Information field theory

Torsten Enßlin

MPI for Astrophysics

Ludwig Maximilian University Munich

IFT Team: Philipp Arras, Michael Bell,Vanessa Böhm, Sebastian Dorn, Philipp Frank, Maksim Greiner, Sebastian Hutschenreuter, Henrik Junklewitz, Jakob Knollmüller, Reimar Leike, Marco Selig, Theo Steininger, Niels Oppermann, Natalia Porqueres, Daniel Pumpe, Martin Reinecke, Valentina Vacca, ...many more

Information Field Theory

signal field: ∞ degrees of freedom
data set: finite
→ additional information needed

information:
physical laws, symmetries, continuity,
statistical homogeneity/isotropy, ...

most important:
space, time, & fields
are continuous

combining concrete (data) &
abstract (knowledge) information
→ **information theory for fields**

Information Field Theory

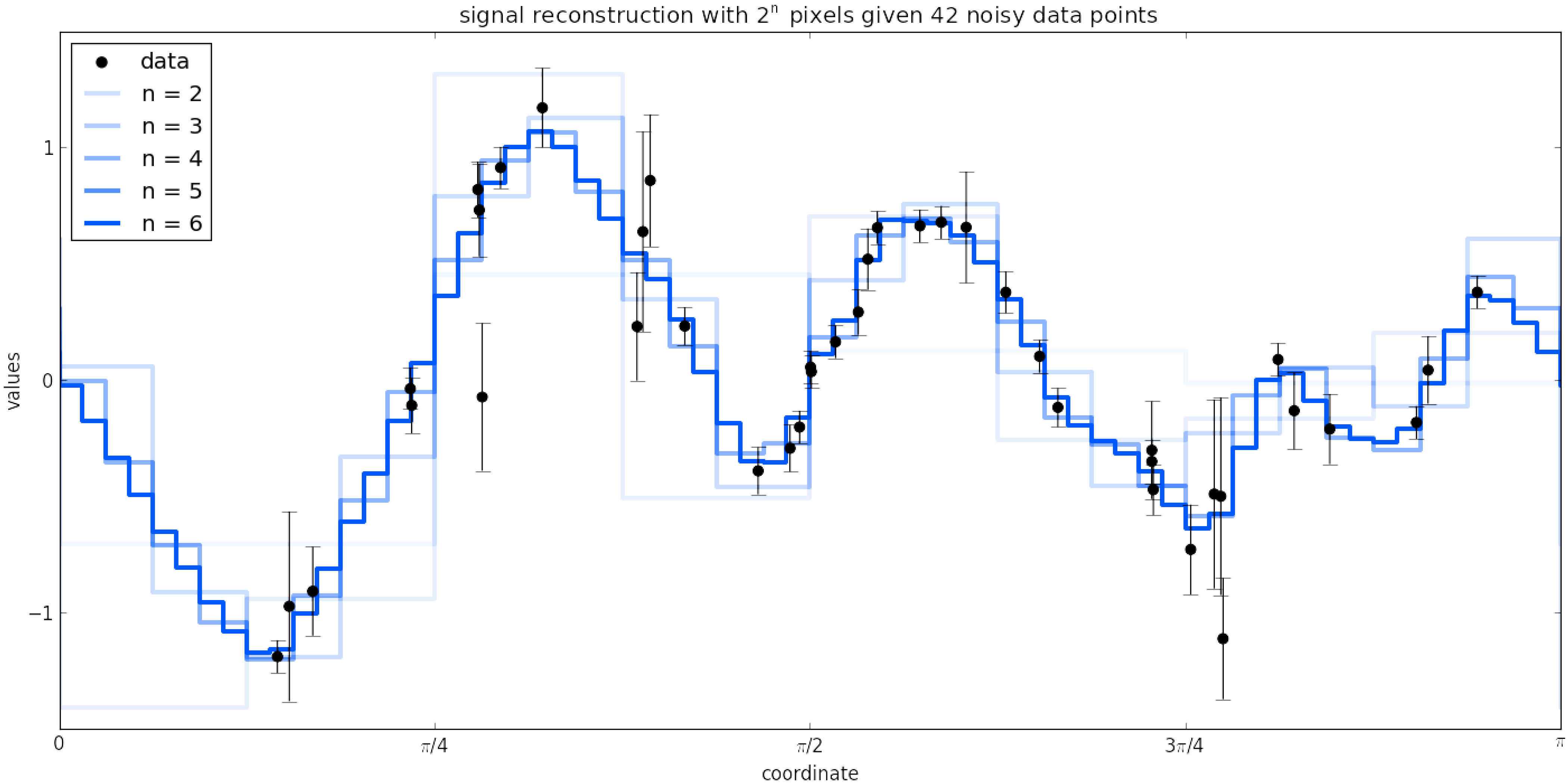
$s =$ signal

$d =$ data

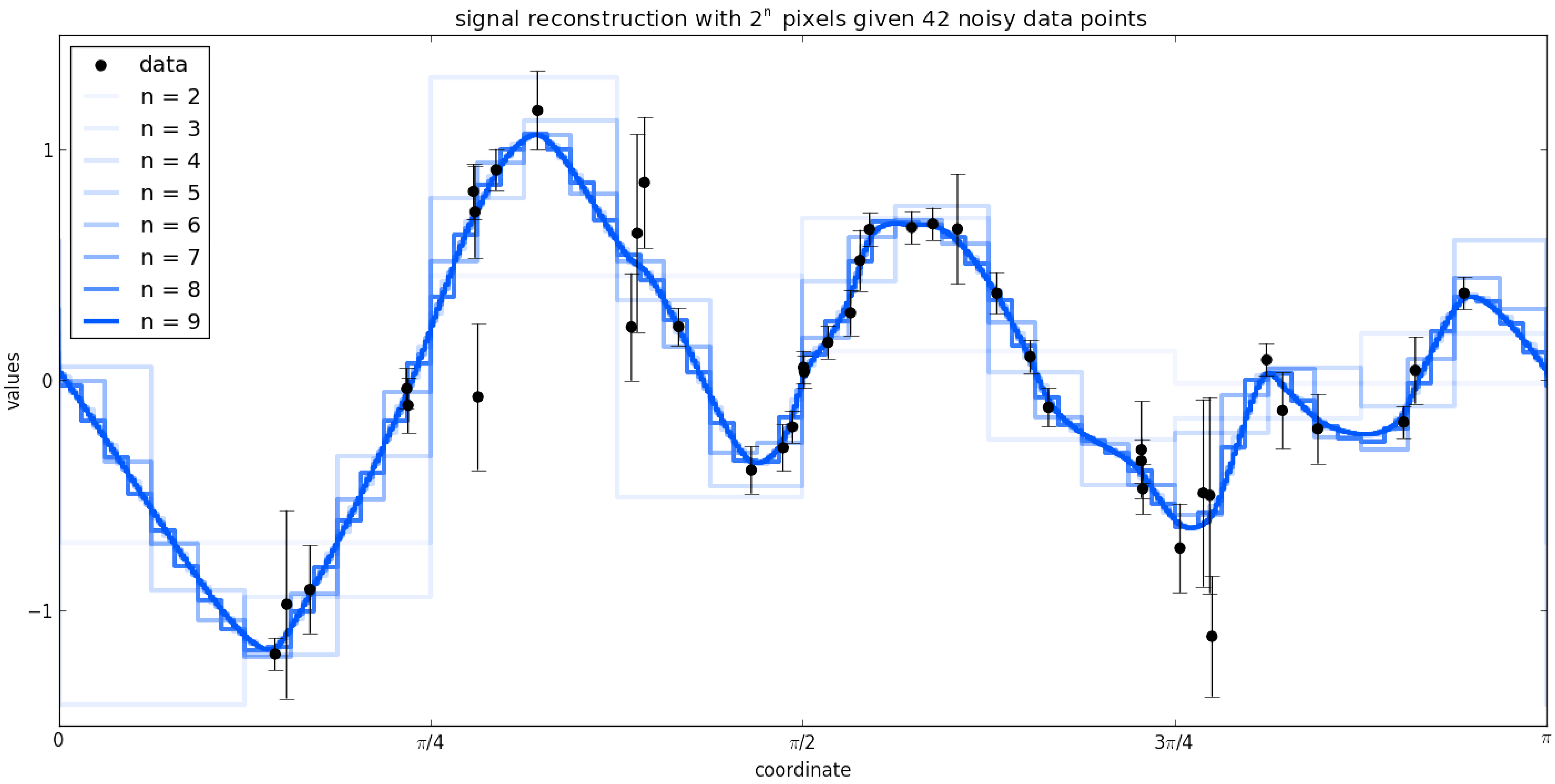
$$\mathcal{P}(s|d) = \frac{\mathcal{P}(d|s)\,\mathcal{P}(s)}{\mathcal{P}(d)}$$

$$\begin{aligned} H(d,s) &= -\log \mathcal{P}(d,s) \\ Z(d) &= \int \mathcal{D}s\, \mathcal{P}(d,s) \end{aligned}$$


Correlations
information field theory
in continuous space
calculations in discretized spaces



signal reconstruction with 2ⁿ pixels given 42 noisy data points
data
n = 2
n = 3
n = 4
n = 5
n = 6
1
0
−1
values
0
π/4
π/2
3π/4
π
coordinate

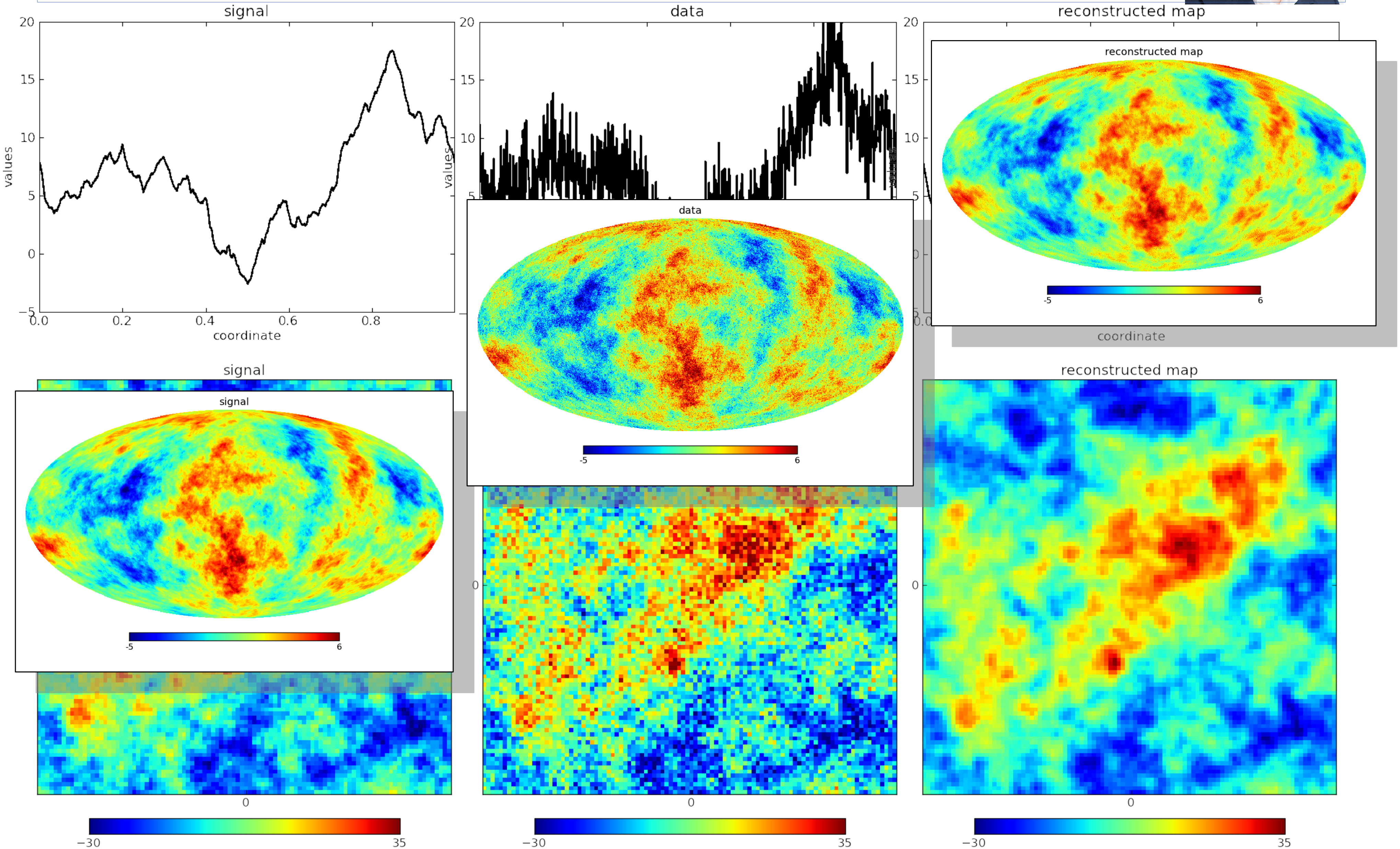


signal reconstruction with 2^n pixels given 42 noisy data points
data
n = 2
n = 3
n = 4
n = 5
n = 6
n = 7
n = 8
n = 9
values
1
0
−1
0
π/4
π/2
3π/4
π
coordinate

NIFTy - Numerical Information Field Theory

Selig et al. (2013), Steininger et al. (sub.)

Code @ https://gitlab.mpcdf.mpg.de/ift/NIFTy




Extended VLA
LOFAR
RESOLVE
ALMA
Junklewitz et al.
Arras et al.
APERTIF
Meerkat
ASKAP

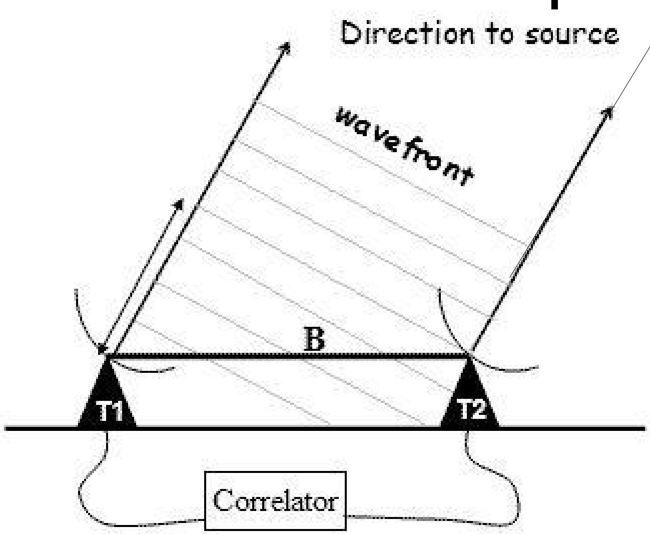
Data model

$$d = R\, e^{\color{red}{s}} + \color{red}{n}$$

known

unknown

known response



Gaussian priors:

$\mathcal{P}(s) = \mathcal{G}(s, \color{red}{S})$ unknown

$\mathcal{P}(n) = \mathcal{G}(n, N)$ known

Information Hamiltonian

$$\mathcal{H}(\boldsymbol{d}, \boldsymbol{s}, \boldsymbol{\tau}) = -\log \mathcal{P}(\boldsymbol{d}, \boldsymbol{s}, \boldsymbol{\tau})$$

likelihood

$$= \frac{1}{2}\left(\boldsymbol{d} - \boldsymbol{R}\mathrm{e}^{\boldsymbol{s}}\right)^{\dagger} \boldsymbol{N}^{-1}\left(\boldsymbol{d} - \boldsymbol{R}\mathrm{e}^{\boldsymbol{s}}\right) + \frac{1}{2}\log\left(\det\left[\boldsymbol{N}\right]\right)$$

prior / regularization

$$+ \frac{1}{2}\boldsymbol{s}^{\dagger}\boldsymbol{S}^{-1}\boldsymbol{s} + \frac{1}{2}\log\left(\det\left[\boldsymbol{S}\right]\right)$$

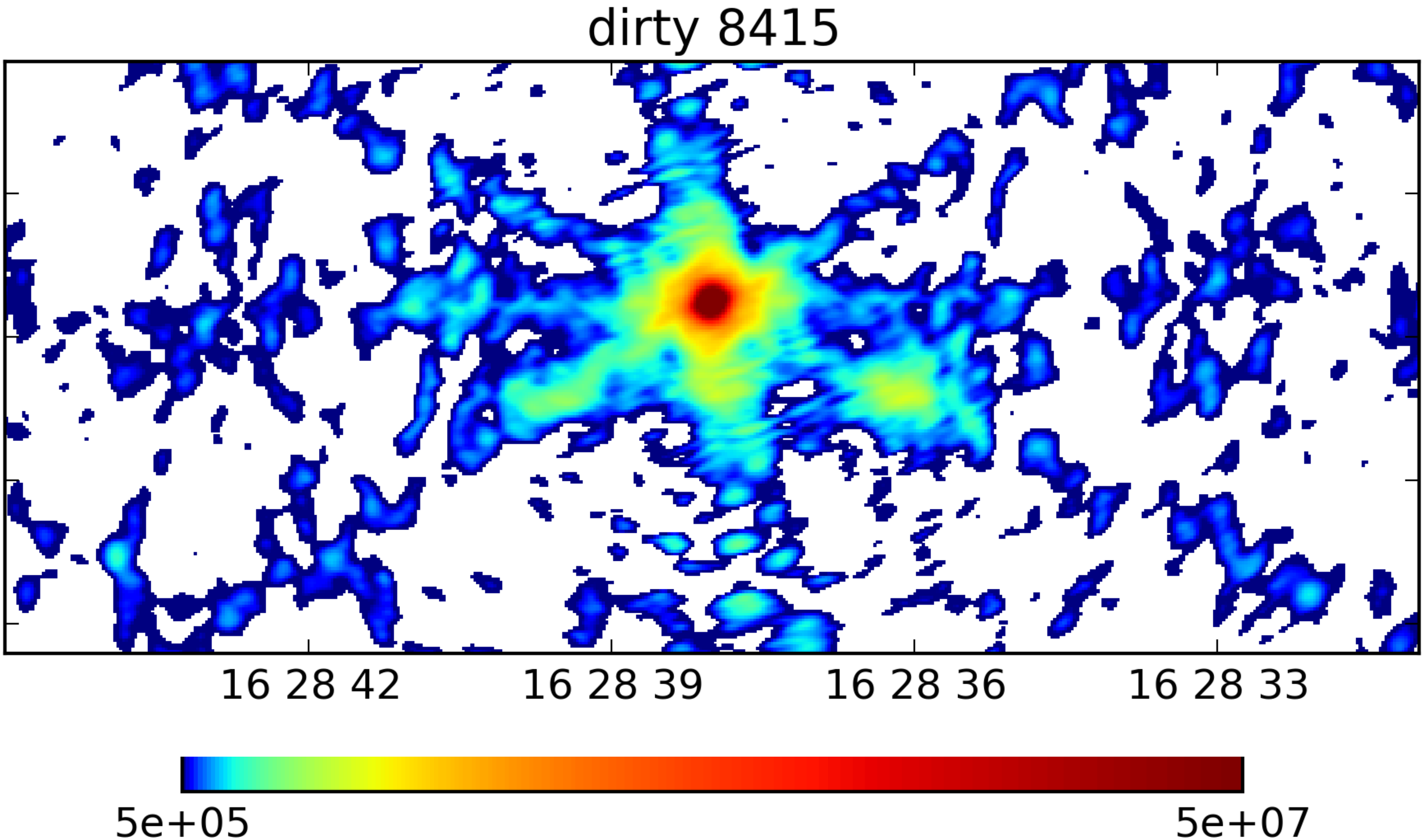
hyper-prior / intelligence

$$+ (\boldsymbol{\alpha} - \boldsymbol{1})^{\dagger}\boldsymbol{\tau} + \boldsymbol{q}^{\dagger}\mathrm{e}^{-\boldsymbol{\tau}} + \frac{1}{2}\boldsymbol{\tau}^{\dagger}\boldsymbol{T}\boldsymbol{\tau}$$

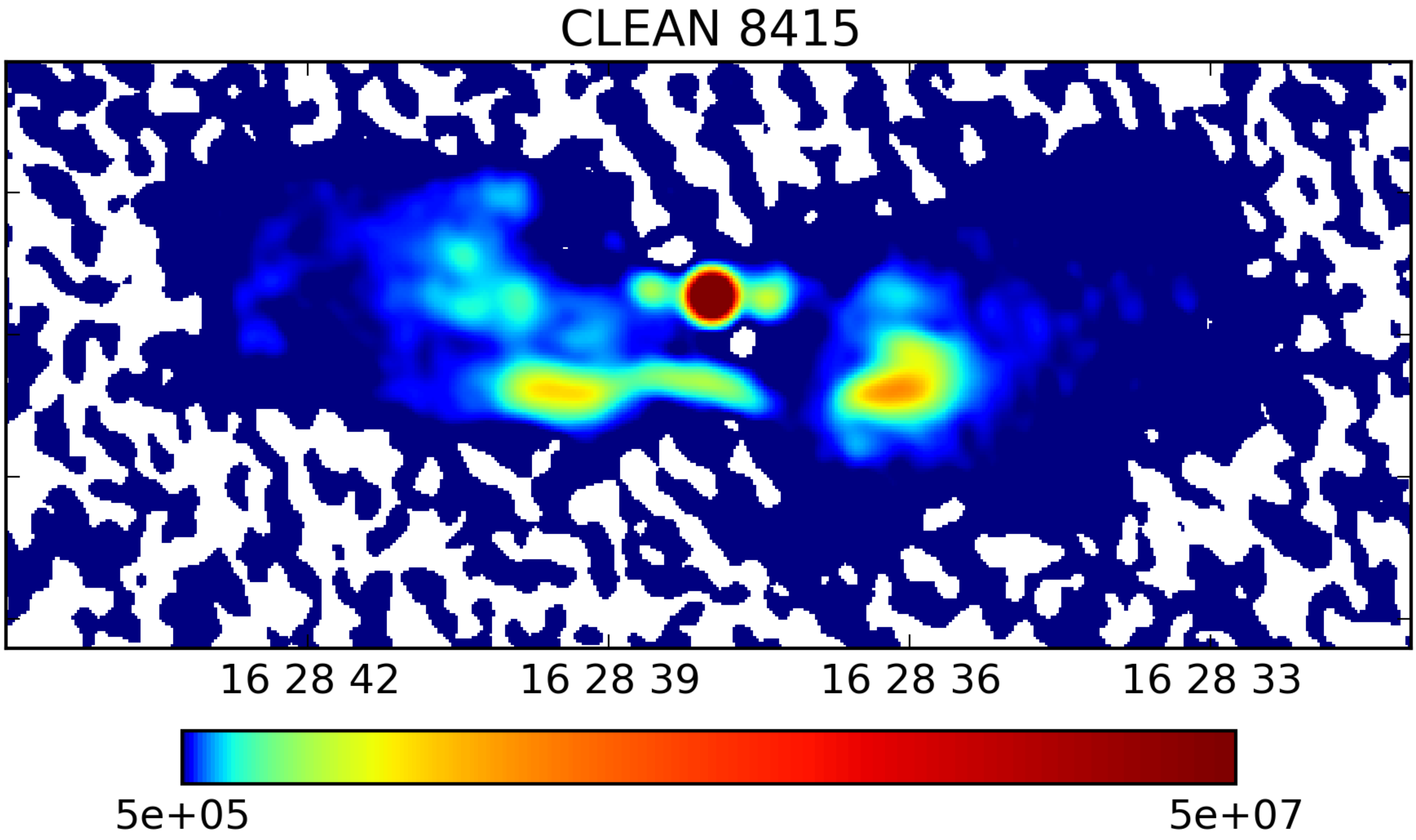
correlation structure

$$\boldsymbol{S} = \sum_{k} \mathrm{e}^{\tau_k}\boldsymbol{S}_k$$

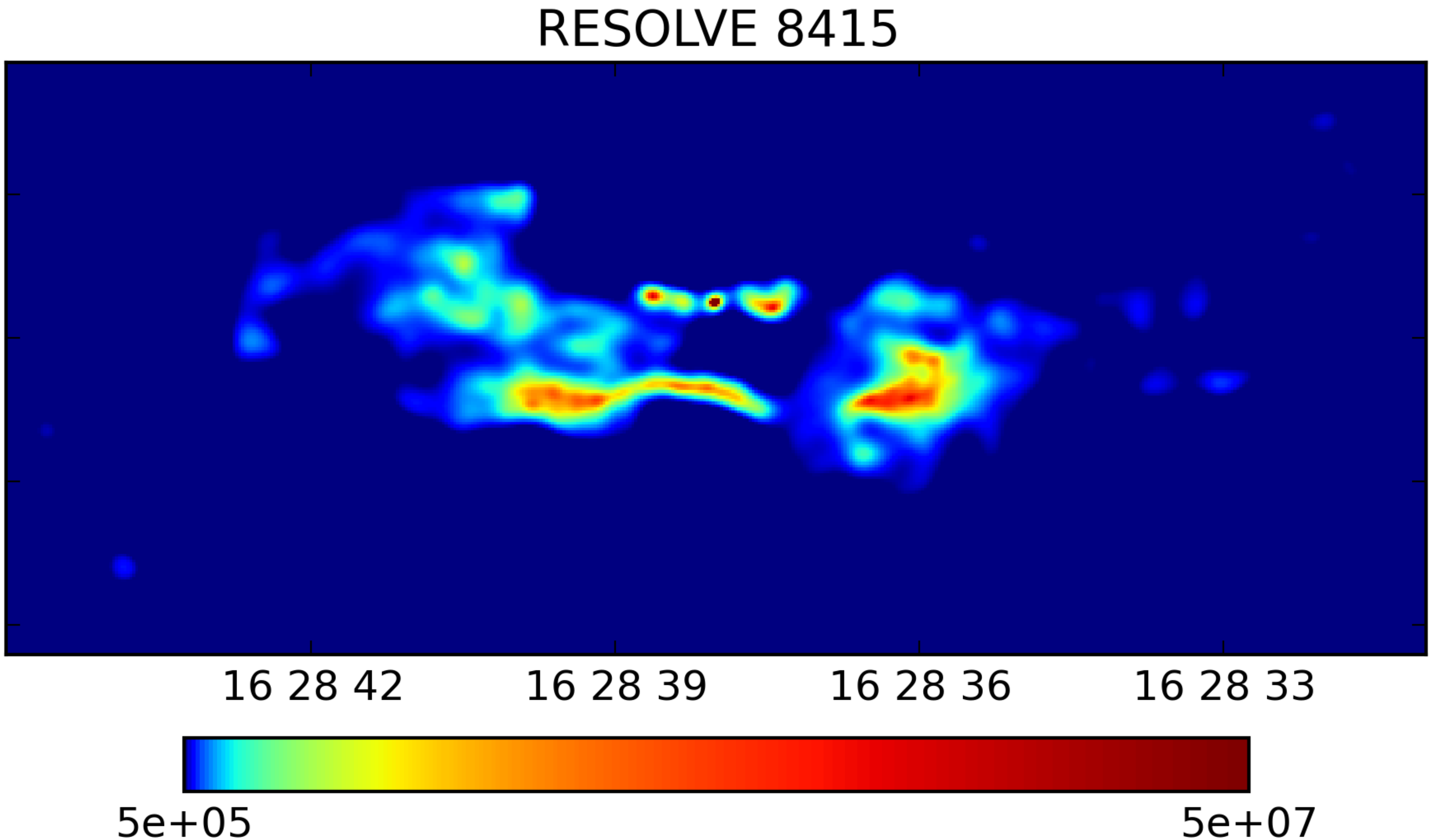
Abell 2219 @ 8415 MHz - data by Valentina Vacca



Abell 2219 @ 8415 MHz - CLEAN map by Valentina Vacca



Abell 2219 @ 8415 MHz – **fast-RESOLVE** map by Maksim Greiner



IFT gang

Maksim Greiner
fastRESOLVE

Data fusion

single dish

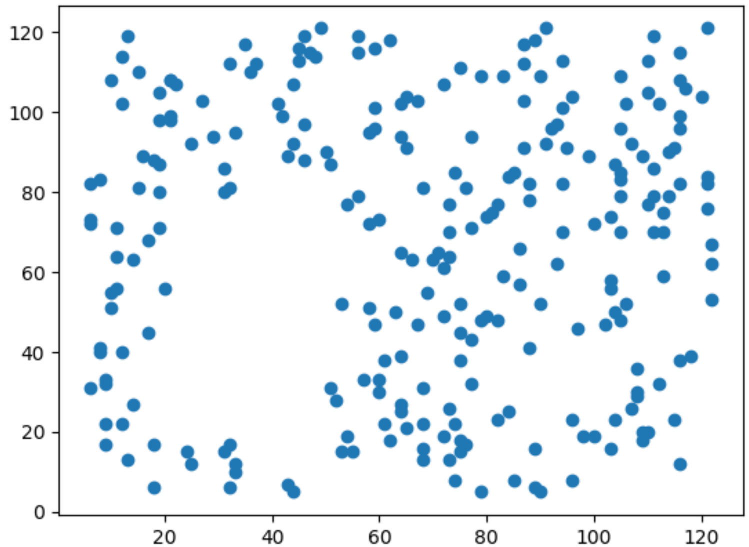
interferometer

Data fusion

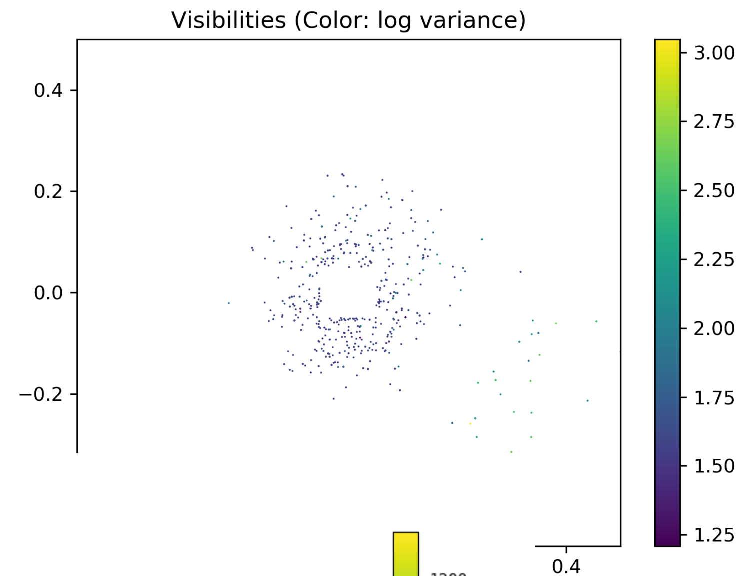
$$\mathcal{P}(d_1, d_2, s) = \mathcal{P}(d_1|s)\,\mathcal{P}(d_2|s)\,\mathcal{P}(s)\cdot\mathcal{H}(s)$$

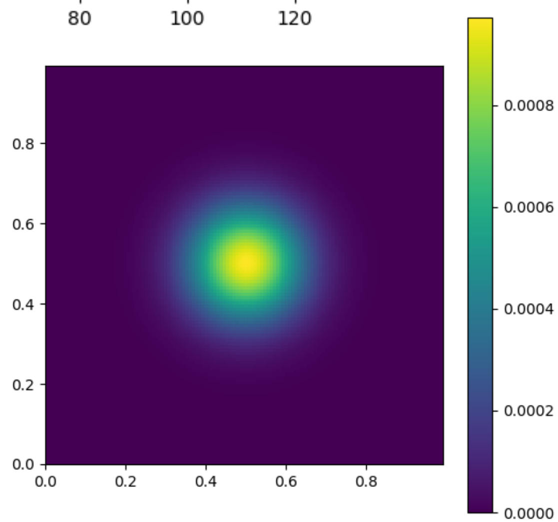
single dish

pointing

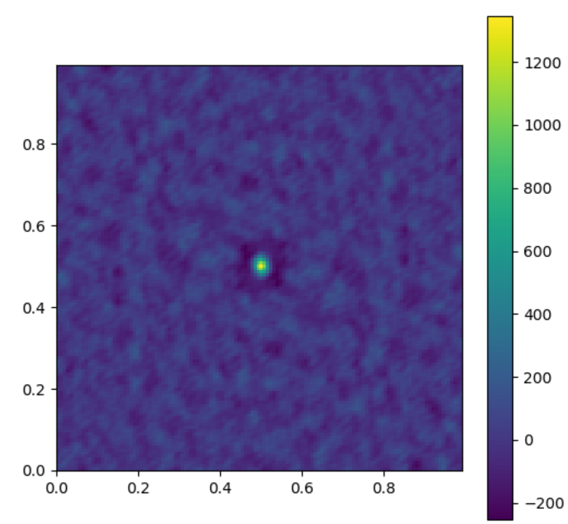


interferometer



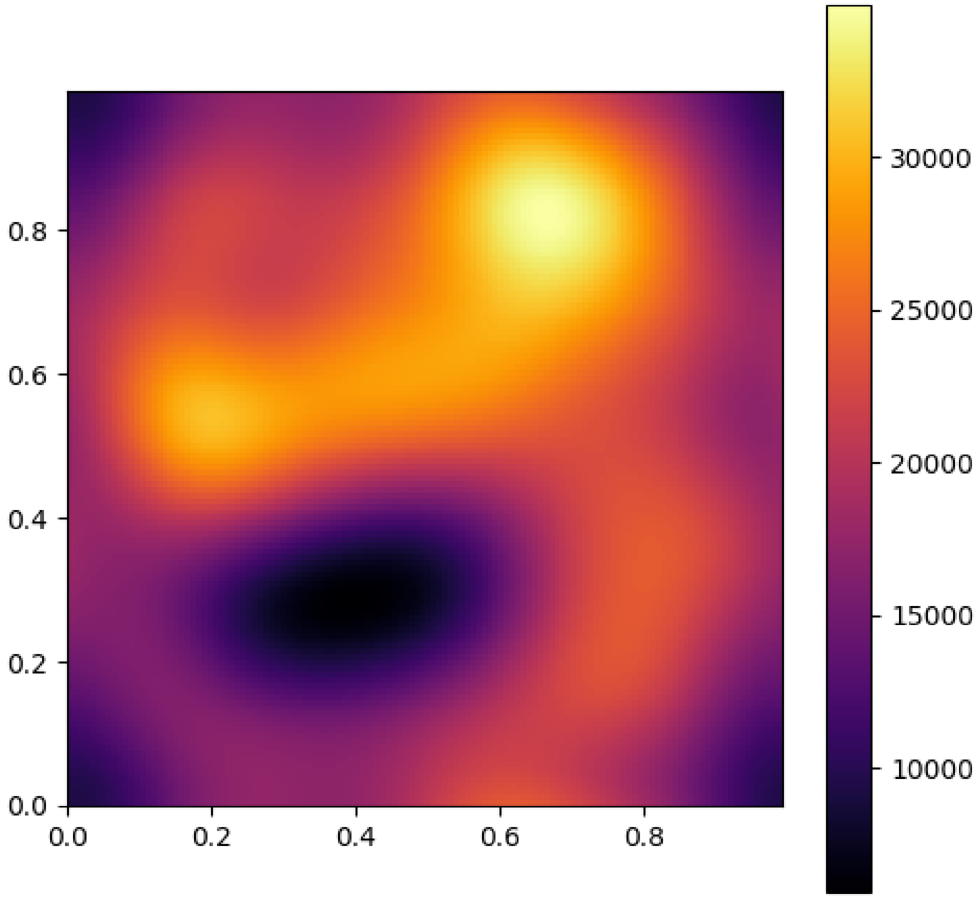


point spread function

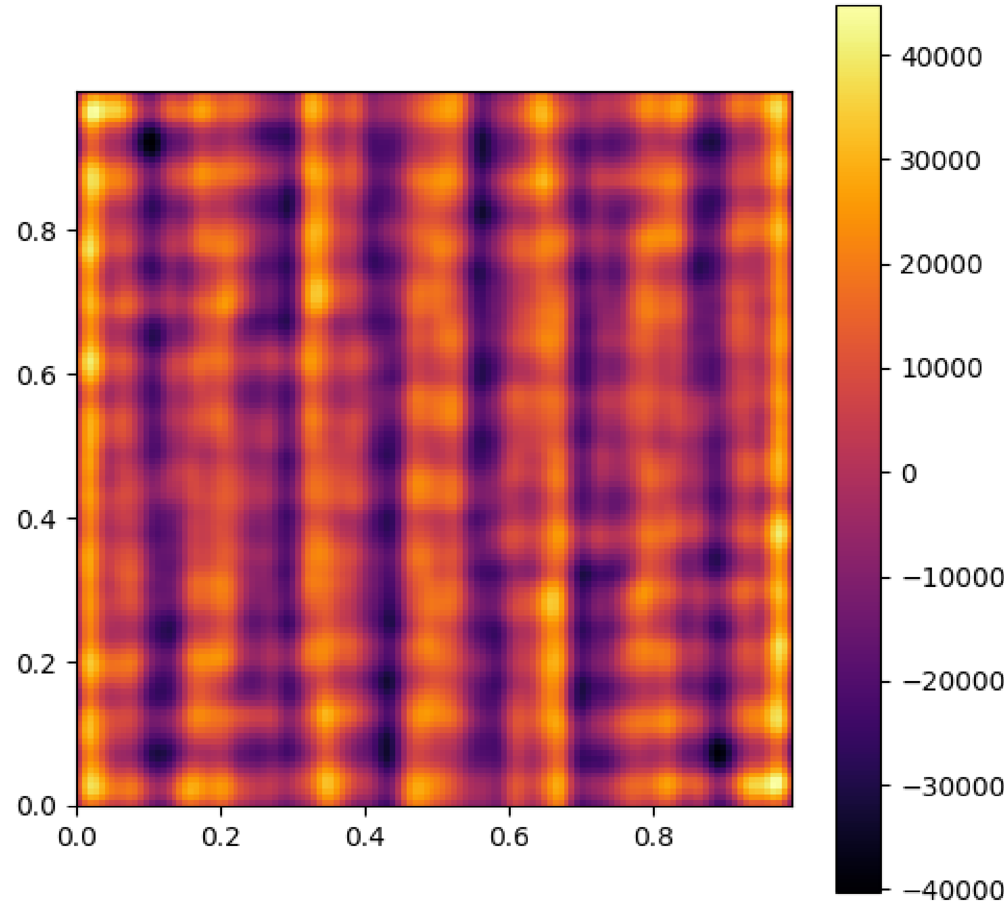


Dirty images

single dish

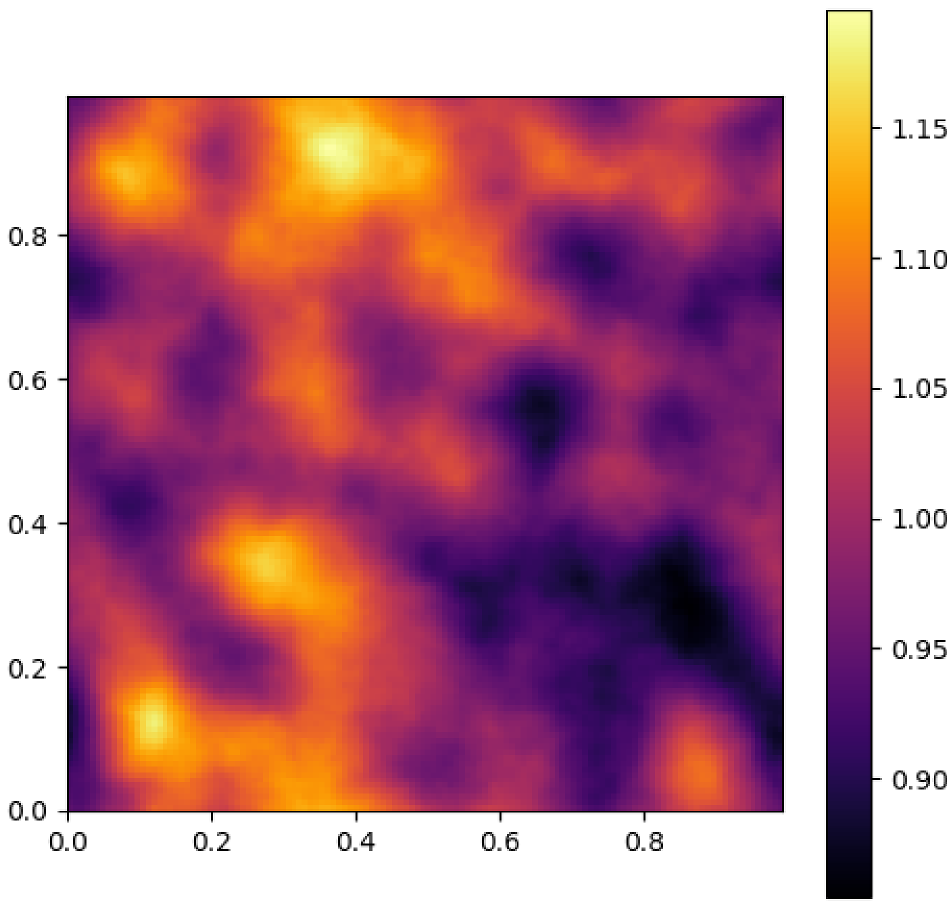


interferometer

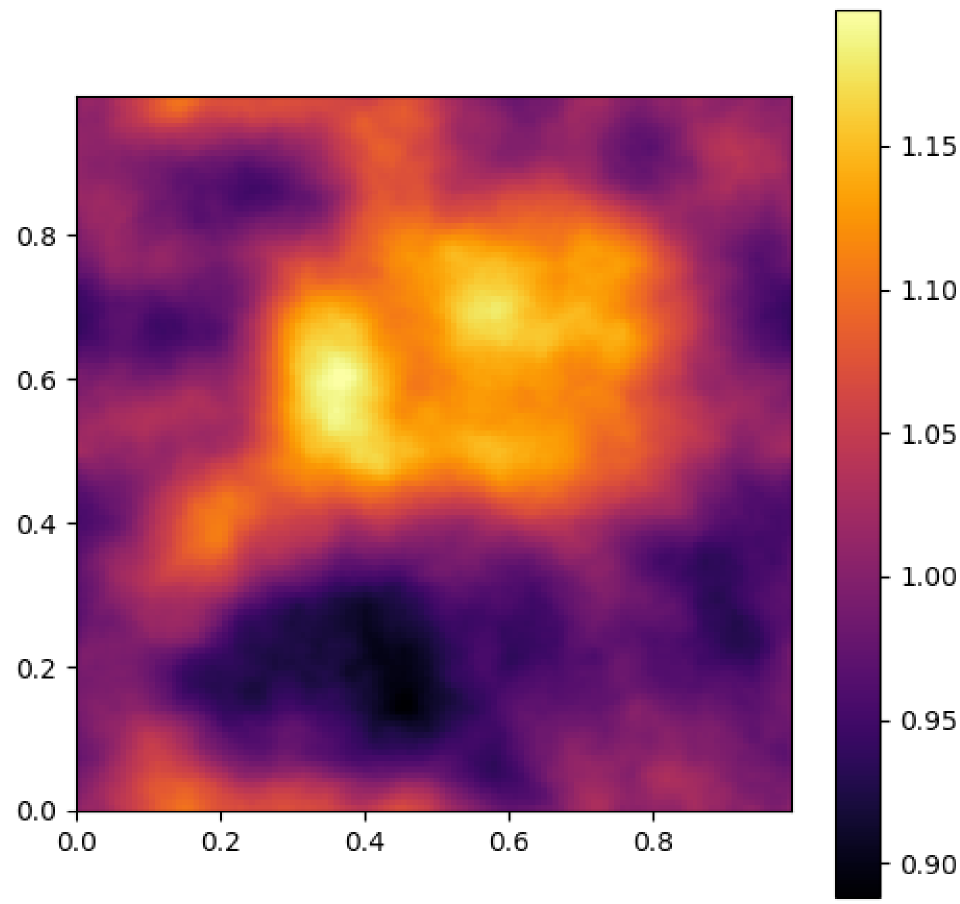


Images

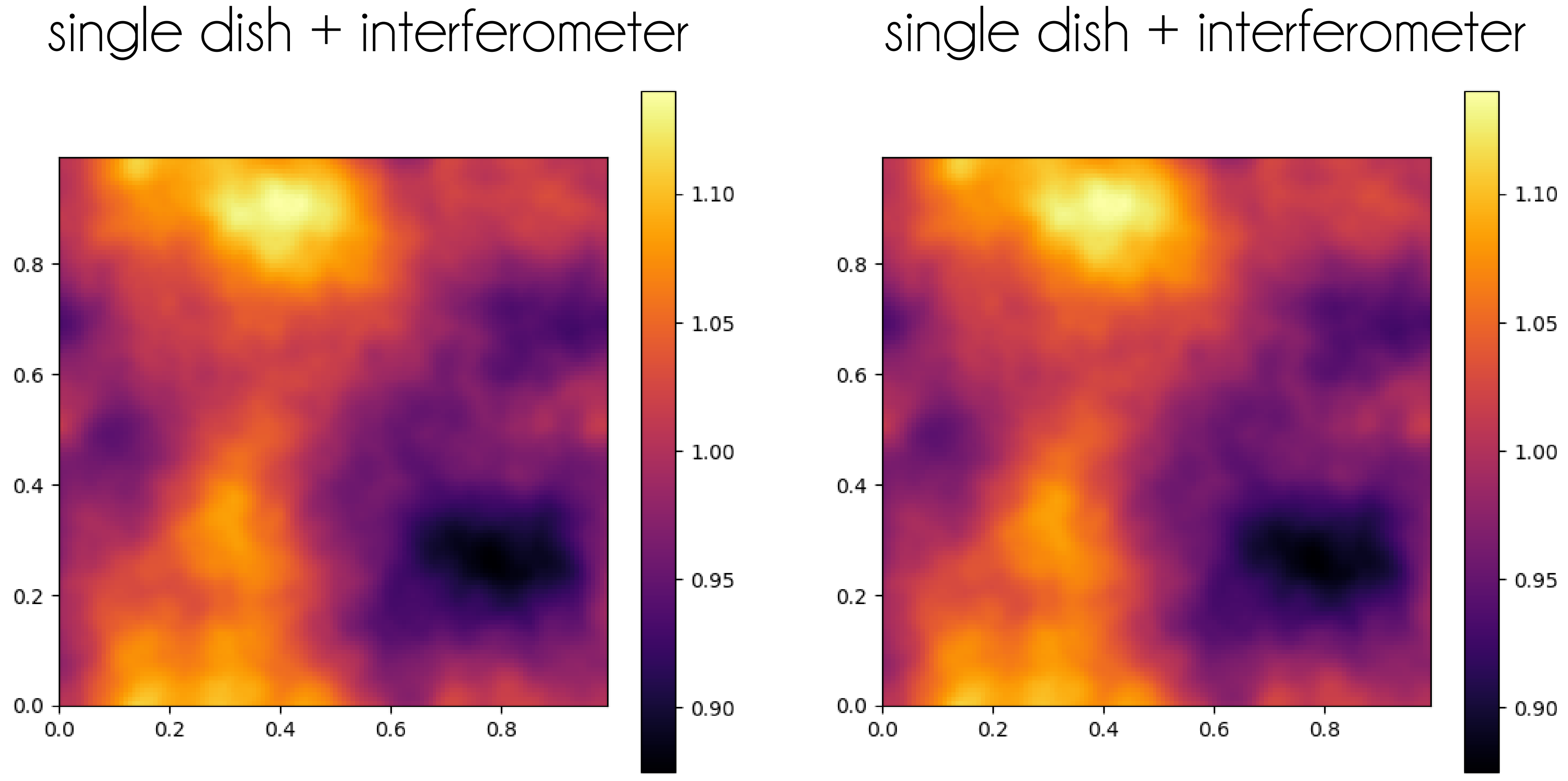
single dish



interferometer

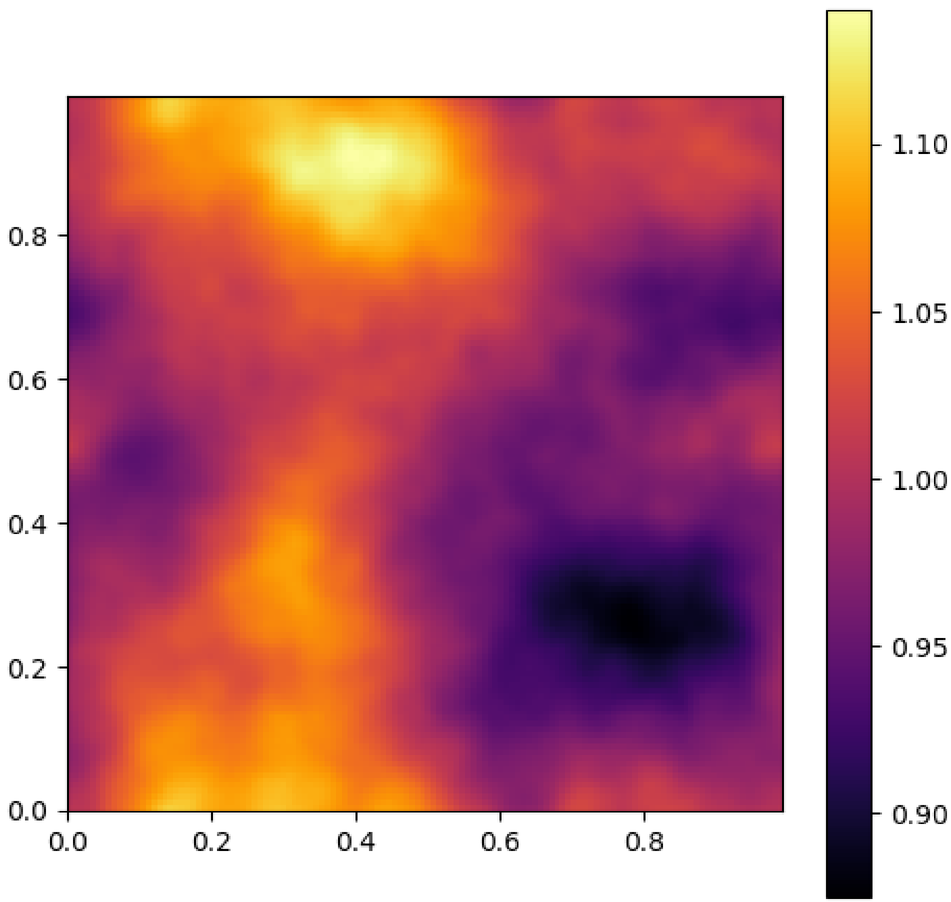


Image

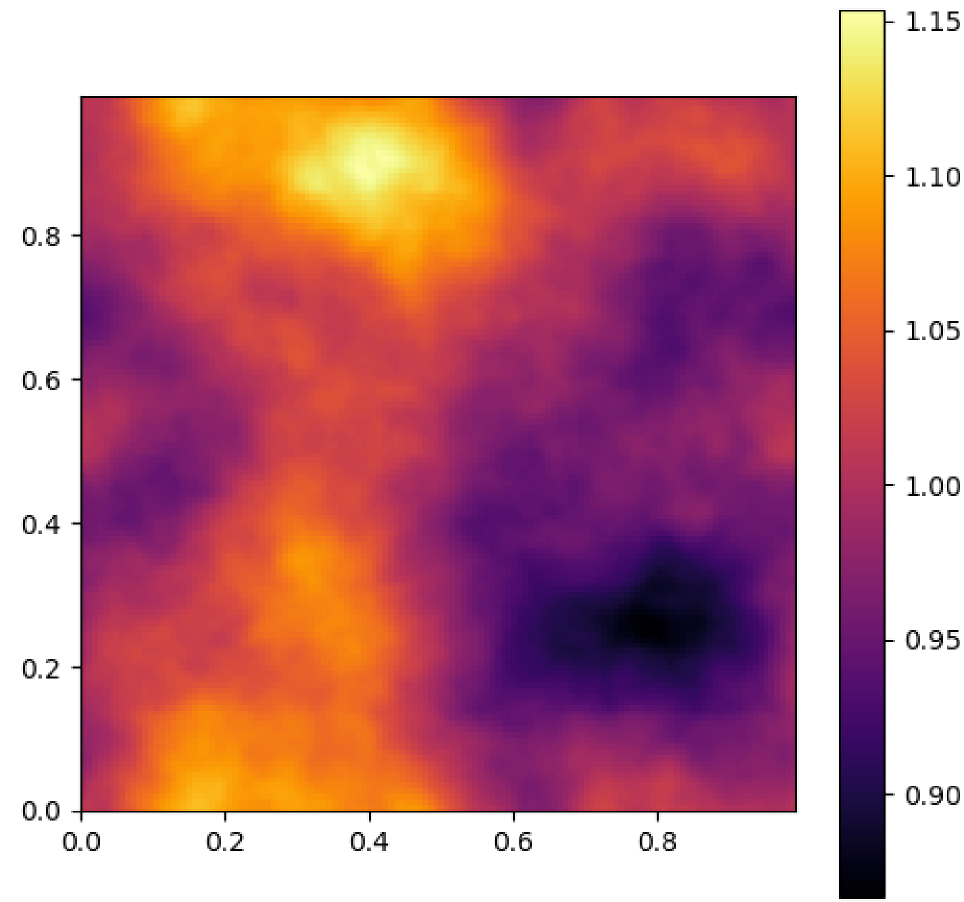

single dish + interferometer
single dish + interferometer
0.0
0.2
0.4
0.6
0.8
0.90
0.95
1.00
1.05
1.10

Image

single dish + interferometer



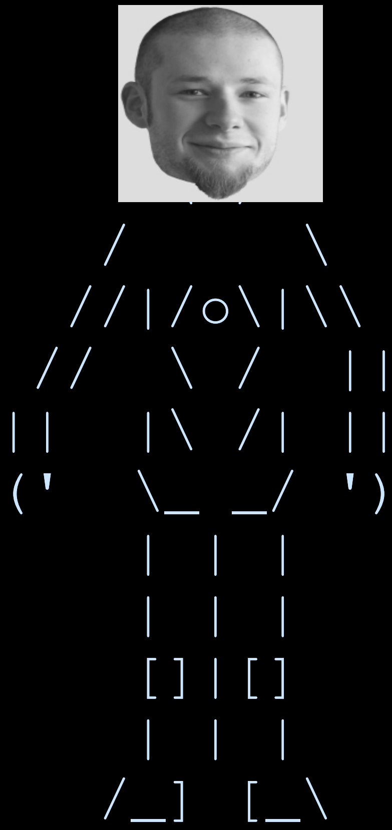
real sky



Denoising, Deconvolving, and Decomposing Photon Observations

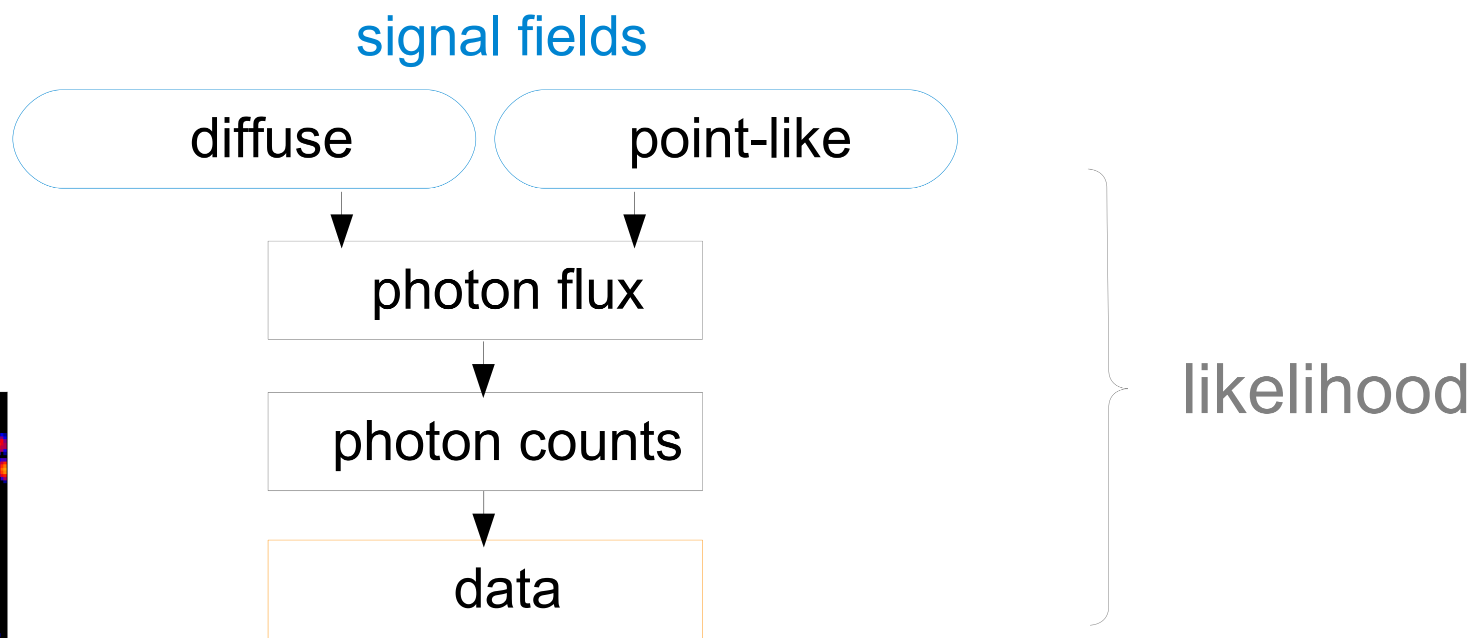
Selig et al. (2014)
www.mpa-garching.mpg.de/ift/d3po

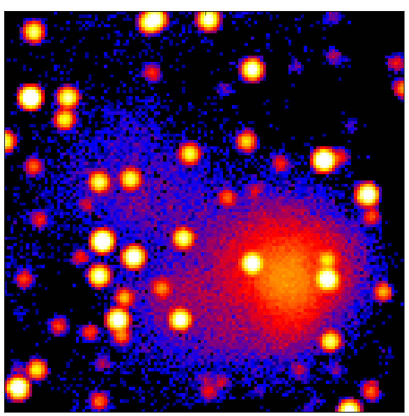
D^{3}PO



D³PO – challenges & assumptions

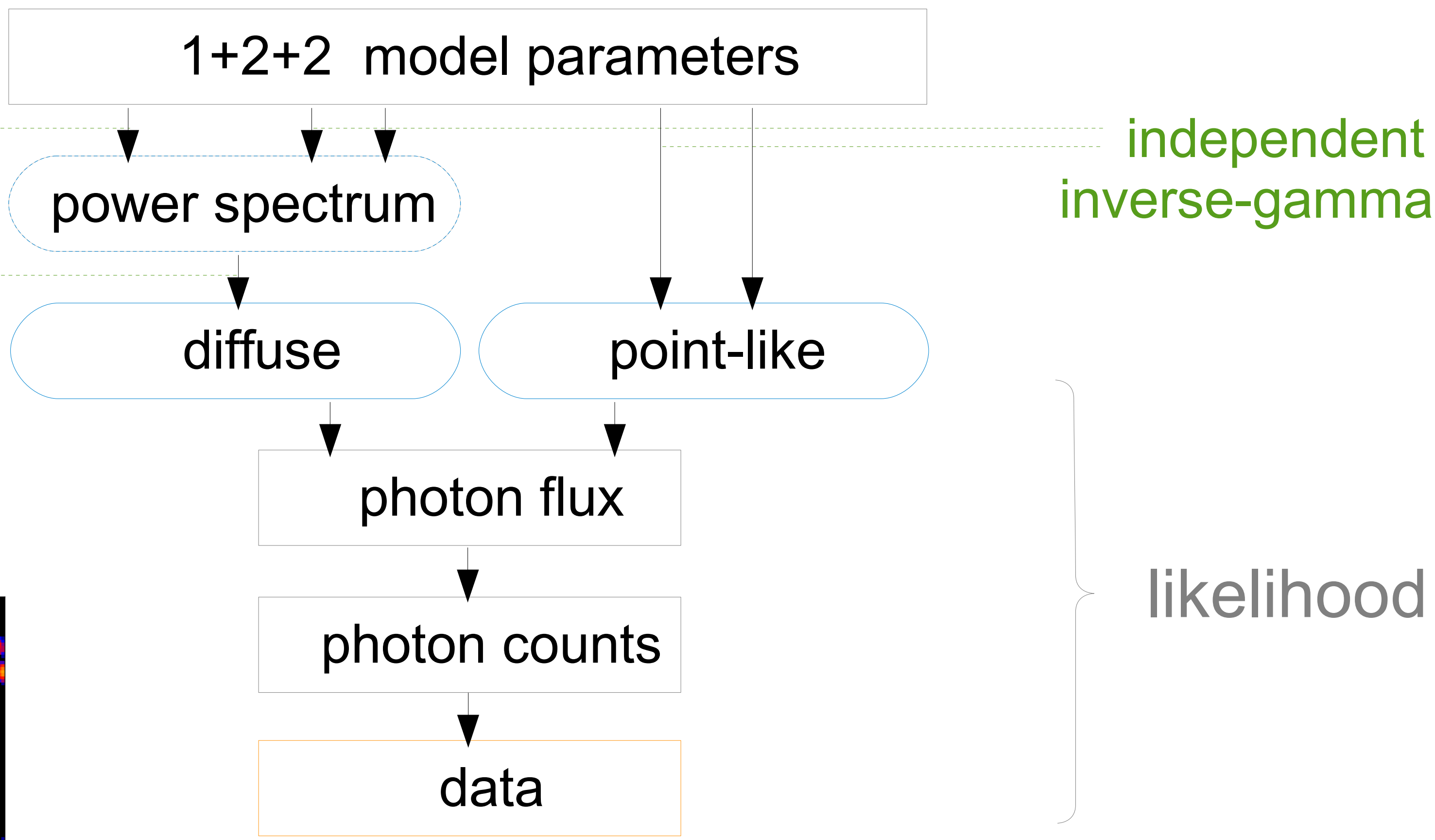
Selig & Enßlin
(2014)
arXiv: 1311.1888

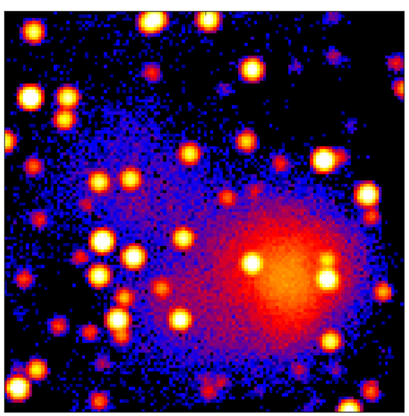




D^3PO – challenges & assumptions

Selig & Enßlin (2014)
arXiv: 1311.1888

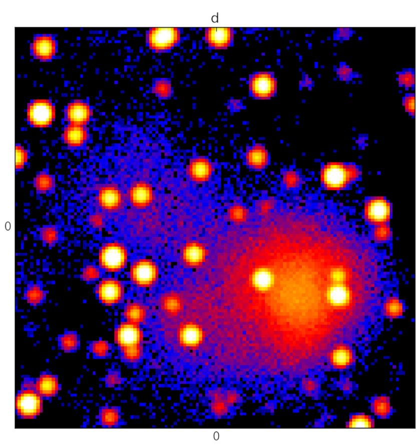




D[3]PO – challenges & assumptions

Selig & Enßlin
(2014)
arXiv: 1311.1888

$$\mathcal{H}(\boldsymbol{s},\boldsymbol{\tau},\boldsymbol{u}|\boldsymbol{d}) = -\log \mathcal{P}(\boldsymbol{s},\boldsymbol{\tau},\boldsymbol{u}|\boldsymbol{d})$$
$$= H_0 + \boldsymbol{1}^\dagger \boldsymbol{R}\left(\mathrm{e}^{\boldsymbol{s}} + \mathrm{e}^{\boldsymbol{u}}\right) - \boldsymbol{d}^\dagger \log\left(\boldsymbol{R}\left(\mathrm{e}^{\boldsymbol{s}} + \mathrm{e}^{\boldsymbol{u}}\right)\right)$$
$$+ \frac{1}{2}\log\left(\det\left[\boldsymbol{S}\right]\right) + \frac{1}{2}\boldsymbol{s}^\dagger \boldsymbol{S}^{-1}\boldsymbol{s}$$
$$+ (\boldsymbol{\alpha} - \boldsymbol{1})^\dagger \boldsymbol{\tau} + \boldsymbol{q}^\dagger \mathrm{e}^{-\boldsymbol{\tau}} + \frac{1}{2}\boldsymbol{\tau}^\dagger \boldsymbol{T}\boldsymbol{\tau}$$
$$+ (\boldsymbol{\beta} - \boldsymbol{1})^\dagger \boldsymbol{u} + \boldsymbol{\eta}^\dagger \mathrm{e}^{-\boldsymbol{u}}$$

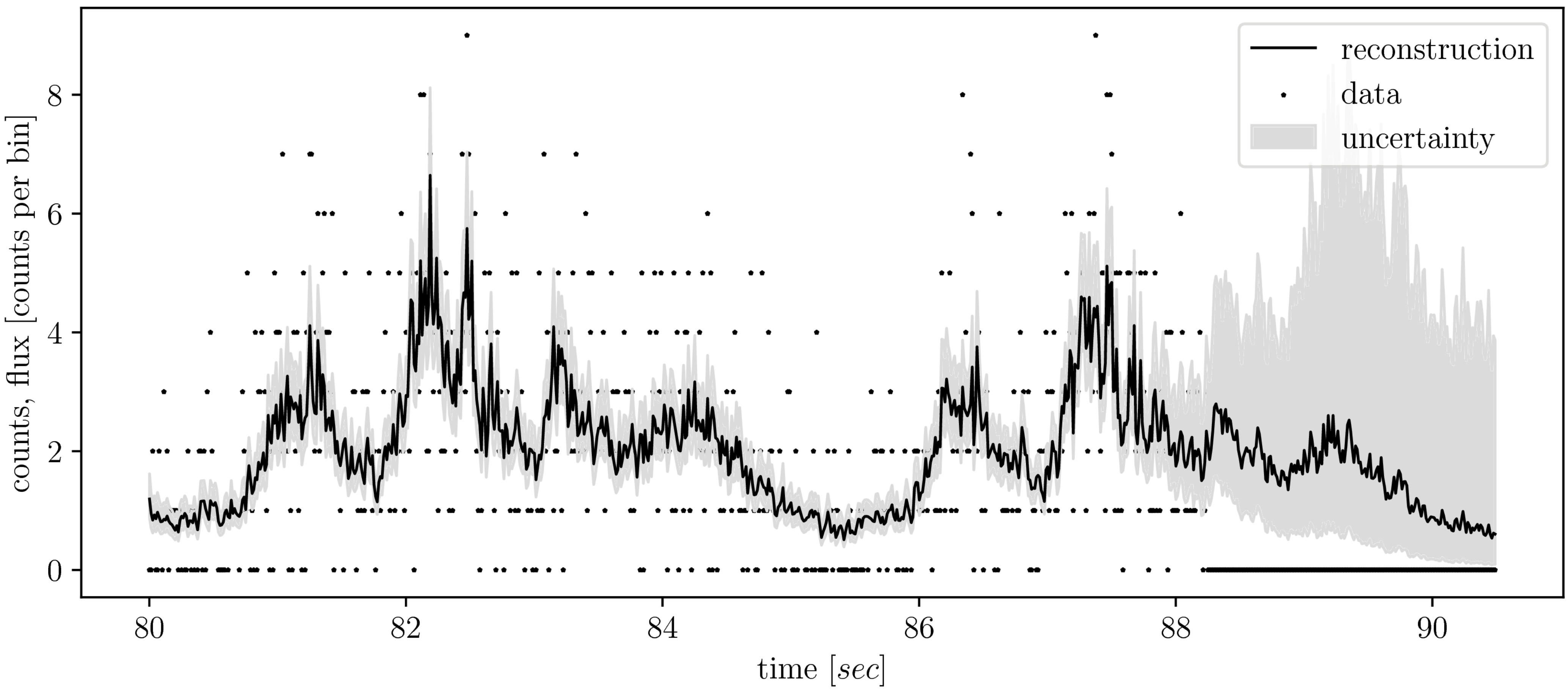


$$\boldsymbol{S} = \sum_k \mathrm{e}^{\tau_k} \boldsymbol{S}_k$$

D³PO in 1D & QPOs

Magnetar flare SGR 1900+14

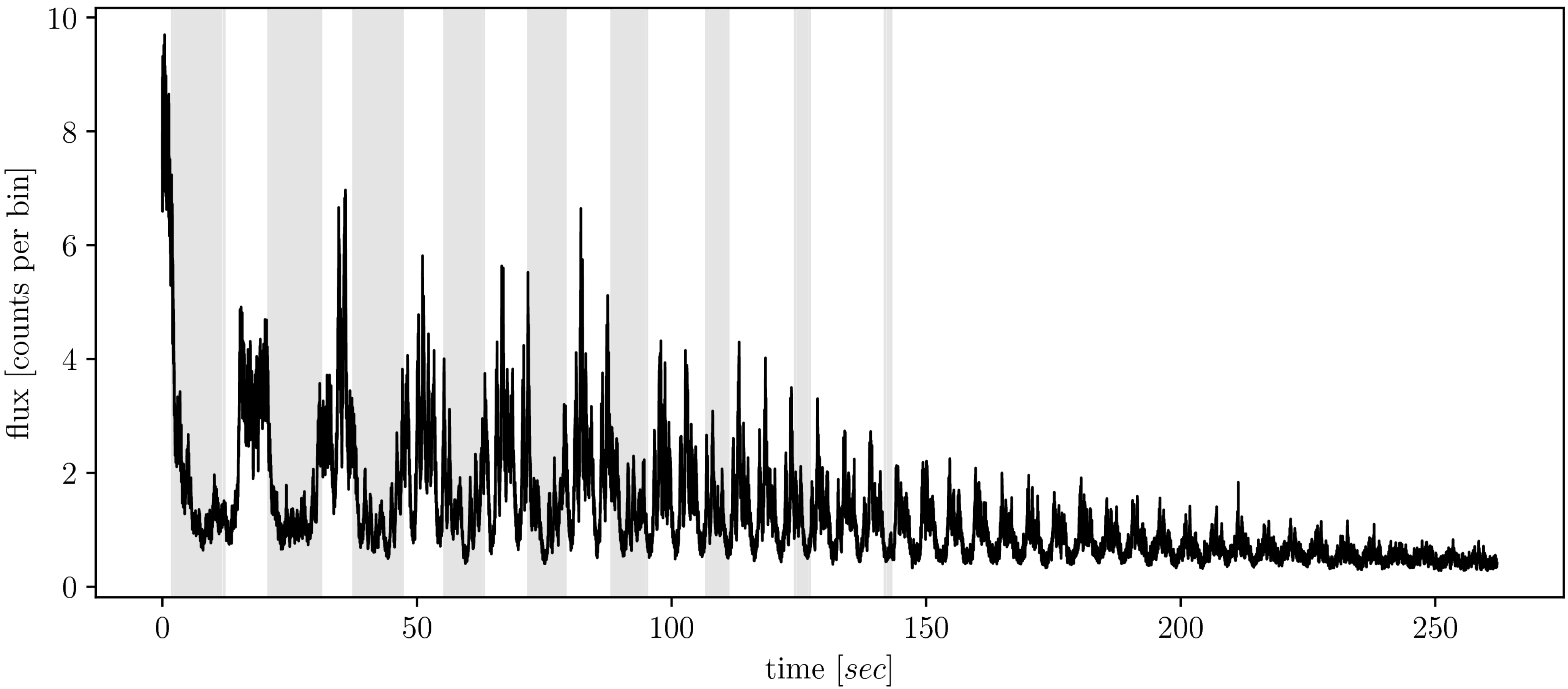
Pumpe et al. arXiv:1708.05702



D³PO in 1D & QPOs

Magnetar flare SGR 1900+14

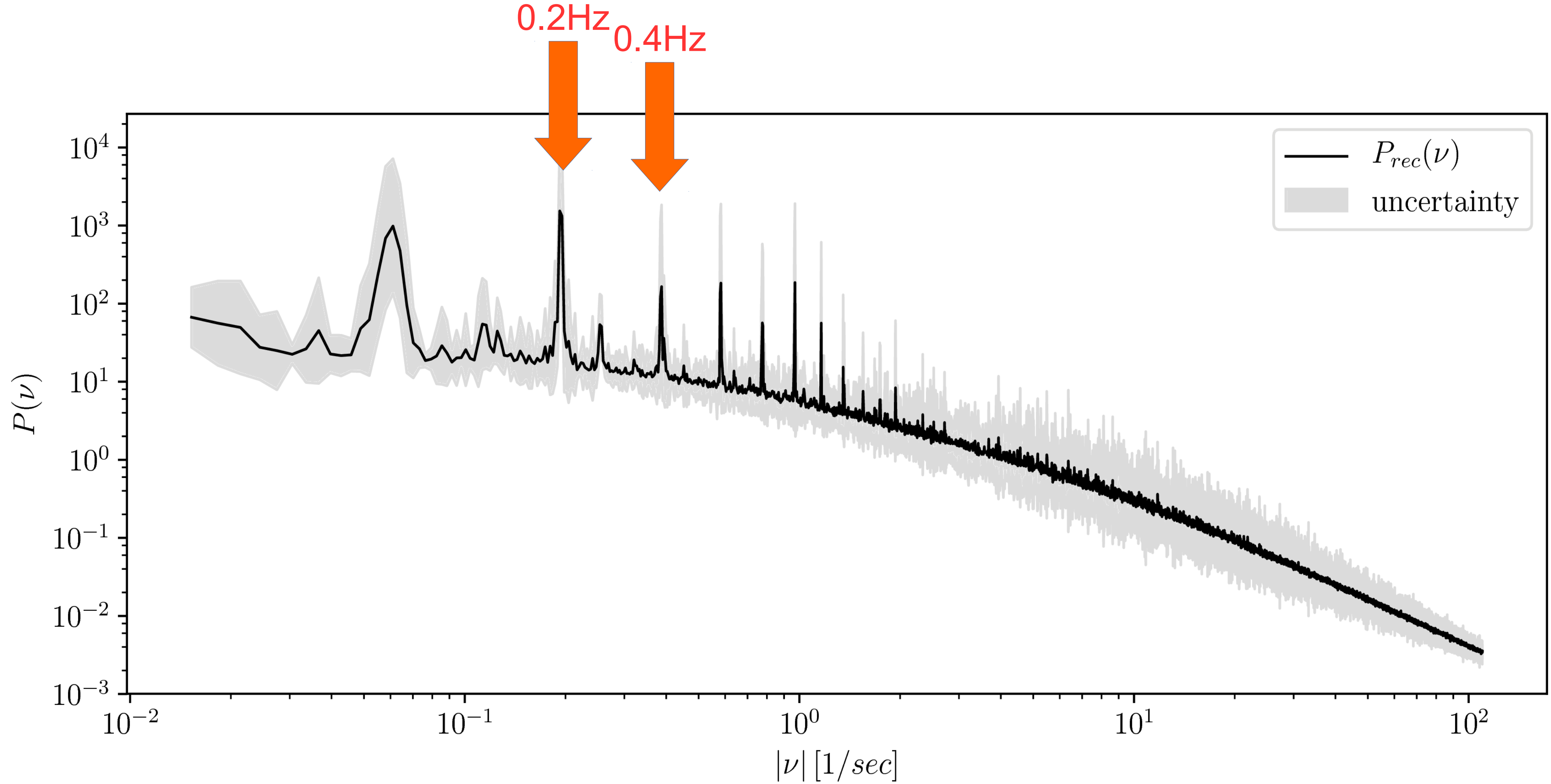
Pumpe et al. arXiv:1708.05702

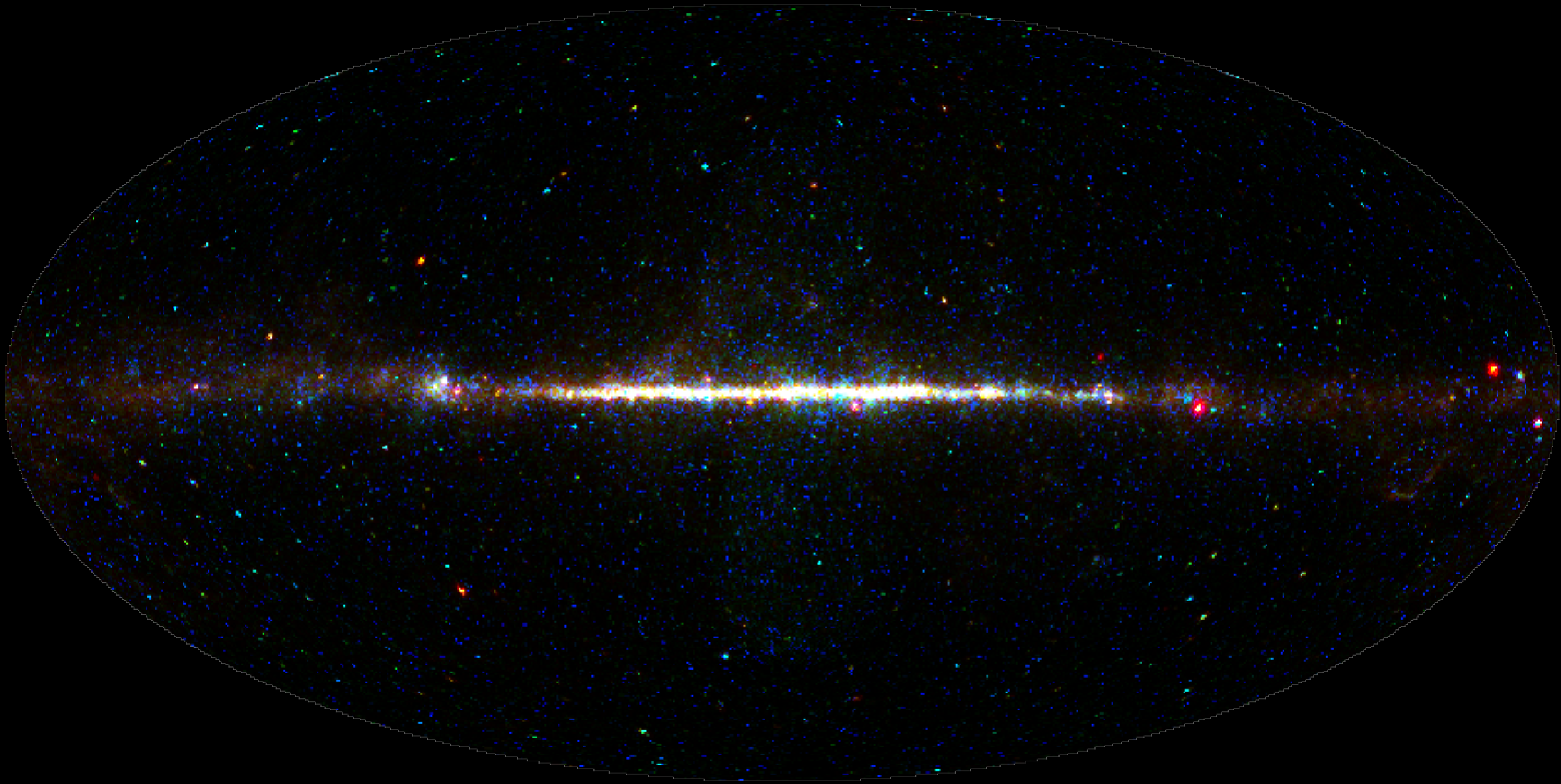


D³PO in 1D & QPOs

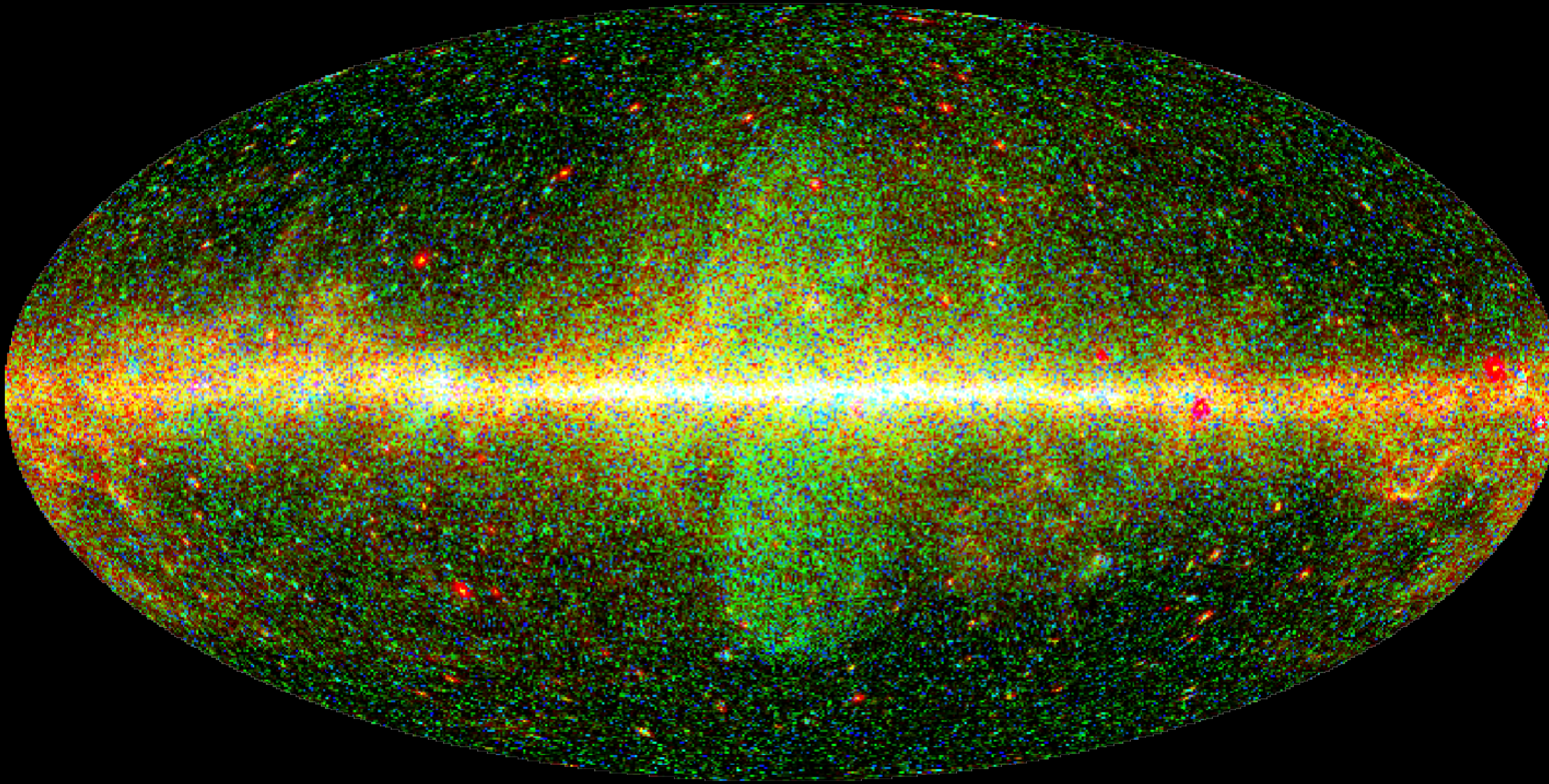
Magnetar flare SGR 1900+14

Pumpe et al. arXiv:1708.05702

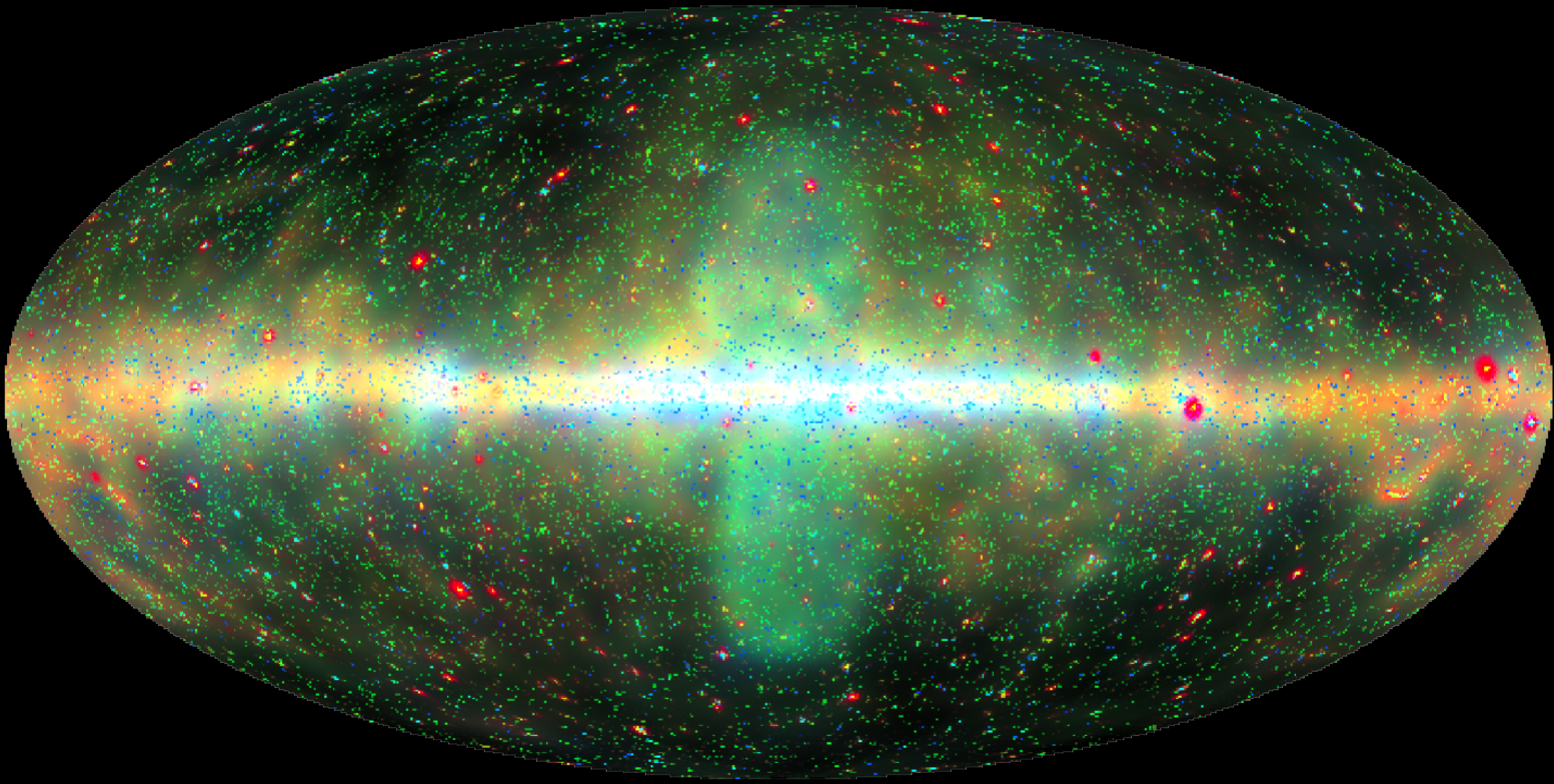




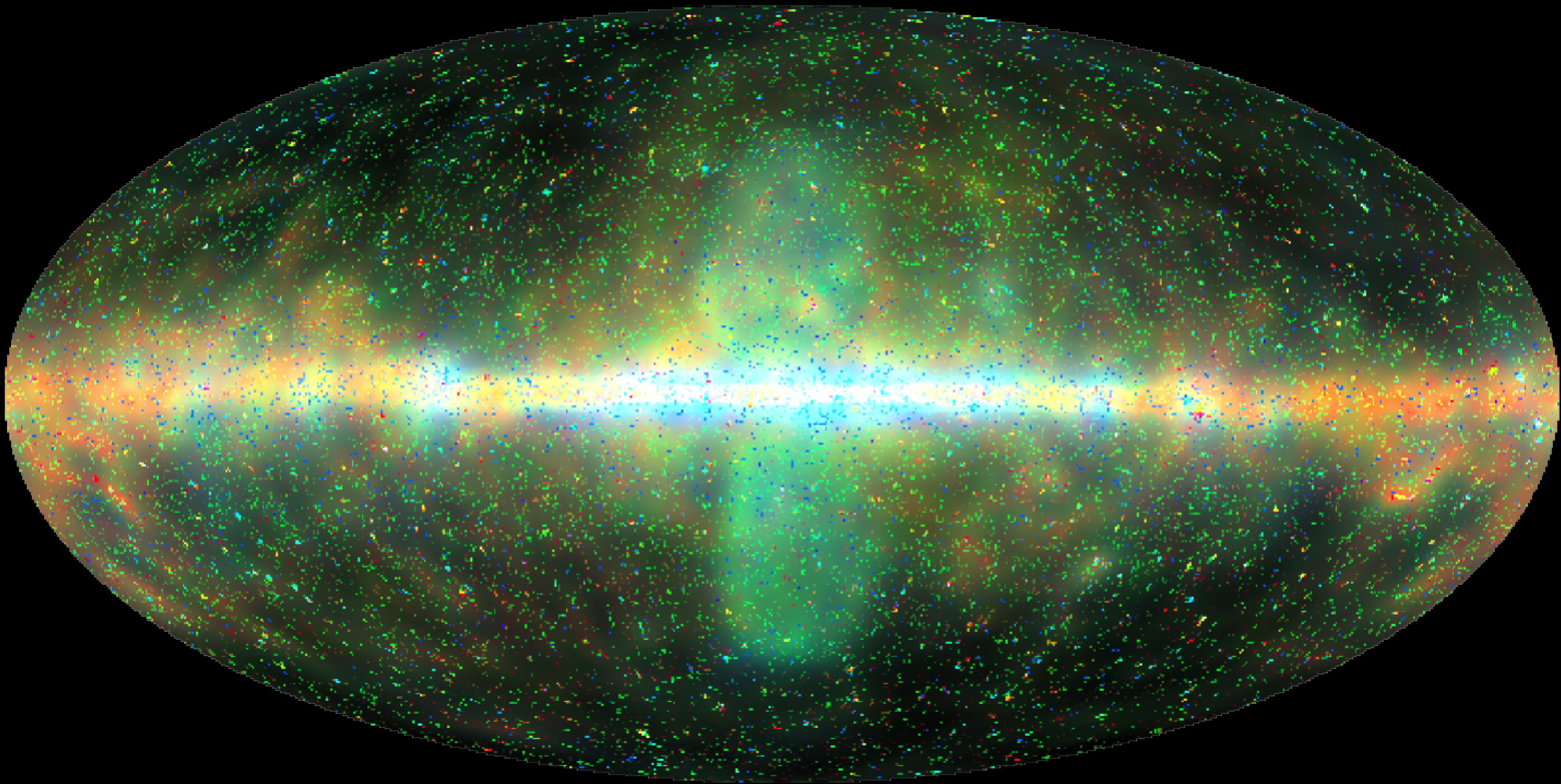
data



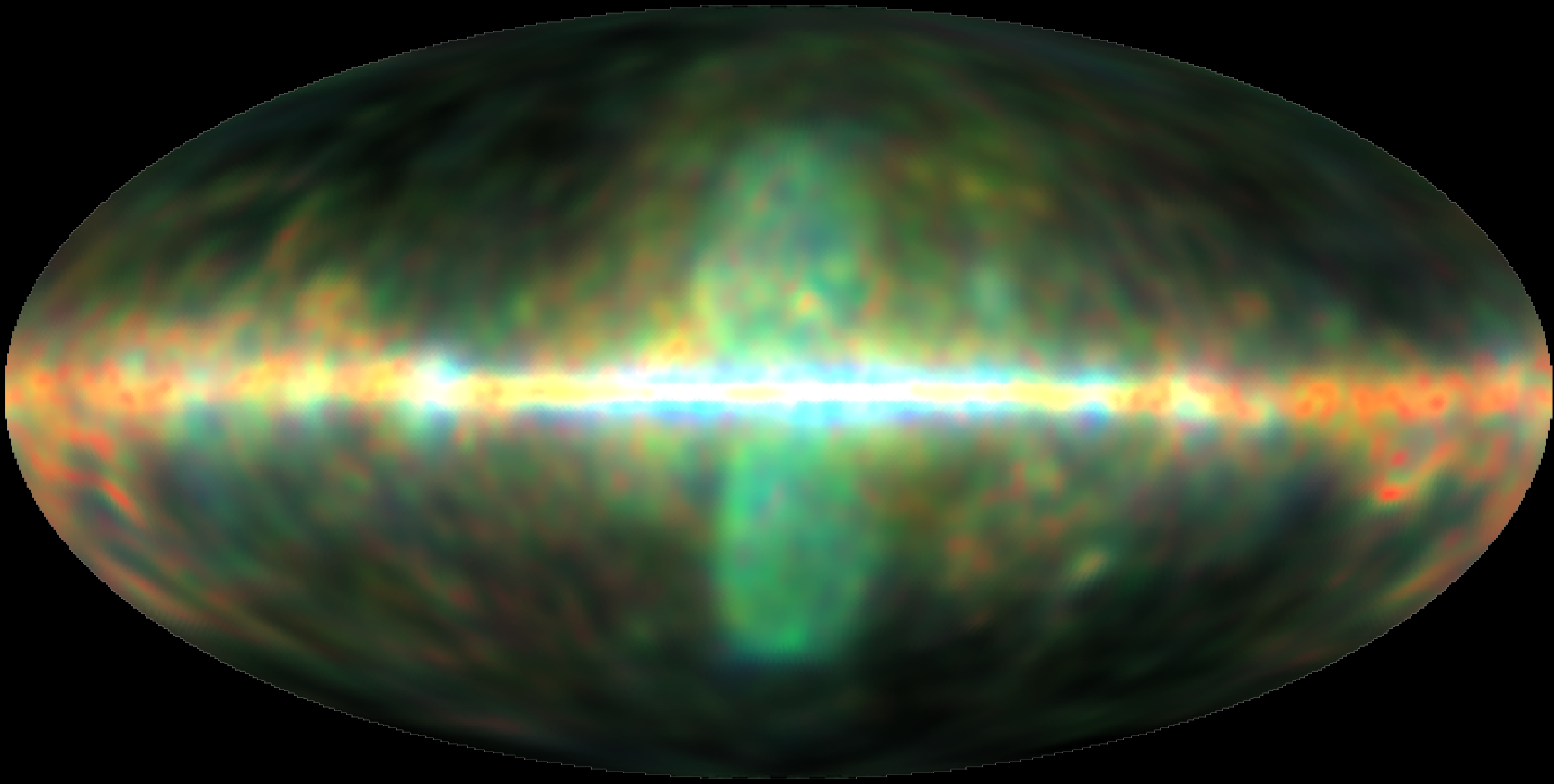
log-data



log-data … denoised

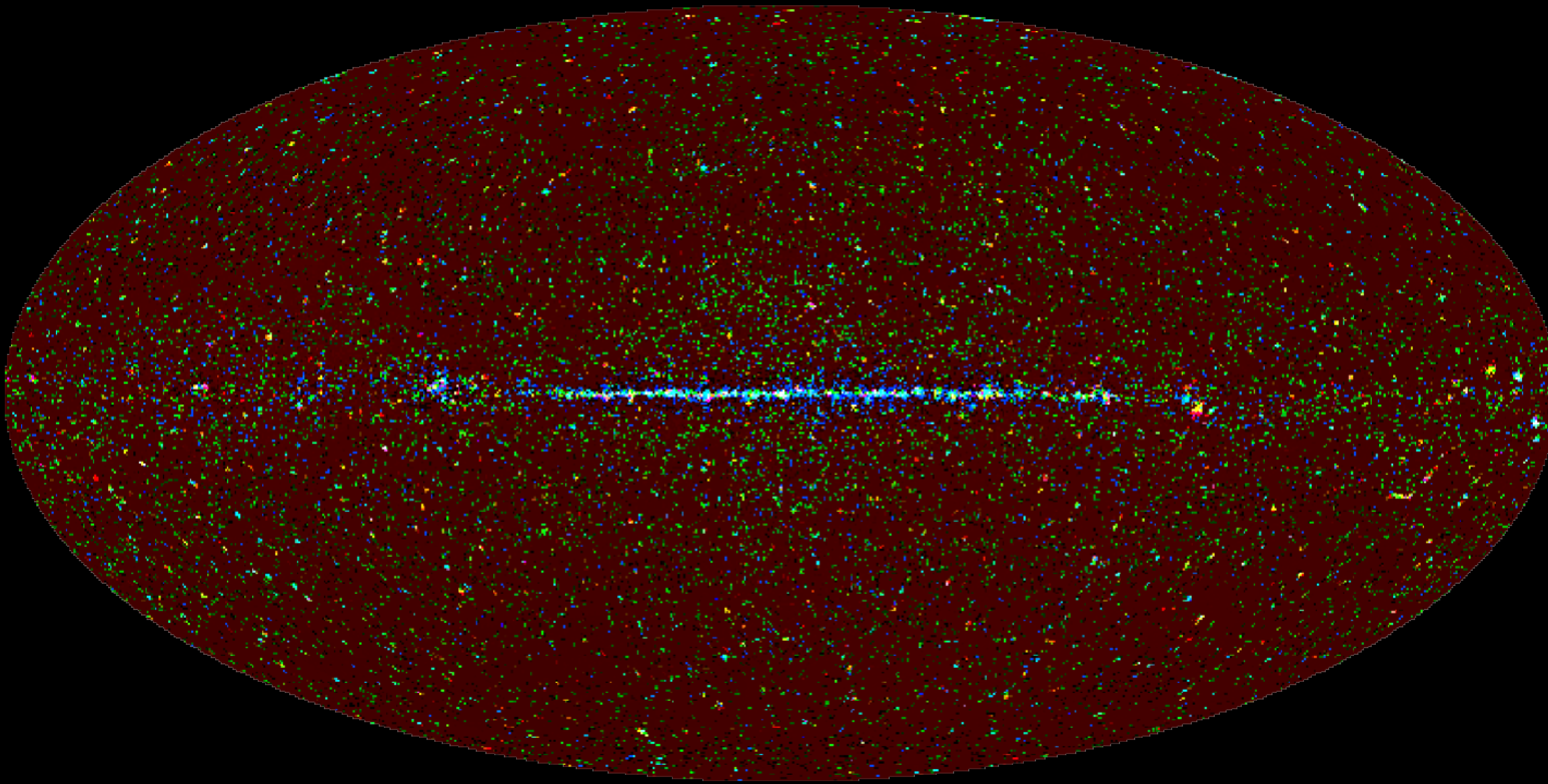


log-data … denoised … deconvolved



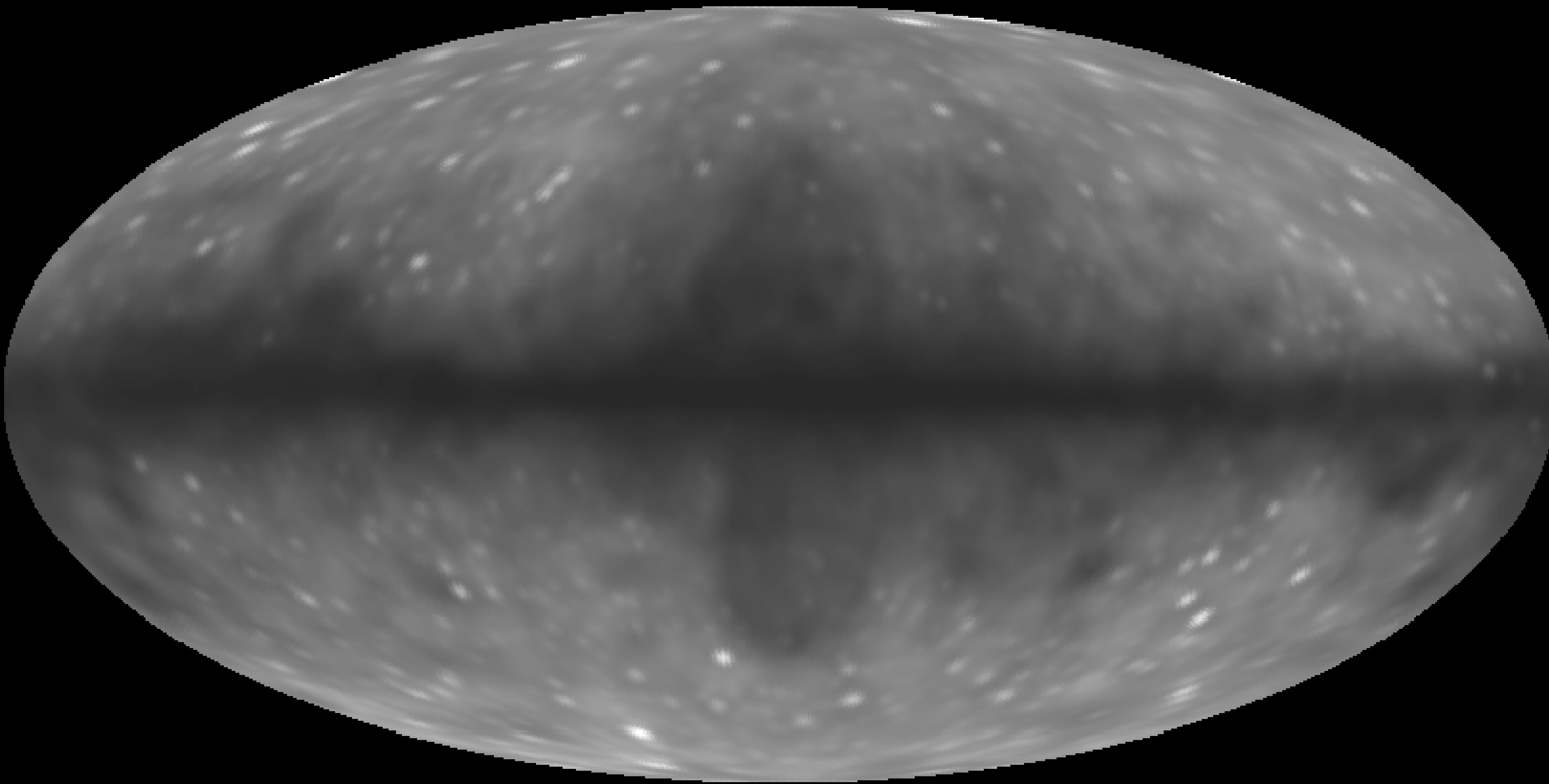
log-data … denoised … deconvolved … decomposed

Selig, Vacca, Oppermann, Enßlin (2015)



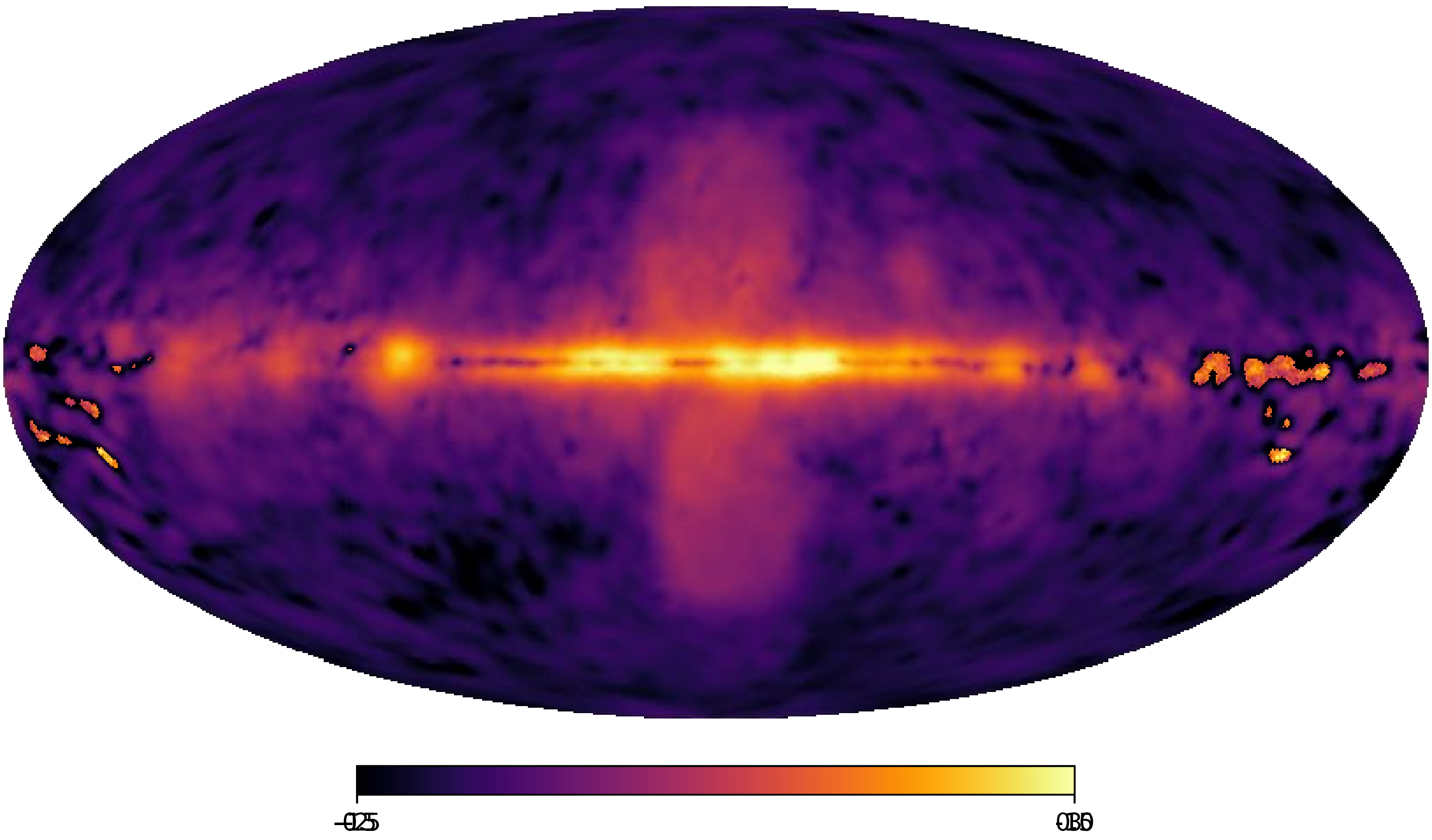
log-data … denoised … deconvolved … decomposed

Selig, Vacca, Oppermann, Enßlin (2015)

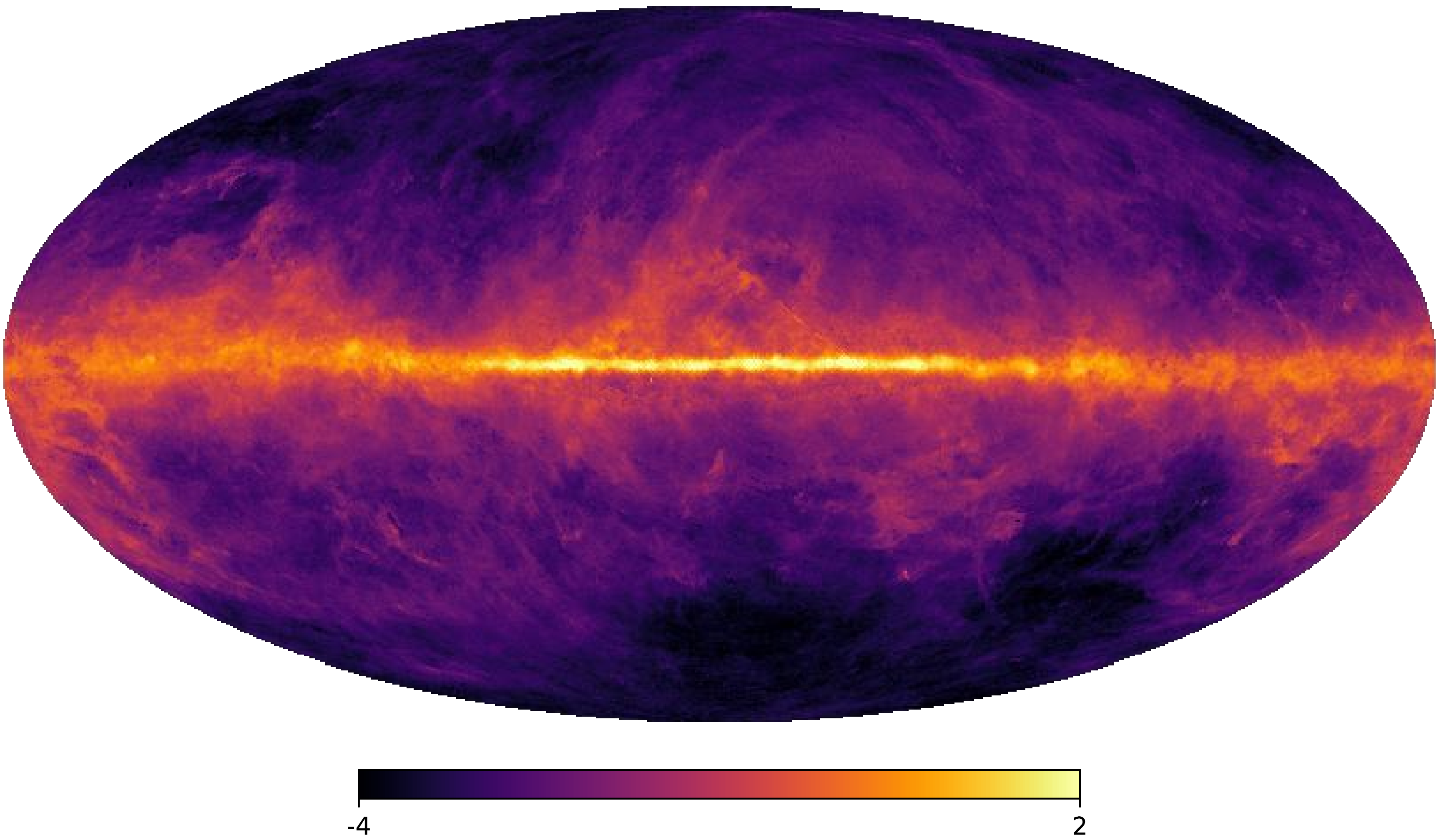


relative uncertainty of diffuse emission

Sharpening up the gamma-ray sky



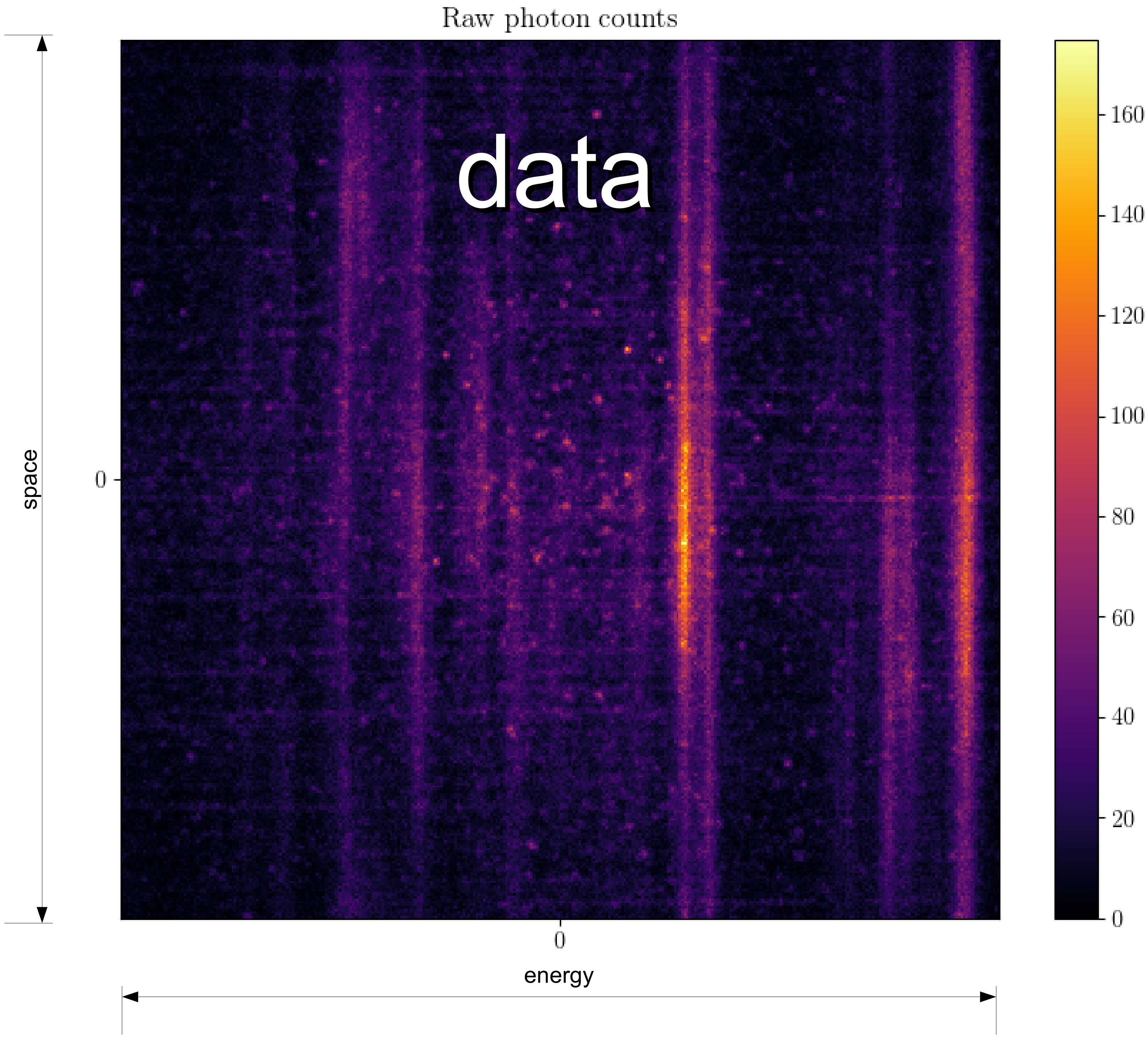
Restoring the hadronic sky



Glimpses into the future

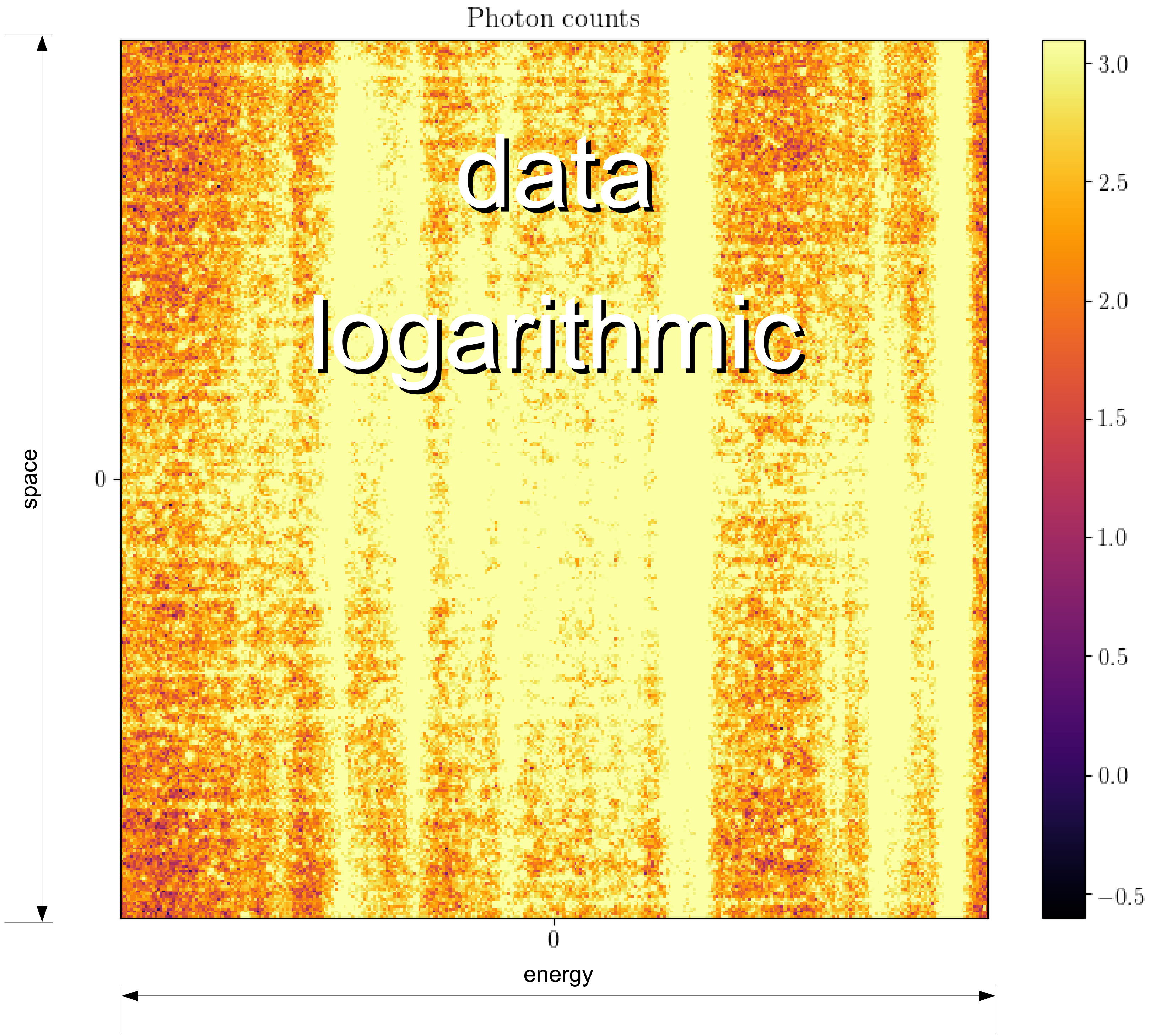
Daniel Pumpe: D^4PO

arXiv:1802.020135

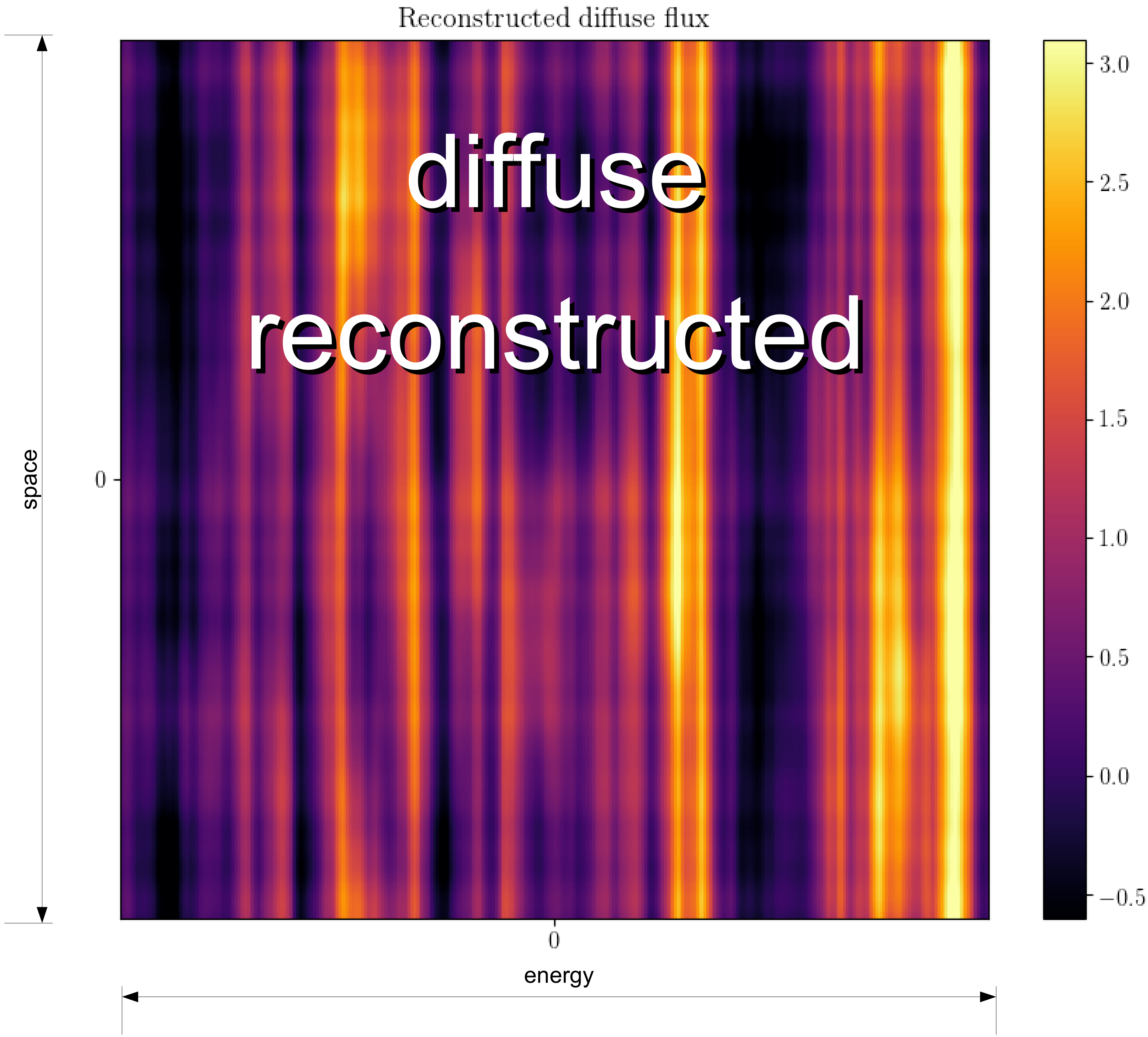


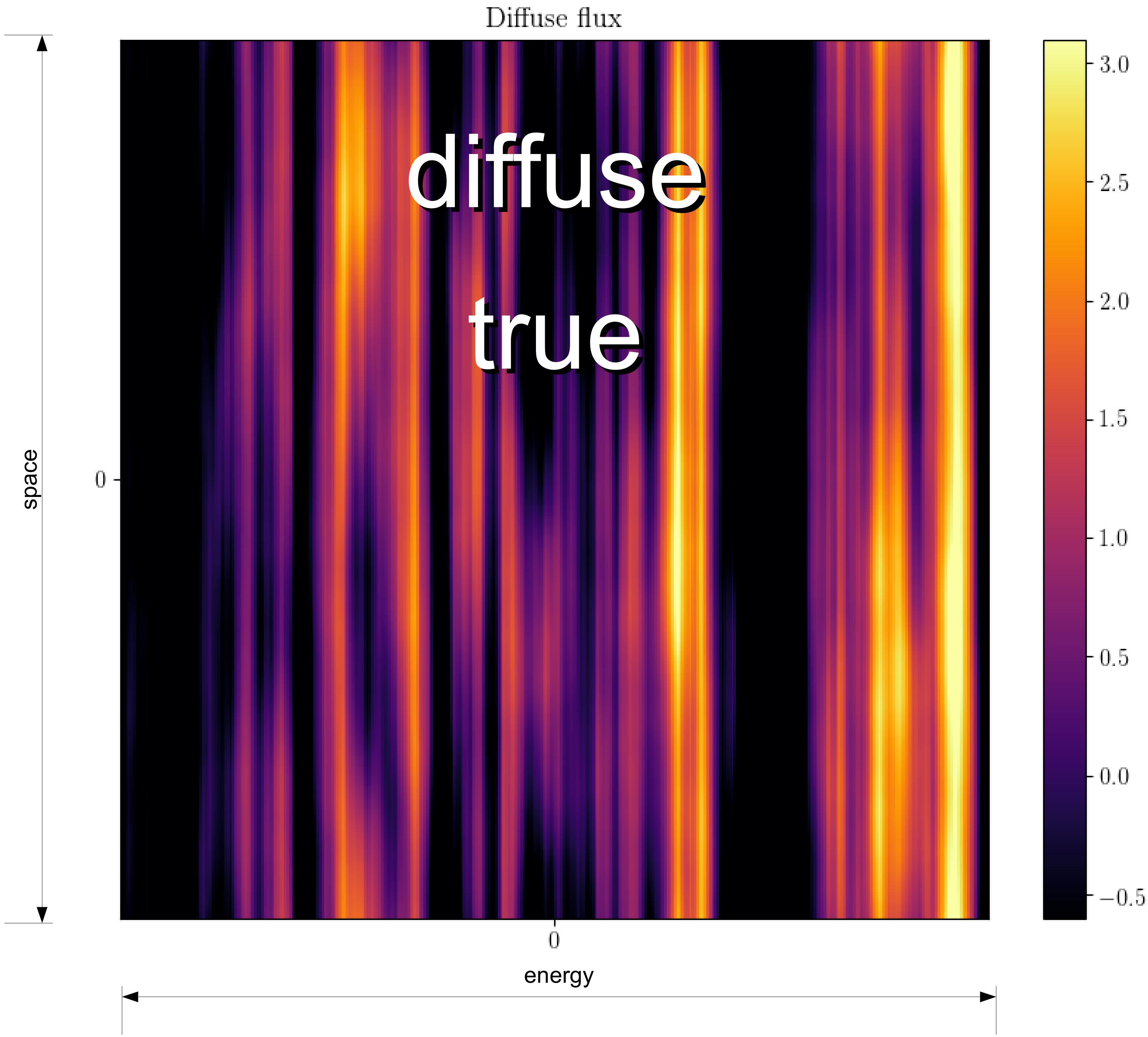
Raw photon counts
data
space
0
0
energy
160
140
120
100
80
60
40
20
0

2018

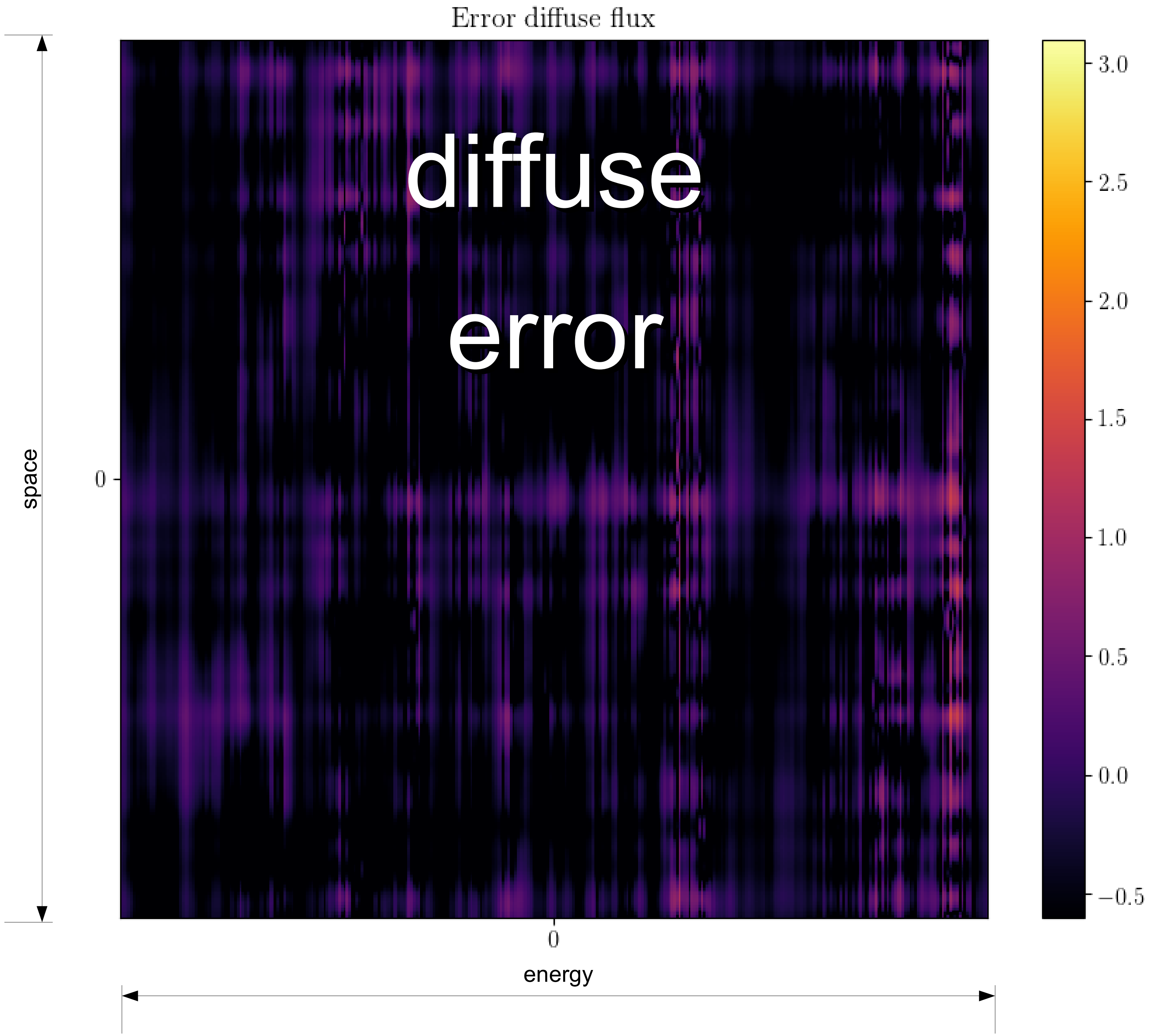


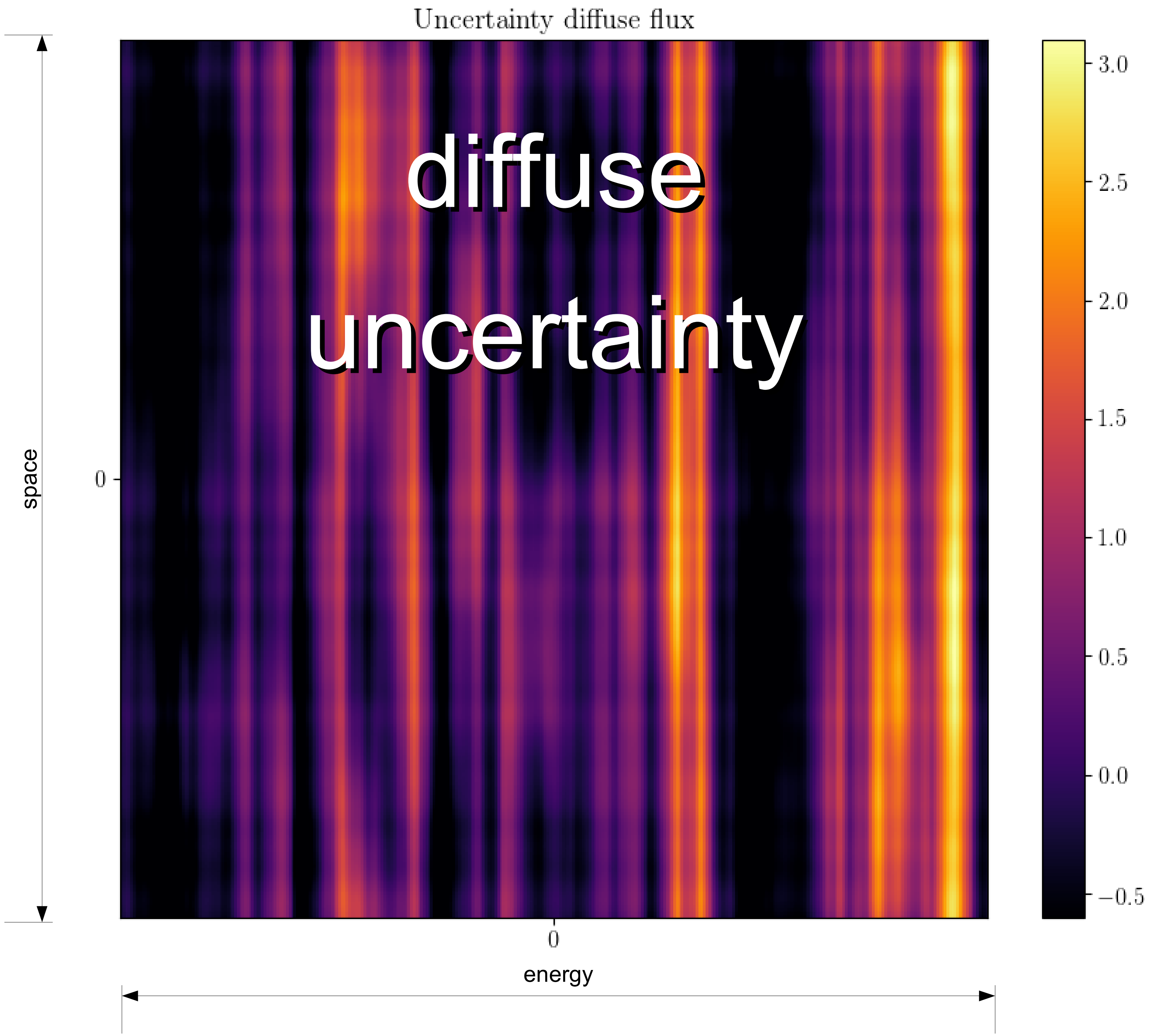
Photon counts
data
logarithmic
space
0
energy
0
3.0
2.5
2.0
1.5
1.0
0.5
0.0
−0.5

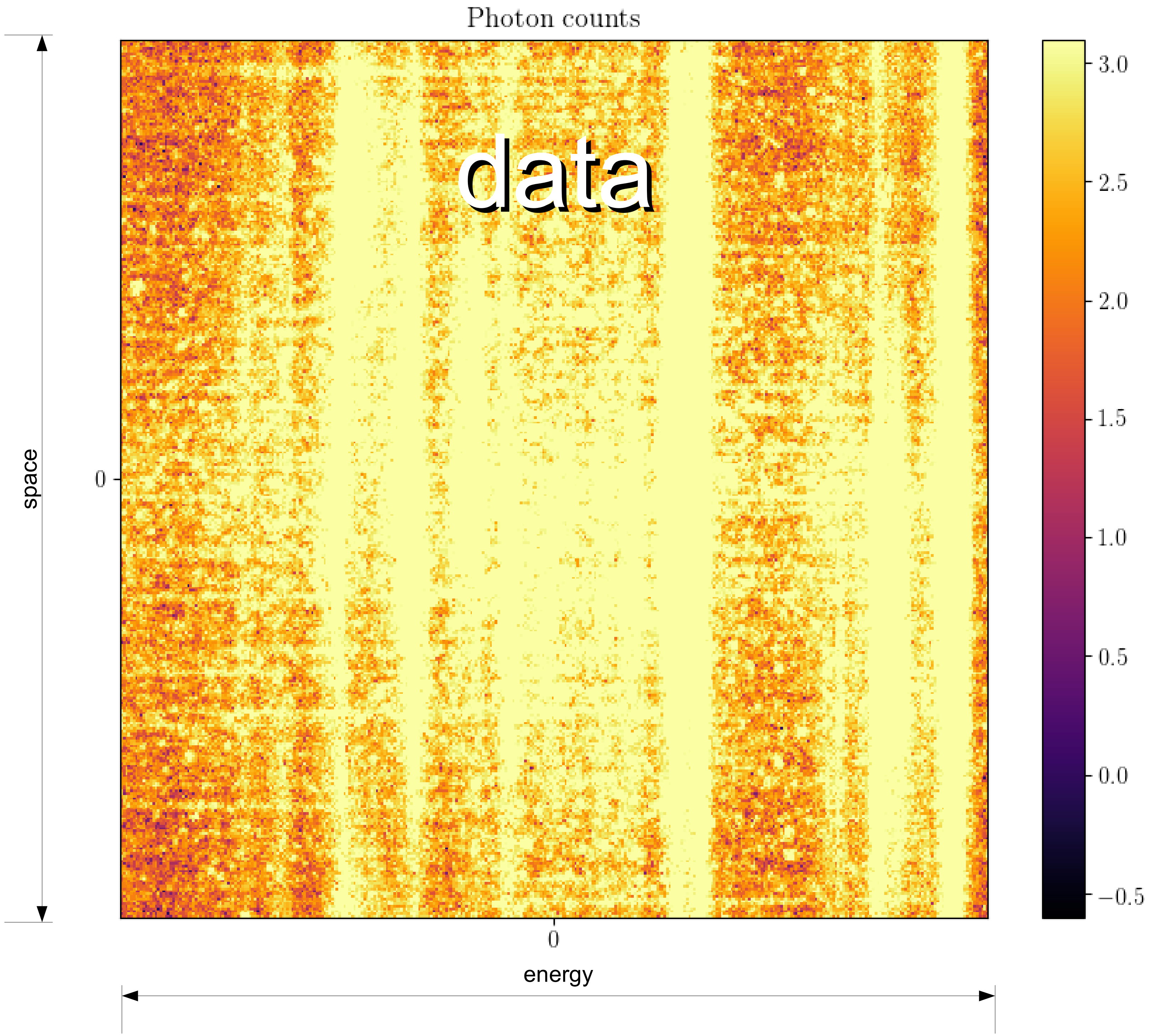

Reconstructed diffuse flux
diffuse
reconstructed
space
0
0
energy
3.0
2.5
2.0
1.5
1.0
0.5
0.0
−0.5

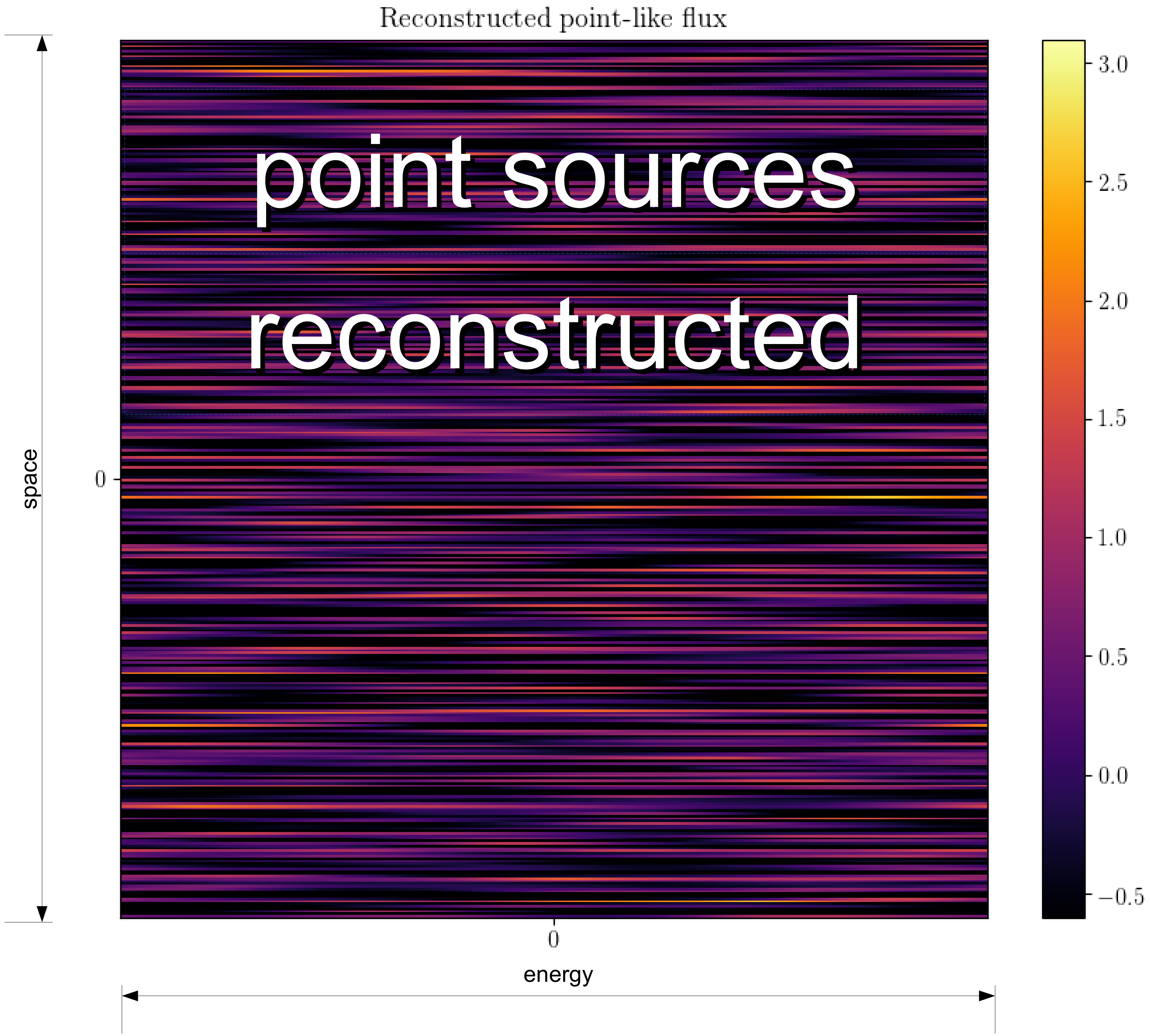


Diffuse flux
diffuse
true
space
0
0
energy
3.0
2.5
2.0
1.5
1.0
0.5
0.0
−0.5

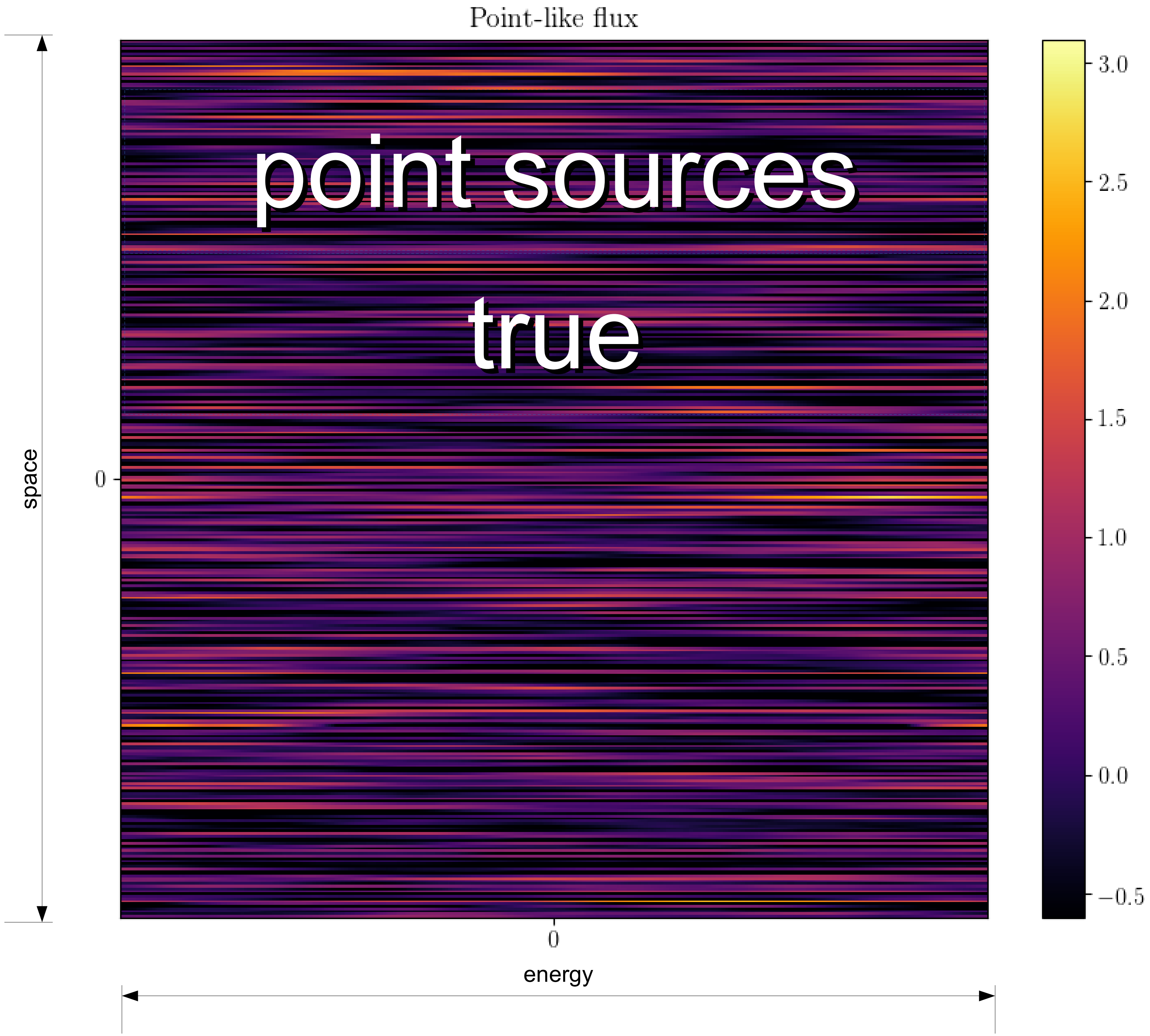

Error diffuse flux
diffuse
error
space
0
0
energy
3.0
2.5
2.0
1.5
1.0
0.5
0.0
−0.5


Uncertainty diffuse flux
diffuse
uncertainty
space
0
0
energy
3.0
2.5
2.0
1.5
1.0
0.5
0.0
−0.5


Photon counts
data
space
0
energy
0
3.0
2.5
2.0
1.5
1.0
0.5
0.0
−0.5

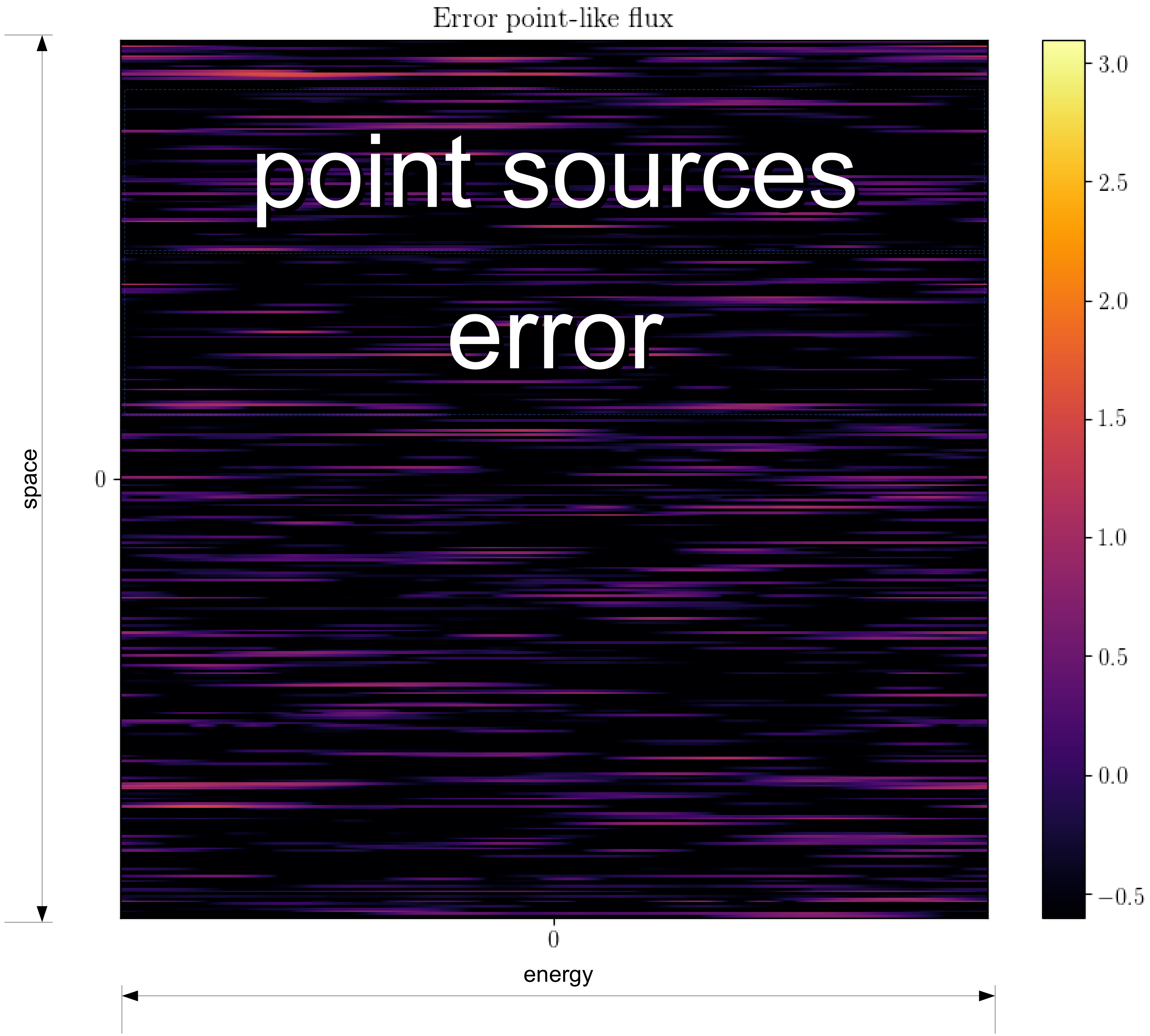

Reconstructed point-like flux
point sources
reconstructed
space
0
0
energy
3.0
2.5
2.0
1.5
1.0
0.5
0.0
−0.5

2018

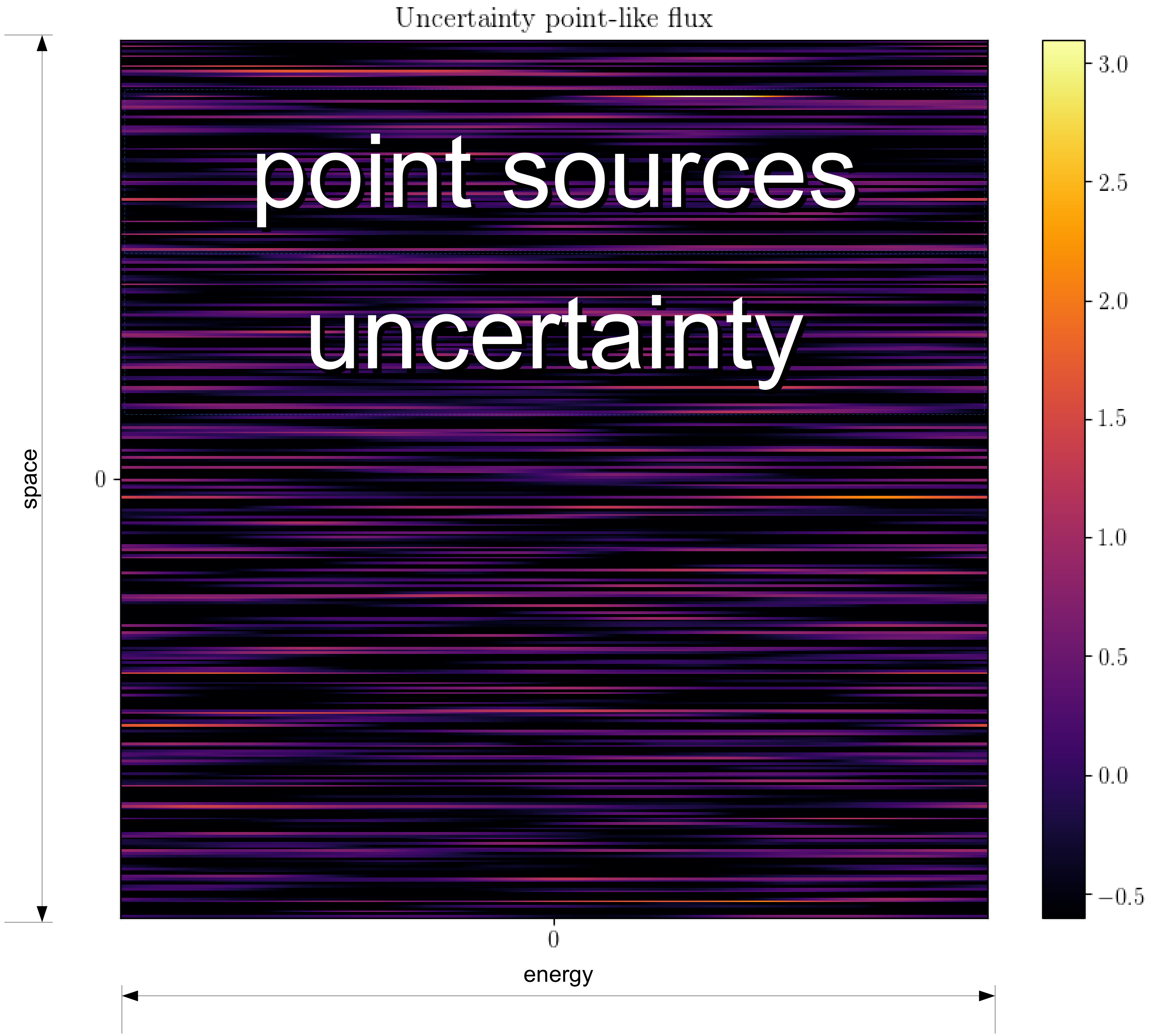


Point-like flux
point sources
true
space
0
energy
0
3.0
2.5
2.0
1.5
1.0
0.5
0.0
−0.5

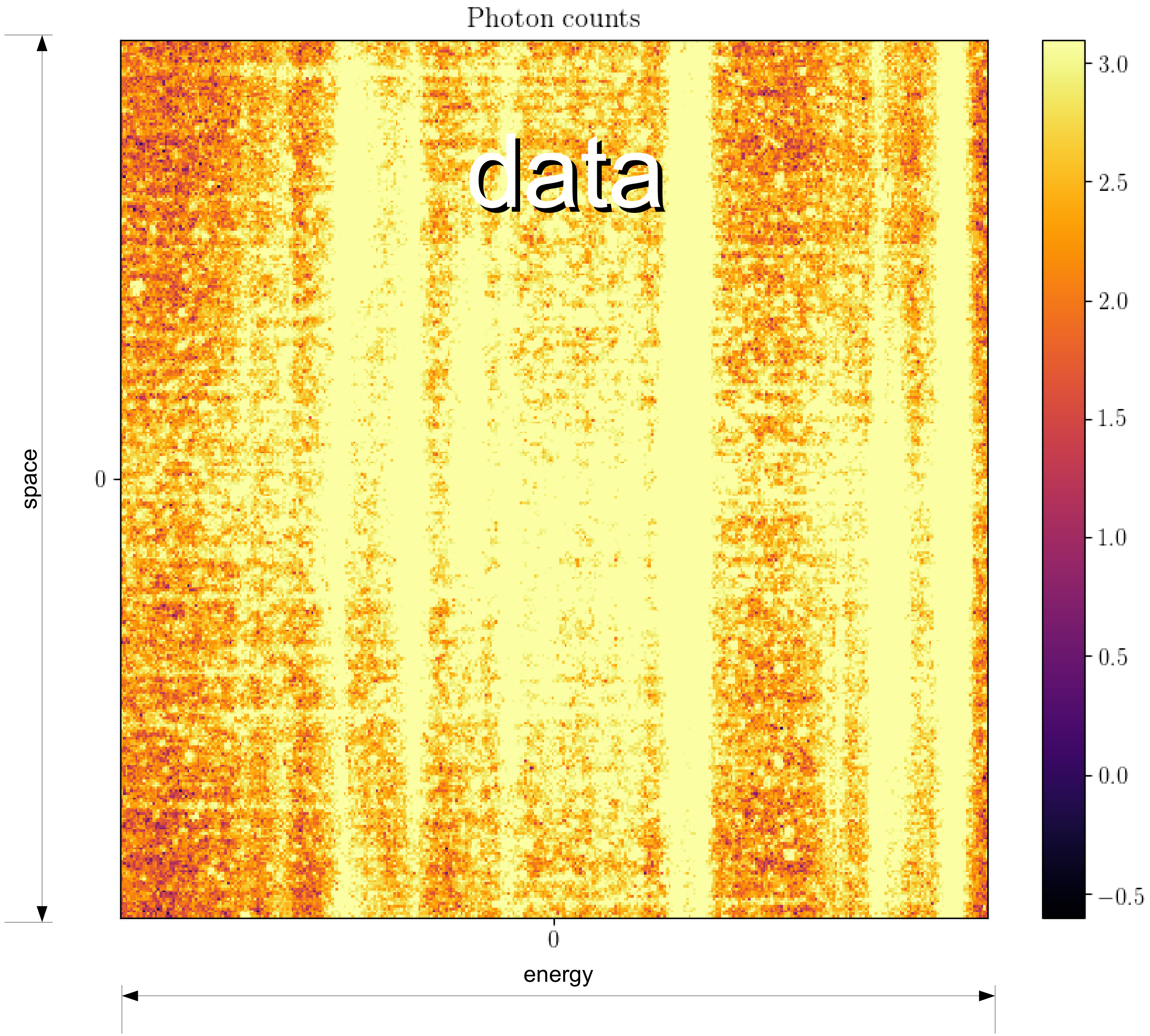
2018

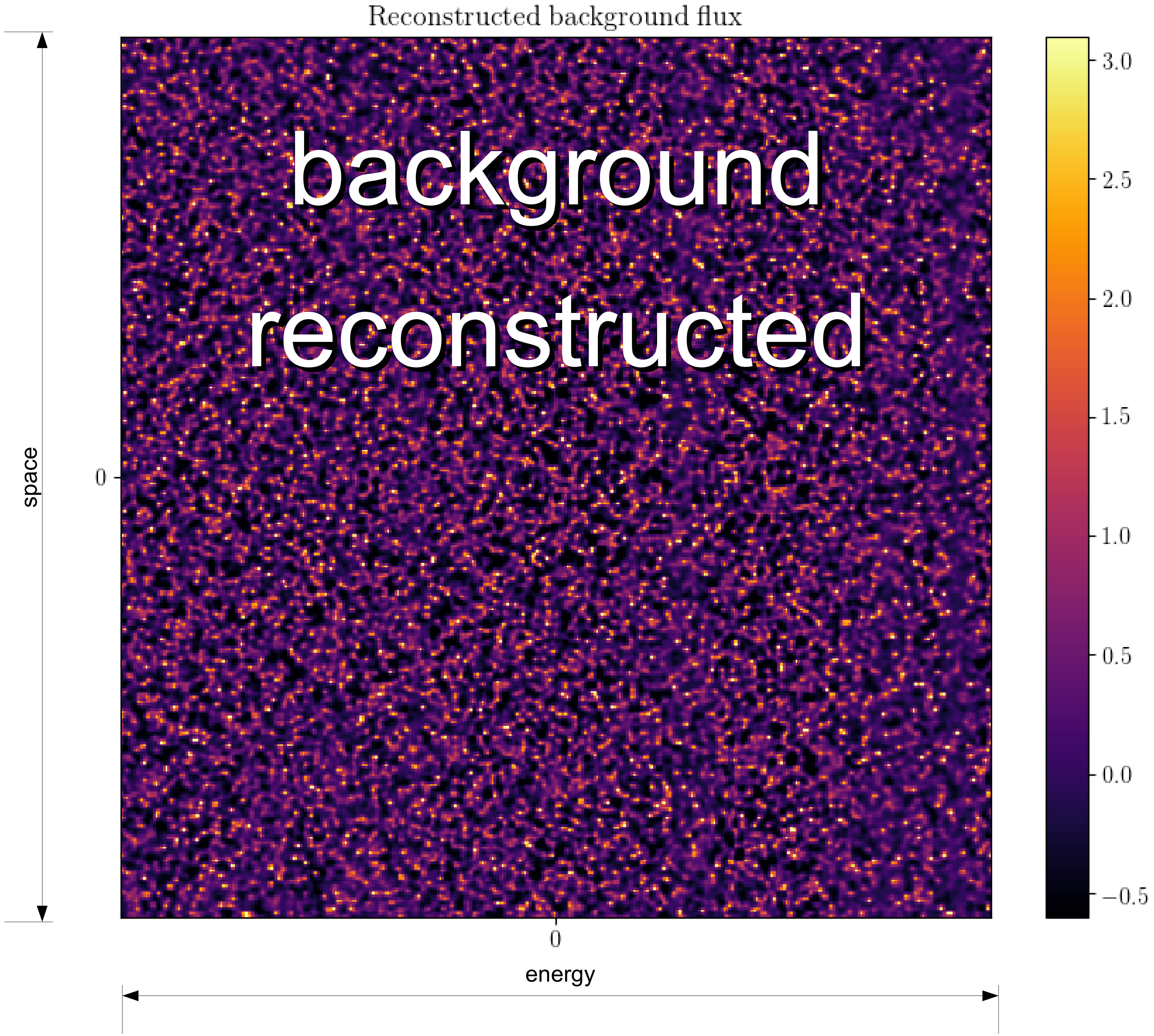

Error point-like flux
point sources
error
3.0
2.5
2.0
1.5
1.0
0.5
0.0
−0.5
space
0
0
energy

2018

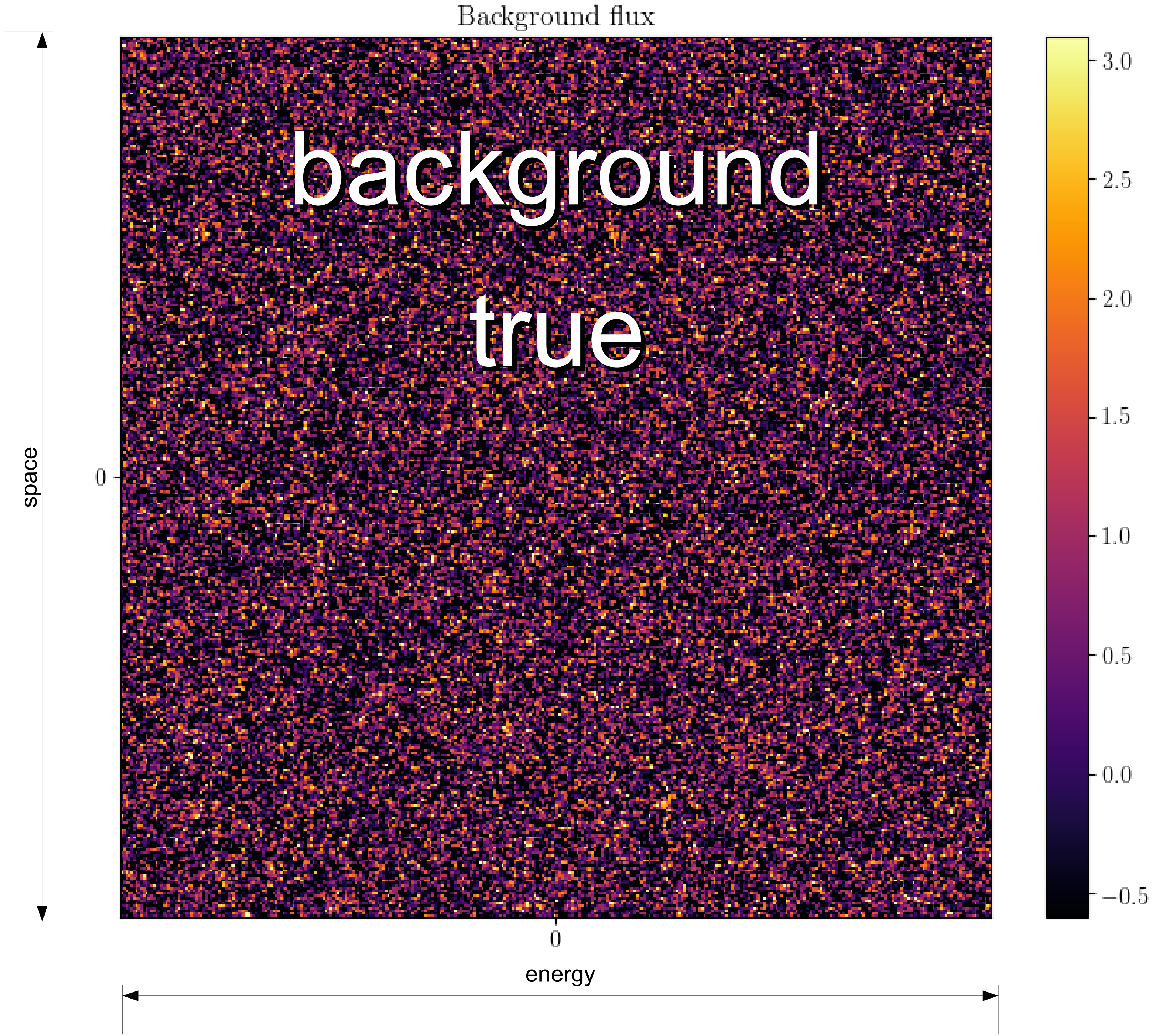

Uncertainty point-like flux
point sources
uncertainty
space
0
0
energy
3.0
2.5
2.0
1.5
1.0
0.5
0.0
−0.5

2018

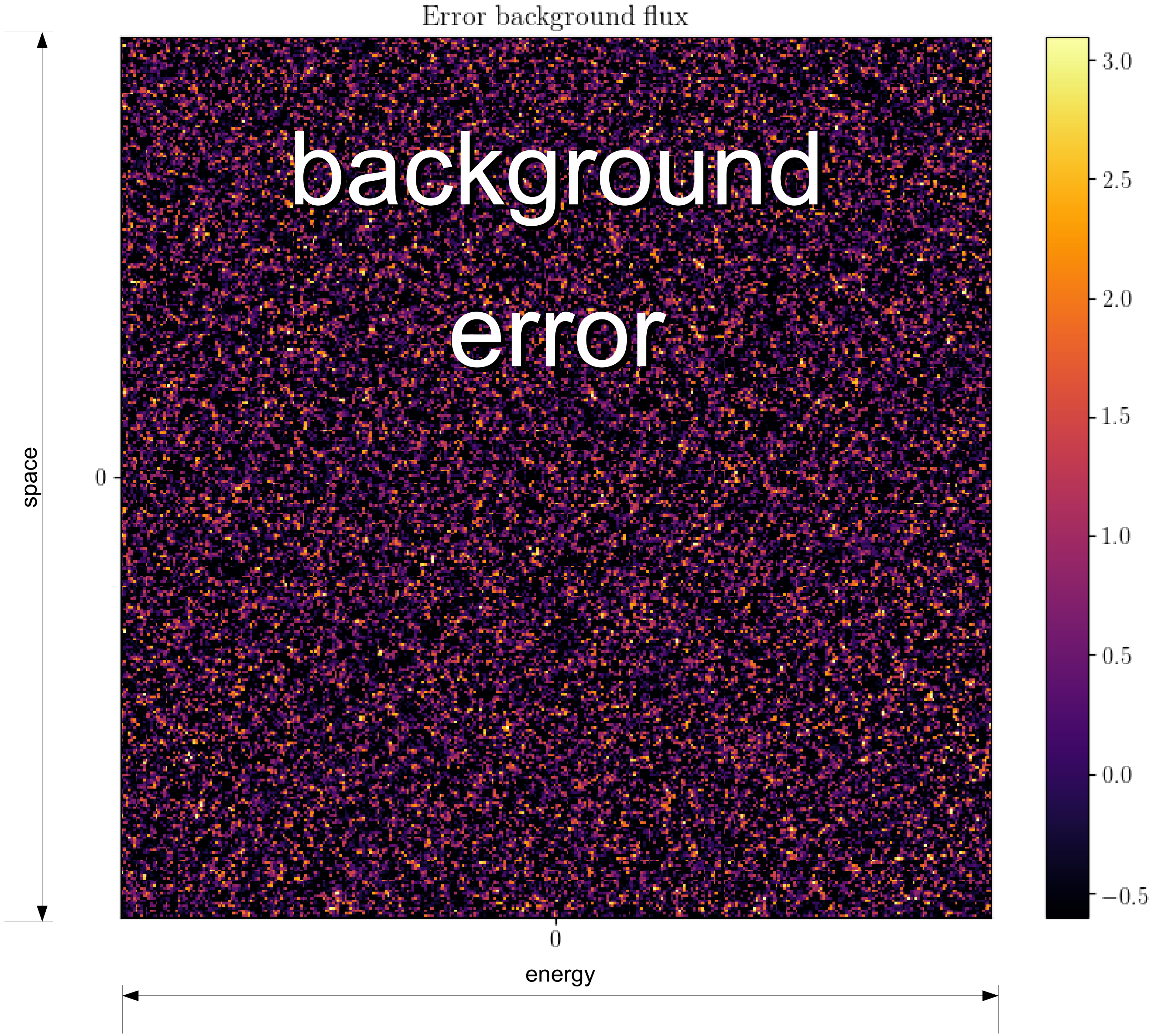

Photon counts
data
space
0
0
energy
3.0
2.5
2.0
1.5
1.0
0.5
0.0
−0.5

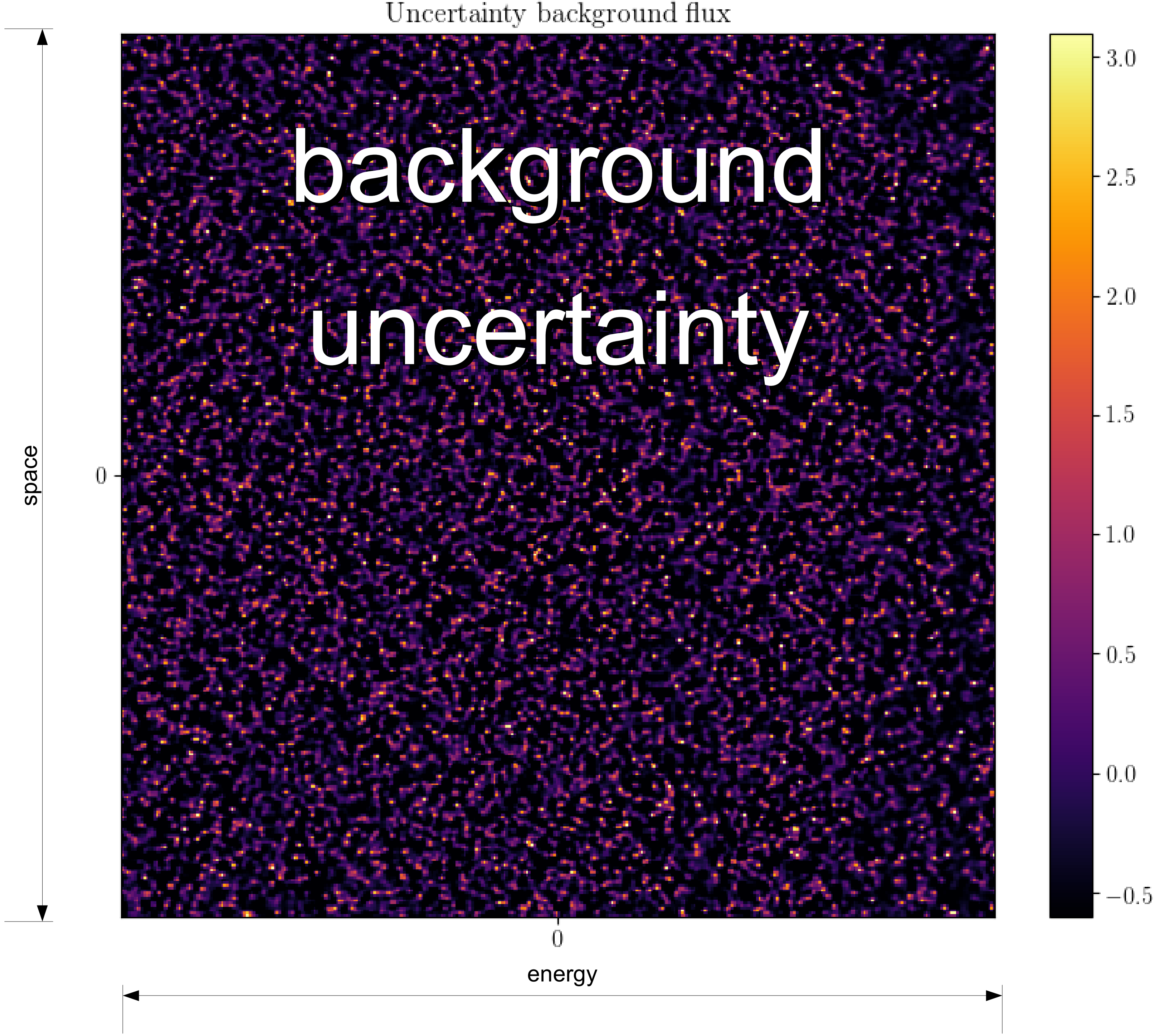


Reconstructed background flux
background
reconstructed
space
0
0
energy
3.0
2.5
2.0
1.5
1.0
0.5
0.0
−0.5



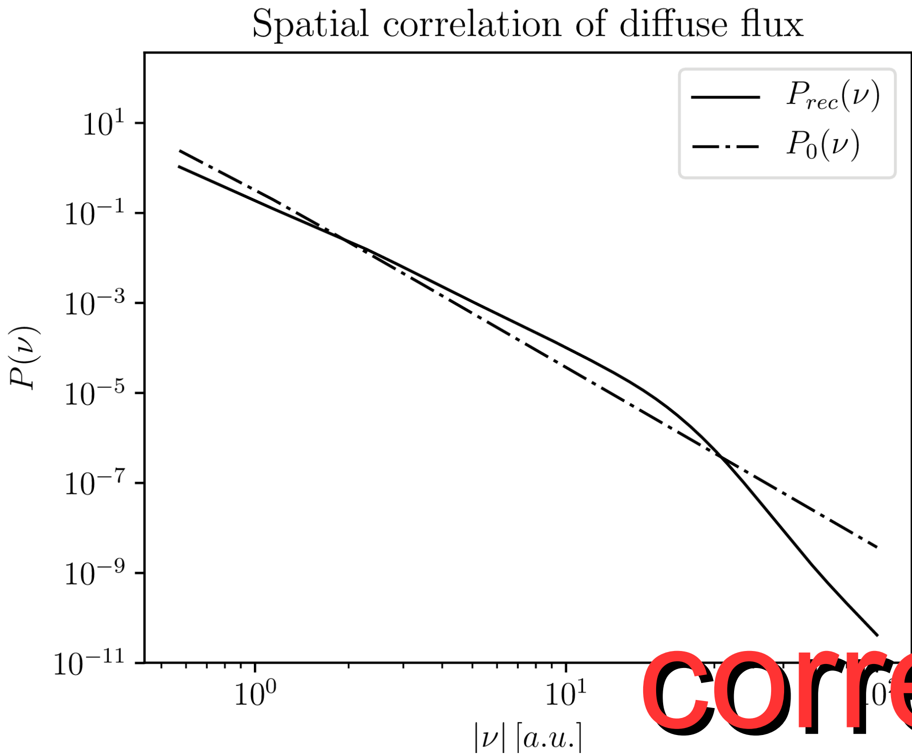
Background flux
background
true
space
0
0
energy
3.0
2.5
2.0
1.5
1.0
0.5
0.0
−0.5

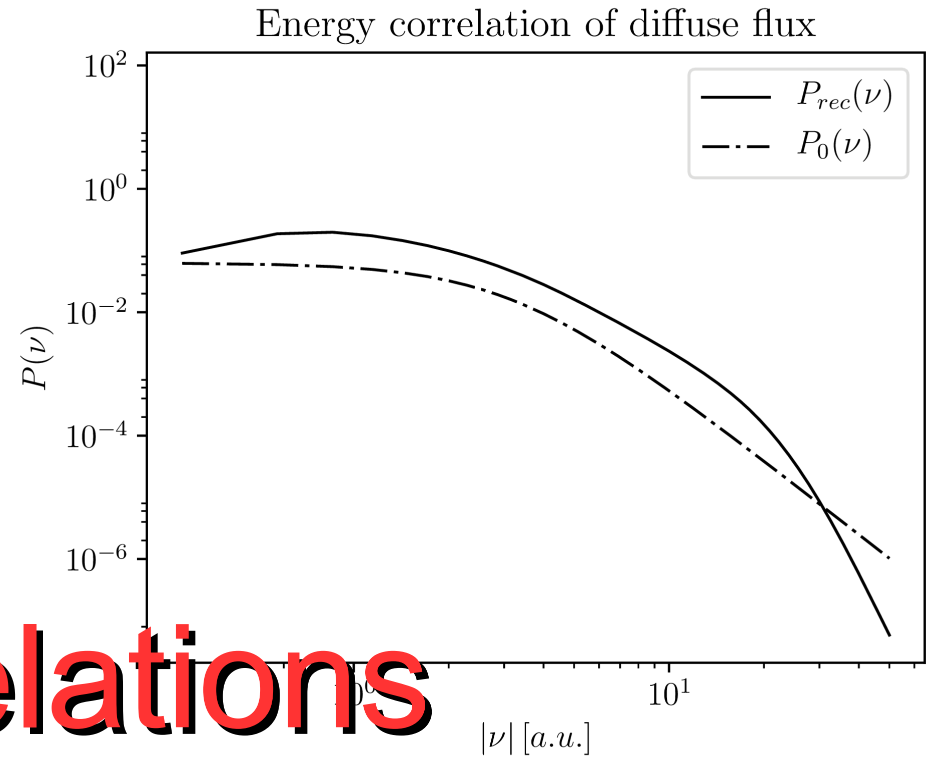

Error background flux
background
error
space
0
0
energy
3.0
2.5
2.0
1.5
1.0
0.5
0.0
−0.5

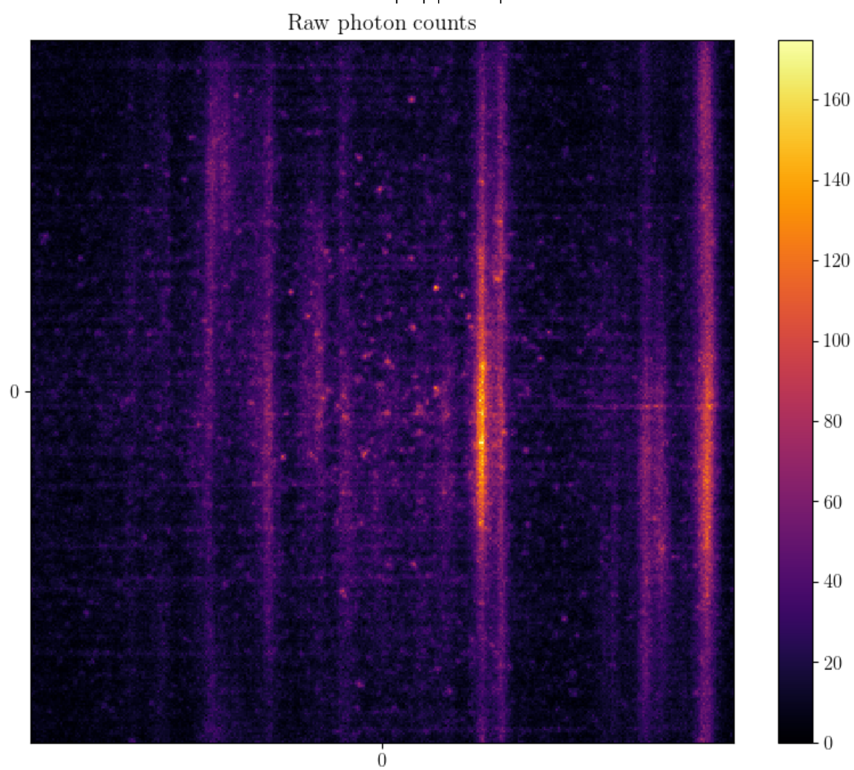


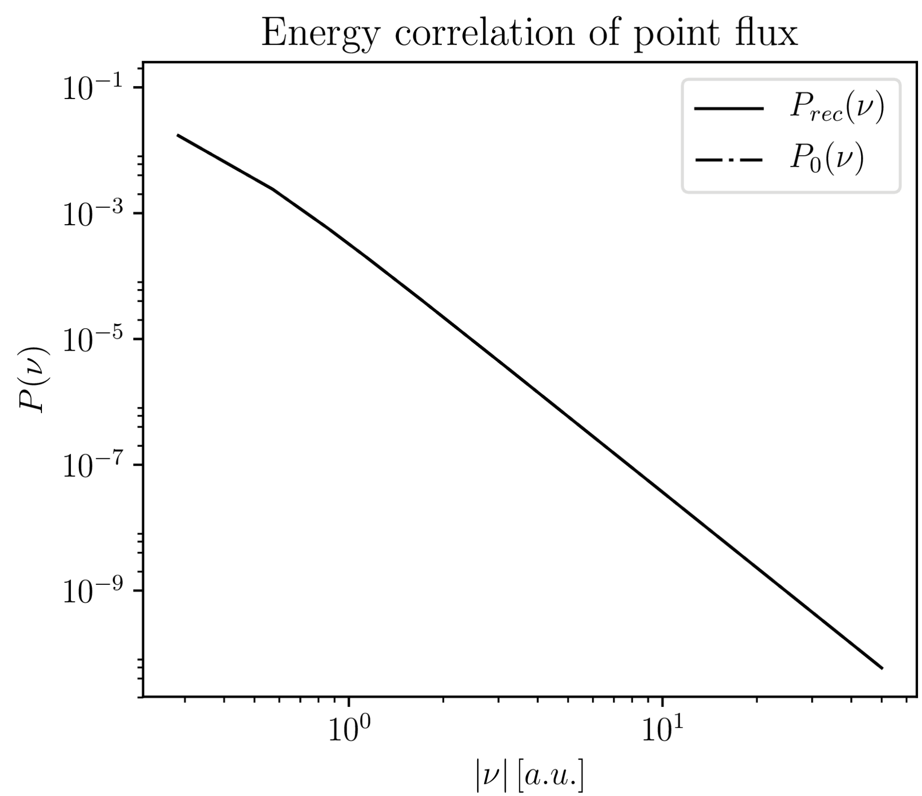
Uncertainty background flux
background
uncertainty
space
0
energy
0
3.0
2.5
2.0
1.5
1.0
0.5
0.0
−0.5

correlations

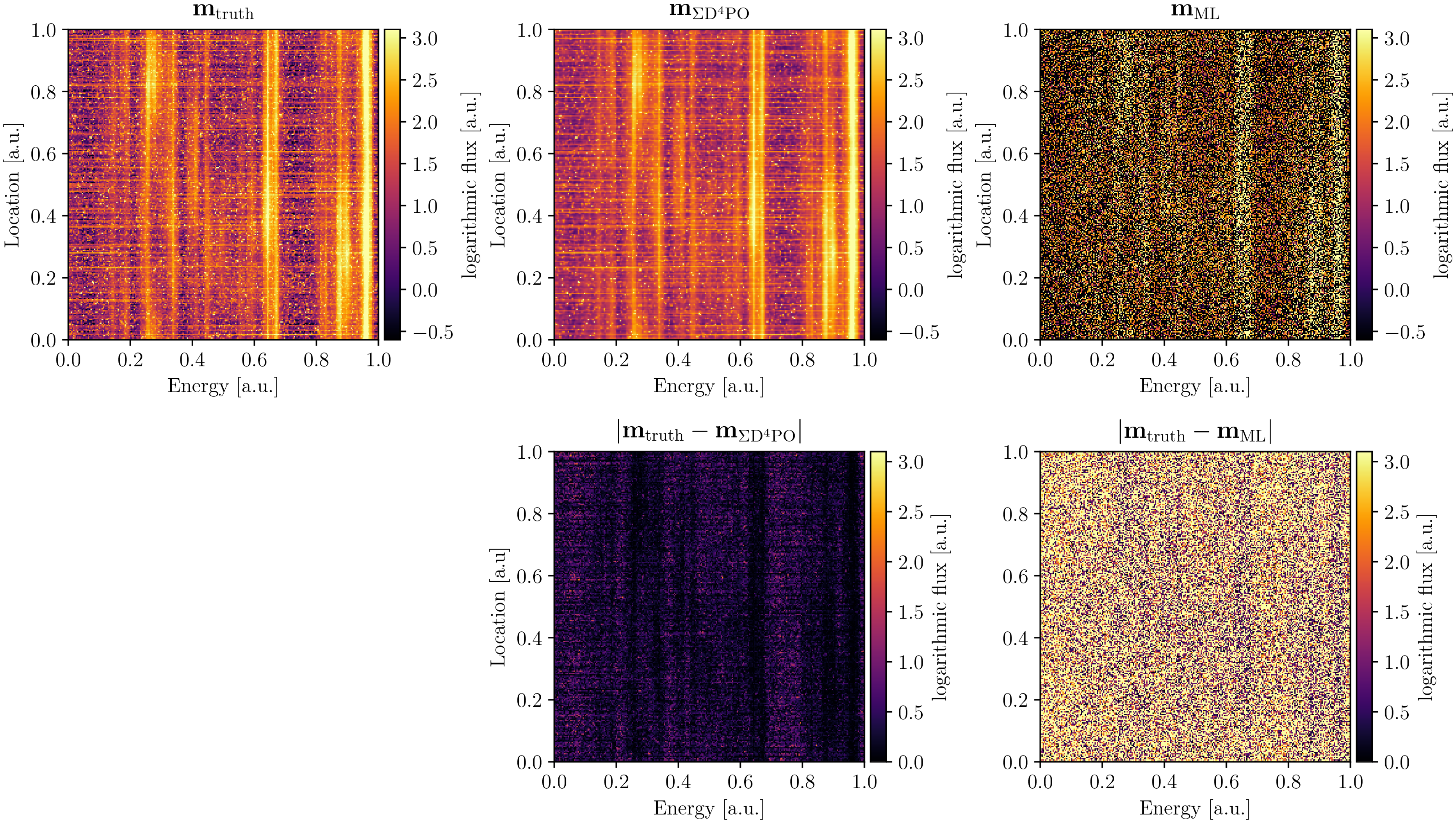








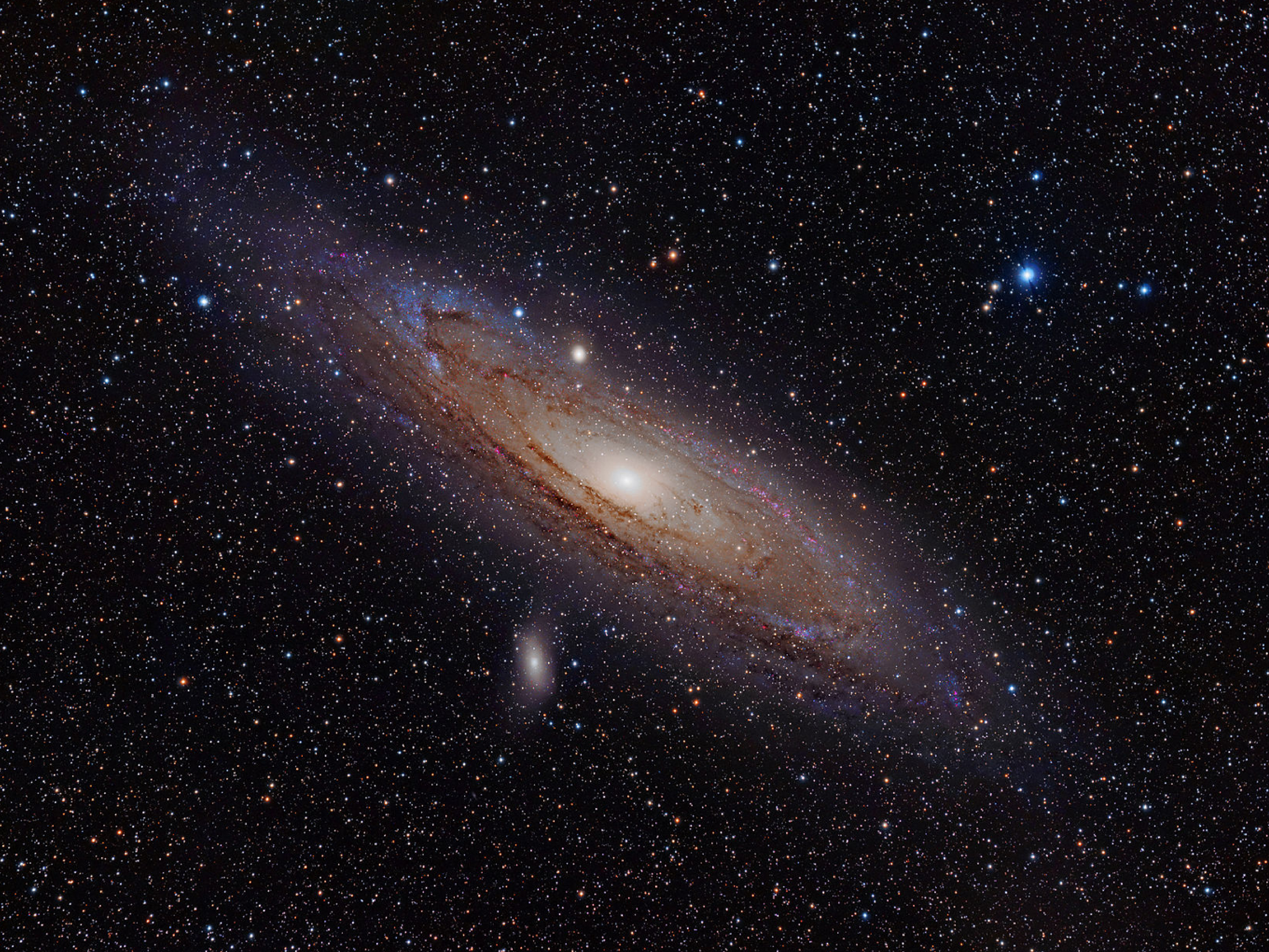
D^{4}PO vs ML

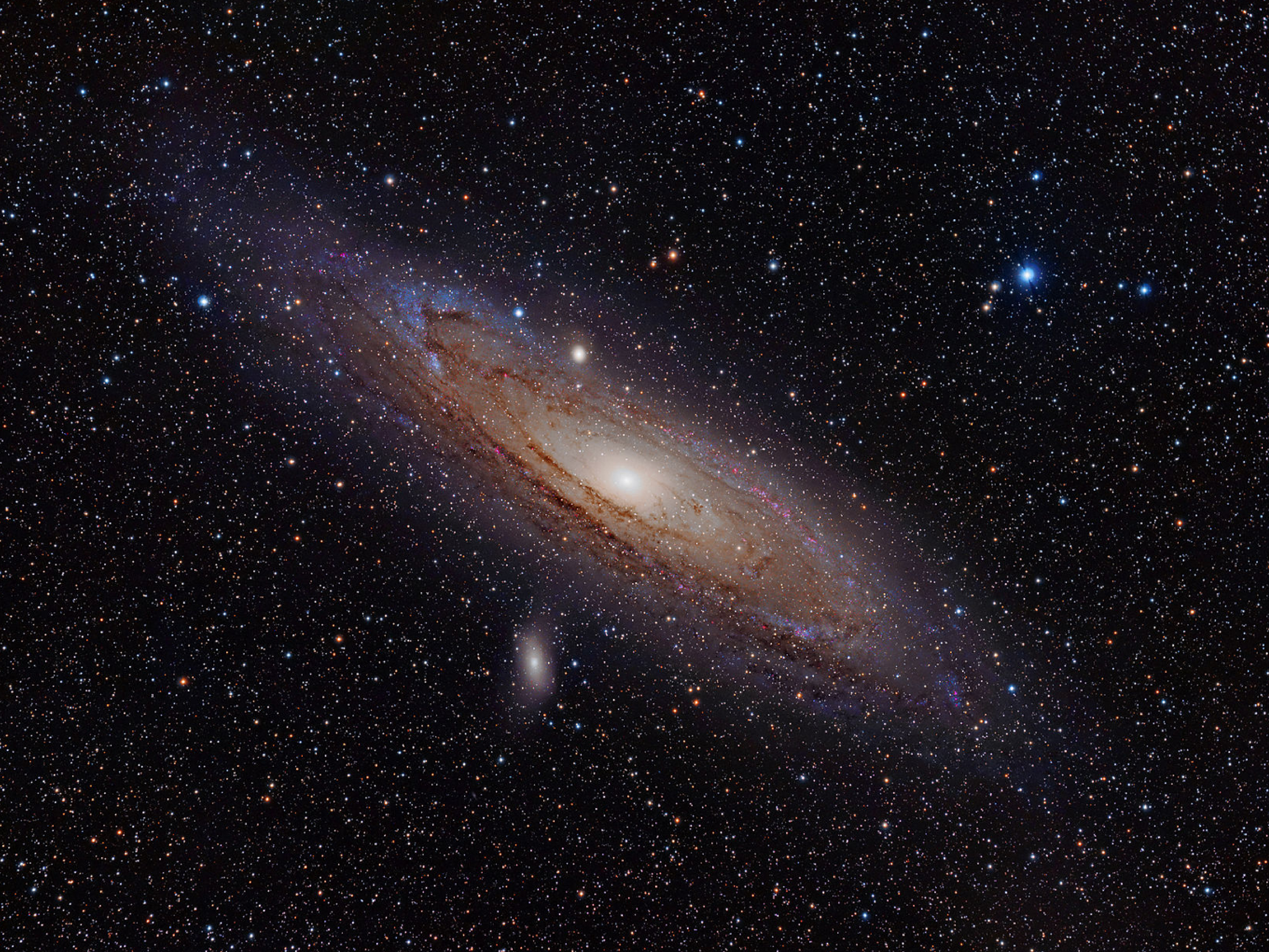


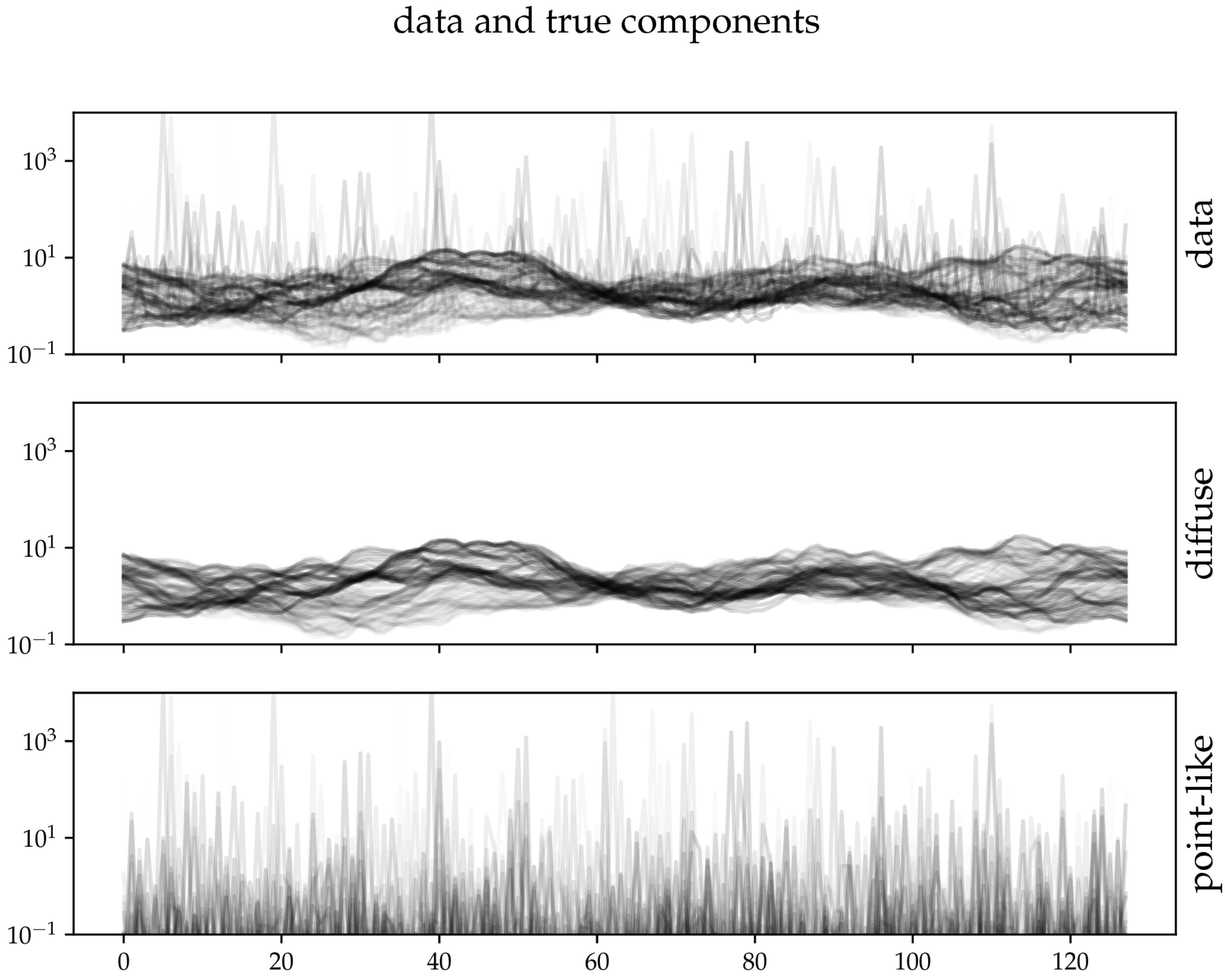
$\mathbf{m}_{\text{truth}}$
$\mathbf{m}_{\Sigma\text{D}^4\text{PO}}$
$\mathbf{m}_{\text{ML}}$
$|\mathbf{m}_{\text{truth}} - \mathbf{m}_{\Sigma\text{D}^4\text{PO}}|$
$|\mathbf{m}_{\text{truth}} - \mathbf{m}_{\text{ML}}|$
Location [a.u.]
Location [a.u]
Energy [a.u.]
logarithmic flux [a.u.]
0.0
0.2
0.4
0.6
0.8
1.0
−0.5
0.5
1.5
2.0
2.5
3.0

Glimpses into the future

Jakob Knollmüller
`starblade`

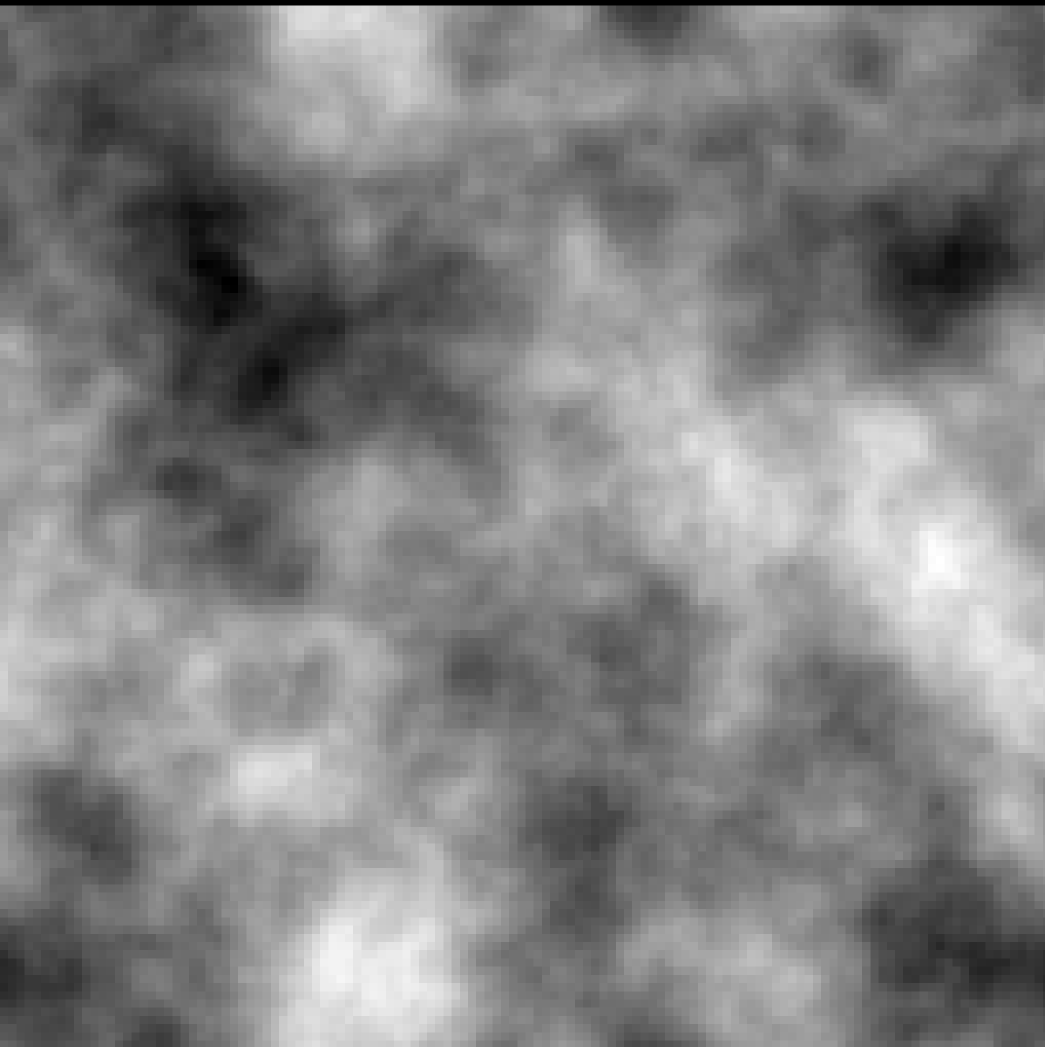




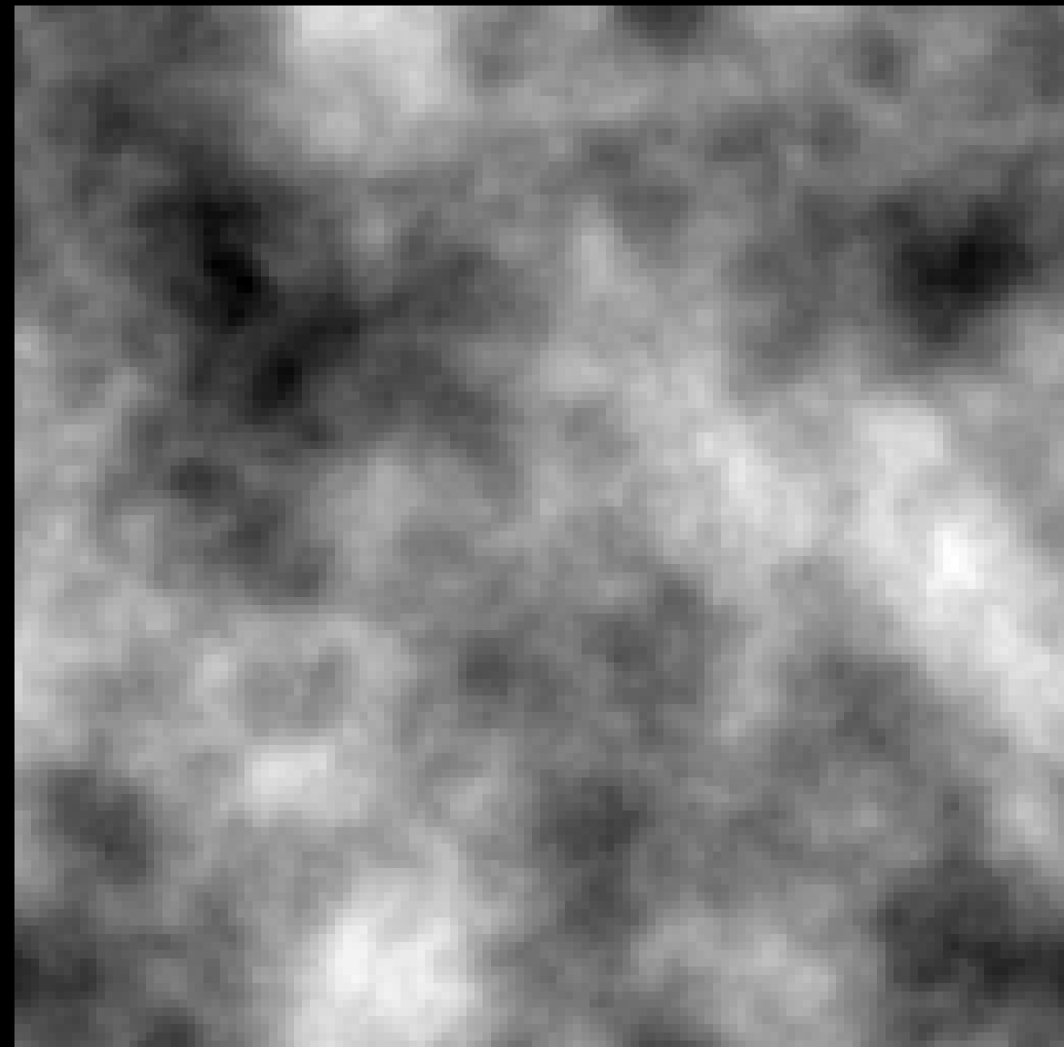


data and true components
data
diffuse
point-like
10^3
10^1
10^{-1}
0
20
40
60
80
100
120

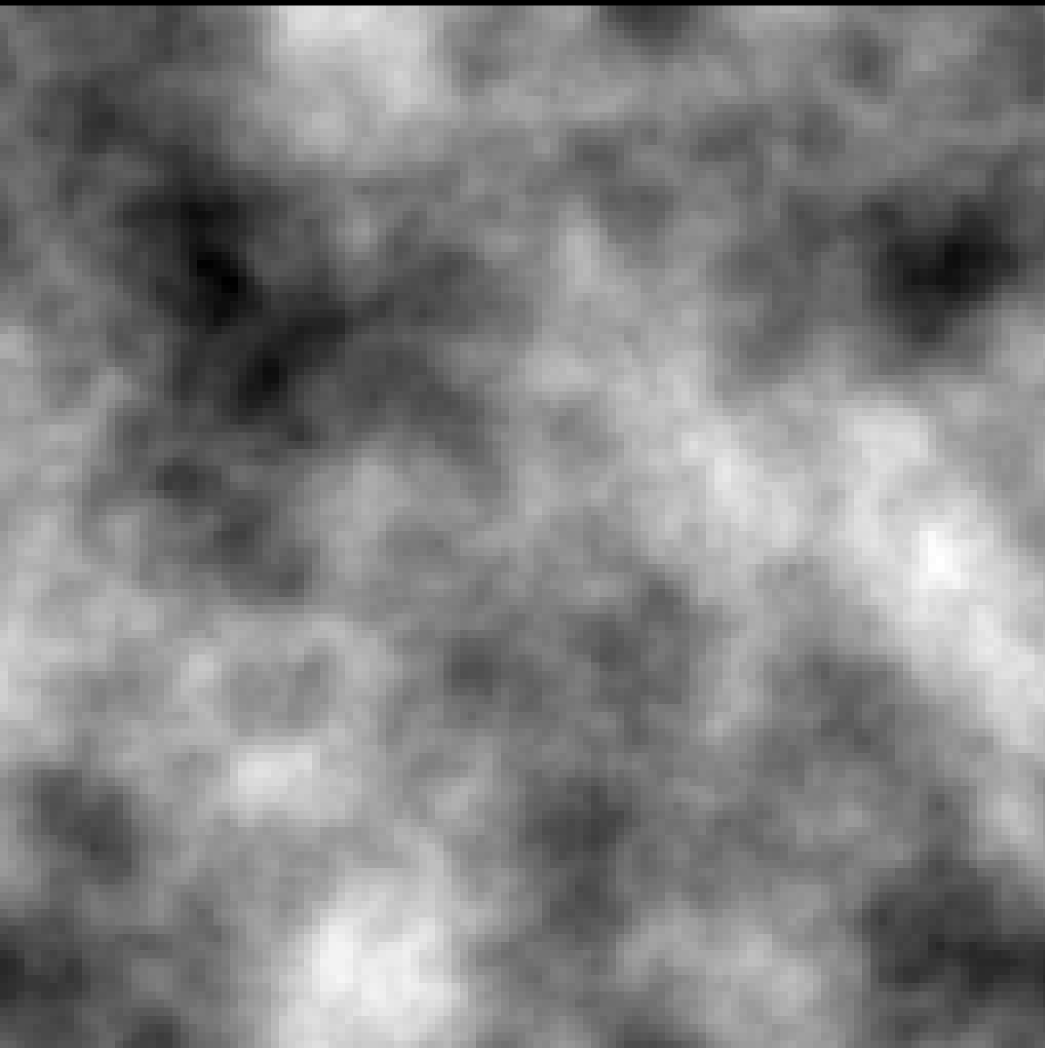
ground truth / starblade



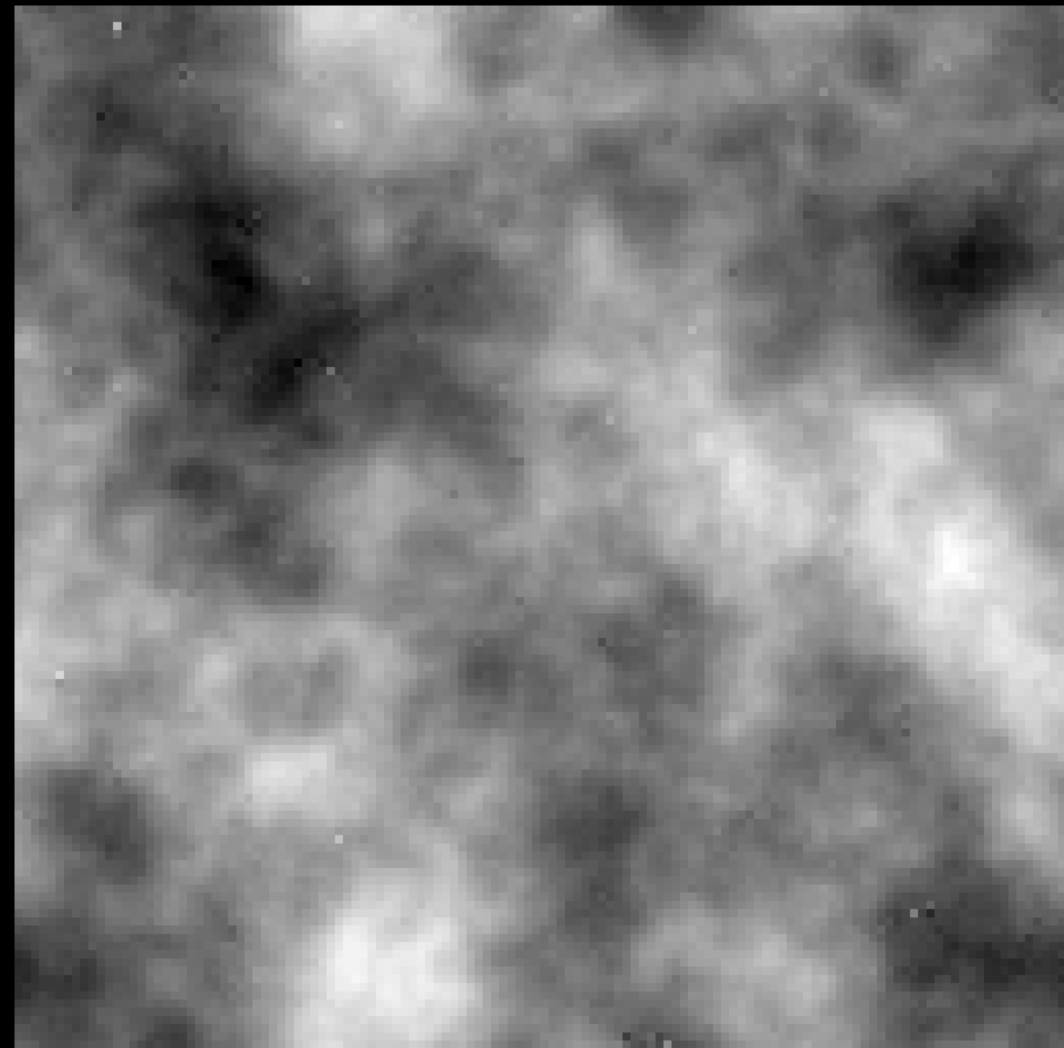
ground truth / autoencoder

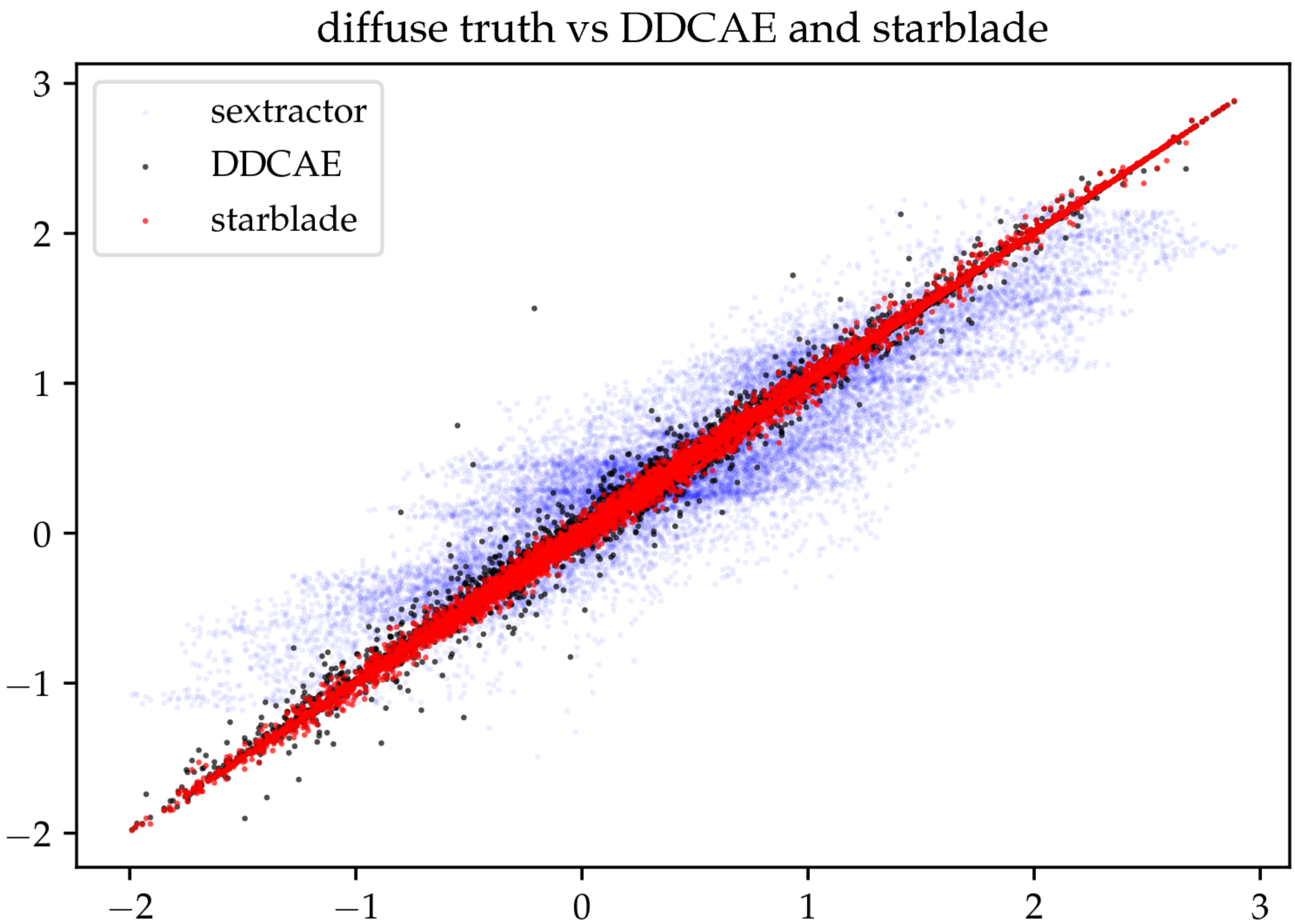


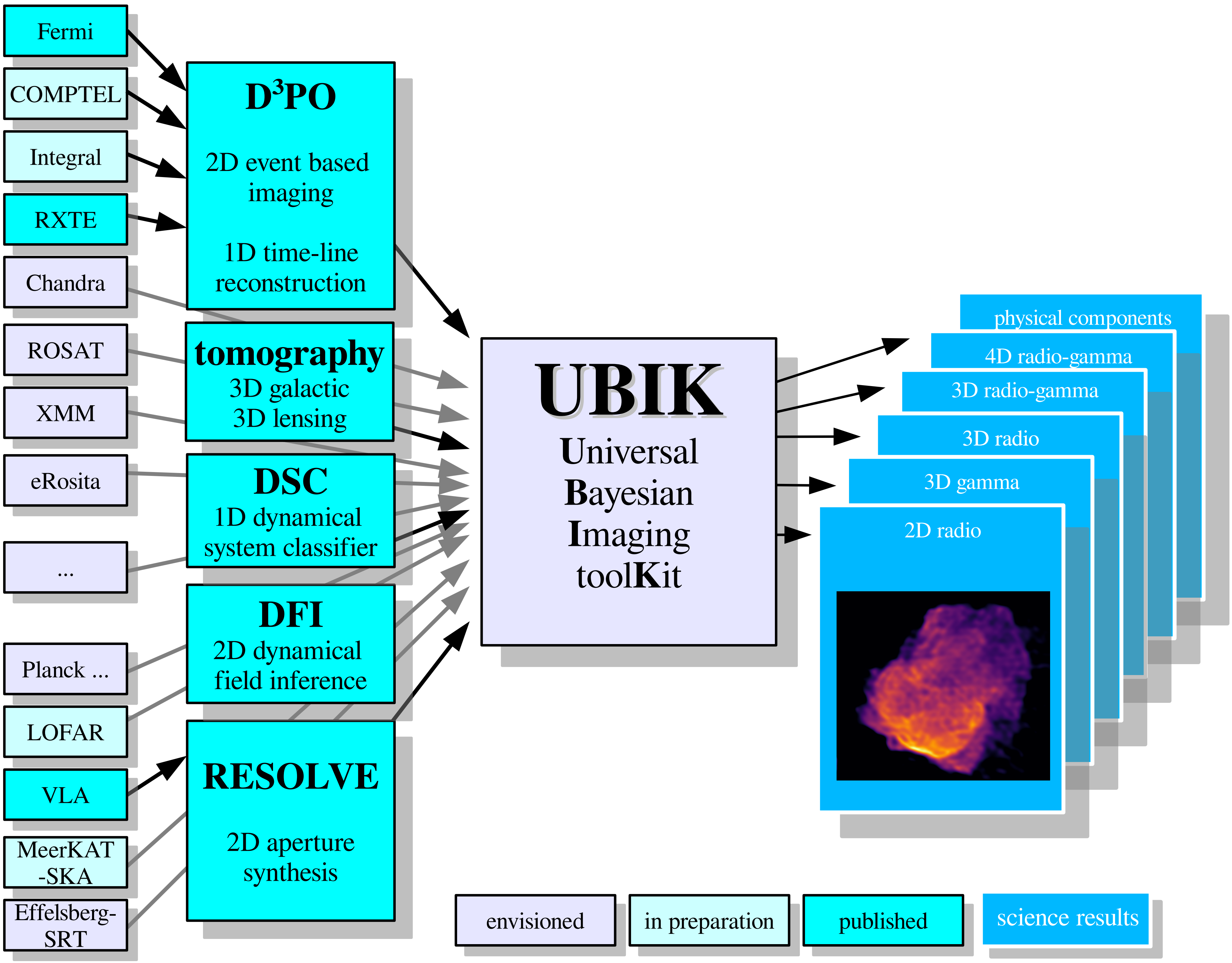
ground truth / starblade

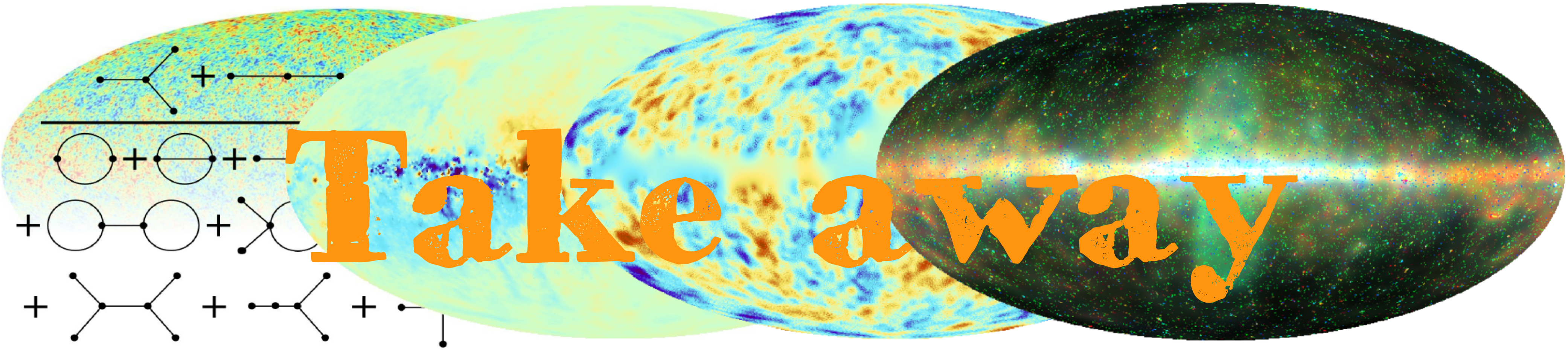


ground truth / autoencoder




diffuse truth vs DDCAE and starblade
sextractor
DDCAE
starblade
3
2
1
0
−1
−2
−2
−1
0
1
2
3

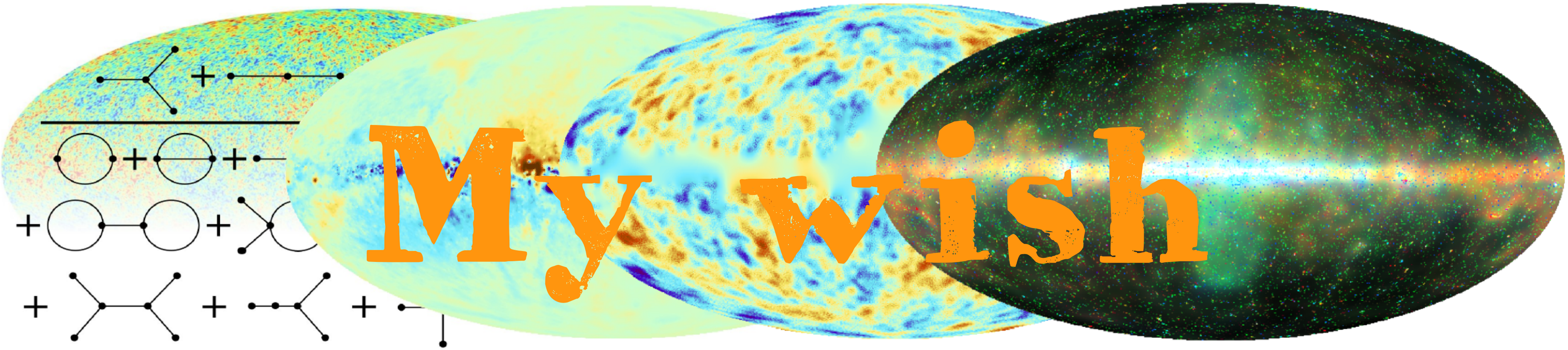

Fermi
COMPTEL
Integral
RXTE
Chandra
ROSAT
XMM
eRosita
...
Planck ...
LOFAR
VLA
MeerKAT
-SKA
Effelsberg-
SRT
D3PO
2D event based
imaging
1D time-line
reconstruction
tomography
3D galactic
3D lensing
DSC
1D dynamical
system classifier
DFI
2D dynamical
field inference
RESOLVE
2D aperture
synthesis
UBIK
Universal
Bayesian
Imaging
toolKit
physical components
4D radio-gamma
3D radio-gamma
3D radio
3D gamma
2D radio
envisioned
in preparation
published
science results



Take away

Imaging goes inference

IFT	-	information field theory
NIFTy	-	numerical IFT in Python
UBIK	-	multi-instrument & -dimension imaging & tomography

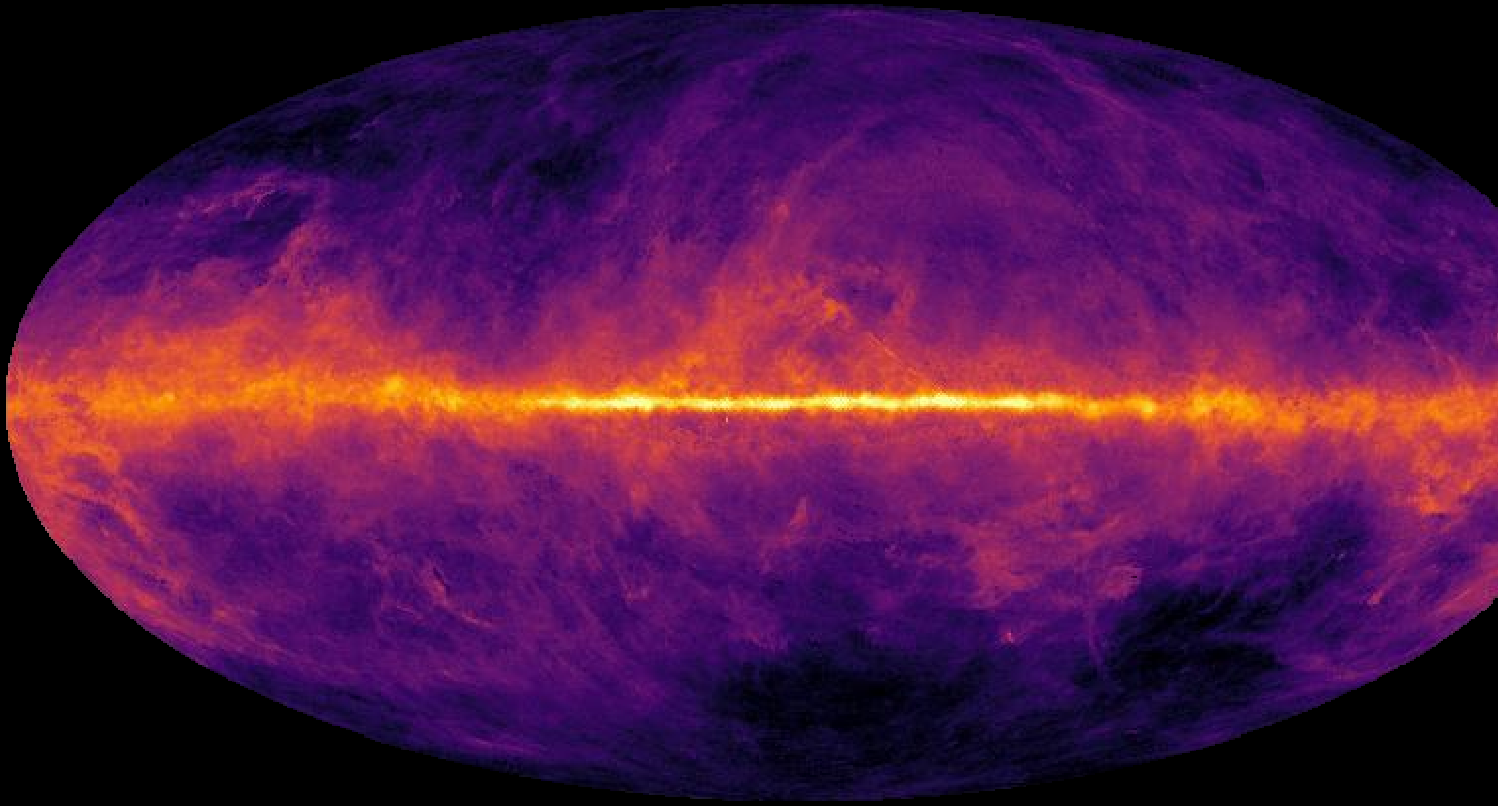


My wish

If you build an instrument please provide

- a fast response code
  R: sky → data
- its derivative
  R' = dR/dsky: (sky, dsky) → ddata
- its adjoint
  $R'^{\dagger}$: (sky, ddata) → dsky

and UBIK will be your friend.



Thank you!

www.mpa-garching.mpg.de/ift