Statistical aspects of X-ray spectral fitting

IACHEC, 13.05.2025

Johannes Buchner

with Peter Boorman, David Homan, and the BXA community

Johannes Buchner

https://astrost.at/istics/

Astronomy path

Chandra & XMM AGN surveys (PhD 2015)

First Compton-thick luminosity function

Interests in Postdoc: Black hole obscuration & demographics

eROSITA (2019-)

chair of the AGN working group

Data science path

Bayesian inference with Markov Chains

pymultinest

scaling to big data &

diverse applications

ultranest

review of nested sampling

review of X-ray spectral analysis

neutron stars

Astro-particles

...

exoplanets

Book chapter

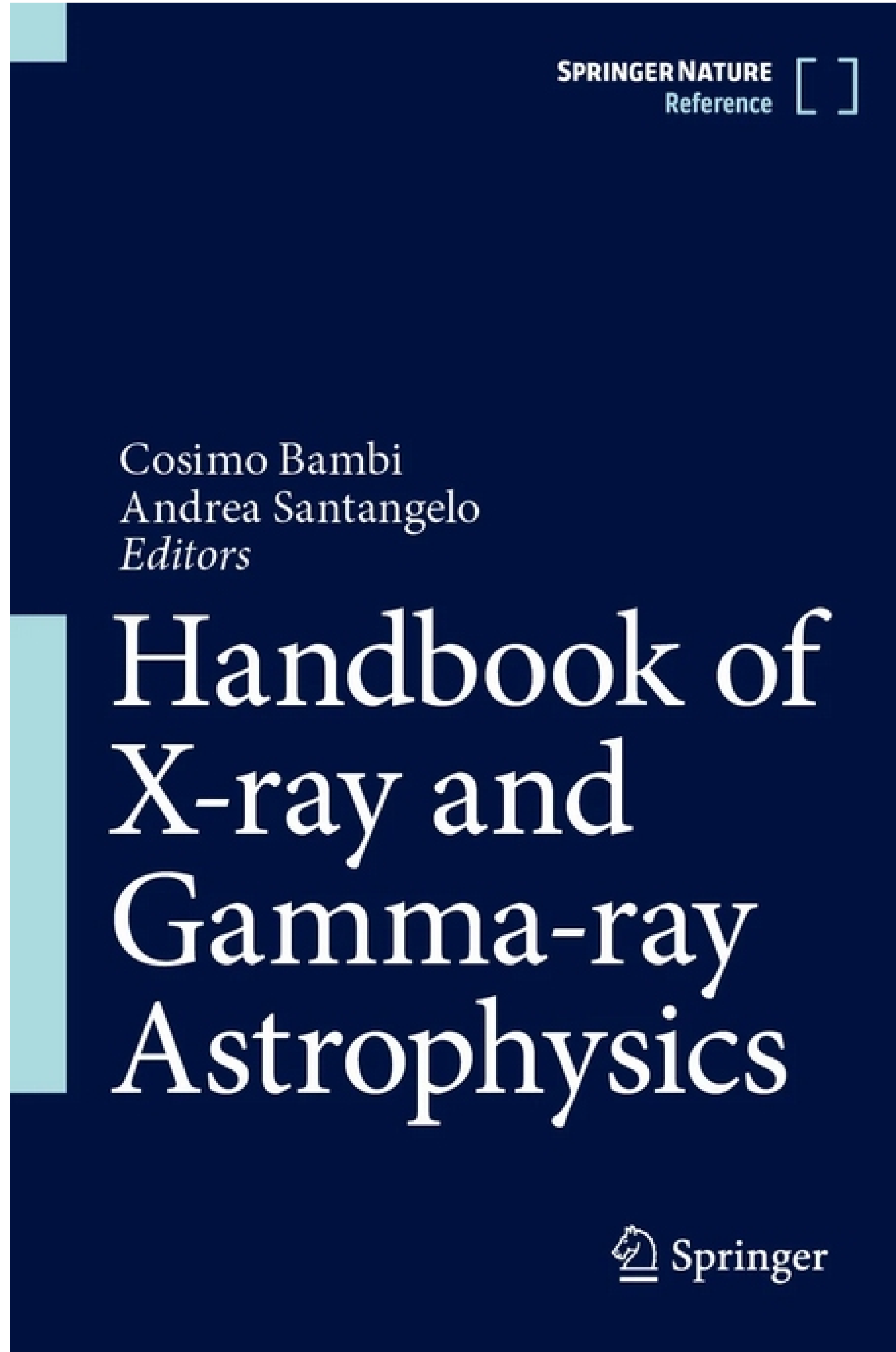
Statistical Aspects of X-ray Spectral Analysis

Johannes Buchner & Peter Boorman

freely available at:

https://arxiv.org/abs/2309.05705

includes hands-on exercises
for both sherpa & xspec



Workshops

X-ray spectral fitting methods workshop

- **When**: 24-25. September 2019
- **Where**: Max Planck Institute for extraterrestrial Physics, Garching, Germany
- **Who**: 44 participants. Speakers: Johannes Buchner, J Michael Burgess, Joern Wil

VIDEO RECORDINGS, SLIDES & LINKS

J Michael Burgess,
Joern Wilms,
JB

SPRINGER NATURE Reference

☆ Hol|
7 - 11 February 2022, Praha Czechia
Welcome to the X-ray Spectral Fitting (XSF) 2022 winter school page.

of

with Peter Boorman

S
X
Jo

freely a
https
incluc

Successful School of Astro-Statistics

26TH JANUARY 2021 • EVENTS, STATISTICS

BiD4BESt has just enjoyed its first online school. Across two weeks five academics shared their knowledge on Bayesian Inference, X-ray Spectral Analysis, Machine Learning and its Applications to Astronomy, and Fitting the Spectral Energy Distributions of galaxies and AGN.

Chandra Data Science

Chandra Data Science:
Novel Methods in Computing and Statistics for X-ray Astronomy

August 2021

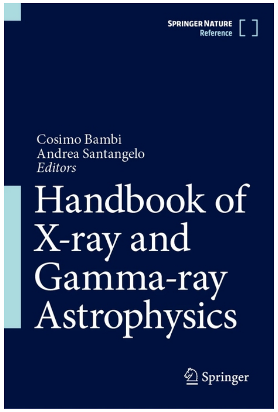
Virtual (online)
Hosted by the Chandra X-ray Center

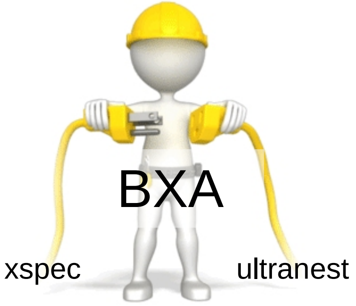
Springer

“jump right to the difficult stuff and make controversial statements” – Vinay

This talk

- book chapter contents & discussion
- Nested sampling & model comparison
- The evolving software landscape





Disclaimer #1:

members of this group have discussed statistical aspects for a long time

Book chapter transfers knowledge from modern statistical literature to X-ray astronomy

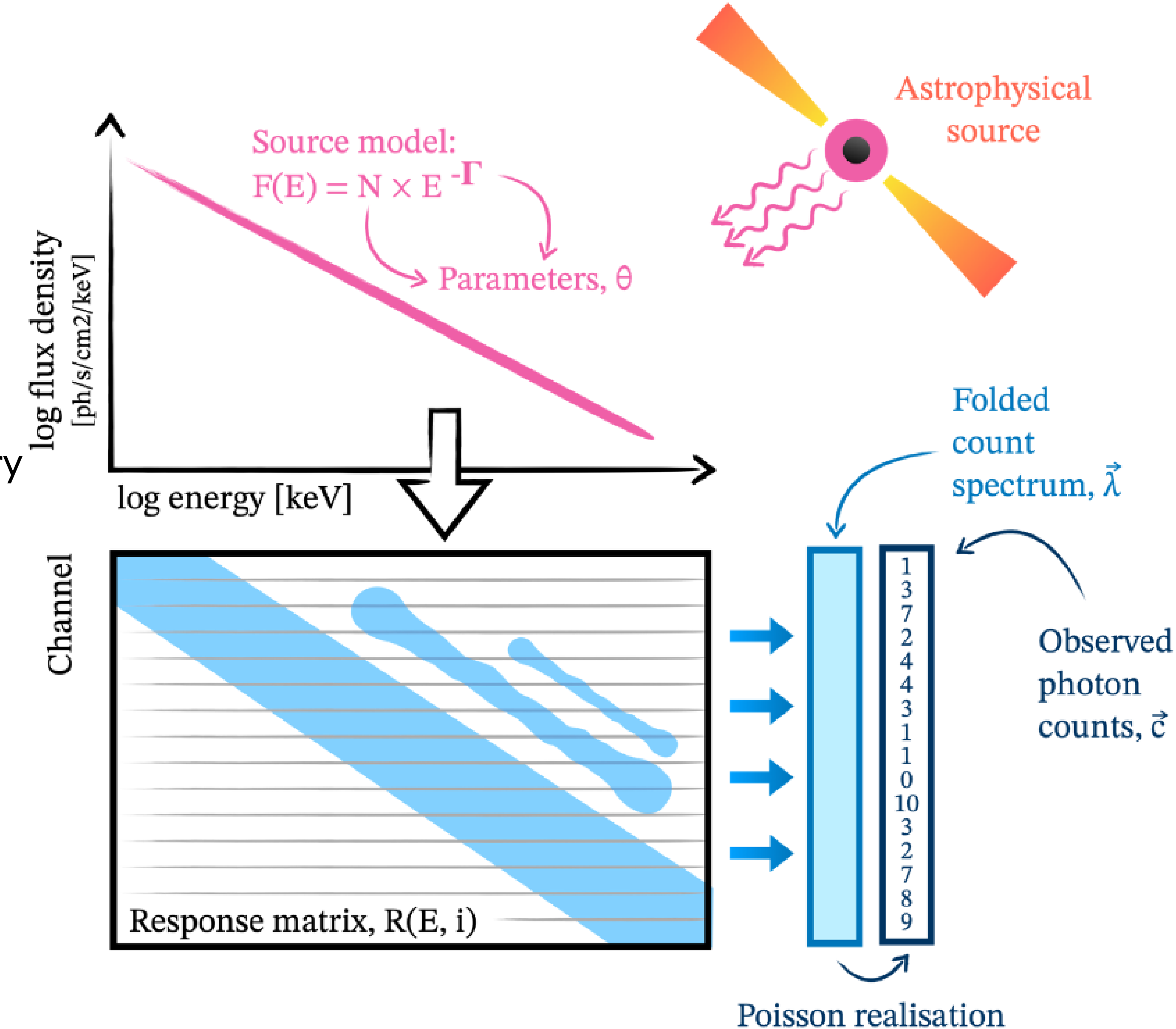
Disclaimer #2:

unknown systematics / calibration can limit the statistical modelling

importance of IACHEC & cross-mission calibration

Selected topics

- X-ray astronomy
  - the linear instrument model
  - Poisson statistics (cstat) & its asymmetry
  - Gaussian (chi²) statistics, persistant biases
  - Background models: physical, per-bin (wstat) and its bias, empirical (with ML)
- Frequentist statistics:
  - Confidence intervals: Calibration of confidence intervals
  - Model comparison: false positive rates, false negative rates
  - Any MCMC analysis assumes a prior

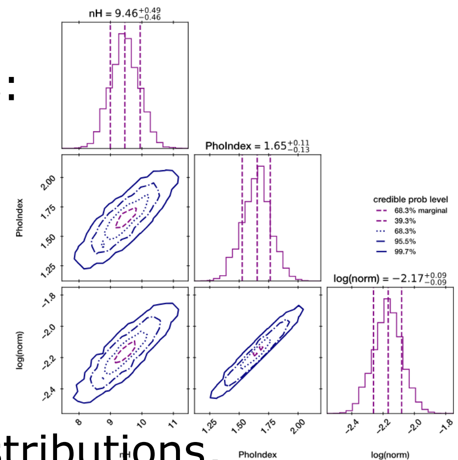


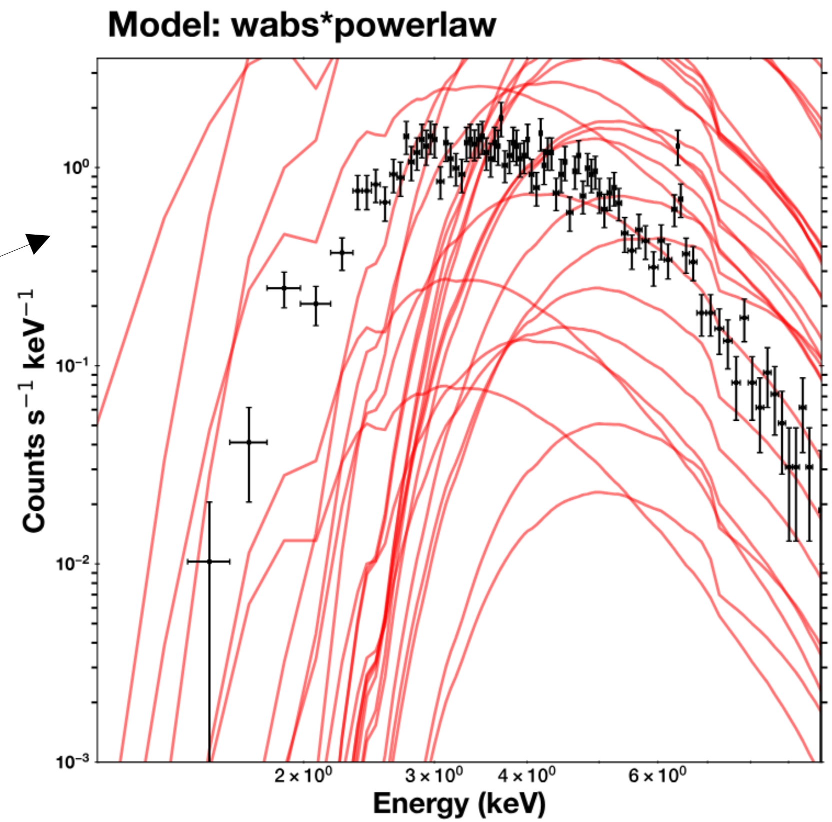

```
def cstat((N, Gamma)):
    F = N * E**Gamma
    lambda = (F * ARF) @ RMF * exposure
    poisson_lp = counts * log(lambda) + lambda
    return -poisson_lp

bestfit = scipy.optimize.minimize(cstat, (1, 1))
```

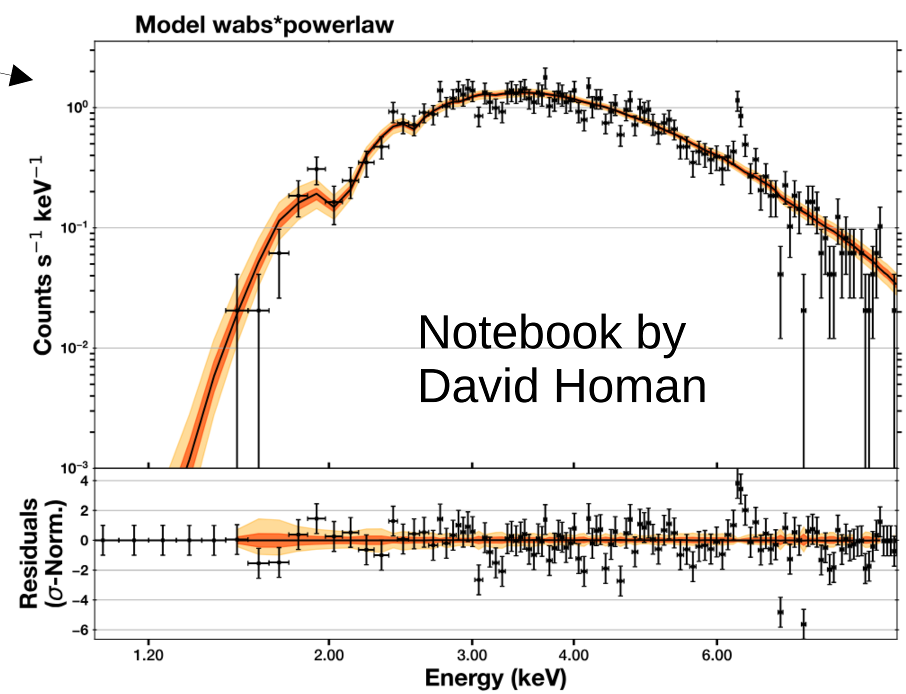
(cont) Bayesian workflow

- designing priors, prior predictive checks
- Monte Carlo Algorithms:
  - MCMC: emcee,
    Slice sampling, HMC
  - nested sampling:
    MultiNest, UltraNest

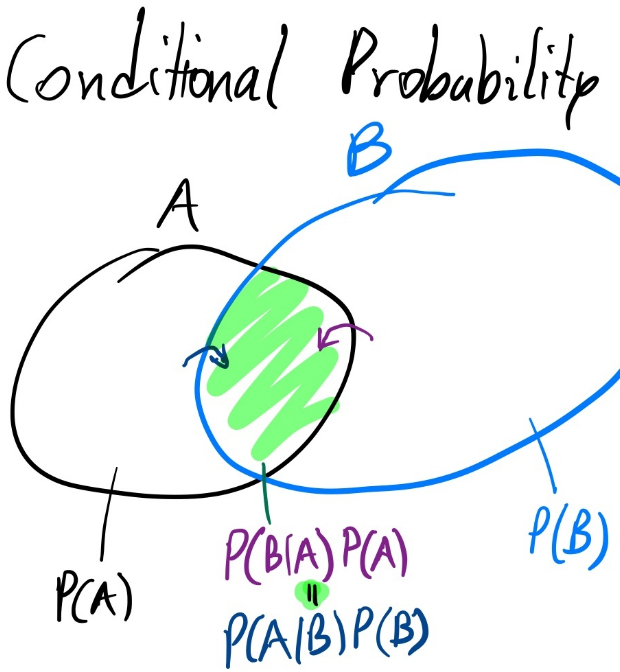




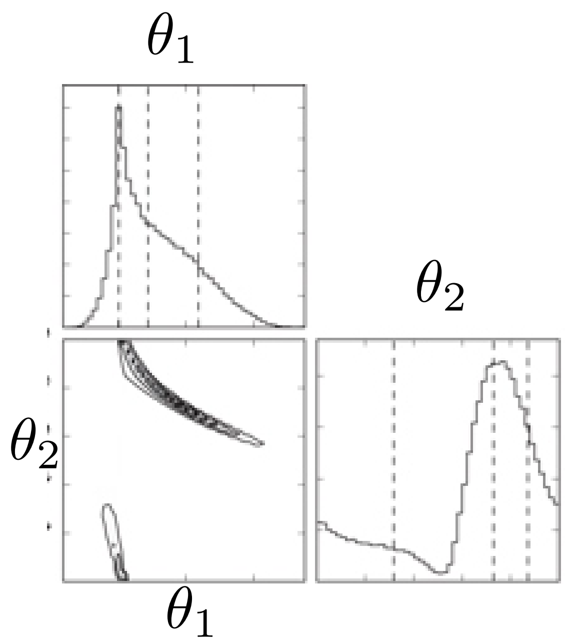
- parameter posterior distributions,
  posterior predictive checks,
  quantile-quantile plots
- Bayesian model comparison:
  built-in Occam's razor
- Frequentist properties of
  Bayesian analyses
- experiment design (for proposals)
- histograms of best-fit parameters,
  hierarchical Bayesian models
  (PosteriorStacker code)



https://johannesbuchner.github.io/BXA/tutorial_usage_plotbxa.html



θ=(L, T, …, physical parameters)



Bayes 101

Bayes theorem:

Posterior distribution

Prior distribution

Likelihood function

$$P(\theta|D) = \frac{\pi(\theta) \cdot P(D|\theta)}{P(D)}$$

Where is most of the probability ?

Evidence

normalizes the posterior

“average” likelihood

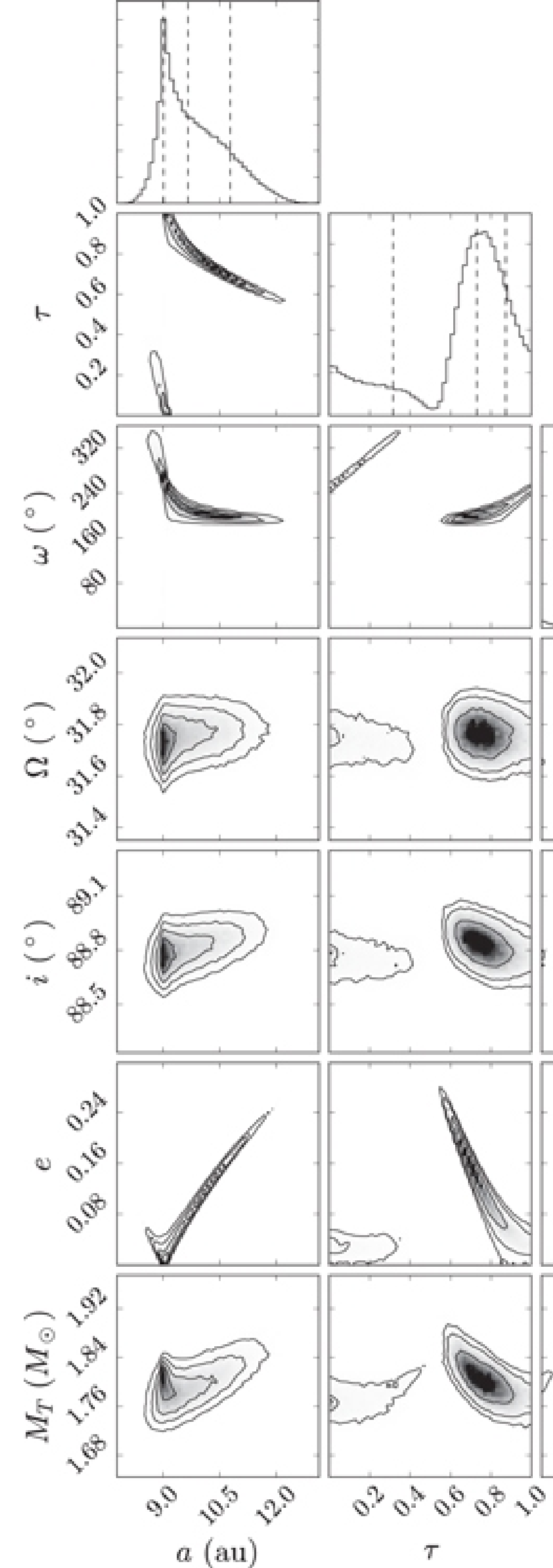
$$P(D) = \int d\theta \pi(\theta) P(D|\theta)$$

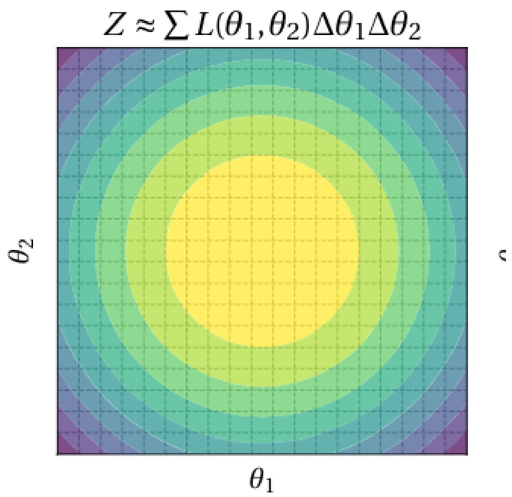
Parameter space exploration

- Local optimization
  - LM, simplex, … (many)
  - Monte carlo optimization
- Local sampling: MCMC
  - Tempering
  - Limitations
- Global optimization
  - Genetic algorithms (DE)
- Global sampling
  - Nested sampling

yields posterior samples

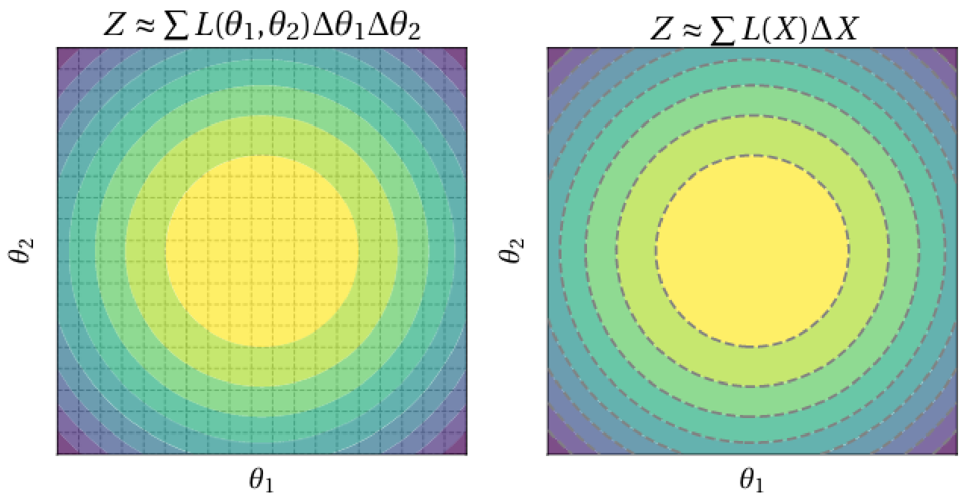
Yields evidence and posterior samples





Riemann integration

Figure 1 | **Illustrations of NS algorithm.**



Riemann integration   Lebesgue integration

Figure 1 | **Illustrations of NS algorithm.**

Ashton, Buchner, et al. 2022, Nature Reviews Methods Primers

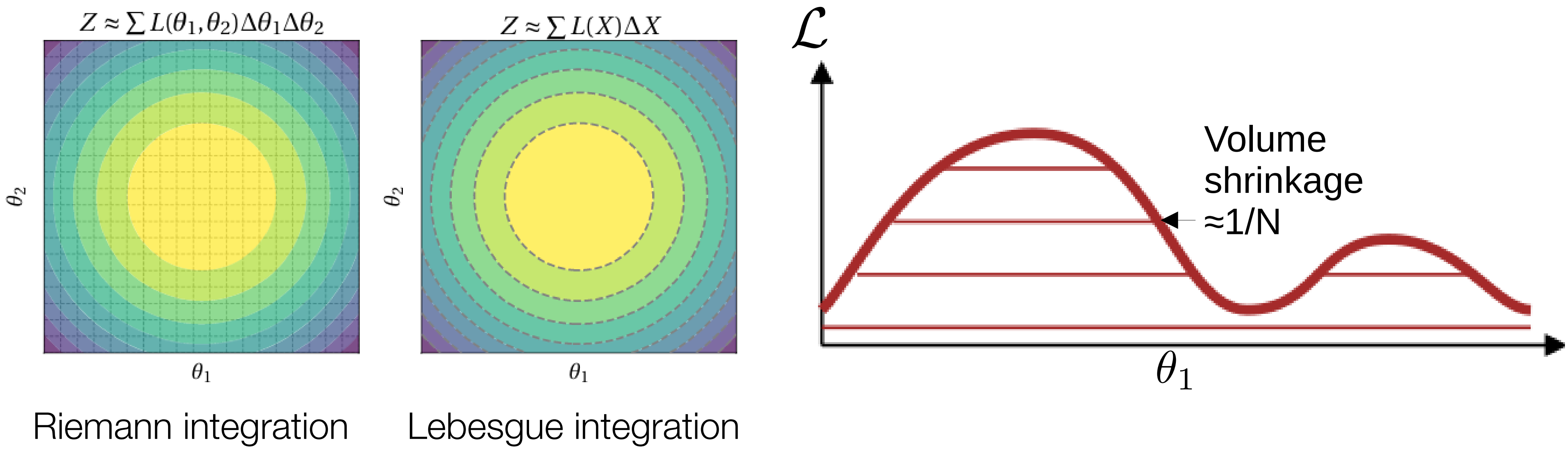
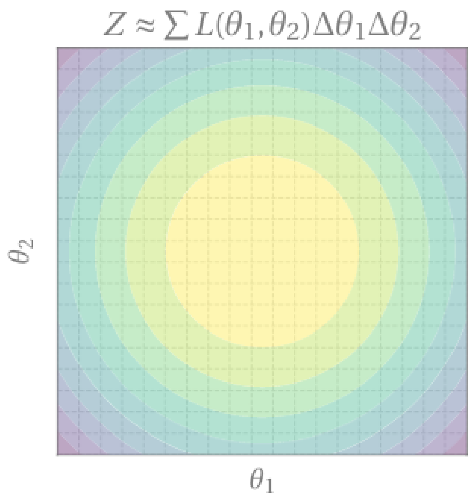
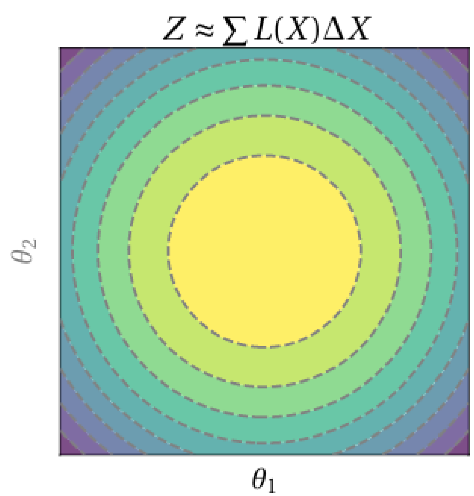


Figure 1 | **Illustrations of NS algorithm.**

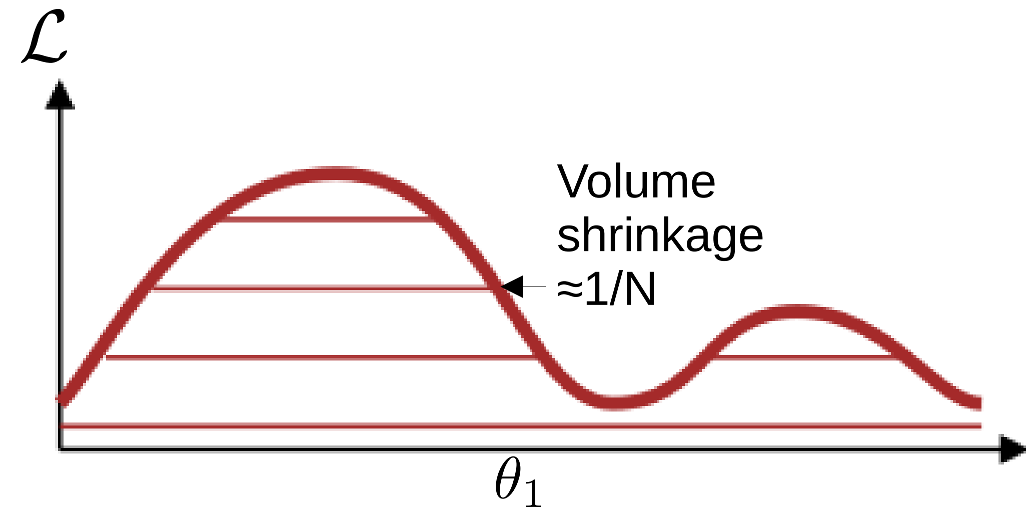
Ashton, Buchner, et al. 2022, Nature Reviews Methods Primers



Riemann integration



Lebesgue integration



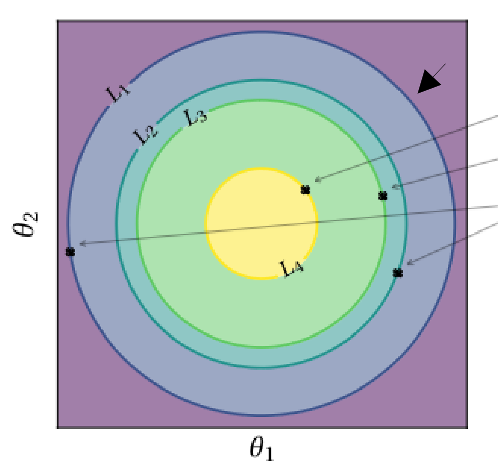


Figure 1 | **Illustrations of NS algorithm.**

Ashton, Buchner, et al. 2022, Nature Reviews Methods Primers

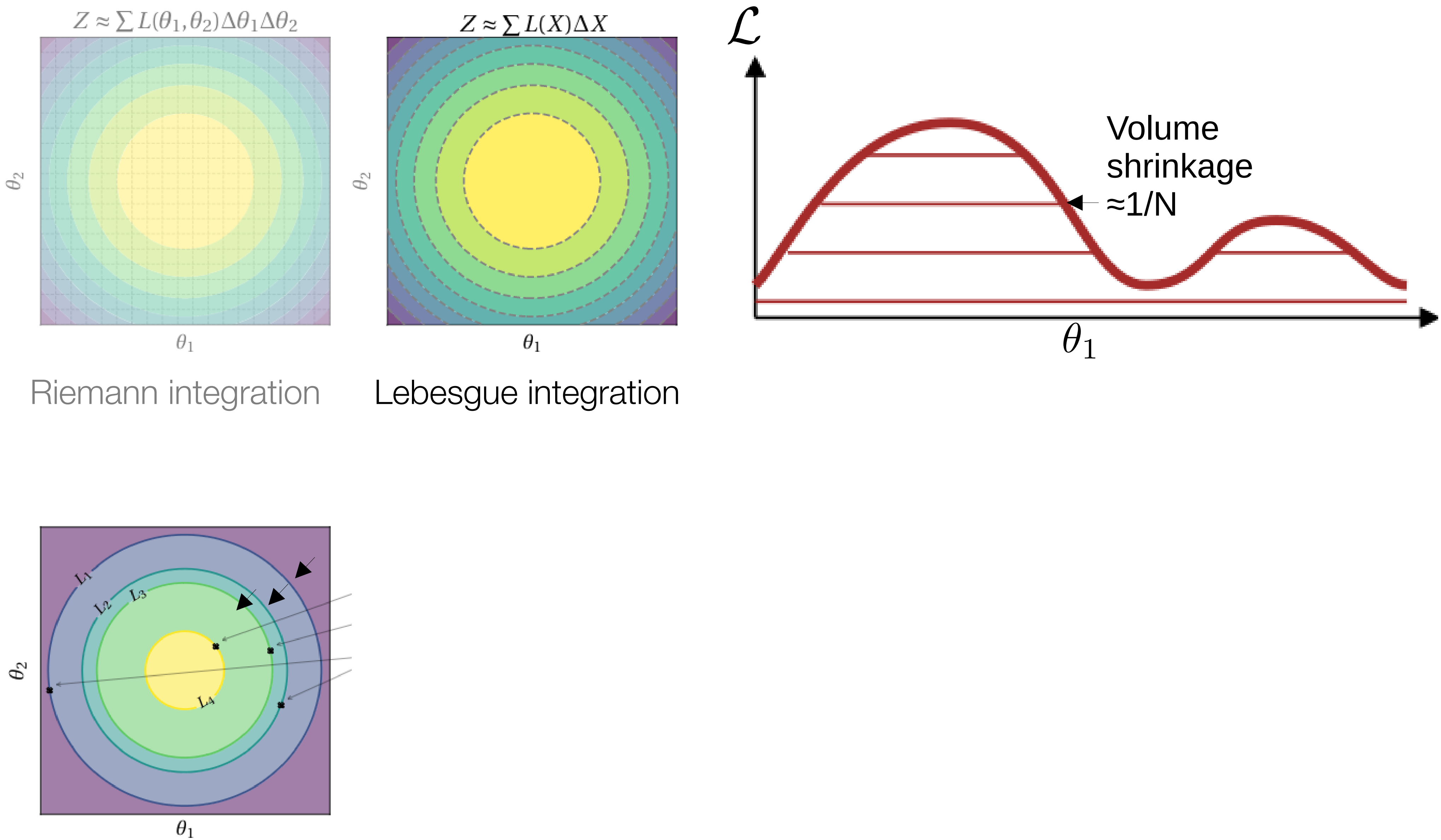
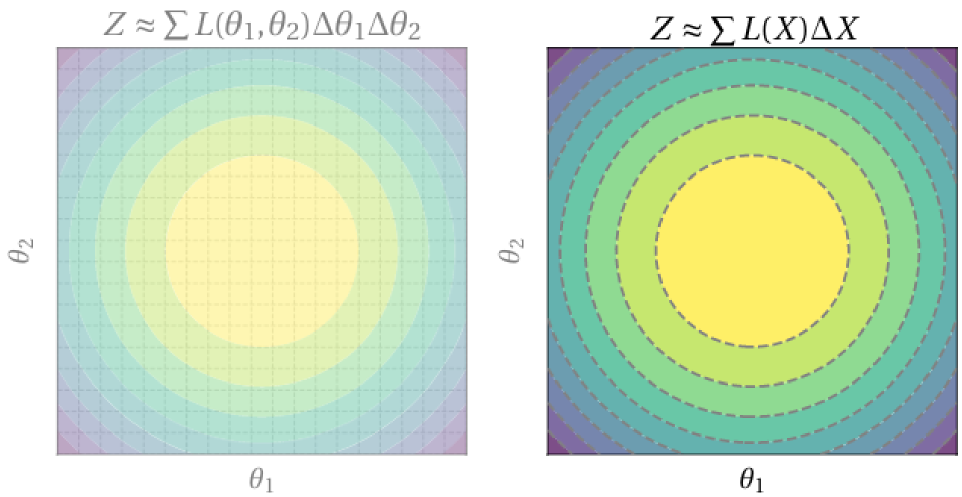
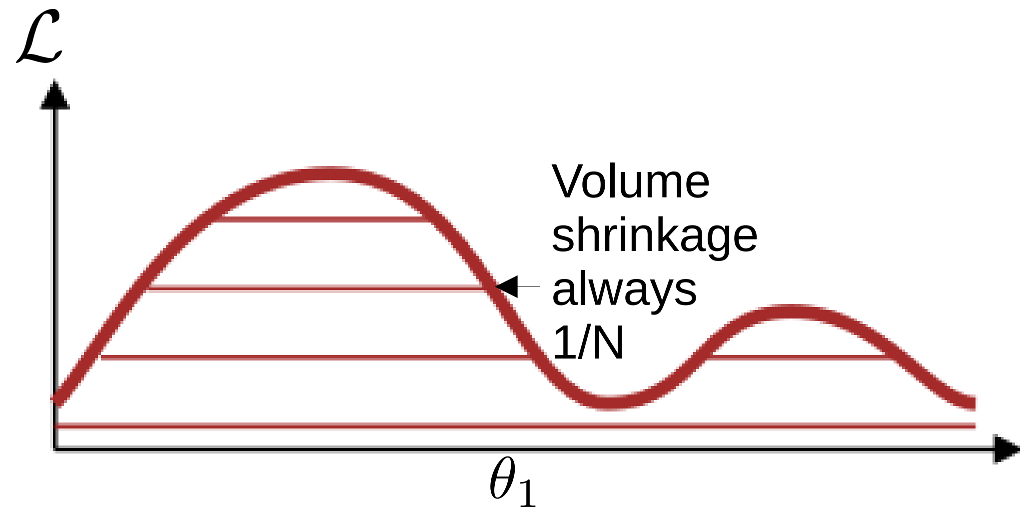


Figure 1 | **Illustrations of NS algorithm.**

Ashton, Buchner, et al. 2022, Nature Reviews Methods Primers



Riemann integration   Lebesgue integration



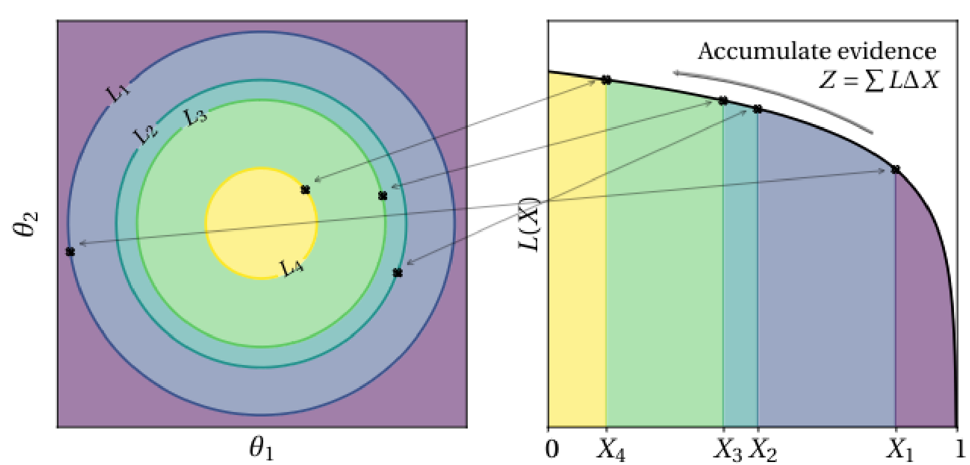
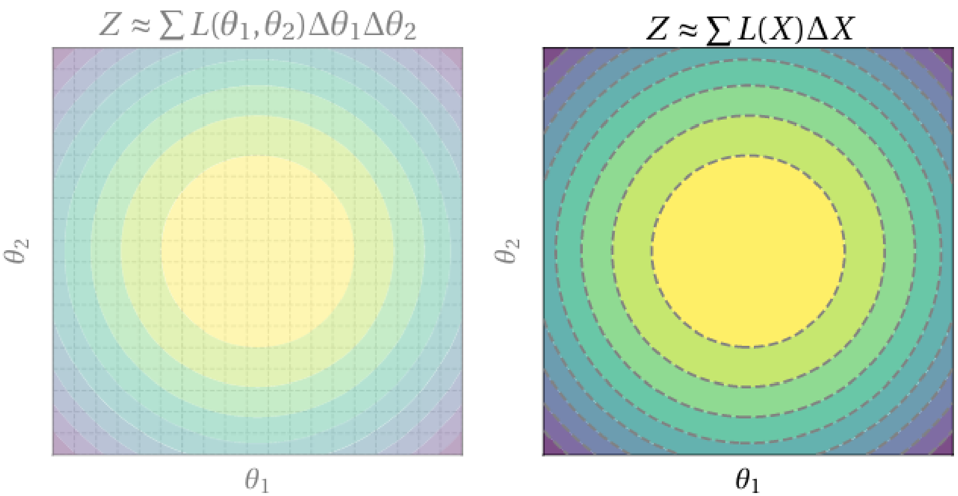
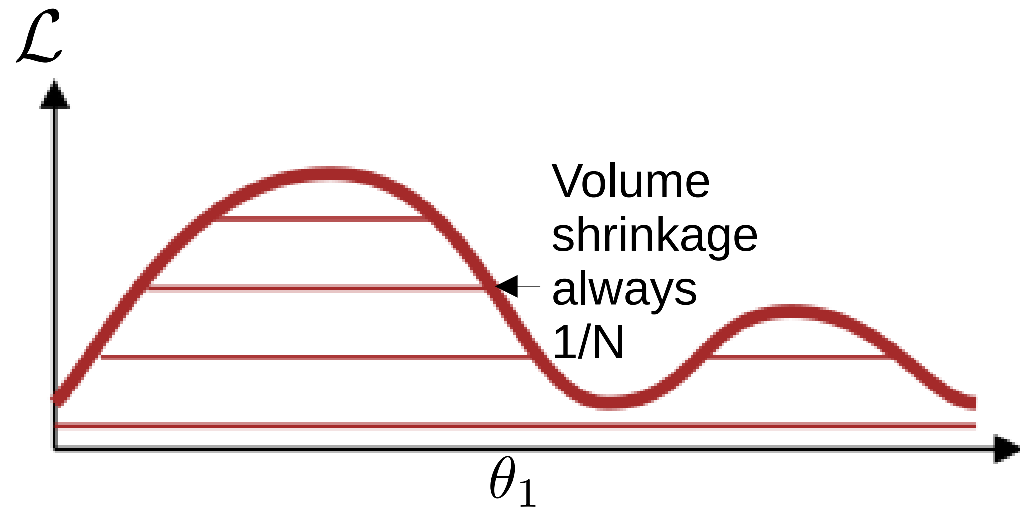


Figure 1 | **Illustrations of NS algorithm.**

Ashton, Buchner, et al. 2022, Nature Reviews Methods Primers



Riemann integration   Lebesgue integration



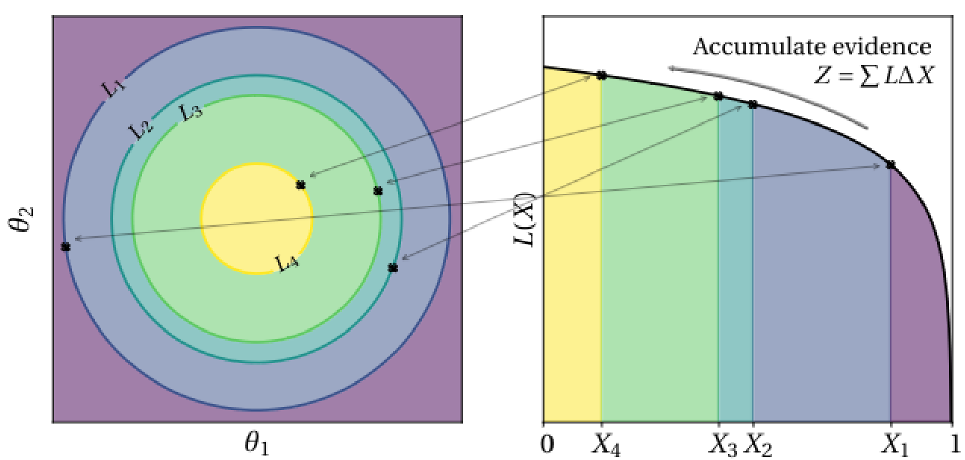


Figure 1 | **Illustrations of NS algorithm.**

Systematic literature review
"Nested Sampling Methods" Buchner (2023)

>100 papers

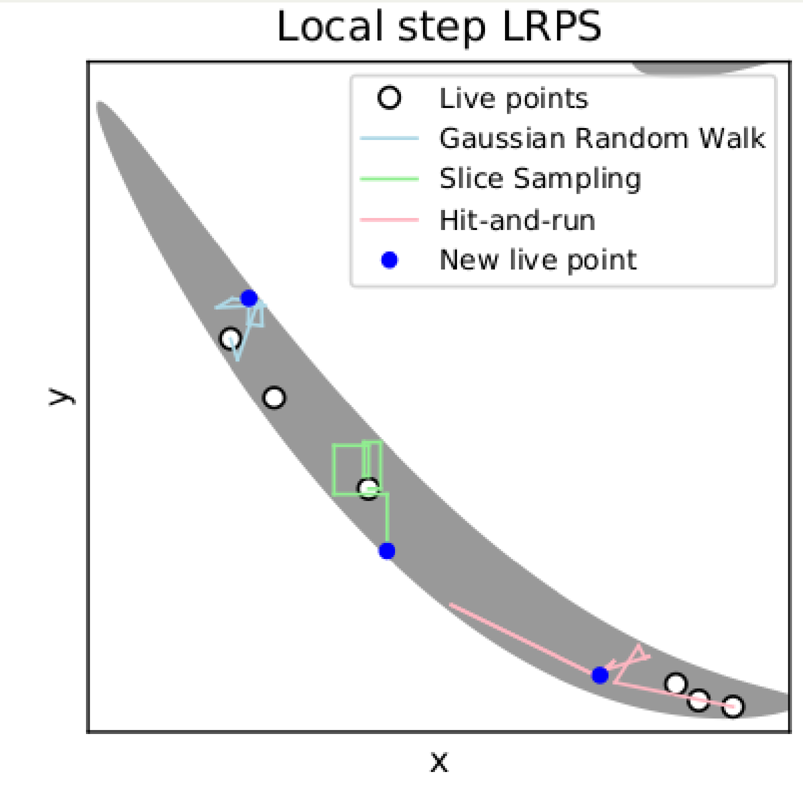
- Theory
- Estimators
- Termination
- Diagnostics
- variations:
  - Soft constraint
  - Variable number of live points
  - Parallelisation
- Likelihood-restricted prior sampling (LRPS)

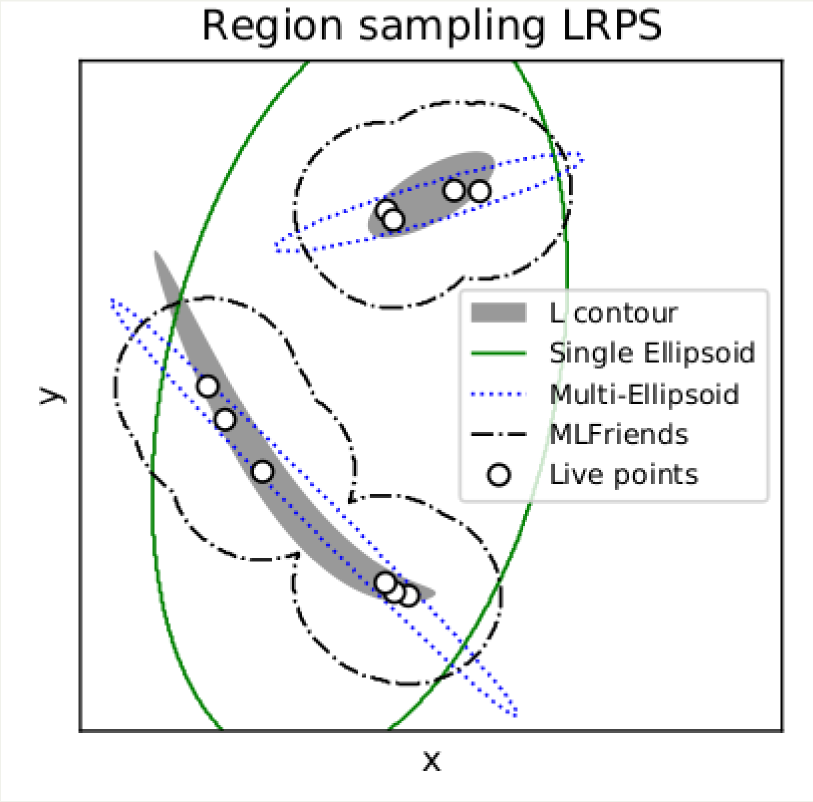
Convergence proof of Z and posterior :
e.g. Evans (2007), Chopin&Robert (2010)

Buchner (2022),
Statistics Surveys

Missing ingredients

- MCMC: Insert tuned transition kernel
- NS: Likelihood restricted prior sampling (LRPS)
  - General solutions: MultiNest, MCMC, HMCMC, Galilean, RadFriends (UltraNest), PolyChord





Animation:
https://johannesbuchner.github.io/mcmc-demo/app.html#RadFriends-NS,standard
(via chi-feng.github.io)

So what is BXA?

Buchner+14 (1326 citations)

Idea: make
physical parameter inference &
model comparison
easy & practical

XMM2Athena
319,565 X-ray sources
processed (Webb+23)

parallelisation,
resuming
sophisticated, robust

inference engine

BXA

community models
fully-fledged
fitting
data formats
environment
sherpa
pyxspec
(threeml)
(spex)

based on nested sampling

MultiNest
UltraNest

+ background models
+ some visualisation tools

An evolving software landscape

- Xspec, Sherpa, SPEX; 3ML

  maintainance is institutional effort

  Xspec models a community focal point

- 2014: BXA: xspec/sherpa plug-in for modern inference algorithms

  Buchner+14

- 2022: Model emulators

  Kerzendorf+22
  Matzeu+22

- 2024: Diff PPL: e.g. jaxspec

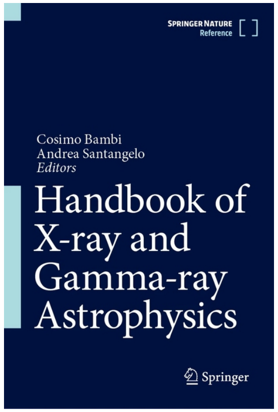
  Dupourqué+24
  Barret+24

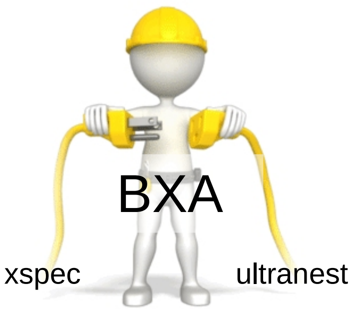
  – Require re-implementing models!

- Missing? partially diff XSF

Take-aways

- book chapter contents & discussion
- Nested sampling & model comparison
- The evolving software landscape





Disclaimer #1:

members of this group have discussed statistical aspects for a long time

Book chapter transfers knowledge from modern statistical literature to X-ray astronomy

Disclaimer #2:

unknown systematics / calibration can limit the statistical modelling

importance of IACHEC & cross-mission calibration