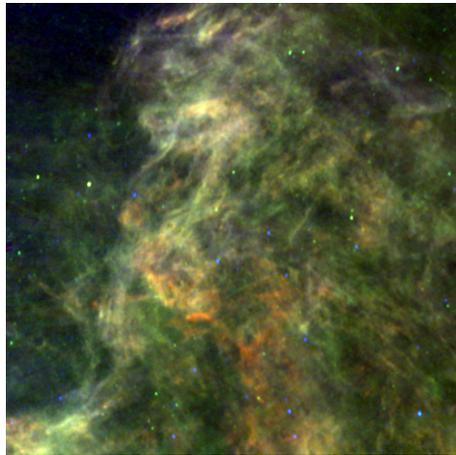


**STAGE DE RECHERCHE, MASTER 2**

Manuel Berthet

**Statistique d'orientation des filaments et du champ  
magnétique dans le milieu interstellaire**



François Levrier, Edith Falgarone

*LERMA/LRA - Observatoire de Paris / Ecole Normale  
Supérieure de Paris*

19/03/2012 - 29/06/2012

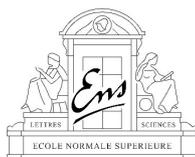


Image sur la page de garde provenant de IRIS - Retraitement des données IRAS [7].

*Polaris Flare* : nom dû à sa forme d'éruption.

Image couleur reconstituée à partir des images à 12, 25, 60 et 100  $\mu m$ .

<http://www.cita.utoronto.ca/~mamd/IRIS/IrisOverview.html>

Document édité avec L<sup>A</sup>T<sub>E</sub>X. Version du 22 juin 2012.

## Résumé

On cherche à contraindre les modèles de formation stellaire via l'étude des filaments de matière dont les modèles actuels font des précurseurs des cœurs denses pré-stellaires. On cherche en particulier à comprendre le lien entre distribution de matière et champ magnétique. Nous disposons pour cela, dans le champ de Polaris, de la polarisation de la lumière d'étoiles d'arrière-plan et de cartes d'émission des poussières montrant des structures de matière filamenteuses. Nous utilisons des algorithmes permettant d'isoler ces filaments et étudions la statistique des angles entre la polarisation de la lumière et les filaments voisins. Nous montrons que la distribution de ces angles n'est pas uniforme, avec un pic autour de  $\pi/10$ .

# Table des matières

<b>Résumé</b>	<b>3</b>
<b>Table des matières</b>	<b>4</b>
<b>Introduction</b>	<b>5</b>
<b>1 Présentation des données</b>	<b>6</b>
1.1 IRIS . . . . .	6
1.2 Polarisation . . . . .	6
<b>2 Traitement pour obtenir l'orientation des filaments</b>	<b>9</b>
2.1 Analyse en curvelets . . . . .	10
2.1.1 Transformées en ondelettes, ridgelets et curvelets . . . . .	10
2.1.2 Morphological Component Analysis (MCA) . . . . .	12
2.1.3 Application . . . . .	15
2.2 Détection des filaments . . . . .	15
2.2.1 DisPerSE, principe et fonctionnement . . . . .	15
2.2.2 Calcul du squelette . . . . .	20
<b>3 Statistique d'orientation</b>	<b>25</b>
3.1 Les filaments . . . . .	25
3.2 Le champ magnétique . . . . .	25
<b>4 Résultats et discussion</b>	<b>27</b>
<b>Conclusion</b>	<b>34</b>
<b>Remerciements</b>	<b>34</b>
<b>Références</b>	<b>35</b>
<b>Annexes</b>	<b>36</b>
<b>A Algorithme de MCA</b>	<b>36</b>
<b>B Oscillations gênantes pour MCALab</b>	<b>36</b>

# Introduction

La matière interstellaire, étant le réservoir depuis lequel sont formées les étoiles, est d'une importance capitale pour le problème toujours ouvert qu'est la compréhension de la formation des étoiles et des systèmes planétaires. La structure filamenteuse se retrouve à de multiples échelles en astrophysique, et en particulier à celles des nuages de gaz interstellaire. On y retrouve deux types de filaments : des filaments massifs, auto-gravitants, qui peuvent être le lieu d'instabilités gravitationnelles et de ce fait former des étoiles, et des filaments plus ténus, trop peu massifs pour être auto-gravitants, comme on en trouve dans le nuage Polaris. La formation des filaments et leur évolution sont donc intimement liées à la formation d'étoiles, et sont encore mal comprises. Le champ magnétique pouvant avoir une influence décisive dans ces processus astrophysiques, il est essentiel de connaître son lien avec les filaments. L'objectif de ce stage a été de mettre en évidence, s'il en est une, la corrélation entre l'orientation du champ magnétique et celle des filaments dans le nuage Polaris.

Dans un modèle simple, les poussières non sphériques du milieu interstellaire tournent sur elles-mêmes sous l'effet de couples engendrés par leur interaction avec les photons. Comme elles sont chargées, cela induit un moment magnétique, qui tend à s'aligner avec le champ magnétique, si bien que l'axe principal de ces grains s'aligne perpendiculairement au champ magnétique local. Il faut bien être conscient, cependant, que les modèles actuels sont autrement plus complexes [1]. Dans ce modèle simple, l'alignement des grains provoque, via une absorption préférentielle de la lumière visible et proche infrarouge polarisée linéairement le long de leur axe principal, une polarisation de la lumière provenant d'étoiles en arrière-plan, parallèlement au champ magnétique dans le nuage [2]. Des données de polarisation de la lumière provenant d'étoiles d'arrière-plan nous donnent donc accès au champ magnétique dans le nuage. Le nuage Polaris est alors tout trouvé pour étudier le rapport entre champ magnétique et filaments, puisque de telles données de polarisation sont disponibles en plus des images infrarouges de l'émission des poussières qui nous viennent du télescope spatial IRAS (InfraRed Astronomical Satellite). Polaris présente de plus l'avantage d'être a priori privé de formation stellaire [3], et donc des perturbations associées, qui fausseraient notre analyse en apportant des contraintes supplémentaires aux filaments. Ces données sont présentées dans la première partie du rapport.

La méthode pour analyser les filaments depuis les images IRAS est analogue à celle utilisée pour les images Herschel du nuage moléculaire IC 5146 [4]. Les images sont tout d'abord décomposées en une partie projetée sur un espace d'ondelettes, et une partie projetée sur un espace de curvelets, exploitant la différence de morphologie des éléments mis en valeur par ces transformées. Les filaments sont ainsi obtenus avec un contraste rehaussé, et privés des sources ponctuelles. L'algorithme `DisPerSE` est alors appliqué à cette image traitée afin d'identifier les filaments les plus importants et les cartographier. Cette méthode est expliquée dans la deuxième partie.

Dans un troisième temps, nous calculons des statistiques d'orientation à partir de ces données pour avoir un premier aperçu des résultats de notre traitement. Nous finissons par une analyse de la corrélation entre l'orientation du champ magnétique et des filaments proches des points de mesure de la polarisation.

# 1 Présentation des données

## 1.1 IRIS

Polaris est un grand nuage moléculaire, situé proche du pôle Nord céleste, et à une haute latitude galactique ( $b \simeq 28$  degrés). Cette haute latitude galactique rend ardue l'estimation de sa distance : il y a moins d'étoiles dans sa direction, rendant difficile la statistique de rougissement habituellement utilisée. Son éloignement est estimé entre 100 et 200 pc [5], voire, plus récemment, plutôt près de la limite des 100 pc, en montrant qu'une partie du nuage, MCLD123.5+24.9, se trouvait en avant-plan de l'étoile polaire, placée à  $132 \pm 8$  pc [6]. On n'y observe a priori pas de formation stellaire [3], on élimine ainsi les potentielles perturbations qui y sont liées. Le télescope spatial IRAS (InfraRed Astronomical Satellite) nous a fourni des images de ce nuage de gaz et de poussières à quatre longueurs d'onde infrarouges, à 12, 25, 60 et 100  $\mu\text{m}$ , avec une résolution d'environ  $4'$  ( $3, 8'$  à 12  $\mu\text{m}$  et  $4, 3'$  à 100  $\mu\text{m}$ ). Rappelons qu'on trouve en effet des poussières dans le gaz interstellaire, restes d'enveloppes d'étoiles froides évoluées, ou de novae et supernovae, dont la taille varie du nanomètre au micromètre. Ces poussières absorbent une partie du rayonnement visible et ultraviolet, et se refroidissent en émettant dans l'infrarouge moyen et lointain. Nous utilisons les données retraitées IRIS (Improved Reprocessing of the IRAS Survey) [7]. Les cartes d'émission des poussières de Polaris sont présentées sous la forme d'images de 500 x 500 pixels pour un champ de 12.5 degrés x 12.5 degrés, ou encore 22 pc x 22 pc (à 100 pc), pour chacun des trois passages du satellite, ainsi qu'une quatrième comprenant la somme des trois passages. La figure 1 montre cette dernière version pour les quatre bandes infrarouges, sans aucune filtration. On voit clairement que la dernière bande, à 100  $\mu\text{m}$ , offre le meilleur contraste pour la structure filamenteuse : c'est donc sur celle-ci que nous baserons notre analyse. On peut déjà noter que l'image présente une structure dense et étendue de filaments facilement distinguables, il sera facile de trouver des filaments proche des points où l'on connaît la polarisation de la lumière stellaire.

## 1.2 Polarisation

Nous disposons de données de polarisation de la lumière en une soixantaine de points du champ, correspondant aux étoiles a priori d'arrière-plan accessibles aux instruments. Les observations ont été effectuées sur le télescope de 1.6 m de l'Observatoire du Mont Mégantic, au Québec, avec le polarimètre nommé "Beauty and the Beast", en mai 2010 [8]. On répertorie dans le tableau 1, tel que fourni dans [8], les coordonnées J2000 de l'étoile visée, sa référence dans les catalogues Hipparcos ou Tycho si elle existe, sa distance et sa magnitude, le degré de polarisation, l'angle de polarisation moyen, et les incertitudes à  $1\sigma$ . L'angle de polarisation est défini comme nul parallèlement à l'axe d'angle horaire nul, puis positivement du Nord vers l'Est (sens anti-horaire), de 0 à 180 degrés. Après discussion avec les auteurs, la définition précisée dans [8] (angle défini par rapport au Nord) n'est pas correcte. Le Nord étant placé près du centre de l'image, on prend un autre point de l'image, d'angle horaire nul, pour établir l'orientation de référence. Les coordonnées précises du Nord céleste sur l'image, (250.1, 238.9) pixels, et d'un point d'angle horaire nul, (252.3, 38.3) pixels, ont été obtenues grâce à la commande `sky2xy` des WCSTools [9], à partir des coordonnées J2000 correspondantes.

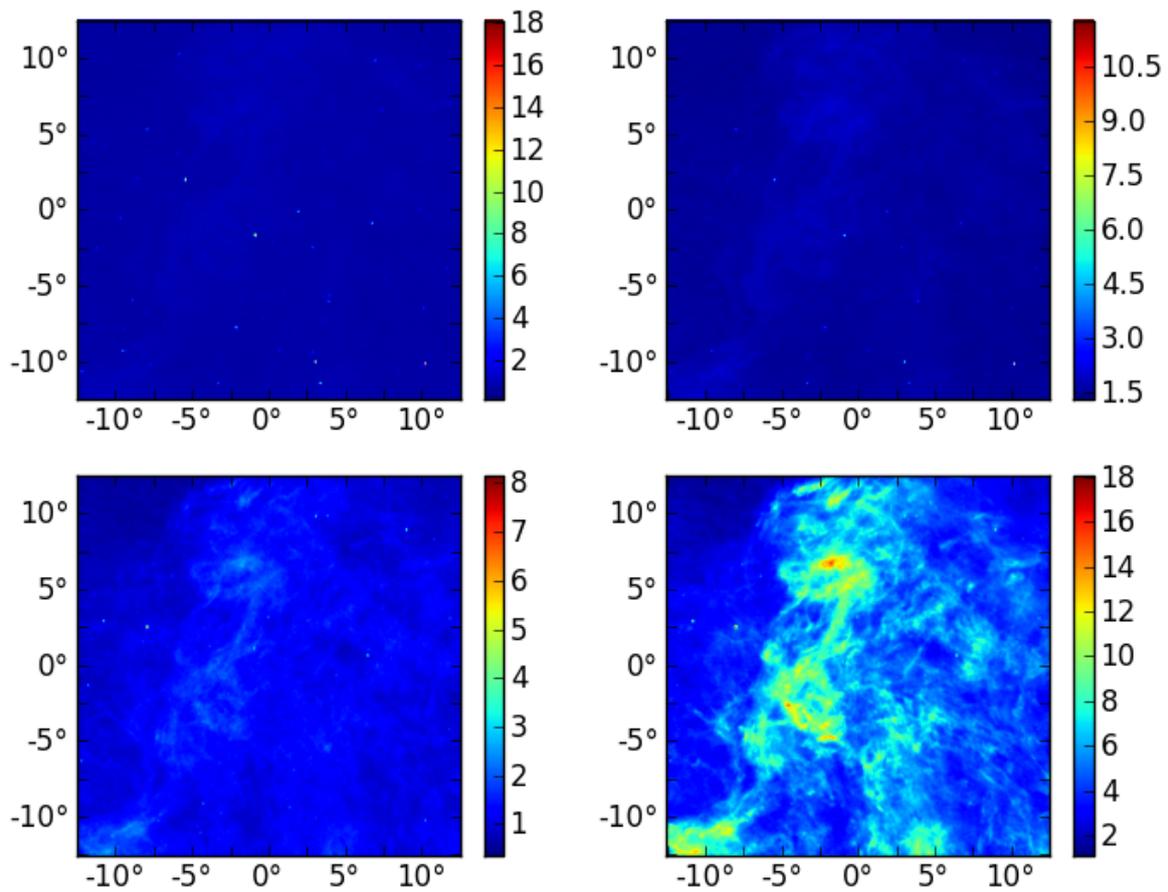


FIGURE 1 – Images IRIS de Polaris (région 430), somme des trois passages du satellite, dans les quatre longueurs d’onde infrarouges 12, 25, 60 et 100  $\mu\text{m}$ , respectivement de gauche à droite, puis de haut en bas. Seule la dernière bande offre un bon contraste pour la structure filamenteuse. En abscisses et en ordonnées, les coordonnées angulaires par rapport au centre de l’image. L’unité de l’échelle est le MJy/sr.

TABLE 1: Degré et angle de polarisation pour différentes lignes de visée traversant le nuage de Polaris.

R.A. 2000 hms	DEC 2000 dms	HIP/Tycho	$d$ pc	$[V]$ mag	$P$ %	$\sigma_P$ %	$\theta$ deg.	$\sigma_\theta$ deg.
00 :17 :02.90	84 :57 :18.20	1359	180.2	7.8	0.08	0.02	173.29	6.43
00 :42 :43.40	85 :14 :14.10	3354	135.0	9.6	0.02	0.02	164.83	23.51
00 :55 :16.20	85 :42 :09.30	4320	235.3	8.1	0.02	0.03	36.50	33.71
01 :03 :39.70	84 :36 :24.70	4965	149.9	6.7	0.02	0.03	35.00	35.02
01 :22 :49.60	83 :48 :33.30	6447	142.9	9.6	0.10	0.04	168.00	10.74
01 :53 :47.60	88 :33 :05.80	8846	234.2	7.9	0.04	0.01	169.80	5.05
02 :02 :18.30	85 :44 :30.10	9504	166.4	9.2	0.01	0.04	0.00	130.97
02 :03 :34.80	86 :55 :54.40	9614	178.9	8.1	0.03	0.05	41.40	57.30
02 :16 :45.50	83 :33 :41.00	10623	131.2	6.4	0.02	0.02	83.00	35.26
02 :18 :59.60	85 :44 :10.10	10800	174.2	7.0	0.02	0.01	71.53	14.12
02 :34 :32.60	88 :28 :15.70	11980	154.1	8.1	0.03	0.01	161.50	12.71
02 :36 :52.30	83 :18 :58.20	12175	233.1	9.6	0.05	0.02	3.60	9.67
02 :37 :27.70	83 :50 :04.40	12232	197.2	6.7	0.03	0.04	78.00	42.98
02 :39 :48.00	85 :03 :33.50	12410	114.3	8.6	0.23	0.03	32.00	3.18
02 :46 :34.30	85 :47 :57.40	12952	154.6	7.8	0.03	0.01	84.47	9.35
03 :04 :07.90	87 :01 :33.90	14285	157.2	7.8	0.02	0.01	1.66	23.34
03 :11 :53.10	82 :54 :13.70	14859	196.5	7.1	0.03	0.02	88.00	24.35
03 :47 :08.90	87 :53 :48.20	17673	117.8	8.8	0.04	0.02	1.44	14.86
03 :56 :30.10	85 :15 :11.50	18436	207.5	9.2	0.07	0.02	10.22	7.50
05 :43 :48.80	85 :40 :06.00	27015	144.9	6.7	0.14	0.03	0.74	6.77
06 :12 :27.00	88 :07 :52.20	29457	101.3	10.6	0.10	0.03	35.78	7.51
07 :28 :38.90	88 :38 :13.50	36324	152.2	9.2	0.15	0.01	165.43	2.58
07 :39 :01.50	88 :27 :49.60	37247	211.4	10.0	0.16	0.05	158.50	9.10
08 :14 :08.20	87 :41 :32.40	40335	177.6	9.1	0.10	0.05	169.50	12.43
09 :04 :48.70	88 :19 :20.70	44563	147.5	9.9	0.11	0.03	177.48	7.49
09 :21 :48.50	88 :34 :13.20	45919	125.6	7.1	0.09	0.03	1.60	8.63
09 :35 :14.50	88 :30 :45.70	47044	138.1	9.6	0.07	0.05	11.80	18.75
09 :46 :20.50	89 :34 :09.90	47953	147.7	9.0	0.03	0.04	8.00	33.06
09 :49 :02.00	86 :51 :07.00	48163	153.6	8.3	0.12	0.01	11.50	3.41
10 :09 :32.30	86 :34 :35.20	49768	232.6	7.8	0.07	0.03	39.00	11.35
10 :18 :30.50	85 :49 :43.80	50482	237.5	8.4	0.11	0.04	15.00	9.76
10 :49 :13.50	87 :51 :32.10	52908	200.4	8.8	0.08	0.03	30.00	9.70
11 :21 :47.90	87 :38 :18.50	55483	205.3	7.5	0.09	0.03	70.00	10.74
11 :30 :16.60	88 :44 :59.10	56124	153.4	9.5	0.13	0.05	42.00	11.46
15 :26 :20.40	87 :26 :07.60	75576	130.4	9.1	0.06	0.04	21.00	17.19
17 :42 :50.00	85 :40 :12.30	86695	114.3	7.6	0.12	0.03	44.00	5.97
17 :54 :23.80	86 :33 :31.10	87663	212.8	7.7	0.09	0.03	25.00	9.97
20 :38 :47.40	89 :24 :23.40	101884	242.1	8.9	0.10	0.05	42.00	14.15
20 :57 :22.30	87 :01 :57.60	103435	150.4	7.4	0.09	0.04	169.50	12.28
21 :52 :31.90	84 :49 :25.30	107982	126.3	9.1	0.08	0.03	10.00	10.58

TABLE 1: Degré et angle de polarisation pour différentes lignes de visée traversant le nuage de Polaris.

R.A. 2000	DEC 2000	HIP/Tycho	$d$	$[V]$	$P$	$\sigma_P$	$\theta$	$\sigma_\theta$
hms	dms		pc	mag	%	%	deg.	deg.
22 :03 :13.60	88 :04 :23.00	108862	156.2	7.4	0.07	0.03	4.00	13.07
22 :13 :10.90	86 :13 :16.90	109694	233.6	6.6	0.06	0.02	0.00	10.42
22 :16 :33.80	84 :30 :15.90	109985	222.7	7.6	0.05	0.03	38.50	18.58
22 :51 :03.00	85 :22 :25.40	112833	140.1	5.9	0.06	0.03	14.50	15.52
23 :24 :20.70	86 :25 :04.00	115550	136.1	6.6	0.09	0.02	25.90	7.47
23 :36 :23.80	83 :12 :14.60	116496	216.9	7.5	0.08	0.03	0.60	8.55
23 :42 :18.40	85 :28 :06.70	116945	242.7	7.7	0.06	0.02	30.60	10.11
23 :43 :12.60	88 :20 :31.30	117008	241.0	9.0	0.04	0.04	33.00	28.65
23 :43 :48.10	87 :21 :17.60	117053	141.0	9.5	0.12	0.04	39.80	8.95
23 :46 :13.80	84 :03 :12.50	117233	243.3	9.2	0.69	0.02	166.84	1.02
23 :47 :37.80	85 :04 :42.20	117350	246.9	7.8	0.06	0.01	35.09	6.52
23 :59 :30.70	86 :42 :23.00	118285	193.8	6.8	0.09	0.04	32.00	12.11
00 :59 :47.60	87 :53 :24.70	0	0.0	0.0	0.19	0.08	27.00	11.46
02 :09 :36.40	87 :40 :52.30	41951	0.0	10.3	0.77	0.06	76.85	2.05
02 :34 :43.30	87 :25 :35.40	12003	770.0	9.9	0.07	0.08	180.00	32.36
01 :24 :29.20	88 :15 :45.60	0	0.0	8.9	0.08	0.02	68.56	6.12
03 :35 :33.20	86 :23 :59.10	16754	147.0	9.0	0.19	0.04	69.00	5.28
00 :59 :05.40	87 :52 :53.10	0	0.0	0.0	0.09	0.06	81.00	20.46
00 :39 :49.70	87 :47 :48.40	3133	317.0	9.0	0.03	0.05	38.00	57.30
02 :48 :58.80	88 :39 :39.40	4628201	0.0	8.8	0.04	0.03	73.50	26.60
02 :33 :18.50	88 :10 :25.50	4628671	0.0	10.7	0.37	0.06	159.00	4.81
05 :23 :20.00	88 :19 :01.80	4629801	0.0	10.7	0.46	0.10	177.50	6.19
03 :31 :39.20	88 :15 :34.30	4628941	0.0	10.5	0.23	0.07	13.00	8.60
03 :45 :14.40	86 :34 :44.30	4624631	0.0	10.3	0.38	0.06	71.50	4.27
01 :48 :19.20	87 :53 :28.50	8397	123.0	9.7	0.08	0.02	83.63	8.08

## 2 Traitement pour obtenir l'orientation des filaments

Notre objectif ici est de fabriquer un masque pour l'image, qui vaudra 1 sur les filaments sélectionnés, et 0 partout ailleurs. Cela nous permettra, par la suite, de calculer facilement l'orientation de branches de filaments et leur distance aux points où l'on connaît la polarisation de la lumière. Dans un premier temps, nous filtrons l'image afin d'éliminer les sources ponctuelles n'appartenant pas au nuage (rappelons qu'il n'y a pas de formation d'étoile dans Polaris) et de faire ressortir les filaments. Nous fabriquons ensuite le masque, en filtrant pour conserver les filaments les plus importants.

## 2.1 Analyse en curvelets

La transformée en ondelettes permet une décomposition en espace et en échelle d'un signal. Découverte il y a plus de 25 ans, elle s'applique maintenant à de nombreux domaines, et en particulier aux images astrophysiques, qui présentent généralement des structures à plusieurs échelles. Cependant, la transformée en ondelettes met principalement en valeur les éléments isotropes : d'autres transformées ont donc été développées par la suite pour décrire les éléments anisotropes [10]. En deux dimensions, les ridgelets décrivent des schémas rectilignes à l'échelle de l'image ; les curvelets, transformées en ridgelets locales sur différentes parties de l'image, mettent l'accent sur les courbes à toute échelle. Nous comprenons alors qu'une transformée en curvelets nous intéresse particulièrement, puisqu'elle fera ressortir les filaments, éléments anisotropes.

### 2.1.1 Transformées en ondelettes, ridgelets et curvelets

Nous nous intéressons maintenant de façon plus quantitative à ces transformées. Commençons par décrire la transformée en ondelettes continue à 1D. C'est une généralisation de la transformée de Fourier : l'atome de base n'est plus une sinusoïde, dont on fait varier la fréquence, mais une fonction vérifiant certaines propriétés, l'ondelette "mère", qui est dilatée et translatée. Pour avoir un aperçu du type de fonction possible, une ondelette mère standard est présentée en figure 2. Les coefficients obtenus ont donc pour paramètres ceux de dilatation et de translation.

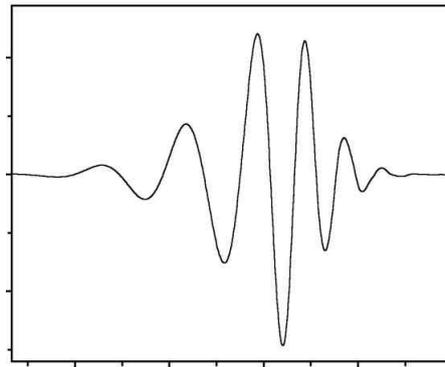


FIGURE 2 – Ondelette mère standard, appelée Daubechies 20.

Pour la suite, nous avons pour référence [11]. On se place dans  $L^2(\mathbb{R}) = \{f : \|f\|_2 < \infty\}$ . La fonction d'ondelette "mère"  $\psi$  doit vérifier la condition d'admissibilité, avec  $\hat{\psi}$  la transformée de Fourier de  $\psi$ ,

$$C_\psi = \int_0^\infty |\hat{\psi}(k)|^2 \frac{dk}{|k|} < \infty \quad (1)$$

La famille d'ondelettes translatées et dilatées s'écrit alors, avec  $a > 0$  et  $b \in \mathbb{R}$ ,

$$\psi_{a,b}(x) = \frac{1}{\sqrt{a}} \psi\left(\frac{x-b}{a}\right) \quad (2)$$

Le paramètre de dilatation est bien sûr  $a$ , mesurant la taille du support de l'ondelette, et le paramètre de translation est  $b$ . La transformée en ondelettes d'une fonction  $f$  est alors simplement

la projection de cette fonction sur la famille d'ondelettes générées (l'étoile signifie le conjugué, dans le cas de valeurs complexes) :

$$\tilde{f}(a, b) = \int_{-\infty}^{\infty} f(x) \psi_{a,b}^*(x) dx \quad (3)$$

La synthèse s'effectue en intégrant le produit de chaque coefficient avec son ondelette associée sur l'espace des paramètres :

$$f(x) = \frac{1}{C_\psi} \int_0^\infty \int_{-\infty}^\infty \tilde{f}(a, b) \psi_{a,b}(x) \frac{da db}{a^2} \quad (4)$$

La discrétisation de cette transformée est faite usuellement de façon logarithmique pour le paramètre d'échelle et linéairement pour le paramètre de translation. On pose  $a_j = a_0^{-j}$ , où  $j \in \mathbb{Z}$  est l'indice d'échelle, et  $b$  est remplacé par  $x_{ji} = ib_0 a_0^{-j}$ . Les ondelettes discrètes sont alors :

$$\psi_{ji}(x) = a_j^{-1/2} \psi\left(\frac{x - x_{ji}}{a_j}\right) \quad (5)$$

Et les coefficients sont alors naturellement :

$$\tilde{f}(j, i) = \int_{-\infty}^{\infty} f(x) \psi_{ji}^*(x) dx \quad (6)$$

On utilise en fait  $a_0 = 2$  et  $b_0 = 1$ , qui font des ondelettes  $\psi_{ji}$  une base orthonormale de  $L^2(\mathbb{R})$  et qui autorisent ainsi la transformée discrète inverse [11]. Les algorithmes de transformée rapide utilisent le concept mathématique appelé analyse multirésolution. Celui employé dans le code que nous avons utilisé est appelé algorithme "à trous", et nous le présentons qualitativement. Une fonction  $f$  peut être approximée à différentes échelles  $j$  en  $f_j$ . Notons que nous prenons maintenant  $j \Leftrightarrow -j$  : plus  $j$  augmente, plus l'échelle est grossière. Notre signal discret est une approximation particulière d'un signal continu, du simple fait de sa discrétisation. La fonction  $f_{j+1}$  est alors une version deux fois plus grossière de la fonction  $f_j$  (on fonctionne avec des échelles de facteur  $a_0 = 2$ ). La différence d'information entre  $f_j$  et  $f_{j+1}$  peut être entièrement représentée par des ondelettes orthogonales : cette information est un signal, exclusivement à l'échelle  $j$ . Une analyse en ondelettes orthogonales peut donc être vue comme une décomposition du signal à des échelles de plus en plus grossières, en enregistrant au fur et à mesure les détails perdus sous forme de coefficients d'ondelettes. Lors de l'itération d'une échelle à l'échelle suivante, le calcul des coefficients à sauvegarder se fait par l'application d'un filtre au signal sous-échantillonné. L'idée de l'algorithme "à trous" est de ne pas faire de sous-échantillonnage successif, mais plutôt d'introduire plutôt des zéros (d'où le nom de l'algorithme) entre les coefficients des filtres discrets (en fait  $2^j - 1$  zéros exactement) avant chaque nouvelle application au signal : on "espace" le filtre, pour prendre en compte le changement d'échelle. Le passage à 2D peut se faire, pour des ondelettes isotropes, simplement en posant  $\psi(\vec{x}) = \psi_{1D}(|x|)$ .

Nous l'avons dit, cette transformée n'est pas efficace quand il s'agit de mettre en valeur des éléments linéaires dans une image. L'idéal pour faire apparaître un élément linéaire dans un signal bruité, serait d'intégrer le signal le long de lignes parallèles à cet élément : on aurait alors le maximum de chance de détecter l'élément via un saut lors du parcours du signal intégré. La transformée en ridgelets s'appuie sur cette observation. Présentons-en sa forme continue à deux dimensions. Prenons une fonction  $\psi$  vérifiant la condition d'admissibilité :

$$C_\psi = \int_0^\infty |\hat{\psi}(k)|^2 \frac{dk}{|k|^2} < \infty \quad (7)$$

On suppose de plus une normalisation telle que  $C_\psi = 1$ . Alors, avec  $a > 0$ ,  $b \in \mathbb{R}$ ,  $\theta \in [0, 2\pi[$  et  $\vec{u}_R = (\cos \theta, \sin \theta)$ , les ridgelets s'écrivent :

$$\psi_{a,b,\theta}(\vec{x}) = \frac{1}{\sqrt{a}} \psi\left(\frac{\vec{x} \cdot \vec{u}_R - b}{a}\right) \quad (8)$$

Une ridgelet est donc constante le long des droites passant par l'origine, et est une ondelette à 1D transversalement à ces droites. La transformée de Radon décrit ces dernières droites, avec  $\delta$  la distribution de Dirac :

$$R_a[f](\theta, t) = \int f(x_1, x_2) \delta(\vec{x} \cdot \vec{u}_R - t) dx_1 dx_2 \quad (9)$$

Une transformée en ridgelets à 2D est donc une transformée en ondelettes à 1D appliquée aux tranches obtenue par la transformée de Radon. Les coefficients des ridgelets sont ainsi :

$$R(a, b, \theta) = \int f(\vec{x}) \psi_{a,b,\theta}^*(x) d\vec{x} \quad (10)$$

Et la formule de reconstruction est :

$$f(\vec{x}) = \int_0^{2\pi} \int_0^\infty \int_{-\infty}^\infty R(a, b, \theta) \psi_{a,b,\theta}(\vec{x}) \frac{da}{a^3} db \frac{d\theta}{4\pi} \quad (11)$$

L'algorithme de cette transformée discrétisée utilise tout d'abord une transformée de Radon. Celle-ci est calculée en effectuant une transformée de Fourier inverse à 1D sur des points pris le long des droites passant par l'origine dans le plan de Fourier 2D [10]. Une transformée en ondelettes comme décrite plus haut est appliquée ensuite à ces droites.

La transformée en ridgelets est optimale pour repérer des lignes de la taille de l'image, typiquement des artefacts prenant la forme de bande. L'idée de la transformée en curvelets est d'appliquer le précédent traitement à des sous-parties de l'image, et ce pour différentes échelles de sous-parties. On retrouve ici l'esprit des ondelettes avec la décomposition de l'image en plusieurs échelles. La transformée en curvelets s'appuie en effet pour commencer sur l'application d'un algorithme d'ondelette "à trous" pour obtenir des images comprenant les détails correspondant à différentes échelles. Ces images sont découpées en sous-images de taille correspondant à l'échelle représentée. Une transformée en ridgelets est finalement appliquée localement en cette sous-image. L'algorithme de cette transformée est schématisé sur la figure 3, tirée de [10]. Les coefficients des curvelets sont donc des coefficients de ridgelets, avec comme paramètres supplémentaires l'indice d'échelle et l'indice de positionnement dans l'image.

### 2.1.2 Morphological Component Analysis (MCA)

L'utilisation de ces transformées pour améliorer la qualité d'une image, c'est à dire pour augmenter le contraste de ce qui nous intéresse, implique habituellement la suppression de certains coefficients dans l'espace de la transformée. Il n'est cependant pas toujours évident de sélectionner les coefficients correspondant à ce que l'on veut enlever, et il est possible d'enlever par la même occasion de l'information qui nous intéresse : ce sont les problèmes majeurs de ces transformées. La Morphological Component Analysis (MCA) a pour objectif la tâche ardue de séparer une image en différentes composantes ayant des morphologies, des

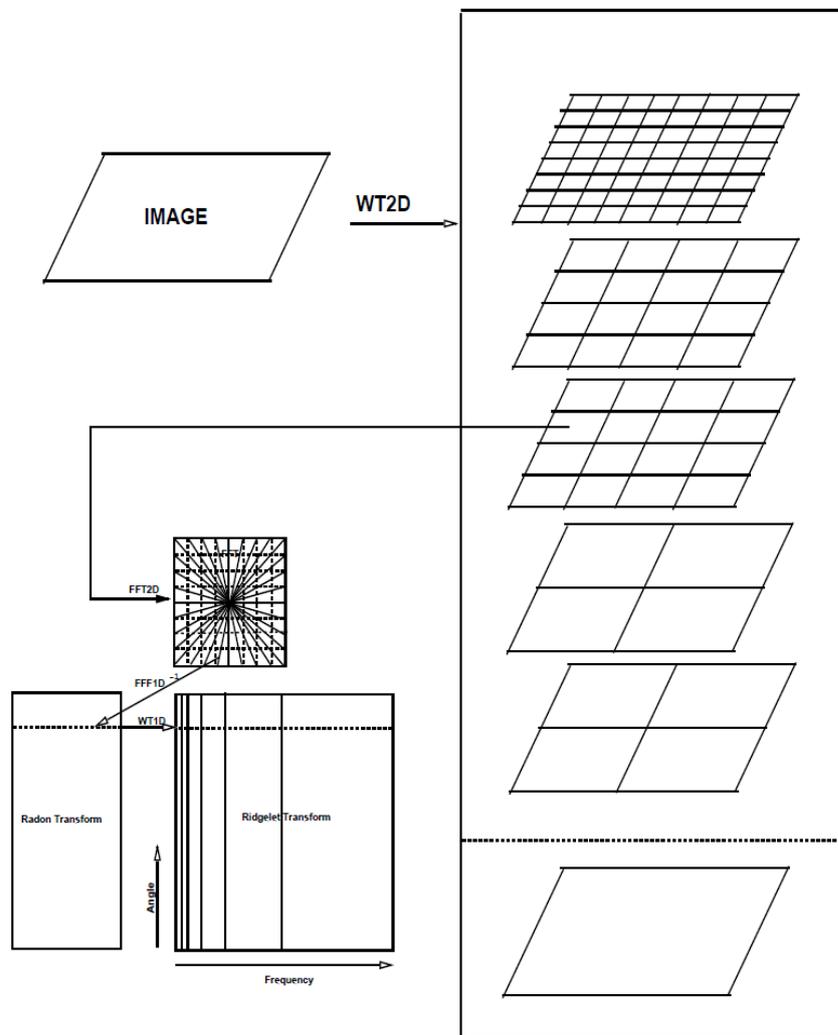


FIGURE 3 – Schéma de l’algorithme de transformée en curvelets. WT signifie transformée en ondelettes. L’image est décomposée en plusieurs échelles avec une transformée en ondelettes, puis chaque sous-image se voit appliquer une transformée en ridgelets : passage dans le plan de Fourier 2D, tracé de lignes et transformée de Fourier inverse le long de ces lignes, et enfin transformée en ondelette le long des lignes.

géométries différentes. L'idée de base est que chaque image est une superposition linéaire de différents signaux, qui sont chacun plus ou moins proches de tel ou tel atome de base d'une transformée. On constitue donc un dictionnaire de transformées susceptibles de mettre en valeur les différentes composantes de l'image que l'on veut séparer. L'application de l'algorithme nous donne une représentation minimale sur les coefficients du dictionnaire les plus appropriés, à partir desquels on peut reconstruire les composantes qui, sommées, donnent l'image originale. Dans notre cas, par exemple, l'idéal serait de séparer notre image en une image contenant les sources ponctuelles, qui n'appartiennent pas à Polaris, car on sait le nuage sans formation stellaire, et en une autre image contenant les filaments.

Voyons maintenant plus en détail comment cet algorithme permet ceci [12]. Supposons que l'on veut séparer une image  $F$  en  $K$  composantes  $(F_1, \dots, F_K)$ . Soit  $\Phi$  le dictionnaire des transformées choisies  $(\Phi_1, \dots, \Phi_K)$ . On note  $\alpha_i = \Phi_i F_i$  le vecteur des coefficients de la composante  $F_i$  pour la transformée  $\Phi_i$ . Si  $F$  contient  $n$  pixels, obtenir les  $K \cdot n$  pixels inconnus à partir de ces  $n$  données présente clairement plus d'inconnues que d'équations. L'équation  $F = \Phi^{-1}\alpha$  est indéterminée, et de multiples solutions sont possibles : on peut par exemple représenter entièrement l'image dans la transformée  $\Phi_1$ , prenant ainsi  $F_1 = F$  et pour tout  $i = 2 \dots K$ ,  $F_i = 0$ , mais ce n'est pas ce que l'on souhaite. On suppose donc que les transformées choisies sont indépendantes. On entend par là que la représentation  $\alpha_i$  de  $F_i$  dans  $\Phi_i$  est assez étendue, ou du moins plus étendue que pour les autres transformées  $\Phi_{l \neq i}$ . Cette étendue est à rapprocher de la largeur spectrale dans le cas de la transformée de Fourier d'un signal périodique. On résout alors l'indétermination en ajoutant la condition d'optimisation suivante :

$$\min_{(F_1, \dots, F_K)} \sum_{k=1}^K (\|\Phi_k F_k\|_p)^p, \text{ tel que } \|F - \sum_{k=1}^K F_k\|_2 \leq \sigma \quad (12)$$

La norme  $(\|\alpha\|_p)^p = \sum_i |\alpha(i)|^p$ , avec  $0 \leq p \leq 1$ , quantifie l'étendue. On se placera souvent dans le cas  $p = 0$ , pour lequel  $\|\alpha\|_0$  est simplement le nombre de composantes non-nulles du vecteur.  $\sigma$  quantifie l'éloignement maximal souhaité de la superposition des composantes séparées à l'image originale (ie le niveau de bruit). Avec cette condition on limite le nombre de coefficients représentant chaque  $F_k$ , tout en essayant de se rapprocher le plus possible de l'image originale, ce qui limite la présence d'éléments géométriques dans des composantes auxquels il ne sont pas destinés. L'algorithme de séparation des composantes prend alors la forme d'une boucle itérative qui sélectionne les coefficients, décrivant l'image à une échelle brute tout d'abord puis de plus en plus fine, des coefficients les plus importants aux moins élevés. Un seuil est fixé à l'initialisation pour les coefficients des transformées de l'image. Tous les coefficients au dessus de ce seuil sont sauvegardés, puis on retire à l'image originale les images correspondant à ces coefficients. On abaisse le seuil, puis on réitère ce procédé, pour chaque transformée, à l'image résiduelle, à laquelle on ajoute l'image séparée déjà sauvegardée pour cette transformée. Ceci est fait jusqu'à atteindre le nombre d'itérations maximal, ou le seuil minimal. Pour une présentation plus quantitative, se référer à l'annexe A où l'algorithme est détaillé.

Résumons les choix à faire pour pouvoir utiliser MCA :

- Le choix du dictionnaire se fait selon ce qu'on l'on souhaite mettre en valeur. Les ondelettes représentent bien les éléments isotropes. Les curvelets correspondent aux courbes de l'image. Les ridgelets sont parfaites pour des lignes de taille proche de la dimension de l'image. On peut imaginer utiliser d'autres transformées pour d'autres types de structure.

- Il faut choisir pour chaque transformée les paramètres qui lui correspondent. Pour des ondelettes ou des curvelets, il convient de choisir la forme de l’ondelette mère, et le nombre d’échelles  $J$ . Il correspond à la taille minimale des sous-images qui vont être inspectées,  $2^{-J}$ . Il faut donc veiller à ce que  $2^{J+1} \leq n_{1D}$ , avec  $n_{1D}$  la taille linéaire en pixels de l’image (pas le nombre total de pixels).
- Le paramètre  $p$  est pris égal à 0 pour la suite.
- Le nombre d’itérations  $N_{iter}$  va influencer la finesse des composantes séparées, mais également le temps de calcul.
- Il faut sélectionner une stratégie de décroissance du seuil  $\lambda$ , linéaire ou exponentielle.
- Le seuil  $\sigma$  en dessous duquel on estime que la différence entre la somme des composantes séparées et l’image originale est négligeable, typiquement le niveau de bruit.

### 2.1.3 Application

Nous avons commencé par utiliser MCALab, la version MATLAB du code MCA [12]. Les résultats obtenus n’étaient pas satisfaisants : il n’était pas possible d’utiliser d’algorithme à trous pour les ondelettes, et les composantes obtenues ne convergeaient pas, mais oscillaient entre deux images, avec des paramètres pourtant très proches de ceux cités ci-dessous. Ceci est présenté en annexe B. Un compte-rendu a été envoyé à Jalal Fadili, auteur de [12], qui nous a redirigé vers la version binaire du code.

Nous avons donc finalement utilisé la version binaire du code MCA, une version plus récente que celle décrite dans [12]. L’image traitée est l’image de Polaris à  $100\mu\text{m}$  présentée en partie 1.1, rappelée en figure 4. La décomposition sur des ondelettes à trous et des curvelets, avec une décroissance exponentielle du seuil, un seuil  $\sigma$  par défaut, et un nombre d’échelles  $J$  de 6 et 7 ont donné des résultats très satisfaisants, présentés respectivement dans les figures 5 et 6. La forme de l’ondelette mère joue peu sur les résultats.

Les ondelettes semblent être mieux représentées pour un nombre d’échelles de 7, offrant plus de détails pour les sources ponctuelles, tandis que 6 échelles réussissent mieux aux curvelets, montrant mieux les filaments. C’est donc la composante en curvelets pour un nombre d’échelles de 6 que nous conservons pour la suite de notre analyse. La puissance de l’image originale récupérée dans cette composante est de 11%, et donc de 89% pour les ondelettes à trous (la puissance est simplement définie comme la somme de la valeur de chaque pixel, toutes ces valeurs étant positives). La puissance récupérée dans le cas de 7 échelles est de 16% pour les curvelets et 84% pour les ondelettes : il y a un léger transfert de puissance.

## 2.2 Détection des filaments

Nous disposons désormais d’une image traitée sur laquelle restent principalement les filaments. Nous pouvons alors mettre en œuvre le logiciel DisPerSE [13] afin de fabriquer un réseau de segments correspondant aux filaments les plus importants : ce réseau est appelé le squelette. Une fois ce squelette calculé, nous pourrions étudier quantitativement les propriétés des filaments.

### 2.2.1 DisPerSE, principe et fonctionnement

DisPerSE est un logiciel qui s’appuie sur la récente théorie de Morse discrète [14] pour analyser les réseaux de filaments à partir de champs de densité discrets ou échantillonnés. La

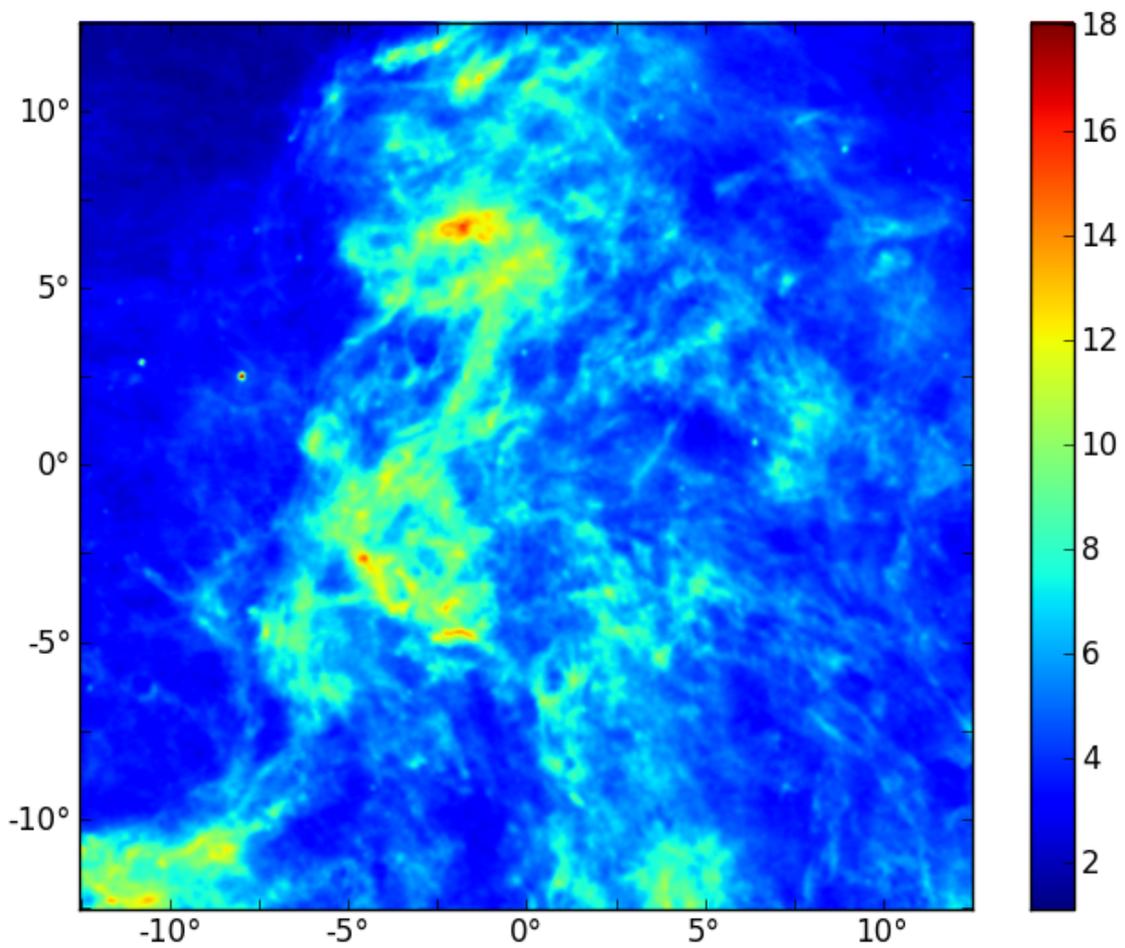


FIGURE 4 – Image de Polaris (région 430) provenant des données IRIS, somme des trois passages du satellite, à  $100\mu\text{m}$ . En abscisses et en ordonnées, les coordonnées angulaires par rapport au centre de l'image. L'unité de l'échelle est le MJy/sr.

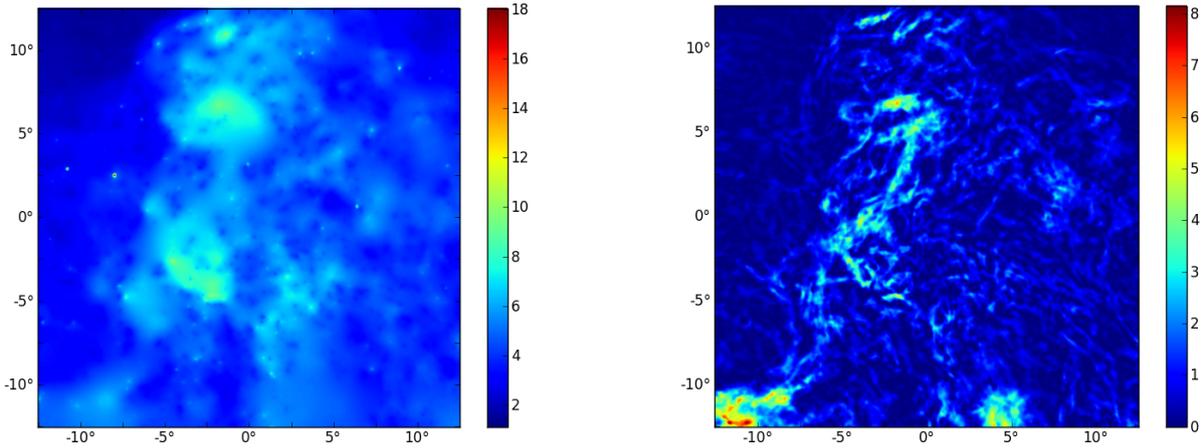


FIGURE 5 – Décomposition de l’image de Polaris en ondelettes à trous (à gauche) et en curvelets (à droite) pour un nombre d’échelles de 6, une décroissance exponentielle du seuil et 50 itérations. En abscisses et en ordonnées, les coordonnées angulaires par rapport au centre de l’image. L’unité de l’échelle est le MJy/sr.

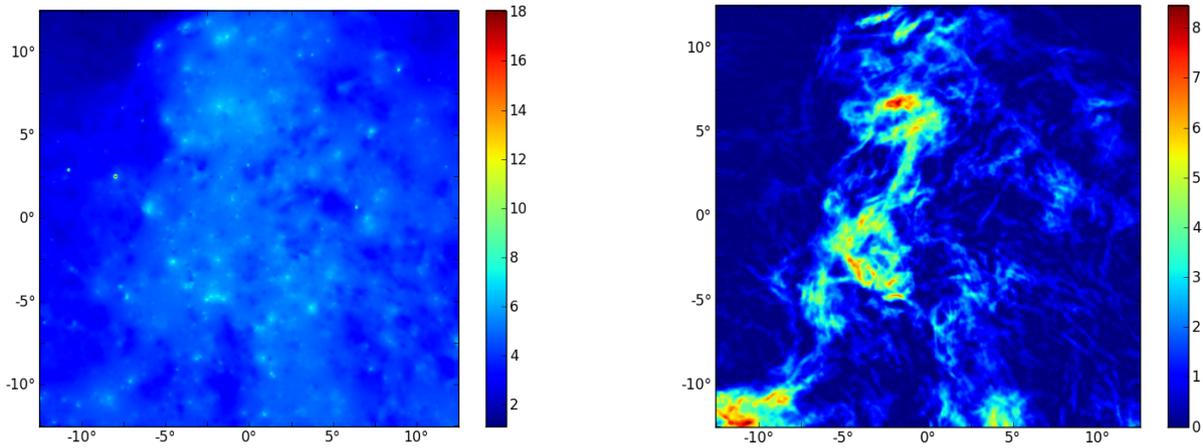


FIGURE 6 – Décomposition de l’image de Polaris en ondelettes à trous (à gauche) et en curvelets (à droite) pour un nombre d’échelles de 7, une décroissance exponentielle du seuil et 20 itérations. En abscisses et en ordonnées, les coordonnées angulaires par rapport au centre de l’image. L’unité de l’échelle est le MJy/sr.

théorie de la persistance permet de prendre en compte le bruit, et de faire apparaître plus ou moins de filaments selon la précision avec laquelle le champ étudié est connu. L’ensemble de la théorie sous-jacente est décrite dans [13]; nous donnons ici un aperçu de la logique du logiciel en résumant le cas continu à deux dimensions. Si on se représente le champ de densité comme un paysage de montagne, les filaments vont correspondre aux crêtes, reliant les sommets via les cols. Le squelette est alors l’ensemble des lignes reliant les maxima, en passant par les points selle. La théorie de Morse continue permet d’exploiter ce constat rigoureusement.

Soit donc un champ scalaire  $f$  "suffisamment régulier" (par exemple de classe  $C^2$ ) défini sur  $\mathbb{R}^2$ . Un point  $P$  de coordonnées  $\vec{p} \in \mathbb{R}^2$  est dit critique si  $\nabla f(\vec{p}) = 0$ . Il est dit d’ordre  $k$  si la matrice hessienne  $\partial^2 f / \partial x_i \partial x_j$  possède exactement  $k$  valeurs propres négatives en  $\vec{p}$ . Les

maxima, les points selle et les minima sont des des points critiques respectivement d'ordre 2, 1 et 0. On suppose qu'il n'y a pas de points critiques dégénérés, c'est à dire que le hessien en ces points ne possède pas de valeur propre nulle. Une fonction vérifiant cette condition est appelée une fonction de Morse. On appelle ligne intégrale, ou ligne de champ, une ligne qui suit la direction du gradient du champ (en dehors des points critiques). On peut montrer que ces lignes intégrales partent toujours et arrivent toujours en des points critiques. Dans le cas d'une fonction de Morse, deux lignes intégrales passant par deux mêmes points distincts ne peuvent se croiser, à moins d'être identiques. L'ensemble des lignes intégrales couvre entièrement  $\mathbb{R}^2$  et chaque point non critique n'appartient qu'à une seule ligne intégrale. On peut donc classer les points de  $\mathbb{R}^2$  selon l'origine, ou la destination, de leur ligne de champ, définissant ainsi les  $n$ -variétés ascendantes et descendantes. Soit  $P$  un point critique d'ordre  $k$ . La  $k$ -variété descendante est un ensemble de dimension  $k$  : c'est l'ensemble des points par où passent des lignes intégrales ayant pour destination  $P$ . De façon analogue, la  $(2 - k)$ -variété ascendante est un ensemble de dimension  $(2 - k)$  : c'est l'ensemble des points par où passent des lignes intégrales ayant pour origine  $P$ . Les  $n$ -variétés définies ainsi sont des points critiques, des lignes, ou des surfaces. Par exemple, si  $P$  est un maximum local du champ, il est au "sommet" d'une 2-variété descendante, et il est sa propre 0-variété ascendante. L'ensemble des variétés ascendantes et descendantes s'appelle le complexe de Morse. Notre réseau de filaments est alors tout trouvé : c'est l'ensemble des 1-variétés ascendantes, chacune reliant un point selle à un maximum. `DisPerSE` utilise la théorie de Morse discrète qui permet de généraliser à des fonctions échantillonnées.

Nous expliquons maintenant comment est utilisée la théorie de la persistance. Le concept est le suivant : pour étudier la topologie d'un champ scalaire, il est intéressant de regarder comment varie la topologie de l'ensemble des points avec des valeurs au dessus d'un certain seuil, appelé ensemble d'excursion, quand ce seuil varie. Une image simple est d'inonder notre paysage montagneux du début de cette section, et de faire descendre petit à petit le niveau d'eau. On peut alors voir apparaître tout d'abord les sommets, comme points isolés, puis les voir s'élargir, puis se connecter entre eux petit à petit. En fait, on remarque que lorsque le seuil traverse la valeur d'un point critique, un élément de l'ensemble d'excursion apparaît ou disparaît, se connecte ou se déconnecte d'autres éléments : on peut le dire créateur ou destructeur. On peut en fait relier les points critiques par paires : celui créant un nouvel élément, et celui le faisant disparaître en le connectant aux autres, par exemple. Autrement dit, une paire de points critiques correspond à la création et la destruction d'une caractéristique topologique de l'ensemble d'excursion. La persistance d'une paire de points critiques est alors la différence de valeur de la fonction étudiée entre ces deux points. Pour reprendre notre exemple imagé, si le niveau de l'océan recouvrant le paysage descend à vitesse constante, la persistance d'une paire est le "temps de vie" de l'élément topologique qu'elle représente ; le temps entre l'apparition d'une île et sa connexion à une autre île. On peut le voir aussi comme la résistance de cet élément topologique à l'ajout de bruit. Ajoutons que, par convention, un point critique connectant deux éléments de l'ensemble d'excursion est appairé au point critique ayant "apparu" le plus récemment, c'est à dire ayant la valeur la plus faible. Toujours avec notre exemple, quand deux îles se rejoignent en un point, le point de jonction est appairé au sommet de l'île la plus basse. La figure 7, issue de [13], illustre simplement ce que nous venons de décrire, à une dimension. Si une paire de points critiques a une persistance inférieure à un seuil voulu, elle est supprimée, et n'est plus prise en compte pour le calcul du squelette. On évite ainsi la prolifération de petits filaments, recouvrant l'intégralité de la carte, correspondant au bruit. Ceci sera illustré en figure 10.

Un lissage du squelette est proposé avec `DisPerSE`. Il est présenté dans [15]. Son principe

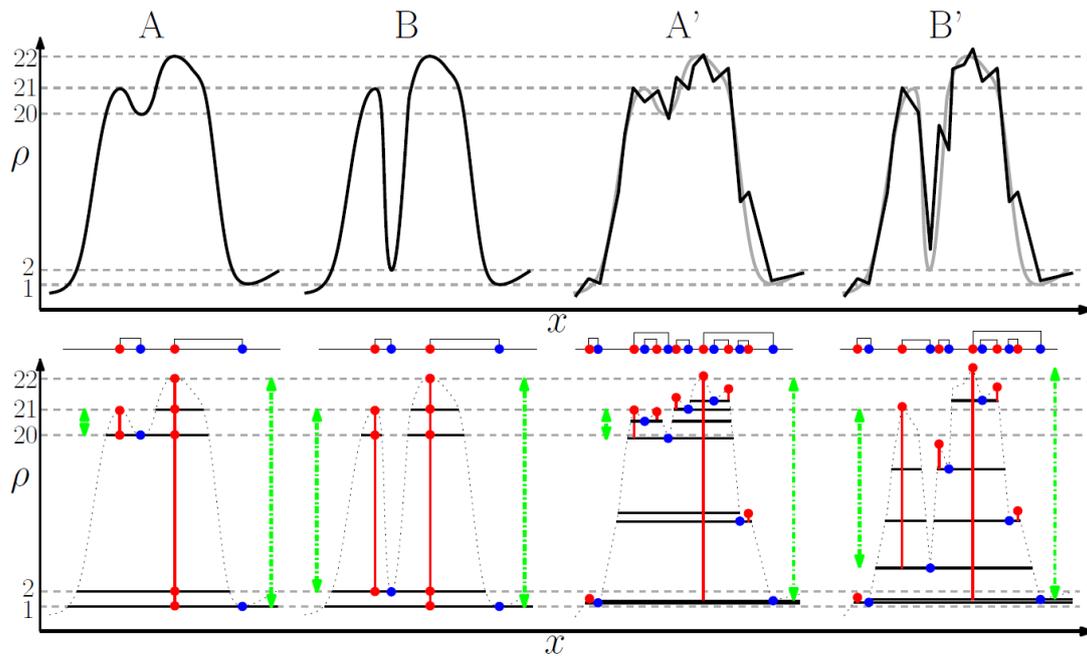


FIGURE 7 – Illustration de la persistance sur un champ 1D.  $\rho$  est la valeur de la densité. La partie haute de la figure illustre différents champs de densité A et B continus, puis les champs correspondant A' et B' discrétisés et avec du bruit. Dans la partie basse de la figure sont représentés ces champs de densité avec les points critiques, en rouge et bleu, et les niveaux de persistance via les flèches vertes en pointillé. Les points rouges sont les points "créateurs", et les points bleus les points "destructeurs", en terme de topologie. Dans ce cas 1D, les points rouges sont des maxima et les points bleus des minima. Les paires de points critiques sont reliés par des lignes noires dans la partie centrale de la figure. La paire de points critiques de gauche, reliant le maximum local de gauche au minimum local central, est bien plus "résistante" au bruit pour le champ B que pour le champ A. Fixer un seuil de persistance de 0.9 (à comparer à la dynamique de  $\rho$ , comprise entre 1 et 22) permet d'effacer les paires de points critiques apparaissant avec le bruit pour les champs A' et B' sans perdre les "vraies" paires, celles présentes sans bruit.

est simple, et il ne comporte qu'un paramètre  $s$  : les points critiques sont fixés, et on lisse le long de chaque filament selon la méthode suivante. Les filaments sont représentés par des segments, d'une taille approximative de un pixel, dont les extrémités, sont, avant lissage, à l'angle d'un pixel, ou au centre d'un de ses côtés. Le lissage fait une moyenne pondérée de la position de l'extrémité d'un segment et de celles de ses deux voisins. Le procédé est répété  $s$  fois afin d'obtenir un lissage sur  $s$  segments. En pratique : un filament est composé de  $N + 2$  segments, mais les deux segments des bouts du filament ont chacun une extrémité fixée, puisque ces extrémités sont des points critiques.  $N$  segments sont donc totalement libres, soit  $N + 1$  extrémités. Notons  $x_j^i$  la  $j$ -ième coordonnée de la  $i$ -ième extrémité, avec  $i$  entre 0 et  $N$ , numérotant dans l'ordre croissant d'un bout à l'autre du filament. Alors les nouvelles coordonnées lissées sur un segment, notées  $y_j^i$ , s'obtiennent avec :

$$y_k^i = \sum_{j=0}^N A^{ij} x_k^j \quad (13)$$

avec :

$$A^{ij} = \begin{cases} 3/4 & \text{si } i = j = 0 \text{ ou } i = j = N, \\ 1/2 & \text{si } i = j \text{ et } i \neq 0 \text{ et } i \neq N, \\ 1/4 & \text{si } i = j + 1 \text{ ou } i = j - 1, \\ 0 & \text{sinon.} \end{cases} \quad (14)$$

Il suffit donc d'appliquer  $s$  fois l'équation 13 pour obtenir un lissage sur  $s$  segments. La résolution obtenue alors pour le squelette est sub-pixel.

Résumons finalement ce que `DisPerSE` accomplit pour nous. Nous lui donnons en entrée une image au format FITS, un seuil de persistance  $p$ , et un nombre de segments  $s$  sur lequel lisser. Nous obtenons en sortie une table au format ASCII contenant les coordonnées en pixels de l'ensemble des extrémités des segments représentant les filaments sélectionnés. Nous pouvons, à partir de cette table, facilement dessiner le squelette sur notre image originale, ou encore calculer la statistique d'orientation des segments. Dans la prochaine section, nous observons les résultats obtenus avec `DisPerSE` sur notre image, et l'influence du choix des paramètres.

### 2.2.2 Calcul du squelette

Sur toutes les images qui suivent, le squelette est représenté en rouge, par dessus l'image originale, non traitée, qui était présentée par exemple en figure 4. La figure 8 représente un squelette "typique", calculé à partir de l'image traitée (ie la composante en curvelets pour un nombre d'échelle de 6, présentée à droite de la figure 5), avec un seuil de persistance  $p = 0.5$  et un lissage sur  $s = 10$  segments. Précisons que le choix du seuil de persistance se fait en "tâtonnant" : le logiciel propose un mode semi-interactif qui permet de voir combien de points critiques sont écartés pour un seuil de persistance donné. Cela donne une idée de la filtration effectuée sur le squelette. L'affichage permet ensuite d'affiner son choix, dont les critères restent très subjectifs. En effet, on cherche à ce que les filaments identifiés correspondent à ce que l'on voit sur l'image. Rappelons cependant que la valeur du seuil de persistance correspond à une intensité bien réelle, et qu'elle doit donc bien entendu être prise supérieure au niveau de bruit. Notons que le bruit sur l'image originale est  $\sigma = 0,06 \pm 0.02$  MJy/sr [7], et que nos seuils de persistance choisis par la suite sont bien au dessus de ce niveau.

La figure 9 illustre l'influence du seuil de persistance. Le squelette en haut à gauche, calculé sans appliquer de seuil de persistance, est clairement noyé par le bruit. Cela montre l'importance

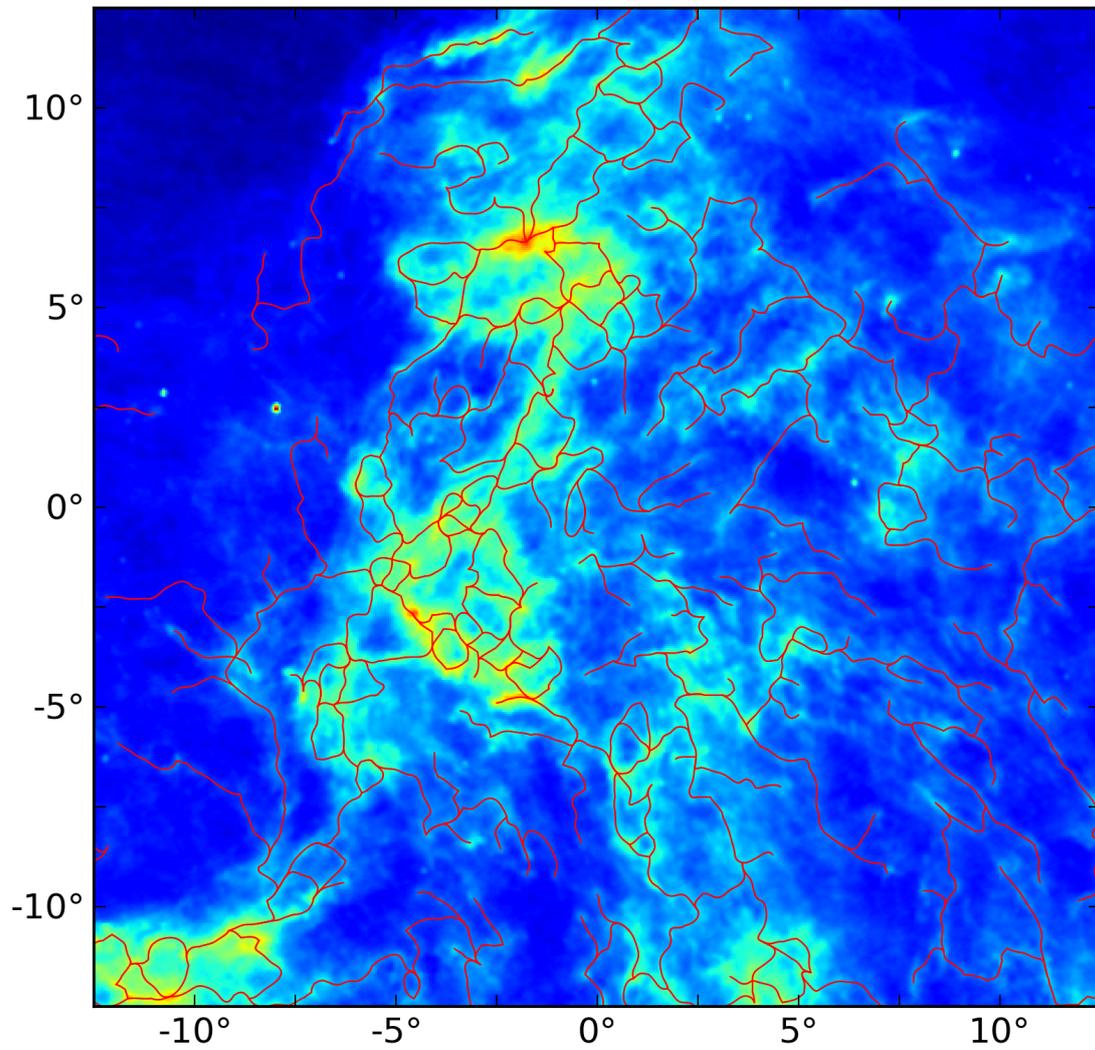


FIGURE 8 – Squelette calculé à partir de la composante en curvelets, un seuil de persistance  $p = 0.5$ , un lissage sur  $s = 10$  segments, tracé en rouge par dessus l'image originale.

de l'utilisation de la théorie de la persistance pour l'identification de ces squelettes. On observe bien un squelette de plus en plus restreint au fur et à mesure que le seuil de persistance augmente. Un seuil de 0.5 semble relativement idéal, puisque le squelette correspondant recouvre la plupart des filaments visibles, en traçant peu de filaments à partir du bruit. On illustre maintenant

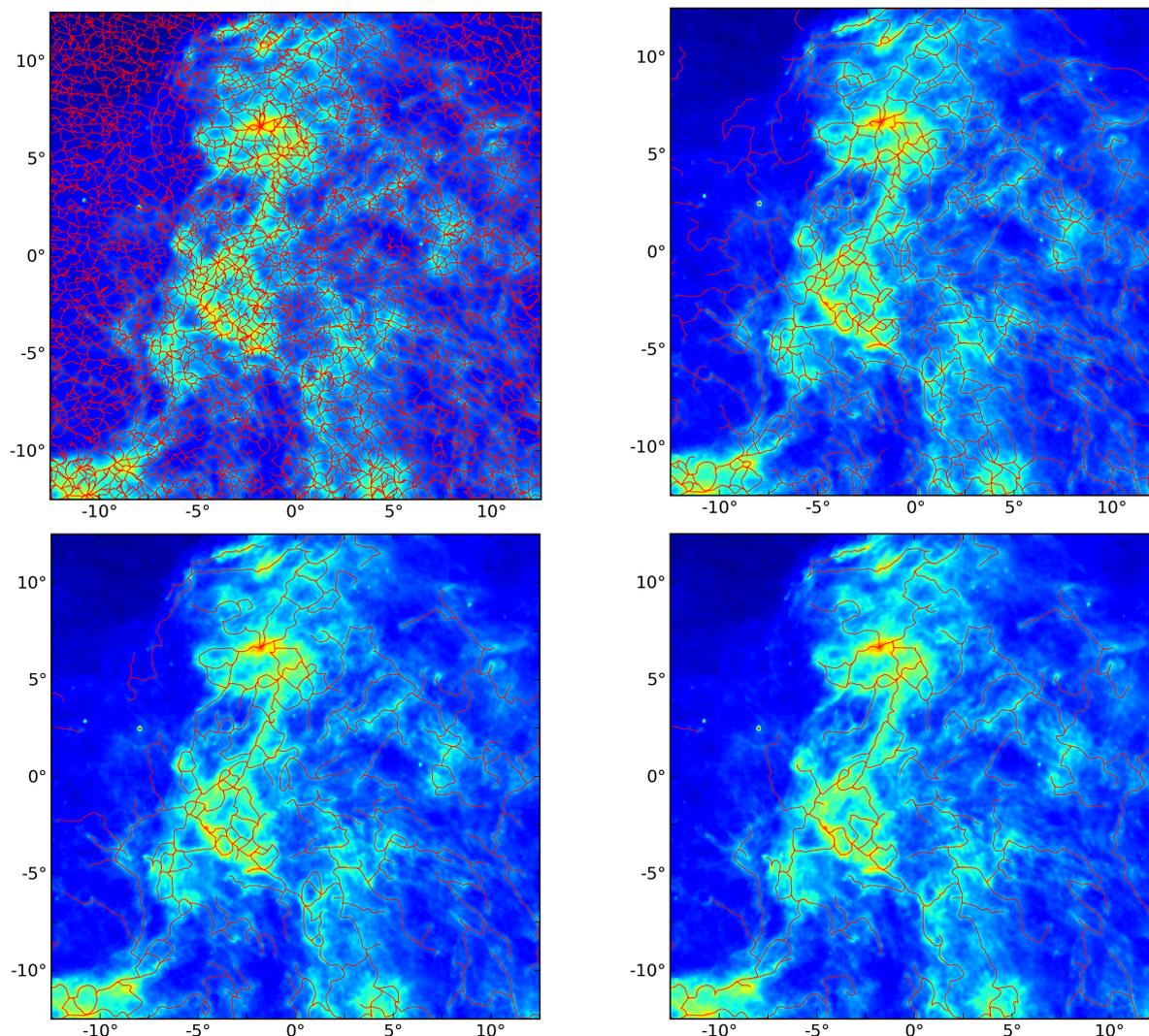


FIGURE 9 – Squelette calculé à partir de la composante en curvelets, un lissage sur  $s = 10$  segments, tracé en rouge par dessus l'image originale, pour un seuil de persistance  $p = 0, 0.25, 0.5,$  et  $0.75$  respectivement de gauche à droite puis de haut en bas.

l'influence du lissage en figure 10. On observe un réel effet en regardant le squelette non lissé, en haut à gauche, par rapport aux squelettes lissés. Les zigzags à l'échelle du pixel sont supprimés avec le lissage au profit d'une orientation plus régulière le long du filament. Ceci est important, puisque nous allons comparer l'orientation de la polarisation avec un segment de filament - le plus proche - donc celui-ci doit représenter la tendance globale du filament près de ce point. L'influence du paramètre du lissage est peu visible après  $s = 3$ . Un lissage de 10 nous paraît satisfaisant : un segment représente son filament à l'échelle d'une dizaine de pixels. On pourrait finalement se demander ce que donne `DisPerSE` sur une image non traitée, c'est-à-dire directement sur l'image originale. Le résultat est présenté en figure 11 pour des

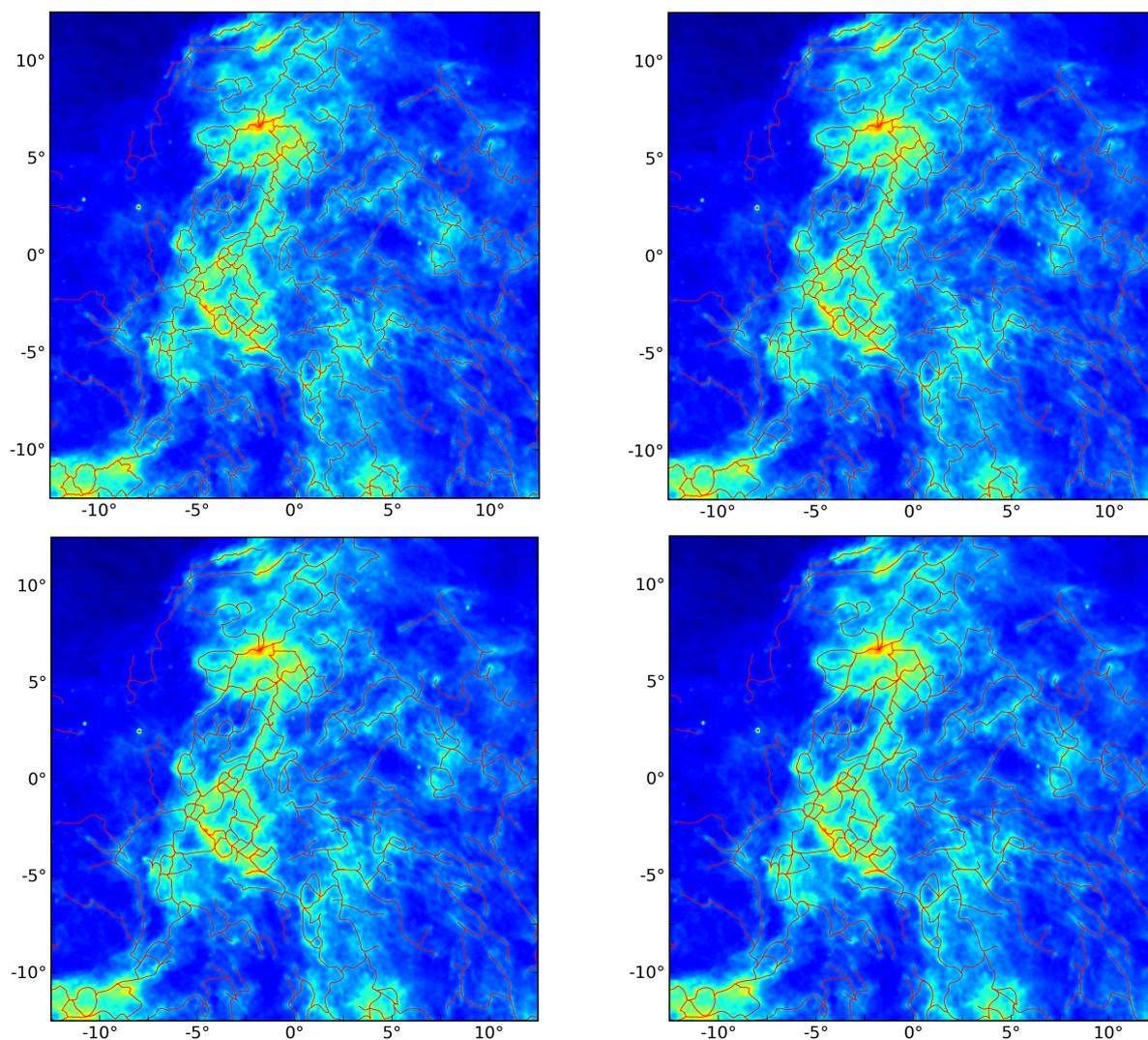


FIGURE 10 – Squelette calculé à partir de la composante en curvelets, un seuil de persistance  $p = 0.5$ , tracé en rouge par dessus l’image originale, sans lissage puis pour un lissage sur  $s = 3$ , 10, 25 segments, respectivement de gauche à droite puis de haut en bas.

paramètres classiques ( $s = 10$  et seuil de persistance faisant apparaître un nombre de filaments équivalent à  $p = 0.5$  pour la composante en curvelets), et est à comparer avec la figure 8. Le résultat est assez bon aussi, et proche de ce qu'on a obtenu précédemment. On peut tout de même reprocher l'attraction systématique des sources ponctuelles, qui sont toujours reliées à des filaments (ce sont en effet des maxima, donc des points critiques), alors qu'elles ne font pas partie de Polaris. C'est un biais qui est évité avec succès en utilisant l'image filtrée en curvelets.

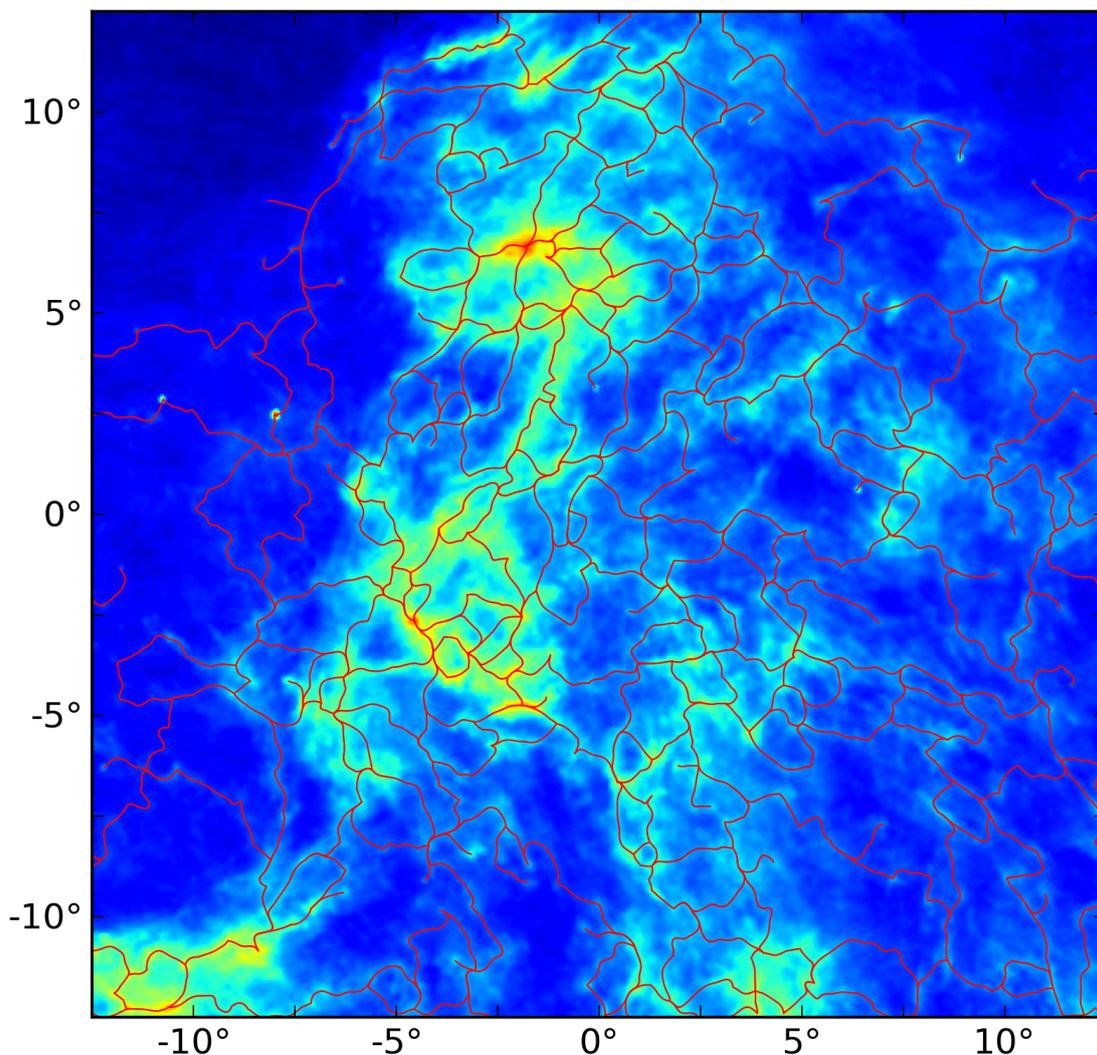


FIGURE 11 – Squelette calculé à partir de l'image originale, un seuil de persistance  $p = 0.64$ , un lissage sur  $s = 10$  segments, tracé en rouge par dessus l'image originale. Le seuil de persistance de 0.64 permet d'obtenir un nombre de filaments apparaissant sur la figure équivalent à celui du squelette calculé à partir de l'image filtrée en curvelets pour un seuil de 0.5, ce qui permet de les comparer pertinemment. On remarque que le squelette est "attiré" par les sources ponctuelles.

Nous sélectionnons finalement le squelette calculé à partir de la composante en curvelets, avec un seuil de persistance de 0.5 et un lissage sur 10 segments, présenté en figure 8, comme étant bien représentatif du réseau filamenteux de Polaris présent sur l'image de la figure 4. Il faut néanmoins garder à l'esprit que ce choix est assez subjectif, et que l'on peut rester souple sur les paramètres choisis dans nos prochaines analyses.

## 3 Statistique d'orientation

### 3.1 Les filaments

Nous disposons maintenant d'un squelette représentatif, et nous nous intéressons à ses propriétés. Dans cette section, nous traçons la statistique d'orientation des filaments pour différents paramètres. Pour ce faire, nous observons l'orientation de chaque segment composant le squelette par rapport à l'horizontale de l'image. La figure 12 montre des histogrammes de ces orientations pour un squelette calculé à partir de la composante en curvelets, avec un seuil de persistance de 0.5 et sans lissage, puis avec un lissage sur 10 et 25 segments. Comme il a été précisé lors de la description de `DisPERSE`, sans lissage, les coordonnées des extrémités des segments ne peuvent prendre que des valeurs discrètes, ces extrémités étant placées soit dans un coin de pixel, soit au milieu d'un côté de pixel, avec des segments d'une taille toujours proche de un pixel. On le vérifie avec le premier histogramme, en haut : les angles ne peuvent prendre que 8 valeurs. On remarque que ces orientations privilégiées par l'algorithme sont encore sur-représentées avec un lissage sur 10 segments (au milieu). C'est un biais que nous souhaitons éviter, et pousser le lissage à 25 segments donne un résultat satisfaisant (en bas).

On présente un histogramme plus clair, avec seulement 10 sous-intervalles, pour les orientations des filaments du squelette calculé à partir de la composante en curvelets, avec un seuil de persistance de 0.5 et avec un lissage sur 25 segments, en figure 13. La répartition des angles est relativement uniforme, on ne trouve pas vraiment d'orientation privilégiée pour les filaments de cette image de Polaris.

Rappelons que le choix du squelette représentatif était quelque peu subjectif. On peut donc à juste titre se demander ce que donneraient ces statistiques en changeant un peu les paramètres. La figure 14 présente des histogrammes en prenant un seuil de persistance de 0.75, puis en traçant le squelette à partir de l'image originale, avec un seuil de persistance de 0.64. Le lissage est toujours fait sur 25 segments, désormais. Les résultats obtenus sont assez similaires à ceux de la figure 13, même lorsque le squelette est calculé à partir de l'image originale, et donc biaisé par la présence de sources ponctuelles.

### 3.2 Le champ magnétique

Nous avons vu que, d'après [2], l'orientation de la polarisation de la lumière est a priori parallèle à l'orientation du champ magnétique dans le plan du ciel. Les données de la table 1 nous indiquent donc la direction du champ magnétique en différents points de l'image. On peut alors accéder à la statistique d'orientation du champ magnétique dans Polaris. Les angles sont recalculés afin d'avoir pour référence l'horizontale, dans l'image, et d'être en radians. Les données manquent de précision pour quelques étoiles, où la polarisation est quelquefois faible : nous décidons donc d'appliquer un filtre pour ne conserver que les données correspondant à une lumière polarisée à plus de 5% et une incertitude à  $1\sigma$  sur l'angle inférieure à 40 degrés. Nous gardons les données provenant de 46 étoiles après cette filtration. La figure 15 présente les histogrammes de l'orientation du champ magnétique à partir des données filtrées et non filtrées. Notons que le champ magnétique, en ces points de Polaris, présente une orientation privilégiée autour de l'axe vertical de l'image (c'est-à-dire des pics à  $\pm\pi/2$ ).

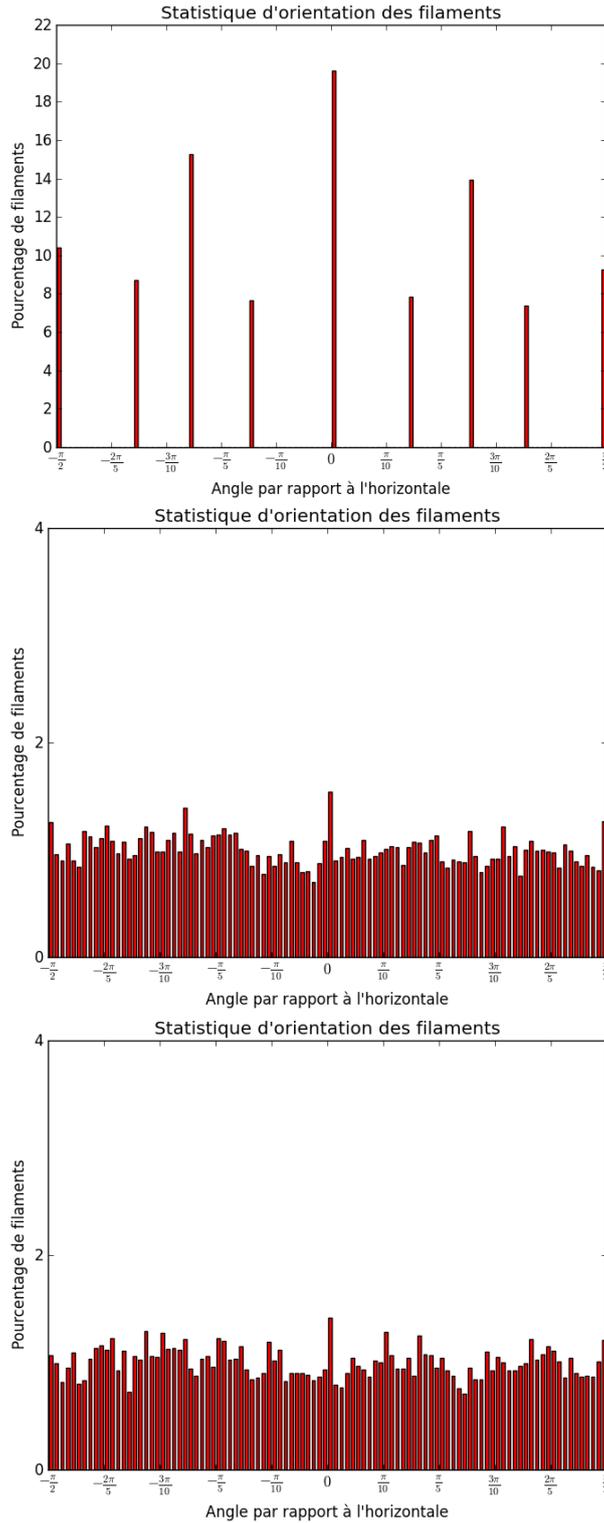


FIGURE 12 – Histogrammes des orientations des filaments du squelette calculé à partir de la composante en curvelets, avec un seuil de persistance de 0.5 et sans lissage, puis avec un lissage sur 10 et 25 segments, respectivement de haut en bas. Les angles sont donnés en radians, définis comme nuls sur l'axe horizontal de l'image, et orientés dans le sens polaire, c'est à dire du Nord vers l'Est. L'intervalle  $[-\pi/2, \pi/2]$  est divisé en 100 parties égales.

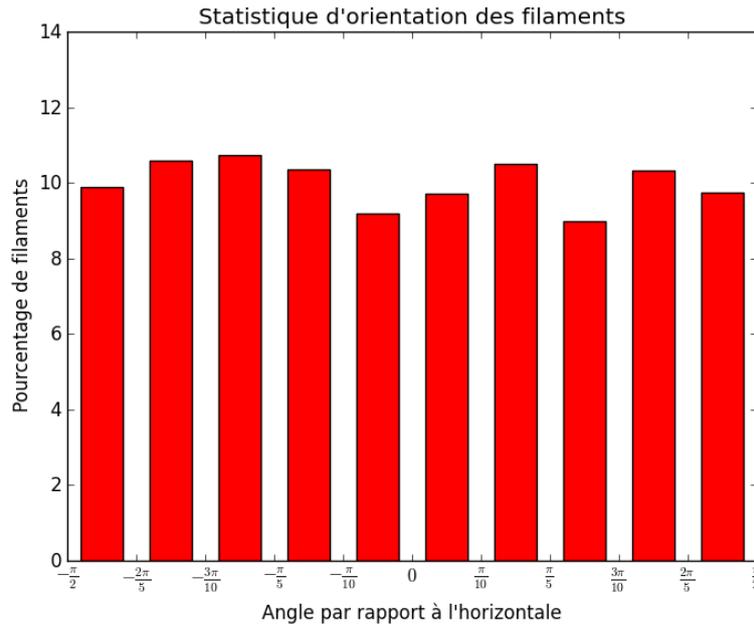


FIGURE 13 – Histogramme des orientations des filaments du squelette calculé à partir de la composante en curvelets, avec un seuil de persistance de 0.5 et avec un lissage sur 25 segments. Les angles sont donnés en radians, définis comme nuls sur l’axe horizontal de l’image, et orientés dans le sens polaire, c’est à dire du Nord vers l’Est. L’intervalle  $[-\pi/2, \pi/2]$  est divisé en 10 parties égales.

## 4 Résultats et discussion

Nous pouvons désormais nous pencher sur le cœur de notre problématique, c’est-à-dire le lien entre l’orientation du champ magnétique et celle des filaments. Il est facile de convertir les coordonnées J2000 de la table 1 en coordonnées sur l’image, toujours avec l’outil `sky2xy` des `WCSTools` [9]. La figure 16 présente l’image originale de Polaris, avec en superposition le squelette calculé à partir de la composante en curvelets, avec un seuil de persistance de 0.5 et un lissage sur 25 segments, en rouge, et en noir les données de polarisation. L’orientation privilégiée sur l’axe vertical remarquée en figure 15 est bien présente.

Les points où l’on connaît la polarisation sont souvent proches d’un filament. On peut donc s’intéresser à la statistique de l’écart d’angle entre les orientations du champ magnétique et celles des filaments les plus proches. Pour ce faire, on a sélectionné, pour chaque étoile pour laquelle on connaissait la polarisation de la lumière, le segment dont le centre était le plus proche de cette position, et calculé l’écart d’orientation (après lissage, bien entendu). Un histogramme de la valeur absolue de cet écart calculé à partir des données utilisées pour la figure 16 est présenté en figure 17. Cette distribution d’angles présente deux pics, en  $\pi/10$  et  $7\pi/20$ .

On peut se demander si ces deux pics sont retrouvés avec des paramètres différents. La figure 18 montre des histogrammes pour un seuil de persistance plus faible, plus élevé, ou encore avec un squelette calculé à partir de l’image originale. Le pic en  $\pi/10$  est quasiment systématiquement retrouvé. Celui en  $7\pi/20$  est bien moins sûr. Pour confirmer que les pics remarqués dans les figures 17 et 18 ne sont pas dûs au bruit, on utilise les quantiles de la statistique du  $\chi^2$ , pour déterminer la probabilité que la distribution des écarts d’orientation, sous-

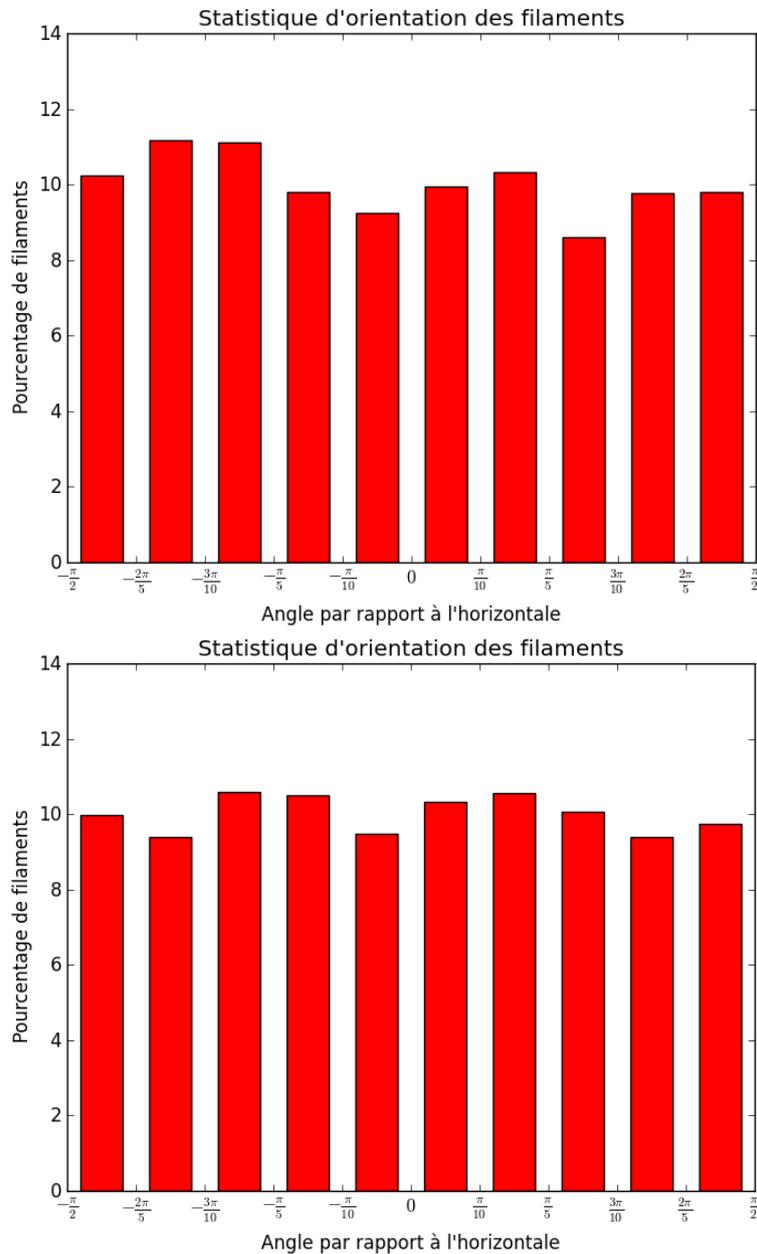


FIGURE 14 – Histogrammes des orientations des filaments du squelette. Les angles sont donnés en radians, définis comme nuls sur l’axe horizontal de l’image, et orientés dans le sens polaire, c’est à dire du Nord vers l’Est. L’intervalle  $[-\pi/2, \pi/2]$  est divisé en 10 parties égales. En haut : Squelette calculé à partir de la composante en curvelets, un seuil de persistance de 0.75, et un lissage sur 25 segments. En bas : Squelette calculé à partir de l’image originale, un seuil de persistance de 0.64, et un lissage sur 25 segments.

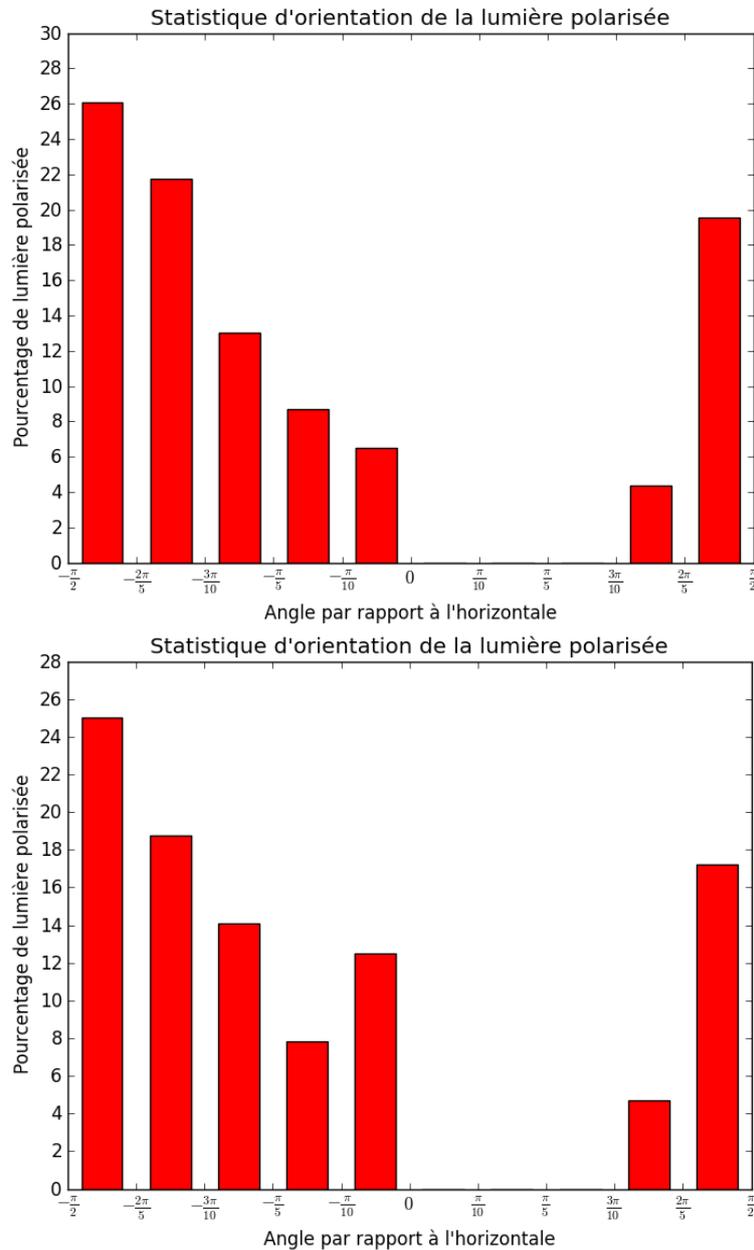


FIGURE 15 – Histogrammes des orientations du champ magnétique en certains points de Polaris. Les angles sont donnés en radians, définis comme nuls sur l'axe horizontal de l'image, et orientés dans le sens polaire, c'est à dire du Nord vers l'Est. L'intervalle  $[-\pi/2, \pi/2]$  est divisé en 10 parties égales. En haut : Données de la table 1 filtrées pour avoir une polarisation supérieure à 5% et une incertitude sur l'angle à  $1\sigma$  inférieure à 40 degrés. En bas : Données de la table 1 brutes.

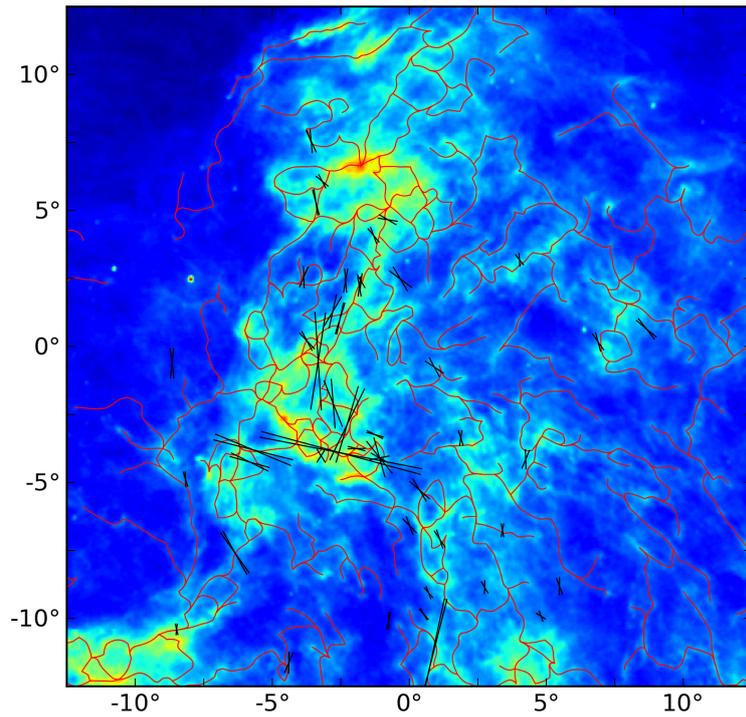


FIGURE 16 – Image originale de Polaris, avec, en superposition, le squelette calculé à partir de la composante en curvelets, un seuil de persistance  $p = 0.5$ , un lissage sur  $s = 25$  segments (en rouge) et les données de polarisation filtrées (en noir). Les deux segments formant une croix représentent les orientations de polarisation avec les angles minimal et maximal mesurés, à  $1\sigma$ . La longueur des segments est proportionnelle à la proportion de lumière polarisée. Pour rappel, la polarisation maximale est de 77%.

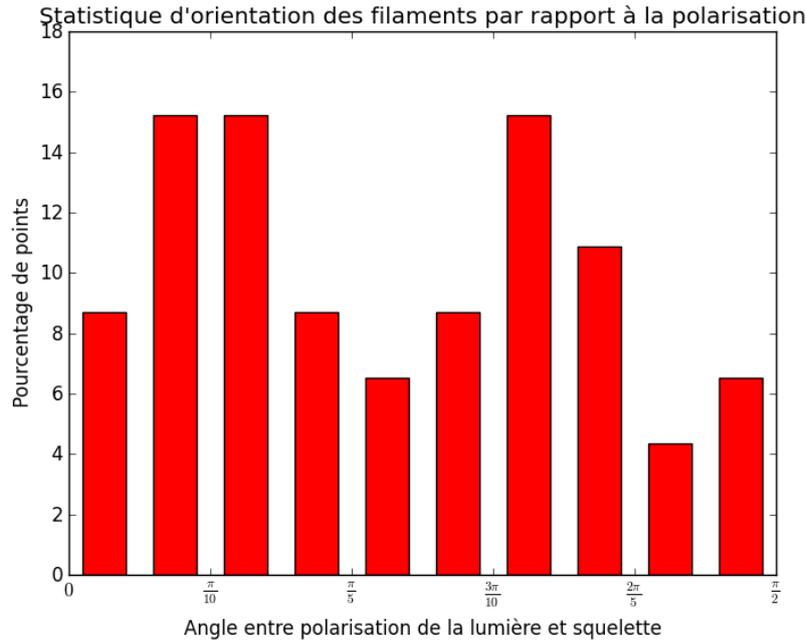


FIGURE 17 – Histogramme de la valeur absolue de l'écart angulaire entre les orientations du champ magnétique et celles du filament le plus proche. Le squelette utilisé est calculé à partir de la composante en curvelets, un seuil de persistance  $p = 0.5$ , un lissage sur  $s = 25$  segments. Les données du champ magnétique sont les données de polarisation filtrées.

jacente aux histogrammes présentés, soit différente d'une distribution uniforme bruitée, c'est à dire qu'elle contienne effectivement de l'information (méthode décrite dans [16]). L'application de cette méthode à ces histogrammes montre que la probabilité que les distributions sous-jacentes ne soient pas uniformes est supérieure à 95%, pour les quatre cas. Nous pouvons donc affirmer que nous avons mis en évidence une corrélation entre l'orientation du champ magnétique et celle des filaments. Les corrélations déjà mises en évidence par d'autres travaux ne sont pas en désaccord avec celle-ci, car très variables. Par exemple, dans Taurus, des zones où les filaments ont tendance à être orientés parallèlement aux lignes de champ magnétique ont été mises en évidence, tout comme d'autres où l'orientation est plutôt perpendiculaire [17].

Nous pouvons noter cependant qu'il est tout à fait possible d'affiner ces résultats. Tout d'abord, remarquons que pour le calcul du squelette, la filtration des filaments via le seuil de persistance n'est pas optimal. En effet, si un filament très brillant et bien défini est trop uniformément brillant, bien que parfait pour être identifié visuellement, il sera écarté par le seuil de persistance si celui-ci est trop haut. Pour s'en persuader, imaginons le cas idéal d'un filament avec une intensité variant uniquement au niveau du bruit quand on le parcourt : la persistance associée à ses paires de points critiques sera nécessairement dans le niveau de bruit, puisque c'est une différence d'intensité au sein du filament. En portant un œil attentif aux figures 8, 9 ou 11, on pourra remarquer que des filaments tout à fait distinguables sont écartés par la filtration. On pourrait peut-être améliorer ceci en fixant un seuil de persistance évolutif selon l'intensité de son point critique le plus bas, par exemple. La filtration serait plus flexible dans les zones lumineuses de l'image. Il est évident aussi que la statistique d'orientation obtenue gagnerait à être comparée à celle obtenue avec la même méthode pour d'autres nuages, afin de

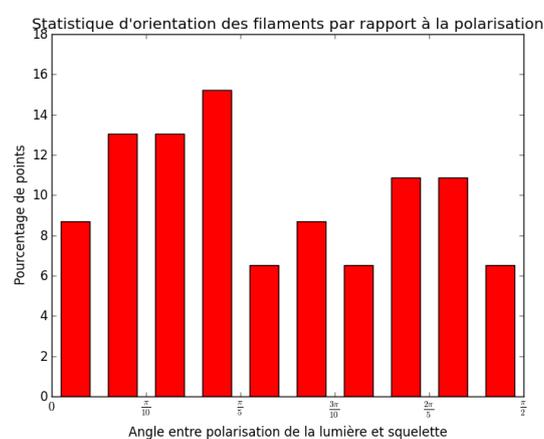
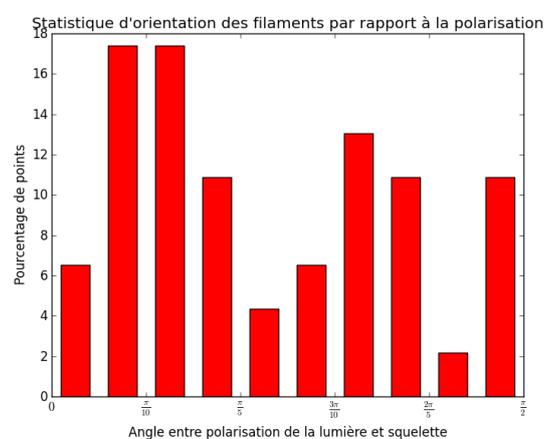
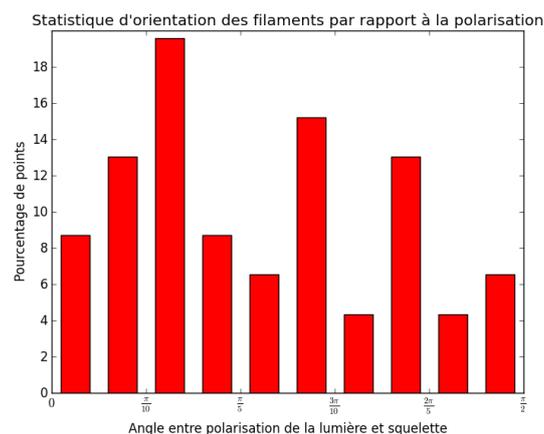


FIGURE 18 – Histogramme de la valeur absolue de l'écart angulaire entre les orientations du champ magnétique et celles du filament le plus proche. En haut : le squelette utilisé est calculé à partir de la composante en curvelets, un seuil de persistance  $p = 0.25$ , un lissage sur  $s = 25$  segments. Au milieu : le squelette utilisé est calculé à partir de la composante en curvelets, un seuil de persistance  $p = 0.75$ , un lissage sur  $s = 25$  segments. En bas : le squelette utilisé est calculé à partir de l'image originale, un seuil de persistance  $p = 0.64$ , un lissage sur  $s = 25$  segments.

la rendre plus précise, voire de vérifier si la corrélation est comparable. Rappelons que nous ne disposons que de moins d'une cinquantaine de données de polarisation, ce qui est peu pour faire des statistiques. Il n'est alors pas envisageable d'écarter les points trop loin de filaments, ce qui pourrait pourtant rendre plus précis nos résultats. Il est cependant difficile d'obtenir de meilleures données pour Polaris. Son éloignement du plan galactique réduit le nombre d'étoiles d'arrière-plan susceptibles d'être analysées, et le polarimètre utilisé était déjà très performant. On pourrait imaginer travailler directement avec les coefficients de la transformée en curvelets, puisqu'ils contiennent des informations de position et d'orientation. Finalement, notons que la difficulté d'accès à une version fonctionnelle et récente de MCA a empêché une utilisation approfondie de ses possibilités via une exploration plus ample des paramètres, et qu'il est peut-être possible d'obtenir une composante en curvelets montrant encore mieux les filaments.

## Conclusion

Finalement, nous avons obtenu la corrélation recherchée entre l'orientation du champ magnétique et celle des filaments au sein du nuage moléculaire Polaris. Elle se présente sous la forme d'un pic de la statistique des différences d'angle entre champ magnétique et filament le plus proche autour de  $\pi/10$ . Ce lien peut servir à contraindre les modèles de formation et d'évolution des filaments [18] [19], et s'inscrit dans le long cheminement vers une meilleure compréhension de la formation stellaire. La méthode mise en place permet une très bonne identification et cartographie des filaments d'une image, et présente quelques évolutions possibles afin de la rendre encore meilleure. Elle offre une perspective intéressante, avec l'arrivée prochaine et attendue des données de polarisation du satellite Planck, couvrant l'ensemble du ciel avec une grande précision et un grand champ, qui vont permettre de réitérer cette analyse dans beaucoup d'autres régions du ciel. Cela permettra de voir si le résultat obtenu s'applique à d'autres régions célestes.

## Remerciements

Ce stage m'a beaucoup apporté, en me donnant une nouvelle expérience du milieu de la recherche, en me laissant aborder un nouveau champ de l'astrophysique passionnant, et en m'apportant tout un éventail de nouvelles techniques et savoirs qui me seront utiles dans la suite de ma carrière. Je suis plus que jamais décidé à faire mon possible pour apporter ma pierre au formidable édifice qu'est la recherche.

Je tiens donc à remercier chaleureusement François et Edith pour l'accueil au sein de leur équipe, pour leur disponibilité, leur encadrement, et pour le partage de leur expérience et leur savoir qui m'ont permis d'avancer tout au long de ce stage. De grands remerciements à l'équipe du master de l'Observatoire de Paris pour m'avoir mis un pied solide dans le monde de la recherche en astrophysique via ses enseignements et les opportunités offertes. Merci aussi à Thierry Sousbie pour la mise à disposition de son algorithme `DisPerSE` avant sa sortie officielle, et son aide précieuse pour l'utilisation du code alors qu'aucune documentation n'était encore écrite. Je suis reconnaissant envers Pierre Didelon pour ses conseils et son temps quand je rencontrais des difficultés avec MCA. Merci enfin aux membres du laboratoire que je ne pourrais citer, avec qui j'ai partagé ces trois mois.

## Références

- [1] T. Hoang & A. Lazarian, *Grain Alignment Induced by Radiative Torques : Effects of Internal Relaxation of Energy and Complex Radiation Field*, **697**, 2, 1316 (2009).
- [2] N.L. Chapman *et al.*, *The Magnetic Field In Taurus Probed By Infrared Polarization*, *ApJ*, **741**, 1, 21 (2011).
- [3] A. Heithausen *et al.*, *On the dynamical state of high-latitude molecular clouds*, *A & A*, **314**, 251 (1996).
- [4] D. Arzoumanian *et al.*, *Characterizing interstellar filaments with Hershel in IC 5146*, *A & A*, **529**, L6 (2011).
- [5] H.W. De Vries *et al.*, *Molecular and atomic clouds associated with infrared cirrus in Ursa Major*, *ApJ*, **319**, 723 (1987).
- [6] F. Zagury *et al.*, *Optical images of MCLD123.5+24.9 : a cloud illuminated by the North star ?*, *A & A*, **352**, 645 (1999).
- [7] M.-A. Miville-Deschênes and G. Lagache, *IRIS : A new generation of IRAS maps*, *ApJSS*, **157**, 302 (2005).
- [8] K. Mohan Narasimhan, encadré par P. Hily-Blant, *Magnetic fields in turbulent molecular clouds*, Rapport de stage, [http://ipag.osug.fr/~philybla/data/rapport\\_kirthika.pdf](http://ipag.osug.fr/~philybla/data/rapport_kirthika.pdf) (2010).
- [9] WCSTools, *Image World Coordinate System Utilities*, <http://http://tdc-www.harvard.edu/wcstools/>.
- [10] J.L. Starck *et al.*, *Astronomical image representation by the cruvelet transform*, *A & A*, **398**, 785 (2003).
- [11] K. Schneider and M. Farge, *Wavelets : Mathematical Theory*, *Encyclopedia of Mathematical Physics*, Elsevier, 426 (2006).
- [12] M.J. Fadili *et al.*, *MCA Lab : Reproducible Research in Signal and Image Decomposition and Inpainting*, *Comp. in Sc. & Eng.*, **12**, 44 (2010).
- [13] T. Sousbie, *The persistent cosmic web and its filamentary structure - I. Theory and implementation*, *Mon. Not. R. Astron. Soc.*, **414**, 350 (2011).
- [14] R. Forman, *Adv. Math.*, **134**, 90 (1998).
- [15] T. Sousbie *et al.*, *The fully connected N-dimensional skeleton : probing the evolution of the cosmic web*, *Mon. Not. R. Astron. Soc.*, **393**, 457 (2009).
- [16] F. Levrier *et al.*, *Fourier phase analysis in radio-interferometry*, *A & A*, **456**, 205 (2006).
- [17] P.F. Goldsmith *et al.*, *Large-scale structure of the molecular gas in Taurus revealed by high linear dynamic range spectral line mapping*, *ApJ*, **680**, 428 (2008).
- [18] J.D. Fiege *et al.*, *Polarized submillimeter emission from filamentary molecular clouds*, *ApJ*, **544**, 830 (2000).
- [19] T. Nagai *et al.*, *An origin of filamentary structure in molecular clouds*, *ApJ*, **506**, 306 (1998).

## Annexes

### A Algorithme de MCA

Ici est détaillé l'algorithme de MCA, dans une présentation très proche de celle faite dans [12].

---

#### Algorithme de séparation des composantes (MCA)

---

**Paramètres :**

L'image  $F$ , le dictionnaire  $\Phi$ , un nombre d'itérations maximal  $N_{iter}$ , un seuil d'arrêt  $\lambda_{min}$ , et une stratégie de décroissance du seuil (linéaire ou exponentielle).

**Initialisation :**

- la solution : pour tout  $k$ ,  $F_k^{(0)} = 0$ .
- le résidu :  $r^{(0)} = F$
- le seuil : on prend  $k^*$  le  $k$  qui maximise  $\|\Phi_k F\|_\infty$ , et alors  $\lambda^{(0)} = \max_{k \neq k^*} \|\Phi_k F\|_\infty$ .

**Itération :**

Pour  $t = 1$  à  $N_{iter}$ ,

- Pour  $k = 1$  à  $K$ ,
    - Calcul de l'image et des résidus pour la composante  $k$  :  $r_k^{(t)} = r^{(t)} + F_k$ .
    - Mise à jour des coefficients plus grand que le seuil :  $\alpha_k^{(t)} = S_{\lambda^{(t)}}(\Phi_k r_k^{(t)})$ .
    - Mise à jour de l'image composante :  $F_k^{(t)} = \Phi_k \alpha_k^{(t)}$ .
  - Mise à jour des résidus :  $r^{(t+1)} = F - \sum_{k=1}^K F_k^{(t)}$ .
  - Mise à jour du seuil  $\lambda^{(t+1)}$  selon la stratégie choisie, linéairement ou exponentiellement, vers 0.
  - Condition d'arrêt : si  $\lambda^{(t+1)} \leq \lambda_{min}$ , sortie de la boucle.
- 

### B Oscillations gênantes pour MCALab

Nous avons rencontrés des oscillations non convergentes gênantes lors de l'utilisation de la version MATLAB de MCA présentée dans [12]. Nous avons utilisé ses fonctions sans passer par l'interface GUI. La figure 19 a été obtenue avec des paramètres très similaires à ceux utilisés avec la version binaire pour la figure 6. Le seul changement majeur a été l'utilisation d'UDWT (Undecimated Discrete Wavelet Transform) à la place de l'algorithme à trous pour les ondelettes. C'est à dire que nous avons traité l'image présentée en figure 4, avec un nombre d'échelle de 7, des ondelettes et des curvelets, 19 et 20 itérations, une valeur de  $\sigma$  et un seuil d'arrêt standard, une décroissance exponentielle du seuil et pas de Total Variation (voir [12]). On remarque clairement que les images obtenues après 19 itérations et 20 itérations sont très différentes, et celle à 20 itérations est très différente de l'image originale, elle semble ne contenir que très peu de coefficients. La figure 20 met en évidence l'oscillation dont nous avons fait part, puisque

c'est l'évolution de la valeur d'un pixel de l'image obtenue en superposant la composante en ondelettes et la composante en curvelets, donné pendant les 20 itérations.

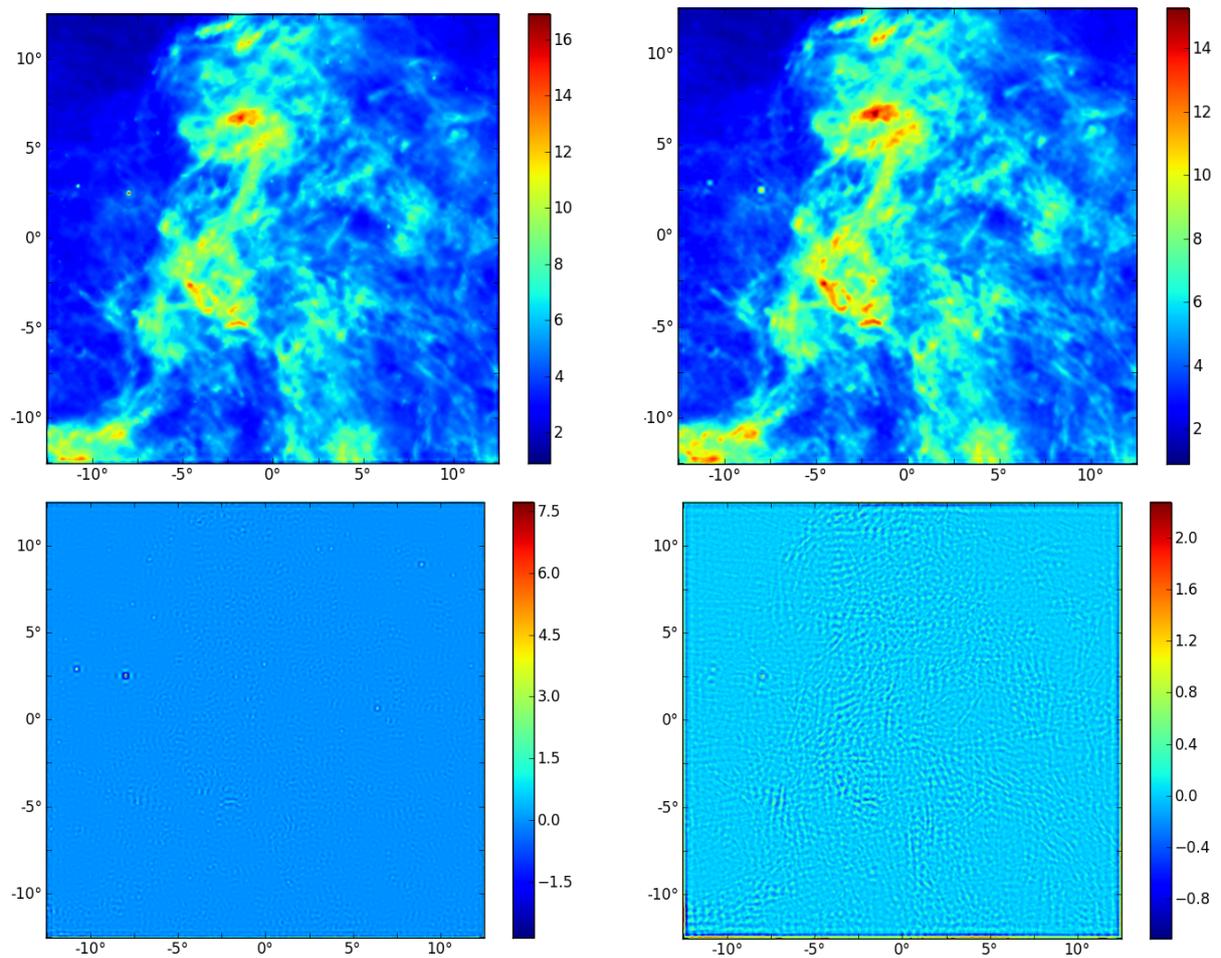


FIGURE 19 – Résultat obtenu avec MCALab dans des conditions similaires à celles de la figure 6 avec 19 itérations (en haut) et 20 itérations (en bas). La composante en ondelette est présentée à gauche, et celle en curvelets à droite.

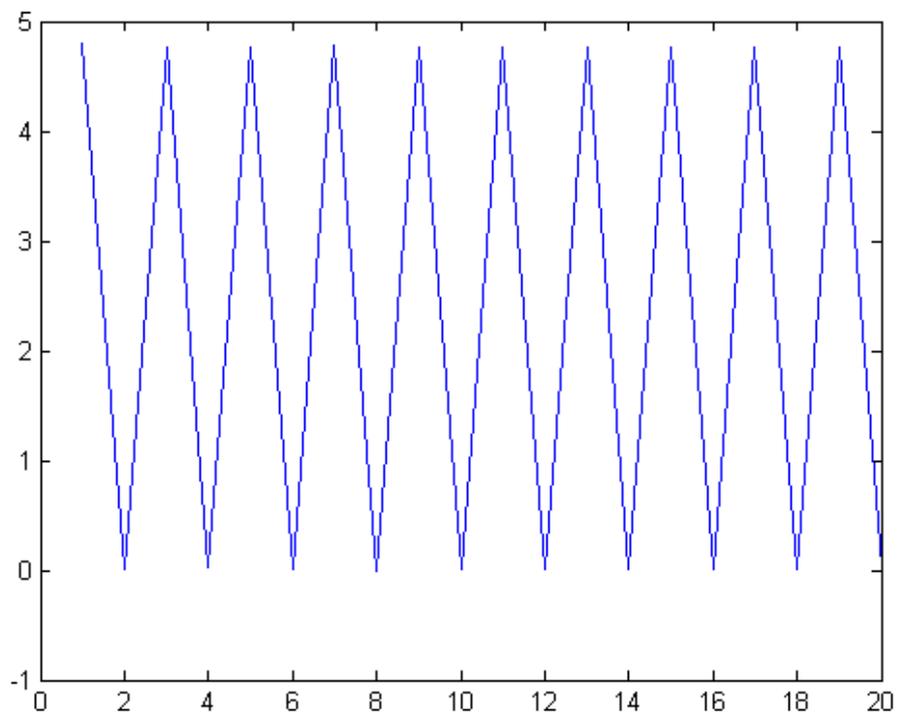


FIGURE 20 – Évolution de la valeur d'un pixel donné de l'image obtenue en superposant la composante en ondelettes et la composante en curvelets suivant les itérations de l'algorithme.