of compressive sampling ?

  [Tipping 01, Ji et al. 07, …]

  - Gaussian truncation error
  - Laplace prior
  - MAP equivalent to ℓ_1-minimization
  - New samples to minimize posterior uncertainty

Non-smooth solutions:

- Sharp gradients/discontinuities
- Sparse approximation in multi-wavelet basis
  - Adaptive sampling strategies

Multi-physics/Multi-scale applications:

- Lithium batteries as a test bed

Acknowledgment

Ph.D. students:

- Ji Peng and Ahad Validi (CU Boulder)

Collaborators:

- Houman Owhadi (Caltech)
- Gianluca Iaccarino (Stanford)
- Kurt Maute (CU Boulder)

Financial support:

- DOE office of Advanced Scientific Computing Research
- Stanford PSAAP center

Thank You!