# Bayesian Analysis at JET and the Minerva framework

J Svensson, O Ford, A Werner, S Schmuck, D C McDonald, M Beurskens, A Boboc, M Brix, A Meakins, M Krychowiak

# Problems with modern large scale scientific experiments.

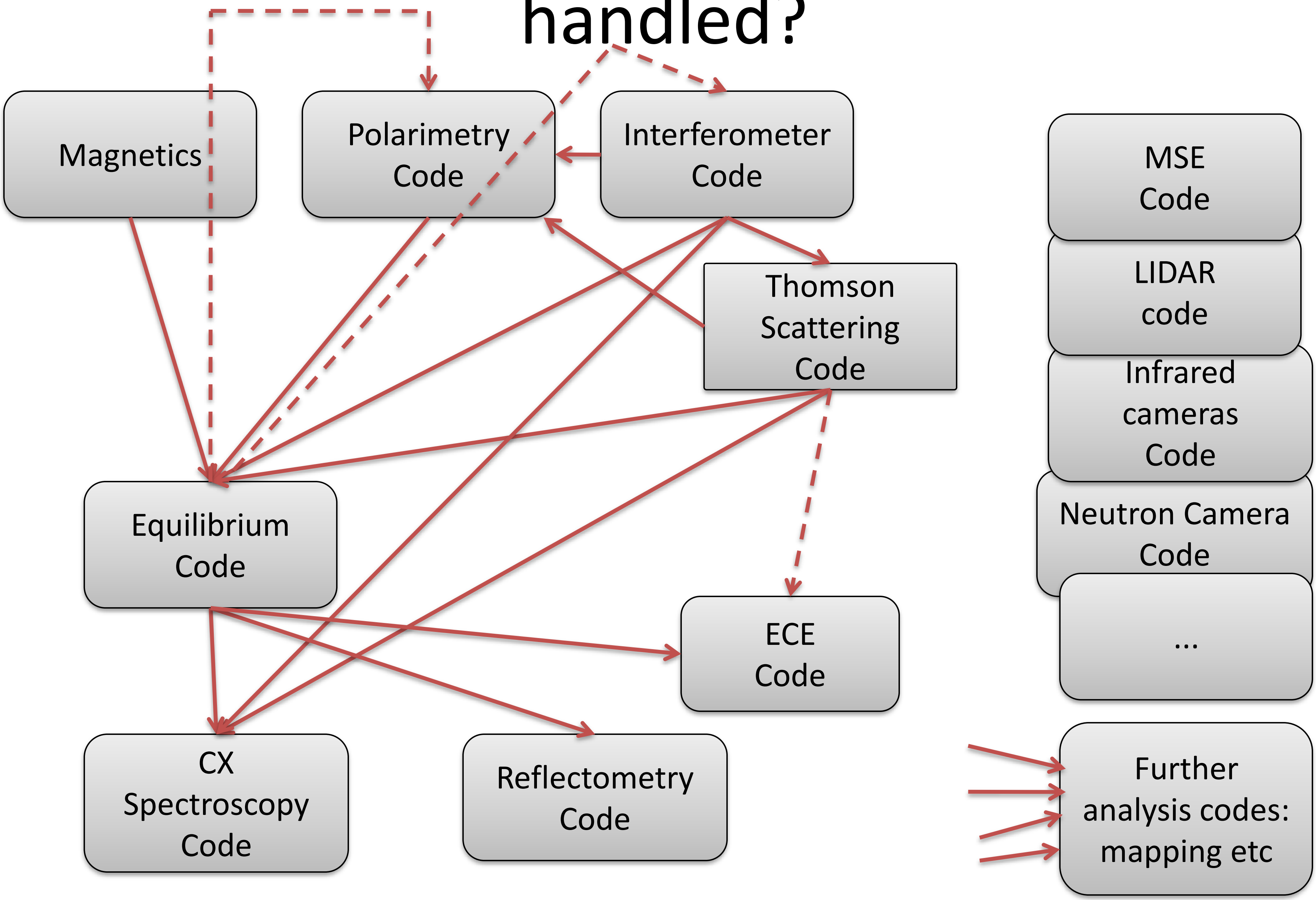
## Drowning in data?

**Solutions**: Better data aquisition systems, databases etc.

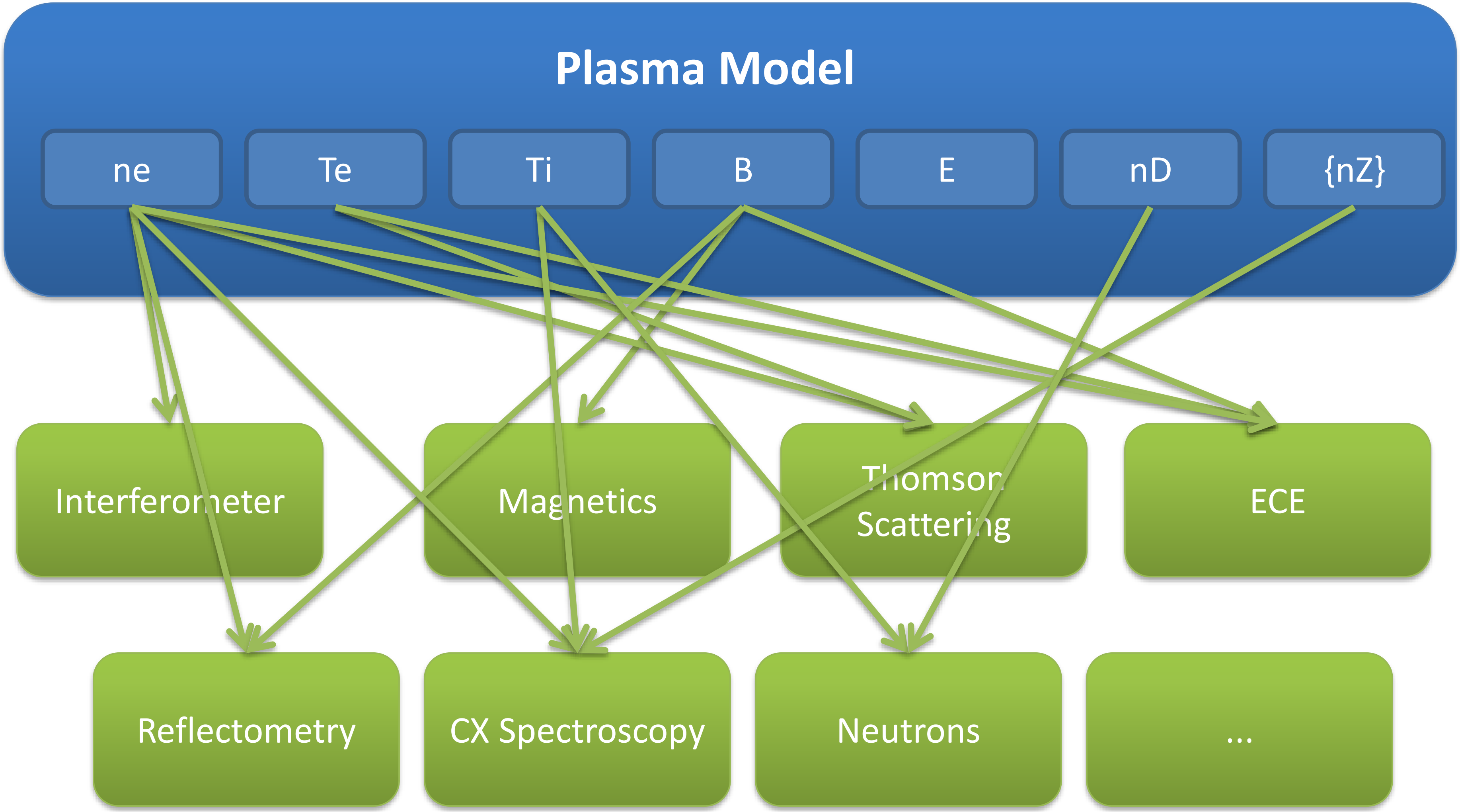
## Drowning in complexity.

**Solutions**: ?

# How is scientific inference currently handled?



# A top-down view on our inference problem.



# The Minerva analysis infrastructure

## Minerva

• A fully modular inference infrastructure (separating diagnostic models/combinations, physics assumptions, inversions, forward models etc)

• Based on Bayesian Graphical Models (handles modularity of complex pdf:s)

• Handles all dependency management (keeps track of every single parameter/option that can change), so user never needs to think about data flow, this also makes automatic caching possible.

• Can fully *declare* a scientific model in a complex experiment.

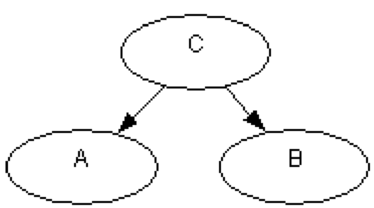
• Written in Java

• Machine independent (used at JET, MAST, W7-X, H-1?)

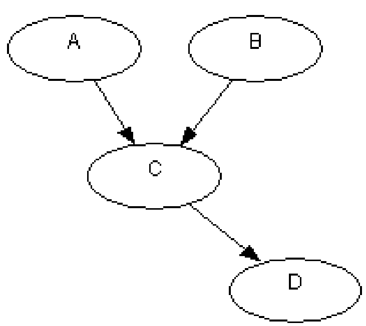
J Svensson, A Werner, *Large Scale Bayesian Data Analysis for Nuclear Fusion Experiments*, WISP 2007

# Probability Engineering: Bayesian Graphical Models:

- A combination of graph theory and probability theory.
- Makes it easier to handle complex probabilistic systems (like our systems!).
- Nodes represent variables: stochastic or deterministic.
- Edges represent dependencies




p(a,b,c)=p(a|c)*p(b|c)*p(c)



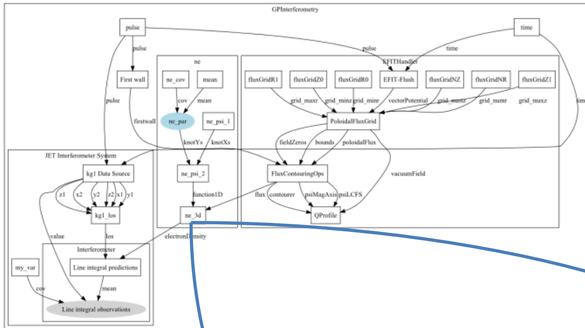

p(a,b,c,d)=p(d|c)*p(c|a,b)*p(a)*p(b)

**Generally**: $$p(n_1, n_2, n_3, \ldots, n_{N_T}) = \prod_{i=1}^{N_T} p(n_i \mid par(n_i))$$

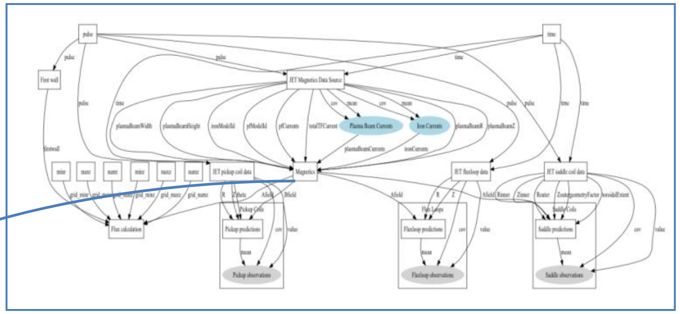
# Minerva

## Example graphs

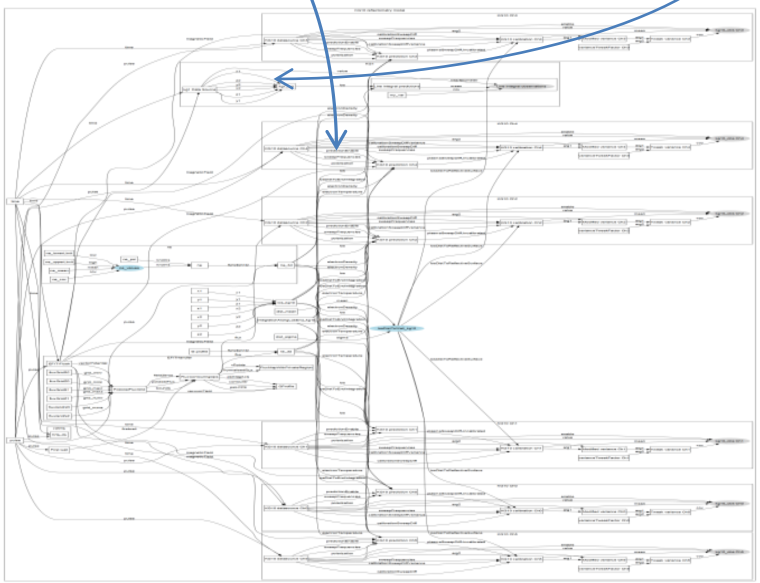
KG1 (interferometry)



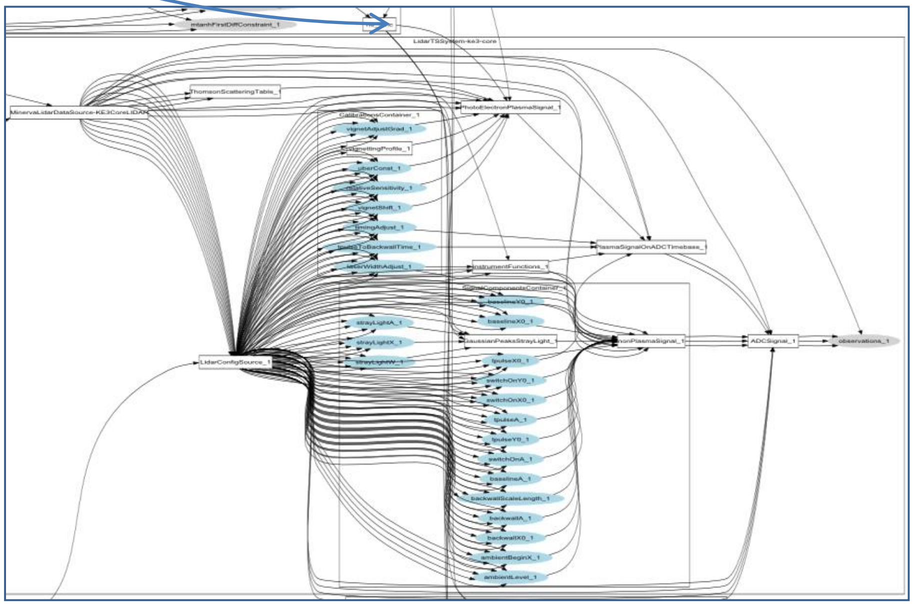
Magnetics ,PF



KG10 (reflecometry)



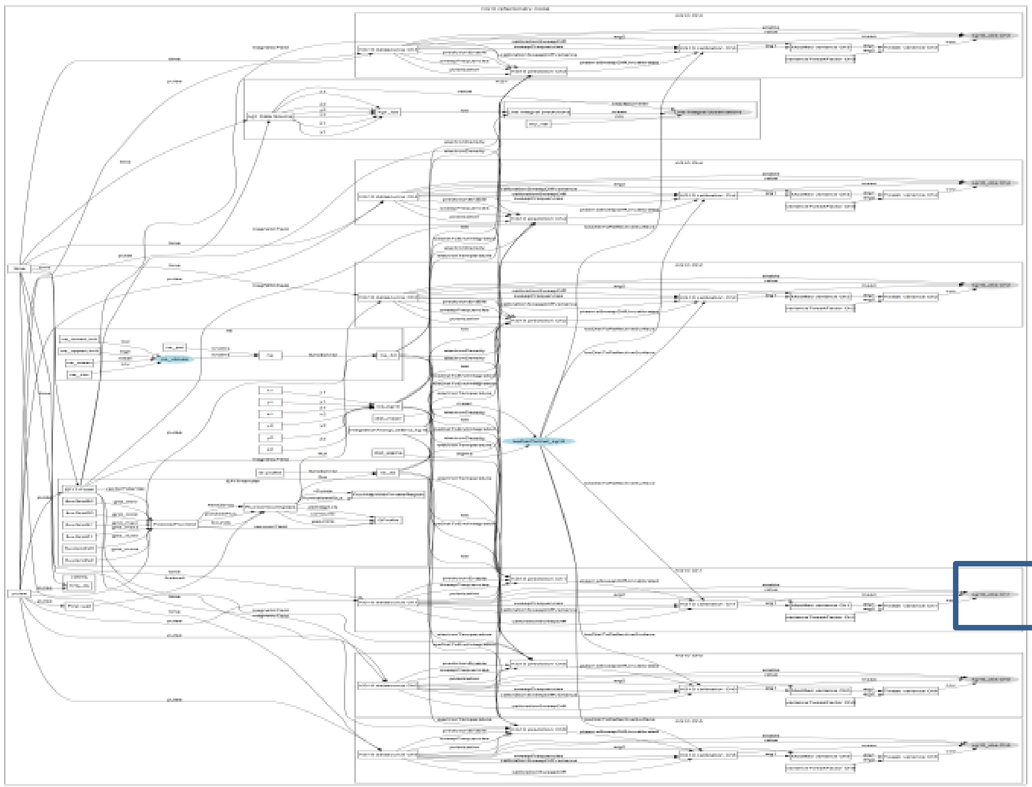
KE3 (LIDAR)

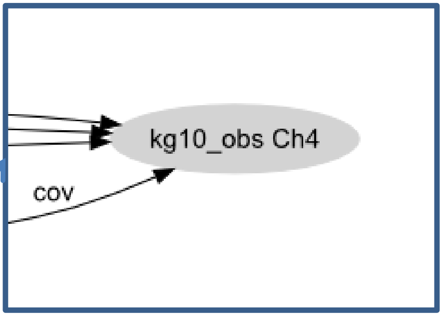


…

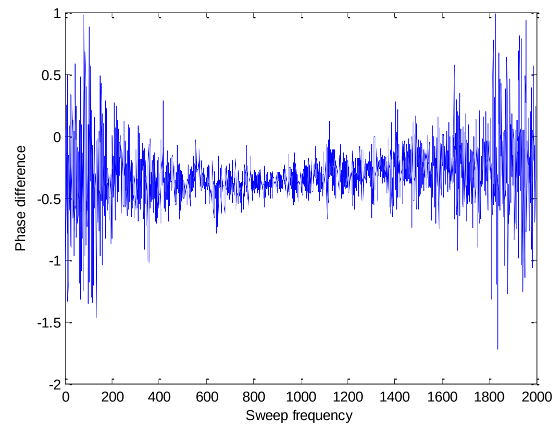
# Minerva
## Prediction

KG10



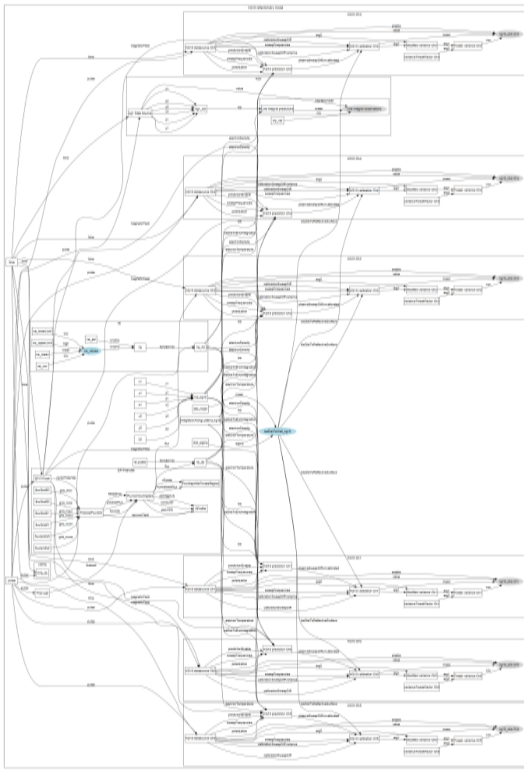



```
.sample()
```




-> Can be used for understanding a diagnostic/ diagnostic “debugging”/diagnostic design/ experimental planning etc

# Minerva
## Inversion



Drop in

MAximum Posterior

Gradient Descent
Conjugate Gradients
Coordinate Descent
Hooke & Jeeves
Nelder-Mead
Genetic Algorithms
...

Markov Chain Monte Carlo

Random Walk
Independent Chain
Adaptive step size
Adaptive covariance
...

Linear Inversion

Automatically finds linear coefficients. Produces full posterior.

...others

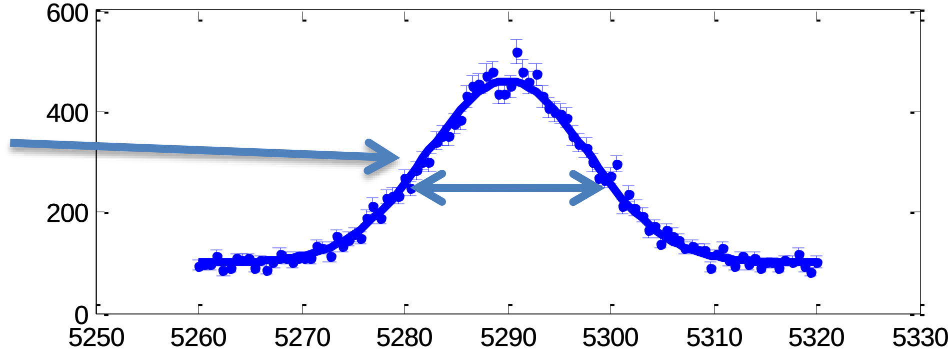
Model Description/ Definition

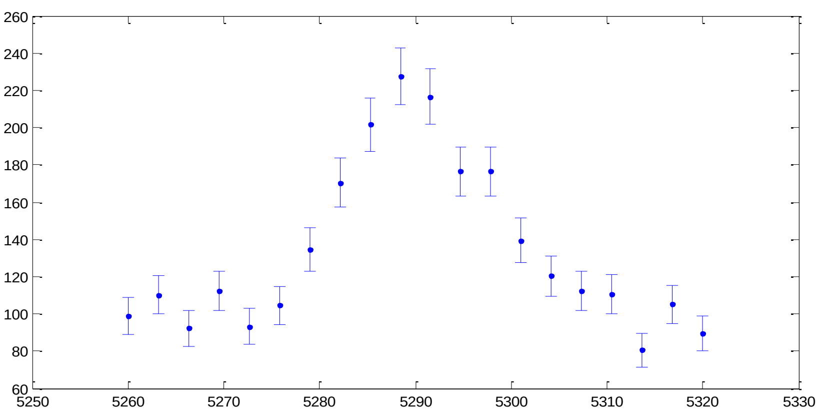
Model Inversion

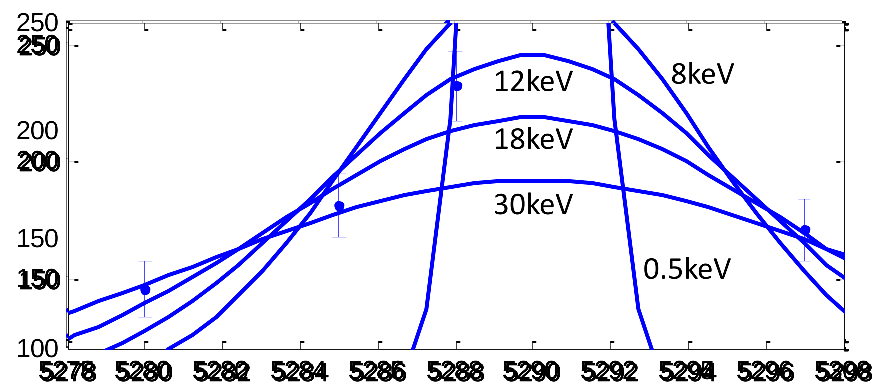
# Bayes intro: An inference example

Extract temperature from Doppler broadening of spectral line:

$$width = 2\lambda_0 \sqrt{\frac{kT}{mc^2}}$$



# Bayes Theorem

Likelihood

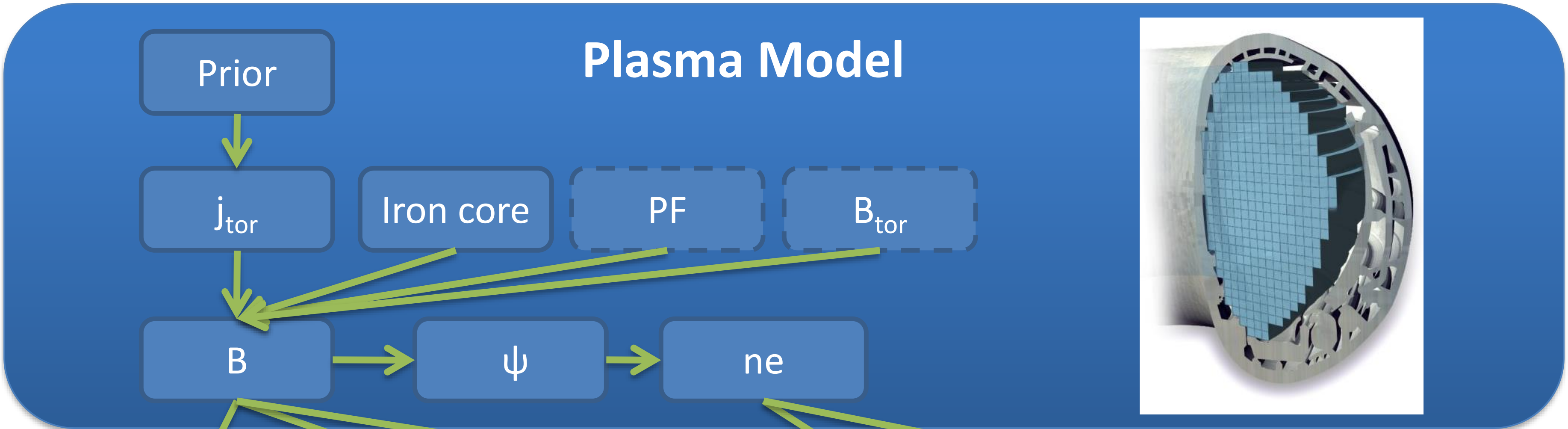
Prior

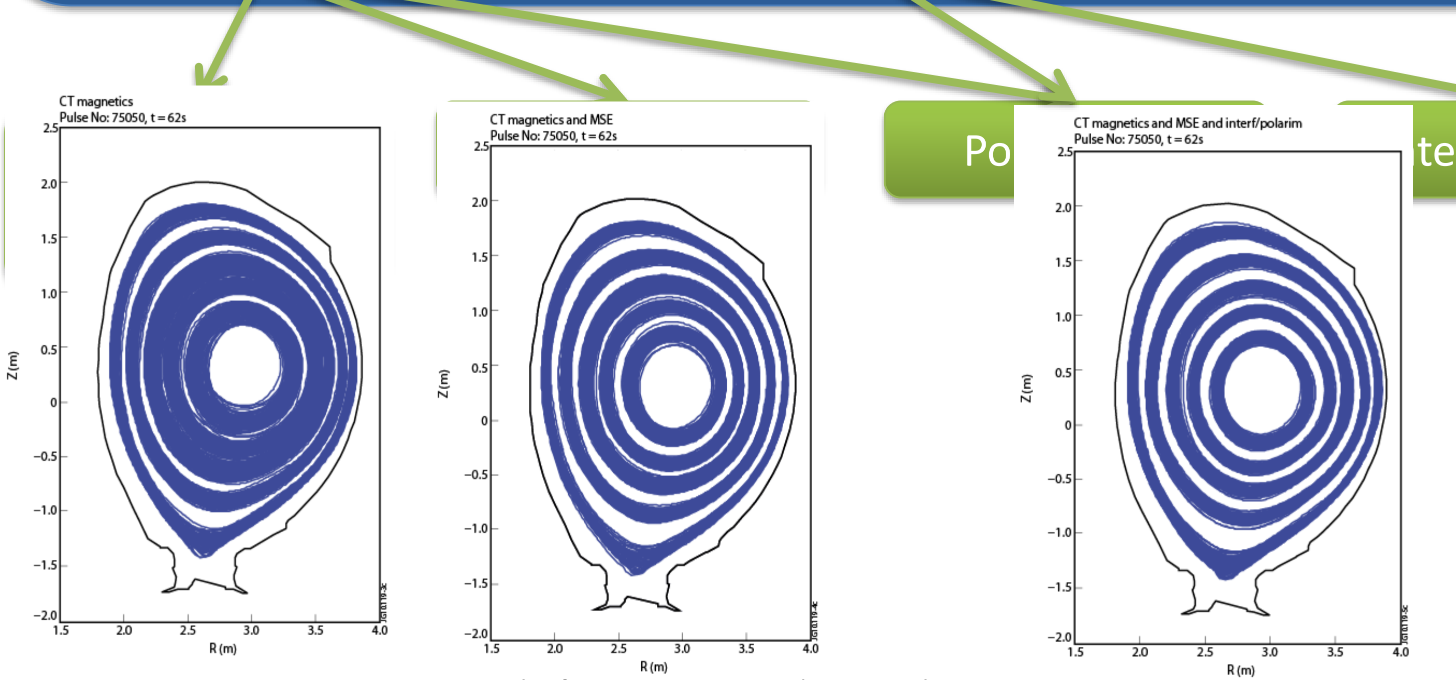
Posterior

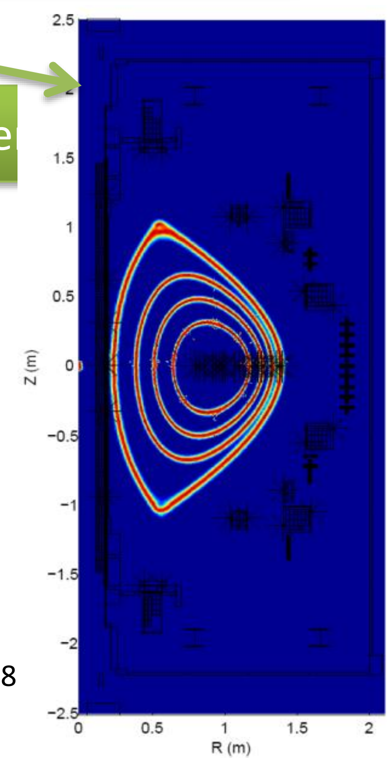
$$p(T \mid d) = \frac{p(d \mid T)\, p(T)}{p(d)}$$

Model Evidence

# Inference on toroidal current distribution





- J Svensson, A Werner, *Current Tomography for Axisymmetric Plasmas*, *Plasma Phys. Control. Fusion* **50** No 8, 2008
- G von Nessi et al, forthcoming

# Maths for current tomography

Forward function: $\overline{D}^{\text{Pred}} = \overline{\overline{M}}\,\overline{I} + \overline{C}$

Likelihood: Gaussian noise

$$p(\overline{D}^{Mag} \mid \overline{I}) = \frac{1}{(2\pi)^{N_D/2} \left|\overline{\overline{\Sigma}}_D\right|^{1/2}} \exp\left(-\frac{1}{2}(\overline{\overline{M}}\,\overline{I} + \overline{C} - \overline{D}^{Mag})^T \overline{\overline{\Sigma}}_D^{-1} (\overline{\overline{M}}\,\overline{I} + \overline{C} - \overline{D}^{Mag})\right)$$

---

Prior: Multivariate Normal over free currents
(prior covariance imposes e.g. smoothing between nearby beams)

$$p(\overline{I}) = \frac{1}{(2\pi)^{N_I/2} \left|\overline{\overline{\Sigma}}_I\right|^{1/2}} \exp\left(-\frac{1}{2}(\overline{I} - \overline{m}_I)^T \overline{\overline{\Sigma}}_I^{-1} (\overline{I} - \overline{m}_I)\right)$$

---

$\Longrightarrow$ Posterior: Multivariate normal over free currents

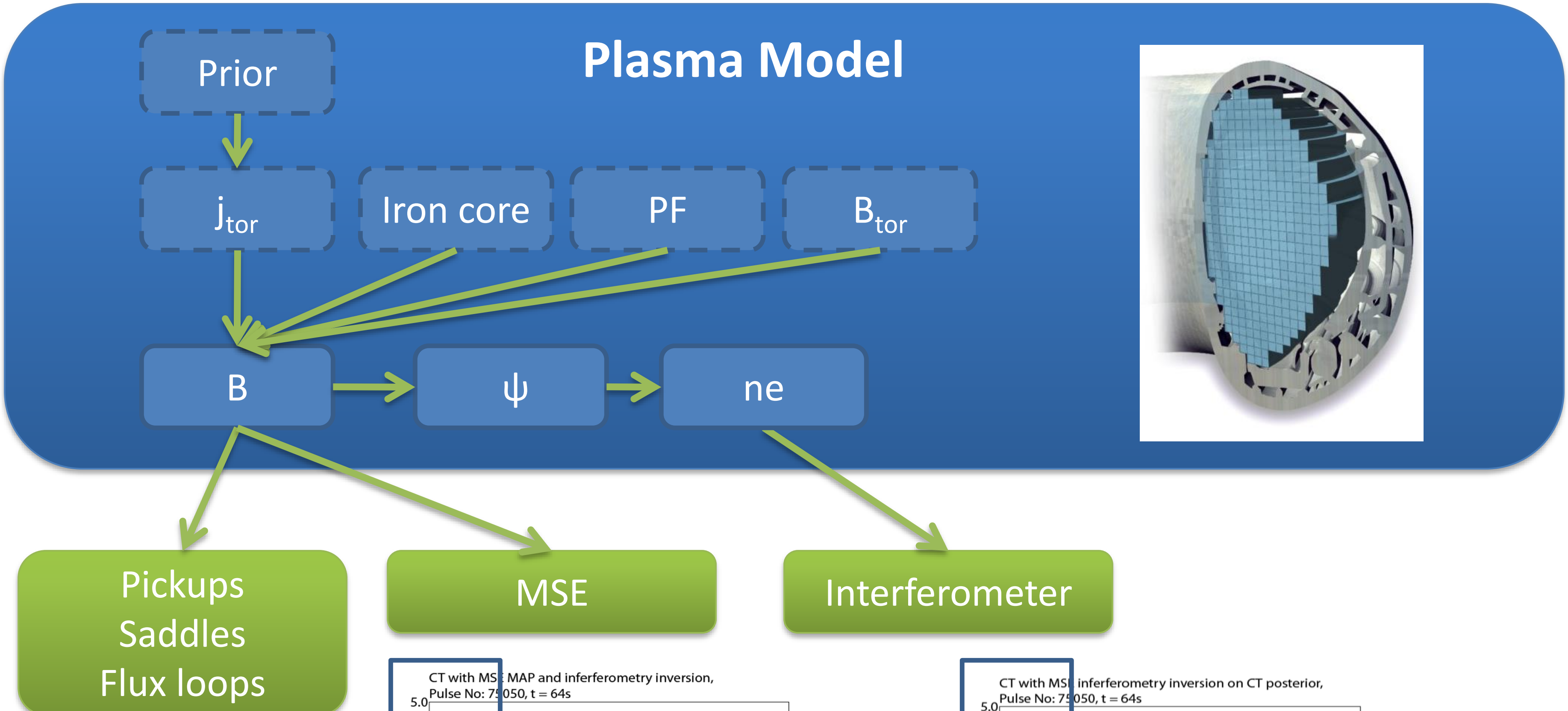
$$p(\overline{I} \mid \overline{D}^{Mag}) = \frac{1}{(2\pi)^{N_I/2} \left|\overline{\overline{\Sigma}}\right|^{1/2}} \exp\left(-\frac{1}{2}(\overline{I} - \overline{m})^T \overline{\overline{\Sigma}}^{-1} (\overline{I} - \overline{m})\right)$$

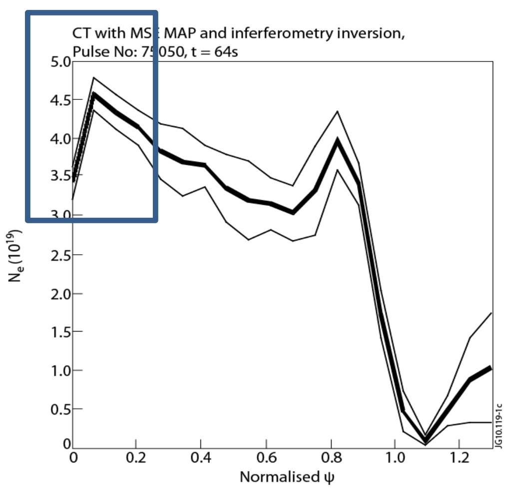
MAP
(MAximum Posterior)

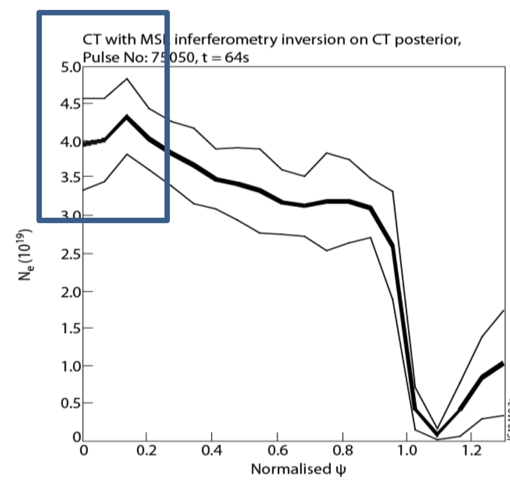
where $\overline{m} = (\overline{\overline{M}}^T \overline{\overline{\Sigma}}_D^{-1} \overline{\overline{M}} + \overline{\overline{\Sigma}}_I^{-1})^{-1} \overline{\overline{M}}^T \overline{\overline{\Sigma}}_D^{-1} (\overline{D}^{Mag} - \overline{C})$

$$\overline{\overline{\Sigma}} = (\overline{\overline{M}}^T \overline{\overline{\Sigma}}_D^{-1} \overline{\overline{M}} + \overline{\overline{\Sigma}}_I^{-1})^{-1}$$

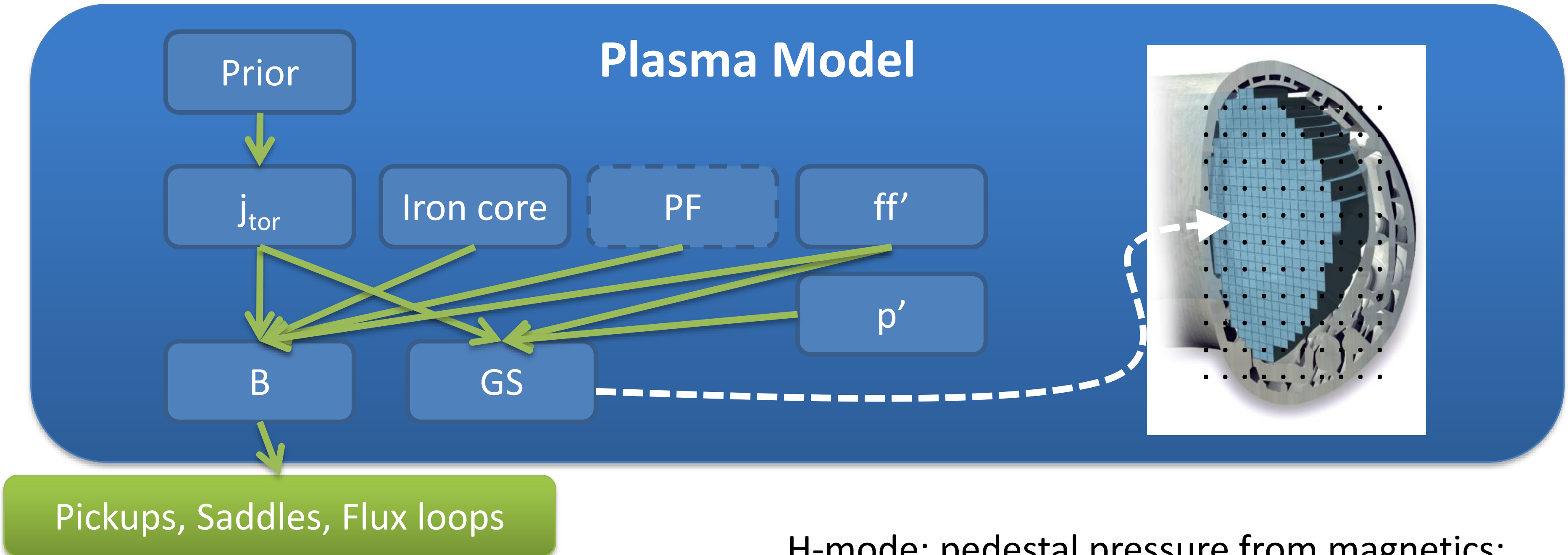
# $n_e$ with uncertainty in mapping



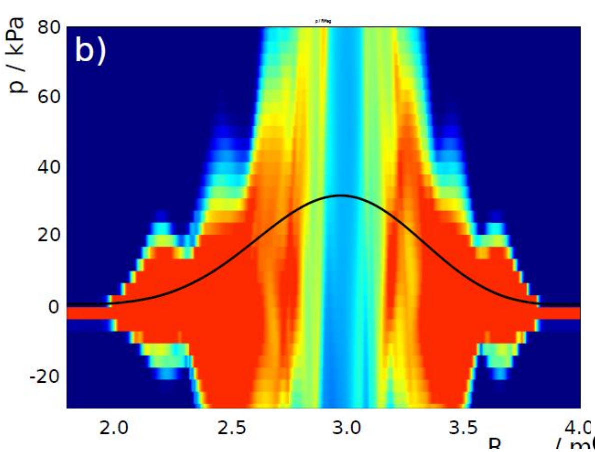



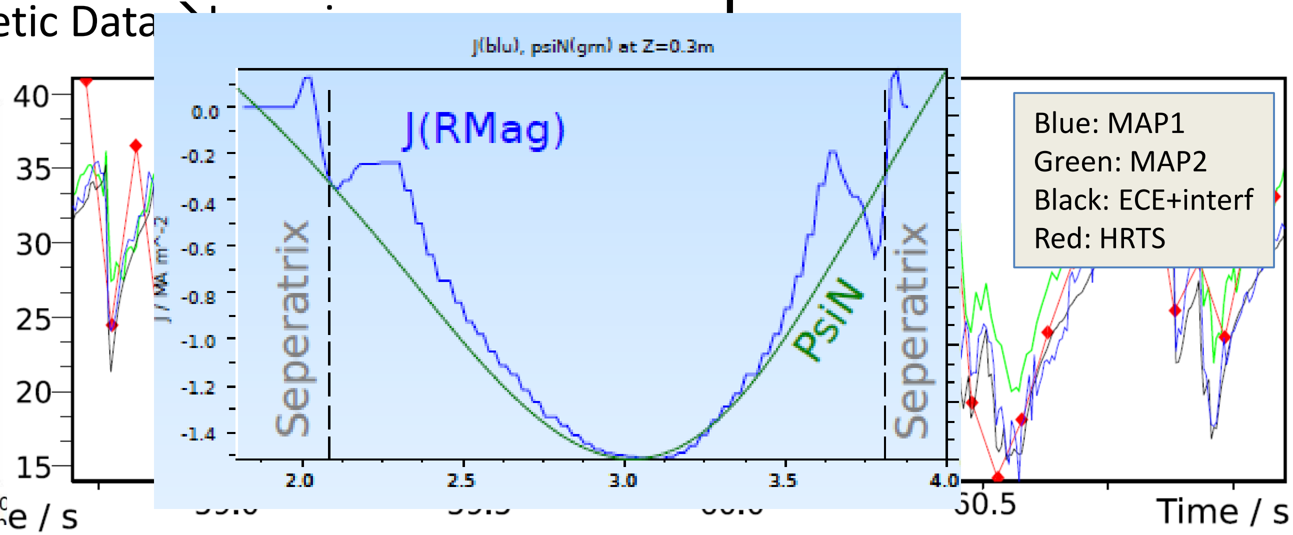
# $j_{tor}$ ,p,with force balance assumption




H-mode: pedestal pressure from magnetics:

L-mode: P(p): EFIT → Synthetic Data →

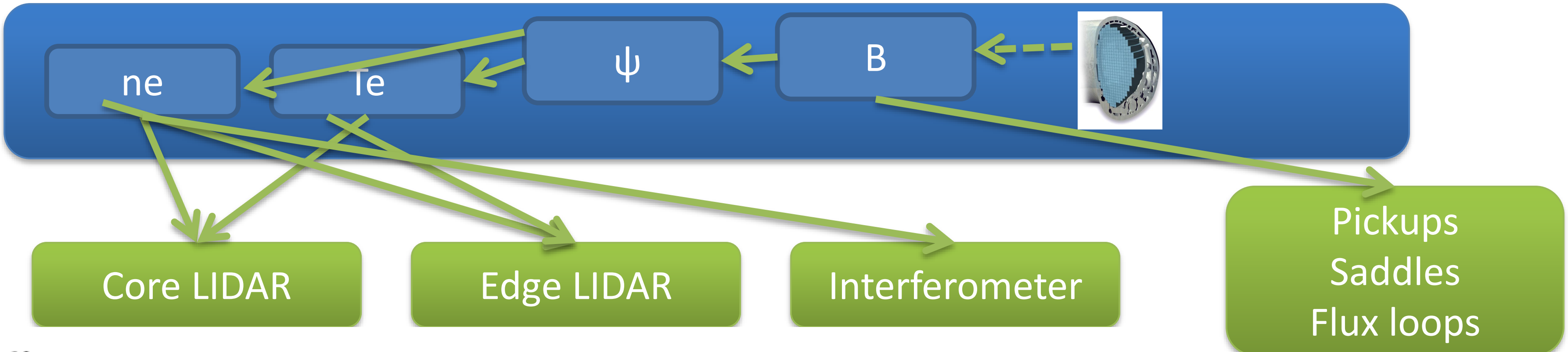


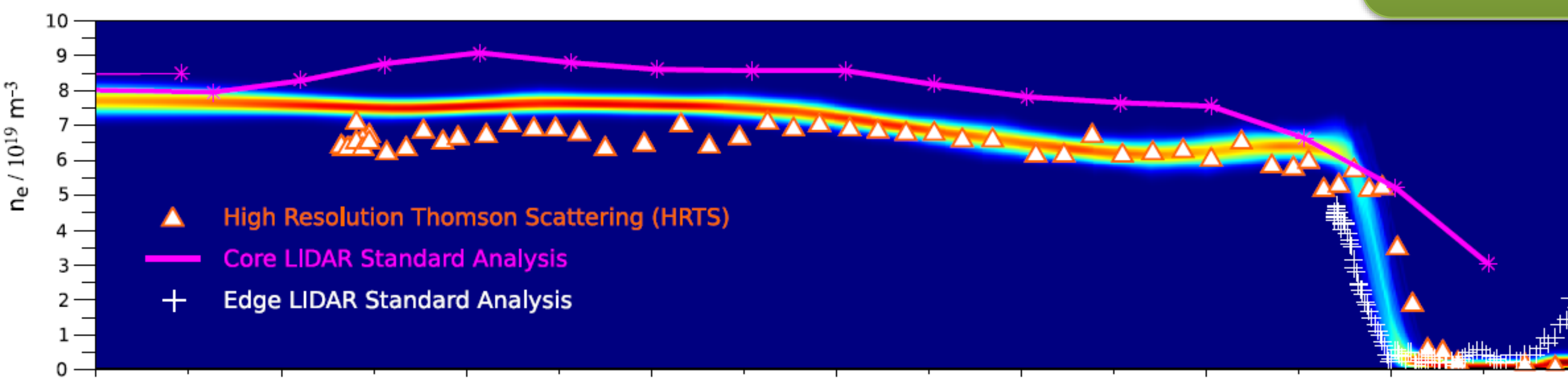



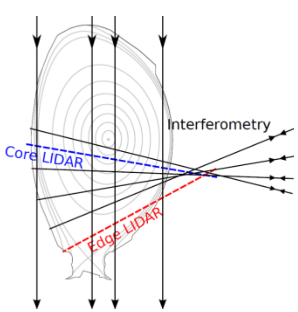

O Ford, PhD thesis 2010, Imperial College London

Pulse 78601

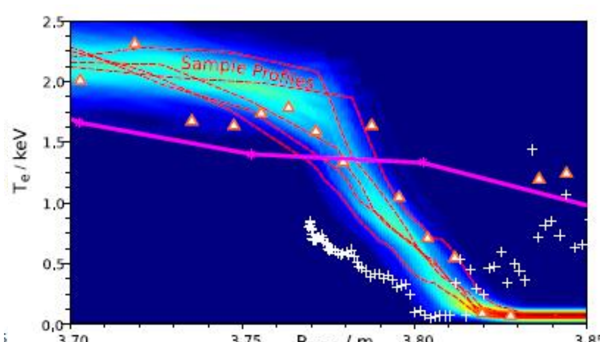
# $n_e$ and $T_e$ profiles from combined core LIDAR, edge LIDAR and interferometer



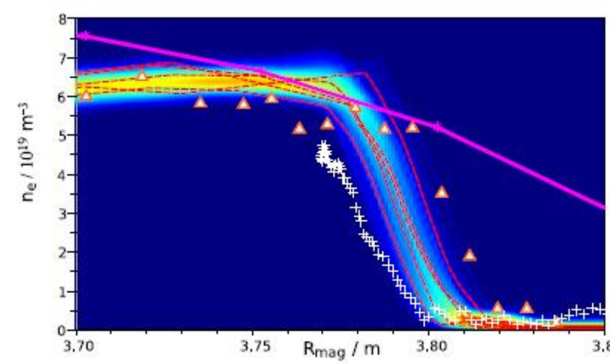




Te:




ne:




O Ford, PhD thesis 2010

# Interim summary and outlook

- **Summary:**
  - Uncertainties on flux surfaces, combined with inversions
  - Bayesian exploration of uncertainties in equilibrium inference
  - Greatly improved accuracy on ne and Te profile inference at JET
  - In total about 10 diagnostic systems modelled in Minerva.
- **Ongoing Minerva projects**:
  - ECE diagnostics (PhD of Stefan Schmuck, KTH, Sweden)
  - Soft-X and Bolometry (PhD of Dong Li, Greifswald University)
  - MAST modelling: Thomson scattering, CX, gen. force balance (with ANU, Australia).
  - Reflectometry (with Antoine Sirinelli, JET),
  - Bremsstrahlung and He-beam (with Maciej Krychowiak, IPP)

# Excursion:
# Inference on infinite dimensional functions

## There are two quite different classes of inversion problems:

1. We have a real underlying parameterised physics-based model.
Example: fitting line spectra to (Gaussian, Voigt etc) line shapes.

2. We want to find an underlying, in principle infinite-dimensional, function, and we do not really have any specific parameterisation we believe represent the underlying physics process.
Example: tomographic inversion, density profiles etc

We tend to use methods developed for (1) also for problems belonging to class (2)!

... by choosing some "flexible" parameterisation such as grids, polynomials/splines etc. then regularizing the solution. But what kind of "smoothness" does a polynomial of A given order express in comparison to a set of Gaussian basis functions or something else?

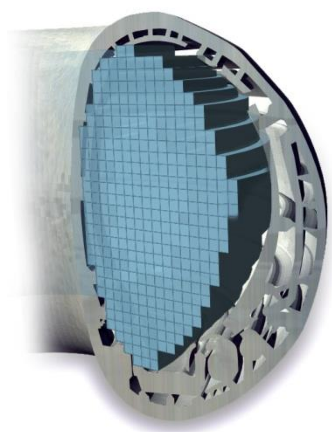
# Difference between prior constraints and parametric constraints.

Parameters “freeze” in or dictates the exact form of all possible functions we could observe, at all positions.

Prior constraints, on the other hand, provide a “loose” probabilistic constraint that can be overridden by the data.

... so should we not prefer to “guide” the inference on infinite dimensional functions by prior constraints rather than parametric constraints?

Could

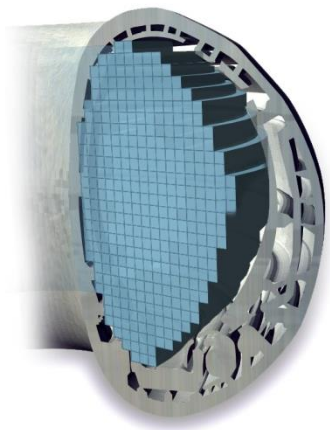


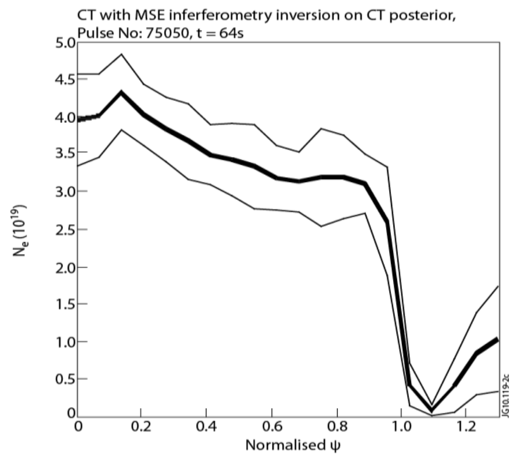
+ prior on discrete beam currents

Be replaced by $j_{tor}(R,Z)$ + prior directly on the *function* $j_{tor}(R,Z)$

# Can we do non-parametric tomography?

What happens if we let the density of beams, or density of profile points -> Infinity?






..the prior mean becomes a function over R,Z and the prior covariance matrix over beams becomes a continous *covariance function* defining the prior covariance between every pair of (R,Z):

$$k(R, Z)=...$$

... the posterior can be calculated directly and gives another posterior mean and a posterior covariance function.

Both prior and posterior are now *Gaussian random processes* rather than pdf:s over discrete parameters.

J Svensson, *Non-parametric tomography using Gaussian processes*, forthcoming

# Non-parametric tomography

$$\boldsymbol{m}_{f*} = \boldsymbol{m}_{\mathbf{pf}*} + \boldsymbol{K}_{*L}(\boldsymbol{K}_{LL} + \boldsymbol{\Sigma}_y)^{-1}(\boldsymbol{y} - \boldsymbol{f}_L)$$

$$\boldsymbol{\Sigma}_{f*} = \boldsymbol{K}_{**} - \boldsymbol{K}_{*L}(\boldsymbol{K}_{LL} + \boldsymbol{\Sigma}_y)^{-1}\boldsymbol{K}_{L*}$$

$$(\boldsymbol{K}_{**})_{kl} = k(\boldsymbol{x}_*^k, \boldsymbol{x}_*^l)$$

$$(\boldsymbol{K}_{L*})_{kl} = \int_{los\ k} k(\boldsymbol{x}, \boldsymbol{x}_*^l)d\boldsymbol{x}$$

$$(\boldsymbol{K}_{*L})_{kl} = \int_{los\ l} k(\boldsymbol{x}_*^k, \boldsymbol{x})d\boldsymbol{x}$$

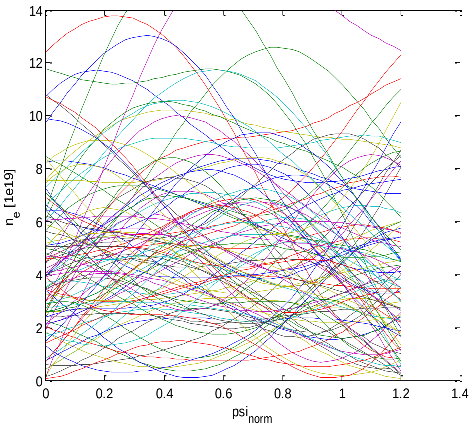
$$(\boldsymbol{K}_{LL})_{kl} = \int_{los\ k}\int_{los\ l} k(\boldsymbol{x}, \boldsymbol{x}')d\boldsymbol{x}d\boldsymbol{x}'$$

*k* – prior covariance function
$f_L$ – predicted observations at prior function mean
$\boldsymbol{x}_*$ – points (R,Z) where we want to evaluate inferred function
y – los observations

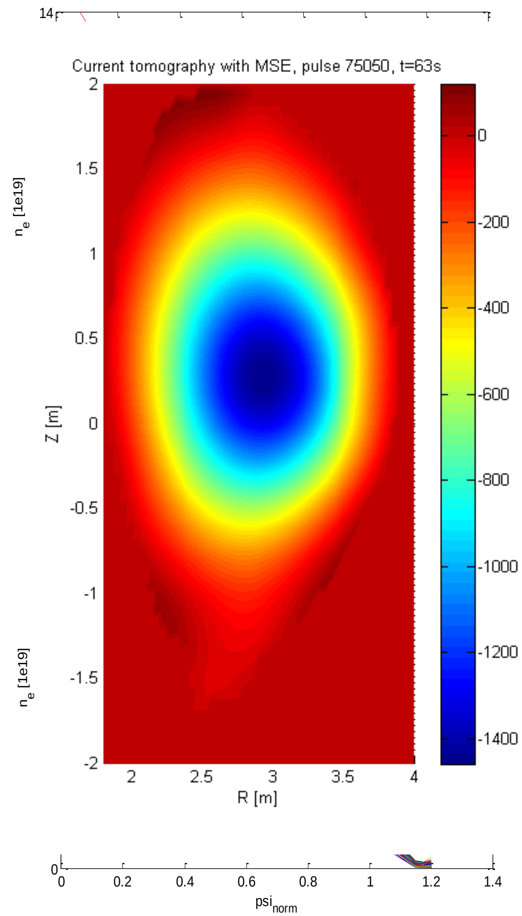
J Svensson, *Non-parametric tomography using Gaussian processes*, forthcoming

# Example: non-parametric interferometry inversion

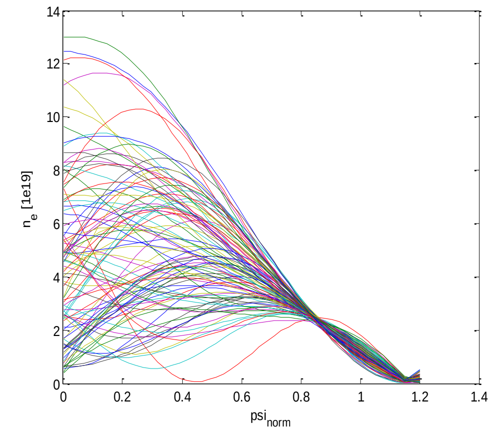
Samples from prior process

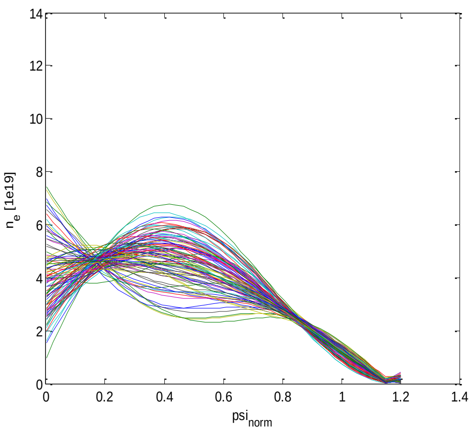



1 los observation



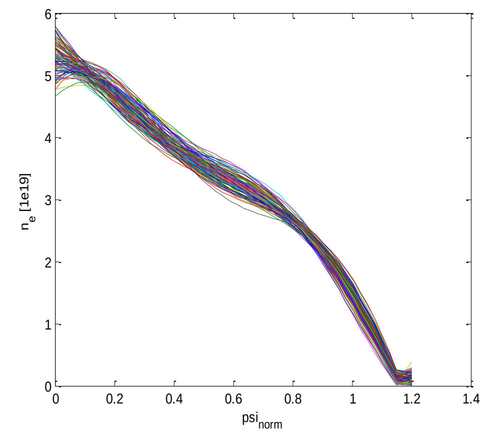

2 los observations




3 los observations

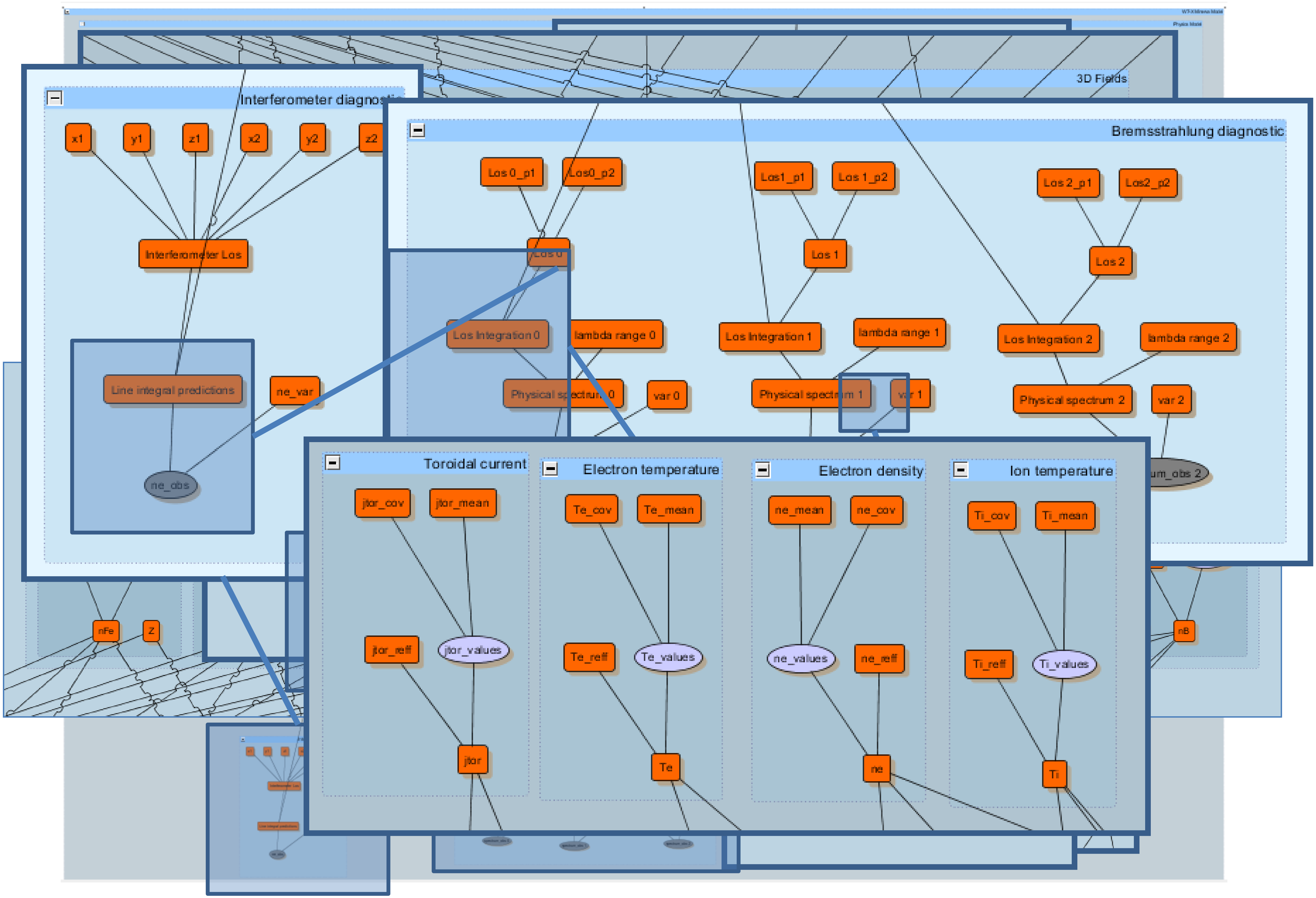
7 los observation

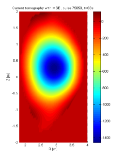



7 los observations, zoomed

J Svensson, *Non-parametric tomography using Gaussian processes*, forthcoming

# Minerva for Wendelstein 7-X




Current tomography with MSE, pulse 75060, t=63s
Z [m]
R [m]
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
2
2.5
3
3.5
4
0
-200
-400
-600
-800
-1000
-1200
-1400

# Thanks!