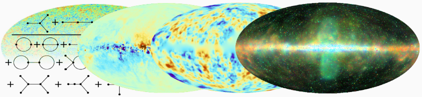


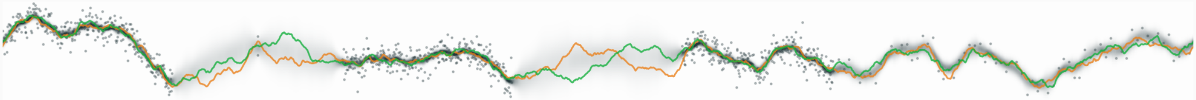
Numerical information field theory

Bayesian imaging using IFT

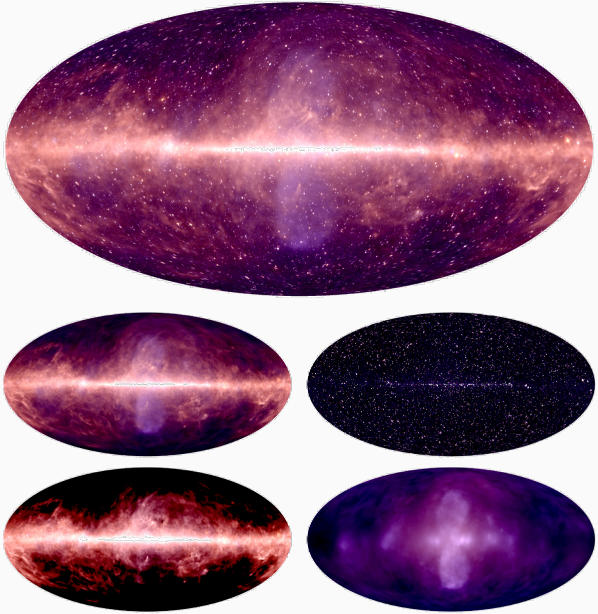
Philipp Frank[1]

The Road to Differentiable and Probabilistic Programming in Fundamental Physics,
Max Planck Institute for Extraterrestrial Physics, Garching, June 28, 2023

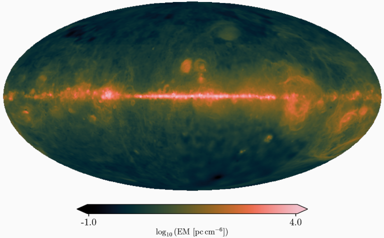
(1) Max-Planck Institute for Astrophysics MPA, Garching, Germany

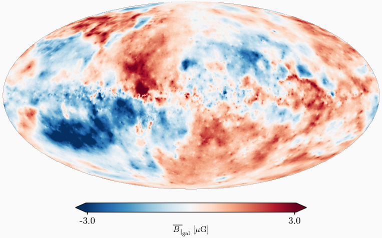


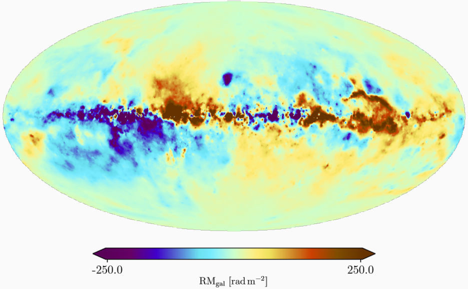
Fermi - Gamma Ray [PKA+22]



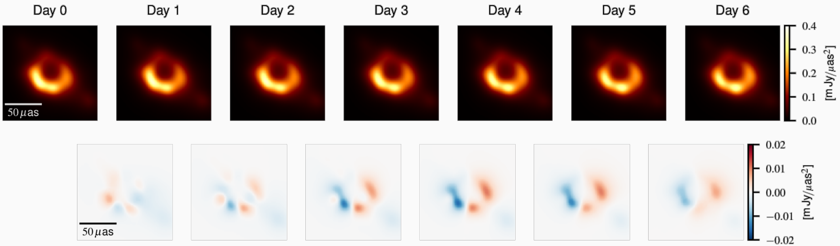
Faraday Tomography [HHF+23]





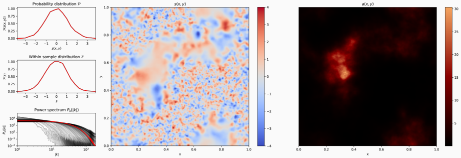


Radio Interferometry (VLBI) - M87* [AFH+22]



NIFTy - Toolkit [AEF+22]

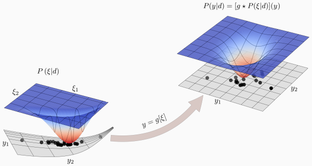
Gaussian & Generative processes



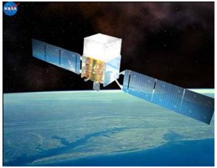
Automatic differentiation

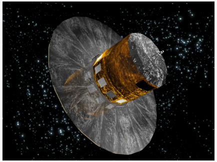
NIFTy [1]

Variational Inference



Common Likelihood & Instrument Models


[2]


[3]

[1] jax.readthedocs.io/
[2] www.nasa.gov/
[3] www.esa.int/

Gaussian Processes

Gaussian processes

- Probability distributions $\mathcal{P}(s)$ over functions $s_x \equiv s(x)$, with $s \in \mathcal{L}^2[\Omega]$, $x \in \Omega \subset \mathbb{R}^N$.

Gaussian processes

- Probability distributions $\mathcal{P}(s)$ over functions $s_x \equiv s(x)$, with $s \in \mathcal{L}^2[\Omega]$, $x \in \Omega \subset \mathbb{R}^N$.
- Gaussian processes are fully specified by their one- and two-point correlation functions:
- Mean field: $m_x = m(x) \equiv \langle s_x \rangle_{\mathcal{P}(s)}$.
- Correlation structure: $C_{xy} = C(x, y) \equiv \left\langle (s_x - m_x)(s_y - m_y)^* \right\rangle_{\mathcal{P}(s)}$.

Gaussian processes

- Probability distributions $\mathcal{P}(s)$ over functions $s_x \equiv s(x)$, with $s \in \mathcal{L}^2[\Omega]$, $x \in \Omega \subset \mathbb{R}^N$.
- Gaussian processes are fully specified by their one- and two-point correlation functions:
- Mean field: $m_x = m(x) \equiv \langle s_x \rangle_{\mathcal{P}(s)}$.
- Correlation structure: $C_{xy} = C(x, y) \equiv \left\langle (s_x - m_x)(s_y - m_y)^* \right\rangle_{\mathcal{P}(s)}$.

- Generative GP: $s(x) = m(x) + \int A(x, y)\, \xi(y)\, \mathrm{d}y \equiv (m + A\xi)(x)$
- With $AA^\dagger \equiv C$ and $\xi \leftarrow \mathcal{N}(\xi; 0, \mathbb{1})$

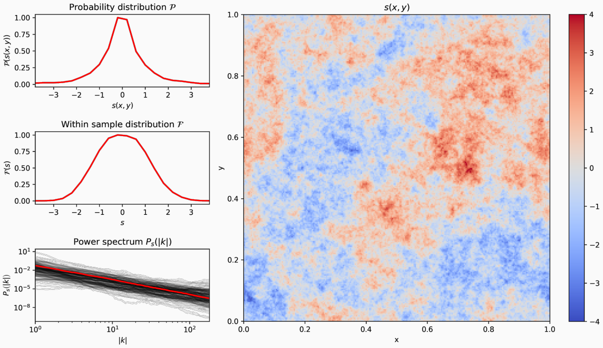
Gaussian processes

- Probability distributions $\mathcal{P}(s)$ over functions $s_x \equiv s(x)$, with $s \in \mathcal{L}^2[\Omega]$, $x \in \Omega \subset \mathbb{R}^N$.
- Gaussian processes are fully specified by their one- and two-point correlation functions:
- Mean field: $m_x = m(x) \equiv \langle s_x \rangle_{\mathcal{P}(s)}$.
- Correlation structure: $C_{xy} = C(x, y) \equiv \left\langle (s_x - m_x)(s_y - m_y)^* \right\rangle_{\mathcal{P}(s)}$.

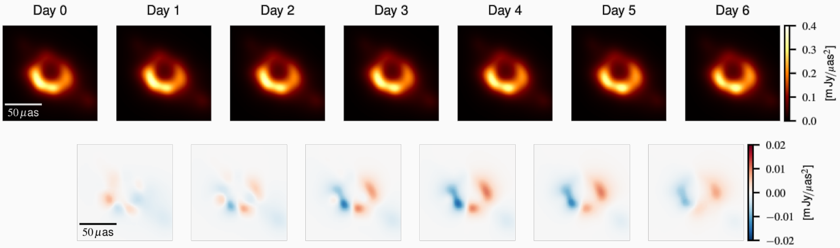
- Generative GP: $s(x) = m(x) + \int A(x, y)\, \xi(y)\, \mathrm{d}y \equiv (m + A\xi)(x)$
- With $AA^\dagger \equiv C$ and $\xi \leftarrow \mathcal{N}(\xi; 0, \mathbb{1})$
- Generative Amplitude: $A(x, y) \equiv A_\sigma(x, y)$
- With $\sigma = \sigma(\xi_\sigma)$ and $\xi_\sigma \leftarrow \mathcal{N}(\xi_\sigma; 0, \mathbb{1})$

GP - Priors

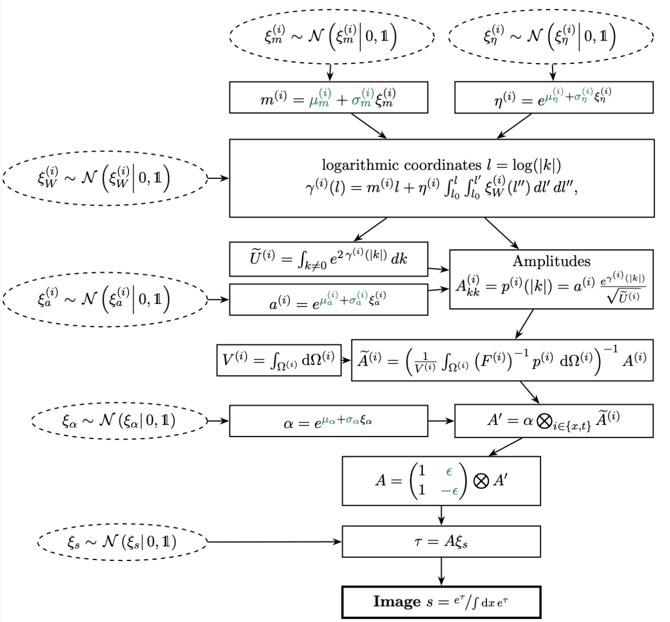
$$s = A\,\xi, \quad \text{with} \quad A \propto \mathcal{F}^{-1}\widehat{\sqrt{P_s}}, \quad P_s(k) \propto e^{\tau(k)} \;.$$



VLBI - M87* [AFH+22]

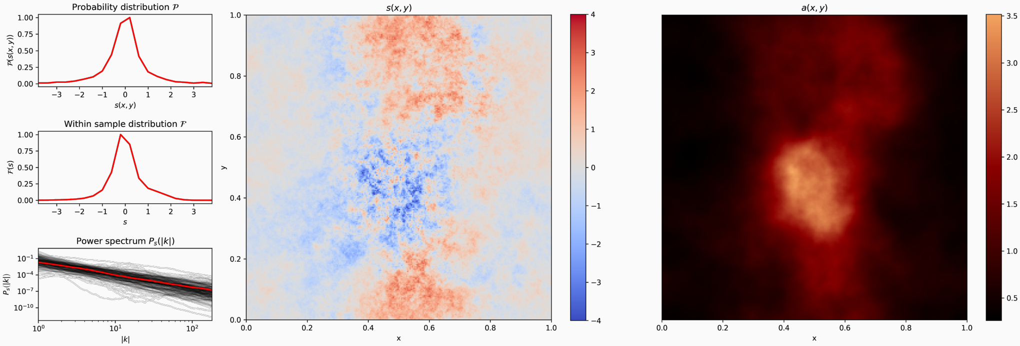


VLBI - M87* [AFH+22]

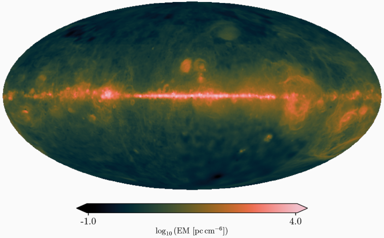


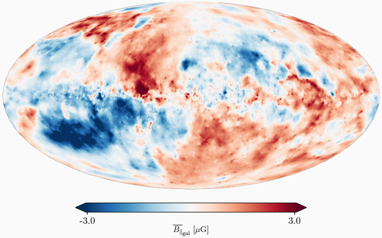
GP - Priors

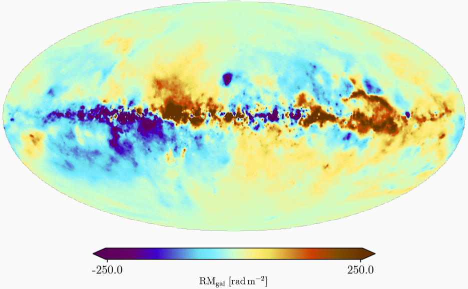
$$s = A\,\xi, \quad \text{with} \quad A \propto \hat{a}\,\mathcal{F}^{-1}\widehat{\sqrt{P_s}}, \quad P_s(k) \propto e^{\tau(k)}\ .$$



Faraday Tomography [HHF+23]

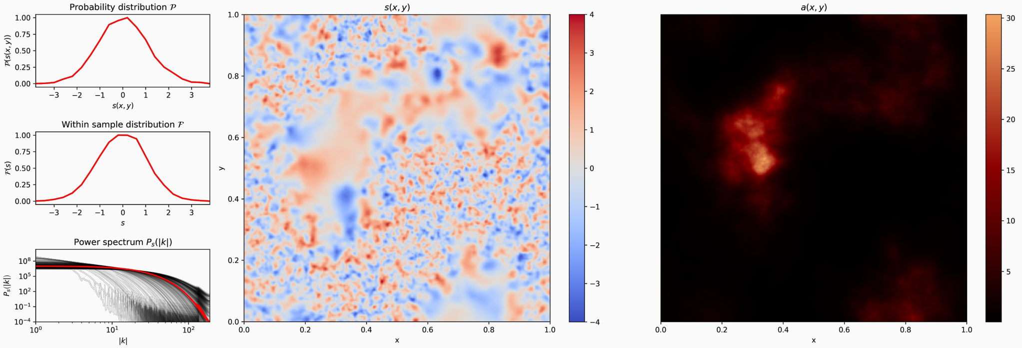




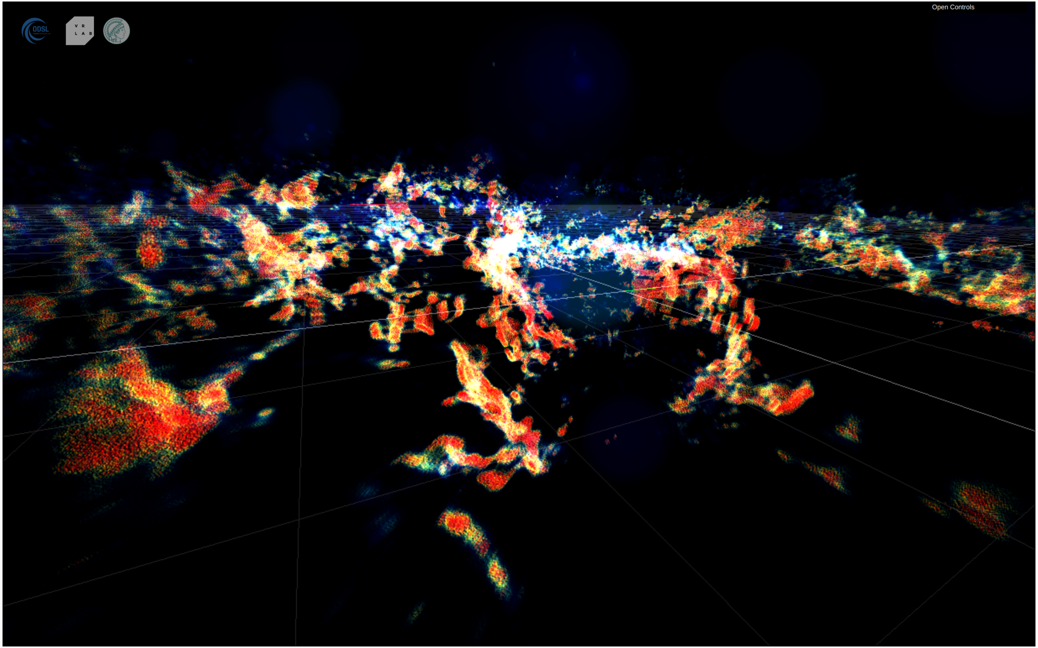


GP - Priors

$$s = A\,\xi, \quad \text{with} \quad A\left(\vec{x}, \vec{x'}\right) \propto {}^{1}\!/\!\left(1+\frac{1}{\sigma(a(\vec{x}))}\left|\vec{x}-\vec{x'}\right|^{2}\right)^{2}.$$

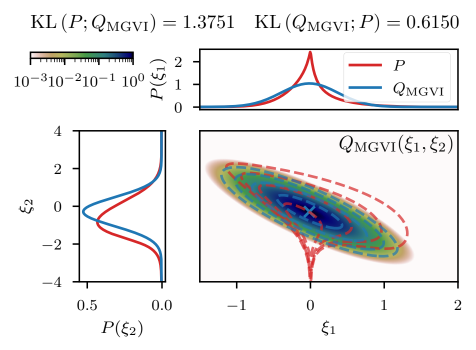


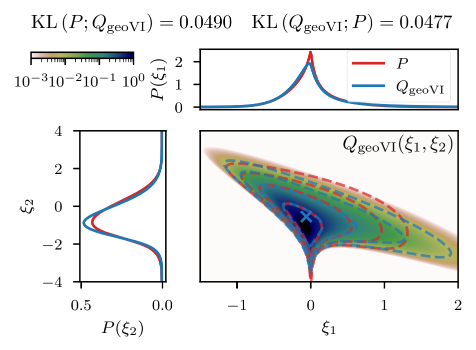
Dust tomography [LEK+22]



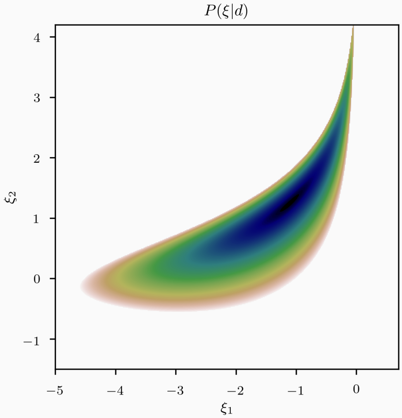
Variational Inference

Variational Inference (VI)



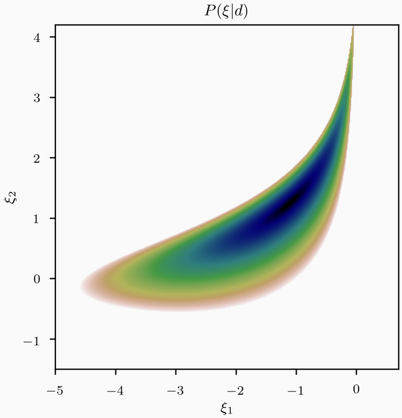


Geometric Variational Inference (geoVI) [FLE21]



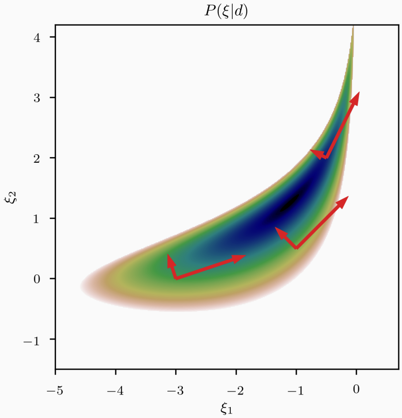
Information Hamiltonian $\mathcal{H}(\xi|d)$: $\quad -\log(\mathcal{P}(\xi|d))$

Geometric Variational Inference (geoVI) [FLE21]



Information Hamiltonian $\mathcal{H}(\xi\|d)$:	$-\log\left(\mathcal{P}(\xi\|d)\right)$
Posterior metric $\mathcal{M}(\xi)$:	$\mathcal{M}_{\mathrm{lh}}(\xi) + \mathbb{1}$
Fisher information metric $\mathcal{M}_{\mathrm{lh}}(\xi)$:	$\left\langle \frac{\partial^2 \mathcal{H}(d\|\xi)}{\partial\xi\partial\xi'} \right\rangle_{\mathcal{P}(d\|\xi)}$

Geometric Variational Inference (geoVI) [FLE21]

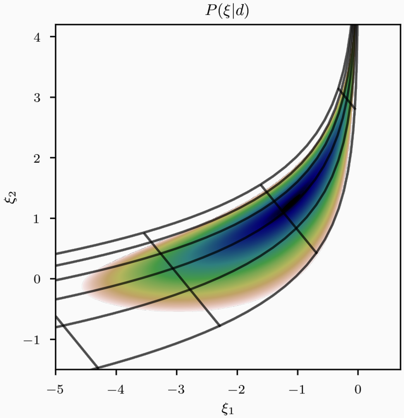


Information Hamiltonian $\mathcal{H}(\xi|d)$: $-\log\left(\mathcal{P}(\xi|d)\right)$

Posterior metric $\mathcal{M}(\xi)$: $\mathcal{M}_{\mathrm{lh}}(\xi) + \mathbb{1}$

Fisher information metric $\mathcal{M}_{\mathrm{lh}}(\xi)$: $\left\langle \frac{\partial^2 \mathcal{H}(d|\xi)}{\partial \xi \partial \xi'} \right\rangle_{\mathcal{P}(d|\xi)}$

Geometric Variational Inference (geoVI) [FLE21]

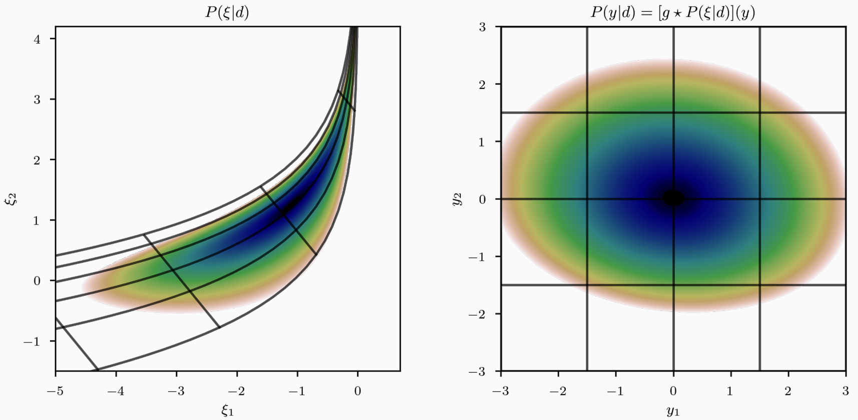


Information Hamiltonian $\mathcal{H}(\xi|d)$: $-\log\left(\mathcal{P}(\xi|d)\right)$

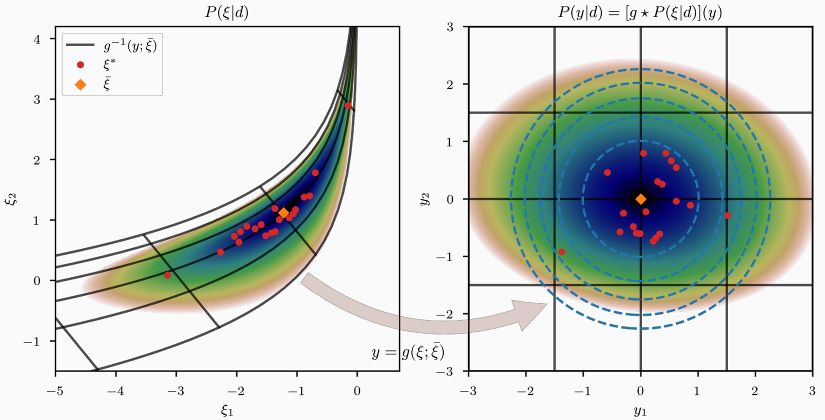
Posterior metric $\mathcal{M}(\xi)$: $\mathcal{M}_{\mathrm{lh}}(\xi) + \mathbb{1}$

Fisher information metric $\mathcal{M}_{\mathrm{lh}}(\xi)$: $\left\langle \frac{\partial^2 \mathcal{H}(d|\xi)}{\partial \xi \partial \xi'} \right\rangle_{\mathcal{P}(d|\xi)}$

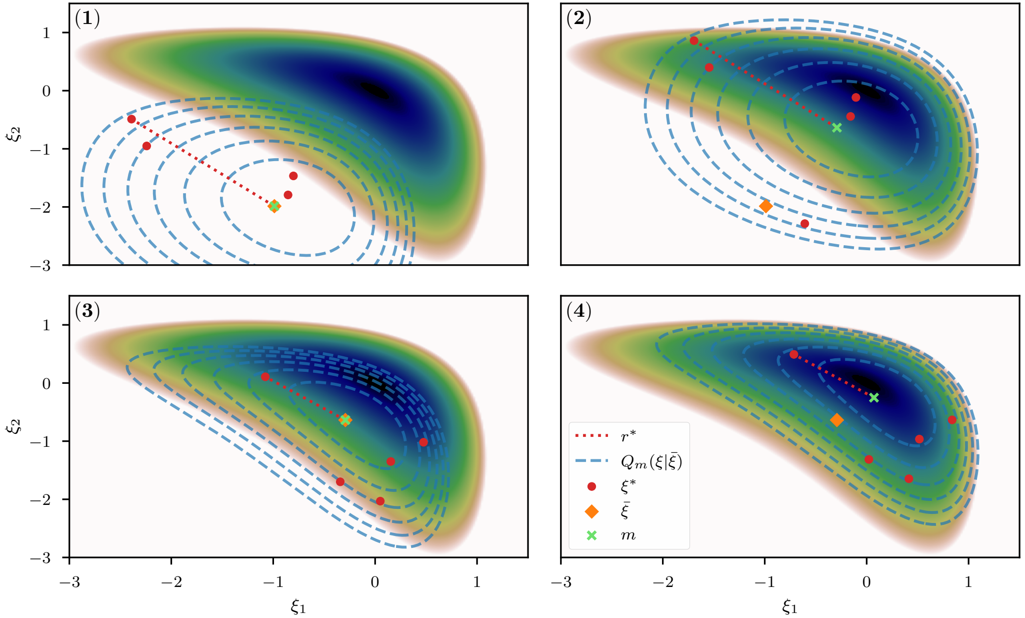
Geometric Variational Inference (geoVI) [FLE21]



Geometric Variational Inference (geoVI) [FLE21]



Geometric Variational Inference (geoVI) [FLE21]



Automatic differentiation

Automatic differentiation

- Numpy based 1st order AD (nifty8) $\rightarrow$ Jax based AD in jifty

Automatic differentiation

- Numpy based 1st order AD (nifty8) $\rightarrow$ Jax based AD in jifty
- NIFTy VI requires both vector-jacobian (vjp) and jacobian-vector (jvp) products
- Posterior metric of the form: $\mathcal{M}(\xi) = \frac{\partial x}{\partial \xi} \left(\frac{\partial x}{\partial \xi'} \right)^{\dagger} \Big|_{\bar{\xi}} + \mathbb{1}$

Automatic differentiation

- Numpy based 1st order AD (nifty8) $\rightarrow$ Jax based AD in jifty
- NIFTy VI requires both vector-jacobian (vjp) and jacobian-vector (jvp) products
- Posterior metric of the form: $\mathcal{M}(\xi) = \frac{\partial x}{\partial \xi} \left(\frac{\partial x}{\partial \xi'} \right)^{\dagger} \Big|_{\bar{\xi}} + \mathbb{1}$
- During Inference & Sampling: Approximate solutions to $\mathcal{M}x = y$
  $\rightarrow$ Many applications of jvp + vjp with same primals $\bar{\xi}$ but different tangents

Conclusion

Numerical Information Field Theory (NIFTy)

NIFTy

Code: https://gitlab.mpcdf.mpg.de/ift/nifty

Docs: https://ift.pages.mpcdf.de/nifty

Contributors

Andrija Kostic, David Outland, Gordian Edenhofer, Jakob Knollmüller, Jakob Roth, Lukas Platz, Margret Westerkamp, Martin Reinecke, Massin Guerdi, Matteo Guardiani, Philipp Arras, Philipp Frank, Reimar Heinrich Leike, Torsten Enßlin, Vincent Eberle
& the entire IFT-Group at MPA

Philipp Arras, Gordian Edenhofer, Philipp Frank, Andrija Kostic, Jakob Knollmüller, Jakob Roth, Lukas Platz, Matteo Guardiani, Martin Reinecke, Reimar Heinrich Leike, Simon Ding, and Vincent Eberle.
NIFTy: Numerical information field theory, 2022.

Philipp Arras, Philipp Frank, Philipp Haim, Jakob Knollmüller, Reimar Leike, Martin Reinecke, and Torsten A. Enßlin.
Variable structures in m87* from space, time and frequency resolved interferometry.
Nature Astronomy, 6(2):259–269, 2022.

Philipp Frank, Reimar Leike, and Torsten A. Enßlin.
Geometric variational inference.
Entropy, 23(7), 2021.

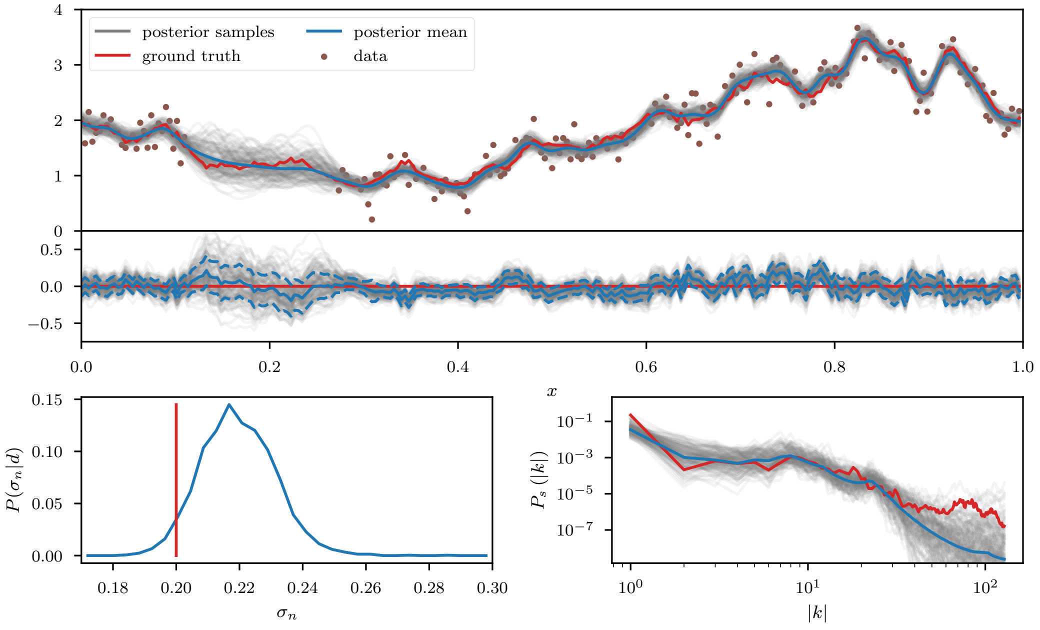
Sebastian Hutschenreuter, Marijke Haverkorn, Philipp Frank, Nergis C. Raycheva, and Torsten A. Enßlin.
Disentangling the faraday rotation sky, 2023.

Reimar Leike, Gordian Edenhofer, Jakob Knollmüller, Christian Alig, Philipp Frank, and Torsten A. Enßlin.
The galactic 3d large-scale dust distribution via gaussian process regression on spherical coordinates.
arXiv, 2204.11715, 2022.

Leike, Reimar H., Glatze, Martin, and Enßlin, Torsten A.
Resolving nearby dust clouds.
A&A, 639:A138, 2020.

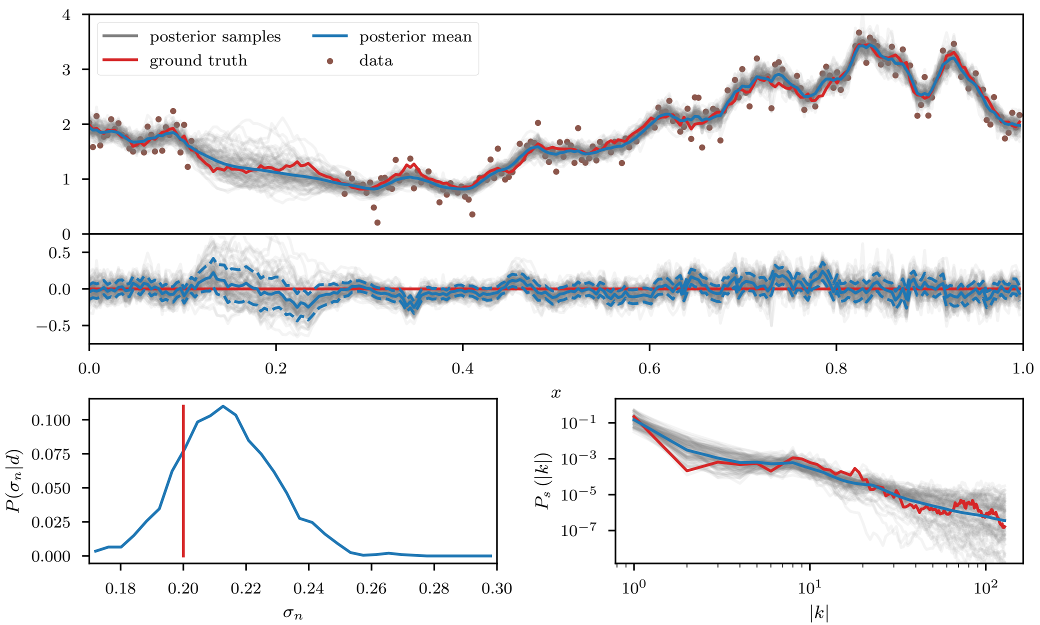
Lukas I. Platz, Jakob Knollmüller, Philipp Arras, Philipp Frank, Martin Reinecke, Dominik Jüstel, and Torsten A. Enßlin.
Multi-component imaging of the fermi gamma-ray sky in the spatio-spectral domain, 2022.

Appendix



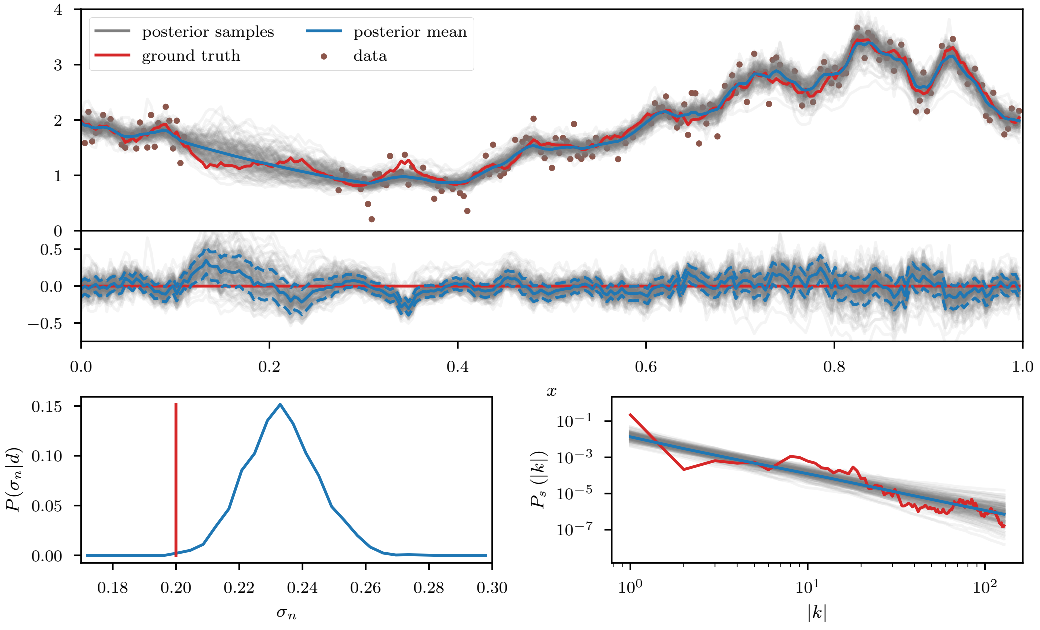
posterior samples
ground truth
posterior mean
data
4
3
2
1
0
0.5
0.0
−0.5
0.0
0.2
0.4
0.6
0.8
1.0
x
0.15
0.10
0.05
0.00
$P(\sigma_n|d)$
0.18
0.20
0.22
0.24
0.26
0.28
0.30
σ_n
10^{-1}
10^{-3}
10^{-5}
10^{-7}
$P_s(|k|)$
10^0
10^1
10^2
$|k|$

Appendix



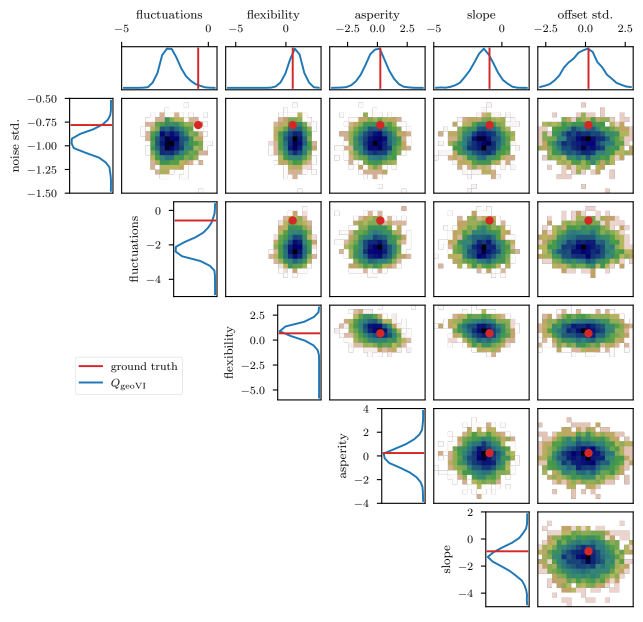
posterior samples
posterior mean
ground truth
data
0
1
2
3
4
0.5
0.0
−0.5
0.0
0.2
0.4
0.6
0.8
1.0
x
$P(\sigma_n|d)$
0.000
0.025
0.050
0.075
0.100
0.18
0.20
0.22
0.24
0.26
0.28
0.30
σ_n
$P_s(|k|)$
10^{-1}
10^{-3}
10^{-5}
10^{-7}
10^{0}
10^{1}
10^{2}
$|k|$

Appendix



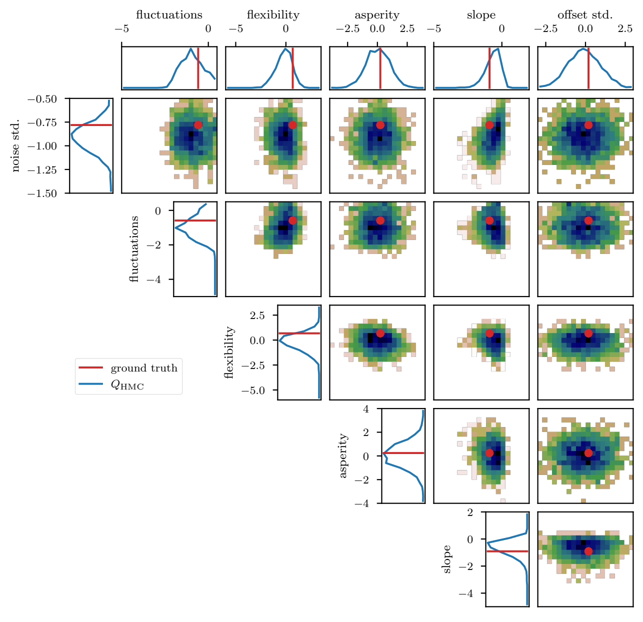
posterior samples
ground truth
posterior mean
data
0.0
0.2
0.4
0.6
0.8
1.0
x
$P(\sigma_n|d)$
σ_n
$P_s(|k|)$
$|k|$

Appendix



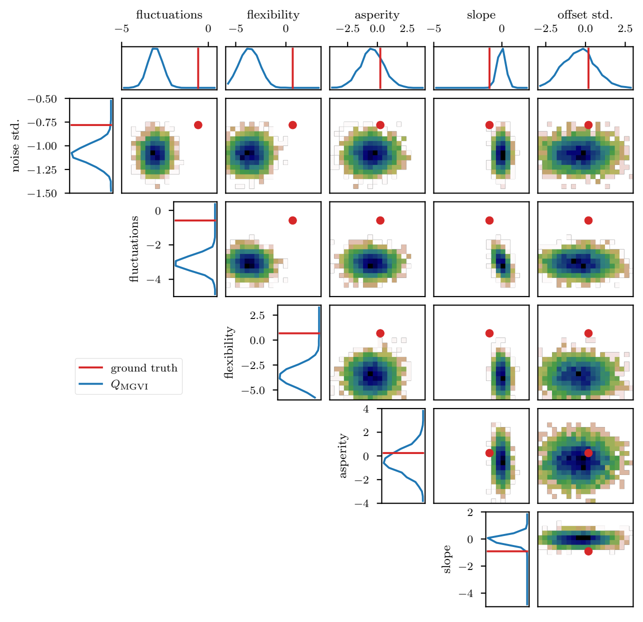
fluctuations
flexibility
asperity
slope
offset std.
noise std.
fluctuations
flexibility
asperity
slope
ground truth
Q_{geoVI}

Appendix

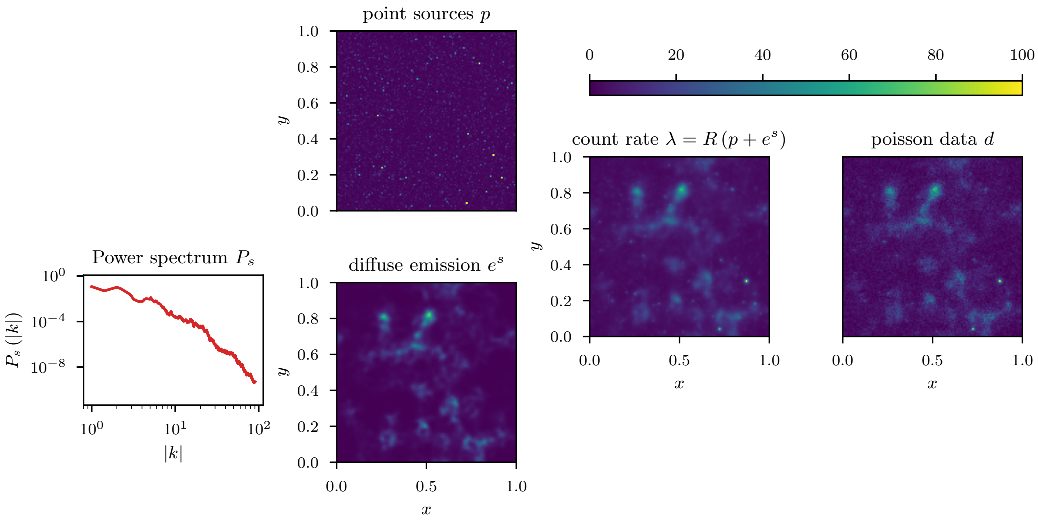


fluctuations
flexibility
asperity
slope
offset std.
noise std.
ground truth
Q_{HMC}

Appendix

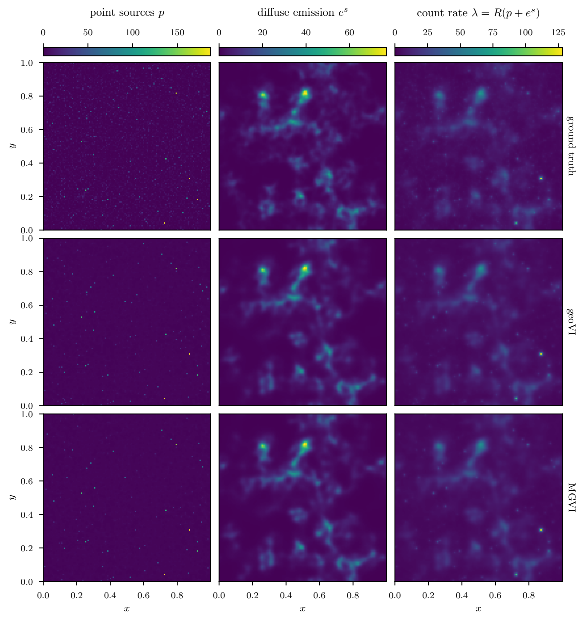

fluctuations
flexibility
asperity
slope
offset std.
noise std.
ground truth
Q_{MGVI}

Appendix



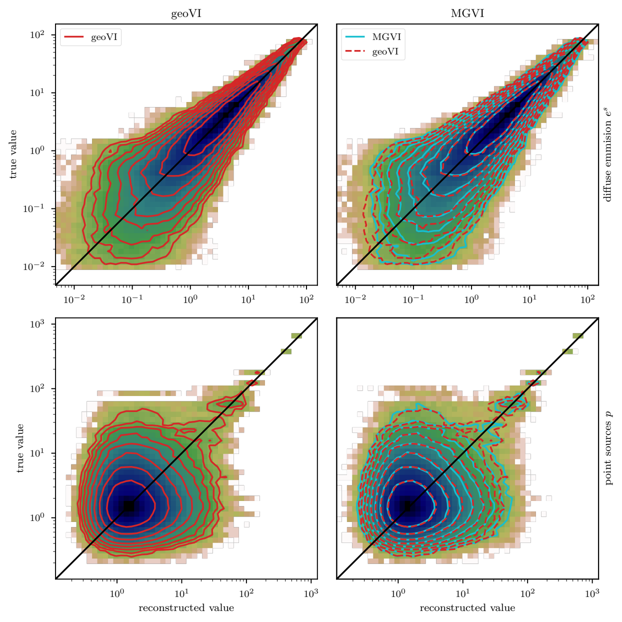
Power spectrum P_s
10^0
10^{-4}
10^{-8}
$P_s(|k|)$
10^0
10^1
10^2
$|k|$
point sources p
1.0
0.8
0.6
0.4
0.2
0.0
y
diffuse emission e^s
1.0
0.8
0.6
0.4
0.2
0.0
y
0.0
0.5
1.0
x
0
20
40
60
80
100
count rate $\lambda = R(p + e^s)$
1.0
0.8
0.6
0.4
0.2
0.0
y
0.0
0.5
1.0
x
poisson data d
0.0
0.5
1.0
x

Appendix



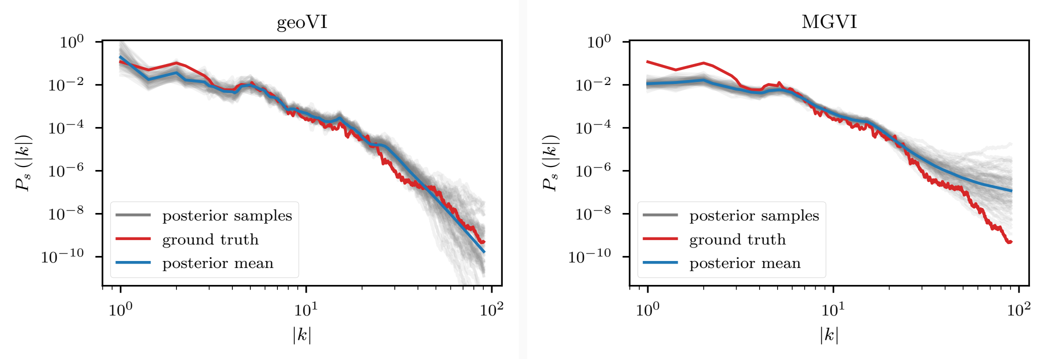
point sources p
diffuse emission e^s
count rate $\lambda = R(p + e^s)$
ground truth
geoVI
MGVI
x
y

Appendix



geoVI
MGVI
geoVI
MGVI
geoVI
true value
true value
diffuse emmision e^s
point sources p
reconstructed value
reconstructed value

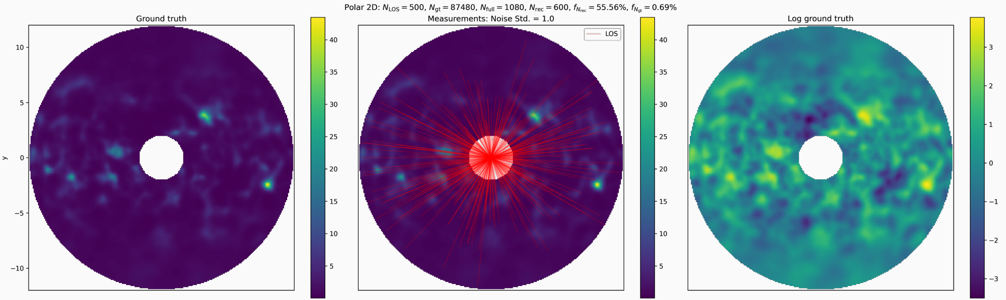
Appendix



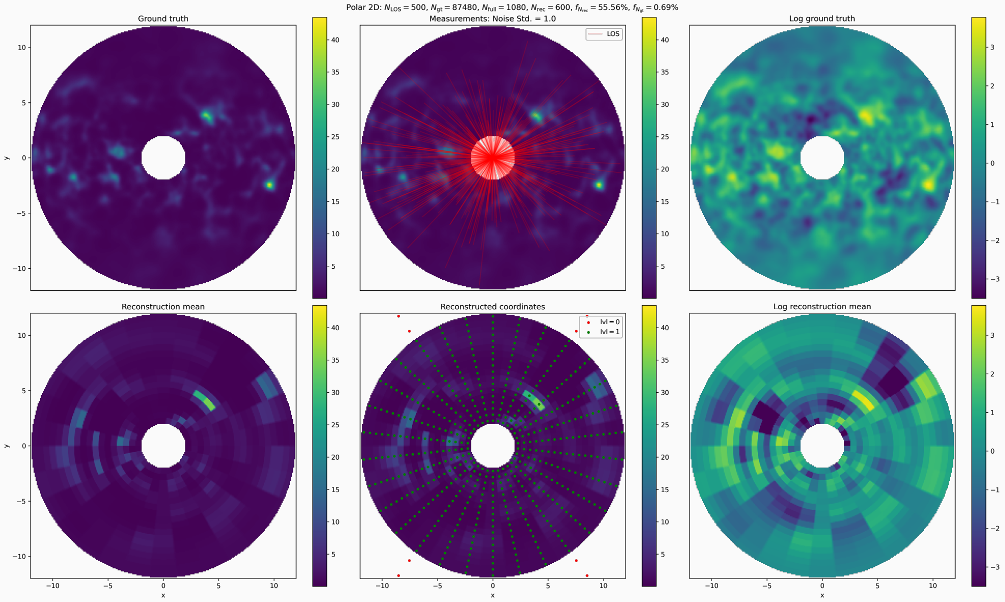
geoVI
MGVI
$P_s(|k|)$
$|k|$
10^0
10^{-2}
10^{-4}
10^{-6}
10^{-8}
10^{-10}
10^0
10^1
10^2
posterior samples
ground truth
posterior mean

Adaptive Resolution - Outlook

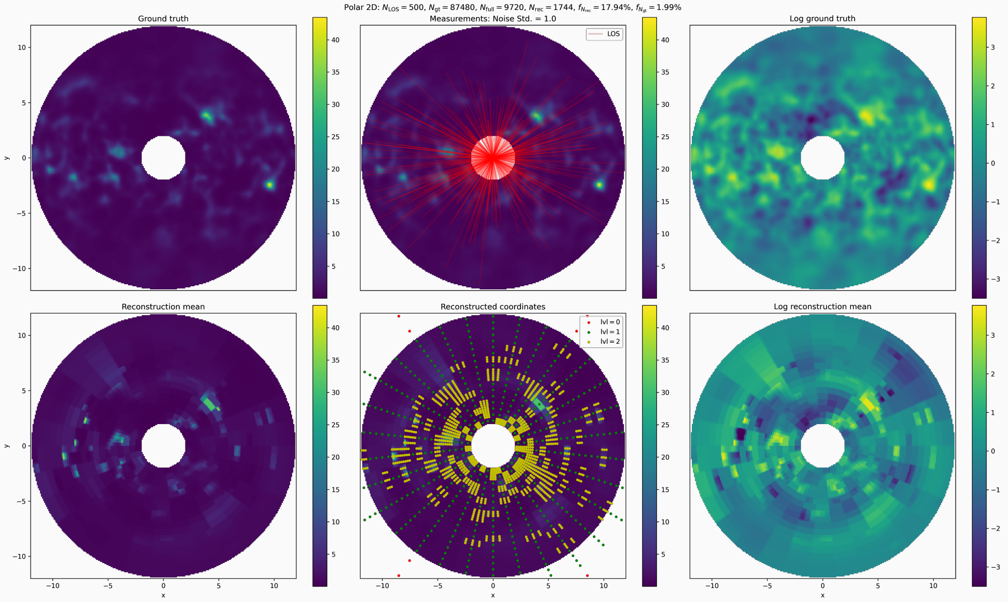
Adaptive Resolution - Outlook (Preliminary)



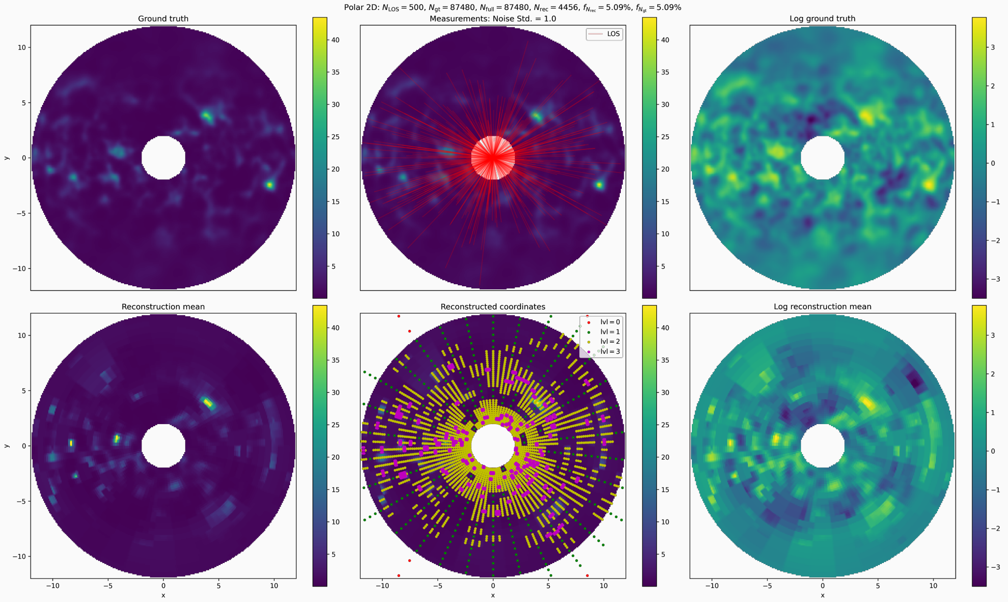
Adaptive Resolution - Outlook (Preliminary)



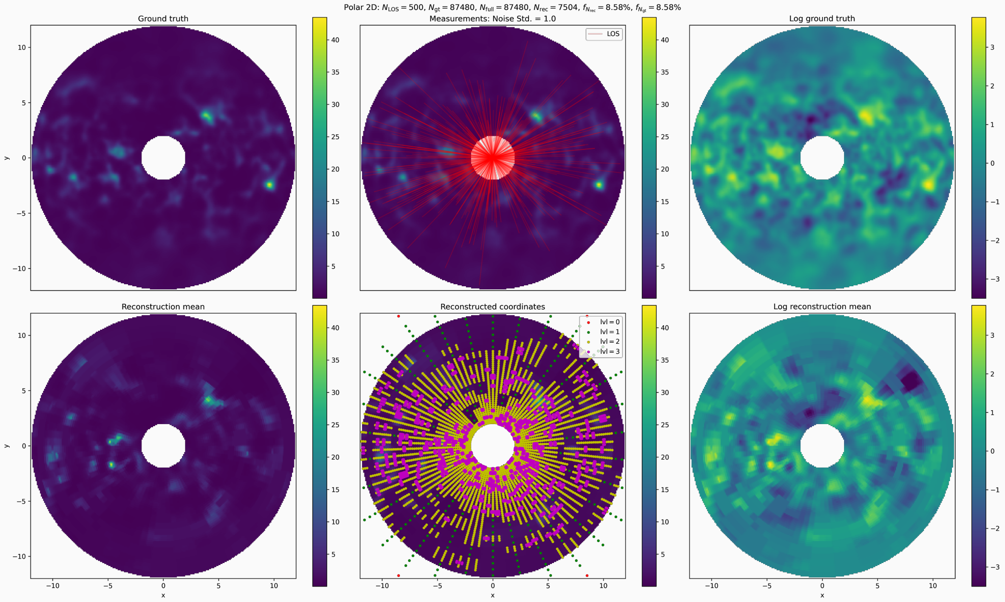
Adaptive Resolution - Outlook (Preliminary)



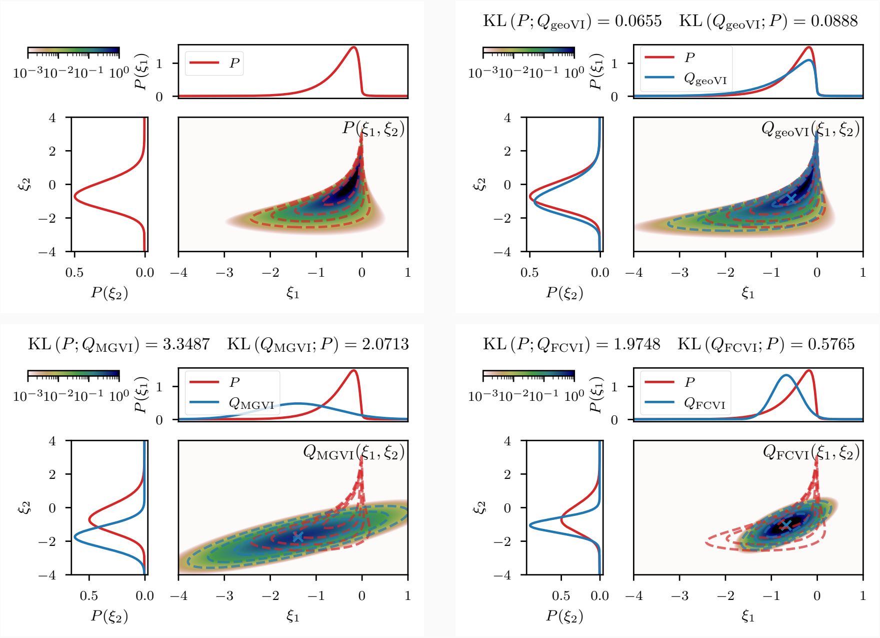
Adaptive Resolution - Outlook (Preliminary)



Adaptive Resolution - Outlook (Preliminary)

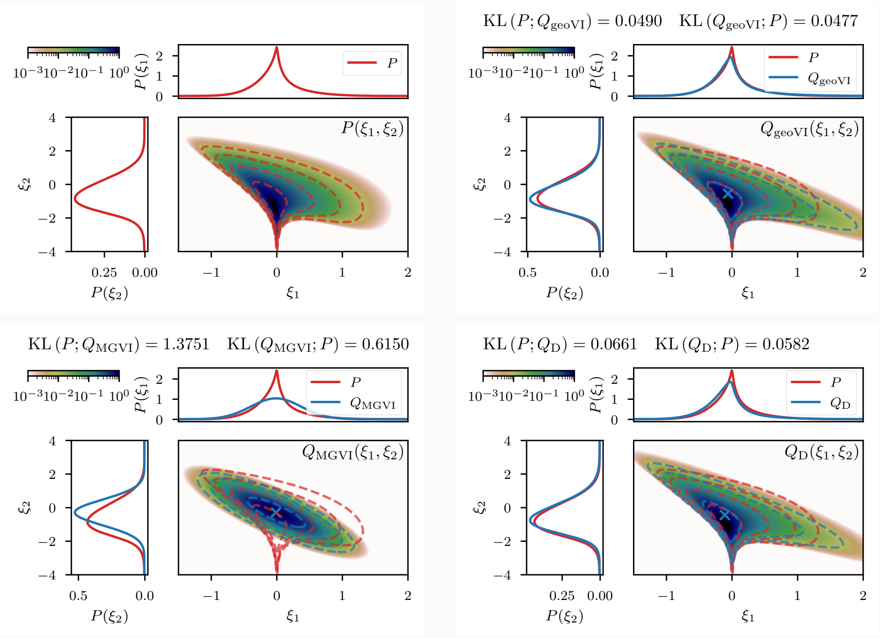


Appendix



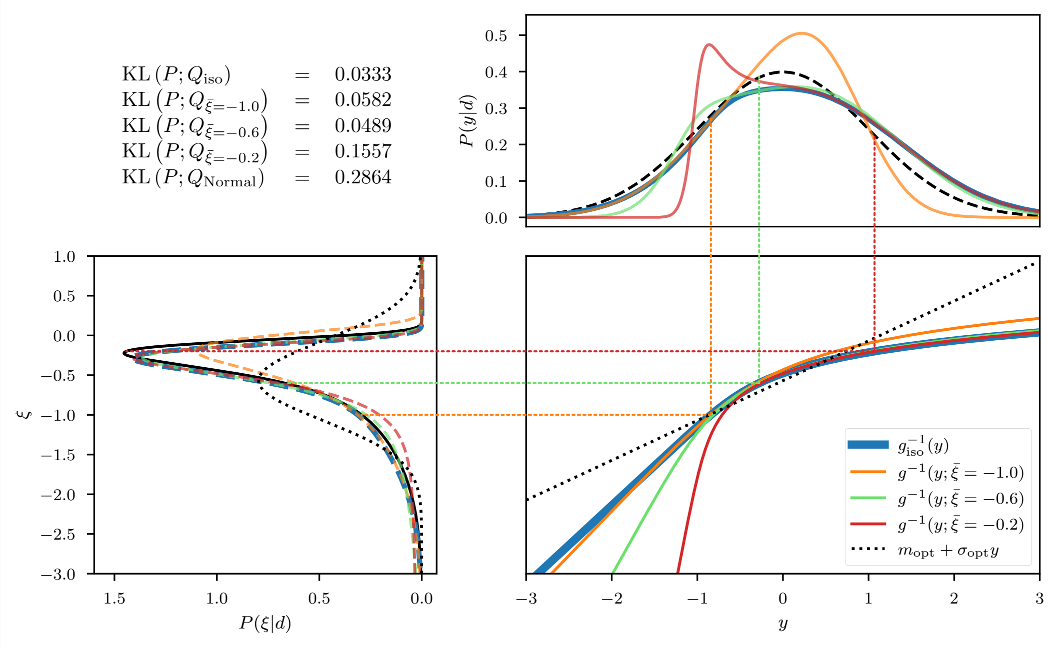
P
$P(\xi_1)$
$P(\xi_1, \xi_2)$
ξ_2
$P(\xi_2)$
ξ_1
$10^{-3}\ 10^{-2}\ 10^{-1}\ 10^{0}$
KL $(P; Q_{\mathrm{geoVI}}) = 0.0655$ KL $(Q_{\mathrm{geoVI}}; P) = 0.0888$
P
Q_{geoVI}
$Q_{\mathrm{geoVI}}(\xi_1, \xi_2)$
KL $(P; Q_{\mathrm{MGVI}}) = 3.3487$ KL $(Q_{\mathrm{MGVI}}; P) = 2.0713$
P
Q_{MGVI}
$Q_{\mathrm{MGVI}}(\xi_1, \xi_2)$
KL $(P; Q_{\mathrm{FCVI}}) = 1.9748$ KL $(Q_{\mathrm{FCVI}}; P) = 0.5765$
P
Q_{FCVI}
$Q_{\mathrm{FCVI}}(\xi_1, \xi_2)$

Appendix



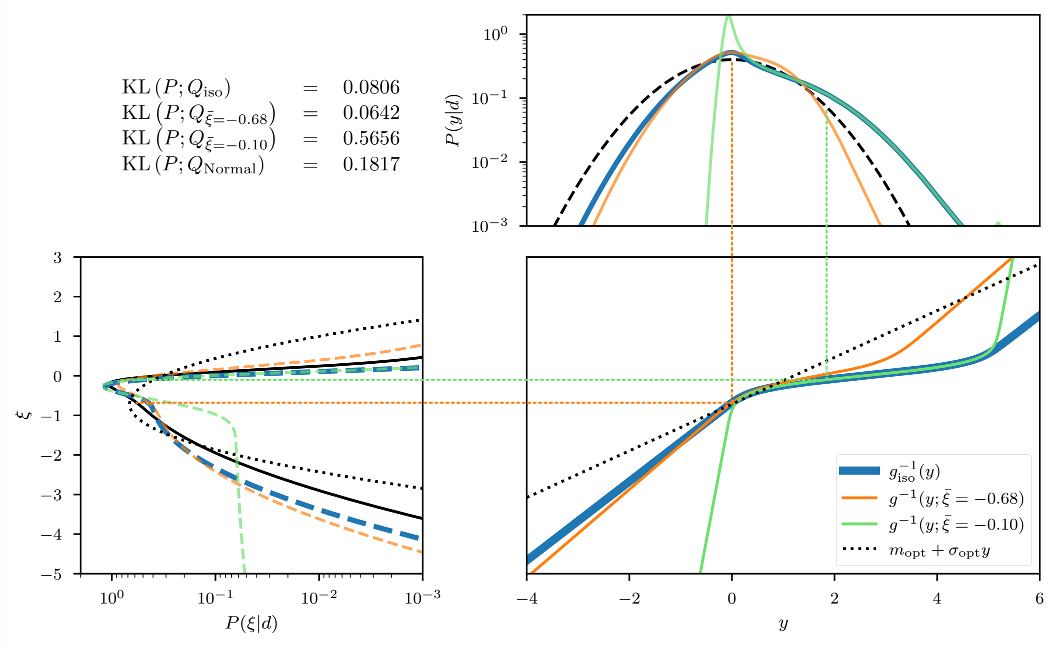
KL $(P; Q_{\mathrm{geoVI}}) = 0.0490$ KL $(Q_{\mathrm{geoVI}}; P) = 0.0477$
KL $(P; Q_{\mathrm{MGVI}}) = 1.3751$ KL $(Q_{\mathrm{MGVI}}; P) = 0.6150$
KL $(P; Q_{\mathrm{D}}) = 0.0661$ KL $(Q_{\mathrm{D}}; P) = 0.0582$
$P(\xi_1, \xi_2)$
$Q_{\mathrm{geoVI}}(\xi_1, \xi_2)$
$Q_{\mathrm{MGVI}}(\xi_1, \xi_2)$
$Q_{\mathrm{D}}(\xi_1, \xi_2)$
$P(\xi_1)$
$P(\xi_2)$
ξ_1
ξ_2

Appendix



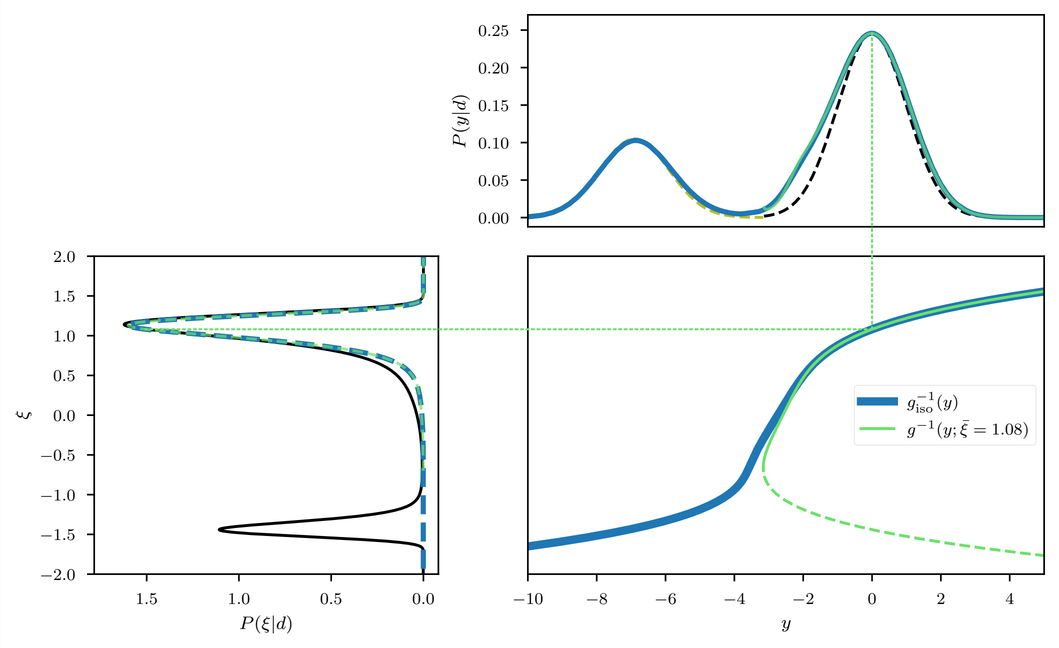
KL (P; Q_iso) = 0.0333
KL (P; Q_ξ̄=−1.0) = 0.0582
KL (P; Q_ξ̄=−0.6) = 0.0489
KL (P; Q_ξ̄=−0.2) = 0.1557
KL (P; Q_Normal) = 0.2864
P(y|d)
0.5
0.4
0.3
0.2
0.1
0.0
1.0
0.5
0.0
−0.5
ξ
−1.0
−1.5
−2.0
−2.5
−3.0
1.5
1.0
0.5
0.0
P(ξ|d)
−3
−2
−1
0
1
2
3
y
g_iso^{−1}(y)
g^{−1}(y; ξ̄ = −1.0)
g^{−1}(y; ξ̄ = −0.6)
g^{−1}(y; ξ̄ = −0.2)
m_opt + σ_opt y

Appendix



$\mathrm{KL}\left(P;Q_{\mathrm{iso}}\right) = 0.0806$
$\mathrm{KL}\left(P;Q_{\bar{\xi}=-0.68}\right) = 0.0642$
$\mathrm{KL}\left(P;Q_{\bar{\xi}=-0.10}\right) = 0.5656$
$\mathrm{KL}\left(P;Q_{\mathrm{Normal}}\right) = 0.1817$
$P(y|d)$
10^{0}
10^{-1}
10^{-2}
10^{-3}
ξ
3
2
1
0
−1
−2
−3
−4
−5
$P(\xi|d)$
y
−4
−2
0
2
4
6
$g_{\mathrm{iso}}^{-1}(y)$
$g^{-1}(y;\bar{\xi}=-0.68)$
$g^{-1}(y;\bar{\xi}=-0.10)$
$m_{\mathrm{opt}}+\sigma_{\mathrm{opt}}y$

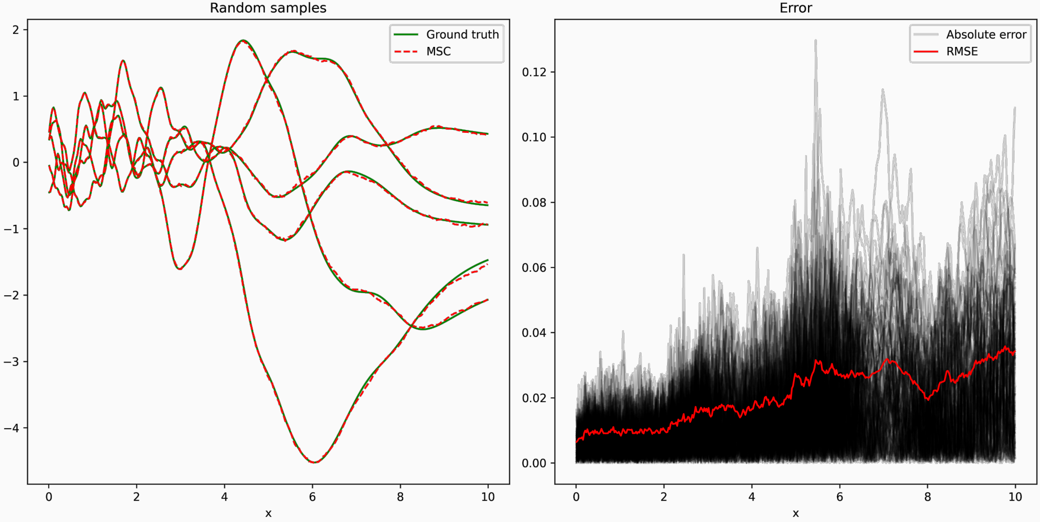
Appendix



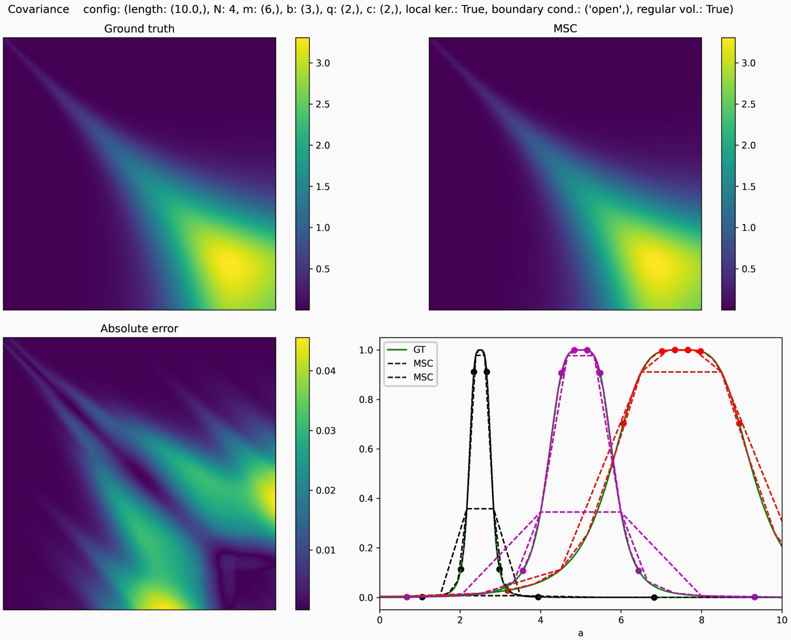
0.25
0.20
0.15
0.10
0.05
0.00
$P(y|d)$
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
−1.5
−2.0
ξ
1.5
1.0
0.5
0.0
$P(\xi|d)$
$g_{\text{iso}}^{-1}(y)$
$g^{-1}(y; \bar{\xi} = 1.08)$
−10
−8
−6
−4
−2
0
2
4
y

Appendix

Samples config: (length: (10.0,), N: 4, m: (6,), b: (3,), q: (2,), c: (2,), local ker.: True, boundary cond.: ('open',), regular vol.: True)



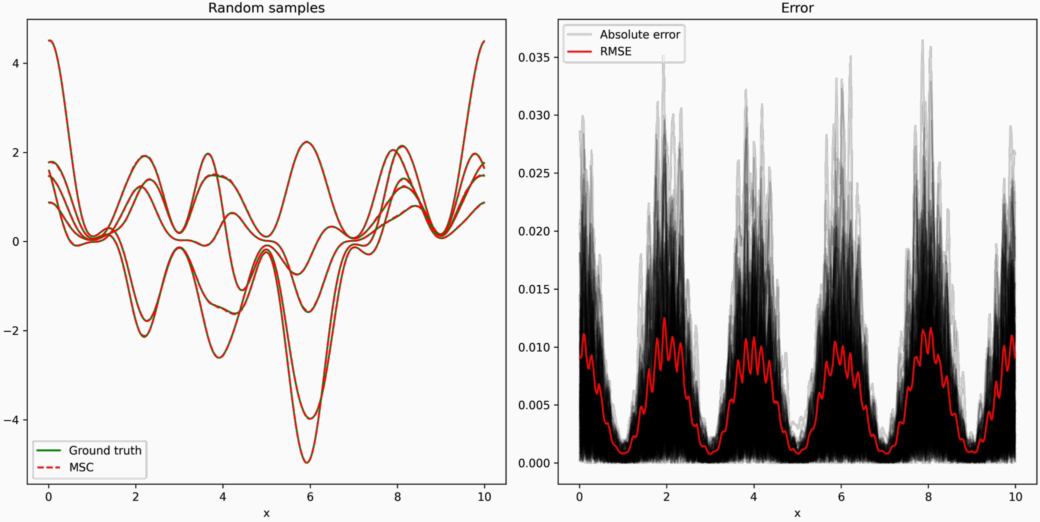
Appendix



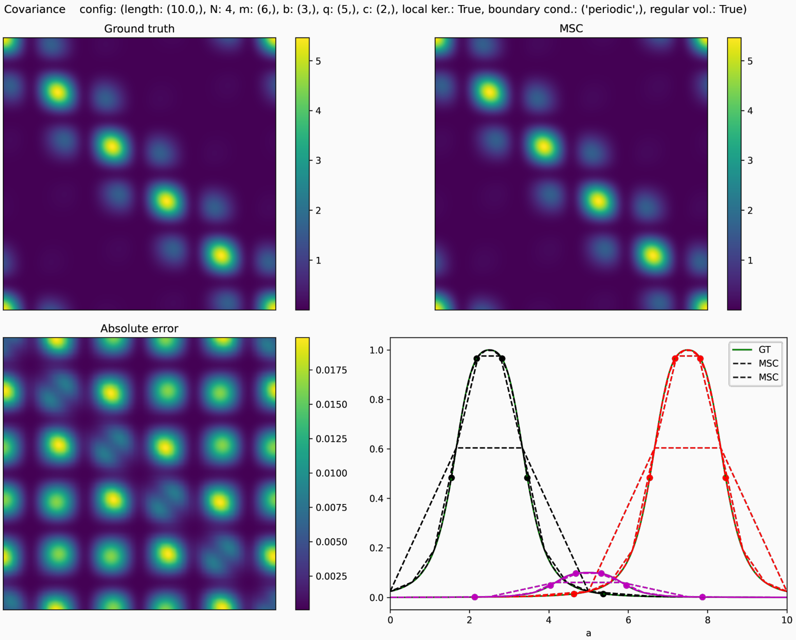
Covariance config: (length: (10.0,), N: 4, m: (6,), b: (3,), q: (2,), c: (2,), local ker.: True, boundary cond.: ('open',), regular vol.: True)
Ground truth
MSC
Absolute error
3.0
2.5
2.0
1.5
1.0
0.5
0.04
0.03
0.02
0.01
GT
MSC
MSC
1.0
0.8
0.6
0.4
0.2
0.0
0
2
4
6
8
10
a

Appendix

Samples   config: (length: (10.0,), N: 4, m: (6,), b: (3,), q: (5,), c: (2,), local ker.: True, boundary cond.: ('periodic',), regular vol.: True)

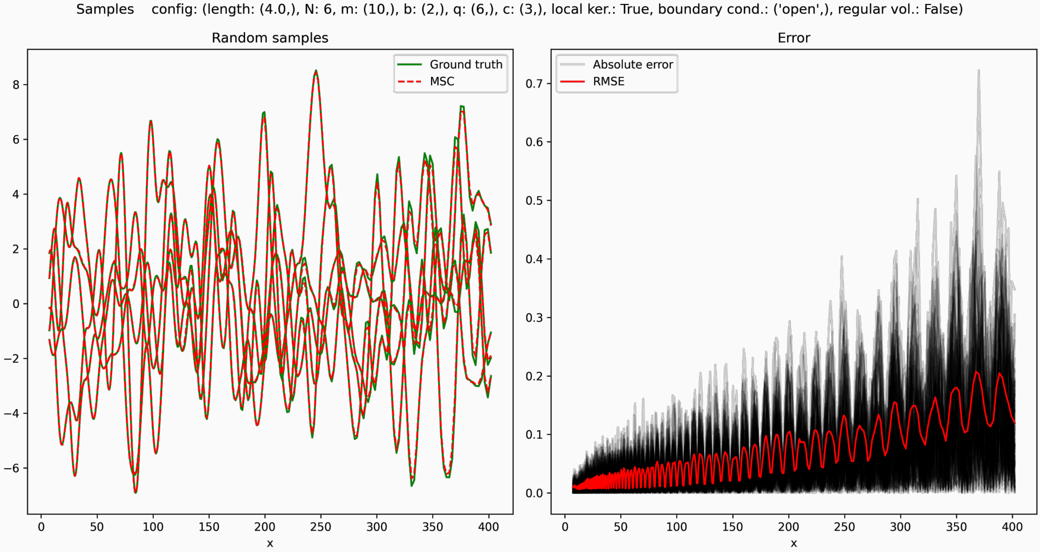


Appendix



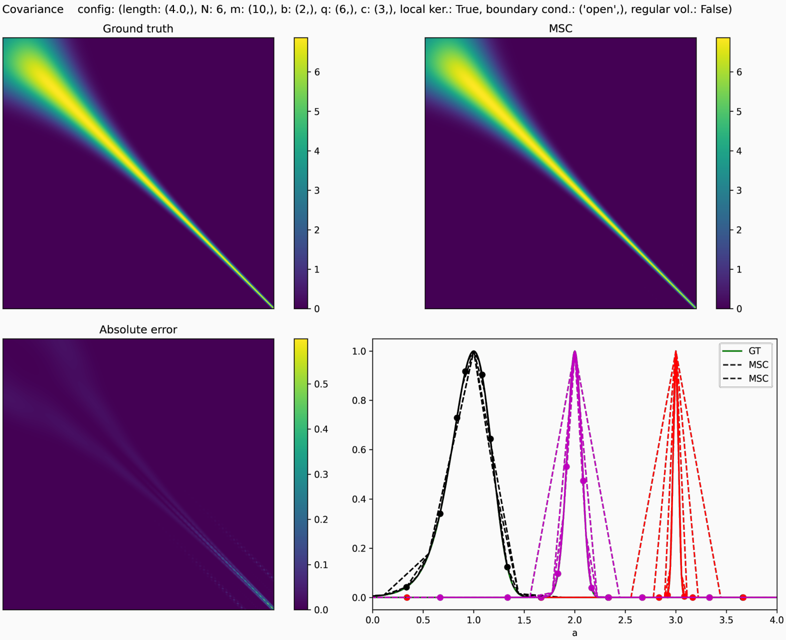
Covariance config: (length: (10.0,), N: 4, m: (6,), b: (3,), q: (5,), c: (2,), local ker.: True, boundary cond.: ('periodic',), regular vol.: True)
Ground truth
MSC
Absolute error
GT
MSC
MSC
a

Appendix



Samples config: (length: (4.0,), N: 6, m: (10,), b: (2,), q: (6,), c: (3,), local ker.: True, boundary cond.: ('open',), regular vol.: False)
Random samples
Ground truth
MSC
x
Error
Absolute error
RMSE
x

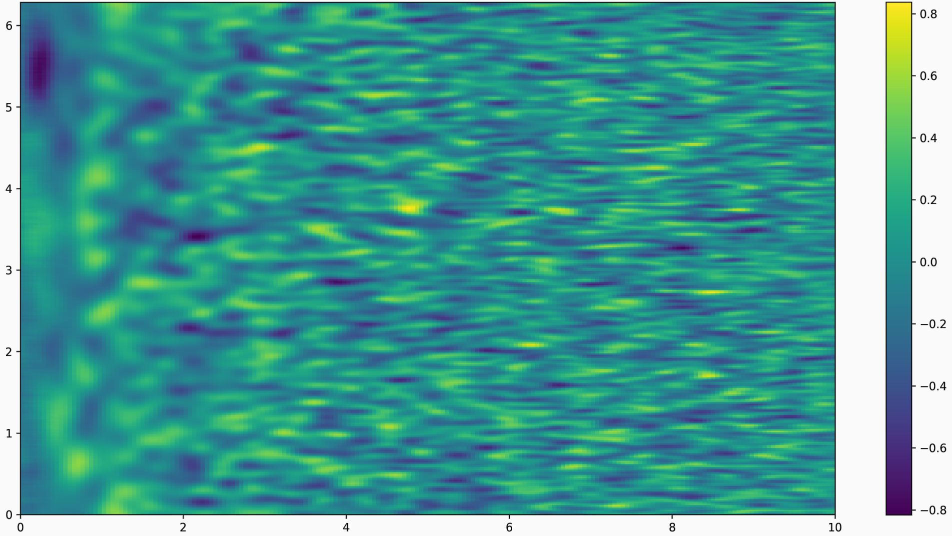
Appendix



Covariance config: (length: (4.0,), N: 6, m: (10,), b: (2,), q: (6,), c: (3,), local ker.: True, boundary cond.: ('open',), regular vol.: False)
Ground truth
MSC
Absolute error
GT
MSC
MSC
a

Appendix

Local 2D config: (length: (10.0, 6.28), N: 3, m: (8, 12), b: (3, 3), q: (3, 3), c: (1, 2), local ker.: True, boundary cond.: ('open', 'periodic'), regular vol.: False)



Appendix

Comp. 2D config: (length: (10.0, 6.28), N: 3, m: (8, 12), b: (3, 3), q: (3, 3), c: (1, 2), local ker.: True, boundary cond.: ('open', 'periodic'), regular vol.: False)