Résoudre un problème inverse combinant mélange de bruit et modèle direct boîte noire avec des méthodes d'échantillonnage - ANR RICOCHET -

Pierre Palud

PhD directed by
Pierre Chainais, Franck Le Petit

with the collaboration of
Emeric Bron, Pierre-Antoine Thouvenin

Ecole Centrale de Lille, CRIStAL, LERMA
financed by CNRS via 80|Prime

Outline

Outline

~~Mostly empty?~~

Mostly Interstellar Medium!

Observations of GMC: Orion B in visible frequencies

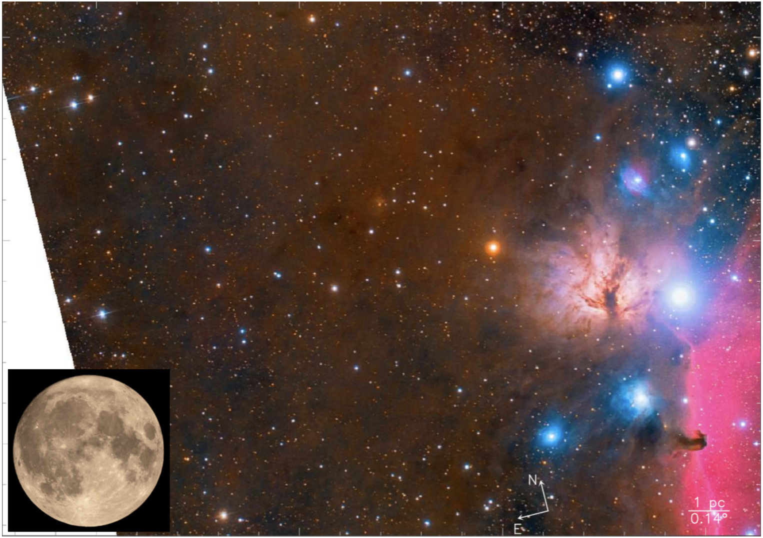


Figure: Image from Pety et al. [2016]

Observations of GMC: Orion B in visible frequencies

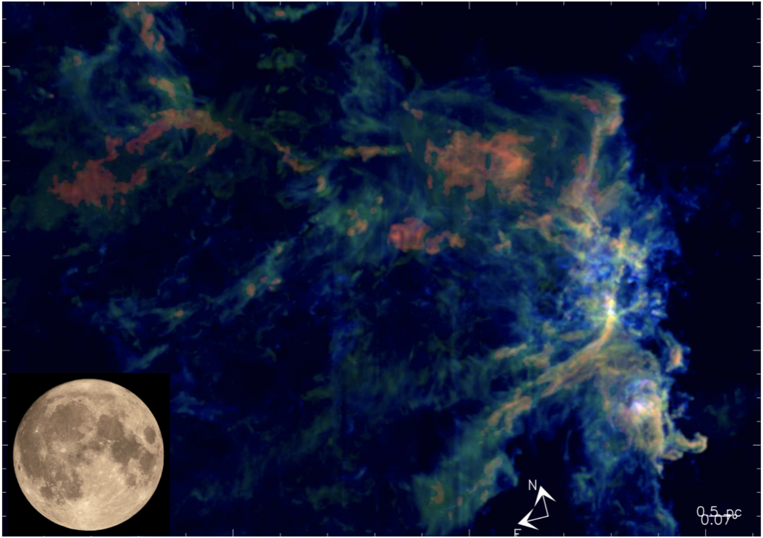


Figure: Image from Pety et al. [2016]
blue: ^{12}CO, green: ^{13}CO, red: $C^{18}O$

Photo-Dissociation Region (PDR)

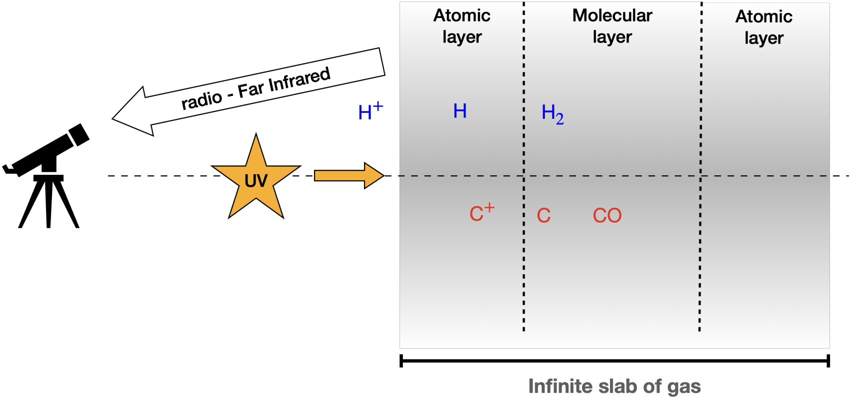


Figure: Structure of a PDR

Photo-Dissociation Region (PDR)

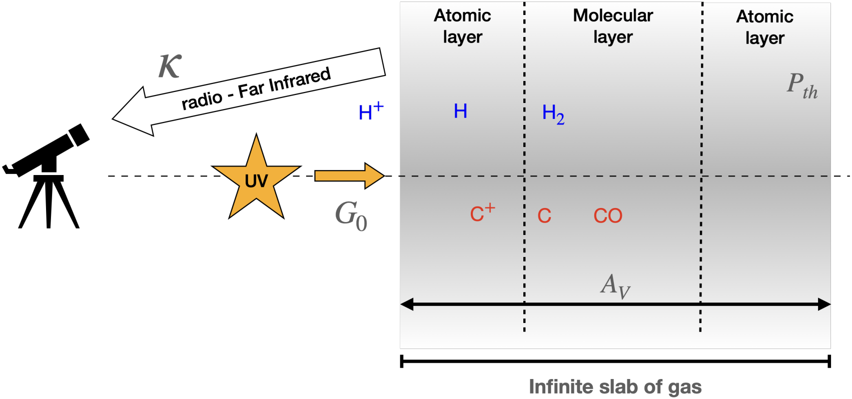
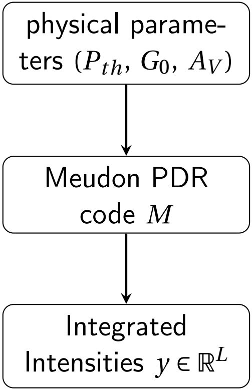


Figure: Structure of a PDR

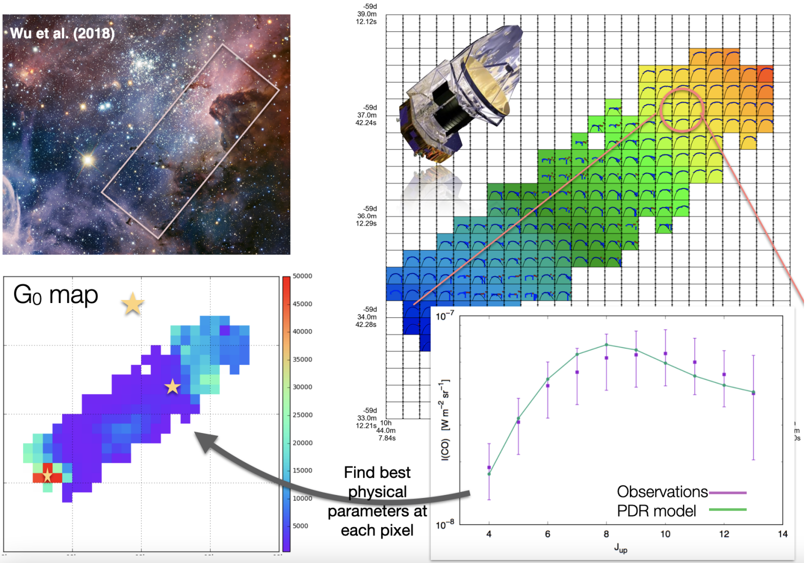
- Introduced in Le Petit et al. [2006].
- for stationary 1D slab of gas, solves:
  1. radiative transfer
  2. chemistry
  3. thermal balance

  that are all **coupled**!



Can we infer (P_{th}, G_0, A_V) from y and M?
With **no ground truth** → with **credibility intervals**

Current state of the art in astrophysics



Current state of the art in astrophysics

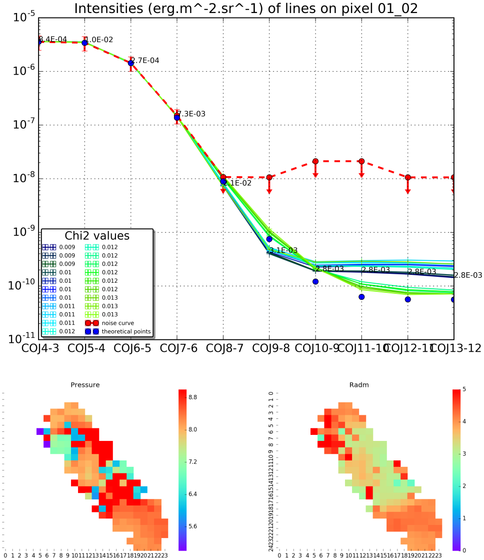


Figure: Images from Nicolas Chabalier internship report

Outline

Bayesian map inversion

inference with credibility interval

$$\Downarrow$$

a posteriori probability distribution $\mathbb{P}[(\kappa, P_{th}, G_0, A_V) \mid y]$

$$\underbrace{\mathbb{P}[(\kappa, P_{th}, G_0, A_V) \mid y]}_{\text{a posteriori}} \propto \underbrace{\mathbb{P}[y \mid (\kappa, P_{th}, G_0, A_V)]}_{\text{likelihood}} \times \underbrace{\mathbb{P}[(\kappa, P_{th}, G_0, A_V)]}_{\text{a priori}}$$

Complex distribution

$\implies$ impossible to manipulate as is

$\implies$ sampling with MCMC

Observation model

Observation: $\mathbf{y} = (y_{n,\ell})$
Parameters: $X = (x_n)$, with $x_n = (\kappa, P_{th}, G_0, A_V)$
Forward Model : $\forall \ell, \quad f_\ell : x \in \mathbb{R}^D \mapsto \kappa M_\ell(P_{th}, G_0, A_V) \in \mathbb{R}_+^*$

Observation Model:

$$\forall n, \ell,\ y_{n,\ell} = \max\left\{\omega,\ \epsilon^{(m)}_{n,\ell} f_\ell(x_n) + \epsilon^{(a)}_{n,\ell}\right\}$$

with

- M_ℓ : Meudon PDR code ⚠non explicit ⚠
- κ_n : scaling factor (includes dilution factor, etc.)
- $\epsilon^{(a)}_{n,\ell}$: additive noise (thermal, instruments)
- $\epsilon^{(m)}_{n,\ell}$: multiplicative noise (calibration error)
- ω : minimum detectable value by telescope

How to deal with **non explicit** forward map f ?
How to deal with **both** additive and multiplicative noises?

Approximation of the Meudon PDR code

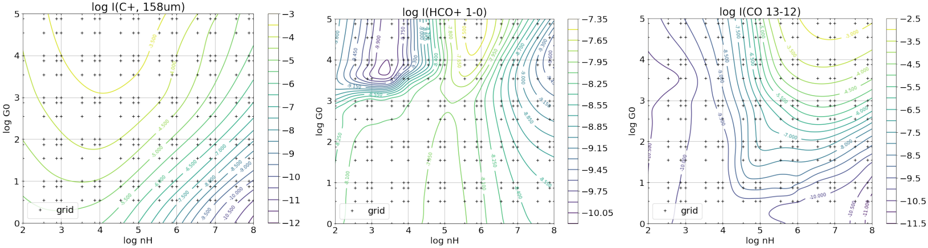
Forward map f :

1. long to evaluate (~2h)
2. non-explicit $\implies$ non accessible gradients

Approximation of the Meudon PDR code

Forward map f :

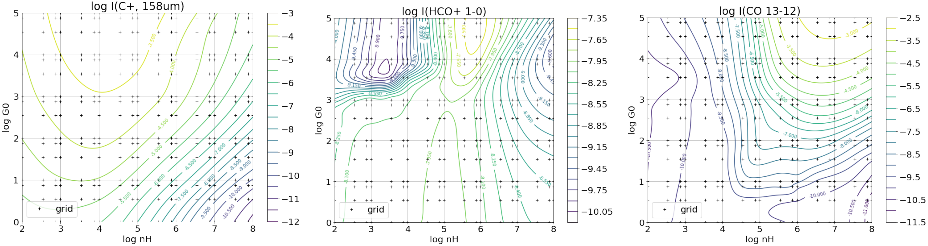
1. long to evaluate (~2h)
2. non-explicit $\implies$ non accessible gradients



Approximation of the Meudon PDR code

Forward map f :

1. long to evaluate (~2h)
2. non-explicit $\implies$ non accessible gradients



model reduction : polynomial regression in the log space on some set $\mathscr{C}$

$$\forall \ell, \quad f_\ell(x) = \exp\left[\log \kappa + P_\ell(P_{th},\, G_0,\, A_V)\right]$$

Observation Model

$$u_{n,\ell} = \epsilon^{(m)}_{n,\ell} f_\ell(x_n) + \epsilon^{(a)}_{n,\ell}$$

Additive approximation

$$u_{n,\ell} \simeq f_\ell(x_n) + e^{(a)}_{n,\ell}$$

$$e^{(a)}_{n,\ell} \sim \mathcal{N}(m_{a,n,\ell},\, s^2_{a,n,\ell})$$

Multiplicative approximation

$$u_{n,\ell} \simeq e^{(m)}_{n,\ell} f_\ell(x_n)$$

$$e^{(m)}_{n,\ell} \sim \log\mathcal{N}(m_{m,n,\ell},\, s^2_{m,n,\ell})$$

Final Likelihood

$$\forall n, \ell,\ y_{n,\ell} = \begin{cases} \omega & \text{if } \epsilon^{(m)}_{n,\ell} f_\ell(x_n) + \epsilon^{(a)}_{n,\ell} \le \omega \quad \text{(censored)} \\ \epsilon^{(m)}_{n,\ell} f_\ell(x_n) + \epsilon^{(a)}_{n,\ell} & \text{else} \quad \text{(uncensored)} \end{cases}$$

$$\Downarrow$$

$$\pi_{\text{mix}}(y_{n,\ell}|x_n) = \left\{ \pi^{(c)}_a (y_{n,\ell}|x_n)^{\lambda_{n,\ell}}\ \pi^{(c)}_m (y_{n,\ell}|x_n)^{1-\lambda_{n,\ell}} \right\}^{[y_{n,\ell}=\omega]} \times \left\{ \pi^{(u)}_a (y_{n,\ell}|x_n)^{\lambda_{n,\ell}}\ \pi^{(u)}_m (y_{n,\ell}|x_n)^{1-\lambda_{n,\ell}} \right\}^{[y_{n,\ell}>\omega]}$$

- $y_{n,\ell} = \omega$: if true, censored data
- $\pi^{(c)}_a$, $\pi^{(u)}_a$: additive approximation
- $\pi^{(c)}_m$, $\pi^{(u)}_m$: multiplicative approximation
- $\lambda_{n,\ell}$: **controls the mixing of the two approximation models**

Choice of mixing parameter λ

How to define λ ? We have that

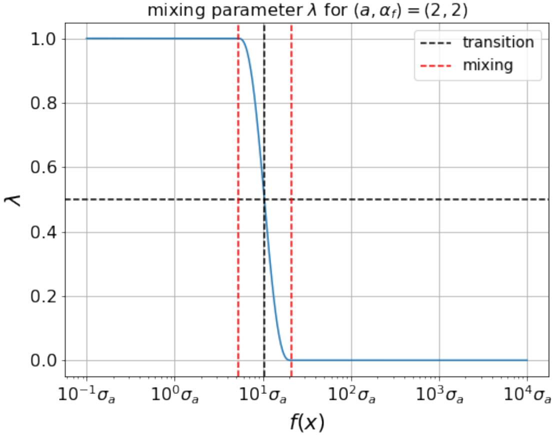
- When $\sigma_m \to 0$, observation model $\rightsquigarrow$ additive approx
- When $f_\ell(x) \gg \sigma_a$, observation model $\rightsquigarrow$ multiplicative approx

Choice of mixing parameter λ

How to define λ ? We have that

- When $\sigma_m \to 0$, observation model $\rightsquigarrow$ additive approx
- When $f_\ell(x) \gg \sigma_a$, observation model $\rightsquigarrow$ multiplicative approx

$\Longrightarrow$ λ parameteric function :



How relevant is our approximation?

$$\forall \ell,\ D^{(\ell)}_{\text{K-S}}\left(\pi, \tilde{\pi}_{(a,\alpha_f)}\right) : x \mapsto \sup_{y \in \mathbb{R}} \left| F^{(\ell)}(y|x) - \tilde{F}^{(\ell)}_{(a,\alpha_f)}(y|x) \right|$$

Assuming $X \sim \text{Unif}(\mathscr{C})$:

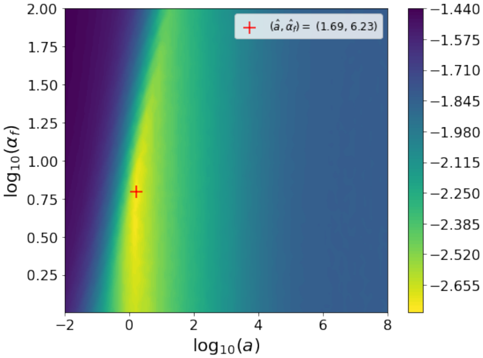
$$(a, \hat{\alpha}_f) \in \underset{a>0, \alpha_f>1}{\arg\min} \frac{1}{L} \sum_{\ell=1}^{L} \mathbb{E}\left[D^{(\ell)}_{\text{K-S}}\left(\pi, \tilde{\pi}_{(a,\alpha_f)}\right)(X) \right]$$

How relevant is our approximation?

$$\forall \ell,\ D^{(\ell)}_{\text{K-S}}\left(\pi, \tilde{\pi}_{(a,\alpha_f)}\right) : x \mapsto \sup_{y \in \mathbb{R}} \left| F^{(\ell)}(y|x) - \tilde{F}^{(\ell)}_{(a,\alpha_f)}(y|x) \right|$$

Assuming $X \sim \text{Unif}(\mathscr{C})$:

$$(a, \hat{\alpha}_f) \in \underset{a>0, \alpha_f>1}{\arg\min} \frac{1}{L} \sum_{\ell=1}^{L} \mathbb{E}\left[D^{(\ell)}_{\text{K-S}}\left(\pi, \tilde{\pi}_{(a,\alpha_f)}\right)(X) \right]$$



Including a priori information

a priori information on $(\kappa, P_{th}, G_0, A_V)$:

- spatial regularization with unknown parameter (Pereyra et al. [2015])

- f_ℓ: estimated from a grid on $\mathscr{C}$
  → **constraint of belonging to a cube**
  ⚠non smooth prior ⚠but can be tempered with a smooth penalty

Sampler

smooth prior + **smooth** likelihood $\implies$ **smooth** posterior
classic MCMC algorithm (e.g., MALA) OK

Sampler

smooth prior + **smooth** likelihood $\implies$ **smooth** posterior
classic MCMC algorithm (e.g., MALA) OK

Other difficulties :

- f and ∇f cover **various decades**
  $\implies$ classic methods **inefficient**
  $\implies$ Preconditioned MALA kernel with RMSProp preconditioner

Sampler

smooth prior + **smooth** likelihood $\implies$ **smooth** posterior
classic MCMC algorithm (e.g., MALA) OK

Other difficulties :

- f and ∇f cover **various decades**
  $\implies$ classic methods **inefficient**
  $\implies$ Preconditioned MALA kernel with RMSProp preconditioner
- **non-convex** negative log-posterior (because of Meudon PDR code)
  $\implies$ need to avoid being trapped in **local minima**
  $\implies$ Multiple-Try Metropolis (MTM) kernel

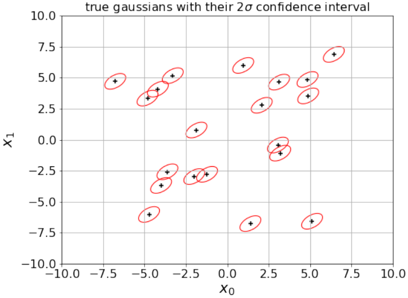
Final sampler : combination of these two kernels

Outline

Mixture of gaussians in a square

Illustration that our algorithm explores interesting **local minima** :

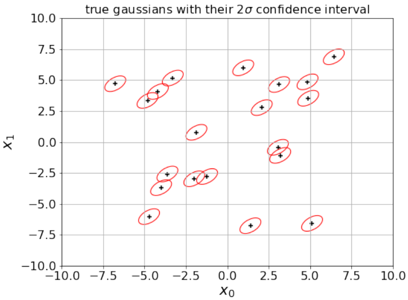
- mixture of 20 gaussians
- with constraint $X \in \mathscr{C} = [-10, 10] \times [-10, 10]$

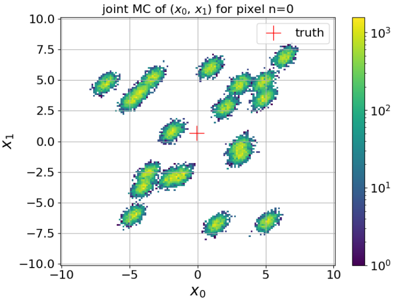


Mixture of gaussians in a square

Illustration that our algorithm explores interesting **local minima** :

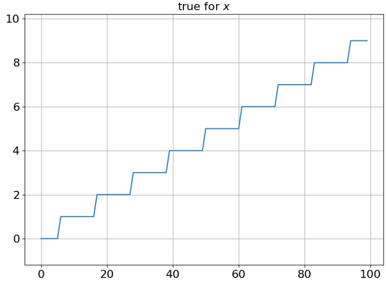
- mixture of 20 gaussians
- with constraint $X \in \mathscr{C} = [-10, 10] \times [-10, 10]$

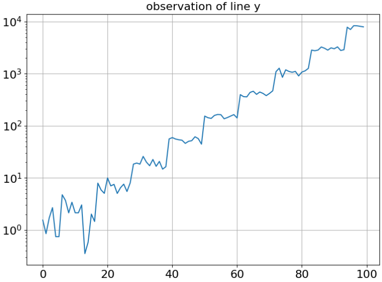




Toy case : Time Series Inversion

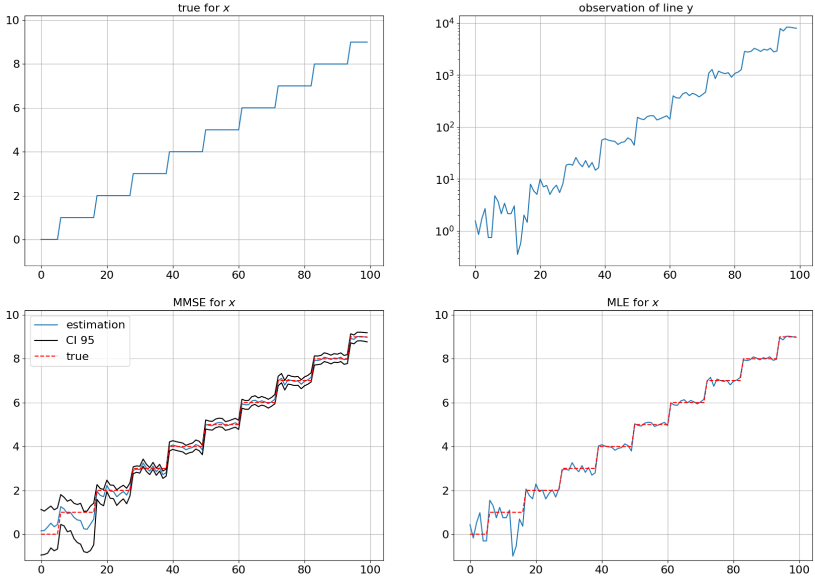
$$y_{n,\ell} = \epsilon^{(m)}_{n,\ell} f(x_n) + \epsilon^{(a)}_{n,\ell} \text{ with } f : x \in \mathbb{R} \mapsto e^x,\ \sigma_a = 1,\ \sigma_m \sim 10\%$$





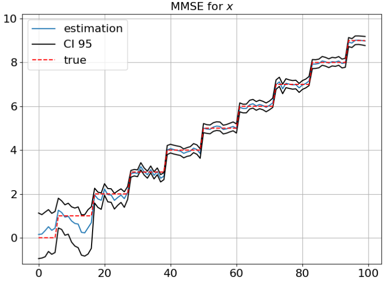
Toy case : Time Series Inversion

$$y_{n,\ell} = \epsilon^{(m)}_{n,\ell} f(x_n) + \epsilon^{(a)}_{n,\ell} \text{ with } f : x \in \mathbb{R} \mapsto e^x,\ \sigma_a = 1,\ \sigma_m \sim 10\%$$



Toy case 1: Time Series Inversion

$$y_{n,\ell} = \epsilon^{(m)}_{n,\ell} f(x_n) + \epsilon^{(a)}_{n,\ell} \text{ with } f : x \in \mathbb{R} \mapsto e^x,\ \sigma_a = 1,\ \sigma_m \sim 10\%$$



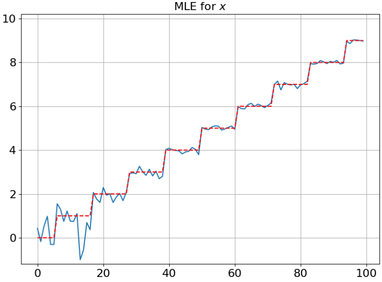
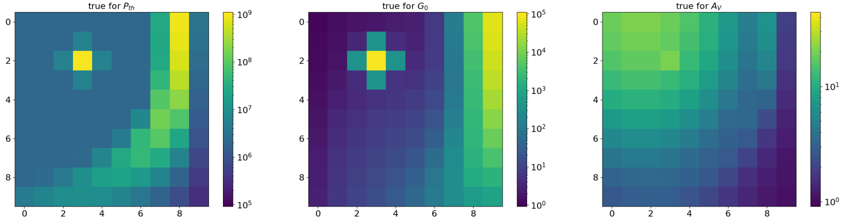


Table: Estimation Summary

estimator	MSE	SNR
MMSE	3.6	28.8
MLE	10.5	24.1

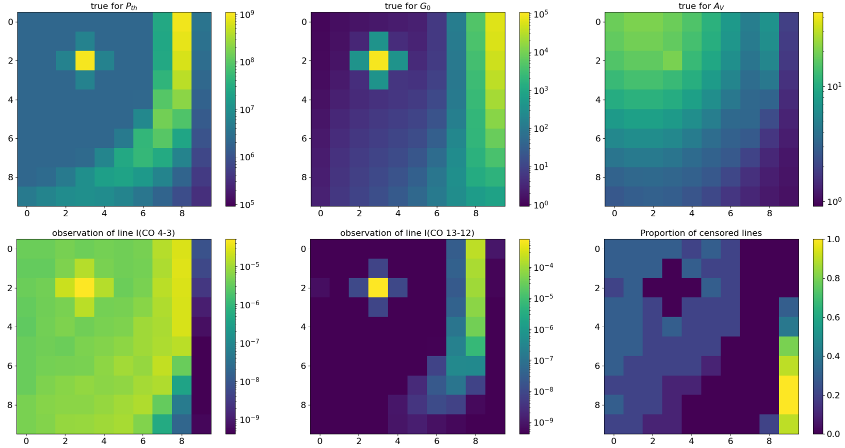
Astrophysical Toy case: Map Inversion

$$y_{n,\ell} = \max\left\{\omega,\ \epsilon_{n,\ell}^{(m)} f_\ell(x_n) + \epsilon_{n,\ell}^{(a)}\right\}$$

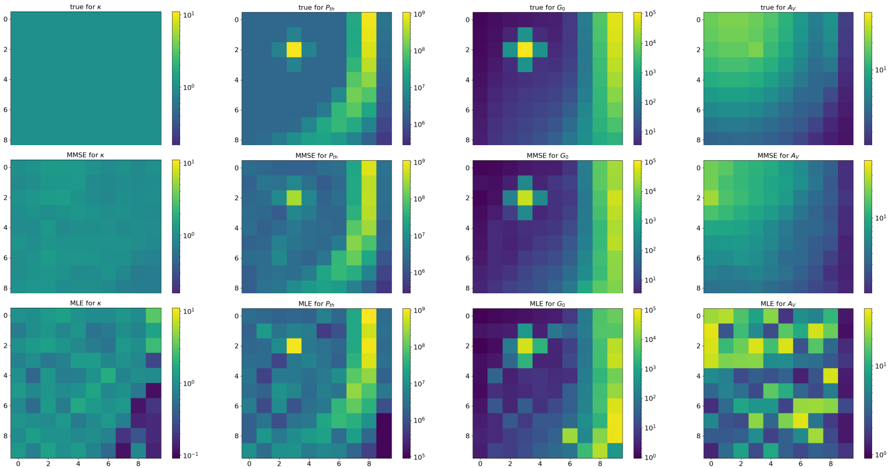


Astrophysical Toy case: Map Inversion

$$y_{n,\ell} = \max\left\{\omega,\ \epsilon^{(m)}_{n,\ell} f_\ell(x_n) + \epsilon^{(a)}_{n,\ell}\right\}$$

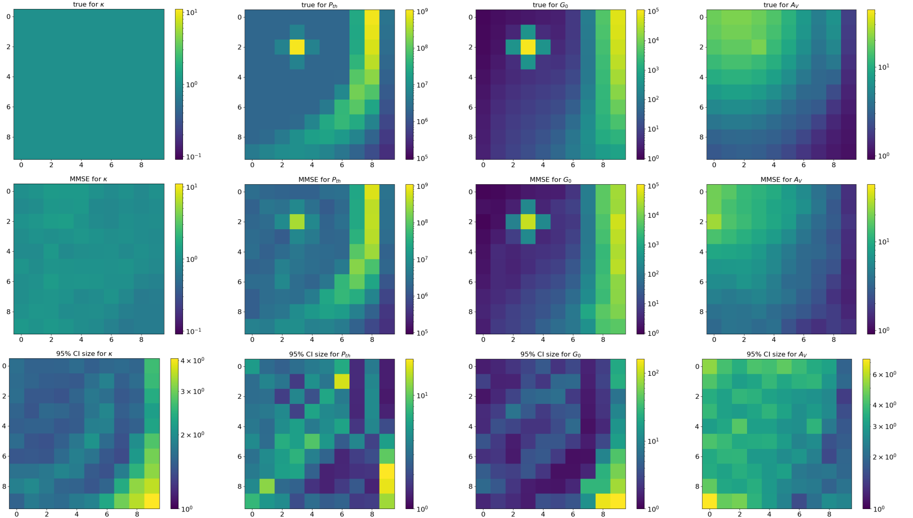


Astrophysical Toy case: Map Inversion



estimator	MSE	SNR
MMSE	13.9	14.6
MLE	146.2	4.4

Astrophysical Toy case: Map Inversion



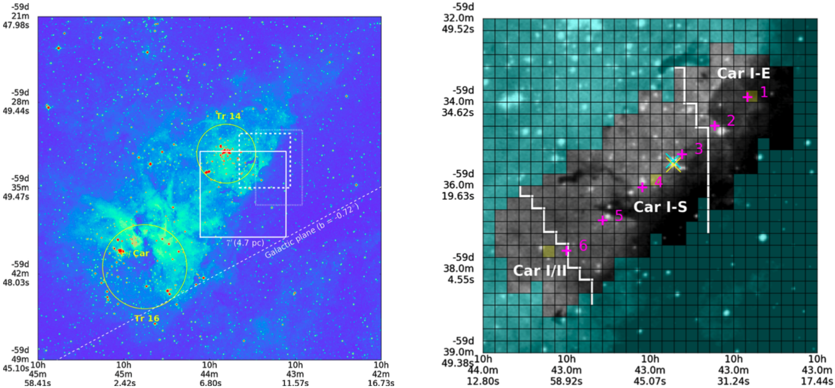


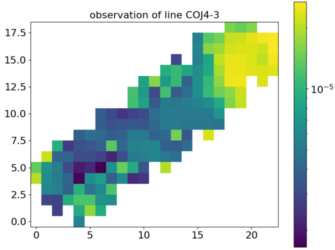
Figure: images from Wu et al. [2018]

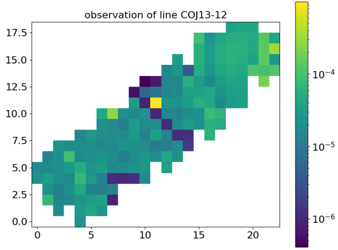
Application to Carina cloud (Car-I)

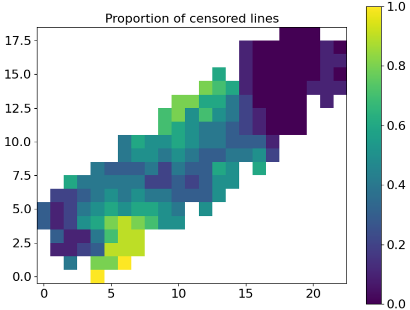
1. one of the most studied PDR
2. $\sim$2kpc ($\sim 5\times$ further to us than Orion B)
3. close to Trumpler-14 star cluster of O and B stars
4. previous estimations of G_0 :
   - $\simeq 10^4$ (Brook et al., 2003) (estimation from stellar composition)
   - 1390 (Oberst et al., 2011) (estimation from PDR model)
   - 3200 (Kramer et al., 2011) (estimation from PDR model)

Note : There is apparently an inconsistency between estimations of G_0

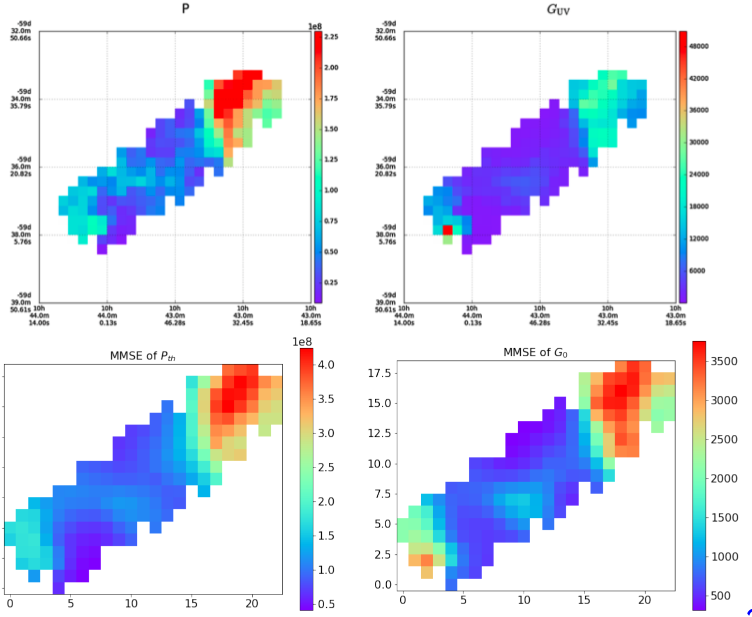
Application to Carina cloud (Car-I)





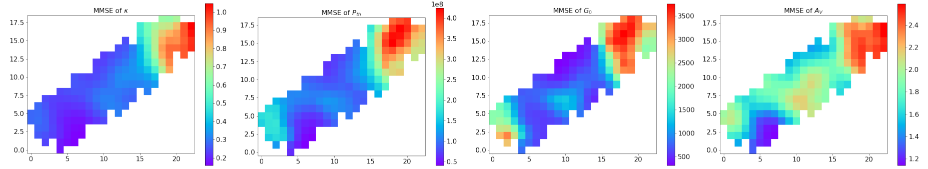


Application to Carina cloud (Car-I)



Application to Carina cloud (Car-I)

Figure: MMSE estimation



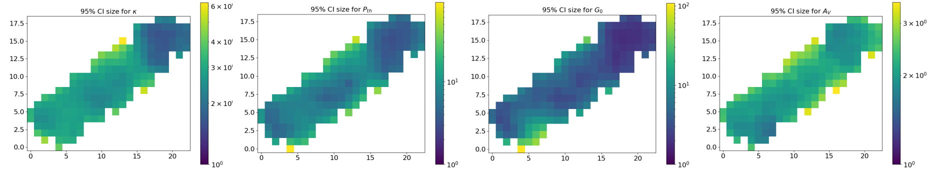


Figure: size of 95% confidence interval $u_{97.5\%}/u_{2.5\%}$

Outline

Summary

- Definition of a likelihood that handles mixture of noises
- Definition of a method of physical parameters inference with
  1. Model reduction to handle « black box » forward model
  2. P-MALA kernel to tackle regularity issues
  3. MTM kernel to tackle the non-log-concavity of the posterior
- Evaluation of the method on toy data
- Application to real world data

Thank you for your attention!

F. Le Petit, C. Nehme, J. Le Bourlot, and E. Roueff. A Model for Atomic and Molecular Interstellar Gas: The Meudon PDR Code. The Astrophysical Journal Supplement Series, 164(2):506–529, June 2006. ISSN 0067-0049, 1538-4365. doi: 10.1086/503252. URL https://iopscience.iop.org/article/10.1086/503252.

M. Pereyra, J. M. Bioucas-Dias, and M. A. T. Figueiredo. Maximum-a-posteriori estimation with unknown regularisation parameters. In 2015 23rd European Signal Processing Conference (EUSIPCO), pages 230–234, Nice, Aug. 2015. IEEE. ISBN 978-0-9928626-3-3. doi: 10.1109/EUSIPCO.2015.7362379. URL http://ieeexplore.ieee.org/document/7362379/.

J. Pety, V. Guzman, J. Orkisz, H. Liszt, M. Gerin, E. Bron, S. Bardeau, J. Goicoechea, P. Gratier, F. Petit, F. Levrier, K. Oberg, E. Roueff, and A. Sievers. The anatomy of the Orion B Giant Molecular Cloud: A local template for studies of nearby galaxies. Astronomy & Astrophysics, 599, Nov. 2016. doi: 10.1051/0004-6361/201629862.

References II

R. Wu, E. Bron, T. Onaka, F. Le Petit, F. Galliano, D. Languignon, T. Nakamura, and Y. Okada. Constraining physical conditions for the PDR of Trumpler 14 in the Carina Nebula. Astronomy & Astrophysics, 618:A53, Oct. 2018. ISSN 0004-6361, 1432-0746. doi: 10.1051/0004-6361/201832595. URL https://www.aanda.org/10.1051/0004-6361/201832595.