AtomDB Workshop 2018 @ CfA

Atomic Uncertainties into Data Analysis

Vinay Kashyap
CHASC/CfA
& the ISSIAHS Collaboration

ISSIAHS Collaboration

International Space Science Institute – Atomic, Heliophysics, and Statistics

To incorporate atomic data uncertainties and statistical uncertainties in estimates of parameters that define coronal structure.

Harry Warren (NRL; PI), *Adam Foster* (CfA), Chloe Guennou (Columbia), *Connor Ballance* (Queen's Univ), *David Stenning* (Imperial), *David van Dyk* (Imperial), Fabio Reale (OAPA, Palermo), Frederic Auchere (Inst. Astr. Spatiale), *Giulio Del Zanna* (Cambridge), Inigo Arregui (Inst. Astr. de Canarias), *Jessi Cisweski* (Yale), Mark Weber (CfA; co-PI), *Nathan Stein* (Spotify), *Randall Smith* (CfA), Veronique Delouille (Royal Obs. Belgium), Vinay Kashyap (CfA), and *Xixi Yu* (Imperial)

ISSIAHS Collaboration

International Space Science Institute – Atomic, Heliophysics, and Statistics

THE ASTROPHYSICAL JOURNAL, 866:146 (20pp), 2018 October 20 https://doi.org/10.3847/1538-4357/aadfdd

CrossMark

Incorporating Uncertainties in Atomic Data into the Analysis of Solar and Stellar Observations: A Case Study in Fe XIII

Xixi Yu[1], Giulio Del Zanna[2], David C. Stenning[1], Jessi Cisewski-Kehe[3], Vinay L. Kashyap[4], Nathan Stein[5], David A. van Dyk[1], Harry P. Warren[6], and Mark A. Weber[4]

[1] Statistics Section, Department of Mathematics, Imperial College London, London SW7 2AZ, UK
[2] DAMTP, Centre for Mathematical Sciences, University of Cambridge, Wilberforce Road, Cambridge CB3 0WA, UK
[3] Department of Statistics and Data Science, Yale University, New Haven, CT 06511, USA
[4] Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138, USA
[5] Spotify, 45 W 18th St, New York, NY 10011, USA
[6] Space Science Division, Naval Research Laboratory, Washington, DC 20375, USA
Received 2018 March 19; revised 2018 August 10; accepted 2018 September 2; published 2018 October 23

Intentions

- Ultimate goal

  $f_\lambda = \int dT\ \varepsilon_\lambda(n_e,T)\ DEM(n_e,T)$

  fit a model of emission measure distribution $DEM(n_e,T)=n_e^2\ dV/dT$ to high-resolution spectra including uncertainties in emissivity ε

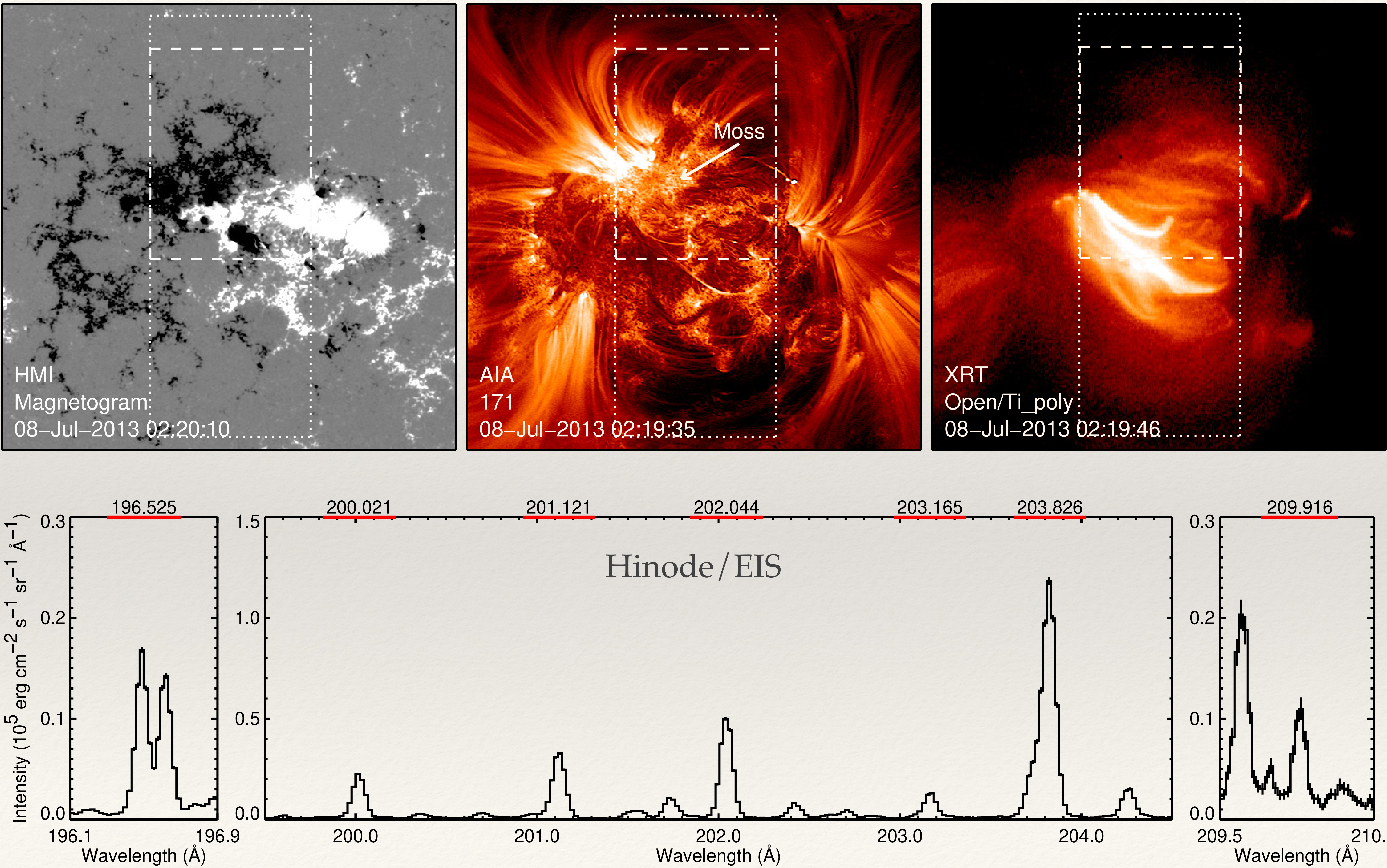
- Baby step

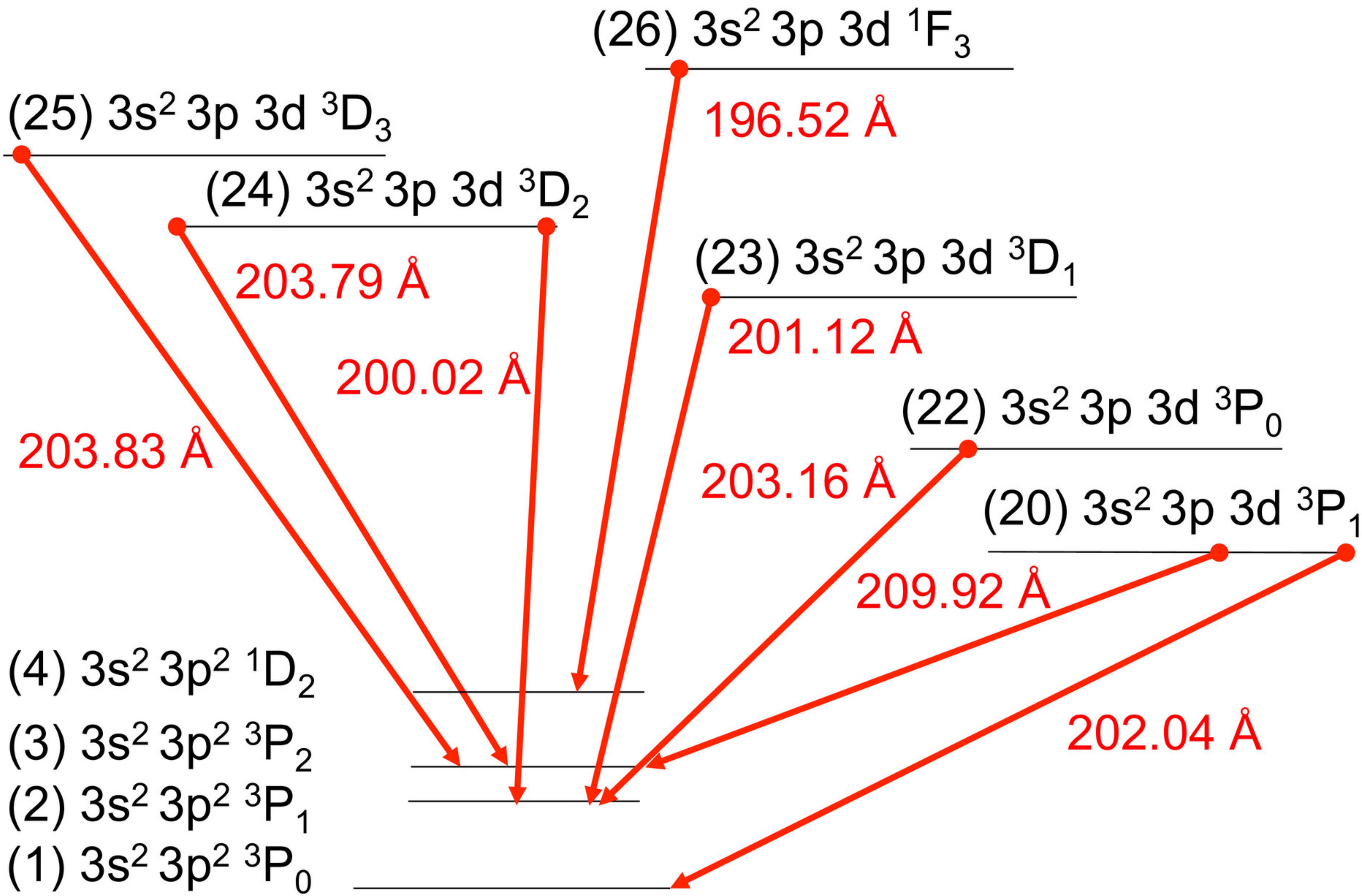
  estimate density n_e and path length δs from Fe XIII lines from the solar corona, which are to a very good approximation not T sensitive

  $f_\lambda = \varepsilon_\lambda(n_e)\ n_e^2\ \delta s$,

  $\{\lambda$=196.5, 200, 201.1, 202, 203.2, 203.8, 209.9 Å$\}$

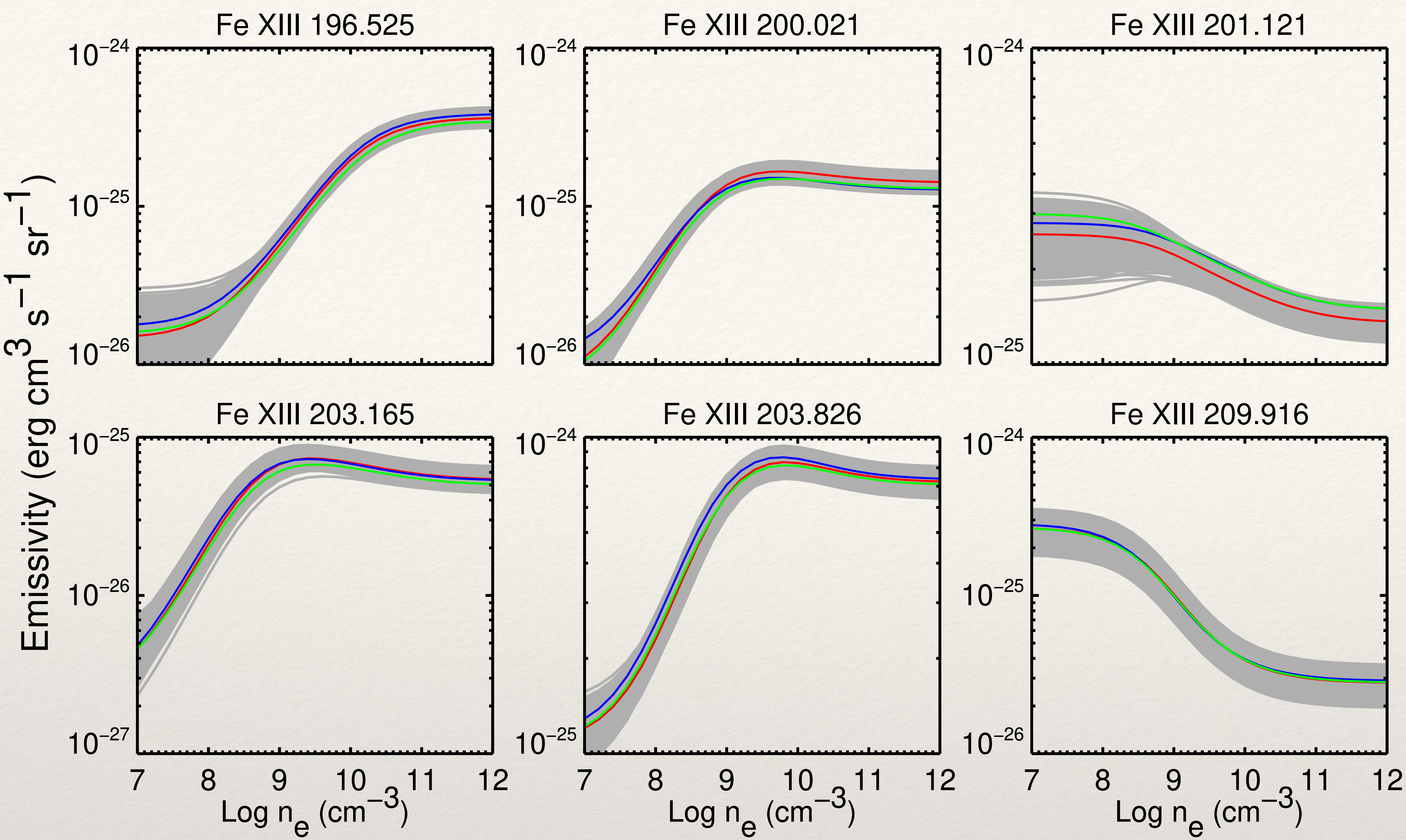
AR 11785 from 8 Jul 2013
from which 1000 pixels were chosen randomly for analysis




(26) 3s2 3p 3d 1F3
196.52 Å
(25) 3s2 3p 3d 3D3
(24) 3s2 3p 3d 3D2
203.79 Å
(23) 3s2 3p 3d 3D1
201.12 Å
200.02 Å
(22) 3s2 3p 3d 3P0
203.83 Å
203.16 Å
(20) 3s2 3p 3d 3P1
209.92 Å
(4) 3s2 3p2 1D2
(3) 3s2 3p2 3P2
202.04 Å
(2) 3s2 3p2 3P1
(1) 3s2 3p2 3P0

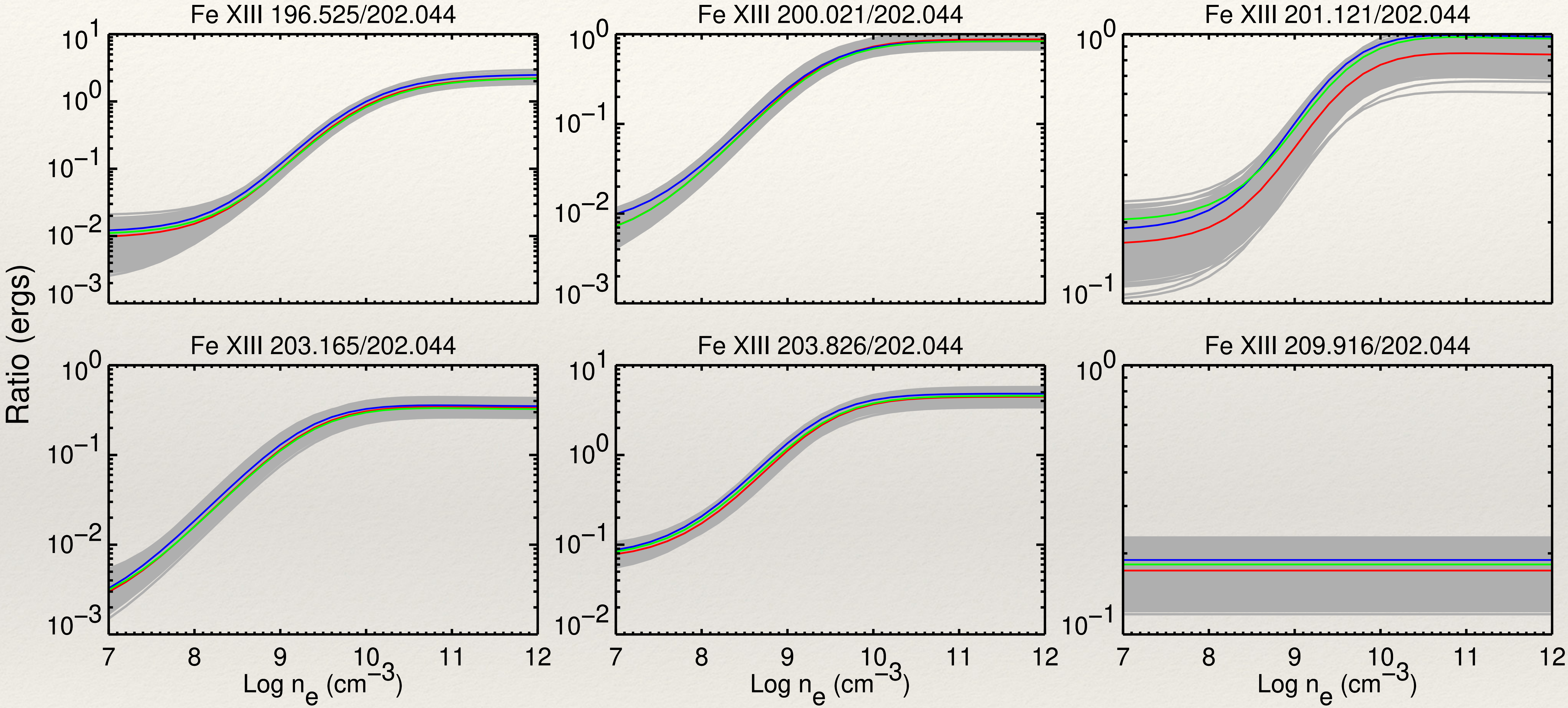
Uncertainties in Atomic Data

- collision strengths (comparisons Storey & Zeippen 2010 vs Del Zanna & Storey 2012, Distorted Wave vs R-matrix)
  - $3s^2\ 3p^2 \rightarrow 3s^2\ 3p\ 3d$: 5% for >1, linear to 0.01, max of 50%
  - other n=3 levels in SZ2010: 10% for >0.1, linear to 0.001, max of 50%
  - to n=4 and other n=3 levels: 20% for >0.01, 50% for <0.01
- A-values
  - 5% for $A>10^{10}$, 10% for $10^8<A<10^{10}$, 30% for $A<10^8$
- generate random deviates to collision strengths and A-values and run through CHIANTI to generate new emissivity curves for each transition



1000 emissivities generated by imputing uncertainties on collision strengths and transition probabilities of Fe XIII levels in CHIANTI, and generating new emissivities by propagating these uncertainties through level population estimates. Red curve is default CHIANTI. Blue curve is #471 (foreshadowing!)

Ratios of sampled emissivities relative to 202.044 Å line



Statistical Strategy

- Stage 1: the usual suspect (χ^2 minimization)
- Stage 2: brute force error estimation
- Stage 3: Bayesian analysis
- Stage 4: Fully Bayesian analysis of a thousand (randomly selected) datasets

Statistical Analysis

- Bayesian analysis, following the same track as Lee et al. 2011 (ApJ 731, 126) and Xu et al. 2014 (ApJ 794, 97)
- *Pragmatic Bayes*, which takes the sample of emissivities as given, and sees what effect it has on the parameter estimates and uncertainties

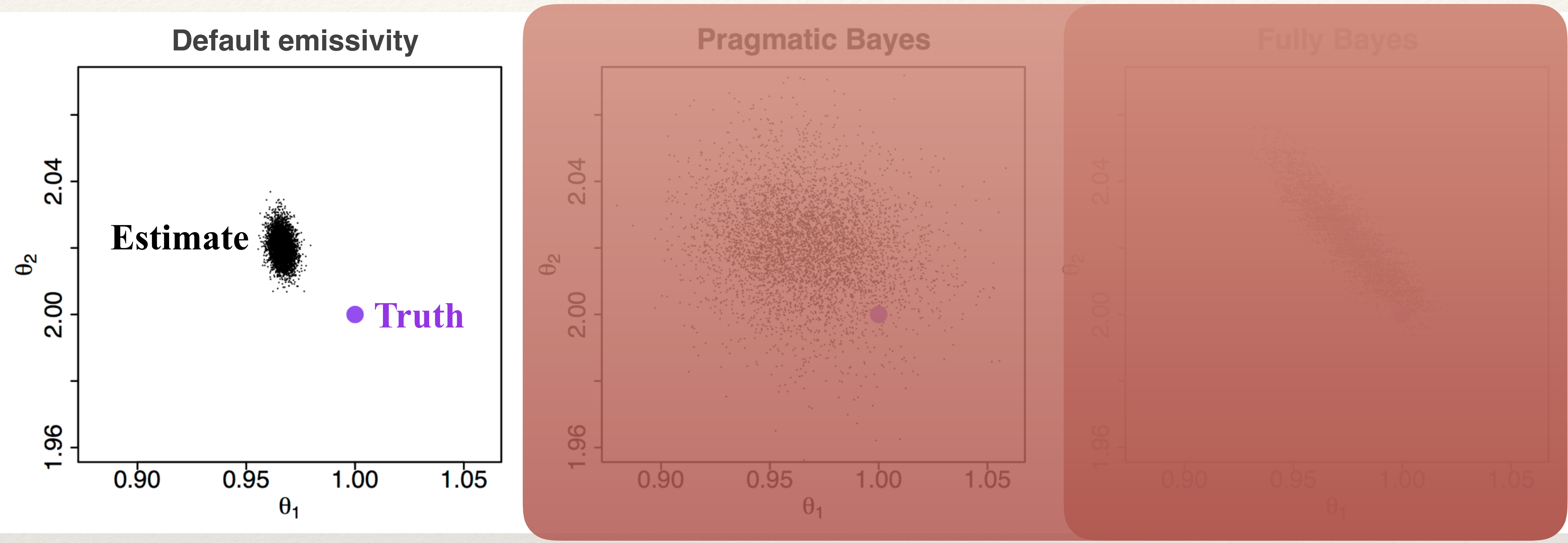
$$\mathbf{p(m,\theta|D) = p(\theta|D,m)\ p(m)}$$

- *Fully Bayes*, which "weights down" instances of emissivity samples that produce bad likelihoods hence additionally selects preferred emissivities

$$\mathbf{p(m,\theta|D) = p(\theta|D,m)\ p(m|D)}$$

Lee et al. 2011, ApJ 731, 126; Xu et al. 2014, ApJ 794 97; Yu et al. 2018, ApJ 866 146

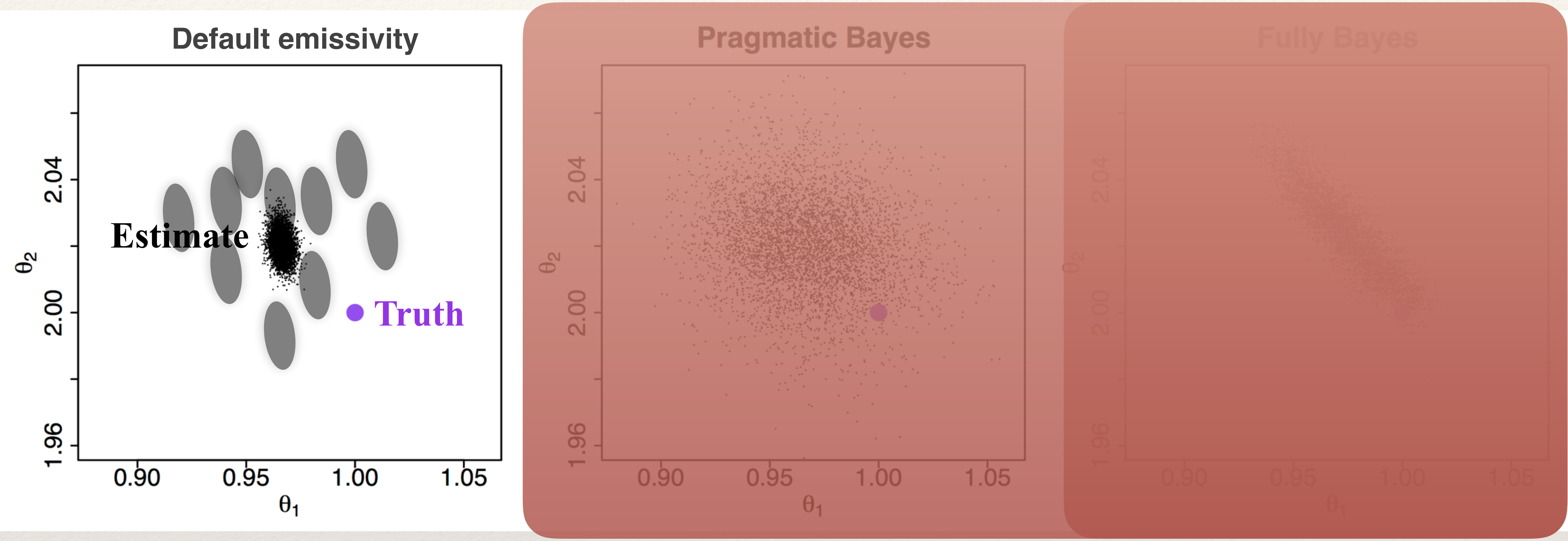
Standard → Prag Bayes → Full Bayes



$p(\theta|D,\varepsilon^{(def)})$

Lee et al. 2011, ApJ 731, 126; Xu et al. 2014, ApJ 794 97; Yu et al. 2018, ApJ 866 146

Standard → Prag Bayes → Full Bayes

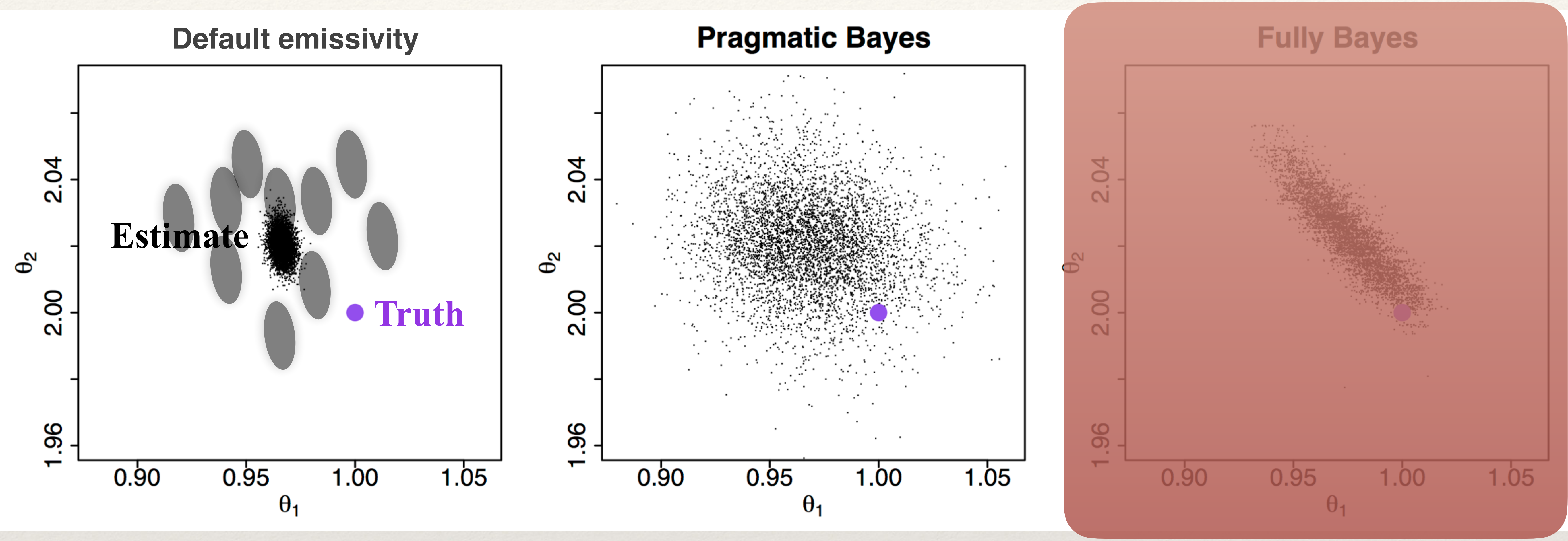


$p(\theta|D,\varepsilon^{(m)})$

$p(\theta|D,\varepsilon^{(def)})$

Lee et al. 2011, ApJ 731, 126; Xu et al. 2014, ApJ 794 97; Yu et al. 2018, ApJ 866 146

Standard → Prag Bayes → Full Bayes



$p(\theta|D,\varepsilon^{(m)})$

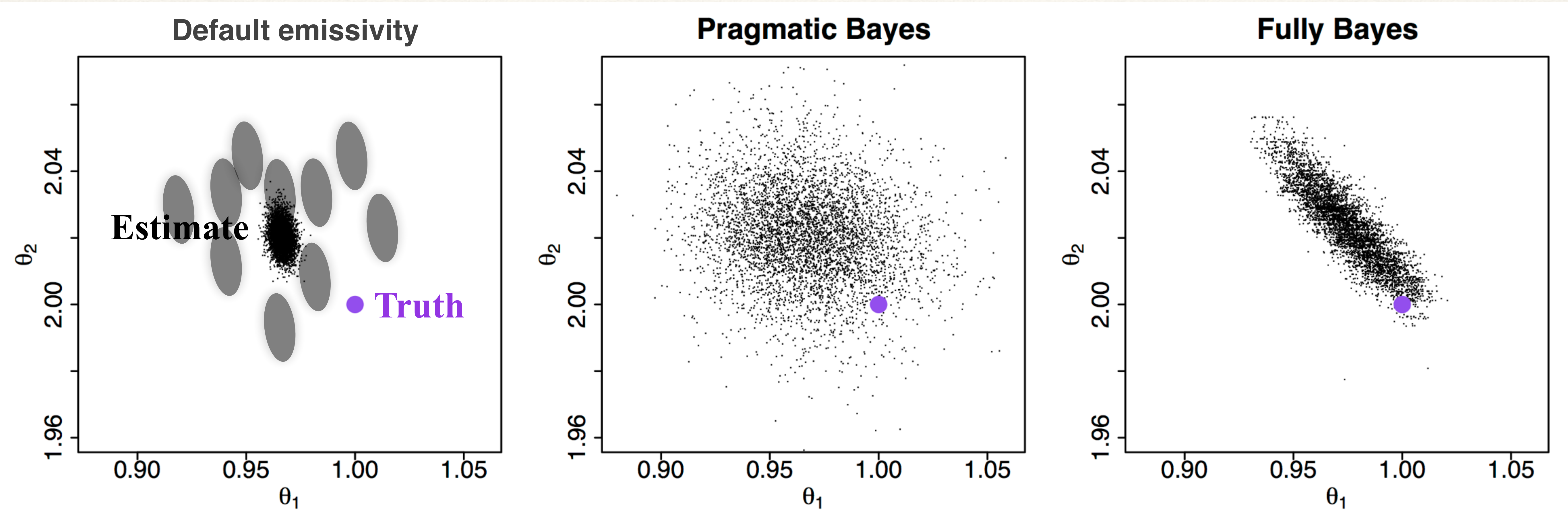
$p(\theta|D,\varepsilon^{(def)})$

$p(\varepsilon,\theta|D)$

$\rightarrow p(\theta|D,\varepsilon) \cdot p(\varepsilon)$

Lee et al. 2011, ApJ 731, 126; Xu et al. 2014, ApJ 794 97; Yu et al. 2018, ApJ 866 146

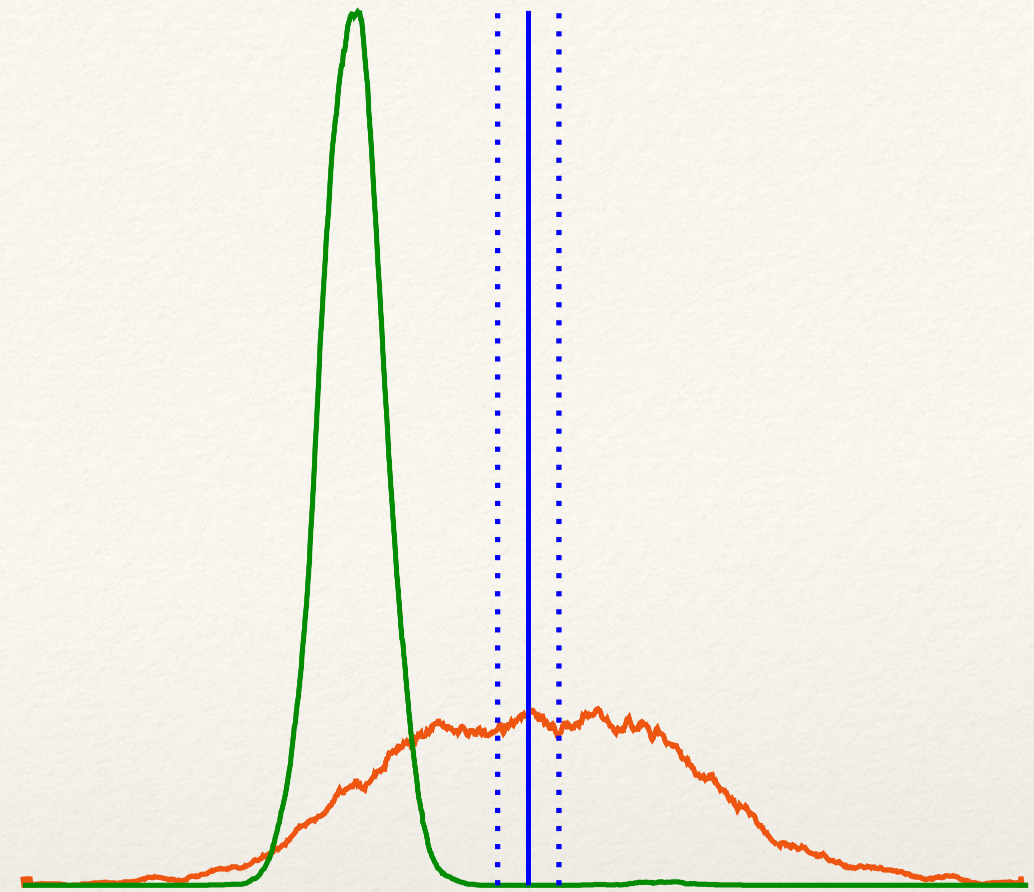
Standard → Prag Bayes → Full Bayes

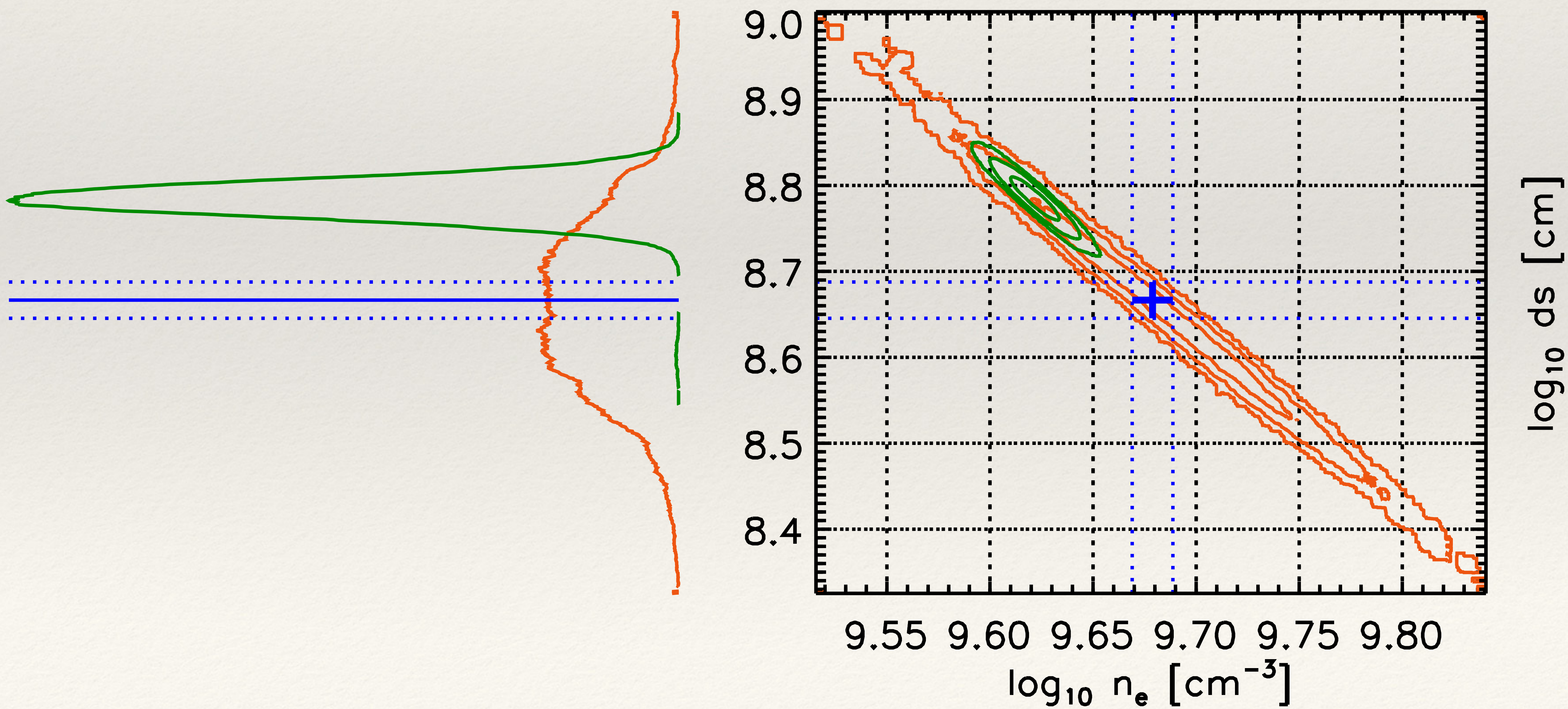


$p(\theta \vert D,\varepsilon^{(m)})$	$p(\varepsilon,\theta \vert D)$	$p(\varepsilon,\theta \vert D)$
$p(\theta \vert D,\varepsilon^{(def)})$	$\to p(\theta \vert D,\varepsilon) \cdot p(\varepsilon)$	$\to p(\theta \vert D,\varepsilon) \cdot p(\varepsilon \vert D)$

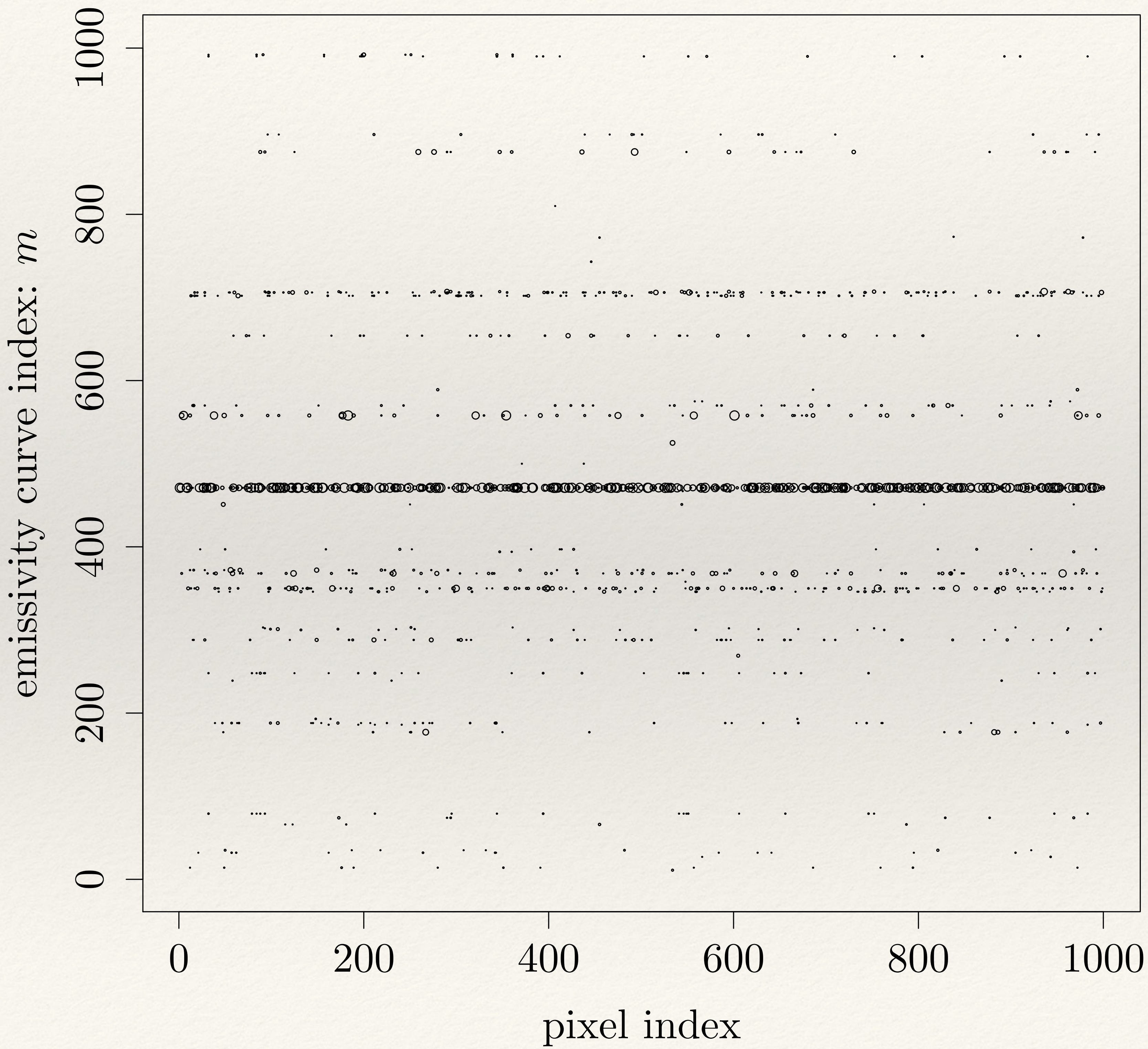
pix 217

$\log_{10} n_e$ [cm^{-3}]		$\log_{10}$ ds [cm]
9.68+−0.010	[std]	8.67+−0.021
9.68+−0.049 EQT_{90}: 9.61 − 9.76	[pragB]	8.66+−0.098 EQT_{90}: 8.50 − 8.81
9.62+−0.011 EQT_{90}: 9.61 − 9.64	[fullB]	8.78+−0.023 EQT_{90}: 8.75 − 8.82

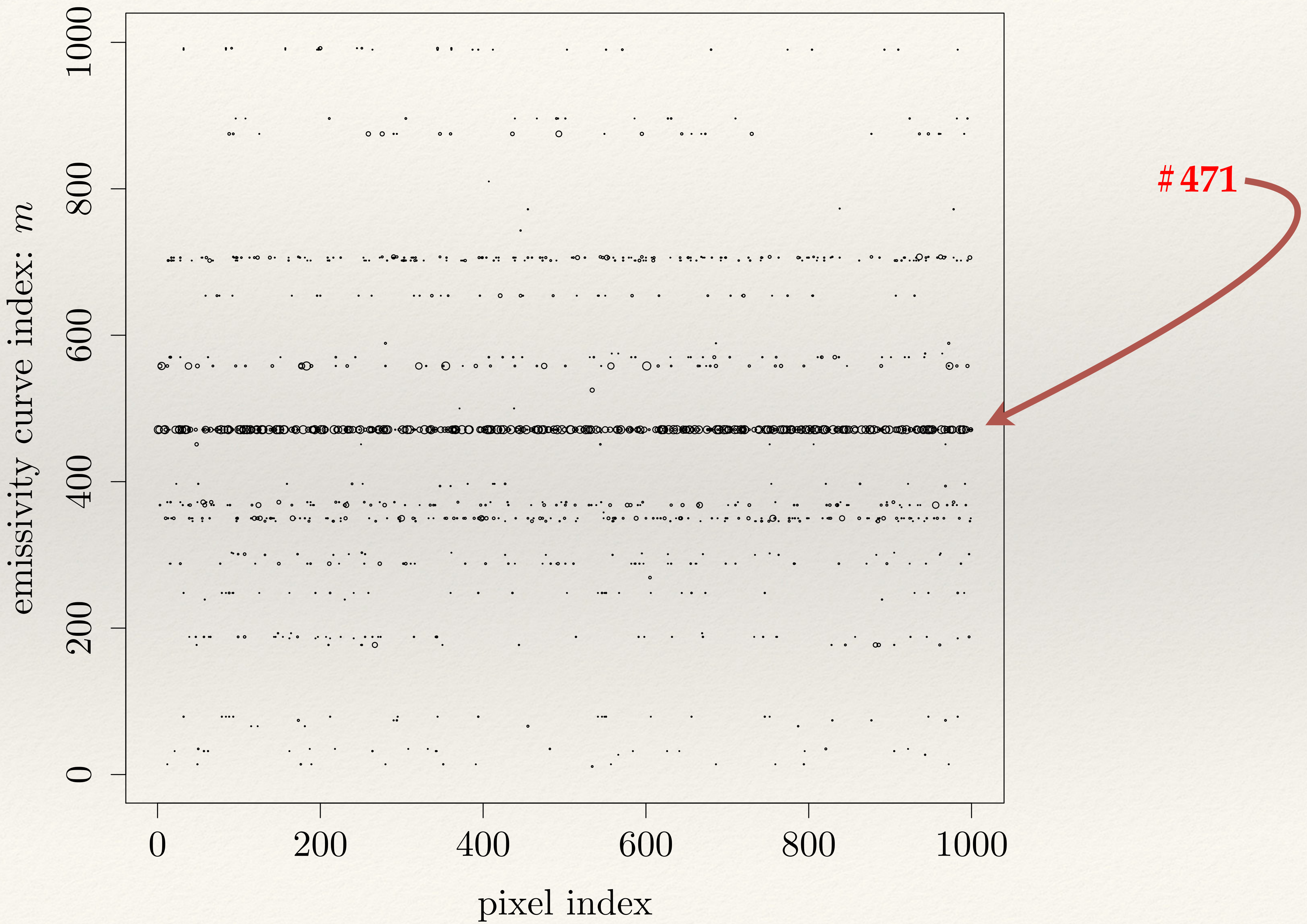




A surprising result



A surprising result



Compare selected emissivities with default

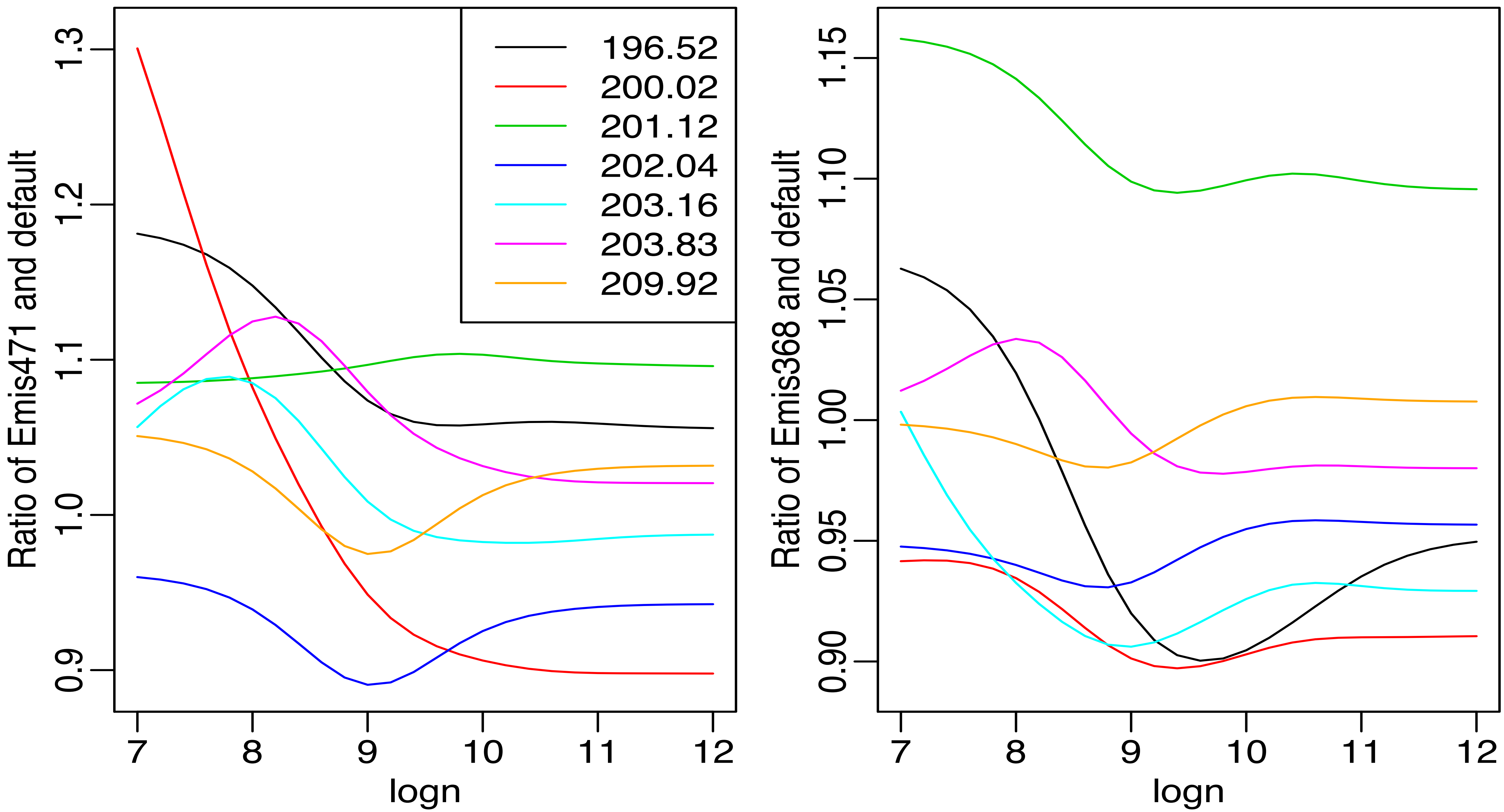


Figure: Plot of ratio of selected emissivities and default CHIANTI over 7 lines.

Compare selected emissivities with default

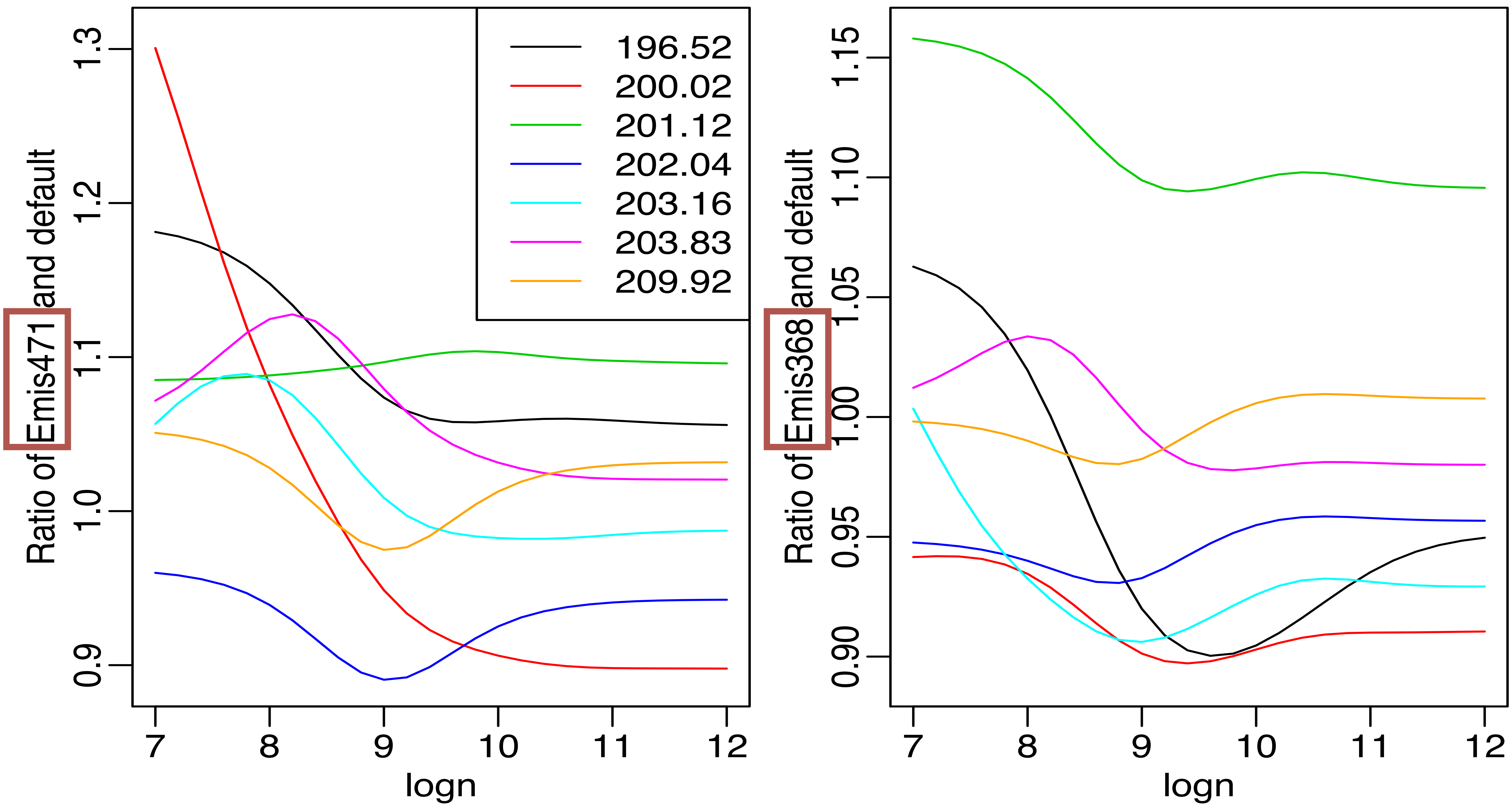
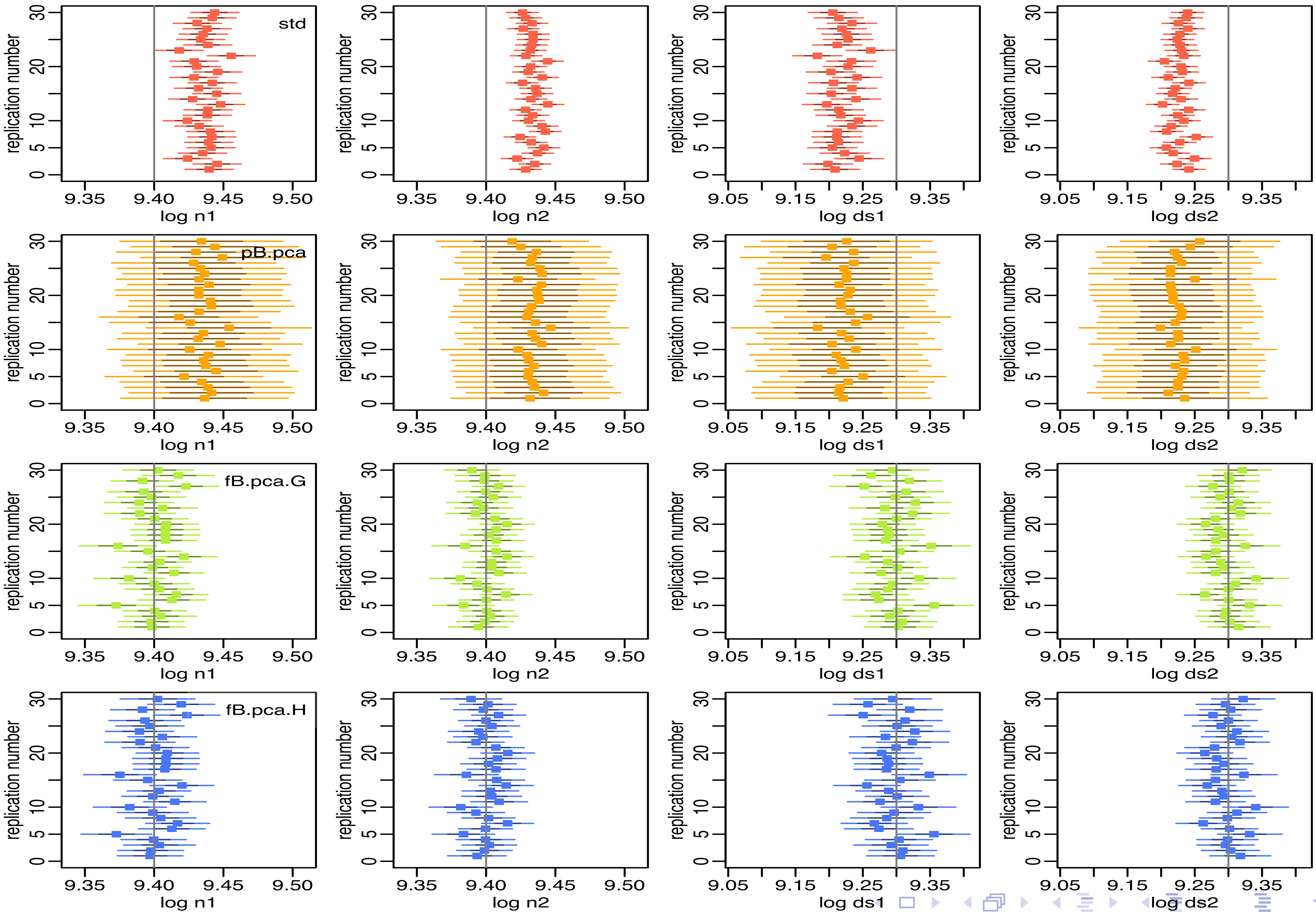


Figure: Plot of ratio of selected emissivities and default CHIANTI over 7 lines.

What's next?

- Better emissivity sampling, esp. energy levels in He-like lines
- PCA to overcome sparsity
- *Chandra* Capella O VII+O VIII+Fe XVII to extend to ion balance uncertainties, Temperature
- Correlations..

Prior II: Output summary, $J = 3$ and $K = 2$



Summary

- Developed a proof-of-concept method to incorporate atomic data uncertainties in data analysis
- Standard analysis underestimates systematic error
- Including atomic uncertainties (pragmatic Bayes) inflates error bars; further allowing astro data to inform atomic data (fully Bayes) shifts the estimates
- Surprising consistency in "optimum" emissivity
  - an error in CHIANTI calculations, or unaccounted problems with data analysis, or arising from model misspecification
- Conceptually complex, computationally costly
  - Right now not many people can generate emissivity samples that reflect atomic data uncertainties
  - Sparse sampling of emissivities requires filling in using PCA, will also speed it up