Error Budgets: A Path from Uncertainty Quantification to Model Validation

Roger Ghanem

Aerospace and Mechanical Engineering
Civil Engineering
University of Southern California
Los Angeles
Acknowledgments

Funding from:
Sandia, ONR, AFOSR, NSF, DARPA, Taisei, Shimizu, ARAMCO

Intellectual contributions:
Graduate students at Buffalo, Hopkins and USC
John Red-Horse - SNL
Habib Najm - SNL
Bert Debusschere - SNL
Christian Soize - UMLV, France
OUTLINE

- Motivation
- Mathematical structure for validation and verification
- Packaging of information using this structure
- Construction of approximation spaces on these structures
- Examples
- Technical challenges and conclusions
A simple prelude

• Quantity of interest: decision with upper bound on risk
  – Flip coin 100 times ➞ calibrated decision tool ➞ 50-50
  – ⊕ observe initial configuration ➞ calibrated decision tool ➞ 70-30
  – ⊕ observe surface tension ➞ calibrated decision tool ➞ 80-20
    --- ➞ --- ➞ ---
  – ⊕ observe quantum states ➞ calibrated decision tool ➞ 99-1

• Use Model-Based Predictions in lieu of physical experiments

Ingredients:
  ➢ calibrate a stochastic plant
  ➢ evaluate limit on predictability of plant
  ➢ refine plant if target confidence not achievable
Interaction of model and data
Interaction of model and data
Interaction of model and data
Interaction of model and data
Interaction of model and data
Error budget

\[ U = \hat{U}|_{h,d,p,m} + \varepsilon_h|_{p,d,m} + \varepsilon_p|_{d,m} + \varepsilon_d|_m + \varepsilon_m \]

\[ \varepsilon_h|_{d,m} : \text{can be reduced through better numerics.} \]

\[ \varepsilon_p|_{d,m} : \text{can be reduced through better statistics.} \]

\[ \varepsilon_d|_m : \text{can be reduced through better data.} \]

\[ \varepsilon_m : \text{can be reduced through better models.} \]
We use probabilistic models for uncertainty

The probabilistic framework provides a packaging of information that is amenable to a level of rigor in analysis that permits the “quantification” of uncertainty.

Although:
“... unacquainted with problems where wrong results could be attributed to failure to use measure theory....” (E.T. Jaynes, published 2003)

An experiment is defined by a probability triple, (a measurable space) \((\mathcal{A}, \mathcal{T}, \mu)\)

- \(\mathcal{A}\) Set of elementary events
- \(\mathcal{T}\) Sigma algebra of all events that make sense
- \(\mu\) Measure on all elements of \(\mathcal{T}\)

A random variable \(X\) is a measurable mapping on a probability space \((\mathcal{A}, \mathcal{T}, \mu)\) to \((R, \mathcal{B})\)

The probability distribution \(P\) of \(X\) is the image of \(\mu\) under \(X\), defining the probability triple \((R, \mathcal{B}, P)\)
What do we mean by things being close?

Modes of Convergence: Let $X_n$ be a sequence of r.v. defined on $(\mathcal{A}, \mathcal{T}, \mu)$ and let $Y$ be another r.v on same probability space:

- **Almost sure Convergence:**
  \[ P[|X_n - Y| > \epsilon] \to 0 \quad \forall \epsilon \]
  $X_n(\omega)$ converges a.e. to $Y(\omega)$.

- **Convergence in mean square:**
  \[ \|X_n - Y\|_{L^2} \to 0 \]

- **Convergence in mean:**
  \[ \|X_n - Y\|_{L^1} \to 0 \]

- **Convergence in probability:**
  \[ P[\rho(X_n, Y) < \epsilon] \to 1 \quad \forall \epsilon \]

- **Convergence in distribution:**
  Let $X_n \in L^0(\mathcal{A}_n, \mathcal{T}_n, \mu_n)$ and $Y \in L^0(\mathcal{A}, \mathcal{T}, \mu)$

  then $X$ converges in distribution to $Y$ if $P_n$ converges to $P$.

Convergence in mean-square permits an $L_2$ analysis of random variables.

First step in a unified perspective in verification and validation.
Characterize $Y$, given characterization of $Z^1, \ldots, Z^p$.

$$Y = \tilde{f}(Z^1, \ldots, Z^p)$$

Each $Z^j$ is a vector-valued random variable:

$$Z^j : (A, \mathcal{T}, P) \rightarrow \mathbb{R}^{m_j}$$

Orthogonalizing the random variables:

$$C_{Z^j} = L_{Z^j}^T L_{Z^j}$$

$$Z^j = m_{Z^j} + L_{Z^j}^T X^j$$

Then:

$$y = f(x^1, \ldots, x^p) : \mathbb{R}^{m_1} \times \ldots \times \mathbb{R}^{m_p} \rightarrow \mathbb{R}^m$$
Assumptions:

- Second-order random variables:
  \[ E \left\{ \| f(X^1, \ldots, X^p) \|^2 \right\} < +\infty \]

- Basic random vectors are independent:
  \[ P_{X^1, \ldots, X^p} = P_{X^1} \otimes \ldots \otimes P_{X^p} \]

- Vector \( X^j \) does not necessarily involve independent random variables:
  \[ P_{X^i} \neq \prod_{k=1}^{P} P_{X^j_k} \]
Functional representations

$$H^{(m)} = L^2_{P_{X^1,\ldots,X^p}} (\mathbb{R}^{m_1} \times \ldots \times \mathbb{R}^{m_p}, \mathbb{C}^m)$$

$$\cong \left( L^2_{P_{X^1,\ldots,X^p}} (\mathbb{R}^{m_1} \times \ldots \times \mathbb{R}^{m_p}, \mathbb{R}) \right) \otimes \mathbb{C}^m$$

$$\cong \left( \bigotimes_{j=1}^p H_j \right) \otimes \mathbb{C}^m$$

$$\cong H \otimes \mathbb{C}^m$$

$$H_{j,k} = L^2_{P_{X^j}} (\mathbb{R}, \mathbb{R})$$

$$H_j = \bigotimes_{k=1}^{m_j} H_{j,k}$$

Given bases of $H_{j,k}$ other bases can be constructed.
Functional representations

For all $x^j = (x_1^j, \ldots, x_{m_j}^j)$ belonging to the support $S_{m_j}$ of $p_{X_j}(x^j)$, Hilbertian basis $\{\phi_{\alpha_j}^j, \alpha_j^j \in \mathbb{N}^{m_j}\}$ of real Hilbert space $\mathbb{H}_j$ is given by

$$\phi_{\alpha_j}^j(x^j) = \left( \frac{p_{X_1^j}(x_1^j) \times \ldots \times p_{X_{m_j}^j}(x_{m_j}^j)}{p_{X_j}(x^j)} \right)^{1/2} \psi_{\alpha_1^j}^1(x_1^j) \times \ldots \times \psi_{\alpha_{m_j}^j}^{m_j}(x_{m_j}^j),$$

where $\{\psi_{\alpha_k^j}(x_k^j)\}_{\alpha_k^j}$ is a Hilbertian basis of real Hilbert space $\mathbb{H}_{j,k}$ and $\alpha^j = (\alpha_1^j, \ldots, \alpha_{m_j}^j)$. 
Some common bases

Infinite-dimensional case:
This is an exercise in stochastic analysis:

• Hermite polynomials: Gaussian measure (Wiener: Homogeneous Chaos)
• Charlier polynomials: Poisson measure (Wiener: Discrete Chaos)
• Very few extensions possible: Friedrichs and Shapiro (Integration of functionals) provide characterization of compatible measures. Segall and Kailath provide an extension to martingales.

Finite-dimensional case: independent variables:
This is an exercise in one-dimensional approximation:

• Askey polynomials: measures from Askey chart (Karniadakis and co-workers)
• Legendre polynomials: uniform measure (theoretical results by Babuska and co-workers)
• Wavelets: Le-Maitre and co-workers
• Arbitrary measures with bounded support: C. Schwab

Finite-dimensional case: dependent variables:
This is an exercise in multi-dimensional approximation:

• Arbitrary measures: Soize and Ghanem.
Another special basis: Karhunen-Loeve expansion

A zero-mean stochastic process

\[
\lim_{N \to \infty} \mathbb{E} \left\| \mathbf{U}(x, \theta) - \sum_{i=1}^{N} \xi_i(\theta) \sqrt{\lambda_i} \phi_i(x) \right\|_{L_2} = 0 \quad \mu = \sum_{i=1}^{\infty} \lambda_i < \infty
\]

where

\[ R_U v = \lambda v, \]

is the eigen-problem associated with the covariance function of the process \( \mathbf{U}(x, \theta) \).

The Karhunen-Loeve expansion:

\[
\mathbf{U}(x, \theta) = \sum_{i=1}^{N} \xi_i(\theta) \sqrt{\lambda_i} \phi_i(x) = \sum_{i=1}^{N} (\mathbf{U}, \phi_i) K \phi_i(x) = \sum_{i=1}^{N} (\mathbf{U}, \xi_i) \xi_i(\theta)
\]

If the covariance kernel of \( \mathbf{U} \) has \( 2r \) continuous symmetric derivative and is \( \alpha \)-Lipschitz, then

\[ \lambda_n = O \left( \frac{1}{n^{2r+1+\alpha}} \right) \]

Reference: J.B. Read (1983)
Management of uncertainty
Management of uncertainty

COORDINATES IN THIS SPACE REPRESENT PROBABILISTIC CONTENT.

SENSITIVITY OF PROBABILISTIC STATEMENTS OF BEHAVIOR ON DATA.
The random quantities are resolved as

$$\alpha(x, \theta) = \sum_{i=1}^{\infty} \alpha_i(x) \psi_i(\{\xi(\theta)\})$$

These could be, for example:

- Parameters in a PDE
- Boundaries in a PDE (e.g. Geometry)
- Field Variable in a PDE

Multidimensional Orthogonal Polynomials in independent random variables

$$\begin{align*}
\psi_0 &= 1 \\
\psi_1 &= \xi_1 \\
\psi_2 &= \xi_2 \\
\psi_3 &= \xi_1^2 - 1 \\
\psi_4 &= \xi_1 \xi_2 \\
\psi_5 &= \xi_2^2 - 1 \\
&\vdots
\end{align*}$$

Arbitrary measures:

- Karniadakis
- Babuska
- LeMaitre/Ghanem
- Soize/Ghanem
The random quantities are resolved as

$$\alpha(x, \theta) = \sum_{i=1}^{\infty} \alpha_i(x) \psi_i(\theta) \quad \alpha_i(x) = \frac{\langle \alpha(x, \theta), \psi_i(\theta) \rangle_{L^2(\Omega)}}{\| \psi_i(\theta) \|_{L^2(\Omega)}^2}$$

These could be, for example:

- Parameters in a PDE
- Boundaries in a PDE (e.g. Geometry)
- Field Variable in a PDE

These decompositions provide a resolution (or parameterization) of the uncertainty on spatial or temporal scales
Representation of uncertainty: Maximum likelihood

The random quantities are resolved as:

\[ \alpha(x, \theta) = \sum_{|i| \leq q} \alpha_i(x) \psi_i(\theta) \]

\[ KL : \eta_i(\theta) = \frac{1}{\sqrt{\lambda_i}} (\alpha(., \theta), \phi_i)_H \]

Let

\[ A = \{ \alpha_i(x), \ |i| \leq q \} \]

Likelihood Function is:

\[ L(\alpha^1, \ldots, \alpha^M; A) = p_{\alpha^1(x), \ldots, \alpha^M(x)}(\alpha^1, \ldots, \alpha^m; A) \times \ldots \times p_{\alpha^1(x), \ldots, \alpha^M(x)}(\alpha^1, \ldots, \alpha^M; A) \]

\[ L(\eta^1, \ldots, \eta^m; A) = p_{\eta_1, \ldots, \eta_m}(\eta^1, \ldots, \eta^m; A) \times \ldots \times p_{\eta_1, \ldots, \eta_m}(\eta^1, \ldots, \eta^m; A) \]

\[ \approx \prod_{j=1}^{\nu} p_{\eta_j}(\eta_j^1; A) \times \ldots \times \prod_{j=1}^{\nu} p_{\eta_j}(\eta_j^m; A) \]

Maximum Likelihood:

\[ \max_A L \]
“essential” dimensionality of a process

Physical object: Linear Elasticity

Convergence as function of “dimensionality”

Stochastic parameters
Representation of uncertainty: Bayesian inference

Covariance matrix of observations

\[
\hat{C} = \frac{1}{M-1} \sum_{i=1}^{M} (a_i - \bar{a})^T (a_i - \bar{a})
\]

Reduced order representation: KL:

\[
a(\omega) \approx \bar{a} + \sum_{i=1}^{\mu} \sqrt{\lambda_i} \eta^{(i)}(\omega) \phi_i
\]

Where:

\[
\begin{align*}
\eta^{(i)}_j &= \frac{1}{\sqrt{\lambda_i}} \langle a_j, \phi_i \rangle_{L_2} \\
E[\eta^{(n)}] &= 0 \quad n, m = 1, \ldots, \mu. \\
E[\eta^{(m)}\eta^{(n)}] &= \delta_{mn}
\end{align*}
\]
Representation of uncertainty: Bayesian inference

Objective is to estimate \( \{ \eta_i(\theta) \} \)

Polynomial Chaos representation of reduced variables:

\[
\hat{\eta}^{(i)} = \sum_{j=1}^{\infty} \gamma_j^{(i)} \Psi_j(\xi_i)
\]

\[
\approx \sum_{j=1}^{p} \gamma_j^{(i)} \Psi_j(\xi_i) \quad i = 1, \ldots, \mu
\]

Constraint on chaos coefficients:

\[
\sum_{j=1}^{p} (\gamma_j^{(i)})^2 = 1 \quad i = 1, \ldots, \mu.
\]

Estimation of stochastic process using estimate of reduced variables:

\[
\hat{a}(\omega) = \bar{a} + \sum_{i=1}^{\mu} \sqrt{\lambda_i} \hat{\eta}^{(i)}(\omega) \phi_i.
\]
Define Cost Function (hats denote estimators):

\[
L[\gamma, \hat{\gamma}] = \begin{cases} 
1 & \text{if } \max_{1 \leq i \leq p} |\gamma_i^{(k)} - \hat{\gamma}_i^{(k)}| > \Delta \\
0 & \text{if } \max_{1 \leq i \leq p} |\gamma_i^{(k)} - \hat{\gamma}_i^{(k)}| \leq \Delta,
\end{cases}
\]

Then Bayes estimate is:

\[
\hat{\gamma} = \arg \min_{\zeta} E \{L[\gamma, \zeta(\hat{\eta}_1^{(k)}, \ldots, \hat{\eta}_M^{(k)})]\}
\]

\[
\approx \arg \min_{\zeta} E \{L[\gamma, \zeta(\eta_1^{(k)}, \ldots, \eta_M^{(k)})]\}.
\]

Bayes rule:

\[
\pi[\gamma | \eta_1^{(k)}, \ldots, \eta_M^{(k)}] \propto P(\hat{\eta}_1^{(k)}, \ldots, \hat{\eta}_M^{(k)} | \gamma) \times \pi(\gamma)
\]

- Use kernel density estimation to represent the Likelihood function
- Use Markov Chain Monte Carlo to sample from the posterior (metropolis Hastings algorithm) \(\rightarrow\) BIMH
Representation of uncertainty: Bayesian inference
Representation of uncertainty: Bayesian inference
Representation of uncertainty: Bayesian inference
Representation of uncertainty: Bayesian inference
Representation of uncertainty: Bayesian inference
Representation of uncertainty: Bayesian inference
Representation of uncertainty: Bayesian inference
Representation of uncertainty: Bayesian inference
Representation of uncertainty: MaxEnt and Fisher information

Observe at N locations and n sets of observations.
Reduce dimensionality using KL.

Moments of observations as constraints:

\[ \hat{\beta}_j = \frac{1}{n} \sum_{k=1}^{n} \left[ \prod_{i=1}^{N} \left( z_i^{(k)} \right)^{m_{ij}} \right], \quad j = 1, \cdots, p. \]

Maximum Entropy Density Estimation (MEDE) results in joint measure of KL variables:

\[ p_{\mathbf{Z}}(\mathbf{Z}) = \exp \left[ -\lambda^T T(\mathbf{Z}) - \xi(\mathbf{\lambda}) \right] \]

\[ T(z_1, \cdots, z_M) = [t_1(\mathbf{Z}), \cdots, t_p(\mathbf{Z})]^T \]

\[ t_j(\mathbf{Z}) = \left( \prod_{i=1}^{M} z_i^{m_{ij}} \right), \quad j = 1, \cdots, p \]

Fisher Information Matrix:

\[ \mathbf{F}_n(\mathbf{\lambda}) = -E \left[ \frac{\partial^2 \ln \ell(\mathbf{\lambda}|\mathbf{Z}_n)}{\partial \mathbf{\lambda} \partial \mathbf{\lambda}^T} \right] \mathbf{\lambda} \]

\[ \mathbf{Z}_n = \begin{bmatrix} Z_1^T, \cdots, Z_n^T \end{bmatrix}^T \]

Then asymptotically (h denotes coefficients in polynomial chaos description of observations):

\[ h_q(\hat{\mathbf{\lambda}}) \approx N(h_q(\mathbf{\lambda}), h_q(\mathbf{\lambda})^T \mathbf{F}_n(\mathbf{\lambda})^{-1} h_q(\mathbf{\lambda})), \quad q = 1, \cdots, N \]
Stochastic FEM

\[ \varepsilon_{p,h} = u|_d - \sum_{i=0}^{p} u_i \psi_i \]
Variational Formulation:

Find \( u \in H \) s.t.: \( B(u, v) = L(v), \forall v \in H \)

Where:

\[
H = H^1_0(D) \otimes L^2(\Gamma)
\]

\( \Gamma \equiv \Gamma_1 \times \cdots \times \Gamma_M, \quad \xi_i : \Omega \rightarrow \Gamma_i \subset \mathbb{R} \)

\[
B(v, w) \equiv \langle \int_D a \nabla v \cdot \nabla w \rangle, \quad \forall v, w \in H
\]

\[
L(v) \equiv \langle \int_D f v \rangle, \quad \forall v \in H
\]

Notice:

\( B(., .) \) should be coercive and continuous
Approximation Formulation:

Find $u_{X,Y} \in X \otimes Y$ s.t.: $\mathcal{B}(u_{X,Y}, v) = \mathcal{L}(v), \quad \forall v \in X \otimes Y$

Where:

$X \equiv X_h \subset H^1_0(D)$
$Y \equiv Y_p = \bigoplus_{i=1}^P \mathcal{P}_i \subset L_2(\Gamma)$

Write:

$u_{X,Y}(x, \theta) = \sum_{i,j} u_{ij} N_i(x) \psi_j(\theta) \quad \ast \quad \mathcal{KU} = \mathcal{F}$

$i^{th}$ Homogeneous Chaos in $\{\xi_i(\theta)\}_{i=1}^M$
Typical System Matrix: 1st Order
Typical System Matrix: 2nd Order
Typical System Matrix: 3rd Order
Efficient pre-conditioners
Sources of Error:

• **Spatial Discretization Error:**
  \[ H^1_0 \rightarrow X \]

• **Random Dimension Discretization Error:**

  \[ L_2(\Gamma) \rightarrow Y \]
  \[ u_i(\theta) = \sum_{j=1}^{\infty} u_{ij} \psi_j(\theta) \rightarrow u_i(\theta) = \sum_{j=1}^{P} u_{ij} \psi_j(\theta) \]

- Joint error estimation is possible, for general measures, using nested approximating spaces (e.g. hierarchical FEM) (Doostan and Ghanem, 2004, 2005)

- Joint error estimation is possible, for special cases:
  - infinite-dimesional gaussian measure: Benth et.al, 1998
  - tensorized uniform measure: Babuska et.al, 2004
Using Components of Existing Analysis Software

\[
\sum_{j=0}^{P} \left( \sum_{i} c_{ijk} K_i \right) u_j = f_k, \quad k = 0, \ldots, P
\]

\[
\sum_{j=0}^{P} K_{jk} u_j = f_k, \quad k = 0, \ldots, P
\]

\[
K_0 u_k = f_k - \sum_{j \neq k} \sum_{i} c_{ijk} K_{jk} u_j - \sum_{j=0}^{P} (1 - \delta_{j0}) K_{jk} u_j
\]

Essentially preconditioning with $K_0$.

Only one deterministic solve required. Minimal change to existing codes.

Need iterative solutions with multiple right hand sides.

Integrated into ABAQUS (not commercially).
Non-intrusive implementation

The solution $u$ is a function of the basic random variables $x_i$:

$$u = f(x_1, \ldots, x_n)$$

where $f(\ldots)$ is a mapping available through an analysis code.

A Polynomial Chaos Decomposition of the solution has the form:

$$u = \sum_i \psi_i u_i, \quad u_i = \frac{\langle u \psi_i \rangle}{\| \psi_i \|^2}$$

where the mathematical expectation is evaluated through a statistical average over a finite number of samples.

Error Estimator:

$$E = \frac{1}{n_s} \sum_{k=1}^{n_s} \| u(a_k) - u^{r,n_s}(a_k) \|^2_C$$
Example: Protein Labeling

Continuity and momentum equations:

\[ \nabla \cdot \mathbf{u} = 0 \]
\[ \frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{u} = -\nabla p + \nu \nabla^2 \mathbf{u} \]

Wall electrostatic forces (Helmholtz-Smoluchowski relationship):

\[ u_w = \frac{\xi \phi_w}{\mu} \]

Species concentrations:

\[ \frac{\partial c_i}{\partial t} + \nabla \cdot \left[ c_i \mathbf{u} + \mathbf{u}_i^e \right] = \nabla \cdot \left( D_i \nabla c_i \right) + \hat{w}_i \]

Electromigration velocity:

\[ \mathbf{u}_i^e = -\beta_i z_i F \nabla \phi \]

Diffusivity:

\[ D_i = RT \beta_i \]

Electrostatic Field Strength:

\[ \nabla \cdot (\sigma \nabla \phi) = -F \sum_i z_i \nabla \cdot (D_i \nabla c_i) \]

\[ \sigma = F^2 \sum_i z_i^2 \beta_i c_i \]

\[ \zeta = \left\{ \frac{1}{2} \left[ 1 + \tanh(5(pH - 7.5)) \right] \right\} (pH - 7.6) \times \left( -2.7 \ln(M + 2.3 \times 10^{-4}) \right) \]
Chaos representations of various stochastic parameters and solutions:

\[ D(\xi) = \sum_{k=0}^{P} D_k \Psi_k(\xi) \]

\[ D_k = \frac{\left< \Psi_k D \right>}{\left< \Psi_k^2 \right>} \]

\[ u(x,t;\theta) = \sum_{k=0}^{P} u_k(x,t) \Psi_k(\theta) \]

\[ c(x,t;\theta) = \sum_{k=0}^{P} c_k(x,t) \Psi_k(\theta) \]

\[ \phi(x,t;\theta) = \sum_{k=0}^{P} \phi_k(x,t) \Psi_k(\theta) \]

\[ \Psi_0 = 1, \; \Psi_1 = \xi, \; \Psi_2 = \xi^2 - 1, \; \Psi_3 = \xi^3 - 3\xi, \ldots \]

Equations governing evolution of Chaos coordinates:

\[ \frac{\partial u_k}{\partial t} + \sum_{i=0}^{P} \sum_{j=0}^{P} C_{ijk} (u_i \cdot \nabla) u_j = -\nabla p_k + \sum_{i=0}^{P} \sum_{j=0}^{P} C_{ijk} \nabla^2 u_j \quad \forall k \]

\[ C_{ijk} = \frac{\left< \Psi_i \Psi_j \Psi_k \right>}{\left< \Psi_k^2 \right>} \]

\[ \frac{\partial c_{m,k}}{\partial t} + \sum_{i=0}^{P} \sum_{j=0}^{P} C_{ijk} \nabla \cdot \left( c_{m,j}(u_j + u_m^e) \right) = \sum_{i=0}^{P} \sum_{j=0}^{P} C_{ijk} \nabla \cdot (D_{m,j} \nabla c_{m,j}) + \hat{w}_{m,k} \quad \forall k \]

\[ \sum_{i=0}^{P} \sum_{j=0}^{P} C_{ijk} \nabla \cdot (\sigma_i \nabla \phi_j) = -F \sum_{m} \zeta_m \sum_{i=0}^{P} \sum_{j=0}^{P} C_{ijk} \nabla \cdot (D_{m,i} \nabla c_{m,j}) \quad \forall k \]

\[ \sigma_i = F \sum_{m} \zeta_m \sum_{j=0}^{P} \sum_{k=0}^{P} C_{ijk} \beta_{m,j} c_{m,k} \]

Implementation issues:

✓ Stochastic toolkit (working on version 2)

✓ Adapted time integration

✓ Adapted spatial discretization
Protein Labeling: some results

L₂ norm of the difference between solutions on successive grids as a function of the fine grid spacing $dx_f$. The slope of the lines shows a second-order spatial convergence rate for various species concentrations as well as the streamwise velocity.

L₂ norm of the difference between solutions at successive time steps as a function of the shorter time step $dt$. The slope of the lines shows a fourth-order temporal convergence rate for various species concentrations as well as the streamwise velocity.
Time evolution of U and L concentrations in a homogeneous protein labeling reaction. The uncertainty in these concentrations, due to a 1% uncertainty in the labeling reaction rate parameters, is indicated by 63s “error bars.”

PDF of the unlabeled protein concentration at different mean values. As the unlabeled protein reacts away, its PDF becomes narrower and more skewed.

Major contributions of individual input parameters to the overall standard deviation in [L] in the area around the reaction zone at t=0.12 s, y=0.5 mm. The uncertainty in the applied voltage potential “DV” has the most dominant contribution to the overall standard deviation in [L].
Mean concentrations of proteins U, L, and dye D at $t=0.12$ s. U and D just met and L is produced at their interface. The values of the contour levels go linearly from 0 (blue) to $1.3\times10^{-4}$ (red) mol/l.

Standard deviation of the protein and dye concentrations at $t=0.12$ s. The values of the contour levels go linearly from 0 (blue) to $1.1\times10^{-5}$ (red). The largest uncertainties are found in the reaction zone.

Mean (top) and standard deviation (bottom) of the labeled protein concentration L at $t=0.50$ s. The initially flat profiles are now severely distorted. The values of the contour levels go linearly from 0 (blue) to $3.2\times10^{-4}$ mol/l (red) in the top plot and from 0 (blue) to $1\times10^{-4}$ mol/l (red) in the bottom plot.

Mean (top) and standard deviation (bottom) of the electrical conductivity of the electrolyte solution at $t=0.50$ s. Annihilation of ions in the labeling reaction results in a significantly lower mean electrical conductivity near the L plug. The values of the contour levels go linearly from $7.1\times10^{-3}$ S/m (blue) to $1.3\times10^{-2}$ S/m (red) in the top plot and from 0 (blue) to $1.5\times10^{-3}$ S/m (red) in the bottom plot.
Mean (top) and standard deviation (bottom) of the electrical field strength in the x direction at \( t=0.50 \) s. Near the L plug, the mean streamwise electrical field strength is about 40% higher than in the undisturbed flow. The values of the contour levels go linearly from 91.4 kV/m (blue) to 146 kV/m (red) in the top plot and from 0.20 kV/m (blue) to 13 kV/m (red) in the bottom plot.

Mean (top) and standard deviation (bottom) of the electrical field strength in the y direction at \( t=0.50 \) s. The magnitude of the mean of this field strength is up to 15% of the initial field strength in the x direction. The values of the contour levels go linearly from 216.3 kV/m (blue) to 16.3 kV/m (red) in the top plot and from 0 (blue) to 5.8 kV/m (red) in the bottom plot.

Mean (top) and standard deviation (bottom) of the streamwise velocity at \( t=0.50 \) s. The local increase in the electroosmotic wall velocity leads to recirculation zones near the L plug. The largest uncertainties are found near the wall. The values of the contour levels go linearly from 6.8 mm/s (blue) to 9.1 mm/s (red) in the top plot and from 5.6e-3 mm/s (blue) to 0.59 mm/s (red) in the bottom plot.

Mean (top) and standard deviation (bottom) of the wall-normal velocity at \( t=0.50 \) s. The mean of this velocity has a magnitude of up to 6% of the initial streamwise velocity. The values of the contour levels go linearly from 20.56 mm/s (blue) to 0.56 mm/s (red) in the top plot and from 0 (blue) to 0.26 mm/s (red) in the bottom plot.
Connection to multiscale analysis

Need innovative uni-scale models that know what to do with other-scale information: (eg. stochastic homogenization; stochastic equation-free; multiscale mechanics)
(reference: Jardak, M. and Ghanem, R. in CMAME)

New stochastic models for processes exhibiting variation over a few scales. Where spectral analysis, correlation analysis does not apply
(reference: Jianxu S., PhD thesis at Hopkins)

Multiscale data assimilation: transform coarse scale measurements into fine scale parameters (both deterministic and stochastic)
(reference: Zou, Y. and Ghanem, R. in SIAM MMS)

Need logic for targeted scale adaptation: signatures of various subscales in stochastic representation at coarse scale.
On the horizon

- **Critical examination of probabilistic models** of data:
  - Physical and mathematical implications of these models.
  - Connection to multi-scale properties of materials and systems.
  - Adapted bases for enhanced convergence.

- **Efficient numerical solvers:**
  - Using existing codes.
  - Very high-dimensional quadrature.
  - Intrusive algorithms.

- **Visualization** of probabilistic information as decision aids.

- **Model reductions** that maximize information content.

- **Optimization** under uncertainty: uncertainty in objective function, decision variables, and constraints.

- **Validation** of complex interacting systems.

- **Error estimation** and refinement: allocation of resources to physical and numerical experiments.

- **Fusion** of experiments and model-based predictions.


