

Un spectrographe 3D : La Spectral Energy Distribution machine

Dans le chapitre précédent, nous avons présenté la collaboration Zwicky Transient Facility et nous nous sommes focalisés sur la caméra principale de 47deg^2 montée sur le P48 au Mont Palomar. Cette caméra permet à ZTF de détecter 10^5 événements transitoires ou variables, en scannant l'entièreté du ciel Nord visible chaque nuit, à la vitesse vertigineuse de $3760\text{deg}^2/\text{heure}$. Parmi ces événements, $\mathcal{O}(10)$ correspondent à de nouveaux événements transitoires non répertoriés : Les Supernovae. Comme expliqué dans le Chapitre ??, seules les Supernovae de type Ia sont d'intérêts dans la cosmologie, de part leur propriété de chandelle standardisable. Il faut donc les classifier. Pour cela, on utilise leur spectre, dont les raies d'absorption/émission sont caractéristiques d'un type à l'autre de SN. Ainsi, ZTF possède également un spectrographe 3D monté sur le télescope P60 au Mont Palomar (Figure ??) spécialement conçu à cet effet. Nous présentons dans ce chapitre ce spectrographe, la Spectral Energy Distribution machine (SEDm).

3.1 Présentation de l'instrument

3.1.1 Principe d'un IFS

Le spectrographe 3D SEDm est ce qu'on appelle un IFS pour Integral Field Spectrograph. Sans surprise, c'est un instrument qui permet de recueillir le spectre du ciel sur un champ de vue bidimensionnel. Ainsi et indépendamment de la méthode utilisée, le produit final avec cet instrument correspond à un cube de données ayant 2 dimensions spatiales ((x, y) ou (RA, Dec)) et une dimension spectrale (longueur d'onde λ ou une vitesse). Dans tout le manuscrit, cette notion de 3D/cube 3D fera systématiquement référence aux dimensions " x, y, λ ".

Un IFS est composé de 2 parties : le spectrographe qui va disperser la lumière incidente, et l'IFU (Integrated Field Unit). Le rôle de l'IFU est de diviser le plan spatial 2D du champ de vue en un réseau continu et concentré de lumière. Ce réseau est ensuite donné en entrée au spectrographe qui va se charger de le disperser sur le détecteur.

Il existe 3 types principaux d'IFU, schématisés dans la Figure 3.1.

- **Le réseau de micro-lentilles** conceptualisé par ? (qui s'apparente aux yeux composites de certains insectes) : C'est le système utilisé par la SEDm, mais également par l'IFS SAURON (?) dans le projet ATLAS3D (?) ou encore SNIFS (?). Dans ce système, l'image bi-dimensionnelle est fractionnée par un réseau de micro-lentilles (le MLA, microlens array). Chaque élément est ensuite concentré et dispersé par le spectrographe (voir Figure 3.1). Pour éviter au maximum le chevauchement des spectres sur le détecteur, le réseau de lentille est légèrement incliné. Le désavantage principal de cette technique est le court intervalle de longueur d'onde dispersable sans induire de chevauchement.
- **Le paquet de fibre** comme avec l'IFS du relevé MaNGA d'SDSS (?) qui peut être utilisé en combinaison (?) ou non (?) de réseau de micro-lentilles. Ici la lumière n'est pas concentrée par des lentilles mais acheminée par un paquet de fibres optiques "à la chaîne" jusqu'à la fente du spectrographe. Le premier avantage est bien évidemment la flexibilité des fibres. Mais en contrepartie l'échantillon du ciel dispergé devient non contigu, à cause de la forme circulaire des fibres. Il est possible de pallier à cet effet en ajoutant un réseau de micro-lentilles (lui contigu) entre le plan focal et le paquet de fibres.
- **Le "trancheur d'image"** qui est la méthode la plus ancienne (?, ?) utilisée par exemple avec le NIFS (near-infrared integral field spectrograph, ?). Cette méthode utilise un miroir segmenté en fines sections horizontales. Chacune de ces sections va diriger la lumière incidente dans des directions légèrement différentes jusqu'à un second miroir segmenté. Ce dernier va réarranger les tranches incidentes non pas l'une au dessus de l'autre, mais de façon étalées, "à la chaîne" comme avec la méthode fibrée. L'agencement est ensuite dispersé par la fente du spectrographe. Cette méthode permet de conserver la contiguïté du champ de vue, mais est en contrepartie couteuse et difficile à concevoir.

Les données brutes obtenues à partir d'un IFS sont ainsi sous la forme de multiples spectres (de plusieurs dizaines à plusieurs milliers) étalés (la trace) sur le détecteur, chacun ayant pour origine un élément individuel de l'IFU. Ces éléments sont en quelques sortes des

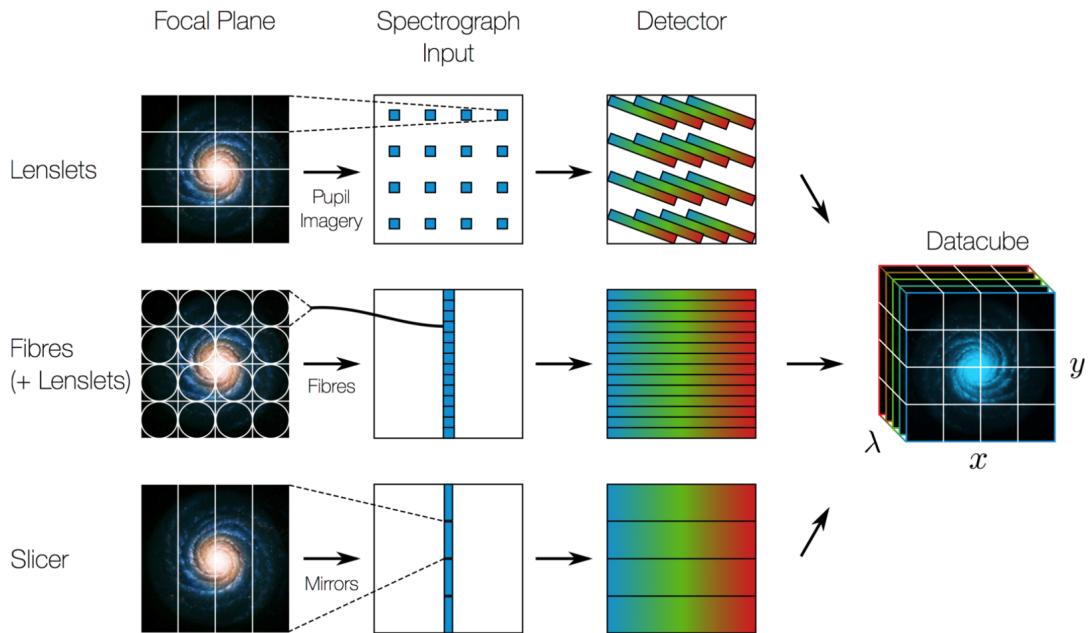


FIGURE 3.1 – Fonctionnement d'un IFS pour différents types d'IFU. La SEDm utilise un système d'agencement de micro-lentilles (cas du haut) (*Crédit M. Westmoquette, adaptée de ?*)

pixels spatiaux, que l'on contracte communément par le terme de spaxels. La reconstruction du cube de données se fait en extrayant chaque spectre du détecteur, et en les réarrangeant dans le même espace géométrique que le plan focal du télescope (nous détaillerons ce processus dans la section suivante).

3.1.2 La SEDm

Focalisons nous maintenant sur notre instrument, la Spectral Energy Distribution machine, présenté par ?. Comme mentionné plusieurs fois, celui ci est monté sur le télescope P60 (Cassegrain) au Mont Palomar depuis Août 2016. Une vue d'ensemble de l'instrument est présenté dans la Figure 3.2, où l'on peut voir qu'il est composé de deux canaux : l'IFU et la “Rainbow Camera” (RC), montés sur un agencement en forme de T. Cette caméra d'acquisition multi-bande est accompagnée de 4 filtres photométriques u' , g' , r' et i' .

Les 2 caméras de la SEDm sont des Princeton Instruments identiques : une PIXIS 2048B et une PIXIS 2048B_Excelon chacun avec 2048×2048 pixels de taille $13,5 \mu\text{m}$.

La Rainbow Camera est utilisée pour le guidage, la calibration, l'acquisition de cible ou encore l'imagerie scientifique. Le champ de vue de $13' \times 13'$ est divisé en 4 quadrants, un pour chacun des filtres $u'g'r'i'$.

L'IFU de la SEDm fonctionne sur la méthode du réseau de micro-lentilles, le MLA. Celui ci couvre un champ de vue de 28×28 arcsec, avec 45×52 lentilles hexagonales. Le faisceau de lumière projeté par ces lentilles passe dans un triple prisme avec une résolution spectrale achromatique de $R = \frac{\lambda}{\Delta\lambda} \sim 100$. Comme illustré dans la Figure 3.2, c'est la RC qui est alignée avec la lumière directe en provenance du Cassegrain. Il faut donc en dévier une partie qui sera transmise à la caméra de l'IFU : cela est effectué avec un prisme

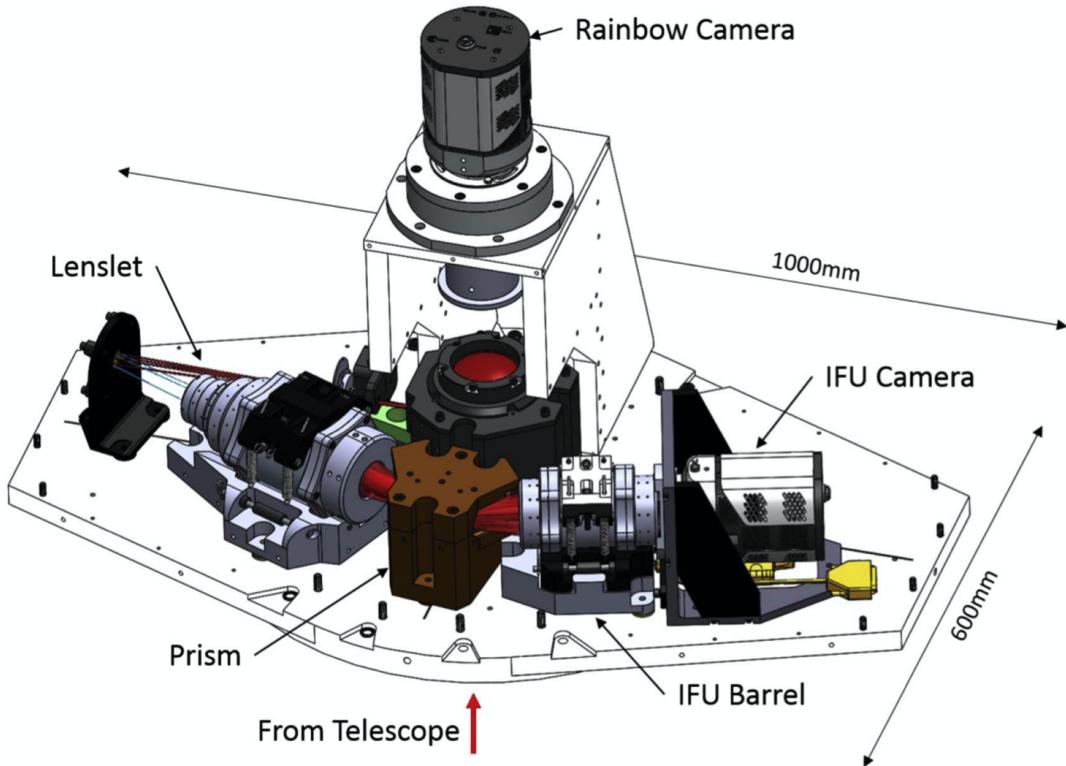


FIGURE 3.2 – Vue d’ensemble de la SEDm (?). La source de lumière du Cassegrain est indiqué par la lumière rouge. L’instrument photométrique, la RC, est situé en haut au centre et récolte directement la lumière. Pour l’IFU, la lumière est redirigée jusqu’à un miroir que l’on peut voir tout à gauche de la représentation. Elle est ensuite réfléchie dans le MLA (Lenslet). La lumière de chaque micro-lentille du MLA passe ensuite dans le prisme (en marron au centre) pour y être dispersé et focalisé par l’optique (IFU Barrel) sur le détecteur (IFU Camera).

d’interception centré sur le faisceau incident, qui va rediriger le champ de 28×28 arcsec vers un miroir. Les images photométriques de la RC ont donc en leur centre un masque qui correspond au champ de vue de l’IFU. Ainsi pour faire l’acquisition d’une cible avec l’IFU, il faut d’abord effectuer une acquisition avec la RC sur laquelle un fitter d’astrométrie est appliquée. Cette étape fournit une information précise sur le pointage du télescope, et permet donc d’appliquer l’offset nécessaire pour positionner la cible au centre du champ de vue de l’IFU.

3.2 Extraction des spectres du CCD et création des cubes de données

Dans cette section nous allons détailler la procédure et le pipeline associé permettant de reconstruire les cubes de données à partir des images brutes 2D obtenues sur le CCD de l’IFU. Ce passage 2D → 3D nécessite d’identifier 3 éléments :

- ◊ La localisation de la lumière des spaxels sur le CCD : la trace.
- ◊ La correspondance entre la localisation de la trace sur le CCD et la position du spaxel dans le plan focal.
- ◊ La cartographie de la dispersion spectral qui permet de convertir chaque pixel du CCD en longueur d'onde (et vice versa).

Ces trois étapes de calibration sont effectuées durant l'après-midi afin d'optimiser le temps d'observation. Durant les acquisitions scientifiques, certaines corrections que nous aborderons plus loin sont effectuées en temps réels.

Toute cette démarche a été implémenté dans le pipeline public **PYSEDM**¹, conçu par ? et la suite de cette section résumera une grande partie de cette publication.

3.2.1 Identification des traces

Commençons par l'identification des traces sur le CCD. Pour cette étape on effectue une exposition du dome qui sera l'équivalent d'un flat 3D pour la calibration de l'IFU (les dark et les bias sont directement corrigés sur les images du CCD).

Rappel des images de calibrations.

- **Le flat** correspond à une acquisition d'une lumière uniforme, ce qui permet de mesurer la réponse relative de chaque pixel d'un détecteur. En divisant les images de science par ce flat, on uniformise alors la réponse de chaque pixel, ce qui permet de corriger d'éventuels effets de vignettage (habituellement dominant sur les bords du détecteur). Pour le mesurer, il faut illuminer le détecteur avec une lumière uniforme.
- **Le dark** lui permet de corriger le bruit dû aux longues expositions. Une caméra CCD (ou un capteur CMOS comme dans les Reflex) génère un bruit thermique, dont l'intensité dépend de la température, du temps de pose et des ISO. Pour diminuer ce bruit, on soustrait le dark aux images de science. On crée les dark dans les mêmes conditions que les observations (temps d'exposition, température, ISO etc). Ces images sont prises dans le noir, avec un couvercle devant l'objectif ou un obturateur devant la caméra par exemple.
- **Le bias** va quant à lui corriger le bruit inhérent à la caméra : le bruit de lecture. Tout comme le dark, il se soustrait aux images de science. Tout comme le dark, ces images ce fond dans le noir, mais avec le temps d'exposition minimal atteignable par la caméra.
- **Le dark-flat** est comme le dark, mais pour les images flat.

Pour chaque type d'image, on procède à l'acquisition d'un certain nombre que l'on combine pour former les Master images (Master bias, Master dark etc).

1. <https://github.com/MickaelRigault/pysedm>

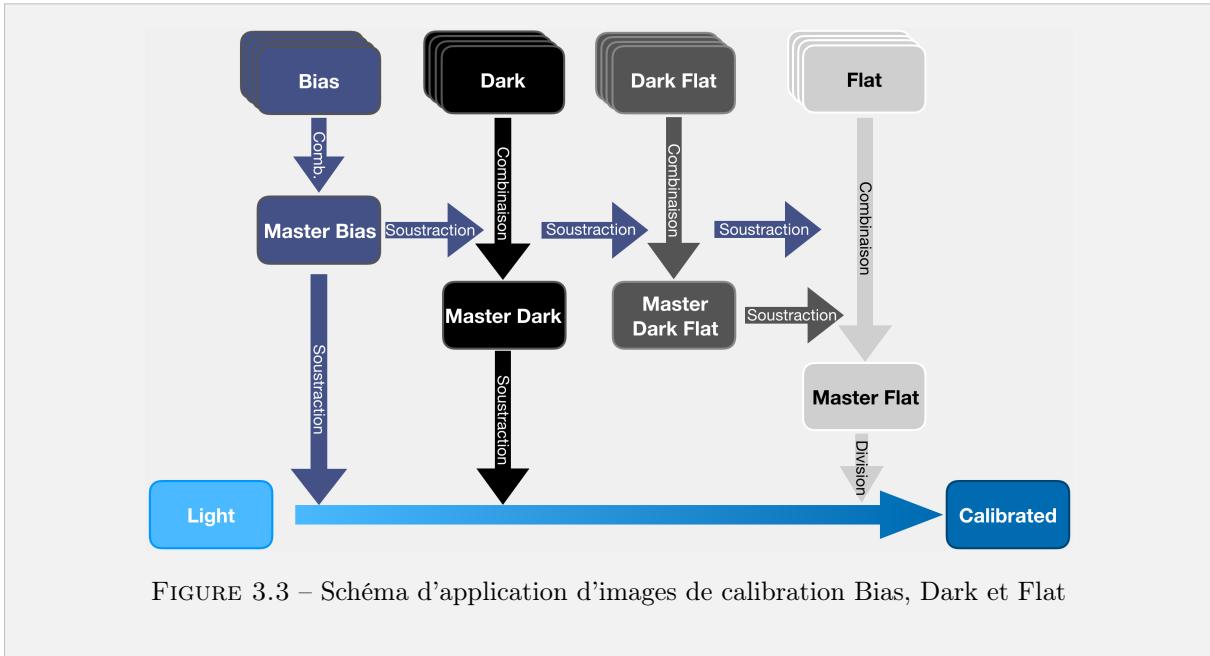


FIGURE 3.3 – Schéma d’application d’images de calibration Bias, Dark et Flat

Une exposition CCD du dome est représentée dans la Figure 3.4a. L’isolation de la lumière dispersée par chaque spaxel se fait en définissant une ellipse via la méthode **EXTRACT** de **SEP**¹ (?) (implémentation python de **SEXTRACTOR** ?). La trace des spaxels est ensuite isolée en considérant un rectangle à partir des informations de l’ellipse, de sorte que 95.5% de la lumière soit encapsulée. L’intervalle de longueur d’onde ainsi isolé va de 3500 à 9500 Å, respectivement de droite vers la gauche. Un masque 2D pondéré est ensuite créé prenant en compte la fraction de l’aire de chaque pixel présent dans le contour défini de la trace. Un spectre pour chaque spaxel est ainsi extrait en unité de “count” par pixel.

3.2.2 Solution en longueur d’onde

On passe maintenant à l’étape qui va permettre d’associer pour chaque pixel une longueur d’onde, et ce pour chaque spaxel indépendamment. Pour cette calibration, on utilise 3 lampes à arc qui émettent de fortes raies d’émissions à des longueurs d’onde connues :

Hg Lampe à mercure avec 4 raies d’émission.

Cd Lampe à cadmium avec 4 raies d’émission.

Xe Lampe à xenon avec 6 raies d’émission.

La procédure est la suivante (et illustré dans la Figure 3.4b) :

- Exposition du CCD pour chaque lampe à arc.
- Extraction du spectre de chaque spaxel (en unité de pixel) pour chaque exposition.

1. <http://github.com/kbarbary/sep>

- (c) Fit indépendant pour chaque spaxel et pour chaque lampe. Pour cela un continuum polynomial de 3^e ordre est utilisé associé à une combinaison de Gaussiennes (autant qu'il y a de raies d'émission).
- (d) Fit conjoint des 14 centroids des raies d'émission en fonction de leur longueur d'onde attendue. Un polynôme de degré 5 est utilisé pour cette étape.

La précision atteinte pour la solution en longueur d'onde est de l'ordre de 3Å au centre de l'IFU, mais peut monter à $\sim 10\text{\AA}$ sur les bords.

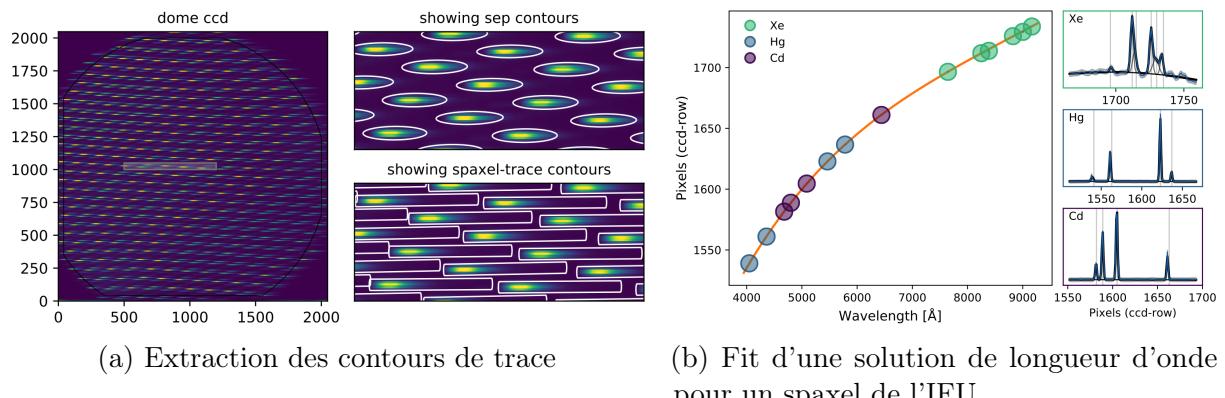


FIGURE 3.4 – Extraction de traces (*à gauche*) et fit de solution en longueur d'onde (*à droite*) pour la SEDm (de ?)

3.2.3 Identification spatiale

La dernière étape avant la construction du cube est l'identification spatiale, afin de récupérer la structure hexagonale du réseau de micro lentilles. Cette procédure est purement géométrique et est détaillée en 8 étapes dans la section 2.1.3 de ?. Le résultat final est une grille hexagonale, avec pour chaque hexagone un identifiant associé au spaxel correspondant ainsi que les coordonnées de chaque vertex. Combinés aux spectres extraits de chaque spaxel, cela permet de cartographier la position de chaque spaxel, et ainsi reconstruire le MLA.

3.2.4 Construction du cube 3D

Vient enfin la reconstruction d'un cube après une acquisition scientifique. Comme expliqué au début de ce chapitre, toutes les observations d'une nuit sont calibrées à partir des informations de calibrations faites durant l'après-midi (bias, dark, flat, solution en longueur d'onde, localisation de trace etc). Cependant certains de ces paramètres peuvent varier au cours de la nuit, c'est pourquoi 2 étapes de corrections sont appliquées à chaque nouvelle exposition : Une pour la localisation des traces, l'autre pour la solution en longueur d'onde. Voici les étapes principales de reconstruction d'un cube 3D à partir d'une image 2D du CCD (on ne détaillera pas ici chacune des étapes mais nous invitons le lecteur à consulter ? pour plus d'information technique) :

- (a) Optimisation de la localisation des traces en effectuant une correction verticale de l'ensemble du réseau.
- (b) Soustraction du background du CCD, construit à partir des pixels qui sont en dehors des traces.
- (c) Extraction des spectres en unité de count par pixel, projection dans l'espace des longueur d'onde grâce à la solution en longueur d'onde, puis création du cube grâce à la l'identification spatiale des spaxels.
- (d) Estimation de la déviation en longueur d'onde en mesurant les raies telluriques et du ciel (connues). Cette correction $\Delta\lambda$ est ensuite convertie en une correction horizontale Δi de pixel.
- (e) Répétition de l'étape (c) en corrigeant de la déviation Δi mesurée dans (d) lors du passage de l'espace des pixels vers l'espace des longueur d'onde.
- (f) Application du cube Flat (exposition du dome) pour corriger la réponse relative des spaxels.

3.3 Actuelle méthode d'extraction de source ponctuelle

3.3.1 Localisation de la cible

Avant d'extraire la source ponctuelle ciblée avec l'IFU, encore faut-il pouvoir la localiser dans le MLA.

Le faible champ de vue de l'IFU de la SEDm ne permet habituellement pas d'observer simultanément de nombreuses sources. En générale, seule la cible et sa galaxie hôte sont visibles dans la MLA, ce qui ne permet pas d'en résoudre la solution astrométrique.

Une première option triviale est de la localiser manuellement. Mais l'automatisation de cette étape est bien entendue privilégiée. Dans le cas de l'observation d'une étoile standard isolée (pour la calibration photométrique), il est suffisant d'estimer sa position à partir des spaxels les plus brillants (ou la médiane de la position de N spaxels pour éviter un outlier comme un rayon cosmique).

Mais dans le cas de l'observation d'un transient, c'est un peu plus subtil. Cette dernière peut en effet être accompagnée de sa galaxie hôte dans le champ de vue, et la méthode des spaxels brillants ne marche en générale plus.

Nous avons mentionné dans la section 3.1.2 que la SEDm était, en plus de l'IFU, équipée d'une Rainbow Camera, un canal photométrique avec un champ de vue bien plus important de $13' \times 13'$. Le prisme d'interception pour l'IFU étant fixé, il est possible de déterminer une solution WCS (World Coordinate System) à partir des images de guidage, et de les projeter dans le MLA à une longueur d'onde de référence arbitraire. La solution WCS permet ainsi de passer de l'espace des pixels/spaxels à l'espace des coordonnées célestes *et vice versa*. Connaissant la position céleste de la cible (par détection photométrique avec la caméra ZTF), on peut ainsi déterminer sa position dans le MLA. Cette méthode est précise à environ 1 – 2 arcsec près.

3.3.2 Extraction de la source ponctuelle

Une fois la position de la cible connue, on peut à présent en extraire le spectre du cube de données. La première étape est de modéliser la source ponctuelle, qui est entièrement définie par sa position, sa fonction d'étalement de point (PSF), et son amplitude. Il faut à cela rajouter une composante de fond, le background, qui doit modéliser le ciel (constante) mais également la lumière de la galaxie hôte, considéré ici comme un background structuré.

Le processus d'extraction est la scission du cube 3D en $N \times 2D$ meta-tranches indépendantes (1 meta-tranche étant un empilement de n tranches pour un meilleur ratio signal/bruit), l'extraction de la source ponctuelle dans chacune d'elles, et enfin l'extrapolation à tout le cube en modélisant la chromaticité. Dans un premier temps le pipeline `PYSEDM` considère un disque de rayon 10 spaxels autour de la position de la cible estimée avec les images de guidage (section 3.3.1). Cette position est une condition initiale, et permet de se limiter à un sous-cube (un cylindre de rayon 10 spaxels).

Raisonnons en 2D. Dans chaque meta-tranche nous avons à déterminer :

- **Le background**, qui est structuré à cause de la présence potentielle de la galaxie hôte dans le champ de vue, est modélisé par un polynôme d'ordre 2, et a ainsi 3 paramètres libres.
- **L'amplitude** de la source ponctuelle, 1 paramètre libre.
- **La position** (x, y) de la source ponctuelle, 2 paramètres libres.
- **La PSF** de la source ponctuelle, modélisée par une combinaison linéaire Gaussienne + Moffat (? , ?). Ce modèle est paramétré avec 3 paramètres libres dans ? : le rayon de la Moffat, celui de la Gaussienne et le poids entre les deux distributions. Il faut également rajouter 2 paramètres libres d'ellipticité (un pour l'angle, l'autre pour l'excentricité).

Tous ces paramètres sont fittés pour chaque meta-tranches prises indépendamment les unes des autres.

Le set de paramètres ainsi obtenu est utilisé pour déterminer la chromaticité de la position et de la PSF. L'amplitude et le background sont des paramètres de nuisance à ce stade. Dans `PYSEDM` , la chromaticité de l'ellipticité, du rayon de la Moffat et du poids entre la Gaussienne et la Moffat sont modélisés par une constante. La chromaticité du rayon de la Gaussienne est modélisé par une loi de puissance. Quant à la position, sa chromaticité est due à la réfraction de la lumière par l'atmosphère. C'est ce qu'on appelle l'ADR (Atmospheric Differential Refraction). Nous détaillerons en détail cet effet et sa modélisation dans la Partie II de ce manuscrit (Section ??). Ainsi, les paramètres des modèles de chromaticités sont à leur tour fittés à partir du set de paramètres obtenu avec les $N \times 2D$ meta-tranches.

Une fois cela effectué, la PSF et la position de la cible sont fixées pour chaque longueur d'onde, et un dernier fit linéaire sur l'ensemble du cube est effectué pour les paramètres du background et de l'amplitude de la PSF. L'extraction de cette amplitude à chaque longueur d'onde du cube de données fournit ainsi le spectre de la source ponctuelle.

3.3.3 Calibration en flux

Le spectre extrait dans la section 3.3.2 étant en unité de pseudo-ADU (Analog to Digital Units), il faut à présent procéder à sa calibration afin de pouvoir l'exprimer en unité de flux physique. Le spectre d'une source astronomique observée peut être exprimé suivant le formalisme suivant (?) :

$$S(\lambda, t, z) = S^*(\lambda, t) \times \mathcal{C}(\lambda, t) \times T_{atm}(\lambda, t, z) \quad (3.1)$$

Avec $S^*(\lambda, t)$ le spectre instrinsèque de la source en unités physiques ($erg/cm^2/s/\text{\AA}$), $\mathcal{C}(\lambda, t)$ la réponse instrumentale, et $T_{atm}(\lambda, t, z)$ la transmission atmosphérique, qui dépend de la masse d'air le long de la ligne de visée. Calibrer un spectre revient donc à déterminer \mathcal{C} et T_{atm} pour isoler S^* .

En toute rigueur, la transmission atmosphérique devrait être décomposée en composantes physiques bien connues comme la diffusion de Rayleigh, la diffusion aérosol, l'absorption de l'ozone et l'absoption telluriques (? , ?, ?).

Dans le cadre de la SEDm, le but n'étant pas de faire une étude spectrophotométrique des sources observées, ce formalisme a été fortement simplifié et est exprimé tel que :

$$S(\lambda, t, z) = S^*(\lambda, t) \times [\mathcal{C}(\lambda, t) + \mathcal{T}(\lambda, t, z)] \quad (3.2)$$

Où \mathcal{T} représente l'absoption tellurique. Les spectres telluriques utilisées sont ceux du Kitt Peak National Observatory ¹, scindés en deux catégories de longueur d'onde : l' O_2 et l' H_2O . L'absoption tellurique est alors exprimée suivant :

$$\mathcal{T}(z) = \mathcal{T}_{\text{O}_2} \times (c_{\text{O}_2} + z^{\rho_{\text{O}_2}}) + \mathcal{T}_{\text{H}_2\text{O}} \times (c_{\text{H}_2\text{O}} + z^{\rho_{\text{H}_2\text{O}}}) \quad (3.3)$$

Où les amplitudes relatives c_i et les dépendances en masse d'air ρ_i sont des paramètres libres. Quant à la réponse instrumentale \mathcal{C} , elle est modélisée par un polynome de Legendre d'ordre 20.

La détermination des composantes \mathcal{C} et \mathcal{T} est alors la suivante :

- (a) On observe des étoiles standard du catalogue Calspec (?) avec la SEDm.
- (b) On en extrait le spectre en pseudo-ADU avec la méthode d'extraction de source ponctuelle (Section 3.3.2).
- (c) On récupère dans les archives calspec ² le spectre spectrophotométrique correspondant à l'étoile standard observée.
- (d) Les composantes de réponse instrumentale \mathcal{C} et d'absoption telluriques \mathcal{T} sont fittées simultanément en minimisant la quantité $(S_{ADU}/S_{ref}) - (\mathcal{C} + \mathcal{T})$.

Au moins une étoile standard est observée chaque nuit. La calibration d'une observation scientifique se fait en considérant la masse d'air associée (l'absoption tellurique est airmass-dépendante). On obtient alors le spectre calibré simplement en effectuant l'opération :

$$S^*(\lambda, t, z) = \frac{S_{ADU}(\lambda, t)}{\mathcal{C}(\lambda, t) + \mathcal{T}(\lambda, t, z)}$$

1. <http://www.noao.edu/kpno/>

2. https://archive.stsci.edu/hlsp/references-atlases/cdbs/current_calspec/

On appelle la quantité $(\mathcal{C} + \mathcal{T})$ la courbe de sensibilité inverse. La précision de la calibration en flux des spectres de science acquis avec la SEDm en utilisant le pipeline **PYSEDM** est de l'ordre de quelques pourcents.

3.4 Classification

Rappelons que le rôle premier de la SEDm au sein de la collaboration ZTF est la classification des événements transitoires.

Les deux classificateurs spectraux principaux existants sont **SUPERFIT** (?) et **SNID** (?). **SUPERFIT** est un software écrit en **IDL** utilisant une méthode de minimisation de χ^2 , et **SNID** est quant à lui écrit en **FORTRAN** et utilise l'algorithme de corrélation croisée de ?. Nous pouvons également mentionner le plus récent classifieur de supernovae **DASH** (?), utilisant une approche de deep learning.

Dans le pipeline de réduction de données de la SEDm, c'est le classificateur **SNID** qui est utilisé. Ce dernier est disponible publiquement¹ et régulièrement mis à jour. Pour effectuer l'analyse des corrélations entre le spectre d'entrée et la base de données, un pré-traitement est effectué (section 2.3 de ?). Celui-ci consiste en (1) Binner le spectre d'entrée en coordonnées $\ln(\lambda)$, (2) Diviser par le continuum du spectre estimé à partir d'une spline d'ordre 13 et enfin (3) Appliquer un filtre passe-bande d'ordre 4.

La fiabilité de la classification est quantifiée par 2 paramètres : le ratio “hauteur-bruit” r qui quantifie l'importance du pic de la fonction de correlation normalisée, et le paramètre de superposition des spectres ($rlap$), qui est par définition compris entre $0 < rlap < \ln(\lambda_1/\lambda_2)$ (où λ_1 et λ_2 sont les extrêmes de l'intervalle de longueur d'onde communs entre le spectre d'entrée et les spectres de la base de données).

Associés, ces deux paramètres forment un paramètre de qualité, le $rlap = r \times ldp$. Dans la section 6.1 de ?, il est montré qu'avec un $rlap \gtrsim 5$ la confusion entre une SNIa et un autre type est quasi non-existent ($\lesssim 2\%$) sans aucune contrainte sur le redshift ou la phase de la supernova.

De ce fait, **PYSEDM** rapporte tout spectre classifié quelque soit le type lorsque le paramètre de qualité $rlap > 5$. Nous présentons dans la Figure 3.5 un exemple d'extraction de source ponctuelle avec **PYSEDM** et sa classification avec **SNID**.

La banque de modèles utilisée comme référence pour la classification est une combinaison de plusieurs sets de données :

- Le set d'entraînement utilisé par **DASH** (?) : celui-ci est composé du **TEMPLATE-2.0** de **SNID**, de **SNIb/c** de ?, ?, ?, ainsi que des spectres du programme SN Ia 7.0 de Berkeley (?).
- Les spectres **SNIIP** de ?.
- Les spectres **SLSN-Ic** de ?.
- Plusieurs **SLSN-I**, **SLSN-IIIn** et **TDE** ajoutés par J.D.Neill.

Les spectres pour lesquels la date du maximum de luminosité ont été retirés. La banque de données finale utilisée pour **SNID** contient 3288 spectres de 312 SNe Ia, 1055 spectres de

¹. <https://people.lam.fr/blondin.stephane/software/snid/>

80 SNe Ib/c, 620 spectres de 33 SNe II, 207 spectres de 35 SLSNe, 29 spectres de 7 TDE, 15 spectres de 3 LBVs, 11 spectres de 11 galaxies, 11 spectres de M-stars, et 1 spectre de 1 AGN.

3.5 Contamination par la galaxie hôte

Il est indéniable que le pipeline **PYSEDM**, de par sa polyvalence et son automatisation, est d'une grande efficacité en ce qui concerne la réduction de données et l'extraction de sources ponctuelles.

Cependant cela n'est rendu possible que lorsque la source ponctuelle est isolée de sa galaxie hôte dans le champ de vue de l'IFU. Dans le cas contraire, nous entrons dans un régime de contamination qui nuit à la qualité de l'extraction, et a fortiori peut empêcher la classification spectral de la supernova. Sachant que la FWHM d'une source ponctuelle est de l'ordre de 2 arcsec pour la SEDm, il est raisonnable de considérer une contamination croissante par la galaxie hôte lorsque la distance angulaire passe en deça des 4 arcsec, jusqu'à être presque confondu dans le coeur de la galaxie hôte en deça de 2 arcsec (n'oublions pas que le bulbe de la galaxie est également étendue). Hors nous pouvons voir dans la Figure 3.6 (?), que près de 60% des SNeIa sont à une distance angulaire inférieur à 4 arcsec, et 40% pour les supernovae à effondrement de coeur. Cette fraction passe à 45% et 25% en deça de 2 arcsec pour les SNeIa et CC SNe respectivement, et environ 20% et 10% en deça de 1 arcsec.

Dans le cadre d'une observation où la source ponctuelle est faiblement contaminée par son hôte, une sensible amélioration de la méthode d'extraction a été apporté par ? avec ses modules **BYECR** et **CONSTEP** :

- **BYECR** a pour rôle de retirer les rayons cosmiques après construction des cubes de données, afin d'utiliser les informations spatiales. Ce module va normaliser le flux de chaque voxel en le divisant par le flux moyen dans un interval de $[-10, 10]$ tranches spectrales. Une comparaison est ensuite effectué entre le flux du voxel considéré et celui de ses 6 (structure hexagonale) voisins. Un écart supérieur à 5σ est considéré comme une présence d'un rayon cosmic, et le voxel en question est ainsi ignoré lors de l'extraction spectral de la source ponctuelle.
- **CONSTEP** Le but ici est de tenter d'ignorer automatiquement les spaxels où le flux est dominé par celui de la galaxie hôte. L'idée est de trouver le contour d'isomagnitude le moins brillant qui sépare la galaxie de la source ponctuelle. ? (Section 2.2 et Figure 2) utilise les images dans la bande r du relevé PANSTARRS (?), dans lesquels une source ponctuelle fictive de magnitude 16 est insérée à la position céleste de détection. Cela permet alors de récupérer les contours d'isomagnitude, et ainsi ne sélectionner "que" les spaxels de la source ponctuelle lors de son extraction.

Vis à vis du problème de la contamination par la galaxie hôte, seul **CONSTEP** est d'intérêt. Seulement, celui ci n'apporte qu'une augmentation de l'ordre de 0.5% du nombre de spectres classifiés (Table 2 de ?), aucun changement notable dans la distribution des rlap obtenus avec **SNID**, et une amélioration globale de la classification de 1.5% en se basant sur celles de l'échantillon BTS comme référence.

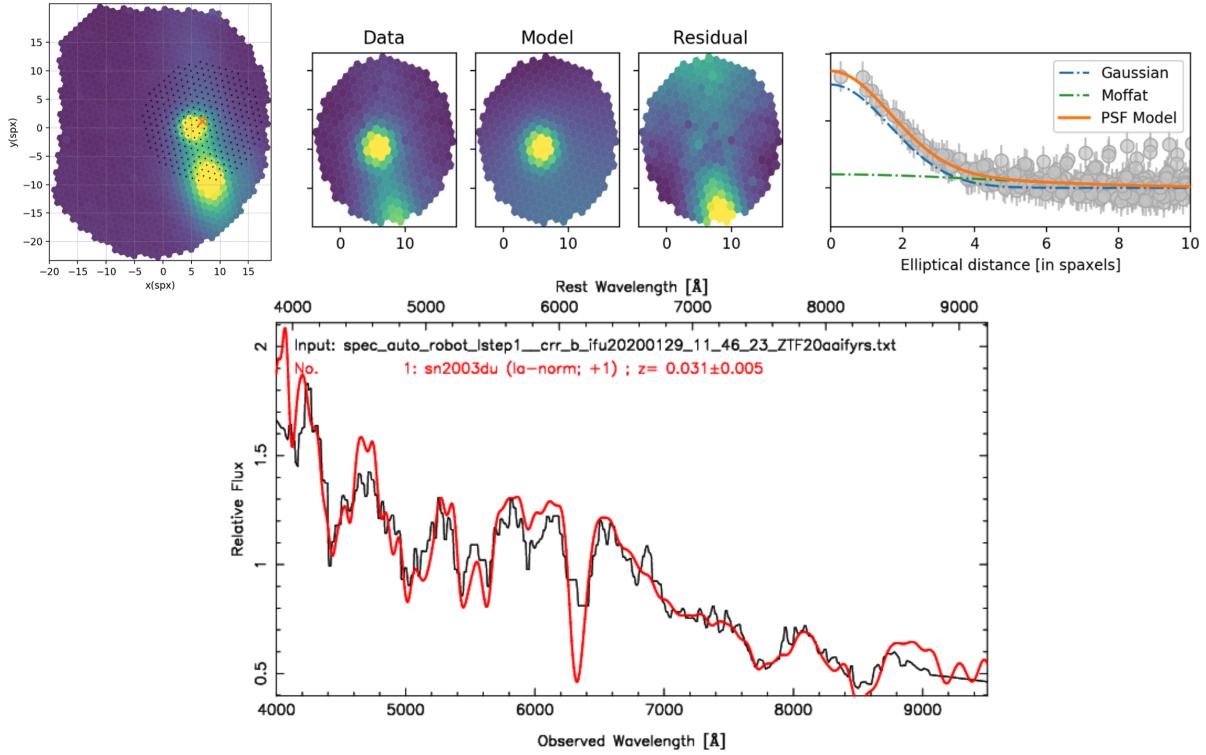


FIGURE 3.5 – Exemple d’extraction de source ponctuelle avec **PYSEDM** : ZTF20aaifyrs. *En haut, de gauche à droite :* (a) Image 2D intégré du cube 3D extrait suivant les étapes de la section 3.2. La croix orange indique la position estimée de la source ponctuelle à partir des images de guidages de la RC (section 3.3.1). Les marqueurs noirs correspondent aux spaxels considérés pour l’extraction de la source. (b) Meta-tranche intégrée entre $\lambda = [5750 - 6167]\text{\AA}$ centré sur la position estimée de la source ponctuelle et de rayon 10 spaxels. (c) Modèle fitté avec la composante de background et la PSF (section 3.3.2). (d) Résidu Data-Modèle, le signal résiduel en bas correspond à un bout de la galaxie hôte. (e) Profil radial de la source ponctuelle, exprimé en pseudo-ADU en fonction du rayon elliptique après soustraction du modèle de background. La courbe bleue représente la composante Gaussienne, la courbe verte la Moffat et la courbe orange le profil modélisé total. Les données sont représentées par les marqueurs gris avec leur barre d’erreur. *En bas :* Spectre extrait en noir, et le meilleur modèle de supernova fitté par **SNID** en rouge. On peut clairement voir ici que la source ponctuelle extraite est une Supernova de type Ia proche de son pic de luminosité.

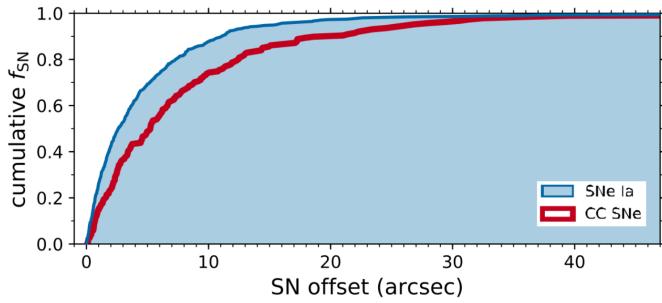


FIGURE 3.6 – Distribution cumulée du décalage angulaire, en secondes d’arc, entre les supernovae BTS et leur galaxie hôte (Figure 6 de ?). L’anagramme “CC” fait référence au terme “Core-Collapse”. Les supernovae de type Ia ont un redshift en moyenne plus élevé, ce qui explique cette distribution de plus faible distance angulaire comparé aux autres types.

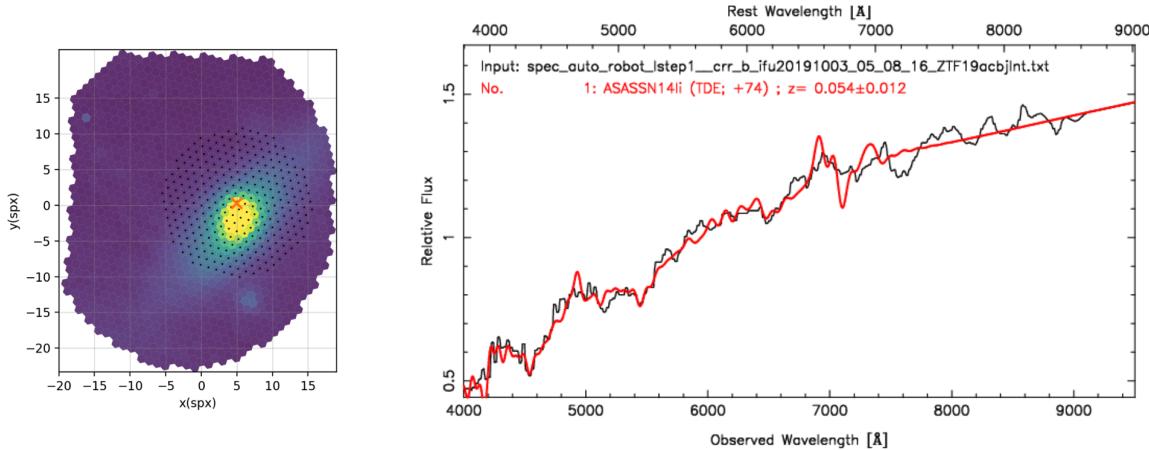


FIGURE 3.7 – Exemple de situation extrême de contamination de supernova par la galaxie hôte : ZTF19acbJnt. À gauche est représenté une image 2D du cube intégré entre $[5000 - 8000]\text{\AA}$. La croix orange indique la position estimée de la supernova à partir des images de guidage de la RC. Les marqueurs noirs indiquent les spaxels considérés pour l’extraction de source ponctuelle. À droite le spectre extrait en noir et la tentative de classification par SNID. On voit ici clairement que c'est à la fois le coeur de la galaxie et la supernova qui ont été extraits. L'extraction avec `CONTSEP` ne change rien au résultat.

En effet, `CONTSEP` n'apporte (presque) aucun bénéfice dans le cas où l'astrométrie dans le MLA est peu précise et/ou quand la source ponctuelle n'est pas séparée du coeur de la galaxie hôte.

La Figure 3.7 illustre une telle situation, dans un cas extrême où la supernova (ici ZTF19acbJnt) explode quasiment dans la ligne de visée du coeur de la galaxie.

Un moyen de lever cette contamination serait d'être en mesure de modéliser la galaxie elle-même, afin de complètement isoler la source ponctuelle dans le cube de données et ainsi procéder à une extraction propre.