



De la possibilité d'ignorer un bruit de calibration lors de la construction d'un estimateur du maximum de vraisemblance en radio-astronomie

GRETISI - 27 Août 2025

Léontine Ségal^{1,2} Antoine Roueff² Claude Jauffret² Jérôme Pety^{1,3} Maryvonne Gerin³
Le consortium ORION-B⁴

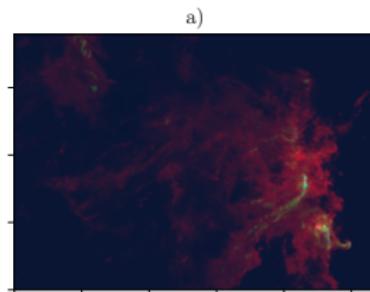
¹ Institut de RadioAstronomie Millimétrique (IRAM) ² Univ. de Toulon, Aix Marseille Univ., CNRS, IM2NP ³ LUX

⁴ <https://www.iram.fr/~pety/ORION-B/team.html>

Introduction

Contexte et objectifs

a) Image composite du nuage géant moléculaire **Orion B** ($J = 1 \rightarrow 0$ de ^{12}CO , C^{18}O , et H^{13}CO^+).



Le *ORION-B Large Program* (P.I. : J. Pety & M. Gerin)



FORMATION DES ÉTOILES dans le milieu interstellaire

- ESTIMER LES PROPRIÉTÉS DU NUAGE
- n [cm $^{-3}$] : densité volumique de H $_2$
 - T [K] : température cinétique
 - ...



Radiotélescope **IRAM-30m** (Sierra Nevada)

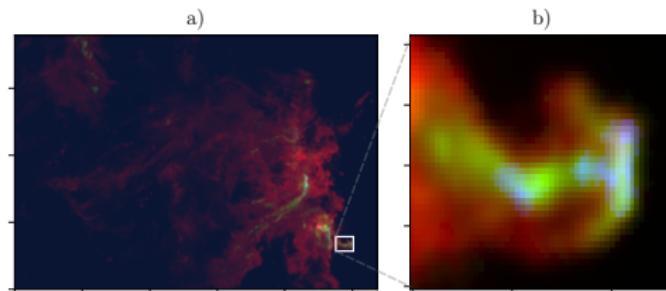
- [72, 116] GHz
- 5 deg 2 , résolution angulaire $\sim 27''$
- $\sim 850\text{h}$ d'observation

\Rightarrow Méthodologie d'analyse **automatisée**

Introduction

Contexte et objectifs

a) Image composite du nuage géant moléculaire **Orion B** ($J = 1 \rightarrow 0$ de ^{12}CO , C^{18}O , et H^{13}CO^+). b) La nébuleuse de la **Tête de Cheval**.



Le *ORION-B Large Program* (P.I. : J. Pety & M. Gerin)



🔍 Formation des étoiles dans le milieu interstellaire

⌚ Estimer les propriétés du nuage

- n [cm $^{-3}$] : densité volumique de H $_2$
- T [K] : température cinétique
- ...

📡 Radiotélescope **IRAM-30m** (Sierra Nevada)

- [72, 116] GHz
- 5 deg 2 , résolution angulaire $\sim 27''$
- ~ 850 h d'observation

1074 × 758 pixels
× 200 000 canaux fréq.

⇒ Méthodologie d'analyse **automatisée**

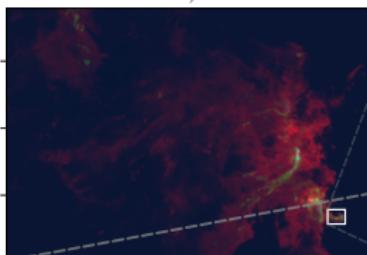
🎥 La nébuleuse de la **Tête de Cheval** ~ 1500 pixels

Introduction

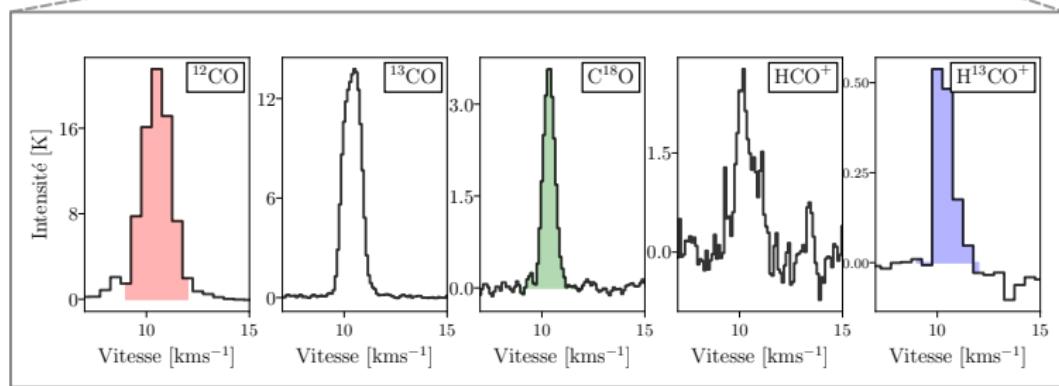
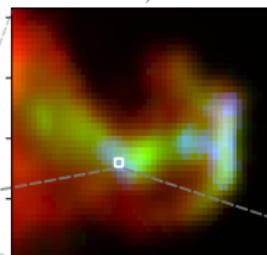
Contexte et objectifs

a) Image composite du nuage géant moléculaire **Orion B** ($J = 1 \rightarrow 0$ de ^{12}CO , C^{18}O , et H^{13}CO^+). b) La nébuleuse de la **Tête de Cheval**.
Conversion fréq. → vitesse (effet Doppler $\sim \frac{f_{\text{obs}}}{f_{\text{repos}}} = 1 - \frac{v_{\text{source/obs}}}{c}$).

a)



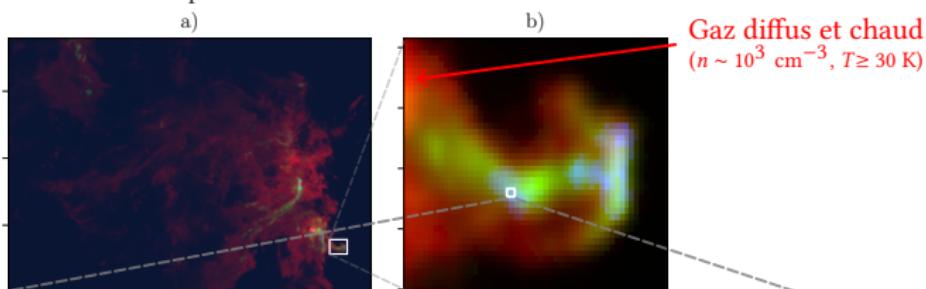
b)



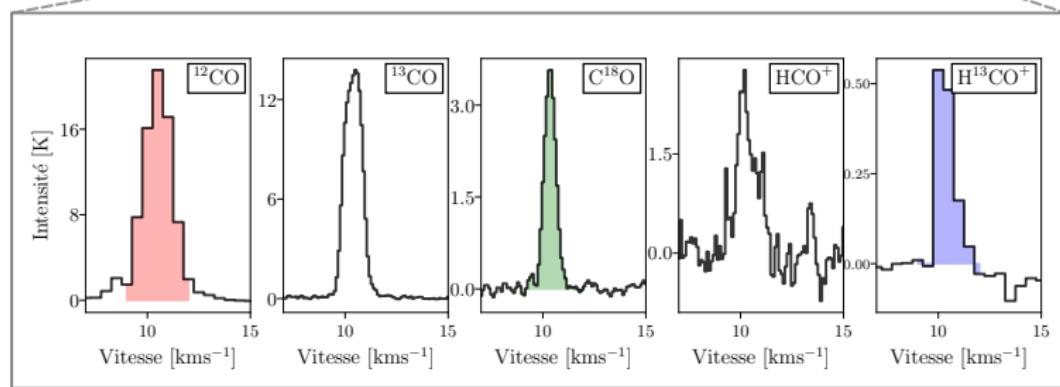
Introduction

Contexte et objectifs

a) Image composite du nuage géant moléculaire **Orion B** ($J = 1 \rightarrow 0$ de ^{12}CO , C^{18}O , et H^{13}CO^+). b) La nébuleuse de la **Tête de Cheval**.
Conversion fréq. → vitesse (effet Doppler $\sim \frac{f_{\text{obs}}}{f_{\text{repos}}} = 1 - \frac{v_{\text{source/obs}}}{c}$).



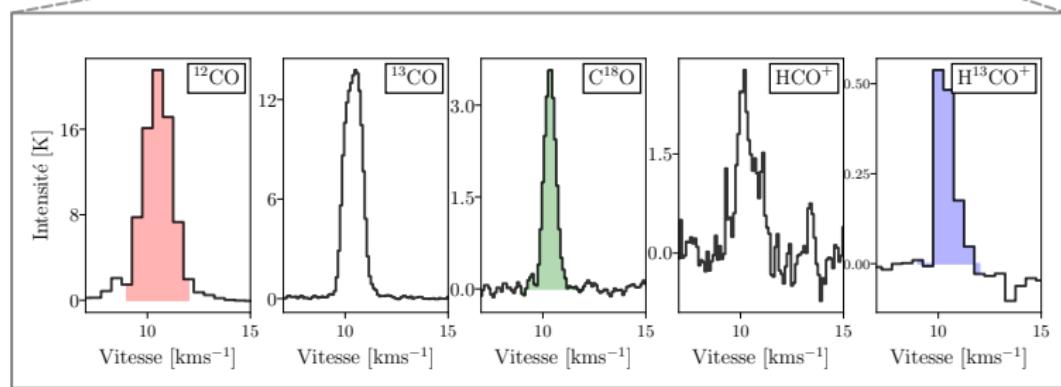
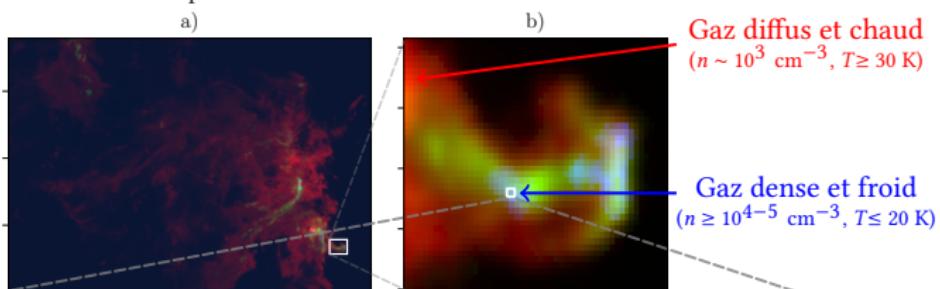
Gaz diffus et chaud
($n \sim 10^3 \text{ cm}^{-3}$, $T \geq 30 \text{ K}$)



Introduction

Contexte et objectifs

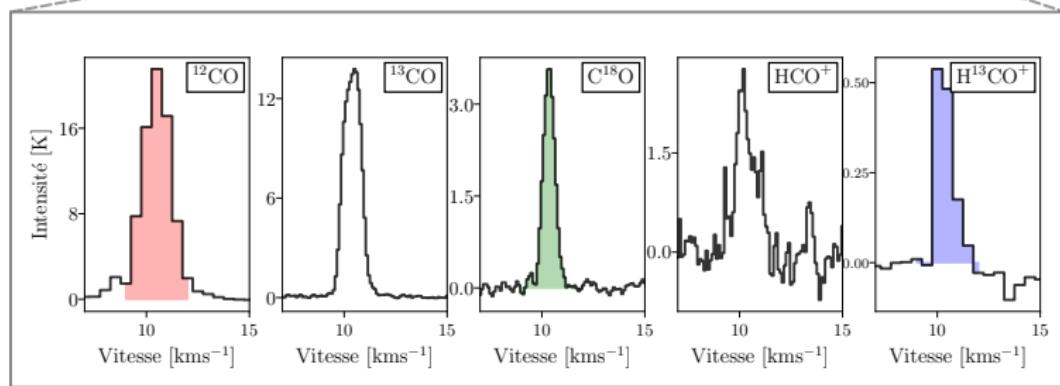
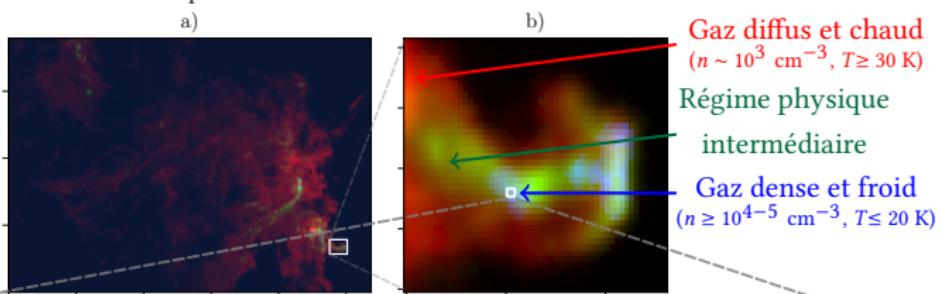
a) Image composite du nuage géant moléculaire **Orion B** ($J = 1 \rightarrow 0$ de ^{12}CO , C^{18}O , et H^{13}CO^+). b) La nébuleuse de la **Tête de Cheval**.
Conversion fréq. → vitesse (effet Doppler $\sim \frac{f_{\text{obs}}}{f_{\text{repos}}} = 1 - \frac{v_{\text{source/obs}}}{c}$).



Introduction

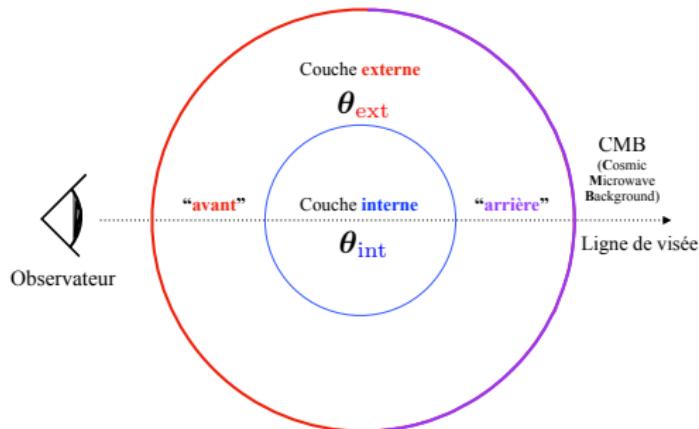
Contexte et objectifs

a) Image composite du nuage géant moléculaire **Orion B** ($J = 1 \rightarrow 0$ de ^{12}CO , C^{18}O , et H^{13}CO^+). b) La nébuleuse de la **Tête de Cheval**.
Conversion fréq. → vitesse (effet Doppler $\sim \frac{f_{\text{obs}}}{f_{\text{repos}}} = 1 - \frac{v_{\text{source/obs}}}{c}$).



Introduction

Modèle *hétérogène* d'émission en chaque pixel (ligne de visée)



Décomposition du spectre en différentes contributions

$$s(\theta) = s_{\text{avant}}(\theta) + s_{\text{interne}}(\theta) + s_{\text{arrière}}(\theta) + s_{\text{CMB}}(\theta)$$

- $\theta = \{\theta_{\text{ext}}, \theta_{\text{int}}\}$, avec $\theta_{\text{ext}} = (T_{\text{ext}}, n_{\text{ext}}, \dots)$ et $\theta_{\text{int}} = (T_{\text{int}}, n_{\text{int}}, \dots) \Rightarrow 14$ paramètres inconnus
- Équation de transfert radiatif $\theta \rightarrow$ RADEX (van der Tak et al. 2010) $\rightarrow s(\theta) \Rightarrow$ pas d'expression explicite

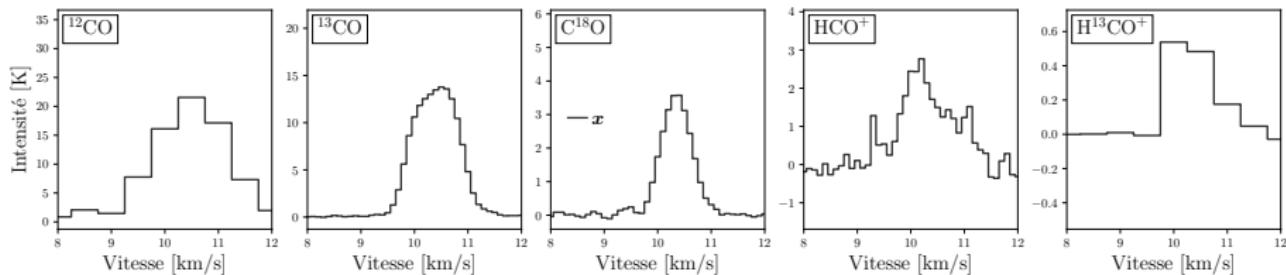
Introduction

Modèle de mesure en chaque pixel (ligne de visée)

Pour chaque spectre \mathbf{x} ,

$$\boxed{\mathbf{x} = s(\boldsymbol{\theta}) + \mathbf{b}} \quad (1)$$

- \mathbf{b} : bruit additif blanc, Gaussien (contribution atmosphérique)
- $\boldsymbol{\theta} \in \mathbb{R}^{14}$: vecteur des paramètres physiques



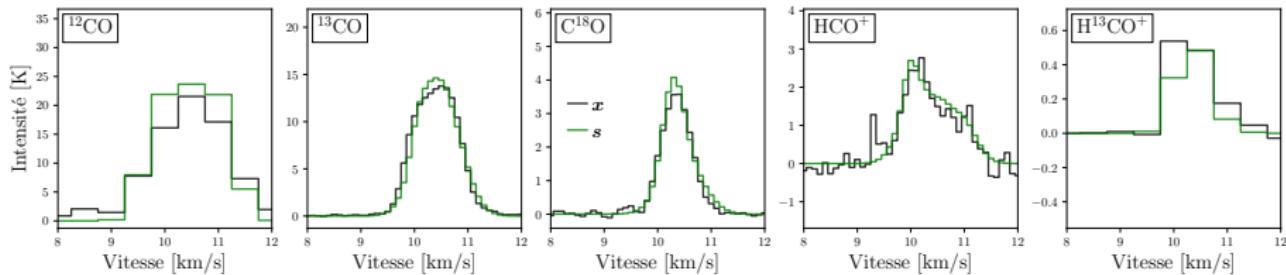
Introduction

Modèle de mesure en chaque pixel (ligne de visée)

Pour chaque spectre \mathbf{x} ,

$$\mathbf{x} = \mathbf{s}(\hat{\boldsymbol{\theta}}) + \mathbf{b} \quad (1)$$

- \mathbf{b} : bruit additif blanc, Gaussien (contribution atmosphérique)
- $\hat{\boldsymbol{\theta}} \in \mathbb{R}^{14}$: vecteur des paramètres physiques estimé par maximum de vraisemblance (EMV), i.e., $\hat{\boldsymbol{\theta}} = \operatorname{argmin}_{\boldsymbol{\theta}} [-\log \mathcal{L}(\boldsymbol{\theta})]$



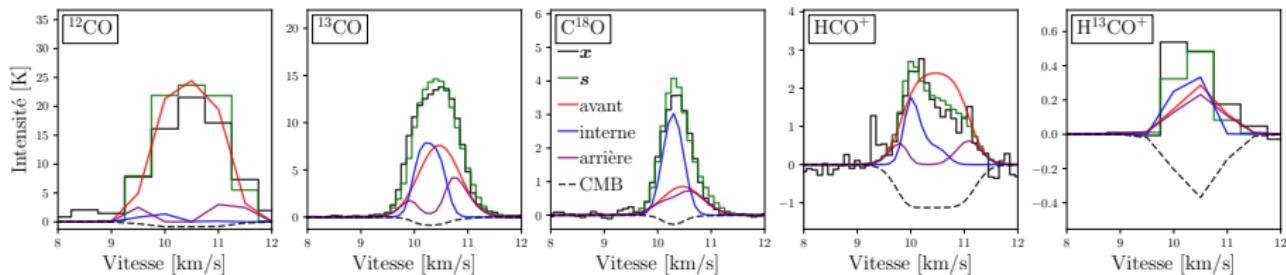
Introduction

Modèle de mesure en chaque pixel (ligne de visée)

Pour chaque spectre \mathbf{x} ,

$$\mathbf{x} = \mathbf{s}(\hat{\boldsymbol{\theta}}) + \mathbf{b} \quad (1)$$

- \mathbf{b} : bruit additif blanc, Gaussien (contribution atmosphérique)
 - $\hat{\boldsymbol{\theta}} \in \mathbb{R}^{14}$: vecteur des paramètres physiques estimé par maximum de vraisemblance (EMV), i.e., $\hat{\boldsymbol{\theta}} = \operatorname{argmin}_{\boldsymbol{\theta}} [-\log \mathcal{L}(\boldsymbol{\theta})]$
- ⇒ décomposition des spectres en différentes contributions \sim interprétabilité des résultats sur la Tête de Cheval (Ségal et al. 2024)



Introduction

Modèle de mesure en chaque pixel (ligne de visée)

Pour chaque spectre \mathbf{x} ,

$$\boxed{\mathbf{x} = \mathbf{c} \times \mathbf{s}(\boldsymbol{\theta}) + \mathbf{b}} \quad (2)$$

- \mathbf{b} : bruit additif blanc, Gaussien (contribution atmosphérique)
- $\hat{\boldsymbol{\theta}} \in \mathbb{R}^{14}$: vecteur des paramètres physiques estimé par maximum de vraisemblance (EMV), i.e., $\hat{\boldsymbol{\theta}} = \operatorname{argmin}_{\boldsymbol{\theta}} [-\log \mathcal{L}(\boldsymbol{\theta})]$
- $\mathbf{c} \in \mathbb{R}$: bruit multiplicatif de “calibration” (chaîne de mesure), $c \sim \mathcal{N}(1, \sigma_c^2)$ où $\sigma_c \in \{0.05, 0.1\}$ (Einig et al. 2023)
⇒ performance de l’EMV construit sur le modèle simplifié (1) ?

Introduction

Modèle de mesure en chaque pixel (ligne de visée)

Pour chaque spectre \mathbf{x} ,

$$\boxed{\mathbf{x} = \mathbf{c} \times \mathbf{s}(\boldsymbol{\theta}) + \mathbf{b}} \quad (2)$$

- \mathbf{b} : bruit additif blanc, Gaussien (contribution atmosphérique)
- $\hat{\boldsymbol{\theta}} \in \mathbb{R}^{14}$: vecteur des paramètres physiques estimé par maximum de vraisemblance (EMV), i.e., $\hat{\boldsymbol{\theta}} = \operatorname{argmin}_{\boldsymbol{\theta}} [-\log \mathcal{L}(\boldsymbol{\theta})]$
- $\mathbf{c} \in \mathbb{R}$: bruit multiplicatif de “calibration” (chaîne de mesure), $\mathbf{c} \sim \mathcal{N}(1, \sigma_c^2)$ où $\sigma_c \in \{0.05, 0.1\}$ (Einig et al. 2023)

Quantification de la dégradation des perf. d'estimation

Huber 1967, Akaike 1973, White 1982 \Rightarrow comportement asymptotique de l'EMV
en présence d'erreur de modèle

Dans cette étude GRETSI

- application sur les données de la **Tête de Cheval**
- application sur un modèle “physique” simple \leadsto expression analytique (cf. papier)

Comportement asymptotique de l'EMV

Soit $\boldsymbol{\theta} \in \mathbb{R}^M$ et $\mathbf{y} = \{\mathbf{x}_1, \dots, \mathbf{x}_L\}$, où

$$\mathbf{x}_l = \mathbf{c}_l \times \mathbf{s}_l(\boldsymbol{\theta}) + \mathbf{b}_l \quad (2)$$

$$\hat{\boldsymbol{\theta}} = \arg \min_{\boldsymbol{\theta}} \{-\ln p_Y(\mathbf{y}|\boldsymbol{\theta})\}$$

Propriétés asymptotiques (e.g., [kay1993](#))

- Gaussien, i.e., $\hat{\boldsymbol{\theta}} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$
- $\boldsymbol{\mu} = \boldsymbol{\theta}_{\text{vrai}} \rightsquigarrow$ non biaisé
- $\boldsymbol{\Sigma} = \mathcal{B}(\boldsymbol{\theta}) \rightsquigarrow$ efficace



Borne de Cramér-Rao (BCR) où $\mathcal{B}(\boldsymbol{\theta}) = \mathcal{I}^{-1}(\boldsymbol{\theta})$ (Information de Fisher)

$$\mathcal{I}(\boldsymbol{\theta}) = \mathbb{E} \left[\nabla_{\boldsymbol{\theta}} \ln p_Y(\mathbf{Y}|\boldsymbol{\theta}) \nabla_{\boldsymbol{\theta}}^T \ln p_Y(\mathbf{Y}|\boldsymbol{\theta}) \Big| \boldsymbol{\theta} \right]$$

Comportement asymptotique de l'EMV

Soit $\boldsymbol{\theta} \in \mathbb{R}^M$ et $\mathbf{y} = \{\mathbf{x}_1, \dots, \mathbf{x}_L\}$, où

$$\mathbf{x}_l = \mathbf{c}_l \times \mathbf{s}_l(\boldsymbol{\theta}) + \mathbf{b}_l \quad (2)$$

$$\hat{\boldsymbol{\theta}} = \arg \min_{\boldsymbol{\theta}} \{-\ln p_Y(\mathbf{y}|\boldsymbol{\theta})\}$$

Propriétés asymptotiques (e.g., [kay1993](#))

- Gaussien, i.e., $\hat{\boldsymbol{\theta}} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$
- $\boldsymbol{\mu} = \boldsymbol{\theta}_{\text{vrai}} \rightsquigarrow$ non biaisé
- $\boldsymbol{\Sigma} = \mathcal{B}(\boldsymbol{\theta}) \rightsquigarrow$ efficace



Borne de Cramér-Rao (BCR) où $\mathcal{B}(\boldsymbol{\theta}) = \mathcal{I}^{-1}(\boldsymbol{\theta})$ (Information de Fisher)

$$\mathcal{I}(\boldsymbol{\theta}) = \mathbb{E} \left[\nabla_{\boldsymbol{\theta}} \ln p_Y(Y|\boldsymbol{\theta}) \nabla_{\boldsymbol{\theta}}^T \ln p_Y(Y|\boldsymbol{\theta}) \right]$$

$$\left. \begin{array}{l} \\ \\ \end{array} \right\} \simeq \hat{\theta}_i \pm \sqrt{\mathcal{B}(\hat{\theta}_i)}$$

BCR \rightsquigarrow intervalle de confiance

Comportement asymptotique de l'EMV

Soit $\boldsymbol{\theta} \in \mathbb{R}^M$ et $\mathbf{y} = \{\mathbf{x}_1, \dots, \mathbf{x}_L\}$, où

$$\mathbf{x}_l = s_l(\boldsymbol{\theta}) + \mathbf{b}_l \quad (1)$$

$$\hat{\boldsymbol{\theta}}_f = \arg \min_{\boldsymbol{\theta}} \{-\ln p_{Y,f}(\mathbf{y}|\boldsymbol{\theta})\}$$

Propriétés asymptotiques (e.g., [kay1993](#))

- Gaussien, i.e., $\hat{\boldsymbol{\theta}} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$
- $\boldsymbol{\mu} = \boldsymbol{\theta}_{\text{vrai}} \rightsquigarrow$ non biaisé
- $\boldsymbol{\Sigma} = \mathcal{B}(\boldsymbol{\theta}) \rightsquigarrow$ efficace



Borne de Cramér-Rao (BCR) où $\mathcal{B}(\boldsymbol{\theta}) = \mathcal{I}^{-1}(\boldsymbol{\theta})$ (Information de Fisher)

$$\mathcal{I}(\boldsymbol{\theta}) = \mathbb{E} \left[\nabla_{\boldsymbol{\theta}} \ln p_Y(Y|\boldsymbol{\theta}) \nabla_{\boldsymbol{\theta}}^\top \ln p_Y(Y|\boldsymbol{\theta}) \right]$$

}

$$\simeq \hat{\theta}_i \pm \sqrt{\mathcal{B}(\hat{\theta}_i)}$$

BCR \rightsquigarrow intervalle
de confiance

Comportement asymptotique de l'EMV mal spécifié

Soit $\boldsymbol{\theta} \in \mathbb{R}^M$ et $\mathbf{y} = \{\mathbf{x}_1, \dots, \mathbf{x}_L\}$, où

$$\mathbf{x}_l = s_l(\boldsymbol{\theta}) + \mathbf{b}_l \quad (1)$$

$$\hat{\boldsymbol{\theta}}_f = \arg \min_{\boldsymbol{\theta}} \{-\ln p_{Y,f}(y|\boldsymbol{\theta})\}$$

Propriétés asymptotiques

- Gaussien, i.e., $\hat{\boldsymbol{\theta}}_f \sim \mathcal{N}(\boldsymbol{\mu}_f, \boldsymbol{\Sigma}_f)$ (Huber 1967)
- $\boldsymbol{\mu}_f = \boldsymbol{\theta}_{ps}$ “pseudo-vrai” \leadsto biaisé
- $\boldsymbol{\Sigma}_f = \mathcal{B}_M(\boldsymbol{\theta})$

$$\arg \min_{\boldsymbol{\theta}} \mathcal{K} \left[p_Y(Y|\boldsymbol{\theta}_{\text{vrai}}) \| p_{Y,f}(Y|\boldsymbol{\theta}) \right] \text{ Divergence de Kullback-Leibler (Akaike 1973)}$$

$$\downarrow$$

Borne de Cramér-Rao “Modifiée” (BCRM) où $\mathcal{B}_M(\boldsymbol{\theta}) = \mathbf{A}^{-1} \mathbf{J} \mathbf{A}^{-1}$ (White 1982, Ren et al. 2015)

$$\mathbf{J} = \mathbb{E} \left[\nabla_{\boldsymbol{\theta}} \ln p_{Y,f}(Y|\boldsymbol{\theta}) \nabla_{\boldsymbol{\theta}}^T \ln p_{Y,f}(Y|\boldsymbol{\theta}) |_{\boldsymbol{\theta}=\boldsymbol{\theta}_{ps}} \right] \quad \text{et} \quad \mathbf{A} = -\mathbb{E} \left[\nabla_{\boldsymbol{\theta}}^2 \ln p_{Y,f}(Y|\boldsymbol{\theta}) |_{\boldsymbol{\theta}=\boldsymbol{\theta}_{ps}} \right]$$

Comportement asymptotique de l'EMV mal spécifié

Soit $\boldsymbol{\theta} \in \mathbb{R}^M$ et $\mathbf{y} = \{\mathbf{x}_1, \dots, \mathbf{x}_L\}$, où

$$\mathbf{x}_l = s_l(\boldsymbol{\theta}) + \mathbf{b}_l \quad (1)$$

$$\hat{\boldsymbol{\theta}}_f = \arg \min_{\boldsymbol{\theta}} \{-\ln p_{Y,f}(y|\boldsymbol{\theta})\}$$

Propriétés asymptotiques

- Gaussien, i.e., $\hat{\boldsymbol{\theta}}_f \sim \mathcal{N}(\boldsymbol{\mu}_f, \boldsymbol{\Sigma}_f)$ (Huber 1967)
- $\boldsymbol{\mu}_f = \boldsymbol{\theta}_{\text{vrai}} \rightsquigarrow$ non biaisé
- $\boldsymbol{\Sigma}_f = \mathcal{B}_M(\boldsymbol{\theta})$

\downarrow $\arg \min_{\boldsymbol{\theta}} \mathcal{K} \left[p_Y(Y|\boldsymbol{\theta}_{\text{vrai}}) \| p_{Y,f}(Y|\boldsymbol{\theta}) \right]$ Divergence de Kullback-Leibler (Akaike 1973)

$$\downarrow$$

Borne de Cramér-Rao “Modifiée” (BCRM) où $\mathcal{B}_M(\boldsymbol{\theta}) = \mathbf{A}^{-1} \mathbf{J} \mathbf{A}^{-1}$ (White 1982, Ren et al. 2015)

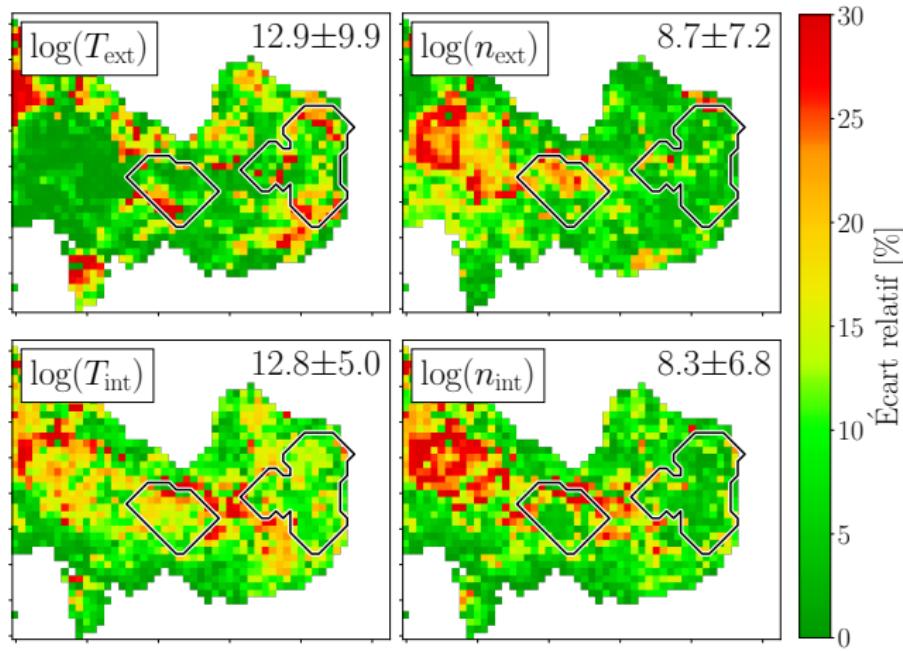
$$\mathbf{J} = \mathbb{E} \left[\nabla_{\boldsymbol{\theta}} \ln p_{Y,f}(Y|\boldsymbol{\theta}) \nabla_{\boldsymbol{\theta}}^T \ln p_{Y,f}(Y|\boldsymbol{\theta}) |_{\boldsymbol{\theta}=\boldsymbol{\theta}_{\text{vrai}}} \right] \quad \text{et} \quad \mathbf{A} = -\mathbb{E} \left[\nabla_{\boldsymbol{\theta}}^2 \ln p_{Y,f}(Y|\boldsymbol{\theta}) |_{\boldsymbol{\theta}=\boldsymbol{\theta}_{\text{vrai}}} \right]$$

○ Analyse de précision $\Rightarrow \mathcal{B}_M$ vs. \mathcal{B}

Application sur la Tête de Cheval

Quantification de la dégradation des perf. d'estimation

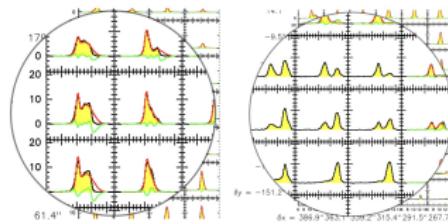
$$\text{Écart relatif } \epsilon, \text{ i.e., } \epsilon(\tau) \triangleq \frac{\sqrt{\mathcal{B}_M} - \sqrt{\mathcal{B}}}{\sqrt{\mathcal{B}}}$$



Conclusion et perspectives

Application sur les données ORION-B

- ① EMV sans bruit multiplicatif → perte de précision 10 – 30%
 - ② EMV avec bruit multiplicatif ↗ risque de **sur-apprentissage**
 - ③ Analyser le reste des données ORION-B ⇒ ⚡ phénomène de **fusion**
- écart au
modèle physique >> bruit multiplicatif



Défis pour d'autres applications

- ⌚ Si écart-type élevé ($\sigma_c \geq 30\%$) ⇒ approximation **Gaussienne** invalide
- Calculs et implémentation de l'EMV + difficiles

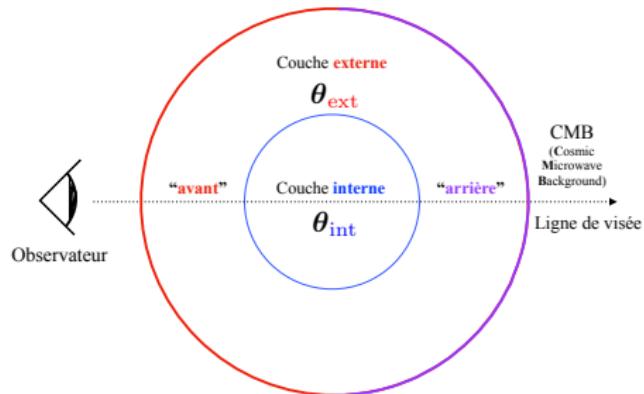
Références I

-  Akaike, H. (1973). "Information Theory and an Extension of the Maximum Likelihood Principle". In: Proceeding of the Second International Symposium on Information Theory. Budapest: Akadémiai Kiadó.
-  Einig, L. et al. (2023). "Deep learning denoising by dimension reduction: Application to the ORION-B line cubes". In: A&A 677, A158. doi: [10.1051/0004-6361/202346064](https://doi.org/10.1051/0004-6361/202346064).
-  Galton, F. (1879). "XII. The geometric mean, in vital and social statistics". eng. In: Proceedings of the Royal Society of London 29.196, pp. 365–367. ISSN: 0370-1662.
-  Huber, P.J. (1967). "The Behavior of Maximum Likelihood Estimates Under Nonstandard Conditions". In: Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, University of California Press, pp. 221–233.
-  Kay, S. M. (1993). "Fundamentals of statistical signal Proc.: estimation theory". In: USA: Prentice-Hall, Inc. Chap. 3. ISBN: 0133457117.
-  Petersen, K. B. and M. S. Pedersen (Oct. 2008). The Matrix Cookbook. Version 20081110.
-  Ren et al. (2015). "Performance bounds under misspecification model for MIMO radar application". In: 2015 23rd EUSIPCO, pp. 514–518. doi: [10.1109/EUSIPCO.2015.7362436](https://doi.org/10.1109/EUSIPCO.2015.7362436).
-  Ségal, L. et al. (2024). "Toward a robust physical and chemical characterization of heterogeneous lines of sight: The case of the Horsehead nebula". In: A&A 692, A160. doi: [10.1051/0004-6361/202451567](https://doi.org/10.1051/0004-6361/202451567).

Références II

-  van der Tak, F. F. S. et al. (Oct. 2010). [Radex: Fast Non-LTE Analysis of Interstellar Line Spectra](#). Astrophysics Source Code Library, record ascl:1010.075.
-  White, H. (1982). “Maximum Likelihood Estimation of Misspecified Models”. In: [Econometrica](#) 50.1, pp. 1–25. ISSN: 00129682, 14680262.

Modèle des spectres en une ligne de visée



$\theta = \{\theta_{\text{ext}}, \theta_{\text{int}}\}$ where $\theta_i = \{\log T_i, \log n_i, \log N_i, \text{FWHM}_i, C_{V,i}\}$, for $i \in \{\text{ext}, \text{int}\}$

Équation de transfert radiatif

$$s = s_{\text{avant}} + s_{\text{interne}} + s_{\text{arrière}} + s_{\text{CMB}}$$

Intensité [K]

$$J(T_{\text{ex}}) = \frac{hv}{k} \frac{1}{\exp\left(\frac{hv}{kT_{\text{ex}}}\right) - 1}$$

$$s_{\text{avant}} = J(T_{\text{ex,ext}}) [1 - \exp(-\Psi_{\text{ext}})]$$

$$s_{\text{interne}} = J(T_{\text{ex,int}}) [1 - \exp(-\Psi_{\text{int}})] \exp(-\Psi_{\text{ext}})$$

$$s_{\text{arrière}} = J(T_{\text{ex,ext}}) [1 - \exp(-\Psi_{\text{ext}})] \exp(-\Psi_{\text{ext}} - \Psi_{\text{int}})$$

$$s_{\text{CMB}} = J(T_{\text{CMB}}) [\exp(-2 \times \Psi_{\text{ext}} - \Psi_{\text{int}})] - J(T_{\text{CMB}})$$

where $T_{\text{CMB}} = 2.73 \text{ K}$

Profile d'opacité [-]

$$\Psi = \tau \exp\left(\frac{-(v - C_V)^2}{2\sigma_V^2}\right)$$

where $\sigma_V = \text{FWHM}/\sqrt{8\ln 2}$

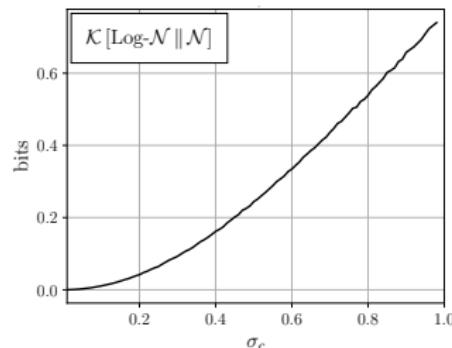
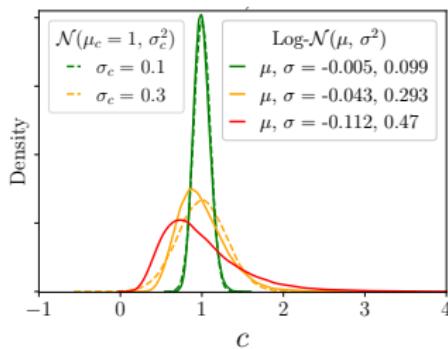
$\theta \rightarrow \text{RADEX (van der Tak et al. 2010)} \rightarrow \{T_{\text{ex}}, \tau\}$



Spécification d'un bruit multiplicatif de faible écart-type

Approximation Gaussienne d'une log-Gaussienne

- bruit multiplicatif $\sim \text{Log-}\mathcal{N}(\mu, \sigma^2)$ (e.g., Galton 1879)
- ajustement Gaussien, i.e., $\mathcal{N}(\mu_c = 1, \sigma_c^2)$
 - ✓ $\sigma_c \lesssim 0.1$ (e.g., le jeu de données ORION-B)
 - ~ $\sigma_c \in [0.1, 0.3]$
 - ✗ $\sigma_c \gtrsim 0.3$



⇒ cadre d'étude : bruit multiplicatif d'écart-type $\lesssim 0.3$
de sorte que $c \sim \mathcal{N}(\mu_c = 1, \sigma_c^2)$

GauSSIannité des observations

$$\mathbf{X} = C \mathbf{s}(\boldsymbol{\theta}) + \mathbf{B} \quad \Rightarrow \quad \begin{pmatrix} \mathbf{X} \\ C \end{pmatrix} = \underbrace{\begin{pmatrix} \mathbf{s}(\boldsymbol{\theta}) & \mathbf{I}_N \\ 1 & \mathbf{0}_N^\top \end{pmatrix}}_{\text{déterministe}} \underbrace{\begin{pmatrix} C \\ \mathbf{B} \end{pmatrix}}_{\text{Gaussien}}$$

La p.d.f. marginale s'exprime donc

$$p(\mathbf{x}|\boldsymbol{\theta}) = \frac{1}{\sqrt{\det(2\pi \boldsymbol{\Sigma})}} \exp \left\{ -\frac{1}{2} \boldsymbol{\epsilon}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\epsilon} \right\}, \quad (1)$$

où $\boldsymbol{\epsilon}(\boldsymbol{\theta}) = \mathbf{x} - \boldsymbol{\mu}(\boldsymbol{\theta})$, et

$$\boldsymbol{\mu}(\boldsymbol{\theta}) = \mathbf{s}(\boldsymbol{\theta}) \quad \text{et} \quad \boldsymbol{\Sigma}(\boldsymbol{\theta}) = \sigma_c^2 \mathbf{s}(\boldsymbol{\theta}) \mathbf{s}(\boldsymbol{\theta})^\top + \sigma_b^2 \mathbf{I}_N. \quad (2)$$

Par ailleurs, $\det(\mathbf{I}_N + \mathbf{u}\mathbf{v}^\top) = 1 + \mathbf{u}^\top \mathbf{v}$ (e.g., Petersen et al. 2008, “*The Matrix Cookbook*”, Eq. (24)). Ainsi,

$$\det(\boldsymbol{\Sigma}) = \sigma_b^{2N} \left(1 + \frac{\sigma_c^2}{\sigma_b^2} \mathbf{s}^\top \mathbf{s} \right).$$

et la formule de *Sherman-Morrison* mène à

$$\boldsymbol{\Sigma}_2^{-1} = \frac{1}{\sigma_b^2} \left(\mathbf{I}_N - \frac{\sigma_c^2 \mathbf{s} \mathbf{s}^\top}{\sigma_b^2 + \sigma_c^2 \mathbf{s}^\top \mathbf{s}} \right).$$

Divergence de Kullback-Leibler

EMV mal spécifié non biaisé

Modèle des mesures :

$$\mathbf{X} = C \mathbf{s}(\boldsymbol{\theta}_{\text{vrai}}) + \mathbf{B},$$

avec

$$\boldsymbol{\mu}(\boldsymbol{\theta}_{\text{vrai}}) = \mathbf{s}(\boldsymbol{\theta}_{\text{vrai}}) \quad \text{and} \quad \boldsymbol{\Sigma}(\boldsymbol{\theta}_{\text{vrai}}) = \sigma_c^2 \mathbf{s}(\boldsymbol{\theta}_{\text{vrai}}) \mathbf{s}(\boldsymbol{\theta}_{\text{vrai}})^T + \sigma_b^2 \mathbf{I}_N. \quad (3)$$

Lorsque le bruit de calibration est négligé, alors

$$\boldsymbol{\mu}_{\cancel{f}}(\boldsymbol{\theta}) = \mathbf{s}(\boldsymbol{\theta}) \quad \text{and} \quad \boldsymbol{\Sigma}_{\cancel{f}} = \sigma_b^2 \mathbf{I}_N. \quad (4)$$

On a (Akaike 1973)

$$\boldsymbol{\theta}_{\text{ps}} = \arg \min_{\boldsymbol{\theta}} \mathcal{K} \left[p_{\mathbf{X}}(\mathbf{X} | \boldsymbol{\theta}_{\text{vrai}}) \| p_{\mathbf{X}, \cancel{f}}(\mathbf{X} | \boldsymbol{\theta}) \right],$$

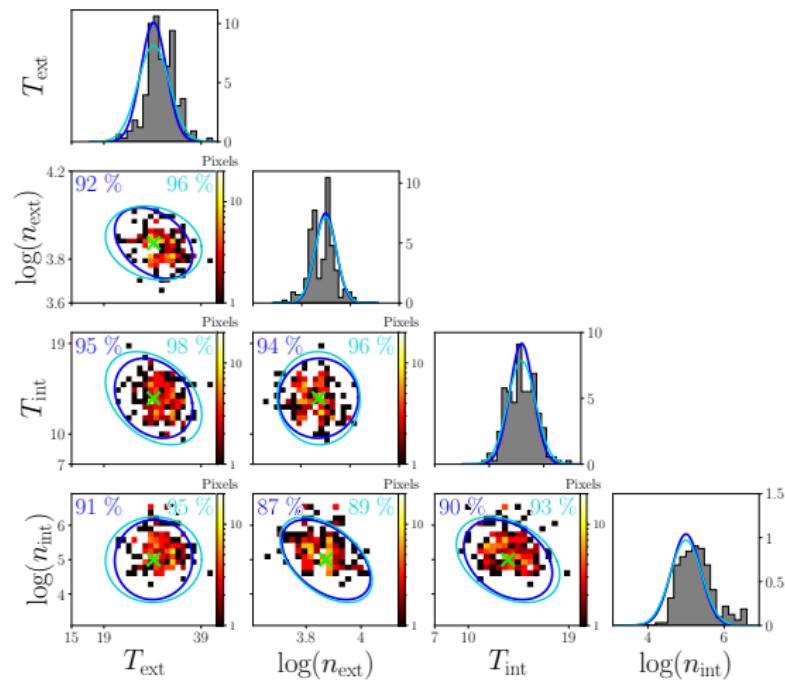
où

$$\begin{aligned} \mathcal{K} \left[p_{\mathbf{X}}(\mathbf{X} | \boldsymbol{\theta}_{\text{vrai}}) \| p_{\mathbf{X}, \cancel{f}}(\mathbf{X} | \boldsymbol{\theta}) \right] &= \frac{1}{2} \left(\text{tr} \left(\boldsymbol{\Sigma}_{\cancel{f}}^{-1} \boldsymbol{\Sigma} \right) - N + \ln \left(\frac{\det \boldsymbol{\Sigma}_{\cancel{f}}}{\det \boldsymbol{\Sigma}} \right) \right) + \frac{1}{2} (\boldsymbol{\mu}_{\cancel{f}} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}_{\cancel{f}}^{-1} (\boldsymbol{\mu}_{\cancel{f}} - \boldsymbol{\mu}) \\ &= \text{cst} + \frac{1}{2\sigma_b^2} (\mathbf{s}(\boldsymbol{\theta}) - \mathbf{s}(\boldsymbol{\theta}_{\text{vrai}}))^T (\mathbf{s}(\boldsymbol{\theta}) - \mathbf{s}(\boldsymbol{\theta}_{\text{vrai}})). \end{aligned}$$

Analyse de l'efficacité

Monte-Carlo sur un cœur dense (simulation)

Histogrammes des estimations (200 échantillons). La croix indique θ_{vrai} . Les ellipses de confiance à 99% sont construites à partir des BCR and BCRM. La proportion (%) des bins contenus dans ces ellipses est indiquée en haut de chaque panneau.



Calcul de la BCR lors d'un changement de variable

Soit le vecteur de paramètres inconnus $\boldsymbol{\theta} = (\theta_1 \cdots \theta_M)^\top$ de BCR correspondante $\mathcal{B}(\boldsymbol{\theta})$. On cherche la BCR sur $\boldsymbol{\alpha} = \mathbf{f}(\boldsymbol{\theta}) \in \mathbb{R}^M$ notée $\mathcal{B}(\boldsymbol{\alpha})$, telle que

$$\forall i \in \{1, \dots, M\}, \quad f_i(\theta) = \begin{cases} 10^\theta & \text{si } \theta \in \{\log T, \log n, \dots\}, \\ \theta & \text{sinon.} \end{cases} \quad (5)$$

D'après [Kay 1993, Sect. 3.8](#),

$$\mathcal{B}(\boldsymbol{\alpha}) = \frac{\partial \mathbf{f}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \mathcal{B}(\boldsymbol{\theta}) \frac{\partial \mathbf{f}(\boldsymbol{\theta})^\top}{\partial \boldsymbol{\theta}},$$

où

$$\left[\frac{\partial \mathbf{f}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right]_{i,i} = \begin{cases} \ln 10 \times 10^{\theta_i} & \text{si } \theta_i \in \{\log T, \log n, \dots\}, \\ 1 & \text{sinon.} \end{cases}$$

D'où, e.g.,

$$\mathcal{B}(T) = (\ln 10 \times T)^2 \mathcal{B}(\log T).$$

La précision est alors donnée par $\sqrt{\mathcal{B}(T)} = (\ln 10 \times T) \sqrt{\mathcal{B}(\log T)}$.