
Les grands précurseurs IRAS et ISO

Le satellite IRAS (satellite astronomique infrarouge²⁹, Neugebauer *et al.* (1984)), construit par une collaboration entre les Etats-unis, le Royaume Uni et les Pays Bas, a été lancé en 1983 par la NASA³⁰. Ce télescope infrarouge d'un diamètre de 57 cm a effectué un relevé de 96% du ciel à 12, 25, 60 et 100 μm . Il était équipé d'un système de refroidissement à l'hélium superfluide (1,6 K) qui a fonctionné 10 mois. Ce télescope a permis de découvrir environ 75 000 galaxies infrarouges, mais aussi les cirrus galactiques (nuages de gaz diffus dans notre galaxie émettant dans l'infrarouge).

Le satellite ISO (Observatoire Spatial Infrarouge³¹, Kessler *et al.* (1996)), lancé en 1995, possédait un miroir de 70 cm par l'ESA³², était refroidi à l'hélium liquide, et observait entre 2 et 200 μm . Contrairement à IRAS qui observait tout le ciel, ISO réalisait des observations pointées. Il a permis de résoudre la majorité du fond infrarouge en sources individuelles à 15 μm (Elbaz *et al.* (1999)). En revanche, à cause de sa résolution angulaire limitée, il ne résolvait qu'une faible part du fond à plus grande longueur d'onde (Dole *et al.* (2001), Héraudeau *et al.* (2004)).

1.4.3 Spitzer

Spitzer (Werner *et al.* (2004)) est un télescope de la NASA, lancé en 2003. Il est placé en orbite héliocentrique avec un demi-grand-axe très légèrement supérieur à celui de la Terre, dont il s'éloigne lentement. Son miroir primaire, mesurant 85 cm de diamètre, est refroidi à 5.5 K. Il observe entre 3.6 et 160 μm , grâce aux instruments IRAC, IRS et MIPS. Le système de refroidissement s'est arrêté en 2009 rendant les observations impossibles au delà de 5 μm (voir Fig. 1.11).

L'instrument IRAC³³ (Fazio *et al.* (2004b)) est une caméra infrarouge observant à 3.6, 4.5, 5.8 and 8.0 μm . Chaque longueur d'onde est associée à un détecteur de 256×256 pixels, chaque pixel observant un angle de 1.2" sur le ciel. Les 4 canaux sont observés de manière simultanée grâce à des séparateurs dichroïques. La PSF (Fonction d'étalement de point³⁴) a une largeur à mi-hauteur de 1.66", 1.72", 1.88" et 1.98" à 3.6, 4.5, 5.8 et 8.0 μm , respectivement.

L'instrument MIPS³⁵ (Rieke *et al.* (2004)) est un photomètre imageur observant à 24,

²⁹En anglais, *InfraRed Astronomical Satellite*

³⁰*National Aeronautics and Space Administration.*

³¹*Infrared Space Observatory*

³²*European Space Agency.*

³³*InfraRed Array Camera*

³⁴*Point Spread Function*

³⁵*Multiband Imaging Photometer for Spitzer*

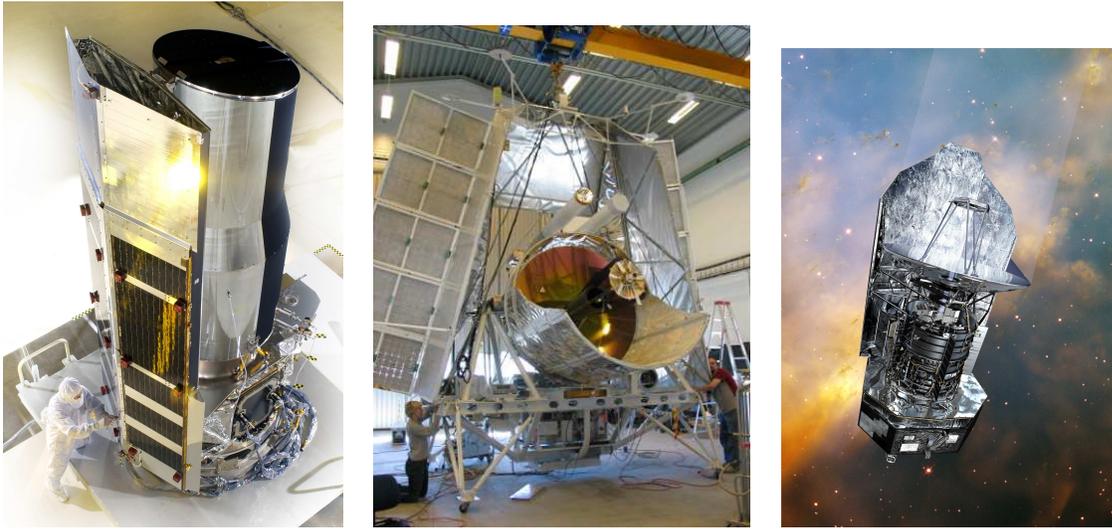


FIGURE 1.12 – De gauche à droite : le télescope spatial *Spitzer* (Crédit : NASA), le télescope porté par ballon BLAST (Crédit : Gaelen Marsden), le télescope spatial *Herschel* (Crédit : ESA).

70 et 160 μm . Les détecteurs possèdent 128 \times 128 pixels de 2.5" projetés sur le ciel, 16 \times 32 pixels de 9.8", et 2 \times 20 pixels de 16", à 24, 70 et 160 μm , respectivement. La PSF a une largeur à mi-hauteur de, respectivement, 6.6", 18" et 40".

L'instrument IRS³⁶ Houck *et al.* (2004) est un spectrographe possédant deux modules de basse résolution ($R\sim 100$) et deux modules de haute résolution ($R\sim 600$). L'instrument possède également deux matrices permettant de réaliser des images à 15 et 22 μm , initialement prévues pour le centrage des sources. Elles ont également servi à réaliser des relevés photométriques à 15 μm (Teplitz *et al.* (2010)).

1.4.4 BLAST

L'expérience BLAST³⁷ (Pascale *et al.* (2008)) a pour but d'observer le ciel submillimétrique à 250, 350 et 500 μm . Il s'agit d'un télescope de 2 m de diamètre porté par un ballon stratosphérique. Il a volé en 2005 en Suède et en 2006 en Antarctique. Les détecteurs sont des matrices de bolomètres *spider web*³⁸ refroidis à l'hélium liquide, précurseurs de l'instrument SPIRE de *Herschel*. Sa PSF a une largeur à mi-hauteur de 36", 42" et 60" à 250, 350 et 500 μm , respectivement.

³⁶*InfraRed Spectrometer*

³⁷Balloon-borne Large-Aperture Submillimeter Telescope

³⁸Toile d'araignée

1.4.5 *Herschel*

Le télescope spatial européen *Herschel* (Pilbratt *et al.* (2010)) a été lancé avec succès en 2009 par l'ESA en même temps que *Planck* (voir section suivante). Il s'agit d'un télescope de 3.5 m de diamètre, placé au point de Lagrange 2³⁹. Grâce à une orientation dos au Soleil et la Terre, son miroir est refroidi passivement à 60 K. Il est équipé de 3 instruments : PACS, SPIRE et HIFI.

L'instrument PACS⁴⁰ (Poglitsch *et al.* (2010)) est capable de faire de la photométrie en bande large et de la spectroscopie basse résolution entre 55 et 210 μm . Ses bandes photométriques sont placées à 70, 100 et 160 μm . La largeur à mi-hauteur de la PSF est respectivement 5.2", 7.7" et 12". Une matrice de 32×16 bolomètres permet d'observer à 160 μm , en parallèle d'une autre matrice de 64×32 observant soit à 70, soit à 100 μm . La sensibilité de PACS est limitée par le bruit statistique des photons émis par le miroir.

L'instrument SPIRE⁴¹ (Griffin *et al.* (2010)) a les mêmes fonctions que SPIRE, mais dans le domaine sub-millimétrique de 200 à 650 μm . Les canaux du photomètre sont centrés autour de 250, 350 et 500 μm . Ces canaux sont équipés de matrices de bolomètres de respectivement 139, 88 et 43 éléments. Des bolomètres aveugles permettent de corriger les dérives en température de chaque matrice. La largeur à mi-hauteur de la PSF est, respectivement, 18", 25" et 37". Avec SPIRE, les observations de champs profonds sont très vite limitées par le phénomène de confusion dû à la faible résolution angulaire du télescope.

Herschel est également équipé du spectromètre hétérodyne de très haute résolution ($R \sim 10^7$) HIFI⁴² (de Graauw *et al.* (2010)). Malheureusement, à cause de sa sensibilité limitée, cet instrument ne peut observer que des objets très brillants et donc très proches.

³⁹Le point de Lagrange 2 est un point d'équilibre du système Terre-Soleil. Il est située à 1.5 millions de kilomètres de la Terre dans la direction opposée à celle du Soleil. Ainsi, le satellite voit le Soleil et la Terre alignés.

⁴⁰*Photoconductor Array Camera and Spectrometer*

⁴¹*Spectral and Photometric Imaging Receiver*

⁴²*Heterodyne Instrument for the Far-Infrared*

1.4.6 Les nouveaux observatoires spatiaux tout-le-ciel : AKARI, WISE, *Planck*

Des relevés tout-le-ciel ont également été réalisés dans l'infrarouge. En général, leur sensibilité ne permet pas de résoudre une partie significative du fond infrarouge. Ces observations sont en revanche cruciales pour les études sur l'Univers local et les fluctuations du fond.

Le satellite japonais *Akari*⁴³ (Ishihara *et al.* (2010)) a été lancé en 2006. Il a réalisé un relevé de tout le ciel dans de nombreuses bandes photométriques entre 6 et 180 μm . Des relevés profonds ont également été réalisés. Le satellite WISE⁴⁴, lancé par la NASA en décembre 2009, a réalisé également un relevé du ciel complet à 3.3, 4.7, 12 et 23 μm . Sa mission s'est achevée en septembre 2010, lorsque l'hydrogène solide du système de refroidissement a été entièrement consommé. Sa sensibilité est très supérieure à celle d'*Akari*.

Le satellite *Planck* (Tauber *et al.* (2010)), lancé en même temps que *Herschel*, réalise un relevé de tout le ciel du domaine sub-millimétrique au domaine radio. L'instrument HFI⁴⁵ (Lamarre *et al.* (2010)) observe à 350, 550, 840, 1380, 2100 et 3000 μm . Sa résolution inférieure à 3' ne permet de résoudre que les objets les plus brillants. En revanche, sa grande stabilité va permettre d'étudier avec une précision inégalée les fluctuations du fond infrarouge.

1.4.7 Télescopes sub-millimétriques et millimétriques au sol : JCMT, IRAM, APEX, SPT...

Au delà de 450 μm , il est possible d'observer depuis le sol dans les principales fenêtres atmosphériques (450 μm , 850 μm , 1.2 mm...). Les télescopes au sol ont de plus grands diamètres que leur homologue spatiaux. Néanmoins, la transmission atmosphérique est faible, et il est préférable d'observer dans des endroits secs en altitude. On peut citer entre autre le JCMT⁴⁶ à Hawaï, l'antenne de 30 m de l'IRAM⁴⁷ près de Grenade, APEX⁴⁸ et ACT⁴⁹ dans le désert de l'Atacama au Chili, et SPT⁵⁰ au Pôle sud.

⁴³Lumière en japonais

⁴⁴*Wide-Field Infrared Survey Explorer*

⁴⁵*High Frequency Instrument*

⁴⁶*James Clerk Maxwell Telescope*

⁴⁷Institut de RAdioastronomie Millimétrique

⁴⁸*Atacama Pathfinder Experiment*

⁴⁹*Atacama Cosmology Telescope*

⁵⁰*South Pole Telescope*

Chapitre 2

Comptages de sources résolues

Sommaire

2.1	Introduction	42
2.1.1	Pourquoi compter les sources infrarouges ?	42
2.1.2	Interprétation des comptages	42
2.2	Relevés	47
2.2.1	Choix des champs observés	47
2.2.2	Observations	47
2.2.3	Construction des cartes infrarouges	47
2.2.4	Champs profonds et champs larges	48
2.3	Extraction de sources	49
2.3.1	La problème de la confusion	49
2.3.2	Détection	50
2.3.3	Photométrie d'ouverture	53
2.3.4	Photométrie par ajustement de PSF	55
2.3.5	Photométrie et filtrage des cartes	58
2.4	Estimation des comptages	60
2.4.1	Validation de l'extraction par méthode Monte-Carlo	60
2.4.2	Complétude	60
2.4.3	Qualité de la photométrie	61
2.4.4	Estimation des comptages et correction des biais d'extractions	63
2.4.5	Estimation des incertitudes sur les comptages	63
2.5	Comptages MIPS	66

2.5.1	Motivations	67
2.5.2	Méthodes	67
2.5.3	Résultats	68
2.6	Comptages BLAST	72
2.6.1	Motivations	72
2.6.2	Extraction de sources aveugle	72
2.6.3	Extraction de sources avec a priori	73
2.6.4	Discussion	75
2.7	Conclusion	75

2.1 Introduction

2.1.1 Pourquoi compter les sources infrarouges ?

On appelle comptages la distribution en flux des sources dans le ciel. Cette observable peut être mesurée à partir d'une simple carte infrarouge. Elle ne nécessite donc pas de rassembler des jeux de données hétérogènes, tel des *redshifts* ou des mesures photométriques issues d'autres bandes spectrales. Les effets de sélections peuvent ainsi être aisément interprétés par les modèles. Il est néanmoins nécessaire de corriger les biais dus à l'extraction des sources. Ce qui peut-être fait avec une très bonne précision. C'est pour cette raison que les comptages sont une contrainte massivement utilisée par les modèles d'évolution des galaxies.

2.1.2 Interprétation des comptages

Comptages intégraux et comptages différentiels

Les comptages de galaxies peuvent être exprimés en utilisant deux conventions différentes : comptages intégraux et comptages différentiels. Les comptages intégraux (notés $N(>S)$) correspondent au nombre de sources plus brillantes qu'un flux donné par unité d'angle solide. En général, $N(>S)$ est calculé pour une série de flux S . Cette convention a l'avantage de permettre de déterminer très rapidement le nombre de sources qu'il est possible de détecter en réalisant un relevé de sensibilité donnée et de taille donnée. Le gros désavantage des comptages intégraux est le fait que les points de mesures sont fortement corrélés. De plus, $N(>S)$ peut varier de plusieurs ordres de grandeur sur un même ordre de grandeur en flux. Ceci peut être problématique lors du tracé des courbes, et en particulier pour réaliser une comparaison fine entre observation et modèle.

Les comptages différentiels (noté $\frac{dN}{dS}$), correspondent au nombre de sources par unité d'angle solide et par intervalle de flux. Dans ce cas, les corrélations entre points sont plus faibles¹. Il est également possible de tracer $dN/dS \times S^{2.5}$ en fonction de S dans les graphiques afin de limiter le nombre d'ordres de grandeur à représenter en ordonnées.

Comptages dans un Univers euclidien sans évolution

Dans un Univers euclidien, infini et sans évolution, les comptages peuvent se calculer analytiquement. On considère ici une seule population de galaxies de luminosité L ayant une densité ρ_0 . Le nombre de sources plus brillantes que S_{max} est

$$N(> S_{max}) = \rho_0 V_{max}, \quad (2.1)$$

avec V_{max} le volume dans lequel les sources de flux S_{max} sont détectées. Calculons ce volume. La distance maximale à laquelle des sources sont vus est :

$$D_{max} = \sqrt{\frac{L}{4\pi S_{max}}}, \quad (2.2)$$

et donc,

$$V_{max} = \frac{4\pi}{3} \left(\frac{L}{4\pi S_{max}} \right)^{\frac{3}{2}}, \quad (2.3)$$

d'où,

$$N(> S_{max}) = \frac{\rho_0 L^{\frac{3}{2}} 6 \sqrt{\pi}^{-\frac{3}{2}}}{S_{max}}. \quad (2.4)$$

Ce comportement en $S^{-\frac{3}{2}}$ reste vrai dans un Univers contenant des galaxies de différentes luminosités, à condition que la fonction de luminosité LF n'évolue pas.

Les comptages différentiels sont donc donnés par

$$\frac{dN}{dS} = \frac{\rho_0 L^{\frac{3}{2}}}{4 \sqrt{\pi}} S^{-\frac{5}{2}}. \quad (2.5)$$

Si on multiplie les comptages différentiels par $S^{\frac{5}{2}}$, on obtient donc une constante. On trace souvent $dN/dS \times S^{2.5}$ pour pouvoir mesurer les écarts par rapport aux comptages euclidiens. On parle alors de comptages normalisés.

¹Elles sont même nulles si on suppose que les sources sont distribuées de manière poissonnienne.

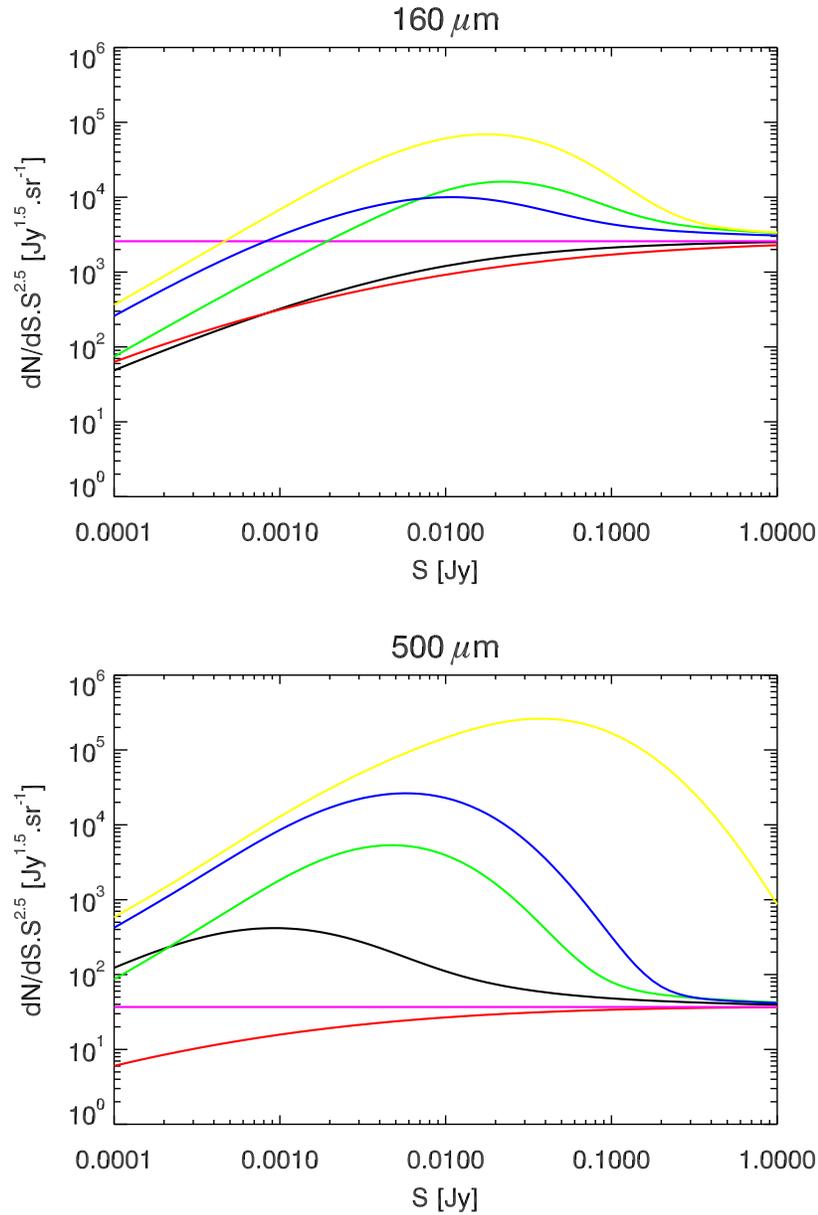


FIGURE 2.1 – Comptages différentiels normalisés à 160 (en haut) et 500 μm (en bas) prédits par le modèle Béthermin *et al.* (2011). Différentes situations sont représentées. En rouge : pas d'évolution, spectres des galaxies plats. En noir : pas d'évolution, spectres réalistes. En bleu : évolution de la luminosité des galaxies en $(1+z)^{2.5}$. En jaune : évolution de la luminosité des galaxies en $(1+z)^{3.5}$. En vert : évolution complexe en densité et luminosité utilisée par le modèle Béthermin *et al.* (2011). En violet : Univers euclidien sans évolution.

Comptages dans un Univers Λ CDM sans évolution

Le calcul précédent supposait que l'Univers était statique et sans expansion, mais, en réalité, l'Univers observable a un volume fini. La Fig. 2.1 montre les comptages à 160 et 500 μm prédits avec les outils présentés au Chap. 6 dans différents cas. La courbe violette correspond au cas du chapitre précédent, c'est à dire, un Univers euclidien, sans évolution. La courbe en rouge représente les comptages dans un Univers en expansion², mais où les spectres de galaxies sont supposés plats³. Les comptages de sources brillantes des deux modèles ($S > 1$ Jy) sont similaires. En effet, les sources brillantes se situent dans l'Univers proche, où les effets de l'expansion sont négligeables. En revanche, à bas flux, il y a moins de sources dans un Univers en expansion. Ce changement de pente des comptages à bas flux est obligatoire dans un Univers fini. En effet, si on intégrait des comptages euclidiens jusqu'à un flux nul, le fond total divergerait⁴.

La prise en compte de la SED réelle des sources infrarouges modifie les comptages à 500 μm , mais quasiment pas à 160 μm (voir courbe noire). Ceci s'explique par l'effet de K-correction, qui est le rapport entre la densité de luminosité d'un objet dans son référentiel propre et dans celui de l'observateur :

$$K(\nu, z) = \frac{L_{\nu(1+z)}}{L_{\nu}} \quad (2.6)$$

La Fig. 2.2 représente la K-correction pour différentes longueurs d'onde d'observation en fonction du *redshift*. Une K-correction grande devant 1 indique qu'une source est *boostée* par le décalage vers le rouge, c'est à dire qu'on la voit plus brillante qu'elle ne le serait dans un Univers euclidien à la même distance (lumineuse). La SED des galaxies infrarouges étant très pentue à grande longueur d'onde, la K-correction est donc très forte. Cet effet produit un excès dans les comptages autour de 10 mJy, qui correspond à des sources à haut *redshift*, vues proches de leur maximum d'émission. La K-correction étant faible à 160 μm , on n'y observe pas d'excès significatif.

Comptages dans un Univers Λ CDM avec évolution

On a jusqu'ici supposé que la distribution en luminosité des galaxies infrarouges ne changeait pas avec le *redshift*. Or, une évolution des populations infrarouges agit fortement

²On suppose une cosmologie Λ CDM avec les paramètres mesurés par WMAP 7 ans (Larson *et al.* (2010)).

³ $\nu L_{\nu} = \text{Constante}$

⁴Ceci est appelé paradoxe d'Olbers. En effet, au XIX^{me} siècle, les scientifiques pensaient que l'Univers était infini et sans évolution. Ceci conduisait au paradoxe d'un fond de rayonnement infini. On pensait alors que l'absorption des photons pouvait résoudre ce problème. En fait, la principale explication est bien sûr l'expansion de l'Univers et la finitude de la vitesse de la lumière (Wesson *et al.* (1987), Dole (2010)).

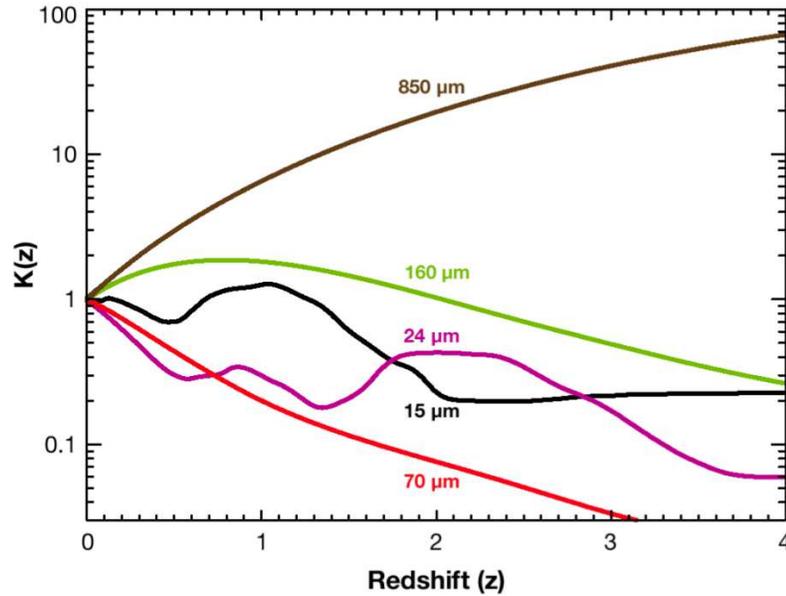


FIGURE 2.2 – K-correction pour différentes longueurs d’onde d’observation en fonction du *redshift* (extrait de Lagache *et al.* (2005)).

sur les comptages. On peut par exemple supposer une évolution simple où la luminosité de toutes les galaxies augmente comme $(1+z)^\alpha$. Le Fig. 2.1 représente les comptages pour une évolution en $(1+z)^{2.5}$ (en bleu) et en $(1+z)^{3.5}$ (en jaune). On constate que les comptages sont très sensibles à de telles évolutions. On remarque également que l’effet est assez différent en fonction de la longueur d’onde. En effet, à cause des effets de K-corrections, les comptages à $500 \mu\text{m}$ sont plus sensibles aux grands *redshift*, où l’impact d’une évolution en $(1+z)^\alpha$ est plus fort. Cet effet montre que l’utilisation de comptages à différentes longueurs d’ondes peut lever certaines dégénérescences dans les scénarios d’évolution des galaxies.

En pratique, une évolution aussi simple ne permet pas de reproduire simultanément les comptages infrarouges à plusieurs longueurs d’ondes. Une évolution de la luminosité et de la densité des galaxies est souvent nécessaire. Les comptages prédits par le modèle Béthermin *et al.* (2011), qui utilise une évolution complexe de la fonction de luminosité, sont tracés en bleu. La démarche de construction d’un tel modèle sera expliquée Chap. 6.

2.2 Les relevés infrarouges

2.2.1 Choix des champs observés

Les comptages infrarouges peuvent être estimés à partir de la carte infrarouge d'une région du ciel. Il "suffit" alors d'estimer le nombre de sources par tranche de flux dans la carte et de le diviser par la surface. L'Univers étant isotrope, le champ choisi peut théoriquement se situer n'importe où dans le ciel. Toutefois, pour être représentative, les coordonnées de la région observée doivent être choisies de façon non biaisée. On ne peut pas, par exemple, utiliser un champ centré sur un super-amas de galaxies, qui contient un excès de sources par rapport à la moyenne. En pratique, le choix des champs à observer est guidé par ces deux éléments : présence de données ancillaires (comme des données optiques) et faible contamination par les avant-plans galactiques.

2.2.2 Observations

Le champ des télescopes infrarouges est souvent relativement petit (par exemple, $5' \times 5'$ pour MIPS à $24 \mu\text{m}$ ou $4' \times 8'$ pour SPIRE). Il n'est donc pas possible de construire une carte du ciel à partir d'une seule observation pointée. En général, le télescope balaye la région du ciel à observer de manière systématique. Chaque point est observé plusieurs fois, afin d'identifier dans les données des parasites tels que les impacts de rayons cosmiques sur les détecteurs. Ceci permet également de corriger une partie des effets dus au bruit basse fréquence des instruments.

2.2.3 Construction des cartes infrarouges

Les données fournies par le télescope sont les suivantes.

- Le signal reçu par chaque détecteur en fonction du temps. Ce signal peut être ou non déjà converti dans une unité de flux.
- La direction dans laquelle observe chaque détecteur en fonction du temps. En pratique, cette information est souvent fournie grâce au pointage d'un seul des détecteurs et d'un angle. Le pointage des autres détecteurs se déduit de la géométrie de la matrice de détecteurs.
- Dans certains cas (comme SPIRE), des bolomètres aveugles⁵ sont placés sur la matrice, afin de corriger les fluctuations de température de celle-ci.

⁵Ne captant pas le signal venu du ciel

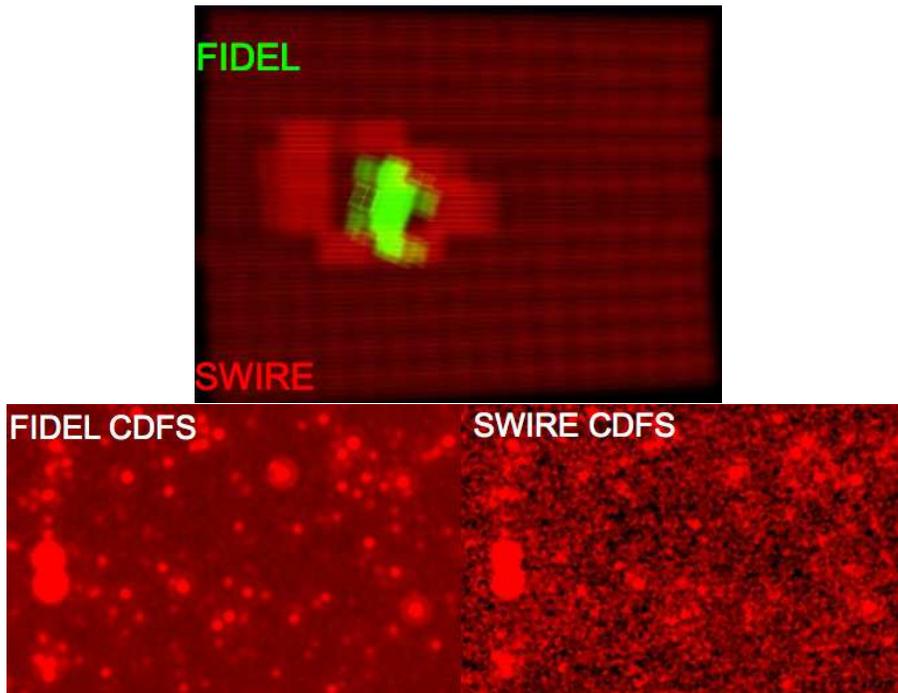


FIGURE 2.3 – Relevés *Spitzer* 24 μm larges et profonds dans le CDFS⁸. (En haut) Carte de couverture du relevé profond FIDEL (en vert) et du relevé large SWIRE (en rouge). (En bas) Image de la même zone du ciel vue par FIDEL (gauche) et SWIRE (droite).

Une première analyse permet de masquer ou corriger dans le signal temporel de chaque détecteur les zones affectées par des effets instrumentaux, comme par exemple les impacts de rayons cosmiques.

Des logiciels permettent de convertir ces informations en une carte de la région observée. La redondance peut être utilisée pour réduire les effets du bruit basse-fréquence des instruments. Le logiciel SANEPIC⁶(Patanchon *et al.* (2008)) réalise, par exemple, un traitement bayésien optimal du bruit corrélé pour produire les cartes BLAST. Toutefois, dans certains cas où le bruit basse fréquence est très fort, comme MIPS, il est soustrait par un filtrage médian multi-fenêtres des données (Gordon *et al.* (2007)).

2.2.4 Champs profonds et champs larges

En l'absence de corrélation, le bruit instrumental décroît comme $1/\sqrt{t_{int}}$, où t_{int} est le temps d'intégration sur un point donné du ciel. Pour détecter les sources les plus faibles,

⁶Signal and Noise Estimation Procedure Including Correlations

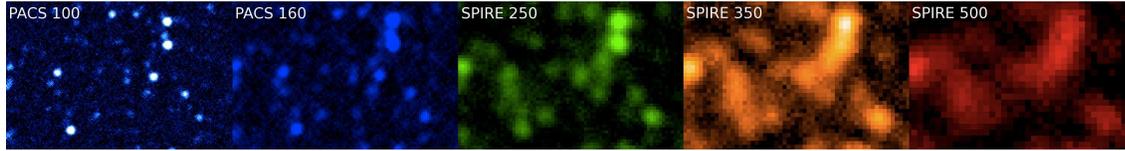


FIGURE 2.4 – Images de la même zone du ciel vue à différentes longueurs d’ondes par le satellite *Herschel* (données H-GOODS).

il est donc nécessaire d’utiliser des temps d’intégration très longs. On parlera alors de champs profonds. Il n’est donc en général pas possible d’obtenir ce genre d’images sur des grandes zones du ciel. Or, les petits champs ne contiennent que peu de sources brillantes. Pour obtenir un échantillon significatif de sources brillantes, il est donc également très utile de réaliser des observations peu profonde sur de larges régions du ciel. Les observations profondes et larges sont donc fortement complémentaires. La Fig. 2.3 illustre la différence entre champs profonds et larges. Il faut également noter que dans l’infrarouge le phénomène de confusion (voir la section suivante) limite fortement la profondeur des relevés. Certains flux faibles restent donc inaccessibles quelque soit le temps d’intégration.

2.3 Extraction des sources infrarouges

2.3.1 La problème de la confusion

A cause du phénomène de diffraction, la résolution angulaire d’un télescope de diamètre donné diminue quand la longueur d’onde augmente. A cause de cela, il est impossible d’extraire d’une carte même sans bruit instrumental les sources plus faibles qu’un certain flux, appelé limite de confusion. En effet, plus on considère des bas flux, et plus la densité de sources est grande. A partir d’un certain flux, il y a trop de sources par PSF, et elles deviennent impossibles à extraire. Ce phénomène limite fortement la possibilité d’étudier les sources faibles aux grandes longueurs d’ondes. La Fig. 2.4 illustre l’effet de la confusion dans le relevé profond *Herschel*-GOODS.

Il existe plusieurs critères permettant de définir le bruit de confusion. Le premier est un critère de densité. Il correspond au flux pour lequel la densité de source devient trop importante. Le choix de cette densité critique n’est pas trivial. On peut par exemple suivre l’approche de Dole *et al.* (2003). Le densité de sources limite N_{SDC} est atteinte quand il y a une probabilité P d’avoir une autre source dans un rayon $k \theta_{FWHM}$, où θ_{FWHM} est la largeur à mi-hauteur de la PSF. La probabilité P qu’une source tombe dans le rayon d’une autre source vaut :

$$P = 1 - e^{-\pi N k^2 \theta_{FWHM}^2}, \quad (2.7)$$

où N est la densité de sources. On en déduit :

$$N_{SDC} = -\frac{\log(1 - P)}{\pi k^2 \theta_{FWHM}^2}. \quad (2.8)$$

Dole *et al.* (2003) a proposé d'utiliser $P=0.1$ et $k=0.8$ en se basant sur des simulations. Ceci correspond à une densité d'environ 17 *beam*/source⁹.

Dans le cas où les comptages ont une forte pente, les fluctuations du fond dues aux sources faibles agissent comme un bruit. On parle alors de bruit de confusion. Ce bruit n'est pas gaussien. Le formalisme du $P(D)$ (voir Chap. 4) permet de calculer, à partir d'un modèle de comptages, l'histogramme d'une carte en l'absence de bruit instrumental. Cet histogramme présente une forme proche d'une gaussienne à bas flux et une queue de distribution en loi de puissance à haut flux due aux sources brillantes. Cette dernière rend non trivial la définition précise du bruit de confusion. Nous avons choisi le définir comme étant l'écart inter-quartile multiplié par 1.349, de manière à ce que dans le cas gaussien le bruit de confusion corresponde exactement à l'écart type. Le bruit de confusion prédit par le modèle Béthermin *et al.* (2011) (voir Chap. 6) est tracé Fig. 2.5 (graphique du haut). On remarque qu'il décroît fortement avec le diamètre de l'instrument.

Pour qu'une source soit résolue, il faut donc qu'elle soit plus brillante que le limite fixée par la densité, mais aussi le bruit de confusion à $5\text{-}\sigma$ ¹⁰. Le graphique du bas représente la fraction du CIB émise par les sources résolues. L'astérisque représente la longueur d'onde pour laquelle on change de régime. En effet, les courtes longueurs d'ondes sont limitées par la densité, et les grandes longueurs d'ondes par les fluctuations. On constate que *Spitzer* ne résout une fraction significative du CIB qu'en dessous de $70 \mu\text{m}$, alors que *Herschel* résout la majeure partie du CIB jusqu'à $160 \mu\text{m}$.

2.3.2 Détection

La première étape à réaliser lors d'une extraction de source est de déterminer leurs positions dans la carte. Il existe de nombreuses méthodes possibles. La plus simple consiste à chercher les pixels qui ont un rapport signal sur bruit supérieur à un certain seuil. Cette méthode est limitée car elle ne prend en compte qu'un seul pixel, alors que le flux d'une source s'étale sur plusieurs. Pour prendre cela en compte, on peut lisser la carte en la

⁹Le *beam* est la taille angulaire de la PSF. Il vaut $\Omega = \frac{\iint PSF(\theta, \phi) d\theta d\phi}{\max(PSF)}$. Pour un profil gaussien ou de Airy, on a $\Omega \approx 1.1 \times \theta_{FWHM}^2$

¹⁰On peut descendre en dessous de $5\text{-}\sigma$, mais les biais d'extractions peuvent devenir importants

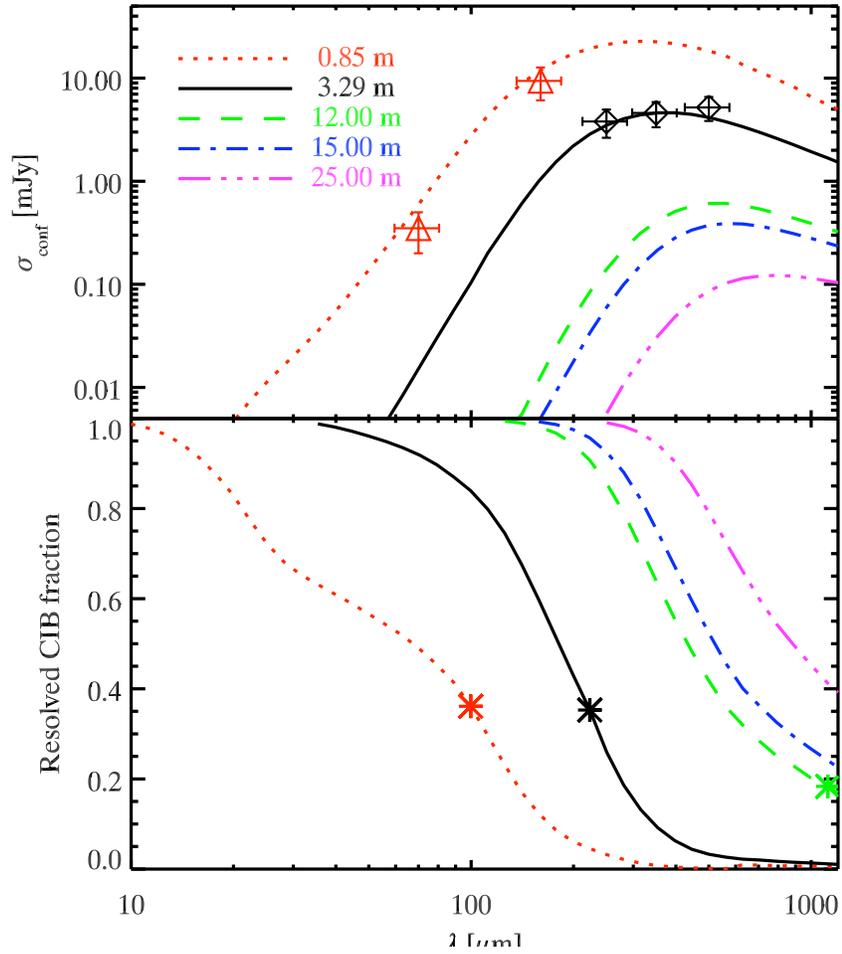


FIGURE 2.5 – (En haut) Bruit de confusion à 1σ en fonction de la longueur d’onde pour différents diamètres de télescopes. (En bas) Fraction du CIB résolue en fonction de la longueur d’onde pour différents diamètres de télescopes. L’astérisque marque le changement de régime de confusion. Extrait de Béthermin *et al.* (2011).

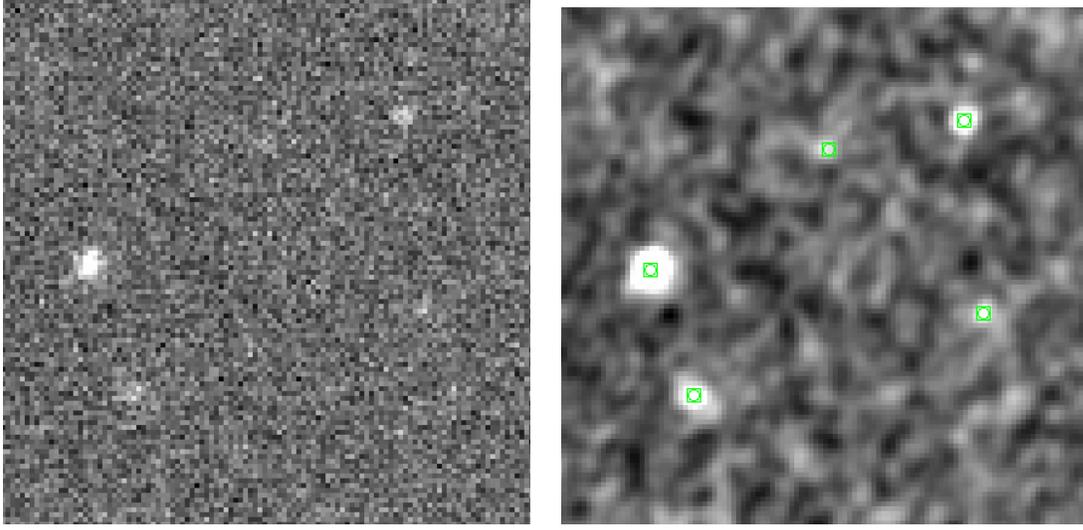


FIGURE 2.6 – (A gauche) Carte d’une région du ciel observée avec BLAST. (A droite) Même région, mais où la carte a été lissée par la PSF. Les carrés vers indiquent les détections à 5σ .

cross-corrélatant avec la PSF. La carte lissée m_l est donc

$$m_l(i_0, j_0) = \sum_{i=-N}^{+N} \sum_{j=-N}^{+N} m(i_0 + i, j_0 + j) \times PSF(i, j), \quad (2.9)$$

avec $(m_l(i_0, j_0))$ le signal de la carte lissée aux coordonnées (i_0, j_0) , m le signal de la carte initiale. La carte de bruit correspondante n_l est donnée par

$$n_l(i_0, j_0) = \sqrt{\sum_{i=-N}^{+N} \sum_{j=-N}^{+N} n^2(i_0 + i, j_0 + j) \times PSF^2(i, j)}, \quad (2.10)$$

où n est la carte de bruit initiale. On en déduit la carte de signal sur bruit en divisant m_l par n_l . On détecte alors les maxima locaux dépassant le seuil demandé, typiquement entre 3 et 5. On peut montrer que cette méthode est optimale pour des sources isolées et un bruit gaussien non-corrélé (pour cela on utilise le formalisme présenté Sect. 2.3.4). La Fig. 2.6 illustre le fonctionnement de cette méthode. On voit clairement que les sources n’émergent que légèrement du bruit dans la carte initiale, alors qu’elles sont parfaitement visibles dans la carte lissée.

La routine `find` de DAOPHOT (Stetson (1987)) utilise une méthode quasi-similaire. Toutefois, la routine cross-correlle la carte par une gaussienne à la place de la PSF réelle. Les pics ne satisfaisant pas certains critères statistiques (symétrie centrale, moment d’ordre

3...) sont éliminés.

Lorsque la densité de source est forte, mais que les fluctuations du fond dues aux sources non résolues sont faibles (comme par exemple dans l'infrarouge proche et moyen), le principal défi est de séparer les paires de sources si proches que leurs PSFs se confondent. Le programme SExtractor (Bertin et Arnouts (1996)), grâce à une étude itérative des maxima locaux dans les cartes, est très efficace dans ces conditions.

Enfin, lorsque les fluctuations du fond sont fortes, comme c'est par exemple le cas pour BLAST, il est préférable d'utiliser une méthode séparant les fluctuations aux grandes échelles des sources (petites échelles). Cette analyse peut être faite par analyse en ondelettes, avec par exemple l'algorithme *atrou* (Starck *et al.* (1999)).

2.3.3 Photométrie d'ouverture

Un fois la position des sources connue, il est ensuite nécessaire de mesurer leur flux. La photométrie d'ouverture est une méthode simple de photométrie. Elle consiste à intégrer le flux reçu dans une ouverture circulaire centrée sur la source étudiée. Le niveau du fond sera estimé dans un anneau situé suffisamment loin de la source pour ne pas être trop contaminé par cette dernière, mais pas trop pour que l'estimation du fond reste fidèle. La Fig. 2.7 schématise le fonctionnement de cette méthode de photométrie.

Le flux mesuré dans l'ouverture n'est pas exactement le flux de la source. En effet, une partie non-négligeable de son flux est diffracté hors l'ouverture. En raison de la présence d'autres sources dans les cartes, il est en général impossible d'utiliser des ouvertures suffisamment grandes pour que ce problème soit négligeable. On peut néanmoins montrer que le flux initial s'obtient en appliquant un simple facteur multiplicatif, appelé correction d'ouverture. On considère une carte m contenant une unique source ponctuelle de flux S sur un fond C :

$$m(r, \theta) = S \times PSF(r, \theta) + C, \quad (2.11)$$

où $PSF(\theta, \phi)$ est la réponse instrumentale à une source de flux unité centrée sur l'origine des coordonnées. Le flux intégré sur l'ouverture de rayon a est

$$S_a = \int_{\theta=0}^{2\pi} \int_{r=0}^a S \times PSF(r, \theta) + C dr d\theta \quad (2.12)$$

$$= S \times \left[\int_{\theta=0}^{2\pi} \int_{r=0}^a PSF(r, \theta) dr d\theta \right] + A \times C \quad (2.13)$$

$$= S \cdot \alpha + A \cdot C, \quad (2.14)$$

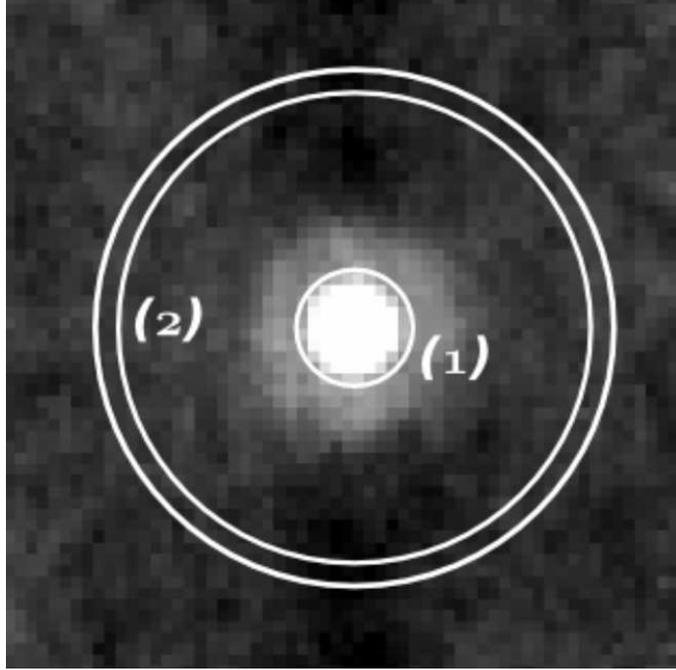


FIGURE 2.7 – Principe de la photométrie d'ouverture. (1) Ouverture circulaire. (2) Anneau d'estimation du fond de ciel. Extrait du rapport de stage de Morgane Cousin.

où A est l'aire de l'ouverture. On note $\alpha = \int_{\theta=0}^{2\pi} \int_{r=0}^{r_{ext}} PSF(r, \theta) dr d\theta$. Le signal provenant du fond dans l'ouverture est estimé à partir du signal dans l'anneau et du rapport des surfaces entre l'ouverture et l'anneau :

$$S_{a,fond} = \frac{A}{A_{anneau}} \int_{\theta=0}^{2\pi} \int_{r=r_{int}}^{r_{ext}} S \times PSF(r, \theta) + C dr d\theta \quad (2.15)$$

$$= S \times \left[\frac{A}{A_{anneau}} \int_{\theta=0}^{2\pi} \int_{r=r_{int}}^{r_{ext}} PSF(r, \theta) dr d\theta \right] + A \times C \quad (2.16)$$

$$= S \cdot \beta + A \cdot C, \quad (2.17)$$

où r_{int} et r_{ext} sont les rayons intérieurs et extérieurs de l'anneau, A l'aire de l'ouverture et A_{anneau} l'aire de l'anneau. On note $\beta = \int_{\theta=0}^{2\pi} \int_{r=r_{int}}^{r_{ext}} PSF(r, \theta) dr d\theta$. Le flux mesuré S_m est donc

$$S_m = S_a - S_{a,fond} = S \times (\alpha - \beta). \quad (2.18)$$

La correction d'ouverture est donc $1/(\alpha - \beta)$. En pratique, elle peut être estimée en mesurant l'inverse du flux d'une PSF unitaire¹¹, avec l'ouverture et l'anneau choisis.

¹¹Image d'une source de flux unité représenté dans les même unité que la carte sur laquelle on travaille.

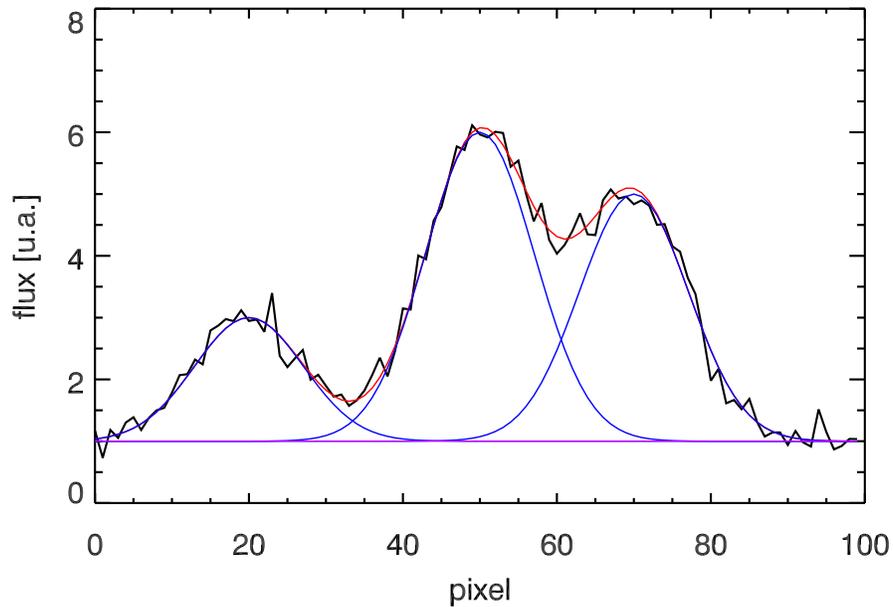


FIGURE 2.8 – Principe de la photométrie par ajustement de PSF. La courbe est une coupe d’une carte simulée où les sources sont supposées gaussiennes. La ligne en noir est le signal observé, la ligne en violet le fond, les lignes en bleu le profil des différentes sources, et enfin la ligne rouge le total des sources et du fond (sans bruit).

2.3.4 Photométrie par ajustement de PSF

Principe

La photométrie d’ouverture est une méthode très efficace lorsque les sources sont isolées. En revanche, elle l’est beaucoup moins lorsque la densité devient élevée. En effet, cette méthode ne permet pas de séparer le flux de deux sources proches dont les tâches de diffraction sont mélangées. Dans ce genre de cas, il est préférable d’utiliser des méthodes d’ajustement de PSF. Cette méthode consiste à ajuster un modèle contenant des sources et un fond. Les paramètres libres de ce modèle sont le flux des sources (et éventuellement leur position) et la valeur du fond. Les flux ajustés sont bien retrouvés même lorsque les sources sont mélangées. La Fig. 2.8 illustre le principe de cette méthode. Il existe de nombreuses routines d’ajustement de PSF, comme DAOPHOT/allstar (Stetson (1987)) ou Starfinder (Diolaiti *et al.* (2000)).

Développement mathématique

Sous certaines hypothèses, le PSF-fitting peut se résoudre à un problème relativement simple. On se place dans le cadre suivant :

- Les sources sont toutes ponctuelles et de positions connues ;
- Le bruit instrumental est gaussien et décorréolé.

La carte infrarouge considérée peut être décrite de la manière suivante :

$$m = \sum_{k=1}^{N_s} S_k \times PSF_{x_k, y_k} + \sigma + \mu, \quad (2.19)$$

où m est le tableau décrivant la carte, N_s le nombre de sources à ajuster, S_k le flux de la k -ième source, PSF_{x_k, y_k} est un tableau de la taille de la carte contenant une seule source de flux unité centrée sur la position (x_k, y_k) , σ un bruit gaussien, et μ un fond constant.

Le vecteur S est le vecteur contenant les S_k . La vraisemblance est alors donnée par :

$$L(m|S) = \prod_{pixels} C(n) \times \exp\left[-\frac{(m - \sum_{i=1}^{N_{sources}} PSF_{x_i, y_i} \times S_i - \mu)^2}{2n^2}\right], \quad (2.20)$$

où $C(n)$ est une constante de normalisation qui ne dépend que de n , la carte de bruit. Le logarithme de cette expression est donc :

$$-\log(L) = \sum_{pixels} \frac{(m - \sum_{i=1}^{N_{sources}} PSF_{x_i, y_i} \times S_i - \mu)^2}{2n^2} + \sum_{pixel} \log(C(n)). \quad (2.21)$$

La valeur de S maximisant la vraisemblance doit vérifier :

$$\forall S_k, \frac{\partial(-\log(L))}{\partial S_k} = 0 \quad \& \quad \frac{\partial(-\log(L))}{\partial \mu} = 0, \quad (2.22)$$

et par conséquent,

$$\forall S_k, \sum_{pixels} \frac{-PSF_{x_k, y_k} (m - \sum_{i=1}^{N_{sources}} PSF_{x_i, y_i} \times S_i - \mu)}{n^2} = 0. \quad (2.23)$$

$$\text{et } \sum_{pixels} \frac{m - \sum_{i=1}^{N_{sources}} PSF_{x_i, y_i} \times S_i - \mu}{n^2} = 0 \quad (2.24)$$

On peut réécrire cela de la façon suivante :

$$\forall S_k, \sum_{\text{pixels}} \frac{\mu \times PSF_{x_k, y_k}}{n^2} + \sum_{i=1}^{N_{\text{sources}}} S_i \times \sum_{\text{pixels}} \frac{PSF_{x_k, y_k} \times PSF_{x_i, y_i}}{n^2} = \sum_{\text{pixels}} \frac{PSF_{x_k, y_k} \times m}{n^2} \quad (2.25)$$

et

$$\sum_{\text{pixels}} \frac{\mu}{n^2} + \sum_{i=1}^{N_{\text{sources}}} S_i \times \sum_{\text{pixels}} \frac{PSF_{x_k, y_k}}{n^2} = \sum_{\text{pixels}} \frac{m}{n^2}. \quad (2.26)$$

Ceci est équivalent à une équation matricielle de taille $(N_{\text{sources}} + 1) \times (N_{\text{sources}} + 1)$:

$$A \cdot X = B \quad (2.27)$$

où X est un vecteur contenant $[S_1, S_1, \dots, S_{N_{\text{sources}}}, \mu]$ et où A et B sont définies par :

$$\forall i \in [1, N_{\text{sources}}], \forall j \in [1, N_{\text{sources}}], A = (a_{ij}) = \sum_{\text{pixels}} \frac{PSF_{x_i, y_i} \times PSF_{x_j, y_j}}{n^2} \quad (2.28)$$

$$\forall i \in [1, N_{\text{sources}}], (a_{i(N_{\text{sources}}+1)}) = (a_{(N_{\text{sources}}+1)i}) = \sum_{\text{pixels}} \frac{PSF_{x_i, y_i}}{n^2} \quad (2.29)$$

$$(a_{(N_{\text{sources}}+1)(N_{\text{sources}}+1)}) = \sum_{\text{pixels}} \frac{1}{n^2} \quad (2.30)$$

$$\forall i \in [1, N_{\text{sources}}], B = (b_i) = \sum_{\text{pixels}} \frac{PSF_{x_i, y_i} \times m}{n^2} \quad (2.31)$$

$$B = (b_{N_{\text{sources}}+1}) = \sum_{\text{pixels}} \frac{m}{n^2} \quad (2.32)$$

La matrice A est symétrique et donc inversible. L'estimateur du maximum de vraisemblance est donc :

$$S = A^{-1}B. \quad (2.33)$$

L'ajustement simultané de N sources peut donc se réduire au problème de l'inversion d'une matrice $(N_{\text{sources}} + 1) \times (N_{\text{sources}} + 1)$.

On peut également estimer les erreurs. La matrice de Fisher de ce problème est :

$$F(S) = (f_{mn})(X) = -E\left(\frac{\partial^2 \log(L)}{\partial X_m \partial X_n} \middle| X\right), \quad (2.34)$$

où $E(X|S)$ est l'espérance de la variable aléatoire X sachant S. On utilise alors les résultats de l'Eq. 2.26 :

$$\frac{\partial(\log(L))}{\partial S_n} = - \sum_{i=1}^{N_{\text{sources}}} S_i \times \sum_{\text{pixels}} \frac{PSF_{x_n, y_n} \times PSF_{x_i, y_i}}{n^2} + \sum_{\text{pixels}} \frac{PSF_{x_n, y_n} \times m}{n^2}, \quad (2.35)$$

puis,

$$\frac{\partial^2(\log(L))}{\partial S_m \partial S_n} = - \sum_{pixels} \frac{PSF_{x_n, y_n} \times PSF_{x_m, y_m}}{n^2} = (a_{nm}). \quad (2.36)$$

On peut aussi montrer que :

$$\frac{\partial^2(\log(L))}{\partial S_n \partial \mu} = (a_{n(N_{sources}+1)}) \quad \& \quad \frac{\partial^2(\log(L))}{\partial \mu \partial \mu} = (a_{(N_{sources}+1)(N_{sources}+1)}). \quad (2.37)$$

A est donc la matrice de Fisher du problème. Les variances sur les flux estimés sont donc les termes diagonaux de l'inverse de la matrice A, la matrice de covariance étant l'inverse de la matrice de Fisher.

$$Var(S_k) = (A^{-1})_{kk}. \quad (2.38)$$

C'est ce formalisme qui est utilisé par la routine FASTPHOT (Béthermin *et al.* (2010b)) que j'ai développé en collaboration avec Morgane Cousin lors son stage de Master 1. Elle a été optimisée pour être extrêmement rapide et s'applique au cas où les sources sont à une position connue (typiquement à partir d'observations à une plus courte longueur d'onde). La Fig. 2.10 montre une carte simulée du ciel et les résidus par rapport au modèle ajusté par FASTPHOT. On voit qu'il ne reste que du bruit instrumental, et que les sources même proches ont bien été extraites. FASTPHOT et sa documentation ont été mis en ligne publiquement sur le site web de groupe galaxies infrarouges de l'IAS (<http://www.ias.u-psud.fr/irgalaxies/downloads.php>).

2.3.5 Photométrie et filtrage des cartes

Dans le cas d'une source isolée et d'une carte à fond nul (ce qui est presque équivalent à une carte à moyenne nulle), A et B sont de simples nombres. On a alors

$$S = \frac{B}{A} = \frac{\sum_{pixels} \frac{PSF_{x_i, y_i} \times m}{n^2}}{\sum_{pixels} \frac{PSF_{x_i, y_i} \times PSF_{x_j, y_j}}{n^2}}. \quad (2.39)$$

Si on suppose également un bruit uniforme, le flux est estimé simplement par

$$S = \frac{B}{A} = \frac{\sum_{pixels} PSF_{x_i, y_i} \times m}{\sum_{pixels} PSF_{x_i, y_i} \times PSF_{x_j, y_j}}. \quad (2.40)$$

Dans une carte cross-corrélée par la PSF et normalisée en appliquant un facteur 1/A, le signal dans le pixel centré sur la source est donc exactement son flux estimé par PSF-fitting. Cette méthode est optimale au sens des moindres carrés dans le cas d'un bruit blanc, mais pas en présence de bruit de confusion où un autre filtrage peut être plus performant. De

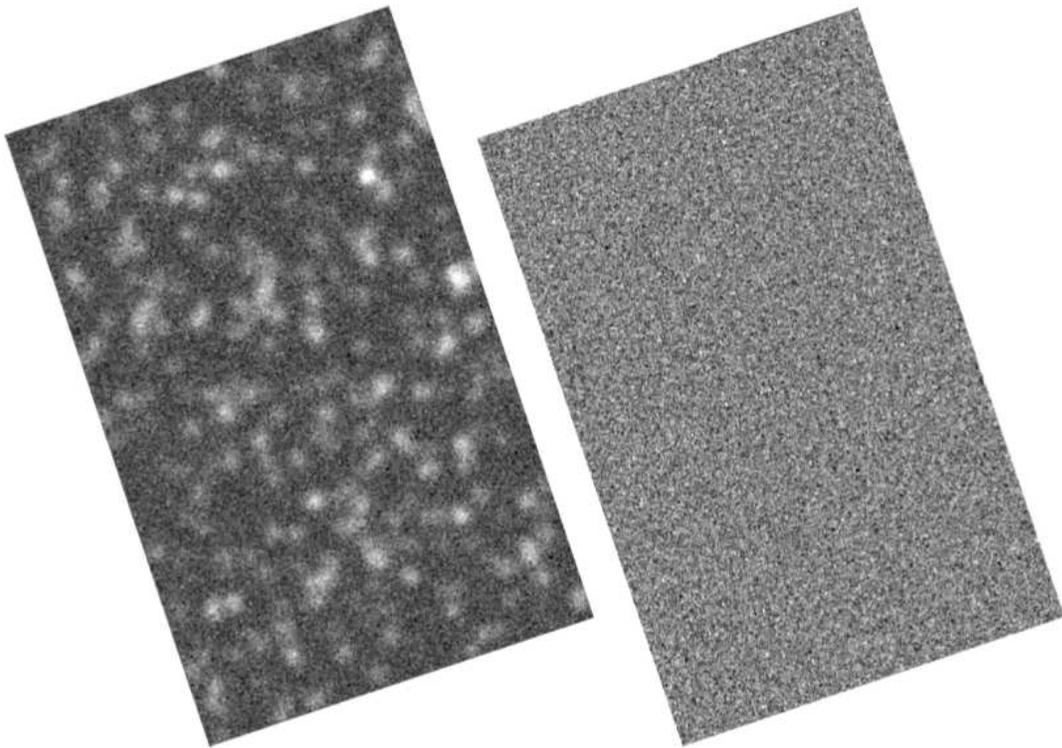


FIGURE 2.9 – (A gauche) Carte simulée. (A droite) Carte de résidus après utilisation de la routine d'ajustement de PSF FASTPHOT. Extrait du rapport de stage de M1 de Morgane Cousin.