Motivation

• Can a simple collisional drift-wave model capture experimental measurements of CSDX linear device?

• To answer this question, we carried out quantitative validation study of a CSDX experimental condition.

• To better incorporate the experimental uncertainties, we are taking advantage of uncertainty quantification (UQ) methods developed for other fields in validation studies of plasma turbulence.
Outline

• Theory of Probabilistic Collocation Method (PCM)

• Discussion on the analytical mapping of probability distributions

• Case study of uncertainty quantification near critical gradient

• Implementing improved uncertainty quantification in a validation study of CSDX linear plasma device experiment.
Types of Uncertainty in V&V Study

Simulation Uncertainties

- Reduced model uncertainty e.g. using minimal model instead of full two-fluid or kinetic description
- Numeric uncertainty e.g. due to limited accuracy of numerical integration, or finite grid resolution
- Input parameters uncertainty due to experimental measurement uncertainties

Experimental Uncertainties

- Plasma reproducibility
- Diagnostic properties
- Experimental fitting evaluations
Validation metrics in terms of increasing quality and utility

Increasing quality of validation metrics

Improved Uncertainty Quantification: Forward propagation of input probability distributions

• Let’s consider flux as a function of temperature gradient,

\[ Q = f(x, L_T) \]

e.g. Temperature gradient

• Knowing the probability distribution of input parameter \( p_{LT} \), the goal is to estimate \( Q(L_T) \) function and to construct the quantity of interest (QoI) probability distribution \( p_Q \).
Improved Uncertainty Quantification: Forward propagation of input probability distributions

• Let’s consider flux as a function of temperature gradient,

\[ Q = f(x, L_T) \]

• Knowing the probability distribution of input parameter \( p_{LT} \), the goal is to estimate \( Q(L_T) \) function and to construct the quantity of interest (QoI) probability distribution \( p_Q \).
Improved Uncertainty Quantification: Forward propagation of input probability distributions

• Let’s consider flux as a function of temperature gradient,

\[ Q = f(x, L_T) \]

• Knowing the probability distribution of input parameter \((p_{LT})\), the goal is to estimate \(Q(L_T)\) function and to construct the quantity of interest (QoI) probability distribution \(p_Q\).

\[ \int p_Q dQ = \int p_{LT} dL_T \]
• Using chaos polynomials (CP) to approximate response function.

\[ Q = f(x, L_T) \approx \hat{f}(x, L_T) = \sum_{n} c_n P_n(L_T) \]

e.g. normal, uniform, logarithmic, etc.

Determining the input parameter joint distribution \( p_{LT} \)
Non-intrusive Probabilistic Collocation Method (PCM)  

• Using chaos polynomials (CP) to approximate response function.

\[ Q = f(x, L_T) \approx \hat{f}(x, L_T) = \sum_{n} c_n P_n(L_T) \]

Determining the input parameter joint distribution \( p_{LT} \)

Creating chaos polynomial expansion

nth order orthogonal polynomial from input PDF. Polynomial order TBD from sensitivity analysis.

e.g. normal, uniform, logarithmic, etc.
Non-intrusive Probabilistic Collocation Method (PCM)  

- Using chaos polynomials (CP) to approximate response function.

\[ Q = f(x, L_T) \approx \hat{f}(x, L_T) = \sum_n c_n P_n(L_T) \]

e.g. normal, uniform, logarithmic, etc.

- Determining the input parameter joint distribution \( p_{LT} \)

- Creating chaos polynomial expansion

- Choose samples from \( p_{LT} \)

  nth order orthogonal polynomial from input PDF. Polynomial order TBD from sensitivity analysis.

  Smarter sampling for faster convergence, e.g. Latin Hypercube, Hammersley sequence sampling, etc.

- Using chaos polynomials (CP) to approximate response function.

\[ Q = f(x, L_T) \approx \hat{f}(x, L_T) = \sum_{n} c_n P_n(L_T) \]

1. Determining the input parameter joint distribution \( p_{LT} \)
2. Creating chaos polynomial expansion
   - nth order orthogonal polynomial from input PDF. Polynomial order TBD from sensitivity analysis.
3. Choose samples from \( p_{LT} \)
   - Smarter sampling for faster convergence, e.g. Latin Hypercube, Hammersley sequence sampling, etc.
4. Solve the model for \( Q \)
   - For the number of samples we chose in the previous step

E.g. normal, uniform, logarithmic, etc.
Non-intrusive Probabilistic Collocation Method (PCM)  

- Using chaos polynomials (CP) to approximate response function.

\[ Q = f(x, L_T) \approx \hat{f}(x, L_T) = \sum_n c_n P_n(L_T) \]

- Determining the input parameter joint distribution \( p_{LT} \)
- Creating chaos polynomial expansion
  - nth order orthogonal polynomial from input PDF. Polynomial order TBD from sensitivity analysis.
  - Smarter sampling for faster convergence, e.g. Latin Hypercube, Hammersley sequence sampling, etc.
- Choose samples from \( p_{LT} \)
- Solve the model for \( Q \)
- Using regression techniques to find the coefficients of CP
  - For the number of samples we chose in the previous step
  - e.g. least square method
Outline

• Theory of Probabilistic Collocation Method (PCM)

• Discussion on the analytical mapping of probability distributions

• Case study of uncertainty quantification near critical gradient

• Implementing improved uncertainty quantification in a validation study of CSDX linear plasma device experiment.
Analytically, if we exactly know the $Q - L_T$ relation, we can map out $p_{LT}$ to $p_Q$

$$Q(L_T) = Q_0 + Q_1 L_T$$
Approximating $Q$ by drawing samples from $p_{LT}$ probability distribution and solving for $Q(L_T)$

$$Q(L_T) = Q_0 + Q_1 L_T$$
Strawman Argument: let’s fit a normal distribution with the mean and standard deviation obtained from samples!

\[ Q(L_T) = Q_0 + Q_1 L_T \]

No of samples = 5
Strawman Argument: Re-arranging the samplings in the higher probability region still does not reproduce the analytical $p_Q$!

$$Q(L_T) = Q_0 + Q_1 L_T$$

No of samples = 5
We can converge to the analytical probability distribution by increasing the number of samplings.

\[ Q(L_T) = Q_0 + Q_1 L_T \]

**No of samples = 100**

This means to get the probability distributions right we need about 100 nonlinear simulations!!!
We can do better: Using Probability Collocation Method to map out the probability distribution with a few number of samplings!

\[ Q(L_T) = Q_0 + Q_1 L_T \]

No of samples = 5
Near critical gradient the probability distribution gets skewed! We need to use expected value and confidence intervals to describe response probability.

$$Q(L_T) = Q_0 \exp(Q_1 \exp(L_T + \beta)^\alpha)$$
Near critical gradient the probability distribution gets skewed! We need to use expected value and confidence intervals to describe response probability.

$$Q(L_T) = Q_0 \exp(Q_1 \exp(L_T + \beta)^\alpha)$$

No of samples = 5
Near critical gradient the probability distribution gets skewed! We need to use expected value and confidence intervals to describe response probability.

\[ Q(L_T) = Q_0 \exp \left( Q_1 \exp \left( L_T + \beta \right)^\alpha \right) \]

No of samples = 5

No of samples = 5
Outline

- Theory of Probabilistic Collocation Method (PCM)
- Discussion on the analytical mapping of probability distributions
- Case study of uncertainty quantification near critical gradient
- Implementing improved uncertainty quantification in a validation study of CSDX linear plasma device experiment.
Case study: looking at critical gradient behavior in a collisional drift-wave model in cylindrical geometry

- 3-field drift-reduced nonlocal fluid model evolving density, vorticity, and electron temperature fluctuations,

\[
\frac{\partial \tilde{n}}{\partial t} - \frac{a}{\rho_s r} \frac{\partial n_0}{\partial r} \frac{\partial \tilde{\phi}}{\partial \theta} = \nabla_{\parallel} \tilde{j}_{\parallel} + D_n \nabla_{\perp}^2 \tilde{n} + S_n + \frac{a}{\rho_s} \{\tilde{\phi}, \tilde{n}\}
\]

\[
\frac{\partial \tilde{\Omega}}{\partial t} = \nabla_{\parallel} \tilde{j}_{\parallel} + \mu_{\perp} \nabla_{\perp}^2 \tilde{\Omega} - \nabla_{\perp} \cdot \left( \nu_{in} n_0 \nabla_{\perp} \tilde{\phi} \right) + \frac{a}{\rho_s} \{\tilde{\phi}, \tilde{\Omega}\}; \quad \tilde{\Omega} = \nabla_{\perp} \cdot (n_0 \nabla_{\perp} \tilde{\phi})
\]

\[
\frac{\partial \tilde{T}_e}{\partial t} - \frac{a}{\rho_s r} \frac{\partial T_{e0}}{\partial r} \frac{\partial \tilde{\phi}}{\partial \theta} = \chi_{\parallel} \nabla_{\parallel}^2 \tilde{T}_e + \frac{2}{3} \frac{T_{e0}}{n_0} \nabla_{\parallel} \tilde{j}_{\parallel} + D_{Te} \nabla_{\perp}^2 \tilde{T}_e + S_{Te} + \frac{a}{\rho_s} \{\tilde{\phi}, \tilde{T}_e\}
\]

\[
\tilde{j}_{\parallel} = \frac{1}{\eta_{\parallel}} \left( 1.71 n_0 \nabla_{\parallel} \tilde{T}_e + T_{e0} \nabla_{\parallel} \tilde{n} - n_0 \nabla_{\parallel} \tilde{\phi} \right)
\]

- Using ad-hoc sources for frozen axisymmetric density and temperature profile,

\[
S_n = -\nu_s \tilde{n}_{Z}; \quad S_{Te} = -\nu_s \tilde{T}_{eZ}
\]
Let’s assume we have large uncertainty in the experimental measurement of equilibrium density steepness

- Let assume we have a very coarse measurement of density equilibrium,
Let’s assume we have large uncertainty in the experimental measurement of equilibrium density steepness.

- Let assume we have a very coarse measurement of density equilibrium,
- Fitting a functional form to the measurements,

\[ n_0(r) = (\bar{n} - n_{\text{base}}) \exp\left(-\left(r/L\right)^{n_\alpha/n_\alpha}\right) + n_{\text{base}} \]
Let’s assume we have large uncertainty in the experimental measurement of equilibrium density steepness.

- Let assume we have a very coarse measurement of density equilibrium,
- Fitting a functional form to the measurements,

\[ n_0(r) = (\bar{n} - n_{\text{base}}) \exp \left( -\frac{r}{L} \frac{n_{\alpha}}{n_\alpha} \right) + n_{\text{base}} \]
- Performing scan of linear growth rates and turbulence fluctuation quantities,
Scenario 1: Well above critical gradient probability distribution of the fluctuations is fairly Gaussian.

Making additional experimental observation to reduce the uncertainty
Scenario 1: Well above critical gradient probability distribution of the fluctuations is fairly Gaussian.

[Graphs and charts showing probability density function (PDF) and distribution of fluctuations.]
Scenario 1: Well above critical gradient probability distribution of the fluctuations is fairly Gaussian.

Making additional experimental observation to reduce the uncertainty.
Scenario 1: Well above critical gradient probability distribution of the fluctuations is fairly Gaussian.

Making additional experimental observation to reduce the uncertainty.
Scenario 2: Concavity in $Q(L_T)$ relation skews the probability distribution.

If we measure less density here
Scenario 2: Concavity in $Q(L_T)$ relation skews the probability distribution.

If we measure less density here
Scenario 3: Marginally above critical gradient the response probability distribution gets skewed.

If density measured is near critical gradient
Scenario 3: Marginally above critical gradient the response probability distribution gets skewed.

If density measured is near critical gradient

Vaezi/IAEA-TM17
• Theory of Probabilistic Collocation Method (PCM)
• Discussion on the analytical mapping of probability distributions
• Case study of uncertainty quantification near critical gradient
• Implementing improved uncertainty quantification in a validation study of CSDX linear plasma device experiment.
Case Study: Applying improved UQ methodology in a rigorous validation study of an experimental case.

- Performing validation study of a CSDX linear plasma device experimental condition at 1kG, with Argon plasma, gas pressure of 3mTorr, 1.5kW m=0 helicon source, and insulating endplates [Thakur et al., Phys. Plasmas, 2013].

- Applying improved UQ methodology (PCM) for this validation study.

- Applying improved UQ methodology (PCM) for this validation study.
We have identified 6 nominal input uncertainties with their range of errors.

\[ T_g(r) = (\bar{T}_g - T_{g_{\text{base}}}) \exp\left(-\frac{(r/l_T)^{T_{g_{\alpha}}/T_{g_{\alpha}}}}{T_{g_{\text{base}}}}\right) + T_{g_{\text{base}}} \]
We can reduce the uncertainties dimensionality by performing linear sensitivity analysis of maximum growth rate.
We can reduce the uncertainties dimensionality by performing linear sensitivity analysis of maximum growth rate.

\[ m = 3 \gamma_{\text{lin}} \left( \frac{c_s}{a} \right) \]

\[ \frac{1}{a} \int dr \left( \frac{1}{L_n} \right) \]

\[ \frac{1}{a} \int dr \left( \frac{1}{L_{Te}} \right) \]

Nonlinear simulations and UQ study over linearly most sensitive uncertainties.
To build up UQ analysis, we need to recognize probability distributions of input parameter uncertainties.

- Normal PDF for $T_{i0} = \text{normal}(\text{mean} = 0.6, \text{std} = 0.18) \text{ eV}$
- Uniform PDF for $T_{g\text{base}} = \text{uniform}(\text{low} = 0.03, \text{high} = 0.1) \text{ eV}$
To build up UQ analysis, we need to recognize probability distributions of input parameter uncertainties.

- Normal PDF for $T_{i0} = normal(mean = 0.6, std = 0.18) \ eV$
- Uniform PDF for $T_{g_{base}} = uniform(low = 0.03, high = 0.1) \ eV$

In this case, we have assumed no correlation between input parameter distributions.
Improved sampling techniques can immensely facilitate faster convergence of an UQ method with smaller number of samples.

- Space filling sampling methods such as Hammersley sequence sampling are more efficient representation of multivariate PDFs [Simpson et al., Int. J. Reliab. Appl., 2001].
More Rigorous Comparison by Comparing the Model Response Probability Against the Experiment

\[ \langle I_{sat}^2 \rangle^{0.5} \]

\[ \langle v_f^2 \rangle^{0.5} \]

Experimental measurements

\[ \langle n v_r \rangle \]

\[ \times 10^{16} \]

\[ \times 10^{16} \]
Synthetic Diagnostics Setup

• Translating back physical simulation quantities into experimentally measured quantities for validation studies [Ricci et al., Phys. Plasmas, 2009].

\[ I_{sat} = 0.61 A_{probe} \varepsilon n_e \sqrt{\frac{k_b (T_e + T_i)}{m_i}} \rightarrow \frac{\tilde{I}_{sat}}{\tilde{I}_{sat}} = \frac{\tilde{n}}{\bar{n}} + 0.5 \frac{\tilde{T}_e}{\bar{T}_e} \]

\[ \nu_{float} = \phi_{plasma} - \Lambda \frac{k_b T_e}{e} \rightarrow \frac{e\nu_{float}}{\bar{T}_e} = \frac{e\phi_{plasma}}{\bar{T}_e} - \Lambda \frac{\tilde{T}_e}{\bar{T}_e} \]

\[ \Gamma_{syn}^{(r,\theta,z)} = \frac{\bar{n}\tilde{I}_{sat(r,\theta,z)}}{\tilde{I}_{sat}} \left( \frac{\tilde{\phi}_{f(r,\theta-1,z)} - \tilde{\phi}_{f(r,\theta+1,z)}}{2Br^2\Delta\theta} \right) \]
More Rigorous Comparison by Comparing the Model Response Probability Against the Experiment

\[ <l_{sat}^2>^{0.5} \]

\[ <v_f^2>^{0.5} \]

Experimental measurements

Individual Simulation

\[ <n v_f> \]

\[ <l_{sat} v_f> \]

Vaezi/IAEA-TM17
More Rigorous Comparison by Comparing the Model Response Probability Against the Experiment

\[ <I_{\text{sat}}^2>^{0.5} \]

\[ <\nu_f^2>^{0.5} \]

1σ confidence interval

Individual Simulation

Expected value

Experimental measurements

\[ <n_{\tau}> \]

\[ <I_{\text{sat}} \nu_f> \]
More Rigorous Comparison by Comparing the Model Response Probability Against the Experiment

Model predicts experimental fluctuations at r=2-3 cm within $\pm 1\sigma$ confidence interval.
More Rigorous Comparison by Comparing the Model Response Probability Against the Experiment

Model overestimates experimental edge fluctuations at r=5 cm within $\pm 1\sigma$ confidence interval.
More Rigorous Comparison by Comparing the Model Response Probability Against the Experiment

within ±1σ confidence interval, the synthetically measured particle flux is distorted from the physical particle flux at r=3-5cm!
Summary

• Use of advanced sampling method results in better resolving of the probability space with a fewer number of simulations.

• Use of probabilistic collocation method results in efficient mapping of input parameter probabilities into validation metric probabilities, enabling higher-order statistics comparison of simulation results against the experimental measurements.

• Considering the stochastic properties of input parameters near critical gradients are important. Metrics expected values and confidence intervals cannot be described with error bars!
Backup Slides
Chaos Polynomials

• Unlike oscillations in Lagrange Polynomials, Polynomial Chaos employs inner product spaces weighted with respect to probability distribution [Fienberg, J. Comput. Sci., 2015],

\[ \langle P_n, P_m \rangle = \int P_n(q)P_m(q)p_I(q)dq = 0 \quad n \neq m \]

• These orthogonal polynomials can be generated either analytically e.g. Gram-Schmidt theorem or numerically e.g. discretized Stieltjes recursion.

\[
P_{n+1} = (q - A_n)P_n - B_nP_{n-1}; \quad P_{-1} = 0; P_0 = 1
\]

\[
A_n = \frac{\langle qP_n, P_n \rangle}{\|P_n\|^2}, \quad B_n = \begin{cases} \frac{\|P_n\|^2}{\|P_{n-1}\|^2} & n > 0 \\ \frac{\|P_n\|^2}{\|P_{n}\|^2} & n = 0 \end{cases}
\]
A sample set of Chaos Polynomials generated in an arbitrary distribution space
Higher number of samples justifies higher number of collocation nodes.
Approximating $Q = f(L_T)$ using Chaos Polynomials

• Conventional regression techniques can be used to find polynomial coefficients,

$$
\begin{bmatrix}
P_0(q_0) & \cdots & P_n(q_0) \\
\vdots & \ddots & \vdots \\
P_0(q_{k-1}) & \cdots & P_n(q_{k-1})
\end{bmatrix}
\begin{bmatrix}
c_0 \\
\vdots \\
c_n
\end{bmatrix}
=
\begin{bmatrix}
\hat{f}(q_0) \\
\vdots \\
\hat{f}(q_{k-1})
\end{bmatrix}
$$

• Expected value and variance form of Chaos Polynomials have a simple form,

$$
\mu_f = \left\langle \sum_{n=0}^{N} c_n P_n \right\rangle = \sum_{n=0}^{N} c_n \langle P_n, P_0 \rangle = c_0
$$

$$
\sigma_f = \sum_{n=0, m=0}^{N} c_n c_m \langle P_n P_m \rangle - \langle P_n \rangle \langle P_m \rangle = \sum_{n=0, m=0}^{N} c_n c_m \langle P_n P_m \rangle - c_0^2 = \sum_{n=1}^{N} c_n^2 \|P_n\|^2
$$
All random variables can with aid of the Rosenblatt transformations be transformed to/from the uniform distribution [Rosenblatt, 1952]
Sampling choice in PCM helps to converge to analytical solution faster!

\[ Q(L_T) = Q_0 \exp(Q_1 \exp(L_T + \beta)^\alpha) \]

No of samples = 5
When uncertainties in the stiff region are very large, PCM approximation with small number of samples can be deficient!

\[ Q(L_T) = Q_0 \exp(Q_1 \exp(L_T + \beta)^\alpha) \]

No of samples = 5
Sensitivity scan on the order of the Chaos Polynomials

\[ Q(L_T) = Q_0 \exp(Q_1 \exp(L_T + \beta)\alpha) \]

No of samples = 5
Polynomial Order = 1
Sensitivity scan on the order of the Chaos Polynomials

\[ Q(L_T) = Q_0 \exp (Q_1 \exp (L_T + \beta)^\alpha) \]

No of samples = 5
Polynomial Order = 2
Sensitivity scan on the order of the Chaos Polynomials

\[ Q(L_T) = Q_0 \exp \left( Q_1 \exp \left( L_T + \beta \right)^{\alpha} \right) \]

No of samples = 5
Polynomial Order = 3
No of samples = 5
Polynomial Order = 4

\[ Q(L_T) = Q_0 \exp \left( Q_1 \exp \left( L_T + \beta \right)^\alpha \right) \]
Use of continuous function and with a continuous derivative is necessary near critical gradient for correct mapping of probabilities.

\[ \rho_Q = \frac{1}{\left| \frac{dQ}{dT} \right|} \rho_{LT} \]
Good sampling technique significantly improves the error convergence, while incorporating PCM uncertainty quantification further increases UQ accuracy.
Cross-phase distortion in Particle Flux Comparison with the Experiment

$T_e$ fluctuation effects get amplified by the sheath factor and distorts $(n, \phi)$ cross-phase

Experimental $(I_{sat} - v_f)$ crossphase

Simulation crossphase

[Burin et al., Phys. Plasmas, 2005]
Dealing with time evolution uncertainties (turbulence burstiness) along with input parameter uncertainty

Time evolution uncertainty adds additional step to obtain the probability distribution of QoI. There different ways to deal with this uncertainty,

1. In very small region of input parameter uncertainties, we can assume that the burstiness as white noise, and independent from QoI distribution. -> Square root sum of uncertainties.

2. Analytically derive input parameter relation with time evolution to map out the probability distributions. -> Can be difficult to obtain.

3. Data driven approach: apply Bayes rule to obtain posterior probability of QoI. -> Can become computationally expensive.
Identifying the experimental uncertainties into the model

- Uncertainties in the driving forces (measured via Langmuir Probe)
  - Density equilibrium profile uncertainty
  - Electron temperature equilibrium profile uncertainty

- Uncertainties related to the damping mechanisms
  - Ion temperature value (measured via LIF)
  - Neutral gas temperature profile
    - Core gas temperature (measured via spectroscopy)
    - Edge gas temperature (unknown)
    - Temperature steepness (unknown)

\[ T_g(r) = (T_g - T_{gbase}) \exp\left(-\frac{r}{l_{T_g}} T_{g\alpha} / T_{g\alpha}\right) + T_{gbase} \]