Information Field Theory

As as Tool in Cosmic Ray data analysis

Simon Strähnz – February 7, 2024

IFT Basics

Karlsruher Institut für Technologie

- Bayes Theorem

$$P(\phi|d) = \frac{P(d|\phi)P(\phi)}{P(d)}$$

likelihood: $P(d|\phi)$; prior: $P(\phi)$; posterior: $P(\phi|d)$; evidence: $P(d)$

- Information Hamiltonian

$$\mathcal{H}(d,\phi) = -\ln(P(d,\phi)) \qquad P(\phi|d) \hat{=} \exp(-\mathcal{H}(d,\phi))$$

- Measurement Equation $d = R(s) + n$

d: Data (measurement) R: Instrument response n: Noise — known

s: Signal — unknown

IFT Basics

- Measurement Equation $d_i = R(s_i) + n_i$
- Information Hamiltonian

$$\mathcal{H}(d, \phi) = \sum_{i=0}^{N} \mathcal{H}(d_i, \phi)$$

- How to solve this?
  - Minimize Hamiltonian
  - Minimize Gibbs free energy
  - Variational inference

Variational Inference

Karlsruher Institut für Technologie

- Approximate posterior with known distribution
  - Often a Gaussian
- Minimize the difference between the approximate and actual posterior
  - Kullback-Leibler divergence
- Metric Gaussian Variational Inference
- Geometric Variational Inference

MGVI

- Start with an estimate of the model parameters $\bar{\theta}$
- Estimate the covariance matrix Θ
- Fix covariance and optimise the Kullback-Leibler divergence
- Set estimate to optimised parameters
- Repeat until no further improvement is made

MGVI

Karlsruher Institut für Technologie

- Kullback-Leibler diveregence:

$$\mathcal{D}_{KL}(Q_\eta(\theta)|P(\theta|d)) \hat{=} \underbrace{\langle \mathcal{H}(d,\theta)\rangle_{Q_\eta}}_{\text{Cross Entropy}} - \frac{1}{2}\underbrace{\ln|2\pi e\Theta|}_{\text{Shannon Entropy}}$$

- Derivatives

$$\frac{\partial}{\partial\bar{\theta}}\mathcal{D}_{KL} = \left\langle \frac{\partial}{\partial\theta}\mathcal{H}(d,\theta)\right\rangle_{Q_\eta}$$

$$\frac{\partial}{\partial\Theta}\mathcal{D}_{KL} = \frac{1}{2}\left\langle \frac{\partial}{\partial\theta\partial\theta^\dagger}\mathcal{H}(d,\theta)\right\rangle_{Q_\eta} - \frac{1}{2}\Theta^{-1}$$

MGVI

- Inverse covariance

not positive definite!

$$\Theta^{-1} = \left\langle \frac{\partial}{\partial\theta\partial\theta^\dagger}\mathcal{H}(d,\theta) \right\rangle_{Q_\eta}$$

$$= \left\langle \frac{\partial\mathcal{H}(d,\theta)}{\partial\theta}\frac{\partial\mathcal{H}(d,\theta)}{\partial\theta^\dagger} \right\rangle_{Q_\eta} - \left\langle \frac{1}{P(d,\theta)}\frac{\partial^2 P(d,\theta)}{\partial\theta\partial\theta^\dagger} \right\rangle_{Q_\eta}$$

- Fisher metric

$$I = I_d + I_\theta$$

$$= \left\langle \frac{\partial\mathcal{H}(d,\theta)}{\partial\theta}\frac{\partial\mathcal{H}(d,\theta)}{\partial\theta^\dagger} \right\rangle_{Q_\eta} + \left\langle \frac{\partial\mathcal{H}(\theta)}{\partial\theta}\frac{\partial\mathcal{H}(\theta)}{\partial\theta^\dagger} \right\rangle_{Q_\eta}$$

Standardisation

Karlsruher Institut für Technologie

- For easier computation the prior can be standardised:

$$\theta_i = f_i(\xi_i) \qquad\qquad \xi_i \hookleftarrow \mathcal{G}(\xi, 1)$$

- Then the Fisher metric becomes:

$$I = J(\xi)^\dagger I_d(f(\xi))J(\xi) + \mathbb{I}$$

- This changes the measurement equation

$$d = R \circ s(f(\xi)) + n$$

- And the Information Hamiltonian becomes

$$\mathcal{H}(d, \xi) = -\ln[\mathcal{N}\{d - R \circ s(f(\xi))\}] + \frac{1}{2}\xi^\dagger \mathbb{I} \xi$$

Sampling

Karlsruher Institut für Technologie

$$\theta_*^i \hookleftarrow \mathcal{G}(\theta - 0, \Theta)$$
$$\eta_*^i \hookleftarrow \mathcal{G}(\eta - 0, 1)$$

- Expectations can be approximated by sampling

$$\langle \mathcal{H}(d, \theta) \rangle_{\mathcal{G}(\theta - \bar{\theta}, \Theta)} \approx \frac{1}{N} \sum_{i=1}^{N} \mathcal{H}(d, \bar{\theta} + \theta_*^i)$$

- Drawing samples from the inverse covariance
  - Eigenbasis: $\Theta = Q \Lambda Q^\dagger$
  - Sample from white noise: $\theta_*^i = Q \sqrt{\Lambda} \eta_*^i \Rightarrow \theta_*^i \hookleftarrow \mathcal{G}(\theta, \Theta)$ ← only have inverse
  - Can draw from inverse: $\phi_*^i = Q \sqrt{\Lambda^{-1}} \eta_*^i \Rightarrow \theta_*^i \hookleftarrow \mathcal{G}(\phi, \Theta^{-1})$
  - Now apply covariance:

$$\Theta \phi_*^i = \theta_*^i$$

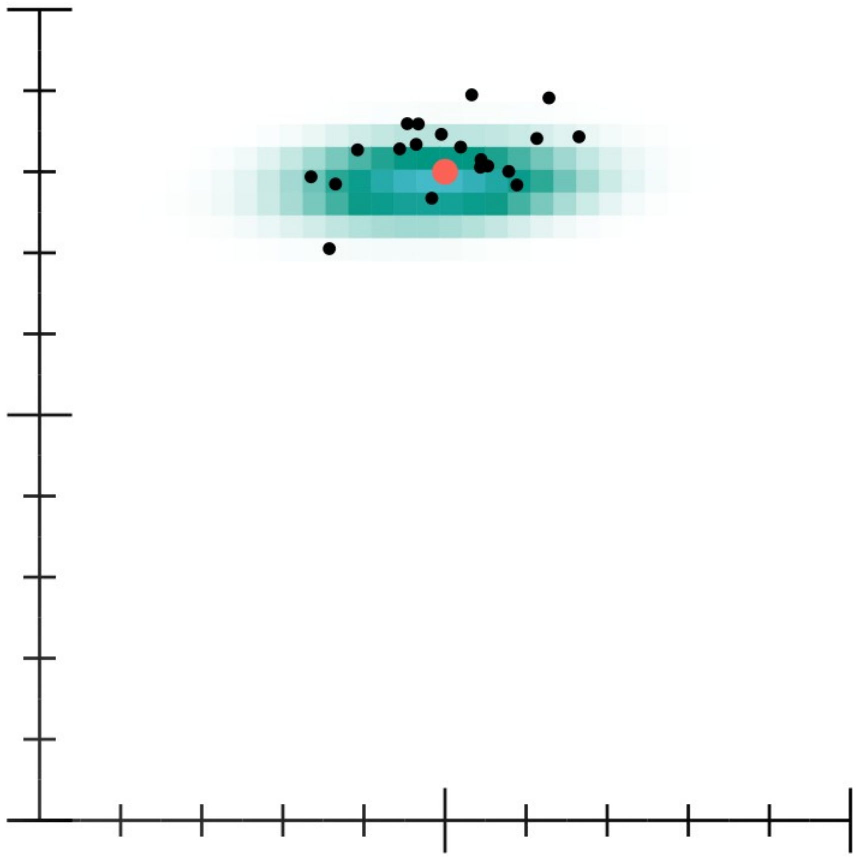
MGVI in practice

- Update covariance
- Optimize KL-divergence
- Update latent mean

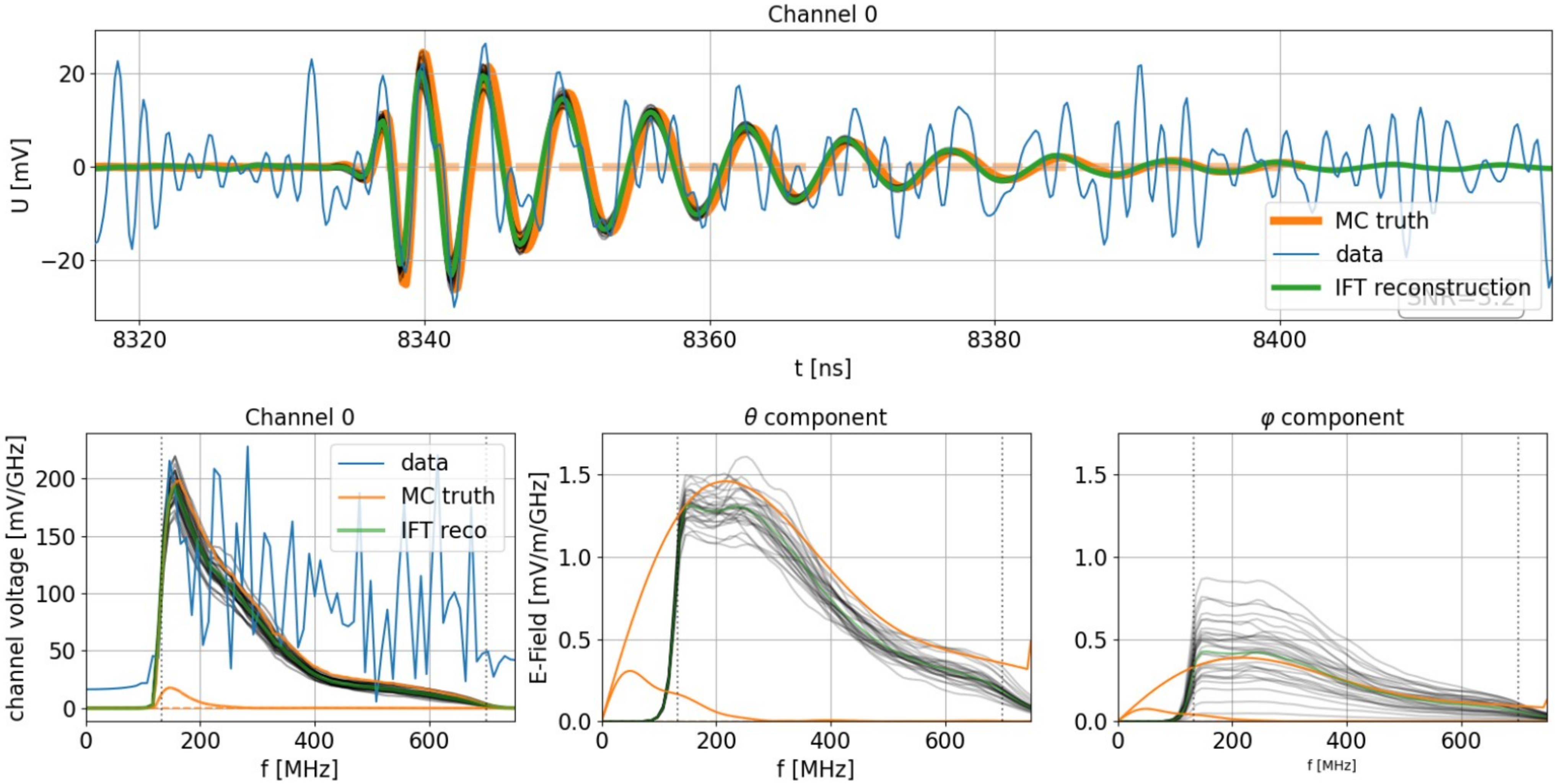
$$\bar{\theta} = (.5, .8, \ldots)$$

$$\xi = (1.82, -0.8, 1.52, \ldots)$$

$$\Theta = \begin{pmatrix} .01 & 0.0006 & \ldots \\ .0006 & 0.0009 & \ldots \\ \vdots & \vdots & \vdots \end{pmatrix}$$

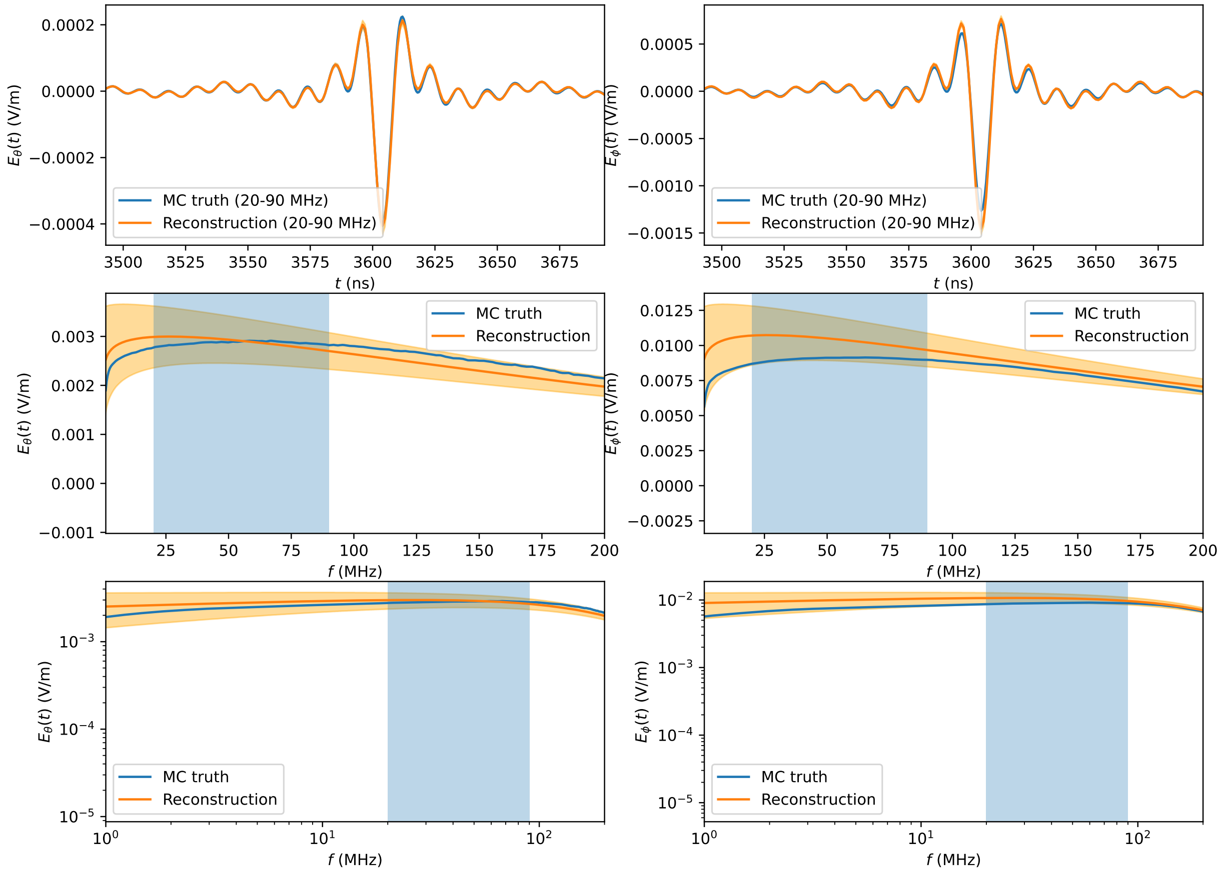


Examples



Welling et al. (2021)

Examples



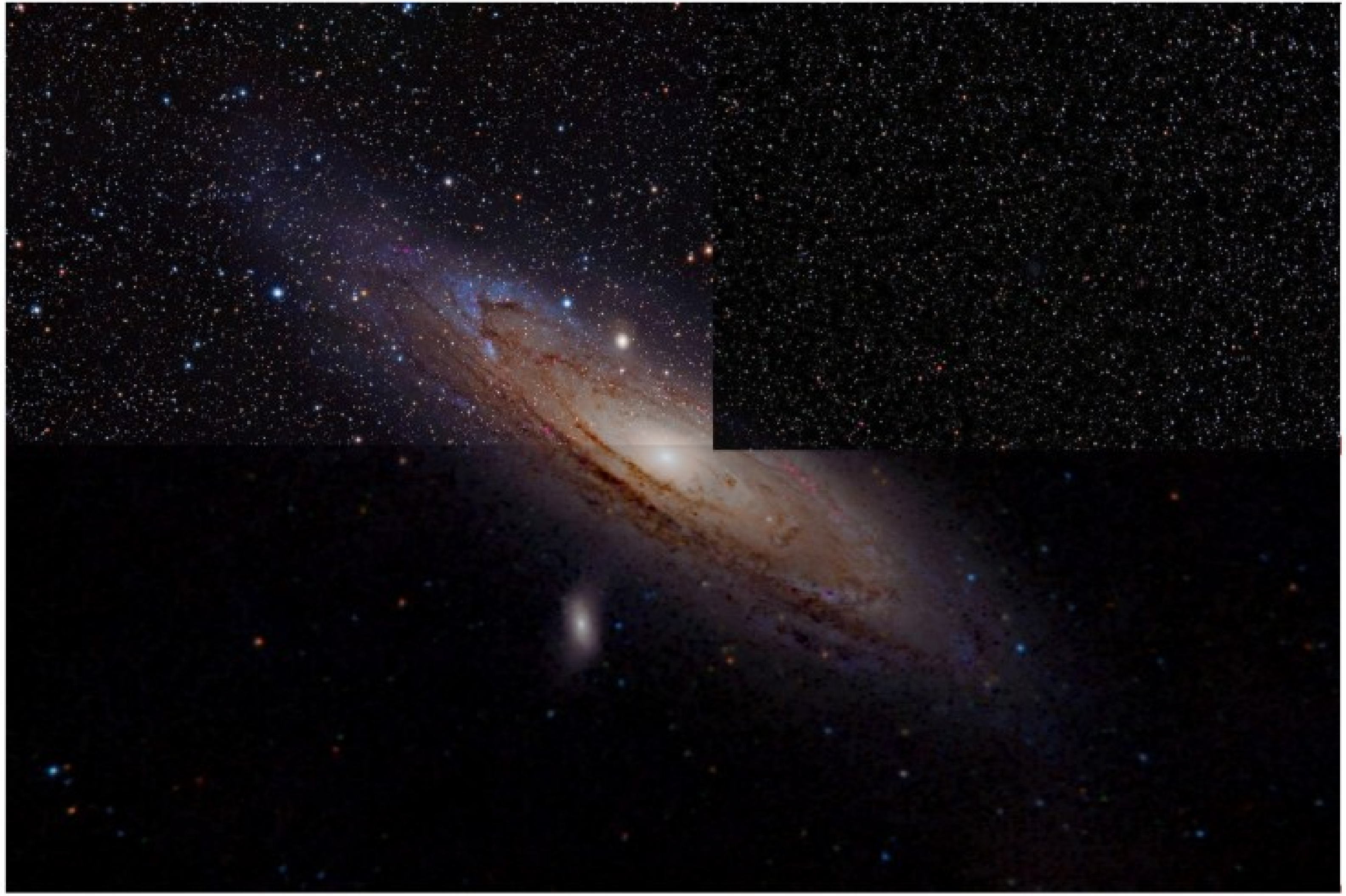
0.0002
0.0000
−0.0002
−0.0004
$E_\theta(t)$ (V/m)
MC truth (20-90 MHz)
Reconstruction (20-90 MHz)
3500 3525 3550 3575 3600 3625 3650 3675
t (ns)
0.0005
0.0000
−0.0005
−0.0010
−0.0015
$E_\phi(t)$ (V/m)
MC truth (20-90 MHz)
Reconstruction (20-90 MHz)
3500 3525 3550 3575 3600 3625 3650 3675
t (ns)
0.003
0.002
0.001
0.000
−0.001
$E_\theta(t)$ (V/m)
MC truth
Reconstruction
25 50 75 100 125 150 175 200
f (MHz)
0.0125
0.0100
0.0075
0.0050
0.0025
0.0000
−0.0025
$E_\phi(t)$ (V/m)
MC truth
Reconstruction
25 50 75 100 125 150 175 200
f (MHz)
10^{-3}
10^{-4}
10^{-5}
$E_\theta(t)$ (V/m)
MC truth
Reconstruction
10^0 10^1 10^2
f (MHz)
10^{-2}
10^{-3}
10^{-4}
10^{-5}
$E_\phi(t)$ (V/m)
MC truth
Reconstruction
10^0 10^1 10^2
f (MHz)

Karlsruhe Institute of Technology

Other use cases

- D^{3}PO and D^{4}PO
  - Reconstruction of photon emission fields
- RESOLVE
  - Interferometry radio imaging
- charm
  - Reconstructs expansions of Universe from type Ia supernovae
- starblade
  - Separates point sources from diffuse emission

Starblade



Knobmüller in Enßlin (2018)

Division V Physics and Mathematics
Institute of Experimental Particle Physics