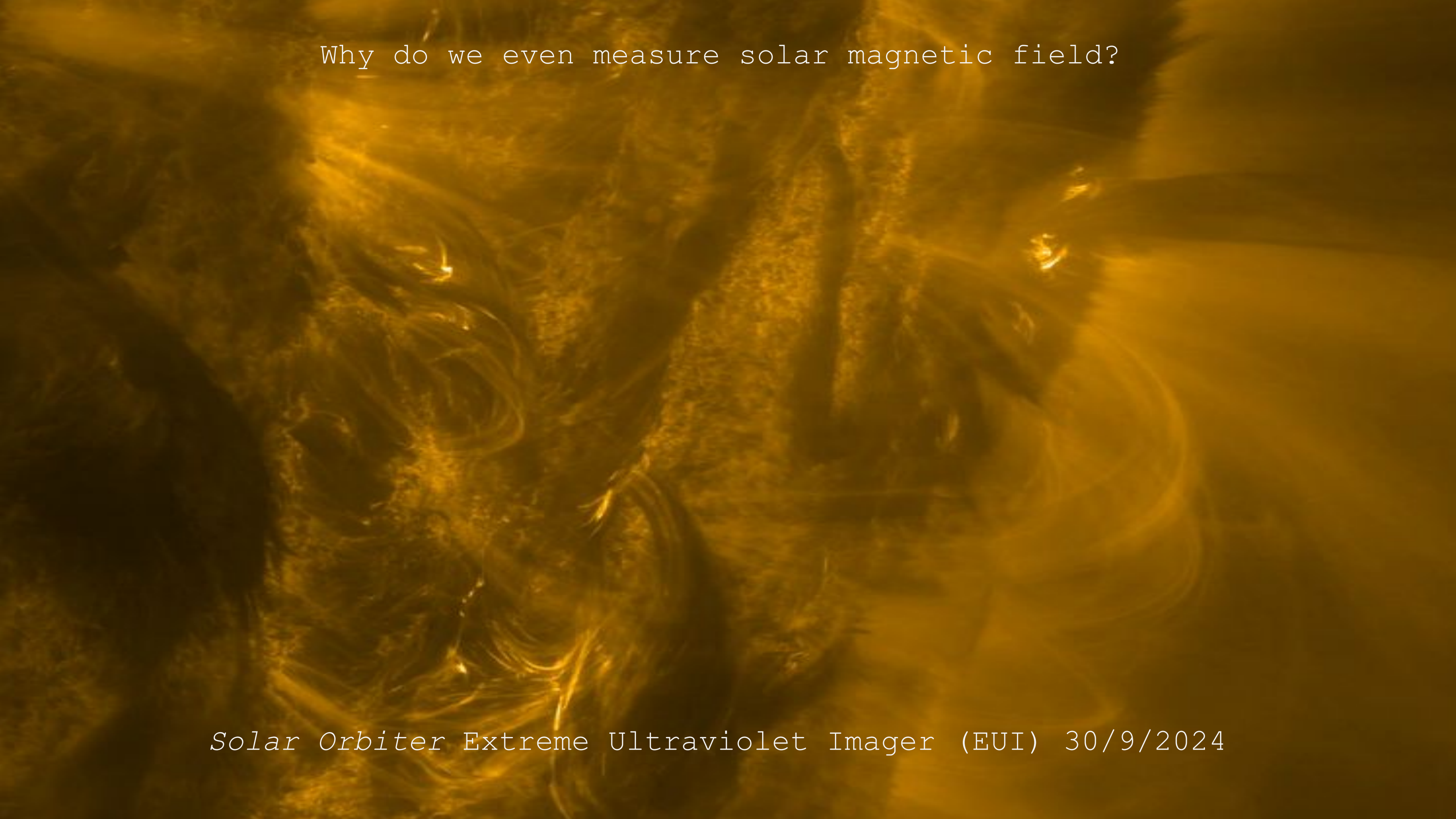
Application of machine learning in spectropolarimetric diagnostics of the Solar magnetic field

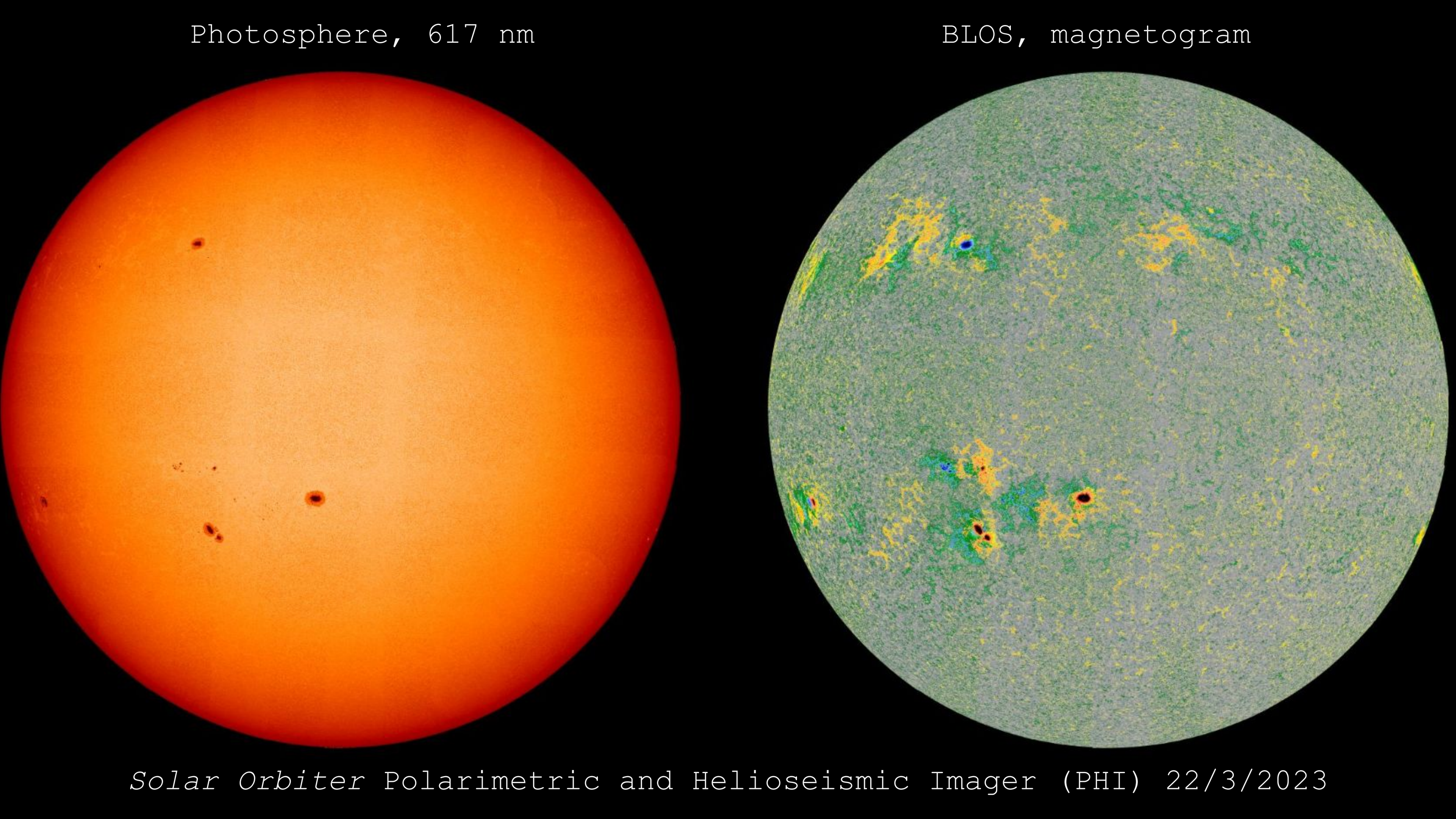
Seminar of the Department of Astronomy, 12th of May 2026

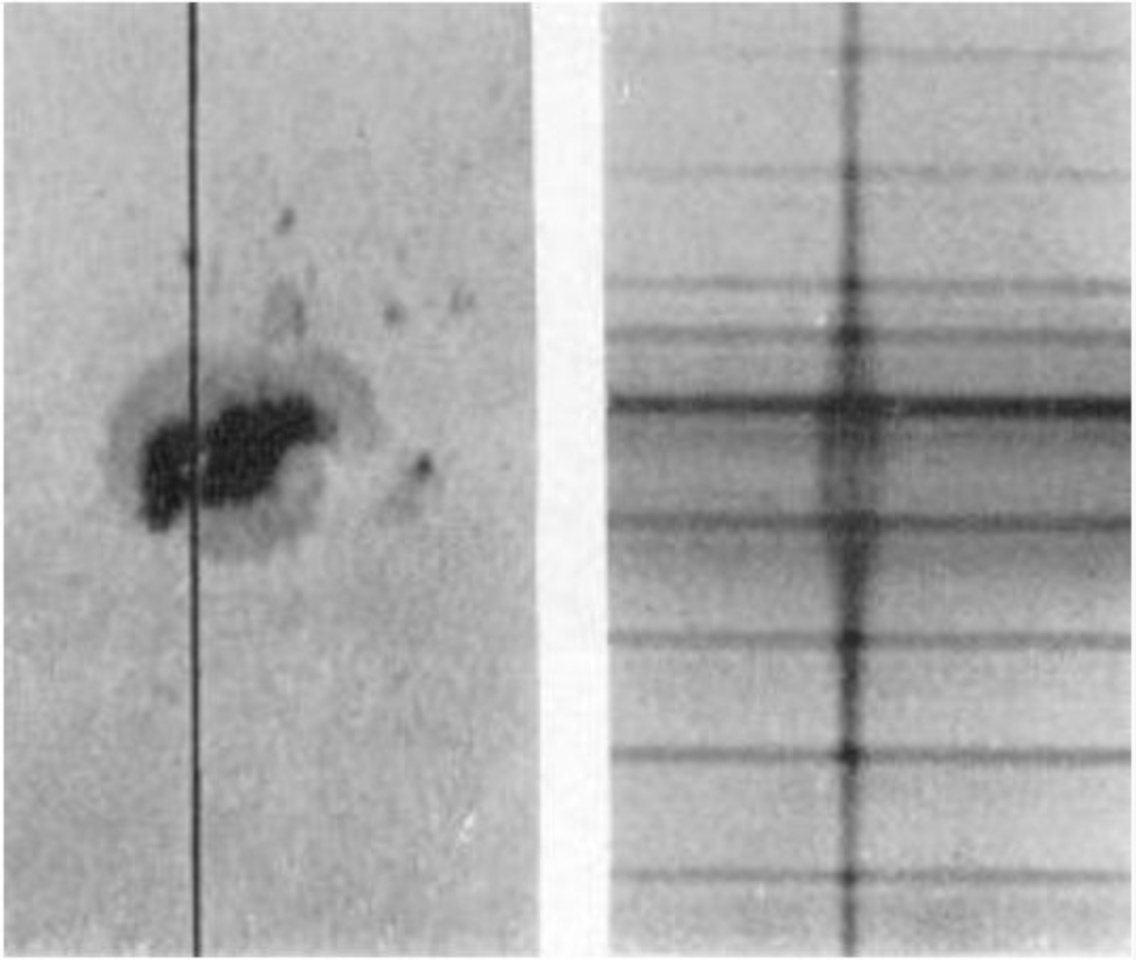
Phd student, Đorđe Mijović

Why do we even measure solar magnetic field?

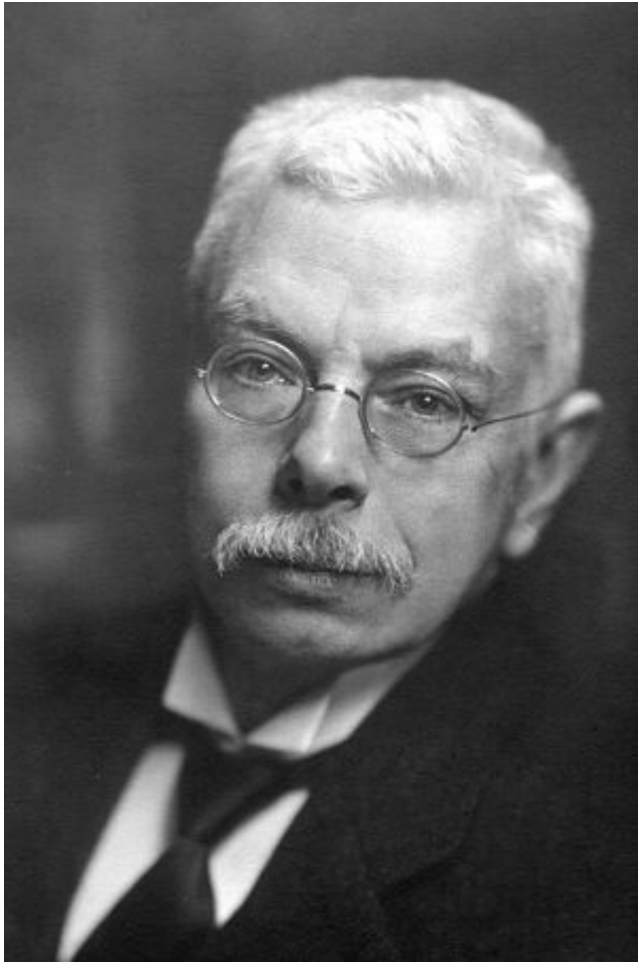


Solar Orbiter Extreme Ultraviolet Imager (EUI) 30/9/2024


Photosphere, 617 nm
BLOS, magnetogram
Solar Orbiter Polarimetric and Helioseismic Imager (PHI) 22/3/2023



G.E. Hale et al.
(1919)

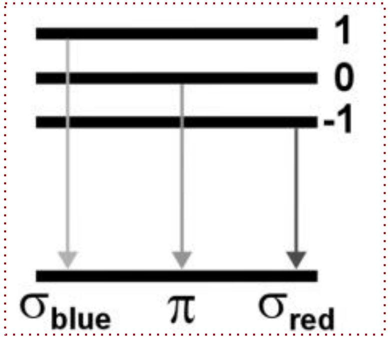


Pieter Zeeman (NP in 1902)

Energy levels splitting in magnetic field

$$J_u = 1,\ m_u = -1, 0, 1$$

$$J_u = 0,\ m_u = 0$$



$$\Delta E = \mu_B g_J m_J B, \quad \mu_B = \frac{e\hbar}{2m_e}$$

Transitions between Zeeman sublevels in simple case of spectral line (adapted from Trujilo Bueno, 2006)

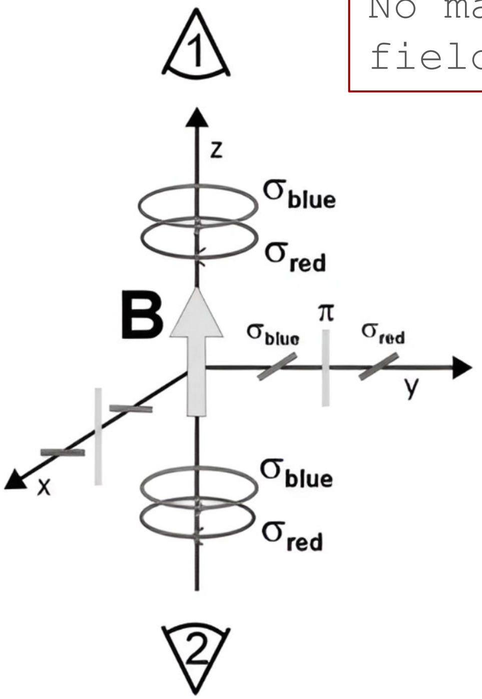
example:

$$\Delta\lambda \propto 10^{-13}\lambda^2 B$$

Fe I 6302.5Å

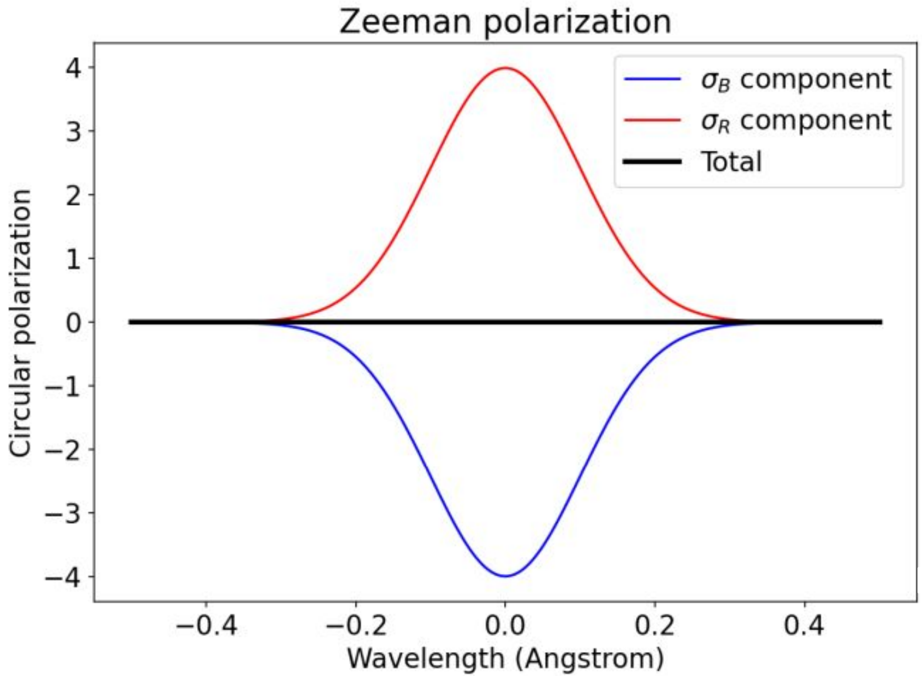
$B \sim 1000\,\mathrm{G}$

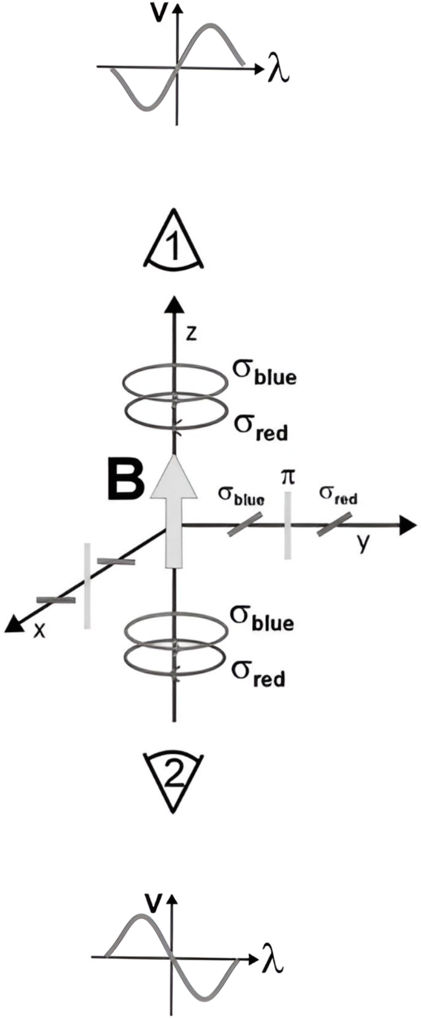
$\Delta\lambda = 0.046$Å

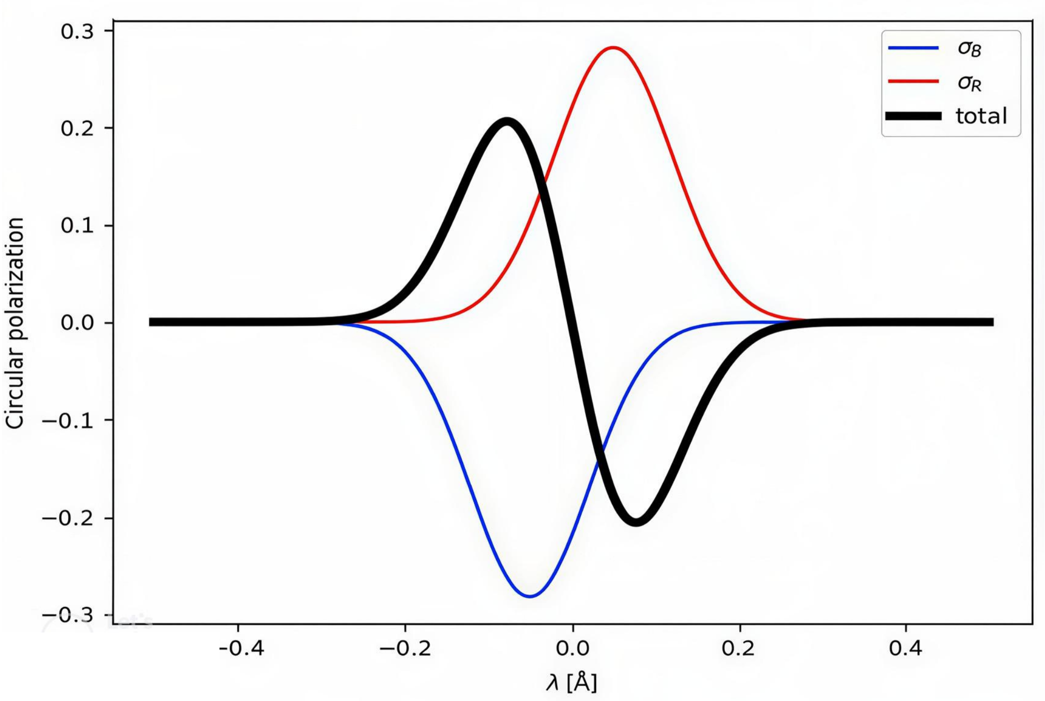


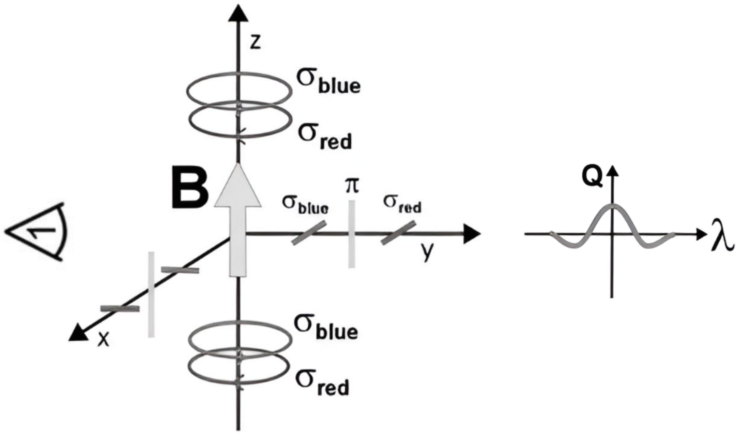
No magnetic field

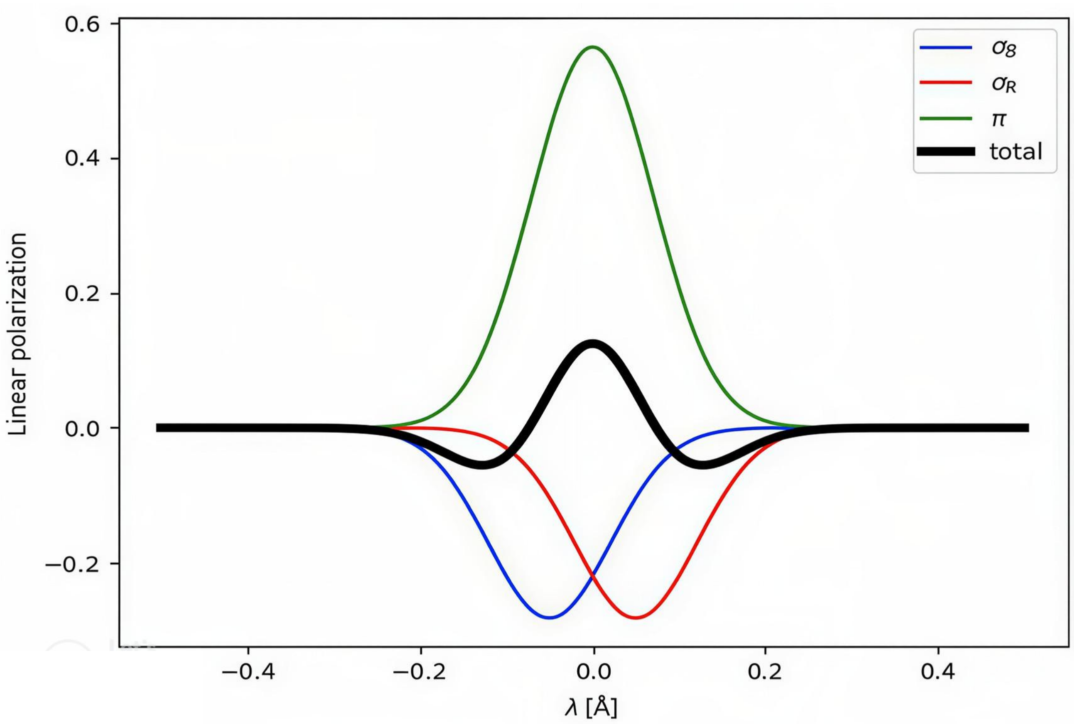
$$V_T(\lambda) = V_{\sigma_R}(\lambda) + V_{\sigma_B}(\lambda)$$




V
λ
1
z
σblue
σred
B
σblue
π
σred
y
x
σblue
σred
2
V
λ


σ_B
σ_R
total
Circular polarization
λ [Å]
0.3
0.2
0.1
0.0
−0.1
−0.2
−0.3
-0.4
−0.2
0.0
0.2
0.4


z
σ_{blue}
σ_{red}
B
σ_{blue}
π
σ_{red}
y
x
σ_{blue}
σ_{red}
Q
λ

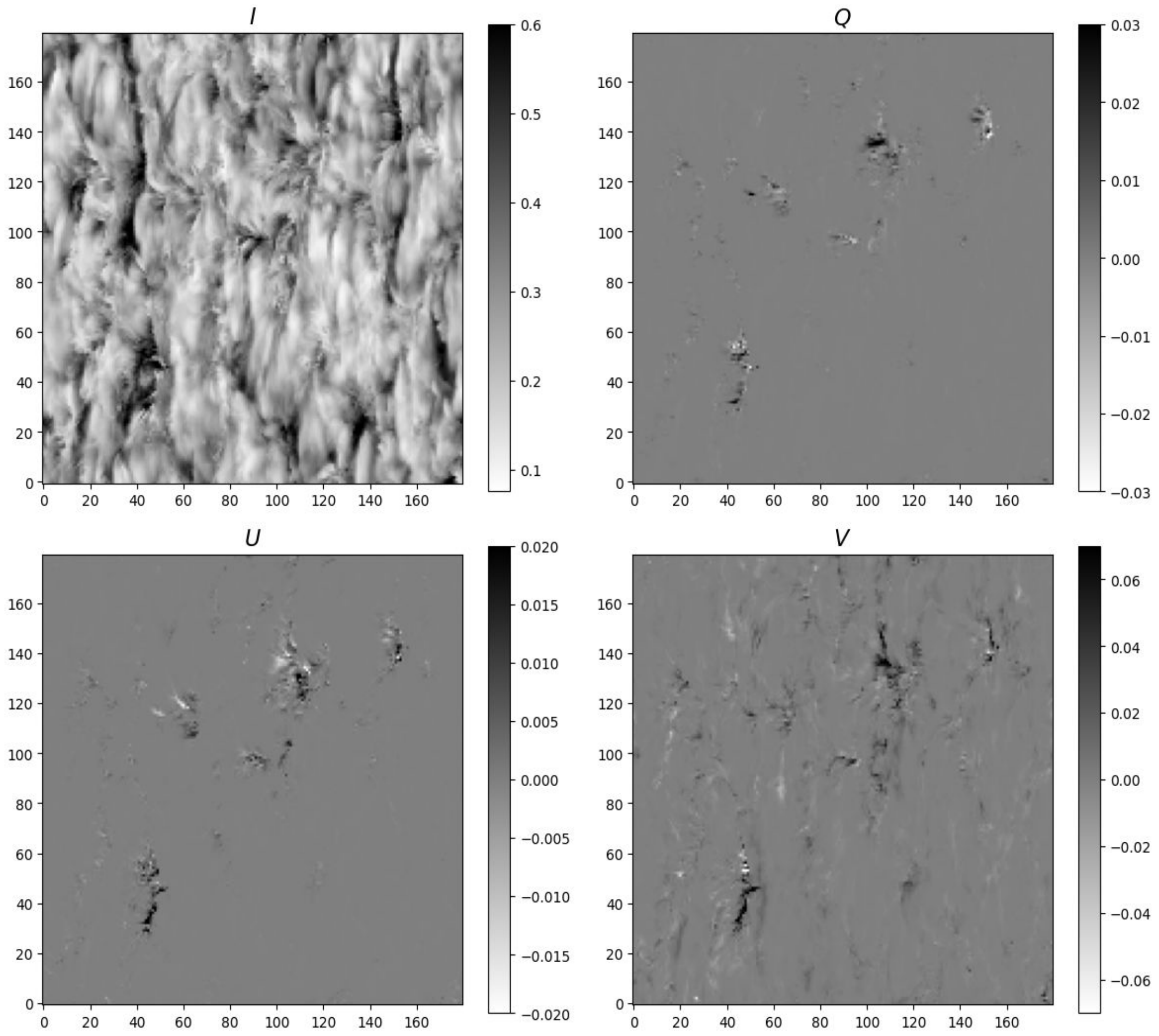

0.6
0.4
0.2
0.0
−0.2
Linear polarization
−0.4
−0.2
0.0
0.2
0.4
λ [Å]
σ_B
σ_R
π
total

Weak field approximation

$$V \propto \lambda^2 B_{\parallel} \frac{\partial I}{\partial \lambda}$$

$$Q \propto \lambda^4 B_{\perp}^2 \cos(2\varphi) \frac{\partial^2 I}{\partial \lambda^2}$$

$$U \propto \lambda^4 B_{\perp}^2 \sin(2\varphi) \frac{\partial^2 I}{\partial \lambda^2}$$



We need to solve radiative transfer equation!

$$\frac{d}{dz}\begin{pmatrix} I \\ Q \\ U \\ V \end{pmatrix}_\lambda = -\chi \begin{pmatrix} \eta_I & \eta_Q & \eta_U & \eta_V \\ \eta_Q & \eta_I & \rho_V & -\rho_U \\ \eta_U & -\rho_V & \eta_I & \rho_Q \\ \eta_V & \rho_U & -\rho_Q & \eta_I \end{pmatrix}_\lambda \begin{pmatrix} I - S_I \\ Q - S_Q \\ U - S_U \\ V - S_V \end{pmatrix}_\lambda$$

Propagation of radiation and interaction with matter in the presence of a magnetic field

- RTE is hard to solve analytically, so we use **Milne-Eddington approximation** (Unno, 1956; Rachkovsky, 1962)

$$\frac{d\vec{I}}{d\tau} = \boldsymbol{K}(\vec{I} - \vec{S})$$

propagation matrix

source function

- Magnetic field (Zeeman effect), velocity, pressure, and other parameters are constant with height, and the source function changes linearly with optical depth

$$\vec{S} \equiv \vec{S}_0 + \vec{S}_1\tau$$

possible to solve RTE analytically, so there is a basis for inversion!

Direct problem

$$\vec{x}(\tau) \equiv \left[T(\tau), p_e(\tau), v_{\mathrm{LOS}}(\tau), B(\tau), \theta(\tau), \varphi(\tau), \ldots\right]$$

atmospheric model

RTE

$$\vec{I}(\lambda) = \begin{bmatrix} I(\lambda) \\ Q(\lambda) \\ U(\lambda) \\ V(\lambda) \end{bmatrix}$$

we obtain Stokes vector

* Radiative transfer equation - RTE

Inverse problem

we define operator

$$\vec{I}_{\mathrm{syn}}(\lambda) = \mathcal{F}[\vec{x}(\tau)]$$

$$\vec{r} = \vec{I}_{\mathrm{obs}} - \vec{I}_{\mathrm{syn}}(\vec{x})$$

we look for a specific model

$$\vec{I}_{\mathrm{obs}}(\lambda) \approx \mathcal{F}[\vec{x}^*(\tau)]$$

problem is posed as minimization

$$\chi^2(\vec{x}) \propto \sum_{\lambda} \sum_{i=1}^{4} \left[I_{\mathrm{i,\,obs}}(\lambda) - I_{\mathrm{i,\,syn}}(\lambda; \vec{x}) \right]^2$$

we use Levenberg-Marquardt algorithm

$$(\mathbf{J}^\top \mathbf{J} + \lambda \, \mathrm{diag}\,(\mathbf{J}^\top \mathbf{J}))\delta\vec{x} = \mathbf{J}^\top \vec{r}$$

$$\mathbf{J} = \partial \mathcal{F} / \partial \vec{x}$$

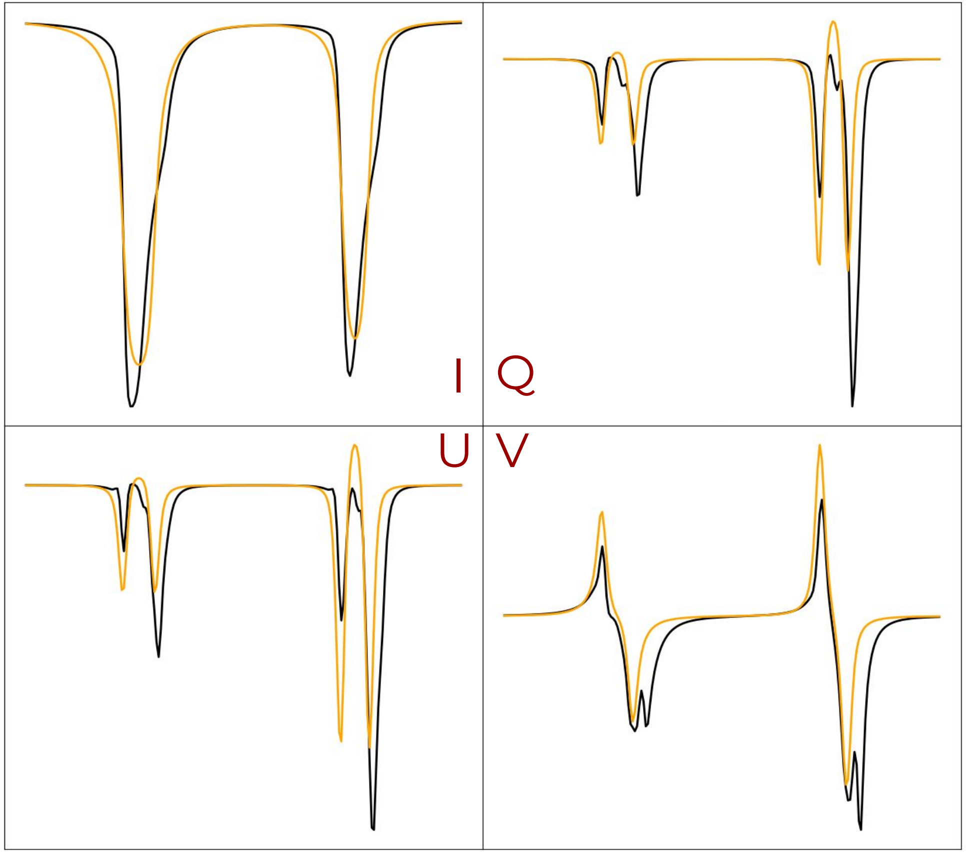
response functions

Milne-Eddington inversion

original
profiles

inferred
profiles

*pixel-by-pixel
inversion



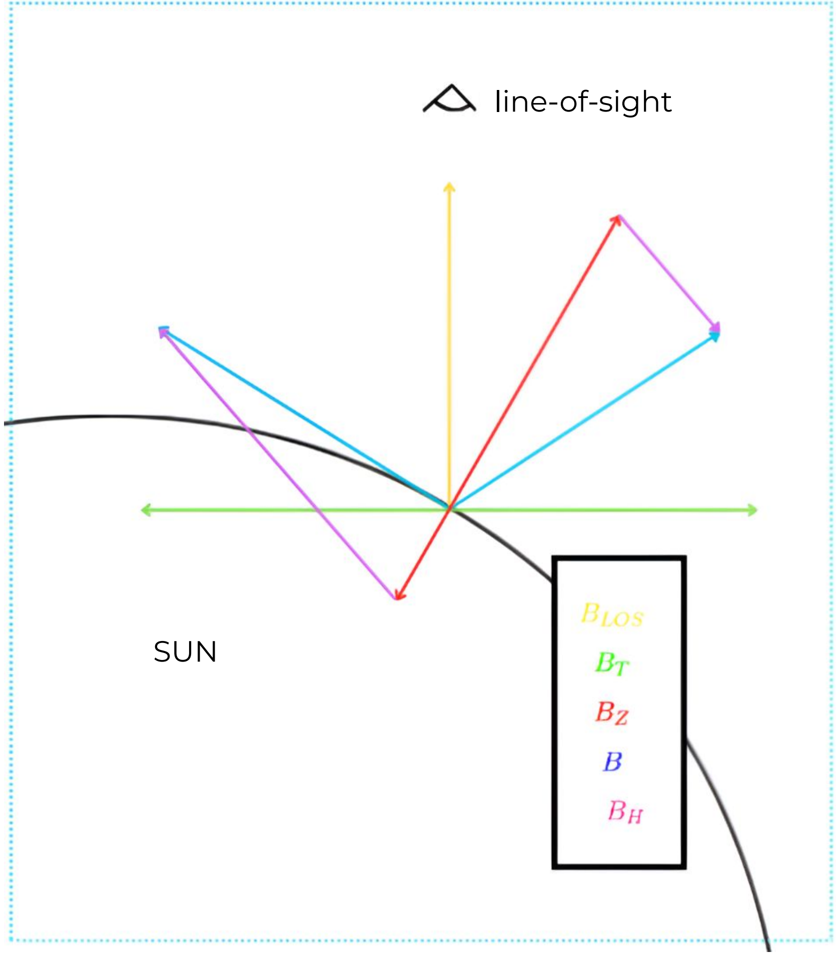
Azimuth ambiguity

$$\varphi \rightarrow \varphi + 180^\circ$$

$$Q \propto \cos(2\varphi)$$

$$U \propto \sin(2\varphi)$$

$$B_\perp \propto (Q^2 + U^2)^{1/4}$$



Resolving azimuthal 180° ambiguity

All methods are based on additional approximations or assumptions; There are two general approaches:

a) Choosing the solution closest to a reference field (e.g., potential field)

b) Minimization of a functional

Our approach: a differentiable surrogate of so-called Minimum energy method. Reasons for this choice are:

a) Global method - disambiguation informed by spatial context, not pixel-isolated

b) Physical basis - acute angle (potential reference) + Jz^2 (current minimization) + dBz/dz (divergence constraint)

c) Proven superiority - MEM was the top performer in the Metcalf et al. (2006) comparison

Minimum *energy* method

- Minimization of divergence yields a physically-based solution

- Minimization of the vertical component of the current density yields a smoothness condition

Horizontal derivatives → finite differences

Vertical derivative → the so-called spectral method

$$E = \sum_{i,j}^{N_x,N_y} \left(\left| \frac{\partial B_x}{\partial x}(i,j) + \frac{\partial B_y}{\partial y}(i,j) + \frac{\partial B_z}{\partial z}(i,j) \right|^2 + \eta \left| \frac{\partial B_y}{\partial x}(i,j) - \frac{\partial B_x}{\partial y}(i,j) \right|^2 \right)$$

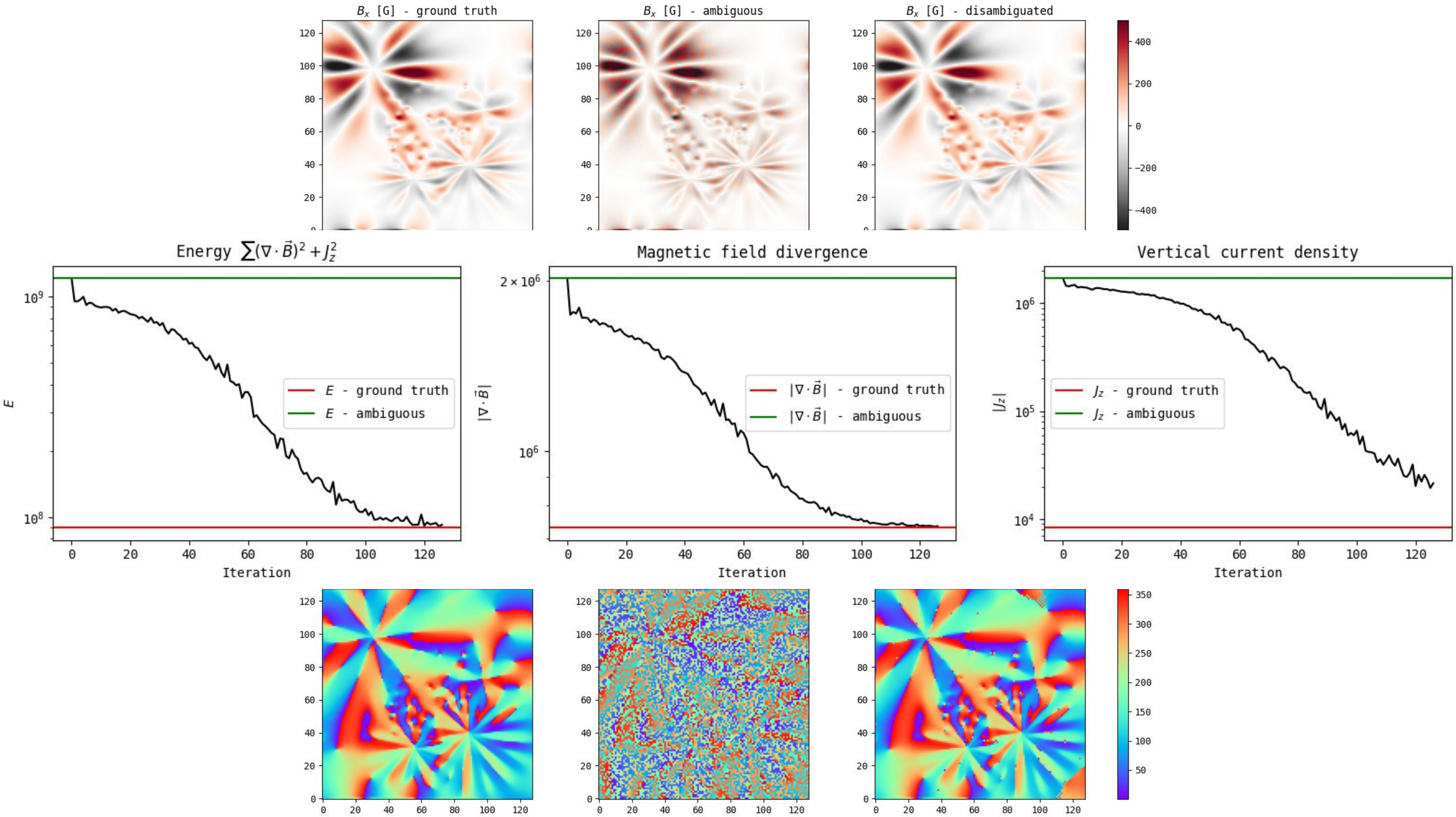
Minimum *energy* method
→ **simulated annealing**

- For every pixel, adding 180° to the azimuth is proposed
- if change decreases energy it is accepted
- if change increases energy, it is accepted with probability that depends on *temperature*

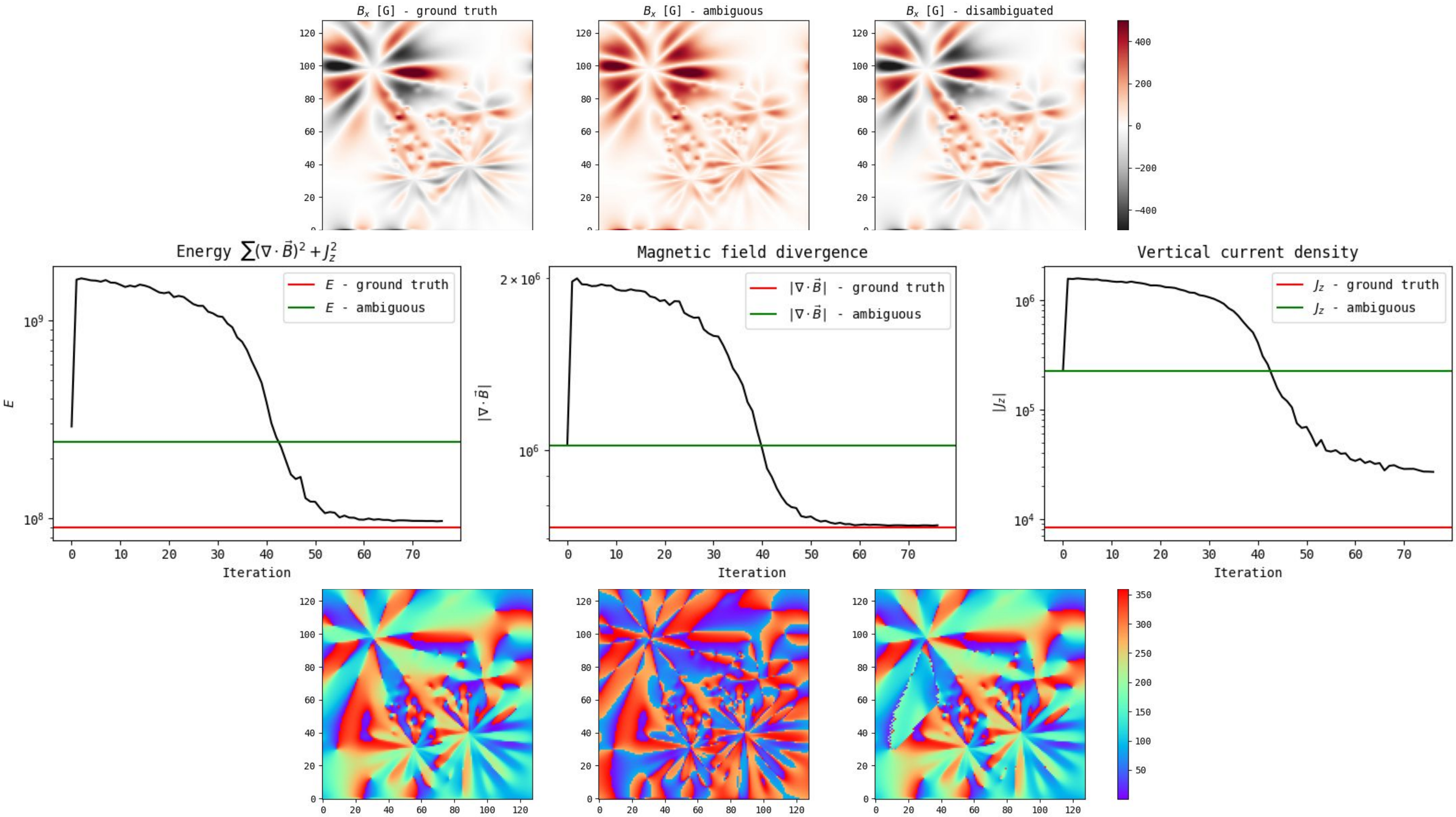
$$\Delta E = E' - E$$

$$p(\Delta E) = e^{-(\Delta E / T)}$$

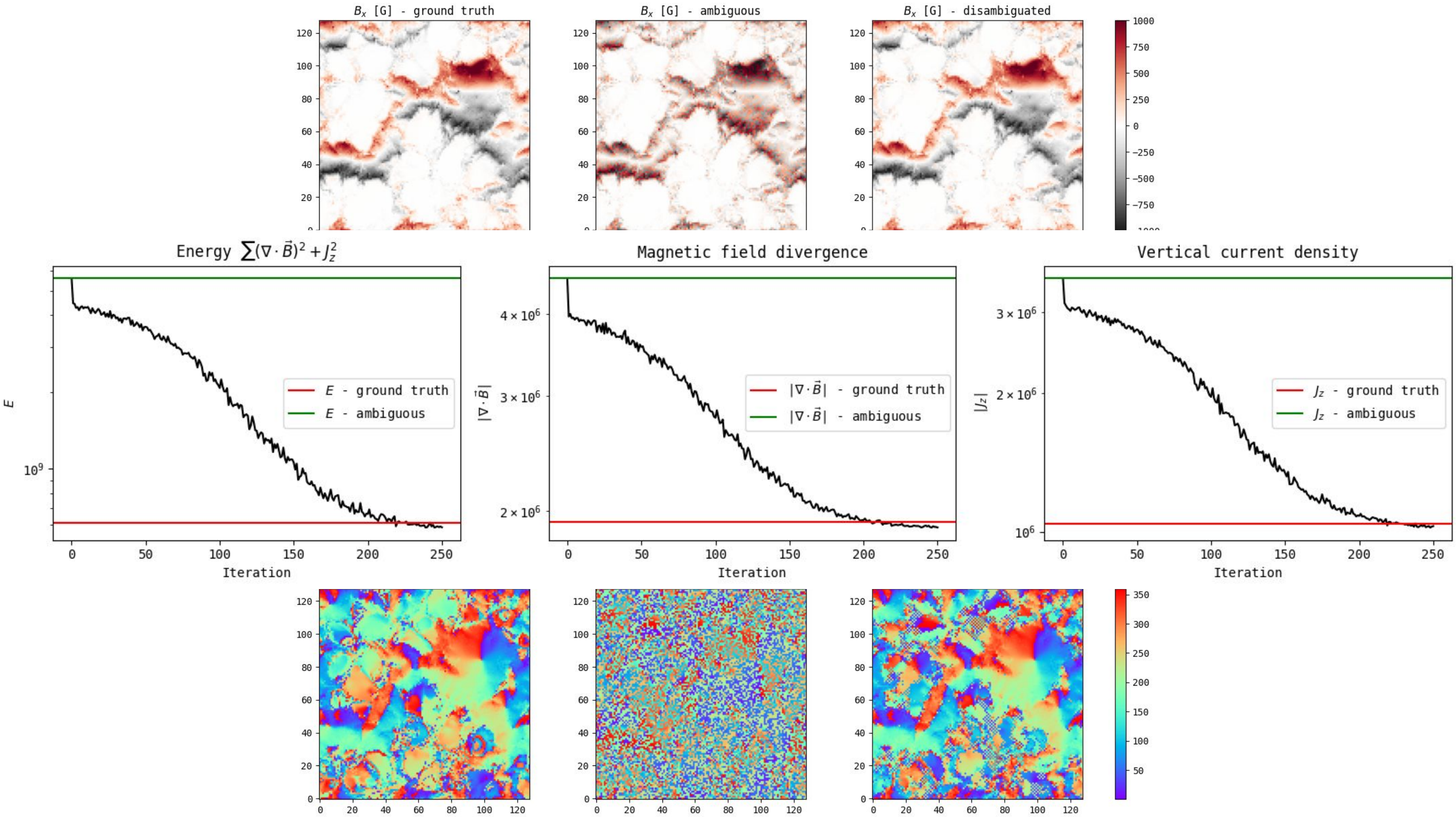
The system is gradually cooled by decreasing the *temperature*!



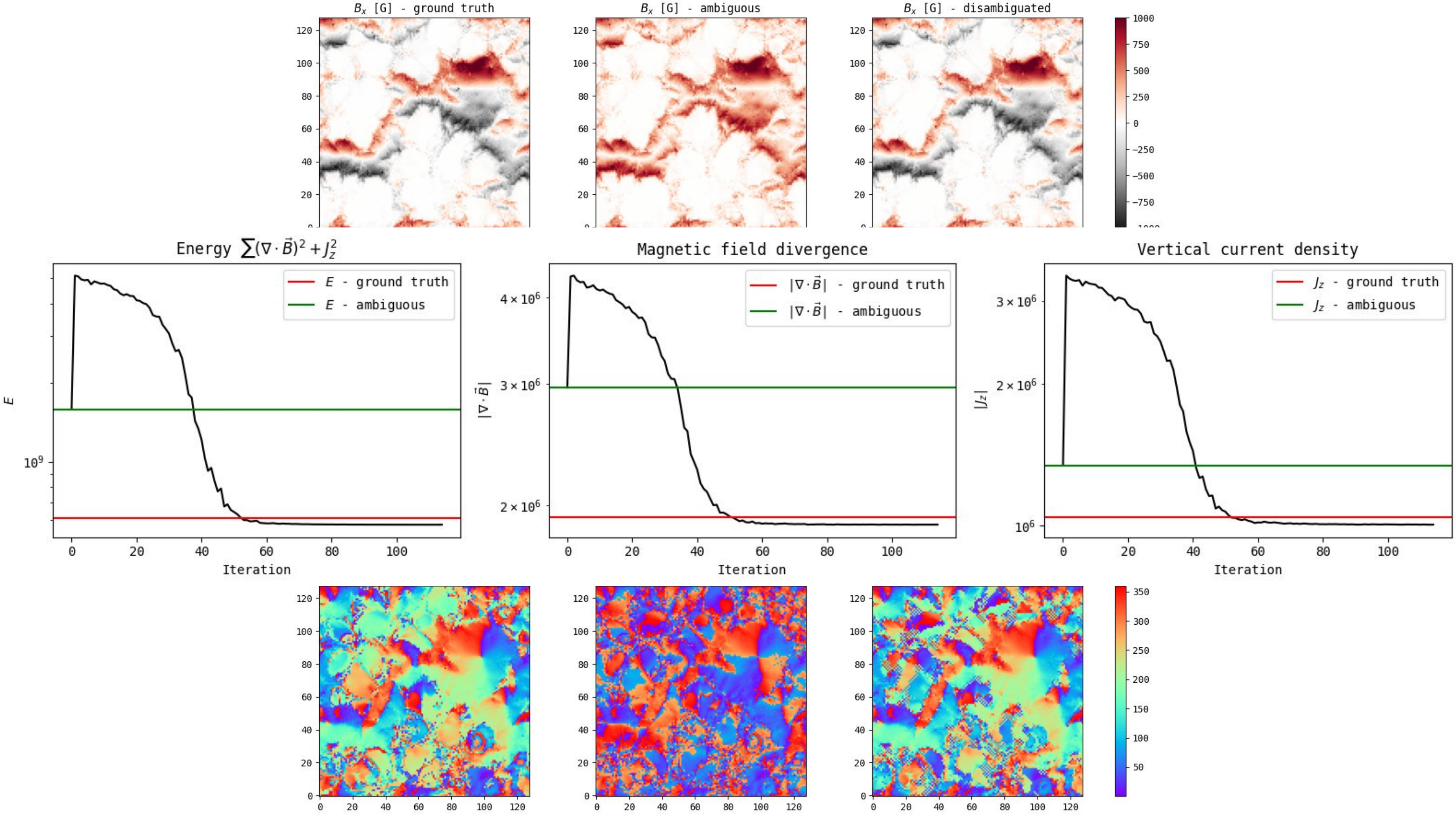
B_x [G] - ground truth
B_x [G] - ambiguous
B_x [G] - disambiguated
Energy $\sum(\nabla\cdot\vec{B})^2 + J_z^2$
Magnetic field divergence
Vertical current density
E - ground truth
E - ambiguous
$|\nabla\cdot\vec{B}|$ - ground truth
$|\nabla\cdot\vec{B}|$ - ambiguous
J_z - ground truth
J_z - ambiguous
E
$|\nabla\cdot\vec{B}|$
$|J_z|$
Iteration



B_x [G] - ground truth
B_x [G] - ambiguous
B_x [G] - disambiguated
Energy $\sum(\nabla\cdot\vec{B})^2+J_z^2$
E - ground truth
E - ambiguous
Magnetic field divergence
$|\nabla\cdot\vec{B}|$ - ground truth
$|\nabla\cdot\vec{B}|$ - ambiguous
Vertical current density
J_z - ground truth
J_z - ambiguous
Iteration



B_x [G] - ground truth
B_x [G] - ambiguous
B_x [G] - disambiguated
Energy $\sum(\nabla\cdot\vec{B})^2+J_z^2$
E - ground truth
E - ambiguous
Magnetic field divergence
$|\nabla\cdot\vec{B}|$ - ground truth
$|\nabla\cdot\vec{B}|$ - ambiguous
Vertical current density
J_z - ground truth
J_z - ambiguous
Iteration



B_x [G] - ground truth
B_x [G] - ambiguous
B_x [G] - disambiguated
Energy $\sum(\nabla\cdot\vec{B})^2 + J_z^2$
E - ground truth
E - ambiguous
Magnetic field divergence
$|\nabla\cdot\vec{B}|$ - ground truth
$|\nabla\cdot\vec{B}|$ - ambiguous
Vertical current density
J_z - ground truth
J_z - ambiguous
Iteration

Why machine learning and what's been done so far?

Socas-Navarro (2005) - Strategies for Spectral Profile Inversion Using Artificial Neural Networks

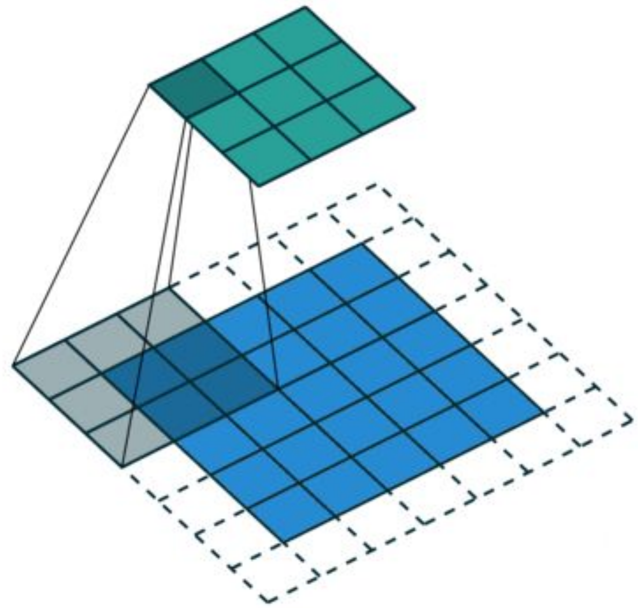
Asensio Ramos & Díaz Baso (2019) - Stokes Inversion based on Convolutional Neural Networks

Milić & Gafeira (2020) - Mimicking spectropolarimetric inversions using convolutional neural networks

Higgins et al. (2022) - SynthIA: A Synthetic Inversion Approximation for the Stokes Vector Fusing SDO and Hinode into a Virtual Observatory

Jarolim et al. (2025) - PINN ME: A Physics-Informed Neural Network Framework for Accurate Milne-Eddington Inversions of Solar Magnetic Fields

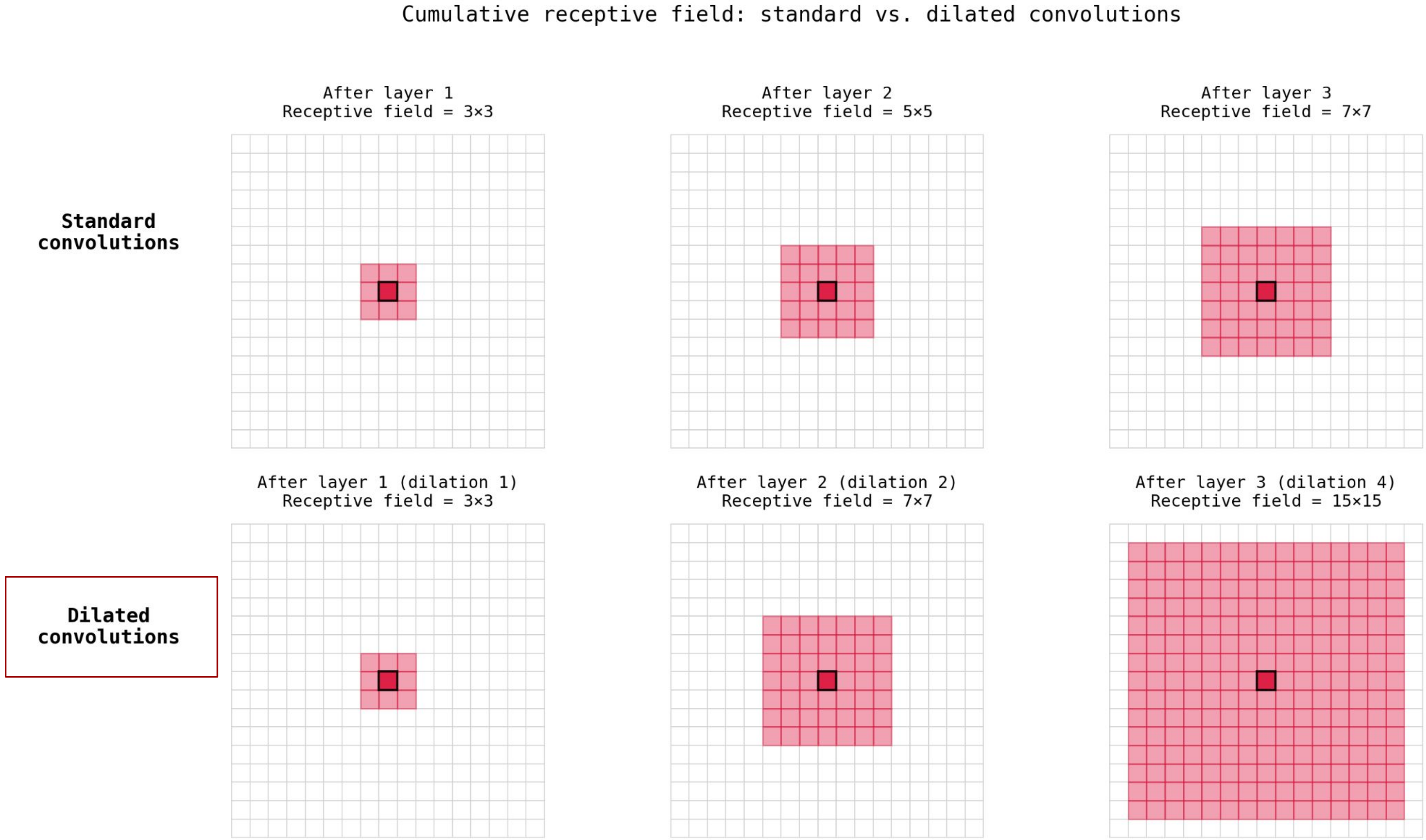
Convolutional Neural Networks



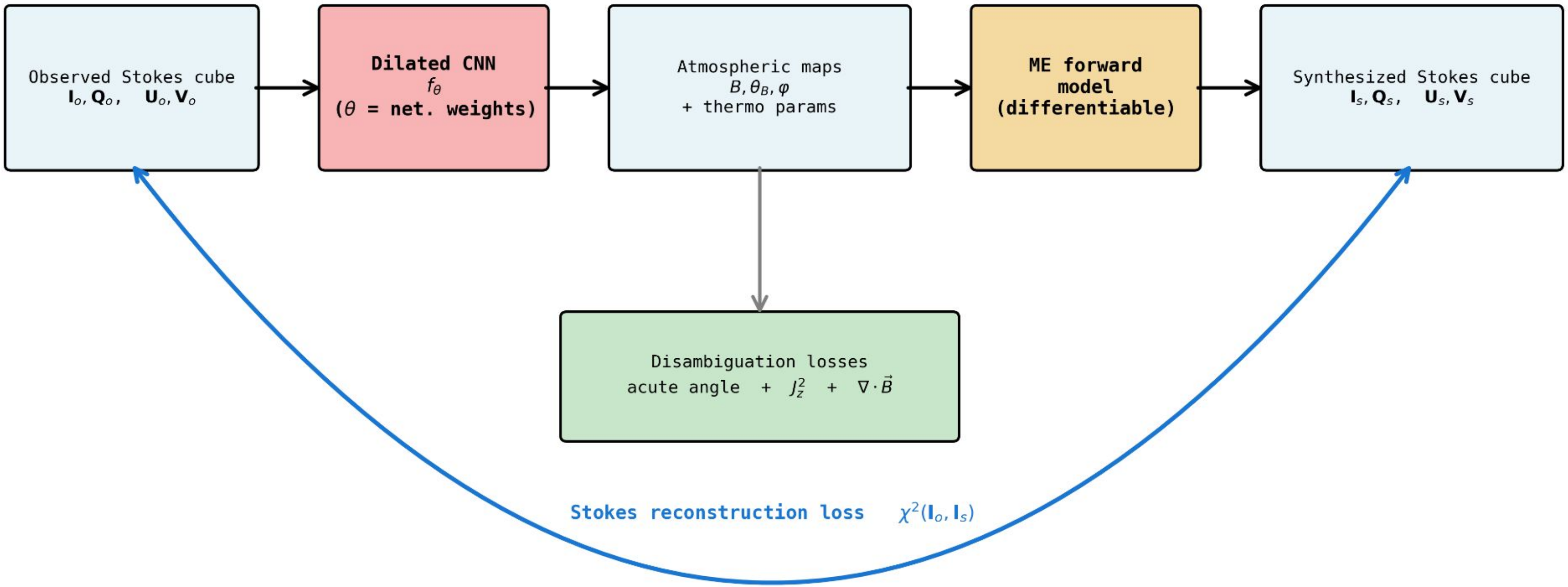
Convolution - a small kernel (e.g., 3×3) slides across the input, applying the *same weights* at every spatial position

Translation invariance - the operation is identical everywhere; physical laws don't care where in the field of view we look

Receptive field - each output pixel "sees" only a small neighbourhood of the input; stacking layers expands this neighbourhood


Cumulative receptive field: standard vs. dilated convolutions
Standard convolutions
After layer 1
Receptive field = 3×3
After layer 2
Receptive field = 5×5
After layer 3
Receptive field = 7×7
Dilated convolutions
After layer 1 (dilation 1)
Receptive field = 3×3
After layer 2 (dilation 2)
Receptive field = 7×7
After layer 3 (dilation 4)
Receptive field = 15×15

End-to-end pipeline



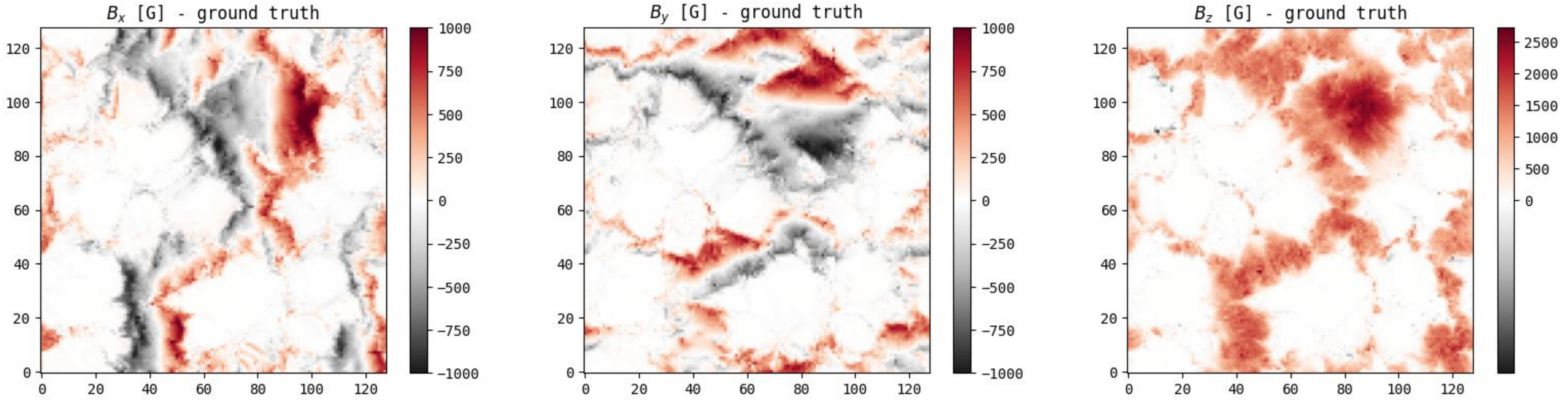
$$\mathcal{L} = \mathcal{L}_{\text{Stokes}} + \mathcal{L}_{\text{disambig}}$$

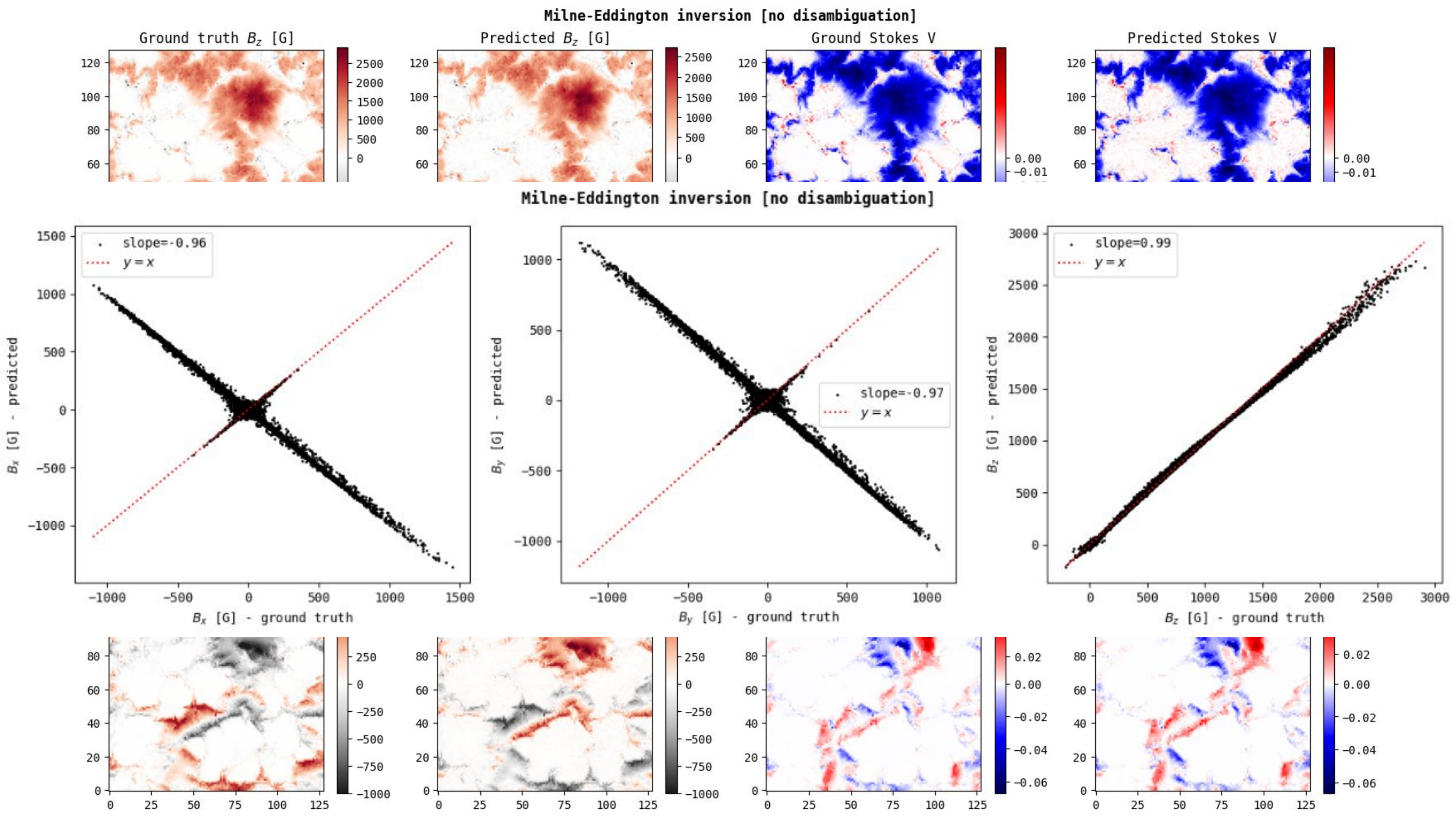
$$\mathcal{L}_{\text{Stokes}} = \frac{1}{4 N_\lambda N_{\text{pix}}} \sum_{i,j} \sum_{k=1}^{4} \sum_{\lambda} \left[I_{k,\text{obs}}(\lambda; i, j) - I_{k,\text{syn}}(\lambda; i, j) \right]^2$$

$$\mathcal{L}_{\text{disambig}} = \lambda_1 \mathcal{L}_{\text{acute}} + \lambda_2 \mathcal{L}_{J_z^2} + \lambda_3 \mathcal{L}_{\nabla \cdot \vec{B}}$$

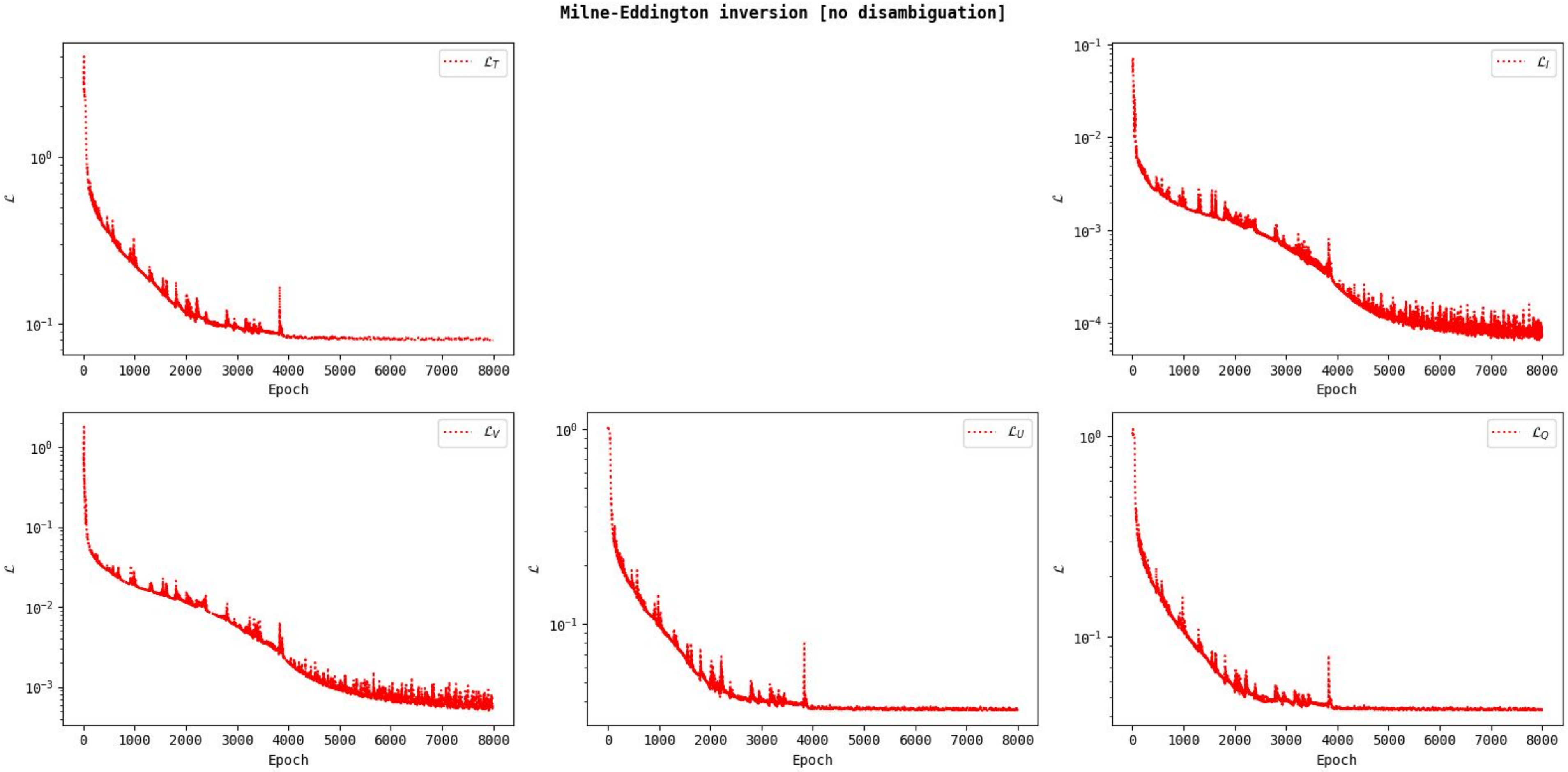
RESULTS

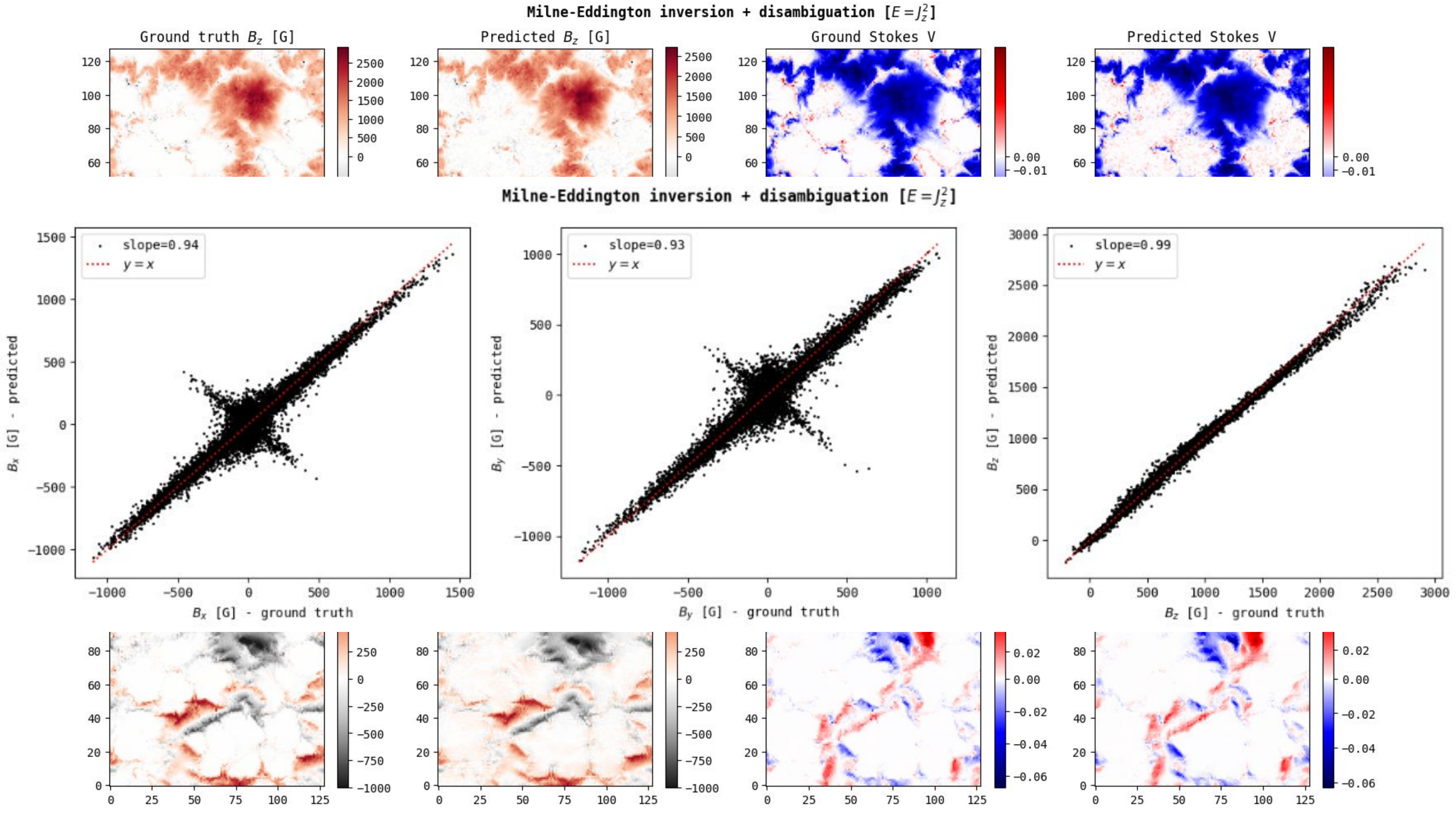
Dataset we test this on



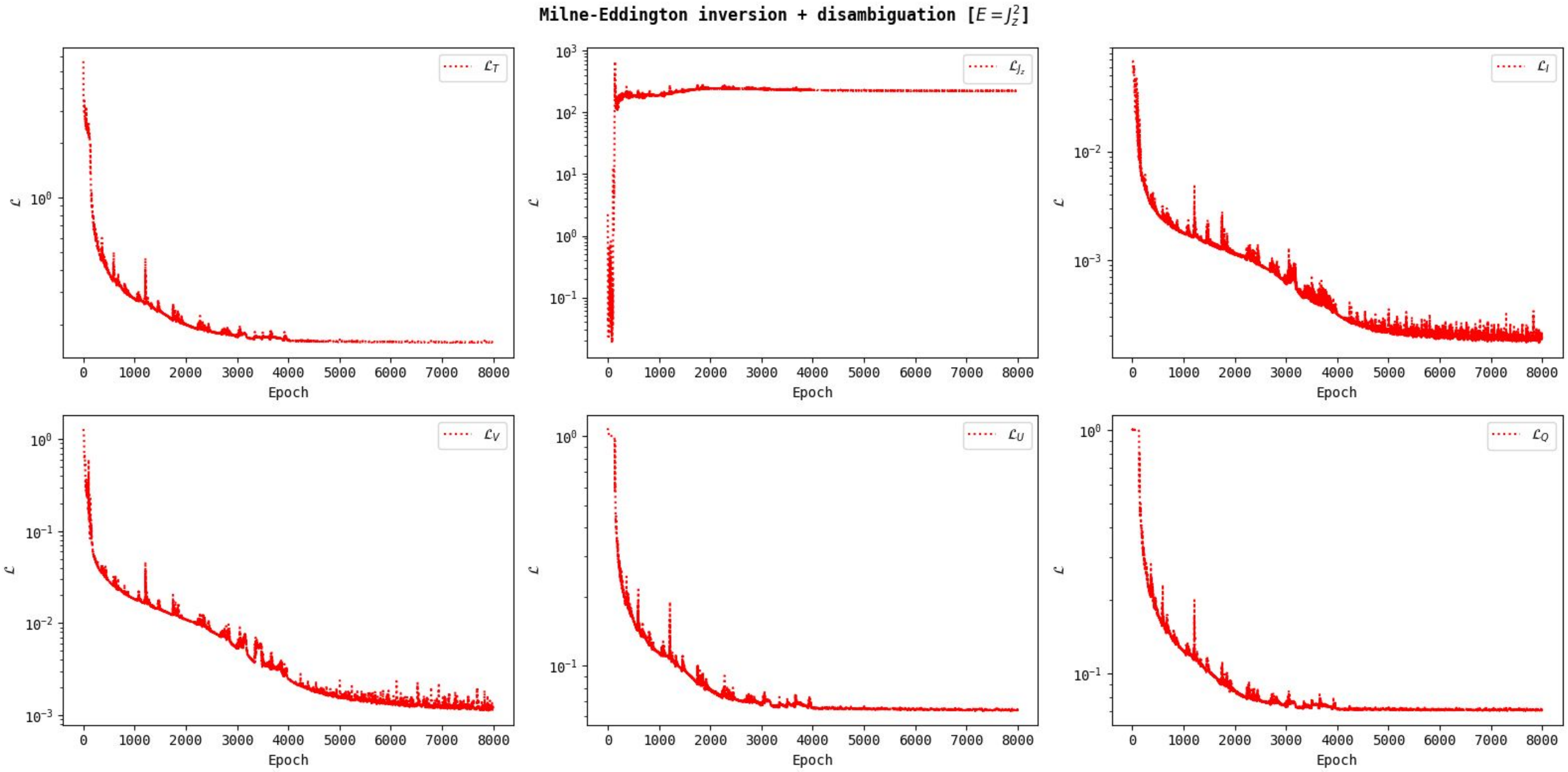


Milne-Eddington inversion [no disambiguation]
Ground truth B_z [G]
Predicted B_z [G]
Ground Stokes V
Predicted Stokes V
Milne-Eddington inversion [no disambiguation]
slope=-0.96
$y = x$
B_x [G] - predicted
B_x [G] - ground truth
slope=-0.97
$y = x$
B_y [G] - predicted
B_y [G] - ground truth
slope=0.99
$y = x$
B_z [G] - predicted
B_z [G] - ground truth

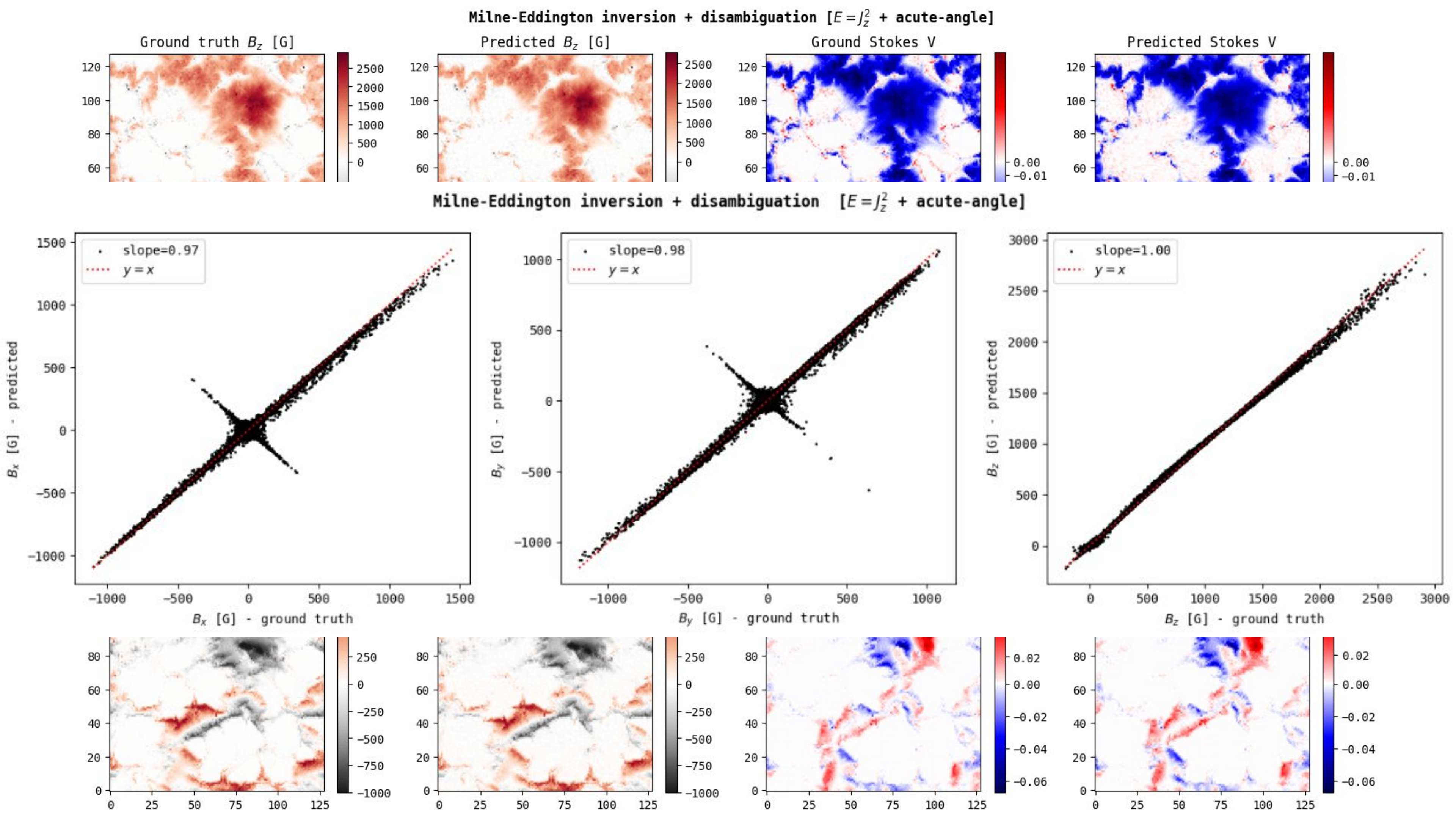

Milne-Eddington inversion [no disambiguation]
$\mathcal{L}_T$
$\mathcal{L}_I$
$\mathcal{L}_V$
$\mathcal{L}_U$
$\mathcal{L}_Q$
$\mathcal{L}$
Epoch

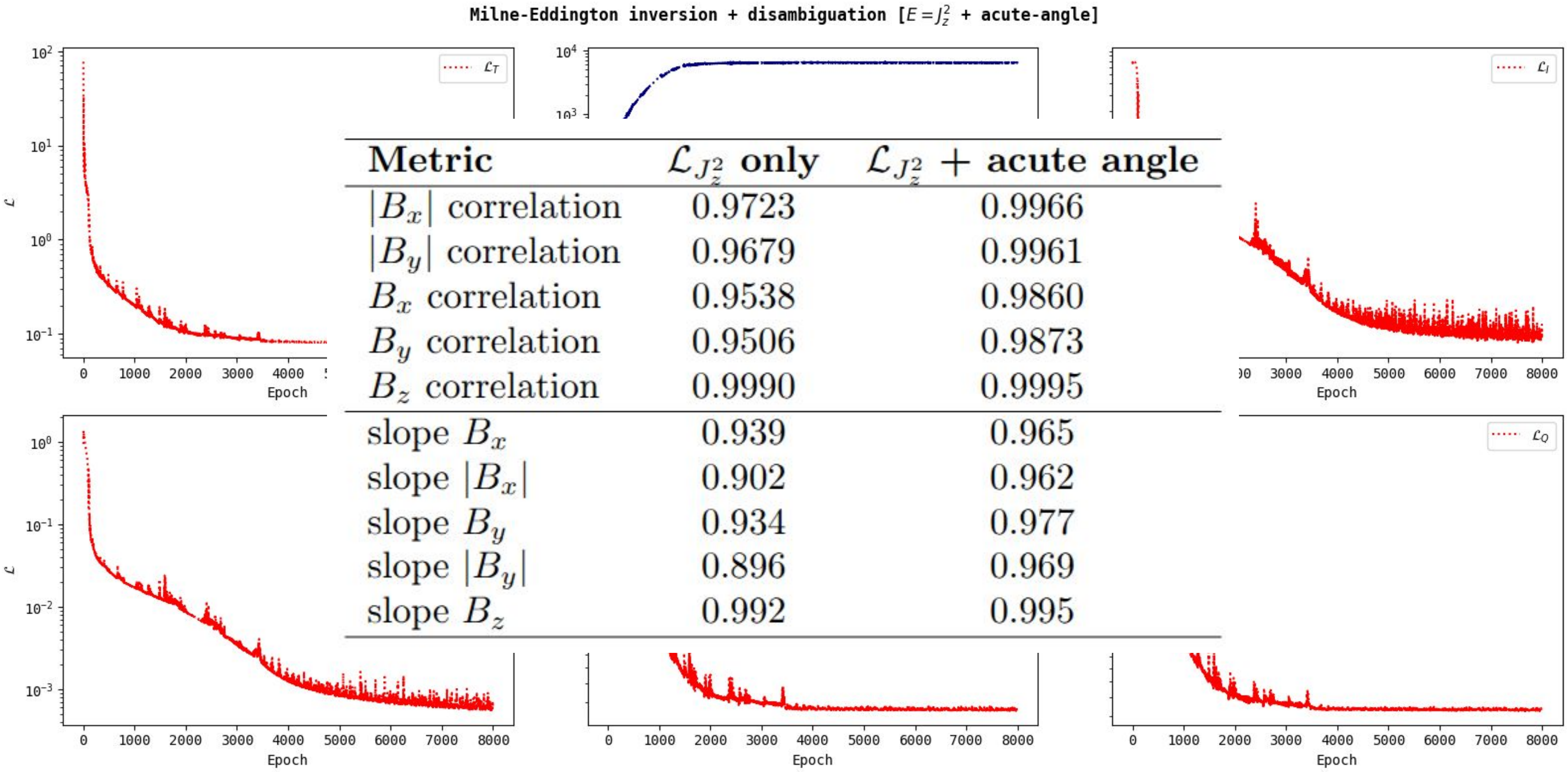


Milne-Eddington inversion + disambiguation [$E = J_z^2$]
Ground truth B_z [G]
Predicted B_z [G]
Ground Stokes V
Predicted Stokes V
Milne-Eddington inversion + disambiguation [$E = J_z^2$]
slope=0.94
$y = x$
B_x [G] - predicted
B_x [G] - ground truth
slope=0.93
$y = x$
B_y [G] - predicted
B_y [G] - ground truth
slope=0.99
$y = x$
B_z [G] - predicted
B_z [G] - ground truth

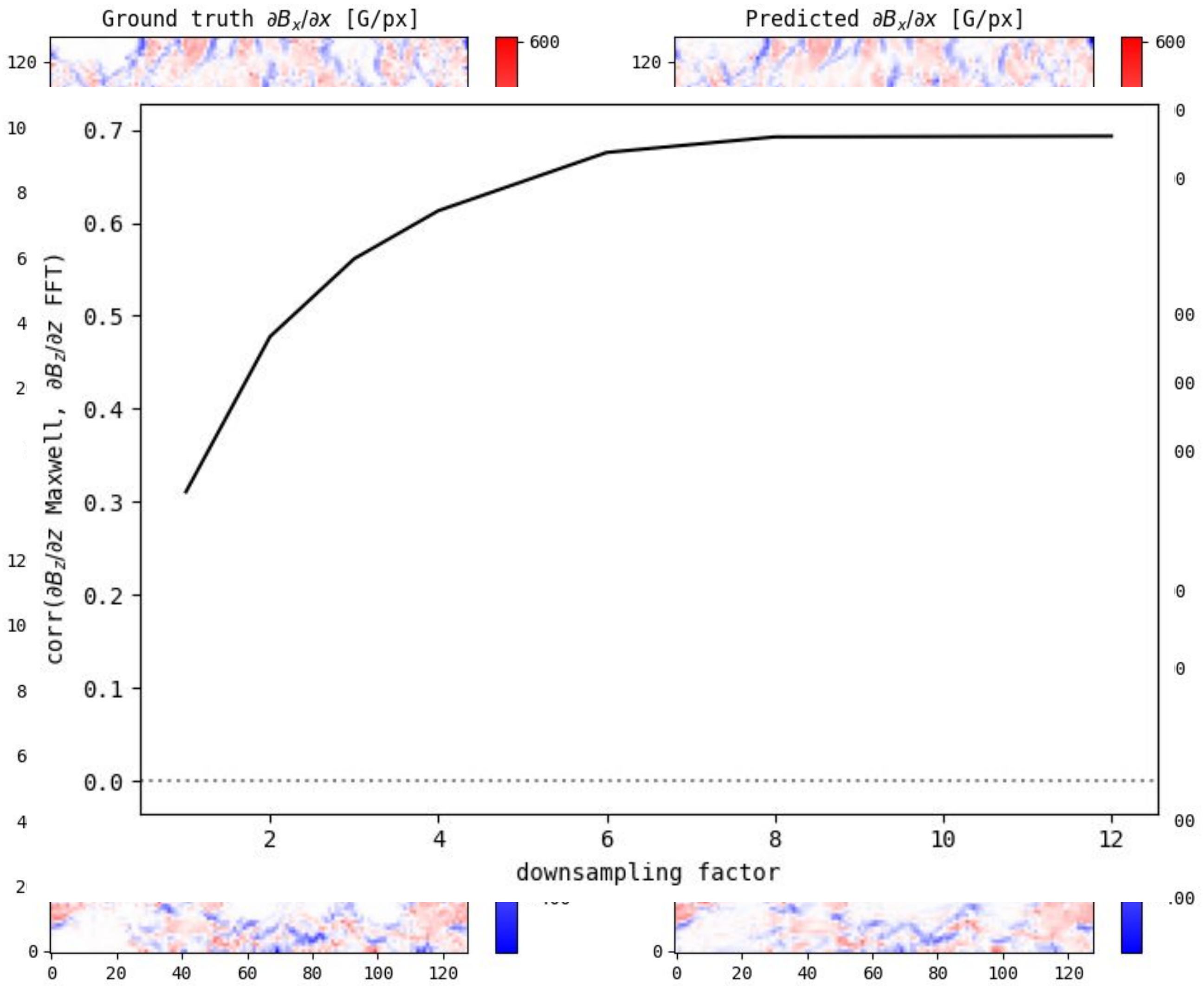


Milne-Eddington inversion + disambiguation [$E = J_z^2$]
$\mathcal{L}_T$
$\mathcal{L}_{J_z}$
$\mathcal{L}_I$
$\mathcal{L}_V$
$\mathcal{L}_U$
$\mathcal{L}_Q$
$\mathcal{L}$
Epoch
0
1000
2000
3000
4000
5000
6000
7000
8000
10^0
10^{-1}
10^{-2}
10^{-3}
10^1
10^2
10^3


Milne-Eddington inversion + disambiguation [$E = J_z^2$ + acute-angle]
Ground truth B_z [G]
Predicted B_z [G]
Ground Stokes V
Predicted Stokes V
Milne-Eddington inversion + disambiguation [$E = J_z^2$ + acute-angle]
slope=0.97
$y = x$
B_x [G] - predicted
B_x [G] - ground truth
slope=0.98
$y = x$
B_y [G] - predicted
B_y [G] - ground truth
slope=1.00
$y = x$
B_z [G] - predicted
B_z [G] - ground truth



Metric	$\mathcal{L}_{J_z^2}$ only	$\mathcal{L}_{J_z^2}$ + acute angle
$\lvert B_x \rvert$ correlation	0.9723	0.9966
$\lvert B_y \rvert$ correlation	0.9679	0.9961
B_x correlation	0.9538	0.9860
B_y correlation	0.9506	0.9873
B_z correlation	0.9990	0.9995
slope B_x	0.939	0.965
slope $\lvert B_x \rvert$	0.902	0.962
slope B_y	0.934	0.977
slope $\lvert B_y \rvert$	0.896	0.969
slope B_z	0.992	0.995



Ground truth ∂Bx/∂x [G/px]
Predicted ∂Bx/∂x [G/px]
600
600
corr(∂Bz/∂z Maxwell, ∂Bz/∂z FFT)
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
2
4
6
8
10
12
downsampling factor
0
20
40
60
80
100
120

CONCLUSION

Self-supervised dilated CNN performs ME inversion and 180°
disambiguation simultaneously!

Acute angle loss is the critical disambiguation component

J_z^2 provides regularization but cannot disambiguate fully alone

$\nabla \cdot \mathbf{B}$ term unreliable at native resolution - ongoing work

Ongoing work:

Quantifying where and how badly the FFT potential-field closure fails across different magnetic structures

Future work:

Multi-height inversions for true $\partial B_z/\partial z$

Application to Solar Orbiter PHI data

Uncertainty quantification