

# Mise en perspective didactique d'un dossier de recherche

Constant Auclair

Concours externe spécial de l'agrégation de physique-chimie option physique  
Session 2025

## 1 - Formations et expériences

### Formation académique

**2024 - 2025 :** Préparation à l'agrégation de physique au centre interuniversitaire de Montrouge, *École Normale Supérieure de Paris, Sorbonne Université, Université Paris-Saclay*

**2021 - 2024 :** Thèse de doctorat d'astrophysique de l'Université PSL, intitulée *Modélisation statistique des avant-plans interstellaires au fond diffus cosmologique*, préparée à l'*École Normale Supérieure de Paris* sous la direction de François Boulanger et Erwan Ally

**2020 - 2021 :** Master 2 International Center for Fundamental Physics (ICFP), parcours Theoretical Physics, *École Normale Supérieure de Paris*

**2019 - 2020 :** Master 1 Physique Fondamentale, *Université Paris-Cité*

**2018 - 2019 :** Licence 3 Physique, *Université de Nantes*

**2016 - 2018 :** Classe Préparatoire aux Grandes Écoles (CPGE), filière MPSI/MP, *Lycée Charlemagne*

### Expériences d'enseignement

**2021 - 2024 :** Mission d'enseignement (64h/an) à l'Observatoire de Paris

- Travaux Dirigés de Physique Statistique, *Master 1*
- Travaux Pratiques d'optique, *CPES*
- Stage de formation à l'astrophysique pour des professeurs de lycée
- Tutorat pédagogique pour des étudiants du Diplôme d'Université *Lumières sur l'Univers*
- Accompagnement d'élèves de primaire lors de la Fête de la Science

**2019 - 2023 :** Cours particuliers et stages

- Stages de mathématiques, niveau lycée, *Les cours du Parnasse*
- Stages de statistiques, niveau licence, *Les cours d'Assas*
- Cours de mathématiques, niveau CPGE, *Groupe Réussite*
- Tutorat de mathématiques, niveau collège, *Alveus Club*

**2019 - 2021 :** Colles de physique en MPSI, MP, PC et PC\*, *Lycée Charlemagne*

**2014 - 2015 :** Soutien scolaire bénévole pour des élèves de primaire, *Association Socrate*

## 2 - Enseignement

L'enseignement a pris une part importante de mon temps depuis que je suis au lycée. D'abord mené en parallèle de mes études, il a fini par y être inclus sous forme de missions doctorales pendant ma thèse. La prise de recul nécessaire pour enseigner une notion a toujours été pour moi un moteur de compréhension fine. Apprendre et enseigner sont donc allés de pair dans mes études, chacun bénéficiant de l'autre. Par ailleurs, au-delà des progrès disciplinaires des élèves, je considère que l'objectif d'un enseignement scientifique est aussi de donner le goût de la science et d'inculquer la rigueur de la démarche scientifique. Je vais ici revenir plus en détail sur trois de mes expériences d'enseignement qui me semblent intéressantes à présenter dans le contexte de ce dossier.

### Stage d'introduction à l'astrophysique

Dans le cadre de mes missions doctorales à l'Observatoire de Paris, j'ai participé à l'organisation d'un stage d'introduction à l'astrophysique pour des professeurs de lycées de l'académie de Créteil. Ce stage, d'une durée de trois jours, permettait aux professeurs d'acquérir des connaissances en astrophysique mais aussi de voir des exemples d'activités et de projets qui pourraient être menés avec leurs élèves. J'étais chargé de deux travaux pratiques et un cours. Le premier TP consistait à retrouver la loi de Hubble à partir d'observations spectroscopiques de galaxies lointaines, le second à caractériser une lunette astronomique et le cours était dédié aux instruments d'astrophysique. Enseigner à des professeurs était une tâche complexe, de par l'expérience qu'ils avaient, mais très enrichissante puisque leur approche des cours était orientée vers une adaptation future à leurs élèves. Nous réfléchissions donc ensemble à la façon de formuler les énoncés, les pièges dans lesquels les élèves pouvaient tomber et l'adaptation des activités au format d'un cours de lycée. Par exemple, les professeurs m'ont fait prendre conscience que le TP sur la loi de Hubble que je leur ai proposé ne pouvait pas du tout être proposé tel quel à des lycéens. En effet, bien que les calculs soient simples, le sujet abordé était complexe, à la fois dans les documents utilisés mais aussi dans la conclusion que les élèves devaient en tirer : il fallait d'abord savoir interpréter un spectre de galaxie, comprendre le lien entre le décalage vers le rouge et l'éloignement des galaxies, et enfin en déduire la loi de Hubble. Pour proposer un sujet de ce type à une classe de lycée, je veillerais à soigneusement identifier tous les éléments qui ne sont pas au programme et assurer leur compréhension à l'aide de questions intermédiaires guidées. En tant que jeune professeur docteur au sein d'une équipe pédagogique, je bénéficierais de l'expérience des professeurs aguerris tout en apportant mes connaissances pour la construction de projets ou séquences sur le thème de l'astrophysique.

### Encadrement de la Fête de la Science

Toujours dans le cadre de mes missions doctorales, j'ai participé à la Fête de la Science à l'Observatoire de Paris-Meudon. Pendant une journée, j'ai accompagné une classe de CM2 à travers différents ateliers présentés par des chercheurs et chercheuses de l'Observatoire. Nous avons commencé par visiter le spectrohéliophotographe, qui fournit quotidiennement des images du Soleil. Pour les élèves, dont certains dessinaient encore notre étoile comme une étoile à cinq branches, réaliser que le Soleil est une "boule de feu" constituait un important changement de paradigme. Ensuite, la classe participa à un atelier sur l'exploration de Mars. Le concept était de réfléchir, par petits groupes, à tout ce dont on pourrait avoir besoin pour explorer la planète rouge. La chercheuse qui encadrait l'atelier déconstruisait au fur et à mesure ce que les élèves imaginaient d'une autre planète : la pesanteur est-elle la même que sur Terre ? l'atmosphère est-elle constituée des mêmes choses ? fait-il la même température ? Petit à petit, les élèves réalisaient à quel point notre planète est accueillante au sein du système solaire. L'atelier se terminait donc sur une sensibilisation à la protection de l'environnement, axée sur les risques du changement climatique mais aussi sur la préservation de la biodiversité. Enfin, le dernier atelier reposait sur une maquette du système solaire rarement vue, puisqu'elle était à l'échelle. Le Soleil et les planètes, dont les tailles variaient d'une vingtaine de centimètres à quelques millimètres, étaient donc dispersés sur plusieurs dizaines de mètres dans le jardin de l'Observatoire. Au-delà des présentations du Soleil et des planètes par l'encadrant, ce sont les distances qui ont le plus marqué les élèves, habitués aux affiches condensées accrochées dans les classes. En tant que professeur, je veillerai à sensibiliser les élèves sur les nombreuses simplifications de représentations que l'on rencontre en physique, du modèle planétaire de l'atome jusqu'au système solaire. En particulier, l'astrophysique

étant une porte d'entrée classique dans la physique, c'est un domaine où je serai amené à devoir déconstruire des préconceptions erronées des élèves. Pendant cette journée, mon rôle a été de faire l'intermédiaire pédagogique entre les élèves et les chercheurs, dont le discours était généralement peu adapté au niveau des élèves. Il fallait donc trouver les mots justes pour simplifier les notions sans pour autant perdre trop en richesse scientifique. Cette expérience a été pour moi très révélatrice de l'importance de s'adapter à son auditoire, aussi bien dans un cadre de vulgarisation que d'enseignement.

## Tutorat de mathématiques

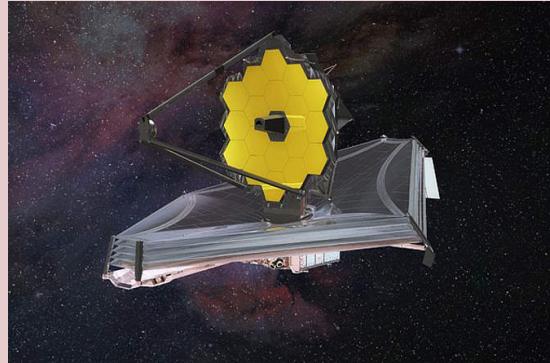
Enfin, pendant ma première année de master, j'ai travaillé pour une entreprise qui proposait des séances de tutorat pour des petits groupes de collégiens. Les élèves venaient, plusieurs fois par semaine, avec leurs devoirs ou avec une notion à retravailler, principalement en mathématiques. Mon rôle était de les accompagner individuellement dans leur travail mais aussi de créer une dynamique de groupe dans laquelle chacun trouverait sa place. Pour ce faire, j'ai mis en place différentes formes sociales de travail au service de différentes modalités d'apprentissage. Par exemple, j'ai instauré dès le début de l'année un système de correction croisée, dont l'objectif était que les exercices faits soient d'abord soumis au regard d'un camarade. Les élèves devaient juger, s'ils en étaient capables, de la justesse des réponses proposées par leur camarade mais aussi de la qualité de la rédaction et de la présentation. Cela permettait de faire collaborer des élèves qui n'avaient pas forcément le même niveau disciplinaire mais qui pouvaient quand même s'aider à progresser. Nous avons aussi beaucoup travaillé sous forme de classe inversée, dans laquelle un élève expliquait au reste du groupe une notion qu'il estimait maîtriser. Souvent, la difficulté à mettre des mots rigoureux sur des idées pointait les zones fragiles de la compréhension de la notion. L'élève jouant le professeur apprenait donc tout autant que ceux qui l'écoutaient. Enfin, je commençais chaque séance par un quiz d'une dizaine de minutes sur des questions du concours Kangourou. Plutôt que les élèves répondent sur leur ardoise, ils votaient pour la réponse correcte à chaque question et débattaient ensemble en cas de désaccord. Afin de rendre les débats constructifs, je proposais à chaque parti de fournir une démonstration de leur réponse ou un contre-exemple à la réponse adverse. Cette expérience d'enseignement a vraiment été pour moi un laboratoire pédagogique dans lequel j'étais très libre et j'ai donc pu mettre en place beaucoup d'ateliers aux pédagogies variées. Bien que ces tutorats aient été menés en mathématiques et avec des petits groupes d'élèves, je pense que la majeure partie des méthodes pédagogiques évoquées pourront être utilisées dans ma classe de physique-chimie, en prenant soin de les adapter en fonction du niveau, de l'effectif et du contexte de la classe. Je pense notamment aux corrections croisées et à la classe inversée. Cependant, afin d'avancer au rythme soutenu exigé par le contenu des programmes, ces ateliers seront mis en place de façon réfléchie et parcimonieuse.

## 3 - Activités pédagogiques

Au regard de mon expérience en astrophysique, j'aimerais proposer au sein de ma classe des activités pédagogiques qui permettraient aux élèves de s'exercer sur des notions du programme tout en éveillant leur curiosité scientifique. Je vous propose les deux activités suivantes, l'une en CPGE et l'autre au lycée.

### Activité 1 : *James Webb Space Telescope*

L'objectif de cette activité est d'étudier le *James Webb Space Telescope* (JWST) sous l'angle de différents aspects de la physique. On s'intéressera aux propriétés optiques, mécaniques et thermodynamiques de ce satellite ancré dans l'actualité astrophysique. Cette activité, qui se déroulerait en fin d'année, prendrait la forme d'un travail de 3 heures par groupe de 2 ou 3 élèves. Elle permettrait aux élèves de s'entraîner sur des notions apprises dans différents chapitres vus en cours tout au long de l'année. Après la mise en commun des résultats obtenus, la classe pourrait recevoir un astrophysicien qui leur présenterait ses recherches en lien avec les découvertes actuelles du JWST.



**Niveau :** CPGE 2<sup>ème</sup> année.    **Thèmes :** Optique, Mécanique, Thermodynamique.

**Présentation :** Le *James Webb Space Telescope* (JWST) est un télescope spatial développé par la NASA avec la participation de l'Agence Spatiale Européenne. Depuis son lancement fin 2021, le JWST collecte des images infrarouges de l'Univers qui constituent une mine d'information pour les astrophysiciens. On se propose d'étudier les propriétés optiques, mécaniques et thermodynamiques du JWST.

**Propriétés optiques :** On propose aux élèves une image des Piliers de la Création, une célèbre nébuleuse, prise par le JWST. On leur fournit comme informations la distance entre le Soleil et cette nébuleuse, sa taille ainsi que les dimensions de l'image en pixels. Les élèves doivent alors calculer la résolution angulaire, en degré, du JWST. On leur propose ensuite de comparer cette résolution avec celle d'un œil humain. Enfin, on propose une autre image des Piliers de la Création, prise cette fois-ci par le satellite Hubble. Que remarque-t-on en comparant les étoiles dans ces deux images ? Les élèves constatent que les tâches de diffraction des deux télescopes sont différentes, et peuvent relier cette observation à leur différence de structure.

**Propriétés mécaniques :** On propose aux élèves d'étudier l'orbite du JWST. On commence par définir les points de Lagrange, et on demande aux élèves de calculer la position des trois qui sont sur l'axe Soleil-Terre. D'après vous, sur lequel de ces trois points a-t-on intérêt à placer le JWST ? Les élèves écrivent ensuite un programme Python pour calculer la position des deux autres points de Lagrange.

**Propriétés thermodynamiques :** Enfin, on propose d'étudier le bouclier thermique du JWST. On fournit aux élèves un schéma légendé du télescope sur lequel sont données les tailles des différents éléments. Dans le cas où le bouclier n'est pas déployé, évaluer la température du JWST en considérant qu'il est éclairé par le Soleil. Pourquoi a-t-on alors besoin d'un bouclier thermique ? Question à prise d'initiative : de quel effet physique connu se rapproche le principe de fonctionnement du bouclier thermique ? Sachant que le bouclier du JWST est composé de 5 couches et que la température du télescope est d'environ 50 K quand il est protégé par le bouclier, estimer la fraction du flux lumineux incident que transmet une couche du bouclier.

## Activité 2 : Effet Doppler et vitesse de la galaxie d'Andromède

L'objectif de cette activité est de présenter une application de l'effet Doppler en astrophysique. L'idée est de sortir des exemples habituels, type ambulance, pour intéresser les élèves et susciter des questions. On s'intéresse dans un premier temps à la mesure de la vitesse radiale de la galaxie d'Andromède puis on en déduit une méthode pour mesurer sa vitesse de rotation. Cette activité prendrait la forme d'un travail par groupes de 3 ou 4 élèves, pour une durée d'environ une heure, correction comprise. Des éléments de réponse sont donnés entre crochets. La question 5, dont la résolution nécessite une prise d'initiative, est proposée avec trois niveaux de différentiation.



**Niveau :** Terminale Générale.      **Thème :** Ondes et signaux.

La galaxie d'Andromède est la galaxie la plus proche de la Voie Lactée. Dans le milieu galactique, des atomes d'hydrogène sont présents sous forme de gaz. Ils émettent des ondes électromagnétiques dont la longueur d'onde vaut  $\lambda_0 = 21,00 \text{ cm}$ . Ces ondes sont captées par des télescopes, et fournissent aux astrophysiciens des informations précieuses que nous allons étudier ici.

**Question 1 :** Des astrophysiciens remarquent que les ondes électromagnétiques émises par les atomes d'hydrogène de la galaxie d'Andromède arrivent dans leur télescope avec une longueur d'onde  $\lambda$  différente de  $\lambda_0$ . Comment expliquer cette différence ? [Effet Doppler dû au mouvement relatif entre les deux galaxies.]

**Question 2 :** Déterminer la vitesse relative entre la Terre et la galaxie d'Andromède en fonction de  $\lambda$  et  $\lambda_0$ . Faire l'application numérique pour  $\lambda = 21,03 \text{ cm}$ . [ $v = c(\frac{\lambda_0}{\lambda} - 1) = -428,0 \text{ km/s.}$ ]

**Question 3 :** Comment interpréter le signe de cette vitesse ? [Andromède s'approche de la Terre.]

**Question 4 :** La galaxie d'Andromède est à environ 2,5 millions d'années-lumière de la Terre. En considérant qu'elle se déplace à vitesse constante, dans combien de temps arrivera-t-elle à proximité de la Terre ? [ $t = d/v \simeq 1,8 \text{ milliards d'années.}$ ]

**Question 5 :** La vitesse déterminée à la question 2 est en fait une vitesse moyenne calculée sur l'ensemble de la galaxie d'Andromède, dont le diamètre  $d_A$  vaut environ 220000 années-lumière. La région qui s'approche le plus vite de la Terre se situe à une extrémité de l'ellipse sous laquelle Andromède nous apparaît. On mesure avec la méthode précédente une vitesse de  $v' = -1288 \text{ km/s.}$

→ **Niveau débutant :** À l'aide d'une soustraction entre deux vitesses, calculer la vitesse d'un point du bord d'Andromède par rapport au centre. Une révolution d'Andromède correspond à un tour complet de ce point. Quelle distance doit-il parcourir en années-lumière ? Sachant qu'une année-lumière est la distance parcourue par la lumière en une année, convertir cette distance en km. En utilisant la vitesse et la distance obtenues, déterminer la période de révolution  $T_A$  de la galaxie d'Andromède.

→ **Niveau intermédiaire :** Déterminer la vitesse d'un point du bord d'Andromède par rapport à son centre. Quelle distance doit-il parcourir pour en faire le tour ? En déduire la période de révolution  $T_A$  de la galaxie d'Andromède.

→ **Niveau expert :** Estimer la période de révolution  $T_A$  de la galaxie d'Andromède.  
[On calcule le périmètre de la galaxie et on divise par la vitesse d'un point du bord de la galaxie.]

## 4 - Travaux de recherche

### La première lumière de l'Univers

Après le Big Bang, l'Univers était tellement chaud que les particules élémentaires formaient une sorte de soupe, incapables de se lier les unes aux autres. En effet, l'énergie d'agitation thermique était supérieure à l'énergie de liaison de l'atome d'hydrogène, et les atomes étaient donc complètement ionisés. Ces particules chargées empêchaient la lumière de voyager sur de grandes distances, et l'Univers était donc opaque. Suite à son expansion, la température de l'Univers diminue progressivement jusqu'à atteindre, environ 380 000 ans après le Big Bang, une température suffisamment basse pour que les particules puissent se lier entre elles et former des atomes. L'Univers devient donc transparent et la lumière peut enfin s'y propager librement. Après avoir été fortement décalée vers le rouge par l'expansion de l'Univers, cette première lumière est visible aujourd'hui dans le domaine des micro-ondes sous le nom de Fond Diffus Cosmologique (CMB pour *Cosmic Microwave Background*).

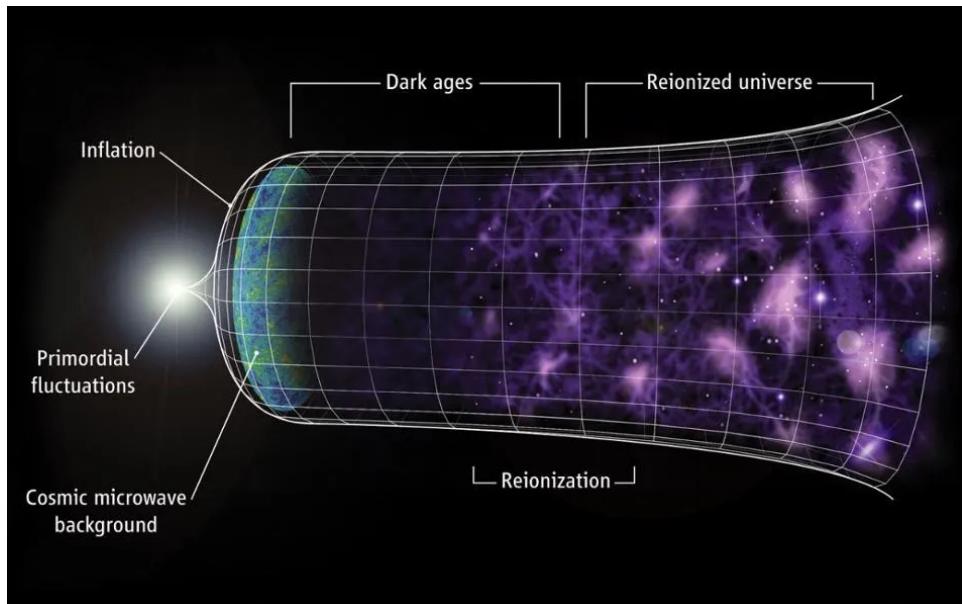


FIGURE 1 – Schéma de l'évolution de l'Univers, du Big Bang à aujourd'hui. La phase d'inflation est représentée par un brutal élargissement de la structure de l'Univers. Figure extraite de [Faucher-Giguère, 2008].

Le CMB a un spectre de corps noir à une température extrêmement uniforme sur le ciel de 2,75 K. Cette uniformité suggère qu'à un certain moment, toutes les régions de l'Univers étaient suffisamment proches pour qu'elles puissent interagir. Une explication possible est proposée par les modèles dits d'inflation, pour lesquels l'Univers a connu une phase d'expansion accélérée juste après le Big Bang, comme illustré sur la Figure 1. Le CMB n'est en réalité pas parfaitement uniforme, mais présente de minuscules anisotropies d'environ 100  $\mu\text{K}$ . Ces fluctuations de température, prévues par les modèles d'inflation, ont été observées par le satellite *Planck* et sont présentées sur la Figure 2. Mais le CMB est un signal polarisé, et les modèles d'inflation affirment que l'expansion rapide de l'Univers devrait avoir laissé une marque sur la polarisation du CMB. Cette signature devrait apparaître sous forme de motifs particuliers de polarisation rectiligne appelés modes E et B. Attention, ces modes sont nommés ainsi en raison de similitudes avec les comportements de champs électrique et magnétique mais n'y sont pas directement liés. Si la température et les modes E de polarisation du CMB ont aujourd'hui été mesurés avec une très bonne précision, les modes B sont pour l'instant toujours hypothétiques. Et ce n'est pas étonnant : on s'attend à ce qu'ils soient d'amplitude très faible, au moins 100 fois plus faible que les modes E ! La mesure de leur amplitude permettrait de valider et contraindre les modèles d'inflation, ce qui constitue un enjeu majeur de la cosmologie moderne. Leur détection est donc à la fois un enjeu théorique pour la compréhension de l'histoire de l'Univers mais aussi un immense défi technique et d'analyse de données.

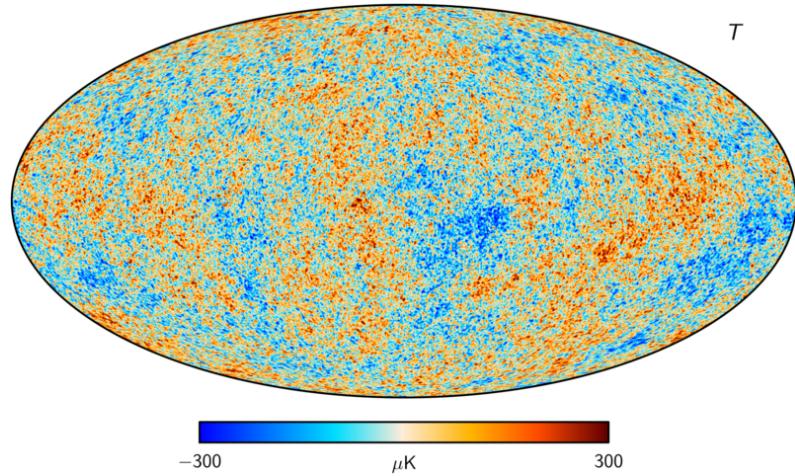


FIGURE 2 – Anisotropies de la température du CMB sur le ciel. Figure extraite de [Planck Collaboration, 2020].

## Les défis de l'analyse des données CMB

Pourquoi les télescopes n'ont-ils toujours pas, même avec une moins bonne précision que les modes E, détecté les modes B du CMB ? La réponse se trouve dans le grand drame des observations astrophysiques : quand on pointe un télescope dans une direction, on n'observe pas uniquement ce qu'on a envie d'observer mais aussi tous les signaux émis par notre environnement galactique. Ces signaux, de différentes origines, apparaissent au premier plan dans les observations, et sont ainsi qualifiés d'avant-plans.

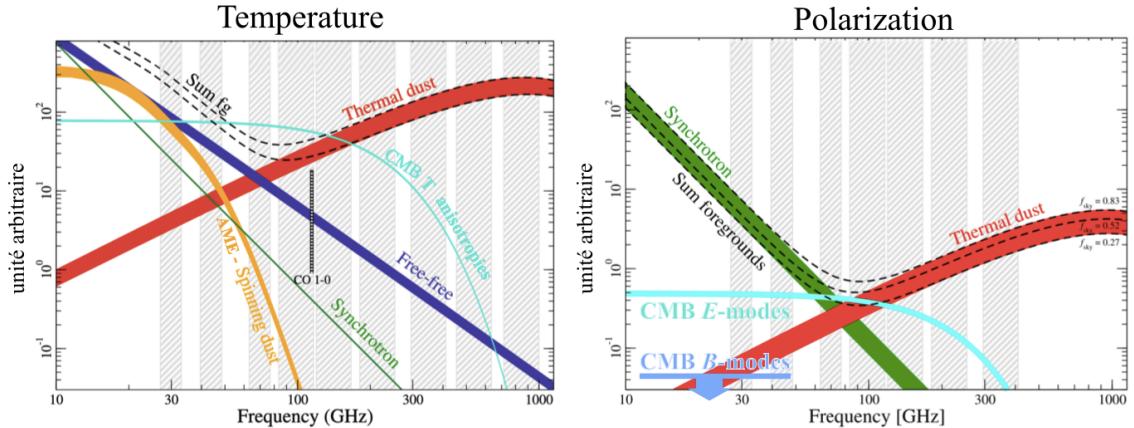


FIGURE 3 – La brillance des différents avant-plans du CMB en fonction de la fréquence d'observation, en température (à gauche) et en polarisation (à droite). Figure empruntée à Jonathan Aumont, adaptée de [Planck Collaboration, 2020].

La Figure 3 présente l'intensité relative des différents avant-plans du CMB en fonction de la fréquence, à la fois pour la température (à gauche) et pour la polarisation (à droite). Les courbes noires pointillées *Sum fg* représentent la somme de tous les avant-plans. Pour la température, on voit que les anisotropies du CMB (en cyan) dominent les avant-plans autour de 70 à 100 GHz. Cette bande de fréquence où le CMB est accessible a permis de construire des cartes comme celle présentée en Figure 2. De la même façon, les modes E de la polarisation (en cyan) ont une intensité presque similaire à celle des avant-plans entre 70 et 100 GHz. En utilisant les données obtenues dans des régions du ciel où les avant-plans galactiques sont moins intenses, les modes E du CMB ont donc pu être mesurés avec une bonne précision. Par contre, comme illustré sur la figure, les modes B (en bleu ciel) sont largement dominés par les avant-plans à toutes les fréquences. Leur détection nécessite donc l'utilisation de méthodes de traitement de données que l'on désignera par la suite par « séparation de composantes ».

Les séparations de composantes visent à caractériser un ou plusieurs signaux d'intérêt à partir d'un mélange, ce qui est nécessaire dans de nombreux cadres expérimentaux. C'est particulièrement le cas pour les observations astrophysiques, qui capturent généralement plusieurs émissions provenant de différents processus physiques ou régions, ainsi que des effets instrumentaux. Le développement de méthodes de séparation des composantes est souvent une étape cruciale dans l'analyse des données astrophysiques. La Figure 4 présente une carte de l'intensité du ciel entier dans le domaine des micro-ondes, observée par le satellite *Planck*. Cette carte illustre bien le problème de la séparation de composantes : on y voit à la fois l'émission de la poussière Galactique, représentée en bleu, très intense le long du grand axe de l'ellipse, mais on devine aussi le CMB, en rouge, dans les régions où l'émission de la poussière est faible.

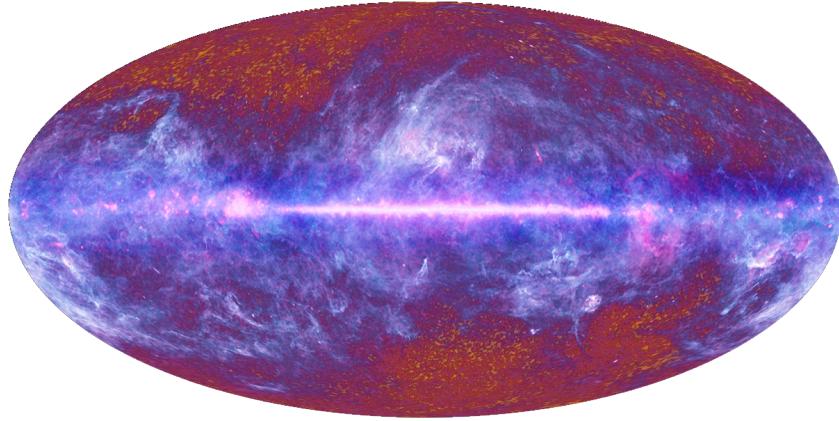


FIGURE 4 – Notre ciel dans le domaine des micro-ondes, vu par le satellite *Planck*.

En fait, des méthodes de séparation de composantes ont déjà été utilisées dans le contexte de l'analyse des données *Planck* pour cartographier la température et les modes E du CMB. Elles ont par exemple permis d'obtenir la carte présentée en Figure 2 à partir des données de la Figure 4. Certaines de ces méthodes reposent sur la modélisation des avant-plans : si l'on est capable de bien les modéliser, on devrait pouvoir être capable de les soustraire et de détecter les signaux d'arrière-plan. Ce genre d'algorithme a bien fonctionné pour la température et les modes E du CMB, mais les modèles d'avant-plans se sont révélés trop peu réalistes pour la séparation d'un signal aussi faible que les modes B. La possible détection des modes B est donc devenu un défi de modélisation des avant-plans, et en particulier de l'émission thermique de la poussière. **Pendant ma thèse, j'ai construit des modèles statistiques de l'émission infrarouge de la poussière.**

## Modéliser l'émission de la poussière

La poussière de notre galaxie représente une part importante de ce qu'on appelle le milieu interstellaire. Ce milieu, comme son nom l'indique, désigne tout ce qui se trouve entre les étoiles, et est principalement composé de poussière et de gaz. Le milieu interstellaire est un milieu très complexe, structuré par les interactions non-linéaires entre de nombreux processus physiques comme la gravitation, la turbulence, le magnétisme, la radiation et les réactions chimiques. La modélisation du milieu interstellaire et en particulier de l'émission de la poussière est donc un véritable défi.

Le CMB suit dans une très bonne approximation une distribution gaussienne, ce qui signifie que chaque échelle du signal est indépendante statistiquement. Pour faire une modélisation statistique du CMB, il suffit donc de connaître la puissance de ce signal à chaque échelle spatiale, c'est ce qu'on appelle le spectre de puissance. Au contraire, à cause de la physique très non-linéaire du milieu interstellaire, l'émission de la poussière est un signal fortement non-gaussien qui comporte de forts couplages entre échelles spatiales. Sa modélisation statistique ne peut donc être basée uniquement sur le spectre de puissance et nécessite des outils qui vont au-delà en caractérisant les couplages entre échelles.

Ces dernières années, des algorithmes d'intelligence artificielle qualifiés de « réseaux de neurones » ont été appliqués à la modélisation de champs non-gaussiens, fournissant des résultats d'une impressionante qualité. Cependant, ces algorithmes ont besoin d'être entraînés et leur efficacité dépend fortement du jeu de données utilisé. Appliquer des réseaux de neurones à des problèmes d'astrophysique est donc très difficile, puisque la quantité de données disponible est généralement très limitée, à la fois par la taille du ciel, les effets instrumentaux, et la non-stationnarité des phénomènes observés. La modélisation de l'émission de la poussière se présente donc comme un domaine trop complexe pour les outils statistiques usuels et trop limité en quantité de données pour bénéficier de la puissance des réseaux de neurones.

Récemment, de nouveaux outils statistiques ont été développés dans le domaine de la science des données dans le but de caractériser les champs non-gaussiens. Regroupés sous le terme de *Scattering Transforms*, ces outils sont des fonctions mathématiques inspirées des premières couches des réseaux de neurones mais ne nécessitant pas de phase d'entraînement. Dans les dernières années, les *Scattering Transforms* ont été appliquées avec succès à des données astrophysiques et cosmologiques pour des tâches de classification, d'inférence de paramètre et de modélisation. Par la suite, on utilisera les *Scattering Transforms* sans entrer dans les détails de leur calcul.

## Construire un modèle avec les Scattering Transforms

Les *Scattering Transforms* permettent de construire des modèles génératifs à partir de très peu de données, jusqu'à une seule image. Autrement dit, connaître la valeur des *Scattering Transforms* d'une carte permet de générer des nouvelles cartes qui ont une structure très similaire. Le processus est le suivant : considérons une carte  $x$  dont on souhaite construire un modèle statistique génératif. La construction du modèle en tant que tel tient en fait en une seule étape, qui consiste à calculer les *Scattering Transforms* de  $x$ , que l'on note  $\Phi(x)$ . Pour générer une nouvelle réalisation du modèle, il faut générer une carte  $\tilde{x}$  qui possède les mêmes *Scattering Transforms* que  $x$ , autrement dit telle que  $\Phi(\tilde{x}) \simeq \Phi(x)$ . Pour l'obtenir, on commence par générer un bruit blanc  $u$  de la même taille que  $x$ . Ensuite, on déforme les pixels de  $u$  de façon à ce que  $\Phi(u) \rightarrow \Phi(x)$ . Finalement, on obtient une nouvelle réalisation du modèle  $\tilde{x} = u$ . Concrètement, ce processus est une descente de gradient sous contrainte statistique, c'est-à-dire qu'on modifie itérativement les pixels de  $u$  de façon à minimiser la fonction de coût  $\mathcal{L}(u) = \|\Phi(u) - \Phi(x)\|^2$ .

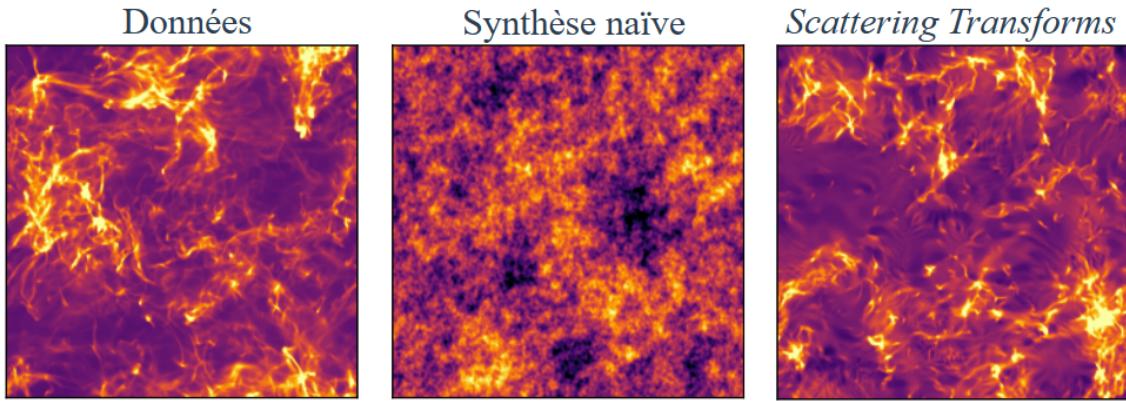


FIGURE 5 – Comparaison entre des modèles statistiques d'une simulation du milieu interstellaire construits à partir de différents outils statistiques. De gauche à droite : simulation numérique du milieu interstellaire (Données), réalisation d'un modèle basé sur le spectre de puissance (Synthèse naïve), réalisation d'un modèle basé sur les *Scattering Transforms*.

Pour illustrer le fait que les *Scattering Transforms* sont des statistiques adaptées pour modéliser la structure non-linéaire de l'émission de la poussière, la Figure 5 compare des réalisations de modèles génératifs d'une simulation numérique du milieu interstellaire basés sur différentes statistiques. Tout d'abord, on constate que la simulation du milieu interstellaire présente une structure complexe, filamentaire, avec des régions très intenses mais aussi de

grandes zones sombres. Ces structures très définies sont le signe de forts couplages entre échelles qui traduisent le caractère très non-gaussien de ce processus. La réalisation obtenue à partir du modèle basé sur le spectre de puissance a perdu toute la structure filamentaire, ne gardant que de grossières variations d'intensité. À l'inverse, la réalisation obtenue à partir du modèle basé sur les *Scattering Transforms* présente une structure fortement filamentaire, beaucoup plus ressemblante à celle de la simulation numérique du milieu interstellaire. Toutes ces considérations visuelles ont évidemment été accompagnées d'analyses statistiques quantitatives mais il apparaît que le critère visuel reste très efficace dans la comparaison de processus non-gaussiens. Les *Scattering Transforms* semblent donc particulièrement adaptées pour construire des modèles statistiques du milieu interstellaire et en particulier de l'émission thermique de la poussière.

Il suffirait donc d'appliquer les *Scattering Transforms* à des observations de l'émission thermique de la poussière pour en construire des modèles. Cependant, comme évoqué dans les sections précédentes, les télescopes n'observent jamais uniquement ce qu'on a envie qu'ils observent. Les observations infrarouges du milieu interstellaire contiennent en fait l'émission thermique de la poussière mais aussi un signal cosmologique, le fond diffus infrarouge (CIB pour *Cosmic Infrared Background*), qui correspond à l'émission infrarouge cumulée des galaxies sur toute l'histoire de l'Univers. L'utilisation des observations pour construire des modèles statistiques basés sur les *Scattering Transforms* nécessite donc de séparer l'émission de la poussière de celle du CIB. Le problème a donc été inversé : l'émission de la poussière nous empêche d'observer le CMB dans le domaine des micro-ondes, on veut donc modéliser cette émission, ce qui nécessite des observations infrarouges, dans lesquelles l'émission de la poussière est contaminée par le CIB.

## Séparation de composantes

La séparation entre l'émission de la poussière et le CIB est un enjeu important pour l'astrophysique Galactique puisque ce dernier contamine toutes les observations infrarouges, biaisant fortement les analyses de la poussière à petite échelle. Puisque le CIB a comme origine l'émission de la poussière des galaxies, sa distribution spectrale d'énergie est très proche de celle de l'émission de la poussière de notre Galaxie. Cela empêche l'utilisation des méthodes de séparation de composantes usuelles en astrophysique basées sur des différences de distributions spectrales d'énergie. Ces deux signaux, bien que spectralement très proches, présentent cependant une structure apparente sur le ciel très différente. En effet, alors que l'on observe la poussière Galactique de façon très détaillée, les structures qui apparaissent dans le CIB résultent de groupes de galaxies non résolues individuellement. L'originalité de ma thèse est d'utiliser la structure sur le ciel comme bras de levier pour séparer deux signaux astrophysique et cosmologique. Cela est rendu possible par la capacité des *Scattering Transforms* à différencier quantitativement les signatures statistiques de l'émission de la poussière Galactique et du CIB.

La méthode de séparation de composantes repose sur un algorithme similaire à celui utilisé pour les modèles génératifs. Considérons une observation astrophysique  $d$  dans laquelle sont présents un signal de poussière  $s$  et le CIB  $c$ . On a alors, dans l'espace des pixels,  $d = s + c$ . Dans la section sur les modèles statistiques, on a vu que la connaissance des *Scattering Transforms* d'un processus suffisait pour en construire un modèle statistique. L'objectif de l'algorithme est de déterminer  $\Phi(s)$  en connaissant  $\Phi(d)$  et  $\Phi(c)$ . Les *Scattering Transforms* étant par essence des statistiques non-linéaires,  $\Phi(s) = \Phi(d - c) \neq \Phi(d) - \Phi(c)$ . Nous allons donc, là encore, faire une optimisation dans l'espace des pixels de façon à déterminer une carte qui minimise une fonction de coût bien choisie.

On considère une carte variable  $u$  que l'on aimerait faire converger vers notre estimation du signal de poussière. On sait que si on ajoutait à  $u$  une carte de CIB, il faudrait que la somme ressemble à l'observation  $d$ . Si c'est le cas, cela veut dire que notre carte  $u$  est proche du signal de poussière  $s$ . La première contrainte que l'on pourrait mettre sur  $u$  est donc  $\Phi(d) \simeq \Phi(u + c)$ . Mais que représente exactement  $c$  dans cette équation ? Nous avons supposé connaître  $\Phi(c)$ , mais on ne connaît a priori pas  $c$ , sinon il suffirait de le soustraire à  $d$  pour trouver  $s$ . Par contre, on peut utiliser  $\Phi(c)$  pour tirer une réalisation du modèle de  $c$ , que l'on note  $c_0$ . On peut donc imposer  $\Phi(d) \simeq \Phi(u + c_0)$ . Le problème n'est en fait pas encore résolu : avec cette contrainte, la carte  $u$  va être déformée

de façon à absorber précisément les structures de  $c_0$  de façon à ce que la somme ressemble à  $d$ . On illustre ici la difficulté de cet algorithme, qui doit en fait sommer des processus statistiques dans l'espace des pixels mais qui ne doit pas dépendre de la réalisation du processus que l'on utilise (ici,  $c_0$ ). La solution mise en œuvre consiste à moyenner cette contrainte sur un grand nombre de réalisations  $c_i$ , de façon à ce que l'évolution de  $u$  ne dépende pas d'une réalisation en particulier. La contrainte utilisée peut donc s'écrire sous la forme d'une fonction de coût

$$\mathcal{L}_1(u) = \langle \|\Phi(d) - \Phi(u + c_i)\|^2 \rangle_i \quad (1)$$

où  $\langle \cdot \rangle_i$  désigne la moyenne sur  $i$ . Pour minimiser cette fonction, la carte  $u$  doit, quand on lui ajoute n'importe quelle carte dont la structure est celle du CIB, être proche de  $d$  en termes de *Scattering Transforms*. Une autre contrainte que doit vérifier  $u$  est que si on la soustrait à  $d$ , la différence doit avoir la structure du CIB. On va alors imposer que  $\Phi(d - u) \simeq \Phi(c)$ . Cette fois-ci, on ne connaît pas uniquement  $\Phi(d)$  mais aussi  $d$  puisqu'il s'agit de l'observation que l'on utilise. Cette contrainte peut aussi s'écrire sous la forme d'une deuxième fonction de coût

$$\mathcal{L}_2(u) = \|\Phi(d - u) - \Phi(c)\|^2 \quad (2)$$

qui nous permet de définir la fonction de coût totale

$$\mathcal{L}(u) = \mathcal{L}_1(u) + \mathcal{L}_2(u) \quad (3)$$

que l'algorithme a pour but de minimiser. Dans l'optimisation, la carte  $u$  est initialisée à  $d$  de façon à hériter des grandes structures de l'émission de la poussière qui ne sont que très peu affectées par le CIB.

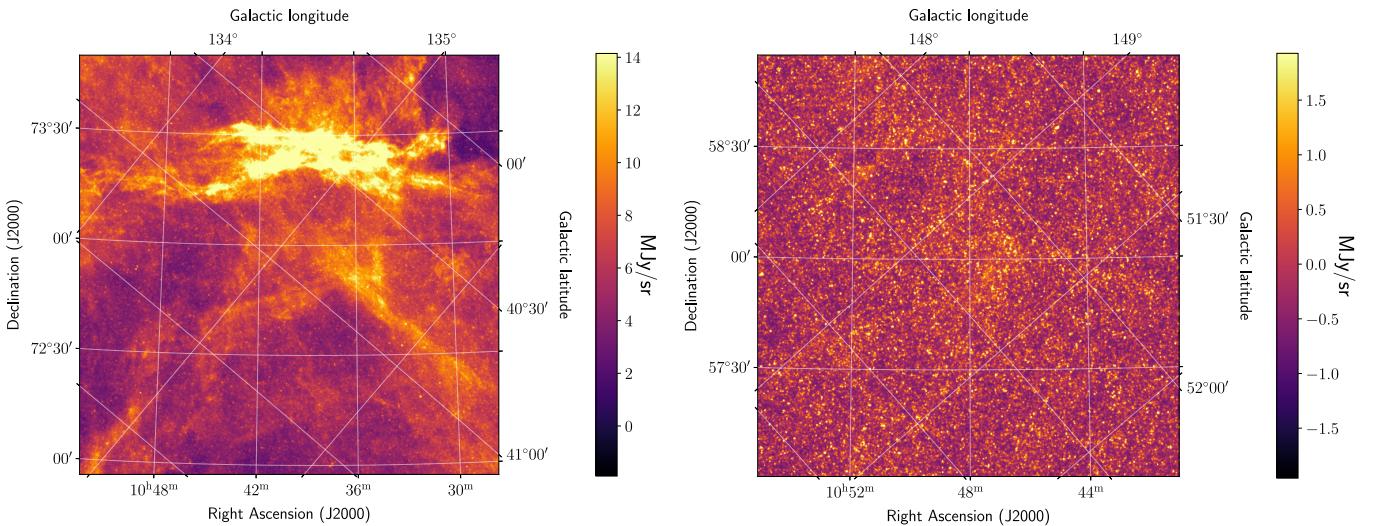


FIGURE 6 – Observations par le satellite *Herschel* à une longueur d'onde  $\lambda = 250 \mu\text{m}$  de deux régions : Spider (à gauche) et Lockman Hole (à droite).

J'ai appliqué cet algorithme à une observation infrarouge du satellite *Herschel* présentée en Figure 6. Sur cette observation  $d$ , on distingue à la fois l'émission de la poussière Galactique  $s$ , qui forme des structures brillantes à assez grande échelle, mais aussi le CIB  $c$  dont la structure à petite échelle s'apparente à un bruit de grenade. Pour obtenir  $\Phi(c)$ , j'utilise une observation d'une autre région, appelée le Lockman Hole, présentée en Figure 6, dans laquelle on sait que l'émission de la poussière est extrêmement faible. Sous hypothèse d'isotropie des statistiques du CIB sur le ciel, qui semble raisonnable en tant que signal cosmologique, on considère que les *Scattering Transforms* de l'observation du Lockman Hole sont une bonne approximation de  $\Phi(c)$ . Toutes les conditions sont donc réunies pour appliquer l'algorithme de séparation de composantes à ces observations. L'algorithme est dans un premier temps appliqué sur de fausses données pour lesquelles on connaît  $\Phi(s)$  et validé quantitativement à l'aide de différentes statistiques non contraintes dans l'optimisation. Une fois validé, il est appliqué aux observations *Herschel*.

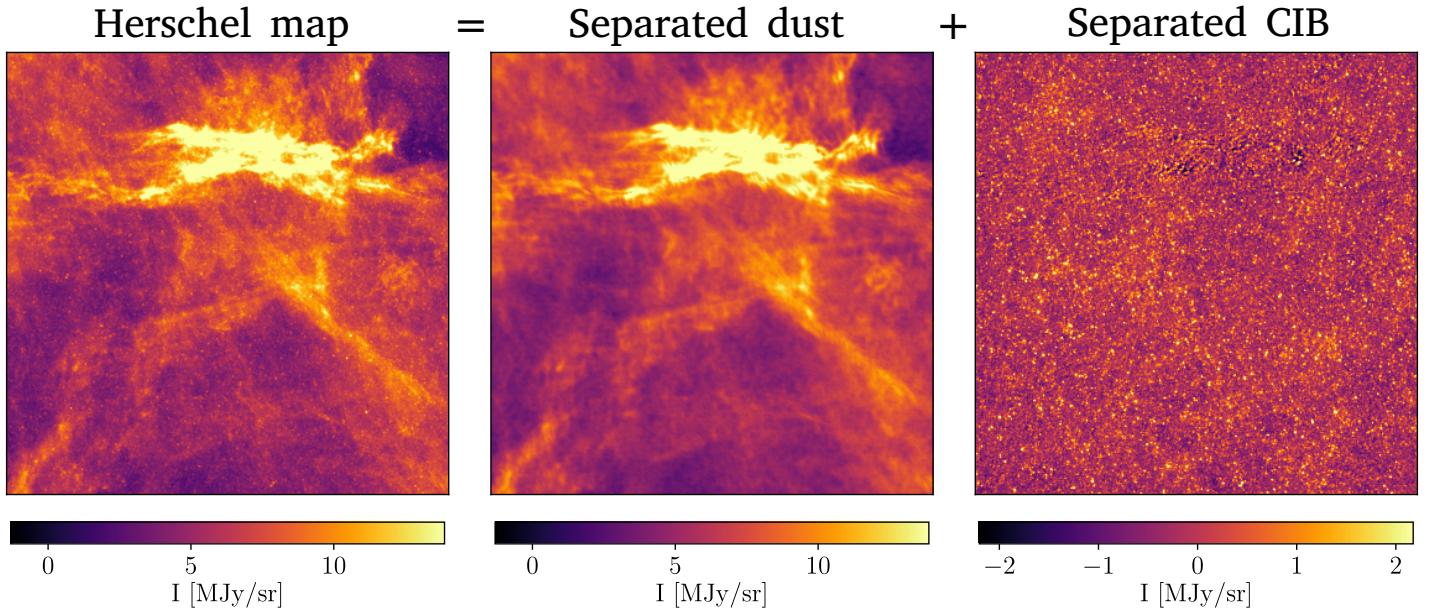


FIGURE 7 – Séparation de composantes entre l'émission de la poussière  $\tilde{s}$  (au milieu) et le CIB  $\tilde{c}$  (à droite) à partir de l'observation *Herschel d* de la région Spider (à gauche).

Même si l'objectif de l'algorithme est de déterminer  $\Phi(s)$ , il produit en fait une carte  $\tilde{s}$  dont les *Scattering Transforms* constituent notre estimation de  $\Phi(s)$ . La Figure 7 présente les résultats de la séparation de composantes au niveau des cartes. L'observation *Herschel d* est décomposée en une carte de l'émission de la poussière  $\tilde{s}$  et une carte du CIB  $\tilde{c} = d - \tilde{s}$ . On voit que le CIB présent dans  $d$  a été séparé et n'est plus visible dans  $\tilde{s}$ . De plus,  $\tilde{c}$  présente bien une structure similaire à celle de l'observation du Lockman Hole présentée en Figure 6. Ces affirmations visuelles ont bien sûr été accompagnées d'analyses statistiques quantitatives qui ne sont pas détaillées ici. **La méthode de séparation de composantes développée permet donc de créer des modèles statistiques de l'émission infrarouge de la poussière Galactique uniquement à partir d'observations.**

## 5 - Conclusion

Finalement, ce travail de recherche m'a permis de développer de nombreuses compétences qui seront utiles à un professeur de physique-chimie. Tout d'abord, la confrontation à un défi scientifique actuel m'a appris la rigueur, que je tiens à inculquer à mes futurs élèves. Je m'attacherais à donner une place importante au sein de mes cours à des activités de recherche et à l'analyse de l'actualité scientifique, par exemple avec les nombreuses découvertes que nous réserve le *James Webb Space Telescope* dans les prochaines années. Parallèlement, la nature bidisciplinaire de ma thèse, partagée entre l'astrophysique et la science des données, m'a appris à expliquer des notions complexes d'une discipline à des spécialistes de l'autre. Les capacités pédagogiques que j'ai dû développer me permettront d'intéresser les élèves aux sujets scientifiques variés qui seront étudiés en classe. Enfin, les interactions scientifiques et humaines au sein de mon laboratoire m'ont enseigné l'ouverture d'esprit et la valeur du travail d'équipe, qui sont cruciales à la fois au sein d'une salle de classe et d'une équipe pédagogique.