Reassessment of impurity transport coefficients in Alcator C-Mod

M.A. Chilenski, M. Greenwald, Y. Marzouk*, J.E. Rice, A.E. White

MIT Plasma Science and Fusion Center/Alcator C-Mod
*MIT Aero/Astro, Uncertainty Quantification Group

November 2, 2016

This material is based upon work conducted using the Alcator C-Mod tokamak, a DOE Office of Science user facility. This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Fusion Energy Sciences under Award Number DE-FC02-99ER54512. This material is based upon work supported in part by the U.S. Department of Energy Office of Science Graduate Research Fellowship Program (DOE SCGF), made possible in part by the American Recovery and Reinvestment Act of 2009, administered by ORISE-ORAU under contract number DE-AC05-06OR23100. The XEUS and LoWEUS spectrometers were developed at the LLNL EBIT lab. Work at LLNL was performed under the auspices of the US DOE under contract DE-AC52-07NA-27344. Some of the computations using STRAHL were carried out on the MIT PSFC parallel AMD Opteron/Infiniband cluster Loki.

Measuring impurity transport and testing simulations: making sure we know what we think we know

Impurity transport coefficients D, V are often used in validation metrics

$$\frac{\partial n_Z}{\partial t} = -\nabla \cdot \Gamma_Z + Q_Z$$

Model impurity flux Γ_Z with diffusion coefficient D, convective velocity V:

$$\Gamma_Z = -D\nabla n_Z + V n_Z$$

D, V are often used to validate impurity transport simulations:
Important to measure D, V properly to have a strong test of the code.

Measuring impurity transport and testing simulations: making sure we know what we think we know

$$\Gamma_Z = -D\nabla n_Z + V n_Z$$

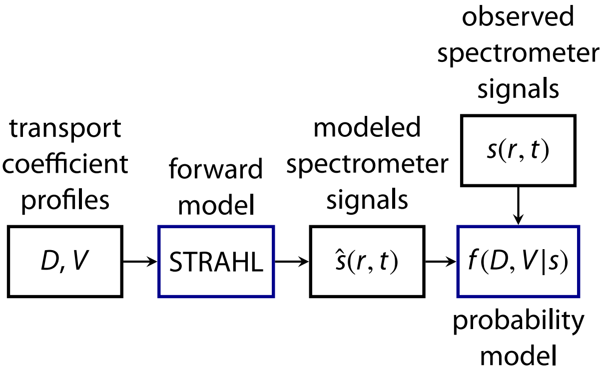
Current approaches for measuring *D*, *V* have considerable shortcomings

- Error bars not consistent with intuition.
- Different starting points give different results:
  - Multiple solutions?
  - Broad region of acceptable solutions?

New approach fixes these issues

- Use advanced inference techniques to find $D(r)$, $V(r)$.
- Rigorous selection of level of complexity in $D(r)$, $V(r)$ is critical.

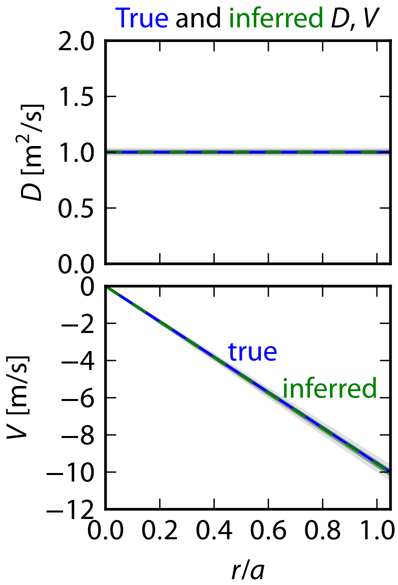
Inferring impurity transport coefficients: a nonlinear inverse problem



- Inject impurity with laser blow-off
- Can only observe s, want to know D, V.
- Need to assume a parameterization for $D(r), V(r)$.
- Two steps:
  1. Find D, V consistent with s for given parameterization.
  2. Find best parameterization.
- Do both steps simultaneously with **MultiNest** [Feroz MNRAS 2008, 2009]: Bayesian inference algorithm.

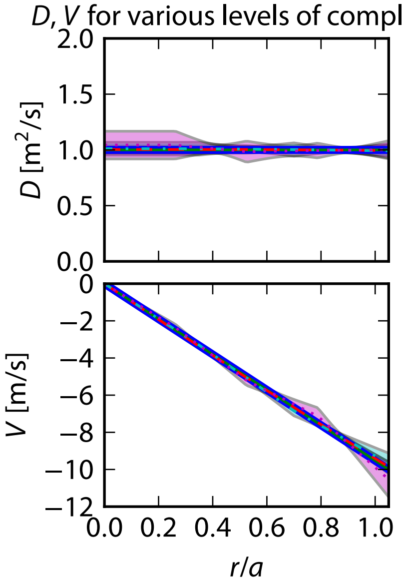
MultiNest successfully reconstructs simple *D*, *V* profiles

(synthetic data)

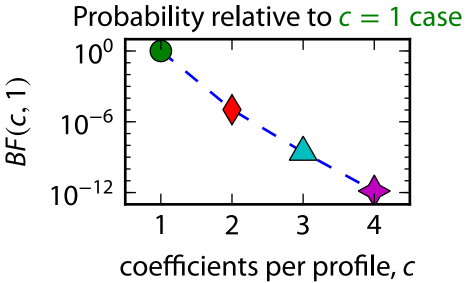


- Five local measurements, uniformly spaced over $0 \leq r/a \leq 1$
- $\Delta t = 6$ ms, 5% noise
- Have also tested with 32 x-ray spectrometer chords: can handle tomographic inversion

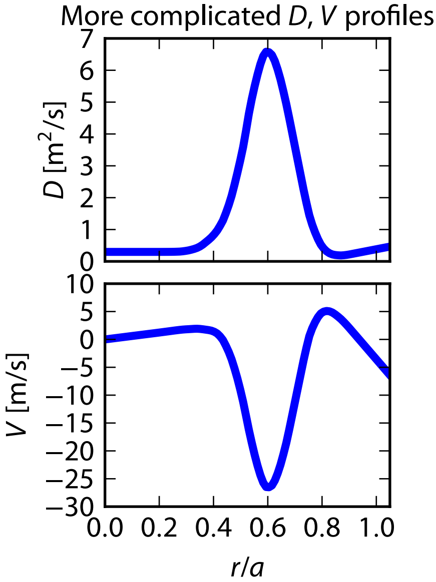
MultiNest successfully determines how many spline coefficients to use



- MultiNest estimates the evidence $f_{s|c}(s|c)$: probability of observing the data given c.
- Ran with various numbers of free parameters: correctly selected $c = 1$ case.
- Bayes factors: $BF(c, 1) = f_{s|c}(s|c)/f_{s|1}(s|1)$

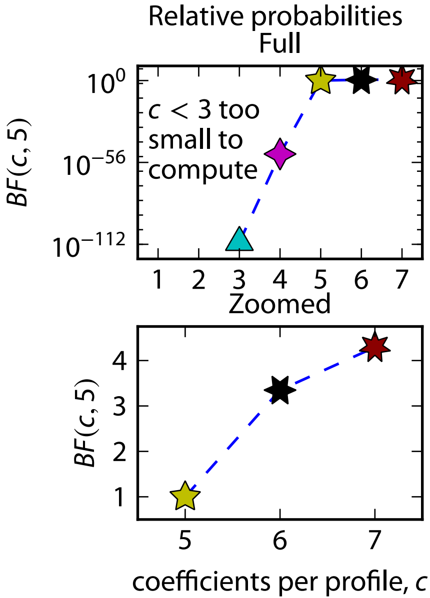


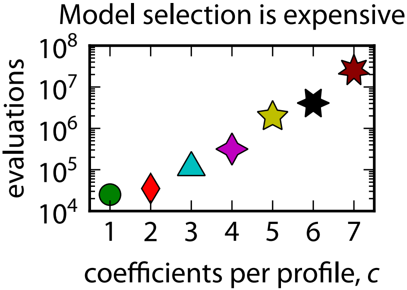
Testing with more complicated synthetic data



- Need to test with data representative of reality.
- Used result from [Howard NF 2012, Chilenski NF 2015] as true profile.
- Realistic diagnostic configuration:
  - 32 x-ray spectrometer chords (Ca^{18+}), 6 ms time resolution, 5% noise
  - 2 VUV spectrometer chords (Ca^{17+}, Ca^{16+}), 2 ms time resolution, 5% noise

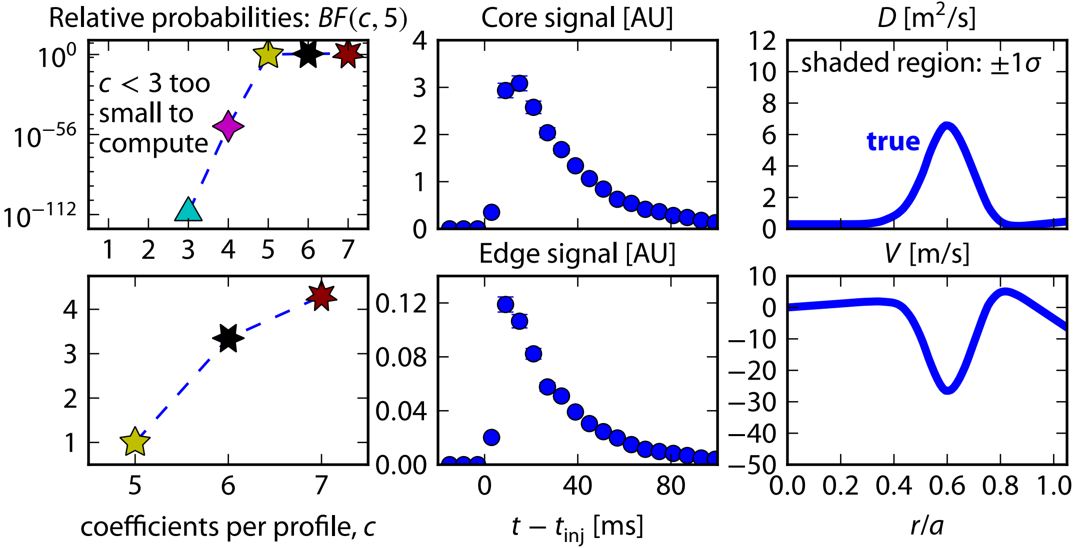
More complicated synthetic data pose a challenge



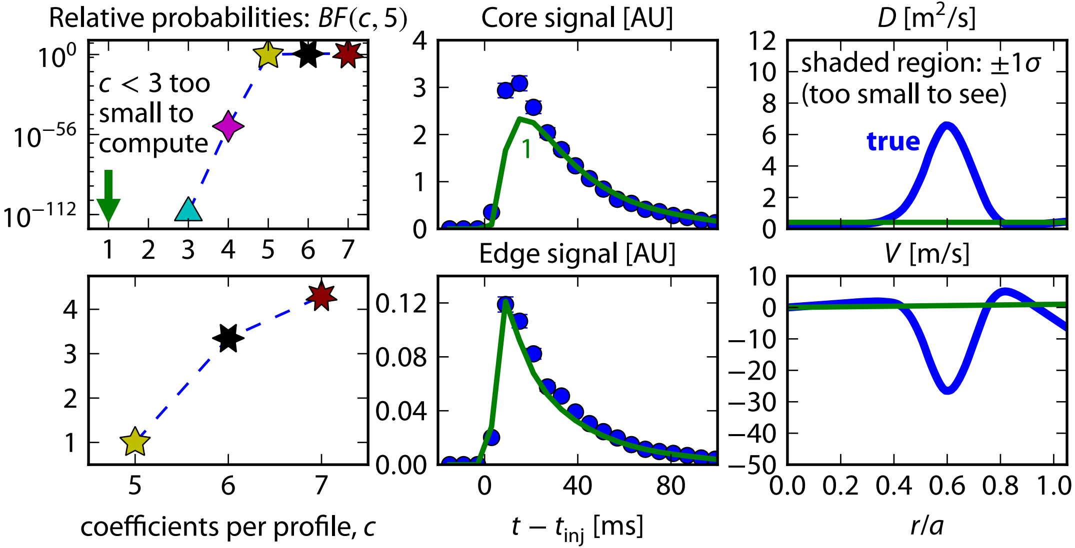


- 7 coefficient case took 7000 CPU-hours = 15 wall-clock days!
- Need to speed up model, deploy on cluster to make this practical.
- (Recall: $BF(c, 5) = f_{s|c}(s|c)/f_{s|5}(s|5)$)

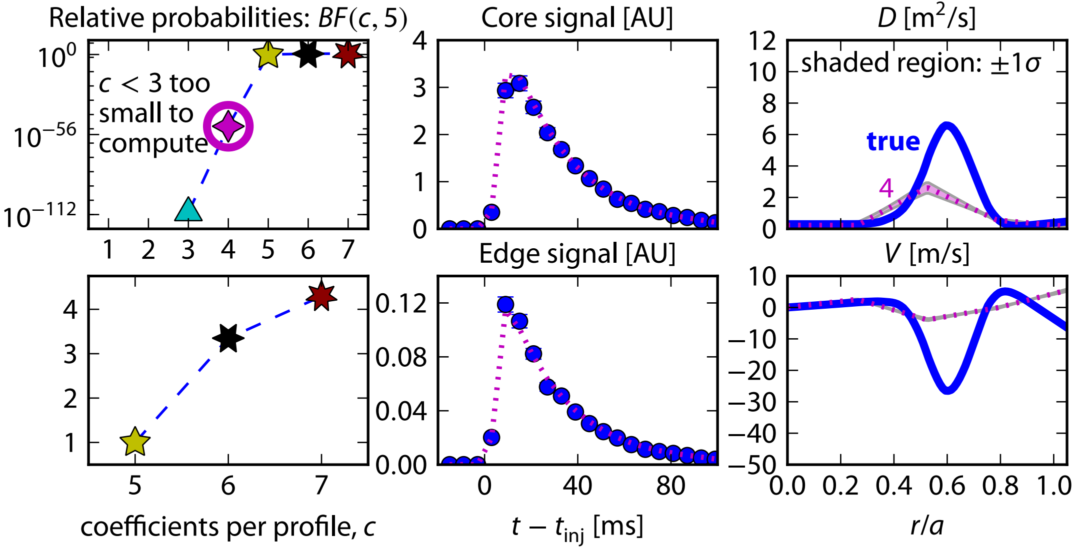
Results only resemble true profile when a minimum level of complexity is obtained…despite “good” match to data: “eyeballing” does not work!



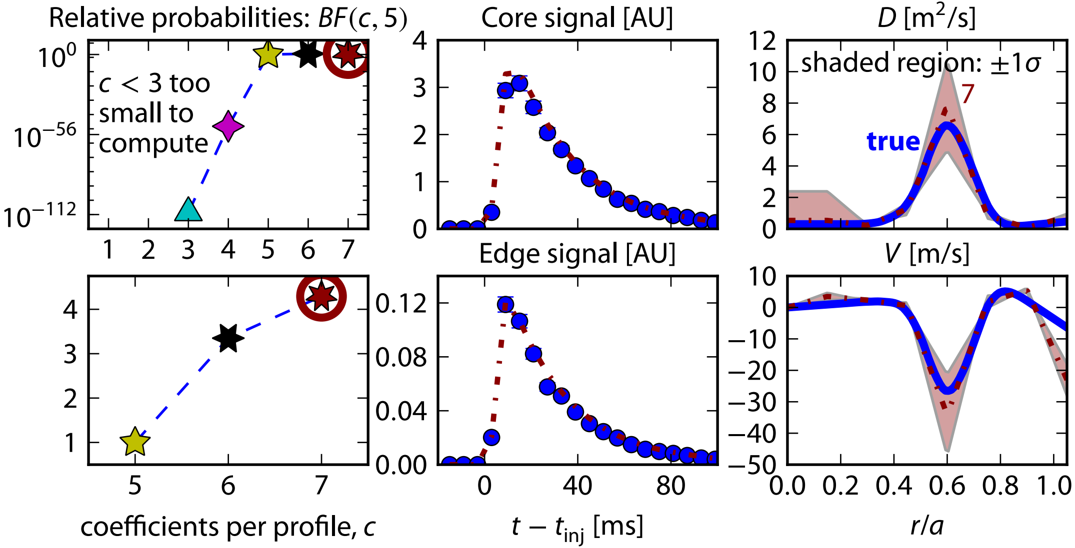
Results only resemble true profile when a minimum level of complexity is obtained…despite "good" match to data: "eyeballing" does not work!



Results only resemble true profile when a minimum level of complexity is obtained…despite "good" match to data: "eyeballing" does not work!



Results only resemble true profile when a minimum level of complexity is obtained…despite "good" match to data: "eyeballing" does not work!



Getting *D*, *V* right requires careful statistical analysis

Conclusions

- ***Need*** to select right level of complexity:
  - Can appear to match data well while not matching the real *D* and *V* at all.
  - New approach rigorously selects the most likely model.
- Can also estimate/verify diagnostic requirements.
- Validation of impurity transport simulations is still an open question, but we have a path forward.

Future work

- Speed up, parallelize analysis.
- Improve handling of sawteeth.
- Develop more efficient ways of selecting basis functions.
- Re-assess previous results from Alcator C-Mod and other tokamaks.

Additional details are in my PhD thesis: markchil.github.io/pdfs/thesis.pdf

Open-source software: github.com/markchil/bayesimp

Backup slides

Inferring impurity transport coefficients: a nonlinear inverse problem

inject impurity → diffusion, convection move impurity → observe signals

Finding D, V with Bayesian inference

$$\underbrace{f_{D,V|s}(D,V|s)}_{\text{posterior distribution}} = \frac{\overbrace{f_{s|D,V}(s|D,V)}^{\text{likelihood: from STRAHL + data}} \; \overbrace{f_{D,V}(D,V)}^{\text{prior distribution } (D>0,\, V(0)=0, \text{ etc.})}}{\underbrace{f_s(s)}_{\text{evidence}}}$$

- f: probability density function
- s: measured signals
- D, V: parameters describing radial profiles of diffusion, convection

Parameter estimation: Find D, V: characterize $f_{D,V|s}(D,V|s)$.

Model selection: Find best way of parameterizing D, V: maximize $f_s(s)$.

$\mathcal{M}$: functional form (model) used to parameterize $D(r)$, $V(r)$.

Use **MultiNest** [Feroz MNRAS 2008, 2009]: samples $f_{D,V|s}(D,V|s)$ *and* estimates $f_s(s)$.

Inferring impurity transport coefficients: a nonlinear inverse problem

inject impurity → diffusion, convection move impurity → observe signals

Finding *D*, *V* with Bayesian inference

$$\underbrace{f_{D,V|s,\mathcal{M}}(D,V|s,\mathcal{M})}_{\text{posterior distribution}} = \frac{\overbrace{f_{s|D,V,\mathcal{M}}(s|D,V,\mathcal{M})}^{\text{likelihood: from STRAHL + data}}\ \overbrace{f_{D,V|\mathcal{M}}(D,V|\mathcal{M})}^{\text{prior distribution } (D>0,\ V(0)=0,\text{ etc.})}}{\underbrace{f_{s|\mathcal{M}}(s|\mathcal{M})}_{\text{evidence}}}$$

- f: probability density function
- s: measured signals
- D, V: parameters describing radial profiles of diffusion, convection

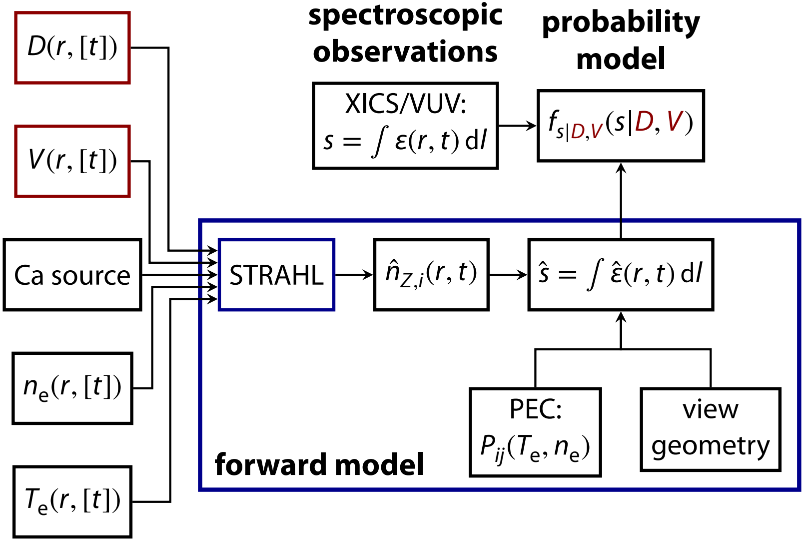
Parameter estimation: Find D, V: characterize $f_{D,V|s,\mathcal{M}}(D,V|s,\mathcal{M})$.

Model selection: Find best way of parameterizing D, V: maximize $f_{s|\mathcal{M}}(s|\mathcal{M})$.

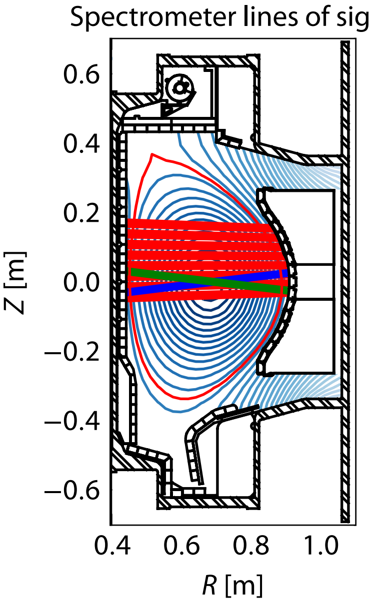
$\mathcal{M}$: functional form (model) used to parameterize $D(r)$, $V(r)$.

Use **MultiNest** [Feroz MNRAS 2008, 2009]: samples $f_{D,V|s,\mathcal{M}}(D,V|s,\mathcal{M})$ *and* estimates $f_{s|\mathcal{M}}(s|\mathcal{M})$.

Inferring impurity transport coefficients is a difficult inverse problem



Spectrometer chords on Alcator C-Mod



HiReX-SR X-ray imaging crystal spectrometer (XICS) with 32 chords split into 8 groups. Views He-like Ca (0.32 nm) with 6 ms time resolution.

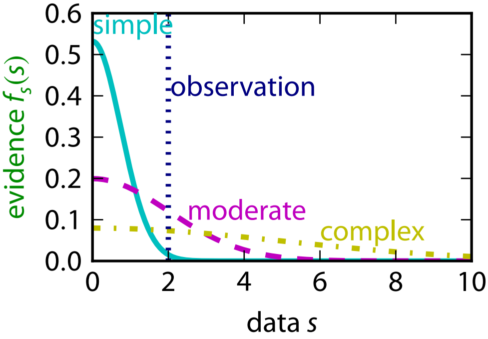
XEUS Vacuum ultraviolet (VUV) spectrometer with one chord. Views Li-like Ca (1.9 nm) with 2 ms time resolution.

LoWEUS Vacuum ultraviolet (VUV) spectrometer with one chord. Views Be-like Ca (19 nm) with 2 ms time resolution.

Model selection using the evidence (a.k.a. the marginal likelihood)

$$\underbrace{f_{D,V|s}(D,V|s)}_{\text{posterior}} = \frac{\overbrace{f_{s|D,V}(s|D,V)}^{\text{likelihood}}\,\overbrace{f_{D,V}(D,V)}^{\text{prior}}}{\underbrace{f_s(s)}_{\text{evidence}}}$$

$$f_s(s) = \int f_{s|D,V}(s|D,V) f_{D,V}(D,V)\,\mathrm{d}D\,\mathrm{d}V$$



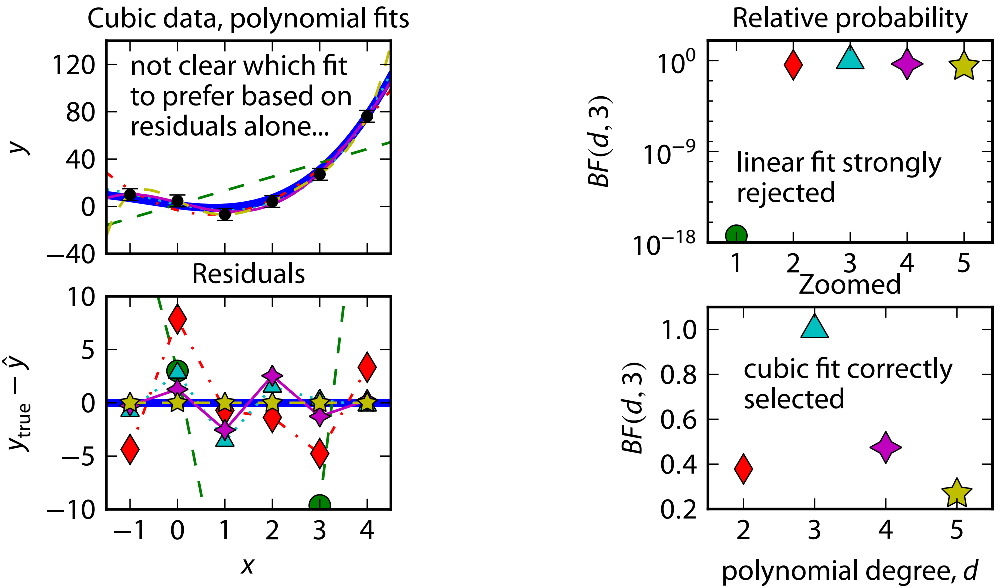
Tradeoff between goodness of fit, complexity [Schwarz AS 1978]:

$$\ln f_s(s) \approx \underbrace{\ln f_{s|\hat{\theta}}(s|\hat{\theta})}_{\text{goodness-of-fit}} - \underbrace{\frac{d}{2}\ln N}_{\text{complexity}}$$

(d is number of parameters, N number of datapoints)

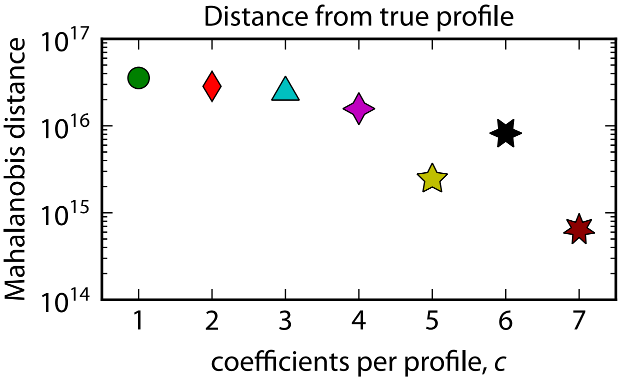
- $f_s(s)$ is maximized by model with "right" level of complexity.
- Simple models can only explain a few data sets, low evidence for most s.
- Complex models can explain many data sets, any given s has low probability.

Simple example: fitting noisy data from $y = x^3 + 2x^2 - 5x + 1$

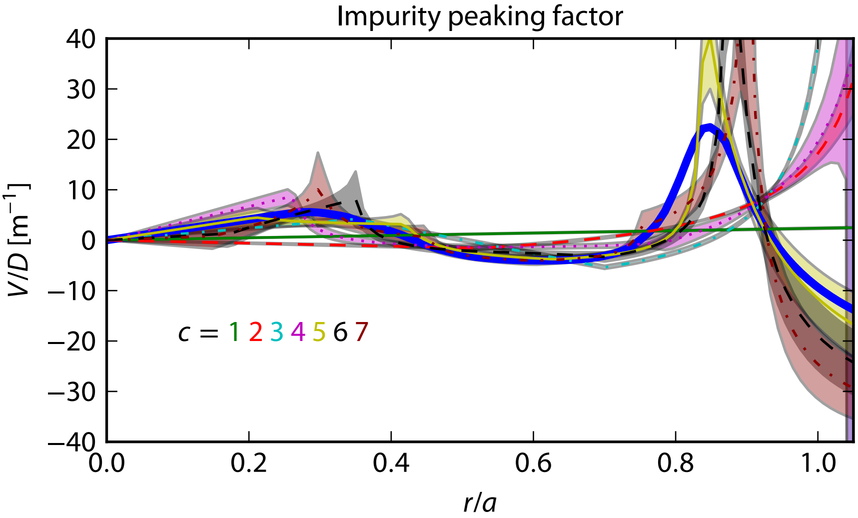


More complex models match the true profiles better

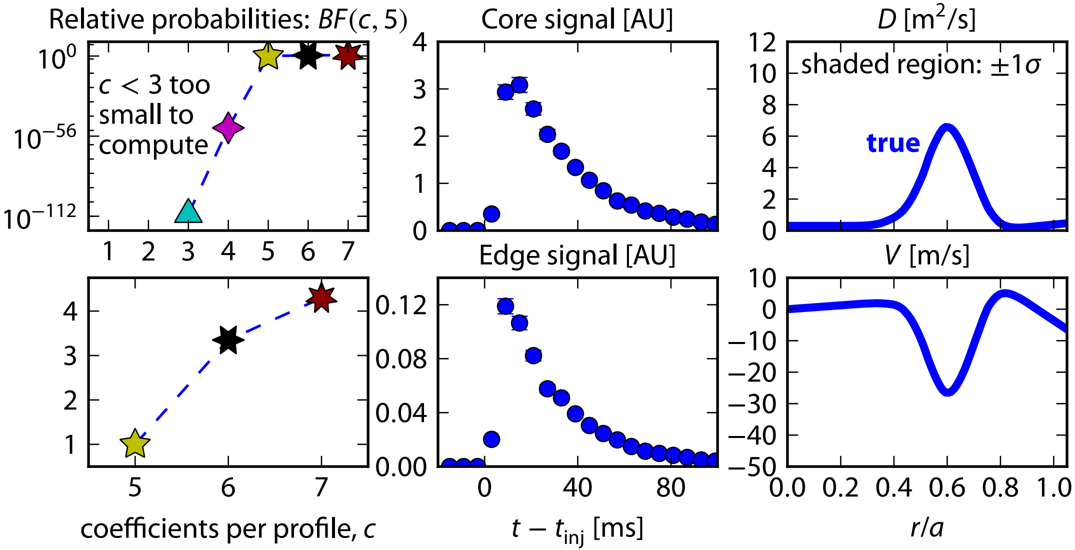
Mahalanobis distance: $M = \sqrt{(\boldsymbol{T}_{\text{true}} - \boldsymbol{\mu})^{\top} \boldsymbol{\Sigma}^{-1} (\boldsymbol{T}_{\text{true}} - \boldsymbol{\mu})}$, $\boldsymbol{T} = [D, V]$



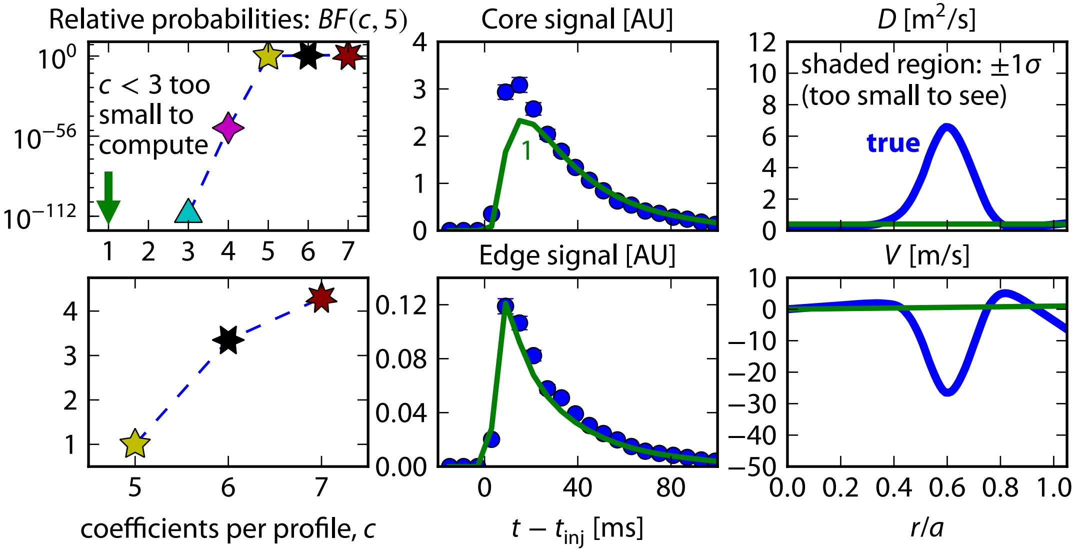
V/D is captured for $r/a \lesssim 0.7$, $c \geq 5$



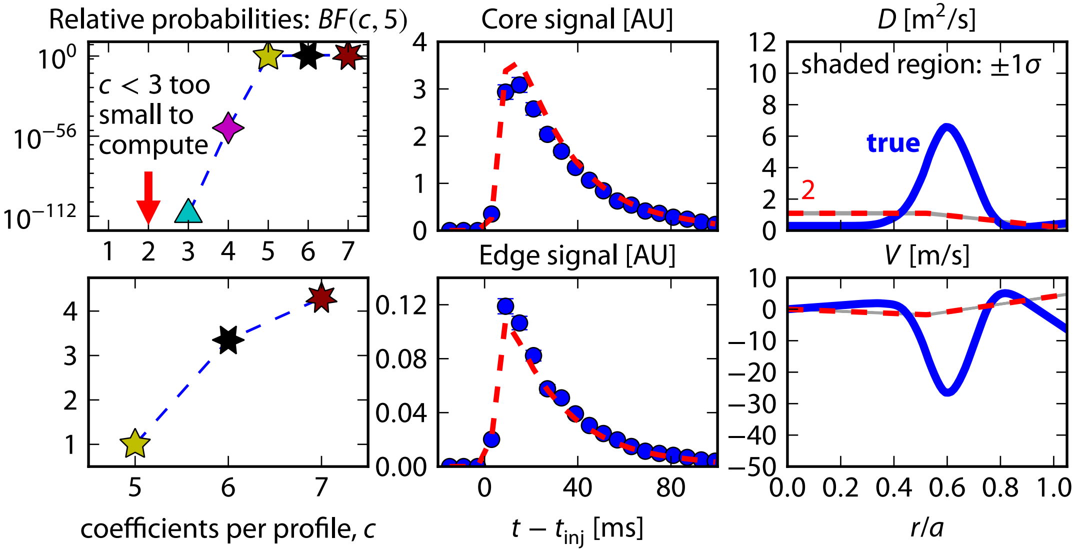
Results only resemble true profile when a minimum level of complexity is obtained…despite "good" match to data: "eyeballing" does not work!



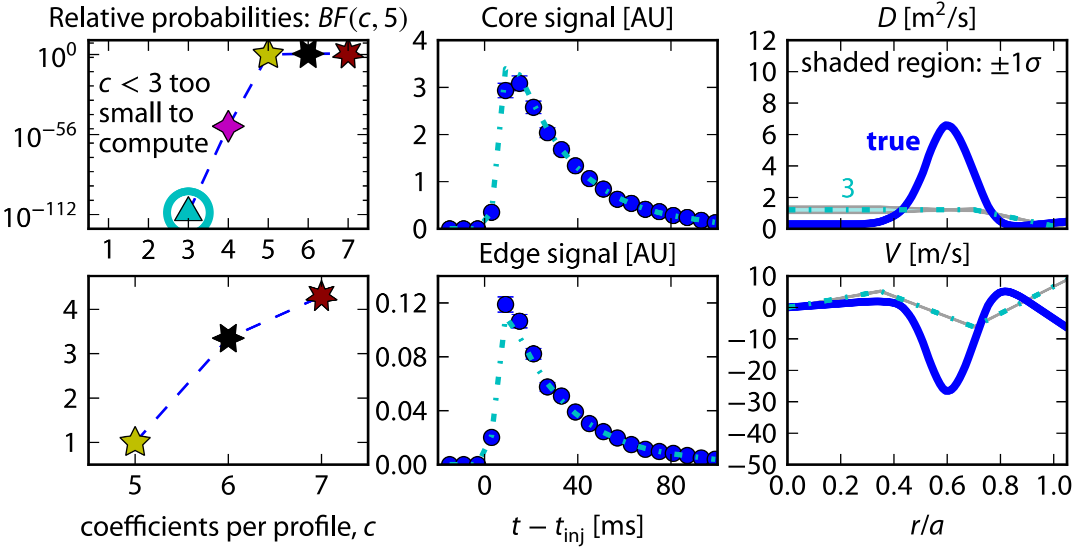
Results only resemble true profile when a minimum level of complexity is obtained…despite "good" match to data: "eyeballing" does not work!



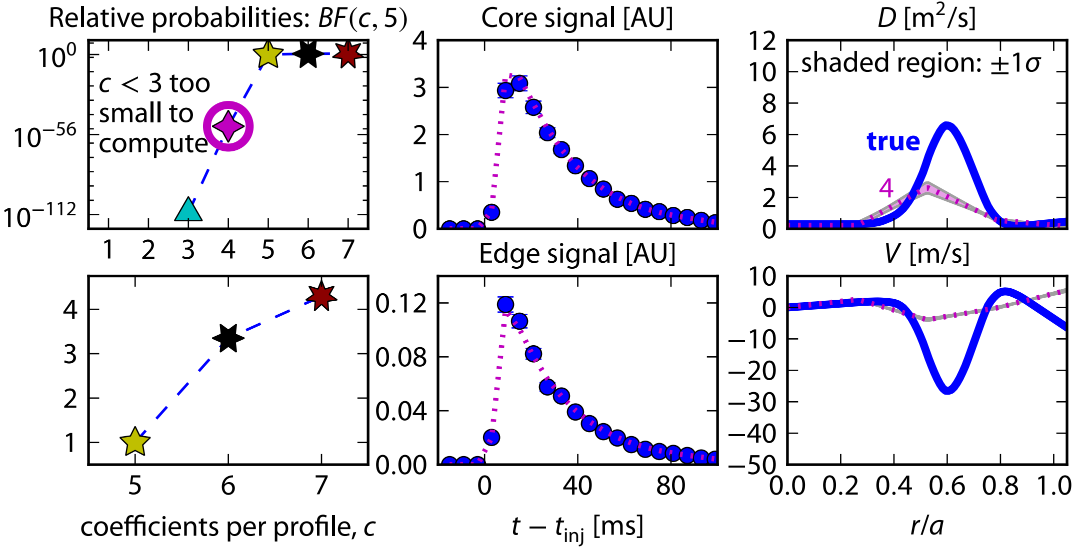
Results only resemble true profile when a minimum level of complexity is obtained…despite “good” match to data: “eyeballing” does not work!



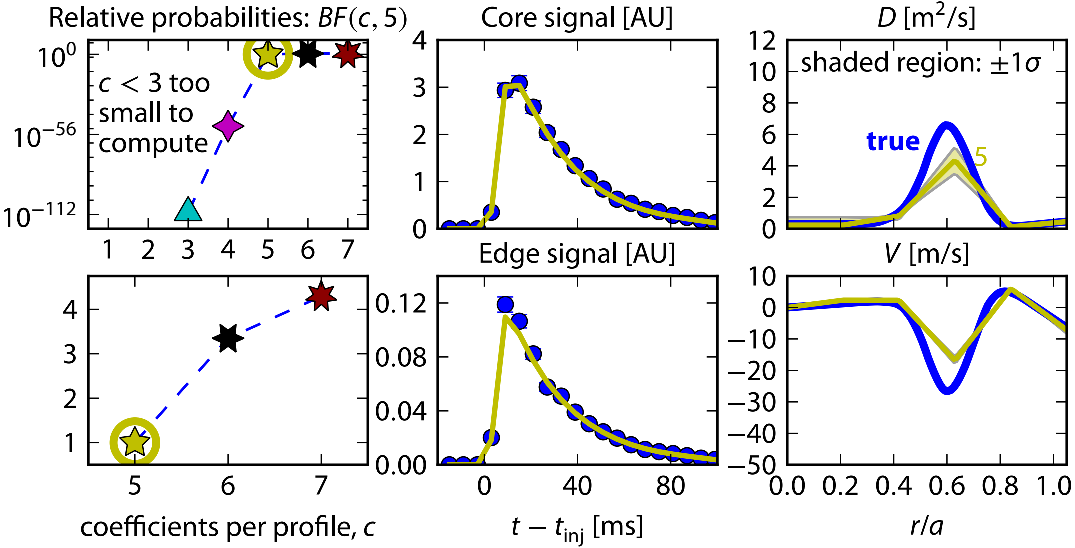
Results only resemble true profile when a minimum level of complexity is obtained…despite "good" match to data: "eyeballing" does not work!



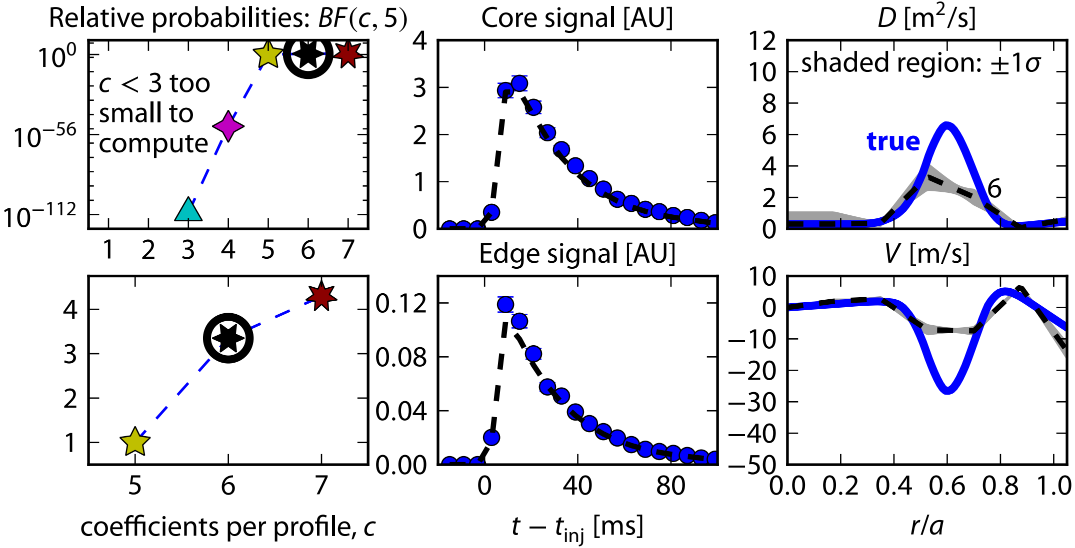
Results only resemble true profile when a minimum level of complexity is obtained…despite "good" match to data: "eyeballing" does not work!



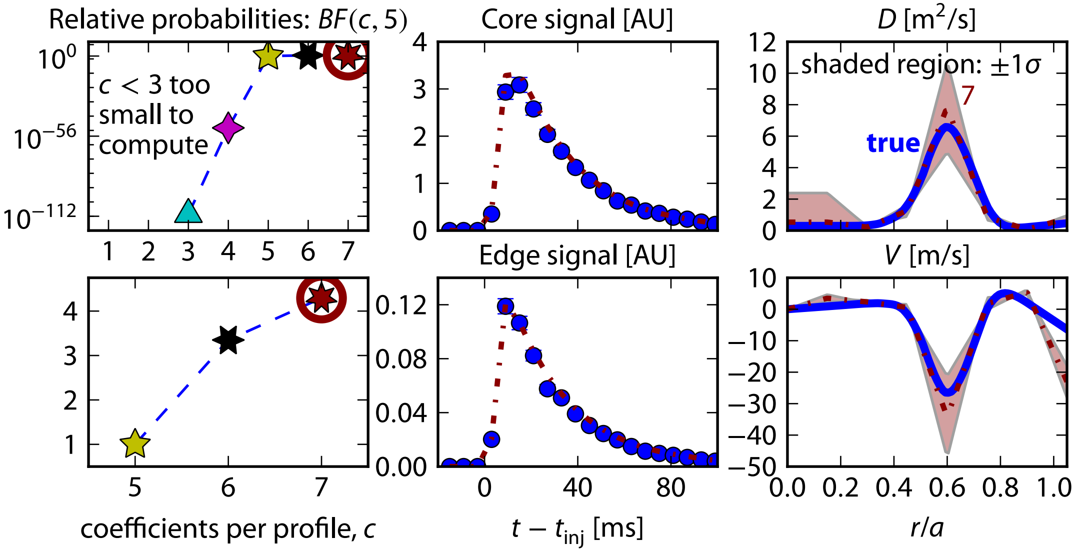
Results only resemble true profile when a minimum level of complexity is obtained…despite "good" match to data: "eyeballing" does not work!



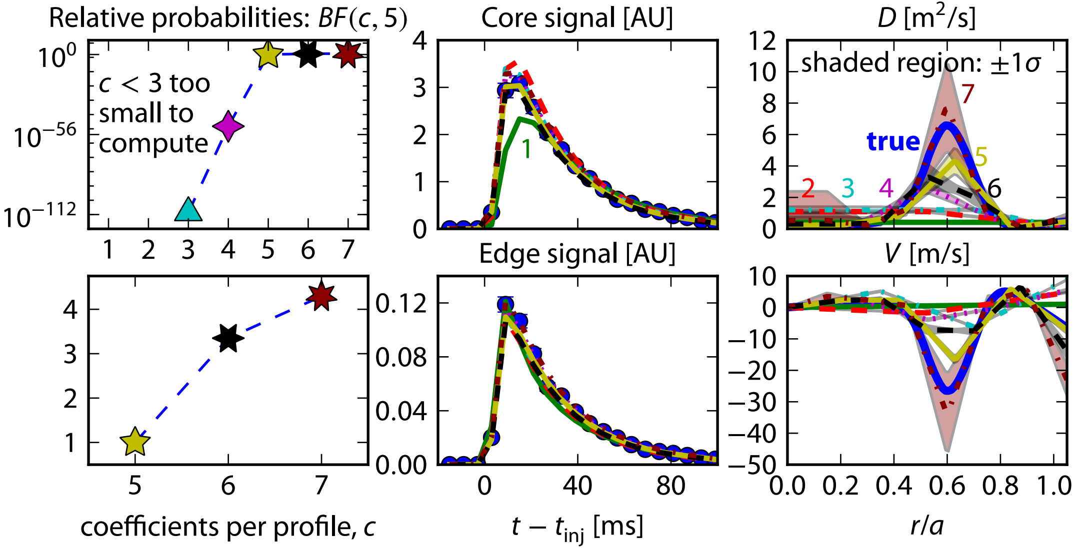
Results only resemble true profile when a minimum level of complexity is obtained…despite "good" match to data: "eyeballing" does not work!



Results only resemble true profile when a minimum level of complexity is obtained…despite “good” match to data: “eyeballing” does not work!

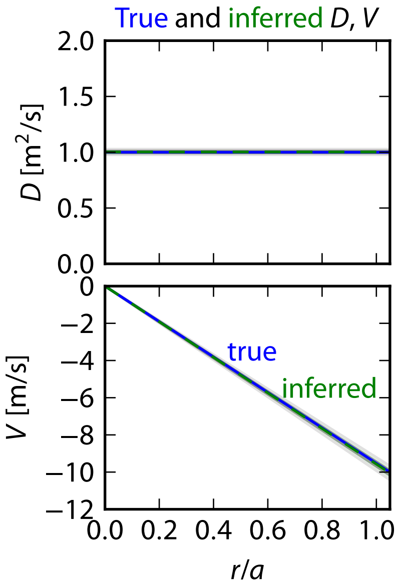


Results only resemble true profile when a minimum level of complexity is obtained…despite "good" match to data: "eyeballing" does not work!

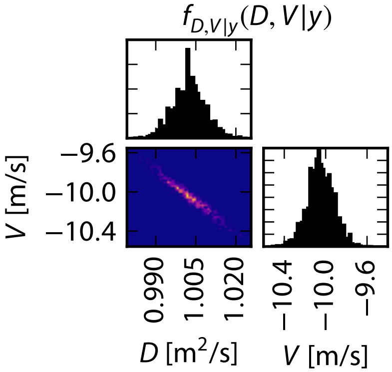


MultiNest successfully reconstructs simple *D*, *V* profiles

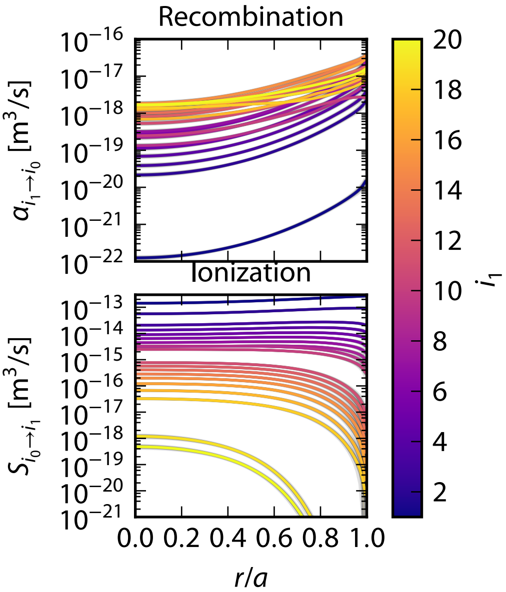
(synthetic data)



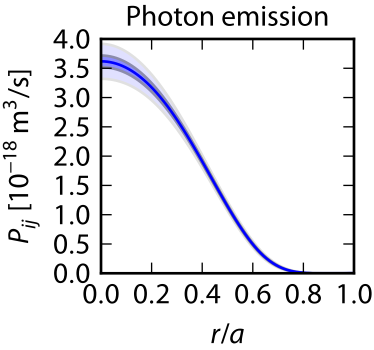
- Five local measurements
- $\Delta t = 6$ ms, 5% noise
- Have also tested with 32 XICS chords



Rate coefficients have low sensitivity to n_e, T_e over their uncertainties



- 1000 samples from n_e, T_e
- $\pm 3\sigma$ band on α, S not even visible!
- These are the only ways that n_e, T_e enter the calculation.

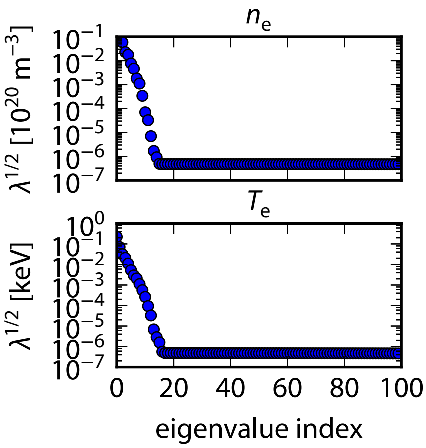


GPR permits efficient propagation of uncertainty [Chilenski NF 2015]

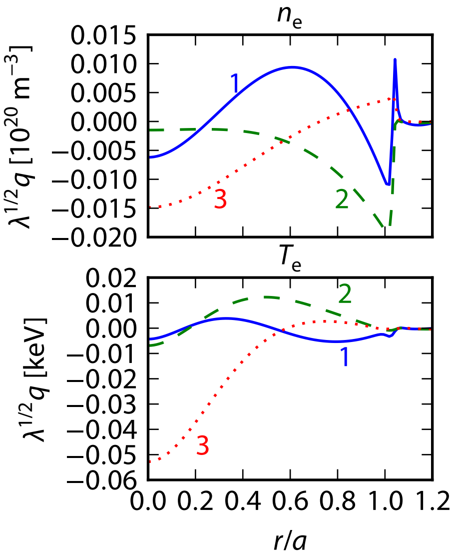
Drawing samples of the profile $\boldsymbol{y}$:

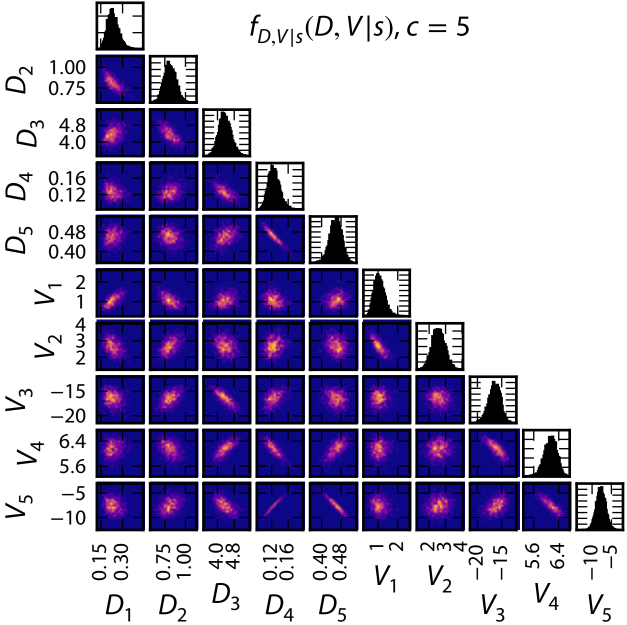
$$\boldsymbol{y} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}), \quad \boldsymbol{\Sigma} = \mathbf{Q}\boldsymbol{\Lambda}\mathbf{Q}^{-1}$$

$$\tilde{\boldsymbol{y}} = \mathbf{Q}\boldsymbol{\Lambda}^{1/2}\boldsymbol{u} + \boldsymbol{\mu}, \quad \boldsymbol{u} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$$

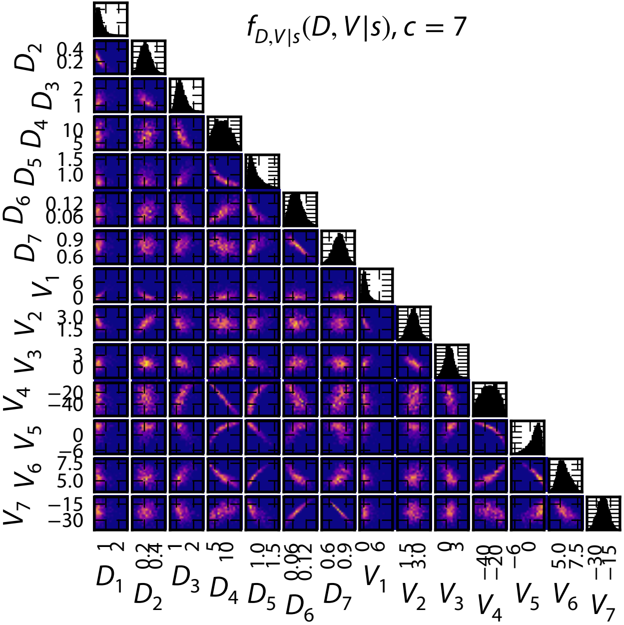


n_e, T_e eigenvectors





$f_{D,V|s}(D, V|s), c = 5$
D_1 D_2 D_3 D_4 D_5 V_1 V_2 V_3 V_4 V_5


$f_{D,V|s}(D,V|s), c = 7$
D_1
D_2
D_3
D_4
D_5
D_6
D_7
V_1
V_2
V_3
V_4
V_5
V_6
V_7