A Tutorial on Bayesian Analysis for (Fusion) Science

Lisa Reusch

Daniel Den Hartog

THE UNIVERSITY of WISCONSIN MADISON

Monday Physics meeting • NSTX, PPPL • Aug 15th, 2016

Outline

- Motivation: Integrated Data Analysis (IDA) collaboration project
- Introduction to Bayesian Analysis
  - Fundamentals of Bayesian analysis
  - Example application of Te measurement using double-filter SXR diagnostic
- Application of Bayesian Analysis to IDA

We aim to improve T_e and Z_{eff} measurement on NSTX-U using integrated data analysis (IDA).

- Over the next three years, we will concentrate on
  - increasing the accuracy, precision, and resolution (both spatial and temporal) of T_e and Z_{eff}
    - Integrate Thomson scattering, Multi-color soft x-ray, charge-exchange recombination spectroscopy, and other diagnostics as appropriate and available
    - Initial priority is Z_{eff}

- This project will also develop IDA expertise, with emphasis on new methods of multi-diagnostic and multi-parameter analysis.

As fusion experiments become more complex, maximum scientific value must be extracted from data.

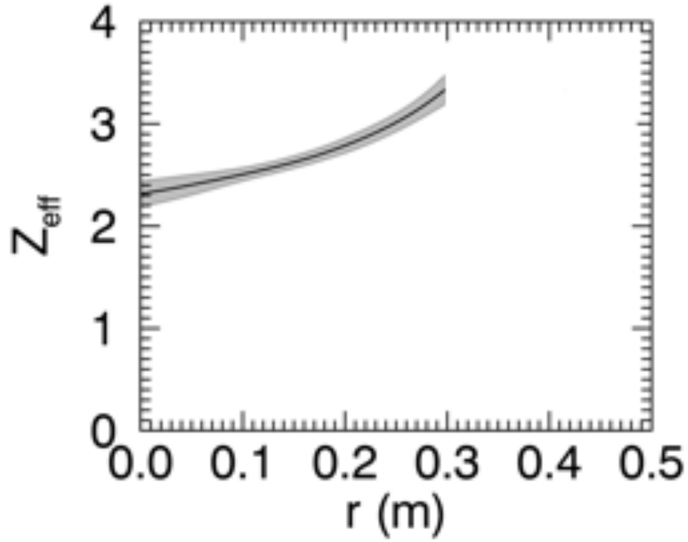
- As we transition to fusion experiments that are full nuclear environments
  - severe limitations will be imposed on diagnostics

- A representative example from the ITPA Diagnostics group of the measurement challenges on ITER:
  - Action Item 20a353, “Can we get Z_{eff} from other measurements?”
    - No dedicated single-chord visible bremsstrahlung measurement is currently planned for installation on ITER

Integrated Data Analysis (IDA) provides a framework to deal with measurement limitations.

- The goal of IDA is to
  - combine data from heterogeneous and complementary diagnostics
  - consider all dependencies within and between diagnostics
  - obtain the most reliable results in a transparent and standardized way.

- Increased measurement precision/resolution is a typical result of IDA.
- IDA additionally enables formation of "meta-diagnostics," which combine information from various instruments to produce unique measurements.
  - Z_{eff} on MST obtained by combining information from CHERS and SXR:



Integrated Data Analysis is often accomplished using a Bayesian statistical framework.

- Highly modular
- Consistent and automatic error analysis
- Can include background/physics information into the analysis quantitatively

Outline

- Motivation: Integrated Data Analysis (IDA) collaboration project
- Introduction to Bayesian Analysis
  - Fundamentals of Bayesian analysis
  - Example application of Te measurement using double-filter SXR diagnostic
- Application of Bayesian Analysis to IDA

Good sources for learning more about Bayesian Analysis

- Sivia D.S. and Skilling J., *Data Analysis: a Bayesian Tutorial 2nd edn* (Oxford: Oxford University), 2006
- Von Toussaint, U., “Bayesian Inference in physics,” *Rev. Mod. Phys*. **83**

Bayesian analysis has an alternate view of data interpretation

- In a classical approach to data analysis
  - An objective Truth exists (i.e. a single right answer)
  - Measurements sample that Truth imperfectly.
  - Repeated measurements reveal the underlying Truth.

- A Bayesian approach
  - Has is a truth, but it is more subjective. (For example, the data itself is the truth—it's the physical entity that can be measured.)
  - Attempts to determine what we can learn from a single measurement or a limited number of measurements.

The foundations of Bayesian Analysis are Bayes' Rule and Marginalization

Bayes' Rule:

$$P(x \mid y, I) = \frac{P(y \mid x, I) P(x \mid I)}{P(y \mid I)}$$

Marginalization:

$$P(x \mid y, I) = \int P(x, \alpha \mid y, I)\, d\alpha$$

Estimation of electron temperature takes the form

$$\underbrace{P(T_e \mid D,\sigma)}_{Posterior} = \frac{\overbrace{P(D \mid T_e,\sigma)}^{Likelihood}\ \overbrace{P(T_e \mid \sigma)}^{Prior}}{\underbrace{P(D \mid \sigma)}_{evidence}}$$

- T_e is electron temperature, the desired parameter
- D is the diagnostic data, i.e. the actual measurement made
- σ is the uncertainty of the system

The posterior is the desired result

$$\underbrace{P(T_e \mid D,\sigma)}_{Posterior} = \frac{\overbrace{P(D \mid T_e,\sigma)}^{Likelihood}\overbrace{P(T_e \mid \sigma)}^{Prior}}{\underbrace{P(D \mid \sigma)}_{evidence}}$$

- Can be interpreted as the result and error bar.
- Can be calculated directly
  - Often involves inversions or fitting routines
  - Often lacks uncertainty information

The Likelihood Function relates the measurement to the parameter of interest through modeling.

$$\underbrace{P(T_e \mid D,\sigma)}_{Posterior} = \frac{\overbrace{P(D \mid T_e,\sigma)}^{Likelihood}\ \overbrace{P(T_e \mid \sigma)}^{Prior}}{\underbrace{P(D \mid \sigma)}_{evidence}}$$

- Involves forward models of the system
  - Physical processes generating signal
  - Instrumentation effects
- Is often easier to calculate than the posterior probability
- Can incorporate systematic and statistical uncertainties in a straightforward way

The Prior probability reflects our background knowledge about the system

$$\underbrace{P(T_e \mid D,\sigma)}_{Posterior} = \frac{\overbrace{P(D \mid T_e,\sigma)}^{Likelihood}\ \overbrace{P(T_e \mid \sigma)}^{Prior}}{\underbrace{P(D \mid \sigma)}_{evidence}}$$

- Often this is a range of values in which we expect the answer to lie.
- Often informed by physical constraints
  - Te must be positive
  - Must be within the measurement range of the diagnostic

The evidence is a normalization factor

$$\underbrace{P(T_e \mid D,\sigma)}_{Posterior} = \frac{\overbrace{P(D \mid T_e,\sigma)}^{Likelihood}\overbrace{P(T_e \mid \sigma)}^{Prior}}{\underbrace{P(D \mid \sigma)}_{evidence}}$$

- Can be ignored for parameter estimation problems
- Often important when choosing between 2 or more models
- Bayes' Rule becomes:

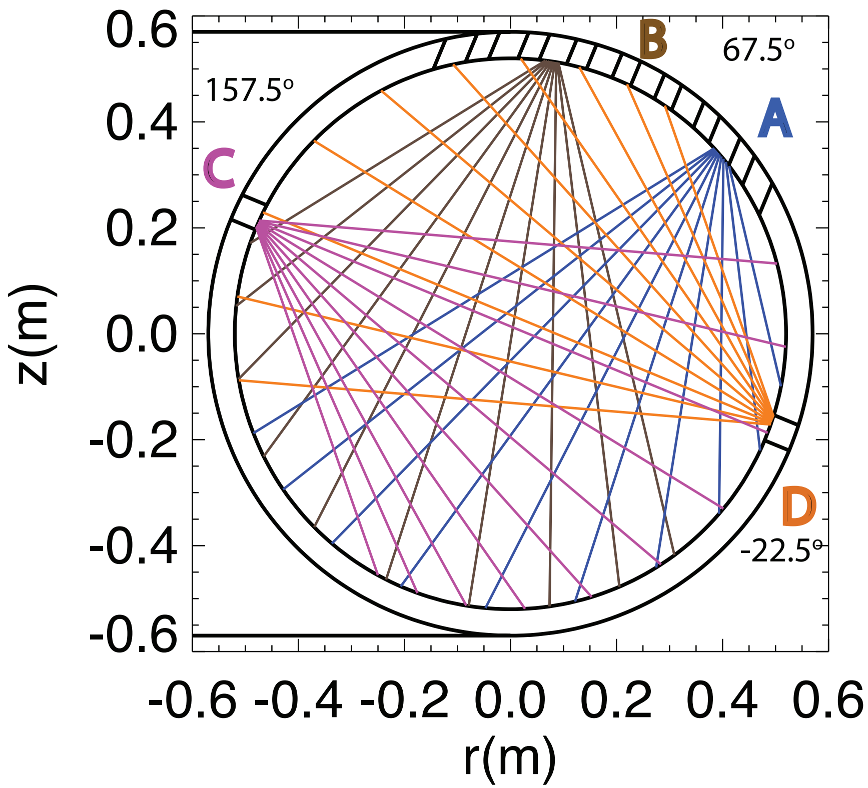
$$\underbrace{P(T_e \mid D,\sigma)}_{Posterior} \propto \overbrace{P(D \mid T_e,\sigma)}^{Likelihood}\overbrace{P(T_e \mid \sigma)}^{Prior}$$

Outline

- Motivation: Integrated Data Analysis (IDA) collaboration project
- Introduction to Bayesian Analysis
  – Fundamentals of Bayesian analysis
  – Example application of Te measurement using double-filter SXR diagnostic
- Application of Bayesian Analysis to IDA

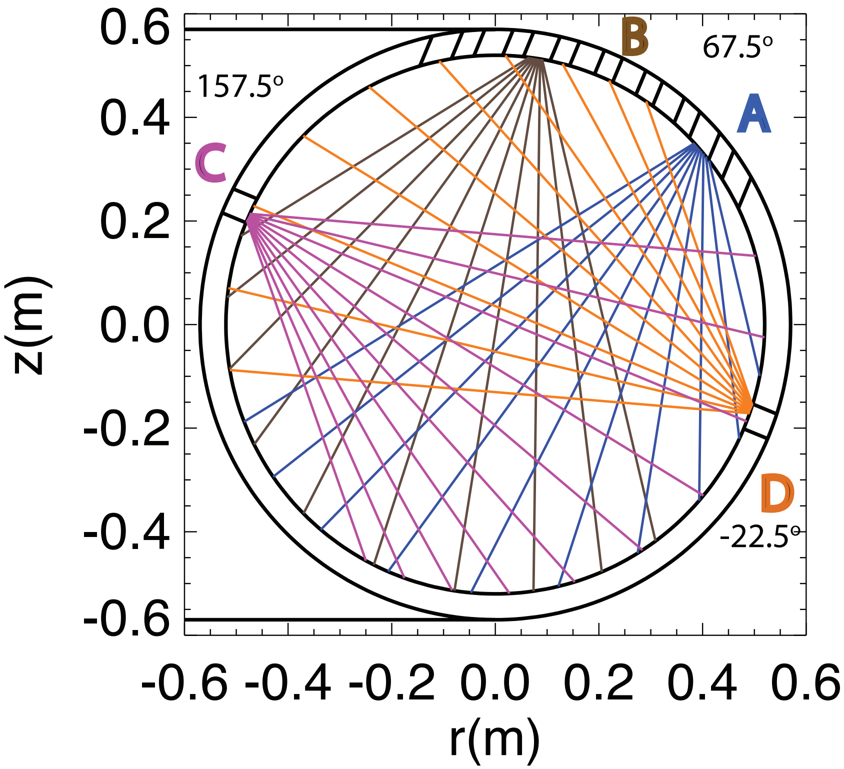
The two-color SXR tomography system contains T_e information in a poloidal cross-section.

- 40 unique lines of sight at one toroidal location
  - Two detectors: 425 µm and 800 µm Be filters
- Thicknesses chosen to block high energy emission lines
- Different thickness filters allow estimation of x-ray spectrum slope



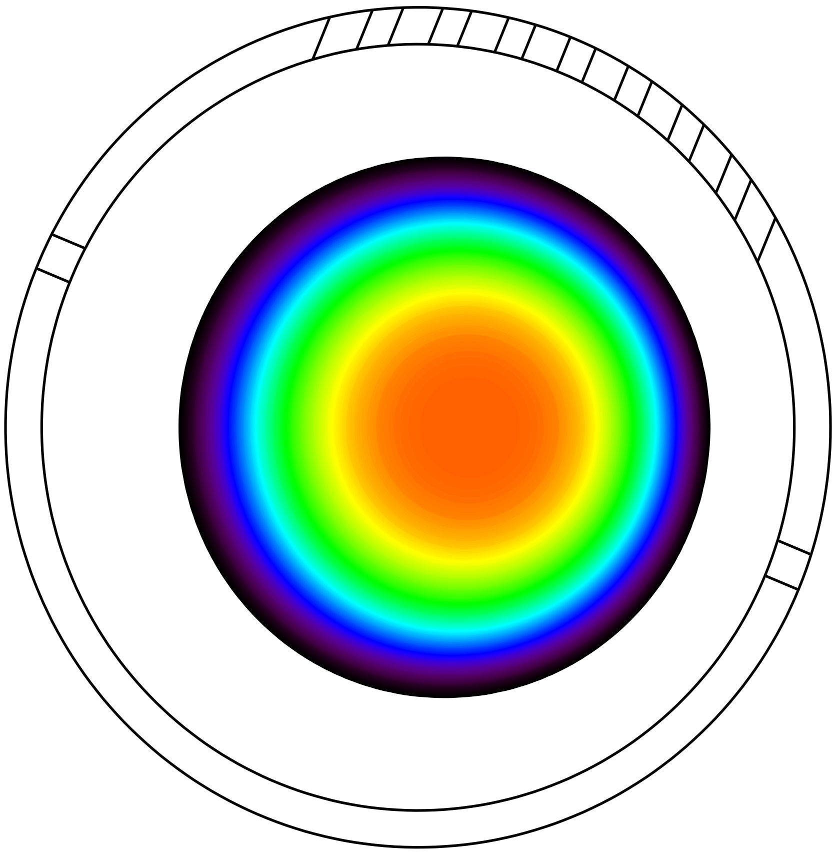
Two-color SXR tomography system is capable of measuring T_e.

- $f \sim n_e Z_{\text{eff}} T_e^{-1/2} e^{-E/T_e}$
  $+ \ recomb$
- Detectors that share a line of site have same n_e, Z_{eff}, and recombination.
- $\frac{f_{thin}}{f_{thick}} \sim T_e$
- The relationship of the measured ratio to Te is not straightforeward.



SXR forward model predicts x-ray emissivity on a 2-dimensional grid.

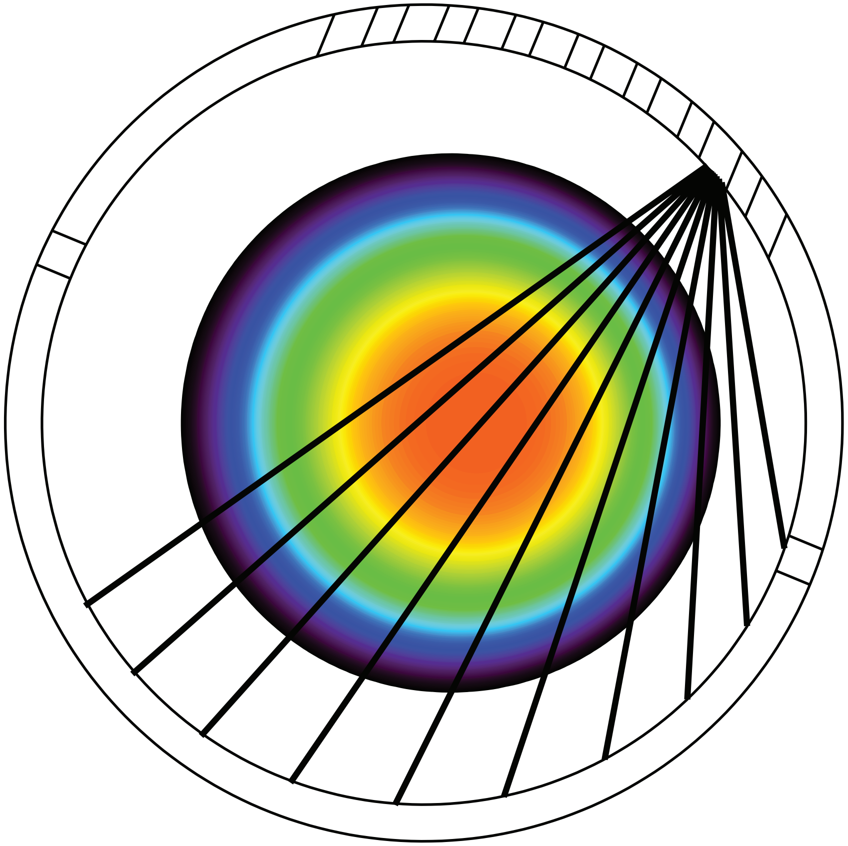
- Three parameter function for T_e profile used (axisymmetric)
  $T_e(r) = T_{e0}(1-(r/a)^{\alpha})^{\beta}$
- Ansatz profiles for density and Z_{eff} used.



Predicted brightness based on x-ray emissivity calculated for each detector

- Predicted brightness takes into account:
  - Geometry of line of sight
  - Detector effects
  - Be filter effects

- Predicted ratio, R_p, calculated from predicated brightness



Likelihoods follow a Gaussian distribution.

- SXR detection relies on sensing photons ➔ Poisson distributions
- In the limit of large numbers of photons, they are Gaussian distributions.

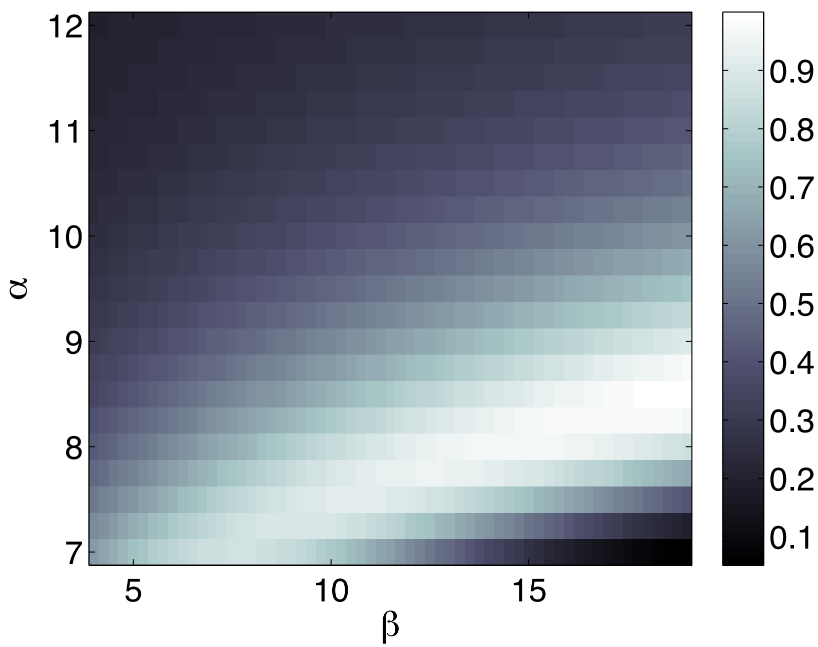
$$P(D_{SXR} \mid T_e(r), \sigma) = \frac{1}{\sqrt{2\sigma^2}} e^{-\chi^2}$$

$$\chi^2 = \frac{\left(R_m - R_p\right)^2}{2\sigma^2}$$

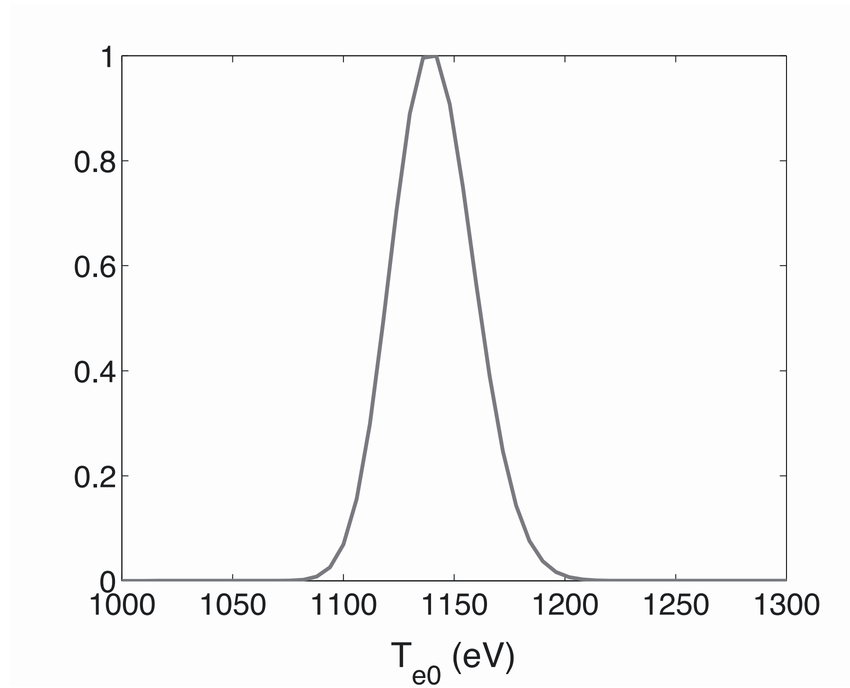
- σ contains both the statistical and systematic sources of uncertainty
- Systematic uncertainties estimated from:
  - Tolerances and Machining precision
  - Filter thickness measurements
  - Filter calibrations

Likelihood function are often multidimensional

$$P(D_{SXR} \mid T_e(r), \sigma) \rightarrow P(D_{SXR} \mid T_{e0}, \alpha, \beta, \sigma)$$



Marginalized core T_e
(T_{e0})



Marginalized edge gradient
(α and β)

Prior distributions can be uniform PDFs over a range informed by experiment

Parameter	Range
T_{e0}	500-2300 eV
α	7-12
β	4-19

$$P(T_{e0}) = \begin{cases} \dfrac{1}{2300-500} & \text{for } 500 \le T_{e0} \le 2300 \\ 0 & \text{otherwise} \end{cases}$$

$$P(\alpha) = \begin{cases} \dfrac{1}{12-7} & \text{for } 7 \le \alpha \le 12 \\ 0 & \text{otherwise} \end{cases}$$

$$P(\beta) = \begin{cases} \dfrac{1}{4-19} & \text{for } 4 \le \beta \le 19 \\ 0 & \text{otherwise} \end{cases}$$

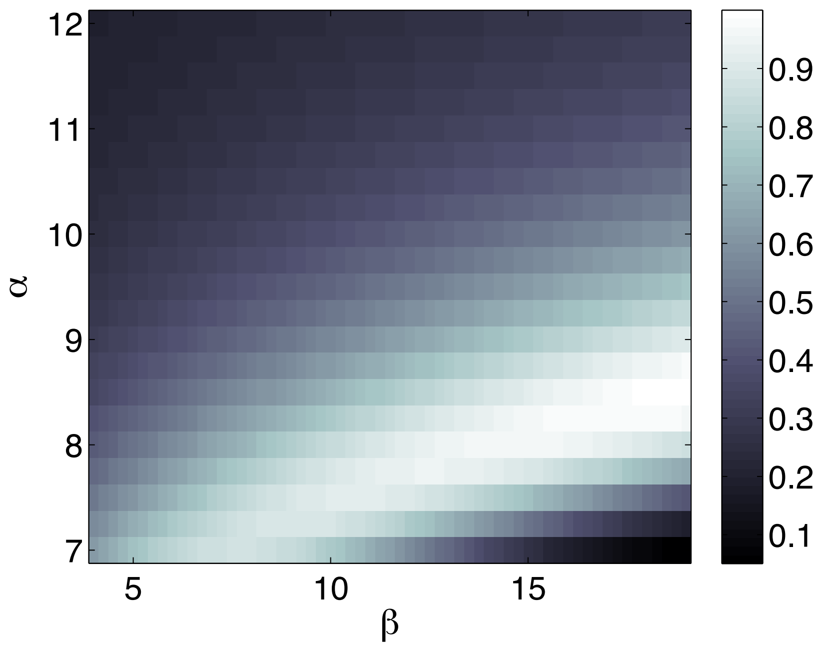
- Likelihood functions calculated on uniform grid of points covering these ranges

The posterior probability is the answer desired

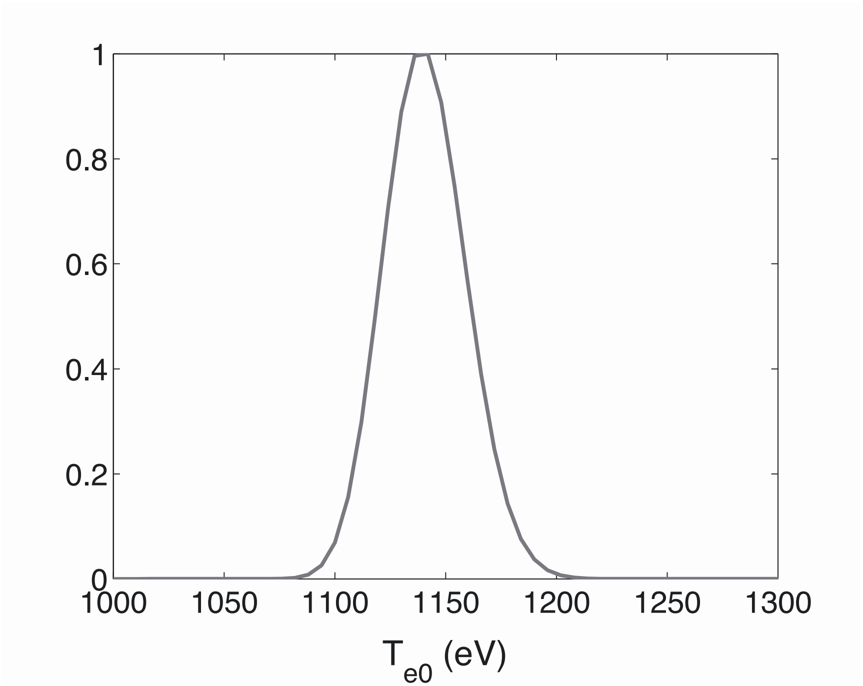
$$P(T_e(r) \mid D_{SXR}, \sigma) \propto \underbrace{P(D_{SXR} \mid T_{e0}, \alpha, \beta, \sigma)}_{\text{Likelihood function}} \underbrace{P(T_{e0}) P(\alpha) P(\beta)}_{\text{Priors}}$$

The posterior probability is the answer desired

$$P(T_e(r)\,|\,D_{SXR},\sigma) \propto \underbrace{P(D_{SXR}\,|\,T_{e0},\alpha,\beta,\sigma)}_{\text{Likelihood function}}\underbrace{P(T_{e0})P(\alpha)P(\beta)}_{\text{Priors}}$$



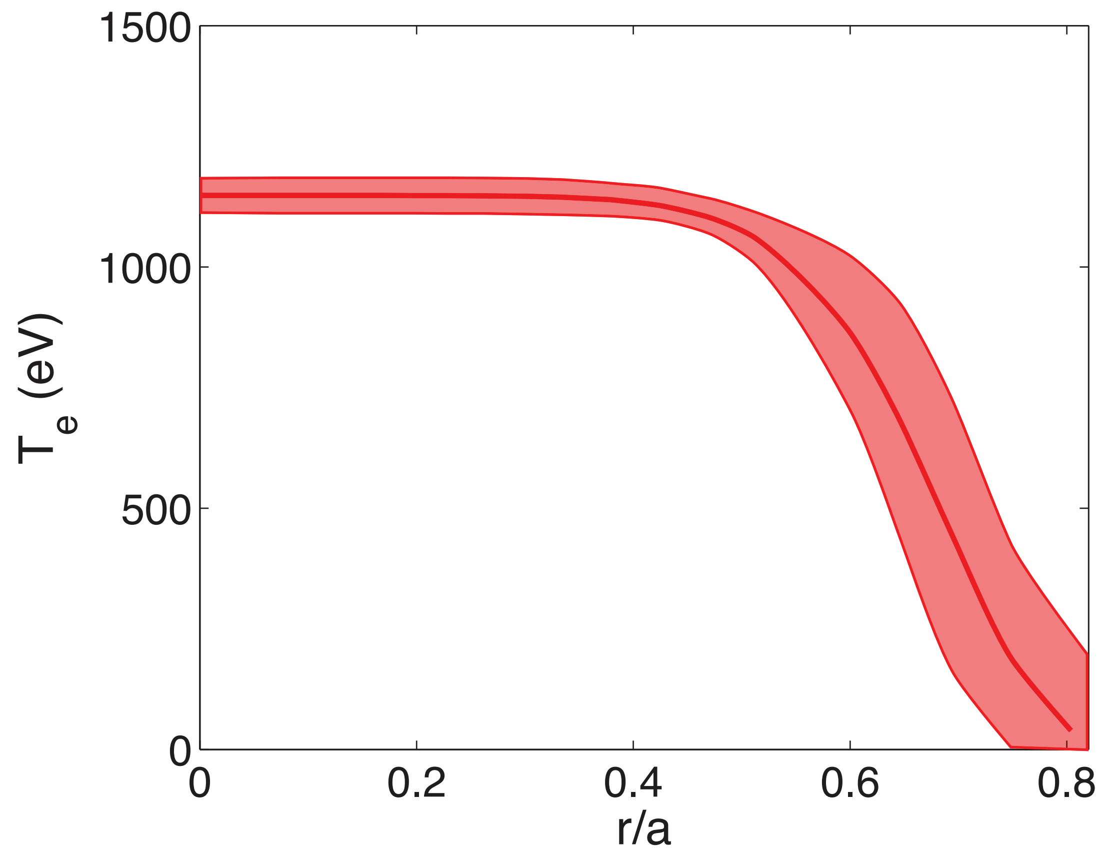
Marginalized core T_e



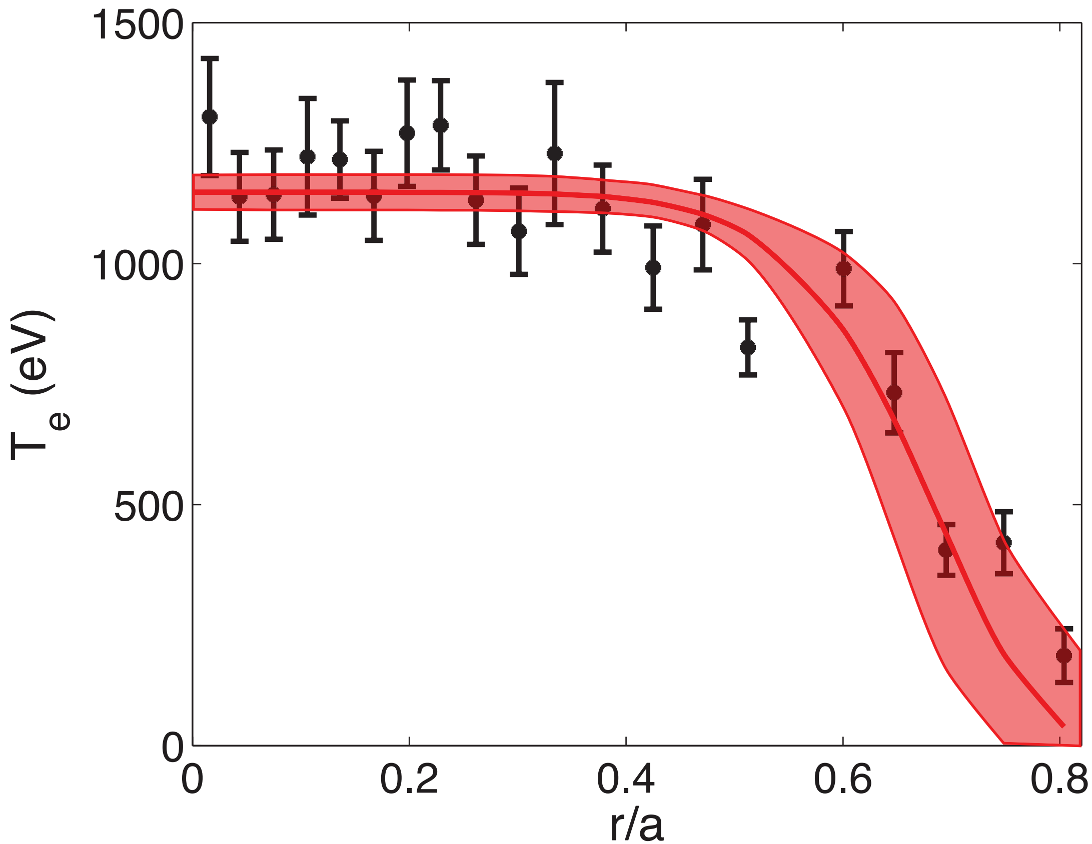
Marginalized edge gradient

The result is the most likely T_e profile

$$\underbrace{P(T_e(r) \mid D_{SXR}, \sigma) \propto P(D_{SXR} \mid T_{e0}, \alpha, \beta, \sigma)}_{\text{Likelihood function}} \underbrace{P(T_{e0}) P(\alpha) P(\beta)}_{\text{Priors}}$$



The most likely T_e profile agrees well with TS.

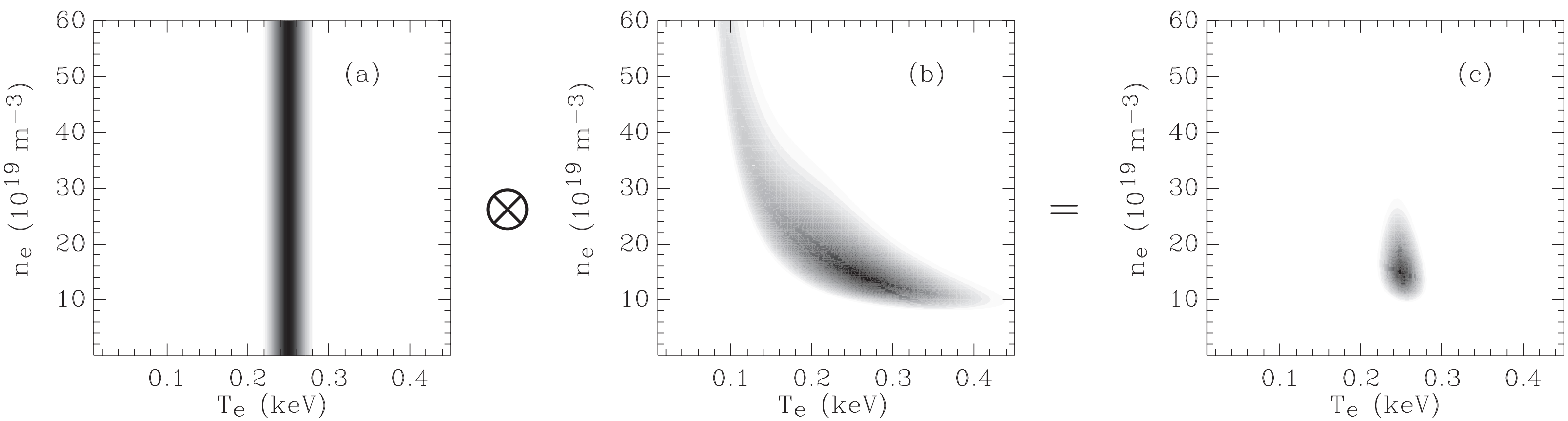


This is the most likely profile assuming the plasma is axisymmetric

Outline

- Motivation: Integrated Data Analysis (IDA) collaboration project
- Introduction to Bayesian Analysis
  - Fundamentals of Bayesian analysis
  - Example application of Te measurement using double-filter SXR diagnostic
- Application of Bayesian Analysis to IDA

Integrated Data Analysis in a Bayesian framework proceeds with a series of steps:



- Identify uncertainties and quantify with probability distribution functions (PDF)
- Combine all relevant information within a probabilistic framework
  - include diagnostic models and prior knowledge
  - develop a forward model for measurement
  - marginalize out nuisance parameters such as systematic effects
- Search parameter space, which is often high-dimensional
- Final result is the posterior PDF of the quantity of interest

R. Fischer, A. Dinklage, and E. Pasch, “Bayesian modelling of fusion diagnostics,” Plasma Phys. Control. Fusion **45**, 1095–1111 (2003).

Using Bayesian Framework to estimate Z_{eff}

- Many Diagnostics are sensitive to Z_{eff}
  - Near infrared / Visible Bremsstrahlung
  - Charge exchange recombination spectroscopy (CHERS)
  - Soft x-ray (SXR)
  - Neutral beam attenuation
  - Thomson Scattering background light
  - Loop voltage
  - …etc.

$$P(Z_{\mathrm{eff}} \mid SXR, CHERS, \ldots) \propto L_{SXR} \times L_{CHERS} \times \ldots \times \mathrm{Priors}$$

$$e.g.\ \ L_{SXR} \sim P(SXR \mid Z_{\mathrm{eff}})$$

Summary

- We have started developing an IDA technique using a Bayesian probability framework to improve T_e and Z_{eff} measurements on NSTX-U
- The fundamental ideas in Bayesian Analysis are:
  - Bayes' Rule: $P(x \mid y,I) = \dfrac{P(y \mid x,I)P(x \mid I)}{P(y \mid I)}$
  - Marginalization: $P(x \mid y,I) = \int P(x,\alpha \mid y,I)\,d\alpha$
- Consideration of all assumptions, background physical knowledge and uncertainties are necessary when developing a Bayesian approach to the analysis of a diagnostic.
- Bayesian Analysis is a natural framework in which to develop IDA due its
  - Modularity
  - Automatic error analysis
  - Ability to include background information in the analysis