Determining the distribution of interstellar gas with information field theory

Andrea Vittino (TTK, RWTH Aachen)
with Philipp Mertsch

Big Data Science in Astroparticle Research Workshop
Aachen, 17-19 February 2020

Outline

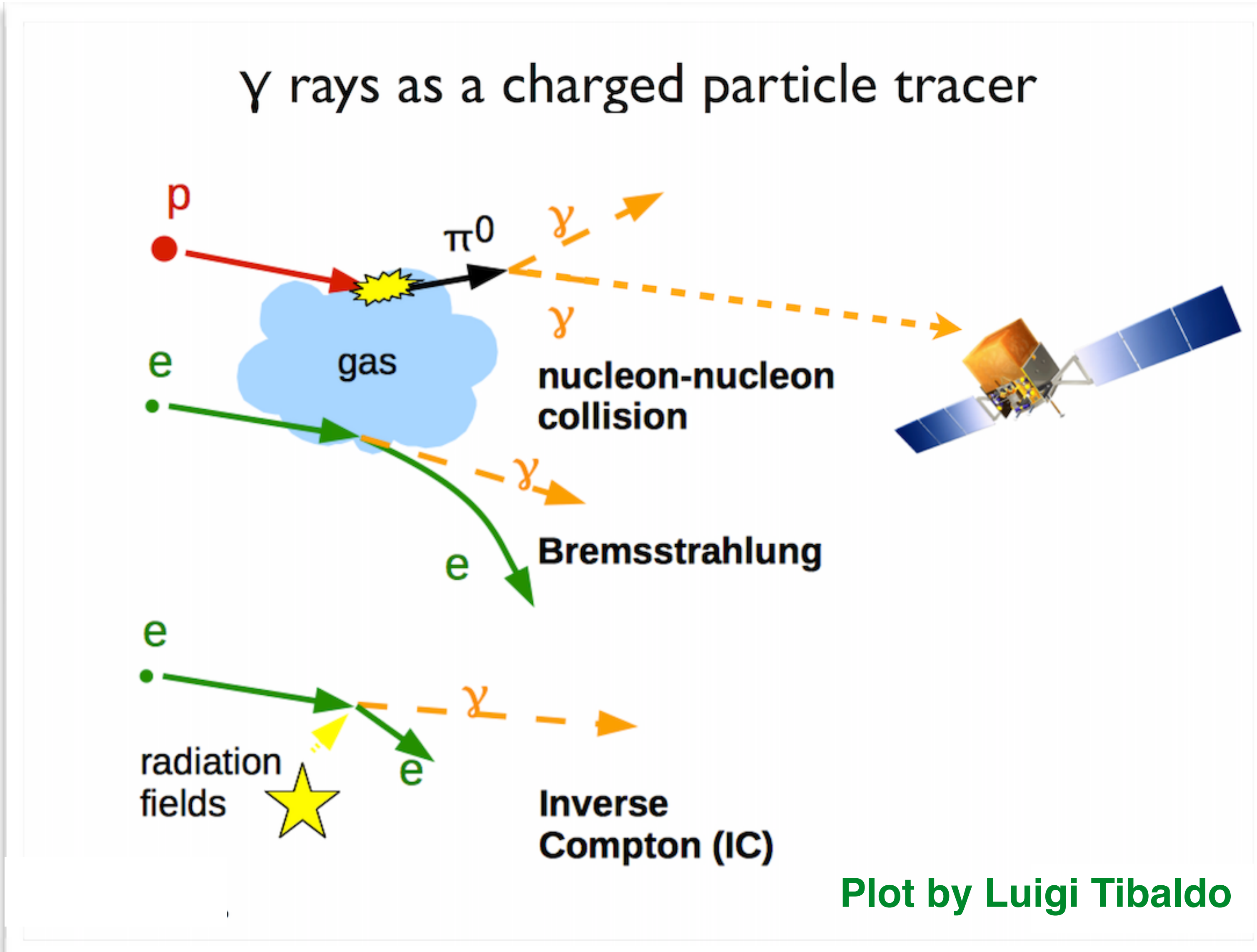
- **Motivation**
  - ✦ Why are we interested in the interstellar gas distribution?
- **The problem** : inferring the gas distribution
  - ✦ How do we observe the gas? How can we use such observation to understand how gas is distributed?
- **Our method** : information field theory
  - ✦ What is information field theory? How do we intend to use it to infer the gas distribution?

Outline

- **Motivation**
  - ✦ Why are we interested in the interstellar gas distribution?
- **The problem** : inferring the gas distribution
  - ✦ How do we observe the gas? How can we use such observation to understand how gas is distributed?
- **Our method** : information field theory
  - ✦ What is information field theory? How do we intend to use it to infer the gas distribution?

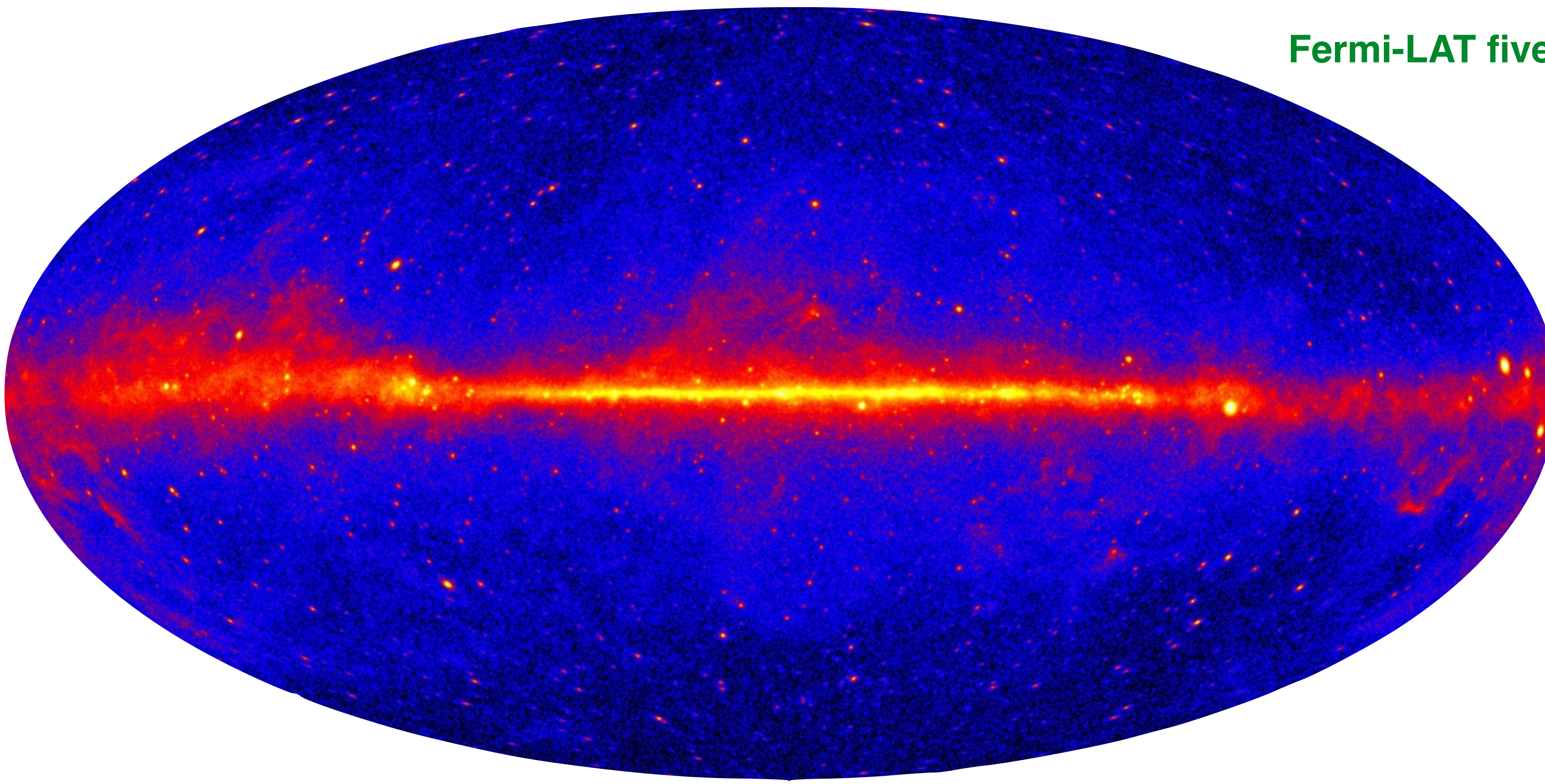
Cosmic rays produce gamma-rays

By interacting with the interstellar gas, **cosmic rays produce gamma-rays**



Observations of **diffuse gamma-ray** emission provide a **direct probe** of **spatial densities and spectra** of CRs in distant locations, **far beyond the reach of direct measurements**

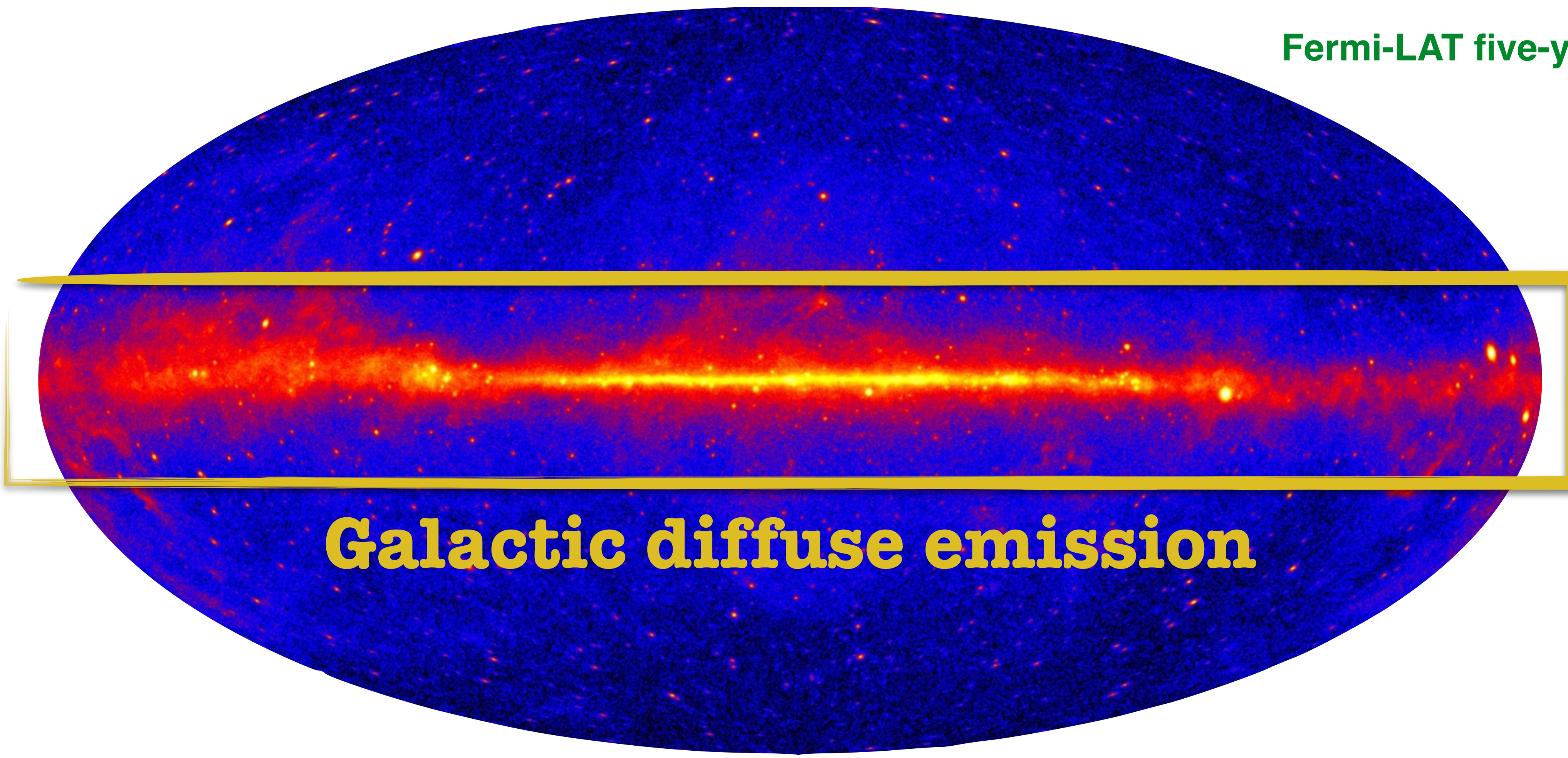
Gamma-ray data

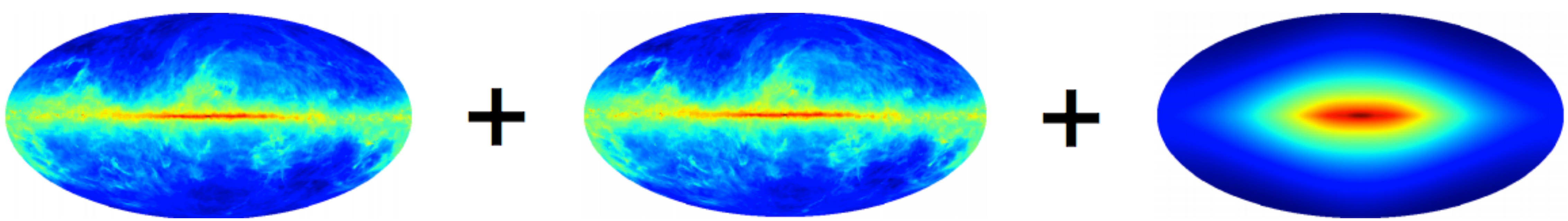


Fermi-LAT five-year skymap

Gamma-ray data

Fermi-LAT five-year skymap

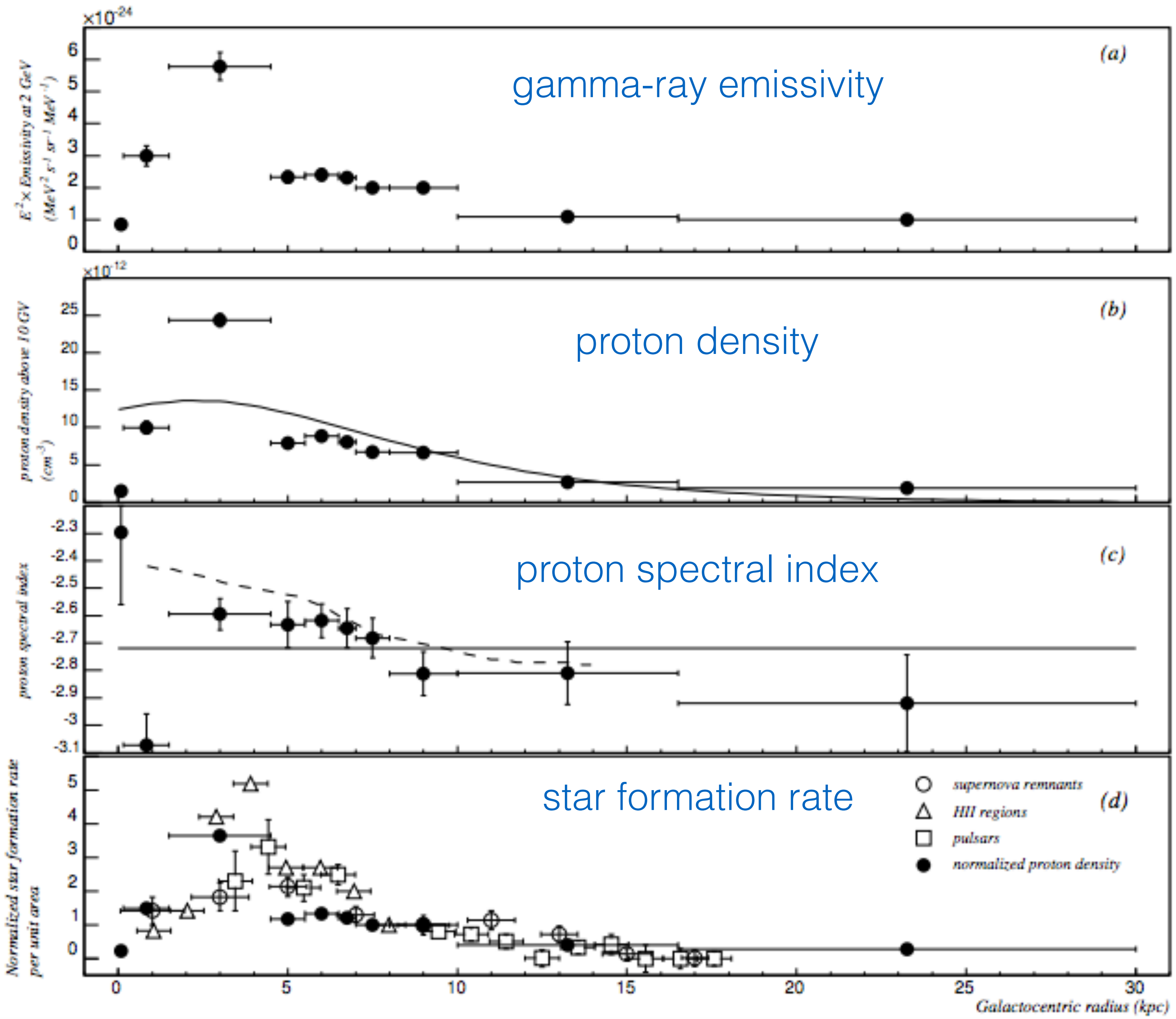




pion decay   Bremsstrahlung   Inverse Compton

Galactic CR proton density

Hadronic emission (pion decay)

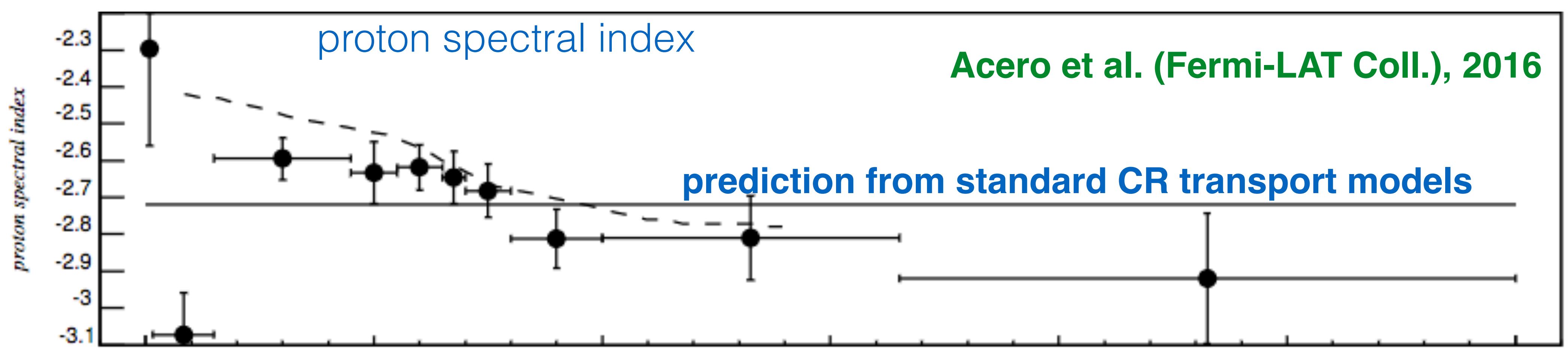


Acero et al. (Fermi-LAT Coll.), 2016

If we know the **gas density**, we can **map CR protons** across the Galactic plane

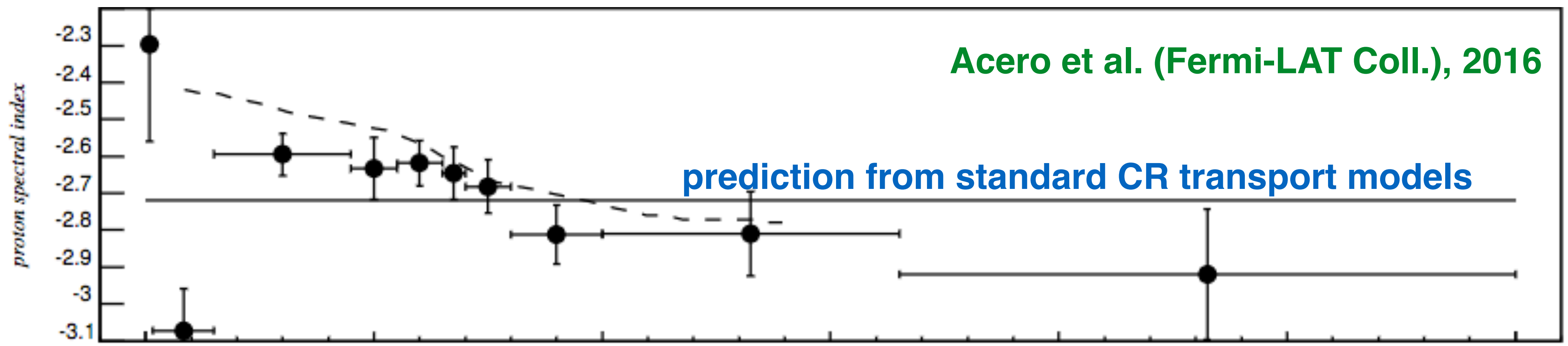
CR proton density - the role of CR transport

Results **do not agree** with the prediction from **standard CR transport models**.

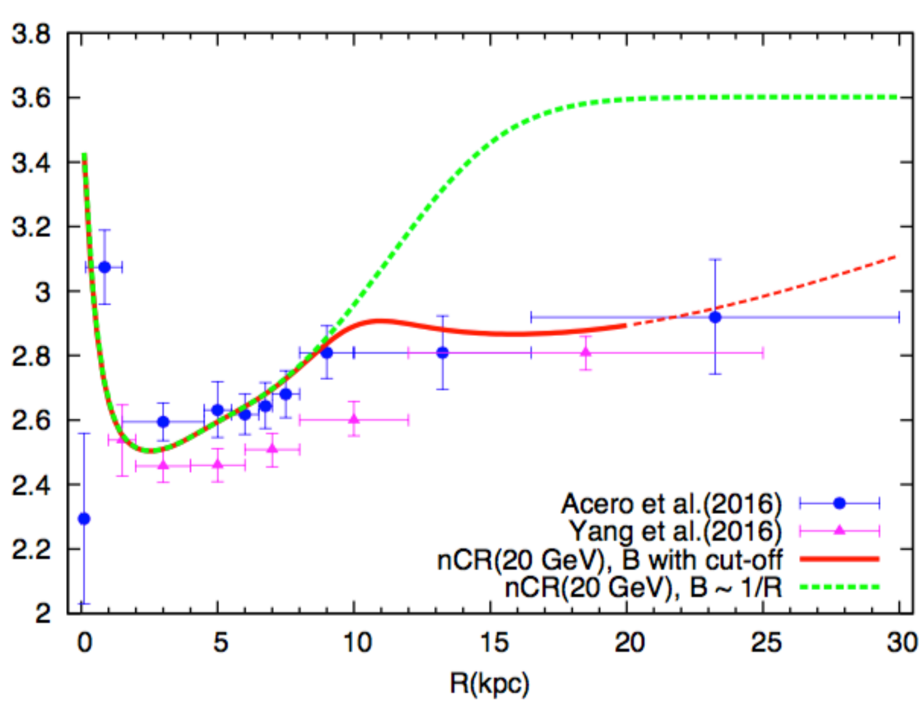


CR proton density - the role of CR transport

Results **do not agree** with the prediction from **standard CR transport models**.

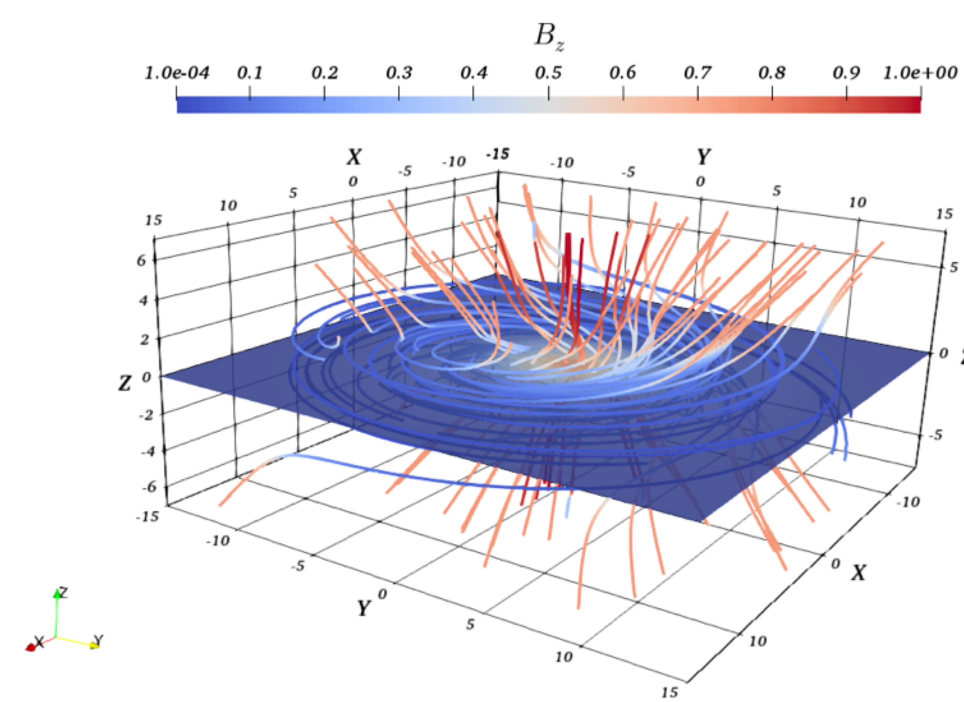


Alternative transport models have been proposed:



CR transport in self-induced turbulence

Recchia, Morlino, Blasi, 2017

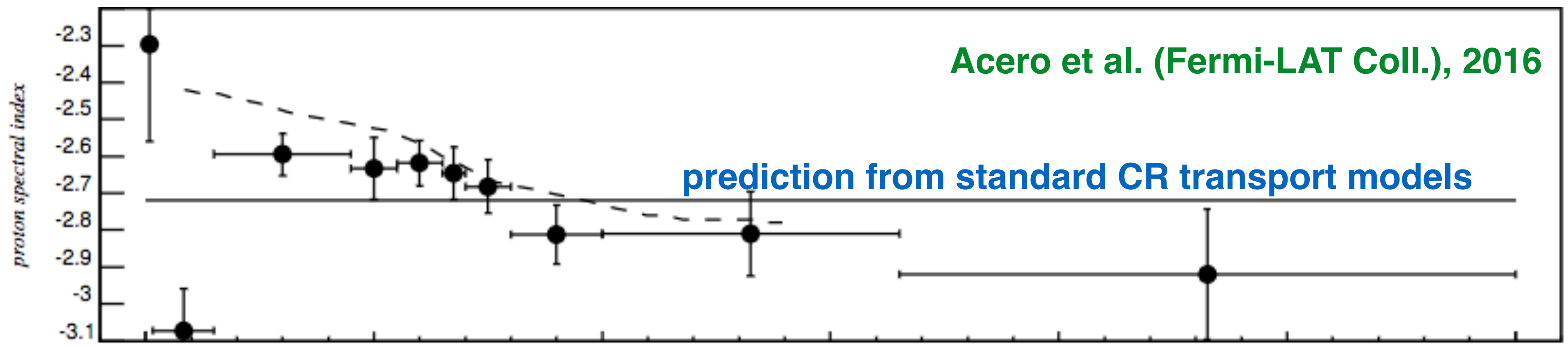


Anisotropic diffusion

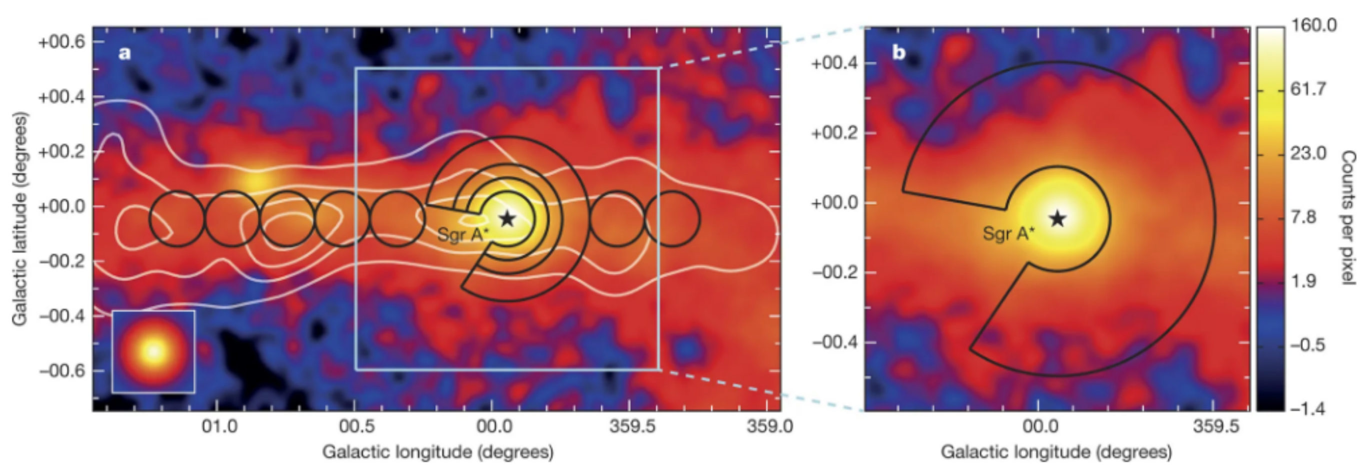
Cerri, Gaggero, Vittino, Evoli, Grasso, 2017

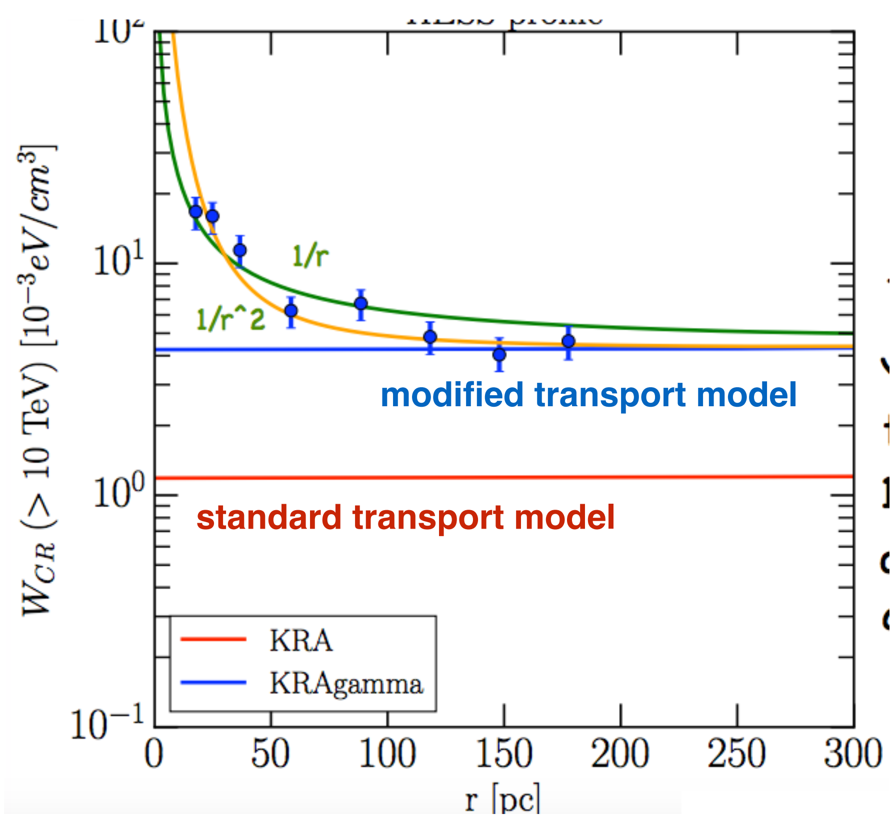
CR proton density - the role of CR transport

Results **do not agree** with the prediction from **standard CR transport models**.



These alternative transport models could play **a major role** in assessing the properties of **CR sources in the inner Galaxy** (e.g., the Pevatron discovered by HESS in 2016)

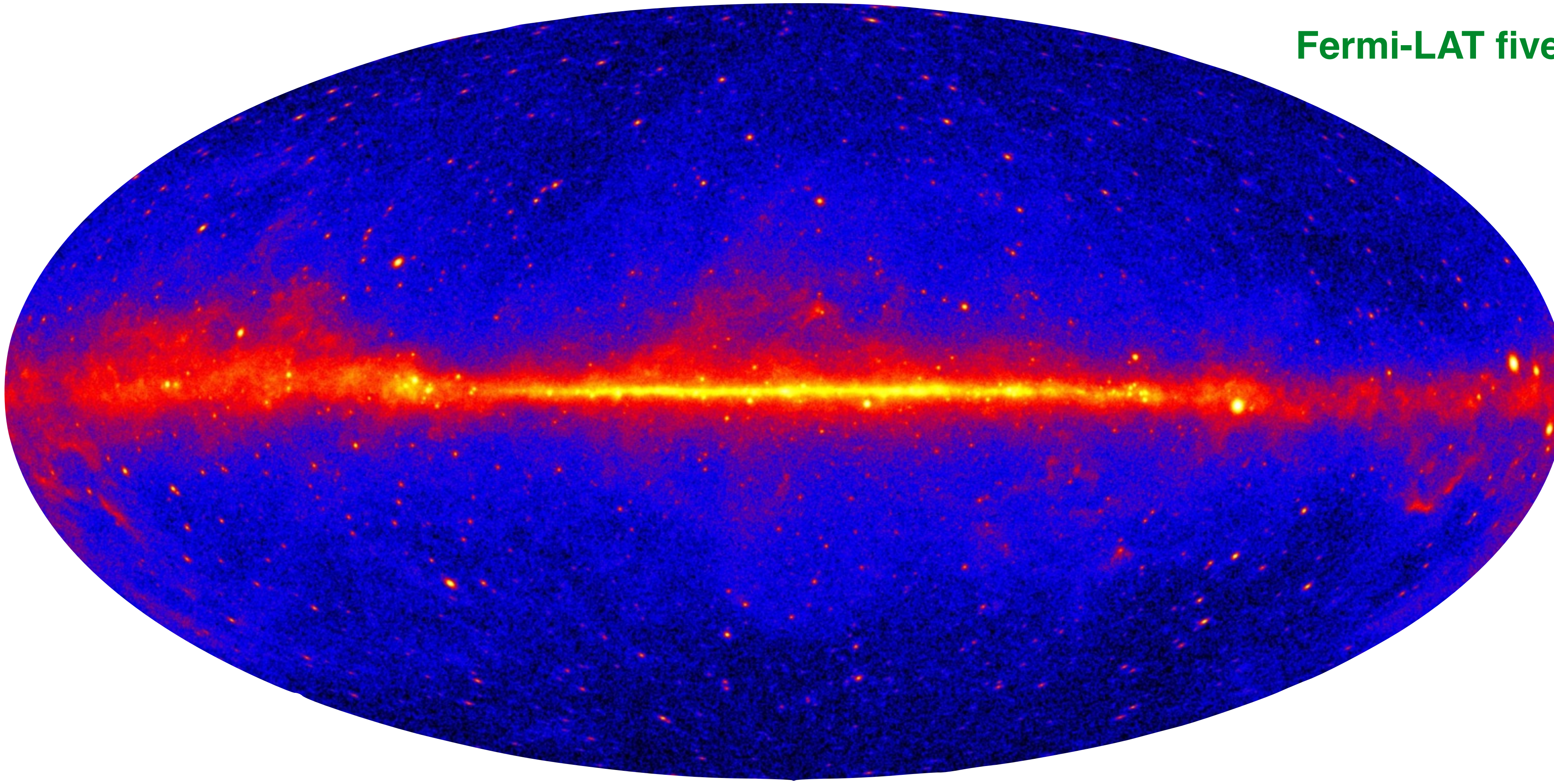




Abramowski et al. (HESS Collaboration), 2016

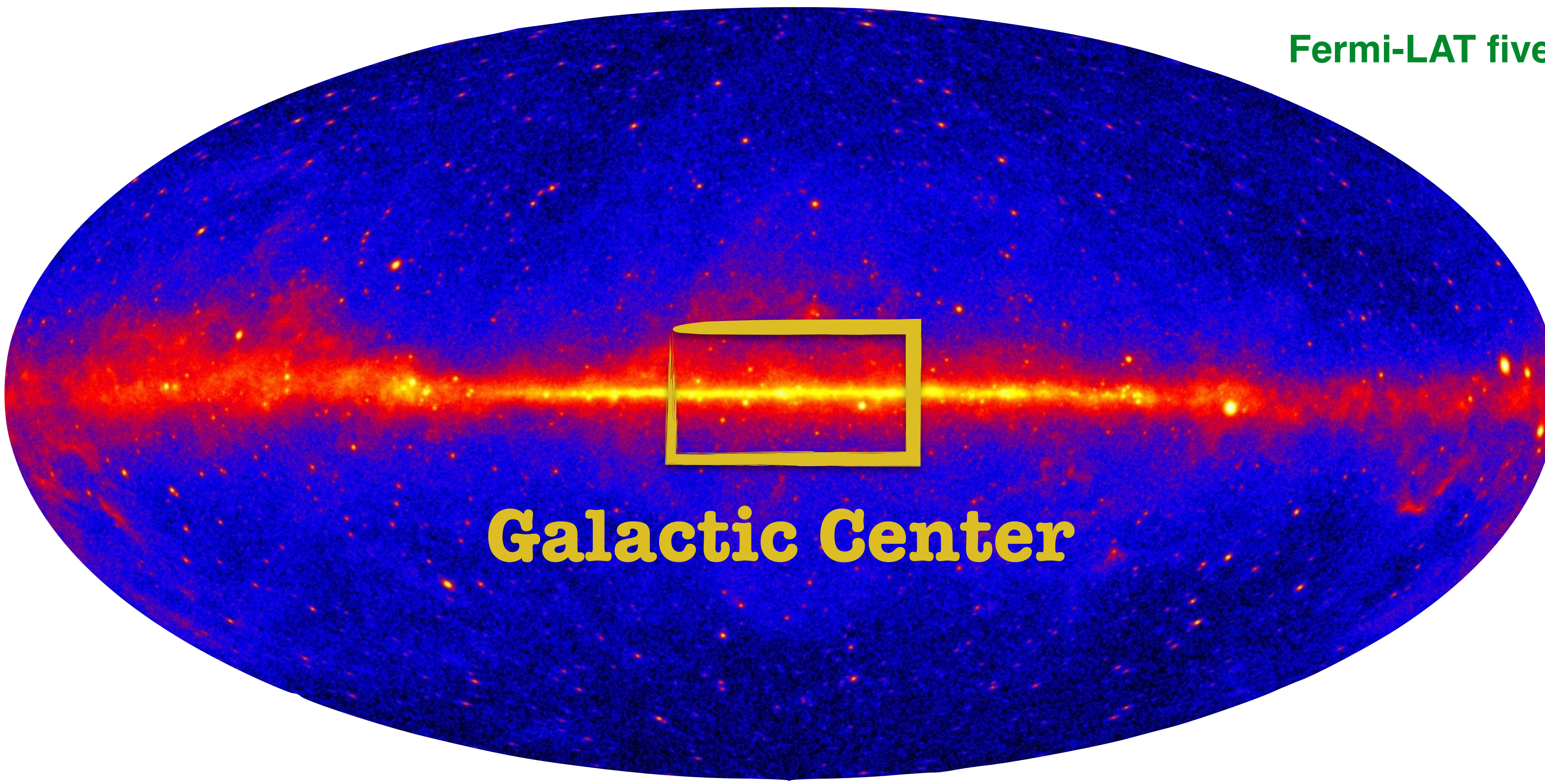
Gaggero, Grasso, Marinelli, Taoso, Urbano, 2017

Gamma-ray data



Fermi-LAT five-year skymap

Gamma-ray data



Fermi-LAT five-year skymap

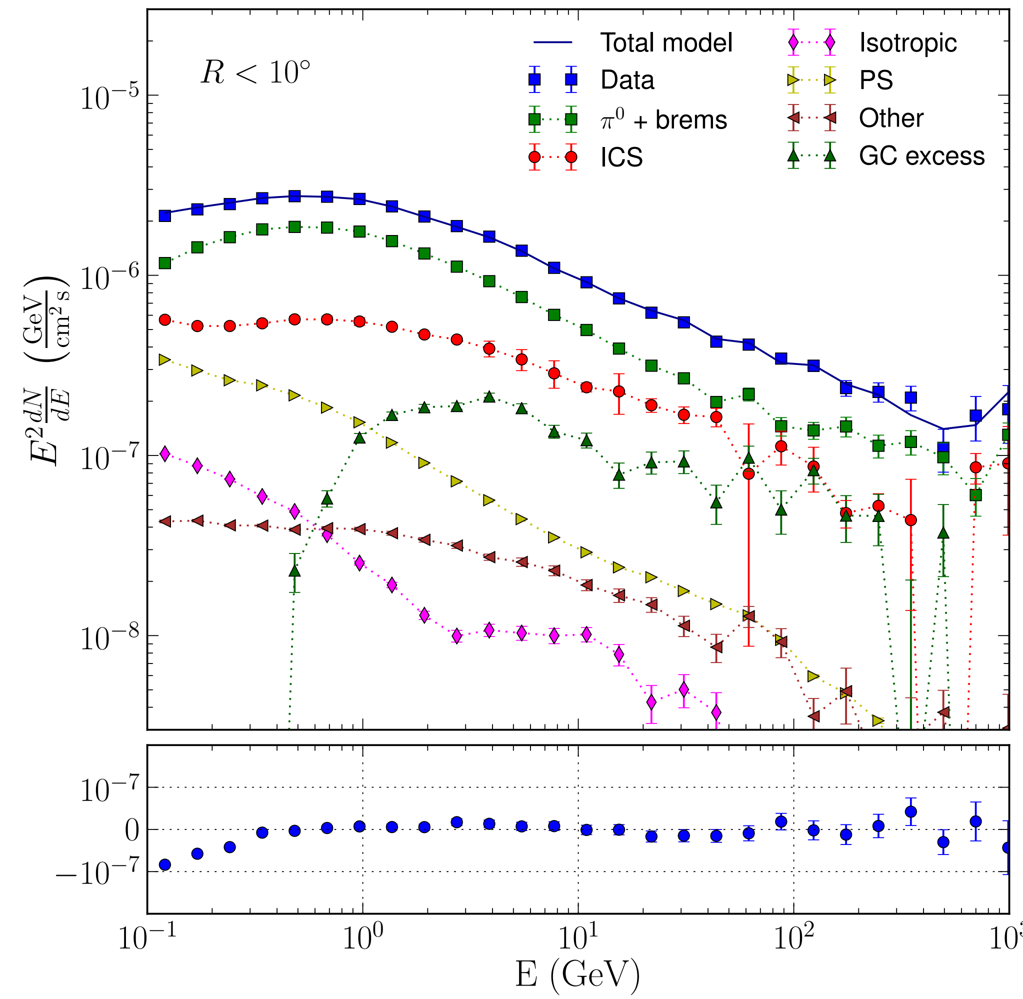
- Highest gamma-ray flux from **dark matter annihilation**
- But, it is a very **complicated environment**
- Claims of an **excess** (with respect to standard astrophysical emissions) since 2009

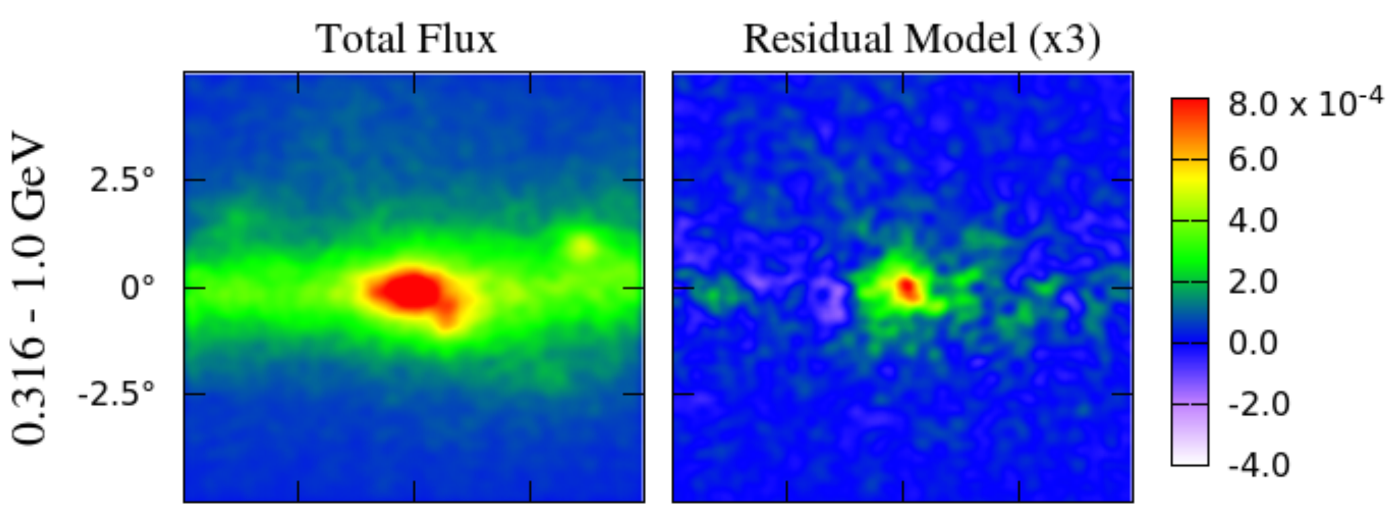
The Galactic Center excess

An important topic in astroparticle research:

Vitale, Morselli 2009, Hooper, Linden 2011, Hooper, Goodenough 2011, Boyarsky, Malyshev, Ruchayskiy 2011, Abazajian, Kaplinghat 2012, Macias, Gordon, 2014, Abazajian et al. 2014, Daylan et al. 2014, Casandjian 2014, Calore, Cholis, Weniger 2015, Huang, Esslin, Selig 2015, Carlson et al. 2015, Ajello et al. 2015, De Boer et al. 2016, Macias et al. 2016, Ackermann et al. 2017, Leane, Slatyer 2019, Cholis et al. 2020

Ackermann et al. (Fermi-LAT Coll.), 2017





Dylan et al. 2014

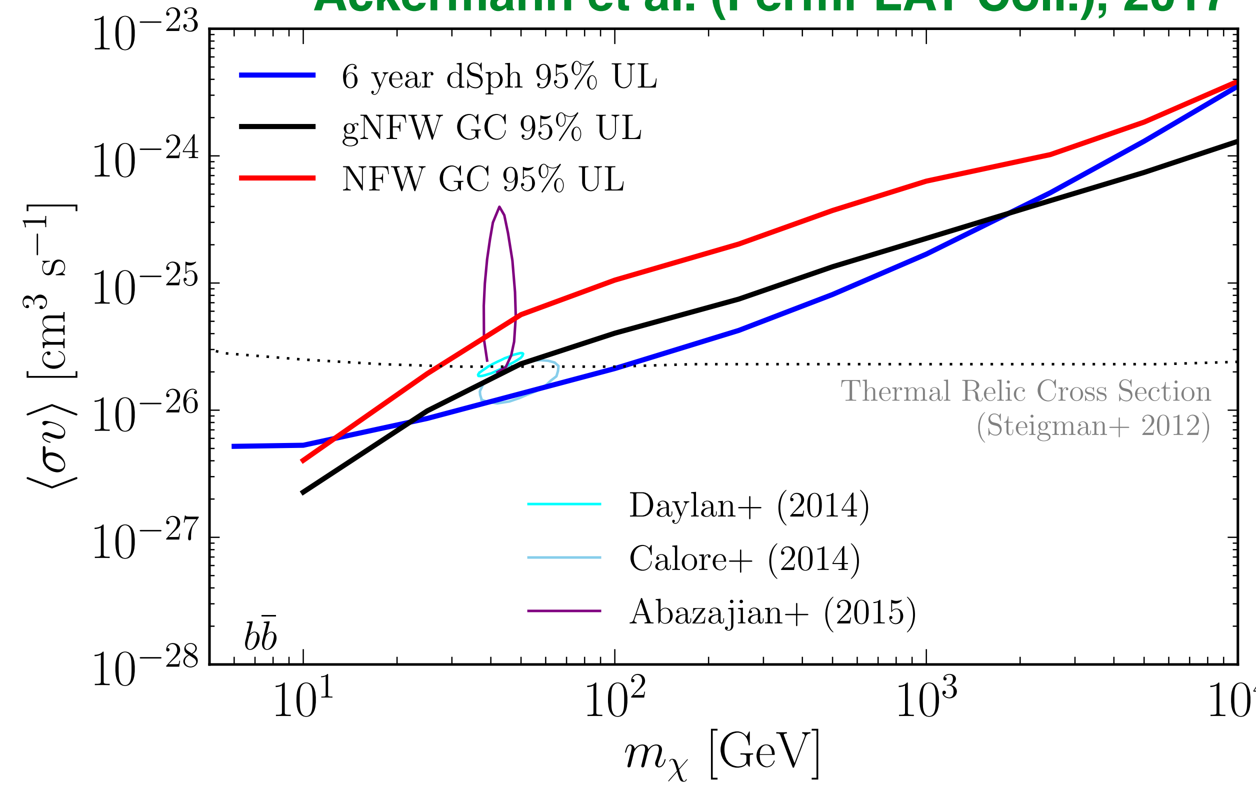
- Removal of astrophysical emission: **spatial and spectral template subtraction**
- Morphology : approximately **spherical**

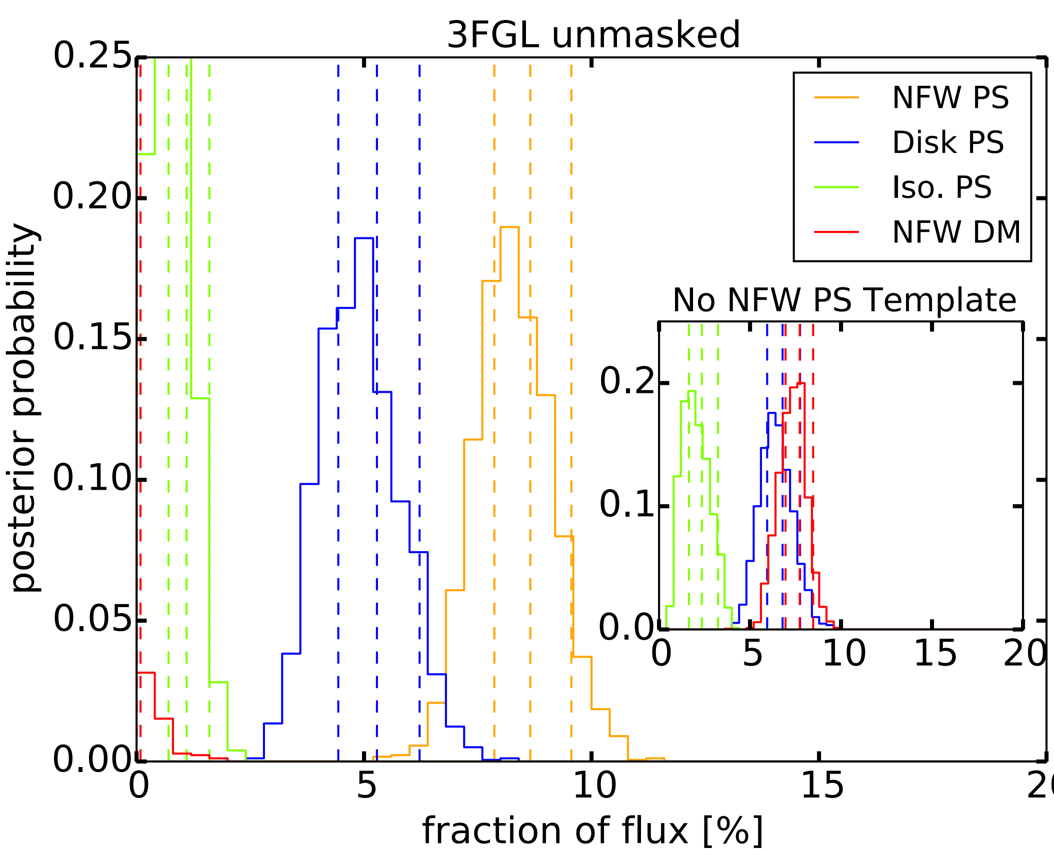
Dark matter vs. point sources

Dark matter

- cross sections **close to the thermal value** seem to provide the best fit to the excess.
- The Fermi-LAT coll. finds similar excesses across the Galactic plane (-> **upper limits**)

Ackermann et al. (Fermi-LAT Coll.), 2017





(Unresolved) point sources

- spectrum and morphology of the excess are also **compatible with millisecond pulsars**.
- Unresolved sources can be **constrained**
  - with **point-count statistics** (**Lee et al. 2015**, but see also **Leane, Slatyer 2019**)
  - with **wavelet analysis** (**Bartels et al. 2015**)

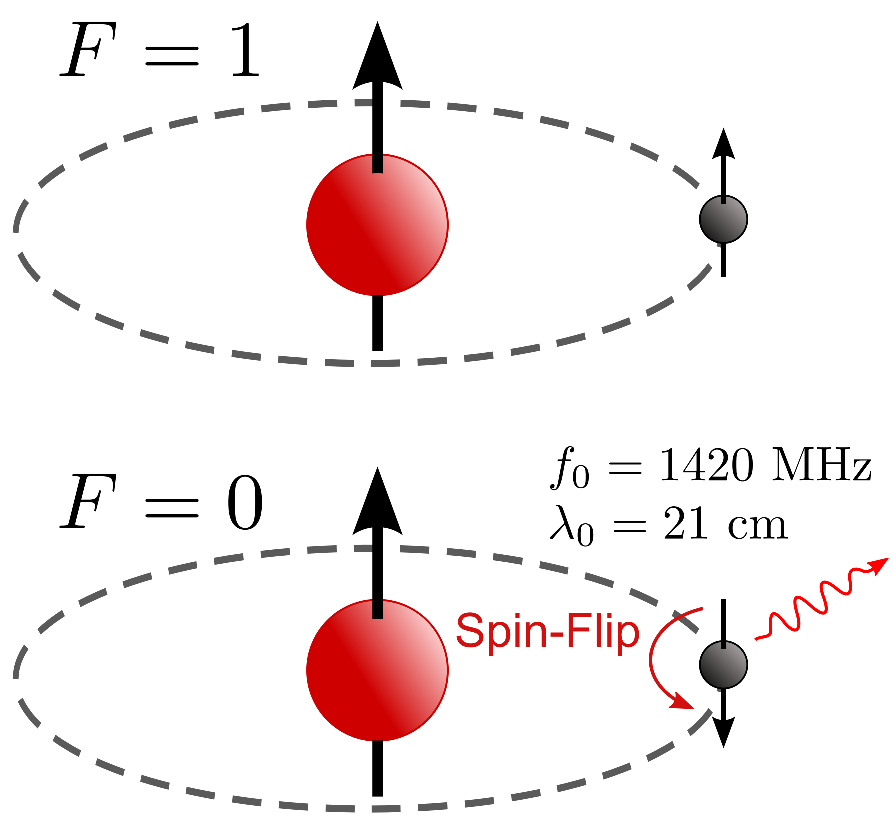
Outline

- **Motivation**
  - ✦ Why are we interested in the interstellar gas distribution?
- **The problem** : inferring the gas distribution
  - ✦ How do we observe the gas? How can we use such observation to understand how gas is distributed?
- **Our method** : information field theory
  - ✦ What is information field theory? How do we intend to use it to infer the gas distribution?

Interstellar gas

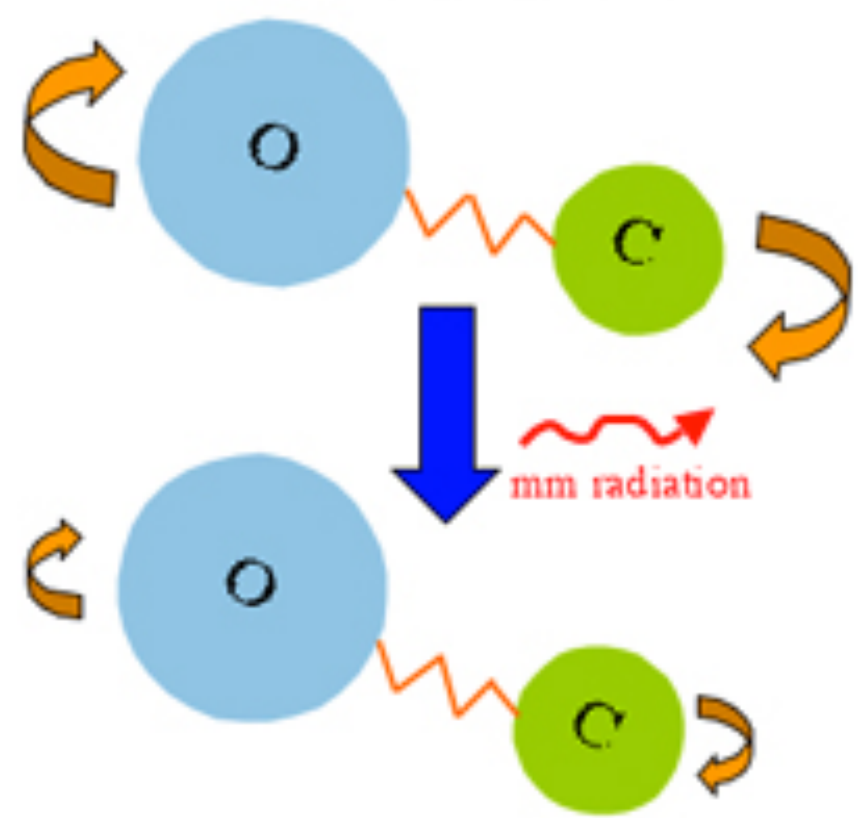
Both the atomic and the molecular hydrogen are traced through **line emissions**

atomic hydrogen



transition between the hyperfine levels of the ground state determine the **21 cm line emission**

molecular hydrogen



H2 does not emit any radiation, so one must use a **tracer** (typically CO).

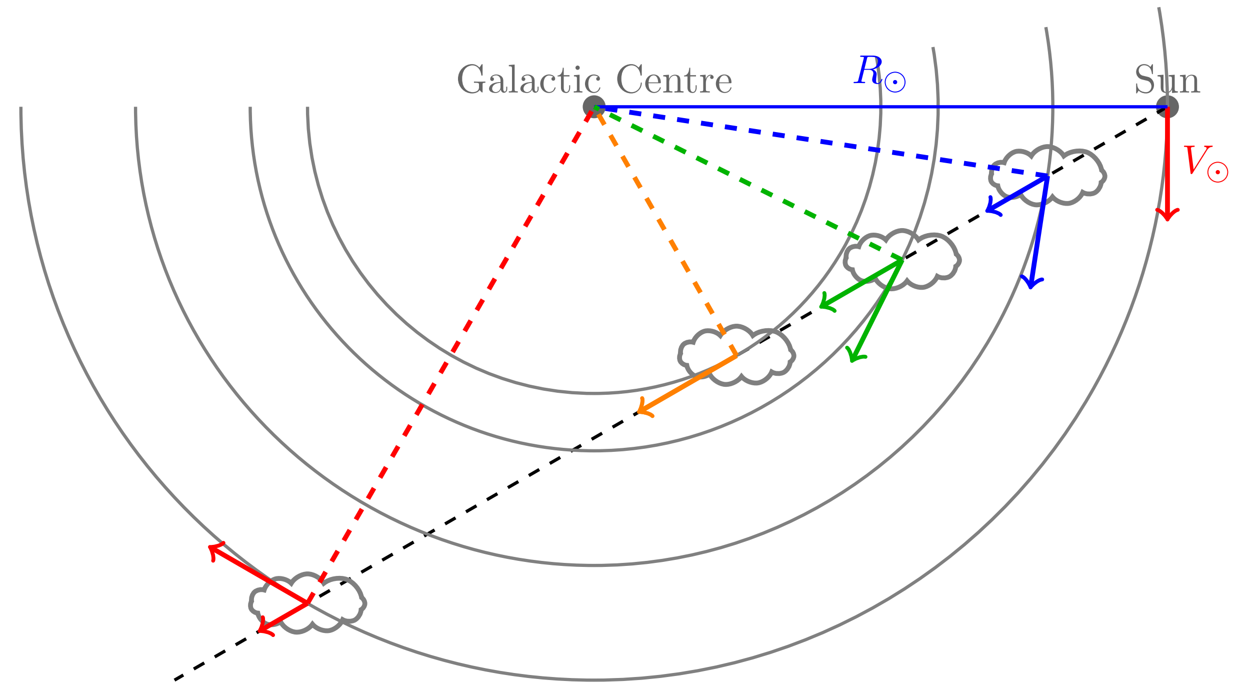
CO emits lines in the **transition between rotational states**

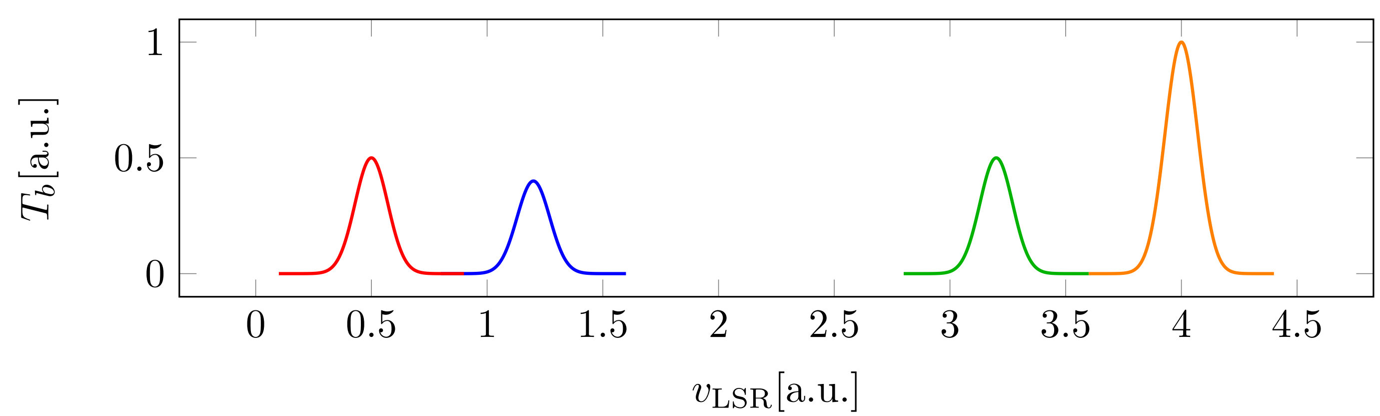
Velocity spectroscopy

- The line emission is **Doppler-shifted** because the gas cloud is rotating

- Different line-of-sight velocities are associated to **different distances**

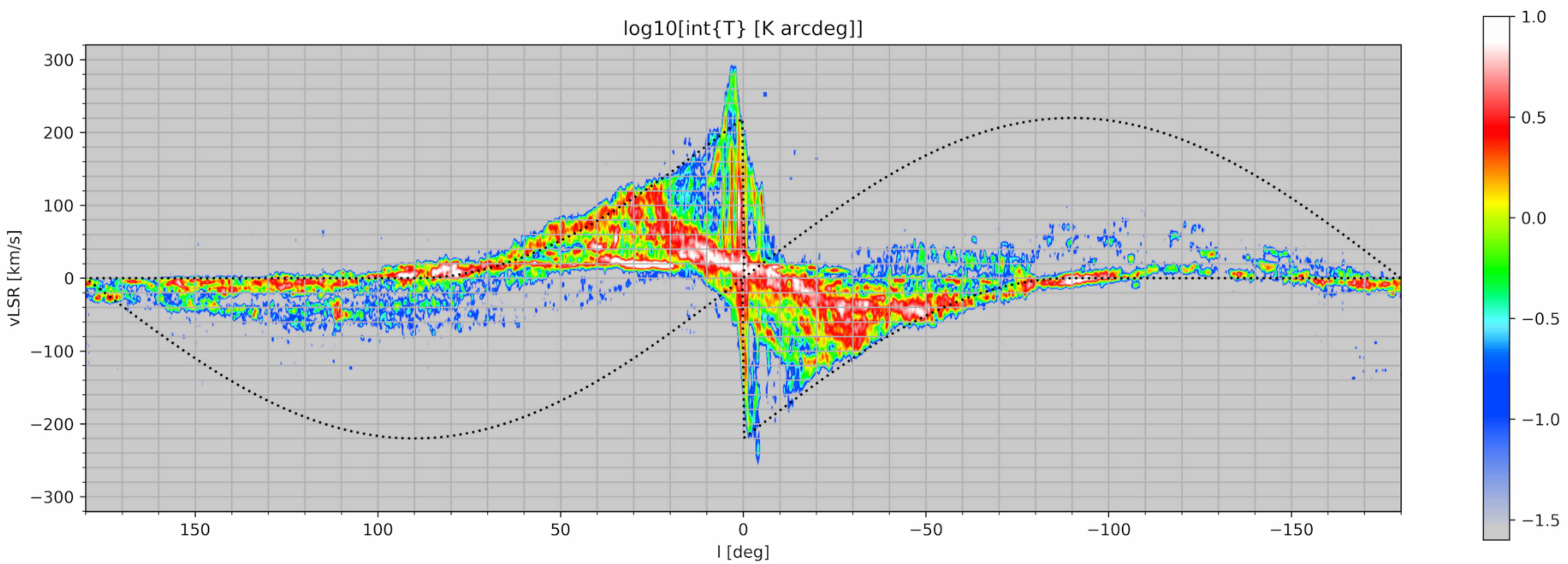
- This allows for a **deprojection** of the observations (from velocity to distance)





CO line data

longitude-velocity map of the **CO line emission** in the Galactic disk

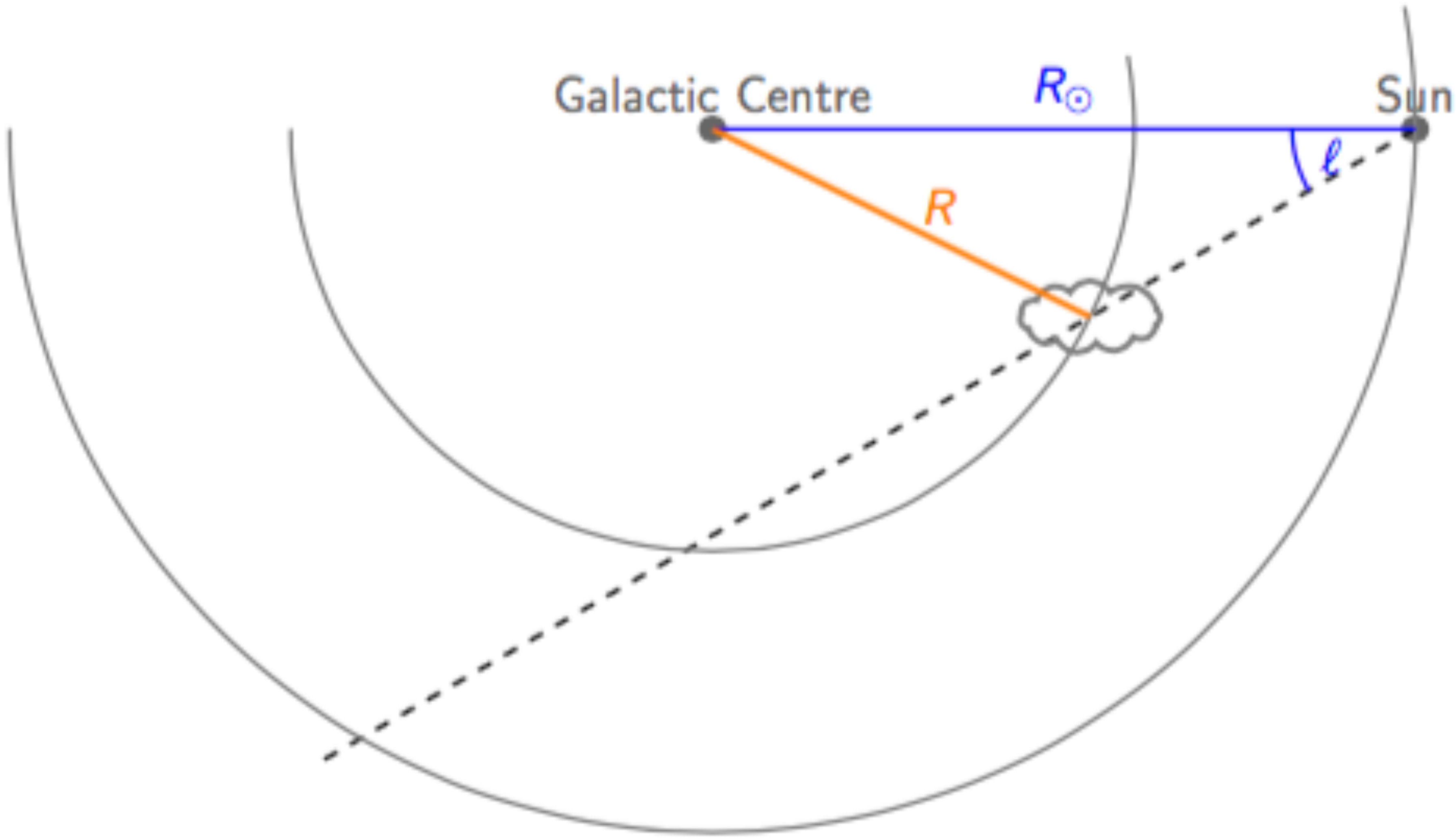


Dame et al. 2001

How do we deproject these data to get a map of the gas responsible for the emission?

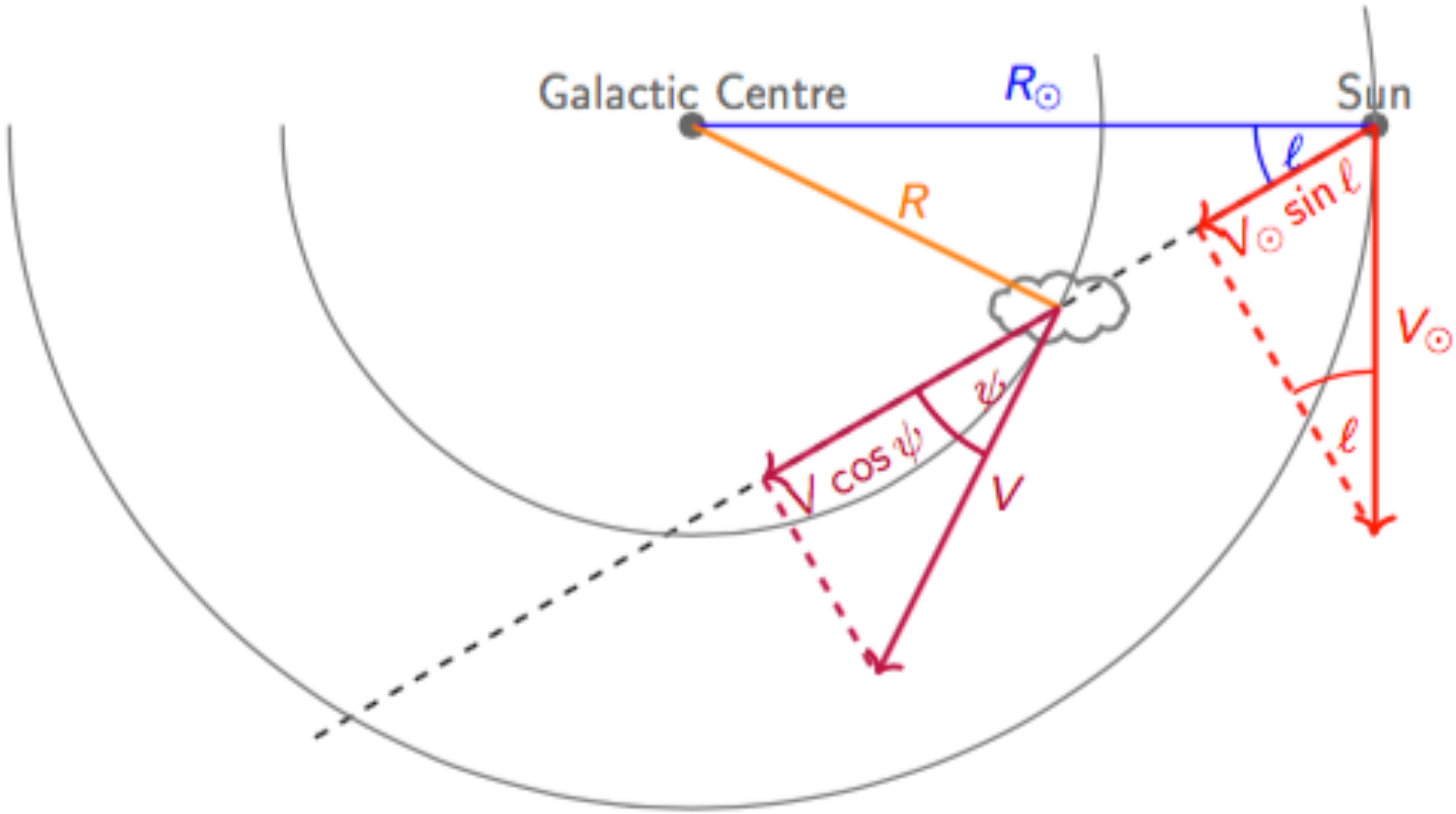
Velocity spectroscopy

Assuming that the motion is **purely circular**:



Velocity spectroscopy

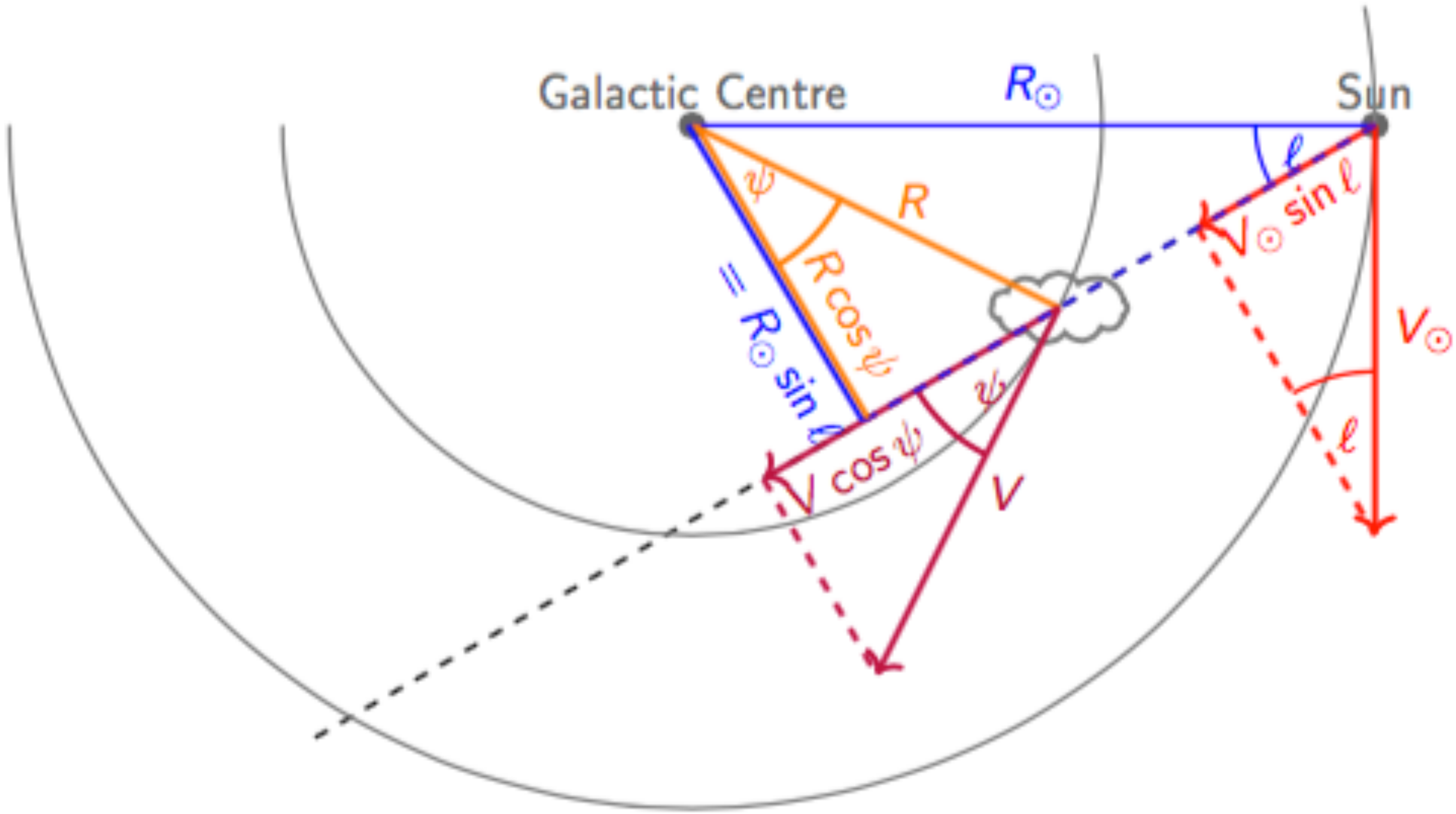
Assuming that the motion is **purely circular**:



$$v_{LSR}(R, \ell) = V(R) \cos \psi - V_\odot \sin \ell$$

Velocity spectroscopy

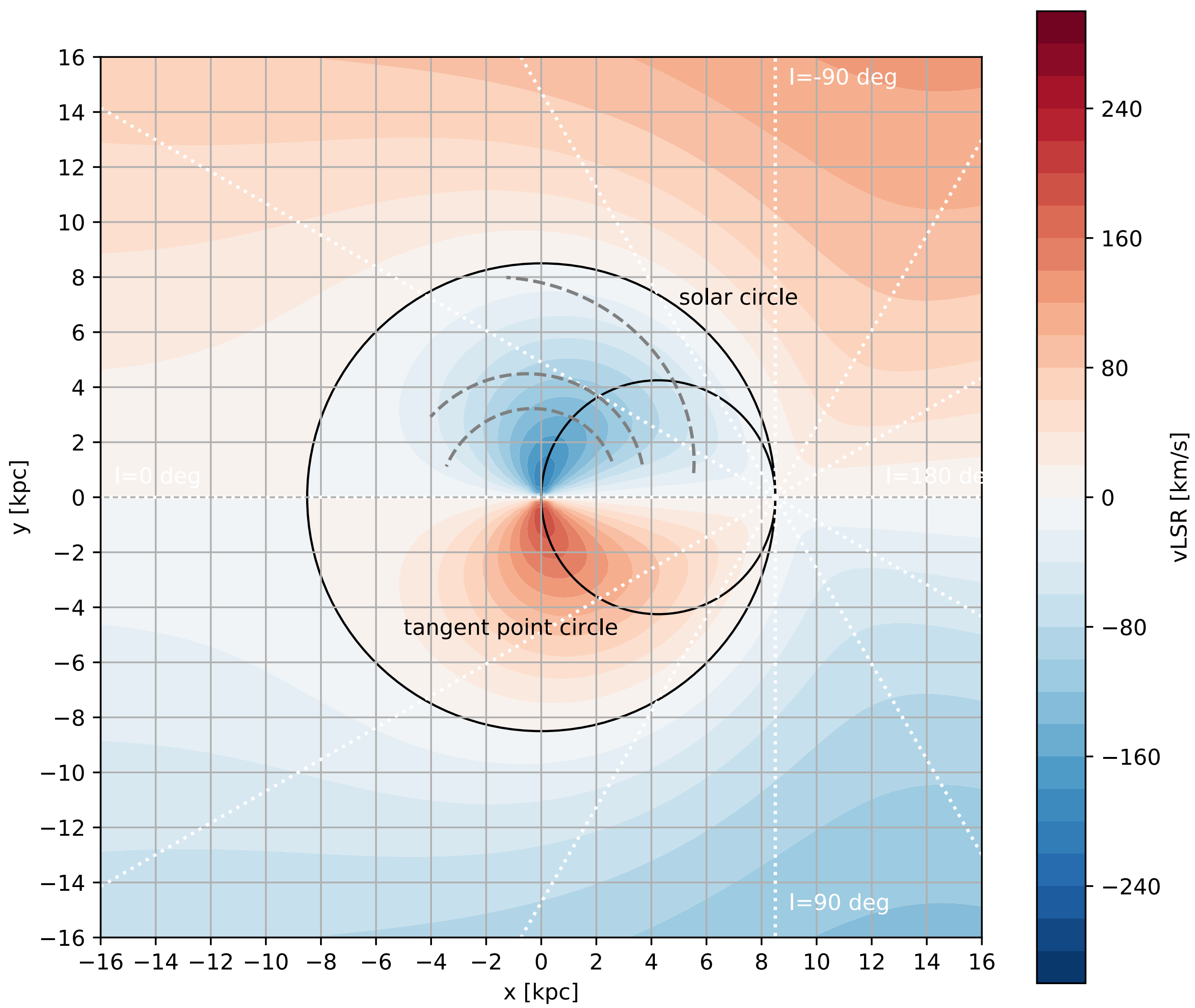
Assuming that the motion is **purely circular**:



$$v_{LSR}(R,\ell) = V(R)\cos\psi - V_\odot \sin\ell = \sin\ell\left(\frac{R_\odot}{R}V(R) - V_\odot\right)$$

Velocity spectroscopy

Given a model of the Galactic rotational curve V(R), one can **associate a velocity to each distance** measured along a line of sight



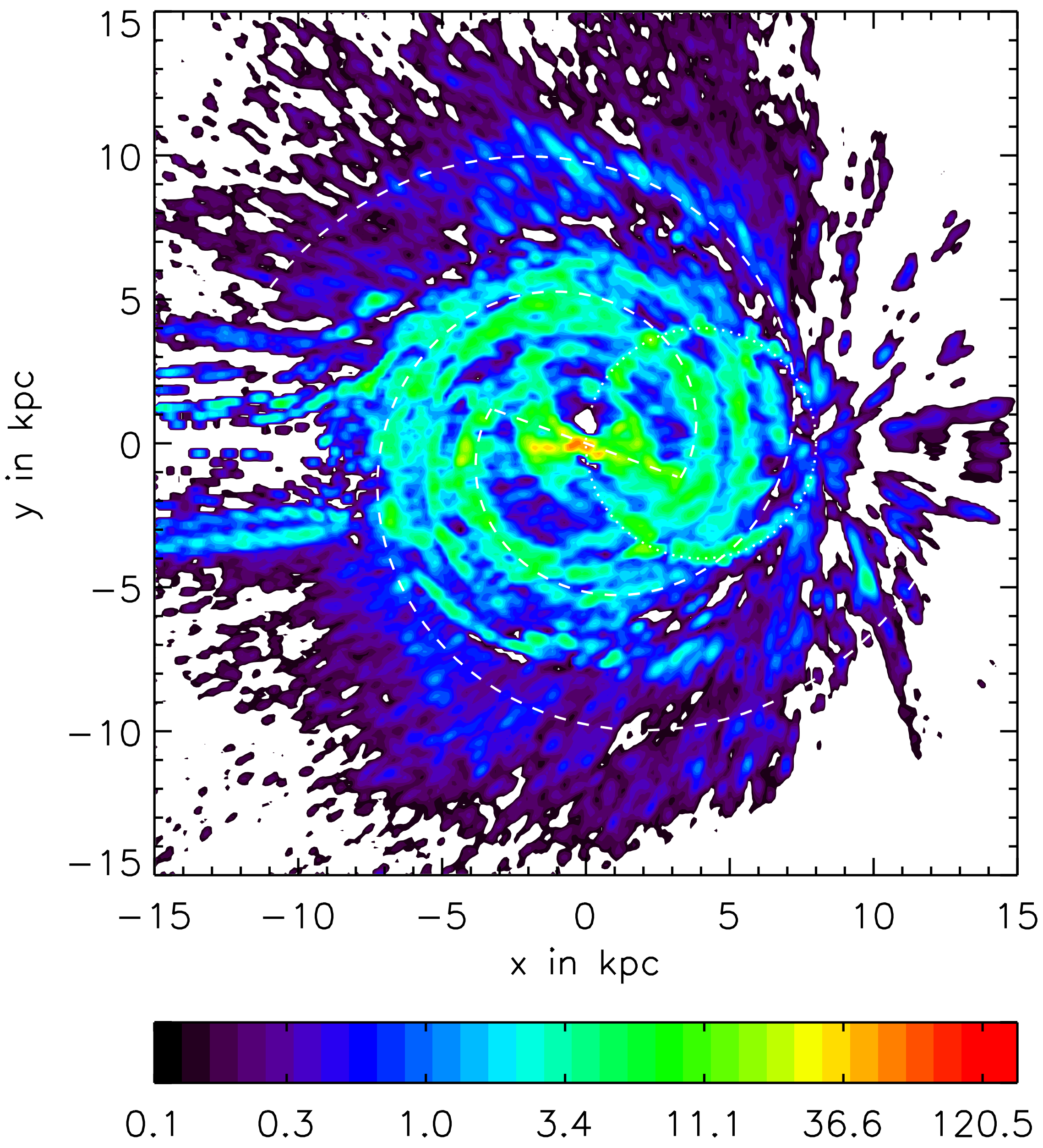
Issues

- **The rotation curve of the Galaxy is uncertain** : in the inner Galaxy, deviations from a purely circular motion might be present.
- **Near - far ambiguity** : within the Solar circle, objects at different distances can have the same velocity.
- **Lack of resolution** for longitudes of 0^{o} and 180^{o}, if motion is assumed to be purely circular

Previous attempts at mapping the gas

Pohl, Englmeier, Bissantz, 2008

- **Rotational curve** derived from a **hydrodynamic gas flow model** (non-circular motion in the inner Galaxy)
- This non-circular motion **fixes the lack of resolution** at the **Galactic Center**
- **Fit** to the longitude-velocity diagram with a **sum of Gaussians**, along **individual line-of-sights**.
- The gaussians in velocity space are then **deprojected in coordinate space**

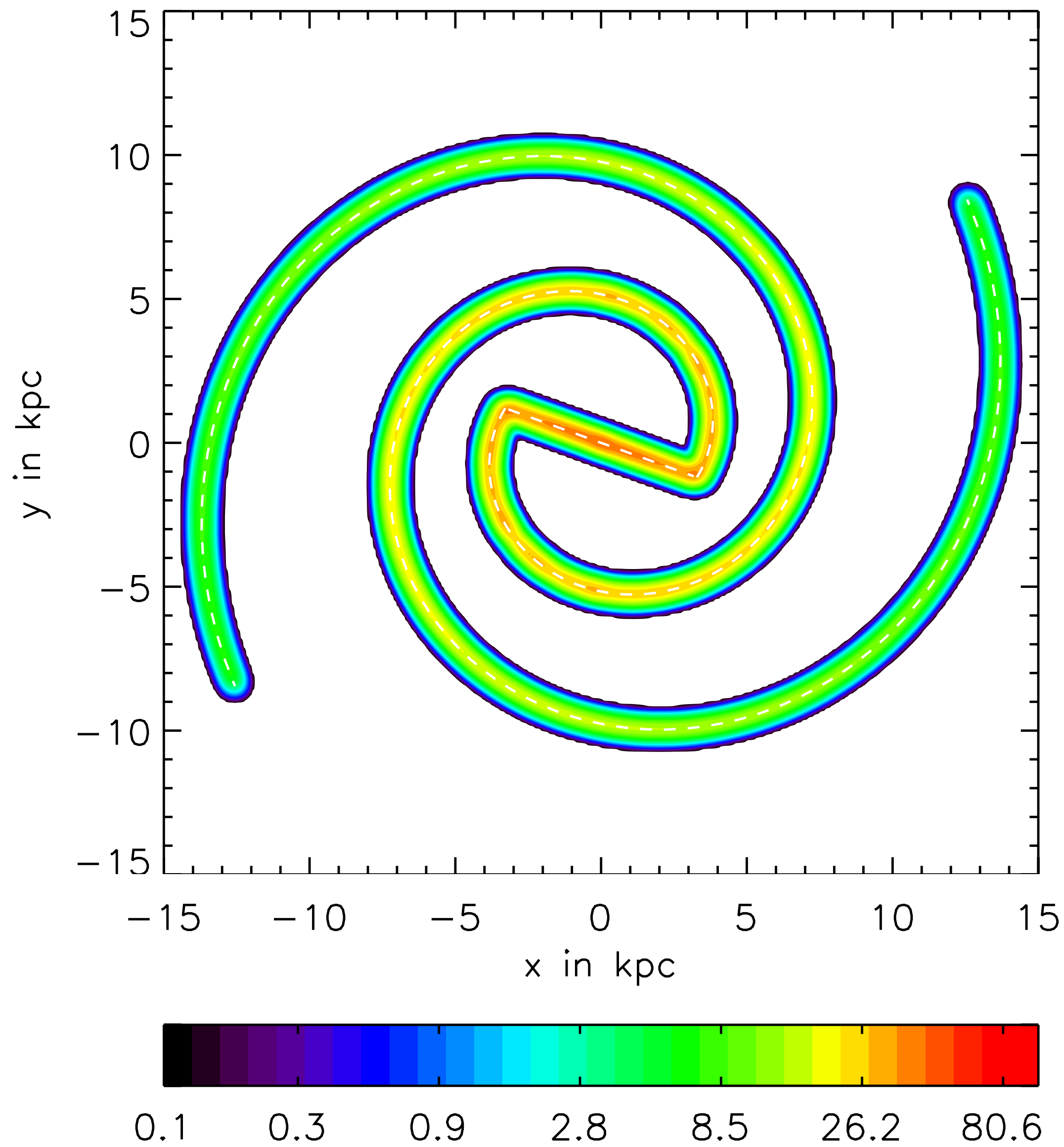


Previous attempts at mapping the gas

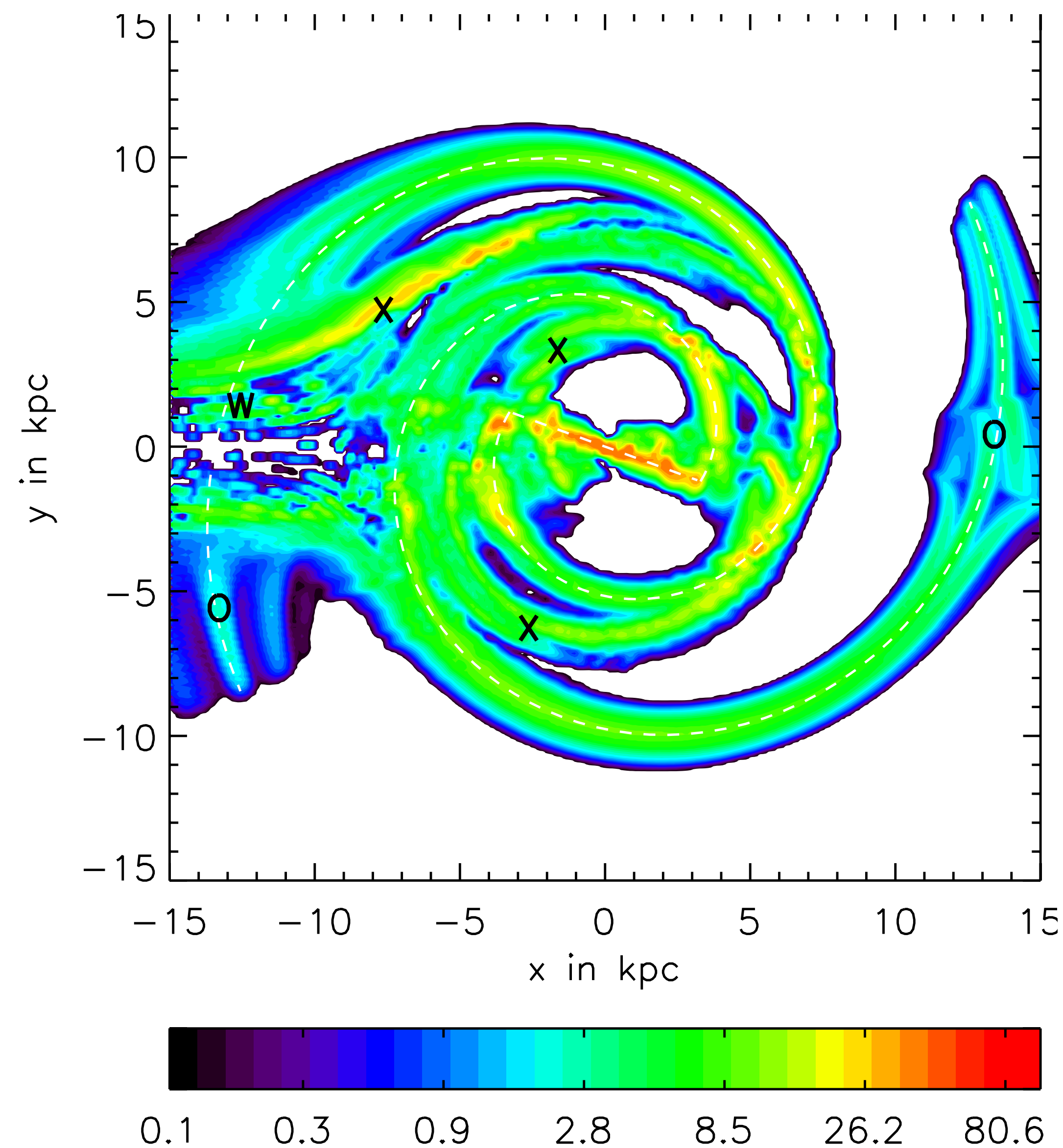
Pohl, Englmeier, Bissantz, 2008

This method introduces **artefacts**, related to the **issues associated to the deprojection technique**.

simulated gas map



reconstructed map

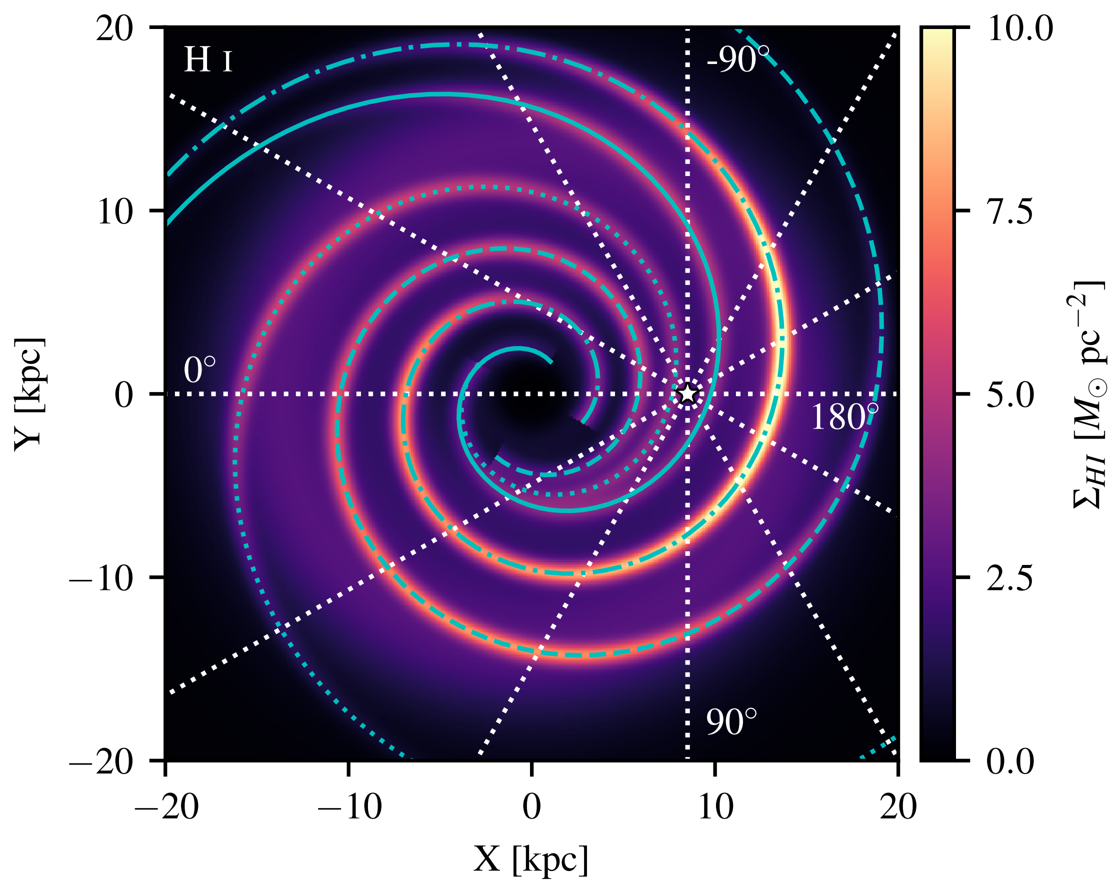


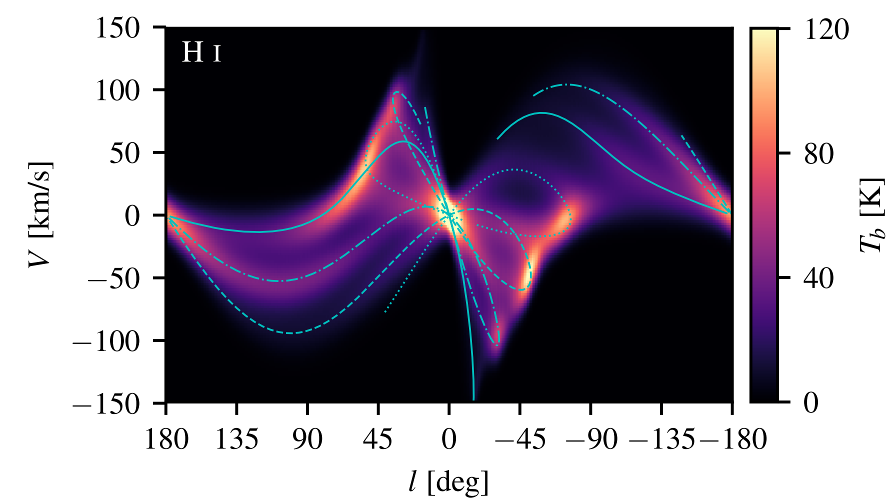
Previous attempts at mapping the gas

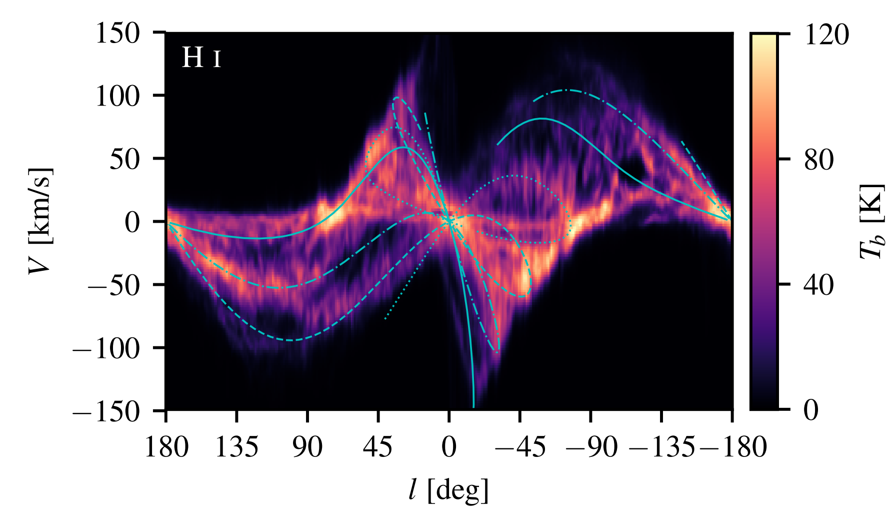
forward modelling

Johannesson, Porter, Moskalenko 2018

- Adopt a **parametric model for the gas density** and predict the longitude-velocity diagram, which is then compared to the actual data
- **PRO** : Address **near-far ambiguity** and **lack of resolution**





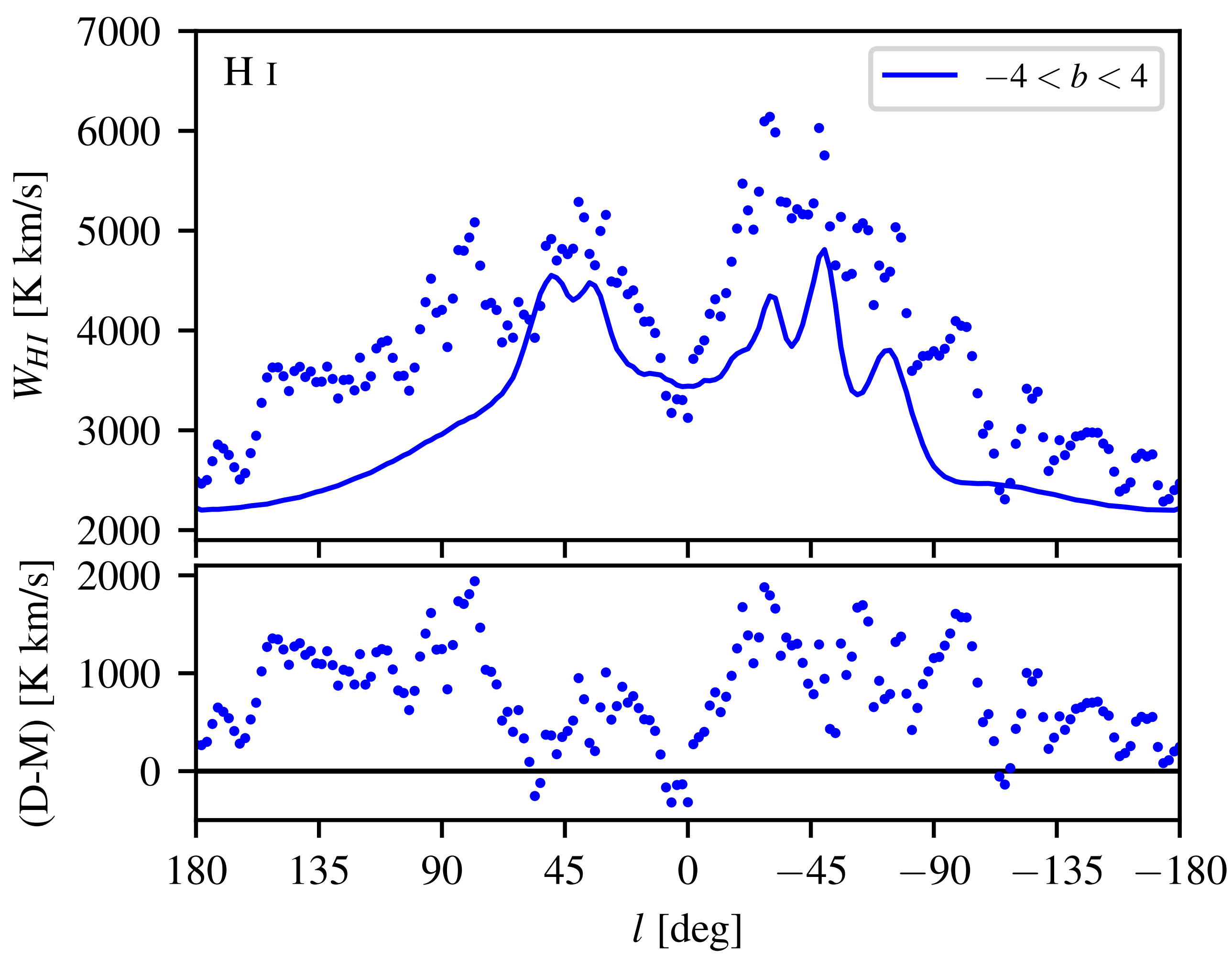


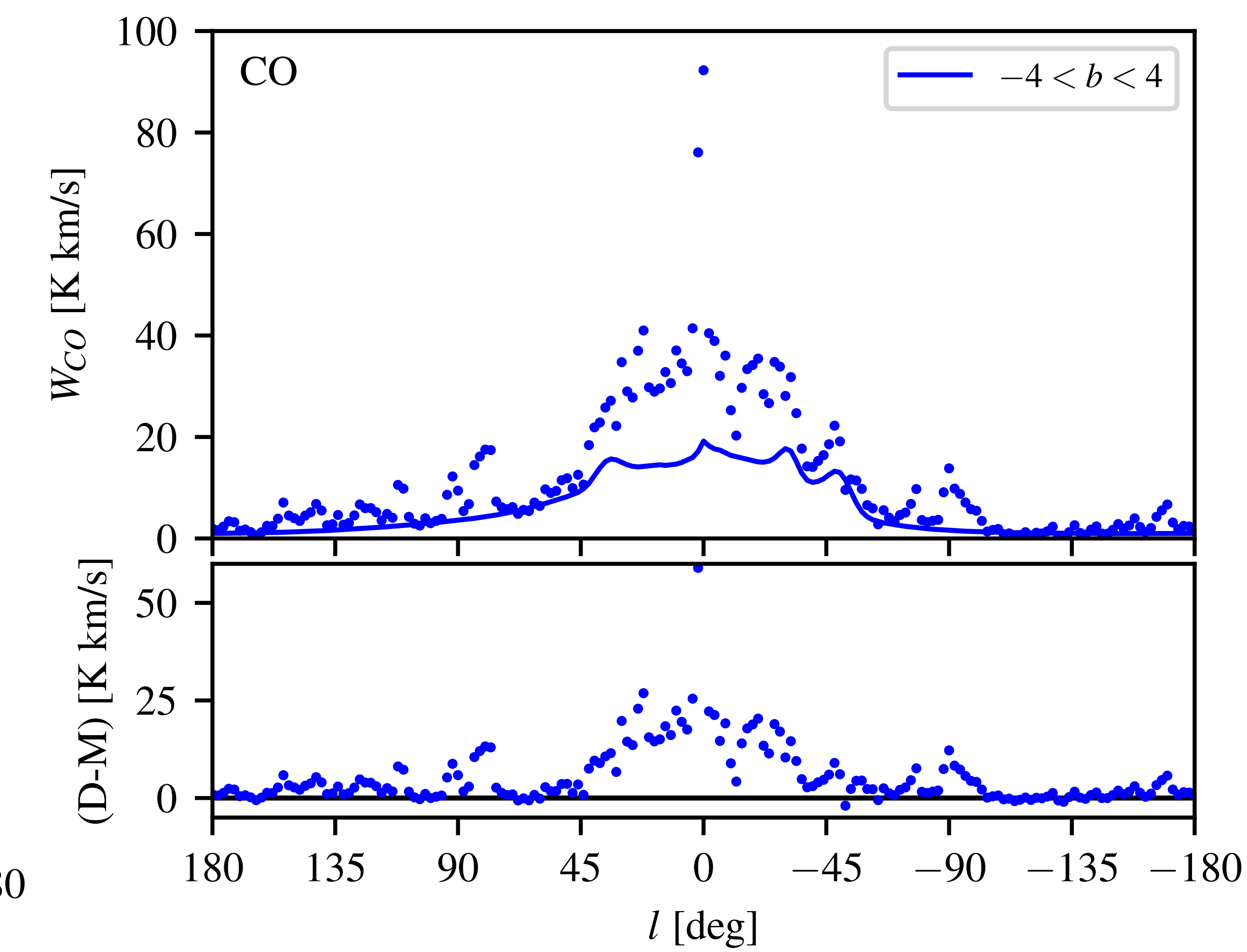
Previous attempts at mapping the gas

forward modelling

Johannesson, Porter, Moskalenko 2018

- Adopt a **parametric model for the gas density** and predict the longitude-velocity diagram, which is then compared to the actual data
- **PRO** : Address **near-far ambiguity** and **lack of resolution**
- **CON** : **Not flexible enough**





Outline

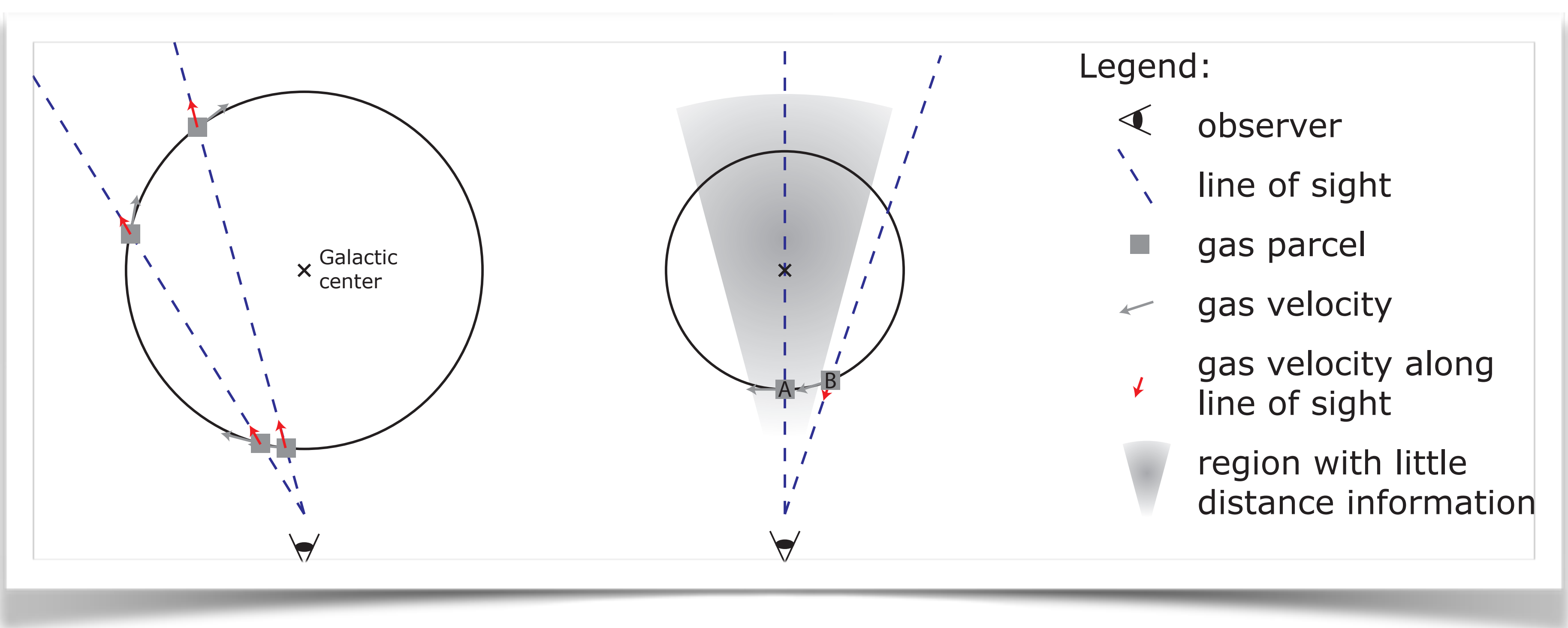
- **Motivation**
  - ✦ Why are we interested in the interstellar gas distribution?
- **The problem** : inferring the gas distribution
  - ✦ How do we observe the gas? How can we use such observation to understand how gas is distributed?
- **Our method** : information field theory
  - ✦ What is information field theory? How do we intend to use it to infer the gas distribution?

Our approach

- Forward methods are **not flexible enough** to be really accurate

- Methods based on the **deprojection** of the longitude - velocity diagrams suffer from **several artefacts**. Moreover, in such methods, **line of sights are treated separately**

- However, **correlations are expected to be present**, as due to:
    - **large scale structure (disk, bulge, spiral arms, etc …)**
    - **gravitational collapse**
    - **turbulence in the interstellar medium**



we need to keep track of correlations and exploit them!

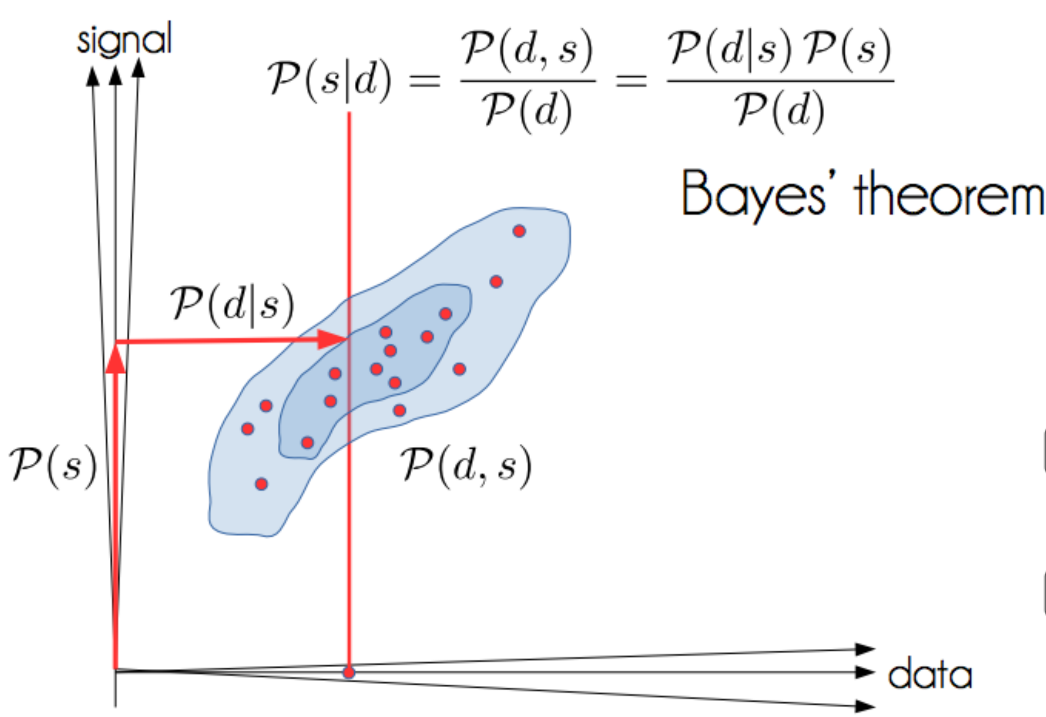
Information field theory

We use **Information field theory**, a Bayesian method to infer a signal from data

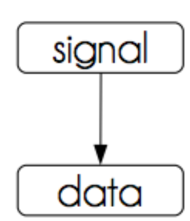
T.A. Enßlin et al. 2009

$$d = f(s) + n$$

- d : data (measured spectra)
- s : signal (gas density)
- n : noise (uncertainty in brightness temperature)



Plot by Torsten Enßlin



$$\mathcal{P}(s|d) = \frac{e^{-\mathcal{H}(d,s)}}{\mathcal{Z}_d}$$

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

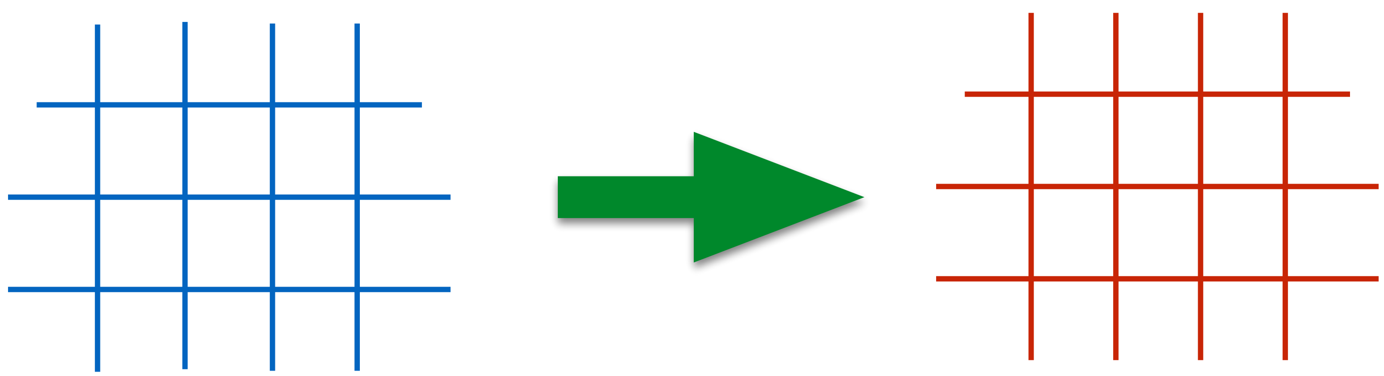
$$\mathcal{Z}_d = \int \mathcal{D}s\, \mathcal{P}(d|s)\, \mathcal{P}(s)$$ Partition function

The response function

We assume the measurement to be **linear** in both the **signal** and the **noise**

$$d = f(s) + n \quad \rightarrow \quad d = R\, s + n$$

The response function maps the signal space into the data space



$s(x_\alpha, y_\beta, z_\gamma)$ $\qquad$ $d(\ell_i, b_j, v_k)$

The response function is built by **mapping each bin** of the discretised signal space into the data space

strong dependence on the assumed **rotation curve** (we assume a **purely circular motion**)

More formally:

$$R^{\alpha\beta\gamma}_{ijk} = \frac{\epsilon}{\Delta\ell\Delta b\Delta v} \int \mathrm{d}\ell \int \mathrm{d}b \int \mathrm{d}v \int \mathrm{d}t\ \theta(\text{if } \vec{r} \text{ in } \{ijk\})\, \delta(v - v_{\mathrm{LSR}}(\vec{r}))|_{\vec{r}}$$

$\mathrm{d}t$: coordinate along the line-of-sight

Example 1 : Wiener filter

If we assume both the **signal and the noise to be Gaussian** with **known covariances**

$$p(s) = \mathcal{G}(s, S) \qquad p(n) = \mathcal{G}(n, N)$$

Then the posterior distribution can be found **easily**:

$$\begin{aligned} p(s|d) &= \mathcal{G}(s - m, D) \\ m &= D\, j \\ D &= \left(S^{-1} + R^{\dagger} N^{-1} R\right)^{-1} \\ j &= R^{\dagger} N^{-1} d \end{aligned}$$

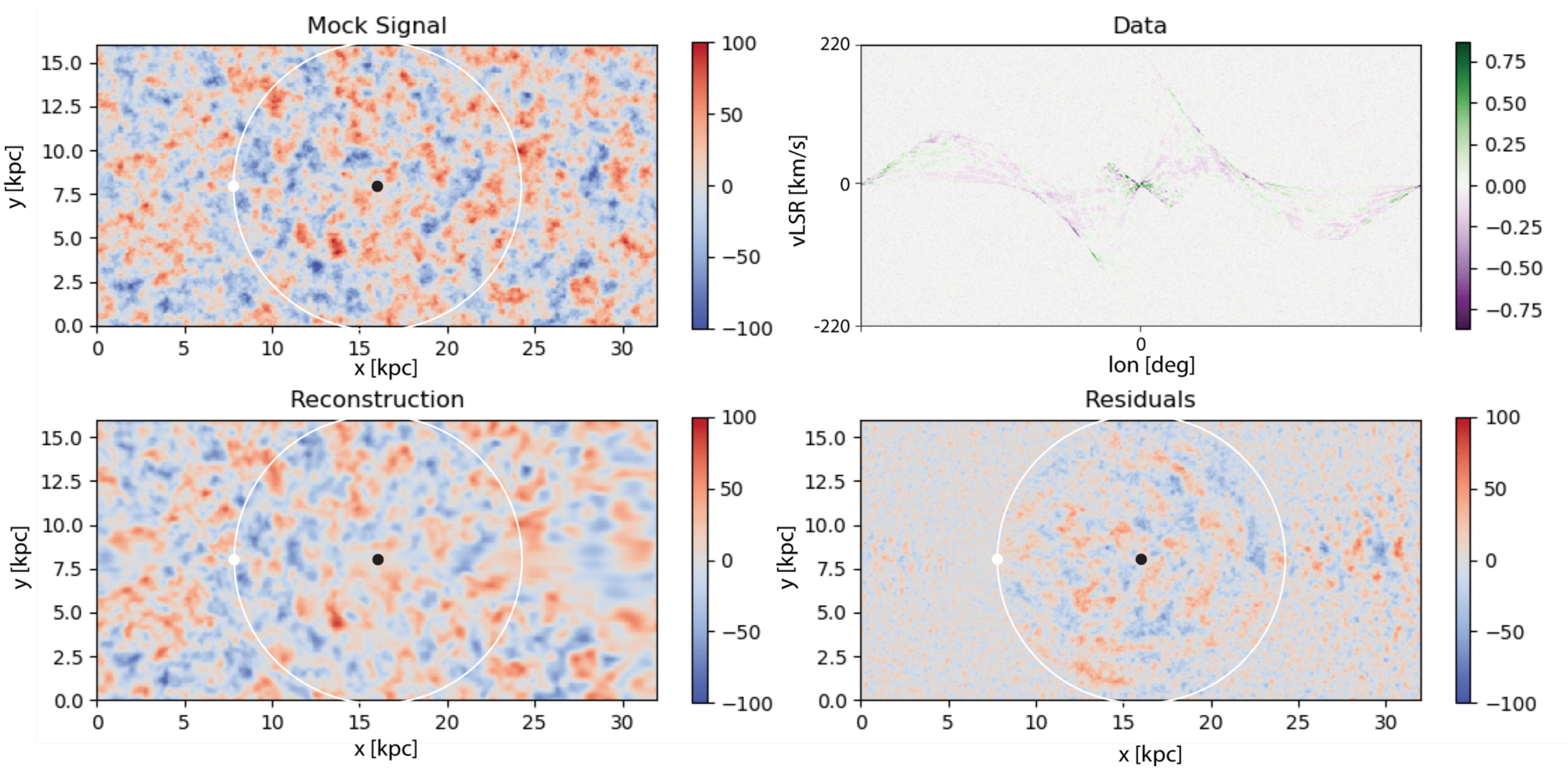
posterior mean (Wiener filter)

information propagator

information source

Example 1 : Wiener filter

mock signal : gaussian random field
signal model : gaussian random field

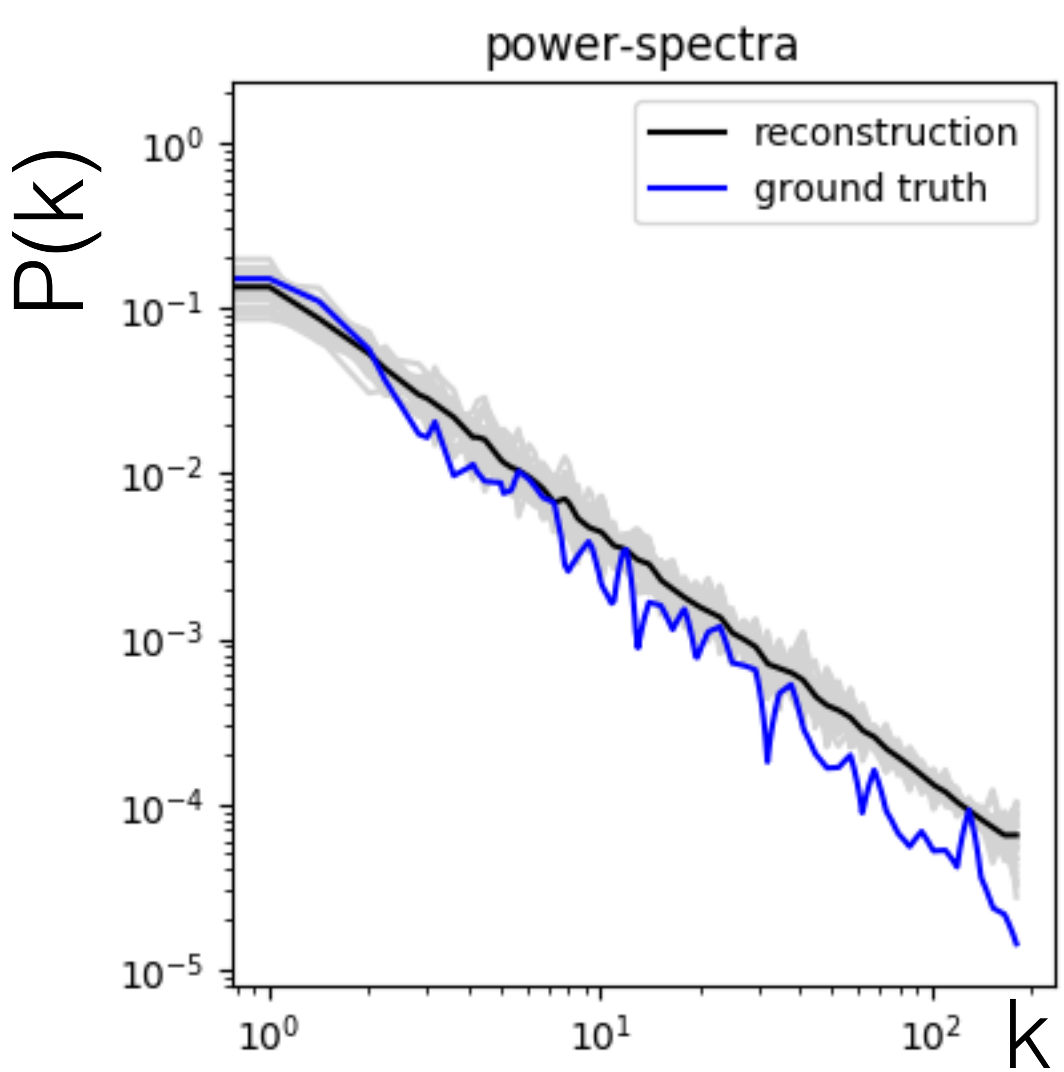


results obtained with **the NIFTy package**

Example 2 : Critical Wiener filter

- The problem with the Wiener filter approach is that it is **only usable if we know the power spectrum of the signal**.

- The solution is to **treat the power spectrum as a random variable**, within the framework of a **generative model**



Example 2 : Critical Wiener filter

- The problem with the Wiener filter approach is that it is **only usable if we know the power spectrum of the signal**.

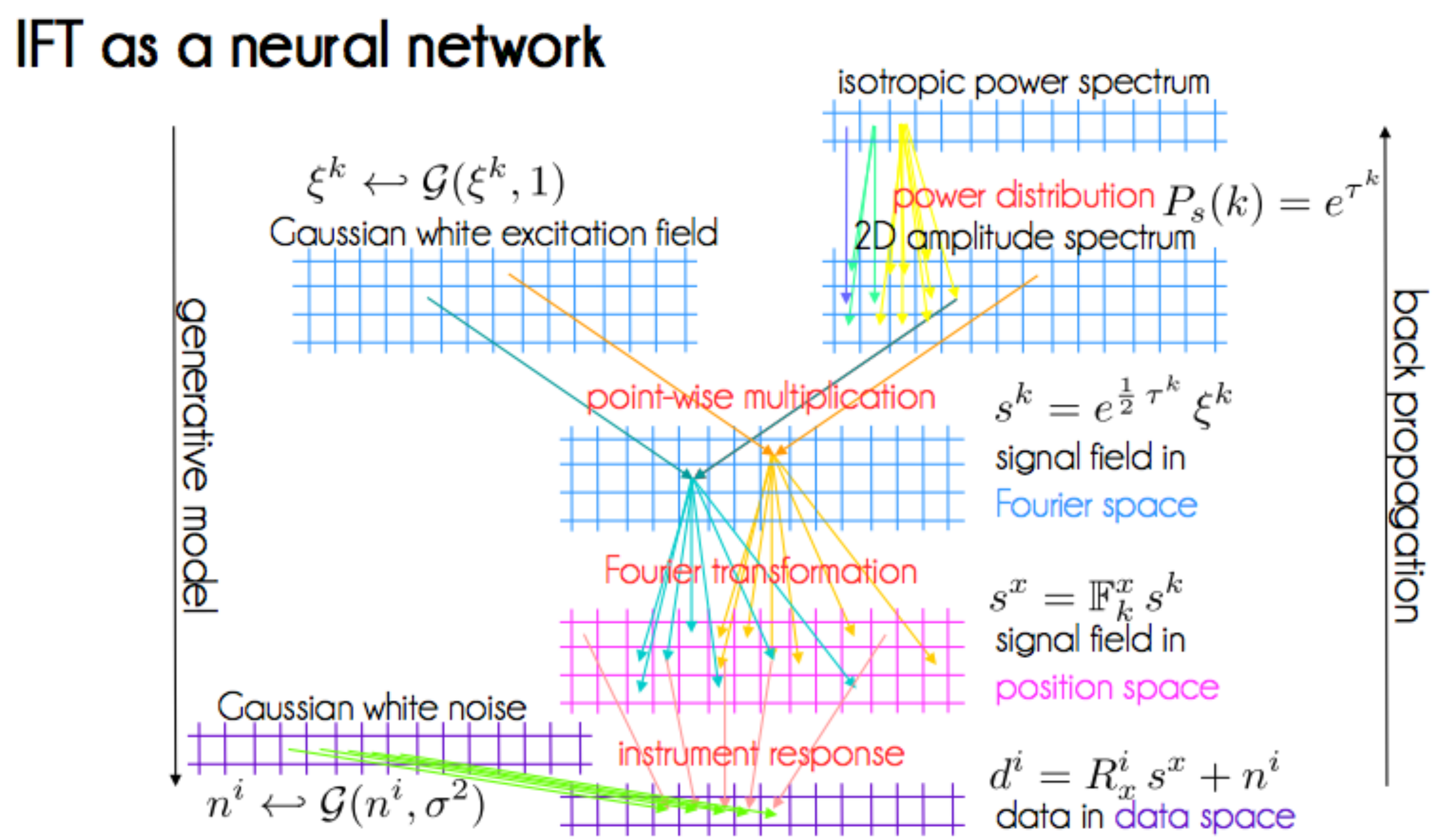
- The solution is to **treat the power spectrum as a random variable**, within the framework of a **generative model**

$$s = A\xi$$

ξ gaussian white noise

A amplitude operator

$\mathrm{cov}(s) = AA^\dagger$



Slide by Torsten Ensslin

Example 2 : Critical Wiener filter

- within the generative model, the posterior distribution is estimated through a **variational inference approach**

- We start by **approximating the unknown posterior with a parametrised distribution** (e.g., Gaussian)

$$\mathcal{P}(s|d) \sim \mathcal{Q}(s|d) = \mathcal{G}(m, D)$$

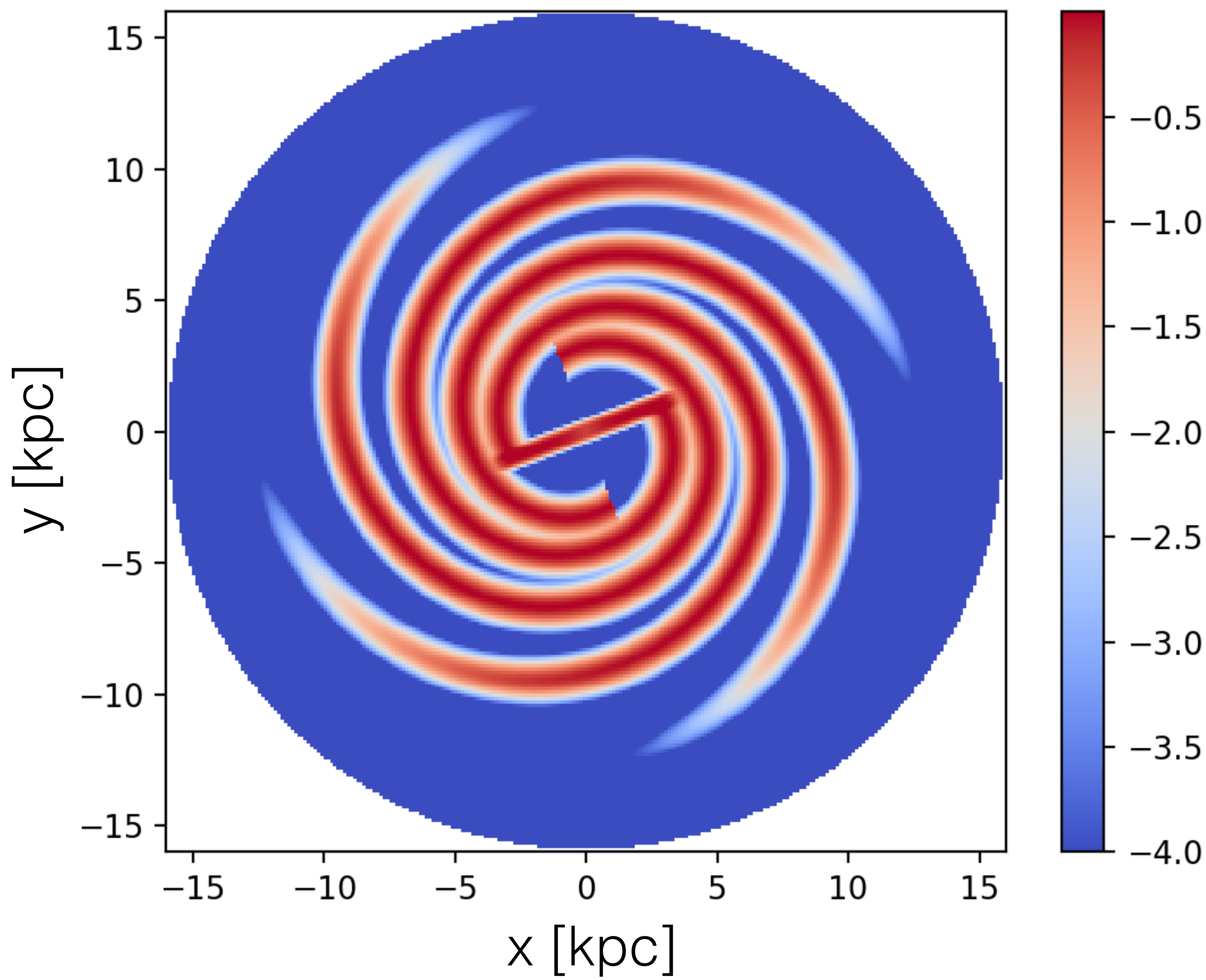
- The parameters of the distribution $\mathcal{Q}$ are determined by **minimising the Kullback-Leibler divergence** between $\mathcal{Q}$ and $\mathcal{P}$:

$$\mathrm{KL}(\mathcal{Q}, \mathcal{P}) = \int \mathcal{D}s \, \mathcal{Q}(s|d) \log \frac{\mathcal{Q}(s|d)}{\mathcal{P}(s|d)}$$

- Numerically, this minimisation is done in **several steps**, which subsequently updates the parameters of $\mathcal{Q}$

Example 2 : Critical Wiener filter

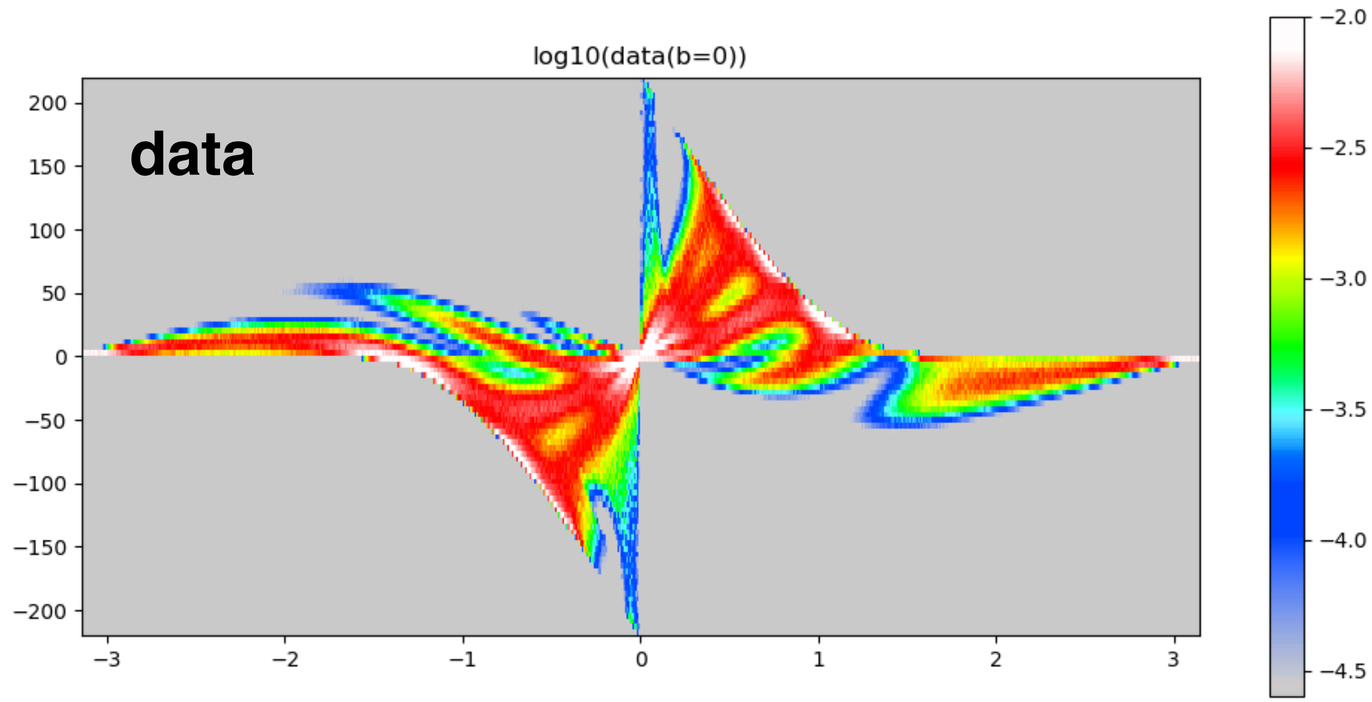
mock signal : four-armed spiral
signal model : log-normal field ($d = Re^s + n$)

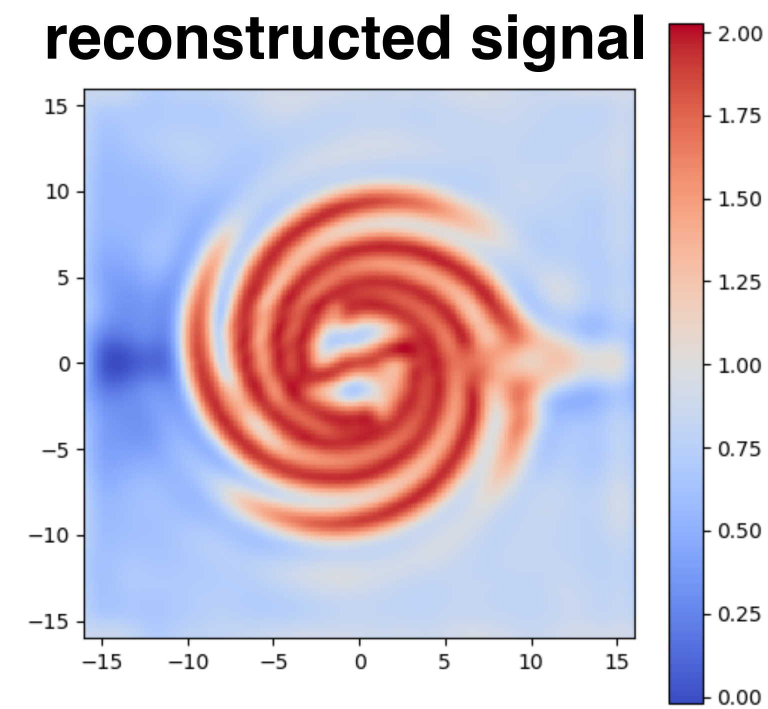


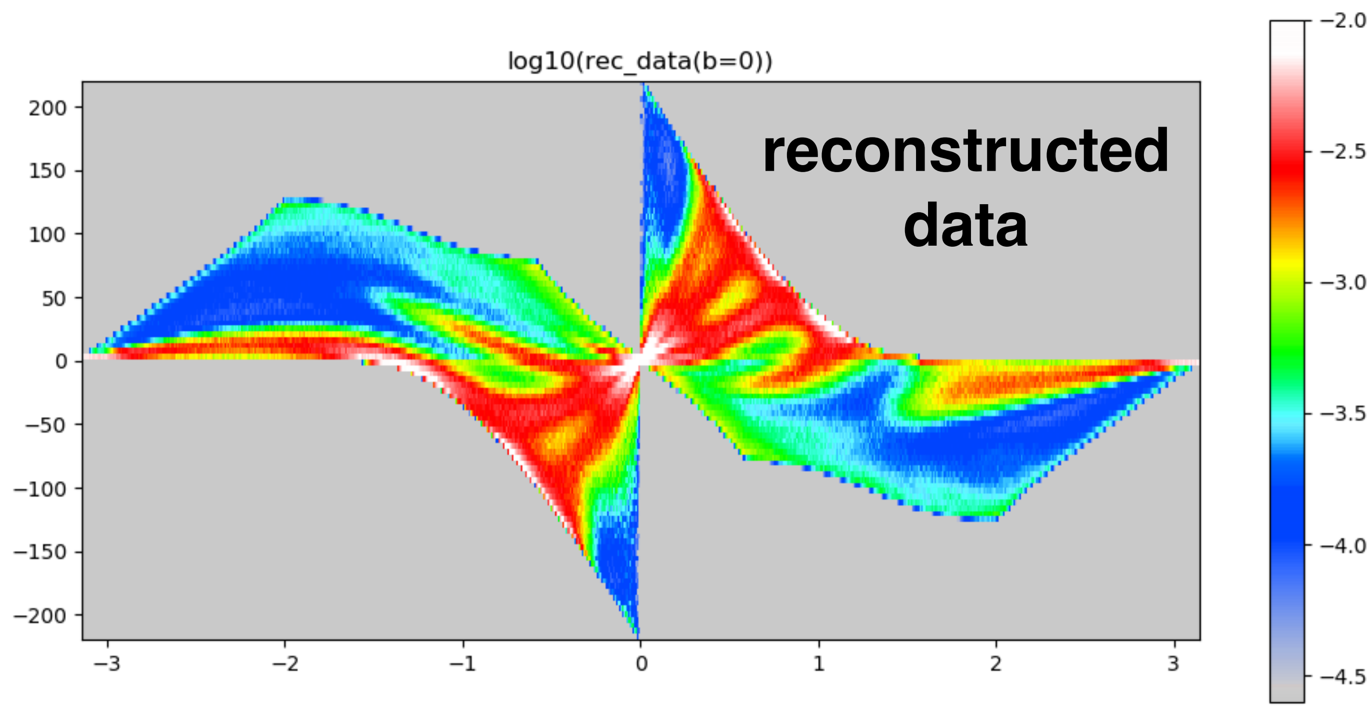
Example 2 : Critical Wiener filter

mock signal : four-armed spiral

signal model : log-normal field ($d = R e^{s} + n$)







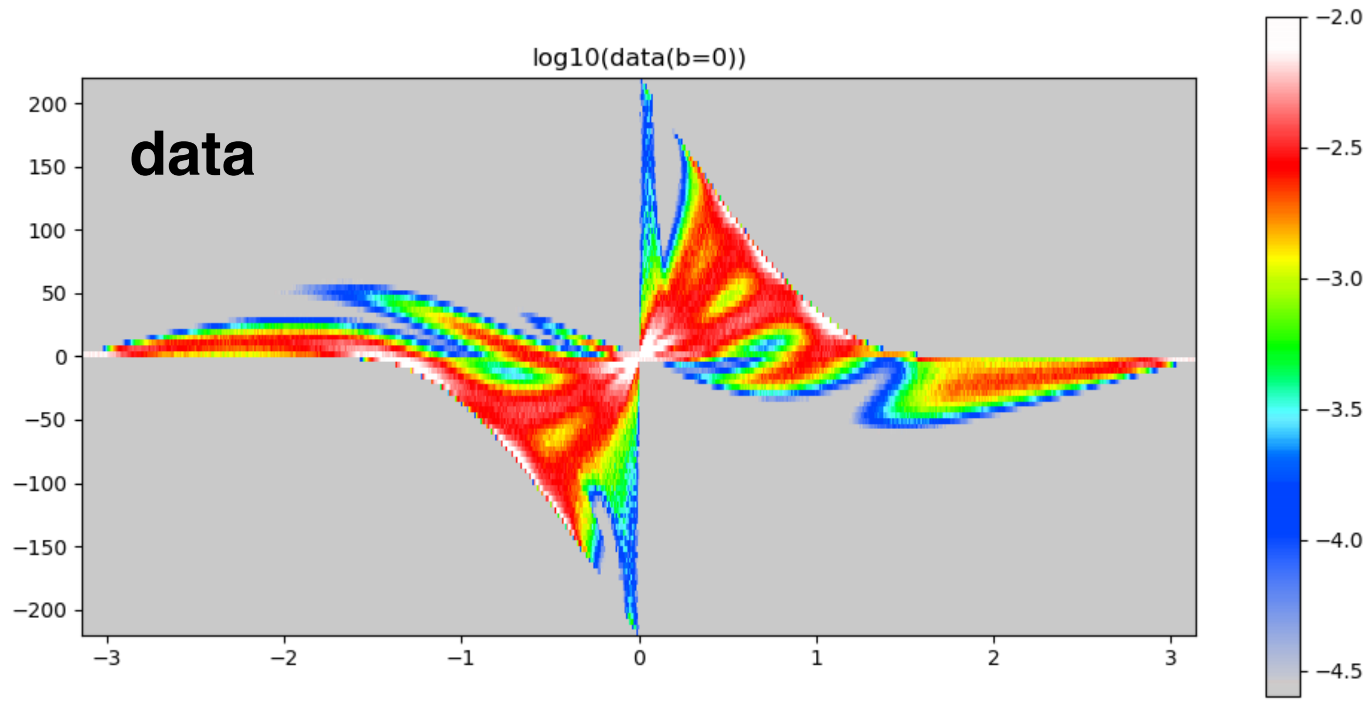
step 1

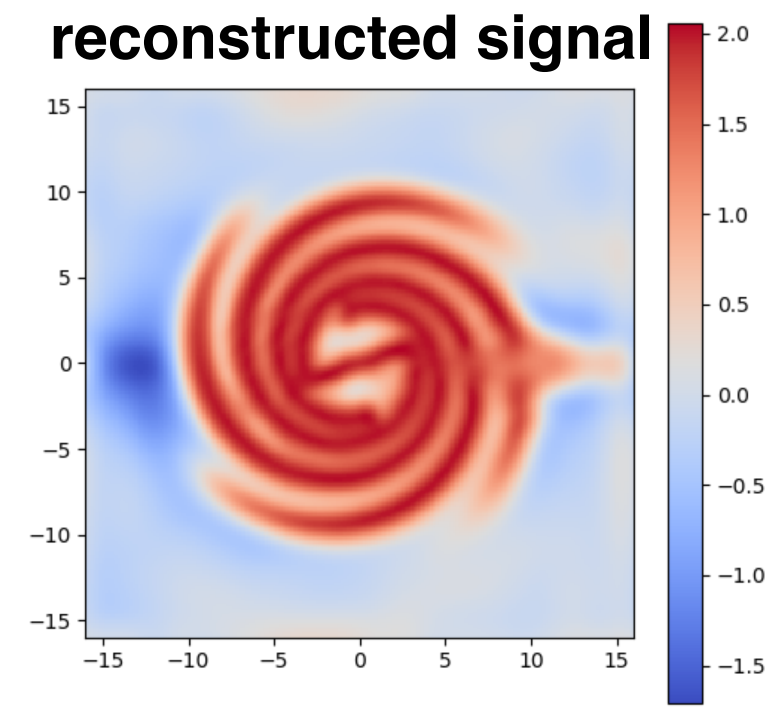
results obtained with **the NIFTy package**

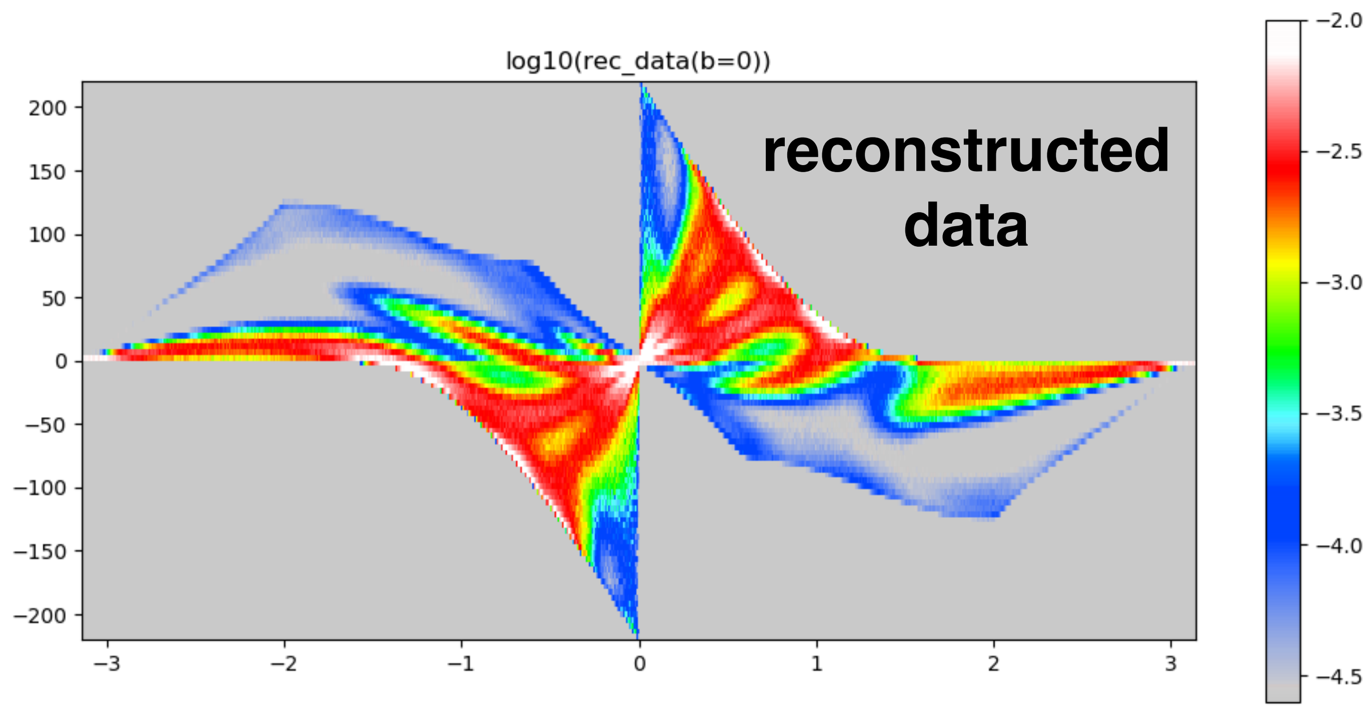
Example 2 : Critical Wiener filter

mock signal : four-armed spiral

signal model : log-normal field ($d = Re^s + n$)







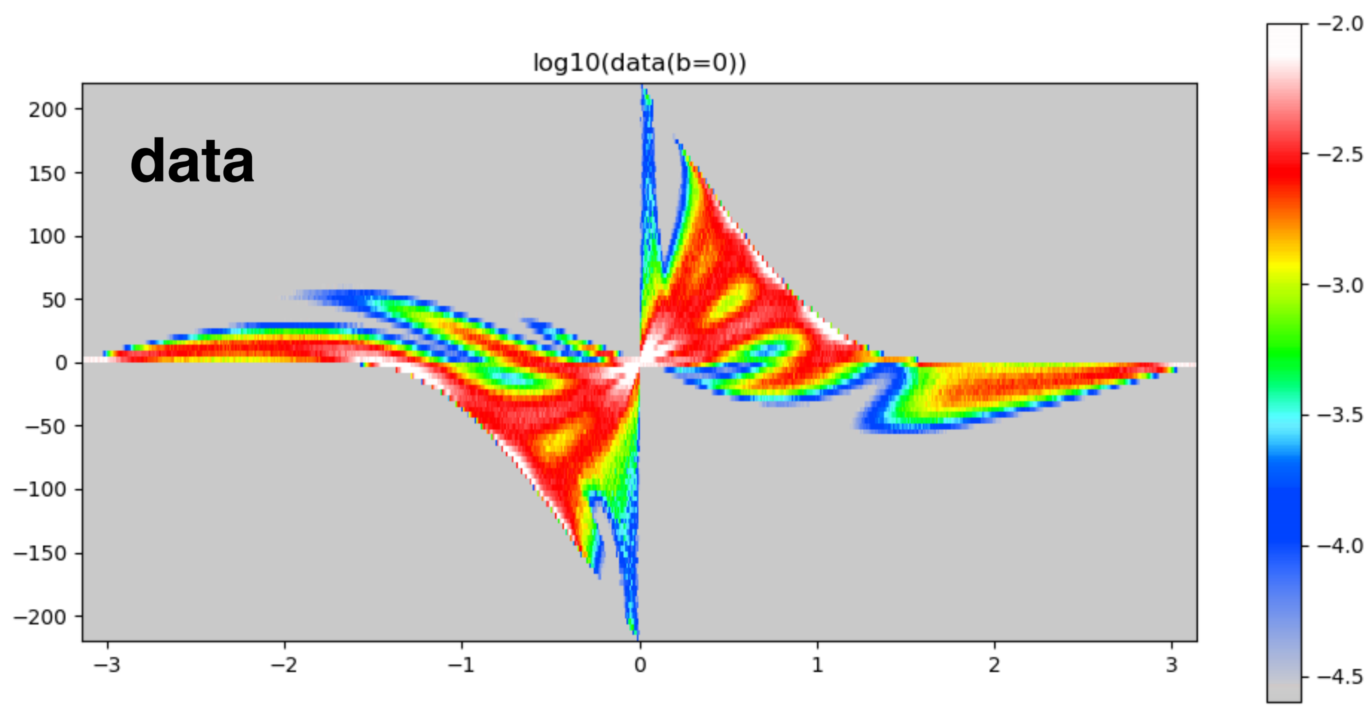
step 2

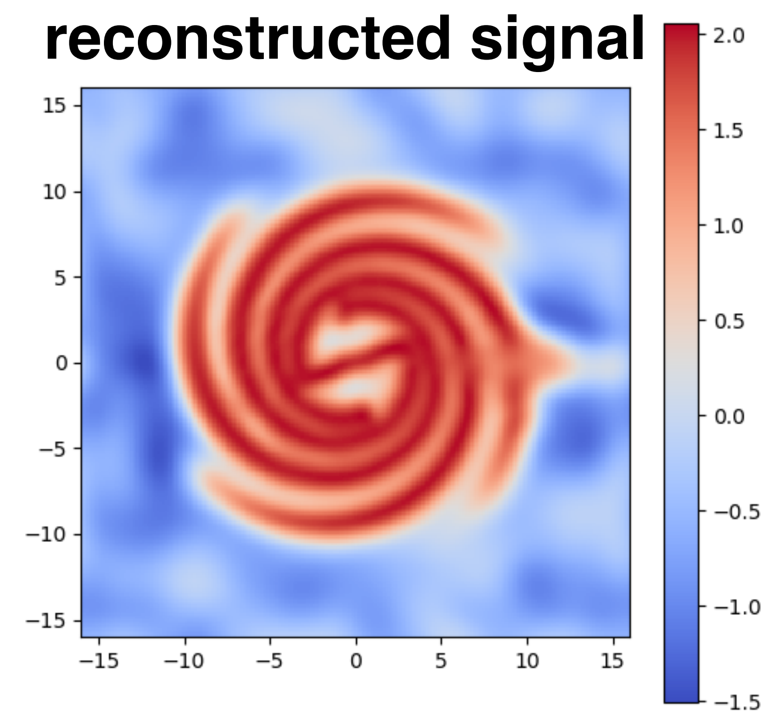
results obtained with **the NIFTy package**

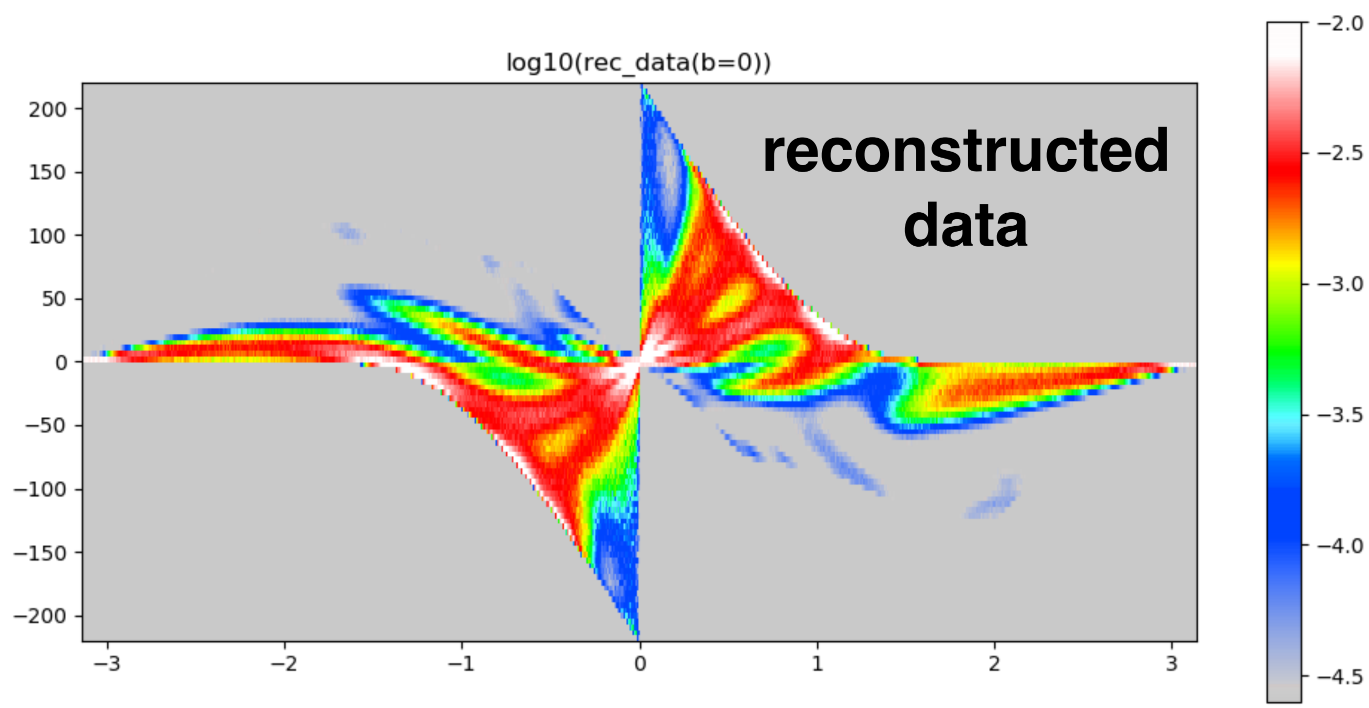
Example 2 : Critical Wiener filter

mock signal : four-armed spiral

signal model : log-normal field ($d = R e^{s} + n$)







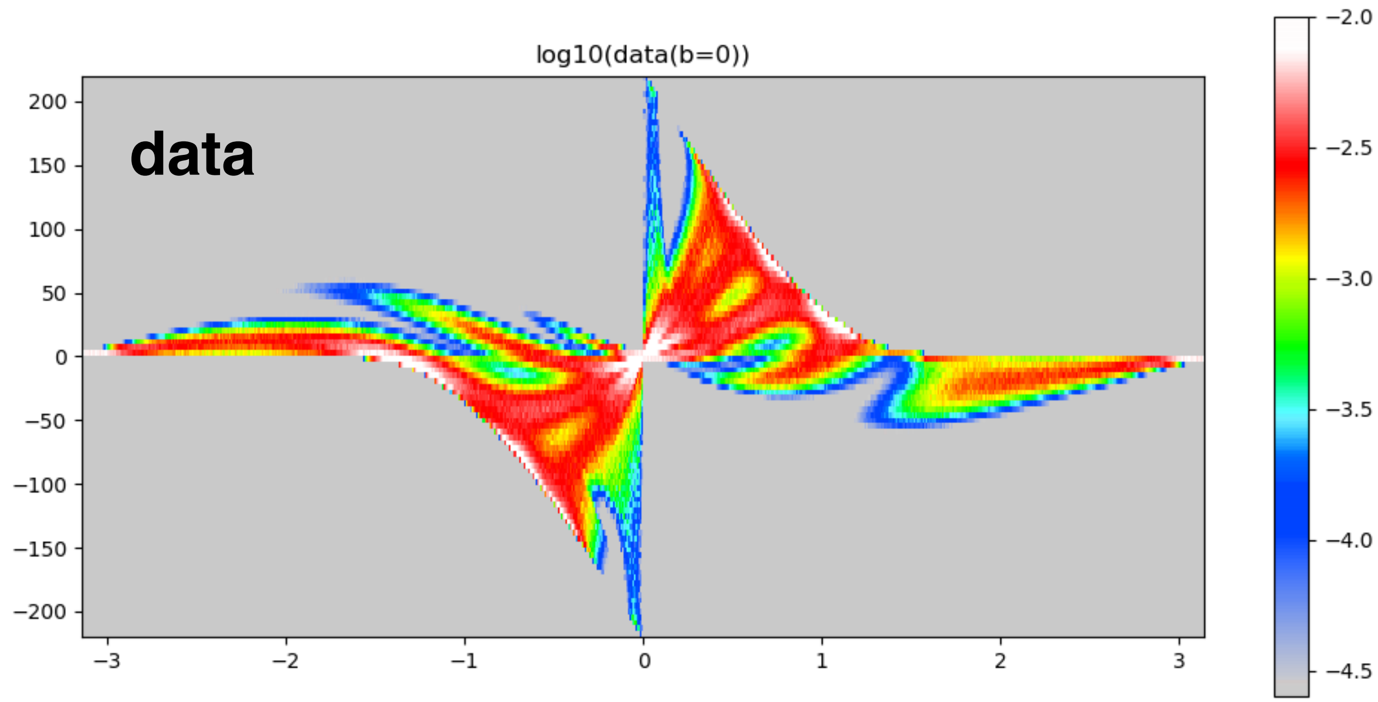
step 3

results obtained with **the NIFTy package**

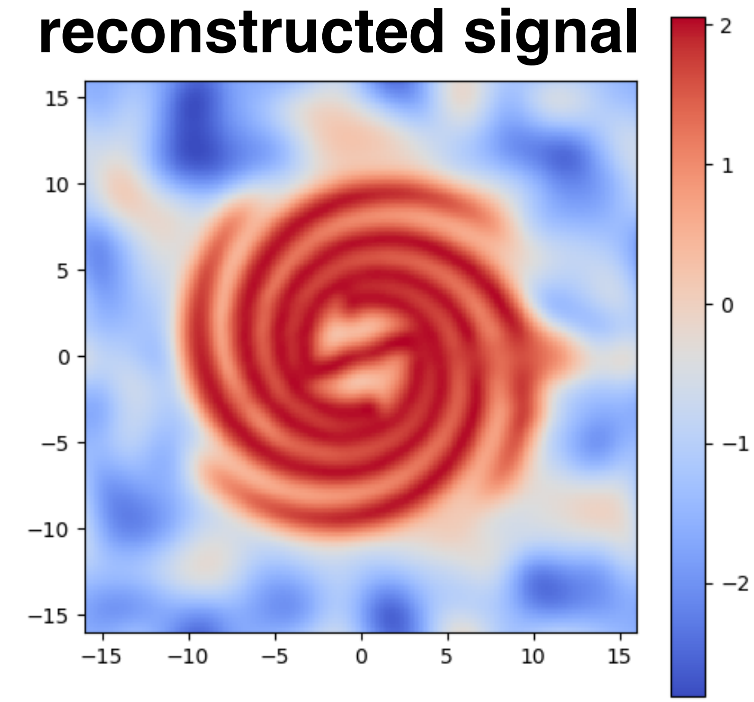
Example 2 : Critical Wiener filter

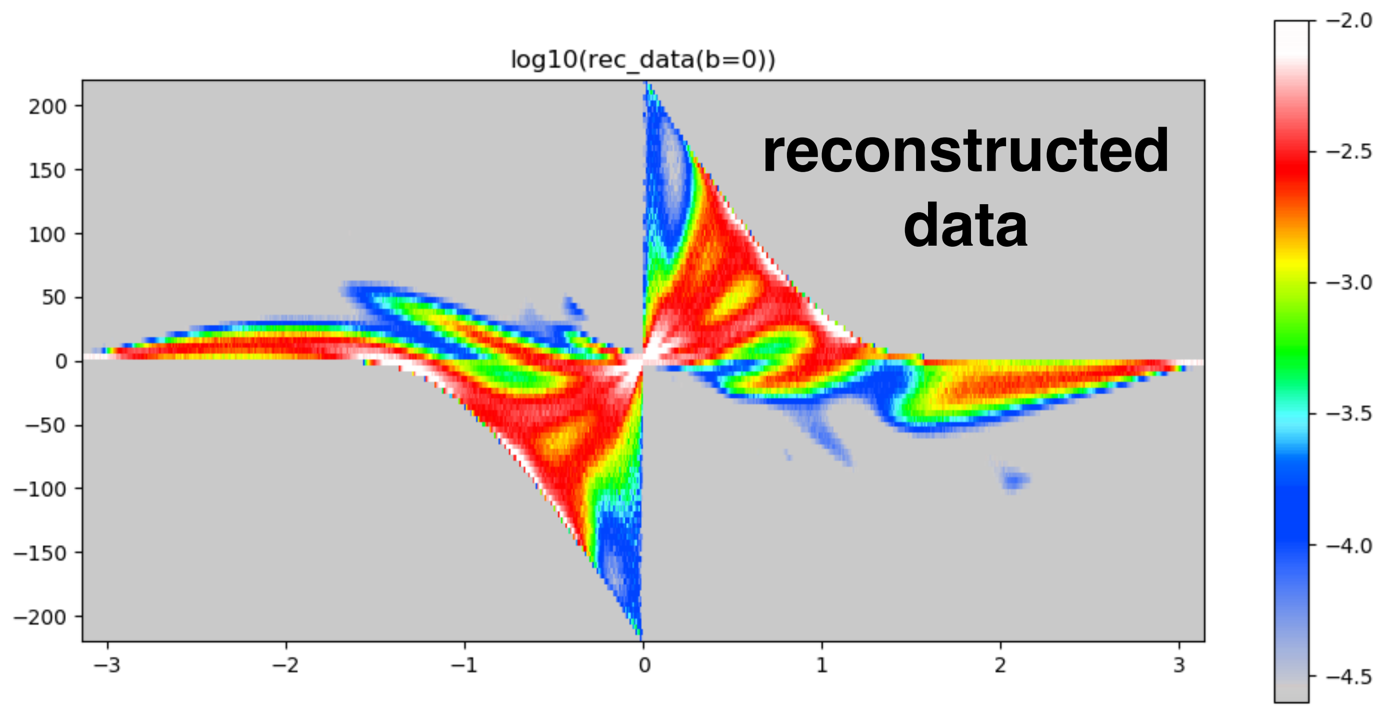
mock signal : four-armed spiral

signal model : log-normal field ($d = Re^s + n$)



reconstructed signal





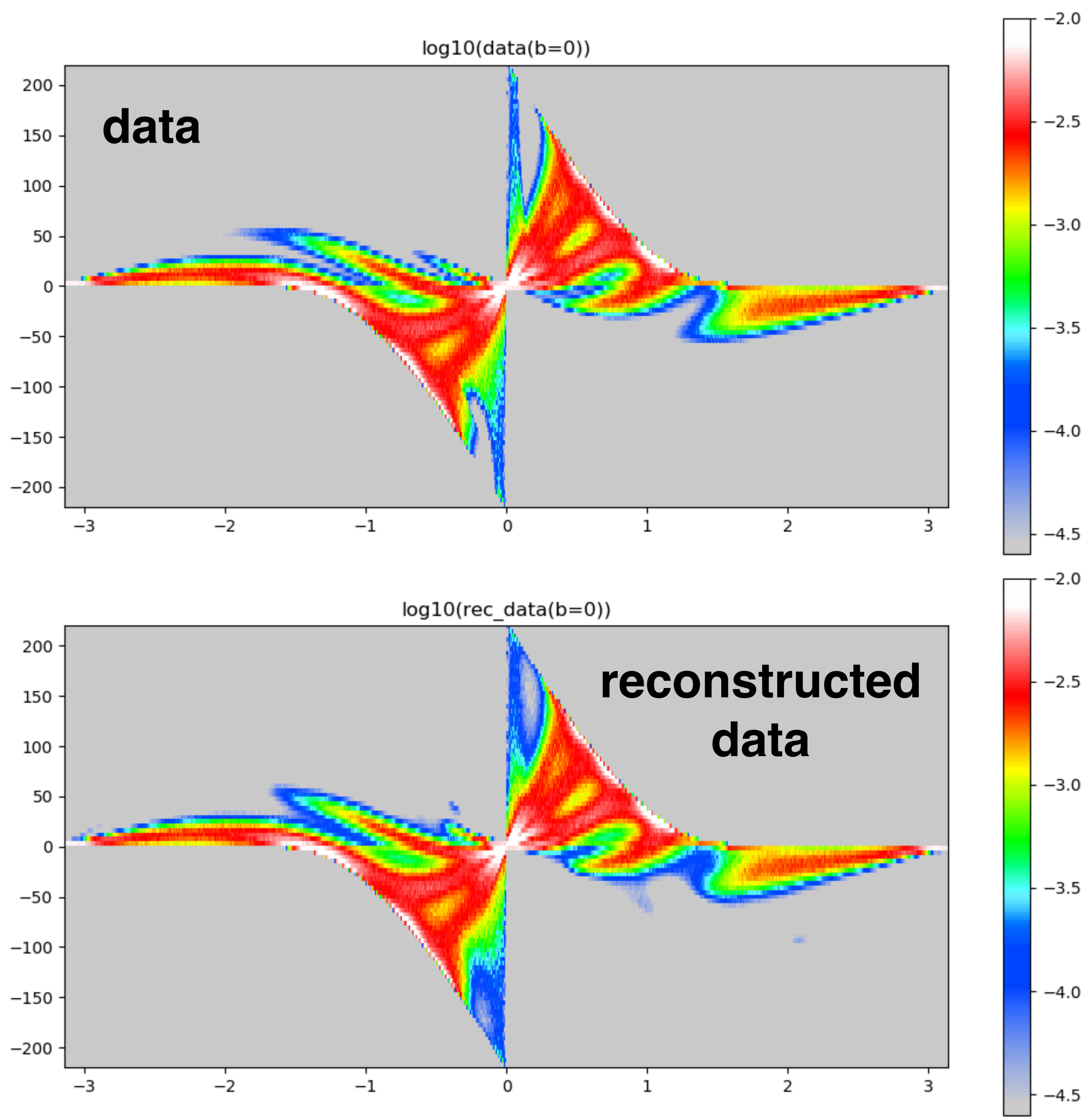
step 4

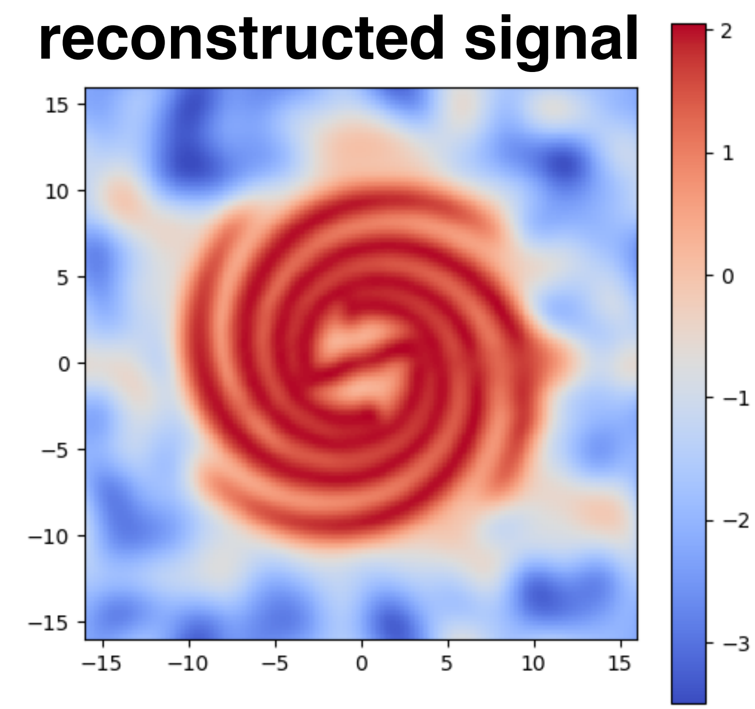
results obtained with **the NIFTy package**

Example 2 : Critical Wiener filter

mock signal : four-armed spiral

signal model : log-normal field ($d = Re^s + n$)





step 5

results obtained with **the NIFTy package**

Conclusions

- We have illustrated how **determining the density of interstellar hydrogen will help us** in shedding light on several issues related to both **cosmic-ray physics** and **dark matter searches**

- We have seen how the methods that have been devised so far to determine the gas density from observations **suffer from several drawbacks**.

- We have seen how **information field theory works promisingly** in reconstructing mock signals.

Outlook

We are now using our approach with the actual CO-line emission data. We plan to do several things; as an example:

- we will study how the **reconstructed signal depends on the rotation curve** we assume (in particular if we consider non-circular motion in the inner Galaxy)
- we will study how the **priors that we impose on the power spectrum of the signal** affect the reconstruction

Outlook

We are now using our approach with the actual CO-line emission data. We plan to do several things; as an example:

- we will study how the **reconstructed signal depends on the rotation curve** we assume (in particular if we consider non-circular motion in the inner Galaxy)
- we will study how the **priors that we impose on the power spectrum of the signal** affect the reconstruction

thank you for your attention!

Information field theory

which signal configuration is the right answer to the inference problem?

- **Maximum a Posteriori (MAP) solution**:

$$\left.\frac{\delta \mathcal{H}(d,s)}{\delta s}\right|_{s=m^{(\mathrm{MAP})}} = 0$$

(easy to calculate, but takes into account only **local information** at the maximum).

- **Posterior mean**

$$m := \langle s \rangle_{\mathcal{P}(s|d)} = \int \mathcal{D}s\, s\, \mathcal{P}(s|d)$$

(it is usually a **better choice**, as it is influenced by the whole posterior distribution, but calculating it is typically **much more difficult**)