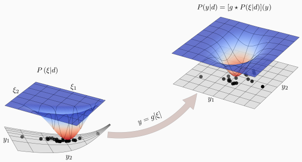
From Bayesian estimators to ML

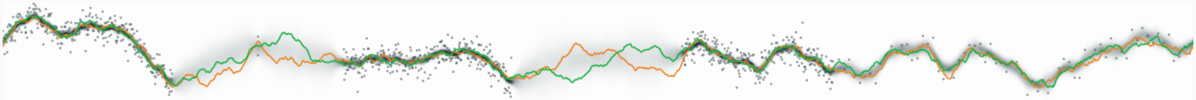
FOR FIELD-LIKE SIGNALS



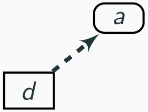
Philipp Frank[1]

1st ErUM-IFT Collaboration Meeting: MPA, Garching, Germany, November 22, 2023

(1) Max-Planck Institute for Astrophysics MPA, Garching, Germany

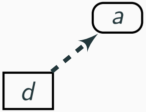


Probabilistic (Bayesian) Estimators



Given data d $\rightarrow$ obtain answers a about a system

Probabilistic (Bayesian) Estimators



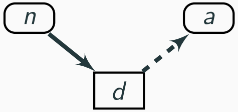
Given data d $\rightarrow$ obtain answers a about a system

Probabilistic estimator

$$\hat{a} = E\left(d; M\right) \quad .$$

With: d = Data, M = Model.

Probabilistic (Bayesian) Estimators



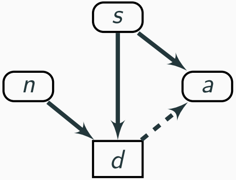
Given data d $\rightarrow$ obtain answers a about a system

Probabilistic estimator

$$\hat{a} = E\left(d; M\right) = \int a\, \mathcal{P}\left(a|d, M\right) \mathrm{d}a \qquad .$$

With: $d =$ Data, $M =$ Model.

Probabilistic (Bayesian) Estimators



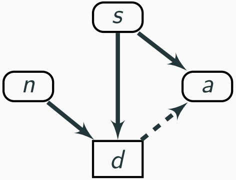
Given data d $\rightarrow$ obtain answers a about a system

Probabilistic estimator

$$\hat{a} = E\left(d; M\right) = \int a\, \mathcal{P}\left(a|d, M\right) \mathrm{d}a = \int a\left(s\right)\ \mathcal{P}\left(s|d, M\right) \mathrm{d}s\ .$$

With: $d =$ Data, $s =$ Signal, $M =$ Model.

Probabilistic (Bayesian) Estimators



Given data d $\rightarrow$ obtain answers a about a system

Probabilistic estimator

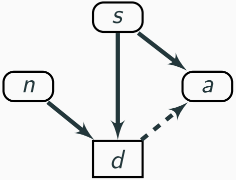
$$\hat{a} = E\left(d; M\right) = \int a\, \mathcal{P}\left(a|d, M\right) \mathrm{d}a = \int a\left(s\right)\ \mathcal{P}\left(s|d, M\right) \mathrm{d}s\ .$$

With: $d =$ Data, $s =$ Signal, $M =$ Model.

Product rule aka Bayes' Theorem

$$\mathcal{P}\left(s|d, M\right) = \frac{\mathcal{P}\left(s, d|M\right)}{\int \mathcal{P}\left(s, d|M\right)\ \mathrm{d}s} \qquad .$$

Probabilistic (Bayesian) Estimators



Given data d $\rightarrow$ obtain answers a about a system

Probabilistic estimator

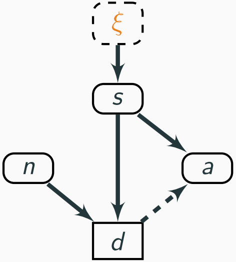
$$\hat{a} = E\left(d; M\right) = \int a\, \mathcal{P}\left(a|d, M\right) \mathrm{d}a = \int a\left(s\right)\; \mathcal{P}\left(s|d, M\right) \mathrm{d}s\;.$$

With: $d =$ Data, $s =$ Signal, $M =$ Model.

Product rule aka Bayes' Theorem

$$\mathcal{P}\left(s|d, M\right) = \frac{\mathcal{P}\left(s, d|M\right)}{\int \mathcal{P}\left(s, d|M\right)\; \mathrm{d}s} = \frac{\overbrace{\mathcal{P}\left(d|s \qquad , M\right)}^{\text{Likelihood}}\; \overbrace{\mathcal{P}\left(s|M\right)}^{\text{Prior}}}{\int \mathcal{P}\left(s, d|M\right)\; \mathrm{d}s}\;.$$

Probabilistic (Bayesian) Estimators



Given data d $\rightarrow$ obtain answers a about a system

Probabilistic estimator

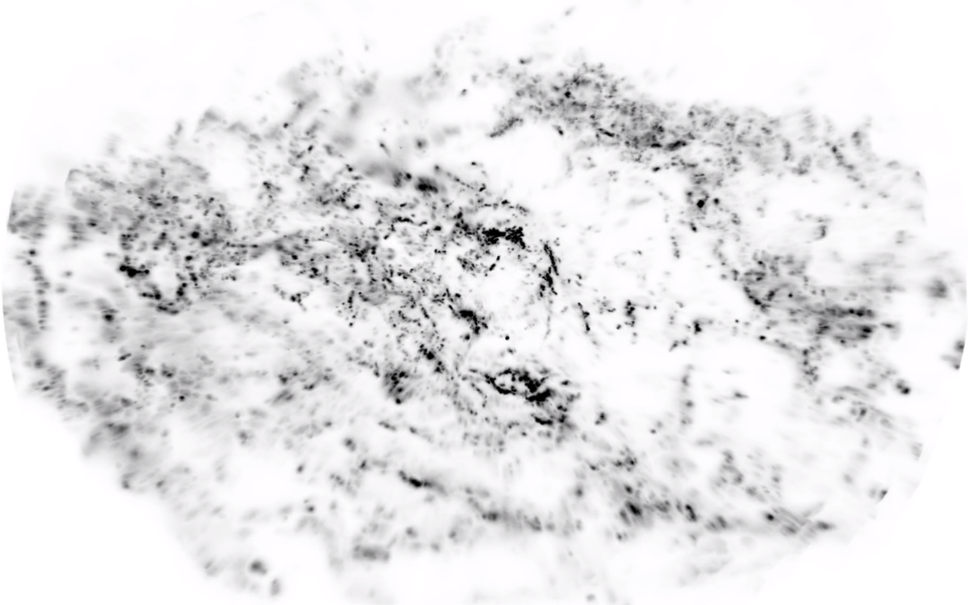
$$\hat{a} = E\left(d; M\right) = \int a\, \mathcal{P}\left(a|d, M\right) \mathrm{d}a = \int a\left(s\right)\, \mathcal{P}\left(s|d, M\right) \mathrm{d}s\ .$$

With: d = Data, s = Signal, M = Model.

Product rule aka Bayes' Theorem

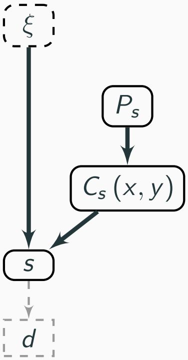
$$\mathcal{P}\left(\xi|d, M\right) = \frac{\mathcal{P}\left(\xi, d|M\right)}{\int \mathcal{P}\left(\xi, d|M\right)\, \mathrm{d}\xi} = \frac{\overbrace{\mathcal{P}\left(d|s\left(\xi\right), M\right)}^{\text{Likelihood}}\ \overbrace{\mathcal{N}\left(\xi; 0, \mathbb{1}\right)}^{\text{Prior}}}{\int \mathcal{P}\left(\xi, d|M\right)\, \mathrm{d}\xi}\ .$$

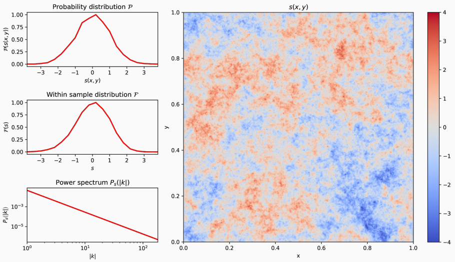
With: ξ = Parameters.



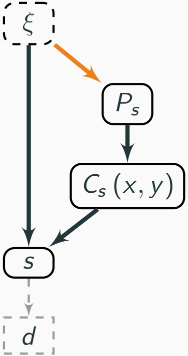
Priors - Gaussian & generative processes

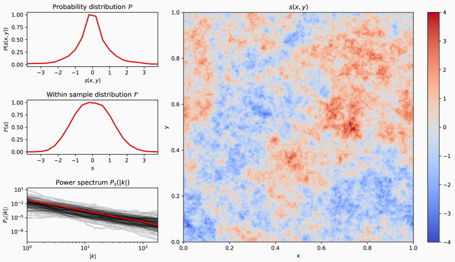
Priors - Gaussian & generative processes



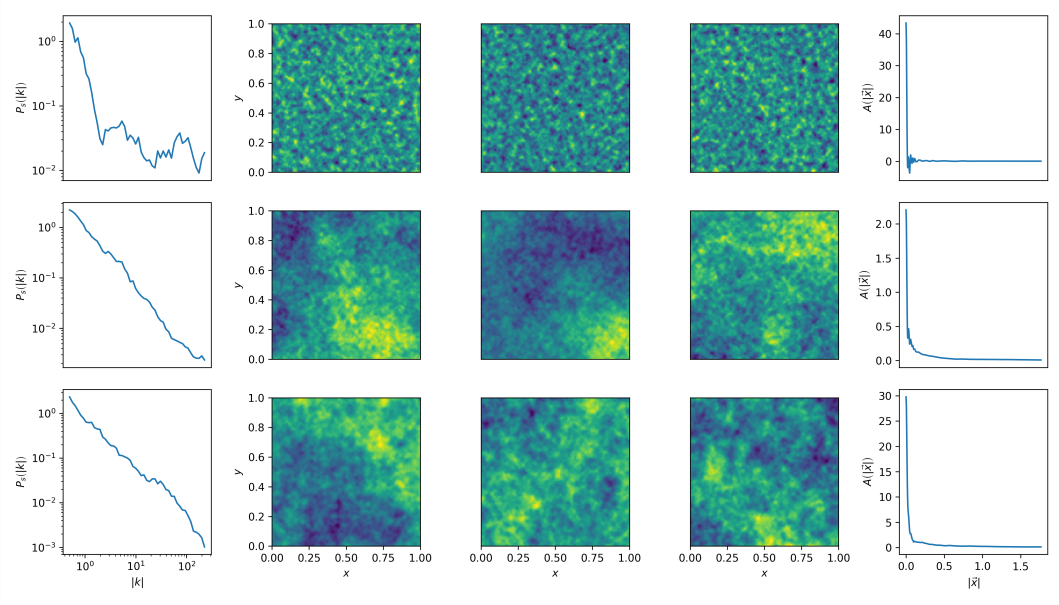


Priors - Gaussian & generative processes

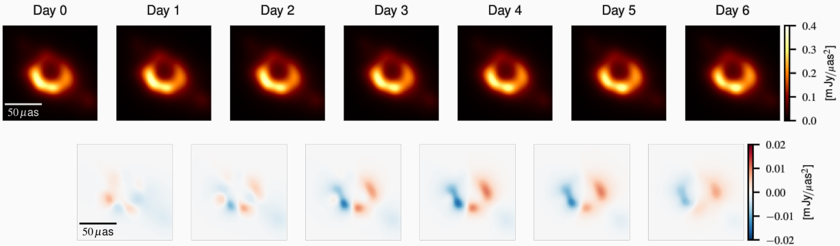


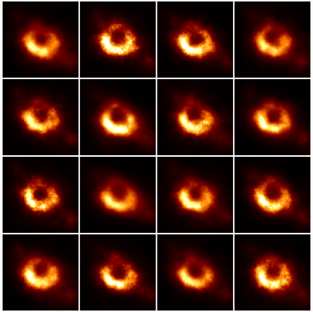


Priors - Gaussian & generative processes



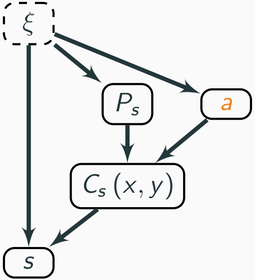
Priors - VLBI imaging of M87[1]

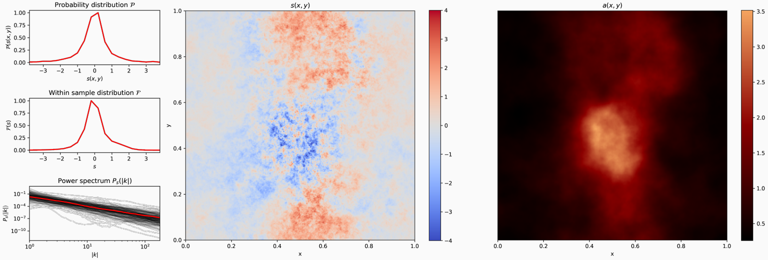




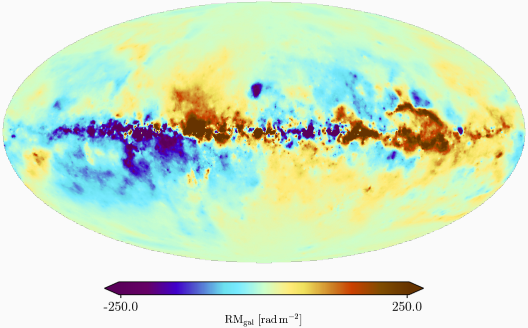
[1]Arras, Frank, Haim, et al. 2022.

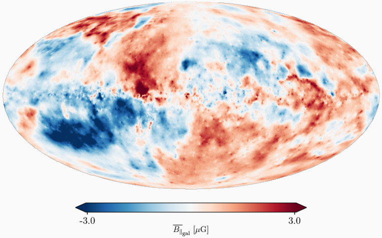
Priors - Gaussian & generative processes

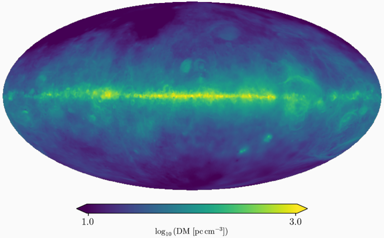




Priors - Faraday tomography[2]

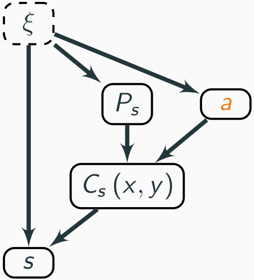


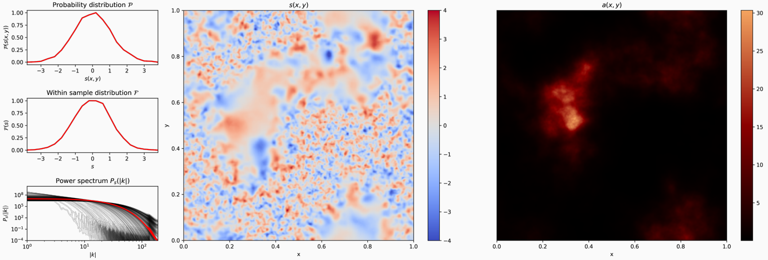




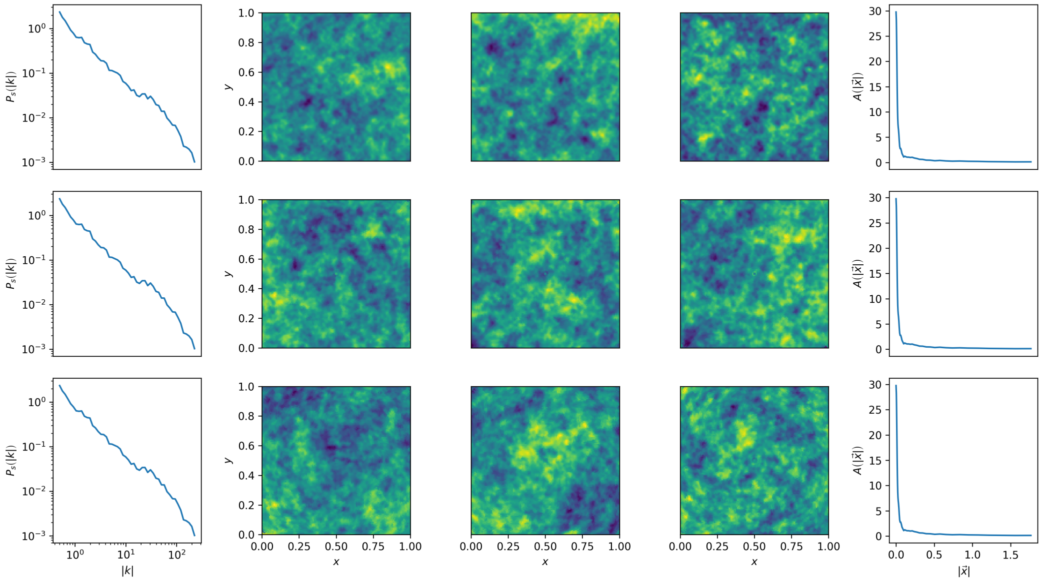
[2] Hutschenreuter, Haverkorn, Frank, et al. 2023.

Priors - Gaussian & generative processes

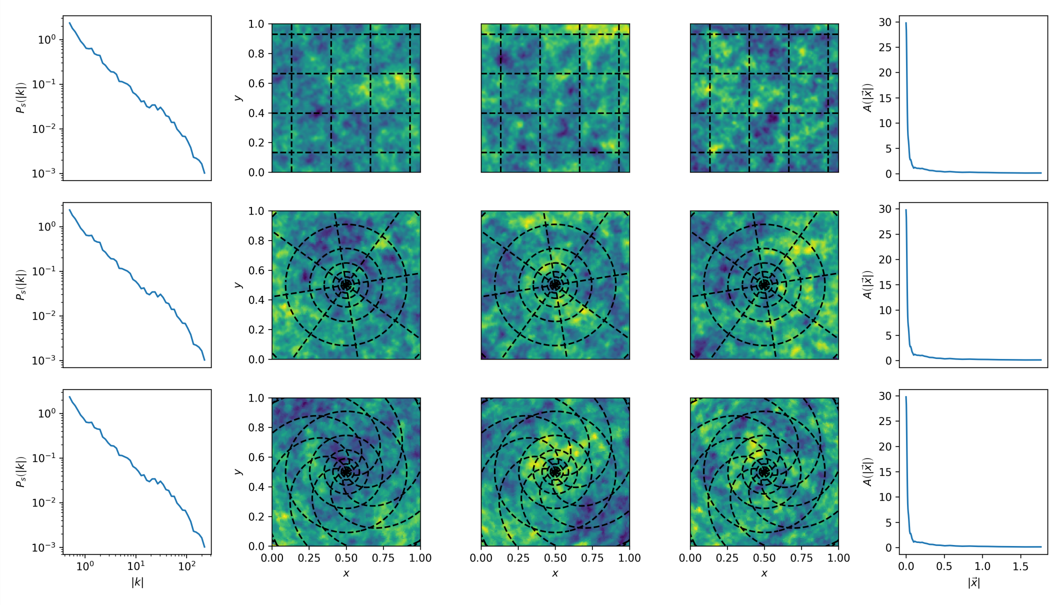




Priors - Gaussian & generative processes

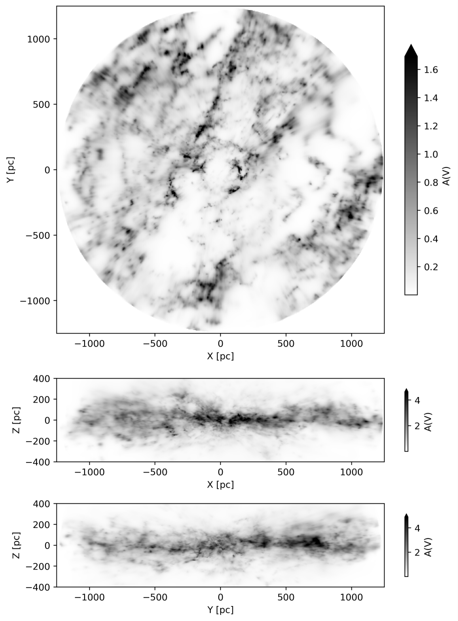


Priors - Gaussian & generative processes



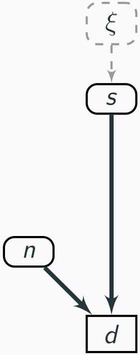
Priors - GAIA 3D dust tomography[3]

- HEALPix angular + Log-radial grid
- Log-Normal process on $\sim 0.66 \times 10^9$ voxels
- $\sim 80 \times 10^6$ measurements



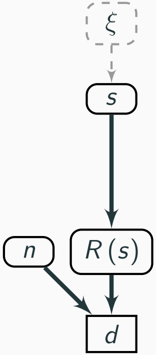
[3] Edenhofer, Zucker, Frank, et al. 2023.

Likelihoods - Instrument response



[4] Credit: https://cxc.harvard.edu/

Likelihoods - Instrument response



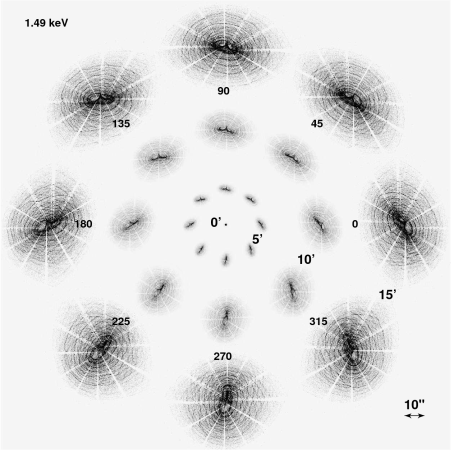
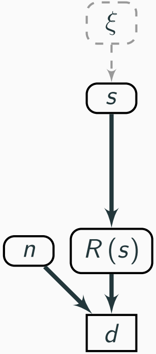


Figure 1: Simulated Chandra PSF[4]

[4]Credit: https://cxc.harvard.edu/

Likelihoods - Instrument response



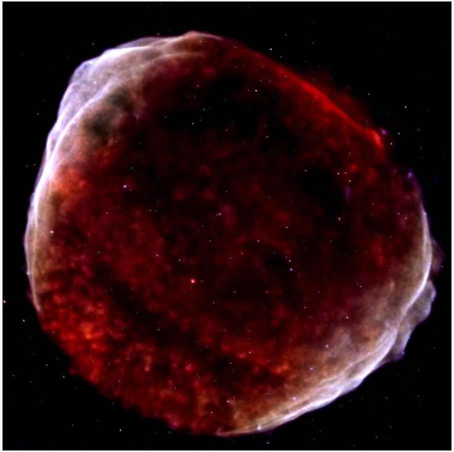
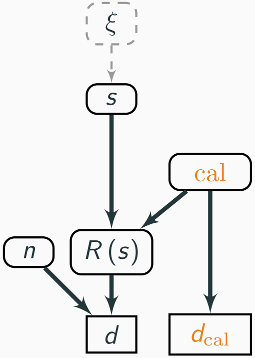


Figure 1: SN1006 from Chandra data[4]

[4]Westerkamp, Eberle, Guardiani, et al. 2023.



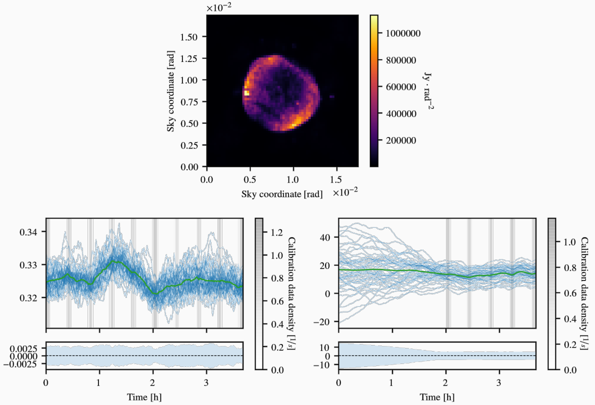
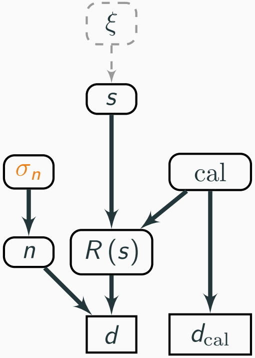


Figure 1: SN1006 from VLA data[4]

[4] Arras, Frank, Leike, et al. 2019.

Likelihoods - Instrument response



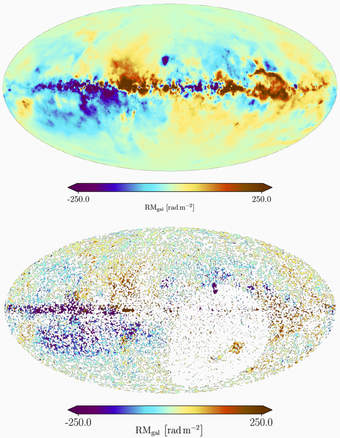
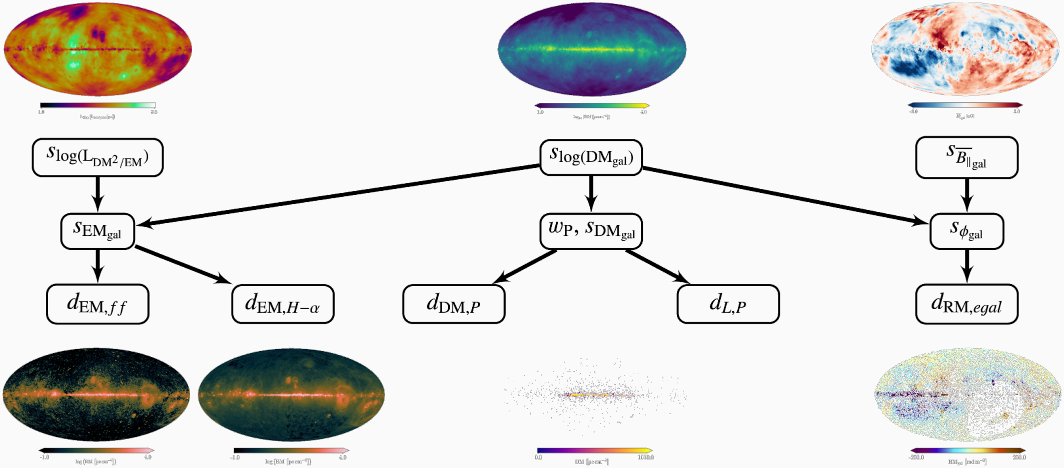


Figure 1: Faraday sky[4]

[4] Hutschenreuter, Haverkorn, Frank, et al. 2023.

Likelihoods - Faraday tomography[5]



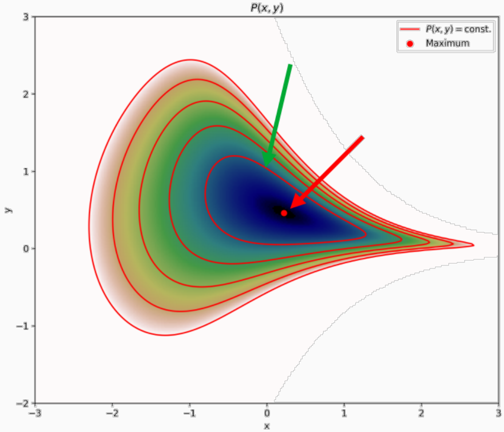
[5]Hutschenreuter, Haverkorn, Frank, et al. 2023.

Approximate Inference

Posterior expectation

$$\langle f(\xi)\rangle_{\mathcal{P}(\xi|d)} = \int f \; \mathrm{d}\mathcal{P} = \int f(\xi)\; \mathcal{P}(\xi|d)\; \mathrm{d}\xi \approx \int f(\xi)\; \mathcal{Q}(\xi|d)\; \mathrm{d}\xi$$

Function: $f(\xi)$; Posterior: $\mathcal{P}(\xi|d)$; parameters: ξ; data: d.

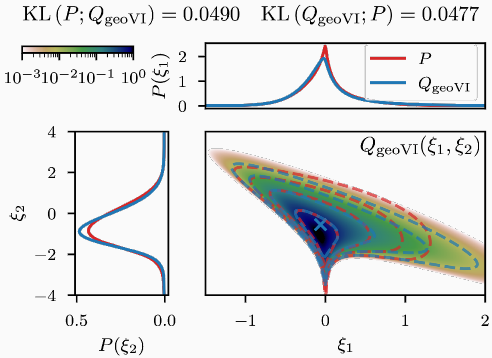


Approximate Inference - Variational Inference (VI)

Kullback-Leibler divergence

$$\mathrm{KL}\left[\mathcal{Q}_\sigma || \mathcal{P}\right] = -\int \log\left(\frac{\mathcal{P}(\xi|d)}{\mathcal{Q}_\sigma(\xi)}\right) \mathcal{Q}_\sigma(\xi)\, \mathrm{d}\xi$$

Posterior: $\mathcal{P}(\xi|d)$; Approximation: $\mathcal{Q}_\sigma(\xi)$; Variational parameters: σ.



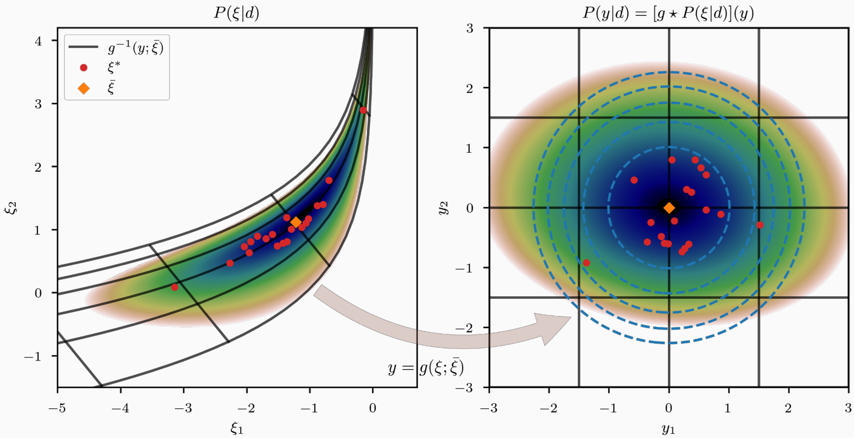
[6]

[6]Frank, Leike, and Enßlin 2021.

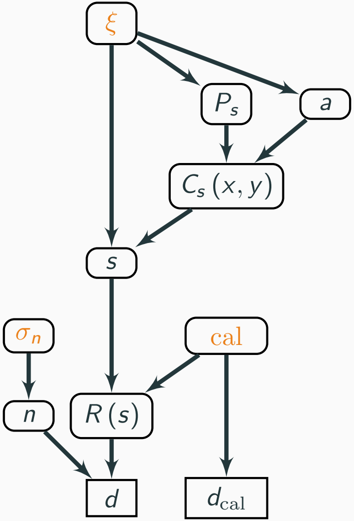
Geometric Variational Inference

Normalizing coordinate transformation $y = g_\sigma(\xi)$ with $\sigma = \bar{\xi}$.

Approximate distribution $\mathcal{Q}(y) = \mathcal{N}(y; 0, \mathbb{1})$

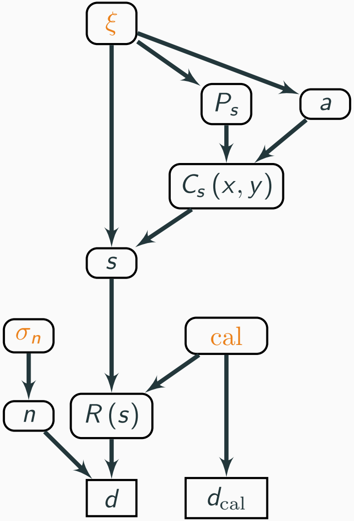


Conclusion - Outlook



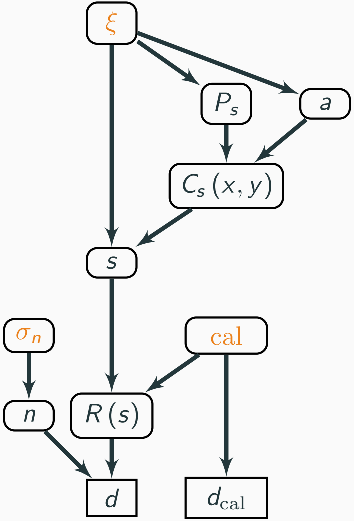
- State of IFT & NIFTy:

Conclusion - Outlook



- State of IFT & NIFTy:
  - (Fairly) generic framework for fields
  - Fast prototyping for models
  - Scalable approximate inference

Conclusion - Outlook



- State of IFT & NIFTy:
  - (Fairly) generic framework for fields
  - Fast prototyping for models
  - Scalable approximate inference
- ML developments
  - Probabilistic ML for fast (repeated) inference
  - Automated model extension / component discovery

References

- Arras, Philipp et al. (July 2019). "Unified radio interferometric calibration and imaging with joint uncertainty quantification". In: *A&A* 627, A134, A134. DOI: 10.1051/0004-6361/201935555. arXiv: 1903.11169 [astro-ph.IM].
- Arras, Philipp et al. (2022). "Variable structures in M87* from space, time and frequency resolved interferometry". In: *Nature Astronomy* 6.2, pp. 259–269.
- Edenhofer, Gordian et al. (Aug. 2023). "A Parsec-Scale Galactic 3D Dust Map out to 1.25 kpc from the Sun". In: *arXiv e-prints*, arXiv:2308.01295, arXiv:2308.01295. DOI: 10.48550/arXiv.2308.01295. arXiv: 2308.01295 [astro-ph.GA].
- Frank, Philipp, Reimar Leike, and Torsten A. Enßlin (2021). "Geometric Variational Inference". In: *Entropy* 23.7. ISSN: 1099-4300. DOI: 10.3390/e23070853. URL: https://www.mdpi.com/1099-4300/23/7/853.

References ii

- Hutschenreuter, Sebastian et al. (Apr. 2023). "Disentangling the Faraday rotation sky". In: *arXiv e-prints*, arXiv:2304.12350, arXiv:2304.12350. DOI: 10.48550/arXiv.2304.12350. arXiv: 2304.12350 [astro-ph.GA].
- Westerkamp, Margret et al. (Aug. 2023). "First spatio-spectral Bayesian imaging of SN1006 in X-ray". In: *arXiv e-prints*, arXiv:2308.09176, arXiv:2308.09176. DOI: 10.48550/arXiv.2308.09176. arXiv: 2308.09176 [astro-ph.HE].